Statistical approach to GPS positioning of mobile robot

Ladislav Jurišica, Anton Vitko, František Duchoň, Dušan Kaštan

Faculty of Electrical Engineering and Information Technology, Institute of Control and Industrial Informatics, Ilkovičova 3, Bratislava, ladislav.jurisica@stuba.sk, anton.vitko@stuba.sk, frantisek.duchon@stuba.sk, dusan.kastan@gmail.com

Abstract: This article analyzes the present state of GPS system, possibilities and constraints of the mobile robot localization under environment influences. Real measured data from GPS receiver are statistically evaluated.

Keywords: GPS, Kalman filter, moving average

1. INTRODUCTION

Navigation of a mobile robot in an unknown and uncertain environment requires knowledge about the robot’s current position. A natural approach to manage the task is using GPS (Global Positioning System) [1]. Because the GPS’s signals are commonly corrupted by various disturbances like signal bounces, inaccurate receivers etc. the disturbances should be adequately mitigated at the highest possible rate. The most effective means of doing this seems to be statistical ones.

2. THE GPS SYSTEM

GPS serves for position of an object on Earth independently of actual meteorological conditions. Position of a measured point is given by intersection point of spherical surface of the radius given by distances between satellites and measured point. From geometric point of view, for determination of the position it is needed to know positions of three satellites minimally. Because the distance between satellite and measured point must be defined at the same time, for exact position identification of measured point it is needed to know positions of four satellites. To achieve the high accuracy of position determination, it is important to use the maximum possible number of visible satellites that should be properly distributed on the sphere. Applications based on the GPS technology are almost unlimited. They cover virtually all the areas of humane activities. Application area of GPS is continually widening. For instance, pilots can utilize GPS for searching airports, marines for searching ports, tourists can orient themselves in an unknown country, fish men can find out suitable time for fishing, land surveyors can identify position of a point with high accuracy and mobile robots can localize themselves in the environment. [2].

The GPS in comparison with conventional measuring methods has several advantages:

- among particular measuring points may not be immediate visibility
- it has high accuracy
- the measurement is quick and provides results in the united world coordinate system WGS 84
- it provides three-dimensional position data
- it works regardless of weather, daily or night time [4].

But it suffers from several disadvantages:

- it can’t be used for measuring in the underground
- it provides worse results in dense vegetation (forest etc.)
- it is required for satellite to be immediately visible and the sky should be visible from measurement point within 15° above the horizon and more in all other directions
- problematic measurement in heavy inhabited areas (town with narrow streets)
- problematic measurement in valleys

2.1 The system navstar gps

The system consists of three segments [1]:

1. Cosmic segment

Nominal constellation of GPS consists of 24 satellites deployed above the ground at the height of 22 200 km. It is required minimal number of 24 operating satellites. Satellites are deployed on 6 orbits (4 satellites on orbit minimally), that are equally spaced (60 degrees from each other) and inclined 55 degrees with regard to equatorial plane. This constellation ensures that from an optional point on Earth a user is visible from 5 to 8 satellites. Orbit time of one satellite from GPS system is 12 h of star time what means that satellites have identical configuration during 11 h 58 min of solar day.

2. The control segment

Tasks of the control segment of GPS are: non-stop monitoring and control of satellite system, definition of the system time of GPS, prediction satellites’ paths and clock operation on satellites and regular regeneration of navigation report of each satellite [3]. The segment consists of five monitoring stations (Hawaii, Kwajalein, Ascension Island, Diego Garcia, Colorado Springs), four ground antennas
Ground monitoring stations accept signals of all visible satellites. Data are sent into Master Control Station where orbital elements of satellites are defined, the correction of atomic clocks is performed and navigation report is arranged [3]. Navigation report is then sent on particular satellites by ground transmitting antennas. Satellites then send back their orbital elements and exact time to the Earth. Transmitting antennas are deployed in such a way that the connection with each satellite is possible at least three times per day. [2].

3. User segment
This segment is composed of the GPS receivers. By way of utilization they can be divided into navigational (ground, naval, aerial and others navigation), geodetic (single-frequency and two-frequency devices, RTK systems etc.) and receivers for time synchronization (astronomic measurements and telecommunications).

