Journal of Global Positioning Systems (2011)
Vol.10, No.1 :11-18
DOI: 10.5081/jgps.10.1.11

Journal of Global Positioning Systems
2011
© CPGPS

Visual-aided Two-dimensional Pedestrian Indoor Navigation with a Smartphone

Laura Ruotsalainen, Heidi Kuusniemi, Ruizhi Chen
Department of Navigation and Positioning, Finnish Geodetic Institute (FGI), Finland
Email: firstname.lastname@fgi.fi, fax +358-9-29555200

Abstract

Indoor pedestrian positioning sets severe challenges for a navigation system. To be applicable for pedestrian navigation the platform used has to be small in size and reasonably priced. Smartphones fulfill these requirements satisfyingly. GNSS signals are degraded indoors and in order to obtain accurate navigation aiding from other sensors is needed. Self-contained sensors provide valuable information about the motion of the pedestrian and when integrated with GNSS measurements a position solution is typically obtainable indoors. The accuracy is however decreased due to errors in the measurements of the self-contained sensors introduced by various environmental disturbances. When the effect of the disturbance is constrained using visual-aiding the accuracy can be increased to an acceptable level. This paper introduces a visual-aided two-dimensional indoor pedestrian navigation system integrating measurements from GNSS, Bluetooth, WLAN, self-contained sensors, and heading change information obtained from consecutive images. The integration is performed with an Extended Kalman filter. Reliability information of the heading change measurements calculated from images using vanishing points is provided to the filter and utilized in the integration. The visual-aiding algorithm is computationally lightweight taking into account the restricted resources of the smartphone. In the conducted experiment, the accuracy of the position solution is increased by 1.2 meters due to the visual-aiding.

Keywords: pedestrian, navigation, visual-aided, smart-phone, indoor

---

1. Introduction

Pedestrian indoor navigation sets challenges for the positioning equipment. The system has to be accurate, small enough to be carried by a human, and reasonably priced. Outdoors, GNSS receivers fulfill all demands set for a pedestrian navigation system, but indoors the accuracy decreases substantially or the position information becomes impossible to obtain. There is not a single comprehensive sensor for indoor navigation such as GNSS outdoors, and therefore measurements from different sensors have to be integrated.

Wireless radio sensors, like Bluetooth or Wireless Local Area Network (WLAN) are often used for indoor positioning. Their drawback is however the limited availability of the location information provided as well as the need for a pre-installed infrastructure. Self-contained sensors measure the motion of the pedestrian and provide a relative position; the attitude and distance traveled (Collin, 2006), (Retscher, 2007). The attitude may be measured with a gyroscope or a digital compass and the distance with an accelerometer. The gyroscope suffers, however, from cumulative drift errors (Saarinen, 2009), and the electric devices in indoor environments, like elevators and printers, cause the measurements from a digital compass to be erroneous. A camera is independent from other sensors and the noise in images is not cumulating over time. The camera is also free from an infrastructure installation. The motion of the camera may be calculated from consecutive images. When the camera is carried by a pedestrian, the motion of the camera relates to the motion of the pedestrian. Integrating the motion information obtained from the images derives a navigation system with increased availability and accuracy.

Visual-aiding has been used already for decades in navigation of robots and unmanned vehicles (Corke et al. 2007). In pedestrian navigation, the focus of the research has been mainly in systems using a priori formed databases. The databases contain images of recognizable features in the surroundings attached with position information (Aoki et al., 1999), (Robertson and Cipolla, 2004), (Zhang and Kosecka, 2006), (Steinhoff et al., 2007). When a match between images in the database and the ones taken by a pedestrian is found, the absolute position may be obtained. The database based procedure is though laborious due to the a priori preparations and is restricted to the predefined region.

Integrating visual-aiding information with the measurements from other sensors has become a research topic only in the latest years. Kourogi and Kurata (2003) corrected the position measurement obtained with self-contained sensors using image matching with a database. A procedure similar to the one discussed in this paper was introduced by Diel et al. (2005) and Kessler et al. (2010). They calculated motion information from consecutive images and used it to visual-aid the navigation system.

Integrating the measurements from different sensors increases the accuracy, availability, continuity, and integrity of a pedestrian navigation system. This can't though be done with the cost of the system usability, meaning the compactness, lightweight and price of the equipment. A smartphone is an appropriate platform for the system. It is reasonably priced and lightweight. Almost all smartphones contain most of the sensors needed; camera, wireless radio, GNSS, digital compass, accelerometer, some also a gyroscope. The restricted processing power and storage space set challenge to a navigation system running on a smartphone.

