Mitigation of Attitude and Gyro Errors through Vision Aiding

Laura Ruotsalainen*, Jared Bancroft, Gérard Lachapelle
Department of Geomatics Engineering
Schulich School of Engineering, University of Calgary, Canada

*Visiting Researcher from the Finnish Geodetic Institute

***Abstract*—Accurate positioning of first responders, electronic monitoring, and military personnel is often critical in GNSS denied environments. In such environments, inertial navigation systems (INS) are typically the preferred tool to be used for navigation. However, the gyros suffer from errors including biases, scale factors and g-dependent errors being the most significant ones. In order to sustain an accurate navigation solution for long durations, the gyroscope errors have to be measured and mitigated. Ideally, this calibration is done in situ.**

The attitude obtained using visual information is independent of the errors affecting the gyroscope. Human-made environments are commonly full of straight and parallel lines found in orthogonal directions. Perspective projection mapping transforms three-dimensional scenes into two-dimensional images. The process maintains the straight lines but modifies their parallelism resulting in an apparent point intersection of the lines. This point is called the vanishing point. Lines in three orthogonal directions constitute three vanishing points. The vanishing point locations are dependent on the camera rotation, but not camera translation. By monitoring the motion of the vanishing point locations in consecutive images, the relative roll, heading and pitch attitudes may be obtained. The absolute attitude, known from some *a priori* knowledge of the building layout, is then used to update the inertial navigation filter. Over time the visual measurements mitigate the cumulative errors of the gyro bias, scale factor and g-dependent bias.

The performance of the vision-aided INS based navigation approach is evaluated herein. A camera is attached to a backpack and foot of a user moving through typical pedestrian based environments. The case of a foot-mounted camera is unique because of the high accelerations experienced during the human gait. The visual-aiding correction is found to significantly improve the attitude accuracy, especially the heading. Using the body solution, namely the camera and INS attached to a backpack, the vision-aiding yielded a 93 % improvement in the heading error during evaluation tests. With a foot-mounted solution, namely the INS and camera attached to the ankle of the user, the horizontal position error decreased by 34 %.

Keywords - visual-aiding; vanishing point; attitude; gyroscope; inertial navigation.

I. INTRODUCTION

Accurate indoor navigation is a challenging task to accomplish given current Global Navigation Satellite Systems (GNSS) limitations. The circumstances under which first responders, electronic monitoring and military personnel operate make the requirements set for a navigation system particularly demanding. The communication and RF infrastructures may be inoperable or have degraded functionality [1]. With these restrictions, self-contained sensors carried by the user are desirable equipment for positioning. With a known initial position, the current position may be propagated using a triad of gyroscopes and accelerometers for a limited time [2]. The propagation is done using standard inertial algorithms incorporating the attitude obtained using the integration of the gyroscopes and the double integration of the accelerometer. The limitation of the self-contained sensors is the cumulative measurement errors that affect the accuracy of the attitude obtained from the gyroscopes [3]. When the navigation system users are first responders whose movements follow an irregular path, their need for accurate position information lasts for up to 30 minutes, causing unaided INS methods to fail; the INS errors can be corrected by re-aligning the system using sensor nodes, namely ultrasonic and radio frequency (RF) nodes, placed on the ground by the user [4]. The method provides positioning by correcting the solution using the nodes while the user is hovering around or returning following the route walked previously. Herein a method of updating the navigation filter's attitude using vision-aiding and thereby providing accurate absolute user position is presented. The method uses only equipment carried by the user, thus not leaving any marks behind or requiring infrastructure, which is desirable especially for military personnel and electronic monitoring.

When the user is experiencing large accelerations the g-dependent bias resulting from mass imbalances occurring from manufacturing becomes an important source of the gyro errors in addition to the bias and scale factor. The acceleration amplitudes are higher in the vertical direction than in both horizontal directions. The total acceleration of a person can rise up to a maximum of 5 g when measured on the back to 12 g on the ankle [5]. If the error is not compensated, it can affect the micro-electromechanical measurement systems (MEMS) rate gyros to 100 °/hr/g [6]. Vision-aiding is a complementary

978-1-4673-1954-6/12/$31.00 ©2012 IEEE

method mitigating the gyroscope errors through attitude updating because it suffers from different error sources. The camera rotation may be obtained using features in images called vanishing points; namely the points where the lines parallel in the scene seem to intersect. When the camera orientation with respect to the gyroscope is known, the location change of the vanishing points in consecutive images may be transformed into attitude of the gyroscope. By integrating the image-based rotation with the rate gyroscope's own measurements, the gyroscope is calibrated and the errors are mitigated during navigation. The method has been used before by Prahl and Veth [7] for UAVs and Kessler et al. [8] for pedestrian indoor navigation. The contribution of this paper is the accommodation of unrestricted pedestrian motion by using error modeling for the obtained vanishing points and fault detection for the integration of the measurements. The feasibility of the method presented is also evaluated through experiments with extended duration and challenging environments.

This paper is organized as follows; first the principle of calculating the camera rotation using the vanishing point location information in the images is explained. Secondly, the method of integrating the image based rotation measurements with the Kalman filter is introduced. Finally, the results from two experiments evaluating the vision-aided navigation solution accuracy with firstly mounting the camera and IMU on the backpack of the user and secondly on the foot are presented.

