Pedestrian Inertial Navigation System Augmented by Vision-Based Foot-to-foot Relative Position Measurements

1st Chi-Shih Jao
MicroSystems Lab
University of California, Irvine
Irvine, USA
chishihj@uci.edu

2nd Yusheng Wang
MicroSystems Lab
University of California, Irvine
Irvine, USA
yushengw@uci.edu

3rd Andrei M. Shkel
MicroSystems Lab
University of California, Irvine
Irvine, USA
ashkel@uci.edu

***Abstract*—In this paper, we investigate how self-contained pedestrian navigation can be augmented by the use of foot-to-foot visual observations. The main contribution is a measurement model that uses Zero velocity UpdaTe (ZUPT) and relative position measurements between the two shoes obtained from shoe-mounted feature patterns and cameras. This measurement model provides directly the compensation measurements for the three position states and three velocity states of a pedestrian. The involved features for detection are independent of surrounding environments, thus, the proposed system has a constant computational complexity in any context. The performance of the proposed system was compared to a standalone ZUPT method and a relative-distance-aided ZUPT method. Simulation results showed an improvement in accumulated navigation errors by over 90%. Real-world experiments were conducted, exhibiting a maximum improvement of 85% in accumulated errors, verifying validity of the approach.**

***Index Terms*—Sensor Fusion, Inertial Navigation, Pedestrian Navigation, Vision-aided Inertial Navigation**

I. Introduction

In Global Navigation Satellite System (GNSS) challenging environments, there is a demand for an accurate pedestrian navigation system when GNSS is not available to benefit applications such as localization of first responders, firefighters, and rescuers. In such environments, the positioning often may be also achieved with radio navigation systems using signals including Wireless Local Area Networks (WLAN), Bluetooth, or Long-Term Evolution (LTE) [1]–[3]. However, these systems require a pre-installed infrastructure or surveying the areas of interest for reference signals in advance. In some scenarios, however, assumptions about infrastructural information might not be realistic. Pedestrian Dead Reckoning (PDR), or pedestrian Inertial Navigation Systems (INS), using environment-independent self-contained sensors, are needed for this purpose [4], [5].

The successful development of Micro Electro-Mechanical System (MEMS) technology has enabled pedestrian INS to use small-size Inertial Measurement Units (IMUs). Nevertheless, navigation systems based solely on dead reckoning measurements obtained from Commercial Off-The-Shelf (COTS) MEMS-based IMUs have high drift in estimation of positions because the sensors suffer from a high noise level and uncompensated in-run drift. Zero velocity UpdaTe (ZUPT) is a method widely applied to assist pedestrian INS allowing to effectively constrain error growth in velocity estimations without increasing the system complexity [6]–[9]. The ZUPT is designed based on the observation that during a stance phase of a human gait cycle, the velocity of a foot is very close to zero, and the information is utilized to correct the velocity states in an Extended Kalman Filter framework. However, the ZUPT has systematic errors contributed from two sources: 1) the ZUPT threshold needs to be pre-determined precisely and 2) the actual velocity of a foot in the stance phase is very small, but a nonzero number. These systematic errors can accumulate gradually in long-term navigation [10].

An enhancement mechanism to augment ZUPT for pedestrian INS is to employ a dual foot-mounted IMU system, which has shown to significantly improved navigation results compared to a single foot-mounted IMU and a standalone ZUPT algorithm. Dual-IMU INS benefits from additional measurements derived from the relative motions between the two shoes of a pedestrian, which information will be stacked with the pseudo-measurements from ZUPT in a Kalman filter-based algorithm [11]–[17]. Several different measurements derived from the relative shoe motion were reported. One branch of measurement methods imposes statistical constraints on maximum or minimum distance between the two shoes [11], [15], [16]. These methods formulate the constraint conditions based on a single hypothetical foot motion model, which in general is different for each individual. Another branch uses actual ranging sensors to acquire real-world foot-to-foot distance measurements. Brand et al. proposed a personal navigation system aided by foot-to-foot ranging measurement, which is the relative distance between the two shoes [13]. Their system was later tested by Laverne et al. with foot-to-foot range measurements obtained from two pairs of shoe-mounted SONARs [14]. Although this system has been demonstrated

This work is performed under the following financial assistance award: 70NANB17H192 from US department of Commerce, National Institution of Standards and Technology (NIST).

Authorized licensed use limited to: Access paid by The UC Irvine Libraries. Downloaded on June 12,2020 at 04:00:11 UTC from IEEE Xplore. Restrictions apply.

with excellent navigation results, their measurement model had to be linearized for use of the Extended Kalman Filter (EKF); their model assumed the mismatch between position estimations from IMUs and distance measurements from SONARs to be small, which might not always be the case. In addition to relative distance measurements, the system proposed by Bancroft et al. also used relative positions between the two shoe as updates in the EKF [12]. However, their measurements were obtained based on the assumption that the width between the two shoes is equal to a nominal shoulder width throughout their experiments and the relative height measurements were always zero, which did not sufficiently account for real-world gaits. Our proposed visual-aided dual-IMU Pedestrian INS uses foot-to-foot relative positions between the two shoes. The relative position measurements are obtained by implementing pose estimation algorithms based on consecutive images of a feature pattern mounted on one shoe of our system, captured by a camera mounted on the other shoe. Our system does not impose artificial constraints on the Kalman Filter states and the measurement model does not need to be linearized.