The navigation environment sets constraints for the visual-aiding. The methods based on calculating the motion of the camera from consecutive images are based on finding features from images that can be followed in the subsequent ones. The indoor surroundings where navigation is mostly needed, like offices and public buildings, are poor with features. The lighting is also often constricted. Fortunately these indoor environments contain many straight lines, like borders of floors and walls that are also in some extent invariant for changes in lighting. The lines may be used for calculating vanishing points, features arising from the projections of three-dimensional objects in the field of vision of the camera into two-dimensional image points. These features may be used to define the change in the orientation of the camera relating to the orientation of the pedestrian.

The algorithms for vanishing point calculations presented in previous research are computationally heavy for a smartphone solution. This paper introduces an indoor pedestrian navigation system using a rapid algorithm utilizing visual-aided heading of the motion. The algorithm calculates the change in the heading using vanishing points. The calculations are performed with a frequency of 1 Hz which fulfills the real-time requirements set for navigation. The system integrates the visual-aiding information with measurements obtained from other sensors using an Extended Kalman filter. The paper presents an evaluation of the accuracy and reliability of the visual-aiding - information needed to construct a favorable integration filter.

2. Resolving the Heading of Motion with Computer Vision Based Methods

An image is a perspective projection of three-dimensional objects in the field of vision of the camera into two-dimensional figures. The projection loses valuable information like the depth of the scene. Though straight lines stay straight in projections, the parallel ones don't stay parallel but seem to intersect in a point. This point is called the vanishing point.

The change of the orientation of the camera is obtained by calculating the change of the vanishing point coordinates in consecutive images. When the rotations in pitch and roll are restricted, the orientation change is the change in the heading. The following describes the definition of the vanishing points and various computer vision based methods for reducing noise in the images, finding the parallel lines and the vanishing point as well as calculating the heading change.

2.1 Vanishing points

Vanishing point is a point in the image where all parallel lines seem to intersect. The projection of a line $\mathbf{l}$ to the image plane is shown in Fig. 1. A point $\mathbf{p}_1$ on the line is projected to the plane as $\mathbf{p'}_1$. The projection may be expressed using the camera matrix $\mathbf{M}_c$ with $\mathbf{p'}_1=\mathbf{M}_c\mathbf{p}_1$ (Kessler et al., 2010). The camera matrix $\mathbf{M}=\mathbf{KR}$ consists of the camera calibration matrix $\mathbf{K}$ and orientation of the camera $\mathbf{R}$. Similarly the line $\mathbf{l}$ with direction $\mathbf{d}$ is projected as $\mathbf{l'}$. All lines passing through the camera center with direction $\mathbf{d}$ and parallel to $\mathbf{l}$ intersect a plane at infinity in the infinite point $\mathbf{p}_\infty$. The vanishing point $\mathbf{v}$ is the projection of this point to the image plane. This reconstruction enables resolving the change in the direction of the lines in consecutive images and that way in the orientation of the camera $\mathbf{R}$. Let $\mathbf{v}_1$, $\mathbf{v}_2$ be the vanishing points in the first and second consecutive images and $\mathbf{d}_1$, $\mathbf{d}_2$ their directions respectively. By defining a normalizing factor $\|\mathbf{K}^{-1}\mathbf{v}_i\|$, the directions $\mathbf{d}_i$ may be calculated with (1) using the coordinates of the vanishing points (Hartley and Zisserman, 2003)

$$\mathbf{d}_i = \mathbf{K}^{-1}\mathbf{v}_i / \left\|\mathbf{K}^{-1}\mathbf{v}_i\right\| . \tag{1}$$

The calibration matrix $\mathbf{K}$ contains information of the intrinsic camera parameters, the focal length (f_x, f_y), principal point (u,v), aspect ratio and skew (S). In order to make the motion calculations easier, the aspect ratio is usually assumed to be 1 and skew 0 in computer vision algorithms. With these assumptions, the calibration matrix (2) may be calculated even from a single image using vanishing points (Kosecka and Zhang, 2002).

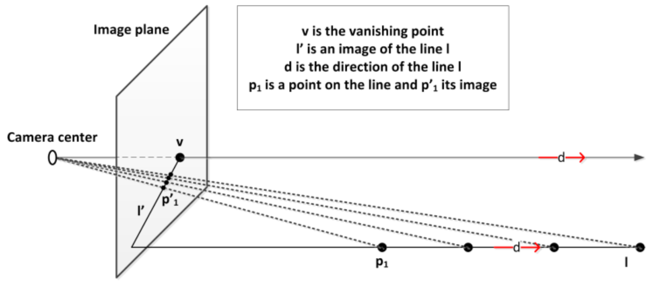

Figure 1: The vanishing point v of line l shown on the image plane. Figure is based on (Kessler et al. 2010).

$$\mathbf{K} = \begin{bmatrix} f_x & S & u \\ 0 & f_y & v \\ 0 & 0 & 1 \end{bmatrix}. \qquad (2)$$

2.2 Retrieving the lines from an image

Parallel lines in the image must be retrieved for the calculation of their intersection point, the vanishing point. This is done by first looking for the edges of objects in the image and then identifying the straight lines among them.

