Indoor Pedestrian Localization With a Smartphone: A Comparison of Inertial and Vision-Based Methods

Wael Elloumi, Abdelhakim Latoui, Raphaël Canals, Aladine Chetouani, and Sylvie Treuillet

***Abstract*— Indoor pedestrian navigation systems are increasingly needed in various types of applications. However, such systems are still face many challenges. In addition to being accurate, a pedestrian positioning system must be mobile, cheap, and lightweight. Many approaches have been explored. In this paper, we take the advantage of sensors integrated in a smartphone and their capabilities to develop and compare two low-cost, hands-free, and handheld indoor navigation systems. The first one relies on embedded vision (smartphone camera), while the second option is based on low-cost smartphone inertial sensors (magnetometer, accelerometer, and gyroscope) to provide a relative position of the pedestrian. The two associated algorithms are computationally lightweight, since their implementations take into account the restricted resources of the smartphone. In the experiment conducted, we evaluate and compare the accuracy and repeatability of the two positioning methods for different indoor paths. The results obtained demonstrate that the vision-based localization system outperforms the inertial sensor-based positioning system.**

***Index Terms*— Indoor pedestrian navigation assistance, vision, inertial sensors, smartphone.**

I. Introduction

DESPITE the growing interest of research related to the problem of indoor pedestrian navigation assistance, this topic remains a highly challenging issue. Outdoor positioning mainly relies on the global positioning system (GPS). However, this type of system cannot be exploited in indoor environments or even in urban areas where GPS signals are blocked or degraded by multipath effects. Alternative solutions are to use other navigation sensors such as wireless radio sensors, IMU (Inertial Measurement Unit) sensors and/or video camera sensors (vision-based navigation). However, these technologies face other challenges for indoor navigation such as accuracy, mobility, cost, weight and size. For a pedestrian navigation application, an indoor localization system should be accurate enough to guide a person. It is also desirable that it require no additional infrastructure that would restrict its applicability. Furthermore, it should be inexpensive, small in size and lightweight so as to be handheld.

Nowadays, smartphones have become ubiquitous devices that facilitate many aspects of our daily lives. They are equipped with various connection modes, tiny sensors (GPS, MEMS, vision sensors, etc.) that can be used for navigation, and their price makes them affordable.

Our goal was to develop an infrastructure-independent navigation system using a smartphone. A major contribution is the development of a vision-based localization system while reducing computational complexity. Our solution relies exclusively on the smartphone resources and provides instantaneous localization and orientation estimation (6 DoF). To overcome the drift problem (cumulative errors over time) revealed by our previous feasibility study [1], the pedestrian's position is obtained by matching the current image produced by the smartphone camera with geo-referenced images extracted from a database while their orientation is estimated thanks to the 3 orthogonal directions of the vanishing points (VP). The estimated position is then refined by using an Extended Kalman Filter (EKF). Another contribution of this paper is the development of a second localization system using the smartphone's inertial sensors. This system does not require any anchor or physical map information to provide indoor localization. Finally, a detailed study and a comparison of the two localization solutions was conducted. This assessment study is worthwhile since it enables us both to provide an outlook for the future and to find the relevant cooperation of the two technologies on the same system.

The paper is organized as follows. A review of related work is presented in Section II. Sections III and IV describe the two proposed localization approaches. Experimental results are shown in Section V, followed by the conclusion.

II. Related Work

Existing pedestrian navigation systems rely on several technologies which fall into two main categories. The first one requires augmenting the physical infrastructure with a network of fixed transmitters or tags and a receiver, such as GPS, Ultra-Wide Band (UWB), Wireless Local Area Networks (WLAN), Bluetooth, Radio-Frequency Identification (RFID), etc. The radio frequency signal generated is processed by the receiver in order to determine its position [2]–[7]. The main drawback

Manuscript received March 17, 2016; revised May 3, 2016; accepted May 3, 2016. Date of publication May 10, 2016; date of current version June 2, 2016. This work was supported by the Région Centre through the AZIMUT Project. The associate editor coordinating the review of this paper and approving it for publication was Prof. Danilo Demarchi.

W. Elloumi is with Worldline, Blois 41000, France (e-mail: wael.elloumi@worldline.com).

A. Latoui is with the Department of Electronics, Faculty of Science and Technology, Université Mohamed El Bachir El Ibrahimi, Bordj Bou Arréridj 34030, Algeria (e-mail: a.latoui@univ-bba.dz).

R. Canals, A. Chetouani, and S. Treuillet are with the Mechanic and Energetic Laboratory, Institute of Research in Systems Engineering, University of Orléans, Orléans 45067, France (e-mail: raphael.canals@univ-orleans.fr; aladine.chetouani@univ-orleans.fr; sylvie.treuillet@univ-orleans.fr).

Digital Object Identifier 10.1109/JSEN.2016.2565899

of the aforementioned navigation solutions is their dependency on pre-installed equipment which is relatively expensive and is not available in all network infrastructures.

The second category of navigation systems does not require any pre-installed infrastructure and mainly uses embedded sensors. These sensors are used to detect visual features (camera) or to estimate human motion (accelerometer, gyroscope, etc.) in order to determine the position of a given person. The associated methods fall into two broad groups depending on the type of sensors employed: camera sensor or inertial sensors.

Computer vision techniques are based on a camera sensor. The images captured by the camera are analyzed to detect some objects or features which are used as references to find the image position in the scene.

Real-time Structure From Motion (SFM) approaches, also called Simultaneous Localization And Mapping (SLAM), have been extensively used in robot navigation to estimate the motion of a robot in real-time. A map of the environment concerned is constructed by using the captured images [8]. Using feature correspondences detected on multiple frames, geometric constraints among the different views can be established and the motion of the camera can be estimated. This kind of method was applied for pedestrian localization in [9] and [10].

Offline SFM methods have also been proposed. Referred to as teach and replay or visual memory approaches, they work in two main phases: training and localization. Learning is performed offline and aims to create a map of the surrounding area (environment map building). The localization phase consists of exploiting the map thus obtained in order to localize the camera position on the map [11].

Along the same lines, recent approaches have shown that feature-based 6DoF pose estimation (position and orientation) can be done in real-time on mobile devices. Thus an approach for wide-area 6DoF pose estimation relying on a previously acquired 3D feature model (image-to-map matching) was presented in [12]. Other authors [13] use binary features combined with Locality Sensitive Hashing (LSH) to estimate the pose of the camera within a Parallel Tracking and Mapping (PTAM) algorithm. For both methods, the obtained pose estimate is further refined using bundle adjustment approaches.

However, SFM algorithms are often sensitive to degenerate camera motions. In addition, quick and abrupt movements, as well as environmental conditions in the case of pedestrian localization, are still challenging for these approaches.

On the other hand, an alternate group of visual systems uses preformed databases to determine the absolute pedestrian position. They mainly rely on feature recognition in images with referenced position [14]–[16]. Some descriptors are extracted from images and stored with their corresponding positions on the map. The absolute position is then given by matching the frame acquired by the pedestrian camera and images from the database. The challenging task remains the implementation of an efficient and fast comparison method to recognize already visited places despite changes in point of view and lighting. A scalable streaming approach for client-server-based visual localization can be found in [17]. To address the limited computational resources of mobile devices, the authors proposed a modified version of the state-of-the-art Vector of Locally Aggregated Descriptors image signature, which is based on binary features. The resulting vectors are binarized for rapid matching. Although this approach preemptively loads image signatures of reference images in the near vicinity of the user onto the mobile device to mitigate the effect of network latency, it still requires a network connection which is not always available. The indoor navigation system proposed by [18] uses a smartphone with a panoramic camera and a server architecture for image indexing. The authors proposed employing a multi-core GPU to parallelize and to speed up the localization step.

To improve their navigation solution, some authors proposed integrating vision in a multi-sensor navigation system [19]–[24]. These studies show that vision algorithms can significantly increase the localization accuracy and limit the drift of the other embedded technologies. A survey of optical indoor positioning systems was carried out by Mautz et al. in [25]. The study separates optical positioning approaches based on the way images are referenced to the environment. Vanishing points have also been exploited to estimate the pedestrian heading and aid vision-based localization [19], [23].

In line with these studies, we consider that vanishing points are relevant features to use inside buildings since images captured indoors generally contain large uniform areas (without texture) but the building's geometric structures (walls, doors, ceilings, windows, etc.) remain clearly visible in all lighting conditions [26].

The second broad group of methods concerns Dead-Reckoning (DR) techniques which are founded on inertial navigation and draw on MEMS technology, such as accelerometer (motion sensor), magnetometer (compass sensor), pedometer (stride sensor), and gyroscope (rotation sensor). Various approaches have been proposed in this area for Pedestrian DR (PDR), especially for people with visual impairments, in order to guide them to reach their destinations by using their smartphone. However, these methods are subject to cumulative errors which are essentially due to the material inaccuracies of low-cost smartphone-embedded sensors. In order to attenuate the effects of the bias of accelerometers and subsequent drift in velocity and position, a Zero Velocity Update (ZUPT) algorithm was introduced in [27]. However, the inertial sensor employed has to be mounted on the user's foot, thereby limiting its use to some specific pedestrian applications, such as finding and rescuing firefighters or other emergency first responders.

