A Kalman filter approach to reduce position error for pedestrian applications in areas of bad GPS reception

Mattias Eliasson

Mattias Eliasson
Spring 2014
Degree Project, 15 hp
Supervisor: Lars-Erik Janlert
Examiner: Pedher Johansson
Bachelor's programme in Computing science, 180 hp
Abstract

The number of GPS-enabled devices are growing rapidly. A large segment of the growth is coupled to the growth of smartphones. Some location-based applications are relatively simple, requiring only a rough position estimate. Other applications provide services where the usefulness of the application is directly connected to the accuracy of GPS positioning. Lifelogging, fitness, and navigation are some types of applications where precise location estimation greatly benefits the user.

The GPS technology is available worldwide and 24 hours per day. Its accuracy is not uniform, varying over time of day and place. Buildings can reflect or block the signal, the atmosphere delays it, satellite clock and orbit errors introduce bias. These are some of the error sources affecting GPS positioning.

Many applications with need of high accuracy are used in everyday life. Users will eventually venture into areas unsuitable for GPS positioning. In these situations, these applications may not function sufficiently well.

In this thesis, a data fusion method called the Kalman filter is evaluated as a means of improving the GPS positioning. A simple motion model is employed, tracking the position and velocity. The motion model utilizes sensors commonly available in a modern day smartphone. The Kalman filter will be evaluated through comparison to the raw data and a simple moving average filter.

The results show that the Kalman filter is able to significantly reduce the variance compared to the raw data, but not significantly lower than the moving average filter.
Acknowledgements

I would like to acknowledge the help from my supervisor, Lars-Erik Janlert, for nudging me in the right direction whenever it was necessary.

Peter, for the late night phone calls to discuss things that felt out of place. Thank you.

Finally, I would like to thank Lenita for having patience with me during these months, allowing me to constantly think out loud about this project, and for all the help and worthwhile discussions.
Contents

1 Introduction 1
   1.1 Problem Specification 1
   1.2 Combining GPS and INS 2
   1.3 Purpose 2

2 Global Positioning System 5
   2.1 Position estimation 6
   2.2 Speed estimation 7
   2.3 Sources of Error 7
   2.4 Differential GPS signal correction 10

3 Filter methods 11
   3.1 Kalman filter 11
   3.2 Moving average 13

4 Data collection 15
   4.1 Test environment conditions 15
   4.2 The measurement application 16
   4.3 Statistical methods 17
   4.4 The data set 18

5 Results 19
   5.1 Descriptive statistics 19
   5.2 Comparison tests 20

6 Conclusion and Discussion 21
   6.1 Regarding GPS reception conditions 21
   6.2 Regarding the measurement path 22
   6.3 Ethical aspects 23
6.4 Societal aspects  

6.5 Future work  

References
1 Introduction

The need to accurately determine the position and direction has been present for human explorers for millennia. Throughout the centuries, many clever ways to navigate have been developed. Angular measurements of the sun and stars by hand or with astrolabes, and later sextants, dates back six thousand years. This allowed sailors and travelers to measure the latitude. The compass enabled explorers to keep a straight heading over long distances, without the need for observing celestial bodies. To precisely determine the longitude, however, was an unsolved problem for many centuries. The problem of estimating the longitude eventually became a synonym for performing an impossible feat. The trick lay in keeping accurate time, both local time and the time at a port, from which the difference in longitude could be calculated. Portable clocks could not be made accurate enough to precisely keep time at sea. Other attempts for measuring time were used. One method was to keep calendars of astronomical observations, and by observing the transit of the moon and the position of the moons of Jupiter, the time could be estimated. The calculations needed were complex and resulted in large errors. A better solution came in 1761 when John Harrison invented the chronometer. By creating an almost friction free clock, without a pendulum and with metal alloys to compensate for metal expansions in changing temperatures, a sufficiently correct clock rate could be kept. [13]

1.1 Problem Specification

The number of devices with integrated support for the Global Positioning System (GPS) are growing at a rapid rate. Today the technology is applied in a varied range of industries. For example: aviation, agriculture, traffic systems, emergency systems, surveying, environmental protection, recreation and many others all successfully employ GPS. A large segment of the current growth of GPS devices is coupled to the growth of the smartphone market. Some location-based applications for the smartphone market are relatively simple, aimed at determining a rough position of the user. Other applications provide services for which the user experience is enhanced as the position accuracy increases. One example is lifelogging applications, where the location and movements of the user is tracked continuously, measuring not only the location, but also in what manner the user moves about and how far the user has moved. Another is high-precision geofencing, virtual perimeters where the user gets notified in the event of the device entering or leaving the area. Others are context and location aware applications, taking relevant information from the surroundings and supplies this to the user in appropriate ways. All of these examples are applications used in part in a pedestrian environment, where the device is carried on the person moving about on foot. Many of them are also meant to be used in everyday life, where users eventually will venture into areas unsuitable for GPS positioning and the necessary level of accuracy cannot be achieved. The GPS accuracy are affected by many different factors, including satellite clock
drift, atmospheric interference and signal reflection. The solution to tracking the location of an object when GPS cannot be relied on involves a suitable inertial navigation system. Inertial Navigation Systems (INS) are systems where the device monitors its own position, only utilizing internal motion sensors. Devices used in a pedestrian environment introduces limitations to what kind of INS can be used, so previously successful INS models for use in vehicles are not necessarily applicable, as their weight, size and ergonomics needs to be considered. Furthermore, how the device is placed on the body has an influence on the sensor readings [1]. As smartphones has many other uses than tracking user location, changes in orientation and acceleration not directly related to user movement must be accounted for.

