Inertial Navigation Systems for User - Centric Indoor Applications

Alberto Serra[3], Tiziana Dessì[1,2], Davide Carboni[3], Vlad Popescu[2], Luigi Atzori[2]

[1] GeoWeb Laboratory, Sardegna Ricerche, Pula, Italy; [2]University of Cagliari, Cagliari, Italy; [3]CRS4, Pula, Italy

E-mail: alserra@crs4.it, tiziana.dessi@diee.unica.it, dcarboni@crs4.it, vlad.popescu@diee.unica.it, l.atzori@diee.unica.it

Abstract: **Inertial navigation systems (INS) are aimed at supporting the navigation of objects and/or users making use of a computer equipped with motion (accelerometers) and rotation (gyroscopes) sensors to continuously calculate their motion and position. These are quite useful when the GPS signal is not available, such as in indoor scenarios. Thanks to the exponential growth in the number of smartphones equipped with accelerometers and gyroscopes, more and more users in the consumer market can benefit of the INS installed in their handsets, which can leverage the deployment of location - aware applications, from the navigation of public indoor buildings to buddy - mapping social networking.**

This paper describes the architecture and main components of an INS deployed on common smartphones. Furthermore provides some interesting experimental results that highlight the limits of the technologies currently available in the consumer market. We then provide some crucial mechanisms that are needed to improve the accuracy of the system, such as the recalibration of the system with geo-referenced points.

Keywords: Pedestrian navigation, personal positioning, indoor navigation, plane homography, Hough transform.

1 INTRODUCTION

In recent years, the mobile telephony market has been characterized by an exponential growth of computational power and sensing capabilities of the handsets, currently referred to with smartphone, equipped with embedded motion (accelerometers) and rotation (gyroscopes) sensors, Internet connection and high-resolution cameras. All the sensing and computing technologies available in a common smartphone makes it ideal for INS (Inertial Navigation System) applications aiming at supporting the navigation of objects and/or users in an indoor environment where common localization systems, such as GPS (Global Positioning System), fail due to severely attenuation or obscuration of the satellite's signal. In fact, the GPS solution is suitable only if at least three satellites are in the line of sight, in the other cases, such as indoor application or urban "canyon", we need to use alternative positioning techniques.

In inertial navigation systems, localization/orientation estimation is source-independent. The user's position is calculated in relation to a known starting position using a dead reckoning algorithm. The whole system makes use of the previously mentioned sensors: accelerometers are used to calculate the distance travelled and the gyroscopes/magnetic compass to determine the direction. The uncertainty in the estimated position grows with time from the initial known starting position since the errors introduced by estimating the user/object movements are additive, increasing the total inaccuracy. This demands for a periodic recalibration of the system to reduce the cumulative error.

The present paper focuses on this problem by proposing an early development of an Indoor Navigation System based solely on the capabilities of a typical modern smartphone equipped with accelerometers, compass, camera and Internet connectivity. The user initially takes a photo of a geo-referenced 2D-bar code in order to acquire the map of the building and the initial position. The system then estimates the movement calculating the number of user's steps from the starting point using the accelerometers and the direction using the compass. The considered scenario includes the presence of geo-referenced 2D-tags placed in some known, key positions of the site to be visited. By taking a photograph of the tags, the system is able to initialize and subsequently re-calibrate the location data. To improve the calibration accuracy, we have focused on the problem of computing the exact position of the user (based on the known position of the tag) in terms of orientation and distance from the reference point using plane homography and affine transformation [3]. This allows us to correct perspective and projective distortion from the taken photo and derive information about the viewing angle (the user's orientation) and distance between camera and object.

Inertial Navigation System are used in many different kind of application involving moving objects, including vehicles, such as for example aircraft and submarines for navigation purpose. Recently, some research has proposed its use as assistive mobility technology for people with some kind of disabilities. Furthermore, indoor navigation can support commercial activities such as the retrieval of products in a large mall, but can also be deployed for security reasons: evacuation of complex buildings, route identification for visitors etc.

Corresponding author: Tiziana Dessì, University of Cagliari, piazza d'Armi, +39-0706755903, tiziana.dessi@diee.unica.it

The paper is organized as follows: in the next section we describe the background and related work in the field of indoor navigation systems. The third section introduces the developed prototype and presents the preliminary test results. Then, in the fourth section an overview of the plane homography techniques applied to the indoor navigation problem is presented. Finally, in the fifth and last section, we draw conclusions and present the planned future work.

