Martti Kirkko-Jaakkola
Contributions to Positioning Methods on Low-Cost Devices

Julkaisu 1148 • Publication 1148

Tampere 2013
Martti Kirkko-Jaakkola

Contributions to Positioning Methods on Low-Cost Devices

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 30th of August 2013, at 12 noon.
ABSTRACT

Global Navigation Satellite System (GNSS) receivers are common in modern consumer devices that make use of position information, e.g., smartphones and personal navigation assistants. With a GNSS receiver, a position solution with an accuracy in the order of five meters is usually available if the reception conditions are benign, but the performance degrades rapidly in less favorable environments and, on the other hand, a better accuracy would be beneficial in some applications.

This thesis studies advanced methods for processing the measurements of low-cost devices that can be used for improving the positioning performance. The focus is on GNSS receivers and microelectromechanical (MEMS) inertial sensors which have become common in mobile devices such as smartphones. First, methods to compensate for the additive bias of a MEMS gyroscope are investigated. Both physical slewing of the sensor and mathematical modeling of the bias instability process are considered. The use of MEMS inertial sensors for pedestrian navigation indoors is studied in the context of map matching using a particle filter. A high-sensitivity GNSS receiver is used to produce coarse initialization information for the filter to decrease the computational burden without the need to exploit local building infrastructure. Finally, a cycle slip detection scheme for stand-alone single-frequency GNSS receivers is proposed.

Experimental results show that even a MEMS gyroscope can reach an accuracy suitable for North seeking if the measurement errors are carefully modeled and eliminated. Furthermore, it is seen that even a relatively coarse initialization can be adequate for long-term indoor navigation without an excessive computational burden if a detailed map is available. The cycle slip detection results suggest that even small cycle slips can be detected with mass-market GNSS receivers, but the detection rate needs to be improved.
This thesis is based on research conducted at the Department of Computer Systems, Tampere University of Technology (TUT), in 2009–2012. The research has been funded by the Doctoral Programme of TUT’s President and by the European Space Agency (ESA), under the NPI programme, which are gratefully acknowledged as well as the additional financial support from Ulla Tuomisen säätiö, Jenny ja Antti Wihurin rahasto, A. R. Winterin muistosäätiö, HPY:n tutkimussäätiö, and Tampere Doctoral Programme in Information Science and Engineering.

First of all, I would like to thank my supervisor, Professor Jarmo Takala, for the opportunity to carry out this research at the department and for his professional insight and advice that are needed in order to get the articles published. I am equally grateful to my instructor, Dr. Jussi Collin, whose seemingly endless supply of ideas has been a significant factor for this research. I would also like to thank the rest of our research group for fruitful conversations and code exchange during these years. Furthermore, Dr. Thomas Pany and Dr. Jari Saarinen deserve my gratitude for their efforts in reviewing this thesis and for the constructive feedback.

I have had the opportunity to spend a significant amount of time abroad during this research work. I would like to thank Prof. Dr.-Ing. Florian Holzapfel and Dr.-Ing. Johannes Traugott for the kick-start of my postgraduate studies in Bavaria. Dr. Gustavo López-Risueño is acknowledged for supervision and advice from and in the Netherlands. I am also grateful to Prof. Robert Piché for making it possible for me to work with the ESA.

Last but not least, the supportive attitude of my friends and family—regardless of the varying distance—has not been unnoticed. Bedankt.

Noordwijk, July 2013

Martti Kirkko-Jaakkola
## CONTENTS

*Preface* ................................................................. iii

*List of Publications* ........................................... vii

*Abbreviations and Symbols* ..................................... ix

1. *Introduction* .................................................... 1
   1.1 Scope of the Research ........................................ 3
   1.2 Main Contributions .......................................... 3
   1.3 Author's Contribution ....................................... 4

2. *Preliminaries* .................................................. 5
   2.1 Global Navigation Satellite Systems ....................... 5
      2.1.1 Overview of Signal Processing .......................... 6
      2.1.2 Basic Measurement Models .............................. 7
   2.2 Inertial Sensors and Dead Reckoning ..................... 9
      2.2.1 Accelerometers ........................................ 10
      2.2.2 Gyroscopes ........................................... 11
      2.2.3 Inertial Navigation Mechanization ................... 12
   2.3 Signals of Opportunity and Non-Inertial Sensors ......... 14
   2.4 Bayesian Filtering ......................................... 15

3. *Inertial Sensor Error Estimation and Calibration* ...... 17
   3.1 Error Processes and Allan Variance ...................... 17
   3.2 Calibration Strategies ..................................... 18
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3</td>
<td>Mathematical Modeling of Sensor Errors</td>
<td>19</td>
</tr>
<tr>
<td>3.4</td>
<td>Error Elimination by Carouseling</td>
<td>21</td>
</tr>
<tr>
<td>3.5</td>
<td>Relation of Thesis to Related Work</td>
<td>22</td>
</tr>
<tr>
<td>4.</td>
<td><strong>Using Map Information in Indoor Navigation</strong></td>
<td>25</td>
</tr>
<tr>
<td>4.1</td>
<td>Choice of Filter</td>
<td>25</td>
</tr>
<tr>
<td>4.2</td>
<td>Map Matching Methods</td>
<td>26</td>
</tr>
<tr>
<td>4.3</td>
<td>Fusion Filter Initialization Approaches</td>
<td>28</td>
</tr>
<tr>
<td>4.4</td>
<td>Different Levels of Map Information Detail</td>
<td>29</td>
</tr>
<tr>
<td>4.5</td>
<td>Relation of Thesis to Related Work</td>
<td>30</td>
</tr>
<tr>
<td>5.</td>
<td><strong>Cycle Slip Detection for Low-Cost GNSS Receivers</strong></td>
<td>33</td>
</tr>
<tr>
<td>5.1</td>
<td>Differential Measurements</td>
<td>34</td>
</tr>
<tr>
<td>5.2</td>
<td>Phase Prediction Approaches</td>
<td>35</td>
</tr>
<tr>
<td>5.3</td>
<td>Relation of Thesis to Related Work</td>
<td>36</td>
</tr>
<tr>
<td>6.</td>
<td><strong>Conclusions</strong></td>
<td>39</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary of Results</td>
<td>39</td>
</tr>
<tr>
<td>6.2</td>
<td>Discussion</td>
<td>40</td>
</tr>
<tr>
<td>6.3</td>
<td>Future Work</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td><strong>References</strong></td>
<td>43</td>
</tr>
<tr>
<td></td>
<td><strong>Publications</strong></td>
<td>55</td>
</tr>
</tbody>
</table>
LIST OF PUBLICATIONS

This thesis consists of an introductory part and the following articles which have been published in open literature:


The publications are reproduced here with kind permissions from the publishers.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>Access Point</td>
</tr>
<tr>
<td>AR</td>
<td>Autoregressive</td>
</tr>
<tr>
<td>B-frame</td>
<td>Coordinate frame fixed to the sensor housing (body)</td>
</tr>
<tr>
<td>C/A</td>
<td>Coarse Acquisition</td>
</tr>
<tr>
<td>$C/N_0$</td>
<td>Carrier-to-Noise-Density Ratio</td>
</tr>
<tr>
<td>DR</td>
<td>Dead Reckoning</td>
</tr>
<tr>
<td>E-frame</td>
<td>Global ECEF frame</td>
</tr>
<tr>
<td>ECEF</td>
<td>Earth-Centered Earth-Fixed</td>
</tr>
<tr>
<td>EKF</td>
<td>Extended Kalman Filter</td>
</tr>
<tr>
<td>ENU</td>
<td>East–North–Up</td>
</tr>
<tr>
<td>GM</td>
<td>Gauss–Markov</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global Navigation Satellite System</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>IMU</td>
<td>Inertial Measurement Unit</td>
</tr>
<tr>
<td>KF</td>
<td>Kalman Filter</td>
</tr>
<tr>
<td>L-frame</td>
<td>Local ENU frame</td>
</tr>
<tr>
<td>L1</td>
<td>GPS carrier frequency 1.57542 GHz</td>
</tr>
<tr>
<td>MEMS</td>
<td>Microelectromechanical System</td>
</tr>
<tr>
<td>PDR</td>
<td>Pedestrian Dead Reckoning</td>
</tr>
<tr>
<td>PF</td>
<td>Particle Filter</td>
</tr>
<tr>
<td>PLL</td>
<td>Phase-Locked Loop</td>
</tr>
<tr>
<td>PPP</td>
<td>Precise Point Positioning</td>
</tr>
<tr>
<td>PVT</td>
<td>Position, Velocity, and Time</td>
</tr>
<tr>
<td>RAIM</td>
<td>Receiver Autonomous Integrity Monitoring</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RTK</td>
<td>Real-Time Kinematics</td>
</tr>
<tr>
<td>SLAM</td>
<td>Simultaneous Localization and Mapping</td>
</tr>
<tr>
<td>TOA</td>
<td>Time of Arrival</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
<tr>
<td>6DOF</td>
<td>Six Degrees of Freedom</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>$\sigma^2_A(\tau)$</td>
<td>Allan Variance at averaging time $\tau$</td>
</tr>
<tr>
<td>$\rho^S(t)$</td>
<td>Pseudorange to satellite $S$ at time $t$</td>
</tr>
<tr>
<td>$\phi^S(t)$</td>
<td>Carrier phase for satellite $S$ at time $t$</td>
</tr>
<tr>
<td>$f^S_D(t)$</td>
<td>Doppler shift for satellite $S$ at time $t$</td>
</tr>
<tr>
<td>$x(t)$</td>
<td>Position of the user at time $t$</td>
</tr>
<tr>
<td>$x^S(t)$</td>
<td>Position of satellite $S$ at time $t$</td>
</tr>
<tr>
<td>$x_t$</td>
<td>Filter state at time $t$</td>
</tr>
<tr>
<td>$f_t$</td>
<td>State transition function at time $t$</td>
</tr>
<tr>
<td>$q_t$</td>
<td>Process noise at time $t$</td>
</tr>
<tr>
<td>$y_t$</td>
<td>Measurement made at time $t$</td>
</tr>
<tr>
<td>$h_t$</td>
<td>Measurement model function at time $t$</td>
</tr>
<tr>
<td>$r_t$</td>
<td>Measurement error at time $t$</td>
</tr>
<tr>
<td>$|\cdot|$</td>
<td>Euclidian norm</td>
</tr>
<tr>
<td>$c$</td>
<td>Speed of light</td>
</tr>
<tr>
<td>$\delta_t(t)$</td>
<td>Receiver clock bias at time $t$</td>
</tr>
<tr>
<td>$\dot{\delta}_t(t)$</td>
<td>Receiver clock drift at time $t$</td>
</tr>
<tr>
<td>$\delta^S(t)$</td>
<td>Clock bias of satellite $S$ at time $t$</td>
</tr>
<tr>
<td>$I^S(t)$</td>
<td>Ionospheric delay for satellite $S$ at time $t$</td>
</tr>
<tr>
<td>$T^S(t)$</td>
<td>Tropospheric delay for satellite $S$ at time $t$</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Additive random noise</td>
</tr>
<tr>
<td>$N^S$</td>
<td>Carrier phase integer ambiguity for satellite $S$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Wavelength</td>
</tr>
<tr>
<td>$f$</td>
<td>Frequency</td>
</tr>
<tr>
<td>$g$</td>
<td>Magnitude of the gravitational acceleration</td>
</tr>
<tr>
<td>$g^L$</td>
<td>Three-dimensional gravitational acceleration in the $L$-frame</td>
</tr>
<tr>
<td>$f^B(t)$</td>
<td>Specific force measurement vector in the $B$-frame at time $t$</td>
</tr>
<tr>
<td>$S$</td>
<td>Scale factor and cross-coupling matrix</td>
</tr>
<tr>
<td>$a^B(t)$</td>
<td>True acceleration vector in the $B$-frame at time $t$</td>
</tr>
<tr>
<td>$C^B_L$</td>
<td>Rotation matrix from the $L$-frame to the $B$-frame</td>
</tr>
<tr>
<td>$b(t)$</td>
<td>Additive sensor bias at time $t$</td>
</tr>
<tr>
<td>$\omega^B(t)$</td>
<td>Output of a gyroscope triad in the $B$-frame at time $t$</td>
</tr>
<tr>
<td>$\omega_{true}^B(t)$</td>
<td>True angular rate in the $B$-frame at time $t$</td>
</tr>
<tr>
<td>$\Omega_E$</td>
<td>Magnitude of the Earth rotation rate</td>
</tr>
<tr>
<td>$\Omega^E_E$</td>
<td>Earth rotation vector expressed in the $E$-frame</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

Positioning equipment are becoming a part of the everyday life of individuals because of a rapid development during the past decades. No longer is the word “navigation” only related to maritime travel which has been a significant driving force of navigation research for centuries. As the performance of navigation systems has improved and the cost of such systems has decreased, new applications of position information have arisen. While a positioning uncertainty of, e.g., 100 meters is much better than a 15th-century navigator would ever have needed, it is excessive for a personal navigation assistant providing turn-by-turn directions to a car driver in a downtown environment.

The rise of the personal navigation market was made possible by the development of Global Navigation Satellite Systems (GNSSs), most notably the U.S. Global Positioning System (GPS) and the deactivation of the intentional degradation of its freely accessible signals. Although satellite navigation equipment nowadays are small, power-efficient, and low-priced enough to be integrated into devices such as cellular phones, the satellite signals themselves are subject to certain inherent limitations. The most significant limitation is the fact that the signals are broadcast from a medium Earth orbit altitude of approximately 20000 kilometers at a power of a few dozen watts, which implies that the signals are very weak when they reach the Earth. This makes them particularly susceptible to interference and attenuation; most GNSS receivers are unable to acquire and track signals that have penetrated walls and, therefore, cannot operate indoors.

Despite the limitations of satellite navigation, demand for indoor positioning exists. For instance, the possibility of tracking first responders on mission could save the life of a firefighter who gets injured. As a more commercial example, applications that help customers navigate in a shopping mall to the shop or product they are looking for can create revenue by advertising other products and services located close to the user. In general, people spend a great deal of their time indoors
and expect location-based applications to work seamlessly.

In order to reach the level of performance needed by such applications, navigation methods complementary to satellite signals are needed. Inertial sensors based on microelectromechanical system (MEMS) technology have promising size, cost, power consumption, and durability properties considering the requirements of the mass market. However, despite immense development during the past years, their output contains significant error components which prevent their direct application for navigation purposes. Nevertheless, given the fact that these sensors are immune to external factors such as interference or jamming, it is intriguing to investigate how such sensors could be used to assist other methods of navigation.

Current indoor positioning solutions are mostly based on detecting wireless local area network (WLAN) signals. Knowing the position of the transmitting access point (AP) makes it possible to estimate the position of the receiver with a variety of methods. The same principle can be used in a larger scale with cellular phone networks, but with a lower accuracy. An obvious downside of these signals is the need of a database of AP or cell tower locations; such databases are tedious to build and maintain, but commercial solutions are available [1, 2]. A more serious drawback of relying on WLAN is the fact that the operation of APs can be interrupted by, e.g., a power failure, making the approach less favorable for the use case of first responders.

Limited availability indoors is not the only area of improvement foreseen for satellite based personal navigation solutions, but the positioning accuracy has potential for enhancement. Centimeter-level positioning accuracies have been achieved for years using professional GNSS equipment, but similar techniques have not gained widespread success in the consumer market. Although an equivalent accuracy might be impossible to attain using miniaturized low-cost antennas and other hardware, an accuracy in the order of, e.g., 10 centimeters would be sufficient for identifying which lane a car is driving on.

However, high accuracies demand robust algorithms or precise identification of any erroneous measurements—not only with satellite signals, but any type of observation. When adapting traditional error detection algorithms to such scenarios, a measurement error of, e.g., 10 cm can be regarded as an outlier although it normally would be adequate to detect errors 100 times as large.
1.1 *Scope of the Research*

The performance of a navigation system does not equal its positioning *accuracy* but also comprises other important factors such as *reliability* and the *availability* of a position solution. The system is reliable if the reported positioning accuracy can be considered trustworthy, i.e., no unmodeled gross measurement errors are biasing the solution. Availability refers to the percentage of time during which the system is able to report a position solution; for instance, using standalone GNSS yields a poor availability when the user is located indoors.

The objective of this research is to investigate methods of processing measurements from low-cost consumer-grade personal positioning equipment to enhance the availability and reliability of the position solution. The equipment under consideration are GNSS receivers and MEMS sensors. The focus with MEMS sensors is on calibrating the sensors, particularly gyroscopes; different mechanization approaches are not investigated in detail.

1.2 *Main Contributions*

The contributions of this thesis can be summarized as follows.

- A proof-of-concept measurement setup and error model to achieve an accuracy sufficient for North seeking using a low-cost MEMS gyroscope only [I]. The ability to replace optical gyroscopes with MEMS sensors in high-precision applications leads to significant reductions in the size and cost of the equipment.

- New algorithms for estimating the additive bias of a MEMS gyroscope [II]. More accurate estimation of the bias leads to a lower heading drift in dead reckoning systems. The proposed algorithms are based on the spectral properties of the bias instability.

- Demonstrating the use of map-matching with low-cost inertial MEMS sensors for long-term indoor using GNSS initialization as the sole external information [III]. The method is independent of WLAN measurements or other
signals of opportunity based on the availability of local infrastructure. Therefore, the applicability of the method is not limited, e.g., in emergency situations during a power outage.

- A novel method to detect and correct for measurement anomalies in the carrier phase observations of a stand-alone low-cost GNSS receiver [IV]. The method is based on receiver autonomous integrity monitoring (RAIM) methodology onto time-differenced carrier phase measurements. Since only one receiver and single-frequency measurements are assumed, the method is suitable for error detection in various precise positioning use cases, thus improving the positioning reliability.

1.3 Author's Contribution

The main results are reported in four publications and the author’s contribution to each publication is described in the following.

[I] The author designed and implemented the Extended Kalman Filter. He also worked out the error analysis. Furthermore, the author wrote the manuscript in collaboration with L.I. Iozan. The results of this publication have been partially discussed in Dr. Iozan's dissertation [3], but the error analysis was not covered in [3].

[II] The author worked out the theory, designed and implemented the Kalman Filter method, planned and ran the tests with both proposed methods, and wrote the manuscript.

[III] The author designed and implemented the algorithm, collected and processed the experimental data, and wrote the manuscript.

[IV] The author developed the proposed method with J. Traugott and validated its performance with the test data. He wrote the manuscript in collaboration with J. Traugott. The results of this publication have been discussed in Dr. Traugott's dissertation [4].
2. PRELIMINARIES

In this chapter, the positioning systems considered in this thesis are described. Furthermore, measurement models are presented for those observation\(^1\) types that are studied in the thesis. Furthermore, a short overview is given about the most common signals of opportunity and non-inertial motion sensors, although they are not investigated in the scope of the thesis. Finally, the concept of Bayesian filtering, a popular framework for combining measurements from various sources, is presented.

2.1 Global Navigation Satellite Systems

Modern GNSS positioning is based on the principle of trilateration. Determining the time of arrival (TOA) of a signal whose transmit time is known makes it possible to estimate the distance between the transmitter and the receiver. However, the clock of the receiver is not synchronized to the satellites’ clocks; therefore, the measured ranges have a common additive bias and are called *pseudoranges*. Therefore, four satellites are needed to solve for the three position coordinates and the additive clock bias. TOA-based ranging was made possible by advances in space-capable atomic clock technology; earlier systems, such as Transit, were based on long observations of the Doppler shift of low Earth orbit satellite signals [5, 6].

GNSS signals are broadcast on the L-band. The satellites transmit signals on multiple frequencies, but not all signals are available to the public. For instance, in the GPS system, there is currently only one publicly available signal that is broadcast by each satellite; this is the coarse acquisition (C/A) signal that is transmit at the L1 frequency of 1.57542 GHz. Other current GPS signals are either supported

\(^1\) In this thesis, the words ‘measurement’ and ‘observation’ are used as synonyms
by a limited number of satellites only or restricted to the use of the U.S. military and other parties approved by the U.S. Department of Defense. Although GPS L1 is the most widely used signal, the Russian GLONASS system has reached full operational capability and is utilized by many professional and mass-market GNSS receivers on the market today. The modernization of these operational systems and the deployment of new GNSSs such as the European Galileo and Chinese BeiDou will bring new signals and frequencies for civilian users too, but it remains to be seen how long it will take until a majority of receivers support them.

