International Journal of Control, Automation, and Systems (2011) 9(4):716-725 http://www.springer.com/12555
DOI 10.1007/s12555-011-0413-y

Localization using GPS and VISION aided INS with an Image Database and a Network of a Ground-based Reference Station in Outdoor Environments

Ji-Hoon Choi, Yong-Woon Park, Jae-Bok Song*, and In-So Kweon

Abstract: GPS/INS integrated systems do not guarantee robustness and accuracy of localization, because GPS has vulnerability to external disturbances. However, the overall performance and reliability of the system can be significantly improved by fusing multiple sensors with a different operating principle. In outdoor environments where GPS may be blocked, there are many features compared to the open space and these features can provide much information for UGV localization. Thus, this paper proposes an improved localization algorithm based on the hierarchical federation of three measurement layers, i.e., GPS, INS, and visual localization, to overcome the shortcomings of GPS/INS integrated systems. The proposed algorithm automatically switches the operation modes according to GPS status and a network of a ground-based reference station. A vocabulary tree with SURF is used in the visual localization method. In the data fusion of visual localization and INS, an asynchronous and time-delayed data fusion algorithm is presented because visual localization is always time-delayed compared with INS. By using DGPS to obtain the reference position under the dynamic conditions of the reference station, the restrictions of the conventional DGPS are overcome and all UGVs within WiBro communication range of the reference station can accurately estimate the position with a common GPS. The experiment results with a predefined path demonstrate enhancement of the robustness and accuracy of localization in outdoor environments.

Keywords: Image matching, localization, operation mode, reference station, vocabulary tree.

1. INTRODUCTION

Localization plays an important role in the autonomous navigation of unmanned ground vehicles (UGVs). INS (Inertial Measurement System) and GPS (Global Positioning System) are commonly used in outdoor environments. INS provides high-rate position, velocity, and attitude data with good short-time stability while GPS provides position and velocity data with good long-term stability. With the data fusion of GPS and INS, highly accurate localization can be achieved. If GPS is subject to signal outage or jamming, however, the performance of an INS/GPS integrated system may be degraded and localization error will increase steeply. Various approaches to overcome this problem have thus been suggested.

The fusion of INS/GPS/odometer information is efficient and cost-effective. Since an odometer is a velocity sensor that is self-contained and is relatively disturbed, it compensates the navigation error even in the environments where GPS is unavailable. However, it is very sensitive to error such as wheel slippage [1]. A magnetometer can be an efficient solution for vehicle control with an INS/GPS integrated system. The trajectory upon which the vehicle moves is formed by magnets. Since the magnetometer measurements compensate the INS position, magnetometer aided INS accurately maintains the vehicle lateral position even in tunnels or urban canyon situations when GPS in the direction normal to the trajectory is temporarily unavailable [2]. A digital map can also be used to improve vehicle localization accuracy. Using the road geometry information obtained from a digital map database (DB), some constraints between the vehicle states can be obtained. The vehicle states are estimated by a GPS/INS integrated system, and the estimates are then projected into the state constraints derived from the digital map DB [3]. INS can be fused with multi-aiding information such as vehicle constraints [4], used as a "virtual sensor," together with the encoders and a single-axis gyroscope. The aiding systems provide not only velocity observations from the encoders and the virtual sensor, but also attitude observation from the gyroscope.

Manuscript received November 8, 2009; revised August 5, 2010 and March 28, 2011; accepted May 2, 2011. Recommended by Editorial Board member Sooyeong Yi under the direction of Editor Hyun Seok Yang. This work was partly supported by the Dual-Use Technology Program of DAPA/DUTC and MIC/IITA in Korea

Ji-Hoon Choi is with the Department of Mechanical Engineering, Korea University, 5, Anam-dong, Sungbuk-gu, Seoul 136-713, Korea, and he is also with UGV Technology Directorate at Agency for Defense Development, Yuseong P. O. BOX 35-5, Daejeon 305-600, Korea (e-mail: ji_hoon_choi@add.re.kr).

Yong-Woon Park is with UGV Technology Directorate, Agency for Defense Development, Yuseong P. O. BOX 35-5, Daejeon 305-600, Korea (e-mail: yongpark@add.re.kr).

Jae-Bok Song is with the Department of Mechanical Engineering, Korea University, 5, Anam-dong, Sungbuk-gu, Seoul 136-713, Korea (e-mail: jbsong@korea.ac.kr).

In-So Kweon is with the Department of Electrical Engineering, KAIST, 373-1, Guseong-dong, Yuseong-gu, Daejeon 305-701, Korea (e-mail: iskweon@ee.kaist.ac.kr).

* Corresponding author.

Fusing the multi-aiding information with INS, the INS errors can be constrained to a reasonable level [5].

This paper is focused on the development of a navigation system capable of operating autonomously and accurately under the challenges of outdoor environments, which will be applied to a UGV, Dog-Horse Robot (DHR). The major requirement for localization in this project is as follows: localization should be robust and accurate even in cases where GPS is unavailable in known environments.

In outdoor environments where GPS may be blocked, there are many features compared to the open space. Since these features can provide very useful information for UGV localization, visual localization based on image matching and pose estimation can lend an important contribution in these environments.

This paper proposes an improved localization algorithm based on the hierarchical federation of three measurement layers, i.e., GPS, INS, and visual localization, to overcome the shortcomings of GPS/INS integrated systems. INS is a reference system, and GPS and visual localization become aiding sources to bound the INS errors. This algorithm automatically switches the operation modes according to GPS status and a network of a ground-based reference station: In the open space, GPS aided INS is used for localization. If the uncertainty from GPS becomes large, VISION aided INS is used. A vocabulary tree [6] with SURF (Speeded Up Robust Feature) [7] is used for fast and efficient matching and the searching process in visual localization. The localization system also includes a FDI (fault detection and isolation) algorithm for system safety and Bierman's U-D factorization algorithm [8] for numerical stability.

Time synchronization between sensor data should be considered very carefully in filter design. GPS and INS estimate the real-time position at 1Hz and 100Hz, respectively. However, visual localization estimates the time-delayed position due to the processing time for image matching and pose estimation. As the velocity of the UGV increases, the travel distance of the UGV during the delay time will strongly affect the localization performance of VISION aided INS. This paper presents a backward processing algorithm for asynchronous and time-delayed data fusion between the real-time system and the time-delayed system as follows: sensor data and filter data are temporarily stored in a buffer whenever the stereo camera acquires a query image. With the output of visual localization, measurement update of the Kalman filter is performed with the initial INS data in the buffer. After correcting the initial INS errors, all INS and filter data in the buffer are updated in order by using the inertial navigation algorithm and time update of the Kalman filter. The INS and filter data of the integrated system are then reset by the final result of the backward processing algorithm.

