Vision–IMU Integration Using a Slow-Frame-Rate Monocular Vision System in an Actual Roadway Setting

Duminda I. B. Randeniya, *Member, IEEE*, Sudeep Sarkar, *Senior Member, IEEE*, and Manjriker Gunaratne

***Abstract*—We present results of an effort where position and orientation data from vision and inertial sensors are integrated and validated using data from an actual roadway. Information from a sequence of images, which were captured by a monocular camera attached to a survey vehicle at a maximum frequency of 3 frames/s, is fused with position and orientation estimates from the inertial system to correct for inherent error accumulation in such integral-based systems. The rotations and translations are estimated from point correspondences tracked through a sequence of images. To reduce unsuitable correspondences, we used constraints such as *epipolar lines* and *correspondence flow directions*. The vision algorithm automatically operates and involves the identification of point correspondences, the pruning of correspondences, and the estimation of motion parameters. To simply obtain the geodetic coordinates, i.e., latitude, longitude, and altitude, from the translation-direction estimates from the vision sensor, we expand the Kalman filter space to incorporate distance. Hence, it was possible to extract the translational vector from the available translational direction estimate of the vision system. Finally, a decentralized Kalman filter is used to integrate the position estimates based on the vision sensor with those of the inertial system. The fusion of the two sensors was carried out at the system level in the model. The comparison of integrated vision–inertial-measuring-unit (IMU) position estimates with those from inertial–GPS system output and actual survey demonstrates that vision sensing can be used to reduce errors in inertial measurements during potential GPS outages.**

***Index Terms*—Inertial navigation, intelligent vehicles, multisensor fusion, vision–inertial–Globall Positioning System (GPS) fusion.**

I. Introduction

LONG-TERM error growth in the position and orientation (*pose*) estimates obtained from the integration of the accelerations and angular rates of inertial systems is a major issue that limits the accuracy of inertial navigation. However, inertial systems are known for their precision in short-term and high-frequency data rates. To leverage these advantages, many techniques based on Differential Global Positioning Systems (DGPSs) have been proposed by researchers to be used in conjunction with inertial systems to overcome long-term error growth [1]–[4]. However, intermittent loss of the GPS signal is a common problem encountered in navigation based on GPS-integrated inertial systems [3]. Hence, there is a need for an alternative technology that would ensure smooth and reliable inertial navigation *during GPS outages*.

Due to advances in computer vision, potentially promising studies that engage vision sensing have been carried out in the areas of intelligent transportation systems (ITS) and automatic highway systems [5], [6]. These studies have used sequences of images obtained from a forward-view camera, which was rigidly installed on a vehicle, to estimate the position and orientation (*pose*) of the vehicle [7], [8]. In these studies, the vision system has been used as a supplementary data source to overcome the issue of time-dependent error growth in inertial system estimations.

Most vision-augmented inertial integration has been performed in controlled environments such as laboratories or indoor domains [9], [10]. They used artificial landmarks to simplify the establishment of feature correspondences needed for motion estimation [10], [11]. However, correspondence methods developed indoors do not generalize to outdoor conditions due to the unstructured nature of outdoor scenes. Thus, one cannot expect the algorithms designed for indoor conditions to perform with the same degree of accuracy in an uncontrolled environment. Therefore, it is essential that algorithms designed for uncontrolled (outdoor) environments also incorporate tools to filter out erroneous correspondences. To address this, one needs to employ an additional validation criterion to filter the correspondences based on epipolar geometry and correspondence-based motion fields [12].

Some researchers [13]–[16] have experimented with vision sensors to aid autonomous navigation without fusing the vision sensor to other sensor systems that measure absolute position, such as GPS. However, they do not incorporate inertial navigation system measurements to estimate the position of the vehicle. In Table I, we present a summary of the works that integrate an inertial measuring unit (IMU) with vision for land navigation and differentiate our current work from them.

Roumeliotis *et al.* [17] designed a vision–inertial fusion system for use in landing a space vehicle, using aerial photographs and an IMU. This system was designed using an indirect Kalman filter that is based on the errors in the estimated measurement instead of the direct measurements from camera

Manuscript received June 4, 2007; revised February 3, 2008, July 14, 2009, and November 8, 2009; accepted November 23, 2009. Date of publication January 8, 2010; date of current version May 25, 2010. The Associate Editor for this paper was A. Zelinsky.

D. I. B. Randeniya is with Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA (e-mail: randeniyadi@ornl.gov).

S. Sarkar is with the Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: sarkar@cse.usf.edu).

M. Gunaratne is with the Department of Civil Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: gunaratn@eng.usf.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2009.2038276

TABLE I
PUBLISHED WORK ON IMU–VISION INTEGRATION (X—NO AND √—YES, KF, EKF, MLKF, AND PMM STAND FOR KALMAN FILTER, EXTENDED KALMAN FILTER, MODIFIED LINEAR KALMAN FILTER, AND PARAMETERIZED MODEL MATCHING, RESPECTIVELY)

Reference	Augmented Realty	IMU for camera *pose*	*pose* from camera	INS vision fusion	Filter	IMU nav. equations	Outdoor testing	Camera speed	Control data	Results compared
Roumeliotis et al. (2002)	x	x	√	√	KF	x	x	-	x	-
Chen et al. (2004)	x	x	-	√	KF	√	x	-	√	-
Foxlin et al. (2003)	√	x	-	√	KF	√	x	24 Hz	√	Real-data
You et al. (2001)	√	x	√	√	EKF	√	x	30Hz	√	-
Dial et al. (2005)	x	Rota only	T only	√	MLKF	√	x	10 Hz	x	Raw (Un-aid) data
Huster et al. (2003)	x	√	X	√	EKF		√	-	x	-
Hu et al. (2004)	x	x	√	√	PMM	x	√	-	x	Back projection
This paper	x	x	√	√	KF	√	√	3 Hz	x	Real & Survey data

and IMU systems. According to Roumeliotis *et al.* [17], the Kalman filter algorithms used in conjunction with IMUs can be categorized into two different types, namely, direct and indirect Kalman filters. The aforementioned categorization is based on the measurement vector of the filter. If the filter is directly based on the measurements from the instruments, then it is called a direct Kalman filter. On the other hand, in an indirect filter, the measurement differences, i.e., differences between IMU velocities and GPS velocities, are used as inputs. Two disadvantages of using the direct Kalman filter are the following: 1) The filter has to maintain explicit and accurate awareness of the vehicle's angular motion and attempt to suppress noisy and erroneous data at relatively high frequencies, and 2) the entire navigation algorithm will fail if the filter fails. More information on indirect Kalman filters can be found in [18]–[20]. However, the fusion was performed on the relative *pose* estimated from the two sensor systems, and for this reason, a much simpler inertial navigation model (compared to an absolute *pose* estimation system) was used. Testing was performed on a gantry system designed in the laboratory.

Chen and Pinz [21] investigated motion estimation using a camera, structure from motion, and an inertial system, without updating the IMU error accumulation. The main task of this fusion was to estimate the motion parameters using the vision system as the primary sensor and not the IMU, as is our case.

Foxlin and Naimark [10] used an inertial–vision integration strategy to develop a miniature self-tracker that uses artificial fiducials. Fusion was performed using a bank of Kalman filters designed for acquisition, simple tracking, and hybrid tracking of these fiducials. The IMU data were used to predict the vicinity of the fiducials in the next image. On the other hand, You and Neumann [22] developed an integration method that could be used in augmented-reality applications. The authors used a vision sensor to estimate the relative position, while the rotation was estimated using gyroscopes. Dial *et al.* [9] used an IMU and vision integration system for robot navigation under indoor conditions. Gyroscopes were used to estimate the rotation of the cameras, and the main target of the fusion was to interpret the visual measurements. Finally, Huster *et al.* [8] used the vision–inertial fusion to position an autonomous underwater vehicle (AUV) relative to a fixed landmark. Only one landmark was used in this process, making it impossible to estimate the *pose* of the AUV using a camera, and the IMU system was used to fulfill this task.