2.1.1 Overview of Signal Processing

The main functional blocks of a GNSS receiver are illustrated in Fig. 2.1. The radio frontend filters and converts the analog radio signal to digital in-phase (I) and quadra-phase (Q) samples [7]. The baseband block uses these samples for signal acquisition and tracking, and extracts measurements based on which the position, velocity, and time (PVT) solution is computed. The baseband block can also output other information such as the carrier-to-noise-density ratio ($C/N_0$); in optimal reception conditions the $C/N_0$ can exceed 50 dB-Hz, but indoors it can decrease below 10 dB-Hz; there exist receivers that can track signals at that power level, e.g., [8], if they have first acquired the signal at a higher power.

A GNSS signal consists of three components: a sinusoidal carrier wave, a binary ranging code, and navigation data bits. Ranging codes are sequences that have special correlation properties: the correlation between two different codes is negligible, and so is the correlation of a code with an index-shifted version of itself. The duration of one code bit, or chip as they usually are called because they contain no information, affects the ranging precision: the GPS C/A code has a chipping rate of 1.023 MHz while the more precise restricted-access signals have a ten times higher chipping rate.

The navigation data bits contain the information that is needed to transform a set of pseudoranges into a position solution. The most important information are the
ephemeris, i.e., a prediction of the position of the satellite as a function of time, and the satellite clock error model which is needed to account for the offset errors of the clocks in individual satellites. Furthermore, the navigation message contains the time stamps that are needed to determine the time of transmission. However, as opposed to the ranging code, navigation data bits are not known beforehand and, therefore, a bit transition is seen to the receiver as a 180° shift in the carrier phase. Although this is not a problem for most applications, the effect of bit transitions is detrimental in high-sensitivity processing where the correlation peak is accumulated coherently over a long period of time [9]. Fortunately, the repetitive nature of the navigation message stream makes it possible to predict the bit values [10]. Furthermore, many of the modern signals contain a separate pilot channel where no navigation data modulation is present, but only a handful of satellites transmitting such signals are currently operational.

2.1.2 Basic Measurement Models

In order to track a satellite signal, the receiver must follow the phase of the ranging code and either the frequency or the phase of the carrier wave. The code phase measurement can be transformed into a pseudorange. The models of these observables are presented in the following sections.

Pseudorange

The pseudorange is the basic measurement in satellite positioning. Derived from the TOA, the pseudorange to satellite \( S \) at time \( t \) can be modeled, in units of meters, as

\[
\rho^S(t) = \| x^S(t) - x(t) \| + c \left( \delta t(t) - \delta t^S(t) \right) + I^S(t) + T^S(t) + \epsilon^S(t)
\]

(2.1)

where, omitting the time indices, \( x^S \) and \( x \) are the positions of the satellite and the user, respectively; \( c \) denotes the speed of light; \( \delta t \) and \( \delta t^S \) are the additive clock biases of the receiver and the satellite, respectively; \( I^S \) and \( T^S \) are the propagation delays occurring while penetrating the ionosphere and troposphere, respectively; and \( \epsilon^S \) contains all other errors, such as multipath propagation, i.e., the interference occurring when multiple reflections of the signal are received, errors in the broadcast ephemeris, and measurement noise.
Table 2.1: Approximate magnitudes of pseudorange error sources [10]

<table>
<thead>
<tr>
<th>Error source</th>
<th>Contribution [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual satellite clock error</td>
<td>1</td>
</tr>
<tr>
<td>Ionospheric delay</td>
<td>7</td>
</tr>
<tr>
<td>Tropospheric delay</td>
<td>2</td>
</tr>
<tr>
<td>Multipath</td>
<td>0.2</td>
</tr>
<tr>
<td>Receiver noise</td>
<td>0.1</td>
</tr>
<tr>
<td>Other errors</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2.1 shows an exemplary error budget for a pseudorange measurement. The values given for multipath and noise correspond to favorable reception conditions. However, in harsh conditions, such as indoors, their contributions are significantly higher. In the worst case, the line-of-sight signal path is too attenuated to be detected at all, and only reflections of the signal are tracked, leading to severely erroneous observations. It can be concluded that pseudorange measurements can achieve an accuracy no better than in the order of decimeters if deterministic errors are corrected for, but normally in the order of meters.

Carrier Phase and Doppler

If the receiver has a phase-locked loop (PLL), the information about the carrier phase $\Phi^S(t)$ can be used to construct a measurement reminiscent of the pseudorange $\rho^S(t)$, modeled in units of carrier cycles as

$$\Phi^S(t) = \lambda^{-1} \left( \| x^S(t) - x(t) \| + c \left( \delta t(t) - \delta t^S(t) \right) - I^S(t) + T^S(t) \right) + N^S + \epsilon^S(t) \tag{2.2}$$

where $\lambda$ denotes the signal wavelength. The key difference with respect to the pseudorange model (2.1) is the appearance of a new term $N^S$, known as the integer ambiguity. Furthermore, the sign of the ionospheric error is changed because of the dispersive nature of the ionosphere. The multipath and noise characteristics are notably different from their code-based counterparts: phase tracking noise is in the order of fractions of a carrier cycle, corresponding to millimeters or centimeters in the distance domain, and the multipath error is limited to no larger than a quarter cycle [11].

The reason for the appearance of the integer ambiguity term is the fact that when
the tracking loop locks onto the phase, it only measures the fractional phase but cannot know the amount of full carrier cycles between the receiver and transmitter antennas. Actually, depending on the tracking loop architecture, this ambiguity can be in halves of carrier cycles instead of full wavelengths: if the receiver is using so-called Costas discriminators to overcome the $180^\circ$ phase shifts caused by the navigation data bits, there is a half-cycle ambiguity until the receiver manages to decode navigation message frames [10]. An important property of the integer ambiguity is the fact that it remains constant during continuous phase lock, which is why it lacks the time index in (2.2). A change in the integer ambiguity caused by a temporary loss of lock is called a cycle slip, and detecting such situations is vital in precise GNSS positioning; we will study this problem in Chapter 5.

If the receiver only tracks the frequency of the carrier but not its phase, instantaneous estimates of the Doppler shift are obtained. Carrier phase observations from a PLL also make it possible to estimate the Doppler shift by dividing the difference of two consecutive phase observations by the duration of the measurement interval, assuming that no cycle slips occurred. Therefore, the Doppler shift is directly connected to the derivative of the pseudorange and carrier phase:

$$f_D^S(t) = \lambda^{-1}\left(\frac{d}{dt}\|x_S^S(t) - x(t)\| + c\dot{\delta}t(t)\right) + \epsilon^S(t). \tag{2.3}$$

In this formulation, the derivatives of the atmospheric errors and the satellite clock bias have been omitted because these errors do not typically change rapidly. The time derivative of the clock bias, $\dot{\delta}t$, is usually called the clock drift. The Doppler noise typically has a smaller variance than the pseudorange noise and, therefore, the Doppler measurement offers a way to estimate the velocity of the receiver more precisely than by differencing consecutive pseudorange-based position solutions.

## 2.2 Inertial Sensors and Dead Reckoning

**Dead reckoning** (DR) is a method of position estimation where measured changes in position are accumulated; the initial position needs to be known by other means. The displacements can be measured using, e.g., an odometer or a radar in land vehicles, or inertial sensors; laser scanners are another option [12], but such equipment are more commonly encountered in the field of robotics than in personal positioning.
Preliminaries

(a) At rest  

(b) Accelerating rightward  

(c) At rest

Fig. 2.2: Schematic of a spring–mass accelerometer

It should be noted that DR is not the sole means of utilizing sensor measurements; for instance, they can be used for motion state recognition to verify the applicability of a specific GNSS positioning algorithm [13]. In the following sections, we discuss the basics of inertial sensors, i.e., accelerometers and gyroscopes; an inertial measurement unit (IMU) typically consists of three accelerometers and three gyroscopes. Finally, approaches for mechanizing the output of inertial sensors are addressed.

2.2.1 Accelerometers

Accelerometers measure linear accelerations, and a typically construction of such a sensor is a spring-and-mass systems where the inertia of the mass causes the spring to extend or contract when subject to an external force, as predicted by Newton’s laws. Strictly speaking, the output of an accelerometer is not the acceleration but a measurement of specific force: both the test mass and the sensor housing are subject to gravity, but normal forces only act on the housing, which causes the sensor to measure the normal force but not the gravitational acceleration. This phenomenon is illustrated in Fig. 2.2. Consequently, the force sensed by an accelerometer in free fall is zero.

Because of the gravitational acceleration component in the measurement, one must know and compensate for the gravity by other means in order to estimate the velocity and position by integrating the measurements. However, the gravitational component is not always a nuisance: for instance, it allows a smartphone to use accelerometers to determine if it is being held in portrait or landscape orientation.

The output of an accelerometer triad can be modeled in the sensors’ body frame $B$
2.2. Inertial Sensors and Dead Reckoning

![Schematic of a Coriolis vibratory gyro. A test mass is vibrated along the drive axis and its Coriolis acceleration is observed along the sensing axis.](image)

in vector notation as

\[
f^B(t) = S a^B(t) - C^B_L g^L + b(t) + \epsilon(t)
\] (2.4)

where \(S \in \mathbb{R}^{3 \times 3}\) is the scale factor and cross-coupling error matrix; \(a^B\) is the true three-dimensional acceleration vector as expressed in the body frame; \(C^B_L\) is the rotation matrix from the local East–North–Up (ENU) frame \(L\) to the body frame \(B\) and \(g^L = [0 \ 0 \ -g]^T\) is the gravitational acceleration vector in the \(L\)-frame; \(b\) denotes the additive biases of the sensors; and \(\epsilon\) represents measurement noise. In reality, the distinction between the bias \(b\) and noise \(\epsilon\) is far from trivial because of stochastic error processes that exhibit correlation in time; we will focus on this problem in Chapter 3.

2.2.2 Gyroscopes

While accelerometers measure linear accelerations, rotations are sensed using gyroscopes, which are commonly called gyros. The gyroscope equivalent of the mechanical spring–mass accelerometer is a heavy spinning test mass; it resists external rotations in accordance with the conservation of angular momentum. However, such a construction is often impractical, and thus, other types of gyros, such as optical and MEMS gyroscopes, have gained popularity. Optical gyros are based on the Sagnac effect whereas MEMS gyros make use of the Coriolis effect [14]. The principle of operation of a Coriolis gyro is depicted in Fig. 2.3 where a test mass is made to vibrate along the drive axis. If the system is rotated around an axis perpendicular to the drive and sensing axes, the Coriolis effect will cause the test mass to move along the sensing axis in proportion to the rate of rotation, thus making it possible to estimate the rotation rate by measuring the displacement.
The output of a terrestrial gyroscope includes a component due to the rotation of the Earth, which, in principle, makes it possible to directly estimate the heading using gyroscopes in a procedure known as *gyrocompassing*. However, the Earth rotation is a weak signal and typically buried under the measurement noise of consumer-grade gyros, but it is used in the initial alignment of high-grade IMUs.

To describe the angular rate output $\omega$ of a gyro triad, we will use the following model which is analogous to the accelerometer model (2.4):

$$\omega^B(t) = S \omega_{true}^B(t) + C^B_E \Omega^E_E + b(t) + \epsilon(t)$$  \hspace{1cm} (2.5)

where $\omega_{true}^B(t)$ is the true angular rate in the $B$-frame, $C^B_E$ is the rotation matrix from the Earth-Centered Earth-Fixed (ECEF) frame $E$ to the body frame, and the Earth rotation rate is expressed in the vector form $\Omega^E_E = [0 \quad 0 \quad \Omega_E]$.

In the case of MEMS gyroscopes, the bias term $b(t)$ changes when the sensor is subject to a linear acceleration; this property is often referred to as the *g-sensitivity*. Furthermore, the biases of MEMS sensors tend to be sensitive to the ambient temperature; this problem can be circumvented, e.g., by placing the sensor into a temperature-controlled oven [15] but this increases the physical size, power consumption, and manufacturing cost of the unit.

### 2.2.3 Inertial Navigation Mechanization

The process of transforming the outputs of accelerometers and gyros to position information is called *mechanization*. With three accelerometers and three gyros, the most intuitive mechanization is to directly integrate the measurements in order to yield position, velocity, and attitude information; this is known as the six degrees of freedom (6DOF) inertial navigation mechanization. An IMU that uses gimbals to retain its orientation fixed with reference to an inertial space is called a stabilized platform system whereas a unit without such mechanics is called a *strapdown* IMU.

Unfortunately, the 6DOF mechanization has severe shortcomings. Integrating the sensor outputs also causes the bias errors to be integrated. Thus, for instance, a constant acceleration estimation bias turns into a position error that increases in proportion to time squared. Unless the sensors are of very high accuracy, the error growth limits the applicability of the 6DOF mechanization to short time spans.
only in the absence of other position information. Another problem is the need to compensate for the gravitational acceleration which itself is a function of the altitude, constituting a positive error feedback loop [16].

Therefore, in order to make use of low-cost inertial sensors for positioning, alternative mechanization approaches are needed. If the user is known to be moving on foot, it is possible to detect steps from the acceleration signal and propagate the position estimate in discrete steps; this mechanization is called pedestrian dead reckoning (PDR) [17]. PDR implementations differ, e.g., in the mounting point of the IMU: mounting it onto the foot of the user [18] makes it possible to estimate the stride length by integrating the sensor output, whereas body-mounted sensors require an external model for the step length but are a much more practical scenario when considering, e.g., mobile devices such as cellular phones.

PDR has significant advantages over the 6DOF mechanization:

- the position error is proportional to the accumulated heading error and the number of steps taken, i.e., linear in time instead of quadratic growth,

- detecting steps does not require compensation of the gravitational acceleration or the accelerometer biases, and

- there is no need to track the roll and pitch angles accurately, only the heading is used for navigation; roll and pitch determined using accelerometers are typically sufficient for resolving the heading rate if the IMU is not leveled, such as in the case of hand-held mobile devices.

On the other hand, PDR does not improve the stability of the heading estimation, and especially with body-mounted sensors, stride length models often require user-specific calibration and vertical motion can cause problems, e.g., in stairs. Nevertheless, PDR has been shown to be feasible with low-cost sensors as long as external information to limit the heading error growth is available.

Other inertial navigation mechanizations exist as well. For instance, imposing constraints on the navigation solution, e.g., a fixed ground speed, leads to different mechanizations [19].
2.3 Signals of Opportunity and Non-Inertial Sensors

Today’s mobile devices are capable of receiving signals whose purpose is not positioning but which still can yield position information; they are called *signals of opportunity*. Furthermore, the devices often contain other sensors than accelerometers and gyros that give information about the motion or heading of the user. Although such sources of information are not considered in this thesis, they are especially popular in indoor-based applications and, therefore, a short overview is given on the most common signals of opportunity and non-inertial sensors.

Probably the most popular method of indoor positioning is to use WLAN signals. Because of their relatively short range, observing a signal from an AP whose location is known already yields a coarse estimate of the location of the receiver, and the precision is improved when multiple APs are observed and the received signal strengths (RSS), which depend on the distance between the transmitter and the receiver, are considered. The most popular strategies are triangulation and fingerprinting. In the former, the signal strengths are transformed to distance estimates using a path loss model, which allows to resolve the user location. The positioning performance then depends on the accuracy of the path loss model; unfortunately, the attenuation of radio frequency (RF) signals in indoor environments is difficult to predict. In contrast, the fingerprinting approach relies on a radio map constructed by recording the observed APs and signal strengths at known locations. The accuracy of this method depends on the granularity of the radio map, but the relation is nonlinear [20, 21]. Nevertheless, collecting a densely spaced radio map is a tedious task.

Cellular network signals can be used in similar manners as the WLAN processing strategies described above, but the achievable position accuracy is often inadequate. Fortunately, dedicated positioning signals are available in long-term evolution cellular networks, which can achieve a position accuracy in the order of 10 meters or better [22, 23].

Many cellphones and other mobile devices contain a magnetometer, i.e., a digital compass, which allows to estimate the absolute heading. Unfortunately, magnetometers are prone to large heading errors in indoor environments due to a significant amount of nearby electrical equipment and ferromagnetic objects. However, they can be used indoors in a fingerprinting manner to estimate the absolute po-
2.4 Bayesian Filtering

The goal of Bayesian filtering is to estimate the posterior probability distribution \( p(x_t|x_{0...t-1}, y_{1...t}) \) of the unknown state variables \( x \) in the light of certain measurements \( y \), given the prior distribution of the state \( x_{t-1} \); thus, a time series is estimated. As the name suggests, a key principle in filtering the posterior distribution is Bayes’ theorem. The term “optimal filtering” is sometimes used; it refers to optimality in the sense of estimation error variance. A Bayesian filter operates based on a state transition model

\[
x_t = f_t(x_{t-1}) + q_t
\]

(2.6)

where \( f_t \) is a state transition function and \( q \) is a random variable denoting process noise, and measurements

\[
y_t = h_t(x_t) + r_t
\]

(2.7)

where \( h_t \) is a measurement function and \( r \) represents measurement noise. In the above equations, the time \( t \) is indexed using a subscript to emphasize that we are operating in discrete time. Generally, measurement and process noises are assumed to be zero-mean, white, and independent of the state \( x_t \). In navigation applications, a classical use case for a Bayesian filter is the integration of a GNSS receiver with an IMU [14].

The posterior distribution cannot be solved for generally in closed form. However, when the transitional and measurement models are linear and the underlying probability distributions are Gaussian, a closed-form solution does exist; it can
be computed using an algorithm known as the Kalman filter (KF) [31]. With significantly nonlinear and/or non-Gaussian models, Monte Carlo methods can be the most straightforward option.

When operating on Gaussian distributions, it is only necessary to compute the mean and covariance in order to completely characterize the posterior distribution; this is what the KF does. Furthermore, with linear transitional and measurement models, the filtering of these moments is carried out by means of fairly simple matrix algebra; therefore, the KF is computationally relatively efficient [32].

However, most measurement models related to navigation are nonlinear by nature, which prevents direct application of the KF. Probably the most widely used solution is to use the KF on the first-order Taylor series of the nonlinear functions; this method is known as the Extended Kalman Filter (EKF) [33]. For instance, linearizing a GNSS pseudorange equation yields a locally fairly accurate approximation because of the long distance to the satellites, and therefore, EKF has been observed to function well in GNSS positioning. Nevertheless, it should be kept in mind that the EKF is not an optimal filter [34]—after all, it is the exact solution for a different (linearized) model, not even an approximation of the Bayesian solution for the original nonlinear models.
In most use cases, the output signal of inertial sensors is corrected for certain errors before being used for position computations. In this chapter we characterize the most significant error processes that are present in the output of inertial sensors, focusing on the case of MEMS. Furthermore, methods of estimating and compensating for them are discussed. Finally, the differences between the methods proposed in this thesis and the state-of-the-art methods studied in this chapter are described.

3.1 Error Processes and Allan Variance

In Chapter 2, it was mentioned that the distinction between uncorrelated noise and a constant or slowly changing bias is not always simple. The effect of uncorrelated noise can be reduced by averaging consecutive samples, but averaging a changing bias can cause loss of accuracy.

A popular method of finding optimal averaging times is to compute the Allan variance [35] of the sensor output. Allan variance was originally developed for assessing the stability of atomic clocks but is applicable to many other systems as well. It is computed as a function of averaging time $\tau$: the data are partitioned into $N$ disjoint bins of length $\tau$. Denoting the average value of the $i$th bin by $\bar{y}(\tau)_i$, the Allan variance is computed as

$$\sigma^2_A(\tau) = \frac{1}{2(N-1)} \sum_{i=1}^{N-1} \left(\bar{y}(\tau)_{i+1} - \bar{y}(\tau)_i\right)^2; \quad (3.1)$$

the square root of the Allan variance is called the Allan deviation.

It is customary to visualize the Allan variance as a log–log graph; an example is shown in Fig. 3.1. The concave shape of the curve reveals many error characteristics [36], but we focus on the three most significant ones. At short averaging times,
3. Inertial Sensor Error Estimation and Calibration

The Allan variance is decreasing, theoretically with a slope of $-1$; this is caused by the averaging of white (uncorrelated) noise. As the averaging time gets longer, the Allan variance eventually starts to increase, which is due to a slowly drifting additive bias; in the context of gyros, this error component is called rate random walk. The flat section between these two slopes is caused by $1/f$ (flicker) noise; the name refers to the fact that the power spectral density of the process is proportional to $|f|^{-\gamma}$ for some $0 < \gamma < 2$ [37] with $f$ referring to the frequency.

