Map-building and Localization by Three-dimensional Local Features for Ubiquitous Service Robot

Youngbin Park[1], Seungdo Jeong[2], Il Hong Suh[1], Byung-Uk Choi[1]

[1] College of Information and Communications, Hanyang University
17 Haengdang-dong, Sungdong-gu, Seoul, 133-791, Korea
ybpark@mlab.hanyang.ac.kr, {ihsuh, buchoi}@hanyang.ac.kr
[2] Department of Electrical and Computer Engineering, Hanyang University
17 Haengdang-dong, Sungdong-gu, Seoul, 133-791, Korea
sdjeong@hanyang.ac.kr

Abstract. In this work, we propose a semantic-map building method and localization method for ubiquitous service robot. Our semantic-map is organized by using SIFT feature-based object representation. In addition to semantic map, a vision-based relative localization is employed as a process model of extended Kalman filters, where optical flows and Levenberg-Marquardt least square minimization are incorporated to predict relative robot locations. Thus, robust map-building performances can be obtained even under poor conditions in which localization cannot be achieved by classical odometry-based map-building. To localize robot position and solve kidnap problem, we also propose simple, but fast localization method with a relatively high accuracy by incorporating our semantic-map.

1 Introduction

To navigate and plan a path in an environment, the intelligent ubiquitous service robot has to be able to build a map of the environment and to recognize its location autonomously. SLAM(Simultaneous Localization And Mapping) is a popular technique to accurately localize the robot and simultaneously build a map of the environment. Numerous studies of SLAM have been performed, as this has been one of the important issues in the intelligent mobile robot community for a long time. Since the 1990s, methods using extended Kalman filters have been focused on by a number of researchers interested in SLAM [3, 8]. Most algorithms using extended Kalman filters combine two methods. One is relative localization, which is the method used to compute the current position with respect to an initial location. The other is the method using a map composed of landmarks.

Odometry is often used for relative localization. However, it has many errors caused not only by systematic factors such as a difference in wheel diameter, inaccurate gauging of wheel size, and others, but also non-systematic factors such as slip, poor condition of the floor, etc. [4] proposed an algorithm to reduce such systematic errors. Their algorithm consists of three steps: odometry error modeling, error parameter estimation using the PC-method, and estimation of the covariance matrix. Even so, it is hard to overcome the error of estimation caused by non-systematic factors.

Building the map of the environment is an issue with SLAM. The map consists of landmarks which are used for ubiquitous service robot localization. Davision uses three-dimensional corner points as landmarks, which are obtained by a Harris corner detector and stereo matching [5]. Se uses keypoints obtained from SIFT(Scale Invariant Feature Transform) [6, 10]. However, simple coordinates or features within the image are not enough to facilitate interaction with humans.

Robot localization has been another issue in robot navigation system. A variety of methods have been proposed to locate robot position. One of those methods are to update robot position by utilizing odometer history. In indoor environments, those methods probably yield good results. But, when the robot moves without the odometer history update as in the kidnapped robot problems, robot cannot locate itself. And, robot position should be determined by other sensor information in the map. In solving the kidnapped robot problems, due to uncertainties and incorrectness of map and sensor information, estimation of robot position yields errors. Determining accurate robot position by estimated values from these incorrect measurements is also challenging task [9].

In this work, we propose the semantic-map building method for ubiquitous service robot. For building the semantic-map, we use vision-based relative localization process as the process model of our extended Kalman filter. This approach is robust enough for environments which cannot be supported by encoders that measure values such as the number of rotations of the robot's wheels. And we also propose a technique to locate robot position in a realtime while ensuring accuracy, where our semantic-map and stereo vision are employed.

2 OFM-based Robot Localization and Semantic-Map Building

In this work, we propose a three-dimensional object feature model (OFM) with essential properties and propose a method to use the OFM for improving the performance of vision-based SLAM. Figure 1 shows the block diagram of our SLAM method using a three-dimensional OFM.

We use images taken by the stereo camera as the only sensor information in this work. Our system is composed of three parts: the real-time relative localization part which estimates robot location in real-time, the landmark recognition part which builds up and/or observes landmarks by using SIFT-based object recognition, and the data fusion part which combines two observation data using an extended Kalman filter.

2.1 Real-time Relative Localization

Image-based relative localization is a method that applies the motion between corresponding points within consecutive image sequences to localize the robot. Thus, matching between feature points is a crucial factor for localization performance.