1.2 Combining GPS and INS

The Kalman filter [2] is a method that can be used for combining a model of a system with a set of noisy sensor measurements to produce an estimate of the underlying state, such as the position. The method is divided into two steps, the time update and the measurement update. In the time update, a motion model is used to predict the future state of the system. The measurement update then produces a new corrected estimate, by combining the prediction with the sensor measurements. The first application of the Kalman filter was by Stanley Schmidt for trajectory estimation in the Apollo program. [3]

Much research has been devoted to inertial navigation systems for pedestrian use. Many use sensors fixed on body parts in order to obtain more accurate and relevant readings. Inertial measurement units placed on feet [4, 5, 6, 7, 8] is a common method. One advantage of using foot-mounted inertial sensors is that sensors can be reset to remove drift at each step. Gablagio et al. [9] used the Kalman filter to augment pedestrian GPS navigation, placing accelerometers and gyroscopes vertically along the thorax and oriented along the walk direction. Sensor placement on the torso are ideal for determining movement direction. Inertial sensors worn at the waist [10, 11] are easily placed and provide accurate readings of movement direction similar to that of torso placement. Head-mounted inertial navigation systems have also been developed. [12]

1.3 Purpose

The purpose of this thesis is to evaluate the Kalman filter method as a means of significantly reducing noise in GPS position estimation. The Kalman filter will be tested under conditions where the GPS signal is considered degraded and the position estimation is affected adversely. The sensors are limited to those commonly available in a smartphone today, carried on the person but not fastened on the body. These sensors are accelerometer, gyroscope and magnetic compass. The environment is supposed to reflect ordinary circumstances for a modern-day smartphone. Under normal circumstances, the device can be used for other purposes simultaneously as the location is being tracked. This can alter the orientation of the device and its acceleration in unpredictable ways not directly related to the direction of movement. As the true positions are unknown, whether the noise is reduced will be observed indirectly by estimating distance along a path of known length. The distance estimation from the Kalman filter is compared to the raw data and a simple moving average filter. The method with better performance will be determined by performing statistical
analysis on the collected data.

This thesis aims to answer the following questions:

- Will a Kalman filter significantly improve the accuracy of distance estimation for GPS in areas of bad signal reception, compared to the unfiltered raw data?

- How large is the improvement by using Kalman filtered data for distance estimation compared to using a simple filter method such as the moving average to reduce noise?
2 Global Positioning System

The Global Positioning System is a satellite-based system for high-accuracy position, velocity and time estimation. It was originally developed by the U.S. for military applications with the objective of being available worldwide, under different weather conditions, and during 24 hours per day. The system can be divided into three segments, the space segment, the control segment, and the user segment.

The space segment consists of the satellites in orbit, dispersed uniformly in six orbits such that at any point in time, three or more satellites are in view anywhere on earth. The satellites are continuously transmitting navigation messages containing information about their location in orbit and the time of transmission, to be acquired by the GPS receivers in the user segment. This one-way broadcasting of data from the satellites to the user segment allows for an unlimited amount of receivers to simultaneously utilize GPS.

The control segment is a global network of ground-level monitoring stations and control facilities, responsible for maintaining the proper functioning of the satellite system. The control facilities monitor their health, the state of the solar arrays, battery power and maneuverability. Due to external factors, the satellites do not perfectly follow a fixed orbit around the earth. The position of the satellites in their orbit are closely observed by the monitoring stations and any deviation between their actual location and the information sent in the navigation message is corrected. The control segment also monitors satellite clock drift, and uploads correction information at regular intervals.

The user segment consists of GPS receiver equipment, capable of processing the transmitted signals to estimate the range to each satellite. With enough satellites in view, the three-dimensional position of the GPS receiver, its velocity and the local time can be estimated. [14, 15, 16]

The coordinate system used in GPS to express a point on earth is called the World Geodetic System (WGS). The current standard, WGS84 [18], was established in 1984 and is an ellipsoidal model of the earth. [15, p.29] The global coordinates consists of three components, latitude, longitude and altitude. The longitude measures the angle in degrees east and west from the Greenwich meridian, spanning 0 – 180° east and 0 – (-180°) west. The latitude measures the angle from the equator towards the poles, with 0 – 90° north and 0 – (-90°) south.

Until May 1, 2000, the GPS signal included intentional pseudo-random noise, called Selective Availability. This noise increased the error margin by up to 100 meters horizontally and 50 meter vertically. The intention of this mechanism was to control the accuracy of navigation and to degrade the signal for users without military receivers and a daily key. By using correction information from reference stations at known locations, Selective Availability could be circumvented to some extent. [14, p.120-121] This has since then been turned off and as a result the commercial and private applications of GPS increased rapidly. The use of GPS is now fast growing, having potential applications in many different types...
of industries. The main application of GPS has been for navigation and surveying on land, air and at sea. More recently GPS has been used in agriculture, emergency systems, robot tracking, traffic systems and for recreational purposes in hunting, sailing, hiking and many others. [17]

2.1 Position estimation

To determine the three-dimensional position, the time, and the velocity of the GPS receiver, a technique called time-of-arrival ranging is used. This amounts to measuring the transit time for a signal to propagate from the satellite to the receiver. The range between the satellite and the receiver is then calculated by multiplying the transit time by the speed of light. Because the transit time measurement includes the unknown receiver clock bias, the distance is considered an estimate. The term pseudorange is used to denote the satellite-to-receiver range before corrections have been made. Finally, through trilateration with the calculated pseudoranges from several satellites, the GPS receiver can estimate its position and local time.