II. VISUAL ANGULAR VELOCITY MEASUREMENTS

Most human made environments, namely indoors and urban outdoor areas, consists of segments forming a cartesian coordinate system with straight lines in three orthogonal directions. The coordinate system is called the Manhattan grid [9] and it provides a good basis for vision-aided urban environment navigation with vanishing points. The vanishing point is a point in the image where the lines parallel in the real world seem to intersect. The straight parallel lines in three orthogonal directions induce three vanishing points. The vision-aided navigation approach consists of three frames, namely the world frame (i.e. the scene), the camera frame and the image frame. The frames are shown in Figure 1. When the camera frame's coordinate axes (C) are precisely aligned with the world frame's axes (W), the central vanishing point ($\mathbf{v}_z$) composed by the lines in the Z-axis direction falls to the principal point (u,v) and the vanishing points in the X- and Y-axis directions fall to infinity on the corresponding axes. When the camera rotates in the world frame, the vanishing points in consecutive images move with respect to the rotation angles. The heading, pitch and roll angular velocity magnitudes may be obtained as follows.

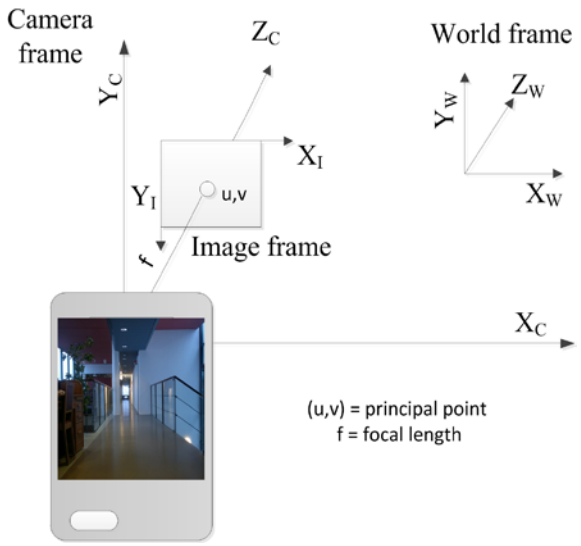

Figure 1. Coordinate frames for vision-aiding; camera (C), world (W) and image (I) frames.

A. *Calculating vanishing points*

Straight lines in images are identified using the Hough Lines algorithm and classified based on their orientation into vertical, horizontal or in the direction of propagation, as shown in Figure 2. Lens distortion correction, explained below, is the reason why the lines in the figure do not always coincide with the lines in the image. Most of the lines are in the direction of propagation and used for calculating the central vanishing point. Heading and pitch, relative to the environment, may be observed using the central vanishing point, but information of the horizontal or vertical vanishing point is needed for roll. Because the horizontal lines are not frequent in the navigation scenes, the roll is calculated using the vertical vanishing points. Central and vertical vanishing point locations are calculated for all images using a voting algorithm. The algorithm identifies the points being intersection points for most of the line pairs in the corresponding direction. Using information of the two vanishing points the heading, pitch and roll angular velocities may be obtained as explained below.

A camera can be rolled over 15 degrees for the purpose of obtaining images with special viewpoints [10], but this is not convenient for vision-aided navigation. Because the calculation of the accurate vertical vanishing point is not always possible (due to noise in the images and shortage of lines) the roll's magnitude is monitored. If the roll's magnitude exceeds 15 degrees, the vertical vanishing point is excluded from the calculations. In these situations the roll is evaluated to be zero and the heading and pitch are calculated more accurately using only the central vanishing point. If the camera is actually experiencing roll when the calculations fail and the roll is estimated to be zero, errors also appear in the heading and pitch.

Fortunately, in most cases these errors are small. In most common cases when the heading changes and roll between two images are less or equal to five degrees, the errors in estimated heading and pitch are 0.6 and 0.1 degrees, respectively. In an

extreme case when the camera is simultaneously experiencing roll and heading change of 15 degrees between two consecutive images, the errors in heading and pitch are 6.1 and 2.8 degrees, respectively. When the camera is otherwise static (i.e. the heading change and pitch are around one degree between consecutive images) even a large roll causes small errors to the observed heading and pitch, namely 0.3 and 0.01 degrees, respectively. Table I summarizes some errors arising from camera motions between two consecutive images when the roll is erroneously estimated to be zero due to vertical vanishing point calculation failure.

TABLE I. EFFECT OF ROLL ERROR ON ANGLE OBSERVATIONS

Real camera rotation (degrees)			**Errors in observation when roll estimated to be zero (degrees)**	
Heading	***Pitch***	***Roll***	***Heading***	***Pitch***
1	1	-15	0.3	0.01
5	5	-5	0.6	0.1
15	-15	15	6.1	2.1
-15	-15	15	5.2	2.8

1) *Error evaluation*

The vanishing point calculations may fail due to the restricted line geometry, meaning there are insufficient lines arising from the scene's floor and ceiling constructions or due to a substantial number of non-orthogonal lines. To avoid errors in the navigation solution, the accuracy of the estimated vanishing point has to be evaluated and supplied to the integration algorithm. The pedestrian motion is unforeseeable and arbitrary and therefore the vanishing point locations can theoretically have no restrictions. That is, even the central vanishing point cannot be assumed to lie inside the image.