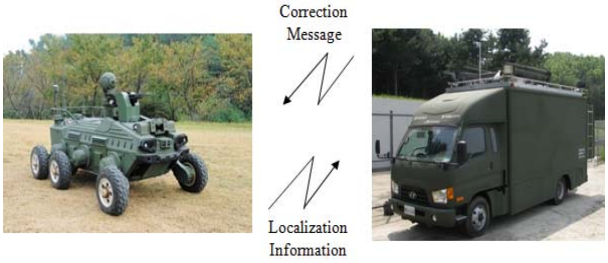
To obtain the submeter accuracy in open space where GPS is available, this paper uses a network of a ground-based reference station to correct GPS errors. The reference station is installed in the C2 (command and control) vehicle and the correction signal is broadcasted using WiBro (Wireless Broadband) communication. With this LDGPS (Local Differential GPS), all UGVs within WiBro communication range of the C2 vehicle can accurately estimate the position with a common GPS. By using DGPS to obtain the real-time reference position under the dynamic conditions of the reference station, the restrictions of conventional DGPS are also overcome.

The remainder of the paper is organized as follows. Section 2 presents the localization system architecture, Section 3 describes the design of localization algorithms, Section 4 describes the integrated localization system, and Section 5 and 6 represent the experimental results and conclusions.

2. LOCALIZATION SYSTEM ARCHITECTURE

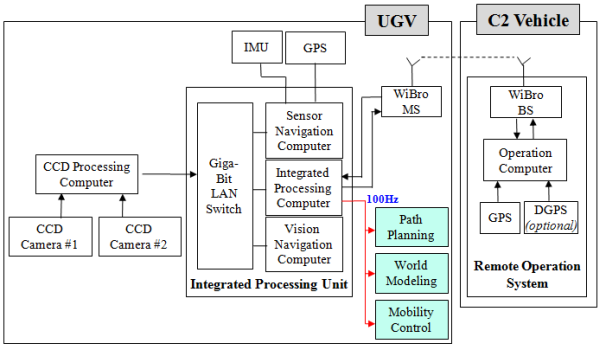
The localization system consists of two subsystems, as shown in Fig. 1(a). Subsystem 1 is the C2 vehicle, which acts as a reference station. In the C2 vehicle, the position and status of the UGV are displayed in the ROS (Remote Operation Station). With the real-time reference position, the GPS errors are estimated even under the dynamic conditions of the reference station and broadcasted using WiBro communication. Subsystem 2 is UGV rover. GPS errors are corrected by the reference station and the UGV position is obtained according to the operation modes.

Fig. 1(b) shows the hardware architecture of the UGV localization system. Subsystem 1 includes a GPS receiver (NovAtel, ProPak-V3-L1), DGPS receiver (Navcom SF-2050M), LDGPS operation computer (2.66GHz Core2Duo, 2GB DDR2), and WiBro BS (base station). Subsystem 2 includes a GPS receiver (NovAtel, OEMV3-L1), IMU (Honeywell, HG1700AG56), stereo camera (Point Grey, Flea2), autonomous navigation



<UGV: DHR> <C2 Vehicle>

(a) System configuration.



(b) Hardware architecture.

Fig. 1. UGV localization system.

computers (SBC: 1.66GHz Core Duo, 1GB DDR2), WiBro MS (mobile station), and Gigabit LAN. The vision navigation computer performs the visual localization and transfers the results to the sensor navigation computer. The sensor navigation computer acquires sensor data, runs the integrated localization algorithm, and transfers the results to the integrated processing computer. The integrated processing computer broadcasts the localization results to the autonomous navigation software components (world modeling, path planning, and motion control) in the UGV and the ROS in the C2 vehicle.

3. DESIGN OF LOCALIZATION ALGORITHMS

3.1. INS error model

There are two approaches to derive INS error models. One is the phi-angle approach or true-frame approach. The other is the psi-angle approach or computer-frame approach. In this paper, the psi-angle error model is implemented since it is generally easier to use for long-term error analysis and filter design for the integrated navigation systems [9]. The INS error model in NED (North-East-Down) coordinates can be written as follows:

$$\begin{aligned}
\delta \dot{L}^n &= -\omega_{\mathrm{en}}^{\mathrm{n}} \times \delta L^n + \delta V^n, \\
\delta \dot{V}^n &= \mathrm{C}_{\mathrm{b}}^{\mathrm{n}} \delta f^b + f^n \times \delta \psi^n - (2\omega_{\mathrm{ie}}^{\mathrm{n}} + \omega_{\mathrm{en}}^{\mathrm{n}}) \\
&\quad \times \delta V^n + \delta g^n, \\
\delta \dot{\psi}^n &= -(\omega_{\mathrm{in}}^{\mathrm{n}} \times \delta \psi^n) - \mathrm{C}_{\mathrm{b}}^{\mathrm{n}} \delta \omega_{ib}^b,
\end{aligned} \tag{1}$$

where δL^n, δV^n, and $\delta \psi^n$ are the position, velocity, and attitude errors vectors, respectively, f^n is the acceleration vector, δf^b is the accelerometer bias error vector, $\delta \omega_{ib}^b$ is the gyroscope bias error vector, C_b^n is the coordinate transform matrix from the body coordinates (b) to NED coordinates (n), ω_{ie}^n is the earth rate vector, ω_{en}^n is the craft rate vector, and ω_{in}^n is the sum of ω_{ie}^n and ω_{en}^n.

3.2. Filter design

Using (1), the system equation can be derived as follows.

$$\begin{aligned}
\dot{x}(t) &= F(t)x(t) + \omega(t), \\
x(t) &= [\delta L^n \ \delta V^n \ \delta \psi^n \ \delta f^b \ \delta \omega_{ib}^b]^T,
\end{aligned} \tag{2}$$

where $x(t)$ is the state vector, $F(t)$ is the system matrix [10], and $\omega(t)$ is the process noise vector.