As an alternative, several step-based tracking methods have been proposed [28], [29]; they generally consist of two basic components, namely step detection and motion direction estimation. For step detection, most techniques reported in the literature process accelerometry data, and more specifically vertical acceleration, and rely on threshold-based algorithms. These techniques include peak detection [30]–[32] and zero-crossing detection [33]. The authors of [34] proposed an end-to-end localization system using inertial sensors on commodity phones. They designed a step detection algorithm

with heuristics constraints and dynamic time warping (DTW) validation to reduce false positives. Spectral analysis has also been proposed to detect steps [28]. For motion direction estimation, a digital compass can be employed. It should be noted that, to compensate for inaccuracies during step detection and heading estimation, some authors have used sequence alignment algorithms to match the number of detected steps with the expected path [35], [36]. Turn detection has also been exploited for path back-tracing using only data collected by iPhone sensors [37]. More recently, a smartphone-based pedestrian dead reckoning system, called SmartPDR, limited the increasing localization errors from low cost noisy inertial sensors by detecting the turning points along the walking path inside the building [38]. This improvement is efficient only if the experimental setup contains closed loops.

In order to improve accuracy and performance, hybrid localization techniques have been explored: Google recently launched its new project Tango [39] to supply Android mobile devices with the ability to navigate the physical world like humans thanks to inertial and vision sensors. However, this technology is still in development and uses specific mobile devices with sophisticated embedded sensors that are not available on commercialized smartphones.

To summarize this section, infrastructure-dependent indoor pedestrian navigation systems need costly installations and are not easy to deploy. In contrast, infrastructure-independent solutions do not require any pre-installed equipment. They fall into two groups of methods depending on the sensors used: DR approaches (inertial sensors) which are subject to cumulative errors, and vision-based solutions (camera sensor) that are time and energy consuming and sensitive to degenerate camera motions and abrupt movements.

In this paper, we investigate two technologies for indoor pedestrian localization taking advantage of smartphone sensors. The first one is a new vision-based localization method that has been implemented in a smartphone to provide real-time positioning. The originality of the method is to estimate the camera orientation from vanishing points to refine the position of the walker along a reference path defined by georeferenced key frames. An approximate but realistic position is estimated in real-time by clinging to the closest key frame and comparing the orientation of the camera in the current image with those determined in this key frame. Unlike image matching based localization systems that have high computational complexity [40], our solution uses a lightweight matching algorithm which represents a good compromise between computation time and robustness [41]. Moreover, it is not sensitive to degenerate camera motions as the pose of the camera (position and orientation) is estimated based on the position of the georeferenced key frames and the orientation of the three orthogonal vanishing points instead of 2D/3D feature matching.

The second positioning technology is premised on low-cost smartphone inertial sensors (magnetometer, accelerometer, and gyroscope). In order to measure the displacement of the pedestrian and to provide his successive relative positions, this approach consists of two stages: step detection and orientation estimation. Both positioning systems are handheld and inexpensive solutions. They require no special infrastructure and depend on cheap and lightweight sensors. Thus such localization systems may be helpful in various types of applications such as visually impaired guidance, emergency intervention, tourism, etc.

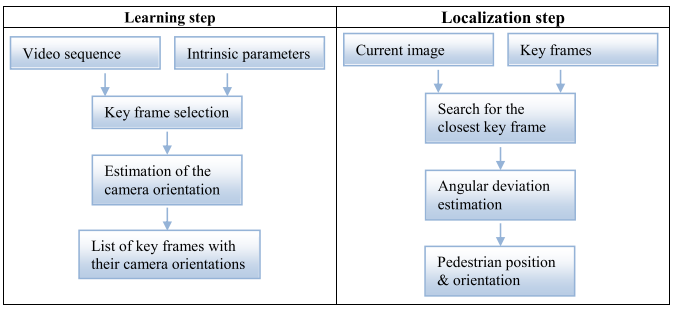


Fig. 1. Localization method based on vision sensor.

To highlight the advantages and drawbacks of each technology as well as to establish their prospects for the future, the present paper includes extended experiments comparing the two positioning technologies, each embedded in a smartphone, in a real world situation. The systems were evaluated along six different walking paths in a building representing a total distance of about one kilometer following corridors, crossing hallways and doors. The two positioning methods were synchronized with a commercial robotic localization system to provide a Ground Truth by walking. The results presented include a comparison in terms of accuracy and repeatability.

The remainder of the paper first presents our vision-based localization method and then that based on inertial sensors, as well as a comparison between them.

III. Vision-Based Localization

As depicted in Fig. 1, our proposed method is based on two main stages: an offline learning phase in which a reference path is recorded by selecting key frames along the way and by computing the camera orientation in these frames. Then, in a localization phase, an approximate but realistic position of the walker is estimated in real-time. This estimation is done by comparing the orientation of the camera in the current image and those determined in the key frames to assist the pedestrian with navigation guidance.

For experimental data collection, we used the commercial Android Samsung Galaxy S III smartphone, which embeds a camera providing images of size 320×240 pixels at 30 frames per second.

A. Learning Phase

The learning algorithm is performed once for each path. It is performed offline and is mainly composed of key frames selection with camera orientation estimation.

1) Key Frame Selection: We first acquire a video sequence of the path. To obtain reference positioning, we synchronized the smartphone camera with a ground truth system that detects ceiling-mounted landmarks. More details about this system are provided in the experimental section V-A. Then the video sequence is processed offline and key frames are selected so

that they are spaced approximately one meter apart. Thus the positions of these key frames are identified to estimate the pedestrian's position in the localization phase.

2) Camera Orientation Estimation Based on Vanishing Points: Once the key frames have been selected, the orientation of the camera is estimated using 3 orthogonal directions associated with the vanishing points and available in man-made environments (also called Manhattan World [42]). More details about this approach can be found in our paper [26].

At the end of this phase, we obtain a visual memory of the travelled path composed of the list of the selected key frames and their camera orientations.

B. Localization Phase

The localization algorithm is performed in real-time. It takes the current image and the list of the key frames as inputs. The pedestrian position and orientation are obtained by applying the following three stages.

1) Search for the Closest Key Frame: This stage is based on image matching. It provides the approximate position of the walker along the reference path.

Harris-Based Matching: we implement image matching to compare the similarity between the current image (video stream provided by the smartphone camera) and the key frames. For this, we use Harris corners [43] as image features and the Zero-mean Normalized Cross Correlation (ZNCC) as descriptor. To have well distributed features, the images are divided into 15×15-pixel buckets. In each bucket, we detect as far as possible corners and select the best ones. These image points are then matched based on the cross correlation measurement to build lists of homologous pairs (one list for each pair of images). Let P_{i1} be a point detected in image I_1 and P_{j2} its corresponding point in image I_2. To match P_{i1} with P_{j2}, we define an 11×11-pixel search region centered on the coordinates of P_{i1} in image I_2. Next we compute a ZNCC score s for each detected point inside the defined search region as follows:

$$s = ZNCC(P_1, P_2) = \frac{\sum_{d \in V_{11}} (I_1(P_{i1}+d) - \bar{I}_1(P_{i1}))(I_2(P_{j2}+d) - \bar{I}_2(P_{j2}))}{\sqrt{\sum_{d \in V_{11}} (I_1(P_{i1}+d) - \bar{I}_1(P_{i1}))^2} \sqrt{\sum_{d \in V_{11}} (I_2(P_{j2}+d) - \bar{I}_2(P_{j2}))^2}} \quad (1)$$

where V_{11} is the 11×11-pixel neighborhood with $d \in V_{11}$ and the notation $\bar{I}$ indicates the average.

To satisfy the uniqueness constraint, we select the pair with the highest score (P_{i1}, P_{j2}, s) and remove all other pairs in the list containing one of the points P_{i1} or P_{j2}. The minimum threshold to select a pair of points is fixed at 0.8. This value was chosen for two reasons: (i) in our experiments, lower values generate numerous bad matches; (ii) higher values do not ensure sufficient numbers of matches necessary for good global matching. Inliers are detected by estimating the epipolar geometry for each pair of images. Finally the key frame with the highest number of inliers is considered as the closest to the current image (Fig. 2). This matching method gives the best compromise between computation time and robustness [41].

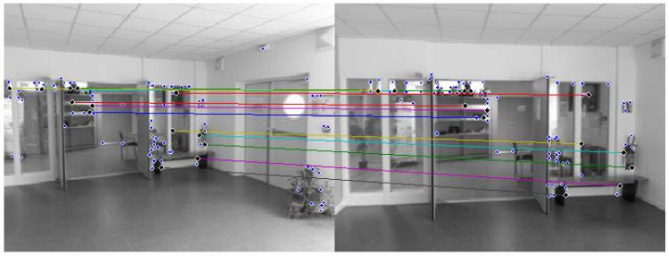


Fig. 2. Image matching.