The accuracy evaluation of the estimated vanishing point is based on the geometry of the lines intersecting at the vanishing point, and a concept of an "lDOP", discussed in more detail in [11], is used. The line geometry evaluation is based on dividing the image into four sections around the estimated central vanishing point. If the lines intersecting at the vanishing point are found from four or three sections, the geometry is good for the calculations, and a low lDOP value is assigned. If lines are present in two sections, their mutual orientation is considered. When the slopes of the lines diverge much, the calculation accuracy is higher than when they are close to each other, and a lower lDOP value is assigned.

When the line geometry is severely restricted, namely lines are found only from one section, the lDOP value is evaluated based on the angle difference between each line pair and the image x-axis. Following the work of [12] on determining the Dilution of Precision for GNSS positioning, the lDOP value is computed as

$$lDOP=\sqrt{\frac{1}{\mathbf{H}}\left(\cos^2(\alpha_1)+\cos^2(\alpha_2)+\sin^2(\alpha_1)+\sin^2(\alpha_2)\right)}, \quad (1)$$

where (α_1) is the angle between the x-axis and the first line and (α_2) between the x-axis and the second line and $\mathbf{H}^{-1}$ is defined as

$$\mathbf{H}^{-1}=\frac{1}{|\mathbf{H}|}\begin{bmatrix}\sin^2(\alpha_1)+\sin^2(\alpha_2) & -\cos(\alpha_1)\sin(\alpha_1)+\cos(\alpha_2)\sin(\alpha_2)\\ \cos(\alpha_1)\sin(\alpha_1)+\cos(\alpha_2)\sin(\alpha_2) & \cos^2(\alpha_1)+\cos^2(\alpha_2)\end{bmatrix} \quad (2)$$

Thus, $|\mathbf{H}|=\sin^2(\alpha_1-\alpha_2)$ is the determinant of **H**. The smallest possible value for lDOP is $\sqrt{2}$, occurring when the two lines are perpendicular.

2) *Lens distortion correction*

The best accuracy for vision-aided calculations is obtained when a camera with a wide angle lens offering an extended field-of-view is used [13]. However, the wide angle lens results in radial distortion in the images. If the distortion is not corrected, the vanishing point calculation accuracy suffers. According to [14], the rectification of the whole image introduces aliasing effects complicating the feature detection. For optimal result, the radial distortion is corrected only for the lines extracted from the images using a model presented in [15].

The radial distance (r_d) of the normalized distorted image points (x_d, y_d) from the radial distortion center, which is in this as in most cases the principal point (u,v) [14], is

$$r_d=\sqrt{x_d^2+y_d^2}\ . \quad (3)$$

Using the radial distance of the distorted image points, the radial distance (r) of the corrected image points (x_u, y_u) is obtained as

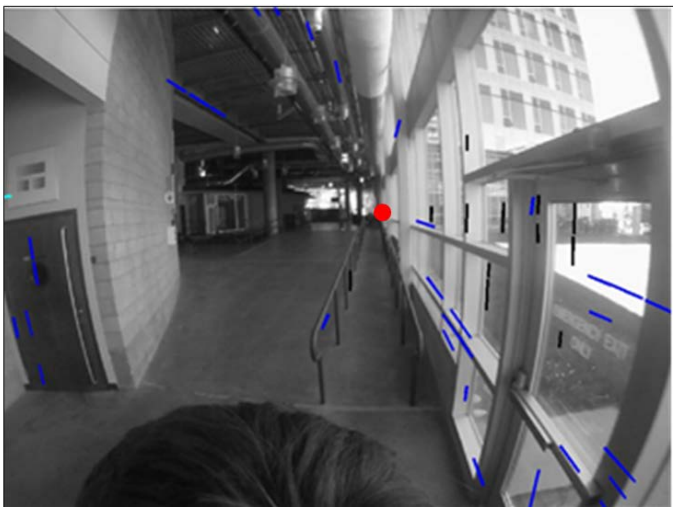

Figure 2. Lines identified from an image are classified as vertical (black), horizontal (turquoise) and in the direction of propagation (blue). The vanishing point is shown as a red dot.

$$r = r_d(1 - k_1 r_d^2 - k_2 r_d^4). \quad (4)$$

The constants k_i are the distortion values specific to the camera and are obtained from calibration. The corrected and distorted image points are related as

$$\begin{aligned} x_d &= x_u(1 + k_1 r + k_2 r^2), \\ y_d &= y_u(1 + k_1 r + k_2 r^r) \end{aligned}. \quad (5)$$

All lines used for vanishing point calculations are corrected using the above equations. Figure 3 shows the effect of the correction on the vanishing point location. On the left, the vanishing point is calculated using the original distorted lines and has deviated from the correct location and on the right the calculation is done using the corrected lines locating the vanishing point to the right position. Also some image parts of the image on the right are corrected to show how much the distortion affects performance. For example the electric box on the wall on the left part of the image has moved many pixels to the left due to the distortion correction.

