Wearable indoor pedestrian dead reckoning system

Jorge Torres-Solis\textsuperscript{a,b,c,d}, Tom Chau\textsuperscript{a,b,c,}\textsuperscript{*}

\textsuperscript{a} Bloorview Kids Rehab, 150 Kilgour Road, Toronto, ON, M4G 1R8, Canada
\textsuperscript{b} Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Canada
\textsuperscript{c} Institute of Biomaterials and Biomedical Engineering, University of Toronto, Canada
\textsuperscript{d} Komodo OpenLab, Toronto, ON, Canada

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Pedestrian indoor navigation
Dead reckoning
Camera-assisted position correction

\textbf{ABSTRACT}

We introduce a wearable pedestrian indoor localization system with dynamic position correction. The system uniquely combines dead reckoning and fiducial marker-based localization schemes, exclusively using widely available, low end and low power consumer hardware components.

The proposed system was tested with various walking patterns inside a building, achieving an indoor positioning accuracy of 3.38\% of the total distance walked. This accuracy is comparable to those obtained with solutions deploying specialized high cost hardware components. The low cost wearable system proposed herein could serve as the foundation for a pervasive solution for indoor way finding and patient tracking.

1. Introduction

The estimation of indoor user position is a key challenge in the development of ubiquitous computing environments because location is a fundamental aspect of a user’s context [1]. A system must be location-aware in order to seamlessly deliver context-relevant information to the user [2]. Therefore, to realize a truly pervasive system, an accurate indoor positioning system is necessary.

In recent years, localization systems based on pedestrian dead reckoning (PDR) have gained interest among the scientific community. In PDR systems, position estimation is achieved by double integration of accelerometer measurements, while the orientation is estimated by integrating the angular velocity obtained from gyroscopes [3] or by obtaining the direction of the geographical north pole using magnetometers. The latter are oftentimes combined with gyroscopes, as they are prone to errors due to metallic structures and electrical appliances in the vicinity [4]. Inertial sensors (accelerometers and gyroscopes) are prone to drift errors. To mitigate such drift errors during pedestrian tracking, the properties of human gait can be exploited, resetting the velocity and acceleration values when the accelerometers are known to be static. This technique is called “zero-velocity updates” or ZUPT [4].

As mentioned in [4], most of the experiments in the previous literature regarding PDR do not consider varied and realistic walking patterns. Recent research has presented PDRs capable of tracking the movements of a pedestrian without imposing restrictions on how and where a person walks. For example, Sagawa et al. [5] propose a method to measure the distance traversed without imposing constraints on speed, or stride length. This system used a combination of 3 accelerometers, a gyroscope and an atmospheric pressure sensor. The system was only tested during straight-path walking and stair climbing tasks, achieving a horizontal distance estimation error of less than 5.3\% of the total distance traveled. A PDR system capable of tracking unrestricted pedestrian movements that achieves an indoor positioning accuracy within 0.3\%
of the distance traveled is presented in [4]. This solution uses a hardware device named InertiaCube3, which is composed of 3 accelerometers, 3 gyroscopes and 3 magnetometers. Localization is achieved by combining the location information provided by the sensors using Kalman filters.

Other authors have considered combining PDR systems with localization systems based on external location references to recalibrate or enhance the accuracy of the dead reckoning navigation system. Recent research proposes the recalibration of inertial navigation systems for vehicles using image processing [6–8]. Of particular interest for this work are indoor pedestrian tracking solutions which combine inertial sensors and image processing. In [9] a camera was used to recalibrate an indoor localization system designed for pedestrians, using pattern classification to determine the movement performed by the user, and Kalman filtering to combine the location data provided by the sensors. The inertial tracking system was custom made, and the hardware setup was not fully described in the article. In [10] a camera was mounted on the head, pointing in the superior direction with respect to the user. Several markers were positioned on the ceiling and walls, arranged in a high density pattern of about 1.7 markers per square metres. InertiaCube2 was used as the inertial measurement device, which contains a sensor array similar to InertiaCube3, described above. The localization accuracy achieved by this system was not reported.

1.1. Motivation

Although indoor pedestrian navigation systems have been previously investigated, prior solutions require custom made hardware [9] or specialized commercial components which are expensive and not readily available to the general public. For instance, the InertiaCube family of devices have been used in [4, 10], producing solutions in the thousands of dollars. Further, both studies only present data from a single indoor localization experiment.