Images contain noise coming from different stages of the imaging process. A considerable source is the varying lighting conditions, especially indoors. The presence of noise disturbs the edge detection and causes errors for the calculations. Thorough pre-processing of the images removes noise and makes the calculations more stable. A pixel with an intensity value diverging notably from its neighbours is suffering from noise with a high probability. The noise may be removed by substituting the intensity of the pixel with a weighted sum of its neighbours'. The process is called convolution (Forsyth and Ponce, 2003). The image $\mathbf{F}_{u,v}$ is convoluted with a kernel $\mathbf{H}_{i-u,j-v}$ replacing the value of a pixel (u,v) with a weighted sum of values of neighbour pixels resulting in image $\mathbf{R}_{i,j}$ as defined with (3)

$$\mathbf{R}_{i,j} = \sum_{u,v} \mathbf{H}_{i-u,j-v} \mathbf{F}_{u,v} . \qquad (3)$$

A Gaussian kernel $\mathbf{H}_{i,j}$ reduces the noise effectively by emphasizing the weights of the nearest neighbours of the pixel when a large standard deviation value is used. The procedure causes some blurring but that doesn't disturb the edge detection notably. The Gaussian kernel may be presented with (4)

$$\mathbf{H}_{i,j} = \frac{1}{2\Pi\sigma^2} \exp(-\frac{((i-k-1)^2 + (j-k-1)^2)}{2\sigma^2} \qquad (4)$$

The size of the neighbour area is $2k+1x2k+1$.

Edges of objects may be found by tracing fast changes in the brightness values. Canny edge detector is one of the most used edge detectors (Canny, 1986). In vanishing point calculations only lines are considerable features and other edges should be left out from the calculations. Hough transform is a robust method for retrieving the lines from the set of edges (Hough, 1962).

2.3 Resolving the heading change

The method used in this paper calculates the vanishing points by voting for the intersection points of all lines and correcting the effect of noise with robust estimation using weighted means (Jepson and Fleet, 2009). When the rotations in the pitch and roll directions are restrained, the change of the vanishing point coordinates in z-direction is defined with (5) (Gallagher, 2005)

$$\mathbf{v}_z = \begin{bmatrix} f_x \sin\theta \\ 0 \\ \cos\theta \end{bmatrix}. \qquad (5)$$

The component of focal length, f_x, is obtained from the camera calibration matrix. The z-direction is the direction of the motion along the z-axis of the world coordinate frame presented in Fig. 2. By calculating the coordinates of the vanishing point in consecutive images, the change in the heading, θ, related to the rotation of the camera along the world XZ-plane, is obtained.

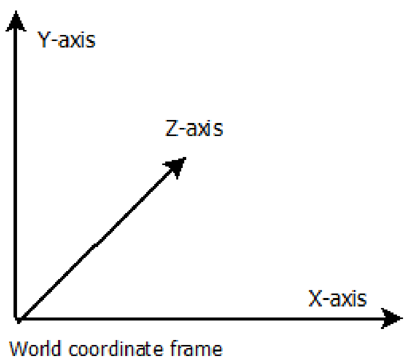

Figure 2: World coordinate frame with z-direction being the direction of motion

3. Integrating Visual Information with Other Measurements

The position information obtained using calculations from consecutive images is relative. The relativity means that only the change in distance and attention may be evaluated. The information has to be integrated with measurements from other sensors to get an operative navigation system. As discussed in the introduction part, there is no one comprehensive sensor to be used indoors. The system presented in this paper consists of different self-contained and radio sensors, represented more closely in Section IV. Measurements obtained with different sensors must be integrated with a filter. In the system discussed the filter is an Extended Kalman Filter. The filter needs estimates of the reliability of the measurements entered as inputs. The evaluation of the accuracy of the heading change obtained with visual-aided methods presented above is explained in the following.