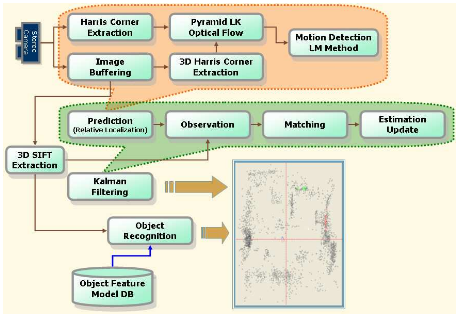


Fig. 1. Block diagram of the proposed map-building.

Tracking of corner features To track the camera motion, it is necessary to obtain the feature points in the current frame corresponding to the feature points in the previous frame. Note that the corner points satisfy the local smoothness constraint. Lucas-Kanade optical flow performs well in tracking corner points with that property [7]. However, Lucas-Kanade optical flow is best suited to small motion tracking. It is not sufficient to follow the movement of the ubiquitous service robot. Thus, pyramid Lucas-Kanade (PLK) optical flow using a Gaussian image pyramid is used for relative localization, where it is known to track a relatively wide area.

Relative localization The feature points extracted from the initial frame are used as 3D point landmarks. Note that we use the stereo camera as the input sensor, and we obtain a stereo image pair. The 3D coordinates of the extracted feature point are calculated using the disparity of the stereo camera images. The stereo camera is previously calibrated. When we know the exact relationship between the 3D coordinates of the landmark and the corresponding 2D coordinates which are projected to the image, the projection matrix indicates the mapping relationship between points of 3D space and points of the image.

The projection matrix is composed of two parts. One is the intrinsic parameter matrix which includes internal information of the camera. This matrix represents the relationship between the camera coordinate system and the image coordinate system. The other part of the projection matrix is the extrinsic parameter matrix. This matrix represents the translational and rotational relationship between the 3D space coordinate system and the camera coordinate system.

Therefore, if the projection matrix is estimated with coordinates of corresponding points in 3D space and in the 2D image, we can obtain the motion of the camera and the relative location of the robot. It is a non-linear problem to estimate the projection matrix with numerous corresponding points. In this work, we estimate the rotation and translation components using the Levenberg-Marquardt least square minimization (LM LSM) method.

2.2 Extended Kalman Filter-based SLAM using Object Landmarks

In most previous research works on SLAM, landmarks were composed of lines of the environment. They were obtained using a range finder or feature points of camera images. These maps only involve coordinates of feature points. Thus, the use of these maps for high level purposes, such as delivery tasks or interactions with humans, requires an anchoring technique to link point-based landmarks with their associated semantics. Alternatively, in this work, objects are recognized using SIFT and then features of recognized objects are registered as landmarks. Therefore, we can create the semantic-map supporting the location of main objects in the environment without any additional anchoring process.

In most SLAM methods using Kalman filters, odometry-based robot kinematics or dynamics are used as the process model. However, those performances are very poor because of systematic factors such as the difference of wheel diameter and non-systematic factors such as slip. Moreover, in the case that a wheel encoder does not exist, such as for a humanoid robot, odometry-based methods are difficult to apply.

Therefore, to resolve such drawbacks, we propose the extended Kalman filter-based SLAM that integrates PLK optical flow and LM algorithm-based relative localization with object landmarks.

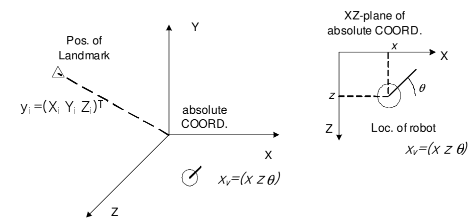


Fig. 2. System coordinate for the robot location and the landmark position.