2 RELATED WORK

In [1], a method of personal positioning for a wearable Augmented Reality System is proposed, allowing a user to freely move around indoor and outdoor. In this system, users are equipped with a communication device with built-in sensors, a wearable camera, an inertial head tracker and display. The method uses the dead reckoning process to detect and measure a unit cycle of walking locomotion and direction achieved by analyzing the acceleration vector, angular velocity and magnetic vector acquired from built-in sensors (in this work the 3DM-G from MicroStrain Inc is used).

The German Aerospace Centre studies a sensor fusion approaches that combines GNSS (Global Navigation Satellite System), foot mounted inertial sensors, electronic compasses, baro - altimeters, maps and active RFID tags. The system consists of a two-layer sensor fusion architecture that operates with a Kalman filter where possible, and fuses other sensors and maps at a higher-level, lower rate, particle filter. In buildings, a few dispersed RFID tags can significantly improve the overall performance of the positioning system [2].

In [3], the authors present a method aimed at removing geometric distortions in taken images of rectangular planar patterns (such as maps, pictures, or posters) without any preliminary knowledge of the calibration information or explicit knowledge of the imaging device. They use plane homography to transform the pictured rectangular object (taken image) into a rectangular object (at some distance) in an object plane parallel to the image plane. Then lens distortion is removed using Hough transform (for line detection) and LSE (Least Square Error) for approximation of the curved lines into straight lines.

In [4] a wayfinding prototype system based on QR codes and social computing for individuals with cognitive impairments is presented. It is based on psychological models of spatial navigation. Thanks to the camera embedded in a common PDA (Personal Digital Assistant), the user can take photos showing directions when he reaches a position on a planned trip. Every such photo is triggered by a geo-coded QR code image. A tracking function is also integrated in the system, allowing recording the person’s ID, timestamping the visited position, elapsed time after leaving the last position and expected arrival time to the next position, in order to activate the support team or family members in case of anomalies.

In [5] the authors present a prototype that allows indoor navigation without any external infrastructure making use only of pictures of indoor "You-Are-Here" maps, built-in cell-phone sensors like the accelerometer as well as compass data and compass deviation for a dead reckoning approach. To make the picture of the indoor map usable for navigation, a reference-sequence is needed to adopt it to the users physiognomy (i.e. the length of legs and size of foot). They do not estimate a step length, but simply assume that user keeps the same step length most of the time. The reference points in the indoor scenario have to be identified by the user on the YAH-map.

A low-cost indoor navigation system running on off-the-shelf camera phones is presented in [6]. The proposed system uses built-in cameras to determine user location in real time by detecting unobtrusive fiduciary markers. The approach allows the continuous scanning of an environment in real time (15 Hz or more) in search of navigation hints. Thus, navigation scales from sparse, strategically placed fiduciary markers to continuous navigation in 3D. The required infrastructure is limited to paper markers (square markers or frame markers) and static digital maps.

Another indoor positioning and localization system that does not require any additional infrastructure in the environment and provides 3D positioning and orientation data is proposed by Köhler *et al* in [7]. The mentioned solution uses a camera to locate and track a grid pattern projected onto surfaces in the camera’s field of view to determine its distance and orientation and uses that information to calculate its 3D position and pose in the room.

3 INS - ARCHITECTURE IMPLEMENTATION ON A MODERN SMARTPHONE

Differently from the systems presented in the previous section, our solution is solely based on the capabilities of a common modern smartphone. The data read from the phone’s sensors, combined with a reference map of the place and a known starting point, gives the actual position of the user. Hence, there is no need to connect to any external or pre-installed positioning system such as GPS, RFID, or to use WiFi trilateration; solely the dead reckoning technique is used instead. Dead reckoning is the process of estimating the current position of an user based upon a previously known position, upgrading this position upon measured or estimated speeds over elapsed time and course.

The prototype of the proposed system uses the data from the motion sensors embedded in the smartphone to compute the correct position of the user based on a known initial location. The smartphone application, still under development, is presented in Figure 1.