Incorporating motion measurements extracted from a sequence of images with INS has been shown to mitigate errors in the INS due to drawbacks of accelerometers and gyroscopes [18]. Such INS, often known as Visual-aided Inertial Navigation Systems (V-INS), couples the displacement and the orientation measurements of an object of interest obtained from camera frames with measurements of an IMU in a Kalman Filter-based framework [19]–[24]. Efforts have been made to apply V-INS to the field of pedestrian navigation [18], [25]–[31]. Nevertheless, most of these methods performed localization based on feature points extracted from feature patterns found in the surrounding environments and did not fully utilize the knowledge of shoe motion. Feature detection could become challenging for these algorithms in less favorable circumstances. The system proposed by Placer et al. [27] adopted a shoe-mounted marker. They used the marker as a landmark and updated the position of the system based on the landmark when the corresponding images showed that the marker was stationary. However, in practice, the foot velocity is rarely zero during a gait cycle. The usage of a shoe-mounted marker as a fixed landmark might introduce systematic errors to navigation solution. They also determined the stationary status of the shoe-mounted marker by manually observing each collected image, which was not favorable for real-time implementation.

In this paper, we propose a vision-aided pedestrian inertial navigation system that includes two sets of shoe-mounted IMUs, cameras, and feature patterns. The system uses shoe-mounted feature points to estimate the relative positions between the shoes and compensates the position drifts caused by estimations based on IMU measurements. The number of feature points to be used in the system is a fixed quantity. As a result, the computational complexity is constant in any context. Our proposed system uses foot-mounted IMUs and implements the ZUPT to limit the error growth in velocity estimation from IMU measurements between each camera frame. Additionally, the system does not use historical measurements for compensation. Thus, it is possible to achieve a real-time implementation. This paper makes the following contributions:

1) presents a measurement model that uses self-contained vision measurements and ZUPT,
2) provides a mechanism to simulate the foot-to-foot relative position measurements obtained by shoe-mounted cameras,
3) verifies the proposed system with real-world experiments by comparing the results with a standalone ZUPT method and the ZUPT aided by foot-to-foot relative distance measurements.

The rest of the paper is organized as follows. Section II presents the proposed visual-aided pedestrian navigation system. The section also includes the mechanisms for the ZUPT detector and measurements of the relative position. Simulation and experimental results are provided in Section III. Finally, Section IV concludes the paper and suggests future research directions.

II. Proposed Vision-Aided Pedestrian Inertial Navigation System

The proposed system aims to simultaneously track positions and orientations of the two shoe-mounted IMUs $\{b_L\}$ and $\{b_R\}$ in the navigation frame $\{n\}$ in an EKF. Two cameras $\{C_R\}$ and $\{C_L\}$ and two feature patterns $\{f_L\}$ are used to extract measurements for the update step of the EKF. The relationships between all the coordinate frame used in this paper are shown in Fig. 1. Earth rotation effect on the navigation frame is also included in our model. The system has two steps. The first step estimates the position and the orientation of the two IMUs based on the strap-down inertial navigation algorithm [32]. In the second step, the system estimates the pose of the IMUs based on images captured by the cameras and determines the status of each foot with a ZUPT detector. Then, these two measurements are fed to the EKF update step.

A. System configuration

The setup of the system includes two IMUs, two feature patterns, and two cameras. All these components are mounted on the heel side of both shoes, shown in Fig. 2. Returning to Fig. 1, the positions and orientations of the two cameras in the two IMU frames are expressed by two vectors $(q_{b_L}^{C_L}, p_{b_L}^{C_L})$ and $(q_{b_R}^{C_R}, p_{b_R}^{C_R})$. The positions and orientations of the feature patterns in the two IMU frames are expressed by $(q_{b_L}^{f_L}, p_{b_L}^{f_L})$ and $(q_{b_R}^{f_R}, p_{b_R}^{f_R})$. $(q_{C_R}^{f_L}, p_{C_R}^{f_L})$ is the position of the feature pattern from the left shoe in the right camera frame and $(q_{C_L}^{f_R}, p_{C_L}^{f_R})$ is the position of the feature pattern from the right shoe in the left camera frame.

B. Structure of the EKF States

The system keeps track of the states of the two shoe-mounted IMUs with an EKF. The EKF state is a 30×1 vector, described as follows:

$$\bar{x} = [q_n^{b_L}, v_n^{b_L}, p_n^{b_L}, b_a^{b_L}, b_g^{b_L}, q_n^{b_R}, v_n^{b_R}, p_n^{b_R}, b_a^{b_R}, b_g^{b_R}],$$

Authorized licensed use limited to: Access paid by The UC Irvine Libraries. Downloaded on June 12,2020 at 04:00:11 UTC from IEEE Xplore. Restrictions apply.

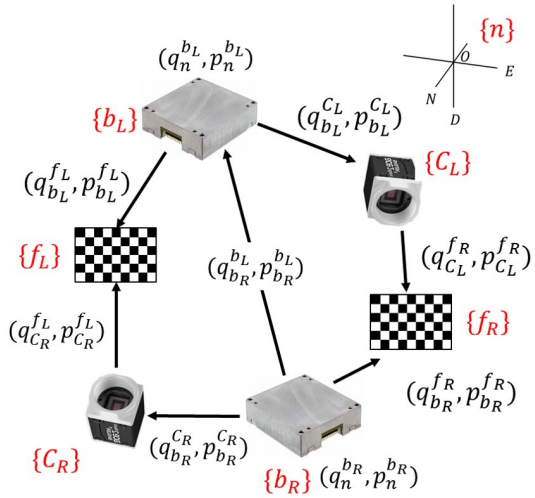

Fig. 1. Relationship between the coordinate frames of different objects in the proposed system.