B. Attitude from camera rotation matrix

When the camera rotates, the roll (β), pitch (ϕ) and heading (θ) form the camera's rotation matrix **R** that is defined as

$$\mathbf{R} = \begin{bmatrix} \cos\beta\cos\theta - \sin\beta\sin\phi\sin\theta & -\sin\beta\cos\phi & \cos\beta\sin\theta + \sin\beta\sin\phi\cos\theta \\ \sin\beta\cos\theta + \cos\beta\sin\phi\sin\theta & \cos\beta\cos\phi & \sin\beta\sin\theta - \cos\beta\sin\phi\cos\theta \\ -\cos\phi\sin\theta & \sin\phi & \cos\phi\cos\theta \end{bmatrix} \quad (6)$$

The angle magnitudes may be obtained using the known locations of the vertical and central vanishing points ($\mathbf{v}_y$,$\mathbf{v}_z$) and the camera calibration matrix **K** using the relation **V**=**KR** [10]. The camera calibration matrix incorporates the camera intrinsic parameters, namely the focal length (f_x,f_y) and the principal point (u,v). The parameters are resolved by calibrating the camera in advance [16][17], or may be approximated using the focal length obtained from the camera manufacturer and the image center point as a principal point. The camera calibration matrix **K** is

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & u \\ 0 & f_y & v \\ 0 & 0 & 1 \end{bmatrix}. \quad (7)$$

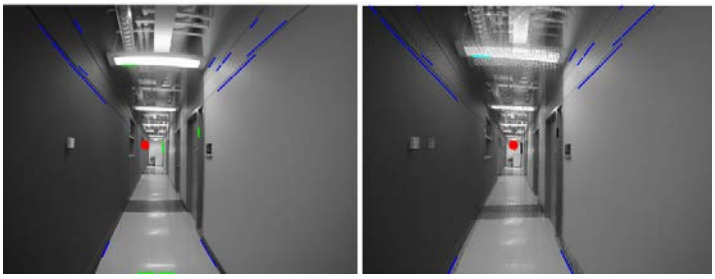

Figure 3. Effect of distortion correction on the accuracy of the vanishing point location and image appearance. The image on the left is the original one and on the right the corrected one.

III. INTEGRATING THE IMU AND VISUAL MEASUREMENTS

The errors in the gyroscopes cause the attitude measurements drift, introducing continuously increasing errors in the navigation solution. The errors consist of the gyro bias, scale factor and non-orthogonalities, the g-dependent error and noise. The error model, discussed in more detail in [18], is

$$\tilde{\boldsymbol{\omega}}_{ib}^b = \mathbf{S}_g \boldsymbol{\omega}_{ib}^b + \mathbf{b}_g + \mathbf{G}\mathbf{f}_{ib}^b + \eta_g, \quad (8)$$

where $\tilde{\boldsymbol{\omega}}_{ib}^b$ is the gyroscope angular velocity measurement, $\mathbf{S}_g$ is a matrix including the scale factors and non-orthogonalities, $\boldsymbol{\omega}_{ib}^b$ is the body (b) turn rate with respect to the inertial (i) frame measured by the gyroscope, $\mathbf{b}_g$ are the gyro biases, **G** is a 3x3 matrix of the g-sensitivity coefficients, $\mathbf{f}_{ib}^b$ is the specific force and η_g is the noise.

The g-dependent bias is introduced by high accelerations, especially affecting sensors attached to the ankle. The g-dependent bias in the gyroscopes is a result of mass imbalances caused by the manufacturing process and can impact the MEMS gyros to 100 degrees/hour/g or more when uncompensated. In order to obtain an accurate navigation solution, the gyro bias, scale factor and g-dependent bias are mitigated by vision-aiding the attitude measurements using the angular velocities obtained from the camera rotation matrix as explained below.

GNSS, IMU and visual data are fused with a tightly coupled 21-state extended Kalman filter (EKF) (23 if GPS receiver clock states are added). The filter consists of linear perturbations of the position, attitude, velocity, gyro and accelerometer bias, three gyro scale factor coefficients and three g-sensitivity coefficients and is defined as

$$
\begin{aligned}
\delta\dot{\mathbf{r}}^e &= \delta\mathbf{v}^e \\
\delta\dot{\mathbf{v}}^e &= \mathbf{N}^e\delta\mathbf{r}^e - 2\Omega^e_{ie}\delta\mathbf{v}^e - \mathbf{F}^e\boldsymbol{\varepsilon} + \mathbf{R}^e_b\left(\mathbf{b}_a\right) \\
\dot{\boldsymbol{\varepsilon}}^e &= -\Omega^e_{ie}\boldsymbol{\varepsilon}^e + \mathbf{R}^e_b\left(\left(I - \mathbf{S}_g\right)\boldsymbol{\omega}^b_{ib} + \mathbf{b}_g + \mathbf{G}\mathbf{f}^b_{ib}\right) \\
\dot{\mathbf{b}}_a &= -\tau_a^{-1}\mathbf{b}_a \\
\dot{\mathbf{b}}_g &= -\tau_g^{-1}\mathbf{b}_g \\
\dot{\mathbf{S}}_g &= 0 \\
\dot{\mathbf{G}} &= 0 \\
c\dot{dT} &= \delta c\dot{dT} + \eta_{cdT} \\
c\ddot{dT} &= \eta_{c\dot{dT}}
\end{aligned}
\tag{9}
$$