There exist other ways to characterize error processes as well, both in the time and frequency domains [38, 39]. However, Allan variance is widely used in the context of inertial sensors, and it is recommended in the standard [36]. Nevertheless, other methods which are more popular in, e.g., the field of radiophysics, can provide more detailed information on certain error characteristics [40].

3.2 Calibration Strategies

Particularly with MEMS sensors, the additive bias tends to change significantly every time the device is powered up and, therefore, an on-the-fly calibration is necessary in most applications. It is necessary to know the value of the true input during the calibration, which is often only possible when the sensor is stationary, i.e., the input is zero; however, if complementary information is available, such as when integrating an IMU with a GNSS receiver, stationarity is not required.

An example of an error estimation scheme based on stationarity detection was
3.3 Mathematical Modeling of Sensor Errors

In order to have optimal estimates of the sensor errors between the calibration epochs as well, their evolution in time should be modeled. Of the three main error components, $1/f$ noise is the most difficult to model. In [47], a wavelet-based model was applied and a filtering scheme was proposed to reduce the effect of $1/f$ noise. Test results were presented for high-accuracy (navigation grade) sensors; the $1/f$ filter improved the attitude accuracy by 20% after low-pass filtering of the white noise component.

proposed in [41] where fuzzy inference was applied to determine periods when a calibration algorithm [42] is applicable. This calibration algorithm estimates the scale factor and bias errors of an accelerometer triad; it finds such error parameter values that cause the Euclidian norm of the three accelerometer measurements to equal the gravitational acceleration $g$. Unfortunately, in many cases, the exact value of $g$ is not known as it is not a universal constant but varies locally depending on, e.g., the latitude and the altitude.

Instead of stationarity, another possibility for calibration is a sequence of known movements such as proposed, e.g., in [43]. This approach has the advantage that it makes scale factor errors observable, but the obvious downside is that performing such calibration steps can be cumbersome, or even practically impossible if a great precision is required, to the user, especially with hand-held equipment; these procedures are meant to be performed in laboratory conditions. With precisely controlled rotations, such as on a turntable, it is possible to observe the centripetal force using the accelerometers and improve the calibration performance with these ancillary measurements [44].

Not all rotation-based methods require high-precision movements; for instance, the algorithm proposed in [45] is able to calibrate an IMU when rotated by hand, but obviously with a lower precision than when the rotations are carried out by a machine. Nevertheless, the performance of the method has been shown to be adequate for short-term navigation [46].
Autoregressive (AR) models of order $k$, in discrete time expressed as

$$x_t = \sum_{i=1}^{k} a_i x_{t-i} + \epsilon_t \tag{3.2}$$

where $\epsilon$ represents random noise, are popular error models; in [48], AR moving average models were used for predicting the evolution of sensor errors. A comparison of AR models up to the fourth order indicated that even with low-order models, the prediction performance does not monotonically increase with model order: third-order models yielded a better performance than their fourth-order counterparts.

A recursive least-squares approach for AR modeling was proposed in [49] where the idea was to estimate the AR model parameters on the fly in order to account for, e.g., temperature sensitivity effects. Test results using low-order models suggested a 20% reduction in prediction error with the recursive approach when compared with a nonrecursive estimate. However, the recursive method depends on a correlation time parameter for which no automated estimation procedure was described.

An extensive analysis of AR modeling of MEMS sensor errors is given in [50, 51] with comparison made to another popular model, the first-order Gauss–Markov (GM) process

$$x_t = e^{-\Delta t/\tau} x_{t-1} + \eta_t \tag{3.3}$$

where $\tau$ is the correlation time constant of the process, $\Delta t$ is the sampling interval, and $\eta$ is Gaussian noise. In a static measurement setup, a fourth-order AR model was observed to yield a positioning accuracy approximately 30% better than the GM model in most test cases but the improvement was only about 5% in some updating frequency and update accuracy configurations. In kinematic tests, the performance of AR was not as significantly superior to GM.

A first-order GM model to combine the outputs of several redundant gyroscopes into one “virtual sensor” was proposed in [52] with the conclusion that the lower complexity of the GM model in comparison with higher-order AR processes is a more significant advantage when multiple sensors are to be modeled. Test results indicated that the virtual gyroscope had a lower Allan variance than the individual sensors, but the availability of redundant sensors is limited to specialized use.
cases only and, therefore, the applicability of the method is quite limited when considering today’s mobile devices and other consumer products.

In [53], a black-box method called parallel cascade identification was used to model error processes of MEMS gyros. Results from many tests suggested that the parallel cascade identification approach yields a performance comparable to AR but clearly outperforms GM models. Artificial neural networks are another black-box method that has been used for modeling sensor outputs. A radial basis network predictor [54] has been observed to yield a better performance than a fourth-order AR model but at the cost of computational resources and training time.

### 3.4 Error Elimination by Carouseling

Instead of estimating the sensor errors in a separate calibration period, it is possible to rotate the sensors during the run in a pattern that cancels the net effect of the errors or at least significantly mitigates them [55]; this approach is called carouseling. Implementing carouseling requires, depending on the rotation pattern, one or more motors and possibly gimbals to turn the sensor, which increases the physical size and weight, power consumption, and manufacturing costs of the device, but on the other hand, it can enable very precise compensation of measurement errors.

Carouseling is a particularly useful approach for observing weak signals, such as the Earth rate, using low-cost sensors. In [56], a measurement setup was proposed to measure the rotation rate of the Earth with a MEMS gyro. This setup relied on a carouseling scheme known as indexing or two-position carouseling, where data are first logged with the sensor remaining stationary, then the gyro is rotated by 180° to invert the sensing axis, after which another set of stationary data are measured. Computing the difference between the average values of these data sets cancels any constant biases while keeping constant inputs, such as the Earth rate, intact. The effect of indexing is demonstrated in Fig. 3.2.

More complex carouseling patterns with rotations made simultaneously around two axes with applications in gyrocompassing were investigated in [57] where the data were recorded continuously while the sensor was slewed, as opposed to indexing. Rotating the sensor around two axes makes it possible to eliminate other
errors, such as scale factors, in addition to canceling the additive bias, although this is not crucial for gyrocompassing because the true input of an East-pointing leveled gyroscope is zero. Anyway, simulation results indicated that the two-axis rotation pattern yielded gyrocompassing results with lower variance than two-point carouseling.

Single-axis carouseling results for gyrocompassing were reported in [58] where a heading precision of 2 milliradians (approx. 0.1°) was achieved. However, the authors of [58] point out that this figure corresponds to heading estimation error standard deviation only, i.e., the absolute heading estimation accuracy was not investigated; the precision was estimated experimentally by means of regression analysis instead of, e.g., closed-form variance propagation.

In [59], carouseling was investigated for PDR purposes where a foot-mounted IMU was continuously rotated back and forth around a single axis. Applying the rotation decreased the maximum position error in the tests considerably, but as the authors of [59] point out, the effect depends on the shape of the trajectory; for instance, the effect of a heading bias is very different on a rectilinear trajectory than when walking in circles.

### 3.5 Relation of Thesis to Related Work

The calibration approaches studied in this thesis were based on knowledge that the sensor is stationary. In [II], the 1/f properties of a MEMS gyroscope were mod-

![Fig. 3.2: The effect of indexing on a leveled gyro at latitude 60°](image-url)
eled using a fractional Kalman filter [60] and by computing moving averages of the sensor output. In order to reduce the effect of random walk, a temperature-controlled IMU [15] was used. As opposed to black-box models [53, 54], these predictors were derived based on the statistical properties of $1/f$ noise; the proposed models also use a longer history of samples than common AR based predictors [48–52] to reflect the slowly decaying correlation of $1/f$ noise. A wavelet domain approach to filtering $1/f$ noise was proposed in [47] whereas [II] investigated time domain methods. An error model that does not account for the correlation structure of $1/f$ noise will fail when operating over prolonged periods of time [37]; as future work, the proposed bias predictors should be compared with, e.g., AR models using a very long set of data to verify the significance of modeling the $1/f$ behavior.

In [I], a MEMS gyro [61] was used for gyrocompassing, unlike [56] where the inverse problem, i.e., estimating the Earth rate given the orientation of the sensor, was studied. The gyrocompassing system of [I] was based on an EKF and indexing; the latter avoids the need of multiple gimbals and other complex mechanics as opposed to [57] where the gyrocompassing performance was improved by carouseling around two axes, and the EKF can provide instantaneous estimates about the accuracy of the estimated North direction throughout the entire gyrocompassing process. A mechanical setup corresponding to that of [I] was used in [58] but with a carouseling approach instead of indexing; unfortunately, absolute heading estimation accuracy figures were not reported in [58].
3. Inertial Sensor Error Estimation and Calibration
4. USING MAP INFORMATION IN INDOOR NAVIGATION

In most vehicular and urban pedestrian scenarios, the position or route of the user is confined by the road network or the room layout of a building. Usually, the position estimates are not accurate enough to match the correct area, such as a street; therefore, for instance, in order to provide navigation directions for a car driver, the map constraints need to be taken into account in the position computation process. The incorporation of map information into position estimation is known as map matching. In this chapter, we first discuss the choice of a filter for implementing map matching, and then present map matching methods. Finally, initialization approaches for an indoor map matching filter are studied.

4.1 Choice of Filter

If the state or measurement models are significantly nonlinear or the underlying distributions are far from Gaussian, using Kalman-type filters can be somewhat risky. Multimodal distributions are a common counter-example: typically, a Kalman-type filter will follow one of the modes but cannot recover if that mode turns out to be false in light of later observations. One approach to cope with such situations is to use a filter based on Gaussian mixture distributions, i.e., convex combinations of individual Gaussian densities [62, 63].

However, some types of information are, by nature, impractical to model as a combination of Gaussian distributions. For instance, incorporating map information often leads to binary measurement likelihoods such as “the user crossed a wall”; these models will be addressed in more detail in Section 4.2. In such situations, one possible solution is to use Monte Carlo methods where the posterior distribution is represented using a set of weighted random samples; an example scenario is shown in Fig. 4.1 where the sample weighting is binary (‘alive’ or ‘discarded’). The particle filter (PF) [64], also referred to as sequential Monte Carlo sampling,
is a numerical approximation of the Bayesian filter where the samples are state vector candidates which are weighted based on their measurement likelihoods. It has gained popularity probably because it is simple and intuitive to implement and because of the increase in computational power available in many types of devices.

Representing the distribution using a set of samples, or cloud of particles as they often are referred to in the context of PFs, has the advantage that the types of the underlying distributions are not limited; they can be Gaussian, multimodal, uniform, or practically anything else. The performance of Monte Carlo computations depends on the amount of samples used, but so does its computational complexity; evaluating the state transition and measurement equations for dozens or even thousands of samples is much more demanding than propagating the mean and covariance of the posterior distribution using a Kalman-type filter.

4.2 Map Matching Methods

The most straightforward approach to map matching is to first resolve the user position ignoring the map and then project the estimate onto the closest admissible location such as a road link [65]. Since the position estimation is done totally inde-
4.2. Map Matching Methods

Another significant difference between vehicular navigation on roads and (pedestrian) navigation indoors is the fact that the heading of a vehicle is somewhat constrained by the orientation of the road link whereas one can wander around quite freely in a room. Therefore, map information can be used directly to prevent the heading drift of a DR system in vehicular applications. However, in this thesis, we will focus on indoor scenarios where the heading is not constrained in any way.

In [66], the applicability of Kalman-type filters on indoor map matching was investigated. However, the maps were coarse in detail, consisting of the outer walls of the building only. One of the approaches used in that article was extended in [67] to accommodate room-level constraints using a link–node transition model between rooms; for each room, the state vector contained a component representing the probability of the user being located in that area. Using WLAN observations as the sole measurement information, a Kalman-type filter was observed to achieve a positioning performance comparable to a PF with a significantly lower computational complexity. However, an obvious drawback is the fact that the number of states increases with building size and map granularity; furthermore, incorporating inertial sensors into the system would require at least a revision of the state model.

In the literature, many solutions for fusing an IMU with indoor maps using a PF have been proposed. For example, such a system with three-dimensional indoor maps was presented in [68]. Practically all solutions share the same principle of incorporating the map constraints: a binary function

$$h_t(x_t) = \begin{cases} 0 & \text{if the particle has crossed a wall} \\ 1 & \text{otherwise} \end{cases}$$

is used as the measurement model in (2.7). From the perspective of operating with
a set of weighted samples, such a measurement function implies that wall-crossing particles are discarded or at least significantly down-weighted, which is intuitive.

4.3 Fusion Filter Initialization Approaches

A cascaded filter approach for combining map matching and PDR was proposed in [69]; the idea was to run an EKF at a high update frequency to estimate DR errors and a PF at a lower update frequency to enforce map constraints. The IMU was foot-mounted, and thus, a Kalman-type filter is a natural choice for estimating the accumulating errors as they are commonly used in 6DOF mechanizations as well [14]. The case where two IMUs are used, one on each foot, was also studied; it was shown that such a setup halves the estimation error variance with respect to the case of a single IMU, but obviously doubles the amount of sensors needed.

In [70], a map matching approach for a foot-mounted IMU was presented along with a “backtracking” scheme. The idea of backtracking is to refine previous state estimates using the current information. This approach is commonly known as smoothing in estimation literature; for the case of the Kalman filter, closed-form smoothing equations exist [71] whereas particle smoothing methods, such as [72], require memorizing previous particle sets, which can be quite memory-intensive. However, smoothing does not yield any new information about the most recent state estimate, which is of the greatest interest in most positioning and navigation applications. Therefore, smoothing is not studied in the scope of this thesis.

However, in [69, 70], the initialization of the filter was not focused on and a brute-force approach was used. Initializing the PF with a massive amount of particles to accommodate the entire building imposes a heavy computational load if the building is large and consists of many floors. Thus, in practice, other position information is required to constrain the initial particle distribution; RF-based measurements are the most popular choice. In [73], a DR and map matching scheme using body-mounted sensors was investigated assuming that the initial position is known fairly precisely (with 1 m standard deviation in the coordinates).

A WLAN based initialization scheme was investigated in [68] and extended to make use of continuous WLAN updates in a subsequent paper [74]. The WLAN positioning algorithm was based on a fingerprinting approach; the authors demonstrated
a relatively fast method of constructing the radio map of a building, but nevertheless, this task needs to be carried out offline.

Simultaneous localization and mapping (SLAM) is a classical problem in the field of robotics [75] and has been demonstrated to be able to use WLAN observations as landmarks [76–78]. It has even been shown to be feasible for room layout mapping using PDR only [79]; however, the localization and mapping in SLAM are usually done in terms of an arbitrary reference frame, which is incompatible with the case where the map and coordinate system are obtained from elsewhere. Although it is possible to align a SLAM-generated map with a global map, this does not necessarily lead to a unique solution if the map landmarks are not uniquely identifiable, which can be the case for a relatively short trajectory in a building whose room layout exhibits significant symmetries [80].

In [81], dedicated Bluetooth transmitters were used instead of WLAN APs as sources of radionavigation measurements. Bluetooth signals are expected to be less prone to outlier measurements in large open halls than WLAN observations because Bluetooth signals are transmitted at a significantly lower power, thus reducing the propagation range. On the other hand, shorter-range transmitters need to be more densely deployed, and they are not readily available as part of other infrastructure as opposed to WLAN APs. It is noteworthy that the transmitters used in [81] consist of multiple patch antennas mounted on a hemispherical surface to render the angle of transmission observable along with the TOA, which can be beneficial when the transmitter is located in a large open hall.

4.4 Different Levels of Map Information Detail

The effect of indoor map detail level on positioning accuracy is a relevant question and has been studied in [70,82]. Obviously, the more details the map has, the more accurate constraints can be imposed on the particles, but the availability of a very accurate map cannot be taken for granted. For instance, in [83] investigated the possibility of constructing the map from a photograph of a building's evacuation plan. An interesting feature of evacuation plans is the fact that the current location is always indicated on the map, but unfortunately not with a standardized symbol; other problems are the lack of scale and global coordinate frame information.
An extreme case is to only use the outer walls of the building as constraints; such information can be extracted, e.g., from an aerial photograph. In addition to using them directly for map matching according to (4.1), other algorithms have been proposed. Assuming a rectangular room layout aligned parallel to the outer walls, the average heading of a person walking along a (long) hallway can be constrained to one of four principal directions [84]. Significant improvements in positioning accuracy in comparison with a stand-alone foot mounted IMU are reported, but an obvious downside of the method is that it relies heavily on assumptions about the building geometry, and a method for verifying their validity needs to be developed.

A set of wall segments is not the only possible representation of indoor map information. For instance, the position of the user in a building can be constrained onto a finite connected graph of possible positions, which resembles the concept of a road network map [85]. Such a structure is useful for, e.g., route planning purposes because standard graph search algorithms, such as Dijkstra’s [86], are directly applicable, but the same is possible with a wall segment representation as well if a room structure is identified.

4.5 Relation of Thesis to Related Work

In [III], map matching was used to augment PDR for indoor navigation. The initial position distribution was obtained using a deeply coupled high-sensitivity GNSS and IMU system [87] with support for coherent integration over several seconds to overcome heavy signal attenuation; thus, no WLAN APs [74], dedicated short-range transmitters [81], or other local infrastructure was used, yet the number of particles was relatively low in comparison with, e.g., [68, 69]. However, the altitude accuracy of high-sensitivity GNSS was inadequate for identifying the floor in a multi-storey building. Therefore, other means, e.g., a barometer, are needed to resolve the floor number; with short-range RF measurements, the floor is easier to identify. Moreover, the high level of noise in high-sensitivity GNSS measurements causes a need of filtering the initial position solution over a period of time instead of being able to operate in a snapshot manner such as short-range RF measurements, and there are places where GNSS measurements are simply impossible, such as underground facilities.
Furthermore, the IMU was body-mounted in [III] as opposed to [68–70] where foot-mounted sensors were used. It was seen that even though PDR with body-mounted sensors is not as accurate as with a foot-mounted IMU, the relatively coarse GNSS solution was sufficient for PF initialization in the absence of initial heading knowledge without resorting to an unreasonable number of particles; the scenario with body-mounted sensors and precise knowledge of initial position has been studied earlier in [73]. The capability of converging with relatively imprecise initial conditions is mainly due to the efficient constraints given by the detailed map, which limit the accumulation of both heading and along-track errors. As opposed to SLAM [79], a readily available map was utilized in [III] instead of constructing it on the fly.
4. Using Map Information in Indoor Navigation
Positioning accuracy estimates are only valid when the assumptions made about the underlying measurement models hold. Therefore, it is essential to screen the data to detect and exclude biased observations. In this chapter, this problem is studied from the perspective of carrier phase based positioning where a major source of errors are cycle slips.

As mentioned in Chapter 2, a fundamental property of the carrier phase observable is the fact that the integer ambiguity term $N^S$ remains constant as long as a phase lock is maintained on the signal. However, if the phase tracking is interrupted for even a fraction of a second, e.g., because of high receiver dynamics or temporary signal shadowing, the integer ambiguity may change; this phenomenon is illustrated in Fig. 5.1. Cycle slips can also be caused, e.g., by tracking errors due to a low $C/N_0$ or a bug in the receiver software [88].

Because the signal wavelength at the L1 frequency is approximately 19 centimeters while the carrier phase tracking noise is in the order of millimeters, it is clear that

![Fig. 5.1: Cycle slips in the carrier phase measured by a stationary GPS receiver](image-url)
a cycle slip of even a couple of carrier cycles—or in the most subtle case, a half cycle—will cause a significant loss of precision. Therefore, it is important to detect them. Fortunately, if correctly detected, cycle slips can be corrected for because of the discrete nature of the integer ambiguity.

In this chapter, we discuss two approaches suitable for detecting cycle slips using a stand-alone single-frequency receiver in real time: computing measurement differences and predicting the carrier phase values.

5.1 Differential Measurements

In order to make the cycle slips better detectable, other sources of variation of the carrier phase measurement in (2.2) need to be eliminated. In the absence of external information, the most straightforward way is to attempt to cancel the sources of variation by computing measurement differences. An obvious strategy is to take the difference of two consecutive carrier phase measurements: if a cycle slip took place between them, it can be seen as a peak in the difference. An example is shown in Fig. 5.2. Otherwise, the difference is a scaled version of the Doppler measurement in (2.3).