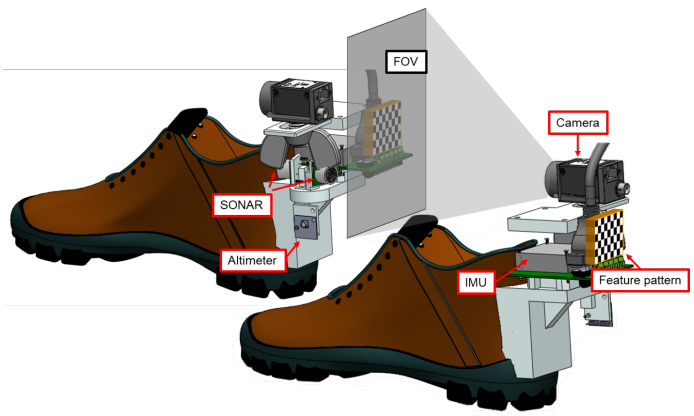

Fig. 2. Lab-on-Shoe system for investigation of self-contained navigation.

where $q_n^{b_L}, v_n^{b_L}, p_n^{b_L}$ are the attitudes, velocities, and positions of the IMU mounted on the left shoe expressed in navigation frame, respectively. $b_a^{b_L}$ and $b_g^{b_L}$ are the bias of the accelerometers and gyroscopes of the left IMU. $q_n^{b_R}, v_n^{b_R}, p_n^{b_R}, b_a^{b_R}$, and $b_g^{b_R}$ indicate the attitudes, velocities, positions, and the bias of the IMU mounted on the right shoe. The error state that are used in the EKF update step is expressed as

$$\delta\bar{x} = [\delta\theta_n^{b_L}, \delta v_n^{b_L}, \delta p_n^{b_L}, \delta b_a^{b_L}, \delta b_g^{b_L}, \delta\theta_n^{b_R}, \delta v_n^{b_R}, \delta p_n^{b_R}, \delta b_a^{b_R}, \delta b_g^{b_R}],$$

Note that the attitude states are expressed in terms of the Euler angle (roll, pitch, yaw) in the error state because we assume that the true attitudes and the estimated attitudes only differ by a small amount. Therefore, according to [20], the error quaternions δq can be approximated by

$$\delta q = [\frac{1}{2}\delta\theta^T, 1]^T,$$

C. Strapdown Inertial Navigation using Dual IMUs

In the prediction step of the EKF, the states of each IMU are propagated according to the standard Strapdown Inertial Navigation [32] and the motions of the two feet are considered to be independent of each other. The linearized continuous-time model of the system state is expressed as follows:

$$\dot{\bar{x}} \triangleq A(t)\bar{x} + B(t),$$

where

$$A(t) = \begin{bmatrix} A_L(t) & O_{15\times15} \\ O_{15\times15} & A_R(t) \end{bmatrix}, B(t) = \begin{bmatrix} B_L \\ B_R \end{bmatrix},$$

with

$$A_L(t) = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & -C(q_n^{b_L}) & 0_{3\times3} \\ [\overrightarrow{f_L}^n\times] & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & C(q_n^{b_L}) \\ 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \end{bmatrix},$$

$$A_R(t) = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & -C(q_n^{b_R}) & 0_{3\times3} \\ [\overrightarrow{f_R}^n\times] & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & C(q_n^{b_R}) \\ 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \end{bmatrix},$$

$$B_L(t) = \begin{bmatrix} C(q_n^{b_L})ARW \\ C(q_n^{b_L})VRW \\ 0 \\ RRW \\ AcRW \end{bmatrix}, B_R(t) = \begin{bmatrix} C(q_n^{b_R})ARW \\ C(q_n^{b_R})VRW \\ 0 \\ RRW \\ AcRW \end{bmatrix},$$

where $[\overrightarrow{f_L}^n\times]$ and $[\overrightarrow{f_R}^n\times]$ are the skew-symmetric cross-product-operator of the accelerometer outputs of the left IMU and the right IMU, expressed in the navigation frame, respectively, ARW is Angle Random Walk of the gyroscopes, VRW is the Velocity Random Walk of the accelerometers, RRW is Rate Angle Walk of the gyroscope, and AcRW is the Acceleration Random Walk of the accelerometers. $C(q)$ is the Directional Cosine Matrix (DCM) corresponding to the quaternion q.

D. Measurement Model

We now present the measurement model for the EKF. The model consists of measurements from the ZUPT and the relative position between the two shoes obtained by the camera capturing the feature pattern mounted on the other shoe.

The measurement extraction methods are described as follows.

Zero velocity UPdaTe (ZUPT):

A stance phase detector is used to determine if a zero velocity measurement should be fed to the system. Different mechanisms to build a stance phase detector have been explored, including using IMU with a pre-determined threshold, IMU and additional sensors, or IMU with adaptive threshold [10]. In this paper, we use the mechanism such that a stance phase

Authorized licensed use limited to: Access paid by The UC Irvine Libraries. Downloaded on June 12,2020 at 04:00:11 UTC from IEEE Xplore. Restrictions apply.

is detected when the summation of variances of gyroscope readout σ_g and that of accelerometer readout σ_a are lower than a specified threshold γ. It can be described as

$$ZUPT\ status = H(\frac{\sigma_a}{\bar{\sigma}_a} + \frac{\sigma_g}{\bar{\sigma}_g} - \gamma),$$

where H() is a Heaviside function, $\bar{\sigma}_g$ and $\bar{\sigma}_a$ are normalized amplitudes of VRW and ARW.

Foot-to-Foot relative position measurement:

We use images that contain the feature pattern to estimate the pose of the camera mounted on the opposite shoe. This method considers a valid measurement when the feature pattern is fully present inside the field of view (FOV) of the camera. For each frame taken by the camera, the position of the feature pattern relative to the camera can be estimated based on the features detected by computer vision-based methods and deduce the relative position between the camera and the feature pattern $(q_{C_R}^{f_L}, p_{C_R}^{f_L})$ and $(q_{C_L}^{f_R}, p_{C_L}^{f_R})$.

We can relate the positions of the two IMUs with the cameras and the feature patterns, shown in Fig. 1. The foot-to-foot relative position measurements, which is the difference in positions of the left IMU and the right IMU expressed in the navigation frame $p_n^{b_L} - p_n^{b_R}$, are derived as follows.