Estimating the pseudorange from one satellite defines a spherical area with possible positions located on its surface, given that the pseudorange to the satellite is correct. Using data from a second satellite reduces the number of possible locations to those on a circle at the intersection between the two spheres. To solve for three unknowns, longitude, latitude and altitude, information from three visible satellites are required. However, as the signal travels at the speed of light, very precise time synchronization is critical to accurately estimate the distance to each satellite. The atomic clocks on the satellites are extremely accurate, but the receiver clock can usually not provide accuracy at the same level, so without correction the estimated position would not be useful. A discrepancy of 1 ms between satellite time and receiver time yields a pseudorange offset of 300km. To determine the receiver clock bias, together with the three position estimates, a fourth satellite is used. If the altitude is known or can be assumed to be near sea level, as for example with ships, three satellites in view will suffice to obtain accurate results.

Each GPS satellite continuously broadcasts navigation messages on two frequencies, $L_1$ and $L_2$. $L_1$ contains two codes, the public Coarse/Acquisition (C/A) code and the encrypted Precision (P(Y)) code. The signal on the $L_2$ frequency only contains the P(Y) code, except for more modern satellites transmitting a civilian code on the $L_2$ frequency, called L2C. The P(Y) code enables higher position accuracy, but is restricted to military GPS receivers. The C/A code is a pseudorandom (PRN) code that repeats every millisecond. To determine the transit time, the GPS receiver first generates its own PRN code. By comparing the receiver generated PRN code with the incoming satellite signal, the time shift can be determined. This time shift corresponds to the propagation time of the navigation message, from which the range can be calculated. The navigation data can also be extracted from the signal. The navigation message contains information regarding the satellite position in orbit at the time of message transmission, time correction data and orbit almanac. The almanac contains course long-term orbit information. This information is valid for roughly 180 days and allows for quick satellite acquisition at receiver startup.
2.2 Speed estimation

The velocity of the receiver can be determined by measuring the change in position over time with satisfactory results given that the receiver velocity is near constant during the time period. In modern GPS receivers, other more precise measurements of the velocity are available by estimating the Doppler frequency of the received signal. The Doppler shift is caused by the relative motion of the satellite compared to the receiver. [15, p.58] The accuracy of the velocity computation is dependent on correct information regarding satellite position in orbit, satellite velocity, and the accuracy of the user time and position estimates. Therefore, in order to obtain a good estimation, four or more visible satellites are required. [15, p.61]

2.3 Sources of Error

The satellite signal is subject to various sources of error before reaching the receiver. These error sources degrade the signal, decreasing the accuracy of the pseudorange and subsequently the obtained position estimate. The errors can be systematic in that the error resulting from these sources are more or less a constant bias, which effect persists over a longer period of time, or they can be random sources of error, contributing to signal noise and changing rapidly. Various methods can be employed to mitigate these effects, to a varying degree of success depending on the surrounding environment.

2.3.1 Geometric dilution of precision

The geometry of how the satellites are arranged in their orbits relative to the receiver has an effect on the position estimate. The ideal situation is one satellite above and the remaining dispersed evenly around the receiver near the horizon. Satellites clustered together will yield nearly equal pseudorange estimates and will not provide additional information. Small errors are therefore greatly increased. The effect of the geometric satellite constellation is called geometric dilution of precision (GDOP), and has a multiplicative effect on range errors. [14, p.39] The effect of satellite clustering on the position estimate can be explained visually using a two-dimensional position example. A GPS receiver is placed on a 2D plane and measures the pseudorange to one visible satellite. The pseudorange has an error margin, so the possible positions are located on a band of points surrounding the satellite at a certain distance. Pseudorange measurements from a second satellite at a favorable location, with respect to GDOP, reduces the possible positions to a small area at the intersection of the two bands. This is illustrated on the left side of Figure 1. On the right side of Figure 1 the arrangement of the satellites have changed so that they are closer together, and as a result, the error has grown in comparison to the prior satellite arrangement. More satellites in view can improve the GDOP value, either by using more than the four required satellites to solve the position equations or allow the receiver to select a more favorable subset of the satellites [19].
2.3.2 Ephemeris errors

Over time, the satellites are affected by solar radiation and the gravitational pull of the sun and the moon, gradually altering their orbits. These changes are continually observed by the control segment from monitor stations at precisely known locations. This allows the monitor station to perform inverted positioning as if the satellites themselves were users. [14, p.185] Model-based corrections of the future orbits are predicted and uploaded to the satellites to be further transmitted to the receivers in the navigation message. The prediction residual on the satellite orbit position is in the range of 1-6 m. However, the magnitude of the satellite orbital position error does not directly correspond to pseudorange error. The average pseudorange error due to ephemeris prediction error is 0.8 m. [15, p.305]

2.3.3 Satellite clock errors

The atomic clocks onboard the satellites are required to be very accurate, but small drifts do occur over time. This drift would have a large effect on the location estimate error margin if not corrected for at regular intervals. An error of 10ns in the satellite clock translates to a 3 m range error at the receiver. The satellite clock drift is monitored by observation facilities in the control segment. Due to the difficulty of synchronizing the time across all satellites, clock correction data is generated and instead sent along in the GPS navigation message to allow the receiver to correct for the clock drift. [14, p.185-186] The residual clock error after correction at the receiver ranges from 0.3-4 m depending on satellite and time since the last clock drift update. [15, p.304-305]

2.3.4 Atmospheric errors

As the satellite signal is transmitted through the different layers of the atmosphere, certain factors affect the speed of the signal. Factors such as the refraction index of the medium the signal propagates through, and the distance of air mass the signal must pass through before

---

Figure 1: Visual representation of how satellite clustering adversely affects the position estimate.
it reaches the receiver. The amount of airmass is greater for satellites nearer the horizon relative to the receiver, compared to satellites at larger angles. Given that a rough position of the receiver and some meteorological parameters are known, this error can be modeled and be largely compensated for.

