March 27 - 29, 2012

Stuttgart, Germany

3rd International Conference on Machine Control & Guidance

Proceedings

Ed. by:

Volker Schwieger — Institute of Engineering Geodesy, University of Stuttgart

Stefan Böttinger — Institute of Agricultural Engineering, University of Hohenheim

Bimin Zheng — Institute of Engineering Geodesy, University of Stuttgart

3rd International Conference on Machine Control & Guidance

Proceedings

27th – 29th March 2012

University of Stuttgart, Germany

Institute of Engineering Geodesy

University of Hohenheim, Germany

Institute of Agricultural Engineering

Edited by:

Volker Schwieger, Stefan Böttinger and Bimin Zheng

These proceedings can be ordered from:
sekretariat@ingeo.uni-stuttgart.de

ISBN: 978-3-00-037295-7

www.uni-stuttgart.de/ingeo/mcg2012

Printed in Germany

Preface

Due to some typical trends in highly industrialized countries, automation will be a key topic in the future. The ageing of society, the lack of a qualified workforce, and the pressure towards cost-efficient solutions will lead to a higher demand for automated solutions.

The field of machine guidance and control is emerging and innovative for agricultural as well as construction applications. Interdisciplinary work is needed to fill the challenging tasks like automatic driving, seeding and reaping as well as milling and constructing. The topic even shows interfaces to robotics and driver assistance systems for intelligent transport systems. All these applications, developments, and research need geometric information like three-dimensional positions and altitudes to enable control and guidance of machines. This is the point, where engineering geodesy has its field of application and can help the other disciplines reaching the control aim. The interdisciplinarity is obvious: engineers and scientists from control engineering, mechanics, electronics, informatics, agriculture, construction, and engineering geodesy have to work together.

There are already many guidance and control solutions on the market, ready-to-use for the customers or customized for the newest applications; e.g. autonomous asphalt paver systems or tractors. On the other hand, highest accuracy demands like precision seeding or curb and gutter pavers as well as the full automation of operational-safety-related machines like diggers are still issues of research. For this reason the conference is a discussion platform and an interchange fair for practitioners as well as researchers, both learning from each other.

In 2008, the ETH Zürich launched the first conference on Machine Control and Guidance that was an overwhelming success, showing an interdisciplinary and international feedback not being expected. This was a good reason for the second conference in 2010 at the University of Bonn, showing again the high interest of the community. This year the conference is jointly organized by the Universities of Hohenheim and Stuttgart, represented by the Institute of Agricultural Engineering and the Institute of Engineering Geodesy.

In this conference the program focusses on:

- Global Navigation Satellite Systems (GNSS)
- Kinematic Measurements
- Sensor Integration
- Data Management and Communication
- Control Algorithms
- Construction Applications
- Agricultural Applications

The organizers like to thank all authors and members of the scientific committee for making this book a success. Finally, we want to thank the presenters and the audience in advance for contributing to the conference. We hope that interdisciplinary discussions, co-operations and projects will be an outcome of this conference.

Stuttgart, February 2012

S. Böttinger   Volker Schwieger   Bimin Zheng

Stefan Böttinger   Volker Schwieger   Bimin Zheng

Content

Global Navigation Satellite Systems I

Real-Time PPP using Open CORS Networks and RTCM Standards

Andrea Stürze[1)], Leos Mervart[2)], Wolfgang Söhne[1)], Georg Weber[1)] , Gerhard Wübbena[3)]
[1]Federal Agency for Cartography and Geodesy, Frankfurt, Germany
[2]Geodetic Institute, Czech Technical University, Prague
[3]Geo++ GmbH, Garbsen, Germany

Abstract
Using GNSS is an appropriate method for positioning and navigation even in the engineering sector. For absolute point positioning, having only one single station or equipment, different methods and services exist, which provide corrections for improvement of the kinematic or static position. Globally or regionally available, the services have in common that they are either commercial and / or they are using proprietary formats for corrections. However, relative point positioning in the form of differential GNSS or (Network) RTK is well standardized. Nevertheless, both approaches have their limitations.

The RTCM SC 104 (Radio Technical Commission for Maritime Services, Special Committee 104) is in progress now to release a new format combining the advantages of Precise Point Positioning (PPP) and Real-Time Kinematic (RTK) to one concept that was termed very early PPP-RTK by WÜBBENA et al. (2005). The so-called State Space Representation (SSR) approach is defining corrections, which will be transferred to the user e.g. via NTRIP (Network Transport of RTCM via Internet Protocol). Depending on the level of refinement, different accuracies will be reachable. The convergence times to reach a certain accuracy level will depend on the density of the network.

This paper describes the logic behind the upcoming RTCM standard, a pilot project, which so far exists under the umbrella of the IAG (International Association of Geodesy) services, and finally shows some results using already existing satellite orbit and clock corrections.

Keywords
GNSS, Real-Time, RTCM, IGS, PPP, RTK

1 INTRODUCTION

1.1 PPP

Precise absolute positioning is associated with the term Precise Point Positioning (PPP) using precise satellite orbit and clock parameters derived from a global network of reference stations. In other words, the principle is to provide information on individual error sources which can be referred to as State Space Representation (SSR). Applications are available in post processing using the IGS (International GNSS Service) products as well as in real time using proprietary data formats and commercial systems and services achieving accuracies in the sub-meter or even sub-decimeter level with dual frequency receivers on a global scale.

Precise absolute positioning with IGS products is based on the un-differenced ionosphere-free linear combinations of dual-frequency GNSS pseudo range and carrier phase observations and applies the orbit and clock information among other corrections. The user's position and a parameter describing the troposphere effect have to be estimated. Because the linear combination coefficients are no integer values and the available state information derived from global networks is insufficient to preserve the integer characteristics of the ambiguities, it is impossible to fix the ambiguities to integer values. Hence, for higher accuracy levels long observation times are required.

1.2 RTK and Network RTK

Differential GNSS techniques using a station with known coordinates as reference station allow centimetre-level positioning in real-time with dual-frequency carrier phase observations by fixing their double-differenced ambiguities to the correct integer values. This technique is referred to as Precise DGPS or RTK. Differential techniques degrade with increasing distance to the reference station because of distance-dependent biases such as orbit errors or atmospheric refraction. Therefore, centimetre-level RTK systems are restricted to distances of about 10 km.

A possibility to overcome the problem of distance dependent errors is the interconnection of reference stations and an estimation of the error state within the working area. Because this so called networked RTK in local or regional reference station networks has proven to be an efficient technology for high accuracy DGNSS positioning, several concepts have been developed and standardized over the last years. The latest RTCM Version 3.1 standard defines a variety of Network RTK messages. RTK operation using the RTCM standard is currently based on corrections or raw measurements from single or multiple reference stations. Hence, this approach can be called Observation Space Representation (OSR).

2 RTCM STATE SPACE REPRESENTATION MESSAGES

2.1 Background and Motivation

In view of the completion of the GLONASS system as well as the upcoming European Galileo system and the GPS modernization program, the availability of a large number of new observables is expected. A disadvantage of the conventional differential positioning approaches is that significantly more data need to be transferred between reference stations and rovers when using all observables to obtain the best possible accuracy. On the one hand, a service provider may not be able to track such a diversity of signals; on the other hand, mixing of different types of frequencies and codes on provider and user side will cause problems when using the mentioned observation space representation.

One way to reduce the data transfer and to solve this problem would be to apply positioning algorithms capable of using state space information. This means that current state parameters are transmitted to the rover and the user corrects his observations with corrections computed from these parameters representing the complete GNSS state to perform improved positioning. The state parameters must not necessarily be derived from compatible observations.

The state space approach would have several further advantages. So, along with the redundancy within a network, residual local errors of an individual reference station are eliminated or significantly reduced and missing observations do not result in missing state space information. Because satellite dependent parameters are globally valid but atmospheric parameters have global, regional, or even local validity depending on the aspired accuracy level, the state space representation approach opens the possibility to use hierarchical reference station networks. However, the consistency of different correction data sets has to be ensured. Using an observation space approach, the update rate will be a compromise between the requirement of the parameter having the highest dynamic as well as the endeavour to reduce the transmission bandwidth as much as possible. But, using the state space approach, for each parameter individual update rates can be defined. As orbits and troposphere have high correlation in time, their update rates can be low. Satellite clocks and parameters describing the ionosphere however require higher update rates. Defining state parameters as corrections to reference parameters or models allows a further bandwidth reduction.

The major disadvantages are a higher standardization effort and the more complex rover algorithm. The rover has to apply several additional corrections, such as relativistic corrections, phase wind-up, earth tides, etc. in order to realize a consistent modelling with the service provider and to take account of such an absolute positioning approach.

2.2 Principle

The principle of the State Space Representation concept is to provide information on individual GNSS error sources. The respective state vector should consist at least of the following parameters:

- satellite orbit errors,
- satellite clock errors,
- satellite signal biases,
- ionosphere delay parameters,
- troposphere delay parameters

as well as quality indicators for all the state parameters.

Additional issues are relevant to SSR in order to ensure a consistent modelling. Among them are the treatment of site displacements, the standardization of correction models regarding troposphere, the consideration of different coordinate reference systems, the consideration of higher order ionosphere effects, satellite attitude models, satellite phase centre variation (PCV) and offset (PCO) models, and others. The International Earth Rotation and Reference Systems Service (IERS) conventions (PETIT and LUZUM, 2010) can serve as a reference for their standardization.

On its meeting in February 2007, the RTCM SC-104 decided to establish a working group “State Space” which is chaired by Dr. Gerhard Wübbena. The goal of the working group is the development of State Space Representation concepts and messages for all levels of targeted accuracies including RTK. The work plan consists of three major steps:

1. The development of messages for precise orbits, satellite clocks and satellite code biases. This is compatible to the basic PPP mode using IGS products. Such messages will enable real-time PPP for dual frequency receivers.
2. The development of vertical TEC (Total Electron Content) messages (VTEC). This will enable real-time PPP for single frequency receivers.
3. The development of slant TEC messages (STEC), troposphere messages and satellite phase biases messages. This will enable real-time PPP-RTK.

Meanwhile, RTCM has completed the development of Version 3 messages to transmit satellite orbit and clock corrections to broadcast ephemeris as well as code biases and user range accuracies for GPS and GLONASS as mentioned in step one. In July 2011, the following State Space Representation messages became an RTCM Recommended Open Standard:

- Message type 1057: GPS orbit corrections to Broadcast Ephemeris
- Message type 1058: GPS clock corrections to Broadcast Ephemeris
- Message type 1059: GPS code biases
- Message type 1060: Combined orbit and clock corrections to GPS Broadcast Ephemeris
- Message type 1061: GPS User Range Accuracy (URA)
- Message type 1062: High-rate GPS clock corrections to Broadcast Ephemeris
- Message type 1063: GLONASS orbit corrections to Broadcast Ephemeris
- Message type 1064: GLONASS clock corrections to Broadcast Ephemeris
- Message type 1065: GLONASS code biases
- Message type 1066: Combined orbit and clock corrections to GLONASS Broadcast Ephemeris
- Message type 1067: GLONASS User Range Accuracy (URA)
- Message type 1068: High-rate GLONASS clock corrections to Broadcast Ephemeris

Based on the respective RTCM PAPER (142-2011-SC104-STD) as well as on WÜBBENA et al. (2010), the following sections describe these messages in more detail.

2.3 Orbit and Clock Messages

The RTCM-SSR data sets for precise satellite orbit information contain parameters for orbit corrections $\delta\boldsymbol{O}$ in radial, along-track and cross-track component (**Figure 1**). These orbit corrections are used to compute a satellite position correction $\delta\boldsymbol{X}$, to be combined with satellite position $\boldsymbol{X}_{broadcast}$ calculated from broadcast ephemeris as follows:

$$\boldsymbol{X}_{orbit} = \boldsymbol{X}_{broadcast} - \delta\boldsymbol{X}$$

With

$\boldsymbol{X}_{orbit}$	satellite position corrected by SSR Orbit Correction message,
$\boldsymbol{X}_{broadcast}$	satellite position computed according to corresponding GNSS ICD ,
$\delta\boldsymbol{X}$	satellite position correction.

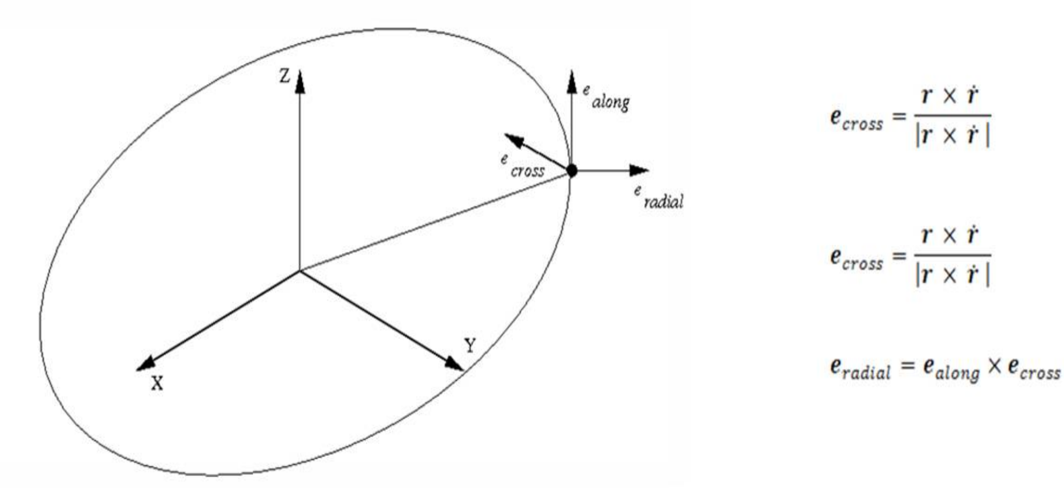


Figure 1: Radial, along-track and cross-track orbit components (RTCM Paper 142-2011-SC104-STD)

The satellite position correction $\delta\boldsymbol{X}$ can be computed according to

$$\delta\boldsymbol{X} = [\boldsymbol{e}_{radial} \quad \boldsymbol{e}_{along} \quad \boldsymbol{e}_{cross}]\, \delta\boldsymbol{O}$$

Using

$\boldsymbol{r} = \boldsymbol{X}_{\boldsymbol{broadcast}}$	satellite broadcast position vector,
$\dot{\boldsymbol{r}} = \dot{\boldsymbol{X}}_{\boldsymbol{broadcast}}$	satellite broadcast velocity vector,
$\boldsymbol{e}_i$	direction unit vector, i = {radial, along, cross},
$\delta\boldsymbol{O}$	orbit correction vector.

The complete orbit correction vector $\delta\boldsymbol{O}$ is computed from the individual correction terms and their velocities:

$$\delta\boldsymbol{O} = \begin{bmatrix} \delta O_{radial} \\ \delta O_{along} \\ \delta O_{cross} \end{bmatrix} + \begin{bmatrix} \delta \dot{O}_{radial} \\ \delta \dot{O}_{along} \\ \delta \dot{O}_{cross} \end{bmatrix} (t - t_0)$$

With

t	time,
t_0	reference time obtained from SSR Orbit Correction message,
$\delta O_i, \delta \dot{O}_i$	orbit correction terms from SSR Orbit message, i = {radial, along, cross}.

Orbit representation requires the definition of a coordinate reference system. For global services the coordinate system should be related to the ITRS. For regional services a reference system related to the tectonic plate of the region is often used. The RTCM-SSR data sets for precise satellite orbit information allow orbits transformed from ITRF to a global coordinate system close to ITRF (e.g. ETRF). Hence, it is not necessary for the rover to perform the corresponding transformation.

The RTCM-SSR data sets for precise satellite clock information contain parameters to compute the clock correction δC that can be δCapplied to the broadcast satellite clock $\mathrm{t_{broadcast}}$ as follows

$$t_{satellite} = t_{broadcast} - \frac{\delta C}{Speed\ of\ light}$$

Using

$t_{broadcast}$	satellite time computed according to corresponding GNSS ICD,
$t_{satellite}$	satellite time corrected by SSR Clock Correction message,
δC	clock correction obtained from SSR Clock Correction message.

The polynomial is computed according to

$$\delta C = C_0 + C_1(t - t_0) + C_2(t - t_0)^2$$

With

t	time,
t_0	reference time obtained from SSR Clock Correction message,
C_i	polynomial coefficients from SSR Clock Correction message, i = {0, 1, 2}.

The respective reference time t_0 is computed from the GNSS epoch time plus half the SSR update interval. Exception is SSR update interval “0”, which uses the GNSS epoch time as reference time.

2.4 Code Bias Messages

Per convention, satellite clocks are determined from ionosphere-free signals derived from observations used by the service provider. Such observations are affected by delays introduced through the satellite hardware. For example, GPS broadcast clocks are referenced to the ionosphere-free linear combination of the P codes on L1 and L2, ignoring any code biases. Hence a provider of RTCM-SSR corrections has to ensure a consistent transmission of clock and code bias parameters. A rover must then consistently apply the code biases and clock corrections.

The RTCM-SSR Code Biases have to be added to the pseudo range measurements of the corresponding code signal to get corrected pseudo ranges. The Code Bias message contains absolute values, but also enables the use of differential code biases by setting one of them to zero.

2.5 Consistency of parameters and processing

As mentioned above, the proposed concept shall enable the common use of different RTCM-SSR data sets to support different applications, update rates, and accuracy requirements. Therefore, the consistency of the different RTCM-SSR data sets is of importance. Because in real-time applications messages may be lost or delayed, the message chronology itself cannot ensure the consistency. Regarding the satellite orbit for example, the IOD / IODE (Issue of Data / Issue of Data, Ephemeris for GPS) can be used to realize consistency of both, computation and application of the respective RTCM-SSR messages.

In order to realize the consistency between different RTCM-SSR data sets in general, the validity interval defined through reference time and update interval has to be considered for the combination of different data sets. In other words, only data sets with a common validity interval are allowed to be combined. An example is displayed within **Figure 2**.

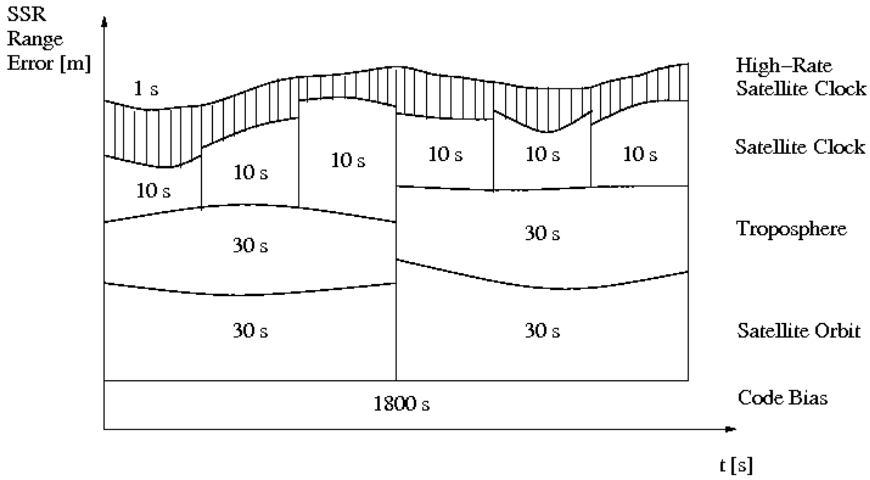


Figure 2: Concept of RTCM-SSR update interval to ensure the consistency of different SSR parameters (RTCM Paper 142-2011-SC104-STD)

Assuming a satellite orbit message that is defined for a 30 seconds period and three satellite clock data sets are generated for the same period with a 10 seconds update rate, consistency is ensured as long as one of these three clock data sets is combined with the overall satellite orbit message. Hence, three different, but consistent combinations of data sets are possible. Nevertheless, the latest combination should be the first choice.

3 REAL-TIME IGS PILOT PROJECT

As part of its strategic plan (IGS, 2008) the International GNSS Service is developing real-time products in order to support e.g. PPP applications. An IGS Real-time Pilot Project (IGS-RT PP) was initiated in 2007 (http://www.rtigs.net). The key project activities include network management, format specifications, product generation and distribution with the aim to improve the product accuracy and the real-time precise point positioning results. Through its close cooperation with RTCM, the Pilot Project is helping to develop and design GNSS data and correction formats. IGS Analysis Centres (AC) such as CODE (Center for Orbit Determination in Europe) or DLR (German Aerospace Center), are also generating for example ionosphere VTEC maps (SCHAER et al., 1998). A consideration of TEC values is foreseen in step two of the RTCM-SSR state space concept. Furthermore, the Pilot Project participants are exchanging data and real-time products using the formats that are under development and discussion.

Within the Pilot Project, IGS real-time analysis centres (ACs) are computing GNSS orbit and clock corrections using IGS ultra rapid predicted orbits and real-time observation streams from over 100 stations in the IGS network. Computed corrections are encoded in RTCM-SSR messages and disseminated over the Open Internet using the NTRIP transport protocol.

As of November 2011, nine ACs are contributing with a more or less routinely provided solution. Moreover, some ACs are processing a second solution in parallel. This second solution may use GLONASS in addition to GPS or may be processed on a different host. As mentioned, the major application of the Pilot Project products is real-time Precise Point Positioning, which is possible nowadays with accuracy better than one decimetre after about 15 minutes or even less convergence time (**Figure 3**).

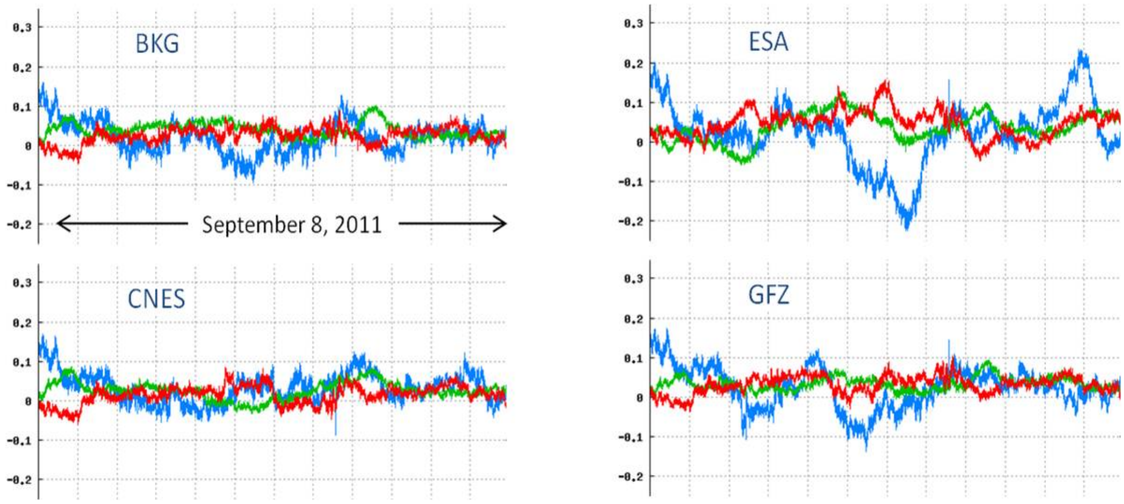


Figure 3: 24h time series of PPP applying RTCM-SSR of different Analysis Centres, North, East, Up (NEU) displacements, 4 examples, September 8, 2011 (WEBER and MERVART, 2011).

During 2011, a main focus of the Pilot Project was directed towards the development of combination solutions. As for almost all other products, the derivation of a reliable combination product is a major goal of IGS. A potential user should be put into the situation to access an unambiguous product. Different approaches for a combined IGS clock correction product are under development which may have their pros and cons with regards to e.g. robustness, outlier detection etc. (WEBER and MERVART, 2011 and WEBER et al., in preparation).

Taking into account the number of real-time RTCM-SSR products resulting from individual AC solutions as well as from combination approaches, BKG has established the broadcaster (http://products.igs-ip.net) that provides all IGS-RT PP products under an open data policy.

One important future goal of the IGS-RT PP and RTCM's State Space woking group is the correction streams to be implemented into manufacturers receiver firmware. Currently, there are only a few Open Source software packages available, which could be used by the geodetic community. The BKG NTRIP Client (BNC), for example has been developed for the Federal Agency for Cartography and Geodesy (BKG) within the framework of EUREF's Real-time GNSS Project and the Real-Time IGS Pilot Project (Söhne et al., 2009). BNC is a program for simultaneously retrieving, decoding and converting real-time GNSS data streams from NTRIP broadcasters like http://www.igs-ip.net/home. It is written under GNU general public license and can be downloaded via http://igs.bkg.bund.de/ntrip/download.

4 CONCLUSION AND OUTLOOK

Combining the advantages of both, Precise Point Positioning (PPP) and Real-Time Kinematic (RTK) to a common concept termed PPP-RTK opens the perspective to reach centimetre accuracy for an unlimited number of static or kinematic users.

In this paper, the underlying principle to use state space information derived from global, regional and local reference station networks is described. Furthermore, standardized messages for precise orbits, satellite clocks and satellite code biases and some application results are presented. The next steps required in SSR development for reaching RTK accuracy level are outlined.

Beneath the messages describing the GNSS error states the so called Multiple Signal Message (MSM) format is under discussion and development within RTCM in order to fully support operational, planned and future GNSS and their signals. Because of the similar nature of the observables corresponding to each of the currently known GNSS it is possible to present all observables for each

GNSS in a universal form. That means, the MSMs are designed to be valid for all existing and future GNSS signals and to be maximum compatible with RINEX (Receiver Independent Exchange Format) version 3 as well. The MSM set for each GNSS can be divided into so called standard precision and high precision messages. Both represent the same data, but high precision messages provide them with higher resolution (RTCM PAPER 195-2011-SC104-669, 2011).

ACKNOWLEDGEMENTS

The authors are greatly indebted to GNSS Analysis Centres operated by BKG/CTU, CNES, DLR, ESA, GFZ, GMV, NRCan, and Wuhan, which support the Real-time IGS Pilot Project through NTRIP broadcasts of satellite orbit and clock corrections. This work would not have been possible without their commitment to an open data policy.

REFERENCES

IGS, http://igscb.jpl.nasa.gov/igscb/resource/pubs/IGS_StrategicPlan_proof5.pdf, last accessed on December 05, 2011.

PETIT, G. and LUZUM, B. (eds.): *IERS Conventions (2010)*. IERS Technical Note 36, Frankfurt Main, Verlag des Bundesamts für Kartographie und Geodäsie, p. 179, 2010.

RTCM PAPER 142-2011-SC104-STD: *Amendment 5 to RTCM standard 10403.1 Differential GNSS (Global Navigation Satellite Systems) Services – Version 3 developed by RTCM Special Committee No. 104*, 6/ 2011.

RTCM PAPER 195-2011-SC104-669: *Comments on Committee Draft for Vote Amendment 5 to RTCM standard 10403.1 Differential GNSS (Global Navigation Satellite Systems) Services – Version 3*, 3/2011.

SCHAER, S., GURTNER, W., FELTENS, J.: *IONEX: The IONosphere Map EXchage Format Version 1*. Proceedings of the IGS AC Workshop, Darmstadt, Germany, 2/1998, p. 9-11, 1998.

SÖHNE, W., STÜRZE, A., WEBER, G., MERVART, L.: *Prospects of Real-time Satellite Positioning Using EPN Resources.* Bollettino di Geodesia e Scienze Affini, Volume LXVIII, no. 2, 2009.

WEBER, G. and MERVART, L.: *Real-time Combination of GNSS Orbit and Clock Correction Streams Using a Kalman Filter Approach*. Proceedings of the 24st International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2011), 2011.

WEBER, G., MERVART, L. AGROTIS, L., STÜRZE, A.: *Real-time Combination of GNSS Orbit and Clock Correctors for Precise Point Positioning.* Proceedings of the XXV IUGG General Assembly, Melbourne, Australia, June 28 – July 7, in preparation.

WÜBBENA, G., SCHMITZ, M. and BAGGE, A.: *PPP-RTK: Precise Positioning Using State-Space Representation in RTK networks.* Proceedings of the 18st International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2005), 2005.

WÜBBENA, G., SCHMITZ, M. and BAGGE, A.: *Communication via email,* 2010.

Online GNSS Service with Scalable Positioning Accuracy

Jürgen Alberding[1], Jens Wickert[2]
[1]Alberding GmbH
[2]Deutsches GeoForschungsZentrum GFZ

Abstract
The main project focus is the development and test of a new online GNSS processing algorithm using the Precise Point Positioning (PPP) approach. PPP takes advantage of global GNSS satellite information derived from a worldwide network of continuous operating GNSS reference stations. Based on real-time GPS clock and orbit solutions complemented with regional and local information the service will provide multiple GNSS correction data streams to the users. Depending on the quality of the GNSS user equipment, the selected service information and the processing algorithm the service will provide a scalable positioning accuracy in the range of 0.05–0.5m with minimal initialization time world-wide. The online GNSS correction data will be distributed via internet using the standardized Ntrip protocol. Access to the data streams will be provided by a dedicated user-box and via software tools (DLL) which could easily be integrated in field software running on Windows operating systems.

1 GNSS AUGMENTATION SYSTEMS

Since more than 15 years GNSS augmentation systems are used to increase the GNSS positioning accuracy and to control the GNSS system itself. Improvements of the position accuracy are achieved by using differential observation techniques. Relying on the fact that systematic GNSS errors are eliminated by differencing simultaneous measurements the existing GNSS services are based on continuously operating reference stations which are locally installed with known positions. Starting with the single reference station approach most of the systems have moved to network operation to provide more reliable and homogeneous information to the users.

The advantage of the existing DGNSS services is the standardized technique and the availability of dedicated product solutions on the market. The majority of these services can be differentiated into two groups: Code differential ("DGPS or DGNSS") and carrier phase differential services ("RTK networks", Real-Time Kinematics). Today code differential services are offered mainly by governmental institutes for larger regions. Observing corrections from regional DGNSS services the GNSS users get a position accuracy with professional DGNSS receivers (> EUR 500) in a range of 0.5 – 1m in real-time. The DGNSS code corrections are an essential part of GNSS services as EGNOS or WAAS.

The effort to establish and operate a RTK network (NetRTK) service is significantly higher because of expensive reference station hardware and networking software. The recommended distance between the continuous operating reference stations is 50 to 80km. For a region like Germany more than 150 GNSS reference stations are required to reach maximum service quality. The access to a NetRTK service is mainly provided via mobile internet. The costs of the GNSS hardware to achieve a real-time accuracy of 1-3cm are > EUR 7,500.

2 MARKET REQUIREMENTS FOR GNSS POSITIONING

The existing GNSS services cover already a wide area of the GNSS user demands. In the DGNSS market EGNOS provides a service free of charge and the corrections can be received without additional hardware at the user site via the GPS/GNSS antenna. In urban canyons EDAS (EGNOS Data Access Service) data is available via mobile internet and system integrity information will be available as well. The German waterway authority as a part of the IALA (International Association of Lighthouse Authorities) group provides DGNSS (GPS and GLONASS) corrections via radio beacons, AIS (Automatic Identification System) and mobile internet free of charge. Currently there is no requirement for additional regional DGNSS services.

The main challenges in the DGNSS market currently are: a global service approach, the integration of additional GNSS systems like GLONASS or Galileo and the price reduction of the GNSS hardware. If DGNSS accuracies of 0,5 to 1m could be achieved with low cost GNSS receivers (< EUR 100) the DGNSS market would increase dramatically. To bring DGNSS out of the niche market there is a need to provide global DGNSS corrections with unique access and to improve the quality of the hardware (antenna) and the processing algorithm inside the GNSS receivers.

In the high-end GNSS market the end user hardware price (> EUR 7.500) of professional GNSS receivers is one reason for the limited use of GNSS in many applications. The acceptance and the use of GNSS in the agriculture or the construction market would be much higher if the price of a GNSS system would be less than EUR 3.000. Besides the purchase price there are two other limitations: the costs of GNSS corrections for RTK accuracy and the missing unique service for a larger region. Suppliers of construction or agricultural machines are looking for a world-wide GNSS service which delivers positioning information of same quality. Currently there are a few commercial services mainly targeting the offshore oil market that provide 0.1-0.2m positioning accuracy after 20-30min initialization time via a world-wide network.

Because of the required dense GNSS base station network today the NetRTK services are of local or regional character. Besides governmental institutions, especially the national surveying organizations, more and more NetRTK services are provided by GNSS receiver manufacturers. In combination with their regional distributors GNSS receiver manufacturers have the opportunity to deliver a complete system to the end user including the service, the GNSS hardware and the application software. There might be a risk that powerful GNSS manufacturers are using this service to push their complete product portfolio by giving external suppliers a limited access (price, information, accuracy) to their GNSS service or their hardware products.

3 NEW APPROACH: PRECISE POINT POSITIONING (PPP) PLUS REGIONAL INFORMATION

The new precise point positioning (PPP) approach is basically different to existing DGNSS techniques which rely on correction data from permanent GNSS reference stations (or networks). In contrast to PPP DGNSS algorithms do not treat the GNSS error sources (satellite, atmosphere or receiver) separately. DGNSS combines the errors in just one correction per satellite which blows up the data protocol with more and more satellites and requires an update rate of 1 second for centimeter accuracy.

The new PPP algorithm is a single point processing method which relies on global GNSS information like corrections for the GNSS satellite orbits and clocks. These orbit and clock information is derived from a worldwide network of continuously operated reference stations (50-100 stations world-wide) and will be provided by research institutes as the German Research Center for Geosciences GFZ at Potsdam (www.gfz-potsdam.de) in real-time.

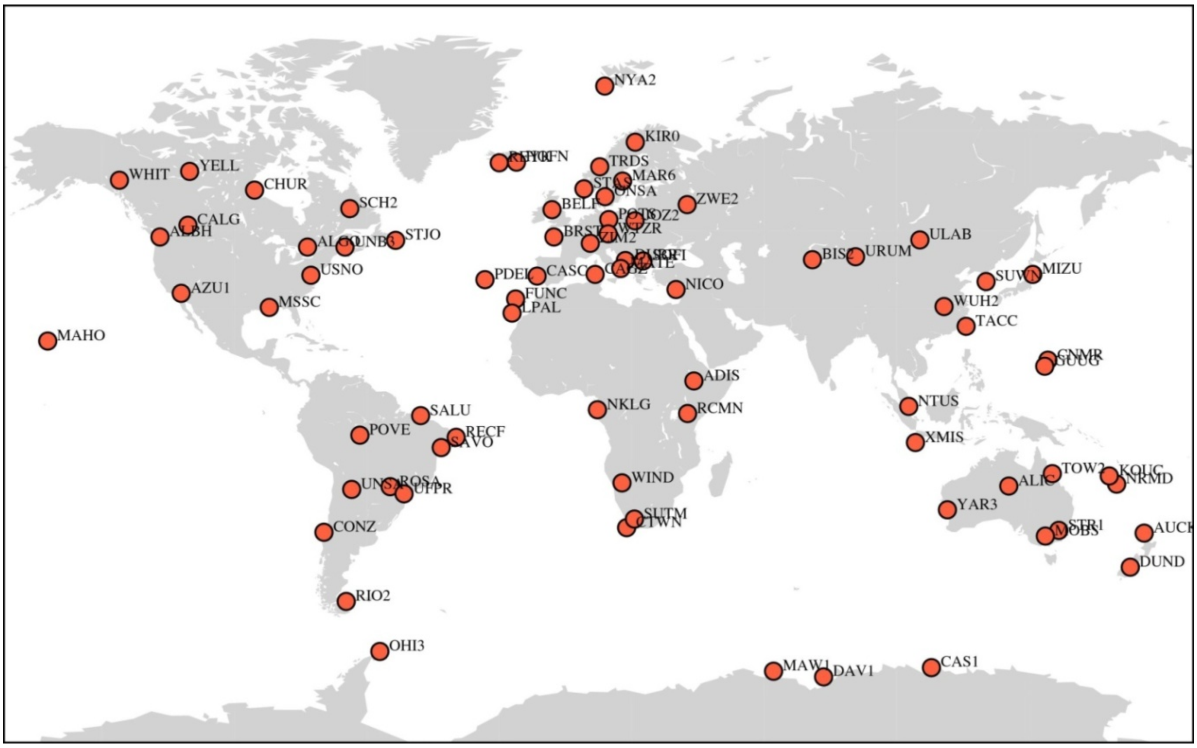


Figure 4: Real-time network for GNSS clock and orbit determination (Ge et al., 2011)

Based on global GNSS satellite clock and orbit corrections a GNSS service provider has now the opportunity to provide a world-wide GNSS service without any direct access to GNSS reference stations at all. Compared to NetRTK the PPP approach has significant advantages in the world-wide availability and the required update rates of the GNSS orbit and clock corrections. In combination with dual frequency GNSS receivers today PPP achieves a positioning accuracy of 0.1-0.2m after 20 to 30 minutes initialization time. The advantage of NetRTK is the positioning accuracy (RTK: 1-3cm vs. PPP: 1-2dm) and the faster initialization time (RTK: 1-2 min vs. PPP: 20-30 min). Therefore a major challenge of our project is to develop and improve the PPP algorithms in order to:

- Minimize the initialization time
- Improve the accuracy

Simulations and test measurements on permanent stations look very promising that it will be possible to reduce the initialization time to less than 5 min and to improve the accuracy to better than 10cm. If it will be possible to fix the carrier phase ambiguities the horizontal positioning accuracy will be in the range of less than 5cm.

4 GNSS SERVICE WITH SCALABLE POSITIONING ACCURACY

The basic idea of our project is to optimize a service on the server site in combination with a processing algorithm on the mobile user site. It has to be defined what kind of information a GNSS service has to provide in a required update rate to achieve the optimal positioning accuracy calculated by the algorithms on the mobile site. In addition to the global “clock and orbit” information it will be possible to integrate regional GNSS information in the service to provide higher accuracies in regions of interest. Additional information could be the data of regional GNSS reference stations and/or regional or local atmospheric models.

The main task of the project is the development and optimization of the processing algorithms and their integration in embedded hardware solutions for machine applications. Depending on the regional service data level, the processing method and the user equipment the algorithms will provide different positioning accuracy levels which makes the service scalable. The developed algorithms for servers have to be ported to less powerful platforms and Microsoft Windows operating systems for field applications.

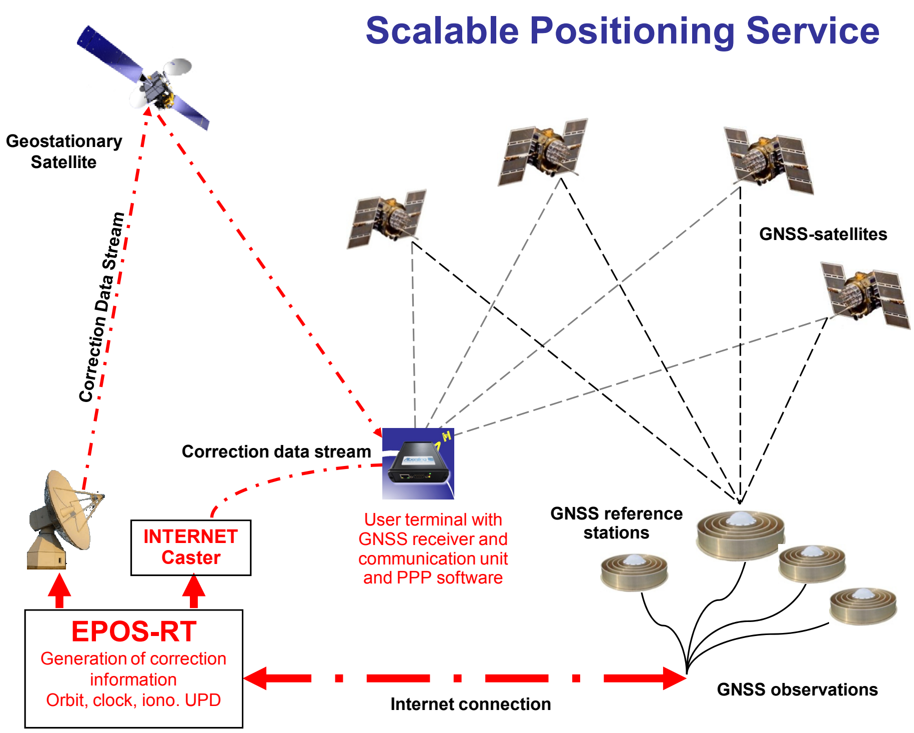


Figure 5: Schematic overview on the new scalable positioning service (Alberding and Wickert, 2010).

In contrast to existing services of GNSS receiver manufacturers the idea of the planned service is to input GNSS receiver raw data information, to process this data with the new developed algorithms and to output the position information in standardized formats. From the service providers viewpoint the "raw data" approach guaranties nearly the same accuracy level with all GNSS receivers. Because of dedicated input and output data streams for all systems the support effort concerning the GNSS users will be reduced.

The approach of the service to concentrate on the GNSS processing gives GNSS users the opportunity to choose between GNSS receiver manufacturers on one hand and between solution providers on the other hand. Compared to a complete package (hardware, service, software) of GNSS receiver manufacturers the approach in this project leaves room for competition between suppliers.

5 ACCESS TO THE SERVICE

Besides the development and improvement of the processing algorithms the mobile access to the service has to be guaranteed. Software solutions for the data transfer, encryption, GNSS raw data management, etc. have to be developed and integrated in the system software. Mobile internet using the Ntrip protocol will be the standard data transmission channel. Depending on user requirements it might be useful to integrate additional data channels like geostationary satellites or dedicated transmitters which can locally rebroadcast the data in regions without any mobile internet access.

For testing purposes a few hardware prototypes ("user box") will be developed. The prototypes consist of a low cost GNSS receiver, a GPRS modem and a powerful processor. The algorithms should provide a positioning accuracy of < 0,5m with the low cost GNSS receiver in the user box. Professional GNSS receivers and other sensors can be adapted to the user box via cable or wireless. In combination with external professional GNSS receivers the algorithm inside the user box should provide an accuracy of less than 10cm with an initialization time of less than 10 min. If there is no loss of the GNSS signals at the mobile site the accuracy from 10cm will be reduced to < 5cm after additional 20 min.

Because of existing communication equipment (mobile internet) and processing power on machines or mobile personal applications the algorithms for the GNSS data processing and the access to the service will be provided as a DLL to integrate them into existing software solutions.

6 STATUS OF THE PROJECT

The project started at November 1, 2010 and will be finished in March 2012. Currently we are focusing on the development of the GNSS processing algorithms and on the communication and server part on the other hand. The integration of the algorithms and the client software into the mobile units will be done by the end of December 2011. First field tests with the user box are planned for January 2012. For the first quarter 2012 it is planned to intensify the tests and to optimize the service and the algorithms.

By the end of Q1/2012 we expect to finish the developments successfully. It is planned that Alberding GmbH will be offering a professional scalable PPP service by mid 2012.

REFERENCES

ALBERDING, J., and J. WICKERT, *Online GNSS service with scalable positioning accuracy*, proposal to Federal Ministry for Economy, Wildau/Potsdam, 2010.

GE, M., X. LI, J. CHEN, J. DOUSA and J. WICKERT, *Global Real-Time Precise Point Positioning Service and Its Regional Augmentation*, talk at University Colloquium Satellite Navigation, Technical University Munich, Feb. 1, 2011.

Post-Processed Kinematic Low-Cost GPS

Annette Scheider, Alexander Beetz, Volker Schwieger
Institute of Engineering Geodesy (IIGS), University of Stuttgart, Germany

Abstract
Several investigations have been carried out already concerning the use of low-cost GPS receivers for static measurements. To some extent, the results achieved with these receivers in monitoring of glacial movements or landslides were very good. So, obviously the next step forward should be the investigation of tracking and monitoring of moving objects, such as motor vehicles. Up to now, in most cases expensive geodetic GPS or GNSS receivers have been used for this task. This paper focuses on investigations regarding the possibility to use low-cost GPS receivers for tracking and recording of vehicle trajectories with sub-decimeters accuracy. To obtain this accuracy, a low-cost receiver must be able to receive and give out the raw data of at least the carrier frequency L1. One receiver with this ability is the u-blox AEK-4T which was compared with a Leica 1200 GNSS system. For the evaluation of both receivers a total station Leica TCRP 1201 was measuring additionally the positions of several test trajectories simultaneously in kinematic mode. To do GPS and tachymeter measurements at the same time and with respect to the same point, the rover antenna was placed centered on a 360° prism.

The first step of the evaluation was to calculate baselines with a SA*POS*® station as reference station in post-processing. The accuracies of the kinematic survey of the geodetic and of the low-cost GPS receiver are compared. The lateral deviations of each GPS trajectory to the tachymeter reference trajectory were computed. To assess the position accuracy of the GPS receivers, the mean values of these lateral deviations were used. Both receiver types achieved a mean value worse by a factor of 3.5 than the tachymeter accuracy. In a second step, the baseline evaluation was done with a u-blox receiver as reference station. Compared to the first results, the accuracies were slightly worse by a factor of 1.7.

Keywords
low-cost GPS, DGPS, tracking, kinematic measurement, vehicle position

1 INTRODUCTION

For kinematic applications, only expensive geodetic GPS or GNSS receivers have been used to determine a position with a quality of a few centimeters for a long time. But the high prices of these receivers increase the costs for monitoring tasks. On the other hand, low-cost receivers are much cheaper, but they use only code measurements to determine a position. The latest low-Cost GPS modules like u-blox LEA-6T promise a horizontal position accuracy of 2.5 m for code measurements in UBLOX (2011b). The accuracy of these receivers can be improved by using them in DGPS applications. A 3D-accuracy of 2.7 cm was reached in SCHWIEGER (2009) by using carrier phase data for static measurements and a measuring epoch of 30 minutes. Low-cost GPS receivers have already been employed in geomonitoring networks with several permanent GPS stations measuring in static mode. Such networks are used to detect slow movements, e.g. glacial movements or landslides as described in GLABSCH et al. (2010a) and LIMPACH (2009). For those measurements, one receiver is working as a reference station, other receivers are rovers. For short baselines, accuracies of 1 mm (horizontal) and 5 mm (vertical) are reached in LIMPACH (2009). GLABSCH et al. (2010b) describe a mean standard deviation of 3.2 mm (horizontal) and 6.7 mm (vertical). In LIMPACH (2009) the data are also analyzed in kinematic mode. A coordinate accuracy of 1 cm (horizontal) and 1.5 cm (vertical) could be realized in this mode for such slow moving areas. But also single low-Cost GPS receivers can be used to detect movements with high precision. For example, they can monitor different types of moving objects, such as vehicles. Especially, if several vehicles shall be tracked at the same moment, low-cost GPS receivers are a cheaper solution than geodetic GNSS receivers. The position accuracy of this application does not reach a level of sub-decimeters at the moment.

This paper focuses on the use of low-cost GPS receivers for precise tracking applications of vehicle trajectories. Single positions along the trajectory are calculated in the post-processing evaluation by creating baselines. For that purpose, a rover station is placed on the vehicle and the reference station is positioned within the measuring area. This research tests the suitability of DGPS reference services (e.g. SA*POS*®) and local reference stations. The local reference station is realized by a second low-cost GPS receiver. To improve the quality of the trajectory, different time epochs for initializing have been tested before the vehicle was moved. The accuracy of this tracking procedure is considered to be better than 1 dm. To get a reliable result, several kinds of trajectories were driven on a free field. Shadowing and multipath effects of large objects around do not affect the measurement.

2 EQUIPMENT

For DGPS applications, several GPS receivers are necessary. At least two receivers must be activated at the same time. Two types of receivers are used for this investigation: One frequency GPS-receivers and precise geodetic GNSS two frequency receivers. Instead of using two GPS receivers on location, a reference station can also be realized by a satellite positioning service. These services (here: SA*POS*®) offer data of fixed reference stations. They have been utilized to evaluate the results in this investigation.

Next to the GPS receivers, a total station and other additional equipment have been used.

2.1 GPS component

2.1.1 Low-cost GPS receiver

The low-cost elements are two ANTARIS GPS-Timing Evaluation Kits (AEK-4T) from u-blox. This kit exists of an antenna AMN-MS, the receiving module LEA-4T, accessories, and the software u-center to record the measured data (**Figure 6**). The module has a maximum position update rate of 10 Hz to get raw data. UBLOX (2011a) explains that this kit can register only the L1-signal of GPS and that it is using 16 channels. Normally, only a code solution is available for navigation processes. But it is possible to store raw data of carrier frequency L1 on a computer via usb-interface. Officially, they are used to smooth the code signal. But here, they are used to get a phase solution in the post-processing evaluation. AEK-4T itself cannot store any data. The software u-center must be installed on a computer, which is connected to the GPS module. It can store all raw data in a selected file format. SCHWIEGER (2009) describes how the data are stored in the native u-blox format *ubx*.

Figure 6: AEK-4T Evaluation Kit of u-blox, UBLOX (2011)

The quality of the measured trajectory is judged by comparing it to a reference trajectory. Therefore the GPS antenna must be placed centered on a 360° prism. As there is no special adapter to fix the antenna on such prisms, an adapter has been constructed by IIGS as it is mentioned in SCHWIEGER (2009). For this research, a shielding plate was added to this adapter as it is shown in **Figure 7**.

The quality of data measured by the low-cost GPS receivers is supposed to have nearly the same accuracy level as geodetic GNSS receivers. For that reason only short baselines should be realized and good ionosphere models must be used. So the influence of errors is minimized.

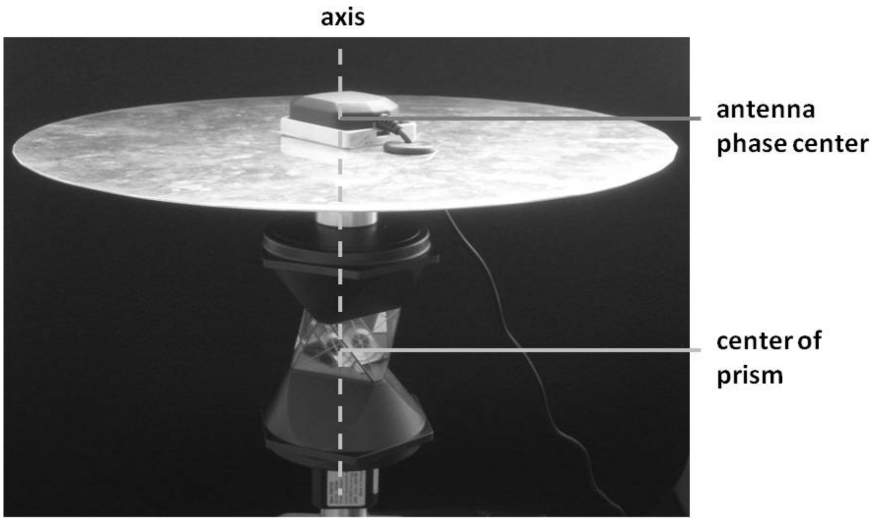


Figure 7: Test arrangement - u-blox antenna with shielding plate, mounted on a 360° prism

2.1.2 GNSS receiver

The used geodetic GNSS receivers are products of the Leica system 1200 (GPS1200+ GNSS). These receivers are able to work with three GPS frequencies (L1, L2 and L5) and also with GLONASS and, in future, Galileo. They have a maximum measuring frequency of 20 Hz and are using up to 120 channels (LEICA 2011). For this project, only the GPS signals are used. The GNSS antennas have already a matched thread; they can be fixed directly on a screw fitting or on a tripod adapter. The following accuracies in **Table 1** are given by manufacturer and can be achieved with using DGPS configurations. The horizontal accuracy of this system according to the length of baselines is shown in **Figure 8**.

Table 1: Accuracy Leica GPS1200+ according to LEICA (2011)

	static mode	**kinematic mode**
horizontal	5 mm + 0.5 ppm	10 mm + 1 ppm
vertical	10 mm + 0.5 ppm	20 mm + 1 ppm

2.1.3 DGPS service

For DGPS application, reference stations of the DGPS service SA*POS*®[1] were used. SA*POS*® offers different services according to the desired accuracy and the timing mode: Real-time or Post Processing. Here, a post processing procedure was chosen and the geodetic post processing service (GPPS) was used. This GNSS service actually offers a horizontal accuracy of 1 cm and different clock rates (1 s to 60 s according to SAPOS-BW (2011)). Data can be downloaded in RINEX format. This service is fee required. The procedure is described on SAPOS (2011).

[1] *Satellite Positioning Service of the German State Survey* ("Satelliten*positionierungsdienst* der deutschen Landesvermessung")

Not only the real reference stations can be used; there is also the possibility to create virtual reference stations (VRS) in the surveying area. For this purpose, the user sends approximated reference coordinates inside the surveying area to the SA*POS*® computer center. There, the influence of error is calculated for this position by using data of real SA*POS*® reference stations and a correction model (SAPOS-BW, 2011), (BAUER, 2011). The corrected data of the approximated position are sent back to the user and can be evaluated as reference data. VRS are used to realize short baselines if SA*POS*® reference stations are too far away.

2.2 Total station

A total station was used to track an independent reference trajectory. Here, a Leica TCRP 1201 total station was used. This total station has an angular accuracy of 0.3 mgon. The distance accuracy is 2 mm + 2 ppm in static mode and 5 mm + 2 ppm in tracking mode according to LEICA (2004). Hence, the horizontal accuracy of a measured point depends on the distance between instrument and reflector as it is shown in **Figure 9**. In tracking mode, a total station needs less than 0.15 s for one measurement (LEICA 2004). In fact the mean measuring frequency is approximately 10 Hz in this case.

A prism was placed on the roof of the vehicle, so the trajectory of the vehicle can be monitored by focusing the prism. With the used 360° prism, the position of a moving object, like a vehicle, can still be tracked, also if the driving direction is changed. The assembled prism (Leica GRZ122) has a screw fitting on top, so a GPS antenna can be fixed on it.

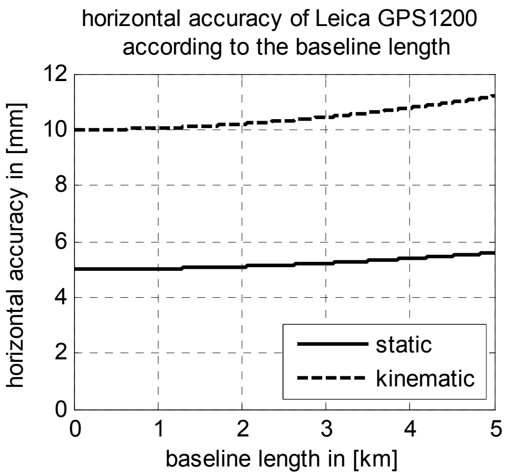


Figure 8: Horizontal accuracy of GPS points according to the baseline length, measured with Leica GPS1200+

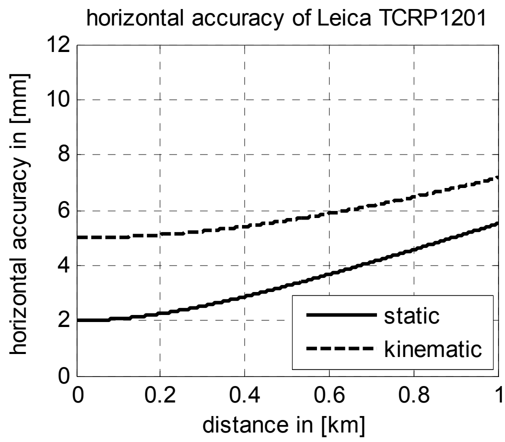


Figure 9: Horizontal accuracy of Leica TCRP1201 depending on the measured distance (without accuracy of the station)

2.3 Recording equipment

To record the measured data, the total station and the low-cost GPS receivers were connected to a Parasonic Toughbook via usb-cable. The u-blox receivers were controlled with software u-center. The native ubx-format cannot be used for baseline analysis, so the data are transferred to the independent RINEX format. For this purpose, the freeware TEQC is used.

The measurement process of the total station is controlled by a LabView™ Tool which is able to store the measured data in a database and also to synchronize the internal time of the total station with the GPS time. The total station as well as GPS receivers are storing the time of measurement. It is possible to get measurements at comparable time epochs.

3 MEASUREMENTS

To place the total stations in space, a free stationing was done using some fixed points. These fixed points have been created with geodetic GNSS receivers (Leica GPS1200+ GNSS) in static measurements. They are not only used as connecting points for the total station, but also as station for local reference stations. BAUER (2010) determines the accuracy of these points with approx. 5 mm, according to the measuring time.

Both measuring systems, the GPS receivers and the total stations, must use synchronized time. At each position, points of the GPS trajectory can be compared to single reference positions. Therefore every computer was connected to a u-blox antenna and synchronized with the PPS signal (Pulse per second). This signal is included in the GPS signal. It is useful to synchronize the time before each measuring session because the computer time is drifting with regard to their charge conditions. The GPS receivers have measured with a frequency of 1 Hz.

Three drives have been carried out to test the quality of a low-cost GPS-receiver as a rover for kinematic evaluation. Different types of trajectories have been driven each time: straight lines, U-turns and sinusoids (see **Figure 10**)

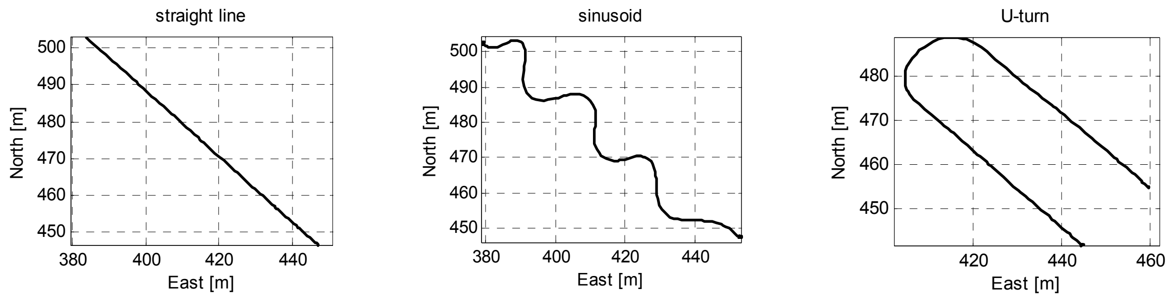


Figure 10: Different types of driven trajectories

At the beginning and at the end of every session, the GPS receiver should stay on a position for a while to get a precise and reliable static position determination. Therefore, the vehicle stopped moving for initialization before a measurement session started and after it was finished. The complete measurement periods of the single trajectories varied between 12 minutes and more than 2 hours.
For the evaluation, a SA*POS*® reference station (VRS) and a local reference station, realized by a u-blox receiver, were used. The distances to the next fixed reference station are less than 10 km. It would also be possible to use this original reference station for evaluation.

The GPS antenna (rover) was mounted on the roof of a vehicle. To compare the data quality of the low-cost rover to that of a geodetic GNSS receiver, the u-blox antenna was replaced by a GPS1200+ receiver for some drives.

Based on the test preparations, the rover was not placed directly on the roof. Multipath effects are expected to have a strong influence on the position quality of a u-blox antenna. To test those effects, the u-blox antenna has been shielded with respect to ground reflections by a plate for some drives only.

4 EVALUATION

In BAUER (2010), several software packages (RTKLIB, Leica Geo Office, Wa1) were tested for the Post Processing evaluation of the kinematic DGPS data. Finally, software Wa1 of WaSoft (WASOFT, 2011) was chosen to get a baseline solution for this application. As it is described in BAUER (2010), this software solves the ambiguities in the best way and provides the best data quality. For a detailed analysis of the measured points, the software Matlab® was used. Characteristics to judge the quality of the evaluation process are the lateral and the longitudinal deviation, because only the horizontal coordinates were considered. The quality of the trajectory is evaluated by the following parameters: Mean value of lateral deviation $\overline{q}$ and the scale factor m. They are calculated by comparing them to a trajectory measured by the total station. This reference trajectory is smoothed to eliminate outliers.

$$\overline{q} = \frac{1}{n}\sum_{i=1}^{n} q_i \tag{1}$$

$$m = \frac{s_{\text{total station}}}{s_{\text{GPS}}} \tag{2}$$

$s_{\text{total station}}$: distance in between two points measured by total station
s_{GPS}: distance between two points measured by GPS

Because of the different measuring frequencies of the instruments, points were interpolated on the trajectory of the total station corresponding to the time of GPS measurements (see **Figure 11**).

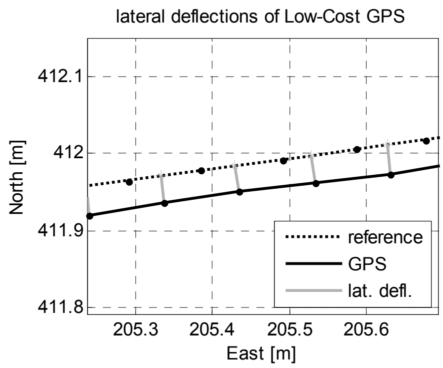


Figure 11: Lateral deviations of low-Cost GPS measurements compared to the reference trajectory

The first evaluation was concentrated on the use of a shielding plate for the u-blox rover. Without shielding plate, the GPS-trajectories have an offset and are rotated, compared to the reference trajectory. The differences between the reference trajectory and the GPS trajectory are in general larger than 1 dm (**Figure 12**). If one considers only the shape of a trajectory, a good relative accuracy can be shown. This purely relative accuracy is comparable to the absolute accuracy of the trajectory in case a shielding plate is used. On average the mean lateral deviation is worse by 66% compared to the same average value measured with shielding plate according to BAUER (2010). The deviation of the scale factor from an ideal factor 1.0 is larger by a factor 13.4 than with shield. The shielding plate avoids offsets and rotations of the GPS trajectory. The causes for these deviations will not be discussed in detail here. The authors assume multipath effects. When it is used, a good absolute accuracy compared to the geodetic GPS receivers is achieved. Both types of receivers are evaluated with a real local reference station and a VRS using SA*POS*®.

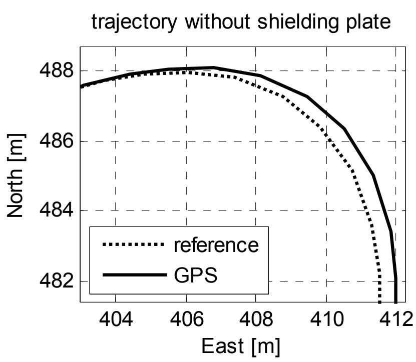


Figure 12: GPS measurements without shielding plate - translated and rotated GPS trajectory

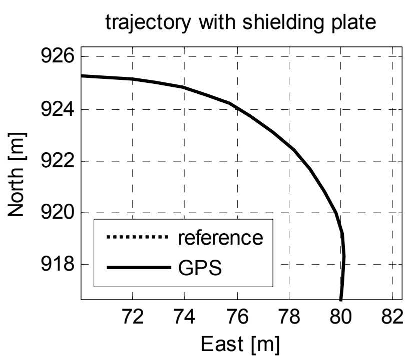


Figure 13: GPS measurements with shielding plate

Table 2: Factors of deviation - results of GPS measurements compared to the accuracy of the used total station

Reference	Rover	Configuration	$\overline{q}\,/\,\sigma_{\text{total station}}$	*m*
VRS	u-blox	without shielding plate, relative accuracy	5.8	0.9995956
	u-blox	with shielding plate, absolute accuracy	3.5	0.9999697
	u-blox	with shielding plate, 10 min for initialization, absolute accuracy	4.9	0.9999711
	Leica	absolute accuracy	3.5	1.0000188
u-blox	u-blox	with shielding plate absolute accuracy	6.0	1.0000715
	Leica	absolute accuracy	3.5	0.9999697

Compared to a geodetic GNSS receiver, the u-blox kit, evaluated with a VRS as reference station, reaches nearly the same position quality, if a shielding plate is used. The mean lateral deviations of both positioning systems are nearly the same value; the difference is less than 1% as it is shown in **Table 2**. If also its root mean square (RMS value) is considered, it is only 24% worse than that of the Leica GNSS system (BAUER, 2010). The results of a u-blox receiver are more scattered. The deviation from the ideal scale factor 1 for the u-blox receiver is 61% worse compared to the deviation of the geodetic GNSS receiver. For this evaluation results, a measuring epoch of more than two hours was evaluated. After a long period of initialization, the vehicle was moved for one hour.

In a next step, the initialization time was abbreviated step by step in BAUER (2010) by taking only specified epochs of the original epoch. For initialization epochs of 60 to 10 minutes, the mean value of lateral deviations is changing in general by less than 41% compared to the full epoch of initialization time. For the scale factor, the deviation is even less than 5%. This means, that in most cases the duration of initialization time does not have a large effect on the lateral deviation or on the scale factor. Also if the trajectories were split into several shorter parts, the initialization time does not affect the deviations a lot. In comparison to the original value, the mean lateral deviation is changing by less than 50% and the deviation to the ideal scale factor 1.0 is larger up to factor 2.5. Those variations and some outliers are originated by a bad satellite configuration. Bad satellite configurations can be detected by the PDOP value (Position Dilution of Precision) which is available in cases with using u-center. In those cases the mean lateral deviation and the scale factor are much worse than those with a longer driving time.

The horizontal accuracy of the trajectory is still good, if a second u-blox receiver is used as a local reference station. The deviations are better than the required value of 1 dm as it is described in **Table 2**. In BAUER (2010), the result of the u-blox receiver evaluated with a local reference station is worse by 71% compared to the trajectory with SA*POS*® as reference station. Regarding the scale factor, the deviation from factor 1.0 is worse by factor 2.4 compared to the deviation with SA*POS*®. If the results of geodetic GPS receivers evaluated with a local u-blox reference are compared to those evaluated with SA*POS*®, they have the same mean value of lateral deviation. The deviation is only 1%. The deviation from the ideal scale factor 1.0 is worse by factor 1.6 compared to a reference service.

5 CONCLUSIONS

A low-cost GPS receiver like the u-blox AEK-4T is applicable for kinematic position determination in post-processing mode. For this, adequate software is needed. Here, TEQC and Wa1 were used for data transfer and GPS baseline evaluation. The results of a low-cost GPS receiver evaluated with Wa1 have almost the same accuracy than positions measured with a geodetic GNSS receiver. To achieve such results, the use of a shielding plate is necessary for the u-blox antenna to reduce multipath effects. Another possibility would be to place the antenna directly on the roof of a vehicle. As reference station either reference services like SA*POS*® or a local reference station can be used. For post-processed kinematic evaluations with reference stations of a DGPS-service, the u-blox LEA-4T module reaches approximately the same mean lateral deviations and a 60 % worse longitudinal deviation than a Leica GPS1200+ receiver. In comparison to the total station, the lateral deviation is worse by factor 3.5. It is even possible to use a second low-cost receiver as reference station, if it is positioned on a point with known coordinates. The lateral deviation differs by a factor of 6.0 from the accuracy achieved by a Leica total station that was used for reference measurements during the same setup. Even if those factors of deviation seem to be large, it must be considered that all parameters are within the required deviation of 1 dm. Therefore, the low-cost GPS receivers reach a good position quality in this research. These good results can be achieved because of the short baselines. The influence of the ionosphere is nearly the same on both observation points.

Low-cost GPS receivers like the u-blox receiver are only single frequency receivers. They react more sensitive to bad measurement conditions. Bad satellite configurations, detectable by PDOP value or different conditions on the ground affect the quality of position much more than with a geodetic two frequency receiver, since the antenna gain pattern and the evaluation algorithms are better suited for high quality measurements.

There is no special requirement concerning a defined initialization time before and after driving, the evaluated data do not differ much in their accuracy level. But it is recommended to have anyway a short stop period at the beginning and at the end of the trajectory. It avoids that the measurements are starting or ending in a period of bad satellite configuration like it is described in BAUER (2010). An accurately and reliable determined start and end point increase the accuracy of the kinematic positions.

ACKNOWLEDGEMENTS

The authors would like to thank Ms. Nina Bauer for preparing, realizing and analyzing the measurements within her diploma thesis.

REFERENCES

Books:

BAUER, M.: *Vermessung und Ortung mit Satelliten – Globale Navigationssatellitensysteme (GNSS) und andere satellitengestützte Navigationssysteme.* Wichmann, Berlin, 2011.

BAUER, N.: Diploma thesis at the Institute for Applications of Geodesy to Engineering (today: Institute for Engineering Geodesy), Universität Stuttgart, not published, 2010.

GLABSCH, J., HEUNECKE, O., SCHUHBÄCK, S.: *Überwachung von Rutschhängen mittels Low-Cost-GNSS-Empfängern im Near Real Time Processing.* Int. Kurs für Ingenieurvermessung 2010 in München, Wichmann, Berlin, 2010.

GLABSCH, J., HEUNECKE, O., PINK, S., SCHUHBÄCK, S.: *Nutzung von Low Cost GNSS Empfängern für ingenieurgeodätische Überwachungsaufgaben.* GNSS 2010 – Vermessung und Navigation im 21. Jahrhundert, 100. DVW-Seminar, 4./5. Oktober 2010 in Köln, Wissner, Augsburg, 2010.

LEICA GEOSYSTEMS: *LEICA TPS1200 Serie – TPS1200 Technische Daten.* 2004.

Journal articles:

SCHWIEGER, V.: *Accurate High-Sensitivity GPS for Short Baselines.* FIG Working Week, Eilat, March 3-8, 2009.

LIMPACH, P.: *Rock glacier monitoring with low-cost GPS: Case study at Dittu glacier, Mattertal.* AHORN 2009, 5./6. November 2009 in Zürich, 2009.

Links:

http://www.igp.ethz.ch/AHORN/downloads/PhilippeLimpach_AHORN2009.pdf, last accessed on November 04, 2011

http://www.u-blox.com/en/download/documents-a-resources/antaris-4-gps-modules-resources.html, last accessed on November 21, 2011

http://www.u-blox.com/images/downloads/Product_Docs/LEA-6_DataSheet_%28GPS.G6-HW-09004%29.pdf, last accessed on November 21, 2011

http://www.leica-geosystems.com/downloads123/zz/gps/general/brochures/GPS1200_brochure_de.pdf last accessed on November 21, 2011

http://www.sapos.de, last accessed on November 21, 2011

http://www.sapos-bw.de/, last accessed on November 21, 2011

http://www.wasoft.de/e/wa1/index.html, last accessed on November 21, 2011

Global Navigation Satellite Systems II

This Paper has been accepted as a "peer reviewed paper" by the scientific committee

Single-Epoch Ambiguity Resolution for Kinematic GNSS Positioning

Christian Eling, Philipp Zeimetz, Heiner Kuhlmann
Institute of Geodesy and Geoinformation, University of Bonn, Germany

Abstract
Automatic machine control requires accurate and reliable information about the latest attitude and position of the vehicle. In addition to inertial sensors and odometer Global Navigation Satellite Systems (GNSS) are well established in the determination of these parameters.

Besides code observations GNSS additionally provide carrier-phase measurements, which should be used to achieve high accuracies. Certainly, the key to GNSS carrier-phase positioning is the ambiguity resolution. This is the process resolving the unknown number of integer cycles in the carrier phase data. Principally, different approaches exist to resolve the ambiguities. Since multi-epoch techniques lead to a substantial loss of possible solutions, a single-epoch ambiguity resolution should be aimed at. A common procedure that enables an ambiguity resolution for every single epoch is the Ambiguity Function Method (AFM). By means of a cost function the AFM tests candidates corresponding to a generated search space, including possible rover positions. However, by use of this approach disadvantages occur due to the computation time, increasing with the size of the search space, and the reliability, depending on the decidedness of the complicated multipeak-function. Accordingly, the candidates in the search space have to be selected carefully. For this purpose, position approximations can be achieved by use of GNSS-velocities, a vehicle motion model, differential code-solutions as well as Kalman filtering. Therefore, through the combination of these tools it was possible to develop a single-epoch ambiguity resolution algorithm that also shows good performances in urban areas with a success rate of 96.59 %.

Keywords
GNSS, ambiguity resolution, Kalman filtering, GNSS velocities

1 INTRODUCTION

In the field of automated control as well as supported navigation of machines Global Navigation Satellite Systems (GNSS) are of major importance, since they allow for the determination of absolute positions, attitudes and velocities. On this account most of the land- and construction-machines are nowadays already equipped with at least one GNSS antenna. Besides pseudoranges GNSS also provide more precise carrier phase measurements. Within differential positioning based on double differenced carrier phases, sub-centimetre-level precision GNSS positioning becomes possible. However, it is well known that GNSS double differenced carrier phase measurements are ambiguous by an unknown number of integer cycles. To fully exploit the high accuracy of the carrier phase observables, the ambiguities must be resolved to their correct integer value (Hoffmann-Wellenhof et al., 2008). Especially in urban areas obstacles like street canyons, bridges or vegetation lead to frequent losses of lock, which always necessitate a new ambiguity resolution. Therefore, for kinematic applications, the duration of the ambiguity resolution is of particular importance.

In this contribution we will first give a short overview of the existing ambiguity resolution techniques. Afterwards we will present a single-epoch ambiguity resolution method for kinematic positioning, which is based on the combination of approaches like GNSS velocity determination and Kalman filtering with an instantaneous ambiguity resolution technique. By means of test runs, the procedure was tested successfully.

2 AMBIGUITY RESOLUTION BACKGROUND

In the last decades integer ambiguity resolution was the focus of many researchers, since the ambiguity resolution is the key to precise GNSS positioning. Therefore, many different techniques exist to determine the unknown integer cycles of the observed double differences. Generally, these approaches can be classified into three categories: ambiguity resolution in the measurement domain, search technique in the ambiguity domain and search technique in the coordinate domain (Kim and Langley, 2000).

The ambiguity resolution in the measurement domain is principally based on code observations. Since code observations are more inexact than carrier phase measurements, these approaches are ordinarily not very suitable. Only the processing of interfrequency linear combinations enables a more or less reliable ambiguity resolution. However, this requires the observation of at least 2 frequencies.

The second class of ambiguity resolution techniques contains approaches searching in the ambiguity domain. Generally, they are based on the so called integer least squares (ILS) theory (Teunissen, 1993). The ILS-approaches consist of three steps. By means of a float solution the cycles are estimated via a least squares adjustment. Using the resulting float ambiguities and the variance-covariance matrix the odd number of ambiguities can afterwards be fixed within a search process in the integer ambiguity estimation step. As soon as the ambiguities are set to integer values a fixed solution follows to determine the precise baseline parameters. The most famous and reliable ILS ambiguity resolution technique is the LAMBDA method (e.g. Teunissen, 1995). Further well known approaches are the FASF (Chen and Lachapelle, 1995), the FARA (Frei and Beutler, 1990), the OMEGA (Kim and Langley, 1999) and the LSAST (Hatch, 1990) method. Except the LSAST method, all of these search techniques are multi-epoch approaches. This is because the float solution step necessitates the usage of observations from more than one epoch, since the number of parameters definitely exceeds the number of available double differences in one single epoch.

The third class of ambiguity resolution techniques contains approaches searching in the coordinate domain. The most famous of these methods and simultaneously one of the earliest ambiguity resolution search techniques in general is the Ambiguity Function Method (AFM) (Counselman and Gourevitch, 1981; Remondi, 1984; Mader, 1990). On the basis of appropriate criteria, candidates of a predefined search space have to be tested, only regarding the fractional part of the observed carrier phases of one single epoch. In this paper, this method is the basis for further investigations.

3 INSTANTANEOUS AMBIGUITY RESOLUTION

In case of kinematic applications the rapidity of the ambiguity resolution is of particular importance. Therefore, fixing the integer number of cycles of the observed double differences within one single epoch should be aimed at. Moreover, not only long term signal interruptions appear very often during kinematic applications. Even cycle slips and interruptions to the signal between epochs occur, which result in new sets of integers (Corbett and Cross, 1995). To avoid additional processing during cycle slip detection a single epoch ambiguity resolution is inevitable.

Since the AFM is not in need of float ambiguities or a variance-covariance matrix it is well-suited for instantaneous ambiguity resolution and therefore also resistant to cycle slips. For these reasons, we now like to introduce the basic principle of the AFM.

As mentioned above, the ambiguity search using the AFM takes place in the coordinate domain. Maximizing the Ambiguity Resolution Function (ARF) enables the assessment of candidates of a predefined search space, containing possible rover positions.

The ARF, for the single frequency case, can be written as (Lachapelle et al., 1992):

$$ARF\left(X_C, Y_C, Z_C\right) = \sum_{m=1}^{N-1} \cos 2\pi \left(\nabla\Delta\phi_{obs}^{kj}[E1 \mid E2] - \nabla\Delta\phi_{calc}^{kj}[E1 \mid X_C, Y_C, Z_C]\right) \tag{1}$$

whereas $\nabla\Delta\phi^{kj}{}_{obs}$ is the observed and $\nabla\Delta\phi^{kj}{}_{calc}$ a calculated double difference of the satellites *k* and *j*. Since *E1* is the known position of the master antenna, the ARF is only dependent on the coordinates of the candidates, which are located in the search space around the true rover position *E2* (see **Figure 14**). In case the position of the candidate (X_C, Y_C, Z_C) is similar to *E2*, the difference between the observed and the calculated double difference corresponds to the unknown ambiguities in order that the result of the cost function is equal to 1. Considering the carrier phase measurements of *N* satellites on *w* frequencies the outcome of the ARF would ideally be *(N-1)·w*, taking into account that due to multipath and receiver noise this maximum will never be reached.

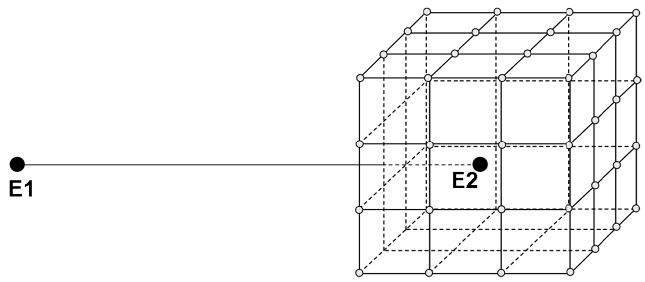


Figure 14: Depiction of a possible search space for the AFM

However, there are two drawbacks of the AFM. First, the computational efficiency is highly dependent on the size of the search space, which is defined by the accuracy of the approximate position. Therefore, the computation time can potentially be very long. And second there may be several maxima points that the AFM must discriminate between to find the optimal solution (Han and Rizos, 1995). Consequently, the basic AFM approach has to be improved to make it suitable for practice.

4 EFFICIENCY IMPROVEMENT OF THE AFM

The key to reliable and fast ambiguity fixing by means of the AFM is a reduction of the size of the search space. In so doing, the computation time decreases heavily. Furthermore, false candidates can be excluded with the result that the decision-making of the AFM will be simplified.

In the creation of the search space two cases can be decided. On the one hand the first time initialization or re-initialization after a GNSS gap and on the other hand the transfer between two epochs as long as GNSS is available.

4.1 First-time initialization or re-initialization of the ambiguities

In the beginning of an application as well as in consequence of a gap of the GNSS signals, a first-time initialization or a re-initialization of the ambiguities is necessary. The starting point for this search process is an approximate rover position to generate a search volume. Without the use of additional sensors, there are only few opportunities for determining these preliminary coordinates. In most cases code observations are preferably used to cope with this task. By means of differenced code signals accuracies in the range of a few decimetres to metres are achievable. Linear combinations such as the wide lane also enable the determination of an approximate position (Abidin, 1994). However, this requires the observation of at least two frequencies, which is not always given, e.g. low-cost receivers. Therefore, we use differential code-observations for the determination of an approximate rover position in case of first-time or re-initialization of the ambiguities. Even if the ambiguities could be fixed correctly in the first epoch after a loss of lock, this initialization process will be delayed for more than one epoch to avoid an incorrect first-time or re-initialization. During this time the preliminary coordinates are filtered in an Extended Kalman Filter (EKF) to improve the reliability of the ambiguity resolution (see chapt.5).

In order to consider the balance between the computational effort of the search process and the size of the search space, which has to be large enough to contain the true rover position, the configuration of the candidates should be carefully selected. Since the candidates vary on the basis of different ambiguities, the possible rover positions in the search space are also dependent on different sets of integer ambiguities. Therefore, the approximate rover position is used to determine the ambiguities of the observed double differences, which can be rounded to integer cycles. Afterwards these ambiguities need to be varied for different values to determine possible rover positions. In order to limit the number of candidates in the search space, not all double difference ambiguities, but only three should be used for generating the search space. The selection of this three primary observations occurs in consideration of the position dilution of precision (PDOP), the elevation angles as well as residuals of the code solution. In case the range of the ambiguities is set from minus five to plus five values the search volume consists of 1331 ($[(2*5)+1]^3$) sets of ambiguities, which lead to possible positions, including the true rover position (Corbett and Cross, 1995). Depending on the PDOP, the edges of the cubic search space reach lengths up to 10 m. Therefore, the search space is first defined in the ambiguity domain before it is used to generate a physical space, defined in the coordinate domain. Finally, the ambiguity resolution of all double differences occurs by testing the candidates by means of the AFM. Since the ARF is a multi peak function the results are not inevitably unambiguous (see chapt.4.3). In case the maximum of the ARF cannot be clearly distinguished from side-lobes, further investigations are necessary to find the correct set of ambiguities. Criteria are for example the variances of a least squares adjustment in the determination of the baseline parameters.

4.2 Position update using GNSS-velocities

Once the ambiguities were fixed and at least four GNSS satellites are visible in two successive epochs, the determination of an approximate rover position can occur by use of integrated GNSS-velocities in combination with a Kalman filter. This is because the GNSS-velocities are also based on precise carrier phase measurements, whereas they are not in need of an ambiguity resolution. This becomes obvious regarding the observations used for the velocity determination, consisting of the first order difference approximation of the carrier-phase observations (Serrano et al., 2004):

$$\dot{\phi}_i^j \approx \frac{\phi_i^j - \phi_{i-\Delta t}^j}{\Delta t} \tag{2}$$

where ϕ is the fractional carrier-phase of the satellite j, i is the observation epoch and Δt the sampling rate. The velocity determination occurs by use of a least squares adjustment based on the following objective function:

$$\dot{\phi}_i^j = \boldsymbol{h}_i^j \cdot \left(\boldsymbol{v}_i^j - \boldsymbol{V}_i\right) + \dot{B}_i + \varepsilon_i^j \tag{3}$$

Where $\boldsymbol{V}$ is a vector containing the unknown receiver velocities (V_x, V_y and V_z). $\boldsymbol{v}$ represents the satellite velocity vector, which can be computed by use of the ephemeris. $\boldsymbol{h}$ stands for the directional cosine between the receiver and the satellite:

$$\boldsymbol{h}_i^{jT} = \frac{\left(\boldsymbol{S}_i^j - \boldsymbol{X}_i\right)}{|\, \boldsymbol{S}_i^j - \boldsymbol{X}_i \,|} \tag{4}$$

whereas $\boldsymbol{S}$ is the position vector of satellite j and $\boldsymbol{X}$ the receiver position vector. Furthermore, $\dot{B}$ are the receiver clock drift and ε the receiver noise.

By means of equations (2)-(4) the receiver velocity can be determined in every observation epoch *i*. Of course the resulting velocities are not the true ones for the actual epoch, since the first order difference approximation in equation (2) enables the estimation of the mean velocity between the epochs *i* and *i-Δt,* but in case of sampling rates higher than 1 Hz, they are still accurate enough to deliver a suitable approximate position for the AFM.

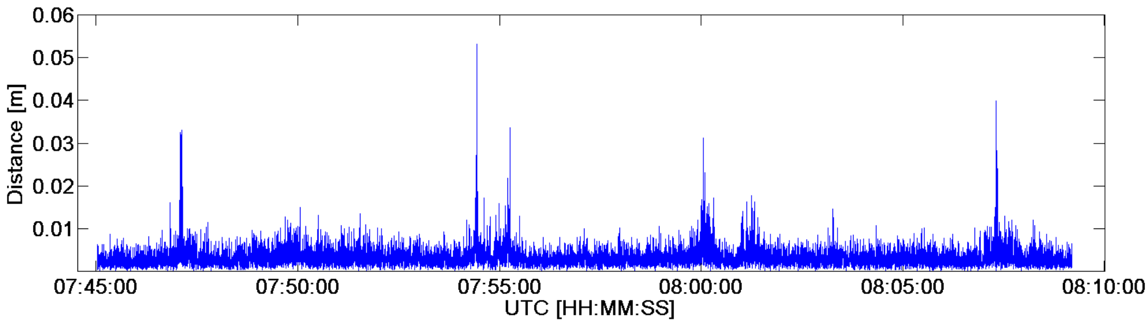


Figure 15: Distances between integrated GNSS velocity positions and final positions.

To underline this, the distances between the approximate positions, calculated by integration of the velocities, and the final positions during a kinematic experiment are presented in **Figure 15**. Except for a few outliers, the total deviations are mostly less than 1 cm. The mean of the distances between the approximate and the final positions is about 3 mm. The clearly visible outliers are based on poor GNSS measurement conditions. However, these deviations are less than 6 cm whereas the wavelengths of GNSS signals are in the order of 20 cm. According to this, the filtered approximate positions are still well suited to reduce the size of the search space. Therefore, as long as GNSS is available the GNSS velocities are used to define the search volume.

4.3 Reducing the number of maxima in the AFM

One drawback of single-epoch ambiguity resolution approaches is the susceptibility in case of poor GNSS measurement conditions. Biases like multipath, residual atmospheric effects and satellite orbit errors lead to deviations in the observed carrier phases (Kim and Langley, 2000). Therefore, it cannot be excluded that false ambiguities will be selected from the set of candidates in the search space, since the multi peak ARF does not allow for an unambiguous decision in such epochs. To improve the performance of single-epoch ambiguity resolution techniques approaches employing linear filters for the residuals or time averages for the objective function showed good performances in earlier studies (e.g. Borge and Forssell, 1994; Martin-Neira et al., 1995). Therefore, the decision-making of the AFM should also be improvable.

In our single-epoch ambiguity resolution approach we use a Kalman filter to predict the residuals of the observed double differences in every epoch. Since the residuals cannot be described by any motion behaviour a random-walk-process is applied as motion model in this procedure. Accordingly, the discrete system equation follows:

$$\begin{vmatrix} x_R(k+1) \\ \Delta x_R(k+1) \end{vmatrix} = \begin{vmatrix} 1 & \Delta t \\ 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} x_R(k) \\ \Delta x_R(k) \end{vmatrix} + \begin{vmatrix} \Delta t \\ 1 \end{vmatrix} \cdot w(k) \tag{5}$$

$$\boldsymbol{x}(k+1) = \boldsymbol{T}(k) \cdot \boldsymbol{x}(k) + \boldsymbol{S}(k) \cdot w(k) \tag{6}$$

whereas $\boldsymbol{x}(k)$ is the state vector, containing the filtered residual x_R and the derivative of the residual Δx_R. $\boldsymbol{T}$(k) is the transition matrix, $\boldsymbol{S}$(k) the system noise coupling and w(k) the system noise. By use of this Kalman filter the residuals of every observed double-difference can be predicted from epoch *i-1* to epoch *i*. Therefore, the performance of the ambiguity resolution can be improved by reducing the deviations of the observations in the AFM using the predicted residuals.

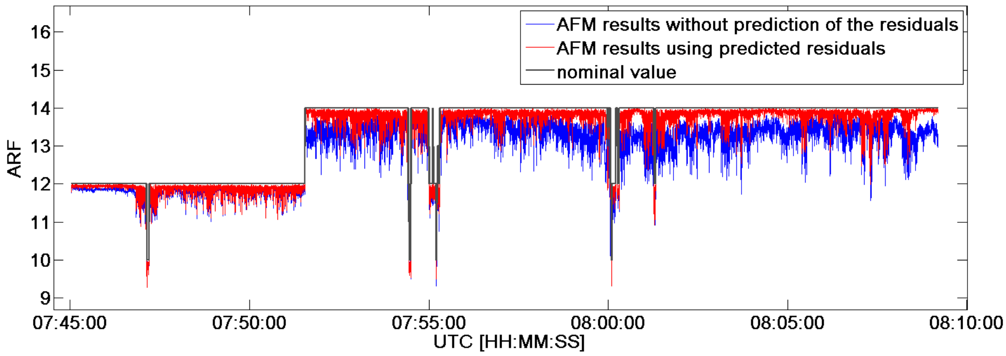


Figure 16: Comparison of results of the ARF using and disregarding the prediction of the residuals.

In **Figure 16** a comparison of the outcome of the AFM disregarding and using the residuals prediction during a kinematic test is presented. According to this, the maxima of the ARF increase by use of the Kalman filter, with the result that they obviously come closer to the nominal value of *(N-1)·w*. However, this does not imply a simplification in the discrimination of false and correct solutions in the search space, since a simultaneously increase of the side-lobes is also possible. In order to demonstrate the actual impact of the filter process, the outcome of the ARF for every candidate of the search space during one epoch is presented for two cases, using and disregarding the filtered residuals, in **Figure 17**. Therefore, not only the size of the maxima but also the difference to side-lobes increases by use of the prediction step. Summarizing this section, there are two reasons why the single-epoch ambiguity resolution has become better. On the one hand, the maximum of the outcome increases. And furthermore, the correct solution is now in greater contrast to incorrect sets of ambiguities. Therefore, there are no more investigations necessary to come to a decision, which candidate of the search space leads to the best and correct ambiguities.

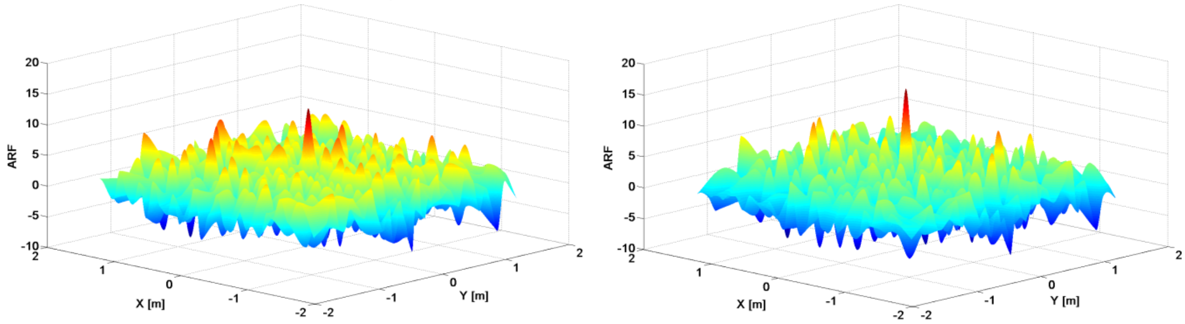


Figure 17: Results of the ARF disregarding (left) and using (right) the residuals prediction for one epoch.

5 EXTENDED KALMAN FILTER

As mentioned above, an EKF is used to improve the reliability of the determined rover positions in this system. Generally, an EKF is a recursive algorithm that enables the combination of noisy measurements with a priori known motion behaviour of a vehicle, to estimate an optimal state vector as time progress on the basis of external observations, (e.g. Grewel et al., 2007). Depending on the correctness of the used motion model the EKF is well suited to reduce white noise and to detect outliers. Since vehicles mostly move on streets, which are compiled of the basic elements straight, arc and clothoide, we assume a uniform circle movement as system dynamics model (Aussems, 1999; Eichhorn, 2005):

$$\bar{\boldsymbol{x}}_{k+1} = \begin{bmatrix} \begin{vmatrix} \hat{x}_k^G \\ \hat{y}_k^G \\ \hat{z}_k^G \end{vmatrix} + R_L^G \begin{vmatrix} \Delta East_k \\ \Delta North_k \\ \Delta Up_k \end{vmatrix} \\ \hat{\alpha}_k + \Delta\hat{\alpha}_k \\ \hat{v}_k \\ \Delta\hat{\alpha}_k \\ \Delta\hat{h}_k \end{bmatrix} \qquad \begin{aligned} \Delta East_k &= \frac{-\Delta t \cdot \hat{v}_k \cdot \left(\cos(\hat{\alpha}_k + \Delta\hat{\alpha}_k) - \cos(\hat{\alpha}_k)\right)}{\Delta\hat{\alpha}_k} \\ \Delta North_k &= \frac{\Delta t \cdot \hat{v}_k \cdot \left(\sin(\hat{\alpha}_k + \Delta\hat{\alpha}_k) - \sin(\hat{\alpha}_k)\right)}{\Delta\hat{\alpha}_k} \\ \Delta Up_k &= \Delta\hat{h}_k \end{aligned} \tag{7}$$

whereas *x,y,z* are the cartesian coordinates from the epoch *k*, *α* is the heading, *Δα* the heading change, *v* is the velocity in driving direction and *Δh* is the altitude change. R_L^G represents the rotation matrix to transform the coordinate changes *ΔEast, ΔNorth* and *ΔUp* from the local level frame to the global geocentric coordinate frame (Hoffmann-Wellenhof et al., 2008):

$$R_L^G = \begin{vmatrix} -\sin\lambda & -\sin\varphi\cos\lambda & \cos\varphi\cos\lambda \\ \cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi\sin\lambda \\ 0 & \cos\varphi & \sin\varphi \end{vmatrix}. \tag{8}$$

Thereby λ stands for the longitude and φ for the latitude.
According to this, the transition matrix $\boldsymbol{T}$ consists of the derivatives of equation (7) with respect to the states $\boldsymbol{x} = |x^G\ y^G\ z^G\ \alpha\ v\ \Delta\alpha\ \Delta h|^T$. The system noise coupling $\boldsymbol{S}$ contains the derivatives of equation (7) with respect to *v*, *Δα* and *Δh*, integrated over the sampling rate *Δt*, with the result that the discrete system equation (6) allows for accelerations, updates of the heading changes as well as updates of the altitude changes. The measurement equation (9) establishes the connection between the state vector $\boldsymbol{x}$ and the observations, given by the design matrix $\boldsymbol{H}$ and a white noise *ε*. The observation vector $\boldsymbol{l}=|x^{GPS}, y^{GPS}, z^{GPS}, v^{GPS}|^T$ consists of the GPS rover position and the rover velocity. This rover velocity is the norm of the East and North component of the transformed GPS rover velocity vector $\boldsymbol{V}$, which can be estimated by means of eq. (3). $\boldsymbol{D}$ is a unit vector $\boldsymbol{D} = |1\ 0\ 0|$.

$$\begin{vmatrix} x_{k+1}^{GPS} & y_{k+1}^{GPS} & z_{k+1}^{GPS} & v_{k+1}^{GPS} \end{vmatrix}^T = \begin{vmatrix} I_{3x3} & 0_{3x4} \\ 0_{1x4} & D \end{vmatrix} \cdot \begin{vmatrix} \bar{x}_{k+1}^G & \bar{y}_{k+1}^G & \bar{z}_{k+1}^G & \bar{\alpha}_{k+1} & \bar{v}_{k+1} & \Delta\bar{\alpha}_{k+1} & \Delta\bar{h}_{k+1} \end{vmatrix}^T + \varepsilon \tag{9}$$

Outliers, which are attributable to multipath or incorrect fixed ambiguities, can be detected by an innovation test. Besides the reduction of white noise as well as the detection of outliers, the system dynamics model of the EKF also allows for bridging GNSS gaps. In **Figure 18** two examples for this type of application are shown. In the left chart, the GNSS gap only lasts 25 epochs, with the result that the prediction is working very well. Therefore, in case of a short term GNSS outage, which can for example be caused by a tree, it is well suited to detect outliers, which are not uncommon during the first epochs after a signal interruption. However, in the right chart, a long term signal interruption is presented (146 epochs). In this case, the prediction does not completely agree with the true motion behaviour, in order that the EKF has to be restarted, once the ambiguities are fixed.

By adding further sensors, like gyroscopes or odometer, the bridging of long term GNSS gaps can also still be improved, in order that the deviations to the true position would only increase very slowly.

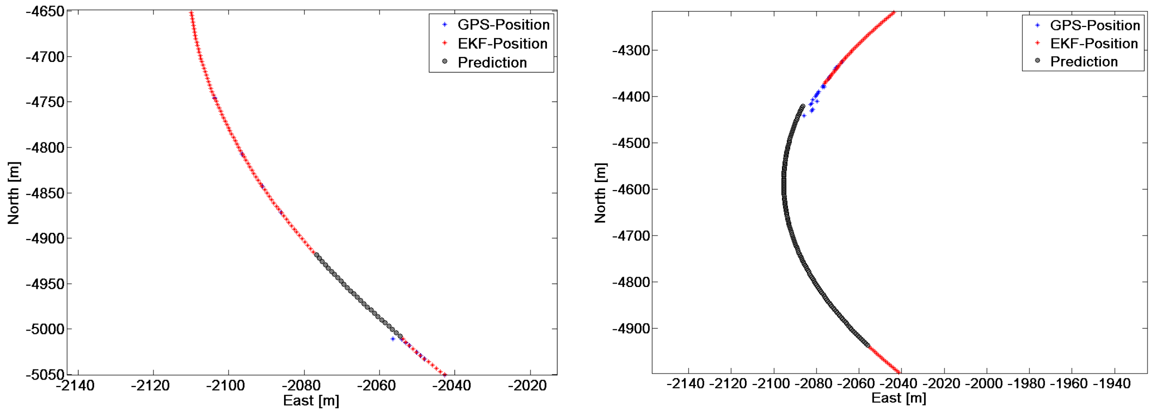


Figure 18: Prediction of the states during GNSS gaps by use of the system dynamics model.

The entire single epoch ambiguity resolution approach for the determination of GNSS positions during kinematic applications is presented in a flow chart in **Figure 19**. Accordingly, the ambiguity resolution either occurs by use of a Kalman filtered code position or a position update based on GNSS velocities. In case no GNSS observations are available, the states are predicted by the system dynamics model presented in equation (7).

6 RESULTS

The developed single epoch ambiguity resolution approach for kinematic GNSS positioning was tested by means of different experiments. In order to identify the correctness of the ambiguity resolution, the results were compared to commercial software. The outcome of this commercial software of a test run, carried out on freeways and city avenues in and near Bonn, is presented in **Figure 19**. Especially in urban areas this post processing led to less pleasant results, since most of the positions are based on imprecise code observations. The reason is that the commercial software could not fix the ambiguities quickly after each of the numerous signal interruptions. Therefore, this example emphasises the necessity of a single-epoch ambiguity resolution in case of kinematic applications, since a multi-epoch approach leads to long-term GNSS gaps after signal interruptions.

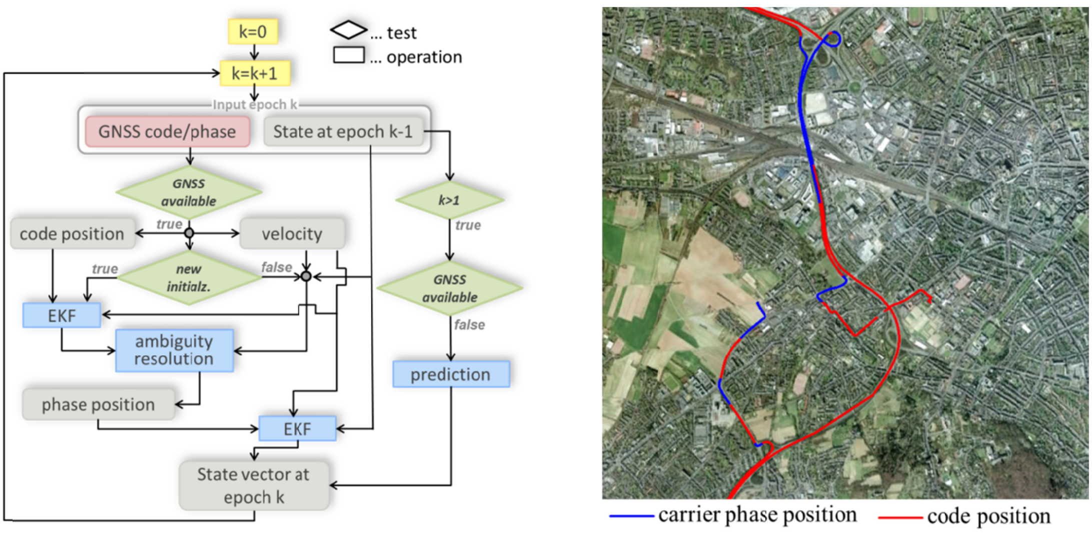


Figure 19: Flow chart of the developed approach (left) and test run processed by com. software (right)

The same dataset was also analyzed by means of the presented single-epoch approach. The sampling rate during this test run, in which speeds up to 140 km/h were reached, was 10 Hz. In **Table 3** the success rates of the ambiguity resolution for all of the 20294 epochs are shown. As expected, the reliability of the ambiguity resolution depends on the number of visible satellites. This is mainly because of the distinctness between correct and false sets of ambiguities in the AFM. The less observations are available the increased is the influence of multipath effects of individual signals. However, still 89.46% of the epochs, in times only four satellites were visible, led to correct ambiguity resolutions. In case more than six satellites were visible, the success rate increases 98 % whereas in 99.65% of all the epochs, where 9 satellites were visible, the ambiguities could be fixed correctly.

Table 3: Success rates of the ambiguity resolution by use of the single-epoch approach.

visible satellites	epochs	incorrect resolutions	success rate
4	2240	236	89.46 %
5	2320	183	91.98 %
6	3153	136	95.69 %
7	4492	85	98.11 %
8	4961	42	99.15 %
9	3128	11	99.65 %
sum	20294	693	96.59 %

Besides the success rates the time to fix the ambiguities is also of interest, since the objective of the single-epoch approach is to find the correct set of ambiguities as fast as possible after a signal interruption. In **Table 4** the number of epochs needed to find the true ambiguity resolution is presented for different re-initializations during the kinematic experiment. For comparison, the number of epochs elapsed until re-initialization is also shown for the outcome of the commercial software. In case of the single-epoch approach it should be noticed that not all of these epochs led to incorrect ambiguities, but rather the ambiguity resolutions were inconstant during these listed epochs.

Table 4: Comparison of epochs needed to re-initialize the ambiguities after GNSS gaps.

	epochs elapsed until re-initialization	
re-initialization	single-epoch approach	commercial software
1	2	52
2	3	324
3	27	85
4	2	37
5	3	51
6	3	58
7	2	58
8	37	20
9	12	50
10	10	403
11	20	47
12	14	85
13	7	38
14	5	206
15	2	57
16	10	73
17	2	278

According to this, the single epoch approach is mostly a lot faster than the commercial software. In many cases, it was possible to fix the ambiguities in the second epoch beyond a signal interruption. Especially in case of the second, the tenth, the fourteenth and seventeenth re-initialization the difference between both approaches is particularly evident, since the commercial software requires above two hundred epochs more to make the user carrier phase positions available again. Instead, the number of epochs elapsed until re-initialization is not of interest, if the mobile object is moving very slowly or the sampling rate is very high. Therefore, to illustrate the advantage of the developed ambiguity resolution procedure, the positions of the seventeenth re-initialization process are presented in the left chart of **Figure 20**. Whilst the single-epoch approach could fix the ambiguities in the second epoch, the commercial software only provided code positions for 277 epochs until the ambiguities could be fixed. During this time, the vehicle covered a distance of almost 400 metres.

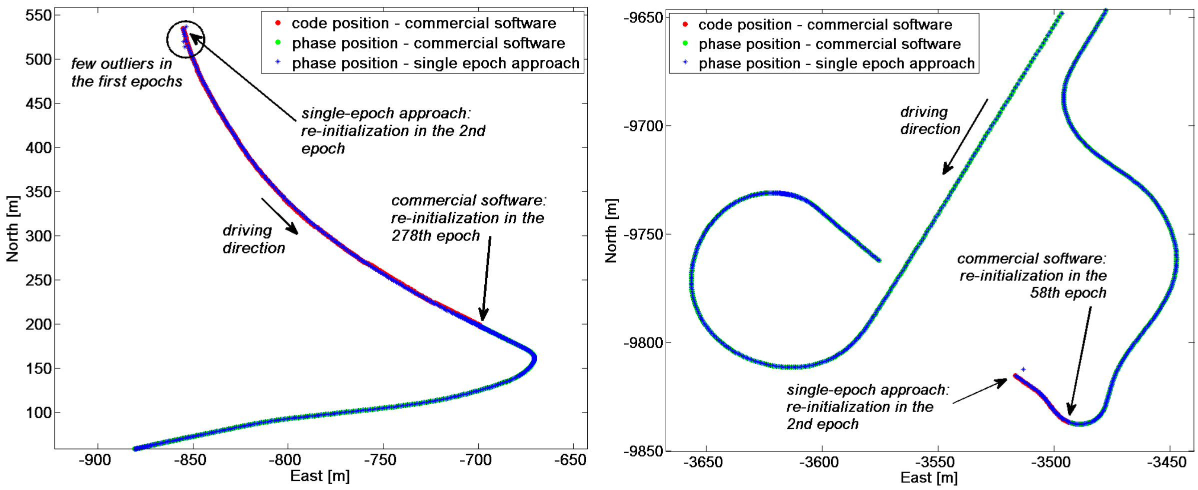


Figure 20: Comparison of results of the single-epoch approach to the outcome of a com. software.

However, it is also conspicuous that there are often few outliers in the first epochs after a signal interruption visible, which are based on poor GNSS conditions due to the proximity to the obstacle, which previously produced the loss of lock. Furthermore, the receiver generally needs different lengths of time to allocate the current carrier-phases. Hence, there are mostly only few observations available in the first epochs after a loss of lock. Certainly, regarding the Kalman filter as well as the variances of the positions, discontinuities become apparent. This is also true for the seventh re-initialization process, which is shown in the right chart of **Figure 20**. In the first epoch after the gap only three poor double differences were available with the result of an incorrect ambiguity resolution. Nevertheless, already one epoch later these ambiguities were fixed correctly, whereas the commercial software required 58 epochs (ca. 32 metres) to provide carrier-phase positions.

Concluding, by combining the approaches to limit the search space with the prediction of the residuals, a fast and reliable procedure to fix the ambiguities could be implemented. Furthermore, the procedure works independently of the application field (low speed in urban areas or high speed on freeways). In most cases only less than 5 epochs are required to find the correct ambiguity resolution, which leads to an ambiguity resolution success rate of 96.59%.

7 CONCLUSIONS

In this contribution we presented an ambiguity resolution approach for kinematic GNSS positioning. Especially in urban areas obstacles lead to frequent losses of lock, which always necessitate a re-initialization of the double-difference ambiguities. Therefore, we developed a single epoch ambiguity resolution to provide carrier-phase positions as fast as possible. In our approach we used the AFM for the determination of the integer number of unknown cycles. To overcome deficiencies caused by the computational efficiency of this procedure we used a combination of an ambiguity resolution in the ambiguity and the coordinate domain for first-time and re-initializations as well as GNSS velocities as long as no interruption occurred, to generate a well-suited search volume. Furthermore, by prediction of the residuals, the unambiguousness of the ARF could be improved. Considering experiments during different applications the approach was tested successfully. Despite frequently poor GNSS conditions an average ambiguity resolution success rate of 96.59% was reached. With few exceptions the integer cycles were re-initialized during the first 10 epochs after every signal interruption, in order that the approach leads to a reliable and fast ambiguity resolution.