Note that during the process initialization, the first frame from the video stream is compared with all the key frames of the learning sequence so that we can determine an initial localization. For the following images, in order to reduce computation time, the search for the closest key frame is limited to a window of 4 images around the previous selected key frame (1 image before and 3 after).

Approximate Position Estimation: once the closest key frame has been found, the approximate position of the person along the reference path can be associated to the known position of this closest key frame. This position is then refined by an Extended Kalman Filter (EKF).

2) Pedestrian Position Estimation: To estimate the current user's location along the reference path, an EKF uses a dynamic system model to make a prediction of the system state in the next time increment. Then it compares the predicted and observed states using an observation model. The positioning equations for a pedestrian in the horizontal plane are:

$$x_k = x_{k-1} + \Delta t \dot{x}_{k-1} \quad (2)$$

$$y_k = y_{k-1} + \Delta t \dot{y}_{k-1} \quad (3)$$

$$\dot{x}_k = \dot{x}_{k-1} + \frac{\cos \beta_{k-1}}{\Delta t} \quad (4)$$

$$\dot{y}_k = \dot{y}_{k-1} + \frac{\sin \beta_{k-1}}{\Delta t} \quad (5)$$

$$\beta_k = \beta_{k-1} \quad (6)$$

where x_k, y_k are the coordinates of the key frame closest to the current image k; $\dot{x_k}, \dot{y_k}$ are the velocity components and β_k is the yaw angle of the camera which gives the pedestrian heading. The variable Δt represents the time between two successive images. The state vector $\boldsymbol{X_k}$ and the state transition matrix $\boldsymbol{A_k}$ are defined as:

$$X_k = (x_k \quad y_k \quad \dot{x_k} \quad \dot{y_k} \quad \beta_k)^T \quad (7)$$

$$A_k = \begin{bmatrix} 1 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & \frac{\cos \beta_{k-1}}{\Delta t} \\ 0 & 0 & 0 & 1 & \frac{\sin \beta_{k-1}}{\Delta t} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

At time k, the general observation model form is defined according to the information provided by our vision algorithm.

$$Z_k = H_k X_K + V_k \quad (9)$$

TABLE I

COMPUTATION TIME DISTRIBUTION FOR VISION-BASED LOCALIZATION ALGORITHM

Task	Computation time (ms)
Position estimation (stages 1 and 2)	986
Camera orientation estimation (stage 3)	90
Global localization time	1076

where Z_k is the observation (measurement) vector, H_k is the observation model which relates the true state space to the observed space, and V_k is the observation noise which is assumed to be zero-mean Gaussian white noise, $V_k \sim N(0, R_k)$, with the covariance matrix R_k. The full-scale measurement vector is:

$$Z_k = (x_{CAM} \quad y_{CAM} \quad \beta_{CAM})^T \tag{10}$$

where x_{CAM}, y_{CAM} and β_{CAM} are respectively the position and the heading of the user provided by the embedded camera.

The overall process to localize an image of size 320×240 takes about 1 second with non-optimized code. The distribution of the computation time is shown in Table I.

3) Angular Deviation Estimation: This stage consists of computing the angular deviation of the camera attitude (yaw angle) between the current image and the closest key frame. For this, we first estimate the camera orientation of the current image using the 3 orthogonal VPs, as presented in section III-A-2. Then we compare its attitude to that of the closest key frame by:

$$\gamma_i = \beta_{c_i} - \beta_k \tag{11}$$

where β_{c_i} is the yaw angle of the closest key frame and β_k is the yaw angle of the current image. This angle enables guidance of the person by comparing it to a deviation threshold T from the reference path. In our experiments the value of T is set empirically to 15°. So when $|\gamma_i| > 15°$, we provide the user with guidance instructions to avoid him straying from the reference path.

IV. INERTIAL SENSOR-BASED LOCALIZATION

Inertial sensor-based localization algorithms can be divided into two main categories: integration methods and step counting methods [44] (see section II-B). The integration scheme is based on: (i) angular velocity data integration, collected from a gyroscope, to find the orientation; (ii) position is estimated by integrating acceleration collected from the accelerometer. In fact, the Zero Velocity Update (ZUPT) algorithm is widely used to decrease the accumulating error from double integration for a pedestrian tracking system. However, the sensor needs to be on the shoe, which causes inconvenience to the user and restricts its use to certain special cases. Moreover, due to the drift error of inertial sensors in a smartphone, the double integration of the small acceleration error produces a progressively larger position error that cubically grows in time in the short term [27]. Thus, counting every step of the user and calculating the length of every step as precisely as possible is a more practical approach to estimate the position of the user using only a smartphone and taking advantage of the inertial sensors integrated within it. So, in our work, we count steps with a 3-axis accelerometer and determine heading with a 3-axis digital compass. In other words, our approach consists of two basic components, called step detection and motion direction estimation. An EKF is then used to estimate the position of the user. These issues are discussed in more detail below.

TABLE II

COMPARISON BETWEEN THE NUMBER OF COUNTED STEPS AND THE NUMBER FOUND BY THE STEP DETECTION ALGORITHM

Walked path	Distance travelled (m)	Actual number of steps	Estimated number of steps
Walk1	35	63	61
Walk2	35	61	60
Walk3	58	93	90
Walk4	58	95	94
Walk5	57	98	94
Walk6	59	95	94

A. Step Detection

In general, implementing a practical step-based PDR system includes several basic stages [29]: noise filtering, step detection, stride estimation, and heading determination. As previously, for experimental data collection, we used a Samsung Galaxy S III smartphone which integrates the 2 types of sensors mentioned earlier.

As noted in [45], there are many effective methods, such as peak detection, zero-crossing, autocorrelation and stance-phase detection, which can be used to detect step occurrences. In [46], the authors empirically found that peaks in the variance of acceleration in the global Z-axis were a robust feature to detect the steps of a given leg. A similar approach was employed in this work. We use the acceleration measurements (Fig. 3a) which have been projected into the vertical direction in order to detect steps. The data are then processed with a sliding window for smoothing (Fig. 3b). The smoothed results are processed for step detection through peak detection (Fig. 3c). It is worth noting that small peaks observed in the acceleration signal are caused by residual sensor noise and the complexities of human walking dynamics. To overcome these effects in step counting, a relative threshold detection scheme was employed. Table II shows a comparison between the counted number of steps (Ground Truth) and that found by the step detection algorithm for 6 different walks. Several tests were conducted repeating each walk three times with a total length of 855m and 1516 counted steps; we found that the step detection errors were about 3.5% of the total counted steps. Therefore, mounting the smartphone on a body harness has an insignificant influence on step detection (Fig. 6).

B. Heading Determination

In this study, heading was directly measured with the digital compass of the smartphone. According to the smartphone Software Development Kit (SDK) documents [47], the built-in compass gives the azimuth of the device in degrees from the magnetic north, in a clockwise direction with respect to the X-axis. Fig. 4 shows an example of azimuth angles collected in our smartphone which is held in the hand and facing the user.

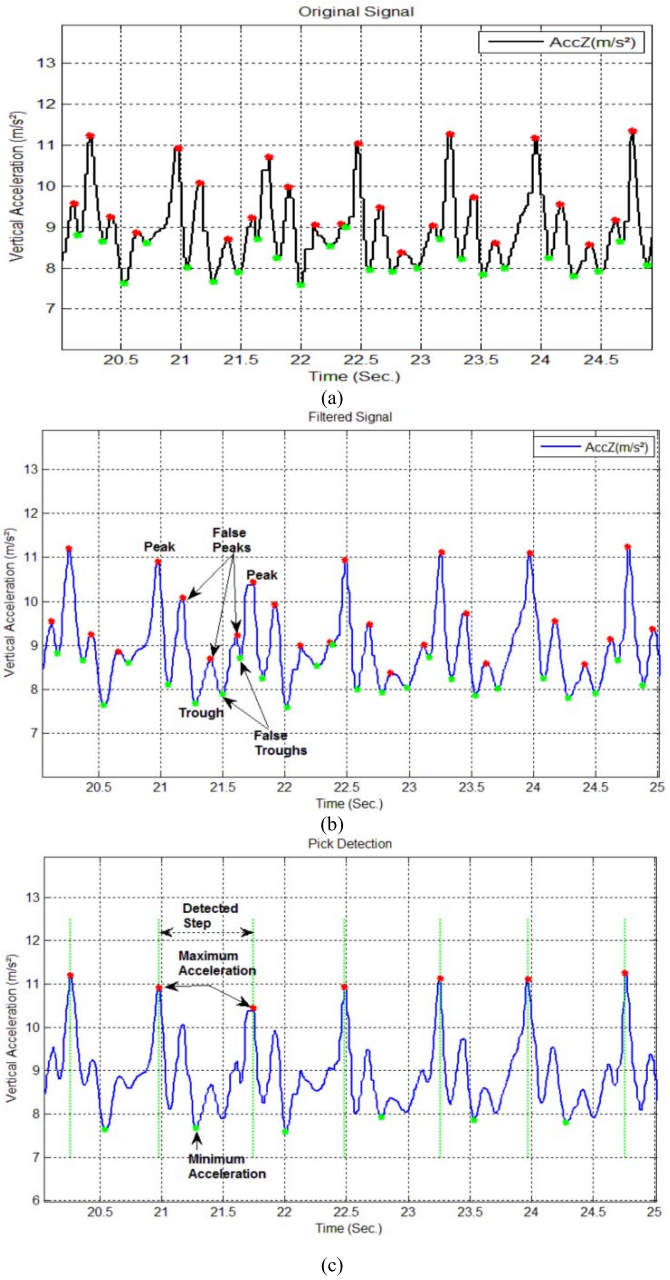


Fig. 3. Step detection based on acceleration variations in the vertical direction. (a) Original signal (Z-axis) obtained from the Samsung Galaxy S III smartphone (b) Filtered signal. (c) Step detection using peak detection method.