Two measurement equations are designed since GPS and visual localization are operated selectively according to operation modes in this system. Generally, the measurement equation can be written as follows:

$$z_i(t) = H_i(t)x(t) + v_i(t), \tag{3}$$

where $z_i(t)$ is the i^{th} measurement vector, $H_i(t)$ is the i^{th} measurement matrix connecting measurement and states, and $v_i(t)$ is a measurement noise vector.

In the first measurement equation design, z_1 is the difference in position and velocity between GPS and INS in NED coordinates. Thus, the first measurement equation is written as follows:

$$\begin{aligned}
z_1(t) &= H_1(t)x(t) + v_1(t), \\
z_1(t) &= \begin{bmatrix} \delta L_N & \delta L_E & \delta L_D & \delta V_N & \delta V_E & \delta V_D \end{bmatrix}^T, \\
H_1(t) &= \begin{bmatrix} I_{6\times 6} & 0_{6\times 9} \end{bmatrix}.
\end{aligned} \tag{4}$$

In the second measurement equation design, z_2 is the difference in position between INS and visual localization in NED coordinates. Therefore, the second measurement equation is written as follows:

$$\begin{aligned}
z_2(t) &= H_2(t)x(t) + v_2(t), \\
z_2(t) &= \begin{bmatrix} \delta L_N & \delta L_E & \delta L_D \end{bmatrix}^T, \\
H_2(t) &= \begin{bmatrix} I_{3\times 3} & 0_{3\times 12} \end{bmatrix}.
\end{aligned} \tag{5}$$

In the Kalman filter design, it is assumed that ω and v_i are independent, zero-mean, white Gaussian sequences with covariance Q and R_i, respectively. The initial system state $x_i(t_0)$ is a Gaussian distributed random variable with an initial value $x_{i,0}$ and covariance $P_{i,0}$ and is independent of the noise. The Kalman filter algorithm is as follows [11]:

Time Update (Predictor, effect of dynamics):

$$\hat{x}_i(t_k^-) = \Phi_i(t_k, t_{k-1})\hat{x}_i(t_{k-1}^+), \tag{6}$$

$$P_i(t_k^-) = \Phi_i(t_k, t_{k-1})P_i(t_{k-1}^+)\Phi_i^T(t_k, t_{k-1}) + Q(t_{k-1}). \tag{7}$$

Measurement Update (estimator, effect of measurement):

$$K_i(t_k) = P_i(t_k^-)H_i^T(t_k)[H_i(t_k)P_i(t_k^-)H_i^T(t_k) + R_i(t_k)]^{-1}, \tag{8}$$

$$\hat{x}_i(t_k^+) = \hat{x}_i(t_k^-) + K_i(t_k)[z_i(t_k) - \mathrm{H}_i(t_k)\hat{x}_i(t_k^-)], \tag{9}$$

$$P_i(t_k^+) = [I - \mathrm{K}_i(\mathrm{t}_k)\mathrm{H}_i(\mathrm{t}_k)]P_i(t_k^-), \tag{10}$$

where $\Phi_i\left(= \exp\left(\int_{t_{k-1}}^{t_k} F(t)dt\right)\right)$ is the system transition matrix, K is the Kalman gain, the superscript '+' denotes an a posteriori estimate, and the superscript '–' denotes an a priori estimate.

The Kalman filter is known to provide optimal estimation performance. However, it is sensitive to the round-off error of the computer and the numerical accuracy. In order to avoid this problem, a square root filtering technique, Bierman's U-D factorization, is implemented.

For the safety of the localization system, the faults in the local filter are detected by including a Chi-square (χ^2) in the FDI algorithm, which monitors the statistics of the residual as follows.

$$\zeta(t_k) = \sum_{i=k-N+1}^{k} \upsilon_i^T S_i^{-1} \upsilon_i \sim \chi_{Np}^2, \tag{11}$$

where $\zeta(t_k)$ is the test statistic, which has a χ^2 distribution with N·p degrees of freedom; N and p denote the window size and dimension of the residual, respectively, $\upsilon(= z(t_k) - \mathrm{H}(t_k)\hat{x}(t_k^-))$ is a residual, and $S(= H(t_k)P(t_k^-)H^T(t_k) + R(t_k))$ is a residual covariance. In this paper, the window size N is chosen as 1 and a confidence level of 95% is chosen for the Chi-square. The confidence level is first chosen after the quality of each measurement is checked in the environment where GPS and visual localization are normal, respectively, and then adjusted through a trial and error method.

3.3. Visual localization design

The structure of visual localization is shown in Fig. 2. It is classified into off-line and on-line processes. In the off-line process, stereo camera and DGPS aided INS are used to collect images and the position data. After the feature extraction, a tree-based DB is created. The on-line process is for recognition, which includes searching through the tree and matching between images. When a query image is provided, a ranked list of matched images is computed. Relative pose estimation between the query image and DB images including the best matched image is then performed for localization.

In the feature extraction stage, SURF is used. SURF requires less computing resources compared with SIFT (Scale-Invariant Feature Transform), since it applies an approximate box filter to speed up its detecting performance.

In the search tree generation and the image retrieval stage, the vocabulary tree-based structure is used, because it is very fast and provides an efficient matching and searching process between a query image and DB images. It is built through unsupervised training by clustering the training set of feature descriptors using hierarchical *k*-means, where *k* is the branching factor of the tree. Fig. 3 shows the process of image retrieval and matching. Features from a query image are propagated down the tree until reaching the leaf ends and is scored by a single integer. The scoring process is shown in [6] as follows. Given a query image, define both the query vector q_i and the database vector d_i.

$$q_i = n_i \omega_i, \tag{12}$$

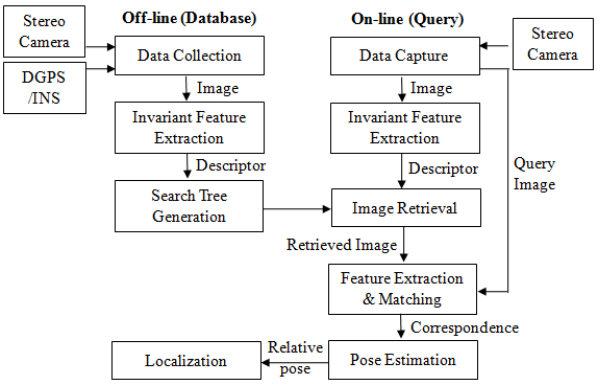


Fig. 2. Block diagram of visual localization algorithm.

$$d_i = m_i \omega_i, \tag{13}$$

where n_i and m_i are the number of descriptor vectors of query and database image, respectively, with a path through node *I*, and the weight ω_i is determined by (14).

$$\omega_i = In\frac{N}{N_i}, \tag{14}$$

where *N* is the number of images in the database, and N_i is the number of images in the database with at least one descriptor vector path through node *i*. Each database image is then given a relevance score *s* based on the normalized difference between the query and database vectors:

$$s(q,d) = \left\| \frac{q}{\|q\|} - \frac{d}{\|d\|} \right\|. \tag{15}$$

After the scores are calculated at each leaf, the image with the highest score is assumed to be the best match and the absolute position of the query image is coarsely estimated.