Pedestrian localization applications in the fields of rehabilitation, patient monitoring or personnel tracking require solutions that are economically attractive and readily available on the market. These requirements are particularly relevant for technological solutions in developing countries, where specialized hardware is difficult to acquire or prohibitive in terms of cost [11, 12]. In light of these observations and the review above, it is clear that further research is needed to develop cost effective localization systems that allow indoor pedestrian tracking for unrestricted navigation in real-time. We hypothesized that an indoor pedestrian navigation system can be built using mainstream hardware components, while achieving position errors comparable to previously published solutions.

The specific objective of this work was to investigate an indoor pedestrian navigation system that:

- Can be easily built and deployed by using mainstream hardware components,
- Does not require major infrastructure alterations, such as the installation of wireless network equipment,
- Requires less than 1 fiducial marker per 10 square metres, and
- Yields real-time location estimates with an accuracy comparable to that of solutions found in recent literature.

The reason for limiting the amount of fiducial markers in the environment is to minimize the size of the database associated with the fiducial markers and the visual clutter, which would be unacceptable in health care facilities and public buildings. In this paper, we introduce novel features to the PDR field by proposing the usage of mainstream, low-cost hardware components in the creation of an indoor PDR system with image processing-based recalibration.

2. System design

2.1. Hardware set-up

Our PDR system is composed of a unique combination of sensors, namely:

- A Nintendo Wii remote control to obtain tri-axial accelerometry data,
- A Gyrosense Gyropoint air mouse to obtain angular velocity from its dual-axis gyroscope,
- A Logitech QuickCam Pro 4000 webcam to detect visual markers fixed in the environment, and
- A small form factor computer with a Bluetooth interface and USB ports.

To record the data acquired by the sensors, the small form factor computer was mounted on an Ergotron mobile workstand cart. A Bluetooth dongle and the Gyrosense wireless receiver were connected to the computer. The laptop was pushed along by an experimenter while the subject walked on the predefined circuit. The Wii remote control was mounted on the lateral side of the right ankle. The gyromouse was mounted anterior to the abdomen, with one sensitive axis in the superior–inferior direction and the other in the mediolateral direction. The camera was mounted on the left shoulder of the participant. Although the camera location differed from the foot location, they were spatially correlated. The data processing stage (Section 5) specifically compensated for this location difference. The hardware arrangement is presented in Fig. 1.

Thirty distinct visual markers were positioned at the corners of a predefined polygonal indoor circuit. We created a database containing the positions of each marker within the building. These marker positions were obtained with a laser distance measurement tool. The distances were measured twice and then averaged to correct for measurement deviations. A marker consisted of a letter-sized sheet of paper with a unique $17 \times 17$ cm$^2$ symbolic image. The visual markers and the image processing algorithms used for image localization are described in detail in [13].
Fig. 1. Pedestrian dead reckoning system showing wearable sensing and computing elements as well as a typical fiducial marker in the environment. The wearable computer was mounted on a mobile cart for this particular experiment (not shown).

Fig. 2. Reference frames. \((X, Y, Z)\) are the vectors that define the building reference frame. \((u, v, w)\) are the vectors that define the Wii reference frame.

2.2. Reference frames

The Wii controller was attached to the leg using a spandex sleeve to ensure intimate coupling with the limb. During gait, the tibial angle changed with respect to the axis defined by the gravitational force. Therefore, the orientation of the accelerometer axes changed with respect to the gravitational axis.

We defined two reference frames, one that was fixed with respect to the building, aligned with the gravitational force, which we call building reference frame. We use the letters \(X\), \(Y\) and \(Z\) to label the axis of this reference frame, with \(Z\) parallel to the gravitational axis and \(X\) pointing north.

The second reference frame was fixed with respect to the Wii controller, and we call this the Wiimote reference frame. We use the letters \(u\), \(v\) and \(w\) to label the axis for this reference frame, with \(w\) parallel to the tibia. We also define the angle \(\phi\) as the angle between the building \(Z\) axis and the Wii remote \(v\) axis.

The reference frames for this study are shown in Fig. 2.

2.3. Software

The most important software components of the system are listed below. All software was written in C, unless specified otherwise.

- ARToolkit: library for augmented reality applications. A program was created using this library to obtain the position of the webcam relative to the markers present in the environment.
Table 1
Localization metrics for 8 different walking tasks performed by the male participant. Abbreviations: CW = clockwise; CCW = counter-clockwise; UC = uncorrected measure using the Inertial sensors only; C = corrected measure using a combination of inertial sensors and image processing-based recalibration.