A cycle slip detection algorithm based on comparing the time-differenced carrier phase with the native Doppler measurement was presented in [89] where an integer optimization technique [90] were used to estimate the magnitudes of the slips; the algorithm was reported to have succeeded in 50–75 % of the time in field tests. A similar approach was taken in [91] where time-differenced pseudorange observations were used along with Doppler and time-differenced carrier phase values.

Higher order measurement differences are possible as well. One approach for finding an optimal order of differentiation was proposed in [92]; the idea was to find a trade-off between the amplification of random noise and the cancellation of slowly varying components. This way, cycle slips can be detected using simple threshold criteria on the differenced data, but the authors point out that the method is not efficient at detecting small cycle slips.

Real-time kinematics (RTK) is an important application based on the carrier phase measurement, and many cycle slip detection algorithms have been developed specifically for its purposes. The goal of RTK is to estimate the position of the receiver
with respect to a reference receiver whose location is precisely known. Computing the difference between the carrier phase observations made at the different receivers cancels a significant part of the variation in the carrier phase due to satellite motion; when further differenced in time and between different satellites, cycle slips are usually the most significant source of variation left [93], as illustrated in Fig. 5.2b; note the appearance of a new cycle slip because of the incorporation of another satellite and another receiver. If dual-frequency measurements are available, as often is the case in RTK, cycle slip detection can also be facilitated by combining the different frequencies in a suitable way, e.g., as described in [94].

5.2 Phase Prediction Approaches

Instead of only using the most recent measurements for detecting cycle slips, it is also possible to compare the current measurements to values predicted based on other measurement epochs. A straightforward approach is to use polynomial interpolation or regression [95]; the problem with interpolation is that it needs to know the values of future measurements, thus preventing use in real-time applications. Another problem in such methods is how to avoid the use of erroneous measurements for constructing the regression estimate.

Using a KF as a measurement predictor is a common approach in the quality control of pseudorange based positioning [96] and has been applied for error detection in RTK [97, 98]. The problem with filter-based error detection is that if an undetected erroneous observation is used to update the state estimate, the false
information will affect all subsequent error detection attempts; this is most critical during the initial convergence period of the filter. This scenario was accounted for in [97] by modeling and statistical testing of the occurrence of such a fault.

A cycle slip detection method for a stand-alone receiver based on a Precise Point Positioning (PPP) filter was introduced in [91] along with the time-differential approach; PPP is based on using precise correction data for as many error sources as possible, which can be utilized for cycle slip detection without the need for computing measurement differences. The time-differential approach was observed to outperform the PPP based method, but the authors of [91] propose that it would be possible to use both simultaneously.

If inertial sensors are available, they are a suitable means for predicting the evolution of the carrier phase. Such an approach has been studied for RTK purposes in many articles, e.g., [99–101], and an algorithm for stand-alone dual-frequency receivers was proposed in [102].

The use of an artificial neural network for cycle slip detection was studied in [103]; the method also made use of a wavelet transform whose applicability on cycle slip detection has been studied earlier [104]. However, the test results presented in [103] were not extensive enough to demonstrate the performance of the algorithm.

5.3 Relation of Thesis to Related Work

In [IV], cycle slips were detected from first-order carrier phase time differences by verifying their consistency. Such methods are commonly applied on pseudorange based positioning in the RAIM process [105, 106]. The algorithm presented in [IV] is based on observations on a single frequency and did not assume the availability of inertial sensors or a reference receiver, which gives a wider applicability than RTK-specific algorithms; the method was tested with a time-differential flight path reconstruction algorithm [107], and implementing PPP with a single-frequency receiver has been a popular research topic during the past years [108].

The approach of [IV] consists of two steps, i.e., detection and identification, and the latter is only performed if errors are detected. Thus, it is computationally lighter than the more recent algorithms [89, 91] which need integer optimization.
On the other hand, RAIM methods are susceptible to failure or degraded error detection resolution when multiple errors occur simultaneously [109]; integer optimization methods might perform better in this sense.
5. Cycle Slip Detection for Low-Cost GNSS Receivers
6. CONCLUSIONS

In this chapter, conclusions regarding the research are drawn. We first summarize the results obtained and discuss them, and finish with suggestions for future work.

6.1 Summary of Results

In [I], a two-sigma gyrocompassing accuracy of 4.03° was obtained with an experimental setup consisting of a leveled MEMS gyroscope and a step motor to implement indexing. Eight hours of time were required for the filter to achieve this accuracy, but this is affected by many factors, such as the initial heading uncertainty and the choice of the indexing period. In particular, it is noteworthy that the experiments were conducted at a relatively high latitude (approx. 60.5°) where the horizontal component of the Earth rotation constitutes less than half of the whole rotation rate; a better performance would be achieved closer to the Equator.

Simulations and experimental tests conducted in [II] suggested that taking into account the $1/f$ properties when predicting the additive bias of a MEMS gyroscope decreases the root-mean-square prediction errors. As the correlation function of $1/f$ noise decreases very slowly, optimal prediction of the process becomes computationally expensive; for this reason, a suboptimal but simpler predictor based on moving averages was also investigated.

Initializing a map matching and PDR fusion filter using high-sensitivity GNSS measurements was studied in [III]. It was found out that once the GNSS solution had converged, it could be utilized for initializing the map matching particle filter without the need for further position updates. However, because of the imprecision of the GNSS position estimate and no knowledge of the initial heading, a relatively large amount of particles was needed for the PF to converge to the correct position and heading, but the number of particles was still significantly smaller than
6. Conclusions

in the approach of covering the entire map with particles to initialize the filter.

In [IV], an algorithm for cycle slip detection using a stand-alone single-frequency GNSS receiver was proposed. Test results based on authentic flight data suggest that the method can detect even small cycle slips, but due to the nature of the test data, the detection results could be verified by hand in less than 30% of the cases only, thus hindering any definitive conclusions. Anyway, in the verifiable cases, the performance of the method was comparable to an RTK based algorithm [98]; the method yielded erroneous results most prominently when multiple cycle slips occurred at one epoch.

6.2 Discussion

According to the results obtained, the performance of low-cost MEMS sensors can be significantly improved with suitable calibration and assistance. The gyrocompassing results, achieved with a setup of one gyro and a step motor, do not yet meet the requirements of commercial gyrocompassing applications such as borehole surveying where the time to reach a degree-level accuracy should be a couple of minutes. Nevertheless, the results demonstrate a significant performance potential: although carouseling around one axis does need additional mechanical components, it could be feasible for vehicular navigation where the improvement in heading estimation stability could be significant in long-term autonomous DR when combined with a ground speed sensor.

Obviously, carouseling is not practical in hand-held-size equipment. In that case, the sensor error accumulation can be mitigated by error process modeling. The $1/f$ noise based bias predictors studied in this research suggest that modeling the bias instability is beneficial, but they require a separate treatment for the rate random walk process which becomes more significant when the time between calibration points, i.e., stationary periods, becomes longer. Using other information, such as RF measurements or a map, can decrease the requirements of sensor modeling accuracy; for instance, indoor map constraints can be efficient enough to work with very simple gyro error models.

With map matching, the major problem is how to find the initial position and heading. While this is not usually difficult outdoors where strong GNSS signals
are available, other means, such as WLAN AP position information or specialized GNSS receivers, are required to accomplish the task indoors without the need to evaluate the filtering equations for an excessive amount of particles. However, when a map and inertial sensors are available, the frequency of GNSS updates could be reduced; with a hardware-optimized PDR and map matching implementation, the reduced need for GNSS processing could lead to savings in power consumption [110, 111].

Although single-frequency RTK and PPP have been popular research topics during the past years and some correction data for PPP can now be obtained in real time, carrier phase based positioning applications have not become popular in the consumer market. The difference in antenna quality alone makes it unreasonable to expect miniature receivers to achieve the same performance as professional equipment, but sub-meter accuracy would still be useful in certain consumer use cases: for example, it would make it possible for a personal navigation assistant to identify the current driving lane. Obviously, detecting cycle slips is essential in such applications.

The experimental results obtained with an off-the-shelf commercial receiver, and those of the related research, imply that detecting and correcting for cycle slips is feasible even with low-cost devices. However, more research is needed to improve the detection rate; for a stand-alone receiver, the incorporation of well-calibrated MEMS sensors and a detailed map could help to improve the measurement redundancy in order to make cycle slips better detectable.

6.3 Future Work

The PDR and map matching results, along with the observations made in related studies, suggest that, at least in some room layouts, RF updates are not critical after the initial convergence. However, it should be studied how much a tight or deep coupling of GNSS and the PDR and map matching filter could help the convergence to the correct position and heading and reduce the number of particles needed.

The importance of different information sources—GNSS, IMU, WLAN and other signals of opportunity, map, etc.—in different use cases is an interesting ques-
tion when considering, e.g., the robustness, power consumption, and likelihood of measurement availability. For safety-critical applications, such as the tracking of first responders, the requirements are obviously different than for a passenger who is trying to find his way in an airport terminal and does not want to drain the battery of his cellphone with frequent WLAN scanning and GNSS processing.
REFERENCES


References


[61] SCC1300-D02 combined gyroscope and 3-axis accelerometer with digital SPI interfaces, data sheet rev. 2.1, Murata Electronics Oy.


[70] Widyawan, M. Klepal, and S. Beauregard, “A backtracking particle filter for fusing building plans with PDR displacement estimates,” in 5th Workshop


References


[89] S. Carcanague, “Real-time geometry-based cycle slip resolution technique for single-frequency PPP and RTK,” in *Proc. 25th International Technical
References


[98] D. Odijk and S. Verhagen, “Recursive detection, identification and adaptation of model errors for reliable high-precision GNSS positioning and at-


© 2012 IOP Publishing Ltd. Reprinted, with permission, from
gyroscope to measure the Earth's rotation for gyrocompassing applications,” Measurement Science and Technology, vol. 23, issue 2, February 2012 (8 pp).
doi:10.1088/0957-0233/23/2/025005
Using a MEMS gyroscope to measure the Earth’s rotation for gyrocompassing applications

L I Iozan1, M Kirkko-Jaakkola2, J Collin2, J Takala2 and C Rusu1

1 Technical University of Cluj-Napoca, Baritiu 26-28, 400027 Cluj-Napoca, Romania
2 Tampere University of Technology, PO Box 553, 33101 Tampere, Finland
E-mail: Lucian.iozan@bel.utcluj.ro, martti.kirkko-jaakkola@tut.fi, jussi.collin@tut.fi,
jarmo.takala@tut.fi and Corneliu.Rusu@bel.utcluj.ro

Received 16 May 2011, in final form 28 November 2011
Published 11 January 2012
Online at stacks.iop.org/MST/23/025005

Abstract
In this paper, a method and system for gyrocompassing based on a low-cost micro-electro-mechanical (MEMS) gyroscope are described. The proposed setup is based on the choice of a gyroscope with specified bias instability better than 2 deg h\(^{-1}\) and on careful error compensation. The gyroscope is aligned parallel to the local level, which helps to eliminate the g-sensitivity effect but also sacrifices a fraction of the Earth’s rotation rate that can be observed. The additive bias is compensated for by rotating the sensor mechanically and by extended Kalman filtering. In this paper, it is demonstrated that the proposed system is capable of observing the Earth’s rotation, and the north finding results show that a two-sigma accuracy of 4.03\(^{\circ}\) was attained at latitude 61\(^{\circ}\)N. With current MEMS gyroscopes, the system requires hours of time to achieve this accuracy, but the results demonstrate the theoretical accuracy potential of these small self-contained, low-cost sensors.

Keywords: gyrocompass, MEMS gyroscope, bias compensation

1. Introduction
Accurate heading is a key factor in many location systems and is typically obtained by using digital magnetic compasses (DMCs). These are compact, low-cost instruments capable of achieving sub-degree heading accuracies [1]. However, the heading accuracy of DMCs can be very easily degraded by nearby ferrous materials or by electromagnetic interference. For these reasons, DMCs are not very reliable indoors. In contrast, inertial gyrocompasses are unaffected by such factors and capable of maintaining a high level of accuracy regardless of the environment [2, 3]. Gyrocompasses operate by observing the Earth’s rotation rate. This requires highly accurate sensors, and only recently have micro-electro-mechanical system (MEMS) gyroscopes been considered accurate enough to measure it [4, 5]. The ability to measure such slow rotation using low-cost, low-power, small-size sensors is very interesting from both academic and application viewpoints.

In this paper, we discuss the most significant error sources that degrade the performance of MEMS gyroscope sensors and propose ways of compensating for them in the context of gyrocompassing. This is done in order to maximize the performance of these sensors, which are usually designed as electronically driven resonators, often manufactured from a single piece of quartz or silicon. Such gyroscopes operate in accordance with the dynamic theory, which states that when an angular rate is applied to a translating body, a Coriolis force is generated [2, 6]. The force is proportional to the applied angular rate and causes a displacement that can be measured capacitively in a silicon instrument or piezoelectrically in a quartz instrument. Theoretically, if the errors of such gyroscopes are properly modeled and compensated for, it is possible to measure very small angular rates, like the Earth’s rotation rate. In practice, however, this requires averaging a long sequence of observations, and if the bias instability of the gyroscope is not good enough, the small angular rate will be buried under bias drift. Consequently, the sensor must be chosen carefully.
The direct measurement of the Earth’s rotation rate has been performed using a MEMS gyroscope in [4], while [7–9] report progress in the development of high-accuracy MEMS gyroscopes with north seeking capabilities. In [4], the sensitive axis of the gyroscope was parallel to the Earth’s rotation axis, and long-term measurements were carried out in order to compensate for external factors (e.g., gravity) from the measurement data; in [8, 9], the errors were eliminated by means of carouseling.

This paper proposes an approach where the sensitive axis of the gyroscope is always parallel to the Earth surface. There are two main reasons for this: first, when seeking north, we are interested in the horizontal direction; thus, by keeping the sensor aligned with the local level, we need only rotations around a single axis, which keeps the mechanics simple. Secondly, the horizontal plane is perpendicular to the gravity vector, which causes the effect of $g$-sensitivity to remain constant after each rotation; this way, it is more straightforward to eliminate. On the other hand, the component of the Earth’s rate that can be detected by a leveled sensor depends on the latitude and is less than half of the true rate in our tests at latitude 61.449° N. Furthermore, we eliminate the additive bias by periodical mechanical changes in the sensor orientation (also known as two-point gyrocompassing [9]), and the north direction and required correction values are estimated using an extended Kalman filter (EKF).

The paper is organized as follows. In section 2, the theoretical background including the EKF model and an analysis of the attainable accuracy is presented. The measurement setup is described in section 3, whereas section 4 demonstrates the theoretical and practical north finding capability of the proposed system. Section 5 concludes the paper.

2. Theoretical background

According to the World Geodetic System 1984 (WGS84) model, the magnitude of the Earth’s rotation rate is $\Omega_0 = 7292115 \times 10^{-11} \text{ rad s}^{-1} \approx 4.178 \times 10^{-3} \text{ deg s}^{-1}$ [10]. Since the input range of a typical MEMS gyroscope is approximately $\pm 100 \text{ deg s}^{-1}$, it is clear that measuring the Earth’s rotation using such equipment requires a careful error analysis and filter design. Scale factor errors, typically in the order of 1–2% and possibly asymmetric, do not have a significant effect with such weak signals, and temperature effects can be mitigated by keeping the unit at a constant temperature and allowing it to warm up prior to the tests. In this paper, we will concentrate on the most significant error sources, i.e. angular rate random walk, in-run bias instability and the temperature-dependent bias. A good measure to characterize the first two errors is the Allan variance, which is described in the following section.

2.1. Allan variance

Named after Dr David W Allan, the Allan variance is a measure for characterizing noise and stability in clock systems [11]. It analyzes a sequence of observations and estimates the intrinsic frequency noise of the system as a function of averaging time.

\[ \text{AVAR}(\tau) = \frac{1}{2(n-1)} \sum_{i=1}^{n} \left[ y(\tau)_{i+1} - y(\tau)_i \right]^2, \]  

(1)

Although the method was originally developed for clocks, it can easily be adapted for different types of data.

The computation of Allan variance starts by taking a long sequence of data and dividing it into disjoint bins based on averaging time \( \tau \). The equation for the Allan variance is as follows:

\[ \text{AVAR}(\tau) = \frac{1}{2(n-1)} \sum_{i=1}^{n} \left[ y(\tau)_{i+1} - y(\tau)_i \right]^2, \]

where \( \text{AVAR}(\tau) \) is the Allan variance as a function of the averaging time, \( y_i \) is the average value of the measurements in the bin \( i \) and \( n \) is the total number of bins. Analogously to statistical terminology, the square root of Allan variance is known as the Allan deviation.

Figure 1 shows three sets of gyroscope data and their Allan deviation plots; the relation between precision and Allan deviation is easy to see. For short averaging times, the Allan deviation is dominated by uncorrelated noise in the sensor output; in fact, the power of white noise is directly related to the value of the Allan deviation at small \( \tau \) [12]. Initially, the Allan deviation decreases as the averaging time gets longer. Then, at some point, the Allan deviation levels off because of $1/f$ noise [13], and eventually, with long averaging times,
it starts to increase. This is due to bias drift or rate random walk in the sensor output, a process which is clearly non-
stationary\(^1\) [12]. The power of 1/f noise is commonly used to define bias instability [14] and thus, the minimum value of
the Allan variance can be used as a measure of bias instability.

Allan variance is a very useful tool in the error analysis of
MEMS-based navigation where the 1/f noise and random walk components are typically very strong. For example,
consider a problem where the gyroscope bias needs to be
estimated from the output of a stationary gyroscope. If the
sensitivity axis of the gyroscope is aligned with the east–west
direction, it is known that the actual input is zero. To estimate
the bias, one can then obtain a sequence of measurements and
take the average; the minimum point of the sensor’s Allan
variance plot indicates the best choice of averaging time.
When the gyroscope output is used to compute the attitude,
the estimated bias is subtracted from the measurements and
the difference is integrated (i.e. summed as the measurements
are discrete) to obtain the change in orientation over time.
The estimate of the bias obtained as described above is not
necessarily optimal, as the error process is correlated in time
[15], but suffices for our purpose.

2.2. Modeling the gyroscope measurement
In this study, the aim is to observe the horizontal component
of the Earth’s rotation rate in order to find the true north
orientation using a single gyroscope sensor aligned parallel
to the local level. The output of the gyroscope is an angular
rate measurement which we will model as

\[
\omega^+ (\varphi, \psi) = \Omega_e \cos \varphi \cos \psi + b \quad (2a)
\]

\[
\omega^- (\varphi, \psi) = -\Omega_e \cos \varphi \cos \psi + b, \quad (2b)
\]

where \(b\) is the gyroscope bias and \(\varphi\) is the latitude. The
observable component of the Earth’s rate depends on the
latitude because of the horizontal alignment of the sensor and
on the angle \(\psi\) between the sensing axis and north direction.
For example, when the sensing axis is parallel to the east–
west direction (\(\psi = \pm 90^\circ\)), the Earth’s rotation is not sensed
at all.

Model (2b) is obtained by rotating the measurement setup
by 180\(^\circ\) and knowing that \(\cos (\psi + 180^\circ) = -\cos \psi\). The
motivation for the rotation is to be able to cancel the gyroscope bias:
for each measurement orientation \(\psi\), the magnitude of
the Earth’s rotation rate \(\Omega_e\) can be estimated using (2a) and
(2b) as follows:

\[
\Delta \omega (\varphi, \psi) = \frac{\omega^+ (\varphi, \psi) - \omega^- (\varphi, \psi)}{2} = \Omega_e \cos \varphi \cos \psi + \delta b,
\]

(3)

where the term \(\delta b\) has been introduced to denote the
possible change in the gyroscope bias \(b\); since the opposite angular rate measurements (2a) and (2b) cannot be
made simultaneously, the gyroscope bias \(b\) may have fluctuated
between the measurements. In particular, the biases of MEMS
gyroscopes are sensitive to temperature changes [16].

It is noteworthy that since the sensing axis of the gyroscope is perpendicular to the gravitational acceleration
vector \(g\) in both (2a) and (2b), the components of \(g\) at
the sensing and drive axes of a vibratory gyroscope remain
unchanged when rotating the gyroscope from model (2a) to
(2b) [2], and therefore the \(g\)-sensitivity effect is eliminated
in difference (3). This is advantageous as the impact of the
\(g\)-sensitivity on the measurements of a MEMS gyroscope can be
significant [5].