3.1 Evaluation of errors in visual-aiding information

The information obtained from images with visual-aiding methods is noisy, especially in environments with varying lighting conditions such as indoors. The noise may be reduced using an appropriate kernel for filtering the image as explained in the previous section. The appearance of unessential lines due to e.g. shadows in the image caused by the sunlight coming from the windows or texture with lines that are not parallel, introduce errors into the vanishing point calculations that aren't easy to correct for. The lack of sufficient light diminishes the number of appropriate lines found for the calculations. In order to get an accurate position solution, the reliability of the measured vanishing points has to be evaluated after each calculation and the confidence for its correct location taken into account in the integration procedure. The heading change computation using the vanishing points with the previously presented method was evaluated to give an accuracy of 1.5 degrees with a stationary camera and for some extent stable lighting conditions; more details are in (Ruotsalainen et al., 2011). The situation in the navigation system presented in this paper is, however, more difficult, since the motion induces noise to the images and the lighting conditions vary heavily.

The reliability of the vanishing point calculations may be evaluated based on the geometry of the lines found. The lines found from the same plane, e.g. the floor and from the same side in regard to the camera, don't intersect at the vanishing point. That is, the slope of all same-plane lines is either positive or negative. This is shown in Fig. 3. The vanishing point, a red circle, is calculated using lines colored with blue. The green lines are lines orthogonal to the direction of the motion and therefore they don't intersect in the vanishing point calculated in the direction of motion. Heading rate information from images with vanishing points deemed erroneous is excluded from the integration and only measurements from the other sensors are used at those epochs.

The y coordinate of the vanishing point should be the same within a threshold for all images taken with a camera having unchanged pitch and roll angles as was discussed in the previous section. If the line geometry is satisfying the correctness of the vanishing point may be evaluated by examining its coordinates. If the y coordinate doesn't comply with the rule depicted above, the visual-aiding measurement is included in the integration with decreased reliability and weight.

The statistics of the errors in the heading change were calculated with a test using a visual-aided navigation system presented in the following section. The results were compared with a ground truth reference. The distribution of the errors was found close-to zero-mean Gaussian with the standard deviation (σ) being 1.9 degrees. A decreased trustworthiness was found for three images out of the total of 50. Fig. 4 shows the distribution of the errors. It is noted that the test performed is short in time, limited with images obtained, and contains only a particular indoor environment scenario, but it serves the demonstration purpose well and is expected to have generalization potential.

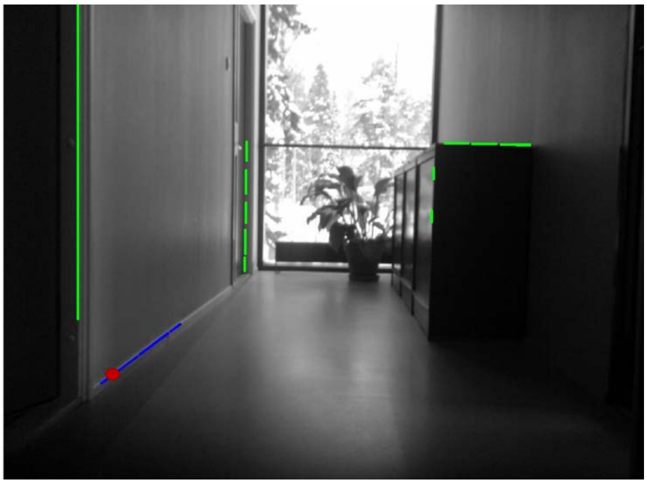

Figure 3 An erroneous vanishing point (shown with a red dot) due to bad geometry of the lines found.

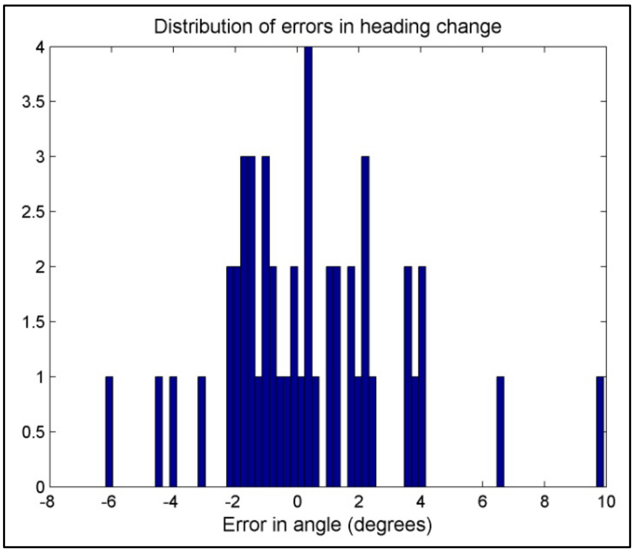

Figure 4 Distribution of the heading change errors

3.2 Integrating measurements with a Kalman filter

A filter is needed for the integration of the measurements from different location and motion sensors and the visual-aiding information. A Kalman Filter, e.g. (Welch and Bishop, 2006), (Grewal and Andrews, 2008), estimates the state $\mathbf{x} \in \Re^n$ of a discrete-time process with the equation (6)