C. Heading Pedestrian Position Estimation

As in the vision-based localization algorithm, we used an EKF to estimate the pedestrian position. The state X_k in (12) describes the current position and heading at time k, where x_k, y_k are the east and north coordinate components in the locally horizontal (East-North-Up, ENU) plane and β_k is the heading.

$$X_k = \begin{pmatrix} x_k \\ y_k \\ \beta_k \end{pmatrix} \tag{12}$$

When a step is detected at time k, the concurrent derived heading is sent as an input to the filter. The update of the

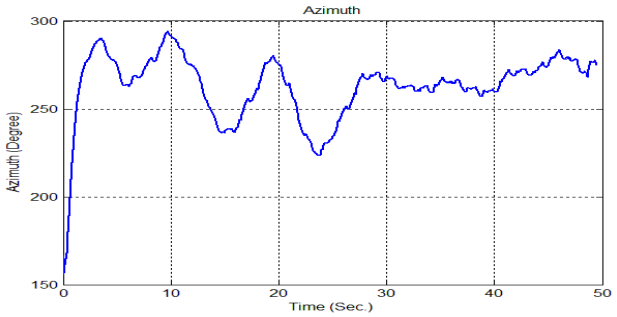


Fig. 4. Azimuth angle extracted from sensor orientation of our smartphone while walking.

a priori estimate state $\hat{X}_k^-$ is given by:

$$\hat{X}_k^- = \begin{pmatrix} x_{k-1} \\ y_{k-1} \\ \beta_{k-1} \end{pmatrix} + \begin{pmatrix} \cos(\beta_{k-1}) \\ \sin(\beta_{k-1}) \\ 0 \end{pmatrix} u_{k-1} \tag{13}$$

where u_{k-1} is the control signal corresponding to the step length. This step length can be selected variable or fixed [30], [48]. A variable value takes into account the normal step length unique to each person, but also the speed variation (straight or curved) deduced from the accelerometer signal frequency. This requires supplementary processing to estimate the instantaneous step length of a person and represents an additional computational cost. Moreover it produces good results only if the inertial sensor is attached to the ankle or foot. It also appears that for normal walking, stride length varies slightly from one person to another and for the same person (e.g. there is a slight difference between walking along straight and curved paths). A fixed value of the step length is then a good approximation of the average value of this length [46], [49]. This is the option which was selected: after some tests, the step length S_l was estimated at 65 cm.

In general, step length is not as critical as the determination of heading [44] since the accuracy of the compass may be affected by any nearby ferromagnetic materials such as doors, elevators, and any other iron object. The compass must be calibrated to compensate for errors mainly due to magnetic declination and hard / soft iron effects from ferrous materials in the platform. In this paper, calibration was performed as described in [50].

At this stage, we can calculate the Jacobian matrix F_k, which is necessary for the covariance propagation, as follows (Eq.14):

$$F_k = \begin{pmatrix} 1 & 0 & -\sin(\beta_{k-1} - 1).Sl \\ 0 & 1 & \cos(\beta_{k-1} - 1).Sl \\ 0 & 0 & 1 \end{pmatrix} \tag{14}$$

Since the time between detected steps is not taken into account, the process noise Q_k is a constant value ($I_{3\times3}$ matrix) and is carried out in advance. Then the estimated error covariance P_k^- is calculated as:

$$P_k^- = F_k P_{k-1} F_{k-1}^T + Q_k \tag{15}$$

However, the measurement input, in our case, is considered as the 1×3 matrix H_k such that $H_k = (001)$ since the filter has

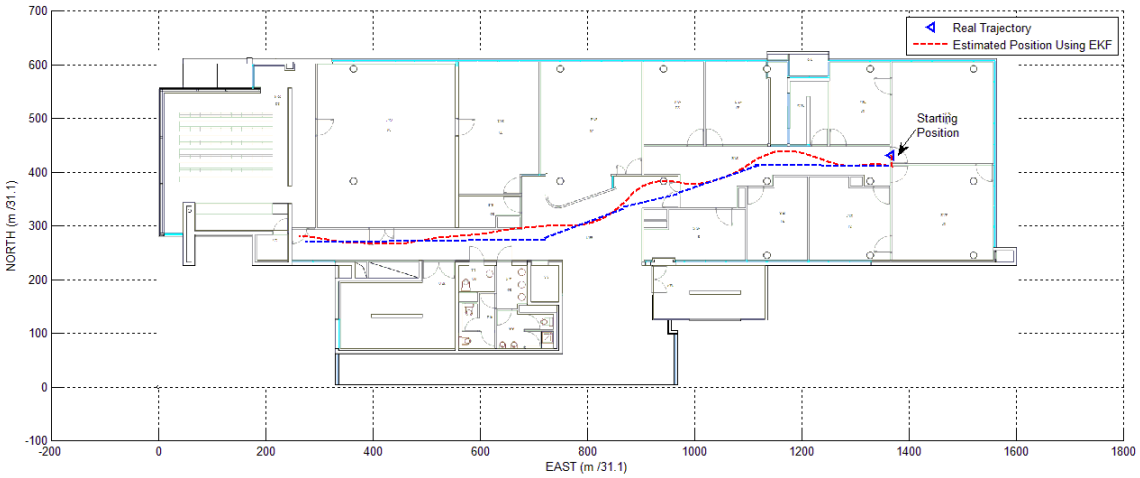


Fig. 5. Plot of user displacement inside Polytech Orléans building estimated by EKF.

Fig. 6. Our pedestrian system.

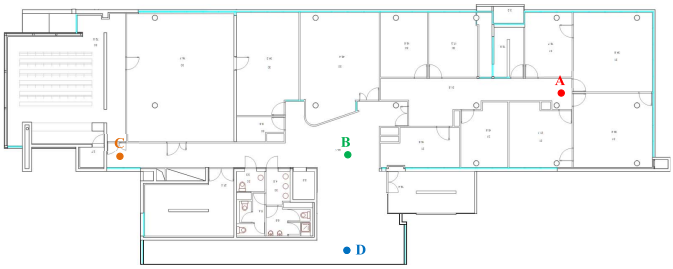


Fig. 7. The map of the third floor of Polytech Orléans building, site Galilée with map nodes.

only one measurement which will be of course the heading β_k. The measurement input can thus be expressed merely by $Z_k = \beta_k$. We should note that the measurement noise R_k is also considered to be constant and is set here in advance. After conducting extensive testing to determine the best value, we found that the value $R_k = 0.075$ gave good results.

Now, by deriving the Kalman gain K according to Eq.16, the measurement update of state $\hat{X}_k$ and covariance P_k can be calculated as shown in Eq.17 and 18 respectively.

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \quad (16)$$

$$\hat{X}_k = \hat{X}_k^- + K_k(Z_k - H_k \hat{X}_k^-) \quad (17)$$

$$P_k = (I_{3\times 3} - K_k H_k) P_k^k \quad (18)$$

where $I_{3\times 3}$ is the unit matrix of size 3.

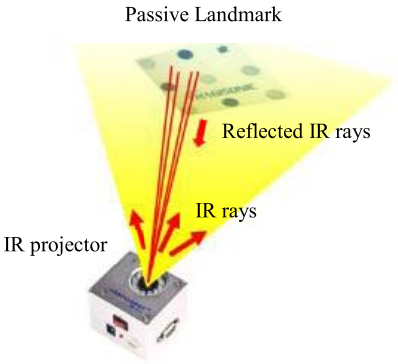


Fig. 8. Illustration showing how the StarGazer device works.

Fig. 5 shows the real trajectory (blue) and the one estimated by the EKF (red) with the above model.

V. Experiments

Now that the two localization methods (vision-based and inertial-based) have been presented, we need to assess their performance. We evaluated the localization results obtained with each of the two methods compared to a Ground Truth, and also compared one with the other.