where $\mathbf{r}^e$, $\mathbf{v}^e$ are the position and velocity vectors in the earth centered earth fixed (ECEF) frame, $\boldsymbol{\varepsilon}$ is the perturbation of the Euler angles relating the body frame to the ECEF frame and $\mathbf{b}_a$ and $\mathbf{b}_g$ are the biases of the accelerometer and gyro. The inertia tensor is denoted with $\mathbf{N}^e$, the skew symmetric forms of the earth rotation vector Ω^e_{ie} and specific force measurement $\mathbf{F}^e$. The rotation matrix $\mathbf{R}^e_b$ rotates the specific force and angular velocity from the body to ECEF frame. Receiver clock errors are denoted as *cdT*. The integration is discussed further in [18] and [19].

The navigation filter attitude is updated using the visual heading, pitch and roll measurements. Only the measurements having an lDOP value below a threshold are used for the update. However, some errors arising from the environments that are non-suitable for the vanishing point based method are not identified by the error detection algorithm and have to be discarded using a fault detection algorithm. The fault detection is applied by accepting only the standardized visual measurement values w_i that do not surpass a pre-defined threshold value [20]. The standardized visual measurements are obtained from the innovations of the heading, pitch and roll values v_i and their corresponding estimated standard deviation, σ_v, as

$$
w_i = \left|\frac{v_i}{\sigma_v}\right|, i = 1{:}n \tag{10}
$$

In some cases only the visual heading measurement is found faulty and discarded, while the pitch and roll measurements are used in an update.

IV. Experimental results

The method of updating the navigation filter attitude with visual measurements was evaluated through two experiments; for the first round the test equipment was attached to the backpack and for the second to the foot, specifically the ankle, of the user. The experiments were conducted on the University of Calgary campus, mainly inside buildings.

The equipment used for the experiments consisted of Analog Devices ADIS16488 inertial measurement units [21] and a GoPro Hero helmet camera [22]. The ADI IMU is a high grade MEMS IMU, with a 12 °/hr in-run bias stability and 1620 °/hr noise level. The GoPro camera has a wide-angle lens and provides tall HD video stream. The MEMS IMU data rate was 200 Hz and camera 10 Hz rate. The NovAtel SPAN-SE GPS/GLONASS receiver with a Northrop Grumman's tactical grade LCI IMU was used as a reference system and carried in the backpack for both experiments.

A very challenging environment for the visual measurements was selected as a test area to assure the usability of the method in a real-life navigation situation. The area consisted of numerous sharp turns, wide regions such as cafeterias and outdoor garden areas. The experiment was conducted during office hours adding many moving humans into the images. The route for the experiments is shown in Figure 4.

The visual heading, pitch and roll algorithms are obtained based on the assumption that the walls and floor are completely orthogonal and planar. When the camera is totally aligned with the scene, namely the world frame, the central vanishing point lies in the image principal point and the visual heading magnitude corresponds to the scene heading and pitch and roll are zero. The weakness of the vanishing point based vision-aiding is that the method fails in sharp turns. This is due to a temporary loss of sight of the lines in the direction of propagation. Because the continuous transition from one scene to another when turning around a corner fails, the heading of the new scene is unknown and has to be measured again. Herein the scene heading is detected using the building layout. The heading obtained from the building layout is used only as the initial heading after a turn and, until the next turn, the visual heading is obtained from the central vanishing point location change only.

In real-life navigation the orthogonality and planarity assumptions do not always hold. These situations introduce errors in the visual attitude measurements not identifiable by the error detection algorithm. The erroneous measurements are discarded from the updates using the fault detection method described earlier. One such situation is when the user is walking along a ramp. In that case all visual pitch measurements deviate from the real pitch, as shown in Figure

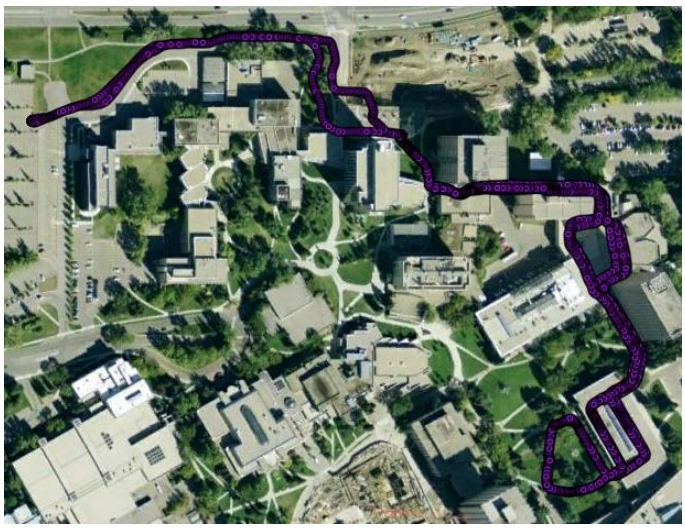

Figure 4. Route for the experiments.