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \tag{6}$$

with a measurement $\mathbf{z} \in \Re^m$ that is

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k \,. \tag{7}$$

The variables $\mathbf{w}_k$ and $\mathbf{v}_k$ represent noise. A covariance matrix **Q** defines the process noise and **R** the measurement noise. The matrix **A** propagates the state from the previous step *k-1* to the current step *k*. The matrix **B** is the control-input model that is applied to the control vector $\mathbf{u}_k$. The matrix **H** relates the state to the measurement $\mathbf{z}_k$. The Kalman Filter estimates the process state first and then, after obtaining the measurements, adjusts the state estimate and covariance using a variable called the Kalman gain.

The Kalman Filter was created for linear equations; some adjustments have to be made to be able to use it for usually non-linear navigation applications. The Extended Kalman filter is a suitable modification of the original Kalman filter: the state estimate is linearized around the current estimate. This may be done by using partial derivatives of the measurements and the process equations (Welch and Bishop, 2006). However, it has to be noted, that if the initial estimate of the state is incorrect, or if the process is modelled inexactly, the filter may quickly diverge due to the linearization. The measurement error estimates describing the confidence on the observations are fed to the filter through the measurement covariance matrix **R**.

4. Results with the Visual-aided Navigation System

Results obtained with the two-dimensional pedestrian indoor navigation system utilizing visual-aiding are shown in the following.

4.1 System construction

The visual-aided pedestrian indoor navigation system consists of wireless local area network (WLAN), Bluetooth (BT), and GPS location information for retrieving absolute positioning. The system is augmented with self-contained sensors; accelerometer offering information about the translation and two digital compasses about orientation, as well as with a Nokia N8 smartphone camera for the visual-aiding in form of the heading change rate measurement. The measurements are integrated using the Kalman filter construction discussed in (Kuusniemi et al., 2011) and presented briefly in the following.

The positioning equations for a pedestrian in a horizontal plane are presented in (8), where φ is the East coordinate, λ the North coordinate, both in meters, S (m/s) is the speed, θ (degrees) the heading defined with the origin East and counter-clockwise positive, and $\dot{\theta}$ (degrees/s) the heading change rate.

$$\begin{aligned}
\varphi_{k+1} &= \varphi_k + S_k \cdot \cos(\theta_k) \cdot \Delta t_k + w_1 \\
\lambda_{k+1} &= \lambda_k + S_k \cdot \sin(\theta_k) \cdot \Delta t_k + w_2 \\
\dot{\theta}_{k+1} &= \dot{\theta}_k + w_3 \\
\theta_{k+1} &= \dot{\theta}_k \Delta t_k + \theta_k + w_4 \\
S_{k+1} &= S_k + w_5
\end{aligned} \tag{8}$$

The variable Δt presents the time between two epochs and w_{i} is the noise in the measurements.

The state vector is defined as

$$\mathbf{x}_k = \begin{bmatrix} \varphi \\ \lambda \\ \dot{\theta} \\ \theta \\ S \end{bmatrix}_k \tag{9}$$

and thus the state model and measurement model are defined as

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_k \tag{10}$$

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \tag{11}$$

where the process noise is $\mathbf{w}_k \sim N(0, \mathbf{Q}_k)$, the measurement noise is $\mathbf{v}_k \sim N(0, \mathbf{R}_k)$. To avoid linearization, the state transition matrix is defined here simplified as

$$\mathbf{A}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & \cos(\theta_{k-1})\Delta t_k \\ 0 & 1 & 0 & 0 & \sin(\theta_{k-1})\Delta t_k \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \Delta t_k & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{12}$$

$\mathbf{A}_k$ is approximated as a constant matrix at every time epoch k. The corresponding process noise matrix $\mathbf{Q}_k$ is

$\mathbf{Q}_k =$

$$\begin{bmatrix} q_1\Delta t + \frac{b^2}{3}q_5(\Delta t)^3 & b\frac{a}{3}q_5(\Delta t)^3 & 0 & 0 & \frac{b}{2}q_5(\Delta t)^2 \\ b\frac{a}{3}q_5(\Delta t)^3 & q_2\Delta t + \frac{a^2}{3}q_5(\Delta t)^3 & 0 & 0 & \frac{b}{2}q_5(\Delta t)^2 \\ 0 & 0 & q_3\Delta t & \frac{(\Delta t)^2}{2}q_3 & 0 \\ 0 & 0 & \frac{(\Delta t)^2}{2}q_3 & \frac{(\Delta t)^3}{4}q_3 & 0 \\ \frac{b}{2}q_5(\Delta t)^2 & \frac{b}{2}q_5(\Delta t)^2 & 0 & 0 & q_5 \end{bmatrix}_k \quad (13)$$