REFERENCES

ABIDIN, H.Z.: *On-the-Fly Ambiguity Resolution*, GPS World, April 1994, p. 40-49, 1994

AUSSEMS, T.: *Positionsschätzung von Landfahrzeugen mittels Kalman-Filterung aus Satelliten- und Koppelnavigationsbeobachtungen*, Veröffentlichung des Geod. Institutes der Rheinisch-Westfälischen Technischen Hochschule Aachen, Nr. 55, 1999

BORGE, T.K., FORSELL, B.: *A new real-time ambiguity resolution strategy based on polynomial identification*, Proc. of KIS'94, Banff, Canada, p. 233-240, 1994

CHEN, D., LACHAPELLE, G.: *A comparison of the FASF and least-squares search algorithms for on-the-fly ambiguity resolution*, Navigation: Journal of the Institute of Navigation, Vol. 42, No.2, p. 371-390, 1995

COUNSELMAN, C.C., GOUREVITCH, S.A.: *Miniature interferometer terminals for earth surveying: ambiguity and multipath with Global Positioning System*, IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-19, No. 4, p. 244-252, 1981

CORBETT, S.J., CROSS, P.A.: *GPS single epoch ambiguity resolution*, Survey Review, Vol. 33, No. 257, 1995

EICHHORN, A.: *Ein Beitrag zur parametrischen Identifikation von dynamischen Strukturmodellen mit Methoden der adaptiven Kalman Filterung*, DGK, Reihe C, Heft 585, München, 2005

FREI, E., BEUTLER, G.: *Rapid static positioning based on the fast ambiguity resolution approach "FARA": theory and first results*, Manuscripta Geodaetica, Vol. 15, No. 4, p. 325-356, 1990

GREWAL, M.S., WEILL, L.R., ANDREWS, A.P.: *Global Positioning System, Inertial Navigation and Integration*, Second Edition, John Wiley and Sons, 2007

HAN, S., RIZOS, C.: *Improving the efficiency of the ambiguity function algorithm*, Journal of Geodesy (1996) 70, p 330-341, 1995

HATCH, R.: *Instantaneous ambiguity resolution*, Proc. of KIS'90, Banff, Canada, pp. 299-308, 1990

HOFMANN-WELLENHOF, B., LICHTENEGGER, H., WASLE, E.: *GNSS – Global Navigation Satellite System*, SpringerWienNewYork, Graz, 2008

KIM, D., LANGLEY, R.B.: *A search space optimization technique for improving ambiguity resolution and computational efficiency*, In: Earth, Planets and Space, Vol. 52, No.5, p. 807-812, 1999

KIM, D., LANGLEY, R.B.: *GPS Ambiguity Resolution and Validation: Methodologies, Trends and Issues*, Proc. of the 7th GNSS Workshop – Intern. Symposium on GPS/GNSS, Seoul, Korea, 2000

LACHAPELLE, G., CANNON, M.E., ERICKSON, C., FALKENBERG, W.: *High Precision C/A Code Technology for Rapid Static DGPS Surveys*, Proc. of the 6th Intern. Geodetic Symposium on Satellite Positioning, Columbus, Ohio, USA, p 165-173, 1992

MADER, G.L.: *Ambiguity function techniques for GPS phase initialization and kinematic solutions*, Proc. of 2nd Intern. Symposium on precise positioning with the global positioning system, Ottawa, Canada, p. 1233-1247, 1990

MARTIN-NEIRA, M., TOLEDO, M., PELAEZ, Z.: The *null space method for GPS integer ambiguity resolution*, Proc. of DSNS'95, Bergen, Norway, 1995

REMONDI, B.W.: *Using the Global Positioning System (GPS) phase observable for relative geodesy: modelling, processing and results*, PhD thesis, Center for Space Research, University of Texas, Austin, 1984

SERRANO, L., KIM, D., LANGLEY, R.B., ITANI, K., UENO, M.,: *A GPS Velocity Sensor: How Accurate Can It Be? – A First Look*, Proc. of ION NTM 2004, San Diego, p. 875-885, 2004

TEUNISSEN, P.J.G.: *Least-squares estimation of the Integer GPS ambiguities*, Invited lecture, Section IV Theory and Methodology, IAG General meeting, Beijing, China, 1993

TEUNISSEN, P.J.G., DE JONG, P.J., TIBERIUS, C.C.J.M.: *The LAMBDA-method for fast GPS surveying*, Proc. of Intern. Symposium 'GPS technology applications', Bucharest, p. 203-210, 1995

Swept Path Determination by Means of PDGNSS

Jessica Glabsch[1], Otto Heunecke[1], Stefan Schuhbäck[1], Wolfgang Wirth[2]
[1] Institute of Geodesy, [2] Institute of Transport and Spatial Planning
University of the Federal Armed Forces Munich

Abstract
Swept paths are movements and tracks of different parts of a vehicle when the vehicle is undertaking turning manoeuvres. All road transport infrastructure is planned with respect to this circumstance but especially large trucks with trailers regularly cause damaged curbs and shoulders in small roundabouts and tight curves. Additionally damage at the rear wheels of trucks are the result and other road users are endangered if the rear part of such vehicles impede the opposing traffic, pedestrians and cyclists. Calculation of swept paths is not trivial and a complex function of winding geometry, roadway width, vehicle geometry and vehicle characteristics like maximum steering angle, wheel bases and slip. Finally driver skills play an important role. Road design software often shows some weakness regarding swept path considerations especially for larger vehicle combinations, however. Beside theoretical modelling practical swept path determination is requested to prove a roundabout or tight curve with respect to clearance, traffic safety etc. Inter alia, new kinds of trucks, e.g. the so-called Gigaliners, mandatory must give evidence that legal requirements are fulfilled in minimum turning curves and roundabouts. The behaviour of moving vehicles can be measured by augmented kinematic satellite-based positioning – generally termed kinematic PDGNSS[2] – whereby simultaneously antennas on the tractor and the trailer(s) of large vehicles have to record data. The geo-referenced tractrix of the vehicles then can be visualized and analyzed in detail. The paper reports on such experiments performed together with a German vehicle manufacturer.

Keywords
EuroCombi, geo-referenced tractrix, Gigaliner, kinematic GNSS positioning, swept path visualization

1 INTRODUCTION

In the next years there will be a significant strengthening of already existing transport infrastructure in Europe (e.g. BAST, 2008). There are prognoses that road freight will increase more than a third in Germany. It is undisputed that during a short time horizon of some years through expansion neither of the railway network or inland waterways nor by intensified development of the road network the traffic problem can be solved satisfactorily – beside other relevant questions in densely populated countries like additional land consumption, noise pollution etc. From the viewpoint of energy and emission balances trucks with up to 60 tons total weight – which are commonly known as EuroCombi or Gigaliner – are advocated as a possible solution, especially for European long haulage transport. Other often used terms for such kind of road trains are Long Combination Vehicles (LCV) or European Modular System (EMS). Such kind of large trucks – mentioned as Gigaliners in the following – usually are seven- or eight-axle vehicle combinations with 25.25 meters in total length. Vehicle width is about 2.55 meters. So far, in accordance with the EU Directive 96/53/EC only trucks with a total length of 18.75 meters and a total weight of 40 tons are permitted in Europe. The pros and cons and the consequences of accreditation of these new large trucks are discussed controversial and emotional between the different parties involved and – via the media – within the entire public. For some more information refer to e.g. ALBERS, 2006, BAST, 2006, 2008, and STEIGER, 2007.

Recently the German government decided to have a far created field test for five years beginning with January 1 2012 to verify whether the use of large trucks (with reduced max. weight of 44 tons) worth. It is no subject in the following taking sides in the discussion of all the pros and cons of Gigaliners. In

[2] PDGNSS: Precise Differential Global Navigation Satellite Systems.

the following evidence is given on their manoeuvring behaviour with respect to roundabouts and turning circles by some test runs at one of Europe's leading manufacturers of trucks.

2 MEANING OF SWEPT PATHS

When vehicles, e.g. trucks, cars, airplanes at airports or even motorcycles, are undertaking turning manoeuvres wheels at the rear axles turn at a smaller curve than the front wheels. The path taken by each wheel including the space in plane projection needed by the entire vehicle body during the turn describes a swept path, tractrix or driving cycle. Thus, a swept path in transportation planning is the plane envelope of a vehicle for a particular sequence of turning movements. See **Figure 21** top right for a graphical representation of the relevant terms. Swept paths are indispensable tools of road design with respect to traffic ability of roads especially for large vehicles. Damaged curbs and shoulders in small roundabouts and tight curves, tire tracks on verges, buckled posts and damaged kerbs are often outcomes of wrong operation by larger vehicles and inexperienced drivers due to the fact that the real covered area usually is about 30% higher than the computed area by road design software (KLEIN & SIMMEN, 2009).

Although mathematically uniquely parameterized as a line ("radiodrome"), the determination of swept paths as envelopes for vehicles is not trivial. It should be mentioned that the real geometrical relationships are more complex than depicted with **Figure 21** top right. Shape and size of swept paths here are subject to numerous factors such as curve progression, lane width, vehicle geometry (length, width, overhangs) and vehicle characteristics (wheel base, steering, and slip). Experience, proficiency and character of the drivers (e.g. aggressive versus calmly driving style) also affect the respective driven real path. Some more complicated constellations of joint vehicle combinations with steerable axles are not yet solved exactly analytically, however.

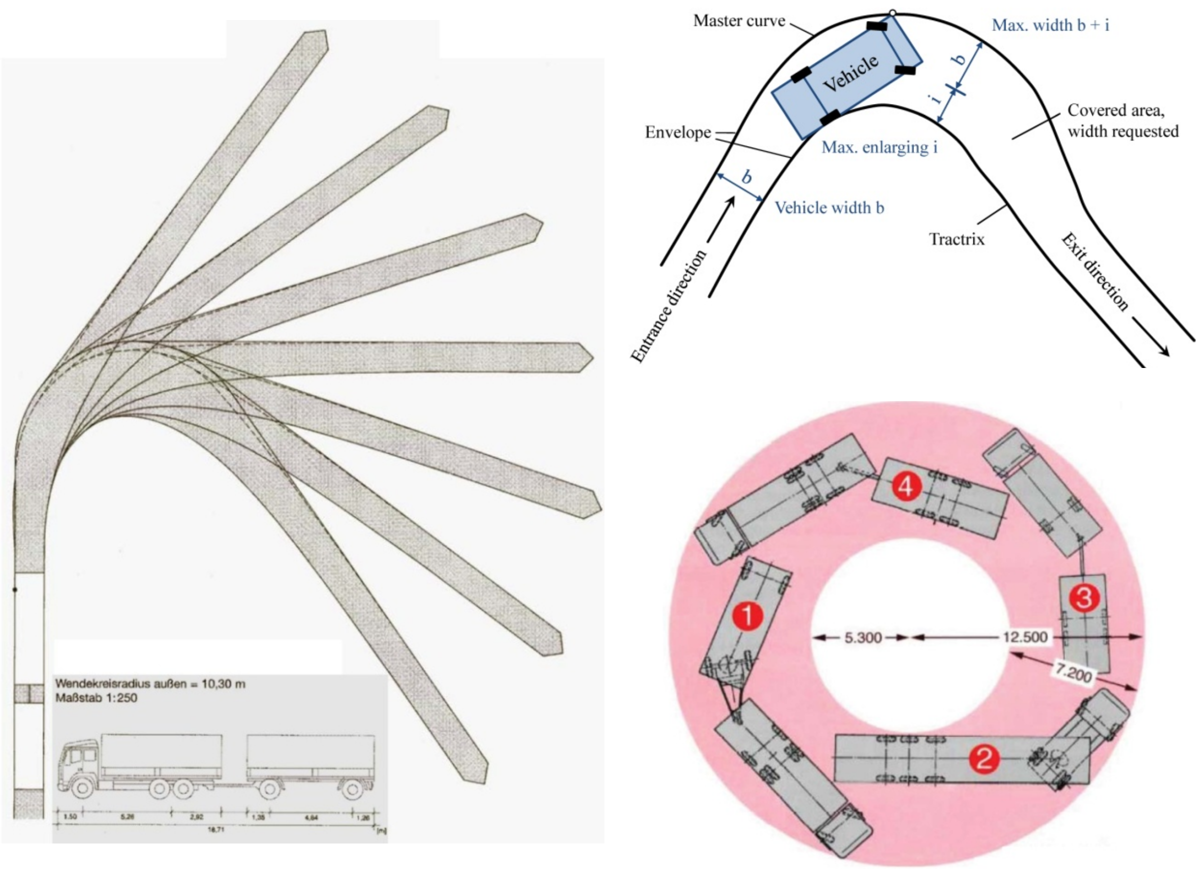


Figure 21: left side: Example of a swept path template for a standard trailer (taken from FGSV, 2001); top right: Master curve, envelope, tractrix and other terms of a driving curve (see GRÄFE et al., 2001); bottom right: Turning circle according to German road traffic regulations (taken from ALBERS, 2006);

There are three common methods of determining swept paths (GRÄFE et al., 2001, MESCHIK, 1992):

- Theoretical approaches, to distinguish between analytical and graphical principles;
- Model tests with vehicles on a smaller scale;
- Practical driving test on scale 1:1.

An analytical calculation of a swept path is only possible if the so-called "master curve" is pre-defined by the traffic planner along a certain point of the vehicle is moving (e.g. the left front edge of the vehicle). Driving skills of the vehicle drivers, vehicle speed, road conditions, traffic utilization etc. usually are not taken into account during planning. Thus, theoretical approaches are always mathematical approximations of the real situation and often show some weakness, especially if the bending is changing. This occurs for instance while entering and while leaving a roundabout.
For graphical construction in the past usually templates are used like to be seen in **Figure 21**, left side. Developed templates for Gigaliners – stage of development 2006, which is partly outdated due to recent automotive technical innovations – can be found in BAST (2006). Available software for swept path calculation and their suitability is investigated e.g. by SOBOTTA (2007) and KLEIN & SIMMEN (2009).

According to German road traffic regulations, in particular StVZO §32d operating characteristics in curves, conditions of approval for use on the road in Germany is (since 1986) that a normal licensed vehicle must be able to fulfil a 360° turning circle within the area of an outer radius of 12.50 meters and a width smaller than 7.20 meters. This is known as the so-called "BO-Kraftkreis", see **Figure 21** right bottom. This requirement is valid for all kind of trucks like standard trailer (no. 1), large semi trailer (no. 2), tandem trailer (no. 3) and axle trailer (no. 4).

Only practical driving tests have the advantage that all influences are implicitly covered in the turning manoeuvres to be tested and can give evidence that a certain vehicle really fulfils the legal requirements. But according to BAST (2008, page 43) it yields: "Since the number of variants of heavy truck combinations is tremendously large, each truck combination is unique. So the measured result is valid only for the tested vehicle or combination and the transition of the results to obviously similar vehicle combinations is, especially for heavy trucks, not possible. Therefore, it is not possible to use these test methods and the test results for regulation purposes."

Practical driving tests in the past (see e.g. BAST, 2008) often were solved rather archaic for instance by leaking paint or water of devices attached to the investigated vehicle. It should be noted that a fitting of markers in the vehicle corners leads only to correct results for the outer boundary of the envelope. The inner boundary of the envelope during turning manoeuvres results of the truck's side line and not the rear corner or rear axle. For uniform circular passages with a constant steering angle the tangent point that defines the envelope at the inner curve can be calculated in order to install the devices containing the marker like water or colour. Once the steering angle e.g. varies at a roundabout passage this point moves on the sideline of the vehicle. Only if the complete vehicle body is considered at the driving tests a complete picture of the real situation evolves. The use of GPS as a suitable method for swept path determination is already introduced by WIRTH (2001), GRÄFE et al. (2001) and NUSSRAINER (2002)

3 SWEPT PATH DETERMINATION TEST SETUPS

3.1 Test vehicle and GNSS instrumentation

The investigated vehicle is a truck with tractor (Krone Cool Liner) and a trailer (Krone Profi Liner) with a mechanically steered dual-axis dolly, see **Figure 22**. Such a LCV is respectively will be found most frequently in Europe.

For the recording of the manoeuvring behaviour the test vehicle was equipped with four GNSS antennas mounted at the corners of the tractor and the trailer, see **Figure 23**. Special adapters ensure a safe attachment of the antennas at the vehicle in which the receivers and batteries were placed nearby using backpacks. Additionally, manual distance measurements of the antennas to the corners of the vehicle are taken in order to calculate the outer contour lines of the vehicle bodies as rigid rectangles (**Figure 24**, right side). If necessary the outer contour lines also can be described by the exterior mirrors of the vehicle or any other specification. Leica System 1200 receivers (GPS & GLONASS) in combination with GSM modems for SAPOS® correction data receiving are selected to describe the tracks of the contour lines of both vehicle bodies in plane, geo-referenced coordinates during the different manoeuvres. Data recording frequency was set to 10 Hz.

Figure 22: Test vehicle with dimensions

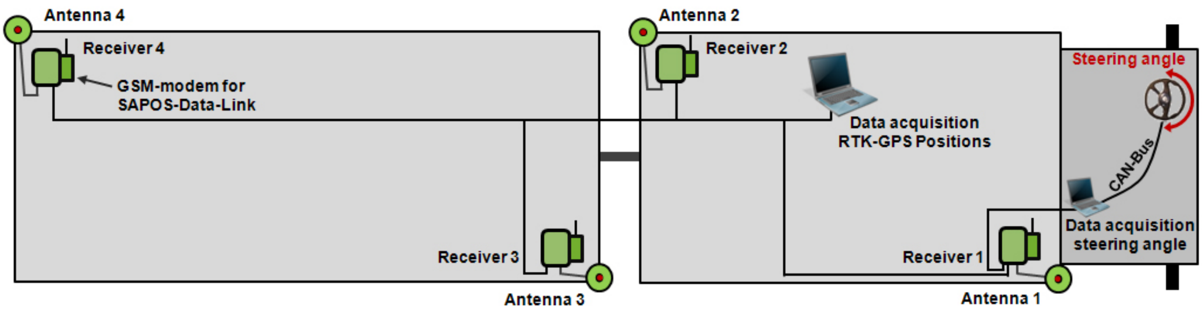


Figure 23: GNSS instrumentation of the test vehicle

Figure 24: left side and middle: Mounting of the GNSS antennas on the vehicle body (suction cup mount, clamp); right side: determination of distance to the vehicle edges

SAPOS® HEPS[3] is used instead of a local reference station for an easy way of obtaining geo-referenced data. The high precision real time positioning service HEPS is specified with an accuracy of 1-2 centimetres for plane coordinates. Furthermore directly Gauß-Krüger or UTM coordinates are processed. Taking the distances to the vehicle edges into account, the tracks of the vehicle bodies then can be computed and visualized. The envelope according to **Figure 21** top right is given by plotting the vehicle bodies at a sequence of measurements (every 0.1 second) during the runs. Thus, it is easily possible to visualize the swept paths in orthophotos or other kinds of digital maps if requested. However, effects of inclination of the vehicle (with the antennas on the roof) during the manoeuvres are neglected. Of course these effects can be some centimetres, especially when the roadway is inclined.

Via serial port and a serial device server the NMEA strings were converted to Ethernet and all four receivers were structured to a Local Area Network (LAN). Wireless Bluetooth communication was avoided due to some bad experience in former test runs. However, the cable across the shaft had to be kept flexible according to the expected vehicle movements. Using the LAN all data was sent to an onboard computer. Beside the tracks the steering angle of the truck can be recorded simultaneously for some detailed investigations. Therefore a second onboard computer was used to read-out the steering angle via the vehicle's CAN bus. Synchronisation is achieved with the NMEA message of receiver no. 1 and the second serial port by forwarding the time signal to the second computer (see **Figure 23**). Total setup time for the runs including all technical tests on site was about two hours. At the test site there were only few obstacles and practically no other disturbing factors for GNSS positioning.

The recording of the positions of all four receivers and the visualisation of the swept paths is obtained with self-developed software based on Labview®, National Instruments. Supervised by the operator all position data, track status (name of run, day & time; start/stop), data quality etc. can be checked during the runs. Thus, it is possible to react to potential problems during the track recording immediately. A simple check is given e.g. by the computed distances between the antennas on the tractor and the trailers which should be a constant value (see **Figure 27**). All data is recorded for a post processing option or an export of the tracks to road design software.

3.2 Task description

During the driving tests several different traffic situations were simulated with the test vehicle:

- Passage of a "big" roundabout with turning rates of 180° and 270°;
- Passage of a "small" roundabout with turning rates of 180° and 270°;
- Passage of the so-called BO-Kraftkreis, see **Figure 21** bottom right, in right and left turn.

The corresponding roadways including road splits at the entrances and exits were marked by some wooden blocks and guard signals at the test site of MAN Munich, see **Figure 25**. Shape and direction of the entrances and exits have a meaningful influence on the swept path in a roundabout because of the changing curvature. The laid-out marks were sufficient for the driver to orient like in a real traffic environment, e.g. a roundabout like to be seen in **Figure 26**. All mentioned trials were performed several times.

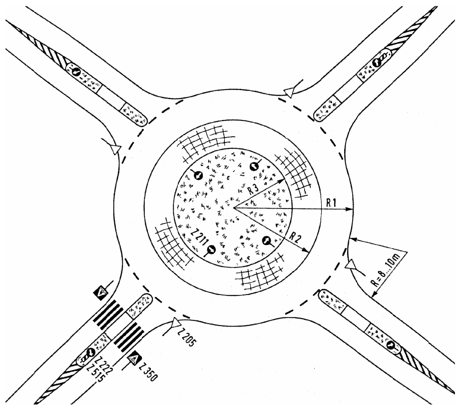
Figure 25: Marking of a roundabout at the MAN test site

[3] SAPOS HEPS: SAtelliten POSitionierungsdienst der deutschen Landesvermessung, Hochpräziser Echtzeit Positionierung Service, more details see www.sapos.de.

The geometrical parameters of the above listed traffic situations are shown in **Table 5**. These parameters are the basis of all background visualization of the test runs shown in the figures below.

Table 5: Geometrical parameters of simulated traffic situations

simulated traffic situation	inner radius [m]	outer radius [m]	width of entrance [m]	width of exit [m]
“big” roundabout	13.50	20.00	3.50	3.75
“small” roundabout	4.00	13.00	3.50	3.75
BO-Kraftkreis	5.30	12.50	-	-



According to FGSV (2006) the investigated roundabouts specified in **Table 5** fall in the category of so-called small roundabouts in Germany once with minimal external radius (termed as “small” throughout this paper) and once with the maximum external radius (termed here as “big”). The width of the circular ring is 9 meters for the “small” roundabout respectively 6.5 meters for the “big” roundabout, see **Table 5**.

Figure 26: Schematic representation of a roundabout with external radius R1 and inner radius R2 (from SOBOTTA, 2007)

3.3 Results of test runs

The analysis and visualization of the measurements is performed with a Matlab® program originally developed for the Institute of Transport and Spatial Planning by the company 3D Mapping Solutions GmbH. To determine the desired coordinates of the edges of each vehicle body during the runs two input files must be provided in addition to the measurement data. One file contains the definition of the vehicle’s coordinate system as well as information of the body length and width. In the second file the antenna offsets to the edges determined after the installation (see **Figure 24** on the right side) are considered. Using this information, all evaluation can be performed automatically. If the traffic situation (roundabouts with inner and outer radii, entrances and exits) should be shown in the plot and there is no digital map available, a corresponding survey of the marked points for simulation has to be provided. This additionally can be loaded by the developed program and the road boundaries are directly integrated in the Matlab® program.

Before starting with the presentation of some results the quality check by computing the spacing of the two antennas on tractor and on the trailer is discussed. This spacing should be a constant value, here 8.16 meters for the tractor and 13.80 meters for the trailer. The result is shown exemplarily for one of the test runs in **Figure 27** (on left side the tractor, on the right hand side the trailer). Depicted are the measurements recorded with 10 Hz. Only a few outliers exceeding the limits of 1 - 1.5 centimetres show the reliability of the results. Obviously proper measurements are obtained at the lower centimetre-level for the test runs.

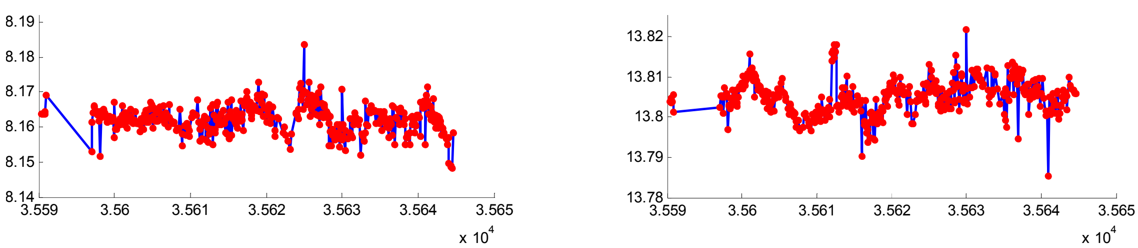


Figure 27: Verification of the antenna distances on the vehicle bodies; on left side antenna distances of the tractor, on right side antenna distances of the trailer; depicted for an exemplarily test run

As the detailed analysis of the runs of all roundabouts shows, a satisfactorily go through of the simulated traffic situations without a direct crossing or over coating of road periphery is possible from the technical and driving dynamic point of view for the tested vehicle. It comes up, however, that a collision or non-collision with road periphery strongly depends on the situation assessment by the driver.

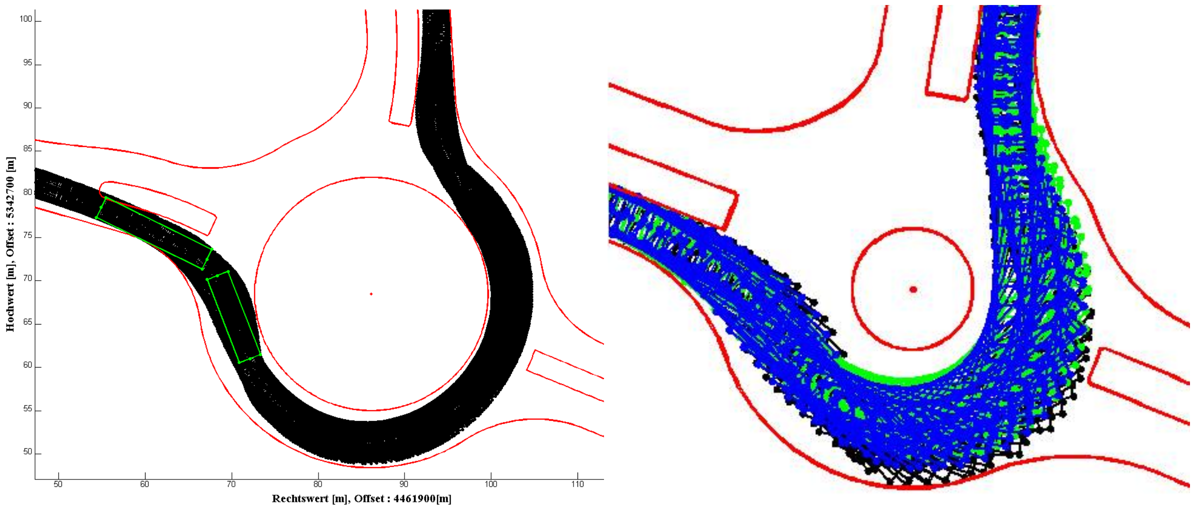


Figure 28: Visualization of the envelope of the test vehicle; left side: "big" roundabout with a turning rate of 270°; right side: all 3 test runs of the "small" roundabout with a turning rate of 270°

Figure 28 left hand side shows the effects of a non-optimal succeeded entrance at the "big" roundabout. The divisional island at the entrance was touched by the rear of the trailer (depicted by the green contour lines of the vehicle bodies at this moment). The exit of the roundabout is not optimally managed too. Here again is a light touch of the divisional island. By **Figure 28** on the right side in blue a successful drive through of the "small" roundabout with a turning rate of 270° is depicted. In black and green there are two further runs for some comparisons.

As it can be seen it depends very much on how the driver decided to manage the passage right from the beginning. The outer circle is slightly touched at the bottom in the first run (in black). Then – by situation awareness – the test driver acted much better in the second (green) and third (blue) attempt. Also the investigations by BAST (2008) confirmed that passages of such a roundabout with new vehicle combinations is possible but utilization of safety zones and side rooms normally is not to avoid.

Passing the BO-Kraftkreis is successfully fulfilled as it is shown in **Figure 29** on bottom left (depicted is a four-time drive through). Essentially here is a steerable (mechanical or electrical) dolly. Otherwise the trailer would pass over the inner circle and the Gigaliner would not meet the stringent requirements for trucks in Germany. An exemplarily enlarging i according to **Figure 21** in this simulated situation has a size of approximately 4.0 meters (highlighted in red at **Figure 29** bottom left, in green the vehicle width *b*). At what different path each part of the vehicle (tractor and trailer) runs along when driving through a circle is shown with **Figure 29** on the bottom right. Pointed in green is the tractor, in blue the trailer.

The trailer passes the roundabout in a much tighter inner radius than the tractor. In addition, it can be seen the already mentioned fact that the inner boundary of the envelope does not result of the vehicle's edges (rear inner corner), compare to the dashed black circle in **Figure 29** bottom right. The error in the determination of the envelope with reference to the corners of the vehicle would be about 2 meters for the BO-Kraftkreis in this case (highlighted in red, **Figure 29** bottom right).

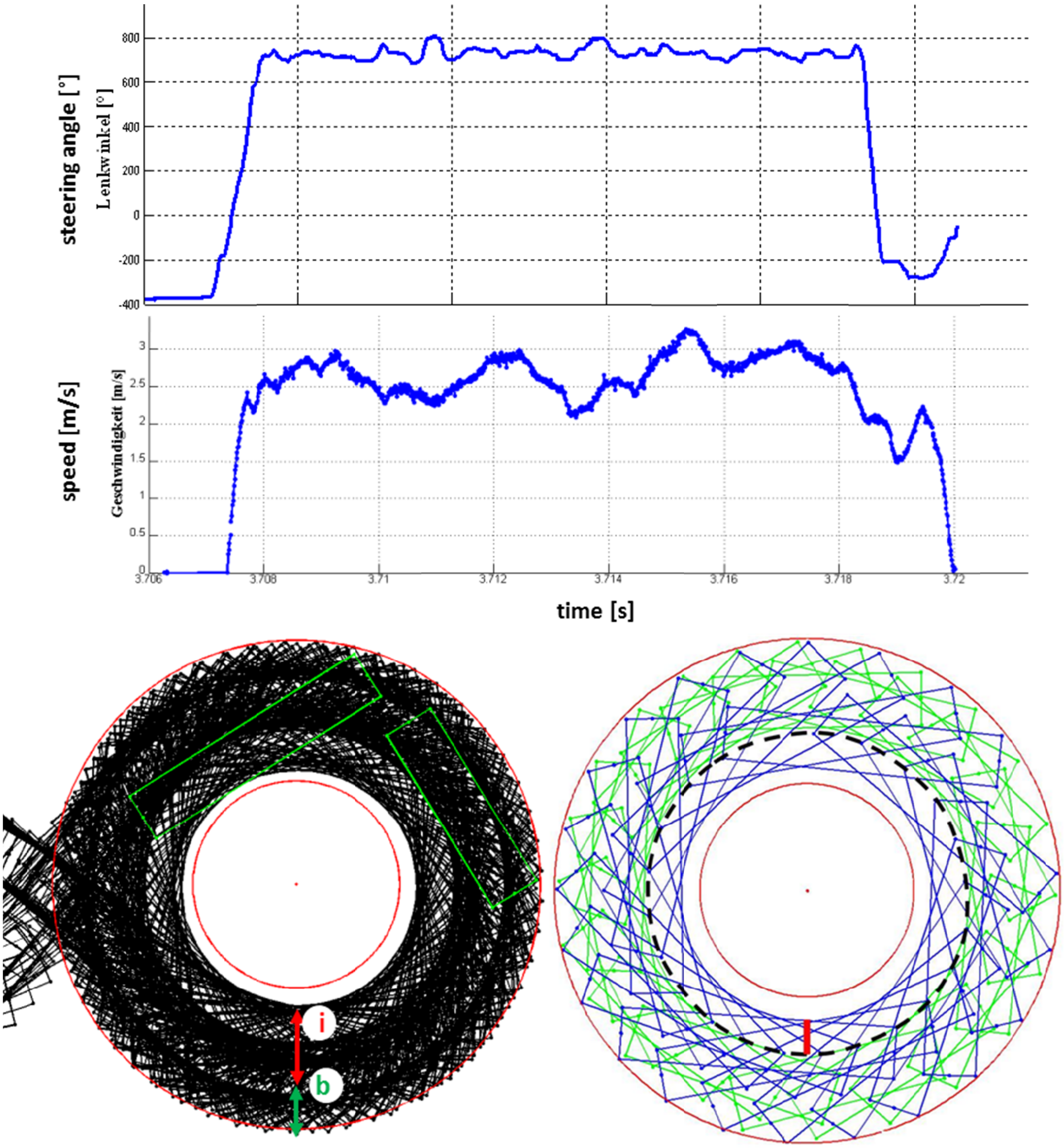


Figure 29: top and middle: Steering angle and speed of the vehicle while passing the BO-Kraftkreis; bottom: Passage of the BO-Kraftkreis (left) and separate visualization of a sequence of vehicle bodies with tractor in green and trailer in blue for one run only (right).

Beside the positioning data the steering angle and the speed of the vehicle is also recorded and can optionally be plotted. This is essential since other studies have shown that the velocity the test vehicle goes through the BO-Kraftkreis affects significantly the covered area (BAST, 2008). A higher speed leads to a stronger centrifugal force so that the trailer is pulled more strongly towards the outside of the curve. During the tests the BO-Kraftkreis was passed with an average speed of about 2.75 meters per second (see **Figure 29** in the middle). Thus, with a recording frequency of 10 Hz every 0.27 meters there is a measured point of the track. The covered area identified in this trail with less than 7.2 meters corresponds to the requirements and is comparable to the results of the BAST-study in 2008. As **Figure 29** shows on the bottom left, the vehicle was slightly over coating the predetermined external radius of 12.5 meters in some areas.

4 CONCLUSIONS

Manoeuvrability of the test vehicle could be proven by several test runs in simulated environment. Due to some recently developed high sophisticated steering techniques trucks with a total length of 25.25 meters today nearly can turn at circles like the common 18.75 meters trucks. According to BAST (2006, page 62) this was not possible a few years ago. Of course these results are only valid for the specific investigated vehicle and should be performed for every long vehicle combination individually to prove the respective manoeuvring behaviour, however.

The whole developed measuring system for swept path determination can be extended to 6 six or more receivers if there is for instance a second trailer, adapted to nearly all kind of vehicles easily (e.g. heavy load vehicles with excess width, airplanes). Furthermore the system inter alia has the great advantage that no complex preparations for the test sites are necessary and tests on public roads are possible without blocking traffic junctions etc. Many manual activities such as the determination of the points of the resulting envelope during swept path determination with archaic methods (leaking paint or water of devices attached to the vehicle) – which also have a high vulnerability to weather conditions like wind – are superfluous. It could be shown that with the above-described method the determination of envelopes at the inside of a curve, which is not covered by the vehicle corners but by the line connecting the vertices, is possible easily and reliable. Due to the fact that all data is geo-referenced a depiction of digital maps in the background is a possible option.

For more detailed investigations the corresponding steering angles can be recorded synchronously to the positions of the vehicle as well as the velocity of the vehicle is present at every sequence during the manoeuvres. Results are available immediately after the driving tests. Finally it can be stated, that the effort for driving tests with high quality results at the centimetre-level is reduced significantly using kinematic PDGNSS and offers many more opportunities for detailed analysis. This should be exploited for the benefit of road planning and – at the end – safety of public traffic.

ACKNOWLEDGEMENTS

All testing took place on the premises of MAN Munich in September 2011. The tested vehicle combination was provided by the companies MAN and Krone.

REFERENCES

ALBERS, F.: *Europäisches Modulares Nutzfahrzeugsystem – Technische Aspekte und Testfahrt mit dem Giga Liner durch Deutschland*. Presentation May 31 2006, www.euro-combi.de/dwl/EuroCargo-GigaLiner.pdf .

BAST: *Auswirkungen von neuen Fahrzeugkonzepten, Schlussbericht.* Bundesanstalt für Straßenwesen, Bergisch Gladbach 2008, see publications at http://www.bast.de.

BAST: *Auswirkungen von neuen Fahrzeugkonzepten auf die Infrastruktur des Bundesfernstraßen-netzes, Schlussbericht.* Bundesanstalt für Straßenwesen, Bergisch Gladbach 2006, see publications at http://www.bast.de.

FGSV: *Bemessungsfahrzeuge und Schleppkurven zur Überprüfung der Befahrbarkeit von Verkehrs-flächen*. Forschungsgesellschaft für Straßen- und Verkehrswesen, Arbeitsgruppe „Straßenentwurf“. FGSV publish house, Cologne 2001.

FGSV: *Merkblatt für die Anlage von Kreisverkehren.* Forschungsgesellschaft für Straßen- und Verkehrswesen, Arbeitsgruppe „Straßenentwurf“. FGSV publish house, Cologne 2006.

GRÄFE, G., HEISTER, H., MESCHIK, M., NUSSRAINER, C., SOBOTTA, R., WIRTH, W.: *Schleppkurven-Symposium 2001.* Schriftenreihe des Instituts für Verkehrswesen und Raumplanung, Universität der Bundeswehr München, Issue 45. ISSN 1436-9982, 2001.

KLEIN, S., SIMMEN, S.: *Überprüfung und Befahrbarkeit bei Straßenentwurfsplanungen*. Straße und Autobahn, Vol. 9, p. 589-592, 2009.

MESCHIK, M.: *Simulation von Schleppkurven verschiedener Fahrzeuge*. PhD, Institut für Verkehrswesen der Universität für Bodenkultur, Issue 22, Wien 1992.

NUSSRAINER, C. (2002): *Learning curve*. Traffic Technology International. Issue 10, 2002.

SOBOTTA, R.: *Überprüfung von Entwurfsparametern für Kreisverkehre mit empirisch ermittelten Schleppkurven*. Schriftenreihe des Instituts für Verkehrswesen und Raumplanung, Universität der Bundeswehr München. Issue 49, 2007

STEIGER, D.: *Pro und Contra Gigaliner. Projekt Vorstudie EU-Policy. Projektarbeit 2007,* www.mobifair.eu/Projekte/07_03_20_Gigaliner.pdf*, 2007*

WIRTH, W. (2001): *Fahrkurven ja, aber die richtigen – Ein Beitrag zur Entmythlogisierung der Schleppkurventhematik*. Internationales Verkehrswesen (53), Issue 6, p. 288-291, 2001

Links:
http://www.euro-combi.de, last accessed on November 21, 2011

http://www.sapos.de, last accessed on November 23, 2011

http://de.wikipedia.org/wiki/EuroCombi, last accessed on November 25, 2011

Uncertainty of Motion Detection from Coordinate Differences

Andreas Wieser
Institute of Geodesy and Geophysics, Vienna University of Technology, Austria

Abstract
The distance computed from measured coordinate differences is a biased estimate of the true distance. The bias exceeds the standard deviation of the coordinate differences if the true distance is short. Therefore, it is necessary to take the uncertainty of the measurements properly into account when motion is to be distinguished from static periods using estimated coordinate differences.

In this contribution we focus on the stochastic properties of the coordinate differences and the vector norm. We give a suitable statistical test which allows correctly detecting motion with a chosen probability if the amount of motion exceeds a chosen minimum. We also quantify the bias of the computed distance and present the Normal distribution as a suitable approximation if the true distance exceeds the level of the standard deviations by a factor of 10 or more. This approximation may be useful in practical applications to express the amount of actual motion using confidence intervals.

Keywords
Motion detection, distribution of vector norm, distance bias.

1 INTRODUCTION

Certain applications require reliable distinction between motion of a vehicle and static periods. Accomplishing this may be a non-trivial task because of environmental conditions, ambient noise and the uncertainty of the measurements on which the distinction must be based. It is particularly difficult if motion has to be detected while the distance travelled is still small compared to the uncertainty of the measurements. Herein, we shall focus on the impact of the statistical uncertainty of the measurements and on using estimated coordinate differences $\Delta\hat{\mathbf{x}}$ or the corresponding distance (Euclidean norm)

$$\hat{d} := \|\Delta\hat{\mathbf{x}}\| \tag{1.1}$$

for motion detection. The coordinate differences may be obtained using Global Navigation Satellite System (GNSS) position estimates referring to different epochs. They could also be obtained using other positioning sensors (e.g., total stations) or velocity sensors (GNSS, optical sensors).

The task can be solved by statistically testing the null-hypothesis

$$\mathrm{H}_0: \quad \mu = 0 \tag{1.2}$$

against the alternative

$$\mathrm{H}_\mathrm{a}: \quad \mu > 0\,, \tag{1.3}$$

with

$$\mu := \|\boldsymbol{\mu}\|\,,\ \boldsymbol{\mu} := E\{\Delta\hat{\mathbf{x}}\}\,. \tag{1.4}$$

Motion is detected if H_0 has to be rejected and the statistical test thus indicates that the true length μ of the vector of coordinate differences is very likely greater than zero.

Some complications arise because the Euclidean norm

$$\|\Delta\hat{\mathbf{x}}\| := \sqrt{x_1^2 + x_2^2 + \ldots x_n^2} \tag{1.5}$$

of a random vector $\Delta\hat{\mathbf{x}}$ with dimension

$$\dim \Delta\hat{\mathbf{x}} = n \times 1 \tag{1. 6}$$

expectation as of eq. (1.4), and dispersion

$$D\{\Delta\hat{\mathbf{x}}\} := \boldsymbol{\Sigma} . \tag{1. 7}$$

is a biased estimate of the norm of the vector, i.e.,

$$E\{\|\Delta\hat{\mathbf{x}}\|\} \neq \mu . \tag{1. 8}$$

This is visualized in **Figure 30** for the special case of n=1, and a standard normal variable $\Delta x \sim N(0,1)$: while the true length of the vector is 0, the expected value of the magnitude is 0.8. It is necessary to take this bias into account when using the magnitude of coordinate differences for motion detection.

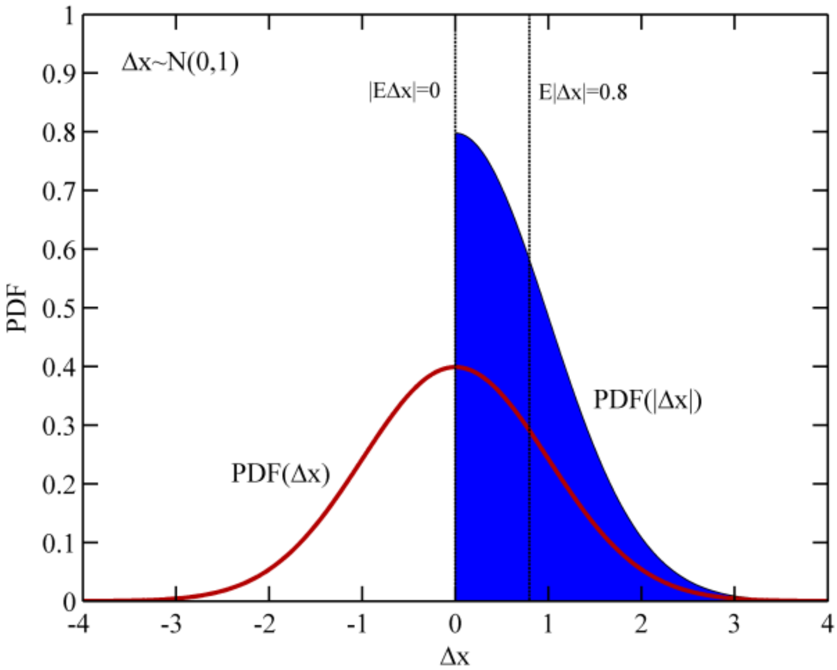


Figure 30: Probability density function for a standard Normal variable Δx and its magnitude.

We will subsequently review a standard test for rejecting the null hypothesis (1.2), outline the shortcomings of such an approach when applied to motion detection and offer a solution which is based on minimum motion detection requirements. In section 3, we briefly discuss the distribution of the vector norm and its bias.

2 TESTING THE RANDOM VECTOR

2.1 Standard test

Assuming that the components of the random vector are normally distributed with known and regular variance-covariance matrix i.e.,

$$\Delta\hat{\mathbf{x}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}),\ \mathrm{rk}\,\boldsymbol{\Sigma} = n , \tag{2. 1}$$

we know that the quadratic form

$$q := \left(\Delta\hat{\mathbf{x}} - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1} \left(\Delta\hat{\mathbf{x}} - \boldsymbol{\mu}\right) \sim \chi_n^2 \tag{2. 2}$$

is χ^2-distributed with n degrees of freedom, see e.g., Sachs (2003), Koch (1988, pp. 144). A reasonable and well established test for deciding between the null-hypothesis

$$\mathrm{H}_0: \quad E\{\Delta\hat{\mathbf{x}}\} = \mathbf{0} \tag{2. 3}$$

and the alternative

$$\mathrm{H}_\mathrm{a}: \quad E\{\Delta\hat{\mathbf{x}}\} \neq \mathbf{0} \tag{2. 4}$$

is thus based on the test statistic

$$\hat{\chi}_n^2 := \Delta\hat{\mathbf{x}}^T \mathbf{\Sigma}^{-1} \Delta\hat{\mathbf{x}} \tag{2. 5}$$

Selecting an appropriate error probability α, e.g. 5%, and taking into account (2.2), (2.3) and (2.5) H_0 is rejected in favor of H_a if the test statistic exceeds the critical value ($1-\alpha$ percentile of the χ^2-distribution):

$$\hat{\chi}_n^2 \overset{\mathrm{H}_\mathrm{a}}{>} \chi_{n,1-\alpha}^2 . \tag{2. 6}$$

The hypotheses (2.3) and (2.4) are perfectly equivalent to (1.2) and (1.3) with (1.4). So, this standard test can be applied for motion detection. However, the problem is choosing an appropriate critical value or equivalently an appropriate numeric value of α. This is not straightforward because the goal is to correctly detect motion if it occurs, whereas the selection of α is related to the detection of stationarity if it occurs. The probability γ of correctly detecting motion, i.e., the power

$$P\{\hat{\chi}_n^2 > \chi_{n,1-\alpha}^2 \,|\, \mathrm{H}_\mathrm{a}\} := \gamma , \tag{2. 7}$$

is related to α only through the critical value and the alternative hypothesis. So, unless the alternative is specified more strictly than in eq. (2.4) and a desired minimum probability of correct detection is specified, the motion detection problem cannot be solved.

2.2 Motion detection

The problem is now stated more precisely as testing H_0 against H_a with

$$\begin{cases} \mathrm{H}_0: & \mu = 0 \\ \mathrm{H}_\mathrm{a}: & \mu > d_{\min} \end{cases} \tag{2. 8}$$

and with a minimum probability $\gamma_{\min}$ of rejecting H_0 in case H_a is true. This means that motion by a distance of at least $d_{\min}$ is to be correctly detected with a probability of at least $\gamma_{\min}$.

Let

$$\Delta\hat{\mathbf{x}} := \delta\hat{\mathbf{x}} + \boldsymbol{\mu} \tag{2. 9}$$

with

$$\delta\hat{\mathbf{x}} \sim N(\mathbf{0}, \mathbf{\Sigma}) \tag{2.10}$$

Then we have

$$\hat{\chi}_n^2 = \Delta\hat{\mathbf{x}}^T \mathbf{\Sigma}^{-1} \Delta\hat{\mathbf{x}} = \delta\hat{\mathbf{x}}^T \mathbf{\Sigma}^{-1} \delta\hat{\mathbf{x}} \sim \chi_n^2 \,|\, \mathrm{H}_0 \tag{2.11}$$

and

$$\hat{\chi}_n^2 = \Delta\hat{\mathbf{x}}^T \mathbf{\Sigma}^{-1} \Delta\hat{\mathbf{x}} = \delta\hat{\mathbf{x}}^T \mathbf{\Sigma}^{-1} \delta\hat{\mathbf{x}} + 2\delta\hat{\mathbf{x}}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu} + \boldsymbol{\mu}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu} \sim \chi_{\lambda,n}^2 \,|\, \mathrm{H}_\mathrm{a} \tag{2.12}$$

with

$$\lambda = \boldsymbol{\mu}^T \mathbf{\Sigma}^{-1} \boldsymbol{\mu} \tag{2.13}$$

So, the test statistic of the standard test follows a *non-central* χ^2-distribution with non-centrality parameter λ if there is motion. Alas, the non-centrality depends on the actual length of the vector and on its direction. However, since the variance-covariance matrix is symmetric and of full rank, and the eigenvalues are the extreme values of quadratic forms of the type (2.13) see e.g., Koch (1988, pp. 52), we find the following limits for the non-centrality:

$$\frac{\mu^2}{\nu_{\max}} \le \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \le \frac{\mu^2}{\nu_{\min}}, \tag{2. 14}$$

where $\nu_{\min}$ and $\nu_{\max}$ are the minimum and maximum eigenvalues of $\boldsymbol{\Sigma}$, respectively. The above minimum distance $d_{\min}$ causes a non-centrality

$$\frac{d_{\min}^2}{\nu_{\max}} = \lambda_{\min} \le \lambda \le \lambda_{\max} = \frac{d_{\min}^2}{\nu_{\min}}. \tag{2. 15}$$

Consequently, the critical value for the above statistical test can be obtained from

$$\chi^2_{n,1-\alpha} = \chi^2_{\lambda_{\min},n,1-\gamma_{\min}}. \tag{2. 16}$$

If the actual motion is larger than $d_{\min}$ or if it occurs in a direction different from the one corresponding to the maximum eigenvalue of $\boldsymbol{\Sigma}$ then the probability of correct detection is greater than $\gamma_{\min}$. It is γ^* and can be computed from the critical value which is also

$$\chi^2_{n,1-\alpha} = \chi^2_{\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu},n,1-\gamma^*}. \tag{2. 17}$$

In real applications it may not always be possible to evaluate the above equations using the correct variance covariance matrix because this matrix may not be known for each vector of coordinate differences. In such cases it is advisable to roughly estimate the maximum eigenvalue from experience or from a representative test data set. This is possible because the maximum eigenvalue is the maximum uncertainty in any spatial direction expressed as the square of the corresponding standard deviation.

2.3 Numerical example

For a numeric example, let

$$d_{\min} = 0.2 \text{ m}$$
$$\gamma_{\min} = 80\%$$

So, motion of at least 0.2m has to be detected correctly with a probability of at least 80%. Furthermore, let the estimated coordinate differences and their variance covariance matrix be

$$\Delta\hat{\mathbf{x}} = \begin{bmatrix} 0.095 \\ -0.036 \end{bmatrix} \quad \text{m}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 0.0085 & 0.0030 \\ 0.0030 & 0.0040 \end{bmatrix} \quad \text{m}^2$$

The maximum eigenvalue of this variance covariance matrix is 0.0100 m^2, and the minimum non-centrality associated with $d_{\min}$ is thus $\lambda_{\min}$ = 0.04/0.01 = 4. So, the critical value to be used within the test is

$$\chi^2_{\lambda_{\min},n,1-\gamma_{\min}} = \chi^2_{4,2,0.2} = 2.164,$$

and the error probability α (false alarm rate) is 33.9% because

$$\chi^2_{n,1-\alpha} = \chi^2_{2,0.661} = 2.164\,.$$

The test statistic is

$$\hat{\chi}^2_n = \Delta\hat{\mathbf{x}}^T \mathbf{\Sigma}^{-1} \Delta\hat{\mathbf{x}} = 2.705 > 2.164\,,$$

and the null-hypothesis is rejected for this vector i.e., the test indicates motion.

This may seem astonishing since the length of the vector is only $\|\Delta\hat{\mathbf{x}}\| = 0.102\,\mathrm{m}$ and thus much shorter than the selected minimum distance. However, the direction of maximum uncertainty (eigenvector corresponding to the maximum eigenvalue) is about 30 gon, and the chosen minimum power of 80% refers to that direction. The test is much more sensitive in other directions, as is visualized using empirical results from a numerical simulation in **Figure 31**.

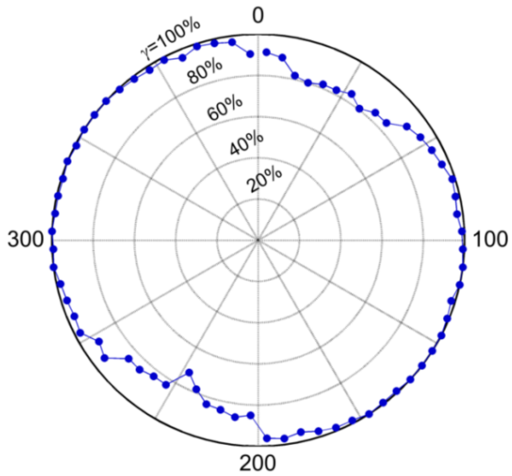


Figure 31: Relative frequency of correct detection of motion as a function of azimuth; results based on Monte-Carlo simulation with 10k samples of normally distributed vectors (true length $\mu = d_{\min}$, true azimuth equally distributed, coordinate errors according to Σ in above numeric example), evaluation using 5-deg azimuth bins.

3 THE DISTRIBUTION OF THE VECTOR NORM

In order to detect motion from the estimated distance as of eq. (1.1) directly, the distribution of the vector norm is required. The vector $\Delta\mathbf{x}$ of estimated or measured coordinate differences is a random vector. If its components are independent and identically normally distributed, i.e.,

$$\Delta\mathbf{x} \sim N_n(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n) \tag{3. 1}$$

with $\boldsymbol{\mu} \neq \mathbf{0}$, then the random variable

$$y := \frac{\|\Delta\mathbf{x}\|}{\sigma} = \sqrt{\frac{\Delta x_1^2}{\sigma^2} + \frac{\Delta x_2^2}{\sigma^2} + \ldots + \frac{\Delta x_n^2}{\sigma^2}} \tag{3. 2}$$

is non-centrally chi-distributed with n degrees of freedom and non-centrality

$$\lambda = \frac{\|\boldsymbol{\mu}\|}{\sigma}\,. \tag{3. 3}$$

The corresponding probability density function is given by

$$f_{\chi(n,\lambda)}(y) = \frac{y^{n/2}}{\lambda^{n/2-1}} \cdot e^{-\frac{1}{2}\left(y^2+\lambda^2\right)} \cdot I_{n/2-1}(\lambda y)\,, \tag{3. 4}$$

where $I_\nu(z)$ is a modified Bessel function of the first kind. As $\lambda \to 0$ this function approaches the density function of the chi-distribution with n degrees of freedom:

$$f_{\chi(n)}(y) = \frac{1}{2^{n/2-1}\Gamma(n/2)} \cdot y^{n-1} \cdot e^{-y^2/2}. \tag{3. 5}$$

More information on the chi-distributions can be found in Johnson et al. (1994, 1995), and Evans et al. (2000).

The error of the distance computed as the norm of a random vector of dimension n is then distributed as

$$\varepsilon_d := \|\Delta\mathbf{x}\| - \mu = d - \mu \sim f_{\chi(n,\lambda)}\left(\frac{\varepsilon_d + \mu}{\sigma}\right). \tag{3. 6}$$

This is visualized in **Figure 32** (left). The plot shows the probability density of the distance errors for three different true lengths of a 3-dimensional vector: 0, 1σ and 10σ along with the corresponding expected values of the errors, i.e., the respective distance bias. We see that the distance bias in this 3d case reaches 140% of the standard deviation if the true distance is 0, i.e., if no motion occurs. The bias decreases with increasing length. **Figure 32** (right) shows the bias as a function of true length for the 1d, 2d and 3d cases. Obviously, the bias may be negligible if the length of the vector exceeds about 2 to 3σ. If distances computed from coordinate differences are further processed, e.g. averaged or accumulated, the bias may still be significant even if the length of the individual vectors exceeds 10σ, see Wieser (2007) for more information.

In practical applications, the components of the random vector are often not independent and not identically distributed. The corresponding vector norm is then not chi- or non-centrally chi-distributed. Furthermore, the distribution function cannot be expressed using simple functions in this general case, see Wieser (2007, p.122). So, from a practical point of view, it is not advisable to test for motion using directly the distances estimated from coordinate differences, because the required percentiles of the related distribution functions are not readily available and their calculation is not straightforward.

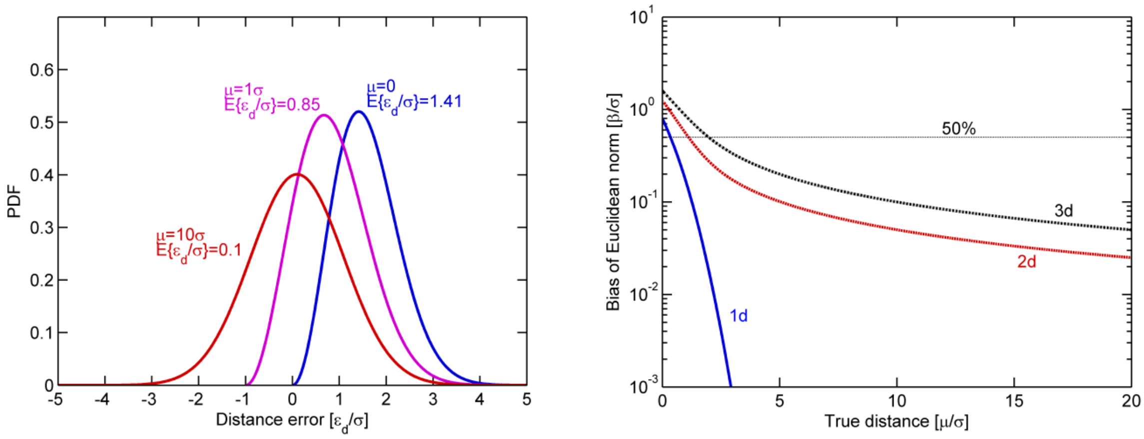


Figure 32: Probability density function of distance error for 3d-vectors distributed as of eq. (3.1) with three different true lengths; error and true lengths given in units of σ for convenience (left). Distance bias corresponding to vector norm of 1d, 2d, and 3d vectors with i.i.d. normally distributed components (right).

However, as shown in Wieser (ibid), the norm of a normally distributed vector is also approximately normally distributed as of

$$\|\Delta\mathbf{x}\| \underset{\text{approx.}}{\sim} N(\mu + \frac{\operatorname{tr}\mathbf{\Sigma} - \sigma_l}{2\mu}, \sigma_l^2) \tag{3. 7}$$

if the true length exceeds the standard deviations by a factor of at least 10, see also **Figure 32** (left). In case the standard deviations corresponding to different spatial directions differ greatly (by more than an order of magnitude), the standard deviations perpendicular to the vector are relevant for assessing whether or not the approximation is sufficiently accurate. In eq. (3.7), the standard deviation σ_l along the vector is needed, which can be obtained by rotating the vector into an arbitrary coordinate system with one of its axes parallel to the vector.

This approximate normal distribution can obviously not be used for motion detection because it only holds if the true vector is significantly longer than the standard deviation whereas the null-hypothesis for motion detection represents the case μ=0. However, eq. (3.7) can be used to calculate an approximate confidence interval for assessing the precision of the calculated distance in case motion has been detected.

4 CONCLUSION

We have presented a statistical test for motion detection using estimated coordinate differences. In section 2.2 we have proposed a criterion for selecting the critical value in order to detect motion with a predefined minimum probability in case it actually occurs and exceeds a given minimum distance. In section 3 we have given the probability density of the vector norm for the special case of i.i.d. normally distributed vector components and quantified the bias of the vector norm. This quantification may help to decide for practical applications how the distances computed from vectors of coordinate differences are further processed and interpreted. The norm of long vectors is approximately normally distributed, and the parameters given in section 3 can be used to quantify the uncertainty of the estimated distance once motion has been detected.

REFERENCES

EVANS, M., HASTINGS, N., PEACOCK, B.: *Statistical Distributions*. 3rd ed, John Wiley & Sons, New York, p.221, 2000.

JOHNSON, N.L., KOTZ, S., BALAKRISHNAN, N.: *Continuous Univariate Distributions, vol 1*. 2nd ed, John Wiley & Sons, New York, p.784, 1994.

JOHNSON, N.L., KOTZ, S., BALAKRISHNAN, N.: *Continuous Univariate Distributions, vol 2*. 2nd ed, John Wiley & Sons, New York, p.752, 1995.

KOCH, K.-R.: *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer Verlag, p.378, 1988.

SACHS, L.: *Applied Statistics. Application of Statistical Methods*. 11th ed, Springer Verlag, p.889, 2003 (in German).

WIESER, A.: *GPS based velocity estimation and its application to an odometer.* Engineering Geodesy - TU Graz, Shaker Verlag, Aachen, 2007.

This Paper has been accepted as a "peer reviewed paper" by the scientific committee

Retransmission of Differential Data Corrections Concept Applied on Machine Control

Andreea Florina Jocea[1], Aurelian Stelian Buda[2], Ştefan Suba[2], Sorin Nistor[2], Norbert Szabolcs Suba[2], Paul Daniel Dumitru[1]
[1]Faculty of Geodesy, Technical University of Civil Engineering of Bucharest, Romania
[2]Faculty of Architecture and Construction, University of Oradea, Romania

Abstract
This paper will present the importance of using the next generation components, which is supporting the access to differential data corrections in areas where this was impossible before. It is well known that differential data corrections can be transmitted from the casters to the users in two ways: radio transmission and mobile Internet transmission.

The most common method of transmitting the differential data corrections is by using the UHF band. This method has a disadvantage: the frequencies chosen depend on the restriction imposed by the governmental agencies. This restriction has a direct impact to the distances between the rover and transmitter. Another way of transmitting differential data is the Mobile Internet by using the NTRIP protocol which has the disadvantage that the mobile services providers do not have enough coverage (data transmission is not the main priority of mobile services providers).

The problems presented above may have been developed a prototype system in which the radio transmitters and mobile Internet technology are "combined" – named retransmission concept.

This concept was applied for the first time in Romania, to Machine Control systems equipped with 3D GNSS receivers which suppose the installation of a permanent station on the building site for the receiving of differential data.

The concept consists in replacing this station with a cheap and effective solution which allows the reception of differential data from the national permanent network stations in order to obtain similar results.

Keywords
Machine Control, RTK, ROMPOS

1 INTRODUCTION

During the last years, the scientific and technological development in terrestrial measurements field, accompanied by a cost reduction, has led, to a fast improvement of the precision requirements.
The use of different combinations of equipment and measurement techniques has generated a transformation of productive processes in various application sectors. An innovation, in the geodetic field, is represented by the integration of the navigation and positioning sensors with other measuring equipment, leading to very good and accurate results.

This technological development has also focused on construction activity, particularly on construction machinery sector activity, assisted by GNSS receivers integrated with sensors for slope determination.
In Romania, the use of machine control systems is still sparse. Considering the advantages, for instance, cost reduction and the quality deriving from such a technique, a significant increase of their use is expected in the near future.

In the dozer type machine, the system consists in a GNSS antenna which is mounted on the front blade. To compute the transversal inclination of the blade, a second GNSS antenna or slope sensors are installed.

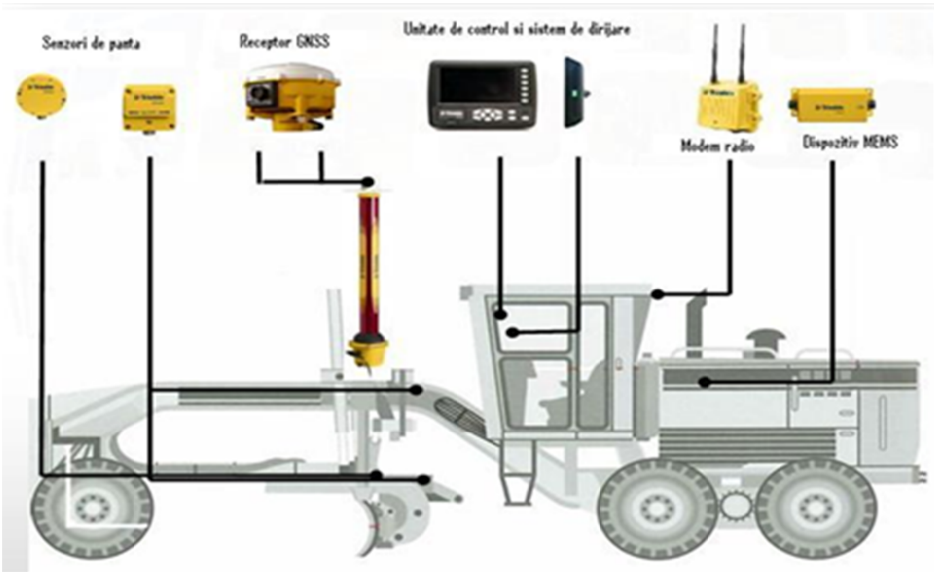


Figure 33：The main parts of a machine control system in case of dozer type machine (Trimble, 2009)

The current organization of GNSS machine control, even realizing a relevant innovation with respect to the traditional man-assisted way of operating in building sites, still presents some problems and limitations: costs for the GNSS base station; communication problems between base station and rover on the machine, usually based on modems, which in countries like Romania are subjected to power restrictions; affordability and integrity problems for the correction data coming from one station only [1].

In this paper, we will present the possibility of using differential data corrections in RTCM format provided by the Romanian Permanent Stations Network – ROMPOS on Machine Control System case. It is studied the possibility of differential data reception provided by ROMPOS using a dedicated application and a computer connected to the internet through a mobile modem and the instant retransmission of these to the machine using a radio modem.

2 BASIC CONCEPT

GNSS Machine Control is based on real time global positioning technologies (RTK), which allow the increasing of the productivity process in a short time and with low cost.

This method of measurement is used - RTK – all over the world because of the working speed and accuracy provided. The main advantages are the velocity of the position determination (in a few seconds) and the accuracy, which is of the order-centimeters and can even reach millimeter accuracy. This working method requires the use of dual frequency receivers which allow the use of differential corrections. The way of obtaining these differential data corrections depends on the working method: obtaining differential corrections provided by the permanent stations and VRS method.

The transfer of differential data corrections from the casters to users can be done in various ways. The most common are: transfer via radio transmitters, mobile communication systems such as: GSM/GPRS/SMS or mobile Internet transmission.

2.1 Transmission of differential data corrections

The next part will present two ways of how differential corrections can be transmitted.

2.1.1 Transmission of differential data corrections using radio waves

The most common method of transmitting the differential data corrections is by using the UHF radio band (and sometimes the VHF radio band), with the data flow rate of 9600 bps. The main advantage of this method is that these electromagnetic waves can travel a long distance, which allows the coverage of a wider area where the differential corrections can be received. The frequencies chosen are dependent on restrictions imposed by the governmental agencies. We would like to highlight the emitting power restriction, due to the fact that this has a direct impact on the possible distance between the rover and transmitter [2].

Radio based transmissions have other several significant disadvantages when used for positioning and navigation. First, radio has a very short data transmission range limited to line of sight (LOS) between the base and rover radios. Second, the radio waves in UHF commercial band are crowded and noisy: the band is designed for public and there is lack of an effective frequency resource sharing strategy like cellular technology to avoid frequency collision.

In urban area, where the radio wave environment is harsh, the radio effective ranges are usually limited to a few kilometers. The coverage is also reduced in geographic regions with hills and mountains, and the existing power restrictions mean, that even in open fields, we will have to deal with a lesser working range than maybe desired. Also, radio-based data links will always suffer from shadowing and signal loss.

2.1.2 Transmission of differential data corrections using mobile Internet

Due to the bandwidth made available by mobile phone operators for the use of data transmission, researchers managed to develop a protocol that allows transmission of differential corrections to GPS receivers via mobile Internet.

Differential data corrections are transmitted via the Internet using the NTRIP protocol (Network Transport of RTCM via Internet Protocol). This protocol was developed by the Federal Agency for Cartography and Geodesy of Germany. Using this protocol, there are three classes of objects that are communicating between each other:

- GPS station servers that assure differential corrections;
- Transmitters that compute differential data corrections and transmit them;
- Users that have access to these differential data corrections.

At the beginning, data was transmitted with the help of the mobile Internet using the GPRS connection (General Packet Radio Service). In the following years, new ways were developed to transmit the data. These are EDGE (Enhanced Data rates for GSM Evolution), CDMA2000 (also known as IMT Multi-Carrier (IMT-MC)) and UMTS (Universal Mobile Telecommunications System) - all of these being third-generation (3G) mobile telecommunications technologies.

Development of these technologies assures a higher data transfer speed, but mobile telecommunication companies don’t offer the desired coverage for RTK surveying purposes. It is understood that the low demand in rural areas do not lead to a demand priority.

2.1.3 The retransmission of differential data concept

Considering that the two methods of transmitting differential corrections data do not accomplish all the time the users’ requirements, a concept named retransmission of differential data concept [2] which integrates them was developed. It can be used due to a software developed by Federal Agency for Cartography and Geodesy of Germany.

The integration of the two methods consists in:

- the identification of an area relatively close to the working place, an area where we have access to the Internet,
- reception and the transmission of differential data corrections via radio waves.

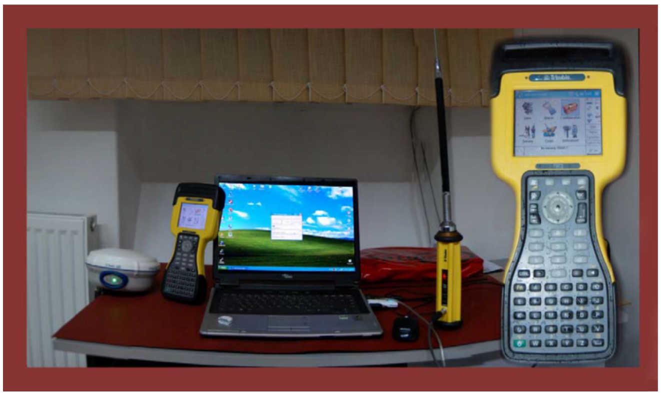


Figure 34：The equipment used to apply the retransmission concept

3 CASE STUDY

The use of GNSS Machine Control systems, besides an innovation related to the traditional way, present some problems described above.

We have used the national network of permanent stations ROMPOS for determining the position in real time in order to remove some of the issues presented above. In this case, there is no need to install a permanent reference stations in the area of work, communication and positioning solutions being much more accessible and stable.

Based on the above considerations, in pursuit of the study case and evaluation of the performance of a machine control system, we experienced such a system installed on a motor grader.

The motor grader HBM NOBAS BG 160TA-T4 equipped with a Trimble GNSS 3D GCS900 system (**Figure 33**) has been tested on a building site organized for road rehabilitation in Certesti area, Galati County, Romania.

Figure 35：Test area

Figure 36：Permanent station installed on Certesti building site

For testing purposes, an interested area has been selected. There was produced an accurate local DTM. A sample design of a road track including a curved part has been set up for test (**Figure 35**).

In a preliminary phase for the determination of the ground points, we need a GNSS base station that should be mounted daily. In order to simplify the procedure, the base station is often georeferenced in an approximate way provided that the monumentation and the assigned coordinates do not change from one day to another [1]. The ETRS89 position of the ground points were obtained by RTK GNSS surveying.

These points have a double set of coordinates: in local system and in reference station system. The system software provides the estimation of a set of transformation parameters between these two systems. From now on, the operator works in the local system, following the design DTM and drawings loaded on the PC and visible on the screen. The machine position, computed in RTK mode in the ETRS89 datum, is automatically converted by the software into the local system [1].

3.1 Land leveling using a machine control system equipped with GNSS receivers using local reference station

When the system receivers are initialized, on the control unit display appear the machine position and the excavation/ filling values required for the land to take the designed shape.

Figure 37：Machine position and excavation value on the control unit display

GNSS receivers receive the information transmitted via radio, by the reference station, on the UHF band, for determining the position of the blade machine in real time (using the technique OTF) with an accuracy of 2 - 3 cm. This position is compared in real time with the position of the model design.

3.2 Land leveling using a machine control system equipped with GPS receivers using national network of permanent station ROMPOS

It was necessary to modify a part of the machine control system in order to use the national network of permanent stations ROMPOS on this test.

In order to satisfy the desires, the system changes consisted in adding additional equipment such as a dual frequency receiver Trimble R6, a PDL450 radio transmitter, a radio modem and a laptop on which was previously installed GNSS Internet Radio software.

The dual frequency receiver Trimble R6 was necessary to receive differential data corrections transmitted by the NTRIP protocol, by using an internal GPRS modem.

The laptop and Trimble R6 receiver device has been installed near the testing area for the test, connecting it through a serial cable to a PDL450 radio transmitter, with the purpose of distributing the network differential data corrections all over the area, reaching the operating machine.

In order to connect directly to a permanent station to obtain differential correction, we chose the VRS method for different reason: allows the correction data stream to by optimize during connection – every second we obtained an update and is providing additional information on ionospheric and geometric error components in the area near the rover location.

During the test, some problems have been encountered: a poor GPRS coverage over the test area and some interference on the radio transmission. Both have been overtaken: the first by changing the device location and the second by changing the radio transmission frequency [1]. Another problem was the mode of transmission of differential correction data stream by the national network of permanent stations ROMPOS.

For this test we chose the transmitting stream correction RTCM 3.0 because of the compact and transfer speed and also for the necessity of a bandwidth of 2742 bps/s.

Even if the motor grader equipped with GNSS receivers has leveled the test area using differential correction data provided by the base station, another leveling operation has been started; this time using differential corrections provided by the national network of permanent stations (VRS NRTK) in order to check the correspondence between the executed and designed digital terrain model for the test area.

The mean differences between the two digital models, the one carrying out and the design is about 5 cm, as it can be noticed in the **Figure 38**. This difference can be justified by the fact that the VRS software is selecting the closest permanent station. In this situation the closest permanent station (Targu Bujor) is situated at 50 km by the test area.

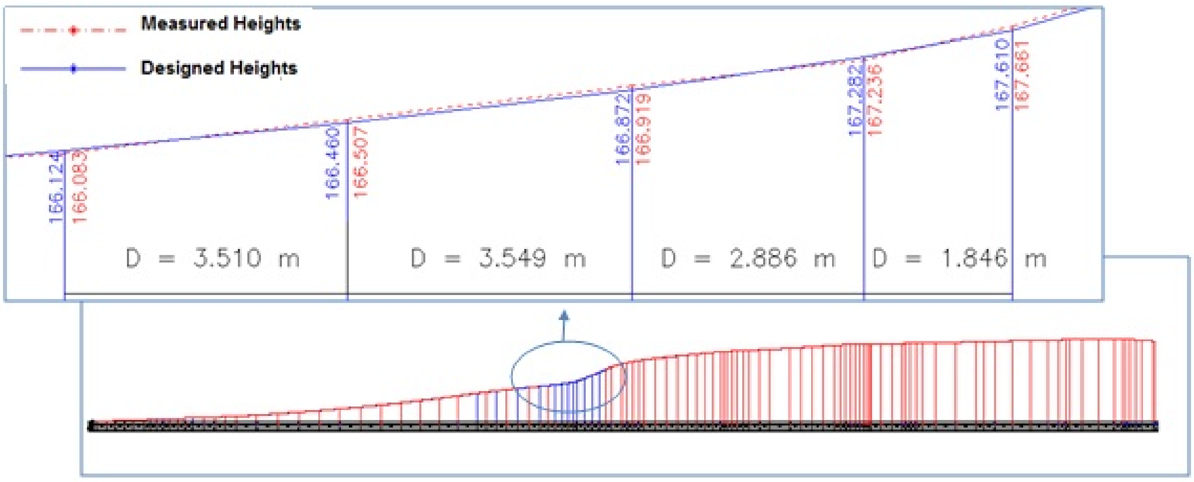


Figure 38 : The distribution of level differences

4 CONCLUSIONS

The experiment performed and referred above represents the first test realized in Romania in order to evaluate the possibility of using the real time positioning services supplied by the ROMPOS permanent network for machine control applications.

The result of the test performance confirms that the differential data corrections provided by the Romanian Positioning Service in machine control systems have some advantages [1]: the elimination of the reference stations installed on the working place (one for every working place) reducing cost, the possibilities of using more than one GNNS machine control systems in every area where there is a permanent stations network coverage, cost reduction regarding the area prospections and quality control process.

REFERENCES

Journal articles:

[1] FASTELLINI, G., RADICIONI, F., STOPINI, A.: *Network corrections for Machine Control, http://www.gpsworld.com/machine-control-ag/network-corrections-machine-control-8479*

[2] SUBA, S., NISTOR, S., JOCEA A.: *The analyze of the transmission possibilities of differential data corrections in the given radio infrastructure context,* RevCAD – Journal of Geodesy and Cadaster, University “1 Decembrie 1918” Alba Iulia, Romania, 10/2010, p. 177-184, 2010

Links:
www.rompos.ro, last accessed on September 30, 2011

www.trimble.com, last accessed on August 8, 2010

Data Management and Communication

Digital Construction Sites Connected Community Powered by Trimble

Norbert Mattivi, Business Development
Trimble, Germany

Abstract

The Connected Community from Trimble is a web-based service that enables construction businesses to manage and share information via their own unique web site, connecting the office to the field. Ideal for organization and project management, the Connected Community introduces communication and collaboration capabilities to significantly increase efficiency and productivity.

By hosting all of an organization's information centrally, the Connected Community connects people, devices and systems. Each organization can grant access to internal and external users. Members can then communicate and collaborate instantly regardless of their location.

Two-way data synchronization via rugged machine gateway allows productivity data to be retrieved from the machine, new design data to be sent from Business Center - HCE wirelessly to the machine and in-field survey equipment to be synchronized. It extends GNSS coverage with Internet Base Station Service (IBSS). Additionally, remote virtual support with Trimble Assistant for reduces downtime.

Asset Management and Fleet Management: The VisionLink(TM) fleet and asset management solution features management tools combined with GPS-based positioning and cellular technology to provide near real-time equipment performance information, as well as site productivity data. VisionLink provides an overview of machine health, fuel management, working utilization and site productivity.

Keywords

Connected Site, Information process, Integrated workflow, Real-time communication and collaboration

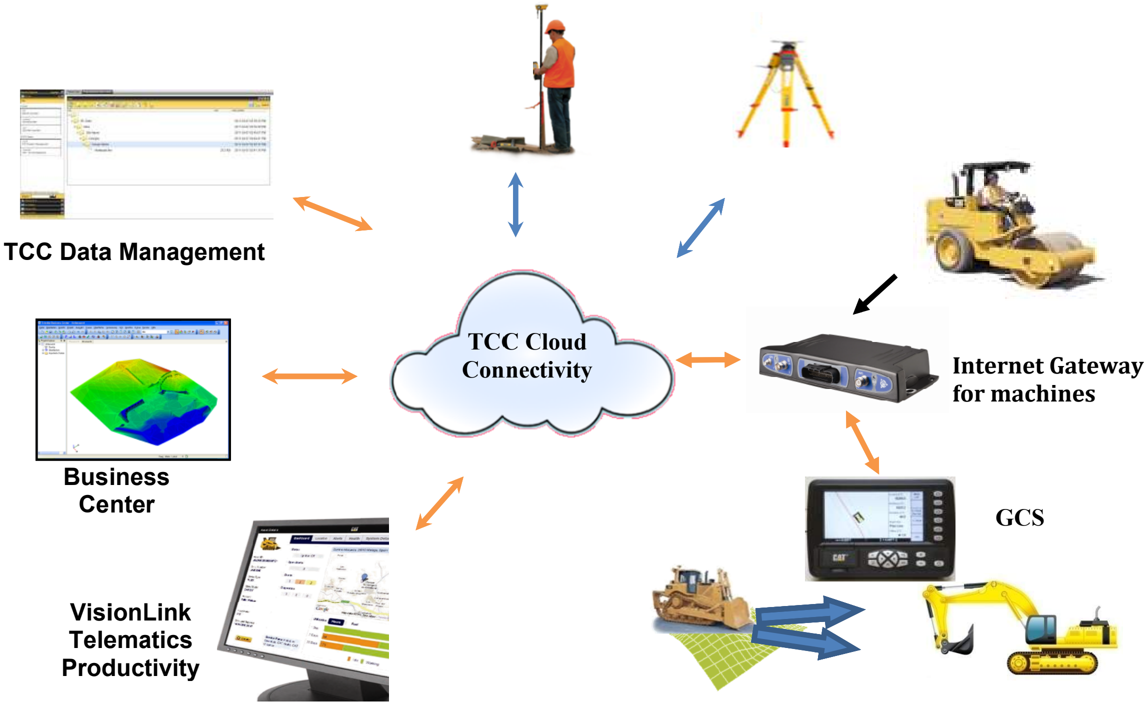
1 INTRODUCTION

Over the last 30 years the use of software in the construction industry has grown as a set of discrete applications that create efficiencies in specific areas and solve individual problems. Today, with an ever-increasing search for improved efficiency, these discrete applications are not enough. Through the combination of increased computing power, high speed graphics, the Internet, wireless communications and GPS it is now possible to integrate the workflow and provide collaboration and information tools that create a new level of productivity.

Connected Community from Trimble is a web-based service that enables construction businesses to manage and share information via their own unique website. Ideal for organization and project management, Connected Community introduces communication and collaboration capabilities to significantly increase efficiency and productivity.

- Reduce or eliminate field to office drive-time
- Information accessible by members anywhere, anytime
- Project progress and site activity tracked in real-time
- Secure online storage of all organization and project information

2 THE TRIMBLE CONNECTED SITE CONCEPT

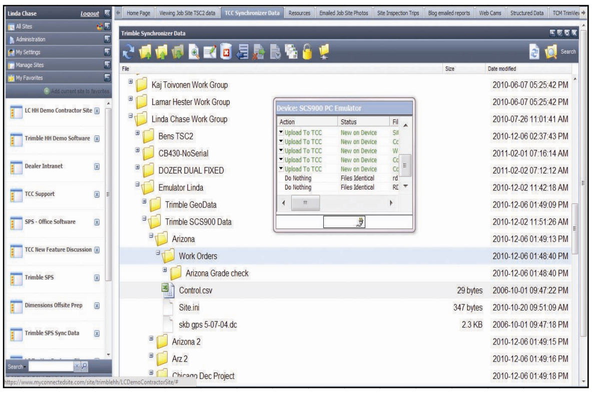


What TCC Trimble Connected Community provides

Real-Time Communication and Collaboration
By hosting all of an organization's information centrally, the Connected Community does what no other information management system in the construction industry can—it connects people, sites and devices. Each organization can grant access to internal and external users. A contractor can set up varying levels of access to project partners or guests such as engineers, subcontractors, suppliers, the head office and clients. Members can then communicate and collaborate instantly regardless of their location.

Connecting Field and office with real-time information

- Two-way data synchronization back-bone
- Wireless work order synchronization between the office and Trimble® SCS700 and SCS900 Site Controller Software
- Send Business Center - HCE designs wirelessly to machines using Trimble GCS900 Grade Control System and receive productivity data back from the field
- Extend GNSS coverage to remote or offsite areas of projects with Internet Base Station Service (IBSS)



Notifications and Alerts
A Connected Community site can be set up to notify members of new information. This feature spares members from manually checking for updates, and ensures critical information is communicated promptly

Visibility, Transparency…Control
The Connected Community lets you view every corner of your operations. It can also make your own select practices transparent, thereby providing more open communication with partners and clients.

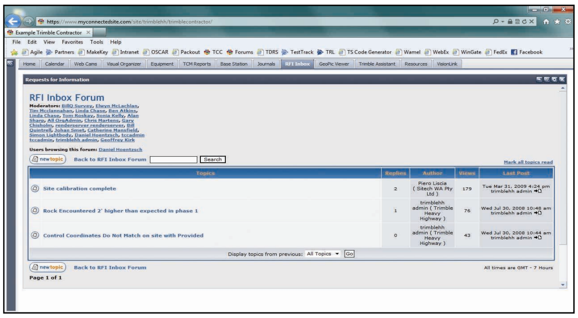
Because all information and communication is stored and recorded, the Connected Community prevents miscommunication, ensures accountability, and enables conflict resolution. For example, a Request for Information (RFI) is made by site staff to the engineer using a Connected Community RFI Forum. The engineer responds with an answer and the entire communication thread is recorded.

Data visualization

- Visualize site measurements and stake-out information
- Site inspections and site visits recorded
- Monitor projects with photographs
- Site camera integration and real-time data streaming
- Overlay Business Center designs and cut/fill maps onto Google Maps or digital imagery
- Real-time personnel and device tracking

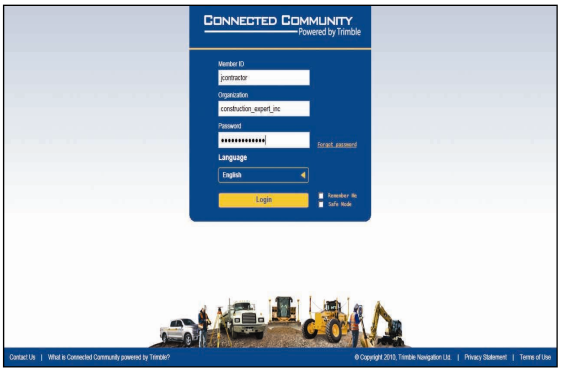
Online collaboration

- Email notifications and alerts
- File sharing and remote collaboration
- Site file synchronization
- Manage and respond to Requests for Information



Comprehensive security

- Data is securely hosted
- Authentication prevents unauthorized access
- Secure Internet log-in
- Encrypted data transmission ensures your data stays secure even while in transit
- Disaster recovery
- Unlimited data retention and upgradeable storage plans
- Data back-up for your business servers



3 VISIONLINK – ASSET MANAGEMENT AND PRODUCTIVITY

The use of advanced asset management tools, such as the VisionLink™ system, monitor machine health, getting the critical information to your equipment manufacturer's dealer or service agent to enable preventative maintenance to occur and reducing unscheduled maintenance. Additionally, today's telematics systems also provide equipment utilization information that enables more efficient use of expensive equipment fleets.

Fleet and asset Management

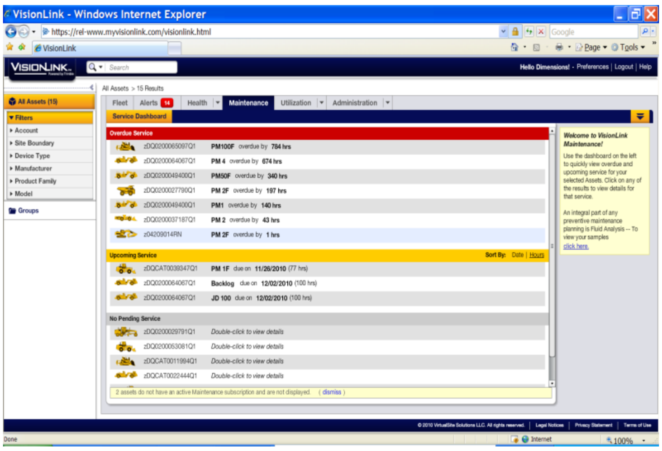
The VisionLink fleet and asset management solution from Trimble allows users to quickly and easily view all of their equipment, regardless of make or model, in a secure, web-based application hosted by Trimble.

- Know when and where equipment is working
- Improve logistics for fuel, transportation or service dispatch with equipment location features
- Manage costs by monitoring fuel usage
- Manage repair and maintenance
- Improve bidding accuracy by profiling machines

VisionLink features user-friendly management tools combined with GPS-based positioning and cellular technology hardware to provide real-time equipment information to senior executives, fleet managers and rental equipment managers.

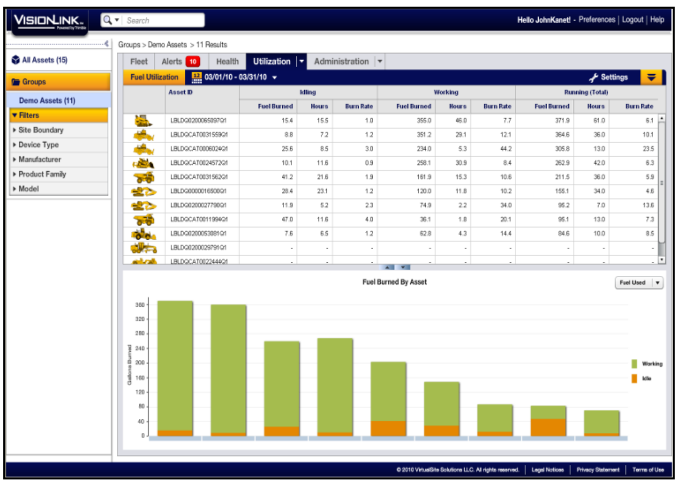
Maintenance Management

The maintenance management functionality in VisionLink can help reduce owning and operating costs by applying an effective and efficient maintenance program. Inspections, backlogs and planned component services can be managed and documented. Users receive alerts based on recommended maintenance intervals, so they can see what's been done, what's due now and avoid overdue service.



- Setup Schedules
  - Hour interval
  - Parts list
  - Service item checklist
  - Maintenance history
- Monitor Machines
  - No Service
  - Upcoming Service
  - Overdue Service
- Improve Planning

Utilization and Fuel Burn



- Utilization
  - Run time
  - Planned run time
  - Idle time / Working time
- Fuel Burn Rate
  - Idling / Working / Running
- Compare
  - Same type of machine
  - Different types of machine
  - Operators
- Create best practices

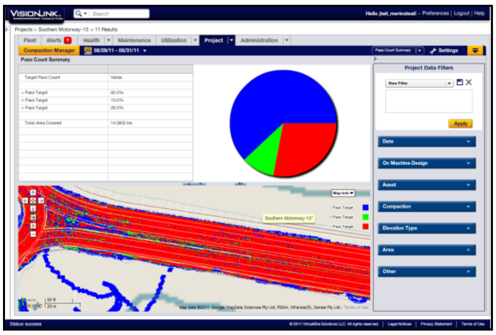
- Improve fuel cost management

Project Progress and Activity Tracking – 3D Project Monitoring

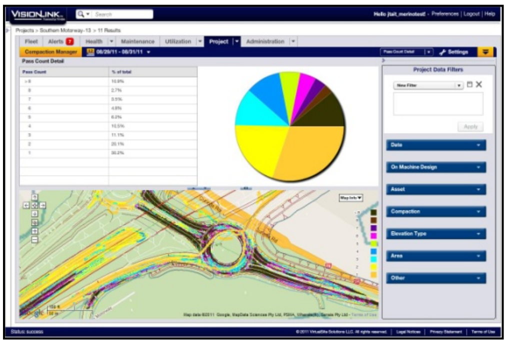
The Trimble Connected Site™ offers multiple ways to monitor project progress. Tools such as journals, calendars, forums, digital image management and site cameras enable members to record and view activity. All records are securely stored and archived for easy access at any time.

For example, if a fleet manager identifies a reduction in productivity, they can "virtually visit" the job site using site cameras, or review weather data and supervisor journal entries to help diagnose the productivity problem.

Intelligent compaction



For example, soil compaction has been tested for many years, based on a sample points across a site. Today's **intelligent compaction** systems will measure the number of passes, the thickness of the layer and the degree of compaction for every square meter across the site, allowing the operator to see in real time that the target compaction has been achieved. The project supervisor can then extract reports on the quality and any areas of undercompaction prior to authorizing the next step in the process.



- Tracks compaction by layer
- Monitors pass count over entire surface from one or multiple machines against a Target Pass Count value
- Highlights over- or under-compaction areas
- CMV values
- Monitors the Compaction Meter Values (CMV) over the entire surface from one or multiple machines
- Highlights compaction inconsistencies such as softspots

Volumes Manager

The Volumes Manager functionality reduces survey work for periodic surveys by using machine captured data for calculation of volumes. It allows surveyors to focus on more value add work.
Near real-time project status for earthworks, grading and finishing operations are provided online. With Volume Manager a contractor can track the entire earthmoving operation, including over-excavation volumes and overfill / cutback or surcharge which can be missed by traditional survey methods. It helps manage machine use, deployment and reallocation

- Volumes Manager provides surface model creation from machine measured data that can be used to display
- Coverage / Elevation maps
- Cut/fill maps (measured compared to design)

- Volumes Manager also provides the ability to compute volumes between measured and design data, or between measured and measured data at different times
- Surface models are improved through use of Trimble's blade on-ground indication system for graders and dozers

Connected Machine – GNSS Corrections (VRS/IBSS)

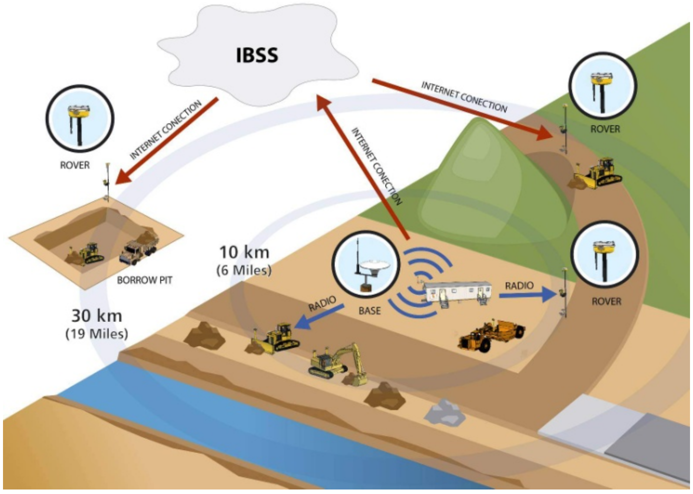
GNSS correction data for machine guidance and surveying can be received via radio or by using a position service like VRS. VRS eliminates the need for a base station by utilizing commercial or government sponsored VRS networks.

IBSS (Internet Base Station Subscription) is a new web service delivered through the Trimble Connected Community (TCC). The service provides Trimble GPS base station owners with the ability to simply connect their base stations to the Internet Base Station Service (IBSS) using LAN or cellular modem technology and serve GPS corrections over longer distances than normally possible with radio and repeater technology.

IBSS is a single baseline RTK solution. The errors associated with IBSS are Horizontal 1cm + 0.5ppm and Vertical 1.5cm + 1ppm

On a large site with a dedicated base station the rovers and machines can work locally within range of the 450MHz radio, this is usually limited to about 1.5-3 km but can be extended up to 8-9 Km in some circumstances by using repeaters.

With the addition of IBSS the base range can be extended to cover auxiliary sites such as borrow pits which are beyond the 450 MHz radio reception up to 30 km.



4 CONCLUSIONS

With the use of construction-specific collaboration tools such as Connected Community from Trimble, transfer of new and completed work orders, the latest design changes and as-built measurements or machine production data between field workers, machine operators or office managers can be performed, anywhere and anytime.

With the potential to save time and cost at every stage, and virtually eliminate some steps in the plan design, construct and operate process, the use of an integrated software-based digital workflow can improve the efficiency and sustainability of highway construction, resulting in the earlier completion of a higher quality highway at a lower cost.

Electronic Assistance for Straw and Hay Harvesting with Large Square Balers

Jan Horstmann
Maschinenfabrik Bernard KRONE GmbH, Spelle, Germany

Abstract
Electronic driver assistance, data management and process improvement become more and more important for agricultural purposes. By help of the ISOBUS standard comprehensive processes can be supported by electronics and IT-components.

The manuscript contains a project presentation of operation planning, electronic field- and job management, field navigation, machine data collection and bale- and yield traceability with large square balers. Linked to the project iGreen KRONE established a process and information system to cover all aspects of hay and straw harvesting.

The script contains a report of a field evaluation for electronic assistance for straw and hay harvesting with digital operation planning, job management, field navigation, machine data management, bale tagging and logistics planning.

Keywords
Agriculture, straw, hay, harvesting, GPS, construction industry, KRONE, BiG Pack, large square baler, field navigation, job management, organisation, bale tagging, traceability

1 INTRODUCTION

Straw and hay harvesting is a very dynamic process with short harvesting times and huge workloads for drivers and machinery. The company Maschinenfabrik Bernard KRONE GmbH in Spelle developed and combined several tools and techniques by help of electronic- and IT-systems to support and optimize straw and hay harvesting.

The typical workflow of straw and hay harvesting starts with a farmer calling his service provider (contractor). An order for harvesting is placed by telephone together with a rough description of the field location. The dispatcher tries to find out the next available baler and gives the driver a short description of what to do by phone. The driver searches the new field location and once reached the driver directly starts working. The documentation of the work is done manually on paper. Back in the office the driver provides the accountant with the paperwork. Afterwards the accountant has to create bills and has to copy all data into an IT-based program. This typical process contains high error potential, time wasting because of phone calls and misleading instructions of field locations. Documentation has to be made manually and traceability of the yields and its quality is not possible.

In order to sharpen this process and to provide automatic documentation with traceability of yields KRONE uses internet technologies, geo-information, GPS, GSM and RFID components to provide the involved persons an optimum of support and assistance.

This lecture will give an overview of the actual situation, technical possibilities, a report of field evaluations and will conclude with a vision for the future.

2 DEFINITION OF THE PROCESS

In this chapter field evaluations of the newly developed process and assistance system are described. In order to structure the way of harvesting operations the field borders are the first entrance point for process improvements. The farmer has to provide the field border electronically (or the contractor has to draw them by help of aerial photos or satellite pictures). The farmer places the order at his contractor either by phone, by GeoFormular or by smartphone app.

The campaign planner at the contractor's site collects all jobs and assigns a suitable machine and drive and arranges the jobs in a proper order.

The new taskset can be sent online to the machine terminal or can be transferred by USB-stick. The driver imports the taskset and starts the navigation to the field and the documentation is done automatically based on the planned tasks.

Finished tasks can be transferred back to office including all machine data and times used for billing.

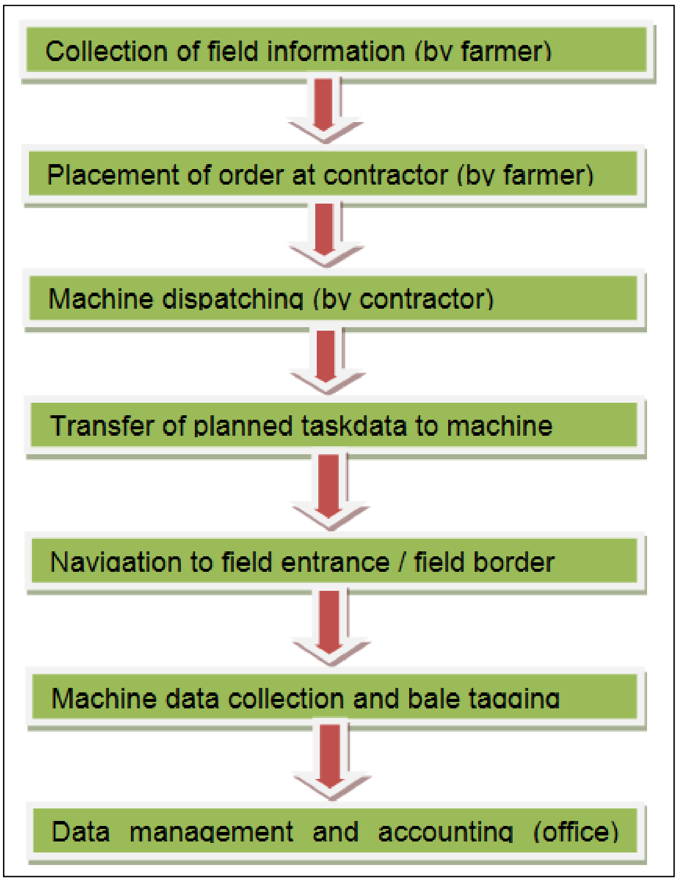


Figure 39: list of actions for data management

The following picture shows the data flow of the newly developed process.

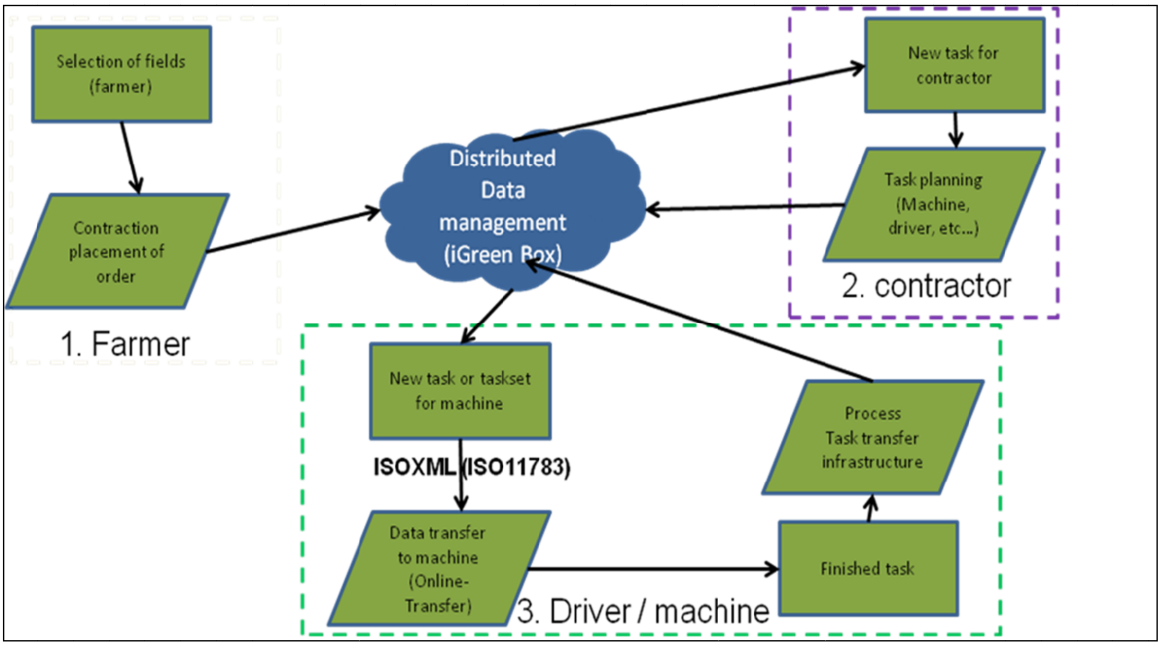


Figure 40: process overview

3 FIELD EVALUATION OF THE PROCESS

To initiate the process of electronic assistance for hay and straw harvesting all available field borders with the field names and numbers have to be collected into one data storage. Every farmer knows his fields best and over 90% of the agricultural cultivation subsidy applications are already sent out electronically. If farmers or service providers are not willing to handle with the cultivation application data the field borders can be set by manually clicks on aerial images or satellite pictures. The screenshot below shows exemplarily how to record field borders by help of the GeoEditor developed within the iGreen project by IIS.

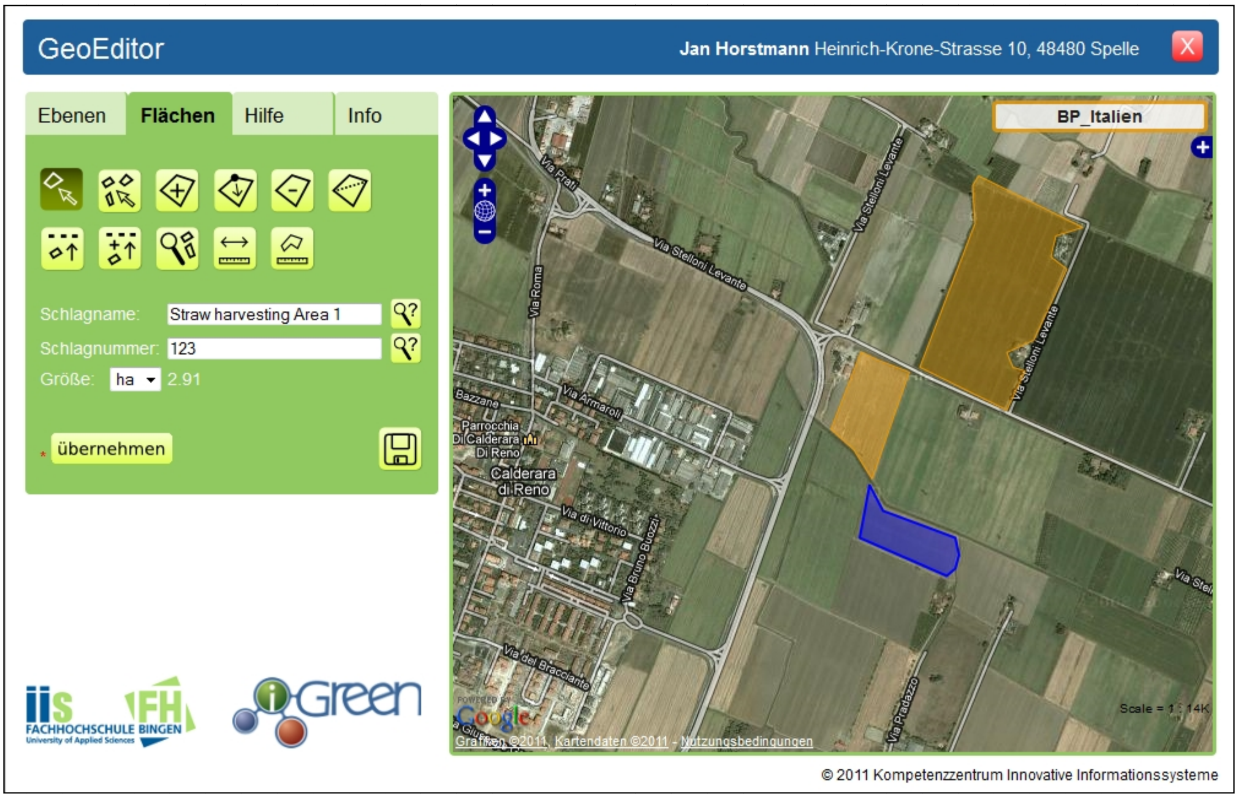


Figure 41: collecting the field borders

The farmer can place his order at the contractor either by telephone (therefore the field borders have to be transmitted before the call) or by help of a smartphone app. The best and most appropriate way to place the order is be help of the GeoFormular. By help of the GeoFormular all relevant harvesting information can be set, e.g.

- Bale size
- Number of bales wished
- Preshop (Yes / no)
- Comments and hints
- etc…

The information of the GeoFormular can easily be used to configure the machines and to help the driver finding the field and doing a qualified documentation.

After the order is placed at the contractor the campaign planner gets an email and is able to do the assignment of drivers and machinery. The order of the fields to be harvested can already be planned and optimized before starting the work. The data (tasksets) can be transferred online to the machine terminals.