When the left feature pattern is presented in the FOV of the right camera, the position of the left IMU in the navigation frame can be rewritten as

$$\begin{aligned} p_n^{b_L} &= C(q_n^{b_R})p_{b_R}^{b_L} + p_n^{b_R} \\ &= C(q_n^{b_R})\{p_{b_R}^{C_R} + C(q_{b_R}^{C_R})[p_{C_R}^{f_L} \\ &\quad - C(q_{C_R}^{f_L})C^T(q_{b_L}^{f_L})p_{b_L}^{f_L}]\} + p_n^{b_R} \end{aligned}$$

Similarly, when the right feature pattern is presented in the FOV of the left camera, the position of the right IMU in the navigation frame can be expressed as

$$\begin{aligned} p_n^{b_R} &= C(q_n^{b_L})p_{b_L}^{b_R} + p_n^{b_L} \\ &= C(q_n^{b_L})\{p_{b_L}^{C_L} + C(q_{b_L}^{C_L})[p_{C_L}^{f_R} \\ &\quad - C(q_{C_L}^{f_R})C^T(q_{b_R}^{f_R})p_{b_R}^{f_R}]\} + p_n^{b_L} \end{aligned}$$

Thus, the foot-to-foot relative position in the navigation frame is described as

$$\begin{aligned} & p_n^{b_L} - p_n^{b_R} \\ &= C(q_n^{b_R})\{p_{b_R}^{C_R} + C(q_{b_R}^{C_R})[p_{C_R}^{f_L} \\ &\quad - C(q_{C_R}^{f_L})C^T(q_{b_L}^{f_L})p_{b_L}^{f_L}]\} \\ &= -C(q_n^{b_L})\{p_{b_L}^{C_L} + C(q_{b_L}^{C_L})[p_{C_L}^{f_R} \\ &\quad - C(q_{C_L}^{f_R})C^T(q_{b_R}^{f_R})p_{b_R}^{f_R}]\} \end{aligned}$$

Note that the above relation includes two equations. The first equation is used when the left feature pattern is present in the FOV of the right camera, and the second equation is used when the right feature pattern is present in the FOV of the left camera.

The measurement model for the relative position measurements and the ZUPT detector for both IMUs are described as follows:

$$z_{feet} = \left[p_n^{b_L} - p_n^{b_R}\right], z_{ZUPT_L} = \left[v_n^{b_L}\right], z_{ZUPT_R} = \left[v_n^{b_R}\right],$$

and the corresponding measurement matrices are

$$\begin{aligned} & H_{feet} = \begin{bmatrix} O_{3\times6} & I_{3\times3} & O_{3\times6} & O_{3\times6} & -I_{3\times3} & O_{3\times6} \end{bmatrix}, \\ & H_{ZUPT_L} = \begin{bmatrix} O_{3\times3} & I_{3\times3} & O_{3\times9} & O_{3\times15} \end{bmatrix}, \\ & H_{ZUPT_R} = \begin{bmatrix} O_{3\times15} & O_{3\times3} & I_{3\times3} & O_{3\times9} \end{bmatrix} \end{aligned}$$

The form of the EKF measurement model employed depends on what measurements are available. When the measurements are available, we stack them in one measurement vector z to form a single batch-form update equation. Similarly, the batch measurement matrix is formed by stacking the measurement matrices corresponding to the available measurements. For example, in the case where relative position and ZUPT of both feet have measurements, the measurement model is

$$z = \begin{bmatrix} p_n^{b_L} - p_n^{b_L} \\ v_n^{b_L} \\ v_n^{b_R} \end{bmatrix}, H = \begin{bmatrix} H_{feet} \\ H_{ZUPT_L} \\ H_{ZUPT_R} \end{bmatrix}$$

III. Simulation and Experimental results

A. Simulation Results

To validate the proposed visual-aided pedestrian INS, we have performed a series of numerical simulations and compared the results with those using standalone ZUPT and using ZUPT aided by foot-to-foot relative distance. The foot-to-foot relative distance measurements were assumed to be obtained from shoe-mounted SOund Navigation Ranging (SONAR) sensors. In the simulation setup, two hypothetical foot trajectories were generated based on a foot motion model of a pedestrian walking at regular speed straight toward the North for 100 steps, resulting in a total length of $154m$ in $107s$. We made four assumptions for the foot motion: 1) each of the steps for the two feet was utterly identical, 2) the left foot started first, 3) the initial separation distance between the two shoes were $20cm$, and 4) the foot velocities were zero during the entire stance phase. In the simulation model, we also included mismatches of g-sensitivity in the IMUs, which leads to the effect that the trajectory of left shoe drifted towards the west, and the trajectory of the right shoe drifted towards the east. This phenomenon was observed in the experiments reported in [33].

The nominal final locations of the two shoes in the navigation frame were $[154.3, 0, 0]^T\ m$ for the left shoe and $[153.5, 0.2, 30]^T\ m$ for the right shoe. We then produced simulated IMU readouts for both shoes based on the generated paths. The IMU noise characteristics were the same as those of Analog Device ADIS-16485 IMUs used in the experiments (see Section III-B). The relative position measurements were derived by converting the difference in positions of the two shoes in the navigation frame to the body frame of each IMU.

Authorized licensed use limited to: Access paid by The UC Irvine Libraries. Downloaded on June 12,2020 at 04:00:11 UTC from IEEE Xplore. Restrictions apply.

The generated relative position measurements were considered valid only if the position was within the FOV region. In this simulation, we assumed that the cameras have a FOV of 75° and a frame rate of $60Hz$. These two parameters directly affect the amount of valid relative position measurements. The adopted standard deviation of the measurements was $1mm$. For relative distance measurements, we assumed that the SONARs have the same characteristics as those of Devantech SRF08 Ultrasonic Sensors.