2.2 GPS systems with higher accuracy

1. Inertial Navigation System (INS)
By addition of the INS device to the GPS receiver it is possible to achieve accuracy of position determination up to one meter. In this configuration the GPS provides a short-term accuracy as far as the INS provides long-lasting stability. Outputs of both systems are compared and properly filtered, whereby corrections on both systems are executed. Corrections are done by using Kalman filter, which combines two estimates and provides the most probable estimation.

2. Differential GPS (DGPS)
By using DGPS the accuracy of position determination can be improved up to one meter. DGPS includes second receiver on a fixed position. The second receiver is set to calculate correction for GPS data. Many free services based on DGPS corrections exist, though paid services have better accuracy.

3. Wide Area Augmentation System (WAAS)
WAAS is extremely precise navigation system developed for civilian navigation. System WAAS allows horizontal and vertical navigation for exact operations (e.g. aircraft landing). WAAS is available in USA and in some Pacific areas. In Europe it is known as the system EGNOS [10]. It improves GPS accuracy of position determination up to three meters. From the above mentioned systems the WAAS is most suitable system for mobile navigation. DGPS isn't accessible everywhere and on the top of this it is a paid service. INS sensors and their implementations are financially demanding. WAAS technology is in the newest models of GPS receivers integrated as a standard. It depends only on the user whether he utilizes this possibility. Ideal solution is usage of GPS with WAAS, whereby position data are further improved with Kalman filter and combined with data about local robot position, e.g. information from incremental sensors or digital compass [6].

2.3 GPS standards
GPS receivers and their large proliferation represent enormous source of data. But there can be problem with data transmission from GPS receiver to the program of measurement evaluation, which was manufactured by a different producer. Receivers often use their own (often undocumented) formats. For this reason the communication standards with GPS receivers such as NMEA 0183, RTCM SC-104 or RINEX were created.

2.4 GPS errors

1. Selective Availability (SA) error
SA error is the main factor, which influences accuracy of GPS. It is an intentionally transmitted error with the aim to decrease accuracy of civilian GPS receivers. An objective is to make utilization of GPS system for enemy armies and terrorist organizations impossible. [1]. This error can lower accuracy up to cca. 100 m. The calculated falling-off GPS accuracy for civilian sector ended 2nd May 2000. At present the GPS receivers from civilian sector reach accuracy around five meters.

2. Satellite geometry
Satellite geometry specifies satellites positions relatively to view of GPS receiver. If GPS receiver sees four satellites situated to the north and west it is possible that it will not be able to identify position, because all measurements are incoming from the same direction. Triangulation is becoming imprecise and resulting plane on which the measurements intersect is very large. Following this error, the variance of a real position can be from 100 to 150 m. If each of assuming visible satellites is on other world's side (90° angle between satellites) the measurement accuracy is advanced because measurements come from different world's sides and the resultant plane is much smaller. Influence of satellite geometry make the measurement worse for cars near high buildings or in uneven terrains. High-end GPS receiver then does not show only the satellites from which it is able to accept signals, but also their positions on sky (azimuth and elevation).

Quality of the satellite geometry can be evaluated by the parameter “Dilution of Precision – DOP”, which is an explicit indicator of position determination quality. Calculation of DOP rests in determination of the relative position of each visible satellite with respect to other visible satellites. Smaller value of DOP corresponds to higher accuracy. There are several types of DOP. They indicate influences of various parameters on accuracy. To the family of DOPs belong the following ones: the relative (RDOP, relative position error), the positional (PDOP, horizontal and vertical measurement), the horizontal (HDOP, horizontal measurement), the vertical (VDOP, height measuring) and the time DOP (TDOP, time shift).
3. Atmospheric effects