<table>
<thead>
<tr>
<th>Walking pattern</th>
<th>Mean out-of-bounds error (m)</th>
<th>Maximum out-of-bounds error (m)</th>
<th>Final position error (m)</th>
<th>Final bearing error (degrees)</th>
<th>Traveled distance (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UC</td>
<td>C</td>
<td>UC</td>
<td>C</td>
<td>UC</td>
</tr>
<tr>
<td><strong>CW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>3.47</td>
<td>2</td>
<td>8.3</td>
<td>8.28</td>
<td>11.05</td>
</tr>
<tr>
<td>Zigzag</td>
<td>2.65</td>
<td>0.32</td>
<td>7.25</td>
<td>4.22</td>
<td>14.83</td>
</tr>
<tr>
<td>6 steps, 5 s</td>
<td>0.42</td>
<td>0.24</td>
<td>1.75</td>
<td>2.51</td>
<td>3.57</td>
</tr>
<tr>
<td>6 steps, turn</td>
<td>1.02</td>
<td>0.99</td>
<td>4</td>
<td>4.85</td>
<td>4.89</td>
</tr>
<tr>
<td><strong>CCW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>4.34</td>
<td>0.74</td>
<td>13.44</td>
<td>3.03</td>
<td>25.75</td>
</tr>
<tr>
<td>Zigzag</td>
<td>0.51</td>
<td>0.22</td>
<td>3.01</td>
<td>2.84</td>
<td>6.83</td>
</tr>
<tr>
<td>6 steps, 5 s</td>
<td>0.68</td>
<td>0.91</td>
<td>3.65</td>
<td>5.84</td>
<td>8.39</td>
</tr>
<tr>
<td>6 steps, turn</td>
<td>2.42</td>
<td>1.77</td>
<td>6.6</td>
<td>7.5</td>
<td>21.81</td>
</tr>
<tr>
<td><strong>Averages</strong></td>
<td>1.94</td>
<td>0.9</td>
<td>6</td>
<td>4.88</td>
<td>12.14</td>
</tr>
</tbody>
</table>

Table 2
Localization metrics for 8 different walking tasks performed by the female participant. Abbreviations: CW = clockwise; CCW = counter-clockwise; UC = uncorrected measure using the Inertial sensors only; C = corrected measure using a combination of inertial sensors and image processing-based recalibration.

<table>
<thead>
<tr>
<th>Walking pattern</th>
<th>Mean out-of-bounds error (m)</th>
<th>Maximum out-of-bounds error (m)</th>
<th>Final position error (m)</th>
<th>Final bearing error (degrees)</th>
<th>Traveled distance (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UC</td>
<td>C</td>
<td>UC</td>
<td>C</td>
<td>UC</td>
</tr>
<tr>
<td><strong>CW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>4.92</td>
<td>1.98</td>
<td>11.59</td>
<td>5.92</td>
<td>7.07</td>
</tr>
<tr>
<td>Zigzag</td>
<td>2.64</td>
<td>1.28</td>
<td>6.77</td>
<td>5.65</td>
<td>11.48</td>
</tr>
<tr>
<td>6 steps, 5 s</td>
<td>1.65</td>
<td>0.26</td>
<td>6.88</td>
<td>2.42</td>
<td>11.82</td>
</tr>
<tr>
<td>6 steps, turn</td>
<td>3.11</td>
<td>1.76</td>
<td>12.28</td>
<td>11.31</td>
<td>14.22</td>
</tr>
<tr>
<td><strong>CCW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>5.99</td>
<td>1.61</td>
<td>13.27</td>
<td>6.8</td>
<td>32.7</td>
</tr>
<tr>
<td>Zigzag</td>
<td>1.85</td>
<td>1.57</td>
<td>7.97</td>
<td>8.25</td>
<td>27.66</td>
</tr>
<tr>
<td>6 steps, 5 s</td>
<td>2.82</td>
<td>0.92</td>
<td>6.97</td>
<td>4.92</td>
<td>5.27</td>
</tr>
<tr>
<td>6 steps, turn</td>
<td>1.97</td>
<td>1.1</td>
<td>5.2</td>
<td>3.37</td>
<td>5.46</td>
</tr>
<tr>
<td><strong>Averages</strong></td>
<td>3.12</td>
<td>1.31</td>
<td>8.87</td>
<td>5.97</td>
<td>14.46</td>
</tr>
</tbody>
</table>

- Libcwiid: library to communicate with the Wii controller. A program was written to capture data from the Wii controller (accelerometry signals).
- Program to capture data from the Gyropoint air mouse, using the mouse data obtained from the operating system, written in C.
- Program to coordinate all the data capturing routines in a synchronized fashion, and
- Software for data analysis and processing, written in Matlab.