We collected 30 sets of simulations and compared the estimated results from our proposed visual-aided INS with those estimated by standalone ZUPT algorithm and by ZUPT aided by foot-to-foot relative distance. Fig. 3 shows the results. The average accumulated errors $\bar{e}_L$ and $\bar{e}_R$ and the covariances along the east direction $\sigma_{x,L}$ and $\sigma_{x,R}$ and the north direction $\sigma_{y,L}$ and $\sigma_{y,R}$ resulting from different methods are summarized in TABLE I. The proposed system showed an average improvement in accumulated errors of more than 90% for both feet compared to the relative distance aided ZUPT method and standalone ZUPT method.

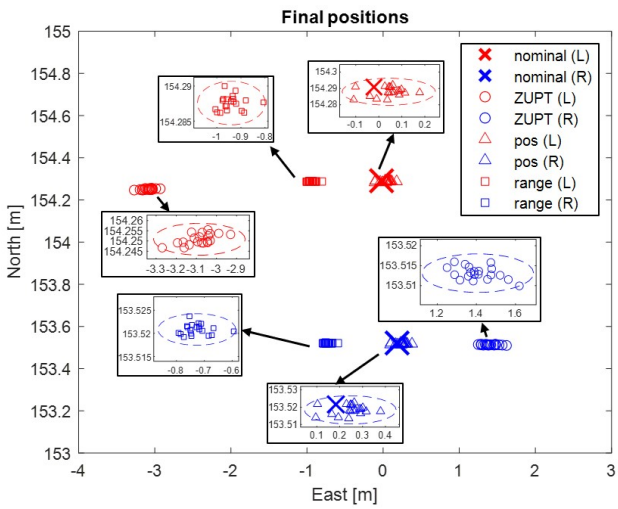

Fig. 3. Simulated results correspond to standalone ZUPT (ZUPT), ZUPT aided by relative distance (range), and ZUPT aided by relative position (pos). Data in red were the estimated final positions of the left shoe, and those in blue were the estimated final position of the right shoe. Zoomed-in views of the data set corresponding to each of the methods are shown next to the data set. The dashed circle around each data set indicates the 3σ limit.

TABLE I
ACCUMULATED ERRORS AND COVARIANCES OF THE SIMULATION DATASET.

Unit [m]	ZUPT aided Measurement models		
	Standalone ZUPT	**Relative distance**	**Relative position**
$\bar{e}_L$	2.9687	0.8687	0.0653
$\bar{e}_R$	1.2354	0.8598	0.0646
$\sigma_{x,L}$	0.0962	0.0654	0.0817
$\sigma_{y,L}$	0.0027	0.0012	0.0014
$\sigma_{x,R}$	0.0841	0.0647	0.0809
$\sigma_{y,R}$	0.0016	0.0012	0.0014

B. Experimental Results

To demonstrate the validity of our system in realistic situations, experiments were conducted with a flexible system integrated with cameras and feature patterns [34]. The performance of our system was evaluated against the standalone ZUPT method and the INS using both ZUPT and foot-to-foot relative distance measurements. The flexible system adopted Analog Device ADIS16485 IMUs and Devantech SRF08 SONARs. The cameras employed in this paper were Gigabit Ethernet (GigE) camera acA800-200gc from the Basler Camera, and the lens of choice had a $4mm$ focal length and a FOV of 73°. The cameras were calibrated by the standard camera calibration method provided by the *MATLAB Toolbox* [35]–[37]. The resolution of the images was 800×600 pixels. Images were recorded at a frame rate of 60 Hz while the IMU provided measurements at 120 Hz. The relative positions between cameras and IMUs, and between feature patterns and IMUs, were determined by computer-aided design (CAD). The SONARs had a sampling rate of 25 Hz and a line-of-sight of 60°.

The feature pattern employed in the experiments was a 6×9 scaled-version checkerboard, which is often used in the standard camera calibration process. We should point out that other reference geometries or features could also be used for detection. The physical size of each grid on the checkerboard was 5×5 mm. The features used for detection were the 40 intersection points on the checkerboard. The feature detection method that we adopted was described in [38]. For the 40 points detected on each frame, we estimated the camera extrinsic matrix by the method presented in [35], and from the extrinsic matrix, we deduced the relative positions between the two shoes. A slice of the sequence of images recorded during one of our indoor walking experiments is shown in Fig. 4.

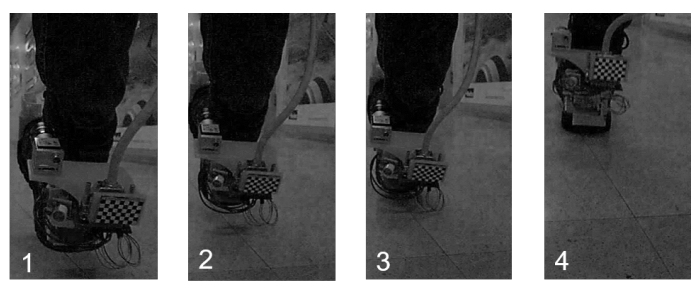

Fig. 4. An example of consecutive images captured by the camera during a walking experiment.

Two sets of experiments with different nominal trajectories were conducted on the second floor of the Engineering Gateway Building at the University of California, Irvine. In the first set of experiments, we conducted 5 runs of indoor experiments of walking straight toward the north for 53 meters. At the end of the experiment, the nominal distance between the two shoes was 30 cm and the nominal final locations were $[53, -0.15, 0]^T$ m for the left shoe and $[53, 0.15, 0]^T$ m for the right shoe. We compared this result to the case of

Authorized licensed use limited to: Access paid by The UC Irvine Libraries. Downloaded on June 12,2020 at 04:00:11 UTC from IEEE Xplore. Restrictions apply.