**Ionospheric errors** The ionosphere is the upper region of the atmosphere, extending from 85 km of the earth’s surface up to 1000 km, consisting of gases ionized by solar radiation. These ionized gases disperse the GPS signal. The error due to this dispersion is proportional to the amount of ionization, or total electron content (TEC). The TEC varies depending on the amount of solar radiation and latitude. The daily cycle of solar radiation reaches its minimum a few hours after midnight, and increases with higher latitudes with a maximum at the poles. Further variations depend on solar activities in longer cycles. The TEC in the ionosphere is not homogenous, so local variations do occur as well. The ionospheric errors mainly affect the transmission speed of the signal, causing a delay, with the error inversely proportional to the frequency of the signal. By comparing the different times of arrival of the L\textsubscript{1} and L\textsubscript{2} frequencies, the ionospheric delay error can be estimated and compensated for with high accuracy. [14, p.152] For users with single frequency GPS receivers, the Klobuchar single frequency ionospheric model [20] is used to estimate the delay. The Klobuchar model can reduce residuals to the ionospheric error by up to 50 %. [14, p.153] Averaged out across elevation angles and over the globe, the residual correspond to 7m error. [15, p.314]

**Tropospheric errors** The troposphere is the lower region of the atmosphere, consisting of wet and dry gases. In contrast to the ionosphere, the effect of the gases in the troposphere on the GPS signal is not frequency dependent, so signal delay between the two frequencies cannot be used to model the refraction effect and mitigate the pseudorange error. The dry component accounts for 90 % of the error and can be predicted accurately. The wet part is harder to model due to unpredictable water vapor fluctuations in the atmosphere. For receivers at sea level, if uncompensated, the error is in the range 2.4-25m, depending on the angle to the satellite from the view of the receiver. The maximum error for satellites located near the horizon and minimum with satellites at the zenith. [15, p.314] Given knowledge of temperature, humidity and air pressure along the signal path, tropospheric refraction can be modelled to compensate for its effect. [21, 22] These parameters do however vary over time and location, and accurate measurements from meteorological sensors is not necessarily readily available for public users with hand-held devices. Standard empirical models [23] of the tropospheric error can be applied, reducing the need for meteorological measurement.

### 2.3.5 Multipath effects

The environment surrounding the receiver might contain objects that can reflect or diffract the incoming signal, such as buildings, mountains, dense foliage or hard ground. This type of errors are called multipath errors. Compared to the direct signal, reflected signals travel a longer path before arriving at the receiver. The multipath signals are therefore delayed in proportion to the increased length of the path. This delay can cause inaccuracies in estimating the location. For longer delays the inaccurate signal can be detected and discarded without adverse effects on performance. For shorter delays, such as when the signal reflects on the ground near the receiver, the signal is superimposed on the direct
signal, distorting the estimate. The size of the error depends not only on the time delay but also on the power of the multipath signal compared to the direct signal. In some cases where the direct signal is partially or entirely blocked by buildings or trees, only the reflected signal might be available for the receiver. [15, p. 279-280] To mitigate multipath effects where signals are reflected off the nearby ground the GPS receiver can be placed lower, in order to reduce the amount of reflected signals and their delay. This method might be unsuitable if the surrounding terrain is not open and provides a large clear view of the sky. Other ways to reduce multipath errors is to use antennas that only record incoming signals from where they are expected to arrive, mainly the sky at angles near or above the horizon, and not from below. [15, p. 293]

2.4 Differential GPS signal correction

Differential GPS (DGPS) is a method to reduce the error in GPS position estimates. For two GPS receivers in operation relatively nearby, the error in the position estimates can be assumed to be similar. If one of the receivers is located at a known location, correction information can be transmitted to reduce the error for other receivers. Some errors, such as satellite clock drift and ephemeris error can effectively be removed if the same satellites are in view for both receivers. The effectiveness of the correction for some of the other sources of error depend on the distance between the reference site and the DGPS receiver. Ionospheric and tropospheric errors are not as effectively reduced by DGPS, as the transmission path to the receivers differ, with different length of atmosphere to pass through, unless the two receivers are in close proximity. [15, p. 381]
3 Filter methods

In this chapter the theoretical concepts of the methods used to filter the raw GPS data are explained.

3.1 Kalman filter

The Kalman filter \[2\] is a data fusion algorithm for combining a stream of noisy sensor measurements over time with a model based prediction. The filter method is then able to obtain an optimal estimate (minimum mean squared error) of the state of an uncertain and dynamic linear system. The sensors used in the algorithm can for example be GPS receivers, accelerometers, gyroscopes, compass, or any other input that can be relevant for determining the state of the system. The state can be the position, velocity, acceleration, altitude or some other aspect of interest. The method alternates between two steps, the time update step and the measurement update step. In the time step the state of the system is predicted forward in time using a model based prediction given the current system state as input. In the measurement update step, the prediction is corrected using an weighted average of the noisy sensory input based on the noise and estimated confidence for each sensor. \[26\] The equations for the time update step are listed next.

\[
\begin{align*}
\hat{x}_{t}^- &= A\hat{x}_{t-1} + Bu_t \\
P_{t}^- &= AP_{t-1}A^T + Q_t
\end{align*}
\]

In Equation 3.1, \(t\) denotes the current time step, \(\hat{x}_t\) is the estimated state vector at time \(t\), \(A\) is the system state transition matrix, transforming the state vector at time \(t-1\) to the state vector at time \(t\) according to the motion model. The vector \(u_t\) is control input and \(B\) is the control transition model, transforming the control input to state vector units. In Equation 3.2, \(P_t\) is the state vector covariance matrix. The noise in the process motion model is assumed to be distributed according to a multivariate normal distribution with zero mean and covariance \(Q_t\), which is the process variance for each state vector parameter. The superscript (-) denotes that the variable is predicted using prior estimates, and the hat notation (\(^\hat{\cdot}\)) denotes the variable is an estimate. The next step in the Kalman filter algorithm is the measurement update, with following equations:

\[
\begin{align*}
K_t &= P_t^H (HP_t H^T + R_t)^{-1} \\
\hat{x}_t &= \hat{x}_t^- + K_t(z_t - H\hat{x}_t^-) \\
P_t &= (I - K_t H)P_t^-
\end{align*}
\]
In Equation 3.3, \( K_t \) is the Kalman gain, corresponding to the confidence given to each sensor. \( H \) is the model of how the sensor measurements affect the system state, a transition matrix transforming measurements to state vector parameter units. \( R_t \) is the sensor noise covariance matrix. In Equation 3.4, \( z_t \) is the measurements vector obtained from sensor input.