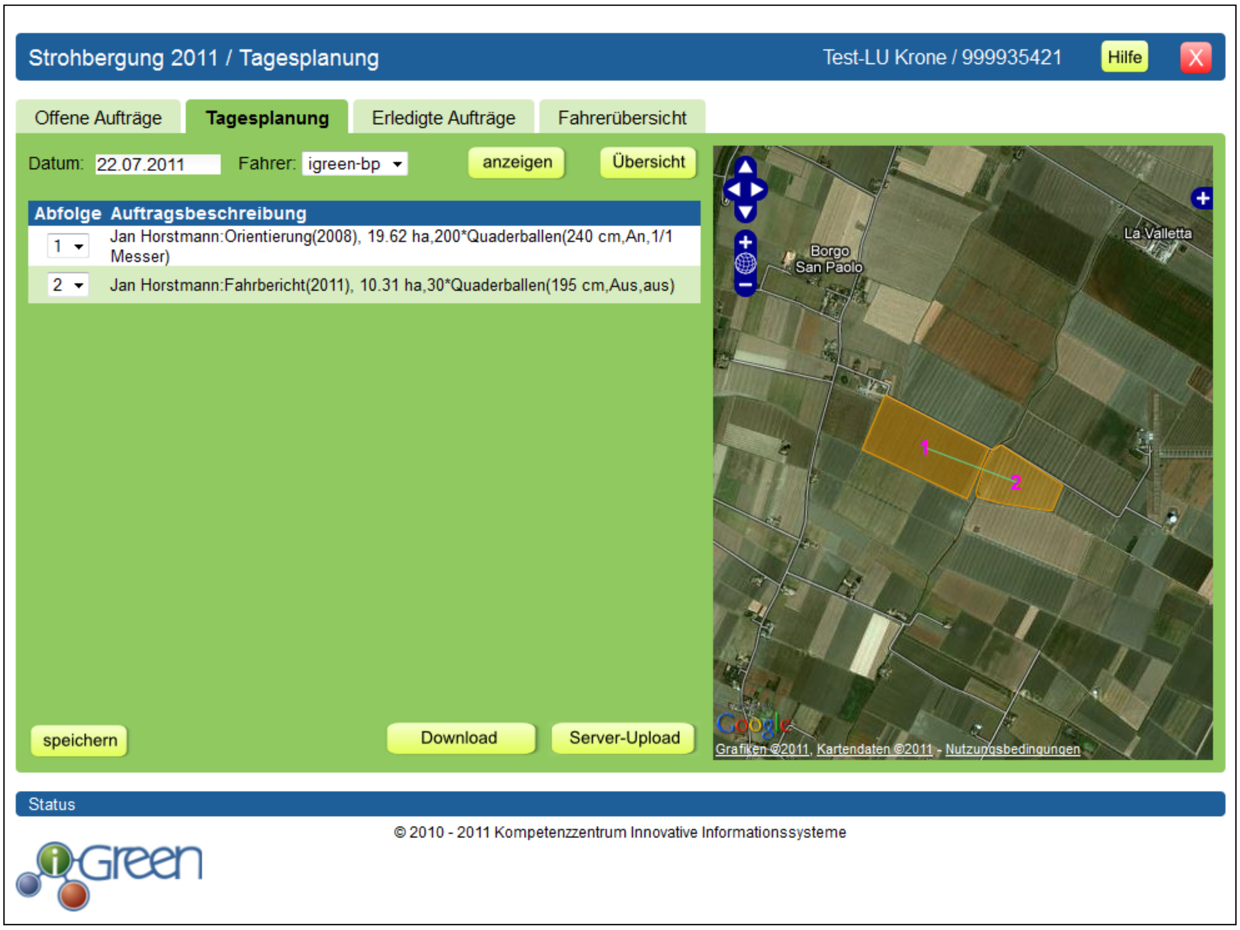


Figure 42: ordering and gathering the jobs for in the office

Process improvement at straw and hay harvesting can only be done by using standardized equipment. Therefore KRONE equipped the large square baler with an ISOBUS terminal (CCI 200), a GPS receiver and a GSM-Box. The ISOBUS terminal has special applications (app concept) installed to allow operating the machine, doing field navigation, online connectivity, task management and machine data collection (ISOBUS virtual terminal VT and task controller TC). Data exchange between machine terminal and farm management systems follow the ISO11783 standard called ISOXML. A taskset with planned tasks is transferred by an internet connection from the office to the machine.

Finished task sets can be transferred from the terminal back to office by the internet connection as well.

Figure 43: equipment

The CCI Terminal is not only used for steering and controlling of the machine. It is increased by additional functionalities like job management, internet connection management and field navigation. The CCI Terminal can be equipped with different plugins / apps which enable the customer driver assistance and support for the daily work.

By help of the task controller of the ISOBUS terminal the driver becomes his instructions electronically. Paperwork and writing down all information is not necessary. By help of the planned tasks the driver can active the machine data collection by starting the tasks. Once a task is started all relevant information like start time, end time, bale counter, bale weights, bale moisture and bale positions are recorded automatically.

Figure 44: Tasklist at CCI Terminal

Lots of important information can be extracted out of the ISOXML task sets which where transferred to the machine terminal. Every task contains the customer with all master data together with the fieldname, field size and the field borders. By help of this information a field navigation system helps to find the correct area. KRONE integrated the field navigation into the ISOBUS task controller of the CCI terminal. By help of special agricultural maps with truck navigation increased by dirt tracks the driver can be navigated to the field border or to the field access point. The route path takes under consideration the weight, length, heights and other regulations like speed requirements.

The following picture shows the Fieldnav navigating to a planned ISOXML task for straw harvesting.

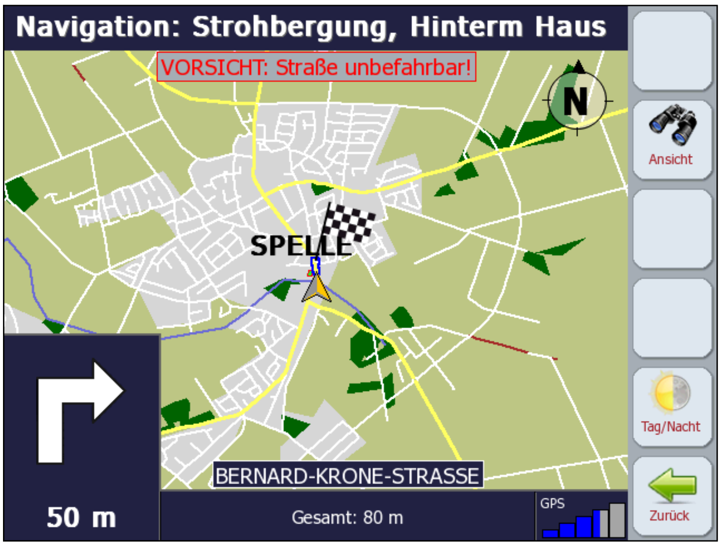


Figure 45: field navigation system

The ISOBUS terminal is equipped with a mobile internet connection via GSM. If the large square baler works on areas where GSM networks are available the terminal can sent the actual position and the working states of the machine to a fleet management portal. Dispatchers are able to monitor the working progress and can easily decide how to plan further tasks. By help of smartphones mobile dispatchers can find the balers without phone calls to the drivers. Mobile documentation on smartphones together with picture documentation of the yields is possible as well. Fleet management reduces the amount of phone calls and enables the driver to concentrate on the harvesting work.
The actual working positions can be seen in a web based fleet management. By help of modern smartphones the baler could be followed or reached as target for a navigation system (e.g. to deliver consumables like twine or foil). Logistics planning of trucks and telehoist load luggers could be optimized by help of this information.

Figure 46: fleet management by web portal

To be capable of the rising duty of documentation and traceability as well as distinguishing between qualified yields and poor yields RFID chips are used to label the bales with an bale-id, bale weight, bale moisture and it's creating position, date and time.

Additionally to the yield documentation done by ISOBUS task controller all straw and hay bales can be tagged with a RFID chip. A special tagger system bonds a RFID chip to the twine of the bale and writes all relevant data like weight, creating position and moisture on the RFID chip. By help of the tags the yield can be categorized into A class and B class yields. Service providers and contractors can proof their work and the quality of the bales.

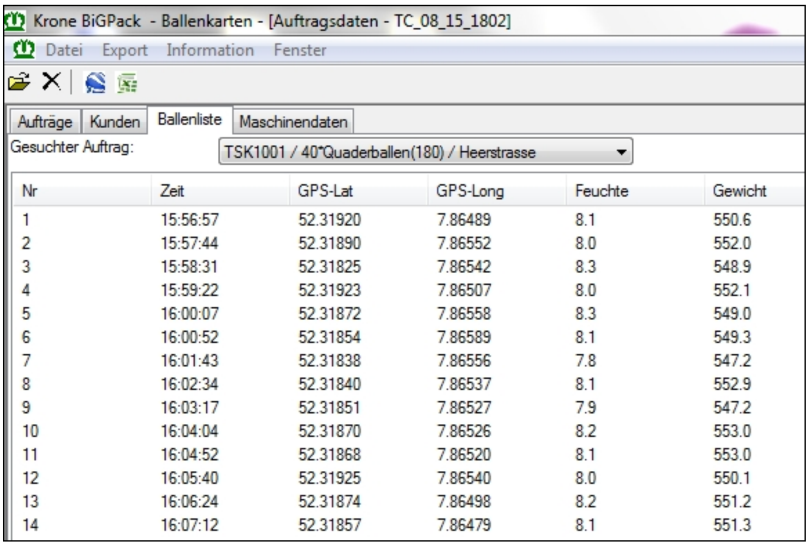
Figure 47: RFID-Tag written by antenna at baler chute

4 CONCLUSIONS AND VISIONS FOR THE FUTURE

During the straw and hay harvesting the ISOBUS terminal collects relevant information from the balers' job computer. This data can be transferred by USB stick or internet connection into the office. Yield mapping with detailed positions of the bales help to plan bale accumulation and help to optimize logistics. Bale weights and moisture data help to categorize the yields. All relevant information together with fieldname, customer name and further more can be used to do the accounting and billing. The picture stated below shows yield mapping for large square balers.

Figure 48: bale map

The bale map can be transferred to a bale list and can be provided to Excel-sheets or billing applications.



Nr	Zeit	GPS-Lat	GPS-Long	Feuchte	Gewicht
1	15:56:57	52.31920	7.86489	8.1	550.6
2	15:57:44	52.31890	7.86552	8.0	552.0
3	15:58:31	52.31825	7.86542	8.3	548.9
4	15:59:22	52.31923	7.86507	8.0	552.1
5	16:00:07	52.31872	7.86558	8.3	549.0
6	16:00:52	52.31854	7.86589	8.1	549.3
7	16:01:43	52.31838	7.86556	7.8	547.2
8	16:02:34	52.31840	7.86537	8.1	552.9
9	16:03:17	52.31851	7.86527	7.9	547.2
10	16:04:04	52.31870	7.86526	8.2	553.0
11	16:04:52	52.31868	7.86520	8.1	553.0
12	16:05:40	52.31925	7.86540	8.0	550.1
13	16:06:24	52.31874	7.86498	8.2	551.2
14	16:07:12	52.31857	7.86479	8.1	551.3

Figure 49: bale list

By help of useful combination of ISOBUS systems KRONE managed to create significant process improvements for hay and straw harvesting. Drivers, dispatchers and involved people can use the process data to improve the documentation, to minimize mistakes and to reach a qualified traceability of yields.

The bale maps can be used to optimize the bale collection, the next step will be to provide the telehoist load lugger driver with a suggestion of bale pickup ways. The routes could be determined by help of the Greedy- or Honeybee algorithm.

The data of the large square baler can be combined in future together with the tractor data; therefore an ISOBUS compliant tractor with taskcontroller client including fuel consumption and PTO speed is necessary.

Existing farm management systems (FMIS) have to open up in order to collect the data out of decentralized data storage (like the iGreen Box). The classical farm management systems have to be able to calculate with ISOXML and should be able to visualize bale maps and extract automatically the billing information out of the balers ISOXML data.

Electronic assistance for straw and hay harvesting helps to optimize the harvesting processes and is a first step to sustainable high productivity agriculture.

REFERENCES

Books:
ISO Office: *ISO11783*, Geneva, 2009.

Links:
iGreen-Project Homepage
www.igreen-projekt.de, last accessed on November 23, 2011

University of applied Sciences Bingen, iGreen GeoEditor
http://lw.igreen.fh-bingen.de/geoEditor3x/, last accessed on November 23, 2011

Competence Center ISOBUS
www.cc-isobus.com, last accessed on November 13, 2011

ISOXML Humanizer iGreen

http://demonstrator.igreen-projekt.de/isoxml-humanizer/, last accessed on November 17, 2011

Lacos Computerservice Fieldnav
http://www.lacos.de/cms_neu/, last accessed on November 14, 2011

Arvato Systems FarmPilot
http://www.farmpilot.de/, last accessed on November 05, 2011

KRONE Software
http://www.krone.de/software/, last accessed on November 24, 2011

Control Algorithms

This Paper has been accepted as a "peer reviewed paper" by the scientific committee

Formation Control Algorithm for a Fleet of Mobile Robots

Pierre Cartade[1], Roland Lenain[1], Benoit Thuilot[2,3], Michel Berducat[1]
[1]Cemagref, 24 avenue des Landais, 63172 Aubière, France
[2]Clermont université, Université Blaise Pascal, Institut Pascal, BP 10448, 63000 Clermont-Ferrand, France
[3] CNRS, UMR 6602, Institut Pascal, 63171 Aubière, France

Abstract
The necessity of decreasing the environmental impact of agricultural activities, while preserving the level of production to satisfy growing population demands requires investigation of new production tools. Mobile robots may constitute a promising solution, since autonomous devices may allow increasing production levels, while preserving the environment thanks to their high accuracy. In this paper, the use of several autonomous mobile robots to perform field operation is investigated. In particular, predictive techniques are also proposed to account for delays induced by low-level actuators. Capabilities of the proposed approach are investigated through full scale experiments.

1 INTRODUCTION

The continuous advances in autonomous mobile robot control (concerning both a single robot [4], as well as multi-robots [1], [7]) offers new possibilities in terms of applications for every-day life improvement. For instance, the development of automated multi-robot fleets can benefit to many applications requiring to cover large areas [5], such as surveillance, cleaning, exploration, etc. It is particularly interesting in environmental applications such as farming, where the use of several light robots in the field may permit to reduce environmental impact while preserving the level of production. This constitutes a challenging problem as stated in [2]. Rather than considering numerous small robots, as in swarm robotics [11], a cooperation framework with a limited number of light machines seems preferable when field treatment is addressed: on one hand, some farming operations such as harvesting require quite large machines to achieve tasks properly, and on the other hand, it appears more tractable from a practical point of view (maintenance, monitoring, acceptability, etc). As a consequence, this paper is focused on formation control of several light robots executing operations in field (as illustrated in **Figure 50**), allowing the use of several autonomous entities instead of driving a sole huge vehicle.

Figure 50: Illustration of the application

In the considered applications, a reference path is defined by a leader vehicle, controlled either manually or autonomously. The shape of the formation is not considered as fixed, since the area covering may require a varying formation (tank unload, maneuvers, etc). Several approaches have been proposed for mobile robot formation control [8], [13], but they are mainly dedicated to structured environments. In contrast, the context of the considered tasks requires a high accurate relative positioning of the robots despite the numerous perturbations encountered in natural environment (skidding, terrain irregularities, etc). This is not addressed by classical approaches.

In this paper, an adaptive algorithm for formation control is proposed, relying on a reference trajectory defining a local relative frame. It decouples longitudinal and lateral dynamics with respect to the desired path: the advance of each robot along the reference path can be addressed independently from the regulation of its lateral deviation with respect to this path. Longitudinal control is based on the regulation of curvilinear inter-vehicle distances, while lateral regulation relies on an observer-based adaptive control approach as has been proposed in [14]. The control of the possibly varying formation gathers both control laws, enabling an accurate formation regulation for field operations, independently from the reference path shape and environment properties. In this paper an adaptive and predictive approach is proposed to reduce lateral overshoots occurring along curves and due to delay introduced.

The paper is presented as following. First the model of a robot including bad grip condition is proposed. As soon as sideslip angles are available by observation, this model can be used for control purpose. The adaptive control of each robot is then investigated in section 3. It permits an accurate servoing in steady state but overshoots occur when transient curvature phase due to neglected actuator setting time. To go further a predictive curvature servoing is developed in section 4 constituting the main contribution of this paper. The efficiency of the proposed control law is finally investigated through full scale experiments.

2 MOBILE ROBOT MODELING

The autonomous control of a fleet of mobile robots is considered with respect to a desired path, used as a reference frame for both longitudinal and lateral positioning of each robot. The objective is to ensure an accurate overall motion of the robots in a desired, but potentially varying, configuration along this chosen trajectory.

2.1 Model of a robot formation

The overall control strategy for the robot formation is based on the modeling proposed in **Figure 51** (two robots among n are shown). In this representation, each robot is viewed as a bicycle, as in the celebrated Ackermann model, see ([12]). The classical rolling without sliding assumption is not satisfied in a natural environment. As they affect robot dynamics significantly, low grip conditions reduce the path tracking accuracy. In order to account for this specific problem, two sideslip angles are added: β^F and β^R, respectively for front and rear axles. These variables are representative of the difference between the tire orientation and the actual tire speed vector direction. Longitudinal sliding is not here accounted, since in the considered applications, longitudinal guidance accuracy is not as critical as the lateral one.

Based on these assumptions, the notations used in the sequel which are depicted in **Figure 51** for the i^{th} robot are:

- Γ is the common reference path for each robot defined in an absolute frame (computed or recorded beforehand).
- O_i is the center of the i^{th} mobile robot rear axle. It is the point to be controlled for each robot.
- s_i is the curvilinear co-ordinate of the closest point from O_i belonging to Γ. It corresponds to the distance covered Γ by robot *i*.
- $c(s_i)$ denotes the curvature of path Γ at s_i.
- y_i is the lateral deviation of robot i w.r.t. Γ.
- δ_i is the i^{th} robot front wheel steering angle.
- l is the robot wheelbase.

- v_i is the i^{th} robot linear velocity at point O_i.
- β_i^F and β_i^R denote the sideslip angles (front and rear) of the i^{th} robot.

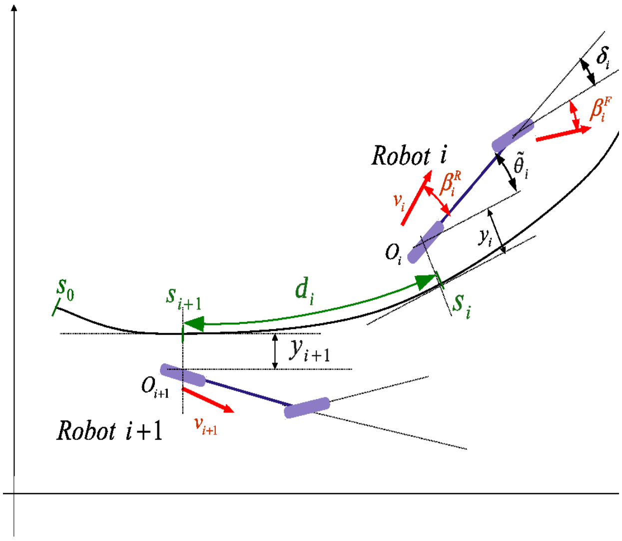


Figure 51: Longitudinal model of a robot fleet

2.2 Sideslip angle estimation

As sideslip angles integrated into robot model are hardly measurable directly, their indirect estimation has to be addressed. The observer-based approach detailed in [6] is here implemented. It follows the algorithm described in **Figure 52**, taking benefit of the duality principle between observation and control.

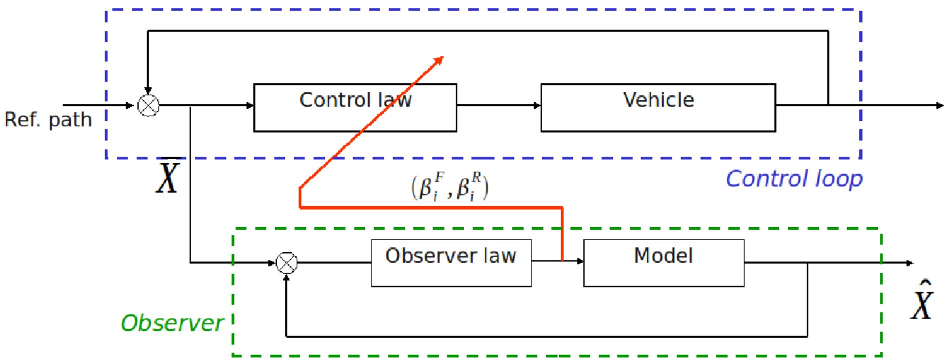


Figure 52: Observer principle scheme

2.3 Model exact linearization for control

Kinematic model has been extended to account for low grip conditions. Nevertheless, it is still consistent with classical kinematic models, such as considered in [12] and [14]. It can consequently be turned into a chained form, enabling then an exact linearization. Both longitudinal and lateral control can then be addressed independently.

3 MOBILE ROBOT FORMATION CONTROL

To address the control of a fleet of mobile robots in a path tracking context, the relative positioning of each robot with respect to the reference trajectory is achieved and then shared within the fleet via wireless communication. The control of each robot aims then at ensuring convergence to desired set points in terms of curvilinear offset (longitudinal control) and lateral deviation offset (lateral control).

3.1 Longitudinal control law

The objective of longitudinal control is to maintain a desired distance (denoted *d*) between curvilinear abscissas of successive vehicles. Each robot is controlled with respect to the curvilinear abscissa s_1 of the leader. This enables avoidance of an oscillating behavior due to error propagation along the fleet. However, for obvious safety reasons, the distance to the previous vehicle has also to be considered. Therefore, as proposed in [3], a composite error x_i equal to the distance to the leader vehicle e_i^1 in the nominal case, and smoothly commuting to the distance to the preceding vehicle e_i^{i-1} when the security distance is approached, is here regulated, see **Figure 53**. The i^{th} robot linear velocity v_i ensuring that x_i converges to 0, so that each vehicle can be controlled longitudinally, whatever the velocity of the leader.

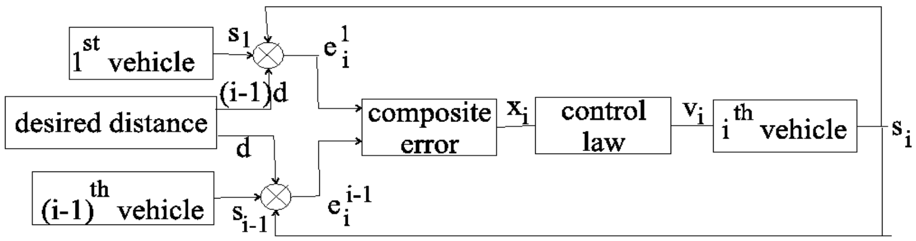


Figure 53: Longitudinal control scheme

3.2 Lateral control law

Once longitudinal control has been achieved, the one of the lateral positions can be addressed. In contrast to the classical path tracking problem, where the error is expected to be null, the lateral deviation of each robot in a formation has to converge to a non-null desired set point.

The steering control law of robot *i*, can be determined using a new variable $y_i^d(s_i)$, representative of its desired lateral deviation. The variable y_i^d permits definition of their lateral positions with respect to the global formation motion. Longitudinal and lateral relative positions of each robot can then be specified in the reference trajectory frame independently. The set point y_i^d has to be constructed to regulate a desired formation, in order to achieve a multi-robot task.

A first mode consists in taking $y_i^d(s_i) = d_i^y$, with d_i^y a constant chosen w.r.t. implement widths. It is completely satisfactory as long as vehicles are never side-by-side.

In contrast, when robots have to work side-by-side we propose the following definition of $y_i^d(s_i)$

$$y_i^d(s_i) = d_i^y + \sigma(y_{i-1})\left[y_{i-1} - d_{i-1}^y\right]$$

where σ is the smooth commutation function shown in **Figure 54**. Thus Robot *i* reproduces robot $i - 1$ deviation, if the latter exceeds a pre-specified threshold. Such a behavior permit to keep the formation when an important deviation is recorded while preserving the global formation free oscillating behavior.

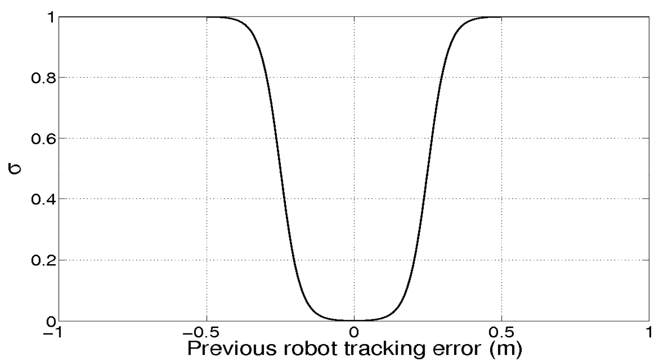


Figure 54: Shape of commutation function

4 PREDICTIVE CONTROL

When a vehicle enters a curve we observe transient overshoots in lateral deviations. They are mainly due to delays induced by low-level actuators, the delays depending of intrinsic properties of the actuators. To reduce such overshoots, we use predictive techniques.

More precisely, assuming that the overshoots are only generated by delays of the actuators in response to fast variations of the curvature, a predictive algorithm is designed, focused on the part of the control law linked to the curvature of the path.

4.1 Splitting the control law

In this purpose the control law of each robot can be split into additive terms:

$$\delta_{i=}\delta_{iTraj}+\delta_{iDeviation} \quad \text{with} \begin{cases} \delta_{iTraj} = \arctan(u_i) \\ \delta_{iDeviation} = \arctan\left(\frac{v_i}{1+u_i v_i+u_i^2}\right) - \beta_i^F \end{cases}$$

The first term (δ_{iTraj}) ensure the convergence of robots curvature to the reference path curvature. As the reference path curvature (or leader path) is known, the curvature variable can be anticipated. The second term ($\delta_{iDeviation}$) cannot be used in the predictive algorithm since the sliding and the resulting deviations are unpredictable phenomena.

4.2 Identification of the low-level dynamics

We propose a simplified model omitting the inertial phenomena. In this case, the low-level process that controls the orientation of the wheels can be considered as a second order process. Its properties can be defined by identifying the response to a step function of the steering angle. The instruction sent to the front wheel at instant n denoted by $\delta_{[n]}^C$ and the real steering angle denoted δ_n^R are linked by the following state equations

$$\begin{cases} X_{[n]}^{\delta} = F X_{[n-1]}^{\delta} + K\delta_{[n-1]}^{C} \\ Y_{[n]}^{\delta} = C Y_{[n]}^{\delta} \end{cases}$$

with $X_{[n]}^{\delta} = \begin{bmatrix} \delta_{[n]}^R \\ \delta_{[n-1]}^R \\ \delta_{[n-1]}^C \end{bmatrix}$, C=$[1 \quad 0 \quad 0]$, F is a 3×3 matrix and K is a 3×1 matrix.

The matrices F and K depend of the sample time and the response of the low-level to the step function of the steering angle.

4.3 Servoing of the curvature by model predictive control

We will use the following notations. For each instant n we choose a predictive horizon that in our case is an integer n_H corresponding to the total number of iterations needing to be performed in the future (corresponding to low level settling time). We denote by δ^{Obj} the future instruction we want to reach, which is define by

$$\delta_{[n+j]}^{Obj} = \arctan(Lc(s(n+j)) \quad (\text{for each } j \in [\![0, n_H]\!])$$

Where $s(n+j)$ is the predicted curvilinear abscissa deducted from the robot velocity and

$$\delta_{[n+j]}^{Ref} = \delta_{[n+j]}^{Obj} - \gamma^j \left[\delta_{[n]}^{Obj} - \delta_{[n]}^{R}\right].$$

which is the desired behaviour for the steering angle. $\gamma \in [0,1]$ determine the shape of this desired behaviour.

4.4 Design of the predictive the control law

In order to build a predictive control, the future control law at the instant $n+j$ ($0 \leq j \leq n_H$) is defined as a linear combination of basis functions. They are denoted by

$$\left(\delta_{B_k}^{C}\right) \qquad (1 \leq k \leq n_B)$$

In this paper we choose a basis of polynomial function:

$$\delta_{B_k}^{C}(j) = j^{k-1}$$

with the convention $0^0=1$.
The generic control is then a linear combination of this basis function $\delta_j^C = \sum_{k=1}^{n_H} \gamma_k \delta_{B_k}^{C}(j)$. The objective is then to find coefficients γ_k which minimizes the difference between desired behaviour and predicted evolution of steering angle.

4.5 Final control law

The result of the minimization process constitute the predictive term of the control law:

$$\delta_{iTraj}^{pred} = \left[\left[\sum_{j=0}^{n_H} \left(\hat{\delta}_B^R(j)\hat{\delta}_B^R(j)^T\right)\right]^{-1} \left[\sum_{j=0}^{n_H} \left(d(n+j)\hat{\delta}_B^R(j)\right)\right]\right]^T \delta_B^C(0)$$

where

$$d(n+j) = \delta_{[n+j]}^{Ref} - CF^j X_{[n]}^{\gamma}.$$

This control low attached to curvature servoing is computed from the minimization of the quadratic function

$$D(n) = \sum_{j=0}^{n_H} \left\{\mu(n)^T \hat{\delta}_B^R(j) - d(n+j)\right\}^2.$$

Finally the control to be applied is the sum of the reactive term (unchanged) and the predictive curvature servoing such as:

$$\delta_{i=}\delta_{iTraj}^{pred}+\delta_{iDeviation}$$

5 EXPERIMENTAL RESULTS

5.1 Experimental setup

The electric off-road vehicles depicted in **Figure 55** are used as an experimental platform. On this picture the leader is RobuFAST and the follower is named Arocco, they are designed for mobility and they can climb slopes up to 45°.

Table 6: Main parameters of experimental robots

Robots	RobuFAST	Arocco
Total mass	*m*=350 kg	*m*=620 kg
Wheelbase	*L*=1.2 m	*L*=1.2 m
maximum speed	8 m s^{-1}	4 m s^{-1}

The main exteroceptive sensor on-board on the two robots is a RTK-GPS receiver, which supplies absolute position measurement with an accuracy of 2 cm at a 10 Hz sampling frequency. The communications between vehicles are made by WiFi communication.

Figure 55: Experimental platform

5.2 Path tracking results

The experiments for the algorithm's validation consist in following the path depicted on **Figure 56**. This path has been recorded beforehand, when the robot was steered manually at 1 m s^{-1}. It is composed of two straight lines and a turn; half the trajectory is on a sloping ground and the other on a level ground. On **Figure 57** and **Figure 58**, one iteration corresponds to 0.1 s.

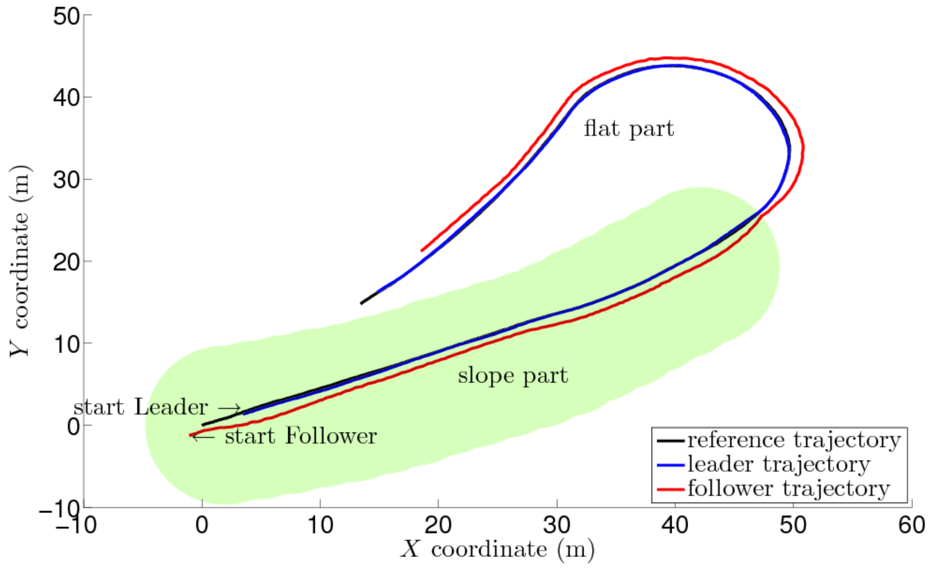


Figure 56: Robots trajectories

The leader moves at 2 m s^{-1}, and has to follow the reference trajectory. The follower has to maintain a lateral distance of 1m and a longitudinal distance of 10 m with the leader. As regards the lateral error on **Figure 57** we can consider the objective is achieved, as it can be seen that after an initializing phase (after iteration 250) the lateral error does not exceed 20 cm with respect to desired deviations: 0m for the leader and 1m for the follower. An overshoot can be observed at iterations 400 and 450 (resp. for the leader and the follower) corresponding to a motion through a bump (slope to flat ground part). This indeed generates a roll motion explaining the variation in lateral error (GPS antennas are placed in the top of robots, see **Figure 55**), which does not correspond to an actual robot motion. Despite this perturbation, the control algorithm stays stable, and provides a level accuracy compatible with actual field operations.

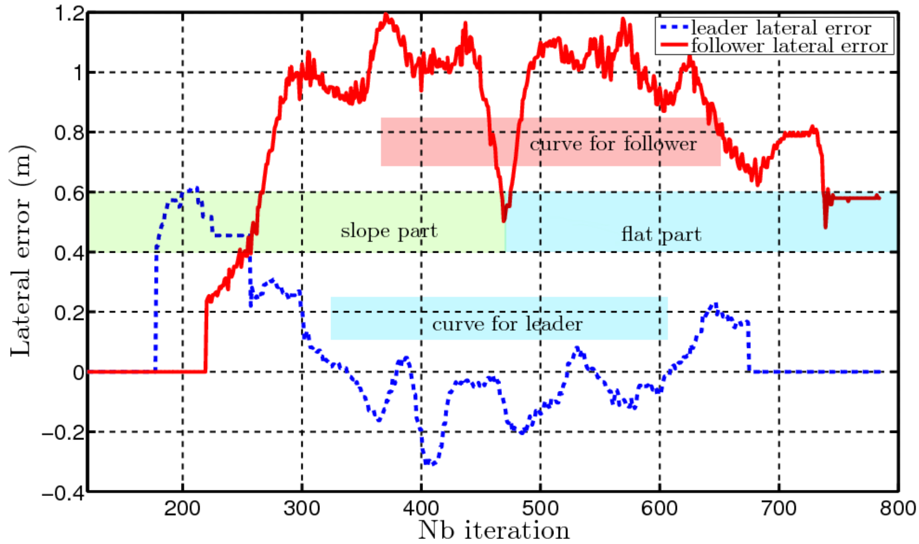


Figure 57: Robots lateral errors

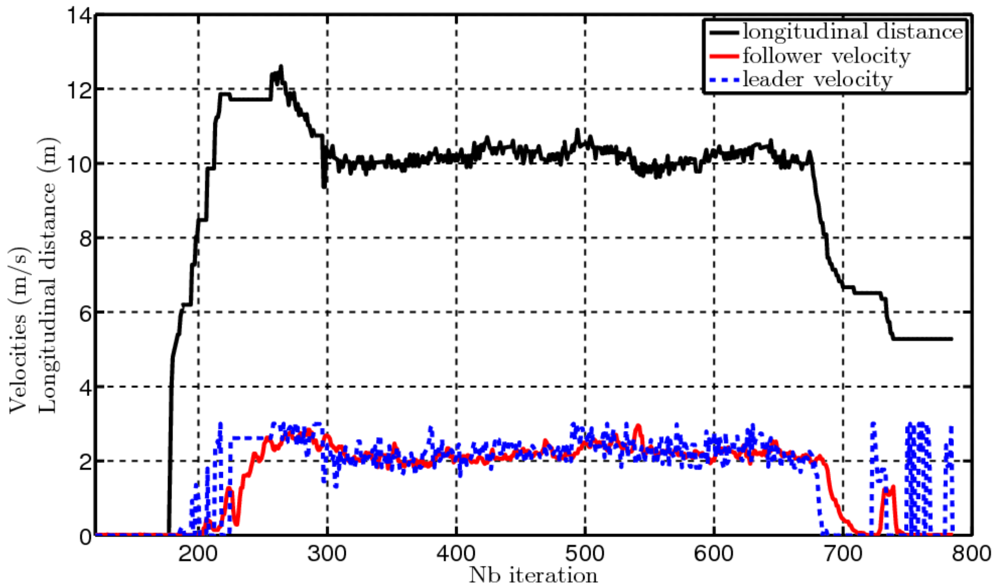


Figure 58: Longitudinal distance and velocity of robots

Figure 58 shows a comparison plot of velocity of robots and longitudinal distance. It can be seen at the start a 2m longitudinal error and at the end another one of more than 3 m. It can be explained by the long time the follower requires to accelerate at the beginning and decelerate at the end. Moreover we note that the longitudinal distance oscillates when robots take the turn and when they reach the flat ground. These inaccuracies occur when fast speed variations are required. Nevertheless, during steady state period, the curvilinear distance between robots is well regulated on the desired value of 10m.

6 CONCLUSION AND FUTURE WORKS

This paper proposes an algorithm for the accurate control of a mobile robot formation moving off-road. This approach considers the formation control as the combination of (i) a platooning control and (ii) an extension of the path tracking problem to a non-null lateral deviation regulation. As a result, the control of each vehicle is decomposed into longitudinal and lateral control with respect to a reference path. An adaptive control strategy is designed. It allows to take into account for low grip conditions, as well as other phenomena encountered off-road and depreciating the accuracy when using classical algorithms. In addition, a predictive curvature servoing has been designed in order to anticipate for overshoots, due to steering actuator settling time. The relative positioning of each robot with respect to a possibly varying formation can then be regulated with a few centimeter accuracy, whatever the shape of the reference trajectory and the grip conditions. The efficiency of the approach has been tested through actual experiments with two off-road mobile robots.

In addition, the proposed strategy is focused on the regulation of a formation with respect to a reference trajectory supplied beforehand. Such an algorithm has now to be extended in order to manage automatically the formation (modification of the formation at the end of the field in order to operate an U-turn, mobile robot entering/leaving the fleet, leader manually controlled, obstacle avoidance, etc). In order to improve longitudinal regulation with respect to follower acceleration performances, a predictive step is under development to anticipate for leader fast speed variation.

REFERENCES

[1] T. Balch and R. Arkin, *Behavior-based formation control for multirobot teams*, IEEE Transactions on Robotics and Automation **14(6)**, 926–939, 1998.

[2] S. Blackmore, B. Stout, M. Wang, and B. Runov, *Robotic agriculture - the future of agricultural mechanisation*, 5th European Conference on Precision Agriculture (ECPA), Upsala (Sweden), 2005.

[3] J. Bom, B. Thuilot, F. Marmoiton, and P. Martinet, *A global control strategy for urban vehicles platooning relying on nonlinear decoupling laws*, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Edmonton (Canada), p. 2875–2880, 2005.

[4] C. Canudas de Wit, G. Bastin, and B. Siciliano, *Theory of robot control*, Springer-Verlag, New York, USA, 1996.

[5] Y. Cao, A. Fukunaga, and A. Kahng, *Cooperative mobile robotics: Antecedents and directions*, Autonomous Robots **4(1)**, 7–27, 1997.

[6] C. Cariou, R. Lenain, B. Thuilot, and M. Berducat, *Automatic guidance of a four-wheel-steering mobile robot for accurate field operations*, Journal of Field Robotics **26(6-7)**, 504–518, 2009.

[7] J. Desai, J. Ostrowski, and V. Kumar, *Controlling formations of multiple mobile robots*, IEEE International Conference on Robotics and Automation (ICRA), Leuven (Belgium), p. 2864–2869, 1998.

[8] J. Fax and R. Murray, *Information flow and cooperative control of vehicle formations*, IEEE Transactions on Automatic Control **49(9)**, 1465–1476, 2004.

[9] R. Lenain, B. Thuilot, C. Cariou, and P. Martinet, *High accuracy path tracking for vehicles in presence of sliding. application to farm vehicle automatic guidance for agricultural tasks*, Autonomous robots **21(1)**, 79–97, 2006.

[10] P. Morin and C. Samson, *Commande de véhicules sur roues non holonomes, une synthèse.*, Actes des troisièmes journées nationales de la recherche en robotique (JNRR), Giens (France), 2001.

[11] E. Sahin, *Swarm robotics: from sources of inspiration to domains of application*, Swarm Robotics, Proceedings of the SAB 2004 International Workshop, Lecture Notes in Computer Science, 2004.

[12] C. Samson, *Control of chained systems. application to path following and time-varying point stabilization of mobile robots*, IEEE Transactions on Automatic Control **40(1)**, 64–77, 1995.

[13] H. Yamaguchi, T. Arai, and G. Beni, *A distributed control scheme for multiple robotic vehicles to make group formations*, Robotics and Autonomous Systems **36(4)**, 125–147, 2001.

[14] R. Lenain, M. Berducat, P. Cartade and B. Thuilot, *Co-ordination of Multiple robots in outdoor environments*, 8th European Conference on Precision Agriculture (ECPA), Prague (Czech Republic), 2011

This Paper has been accepted as a "peer reviewed paper" by the scientific committee

Evolutional Development of Controlling Software for Agricultural Vehicles and Robots

Tsuneo Nakanishi[1], Claes Jæger-Hansen[2], and Hans-Werner Griepentrog[2]
[1]Faculty of Information Science and Electrical Engineering, Kyushu University, Japan
[2]Institute of Agricultural Engineering, University of Hohenheim, Germany

Abstract
Agricultural vehicles and robots expand their controlling software in size and complexity for their increasing functions. Due to repeated, *ad hoc* addition and modification, software gets structurally corrupted and becomes low performing, resource consuming and unreliable. This paper presents an evolutional development process combining *Software Product Line* (*SPL*) and *eXtreme Derivation Development Process* (*XDDP*). While SPL is a promising paradigm for successful reuse of software artefacts, it requires understanding of the whole system, a global and future view of the system, and preparation of well managed core assets. By contrast, while XDDP is a less burden process which focuses only on the portion to be changed in the new system, it never prevents software structure from corrupting due to absence of the global view of the system. The paper describes an adoption process for SPL, with an example of the autonomous tractor, that applies XDDP initially for addition and modification of functions, accumulates core assets and cultivates a global view of the system through iterated development with XDDP, and finally shifts to SPL development.

Keywords
Software Evolution, Software Development Process, Software Product Line, Derivative Development

1 INTRODUCTION

The agricultural vehicle has been getting to provide more operator friendly services. Its evolution toward the unmanned vehicle is a definite trend and its final goal should be the autonomous robot with intelligence. In this forecasted evolution, embedded software controlling agricultural vehicles and robots will play an important role more than before. Most of intelligent and attractive functions to automate agricultural tasks are implemented mainly by software. These functions must analyse data from various sources including on-board sensors, GPS, other vehicles or robots, base stations, databases *etc*.; make sophisticated decisions; and drive multiple mechanical devices such as engine, brake, various implements, *etc*. in a coordinated manner through networks such as CAN, ISOBUS *etc*. Software implementing these functions grows quite easily in size and complexity. In fact, for the decade and more, automotive industry has experienced steep increase in size and complexity of software brought by integrated functions. According to Broy, 2006, in the general passenger vehicle, more than 2,000 functions were controlled by software; the size of the source code was over ten million lines; and 50 to 70% of development cost was dedicated for software. It is almost impossible to construct correctly working software of such large scale by code centric development without well-defined sound process.

Besides size and complexity, variability can be a big issue in agricultural vehicles and robots. Agricultural vehicles and robots perform different tasks for different crops under different geographical, climatic, and economic environments. They employ different technologies, namely hardware and mechanical devices, and the technologies themselves will evolve. These diversities in agriculture and technology finally result in a huge amount of variability in software.

Basically, repeated additions and modifications are applied to the existing software in evolution of software. As *ad hoc* additions and modifications are repeated, software is structurally corrupted and become low performing, resource consuming and unreliable.

This paper discusses evolutional development of software. The authors propose to introduce *software product line* (*SPL*) [Clements & Northrop, 2001; Pohl *et al.*, 2005], a paradigm of software reuse for different products, for steady evolution without corruption of software structure. However, it is often difficult to adopt SPL without preparation even for development sites having concrete development processes for single product development. Moreover, SPL requires a global view of the current and future agricultural vehicles and robots, which is difficult to foresee for a long term. Therefore, the authors also propose to perform some iterations of the *extreme derivative development process* (or *XDDP* for short) [AFFORD] until SPL development gets applicable.

This paper is organized as follows: Section 2 gives some fundamental concepts on SPL. Section 3 describes XDDP in comparison with SPL. Section 4 shows our ideas on evolutional development starting from XDDP and shifting toward SPL. Finally, Section 5 concludes the paper.

2 PARADIGM OF SOFTWARE PRODUCT LINE

SPL development enables production of various software systems with different functionality and quality, namely software product line, in a strategic and planned manner by optimally constructing and reusing core assets shared among the systems.

SPL is absolutely not a development method to produce different products by using libraries, which store codes reusable in other products, in an *ad hoc*, code centric, individual skill dependent manner. SPL development is driven by business and technical plans of the product line. Essential plans are the scope and the road map of the product line that define which products are in and out of the product line at a certain time. Artefacts steadily reused in the products in the plans are constructed and maintained as the core assets of the product line. The core assets include not only codes but also artefacts in upper sub-processes such as requirements, specifications and designs. They are reused to construct each product by a prescribed manner, not by an individual manner.

SPL is a paradigm of software reuse among different products, rather than a certain software development methodology. A lot of methodologies and case studies based on the SPL paradigm have been presented and reported by academic researchers and industrial practitioners for the last ten to fifteen years. Some fundamental concepts in the SPL paradigm, which are described below, are introduced in these works:

Separation of domain engineering and application engineering: Domain engineering is a set of activities to construct and maintain core assets for the whole product line. Application engineering is a set of activities to develop each product by reusing core assets. These are clearly distinguished in SPL development. Moreover, management to coordinate domain engineering and application engineering is also essential. **Figure 59** shows an instance of the SPL development process.

Separation of commonality and variability: Commonality and variability among products are analysed in SPL development. Commonality and variability are often described in terms of *features*, which can be defined as any prominent and distinctive concepts or characteristics that are visible to various stakeholders of the system [Kang *et al.*, 1990; Lee *et al.*, 2002]. Analysed features are categorized in terms of constraint of its selection in each product and organized as a *feature model* [Kang *et al.*, 1990]. Each product is distinguished by its equipping features.

Figure 60 shows an illustrative example feature model of an imaginary autonomous tractor product line. Each node of the feature model represents a feature. A node without any decoration represents a *mandatory* feature, which should be equipped by all the products. A node with circular decoration represents an *optional* feature, which may or may not be equipped by each product. A set of nodes bundled by an arc represents *alternative* features such that one of them is alternatively equipped by each product. Regardless of its category on selection constraint, the feature is not equipped by a product if its parent feature is not equipped in the product. The edge between nodes represents semantic relationship between corresponding features. *Consists-of* relationship means that the parent feature consists of the child feature; that is, the child feature forms a part of the parent feature. *Generalization* relationship means that the parent feature is a generalized concept of the child feature. *Implemented-by* relationship means that the parent feature is realized by the child feature. Consistent selection of features on the feature model specifies a product.

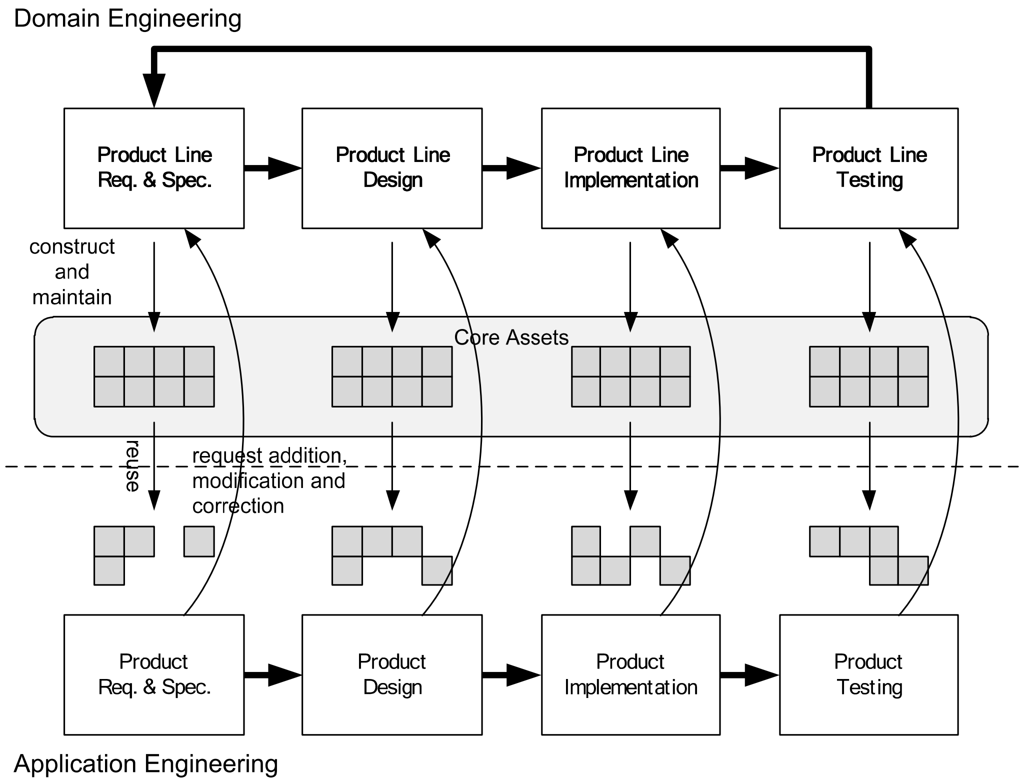


Figure 59: SPL development process

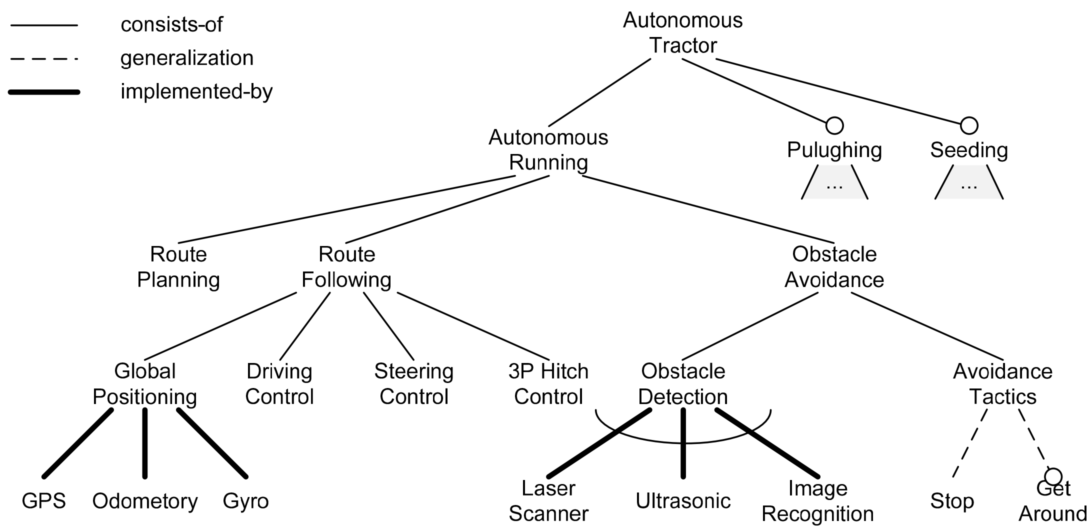


Figure 60: Feature Model

Architecture centric development: In SPL development, software architecture is established with considering the results of commonality and variability analysis to enable comprehensive and disciplined product derivation in application engineering. Core assets are also constructed to be applicable to the architecture. Application engineering is allowed to reuse core asset components at predefined points, referred to as *variation points*, in the software architecture in a prescribed manner.

Separation of problem and solution spaces: The problem space is a space storing variability models of the product line. The feature model is a representative artefact in this space. On the other hand, the solution space is a space storing other artefacts across various abstraction levels including requirements, specifications, designs, implementations, and testing. Traceability from the problem space to the solution space is somehow kept in SPL development. For example, requirement, specification, design, implementation and testing artefacts in the solution space are tagged by a feature at the portions in where the feature is realized. This traceability eases product derivation based on feature selection in application engineering.

Inherently, SPL is a development paradigm which requires a global and future view of the product line and definition of software architecture comprehending the whole product line. Most of development sites already have some working systems before introducing SPL. It is essential to understand the whole system to define the software architecture of the product line. However, that makes introduction of SPL prohibitive for large and complicated systems due to excessively growing cost of domain engineering, time limitation, human resource limitation, lack of reliable documents *etc*. Moreover, it is often too difficult to foresee future evolution of the product line for innovative products such as agricultural robots.

Another adoption problem of SPL is the maturity level of the development site. At least, it is hopeless for development sites continuing code centric development to introduce SPL successfully. It requires a sound development process and documents in enough quality and quantity to perform domain engineering.

To alleviate these adoption barriers, the authors propose to introduce a derivative development process and then shift to SPL development.

3 XDDP: A DERIVATIVE DEVELOPMENT PROCESS

XDDP, which stands for *eXtreme Derivative Development Process* [AFFORD], is a derivative development process introduced to development sites in Japanese industries [Kobata, 2010]. XDDP is a development method to produce new products by adding and modifying an existing product. XDDP can be used as a development process to produce different products with commonality and variability likewise for SPL. However, XDDP is established independently from SPL and, in fact, it does not have fundamental concepts of SPL described in Section 2. For example, XDDP does not have the concept of the core asset. XDDP modifies the base product to construct a new product, instead of combining core assets. **Figure 61** shows the overview of XDDP. Each circular node in the figure represents a sub process of XDDP. Due to page limitation, the details of each sub process are omitted.

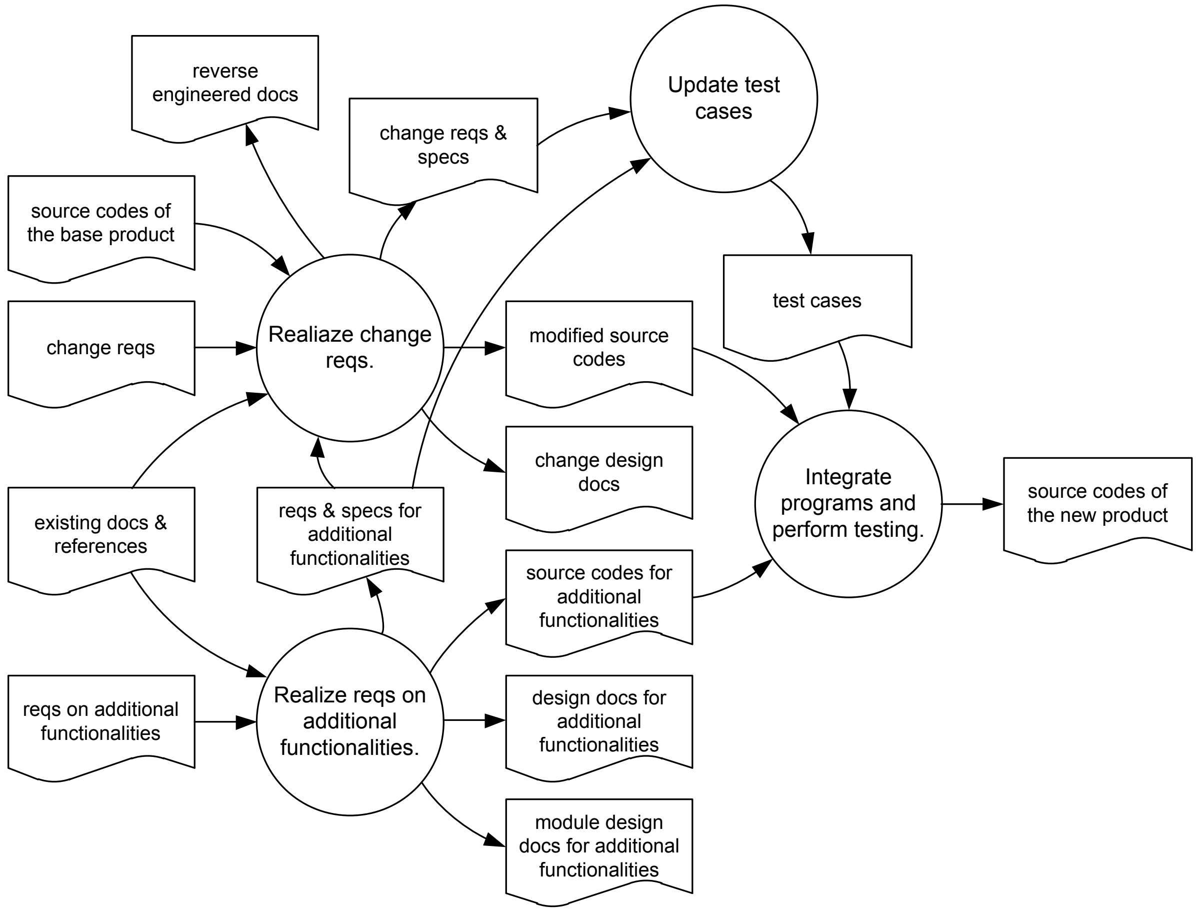


Figure 61: Overview of XDDP

XDDP starts development of a new product from describing change requirements and specifications to the base product as well as requirements and specifications for additional functions. Change requirements and specifications are different from commonly known requirements and specifications in that they regard desired changes, not desired functionality and quality, for the new product as requirements and specifications. Change requirements are description on what the stakeholder of the new product wants to be changed from the base product. Change specifications are description specifying, namely describing without any ambiguity, how the base product should be modified to satisfy the change requirements. On the other hand, requirements and specifications for additional functions are same as commonly known requirements and specifications except they are only for additional functionalities. Note that change requirements and specifications can also include all the modifications to the base product caused by additional functions.

Table 7 shows an example of change requirements and specifications. The example is error correction in longitude and latitude observed by GPS due to tilt of the tractor. (See [Eriksen & Jæger-Hansen, 2010] for the details.) Change requirements and specifications are described with keeping their correspondence. A change requirement is followed by the change specifications satisfying it. Moreover, the change requirement is annotated by its reason to make the context of the requirement explicit.

Table 7: Change requirements and specifications

TiltComp	Change Req	Want to correct the error in longitude and latitude observed by GPS due to tilt of the tractor.
	Reason	Errors non-negligible for precise agricultural tasks are produced depending on tilt of the tractor, because the GPS antenna is attached at a distant and higher position from the reference point in the tractor.

TiltComp.1	Change Spec	Add a task to interface the inclination sensor, get the roll and pitch angles of the tractor and apply LPFs to the observed roll and pitch angles for noise reduction.
TiltComp.2	Change Spec	Let the Kalman filter in the global positioning task, which is used for better estimation of the position, use compensated longitude and latitude for its input, instead of raw longitude and latitude from GPS. Let h_{ant}be the height of the antenna, θ_{Kalman} the angle of the tractor coordinate system to the global coordinate system estimated by the Kalman filter, and θ_{roll} and θ_{pitch} the filtered roll and pitch angles respectively. The errors to real latitude and longitude due to tilt of the tractor, denoted by $\varepsilon_{\text{long}}$ and ε_{lat} respectively, are expressed as follows: $$\varepsilon_{\text{long}} = h\cos(\theta_{\text{Kalman}} + \theta_{\text{P-R}})$$ $$\varepsilon_{\text{lat}} = h\ \sin(\theta_{\text{Kalman}} + \theta_{\text{P-R}})$$ where $h = h_{\text{ant}}\sqrt{\sin^2\theta_{\text{pitch}} + \sin^2\theta_{\text{roll}}}$and $\theta_{\text{P-R}} = \tan^{-1}\frac{\sin\theta_{\text{roll}}}{\sin\theta_{\text{pitch}}}$. These errors are added to longitude and latitude observed by GPS for compensation.

Change design documents describes necessary modification to the existing design to satisfy the change specifications, namely how the modules of the base product should be modified. The traceability matrix makes the change design documents traceable from the change requirements and specifications. **Table 8** is an example of the traceability matrix after module design for the change is finished. The traceability matrix shows which module corresponding to a column should be modified to realize each change requirement or specification corresponding to a row by the check mark "X".

It will be necessary to engineer the current implementation reversely to describe change specifications and the traceability matrix, if the documents on the current implementation are not available. The results of reverse engineering are documented and referenced to describe change designs.

In the final stage, existing source codes of the base product are modified with referencing traceability matrix and change design documents to satisfy change requirements and specifications. *Ad hoc* modification of existing codes often causes newly introduced bugs. Objective of this *lazy* and planned code modification strategy is to avoid unnecessary waste of time for repeated correction due to the bugs newly introduced by modification.

Table 8: Traceability matrix

Req/Spec ID	GetGlobalPos	GetOdometry	GetTilt [New]	GetImprovedGlobalPos	CalcSteeringAngle.	DriveSteering	…
TiltComp			X	X			
TiltComp.1			X	X			
TiltComp.2				X			

While SPL is plan driven as described in Section 2, XDDP is basically change driven. XDDP focuses only on changes to the base product. Documents are produced only for the changes. It can safely state that, although XDDP has less adoption barriers than SPL, XDDP will not prevent software structure from corrupting if it is repeatedly applied without any global and future view of the product line.

4 EVOLUTIONAL DEVELOPMENT TOWARD SPL

XDDP assumes existence of neither core assets nor software architecture. It accepts the current software architecture of the base product and modifies only the portions of the base product to be changed for the new product. Naïve derivation of new products by XDDP does not accumulate core assets, recover the software architecture, and bring reform to SPL development. For steady and affordable shift toward SPL development, the authors tailor XDDP to facilitate mining of core assets from existing artefacts and cultivate the global view through its iterations. The tailored XDDP, which is named as XDDP4SPL here, follows the process described below.

Describing requirements and specifications before and after changes: The original XDDP describes change requirements and specifications for derivation of a new product. XDDP4SPL additionally describes requirements and specifications before and after changes in separate. The *before*-requirements and specifications are imported from existing ones of the base product, or engineered reversely from the source codes of the base product if no document on requirements and specifications is available. The *after*- requirements and specifications are newly described based on change requirements and specifications.

Although readers may think that change requirements and specifications are no longer necessary, they should be kept with *before*- and *after*- requirements and specifications for some reasons.

One reason is that desire for succeeding products is often described as changes to the proceeding products at first. Moreover, the changes are described in various abstraction levels by various stakeholders of the product line. Some changes may be described by users at the abstraction level of requirement as additional or improved functions. Other changes may be described by engineers at the abstraction level of specification without explaining why the changes are needed.

Another reason is that change requirements and specifications describe why one function is newly introduced and record evolution of the product line. These documents are helpful for engineers newly involved in the project to understand the product line better than each function is explained solely.
Before- and *after*- requirements and specifications should be traceable from their corresponding change requirements and specifications. **Table 9** shows a possible description. *Before*- and *after*-requirements and specifications are traceable by requirement and specification IDs in this description.

Table 9: *before*- and *after*- requirements and specifications

TiltComp	Before Req	Want to know the current position of the tractor **<u>without</u>** taking account of tilt of the tractor.
TiltComp	After Req	Want to know the current position of the tractor **<u>with</u>** taking account of tilt of the tractor.
TiltComp.1	Before Spec	
TiltComp.1	After Spec	Get the roll and pitch angles of the tractor periodically and apply LPFs to the observed roll and pitch angles for noise reduction.
TiltComp.2	Before Spec	The Kalman filter in the global positioning task uses **<u>raw longitude and latitude</u>** from GPS for its input.
TiltComp.2	After Spec	The Kalman filter in the global positioning task use **<u>compensated longitude and latitude</u>** for its input. (See **Table 8** for the details of the compensation.)

Performing local variability modelling: XDDP4SPL performs local variability modelling for the limited portion of the system to be changed for a new product. With comparing *before*- and *after*-requirements and specifications, it becomes easier to identify common and different aspects such as structures, behaviours, and properties among products and define features. Different description between *before*- and *after*- requirements and specifications, which are in bold and underlined texts in **Table 9**, is a basis to identify features. Features indirectly related to the changes do not appear in *before*- and *after*- requirements and specifications. Instead, they may found in reverse engineered documents. Variability possibly introduced in future should be identified during local variability modelling.

The primary object of this partial feature modelling is better separation of variability, which will bring better modularization and interface design for reuse among products. Features should be identified such that commonality and variability are cleanly separated. A common feature must not include variable aspects and *vice versa*. Moreover, variable features should be orthogonally separated. A variable feature should not include multiple aspects which are in different concepts or abstraction levels. Guidelines on feature modelling [Lee, 2002] are also helpful for good feature modelling. Other objectives of local variability modelling are better understanding and intuitive representation of the portion directly and indirectly related to the changes.

Figure 62 shows an example of the local feature model. The features identified from *before*- and *after*-requirements and specifications are *Tilt Compensation* and *Inclination Sensing*. *Global Positioning* is identified as the parent feature of *Tilt Compensation* in *consists-of* relationship, since *Tilt Compensation* is for *Global Positioning*. *Kalman Filtering* is a feature identified in reverse engineered documents. It is also a sub feature of *Global Positioning* in *consists-of* relationship. Both *Kalman Filtering* and *Tilt Compensation* are modelled as optional features to enable core assets to be reused for tractors without gyro, odometer, and inclination sensors.

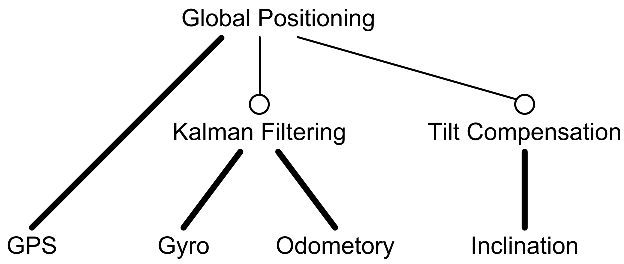


Figure 62: Local Feature Model

Describing partial requirements and specifications for the product line: Based on the *before*- and *after*- requirements and specifications, the local feature model, and reverse engineered documents, partial requirements and specifications for the product line are described. Local variability modelling and description of partial requirements and specifications can be performed iteratively. The partial requirements and specifications become core assets of the product line.

Table 10 shows an example of partial requirements and specifications. The conditional expression written in the brackets ([]) is a guard expression representing a feature selection such that the specification is activated. The term in the expression becomes true if and only if the feature of the same name is selected in the product.

Table 10: Partial requirements and specifications

GlobalPos	Req	Want to know the current position of the reference point.
GlobalPos.1	Spec	Get the current position, namely longitude and latitude, of the tractor from the GPS receiver.
GlobalPos.2	Spec	Get the roll and pitch angles of the tractor from the inclination sensor and apply LPFs to the observed angles for noise reduction. [Inclination]
GlobalPos.3	Spec	Compute errors in longitude and latitude observed by GPS to the reference point due to tilt of the tractor (See **Table 8** for the details of the compensation.) and add them to the current position of the tractor from the GPS receiver for compensation. [Tilt Compensation]
GlobalPos.4	Spec	Get the direction of the tractor from the gyro sensor and apply LPFs to the observed direction for noise reduction. [Gyro]
GlobalPos.5	Spec	Get the odometory data of the tractor from the odometory sensor and apply LPFs to the observed data for noise reduction. [Odometory]

GlobalPos.6a	Spec	Input the raw or tilt compensated current position, the direction, and the odometory data of the tractor into the Kalman filter for better estimation and output the results as the current position of the reference point. [Kalman Filtering]
GlobalPos.6b	Spec	Output the results as the current position of the reference point. [!Kalman Filtering]

Performing additional design and implementation and refactoring existing artifacts: Design and implementation for partial requirements and specifications should be performed. Design and implementation for additional functionalities are constructed newly because there is no asset for them. Existing design artefacts relating to the changes are refactored if they are available, or engineered reversely from the codes otherwise. Existing codes relating to the changes are also refactored. It is essential to introduce variation mechanism, which enables product derivation by combination of core assets depending on feature selection, such as parameters, conditional compilation, common interface, inheritance, *etc*. [Anastasopoulos & Gacek, 2001; Gomaa & Webber, 2004]

Figure 63 shows the overview of XDDP4SPL. Iteration of XDDP4SPL, which is driven by changes, accumulates core assets including partial feature models, partial requirements and specifications, refactored design artefacts and codes, and reverse engineered documents. These locally mined or produced core assets should be sooner or later integrated in a global framework of the product line.
To guide this integration of core assets and facilitate shift toward SPL development, the authors present a status model of the feature. The status model defines *visible* and *invisible* features. The visible feature is one that the modeller has recognized (and thus the visible feature can be modelled in the feature model). Invisible features, which the modeller has not recognized yet, are concealed in the explored portion of the existing system. The invisible feature becomes a visible feature, when it is exposed by reverse engineering work, expert knowledge *etc*.

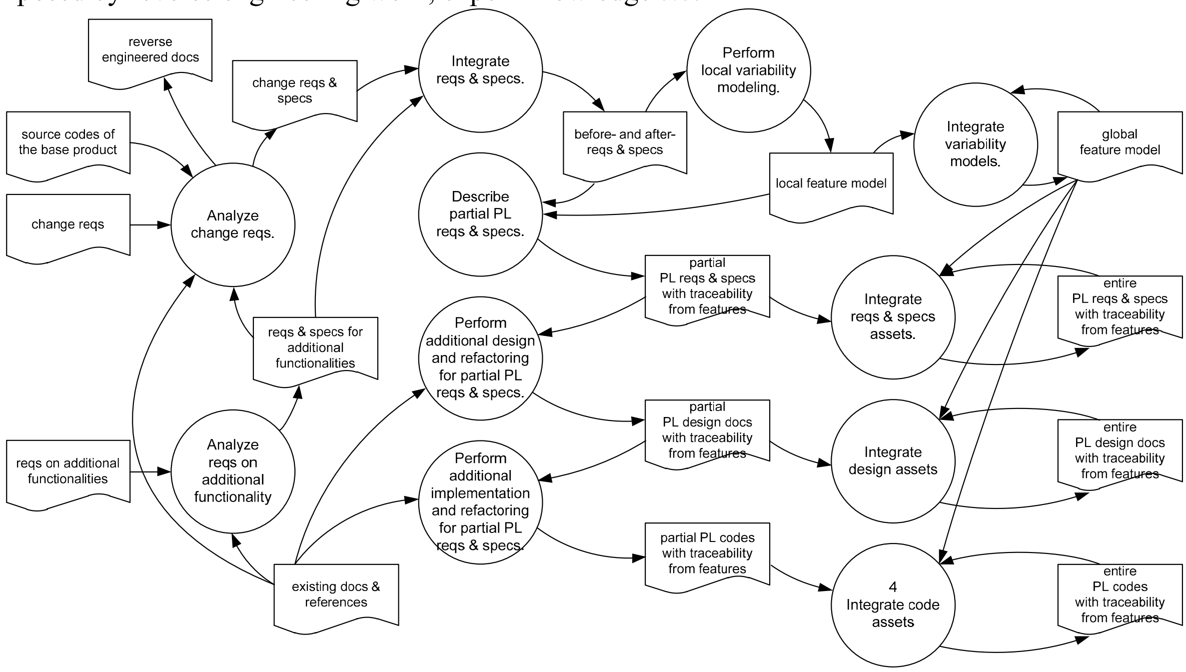


Figure 63: Overview of XDDP4SPL

Moreover, for the visible feature, the status model has two dimensions: scope of feature identification and traceability to core assets. In terms of scope of feature identification, the feature is categorized into *locally identified feature* or *globally identified feature*. The locally identified feature is one identified in a limited scope of the product line. The globally identified feature is one identified in the full scope of the product line. In terms of traceability to core assets, the feature is categorized into *core asset traceable feature* or *core asset untraceable feature*. The core asset traceable feature is one that is traceable to its related core assets. The core asset untraceable feature is one that is not traceable to its related core assets. Thus, each visible feature has four kinds of status in our model.

The feature identified in XDDP4SPL becomes a *locally identified & core asset traceable* feature, since its related artefacts are incorporated in core assets in a traceable manner. The *locally identified & core asset traceable* feature promotes to a *globally identified & core asset traceable* feature, when its position in the global feature model is determined with understanding of the whole product line.

The feature identified in the top-down manner based on expert knowledge initially becomes a *globally or locally identified & core asset untraceable* feature depending on the scope of the expert knowledge, since its related artefacts are not clear at all. The *globally or locally identified & core asset untraceable* feature promotes to a *globally or locally identified & core asset traceable* feature, when its related artefacts are incorporated in core assets in a traceable manner after XDDP4SPL iteration or reverse engineering work.

This categorization of the feature is used to separate the portion to where SPL development is applied and the portion to where derivative development is applied in the system. Shift to perfect SPL development is achieved when i) all the visible features are *globally identified & core asset traceable* and ii) invisible features are believed to be wiped out.

5 CONCLUSION

This paper presented an evolutional development process combining SPL and XDDP, which the authors referred to as XDDP4SPL. The process is as follows: i) Change requirements and specifications are described with focusing on changes to the base system as XDDP does. ii) Based on the change requirements and specifications, *before*- and *after*- requirements and specifications are described to make commonality and variability between the base system and the new system. iii) Local variability modelling, which constructs a local feature model, is performed for better separation of variability and better understanding and intuitive representation of the portion related to the changes. iv) Partial requirements and specifications, which are incorporated in core assets, are described with establishing traceability from features. v) Existing design artefacts and codes are refactored with introducing variability mechanisms to enable product derivation by combination of core assets.

Core assets accumulated through some iterations of XDDP cultivates global view of the system and enables shift to SPL development. To guide integration of core assets and facilitate shift toward SPL development, the paper presented a status model of the feature in the existing system. The feature is categorized into invisible and visible features. The visible features is categorized into four classes, namely { *locally-identified*, *globally-identified* } × { *core asset traceable*, *core asset untraceable* }.

The future works include application and evaluation of XDDP4SPL to develop additional functions for the autonomous tractor.

ACKNOWLEDGMENT

This work was supported by Grant-in-Aid for Young Scientist (B), KAKENHI (No. 21700035).

REFERENCES

Books:

CLEMENTS, P. and NORTHROP, L.: *Software Product Lines: Practice and Patterns*, Addison-Wesley, 2001.

POHL, K., BÖCKLE G., and LINDEN, F. v. d.: *Software Product Line Engineering: Foundations, Principles and Techniques*, Springer, 2005.

Conference papers:
ANASTASOPOULOS, M. and GACEK, C.: *Implementing Product Line Variabilities*, Proc. Symp. on Software Reusability (SSR) '01, p.109–117, 2001.

BROY, M.: *Challenges in Automotive Software Engineering*, Proc. 28th Int. Conf. on Software Engineering, p.33–42, 2006.

GOMAA, H. and WEBBER, D. L.: *Modeling Adaptive and Evolvable Software Product Lines Using the Variation Point Model*, Proc. 37th Hawaii Int. Conf. on System Sciences, 2004.

LEE, K., KANG, K., and LEE, J.: *Concepts and Guidelines of Feature Modeling for Product Line Software Engineering*, Proc. 7th Int. Conf. on Software Reuse, p.62–77, 2002.

KOBATA, K., NAKAI, E., and TSUDA, T.: *Process Improvement Using XDDP: Application of XDDP to the Car Navigation System*, Proc. 5^{th} World Congress for Software Quality, Nov. 2011.

Technical reports:
KANG, K., COHEN, S. G., HESS, J. A., NOVAK, W. E., and PETERSON, A. S.: *Feature-Oriented Domain Analysis (FODA): Feasibility Study*, Software Engineering Institute, Carnegie Mellon University, CMU/SEI-90-TR-222, Nov. 1990.

Thesis:
ERIKSEN, J. and JÄGER-HANSEN, C.: *GPS-Styring af Malerobot* (*GPS Control of Painting Robot*, in English), Department of Electrical Engineering, Technical University of Denmark, Aug. 2010. (in Danish)

Links:
AFFORDD: Association for Facilitation of Rational Derivational Development, http://www.xddp.jp, last accessed on October 10, 2011. (in Japanese)

This Paper has been accepted as a "peer reviewed paper" by the scientific committee

EXPERT: A Driver Assistance System for Fuel Efficient Driving

Tianyi Guan, Christian W. Frey
Department for Measurement, Control and Diagnostics (MRD),
Fraunhofer Institute for Optronics, System Engineering and Image Exploitation (IOSB),
Germany

Abstract

With ever rising resource prices, simultaneous expansion of trade and transportation as well as increased awareness of environmental protection, it is important to reduce the fuel consumption of trucks, transporters, agricultural vehicles and other heavy-duty vehicles. While new generations of vehicles like electric cars are still in in their infancy, it is already possible to reduce fuel consumption through fuel economic driving behavior. The EU funded project "EXPERT" (**EXP**ert System for a more **E**fficient **R**oad **T**ransportation), a consortium of European partners including Fraunhofer IOSB, is investigating ways to support the truck driver to accomplish this goal. The main focus of this paper is on the "Driving Efficiency Module", created by Fraunhofer IOSB, within EXPERT system which uses an adaptive vehicle model to generate fuel efficient online guidelines for the driver. An inverse vehicle model and a partial power train model formulation are presented and compared with each other. System identification methods, solely using CAN-Bus data, are proposed to estimate the unknown model parameters that will enable the model to adapt to a wide range of different vehicles. The adaptive model is then used in an optimization routine that generates fuel efficient guidelines depending on the current vehicle and the current state. Aided by the fuel efficient guidelines, the driver will be able to adopt a more fuel efficient driving behavior. This paper concludes with an evaluation of the proposed ideas based on real world data and simulated data.

Keywords
Driver assistance system, fuel economic driving, road transportation, system identification

1 INTRODUCTION

Several authors have published solutions in the area of fuel efficient driving. The approaches range from complete vehicle control to passive driver assistance systems. Active approaches include the works of [6-12]. In [7] Hellström proposed a model predictive approach, in which a predefined vehicle model is used in combination with height maps. The dynamic programming optimization routine computes a fuel efficient velocity and gear change profile for the upcoming road based on height maps. In [10] Terwen also used a model predictive approach. In addition to height maps he used a radar sensor to detect other traffic participants in front of the own vehicle which are also incorporated into the optimization. His optimization routine is based on an efficient sequential quadratic programming method to calculate an optimal velocity profile for the upcoming road. The gear change optimization is done in an external search routine due to its non-linear nature. An example for passive driver assistance systems is the proposal of Filev and Syed [13]. In contrast to the model predictive approaches, they designed a Fuzzy controller that does not make use of models and only concentrates on the present situation. Their optimization is based on a maximum acceleration pedal position that the driver should not exceed. If he ignores the advice, the system will issue a warning. Depending on the driver's rejection rate, the system is able to adapt to the driver's behavior. A complementary publication by Casavola et al. [14] did not deal with the acceleration pedal position but rather with optimal gear selection. They used an efficiency map and alternatively a Fuzzy controller based approach to find the currently optimal gear. More rudimentary assistance systems can be found among Smartphone apps [16][17] that use the acceleration sensor of the smartphone to notify the driver of abrupt acceleration and braking. Different to model dependent and therefore often manufacturer specific approaches [9-11], EXPERT will be applicable to a great variety of vehicles due to an internal online multivariable adaptive vehicle model, which at the same time takes more vehicle specific characteristics into account than knowledge based approaches [13] or Smartphone app approaches [16-17]. The main objectives of the EXPERT project [18] are the development of algorithms and components for road transportation in order to reduce fuel consumption. Therefore an on-board assistance system is developed. The computations are done in the on-board EXPERT components within every individual vehicle. The EXPERT system consists of different sub-systems, of which Fraunhofer has designed the driving efficiency evaluation and driving guideline generation. These functionalities constitute the “Driving Efficiency Module”. The “Driving Efficiency Module” uses an adaptive power train model to generate fuel efficient online guidelines for the driver. In this paper, an inverse vehicle model and a partial power train model formulation are presented. System identification methods are proposed to estimate the unknown model parameters that will enable the model to adapt to a wide range of different vehicles. The adaptive model is then used in an optimization routine that generates fuel efficient guidelines depending on the current vehicle and the current state (e.g. current velocity, current engine speed). Because no height maps, slope information, object detection, traffic light information, speed limit information or route information is available, the system identification and the fuel efficiency guideline generation is solely based on CAN-Bus data. Fuel efficient route selection is not the topic of this paper, but the assistance of the driver during actual driving.

2 THE EXPERT SYSTEM

Before the “Driving Efficiency Module” is explained in further details, an overview of the entire EXPERT system is given. The EXPERT system is primarily designed for freight forwarding companies, which want to keep track of their fleet and increase the cost efficiency or rather the fuel efficiency of their drivers. An overview of the system is given in **Figure 64**. From the figure the reader can see that EXERT is a collaboration of different components. The “EXPERT Information System” is a fleet management system, which resides at the management headquarters and communicates with vehicles within the fleet via the “Green Box”. The “Green Box” is the main communication hub of every EXPERT assisted vehicle that enables the communication of all other EXPERT hardware components. It is installed into every vehicle’s cabin. A HMI device (e.g. a Tablet computer) is fixed to the dashboard of each vehicle. It contains the “Driving Efficiency Module” software designed by Fraunhofer IOSB. The module receives information from other EXPERT components via the “EXPERT Co-Pilot” HMI software, which is also integrated into the HMI device. The “Driving Efficiency Module” consists of an adaptive vehicle model that serves as state estimation for its counterpart in the real world and an optimization sub-module, which generates fuel efficient driving guidelines for the driver. During operation, the driver retains control of the vehicle at all times.

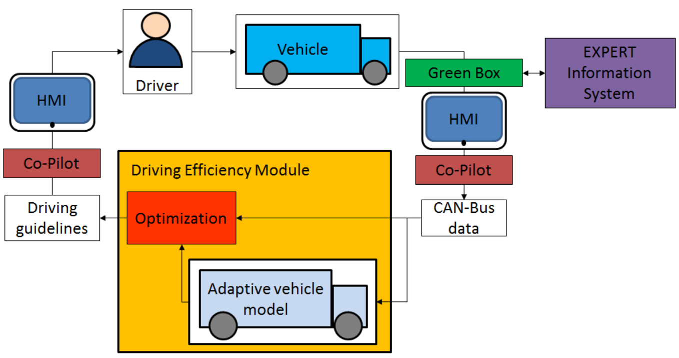


Figure 64: EXPERT system

3 DRIVING EFFICIENCY MODULE

In this chapter the “Driving Efficiency Module” will be explained in further detail. Different adaptive models and fuel efficiency guideline generation strategies will be presented. At the end of this chapter the authors will present the best trade-off with regard to the technical specifications of the EXPERT system.

3.1 Adaptive model

Due to the complex interaction of internal and external propelling and resistant forces in the different vehicle components, it is necessary to simplify the vehicle and the environment to obtain a unified model with a small amount of unknown parameters that can be adapted and used for a great variety of different vehicles. The authors have investigated two different approaches with the aim to create a vehicle model that is sufficiently accurate for optimization purposes, but at the same time enables parameter adaptation of real world vehicles with limited sensor information. The two models share several simplifications. First of all, it is assumed that the vehicle and the environment are quasi stationary, i.e. there are no state changes within a short period of time. The state can only evolve from one time instant to the next. Furthermore vehicle internal slip and vehicle internal friction of any kind are neglected. In order to further simplify the vehicle model, a gray box approach is used. This means in our case that less complex model components are stated as equations known from fundamental physics, while the complex components of the engine are described by static characteristic maps. All model components contain specific unknown model parameters that need to be estimated. The model parameter adaptation must be conducted in a way such that the resulting model output is in compliance with the corresponding CAN-Bus measurements. In order to simplify parameter estimation, it is beneficial to analyze sub-systems whenever possible.

3.1.1 Inverse power train model

The first approach is an *inverse power train model*, which has been inspired by Guzzella et al. [1]. The common chain of effects, i.e. the effect of acceleration pedal change on the overall kinematic behavior of the vehicle, is reverted. The input variables are vehicle velocity v, gear choice G and external resistance forces, which are propagated through the power train and result in a certain fuel consumption that is needed to overcome the resistance forces and realize the proposed velocity. The model is depicted in **Figure 65**.

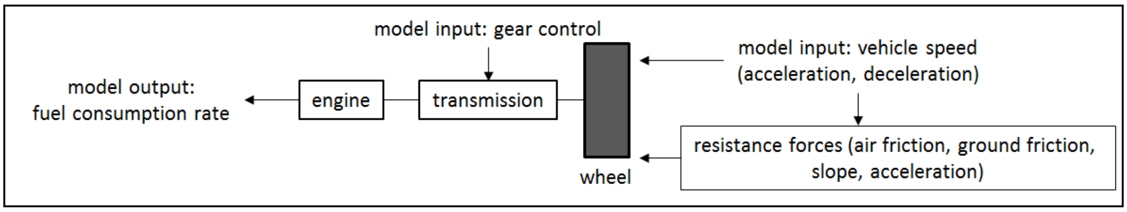


Figure 65: Inverse power train model formulation

<u>Resistance forces</u>

Only the main *longitudinal resistance forces* on the vehicle are regarded (3.1-1 to 3.1-4). They are defined as the gradient force F_{slope} when driving on an uphill slope, the air friction resistance F_{air}, the ground friction resistance F_{ground} and the additional resistance during acceleration F_a. m is the vehicle's mass, μ_{air} is the combined air friction coefficient, μ_{ground} is the ground friction coefficient, v is velocity and φ is the slope angle.

$$F_{air} = \mu_{air} v^2 \tag{3.1-1}$$

$$F_{ground} = \mu_{ground} mg\cos(\varphi) \tag{3.1-2}$$

$$F_{slope} = mg\sin(\varphi) \tag{3.1-3}$$

$$F_a = m\dot{v} \tag{3.1-4}$$

<u>Transmission</u>

The transmission model is described by the static *transmission ratio* $i_t(G)$. It enables the estimation of the current engine speed depending on the current vehicle velocity and currently selected gear G. Internal inertia and clutch dynamics are neglected. Thus, measurement samples during gear shifts are not used for estimation. The transmission ratio is the ratio of the input transmission rotation speed $\omega_{t,in}$ and the output transmission rotation speed $\omega_{t,out}$ dependent on the currently selected gear level. Note that $\omega_{t,in}$ is assumed to be equivalent to the current engine speed ω_e.

$$\omega_e = \omega_{t,in} = \omega_{t,out} i_t(G) \tag{3.1-5}$$

$\omega_{t,out}$ is calculated from the current vehicle speed v and divided by the wheel radius r_W. If the wheel radius is unknown, it can be set to any positive non-zero value because the relation between vehicle speed and engine speed will still remain the same.

$$v = \omega_{t,out} r_W \tag{3.1-6}$$

In order to estimate the transmission ratios, the engine speed and vehicle velocity measurements are filtered by an FIR. Since FIR filters have constant sample delays, the different measurements can be easily synchronized. The calculated ratios are collected over a period of time and sorted according to the current gear level. Finally the ratio median is taken in every gear level category. Because the transmission hardware changes very little over time, it is sufficient to update the transmission ratios every minute once a large data collection (e.g. over an hour) has been collected. Since the designated HMI devices (see section 2) are smartphone devices or tablets, which are not designed to handle huge amounts of historic data in real time, only the currently estimated ratios are planned to be stored on the hard-disk in the final system for later use. By calculating the average of the historic ratio estimate and the current ratio estimate, the transmission ratio estimation is expected to improve over time.

Engine

The engine model is composed of the engine torque T_e and the fuel consumption rate dV_{fuel}/dt. The engine torque (3.1-7) is calculated from the resistance forces and the estimated transmission ratio. The transmission efficiency η_t describes the percentage of the engine torque that is actually available at the wheels (scaled by transmission ratio and wheel radius).

$$T_e = \frac{r_W (F_{air} + F_a + F_{slope} + F_{ground})}{i_t(G)\eta_t} \tag{3.1-7}$$

In order to avoid detailed engine models, the fuel consumption rate is described by a characteristic map dependent on engine torque and engine speed ω_e. Inspired by [1], this *fuel consumption rate map* is defined as a polynomial $P(T_e, \omega_e)$ of low degree (e.g. first or second degree). The equation is stated in (3.1-8). For estimation purposes, fuel consumption rate, engine torque and engine speed measurements are collected and filtered using an FIR. The polynomial coefficients are then estimated using least squares.

$$\frac{dV_{fuel}}{dt} = P(T_e, \omega_e) = a_1 + a_2 T_e + a_3 \omega_e + \ldots + a_{last} T_e^{\,N} \omega_e^{\,N} \tag{3.1-8}$$

Since the designated HMI devices are smartphone devices or tablets, which are not designed to handle huge amounts of historic data, only the currently estimated polynomial coefficients are planned to be stored on the hard-disk in the final system for later use. By calculating the average of the historic coefficients and the current coefficient estimates, the fuel consumption rate map estimation is expected to improve over time.

<u>Remaining power train parameters</u>

After the estimation of transmission and fuel consumption rate map parameters, it is possible to estimate the remaining parameters which are necessary to construct (3.1-7). The remaining parameters are the resistance parameters from (3.1-1 to 3.1-4) and the transmission efficiency from (3.1-7). Due to limited measurement and data acquisition possibilities, the estimation of power train parameters only allows a limited amount of sub-system estimations. For instance it is not possible to isolate the estimation of ground friction coefficient because the necessary measurements directly linked to ground friction are not available. Therefore, the remaining power train parameters of the inverse model must be estimated in unity. As ground friction and slope may change over time, the resistance force parameters must be frequently updated (e.g. every second). To accomplish this, the model is linearized with respect to the unknown parameters. A least squares approach is used for estimation, with velocity as input measurement and fuel consumption rate as output measurement (**Figure 65**). Samples during gear shifts are not taken into account. Initial attempts using regular least squares in batch process lead to large differences when comparing with the original measurements due to high dynamics within the fuel consumption rate measurement. Thus, the estimation is only carried out on the current measurement samples. In order to avoid strong parameter fluctuations, the regular least squares error function is added by a mahalanobis distance measure that penalizes strong parameter changes from one time stamp to the next. Additionally an active set search is performed on the parameters to avoid unrealistic estimation results (e.g. vehicle mass is never negative). In order to avoid local minima the active set search is not iterative as commonly described in text books [5]. Instead, all feasible solutions are computed. The optimal solution is deemed as the solution that yields minimal deviation from the corresponding measurement and leads to minimal change in parameters. The optimal trade-off is described by a cost function.

The advantages of the inverse formulation are that the most important power train components are regarded (with simplifications) and that the acceleration pedal dynamics has little significance. It is best suited for optimization strategies (subsection 3.2) that attempt to increase fuel efficiency through velocity adaptation instead of changing the acceleration pedal. Actually even approaches that model the chain of effects from acceleration pedal to vehicle acceleration sometimes optimize through velocity and retrieve the corresponding acceleration pedal position from the optimization result in order to avoid integration errors [7]. Through the inverse formulation this detour is not necessary. The drawback of the inverse power train formulation is that it needs to learn the manufacturer specific gear shift program in case of an automatic transmission. These gear shift programs can be highly non-linear and heuristic [3]. Therefore correct classification is difficult to assure. Furthermore, a great amount of unknown parameters must be estimated.

3.1.2 Forward partial power train model

The second approach is a *forward partial formulation* that does not consider the entire power train as illustrated in **Figure 66**. It's only focus is on the present. Apart from the engine the forward partial formulation only describes the transmission in case of non-automatic transmissions. Thus, this model can also be used for vehicles with automatic transmissions if only the present situation is of interest. The effect of resistance forces are also left out because only a part of the entire power train is regarded. The drawback of the partial model is that it is largely a static model because state changes (e.g. vehicle acceleration) can only be formulated, if the power flow from the engine to the wheel and the resistance forces are described.

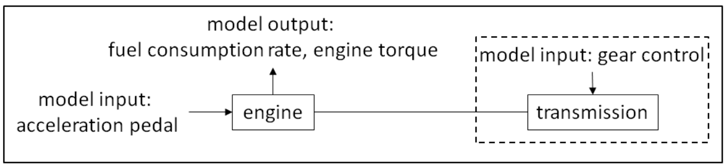


Figure 66: Forward partial power train model formulation

Transmission

The transmission model is described as in the case of the inverse power train (3.1-5 to 3.1-6).

Engine

The engine is described by the same fuel consumption rate map as in the case of the inverse model (3.1-8). Different to (3.1-7), the engine torque is calculated from an *engine torque map* that is dependent on acceleration pedal position u and engine speed ω_e [7]. The engine torque map can be described by a polynomial of 1st and 2nd degree if only little data is available. If large data sets are available (e.g. collected over several hours), it is beneficial to use a three segment spline $S(u,\omega_e)$. The segments are divided at certain acceleration pedal values, e.g. at 20% and 80%. Thus, the spline segments can use different types of polynomials to account for local engine torque map characteristics.

Before parameter estimation is applied it is usually necessary to conduct a histogram test on the collected measurement data. The measurement data is partitioned into a grid depending on torque and acceleration pedal position. From this partition the average measurement point density can be calculated. All partition cells that have fewer points than the average point density are of no further interest. The spline coefficients are then estimated using equality constrained least squares, that takes the common spline intersection equality constraints into account [4]. The different polynomials are developed at different starting points. In our case, $u_h = 100\%$, $u_m = 50\%$. $\omega_{e,m}$ is the engine speed that maximizes the engine torque under full load conditions. It can be retrieved by fitting a parabola through the measurements under full load conditions. All these methods support the stability of the engine torque map and discourage unfavorable interpolation and extrapolation behavior.

$$T_e = S(u,\omega_e) = \begin{cases} a_0 u^2 & ,u < 10\% \\ b_0 + b_1(u-u_m) + b_2(\omega_e - \omega_{e,m}) + \ldots + b_{15}(u-u_m)^3(\omega_e - \omega_{e,m})^3 & ,10\% \le u \le 70\% \\ c_0 + c_1(u-u_h)^4 + c_2(\omega_e - \omega_{e,m})^2 + c_3(u-u_h)^4(\omega_e - \omega_{e,m})^2 & ,u > 70\% \end{cases} \quad (3.1\text{-}9)$$

Since the designated HMI devices are smartphone devices or tablets, which are not designed to handle huge amounts of historic data, only the currently estimated spline or polynomial coefficients are planned to be stored on the hard-disk in the final system. By calculating the average of the historic coefficients and the current coefficient estimates, the torque map estimation is expected to improve over time.

3.2 System constraints

There are currently several technical and economic constraints imposed on the EXPERT system. First of all, routes and position dependent velocity limits cannot always be provided by fleet headquarters. Furthermore, it can generally not be expected that the vehicles using EXPERT possess advanced object detection sensors like radar or camera systems. This makes velocity dependent guidelines difficult to implement in practice. Precise elevation maps are usually subject to fees [20] while free of charge elevation maps have a low resolution [19]. The inverse vehicle model is able to analyze the state evolution from one time instant to the next and is therefore superior to the partial power train model in that regard. But since not even speed limits are available, reliable and sensible model integration during the optimization process is not assured. Furthermore, the inverse model cannot be used on vehicles with automatic transmissions. Thus, the currently planned approach is to use the partial forward adaptive model for fuel efficiency guideline generation.

3.3 Optimized fuel efficiency guideline generation

The optimized fuel efficiency guidelines for the driver are calculated from the vehicle model and the current vehicle internal measurements (CAN-Bus data). The optimization uses cost function minimization, which penalizes unfavorable driving behavior. In this paper, an unfavorable driving behavior is regarded as a driving behavior that leads to overall high fuel consumption and high attrition to the vehicle. Naturally, not driving at all is the best way to accomplish both goals. But at the same time, excessive deviation from the driver's wishes is also unwanted, which includes punctual delivery. All these aspects can be highly contradictory to each other. Thus, the goal is to find an optimal trade-off.

Due to the previously stated system constraints, the partial power train model is used for optimization. The optimization can be conducted through cost function minimization which yields optimal set values for the driver. The proposed cost function consists of two individual cost functions (3.2-1 to 3.2-2) that work best with the partial forward model. The first cost function C_u refers to the acceleration pedal position u. It's first cost term C_{fuel} penalizes high fuel consumption rate. Depending on the current engine speed $\omega_e(t)$ and a possible acceleration pedal candidate u_{cand}, an engine torque candidate $T_e(u_{cand}) = T_e(u_{cand}, \omega_e(t))$ is computed using the estimated torque map (3.1-9). The torque candidate and the current engine speed will lead to a fuel consumption rate candidate $dV_{fuel}(u_{cand})/dt = dV_{fuel}(T_e(u_{cand}), \omega_e(t))/dt$ calculated from the estimated fuel consumption rate map (3.1-8). At the same time, a second cost term C_{torque} encourages engine torque maximization to allow necessary vehicle acceleration. Because the two cost terms cannot be directly compared to each other, normalization values need to be introduced. The fuel consumption rate is normalized by a high fuel consumption rate value $dV_{\max}/dt$. Naturally, it can be defined as the maximum value ever encountered in the collected CAN-Bus data. But in order to avoid excessive measurement errors or high values that rarely occur, it is beneficial to define $dV_{\max}/dt$ as the sum of the expectation and a multiplicity of standard deviation of the available data set. The normalization of engine torque $T_{e,\max}$ is defined in the same way. The minimization of the second cost function C_G leads to a fuel economic gear choice. An estimated engine speed (3.1-5 to 3.1-6) is computed using the current velocity and a possible gear candidate G_{cand}. This engine speed candidate then leads to a fuel consumption candidate depending on the current acceleration pedal position $u(t)$. In order to minimize attrition, additional smoothness terms $C_{smooth,u}$ and $C_{smooth,G}$ with respect to change in acceleration pedal position and gear level are added to both cost functions. They discourage temporal oscillations in the optimal solution.

$$C_u = C_{fuel} + C_{torque} + C_{smooth,u} = \frac{dV_{fuel}(u_{cand})/dt}{dV_{fuel,\max}/dt} + \frac{T_{e,\max}}{T_e(u_{cand})} + C_{smooth,u} \tag{3.2-1}$$

$$C_G = C_{fuel} + C_{smooth,G} = \frac{dV_{fuel}(G_{cand})/dt}{dV_{fuel,\max}/dt} + C_{smooth,G} \tag{3.2-2}$$

The optimized driving guidelines consist of a currently tolerable maximum acceleration pedal position $u_{\max,opt}$ and a currently optimal gear G_{opt} (3.2-3 to 3.2-4). The gear choice must not lead to an engine speed that is below the fuel cut off engine speed $\omega_{e,cut}$ and above the long-time sustainable maximum engine speed $\omega_{e,\max}$. While $\omega_{e,\max}$ and the idle engine speed $\omega_{e,idle}$ can be drawn from the statistics of the collected measurement data, it is currently not possible to retrieve $\omega_{e,cut}$ from the available data records. Therefore $\omega_{e,cut}$ is currently defined as 50% higher than $\omega_{e,idle}$. Threshold violation is detected using engine speed estimation (3.1-5 to 3.1-6). A competing torque maximization term for (3.2-2) is currently not considered because it is not always certain if the driver demands high torque. Since it cannot be expected that the driver will always correctly apply both guidelines at the same time, the two variables are optimized independently from each other. Thus, it is possible to perform the cost function minimization with a direct discretized search in real time. Note that in this paper, real time refers to the perception and reaction time of the driver. The authors define the minimum update rate capability as 4Hz.

$$u_{\max,opt} = \min_u C_u \tag{3.2-3}$$

$$G_{opt} = \min_G C_G, \quad \omega_{e,cut} \le \omega_e(G_{opt}, v(t)) \le \omega_{e,\max} \tag{3.2-4}$$

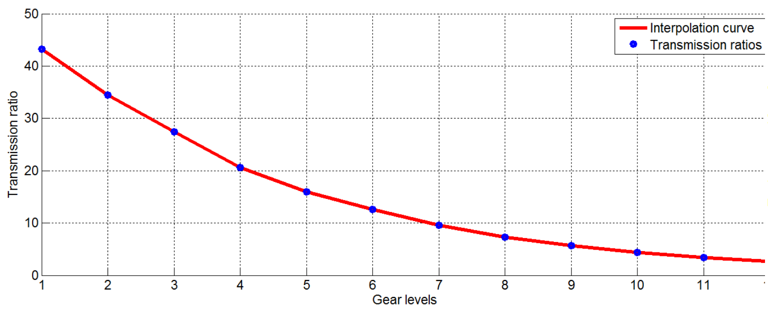
Compared to velocity focused cost functions, $u_{\max,opt}$ (3.2-3) gives the driver more leeway, because there is no precise velocity that must be followed. The driver can choose the preferred velocity and will only get a warning if the acceleration pedal is pressed down more heavily than necessary ($u > u_{\max}$), which also leads to a gradual acceleration behavior. Naturally as a consequence, this new pair of driving guidelines can be less fuel efficient than velocity focused optimization methods. Furthermore, the forward partial model only focuses on the current time instant and acceleration cannot be considered. But the approach is more accessible for the user, only needs CAN-Bus data and can also be applied to vehicles with automatic transmissions (only acceleration pedal optimization).

4 RESULTS

In this chapter the authors would like to present results based on real world CAN-Bus data and simulated environment. The CAN-Bus data are records of a truck delivery that stretches over six hours. It is used to evaluate the performance of model adaptation. Since the manufacturer specific reference data is not available it is not possible to do a quantitative evaluation. The effect of EXPERT fuel efficiency guidelines will be shown in a Matlab/Simulink environment as the final implementation phase of the project has not begun yet.

4.1 Model adaptation

The transmission ratio estimation result is shown in **Figure 67 a**. The highest ratio is approx. 43 while the lowest ratio is approx. 2. The red interpolation line illustrates the exponentially decreasing nature of the ratio values. The decrease is smooth and there are no significant outliers. Using the estimated gear ratios, the engine speed can be predicted depending on the vehicle speed and selected gear. This is important, e.g. if the optimization routine is evaluating the feasibility of the engine speed after a possible gear shift. The result on a randomly chosen sequence is shown in **Figure 67 b**, which shows that the estimated engine speed can mostly follow the engine speed measurement with slightly stronger deviations during gear shifts, because the transmission model is static.



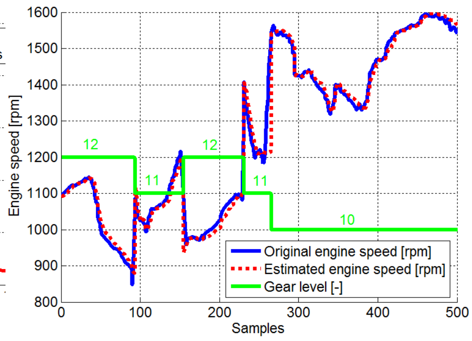
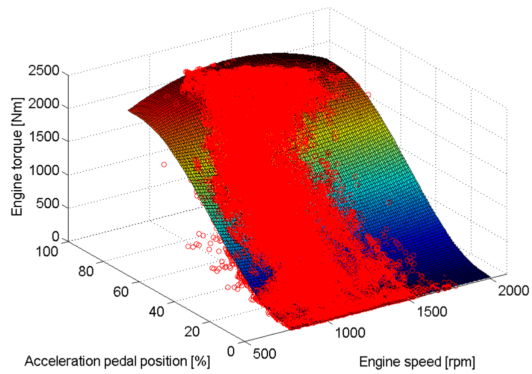


Figure 67: a, b (best viewed in color): Estimated gear ratios (a), estimated engine speed (b)

The estimation result of the engine torque map is displayed in **Figure 68 a**. The measurement points which remain after the histogram outlier test are illustrated as red circles. Note that the engine torque provided by the CAN-Bus is the torque developed in the cylinders according to SAE J1939 specification [21]. It is therefore never negative. The three segment spline is fitted to the measurement points and shows stable interpolation and extrapolation behavior. The spline equality constraints have been imposed on eight different support points along the acceleration pedal positions of 20% and 80%. Several outliers from the original measurement collection could not be discarded. But compared to the main measurement point concentration they are only few in numbers. Using the estimated torque map, the engine torque can be estimated. An example is given in **Figure 68 b**. The estimated engine torque qualitatively follows the recorded engine torque signal, but deviations of more than 10% can occur because the engine map is static and is composed of polynomials of low degree (not more than 3) in order to avoid instabilities.



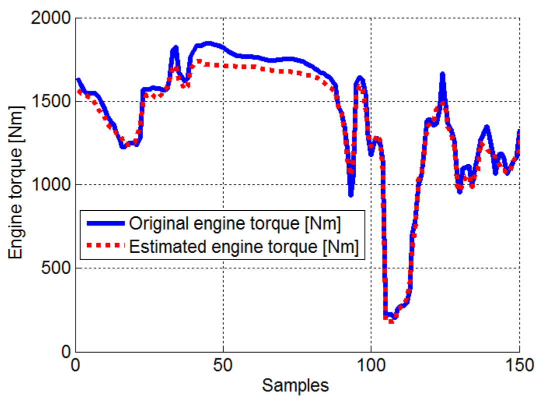


Figure 68: a, b (best viewed in color): Estimated engine torque map (a), estimated engine torque (b)

The estimation result of the fuel consumption rate map is displayed in **Figure 69 a**. The original measurement points are illustrated as red circles. A polynomial of first degree is fitted to the measurement points. It shows good interpolation and extrapolation behavior. Using the estimated fuel consumption rate map, the fuel consumption rate can be estimated. An example is given in **Figure 69 b**. The estimated fuel consumption rate qualitatively follows the fuel consumption rate measurement. Deviations of more than 10% can occur because the fuel consumption rate map is a static polynomial of first degree, but are usually lower than in the case of engine torque estimation.

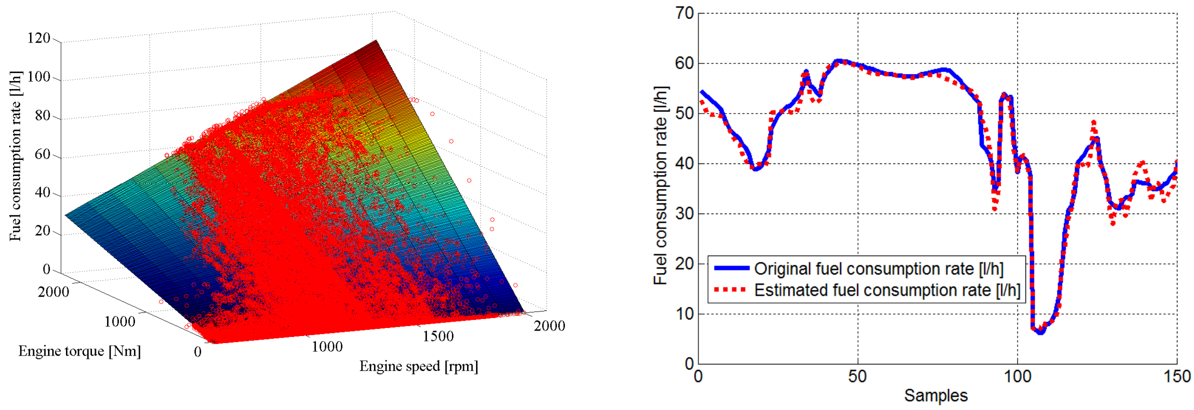


Figure 69: a, b (best viewed in color):
Estimated fuel consumption rate map (a), estimated fuel consumption rate (b)

Note that theoretically, an engine efficiency map can be calculated from the estimated fuel rate map (3.1-8) which describes the ratio of the consumed chemical energy to the available output energy of the engine. The available measurement data for the project was recorded on a vehicle with automatic transmission and was therefore highly concentrated within a certain operation area (**Figure 69 a**), which currently makes a more precise estimation of the fuel rate map impossible.

The fuel consumption rate estimation result based on the estimation of the remaining power train parameters is shown in **Figure 70**. The model output with adapted parameters (red) is usually very close to the filtered fuel consumption rate measurement (blue). This can also been seen in their relatively small absolute difference. The estimated power train parameters (not displayed here for lack of space) can sometimes change significantly over time (e.g. ground friction coefficient). Further investigations are necessary here. Note that this estimation is only necessary for the complete inverse power train formulation.

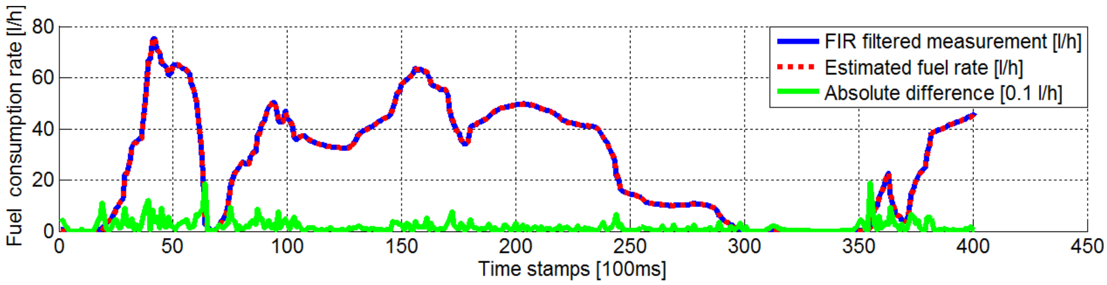


Figure 70: (best viewed in color): Estimated fuel consumption rate (inverse power train model)

4.2 Effect of fuel efficiency guidelines

The effect of fuel efficiency guidelines are demonstrated in a Matlab/Simulink simulation environment. In order to simulate a real world vehicle with manual transmission, a reference model based on the works of [1][15] is created that can describe the interaction between propelling and resistance forces, vehicle internal inertia and clutch dynamics. The simulation environment currently lacks some qualities of a real world vehicle and real world driver. It does not include brakes. The driver can only decrease speed by releasing the acceleration pedal. Furthermore the power flow is not interrupted during gear shifts. The driver is simulated with a PI controller, which tries to retain a set speed of 80km/h and heuristically shifts gears according to the current velocity. Due to the simplicity of the driver model, certain operation areas of the vehicle are rarely reached, which makes model adaptation difficult. Therefore in this subsection, the authors would like to focus on the evaluation of the optimized fuel efficiency guidelines and assume all relevant parameters are known. The simulated vehicle has a mass of 4000kg and 5 gear levels. The road topology has a length of 1800m with two hills. The first hill has a slope of 11% and the second one has 5.5%. It is displayed in **Figure 71**. The vehicle starts with 60km/h and gear level 3. The topology is given in **Figure 71**. The optimization is conducted with the partial forward model, creating a maximum tolerable acceleration pedal position and an optimal gear choice for each time stamp. The results are given in **Figure 72** to **Figure 73**.