where

$b = \cos(\theta_{k\text{-}1})$ and $a = \sin(\theta_{k\text{-}1})$

The number of measurements fed to the filter is varied on an epoch-to-epoch basis based on the availability of the sensors; each source of measurement information has a different output and availability rate. The non-availability situations of the visual aiding were discussed in the previous section. The full-scale measurement vector is $\mathbf{z}_k = (\varphi_{GPS}\ \lambda_{GPS}\ \varphi_{WLAN}\ \lambda_{WLAN}\ \varphi_{BT}\ \lambda_{BT}\ \dot{\theta}_{CAM}\ \theta_{DC1}\ \theta_{DC2}\ S_{ACC}\ \ S_{GPS})^T$, where *BT* is the Bluetooth, *WLAN* the wireless area network, *CAM* the camera, *DC1* and *DC2* digital compasses, and *ACC* the accelerometer. The speed from the accelerometer, S_{ACC}, is obtained through assessing the acceleration pattern by applying speed detection and assumptions about the motion. The WLAN and BT position results are obtained by a fingerprinting database approach utilizing Bayesian estimation and a maximum likelihood algorithm, more details can be found in (Kuusniemi et al., 2011).

The measurement covariance matrix $\mathbf{R}_k$ is a diagonal matrix with its size depending on the amount of measurements at the corresponding epoch *k*

$$\mathbf{R}_k = diag(\sigma^2_{\varphi_{GPS}}, \sigma^2_{\lambda_{GPS}}, \sigma^2_{\varphi_{WiFi}}, \sigma^2_{\lambda_{WiFi}}, \sigma^2_{\varphi_{BT}}, \sigma^2_{\lambda_{BT}}, \sigma^2_{\dot{\theta}_{CAM}}, \sigma^2_{\theta_{DC_1}}, \sigma^2_{\theta_{DC_2}}, \sigma^2_{S_{ACC}}, \sigma^2_{S_{GPS}}).$$

The visual-aided heading change measurement error estimate is derived by assessing the line geometry as introduced in the previous section. More details about the other measurement noise values as well as the process noise values used can be found in (Kuusniemi et al., 2011).

4.2 Results

The system was tested inside a typical office building. The ground truth was obtained with NovAtel's high-accuracy SPAN GPS/INS reference solution. The tester walked along the corridors of an office building pushing the cart that carried the system. A comparison was made between the system with and without the visual-aiding. The results of positioning with GPS, Bluetooth, and WLAN solely were also retrieved for comparison purposes. The visual-aiding based on tracking vanishing points fails when the camera is facing a plane and no lines in the field of the vision of the camera are present (Ruotsalainen et al., 2011). Fig. 5 shows the position result for different constructions of the navigation system. The SPAN ground reference is shown with a green line, the fused navigation system without visual-aiding in red and with the visual-aiding in blue. The accuracy of the position obtained without visual aiding is 4.6 meters and with visual-aiding 3.4 meters. The accuracy while navigating with GPS, Bluetooth, and WLAN solely is 13.1 meters, 6.3 meters, and 8.9 meters, respectively.

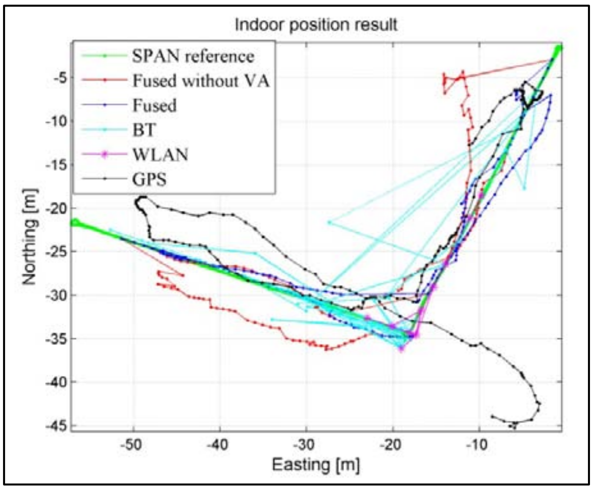

Figure 5: Position result with different constructions of the navigation system.

Due to that the visual-aided heading rate retrieval fails when the camera faces planes or no lines can be found, the visual-aiding solution is tested while walking along the corridor only and not during the turns at the end of the corridors. This decreases the availability of the system to 82 % with visual-aiding, while it is 99 %, 52 %, and 27 % for the GPS, Bluetooth, and WLAN-only systems, respectively. The accuracy and availability statistics are summarized in Table I.