As discussed in section 2.1, the noise process of a MEMS
gyroscope is complicated. It has a 1/f component and another
part of it is uncorrelated (white). However, modeling the
1/f noise component is tricky and can lead to complex
fractional-order state models [17]; therefore, we approximate
the noise process as a combination of white Gaussian noise and
random walk. The random walk component is modeled in the
unknown \(\delta b\).

2.3. Estimating the north direction

Given the latitude \(\varphi\) and the Earth’s rate \(\Omega_e\), the offset
angle \(\psi\) can be estimated from (3). In this paper, an EKF
is used to estimate the offset angle \(\psi\) from a sequence of
such measurements with known shifts in \(\psi\) between the
observations. The filter estimates a bivariate state vector

\[
x_k = \begin{bmatrix} \psi_k \\ \delta b_k \end{bmatrix}
\]

(4)

corresponding to the unknown quantities in (3) and \(k\) denoting
the time. The offset angle \(\psi\) is modeled as piecewise constant
with deterministic changes due to intentional rotations of the
gyroscope setup, while \(\delta b\) is modeled as random walk plus
linear correlation with the ambient temperature. The state
propagation is modeled in discrete time as

\[
x_k = x_{k-1} + \begin{bmatrix} \gamma_k \\ 0 \end{bmatrix} + q_k,
\]

(5)

where \(\gamma_k\) denotes the gyroscope orientation change between
the measurements \(k\) and \(k-1\) and \(q_k\) is a zero-mean Gaussian
random vector with known covariance.

For updating the filter, we assume that temperature
information is available along with the gyroscope output.
We take the differenced angular rate (3) and the average
sensor temperature difference between the measurements as
the observations, yielding the Jacobian matrix

\[
H = \begin{bmatrix} -\Omega_e \cos \varphi \sin \psi & 1 \\ 0 & \theta \end{bmatrix},
\]

(6)

where \(\theta\) is the temperature coefficient of the gyroscope bias.
In this study, the value of \(\theta\) was determined empirically.
Given initial conditions and the state propagation and measurement
noise covariances, we can now construct an EKF [18].

It should be noted that the offset angle obtained using
(2a) and (2b) is ambiguous: these equations do not give
information about the sign of the angle \(\psi\). Thus, the quadrant
of \(\psi\) must be resolved when the initial state vector \(x_0\) and its
covariance are determined for the EKF. This can be done, e.g.,
by rotating the sensor 360\(^\circ\) at a certain spacing and fitting a
sine curve with known frequency but unknown phase to the measurements. However, this approach needs a considerable number of samples in order to work accurately enough, which means that it requires a long time to determine the initial state.

### 2.4. Error analysis

As long as the quadrant is known, the orientation of the gyroscope can be resolved from (3) as

\[
\hat{\psi} = \arccos \left( \frac{\Delta \omega(\psi, \hat{\psi})}{\Omega_\epsilon \cos \psi} \right) = \arccos \left( \cos \psi + \frac{\epsilon}{\Omega_\epsilon \cos \psi} \right),
\]

where \( \hat{\psi} \) denotes the heading estimate and \( \epsilon \) is the uncompensated measurement error caused by gyroscope bias instability and noise. Computing the first-order Taylor series one would first collect a set of EKF covariance propagation equations. According to (8), the heading error due to uncompensated measurement errors is at its smallest when the sensor is in an east–west orientation (i.e. \( \cos \psi = 0 \)), and the error is amplified as the sensor approaches the north–south direction.

When filtering the heading using the EKF, the precision of the orientation estimate improves with time, until it reaches a steady state. Given the variance of \( \epsilon \) (essentially, the gyroscope bias instability) and the precision of the rotary stage movements, it is possible to estimate the attainable heading variance. Since the steady-state covariance matrix is not changed after taking a new filtering step, we have [18]

\[
P = (I - KH)(P + Q), \tag{9a}
\]

\[
K = (P + Q)H^T(H(P + Q)H^T + R)^{-1}, \tag{9b}
\]

where the time step indices \( k \) have been omitted, assuming that the state transition and measurement uncertainty covariances \( Q \) and \( R \), respectively, are constant in time. Substituting (9b) into (9a) yields the discrete algebraic Riccati equation

\[
Q - (P + Q)H^T(H(P + Q)H^T + R)^{-1}H(P + Q) = 0,
\]

which can be solved for the steady-state covariance \( P \). It should be noted that the Jacobian \( H \) changes depending on the current heading estimate; therefore, we can use the Jacobian of the optimal situation \( \psi = 90^\circ \) to compute the lowest attainable estimate covariance. It is also possible to predict the lowest attainable covariance after \( n \) filter steps using the traditional EKF covariance propagation equations.

If the heading resolution algorithm were to be implemented as a batch least-squares solution instead of filtering, one would first collect a set of \( n \) measurements \( \Delta \omega(\psi, \psi_i + \psi) \) with \( i = 1, \ldots, n \) and known \( \psi_i \), and then use, e.g., the Gauss–Newton least-squares method to resolve the sensor orientation \( \hat{\psi} \). Assuming independent and identically distributed measurement errors with variance \( \sigma^2 \), the variance of the resulting heading estimate \( \hat{\psi} \) would be approximately equal to \( \sigma^2(J^TJ)^{-1} \) with \( J \) representing the Jacobian of the system of equations:

\[
J = -\Omega_\epsilon \cos \psi(\sin(\psi_1 + \hat{\psi}) \cdots \sin(\psi_n + \hat{\psi}))^T.
\]

Consequently, the variance of the heading estimate can be calculated as

\[
\text{var} \hat{\psi} \approx \sigma^2 \left( \Omega_\epsilon^2 \cos^2 \psi \sum_{i=1}^n \sin^2(\psi_i + \hat{\psi}) \right)^{-1},
\]

which is minimized when all the squared sine terms equal 1, i.e. if all the observations were made with the input axis pointing to the east or west. An advantage of the least-squares method would be that it does not require knowledge of the distribution of the initial state. On the other hand, the EKF natively adopts to the state propagation model (5): for example, it accounts for the random walk error caused by the imprecision of the rotary stage movements by increasing the variance of the heading estimate accordingly after each rotation.

### 3. Hardware implementation

In this section, the gyrocompassing system proposed in this paper is presented. Since the key component of a gyrocompassing system is, of course, the gyroscope, we first discuss the choice of this sensor. Then, other necessary components and the structure of the system are presented.

#### 3.1. Choosing the gyroscope

As discussed in section 2.1, bias instability is a major source of errors in the output data of a gyroscope sensor. For gyrocompassing applications, the bias instability of the gyroscope must be significantly lower than the Earth’s rate. Figure 1(b) presents the Allan deviations of three gyroscopes based on different technologies; these sensors are described in table 1. From the figure it can be seen that the bias instability of each of the three sensors lies well below the Earth’s rate (dashed line). However, relation (3) between the latitude, the orientation of the sensor and the observable Earth’s rate should be borne in mind; for instance, this study was conducted in an area where only half of the Earth’s rate is observable for a leveled gyroscope. Nevertheless, at an averaging time \( \tau = 1000 \text{ s} \), the Allan deviation of the SCC1300-D02 sensor is approximately 0.000 34 deg s\(^{-1}\) which is less than one-tenth of the theoretical Earth rate. Obviously, the closer to the poles

### Table 1. The gyroscopes appearing in figure 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Technology</th>
<th>Grade</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIST-350</td>
<td>MEMS, oven-conditioned</td>
<td>Mobile devices</td>
<td>[19]</td>
</tr>
<tr>
<td>SCC1300-D02</td>
<td>MEMS</td>
<td>Automotive</td>
<td>[20]</td>
</tr>
<tr>
<td>HG1700</td>
<td>Ring laser</td>
<td>Tactical</td>
<td>[21]</td>
</tr>
</tbody>
</table>

...
one goes in latitude, the better bias instability is required [2]; in the extreme case of $\psi = \pm 90^\circ$, there is no horizontal Earth’s rotation component and a north direction does not even exist.

According to figure 1, the ring laser gyroscope is clearly the best, which is not surprising. When comparing the two MEMS sensors, the SCC1300-D02 is better suited for gyrocompassing than the temperature-stabilized AIST-350 gyroscope because of the lower Allan variance; thus, we choose to use the VTI SCC1300-D02 in the gyrocompassing system. The most important performance values for this gyroscope are listed in table 2. It can be seen that the quantization of the sensor output is coarser than the Earth’s rate. However, the gyroscope noise will cause stochastic dithering of the measurement prior to the analog-to-digital conversion, and thus it should be possible to estimate the Earth’s rate given a sufficient number of samples [22].

### 3.2. Gyrocompassing system

The most essential hardware components of the measurement setup are described in table 3. In addition, for including the SCC1300-D02 sensor in the setup, a custom-printed circuit board was designed. In principle, the system is operated by a laptop computer which first reads data from the gyroscope, and then commands the rotary stage to rotate the sensor. A block diagram depicting the connections between the system’s main components in more detail is presented in figure 2, and the entire measurement system is illustrated in figure 3.

For processing the data and also for controlling the rotary stage, we used MATLAB R2008a software. Furthermore, in order to collect and save the data from the SCC1300-D02 sensor, a software application was developed in Microsoft Visual Studio 2008. This application uses the NI USB-8451 device as a hardware interface between the sensor and the laptop. The maximum sampling rate attained for reading the gyroscope output was approximately 1 kHz; this value was limited by the NI USB-8451 SPI interface.

In addition, the temperature of the gyroscope was measured at a sampling frequency of 2 Hz using the internal temperature sensor of the SCC1300-D02 gyroscope. The performance figures of the temperature sensor are not given in the gyroscope’s specification sheet, but experiments have shown that the standard deviation of the short-term temperature measurement noise is in the order of 0.1 °C in the test environment; the sensor temperature varied between 25 and 28 °C during an observation period of 4 h.

### 4. Experimental results and discussions

In the previous section, a measurement setup was proposed, and it was shown that its specifications meet the theoretical requirements of gyrocompassing. In this section, we demonstrate that the setup is capable of gyrocompassing in practice. First, we show that the gyroscope is precise enough to observe the Earth’s rotation; then, we evaluate its performance in the inverse problem, i.e. estimation of the orientation of the sensor given the Earth’s rotation rate. All the measurements presented in this section were conducted in an office environment in Tampere, Finland, at latitude $\phi \approx 61.449^\circ$ N.

#### 4.1. Measuring the Earth’s rotation

In order to implement a gyrocompassing system, the gyroscope must be able to measure the rotation of the Earth to a sufficient accuracy. This capability can be verified in an experiment where the gyroscope is set sequentially into four different orientations—namely the principal compass points—and the output is logged at each position for a certain time. Then, since the offset angles $\psi$ are known, the magnitude of the Earth’s rate can be estimated.

The experiment was carried out such that the sensor stayed at each orientation for 5 min and was then automatically rotated 90° clockwise to the next measurement position. This process was continued for several hours. The sensing axis remained

Table 3. Components of the measurement setup.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gyroscope</td>
<td>VTI SCC1300-D02 sensor [20]</td>
</tr>
<tr>
<td>USB/SPI interface</td>
<td>National Instruments NI USB-8451 device [23]</td>
</tr>
<tr>
<td>Power supply</td>
<td>Hewlett–Packard E3611A dc power supply, 0–20 V and 1.5 A</td>
</tr>
<tr>
<td>Regulator</td>
<td>Self-manufactured voltage regulator, 5 Vdc and 3.3 Vdc</td>
</tr>
<tr>
<td>Laptop</td>
<td>For reading (and saving) data and for controlling the rotary stage</td>
</tr>
<tr>
<td>Rotary stage</td>
<td>Velmex B5990TS rotary stage for changing the orientation of the gyroscope</td>
</tr>
<tr>
<td>Driver</td>
<td>For controlling the motor of the rotary stage</td>
</tr>
</tbody>
</table>

Figure 2. Block diagram showing the connections between the main components.
parallel to the local level, as depicted in figure 4, for the entire time. Finally, the collected data were post-processed using a Kalman filter [5]. The underlying models used in the filter are basically similar to the equations presented in section 2.2, but the unknowns, i.e. the state vectors, are different, hence the measurement matrices are too.

The resulting estimates of the Earth’s rate are presented in figure 5. Despite the fact that less than half ($\cos 61^\circ < 0.5$) of the Earth’s rotation could be sensed by the gyroscope due to its horizontal alignment, the Earth’s rate was still observed without difficulty. As can be seen from figure 5, the estimated Earth’s rotation is very close to the theoretical value; after 16 h of measurements, the $2\sigma$ deviation of the estimate is $0.1932 \times 10^{-3}$ deg s$^{-1}$, which is approx. 10% of the theoretical observable value $1.997 \times 10^{-3}$ deg s$^{-1}$. Arnaudov and Angelov [4] achieved an estimation error of 8.7%, but it should be noted that they were measuring the full Earth’s rate, i.e. not only its horizontal component, which is approximately twice as large as our theoretical reference. Furthermore, figure 5 shows that the $2\sigma$ interval is overestimating the actual error; hence, it can be concluded that given a long measurement time, the proposed setup can measure the Earth’s rate relatively well.

4.2. Gyrocompassing results

Since the SCC1300-D02 gyroscope was observed to be capable of measuring the Earth’s rotation, a gyrocompassing system was constructed using the EKF described in section 2.3. This system is similar to the system used for estimating the Earth’s rate; however, now the sensor is not rotated in a deterministic way (i.e. in 90$^\circ$ steps) but based on the measured Earth’s rate.

Although the word ‘gyrocompassing’ usually refers to seeking the north direction, it is not obvious that it would be the easiest direction to find. In fact, when the offset angle $\psi$ is zero, the Jacobian matrix (6) becomes singular and the covariance estimate sequence fails to converge; the Riccati equation (10) has no solution in that case. Figure 6 shows...
the differenced measurements (3) from a rotation of 450° spaced by 10° with $\omega^+ (\varphi, \psi)$ and $\omega^- (\varphi, \psi)$ both averaged for 5 min at each position. It can be seen that the slope of the measurements is at its steepest in the east–west direction; this is in accordance with the observation based on (8) that heading errors are minimized in this orientation. Hence, the system was tuned to seek east instead of north—obviously, it is trivial to find north given the east direction.

The EKF was used as a control loop for the Velmex rotary stage. The algorithm consists of two steps. First, the magnitude of the Earth’s rotation rate was measured according to (3). Then, the measurement was fed to the EKF and the resulting state estimate was used to rotate the gyroscope’s sensitivity axis toward east. Finally, the state estimate was propagated according to (5) with $\gamma_k$ equal to the rotation. This procedure was repeated for several hours. As discussed in section 2.4, this approach yields the minimal variance because it aligns the input axis of the gyroscope with the east–west direction.

Figure 7 shows the gyrocompassing results for two separate measurement runs. In both cases, the system was initially offset from the true east direction by 90°, and at each position, the gyroscope output was averaged 5 min for $\omega^+$ and another 5 min for $\omega^-$. The results show that the system reached the true east orientation after 2 h, but the convergence of the covariance estimate takes more time. In fact, the covariance matrix did not reach a steady state in 8 h, but it can be seen that the 2σ confidence intervals get narrower as time passes.

After 2 h in both the tests shown in figure 7, the true east direction stays within the 2σ bounds corresponding to the EKF covariance estimate, although at the end of case (a), the true value is fairly close to the boundary. Initially, the true orientation (zero error line) does not lie between the 2σ lines, which is due to too optimistic an initial covariance estimate. Nevertheless, the accuracy of the system was in both cases better than 5° in terms of both absolute accuracy and the 2σ bounds; the final 2σ deviation was 4.03° in both runs. In autonomous dead reckoning, a heading error of 5° would cause a cross-track error of around 9 m after 100 m of navigation; obviously, this accuracy is inadequate for long-term autonomous navigation. Furthermore, an initialization time of 8 h would be infeasible for most applications. However, in light of these results it is clear that gyrocompassing is possible using a low-cost MEMS gyroscope and a simple rotary stage even at relatively high latitudes.

5. Conclusions

In this paper, a gyrocompassing system based on a single low-cost VTI SCC1300-D02 MEMS gyroscope sensor was presented. The most significant error sources that deteriorate MEMS gyroscope measurements were compensated for by keeping the sensor aligned with the horizontal plane, by rotating it according to a certain sequence, and by error estimation using an EKF. This enabled observing very small angular rates, such as the Earth’s rotation rate.

Since the tests were conducted in an office environment, rapid temperature fluctuations did not occur. Therefore, temperature compensation based on the measured sensor temperature and a temperature coefficient $\theta$ was regarded as sufficient along with the bias compensation based on mechanical rotations. The temperature of the gyroscope can change rapidly during the initial warm-up phase, but...
this was not observed to have caused any problems in the experiments.

The mechanical structure of the experimental setup was fairly simple, consisting mainly of a leveled rotary stage with a gyroscope mounted on it, and a controller to control the stage. Thus, all necessary rotations were made around one single axis, avoiding the need for complex mechanics and gimbals. The horizontal orientation of the rotary stage could be verified, e.g., using accelerometers [2]; the SCC1300-D02 does include a three-axis accelerometer in addition to the gyroscope, but leveling is beyond the scope of this paper.

In the paper, it was shown that the SCC1300-D02 gyroscope is accurate enough to measure the Earth’s rotation. The results show that with the implemented EKF-based control loop, two-sigma accuracies of approximately 4° were attained. A drawback of aligning the sensor parallel to the local level is that a fraction of the observable Earth’s rate is lost, but it was proven that the proposed system can still work at a latitude of 61°. Anyway, the performance of the proposed system depends on the latitude.

As future work, more sophisticated initialization algorithms to obtain the initial state for the EKF and to resolve the sign ambiguity should be developed. In the initialization phase, the sinusoidal regression approach mentioned in this paper requires many samples in order to work reliably. The proposed setup needs hours of time to reach a five-degree-level accuracy, which is infeasible for real-life applications. However, in the future, as the MEMS gyroscope technology progresses, shorter and shorter averaging times can be used, enabling faster north seeking. In conclusion, the results obtained in this study demonstrate the accuracy potential that can be attained using a low-cost, thumbnail-size sensor with small power consumption.

References

[20] VTI Technologies 2010 SCC1300-D02 combined gyroscope and 3-axis accelerometer with digital SPI interfaces
[22] Aumnal Ö 1996 Turning interference and noise into improved resolution Measurement 19 41–8
[23] National Instruments 2007 NI USB-8451 data sheet
PUBLICATION II

© 2012 IEEE. Reprinted, with permission, from

M. Kirkko-Jaakkola, J. Collin, and J. Takala, “Bias prediction for MEMS gyroscopes,”
doi:10.1109/JSEN.2012.2185692

In reference to IEEE copyrighted material which is used with permission in this thesis, the
IEEE does not endorse any of Tampere University of Technology’s products or services. In-
ternal or personal use of this material is permitted. If interested in reprinting/republishing
IEEE copyrighted material for advertising or promotional purposes or for creating new col-
lective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.
Bias Prediction for MEMS Gyroscopes

Martti Kirkko-Jaakkola, Jussi Collin, and Jarmo Takala, Senior Member, IEEE

Abstract—MEMS gyroscopes are gaining popularity because of their low manufacturing costs in large quantities. For navigation system engineering, this presents a challenge because of strong nonstationary noise processes, such as $1/f$ noise, in the output of MEMS gyro. In practice, on-the-fly calibration is often required before the gyro data are useful and comparable to more expensive optical gyroscopes. In this paper, we focus on an important part of MEMS gyro processing, i.e., predicting the future bias given a calibration data with known (usually zero) input. We derive prediction algorithms based on Kalman filtering and the computation of moving averages, and compare their performance against simple averaging of the calibration data based on both simulations and real measured data. The results show that it is necessary to model fractional noise in order to consistently predict the bias of a modern MEMS gyro, but the complexity of the Kalman filter approach makes other methods, such as the moving averages, appealing.

Index Terms—$1/f$ noise, calibration, gyroscopes, microelectromechanical systems, navigation, stochastic processes.

I. INTRODUCTION

ALTHOUGH today’s MEMS gyroscopes are inferior to optical sensors from the accuracy point of view, their low cost, tiny size, low power consumption, and suitability to production in large quantities are undeniable advantages that have enabled their integration into a variety of low-cost consumer devices [1]–[3]. It is clear that the traditional inertial navigation mechanism [4] is not directly suitable for MEMS-based inertial units, but by adding on-the-fly calibration [5], nonholonomic constraints [6], and, preferably, sensors measuring the traveled distance directly, the accuracy of the navigation solution approaches a level that is suitable for many applications. Recently introduced temperature-conditioned MEMS gyro [7] avoid the most significant external error source, i.e., the effect of the ambient temperature in the gyro bias [8]. Studying the noise processes of temperature-controlled MEMS sensors is of great interest because the external factors are negligible and the remaining processes can be considered purely stochastic.