In the pose estimation stage, the coarse position is refined. Features between the query image and 3 subsequent DB images (search result, result-1, result +1) are matched using the Harris corner [12] and NCC (Normalized cross-correlation). After the feature matching, 3D points, which are the 3D position of the 2D points in the camera coordinate system, are generated using the stereo images, which are the results of the image matching process using the vocabulary tree. With the projected coordinates of these 3D points in the query image, the relative position between the query image and the DB images is determined using the perspective 3-point (P3P) algorithm [13]. Fig. 4 shows the process of

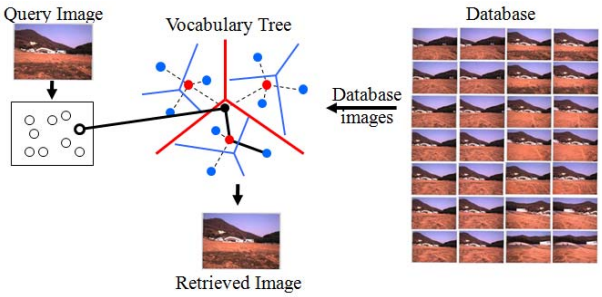


Fig. 3. The process of image retrieval and matching.

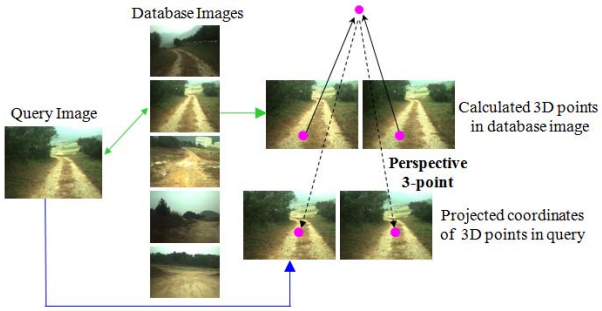


Fig. 4. The process of relative pose estimation.

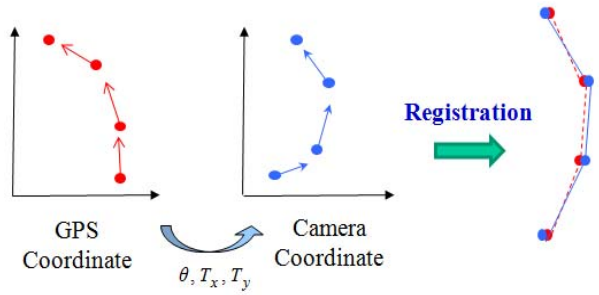


Fig. 5. Registration of two coordinates.

relative pose estimation.

When a stereo camera acquires the image in the off-line process, the absolute position corresponding with each image is also obtained. Therefore, registration between the absolute coordinate (GPS coordinate) and camera coordinate in the 2D plane is needed to transform the relative position of the query image into the absolute position. Since GPS and the camera have their own coordinate systems, the absolute position (longitude and latitude) is converted into the Cartesian coordinate to make its unit equal to that of the camera coordinate. Since the unknowns for registration between two coordinates are the rotation and translation, the cost function to be minimized is as follows.

$$\sum_{n=1}^{N}\left\| \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X_n \\ Y_n \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} - \begin{bmatrix} x_n \\ y_n \end{bmatrix} \right\|^2, \tag{16}$$

where θ is 2D rotation, $[T_x\ T_y]^T$ is 2D translation, N is image number, $[x_n\ y_n]^T$ is an optical center location (relative to 1st camera), and $[X_n\ Y_n]^T$ is the position of the n^{th} image with respect to the position of the first image in the Cartesian coordinate transformed from the absolute coordinate. With the registration between two coordinates as shown in Fig. 5, the relative position of the query image is transformed into the absolute coordinate.

3.4. LDGPS design

The position accuracy of GPS depends on that of the pseudo-range measured using the signal transmitted from satellites. However, the signals from satellites include abundant errors such as ionospheric and tropospheric delay. To reduce these errors, DGPS is generally used in a fixed reference station. In some situations such as communication interruption, however, the reference station is moved and consequently the conventional DGPS is impossible in real-time. To cope with this problem under the dynamic conditions of the reference station, a DGPS receiver is utilized to obtain the reference position in C2 vehicle. With the algorithm shown in Fig. 6, LDGPS is applicable in any environments where GPS is available even when the reference station moves.

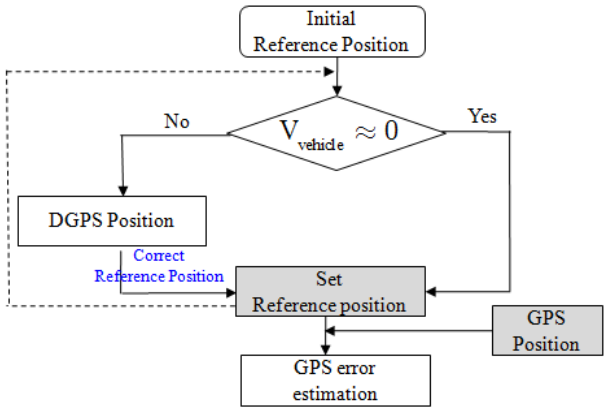


Fig. 6. LDGPS for the real-time reference position.

4. INTEGRATED LOCALIZATION SYSTEM

The integrated localization system uses the hierarchical federation of three measurement layers, i.e., GPS, INS, and visual localization. Therefore, the time synchronization between the localization systems should be considered very carefully in the filter design.