A. Setup

Our system was evaluated by using a Samsung Galaxy S-III smartphone (Android OS Jelly Bean 4.1.2, Quad core CPU 1.4 GHz and 1 GB of RAM) fixed on a body harness such as shown in Fig. 6. It is of critical importance to note that we chose this configuration, instead of attaching the smartphone to the foot of the user, to provide the user with both audio (guidance instructions) and visual outputs (e.g. their instantaneous position on the map). The results previously given in Section IV.a (Table II) show that mounting the smartphone on a body harness does not have a significant influence on the precision of the step detection algorithm. Our vision method was implemented in Java / C++ using the OpenCV library and OpenCv4Android SDK. Intrinsic camera parameters were

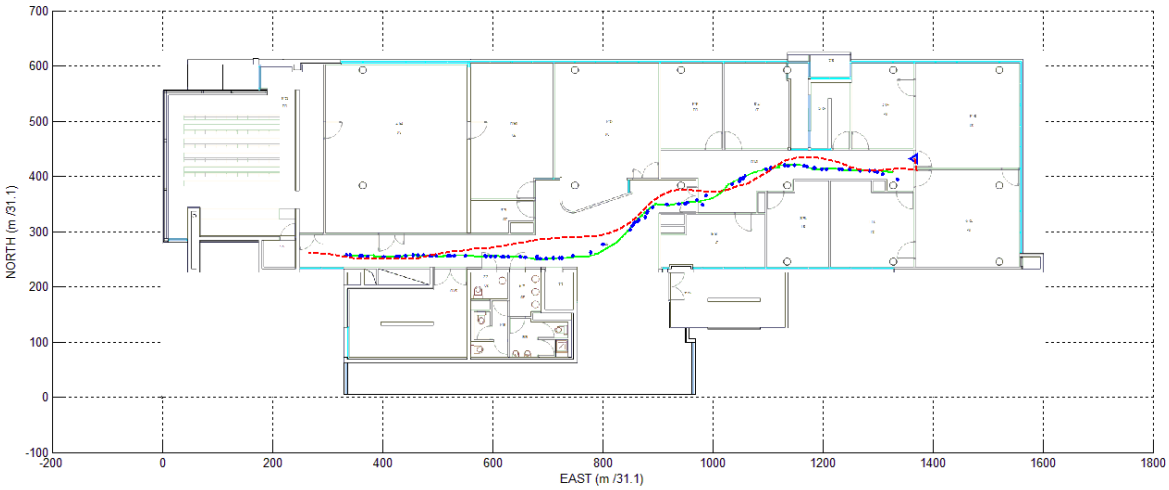


Fig. 9. Localization results of path 1: the reference trajectory (green), the inertial localization trajectory (red) and vision localization trajectory (blue).

Fig. 10. Localization results of path 3: the reference trajectory (green), the inertial localization trajectory (red) and vision localization trajectory (blue).

estimated by prerequisite calibration while the inertial method was implemented in Java for the Android mobile platform.

A map of the 3rd floor of the Polytech Orléans building was created for our localization tests (Fig. 7). It includes corridors, door crossing, 2 hallway intersections and 2 large glazed parts well exposed to the sun. Different possible paths were described through this space. To describe the way to go, we stored for each reference path a series of key frames as well as their location on the map and navigation heading.

To capture the Ground Truth (GT), a commercial robot localization system called Hagisonic StarGazer was used. This device includes an infrared camera and a set of passive landmarks with an independent ID placed on the ceiling, as shown in Fig. 8. The LEDs around the camera (IR projector) send infrared rays to the landmark on the ceiling. The reflected infrared rays are then turned into an image which is analyzed in order to provide the position and the angle of the Hagisonic StarGazer. The type of landmark depends on the distance between the system and the landmark attached to the ceiling (height). For our experiments, we used 34 landmarks of type HLD1-L: 4×4. The Hagisonic StarGazer has two possible operation modes (alone and map modes). In alone mode, the system operates under an independent coordinate system for each landmark, while in map mode, the system operates under a single common coordinate system for all landmarks. The system builds a map defined by regarding the placement of a reference ID as an origin. For this study, only the map mode was considered. The repetitive accuracy of the system is 2 cm and its measurement rate is 10 times/sec. The Hagisonic StarGazer needs to be linked to a laptop for data recording. Its output data are presented as follows:

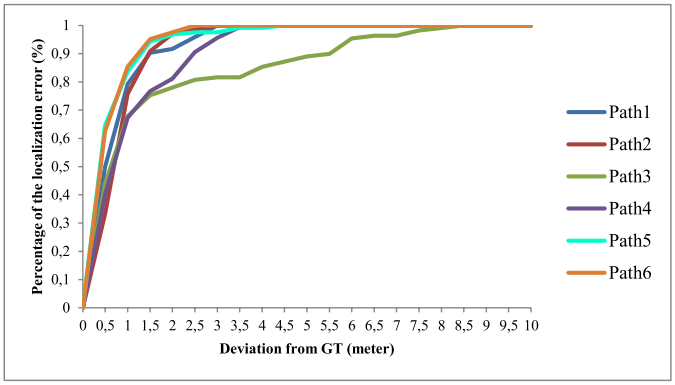
$$< \text{ID},\ \text{Angle [deg]},\ \text{X [cm]},\ \text{Y [cm]},\ \text{Z [cm]},\ \text{Time} >$$

where Angle, X, Y and Z define the sensor angle and position relative to the reference landmark. For each measure, we added the acquisition time in order to synchronize StarGazer output data with those of the inertial and vision systems as well as to detect moments when the signal is lost.

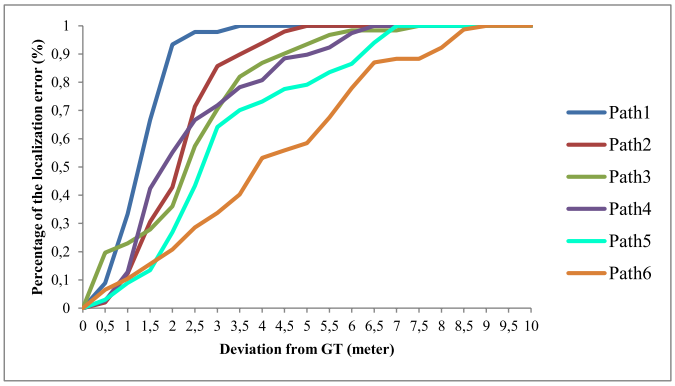
This kind of localization system is designed to be used for robots. To use it in our context, the camera and its battery were installed on a flat surface attached on a small wheeled cart. To reduce noise, the Ground Truth was smoothed by

TABLE III
EXPERIMENTAL PATHS

Path name	Map nodes	Distance (m)	Time traveled (s)
Path 1	A-B-C	35	46.03
Path 2	C-B-A	35	45.53
Path 3	A-B-D-B-C	58	79.36
Path 4	C-B-D-B-A	58	79.70
Path 5	A-B-D-B-A	57	82.36
Path 6	C-B-D-B-C	59	82.20



(a)



(b)

Fig. 11. Localization accuracy evaluation: cumulative histograms of localization errors from the GT: (a) for the vision system and (b) for the inertial system.

removing outliers. Note that the Hagisonic Stargazer linked to the laptop and the two smartphones used (one for the vision-based method and one for the inertial-based method) were manually synchronized. To achieve this, the system clocks of the laptop linked to the StarGazer and those of the two smartphones were synchronized.

B. Results

To evaluate and compare the two proposed localization systems, 6 different paths with different configurations were tested (Table III). The main map nodes of these paths are depicted in Fig. 7. In the following, accuracy and repeatability results of both vision and inertial-based systems are presented.

1) Localization Accuracy: The accuracy of the vision and the inertial sensor-based localization algorithms was evaluated by comparing the qualitative and quantitative results.

TABLE IV
VISION-BASED LOCALIZATION ALGORITHM PERFORMANCE

Path name	Number of key frames	Number of positionings	Correct key frames
Path 1	35	72	90.28%
Path 2	35	66	90.91%
Path 3	58	109	75.23%
Path 4	58	159	76.73%
Path 5	57	122	94.26%
Path 6	59	123	95.12%

TABLE V
RESULTS OF THE VISION-BASED LOCALIZATION ACCURACY

Path name	Average error (m)	STDEV	Error<2m	Error<1m
Path 1	0.685	0.657	91.67%	79.17%
Path 2	0.764	0.547	96.97%	75.76%
Path 3	1.503	2.029	77.98%	67.89%
Path 4	0.975	0.914	81.13%	67.30%
Path 5	0.549	0.684	96.72%	83.61%
Path 6	0.519	0.478	97.56%	85.37%

TABLE VI
RESULTS OF THE INERTIAL SENSOR-BASED LOCALIZATION ACCURACY

Path name	Average error (m)	STDEV	Error<2m	Error<1m
Path 1	1.276	0.588	93.33%	33.33%
Path 2	2.119	0.984	42.86%	12.24%
Path 3	2.371	1.512	36.07%	22.95%
Path 4	2.363	1.566	55.13%	12.82%
Path 5	3.107	1.810	26.87%	8.96%
Path 6	4.146	2.306	20.78%	10.39%

Fig. 9 and 10 depict two representative samples of qualitative comparison. Fig. 9 shows an example of a simple path (Path 1), while Fig. 10 corresponds to Path 3 which is more complex in terms of distance and direction changes. The green curve represents the GT trajectory (obtained by the Hagisonic StarGazer system), while the blue and red curves correspond to the trajectories obtained by the vision-based localization algorithm and the inertial sensor-based localization algorithm respectively.