The contributions of our work are threefold. First, the vision system used in this work uses a slow-frame-rate monocular camera, which differentiates this work since most research has carried out the use of high-frame-rate cameras and, in some cases, more than one camera. Our constraint introduces a major challenge in ITS applications, since the movement of the camera between two consecutive images is significant, forcing one to use feature-tracking approaches, as opposed to optic flow, to estimate the motion. The motivation behind using slow frame rates lie in that it would act as a baseline accuracy measurement for high-frame-rate cameras and would give necessary tools to perform cost–benefit analysis at implementation. Additionally, high frame rates come at the cost of high power and memory-storage requirement, which can be an issue in many applications such as space exploration.

The second contribution is the introduced fusion architecture, i.e., fusing the vision and the inertial information at the system level, incorporating the error dynamics of both systems. Integrating the two systems at the system level adds two important benefits: 1) the ability to incorporate the nonlinearities of both systems into the fusion algorithm and more accurately estimate the position and orientation of the vehicle and 2) enabling flexibility to use different architectures in the fusion process such as loosely coupled, tightly coupled, and ultratightly coupled architectures. Although fusion using the system approach provides more accurate results for the IMU–GPS systems than with other approaches [23], [24], after an exhaustive literature review, we could not find any literature on the fusion of vision with IMUs that uses this system approach. We used a Kalman filter to fuse the two systems for two reasons: 1) It has extremely efficiently and accurately performed in many other nonlinear processes, exhibiting only small differences when compared with nonlinear filters applied to the same systems, and 2) as this is a preliminary work on fusing vision and inertial systems at the system level, it was desirable to know what a linear filter would yield before resorting to nonlinear filtering algorithms.

The third contribution of this paper is the validation of the fusion algorithm based on data obtained from an actual roadway

setting. Only some literature [8], [10] on vision inertial research reports results in data obtained from an actual setting. The test data for our study were obtained from an unstructured roadway setting with both moving and stationary landmarks, including traffic. Under these conditions, it is impossible to guarantee correct feature correspondences all the time. Errors in correspondences will result in erroneous vision estimates. In this regard, we have found two constraints based on *epipolar lines* and *correspondent flow directions* to be very effective in filtering potentially misleading features.

II. Background for Inertial Navigation

A. Coordinate Frames

The inertial frame (i-frame) is the fundamental reference frame for any navigation. It can be defined as a right-handed coordinate frame based on Newtonian mechanics [25]. The Earth-centered Earth-fixed frame [24], [26], or the e-frame, is a right-handed coordinate frame that is centered at the Earth's center of mass with the third axis parallel to the mean and fixed polar axis and the first axis as the axis connecting the center of mass of the Earth and the intersection of the prime meridian (zero longitude) with the equator.

On the other hand, the navigation coordinate frame (n-frame) is a locally leveled right-handed reference frame with the third axis of the system aligned with the local normal to the Earth's surface (gravity direction), and the first axis is set along the local tangent to the meridian (north).

We express the IMU and vision measurements with respect to the right-handed orthogonal coordinate frames defined for the respective sensor system. The reference frame relevant to the IMU measurements is the b-frame, which has its origin at a predefined location on the sensor and has its first axis toward the forward direction of the vehicle, its third axis toward the direction of gravity, and its second axis toward the right side of the navigation system, composing a right-handed coordinate frame. Similarly, the reference frame for the camera system (c-frame) has its third axis along the principal axis of the camera system, its first axis toward the right side of the image plane, and its second axis set up to make the c-frame a right-handed frame.

B. Inertial Navigation

The IMU in the survey vehicle that we use is a strap-down microelectromechanical system unit [24], [27]. When the vehicle is in motion, the accelerometers measure the specific forces, while the gyroscopes measure the rates of change of rotations of the vehicle in the respective direction and senses [24], [26]. Therefore, it is clear that to geolocate the vehicle, one has to integrate the outputs of the accelerometers and gyroscopes from a known initial position (in a fixed coordinate frame, i.e., e-frame).

The angular rates, which are measured by the gyroscopes, are the rates of change of rotations in the b-frame with respect to the i-frame, i.e., ω_{ib}^b. These can be transformed to rotations with respect to the n-frame [24] by

$$\omega_{nb}^b = \omega_{ib}^b - C_n^b \omega_{in}^n \tag{1}$$

where $\omega_{nb}^b = (\omega_1 \quad \omega_2 \quad \omega_3)^T$, with ω_j representing the angular rates about the jth axis ($j = 1$, 2, and 3), ω_{ib}^b in (1) is the angular rate of the b-frame (IMU) with respect to the i-frame given in the b-frame, and n and b represent the n-frame and the b-frame, respectively. The term C_n^b denotes the transformation matrix from the n-frame to the b-frame. The angular rates of the n-frame with respect to the i-frame, i.e., ω_{in}^n, can be estimated using geodetic coordinates as

$$\omega_{in}^n = \Big((\dot{\lambda} + \omega_e)\cos(\eta) - \dot{\eta} - (\dot{\lambda} + \omega_e)\sin(\eta)\Big)^T \tag{2}$$

where $\dot{\lambda}$ and $\dot{\eta}$ denote the rates of change of the longitude and latitude during vehicle travel, and ω_e is the Earth's rotation rate. The transformation between the n-frame and the b-frame, i.e., C_n^b, can be found in terms of quaternions using the following time propagation equation of quaternions:

$$\dot{q} = \frac{1}{2} A q \tag{3}$$

where q is any unit quaternion that expresses C_n^b, and $\boldsymbol{A}$ is a skew-symmetric transformation matrix [24]

Finally, using (1)–(3), one can obtain the transformation (C_n^b) between the n-frame and the b-frame, in terms of Euler angles, from the gyroscope measurements. However, due to problems inherent in the Euler format, such as singularities at poles and the complexity introduced due to trigonometric functions, quaternions are commonly preferred in deriving the differential equation (3).

The accelerometers inside the IMU measure the specific forces expressed as

$$\ddot{x}^i = g^i(x^i) + a^i \tag{4}$$

where a^i is the specific force measured by the accelerometers in the i-frame, and $g^i(x^i)$ is the acceleration due to the gravitational field, which is a function of the position x^i. From the C_n^b estimated from gyroscope measurements in (3) and the specific force measurements from accelerometers in (4), one can deduce the navigational equations of the vehicle in any frame. Generally, what is desired in terrestrial navigation are 1) the final position; 2) the velocity; and 3) the orientations expressed in the n-frame, although the measurements are made in another local frame, i.e., b-frame. This is not possible since the n-frame also moves with the vehicle, making the vehicle horizontally stationary on this local coordinate frame. Therefore, the desired coordinate frame is the fixed e-frame. Hence, all the navigation solutions are given in the e-frame but along the directions of the n-frame.

Considering the effects of the fictitious forces, i.e., forces due to acceleration in non-i-frames, the equations of motion can be written in the n-frame [24], [26] as

$$\text{Acceleration} \quad \frac{d}{dt} v^n = a^n - (\Omega_{in}^n + \Omega_{ie}^n)\, v^n + g^n \tag{5}$$

$$\text{Velocity} \quad \frac{d}{dt} x^n = v^n. \tag{6}$$

The second and third terms in (5) are, respectively, the Coriolis acceleration and the gravitational acceleration of the vehicle.

The vector multiplication of the angular rates is denoted as $\Omega(=[\omega\times])$ [25]. On the other hand, the orientation of the vehicle can be obtained [24], [26] by

$$\frac{d}{dt}C_b^n = C_b^n \Omega_{nb}^n. \tag{7}$$

In (7), Ω_{nb}^n can be obtained using the ω_{nb}^b estimated in (1). Therefore, once the gyroscope and accelerometer measurements are obtained one can set up the complete set of navigational equations by using (5)–(7). Then, one can estimate the traveling velocity and the position of the vehicle by integrating (5) and (6). The gravitational acceleration can be estimated using the definition of the geoid given in WGS-84 standards [25].