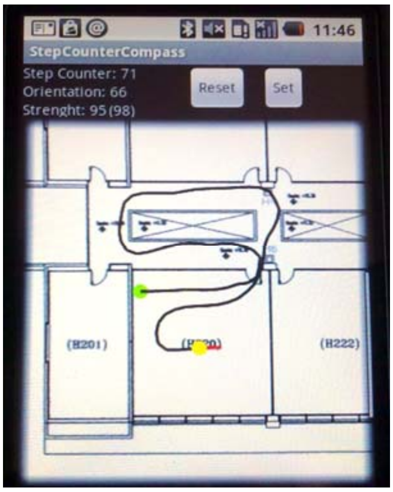

Figure 1. Screen of the application with a pedestrian route example.

The initial position of the user, the only certain information on which the system relies on for further calculation, is retrieved using the integrated photo camera of the smartphone scanning and decoding a datamatrix (2D barcode) placed aside the map of the floor (see figure 2)[8].

Based on the URL encoded in the datamatrix, the application downloads from a dedicated server the indoor vector map for the specific floor together with the initial position of the user on the map (corresponding to the point where the user stands when scanning the datamatrix).

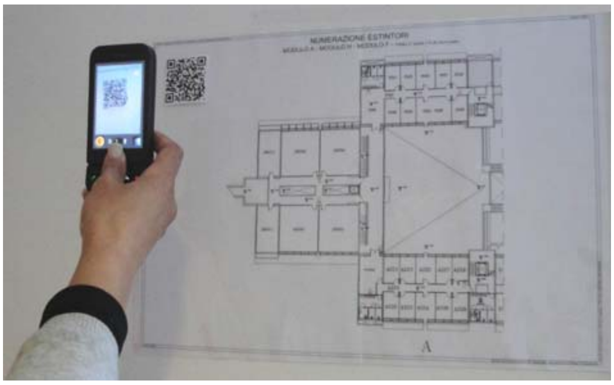

Figure 2. User reads a 2D datamatrix to download the map and his starting position.

When the user starts walking, the application draws step by step the position of the user, as a continuous line, over the downloaded map of the building floor.

The application tracks the number of steps taken by the user based on the numerical values returned by the smartphone's accelerometers. The accelerometers can be polled to return three values, one for each of the x, y and z-axes. We are using all three axes in order to have a combined value of the acceleration valid for any possible position of the smartphone.

The combined acceleration is the module of the three acceleration values. One step is detected when the value of the module is above a high threshold (*Th_high)* and successively is below a *Th_low* value. The following figure is an example graph of the accelerometer's measurement for a step length of 70cm, having the absolute values $Th\_high = 109$ and $Th\_low = 97$.

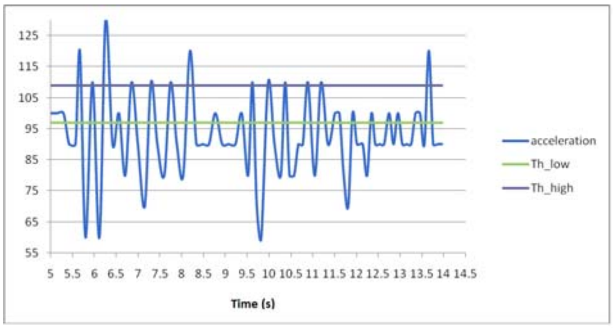

Figure 3. Normalized Acceleration

The current orientation of the user is measured by the smartphone's digital compass (the parameter 'Orientation' in Figure 1). The initial orientation, in this implementation, is set when the user scans the 2D barcode, being perpendicular (within a certain angle) to the floor plan. The relative position of the device with respect to the user (e.g. in a pocket) does not influence the dead reckoning estimation. If the device is held in front of the user, the magnetic compass provides the step-by-step heading improving the overall accuracy of the positioning method. This approach does not consider the user's distance from the object, therefore it's not too accurate. This problem is being addressed in our research making use of plane homography techniques to estimate the relative position of the user with respect to the geo-referenced tags, as it is explained in Section 4.

3.1 First experimental results

Before starting the application, the compass needs an accurate recalibration. This recalibration is necessary because the compass is subject to several errors: initially it has an inaccuracy of maximum 5 degrees, depending also on the used device and on the presence of electromagnetic interferences.

The step counter module based on the accelerometer data was tested and validated after thorough tests, performed in an indoor environment using both men and women with different physical features. The mean placement error was 3,8% on a series of 20 runs consisting of an average step count of 40 steps.