We can see that the vision localization trajectory and the GT trajectory have the same shape in both cases (Path 1 and Path 3). The inertial sensor-based localization trajectory seems to follow the GT trajectory with some sensor drifts in the case of a simple path (Fig. 9). However, for Path 3, inertial sensor-based localization errors are far more significant, which is essentially due to incorrect direction estimation.

The performance results of the vision-based localization algorithm are presented in Table IV. The average percentage of correct key frames found over all the tested paths (651 localizations) was about 87%.

The quantitative accuracy results for the vision and inertial-based systems are respectively presented in Table V and VI. For each path, the localization error is relative to the reference path (green curve in Fig. 9 and 10). The error corresponds to the instantaneous Euclidian distance between the

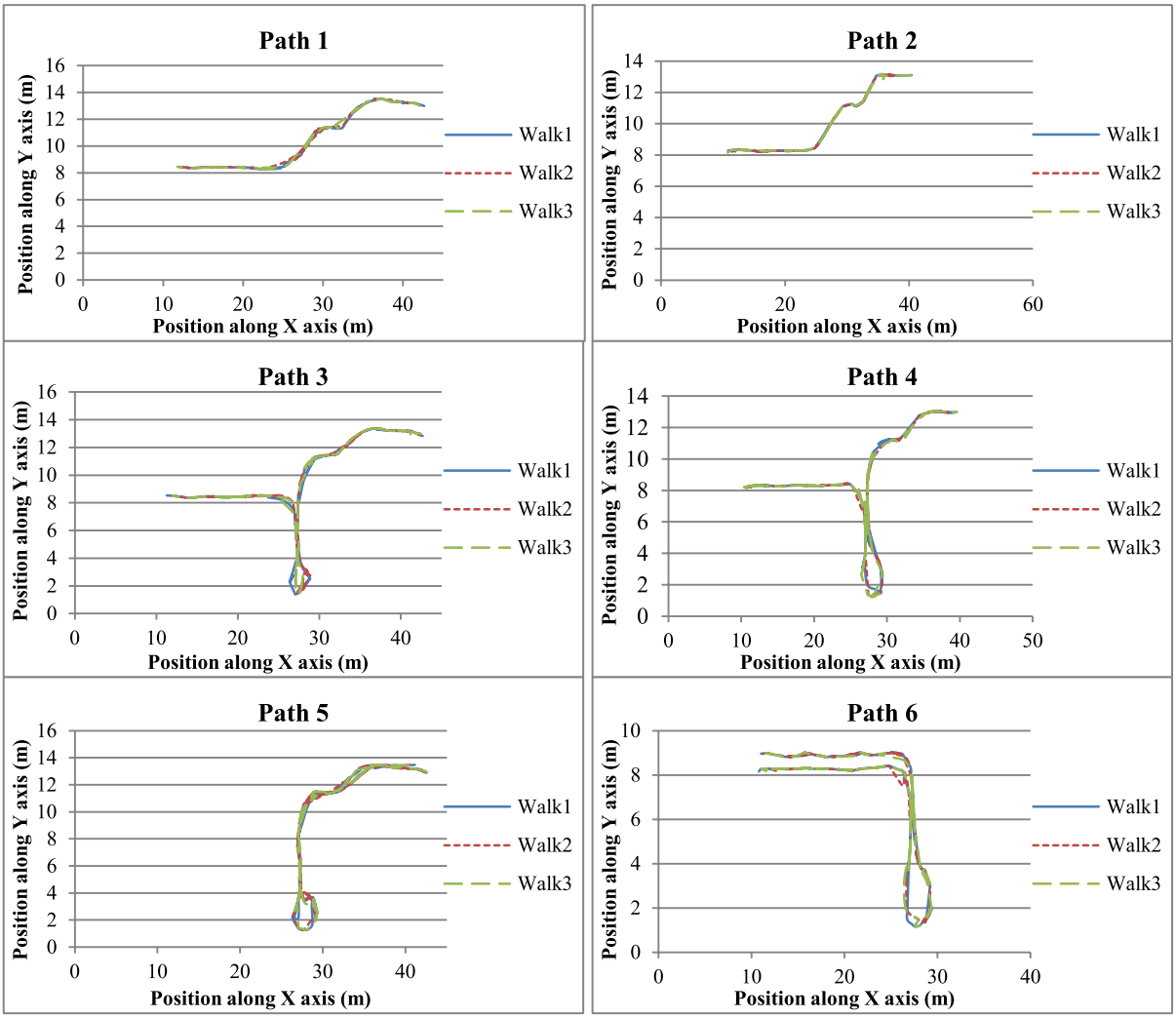


Fig. 12. Repeatability results of the vision-based localization algorithm for the six experimental paths.

two trajectories. These tables present the average and the standard deviation of localization error. We also give the percentage of the localization errors that are lower than a fixed threshold (1 and 2 meters).

For the vision-based system, the localization average error is similar for all the paths, whatever their length or complexity. Most of the instantaneous localization errors are less than 1 meter, which is a good result if we consider that this deviation is equivalent to 1.5 times the human step length. More than 80% of errors are less than 2 meters.

For the inertial-based system, the average localization error increases significantly with the path length. This cumulative deviation is still more visible in the percentage of errors. Contrary to the vision-based system, the results show that the majority of the inertial sensor-based localization errors exceed 1 meter, since more than 70% of errors are greater even for short paths.

These results are due to several factors including the poor quality of the smartphone inertial sensors (drift error and double integration on the small acceleration error), disturbances caused by building infrastructure interference and the problem of pedestrian orientation estimation with the magnetometer.

Clearly, the vision-based system provides a more accurate localization position for all the paths without the drawback of cumulative errors inherent to inertial sensors. This is the advantage of clinging to key frames.

The distribution of the instantaneous localization errors can be seen on Fig. 11, respectively for the vision-based system (a) and the inertial-based system (b). Note that these curves correspond to only one walk for each path. Except for Path 3, the maximum error for the vision-based system remains approximately less than 3 meters. High errors may occasionally be obtained with the vision-based system when the search for the closest key frame fails, especially when in front of a lightly textured or glazed wall (which is well exposed to the sun). In such cases, the number of image features decreases considerably, which affects the matching quality.

The curves on Fig. 11b confirm that the inertial-based system results in much larger errors. For the two simple paths (Path 1 and Path 2) without significant orientation changes, this error is lower.

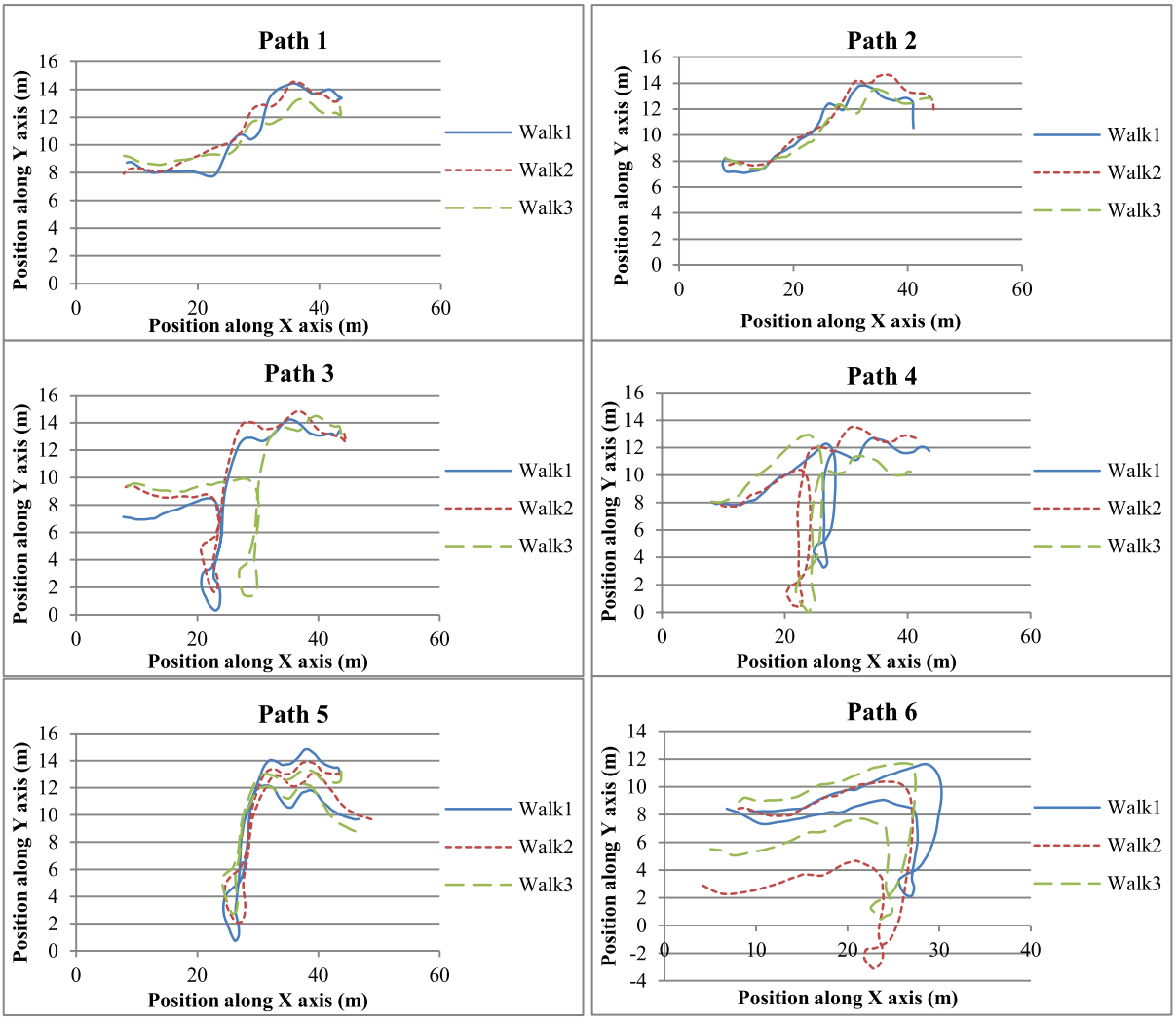


Fig. 13. Repeatability results of inertial sensor-based localization algorithm for the six experimental paths.