Table I: Horizontal error and availability statistics for positioning with a fused system with and without visual-aiding, GPS, Bluetooth, and WLAN only.

Statistics	mean horizontal error (m)	availability (%)
Visual-aided	3.4	82 %
Without visual	4.6	82 %
GPS	13.1	99 %
Bluetooth	6.3	52 %
WLAN	8.9	27 %

The accuracy of the heading change measurements obtained with visual-aiding was also evaluated. The results shown in (Ruotsalainen et al., 2011) were obtained with a static camera and fairly stable lighting conditions: the mean error of the heading change being 1.5 degrees in the stationary case. In the solution described in this paper, the situation is more demanding. The motion of the camera introduces noise to the images as well as the varying lighting circumstances while moving along the corridors of the office building. Fig. 3 presented earlier shows the demanding lighting conditions found indoors. The pedestrian test discussed was done by taking 50 images while moving, and calculating thereafter the vanishing points for obtaining the heading change information. Only three vanishing points were deemed erroneous based on the geometry rules introduced earlier. These images were fully discarded in addition to the solution unavailability during turns. Fig. 6 presents the heading during the pedestrian indoor test. The turns were intentionally left out of the scope of the analysis and will be subject of future research. The mean error of the heading change was 2.1 degrees. The statistics of the heading change errors are shown in Table II.

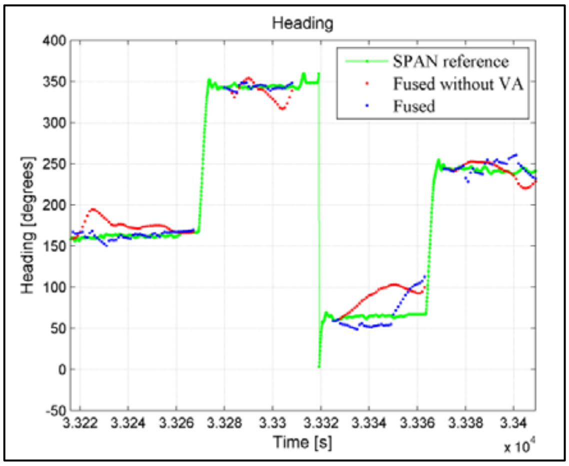

Figure 6: The heading in degrees during the test. The SPAN reference is shown with green, the fused solution without visual-aiding with red, and the visual-aided solution with blue.

Table II: Error statistics of the visual-aided derived heading change.

Statistics	min error (deg)	max error (deg)	mean error (deg)	std of error (deg)
	0.005	9.9	2.1	1.9

5. Conclusions

We have presented a two-dimensional pedestrian indoor navigation system using computer-vision based processing methods for providing the heading change measurements of a moving object. The change of the heading angle obtained by monitoring the change of vanishing point coordinates calculated from consecutive images is accurate, with the mean error being 2.1 degrees in the test conducted. The algorithm used in the system presented is suitable for navigating with a smartphone as a platform due to its 1 Hz computation frequency obtainable with a lightweight vanishing point calculation algorithm. The availability of the fused navigation system is also highly increased compared to navigation with especially Bluetooth or WLAN only. We have presented a method for evaluating the reliability of the vanishing point calculations based on the image geometry. The reliability estimates direct the emphasis given to the visual-aiding information while integrating the measurements obtained with the different navigation sensors. The accuracy of the visual-aided position solution (3.4 meters) is increased more than one meter compared to the solution without visual aiding (4.6 meters), and much compared to navigating with GPS, Bluetooth or WLAN only. The presented method provides the user with a two-dimensional position solution, which may be extended into three dimensional positioning by adding information from a supplementary sensor source. The source of height may be e.g. a barometer providing the altitude information or a wireless network access point providing floor-level information.

References

Aoki H., Schiele B. and Pentland A. (1999), ***Realtime personal positioning system for wearable computers***, Proceedings of the 3rd IEEE International Symposium on Wearable Computers, ISWC 1999, Washington, DC, USA pp. 37-44.

Canny J. F. (1986), ***A computational approach to edge detection***, IEEE Transactions on Pattern Analysis and Machine Intelligene, Vol PAMI-8, No 6, pp. 679-698.

Collin J. (2006), ***Investigations of self-contained sensors for personal navigation***, Doctoral Dissertation, Tampere University of Technology.

Corke P., Lobo J., and Dias J. (2007), ***An introduction to inertial and visual sensing***, The International Journal of Robotics Research, Vol 26, No 6., pp. 519-535.