The ARToolkit identified fiducial markers placed in the environment. Each fiducial marker was composed of a unique arrangement of black and white squares within a 4 by 4 grid, enclosed in a black frame. Through image processing, ARToolkit calculated the distance and orientation of the camera relative to the marker. We created a database of markers along with their position and orientation. In this way, we were able to calculate the absolute position of the camera in the building.

3. Experiments

An 84.4 m indoor circuit was traversed 4 times in each direction, clockwise (CW) and counter-clockwise (CCW), yielding a total of 8 trials per participant. For each trial, the participant followed a specific walking pattern (see Tables 1 and 2). Therefore, each trial or traverse of the circuit will also be referred to as ‘walking task’. Before each task, we set initial conditions for the experiment, and obtained calibration constants for the inertial sensors. The camera was calibrated only once, and recalibration was not required as long as the same camera was used for all walking tasks.

The system was tested by two participants walking at a comfortable pace. The first participant was a 30 year old, 1.81 m tall male, weighting 72 kg. The second participant was a 19 year old, 1.60 m tall female, weighting 48 kg.

3.1. Experiment initialization

A start point and an initial body bearing were defined for each walking direction (clockwise and counter-clockwise). Pieces of tape were used to mark these starting points on the floor. The user had to position the toes of the right foot on the defined markers. This allowed us to define values for the top-view startup position (Pos_0, Pos_y(0)) and the initial user bearing (\(\theta(0)\)) in building coordinates.
Since the user started the experiment in a static position, the acceleration and velocity of the ankle relative to the building are initially equal to zero.

\[ V_{\text{Ank}}[0] = A_{\text{Ank}}[0] = 0. \]

### 3.2. Sensor calibration

On startup, the user is requested to stand in a vertical position for 5 s. This allowed the system to obtain calibration values from the accelerometers. The values obtained during this calibration are recorded as offset constants for the accelerometers, and are named \( A_{\text{u OFFSET}} \), \( A_{\text{v OFFSET}} \) and \( A_{\text{w OFFSET}} \).

The sensitivity of the accelerometers can be calculated from calibration values which are stored in the Wii remote EEPROM, in memory locations \( 0 \times 16 \) to \( 0 \times 1A \). These values contain the calibrated readings for each sensitive axis when subjected to 0g and 1g accelerations, where \( 'g' \) is the acceleration due to the gravity. We obtained the numerical relationship between Wii acceleration units and \( 'g' \) for each sensitive axis from these values: 27 Wii accelerometer units equal 1g, or 9.8 m/s\(^2\). We defined a conversion constant called ‘Wii2MSS’, to convert Wii acceleration units to physical units of m/s\(^2\).

\[
\text{Wii2MSS} = \frac{9.81}{27}. \tag{1}
\]

To calibrate the gyroscopes, the user was requested to rotate the body, changing his yaw by 90 degrees to the right, and then to rotate his body back to the original orientation. The mouse ticks reported by the Gyrosense mouse upon the rotation to the right and left were accumulated in the variables \( \text{Ticks}_R \) and \( \text{Ticks}_L \), respectively. We calculated the sensitivity of the Gyrosense mouse in degrees per mouse tick for left and right rotations in the following fashion

\[
\text{GyroSens}_R = \frac{90}{\text{Ticks}_R}; \quad \text{GyroSens}_L = \frac{90}{\text{Ticks}_L}. \tag{2}
\]

The camera localization system was calibrated using the software utilities provided by ARToolkit, which are explained in detail in [13].

### 3.3. Walking tasks

A walking task consisted of one complete traverse of the predefined circuit. For each task, the user conformed to one of the following walking patterns (each pattern was performed in both directions):

1. Walk continuously along the circuit in a natural straight path, turning only at the corners of the circuit, until the circuit is completed (the starting point is reached again).
2. Walk continuously along the circuit in a zigzag path until the circuit is completed.
3. Take 6 consecutive steps then stop for 5 s. Continue walking in this fashion until the circuit is completed.
4. Take 6 consecutive steps then stop and rotate your body to look left and right and then start walking again. Walk in this fashion until the circuit is completed.

### 4. Position and orientation estimation

The data obtained from the inertial sensors were processed using a localization algorithm with low memory and processing power requirements. If improved localization accuracy is required in the future, these data can be processed using advanced techniques such as Kalman or particle filtering.