Finally, the positions can be obtained by integrating velocities, which can then be converted to the geodetic coordinate frame as

$$\phi_{(k+1)} = \phi_{(k)} + \frac{(v_N^n)_k \, \Delta t}{(M_k + h_k)} \tag{8}$$

$$\lambda_{(k+1)} = \lambda_{(k)} + \frac{(v_E^n)_k \, \Delta t}{(N_k + h_k)\cos(\phi_k)} \tag{9}$$

$$h_{(k+1)} = h_{(k)} - (v_D)_k \Delta t \tag{10}$$

where v_N, v_E, and v_D are the respective velocities in the n-frame, while ϕ, λ, and h are, respectively, the instantaneous latitude, longitude, and altitude. Moreover, M and N are, respectively, the radii of curvature of the Earth at the meridian and the prime vertical passing through the point on the Earth where the vehicle is located. They are given as follows [24]:

$$N = \frac{p}{\sqrt{(1 - e^2 \sin^2 \phi)}} \tag{11}$$

$$M = \frac{p(1 - e^2)}{(1 - e^2 \sin^2 \phi)^{3/2}} \tag{12}$$

where p is the semimajor axis of the Earth assumed as an ellipsoid, and e is the first eccentricity.

C. IMU Error Model

To develop an accurate navigation solution that accounts for the various biases in accelerometers and gyroscopes, it is important to model the error characteristics of the system. In this paper, we consider only the first-order error terms, implicitly assuming that the higher order terms [25] contribute only to a minor portion of the error. This assumption, which has been used by others as well [24], [25], allows the use of Kalman filtering for the fusion of the vision and inertial systems.

The error dynamics used in this paper were obtained by differentially perturbing the navigation solution [25] by small increments and then considering only the first-order terms of the perturbed navigation solution. Therefore, by perturbing (5)–(7), one can obtain the linear error dynamics for the IMU in the following form [26]:

$$\delta\dot{x} = -\omega_{en}^n \times \delta x^n + \delta\varphi \times v^n + \delta v^n \tag{13}$$

where δ denotes the small perturbation introduced to the position differential equation (6), and φ denotes the rotation vector for the position error. "$\times$" is the vector multiplication of the respective vectors

$$\delta v^n = C_b^n \delta a^b + C_b^n a^b \times \varepsilon + \delta g^n - (\omega_{ie}^n + \omega_{in}^n)\, \delta v^n - (\delta\omega_{ie}^n + \delta\omega_{in}^n) \times v^n \tag{14}$$

where ε denotes the rotation vector for the error in the transformation between the n-frame and the b-frame. The first two terms on the right-hand side of (14) are, respectively, due to the errors in specific force measurement and those in transformation between the two frames, i.e., errors in gyroscope measurements

$$\delta\dot{\Psi} = -\omega_{in}^n \times \varepsilon + \delta\omega_{in}^n - C_b^n \delta\omega_{ib}^b. \tag{15}$$

Equations (13)–(15) are linear with respect to the error of the navigation equation. Therefore, they can be used in a linear Kalman filter to statistically optimize the error propagation.

III. Rotation and Translation From Vision

In revising the IMU *pose* estimation data with respect to their inherent error accumulation problem, one needs to obtain the position and orientation at each camera location. Since the camera is rigidly fixed to the vehicle, these camera positions and orientations can be considered as the vehicle's navigation parameters as well. However, it is essential that the camera be first calibrated for its intrinsic parameters such as the focal length, principal point, and distortion parameters. Furthermore, the fixed transformation between the b-frame and the c-frame [28], [29] must be determined. Once the camera is calibrated for its intrinsic parameters, it can be used in the estimation of motion parameters from a sequence of images. These calibration parameters were estimated in two separate runs performed on two test sections other than the roadway sections that were used to check the validity of the fusion algorithm based on the procedure described in [28] and kept constant in all the subsequent runs. The estimation of motion from a sequence of images is performed by establishing objects that are common to and visible in consecutive images (correspondences). The motion estimation process involves two essential stages [11]. They are

1) the determination of correspondences;
2) the estimation of translation and rotation.

In this paper, we used point features since they are easily detected in a computationally efficient manner. These point features are points in the image that can clearly be distinguished from the rest of the image and can also easily be located in the subsequent images. To establish correspondences between point features, we used the well-established Kanade–Lucas–Tomasi (KLT) feature tracker [30]. This program is completely automated, and the point features are tracked in the sequence of images with replacement. In addition, the parameters have not been customized in this standard and established vision algorithm. The only parameters that we changed were the parameters of the camera, such as the focal length and the number of correspondences. It is

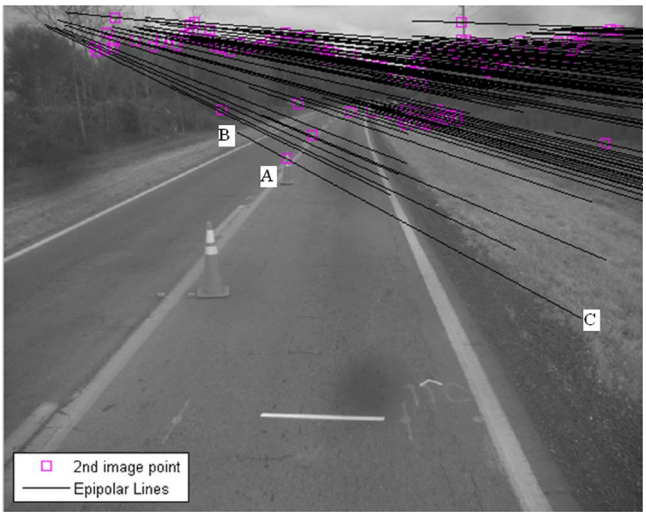


Fig. 1. Epipolar lines drawn for the second-image points.

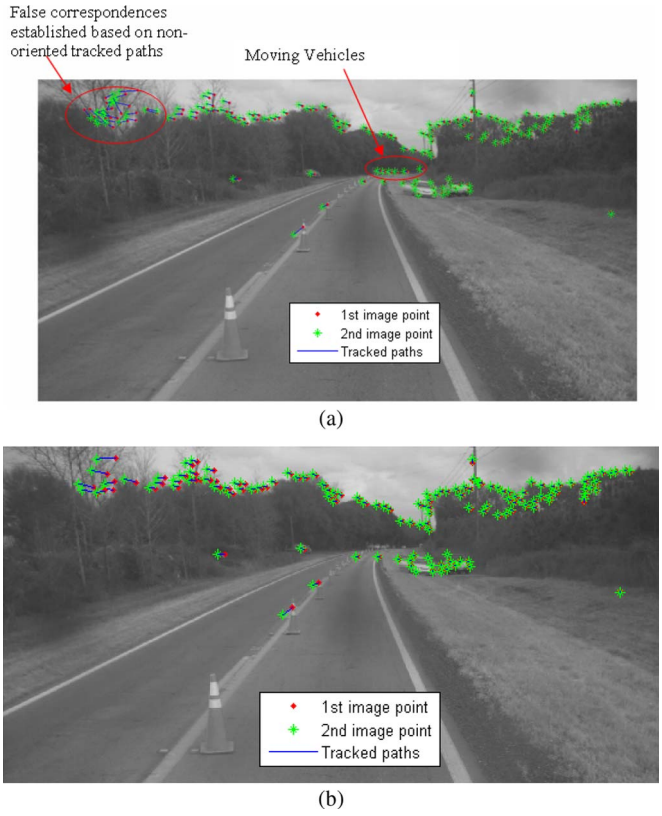


Fig. 2. Filtering point correspondences using motion tracking (a) before (b) after.

understood that there would be some ambiguity involved in determining correspondences arising from the confusion among the features, as well as unreliable tracking constraints [12]. To reduce this ambiguity, one can exploit several constraints. We used two constraints to remove unsuitable correspondences; the constraints were based on 1) epipolar lines (see Fig. 1) and 2) correspondence motion fields [see Fig. 2(a) and (b)]. These methods compare the general motion pattern of the correspondence points in two consecutive images and then remove correspondences with erroneous tracks.