4 CALIBRATION OF THE INS

Based on the first experimental results presented in the previous paragraph, we identified the need to adjust the data returned by the smartphone's sensors and to improve the acquisition of the user initial position.

Therefore, in this section we present the technique that we are investigating for improving the calibration of the system while acquiring the image of the 2D tags mentioned before. The problem we are focusing on is to determine the user's position with reference to the 2D tags in terms of the distance and orientation angle.

To this we are making use of plane homography techniques: in a central projection camera model, a three-dimensional point in space is projected onto the image plane by means of straight visual rays from the point in space to the optical centre. Mathematically this process can be described using a 3x4 projection matrix P, which takes a point in 3-D space in homogeneous coordinates $(X, Y, Z, 1)^T$ and transforms it into a point on the 2-D image plane $(x, y, 1)^T$.

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = [P] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \qquad (1)$$

The projection matrix P can be computed from the internal and external camera parameters:

$$P = K \; [R \mid T] \qquad (2)$$

where K is a 3 x 3 upper triangular matrix, called the camera calibration matrix, including the intrinsic camera parameters (focal length, aspect ratio and skew) and $[R \mid T]$ defines the Euclidean transformation between camera and world coordinates (in general rotations followed by translations), including the external camera parameters, i.e. its position and orientation. In the case where planar surfaces are imaged ($Z = 0$), the transformation is called a plane-to-plane homography:

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = [H] \begin{pmatrix} X \\ Y \\ Z = 0 \\ 1 \end{pmatrix} \qquad (3)$$

The 3x3 transformation matrix, usually called the homography matrix H, has a simpler form than P, but it can be also reduce to:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = [K] \; [\mathbf{r_1} \quad \mathbf{r_2} \quad \mathbf{t}] \qquad (4)$$

where K is the camera's matrix, $\mathbf{r}_1$ and $\mathbf{r_2}$ are the correspondent columns of the rotation matrix R and $\mathbf{t} = -RC$ with C the camera center.

For this particular case we are dealing with the acquisition of a planar surface. Figure 4 shows the mapping between a 2-D point x' in the object plane π' where the tag lies into a 2-D point x in the image plane π that represents the image acquired by the camera.

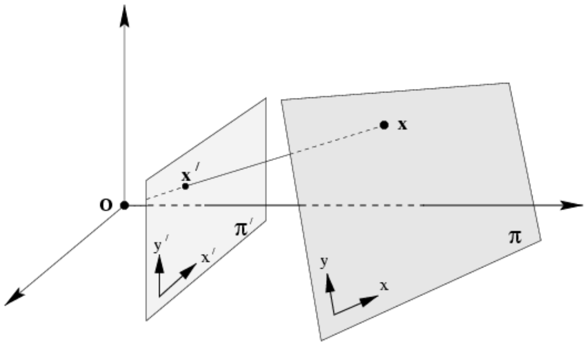

Figure 4. Mapping between planes

This process can be described mathematically by a homography matrix H:

$$P_i^{'} = HP_i \qquad (5)$$

where P and P' are 3 x 1 vectors that could correspond to the images of the same points, the former in the plane of the tag and the latter in the plane of the image, while H is the transformation matrix.

If the homography between a plane in the scene and the plane of the image is known, then the image of the planar surface can be rectified into a front-on view. Given four points on the scene plane, with no more than any 2 points collinear, and their corresponding positions in the image (8 equations), H is uniquely determined. Let $P_1^{'}(x'_1, y'_1)$, $P_2^{'}(x'_2, y'_2)$, $P_3^{'}(x'_3, y'_3)$ and $P_4^{'}(x'_4, y'_4)$ be the four corner points of the rectangular object and $P_1\ (x_1, y_1)$, $P_2\ (x_2, y_2)$, $P_3\ (x_3, y_3)$ and $P_4\ (x_4, y_4)$ their projections obtained using a plane homography transformation.

Corresponding points in two images related by homography are then:

$$x_i' = \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}}$$

$$y_i' = \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \tag{6}$$

Our objective is to remove affine and projective components in order to obtain a similarity transformed such for example a rotated, scaled and/or translated version of the original image.

To do this, we firstly have to identify the four corners in the taken photo, then, using plane homography transformation, map each vertex of the quadrilateral to the corresponding vertex in the known rectangle. Using equations (6) we can find the coefficients of the homography matrix H and finally rectify the image to the frontal view. Once the frontal view is recovered from the knowledge of the calculated homography's matrix coefficients, we can decompose the H matrix, using QR Factorization, in its orthogonal ($[\mathbf{r_1} \quad \mathbf{r_2} \quad \mathbf{t}]$) and upper triangular matrix $[K]$.