To accurately track motion of a rigid body in 3D space with 6 degrees of freedom, an inertial unit, namely a combination of 3 accelerometers and 3 gyroscopes, is required [3]. In the case of an articulated body, an inertial unit has to be mounted on each articulated section of the body [14, 15]. We simplified the system by making the following assumptions:

- The ankle is articulated with respect to the pelvis, but in general the distance between the ankle and the pelvis will remain within certain limits. The movement of the torso (where the gyroscope is attached) and the legs (where the accelerometers are attached) is assumed to be spatially correlated.
- Three accelerometers can be used to calculate the total acceleration of the ankle due solely to inertia, eliminating the contribution of gravity. We can then extract forward and lateral components of the movement of the leg by using the acceleration components in each axis.
- The user of the system will be walking upright with minimal pelvic tilt. We can then use one single gyroscope to estimate the user’s bearing. If this condition is not met, 3 gyroscopes would be required.
- We assume that the user walks over a horizontal plane. This means that the accumulated acceleration of the foot during a full gait cycle (i.e. heel-off to heel-off) must be equal to zero, resulting in no net vertical displacement of the foot.
4.1. Emulation of uniform data sampling

The gyromouse and the Wii remote controller have non-uniform sampling rates. We created uniformly sampled versions of the data from these sensors, simulating a sampling frequency of $F_s = 20$ Hz. The sampling time is the inverse of the sampling frequency (i.e. $T_s = 1/F_s$).

The Wii controller does not report data updates unless the values have changed since the last values reported (i.e. it will not send data when it is static or its movement speed is constant). We generated a uniformly sampled version of the accelerometry signals sent by the Wii controller for each uniform period $i = 1, 2, 3, \ldots, N$ in the following fashion:

- If there were one or more Wii data reports per uniform period $i$, the last acceleration values received during that period are recorded as $A_u[i], A_v[i]$ and $A_w[i]$;
- If the Wii remote did not send any reports during the uniform period $i$, copy the values stored from the previous period into the values of the current period

\[ \begin{align*}
A_u[i] &= A_u[i-1]; \\
A_v[i] &= A_v[i-1]; \\
A_w[i] &= A_w[i-1].
\end{align*} \]

The gyromouse only reports data if the angular velocity of any of its axis is different from zero. We generated a uniformly sampled version of the data reported by the gyrosopes in the following fashion:

- If one or more reports are received during one sampling period $i$, we assign the accumulated value of the samples received from the superior–inferior and the mediolateral sensitive axes during the period to $R_{SI}[i]$ and $R_{ML}[i]$ respectively;
- If we do not receive a report from the gyrosopes during a period $i$, we assign a value of zero to $R_{SI}[i]$ and $R_{ML}[i]$.

4.2. Acceleration, velocity and distance traveled by the ankle

As the orientation of the sensitive axis of the Wii remote changes with respect to the gravitational field during a gait cycle, we need to consider the total inertial acceleration of the Wii remote, in order to determine the distance traveled. The three accelerometers have identical sensitivity values. This allowed us to calculate the inertial acceleration magnitude, which is independent of the orientation of the Wii remote.

\[
A_{\text{Ank}}[i] = |A_u[i] - A_{w,\text{OFFSET}}| + |A_v[i] - A_{v,\text{OFFSET}}| + |A_w[i] - A_{w,\text{OFFSET}}|. \tag{3}
\]

Technically, the ankle acceleration should be obtained via the vector sum of the orthogonal acceleration vectors. For the sake of computational efficiency however, we summed the absolute values of the magnitudes of individual acceleration components. Based on the triangle inequality, this sum is an upper bound for the magnitude of the vector sum of the three acceleration components. The absolute value bars are intentionally omitted from the second term to eliminate the accumulation of small aberrant vibrations and sensor noise in the $v$ axis.

We calculated the ankle speed relative to the ground by integrating Eq. (3)

\[
V_{\text{Ank}}[i] = A_{\text{Ank}}[i] \cdot T_s + V_{\text{Ank}}[i-1] \tag{4}
\]

where $T_s$ is again the sampling period. We then calculated the distance traveled by the ankle by double integration of Eq. (3). We converted the calculated distance to metres by multiplying the resulting value by the conversion factor ‘Wii2MSS’, defined in Eq. (1).

\[
D_{\text{Ank}}[i] = (A_{\text{Ank}}[i] \cdot T_s^2/2 + V_{\text{Ank}}[i-1] \cdot T_s) \cdot \text{(Wii2MSS)}. \tag{5}
\]

4.3. Zero-velocity updates for accelerometry signals

To mitigate the measurement drift of the accelerometers, we used a zero-velocity update (ZUPT) technique, which consists of resetting the speed and acceleration to zero when the accelerometer is known to be static, i.e., between heel-strike and toe-off of the instrumented foot. Originally we modified the Wii controller and connected a force sensor (Interlink electronics FSR, model 406) to one of the Wii controller buttons. The force sensor was placed under the heel, and was used as a foot switch. When the sensor was depressed, the acceleration and velocity values were reset to zero.