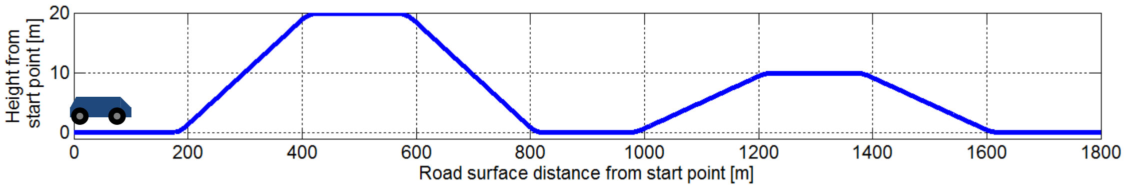


Figure 71: Topology of simulated test road

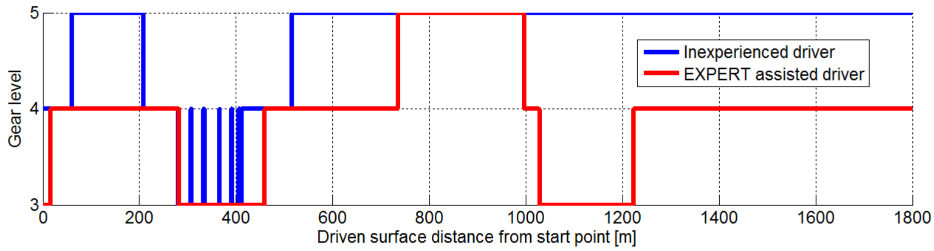


Figure 72: (best viewed in color): Applied gear

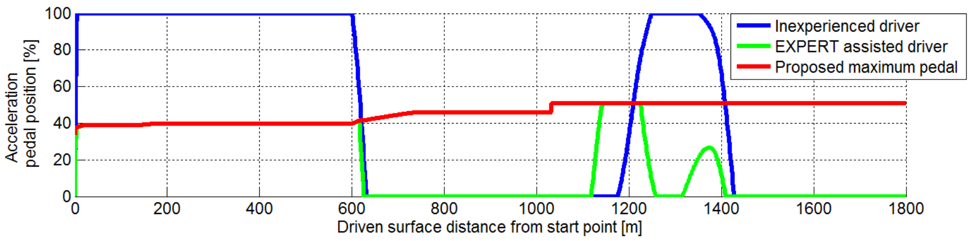


Figure 73: (best viewed in color): Applied acceleration pedal

We can see from **Figure 72** that the inexperienced driver (blue), who solely shifts gears according to velocity, shifts more frequently than necessary, e.g. on the 1.ascent (250m < driven distance < 400m). It is also possible that he shifts at the wrong time or alternatively does not shift at all, e.g. 2.ascent (1000m < driven distance < 1200m). From **Figure 73** we can see that the inexperienced driver (blue) often prefers to press down the acceleration pedal. It can be inferred from the engine's characteristic maps in **Figure 68** and **Figure 69** that a high pedal position above 80% leads to little increase in engine torque but significantly higher fuel consumption. The EXPERT assisted driver rarely goes above 50%. By integrating the fuel consumption over the driven distance, the simulation results show that EXPERT can help the driver reduce overall fuel consumption by up to 31%. Since the current approach does not directly penalize travel time or deviation from a certain set speed, the EXPERT guided driver needed 3% more time to traverse the simulated test road, which is still small compared to the benefit in fuel saving. Simulations on automatic transmissions have also shown a decrease in fuel consumption, which can be as high as 15%. Due to the current limitations on driver simulation, real world experiments will likely lead to smaller fuel savings. Further investigation is necessary in this area.

5 CONCLUSIONS

In this paper, the authors have presented the “Driving Efficiency Module”, which is used in the EXPERT system that has also been briefly introduced. One of the main goals of EXPERT is to provide the driver with an assistance system through the “Driving Efficiency Module” that monitors his or her driving behavior and generate fuel efficiency guidelines to improve fuel economic driving. Two online adaptive vehicles have been presented that can adapt to different vehicles through CAN-Bus data based system identification. The first one is the “inverse power train model”, which describes a simplified power train. Although state changes and resistance forces are incorporated, it requires knowledge of the gear shift control software in case of an automatic transmission. The second model is the “partial forward model”, which only considers a part of the power train. It cannot consider resistance forces, but does not require any transmission models in case of an automatic transmission. The fuel efficiency guidelines consist of a currently sensible maximum acceleration pedal position that should not be exceeded and a currently optimal gear in case of a vehicle with manual transmission. The guidelines are obtained from the minimization of a pair of cost functions which use the “partial forward model”. This optimization strategy only requires CAN-Bus data and no knowledge of the environment or object detection. Simulations have shown that significant fuel savings can be achieved compared to a driver with little experience in fuel efficiency driving.

Future works include the improvement of the simulation framework. The driver simulation will also be improved to emulate a more experienced driver, e.g. by using a Fuzzy controller or by creating a driving simulator. Although the inverse power train model is currently not used by EXPERT, improvement of individual parameter estimation (e.g. slope or vehicle mass) is of great scientific interest. A possible approach is the use of constrained unscented Kalman-filter or constrained particle filter which can estimate the unknown parameters as well as their covariance. A second issue is the optimal gear choice selection which currently only penalizes high fuel consumption rate, which usually leads to a high gear proposal. While high gears are beneficial during coasting and in situations with low torque demand, they may not be the best choice on a steep hill ascent where high torque is demanded. A possible solution is to make use of the current acceleration pedal position to determine if the driver requires high torque. Finally model adaptation can be further improved. The fuel cut off engine speed needs to be estimated and a more precise fuel rate map would enable the application of an engine efficiency map. A precise efficiency map may greatly simplify the optimization procedure, because the engine’s peak efficiency can be directly retrieved from the map.

REFERENCES

[1] GUZZELLA, L., SCIARRETTA, A.: *Vehicle Propulsion Systems*. Springer, Heidelberg, 2005.

[2] HALFMANN, C., HOLZMANN, H.: *Adaptive Modelle für die Kraftfahrzeugdynamik*. Springer, Heidelberg, 2003, (German).

[3] KLEMENT, W.: *Fahrzeuggetriebe*. Carl Hanserverlag, München, 2007, (German).

[4] AHLBERG, J.H.: *The Theory of Splines and their Applications*. Academic Press, New York, 1967.

[5] NOCEDAL, J., WRIGHT, S.: *Numerical Optimization*, Springer, New York, 2006.

[6] FRÖBERG A., HELLSTRÖM E., NIELSEN L.: *A real time fuel optimal cruise controller for heavy trucks using road topography information*, SAE Technical Paper Series 2006-01-0008, 2006.

[7] HELLSTRÖM, E.: *Look-ahead Control of Heavy Trucks utilizing Road Topography*, PhD thesis, Linköpings universitet, 2007.

[8] HELLSTRÖM, E., IVARSSON, M., ASLUND J., NIELSON L.: Look-ahead control for heavy trucks to minimize trip time and fuel consumption, Control Engineering Practice, vol. 17, no. 2, pp. 245-254, 2009.

[9] TERWEN, S., BLACK, M., KREBS V.: *Predictive powertrain control for heavy duty trucks*, 1st IFAC Symposium on Advances in Automotive Control, pp. 121-126, 2004.

[10] TERWEN, S.: *Vorausschauende Längsregelung schwerer Lastkraftwagen*, PhD thesis (German), Karlsruhe Institut für Technologie, 2009.

[11] KEULEN, T., NAUS, G., JAGER, B., MOLENGRAFT, R., STEINBUCH, M., ANEKE, E.: *Predictive Cruise Control in Hybrid Electric Vehicles*, ESV24 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium, Stavanger, Norway, 2009.

[12] HUANG, W., BEVLY, D., SCHNICK S., LI, X.: *Using 3D road geometry to optimize heavy truck fuel efficiency*, Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems, October 12-15, 2008.

[13] FILEV, D., Syed, F.U.: *Applied intelligent systems: blending fuzzy logic with conventional control*, International Journal of General Systems, 4, p. 395-414, 2010.

[14] CASAVOLA, A., PRODI, G., ROCCA, G.: *Efficient Gear Shifting Strategies for Green Policies*, American Control Conference, p. 4331-4336, 2010.

[15] ABASS, A., SHENTON, A.T.: *Automotive driveline modeling, inverse-simulation and compensation*, International Conference on Intelligent Systems, Modelling and Simulation, 2010.

[16] http://www.idsc.ethz.ch/Downloads/qss, last accessed on October 04, 2011

[17] http://www.aglassofwater.org/en, last accessed on October 04, 2011

[18] http://www.blisstrek.com/, last accessed on October 04, 2011

[19] http://www.expertdriving.eu, last accessed on October 04, 2011

[20] http://www.iosb.fraunhofer.de/servlet/is/10772, last accessed on October 04, 2011

[21] http://www2.jpl.nasa.gov/srtm/, last accessed on October 04, 2011

[22] http://www.intermap.com/, last accessed on October 04, 2011

[23] http://www.sae.org/, last accessed on October 04, 2011

[24] http://www.mathworks.de/products/stateflow/demos.html?file=/products/demos/shipping/simulink/sldemo_autotrans.html, last accessed on October 04, 2011

[25] http://rem2030.de/rem2030-de/sitemap/index.php, last accessed on October 04, 2011

Construction I

Innovations in Dredging and Marine Engineering

Michael Schulz[1], Jürgen Reineke[2]
[1]ALLSAT GmbH, Hannover, Germany
[2]Prolec Ltd., Poole, United Kingdom

Abstract - Innovations in Dredging and Marine Engineering
ALLSAT is one of the leading companies in the sector of GNSS applications. The company focuses on the distribution of GNSS equipment and solution engineering.

Prolec has established itself as one of market leaders in providing machine control systems for excavators involved in both dredging and marine engineering activities. Prolec's success in these most demanding application areas has been achieved because of its industry leading sensors technology.
Sensors used in this application are required to be both robust and accurate, e.g. the AS8 marine grade angle sensor meets both of these requirements. The AS8 is made from marine grade stainless steel and when installed offers a true IP68 installation. While other sensors claim to be suitable for marine use, they are often made from materials more suited to land applications, such as mild steel or aluminum and fitted using traditional cable methods. AS8s are installed using hydraulic hose and all connections are made using hydraulic fittings. The sensors are also extremely accurate at 0.09', which is more than twice as accurate as standard accelerometer based sensors.

The second key element in this area is the application specific software Prolec offers. pcX-Pro has several 'off the shelf' variants, but customization can be offered to meet specific requirements. The Systems have been used very successfully in marine applications and offers an open architecture to communicate with several industry standard GNSS input strings. Here the expertise from ALLSAT helps in providing the specific GNSS components. Very good experiences have been made using JAVAD GNSS sensors in RTK-Heading mode in combination with GNSS reference network services, which increases the flexibility and reduces setup times. The presentation will give some insights on the requirements and challenges for these 3D machine control systems.

Prolec systems are commonly found in dredging applications, both land based and barge mounted, and are installed on standard excavators, grab dredgers and rope cranes. One of the biggest increases in use of Prolec systems has been for breakwater and armour rock constructions, with Prolec's products being the choice for prestigious projects such as the Palm Island project in Dubai, the Pearl of Qatar and some others which will be used in our presentation.

Keywords
Machine Guidance, GNSS, Construction Industry

1 INTRODUCTION

ALLSAT is one of the leading companies in the sector of GNSS applications. The company focuses on the distribution of GNSS equipment and solution engineering.

Prolec has established itself as one of market leaders in providing machine control systems for excavators involved in both dredging and marine engineering activities. Prolec's success in these most demanding application areas has been achieved because of its industry leading sensors technology.

1.1 History

Prolec introduced:
1st Can Bus Products in Machine Guidance
1st Excavator Graphical Depth Monitor

1st 3D Excavator RTK Machine Guidance (2001)
ALLSAT introduced:
Graphical Fieldbook including RTK corrections over GSM (1997)
First privately run GNSS reference network (1999)
Survey Software on Windows CE based controller (2000)

1.2 Angle Sensors

Sensors used in this application are required to be both robust and accurate, e.g. the AS8 marine grade angle sensor meets both of these requirements. The AS8 is made from marine grade stainless steel and when installed offers a true IP68 installation. While other sensors claim to be suitable for marine use, they are often made from materials more suited to land applications, such as mild steel or aluminium and fitted using traditional cable methods. AS8s are installed using hydraulic hose and all connections are made using hydraulic fittings. The sensors are also extremely accurate at 0.09', which is more than twice as accurate as standard accelerometer based sensors.

Figure 74: AS8 marine grade angle sensor

1.3 Software

The second key element in this area is the application specific software Prolec offers. pcX-Pro has several 'off the shelf' variants, but customisation can be offered to meet specific requirements. The Systems have been used very successfully in marine applications and offers an open architecture to communicate with several industry standard GNSS input strings.

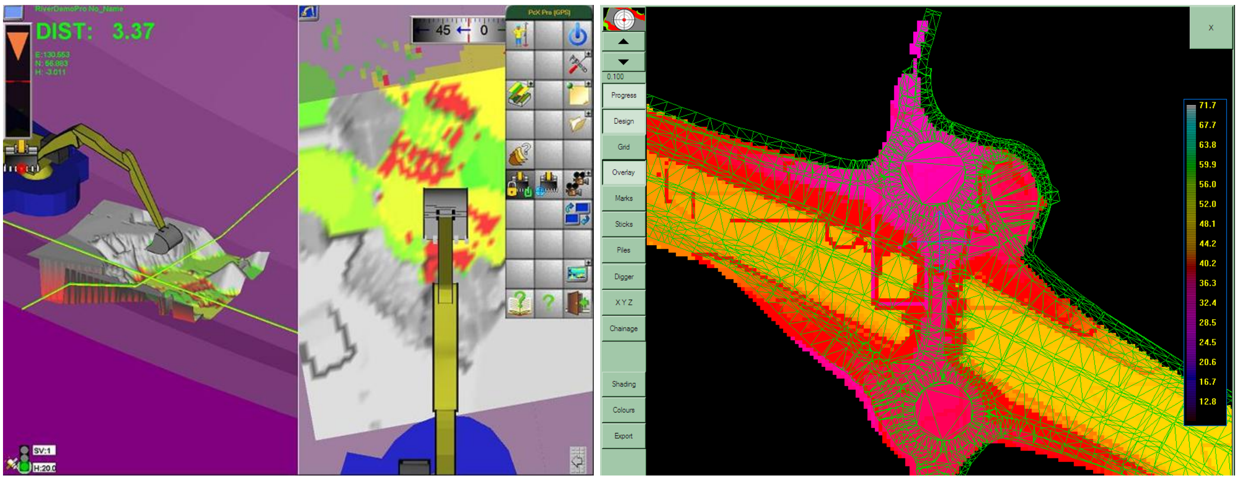


Figure 75: pcX-Pro software

1.4 GNSS Sensors

Here the expertise from ALLSAT helps in providing the specific GNSS components. Very good experiences have been made using JAVAD GNSS sensors in RTK-Heading mode in combination with

GNSS reference network services, which increases the flexibility and reduces setup times. The presentation will give some insights on the requirements and challenges for these 3D machine control systems.

Figure 76: JAVAD GNSS sensor and antenna

2 APPLICATIONS

Prolec systems are commonly found in dredging applications, both land based and barge mounted, and are installed on standard excavators, grab dredgers and rope cranes. One of the biggest increases in use of Prolec systems has been for breakwater and armour rock constructions, with Prolec's products being the choice for prestigious projects such as the Palm Island project in Dubai, the Pearl of Qatar and some others which will be used in our presentation.

Figure 77: Backhoe Dredger working for via donau

3 CONCLUSIONS

The use of Angle and GNSS sensors for marine dredging applications is common practice. However, the seamless integration of the different sensors in one common software requires a profound knowledge of the customer requirements. Future improvements will be done by 3 axis dual redundant sensors, Multi-constellation GNSS sensors and integration of safety applications.

REFERENCES

Links:

ALLSAT GmbH (www.allsat.de)

Prolec Ltd. (www.prolec.co.uk)

Curb & Gutter Concrete Paving with A Wirtgen SP 15 and Wirtgen AutoPilot

Matthias Fritz, Wirtgen GmbH,Windhagen, Germany

Abstract

Wirtgen is building concrete paving machines for over 30 years and has released a brand new series of slipform pavers in 2010. All these machines are highly intelligent and can be run in full 3D operation, meaning steering, elevation and slope of the machine are fully automated. Wirtgen has also introduced their first in-house development of a 3D control system which is deeply integrated into the machine. It allows the machine to operate without any string-lines and fully utilizes the flexibility of the machine's geometrical design. Of all paving machines in the market, the small pavers for curb&gutter applications are the most complex ones and require a great deal of sensors and onboard intelligence in order give a good result to the user. These machines are required to follow a perfectly straight line e.g. for safety barrier work on highways as well as paving concrete in very small radii such as 60cm. This is very demanding for the machine and the control system.

The Usage of Machine Control in 2012: Status-quo and Practical Implementation Challenges

Hellmut Billinger

Abstract

- State of development of MCG from the perspective of construction site
- Development opportunities for MCG of construction site
- Improvement of construction processes through automation

1 INTRODUCTION

Machine control has been around for a long time. When I started to work on machine control systems in the late 1970s, the technology was neither established in the construction industry nor offered by construction machine producers. Since then, things have drastically changed. In this presentation, I will briefly summarize my conclusions from working on machine control systems for more than 30 years.The presentation will cover the evolution of the technology's premises. I will discuss the changing goals that the construction industry was trying to achieve while adopting machine control systems. I will then talk about four implementation challenges that machine control and machine guidance face in Germany today. To illustrate these challenges, I will present a recent successful application of machine control. I will also present a small scenario that shows how machine control could work in the future. I will conclude my presentation with some take-always for academic research and development.

2 MACHINE CONTROL AND MACHINE GUIDANCE: HOW THE PREMISES CHANGED

Machine control was introduced on construction sites many years back. Today it is an important technology that everyone uses in order to reduce time and cost. When I started to work on machine control in the late 1970s and early 1980s, the main motivation for working with this technology was different. Our central aim was to improve production processes. A central aspect, for example, was the evenness of surfaces. Efficiency and quality management concerns were present, but more of secondary importance. Also, during these times it was usually difficult to convince senior management that the adoption of machine control was beneficial for the firm. Fortunately, we do not face this problem anymore.

Today the goals that construction firms pursue when implementing machine control are clearly efficiency and quality management. These goals will also be present in the future, and every implementation of machine control and machine guidance will be rooted in efficiency considerations. This is because this technology cannot be used to justify an increase of "return of investment" based on higher prices and higher margins. The economic baseline condition is well defined for machine control and machine guidance.

This brings me to one important aspect: To achieve additional efficiency, it is very important to understand complete construction scenarios with the underlying administrative processes. Without this understanding, the technology's role in practice can easily be misunderstood. On the following slides, I will illustrate aspects of this by discussing the current implementation challenges.

3 IMPLEMENTATION CHALLENGE 1: DATA INTEGRATION

There are many reasons why the underlying administrative or business processes cannot be used today. First, data standards and data interfaces are often outdated. It may seem odd, but some basic data structures still derive from keypunching standards, which are certainly not state-of-the-art. Second, standards and interfaces are often not defined well enough in order to fulfill today's needs. Third, there are many different regulations on the national and the international level that currently prevent more advanced solutions. Fourth, machine manufacturers often follow their own interests and do not agree on join standards. Future regulation should consider these aspects and aim at universal data standards and data formats. This is likely to ease the implementation of technology and reduce overall cost.

4 IMPLEMENTATION CHALLENGE 2: DATA IS BOTH, NOT UNDERSTOOD AND HIGHLY STRATEGIC

The second implementation challenge concerns the roles that the various players in the construction industry have. In a simple model we can assume that a typical construction site has a contracting body, planners and architects, a principle contractor which is often the construction firm, and some suppliers who, for instance, rent machines with control systems to the construction firm.

Most remarkable is that among the various players in this value chain, we find completely different attitudes and levels of understanding. On one hand, some planners and architects are often not only unfamiliar with advanced data interfaces but are also very reluctant to learn about the potential benefits of machine control. On the other hand, we have players in the value chain that are very aware of the technology's potential. They use this knowledge to their advantage, as a bargaining chip, in order to strategically maximize their profits. For the implementation of machine control, these aspects are critical when it comes to realizing modern machine control solutions.

5 IMPLEMENTATION CHALLENGE 3: THE HUMAN-MACHINE INTERFACE

Another important implementation challenge is the human-machine interface. Already with the first implementations in the 1980s, the success was greatly determined not by the engineers or the project manager. It was determined by the machine operators who only accept changes in their work procedures if they make sense to them. Only if machine operators see their own benefit, technological advances can be implemented. During the years, I have often heard statements, such as "…now I see what I do…", or "…I have additional opportunities to improve my work…" These statements show that machine control can make daily work more interesting and thereby also more valued by everyone involved.

Also important is that machine operators in Germany are required to pursue a 3-year apprenticeship. This education is important for machine control and must not be neglected. Also, it is critical to offer state-of-the-art training programs as the technology develops over the years.

6 IMPLEMENTATION CHALLENGE 4: INTRA-ORGANIZATIONAL ACCEPTANCE

I have already mentioned that senior managers are meanwhile open to implement advanced machine control systems. Also, the statements from the operators above show a positive attitude towards the machine control. However, it is important to note that the actual adoption of machine control can still encounter implementation challenges within the organization. Particularly middle managers sometimes view possible solutions as threads. Typically machine control then becomes risky because it can challenge the organizational status of these middle managers. For the actual implementation, it is therefore important to also consider the intra-organizational consequences of technology implementation.

I now would like to illustrate some of the implementation challenges using a recent example.

7 A RECENT EXAMPLE: HOW GPS-BASED MACHINE CONTROL CAN BE USED

The Deutsche Bahn was confronted with a major task at the main train station in Munich in late 2011. All 272 analogue signals had to be replaced with digital signals. For this task, the main train station was shut down for 78h. A project team consisting of 265 workers completed this task following a precise work schedule. Critical for the realization was a real-time animation of the current status of all machines involved, that is, 9 excavators, 6 service trains and 3 mobile cranes. For all these machines, it was important to know the exact location at all times. This means, in this context, a location accuracy of less than 10cm and a data update every 10 seconds. For the realization, we used state-of-the-art GPS technology, a centralized server, and a projector in the project's command center which showed the actual position of each machine on a large screen. On the screen the layout of Munich's railroad system as well as the various machines that are active during the project could be seen at any given point in time. For all areas of the map, it was possible to zoom in and zoom out in order to monitor the exact positions of all machines.

Because the machine control was openly visualized and could be seen from everyone involved, the coordination between the project manager and his assistants, the various work teams and other people involved was facilitated. Moreover, this approach did not only prevent accidents, it also allowed for a reduction of inefficiencies. One can also imagine how such a system could easily be further developed to analyze machine utilization and other operations management measures.

What was interesting in this example is that the initial project proposal seemed over-engineered. However, the actual project realization turned out to be very successful. Everyone had a clear overview and relevant information was made available in real-time. Further, the actual data had a high accuracy and allowed for firefighting in critical situations during the project. Also important, all deadlines were kept and the overall project was successful.

While all the above may seem obvious, there are a few aspects that I would like to point out. First, the data was more accurate and all of the data could easily be saved. This allows for a precise documentation of the actual construction process, which is very useful for the accounting of the project. Also, as you have seen, we now know exactly who was where at what time, which can be used to improve business processes. Second, the visualization of machine-controlled vehicles turned out to be very useful when the regular train traffic was re-established at the end of the project. This created a significant benefit for the re-opening of regular operations at the railway control center. Several of these benefits were not considered initially, they were unexpected. This highlights the importance of thinking in scenarios, which I would like to illustrate using another example.

8 THE ROLE OF SCENARIOS IN DEVELOPING FUTURE MACHINE CONTROL APPLICATIONS

Let me develop a scenario of a simple construction task: Highway-stone laying in several layers. It requires one grader, several compactors and several trucks. Although this task is done very often, we are yet far from having an advanced application of machine control and machine guidance. The scenario goes like this: First, before the actual construction can start, we need an automatic assessment of the existing surface. This assessment includes a comparison with the plans and identifies the differences between the existing and the planned surface. The grader then lays the first layer of stones, thereby accounting for the identified differences while constructing the target surface. The grader

monitors its progression, including a permanent control of its stone consumption. Second, compactors compress the stones. They monitor the compression rates and how they change for different areas. Third, the grader continues its operation, thereby using latest information from the compression and drawing the required supply that is provided by the trucks. Overall, this scenario seems simple.
However, we are far from having what I would consider a satisfactory solution, mainly because of the implementation and unsolved technical challenges. At the same time, today it is already clear that this task's future scenario will look like the one that I just described. The scenario may differ slightly, but increasing demands in quality management, e.g. because of liability reasons, will require a solution that builds on sophisticated construction documentation based on advanced machine control and guidance.

I also would like to give you some additional perspectives on why the described scenario is not working at this point. Some reasons draw from what I said initially. One implementation challenge that we face is that software and hardware have been developed to a level where they satisfactorily work within machines, but not across machines, especially when machines come from different manufacturers. Also note that some of the key sensor technology, like sensors that monitor compaction, is not developed sufficiently.

The second major issue is that planning data cannot effectively be used on the various machines. Data standards are often not compatible and real-time solutions are not feasible. In addition, and this is a technical restriction, the measurement of altitude is problematic with current GPS solutions.

9 TAKE-AWAYS FOR ACADEMIC RESEARCH AND DEVELOPMENT

Let my summarize some insights and discuss take-always: Data integration is crucial and data standards and data interfaces need to be defined. Data from the planner should easily be transferable to the construction firm which should easily be able to give it to project management and the operator on the machine. At the same time, data describing the actually constructed results should also be able to go exactly same way back. To reap the real benefits of this data exchange, all these steps should be possible in real-time. This also means that some of the existing regulations in construction need to be changed.

The motivation for such changes must also be clear: It is accurate documentation of complete construction processes, which will allow for higher quality, quicker processing, better documentation and finally more efficiency in construction.

For research to be effective, it is essential that full scenarios are considered and bottlenecks in the overall process are identified. I therefore encourage you to think in complete scenarios, disaggregate them and develop realistic coherent solutions for the bottlenecks. The measurement of altitude with GPS and the accurate adjustment of compaction are two examples.

Finally, let me summarize what I think is mandatory for the future development of machine control and machine guidance. This technology is the technical interface between the construction process and the administrative processes of the firm. The connection between these processes is important and largely determines whether this technology can be effectively put to work in practice. With my presentation, I hope that I have provided you some ways of thinking about the next steps of machine control and machine guidance in academic research and development.

Construction II

Machine Control and Guidance Projects in Engineering Surveying

Ulrich Völter
intermetric GmbH, Industriestraße 24, 70565 Stuttgart, Germany

Abstract
For a long time surveying experts have delivered current, accurate and precise geometrical data for the planning and construction industry as well as for the managers of buildings and operators of infrastructure of all kind.

The modern construction industry has increasing demands on precision and construction speed. In this article various projects of the slab track and tunnelling construction industry are described and illustrate how the modern engineering surveyor can respond to these requirements. In each of the depicted projects a machine control or machine guidance system, developed by the author's company, helped to fulfil the demanding technical and economical goals.

Keywords
Multi sensor systems, total station, near photogrammetry, inclinometer, laser triangulation sensors, construction industry, slab track, tunnelling, machine guidance systems, machine control systems

1 INTRODUCTION

For a long time surveying experts have delivered current, accurate and precise geometrical data for the planning and construction industry as well as for the managers of buildings and operators of infrastructure of all kind.

Surveying experts provide the base data for planning new buildings and constructions. If used properly and if used at all, the geometrical data can help to make sure that the newly planned architecture will suit the needs, fit to the existing infrastructure and harmonize with the neighbourhood. During construction surveyors will translate the construction plans into the real world by staking out reference points, lines and planes, making sure, that dimensions and orientation of the new building comply with the planning. Finally an as built survey documents all relevant features of the new construction in suitable detail. With modern 3D-Laserscanners billions of precise 3D-surface coordinates can be obtained in a single day. 2D plans such as layouts and sections, 3D models and GIS systems which enrich geometrical data with technical data can provide the information needed to the operator.

This is the traditional way of seeing surveying services: They are links in a one dimensional process:

idea - achieving base data - planning - staking out - building (- checking) - staking out - building ... - documenting - using

Note that the multiple staking out / building / checking is not a process loop because the single tasks usually do not interact with each other but different parts of the construction are staked out at different times e.g. construction pit, basement, cellar, first floor, roof, fence.

Economic efficiency, high train speeds, complex architecture, inner city constructions and environmental issues like route bundling rise more and more the demand on more precise constructions and more precise construction processes. The precision needed can often be achieved only by an integrated task where the construction is controlled by the surveying. As machine time is usually very costly, the idle time for the machine waiting for the adjustment data shall be as short as possible. Therefore a high grade of automation in the measuring - evaluating - result presenting - process is required.

The surveying instrument industry provides good instruments with multiple measuring and automation features and most important with interfaces. Since the first appearance of these sensors, experienced surveying engineers have quickly integrated them in even more complex integrated machine control and guidance systems. After some time the industry has adapted a number of these systems for the mass market.

Pioneering projects however are still realized by surveying engineering companies with operating experience in the respective area of the construction industry and with profound know-how in programming, electronics and automation.

2 PROJECTS

The intermetric GmbH is a Stuttgart based surveying engineering company with 7 offices in Germany and surveying projects in Germany, Europe and around the world. intermetric has more than 40 years of experience, is a market and technology leader in engineering surveying when it comes to slab track and tunnel surveying and has as such developed a number of integrated machine control and guidance systems.

2.1 Slab Track

The first German high speed routes, Hanover - Wurzburg and Mannheim - Stuttgart, both fully put into service in 1991, were built with conventional ballast technology. The high axle loads of the freight trains combined with the high speeds of up to 280km/h of the passenger trains resulted in early wear of the ballast and as a result in high maintenance costs for the tracks. In consequence a new, ballast less track construction type was studied: slab track.

The basic idea of slab track is to fix the rails once and for all in a close to ideal position in order to get a maintenance free track for many years. One problem is the fixation of the track. This was faced by the construction industry and a series of systems emerged: Rheda 2000 by Rail.One, LVT by vigier rail, ÖBB Platte by Porr, FF Max Bögl, FF Züblin, etc. A second problem is to bring the rails or their fixation points in a close to ideal position prior to fixation. The techniques used in conventional ballast track construction could not be used because the tamping machines require ballast in order to operate.

This second problem is a geometrical challenge and as such a task for engineering surveying.

The demands on accuracy are very high - in fact amongst the highest in engineering surveying: the requirements vary from project to project but generally speaking, the final result, i.e. the operational track, must not differ from the design more than few millimeters. Therefore the surveying precision must be some tenths of a millimeter. This precision is to be provided in open air, under varying climate conditions, repeatedly for each single sleeper or slab and integrated into the construction process. Therefore intermetric studied a number of machine guidance and machine control systems that support the positioning processes.

2.1.1 intermetric highPos

In the mid 1990s the new high speed line Hanover - Berlin was built. Züblin was one of the contractors, assigned with the construction of the slab track sections. In order to automate the elaborate process, Züblin had constructed a machine that could take 10 concrete sleepers with its clamps and vibrate them in a bed of stiff but fresh concrete. Once the vibration stopped, the sleepers would not move due to the stiffness of the surrounding concrete. intermetric‘s job was to guide the vibrating machine into the exact position. In a first attempt this was realized with a total station. It resulted to be effective but not very efficient. This was due mainly to the fact that the construction process and the surveying had to wait for each other too long: During all the preliminary construction tasks needed to grab the sleepers and to bring them into a rough position over the next construction position, there was not much to do for the surveying team. During the time consuming measuring and positioning process the construction team had to wait for the results. Therefore intermetric studied and implemented a different approach during the project: The frame holding the sleepers should position itself relative to pre-positioned targets on the ground. That would de-couple the two processes and bring therefore a significant speed-up in productivity.

Despite the short development time and on-site testing, the system was fully operational and in fact gaining significant increases in speed. Unfortunately the last two sleepers of each section were affected by the vibration in the next section despite being held with a fixation frame. This resulted in considerable reworking. For the project „Frankfurter Kreuz“ the process was modified and the system could again demonstrate its capacity. But the problems with the displaced sleepers due to the vibration in the following section continued. In consequence Züblin abandoned the sleeper-positioning machine. The positioning system, called intermetric highPos, however had a revival on the high speed line Köln - Rhein/Main:

The company Walter-Heilit Verkehrswegebau had been assigned with the construction of the slab track in lot A of the new high speed line Köln-Rhein/Main. Due to previous delays, the 83km of slab track had to be built in only 5 months in the year 2000. The intermetric was awarded with the surveying work. Most of the work was done in a well-trained manual process. But Walter-Heilit had aimed for an automation of the process and - as previously Züblin - constructed a sleeper-positioning machine. The concept differed however: At first the machine positioned the sleepers and secondly the concrete was poured in the section in order to fix the sleepers once the machine loosened the clamps. Therefore no vibration should influence the previous section. As an additional feature, the frame holding eleven sleepers was constructed in a way that it could be bent in order to form an arch that resembled the curved sections.

Again intermetric was awarded with the design and delivery of the machine guidance system. Whilst the idea of highPos was maintained, the system was largely extended and newly developed: Each corner of the frame was equipped with a sensor box capable to measure the relative position (x, y, z) to a target placed on the ground as well as the absolute longitudinal and lateral inclinations.

On the shoulders of the concrete trough pre-calculated coordinates were staked out to the centimeter and marked out with brass marks. Their effective 3D coordinates were determined by a network measurement with subsequent geodetic adjustment. The targets were concentric black and white circles on an aluminium plate mounted in a tripod in order to place them exactly horizontal over the ground mark.

The sensor boxes were equipped with video cameras, laser distance meters and two inclinometers plus light and some electronics.

The measurement values, the predetermined coordinates of the ground based targets and the machine parameters were put into a least square adjustment in order to derive the absolute position of the machine frame in the global coordinate system. This position was then compared with the track design and the displacement values for the motors were calculated and directly transmitted to the motor control system.

2.1.2 intermetric Track Lifter

In 2005 Rheda 2000 vof, a consortium of the dutch BAM NBM and Pfleiderer Track Systems (today: Rail.One) was assigned with the construction of 186 km of slab track on the HSL Zuid, the dutch high speed line connecting Amsterdam and Rotterdam via Belgium to the European high speed network.

Previous to the track work, a concrete sub structure, the so-called civil-plates, bridges or tunnels had been constructed by other contractors. On this sub structure the actual slab track had to be built. in general the process was as follows: A geotextile, the longitudinal reinforcement and the sleepers were laid out. 15 m long installation rails were mounted on the sleepers in such a way, that each pair of rails and 23 sleepers formed an independently movable track panel. The track panels were lifted and rough-adjusted by the track lifter and placed on spindles. Then additional reinforcement and the formwork were mounted. Fine adjustment with the intermetric Laser alignment process brought the rails in the final position. Then the concrete was poured in the formwork to form the slabs.

So Rheda 2000 vof followed a different approach than Züblin and Walter-Heilit Verkehrswegebau: Instead of a complete automation only the heavy pre-positioning, the lifting and rough adjustment was supported by the automatic track lifter.

The track lifter comprises four identical, mechanically independent portals, pulled by a tractor unit, which was equipped with a diesel generator and central components of the motor control system. Each of the portals could move up- and downwards, had a horizontally movable sledge and mounted on the sledge a tilt-able beam equipped with clamps to grab the track panel. Every movement was effectuated by electronically controlled precision motors and gears. The intermetric Track Lifter control system measured the current position and inclination of each beam and calculated the displacement values for the vertical, horizontal and tilt movements needed to bring the track section into the desired position.

The intermetric Track Lifter system (iTL) is composed of a remotely controlled total station, one inclinometer and one prism with remotely movable prism cover on each of the four frames, electronic equipment for analogue-digital-conversion and data transmission and a central controller featuring a ruggedized PDA with the iTL-software that collects all sensor data, does the number crunching and presentation of the displacement values.

The machine guidance system iTL has proven to be very effective. In fact it operated on the HSL Zuid on four track lifter sets. As a follow-up of the HSL Zuid, Rheda 2000 sold a big share of their technology to China. Amongst this were track lifter machines which resulted in a delivery of 13 iTL systems to China.

2.2 Tunnel surveying

Tunnel surveying is a second field of engineering surveying where more and more processes are being automated and more machines get precisely guided or controlled by geodetically determined information.

2.2.1 intermetric Tunnel Checker

As most other industries, the tunneling industry is under constant pressure to deliver high quality standards at low costs if not higher quality at lower costs. One of the challenges that are to be faced in every tunnel construction is to deliver a tunnel profile not smaller than the design, the inner shell shall not be thinner and only a little thicker than designed. Therefore the profile of the outer shell must also be close to design and must not be too narrow.

The way to achieve this is to build the excavation a good share larger than needed and to compensate for with the outer shell concrete. The drawback of this procedure is that superfluous excavation is expensive as well due to the extra excavation as due to the extra material needed for the outer shell. Therefore the goal is to build a tunnel just big enough to produce a profile free outer shell.

But as always when the limits are tested, sometimes they are exceeded. For these cases reworking is needed. At first the areas that require reworking need to be identified. This is done by 3D laser scanning: With a 3D laser scanner the surface of the outer shell is measured and by the use of specialized software, the local scanner coordinates are transformed onto the design profile in order to derive a contour map which displays the over and under profiles in the scanned tunnel section. Now that the rework areas are identified on the plot, they need to be identified also in the tunnel which is much more difficult and further more the operator of the reworking machine needs some indication when to continue working and when to stop. Without a measuring tool the operator can decide only based on his experience or guessing. This inevitably leads or to over-rework or to under-rework which leads to re-rework or, if omitted, to poor quality.

This is where the intermetric Tunnel Checker comes into play. iTC is a system consisting of a remotely controlled total station, a controlling ruggedized tablet computer to be mounted in the reworking machine and a box containing communication and power supply electronics.

The reworking can be prepared in the office: The responsible engineer loads the evaluated contour plots into the iTC software and marks the areas to be reworked. These areas and the contour plot are transferred to the system mounted on the machine via USB-Stick or any other available means of communication.

In the tunnel the machine operator finds the reworking area due to the chainage markings in the tunnel. He places the total station close to the reworking area and performs a setup using the special free stationing component of intermetric, which identifies the points and their coordinates based on the configuration of the measured fix points. Once the setup is completed and he is back in his cabin, the machine operator starts iTC and selects the area to be reworked.

The total station starts in reflectorless mode to continuously measure the marked area in a defined raster, typically 50 cm x 50 cm. The laser of the reflectorless measurements is clearly visible in the dark tunnel. Therefore the operator can easily identify the area to be reworked. He starts with his milling. The total station continues to measure the marked area repeatedly during the milling. Every measurement gets immediately transformed on the tunnel surface and the new distance to the design is directly displayed, keeping the display always updated with the changing reality. In that way the operator knows exactly when and where he has reached the design and where there is still some under profile to be removed.

The same system works also for applying shotcrete: iTC continuously monitors the area where shotcrete is applied and makes sure that the profile is exactly reached without subsequent rework.

2.2.2 intermetric Drilling Control

At Victoria Station, London, a new passenger tunnel is being built. This does not sound very spectacular at first sight but taking into account that Victoria Station is in the heart of a 2000 year old metropolis with all kind of archeological, historical and technical surprises waiting under the ground and taking into further account that more than 90 double decker busses an hour and more than 60.000 pedestrians a day pass the construction site it becomes clear that the project is not standard.

The contractor decided to build a waterproof concrete tube in the ground by jet-grouting hundreds of mini piles from the surface.

In order to position the drill rig such that the jet is placed at the desired position and produces the planned mini pile, the contractor decided to use the intermetric Drilling Control - iDC system that is especially designed for the geometrical guidance of inclined boreholes.

The iDC system consists of a total station featuring remote control, permanently machine-mounted prisms, calibration and configuration data and the on-board software intermetric Drilling Control. The total station is preferably mounted on a fixed console. On the machine there are only two passive 360° prisms needed. Both are mounted on the drill rig, one close to the bottom, the second close to the top of the rig. Their position relative to the drill is determined in a one time machine configuration process and stored in a configuration file.

The iDC software is designed to be mainly used by the drilling machine operator. The standard procedures „stake out ground points“ and „adjust drill rig“ are understood after short introduction. Therefore the system is particularly flexible.

Typically the machine operator checks the machine configuration in the „Settings“ dialogue, and selects a „borehole definition file“. This ASCII file contains the coordinates of the end point and direction point for every borehole.

One typical problem for inclined borings is to find the starting point or ground point of the boring. This becomes clear when taking into consideration that for inclined borings the horizontal position of the penetration point depends on the height of the base grade and that the base grade on construction sites is normally not exactly defined.

Therefore the function „stake out ground points“ guides the operator in an iterative process to the penetration point or ground point of the boring. Typically the operator stakes out and marks the ground points for a number of borings in one go.

In the „adjust drill rig“ dialogue, the machine operator selects the current boring. He sets up the drill rig with the tip of the drill precisely on the staked out ground point. iDC automatically finds the two prisms mounted on the drill carriage based on the machine configuration. Immediately after having measured the prisms, the displacement values are calculated and continuously displayed. If the upper end of the carriage is shifted by this value, it will be precisely aligned in the drilling axis. Normally this is an iterative process with two or at most three repetitions.

Even during drilling, these displacement values are updated at one-second intervals so that the drill rig can be continuously (re-)aligned to the drilling axis. Additionally the software displays the distance of the boreholes end point from the design, presuming the current direction is maintained.

One of the technically less but contractually more important features is the complete, operator based documentation of all effectuated operations, including stationing, measuring control points, staking out ground points and of course adjusting the drill rig.

3 CONCLUSION

Engineering surveying does, as ever, collect measurements, derive local, global and problem specific coordinates, and present them in a suitable way. This was, is and will be the core competence of well educated engineering surveyors. But nowadays this set of core competences needs to be redefined in order to comply with the more and more demanding and more and more complex and more and more changing requirements of our surrounding world.

We engineering surveyors have the skills, high technology gets easier to use and the requirements are out there somewhere. Therefore: let`s go for it.

ACKNOWLEDGEMENTS

The authors would like to thank Markus Müller, Markus Federmann, Roman Pils, Oliver Höpfer and Jürgen Kienle with their teams to successfully work on the projects mentioned and again Jürgen Kienle to give me some hints regarding the language.

An FPGA-based Hybrid Positioning Platform for Machine Control Applications

Qizhi Shi[1], Xiaolin Meng[1], Yiqun Zhu[1], Alan Dodson[1], Karl Soar[2]
[1]The University of Nottingham, Nottingham, United Kingdom
[2]Hexagon Machine Control Division, Nuneaton, United Kingdom

Abstract
An FPGA based hybrid positioning platform is herein proposed for construction machine control applications. It is known that provision of precise vertical positioning and attitude determination are two most important aspects in machine control. The proposed hybrid positioning platform consists of one Xilinx Virtex-6 FPGA board, three different GNSS (GPS, Galileo and GLONASS) receivers, a triaxial accelerometer and a GPRS/3G communication module. On this platform, these GNSS receivers, accelerometer and GPRS/3G communication module are connected to the Xilinx Virtex-6 FPGA board directly. In terms of functionalities, these three GNSS receivers are working together with the GPRS/3G communication module to connect to correction resources for conducting real-time kinematic (RTK) positioning with an attempt to improve the vertical positioning accuracy to be better than 5 cm; these three GNSS receivers are also used to facilitate instant attitude determination for machine control; and the accelerometer is used to cancel out the effects of the vibration characteristics of construction machinery, for the determination of improved position and attitude estimates.

The following research methodologies are adopted in this paper. Firstly, the FPGA board acts as a data acquisition board to acquire the data streams from all these sensors and devices. Secondly, with the acquired data, Kalman filter based Matlab scripts are developed for the data processing to implement the functionalities. Finally, Matlab scripts will be converted into C/C++ code and migrated to the FPGA MicroBlaze processor to form an FPGA based hybrid prototype positioning platform. Furthermore, this code will be embedded into hardware cores by using FPGA hardware logics and parallel processing capacity to achieve optimal real-time performance.

Keywords
GNSS, FPGA, machine control

1 INTRODUCTION

The Global Navigation Satellite Systems (GNSS) are playing an increasing role in positioning, navigation and autonomous machine control applications. Recently, the Global Positioning System (GPS) has been integrated into the design of bulldozers, motorgraders, drills, excavators, pavers, agricultural equipment produced by the major manufactures for mining, construction, agricultural and environmental applications.

For today's contractors, GNSS-augmented machine control is of growing importance; the benefits offered are apparent, in terms of precise, timely and efficient work delivery, automated job and data management, and effective and accurate billing. When GPS is used for machine control, its major roles include both positioning and navigating the machine's prime-mover and determining the attitude of the attached working tool (blade, bucket etc).

However, the practical limitations of space-based positioning technology are well-documented; on opencast mining sites for example, where the depth can often range from two hundred to six hundred metres with its wall slopes exceeding 55 degrees, tracking a sufficient number of well-distributed GNSS satellites proves to be extremely difficult and the signal degradation caused by diffraction or deflection sometimes makes positioning impossible. Use of the new GNSS signals and emerging proprietary augmentation technologies such as "pseudolites" is considered essential to overcome the difficulties that are faced by GNSS-only positioning systems, for the provision of accurate, reliable and continuous position and navigation solutions in real time. However, the current state-of-the-art integrated system is not necessarily a simple, robust, ready to use and profitable solution for all commercial application segments – i.e. compromises in terms of accuracy, precision and applications which can be served with GNSS technology at commercially-sustainable market pricing are typically required.

By around 2014, Galileo will approach its initial operational capability, China's Compass will start the deployment of its final MEO satellites, GLONASS will be fully operating, and L2C and L5 signals of GPS will be available. These multi-constellations and new signals will certainly improve the overall performance of GNSS positioning and navigation. However, the same problems for positioning under dense canopy, crowded urban environments or in deep pits will still exist and multipath will remain a major constraint on the provision of precise and real-time position solutions. Furthermore, the resultant degraded vertical positioning accuracy will not be improved significantly if a GNSS-only system is utilized for machine control.

This paper proposes an FPGA Based hybrid Positioning Platform for construction machine control. This platform consists of multiple sensor based positioning, communication and a data processing modules. Once the system is implemented it will be tested through use of a GNSS simulator-Spirent GSS 8000 0 at the Nottingham Geospatial Institute (NGI) in the University of Nottingham and its performance will be fully assessed against reliability, continuity, accuracy and integrity parameters, before a series of in-situ tests.

The paper is organized as follows: the current issues for machine control are discussed first in Section 2. Section 4 and 4 briefly review precise attitude determination and state-of-the-art FPGA technology, respectively. The proposed FPGA based hybrid positioning platform is described in Section 5. The conclusions are summarised in Section 6.

2 CURRENT ISSUES FOR CONSTRUCTION MACHINE CONTROL AND MULTIPLE GNSS CONSTELLATIONS

2.1 Vertical Accuracy

It is well documented that GNSS generally provides a higher degree of horizontal accuracy than vertical. For most construction applications employing motorgraders, bulldozers or excavators, an accuracy of 1-3 cm in height and 2-5 cm for 3D position at a velocity of 12-35 km/h is sufficient 0. These accuracies are achievable using GPS or a total station (tachymeter). A higher accuracy is required for paving applications, typically better than 5 cm at a velocity of 0.2 – 0.6 km/h 0. Such a vertical accuracy requirement could be achieved from using long-term static measurements, but is considered highly challenging for kinematic positioning techniques.

2.2 Multipath and Lack of Signals

GNSS signal quality in multipath-prone construction environments is significantly degraded due to signal attenuation and fading effects. Signal attenuation is caused by surroundings where the signals are either refracted or reflected, introducing total attenuation from 0 to 30 dB, or even more, with respect to line-of-sight signals (LOS) 0. In an open environment the direct path would be best represented by the line-of-sight transmission of the signal. However, in an urban environment, this LOS signal is often either polluted by multipath or completely obstructed. By their nature, construction sites often congested and most are located in built-up areas, so mitigating multipath and increasing signal availability are two major issues that need to be resolved in a practical embodiment of a GNSS-based machine control system.

2.3 Operating Speed and Vibration

In construction operations, contractors seek maximum productivity. Machine operators must complete work tasks efficiently and without compromising the quality of the surfaces produced. However, when the machine is "working" (i.e. pushing dirt, digging etc.), this is usually at a much slower speed than the maximum travel speed. For example, a CAT 836H landfill compactor can travel in gear one at 6.4 km/h and gear two at 11.4 km/h 0. When it works in a landfill site, it can probably only work in gear one at the lower speed because of the weight of the pushed materials, as well as safety considerations.

Shock and vibration is a further challenge for reliable machine control applications. **Figure 78** shows a bulldozer equipped with an example of the current state-of-the-art GNSS machine control - in this case a Leica PowerBlade 3D GNSS system. It can be seen that the GNSS antenna is installed on a tall mast, well above cab height, due to the need for "clear sky" and to be located well away from the hazardous working area.

This is a simple, practical but inelegant solution, and creates an additional problem for the GNSS positioning - the flexibility of the mast, coupled with shocks by hitting obstructions, or starting/stopping the machine typically introduces some "whip" effect to the antenna. Clearly, the mast length (cantilever), tube diameter, material characteristics, and amplitude of shock, all contribute to the actual amount of whip. Bulldozers and motorgraders are considered problematic in this respect.

Figure 78: The Bulldozer with Leica GNSS Technology

Furthermore, it must be considered in closed-loop control applications, constraints exist on the resultant accuracy and quality of graded surfaces when machine travel speed, GNSS update rate and accuracy are coupled with the varying responsiveness of machine hydraulic systems (hysteresis, latency) and the machine's kinematics (i.e. working tool location with respect to the prime mover, wheelbase length, centre of gravity etc).

Ignoring the additional influence of the characteristics of the material being worked (soil type, homogeneity, particle size distribution, density, moisture content, granularity/cohesiveness etc), put simply, the faster the machine is required to work (to suit commercial imperatives) the greater the risk a GNSS-only closed-loop control system is unable to provide the desired performance in the widest range of use cases.

2.4 Multi-constellation GNSS and new GNSS signals

GNSS is heading towards a multi-constellation and multi-frequency world, in which new L2C and L5 signals will be available for positioning, together with exciting perspectives of the new global systems Galileo and Compass, which are getting closer to reality.
Regional augmentation systems such as EGNOS, QZSS or IRNSS are also on the starting line 0, and previously unobtainable accuracy through single-frequency RTK techniques are emerging.

There will be 120 navigational satellites operating concurrently when Compass and Galileo reach full operation capability, currently estimated around 2020. As a result, the number of visible satellites can reach 30 at any moment. Thus the positioning accuracy will be improved greatly 0.

The obvious advantages offered by multi-constellation and new signals are higher reliability and higher accuracy, increasing the coverage of space and time to realize continuous navigation, etc. However, development of a cost-effective multi-constellation GNSS receiver that makes full use of new L2C and L5 GPS signals and other signals of opportunity still requires significant and timely research and development, to significantly benefit machine control applications.

3 ATTITUDE DETERMINATION

Attitude determination is the representation of a three-axis body frame, such as working tool, rigidly attached to the vehicle's prime mover, with respect to a reference frame, typically a local level frame 0. Traditionally this has been described by using roll, pitch and yaw to signify the rotation about each axis in a right hand system. For a road vehicle, roll represents a rotation around the front-rear axis, i.e. the "lean" experienced whilst driving round a corner. Pitch represents a rotation around the transverse axis, describing the rotations experienced under acceleration or deceleration. Yaw represents a rotation about the vertical axis, which is perpendicular to both the roll and pitch axes, and is synonymous with the vehicle's heading (or in some cases bearing).

Previous to the development of the Global Navigation Satellite Systems (GNSS), instruments used to determine the attitude of a vehicle included magnetometers, compasses, tilt sensors and high grade gyroscopes. GPS was first proposed for use as a solution to attitude determination in 1976 0. The main focus of development was in the aerospace industry, as the attitude of an aerial vehicle is critical to its control. Today, attitude determination is applied to marine and land vehicles, as well as used in ballistic missile guidance and deformation monitoring.

Traditional methods of attitude determination use Inertial Navigation Systems (INS), typically a combination of accelerometers, compasses and gyroscopes. Many of the mechanical instruments can be influenced by magnetic disturbance of nearby objects or other instruments. Some instruments may be slow or sensitive during initialisation. Fundamentally, however the most accurate instruments tend to be prohibitively expensive, and only adopted for use in the largest aerospace and marine vehicles. As GNSS equipment used for attitude determination does not suffer from these disadvantages, systems integrating both GNSS technology and traditional sensors are thus very common for attitude determination.

Several different algorithms can be used for integrating GNSS and IMU (Inertial Measurement Unit) information together to determine the vehicle attitude and obtain an accurate and reliable position solution. The most common integration algorithms are loosely-coupled closely-coupled or tightly-coupled Kalman Filters. The principal difference between these is the level of information to be integrated. Specifically, a loosely coupled GNSS/INS integration combines inertial measurements with position and velocity measurements calculated by the GNSS receiver, which will then fed back to the inertial processor that calibrates the IMU to remove effects from biases, scale factors and/or misalignment. A closely-coupled integration combines the inertial measurements with range information to the GNSS satellites provided by the GNSS receiver, such as pseudo range and carrier phase. A tightly-coupled integration uses the inertial measurement to aid the GNSS receiver's tracking loops.

The loosely-coupled integration can be implemented based on any commercial GNSS receivers and IMU sensors, so it is a cost-effective solution. However, the main disadvantages of this approach include that it requires tracking a minimum of four GNSS satellites and the performance is largely dependent upon the position accuracy of the GNSS receiver. For example, the overall performance will be significantly affected by cycle slips, multipath errors, and etc. For the closely-coupled integration, it could be still operational even less than four satellites tracked because this approach is based on the range information rather than the navigation solution. The major drawback of this method is that the size of the range measurement vector is considerably large and changes with the number of the satellites tracked. Hence, higher computational resources are required. The tightly-coupled integration is the most complex one because the integration occurs at the GNSS tracking loop. Therefore, it can significantly improve the GNSS tracking performance and hence the performance of the GNSS receiver and the overall system 0.

A multi-antenna GNSS configuration has been chosen for the system because it can produce a high accuracy attitude solution. The advantages of a GNSS only system are the flexibility in the selection of GNSS sensors, the cost-effectiveness of the system with the emergence of low-cost high performance GNSS products and the increased productivity of the equipment that can be assembled into a multi-antenna system or dissembled for other positioning purposes. Furthermore, it has the potential to avoid using to some extent, some sophisticated and expensive attitude sensors, such as Inertial Navigation Systems (INS).

4 FPGA TECHNOLOGY

It is known that most modern GNSS receivers are implemented using ASIC (Application Specific Integrated Circuit) devices, which are bespoke integrated circuits designed and optimized for the specific task of GNSS signal correlation. However, the cost of Non-Recurring Engineering (NRE) for ASICs has been so high that only big corporations can afford to create such chips 0.

Field Programmable Gate Arrays (FPGAs) are essentially very large Application-Specific Integrated Circuits (ASICs), which can be programmed to implement algorithms for digital signal processing with the unique feature of arbitrary parallel processing capability. After a development period of around two decades, they are now capable of supporting the implementation of bespoke signal-processing functions for an entire GNSS receiver.

Moreover, an FPGA-based platform is highly reconfigurable because it can include or exclude any sensors as required for a particular application, but importantly, the nature of the FPGA technology also allows reconfiguration of the entire implementation without any change to the physical printed circuit board (PCB) design. For example, some applications may not need network RTK corrections. In that case, one can simply exclude the GPRS/3G Transceiver and re-configure the FPGA with a design without Network RTK.

Therefore, FPGAs are considered highly attractive for implementation of novel GNSS signal processing techniques to achieve real-time performance, as both a rapid prototyping platform, and a path to cost-effective production variants. However, the real-time performance from FPGA technology comes at a cost, as more FPGA logics are required for the implementation of parallel processing to achieve the real-time performance, which means that use of larger and more expensive FPGA devices is essential. Therefore, we need to use hardware/software co-design to identify which functional modules should be implemented on FPGA logics and which can be run on the on-chip processor to achieve an optimal performance for a particular application. For example, considering Kalman filtering in our project, we could use an on-chip processor for initialization, and implementation of the Kalman filter loop on FPGA logics because filter initialization involves inverse matrix calculation, which requires significant FPGA logics, whereas Kalman filter loop only involves matrix addition and multiplication, which can be easily paralleled with relative low FPGA logics. Furthermore, the processing speed of the Kalman filter loop is much more critical than the initialization, in terms of real-time performance.

FPGA technology will thus be used for developing this system. A Xilinx Virtex-6 FPAG ML605 evaluation board has been evaluated and selected to be responsible for the signal processing and the signal interfaces required by all the sensors 0.

5 DESCRIPTION OF THE MODULAR DESIGN OF THE PLATFORM

The proposed FPGA-based hybrid positioning platform is shown in **Figure 79**, in which the modules enclosed by the broken line are implemented in a single Xilinx Virtex-6 FPGA device. The platform also includes three NovAtel multi-constellations, multi-frequency GNSS receiver modules 0, one u-blox GSM/GPRS module and one accelerometer 0 0.

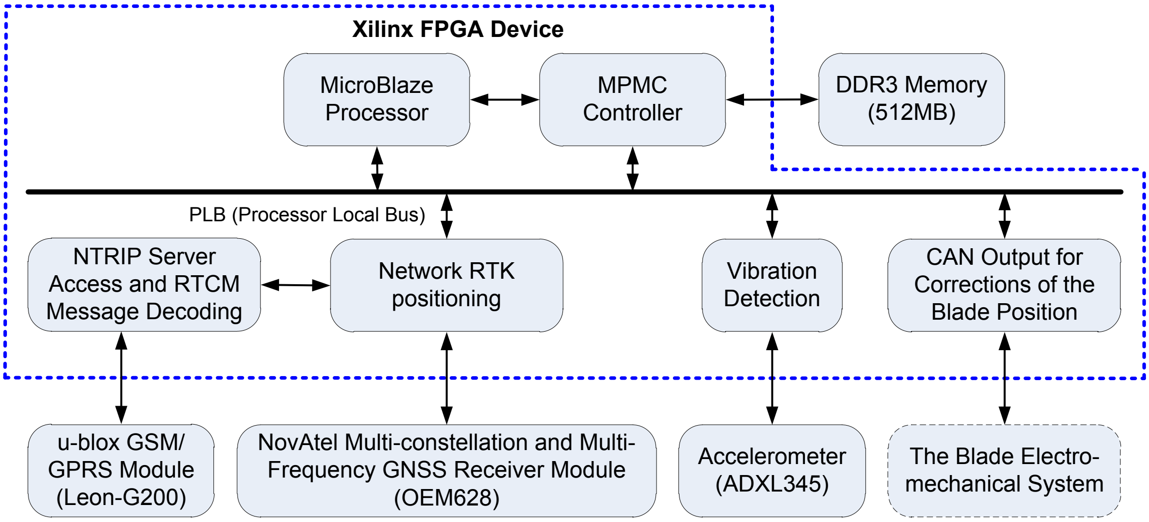


Figure 79: The Positioning Platform based on an FPGA Embedded System

In the platform, the GNSS receiver modules will supply multi-frequency GNSS code and phase solutions, the u-blox module will receive the GNSS correction messages (in this case, Radio Technical Commission for the Maritime Services (RTCM) protocol is proposed). Network corrections delivering over Internet Protocol (NTRIP) are to be implemented and the accelerometer will be integrated into the system to assist with mitigating vibration effects. Kalman filtering will play an important role to integrate the inputs for all the sensors to control the real-time correction process of the construction machine positioning.

It is assumed that all modules except the Kalman filtering based processing module are required to be implemented as hardware IP cores and only one MicroBlaze processor core is required to implement the Kalman filtering processing module, configure the relevant parameters for each hardware IP core and receive the results from the hardware IP cores. In that case, it can achieve high performance, in terms of processing speed, but will require high hardware resource usage. In order to offset the processing speed for the hardware resource usage, some modules can be implemented by software to be run on the MicroBlaze processor, such as network RTK positioning and vibration detection. Especially, in the course of development, the implementation will start from running all modules in MicroBlaze processor as software modules and then continue to migrate them to the corresponding hardware cores one by one.

The entire system development consists of two parts, which are described in the following section.

5.1 GNSS Receiver with Network RTK Solutions

This part is to define the cabin roof position and attitude by using GNSS positioning technology with three antennas.

In order to achieve high accuracy in the system, the multi-constellation and multi-frequency GNSS receiver is adopted. Compared with a single frequency GPS receiver, multi-frequency receiver is able to provide more accurate position solution since the ionospheric error can be mostly cancelled out with two frequency signals from the same satellite. Multi-constellation obviously will provide more satellites for positioning. Therefore, the multi-frequency and multi-constellation system will be more reliable and the positioning solution will be more accurate than a single constellation system.

This part will include the following steps:

1. Set up network connections through mobile phone network by using GPRS (General Packet Radio Service). RTCM messages will be received from the network reference station via an NTRIP server and will be decoded for network RTK positioning.
2. Integrate multi-frequency GNSS code and phase solutions with the network RTK correction messages to obtain even more accurate machine position.
3. Integrate multi-frequency and multi-constellation code and phase solutions with the network RTK correction messages to further improve the accuracy of the machine position.

5.2 Accelerometer Integration

An accelerometer is a device that measures weight per unit of test mass, a quantity also known as g-force. For example, the force caused by vibration or a change in motion (acceleration) causes the mass to push the sensors which will calculate the measurement of acceleration according to the force. Then the measurement of acceleration will be used as an input for application systems. As a result, the measured acceleration will be used to define the dynamic conditions, which is required by the Kalman filter 0.

Accelerometers are usually a popular choice for machinery vibration monitoring. The transducer is typically attached to the outer surface of the prime mover, so it can provide continuous or periodic sensing of absolute case motion (vibration relative to free space), in terms of acceleration 0.In this stage, an accelerometer will be adopted to measure the acceleration rates and send them to the FPGA for further processing.

An accelerometer is adopted in the system to efficiently detect the vibration effect. It can be seen from Eq. 1, an accelerometer can be used to cancel out the vibration effect by integrating acceleration to velocity and further integration to give position movement.

$$v = \int a * dt\,, \qquad s = \int v * dt \qquad [1]$$

In other words, the movement that caused by vibration can be calculated from the sensed accelerometer measurements caused by the machine’s vibration. It is anticipated that, in a further refinement of the basic embodiment, provision will be made to support multiple accelerometers - installed on both the prime over and the working tool subframe to improve vibration and shock mitigation.

6 CONCLUSIONS

In this paper, an FPGA based hybrid positioning platform has been proposed for machine control applications, to address the drawbacks in the current machine control systems, which include low overall GNSS positioning performance and particularly a lower vertical accuracy, multipath, signal availability and vibration. In this platform, three GNSS receivers, one accelerometer and a GPRS/3G communication module are connected to the Xilinx Virtex-6 FPGA board directly. In terms of functionalities, these three multi-constellation and multi-frequency GNSS receivers are working together with the GPRS/3G communication module to conduct real-time kinematic (RTK) positioning with an attempt to improve the vertical positioning accuracy to be better than 5 cm; these three GNSS receivers are also used to facilitate the instant attitude determination for construction machine control; the accelerometer is used to cancel out the random vibration of the construction machine. The authors of this paper presented a proposal on sitting GNSS antennas away from working tools to the prime mover to mitigate both damage hazards and whipping effect caused by the high antenna pole. From our research, both the location of the machine cabin roof and attitude could be precisely determined with three GNSS antennas. We are currently working on hardware/software co-design for the platform and evaluating sensing technologies for measuring the position and attitude of the working tool with respect to, and in the same coordinate system as, of the prime mover.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support provided by the UK's Engineering and Physical Sciences Research Council (EPSRC) and Hexagon Machine Control Division.

REFERENCES

Books:

ALEXANDER BEETZ and VOLKER SCHWIEGER: *Integration of Controllers for Filter Algorithms for Construction Machine Guidance, in the proceeding of the 1st International Conference on Machine Control and Guidance 2008*, p.263-272, 2008

SEYED NIMA SADRIEH, ALI BROUMANDAN, and GERARD LACHAPELLE: *Spatial/Temporal Characterization of the GNSS Multipath Fading Channels*, in the proceeding of the 23rd ION GNSS 2010 Conference,p.393-401, 2010

B. BONET, I. AlCANTARILLA, D. FLAMENT, C. RODRIGUEZ, and N. ZARRAOA: *The benefits of multi-constellation GNSS: Reaching up even to single constellation GNSS users*, in the proceeding of the 22nd ION GNSS 2009 Conference, p.1268-80, 2009

ZHANG MIAOYUAN, HANG JUN, and QIN YONG: *Satellite Selection for Multi-constellation*, IEEE/ION Symposium on Position, Location and Navigation, 5/2008, p.1053-59, 2008.

A. CHO and H. NO: *The performance improvement of a low cost GPS/INS integration system using a single-antenna GPS based attitude*, in Proceedings of the 23rd International Technical Meeting of the Satellite Division of The Institute of Navigation, Portland, OR, United States, p. 1498-1504, 2010.

V.W.SPINNEY: *Application of the Global Positioning Wystem as an Attitude Reference for Near Earth Users*, in The 1976 National Aerospace Symposium "New Frontiers in Aerospace Navigation", Naval Air Development Centre, Warminster, PA, 1976.

D.M.BEVLY, S.COBB: *GNSS for Vehicle Control*, Artch House, 2010.

H. STEWART COBB: *FPGA-Optimized GNSS Receiver Implementation Techniques*, in the proceeding of the 21st ION GNSS Conference, p. 2326-31.

Links:

http://www.uk.cat.com, "Products and applications, 836H Landfill compactor"

http://www.spirent.com, "GSS8000 Multi-GNSS Constellation Simulator"

http://www.xilinx.com, last accessed on October 12, 2010, "Virtex-6 FPGA ML605 Hardware User Guide", UG534.

http://www.novatel.com/assets/Documents/Papers/OEM6_Development_Kit.pdf, NovAtel OEM 628 reciever,.

http://www.u-blox.com/en/evaluation-tools-a-software/wireless-evaluation-kits/evk-g26h.html, U-blox EVK-G26H.

http://www.analog.com/static/imported-files/data_sheets/ADXL345.pdf, ADXL 345 Accelerometer.

http://en.wikipedia.org/wiki/Accelerometer, Accelerometer.

http://www.stiweb.com/appnotes/accel.htm, STI Vibration Monitoring Inc.

Cost Efficient Camera System to Acquire Data for Pavement Maintenance Management System (PMMS)

AL-Mistarehi Bara' [1], Volker Schwieger[1], Mohammed Obaidat[2]
[1]Intitute of Engineering Geodesy, University of Stuttgart, Germany
[2]Institute of Civil Engineering, Jordan University of Science and Technology, Jordan

Abstract
The main objective of this research work was to investigate the potential of a cost-efficient manual camera system for the purpose of flexible pavement, distresses classifications and maintenance priorities. The classification process includes distress type, distress severity level, and distress quantity. Many traditional systems were used to collect pavement surface distresses information. These old techniques are characterized by manual data collection, time consuming, and the lack of archiving capabilities. In this study, a new methodology for collecting, manipulating, retrieving and archiving pavement condition data inventory will be presented. It is anticipated to use digital images (cost-efficient manual camera system), instead of using manual hand odometer method for data collection. This system will give distresses data collection potential. However, it was anticipated to use paver pavement management system, which provides a systematic and consistent method for selecting maintenance and repair (M&R) needs, priorities, and determining the optimal time of repair by predicting PCI value for every pavement section, and future pavement condition.

Obviously, the validation process is deviated by 3.1% (relative RMS) for alligator cracking, and by 2.6% (relative RMS) for longitudinal and transverse cracking due to different given accuracies of the used instruments in the measurement mode using this system. Therefore, low cost-efficient manual camera system has been used in this research work for the purpose of data collection.

Keywords
Cost-efficient, manual camera system, paver, PMMS, pavement distresses, distresses classification, PCI.

1 INTRODUCTION AND BASICS

1.1 Introduction

The main objective of this paper was to present the feasibility to use a simple low cost-efficient camera system in order to build digital distress information linked to locations, useful surface information, and pavement conditions databases in order to collect and analyze different distresses data. Arterials (primary roads for relatively high volumes of traffic), collectors (secondary roads that provide access to higher type roads), and locals roads (land access roads that don't serve through traffic) of Irbid-Jordan city were taken for this prototype study.

1.2 Classification of basic pavement distress

Pavement distress information is needed to assess maintenance requirements .The distresses of asphalt concrete pavement are any defects or deterioration in the pavement, and it can be grouped into the following general categories (SHAHIN 1998):
1. Cracking
2. Distortion
3. Disintegration
4. Skid hazard
5. Lane/Shoulder drop-off.

In the following more details regarding the distress groups is given.

1. Cracking: the following are the major crack types:

a. Alligator cracks: alligator or fatigue cracking is a series of interconnecting cracks caused by fatigue failure of the asphalt concrete surface under repeated traffic loading.

b. Edge cracks: cracks parallel to and usually within 0.3-0.6 m of the outer edge of pavement.

c. Joint reflection cracking: cracks occur only on asphalt surfaced pavement that lay over slab.

d. Longitudinal and transverse cracking: longitudinal cracks are parallel to the pavement centreline or lay down direction and transverse cracks extend across the pavement at approximately right angle to the pavement centreline

e. Block cracking: interconnected cracks that divide the pavement into rectangular pieces range in size from approx. $(0.3m)^2$ to $(3m)^2$.

f. Slippage cracking: a crescent or half-moon shaped cracks, which are produced when braking or turning wheels cause the pavement to slide or deform.

2. Distortion: this type of distresses caused mainly because of subgrade weakness or poor construction like poor compaction. The following are the main distortion distresses:

a. Rutting: a surface depression in the wheel paths noticeable only after a rainfall as a result of permanent deformation in any of pavement layers or subgrades.

b. Bumps and sags: bumps are small, localized, upward displacement of pavement surface & sags are small, abrupt downward displacement of pavement surface.

c. Corrugation: a series of closely spaced ridges and valleys occurring at fairly regular intervals usually less than 3m along the pavement.

d. Shoving: a permanent, longitudinal displacement of a localized area of the pavement surface when traffic pushes against the pavement usually occurred near intersections.

e. Depression: localized pavement surface areas with elevations lower than surrounding pavement.

f. Swell: upward bulge in the pavement surface along, gradual wave longer than 3m that can be accompanied by surface cracking.

g. Patching and utility cut patching: an area of pavement that has been replaced with new materials to repair existing pavement or installation of underground utilities.

h. Railroad crossing: defects that are depressions or bumps around the and/or between tracks.

3. Disintegration: it is the case at which the pavement surface is fragmented into small, loosely pieces due to breaking up of pavement. Two types of disintegration are defined:

a. Potholes: small (usually less than 0.9 m in diameter) bowl-shaped depressions in the pavement surface with sharp edges near the top of the hole.

b. Weathering and ravelling: wearing away of the pavement surface due to a loss of asphalt or tar binder and dislodged aggregate particles.

4. Skid hazard: two types of distresses under this category can be defined:

a. Bleeding: a film of bituminous material on the pavement creates a shiny, glasslike reflecting surface that usually becomes quite sticky.

b. Polished aggregate: the aggregate in the surface becomes smooth to the touch, adhesion with vehicle tires is considerably reduced and the portion of aggregate extending above the surface is small

5. Lane/Shoulder drop-Off: a difference in elevation between pavement edge and shoulder.

1.3 Basic of pavement condition index (PCI)

The first step in selecting the best maintenance and repair needs is to identify the condition of pavement. The most practical and easy way to predict the pavement condition, is PAVER system, which is a pavement management system (PMS) that can be used to manage roads, streets, parking lots and airfield pavement. This system was developed by U.S. army Construction Engineering Research Laboratory (CERL). PAVER uses pavement condition index (PCI) as a measure of pavement condition which is an objective, repeatable rating system for identifying the present condition of the pavement. Calculation of distress types, severity levels, and quantity, pavement condition can be represented in pavement distress index (PDI) which can be used for management purpose (THEODORAKOPOULOS et al. 2007). This index depends on individual distress ratings along severity-extent combinations (UDDIN 2006). A knowledge-based system has been developed based on PAVER system in order to automate distress classification using data extracted from personal computer (PC) vision system (GHUZLAN 1995).

The most useful feature of an effective PMS is the ability to both determine the current condition of a pavement network and predict pavement condition some time into the future. The simplest method to carry out this task is the PAVER system. The PAVER System uses the PCI, a numerical index ranging from 0 for a failed pavement to 100 for a pavement in perfect condition, as its pavement condition rating (SHAHIN et al. 2006). The PCI is calculated based on the results of a visual condition survey in which distress type, severity, and quantity are defined. Field verification of the PCI inspection method has shown that the index gives a good indication of a pavement structural integrity and operational condition (UDDIN 2006).

2 DATA ACQUSITION

2.1 Manual data acquisition method

Old hand odometer data collection system could have tedious data collection procedures for the implementation of pavement management system. Moreover, it requires professional expertise from end-users. It also doesn't have the capability of digital and spatial mapping as well as scheduling capabilities. Thus, the situation of pavement management system would be rather complicated especially for big cites of huge roads' networks (HOWE et al. 2008). The most important factors that influence the pavement condition rating manual was the crack width. Bleeding severity levels depend on the number of days from the whole day that the pavement sticks to shoes or vehicles, polished aggregate have no severity levels. Whereas depression and rutting severity levels are related to the distress depth maximum or mean value. The degree of wearing away of the pavement surface is the most importance factor that describes the severity level of weathering and raveling. For potholes severity, depth and diameter are the control factors (GHUZLAN 1995). These distresses types could be easily defined and classified manually (SHAHIN et al. 2006). On the other hand, a procedure should be developed in order to identify and classify them automatically in real-time.

2.2 Automatic data acquisition method

Several automated distress data collection are now available throughout the world. Japan has developed image collection device with on-based processing capability using laser scanners, this vehicle is used to automatically check a road surface for its condition, seeking out any signs of cracking, wheel rutting and longitudinal unevenness (SHUBINSKY 2000). A photogrammetric system that depends on a video camera was developed by SMITH (2005) to route cracks and sealing it using computer software, which save money and improve safety. A developed system which allows the identification, classification, and quantification of commonly occurring pavement distress types in terms of severity and extent. It must be noted that the identification and classification of distress is limited to distress type that can be quantified by length, geometry or area covered by these distresses (WAUGH 2007).

Other studies that discussed field data acquisition technologies were developed. A knowledge-based system which incorporates a set of if-then-rules was developed. The system suggested three stages of distresses classification: global distresses classification, severity level, and options for repair. The developed knowledge-based system used PAVER as input to construct the knowledge-base. New parameters were introduced in order to facilitate the distresses classification process including: Shape parameters, orientation, and geometric measurements. A software package called K-PAVER was developed which incorporates stereo vision input with the knowledge-based system. The system showed a potential to automatically classify the distress types, minor measurements differences were found between surface measurements extracted from stereo vision and actual measurements. These differences were mainly dependent on stereo image configuration, scale, and image resolution. The developed system is proven to be reliable, ease-of-use, and automate much of the routine functions of distresses classification (GHUZLAN 1995).

Finally it can be said that low cost video image pavement distress analysis is possible and it is within the reach of implementation with currently available equipment and within a short time (GHUZLAN 1995).

Using of photogrammetric methodology was anticipated to open the door to fully low cost technology applications for distresses data collection and pavement surface road conditions, mapping, classification, prediction and analysis. This technology becomes widely popular due to its effectiveness in carrying out different research activities economically and safely (JASELSKIS 2005). Further, researchers are capable to perform real-time operations, extract highly accurate data, present spatially inventory data, introduce numerous analytical techniques and develop highly technological systems. Maintenance and operation engineers were also anticipated to use the findings and guidelines of this research work to automate most of their routine decision making activities (Al-MESTAREHI 2011).

2.3 Cost efficient manual camera system method

Images captured by a Nikon D70s 10 Megapixel digital camera are used in this research work. Nikon D70s is an easy-to-use, compact digital camera with a large range of functions incorporated into a card-sized body. For all photographs used in measurement, the focal length was fixed at 28 mm, 50 mm, 105 mm and the focus at infinity. Camera calibration was not carried out, because image location is closely controlled through the use of global positioning, with geo-referenced position 3D accuracy of the measurement mode was about (2cm to 3cm). The Nikon D70s has a 23.7mm by 15.6mm sensor with 3008x2000 pixels resulting in a pixel size of approximately 8µm; the camera is synchronized and maintains a precision of 2 m. **Figure 80** shows a Nikon D70s 10 Megapixel camera.

Figure 80: Nikon D70s 10 Megapixel camera (AL-MESTAREHI 2011)

3 PCI COMPUTATION AND REALISATION

3.1 General method (Paver procedure)

The most useful feature of an effective PMS is the ability to both determine the current condition of a pavement network and predict pavement condition some time into the future. To predict condition reliably, an objective, repeatable rating system for identifying the pavement’s present condition must be used. This PMS is called the PAVER system, where Micro PAVER system uses pavement condition index (PCI), as pavement condition rating. The selected roads are divided into branches which are a single entity and have a distinct function. The selected branches that are divided into smaller components called sections. The following factors were considered when dividing branches into sections:

1. Pavement structure: the structural composition (thickness and materials).
2. Traffic: the volume and intensity of traffic.
3. Construction history: the pavement sections should have the same construction history.
4. Pavement rank: the functional classification (arterial, collector, local).

5. Drainage facilities: the drainage facilities and shoulders should be consistent throughout the pavement section.

The selected pavement sections are divided into sample units with an area of 233 ± 93 m^2 (SHAHIN et al. 2006). The minimum number of sample units to be surveyed is determined based on the total number of sample units and the PCI standard deviation which is assumed to be 10 for asphalt surfaced pavements (LEE et al. 2001). **Figure 81** shows the minimum number of sample units to be surveyed.

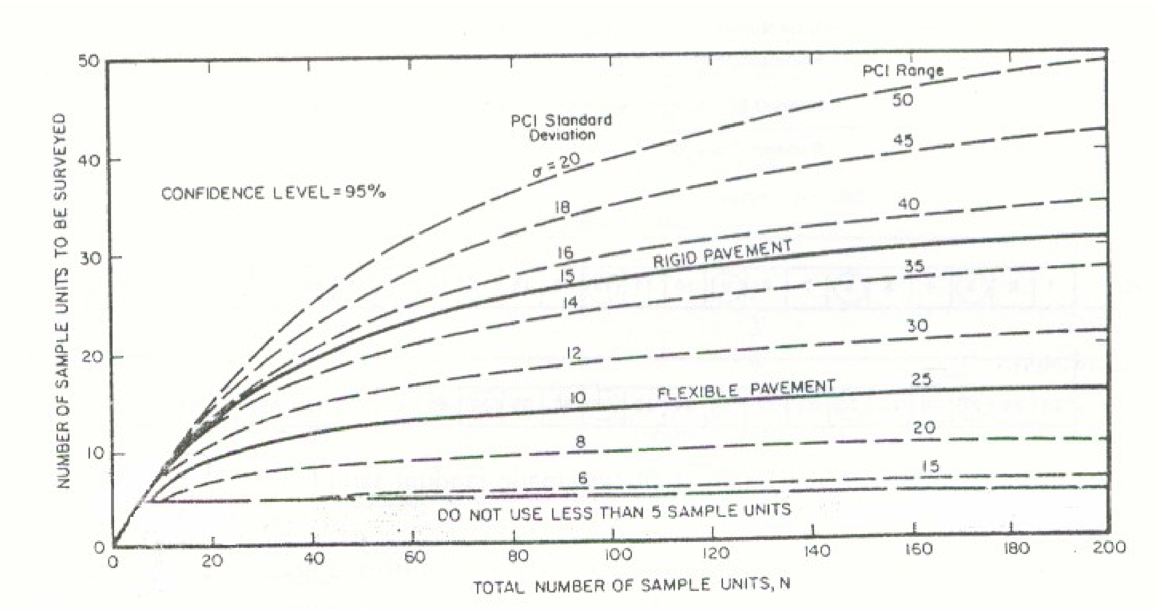


Figure 81: **Selection of the minimum number of sample units (SHAHIN 2007)**

Different instruments (such as digital camera, laser scanner sensor, hand odometer, etc.) are used to measure distress length, area, and the depth of ruts or depression. The distresses inspection is conducted by walking over the sample unit, measuring the distress type and severity according to PCI distress manual, and recording the data on the flexible pavement survey sheet. One data sheet is used for each sample unit through field inspection procedure.

3.2 PAVER software usage for PCI determination

The following steps were used to compute PCI:

1. Defining the pavement inventory (network, branches, or sections).
2. Entering the edit inspection dates and edit sample units information (sample number, area, and type of sample units (here: random)).
3. Entering information on distress (type, severity, and quantity)
4. PCI computation displays five tabs as following:
   a. All indices: this tab displays the condition value for all conditions associated with the current section.
   b. Individual distresses: this tab displays distress information for every distress entry in every sample of the current inspection date. Added to the information is the density of that distress and the corresponding deduct value (value depends on the distress density and severity, and must be adjusted to calculate PCI).
   c. Extrapolated distresses: this table shows each distress type (all identical distress types and severity levels from the previous tab are grouped together here), with the quantities adjusted to reflect the extrapolated value. With random samples, paver extrapolates the distress quantity across the entire section.
   d. Sample/Distress summaries: the left side of this tab displays the number of samples surveyed and compares the total number of samples to the recommended number for a project level inspection. The right side of this tab groups all distresses by type and calculates the corresponding

percentages.

e. Sample info: general data about each sample involved in the inspection is included here. **Figure 82** shows pavement condition results, where in this figure for example PCI is 90 and the section area is 1.605 m^2 with repeatability accuracy of 10.32 m^2 which is quite precise, and means that the rating condition of this pavement section is good.

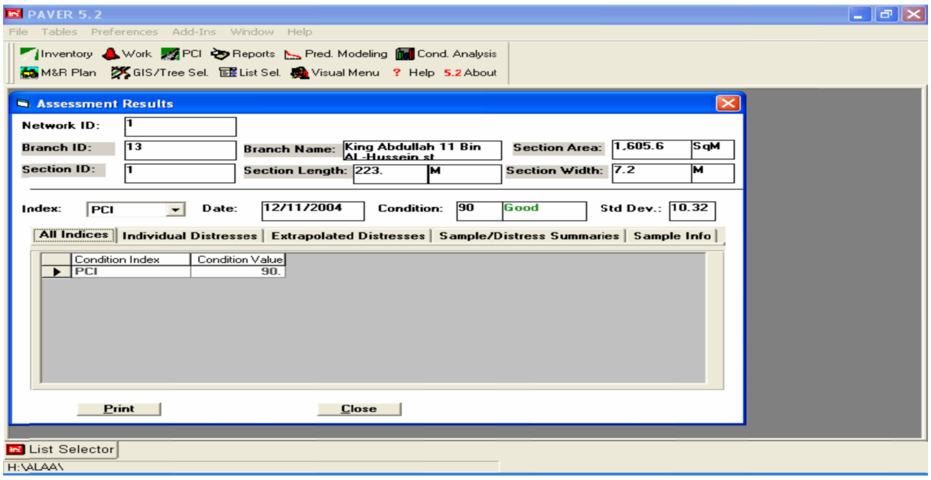


Figure 82: Pavement condition results (AL-MESTAREHI 2011)

4 CASE STUDY OF THIS REASEARH WORK

4.1 Case study description

To achieve the objectives of this study, an integrated database related to 35 arterials, 24 collectors, 31 locals of Irbid-Jordan city were developed. The selection criterion was dependent on covering variables having different pavement and traffic conditions. However, **Figure 83** shows a view of the selected branches in the study area.

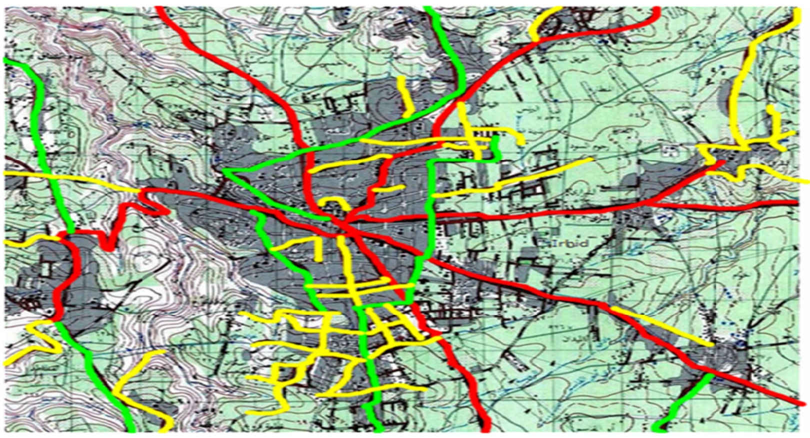


Figure 83: The selected branches (marked in legend) in the study area (AL-MESTAREHI 2011)

4.2 Case study data acquisition

Images captured by digital cameras were used in this new methodology. Mainly, the quantitative measurements were extracted for the longitudinal and transverse cracks and alligator cracking, while, manual procedures were used for the rest of distresses. **Table 11** shows that the cost efficient manual camera system analysis was used too to extract quantitative surface measurements. Both procedures are described in the following.

Table 11: Distresses Measurements methods

Distress#	Distress Type	Measurement Unit	Measurement Method
1	Alligator Cracking	Square Meter	Camera Analysis
2	Bleeding	Square Meter	Manually
3	Bumps & Sags	Meter	Manually
4	Corrugation	Square Meter	Manually
5	Depression	Square Meter	Manually
6	Edge Cracking	Meter	Manually
7	Lane/Shoulder Drop-off	Meter	Manually
8	Longitudinal & Transverse Cracking	Meter	Camera Analysis
9	Patching & Utility Cut Patching	Square Meter	Manually
10	Polished Aggregate	Square Meter	Manually
11	Potholes	Number	Manually
12	Rutting		Manually
13	Shoving	Square Meter	Manually
14	Slippage Cracking	Square Meter	Manually
15	Swell	Square Meter	Manually
16	Weathering & Raveling	Square Meter	Manually
17	Joint Reflection Cracking	Square Meter	Manually
18	Block Cracking	Square Meter	Manually
19	Rail Road Crossing	Square Meter	Manually

The following methodology was used to achieve low cost efficient manual camera procedure of distresses measurement that can be summarized in the following steps (Al-MESTAREHI 2011):

1. Define distresses area.
2. Capture digital image for distress area using a Nikon D70s 10 Megapixel camera (cf. section 2.3).
3. Use camera optical axis perpendicular to mapped area.
4. Use at least two reference points of known distance in between for scale purposes.
5. Scale images using the control points.
6. Measure lengths and areas of distresses.
7. Use extracted surface images for pavement condition evaluation.

While, the manual methodology was used for the rest of distresses, and could be summarized in the following steps (Al-MESTAREHI 2011):

1. Paver procedure (cf. section 3.1). Therefore, an integrated database related to 35 arterials, 24 collectors, 31 locals of Irbid-Jordan city is developed, using hand odometer with a given accuracy of 0.02%. This means that for a distance of 2 km, we have an accuracy of 40 cm.
2. Calculation of PCI (cf. section 3.2). Therefore, PCI values related to 35 arterials, 24 collectors, 31 locals of Irbid-Jordan city are developed.

4.3 Data analysis and interpretation

For the purpose of checking the precision and accuracy of the collected distress data that include distress type, distress quantity and severity level of the observed distresses, a sample of 44 pavement condition images distributed over the surveyed roads were used. **Table 12** shows the validation procedures results.

This analysis procedure contains the following information:

1. Types of distresses visible through images associated with their severity levels.

2. Distresses quantity obtained through camera analysis in m, m² or number units.
3. Distresses quantity obtained manually in m, m² or number units
4. Difference between manual and camera collected distress quantities:

$$\Delta q \equiv \left| q_{manual} - q_{camera} \right| \tag{4.1}$$

Δq ....Absolute value of difference between manual and camera collected distress quantities [m] or [m^2].

q_{camera} … camera based quantity [m] or [m^2].

q_{manual} ….manual based quantity [m] or [m^2].

5. Difference percentages B [%] of the camera collected data with respect to manual data:

$$B = \left[\frac{\Delta q}{q_{manual}} \right] \times 100\% \tag{4.2}$$

6. Absolute average difference [AAB] of the camera collected data in [m] or [m^2]:

$$AAB \equiv \left| \frac{\sum \Delta q}{N} \right| \tag{4.3}$$

N……... number of the inspection sample units.

7. Average relative difference [ARB] of the camera collected data with respect to manual data in [%]:

$$ARB \equiv \frac{\sum B}{N} \tag{4.4}$$

8. Absolute root mean square difference [ARMS] of the camera collected data in [m] or [m^2]:

$$ARMS \equiv \left| \sqrt{\frac{\sum \Delta q^2}{N}} \right| \tag{4.5}$$

9. Relative root mean square difference [RRMS] of the camera collected data with respect to manual data in [%]:

$$RRMS \equiv \sqrt{\frac{\sum B^2}{N}} \tag{4.6}$$

Table 12: Validation process accuracy results

Parameter	Alligator cracking	Longitudinal and transverse cracking
AAB [m or m^2].	0.28	0.23
ARB [%]	2.7	2.3
ARMS [m or m^2].	0.35	0.27
RRMS [%]	3.1	2.6

At first view, it is obvious from , that the parameter values of alligator cracking (e.g. ARB=2.7%) are larger than the parameter values of longitudinal and transverse cracking (e.g. ARB=2.3%). This can be explained by larger deviations in the determination of alligator cracking areas quantities, compared to the longitudinal and transverse cracking length and width quantities. The main reason for this deviation is that the detection of areas with alligator cracking is squared compared to the detection of the length and width of linear cracking. Furthermore, if also alligator cracking absolute average difference (AAB value) is considered, it's only 0.28 m^2, whereas alligator cracking absolute root mean square difference (ARMS) is only 0.35 m^2, this parameter is considered as a statistical measure of magnitude of the varying quantity. Because these two numbers are very close to each others, it can be noticed, that the area quantities of alligator cracking are scattered slightly. On the other hand, the calculated values show also an enhancement of areas and distances quantities determinations through digital images methodology. We have average deviations of 3.1% and 2.6% respectively (RRMS) in this validation process, the reason of this deviation could be found in the different given accuracies of the used instruments (40 cm used hand odometer accuracy, 2cm-3cm used camera accuracy for a distance of 2 km). There's a clear tendency that manually determined areas and distances are smaller than those of the camera based system, nevertheless, the differences between the camera based quantity q_{camera} and the manual based quantity q_{manual} are not large (minor measurements differences) (Al-MESTAREHI 2011). Perhaps, the reason of this tendency can be found in the fact, that these are errors of camera calibration, or scale errors of the hand odometer, where both of the measured systems are mainly dependent on the scale, resolution and the accuracy of the used apparatus in the data collection stages. Therefore, the low cost-efficient manual camera system has been used in this research work for the purpose of data collection. In fact, the task performed in this research for distresses classification and recognition was not fully automated. Visualization of the distresses was done manually and the distresses measurements such as distress quality and severity were preformed from digital images. This system is a step towards full automation in order to relieve the entire hazard that accompanied the traditional manual data collection techniques.

5 CONCLUSIONS

In this paper it could be shown that a cost-efficient manual camera system works very well for pavement distresses data acquisition, where the deviation of the validation process is determined to 3.1% (relative RMS) for alligator cracking, and to 2.6% (relative RMS) for longitudinal and transverse cracking due to different given accuracies of the used instruments in the measurement mode usage of this system. The advantages of the presented procedure are on the one hand the time efficiency, unlike manual data collection which consumes time, money and labour, and on the other hand the possibility to identify sections with certain types of distresses associated with their severity levels and quantities. This would help in assigning the concentrations of severe types of distresses over the networks. Therefore, pavement weakness could be improved, rehabilitated and located. Furthermore this low cost system was of great help for pavement distresses data collection, analysis, manipulation, displaying and classification. The system's development is a step towards real-time distresses classification. In addition, the PCI is calculated based on the results of a visual condition survey in which distress type, severity, and quantity are defined. Field verification of the PCI inspection method has shown that the index gives a good indication for pavement structural integrity and operational condition.

REFERENCES

Books:

Al-MESTAREHI, B.: *Integration of GIS, and PAVER Systems to award Pavement Maintenance Management System (PMMS)*. VDM Verlag Dr. Müller GmbH & Co. KG publishing house ltd, Saarbrücken, 2011.

Journal articles:

GHUZLAN, K.: *A Stereometric Knowledge-Based System for Maintenance of Street Network in Urban Areas.* M.Sc Thesis. Civil Engineering, Jordan, University of Science and Technology, not published, December 1995.

HOWE, R., CLEMEA, G.: *Feasibility of an Automated Pavement Distress Survey System Incorporating Digital Image Processing.* Full report, the media Team, VTRC Report No. 98-R1, 2008.

JASELSKIS, E.: *Field Data Acquisition Technologies for Iowa Transportation Agencies.* Iowa DOT Proj HR-366, ISU-ERI-Ames-94409, Pagination: p.193, 2005.

LEE, H., LEE, N., JITPRASITHSIRI, S., SORCIC, R.: *Development of Geographic Information System-Based Pavement Management System for Salt Lake City*. Transportation Research Record, Issue: 1524, 2101 Constitution Avenue, NW, Washington, DC, 20418-USA, p. 16-24, 2001.

SHAHIN, M.: *Pavement Management for Airports, Roads and Parking Lots*. Klluwer Academic Publishers Group, 101 Philip Drive, Assinippi Park, Norwell, Massachusetts 02061, p. 233-270, 1998.

SHAHIN, M., CLINE, G., BURKHALTER, J.: *Automated Data Collection for Pavement Condition Index Survey*. Fourth International Conference on Microcomputers in Transportation. Published by American Society of Civil Engineers, 345 East 47th Street, New York NY 10017-2398, USA, p. 572-583, 2006.

SHAHIN, M., WALTHER, J.: *Pavement Maintenance Management for Roads and Streets Using the PAVER System*. USACERL Technical Report M-90/05, p. 280, 2006.

SHUBINSKY, G.: *Visual and Infrared Imaging for bridge Inspection*. Northwestern University, BIRL Basic Industrial Research Laboratory, 2000.

SMITH, D.: *Advanced Technology and the Highway Maintenance*. Transportation Research Circular, 12th Equipment Management Workshop Conference, p. 4-14, 2005.

THEODORAKOPOULOS, D., CHASSIAKOS, A., MANARIOTIS, I.: *Improvements in Highway Maintenance Management in Greece*. Ninth, AASTHO/TRB Maintenance Management Conference, Published by Transportation Research Board, 2101 Constitution Avenue, NW, Washington, DC, 20418 USA, p. 11, 2007.

UDDIN, W.: *Microcomputers in Transportation for a maintenance Management System for Highways and Urban Streets*. Fourth International Conference on Microcomputers in Transportation, Published by American Society of Civil Engineers, 345 East 47th Street, New York NY 10017-2398, USA, p. 572-583, 2006.

WAUGH, J.: *Data Collection and Analysis Techniques*. XÍÍÍth World Meeting of the International Road Federation, 525 School Street, SW, Washington, DC, 20024, USA, 2007

Construction III

3D-MC^2 for Motor Grader: Motor Grader Blade Control with Slope Sensor and MEMS Gyroscope

Hiro Konno, Joe Brabec
Topcon Positioning Systems, Inc., U.S.A.

Abstract
A typical approach to automatically controlling the blade of a motor grader is to use a positioning device such as GNSS for the elevation control and a slope sensor for the slope control. Slope sensors, in particular, are an excellent device which provides accurate blade slope measurements and have been well accepted in the machine control market more than 25 years. However, slope sensors have two major drawbacks. First, slope sensors show slow response to rapid and large changes of the slope. This slow response is due to the internal filters used to reduce noise. Second, slope sensors work properly only under a limited range of dynamic motion. As slope sensors encounter a high dynamic motion, the slope measurement becomes erroneous because of the additional acceleration induced by the dynamic motion. To resolve those problems, this paper introduces a new method in which a MEMS IMU and a slope sensor are combined. IMUs provide slope measurements with extremely short delay. Moreover they are immune to disturbances due to high dynamic motions. A drawback, however, is that any sensor errors are accumulated as time passes. When a slope sensor and an IMU are integrated, those sensors work complementarily and provide accurate, fast, and stable slope measurements. The method is implemented in Topcon's 3D-MC^2 for Motor Grader. A field test showed that the system finished more than 90,000 m^2 in a day, which is larger than five times the average area finished by the previous system.

Keywords
Automation, Motor Grader, Blade Control, Slope Sensor, MEMS IMU, Kalman Filter

1 INTRODUCTION

In the last decade, the growth of the market of automation of earthmoving machines is significantly higher than those of the previous decades. A critical milestone for the rapid growth is the use of GNSS. In the late 1990s, machine control manufacturers began the development of machine control systems using precise GNSS positioning technologies, which paved the way for accurate, reliable, cost/time effective 3D machine control systems. A market analysis forecasts machine control applications using precision GNSS to grow at a Compound Annual Growth Rate (CAGR) of 23 – 28% during 2008 – 2012 [1].

Following GNSS, nowadays, a new technology which has potential to trigger the next innovation in the machine control market begins to emerge—the MEMS sensors. The micro-electromechanical system (MEMS) is an enabling technology whose global market revenue is expected to increase at a CAGR of 25.5% during 2009 – 2015 [2]. This strong market is due to the versatility of the MEMS technologies that cover a wide range of products demand from various markets including the consumer electronics and the automotive applications. Among various MEMS sensors, inertial sensors such as gyroscopes, accelerometers, and inertial measurement unit (IMU) are of great interest for the earthmoving machine control. In 2008, Topcon Positioning Systems introduced the world first dozer blade control system integrating a MEMS IMU into the conventional GNSS-based 3D machine control [3]. This system is called *3D-MC^2* (twice the speed, twice the accuracy). In this system, the combination of GNSS and a MEMS IMU detects the position of the dozer blade at the rate of 100 Hz (10 times higher than a typical GNSS receiver) with an extremely short delay. This high-speed and short-delay position determination enables the stable and accurate blade control even for high velocity operations such as 3rd gear.

After the successful introduction of 3D-MC^2, Topcon decided to develop a novel motor grader control system by integrating a MEMS IMU into the conventional system. A basic requirement for motor grader control systems is to control the blade elevation and the blade slope. Topcon's system uses one of the two lift cylinders, which move the blade up and down, for the blade elevation control and the other cylinder for the blade slope control. The sensor for the elevation control is a positioning sensor such as GNSS. The sensor for the slope control is a slope sensor mounted on the blade. Although the system has been well accepted in the market, there are two major problems. First, the response of the slope control is notably slower than that of the elevation control. The machine speed is, therefore, governed by the slope control, even though the elevation control has potential to control the blade with higher grading speed. The slow response of the slope control is due to the delay in the slope sensor. The slope sensor has an internal low-pass filter to reduce measurement noise. As the cost of the noise reduction, this filter induces the delay. The second problem is inappropriate blade motions which are often observed when the motor grader is moving with a high dynamic motion such as a sharp turn and a sudden brake. This problem is due to the vulnerability of the slope sensor against high dynamic motions. The slope sensor senses the gravitational force to measure the blade slope. A high dynamic motion induces additional acceleration (disturbance) which perturbs the sensing of the gravitational force, resulting in slope measurement errors. **Figure 84** is an example of the inappropriate blade motions. Shown is a motor grader turning a sharp curve. Because of the slope sensor error caused by the dynamic motion, the blade was inappropriately controlled to dig into the ground. But the ground was too hard to dig in; consequently, the machine was pushed up. The picture shows the right-rear wheel lifted up from the ground.

To resolve those problems, Topcon's new motor grader control system, *3D-MC^2 for Motor Grader*, integrates a MEMS IMU with the slope sensor. The delay of the IMU is significantly shorter than that of the slope sensor. Moreover the attitude estimation retains the accuracy regardless of the blade motion. A drawback is, however, that any sensor errors are accumulated in the attitude estimation; consequently, the attitude error exceeds the acceptable degree as time passes. When the IMU and the slope sensor are integrated, those sensors work complementarily. The slope sensor that has long-term accuracy and stability compensates for sensor errors in the IMU. The IMU, in turn, provides attitude estimation with a short delay and, therefore, high dynamic response. A disturbance detector that detects disturbances on the slope sensor under high dynamic motions is also implemented in the system. When a disturbance is detected, the system halts using the slope sensor for the IMU compensation until the disturbance passes. Accordingly, no disturbed measurement contaminates the attitude estimation.

This paper describes the outline of 3D-MC^2 for Motor Grader and introduces test results. A field test showed that the system finished more than 90,000 m^2 in a day, which is larger than five times the average area finished by the previous system.

Figure 84: Inappropriate Blade Motion

2 OUTLINE OF 3D-MC^2 FOR MOTOR GRADER

2.1 System Outline

The bottom line of the development is that the new system should be an add-on system to the previous motor grader control system, rather than a replacement. The emphasis is, therefore, on keeping as much of the hardware configuration of the previous system as possible. By restricting the hardware modification to a small portion, customers who have the previous system can easily update to the new system and gain the performance improvement.

Figure 85 shows the hardware configuration of the new system. The part shown in the red box is the hardware added in the new system. As shown, only a MEMS IMU (MC^2 Sensor) is added on the previous system with extra cables. This IMU is the same IMU used in 3D-MC^2 for dozers. MC-R3 is the main controller of the system. It processes measurements from all the sensors and generates control signals for each of the hydraulic valves connected to the elevation-side and the slope-side cylinders. GX-60 is the Graphical User Interface (GUI) unit that provides the operator with an interface and primary control for the system. The touch screen display shows the guidance for the grading such as the cut/fill value indicator and the position of the machine in the field. In addition, various parameters that adjust the machine control performance can be tuned by using the touch screen display. For example, the control gain is one of those parameters.

Units shown in the dashed box are the sensors used for the slope control. The locations at which those sensors are installed are shown in **Figure 86**. The Slope Sensor and the MC^2 Sensor are together installed on the blade and estimates the blade slope. The Slope Sensor contains an electronic level vial that acts like a precision carpenter's level. The MC^2 Sensor is a MEMS IMU that measures rotation rates and accelerations in three orthogonal directions with the sampling rate of 100 Hz. The algorithm to estimate the blade slope with those sensors is described in the next section in detail. The Rotation Sensor is an electric potentiometer installed on the hydraulic swivel of the blade. This sensor measures the rotation angle of the blade with respect to the front-frame structure of the motor grader. The Mainfall Sensor is a slope sensor that is installed on the front-frame structure. As the machine travels uphill or downhill, the Mainfall Sensor measures the longitudinal slope of the machine. Since the desired cross-slope given by the design file is perpendicular to the machine travel, the blade slope that realizes the cross-slope is different from the cross-slope when the machine is travelling uphill or

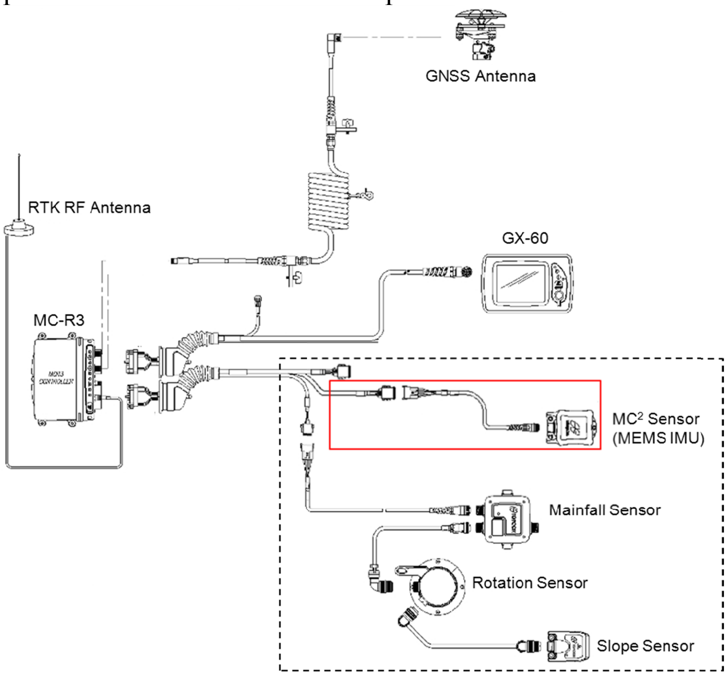


Figure 85: 3D-MC^2 for Motor Grader Hardware Configuration

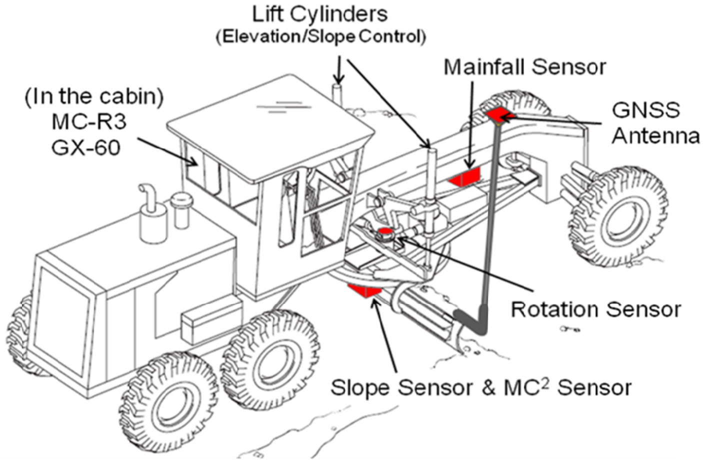


Figure 86: Motor Grader with 3D-MC^2 Motor Grader Control System

downhill with its blade rotated (see **Figure 87**). When the system computes the target blade slope, this difference is compensated by measurements from the Mainfall Sensor and the Rotation Sensor.

The GNSS Antenna and the RTK RF Antenna are for GNSS precision positioning used for the elevation control. The GNSS Antenna is affixed to the top of the vibration pole that is mounted on the blade, and the RTK RF Antenna is affixed somewhere on the machine. The system also allows, albeit not shown in **Figure 85**, to use Topcon's various elevation sensors such as 2D Laser Control, 2D Sonic Control, and mmGPS [4][5].

2.2 Outline of the Blade Control

This section discusses the control algorithm implemented in 3D-MC^2 for Motor Grader. First, let us define the "attitude" of the blade. An Euler-angle representation [6] is used to define the blade attitude with respect to the local navigation frame based on which the design surfaces are described (see **Figure 88**). The rotation order of the Euler angles is yaw (ψ), pitch (θ), and roll (ϕ).