As the present study does not include any control input and the measurements have a one-to-one correspondence to the state vector parameters, by setting \( u_t \) to 0 and \( H \) to 1, the time and measurement equations can be simplified in the following way:

\[
\hat{x}_t^- = A\hat{x}_{t-1} \\
P_t^- = AP_{t-1}A^T + Q_t \\
K_t = P_t^- (P_t^- + R_t)^{-1} \\
\hat{x}_t = \hat{x}_t^- + K_t(z_t - \hat{x}_t^-) \\
P_t = (I - K_t)P_t^-
\]

In Equation 3.8 we can note the influence of the sensor measurement noise matrix \( R_t \) on the Kalman gain. If the noise is large in comparison to the state vector covariance \( P_t \), the Kalman gain approaches zero. For the opposite case, the Kalman gain approaches one. This further affects the state vector update in Equation 3.9 and state vector covariance update in Equation 3.10. For large measurement noise the Kalman gain will pull the state vector update closer to the predicted state. A Kalman gain approaching one will instead pull the corrected state toward the sensor measurements. In this study, the state vector parameters of interest are the position on the x and y axes and the velocity along these axes. The state vector and transition model in Equation 3.6 can be described as follows:

\[
x_t = \begin{bmatrix} x_t \\ \dot{x}_t \\ y_t \\ \dot{y}_t \end{bmatrix} \\
A = \begin{bmatrix} 1 & \delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \delta t \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

The dot-notation in Equation 3.11 represents velocity along respective axis. For the transition model in Equation 3.12, \( \delta t \) is the time elapsed since the previous iteration. The transition model states that the position at the next time step is predicted to be the current position plus the current velocity multiplied by the elapsed time between the time steps.

Similarly to the state parameter vector, the measurement vector is defined as
\[
\begin{bmatrix}
  x_t \\
  \dot{x}_t \\
  y_t \\
  \dot{y}_t
\end{bmatrix}
\]

where the position coordinates are in cartesian coordinates and the speed is divided into respective x and y direction components prior to being input to the Kalman method in the measurement update step.

The change in velocity is not accounted for in the motion model. A constant velocity is in this case not a realistic model, so the acceleration will instead be incorporated into the model through the process noise matrix \(Q_t\). The influence of acceleration noise on the state vector during the time period \(\delta t\) can be defined as in Equation 3.13.

\[
Gv = \begin{bmatrix}
\frac{\delta^2}{\delta t^2} & 0 & 0 & 0 \\
0 & \frac{\delta^2}{\delta t^2} & 0 & 0 \\
0 & 0 & \frac{\delta^2}{\delta t^2} & 0 \\
0 & 0 & 0 & \frac{\delta^2}{\delta t^2}
\end{bmatrix} \begin{bmatrix}
  v_x \\
  v_y
\end{bmatrix}, \tag{3.13}
\]

In Equation 3.13, \(v\) is the acceleration noise magnitude. \(Q_t\), the variance of the acceleration over time period \(\delta t\), is then defined as in Equation 3.14.

\[
Q_t = G\sigma^2_{acc} G^T = \sigma^2_{acc} \begin{bmatrix}
\frac{\delta^4}{\delta t^4} & \frac{\delta^3}{\delta t^3} & 0 & 0 \\
\frac{\delta^3}{\delta t^3} & \frac{\delta^2}{\delta t^2} & 0 & 0 \\
0 & 0 & \frac{\delta^2}{\delta t^2} & 0 \\
0 & 0 & 0 & \frac{\delta^2}{\delta t^2}
\end{bmatrix} \tag{3.14}
\]

At this point the definition of the Kalman filter with a two-dimensional position and constant velocity motion model is complete. What needs to be user-specified is the acceleration variance magnitude, \(\sigma^2_{acc}\), the measurement noise covariance matrix \(R_t\), the elapsed time between time steps, \(\delta t\), and the measurement vector \(z_t\).

### 3.2 Moving average

A central moving average is a method used to smooth a sequence of data points by replacing each point by an average calculated over a small subsequence around the data point. The moving average, \(s_i\) for the sequence of data points \(x_j, j = 1, \ldots, N\), with \(n\) as the size of the subsequence, is defined in Equation 3.15.

\[
s_i = \frac{1}{n} \sum_{j=i-[n/2]}^{i+[n/2]} x_j \tag{3.15}
\]
The central moving average is only valid for odd numbered subsequence sizes and where \( i > \lfloor \frac{n}{2} \rfloor \) and \( i < N - \lfloor \frac{n}{2} \rfloor \). For data points close to the start and end points where the number of neighboring points prior and after the center is not sufficient to span the entire prespecified length of the subsequence, the size of the subsequence is adjusted. This study used a subsequence of size five datapoints for smoothing. For data points \( x_j \), \( j = 1, \ldots, 10 \), using a subsequence size of five, the first four data points in the moving average sequence, \( s_i \), are calculated as in Equation 3.16.

\[
\begin{align*}
    s_1 &= x_1 \\
    s_2 &= \frac{1}{3}(x_1 + x_2 + x_3) \\
    s_3 &= \frac{1}{5}(x_1 + x_2 + x_3 + x_4 + x_5) \\
    s_4 &= \frac{1}{5}(x_2 + x_3 + x_4 + x_5 + x_6)
\end{align*}
\]
4 Data collection

This chapter describes the process of performing the measurements, how the data is collected and which statistical methods are conducted. To evaluate the Kalman filter, an Android application is developed, capable of making GPS measurements at regular intervals and collecting the data. The raw data is used to feed the Kalman filter and the moving average filter, to provide three distance estimates. The true distance walked for each observation is measured using a measuring wheel. In total, the distance measurements on the selected path are repeated ten times.