GPS and INS estimate the position in real-time at 1Hz and 100Hz, respectively, and the synchronization between GPS and INS is straightforward due to the fact that the GPS receiver supplies a 1PPS signal pulse aligned with each GPS second. As shown in Fig. 7, however, visual localization is time-delayed because localization is done at $t = k$ after performing image processing via image matching and pose estimation when the stereo camera acquires the image at $t = i$. In the VISION aided INS, visual localization estimates the position at $t = i$ while INS estimates the position at $t = k$. Thus, the distance the UGV moves during the delay time ($\Delta t = k - i$) can affect the localization accuracy of VISION aided INS. The reason for this is that the position error between two systems may be increased substantially since it depends on the velocity of the UGV as well as the image processing time. Also, the output of visual localization is asynchronized with INS and GPS, because the image capture is not derived from the GPS receiver timing circuit. The CCD processing computer captures images at 30fps (frame per second) for the world modeling of stereo vision and transfers images into the Vision navigation computer by a Giga-Bit LAN switch, as shown in Fig. 1(b).

Therefore, this paper presents a novel asynchronous and time-delayed data fusion scheme to increase the localization accuracy of the integrated system by

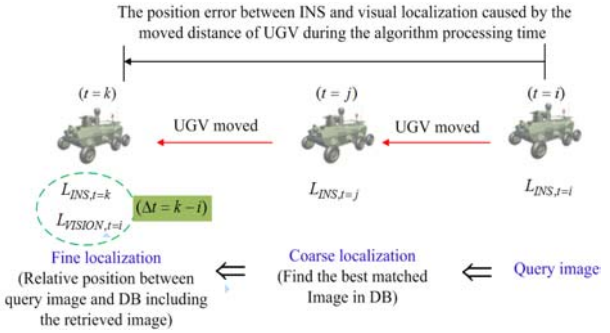


Fig. 7. Position error between two systems.

removing the position error of visual localization in the filtering process. This is called the backward processing algorithm.

1) **Data acquisition:** Data are temporarily stored in buffer as a vector d_j whenever the stereo camera acquires an image.

$$d_j = (t_j \quad IMU_j \quad INS_j \quad GPS_j \quad VISION_j \quad KF_j), \quad (17)$$

where t is time, j is the INS sampling time (10ms), *IMU* is data from IMU (three-axis accelerations and angular velocities), *INS* is data from INS (position, velocity, attitude), *GPS* is data from GPS (HDOP), *VISION* is data from the visual localization process (position), and *KF* is a data from the Kalman filter (state, covariance).

2) **Time synchronization and data fusion:**

Step 1: With the output of visual localization, measurement update is performed using the initial data of *INS* and *KF* in the buffer.

Step 2: The faults of the measurement system are detected and isolated by the FDI algorithm.

Step 3: The errors of the initial INS data in the buffer are corrected by the states of Step 2.

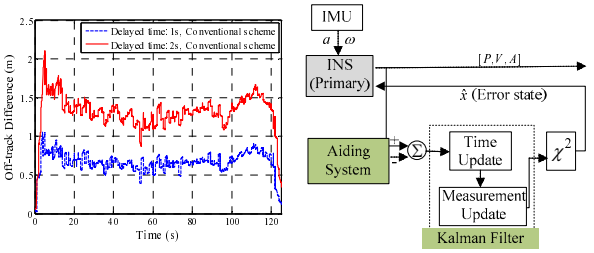
Step 4: INS data in the buffer are updated by using the inertial navigation algorithm in order from the corrected initial data to the final data.

Step 5: Time update is performed with the updated INS data.

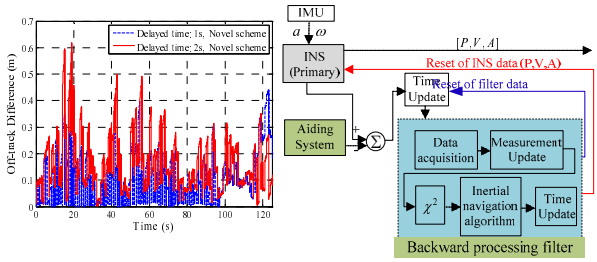
Step 6: Delete all data except for the final data and finish the backward processing algorithm.

3) **Data reset:** The INS and filter data in the integrated localization system are reset by the results of 2) **Time synchronization and data fusion.**

To verify the effects of the backward processing algorithm, a simulation is performed with the experimental results of a GPS/INS integrated system. GPS data is delayed a few seconds to imitate the time delay of visual localization. Fig. 8 shows the off-track (horizontal) differences with respect to the original experimental results. In the conventional Kalman filter, the off-track difference is increased as the delay time of the measurement system is increased. In the backward processing algorithm, however, the off-track difference is not increased significantly, although the delay time is increased. These results show that the proposed fusion scheme is very useful in cases where the output of the aiding systems is time-delayed with respect to the reference system. Since the computation time of the backward processing algorithm is less than 10ms (INS sampling time) in our navigation computers, this algorithm corrects INS errors in real-time.



(a) Conventional filter.



(b) Proposed algorithm.

Fig. 8. Effect of the novel fusion scheme.

In the application of the integrated localization system, it is more efficient to use LDGPS and visual localization selectively as a measurement system according to the operation environment. In open space, LDGPS is very accurate but visual localization is not accurate due to a lack of features and lighting conditions. In the vicinity of trees and building, however, visual localization is accurate but LDGPS is not accurate, because the GPS signal may be blocked. Therefore, this paper designs a navigation manager as shown in Fig. 9. As the UGV moves in open space, LDGPS aided INS is only used for localization. If the uncertainty from GPS becomes large, VISION aided INS is used. In this paper, RTCM (Radio Technical Commission for Maritime Services) messages from the C2 vehicle and HDOP (Horizontal Dilution of Precision) are used to check the uncertainty. The threshold (ε) of HDOP is 3 in this paper. The operation principle is as follows.

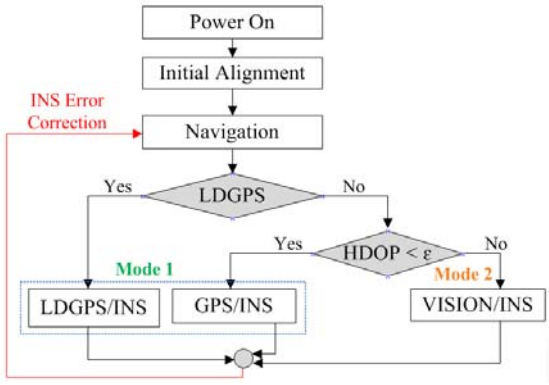


Fig. 9. Navigation manager: operation mode.