One of the most common constraints in establishing accurate point correspondences in two images is to employ the correspondences that are only matched by epipolar lines. Epipolar lines are established using the first point of a correspondence pair and information from the eight-point algorithm [31]. Such epipolar lines pass through a common point (*epipole*) that corresponds to the image of the center of the camera at the first position on the second image plane. Then, the second point of any correspondence pair (such as a point A in Fig. 1) must lie on the epipolar line (BC in Fig. 1) obtained from its partner point in the first image with the *epipole*. This is only applicable in situations where the locus of the c-frame center is not parallel to any of the consecutive image planes considered, which is a condition that is generally satisfied in autonomous driving tasks. Thus, the importance of an epipolar line lies in that it restricts the placement of any correspondence point on the second image to a straight line obtained using its partner point in the first image (some examples are shown in Fig. 1). Therefore, epipolar lines can be used to filter out the erroneous correspondences from the second image.

The data used in this paper were collected in an outdoor setting, and the captured images contained moving vehicles and other roadway landmarks. Although some of the inaccurate correspondences were removed using the epipolar lines (see Fig. 1), nonstationary features were not completely removed. Hence, the tracked paths of correspondences in consecutive images were utilized. Fig. 2(a) and (b) illustrates how the correspondence flow direction (tracked paths) can be utilized to remove erroneous features from the sequence of images. Once the erroneous feature correspondences are removed, the remaining features that are tracked in more than five images are identified and subsequently used as input to the motion-estimation algorithm to obtain the vehicle's rotation and translation between each of the frames.

A. Motion Parameters From Point Correspondences

The algorithm used to estimate the position and orientation (*pose*) requires at least eight point correspondences that are noncoplanar. A detailed description of the aforementioned eight-point algorithm can be found in [12], [32], and [33].

However, the output of the eight-point algorithm is extremely noisy to be directly used in the fusion algorithm. Therefore, in this paper, a separate Kalman filter was specifically designed for the vision system to filter out the noise in rotations and translations. We will refer to this as the "vision-only Kalman filter" in the following sections.

B. Position Residual From the Two Sensor Systems

Due to the perspective projection of the 3-D points onto the image plane [12], the recovery of the absolute depth is not possible in the vision sensor. Therefore, the translations derived from the vision algorithm are just normalized translations or unit vectors, providing only directional information and not the actual magnitude. However, one of the inputs to the Kalman filter that executes the fusion between the IMU and vision information is the position residual estimated by those

two sensors. Of these, the IMU provides a vector expressing the actual translations, while the vision sensor expresses the translation as a unit vector. Thus, the fusion of vision and inertial measurements requires the aforementioned vision-based translation to be expressed as a regular vector rather than a unit vector. Hence, a special technique had to be devised to obtain the position residual from the unit translations obtained from the vision system.

First, we transform both measurements (IMU and vision) into the e-frame. Then, the translations measured by both sensors are projected onto a unit sphere. Hence, the difference between these two unit vectors that would produce the measurement residuals can be estimated. Finally, as needed by the input to the fusion algorithm, this unit residual has to be multiplied by a distance term. We made this distance a state in the Kalman filter algorithm. The reason for using this distance term as a state of the system model is to account for possible correlations with any sensor measurement when estimating the position residual from the two sensors and, thus, reduce the possibility of error propagation during estimation.

Let P_{IMU} represent the position vector between two consecutive camera locations in the e-frame as estimated by the IMU, and let U_{vis} and U_{IMU} denote the unit translation vectors estimated from the vision system and the IMU, respectively, transformed into the e-frame. First, the transformation of the vision system measurements from the c-frame into the e-frame can be performed as

$$U^e_{\mathrm{vis}}(t_k) = C^e_n(t_k)C^n_b(t_k)C^b_c T^c \tag{16}$$

where C^b_c is the transformation between the c-frame and the b-frame, which can be obtained using the vision system calibration procedure outlined in [28], and C^n_b is estimated using (3). The superscript e indicates that the quantities are expressed in the e-frame. The transformation between the n-frame and the e-frame can be obtained by considering the instantaneous latitude and the longitude. It can be deduced that

$$C^e_n = R_3\left(-\lambda(t_k)\right) R_2\left(\pi/2 + \phi(t_k)\right) \tag{17}$$

where R_i represents the Euler rotation about the ith axis $(i = 2, 3)$. The first camera location can be established as the corresponding IMU position in the e-frame. Then, the position at any other time is estimated using (16). Similarly, the IMU translations can be transformed into the e-frame by

$$P^e_{\mathrm{IMU}}(t_k) = C^e_n(t_k)C^n_b(t_k)T^b_{\mathrm{IMU}} \tag{18}$$

where the term T^b_{IMU} is the translation estimated by the IMU measurements between two consecutive camera locations. The IMU position vector obtained from (18) is then normalized, and the unit vector $U^e_{\mathrm{IMU}}(t_k)$ associated with the IMU is determined. Once the two unit vectors are estimated, the measurement residual, which is an input to the fusion filter, is obtained by

$$dT^e = |\rho^e(t_k)|\left(U^e_{\mathrm{vis}}(t_k) - U^e_{\mathrm{IMU}}(t_k)\right) \tag{19}$$

where $|\rho^e(t_k)|$ is the magnitude estimated from the distances obtained at each time instance the fusion algorithm outputs a position estimate, and dT^e is the required translation residual in the e-frame. $\rho^e(t_k)$ can be given as

$$\rho^e(t_k) = \rho^e(t_{k-1}) + \delta\rho^e(t_{k-1}). \tag{20}$$

$\delta\rho^e(t_{k-1})$ in (20) can be estimated from

$$\left(\delta\rho^e(t_{k-1})\right)^2 = \left(\delta p_\phi(t_{k-1})\right)^2 + \left(\delta p_\lambda\left(t_{k-1}\right)\right)^2 + \left(\delta p_h\left(t_{k-1}\right)\right)^2 \tag{21}$$

where δp_ϕ, δp_λ, and δp_h denote distance differences in the latitude, the longitude, and the altitude, respectively, estimated by the fusion algorithm for time step $(k-1)$.

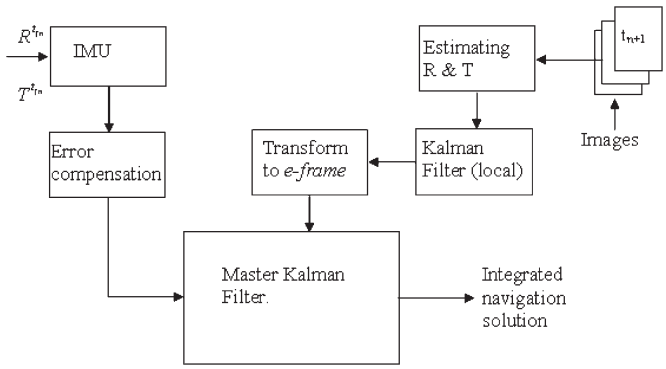


Fig. 3. Schematic of the fusion procedure.

IV. Design of the Kalman Filter

To minimize the error growth in the IMU pose estimations, these IMU estimations have to be updated by an independent measurement at regular intervals. In this paper, vision-based translations and rotations (see Section IV) and a master Kalman filter are used to achieve this objective, thus modeling only the first-order error dynamics for inertial and vision systems. The architecture for this Kalman filter is illustrated in Fig. 3.