Figure 89 is a high-level block diagram of the blade control. The left side and the right side of the diagram correspond to the elevation control and the slope control, respectively. The elevation control uses a proportional controller (P-controller). The module called "Cutting Edge Position Estimator" estimates the current elevation of the blade cutting-edge based on the position of the GNSS

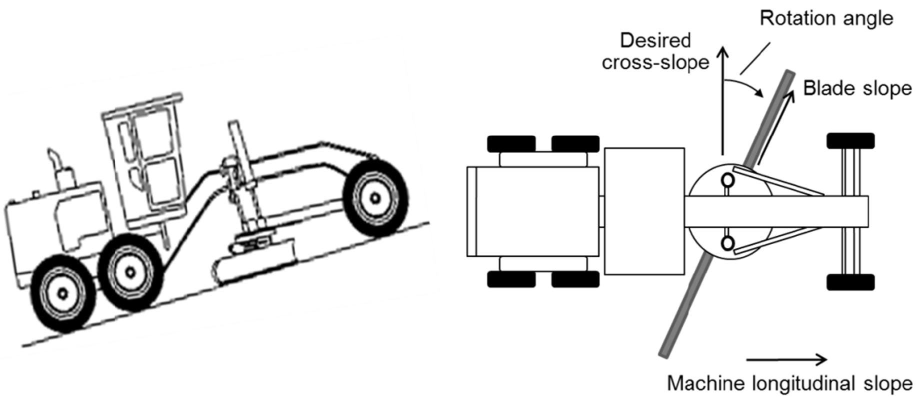


Figure 87: Functions of Rotation Sensor and Mainfall Sensor (*To compute the blade slope that realizes the desired cross-slope, the rotation angle and the longitudinal slope have to be taken in account.*)

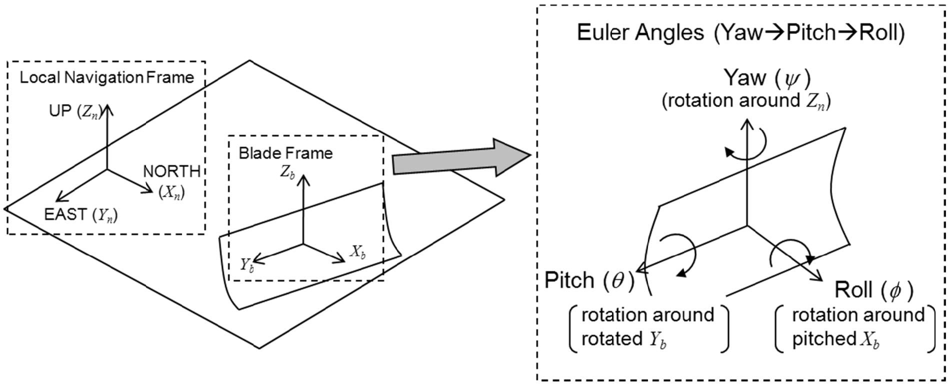


Figure 88: Local Navigation Frame, Blade Frame, and Euler Angles

antenna and the Euler angles of the blade. Among the Euler angles the yaw angle is estimated based upon the current blade-rotation and the history of the GNSS antenna position, while the roll and the pitch angles are estimated in the module called "Blade Attitude Estimator" (which is explained in the next section in detail). The following equation gives the position (therefore the elevation) of any point on the cutting-edge, based on the GNSS antenna position and the Euler angles.

$$P_{edge}^{nav} = P_{antenna}^{nav} - C_{blade}^{nav} l \tag{1}$$

where P_{edge}^{nav} is the position of the point of interest on the cutting-edge (which is called "blade control point") expressed in the navigation frame, $P_{antenna}^{nav}$ is the position of the GNSS antenna expressed in the navigation frame, C_{blade}^{nav} is the rotation matrix from the blade frame to the navigation frame given by the Euler angles, and l is the lever-arm from the blade control point to the GNSS antenna expressed in the blade frame. The lever-arm is computed from the geometrical information of the blade which is manually inputted to the system when the system is installed on the motor grader. A typical example of the blade control point is a tip of the cutting-edge.

The estimate of the cutting-edge elevation is then compared with the target elevation provided by the design file, yielding the elevation error. The control signal fed into the elevation-control valve is generated by multiplying the elevation error with the proportional gain, K_{ep}.

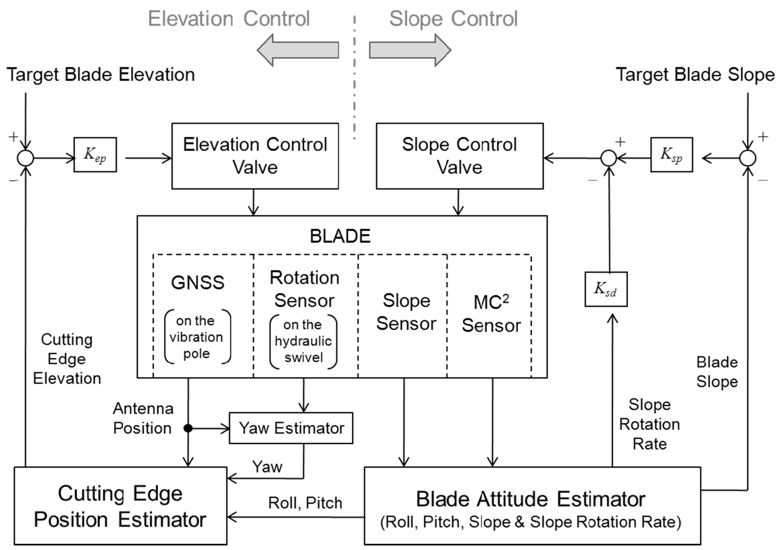


Figure 89: High Level Block Diagram of Motor Grader Blade Control

In contrast with the elevation control, the slope control uses a proportional-and-derivative controller (PD-controller). In this controller, the Blade Attitude Estimator module provides the estimates for the current blade slope and the slope rotation rate. The blade slope estimate is subtracted from the target blade slope provided by the design file, yielding the blade slope error. This blade slope error is then multiplied by the proportional gain (K_{sp}), yielding the proportional part of the control signal. The derivative part is, on the other hand, the product of the slope rotation rate and the derivative gain (K_{sd}). Note that, as discussed in section 2.1, when the system sets the target blade slope, the measurements from the Mainfall Sensor and the Rotation Sensor are used to compensate for the difference between the desired cross-slope and the target blade slope that realizes the cross-slope. Note also that, although the roll and the blade slope appear to represent the same angle, there is a slight difference between them. The blade slope is defined as an angle with respect to the level surface, while the roll is defined as a roll angle about the pitched surface. Hence those values are different for non-zero pitch values. The blade slope is computed from the roll and pitch angles as follows.

$$slope = \operatorname{atan}\left(\frac{\sin(\phi)\cos(\theta)}{\sqrt{\cos^2(\phi)+\sin^2(\phi)\sin^2(\theta)}}\right) \tag{2}$$

In 3D-MC^2 for Motor Grader, the proportional gain for the elevation control (K_{ep}) and that for the slope control (K_{sp}) can be manually tuned by the operator, while the derivative gain (K_{sd}) for the slope control is automatically selected by the system depending upon the setting of the slope proportional gain.

2.3 Blade Attitude Estimator with Slope Sensor and MC^2 Sensor

This section focuses on the Blade Attitude Estimator module in the block diagram shown in **Figure 89**. As described in the previous section, this module provides the blade slope and the slope rotation rate to the slope control and the roll and pitch angles to the elevation control. Two important features of this module are (1) increasing the speed of the slope estimation with keeping the accuracy and (2) eliminating estimation errors due to disturbances applied to the Slope Sensor under high dynamic motions. To realize those features, this module blends the measurements from the Slope Sensor and the MC^2 Sensor.

Figure 90 shows a block diagram of the Blade Attitude Estimator module. Roughly, the top half corresponds to the diagram for estimating the blade slope, slope rotation rate, and the roll angle (let us call it "slope portion"), while the bottom half shows the diagram for the pitch estimation (let us call it "pitch portion"). The core to increase the speed of the slope estimation is the Attitude Update module (module (1) in **Figure 90**) that is used by both the slope potion and the pitch portion. This module updates the blade attitude every time the three-axis gyroscope in the MC^2 Sensor outputs the rotation rate of the blade (100 Hz). A general attitude update algorithm for the inertial navigation system [6] is used to update the roll and pitch angles. Using the roll and pitch angles, the blade slope is computed with Equation (2). A major error source in this attitude estimation is the bias errors on the gyro sensors which could slightly change depending upon environmental conditions such as temperature. As discussed later, the Blade Attitude Estimator module has a mechanism to estimate the current bias values for the X-axis gyro and Y-axis gyro. The Attitude Update module compensates the gyro biases using those estimates.

In the slope portion, the blade slope computed by the Attitude Update module (S_{gyro}) is compared with the Slope Sensor measurement (S_{ss}), yielding the slope discrepancy (ΔS). As discussed in Section 1, the Slope Sensor has a long delay between the time at which measurements are taken, known as the *time-of-validity* [7], and the time at which those measurements are outputted. In contrast, the delay in the Attitude Update module is negligibly short. When the slope values from those two sources (S_{ss} and S_{gyro}) are compared, it is important to ensure that those values have time-of-validities that are sufficiently close each other. The Delay Hander module (module (2) in **Figure 90**)

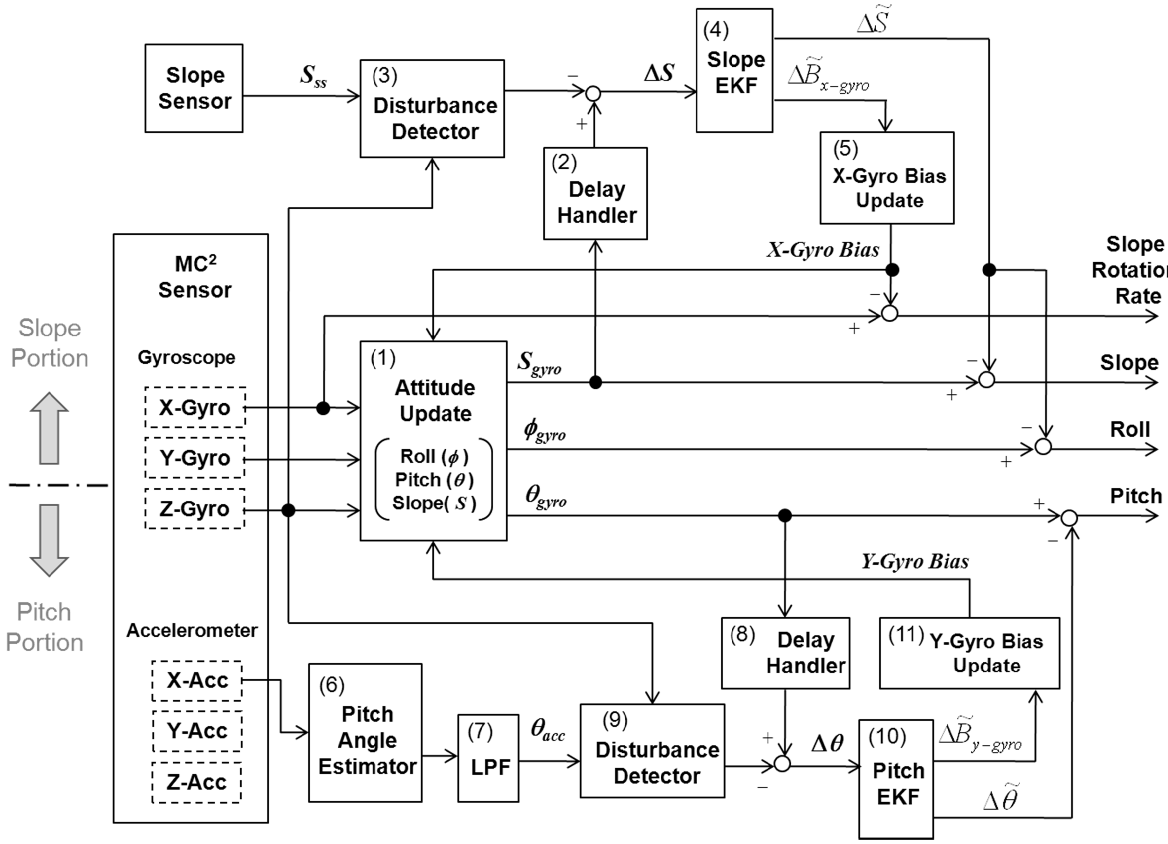


Figure 90: Block Diagram of Blade Attitude Estimator

handles the difference of the delays such that the slope values having close time-of-validities are compared.

The module called "Disturbance Detector" (module (3)) eliminates erroneous slope sensor measurements which are perturbed by a high dynamic motion. This module uses several metrics to detect disturbances. One of them is measurements of the Z-axis gyro. The Z-axis gyro senses the rotation of the blade around the vertical axis (more precisely, Z-axis of the Blade Frame). When the blade is rotated around the vertical axis with a certain speed, the centrifugal force perturbs the slope sensor. Large Z-axis gyro measurements are, therefore, a sign of a disturbance. The slope sensor measurements are discarded when Z-axis gyro measurements exceed the predefined threshold.

After the delay handling and the disturbance screening, the slope discrepancy (ΔS) is fed into an extended Kalman filter (EKF) which is shown as module (4) in **Figure 90**. This EKF estimates the correction to the blade slope ($\Delta\tilde{S}$) and the correction to the X-axis gyro bias ($\Delta\tilde{B}_{x-gyro}$). The slope estimate provided by the Attitude Update module is corrected with the slope correction. The corrected slope estimate is then outputted from the Blade Attitude Estimator for being used in the slope control. The roll angle estimated in the Attitude Update module is also corrected by the slope correction, and the corrected roll angle is outputted for being used in the elevation control. The correction to the X-axis gyro bias is used to update the estimate of the X-axis gyro bias. The updated X-axis gyro bias is then fed back to the Attitude Update module to compensate the bias error in the module. The raw measurements from the X-axis gyro are also corrected with the bias estimate and then outputted from the Blade Attitude Estimator module as the slope rotation rate, which is used in the slope control.

The software architecture of the pitch portion is similar to that of the slope portion. The main difference is that, instead of using the slope sensor, the pitch portion uses measurements from the X-accelerometer in the MC^2 Sensor to measure the pitch angle. The row measurements from the X-accelerometer are inputted into a module called Pitch Angle Estimator (module (6) in **Figure 90**) that transforms the acceleration into the pitch angle. Outputs from this module are then fed into a low-pass filter to reduce noise. As a cost for the noise reduction, this low-pass filter induces a significant delay (the same mechanism as the delay in the slope sensor). This accelerometer-based pitch measurement (θ_{acc}) is then compared with the pitch estimate by the Attitude Update module (θ_{gyro}). The functions of the Delay Hander (module (8)) and the Disturbance Detector (module (9)) are the same as those in the slope potion. The pitch discrepancy ($\Delta\theta$) between the accelerometer-based pitch measurement and the pitch estimate by the Attitude Update module is fed into an EKF (module (10)) that estimates the correction to the pitch angle ($\Delta\widetilde{\theta}$) and the correction to the Y-axis gyro bias ($\Delta\widetilde{B}_{y-gyro}$). The corrected pitch angle is then outputted from the Blade Attitude Estimator for being used in the elevation control. The updated Y-axis gyro bias is fed back to the Attitude Update module.

3 TEST RESULTS

Tests were conducted in Topcon's California Solution Site. The system was installed on a CAT 140H motor grader. The first test was to see the improvement of the speed of slope estimation. The motor grader was parked, and the blade was lifted up about 10% with the manual mode. The target blade slope was set to 0%. Then the system was turned to the automatic mode. This process introduces a step input in the slope control. **Figure 91** shows the response as a function of time. The blue curve shows the slope estimation by the new method described in this paper, while the red curve shows the raw measurements from the slope sensor. As shown, the slope estimation with the new method is significantly faster than the slope sensor. The maximum delay of the slope sensor with respect to the new method is about one second. This difference is significant for the blade control. In the previous system, when the slope control gain was set high, the system tended to be unstable due to the sensor delay. To guaranty the stability, the control speed had to be compromised. In the new system, the control gain can be set lager than three times a typical gain for the previous system, increasing the speed on the slope control without losing the stability.

The next test was to see the effect of the disturbance elimination. The motor grader travelled a test course having a 90-degree turn. **Figure 92** shows the slope measurements as a function of time. The blue and red curves correspond to the slope estimation by the new method and the raw slope sensor measurement, respectively. The cyan curve shows the target blade slope given by the design file. The period of the 90-degree turn is specified in the plot. It took about 20 seconds to pass through the turn. As shown in the Figure, the slope estimation by the new method accurately traces the target slope, implying that the blade was accurately controlled. On the other hand, the slope sensor measurement has a large error during the turn. This error was caused by the centrifugal force applied to the slope sensor during the turn. If the disturbance detector did not halt using the slope sensor for the IMU compensation, the blade slope estimation must have been contaminated by this erroneous measurement, resulting in an inappropriate blade motion. But without using the slope sensor during the turn, no inappropriate motion was observed.

The accuracy of finished surfaces was also evaluated. The test was conducted with 2D control mode. Topcon's system allows manually setting the cross-slope target in the 2D control mode. In the test, both the old system and the new system finished grade with setting three cross-slope targets: 0%,

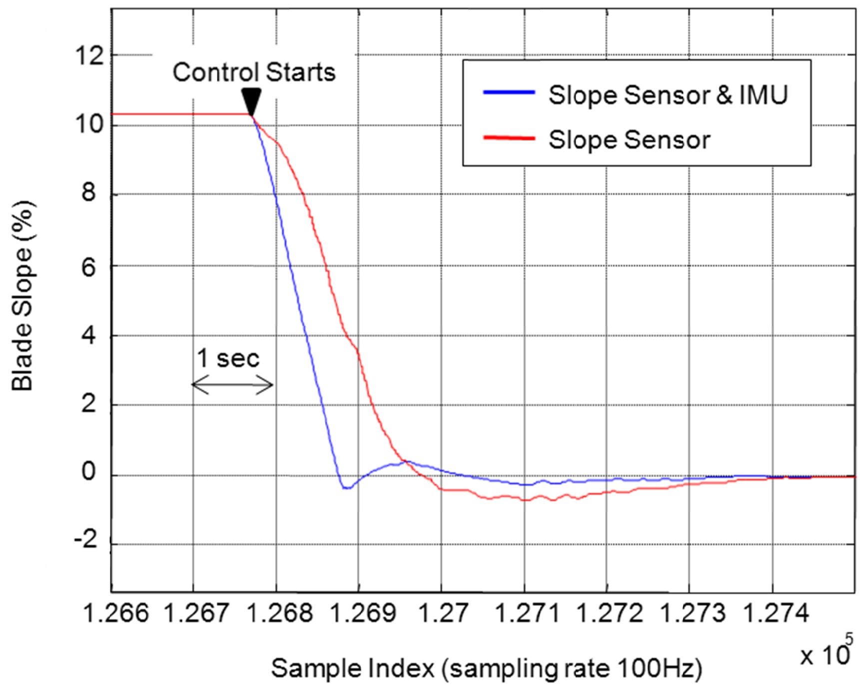


Figure 91: Step Response of New Method and Slope Sensor

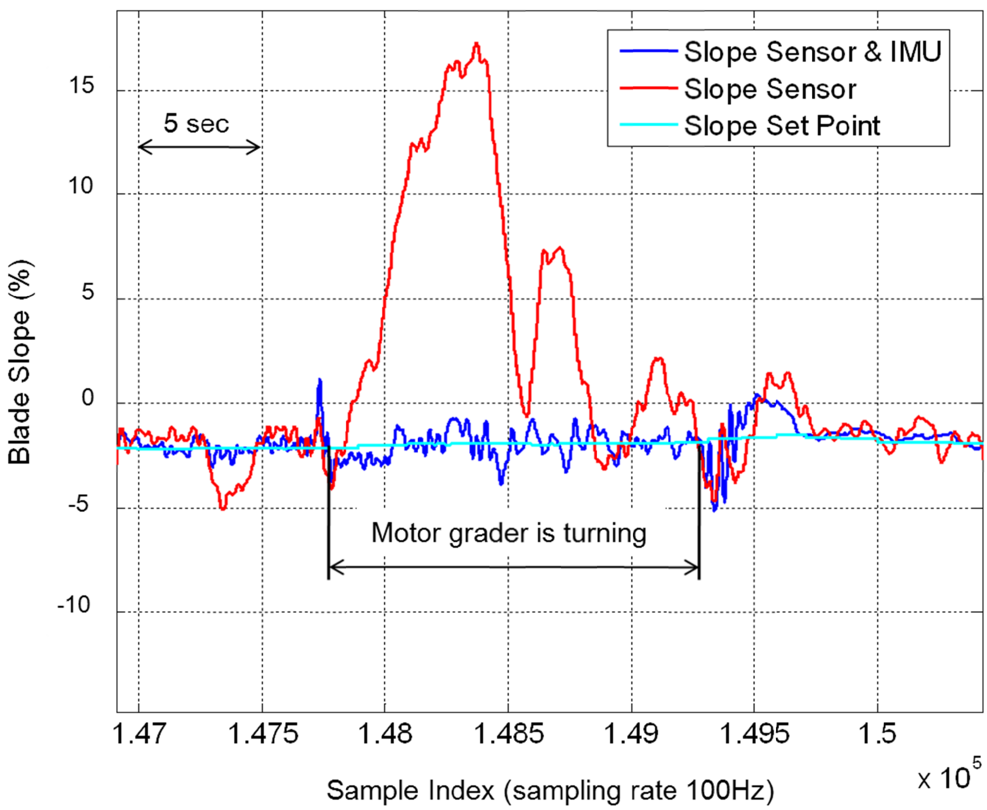


Figure 92: Slope Estimation during Turn

5%, and 20%. The finished surfaces were then measured by a smart level mounted on a 3-m aluminium bar, and RMS values of the cross-slope errors were compared between the old system and the new system. **Table 13** summarizes the accuracy improvement of the new system. For all the cases, the new system improves the cross-slope accuracy more than 40%. Moreover, no inappropriate blade motion was observed in the new system during the test, while occasional inappropriate motions were observed in the old system.

In addition to the in-house engineering tests, beta-tests were conducted by selected customers. Those customers evaluated the system by actually using the system in their job sites and provided a report. One customer reported that about one million square feet were finished in one day with the finishing accuracy of 0.018 meters (with GNSS accuracy check).

4 CONCLUSIONS

Topcon's 3D-MC^2 for Motor Grader was designed to achieve two goals: increasing the speed of the slope control, and improving the performance under high dynamic motions. The results of the engineering tests and the beta tests were promising. After those tests, the system was finally released in August 2011. After the release, valuable feedback has already been provided by customers. Among them, of great interest is the request for increasing the speed of the elevation control. In the previous system, it was the slope side which customers complained about for its slow response. However, with the slope control significantly speeded up, the elevation side becomes the slower side. To satisfy the demand for higher productivity, Topcon has already begun the development of the elevation enhancement. The idea is integrating GNSS with the IMU to increase the speed of the elevation control. The technology used in 3D-MC^2 for Dozer can be used for this improvement.

Another interesting comment is a question asking why the slope sensor is necessary in addition to the IMU. As discussed in this paper, the slope sensor is crucial to provide stability and accuracy for the blade slope estimation. However, the Y-axis accelerometer in the IMU can play the same role as the slope sensor, providing accurate and stable blade slope values. By using the Y-axis accelerometer

Table 13: Slope Accuracy Improvement in RMS Values

Target Cross-Slope	Improvement of the new system over the previous system
0%	41%
5%	48%
20%	42%

instead of the slope sensor, the slope sensor can be removed, reducing the cost and simplifying the installation and the system calibration. The reason to keep using the slope sensor is its high reliability. Slope sensors using an electric level vial have been showing high reliability in the earthmoving machine control industry for more than 25 years. In contrast, MEMS accelerometers are emerging technology, and it is still unclear if their reliability parameters such as MTBF are as high as those of the slope sensors. It is a future task to evaluate and improve (if necessary) the reliability of the MEMS accelerometer and remove the slope sensor from the system.

All comments and feedback from customers are crucial. Topcon keeps working to improve our systems to increase customers' productivity.

REFERENCES

LORIMER, R. and GAKSTATTER, E.: *GNSS Precise Positioning Markets 2008 – 2012*. Position One Consulting, 2008.

GLOBAL MARKETS DIRECT: *Global Micro Electro Mechanical Systems (MEMS) Market Analysis and Forecasts to 2015*.
http://www.globalmarketsdirect.com/Report.aspx?ID=Global-Micro-ElectricalMechanical-systems-

(MEMS)-Market-Analysis-and-Forecasts-to2015&ReportType=Industry_Report&coreindustry=ALL &Title=Clean_Technology

STURM, A.: *3DMC2 inertials—State of the Art dozer control solution*, Proceedings of 2nd International Conference on Machine Control and Guidance, 2010.

KAAK, M.: *mmGPS for Pavers—Complete Paver Control*, Proceedings of 2nd International Conference on Machine Control and Guidance, 2010.

GROVES, P. D.: *PRINCIPLES OF GNSS, INERTIAL, AND MULTISENSOR INTEGRATED NAVITAGION SYSTEMS*. Artech House, 2008.

TITTERTON, D. H. and WESTON, J. L.: *Strapdown Inertial Navigation Technology (Second Edition)*. AIAA, 2004.

Links:
http://www.topconpositioning.com/products/machine-control, last accessed on November, 2011.

Machine Control & Guidance for Large Type Tractors

Eric Durand
Caterpillar

Abstract

Several aftermarket systems exist in the market today, but a system designed by an equipment manufacturer can provide the benefits that only full integration with the machine systems makes possible. Depending on machine family and the specific application of interest, many of the components necessary for automation are already present on the machine.

Cat Grade Control design leverages these components and optimizes the design including sensors, displays and on-board data processing.

Many customers would explain that blade mounted automatics on large dozers are only used, at best, 10% of the time the machine is in operation. Cat Grade Control for the Cat D8T and D9T is a fully functioning, cab mounted, factory integrated MC&G system (3D GNSS), which differentiates the track-type tractor product with never before seen features. By combining the AutoCarry features into Cat Grade Control, the system provides automatics and guidance support the other 90% of the time and allows the machine to be used more productively with fuller blade loads.

Due to its integration the Cat Grade Control GNSS system is a first of its kind 3D automatic blade positioning system without the need for blade mounted components and is ready for operation from the moment it leaves the factory.

Keywords

CAT, Construction industry, GNSS, GPS, Grade Control, RTK, large dozer, Positioning Sensor Cylinder PSC, Type track tractor.

1 INTRODUCTION

1.1 History

Type Track Tractors always have been one of the first earthmoving machines equipped with 3D GNSS systems. Back to the mid nineties, the first GPS systems with rugged antennae were installed on top of mast located directly on the blade of a track type tractor. If the first systems were “indicate” only, very quickly automatic control appeared to get the full advantage of the positioning system.

In 1995, Caterpillar & Trimble established a partnership to develop CAES (Computed Aided Earthmoving System) system for mining industry. Based on this relation, a 50/50 Join-Venture is created on April 2002 - CTCT Caterpillar Trimble Control Technologies to combine forces for development and distribution to provide solutions for the construction industry.

1.2 First Steps of integration

The major suppliers of Machine Control & Guidance systems are Trimble, Topcon and Leica Geosystems. Other smaller suppliers are also engaged in this growing segment. Those companies provide the machine control components including, electrical harness, hydraulic kit, positioning sensors and an onboard computer using the software to display the guidance information to the operator and drive the valves in automatic mode.

Concerning OEM’s, Original Equipment Manufacturers, Caterpillar has been a pioneer in term of integration by offering in 2004 the Attachment Ready Option (ARO) on new machines. This ARO consists in the integration of the machine control harness directly in the main electrical harness of the machine. The valve module is directly the machine ECM (Electronic Control Module) and bracket supports are integrated from the factory. Then, the CAT AccuGrade™ 2D or 3D positioning system components can be plugged directly to the available connectors.

The second generation of integration consist in Grade Control Ready (CGR) making any machine capable to support MC&G for more flexibility and an easier dealer inventory management. The machine harness is already designed for grade control. By simply adding brackets and external piece of harness according machine type, a “Dealers Install Kit” is available to make any standard machine ARO without any welding, ready to plug a system.

The third generation, which is especially the one detailed in this document is the Factory Fit solution named CAT Grade Control available today on Large tractors D8T, D9T, Scrapers and Asphalt Pavers.

2 COMPONENTS

Factory integrated and installed harnesses and sensors allow for real-time blade tip positioning. As compared to traditional aftermarket externally mounted sensors, the Cat Grade Control sensors are protected from the harsh working environments to which the machine is subjected.

With the use of position sensing cylinder technology, Cat Grade Control is a chassis mounted system that moves the traditional blade edge positioning technology off the blade and onto the cab.

Figure 93: D8T with Cat Grade Control option

Table 14: Machine mounted Components

	1 GNSS receiver located on top of the Cab
	1 Inclination sensor to provide the pitch and roll of the machine body
	2 PSC Positioning Cylinder Sensor located in the blade cylinders.
	Factory standard machine implement Electronic Control Module (ECM) includes integrated software necessary to operate the system.

	Keypad to select operating modes in conjunction with the remote switches on the joystick
	Display using the 3D design to provide Cut/Fill information for both blade tips.
 or 	Radio or Cell modem to receive the base station information and transmit productivity data. *(Not included from factory due to different country regulations)*

A Controller Area Network (CAN) provides the communication network for all components. Position sensing cylinders (PSC) are a major new component in the electro-hydraulic and programmable hydraulic systems, which are being used increasingly in Caterpillar equipment.

Electronics within the cylinder allow the hydraulic system to be linked into the machine's electronics. This link enables the system to track the cylinder stroke.

<u>How Position Sensing Cylinders Operate</u>

The Position Sensing Cylinder features a cylinder rod, which has been drilled to allow a magnetic pulse to be constantly measured through the full length of the cylinder. The measurement of the pulses allows the system to know where the cylinder is positioned at all times. The encapsulated design protects the electronics from environmental conditions offering a cost effective, reliable and durable product.

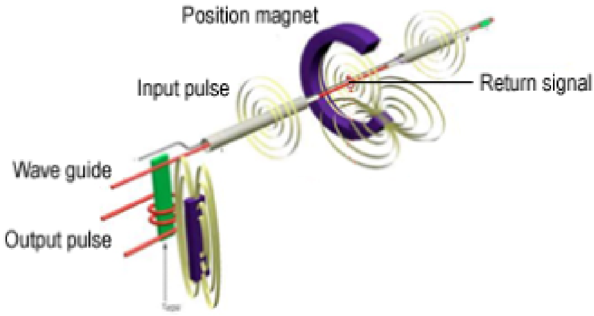


Figure 94: traditional grade control

3 APPLICATIONS FOR LARGE TYPE TRACK TRACTORS

Earthwork is one of the major works involved in large construction project. The large type track tractors is one of the most important items of equipment used for pioneering, clearing, bulk earthmoving, slot dozing to grade, rough grading in rugged environments.

If Machine Control and Guidance products are more and more accepted and utilized, only few systems are installed on heavy equipment. The MC&G products are more commonly used as fine grading tools and usually for the final and expensive layers. However, there is still an opportunity for contractors to improve their earthwork and get jobs to be completed more quickly, at lower cost and with the highest degree of accuracy.

3.1 Traditional grade control

For large dozers, many aftermarket systems exist in the market today, but they are limited to traditional grading operations mainly used in automatic mode. The operator starts each pass back of the cut. Each pass uses the entire length of the cut at a uniform depth. The efficiency and productivity suffer because the machine travels the entire length of the cut in both directions with each pass.

Figure 95: traditional grade control

3.2 CAT® Grade Control

The most efficient way for dozing is the “Back to front” technique where the operator starts the cut from the front to the back with full blade load. The push distance increases with each pass. This technique uses more efficient downhill blade loading and a slot is created and utilized throughout the cut to maintain the material.

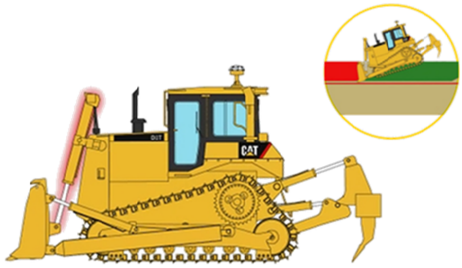


Figure 96: Cat Grade Control

3.2.1 Rough grade feature

Rough Grade Control is an operating mode that increases the machine productivity by allowing the operator to perform manual dozing above grade with comfort in knowing when the blade approaches grade, the system will begin controlling depth. The feature still allows the operator to tilt steer while automatics control lift to design limiting large cuts below the intended design. It uses automatic blade lift control when the blade is in contact with the design plane, or below the design plane.
Manual commands are used whenever the blade is above design. The operator always controls tilt control.

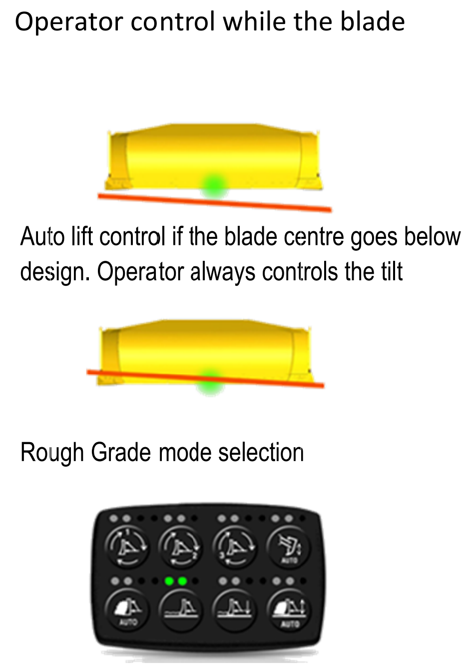


Figure 97: Rough Grade

3.2.2 Grade protection

Grade Protection is an operating mode that prevents the blade tips from going below the site design plan. It allows the operator to perform manual tasks with the added security of knowing that they will not go below the site design and will override/block manual commands when necessary to limit blade travel. A typical application for Grade Protection would be site grooming.

3.2.3 Additional features

3.2.3.1 AutoCarry

AutoCarry is a feature that automates blade lift control to maintain a desired blade load, improve job consistency and to reduce operator fatigue. The feature utilizes velocity data from the GNSS receiver (replacing the ground speed radar from the past) and an updated control algorithm to improve overall performance. The feature works seamlessly in conjunction with Cat Grade Control and Auto Blade Assist.

3.2.3.2 Auto Blade Assist

Auto Blade Assist is a feature that automates blade pitch and lift control during a typical dozing cycle to help reduce operator fatigue. Auto Blade Assist functionality can be operated in conjunction with AutoCarry to help further reduce fatigue.

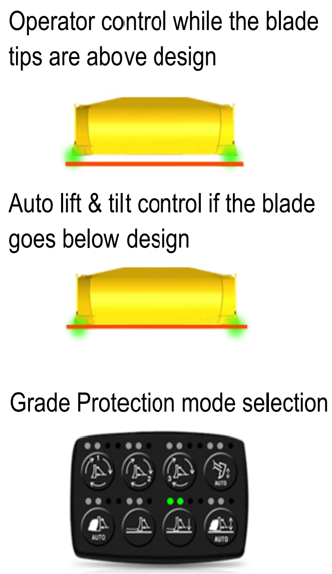


Figure 98: Grade Protection

3.2.3.3 Automatic Ripper Control

Automatic Ripper Control is a system that assists the operator by automatically controlling the ripper depth. The feature automatically changes the depth of the ripper shank based on track slip. This automated feature allows the operator to concentrate and plan out what will be done next rather than focus specifically on the ripper shanks, maintains maximum productivity and minimum wear and tear on the machine helping to reduce operator fatigue.

4 CONCLUSIONS

Due to its integration, the Cat® Grade Control GNSS System is a first of its kind 3D blade positioning system without the need for blade-mounted components. With the Grade protection functions, the system allows the operator to work in manual mode while the blade is above the design, which represents about 90% of the time. Once the grade is reached, the automatics take control of the blade to prevent any over cut.

This new generation in integrated machine control and guidance provides additional advantages over traditional systems available and makes the track-type tractor not only a grading machine, but also a high productivity solution.

REFERENCES

Links:
http://www.cat.com/equipment/track-type-tractors/large-track-type-tractors, last accessed on August 25, 2011

This Paper has been accepted as a "peer reviewed paper" by the scientific committee

Sensor Supported Surveillance of Ground to Control the Stability of Mobile Construction Machines

Jürgen Frick[1], Michael Schreiner[1], Gottfried Sawade[1],
María A. Gonzalez-Huici[2], Udo Uschkerat[3]
[1] Materials Testing Institute University of Stuttgart – MPA Stuttgart, Germany
[2]Centre for Sensor Systems, University of Siegen – ZESS, Germany
[3]Fraunhofer Institute for High Frequency Physics and Radar Techniques – FHR, Germany

Abstract
The daily progress on construction sites worldwide is driven be the use of mobile construction machines like concrete pumps, truck-mounted cranes, and mobile elevating work platforms. The stability of such machines is ensured by supporting legs. The upset of a mobile construction machine causes major damage to persons and property. 23 % of upsets of such machines is caused by soil collapse under the supporting legs [Beutinger, 2005]. Therefore the aim of the project is the development of a sensor solution for the surveillance of the load bearing capacity of soil and the detection of dangerous unknown underground cavities. At the best the two sensor solutions could be integrated in the positioning process of the mobile construction machine.

The surveillance of load bearing capacity of soil is based on the measurements of setting and load force at the supporting legs. The load bearing capacity of soil could be extrapolated online from the measured data according to the theory of soil behaviour from Kondner and Zelasko [Kondner, Zelasko, 1963]. Cavity detection will be done by the use of adapted radar antenna and specialised algorithms based on the modelling of test measurements.

Keywords
load bearing capacity, settlement, accelerometer, GPR, cavity detection, construction machines

1 INTRODUCTION

This paper is divided in two main parts. The first, which corresponds to sections 2 and 3, focuses on the determination of the load bearing capacity of the soil along with the settlement and maximum load from the acceleration and tilt measurements when applying a certain pressure. In the second part, we deal with the detection of potentially dangerous cavities using an ultra-wideband GPR system. In particular, in Section 4 we briefly discuss the principles of Ground Penetrating Radar in general and address the important issues to design a GPR system adapted for deep cavity detection. Next, in Section 5, the prototype system is described and some laboratory measurements of the signal generator and antenna element performance are illustrated. In Section 6 we present the processing techniques and algorithms that can be applied to detect and identify cavities. Finally, conclusions follow.

2 SURVEILLANCE OF LOAD BEARING CAPACITY

The load bearing capacity of soil influences its load-settlement behaviour. Therefore, it is possible to predict the load bearing capacity of soils on the basis of its deformation behaviour. To obtain impartial load-settlement relationships, the measurements of the settlements of soils must be made in relation to a global coordinate system, i.e. independent of a reference point.

It is not possible to get impartial information respective to the settlement in soil base from ascending a supporting leg. Hence, and since the machine movements in setting up the mobile machinery are significantly larger than the ground settlements, it becomes necessary to measure the settlement of any independent reference point. In practice, the setting up of test frames is too time-consuming and error-dependent; therefore the only solution is a sensor that responds only to changes in the global reference system. Accelerometers and vibration pick-ups are suitable for this purpose to register the "absolute" motion in relation to the global reference system "earth". The determination of the settlement is carried out through the integration of the measured signal.

3 EXPERIMENTAL APPROACH

3.1 Measurement system

The developed measuring system for determining the load bearing capacity of the area contains the following elements: a regenerated capacitive acceleration sensor (Honeywell, type: QA1400), inclination sensors with a tilt range from +15 to -15 degree (Seika, type: N3) and a load cell which is based on electrical strain gauges (DMS).

3.2 Determination of settlement

In the simplest case where the orientation with respect to the vertical position remains unchanged in the course of erecting, the settlement is calculated by double integration versus time [Beutinger, 2005]:

$$s(t) = \int_0^{t''} \int_0^{t'} (a - a_0) dt'' dt' \tag{1}$$

with a the measured acceleration and a_0 the measured initial acceleration.

Another approach takes into account the tilt of the vertical load sensor during erecting. Here, the measured accelerations are tilt corrected, hence the initial acceleration doesn't need to be provided. This consideration leads to the following approach:

$$\mathrm{s(t)} = \int_0^{t''} \int_0^{t'} (a_{calc}(t')dt'dt'' \tag{2}$$

where the calculated acceleration is given by $a_{calc} = a_1 - a_{tilt}$ and the condition for the velocity at the initial time is $v(0) = 0$. Then, measurement data from the acceleration sensor 1 and the inclination sensors n_1 and n_2 are collected and the overall inclination is averaged over a period of one second.

From the so calculated total tilt (deviation of the sensor to the vertical) using equation (1) the tilt corrected acceleration is determined a_{tilt}.

$$a_{neig} = g^*\cos(\beta) \tag{3}$$

where α is the tilt angle between the sensor and the vertical direction.

Next, we show an example of such a measurement. The sensor was moved in a calibration stand vertically with a hydraulic cylinder. At the initial position the sensor had a small angle of about 9 ° in respect to the vertical. In the final position the inclination was approximately 0.4 °. **Figure 99** shows the measured acceleration and the acceleration calculated from the inclination sensors a_{tilt}. The individual impacts of the hydraulic pump can be observed clearly from the tips of the measured acceleration.

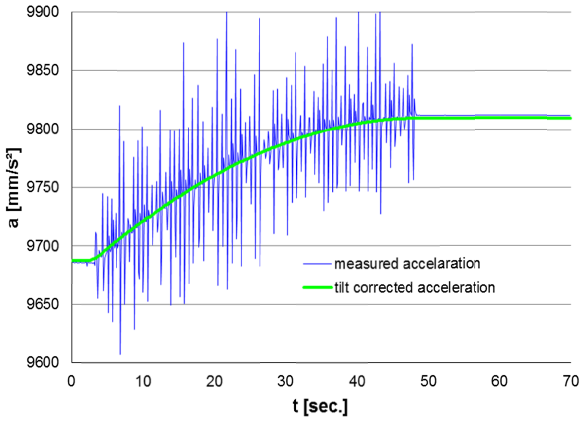


Figure 99: Measured acceleration and acceleration calculated from the inclination sensors.

Figure 100 shows the calculated vertical displacement of the sensor according to the integration with equation (2). As control, the vertical (actual) displacement of the three displacement transducers is shown.

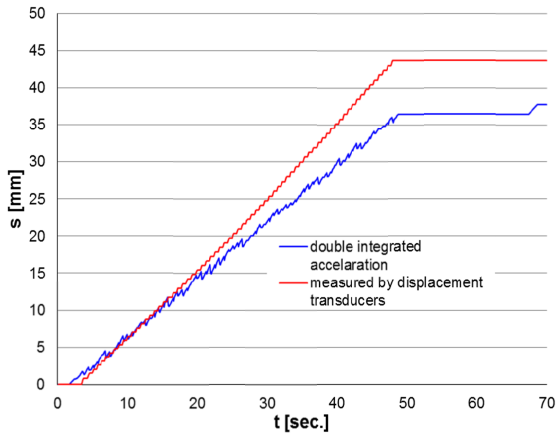


Figure 100: Comparison of equation (6) from the calculated and the measured displacement.

It is evident that the integration of equation (2) in the initial area of about 20 seconds, supplies a sufficient match with the actual displacement. So that load-settlement curves in the initial period of exposure are recorded with sufficient accuracy.

3.3 Determination of the failure load

From the knowledge of the partial increase of the load settlement curve the expected maximum load can be determined using the hyperbolic-equation of the load settlement curve according to [Kondner, Zelasko, 1963], [Beutinger, 2005].
From the expression

$$F(s) = F_{\max} \frac{s}{(b+s)} \tag{4}$$

follows for the size $y = \dfrac{s}{F(s)}$ the relationship:

$$y = \frac{b}{F_{\max}} + \frac{1}{F_{\max}} \cdot s \tag{5}$$

with $F_{\max}$, the ultimate load and b, a parameter depending on the initial stiffness.

The maximum load can thus be obtained from the slope of the linear equation. This is demonstrated by the following example, where we show the calculated load-settlement-course from a load determined by the sensor (see **Figure 101**). In the illustration the settlement obtained by three transducers is also displayed. Determining the maximum load, using equation (5) yields a value of 170 kN; the actual peak load (120 kN) is overestimated. For application in practice, the maximum load has to be reduced by a safety factor. The consideration of the entire load-settlement curve is only possible during the set-up of the machine. A permanent monitoring of stability, however, is not possible.

Another possibility of this system is to capture the settlement changes during a short time loading. For this purpose, defined short impulses are initiated in the support leg and the corresponding change in the settlement is calculated according to equation (2). **Figure 102** illustrates the change of load and the corresponding settlement due to transient impacts. The corresponding dynamic consistency modulus Esdyn results from the quotient of the maximum of the load difference and settlement difference.

The deformation modulus is proportional to the system stiffness. Changes of the system stiffness show the appearing of the ultimate limit state. Due to the short-term load from an idle state, the computation of the deformation modulus can be calculated at any point of the load, because the integration of acceleration is performed over a small period and the initial condition v = **0** is satisfied always sufficiently accurate. Therefore, this method is particularly suitable for monitoring the stability, even during the operating phase of a mobile construction machine.

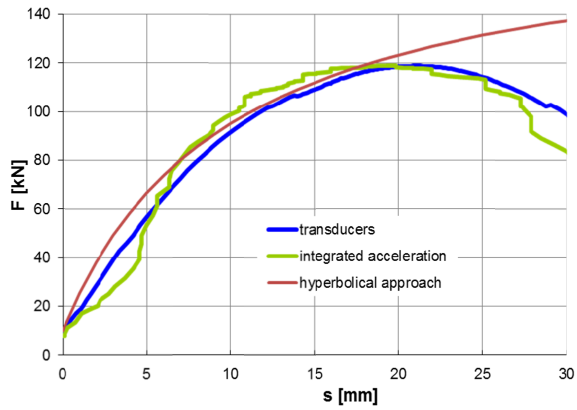


Figure 101: Extension of the support leg on dense sand: force - settlement curve.

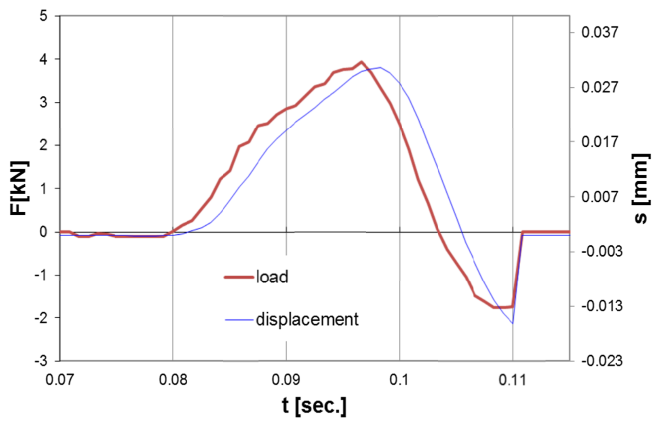


Figure 102: Force-time curve and setting-time course for short-term impact.

Figure 103 shows the results of impact loads, which were obtained with a stress test in sand. The respective dynamic consistency moduli are plotted against the soil pressure. The deformation modulus was determined from the relationship:

$$E_S = 1.5 \cdot d \cdot \frac{d\sigma(\sigma)}{d\,s} \tag{6}$$

where d is the diameter of the loading plate.

The dependence of the consistency modulus of the soil pressure according to usual behaviour of soil [Wichmann and Triantafyllidis, 2006] is clearly visible. It seems to be possible to predict soil failure from the change of the slope of the deformation moduli.

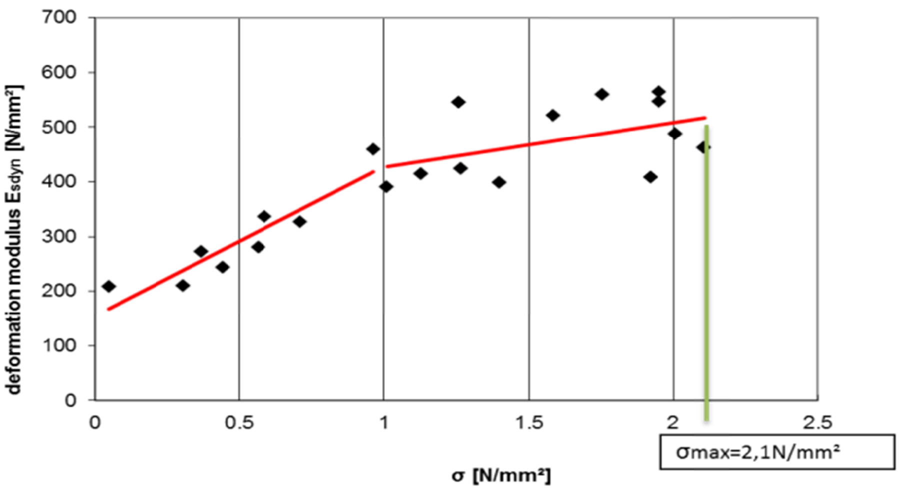


Figure 103: Dynamic deformation modulus of sand depending on the soil pressure.

4 GROUND PENETRATING RADAR

Ground Penetrating Radar (GPR) is a non-destructive geophysical technique that can be applied to investigate near-surface underground structures. The measurement principle basically consists of the emission of a high-frequency (50 MHz to 1 GHz) electromagnetic pulse by a transmitting antenna whose travel time and amplitude are subsequently recorded by a receiving antenna. The kind of objects and features investigated can vary from a few centimetres to hundreds of meters deep depending on the operating frequency.

The radar wavelet propagates through the soil with phase velocity dependent on the dielectric properties of the ground,

$$v = \frac{1}{\sqrt{\varepsilon\mu}} \approx \frac{c_0}{\varepsilon_r}$$
$$\varepsilon_r = \frac{\varepsilon}{\varepsilon_0} \tag{7}$$

where, ε is the permittivity of the material, ε_0 is the free space permittivity (8.854x10^{12} F/m), μ is the magnetic permeability, μ_0 the permeability of the free space (12.57x10^{-7} H/m), and c_0 is the velocity of light in vacuum. The approximation above is valid when the magnetic properties are close to the values of the free space.

Along its travel path the radar wavelet is partially reflected and refracted at the different boundaries (soil layers or distinct objects) where there is a change of the relative dielectric permittivity of the media. Together with the dielectric contrast between background soil and scatterer, the shape and magnitude of the signal reflected by an object are mostly influenced by the reflector size and depth, the system operating frequency and bandwidth, and the signal attenuation in the host material. All these questions will be addressed in the next sections.

4.1 Measurement Setup

There are different acquisition modes but probably the simplest and most popular of them is the common offset operation mode. In this mode the transmitter and receiver antennas are kept at a constant distance and pulled along the survey track simultaneously as illustrated in **Figure 104** (right). The pulses are emitted at a constant repetition rate while the traces are triggered at a fixed separation interval controlled by a survey wheel. This results in a series of traces which are finally displayed by the measurement software as a function of position and time in a so-called radargram. Due to the beam-width of transmit and receive antennas and the differences in round-trip travel time of the pulse caused by the movement of the antenna along the measurement line, the reflections from scatterers will appear as hyperbolic curves in the recorded data (see **Figure 104** left). The scattering hyperbola is given by:

$$s = vt = 2\sqrt{d^2 + \left(\frac{a}{2}\right)^2} = \sqrt{4d^2 + a^2} \tag{8}$$

with s the travel path, t the two-way travel time, d the depth of the scatterer and a the distance between transmitter and receiver.

In the presence of horizontal or almost horizontal reflectors a multi-channel GPR measurement allows to directly estimate cavity depth d and relative permittivity ε_r from the travel-time difference. Besides a second receiver channel provides some extra information that may be useful to identify subsurface voids. Therefore, we have built a prototype impulse GPR system consisting of one transmitter antenna and 2 receiving antennas as shown in **Figure 104** left.

4.2 Penetration Depth

Another important parameter to take into account when imaging deep objects are the energy losses of the electromagnetic waves on its way through the soil. The largest amount of energy loss results from damping of free charge carrier movement.

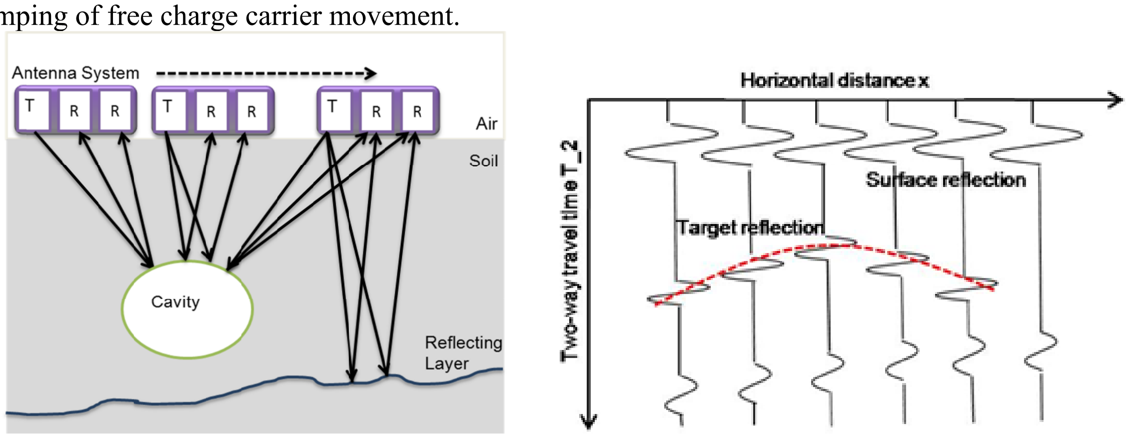


Figure 104: Common Offset with 2 receiver antennas configuration and hyperbola formation.

The material dependent attenuation is determined by the direct current electric conductivity σ_{DC} of the investigated medium. Depending on the travelled distance x the amplitude E of the electromagnetic wave decreases exponentially with respect to its starting value E_0:

$$E = E_0 e^{-\beta x}, \beta = \frac{\sigma_{DC}}{2c_0\varepsilon_0\sqrt{\varepsilon_r}} \tag{9}$$

with β as the attenuation loss. Then, the increasing conductivity of the soil reduces drastically the penetration depth, which is given by: $\delta = 1/\beta$ (see **Figure 105**). Furthermore, losses are also induced as a consequence of reflection and transmission processes at the interfaces between boundaries.

In addition, so-called spherical losses happen which are caused by the geometrical spreading of the energy with the inverse of the square of the traveled path.

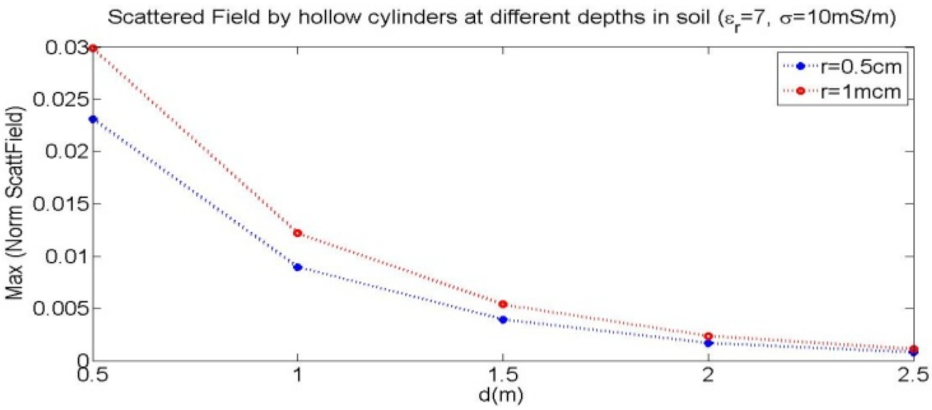


Figure 105: Amplitude attenuation with depth for 1V incoming plane wave.

4.3 Resolution

The successful detection of underground objects or structures is primarily restricted by the resolution of the measurement. This is the ability to differentiate two structures or signals which are temporally and/or spatially close to each other. For an impulse radar the vertical resolution mainly depends on the duration of the radar pulse, which is related to the center frequency of the antenna. The shorter the pulse duration (which is inversely proportional to the associated bandwidth, i.e., 2 ns pulse corresponds to 500 MHz bandwidth), the better its resolution will be. It is generally accepted that two nearby events can be distinguished if they are separated in time by a difference of half the effective pulse duration τ, which is obtained from the width of the signal envelope at its -3 dB level. Therefore the expected spatial vertical resolution can be calculated from the effective duration τ of the radar pulse and the wave propagation velocity v in the medium as follows [Annan, 2003]:

$$\Delta V = \frac{\tau c_0}{4\sqrt{\varepsilon_r}} \qquad (10)$$

In particular, for scatterer classification it is crucial to resolve top and bottom reflections, since the amplitude and travel-path difference between both signals may help to identify a potential cavity.
The relative permittivity of dry geologic materials ranges from 3 to 6 and the permittivity increases with the water content. Then, assuming the worst case scenario (dry sand with permittivity equal to 3), the range resolution for a 3 ns pulse would be around ~20 cm, which in principle should be enough for resolving "dangerous" cavities.

5 THE GPR SYSTEM

The design of a GPR system can be classified into different categories according to its hardware implementation. There are two main classes: time domain and frequency domain. Time Domain GPRs (called also impulse radars) transmit an impulse and receive and process the backscattered signal using a sampling receiver. They can be further separated into two main categories: amplitude modulated and ultra-wide band (UWB), but since we need to combine range with high resolution, a large relative bandwidth is necessary and we will just focus on the second type.

An UWB system produces carrierless pulses. The shape of the pulse is in most cases Gaussian-like and its duration should be as short as possible (typical widths are 0.25-2 ns). The block diagram describing the system presented, here, is illustrated in **Figure 106**. This is the standard diagram corresponding to a time domain UWB system with two receiver channels/antennas and Low Noise Amplifiers at both transmitter and receiver channels.

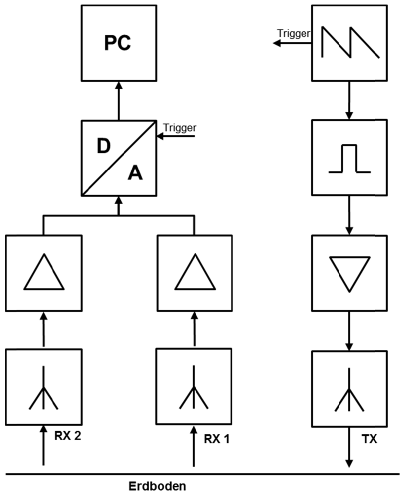


Figure 106: Block diagram of an UWB Impulse GPR with two channels.

5.1 Hardware

The key hardware component which determines the system resolution and penetration depth is the pulse generator. Two types of pulse generators were built: one with a Field Effect Transistor (FET) as principal component and the second one with a Step Recovery Diode (SRD). The performance in terms of generated pulse amplitude, width and late-time ringing has been investigated through several laboratory measurements. In general, the FET alternative gives a narrower pulse while the SRD produces a pulse with higher amplitude (see **Figure 107**). In both cases, the building blocks and feed voltage of the signal generators have been optimized to reduce the ringing. The achieved pulse width with the SRD is not totally satisfactory and there is some ringing that may obscure the cavity response. On the other hand, the FET generator produces a ~1.2 ns quite clean Gaussian pulse and little ringing. Regarding the A/D converter, a two-channel portable oscilloscope is employed as sampling device. This allows connecting and recording data from two receiver antennas simultaneously.

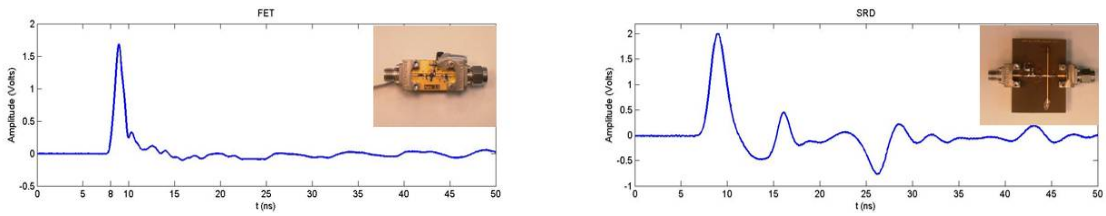


Figure 107: Generated Pulse with FET (left) and SRD (right).

5.2 Antenna element

For our particular application a bandwidth of at least 0.3 Mhz – 1 GHz is desirable. The required antenna should combine compact dimensions with satisfactory performance over the whole frequency band, high gain and stable transfer function over the whole bandwidth. A bow-tie dipole antenna, which combines an ultrawideband characteristic with compact size, seems to be a good choice. Adding a metallic shielding is a quite typical solution to isolate the system from other radiation sources, and in particular to focus and increase the gain of the antenna in the downward direction. The selected reflector for our system has an optimized trapezoidal shape to minimize the sidelobes and maximize the directivity.

Some measurements of the antenna system transfer function and the radiation pattern for several antenna-reflector distances were performed in our anechoic chamber. In general, the radiation patterns showed a good performance (gain around 10 dB) for the required frequency band (350 MHz – 1.1 GHz). As expected, the distance between dipole and reflector affects slightly to the transfer function: shorter distances enhance the high frequency gain while longer distances enhance the low-frequency contribution.

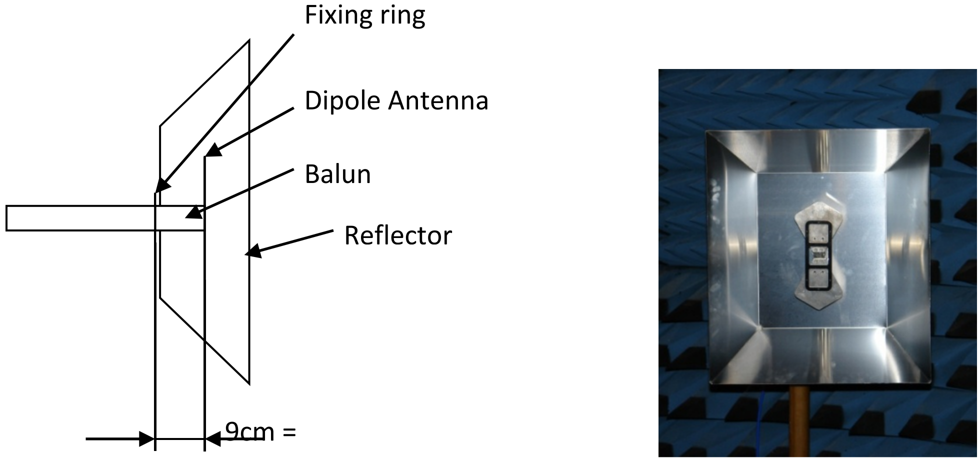


Figure 108: Antenna System with bow-tie dipole and shielding reflector.

6 PROCESSING

A relevant issue to consider is that the recorded response not only contains the scattered signal by the potential target but also other interface/soil reflections and undesirable effects from the direct coupling between antennas and system ringing. These contributions can be at least partially removed by an appropriate pre-processing of the radar data. Some standard pre-processing techniques that may help to increase the Signal to Noise Ratio and hence facilitate cavity detection are DC removal, bandpass filtering, Hilbert transform and average background removal. Another more advanced technique would be to apply a matched filter and deconvolve the free space antenna transfer function measured in the anechoic chamber. In **Figure 109** (bottom) we show a stationary radargram (amplitude and phase) taken with the GPR antenna placed next to the pressure foot during a pressure application test in a sand box where two empty plastic pipes were buried 32 cm deep. Average background removal, a bandpass filter along with Hilbert transform was applied. As we can see some events are monitored (indicated by arrows) which are visible in the radargram as well as in the curves recorded by the force and acceleration sensors.

For the post-processing, we may consider diverse techniques to successfully detect and identify subsurface voids depending on the particular circumstances.

6.1 Hyperbola fitting and travel-path difference

As shown above, diffraction hyperbolas are given by Eq. (8). Then, when the scattering hyperbolas are well resolved and either the soil permittivity or the scatterer depth is known, hyperbola fitting may be a method to obtain missing information. If none of this information is available, from the travel path difference between both channels, it is possible to estimate the propagation velocity (i.e soil permittivity) though a given layer, and thus, estimate the reflector depth. Moreover, comparing the travel paths between top and bottom reflections associated to a given scatterer, it may be possible to determine the permittivity of the scatterer from the travel times and hence, identify cavities.

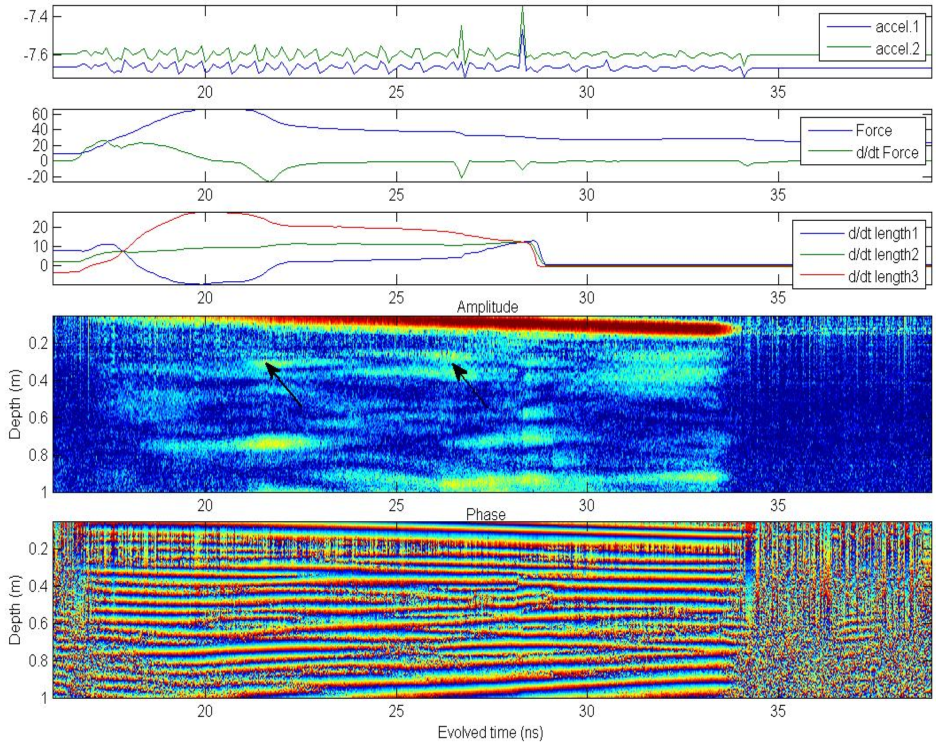


Figure 109: Continuous monitoring during pressure application.

In general, given the distances between transmitter and receivers 1 a 2 (TR1 and TR1) and the two-way travel times t_1 and t_2, the depth d of the reflector and permittivity of the soil are given by:

$$d = \frac{1}{2}\sqrt{\frac{(TR_2)^2(t_1)^2 - (TR_1)^2(t_2)^2}{(t_2)^2 - (t_1)^2}}$$
$$\varepsilon_r = \frac{c_0 t_1}{4d^2 + (TR_1)^2} \qquad (11)$$

6.2 Focusing Techniques: Migration

Migration techniques [Gazdag, 1978], [Schneider, 1978], [Stolt, 1978] are popular Fourier transform based processing algorithms to focus the above mentioned diffraction hyperbolas into the real position of the scatterer, thus improving the SNR and the target detection. To migrate the data successfully a correct estimation of the dielectric permittivity of the propagation medium needs to be done. However, as demonstrated in [Gonzalez-Huici, 2010], migration may be an indirect way of determining the soil permittivity in an iterative manner via contrast maximization of the corresponding image. Some row- and migrated-data radargrams from the experiment referred above are illustrated in **Figure 110**. Here the potential of the method to reconstruct the object real dimensions and depth are demonstrated.

6.3 Recognition Methods

Other algorithms that look for features characteristic of the scattering by cavities in 1D and 2D may be considered to help in the identification process. Some observations of special features associated to the scattering by voids have been already reported in literature [Kofman et al., 2006].

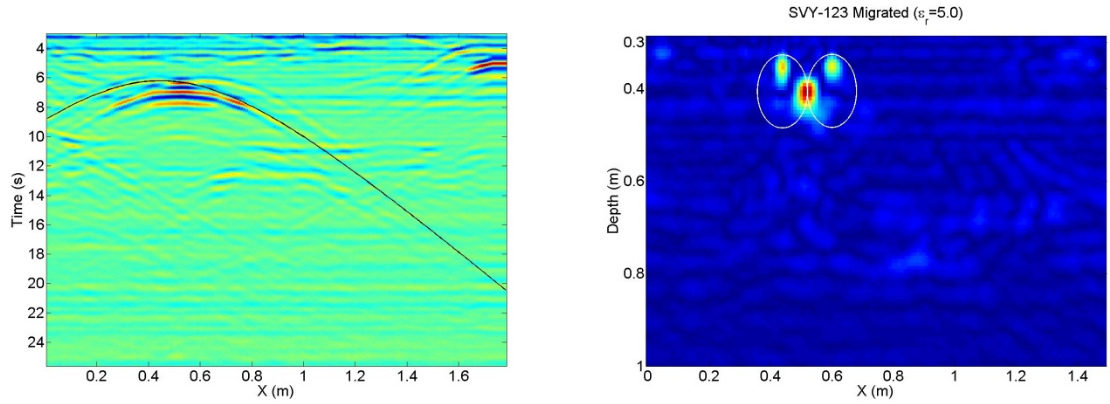


Figure 110: Scattering by two nearby empty cylinder. Raw (left) and migrated (right) data.

7 CONCLUSIONS

This paper addresses the problem of the safe setting up of supporting legs of mobile construction machines. By measuring of the load – settlement characteristic of the soil, using acceleration sensors in combination with inclination sensors the ultimate load can be estimated. These measurements are performed without any reference points. In the presence of underground cavities, however, these predictions may not be accurate. To cope with this problem the use of an adapted radar system for void detection is proposed. The system design as well as some processing techniques for cavity identification is presented.

ACKNOWLDGEMENTS

This work is supported by the Federal Ministry of Economics and Technology under Grant Agreement AIF_Nr. 15848 N. We thank the Construction Equipment and Building Material Machinery Association (FVB) and the industrial advisory group for their strong support.

REFERENCES

GAZDAG, J.: *Wave equation migration with the phase shift method*.Geophysics, 43,1342–1351, 1978.

GONZALEZ-HUICI, M. A.: *Adaptive Stolt migration via contrast maximization for GPR applications*, In Proceedings of the 6th IWAGPR 2011*,* Aachen, Germany.

KOFMAN, L., RONEN, A., AND FRIDMAN, S.: *Detection of model voids by identifying reverberation phenomena in GPR records*, Journal of Applied Geophysics, 59, 284-299, 2006.

SCHNEIDER, W. A.: *Integral formulation for migration in two and three dimensions*, Geophysics, 43, 49–76, 1978.

STOLT, R. H.: *Migration by transform*. Geophysics, 43(1), 23–48, 1978.

KONDNER, R.L., ZELASKO, J.S.: *A hyperbolic stress-strain formulation for sand.* In: Proceedings of the 2nd Panamerican Conference on Soil Mechanics and Foundation Engineering,Sao Paulo, vol. 1, p. 290-324, 1963.

BEUTINGER, P.: *Ein geotechnischer Beitrag zur Standsicherheit mobiler Baumaschinen,* Dissertation Universität Stuttgart*,* 2005.

WICHMANN, T., TRIANTAFYLLIDIS, T.: *Über die Korrelation der ödometrischen und der dynamischen Steifigkeit nichtbindiger Böden,* Bautechnik, 83(7), p. 482-491, 2006.

Kinematic Measurements

Investigation on the Performance of Low-Cost Single-Frequency GPS

Zhenzhong Su, Alain Geiger, Philippe Limpach,
Institute of Geodesy and Photogrammetry,ETH Zürich, Switzerland

Abstract
In the last few years, investigations and researches show a promising future for applying low-cost single-frequency GPS equipment to geodetic and engineering tasks (Schwieger, 2009). The low-cost GPS equipment used in our X-sense project has shown good performance. Static/quasi-static positions of daily solutions with millimeter accuracy and kinematic solutions with centimeter-level accuracy have been achieved and used to measure mass movements in alpine areas (Wirz et al., 2011, Limpach and Grimm, 2009). Besides, we would also like to exploit the real-time high accuracy capability of the low-cost single-frequency GPS equipment in the near future.

In this paper one accuracy limiting factor for GPS positioning, antenna phase center variations (PCV) will be discussed. The PCV determination approach is presented and the improvements of position accuracy by applying PCV are shown. Results of this paper show a increasing potential to achieve high-accuracy real-time kinematic positions from low-cost single-frequency GPS measurements.

Keywords
Low-cost, GPS, high precision, Phase Center Variations

1 INTRODUCTION

1.1 Error sources of GPS signal

Error sources of GPS observations are clock errors from both satellite and receiver, satellite orbit error, atmospheric delay from ionospheric and tropospheric refraction, antenna phase center variations for both satellite and receiver, multi-path, and receiver noise. By double differencing, clock errors of satellite and receiver are eliminated to large extent. For short baselines, the ionospheric and tropospheric refractions for both reference and rover are considered to be the same and most of the effects are eliminated, especially for ionospheric refraction due to the fact that ionosphere is in the upper part of the atmosphere and signals to reference and rover experience almost the same length of ionospheric travel path. When a dual-frequency GPS receiver is used, ionospheric refraction is eliminated by L1&L2 linear combination. The remaining biases from atmosphere are relative troposphere biases caused by errors of (mis-modeled) tropospheric refraction at one endpoint of a baseline relative to the other endpoint (Dach et al., 2007). These relative biases are minor for short baseline. Multi-path can be mitigated by carefully choosing the observing location.

1.2 Phase center offset and Variations

Therefore, the only bias still remaining unmodeled is receiver antenna phase center variations. Most of the geodetic grade GPS antennas have been well calibrated. The antenna offset and PCV for both satellites and receivers are available in igs08.atx file. It provides absolute elevation and azimuth dependent corrections from robot calibrations. For more details about absolute and relative phase center corrections please read (Schmid et al., 2005 & Rothacher et al., 2002). Unfortunately, no antenna calibrations are available for low-cost antennas like the Trimble Bullet III. Thus an approach to estimate a low-cost rover antenna PCV regarding to an absolute calibrated reference antenna is used to investigate the quality of Trimble Bullet III antenna.

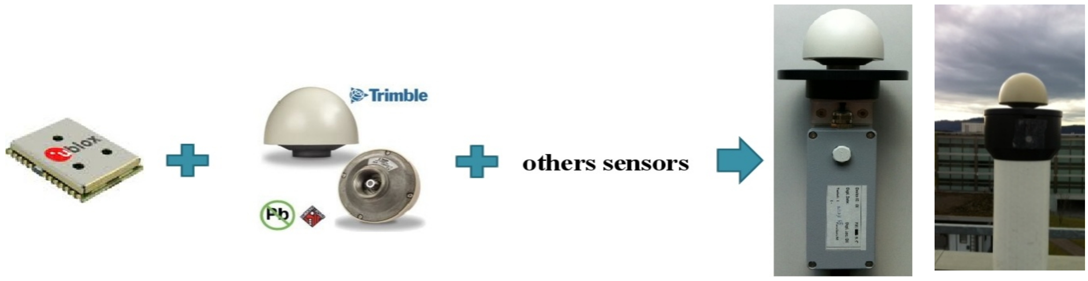


Figure 111: low-cost GPS in X-Sense project (right) & setup for PCV estimation (left)

The calculated GPS receiver position refers to a point which is called antenna mean phase center (Geiger 1988). In most case the mean phase center differs from the geometric center of antenna. Generally, a reference point is marked on the antenna and the offset values from reference point to mean phase center are provided.

Figure 112 shows the reference station ETHZ (with a trimble TRM29659.00 antenna) used for rover antenna PCV estimation. The phase center offset and absolute PCV of TRM29659.00 (marked asTRM29659.00 NONE in igs08.atx) are obtained from igs08.atx.

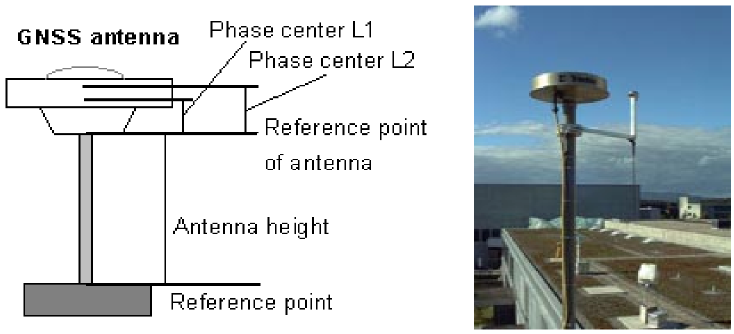


Figure 112: TRM29659.00 antenna of reference station ETHZ (©AGNES, SWISSTOPO)

Antenna phase center variations are estimated with respect to mean phase center. It is used to describe and model the phenomenon that GPS signals from individual satellites of different directions arrive at the receiver antenna at different points, which means each signal, depending on the angle of arrival, carries a corresponding path delay. The antenna PCV can be interpreted as the corrections from the signal surface to the mean phase center (see **Figure 113**).

The antenna PCV can be represented by spherical harmonic functions:

$$\Delta\varphi(\alpha,z)=\sum_{n=1}^{n_{max}}\sum_{m=0}^{m_{max}}\tilde{P}_{nm}(\cos z)\left(a_{nm}\cos m(\alpha-\alpha_0)+b_{nm}\sin m(\alpha-\alpha_0)\right) \quad (1)$$

where $\tilde{P}_{nm}$ are normalized associated Legendre functions of degree n and order m, a_{nm} and b_{nm} are the coefficients of the harmonic series development, α is the azimuth, α_0 is antenna orientation and z is the zenith angle of the satellite line of sight.

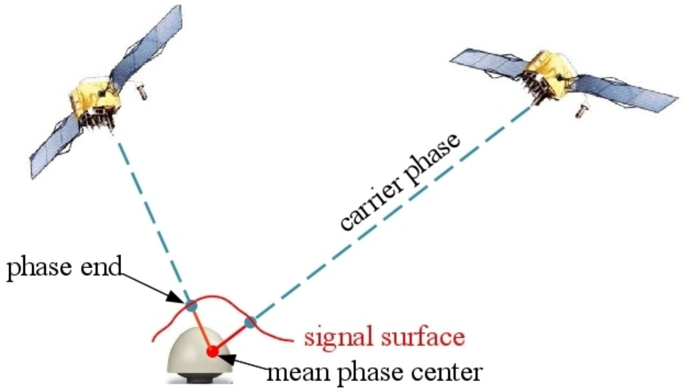


Figure 113: illustration of PCV

2 APPROACH

2.1 Estimation of phase center variations

The PCV talked about in this paper are estimated under nearly "zero baseline" condition and the reference antenna's absolute PCV value is known. The rover station with low-cost equipment is set up near the reference station ETHZ. The length of baseline is 27 meters and double difference observations are formed to eliminate clock errors from satellite and receiver, ionospheric error and tropospheric error. By choosing a measurement site with no conductive obstructions above antenna's horizon, one might reduce the risk of massive multi-path. However, ground multi-path could still occur occasionally. The differenced carrier phases in normal equation (2) are guaranteed to contain only biases induced by PCVs.

$$\Delta\varphi = \Delta\sigma + \Delta\phi + \lambda\Delta N + e \qquad (2)$$

where $\Delta\varphi$ is double differenced carrier phase, $\Delta\sigma$ is differenced geometric distances between satellites and receivers, Δ represents the PCV , ΔN is ambiguity integers, e is noise. The flow chart of PCV estimation procedure is shown in **Figure 114**.

In the first step, the mean phase center of rover antenna is calculated from 24 hours static observations by Bernese software. In the second step, positions of reference station and rover station are both fixed as known parameters. They are imported into double difference equations. With precise satellite positions the second term, geometric distance $\Delta\sigma$ in equation (2) can be calculated as well. Then the program will solve the ambiguity integers and store them as known parameters. Finally the only unknown term is PCV and it is estimated by standard least square adjustment.

A Bernese Processing Engine (BPE) script "PCVEST" (see **Figure 115**) is designed for PCV estimation based on Bernese Software 5.0. The outputs are the estimated coefficients of the harmonic series in formula (2) and they are stored in Bernese format.

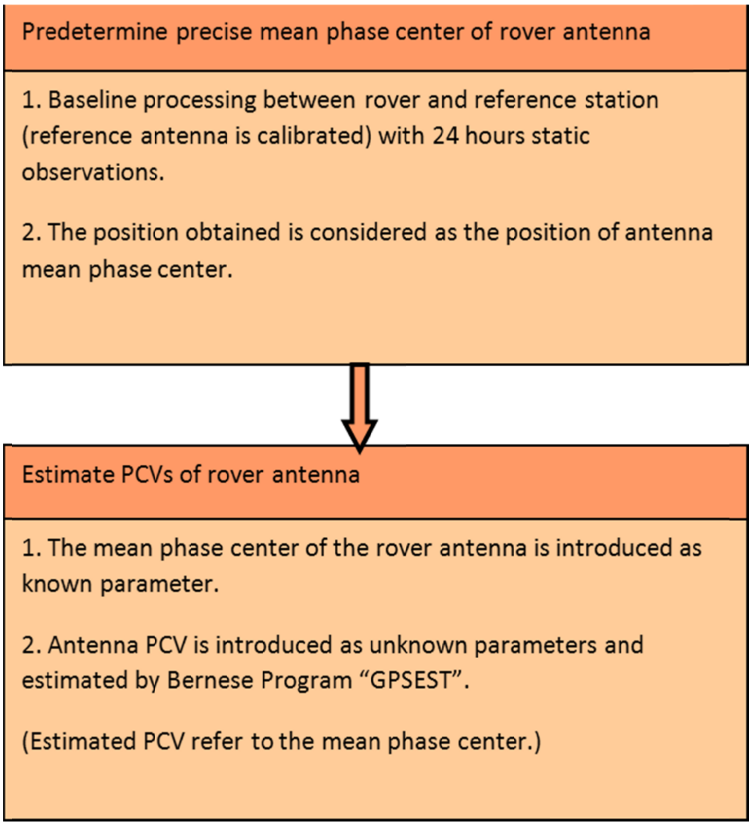


Figure 114: PCV estimation procedure

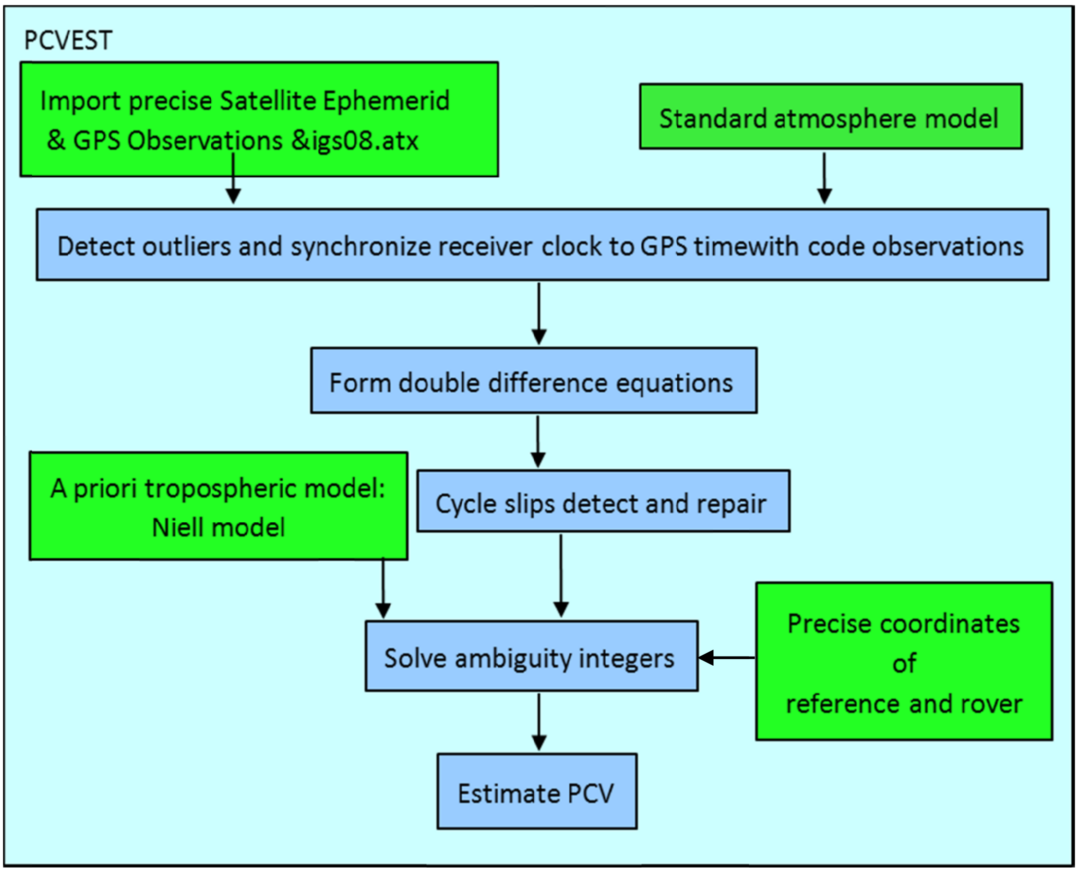


Figure 115: BPE script: PCVEST

3 RESULTS

3.1 Estimated Antenna PCV for Trimble Bullet III

The first kind of PCV (**Figure 116**) is estimated with a fixed antenna orientation and the north gap of GPS constellation is not covered. Thus the estimated PCV for north and near north directions are less reliable. However, the estimated PCV can be used to correct GPS measurements as long as the antenna keeps the same orientation, since there is no satellite in the gap and PCV values for those directions are not used at all. In this situation, the antenna offset is not of interest neither, since the PCV corrections are always applied according to the mean phase center.

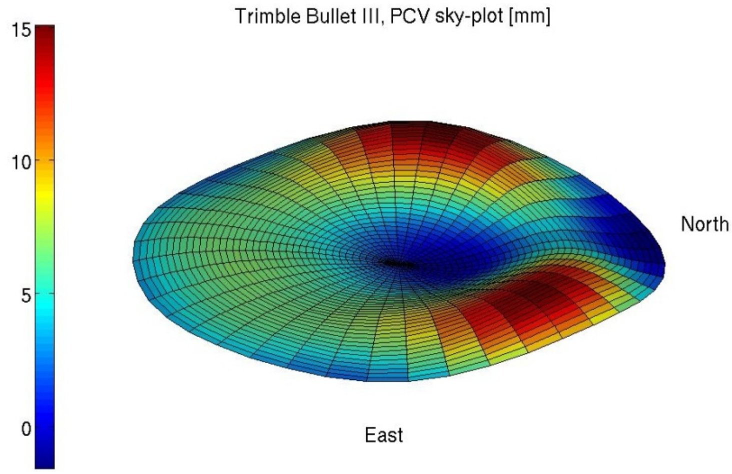


Figure 116: Antenna PCV (the north gap is constrained with small weight during estimation)

The second kind of PCV will be estimated by rotating the antenna, so that the north gap is covered and antenna PCV pattern with all direction coverage is obtained. With the help of antenna offset, reference mark and north mark of the antenna, it can be used to correct data for the same kind of antenna even though the orientation is changed. This will be done with rotating robot from GGL soon.

3.2 Positioning accuracy improvement

The influence of PCV is prominent especially for sub-daily coordinate resolution because of the satellite constellation change. To investigate the magnitude of PCV influence, the estimated PCV for rover antenna is applied to correct GPS observations measured by rover receiver. 2-hour solutions are calculated with and without using PCV corrections. **Figure 110** shows that by using PCV corrections the calculated antenna positions are closer to real value. Each point corresponds to a rover position calculated by 2 hours GPS observations. The position differences against real antenna position are plotted in the graph. Numerical results are summarized in **Table 15.**

Table 15: 2-hour solutions from Day 298

Axis	RMS[mm] (solution without PCV)	RMS[mm] (solution with PCV)	Improvement [%]
N	2.9	1.0	64
E	1.4	0.7	45
U	4.3	2	53

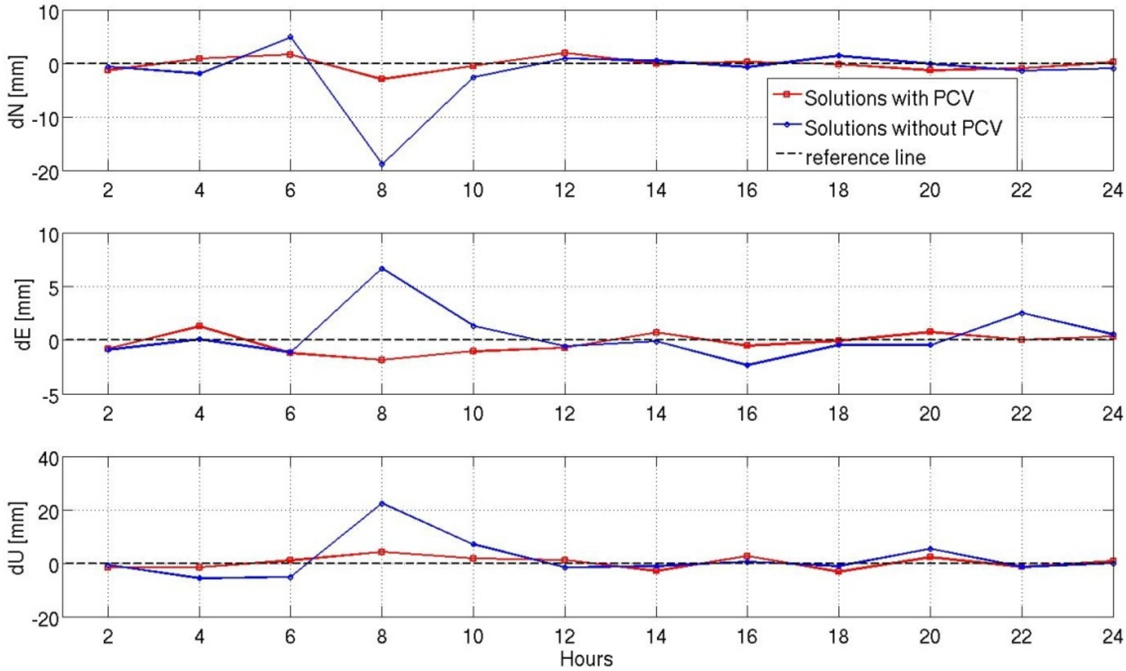


Figure 117: 2-hour solutions from Day 298

4 CONCLUSIONS

From one side, it has been shown that by estimating PCV and applying PCV, GPS positioning with short observational timespan can be improved. From the other side, the estimated PCV for Trimble Bullet III is within 1.5 cm and for most directions is within 1 cm. It is small enough that even without applying PCV in real time positioning case, the ambiguity integers can be fixed correctly, which indicates that using low-cost single-frequency GPS equipment for real time positioning is feasible.

ACKNOWLEDGEMENTS

The authors would like to thank Nano-Tera, a swiss national foundation program which co-finances the Project X-Sense. And sincere appreciation is given to Prof. Dr. Markus Rothacher for his support regarding Bernese software and Bernhard Buchli from TIK, ETH Zurich for hardware support.

REFERENCES

DACH, R., HUGENTOBLER, U., FRIDEZ, P., MEINDL, M.: *Bernese GPS Software Version 5.0.* Astronomical Institute, University of Bern, 2007.

BUCHLI, B., SUTTON, F., BEUTEL, J.: *GPS-equipped Wireless Sensor Network Node for High-accuracy Positioning Applications*. Proc. of 9th European Conference on Wireless Sensor Networks, 2012.

GEIGER, A.: *Modelling of Phase Centre Variation and its Influence on GPS-Positioning.* GPS-Techniques Applied to Geodesy and Surveying, Lecture Notes in Earth Sciences, Springer-Verlag: 210-222, 1988.

LIMPACH P. and GRIMM D.: *Rock glacier monitoring with low-cost GPS receivers*. Abstract Volume 7th Swiss Geoscience Meeting. Neuchatel, Switzerland, November 2009.

ROTHACHER, M.,: *Receiver and Satellite Antenna Phase Center Offsets and Variations.* Position Paper of the "Antenna Session" IGS Workshop. Ottawa, 2002

SCHMID, R., ROTHACHER, M., THALLER, D., STEIGENBERGER, P.: *Absolute phase center corrections of satellite and receiver antennas*. Impact on GPS solutions and estimation of azimuthal phase center variations of the satellite antenna.GPS Solutions, 9, 4, 283-293, 2005.

SCHWIEGER, .V: *Accurate High-Sensitivity GPS for Short Baselines.* FIG Working Week, Eilat, Israel, 03-08.05.2009.

WIRZ, V., GRUBER, S., BEUTEL, J., BUCHLI, B., LIMPACH, P., GEIGER, A.: *Temporal characteristics of different mountain cryosphere-related debris slope movements.* ISPRA (Italian National Institute for Environmental Protection and Research), Rome Italy, 2011.

Modular Imaging Total Stations – Sensor Fusion for High Precision Alignment

Stefan Hauth, Martin Schlüter, Florian Thiery
i3mainz - Institute for Spatial Information and Surveying Technology,
Fachhochschule Mainz - University of Applied Sciences, Mainz

Abstract
Initialized in 2009, the Institute for Spatial Information and Surveying Technology (i3mainz), Mainz University of Applied Sciences, forces research towards modular concepts for imaging total stations. On the one hand, this research is driven by the successful setup of high precision imaging motor theodolites in the near past; on the other hand it is pushed by the actual introduction of integrated imaging total stations to the positioning market by the manufacturers Leica Geosystems, Sokkia, Pentax, Topcon and Trimble.

Modular concepts for imaging total stations are manufacturer independent to a large extent and consist of a particular combination of accessory hardware, software and algorithmic procedures. The hardware part consists mainly of an interchangeable eyepiece adapter offering opportunities for digital imaging. An easy assembly and disassembly in the field is possible allowing the user to switch between the classical and the imaging use of a robotic total station. The software part primarily has to ensure hardware control, but several level of algorithmic support might be added and have to be distinguished. Algorithmic procedures allow reaching several levels of calibration concerning the geometry of the external digital camera and the total station. Here the resulting resolution capacity of our sensor fusion and also the accuracy of the system are presented based on examples. We deliver insight in our recent developments and quality characteristics.

The sensor fusion between camera and polar measuring system allows detecting and measuring different types of targets with high precision. MoDiTa is used to calibrate inclination sensors and to control the long-term stability of laser and tripods.

Keywords
Imaging total station, sensor fusion, image processing

1 INTRODUCTION

The combination of polar measurement systems using modern digital industrial cameras keeps growing in the area of measurement technology. Some commercial producers of polar measurement systems like Pentax, Trimble, Topcon or Leica Geosystems now offer this combination as an integrated system; cf. HAUTH/SCHLÜTER (2010). In addition to the aforementioned manufacturers, research institutes also develop solutions and applications of modular imaging total stations, which carry a camera in front of the eyepiece. A prototype of a theodolite combined with a digital camera, developed at the i3mainz, shows the successful use of technical precision measuring SCHLÜTER ET AL. (2009). The MoDiTa – Modulares DigitalkameraTachymeter (modular imaging total station) of i3mainz is a consistent continuation of the previously mentioned theodolite combined with a digital camera on a modular base, which is examined in the BMBF-funded (Federal Ministry of Education and Research) research project "Modulare DigitalkameraTachymeter".

WASMEIER 2009 from the University of Technology München presents the prototype IATS2 based on total station Leica Geosystems TCRA1201. IATS2 replaces the reticule with a CMOS-chip and removes the eyepiece. The camera is fixed integrated in the total station and complicates an adaptation for different applications or the changing back to a “classic” total station. An advantage is the good compactness of the system.

BÜRKI 2010 from the Institute of Geodesy and Photogrammetry at the Swiss Federal Institute of Technology in Zürich shows with the system DAEDALUS a combination between digital camera and total station, too. This replaces the eyepiece with a CCD-chip and requires no additional optical components between chip and reticule. The image is no longer imaged precisely in the plane of the reticule but in the plane of the CCD-chip. This small shift is eliminated by the focus of the total station (for objects up to 13 meters) or by adding an additional lens in front of the telescope.

MoDiTa from the i3mainz demand only small changes by the total station. The eyepiece is fitted with a digital camera, which as well as the polar measurement system is controlled by a software module. In this case the digital camera takes images directly from the reticule plane, so no change in the optical path in the telescope prevails. By that modular structure, which requires only a few special products, an easy changing between classical eyepiece and the MoDiTa eyepiece in the field is possible, which causes a large variability (**Figure 118**). The modular concept of MoDiTa allows replacing several hardware components, e.g. different telescopes (theodolite and total station), several cameras with different resolutions, and various magnifications of the reticule. Thus, this flexible system exploits the respective strengths and weaknesses of each component optimally. Those advantages and application possibilities of the hardware components are discussed in the following chapters.

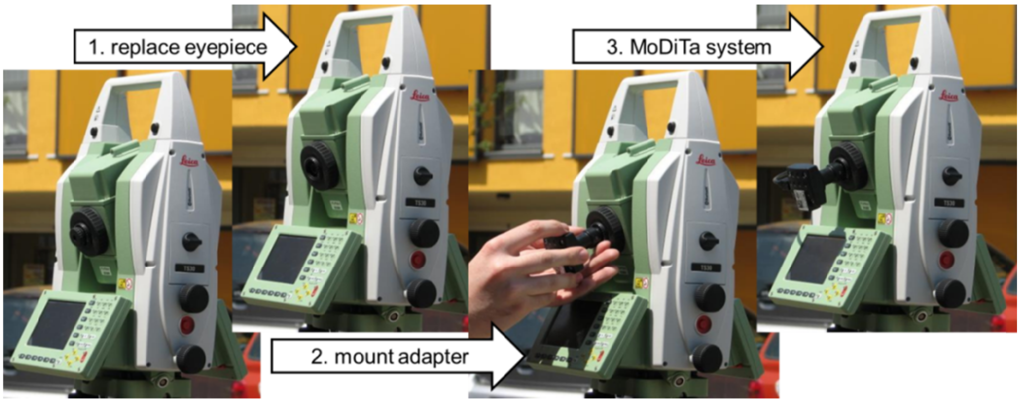


Figure 118: Imaging total station (Leica Geosystems TS30), mounting the adapter

2 THE SENSOR FUSION

The MoDiTa system consists of a sensor fusion between a polar measuring system and a modern industrial camera. The following chapters are describing the configuration and the reachable accuracy and resolving capacity depending on different resolutions of the used cameras.

2.1 Configuration and components of MoDiTa

On the hardware side the specific MoDiTa component is the special eyepiece adapter. This adapter replaces the normal eyepiece and is used as a mount for camera and lens. Camera and optics are standard components and no special developments for MoDiTa. So we use S-Mount and C-Mount ports for camera and optics, and serial ports respective USB for power and data transfer. This offers the advantage that the components are inexpensive to purchase and can be replaced quickly and easily. An optional extension for a focus control is possible. **Figure 119** shows the schematic structure of the MoDiTa eyepiece in cross section for mounting the camera and polar measuring system. The standard lens with various fixed apertures is fixed by S-Mount in the eyepiece adapter. This eyepiece holder is connected to the bayonet mount for the eyepiece of the polar measurement system (here Leica Geosystems). The length of the adapter determines the reproduction scale and controls the field of view. The used cameras are producer repetition parts with C- or CS-Mount and data respective power communication by USB 2.0.

MoDiTa can use different polar measurement systems, cameras, resolutions, optics and magnifications of the reticule. This allows to combining the different advantages of the components. At the moment, the MoDiTa adapter can be used with current Leica Geosystems total stations and theodolits and some Kern theodolits. This allows using the Leica Geosystems TM5100 with MoDiTa. TM5100 is an industrial theodolite with high angular measurement accuracy without an EDM. The advantage of a theodolite telescope is the lack of optical components in the telescope and so less disruptive effects of laser beams. Total stations like TS30 and TCRM1103 are flexible devices with EDM and automatic target aiming.

The adapter of MoDiTa can mount C- and CS-Mount-cameras and so different kind of chips. For most applications colour and monochromatic CMOS-chips are useful, but some tasks require a special dynamic range. In area of extreme exposure is it helpfully to use a HDR-chip. THIERY (2011) shows a benefit of HDR-cameras by observing laser beams.

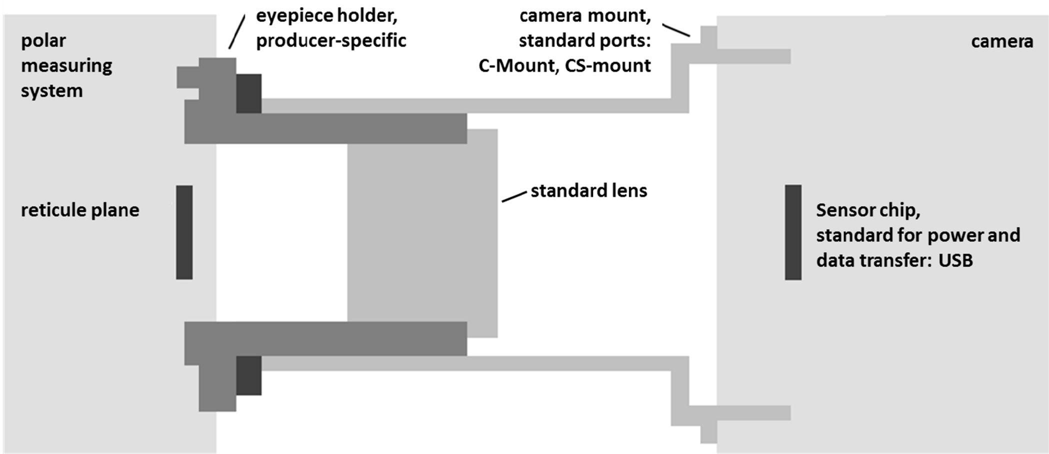


Figure 119: Schematic cross-section of the MoDiTa eyepiece

The resolution can be controlled by chip-size and settings like binning or subsampling. For the maximal accuracy and resolving capacity MoDiTa uses actually a camera resolution of 2560 x 1920 pixels. The disadvantage of this setting is that USB 2.0 limits here the transfer rate and the frame rate. That is no problem by static or slowly moving targets (6 fps). A resolution of 1280 x 960 raises the frame rate of 19 fps, but the resolving capacity and accuracy decrease slightly (chapter 2.3).

Actually MoDiTa provides two kinds of field of views. The full-view images the complete reticule plane and the magnification-view scales up the centre of reticule plane (**Figure 120**). The difference between full-view and magnification-view is about 76%. In principle the magnification-view has a better accuracy and resolving capacity.

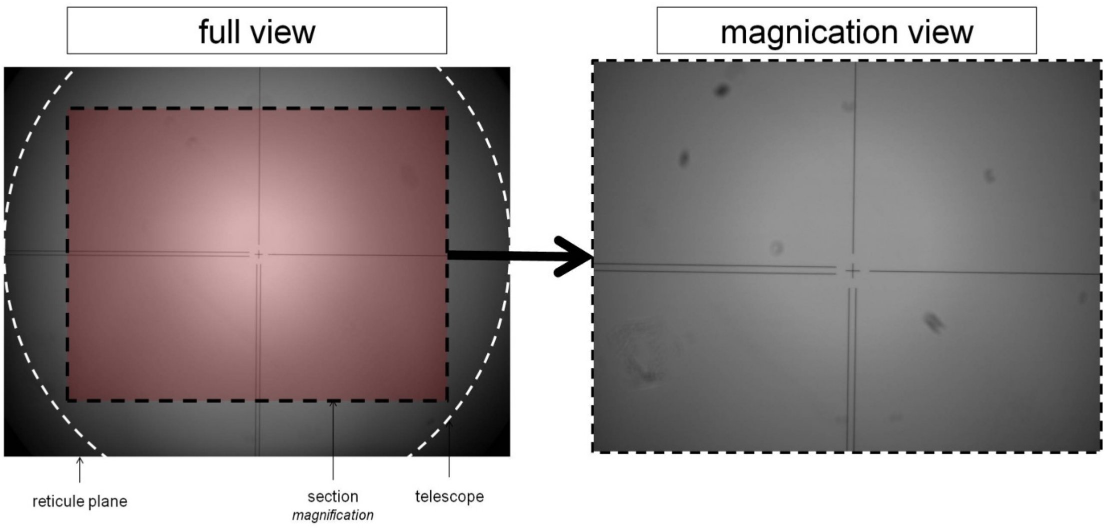


Figure 120: Field of views

2.2 Calibration of field of view

The calibration of field of view allows transforming every pixel coordinate to one angular position. MoDiTa can measure all targets in the field of view without boning with the crosshair (**Figure 121** right). The field of view is determined by a self-calibration. The self-calibration scheme using the photogrammetric collinearity relationship in addition of the distortion parameters is conform to SCHLÜTER ET AL. (2009).

For the self-calibration, a fixed target is repeatedly measured, that the target is well distributed over the calibrating area (**Figure 121** left). Every measurement is forced angular positions, resulting two correction equations, each new combination of reading and image coordinates. The angular readings following mathematically from the pixel coordinates of the reference cross hairs.

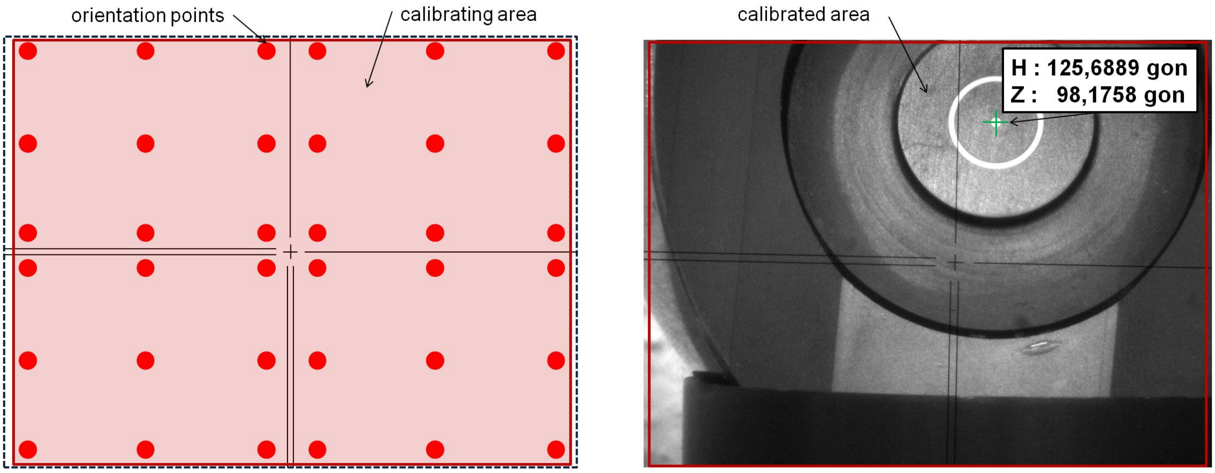


Figure 121: Self-calibration (left) and transforming pixel coordinate to angular position (right)

Beginning with the transition from a levelled local coordinate system to a telescope fixed coordinate system, the parameters are provided solely by the theodolite side compensator and graduated circle read out. The subsequent transition from the telescope fixed coordinate system to the camera system by parameters determining in a self-calibration. The photogrammetric collinearity relationship is chosen as a model for the transformation between direction vectors the telescope system and the image coordinates of digital images. As part of the self-calibration with a fixed scale factor, the angle of rotation matrix (R_K), the camera constant (c_K) and the coefficients of distortion polynomials (Δx, Δy with image coordinates X_B and Y_B) are determined as unknown parameters in a least-square adjustment. Radial symmetric distortion, caused by refraction changes in the lenses of the lens, provides the greatest impact among the mentioned distortions. The radial asymmetrical distortion and tangential distortion be caused by a decentring of the lenses in the lens and is less than the radial symmetric distortion. As the last part of the distortion affinity and angular affinity describing deviations from orthogonality and different scales of the image coordinate axes. The result of the least-square adjustment is a transformation matrix with 13 parameters (3 rotations, 2 image coordinates of the principal point, 1 calibrated focal length, 3 radial symmetric distortions, 2 radial asymmetrical and tangential distortions and 2 affinity and angular affinity).