The state vector and covariance We assume that the robot is located on a plane as in Fig. 2, so position and orientation of the robot is represented by $x_v = (\,x\ z\ \theta\,)^T$. The position of each landmark is denoted as $y_i = (\,X_i\ Y_i\ Z_i\,)^T$. Here, SIFT keypoints of objects recognized by the three dimensional OFM are represented as absolute coordinates. To regulate uncertainty of landmarks and relationships among them we use the system state vector and covariance model proposed by [1]. Thus, we can represent the state vector p and covariance matrix Σ as

$$p = \begin{pmatrix} x_v \\ y_1 \\ y_2 \\ \vdots \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma^2_{x_v x_v} & \sigma^2_{x_v y_1} & \sigma^2_{x_v y_2} & \cdots \\ \sigma^2_{y_1 x_v} & \sigma^2_{y_1 y_1} & \sigma^2_{y_1 y_2} & \cdots \\ \sigma^2_{y_2 x_v} & \sigma^2_{y_2 y_1} & \sigma^2_{y_2 y_2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{1}$$

Three dimensional object feature model The three dimensional object feature model(3D OFM) is defined as the model which is composed of the three dimensional SIFT keypoints extracted from object images. To make the three dimensional object feature model, we rotate and object by 20 degrees with respect to the center of gravity and then take the images using the calibrated stereo camera. This process is repeated to get 18 views images. SIFT keypoints are extracted from each image and are given three dimensional coordinates calculated with the stereo vision technique. SIFT keypoints having three dimensional coordinate are called as the three dimensional SIFT keypoints. The three dimensional SIFT keypoints in objects recognized by using 3D OFM are used as landmarks.

Process model In EKF, we define the process model as the estimation of current location and its covariance for the robot and landmarks referring to the state vector and its covariance matrix for the previous system. We apply the displacement $\Delta x, \Delta z, \Delta\theta$ obtained by PLK optical flow and LM algorithm-based relative localization to the process model. Let the process model of our work be given as

$$x_v(k+1|k) = f(x_v(k|k), u(k)) = \begin{bmatrix} x \\ z \\ \theta \end{bmatrix} + \begin{bmatrix} \Delta x_r \\ \Delta z_r \\ \Delta\theta \end{bmatrix}. \tag{2}$$

In (2), Δx_r and Δz_r are given as

$$\begin{bmatrix} \Delta x_r \\ \Delta z_r \end{bmatrix} = \begin{bmatrix} cos(\theta+\Delta\theta) & sin(\theta+\Delta\theta) \\ -sin(\theta+\Delta\theta) & cos(\theta+\Delta\theta) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix}. \tag{3}$$

The current location for the robot $x_v(k+1|k)$ is estimated by function f as shown in (2). $x_v(k|k)$ and $u(k)$ represent displacement of the previous location and motion of the robot respectively. Δx_r and Δz_r are the displacement which are transformed to the reference coordinate system of the map from the estimated displacement Δx and Δz referring to the camera coordinate system of the previous location.

$$y_i(k+1|k) = y_i(k|k), \forall i. \tag{4}$$

In (4), $y_i(k+1|k)$ represents the estimated current position of the i-th landmark, and we assume that landmarks are fixed.

The covariance for the system $\sigma^2_{x_v x_v}, \sigma^2_{x_v y_i}$ and $\sigma^2_{y_i y_j}$ are obtained as

$$\begin{aligned} &\sigma^2_{x_v x_v}(k+1|k) = \nabla_{x_v} f \sigma^2_{x_v x_v}(k|k) \nabla_{x_v} f^T + \nabla_u f \sigma^2_u(k|k) \nabla_u f^T, \\ &\sigma^2_{x_v y_i}(k+1|k) = \nabla_{x_v} f \sigma^2_{x_v y_i}(k|k), \\ &\sigma^2_{y_i y_j}(k+1|k) = \sigma^2_{y_i y_j}(k|k), \end{aligned} \tag{5}$$

where $\nabla_{x_v} f$ and $\nabla_u f$ represent the Jacobian of the estimation function f for the state vector of the robot and the displacement respectively. These are given as