A. Design of the Master Kalman Filter

The Kalman filter designed to fuse the IMU and vision pose estimation measurements continuously evaluates the error between the two sensor systems and statistically optimizes it. A decentralized Kalman filter architecture [2], [34], [35] is used in this paper to fuse the two sensor systems. The primary reason for using a decentralized architecture was the necessity for a separate error-estimation process (local Kalman filter) for the vision sensor system. This architecture allows one to integrate the processed vision system parameters with IMU measurements. Since the main aim of the integration of the two systems is to correct the errors in high-frequency IMU pose estimates, the vision system is used due to the stability of the vision system from error accumulation over time. The IMU system is the *process* of the Kalman filter algorithm. Since the two sensor systems possess two distinct data-acquisition frequencies, the multirate fusion approach [36] has been used to fuse the IMU and vision systems. On the other hand, the function of the *vision-only Kalman filter* (local Kalman filter; see Fig. 3) is to remove the significantly high noise associated

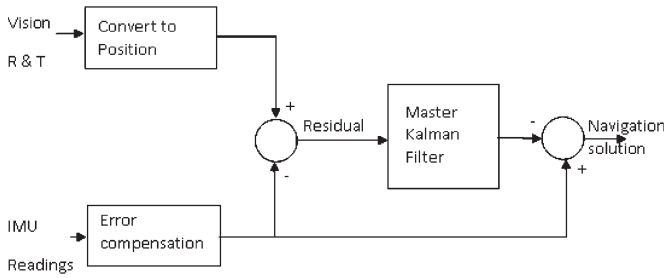


Fig. 4. Master Kalman filter.

with the vision reading because of the relatively high accuracy of measurements demanded by the fusion algorithm. The system architecture of this master Kalman filter is shown in Fig. 4.

The typical inputs to update the master Kalman filter consist of the positions (the e-frame) and the orientations of the b-frame and the c-frame with respect to the n-frame. Since the vision system provides rotations and translations between the relative motion parameters of the camera frames, one needs the absolute position and orientation of the first camera location. For convenience, these can be considered as, respectively, the IMU position in the e-frame and the orientation between the b-frame and the n-frame. The orientation update of the camera between two consecutive images (at t_{k-1} and t_k) with respect to the n-frame can be given as

$$VC_c^n(t_k) = VC_c^n(t_{k-1})R(t_k) \tag{22}$$

where $VC_c^n(t_k)$ is the transformation matrix between the camera orientation with respect to the n-frame at t_k, and $R(t_k)$ is the rotation estimated by the eight-point algorithm at t_k [12]. The IMU used in the test vehicle is a navigational grade IMU that was assumed to be calibrated and aligned. Therefore, the main source of error that could occur in the IMU measurements is the biases of gyroscopes and accelerometers. A more detailed explanation of inertial system errors can be found in [37]. In this paper, only bias errors were considered in error compensation for the IMU. To correct the measurements of gyroscopes and accelerometers, the manufacturer-specified bias terms were used. These bias terms were considered to be propagating in time as

$$b_i(t_{k+1}) = b_i(t_k) + w(t_k) \tag{23}$$

where $b_i(t_k)$ denotes the bias of the ith sensor (accelerometer or gyroscope) at time t_k, and $w(t_k)$ is a random number. The processing system of the Kalman filter consists of the error terms obtained by the perturbation analysis described in Section II. Since, in this paper, only the first-order (linear) errors are considered, the standard Kalman filter equations can be utilized without any linearization. There are 17 system states used for the Kalman filter employed in the IMU–vision integration. These are 1) three states for the position; 2) three states for the velocity; 3) four states for the orientation, which are given in quaternions; 4) six states for accelerometer and gyroscope biases; and 5) distance estimation. Therefore, the entire state vector for the system takes the form (in quaternions) in (24), shown at the bottom of the page, where δ denotes the estimated error in the state, and v_n, v_e, and v_d are, respectively, the velocity components along the n-frame directions, while ϕ, λ, and h are the latitude, the longitude, and the altitude, respectively. The error in the orientation is converted to the quaternion form, and its elements are represented as q_i, where $i = w, x, y, z$. Furthermore, the bias terms in both accelerometers and gyroscopes, i.e., $i = a, g$, along three directions, i.e., $j = x, y,$ and z, are given as b_{ij}. The term ρ^e represents the distance estimated in the Earth's frame. The system equations in the form of the following are used in the Kalman filter process since the measurements from both the IMU and the vision system are discrete:

$$x_k = \boldsymbol{\varphi}_k x_{k-1} + u_k, \quad y_k = \mathbf{H}_k x_k + v_k \tag{25}$$

where x_k is the state matrix, y_k is the measurement at kth time step, and $\boldsymbol{R}_k$ and $\boldsymbol{Q}_k$ are defined as follows:

$$u_k \sim N(0, \mathbf{Q}_k), \quad v_k \sim N(0, \mathbf{R}_k). \tag{26}$$

$\boldsymbol{\varphi}_k$ is the state transition matrix, and $\boldsymbol{H}_k$ is the measurement sensitivity matrix.

The state transition matrix for this problem would be a 17 $\times$ 17 matrix with the terms obtained from (13)–(15). The measurement equation is similarly obtained considering the measurement residual as

$$\mathbf{y_k} = [(\mathbf{P_{vis}} - \mathbf{P_{IMU}}) \quad (\mathbf{\Psi_{vis}} - \mathbf{\Psi_{IMU}})]^T \tag{27}$$

where $\mathbf{P}_i$ and $\mathbf{\Psi}_i$ represent the position vector (3 $\times$ 1) given in geodetic coordinates and the orientation quaternion (4 $\times$ 1), respectively, measured using the ith sensor system ($i =$ vision or IMU). Then, the measurement sensitivity matrix would take the form

$$\mathbf{H_k} = \begin{bmatrix} \mathbf{I_{3\times 3}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I_{4\times 4}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{28}$$

The last critical step in the design of the Kalman filter is to evaluate the process (R_k) and measurement (Q_k) variances of the system. These parameters are quite important in that they define the reliability of the Kalman filter on the system and the measurements [38]. The optimum values for these parameters must be estimated based on the accuracy of the navigation solution. If not, the noisy input will dominate the filter output, making it erroneous. In this paper, to estimate R_k and Q_k, we used a separate data set; one of the three trial runs on the same section that was not used for the subsequent computations. For this purpose, the same Kalman filter was used as a smoother, which was specifically important for the vision measurements since they involve more noise in their measurements.

$$X_k = [\,\delta\phi \quad \delta\lambda \quad \delta h \quad \delta v_n \quad \delta v_e \quad \delta v_d \quad q_w \quad q_x \quad q_y \quad q_z \quad b_{ax} \quad b_{ay} \quad b_{az} \quad b_{gx} \quad b_{gy} \quad b_{gz} \quad \rho^e\,]^T \tag{24}$$

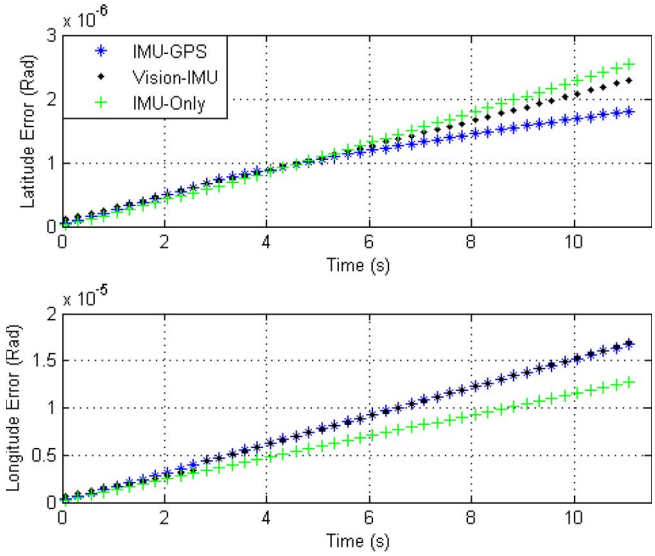


Fig. 5. Comparison of (a) the latitude error and (b) the longitude error for the curve test section.