The quality of a MEMS gyro is often defined by the magnitude of the constant additive bias. In positioning and navigation applications, the angular rate measurement output by the gyro is integrated to obtain an angle measurement; any constant bias error in the angular rate is then integrated into a linearly increasing angle error. The bias can be eliminated by means of carousseling [9] or direct estimation. In strapdown applications, carousseling is not an option because of the high power consumption, large size, and extra weight of the gimbal assembly required for slewing the sensor.

If the gyro error consisted of a constant bias and additive white noise only, the bias could be calibrated out by recording a long sequence of data with known input. However, this does not work in practice because the bias of a MEMS gyro has a significant day-to-day component that changes every time the device is powered up. Another reason is that the MEMS gyro error processes are clearly nonstationary and, therefore, cannot be separated into a constant part and a white noise part. Thus, the calibration should be done whenever possible, i.e., whenever the input rotation rate is known. For land vehicle navigation, a practical scenario is to calibrate the gyro bias whenever the vehicle is at standstill; stationarity can be detected based on, e.g., an odometer. In this paper we will focus on this scenario, the bias calibration problem then being a prediction problem given the gyro data of the standstill period.

Bias instability can be defined as “the random variation in bias as computed over specified finite sample time and averaging time intervals. This nonstationary (evolutionary) process is characterized by a $1/f$ power spectral density” [10]. A large amount of literature is devoted to understanding $1/f$ (flicker) noise and other fractional noise processes [11]–[14]. As shown by Voss [15], the Allan variance of $1/f$ noise is constant; without temperature control, external temperature effects mask this behavior, as can be seen in Fig. 1. The Allan variance of the MEMS gyro without temperature control shows an increasing trend at long averaging times. Because of the low availability of temperature-conditioned sensors, MEMS navigation research has been concentrating on other types of errors.

The contribution of this article is to link the properties of $1/f$ noise and the prediction problem involved in gyro bias calibration. We first derive the optimal Kalman predictor for the $1/f$ noise model. Then, noting the computational complexity of the Kalman predictor, we derive a simpler predictor based on moving averages and compare the performance of these prediction algorithms. According to the experimental results, taking the $1/f$ characteristic of bias instability into account enables more accurate predictions of the gyro bias, thus improving the accuracy of the navigation system for which the gyro is being used.

II. RELATED WORK

Calibration methods for gyroscopes have been extensively researched in the literature. For example, when integrating a traditional inertial navigation mechanization with satellite positioning, the inertial sensor errors can be estimated on
the fly using a Kalman filter [16]. In the context of MEMS gyros, day-to-day errors are often so significant that an initial calibration is necessary. In principle, estimating the gyro bias is straightforward given the true angular rate. A common approach is to average a sequence of data measured while the sensor is standing still [17]–[19]; this way, the input is constant (equal to the component of Earth rotation parallel to the sensing axis) and the obtained average is used as a prediction of the future bias. While modeling the bias this way, i.e., as piecewise constant, can be optimal for stationary errors, the nonstationarity of the 1/f component of the gyro signal causes problems with this approach.

Another approach to model the bias instability process is to use an autoregressive (AR) model. For instance, [20] investigates fitting AR models of order up to four on data where wavelet denoising has been applied. A least-squares AR model fitting procedure is described in [21] where using a fourth-order AR model is proposed. In [22], an AR model of order 120 was constructed using system identification methods.

In the articles described above, however, the AR model was not tuned for 1/f noise only, but also for the effect of temperature changes and other factors that tend to dominate the 1/f phenomenon at long averaging times. In contrast, in this paper, the model is built based on the properties of 1/f noise instead of fitting coefficients on measured data, and we use a temperature-conditioned gyro to minimize the effect of long-term errors that would otherwise mask the 1/f behavior.

III. GYROSCOPE NOISE PROCESSES

The output of a gyro is an angular rate measurement which can be modeled at time $t$ as [23]

$$y(t) = M\omega(t) + \epsilon(t)$$ (1)

where $M$ is the $1 \times 3$ cross-coupling and scale factor error matrix, $\omega(t)$ is the angular rate vector between the sensor body and an inertial reference frame, and $\epsilon(t)$ is the additive measurement error. In this article, we neglect the cross-coupling and scale factor effects and divide the additive error $\epsilon(t)$ into three components: constant bias, uncorrelated (white) noise, and 1/f noise. The influence of the most significant error sources can be characterized by computing the Allan variance of the sensor signal, which is discussed below. It should be noted that Allan variance is not the only available method for identifying gyro error structures [24]; nevertheless, it is recommended in [25].

A. Allan Variance

Also known as the two-sample variance, the Allan variance was originally developed for describing the stability of frequency standards [26]. The Allan variance $\sigma^2_\alpha$ is a function of the averaging time $\tau$, computed as

$$\sigma^2_\alpha(\tau) = \frac{1}{2(N - 1)} \sum_{i=1}^{N-1} (\bar{y}(\tau),_{i+1} - \bar{y}(\tau),_i)^2$$ (2)

where the values of $\bar{y}(\tau),_i$ are obtained by dividing the data $y$ into $N$ disjoint bins of length $\tau$; $\bar{y}(\tau),_i$ is the average value of the $i$th bin. Allan variance is usually visualized as a log-log graph. Given the Allan variance plot of a sensor, it is easy to find the optimal averaging time that minimizes the effect of uncorrelated and time-correlated errors: at short averaging times, uncorrelated noise dominates the output, whereas long averaging times are prone to drifting errors. Fig. 1 shows two sets of data and the corresponding Allan variance curves; it can be seen that for both data sets, an averaging time of 1000 seconds would be a reasonable choice.

B. Gyroscope Noise Components

The accuracy of MEMS gyroscopes is degraded by many error sources and the most significant is bias. A comprehensive description of the different error sources is given in [25, Annex C] in the context of laser gyros but the same error sources apply to MEMS gyros as well. Long-term bias can

![Fig. 1. Effect of temperature control on Allan variance. (a): Angular rate as measured by two gyroscopes; (b): the corresponding Allan variances.](image-url)
be calibrated fairly accurately but there is always a day-to-day bias component. When the actual input, i.e., the angular rate, is zero, the bias can be estimated by averaging the sensor output. However, averaging cannot work perfectly because of certain fluctuating (not necessarily zero-mean) errors in the signal.

White noise (also referred to as angle random walk [25]), originating from, e.g., thermal noise [27], is a simple stochastic process that consists of independent and identically distributed samples. Its variance is inversely proportional to the averaging time, thus the effect of white noise can be seen as an initial negative slope in the Allan variance graph.

1/f noise, however, is a more complicated stochastic process. Although discovered almost a century ago [11], [28] and encountered in many different contexts, such as semiconductors, river Nile’s flood levels, and pitch fluctuations in music, its origin is not exactly known [15]. As opposed to uncorrelated white noise, 1/f noise is a long-memory process, i.e., the mutual correlations between samples decay slowly [29]. It can be modeled as an ARFIMA(0, d, 0) process (autoregressive fractionally integrated moving average) whose dth difference is a white noise process. The dth difference of the sequence \( x_0, x_1, \ldots \) is computed as [14], [29], [30]

\[
\Delta^d x_k = \sum_{i=0}^{k} (-1)^{k-i} \frac{\Gamma(k-i+d)}{\Gamma(k-i+1)\Gamma(d)} x_i
\]

(3)

where \( \Gamma(\cdot) \) denotes the gamma function. Another quantity to characterize a long-memory process is the Hurst exponent \( H \), which is related to the order \( d \) by [14]

\[
d = H - \frac{1}{2}.
\]

(4)

The Hurst exponent is commonly encountered in literature where long-memory processes are discussed but, in this article, \( d \) will be used as the memory parameter.

Another error source in the sensor output is rate random walk (RRW), which is caused when the bias changes (slowly) with time. This can be due to, e.g., aging of the sensor components or changes in the temperature of the gyro; MEMS sensors are typically highly sensitive to temperature variations [8]. RRW causes a positive slope to the Allan variance at long averaging times. Unlike 1/f noise, RRW is a Markov process, i.e., memoryless: its value at step \( k + 1 \) only depends on the value at step \( k \), not on any other (past or future) values.

Since white noise causes imprecision at short averaging times and RRW degrades the accuracy at long averaging times, the optimal averaging time is found between these two. The minimum of the Allan variance of a gyro is called the bias instability and can be used as a measure of the power of 1/f noise [25].

IV. Proposed Gyroscope Error Predictors

In this section, we present two approaches of predicting the evolution of the bias: one based on Kalman filtering and one based on computing a moving average. In Section V, these predictors will be compared to the simple approach of averaging the entire calibration data.

A. Kalman Filter Approach

The Kalman filter (KF) [31] is a tool for analyzing dynamic systems and is widely used in positioning and navigation, along with its nonlinear extensions. Given the observation and system evolution models and an initial state, it estimates the state of the system. In this study, the KF is employed for tracking 1/f noise.

As state variables we take the current and previous noise realizations. This means that the length of the state vector increases at each time step; the solution cannot be optimal unless the state vector is infinite-dimensional [32]. It has been shown that one state variable per decade of samples (i.e., one for the most recent sample, one for the latest 10 samples, one for the last 100 etc.) are sufficient to characterize a 1/f process [13], [33]. Nevertheless, in this study, one state variable was used for each sample. Therefore, the state is propagated from sample \( (j-1) \) to \( j \) as a fractional integral

\[
x_j \approx \left[ I \right]^{1/(j-1)} x_{j-1}
\]

(5)

where \( I \) is the identity matrix, \( \theta \) denotes the transpose, and the \( i \)th element of the vector \( f_j \in \mathbb{R}^{i} \) is obtained as [14], [29]

\[
f_{ji} = \frac{\Gamma(j-i-d)}{\Gamma(j-i+1)\Gamma(-d)}
\]

(6)

The formula is remarkably similar to (3)—only the sign of \( d \) is different. This formula is, in principle, only valid for \( -1/2 < d < 1/2 \), i.e., when the process is stationary and invertible; for \( d > 1/2 \), these conditions do not hold [14]. Therefore, the prediction obtained using the values (6) may be biased if \( d = 1/2 \) is used. For covariance propagation, the variance of the driving noise of the 1/f noise needs to be known.

The coefficients obtained from (6) have absolute value less than one. Thus, as time passes, the forecast eventually decays towards zero. Nevertheless, it would be desirable that the prediction would tend towards the sample mean instead. This can be changed by taking the sample mean as an additional state variable and modeling the bias as the sum of 1/f noise and the sample mean. This makes observation updates straightforward as well, as long as the variance of white noise is known. If it is not given in the specifications of the gyro, it can be estimated from the Allan variance.

B. Moving Average Approach

As the optimal prediction using Kalman filter is too complex for practical implementation, we introduce simpler predictor based on moving averages. Assuming that the Allan variance of the noise process is constant, i.e.,

\[
\frac{1}{2(\bar{N}-1)} \sum_{i=1}^{\bar{N}-1} (\bar{y}(\tau)_{i+1} - \bar{y}(\tau)_{i})^2 = \beta \quad \forall \tau,
\]

(7)

and further assuming that the increments in the block averages \( \bar{y}_{i} \) are mutually independent, the best predictor for \( \bar{y}(\tau)_{i+1} \) given \( \bar{y}(\tau)_{i} \), would be \( \bar{y}(\tau)_{i} \) itself. Thus the
prediction of \( \tau \) future samples would be the mean of past \( \tau \) samples. Generalizing this leads to the simple predictor
\[
\hat{y}_j = \frac{1}{k} \sum_{i=j-k+1}^{0} y_i
\]
where \( j = 0 \) is the time index of the last calibration sample. This is by no means an optimal predictor for \( 1/f \) noise, as the second assumption (independence of the increments) does not hold. However, the predictor is very easy to implement and the results presented in Section V show that its prediction performance is not far from the Kalman predictor values. If the calibration period is shorter than the navigation mission (as often is the case), the simple average of the calibration data
\[
\hat{y}_j = \frac{1}{k} \sum_{i=j-k+1}^{0} y_i, \quad j \geq k
\]
where \( k \) is the length of calibration data, can be used to extend the moving average.

V. Experiments

The performance of the proposed predictors was evaluated by running them on a large number of data sets with zero angular rate input; there were both artificial data sets and authentic gyroscope data. In each run, the first half of the data were used for calibration, based on which the evolution of the bias was predicted in the latter half. This resembles the scenario of driving a vehicle: occasionally, the vehicle must stop, e.g., at crossroads, which gives an opportunity for a zero-velocity update. Another application is pedestrian walking [34] where there is a stance phase followed by a swing phase, although the durations of the phases considerably shorter.

The artificial data were generated as ARFIMA(0, 1/2, 0) sequences according to (3) and the authentic data originated from an immobile gyro; the data sets used in Section V-B contained 2,400 samples each, whereas the results of Section V-C are based on 2,000-sample-long sets. The angular rate caused by the rotation of Earth was not treated in any way—since the gyro was static during the experiments, the effect of Earth rotation was constant and can be seen as a part of the bias. In all test runs, an integration order \( d = 1/2 \) was used for modeling \( 1/f \) noise; this is not necessarily optimal for the real data as some other choice of \( d \) could model the gyro noise correlations better. However, the exact local Whittle estimate [12] of the gyro data used is 0.535, which suggests that the chosen value of \( d \) is reasonable.

In theory, because of the zero angular rate, the integral (in discrete time, the sum) of the difference between the data and the predicted bias should equal zero. Therefore, we use two error metrics to quantify the prediction error:

- Integrated error (IE): the sum of the prediction errors, i.e., the difference between the data and the predicted bias
- Root sum of squared errors (RSSE): the sum of the squares of the prediction errors.

Obviously, the IE describes the accumulated angle error after the prediction phase, and the absolute value of the IE is what we ultimately want to minimize. The RSSE is computed for validation purposes: the KF should, in theory, minimize the error variance among linear estimators. Therefore, if the KF does not yield the lowest RSSE, the models used in the KF are probably incorrect.

In the following sections, we first discuss the choice of the gyroscope. Then, we study the behavior of the proposed predictors in example runs based on both artificial and authentic data. Finally, the performance of the predictors is compared in an extensive batch test in order to assess how the methods perform in general.

A. Choosing the Sensor

In order to get gyroscope data whose power spectral density obeys a \( 1/f \) shape as well as possible (when properly averaged), we need a gyroscope with negligible RRW. As shown in Fig. 1, this can be achieved even using a MEMS sensor if it is temperature-conditioned: in the Allan variance curve of the temperature-conditioned gyro, no RRW slope is visible at averaging times smaller than \( 10^5 \) seconds and the Allan variance is almost constant for averaging times longer than \( 10^2 \) seconds. The same figure shows that this is certainly not the case for the gyroscope that is not temperature-conditioned: its Allan variance starts to increase at averaging times longer than approx. \( 10^4 \) seconds because of RRW. The temperature-conditioned gyro appearing in Fig. 1 is [7] and the other gyro is [35].

In applications, temperature-conditioning has its drawbacks; for instance, the power consumption and physical size of the oven where the sensor is located may prevent using the technology in mobile devices. However, these considerations are beyond the scope of this paper.

B. Example Runs

Fig. 2 shows the results of applying the estimation methods discussed in Section IV on artificial data. The samples were generated by half-integrating Gaussian white noise according to (3) and adding a constant bias; no other components, such as additive white noise or random walk, were introduced. It can be seen that the KF estimate (Fig. 2a) decays smoothly towards the mean of the calibration data while the moving average (Fig. 2b) is more oscillatory, especially at the beginning where the average is computed over a smaller number of samples.

With this data set, the KF performs better in terms of both IE and RSSE: the KF error integrates to \( -201 \) (dimensionless; no units assumed for the simulated data) while the integrated moving average error is 535. The RSSE values are 123 and 127, respectively. For comparison, the errors obtained by using the mean of the calibration data as the sole estimate of the bias were \( -2910 \) (integrated) and 146 (RSSE). Hence, the improvement is obvious in this case. Fig. 2a also shows the two-sigma confidence interval of the KF estimate; the bounds seem to agree with the data.

An example of estimator performance using real gyro data is shown in Fig. 3. To be able to use the KF, the standard deviations of additive white noise and the driving noise of the \( 1/f \) process had to be estimated first. The values 45°/h and 14°/h, respectively, were found to be appropriate, which
is reflected by the consistency of the two-sigma confidence interval estimate in Fig. 3a. It should be noted that these are the confidence bounds of the signal including white noise; the confidence interval for the $1/f$ component only would be narrower.

The KF outperforms the moving average and calibration data mean with this set of data, too. The IE and RSSE for the KF estimate are $2.46^\circ$ and $0.410^\circ$, respectively. The corresponding values for the moving average estimate are $3.13^\circ$ and $0.424^\circ$, and for the mean of the calibration data, $3.55^\circ$ and $0.415^\circ$. In this test case, the mean of the calibration data was a better estimator in the RSSE sense than the moving average.

C. Batch Test

In the example cases described above, it was only known that the prediction model was correct in the case of Kalman filter with simulated data. To validate the modeling, similar tests with different sets of data—1,000 simulated and 1,000 authentic sets—were conducted in a batch and the performance of the three approaches were logged. A new set of data was half-integrated for every simulation run whereas the authentic data sets were obtained by taking consecutive disjoint sections of 2,000 samples from a long set of data. In each run, the first 1,000 samples were used for calibration and the rest 1,000 samples for prediction.

The results of the batch runs are shown in Table I. It can be seen that both of the proposed methods yield a lower
average IE than the traditional CM approach with authentic data whereas CM shows best performance in simulations, but the error standard deviations are fairly large. Dividing the error standard deviations by $\sqrt{1000}$ gives the standard deviations of the sample means; the differences in IE are smaller than these standard deviations. Therefore, these differences cannot be regarded as significant.

In contrast, the differences in RSSE between KF and the other approaches are more considerable. Between the MA and CM methods there is a one-sigma difference in favor of CM in the simulations and vice versa with authentic data, but the KF mean RSSE is smaller by two sigmas in the simulated runs and by four sigmas with real data than the CM mean RSSE. This is in accordance with the optimality of KF among linear estimators and, on the other hand, suggests that the models used in the KF are appropriate.

### VI. Conclusions

In this article, two methods for predicting the evolution of the additive bias of a MEMS gyro were presented and compared to the traditional approach of plain averaging. The results suggested that one of the proposed methods, i.e., the KF approach, yields more accurate estimates in the RSSE sense, but the integral errors, which are of more importance in gyro applications, were not as significantly smaller. Another advantage of the KF is that it directly allows for computation of confidence intervals; the results showed that the obtained bounds are consistent with the data. Knowing the variance of the bias estimate is important when the measurements are used for position computations because the uncertainty in the gyro bias estimate affects the uncertainty of the resulting position solution.

The optimality property of the Kalman filter is only valid given that the models used are correct. Therefore, it is important to ensure that the environmental conditions are not changing during the calibration and prediction phases; for instance, temperature fluctuations or sudden vibrations may affect the gyro bias. In this study, these factors were mitigated by using a temperature-conditioned sensor with vibration isolation.

As future work, the fractional integration model, possibly with different choices of $d$, could be compared with AR models obtained using system identification algorithms, such as in [22], in terms of both model structure and prediction performance. A performance comparison could be made for the moving average estimator as well in order to see if the extra complexity of the AR models pays off. Moreover, confidence interval estimators should be derived for these suboptimal predictors. Given the estimate variance of the AR model it would be easier to decide if for a particular application it is sufficient to use, e.g., the simpler MA method instead of the computationally demanding KF approach, despite the degraded prediction performance.

### REFERENCES


Martti Kirkko-Jaakkola received his M.Sc. (Hons.) degree from Tampere University of Technology, Finland, in 2008, majoring in mathematics. Currently, he is a graduate student and works as a Researcher at the Department of Computer Systems, Tampere University of Technology, where his research interests include precise satellite positioning, low-cost MEMS sensors, and indoor positioning.

Jussi Collin received his M.Sc. and Dr.Tech. degrees from Tampere University of Technology, Finland, in 2001 and 2006, respectively, specializing in sensor-aided personal navigation. He is currently a Research Fellow at the Department of Computer Systems, Tampere University of Technology. His research is focused on statistical signal processing of MEMS motion sensors and novel MEMS-based navigation applications. He is named inventor on several patents in the field of navigation.