The force sensor was discontinued from the system because it is not a mainstream hardware component, and because it failed as a result of mechanical stress after approximately 170 ms of walking.

To substitute for the force sensor, we generated a synthetic foot switch signal by estimating the acceleration of the ankle due to inertia ($A_{\text{Switch}}[i]$) as the sum of the acceleration sensed by each of the three accelerometers contained in the Wii controller, namely,

\[
A_{\text{Switch}}[i] = |A_u[i] - A_{w,\text{OFFSET}}| + |A_v[i] - A_{v,\text{OFFSET}}| + |A_w[i] - A_{w,\text{OFFSET}}|. \tag{6}
\]

The synthetic foot switch signal was then generated using the following rule,

\[
\text{FootSwitch}[i] = \begin{cases} 
0, & \text{if } A_{\text{Switch}}[i] > A_{\text{min}} \\
1, & \text{if } A_{\text{Switch}}[i] \leq A_{\text{min}}.
\end{cases} \tag{7}
\]
If the value of $A_{\text{Switch}}[i]$ was lower than or equal to a minimum acceleration threshold, $A_{\text{min}}$, we assumed that the ankle was static on the ground. This minimum threshold value was selected empirically by visually comparing the synthetic foot switch signal against the signal obtained using an actual force sensor under the heel. Normally, the numerical value of the minimum acceleration would be dependent on the resolution of the analog-to-digital conversion and the sensitivity of the accelerometer. In the present study, $A_{\text{min}} = 8$, or approximately 2.9 m/s².

When the synthetic foot switch signal had a value of 1 (the switch is in a pressed state), we updated the acceleration and speed values of the ankle using the following ZUPT conditions:

$$\begin{align*}
A_{\text{Ank}}[i] &= 0, \\
V_{\text{Ank}}[i-1] &= 0, \\
&\text{if } \text{FootSwitch}[i] = 1.
\end{align*} \tag{8}$$

Applying the ZUPT conditions provided in (8) to the estimation of the distance traveled by the ankle (Eq. (5)), resulted in zero displacement whenever the ankle was static on the ground ($D_{\text{Ank}}[i] = 0$, if FootSwitch = 1).

4.4. User bearing

The gyroscope which has its sensitive axis parallel to the superior–inferior direction of the user reported the rate of change of the bearing of the user’s torso. Therefore, the bearing of the user can be calculated by integrating the data received from this gyroscope. We used the calibration data obtained from Eq. (2) to convert the values reported by the gyromouse into degrees, $\theta$, as follows.

$$\theta[i] = \begin{cases} 
R_{\text{Gyro}}[i] \times \text{GyroSens}_R + \theta[i-1], & \text{if } R_{\text{Gyro}}[i] \leq 0 \\
R_{\text{Gyro}}[i] \times \text{GyroSens}_L + \theta[i-1], & \text{if } R_{\text{Gyro}}[i] > 0.
\end{cases} \tag{9}$$

4.5. Position update rule

To calculate the absolute position of the user in the building, we combined the location information obtained independently by the camera and the inertial navigation system. The algorithm for combining the information of both subsystems worked as follows.

(a) If the camera provided reliable localization information through ARToolkit, this information was preferred over the localization information obtained with the inertial navigation system. When a reliable marker was detected by the camera, the position and bearing of the user were updated using the location information from the fiducial marker system. The image obtained from the camera had an inherent delay. We empirically determined that this delay was approximately 0.5 s (or 10 samples at a 20 Hz sampling frequency). During off-line data processing, we compensated for this delay by adding a 0.5 s time offset to the location samples obtained by the camera. If processing in real time, the last 10 position estimates can be updated upon discovery of a recognized marker in the environment.

$$\begin{align*}
\text{Pos}_x[i] &= \text{Pos}_{\text{CAM}}[i+10] \\
\text{Pos}_y[i] &= \text{Pos}_{\text{CAM}}[i+10] \\
\theta[i] &= \theta_{\text{CAM}}[i+10].
\end{align*}$$

A marker was considered recognized if it was not discounted by the following criteria:

i. ARToolkit provides a recognition quality (RQ) estimate for the detected marker. This estimate takes a value between 0 and 1. An RQ value close to 0 indicates that the detected marker location is not recognized. An RQ value close to 1 indicates that the detected marker ID, position and rotation are highly recognizable. We defined an 80% threshold for the RQ estimate. Markers with RQ values under this minimum threshold were discounted.

ii. We discounted detected markers that implied an upside-down orientation of the user (head towards the floor and feet towards the ceiling).