4.1 Test environment conditions

The test environment is chosen so that many GPS error sources are prevalent. The track is placed alongside a building with several wings which can be followed in order to partially cover different areas of the sky during different parts of the track. Covering large areas of the sky increases the GDOP values for the remaining visible satellites, and require several handoffs as satellites enters and leaves line of sight. This environment enables multipath reflection errors to occur as well. Trees are present on parts the selected track, covering parts of the sky. All measurements performed are done in the time span of a few hours during which the cloud coverage is full. Furthermore, the mobile phone used as a GPS receiver and measurement device is kept in the pocket to further reduce the reception, except for starting and stopping the measurements. The track is selected such that the elevation is relatively constant from start to finish. This is done in part to reduce the error as the measuring wheel measures the total distance including altitude changes, and the GPS measurements here includes two-dimensional movement, but also to be able to keep a more constant speed. The entire length of the track from start to finish is roughly 300m. The start of the track is set a few meters from a wall, in order to obtain a relatively good start position, after which the path continues along the wall of the building. The approximate path is lined out in red in Figure 2.

Figure 2: Approximate path along which the measurements are performed.
4.2 The measurement application

To measure the distance along the path, location changes needs to be continually monitored. For this purpose, an Android application is developed, utilizing the Google Maps Location API to track the device location. In this application, the Kalman filter and moving average is implemented. For specifications regarding the LG Nexus 4 smartphone that is used, or the Android operating system, see references [24] and [25] respectively.

Prior to feeding the raw data to the Kalman filter and moving average filter, the estimated GPS coordinates are transformed from the WGS84 coordinate system to a cartesian coordinate system. The first coordinates after starting to record is set as origo in the cartesian coordinate system. For subsequent location updates, the angle and the distance from the initial location to the current location are calculated. From this, the coordinates in the cartesian system can be calculated. With the location in cartesian coordinates together with the speed and heading measurement, the Kalman filter can be supplied with the required sensor measurements.

To display the path taken on the map, the coordinates must to be expressed in latitude and longitude points according to the WGS84 coordinate system. The moving average and Kalman filter data therefore are converted back from the cartesian coordinate system. This is done using Vincenty’s direct formula [28] using the distance and direction to the coordinate from the start location.

The layout of the application consists of three buttons for starting and stopping the recording of measurements, and to reset the measurements. The distance measurements for the raw data, the Kalman filter and the moving average is displayed in the upper left corner overlaid on top of the map where the different position estimates and paths are displayed. A typical view during measurement is illustrated in Figure 3.

![Figure 3: A typical measurement when using the developed application.](image)
4.2.1 Tuning the Kalman filter

In the Kalman filter equations 3.7 and 3.8 the values of the two matrices $Q_t$ and $R_t$ depend on the specific sensors and the environment at runtime, and needs to be specified. One way to specify $R_t$ is to use the error margin of the equipment indicated by the manufacturer. It is likely that this error margin have been estimated under good conditions. Another way to estimate the error margin is to place the device at a known location and measure deviations over a period of time, from which the variance can be calculated. However, this variance estimate is only accurate if the same conditions apply during actual use. As the aim of this study is to test the GPS positioning during bad and changing conditions, neither of these variance estimates are suitable to be applied. The Android Location API includes a method [29] to retrieve the current accuracy of a location in meters. This value corresponds to the standard deviation of the location estimate. The square of this value was supplied as the variance of the position measurement in the $R_t$ matrix. A conservative estimate of the speed variance was used, which was set at 9. For the process covariance matrix, $Q_t$, only the acceleration variance needs to be supplied. A low value for the acceleration variance indicates that the model is a good estimator for the true process, whereas a high value will relax the process model on the Kalman filter estimate. The Kalman filter will then rely more on the sensor measurements. Leading up to the final measurements, several test-measurements were conducted where various settings for the acceleration variance was used. A suitable value of the acceleration variance was selected to be $0.25^2$.

4.3 Statistical methods

For this study, the ideal data to perform statistical analysis on to determine whether either filtration method or the raw data provides better GPS positioning, would be position estimates in tandem with the true position at each measurement along the path. This information is however not available in a real-world field study. Performing comparison tests solely on the difference between individual data points of raw and filtered data will not yield meaningful results, as only the difference between them can be distinguished, not whether either of them are better or closer to the truth. An approximation to the question whether better positioning can be provided at each point could be whether the distance measured along the path is better for either filtration method or the raw data. This measurement can be compared to the actual truth, as the true distance can much easier be measured to a sufficiently accurate degree, using a measuring wheel. The distance measurements for the raw data, Kalman filter and the moving average filter is obtained by taking the sum of distances between successive data points.

The measurements are repeated several times on the same path to produce vectors of observations for statistical analysis. When performing pairwise comparisons between estimators, the observation of interest is the absolute error from the true distance. The absolute errors for the measurements of each estimator is calculated by taking the absolute value of the difference between the true distance measured by a measuring wheel and the estimated value. On these residuals, t-tests are used to determine the statistical significance of difference in means, and F-tests are employed to test for difference in variance. For two estimators with significant difference, the estimator with lower mean or variance is considered having a better performance. The significance level for all tests are $\alpha = 0.05$. All analysis are performed
in the statistical software package R [27].