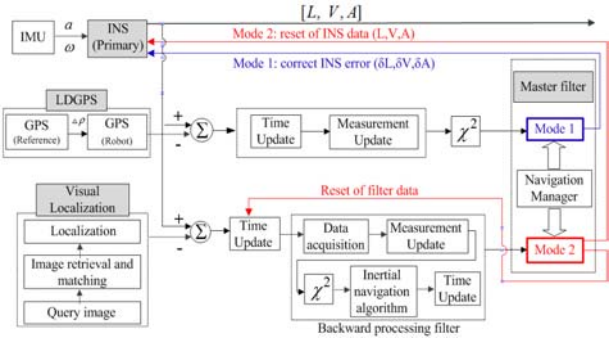


Fig. 10. Integrated localization system architecture.

If message 17 is normal, LDGPS aided INS
Else
  If HDOP < ε, GPS aided INS
  Else VISION aided INS.

Fig. 10 shows the architecture of the proposed integrated localization system. The detailed designs of each module are provided in Sections 3 and 4.

5. EXPERIMENTAL RESULTS

To verify the performance of the proposed integrated localization system, a series of experiments were performed in outdoor environments. First, the mode 1 experiment is carried out to evaluate the localization accuracy of the LDGPS aided INS. Second, the mode 2 experiment is conducted to evaluate the localization accuracy of the VISION aided INS. In the mode 2 experiment, the transition between mode 1 and mode 2 is also verified.

5.1. Localization based on LDGPS aided INS

Since the purpose of this experiment is to evaluate the localization accuracy of the LDGPS aided INS, the experiment is conducted in open space where GPS is available.

Fig. 11 shows the trajectories obtained by GPS, LDGPS, LDGPS/INS, and the reference, respectively. The reference trajectory is generated using CDGPS (Carrier-phase Differential GPS). It shows that the trajectories of LDGPS and LDGPS aided INS besides GPS are very similar to those of reference.

Fig. 12 shows the off-track differences of GPS, LDGPS, and LDGPS aided INS with respect to the reference. In this experiment, HDOP is almost 2.2 and the number of observable satellites is 8. While the off-track difference of GPS is a RMS (Root Mean Square) of 2.313m with a maximum of 2.662m, that of LDGPS is a RMS of 0.383m with a maximum of 0.890m, and that of LDGPS aided INS is a RMS of 0.317m with a maximum of 0.512m. These results show that the GPS errors are

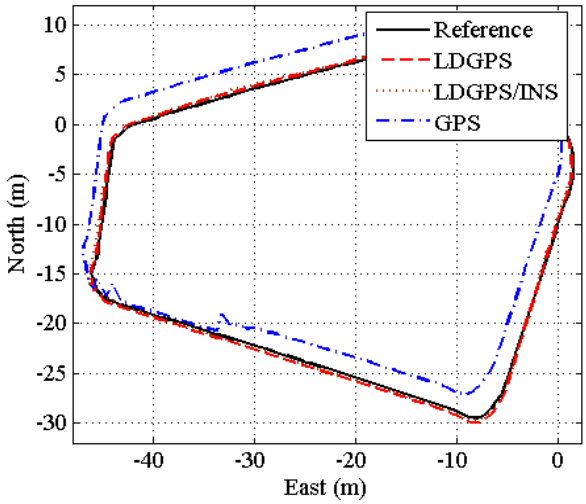


Fig. 11. Trajectory comparison: Mode 1.

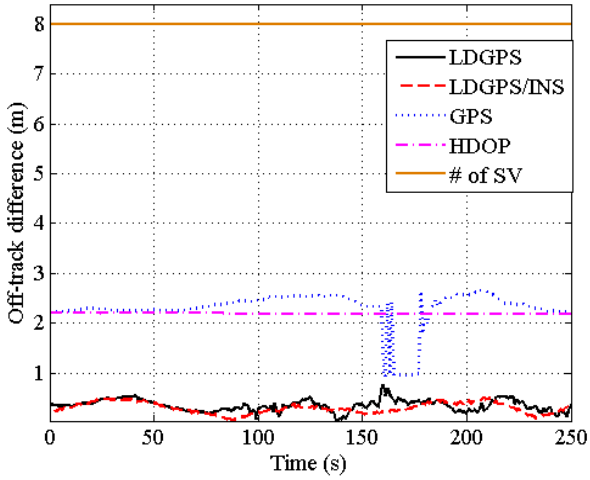


Fig. 12. Off-track differences: Mode 1.

significantly reduced by a network of the ground-based reference station and submeter accuracy can be obtained using LDGPS.

Consequently, it is demonstrated that LDGPS aided INS is sufficient for localization in open space and no additional sensors are needed.

5.2. Localization based on VISION aided INS and the transition between mode 1 and mode 2

In this experiment, the accuracy of visual localization and VISION aided INS is evaluated with the transition between mode 1 and mode 2. The experiment procedure is as follows: First, images with position data are collected while the UGV moves across the reference trajectory and are transformed into a vocabulary tree-based structure. The localization performance is then evaluated while the UGV moves across the trajectory.

Fig. 13 shows two paths of this experiment on the image captured from Google Earth, which provides a virtual globe, map, and geographic information. In the real environment, path 1 has numerous features but path 2 has a few features and open space in some trajectories.

Fig. 14 shows the time delay of visual localization with respect to INS in this experiment. With the UGV computer, the average delay time is 1.15s and the maximum delay time is 1.33s. This means that visual localization estimates the position, on average, 1.15s later relative to INS.

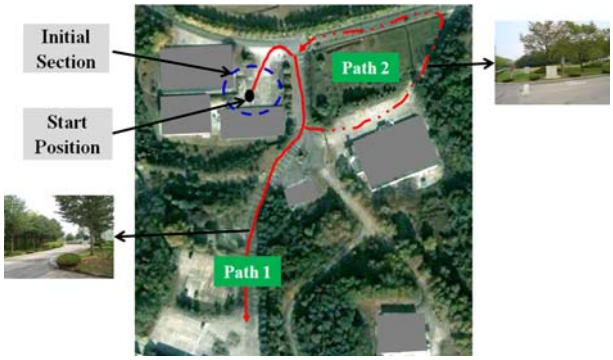


Fig. 13. The experimental environment.