Jarmo Takala received his M.Sc. (Hons.) degree in electrical engineering and Dr.Tech. degree in information technology from Tampere University of Technology, Tampere, Finland (TUT) in 1987 and 1999, respectively. From 1992 to 1995, he worked as a Research Scientist at VTT-Automation, Tampere, Finland. Between 1995 and 1996, he was a Senior Research Engineer at Nokia Research Center, Tampere, Finland. From 1996 to 1999, he was a Researcher at TUT. Currently, he is Professor in Computer Engineering at TUT and Head of the Department of Computer Systems of TUT. His research interests include circuit techniques, parallel architectures, and design methodologies for digital signal processing systems. Prof. Takala was Associate Editor for IEEE Transactions on Signal Processing during 2007–2011, and in 2012–2013 he is Chair of IEEE Signal Processing Society’s Design and Implementation of Signal Processing Systems Technical Committee.
© 2013 IEEE. Reprinted, with permission, from


In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Tampere University of Technology’s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.
Using Building Plans and Self-Contained Sensors with GNSS Initialization for Indoor Navigation

Martti Kirkko-Jaakkola, Jussi Collin, and Jarmo Takala
Department of Pervasive Computing
Tampere University of Technology
P.O. Box 553, FI-33101 Tampere, Finland
Email: {martti.kirkko-jaakkola, jussi.collin, jarmo.takala}@tut.fi

Abstract—This article presents a method for indoor pedestrian navigation based on low-cost inertial sensors and building plan information with Global Navigation Satellite System (GNSS) based initialization. The inertial sensors are utilized in a step-detecting mechanization and the position is estimated with a particle filter using the building as a motion constraint. In the article it is demonstrated that the method is capable of tracking a pedestrian when initialized indoors using a high-sensitivity GNSS receiver. Field test results show that building plan constraints are efficient in preventing the inherent error accumulation of dead-reckoning type navigation systems, thus enabling autonomous dead reckoning for prolonged periods of time. The proposed method does not rely on local infrastructure or signals of opportunity such as wireless LAN observations, and is thus suitable for, e.g., first responders.

I. INTRODUCTION

Using a Global Navigation Satellite System (GNSS) receiver, resolving one’s location is a quick and easy task in most outdoor locations. Nowadays GNSS receivers are integrated in many devices such as cellular phones, which has paved the road for diverse consumer products and applications. Unfortunately, the sensitivity of most GNSS receivers is insufficient to tolerate the severe attenuation of the signals when located indoors. On the other hand, even high-sensitivity receivers specifically designed for acquiring and tracking weak signals cannot typically resolve the user position accurately enough to, e.g., identify the room where the user is located [1]; the obtained accuracy depends on the surroundings of the antenna and on the available assistance information.

Because of these limitations, most current indoor positioning systems rely on other methods than GNSS observations. The most common approach is to scan for wireless local area network (WLAN) signals: since wireless access points (APs) usually stay in the same location for a long time, having a database of the locations and medium access control (MAC) addresses of APs allows estimation of the location of the user; a number of such database services are available. Other popular signals of opportunity include cellular network and Bluetooth [2] signals, but they have certain limitations. Despite practically global coverage, cellular triangulation usually results in a relatively imprecise position solution. Bluetooth signals work at short ranges only, and using them for positioning requires the presence of dedicated transmitters.

There are, however, use cases where a positioning system cannot rely on the availability of local infrastructure. A prominent example is the localization of first responders: if the building is on fire or another factor causes a power outage, it cannot be guaranteed that wireless APs and other electrical devices in the building are operating properly. In these cases, the navigation system must be self-contained or at least only use sources of information that do not depend on signals originating from the building.

Inertial sensors can be used for keeping track of the motion of the user by means of dead reckoning, and microelectromechanical (MEMS) technology is an attractive solution concerning the size, power consumption, and price of the hardware. However, despite immense development in the past years, MEMS inertial sensors are not accurate enough for autonomous dead reckoning for long periods of time. The foremost problem is the bias instability of MEMS gyroscopes which causes the heading estimate to drift away from the true value, resulting in an accumulating position error. In principle, a MEMS magnetometer (compass) could be used for measuring the heading, but the abundance of ferromagnetic objects in typical indoor environments makes these measurements very unreliable in practice.

In this article, we use an indoor map of the building to filter the heading and correct for gyroscope drift with a body-mounted MEMS inertial measurement unit (IMU); for combining these sources of information we use the Monte Carlo approximation of the Bayesian filter called the particle filter (PF). However, dead reckoning always requires initial conditions to be determined by other means, which is usually done using radio frequency (RF) signals. Wang et al. [3] coupled accelerometers and WLAN measurements with indoor maps for pedestrian navigation. It is noteworthy that the setup contained no heading sensors (gyros or magnetometers), giving the system very limited capabilities of operating accurately in the absence of WLAN information.

Woodman and Harle [4] described a general IMU–map-matching framework and demonstrated it using initial conditions obtained from WLAN observations. In a subsequent paper [5], results with continuous WLAN updates were added. Beauregard et al. [6] used a similar setup to investigate the effect of map accuracy to positioning performance whereas Krach and Roberston [7] proposed a cascaded dual filter architecture to the hybridization. In these articles, the IMU
was mounted on the foot of the walker, which enables more accurate step length estimation but is not as practical as mounting the sensors on the body by, e.g., keeping the unit in a pocket. Furthermore, these studies did not focus on the initialization of the particle filter. Using a massive amount of particles to occupy the entire map is only feasible in relatively small buildings, and WLAN is not suitable for all use cases as described above.

Kemppi et al. [2] propose a hybrid indoor positioning system relying on dedicated Bluetooth transmitters. Bluetooth radios are found in various devices such as cellular phones, and in the positioning context Bluetooth has the advantage over WLAN that its signal power is lower, decreasing the probability of outlier observations which tend to occur with WLAN signals in large open halls. However, the obvious downside is the fact that special infrastructure must be deployed in the building. Other dedicated infrastructure, such as RFID tags, have also been proposed [8], but they require special components that are uncommon in today’s mobile devices.

In this work, we use a GNSS receiver to estimate the initial position of the user but assume the initial heading to be unknown. We initialize the PF indoors using a high-sensitivity GNSS receiver with long coherent integration [9]. The advantage of GNSS compared to the related work discussed above is its independence of building-dependent infrastructure whose availability cannot be guaranteed. Test results show that the proposed system is capable of tracking the location of the user indoors without any signals from local infrastructure.

II. THEORETICAL BACKGROUND

In this section, we review the building blocks of the proposed positioning algorithm. First, the basics of high-sensitivity GNSS processing and inertial navigation mechanization for pedestrian use are addressed; then, particle filters and how to apply them for map-matching are discussed.

A. High-Sensitivity GNSS

High-sensitivity (HS) GNSS receivers are capable of acquiring and tracking signals that are significantly weaker than nominally. Most commercially available HS-GNSS receivers can typically cope with signals attenuated by up to 30 dB with respect to the nominal signal power, corresponding to a carrier-to-noise density ratio (C/N0) of approximately 15 dB-Hz. However, in the scope of this article, we will consider a receiver that can acquire and track signals with C/N0 well below 10 dB-Hz which is a typical level indoors, corresponding to a signal that has propagated through a brick roofing or a concrete wall [10].

The key to improved receiver sensitivity is a long coherent integration of the signals. However, the choice of a coherent integration time is subject to certain constraints, the foremost being the navigation message. When using legacy GPS C/A signals without assistance information, the a priori unknown navigation data bits cause coherent integration longer than 20 milliseconds to be prone to correlation loss due to possible signal sign reversals. This problem can be avoided if a pilot channel is available or if the data bit values are known a priori by other means, e.g., retrieved online. If such assistance is not available, the data bits can be estimated on the fly [11].

Another limiting factor for coherent integration time is the instability of the reference oscillator. Although chip-scale atomic frequency standards are nowadays available, they are too expensive to be used in mass-marketed products. Thus, the usual choice of a highly stable oscillator is an oven-controlled crystal oscillator (OCXO) where the inherent temperature sensitivity of crystal oscillators is circumvented by keeping the unit in a constant temperature. This yields a stability superior to temperature-compensated crystal oscillators.

Finally, the motion of the user causes a Doppler shift which is not necessarily constant during a coherent integration interval of, e.g., 1 s. This problem can be overcome by measuring the motion using an IMU and compensating for it in the correlation process, thus decreasing the needed Doppler tracking bandwidth; this can be realized in a tightly or deeply (a.k.a. ultra-tightly) coupled architecture. An additional benefit of narrow-band tracking loops is improved robustness against noise and RF interference.

B. Pedestrian Dead Reckoning

It is well known that the inertial navigation mechanization with six degrees of freedom is inherently unstable. Especially with low-cost IMUs, the double-integrated sensor measurement errors cause rapidly accumulating position and attitude errors. Fortunately, the double integration is not always necessary for dead reckoning. When the user is moving on foot, every step taken introduces a distinctive pattern to the accelerometer output. Therefore, if it is assumed that the user is moving on a flat surface such as on the floor indoors, the positioning problem reduces to two dimensions when the motion is estimated by detecting steps from the acceleration signals. This mechanization is known as pedestrian dead reckoning [12].

Augmenting the above described model with an estimate of the additive gyroscope bias leads to a system with four unknowns, i.e.,

$$\begin{bmatrix} x \\ y \\ \psi \\ \delta \omega \end{bmatrix} = \begin{bmatrix} E \\ N \\ \delta \omega \end{bmatrix}$$

where \( E \) and \( N \) are the East and North position components, respectively, \( \psi \) is the heading, and \( \delta \omega \) is the gyro bias; if necessary, the height (i.e., floor number when navigating inside a building) can be estimated as an additional unknown. We assume that all the steps are taken in the direction of the heading, and a single gyroscope can be used to measure changes in the heading angle. We model the step length as a (known) constant plus white Gaussian noise with standard deviation of 0.1 m. Various step length estimation algorithms have been developed, e.g., [13], [14]; using a less accurate step length model will obviously increase the along-track position error, but we will rely on map-matching to compensate it.
C. Particle Filtering

Bayesian filtering is a framework for estimating a vector of unknown states by combining various sources of information optimally to obtain a minimum variance estimator. However, the Bayesian filtering equations cannot usually be solved in closed form when the transition and measurement models are nonlinear or the underlying distributions are non-Gaussian. Although the well-known Kalman filter implements the Bayesian filter for linear Gaussian models, its nonlinear augmentations such as the Extended Kalman filter are not generally optimal.

A popular approximation of the Bayesian filter is the particle filter [15] which is a Monte Carlo method that approximates the posterior distribution of the states using a set of samples called particles. The particles are weighted based on their likelihood with respect to measurements. The representation of the posterior distribution as a cloud of particles instead of using mean and covariance matrices does not make assumptions on the type of the posterior distribution, which is a significant advantage for map-matching because map constraints lead to significantly non-Gaussian and often multimodal distributions. This is illustrated in Fig. 1 where only an irregular subset of the cloud of particles conforms to the map constraints.

The particle filter operates with a prediction–update principle similar to Kalman-type filters. In the prediction step of time instant $t$, the particles $x_i^{(t)}$, $i = 1, \ldots, N$, are drawn from a proposal distribution $\pi$, i.e.,

$$x_i^{(t)} \sim \pi \left( x_i^{(t)} | x_{i-1}^{(t-1)}, y_1 \ldots y_{t-1} \right)$$  \hspace{1cm} (2)

where $y_1 \ldots y_{t-1}$ comprises all measurements made before time $t$. In this study, the proposal distribution $\pi$ is chosen to be the transition model $p(x_i | x_{i-1})$; we assume that the states constitute a Markov process, i.e., the state $x_i$ depends on the previous state $x_{i-1}$ but not on any other past or future values of the state. During the update step, the weights $w_i^{(t)}$ are updated according to the observation likelihood as

$$w_i^{(t)} = w_i^{(t-1)} \frac{p \left( y_t | x_i^{(t)} \right) p \left( x_i^{(t)} | x_{i-1}^{(t-1)} \right)}{\pi \left( x_i^{(t)} | x_{i-1}^{(t-1)}, y_1 \ldots y_{t-1} \right)},$$ \hspace{1cm} (3)

After the update, it is a common practice to normalize the weights to sum up to unity; then, the expected value of the approximate posterior distribution is easily computed as the weighted sum of the particles.

When running the particle filter, very few or only one of the particles will eventually gain almost all of the weight. Not only is the propagation of particles with negligible weight a waste of resources, but such a sample is not a good approximation of the posterior distribution either. This problem is known as degeneracy and can be avoided by resampling the cloud of particles. Various resampling algorithms exist, but they are not discussed in detail in this article; for more information, see, e.g., [16].

D. Map-Matching

The purpose of map-matching is to improve the accuracy of the position estimate by incorporating information about where the user is likely or unlikely to be located based on a map. Map-matching is common in vehicular navigation where the vehicles are moving along a known road network; in this case, the map information can be utilized, e.g., by projecting the computed position onto the nearest road link [17] or by incorporating the map information into the position computation phase as an additional constraint [18].

In the context of pedestrian indoor navigation, the pedestrian can usually walk around freely inside a room and there are no predefined paths as opposed to a road network. Nevertheless, the walls of a building do constrain the motion of the user; therefore, an indoor map based on the building plan can be used to impose additional constraints onto the state vector. With particle filters, this is intuitively done as an update step with binary likelihood

$$p \left( y_t | x_i \right) = \begin{cases} \epsilon & \text{if the particle crossed a wall} \\ 1 & \text{otherwise} \end{cases}$$  \hspace{1cm} (4)

with $0 \leq \epsilon \ll 1$. Usually, $\epsilon$ is set to zero, causing wall-crossing particles to be discarded; however, a small positive value can be chosen for $\epsilon$ instead to take into account possible errors in the map. Applying the map update after each detected step corresponds to an update rate in the order of 1–2 Hz. Since an indoor map can easily contain thousands of wall segments, efficient searching of walls that lie close to a certain particle requires storing the map in a well designed data structure.

In some occasions it may happen that all particles get trapped into a dead end; this may be caused by errors in the map or in the measured user movement. In order to cope with such situations, we increase the variance of the particle cloud at every resampling step such that the particles are allowed to move across walls (as long as they stay inside the building). In other words, instead of plain duplication of particles, we also add random noise to them during resampling. If the whole cloud of particles hits a dead end, resampling is likely to soon take place, and the added noise allows some particles to pass over the dead end to an adjacent room.
In this section, the operation of the proposed algorithm is demonstrated. We first describe the data collection equipment and the calibration procedures performed. Then, we show results from a real-world experiment.

A. Test Setup

The test data were measured using the Demonstrator for Indoor GNSS Positioning (DINGPOS) platform [9]. It comprises a software HS-GNSS receiver deeply coupled with an Xsens MTi [19] MEMS IMU; there are also other sensors such as a barometer and an INPOS ZigBee transceiver, and the system is operated using a laptop including a WLAN interface, but these measurements are not considered in the scope of this article. For reference purposes, a U-blox EVK-5H commercial off-the-shelf (COTS) GNSS receiver [20] is sharing the antenna with the HS receiver.

In order to enable long coherent integration times, the HS-GNSS receiver makes use of certain assistance data. To overcome the 20 ms limitation, the navigation data bits are recorded and provided by another receiver with an outdoor antenna. Furthermore, the HS receiver is given a coarse initial position for, e.g., predicting the Doppler shift.

Before the test, the clock of the HS receiver was calibrated. First, its offset was compensated for with a precision of ±30 milliseconds. This was done by letting the receiver have a good position fix outdoors and then compensating for the estimated clock bias, but it would have been also possible to obtain a sufficiently precise calibration from the outdoor-located receiver using the network time protocol. After setting the clock, the frequency of the local oscillator of the HS receiver was calibrated; the receiver includes an OCXO whose controller can be fine-tuned. The frequency offset was calibrated to no larger than ±7 Hz using a good GNSS fix as the reference value. After the calibration, the receiver was restarted before beginning the actual test.

When processing the IMU data in the particle filter, one of the three gyroscopes in the IMU was used as the heading gyro without scale factor or axis misalignment corrections. Because the IMU was mounted at the top of a backpack, it is expected that the IMU swings slightly while walking, thus causing cross-coupling between sensor axes. Although these errors are straightforward to compensate for [21], [22], they were intentionally left uncorrected to demonstrate that the proposed algorithm can tolerate these imperfections. The indoor maps were obtained from Nokia Destination Maps.

B. Test Results

The test route started indoors with a ten-minute stationary period, followed by ten minutes of walking and finally returning to the initial position. The COTS receiver did not output any position fixes during the test. The resulting navigation solution is shown in Fig. 2. The user left the room where the receiver initially was and walked two laps along the same route along the hallways without entering any rooms before returning to the initial location at the end of the test. He intentionally zigzagged in some corridor wings; it can be seen that the fluctuating heading did not cause problems with the map-matching filter, but a ground-truth solution is not available for rigorous accuracy quantification.

The HS-GNSS position estimates plotted in Fig. 2 show that the HS-GNSS solution cannot indicate which room the receiver is located in, but the estimates converge to a point less than 20 meters from the true position. $N = 300$ particles were initialized at random locations around the final HS-GNSS estimate, but those located outside the building were given zero weight and the rest were weighted based on the distance to the HS-GNSS estimate. The covariance of the initial particles was chosen based on the precision of the position estimate.

The solution trajectory is imprecise in the beginning which causes random-like patterns in the leftmost wing. This is due to the fact that the particle distribution is multimodal at that time, but the mean values are shown in the figure; clearly, the mean of a multimodal distribution does not necessarily coincide with any of the modes. However, it can be seen that the ambiguity is resolved after turning to the main hallway. The same phenomenon arises later when turning from the wings to the main hallway: when the turn starts, some of the particles actually move to the rooms in the wing, resulting again in a multimodal distribution. These false modes die out gradually, making the average line look like it would be cutting the corner. If necessary, multimodality can be detected using, e.g., clustering methods on the set of particles.

The altitude component of the HS-GNSS position solution was not accurate enough to determine the floor number. Therefore, it had to be assumed that the floor is known by other means, e.g., using a barometer if a reference pressure is known. If there are significantly dissimilar room layouts in different floors of the building, it would be possible to initialize the filter by placing particles in all floors, but this would require a larger number of particles for the filter, thus increasing the demand on computational resources.

IV. Conclusions

In this article, an algorithm for map-assisted pedestrian navigation indoors was described. It was demonstrated using real-world data that a position obtained by means of high-sensitivity GNSS processing was sufficient to initialize the filter; no initial heading information was required. Thus, the algorithm does not need any local infrastructure whose availability cannot be guaranteed, such as WLAN APs. Dependence on electrical building infrastructure is a serious drawback when considering, e.g., positioning of first responders.

After the initial convergence period, the filter was able to follow the pedestrian’s route with a reasonable accuracy. The four-state particle filter was run using 300 particles which cannot guarantee convergence in every Monte Carlo run. More particles could be used to improve the convergence, but a significantly smaller amount will suffice after the false modes have died out; anyway, the running time is obviously proportional to the number of particles. Sometimes there were situations where parallel rooms or hallways caused temporarily
multimodal distributions, implying problems when the final position estimate is computed as the weighted average of the particle cloud. A more sophisticated estimator, e.g., based on clustering the set of particles, could improve the positioning accuracy in ambiguous cases.

In this implementation, the GNSS measurements were only used for determining the initial location. As future work, the map-matching filter is to be coupled with HS-GNSS signal processing in order to get GNSS updates during the navigation. Moreover, for initializing the system inside a multi-story building, a means of identifying the initial floor is needed.

ACKNOWLEDGMENTS
This research has been partly funded by the ESA under the NPI programme, and ESA is also acknowledged for lending the DINGPOS platform to the authors. Furthermore, the authors would like to thank Dr. Gustavo López-Risueño for his feedback on the manuscript.