From the knowledge of the orthogonal matrix we can determine the tilt angle ϕ (rotation around the x-axis), the roll angle ψ (rotation around the y-axis), the pan angle θ (rotation around the z-axis) and the translation along the three axes, thus the orientation of the user/camera in the scene. Given the known values for the focal length *f*, the aperture of the camera embedded in the smartphone, the real dimensions of the 2D datamatrix (*H*) and the corresponding size in pixel (*H'*), we can calculate the user's distance *l* from the datamatrix. The geometric set-up is presented in figure 5 in a simplified version, taking into consideration only the height of the object.

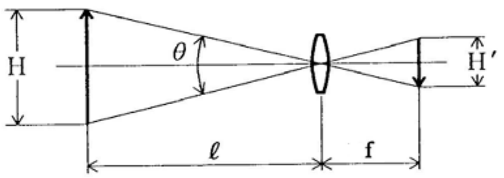

Figure 5. Simplified geometric scheme for distance calculation

Using the results of the calculus (angles and distance), the initial position of the user and his orientation relative to the scanned 2D datamatrix can be more precisely computed, as being of crucial importance for the correct evaluation of the data subsequently generated by the smartphone's sensors, especially in terms of orientation.

5 CONCLUSIONS

In this paper we proposed a new approach and implementation for a pedestrian indoor navigation system. We developed this application on a modern Smartphone and did first experiments in a real indoor environment, measuring the encountered errors. We are investigating the application of the plane homography technique to the indoor navigation problem in order to derive additional information about the relative orientation and distance of the user to the reference point. Based on this supplementary data we are trying to reduce the inherent errors of the dead-reckoning technique implemented in the first prototype.

Future work will include the implementation of the above-mentioned solution together with the improvement of the measurement method of the walking steps to overcome the shortcomings of the current-used fixed-value step length. The estimation of the step length could be obtained by the strength of the step acceleration movement (through a probabilistic algorithm) and the personal information data written previously by the user. The employment of a particle filter or a Kalman algorithm [9] for fixing the accumulating error of the step length will be also subject of a further study. The knowledge of walls, doors, pillars and other elements can be also used for fixing the position errors if these elements are already included as vector data in the floor map, which in this case could be an SVG image.

References

[1] Kougori, M. and Kurata, T. 2003. Personal Positioning based on walking locomotion Analysis with self-contained sensor and a wearable camera. In Proceedings of ISMAR2003, pp. 103-112.

[2] Krach, B. and Robertson, P. 2008. Integration of Foot-Mounted Inertial Sensors into a Bayesian Location Estimation Framework. In Proceedings of 5th Workshop on Positioning, Navigation and Communication 2008 (WPNC 2008), Hannover, Germany.

[3] Wang, X., Klette, R. and Rosenhahn, B. (2005) Geometric and photometric correction of projected rectangular picture', Image and Vision Computing New Zealand, November 2005.

[4] Chang Y.J., Tsai S. K., Wang T. Y and Chou L. D. A novel wayfinding system based on geo-coded QR codes for individuals with cognitive impairments. In Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility, Arizona, USA, 2007, 231 – 232.

[5] Gehring S., Löchtefeld M., Schöning J. and KrügerA. Exploring the Usage of an Electronic Compass for Human Navigation in Extreme Environments. Haptimap 2010: Multimodal Location Based Techniques for Extreme Navigation, In conjunction with Pervasive 2010, Helsinki, Finland, 2010.

[6] Mulloni A., Wagner D., Barakonyi I., Schmalstieg D., Indoor Positioning and Navigation with Camera Phones, IEEE Pervasive Computing, vol. 8, no. 2, pp. 22-31, Apr.-June 2009

[7] M. Kohler et al., Tracksense: Infrastructure Free Precise Indoor Positioning Using Projected Patterns, Pervasive Computing, LNCS 4480, Springer, 2007, pp. 334–350.

[8] http://code.google.com/p/zxing/

[9] Woodman, O. and Harle, R. 2008. Pedestrian Localisation for Indoor Environments. In Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp), Seoul, Korea, ACM 2008.