When the UGV is near the start position of Fig. 13, visual localization itself is not good, because this area is almost open space. Thus, the mode transition is also tested as follows: when the UGV passes the initial section, the GPS signal to the UGV is blocked.

Fig. 15 shows the trajectories of the UGV in path 1. Since this path has numerous features as trees and buildings on each side, visual localization estimates the position accurately except for the initial section. The off-track difference is shown in Fig. 16. In the integrated system, the off-track difference is very small in the initial section because mode 1 is used. After passing the initial section, it is slightly increased with the mode transition from mode 1 to mode 2. In the visual localization, however, it is very large in the initial section, because this section is in open space. After passing the initial section, it is substantially reduced because there are many features along the trajectory and the effect of lighting condition is not significant. In the visual localization, it has a RMS of 3.006m with a maximum of 11.437m. In the integrated localization, it has a RMS of 1.795m with a maximum of 2.945m. Note that the off-track difference of visual localization includes both the image matching error and the position error of the UGV during the time delay of Fig. 14.

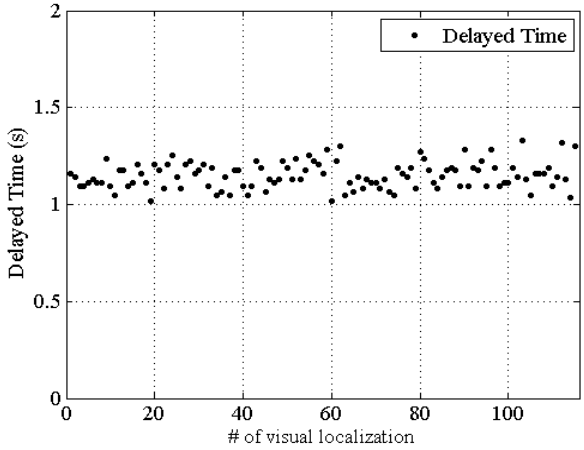


Fig. 14. Time delay of visual localization.

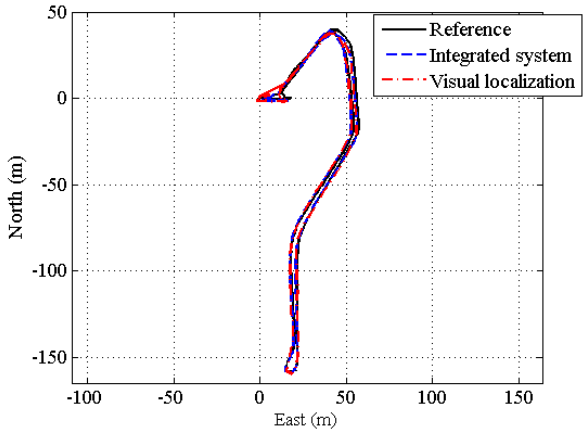


Fig. 15. Trajectory comparison: Path 1.

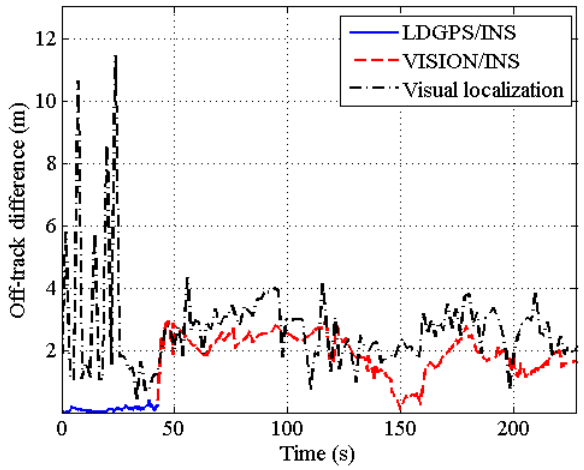


Fig. 16. Off-track differences: Path 1.

Fig. 17 shows the trajectories of the UGV in path 2. As shown in Fig. 13, this path has fewer features than path 1 and has open space in some trajectories. This figure shows that visual localization is affected by the sunlight in some sections. In addition, the off-track difference increases very much, as shown in Fig. 18. However, this large error is rejected in the filtering process, because the integrated localization algorithm uses the FDI algorithm for system safety. The off-track

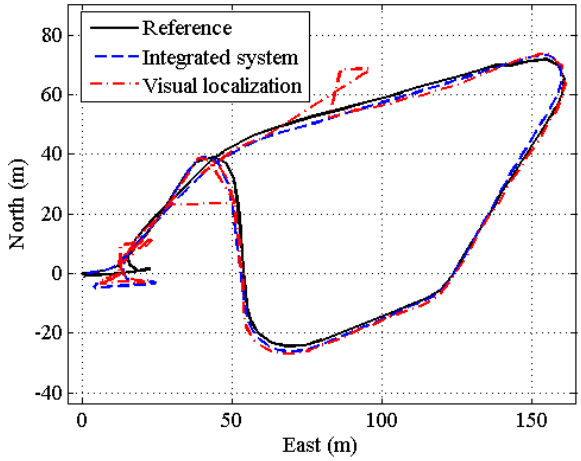


Fig. 17. Trajectory comparison: Path 2.

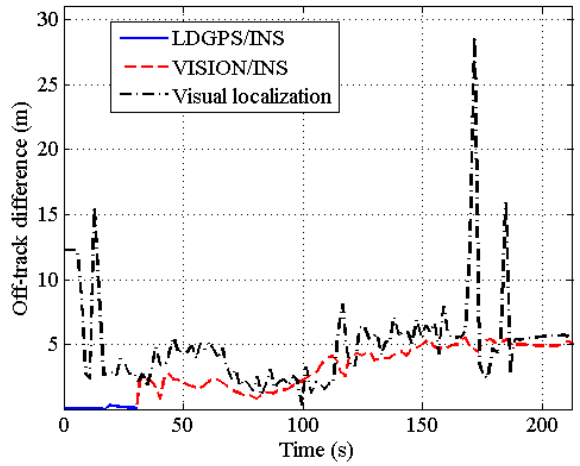


Fig. 18. Off-track differences: Path 2.

difference of visual localization is a RMS of 6.113m with a maximum of 28.757m. The off-track difference of VISION aided INS is a RMS of 3.500m with a maximum of 5.472m. If GPS is not blocked in this experiment, however, the off-track difference will be substantially reduced, because the trajectory with large errors in visual localization is in open space, and as a result, Mode 1 will be operated by the navigation manager.