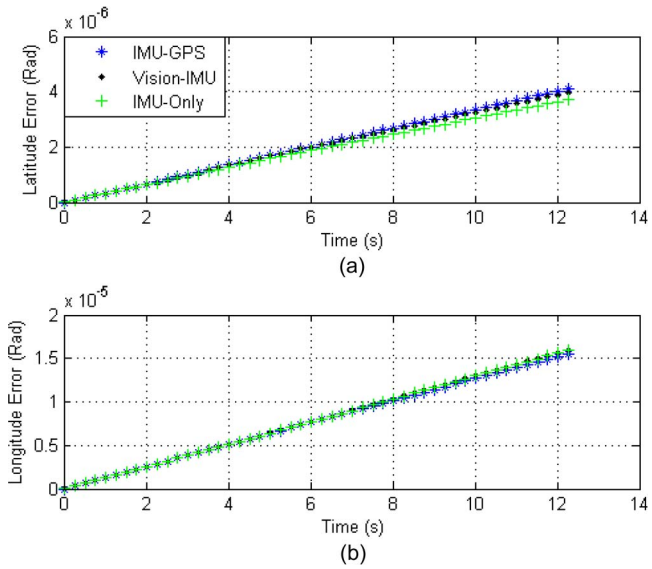


Fig. 6. Comparison of (a) the latitude and (b) the longitude for the straight test section.

V. Results

A. Experimental Setup

The data for the fusion exercise were collected on two test sections on eastbound State Road 26 in Gainesville, FL. One of the test sections was a straight section, whereas the other included a horizontal curve. The curved test section was divided into two separate segments: a short run and another relatively longer run. On the short segment, data from three measuring frames, i.e., survey, vision, and inertial, were collected. The longer section was selected in such a way that it would include the typical geometric conditions encountered on a roadway, such as straight sections, horizontal curves, and vertical curves. Data collected on the longer run were used for the validation of IMU–vision system with IMU–DGPS system data, whereas the short section was used to validate fusion data with the ground truth.

B. Results of the IMU–Vision Integration

The translations and rotations of the test vehicle were estimated from vision sensors using the point correspondences tracked by the KLT tracker on both sections. To estimate the *pose* from the vision system, the correspondences given in Fig. 2 filtered out using the two methods mentioned in Section III were used.

These data were then used in the fusion process to obtain the positions shown in Figs. 5(a) and (b) and 6(a) and (b). From Figs. 5 and 6, we note that the IMU–vision system estimates are closer to those of the IMU–GPS system than the corresponding estimates of the IMU-only system. This is true for both test sections. However, there are some minor deviations associated with the system that can be seen at some occasions. The comparison results are further summarized in Table II in terms of the maximum errors observed at each of the runs.

Table II summarizes the two errors associated with both the IMU–vision and IMU-only systems with respect to the IMU–GPS for the two runs. The data in Table II represent the maximum error occurrence point during the two runs. The number given in the last column is the percentage distance that the IMU-only reading has moved away from the IMU–GPS estimation with respect to the IMU–vision system estimation. It is clear from Table II that the IMU-only data consistently deviate from those of the IMU–GPS system, and this deviation is significant relative to the inertial–vision system. For the curved section, the respective latitude and longitude estimates of the IMU–vision system are 33.6% and 97% closer to the IMU–GPS system than the corresponding estimates of the IMU-only system. For the straight section, the aforementioned figures are 61% and −8.3%, respectively. Hence, the longitude value estimated for the straight section is the only instance that the IMU-only system has better performance. Table II shows that the position estimated by the IMU–vision integration agrees well with those given by the IMU–DGPS integration and consistently performs better than the IMU-only reading.

Figs. 7 and 8 plot the discrepancy between the IMU–vision and IMU–DGPS systems for two test sections. The estimated errors plotted in Figs. 7 and 8 show a maximum error of 12 cm, which is quite acceptable in land-navigation systems. These results clearly show that the IMU–vision system can supplement the IMU measurements without a significant loss in accuracy during a GPS outage.

These results are encouraging since it further signifies the potential use of the vision system as an alternative to GPS in updating IMU errors.

C. Validation of Orientation Results With Ground Truth

Five intermediate points, which were spaced at 11.5 ft, were demarcated and taped in on the shorter test roadway segments in such a way that the vehicle would undergo translations in all three directions and rotations about all three axes (roll, pitch, and yaw) between two consecutive locations.

TABLE II
MAXIMUM ERRORS BETWEEN IMU–GPS, IMU-ONLY, AND IMU–VISION SYSTEMS GIVEN IN FIGS. 5 AND 6

Test Section		IMU-GPS	IMU-Vision		IMU only		Percent Improvement of IMU-Vision system compared to IMU only taking IMU-GPS as the truth
		Value	Value	Difference	Value	Difference	
Curve	Lat	0.5174279	0.5174284	5.068E-07	0.5174286	7.633E-07	33.59595
	Long	-1.4423714	-1.4423713	1.267E-07	-1.4423754	3.971E-06	96.80785
Straight	Lat	0.5174092	0.5174090	1.496E-07	0.5174088	3.902E-07	61.67661
	Long	-1.4424711	-1.44247069	4.531E-07	-1.4424707	4.183E-07	-8.31251

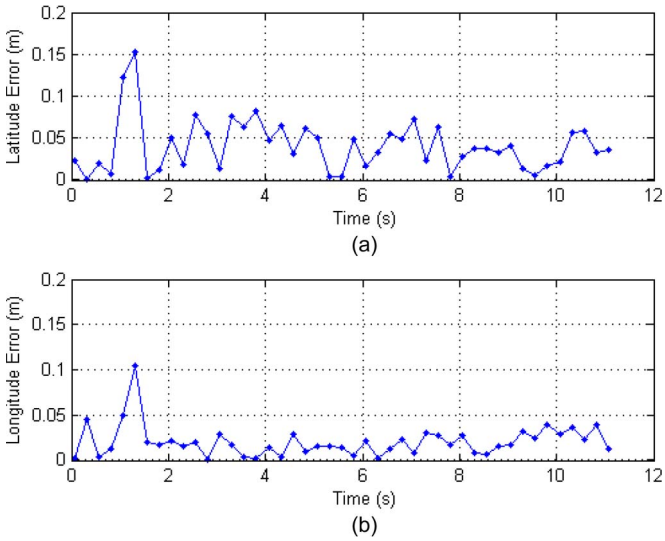


Fig. 7. Error associated with (a) the latitude and (b) the longitude for the curve test section.

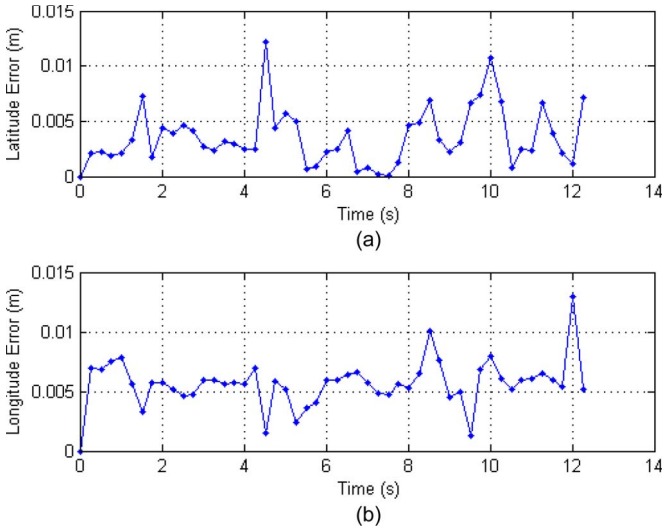


Fig. 8. Error associated with (a) the latitude and (b) the longitude for the straight test section.