5. The figure shows an eight minute clipping of the backpack mounted experiment discussed below. The plot encompasses all visual pitch measurements, including the ones that are excluded from the attitude determination due to the large lDOP value. The reference heading, scaled for visualizing purposes, is shown also in the figure to demonstrate the effect of the vanishing point based pitch measurement failure in the sharp turning situations. The pitch measurements obtained while the user is walking along a ramp are marked with a black square. When the user is experiencing a sharp turn and vision to the lines in the direction of propagation is lost, the error in visual pitch increases significantly. Fortunately most of the erroneous pitch, as well as heading and roll measurements, are discarded using the lDOP based error detection algorithm.

A. Equipment setup on the body

The effect of correcting gyroscope errors by updating the navigation filter attitude through vision aiding was tested by attaching the IMU and camera to a backpack carried by the user. The setup is shown in Figure 6. Data was collected in an experiment of 48 minutes conducted mainly indoors, preceding a 10 minute walk outdoors, thereby allowing the filter to converge. Because the purpose of the research was to assess the vision-aiding performance on gyro errors, GNSS data was only used for three minutes at the start of the experiment to provide an initial position.

The gyroscope's attitude measurements were corrected by updating the navigation filter with visual heading, pitch and roll measurements as explained above. The total number of images taken during the navigation was 29802, of which 16347 were discarded due to large lDOP values.

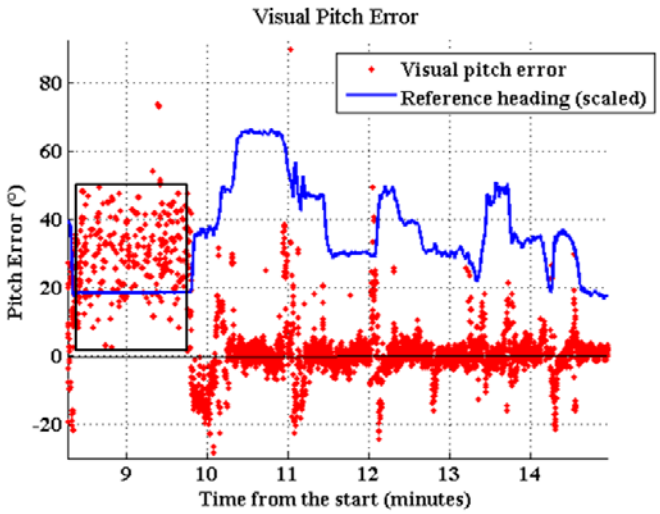

Figure 5. Errors in visual pitch meaurements due to a ramp and turns. All visual measurements, also the ones having a large DOP, are included in the image.

Figure 6. Data set 1 set up with the IMU and the camera located on the backpack.

The fault detection within the navigation filter further rejected 11 % of the remaining images. Visual pitch and roll updates only, with no heading, were accepted from 38% of the remaining images. This resulted into 8337 visual heading updates and 14549 visual pitch and roll updates to the navigation filter.

Figure 7 shows the results of the backpack mounted gyroscope bias estimation. Without vision-aiding updates the gyro biases remain unchanged. The visual updates decrease gyro biases few degrees in all axes. Figure 8 shows the effect of the visual updates on the attitude error. The vision-aiding improves the navigation solution's pitch and roll only slightly, as is shown in the figure and in Table II, namely the pitch root mean square error decreases from 1.7 to 1.4 degrees and the roll from 2.0 to 1.4 degrees. However, the heading improves significantly, namely 93 % as the root mean square error decreases from 29.5 to 2.1 degrees when the navigation filter is updated with visual measurements.

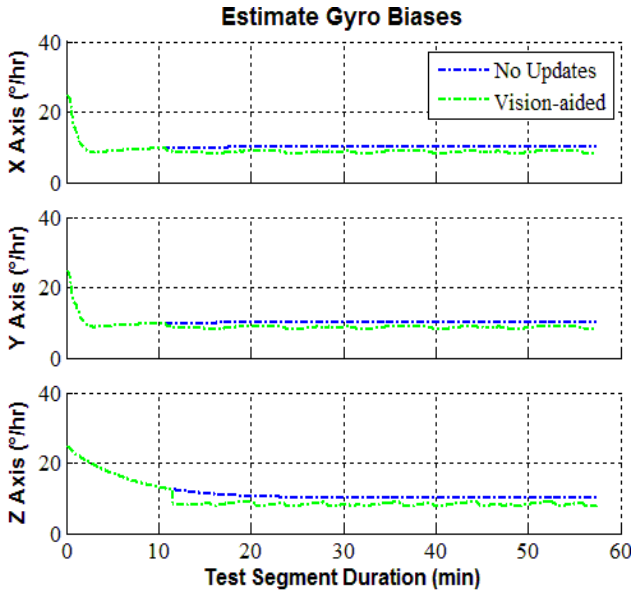

Figure 7. Estimated gyro biases using no updates (blue) and using vision-aided updates (green).