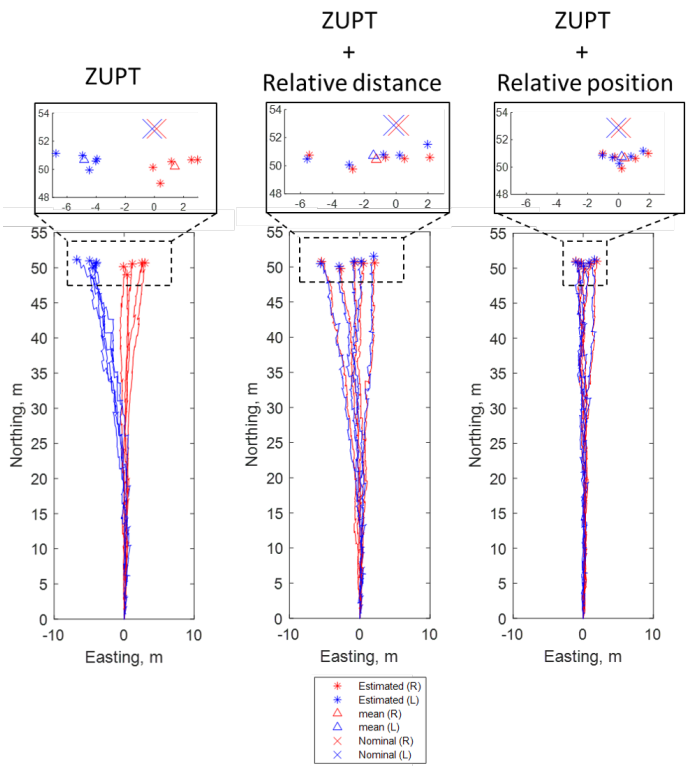

Fig. 5. Estimated results of the first set of experiments from the standalone ZUPT method (ZUPT), ZUPT aided by relative distance measurements (ZUPT + Relative distance), and our proposed system (ZUPT + Relative position). The lower plots show the estimated trajectories, and the upper plots present the corresponding final positions. The triangles in the upper plots indicate the statistical means of each data set.

using relative-distance-aided ZUPT algorithms and that of using standalone ZUPT, shown in Fig. 5. The accumulated navigation errors $\bar{e}_L$ and $\bar{e}_R$ for each case are shown in TABLE II. The proposed system showed improvements in the accumulated navigation error of 55% and 22% for left and right foot when compared with standalone ZUPT; 30% and 31%, when compared with ZUPT aided by relative distance.

TABLE II
ACCUMULATED ERRORS AND COVARIANCES OF THE FIRST SET OF THE EXPERIMENTS

Unit [m]	ZUPT aided Measurement models		
	Standalone ZUPT	**Relative distance**	**Relative position**
$\bar{e}_L$	5.2113	3.3565	2.3589
$\bar{e}_R$	3.2046	3.6173	2.4852

Three things can be noted from the experiments. First, in the plot corresponding to the standalone ZUPT algorithm, we observed that the estimated trajectories of the left shoe drifted toward the west and those of the right shoe to the east. The main factor for this phenomenon was considered to be the mismatch of the g-sensitivity of the IMUs [33]. We did not calibrate the g-sensitivity of the IMUs in this study. Nevertheless, we observed that the trajectories estimated by our proposed system had been shown to mitigate the errors caused by this phenomenon, while the method using foot-to-foot relative distance measurements did not show much improvement. The observation can be explained by the fact that the measurement model of the relative distance in the EKF only optimizes the relative distance between the two shoes. Thus, the effect of such a method is to bring close together the trajectories resulted from the left and the right IMUs. This effect was also observed in Fig. 3. Second, we perceived that all estimated trajectories by the three methods had lengths shorter than $53\ m$. This perception was mainly due to the systematic error in ZUPT, where the velocity of the foot was assumed zero during the stance phase in the gait cycle. While in reality, the foot velocity was not absolutely zero during the stance phase. The false assumption that it is zero led to shorter estimated trajectories. Third, in the experiments, the amount of foot-to-foot relative position measurements obtained in each gait cycle was not consistent. The inconsistency was also another contribution to the difference between experiments and simulations.

In the second set of experiments, we conducted a close loop trajectory, which included more complicated walking motions of four right turns, a ramp, and a short stair. In this experiments, whenever we made a turn, we walked in an arched shape, instead of a direct 90° turn, so that the checkerboard could still be inside the FOV of the camera. The starting point of the left shoe overlapped with the ending point of the right shoe, and vice versa. The nominal total length of the trajectory was $126m$, and the navigation time was $140s$. We compared the estimated results obtained by the three methods. The results are shown in Fig. 6 and the accumulated navigation errors e_L and e_R for each case are summarized in TABLE III. The proposed system showed improvements in the accumulated navigation error of 83% and 85% for each foot when compared to standalone ZUPT and 23% and 54%, when compared with ZUPT aided by relative distance.

TABLE III
ACCUMULATED ERRORS AND COVARIANCES OF THE SECOND SET OF THE EXPERIMENTS.

Unit [m]	ZUPT aided Measurement models		
	Standalone ZUPT	**Relative distance**	**Relative position**
e_L	9.0606	1.9836	1.5337
e_R	5.8524	1.9156	0.8743

We want to point out that the lighting condition in the environment and the exposure time of the camera were key factors in implementation of the proposed system. The lighting condition directly affected the performance of the feature detector. As a result, it is important to have sufficient light sources when the camera is used as a part of the navigator. The exposure time of the camera should also be set appropriately. If the exposure time set too short, the resulting images would not have enough brightness for the feature detection to work well. It cannot be set too long either; since the foot velocity can go up to as fast as $3m/s$, a long exposure time would lead to a quite blurry image. In our experiments, the value was set

Authorized licensed use limited to: Access paid by The UC Irvine Libraries. Downloaded on June 12,2020 at 04:00:11 UTC from IEEE Xplore. Restrictions apply.