A manual survey was performed using a total station that could capture all the locations from one temporary benchmark. At each taped location, four separate points on the cross section of the pavement, i.e., edge, center, and two points in between, were surveyed.

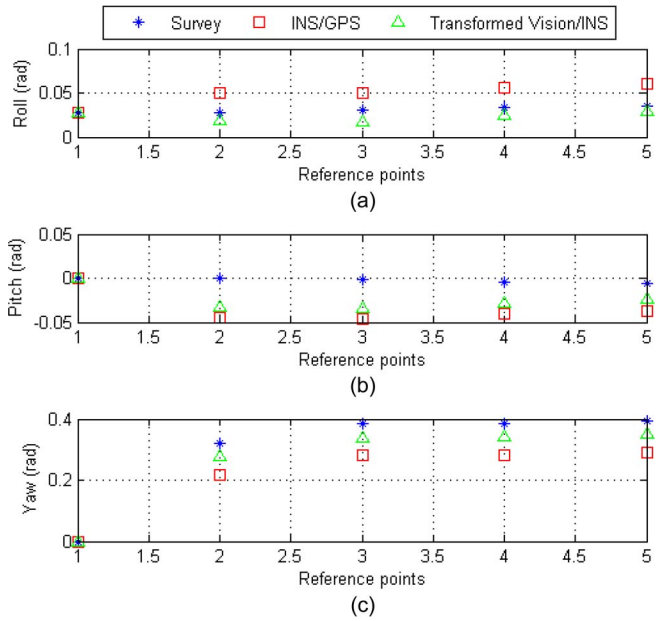


Fig. 9. Comparison of IMU–vision orientations with survey data.

TABLE III
ERROR PERCENTAGES BETWEEN THE IMU–VISION AND IMU–GPS ORIENTATION ESTIMATES WITH MANUAL SURVEY

Rotation	Survey estimate (rad)	IMU/Vision		IMU/GPS	
		Estimate (rad)	Error (%)	Estimate (rad)	Error (%)
Roll	0.0347	0.0296	14.71	0.051	46.97
Pitch	-0.018	-0.023	28.01	-0.030	66.67
Yaw	0.392	0.332	15.31	0.273	30.36

Moreover, at each of these points, the total station measurements were repeated to eliminate any possible errors in the line of collimation. By considering the first surveyed point as the reference, horizontal and vertical angles between each pair of measurement points were estimated. From these measurements, the roll, pitch, and yaw of the vehicle at all consequent measurement points could be estimated with respect to the first point. The aforementioned survey measurements were then compared with the IMU–vision system orientations transformed into the appropriate frame using the transformation found in [28]. Fig. 9 illustrates the comparison between the

TABLE IV
ERROR PERCENTAGES BETWEEN THE IMU–VISION AND IMU–GPS TRANSLATION ESTIMATES WITH MANUAL SURVEY

Survey Point	Distance (ft)			Error (%)		Error in inches	
	Survey	IMU-GPS	IMU-Vision	Survey - IMU/GPS	Survey - IMU/Vision	Survey - IMU/GPS	Survey - IMU/Vision
1	11.614	11.557	11.752	0.494	1.189	0.69	1.66
2	11.535	11.445	11.630	0.777	0.825	1.08	1.14
3	11.606	11.578	11.849	0.246	2.093	0.34	2.92
4	11.541	11.264	11.622	2.402	0.702	3.33	0.97

IMU–DGPS, IMU–vision, and survey orientations for the short test section.

In Fig. 9, one sees that the IMU–vision orientations are relatively closer to the survey orientations than those of IMU–DGPS. Since these results are obtained from a single test on an actual roadway, one must be cautious and employ further experimentation in the future to generalize the aforementioned conclusion. Table III shows the maximum error percentages for the orientations estimated from the IMU–vision and IMU–GPS systems with respect to the survey for the curved test section. It is clear from Fig. 9 and Table III that the maximum percentage errors are quite satisfactory, considering the relatively large distance interval at which the images were captured.

Table III shows that IMU–vision system predicts better orientations than the IMU–GPS system compared with the survey orientations. Table IV shows the errors estimated from the translation predictions of IMU–vision and IMU–GPS systems with respect to the manual survey, which are given as percentages and in inches.

It is clear that the deviation of translations in both cases is small and that the IMU–GPS estimation is closer to the survey measurement in most instances than IMU–vision estimation. However, if one considers the discrepancy in inches, it is mostly within an inch, which is acceptable in autonomous driving tasks under loss of GPS signals. It must be noted that, in the fourth row, the IMU–vision comes closer to the survey reading than the IMU–GPS estimation.

Finally, Tables III and IV also compare the orientation and translation measurements of the IMU–vision and IMU–GPS systems with respect to those of the actual survey. We can see that, for this particular test run, the IMU–vision system estimates of the orientations are closer to the survey measurements than those measured from the IMU–GPS system.

VI. CONCLUSION

It has been shown that a vision system attached to a vehicle can be used to estimate the rotations and translations of that vehicle using a sequence of images. The results have also showed that vision data can be used to update the IMU measurements and control the error growth in such systems. The fusion of IMU–vision measurements has been performed using a sequence of images obtained on an actual roadway and compared with the IMU–DGPS readings, since the main objective of this paper was to explore a reliable alternative system that yields results very close to IMU–DGPS results in situations where the GPS signal is unavailable. The orientations and translations obtained from the fused system have been validated with respect to a manual survey performed at the section. The accuracy of the IMU–vision integrated system can further be improved by a more frequently captured image sequence. The use of accurate correspondences is essential in successfully executing the vision algorithm. We have found that significant improvements can be made by employing constraints such as epipolar lines and correspondence motion fields to prune potentially false correspondences.

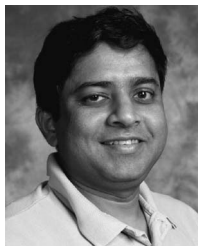
REFERENCES

[1] M. Cramer, GPS/INS Integration, Dec. 2005. [Online]. Available: http://www.ifp.uni-stuttgart.de/publications/phowo97/cramer.pdf
[2] M. Wei and K. P. Schwarz, "Testing a decentralized filter for GPS/INS integration," in *Proc. IEEE Plans Position Location Navig. Symp.*, Mar. 1990, pp. 429–435.
[3] S. Feng and C. L. Law, "Assisted GPS and its impact on navigation transportation systems," in *Proc. 5th IEEE Int. Conf. ITS*, Singapore, Sep. 3–6, 2002, pp. 926–931.
[4] D. Bevly, J. Ryu, and J. C. Gerdes, "Integrating INS sensors with GPS measurements for continuous estimation of vehicle sideslip, roll, and tire cornering stiffness," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 483–493, Dec. 2006.
[5] R. Gregor, M. Lutzeler, M. Pellkofer, K. H. Siedersberger, and E. D. Dickmanns, "EMS-Vision: A perceptual system for autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 1, pp. 48–59, Mar. 2002.
[6] M. Bertozzi, A. Broggi, A. Fascioli, and S. Nichele, "Stereo vision-based vehicle detection," in *Proc. IEEE Intell. Veh. Symp.*, 2000, pp. 39–44.
[7] A. D. Sappa, F. Dornaika, D. Ponsa, D. Geronimo, and A. Lopez, "An efficient approach to onboard stereo vision system POSE estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 3, pp. 476–490, Sep. 2008.
[8] A. Huster and S. Rock, "Relative position sensing by fusing monocular vision and inertial rate sensors," in *Proc. 11th Int. Conf. Adv. Robot.*, Coimbra, Portugal, 2003, pp. 1562–1567.
[9] D. Dial, P. DeBitetto, and S. Teller, "Epipolar constraints for vision-aided inertial navigation," in *Proc. IEEE WACV/MOTION*, 2005, pp. 221–228.
[10] E. Foxlin and L. Naimark, "VIS-Tracker: A wearable vision-inertial self-tracker," in *Proc. IEEE VR*, Los Angeles, CA, 2003, pp. 199–206.
[11] Z. Hu, U. Keiichi, H. Lu, and F. Lamosa, "Fusion of vision, 3D gyro and GPS for camera dynamic registration," in *Proc. 17th Int. Conf. Pattern Recog.*, 2004, pp. 351–354.
[12] O. Faugeras, *Three Dimensional Computer Vision: A Geometric Viewpoint*., 2nd ed. Cambridge, MA: MIT Press, 1996.
[13] M. Sotelo, F. Rodriguez, and L. Magdalena, "VIRTUOUS: Vision-based road transportation for unmanned operation on urban-like scenarios," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 2, pp. 69–83, Jun. 2004.
[14] M. Bertozzi and A. Broggi, "GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 62–81, Jan. 1998.
[15] M. Bertozzi, A. Broggi, and A. Fascioli, "Vision-based intelligent vehicles: State of the art and perspectives," *Robot. Auton. Syst.*, vol. 32, no. 1, pp. 1–16, 2000.
[16] J. C. McCall and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 20–37, Mar. 2006.
[17] S. I. Roumeliotis, A. E. Johnson, and J. F. Montgomery, "Augmenting inertial navigation with image-based motion estimation," in *Proc. IEEE Int. Conf. Robot. Autom.*, Washington, DC, 2002, pp. 4326–4333.