TABLE II. ATTITUDE ERRORS

Attitude error (rms, degrees)			
	Pitch	***Roll***	***Heading***
No updates	1.7	2.0	29.5
Vision-aided updates	1.4	1.4	2.1

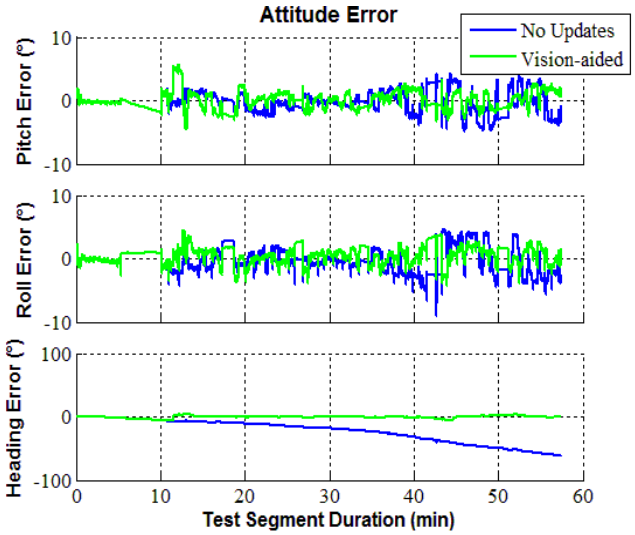

Figure 8. Attitude errors using no updates (blue) and using vision-aided updates (green).

B. Equipment setup at the foot

When the gyro is located on the ankle of the user the vertical acceleration can rise up to the maximum of 12 g causing very large g-dependent errors. The effect of correcting the errors through vision aiding of the attitude was tested by attaching the IMU and camera rigidly to each other and locating them on the ankle of the user. The setup is shown in Figure 9. Data was collected in an experiment of 43 minutes conducted mainly indoors. Because the purpose of the research was to assess the vision-aiding performance on attitude and gyro errors, GNSS data was only used for three periods of two to three minutes during the navigation in low canyons between buildings. A pedestrian navigation solution was obtained by integrating the vision-aided gyroscope attitude measurements and applying zero velocity updates to the inertial navigation filter. The integration was performed using the Kalman filter described above. Due to the lack of a reference system mounted on the foot (the reference system was carried in the backpack), the attitude errors could not be evaluated and the position errors were evaluated instead.

The visual heading, pitch and roll measurements were used as updates to the navigation filter attitude as explained above. The visual measurement calculation was challenging due to large camera movements when attached to the ankle of the pedestrian. The total number of images acquired during the experiment was 25664. Only 18 % received an lDOP value sufficiently low for trusting the visual measurements due to image blurring introduced by the fast motion of the foot and because the camera was pointing straight down to the floor during one step cycle period. Fault detection was used to remove the noise from the visual measurements. The fault detection within the navigation filter further rejected 65 % of the remaining images. Visual pitch and roll updates only, with no heading, were accepted from 18% of the remaining images. This resulted into 785 visual heading updates and 1617 visual pitch and roll updates to the navigation filter.

Table III and Figure 10 show the improvement of the position obtained with the vision-aided foot mounted navigation system. The periods when GNSS was used are shown in the figure with black squares. Vision-aiding improves the horizontal position significantly; the root mean square horizontal position error decreases from 30.9 m to 20 m, yielding an improvement of 34 %. Vision-aiding has no effect on the vertical position error, the error remaining at 68 m.

Figure 9. Data set 2 set up with the IMU, GPS antenna and camera attached. The system was located on the user's ankle.

TABLE III. POSITION ERRORS OF PEDESTRIAN NAVIGATION SYSTEM

RMS Position Errors (m)		
	Horizontal	***Vertical***
No vision-aiding	30.9	67.5
With vision-aiding	20.0	67.7

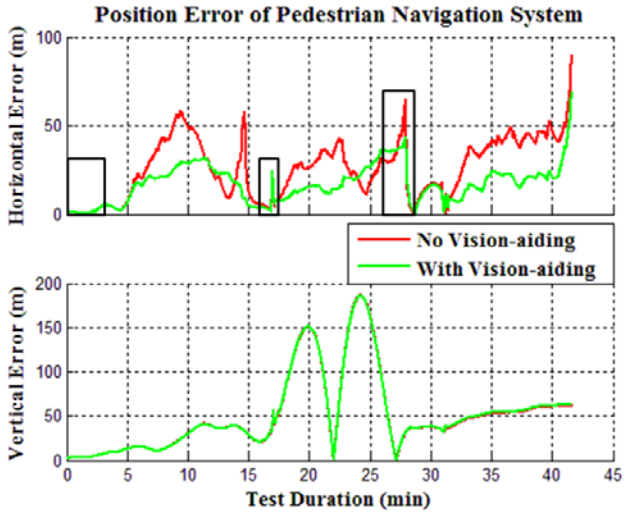

Figure 10. Position errors of the pedestrian navigation system without (red) and with (green) vision-aiding. GNSS was used within three periods (black squares).

V. CONCLUSIONS

The use of vision-aiding in urban environments was shown to significantly improve the attitude by adding updates to the navigation filter which further mitigated the effects of gyro biases, scale factors and g-dependent biases. The motion of the vanishing points arising from urban environments and found in images taken by the user was transformed into attitude information. Using error detection the accurate visual attitude information could be separated from the noise and used as updates to the navigation filter. The method yielded a 93 % improvement to the heading error when the backpack set up was used. The effect on the pitch and roll errors was minor. When the system was located on the ankle of the user, the method improved the horizontal position solution by 34 %. The effect on vertical position was negligible. The vision-aided inertial navigation system using a setup similar to the backpack, i.e. the IMU and camera attached to the body or helmet of the user, is beneficial for example for first responders navigating in buildings without infrastructure for absolute positioning. The foot mounted inertial sensors are coming more common in military, offender tracking and safety of life pedestrian navigation. The experiments showed significant improvement of the navigation solution when vision-aiding the inertial sensor induced attitude in such applications.