Consequently, these experiments show that visual localization can be used effectively in environments where GPS is unavailable. Furthermore, the integrated localization system with the operation modes yields robust and accurate localization in outdoor environments.

6. CONCLUSIONS

This paper proposes an improved localization algorithm based on the hierarchical federation of three measurement layers, i.e., GPS, INS, and visual localization to overcome the shortcomings of INS/GPS integrated systems in outdoor environments. This algorithm automatically switches the operation modes (GPS aided INS, VISION aided INS) with a different operating principle according to GPS status and a network of a ground-based reference station. Since a fast image retrieval process between DB images and query image is possible with a vocabulary tree-based data structure, visual localization can be applicable in large operation environments. In the VISION aided INS, an asynchronous and time-delayed data fusion algorithm between the real-time system and the time-delayed system is presented to improve the accuracy of localization, as the output of visual localization is always time-delayed compared with that of INS. By using DGPS to obtain the reference position even under the dynamic conditions of the reference station, the restrictions of conventional DGPS are overcome and all UGVs within WiBro communication range of the reference station can accurately estimate the position with a common GPS. The experiment results show that the integrated localization algorithm can substantially enhance the robustness and accuracy of localization in outdoor environments.

Consequently, the proposed integrated localization algorithm is well suited for UGVs for perimeter surveillance operation in outdoor environments.

7. FUTURE WORK

This algorithm needs to be further developed to make localization more robust in highly dynamic environments. To this end, the following approaches will be considered: perception-based high level features in addition to low level features, illumination robust feature matching, and dynamic scene management for the classification of static objects in outdoor environments. The application of DEM (Digital Elevation Model) and DSM (Digital Surface Model) will be also considered. This is of fundamental importance since the above approaches will make the UGV localization more robust and accurate even in highly dynamic environments.

REFERENCES

[1] P. Y. Cui and T. L. Xu, "Data fusion algorithm for INS/GPS/Odometer navigation system," *Proc. of IEEE Conf. Industrial Electronics and Applications*, pp. 1893-1897, May 2007.

[2] Y. C. Yang and J. A. Farrell, "Magnetometer and differential carrier phase GPS-aided INS for advanced vehicle control," *IEEE Trans. on Robotics and Automation*, vol. 19, no. 2, pp. 269-282, April 2003.

[3] W. Li and H. Leung, "Constrained unscented Kalman filter based fusion of GPS/INS/Digital map for vehicle localization," *Proc. of IEEE Conf. Intelligent Transportation Systems*, vol. 2, pp. 1362-1367, Oct. 2003.

[4] G. Dissanayake, S. Sukkarieh, E. M. Nebot, and H. Durrant-Whyte, "The aiding of a low-cost strapdown inertial measurement unit using vehicle model constraints for land vehicle applications," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 5, pp. 731-747, October 2001.

[5] B. B. Liu, M. Adams, and I. Guzman, "Multi-aided inertial navigation for ground vehicles in outdoor uneven environments," *Proc. of IEEE International Conf. Robotics and Automation*, pp. 4703-4708, April 2005.

[6] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 2161-2168, 2006.

[7] H. Bay, A. Ess, T. Tuytelaars, and L. Vangool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, June 2008.

[8] T. S. Abuhashim, M. F. Abdel-Hafez, and M.-A. Al-Jarrah, "Building a robust integrity monitoring for a low cost GPS-aided INS system," *International Journal of Control, Automation, and Systems*, vol. 8, no. 5, pp. 1108-1122, 2010.

[9] D. O. Benson, "A comparison of two approaches to pure-inertial and Doppler-inertial error analysis," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-11, no. 4, pp. 447-455, July 1975.

[10] M. F. Abdel-Hafez, D. J. Kim, E. S. Lee, S. B. Chun, Y. J. Lee, T. S. Kang, and S. K. Sung, "Performance improvement of the wald test for GPS RTK with the assistance of INS," *International Journal of Control, Automation, and Systems*, vol. 6, no. 4, pp. 534-543, August 2008.

[11] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, vol. 1, Academic Press, New York, 1979.

[12] C. Harris and M. Stephens, "A combined corner and edge detector," *Proc. of the 4th Alvey Vision Conference*, pp. 147-151, Sep. 1988.

[13] R. M. Haralick, D. Lee, K. Ottenburg, and M. Nolle, "Analysis and solutions of the three point perspective pose estimation problem," *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 592-598, Jun. 1991.

Ji-Hoon Choi received his M.S. degree in Mechanical Engineering from Korea University in 2002 and is now a Ph.D. candidate in Mechanical Engineering at Korea University. He is also a Senior Researcher in UGV technology directorate at the Agency for Defense Development, Korea. His research interests lie mainly in localization based on multi-sensor fusion, obstacle avoidance, vision-based localization, and SLAM.

Yong-Woon Park received his M.S. degree in Mechanical Engineering from Yonsei University in 1987, and his Ph.D. degree in Mechanical Engineering from University of Utah in 1994. He is a Leader in UGV technology directorate at the Agency for Defense Development, Korea. His current research interests lie mainly in autonomous system, path planning, and vision-based navigation.

Jae-Bok Song received his B.S. and M.S. degrees in Mechanical Engineering from Seoul National University in 1983 and 1985, respectively. Also, he received his Ph.D. degree in Mechanical Engineering from MIT in 1992. He is currently a Professor of Mechanical Engineering, Korea University, where he is also the Director of the intelligent Robotics Laboratory from 1993. His current research interests lie mainly in mobile robotics, design and control of intelligent robotic systems, and safe manipulators.

In-So Kweon received his B.S. and M.S. degrees in Mechanical Design and Production Engineering from Seoul National University in 1981 and 1983, respectively, and his M.S. and Ph.D. degrees in Robotics from the Robotics Institute at Carnegie Mellon University, Pittsburgh, U.S.A, in 1986 and 1990, respectively. He is currently a Professor in the Department of Electrical Engineering, KAIST. His current research interests are in computer vision, robotics, pattern recognition, and automation.