Diel D., DeBitetto P., and Teller S. (2005), ***Epipolar constraints for vision-aided inertial navigation***, Proceedings of the 7th IEEE Workshop on Applications of Computer Vision, January 2005, Breckenridge, CO, USA, pp. 221-228, DOI: 10.1109/ACVMOT.2005.48.

Fischler M.A., and Bolles R.C. (1981), ***Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography***, Communications of the ACM, Vol 24, No 6, pp. 381-395.

Forsyth D.A., and Ponce J. (2003), ***Computer vision: a modern approach,*** Prentice Hall.

Gallagher A.C. (2005), ***Using vanishing points to correct camera rotation in images***, Proceedings of the 2nd Canadian Conference on Computer and Robot Vision, 9-11 May, 2005, IEEE Xplore, pp. 460-467, DOI: 10.1109/CRV.2005.84.

Grewal M.S., and Andrews A.O. (2008). ***Kalman filtering, theory and practice using Matlab,*** third edition, John Wiley & Sons, Inc.

Hartley R., and Zisserman A. (2003), ***Multiple view geometry in computer vision***, second edition, Cambridge University Press.

Hough P.V.C., (1962), ***Method and means for recognizing complex patterns***. U.S. Patent 3069654.

Jepson A., Fleet D. (2009), ***Foundations of Computer Vision, Robust estimation,*** lecture notes: http://www.cs.toronto.edu/~jepson/csc2503/index09.html.

Kessler C., Ascher C., Frietsch N., Weinmann M., and Trommer G. (2010), ***Vision-based attitude estimation for indoor navigation using vanishing points and lines***, Proceedings of Position Location and Navigation Symposium (PLANS), 4-6 May, 2010, IEEE Xplore, pp. 310-318, DOI: 10.1109/PLANS.2010.5507247.

Kosecka J., and Zhang W. (2002), ***Video compass***, Proceedings of European Conference on Computer Vision, Springer Verlag, pp. 657-673, DOI: 10.1007/3-540-47979-1_32.

Kourogi M., and Kurata T. (2003), ***A method of personal positioning based on sensor data fusion of wearable camera and self-contained sensors***, Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, MFI2003, 30 July-1 August 2003, pp. 287 – 292, DOI: 10.1109/MFI-2003.2003.1232672 .

Kuusniemi H., Chen L., Ruotsalainen L., Pei L., Chen Y., Chen R. (2011), ***Multi-sensor multi-network seamless positioning with visual aiding***, Proceedings of ICL-GNSS 2011, June 29-30, 2011, Tampere, Finland, pp. 146-151, IEEE Xplore, DOI: 10.1109/ICL-GNSS.2011.5955257 .

Retscher G. (2007), ***Test and integration of location sensors for a multi-personal navigator***, the Journal of Navigation, Vol. 60, pp. 107-117.

Robertson D., and Cipolla R. (2004), ***An image-based system for urban navigation***, Proceedings of the British Machine Vision Conference, September 2004, Kingston-upon-Thames, UK, pp. 260-272.

Ruotsalainen L., Kuusniemi H., Chen R. (2011), ***Heading change detection for indoor navigation with a smartphone camera***, Proceedings of IPIN 2011, September 21-23, 2011, Guimaraes, Portugal, IEEE Xplore (in press).

Saarinen J. (2009), ***A sensor-based personal navigation system and its application for incorporating humans into a human-robot team***, Doctoral Dissertation, Helsinki University of Technology.

Sonka, M., Hlavac, V., Boyle, R. (2008). ***Image Processing, Analysis and Machine Vision,*** third edition, Stamford, Connecticut: Cengage Learning.

Steinhoff U., Omerčević D., Perko R., Schiele B., and Leonardis A. (2007), ***How computer vision can help in outdoor positioning***, Proceedings of the European conference on Ambient intelligence 2007, Springer-Verlag Berlin Heidelberg, LNCS 4794, pp. 124-141.

Welch G., and Bishop G. (2006), ***An introduction to the Kalman filter***, report, UNC-Chapel Hill TR 95-041.

Zhang W., and Kosecka J. (2006), ***Image based localization in urban environments***, The Third International Symposium on 3D Data Processing, Visualization and Transmission, IEEE Computer Society, 14-16 June, 2006, Chapel Hill, NC, USA, pp. 33-40, DOI: 10.1109/3DPVT.2006.80.

Biography

Laura Ruotsalainen is a PhD candidate in the Department of Computer Systems, Tampere University of Technology, Finland. She received her M.Sc. degree from the Department of Computer Science, University of Helsinki in 2003. After working as an information analyst at the VTT Technical Research Centre of Finland, she joined the Department of Navigation and Positioning at the Finnish Geodetic Institute in July 2010 as a research scientist. Her doctoral research is focused on visual-aided seamless indoor/outdoor navigation.