Earth atmosphere is a two-part environment (troposphere and ionosphere) [2][10] with essentially different effects on the broadcasted high-frequency signals. The troposphere is a neutral part without electrically charged elements. For waves of frequency up to 15 GHz the troposphere is a non-dispersive environment. Both types of GPS measurements (phased and encoded) are equally affected by tropospheric refraction, the size of which depends on meteorological parameters of atmosphere (especially temperature, pressure and humidity). The ionosphere is characterized by high capacity of free electrons and ions. Accordingly, it is electrically active. Properties of ionosphere are changing in time, sun activity and changes of Earth magnetic field. As a result, the reflective properties of particular layers are changing. For radio waves the ionosphere has character of non-dispersive environment. In ionized layers of the ionosphere the radio waves are not only bounced back, but they are also absorbed. Besides the electrons and ions, these layers contain also electrically neutral molecules that do not oscillate. However, electrons and ions vibrating by action of electromagnetic field bump on these molecules. This causes the energy losses exhibited as radio wave dumping. Losses are the greater the more no-ionized molecules are in the layer, hence the layer is closer to Earth’s surface [2].

4. Multipath effect

The multipath effect of signal propagation means that the transmitted radio signal is reflected from some objects. GPS satellite does not transmit signals just in the direction toward a certain receiver but into the wide cone [2]. Provided that the reflective plane exists in the surrounding of GPS receiver the signal may approach receiver's antenna also indirectly. The path that signal is required to pass becomes longer and the same goes for the time needed. Due to this the GPS receiver measures longer distance to satellite. As a result, the overall error at position determination is approximately 5 m greater. [7]. There are many ways to reduce influence of multipath effect. The one of the most used consists in a construction modification of GPS antenna, technological improvement of receiver and modification of received signal processing. From among the construction modifications the protective plate below antenna is mainly used, which prevents receiving the signals reflected from Earth or water level and other reflective planes.

5. Relativistic effects

Relativistic effects manifest themselves especially at high speed of surveyed objects and in presence of non-homogeneous gravitation field. Another important factor is accuracy with which it is required to sense given events. In essence those are reasons due to which the relativistic effects are needed to be included into operative equations of GPS measurements. These are mainly the non neglect-able changes of frequency due to quickly moving GPS satellites and changes of signal propagation from quickly moving satellites w.r.t. the rotating Earth as well as their great distance from the Earth. There is also need to include the relativistic disturbance effects on the satellites paths, which are caused by non-homogeneous gravity field of the Earth, as follows from general theory of relativity.

6. Other GPS errors

From among the additional measurement errors one can mention the clock inaccuracy of the GPS receiver (receiver can't contain atomic clock and the error removal is solved by using signals from several satellites), inaccurate determination of parameters of the satellite path (so-called ephemeris error) and number of visible satellites (for higher accuracy of measurement is needed to have higher number of visible satellites).

3. Kalman filter

The term “measurement” is historically related to determination of angles, lengths and elevations by technical means. As early as at the beginning of 19th century German mathematician C. F. Gauss developed method [11] known as the Least Squares Method (LSM). It allows for parameter estimation by minimization of the sum of squares of the differences between individual realizations of the measurement.

At present, thanks to technological progress, we commonly process an enormous quantity of data. Primarily with the arrival and development global satellite systems for navigation and position identification (GNS), the interval of repeated measurements of three-dimensional position has shorten from days to hours, from minutes to seconds or even fragments of the second. Such quantity of data is no longer possible effectively evaluate using LSM, because the LMS necessitates processing large matrices. The problem is solved with recursion filters, which use a part of their outputs as input for next calculation. One of them is Kalman filter [9]. Kalman filter is able to process dynamic data with minimal delay even if the fast dynamic changes occur. The result of Kalman filtering is the system state estimated from measurements corrupted with errors. Let a discrete process of subsequent mobile robot positions is described by the state equation:

\[ x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \]  \hspace{1cm} (1)

with measurements:

\[ z_k = f(x_k) = Hx_k + v_k \]  \hspace{1cm} (2)

Random variables \( w \) and \( v \) represent process noise and measurement noise respectively. It is supposed that these variables are independent of each other and have normal Gauss distributions with covariance matrices \( Q \) and \( R \). The matrix \( A \) represents relations between the system state at time steps \( k \) and \( k-1 \) without influence of control function \( u \) and the signal noise. The matrix \( B \) is a control matrix, which joins the known control vector \( u \) with the system state \( x \). Matrix \( H \) represents relations between the
system state \( \mathbf{x} \) and the measurement \( \mathbf{z} \).