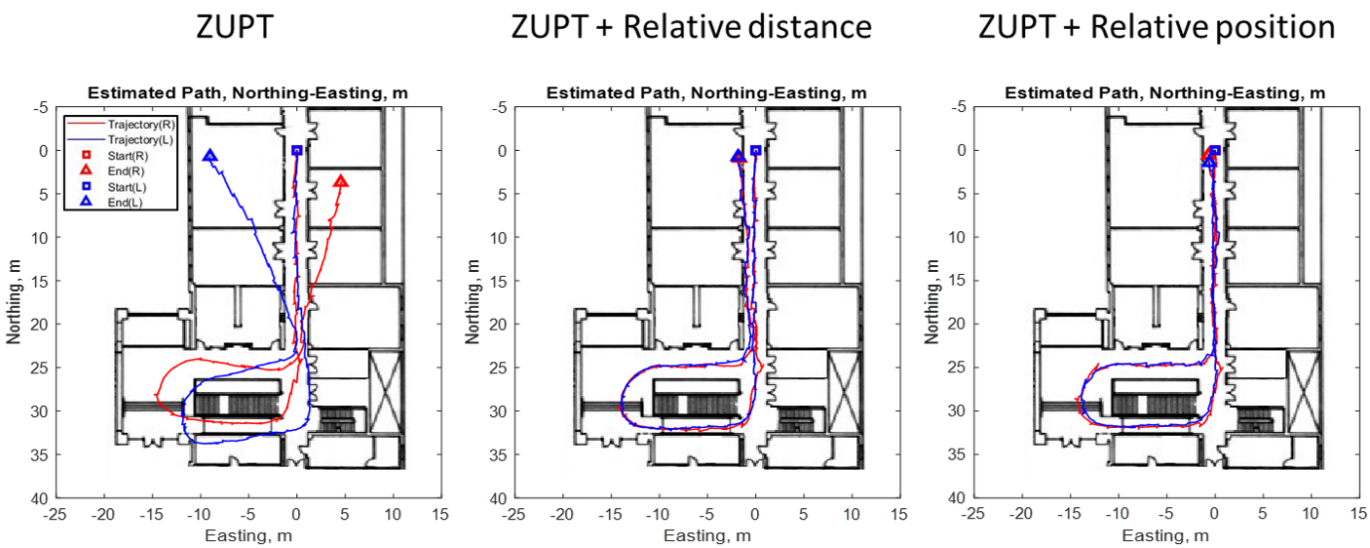

Fig. 6. Estimated results of the second set of experiments from the standalone ZUPT method (ZUPT), ZUPT aided by relative distance measurements (ZUPT + Relative distance), and our proposed system (ZUPT + Relative position).

to 2000 us.

IV. Conclusion

In this paper, we presented a visual-aided Pedestrian INS using foot-to-foot relative position measurements. The main contribution was the measurement model that blends ZUPT and foot-to-foot relative position measurements. The relative position measurements between the two shoes were obtained from shoe-mounted feature patterns and cameras. This measurement model directly outputs compensation measurements for the three position states and three velocity states, and does not need linearization. The proposed system has constant computational complexity in any environment. The simulation results showed an improvement in accumulated navigation errors of over 90%. Experiments were also conducted, where we rendered the shoe-mounted feature pattern to a scaled version of a checkerboard. Experimental results showed a maximum improvement of 85% in accumulated errors, verifying the validity of the proposed system in real-world environments.

References

[1] Y. Chen, R. Chen, L. Pei, T. Kröger, H. Kuusniemi, J. Liu, and W. Chen, "Knowledge-based error detection and correction method of a multi-sensor multi-network positioning platform for pedestrian indoor navigation," in *IEEE/ION Position, Location and Navigation Symposium*, Indian Wells, CA, USA, May 3-6, 2010.

[2] L. Chen, H. Kuusniemi, Y. Chen, L. Pei, T. Kröger, and R. Chen, "Motion restricted information filter for indoor bluetooth positioning," *International Journal of Embedded and Real-Time Communication Systems (IJERTCS)*, vol. 3, no. 3, pp. 54–66, 2012.

[3] A. Abdallah, K. Shamaei, and Z. Kassas, "Indoor localization with LTE carrier phase measurements and synthetic aperture antenna array," in *ION GNSS Conference*, St. Louis, MO, USA, Sep. 21-25, 2019.

[4] X. Yun, E. R. Bachmann, H. Moore, and J. Calusdian, "Self-contained position tracking of human movement using small inertial/magnetic sensor modules," in *IEEE International Conference on Robotics and Automation*, Rome, Italy, Apr. 10-14, 2007.

[5] A. R. Jimenez, F. Seco, C. Prieto, and J. Guevara, "A comparison of pedestrian dead-reckoning algorithms using a low-cost MEMS IMU," in *IEEE International Symposium on Intelligent Signal Processing*, Budapest, Hungary, Aug. 26-28, 2009.

[6] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors," *IEEE Computer graphics and applications*, no. 6, pp. 38–46, 2005.

[7] Y. Wang, A. Chernyshoff, and A. M. Shkel, "Error analysis of ZUPT-aided pedestrian inertial navigation," in *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Nantes, France, Sep. 24-27, 2018.

[8] Y. Wang, D. Vatanparvar, A. Chernyshoff, and A. M. Shkel, "Analytical Closed-Form Estimation of Position Error on ZUPT-Augmented Pedestrian Inertial Navigation," *IEEE Sensors Letters*, vol. 2, no. 4, pp. 1–4, 2018.

[9] Y. Wang, S. Askari, and A. M. Shkel, "Study on Mounting Position of IMU for Better Accuracy of ZUPT-Aided Pedestrian Inertial Navigation," in *IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, Naples, FL, USA, Apr. 1-5, 2019.