2.3 Accuracy

The MoDiTa system allows the use of several components. There are two kinds of field of views, different camera resolutions and total stations. The accuracy of MoDiTa-Systems is determined by the self-calibration on two collimators. The self-calibration is measured with 36 observations in two-face-measurements. Exemplary the **Table 16** shows the results of the least-square adjustment. The used hardware are a Leica Geosystems TS30 total station with an angular measurement accuracy of 0,5’’ (0,15 mgon, manufacturers' instructions) and a monochromatic CMOS-chip. The results of **Table 16** compare two resolutions (2560 x 1920 and 1280 x 960) and two fields of views.

The standard deviations (std.) of all measurements are not significantly different. But the residuals are larger in the full-view significantly than the magnification-view. The differences between the two resolutions are small. Both the standard deviation and the residuals are in the same range.

Table 16: Results of least-square adjustment (self-calibration), 36 observations

	full-view			magnification-view		
2560 x 1920		dHz [mgon]	dV [mgon]		dHz [mgon]	dV [mgon]
	min	-1,0	-0,4	min	-0,1	-0,2
	max	0,7	0,6	max	0,3	0,1
	std.(1σ)	0,20	0,10	std.(1σ)	0,05	0,05
	1 Pixel = 0,6 mgon			1 Pixel = 0,36 mgon		
1280 x 960		dHz [mgon]	dV [mgon]		d□z [mgon]	□V [mgon]
	min	-1,0	-0,5	min	-0,2	-0,3
	max	0,9	1,0	max	0,3	0,2
	std.(1σ)	0,20	0,10	std.(1σ)	0,06	0,06
	1 Pixel = 1,3 mgon			1 Pixel = 0,7 mgon		

The best accuracy for a single observation is magnification-view and the resolution of 2560 x 1920. The resolution of 1280 x 960 provides a similar good result. The adjustment allows to reach the accuracy of the total station (0,5’’ = 0,15 mgon, manufacturers' instructions). The full-view has a maximum residual of 1,0 mgon and the standard deviation of the adjustment is in the range of the double accuracy of the total station.

3 APPLICATIONS

The flexibility of MoDiTa allows using the system in many different scopes of duties. The possibility of collimation measurements and their detections with MoDiTa or an imaging theodolite build by i3mainz, at both collimator crosses, as well as collimating laser beams offer, with different hardware capabilities of a wide variety of applications. The system enables the directional measurement to a moderately moving target, too. This represented target in this paper is a laserdot and a cross of a collimator.

The MoDiTa allows the use of non-contact measurement techniques including image processing and auto-collimation aiming within a calibrated visible range. Through a rapid self-calibration (ca. 2 minutes), followed by automated analysis of those and immediate readiness for measurement in the calibrated range are special tasks, such as the adjustment and calibration of laser terminals during the manufacturing process, see SCHLÜTER ET AL. (2009), very quickly solved. Azimuth determinations by astro-geodetic observations of stars with the given hardware and software are also available. Classical tasks of geodesy, such as monitoring (here landslides, see REITERER (2009) or the documentation of reference ¬ points for laser scanning and photogrammetry, the sensor-fused system also triggers using image processing. MoDiTa also enables the detection of changes in direction, such as tilting or twisting of various hardwares (eg. sensors). In the following exemplary applications which results in projects of the i3mainz, like the calibration of an electronic level and stability testing of a camera tripod through MoDiTa and image processing demonstrated.

3.1 Scale calibration of inclination sensors

Inclination sensors are gaining in importance in the field of geodesy. The electronic collection of data and the high accuracy of small measuring instruments provide in shortest time highly accurate data. The measurement of the slope provides the detection of small movements such as for deformation measurements. The determination of the scale of two inclination sensors (Leica Geosystems NIVEL220 and NIVEL230) is carried out e.g. by AZAR (2009) on behalf of the Wasser- und Schifffahrtsamt Aschaffenburg.

The examination of an inclination sensor, information regarding the producer, takes place in three stages: determining the zero point error, scale error and the detection of the influence of temperature changes. The system of i3mainz is used here only to determine the scale error. The sample is placed on a for flatness granite checked slab, parallel to the axis, compare to **Figure 123**. The combination of sample and collimator is a parallel arrangement of the telescope reaches the level. The granite base is mounted on a tilt table which can be tilted by micrometre screws. With the help of those, putting the sample into the previously determined zero point and gradually tilts the tilt table. Thus, the desired inclination of the level to the actual reading of the level is compared. The MoDiTa-system observes the collimator after every tilt of the table by the micrometre screws. Every observation consists of 36 single observations like a self-calibration (chapter 2.3). This zenith angle performs the tilt angle of the table (inclinometer and collimator).Parallel to the registration of the zenith angle the reading of the inclination sensor (X-and Y-axis separately) is detected.

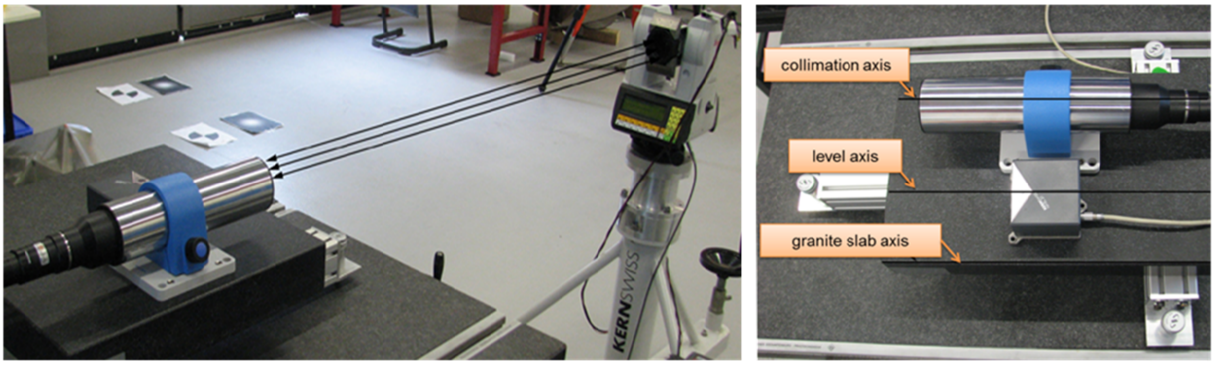


Figure 122: MoDiTa observes collimator (left), collimator and digital level fixed on a tilt table (right)

For this application the MoDiTa-system uses a Leica Geosystems TM5100 theodolite, the camera-resolution 1280 x 1024 and the magnification-view. Magnification-view and resolution is in the range of the angle accuracy from theodolite (0,5'' = 0,002 mrad, manufacturers' instructions). Both test objects (Leica Geosystems NIVEL220 and NIVEL230) have an accuracy of 0,005 mrad (manufacturers' instructions) and thus MoDiTa is suitable for checking the inclination sensors.
The scale is calculated from zenith angle (reference values) and tilt angle of the inclination sensor. The zenith angles are corrected by an average offset, because the collimation axis is not strictly parallel to the surface of the granite base. **Figure 123** shows an exemplary result of alignment an inclination sensor Leica Geosystems NIVEL220.

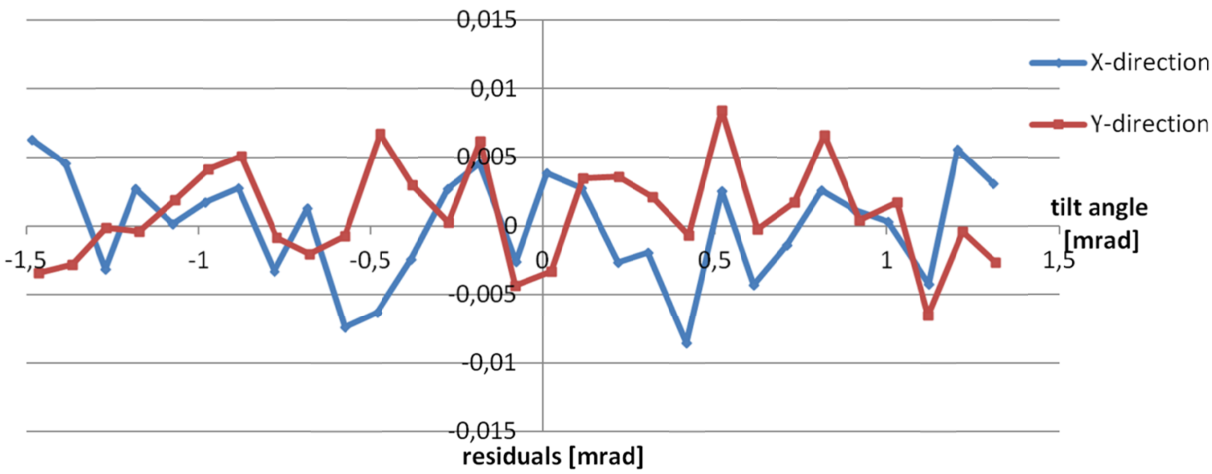


Figure 123: Exemplary result of scale alignment

Here the advantages of MoDiTa are the high repeatability and measurement speed. MoDiTa allows a large number of independent measurements in a relatively short time (36 observation in max. 2 min) with a high accuracy (0,002 mrad corresponds 0,002 mm/m). The measurement has a high degree of automation and a long-term observation is possible.

3.2 Long-term stability of tripods and laser beams

In the precise measurement tripods needed, which have only small drifts in the horizontal direction and vertical angle over a longer period of time, to store testing samples stably. In precision measurement therefore stable tripods made of steel or aluminium are preferred. A long-term stability study intended to show, whether a photo-tripod is suitable for use in precision measurement. Purpose is to detect differences in stability between the tripod head and tripod legs.

Two collimators are used as a target of the imaging total station (**Figure 124**). On the left a collimator is shown, which serves as a reference and mounted on a stable tripod by KERN. On the right hand a further collimator is displayed, where the sample is fixed (here: tripod including tripod head). A successfully self-calibration of the system, is followed by the actual inspection and sub-pixel registration of the horizontal direction and vertical angle of the collimator cross. The analysis covers for each measurement period in a time period of at least 20 minutes (t_0=0min). Starting with a reference measurement to a collimator on the stable KERN-tripod, the collimator crosshair fixed on the sample is being tracked at different time intervals and concluded with a reference measurement. The evaluation is done by calculating residuals in the vertical direction of the reference measurements for each measurement period in order to detect movement on the instrument.

Figure 124: Imaging theodolite aiming a collimator on a Linhof tripod (right)

Based on the results of THIERY (2011) the best configuration for capturing laserdots is a high dynamic range (HDR) camera by using a theodolite without EDM unit to avoid reflections. Theodolite measurements are made only at rest, during a moving target crosses the reticule plane. If the theodolite rests, the direction to the target is registered with high precision, if the target is within the calibrated field of view. For large-scale tracking of a target, the theodolite software supports gradually readjusting. However, in the movement phase, no high-precision measurements are possible. The tracking speed of a moderate moving collimation target corresponds to the recording speed of the used camera.

Figure 125 shows exemplary an image through the eyepiece, along with the recorded measurement collimation image of a laser beam and a resulting image recorded with an HDR camera. If a reference image of the reticule plane is captured and stored, the visibility of that is not necessary to detect the direction. It is only important for detecting movements of the camera. So, matching processes offers the projection of the reticule plane into the image of the laserdot. The centre coordinates of the laser spot can be calculated with subpixel accuracy using digital image processing on an elliptic operator.

Out of the centre coordinates, the theodolite and compensator readings and the data from the self-calibration the (virtual) direction can be determined arithmetical to the laser spot centre.

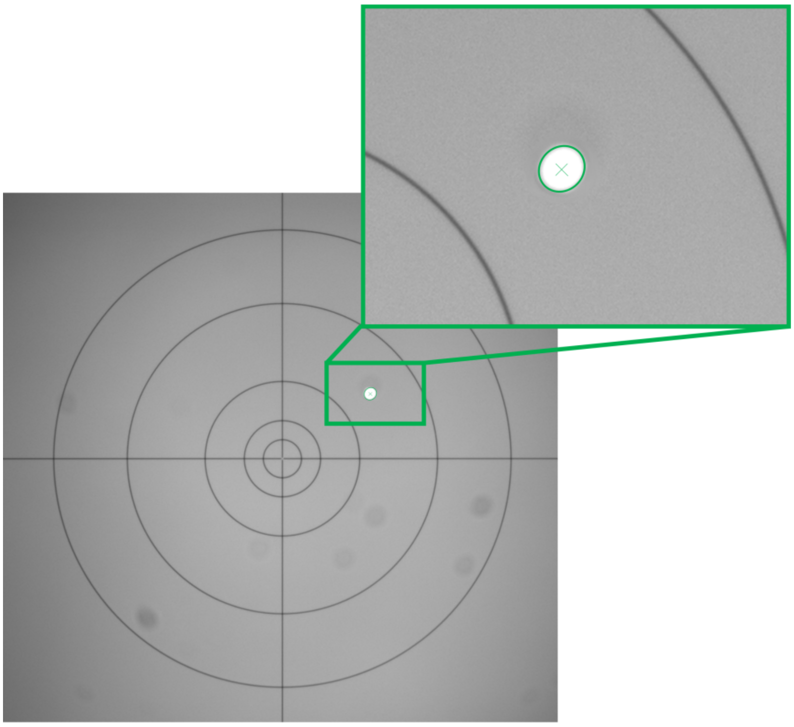


Figure 125: Captured laserdot with reticule

4 CONCLUSIONS

In this paper, we discuss the MoDiTa system, its configuration and some applications. We concentrate on the accuracy and resolving capacity and show that a 1,3 megapixels camera is sufficient to exhaust the angles accuracy of a Leica Geosystems TS30. As the key contribution we present the calibration scheme using the photogrammetric collinearity relationship in addition of the distortion parameters, too. Moreover, we depict that our modular system is able to accomplish and various applications using the best components depending on the problem. In the future, more and more applications for our MoDiTa system should be found. The presented results show a great potential of the given hardware and software.

ACKNOWLEDGEMENTS

The authors would like to thanks Emil Azar and Florian Sauerwein for performing the measurements and Bernd Heppel and Dieter Hahner for hardware support.

REFERENCES

AZAR, E.: *Automatisierte Prüfung von digitalen Libellen mit einem Digitalkameratheodolit*. Bachelorarbeit, FH Mainz, B0035, 2009.

BÜRKI, B./ GUILLAUME, S./ SORBER, P./ OESCH, H.-P.: *DAEDALUS: A Versatile Usable Digital Clip-on Measuring System for Total Stations*. International Conference on Indoor Positioning and Indoor Navigation (IPIN) 2010, Zürich, 2010.

HAUTH, S., SCHLÜTER, M.: *Digitalkameratachymeter - Einsatzmöglichkeiten für modulare und integrierte Systeme*. 30. Wissenschaftlich-Technische Jahrestagung der DGPF 2010, 3-Ländertagung, Wien, Publikationen der DGPF Band 19, 2010.

REITERER, A: *Leica Geosystems Sonderanfertigung Tachymeter 1200 mit integriertem Kamerasystem.* Ausgewählte Forschungsprojekte (Part I): I-MeaS – Intelligent Measurement System, Leica Geosystems Austria GmbH, Wien, 2009.

SCHLÜTER, M., HAUTH, S., HEß, H.: *Selbstkalibrierung motorisierter Digitalkameratheodolite für technische Präzisionsmessungen*. ZfV Jg. 134, 1/2009, p. 22-28, 2009.

THIERY, F.: *Eignung aktueller Industriekameras für modulare Digitalkameratachymeter*. Bachelor-arbeit, FH Mainz, B0066, 2011.

WASMEIER, P.: *Videotachymetrie - Sensorfusion mit Potenzial*. AVN 7/2009, p.261-267, Wichmann Verlag, Heidelberg, 2009

Active Prism for Total Station Measurements

Sebastian Horst, Ilka von Gösseln
Geodetic Institute, Leibniz Universität Hannover, Germany

Abstract
Recent years marked the development of classical theodolites into modern multi sensor total stations. These offer the possibility of motorized control and automatic data collection. Nevertheless, the efficiency and other benefits of these developments are limited by the still passive prisms used in many surveying tasks. Due to the limited aperture angle of a prism of (~ 30°), a time intensive manual alignment to the current total station has to be performed that interrupts the automated workflow. The solution to this problem by using 360° prisms is expensive and associated with loss of accuracy. A way to overcome these disadvantages is to introduce an automatic prism alignment.

This paper presents the development of a low-cost feasibility study for an active rotary prism based on a typical circular prism. The prism is mounted on the motor shaft of a stepper motor. A microcontroller platform is used to calculate and perform the automatic alignment. A wireless connection guarantees a flexible data communication between the total station and the active prism.
With given position and orientation of both - the prism and the total station - the alignment is a straightforward mathematical task. Without this prior information, the alignment has to be carried out on the basis of sensor measurements by compass and GPS. Other additional sensors can be attached to the presented device, too. A temperature sensor is used to deliver atmospheric corrections to increase the distance measurement accuracy.

The paper concludes with test measurements that reflect the capabilities of the new active prism. Based on efficiency analysis of an exemplary network measurement, the advantage of the active prism is shown.

Keywords
Active prism, total station network, efficiency, sensor network, microcontroller, low-cost sensor

1 MOTIVATION

Although GNSS is nowadays available, total station measurements are still used in many geodetic networks, especially to achieve high point accuracies. This is reflected in newer device developments. Modern total stations with motorized control allow a quick and thus efficient data collection. Despite the constant further development of sensors, the network measurement is associated with a high logistic effort. Geodetic networks are based on combined direction and distance measurements among fixed points. To increase the accuracy and reliability, cross-measurements are used on adjacent points, as well as redundant measurements to the same points from different stations. In case of a station-changing the limiting aperture angle of the target prisms requires realignment. This orientation to the current total station position is done manually until now. Of course, this lack of efficiency can be overcome by the adaption of modern 360° prisms. However, the use of these devices has the disadvantage of a lower accuracy (Leica 2011). Further on, in engineering offices there are usually many standard circular prisms available and an exchange of these prisms is very expensive.

Optimization approaches to minimize the driving times within a terrestrial network respectively maximize the efficiency of the network measurements have already been published (Rehr et.al. 2011). In the following, a further optimization approach is presented. The introduction of communication structures between prism and total station and the use of actuators make an automatic alignment of the prisms possible. At this point it must be noted, that the study so far is limited to a horizontal orientation. An extension on the vertical axis is possible by similar funds.

2 ACTIVE PRISM - HARDWARE COMPONENTS

The application of an active prism alignment should be viewed in the context of geodetic sensor networks (Horst et.al. 2011). Communication structures between a group of prisms and total stations have to be implemented for this task. Sensor nodes, respectively active prisms, have restricted resources like link budgets (reachable ranges of communication), power management of the autarkic units and low maintenance (permanent unattended operation). Further on, all components have strict limitations in size and weight (Heunecke 2011).

The following sections describe the components used for the feasibility study (**Figure 126**), taking into account the limitations described above. Based on low-cost hardware it is shown that the implementation of an active prism is possible and also offers an economic advantage. This paper only relates to Leica prisms, as this hardware is available at the Geodetic Institute and used for the study. To realize a flexible adaption on different prisms, most components are adapted on top with a carrier board instead of a target plate. The centering accuracy of the used circular prism GPR121 brings benefit in comparison to 360° prisms. Modified to the active prism it preserves this accuracy which is later proven.

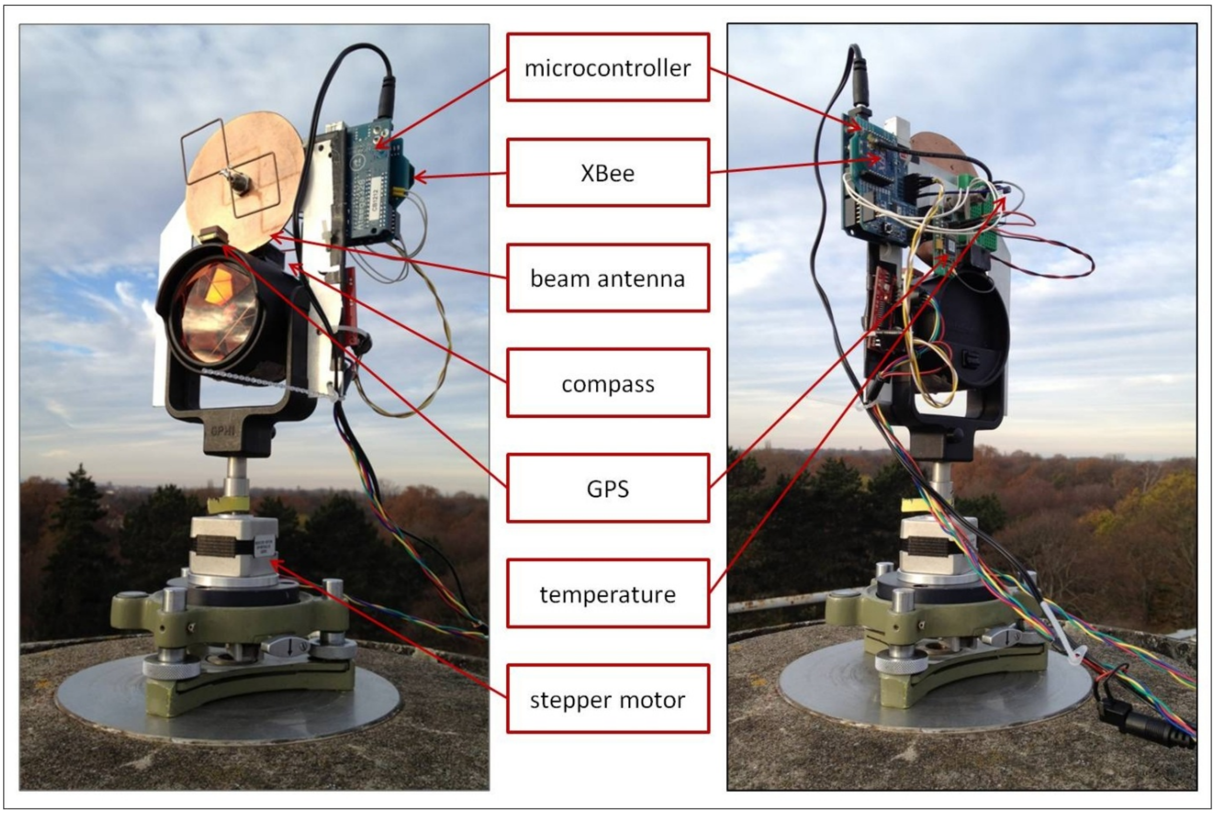


Figure 126: Feasibility study of the active prism

2.1 Controlling

The first precondition to implement an automatic prism alignment is a small and low power controlling unit for all sensors and actuators at the prism. In the presented feasibility study a small size microcontroller board based on the Atmel ATmega328 realizes this controlling. The Arduino platform has been dedicated a special attention by the robotics community worldwide. The Arduino provides 14 digital input/output ports (of which 6 can be used as Pulse-Width Modulation (PWM) outputs) and 6 analog inputs. All pins operate at 5 volts and can provide or receive a maximum of 40 mA. This Arduino board allows the controlling of additional sensors and actuators like stepper motors directly.

For the programming the ATmega328 on the Arduino it has built in a boot loader, which allows uploading new code without the use of an external hardware programmer. The programming language is a variant of C with additional features, such as program flow control and convenience functions. These features are also known from Processing, a Java based rapid-prototyping language for quick and easy design of programs with graphic output (Arduino 2011).

2.2 Communication

Another important requirement for the realization of an active prism is the existence of a suitable wireless transmission method between prism and total station. According to the prism the total station may be controlled by a microcontroller or by a notebook. In the view of existing technologies like WLAN or Bluetooth it can be noticed, that many of these technologies are either too complex, too expensive or consume too much energy. Additionally preconditions like limited weight or the transmission speed and reliability exist. Especially in geodetic environments sensors are mostly battery powered and should survive a long time without battery change. The IEEE 802.15.4 standard has been defined for this constellation.

The first feasibility study of the active prism is equipped with a XBee 802.15.4 1mW Series 1 module (Digi International 2011). This module can be wrapped into a serial command set, consumes only 50 mA of current, has a maximum data rate of 250 kbps and a range up to 1000 m. Basically there are two versions of XBee modules (both in different flavors, e.g. with chip antennas or wire antennas or different ranges and pro-versions). XBee Series 1 implements the basic functionalities and series 2.5 improves these by adding a full-featured mesh networking algorithm which also requires a different chipset to run. Test measurements with the XBee Pro Series 1 modules showed an approximate range of a few hundred meters in open areas with low radio activity, which is consistent to the manufacturer information.

2.3 Positioning and orientation

The automatic alignment requires the position and orientation of the prism as well as the position of the total station. The introduction of position and orientation of the prism can be distinguished between internal and external provision. The external provision describes the determination of the required information by the user. This is entered manually (known station coordinates and initial orientation) or wireless from the central station after the setup. The external determination has to be performed again after any power failure. In order to allow an independent functioning of the prism, the external provision should only be used in exceptional case. The internal provision is characterized by positioning and orientation sensors mounted on the prism. These sensors are controlled by the microcontroller, whereby the determination of required information could be performed at each time. When the current position and orientation of the prism are known, the alignment is carried out with the calculated new orientation angle by a stepper motor.

2.3.1 Stepper motor

A small Mercury stepper motor carries out the alignment of the circular prism. Stepper motors move a known interval for each pulse of power. These pulses of power are provided by a stepper driver and are referred as steps. In each step, the motor rotates by a known angle. This makes them handy devices for repeatable positioning. The used stepper motor has got a motor step increment of 1.8°, a rated voltage at 12V and a holding torque of $2.3\ \mathrm{kg \cdot cm}$. By the adaption of the stepper on a Leica claw plate, the motor can be centered on a tribrach. The prism is mounted on the motor shaft. In this way, the motor rotates the prism around its vertical axis by control commands from the microcontroller.

The centering accuracy of the active prism should mainly depend on the prism accuracy, not on the centering accuracy of the carrier respectively the centricity of the motor shaft. To determine the centricity of the motor shaft, a calibration measurement has been performed. Coordinates of the carrier on the motor shaft in different motor positions of one turn have been determined by a Leica Lasertracker LTD640. The horizontal distances of all measurements to the mean position are shown in

Figure 127. The largest deviation to the mean position is below 0.1 mm. This proves that the centricity of the motor shaft is one power to ten better than the prism accuracy itself. This underlines the benefit of the active prism in comparison to a 360° prism.

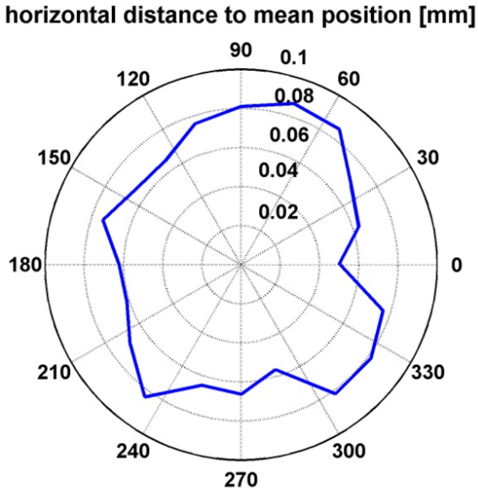


Figure 127: Centricity of the carrier on one rotation

Further on, an optimal alignment of the prism to the target axis of the total station is not necessary. Although the reflective spot shape changes with the angle of incidence, the automatic target recognition of modern total stations allows a relatively wide operating range. Not until the difference between alignment direction and line of sight exceeds 10°, the accuracy of the directional measurement will change for worse (Favre, Hennes 2000). From this it follows that the smallest motor step increment of 1.8° allows a sufficiently accurate alignment.

2.3.2 Positioning with GPS

For the positioning in open areas basically different methods are possible. With regard to an application in classical geodetic networks a low-cost GPS receiver is used here. More precisely, a GPS receiver of the company Libelium with a STA2056 chip is installed. The STA2056 chip combines standard GPS performances with low power consumption. The GPS receiver transmits the NMEA protocol over a serial interface to the microcontroller. Here, the data will be parsed for further use.

The GPS-position of the active prism is used for calculating the orientation angle. For this reason, the low-cost hardware inaccuracies of about 10 m can easily be tolerated. For typical distances in a terrestrial network of at least 100 m (see Section 4.2), position inaccuracies of 10 m produce orientation errors of only 0.1° and can be neglected for this scope.

Next to the prism position, the coordinates of the requesting total station are needed to calculate the orientation angle and perform an automatic prism alignment. According to the prism implementation, this information could get from known station coordinates or additional low-cost hardware. Another method is the use of geodetic GPS receivers on top of the total station, e.g. Leicas Smart Station.

2.3.3 Orientation with beam antenna

In addition to the positioning of the prism, the knowledge about the current orientation of the prism is necessary for the automatic alignment. There are various methods to determine this orientation. In a student project the use of a beam antenna at the prism has been examined to measure the signal strength at the total station. In this case, the prism rotates around its axis in predefined increments and the signal strength is measured at the total station. The stronger the signal is, the better the alignment of the prism. During the examination of this method significant problems in terms of reliability have been noted. Multipath effects interfere with the search for the strongest signal and thereby lead to a wrong orientation. Further on, the proprietary developed antenna doesn't come with an optimized shielding or directivity. Despite the problems encountered during the examination this method should be pursued in the future. The use of a professional beam antenna could lead to better results.

2.3.4 Orientation with digital compass

Another way to implement an autonomous orientation is by using a digital compass module for measuring the geomagnetic field. In the feasibility study of the active prism a LSM303DLH tilt-compensated compass module for the Arduino microcontroller is used. The LSM303DLH combines a digital 3-axis accelerometer and 3-axis magnetometer into a single package that is ideal for making a tilt-compensated compass. The supply voltage is at a low level between 2.5 and 3.3 V. Six independent readings, whose sensitivities can be set in the ranges of ±2 to ±8g for acceleration and ±1.3 to ±8.1gauss for the magnetic field readings, are available through an I²C interface (STMicroelectronics 2011).

In principle, the automatic prism orientation doesn't require a basic 3-axis magnetic sensor due to the precise leveling of the prism. In this case, pitch and roll angles are 0° and the heading angle can be determined directly. Since the sensor in the test setup will be installed on the prism body here, a highly accurate leveling is not always guaranteed. Therefore, the accelerometer is used to measure the tilt angles of pitch and roll for tilt compensation. The magnetic sensor is used to measure the earth's magnetic field. In the last step the tilt-compensated orientation of the active prism with respect to the magnetic north is calculated.

As a simple reliability check of the compass, test measurements on repeated automatic rotation of the active prism were performed. In order to avoid a wrapping of the power cables around the tripod, the rotating direction has been inverted after each round. The measurement result is shown in **Figure 128.**

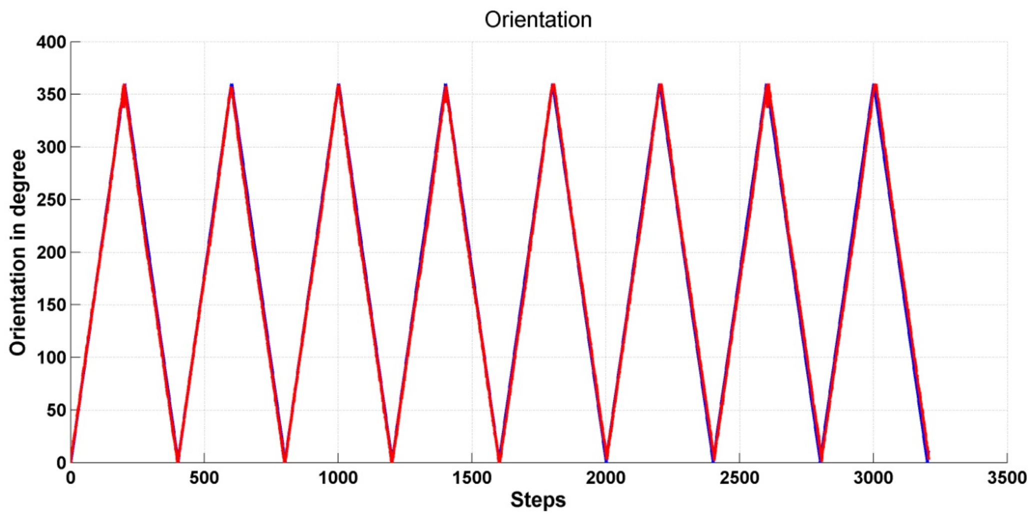


Figure 128: Orientation measurements by few rotations of the active prism

A triangle signal of the orientation angle (red) is clearly visible. To a synthetic reference signal (blue) only small deviations are recognizable. Even with repeated rotation only slight shifts of the two signals appear. Maximum deviations are in the range of about 8°, which are also due to a slip of the stepper motor. These differences remain within the required accuracy of an alignment to the line of sight of the total station (Favre, Hennes 2000). In further work, the compass should be compared to an absolute reference.

2.4 Additional sensors

The existing communication between the prism and total station allows an additional exchange of information for quality improvement. Two sensor components are presented at this point exemplary.

First, a Maxim DS18S20 temperature sensor is used on the active prism. This sensor enables temperatures to be determined at the prism which can be retrieved by the total station. In this way an improved temperature correction of distance measurements can be calculated. The communication between the temperature sensor and the microcontroller is realized via a 1-wire bus. A 1-wire bus requires only one data line (and ground) for communication with the microcontroller board. According to the manufacturer, the sensor has an operating temperature range of –55°C to +125°C and is accurate to ±0.5°C over the range of –10°C to +85°C. In addition, the DS18S20 can derive power directly from the data line, eliminating the need for an external power supply (Maxim 2011).

Furthermore, the acceleration sensors of the compass module can be used for the determination of other quality parameters. Regular monitoring of the prism acceleration allows the assessment of external manipulation of the prism (e.g. through vandalism in remote locations) and thus a recognition of a declined leveling.

3 CENTRAL STATION – CONTROLLING APPLICATION

The interaction between the active prism and the total station is controlled by a central application.

Figure 129 shows the basic communication structure. The total station is controlled via RS232 and common GeoCOM commands. The information from the active prism is received wirelessly via the XBee interface. This interface can also be used to transmit setup parameters to the prism.

An initial registration of all active prisms is made by query parameters (identification number, position, orientation, temperature) which are entered in a database and displayed on a map (see **Figure 129**). Parameters that are not set on the prism can be set by the user and transmitted to the prism. The next step is the initialization of the total station (RS232 communication, station coordinates). In this simplified representation, the position information can be determined via Leica smart station. Subsequent, angle and distance measurements can be performed to the registered prisms. In addition, all distance measurements can be corrected using the averaged temperature on total station and target prism.

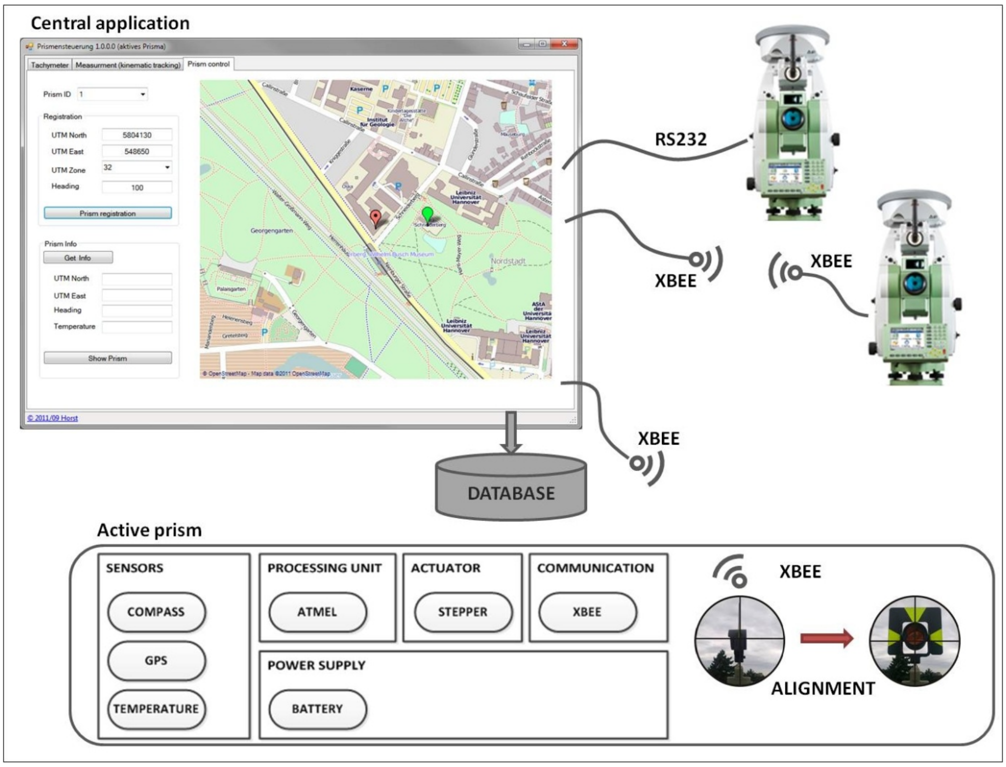


Figure 129: Communication structure

The alignment of the active prisms is implemented within a control loop. The first step is an availability request of the active prism. In a network with more than one total station, the reflector alignment is ambiguous. If the prism is already used by another total station the requesting total station has to wait a defined time before a second request is launched. If the prism is available, the alignment and measurement can be performed. In case of an inaccurate prism alignment a measurement error is responded by the total station. An automatic alignment by rotating the prism has to be initiated. The required rotation angle has to be determined. More detailed information about the decision about an appropriate response of the active prism is published in (Horst et.al. 2011).

4 EFFICIENCY ANALYSIS

Polar network measurements imply a high logistic effort because each point must be visited repeatedly to adjust the prisms to the current total station position. For measuring points with a long access route the expense is very high. These points can be equipped with a 360°-prism if the required accuracy permits. But in most cases a centering accuracy of 5 mm (Leica GRZ4, see **Table 17**) is not accurate enough and a circular prism with a centering accuracy of 1 mm (Leica GPR121) has to be used. A good compromise can be the use of the introduced active prism.

4.1 Costs of an Active Prism

There are different types of prisms, which are suitable for network measurement. Which prisms are used depends on the required accuracy. In this paper we restrict to Leica prisms, as a Leica circular prism is used for the active prism. Leica has two different circular prisms and two different 360° prisms on offer that are basically suitable for network measurements. **Table 17** shows the centering accuracy and the price factors of the prism. For example, the circular prism has the price factor 1 and the 360° prism GRZ122 has the price factor 5. This means that the prism GRZ122 is five times more expensive than the GPR121. Both of the 360° prisms are significantly more expensive than the circular prism GPR121, which is used for the active prism. Even with the additional components, the active prism does not reach the price of a 360° prism. The prices of the 360° prism GRZ4 and the high precision circular prism GPH1p are comparable (three times more expensive than the GPR121) but the centering accuracy of the GPH1p is significantly higher. The centering accuracy of the stepper motor would even allow the use of a GPH1p, with a minimal loss of accuracy.

Table 17: Prism costs (Leica 2011)

prism	GPR121	GPH1p	GRZ4	GRZ122	Active Prism
					
centering accuracy	1 mm	0.3 mm	5 mm	2 mm	1 mm
price factor	1	3	3	5	2

4.2 Simulation of a Network Measurement

In the following it is shown how the use of one or several active prisms affects the measuring duration and the measuring costs of a network measurement.

Figure 130 shows the geodetic network that is used as an example for the time and cost calculation. The network consists of 27 measuring point whereby 13 of them are positions of the total station. The red lines show the observations and the blue ones the road network. There is one point (1600) with a long and difficult access route and three points (2003, 2004, 2005) are located on a water intake tower, which is only accessible via a tunnel. We calculated a simulation of four scenarios: (1) without any active prism, (2) with an active prism on 1600, (3) with tree active prisms on 2003, 2004 and 2005 and (4) with four active prisms on 1600, 2003, 2004 and 2005. The measurement costs include personnel costs and travel costs. A team consists of one surveyor and one assistant, which together earn € 90 per hour. The costs per kilometer driven by car are calculated with 0.30 €/km.

The calculation is done with the software-prototype SimPle-Net (**Si**mulation and **Pl**anning of **e**fficient **Net**work measurements) which is developed within the DFG research project EQuiP (Rehr et.al. 2011). The measurement is calculated as if it is carried out in one piece, without any interruption. In this case it is possible to compare the single scenarios with each other.

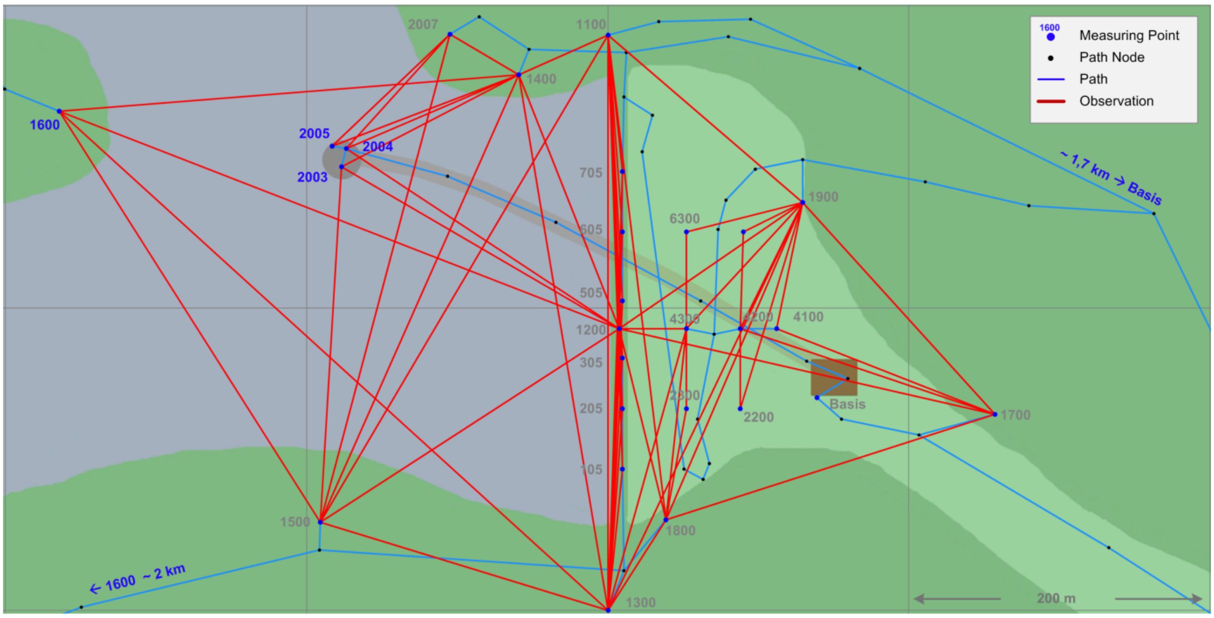


Figure 130: Geodetic network with 27 points (including 13 station positions)

The results of the calculation are shown in **Table 18**. The measurement of scenario 1 without any active prism takes 9:10 hours. The traveled distance is 85 km and the measurement costs amount € 845.

In the second scenario the measuring point 1600 is equipped with an active prism. The distance between the points 1500 and 1600 is more than 2 km and accessible by a bad road where one can only drive very slowly. In carrying out the measurement with one active prism, about 30 minutes and 8 km can be saved. The measurement costs decrease to € 800. The additional costs of approximately € 380 caused by the modification of the prism are amortized over 5 years and 6 measurements per year. The additional costs per measurement and prism thus amount to € 12.70. Therefore the total costs for the second scenario are € 813.

Table 18: Results of the time and cost calculation

# active prisms	pno	# staff	duration [hh:mm]	distance [km]	measurement costs [€]	additional costs active prisms [€]*	total costs [€]
0	--	2	09:10	85	845	0	845
1	1600	2	08:41	77	800	13	813
3	2003, 2004, 2005	2	08:52	83	817	38	856
4	1600, 2003, 2004, 2005	2	08:17	62	761	51	812

* amortization over 5 years with 6 measurements per year (= € 12.7 per prism)

The third scenario shows the results of a simulated measurement with three active prisms on 2003, 2004 and 2005 located on the water intake tower. The benefit of 18 minutes and only 2 km is not as high as in the second scenario because these points are only measured two or three times. On the other hand the point 1600 is measured four times and the access route is even longer. In the third scenario the total costs are even higher than the total cost in the first scenario. The cost for modifying the three prisms cannot be compensated by the reduction of measurement time.

In the last simulation all four points are equipped with active prisms. As expected, in this case the time savings are the greatest. One can save more than 50 minutes and 23 km in comparison with scenario 1, without any active prism. The total costs of scenario 4 (four active prisms) are comparable to the total costs of scenario 2 (one active prism).

In summary it can be concluded that active prisms should only be used for points with difficult access that are often measured on several directions. It must also be noted that the prisms must be used several times per year, so that a modification is worthwhile. We calculated with an amortization over five years and six measurements per year. In the example network it would be advisable to modify only one prism to use it at point 1600.

5 CONCLUSIONS

The proposed active prism is a good alternative to a 360° prism, especially if the accuracy requirements are high. Existing circular prisms with holder can easily be modified with the introduced low-cost components. The prism carrier is replaced by a motor and the other sensors like compass, GPS receiver or beam antenna are placed on the top of the prism. The attachment is carried out via the already existing grooves on the prism holder, which is usually used to fix the target plate.

Such a modification is worthwhile when measurements are often performed in large measurement areas where there are points that are difficult to reach or have a long travel distance. For common network measurement, it is not advisable to modify all the prisms, as the staff is on site anyway for the set-up and dismounting of the prisms and otherwise would be idle during the measurement.

For monitoring tasks a modification is worthwhile if one prism should be measured from different total station positions around the horizon. Another application is the use of active prisms in tunnel network measurement. Here often several prisms are in a line of sight. These must be sequentially disguised, as they could cause faulty measurements (with ART1). Whereas active prisms can receive the command to rotate out of sight and only rotate back to the instrument if they are needed.

To obtain comparable accuracies to a common circular prism with a common carrier, further tests must be performed with the active prism. Taking into account the aperture angle of the prism and inaccuracies of the orientation, borderline cases of the alignment especially the decision about necessity of a new alignment must be proven. To save power on the prism a new alignment has to be carried out only in case of declining position accuracy. Finally, it has to be noted that all components must be adapted for the use in all weather conditions to guarantee certain robustness.

REFERENCES

Journal articles:

FAVRE, C.; HENNES, M.: *Zum Einfluss der geometrischen Ausrichtung von 360°-Reflektoren bei Messungen mit automatischer Zielerfassung*, VPK 2/2000, p.72-78, 2000.

HEUNECKE, O.: *Impact and Benefit of Geo Sensor Network Technology for Monitoring Tasks*, 8th International Sympsium on Field Measurements in GeoMechanics / Proceedings, September 12-16, 2011, Berlin.

HORST,S., ALKHATIB, H., KUTTERER, H: *Automatic response capability in geodetic sensor networks*, 8th International Symposium on Field Measurements in GeoMechanics / Proceedings, September 12-16, 2011, Berlin.

REHR, I. ; RINKE, N. ; KUTTERER, H. ; BERKHAHN, V.: *Maßnahmen zur Effizienzsteigerung bei der Durchführung tachymetrischer Netzmessungen*, AVN 1/2011, p. 2-13, 2011.

Links:
ARDUINO, 2011: http://www.arduino.cc, last accessed on November 17, 2011.

DIGI INTERNATIONAL: www.digi.com, last accessed on November 17, 2011

LEICA, 2011: http://www.leica-geosystems.us/en/Acc-Surveying-Engineering_84780.htm, last accessed on November 25, 2011.

MAXIM, 2011: http://www.maxim-ic.com/, last accessed on November 17, 2011.

STMicroelectronics, 2011: http://www.st.com/internet/com/technical_resources/technical_literature/datasheet/cd00260288.pdf, last accessed on November 17, 2011.

Agriculture

In-Field Route Planning for Agricultural Machines - A Review

Stefan Böttinger
Institute of Agricultural Engineering, University of Hohenheim, Germany

Route planning for agricultural machines in an agricultural field can make their use easier, especially:
- for non-local drivers
- for the simultaneous use of several similar machines
- for the use of cooperating machines
- for autonomous or semi-autonomous machines

With this proposed paper a review will be given about published solutions for this challenge.

This paper has been accepted as a „peer reviewed paper“ by scientific committee

Testing a Mini UAS to Collect Geo-Referenced Data for Agricultural Purposes

Johanna Link Dolezal[1], Klaus Kittmann[2], Dominik Senner[1], Wilhelm Claupein[1]
[1]Intitute of Crop Science, University of Hohenheim, Germany
[2]Institute of Aircraft Design, University of Stuttgart, Germany

Abstract
Precision Farming (PF) aims at taking the given spatial and temporal variability in a field into account, when managing the field. Thus the main focus of PF is to adapt the application of input factors (e.g. nitrogen or pesticides) to the current demand of the plant and adapt the management based on given soil characteristics (e.g. seeding based on soil moisture). PF technology provides sensors and machinery to come up with site-specific management prescription for a given field.

Within the last years, new approaches to identify management zones and application rates were investigated. The usage of an Unmanned Aerial System (UAS) to collect high-resolution data of a field is one of the promising and most challenging ideas in this area. The aim of this study was to implement a mini UAS, which a) is able to follow a given route within a field and b) can be used to collect geo-referenced data for agricultural purposes, which could be used to develop management prescriptions in terms of PF.

The results indicated that it is possible to use a beyond the idea of collecting high-resolution data with a fixed wing UAS, there are some limits and challenges to use mini UAS for agricultural purposes. Thus, further interdisciplinary research in this area is absolutely necessary to bring the idea into practice.

Keywords
UAV, UAS, SenGIS, precision farming, geo-referenced data, flight vehicle design, autopilot, navigation

1 INTRODUCTION

1.1 Precision Farming (PF)

Current agricultural farming systems are confronted with a broad range of economic and environmental issues, which require a reasonable use of given resources (e.g. nitrogen or pesticides) when it comes to management decisions. Beside, management of fields should take the given natural variability into account as growing conditions vary within a field (Bouma, 1997). Precision Farming (PF) represents a management strategy, to manage uncertainty in crop and soil properties “through better understanding and management of spatial and temporal variability” (Dobermann et al., 2004).
Different management strategies were implemented during the last decades to achieve a reasonable use of resources. Besides knowledge and experience from previous years, resource management decisions may also be targeted by information from non-destructive remote sensing data, e.g. spectral data or images (Abuzar et al. 2009; Mistele and Schmidhalter, 2008). Remote sensing may provide information on the current status of plants in the field (e.g. biomass, nutrient status, crop health, yield, product quality etc.) and thus make it possible to adapt necessary inputs to currently existing demand.
In the context of PF, remote sensing applications often require geo-referenced data with high spatial, temporal and spectral resolution. However, data for these purposes are difficult and costly to obtain,

either by ground-truth measurements on one hand, and satellites or manned aircrafts on the other hand. In this context an Unmanned Aerial System (UAS), defined as Unmanned Aerial Vehicle (UAV) equipped with Global Positioning Systems (GPS), autopilot and remote sensing payload, might offer certain advantages and new possibilities to gather high resolution sensor data for agricultural farming systems.

1.2 State of the art

At the moment geo-referenced data for agricultural purposes are mainly gathered by collecting manual ground truth measurements (e.g. rating, measuring of crop status) or by tractor-based sensors (e.g. nitrogen content, biomass index, yield and protein mapping). The collected data will be pre-processed, displayed as spatial maps and used as basis for further management decisions.

During the last years the implementation of micro and mini UAS for agricultural purpose became much more relevant. Herwitz et al. (2004) describe the potential of slow-flying fixed wing UAV equipped with imaging cameras for agricultural surveillance. Grenzdörffer et al. (2008) provide an overview of the potential of several low-cost UAVs in forestry and agriculture. In the last years the major focus was drawn on the implementation of helicopter (e.g. Xiang and Tian, 2011; Merz and Chapman, 2011) and microcopters (Israel, 2011).

While most research with agricultural purpose focuses on image data collection with compact cameras (Rovira-Más et al., 2005; Laliberte et al., 2010) only a few groups use other sensors as multi- or hyperspectral spectrometers (Goel et al., 2003; Nebiker et al., 2008, Overgaard et al., 2010) as basis for further data analysis. Sensors like spectrometers allow a deeper insight and thus a more detailed interpretation of crop status, which provide a new opportunity for improved management decisions.
Besides, the implementation of micro or mini UAS to gather geo-referenced data seems to have some additional advantages. The collection of geo-referenced data from the air (aerial data) makes the data collection independent from existing soil conditions and does avoid additional soil compaction and/or crop damage during data gathering. In case of a well-established data chain from planning the data collection, to the process of data collection itself to the post processing data analysis, a micro or mini UAS provides a powerful tool. The idea behind is, that UAS can help to bridge the spatial and time resolution gap between ground-based observations and remotely sensed data.

1.3 Aim of the study

The aim of this study was to equip a fixed-wing UAV and to test it concerning its suitability to be used as mini UAS for data collection for agricultural purposes. The main focus was a) to test if the mini UAS is able to follow a given route within a field and b) if the mini UAS could be used to collect geo-referenced data, which could be used to develop management prescription in terms of PF.

2 UAS FOR AGRICULTURAL PURPOSE

2.1 General considerations

For the same purpose of data collection (e.g. geo-referenced data for PF), the requirements might be quite different, depending on the end user of the UAS. For an inexperienced end user the equipment needs to be easy to handle, while for an experienced end user the accuracy of flight routes and data collection might be much more important. The key issue to design and manufacture an UAS is to meet the user requirements as well as possible and at the same time to stay inside the restrictions which come from aerodynamics and physics. Normally, the requirements are contrary to each other, so the result is in any case some kind of compromise, which can be optimized towards a certain criteria.
In principle, different approaches to UAS applications for the end users can be distinguished. Either the end user operates and owns an UAS himself or the end user uses an UAS service which can be ordered on demand. Both solutions have advantages and disadvantages. While the access time on a user operated micro or mini UAS is much smaller and can be flown on the basis of a spontaneous

planning, the professional services based UAS can be more complex and can provide a higher performance. On the other side it must be considered that a service based UAS needs more time for customization with designated sensors, which leads to a more inflexible character of such a system.

2.2 Requirements for the mini UAS

In order to implement a mini UAS in agriculture to collect data for PF, different requirements need to be considered.

First of all, agricultural processes are mainly driven by weather conditions, thus the conditions in the field and the development of the crop are variable in time and space. Thus the goal is to own an UAS which is equipped with the user payload and stored ready for flight to be used anytime and anywhere. This required the implementation of a flexible system to guarantee a prompt gathering of geo-referenced data, when weather conditions were optimal. Compared to tractor-based sensors or satellites, the implementation of a fixed wing UAS is independent of current soil conditions and clouded skies, but not completely independent of wind and weather conditions like rain or fog.

To use an UAS for data collection in an agricultural field, it is important to have the possibility of flexible starting and save landing. An UAS which is able to be operated anywhere and anytime should be designed in a way that no infrastructure (e.g. runways) is necessary. For minimal starting and landing requirements a Vertical Take-Off and Landing (VTOL) UAS as a helicopter or microcopter could be used, but both systems normally have a limitation in the covered area per flight and in additional payload. Both, flight duration and payload are important factors for agricultural data collection. One other option would be a fixed wing mini UAS (up to 25 kg) which could be launched by hand or catapult and landed in the field, e.g. as a motor glider or another platform without undercarriage. Using a simple (rubber driven) catapult allows for the start of mini UAS from practically each location. The UAS can also be designed in such a way that it can be landed on extremely small areas and nearly on each ground surface by using no landing gears but a fuselage which is designed to do belly landings. In such a case different surface types (e.g. water, snow, grass, or the agricultural field itself) are also possible which are normally available when performing data collection for agricultural purposes. Due to the higher drag using no landing gear, even the necessary length of a free “runway” is smaller compared to landings on a concrete runway.

To cover a given area (e.g 10 ha) within a reliable time period (e.g. 30 min), it is necessary to guarantee a long flight time for measurements and come up with a reduced number of starts. An UAS which can be variably equipped with the energy source is preferred. Additionally it is helpful if fuel or batteries can be added in case of free payload capacities, to increase the flight endurance. Beside, in terms of flight duration and payload the UAS should possibly have good aerodynamic lift to stay in the air as long as possible with the carried density of energy within the batteries and the required payload. In this study a payload of about 1.3 kg (sensor, battery, and motorization) was accumulated. In order to provide the possibility to use different sensor with the same UAS a big voluminous payload bay which is low restrictive for mechanical integrations is advantageous.

From an acceptance point of view electrical motors have some advantages. Flying UAS is often linked with a question of noise emission. Electrical motors are normally less noisy and the acceptance of third parties is higher and the negative effects on the surrounding environment are reduced.

For research purposes it was very important to insure a high accuracy in position and quality of sensor data. The usage of an autopilot was required to ensure autonomous flights with repeatable flight routes, which was not possible with remotely piloted flights.

Due to regulations in Germany three categories of UAS are differentiated: take-off weight < 5 kg (micro), take-off weight 5-25 kg (mini) and take-off weight 25 – 150 kg (close range). For each class different regulations from the administrative point of view are relevant. As for the first category (< 5 kg) no take-off clearance is required, which would allow a prompt data collection anywhere and anytime, this was also considered as requirement in this study. However, operating UAS is completely different from model flying and hobby applications. Independently of the mass, a UAS needs a certification to be legally operated in Germany. Anyway, keeping below 25 kg makes the certification process slightly easier which might be a design criteria for the UAS for PF purposes.

For a certification it is advantageous to implement redundant systems. Even for mini UAS it is hard to reach a real redundant and reliable system. With several arrangements it is possible to increase the redundancy level of a small carrier. Simple but effective approaches are separated and multiple actuators per function, split and independent control surface segments, multiple communication devices with redundant power supply, multiple engines and flight abort system with parachute to name only a few of them. Using such solutions, it is possible to make the flight vehicle more robust against component failures, human failures (especially during set up and flight preparation) and to reduce the risk of serious damage.

For this study the following requirements to the mini UAS were defined (**Table 19**).

Table 19: Requirements for the fixed wing mini UAS to be used for PF purposes in this study

Start/Landing	without runway (anywhere at any time)
Payload	ca. 1,3 kg
Flight duration	> 15 min.
Take-off weight	< 5 kg
Flight modus	remotely and autonomously piloted
Preparation time	as quick or short as possible, to allow prompt starts
Recovered area	10 ha
Resolution of data	about 10 m in diameter, about 10 m distance
Costs	low

2.3 Requirements for the sensor

As shown in several studies, spectrometers are a suitable tools to gather information about crop status in a field (Reckleben and Isensee, 2005; Barker and Sawyer, 2010). Spectrometers or spectral cameras may provide detailed information about reflectance changes in defined wavelength ranges and thus can hint a nitrogen status (Kitchen et al., 2010; Teal et al., 2006), water content (Zhang et al., 2011) or plant diseases (Graeff et al., 2006) within the field. In order to use a spectral sensor in a mini UAS the following requirements need to be fulfilled in our study (**Table 20**).

Table 20: Requirements for the sensor to be used for PF purposes in this study

Sensor weight	< 1 kg
Sensor dimensions	about 10 cm x 10 cm x 10 cm
Spectral resolution	< 10 nm in visible and near infrared wavelenghts
Spatial resolution at ground	about 10 m in diameter
Interference	low
Costs	low

The sensor weight should be as light and dimensions should be as small as possible in order to integrate the sensor into the UAS. It is necessary to have a high spectral resolution to get reliable information about the relevant wavelength ranges and to provide the possibility to calculate common indices as NDVI which correlate e.g. with biomass or nitrogen. Against the background that the sensor measurements should later on be transferred into management prescriptions for the farmer, it seems to be desirable that the sensor measurements somehow match the working width of the farmer's equipments. Thus in our study a spatial resolution of about 10 m in diameter were required in order to get a reliable number of measurements within the field.

3 CONFIGURATION OF AN FIXED WING MINI UAS FOR AGRICULTURAL PURPOSES

3.1 Specification and realization of an UAS - ATRP (All Terrain Research Platform)

The Universities of Stuttgart (Institute of Aircraft Design) and Hohenheim (SenGIS, https://sengis.uni-hohenheim.de/) performed a study to test a mini UAS for geo-referenced data collection for agricultural purposes and to come up with a suitable mini UAS – a so called All Terrain Research Platform (ATRP). For such applications a robust UAS was necessary which was able to provide a sufficient amount of available endurance and payload capacities (as described above). Additionally, there was an interest in setting up an UAS which is easy to handle and which can be rapidly reactivated even in case of damage. A continuous support of spare parts was also a design criterion.

In order to find a suitable UAS a market study was carried out which mainly focused on the following criteria:

- High payload mass capabilities,
- Compact size to reduce transport and ground handling complexity,
- Multi engine concept to increase reliability,
- Wide fuselage for large payload bay volume to minimize mechanical restrictions during payload installations
- Capability for extreme short start and landing procedures
- Start option by catapult
- Landing option on any surface (ice, water, grass, snow); additional landing possibility in a web.
- Cots products if possible

The result of that study was the choice of a scaled model of the fire fighting airplane Canadair CL 215 manufactured by the company PAF-Flugmodelle (Erftstadt, Germany). The Canadair CL 215 is categorized as fixed wing platform, which afterwards was equipped with an autopilot and sensors to be used as ATRP of the category mini UAS. The following table shows the specifications of the adapted ATRP (**Table 21**).

Table 21: ATRP specifications

Wing span	2000 mm
Take-off weight	3200-4900 g
Maximum take-off weight (tested in flight)	11500 g
Airfoil	Clarc Y mod.
Controls	2 Aileron, 2 Flaps (backfitted), 2 Elevators, 1 Rudder, 2 Motors
Drive	Two brushless electric motors
Energy	2 * 5S LIPO 5000 mAh
Thrust	4,4kg @ 10000 rpms, 18 V

The ATRP provided a big payload bay which has the dimensions of (H, W, L) 205 mm * 190 mm * 400 mm (**Figure 131**). The payload bay was located under the wing section, which was optimal for changing masses. The centre of gravity was not impacted by varying the payload.

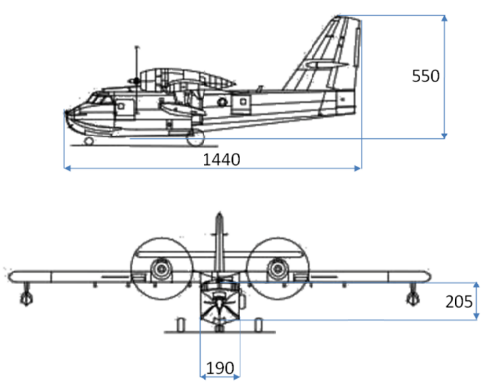


Figure 131: ATRP on the basis of a CL215 scale model – dimensions (Source: modified from http://www.redstar.gr; accessed Sep 2011)

The ATRP was controlled either manually or by an autopilot system. It was equipped with a remote control (RC) receiver, actuators for control surface control, GPS, air data probe, inertial measurement unit (IMU), autopilot and a telemetry modem for transmitting the flight relevant data to the ground control station. This equipment was installed permanently inside the rear part of the fuselage, in a way, that the payload bay was not impacted or reduced by these standard components. Inside the payload bay the aircraft can be equipped with various sensors depending on the flight task.

The ATRP was equipped with two lithium polymer (LiPo) batteries with 5000 mAh each. The available flight time was about 15 minutes. The operating speed was 16.6 m s^{-1}. The theoretical distance which could be flown had a length of 15 km. Depending on the distance of the flight legs and the necessary path between two measurement legs the achievable coverable area can be computed. Typical areas under such conditions (with flight leg distance of 30 m) have a size of about 100 000 m^2. In the current case, the start procedure was made by a trolley, which stays on the ground as soon as the aircraft has taken off. This solution was preferred because it saved weight on the flying structure, which then could be used for payload. The landing took part in form of a belly landing. Catapult start and web landing would be optionally possible.

The autopilot used in that ATRP was a component of the shelf (COTS) product. A detailed description about this autopilot is given by Kittmann et al. (2011). A positioning accuracy of 20 m and an altitude accuracy of 15 m were achievable with this autopilot. The accuracy of this system was limited due to the cheap sensors.

3.2 Sensor properties

Based on the requirements on a sensor used for aerial measurements in terms of PF, a monolithic miniature spectrometer MMS1 from Carl Zeiss (Carl Zeiss GmbH, Jena, Germany) was applied for the collection of geo-referenced data in this study. The MMS1 was implemented in the ATRP, as seen in **Figure 132**. It was selected because of the low mass (about 600 g), low temperature dependence, compact dimensions and the high spectral resolution. The sensor consists of a spectrometer housing fixed glued with an aberration corrected concave grating, a fibre cross section converter as optical input and a diode array. A PDA with GPS antenna (about 150 g) was used to trigger the sensor and to save the data during the flights.

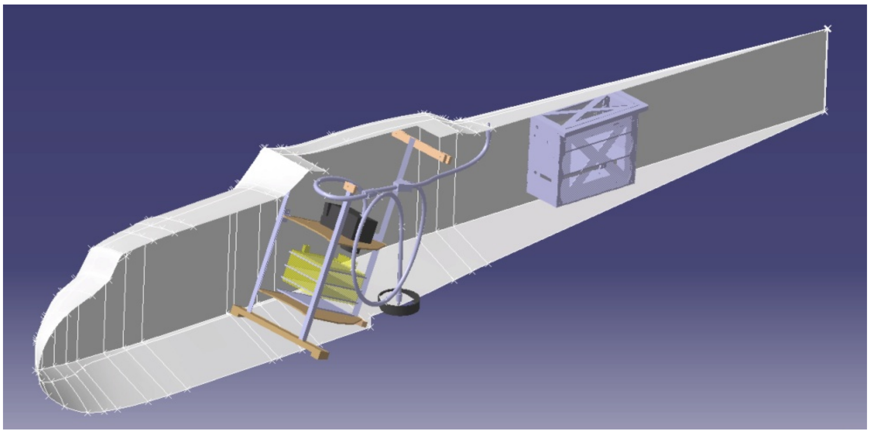


Figure 132: Fuselage segment with sensor installation under the wing section and autopilot in the rear.

The PDA used the internal battery for electric power during the flights for data collection. The power supply for the sensor was done by a LiPo with 3S 2200mAh (about 200 g). The autopilot had a separate power supply, too, with a weight of 182 g. The total weight of the sensor payload was about 950 g (600 g for the sensors system, about 150 g for the PDA plus about 200 g for the battery).
The relevant data from autopilot (position, actual roll- and nick angles and the flight height) were synchronized with the spectral data from the MMS1 via GPS-time and analysed with a GIS-Software (ArcGIS, ESRI, Redlands, California, USA).

4 EXEMPLARY RESULTS

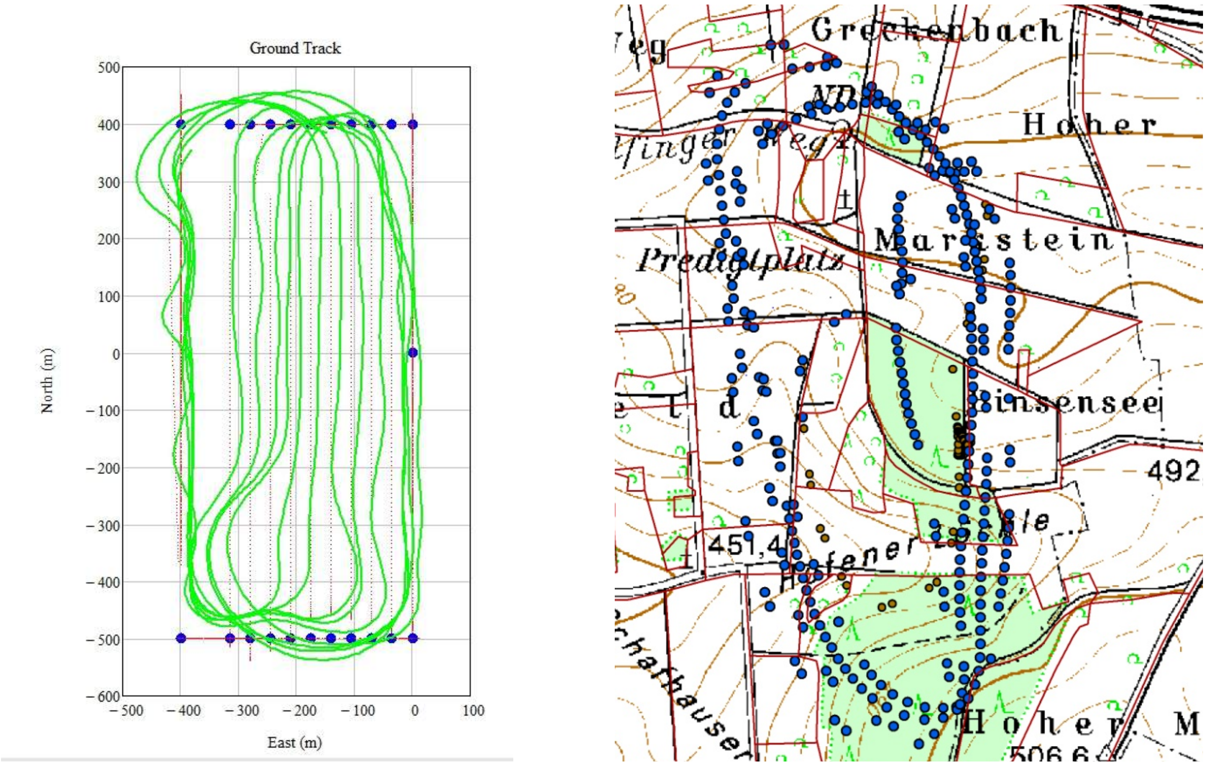
4.1 Navigation

Up to date several test flights to adjust the autopilot were performed, but not finalized at this point in time. During the test flights the ATRP was not equipped with the relevant sensor MMS1, but with sand sacks to simulate the specific weight and the specific weight distribution as under final configuration with MMS1.

In order to use the ATRP as UAS there are still some adjustments necessary. Due to this fact, the following figure shows some exemplary results mapped during the flight with a D40 (**Figure 133a**). The field for data collection was located within the coordinates East -300 to East 0 on the x-axis and within North -200 to North 300 on the y-axis. The grey dotted lines show the target positions across the field, while the green lines give the measured positions during the flight. The blue dots indicate so called waypoints during the flight.

The available results indicated a quite good approach of measured and target positions, when taking the technical limitations of the autopilot into consideration. The performance might be improved, if more accurate and sensitive sensors would be implemented in the COTS autopilot.

Against the background to develop management prescriptions based on the collected data, the working width of the farmers´ equipment needs to be taken into account. As current machinery has working widths of at least 6 m (often the width is 2 or 3 times wider), an accuracy of about +/- 1 m might still be good enough for the purpose in our study.



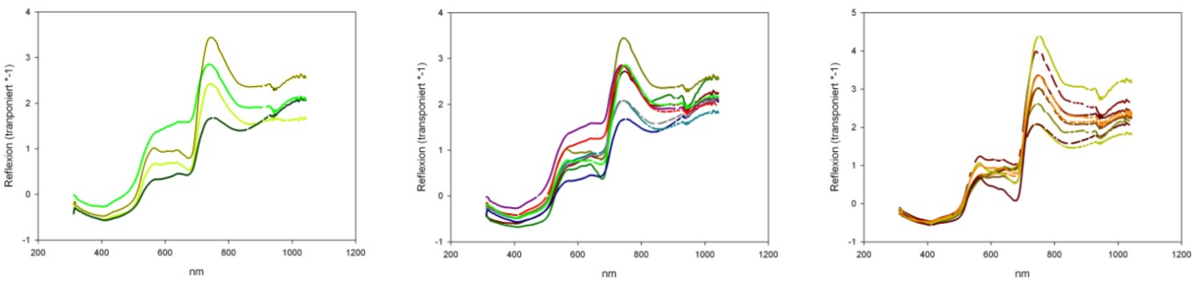
a b

Figure 133: Comparison between actual position of the D40 during flight (green line) and target position in the field (dotted lines). The field covers the area between the coordinates East -300 to 0 and North -200 to 300) (a). Data collection points at the given flight route on a model air field when testing the ATRP (b).

4.2 Sensor data

Several flight tests in form of sensor flights were required to evaluate the ATRP, and will still be performed in the near future.

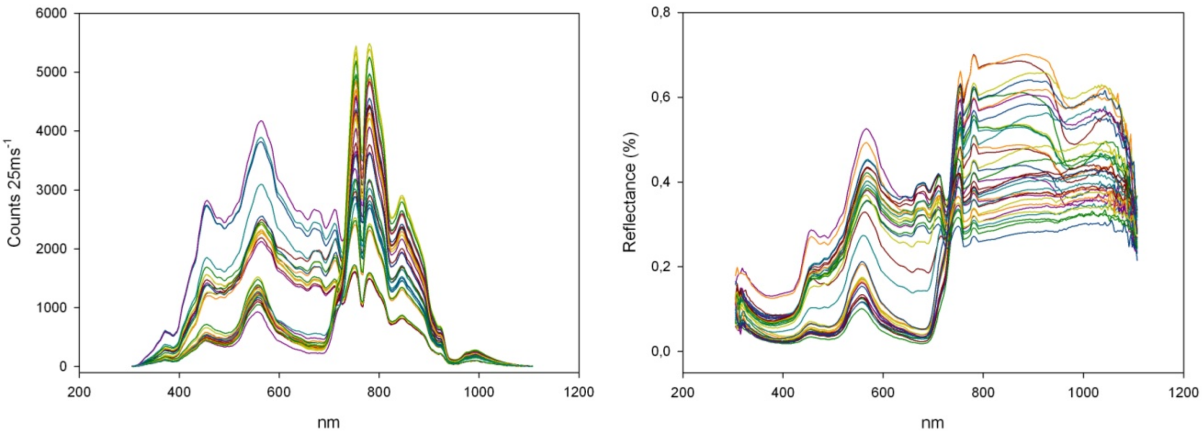
Figure 133b gives an impression of the flight route of the ATRP at the model air field in a height of 100 m above the ground. The data points on the right hand side (in north to south direction) display the three parallel flights routes, which were determined via autopilot for this test. Due to the fact that the GPS of the sensor had to be placed inside of the UAS the data collection pattern shows a lot of missing data. As this problem could be solved easily in the next flights, it seems to be possible to come up with a quite good data set, which covers the area of interest with a sufficient amount of data per area.
Nevertheless the remaining data from this first test flight at the model airfield were categorized concerning the given land cover. The following figures (**Figure 134**) give the transposed reflectance of the sensor measurement for wood (a), grassland (b) and arable land areas (c).



a b c

Figure 134: Transposed raw sensor data (counts 25 ms^{-1}) collected with MMS1 in a flight height of 100 m for forest areas (a), grassland areas (b), and arable land areas (c).

Sensor data taken at a height of 100 m above a wheat field in the current growing season were given in the following figures. **Figure 134a** gives an impression of the collected raw sensor data (counts 25 ms^{-1}) during the flight across the test field. Each line represents another position in the test field. **Figure 135b** shows the reflectance of the crop stand calculated as ratio between measured data at different positions in the field and the measured white reference. The calculation indicates quite remarkable differences concerning reflectance of incoming light within the field. However, further analysis is necessary to evaluate the quality and usability of the sensor data for agricultural purposes, especially for PF.



a **b**

Figure 135: Raw sensor data (counts 25 ms^{-1}) collected with MMS1 in a flight height of 100 m indicating varying growing conditions within the field. Each line represents another location in the field (a). Reflectance (%) of the crop within the test field, calculated from raw data and white reference of MMS1 in a flight height of 100 m. Each line represents another location in the field (b).

5 CONCLUSION

The aim of this study to equip and test an UAV concerning its suitability to be used as UAS for data collection in terms of PF was confronted with a series of challenges and restrictions. The first approach to meet as many requirements as possible (e.g. weight of payload vs. size of UAS, take-off weight vs. flight endurance) resulted in the ATRP, as described above. The exemplarily results of the recorded geo-referenced data from the autopilot indicated, that the ATPR might be able to follow a given route within a field. The performance of the autopilot seemed to be mainly influence by the technical limitations of the implemented sensors. Depending on the accuracy of the management decisions the farmer will do based on UAS data, higher requirements concerning spatial accuracy of the collected data might be favourable. Thus, more sensitive sensors could be included to the autopilot, to improve data collection from this point of view.

First results of measured spectral data showed that an UAS can be used to collect geo-referenced data for agricultural purposes. The collected spectral data hinted at quite remarkable differences in reflectance. Further analysis and correlation with ground truth data from the wheat field need to be performed to evaluate the quality and usability of these data.

Beside, open points for further development are in sensor stabilization, navigation accuracy, more optimal usage of available flight time, sensor weight, and data processing from different systems. Due to the fast development of miniaturisation of sensors and components, there will be a big potential for going a step forward. In order to come up with suitable solutions and to bring the idea behind this study into practice, further interdisciplinary research in this area is absolutely necessary.

ACKNOWLEDGEMENTS

The authors would like to thank the Carl-Zeiss Foundation for the financial support of the SenGIS project. Additional thank goes to Johannes von Helldorf for his studies with the ATRP.

REFERENCES

Books:

BOUMA, J.: *General reflections* In *Precision Agriculture: Spatial and Temporal Variability of Environmental Quality.* LAKE, J.V., BOCK, G.R., GOODE, J.A. 1997.

DOBERMANN, A., BLACKMORE, S., COOK, S.E., ADAMCHUCK, V.I.: *Precision Farming: Challenges and Future Directions*. In*: Proceedings of the 4th International Crop ScienceCongress*. http://www.cropscience.org.au/icsc2004/symposia/4/3/217_dobermanna.htm. 2004.

Journal articles:

ABUZAR, M., O´LEARY, G., FITZGERALD, G.J.: *Measuring water stress in a wheat crop on a spatial scale using airborne thermal and multispectral imagery*. Field Crops Res. 112, p. 55-65. 2009.

BARKER, D.W., SAWYER, J.E.: *Using Active Canopy Sensors to Quantify Corn Nitrogen Stress and Nitrogen Application Rate*. Agronomy Journal, 102, p. 964-971. 2010.

GOEL, P.K., PRASHER, S.O., LANDRY, J.A., PATEL, R.M., BONNELL, R.B., VIAU, A.A., MILLER, J.R.: *Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn*. Computers and Electronics in Agriculture, 38, 2, p. 99-124. 2003.

GRAEFF, S., LINK, J., CLAUPEIN, W.: *Identification of powdery mildew (Erysiphe graminis sp. tritici) and take-all disease (Gaeumannomyces graminis sp. tritici) in wheat (Triticum aestivum L.) by means of leaf reflectance measurements.* Central European Journal of Biology, 1, 2, p. 275-288. 2006,

GRENZDÖRFFER, G., ENGEL, A.: *A comparative study of two micro-UAS's - Perspectives for the current and economic geo-information extraction*. GIS-Zeitschrift für Geoinformatik, 1, p. 17-25. 2008.

HERWITZ, S.R, JOHNSON, L.F., DENAGAN, S.E., HIGGINS, R.G., SULLIAVAN, D.V., ZHENG, J, LOBITZ, B.M., LEUNG, J.G., GALLMYEYER, B.A., AOYAGI, M., SLYE, R.E., BRASS, J.A.: *Imaging from an unmanned aerial vehicle: agricultural surveillance and decision support*. Computers and Electronics in Agriculture, 44, 1, p. 49-61. 2004.

ISRAEL, M.: *A UAV-based Roe Deer Fawn Detection System.* International Archives of Photogrammetry and Remote Sensing, Vol. XXXVIII-1/C22. ISSN 1682-1777. 2011.

KITCHEN, N.R., SUDDUTH, K.A., DRUMMOND, S.T., SCHARF, P.C., PALM, H.L., ROBERTS, D., VORIES, E.: Ground-based canopy reflectance sensing for variable-rate nitrogen corn fertilization. Agronomy Journal, 102, 1. p.71-84. 2010.

KITTMANN, K., BREEMAN, J., SCHMOLLGRUBER, P.: *The NACRE Innovative Evaluation Platform and its Navigation & Control Strategies.* SAE international conference SAE 2011 Aerotech Congress & Exhibition; Toulouse, France. 2011.

LALIBERTE, A.S., HERRICK, J.E., RANGO, A., WINTERS, C.: *Acquisition, orthorectification, and object-based classification of unmanned aerial vehicle (UAV) imagery for rangeland monitoring*. Photogrammetric Engineering and Remote Sensing, 76, 6, p. 661–672. 2010.

MERZ, T., CHAPMAN, S. *Autonomous unmanned helicopter system for remote sensing missions in unknown environments.* International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII-1/C22. 2011.

MISTELE, B., SCHMIDHALTER, U.: *Spectral measurements of the total aerial N and biomass dry weight in maize using a quadrilateral-view optic*. Field Crops Res., 106, p. 94-103. 2008.

NEBIKER, S., ANNEN, A., SCHERRER, M., OESCH, D.: *A Light-Weight Multispectral Sensor for Micro UAS - Opportunities for Very High Resolution Airborne Remote Sensing*. ISPRS Congress. Beijing, China, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XXXVII. Part B1: p. 1193-1199. 2008.

RECKLEBEN, Y., ISENSEE, E.: *Vergleich von Sensorsystemen für die N-Düngung*. Landtechnik 2005, 3, p. 138-139. 2005.

ROVIRA-MÁS, F., ZHANG, Q., REID, J.F.: *Creation of Three-dimensional Crop Maps based on Aerial Stereoimages*. Biosystems Engineering, 90, 3, p. 251-259. 2005.

STEIN, I.Ø., TOMAS, I., KNUT, K., KORSAETH, A.: *Comparisons of two hand-held, multispectral field radiometers and a hyperspectral airborne imager in terms of predicting spring wheat grain yield and quality by means of powered partial least squares regression.* Journal of Near Infrared Spectroscopy, 18, 4, p. 247–261. 2010.

TEAL, R.K., TUBANA, B., GIRMA, K. , FREEMAN, K. W., ARNALL, D. B., WALSH, O., RAUN, W. R.: *In-season prediction of corn grain yield potential using normalized difference vegetation index*. Agronomy Journal, 98, p.1488–1494. 2006.

XIANG, H., TIAN, L.: *Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV)*. Biosystems Engineering 108:174-190. 2011.

ZHANG, L., CLARKE, M.L., STEVEN, M.D., JAGGARD, K.W.: *Spatial patterns of wilting in sugar beet as an indicator for precision irrigation.* Precision Agriculture, 12, 2, p. 296-316. 2011.

This paper has been accepted as a „peer reviewed paper“ by scientific committee

ATV Rollover Prevention System including Varying Grip Conditions and Bank Angle

Mathieu Richier[1], Roland Lenain[1], Benoit Thuilot[2, 3], Michel Berducat [1]
[1]Cemagref, 24 avenue des Landais, 63172 Aubière, France
[2]Clermont université, Université Blaise Pascal, Institut Pascal, BP 10448, 63000 Clermont-Ferrand, France
[3]CNRS, UMR 6602, Institut Pascal, 63171 Aubière, France

Abstract
In this paper, an algorithm dedicated to light ATVs, which estimates and anticipates the rollover, is proposed. It is based on the on-line estimation of the Lateral Load Transfer (LLT), allowing the evaluation of dynamic instabilities. The LLT is computed thanks to a dynamical model split into two 2D projections. Relying on this representation and a low cost perception system, an observer is proposed to estimate on-line the terrain properties (grip conditions and slope), then allowing to deduce accurately the risk of instability. Associated to a predictive control algorithm, based on the extrapolation of rider's action, the risk can be anticipated, enabling to warn the pilot and to consider the implementation of active actions.

Keywords
Light All-Terrain Vehicles (ATV), quad bikes, dynamic stability, Lateral Load Transfert (LLT), lateral rollover, sliding, modelling, observers and slope.

1 INTRODUCTION

Thanks to their high manoeuvrability, quad bikes are more and more popular and especially in agricultural context. Nevertheless, their mechanical properties lead to a significant rollover risk which constitutes the main cause of serious accidents for All-Terrain Vehicles (ATV) (almost 50 of ATV crashes as mentioned in [1] and [2]). If a rollover protective structure (ROPS) may limit health damages, it is not convenient for light vehicles. In contrast, the development of active devices, improving the stability of ATVs, constitutes a promising solution.

In on-road context, some stability systems have been designed in order to improve vehicle stability such as [3], [4] and [5]. As they use a linear tire model, these algorithms are not adapted to large grip condition variations, encountered in off-road context.

Some systems dedicated to off-road mobile robots have been developed like [6], [7], [8] and [9]. However they are hardly transposable to light ATV, since the required sensors remain expensive such as highly accurate INS or RTK GPS. Moreover, the accessibility of GPS data cannot be ensured when the ATV moves in natural environment (trees, mountains, building, etc.).

In previous work [10], a rollover risk prevention system dedicated to high speed ATVs has been proposed, based on low cost sensing equipment. It estimates on-line the tire-ground friction. The Lateral Load Transfer (LLT) has been chosen as a relevant stability criterion among several rollover indicators [11], because low cost sensing equipment is sufficient to estimate it. The main limitation of this system was the assumption of a flat ground, which is not representative of off-road applications. A less important limitation was a singularity in the algorithm, which entails to stop grip conditions update. A new modelling and a new approach to the grip condition observation are proposed in this paper. With low cost sensing equipment composed of a 3-axes accelerometer/gyrometer, a Doppler radar and a steering angle sensor, the LLT can be predicted whatever the grip conditions and the slope. More precisely, a bicycle model is associated to an adapted backstepping observer, which estimates the sliding parameters and the slope. These estimations are then used within a prediction algorithm based on a roll model, in order to anticipate the LLT.

The paper is organized as follows: first, the vehicle modelling is depicted. It allows the rollover metric computation. As the LLT expression requires the knowledge of the sliding parameters, an adapted backstepping observer is developed in the second part. In the same part, a prediction algorithm allowing the anticipation of LLT time-evolution is described. Finally, full scale experiments (with a commercial quad bike) are presented to investigate the capabilities and the applicability of the proposed approach.

2 ROLLOVER METRIC COMPUTATION

2.1 Dynamic model

In order to achieve on-line LLT computation when the ground is uneven, the global vehicle modelling depicted on **Figure 136** is considered. The dynamical model of vehicle is split into two models. The first model represents a yaw 2D projection (shown on **Figure 136(a)**) assuming a flat ground. In order to account for ground variation, a lateral force (P_1) is added. Relying on the state observer described in section III, this model enables the estimation of the sliding parameters (sideslip angles β, α_f, α_r and lateral forces F_f and F_r) and the bank angle (θ), which significantly impact the risk of rollover. These sliding conditions are then injected into the second model: a roll 2D projection (shown on **Figure 136(b)**) used to estimate the LLT.

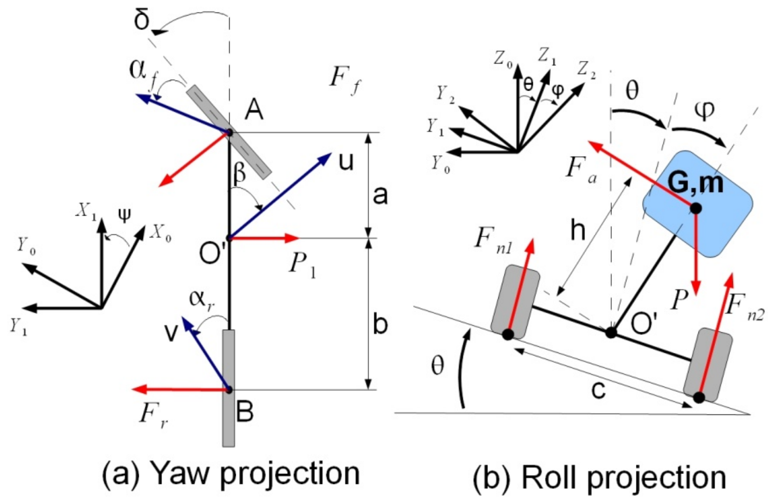


Figure 136: Dynamic bicycle model with sliding parameters.

The variables and parameters used in the sequel, reported on **Figure 136(a)** and **Figure 136(b)** are listed below:

- ψ is the vehicle yaw angle,
- β is the vehicle global sideslip angle,
- α_r is the vehicle rear sideslip angle,
- α_f is the vehicle front sideslip angle,
- θ is the bank angle of the terrain in the roll projection,
- δ is the steering angle,
- v is the linear velocity at the center of the rear axle,

- u is the linear velocity at the roll center O',
- a and b are the front and rear vehicle half-wheelbases,
- c is the vehicle track,
- h is the distance between the roll center and the vehicle center of gravity G,
- I_x, I_y, I_z are the roll, pitch and yaw moments of inertia,
- $P=mg$ is the gravity force on the suspended mass m, with g denoting the gravity acceleration,
- $P_1=mg\sin(\theta)$ is the influence of the gravity force on the lateral dynamics,
- F_{n1} and F_{n2} are the normal component of the tire-ground contact forces on the vehicle left and right sides,
- $F_a(\varphi)$ is a restoring-force parameterized by k_r and b_r, the roll stiffness and damping coefficients:

$$\vec{F_a} = \frac{1}{h}(k_r\varphi + b_r\dot{\varphi})\ \vec{y_2} \tag{1}$$

where φ is the roll angle of the suspended mass associated to the roll dynamics, depicted on **Figure 136**. In section 2.4, a way to calculate φ is given. The parameters k_r and b_r are evaluated previously thanks to a preliminary calibration procedure.

2.2 Contact model

The forces F_r and F_f acting on lateral dynamics widely depend on grip conditions. As a result, a tire-ground model is mandatory. Among several models describing the sliding phenomena (such as Pacejka or LuGre model [12], [13]), the linear model (2) is considered:

$$\begin{aligned} F_f &= C_f(.)\alpha_f \\ F_r &= C_r(.)\alpha_r \end{aligned} \tag{2}$$

Its main advantage lies in the few numbers of parameters to be known. Nevertheless, in order to take into account the non-linearity of the contact and the variations of grip conditions, cornering stiffnesses (C_f and C_r) are considered as varying. They are on-line adapted thanks to the observer detailed in section 3.

2.3 Motion equations in yaw frame

Based on both the linear tire model and the bicycle model representation depicted on **Figure 136(a)**, the equations of motion can be derived using the fundamental principle of the dynamic. In the yaw frame, longitudinal forces as well as roll and pitch motions are neglected. Moreover, the influence of the bank angle is accounted via the addition of a gravity force in acceleration equations. Motion equations are then finally given by:

$$\begin{aligned} \ddot{\psi} &= \frac{-a\cos(\delta)F_f + bF_r}{I_z} \\ \dot{\beta} &= -\frac{F_f\cos(\delta-\beta) + F_r\cos(\beta)}{um} \\ &\quad - \frac{g}{v}\sin(\theta)\cos(\beta) - \dot{\psi}\cos(\beta) \end{aligned} \qquad \begin{aligned} \ddot{\psi} &= \frac{-a\cos(\delta)F_f + bF_r}{I_z} \\ \dot{\beta} &= -\frac{F_f\cos(\delta-\beta) + F_r\cos(\beta)}{um} \\ &\quad - \frac{g}{v}\sin(\theta)\cos(\beta) - \dot{\psi}\cos(\beta) \end{aligned} \tag{3}$$

As this paper deals with dynamic LLT estimation, the velocity is assumed to be always strictly positive. As a result the condition $u\neq0$ is always met.

2.4 Roll motion and LLT computation

The Lateral Load Transfer (LLT) represents the unbalanced repartition of the normal components of the tire-ground contact forces. It is mathematically defined as:

$$LLT = \frac{F_{n1} - F_{n2}}{F_{n1} + F_{n2}} \tag{4}$$

According to definition (4), the LLT reaches ±1 when two wheels on a vehicle's side lift off, which is representative of a rollover risk. In practice a threshold can be chosen above which the vehicle is considered in a hazardous situation. This threshold is chosen as 80% (classical value used in the literature) in order to define a safety margin.

Thanks to the fundamental principle of the dynamic applied to the roll model depicted **Figure 136(b)**, and assuming that $\ddot{\theta} << \ddot{\varphi}$ and $\dot{\theta} << \dot{\varphi}$, dynamics equation for the roll angle φ and for the normal forces are given (5):

$$\begin{aligned}
\ddot{\varphi} &= \frac{1}{h\cos(\varphi)}[h\dot{\varphi}^2 \sin()\varphi) + h\dot{\psi}^2 \sin(\gamma)\cos(\theta) + u\dot{\psi}\cos(\theta)\cos(\beta) \\
&\quad + \dot{u}sin(\beta) + u\dot{\beta}\cos(\beta) - \frac{k_r\varphi + b_r\dot{\varphi}}{mh}\cos(\varphi) + g\sin(\theta)] \\
F_{n1} + F_{n2} &= m[-h\ddot{\varphi}sin(\varphi) - h\dot{\varphi}^2\cos(\varphi) + g\cos(\theta) \\
&\quad - u\dot{\psi}\sin(\theta)\cos(\beta) - \frac{k_r\varphi + b_r\dot{\varphi}}{mh}\sin(\varphi) - \dot{\psi}^2\sin(\gamma)\sin(\theta)] \\
F_{n1} - F_{n2} &= \frac{2}{c}[I_x\ddot{\varphi} + (I_z - I_y)\dot{\psi}^2\frac{\sin(2\gamma)}{2}] - h\sin(\varphi)(F_{n1} + F_{n2})]
\end{aligned} \tag{5}$$

where $\gamma=\theta+\varphi$.

Consequently, as soon as the roll angle φ can be calculated using (5), the LLT can be evaluated thanks to the normal force expressions.

In view of (5), the calculation of φ requires the knowledge of sideslip angle (β) whose value depends on cornering stiffnesses C_f and C_r in view of (3). As quad bikes are expected to move on a natural and slippery ground, grip conditions have an important influence and are moreover varying. Since these variables cannot be measured, their on-line adaptation is then required in order to obtain relevant estimation and prediction of the LLT. Therefore a backstepping observer has been designed to supply on-line their values. Moreover a prediction algorithm is mandatory, if the LLT has to be anticipated in order to prevent the hazardous situations.

3 ROLLOVER PREVENTION

3.1 System overview

The developed system aiming at ATV rollover prevention is summarized on **Figure 137**. It is composed of:

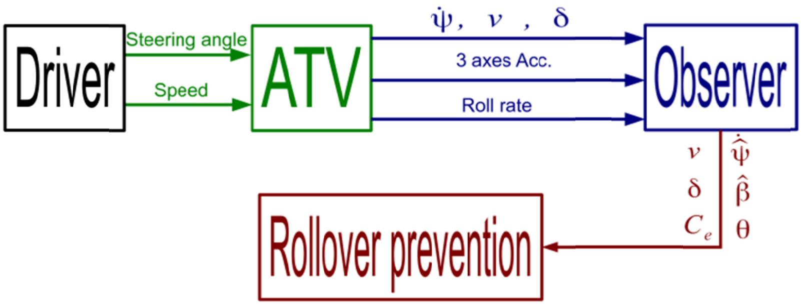


Figure 137: Algorithm overview.

3.1.1 ATV box

The ATV is manually controlled, i.e. the driver specifies the vehicle speed v and steering angle δ. As described in the introduction, the measured data are the roll/yaw rate, the accelerations, the speed and the steering angle.

3.1.2 Observer box

Contact conditions are then on-line estimated. However for observability reasons (see [10]), the two cornering stiffnesses cannot be estimated separately, and are therefore considered to be equal to a global virtual cornering stiffness C_e. An estimation of the sideslip angle $\hat{\beta}$ and yaw rate $\dot{\hat{\psi}}$ is also supplied. Moreover, the bank angle estimation is integrated into the observer thanks to the measure of the lateral acceleration.

3.1.3 Rollover prevention box

Relying on the measured and observed variables (v, δ, C_e, $\dot{\hat{\psi}}$ and $\hat{\beta}$), future LLT values can be predicted on-line, in order to prevent the risk of rollover.
The observer and LLT prediction algorithms are more precisely described in the following sections.

3.2 Observer design

As the front/rear/global sideslip and steering angles can be large, a non-linear system (6) is considered:

$$\ddot{\widehat{\psi}} = a_{11}\dot{\widehat{\psi}} + a_{12}\widehat{\beta} + b_1\delta \qquad \widehat{\alpha}_r = \arctan(\tan(\widehat{\beta}) - \frac{b\dot{\widehat{\psi}}}{v})$$

$$\dot{\widehat{\beta}} = -\frac{\widehat{F} + mg\sin(\theta)\cos(\widehat{\beta})}{vm} - \dot{\widehat{\psi}}\cos(\widehat{\beta})$$

$$\widehat{F} = C_e(\widehat{\alpha}_f\cos(\delta - \widehat{\beta}) + \widehat{\alpha}_r\cos(\widehat{\beta})) \qquad \widehat{\alpha}_f = \arctan(\tan(\widehat{\beta}) + \frac{a\dot{\widehat{\psi}}}{v}) + \delta$$

$$a_{11} = \frac{-(a^2\cos(\delta)+b^2)C_e}{vI_z}, a_{12} = \frac{(b-a\cos(\delta))C_e}{I_z}, b_1 = \frac{a\cos(\delta)C_e}{I_z} \qquad (6)$$

Where $\dot{\hat{\psi}}$, $\hat{\beta}$, $\hat{\alpha}_r$ and $\hat{\alpha}_f$ are respectively the observed yaw rate, global sideslip angle, rear sideslip angle and front sideslip angle. $\hat{F}$ is named the global lateral force. In order to compute the LLT, β and C_e have to be estimated from (6). With this aim, a backstepping approach composed of 4 steps is proposed. An overview is depicted on **Figure 138**.

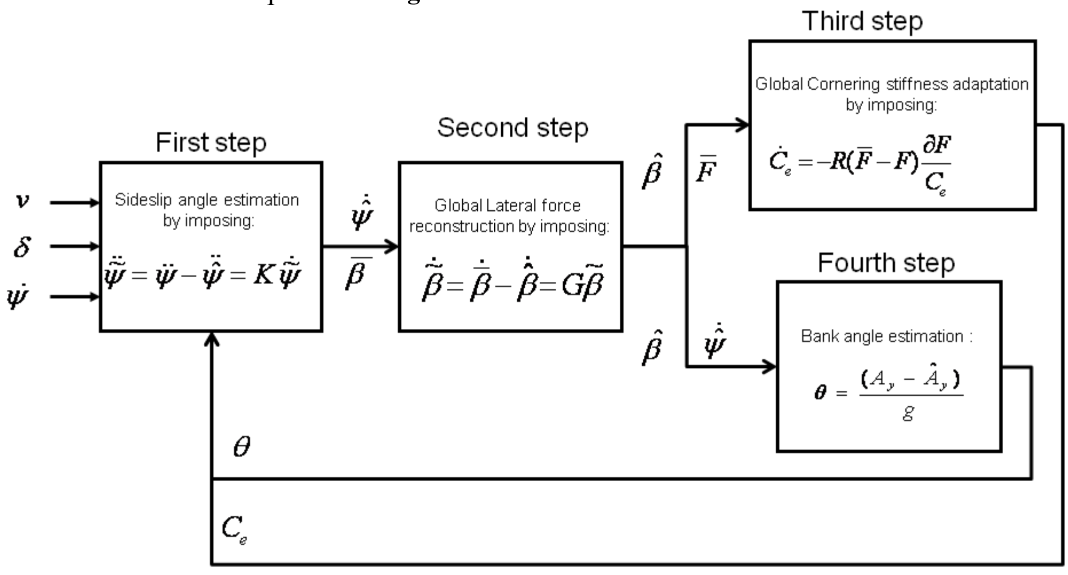


Figure 138: Observer overview.

3.2.1 First step "Sideslip angle estimation"

The first step consists in treating β as a control input (denoted $\overline{\beta}$), to be designed in order to impose the following dynamic on the observed yaw rate error $\dot{\tilde{\psi}}$:

$$\ddot{\tilde{\psi}} = \ddot{\psi} - \ddot{\hat{\psi}} = K\dot{\tilde{\psi}},\ K < 0 \tag{7}$$

where $\ddot{\psi}$ is derived from the measured yaw rate. Injecting (7) into the first equation in (6) leads to the following expression for control variable $\overline{\beta}$:

$$\overline{\beta} = \frac{\ddot{\psi} - K\dot{\tilde{\psi}} - a_{11}(C_e)\dot{\hat{\psi}} - b_1(C_e)\delta}{a_{12}(C_e)} \tag{8}$$

under condition $a_{12} \neq 0$, which is ensured in practice. Since $\overline{\beta}$ ensures that $\dot{\hat{\psi}}$ converges to the actual value $\dot{\psi}$ supplied by the gyrometer, $\overline{\beta}$ can be considered as a relevant estimation of the actual global sideslip angle.

3.2.2 Second step "Lateral force reconstruction"

Just as in the first step, *F* is then treated as a control input (denoted $\overline{F}$) to be designed to impose that $\tilde{\beta} = \overline{\beta} - \hat{\beta}$ converges to 0 with the following dynamic:

$$\dot{\tilde{\beta}} = \dot{\overline{\beta}} - \dot{\hat{\beta}} = G\tilde{\beta},\ G < 0 \tag{9}$$

where $\dot{\overline{\beta}}$ is derived from $\overline{\beta}$. Injecting (9) into the second equation of (6) leads to the following expression for control variable $\overline{F}$:

$$\overline{F} = -mv(\dot{\overline{\beta}} - G\tilde{\beta} + \dot{\hat{\psi}}\cos(\overline{\beta})) - mg\sin(\theta)\cos(\overline{\beta}) \tag{10}$$

Since $\overline{F}$ ensures that β converges to the actual value $\overline{\beta}$, $\overline{F}$ can be considered as an estimation of the actual lateral force.

3.2.3 Third step "Cornering stiffness adaptation"

This step consists in adapting C_e in order to ensure the convergence of *F* to $\overline{F}$ as defined in equation (6). In view of (6), the adaptation of C_e cannot be achieved when F=0, which occurs especially when moving straight ahead on a flat ground.

The adaptation law used in [10] was stopped during the straight line because of a singularity into the equation (division by δ=0), and it might generate a divergence when the observation restarts.

In order to avoid an adaptation interruption in such a case, a MIT rule adaptation [14] is proposed to obtain the convergence:

$$\dot{C}_e = R(\overline{F} - \hat{F})\frac{\partial F}{\partial C_e} = R(\overline{F} - \hat{F})(\hat{\alpha}_f\cos(\delta - \hat{\beta}) + \hat{\alpha}_r\cos(\hat{\beta})) \tag{11}$$

with *R* a strictly positive gain.

As it can be seen on (11), the expression of $\dot{C}_e$ is never singular: when moving in straight line on an even ground, the global lateral force (*F*) tends to zero and consequently $\dot{C}_e = 0$. As a result the cornering stiffness adaptation is frozen in a natural way. This adaptation law (11) does not require to monitor a singularity during straight line motion.

3.2.4 Fourth step "Bank angle estimation"

The principle is to compare the lateral acceleration measured to the one estimated. As the accelerometer is able to measure the constant acceleration like the gravity, the measured lateral acceleration can be modelled into the equation (12).

$$Acc_{\text{measured}} = Acc_{\text{dynamic}} + g\sin(\theta) \tag{12}$$

Moreover, thanks to the yaw and roll representation shown on **Figure 136**, an estimated lateral acceleration (along $\vec{y}_1$) can be computed:

$$\widehat{Acc}_{dynamic} = \underbrace{u\dot{\widehat{\beta}}\cos(\widehat{\beta}) + \dot{u}\sin(\widehat{\beta}) + u\dot{\widehat{\psi}}\cos(\widehat{\beta})}_{\text{Yaw acceleration}} + \underbrace{h\dot{\varphi}^2 - h\ddot{\varphi}}_{\text{negligible roll acceleration}} \tag{13}$$

Finally, the bank angle can be easily estimated with the equation (12) and (13):

$$\theta = \frac{Acc_{\text{measured}} - \widehat{Acc}}{g} \tag{14}$$

3.3 LLT prediction

The previous observer on-line supplies a realist estimation of C_e , β, θ describing respectively the grip conditions, the global sideslip and the bank angle. All variables in model equations (3) and (5) are therefore available, and the LLT can then be predicted by integrating these equations over some temporal horizon *H*. If this prediction reaches a value superior than the threshold (i.e. $LLT_{predicted} \geq 0.8$), the driver is warned of a rollover risk.

More precisely, to perform the integration, the slow-varying variables, i.e. the cornering stiffness C_e and the bank angle θ, are supposed constant over the horizon *H*. On the contrary, the driver's inputs (i.e. the steering angle δ and the speed *v*) have an important influence on the short term evolution of the LLT. Therefore, it is proposed to extrapolate them using a linear function if *v* and/or δ present an evolution raising the instability. Otherwise they are kept constant over the horizon *H*. In this way, the values of the LLT predicted from equations (3) and (5) are at worst overestimated, as it suits for a security device. Precisely, the extrapolation law that has been chosen is:

$$v(t_p) = \begin{cases} v(t) + dt\cdot\dot{v}(t), & \text{if } (v(t),\dot{v}(t)) > (0,0) \\ v(t) \text{ otherwise} \end{cases} \qquad \delta(t_p) = \begin{cases} \delta(t) + dt\cdot\dot{\delta}(t), & \text{if } (\delta(t),\dot{\delta}(t)) > (0,0) \\ \delta(t) + dt\cdot\dot{\delta}(t), & \text{if } (\delta(t),\dot{\delta}(t)) < (0,0) \\ \delta(t) \text{ otherwise} \end{cases}$$
$$with\ t_p = t + dt\ and\ 0 < dt < H \tag{15}$$

4 RESULTS

4.1 Setup testbed

Figure 139: MF400H, Massey Fergusson quad bike used for experiments.

In order to validate the observer and the relevance of the LLT prediction proposed in section 3, experimental results are presented. They have been performed with a quad bike MF400H, manufactured by Massey Fergusson and depicted on **Figure 139**. Its dynamic parameters m, I_z, k_r, b_r, h, a and b have been preliminary calibrated, and it is equipped with the following sensors:

- a Xsens MTI IMU providing accelerations and angular velocities,
- a Doppler radar supplying the linear speed,
- an angular sensor providing the steering angle.

This set of sensors constitutes a low cost perception system (compared to the ATV cost) enabling LLT estimation without requiring for expensive sensors. In addition, dynamometric sensors supplying tire/ground forces have been set up at each wheel. They provide a ground truth, but are not used into the algorithm.

4.2 Experimental results

4.2.1 ATV experimental path

The path described on **Figure 140** has been performed on a mixed flat and sloping wet grass ground, at a speed between 3ms^{-1} and 5 ms^{-1}. It is composed of a straight line part executed on a 5-15° sloping ground, a U-turn on a partially even area, a second straight line on the same sloping ground, a U-turn on an even area and a third straight line on the same sloping ground.