(b) If there were no recognizable markers in the field of view of the camera, the user’s position was updated each sampling period $i$, using the distance traveled by the ankle $D_{\text{Ank}}$ and the user’s body bearing $\theta$ obtained from the inertial sensors in Eqs. (5) and (9). This was done in the following fashion:

i. Calculate the distance traveled by the ankle in the current sampling period $D_{\text{Ank}}[i]$ using Eq. (5), correcting this calculation using the ZUPT procedure presented in Eq. (8).

ii. Calculate the user bearing for the current sampling period $\theta[i]$ using Eq. (9).

iii. Update the top-view position of the user in the building using these values.

$$\begin{align*}
\text{Pos}_x[i] &= D_{\text{Ank}}[i] \cdot \cos(\theta[i]) + \text{Pos}_x[i-1] \\
\text{Pos}_y[i] &= D_{\text{Ank}}[i] \cdot \sin(\theta[i]) + \text{Pos}_y[i-1].
\end{align*}$$
5. Analysis of empirical data

5.1. Performance metrics

We calculated the path traversed by each participant using the proposed system, and we defined six metrics of interest.

1. The mean out-of-bounds error value.
2. The maximum out-of-bounds error value.
3. The difference, in metres, between the final position detected by the proposed system and the real final position.
4. The difference between the final user bearing measured using the proposed system and the real final bearing, in degrees.
5. The total distance measured by the accelerometers during each walking task.
6. The positioning accuracy.

The first two metrics are based on how often the estimated user’s position fell within the boundaries of the predefined circuit during a walking task. An ‘out-of-bounds’ error value was calculated at each sample point. If the detected user position fell within the predefined circuit, the out-of-bounds error was zero. On the other hand, if the detected position was beyond the boundaries of the predefined circuit, the out-of-bounds error was the distance, in metres, to the closest wall of the circuit.

We calculated two versions of the first four metrics, an uncorrected version using solely the inertial sensors to calculate the position of the user, and a corrected version using the camera and the fiducial marker arrangement to adjust the position and bearing estimates during the walking task.

The indoor positioning accuracy was defined as the final position error (metric #3) over the gold standard distance, namely,

\[ \text{Accuracy} = \frac{\text{Final position error}}{\text{Gold standard distance traveled}} \times 100 \]  

where the gold standard distance in the present experiment was 84.4 m. The gold standard distance was obtained as follows. We measured the minimum and maximum distances along the predefined circuit. The minimum distance was measured over the innermost walls of the circuit, using shortcuts (diagonal walking) to minimize the measured distance. The maximum distance was measured along the outermost walls of the predefined circuit. The minimum and maximum distances along the circuit were 74.3 and 94.5 m, respectively. The gold standard distance was taken to be the average of the maximum and minimum distances measured over the predefined path, i.e., \((74.3 + 94.5)/2 = 84.4\).

5.2. Statistical comparisons

To determine if different walking directions, namely clockwise (CW) or counter-clockwise (CCW), had a significant impact on the localization data obtained using the proposed system, we calculated the differences between each pair of corresponding CW and CCW performance metrics per walking task for each participant. Each difference was then divided by the mean of the CW and CCW data for the corresponding walking task. We will refer to these resultant dimensionless numbers as normalized differences. If the performance metrics were similar between walking directions (CW versus CCW), then the expected differences would be close to zero. Therefore, using a t-test, we evaluated the null hypothesis that each set of differences had a mean equal to zero.

We also performed an F-test for the equality of variances of the normalized differences between participants. This would inform us about performance metric variability between subjects.

To ascertain the effect of camera-based location correction, we calculated the difference between each uncorrected performance metric against its corresponding corrected version for each participant. These differences were normalized by dividing by the average of the corresponding corrected or uncorrected measurements. For the first four metrics, we tested the null hypothesis that the mean difference between corrected and uncorrected numbers was zero (t-test), i.e., that the addition of camera correction had no effect on the location estimated by inertial sensors alone.

6. Results

The first five performance metrics for the male and female participants are presented in Tables 1 and 2, respectively. For both male \((p = 0.99)\) and female \((p = 0.44)\) participants, there was no difference between performance metrics in the clockwise or counter-clockwise directions. Further, the variances of the distributions of the normalized differences were statistically similar between participants \((p = 0.29)\). Taken together, these two results suggest that the performance metrics were not statistically associated with the walking direction of the participant.