ACKNOWLEDGMENT

This research has been conducted partly within the project ISPACE (Indoor/Outdoor Seamless Positioning and Application for City Ecosystems) funded by the Finnish Technology Agency TEKES with the Finnish Geodetic Institute, Nokia Inc., Fastrax Ltd., Space Systems Finland Ltd., Bluegiga Ltd. and Indagon Ltd. Laura Ruotsalainen has also received a Nokia Foundation Scholarship for the year 2012 to support her PhD thesis work. The authors would like to thank David Garrett, Summer student in the PLAN Group, for his contribution to the experiments.

REFERENCES

[1] S. Beauregard, "Omnidirectional pedestrian navigation for first responders," in Proceedings of the 4th Workshop on Positioning, Navigation and Communication 2007 (WPNC'07) , 22 March, pp. 33-36, IEEE, 2007.

[2] J. Collin, Investigations of self-contained sensors for personal navigation, Doctoral Dissertation, Tampere University of Technology, 2006.

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

[4] Widyawan, G. Pirkl, D. Munaretto, C. Fischer, C. An, P. Lukowicz, M. Klepal, A. Timm-Giel, J. Widmer, D. Pesch, H. Gellersen, "Virtual lifeline: multimodal sensor data fusion for robust navigation in unknown environments", Pervasive and Mobile Computing, vol 8, pp. 388-401, 2012.

[5] C.V.C. Bouten, K. T. M. Koekkoek, M. Verduin, R. Kodde, and J.D. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity", IEEE Transactions on Biomedical Engineering, vol 44, no 3, pp. 136-147, 1997.

[6] P. D. Groves, Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems, 2008, Artech House.

[7] D.M. Prahl, and M.J. Veth, "Coupling Vanishing Point Tracking with Inertial Navigation to Produce Drift-Free Attitude Estimates in a Structured Environment" in Proceedings of ION GNSS 2011, 19-23

September, Portland OR, pp. 3571-3581, U.S. Institute of Navigation, 2011.

[8] C. Kessler, C. Ascher, N. Frietsch, M. Weinmann and G. Trommer, **"**Vision-based attitude estimation for indoor navigation using vanishing points and lines", IEEE ION/PLANS 2010, pp.310-318, 2010.

[9] J.M. Coughlan and A.L. Yuille, "Manhattan World: Compass direction from a single image by Bayesian interference", in Proceeding of 7th IEEE International Conference on Computer Vision ICCV, September 20-27, 1999, Kerkyra, Greece, 1999.

[10] A.C. Gallagher, "Using vanishing points to correct camera rotation in images", in 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, 2005.

[11] L. Ruotsalainen, J. Bancroft, H. Kuusniemi, G. Lachapelle and R. Chen, "Utilizing Visual Measurements for Obtaining Robust Attitude and Positioning for Pedestrians", ION GNSS 2012, in press.

[12] F. Alizahed-Shabdiz and M. Heidari, "Method of determining position in a hybrid positioning system using a dilution of precision metric", US Patent application 2011/0080317, 19 pages, 2009.

[13] L. Ruotsalainen, J. Bancroft, G. Lachapelle, H. Kuusniemi and R. Chen, "Effect of Camera Characteristics on the Accuracy of a Visual Gyroscope for Indoor Pedestrian Navigation", in Proceedings of the UPINLBS 2012, IEEE, in press.

[14] R. Hartley, A. Zissermann, Multiple view geometry in computer vision, second edition. Cambridge University Press, 2003

[15] L. Ma, Y.Q. Chen, and K.L. Moore, "Analytical piecewise radial distortion model for precision camera calibration", Proceedings in Vision, Image and Signal Processing, vol 153 (4), IEEE, pp. 468-474, 2006.

[16] J. Kosecka and W. Zhang, "Video compass", in Proceedings of European Conference on Computer Vision, Springer Verlag, pp. 657-673, 2002.

[17] J-Y. Bouguet, *Camera Calibration Toolbox for Matlab*, 2010, last accessed July 11, 2012.

[18] J. Bancroft and G. Lachapelle, "Estimating MEMS gyroscope g-sensitivity errors in foor mounted navigation", in Proceedings of the UPINLBS 2012, IEEE, in press.

[19] J. Bancroft, M.H. Afzal, and G. Lachapelle, "High performance GNSS augmented pedestrian navigation in signal degraded environments", International Global Navigation Satellite Systems Society Symposium, 15-17 November 2011, University of New South Wales, Sydney, Australia, 14 pages, 2011.

[20] H. Kuusniemi, "User-level reliability and quality monitoring in satellite-based personal navigation", Ph.D. dissertation, Tampere University of Technology, Tampere, Finland, 2005.

[21] Analog Devices, "*ADIS16375 Data Sheet (Rev. A)*," pp. 28, 2010.

[22] GoPro's web pages http://gopro.com/cameras/hd-helmet-hero-camera/, last accessed June 15, 2012.