Algorithm of Kalman filter is recursive, thus all previous input data are included in the last estimation and also create (except new measurement) the input into the new cycle. In comparison with LSM method there is not need to calculate inversions of large matrices.

Principle of discrete Kalman filter (Fig. 1) consists of two basic steps: prediction and correction (sometimes called update or actualization) [8].

In the prediction step the system state and its covariance matrix at time step \( k \) on basis of system state estimation at time \( k-1 \) is predicted in accordance with eq. (3)

\[
\hat{\mathbf{x}}_k = \mathbf{A}.\hat{\mathbf{x}}_{k-1} + \mathbf{B}.\mathbf{u}_{k-1} \\
\mathbf{P}_k = \mathbf{A}.\mathbf{P}_{k-1} \mathbf{A}^T + \mathbf{Q}
\]

where \( \hat{\mathbf{x}}_k \) stands for the state estimation (position of mobile robot) at time step \( k \). Matrices \( \mathbf{P}_k \) and \( \mathbf{P}_{k-1} \) can be expressed:

\[
\mathbf{P}_k = \text{cov} (\mathbf{x}_k - \hat{\mathbf{x}}_k) \\
\mathbf{P}_{k-1} = \text{cov} (\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1})
\]

Correction step represents improvement of the state estimation based on actual measurements and it is defined as:

\[
\mathbf{K}_k = \mathbf{P}_k^{-1}.\mathbf{H}.(\mathbf{I} - \mathbf{K}_k.\mathbf{H}).\mathbf{P}_k^{-1} \\
\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k + \mathbf{K}_k.(\mathbf{z}_k - \mathbf{H}.\hat{\mathbf{x}}_k) \\
\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k.\mathbf{H}).\mathbf{P}_k
\]

where \( \mathbf{K}_k \) is so-called Kalman gain. It expresses a weight of actual measurement concerning estimated variable. By a simple analysis it becomes clear that with more precise measurement (i.e. with decreasing covariance matrix of measurement noise - \( \mathbf{R} \)) its weight rises:

\[
\lim_{\mathbf{R} \to 0} \mathbf{K}_k = \mathbf{H}^{-1}
\]

On the contrary, if covariance matrix \( \mathbf{P}_k \) approaches zero the weight of actual measurement decreases [4]:

\[
\lim_{\mathbf{R} \to 0} \mathbf{K}_k = \mathbf{0}
\]

The part \( (\mathbf{z}_k - \mathbf{H}.\hat{\mathbf{x}}_k) \) in (5), is often defined as a-priori residuum \( \mathbf{e} \) and represents difference between real value of actual measurement and expected measurement, assigned from last estimation of the state (thus position of mobile robot).

The described procedure requires an initialization. It is a process, which at the beginning of calculation defines values of basic parameters. Vector \( \mathbf{x}_0 \) is designed from initial measurement or it has zero initial value. Covariance matrix \( \mathbf{P}_0 \) at initialization is mostly defined as a diagonal matrix with sufficiently large members, which will have almost no weight in next iteration [9].

4. MOVING AVERAGE

Moving average [12] is in technical applications frequently used because of its simplicity and possibility to combine various moving averages together. Moving averages smoothes data and simplify identification of trends. Many types of moving averages exist but in technical analysis the simple and exponential moving averages are mostly used.

The simple moving average is established by calculation of average value in specific number of periods \( n \):

\[
\text{SMA} = \frac{P_1 + \ldots + P_n}{n}
\]

where \( P_1, \ldots, P_n \) generally stand for measured values (in our case positions of mobile robot). The exponential moving average eliminates delays, which appear within simple moving average. With usage of exponential moving average the delay is reduced by application of a bigger weight \( (\mathbf{K}) \) on recent values relatively to older ones:

\[
\text{EMA}_n = (\mathbf{P}.\mathbf{K}) + \left[ (1 - \mathbf{K}).\text{EMA}_{n-1} \right]
\]

Selection of small calculation period causes that moving averages are more sensitive and generate more signals. Longer calculation period increase reliability but the calculated value may exhibit non-permissible delay (undesirable at estimation of position of mobile robot).