The addition of camera correction did have a significant effect on the first four performance metrics \((p \ll 0.05)\) for both participants. The average indoor positioning accuracy was 3.38\% of the gold standard distance.

Fig. 3 visually depicts location estimations from three typical walks: an iterative “6 steps and stop” walk (a), a zigzag walk (b), and two iterative “6 steps and stop” walk (a) and (c). In all graphs, the thin solid lines delineate the boundaries of the indoor circuit used in the experiments. The thin, light gray line is the estimated position using the proposed mobile dead reckoning system with inertial sensors only. The thick dark gray line is the estimated position using inertial sensors with camera correction. The dashed dark gray segments indicate where fiducial-based position and orientation correction took place.
Fig. 3. Examples of typical location estimation trials showing boundaries of the indoor circuit, the estimated path using inertial sensors alone (thin light gray line), the estimated path with camera correction (thick dark gray line) and the jumps in estimated position due to camera correction (dashed line). (a) clockwise, “6 steps and stop” walk, male participant; (b) counter-clockwise, zigzag walk, male participant; (c) clockwise, “6 steps and stop” walk, female participant.
7. Discussion

7.1. Key findings

The proposed system uniquely combines mainstream hardware to achieve real-time pedestrian indoor localization using a limited number of markers in the environment. The cost of the sensors used in this work is under $150. This makes the solution affordable for clinical applications such as the rehabilitation of indoor topographical orientation or indoor patient tracking, where GPS is not viable due to limited satellite visibility. The proposed localization system does not require major infrastructure reconfigurations, such as the installation of a wireless network. On the other hand, when wireless infrastructure already exists in the building, the accuracy of the proposed solution could be enhanced further using localization data from the wireless network.

The camera-corrected localization metrics prove significantly different from the corresponding non-corrected metrics. This indicates that the localization accuracy improved significantly using camera correction while using a limited number of markers. We also found that the location metrics obtained by the system are not dependent on the user or walking direction. The final positioning accuracy of 3.38% of the total distance is comparable to the positioning accuracies of 3% and 5.3% of the distance traveled reported in [4] and [5] respectively, both of which used Kalman filtering to estimate the position of the user. Hence, our results demonstrate that with simple processing and generic, off-the-shelf consumer hardware, indoor pedestrian localization can be achieved with accuracies comparable to those published in current literature.

7.2. Study limitations

The accuracy of the proposed system could be further improved by using more complex tracking algorithms, such as Kalman filtering as in [4,9] or algorithms that take into account the structure of the building, such as particle filtering. This could be the subject of future research.

The localization algorithm assumed that the user walks on a horizontal surface. If the vertical location of the user needs to be tracked as well, a barometer could be added to the sensor bank.

The dead reckoning localization algorithm used in this study only takes into account the total acceleration experienced by the ankle, assuming that the step was taken forwards. To determine the direction of the step (i.e., steps taken backwards or laterally), the signals obtained from the different axes of the accelerometer need to be analyzed individually. For example, the proposed system may benefit from a step classification algorithm similar to the one used in [9].

For the experiments presented in this study, we asked participants to walk without excessive pelvic tilt. This restriction can be lifted in the future by adding a third gyroscope to the sensor bank.

We noticed that the distance calculated by the system was often shorter than the minimum distance required to fully traverse the predefined circuit. In fact, 6 distance measurements out of 16 were below this minimum distance. Possible causes include inaccuracies in the simulated foot switch used for zero-velocity updates and the difference between the distance traveled by the foot versus the distance traveled by the ankle. Therefore, location estimation might be improved in future research by developing a more accurate foot switch signal from the accelerometry data, and compensating for the difference in the distance traveled by the foot and the ankle.

Our experiments were conducted with a limited number of walking trials and participants. A larger variety of natural walking tasks should be conducted in the future to verify the finding that the quality of location estimation is indeed consistent across a large sample of users and walking directions.

8. Conclusions

The proposed localization system yields indoor positioning accuracies that are comparable to those of previous works on indoor navigation and outdoor GPS-based localization, while being cost-effective and reliant only on mainstream consumer hardware components. Our system uniquely combines dead reckoning sensors for position estimation and fiducial marker-based image processing localization for dynamic recalibration. The proposed system can estimate user locations for four different level-walking tasks without constraining walking speed or stride length. Future research may aim to further improve the localization accuracy of the proposed system and conduct expanded tests with a larger sample of participants.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada, the Canada Research Chairs Program, Bloorview Children’s Hospital Foundation and Conacyt, Mexico.

References