[18] E. J. Lefferts and F. L. Markley, "Dynamics modeling for attitude determination," presented at the AIAA Paper 76-1910, San Diego, CA, Aug. 1976, 10 pp.
[19] S. I. Roumeliotis, "Robust mobile robot localization: From single-robot uncertainties to multi-robot interdependencies," Ph.D. dissertation, Elect. Eng. Dept., Univ. Southern California, Los Angeles, May 2000.
[20] S. I. Roumeliotis, G. S. Sukhatme, and G. A. Bekey, "Circumventing dynamic modeling: Evaluation of the error-state Kalman filter applied to mobile robot localization," in *Proc. IEEE Int. Conf. Robot. Autom.*, Detroit, MI, May 10–15, 1999, vol. 2, pp. 1656–1663.
[21] J. Chen and A. Pinz, "Structure and motion by fusion of inertial and vision-based tracking," in *Proc. 28th OAGM/AAPR Conf.*, 2004, vol. 179, pp. 55–62.
[22] S. You and U. Neumann, "Fusion of vision and gyro tracking for robust augmented reality registration," in *Proc. IEEE Conf. Virtual Reality*, 2001, pp. 71–78.
[23] M. Grewal and A. Andrews, *Kalman Filtering Theory and Practice Using MATLAB*., 2nd ed. New York: Wiley, 2001.
[24] C. Jekeli, *Inertial Navigation Systems With Geodetic Applications*. Berlin, Germany: Walter de Gruyter, 2000.
[25] E. Shin, "Estimation of techniques for low cost inertial navigation," Ph.D. dissertation, Univ. Calgary, Calgary, AB, Canada, 2005.
[26] D. H. Titterton and J. L. Weston, "Strapdown inertial navigation technology," in *IEE Radar, Sonar, Navigation and Avionics*, E. D. R. Shearman and P. Bradsell, Eds. London, U.K.: Peregrinus, 1997, ser. 5, pp. 24–56.
[27] N. Barbour and G. Schmidt, "Inertial sensor technology trends," *IEEE Sensors J.*, vol. 1, no. 4, pp. 332–339, Dec. 2001.
[28] D. Randeniya, M. Gunaratne, S. Sarkar, and A. Nazef, "Calibration of inertial and vision systems as a prelude to multi-sensor fusion," *Transp. Res., Part C Emerg. Technol.*, vol. 16, no. 2, pp. 255–274, Apr. 2008.
[29] S. Nedevschi, C. Vancea, T. Marita, and T. Graf, "Online extrinsic parameters calibration for stereovision systems used in far-range detection vehicle applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 4, pp. 651–660, Dec. 2007.
[30] S. Birchfield, KLT: An implementation of the Kanade–Lucas–Tomasi Feature Tracker, Nov. 2006. [Online]. Available: http://www.ces.clemson.edu/~stb/klt/
[31] D. Randeniya, M. Gunaratne, and S. Sarkar, "Fusion of vision inertial data for automatic geo-referencing," in *Knowledge Discovery From Sensor Data*, A. R. Ganguly, J. Gama, O. A. Omitaomu, M. M. Gaber, and R. R. Vatsavai, Eds. Boca Raton, FL: CRC, Dec. 2007.
[32] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Englewood Cliffs, NJ: Prentice-Hall, 2003.
[33] W. van der Mark and D. M. Gavrilla, "Real-time dense stereo for intelligent vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 38–50, Mar. 2006.
[34] D. J. Allerton and H. Jia, "A review of multisensor fusion methodologies for aircraft navigation systems," *J. Navig.*, vol. 58, no. 3, pp. 405–417, Sep. 2005.
[35] D. J. Allerton and A. J. Clare, "Sensor fusion methods for synthetic vision systems," in *Proc. 23rd DASC*, Oct. 2004, pp. 4.A.3-1–4.A.3-13.
[36] L. Armesto, S. Chroust, M. Vincze, and J. Tornero, "Multi-rate fusion with vision and inertial sensors," in *Proc. Int. Conf. Robot. Autom.*, Apr. 2004, pp. 193–199.
[37] M. Grewal, L. Weill, and A. Andrews, *Global Positioning Systems, Inertial Navigation and Integration*. New York: Wiley, 2001.
[38] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.

Duminda I. B. Randeniya (M'08) received the B.Sc.(Eng.) degree in mechanical engineering from the University of Peradeniya, Peradeniya, Sri Lanka, in 2001, the M.S. degree in applied mathematics, with a concentration on control theory, from Texas Tech University, Lubbock, in 2003, and the Ph.D. degree in intelligent transportation systems (ITS), with a concentration on multisensor fusion, from the University of South Florida, Tampa, in 2007.

Since September 2007, he has been with the Oak Ridge National Laboratory, Oak Ridge, TN. His research interests include ITS, multisensor fusion, network analysis, forecasting models, and risk and vulnerability analysis.



Sudeep Sarkar (S'90–M'92–SM'05) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1988 and the M.S. and Ph.D. degrees in electrical engineering, on a University Presidential Fellowship, from the Ohio State University, Columbus, in 1990 and 1993, respectively.

Since 1993, he has been with the Department of Computer Science and Engineering, University of South Florida (USF), Tampa, where he is currently a Professor. He is a coauthor of the book *Computing Perceptual Organization in Computer Vision* (World Scientific). He is also a Coeditor of the book *Perceptual Organization for Artificial Vision Systems* (Kluwer). He was on the editorial board of the *Pattern Analysis & Applications* journal during 2000–2001 and is currently on the editorial boards of the *Pattern Recognition* journal, *Image and Vision Computing*, and *IET Computer Vision*. His research interests include perceptual organization in single images and multiple image sequences, automated sign language recognition, biometrics, and nanocomputing.

Dr. Sarkar is a Fellow of the International Association for Pattern Recognition. He is the received the National Science Foundation CAREER award in 1994, the USF Teaching Incentive Program Award for undergraduate teaching excellence in 1997, the Outstanding Undergraduate Teaching Award in 1998, and the Ashford Distinguished Scholar Award in 2004. He was on the editorial board of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE from 1999 to 2003. He is currently serving on the editorial board of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, PART B.

Manjriker Gunaratne received the B.Sc.(Engr.) degree from the University of Peradeniya, Peradeniya, Sri Lanka, the M.A.Sc. degree from the University of British Columbia, Vancouver, BC, Canada, and the Ph.D. degree from Purdue University, West Lafayette, IN.

He is currently a Professor with the Department of Civil Engineering, University of South Florida, Tampa. His areas of expertise are geotechnical and highway engineering.