2) Localization Repeatability: The repeatability of the two systems was also tested for the same 6 paths (Table II) by travelling three times as presented in Fig. 12 and 13. Altogether, they have a total distance of about one kilometer following corridors, crossing hallways and doors. For the vision-based localization, very slight differences between the three trajectory tests were obtained while for the inertial sensor-based localization, differences between the trajectory tests of repetitive paths were significant. These differences reached 5 meters in some cases (Walk 1 and Walk 2 of Path 6 in Fig. 13). So the inertial-based system has a serious repeatability deficiency.

VI. Conclusion

In this paper, we have presented a comparison between two different technologies for indoor pedestrian localization, each embedded in a smartphone. The first one is based on vision with a new localization method directly implemented in the smartphone. This method enables estimation of the walker's position and orientation by searching for the closest key frame to the current image and detecting the three orthogonal vanishing points in order to determine the smartphone position relative to this key frame. The second technology relies on low-cost smartphone inertial sensors. It consists of step detection and orientation estimation to measure the motion of the pedestrian and to provide his relative position.

The evaluation and comparison of the two localization solutions conducted in this research allowed us to identify the strengths and weaknesses of each technology and therefore to consider perspectives for the future. It also led us to find the best cooperation of both technologies on the same embedded and autonomous system for future work.

Both localization systems are inexpensive, require no additional infrastructure and use lightweight algorithms. Unlike GPS, which is often ineffective in indoor environments, the vision and inertial-based localization information is always accessible at the rate of one localization per second with non-optimized code. We have also demonstrated that the computer vision solution is able to provide instantaneous localization with an acceptable average error (less than 1 meter in most cases). However, two issues remain for the vision-based system. The first is the drift between two successive key frames

where the pedestrian's position is estimated by a Kalman filter. A second issue is the problem of insufficient matched features between two frames under certain conditions. In this case, the closest key frame search phase fails and therefore the estimated position will be inaccurate. As regards the inertial-based system, it reveals the problem of cumulative errors over time that is intensified by the poor quality of smartphone sensors. Fortunately technologies evolve and new smartphones are equipped with more effective inertial sensors and new software drivers allowing programmers to make better use of these sensors. Our preliminary tests on new smartphone models are indeed promising for the future. We may also consider including a variable step length in our algorithm in order to take into account the characteristics of each person. However it will always be necessary to perform an absolute repositioning of inertial results by another technique (tag, vision, etc.) to limit the drifts inherent in the principle of inertial-based localization.

Another improvement would be to integrate the two localization approaches in the same smartphone and to use a graph to model the major walking paths in the building, especially in wide-open spaces. Future research will investigate the solution of re-initializing the inertial-based localization process using the vision measurement (key frame positions) when the vision-based algorithm provides a high number of inliers (reliable matching). Then the inertial-based system will take over for localization between key frames or in the case of an insufficient number of matched features. For the latter case, we plan to use inertial data to perform a guided search of the closest key frame. In this manner, we expect to improve the accuracy and the performance of our localization system. Last but not least, the global performance of our system could be improved by applying a geometrical model of the indoor environment (learned off-line) that restricts the possible displacement of the pedestrian to the navigable areas.

REFERENCES

[1] W. Elloumi *et al.*, "Indoor navigation assistance with a smartphone camera based on vanishing points," in *Proc. IPIN*, Belfort–Montbéliard, France, Oct. 2013, pp. 1–9.

[2] S. Ertan, C. Lee, A. Willets, H. Tan, and A. Pentland, "A wearable haptic navigation guidance system," in *Proc. ISWC*, Oct. 1998, pp. 164–165.

[3] A. Helal, S. E. Moore, and B. Ramachandran, "Drishti: An integrated navigation system for visually impaired and disabled," in *Proc. ISWC*, Oct. 2001, pp. 149–156.

[4] A. Pradhan, E. Ergen, and B. Akinci, "Technological assessment of radio frequency identification technology for indoor localization," *J. Comput. Civil Eng.*, vol. 23, no. 4, pp. 230–238, Jul. 2009.

[5] V. Kulyukin, C. Gharpure, J. Nicholson, and S. Pavithran, "RFID in robot-assisted indoor navigation for the visually impaired," in *Proc. IROS*, Sep./Oct. 2004, pp. 1979–1984.

[6] E. Mok and G. Retscher, "Location determination using WiFi fingerprinting versus WiFi trilateration," *J. Location Based Services*, vol. 1, no. 2, pp. 145–159, Feb. 2008.

[7] L. Zwirello, T. Schipper, M. Harter, and T. Zwick, "UWB localization system for indoor applications: Concept, realization and analysis," *J. Elect. Comput. Eng.*, vol. 2012, Feb. 2012, Art. no. 849638. [Online]. Available: http://dx.doi.org/10.1155/2012/849638

[8] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 1052–1067, Jun. 2007.

[9] M. Kourogi and T. Kurata, "A method of personal positioning based on sensor data fusion of wearable camera and self-contained sensors," in *Proc. MFI*, Jul./Aug. 2003, pp. 287–292.

[10] R. Jirawimut, S. Prakoonwit, F. Cecelja, and W. Balachandran, "Visual odometer for pedestrian navigation," *IEEE Trans. Instrum. Meas.*, vol. 52, no. 4, pp. 1166–1173, Aug. 2003.

[11] S. Treuillet and E. Royer, "Outdoor/indoor vision-based localization for blind pedestrian navigation assistance," *Int. J. Image Graph.*, vol. 10, no. 4, pp. 481–496, Oct. 2010.

[12] C. Arth, D. Wagner, M. Klopschitz, A. Irschara, and D. Schmalstieg, "Wide area localization on mobile phones," in *Proc. IEEE ACM ISMAR*, Orlando, FL, USA, Oct. 2009, pp. 73–82.

[13] J. Straub, S. Hilsenbeck, G. Schroth, R. Huitl, A. Möller, and E. Steinbach, "Fast relocalization for visual odometry using binary features," in *Proc. ICIP*, Sep. 2013, pp. 2548–2552.

[14] H. Aoki, B. Schiele, and A. Pentland, "Realtime personal positioning system for a wearable computer," in *Proc. ISWC*, Washington, DC, USA, Oct. 1999, pp. 37–43.

[15] W. Zhang and J. Kosecka, "Image based localization in urban environments," in *Proc. 3DPVT*, Chapel Hill, NC, USA, Jun. 2006, pp. 33–40.

[16] U. Steinhoff, D. Omerčević, R. Perko, B. Schiele, and A. Leonardis, "How computer vision can help in outdoor positioning," in *Proc. Eur. Conf. AML*, 2007, pp. 124–141.

[17] D. Van Opdenbosch, G. Schroth, R. Huitl, S. Hilsenbeck, A. Garcea, and E. Steinbach, "Camera-based indoor positioning using scalable streaming of compressed binary image signatures," in *Proc. ICIP*, Paris, France, Oct. 2014, pp. 2804–2808.

[18] F. Hu, Z. Zhu, and J. Zhang, "Mobile panoramic vision for assisting the blind via indexing and localization," in *Proc. 2nd Workshop Assistive Comput. Vis. Robot., Conjunct. (ECCV)*, Zürich, Switzerland, Sep. 2014, pp. 600–614.

[19] L. Ruotsalainen, H. Kuusniemi, and R. Chen, "Heading change detection for indoor navigation with a Smartphone camera," in *Proc. IPIN*, Guimarães, Portugal, Sep. 2011, pp. 1–7.

[20] L. Ruotsalainen, H. Kuusniemi, and R. Chen, "Visual-aided two-dimensional pedestrian indoor navigation with a smartphone," *J. Global Positioning Syst.*, vol. 10, no. 1, pp. 11–18, 2011. [Online]. Available: http://www.gnss.com.au/JoGPS/v10n1/JoGPS_v10n1p11-18.pdf

[21] C. Hide, T. Moore, and T. Botterill, "Low cost vision-aided IMU for pedestrian navigation," *J. Global Positioning Syst.*, vol. 10, no. 1, pp. 3–10, 2011. [Online]. Available: http://www.gnss.com.au/JoGPS/v10n1/JoGPS_v10n1p3-10.pdf

[22] A. Amanatiadis, A. Gasteratos, and D. Koulouriotis, "An intelligent multi-sensor system for first responder indoor navigation," *Meas. Sci. Technol.*, vol. 22, no. 11, p. 114025, Nov. 2011.