### 4.4 The data set

The ten observations that were collected for each distance measurement are illustrated in Figure 4. Three major errors in the GPS distance estimations are recorded at observation 4, 8 and 10. One of these large deviations, observation 8, is shown in Figure 5.

![Distance measurements](image1)

**Figure 4:** Distance measurements for each observation

![Large GPS error](image2)

**Figure 5:** Large GPS error. Raw data in red, Kalman filter in green and moving average in blue.
5 Results

In this chapter, the data collected from measurements is described in detail and the results from statistical analysis is presented.

5.1 Descriptive statistics

The raw data estimator overestimated the true distance on average by roughly 30 meters, whereas both the Kalman filter and moving average underestimated the distance by 20 meters. Both filters had a variance reducing effect on the distance estimate, with the standard deviation of the raw data reduce from 86 meters to 12 meters and 19 meters for the Kalman filter and moving average respectively. Descriptive statistics for the distance measurements are shown in Table 1.

<table>
<thead>
<tr>
<th>Distance estimator</th>
<th>Mean (m)</th>
<th>Standard deviation (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>True distance</td>
<td>302.3</td>
<td>5.0</td>
</tr>
<tr>
<td>Raw data</td>
<td>332.7</td>
<td>85.6</td>
</tr>
<tr>
<td>Kalman filter</td>
<td>279.8</td>
<td>11.7</td>
</tr>
<tr>
<td>Moving average</td>
<td>278.5</td>
<td>19.0</td>
</tr>
</tbody>
</table>

The mean absolute error is lowest for the Kalman filter, but only slightly lower than the moving average. The standard deviations of the filter methods are very close in magnitude. Compared to the filter methods, the absolute error for the raw data is larger both in mean and standard deviation. Descriptive statistics for the absolute errors are presented in Table 2.

<table>
<thead>
<tr>
<th>Distance estimator</th>
<th>Mean (m)</th>
<th>Standard deviation (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw data</td>
<td>55.4</td>
<td>66.0</td>
</tr>
<tr>
<td>Kalman filter</td>
<td>22.5</td>
<td>12.6</td>
</tr>
<tr>
<td>Moving average</td>
<td>25.8</td>
<td>12.5</td>
</tr>
</tbody>
</table>
5.2 Comparison tests

The pairwise mean and variance tests for the absolute errors are described in Table 3. The mean absolute errors for the filter methods are very close in magnitude, and the t-test show that their means are not statistically different. The t-tests comparing the raw data estimator and the filter methods did not show any significant difference, as the variance of the raw data was very large. The F-tests show that the variance is significantly reduced by the filter methods compared to the raw data, and that the variances for the two filter methods are not significantly different.

Table 3 Comparison tests for absolute errors

<table>
<thead>
<tr>
<th>Test</th>
<th>t-test p-value</th>
<th>t-test confidence interval</th>
<th>F-test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw vs Kalman filter</td>
<td>0.15</td>
<td>(-14.7, 80.5)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Raw vs Moving average</td>
<td>0.19</td>
<td>(-18.0, 77.2)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Kalman filter vs Moving average</td>
<td>0.56</td>
<td>(-15.1, 8.5)</td>
<td>0.99</td>
</tr>
</tbody>
</table>
6 Conclusion and Discussion

The results show that during unfavorable conditions for GPS signal reception, the variance of the raw data as a distance estimator is high. Due to its high variance, the mean absolute error of the raw data estimate cannot be statistically differentiated from either one of the filter methods. It can however be concluded, that the variance for the raw data absolute error are significantly higher than that of both filter methods. This high variance makes the raw data unsuitable to be used for distance estimation under harsh GPS conditions, and indicates that even a simple filter method will significantly decrease the variance in the distance estimation.

The Kalman filter and the moving average filter are very close in their measured performance. The Kalman filter did have a lower sample standard deviation, but concerning their mean absolute errors the filters were roughly equal. As the two filter methods were not statistically different, the conclusion is that the Kalman filter in its present form does not significantly improve the distance estimation compared to other simple filter methods such as the moving average. The current motion model and sensor measurements used in the Kalman filter did not provide enough additional information in order to accurately predict the actual movement on the ground. The lack in performance of the motion model caused the method to rely too heavily on the GPS measurement and could therefore not further improve the distance estimate or reduce noise in the position estimate further than the moving average.

6.1 Regarding GPS reception conditions

An environment with consistently harsh GPS conditions was harder to find than expected. During testing, the behavior of the GPS position estimation was much different than when the measurements to perform analysis on were conducted. During the testing, the position estimates had a much higher variance. A GPS reception as seen in Figure 6 would have been much more unfavorable for the raw data and moving average filter estimators. When these deviations were mostly absent, the raw data had a much smoother and more accurate performance. It is also necessary to note, that there was a thin line between having suitably bad GPS reception and having no reception at all. During some of the test measurements, the GPS reception was completely lost for long stretches of time, as the example seen in Figure 7 shows. With more time, one way to go about finding more consistent conditions would be to estimate GDOP values throughout the day to find a time frame in which the conditions remain relatively harsh but also consistent.
6.2 Regarding the measurement path

The relatively few large GPS errors that occurred during the measurements can in part be explained by the GPS reception conditions, but also by the length of the selected path. The effect that bad GPS reception has on the distance estimation might have been more expressed if the distance of the measurement path had been longer. In the present case the distance estimated by the raw data were almost half the time above the true distance, and half the time below. Given a longer path, the bad reception would have had more time to effect position estimates and as such increased the estimation of the distance, giving the raw data an overall larger positive bias. This could have resulted in a statistically significant difference between the raw data and the true distance, and between the raw data and the Kalman filter. However, another factor that had to be considered when determining the length of the path, was the number of times the path had to be traversed. This study used ten measurements. Increasing this number would have allowed for a stronger basis for the statistical analysis, but it would also increase the total length traversed for all measurements. A higher number of observations lowers the limit for how long the path can be before it becomes unfeasible to perform all measurements within a reasonable time frame. In a larger and longer study, more time could have been spent on the measurements, both choosing a longer path and conducting a larger set of observations.