[10] Y. Wang and A. M. Shkel, "Adaptive threshold for zero-velocity detector in zupt-aided pedestrian inertial navigation," *IEEE Sensors Letters*, vol. 3, no. 11, pp. 1–4, 2019.

[11] I. Skog, J.-O. Nilsson, D. Zachariah, and P. Händel, "Fusing the information from two navigation systems using an upper bound on their maximum spatial separation," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sydney, Australia, Nov. 13-15, 2012.

[12] J. B. Bancroft, G. Lachapelle, M. E. Cannon, and M. G. Petovello, "Twin IMU-HSGPS integration for pedestrian navigation," in *ION GNSS*, Savannah, GA, USA, Sep. 16-19, 2008.

[13] T. J. Brand and R. E. Phillips, "Foot-to-foot range measurement as an aid to personal navigation," in *Proceedings of the 59th Annual Meeting of The Institute of Navigation and CIGTF 22nd Guidance Test Symposium*, Albuquerque, NM, USA, Jun. 23-25, 2003.

[14] M. Laverne, M. George, D. Lord, A. Kelly, and T. Mukherjee, "Experimental validation of foot to foot range measurements in pedestrian tracking," in *ION GNSS*, Portland, OR, USA, Sep. 20-23, 2011.

[15] X. Niu, Y. Li, J. Kuang, and P. Zhang, "Data Fusion of Dual Foot-Mounted IMU for Pedestrian Navigation," *IEEE Sensors Journal*, vol. 19, no. 12, pp. 4577–4584, 2019.

[16] R. Girisha, G. Prateek, K. Hari, and P. Händel, "Fusing the navigation information of dual foot-mounted zero-velocity-update-aided inertial navigation systems," in *2014 International Conference on Signal Processing and Communications (SPCOM)*. Bangalore, India, Jul. 22-25, 2014.

Authorized licensed use limited to: Access paid by The UC Irvine Libraries. Downloaded on June 12,2020 at 04:00:11 UTC from IEEE Xplore. Restrictions apply.

[17] Y. Wang, S. Askari, C.-S. Jao, and A. M. Shkel, "Directional ranging for enhanced performance of aided pedestrian inertial navigation," in *IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, Naples, FL, USA, Apr. 1-5, 2019.

[18] L. Ruotsalainen, *Vision-aided pedestrian navigation for challenging GNSS environments*, 2013.

[19] G. Huang, "Visual-inertial navigation: A concise review," in *International Conference on Robotics and Automation (ICRA)*, Montreal, Canada, May 20-24, 2019.

[20] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in *IEEE International Conference on Robotics and Automation*, Rome, Italy, Apr. 10-14, 2007.

[21] T. Qin, P. Li, and S. Shen, "VINS-MONO: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.

[22] R. Mur-Artal and J. D. Tardós, "Visual-inertial monocular SLAM with map reuse," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 796–803, 2017.

[23] E. S. Jones and S. Soatto, "Visual-inertial navigation, mapping and localization: A scalable real-time causal approach," *The International Journal of Robotics Research*, vol. 30, no. 4, pp. 407–430, 2011.

[24] J. Kelly and G. S. Sukhatme, "Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration," *The International Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.

[25] C. Hide, T. Botterill, and M. Andreotti, "Low cost vision-aided IMU for pedestrian navigation," in *2010 Ubiquitous Positioning Indoor Navigation and Location Based Service*, Helsinki, Finland, Oct. 14-15, 2010.

[26] G. Panahandeh and M. Jansson, "Vision-aided inertial navigation based on ground plane feature detection," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 4, pp. 1206–1215, 2013.

[27] M. Placer and S. Kovačič, "Enhancing indoor inertial pedestrian navigation using a shoe-worn marker," *Sensors*, vol. 13, no. 8, pp. 9836–9859, 2013.

[28] L. Ruotsalainen, H. Kuusniemi, and R. Chen, "Overview of methods for visual-aided pedestrian navigation," in *2010 Ubiquitous Positioning Indoor Navigation and Location Based Service*, Helsinki, Finland, Oct. 14-15, 2010.

[29] J. Yan, G. He, A. Basiri, and C. Hancock, "Vision-aided indoor pedestrian dead reckoning," in *2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Houston, TX, USA, May 14-17, 2018.

[30] R. Jirawimut, S. Prakoonwit, F. Cecelja, and W. Balachandran, "Visual odometer for pedestrian navigation," *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 4, pp. 1166–1173, 2003.

[31] P. Corke, J. Lobo, and J. Dias, "An introduction to inertial and visual sensing," *The International Journal of Robotics Research*, vol. 26, no. 6, pp. 519–535, 2007.

[32] D. Titterton, J. L. Weston, and J. Weston, *Strapdown inertial navigation technology*. IET, 2004, vol. 17.

[33] Y. Wang, Y.-W. Lin, S. Askari, C.-S. Jao, and A. M. Shkel, "Compensation of Systematic Errors in ZUPT-Aided Pedestrian Inertial Navigation," in *IEEE/ION Position, Location and Navigation Symposium*, Portland, OR, USA, Apr. 20-23, 2020.

[34] S. Askari, C.-S. Jao, Y. Wang, and A. M. Shkel, "A Laboratory Testbed for Self-Contained Navigation," in *IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, Naples, FL, USA, Apr. 1-5, 2019.

[35] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, 2000.

[36] J. Heikkila, O. Silven *et al.*, "A four-step camera calibration procedure with implicit image correction," in *CVPR*, vol. 97. Citeseer, 1997, p. 1106.

[37] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc.", 2008.

[38] A. Geiger, F. Moosmann, Ö. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot," in *2012 IEEE International Conference on Robotics and Automation*, St Paul, MN, USA, May 14-19, 2007.

Authorized licensed use limited to: Access paid by The UC Irvine Libraries. Downloaded on June 12,2020 at 04:00:11 UTC from IEEE Xplore. Restrictions apply.