5. EXPERIMENTAL RESULTS

Statistical methods were tested by a cheap I-Tec Bluetooth GPS receiver utilizing encoded way of measuring. The receiver works with recording period from 0.1s and uses standard NMEA-0183 at 57 600 b/s. Measurements were carried out by method of the absolute position determination of the area smaller than 30 km². Therefore, measurements were equally loaded by global errors. The only one measuring equipment was used, consisting of sensor I-Tec
GPS and software EvEgps developed for data acquisition from GPS device with standard NMEA 0183.

5.1 Stationary point measurement

The stationary point measurement identifies position of stationary point and was realized without external antenna. A measurement file contains 90 measurements. Measurement is corrupted with errors caused by multipath effect and smaller number of visible satellites. This follows from the selection of placement of stationary point in urban area and also from partial visibility of the sky.

The measurement data were processed by statistical methods. In case of moving average the set up time period was equal to the time of measurement (Fig. 2 b). With Kalman filter the outputs were almost identical with input (Fig. 2 a). As far as the need of repeatability is concerned, after application of moving average with period equal to the time of measurement, it results into different values of position. In the case of using Kalman filter the results are close to the measured values.

5.2 Geometric shape measurement

The second measurement consisted in the measurement of a circular path and it was also realized without application of external antenna. Measuring data file contained 63 points.

By comparison of the measured and real path can be concluded that measurements were influenced by some errors (Fig. 3). The space where measurements were realized was localized in a dense built-up urban area. Errors were caused by insufficient number of visible satellites (during entire measurements the average number of visible satellites were equal to 3), the multipath effect of signal propagation and satellites’ geometry.

From observations of the graph comparing the real and measured path are follows that measurements were influenced with constant (global) error. After application of statistical methods the similar outputs were obtained (Fig. 4) and uncertainties caused by variable (local) errors were eliminated. Global and also local errors of measurement seem to be constant, hence, to discriminate local from global errors it was realized next measurement.
b.) Application of moving average (red colour) on file of measured data (blue colour)

5.3 Measurement of road following a cusped line

A road of the shape of a cusped line was measured without application of external antenna. Measured data file consisted of 200 points.

By comparison of the real and measured path (Fig. 5) it can be pointed out that the measurement was influenced by errors. Measurement was realized in a dense urban area. In the area situated between high-rise buildings were visible at most four satellites. The visibility became better after leaving this area (white dot in Fig. 5) and thereby number of visible satellites increased to seven. This fact becomes more evident when position determination is improved. In case that greater number of satellites was visible the real and measured path was almost equal. On the top of this, no multipath effect was noticed, because there were much less reflexive planes.

In this case the measurement was influenced by errors just in a certain period. These errors can be considered as non-constant and can be removed by application of statistical methods (Fig. 6).

5.4 Road way measurement

The measurement consisted in passing a long road way by car without application of external antenna. It was measured 3010 points.

In this case, measured path was almost equal to the real one (Fig. 7). During measurement there were seven visible satellites on average. With a very long path and sufficient number of visible satellites properly distributed on the sphere the deviations from real path were minimal. Applications of
Statistical methods lead to minimal improvements only (Fig. 8).

Fig. 7. Measured path of road way (red colour)

Fig. 8. Road way measurement
a.) Application of Kalman filter (red colour) on file of measured data (blue colour)
b.) Application of moving average (red colour) on file of measured data (blue colour)

6. CONCLUSIONS

Statistical methods were applied on GPS data in order to improve precision of a commonly available GPS receiver. In the realm of mobile robotics they are mainly used for robot localization in the environment. Both applied methods (Kalman filter and moving average) are characterized by a small delay.

These methods filtered errors caused by non-constant local errors (insufficient number of satellites, multipath effect of signal propagation and satellites geometry). If measurements were influenced also with global errors (atmospheric and relativistic effects) then measurements were corrupted by constant errors. To eliminate global errors it is necessary to use DGPS or WAAS system or employ additional sensors. These systems can eliminate global errors; hence in comparison with standard GPS receivers they are more useful in robot localization in external environments.

7. ACKNOWLEDGEMENTS

This work was supported by MS SR under the contract VEGA 1/0690/09.

REFERENCES