During the measurements, when it became known that the GPS conditions were relatively favorable, a more suitable path could have been chosen. The new path could have been longer and with more cover obscuring a larger part of the sky, reducing the GPS reception. However, doing so would not have been scientifically honest as it would have been selecting the data to fit the theory, instead of the other way around. From the current perspective, the argument can be made that given worse conditions with respect to GPS reception, it seems likely that the estimated distance using raw data would over-estimate the true distance. Fur-
thermore we can argue that sensors with better information from the ground ought to be able to improve the Kalman filter distance estimation, especially at times during bad GPS reception. A perfect motion model would in fact be able to keep an accurate position indefinitely without GPS input, so a better motion model will undoubtedly result in better performance. The moving average filter was here concluded to be a viable alternative filtration method, comparable with the Kalman filter. However, since the simple moving average filter is entirely dependent on the GPS data and does not include any model based estimation and correction, it will not be able to reduce any bias error in the position estimate, only reduce the noise. As in the present case, assuming the raw data bias is low enough, the moving average will perform reasonably well, but this assumption can of course never be expected always to be valid.

6.3 Ethical aspects

A device with enough personal information to uniquely identify an individual, together with a technology allowing a third party to continuously monitor the movement of said individual, presents a host of privacy related issues. A user consenting to give a third party access to sensitive information includes a set of problems the user might not always be aware of. First, the user must trust that the third party is honest, and will not store location information unless explicitly agreed, and that the information is completely removed if third party is asked to do so. But even if the intentions of the third party are completely honest, no database is 100 % secure, so stored information can be stolen. Storage of sensitive information such as this therefore sets a higher requirement on security, something that either lack of competence, time, or money might stop from being correctly implemented and maintained.
6.4 Societal aspects

The proliferation of GPS enables many industries to improve efficiency, but also introduces a vulnerability in that we become more and more reliant on the technology to keep up with competition and economic growth. Being too reliant on the availability and precision of GPS technology might mean moving beyond our ability to cope without it. This might not be too much of a risk or problem for a private consumer, but for a business, being lean and effective can be the key to survival in the current harsh economic environment. After all, GPS is a foreign technology and a space technology, its free availability might be subject to change during war time or if relations between countries change. The Russian equivalent to GPS, GLONASS, just recently made operational, and the European equivalent soon to be made operational, GALILEO, are attempts to reduce the dependency on the GPS technology. In the current world situation, there is nothing indicating that any of these systems might soon become unavailable or have their accuracy severely reduced, but the functionality and availability of these systems cannot be guaranteed indefinitely. So, while still holding an optimistic view regarding the future availability and precision of space-based navigational systems, maintaining some backward compatibility toward older navigational systems will be a sound strategy for business owners and private consumers alike.

6.5 Future work

The aim to evaluate the Kalman filter in a natural environment, where the measurement device could be used for other purposes simultaneously, introduced limitations on the extent to which certain sensors could be relied on. Initially the intent was to utilize the accelerometer for speed estimation and compass for direction of movement. However, the use of accelerometer for speed estimation had to be discarded due to drift. Simpler sensor inputs were used to ensure that the measurements could be conducted. For speed estimation, Doppler shift measurement from the GPS signal was used. This estimate can give accurate results even when the GPS signal is degraded, but cannot function in environments where the signal is completely blocked. Similarly, the heading was also estimated from the GPS data. The motion model and the sensory input used in this study needs to be further improved before the Kalman filter can be considered as a viable method to reduce noise and bias in GPS positioning, or to improve distance estimation. The largest improvement will most likely come from using better sensors from the ground level so the motion model could more accurately predict the position from the speed and direction. For speed estimation, the motion model could use a step detector where the stride length is dynamically adjusted by measuring the step period and acceleration along the direction of movement. The step detector must initially be calibrated, either by the height of the user being supplied, or by letting the user walk a pre-specified distance while the number of steps are counted. This would give an indication of the stride length under normal conditions. A required restriction for the step detector to work would be that the device must be placed in the pocket of the user, in order to detect the sharp change in acceleration as the foot hits the ground. This restriction might be unacceptable depending on the situation. The filter could however be made to adapt to different situations where the reliability of sensors change. For example if the device is being picked up from the pocket and unlocked, the motion model instead switches to the speed estimate from the GPS, until the phone is locked and put back. To determine the direction of movement is trickier, without restricting the orientation of the
phone. In a private consumer application meant to be used in every-day life, the orientation of the device and its compass heading might not correspond to the actual direction of movement. It may change in unpredictable ways at any given time if the device is being actively used. One way to improve the heading estimated by the GPS is to use the changes measured by a compass as control input between subsequent GPS location updates.

In time it is likely that the GPS position estimates will become more accurate as error modelling techniques improves and satellites with more advanced equipment are launched into orbit. However, regardless of the accuracy of GPS, as long as the receiver is dependent on line of sight visibility to the satellites, certain environments will degrade or block the signal to such a degree that the position estimate becomes unusable. In situations like this, in urban areas, forest areas and especially indoors, the Kalman filter will be a useful tool to maintaining a good position estimation.
Bibliography