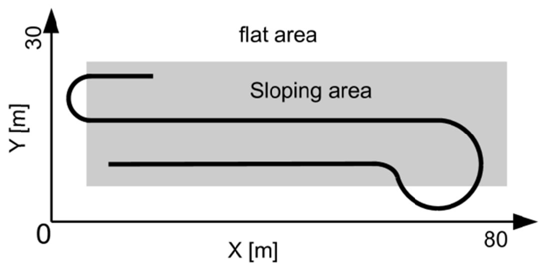


Figure 140: ATV experiment path.

4.2.2 Estimated bank angle

On **Figure 141**, the bank angle profile estimated during experiments is depicted. It corresponds to actual slope values recorded (±15°) previously by an operator. It then constitutes a satisfactory estimation. According to equation (14), θ includes the bank angle estimation and the error in the lateral acceleration estimation.

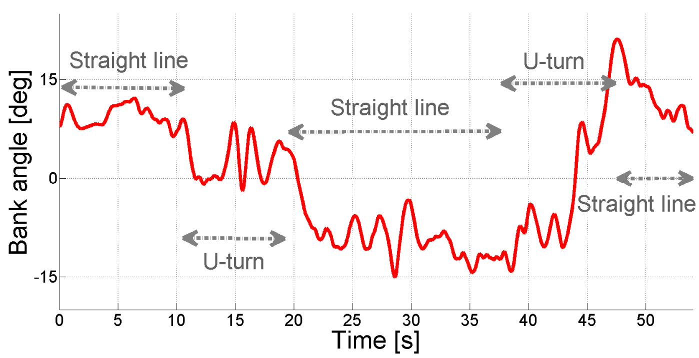


Figure 141: Estimated bank angle.

4.2.3 Observer dynamics

Three experiments have been achieved with different initial conditions for the tire cornering stiffness: 50000 N.rad^{-1}, 20000 N.rad^{-1} and 5000 N.rad^{-1}. The estimated tire cornering stiffnesses are then represented on **Figure 142**.

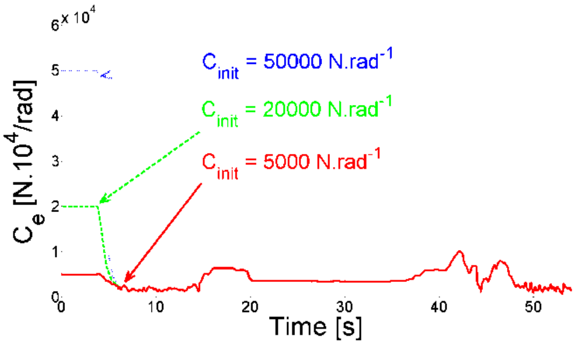


Figure 142: Tire stiffness adaptation.

First, the estimated cornering stiffnesses converge all to the same value. This demonstrates that the choice of the initial condition, which is uneasy, is satisfactorily not crucial. Moreover, the order of magnitude (4000) is representative of the value for wet grass terrains considering a quad bike.

Secondly, the cornering stiffnesses are convergent despite the first straight line (0-10s) because of the small slope. The sideslip generated by the slope is sufficient to adapt the cornering stiffness (see (3)).

Thirdly, the cornering stiffnesses are constant during the straight line (0-10s, 20-37s and 48-54s), as expected in section 3.2.3.

Finally C_e suddenly decreases at 42s (corresponding to an inversion of the steering angle sign), which is representative of non-linear tire behaviour when slip angle changes quickly.

4.2.4 Slope influence on the LLT estimation

On **Figure 143** the LLTs estimated at the current instant (i.e. when H=0s) with and without the bank angle are depicted (respectively in solid and dashed line) and they are compared to the LLT measured thanks to dynamometric sensors. The quad bike enters the slope at 17s, then the U-turn occurs between 36s and 46s, and finally the vehicle comes back on the slope part.

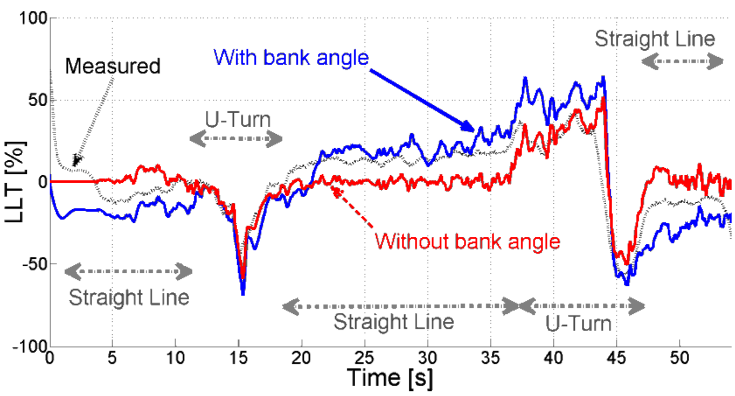


Figure 143: Experiment results of the LLT estimation.

The LLT estimated without accounting for the bank angle stays around 0° in straight line parts, as expected. In this case, θ is indeed mainly responsible for the LLT. On the contrary, the LLT accounting for the bank angle is almost superposed on the actual LLT, especially during the slope parts.

Nevertheless it can be observed some local inaccuracies, which can be due to the driver's behaviour, neglected in the approach. Since quad bikes are light vehicles, the driver mass is important (for this experiment, the driver's mass represents 25% of the total mass), and his behaviour has a significant influence. As demonstrated in [15], the position of the driver may change significantly the location of the center of gravity of the overall system and consequently impacts the LLT values. This can explain the LLT overestimation in the first turn-about (12-18s and 37-43s).

This experiment shows the importance of taking into account the slope to estimate accurately the LLT. But the driver has to be informed of the risk before it appears, therefore a prediction algorithm is mandatory. The next section discusses the efficiency of the prediction algorithm developed in section 3.3.

4.2.5 Rollover risk indicator

The LLT estimated with the bank angle and the measured LLT are plotted again on **Figure 144**, respectively in blue and black lines, and compared to the predicted LLT in red:

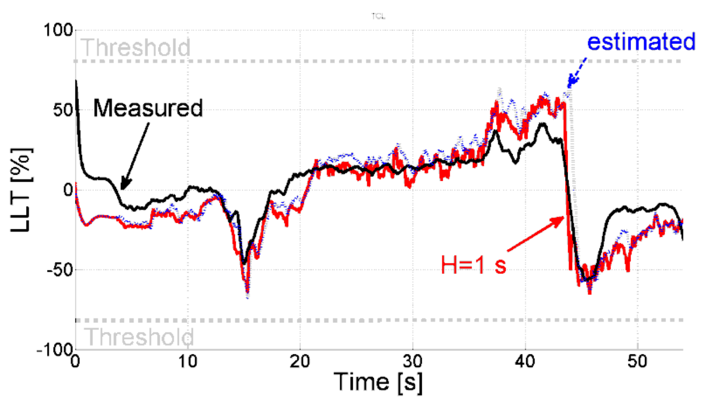


Figure 144: Experiment results of the LLT prediction.

First, it can be noticed that the three curves are almost superposed in steady state conditions: during the straight line parts and during the constant curves. As expected since the ATV motion is then stationary: the predicted LLT is identical to the current value.

In contrast, during the transient phase (42-45s) where rollover may occur, the predicted LLT satisfactorily precedes and overestimates the actual LLT.

Consequently, this rollover indicator is able to prevent the lift-off risk for the ATVs on natural ground.

5 CONCLUSION AND FUTURE WORK

This paper proposes an algorithm able to anticipate a rollover risk for ATVs motion on natural ground. An adapted backstepping observer, based on a bicycle model, has been designed in order to estimate the dynamic variables (sideslip angle, cornering stiffness) allowing to adapt to varying conditions and to estimate the slope. Then, relying on a roll model, the LLT is on-line anticipated. The main contributions lie in the consideration of the terrain slope and in the grip condition adaptation. As demonstrated in the experiments, the LLT can be predicted accurately whatever the terrain conditions (sliding, slope). Moreover the sensing equipment is limited to low cost sensors excluding expensive INS or GPS. Nevertheless the driver's behaviour, which influences the estimated LLT, is not taken into account. Therefore, in order to avoid unnecessary warnings, current developments aim at developing a low cost system which accounts also for the driver's behaviour.

ACKNOWLEDGMENTS
With many thanks for the CCMSA support.

REFERENCES

"*All terrain vehicle enforcement and safety report 2006*," Wisconsin department of natural resources, USA, Tech. Rep., 2007.

CCMSA, "Accidents du travail des salariés et non salariés agricoles avec des quads," Observatoire des risques professionels et du machinisme agricole, Paris, France, Tech. Rep., 2006.

D. BEVLY, J. RYU, and J. GERDES, "Integrating INS sensors with GPS measurements for continuous estimation of vehicle sideslip, roll, and tire cornering stiffness," IEEE Trans. Intell. Transport. Syst., vol. 7, no. 4, p. 483 –493, 2006.

D. PIYABONGKARN, R. RAJAMANI, J. GROGG, and J. LEW, "Development and experimental evaluation of a slip angle estimator for vehicle stability control," *IEEE Trans. Contr. Syst. Technol.*, vol. 17, no. 1, p. 78 –88, 2009.

B. SCHOFIELD, T. HAGGLUND, and A. RANTZER, "Vehicle dynamics control and controller allocation for rollover prevention," in *IEEE International Conference on Control Applications*, p. 149 –154, 2006.

G. BESSERON, C. GRAND, F. BEN AMAR, and P. BIDAUD, "Decoupled control of the high mobility robot hylos based on a dynamic stability margin," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, p. 2435 –2440, 2008.

K. OHNO, V. CHUN, T. YUZAWA, E. TAKEUCHI, S. TADOKORO, T. YOSHIDA, and E. KOYANAGI, "Rollover avoidance using a stability margin for a tracked vehicle with sub-tracks," in *IEEE/RSJ Safety International Workshop on Security Rescue Robotics*, p. 1 –6, 2009.

S. PETERS and K. IAGNEMMA, "An analysis of rollover stability measurement for high-speed mobile robots," in *Robotics and Automation, ICRA*, May 2006, p. 3711 –3716.

M. SPENKO, Y. KURODA, S. DUBOWSKY, and K. IAGNEMMA, "Hazard avoidance for high-speed mobile robots in rough terrain," *Journal of Field Robotics*, vol. 23, no. 5, p. 311–331, 2006.

N. BOUTON, R. LENAIN, B. THUILOT, and P. MARTINET, "A rollover indicator based on a tire stiffness backstepping observer : Application to an all-terrain vehicle," in *Int. Conf. on Intelligent RObots and Systems (IROS)*, Nice, France, 2008.

S. PETERS, "Stability measurement of high-speed vehicles," *Vehicle System Dynamics*, vol. 47, no. 6, p.701–720, 2009.

H. B. PACEJKA, *Tire and vehicle dynamics*. Society of Automotive Engineers, 2002.

C. CANUDAS DE WIT and P. TSIOTRAS, "Dynamic tire friction models for vehicle traction control," in *Decision and Control*, vol. 4, 1999, p. 3746 –3751, 1999.

K. ASTROM and B. WITTENMARK, *Adaptive control (2nd edition)*. New York: Addison-Wesley, 1994.

N. BOUTON, "All-terrain vehicles dynamic stability. new solutions: Application to light vehicles like quad bike," Ph.D. dissertation, Blaise Pascal University, 2009.

Sensor Integration I

Modular Positioning using Different Motion Models

Jürgen Schweitzer
Institute of Engineering Geodesy, University of Stuttgart, Germany

Abstract
The estimation of the current position and orientation of a moving machine on the base of different measurement data, for example from a global satellite navigation system or inertial measurement units, is one of the main topics in navigation, geodesy and other engineering applications. In this field the Kalman Filter is the state of the art in filtering raw measuring observations over time.

In this paper, an application of a Kalman filter approach is presented, in which a fast adaption of different sensor combinations or measurement data is given. The application is designed for measurement data from a multi-sensor system consisting of measurement data from GNSS sensors and additional sensors like odometers, gyroscopes and accelerometers. The sensor data reflect different accuracy levels, especially for GNSS. Therefore, unified interfaces and synchronized measurement values are prerequisite.

Different motion models are integrated to get a problem-related estimation of the position. In detail, motion models assuming a straight line and a circle are used to predict the state vector of the system. Additionally, a model is implemented, for which the two approaches are combined. The choice for the particular model is selected depending on the situation.

The main focus of this paper is the modularity of the approach with respect to motion model, integrated sensor as well as accuracy level of the sensors. The further developed system was named MOPSY (Modular Positioning Systems) 7 years ago (Schwieger et al, 2005). Now, new results regarding different motion models are available. The paper presents results using different motion model approaches. The accuracy level reaches the values of 50 – 150 cm RMS for absolute position with GPS, odometers, gyroscopes and accelerometers. Using precise DGPS instead of GPS, the results of the absolute position are better by a factor of around 10.

Keywords
Kalman Filter, motion models, sensor integration

1 INTRODUCTION

In general, one can say that the Kalman Filter is used to remove the disturbances caused by the measurement instruments and processes, taking into account the dynamic behaviour of the measured object. Particularly in autonomous and assisted navigation, where sensor data like acceleration, velocity, orientation, etc. is used, it is a popular procedure. The Kalman Filter is named after Rudolph E. Kalman, who published his famous paper describing a recursive solution to the discrete-data linear filtering problem in 1960 (Kalman, 1960).

In detail, it is essentially a set of mathematical equations that implement a predictor-corrector type estimator which is optimal in the sense that it minimizes the estimated error covariance – when some presumed conditions are met (Welch and Bishop, 2001).

A typical situation in which a Kalman Filter is used is shown in **Figure 145**. A system with certain system errors has measuring devices to provide the value of certain quantities. Only the "prediction system state" and the "observed measurements" are available for estimation purposes. The filter has now the task to process this information and output an optimal estimate of the system state. Often, the variable of interest (to describe the "state" of a system) cannot be measured directly and so the filter has to convert the measured quantities to the state quantities. For example, a vehicle provides the velocity and the acceleration from which a position can be derived. Furthermore, any measurement will be corrupted to some degree by noise, biases and device inaccuracies. As well the filter has to extract valuable information from a noisy signal. So, a Kalman filter combines all available measurement data, plus prior knowledge about the system and measuring devices, to produce an estimate of the desired variables in such a manner that the error is minimized statistically.

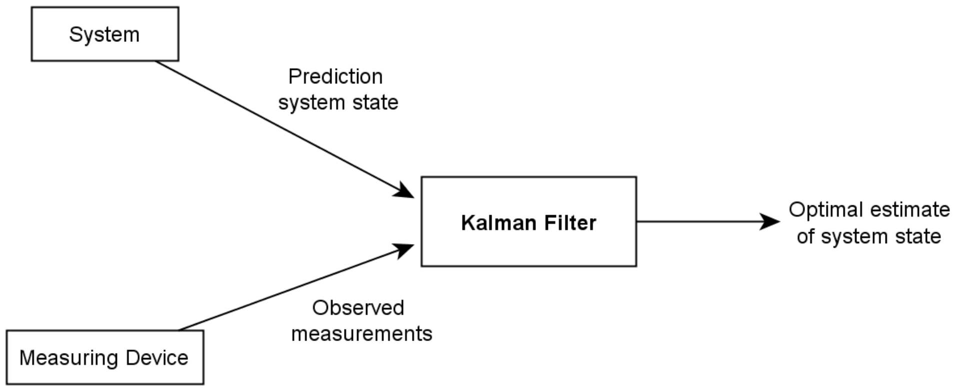


Figure 145: Typical Kalman filter application

In Maybeck (1979) the functionality of the Kalman Filter is explained in a simple way: If we were to run a number of candidate filters many times for the same application, then the average result of the Kalman Filter would be better than the average result of any other.

2 THEORETICAL BASICS

2.1 The Kalman Filter

Based on the linear stochastic difference equation (see Gelb, 1974 and Ramm, 2008)

$$\boldsymbol{x_k} = \boldsymbol{F x_{k-1}} + \boldsymbol{B u_{k-1}} + \boldsymbol{C w_{k-1}} \tag{1}$$

and the measurement

$$\boldsymbol{z_k} = \boldsymbol{H_k x_k} + \boldsymbol{v_k} \tag{2}$$

The Kalman filter is trying to estimate the state $\boldsymbol{x}$ of a discrete time controlled process. The random variables $\boldsymbol{w}$ and $\boldsymbol{v}$, which are both normal distributed, represent the process and measurement noise with the process noise covariance $\boldsymbol{Q_{ww}}$ and the measurement noise covariance $\boldsymbol{Q_{ll}}$ matrices.

$$\boldsymbol{p}(\boldsymbol{w}) \sim \boldsymbol{N}(0, \boldsymbol{Q_{ww}}) \tag{3}$$

$$\boldsymbol{p}(\boldsymbol{v}) \sim \boldsymbol{N}(0, \boldsymbol{Q_{ll}}) \tag{4}$$

The $\boldsymbol{F}$ matrix in the equation (1) is the state transition model which relates the state at the previous time step *k-1* to the state at the current step *k*. The matrix $\boldsymbol{B}$ implies the input control model and applies to the control vector $\boldsymbol{u}$ and with the **C** matrix non-deterministic external error influences are considered. The matrix $\boldsymbol{H}$ in equation (2) is the observation model which maps the state space into the observed space. It is also called design matrix.

The estimation of the state $\boldsymbol{x}$ at some time is divided into two steps called prediction (time update) and correction (measurement update). The prediction is responsible for projecting forward in time the current state and error covariance estimates. The correction is responsible for the feedback (in the form of noisy measurements).

With the assumption that the stochastic modelled variable *E(w)=0*, the prediction can be described with the following equations

$$\overline{\boldsymbol{x}}_{\boldsymbol{k}} = \boldsymbol{F}\widehat{\boldsymbol{x}}_{\boldsymbol{k-1}} + \boldsymbol{B}\boldsymbol{u}_{\boldsymbol{k-1}} \tag{5}$$

$$\overline{\boldsymbol{Q}}_{\boldsymbol{xx,k}} = \boldsymbol{F}\widehat{\boldsymbol{Q}}_{\boldsymbol{xx,k-1}}\boldsymbol{F}^{\boldsymbol{T}} + \boldsymbol{B}\boldsymbol{Q}_{\boldsymbol{uu}}\boldsymbol{B}^{\boldsymbol{T}} + \boldsymbol{C}\boldsymbol{Q}_{\boldsymbol{ww}}\boldsymbol{C}^{\boldsymbol{T}} \tag{6}$$

Where equation (5) is the prediction of the state vector $\boldsymbol{x}$ and equation (6) is the prediction of the cofactor matrix $\widehat{\boldsymbol{Q}}_{xx}$ of the state. Then, the correction is following. Here, the first step is the computation of the Kalman gain matrix **K** (7) which is used to update the state vector (8) and the covariance matrix (9).

$$\boldsymbol{K}_{\boldsymbol{k}} = \overline{\boldsymbol{Q}}_{\boldsymbol{xx,k}}\boldsymbol{H}^{\boldsymbol{T}}(\boldsymbol{H}\overline{\boldsymbol{Q}}_{\boldsymbol{xx,k}}\boldsymbol{H}^{\boldsymbol{T}} + \boldsymbol{Q}_{\boldsymbol{ll}})^{-\boldsymbol{1}} \tag{7}$$

$$\widehat{\boldsymbol{x}}_{\boldsymbol{k}} = \overline{\boldsymbol{x}}_{\boldsymbol{k}} + \boldsymbol{K}_{\boldsymbol{k}}(\boldsymbol{z}_{\boldsymbol{k}} - \boldsymbol{H}\overline{\boldsymbol{x}}_{\boldsymbol{k}}) \tag{8}$$

$$\widehat{\boldsymbol{Q}}_{\boldsymbol{xx,k}} = (\boldsymbol{I} - \boldsymbol{K}_{\boldsymbol{k}}\boldsymbol{H})\overline{\boldsymbol{Q}}_{\boldsymbol{xx,k}} \tag{9}$$

After each time and measurement step the process is repeated with the previous estimates (*k-1*) used to project or predict the new estimates (*k*). For further information about the equations (5) – (9) see Gelb (1974).

2.2 The Extended Kalman Filter

The equations above describe a discrete, linear Kalman filter. For our purposes the process of a vehicle movement (see chapter 3) to be estimated is non-linear. In this case the **Extended Kalman Filter** (EKF) is one option to be used (see Ramm, 2008 or Welch and Bishop, 2001). Here, the terms $\boldsymbol{F}\boldsymbol{x}_{k-1}, \boldsymbol{B}\boldsymbol{u}_{k-1}$ from equation (1) and (5) and $\boldsymbol{H}_k\boldsymbol{x}_k$ from equation (2) and (8) are replaced with non-linear functions $f(\widehat{\boldsymbol{x}}_{k-1}, \boldsymbol{u}_{k-1})$ and $h(\boldsymbol{x}_{k-1}, \boldsymbol{v}_{k-1})$. For the prediction and the update of the covariance matrix, the required linearization is realised by the calculation of a Taylor series for the system and measurement equations. The terms of higher order are neglected and so the equations (1) and (2) are changing while a subscript *k* is attached to the matrices ***F**, **B*** and ***H*** because they are not time-invariant any more. ***F**, **B*** and ***H*** are the Jacobian matrices and can be expressed by the equations (10)-(12).

$$\boldsymbol{F}_{\boldsymbol{k-1}} = \frac{\partial f(\widehat{\boldsymbol{x}}_{k-1}, \boldsymbol{u}_{k-1})}{\partial \boldsymbol{x}_{k-1}} \quad ; \quad \boldsymbol{B}_{\boldsymbol{k-1}} = \frac{\partial f(\widehat{\boldsymbol{x}}_{k-1}, \boldsymbol{u}_{k-1})}{\partial \boldsymbol{u}_{k-1}} \quad ; \boldsymbol{H} = \frac{\partial \boldsymbol{h}(x,v)}{\partial \boldsymbol{x}_k} \tag{10; 11; 12}$$

So the prediction of the system state occurs via the non-linear system equations

$$\overline{\boldsymbol{x}}_{\boldsymbol{k}} = f(\widehat{\boldsymbol{x}}_{\boldsymbol{k-1}}, \boldsymbol{u}_{\boldsymbol{k-1}}). \tag{13}$$

The predicted covariance matrix is calculated similarly to the linear way

$$\overline{\boldsymbol{Q}}_{\boldsymbol{xx,k}} = \boldsymbol{F}_{\boldsymbol{k-1}}\widehat{\boldsymbol{Q}}_{\boldsymbol{xx,k-1}}\boldsymbol{F}_{\boldsymbol{k-1}}{}^{\boldsymbol{T}} + \boldsymbol{B}_{\boldsymbol{k-1}}\boldsymbol{Q}_{\boldsymbol{uu,k-1}}\boldsymbol{B}_{\boldsymbol{k-1}}{}^{\boldsymbol{T}} + \boldsymbol{C}_{\boldsymbol{k-1}}\boldsymbol{Q}_{\boldsymbol{ww,k-1}}\boldsymbol{C}_{\boldsymbol{k-1}}{}^{\boldsymbol{T}} \tag{14}$$

The correction of the system state and the covariance matrix are realized similarly to the linear way (7)-(9)) with the difference that the ***F***, ***B*** and ***H*** are replaced by the Jacobian $\boldsymbol{F}_{k-1}$ and the term $\boldsymbol{H}_x$ in equation is replaced by $h(\boldsymbol{x}_{k-1}, \boldsymbol{v}_{k-1})$.

2.3 Modelling of Vehicle Movement

The state of a moving vehicle in two dimensions can be described in different ways. One possibility is the use of the velocity *v*, the orientation *φ* and the current position *X, Y*. Therefore, the state vector ***x*** may be written as follows:

$$\boldsymbol{x} = [X, Y, \varphi, v, \Delta\varphi] \tag{15}$$

The quantity $\Delta\varphi$ provides redundant information, because it can be derived from the orientation φ, but it is used here to stabilize the filtering.

On the base of this, we can define two exemplary vehicle movement models: **straight line** and **circle** drive. You can find modifications of this vehicle movement models for example in Sternberg (2000).

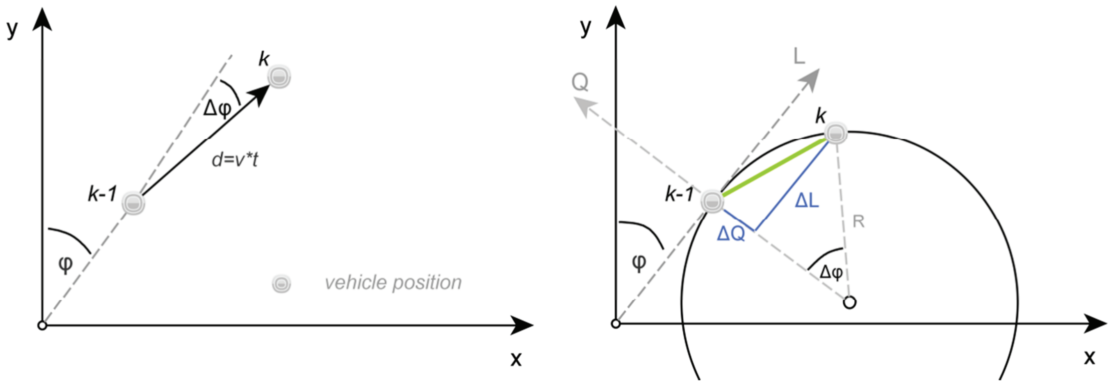


Figure 146: Straight line (left) and circle model (right)

2.3.1 Straight line drive

The straight line drive is an oversimplification of the real movement of a vehicle. One can assume that the vehicle trajectory is consisting of points equal to the measuring epochs and the vehicle is moving between two points on a straight line and the rotation is realized on the point.
The prediction of the position of the next epoch depends on a polar survey method (see **Figure 146**, left). So the prediction $\bar{\boldsymbol{x}}_k = f(\hat{\boldsymbol{x}}_{k-1})$ of the state vector is realized by the following equations

$$\begin{aligned}
\bar{X}_k &= \hat{X}_{k-1} + \hat{v} \cdot dt \cdot \sin(\hat{\varphi} + \Delta\hat{\varphi}) \\
\bar{Y}_k &= \hat{Y}_{k-1} + \hat{v} \cdot dt \cdot \cos(\hat{\varphi} + \Delta\hat{\varphi}) \\
\bar{\varphi}_k &= \hat{\varphi}_{k-1} + \Delta\hat{\varphi}_{k-1} \\
\bar{v}_k &= \hat{v}_{k-1} \\
\Delta\bar{\varphi}_k &= \Delta\hat{\varphi}_{k-1}
\end{aligned} \tag{16}$$

2.3.2 Circle drive

The geometry of the vehicle covered by a trajectory is approximated as a sequence of circular arcs (see **Figure 146**, right). This model is closer to reality than the straight line drive, but it is more difficult to calculate (Eichhorn, 2005). So the prediction $\bar{\boldsymbol{x}}_k = f(\hat{\boldsymbol{x}}_{k-1})$ of the state vector is realized by the following equations

$$\begin{aligned}
\bar{X}_k &= \hat{X}_{k-1} + \Delta L \cdot \sin(\hat{\varphi}) + \Delta Q \cdot \cos(\hat{\varphi}) \\
\bar{Y}_k &= \hat{Y}_{k-1} + \Delta L \cdot \cos(\hat{\varphi}) - \Delta Q \cdot \sin(\hat{\varphi}) \\
\bar{\varphi}_k &= \hat{\varphi}_{k-1} + \Delta\hat{\varphi}_{k-1} \\
\bar{v}_k &= \hat{v}_{k-1} \\
\Delta\bar{\varphi}_k &= \Delta\hat{\varphi}_{k-1}
\end{aligned} \tag{17}$$

with

$$\begin{aligned}
\Delta L &= R * \sin(\Delta\hat{\varphi}) \\
\Delta Q &= R - R * \cos(\Delta\hat{\varphi})
\end{aligned}$$

2.3.3 Difference between circle and straight line

The circle drive can be approximated by straight line if the angle $\Delta\varphi$ is small enough and the radius R exceeds a certain threshold. In normal travel behaviour and time intervals of *Δt=1s (1 Hz)* and *Δt=0.2s (5Hz),* the following positional difference between circle and straight line may occur.

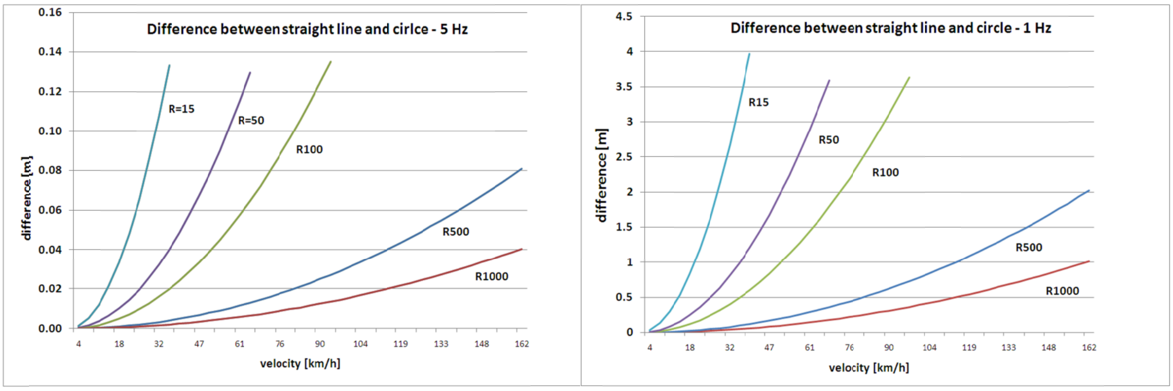


Figure 147: Difference between straight line and circle model with 5Hz (left) and 1Hz (right) data rate

As visible in **Figure 147** for a data rate of 1 Hz there are large differences between the two moving models. At a curvature radius of *15m* and a velocity of about *40 km/h* the difference amounts *4 m*. The smaller the velocity and the larger the radius, the smaller the difference is getting between the two models. If the data rate is increased to for example *5 Hz,* the chart is changed significantly. Now the difference for *R=15* and *v=40km/h* is shrinked to *0.16 m*. So depending on the constraints (data rate, curvature radius and velocity) it may occur that the difference between the two models is negligible.
If the vehicle is moving on an exact straight line, the circle drive model can not be applied! In this case the radius R is infinity and the equations (17) are not solvable!

3 APPLICATION

The Extended Kalman Filter with the two moving models was tested in a real scenario. For data acquisition, the measurement vehicle of the Institute of Engineering Geodesy (see **Figure 148**) with chosen sensors (**Table 22**) is utilized for several test runs.

Figure 148: Measurement vehicle of IIGS with sensors

As input values for the Kalman Filter the measurement data of the GPS receiver, odometer, accelerometer and the gyroscope is used. PDGPS is used as reference for the estimated trajectory resulting from the Kalman filter. The data acquisition was realised with a commercially available computer and the Software LabVIEW[4]. Despite the maximum available frequency of some sensors in **Table 22**, the system works only up to 9 Hz due to the computing power and the operating system (Windows 32 bit). The GPS data from the µblox sensor can only be acquired with 1 Hz. Here the missing values have to be interpolated.

[4] A platform and development environment for a visual programming language from National Instruments.

After the data acquisition the raw data has to be pre-processed. The data has to be synchronised and put in the right format for the Extended Kalman Filter algorithm. The synchronisation is done by the system clock of the computer, of which accuracy is empirically determined to a value of 1 – 5 ms.

Table 22: Sensors

Sensor	**Model**	**Output**	**Frequency**	**σ**
PDGPS	Leica Smartrover1200 (with SAPOS)	*X,Y*	*10 Hz*	*5 cm*
GPS	µblox ANTARIS 4 (AEK4T)	*X,Y*	*1 Hz*	*2 m*
Odometer	Kübler 580X	*Δs*	*50 Hz*	*2 cm*
Accelerometer	MTi XSENS	*a*	*50 Hz*	*0.05 m/s²*
Gyro	Silicon Sensing CRS05	*Δφ*	*50 Hz*	*0.2 °*

Using these sensors the measurement vector **z** can be defined as:

$$\boldsymbol{z} = [X, Y, \Delta\varphi, \Delta s, a] \tag{18}$$

with the associated covariance matrix

$$\boldsymbol{Q_{ll}} = \begin{pmatrix} \sigma_X & 0 & 0 & 0 & 0 \\ 0 & \sigma_Y & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\Delta\varphi} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\Delta s} & 0 \\ 0 & 0 & 0 & 0 & \sigma_a \end{pmatrix} \tag{19}$$

The function $h(\boldsymbol{x}_{k-1}, \boldsymbol{v}_{k-1})$ only depends on $\boldsymbol{x}$ and is linear. So, $h(\boldsymbol{x}_{k-1})$ maps the state space into the observation space. Because of the linearity we can also use the matrix H which is shown in equation (20).

$$\boldsymbol{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & \Delta t^{-1} & 0 \end{pmatrix} \tag{20}$$

The numerical values of $\boldsymbol{Q_{ll}}$ are taken from the accuracy values of **Table 22**. There are no input control terms ($\boldsymbol{B}$, $\boldsymbol{u}$, $\boldsymbol{Q_{uu}}$) defined, but disturbance quantities which are expressed by the matrix $\boldsymbol{C}$ and $\boldsymbol{Q_{ww}}$. The matrix $\boldsymbol{Q_{ww}}$ contains only the disturbance acceleration a_w, so the matrix **C** has the dimension *5x1*, which can be expressed by

$$\boldsymbol{C} = \frac{\partial f(x_{k-1})}{\partial a_{k-1}} = \begin{bmatrix} \frac{\partial X}{\partial a} \\ \frac{\partial Y}{\partial a} \\ 0 \\ \frac{\partial v}{\partial a} \\ 0 \end{bmatrix} \tag{21}$$

On the base of this information the Kalman filter algorithm can be applied. Details see Gao (2011). Test runs were carried out in form of a straight drive, a circular drive and an eight drive (see **Figure 149**).

4 RESULTS

4.1 Numerical results

The result of the evaluation is a trajectory, which consists of a sequence of points. The accuracy of the trajectory is determined by taking the difference of each point resulting from Extended Kalman Filter (X_{EKF}, Y_{EKF}) to the determined reference trajectory (X_{PDGPS}, Y_{PDGPS}) and calculating the RMS (see equation (22).

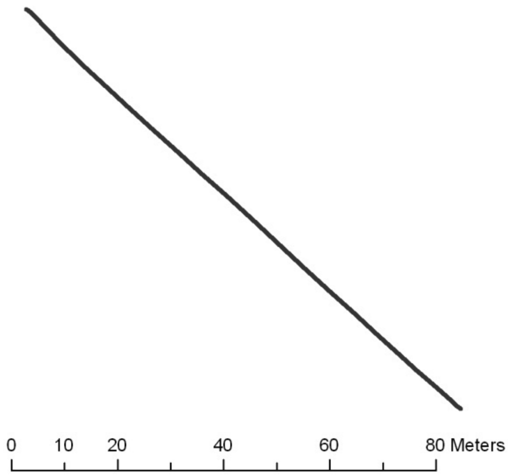
$$\sigma_t = \sqrt{\frac{\sum(X_{PDGPS}-X_{EKF})^2+(Y_{PDGPS}-Y_{EKF})^2}{n}} \qquad (22)$$

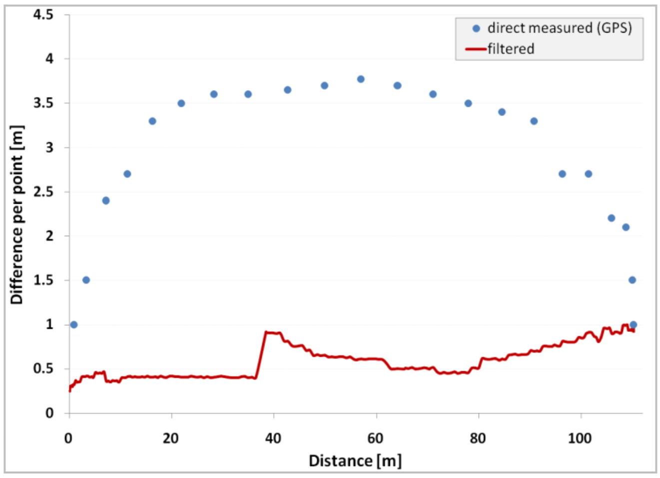
In **Table 23** the results of the three test drives are shown. The disturbance quantity a_w is chosen between 0.5 and $5\ m/s^2$

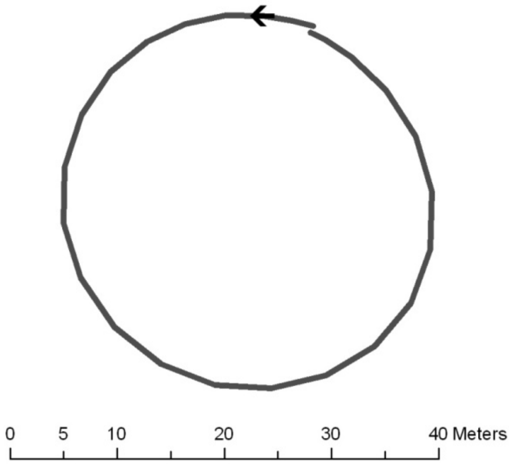
Table 23: RMS for different test drives and different models

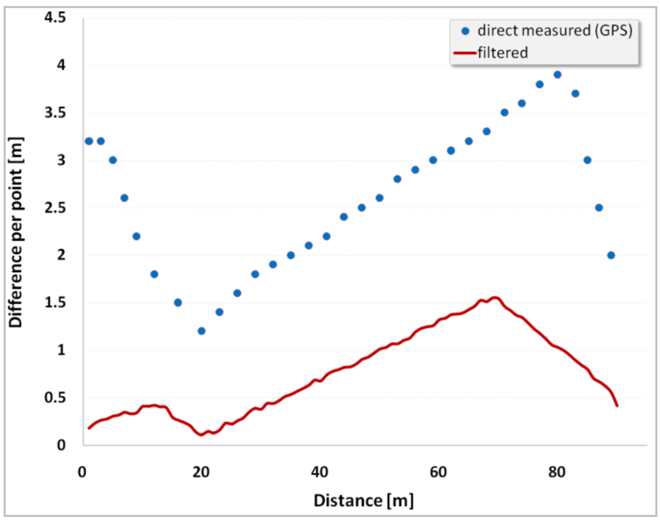
Test drive	σ_t (straight)	σ_t (circle)
Test drive 1 (straight)	0.57 m	-
Test drive 2 (circle)	0.61 m	0.65 m
Test drive 3 (eight)	1.53 m	1.32 m

The test drive 1 is only evaluated with the straight line model. The circle model cannot be applied, because the angle $\Delta\varphi$ is zero (see chapter 2.3). The test drives 2 and 3 are evaluated with both models. At test drive 3, the modularity of the software is obviously visible. For segments, where the vehicle is driving straight forward, the straight line model is applied and segments where the vehicle is driving a circle, the circle model is applied. Changing between the two models is realised during the program runtime.









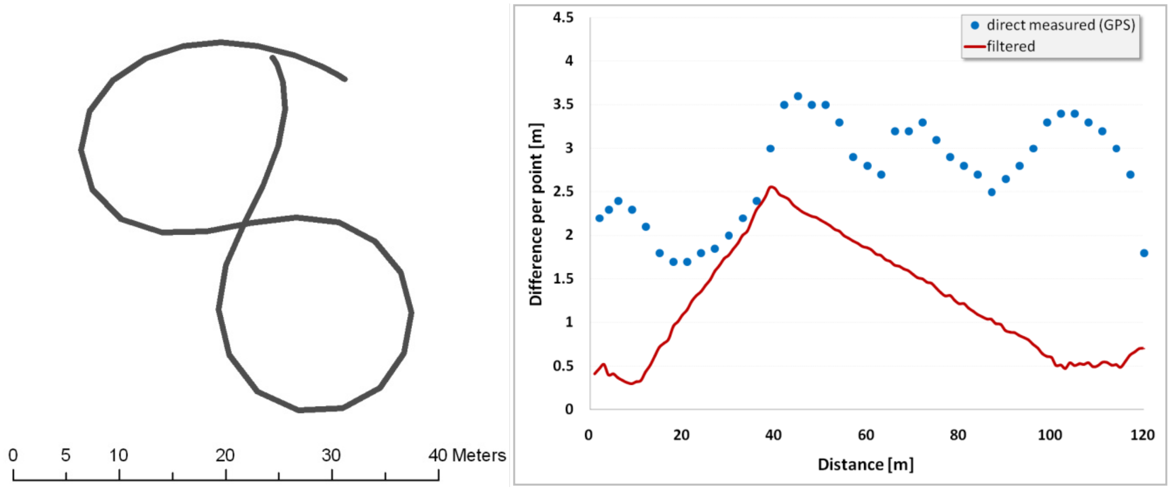


Figure 149: Results of the a) straight line b) circle and c) eight drive

In **Figure 149** the results of the three test drives are presented. The real route, shown on the left side denotes the trajectory resulting from PDGPS. On the right side of each trajectory in **Figure 149** the difference between the PDGPS (nominal value) and the GPS (dots) and the filtered route (line) is plotted. At this point, the improvement that is achieved through the filter can be clearly seen. In **Figure 149a** the maximum improvement starting from GPS is about 3m. The irregularity of the filtered route according to the GPS route (for example at 38m **Figure 149a**) comes from measurement errors of the inertial sensors.

4.2 Evaluation

The RMS of the test drive trajectories reached 50 - 150 cm. These constraints ($R \approx 15m$ $dt=0.011$, $v_{max} \approx 35kmh$) has no practical relevance to differentiate between the two moving models. The maximum positional difference (referring to chapter 2.3.3) between the circle and the straight line drive is about 5 cm. The difference in accuracy between the two moving models does not result from the model, but the inaccuracy of the measurements. So in this case, the straight line algorithm for the entire test drives is preferable, because the computational time is shorter than the circle drive algorithm.

If the constraints (R, dt, v) are changing, so that the speed is increasing and the time interval exceeds 1s or the sensor accuracy is increasing (e.g. change from DGPS to PDGPS), it would make sense to differentiate between these two models.

Additionally one has to mention, that the difference between these two models (shown in chapter 2.3.3) are affecting primarily the prediction of the state vector. So the difference caused by the models has no full impact on the results, because of the filtering (the prediction and the measurements are weighted averaged).

To show the impact of the different models on the results (trajectory), simulated measurement data is created on the base of the circle model. Each measurement epoch contains the measurement data ($X, Y, \Delta\varphi, a, \Delta s$), so that the predicted state vector (see equations (17)) is identical to the updated state vector from measurement data $(\boldsymbol{z}_k = \boldsymbol{H}\overline{\boldsymbol{x}}_k) \rightarrow \ \hat{\boldsymbol{x}}_k = \overline{\boldsymbol{x}}_k$ (compare equation (8)). So, the difference between the nominal trajectory and the results resulting from the circle model has to be zero. Now the difference between the straight-line and the circle model can be easily shown.

With the constraints ($R=15m$ $dt=0.011$, $v=35kmh$) the Kalman Filter algorithm is processed. The difference between the nominal trajectory and the results from the straight line model can be seen in **Figure 150**. The RMS calculated with respect to equation (22) is 1.7 cm. So, in this case the geometric difference between the two models, which is about 5 cm (according to **Figure 147**), is reduced to 1.7 cm by the filter.

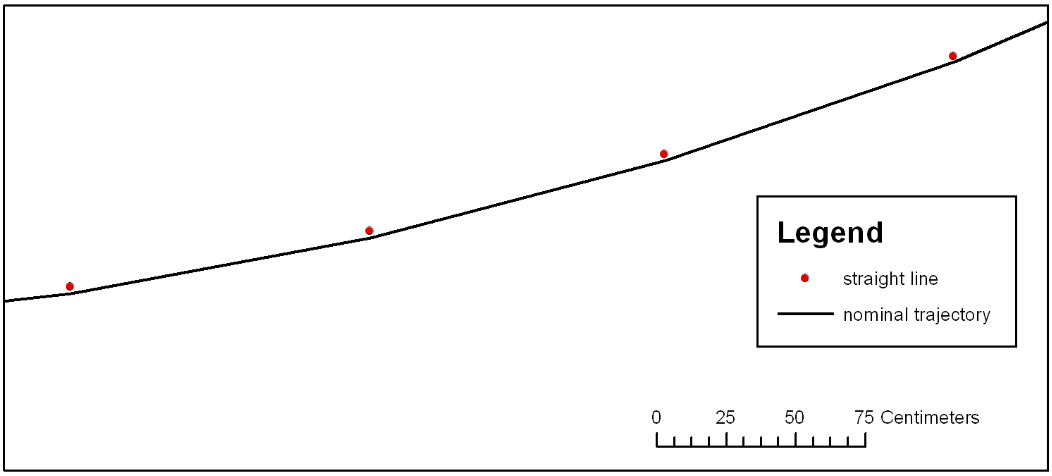


Figure 150: Difference between straight line and circle model after filtering

4.3 Future Improvements

In addition to the geometric model error, there is another fact to be investigated in the prediction phase of the EKF. The prediction of the velocity $\bar{v}_k$ is always the velocity $\hat{v}_{k-1}$ of the last epoch (equations (16) and (17)). In usual traffic flow in inner cities there are a lot of abrupt breaks and accelerations. In this case v is changing very fast in dependence of the accelerations a. In the case of acceleration the predicted point is always behind the point resulting from measurement or in case of a break the predicted point is always in front of the point resulting from measurement, with respect to the direction of travel. To get more dynamics in the filter algorithm, the state vector has to be extended by the acceleration a (see equation (23)). The prediction of v would change to $\bar{v}_k = \hat{v}_{k-1} + \hat{a}_{k-1} \cdot dt$.

$$\boldsymbol{x}^* = [X, Y, \varphi, a, v, \Delta\varphi] \tag{23}$$

Another improvement would be to add an additional disturbance quantity to the disturbance matrix Q_{ww}. An angular acceleration $\ddot{\varphi}_w$ would model also the disturbances in the orientation φ and the orientation change $\Delta\varphi$ in addition to the disturbance acceleration a_w. In general terms the disturbance matrix $\boldsymbol{Q}_{ww}$ and the matrix $\boldsymbol{C}$ are taken into account related to the covariance matrix $\boldsymbol{Q}_{xx}$. With the new disturbance value $\ddot{\varphi}_w$ also a disturbance orthogonal to the direction of travel is possible.

The matrix $\boldsymbol{C}$ has to be extended to the matrix $\boldsymbol{C^*}$ and the matrix $\boldsymbol{Q}$ to $\boldsymbol{Q^*}$ (see equation (24)).

$$\boldsymbol{C}^* = \frac{\partial \boldsymbol{f}(\boldsymbol{x}_{k-1})}{\partial \boldsymbol{f}(\boldsymbol{w})} = \begin{bmatrix} \frac{\partial X}{\partial a} & \frac{\partial X}{\partial \ddot{\varphi}} \\ \frac{\partial Y}{\partial a} & \frac{\partial Y}{\partial \ddot{\varphi}} \\ 0 & \frac{\partial \varphi}{\partial \ddot{\varphi}} \\ \frac{\partial v}{\partial a} & 0 \\ 0 & \frac{\partial \Delta\varphi}{\partial \ddot{\varphi}} \end{bmatrix} \quad \boldsymbol{Q}^* = \begin{bmatrix} a_w & 0 \\ 0 & \ddot{\varphi}_w \end{bmatrix} \tag{24}$$

For this, some of the equations quantities like $\boldsymbol{F}$, $\boldsymbol{B}$, $\boldsymbol{H}$ have to be recomputed in dependence to the new state vector and the disturbance matrix. This is an aim for future work to improve the theoretical background of the filter.With respect to the modularity, more moving models and more sensor combinations should be implemented.

5 CONCLUSIONS

The paper describes the basics of Kalman Filter with respect to positioning with terrestrial multisensor systems. Here, two different motion models (straight line and circle drive) for state prediction in the Kalman filter are presented. This filter is tested on real measurement data, acquired with the IIGS measurement vehicle. A GPS sensor, an accelerometer, a gyroscope and an odometer are used as input measurements for the Extended Kalman filter. The RMS of the filtered trajectories of these test drives are between 50 and 150 cm. In this accuracy level and with the given constraints for data acquisition, the distinction between the two motion models is not required. Only with a lower frequency and/or a higher velocity the distinction between the straight line and circle line will make sense. PDGPS with an accuracy of about 5 cm can also be used, because here an improvement in accuracy about the factor of 10 can be achieved.

ACKNOWLEDGEMENTS

The author would like to thank Ms Yang Gao for the support and the deployment of the measurement data, which are acquired as a part of her master thesis!

REFERENCES

Journal articles:
GAO, Y.: *Positionsbestimmung für verschiedene Bewegungsmodelle mittels Kalman Filter.* Master Thesis at the Institute of Engineering Geodesy, University of Stuttgart, 2011, *unpublished.*

EICHHORN, A.: *Ein Beitrag zur Identifikation von dynamischen Strukturmodellen mit Methoden der adaptiven KALMAN-Filterung*, Verlag der Bayerischen Akademie der Wissenschaften in Kommission beim Verlag C. H. Beck, München, 2005.

GELB, A.: *Applied Optimal Estimation*. The M.I.T. Press, Massachusetts Institute of Technology. Cambridge, Massachusetts and London, England, 1974.

KALMAN, R. E.: *A New Approach to Linear Filtering and Prediction Problems*. Transaction of the ASME – Journal of Basic Engineering, p. 33-45, 1960.

MAYBECK, P. S.: *Stochastic Models, Estimation and Control*, Academic Press Inc., Volume 1, 1979.

RAMM, K.: *Evaluation von Filter-Ansätzen für die Positionsschätzung von Fahrzeugen mit den Werkzeugen der Sensitivitätsanalyse.* Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften / C ; 619, München, 2008.

SCHWIEGER, V., SCHOLLMEYER, R., RAMM, K.: *Integrated Positioning by MOPSY.* Proceedings on 5th European Congress and Exhibition on Intelligent Transport Systems and Services, Hannover, 01.-03.06. 2005.

STERNBERG, H.: *Zur Bestimmung der Trajektorie von Landfahrzeugen mit einem hybriden Meßsystem.* (Dissertation) Schriftenreihe des Studiengangs Vermessungswesen der Universität der Bundeswehr München, Heft 67, 156 S., Neubiberg, 2000.

WELCH, G., BISHOP, G.: *An Introduction to the Kalman Filter*. SIGGRAPH 2001, Course 8, 2001.

This paper has been accepted as a „peer reviewed paper“ by scientific committee

Low-cost Platform for Landmark-based Positioning and Navigation

Sabine Hofmann, Daniel Eggert, Colin Kuntzsch, Monika Sester
Institute of Cartography and Geoinformatics, Leibniz University Hannover, Germany

Abstract
This paper presents a platform based on low-cost components for landmark-based navigation intended for research and teaching purposes. The proposed platform includes a LEGO Mindstorms kit, an Android-based Smartphone as well as a compact laser scanner Hokuyo URG-04LX.

The LEGO Mindstorms kit builds the robot's chassis and provides an interface to the servo motors for movement. As a range sensor, the compact laser scanner gathers range information about the robot's environment. Finally, the Smartphone represents the central data processing and control unit bringing together sensors and actuators. The data processing involves landmark detection from the range data provided by the laser scanner and map matching of the identified landmarks. The required landmark map data structure is not based on single landmarks, but on unique constellations formed by the landmarks.

Keywords
Landmark map, Positioning, Laser scanning, Navigation, Robotics

1 INTRODUCTION

Driving autonomously requires highly accurate and reliable positioning. In order to meet the requirements of different applications, varying solutions might be suitable. In most outdoor applications, global satellite navigation systems are used. In indoor scenarios or where high positioning accuracies are required other solutions are needed. Therefore, alternative positioning systems to GNSS are required, especially to increase the accuracy and to have a complementary data source in areas where GNSS is not available. A possible alternative is the use of landmark maps with natural or artificial landmarks.

The platform for positioning and navigation described in this paper was specifically designed for research and teaching purposes. The idea is to simulate a real-world scenario, such as autonomous driving in an urban scene, on an easy-to-manage scale. Therefore, the used landmark map is designed for an indoor scenario where both spatial extent and number of features are limited.

The paper is structured as follows: In Chapter 3 the map generation process as well as feature detection, localization, routing and navigation are described. Details on the required hardware components for this approach and our prototype are provided in Chapter 4. This is followed in Chapter 5 by an overview of practical problems encountered during design and implementation of the prototype.

2 RELATED WORK

There are currently many map-based approaches such as occupancy grid maps as well as symbolic representations such as line maps or landmark-based maps (Burgard & Hebert 2008). These maps can be generated by the robot itself while driving, using simultaneous localization and mapping (SLAM). One important advantage is that no map needs to be produced in advance. But if landmarks should be used for positioning, appropriate landmarks have to be extracted and proofed, before being saved as a map object (Thrun et al. 2006). Consequently, the accuracy of the produced map directly influences the positioning accuracy of the robot.

Another solution is to provide map data to the robot, which may contain only useable landmarks for positioning tasks. These maps have to be produced in advance, using for example highly accurate laser scanners or cameras, total stations, CAD plans (indoor), cadastral maps or aerial images (outdoor). Using existing map information has the advantage to provide scene interpretation and to supply knowledge of areas which appear outside the robot's field of view, which can be important for path planning. For driver assistance systems, the use of navigational maps was investigated by Blervaque (2008).

Using pole patterns for localization is currently investigated for localization in large outdoor scenarios. Weiss et al. (2005) combined GNSS, odometry and a laser scanner to estimate a vehicle's position. A good overview of point matching algorithms is given by Chui and Rangarajan (2003). Brenner (2009a) describes the use of highly accurate full 3D laser scans and the localization based on local pole patterns which are used as descriptors within the map. Using pole patterns as map and an automotive laser scanner for localization gives promising results (Brenner 2009b).

3 THEORY

In the following sections all steps from map production, feature detection, localization, route planning and finally navigation to a target position are described in detail.

3.1 Landmark Map Generation

In general, for various purposes different types of maps are designed and used, e.g. maps for autonomous vehicles and maps for pedestrian navigation have to meet different requirements with regard to accuracy or feature content. When designing a map for robot indoor navigation one needs to take into account several requirements, such as the choice of objects depending on the used sensors or the uniqueness of feature patterns (e.g. triangles).

For our prototype, we decided on using pole-like objects in our landmark map. In an outdoor environment such pole objects could be tree trunks, poles of traffic lights, sign posts or lamp posts. Using poles (i.e. upright structures with homogenous shape, diameter and position at all heights) as landmarks has many advantages: they can be naturally found in almost all environments and can be described with simple geometric models detectable by different types of sensors. Full 3D laser scanners may show the curvature on the surface of the poles, depending on the angular resolution and point density which can be used to distinguish between poles and other objects. In lower resolution the linear upright structure can be found by using a geometric model or depth-jumps, which appear on both sides of the pole. When scanning only in a few horizontal planes, the pole structure will lead to a stack of horizontal slices. In a single horizontal plane (**Figure 151a**), where the height of the scanners on different vehicles could vary, the same poles can be detected at different heights. Therefore, their extraction from a laser scan-profile is straightforward (see Chapter 3.2). When using other types of sensors, such as radar or cameras, other features should be considered.

Another advantage of using poles is that their 3D structure can be easily encoded into a 2D map, using a representative horizontal section. This way, we represent pole-like landmarks using only 2D center coordinates and landmark radius (**Figure 151b**). Multiple landmark center points can form simple patterns, e.g. geometric shapes like triangles. Self-positioning is now possible through map-matching a number of measured triangles with the triangles implicitly stored within the landmark database (see **Figure 151b**). As matching single triangles may lead to ambiguities among several similar-shaped triangles, we consider all visible triangles, looking for unique landmark constellations (see Chapter 3.3). In order to increase the accuracy of our self-positing approach, a certain minimum density of known landmarks is required, which depends on the used laser scanner's range.

3.2 Landmark Detection

In order to meet the requirements of a low-cost platform, we use a compact 2D laser scanner, scanning in one horizontal plane (**Figure 151a**). Therefore, the landmark detection algorithm has to be adapted to the used sensor. In our approach, poles are extracted from depth-jumps within the laser scan profile. With a range sensor, depth-jumps will be detected if the measured range gap at two consecutive measurements is higher than a defined threshold (**Figure 151c**). For poles, depth-jumps can be found on either side of the scanned object. In order to reduce the influence of outliers, the radius of pole objects, which is known from the map, is used. Therefore, objects which appear to be smaller or wider in radius than the known landmarks are not used for localization.

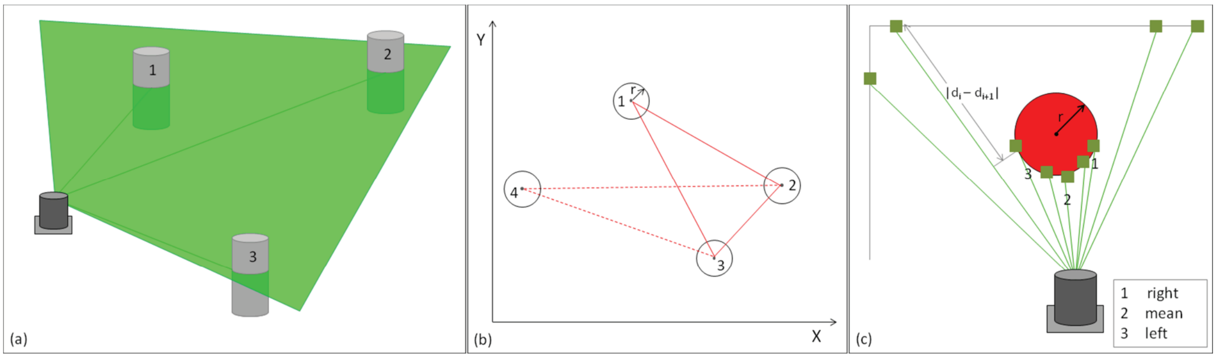


Figure 151: Laser scanner scanning in one plane (a), 2D point pattern from 3D poles (b), Pole detection based on depth-jumps (c).

As only the surface of the objects is scanned, the distance to the centre of any pole has to be computed based on the measurements and the knowledge of the pole's diameter, which is provided by the map. As directional measurement, the mean laser beam in-between the rightmost and leftmost laser beams of a detected pole's surface is used. As distance to the pole's centre, the radius of the pole is added to the range measurement at the mean laser beam (**Figure 151c**).

3.3 Landmark-based Positioning

The positioning process is separated into two parts. Firstly, corresponding poles between all extracted objects and reference landmarks have to be found. Therefore a map matching approach is used. Secondly, a least squares adjustment is performed to estimate the laser scanner's position (which can then be used to determine the robot's position from the fixed scanner location on top of the mobile robot platform).

In our map matching approach, we do not use single landmarks, but unique constellations formed by the landmarks. The correlation step is done by finding corresponding triangles in both the landmark map and the pole configuration from the current laser scan profile. Additional detected poles are then used for verification.

The coordinates of all extracted poles are computed in a local robot coordinate system. With these coordinates, the side lengths of all possible triangles formed by the extracted poles are computed. The map contains the reference poles and thus implicitly the reference triangles. In order to find corresponding triangles, the side lengths of local and reference triangles are compared pair-wise. For comparison of possible triangle-matches between triangles t_{map} within the landmark map and t_{scan} within the current laser scan profile, we define a similarity score as

$$score(t_{map}, t_{scan}) = \left(\bar{a}(t_{map}) - \bar{a}(t_{scan})\right)^2 + \left(\bar{b}(t_{map}) - \bar{b}(t_{scan})\right)^2 + \left(\bar{c}(t_{map}) - \bar{c}(t_{scan})\right)^2$$

where $\overline{a}$ denotes the length of the shortest edge, $\overline{b}$ the length of the longest edge and $\overline{c}$ the length of the remaining edge within the respective triangles.
Depending on the number of visible poles, three cases have to be considered:

1) Three poles: Similarity scores for all reference triangles and the single triangle formed by the visible poles are computed. The reference poles with the lowest squared position deviation are used as corresponding poles.
2) More than three poles: A single local triangle is chosen. The reference triangle with the best similarity score is used to compute an estimate for the robot's position. Based on this preliminary estimation, all additional visible poles are matched with the landmark map. The redundancy of additional poles allows for identification of objects not contained within the landmark map, and vice versa of landmarks known from the landmark map, that are missing within the current observation. Selection of the local triangle and matching all visible poles to the landmark map is repeated for all possible triangles. The solution with the maximum number of successfully matched poles is then accepted.
3) Less than three poles: No map matching is possible, i.e. no position can be estimated from less than three visible pole objects. This leads to requirements in regards to the density of the landmark map.

After the map matching process is completed successfully, the second part, position estimation, is done by resection using the coordinates of the map features and the measurements of the laser scanner (distance and direction) to all visible successfully matched poles. As a result of the adjustment one receives the robot's current position and orientation (heading) as well as belonging accuracies.

The error equations

$$d_i + v_{di} = \sqrt{(X_i - X_R)^2 + (Y_i - Y_R)^2}$$
$$\varphi_i + v_{\varphi i} = \tan^{-1}\left(\frac{Y_i - Y_R}{X_i - X_R}\right) - \gamma_0$$

where d_i, ϕ_i ... distance and direction measurement to pole i,
X_i, Y_i ... coordinates of reference pole i,
X_R, Y_R ... coordinates of robot position,
γ_0 ... heading,

result in the design matrix, which contains all geometric information.

In the least squares adjustment, the positional error of the reference map features must be considered, as the reference positions are not error-free. Therefore, for any visible pole the number of rows of the design matrix increases by four. The stochastic information of the system is considered in a weight matrix, were the measurements are assumed to be uncorrelated.

3.4 Routing

The Routing is based on the Dijkstra Shortest Path (short: DSP) algorithm (Dijkstra 1959). Since DSP determines the shortest path between two nodes within a graph, the landmark map described above has to be converted into such a graph. This conversion (**Figure 152**) includes the following steps:

1) Landmark Map Rasterization (with fixed cell size)
2) Tagging the cells (accessible or not) based on known obstacles (in our case landmarks only)
3) Build the graph based on accessible cells
    a. where the center of each cell represents a node in the graph and
    b. each node is connected by an edge to each accessible neighbor cell (8-neighborhood)

Furthermore, DSP depends not only on a simple graph, but on a graph with weighted edges. The common edge weight reflects the length of the edge, but also other aspects may impact the weight.

Since we are going to use the determined path for autonomous navigation purposes, we also include the riskiness of an edge into its weight. A risky edge is an edge right next to cells tagged as not accessible. That leads to longer but less "dangerous" paths in terms of possible collisions with known obstacles as shown in **Figure 153**.

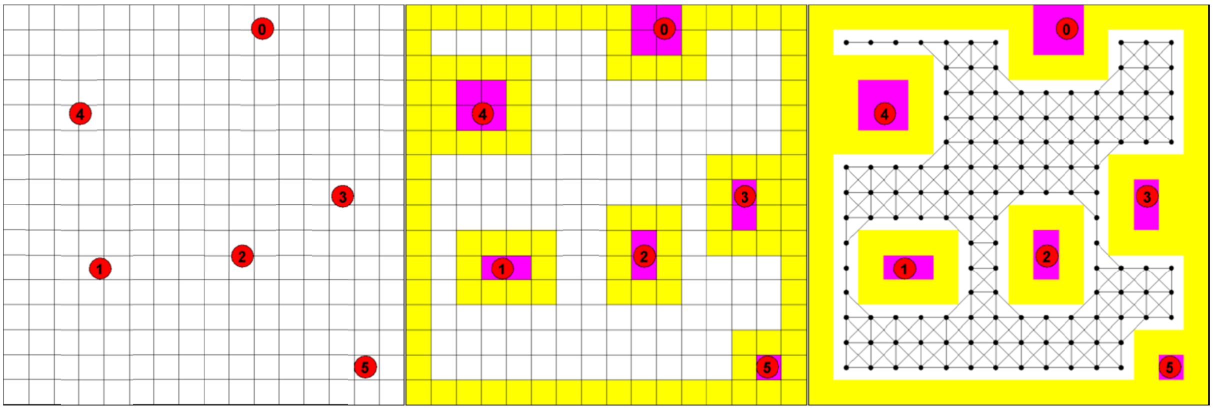


Figure 152: Conversion from Landmark Map to a Graph used by Dijkstra Shortest Path.

Depending on what kind of navigation is used for determination of the route, short route segments may raise problems due to positioning imprecision. In order to address those problems the determined path is altered. Based on the DSP route and the knowledge of what cell is accessible, the route is generalized with the goal of longer route segments. The simplest way of getting longer route segments is to join consecutive segments with no direction changes. This results in long segments for the majority of the route segments. Nonetheless, paths with many direction changes still contain very short segments. Therefore, we simplify the path in another fashion. As mentioned the route is altered not only in respect to the given graph presentation of the landmark map but also with the knowledge of the underlying tagged cell structure. In regard to that, two nodes are directly connected, if the connecting segment does not intersect with cells tagged as not accessible. This results in a virtual line-of-sight route with a minimum of short route segments. As shown in **Figure 153**, the described simplification results in only two long route segments. By comparison, the original DSP route consists of 30 segments, while the simplification of joining consecutive segments with no direction changes still results in eight segments.

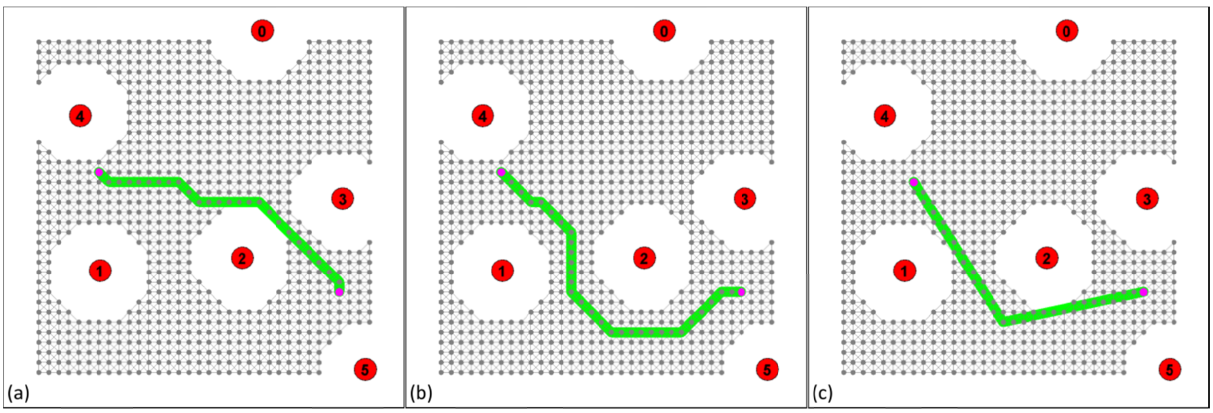


Figure 153: Shortest path in respect to distance (a), riskiness (b) and long route segments (c).

3.5 Navigation

Our Navigation solution is mainly based on navigating along each straight route segment from the robot's current position to the desired segment destination. This task is divided into two steps. The first step is to turn towards the current destination, so the target position is right in front of the robot. The second step is to drive along the straight path until the target position is reached.

4 HARDWARE PLATFORM

In this section we provide details on the physical components of our prototype from the conceptual perspective, defining requirements for the useable hardware components. We then continue to briefly describe specific hardware choices for selected system components that introduce additional capabilities and constraints influencing the performance of the system.

4.1 System overview

Our self-positioning approach requires a moving vehicle carrying a 2D laser scanner and a processing unit dealing with laser scanner result interpretation, matching of detected landmarks with known landmark constellations, self-positioning and routing. The laser scanner needs to be capable of measuring along the horizontal axis parallel to the floor plane. Additional requirement is a database of known detectable objects within this plane whose position in a global geo reference system is known beforehand. This database must be accessible by the vehicle's processing unit (e.g. stored within the processing unit or accessible from a remote data store).

In order for the self-positioning approach to work, it is necessary that at least three objects are detectable by the laser scanner at all times. Larger numbers of visible objects allow for redundancy within position estimation and thus for detection of unknown objects and measurement error reduction. This requires the object database to provide a sufficiently high density of objects in relation to the laser scanner's range along the vehicle's route (or vice versa a sufficient scanner range for a given object density).

4.2 Hardware specifics

The developed prototype consists of three parts. First, the chassis-platform basically built from the LEGO Mindstorms NXT 2.0 Set, including the NXT Brick microcontroller and two servo motors. Second a Hokuyo URG-04LX compact laser scanner mounted on top of the chassis, providing range information for the landmark detection. And finally, a Smartphone mounted at the front as central data processing and control unit. For demonstration purposes, the object database is stored within the processing unit, i.e. stored on the Smartphone. It is complete in regards to the objects contained within our experiment setup.

4.3 Communication interfaces between system components

As the prototype has been developed for teaching and research purposes (e.g. demonstration of the self-positioning approach), we created additional input and output interfaces to access the mobile processing unit, e.g. to be able to manually change system parameters on-the-fly or to inspect intermediate system states. This way, additional control and visualization components can be run on external processing units (e.g. students' computers), accessing intermediary results and interfaces for manual control on the mobile processing unit, while not being required for the operability of the self-positioning/navigation task.

A modular architecture allows exchange or modification of the system components, as well as extension with additional modules using the provided interfaces (see **Figure 154**), as long as the predefined (unidirectional) interfaces between pairs of system components are implemented:

1) The interface between the laser scanner and the on-board smart phone translates a low level data stream from the sensor into a suitable representation for further processing (e.g. depth jump detection). Knowledge of the sensor-specific data format and encoding is required for the interpretation of the measurements (exchange of the sensor hardware thus requires re-implementation of the interface). The current prototype implementation for example translates arrays of distance measures, transmitted in the sequence of measurement (increasing angle in

clockwise direction), into 2D positions of measured points relative to the laser scanner position and furthermore provides parameter-controlled algorithms for depth-jump detection.

2) The interface between the Smartphone and the robot motor controller needs to translate high-level navigation commands into low-level actuator-instructions (in our case setting the rotation speeds and directions for two servo motors). For the prototype, we implemented navigation commands for moving forward/backward on a straight line and turning left/right while standing still. This set of commands is sufficient for navigation in our scenario, while being exchangeable with any other robot movement specification.
3) In terms of system architecture, the on-board Smartphone acts as a data server, allowing an arbitrary number of clients to register onto the streams of laser scanner data, depth-jump detection results and routing results.
4) Additionally, we provide an interface for direct access to the Smartphone, to allow direct access of the high-level robot commands (e.g. manual or by means of alternative routing mechanisms).

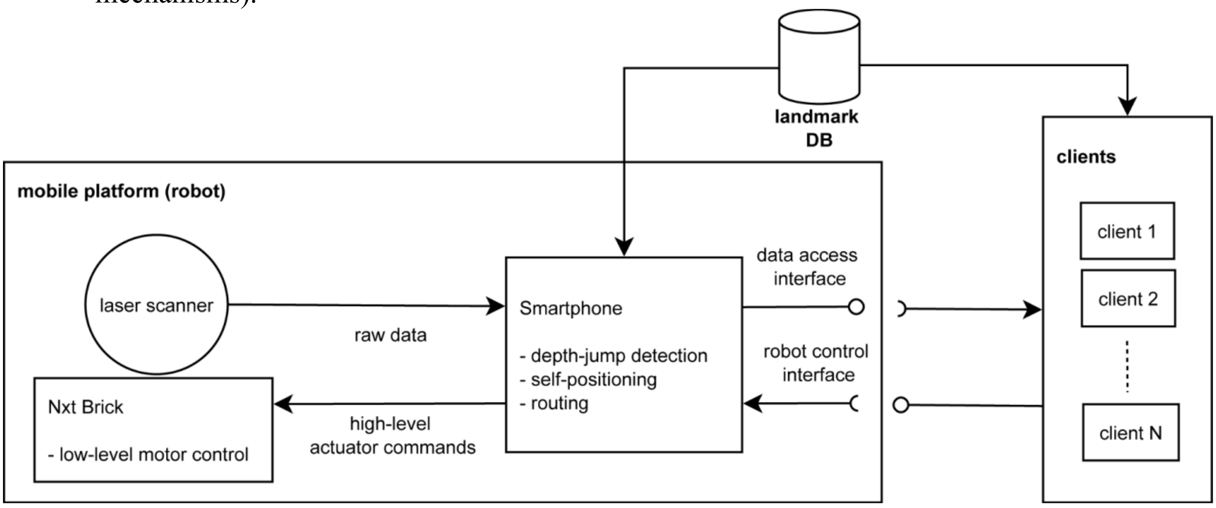


Figure 154: Conceptual view of system components and interfaces.

4.4 Prototype

For demonstration purposes we constructed a wooden square-shaped box with dimensions 2m x 2m. It serves the purposes both to provide a physical boundary for the laser scanner to not detect objects outside the experimental setup as well as a reference frame for definition of a local cartesian coordinate system. Within this local coordinate system we set up a configuration of cylindric obstacles with known positions (both physically placed within our experimental setup as well as stored within our object database). In first experiments we were able to successfully simulate autonomous navigation of the robot within the setup without external positioning assistance or active collision detection.

4.4.1 Visualization client specifics

In order to provide a graphical interface to the system’s internal state for demonstration of the self-positioning approach, we created an example for a possible client application directly accessing the on-board Smartphone in both directions. It allows a live-view on the produced data, i.e. laser scan profiles, detected/matched landmarks, current routing strategy and manual initialization/termination of the navigation task. For better visualization, we installed a static web-cam filming the prototype setup. After manually registering 2D world coordinates to points within the image of the 3D scene, we are then able to directly draw 2D information on top of the web-cam image using a homography transformation (an example for the chosen visualization is shown in **Figure 155**).

4.4.2 Landmark setup specifics

Creation of a suitable landmark-setup for our demonstration prototype proved less trivial than expected, as we needed to consider technical and practical requirements and limitations: depth-jump detection requires a minimum distance between pairs of individual landmarks and a minimum distance of poles from the outer boundary. The self-positioning task needs a sufficient number of landmarks producing unique triangles to allow for redundancy for error detection. On the other hand the setup is constrained by a limited amount of space available for the setup, further complicated by our desire to leave enough space between pairs of individual landmarks for proper navigation. Finally, we developed a semi-automatic setup generator that considers all of those requirements. Landmark data is currently stored locally within the Smartphone and all clients. In future versions of the prototype we are planning to centralize the landmark database in order to realize modular design, with the landmark database being a centralized component.

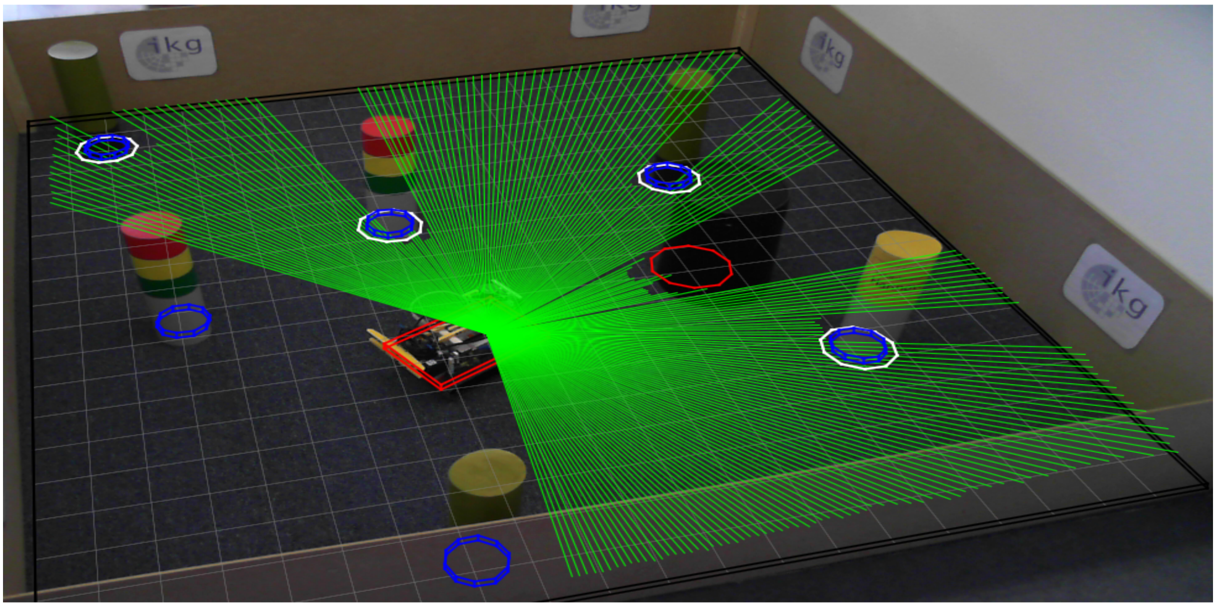


Figure 155: Prototype setup: green = current laser scan measurements, blue = known landmarks, white = successfully matched landmarks detected from laser scan profile, red = depth-jump not corresponding to a known landmark. Screenshot is taken from visualization client implementation.

5 PRACTICAL PROBLEMS AND SOLUTIONS

In this chapter we present some interesting problems encountered during design and implementation and how we faced them.

5.1 Module design

As mentioned above, the described platform was developed for teaching and research purposes, so all modules like landmark detection, landmark based positioning, routing and even navigation are designed to be extensible and exchangeable. The general design of each module is shown in **Figure 156**. All modules work asynchronously to each other, so new available data can be processed by a module while older data can still be used by processing intensive tasks/modules. Each module has to register itself to the module providing the needed input-data, e.g. the positioning module has to register for newly detected landmarks at the landmark detection module. As soon as the landmark detection module finished processing the laser scanner data, it triggers all registered modules (like the positioning module) and provides the most current detected landmarks on demand. The trigger mechanism makes use of binary semaphores, which resume threads from idle state in case new input data is available.

Since the single modules are only coupled over the provided output and needed input data, each module can be easily altered or even exchanged as long as it sticks to the provided output data and the shown module registration and trigger mechanism. So, if a new landmark detection scheme is under investigation, none of the other modules have to be altered, which makes it well suitable for research and teaching purposes.

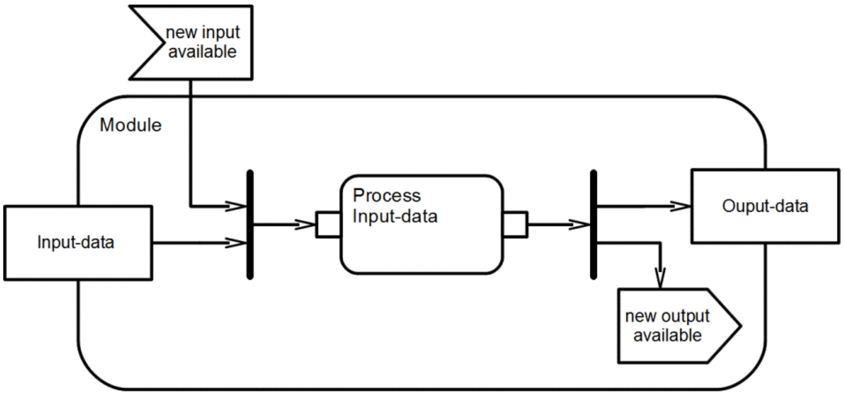


Figure 156: General module design (UML activity diagram).

5.2 Navigation Module Implementation

The navigation implementation navigates the robot based on the given route and the current position along each route segment. As described in Chapter 3.5 this involves two steps, turning to segment destination and driving to segment destination. Since the robot uses only one laser scanner which also has a blind angle of 120 degrees, the navigation module has to deal with the possibility of not being able to determine its current position or heading. In case of the turn step, the robot may not have a heading solution for the destination angle, so it will not stop turning at the desired target heading. In case of the driving step the robot may not have a position solution for its current position because less than three landmarks are visible to the laser scanner. Both problems can be faced by introducing a reverse navigation mode, where the robot simply drives backwards. Another possibility is to mount the laser scanner pivoting, which however requires a more sophisticated cable management. Moreover, a second rear facing laser scanner could be mounted onto the robot, which however conflicts with a low-cost solution.

In addition to the laser scanner's view angle constraints the imprecision of the positioning has to be taken into account in case of navigation. Since it is almost impossible to stop the robot at the exact target heading or at the exact target position some thresholds have to be defined. In our case, we set the angle tolerance threshold to ±10 degrees, meaning that the target heading is considered hit even if the determined heading is 10 degrees below or above the target heading. That heading offset has to be faced in the subsequent driving step. This can be considered as a typical regulation problem. Our solution is to let the robot trail along the route segment by setting different speeds to the driving chains. The second defined threshold is the position tolerance threshold. In our case we set it to 0.1 m, which was derived from the robot's size (0.3 x 0.2 meters) and the cell size used in the routing module (0.05m).

5.3 Real-time Self-positioning

As stated in the previous chapter, the navigation task includes both a turning step as well as a driving step. In the case the NXT servo motors actuators were optimal calibrated both navigation steps might be trivial. It is, however, impossible to perfectly calibrate the servo motors, especially since it relies on the available battery voltage, the used type of drive and of course the current underground. Furthermore, the latency between sending a command and the corresponding actuators reaction leads to small deviations of the actual route from the supposed route.

The scanning frequency of the used Hokuyo laser scanner is 10 Hz, which results in 10 range measurement sets per second. The sparse time-consuming landmark detection and map matching implementations enable us to derive a position solution from each of the provided range measurement sets. Considering the fairly slow speed of the robot, the self-positioning can be considered as a real-time self-positioning solution. This fact alleviates the navigation task, dealing with positioning inaccuracies as well as the shortcomings at the actuator translations described above.

To counter those effects our approach completely omits any actuator calibration and faces this shortcoming with a continuous real-time self-positioning. Thus, with regards to the turning command we let the motors turn until the desired heading is reached. In the same way, driving forward is a continuous process which repeatedly checks, whether the destination has been reached (as opposed to driving forward for a specified distance or time). Finally, the described routing task is not only performed once, but repeatedly whenever a significant deviation from the previous route is detected. This leads to a continuous self-adjustment of the navigation task through continuous comparison of supposed and actual robot position.

6 CONCLUSIONS AND OUTLOOK

This paper proposed a prototype of a low-cost platform for positioning and navigation purposes in an experimental indoor setup. All steps, from map generation to autonomous route planning and navigation were shown. This was followed by a detailed description of the prototype and the software modules. The modular design of the software allows the use of the prototype in teaching and research. Currently we are working on the analysis of achievable accuracies in different map scenarios. Several experiments and practical tests are carried out to demonstrate a stable run. Furthermore, extensions to the modules such as obstacle detection, the detection of missing poles and an incremental update of the map using the robot’s measurements are planned.

REFERENCES

BLERVAQUE, V.: *Prevent Maps and ADAS D12.1 Final Report*, ERTICO – ITS Europe, http://www.prevent-ip.org/download/deliverables/MAPS&ADAS/PR-12100-DEL-080127-v10-ERT-D12.1%20MAPS&ADAS%20Final%20Report.pdf, last accessed on December 14, 2011

BRENNER, C.: *Global Localization of Vehicles using Local Pole Patterns*. DAGM 2009, 31st Annual Symposium of the German Association for Pattern Recognition, LNCS 5748, p. 61-70, 2009a.

BRENNER, C.: *Extraction of Features from Mobile Laser Scanning Data for Future Driver Assistance Systems*, Advances in GIScience: Proceedings of 12th AGILE Conference on GIScience, Lecture Notes in Geoinformation and Cartography, Springer, Berlin, p. 25-42, 2009b.

BURGARD, W., HEBERT, M.: *Springer Handbook of Robotics*. Chapter *World Modeling*, p. 853-869, Springer, 2008.

CHUI, H., RANGARAJAN, A.: *A new point matching algorithm for non-rigid registration*. Computer Vision and Image understanding 89, 2003

DIJKSTRA, E. W.: *A note on two problems in connexion with graphs*, Numerische Mathematik, Volume 1, Number 1, p. 269-271, 1959.

THRUN, S., BURGARD, W., FOX, D.: *Probabilistic Robotics*. The MIT Press, Cambridge, 2005.

WEISS, T., KAEMPCHEN, N., DIETMAYER, K.: *Precise ego localization in urban areas using laserscanner and high accuracy feature maps.* Proc. 2005 IEEE Intelligent Vehicles Symposium, Las Vegas, USA, p. 284-289, 2005.

Educational Robotic Platform "Zero2Nine" for Autonomous Navigation and Tracking based on Imaging Sensor Systems

A. Linz, A. Ruckelshausen
University of Applied Sciences Osnabrück, Germany

Abstract
Robotic platforms are more and more used at engineering educational institutes for teaching various theoretical and practical subjects, such as electronics, computer science, mechatronics, sensor technologies, image processing or artificial intelligence. However, the development of the mechanical platforms as well as the integration of sensors and system technology often limits the experience of combining theoretical concepts with practical experiences. In this paper a commercially available platform is combined with different smart imaging sensor systems and a corresponding system technology. The resulting educational robotic platform "Zero2Nine" allows navigation, tracking and detection applications based on individual imaging devices or sensor fusion algorithms. The name of the robot is related to an application example, the detection of numbers in outdoor environments. The modular system, the robustness and a short getting-started period are important aspects.

As a vehicle the commercially available Fraunhofer VolksBot RT4 platform is applied, which is equipped with 2 Maxon-motors and a motor controller. As imaging sensor systems a color smart camera, a laser scanner and a 3D-Time-of-flight camera are integrated, OpenCV as a computer vision library can be used. The system architecture is based on a PC with Linux as operation system and Ethernet with a standard TCP/IP protocol, thereby implementing a high flexibility with respect to the integration of additional components.

The technology is used for teaching, laboratory experiments and in particular for student projects in the Master and Bachelor modules, such as "Imaging sensor systems", "Sensor systems", "Optoelectronics" and "Seminar Mechatronics". The robot is a helpful facility, especially for "advanced lab", where the students work more intensive on a subject e.g. a smart cam or a laser range finder. It has also been a powerful starting tool for student groups participating robot competitions (example: "International Field Robot Event"). The students are highly motivated and this educational option has a potential with respect to increasing knowledge and practical experience together with team working. Moreover the robot is demonstrated at fairs, examples are "Agritechnica" or "IdeenExpo".

Keywords
Autonomous mobile robotics, Imaging Sensor Systems, University Education

1 INTRODUCTION

Ever more, robots will be present in our daily live. First commercial products are lawn movers or vacuum cleaners (irobot, 2011). Service robots like PR2 (Meeussen et al. 2010) or field robots like BoniRob (Ruckelshausen et al. 2009) are in a development state, but not too far away from a product. At universities robotic platforms are introduced for teaching interdisciplinary aspects of mechanics, electronic and computer sciences. There are a lot of robot platforms with its advantages and disadvantages on the market available. Low cost robots like Lego Mindstorm (Lew et al. 2010) or Asuro (asurowiki, 2011) from the DLR can be afford from the students to get first experience. They are also used in robot competitions for children and pupils. A robot build for outdoor use like optoMaizer (Klose et. al 2006) is more expensive, unique and has fixed soft- and hardware on board. However, VolksBot is an expensive but high expandable and robust platform. That is the reason why the University of Applied Sciences Osnabrück purchased 4 of them for different labs in order to get a technically interdisciplinary approach to autonomous robots. The labs have their focus on Embedded System Engineering, Computer Engineering, Mechatronic System Engineering and Sensor Systems. The focus of this paper is on a Volksbot equipped with optoelectronic sensor systems.

2 CONCEPT

The concept of the mobile robot is to keep the system as modular and flexible as possible. This concerns software, hardware and mechanical construction likewise. Flexibility is given by Ethernet (TCP/IP) as bus system. Sensors, intelligent sensor systems, actuators, controller boards and PCs are interacting on this bus. It is possible to connect to each of this part from “outside” via the WLAN-Bridge or an Ethernet cable. Sensors and actuators can be plugged or unplugged without an influence on the whole system. An example for high flexibility is, that the software can run somewhere in a virtual machine (i.e. Virtual Box with Linux), which connects via LAN or WLAN to all sensors and actuators and controls the robot. The same software can be run on a microcontroller board or a PC, which is physically installed on the robot.

The software on the robot starts a server, which exchanges most data in plain ASCII code. If the connection to the server is established, the client will be able to control the robot or read out sensor data with some simple commands. The easiest way is to connect for example with “PuTTy” (a Telnet client, **Figure 160**) in raw mode and type in a command. A more sophisticated way is to write a program in a high level programming language like JAVA, C#, Perl etc. or use tools like MATLAB or LabVIEW to exchange data and control the robot.

3 HARDWARE

The robotic platform VolksBot is a construction kit developed and distributed by the Fraunhofer Institute IAIS (Intelligent Analyse- und Informationssysteme; volksbot, 2011). The platform is especially constructed for “Physical Rapid Prototyping” of mobile service robots. The modular “Item-system” and 2 Maxon DC motors with 150W makes it extremely robust and reliable. It is designed for outdoor and indoor use respectively.

There are several types of the platform available. Here the 4-wheel version “RT 4” with 2 motors and differential steering (**Figure 157**) is be used. Short datasheet of the Robot is:

- Dimension: 540 x 500 x 260 mm (l x w x h)
- Wight: 13kg (without PC and accumulators)
- 260 x 85 mm wheels
- Gear ratio: 43:1
- Maximum velocity: 2.2 m/s
- Maximum load: 40kg

The basic set of the RT 4 includes a 3-channel motor controller with a RS232- and a CAN-Bus Interface. The RS232-Interface is well documented, so that a self-developed software interface for driving and steering can be written. The motor and the other electronic parts are powered by 2 12V/7.5Ah lead batteries.

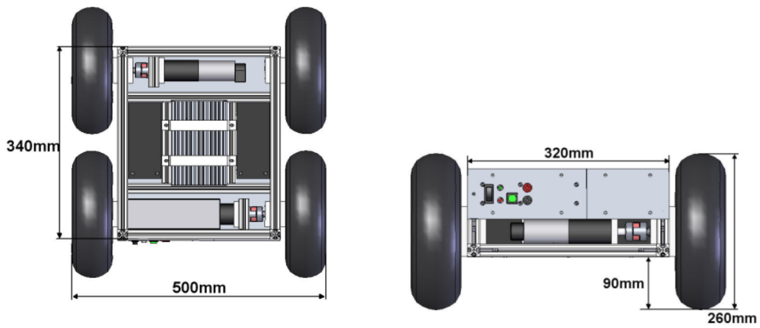


Figure 157: Volksbot® RT 4 from the top and from behind. (volksbot, 2011)

Recently installed sensor systems are:

- The **laser scanner** LMS100 from Sick. It is a very robust and reliable range finder which can scan a maximum distance of 32m in a range of 270° with a resolution of 0.5°. The sample rate is 50Hz.
- The **Time of Flight 3D camera** pmd3D from IFM has a resolution of 64 x 50 and gives back the distance up to 6.5m, and the intensity for every pixel.
- The **smart camera** leanXCam from SCS-Vision has a resolution of 752x512 and is equipped with a 500MHz Blackfin processor on board. There are 2 digital in- and outputs and an RS232 interface, which can be used for debugging purposes. Maximum frame rate of the camera is 30fps. The OpenCV library can be used for image processing.
- The **smart camera** NI1742 from National Instruments is a more professional camera system for industrial applications, equipped with a 533MHz PowerPC from Motorola. The resolution is 640x480. It has a maximum frame rate of 60fps.

All Sensors have an Ethernet interface. The motor controller and the display are connected over a device server (**Figure 158**), which converts TCP/IP to RS232 and back. There are 2 ports left for other devices i.e. GPS receiver.

Figure 158: System Technology

4 SOFTWARE

The software is split in 2 parts, the software part running on the robot (or somewhere in the network) and the software part connecting over the server (described later) to the robot. The first part includes all drivers for sensors and actuators. The second part is used for external visualization and processing of sensor data.

The software on the robot is written in C++ using the GNU C++ compiler. It runs on a 900MHz Mini-ITX with an Ubuntu Linux version 10.04. The development environment is installed on a virtual machine with the same Ubuntu version as on the robot. Thus, it is possible to compile the program on the VM, copy it to the PC on the robot and run it there. Connection from the VM to the robot can be established over VNC or SSH. Because the hardware communicates with TCP/IP, there is no difference if the program is running in the VM or on the robot (as long as the physical connection is fast enough). This makes software developing and debugging much easier.

The concept of the software is that every driver has its own class and runs in its own thread. Data is exchanging over a special class called "*SensorData.cpp*". After a thread is started, it tries to connect to the hardware, initializes it with settings out of the class "*Settings.cpp*" and writes data from the sensor to the data class. The "*WorkingClass.cpp*" class then takes the needed data from the data class, process it and sends drive commands, with the help of the class "*VMC.cpp*", to the motor controller.

The class "*ToolBox.cpp*" supports some functions for reading and converting sensor data and controlling the robot, i.e. "turn 120°" or "drive forward with 25% of maximum speed". **Figure 159** shows all classes used in the program "*SensorServer*".

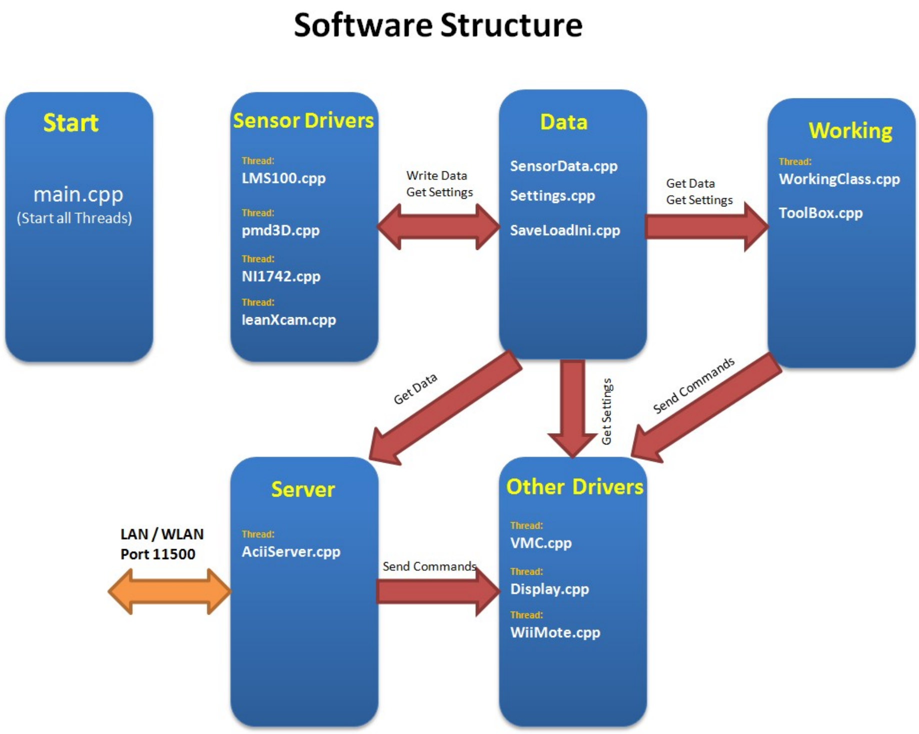


Figure 159: Software Structure

The class "*AsciiServer.cpp*" is also running in a thread and has the special function to communicate over a TCP/IP connection with external applications. The thread opens a port and waits for connection. A connected client can send commands in plain ASCII code. If the command is ok, the server returns an acknowledge "ACK". Otherwise it sends back a "no acknowledge" ("NCK"). **Figure 160** shows a PuTTy-Connection to the server with some commands which can be used by the client.

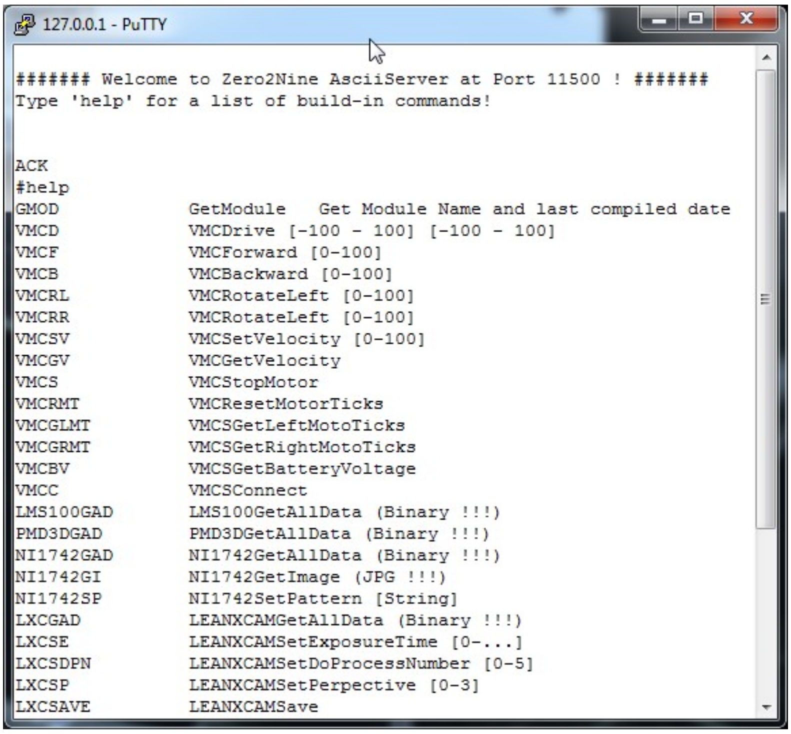
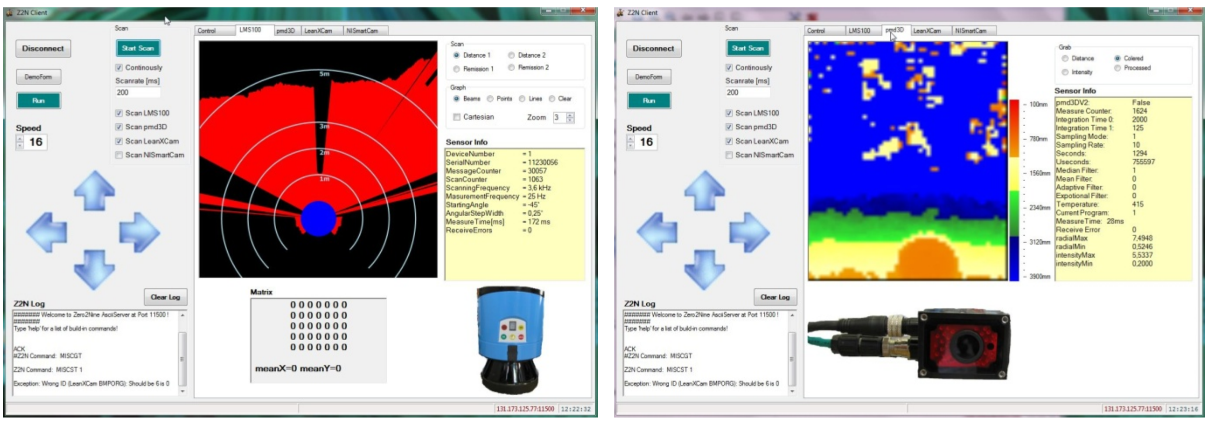


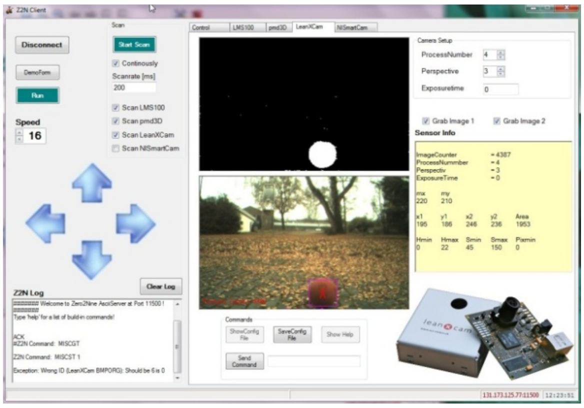
Figure 160: Connection to the “AsciiServer” with PuTTy.

A client can be a command line tool like PuTTy, it can also be a program written in a high level programming language. Operation system and programming language are of no relevance as long as it is possible to establish a TCP/IP connection. Sensor data returned from the server are binary. A data packet starts with *BinaryDataPacketID*, *BinaryDataPacketSize* and *BinaryDataPacketFingerprint* followed by a C-Structure including different types of data. This makes transmitting of the data more easy and secure. However the C-Structure must be known very well, to extract the data in the right way. There are some helpful clients written in C# which are used for development and demonstration.

Every sensor has its own client. The program “Z2N Client” is a collection of all clients together.



a) **b)**



c)

d)

Figure 161: Z2N Client with the visualization of a) the laser scanner, b) 3D-camera, c) leanXcam with color tracking and d) the original sight taken by a photo camera.

In **Figure 161 a** demonstration of the robot driving outdoor and tracking a red ball is shown. The laser scanner and also the 3D-camera give back the distance of the ball and other obstacles in the environment. The leanXCam tracks the color and returns the position of the detected object to the client. If one of the blue arrows will be pressed, the client sends a drive command to the robot i.e. turn left.

5 EDUCATIONAL ASPECTS

5.1 Lectures and laboratory experiments

The robot is used in lectures for explaining sensor systems and sensor fusion. It is helpful to get a "feeling" what a sensor can "see" and where are the limits of it. It is also employed in laboratory experiments, where students can concentrate on one sensor system, sensor fusion or navigation of a mechanical platform. The software structure allows the students a very short getting-started period. The only requirement is to have experience with C++ and maybe other high level programming languages.

In the special format called "advanced lab", students just work out a normal lab experiment e.g. using some function of openCV (an image processing library) on the leanXcam. Second step is to work out a more advanced exercise within the same experiment, given by the supervisors. As a third part the students than conceive an experiment selected by themselves. For example, this could be a tracking experiment with the leanXcam on the robot.

The "advanced lab" results in a higher motivation and a deeper understanding of the subject to the students. An essential part of the exercise is the presentation and discussion of the experimental results. So, the students can get an insight view to the other experiments in the lab. **Table 24** gives an overview of modules where the platform is used.

Table 24: Modules and Learning objectives

Modules	Courses	Learning objectives
Optoelectronics	Bachelor Electrical Engineering	Understanding optical sensor systems, integration in mobile platforms
Imaging Sensor Systems	Master Automation Systems	Image processing, understanding image sensor systems, programming, integration in mobile platforms
Seminar Mechatronic	Master Mechatronic System Engineering	Understanding sensor systems, navigation, programming, integration in mobile platforms

Sensor Systems	Master Mechatronic System Engineering	Understanding sensor systems, programming, integration in mobile platforms
Student projects	Bachelor and Master Courses	Topic defined by the supervisors, by students or proposed by companies

5.2 Student projects

Student projects are running over the whole term with small teams of typically three people. Thus, more complex issues like navigation through rows, object recognition or SLAM (Simultaneous localization and mapping) techniques can be worked on. Presentation and discussion of the results with the other students and the supervisor is also content of the project.

5.3 Bachelor and master thesis

Another field of application are bachelor and master thesis, where students evaluate new sensor systems i.e. Kinect-Camera from Microsoft. Because the camera has an USB-interface, the sensor software can run on a PC or a notebook, which navigates the robot over the "ASCII-Server". **Figure 162** shows an example of a bachelor thesis, where a student programmed a 3D-ToF-Camera from PMD with an USB-interface. For evaluation and development he used MATLAB running on a notebook. MATLAB processes the image data, establishes a TCP/IP connection to the robot and sent the commands for the navigation in plain ASCII-code.

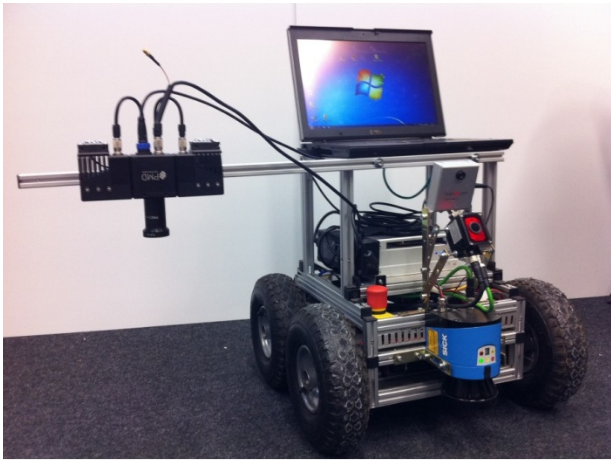


Figure 162: Zero2Nine equipped with a 3D-ToF-camera on the left side. The camera is used for soil properties detection.

5.4 Student competitions

Students use the system as a powerful starting tool for participating at robot competitions. Examples are the yearly International Field Robot Event (Mueller et al. 2006). The robots have to navigate between maize rows as fast and smooth as possible. Moreover, other tasks such as weed control are included. The development of robots and algorithms for outdoor applications results in high-level practical experiences for the students. Teamwork is an important aspect in the competition teams. They have to share different tasks like programming a sensor, preparing mechanical parts or create posters for public relations. There is also a team captain with the function of a supervisor. This is a perfect preparation for professional live. Moreover the students document their work in proceeding of the event (example: Feldkämper et al. 2010).

a)

b)

Figure 163: a) Field robot Team "optomaize Prime" at the Field Robot Event 2010 in Braunschweig, Germany, b) Zero2Nine" at the Ideen Expo 2011 in Hannover, Germany

5.5 Fairs and demonstrations

The robot ran on several fairs like Agritechnica 2009 (go.amazone, 2011) or Ideen Expo 2011 (hs-osnabrueck, 2011) (**Figure 163**) in Hannover for 7 and 10 days respectively. This showed the robustness of the system, because there haven`t been any mentionable hardware or software problems. The integrated Wii-control attracts children (and grownups) because they can control and interact with the robot. If there is an obstacle in front of the robot, they getting a vibration feedback. This is a good starting point to explain the sensors and the algorithm.

There are a lot of guidance tours in the lab for school classes or other research institutes, where Zero2Nine is used as a demonstration tool. Especially the school classes are exited and have a lot of question to the function of the robot.

6 CONCLUSION AND OUTLOOK

The students have a robot platform for evaluating, understanding and programming imaging sensor systems. There is a fast learning curve and a high motivation, because they can test and see the results immediately. The platform is well tested in a lot of projects and at robot competitions. Of course a short introduction to soft- and hardware is needed, but usually there is no more need for support.

A very simple and robust algorithm for driving around with avoiding collision and the visualisation of the sensor data makes Zero2Nine a nice tool for demonstrations in the lab and on fairs. People are getting a better understanding how an autonomous system can navigate by "itself".

Next development step is to integrate additionally ROS (Robot Operation System, ros, 2011). ROS is a widely distributed open source operation system especially for autonomous mobile robots. A lot of driver for different sensors and motor controllers and also more complex software like Kalman filter or SLAM-algorithm just exist. ROS seems to become an important tool for robot development.
Simulation is another topic. First experiences have been made in the lab with Microsoft Robotic Studio (Tsukor et al. 2011). Integration of the platform in MRST (microsoft, 2011) or GAZEBO (playerstage, 2011) is one of the next steps.

REFERENCES

FELDKÄMPER, N., HUFENDIEK, H., JAHN, H., KAMPMEYER, C., KERSSEN, S., LEMKE, C., MOSLER, D., RAVE, P., SIMON, S. TILNEAC, M., WESTERHOFF, J. WUNDER, E.: "*Team Optimaizer, Proceedings of the 8th Field Robot Event 2010*", urn:nbn:de:gbv:084-11100508057, p. 185-201, 2010, URL: http://www.digibib.tu-bs.de/?docid=00041345.