REFERENCES


A RAIM APPROACH TO GNSS OUTLIER AND CYCLE SLIP DETECTION USING L1 CARRIER PHASE TIME-DIFFERENCES

M. Kirkko-Jaakkola\textsuperscript{a}, J. Traugott\textsuperscript{b}, D. Odijk\textsuperscript{c}, J. Collin\textsuperscript{a}, G. Sachs\textsuperscript{b}, and F. Holzapfel\textsuperscript{b}

\textsuperscript{a} Department of Computer Systems, Tampere University of Technology, Finland
\textsuperscript{b} Institute of Flight System Dynamics, Technische Universität München, Germany
\textsuperscript{c} Department of Spatial Sciences, Curtin University of Technology, Australia

ABSTRACT

Cycle slips are a common error source in Global Navigation Satellite System (GNSS) carrier phase measurements. In this paper, the cycle slip problem is approached using Receiver Autonomous Integrity Monitoring (RAIM) methodology. Carrier phase measurements are used here in a single-receiver time-differential positioning method where integer ambiguities are canceled, but any cycle slips remain. The performance of the method was assessed by comparing the detection results to a Real-Time Kinematic (RTK) solution and by manual data examination. Postprocessing results obtained using authentic Global Positioning System (GPS) measurements logged by low-cost single-frequency receivers show that the method is able to reliably detect and identify single errors but fails in an exemplary multiple outlier scenario. As no reference receiver is needed, the method is a potential means to produce cycle-slip-corrected data usable in any postprocessing application.

Index Terms— GNSS, carrier phase, cycle slip, RAIM, time-differential positioning

1. INTRODUCTION

The Global Navigation Satellite System (GNSS) carrier phase observable is a key to high positioning precision. Using it, however, is not straightforward as phase measurements are biased by ambiguities.\textsuperscript{1} During continuous phase lock, the ambiguity remains constant, allowing for cancelation by subtracting two consecutive measurements. This is the key idea in time-differential (TD) positioning \cite{1}. However, a temporary loss of phase lock may cause the ambiguity to change, resulting in a cycle slip. Time-differencing does not cancel cycle slips, so they must be detected as each slipped cycle corresponds, at the GPS L1 frequency, to 19 cm of range error which is much if decimeter-level or better precision is desired.

Cycle slip detection is typically based on geometry-free observables or measurement prediction. Canceling the range component requires either dual-frequency measurements or knowledge on the receiver dynamics from, e.g., inertial sensors \cite{2}. However, these features increase hardware costs. Measurement prediction can be done by simple polynomial fitting or using, e.g., a Kalman filter. The wavelet transform has also been applied to cycle slip detection \cite{3}. In a dynamic application the measurements may, however, be hard to predict or model with wavelets.

Receiver Autonomous Integrity Monitoring (RAIM) has been extensively researched and successfully used in traditional pseudorange-based GNSS positioning, but such a method has also been applied on carrier observables \cite{4}. RAIM is based on measurement redundancy which poses a requirement of an overdetermined system of equations. The least-squares position solution residual can be used as a measure of consistency: if the norm of the residual exceeds some predefined threshold value, an alarm is raised. Further investigations can be made to pinpoint the faulty measurement.

The goal of this work is to use and to assess the performance of a RAIM method on cycle slip detection using standalone time-differenced data. The method could be used as a preprocessing step for any carrier phase processing application as no assumptions are made on, e.g., receiver dynamics.

2. CARRIER PHASE OBSERVABLE AND CYCLE SLIPPING

The GNSS carrier phase measurement for satellite $S$ at epoch $t_i$ can be modeled as

$$\Phi_i^S = \rho_i^S + \delta \rho_i^S + \delta \delta_i^S + T_i^S - I_i^S + \frac{\lambda}{2} \bar{N}^S + \epsilon_i^S \quad (1)$$

where $\rho_i^S$ is the distance between the receiver and satellite $S$, $\delta \rho_i^S$ and $\delta \delta_i^S$ denote the receiver and satellite clock biases, respectively, scaled to units of meters by the speed of light $c$. $T_i^S - I_i^S$ corresponds to atmospheric effects (tropospheric delay and ionospheric advance). $\bar{N}^S$ is the ambiguity which is constant during continuous phase lock. $\lambda$ denotes the carrier wavelength, and $\epsilon_i^S$ comprises modeling errors such as multipath, ephemeris errors, and measurement noise.

\textsuperscript{1}The ambiguity is usually referred to as integer ambiguity as fractional components caused by satellite and receiver phase biases cancel (only) when forming double differences as required for RTK.
Even though the ambiguity is usually referred to in units of full carrier cycles, there may be an additional half-cycle ambiguity due to the navigation message modulated on the signal. As the message is not known beforehand, it cannot be removed to obtain a clean sinusoidal carrier wave to track. For this reason, carrier tracking loops are sometimes constructed as Costas loops which are insensitive to 180° phase shifts caused by, e.g., navigation data bit changes. Such a loop does not know if it is tracking the carrier correctly or off by 180 degrees [5]. The receiver can, however, resolve this ambiguity by examining the decoded navigation data bits, but if it, for some reason, fails to do so, the ambiguous part of the carrier phase measurement is $\frac{1}{2}N'$ as in Eq. (1), while normally it is $\lambda N'$. For this paper, only Costas-type half-cycle-ambiguity receivers were used.

If the signal tracking is subject to a temporary (shorter than the sampling period) discontinuity, the ambiguity $N/S$ may change while the fractional part remains consistent, resulting in a cycle slip. Cycle slips are of persistent nature, i.e., the change in $N/S$ will affect all subsequent measurements to the same satellite as well. Due to the half-cycle ambiguity, the smallest possible cycle slip is $\pm\frac{1}{2}$ cycles, equivalent to approximately 10 cm at the L1 band. The receiver may also be, due to, e.g., multipath, subject to random-magnitude (no integer constraint) temporary measurement blunders, yielding outlier measurements. In the scope of this paper, detected errors are not categorized to slips and errors.

A new observable constructed by forming time differences of (1) is given by Eq. (2):

$$b_i^t D\Phi^S \triangleq b_i^t \Phi^S_i - \Phi^S_0$$

$$= b_i^t D\phi^S + \lambda b_i^t D\Delta N^S + \frac{1}{2} b_i^t D\Delta \Theta^S \quad (2)$$

$$+ b_i^t D\delta^S + b_i^t D\left(T^S - I^S\right) + b_i^t D\epsilon^S$$

with $b_i^t D$ denoting the time difference operator between epochs $t_b$ and $t_i$. The major advantage of this observable is that the integer ambiguity term is canceled—but possible cycle slips and outliers are not. Moreover, atmospheric and other temporally correlated errors are reduced. The price is that uncorrelated errors (noise) are amplified $\sqrt{2}$-fold.

3. METHODS

3.1. Time-Differential Processing

Forming an observable as given in Eq. (2) allows ambiguity cancelation while utilizing precise L1 carrier-phase observations. This fact is taken advantage of to obtain high relative positioning precision when measuring dynamic trajectories using low-cost single frequency receivers. The core navigation algorithm to obtain the relative solution (base vector) between two epochs is based on an iterative least-squares (LS) solution of Eq.(3), similar to code-based single point positioning (no filtering applied).

$$b_i^t D\hat{\Phi} - b_i^t D\Phi \left(\xi_{i,0}, \xi_b\right) = H_{i,0} \Delta \xi.$$  \hspace{1cm} (3)

Here $\Phi = (\Phi^1, \Phi^2, \ldots, \Phi^n)^T$ is the observation vector to all $m \geq 4$ satellites constantly tracked between $t_b$ and $t_i$, and $\xi = (x, y, z, c\delta R)^T$ is the combined position and time solution which is assumed to be known for the base epoch ($\xi_b$) but unknown for the current time ($\xi_i$). The $^\sim$ sign designates measurements while $^\epsilon$ represents modeled quantities. This (over-determined) set of equations once solved, the over-all trajectory can be reconstructed by either accumulating the position increments between subsequent epochs $t_{n-1}, t_n$ or by directly forming over-all differences between the base epoch $t_b$ and the current time $t_i$, compare Fig. 1. Each procedure has particular advantages and shortcomings which are out of this brief paper’s scope.

Note that there is no need for a second, near-by base receiver and no (static) initialization as required by RTK approaches [6]. This advantage does not come for free. The remaining, non-modeled errors (compare last line of Eq.(2)) increase with increasing processing time causing a slow position drift. This fact limits possible processing intervals to a few minutes only. The drift can only be reduced by using external corrections. It cannot be avoided but it can be monitored and quantified. More details are given in [1, 7].

3.2. Receiver Autonomous Integrity Monitoring

As any cycle slips or outliers bias the time differences, they must be detected accordingly. In traditional pseudorange positioning, measurement blunders are usually screened out by means of a RAIM process.

The need for RAIM originates from civil aviation where measurement errors are uncommon but large (dozens or hundreds of meters) and caused by the space segment. In a safety-critical application such as aviation, the user cannot wait—possibly for hours—for the control segment to detect the satellite malfunction and to upload new satellite health data. Thus, the receiver must be able to autonomously detect biased measurements in order to meet required navigation
performance specifications. As the likelihood of a satellite failure is low, many RAIM schemes assume that no more than one error can occur at a time.

The original application of RAIM totally differs from the context of cycle slip detection where errors are small (even sub-meter), occur considerably often, and are caused by, e.g., receiver dynamics and environment, i.e., in the user segment. However, RAIM is based on measurement redundancy in least-squares estimation which is directly applicable on time-differential positioning.

RAIM is based on the assumption that a biased observation in the measurement set should not fit well in the others. The least-squares residual

$$f = \hat{y} - \tilde{y} = \hat{y} - H (H^T H)^{-1} H^T \tilde{y},$$

(4)

with $H$ denoting the (linearized) measurement model and $\tilde{y}$ being the observed measurements, can be used as a measure of inconsistency. According to [8], an outlier alert is issued if the condition

$$\|f\| > T_D; \quad \text{with} \quad T_D = f(\sigma, m)$$

(5)

holds. The test statistic $\|f\|$ is $\chi^2$ distributed with $m - 4$ degrees of freedom. Consequently, $T_{D, m}$ has to be calculated from the inverse $\chi^2$ cumulative density function corresponding to the expected range measurement noise $\sigma$ and the number of used satellites $m$ for a user-defined false alarm rate $F_A, T_D$, as given by [8], for code noise typical at the time and a rather strict false alarm probability as suggested by RTCA is shown in the left-hand side plot of Fig. 2. This value is due to the civil aviation background of RAIM, and is not applicable for cycle slip detection for reasons described above.

Typically, the RAIM procedure is followed by computation of the protection level, i.e., the largest position error that may occur unnoticed with the current RAIM parameters. This is not, however, of much interest at cycle slip detection.

### 3.3. Detection and Exclusion Procedure using Time-Differences

Within this first approach, the goal of time-difference integrity monitoring is limited to outlier detection and exclusion. When using accumulated position increments for reconstructing the over-all trajectory, see Fig. 1, there is no need for cycle slip and outlier discrimination or even cycle slip repair, as each incremental solution is independent from previous measurements. This is the main benefit of the accumulation strategy. Hence, the latter issues are only marginally addressed in the remainder of this article whilst the focus is put on the exclusion task. Instead of $\|f\|$ as proposed by [8], the test statistic $\text{RMS}(f)$ as defined by (6) and (7) is used.

$$\text{RMS}(f) = \sqrt{\frac{\sum_{i=1}^{m} f_i^2}{m - 1}} = \frac{\|f\|}{\sqrt{m - 1}}$$

(6)

$$f_i = n^{-1.4} \text{D} \Phi_i - n^{-1.4} \text{D} \Phi_i$$

(7)

Here subsequent epochs $(t_{m-1}, t_{m})$, spanning intervals between 0.1 and 1 second only depending on the receiver’s measurement rate settings, are used. The effect of drifting errors virtually vanishes for such short time spans causing the residual level to be very low (in the millimeter range). Hence outliers generate distinct spikes in the test statistic. Note that this would not hold when using over-all differences $(t_i, t_b)$ for outlier detection, as the residual level increases with increasing drift of non-modeled errors. The corresponding decision rule is stated in Eq. (8):

$$\text{RMS}(f) > T_D = \text{constant}$$

(8)

$$\|f\| > T_D^{\text{m}} \sqrt{m - 1}$$

(9)

Here $T_D^{\text{m}}$ is independent of the number of satellites $m$. However the effective threshold for $\|f\|$ (i.e. if $\|f\|$ was chosen as test statistic instead of RMS $(f)$) does adapt to $m$, Eq. (9).

As shown by the right-hand side plot of Fig. 2, this adaption is only off by less than 5% from the threshold values proposed by [8] for typical time-difference applications—if $T_D$ is set ‘smartly’. Smartly means: (1) Choose an appropriate false alarm rate, e.g. 0.1% for safety-critical applications. (2) Determine the expected measurement noise. This can be done by empirical experience with the used equipment or by using the variance estimate as dropping out of the least-squares solver: $\sigma_{\Phi, m} = (\sum f_i^2 / (m - 4))^{1/2}$. A phase noise standard deviation of 1 cm is typical for low-cost receivers in dynamic applications. (3) Determine a typical number of satellites of the data to be processed, e.g. 9. (4) Get $T_D$ from the inverse $\chi^2$ cumulative density function for the chosen values (here 5 d.o.f.). (5) Calculate $T_D^{\text{m}}$. This done, a detection and exclusion logic as outlined by Fig. 3 can be executed.

![Fig. 2. Outlier detection threshold vs. number of satellites.](image)

Left: Threshold as initially proposed by [8] when using $\|f\|$ as test statistic for code measurements with standard deviation as high as 33 m and a false-alarm rate $3.33 \cdot 10^{-7}$. Right: Threshold according to [8] and the effective threshold from Eq. (9) (using $T_D = 1.6$ cm) with noise and false-alarm settings appropriate for carrier phase time-differences.

Here subsequent epochs $(t_{m-1}, t_{m})$, spanning intervals between 0.1 and 1 second only depending on the receiver’s measurement rate settings, are used. The effect of drifting errors virtually vanishes for such short time spans causing the residual level to be very low (in the millimeter range). Hence outliers generate distinct spikes in the test statistic. Note that this would not hold when using over-all differences $(t_i, t_b)$ for outlier detection, as the residual level increases with increasing drift of non-modeled errors. The corresponding decision rule is stated in Eq. (8):
4. PERFORMANCE TESTING

The method was tested using various sets of authentic GPS data, all logged for post-processing by low-cost u-Blox L1-only GPS receivers capable of raw-data sampling rates up to 10 Hz. During the first flight test, a stationary base receiver was available for RTK computations. Precise ephemerides, clock corrections, and ionosphere data were used.

4.1. Static Test

The static test was expected to be an easy starting point for validating the correctness of the theory. Stationary receivers are less prone to cycle slips than moving ones and the presence of an outlier is easy to observe from the solution trajectory as it should not contain distinctive jumps. Fig. 4 demonstrates the effect of outliers in the position domain and the top plot of Fig. 5 shows the corresponding test statistic values. The processed interval was short, only 22 seconds, but contained two measurement errors, both on the same satellite (PRN 4), which are well visible in the trajectory computed without error detection. 9 satellites were available, and only the remaining 8 were used after the outlier was detected.

The pattern of the residuals as depicted in the second plot of Fig. 5 reveals a magnitude of \( \frac{1}{2} \) of the measurement errors proposing the presence of half-cycle slips. However, two subsequent epochs are affected in the first case. This indicates the presence of a simple outlier instead of a slip impeding a concluding discrimination. For validation purposes, the non-differenced observations of PRN 4 were interpolated using cubic smoothing splines yielding residuals with a standard deviation similar to the one expected for the range measurements. The excluded measurements were not used when fitting. The fit residuals (bottom plot of Fig. 5) confirm the outlier hypothesis in the first case. In the second case the error is very small which prevents a solid discrimination.

4.2. Flight Test #1: Comparison With RTK

A kinematic test was performed using a data set measured during a test flight starting from Oberpaffenhofen, Germany. The data begins with a 15-minute stationary section to facilitate RTK integer ambiguity resolution, followed by a flight containing circle and dynamic-soaring-like maneuvers. As

![Fig. 3. Outlier detection and exclusion algorithm. Cycle slip discrimination and repairing make more measurements usable but are not addressed in the scope of this paper.](image)

![Fig. 4. Static data: Analysis in position domain. East and North components of the resulting trajectory with and without outlier detection. The initial location is marked with a circle.](image)
opposed to TD processing, RTK is unaffected by error drift but requires two receivers and an initialization to work [6]. The cycle slip detection results of the time-differential solution were compared to those derived from an RTK solution computed at Delft University of Technology, the Netherlands. A total of 165 errors, classified as cycle slips and outliers, were listed in the reference. The time-differential method detected 127 uncategorized errors. For the epochs where the TD method identified an error, time differences of the carrier phase measurements for the suspected satellite were examined to see if an error really had occurred or not. Table 1 shows how many detections were observed to be correct, wrong, or, as was the case for most epochs, uncertain due to, e.g., gaps in the phase data or, more frequently, nonuniform sampling of the phase measurements. The data was logged at 4 Hz but the measurement instants were not spaced by exactly 0.250 seconds. Instead, the measurement timestamps had second fractions .217, .499, .747, and .999. Time-differencing such data yields oscillatory results as every other measurement interval is slightly longer than the others, making visual inspection of the presence of cycle slips difficult, especially if the data originates from a highly dynamic scenario—as was the case. A resampling software was not used to mitigate the zigzag effect.

The detection results of the two methods were not even expected to be identical. Firstly, the RTK solution did not report any half-cycle slips which was due to the software not being configured to do so. However, the time-differential method did detect some half-cycle slips. Secondly, RTK solutions are computed from double-differenced data which causes cycle slips from both the rover and reference receivers to be subtracted from each other. Thus, a slip detected in RTK can have occurred in either of the two receivers. Even though the rover receiver is more prone to cycle slipping, several RTK detections were observed to have occurred in the reference receiver.

In the verifiable cases, the methods achieved a similar performance: supposing that RTK was right in cases where the TD method was verified to be wrong and the methods disagreed, both methods were right in about 50% of the verifiable epochs. Knowing that the flight data was highly dynamic and frequently suffered from a low number of visible satellites, this may be regarded as a fairly good performance, but cannot be considered conclusive.

### 4.3. Flight Test #2: Simultaneous Errors

As a real-life application, the time-differential method was used for estimating the takeoff and landing distances of the Mi30 “Schlacro” aircraft of AKAFLIEG München. The measurement process was repeated six times. During the landings, the aircraft bounced remarkably after hitting the ground, resulting in excessive losses of lock. One landing measurement is taken here to show the behavior of the presented method in the case of multiple simultaneous errors. The reconstructed altitude profile is drawn in Fig. 6. The zoomed version shows an abrupt jump of about 20 cm in the altitude. That section was processed continuously with carrier phases and as the data was logged at 10 Hz, it is not a plausible explanation that the dynamics would have changed suddenly. At the epoch of interest, two cycle slips were detected. It is suspected that these identifications are incorrect.

Fig. 7 shows the time-differenced carrier phases for the used satellites, revealing three half-cycle slips. The error detection algorithm identified satellites 21 and 22 as faulty, but as it can be seen in the figure, PRN 22 is healthy. Thus, two of the remaining cycle slips remain undetected. Forcing PRN 21 to be totally excluded from the computations did not help: a false solution fits well enough in the measurement subset containing two biased measurements, and thus the exclusion al-

---

### Table 1. Comparison of TD and RTK error detection results verified using raw phase data.

<table>
<thead>
<tr>
<th>Detection results</th>
<th>TD Right</th>
<th>TD Wrong</th>
<th>Uncertain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods Agree</td>
<td>6</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Methods Disagree</td>
<td>9</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td>TD Detection only</td>
<td>2</td>
<td>4</td>
<td>31</td>
</tr>
</tbody>
</table>
algorithm terminates after finding two satellites to be excluded. Therefore, the least-squares residual itself is not a sufficient quantity for detecting multiple simultaneous cycle slips in this case, which agrees with the single-outlier assumption in most RAIM schemes mentioned in the above.

5. CONCLUSIONS AND OUTLOOK

In this paper, the problem of cycle slipping was tackled by means of RAIM. The results show that despite the totally different origin, RAIM can detect even small single cycle slips and outliers in carrier phase data. Similar promising results have been observed in other tests [10].

As time-differential positioning and the presented outlier detection approach are relatively simple and easy-to-apply in the field procedures, they can be well used in not-safety-critical applications. Moreover, with an effective discrimination between outliers and cycle slips, the detected cycle slips could be corrected for and the method could be used as a pre-processing step for other phase processing methods. As future work, the performance with multiple simultaneous errors is to be further investigated.

6. ACKNOWLEDGEMENTS

The authors would like to acknowledge Dr. Oliver Montenbruck, DLR Germany, for strongly supporting the work on the time-differential approach.

7. REFERENCES