[23] C. Keßler, C. Ascher, M. Flad, and G. F. Trommer, "Multi-sensor indoor pedestrian navigation system with vision aiding," *J. Gyroscopy Navigat.*, vol. 3, no. 2, pp. 79–90, Apr. 2012.

[24] S. Saeedi, A. Moussa, and N. El-Sheimy, "Vision-aided context-aware framework for personal navigation services," in *Proc. ISPRS*, Melbourne, VIC, Australia, 2012, pp. 1–6.

[25] R. Mautz and S. Tilch, "Survey of optical indoor positioning systems," in *Proc. IPIN*, Guimarães, Portugal, Sep. 2011, pp. 1–7.

[26] W. Elloumi, S. Treuillet, and R. Leconge, "Real-time camera orientation estimation based on vanishing point tracking under Manhattan World assumption," *J. Real-Time Image Process.*, Apr. 2014. [Online]. Available: http://link.springer.com/article/10.1007/s11554-014-0419-9, doi: 10.1007/s11554-014-0419-9.

[27] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors," *IEEE Comput. Graph. Appl.*, vol. 25, no. 6, pp. 38–46, Nov./Dec. 2005.

[28] R. Harle, "A survey of indoor inertial positioning systems for pedestrians," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1281–1293, Jan. 2013.

[29] A. R. Jimenez, F. Seco, C. Prieto, and J. Guevara, "A comparison of Pedestrian Dead-Reckoning algorithms using a low-cost MEMS IMU," in *Proc. WISP*, Aug. 2009, pp. 37–42.

[30] L. Fang *et al.*, "Design of a wireless assisted pedestrian dead reckoning system—The NavMote experience," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 6, pp. 2342–2358, Dec. 2005.

[31] Y. Jin, H.-S. Toh, W.-S. Soh, and W.-C. Wong, "A robust dead-reckoning pedestrian tracking system with low cost sensors," in *Proc. PerCom*, Mar. 2011, pp. 222–230.

[32] J. W. Kim, H. J. Jang, D.-H. Hwang, and C. Park, "A step, stride and heading determination for the pedestrian navigation system," *J. Global Positioning Syst.*, vol. 3, nos. 1–2, pp. 273–279, Feb. 2005.

[33] S. Ayub, X. Zhou, S. Honary, A. Bahraminasab, and B. Honary, "Indoor pedestrian displacement estimation using smart phone inertial sensors," *Int. J. Innov. Comput. Appl.*, vol. 4, no. 1, pp. 35–42, Mar. 2012.

[34] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A reliable and accurate indoor localization method using phone inertial sensors," in *Proc. 14th ACM Int. Conf. Ubiquitous Comput.*, Pittsburgh, PA, USA, Sep. 2012, pp. 421–430.
[35] I. Apostolopoulos, N. Fallah, E. Folmer, and K. E. Bekris, "Integrated online localization and navigation for people with visual impairments using smart phones," in *Proc. ICRA*, Saint Paul, MN, USA, May 2012, pp. 1322–1329.
[36] J. Á. B. Link, P. Smith, N. Viol, and K. Wehrle, "FootPath: Accurate map-based indoor navigation using smartphones," in *Proc. IPIN*, Guimarães, Portugal, Sep. 2011, pp. 1–8.
[37] G. H. Flores, R. Manduchi, and E. D. Zenteno, "Ariadne's thread: Robust turn detection for path back-tracing using the iPhone," in *Proc. IEEE Ubiquitous Positioning Indoor Navigat. Location Based Service (UPINLBS)*, Nov. 2014, pp. 133–140.
[38] W. Kang and Y. Han, "SmartPDR: Smartphone-based pedestrian dead reckoning for indoor localization," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2906–2916, May 2015, doi: 10.1109/JSEN.2014.2382568.
[39] *Project Tango*, accessed on Jun. 5, 2014. [Online]. Available: https://www.google.com/atap/project-tango/
[40] J. Z. Liang, N. Corso, E. Turner, and A. Zakhor, "Image-based positioning of mobile devices in indoor environments," in *Proc. Multimodal Location Estimation Videos Images*, 2015, pp. 85–99.
[41] W. Elloumi, S. Treuillet, R. Leconge, and A. Fonte, "Performance evaluation of point matching methods in video sequences with abrupt motions," in *Proc. VISAPP*, 2010, pp. 427–430.
[42] J. M. Coughlan and A. L. Yuille, "Manhattan world: Compass direction from a single image by Bayesian inference," in *Proc. ICCV*, Sep. 1999, pp. 941–947.
[43] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. AVC*, 1988, pp. 147–151.
[44] Y. Kim, Y. Chon, and H. Cha, "Smartphone-based collaborative and autonomous radio fingerprinting," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 1, pp. 112–122, Jan. 2012.
[45] W. Chen, R. Chen, Y. Chen, H. Kuusniemi, and J. Wang, "An effective pedestrian dead reckoning algorithm using a unified heading error model," in *Proc. PLANS*, May 2010, pp. 340–347.
[46] U. Steinhoff and B. Schiele, "Dead reckoning from the pocket—An experimental study," in *Proc. PerCom*, Mar. 2010, pp. 162–170.
[47] *Sensors Overview for Android Developers. Added in API Level 3, Android 1.5 Platform*, accessed on May 2009. [Online]. Available: http://developer.android.com/guide/topics/sensors/sensors_overview.html
[48] *Using the ADXL202 in Pedometer and Personal Navigation Applications*, Analog Devices Inc., Cambridge, MA, USA, 2002.
[49] J. Liu, R. Chen, L. Pei, R. Guinness, and H. Kuusniemi, "A hybrid smartphone indoor positioning solution for mobile LBS," *Sensors*, vol. 12, no. 12, pp. 17208–17233, 2012.
[50] M. J. Caruso, "Applications of magnetic sensors for low cost compass systems," in *Proc. IEEE Position Location Navigat. Symp. (PLANS)*, San Diego, CA, USA, Mar. 2000, pp. 177–184.

Wael Elloumi received the master's degree in computer science from the INSA of Rouen, France, in 2008, and the Ph.D. degree in computer vision from the University of Orleans, France, in 2012. From 2012 to 2015, he was a Post-Doctoral Researcher with the Image and Vision Team, PRISME Laboratory, University of Orleans. He is currently a Researcher at the R&D team of Worldline company. His research interests include computer vision, image matching, 3-D mapping, indoor navigation, and assistive systems and biometric authentication.

Abdelhakim Latoui received the Electronics Engineering Diploma and M.S. degrees in electronics from the Electronics Institute, University of Ferhat Abbes, Setif, Algeria, in 1996 and 2004, respectively, and the Ph.D. degree in electronics from the University of Ferhat Abbes Setif-1, Algeria, in 2013. He is currently an Assistant Professor with the Department of Electronics, Université Mohamed El Bachir El Ibrahimi, Bordj Bou-Arreridj, Algeria. His research interests focus on design and test and behavioral description of complexes circuits.

Raphaël Canals was born in Thionville, France, in 1966. He received the Dipl.-Ing. degree in electrical engineering and the Ph.D. degree in electronics from the University of Clermont-Ferrand, France, in 1989 and 1993, respectively. His Ph.D. dissertation describes two methods of non-supervised image segmentation and the way to implement them on a parallel processing machine. This work deals with the issues of the fit between algorithm and architecture. In 1994, he joined the Polytechnic School of the University of Orléans, France. He is currently a Teacher and a Researcher with the Laboratory PRISME. His present interests are in biomedical imaging, electronic devices, and real-time signal processing.

Aladine Chetouani received the master's degree in computer science from the University Pierre and Marie CURIE, France, in 2005, and the Ph.D. degree in image processing from the University of Paris 13, France, in 2010. From 2010 to 2011, he was a Post-Doctoral Researcher with the L2TI Laboratory, Paris 13 University. He is currently a Teacher and a Researcher with the Laboratory PRISME. His present interests are in image quality, perceptual analysis, visual attention, and image processing for cultural heritage.

Sylvie Treuillet received the master's degree in electrical engineering from the University of Clermont-Ferrand, France, in 1988, and the Ph.D. degree in computer vision from the University of Clermont-Ferrand, France, in 1993. From 1988 to 1990, she was with a private company in Montpellier, where she was involved in developing image processing and pattern recognition methods in automatics cytogenetics. She then joined the Academic Research Laboratory LASMEA. Since 1994, she has been an Assistant Professor of Computer Sciences and Electronic Engineering with the Ecole Polytechnique and with the PRISME Laboratory, University of Orléans, France. Her research interests include computer vision for 3-D object modeling, pattern recognition, and color image analysis for biomedical or industrial applications.