GRIEPENTROG, H.-W., RUCKELSHAUSEN, A., JÖRGENSEN, R.N., Lund, I.: *Autonomous systems for plant protection, in „Precision Crop Protection – The Challenge and Use of Heterogeneity*“, p. 323-334 (Editors Oerke,E.-C., Gerhards, R., Menz, G., Sikora, R.A.), Springer, 2010.

KlOSE, R., LINZ, A., RUCKELSHAUSEN, A.: *"Robotic Platform for Remote Control and Teleservice in Agicultural Environments"*, Proceedings of CIGR EurAgEng/VDI-MEG, 2006.

LEW, M.W., HORTON, T.B., SHERRIFF, M.S.:*"Using LEGO MINDSTORMS NXT and LEJOS in an Advanced Software Engineering Course*", 23rd IEEE Conference on Software Engineering Education and Training (CSEE&T), 2010.

MEEUSSEN, W., WISE, M., GLASER, S., CHITTA, S., MCGANN, C., MIHELICH, P., MARDER-EPPSTEIN, E., MUJA, M., ERUHIMOV, V., FOOTE, T., HSU, J., BOGDAN RUSU, R., MARTHI, B., BRADSKI, G., KONOLIGE, K., GERKEY, B., BERGER E.:*"Autonomous Door Opening and Plugging In with a Personal Robot*", IEEE International Conference on Robotics and Automation (ICRA), 2010.

MUELLER,J., WALTHER,S., BOETTINGER,S., RUCKELSHAUSEN, A., Van HENTEN,E.J.: *„Field Robot design contest - a showcase in Agricultural Engineering education"*; in: Kataoka T, Noguchi N, Murase H, editors. 3rd IFAC International Workshop on Bio-Robotics, Information Technology and Intelligent Control for Bioproduction Systems 9-10, September 2006. Sapporo (Japan); p. 276-281, 2006.

RUCKELSHAUSEN, A., BIBER, P., DORNA, M., GREMMES, H., KlOSE, R., LINZ, A., RAHE, R., RESCH, R., THIEL, M., TRAUTZ, D., WEISS, U.:“*Boni Rob: an autonomous field robot platform for individual plant phenotyping“*,Joint International Agricultural Conference, 2009.

TSUKOR, V.; KLOSE, R.; RUCKELSHAUSEN, A.: *“Multi-sensor simulation method for outdoor plant phenotyping based on autonomous field robots”*, Proceedings CIGR 2011, to be published.

Links:
http://www.irobot.com/de/store.aspx, last accessed on November 15, 2011

http://www.asurowiki.de/pmwiki/pmwiki.php, last accessed on November 15, 2011

http://www.volksbot.de/rt4-de.php, last accessed on November 15, 2011

http://www.hs-osnabrueck.de/722+M5ea081d3d65.html, last accessed on November 15, 2011

http://go.amazone.de/index.php?news=26, last accessed on November 15, 2011

http://www.ros.org/wiki/, last accessed on November 15, 2011

http://www.microsoft.com/robotics/, last accessed on November 15, 2011

http://playerstage.sourceforge.net/, last accessed on November 15, 2011

Sensor Integration II

Enhancement of the Navigation Data Quality of an Autonomous Flying Unmanned Aerial Vehicle for Use in Photogrammetry

Manfred Bäumker[1], Heinz-Jürgen Przybilla[1], Alexander Zurhorst[2]
[1]Department of Surveying and Geomatics, Bochum University of Applied Sciences, Germany
[2]aerometrics, Landwehrstraße 143, Werne, Germany

Abstract
These days Unmanned Aerial Vehicles (UAV) are beyond the stage of testing. Since 30 years now they have been used for a wide spectrum of applications. The main focus of UAV usage at Bochum University of Applied Sciences is photogrammetric applications. Of course these applications require high quality images so the camera is essential. But first of all the UAV has to be in the right place when the images are acquired – i.e. precise navigation devices are necessary. Up to now typical GPS sensors used in low cost systems do not allow the calculation of precise GPS positions. Therefore to enhance the UAV's navigation precision a realtime kinematic GPS system was build and tested . First results of these tests are presented, amended with the results of classical photogrammetric applications and the alternative use of "Patch-based Multi-view Stereo Software Systems".

Keywords
UAV, autonomous flight, navigation, kinematic GPS, photogrammetry, multi-view stereo system, pointcloud

1 INTRODUCTION

These days unmanned Aerial Vehicles (UAV) is beyond the stage of testing. Since 30 years now they have been used for a wide sprectrum of applications. (Przybilla & Wester-Ebbinghaus 1979, Wester-Ebbinghaus 1980, Eisenbeiss 2009, Eisenbeiss et al. 2011). In the military field these systems are established long since, but in the meantime they get more and more important in the civil area as well. The UAV presented here is a so called „MikroKopter" system which is developed under assistance of an internet community. All its electronic components are standard products, incorporating in sum an efficient yet low-prize UAV.

The main focus for using a UAV at Bochum University of Applied Sciences is photogrammetric applications. These require high quality images as well as precise navigation. But up to now typical GPS sensors used in low cost systems allow the calculation of differential GPS positions with an accuracy of not more than 1-2m in the horizontal and 3-5m in vertical direction. To enhance this accuracy an alternative approach was used, using instead a realtime kinematic GPS device. This solution will yield positioning accuracies better than a decimetre during autonomous flight.

2 TECHNICAL ASPECTS OF THE MIKROKOPTER SYSTEM

As MikroKopter is a sort of "do-it-yourself" project, its assembly requires considerable endeavors at the beginning. A comprehensive project documentation is published under www.mikrokopter.de providing all necessary informations amended by a comprehensive discussion forum.

Figure 164 shows the two different MikroKopter systems built that way which were used during the investigation.

Figure 164: Hexa- and Oktokopter during flight

Table 25 gives an overview of these systems technical parameters.

Table 25: Technical parameters of the MikroKopter system

Parameter	Value
Number of rotors:	4 – 12
Actual load:	250 g – 1000 gr
Weight:	650 g – 1700 gr
Flying time:	7 – 12 min
Distance:	Visual range
Flying height:	Max. 350 m (technically reliable)
Power supply:	Lipo 11,1 V – 14,8 V
Sensors:	Gyroscopes, accelerometers, compass, GPS, barometric altimeter

The collaboration of all electronic components (integrated gyroscopes, accelerometers for levelling the platform and barometric altimeter) is managed by the "Flight-Control" unit. The "Navi-Control" unit expands the autonomous flight options: an integrated GPS-sensor delivers the actual position, an optional telemetry-downlink module working at a frequency range of 868-870 MHz and an operational distance of several hundred meters will transfer all relevant flight and sensor data to the ground segment in realtime. Lastly a 5,8 GHz A/V-transmitter-system allows to broadcast the life image from the camera.

The tests were performed with a Ricoh GXR camera (resolution 3776 * 2832 pixel) with an 18mm lens, resulting in a payload of 750 gr approx. (together with the camera-carrier), (**Figure 165**). The reason for using that camera is its special concept, where lens and CMOS-sensor are combined within a single compact unit. Thus is the mechanical basis is given to achieve a stable Interior Orientation Elements of the camera.

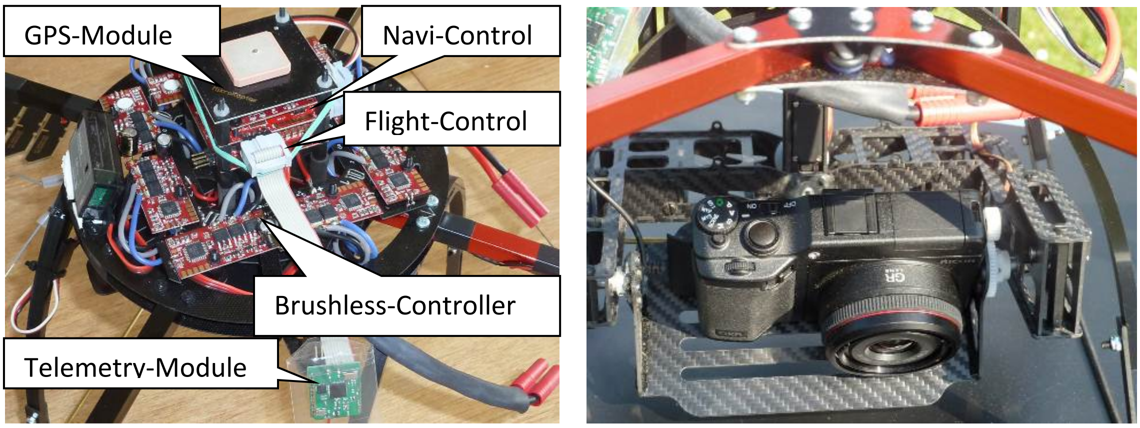


Figure 165: Electronic components (left) and digital camera Ricoh GXR (right)

3 EXAMINATION OF POSITIONING SENSORS

3.1 Actual status of positioning quality using low-cost GPS sensors

Up to now THE typical GPS sensor used in the MikroKopter and similar systems is the u-blox LEA-6S GPS receiver. This system receives 50 channels (GPS-L1/CA-Code, GALILEO OS) and is optionally equipped with an integrated SBAS (satellite based augmentation system) which is able to acquire the correction signals of WAAS, EGNOS or MSAS (cf. www.u-blox.com), (**Figure 166**).

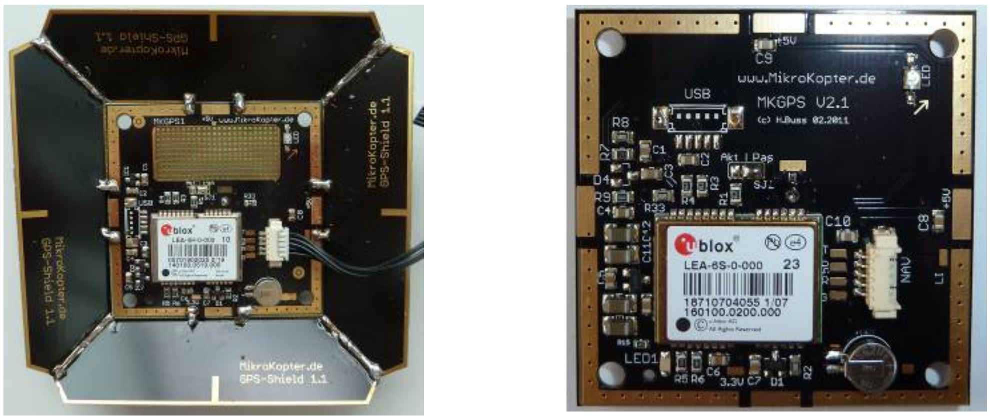


Figure 166: u-blox receivers as part of the GPS circuit board. Left: board surrounded by a protection shield to enhance GPS signal

In Bochum the signals of EGNOS (European Geostationary Navigation Overlay System) satellites Inmarsat AOR-E and IOR-W are available and were used to calculate differential GPS positions, which normally yield accuracies of 1-2m in horizontal and 3-5m in vertical direction. These could be proved by miscellaneous tests at the Bochum aerial testfield (Bäumker & Przybilla, 2011). **Figure 167** shows a typical result while positioning the MikroKopter at a defined waypoint. **Figure 168** shows the differences between pre-calculated and true positions (projection centers) during an autonomous aerial flight.

The procedure of an autonomous flight looks as follows: after launching the UAV and gaining the planned flight level the UAV is switched to autonomous flight mode. It will automatically approach the first waypoint then. After arriving at the waypoint within a certain tolerance distance (which has to be predefined) a countdown (normally some seconds) will start and during that countdown images can be taken. Some problems appear if the UAV doesn't succeed to approach the waypoint closely enough because of oscillation effects and bad GPS-Signal. This is a known problem and first steps to fix the oscillation effects were made. As a result the pre-planned waypoint positions can be reached not better than $\pm$ 3 – 4 m, not considering some outliers. As a consequence of that the planned overlapping of adjacent images often exceeds the tolerable limits (**Figure 169**).

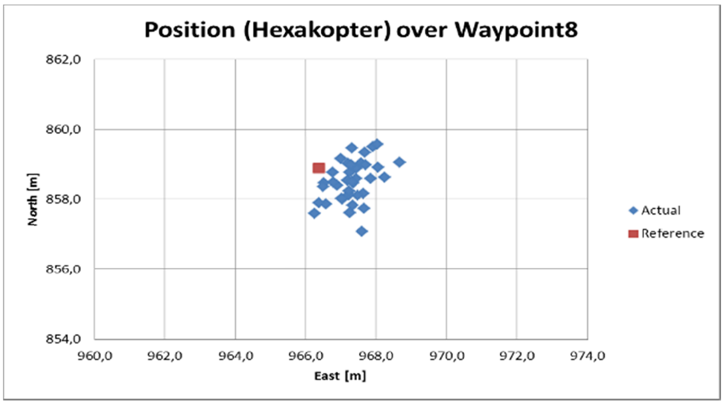


Figure 167: u-blox Copter-Tracking with tachymeter Trimble S6

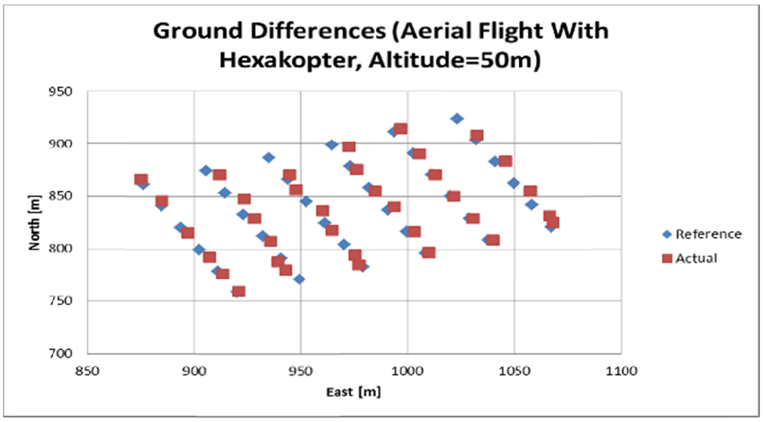


Figure 168: Comparison of reference and realised ground positions during aerial flight

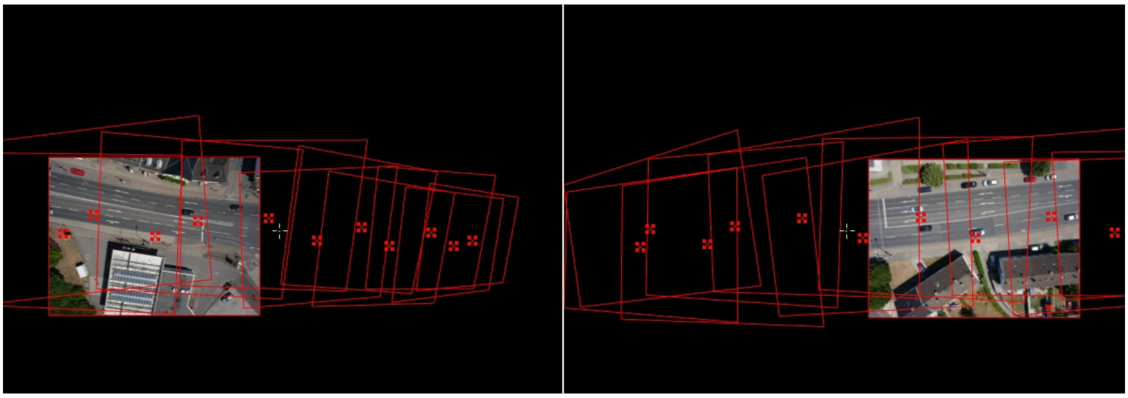


Figure 169: Footprints of an aerial flight showing the overlapping of adjacent images

3.2 Enhancements of positioning quality

As was shown in 3.1 the MikroKopters positioning quality often exceeds the given limits. A possible solution of this problem might be the integration of high performance GPS positioning using real-time kinematic (RTK) technologies. The first requirement for this is the availability of an appropriate GPS-receiver, the second the availability of a likewise appropriate software product calculating the positions. Both requisitions could be met.

The actual development uses a one-frequency receiver of the u-blox series (LEA-6T) in combination with an own reference station. The RTK-calculations are performed with RTKLIB, an open source program package for GNSS positioning (RTKLIB 2012).

The software package is distributed free of charge under GNU GPL v3-license and thus published with source code. It supports standard and precise positioning algorithms with GPS, GLONASS, SBAS, GALILEO (which is enabled but not supported in current version) and QZSS (Japan) as well as various positioning modes with GNSS for both real-time and post-processing (Single-point, DGPS/DGNSS, Kinematic, Static, Moving-baseline etc.). A wide range of standard formats and protocols for GNSS can be processed, like RINEXi, RTCMx, as well as the proprietary messages of several GNSS receivers. The external communication interface supports serial, TCP/IP, NTRIP and FTP/HTTP protocols. All in all RTKLIB is a high-tech solution and predestinated for a low-cost UAV-system.

The GPS-data processing during an autonomous flight is carried out as follows: the u-blox raw-data are sent to the ground control station via telemetry. Realtime calculations of the RTK-solutions are carried out with RTKLIB. Further on the raw-data of the UAV u-blox as well as those of the reference station are stored so it's possible to calculate improved positions with the post-processing module of RTKLIB later on.

First tests were carried out at the Bochum aerial testfield. **Figure 170** shows the test bed with the reference station consisting of the JAVAD Triumph-1 G3T 2-frequency-GNSS-receiver and a notebook. The rover used is a u-blox-LEA-6T (one-frequency).

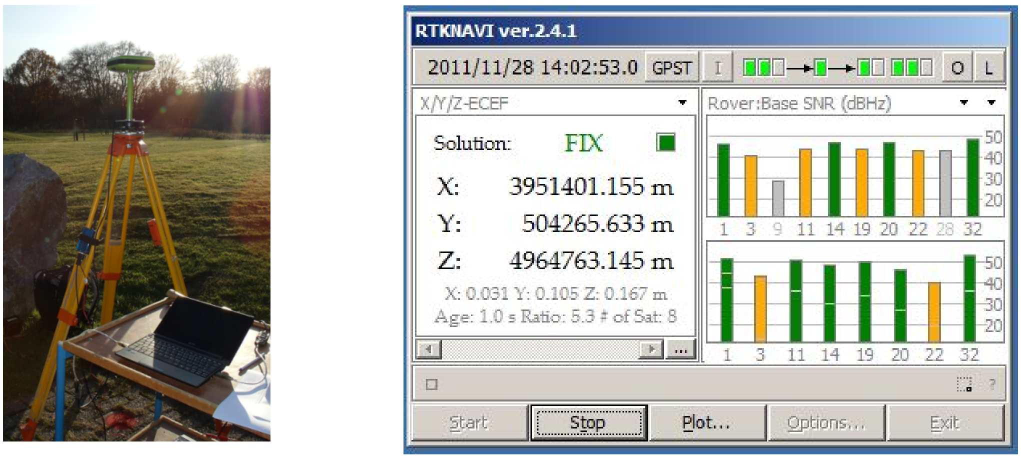


Figure 170: Reference station with JAVAD Triumph-1 G3T (left). RTKLIB screenshot (right)

The results of the realtime calculations done with RTKLIB are shown in **Figure 171**, using the integrated visualization module RTKPLOT. The different colors represent the accuracies achieved:

- red: single GPS-solution: 5 m – 20 m
- orange: float- solution: 0.1 m – 1.0 m
- green: fixed- solution: 0.02 m – 0.10 m (first test: about 90% of the time)

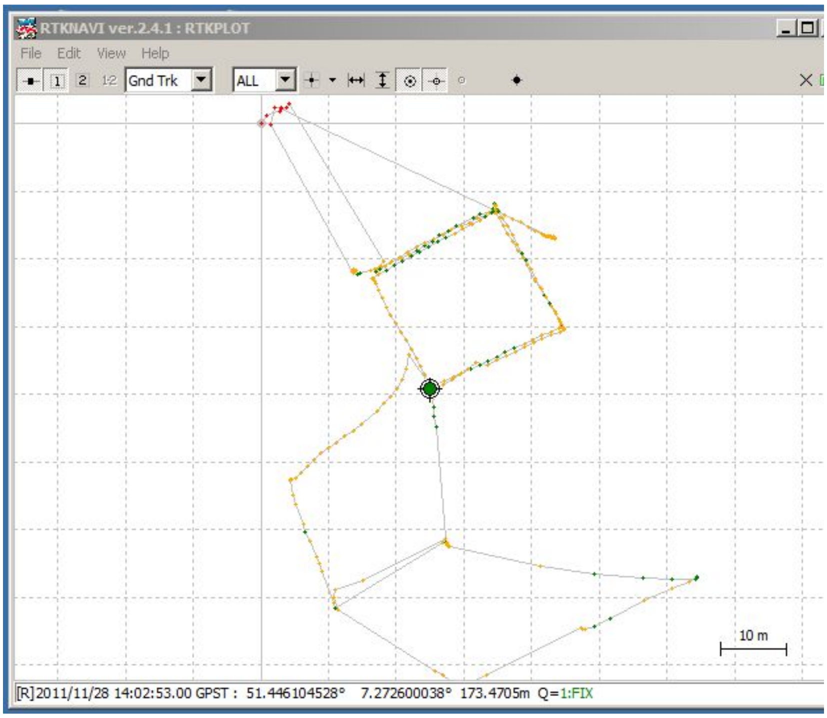


Figure 171: Accuracy of realtime calculations with RTKLIB while positioning the u-blox sensor at the test field

With the exception of the initiation phase lasting 30 seconds the RTK processing software calculates either a fixed solution or a float solution. Instead of an own references station it is possible to use the data of a foreign reference station or a reference service like the German SAPOS HEPS-Service.

4 PHOTOGRAMMETRIC APPLICATIONS

Aerial photogrammetry is a typical application for UAVs. Among the classic aerotriangulation processes and products, like maps and orthoimages, more and more the issue “pointcloud” is of interest. Today software often is developed by internet communities and distributed as open source later on. This is also true for the community interested in computer vision. Their software products enable feature extraction, image orientation and the generation of pointclouds as fully automatical processes, which can be done by means of Web Services or also on the local machine. Web Services familiar to many are “Photosynth” by Microsoft (Microsoft 2012) or - more recently - Autodesk 123D (Autodesk 2012). “PhotoScan” (Agisoft 2012) and “Bundler/PMVS2” (Snavely 2012) are typical representatives of local software solutions.

Bundler is a so called “structure-from-motion (SfM) system” operating on (large) unordered image collections. Its distribution also contains potentially useful implementations of several computer vision algorithms” (Snavely 2012). While Bundler produces sparse point clouds its functional alternative, the PMVS2 software package (Furukawa 2012) calculates denser points.

In context with some UAV investigations at the Bochum aerial testfield diverse oblique images of a neighbouring building were taken (**Figure 172**). Then Bundler/PMVS2 was used to triangulate a block of the oblique images and to generate a pointcloud. Based on 74 images a pointcloud with 537.546 points was computed.

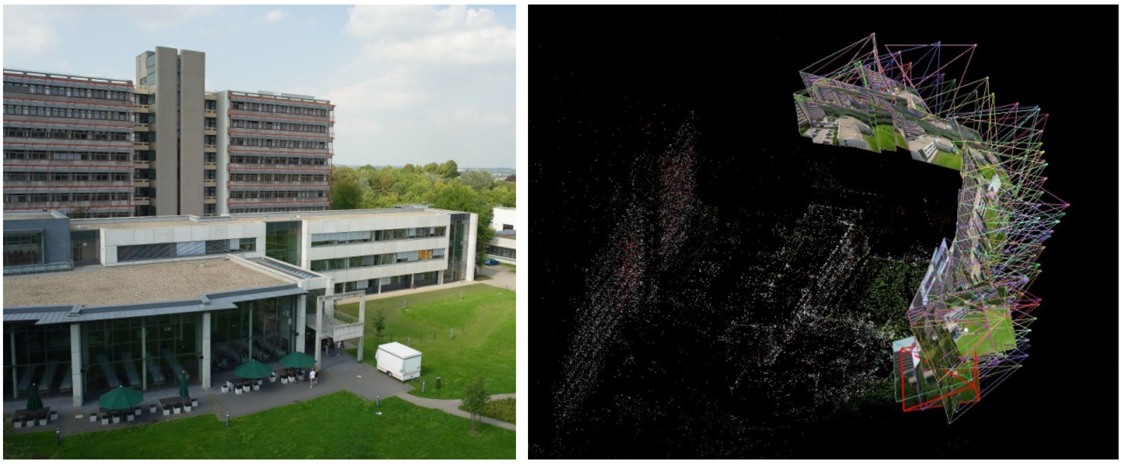


Figure 172: Example of an aerial image taken with UAV (left), Bundler image orientation (right)

After reducing the cloud to the points representing the building 190.462 points remained. These points were transferred to Cyclone (Leica Geosystems) to construct a basic 3D-Model. First results are shown in **Figure 173**. A final validation of that model geometry was not yet done but will be done by comparison with an equivalent 3D-Model, which is still under construction. The preliminary results are promising, because the pointcloud density not only allows to construct planes automatically (via region growth) but also line elements.

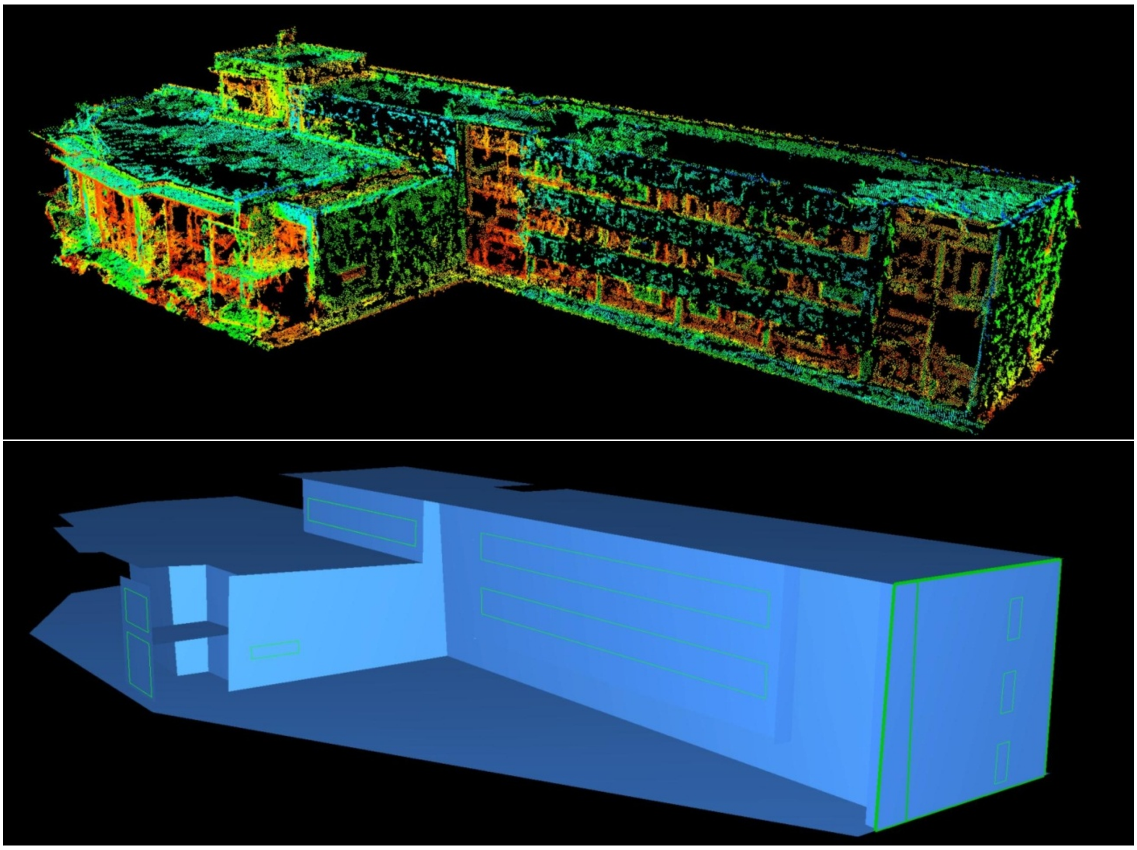


Figure 173: Pointcloud als a result of PVMS2-algorithm (top), 3D-Model (bottom)

5 CONCLUSIONS

The navigation accuracy is an important requirement for a successful operation of an UAVs during its autonomous flight. It has been shown, that a standard low-cost (one frequency) GPS-receiver in combination with free GNNS positioning software can be used to establish a realtime kinematic GPS solution with accuracy within the cm-level.

Is seems possible to replace the local reference station – which is actually needed – with a SAPOS service. This will be tested next. The replacement of the MikroKopters one frequency receiver by a two frequency receiver will be the next important improvement, as thus a partial direct georeferencing of the UAV will be possible.

With regard to the measuring objects with bundles of images a computer vision approach was presented, which allows the establishment of dense pointclouds with an almost automatic procedure. Further investigations have to be undertaken to prove the quality with regard to geodetic applications.

REFERENCES

Books:

EISENBEISS, H.; UAV Photogrammetry. Dissertation ETH Zürich. IPG Mitteilungen Nr. 105, ISSN 0252-9335, 2009

Journal articles:

BÄUMKER, M., PRZYBILLA, H.-J.*: Investigations on the quality of the navigation data of Unmanned Aerial Vehicles using the example of the system MikroKopter;* International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol XXXVIII-1/C22, ISSN 1682-1777, September 14-16, 2011

EISENBEISS, H., KUNZ, M., INGENSAND, H. (EDS*.): Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g),* The International Archives of the Photogrammetriy, Remote Sensing and Spatial Information Science, Vol XXXVIII-1/C22, ISSN 1682-1777, http://www.geometh.ethz.ch/uav_g/proceedings, September 14-16, 2011

PRZYBILLA, H.-J. & WESTER-EBBINGHAUS, W.: *Bildflug mit ferngelenktem Kleinflugzeug.* Bildmessung und Luftbildwesen. Zeitschrift für Photogrammetrie und Fernerkundung. *HerbertWichman Verlag. Karlsruhe, 1979*

WESTER-EBBINGHAUS, W.: *Aerial Photography By Radio Controlled Model Helicopter.* The Photogrammetric Record, Volume 10, Issue 55, p. 85–92, 1980

Links:

http://www.agisoft.ru/products/photoscan/, Agisoft Photoscan, last accessed on January 10, 2012

http://grail.cs.washington.edu/software/pmvs/, Furukuwa, Y: Patch-based Multi-view Stereo Software (PMVS - Version 2), last accessed on January 10, 2012

http://labs.autodesk.com/utilities/photo_scene_editor/, Autodesk® 123D™ Catch Beta (formerly Project Photofly), last accessed on January 10, 2012

http://www.mikrokopter.de, Website for MikroKopter do-it-yourself construction, last accessed on January 10, 2012

http://photosynth.net/, Microsoft Photosynth, last accessed on January 10, 2012

http://phototour.cs.washington.edu/bundler/#S2, Snavely, N., Bundler: Structure from Motion (SfM) for Unordered Image Collections, last accessed on January 10, 2012 http://www.rtklib.com, RTKLIB, an open source program package for GNSS positioning, last accessed on January 10, 2012

http://www.u-blox.com, u-blox 6 receiver description including protocol specification, last accessed on January 10, 2012

Precise Relative Positioning in Machine Swarms

Jan Schattenberg[1], Thorsten Lang[1] , Martin Becker[2], Simon Batzdorfer[2], Ulf Bestmann[2], Peter Hecker[2], Franz Andert[3]
[1]Intitute of Agricultural Machinery and Fluid Power, Technische Universität Braunschweig, Germany
[2]Institute of Flight Guidance, Technische Universität Braunschweig, Germany
[3]Institute of Flight Systems, Unmanned Aircraft, German Aerospace Center, Braunschweig, Germany

Abstract
The increasing automation of mobile working machines and the progressive use of more than one machine up to swarms of machines for cooperative tasks demands information about the relative position between the machines as well as between the machines and their attachments. This is especially necessary and important when carrying out tasks on distributed machines in a formation, like for example the parallel harvesting process in agricultural business or the cooperative search for survivors after a snow slide. Moreover, it is very important to ensure relative position information in case of partial failure or a poor reception of the GNSS receiver, for example to avoid the collision between the machines.

Options to handle this problem are the coupling of Global Navigation Satellite System (GNSS) measurements with measurements of an Inertial Measurement Unit (IMU). A further improvement or stabilisation can be done by vision based systems, like 2D- or 3D-camera system using methods of optical flow for motion estimation. A further possibility for improvement for determining the swarm geometry, which will be described, is the so called swarm positioning method. This method is based on the exchange of the GNSS raw data, i.e. range measurements, between each participant in the swarm using a mobile ad-hoc network. Additionally, GNSS raw measurements and inertial measurements are coupled using multiple filters in order to detect degraded GNSS measurements and exclude these from further data processing.

The challenge of the mobile ad-hoc network is the time variant network structure and the small available transmission rate in combination with a high demand for quick data exchange. Therefore different routing algorithms have to be combined and developed to ensure an information exchange in various scenarios.

The research and developments are part of the joint research project "Next UAV – Navigation for exploration with UAVs at low altitude in disaster Scenarios", which is funded by the Bundesministerium für Wirtschaft und Technologie (BMWi) and is administered by the Space Agency of the German Aerospace Center (DLR) in Bonn.

Keywords
GNSS, INS, Communication, Optical Flow, Cooperative Tasks, Mobile Ad-hoc Network

1 INTRODUCTION

With the increasing automation of mobile machines and the progressive use of machine swarms also the relative positioning of the machines to each other as well as to their attachments, in addition to the absolute position, becomes an increasingly important role. This is particularly necessary in order for machines to perform various tasks in a formation. In addition it is important to cope with degraded GNSS measurements or a complete loss of GNSS information's. Possibilities therefore are to couple the GNSS position solution with the data of the inertial measurement unit (IMU) and with the information of an imaging system, which calculates the movement of the rover by motion estimation.

By using special filters it is possible to detect degraded GNSS signals or measurement errors and exclude these from further processing. One possibility for determining the relative position is the so-called method of swarm location, which is based on the exchange of the GNSS raw data of each participant through the swarm using a mobile ad-hoc network. Therefore, the GNSS raw data of each participant must be exchanged quickly and safely. In addition the system must be respond very flexible to changes in network topology and compensate for example the loss of individual swarm participants. Due to the external conditions also transmit and receive power as well as the antenna technology is limited. Since the individual participants are exchangeable, a system is needed, which is decentralized. For such tasks so called mobile ad-hoc networks (MANet) or mesh networks are predestined. By special routing methods, which respond dynamically to changes in network topology, efficient data exchange can be achieved. The data exchange takes place directly and also indirectly via intermediate stations.

2 TECHNICAL EQUIPMENT

In addition to the ground based flexible research platform comRoBS by the Institute of Agricultural Machinery and Fluid Power ([1],[2]) the unmanned helicopter testbed ARTIS (Autonomous Rotorcraft Testbed for Intelligent Systems) by the DLR-Institute of Flight Systems(0) and quadrocopters by the Ascending Technologies GmbH (**Figure 174**) are used for the tests.

Figure 174: comRoBS, DLR ARTIS, AscTec Pelican Quadrokopter (left to right)

These platforms are equipped for the experiments with a unified system of a navigation computer for data processing, an inertial measurement unit and a communication system for exchange of GNSS raw data, as well as with their own individual sensors and actuators. Moreover they have a wireless connection for data access and debugging. The testbed ARTIS is also equipped with a further computer for image processing and with the corresponding system of a stereo camera as a composite system for motion estimation (**Figure 175**). The GNSS Receiver is based on a µblox LEA-6T-chip which provides timing and raw data output, which is needed for the positioning within the swarm and the integrated navigation. Furthermore the SBAS (Satellite Based Augmentation System) system EGNOS (European Geostationary Navigation Overlay System) can be used for an increasing of the position accuracy and integrity (**Figure 175** left),[6].

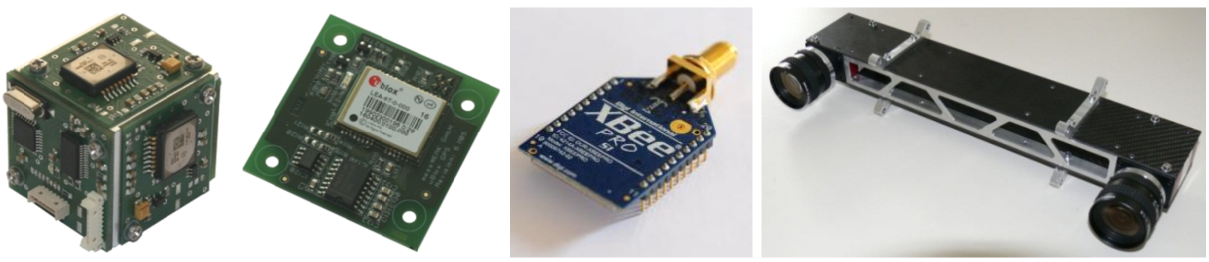


Figure 175: NExt UAV Hardware, IMU, GNSS-Receiver, Radio Module, Stereo camera composite system

Considering the requirements especially of the “Pelican” quadrocopter in respect to the mechanical construction and the limited payload the IMU has been developed and custom-built by the Institute of Flight Guidance of the Technische Universität Braunschweig. The IMU is based on Micro-Electro-Mechanical-System (MEMS) sensors to be small and light weight. The result is a cube with 32mm edge length and a total weight of around 40g. **Table 26** shows the technical data of the used sensors:

Table 26: Technical Data IMU

Quantity	Type	Product	Resolution
2	Dual axis acceleration sensor	Bosch SMB225	16 bit ±4.9 g low-g
3	1-axis gyro sensor	Bosch SMG074	+-187 [°/s], 0,2 [°/s] g-sensity

In order to provide meaningful inertial measurements the IMU has been calibrated using nonlinear calibration models with respect to temperature, as described in [4],[7].

The radio module used for data exchange within the swarm is based on the wireless standard 802.15.4 Xbee and uses the frequency band of 2.4 GHz (ISM-band). It is the pro version of the module type with a radio range of up to 1.6 km under best conditions (outdoor). It is connected by a serial interface with a serial data range of up to 115.2 kBps to the navigation processor board, which is a commercial platform based on a Intel Atom (Z530) processor with 1.6 GHz and the standard Pico ITX dimension of 3.9”x2.0” (100 x 72 mm). The stereo camera composite system consists of two AlliedVisionTec Marlin F-131B cameras with Pentax C815B objectives. The cameras have a resolution of 1280x1024 and 25 fps on full resolution. They are connected by IEEE 1294a interface to the vision pc, which is based on a Core2Duo 9300. The system if fully integrated into the architecture of the ARTIS helicopters.

3 INTEGRATION OF GNSS/INS AND FAILURE DETECTION AND EXCLUSION METHODS

Additionally to a vehicle's position the complete state of the vehicle, i.e. position, velocity, attitude and heading is of interest, when thinking about vehicle guidance and control applications. In order to determine the complete rover state vector, the UGV and UAV described above will be equipped with the presented Inertial Measurement Unit. The vehicle´s accelerations and turn rates measured by the IMU are calculated using a strap down mechanization to get the vehicle's position, velocity, attitude and heading. As GNSS based aiding information GNSS raw measurements, i.e. pseudoranges, are used. Using GNSS raw measurements instead of a pre-processed GNSS position has the advantage that also IMU aiding with less than four measurements is possible.

Post processing of first IMU data conduced in dynamic movement show excellent performance for a MEMS class IMU. As a measure of performance, the filter innovations were chosen that equal the length of the position update vector. (**Figure 176**) The plot over time shows constant navigation accuracy below 20 cm deviation. Furthermore, the absolute error has been evaluated by comparing the results with a couple of known waypoints.

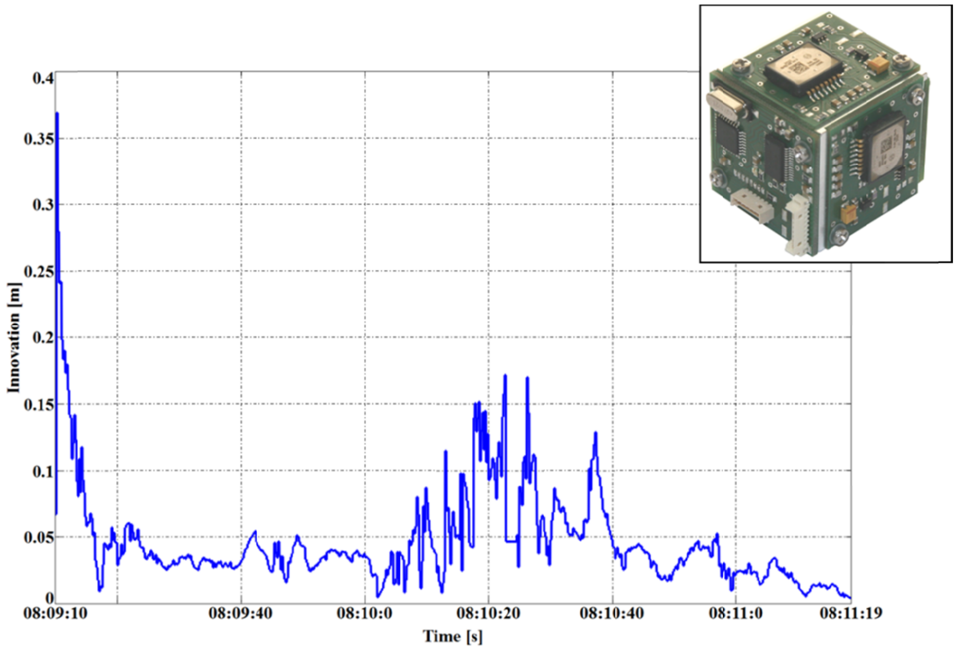


Figure 176: Performance of the coupling filter

In order to detect degraded GNSS signals or measurement errors a GNSS signal monitoring is conducted using inertial measurements (ref. 0). Therefore multiple coupling filters are running parallel forming a filter bank, as depicted in **Figure 177**

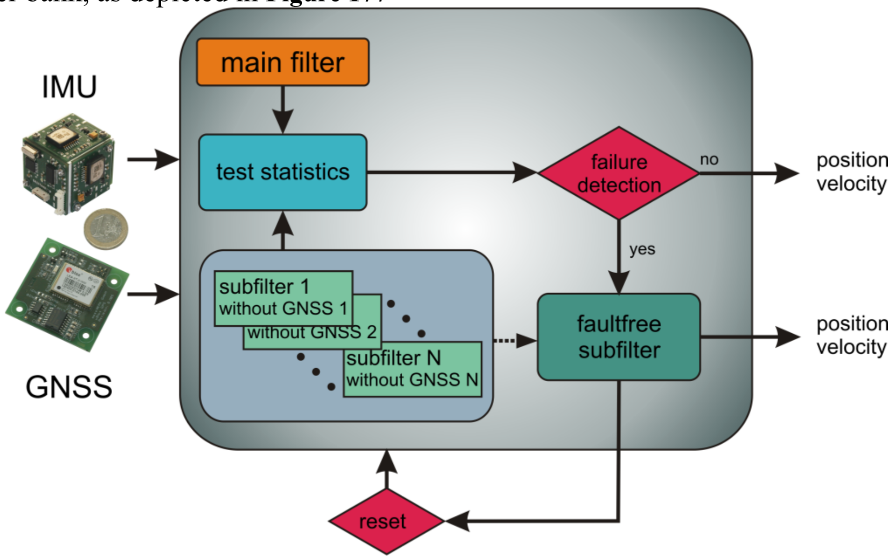


Figure 177: Filter bank architecture

One filter acts as main filter, using all GNSS measurements, the other N filters (N equals the number of GNSS measurements) act as sub filters. Each sub filter omits one measurement so that by applying statistical tests a faulty measurement could be detected and excluded from further processing (*Failure Detection and Exclusion, FDE*). The fault free sub filter parameters, (filter state and covariance matrices) could then be used to reinitialize the corrupted filters in order to avoid additional time needed to reach a steady state of the filter bank.

4 SWARM POSITIONING USING GNSS

Independent from special applications, a dedicated knowledge of the UAV/UGV swarm geometry and position is indispensable. For controlling a swarm and referencing exploration results one need both: the absolute position of the swarm participants as well as the baselines between them to determine swarm geometry. A simple way to obtain the baselines is to compute the vectorial difference of absolute positions one gets using GNSS. In the case of standalone GNSS, i.e. there are no aiding information from reference stations available, the accuracy often doesn't fit the requirements. To cope with this effect, all swarm participants will share their GNSS range measurements using a wireless ad-hoc communication. So every rover will be able to use the measurements of all other to calculate absolute and relative positions of the whole swarm using a differential approach. In contrast to GNSS reference station applications, within the swarm more than two participants are cooperating and that the position of all participants is time variant.

The principle of the approach using GPS code measurements for determining absolute and relative position of swarm participants is shown in **Figure 178**.

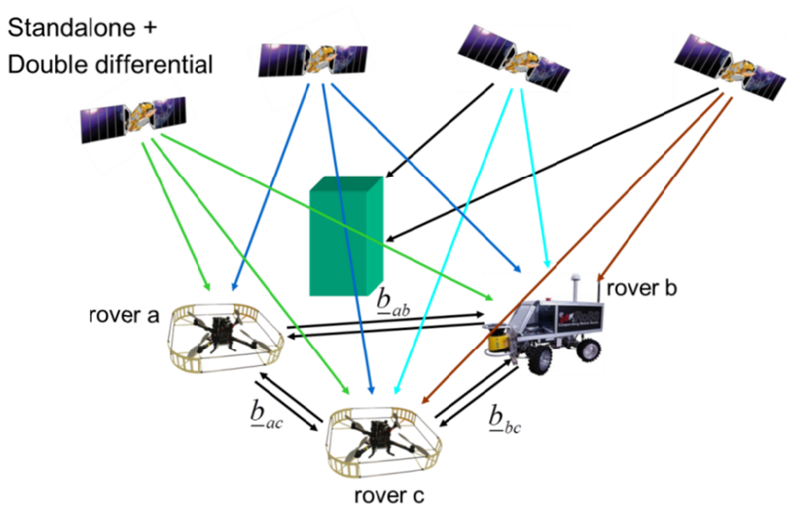


Figure 178: Combined Swarm Navigation Algorithm

The collected GNSS data of a swarm of five mixed UAV/UGV rover are processed using the standalone, double differenced baselines and the proposed combined algorithm. The preliminary results are given by the following **Table 27** as a qualitative comparison of mentioned techniques.

Table 27: Qualitative comparison between different solutions for baseline determination

	noise	processing	complexity	accuracy	availability
Δ standalone	−	+	+	−	+
Δ∇ double differences	−	−	−	−	−
ΔΔ∇ standalone + double diff.	+	−−	−−	−	+
RTK	++	−−	−−	++	−−

Regarding the accuracy of the determined baselines between the swarm participators representing the swarm geometry, a significant improvement has not been achieved. This can be ascribed to the used test set up with relatively short baselines between the UAV/UGV so that the compensating of error components by using the differential approach compared to standalone techniques doesn't became apparent. For further examination of this effect, more trials have to be conducted realizing different stretched formations of the swarm. Concerning the availability the combined approach shows no difference to the usage of standalone based baseline calculation, because all measurements could be used and are not reduced to common satellites tracked by all receivers.

Having a look on the noise of the relative positioning results, the noise level does not change remarkably between the different solutions as long as no carrier phase is included. While the noise of the different positions adds up according to the laws of error propagation, the noise from four measurements is included in each of the double differences. As a result the complex algorithm is expected to result in a significant improvement in noise only if carrier phase measurements are used. In this case the relative positioning has RTK accuracy and the more independent measurements of all receivers improve noise and accuracy of the overall swarm position for scenarios, where no reference station is available.

5 VISION-AIDED RELATIVE LOCALISATION

With small and lightweight cameras and a lot of computational power, image processing techniques are more and more interesting for mobile systems. Basis for image-based localization are modern algorithms that calculate the camera's motion from image sequences in real-time. This is performed by identifying and tracking significant features (e.g. edges and corners with high contrast) over time, known as optical flow. With the characteristics of these feature movements, the ego-motion of the camera can be calculated. With that, the localization filter bank avails an additional sensor to compensate the disadvantages of GNSS and IMU. Actual research investigates techniques and algorithms for an overall integration of camera systems into a reliable and robust navigation system.

By the analysis of image sequences, a set of tracked features with their motion (i.e. optical flow) is calculated first. Since localization should be applicable to unknown environments, these features are not matched with prior knowledge from environmental maps. With that, optical flow gives only information about the relative orientation between two camera frames. Due to that, an image processing system measures only relative movements like an inertial system which means that similar issues like error accumulation must be considered. In addition to that, the relative movements are scale-invariant which means that there are only five of six degrees of freedom measured, with no information about the absolute size of movement. With two or more calibrated cameras, absolute movements can be measured, though, up to a certain distance depending on the camera baselines. The stereo camera system used within this project allows the measurement of distances between approximately 10 m and 40 m, which means that images of objects in this distance field will give full movement information in six degrees of freedom.

In addition to the integration of an optical flow sensor into the navigation filter, research investigations are going to improve the image processing performance and robustness by re-coupling the navigation prediction into the optical flow measurements as depicted in **Figure 179**. This is done because classical optical flow and image-based relative localization only depend on the image quality. In this new approach, the navigation prediction helps to find an initial solution for the image movement and to eventually filter or correct bad optical flow vectors. Additionally, a prediction of optical flow increases the chance of finding the optical flow corresponding to the predicted vehicle movement, and reduces the probability of losing optical flow vectors.

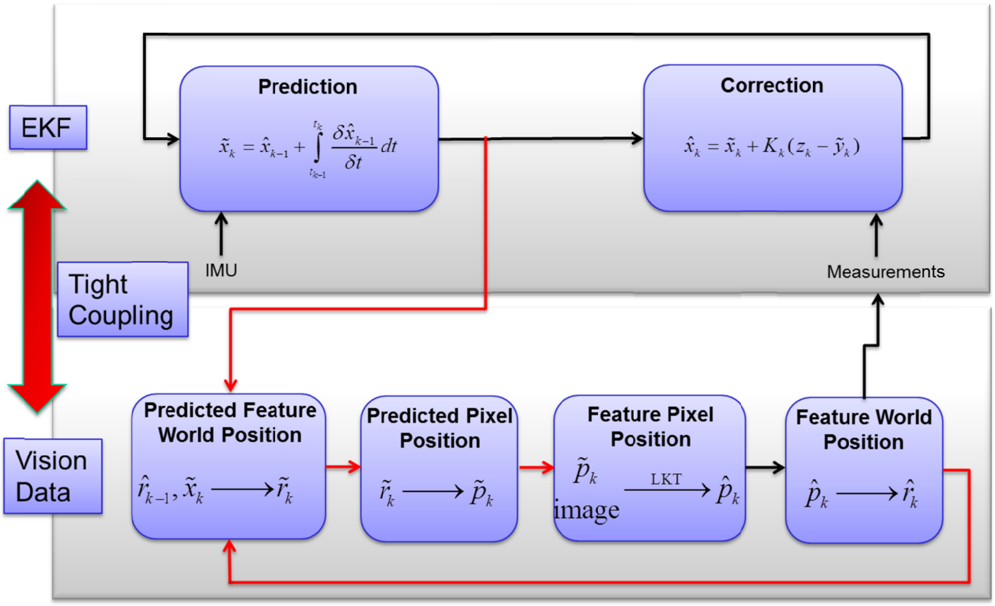


Figure 179: Tight coupling between image processing and navigation filter with a prediction of optical flow

6 REFERENCE LOCALIZATION SYSTEM

A robotic total station is used to benchmark the developed GNSS/INS navigation techniques as seen in **Figure 180**. The localization principle is ground-based azimuth and elevation measuring combined with a laser-based distance measurement which yields the 3D position of the target. Since some "robotic" devices can automatically aim and track moving objects, they are suitable for low-flying aircraft localization. For best measuring results, retro reflection prisms are installed on the vehicles. Typical ranges of application are between 300 m and 1000 m, mainly depending on the size of the reflector and the strength of the laser beam.

Figure 180: Leica Viva Robotic Total station (left), ARTIS helicopter with reflection prism (right)

Typical 3D positioning accuracy is at least one centimetre. 3D vehicle localization is performed with one single device which is easy to set up on a flight test site. The used Leica Viva TS15 device performs automatic vehicle tracking with up to 6 Hz, so that the full 3D path of the vehicle can be measured, to be compared with the on-board navigation solution.

7 MOBILE AD-HOC COMMUNICATION

To improve the relative position within the swarm the GNSS raw data of each participant has to be exchanged as quickly and as safely as possible as mentioned in chapter 4. In addition, the system must be flexible to the dynamic changes of the network topology and react very fast to compensate for example the loss of individual swarm participants without having a network collapse. Due to the external conditions of small and light weight rovers also the transmitting and receiving power as well as the usable antenna technology is limited. Since the individual participants are exchangeable, a decentralized system is needed without a master instance. For such tasks so called “Mobile ad-hoc networks (MANet) or mesh networks are predisposed. Using specialized routing algorithms, which respond dynamically to changes in network topology, an efficient data exchange can be achieved. The data exchange can be thereby done directly and indirectly through intermediate stations.

Figure 181 shows a possible constellation of the test vehicles used in the project, where some participants can communicate directly with each other and others only indirectly, through one or more intermediate stations.

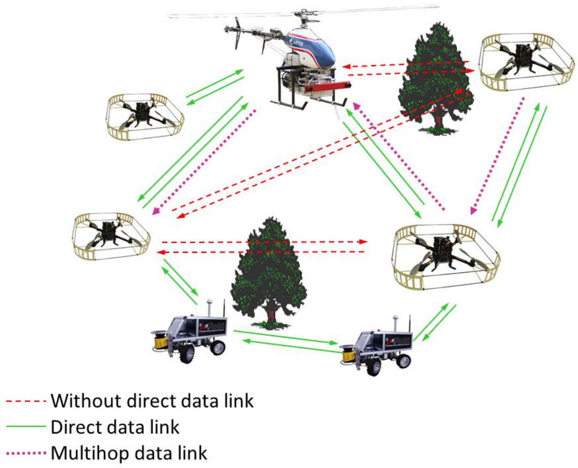


Figure 181: Connection types in an example mobile ad-hoc network configuration

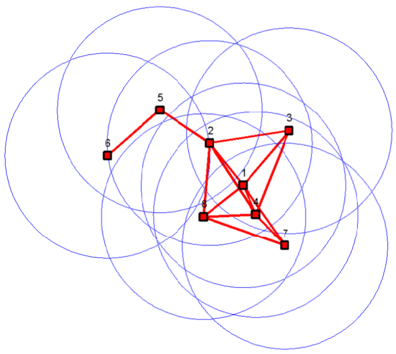
In the field of mobile ad hoc networks, there are various trends and approaches ([6]) for methods to exchange the data. They are usually very specific designed to the particular requirements, which are described for example by the number of participants, the energy consumption of the radio modules, the radio range, etc. Therefore special routing methods have to be used. In first they can be differentiated as proactive and reactive algorithms, and mixed forms, the so-called hybrid methods. A further possible method is a position-based routing process, which uses the information about the position and the state vector of the vehicles to predict the constellation of the network in the next step and to optimise the network discovery. Depending on the situation, all methods have advantages and disadvantages in relation to, for example, the time until the network structure is explored or the data overhead required for the exchange of network information. The objective therefore it is to include multiple routing methods into one system and then dynamically selecting the most appropriate method by recognizing the situation, to be as flexible as possible. A further challenge is the synchronization of the machine swarm, which is required for the real-time processing of the raw data of all participants. The synchronisation therefore is realised by the usage of the GNSS time in combination with a clock which is implemented within the IMU module. For having enough time for data processing the system also has to fulfil the very high demands on a time-effective data exchange. Belonging to these requirements the wireless standard IEEE 802.15.4 is used as basic communication standard, which is particularly suitable because of the low latency for establishing a connection in dynamic networks, the low energy consumption and the small and light-weight modules.

To get a possibility for comparing different algorithms for network discovery and data exchange a simulation was build up using MATLAB® as simulation platform. The simulation is realized in four steps:

1. Generating a random network with a fixed number of nodes within the network and
    a. a variable number of visible satellites (variable data volume which hast to be exchanged)

   b. a fixed number of 12 visible satellites (maximum data volume which has to be exchanged)
2. Simulating the network discovery and calculating the time needed for exploration
3. Calculating the routing path (parallel implementation of different algorithm possible to compare different approaches on the same data base)
4. Simulating the data exchange based on the integrated routing algorithms

Figure 182 shows an example of a random network structure with eight nodes and the corresponding adjacency table which presents the existing connections within the network and therefore is the basement for the routing algorithm.



	1	2	3	4	5	6	7	8
1		x	x	x			x	x
2	x		x	x	x			x
3	x	x		x				
4	x	x	x				x	x
5		x				x		
6					x			
7	x			x				x
8	x	x		x			x	

Figure 182: Random generated network topology and corresponding adjacency table

The simulation of the network discovery itself is realized by some repeating steps until all nodes have the same adjacency table. This is only possible in simulation mode, but gives some statistical results regarding the time needed for exploring the network and can be integrated can be integrated as time criterion for the final system.

Figure 183 depicts the simulation result for 1000 networks and a size of eight nodes. Presented are the percentages of simulations (y-axis) over the overall time needed for exchanging the data on the x-axis.

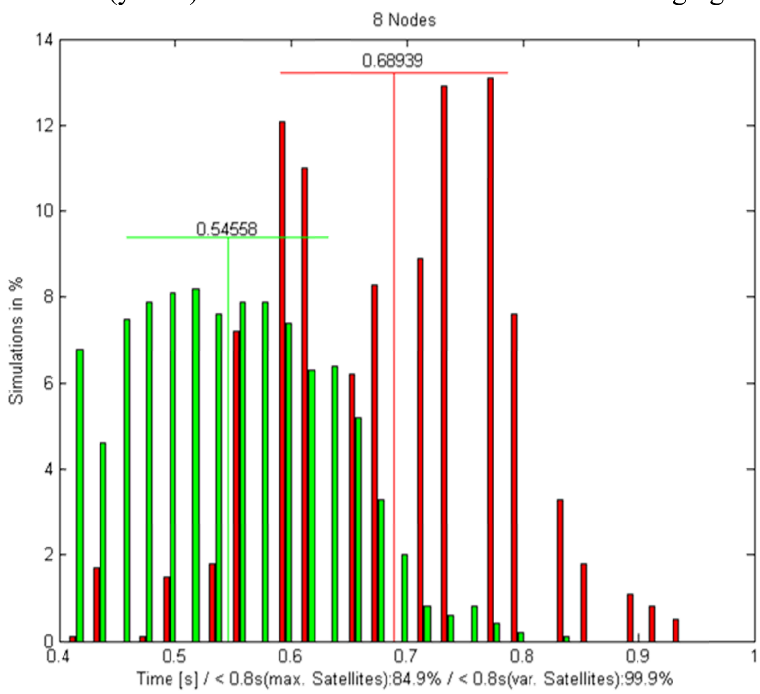


Figure 183: Simulation results for networks with eight nodes and fixed number of satellites (red bars) and variable number of satellites (green bars)

The calculated time includes the time for network discovery as well as for data exchange. Not included is the time which is needed for calculating the routing path. The red bars represents the results for the maximum number of satellites (twelve per node) and the green bars for a variable number of visible satellites which is between minimum four satellites and maximum twelve satellites. Aim is to exchange the data within 0.8 seconds to realise an update rate of 1 Hz and to have some backup time of about 0.2 seconds for data processing etc. As shown in the diagram in the mean 0.69 seconds are needed to exchange the GNSS raw data of 12 satellites visible for each node and 0.55 seconds for a variable number of satellites. Absolute 99.9 % of all simulated networks with the variable number of satellites are within 0.8 seconds and 85% of the networks with twelve visible satellites for each node, which represents the best case for positioning and the worst case for data exchange.

Figure 184 depicts the overall simulation results for networks with a size from four up to twelve nodes.

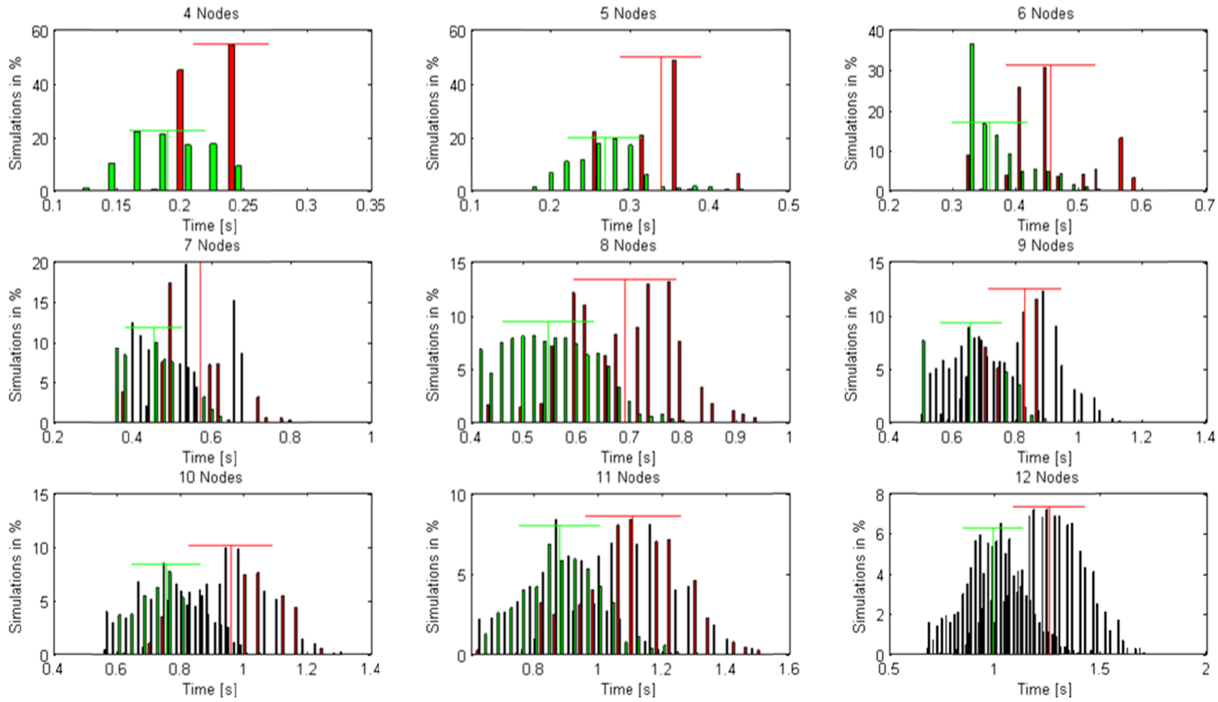


Figure 184: Simulation results for network size from 4 to 12 nodes

It is recognizable that the aim to exchange the data within 0.8 seconds is reachable for a network size of about ten nodes talking about the variable number of satellites and the average time and nine nodes with maximum number of satellites. Further steps to improve the performance are on the one hand to work on data compressing and on the other hand thinking about the need of exchanging the raw data of all participants in a big network or if it is more useful to exchange the raw data only within smaller sub networks.

8 CONCLUSIONS

The technological challenge of the approach presented in this article for relative positioning in machine swarms results in the first-time combination of all presented approaches in one system. In addition, further challenges are the technical limitations of the test vehicle as well as the requirement for implementation with affordable components, as the demand for low weight and low power consumption. Experiences from previous projects have demonstrated the suitability of the individual approaches. The presented hardware is successful integrated into the rovers and the first real tests of the standalone systems were successful.

ACKNOWLEDGEMENTS

The techniques and data shown have been developed within the joint research project "Navigation zur Exploration von tieffliegenden UAV in Katastrophenszenarien" (NEXT UAV). The joint project is funded by the Federal Ministry of Economics and Technology (BMWi) administered by the space agency of the German Aerospace Center (DLR) in Bonn (FKZ 50 NA 50 NA 1002 and 1003).

REFERENCES

SCHATTENBERG, J., ROBERT, M., HARMS, H.-H., LANG, T: *A flexible mobile platform for research on cooperating machinery*, 67th Conference Agricultural Engineering LAND.TECHNIK AgEng 2009, VDI-Verlag, Düsseldorf, p. 173-178, 2009.

SCHATTENBERG, J., DOPICHAY, T., LANG, T.*: comRoBS – Flexible Forschungsplattform für den kooperativen Maschinenbetrieb,* Landtechnik 65, I. 6, p.217-219, 2010

ADOLF, F., ANDERT, F., LORENZ, S., GOORMANN, L., DITTRICH, J.: *An unmanned helicopter for autonomous flights in urban terrain*, Advances in Robotics Research, T. Kröger und F. M. Wahl, ed. Berlin: Springer, P. 275-285, 2009.

BECKER, M.; BESTMANN, U.; SCHWITHAL, A.; HECKER, P.; BOTT, R.; KOHL, P.;KNEDLIK, S.; SCHEYER, H.; V. HINÜBER, E.: *Evaluation of a double equipped MEMS IMU based on real flight trial scenarios*, ION PLANS, 2010

STEEN. M.; HECKER, P.: *GBAS/INS Navigation System: Further Developments of On-Board Integrity Monitoring*, ION ITM, 2010

RTCA DO-229D: *Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment*, RTCA Inc., Washington D.C. Dezember 2006

NExt UAV Wissenschaftsprogramm, 3. Zwischenbericht, 1. Halbjahr 2011, BMWi (FKZ 50NA1002 and FKZ 50NA1003), Institut für Flugführung, Technische Universität Braunschweig and DLR Institut für Flugsystemtechnik, 2011

Combination of Feature-based and Geometric Methods for Positioning

Stefan Niedermayr, Andreas Wieser
Institute of Geodesy and Geophysics, Vienna University of Technology, Austria

Abstract
The location of mobile objects is usually based on angles, distances or distance differences, and on analytical geometry. It typically requires unobstructed line-of-sight between the mobile object and known reference points for "absolute positioning" or integrating measurements of position changes for "relative positioning". The main drawback of the relative methods is the rapid decrease of spatial accuracy over time and distance. The main drawback of the absolute methods is that a direct line-of-sight can often be established only with considerable effort or not at all e.g., in densely built-up areas or with distinctive topography.

We show that the established geometric methods can be augmented by feature-based positioning which is based on the comparison of observed feature vectors with given reference features. Feature-based positioning does not require any line-of-sight, yields "absolute" position with accuracy independent of time and distance travelled, and allows using the inherent location information of a variety of observable features. We give a brief overview of feature-based positioning techniques and present an approach using Bayesian estimation for the combination with classic geometric methods. The real data used for demonstration have been obtained at an airport where a combination of WLAN signal strength measurements and GPS-observations is used to locate vehicles for real-time display in an airport management system.

Keywords
Feature-based positioning, location fingerprinting, Particle Filter, WLAN, GNSS

1 INTRODUCTION

Global Navigation Satellite Systems (GNSS) are typically used for outdoor position determination if accuracies at the decimeter to centimeter level are sufficient. The raw measurements, namely pseudo-angle or carrier-phase, are corrected for clock errors, atmospheric delays and various other effects such that the reduced measurements represent the geometric distances between the satellites and the receiver. The position can be computed from these distances.

Also the raw measurements provided by other "positioning sensors" are normally reduced to their geometric content in terms of distance, distance difference or angle such that the coordinates can easily be computed using analytical geometry. Most geometry-based solutions require an unobstructed line-of-sight (LOS) between the mobile object and known reference points (or satellites, in the case of GNSS). If this line of sight is obstructed the corresponding measurement may not be available (see **Figure 30a**) or its accuracy may be significantly degraded. Consequently, such geometry-based solutions may fail in real-world situations or may only be applicable with considerable effort, e.g., by establishing auxiliary reference marks and carrying out additional measurements.

Geometric measurements representing position *changes* may be attainable without LOS requirement, e.g., using accelerations and rotation rate measurements from an inertial measurement unit (IMU). Based on such measurements, positions can be calculated sequentially by integrating the position changes. However, the error of such "dead reckoning" solutions usually grows rapidly with time or distance travelled. So, special actions or provisions like zero velocity updates or additional sensors yielding absolute position information once in a while (e.g., GNSS) are required to mitigate the inherent drift of relative positioning.

Feature-based positioning techniques represent a totally different approach to position determination. They yield the computed position by comparison of measured location-dependent features with given reference values associated with specific positions (see **Figure 30b**). Mathematically each feature represents a field $F(\mathbf{x})$, i.e., a function of position, and each measurement y_i relating to that feature represents a discrete sample of the field at a particular unknown position $\mathbf{x}_i$, contaminated by an unknown observation error e_i:

$$y_i = F(\mathbf{x}_i) + e_i . \tag{1.1}$$

The unknown position can be estimated by inversion, i.e.

$$\hat{\mathbf{x}}_i = F^{-1}(y_i - \tilde{e}_i) , \tag{1.2}$$

if a reasonable prediction $\tilde{e}_i$ of the observation error is available and if the field can be expressed by an invertible function. However, both will usually not be the case and thus the inversion and estimation of the error are carried out simultaneously by finding the position at which the difference between the value of the field and the actually observed value is a minimum:

$$\hat{\mathbf{x}}_i = \arg\min_{\mathbf{x}} \left| F(\mathbf{x}) - y_i \right|, \quad \tilde{e}_i = y_i - F(\hat{\mathbf{x}}_i) \tag{1.3}$$

Using a single scalar field, e.g., the power of the received signals emitted by a single WLAN access point, eq. (1.3) will usually not yield unique solution. However, if a suitable set of fields is observed simultaneously, e.g., the signal power of several WLAN access points, the actual position can usually be estimated. The reference values i.e., the representation of the fields, can either be derived in advance from georeferenced measurements taken during a mapping phase and stored in a database, or from numerical models.

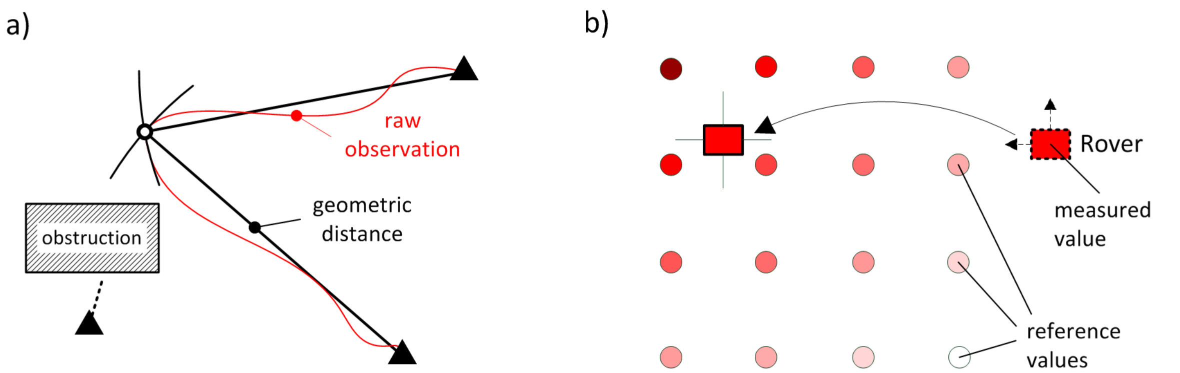


Figure 185: Geometry-based positioning requiring LOS-observations (a), feature-based positioning requiring observable feature with spatial variability (b).

These techniques have become known as “location fingerprinting”. Their most significant advantages are that they do not require LOS observations or reduction of the measurements to their geometric content, and that they may not require dedicated installation of positioning infrastructure but may rather use so-called “signals of opportunity” for positioning, i.e. signals which are available for other purposes anyway like signals used for wireless communication or broadcasting. The most significant disadvantages are that the attainable accuracy is typically low (several meters) and that it may be difficult to obtain, handle and maintain the required reference data.

While location fingerprinting has been discussed extensively in the context of positioning using radio signal strength measurements, see e.g., Bahl and Padmanabhan (2000), Bensky (2008), it need not be restricted to the corresponding type of fields. Any measurable feature which varies with location but is fairly stable or predictable over time can be used for fingerprinting. Even the usual geometric observations (e.g., distances, signal travel times, or angles) can be treated as samples of fields and can be processed using fingerprinting instead of exploiting their geometric content directly. This may facilitate position determination with cm-level accuracy even if the measurements are non-LOS, affected by multipath signal propagation or affected by other systematic effects which can hardly be modeled. In order to emphasize that any type of spatially varying features can be used for positioning we will use the term *feature-based positioning* in this paper.

In view of the respective advantages of the geometry-based and the feature-based approaches we think that a combination of them will be beneficial for a variety of applications. Thus we discuss this combination herein. In section 2 we give a brief introduction to feature-based positioning, in section 3 we discuss the combination of observed features and geometric observations using a probabilistic method, and in section 4 we present results for a tight coupling of WLAN signal strengths and GPS observations for positioning on the apron of a regional airport.

2 FEATURE-BASED POSITIONING

2.1 Feature fields and their representation

A variety of observable features can be used for positioning. Christ et al. (1993) use signal strengths of a proprietary radio frequency infrastructure; Bahl and Padmanabhan (2000) use WLAN signal strengths corresponding to several access points; Storms et al. (2010) use the 3-dimensional magnetic field observable within buildings. Ambient light intensity, ambient light color, ambient noise, and spatial or radiometric variability of the surrounding environment are further examples of fields which may be useful for positioning. As mentioned above, even distances or angles can be interpreted and used as location dependent features. An example can be found in Nygren (2008) where the seafloor topography is used as field and sonar beams are used to sample this topography for navigating underwater vehicles.

The features need to fulfill the following requirements:

The field varies significantly with varying location (the spatial gradient should be high).

The field is constant in time or its temporal variation is predictable.

The feature corresponding to the field is observable and can be uniquely quantified.

Furthermore, the following properties are desirable:

The field is readily available or can be produced and maintained with reasonable effort.

A given combination of feature values should be uniquely associated with a particular location.

If the field is produced artificially, it does not harm humans, animals or equipment.

It may be difficult to find individual features or fields which fulfill all the above requirements. However, we are able to combine several different features – including features of different type – into *feature vectors* $\mathbf{y}$ representing a multi-dimensional vector field which does fulfill the requirements. So we will subsequently assume that feature-based positioning is carried out by comparing observed feature vectors $\mathbf{y}_i$ with vectors $\tilde{\mathbf{y}}(\mathbf{x})$ (reference values) predicted from a corresponding reference field.

The reference field can either be derived by mapping i.e., by measuring the features at distinct known locations $\mathbf{x}_j$ in advance, or by evaluation of numerical models (Bahl and Padmanabhan 2000). The latter approach does not require mapping in advance and is thus less expensive, can easily take changing surrounding and environmental conditions into account and does not suffer from measurement errors during the mapping phase. However, accurate modeling of physical fields in real environment may be very challenging and thus the accuracy of the reference features derived from models instead of measurements may be insufficient. We thus focus on mapping herein; more information on the modeling approach can be found e.g., in Bahl and Padmanabhan (2000).

Reference values can be determined by carrying out measurements at positions corresponding to a regular grid. The continuous field is thus represented by gridded data. A slightly different approach is to record feature values while moving and to determine the corresponding position continuously using a kinematic measurement system. The reference values to be used later are derived using smoothing and 2D-interpolation see e.g., Duvallet and Tews (2008). The continuous field is then represented by the parameters (coefficients) of the selected interpolation. We have chosen a least square interpolation for the results presented in section 4.

Once the reference field is represented by a database or by a set of parameters such that the reference values $\tilde{\mathbf{y}}(\mathbf{x})$ can be calculated for arbitrary positions $\mathbf{x}$, the measurements are compared to the reference values either using deterministic (see sec. 2.2) or probabilistic methods (see sec. 2.3). An outline of various methods is given by Honkavirta et al. (2009).

2.2 Deterministic Fingerprinting

If the reference values of the observed multidimensional field are given as a finite set

$$\mathfrak{R} = \{\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \ldots, \tilde{\mathbf{y}}_N\}, \quad \tilde{\mathbf{y}}_j := \tilde{\mathbf{y}}(\mathbf{x}_j) \tag{2.1}$$

of N feature vectors each of them associated with a specific position $\mathbf{x}_j$, deterministic fingerprinting can be used to estimate the position $\mathbf{x}$ at which the feature vector $\mathbf{y}$ has been observed. The procedures are based on the “distance” between the observed feature vector and the given ones:

$$d_j := \left\| \tilde{\mathbf{y}}_j - \mathbf{y} \right\|, \tag{2.2}$$

where a variety of distance measures can be used, see Honkavirta et al. (2009). However, given the potentially different nature and scale of the individual elements of the feature vector we consider the Mahalanobis distance to be most appropriate. A well established and simple determinstic method is the nearest neighbor method described e.g., by Bahl and Padmanabhan (2000)

$$\hat{\mathbf{x}} = \mathbf{x}_k \quad \text{with} \quad d_k = \left\| \tilde{\mathbf{y}}(\mathbf{x}_k) - \mathbf{y} \right\| = \min_j \left\| \tilde{\mathbf{y}}_j - \mathbf{y} \right\|. \tag{2.3}$$

It is computationally simple but not very accurate. More sophisticated deterministic methods involve calculating the mean of a subset of the k reference vectors closest to the observed one (k-nearest neighbor), a weighted mean of that subset (weighted k-nearest neighbor) or yet another function of the given reference positions, see Honkavirta et al. (2009). However, if the dimension of the vector is high then the algorithms of type nearest neighbor may fail to produce a useful solution (Beyer et al., 1999).

We do not discuss the deterministic methods further because the probabilistic methods are superior with respect to handling the measurement uncertainty and with respect to combination of the feature vectors with prior knowledge and additional information.

2.3 Probabilistic Fingerprinting

Probabilistic fingerprinting uses the Bayes theorem

$$p(\mathbf{x} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{x}) \cdot p(\mathbf{x})}{p(\mathbf{y})} \tag{2.4}$$

(see e.g., Koch, 2000; Roos et al., 2002) for estimating the position $\mathbf{x}$ which is now considered a *random quantity*. The posterior distribution $p(\mathbf{x}|\mathbf{y})$ of the coordinates is calculated from the given observations $\mathbf{y}$, the prior distribution $p(\mathbf{x})$, the likelihood $p(\mathbf{y}|\mathbf{x})$ of the observations and a normalization quantity $p(\mathbf{y})$ which assures that the integral over the range of all admissible positions $\mathbf{x}$ is equal to 1. If no prior information on the actual position is available $p(\mathbf{x})$ is a constant representing a uniform distribution.

The likelihood can be evaluated using various approaches, see Roos et al. (2002), Honkavirta et al. (2009). If the observations are normally distributed, the multivariate normal distribution of the observations $\mathbf{y}$ given the respective reference vector can be used to calculate the value of the likelihood function at each possible position $\tilde{\mathbf{x}}_l$:

$$p(\mathbf{y} \mid \tilde{\mathbf{x}}_l) = p(\mathbf{y} \mid \tilde{\mathbf{y}}_l, \mathbf{\Sigma}_l) = (2\pi)^{-n/2} (\det \mathbf{\Sigma}_l)^{-1/2} \exp\left(-\frac{1}{2} (\mathbf{y} - \tilde{\mathbf{y}}_l)^{\mathrm{T}} \mathbf{\Sigma}_l^{-1} (\mathbf{y} - \tilde{\mathbf{y}}_l) \right). \tag{2.5}$$

$\mathbf{\Sigma}_l$ is the covariance matrix of the reference vector $\tilde{\mathbf{y}}_l$ i.e. of the predicted value of the feature vector at position $\tilde{\mathbf{x}}_l$, and n is the dimension of the feature vector.

The evaluation of the above equations for estimating the position is straightforward if there is a finite and rather small number M of possible positions (e.g., a few thousand grid points or particles, see sec. 3). In this case we have

$$p(\mathbf{y}) = \sum_{l=1}^{M} p(\mathbf{y} \mid \tilde{\mathbf{x}}_l) p(\tilde{\mathbf{x}}_l) \tag{2.6}$$

and the unknown position can be estimated using the maximum-posterior principle as discussed e.g., by Roos et al. (2002):

$$\hat{\mathbf{x}} = \tilde{\mathbf{x}}_k \quad \text{with} \quad p(\tilde{\mathbf{x}}_k \mid \mathbf{y}) = \max_l p(\tilde{\mathbf{x}}_l \mid \mathbf{y}) \tag{2.7}$$

The availability of the posterior distribution within this Bayesian framework is a key to quality control (e.g., by examining a corresponding confidence region). This is a distinct advantage with respect to deterministic fingerprinting. Furthermore, the tight coupling of geometric observations can easily be accomplished within the same framework as we will show.

3 COMBINATION OF METHODS

Geometric observations like distances or angles are typically processed using analytical geometry. For example, distance resection (e.g., Uren and Price, 2010) is carried out by calculating the angles within a planar triangle from the chord lengths between the three corners and using these angles and chord lengths to calculate the new coordinates of one of the points. Often, redundant measurements are taken in order to control and possibly enhance quality. In this case parameter estimation is employed to derive a unique and in a chosen sense optimum solution from an overdetermined and inconsistent system of equations. A common solution is the weighted least-squares estimate within a Gauß-Markov model or an equivalent best linear unbiased estimate of the fixed but unknown coordinates, see e.g., Koch (1999), Grafarend and Schaffrin (1993).

An alternative is to consider the unknown coordinates as stochastic parameters and estimate them using Bayesian estimation as outlined above. The only differences with respect to probabilistic fingerprinting are that $\mathbf{y}$ now represents the observed distances and angles, and $\tilde{\mathbf{y}}_l$ the distances and angles computed from the respective candidate positions $\tilde{\mathbf{x}}_l$ using analytical geometry. This process does not require linearization of the observation equations (as opposed to the linear estimation mentioned above), and allows easily combining geometric observations with observed feature vectors by sequential application of Bayes' theorem. If the geometric observations and the observed features are uncorrelated, this corresponds to simple multiplication of the respective likelihoods, see Koch (2000).

We demonstrate Bayesian processing of geometric observations using a simulated distance resection problem. The coordinates of two anchors (fixed points) P_1 and P_2 are given ([–2.5 m; 0.0 m], [10.0 m; 0.0 m]) in a local coordinate system. The distances from an unknown rover position to these two anchors have been observed ($y_1 = 7.8$ m; $y_2 = 5.0$ m) with different standard deviation ($\sigma_1 = 0.3$ m; $\sigma_2 = 0.2$ m). The possible rover positions are represented by a regular grid (0.04×0.04 m^2) covering the region of interest. Repeatedly using eq. (2.5) the likelihood of each of the two distance observations is computed for all grid points.

Figure 186 (left) shows the likelihood values for one of the distance observations and all grid points. We can clearly see a circular region of (equal) maximum likelihood corresponding to the circle of radius y_1 about P_1. This is the circle defined by the measurement value and the anchor point. The region of decreasing likelihood spreading radially about the circle shows the statistical uncertainty of the measurement and represents the fact that the true rover position need not coincide exactly with the circle defined by the observed value of the distance.

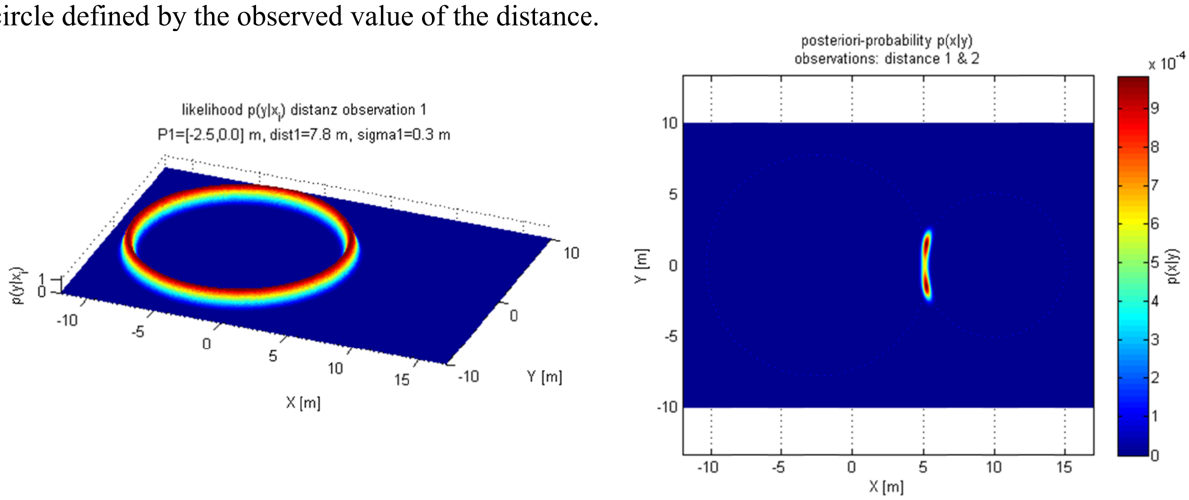


Figure 186: Numeric example of distance resection with Bayes theorem; likelihood of one of the observations (left), posterior distribution of the position (right).

The likelihood corresponding to the second observation is not shown here. Instead, the posterior density of the rover position is shown in **Figure 186** (right). It has been obtained by multiplying the likelihoods of the two uncorrelated observations, assuming constant priori probability, and scaling with a normalization factor as given in eq. (2.6). The figure shows the ambiguity of the non-overdetermined distance resection: there are two points of (equal) maximum posterior density. The coordinates of these points are equal within grid size to those obtained with analytical geometry. The figure also shows that the spatial distribution of the uncertainty is not elliptic (this is visible because the standard deviation of the distance observations is large, and because the solution is based on the correct non-linear model rather than a linearized one).

If the rover moves the accuracy of the estimate can be further improved by taking the dynamics of the movement into account. In particular, the coordinates can be part of a "state vector" which may also include velocity, sensor biases and other quantities describing the rover and the measurement systems. Particle filters, see Simon (2006), Thrun et al. (2006), are numeric implementations of sequential Bayesian estimation which are particularly useful for combining geometric and feature-based positioning for moving objects.

Within a particle filter, the above distributions (prior, posterior) are used to create particles i.e., quasi-random samples of the state vector. These particles then represent the corresponding distributions and can be used for state prediction, computation of the likelihood and for state estimation. Particle filters belong to the class of sequential Monte Carlo methods. The computational steps of the particle filter are summarized e.g., in Simon (2006) p. 468.

Bayesian estimation and a particle filter in particular are well suited for estimating the position of moving objects using observed features. Siddiqi et al. (2003) use a particle filter for Wi-Fi based localization of robots, Duvallet and Tews (2008) use a particle filter for Wi-Fi based navigation of an autonomous vehicle within an industrial environment. We give an example of the combination of observed features and geometric observations within a particle filter in the next section.

4 APPLICATION EXAMPLE

4.1 Testbed

Within the research project SESAAM (Geo-Spatially Enhanced Situational Awareness for Airport Management) several academic and industrial partners cooperate to develop methods for (i) determining the position and motion of all vehicles and movable objects at the apron of an airport in real time, (ii) combine this information with additional sensor data, e.g., on metrology, to derive a complete operational picture (COP) of the apron, and (iii) to visualize this COP in an appropriate and individual way for each actor/entity involved in airport operation and apron activities (Bretz at al., 2011).

A testbed for data collection and experimental investigations has been established as part of SESAAM at Salzburg airport. Several vehicles have been equipped with a linux computer, a ublox LEA 6T GPS receiver with low-cost patch antenna, and two WLAN adapters. One of these adapters is used to scan the signals of all available access points continuously and record the corresponding signal strengths; the other one transmits all collected raw data (WLAN signal strengths; GPS pseudoranges, Doppler, carrier phase and C/N_0 measurements) to a location server (VOS) via the airport LAN. Additionally, GPS data from a nearby fixed reference station are collected and transmitted to the VOS. The VOS estimates the position and motion of the vehicles and transmits the results to the COP server where further geo-processing and visualization are carried out. The inclusion of other moving objects relies on additional measurements but is beyond the scope of this paper and thus not discussed herein.

4.2 Tight coupling of WLAN and GPS

A straightforward way to estimate the position of the vehicles would be to use the recorded WLAN signal strengths for location fingerprinting, and the GPS observations for calculating position and velocity. Then the final estimates would be calculated from these separate results e.g., by weighted averaging. The disadvantage of this loose coupling of WLAN and GPS is that each of the systems needs to provide a sufficiently accurate position solution independently. There is no benefit from the combination for periods with less than 4 visible GPS satellites or with non-unique location fingerprinting result. Even worse if outliers within the signal strength measurements or the GPS measurements contaminate the corresponding individual solution the entire GPS or WLAN information for that epoch needs to be discarded. So, the coupling is not efficient.

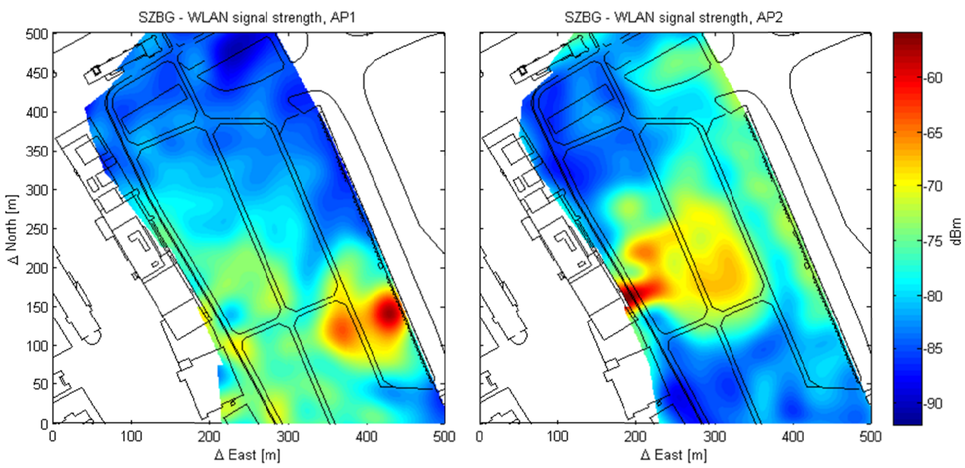


Figure 187: WLAN signal strength reference field associated with two arbitrarily selected access points

We propose tight coupling of the raw observations instead, i.e., the combination at the observation level. This allows for better quality control and higher availability of solutions. The WLAN signal strengths are now treated as features in the above sense and are processed within a particle filter together with the raw GPS observations in order to estimate the position and velocity of the vehicle. For the numeric example presented in this section, the required reference fields comprising signal strengths of about 15 access points and covering a large part of the apron are stored in a database for a dense regular coordinate grid (5×5 m). They have been established by kinematic mapping and subsequent interpolation (see sec. 2.1). As an example, **Figure 187** shows two layers of this multi-dimensional reference field namely the signal strengths associated with two different access points.

The state vector within the particle filter comprises the 3D-position and velocity of the vehicle, and the clock drift of the GPS receiver. The system dynamics are modeled using a second order model with random walk velocity (spectral noise densities chosen as North: 1.5 m^2/s^3, East: 1.5 m^2/s^3, Up 0.1 m^2/s^3). Double differenced pseudorange observations (see e.g., Kaplan and Hegarty, 2006) and undifferenced Doppler observations make up the GPS observation vector $\mathbf{y}^{GPS}$. These observations are assumed to be normally distributed and their variance-covariance matrix is computed using elevation dependent variances. The GPS observation equations (including their geometric content) are used to calculate the predicted observation vector $\tilde{\mathbf{y}}_{(i)}^{GPS}$ for each of the $N = 10.000$ particles. The recorded signal strengths of the access points "visible" at the respective rover position are collected in the WLAN observation vector $\mathbf{y}^{WLAN}$, and the reference values $\tilde{\mathbf{y}}_{(i)}^{WLAN}$ corresponding to the individual particles are computed from the previously established reference field. The signal strength observations are assumed to be independent and normally distributed with a variance of 9 dBm^2 such that their variance-covariance matrix is a scaled identity matrix. The entire observation vector **y**, its variance-covariance matrix, and the predicted observations corresponding to the individual particles are then obtained by stacking these data.

For this demonstration only a single short dataset of GPS and WLAN observations (5 hours) was available, and this set had already been used to derive the reference fields. Thus the numeric results presented subsequently rely partly on simulated observations and are given here only for demonstration of the feasibility of the proposed method. In particular the signal strength measurements associated with 12 access points have been computed using the reference data and normal random noise has been added. The resulting simulated WLAN signal strength measurements have then been processed along with subsets of the real GPS pseudorange and Doppler data, using the particle filter. An investigation based entirely on real measurements and involving independent ground truth is currently carried out and will be presented elsewhere.

The results of processing 5 minutes of data with a vehicle moving on the apron are shown in **Figure 188**. The vehicle starts at position A. A GPS outage of about 1 minute has been simulated between positions B and C by deleting the corresponding observations before processing. During 30% of the outage, less than 4 GPS satellites were still used, during the remaining 60% no satellites were used. At C the vehicle stops for about 1 minute and then continues to position D where the selected dataset ends.

The filled dots represent the estimated position involving WLAN and more than 4 GPS satellites. The circles represent the solutions during GPS outages i.e., periods with less than 4 GPS satellites. During the outages the position is mainly or exclusively determined by the signal strength measurements and the prediction within the particle filter. The estimated position during the outages deviates only by a few meters from the respective solution obtained using all GPS data (crosses). The largest deviation occurs after the vehicle has stopped at C and is due to the filter settings and design (quasi-constant velocity); an adaptive filter would significantly improve that situation. The grey shading shows the areas covered by the particles of all epochs and is thus a first visualization of the uncertainty of the solution.

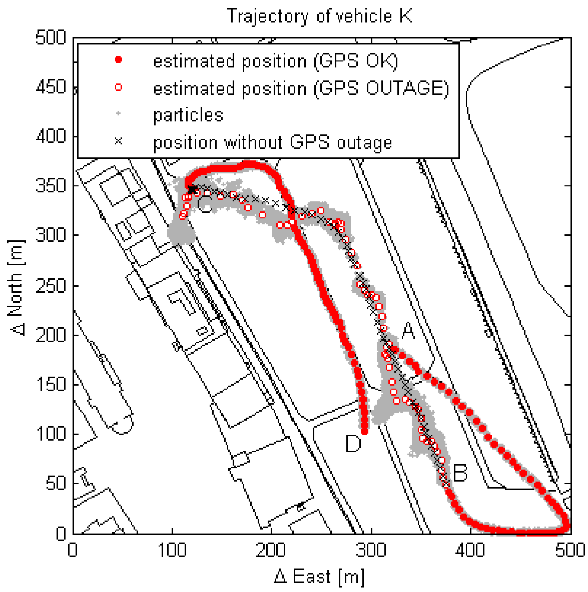


Figure 188: Estimated trajectory of a vehicle (car) during 5 minutes of motion on the apron

These numeric results show that the tight coupling of GPS and feature-based positioning within a particle filter is feasible and will likely help to maintain an accuracy of a few meters even during partial or full GPS outages. As opposed to dead reckoning, the uncertainty of the position estimated with the feature-based approach does not grow steadily during such periods.

5 CONCLUSIONS

We have discussed the advantages and disadvantages of feature-based and geometric positioning and have proposed a tight coupling of both using Bayesian estimation. The numeric example used to illustrate the approach involved WLAN signal strength measurements which may not facilitate sub-meter level accuracies. However, there is a variety of other features, like radio signal travel times (including ultra-wideband signal travel times), which may be exploited using the proposed approach in order to derive position estimates with accuracies of a few centimeters even in environment where geometric positioning alone cannot be applied efficiently because of limited line-of-sight availability.

Further research is required in relation to robustness of the proposed approach, i.e., to outlier detection, identification and mitigation, to adequate spatial representation of the reference features, and to dynamic modeling of the reference features.

ACKNOWLEDGEMENTS

This work has been partially supported by the Austrian research funding society (FFG) under research grant 825613 (SESAAM).

REFERENCES

BAHL, P., PADMANABHAN, V. N.: *RADAR: an in-building RF-based user location and tracking system*. INFOCOM 2000, Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 2, Tel-Aviv, Mar 26-30, p. 775-784, 2000.

BENSKY, A.: *Wireless Positioning Technologies and Applications*. Artech House, Norwood, 2008.

BEYER, K., GOLDSTEIN, J., RAMAKRISHNAN, R., SHAFT, U.: *When is "Nearest Neighbor" Meaningful?* In Proc 7th Intl. Conf. on Database Theory, Jerusalem, Jan. 10-12, 1999, p. 217-235, 2009.

CHRIST, T.W., GODWIN, P.A., LAVIGNE, R.E.: *A prison guard Duress alarm location system*. International Carnahan Conference on Security Technology, Ottawa, Canada, Oct 13-15, p. 106-116, 1993.

BRETZ, B., KAPSER, J., ROTH, C., NIEDERMAYR, S.: *Development of a "complete operational picture" using standard IT-infrastructure: Progress within the FFG-project SESAAM*. Proceedings of 23rd AGIT Symposium, Salzburg, Jul 6-8, p. 670-675, 2011 (in German).

DUVALLET, F., TEWS, A. D.: *WiFi position estimation in industrial environments using Gaussian processes*. IROS 2008, IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, Sep 22-26, Month days, p. 2216-2221, 2008.

GRAFAREND, E., SCHAFFRIN, B.: *Regression in linear models.* BI-Wiss.-Verl. 496 p., 1993 (in German).

HONKAVIRTA, V., PERALA, T., ALI-LOYTTY, S., PICHÉ, R.: *A comparative survey of WLAN location fingerprinting methods*. WPNC 2009, 6th Workshop on Positioning, Navigation and Communication, Hannover, Mar 19, p. 243-251, 2009.

KAPLAN, E. C., HEGARTY C. J.: *Understanding GPS: principles and applications.* Artech House, Norwood, 2006.

KOCH, K. R.: *Parameter Estimation and Hypothesis Testing in Linear Models*. 2nd ed., Springer, Berlin Heidelberg, p. 333, 1999.

KOCH, K. R.: *Introduction into Bayes-Statistics*. Springer, Berlin Heidelberg, 2000 (in German).
NYGREN, I., *Robust and efficient terrain navigation of underwater vehicles.* IEEE/ION 2008, Position, Location and Navigation Symposium, Monterey, May 5-8, p. 923-932, 2008.

RISTIC, B., ARULAMPALAM, S., GORDON, N.: *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, Norwood, 2004.

ROOS, T., MYLLYMÄKI, P., TIRRI, H., MISIKANGAS, P., SIEVÄNEN, J.: *A Probabilistic Approach to WLAN User Location Estimation.* International Journal of Wireless Information networks, vol. 9, p. 155-164, 2002.

SIDDIQI, S., SUKHATME, G.S., HOWARD, A.: *Experiments in Monte-Carlo Localization using WiFi Signal Strength.* Proceedings of the International Conference on Advanced Robotics (ICAR), Coimba, Jun 30 - Jul 3, 2003.

SIMON, D.: *Optimal State Estimation.* John Wiley & Sons, United States of America, 2006.

STORMS, W., SHOCKLEY, J., RAQUET, J.: *Magnetic Field Navigation in an Indoor Environment.* Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), Kirkkonummi, Oct 14-15, p. 10, 2010.

THRUN, S., BURGARD, W., FOX, D.: *Probabilistic Robotics.* MIT Press, Cambridge, 2006.
UREN, J., PRICE, B.: *Surveying for Engineers*. 5th ed., Palgrave Macmillan, p. 816., 2010.

List of authors

Jürgen Alberding
Alberding GmbH
Lilienthalstr. 25
D-12529 Schönefeld, Germany
E-Mail: alberding@alberding.eu

Bara' Al-mistarehi
Institute of Engineering Geodesy
University of Stuttgart
Geschwister-Scholl Str. 24D
D-70174 Stuttgart, Germany
E-Mail:bara.al-mistarehi@ingeo.uni-stuttgart.de

Alexander Beetz
Institute of Engineering Geodesy
University of Stuttgart
Geschwister-Scholl Str. 24D
D-70174 Stuttgart, Germany
E-Mail: alexander.beetz@ingeo.uni-stuttgart.de

Hellmut Billinger
J.- H. Priester Str. 31
D-74564 Crailsheim, Germany
E-Mail: hbillinger@gmx.de

Stefan Böttinger
Institute of Agricultural Engineering
University of Hohenheim
Garbenstr. 9, Institutsgebäude u. UBA, 109/2
D-70599 Stuttgart, Germany
E-Mail: boettinger@uni-hohenheim.de

Aurelian Stelian Buda
Cadastre and Architecture
University of Oradea
O. Ghibu No.28
Oradea, Romania
E-Mail: budaaurelian@yahoo.ro

Pierre Cartade
CEMAGREF / IRSTEA
Dept. TSCA
24, avenue des Landais
B.P. 50085
63172 Aubiere Cedex, France
E-Mail : pierre.cartade@irstea.fr

Eric Durand
CATERPILLAR
44 rue du baricot
33170 Gradignan, France
E-Mail: durand_eric@cat.com

Christian Eling
Institute of Geodesy and Geoinformation
University of Bonn
Nussallee 17
D-53115 Bonn, Germany

E-Mail: eling@igg.uni-bonn.de

Jürgen Frick
MPA Stuttgart
Fraunhofer FHR
Pfaffenwaldring 32
D-70569 Stuttgart, Germany
E-Mail:juergen.frick@mpa.uni-stuttgart.de

Ilka von Gösseln
Geodetic Institute
Leibniz Universität Hannover
Nienburger Str. 1
D-30167 Hannover, Germany
E-Mail: vongoesseln@gih.uni-hannover.de

María A. González Huici
Fraunhofer FHR
Neuenahrer Str. 20
D-53343 Wachtberg, Germany
E-Mail: maria.gonzalez@fhr.fraunhofer.de

Tianyi Guan
Fraunhofer IOSB
Fraunhoferstraße 1
D-76131 Karlsruhe, Germany
E-Mail: tianyi.guan@iosb.fraunhofer.de

Stefan Hauth
i3mainz - Institute for Spatial Information
and Surveying Technology
University of Applied Sciences Mainz
Lucy-Hillebrand-Straße 2
D-55128 Mainz, Germany
E-Mail: stefan.hauth@geoinform.fh-mainz.de

Otto Heunecke
Institute of Geodesy
University of the Federal Armed Forces Munich
Werner-Heisenberg-Weg 39
D-85577 Neubiberg, Germany
E-Mail: otto.heunecke@unib.de

Sabine Hofmann
Institute of Cartography and Geoinformatics
Leibniz University Hannover
Appelstr. 9A
D-30167 Hannover, Germany
E-Mail: sabine.hofmann@ikg.uni-hannover.de

Sebastian Horst
Geodetic Institute
Leibniz Universität Hannover
Nienburger Str. 1
D-30167 Hannover, Germany
E-Mail: horst@gih.uni-hannover.de

Jan Horstmann
Maschinenfabrik Bernard KRONE GmbH
Heinrich-KRONE-Str. 10
D-48480 Spelle, Germany
E-Mail: jan.horstmann@krone.de

Andreea Florina Jocea
Faculty of Geodesy
Technical University of Civil Engineering
of Bucharest
Bdul. Lacul Tei, nr. 124, sector 2
RO-020396, Bucharest, Romania
E-Mail: andreea.jocea@geodezie.utcb.ro

Klaus Kittmann
Institute of Aircraft Design
University of Stuttgart
Pfaffenwaldring 31
D-70569 Stuttgart, Germany
E-Mail: kittmann@ifb.uni-stuttgart.de

Hiro Konno
Engineering Department
Topcon Positioning Systems, Inc.
7400 National Drive, Livermore
CA 94550, U.S.A.
E-Mail: hkonno@topcon.com

Johanna Link-Dolezal
Insitute of Crop Science
University Hohenheim
Fruwirthstr. 23
D-70599 Stuttgart, Germany
E-Mail: johanna.link@uni-hohenheim.de

Andreas Linz
Faculty of Engineering and Computer Science
University of Applied Sciences Osnabrück
Albrechtstr. 30
D-49076 Osnabrück, Germany
E-Mail: a.linz@hs-osnabrueck.de

Norbert Mattivi
Trimble Germany GmbH
Am Prime Parc 11
D-65497 Raunheim, Germany
E-Mail: norbert_mattivi@trimble.com

Xiaolin Meng
Nottingham Geospatial Institute
University of Nottingham, Innovation Park
Triumph Road, Nottingham,
NG7 2TU, United Kingdom
E-Mail: xiaolin.meng@nottingham.ac.uk

Tsuneo Nakanishi
Faculty of Information Science and Elect
Kyushu University
744 Motooka, Nishi
819-0395 Fukuoka, Japan
E-Mail: tsuneo.nakanishi@gmail.com

Stefan Niedermayr
Institute of Geodesy and Geophysics
Vienna University of Technology
Gußhausstraße 27-29/E128-3
A-1040 Wien, Austria
E-Mail: stefan.niedermayr@tuwien.ac.at

Mohammed Taleb Obaidat
Civil Engineering Department
Jordan University of Science and Technology
Irbid - P.O
3030, Jordan
E-Mail: mobaidat@just.edu.jo

Heinz-Jürgen Przybilla
Department of Surveying and Geomatics
Lab. for Photogrammetry
Bochum University of Applied Sciences
Lennershofstr. 140
D-44801 Bochum, Germany
E-Mail: heinz-juergen.przybilla@hs-bochum.de

Jürgen Reineke
Prolec Ltd., Central Europe Office
An der Eichfeldstraße 10
D-32257 Bünde, Germany
E-Mail: jreineke@prolec.co.uk

Mathieu Richier
CEMAGREF / IRSTEA
Dept. TSCA
24, avenue des Landais
B. P. 50085
63172 Aubiere Cedes, France
E-Mail : mathieu.richier@irstea.fr

Jan Schattenberg
Institute of Agricultural Machinery and Fluid Power
Technische Universität Braunschweig
Langer Kamp 19a
D-38106 Braunschweig, Germany
E-Mail: j.schattenberg@tu-braunschweig.de

Annette Scheider
Institute of Engineering Geodesy
University of Stuttgart
Geschwister-Scholl-Str. 24D
D-70174 Stuttgart, Germany
E-Mail: annette.scheider@ingeo.uni-stuttgart.de

Michael Schulz
ALLSAT GmbH
Am Hohen Ufer 3A
D-30159 Hannover, Germany
E-Mail: michael.schulz@allsat.de

Jürgen Schweitzer
Institute of Engineering Geodesy
University of Stuttgart
Geschwister-Scholl Str. 24D
D-70174 Stuttgart, Germany
E-Mail: juergen.schweitzer@ingeo.uni-stuttgart.de

Volker Schwieger
Institute of Engineering Geodesy
University of Stuttgart
Geschwister-Scholl Str. 24D
D-70174 Stuttgart, Germany
E-Mail: volker.schwieger@ingeo.uni-stuttgart.de

Andrea Stürze
Federal Agency for Cartography and Geodesy
Richard-Strauss-Allee 11
D-60598 Frankfurt am Main, Germany
E-Mail: andrea.stuerze@bkg.bund.de

Zhenzhong Su
Intitute of Geodesy and Photogrammetry
ETH Zürich
Schafmattstrasse 34
CH-8093 Zürich, Switzerland
E-Mail: zhenzhong.su@geod.baug.ethz.ch

Ulrich Völter
intermetric GmbH
Industriestraße 24
D-70565 Stuttgart, Germany
E-Mail: ulrich.voelter@intermetric.de

Jens Wickert
Deutsches GeoForschungsZentrum GFZ
Telegraphenberg A 17
D-14473 Potsdam, Germany
E-Mail: wickert@gfz-potsdam.de

Andreas Wieser
Institute of Geodesy and Geophysics,
Vienna University of Technology
Gußhausstraße 27-29/E128-3
A-1040 Wien, Austria
E-Mail: andreas.wieser@tuwien.ac.at