$$
\begin{aligned}
&\nabla_{x_v} f = \left[\, \tfrac{\partial f}{\partial x} \; \tfrac{\partial f}{\partial z} \; \tfrac{\partial f}{\partial \theta} \,\right], \quad \nabla_u f = \left[\, \tfrac{\partial f}{\partial \Delta x} \; \tfrac{\partial f}{\partial \Delta z} \; \tfrac{\partial f}{\partial \Delta\theta} \,\right], \\
&where \quad \partial f/\partial x = \begin{bmatrix} 1 \; 0 \; 0 \end{bmatrix}^T, \quad \partial f/\partial z = \begin{bmatrix} 0 \; 1 \; 0 \end{bmatrix}^T, \\
&\tfrac{\partial f}{\partial \theta} = \begin{bmatrix} -sin(\theta + \Delta\theta)\Delta x + cos(\theta + \Delta\theta)\Delta z \\ -cos(\theta + \Delta\theta)\Delta x - sin(\theta + \Delta\theta)\Delta z \\ 1 \end{bmatrix}, \\
&\partial f/\partial \Delta x = \begin{bmatrix} cos(\theta + \Delta\theta) \; -sin(\theta + \Delta\theta) \; 0 \end{bmatrix}^T, \\
&\partial f/\partial \Delta z = \begin{bmatrix} sin(\theta + \Delta\theta) \; cos(\theta + \Delta\theta) \; 0 \end{bmatrix}^T, \\
&\tfrac{\partial f}{\partial \Delta\theta} = \begin{bmatrix} -sin(\theta + \Delta\theta)\Delta x + cos(\theta + \Delta\theta)\Delta z \\ -cos(\theta + \Delta\theta)\Delta x - sin(\theta + \Delta\theta)\Delta z \\ 1 \end{bmatrix}.
\end{aligned}
\tag{6}
$$

Here, ($\Delta x \; \Delta z \; \Delta\theta$) is the movement of the robot estimated by the LM algorithm.

In (5), σ_u^2 is the covariance matrix due to the noise in the process of camera motion estimation.

Measurement model Let the measurement model be given as $h_i(k+1) = h(y_i, x(k+1|k))$. Observation for landmark is the three dimensional coordinate of landmarks based on the calibrated stereo camera coordinate system. Measurement prediction model of h_i is given as

$$
h_i(k+1) = [R][y_i - x'_v], \quad where \quad x'_v = \begin{bmatrix} x \\ 0 \\ z \end{bmatrix}, R = \begin{bmatrix} cos\theta & 0 & -sin\theta \\ 0 & 1 & 0 \\ sin\theta & 0 & cos\theta \end{bmatrix}. \tag{7}
$$

Three dimensional absolute coordinate of landmark is denoted as y_i. And h_i is the function to transform absolute coordinate to camera coordinate system. x'_v represents planar coordinate of the robot based on the absolute coordinate system. Rotation between the absolute coordinate system to the camera coordinate system is denoted by R. θ in matrix R denotes the orientation of the robot.

Observation and matching Matching between landmarks and observed three dimensional coordinate of SIFT $z_j = [X_j \; Y_j \; Z_j]$ is accomplished through SIFT matching algorithm. To determine searching area for matching, we can use the innovation matrix and its covariance.

The innovation matrix $v_{ij}(k+1)$ between the predicted measurement $h_i(k+1|k)$ for landmark y_i and observation z_j is given as

$$
\begin{aligned}
v_{ij}(k+1) &= [z_j(k+1) - h(y_i, p(k+1|k))], \\
\sigma^2_{IN,ij} &= \nabla_{x_v} h_i \cdot \sigma^2_{x_v x_v} \cdot \nabla_{x_v} h_i^T + \nabla_{x_v} h_i \cdot \sigma^2_{x_v y_i} \cdot \nabla_{y_i} h_i^T \\
&\quad + \nabla_{y_i} h_i \cdot \sigma^2_{y_i x_v} \cdot \nabla_{x_v} h_i^T + \nabla_{y_i} h_i \cdot \sigma^2_{y_i y_i} \cdot \nabla_{y_i} h_i^T + \sigma^2_{R,i},
\end{aligned}
\tag{8}
$$

where $\sigma^2_{R,i}$ represents the covariance of the measurement.

Searching area is determined with the Mahalanobis distance and threshold constant g^2 such as

$$
v_{ij}^T(k+1) \cdot \sigma^{-2}_{IN,ij} \cdot v_{ij}(k+1) \le g^2 \tag{9}
$$

Using SIFT matching algorithm, matching is accomplished between the landmark y_i and the observation satisfying the criterion shown in (9).

Estimation of the system state vector and covariance Kalman gain K can be calculated as

$$K = \sigma^2 \nabla_{x_v} h_i^T \cdot \sigma_{IN_i}^{-2} = \begin{pmatrix} \sigma^2_{x_v x_v} \\ \sigma^2_{y_i x_v} \\ \vdots \end{pmatrix} \frac{\partial h_i^T}{\partial x_v} \sigma_{IN_i}^{-2} + \begin{pmatrix} \sigma^2_{x_v y_i} \\ \sigma^2_{y_i y_i} \\ \vdots \end{pmatrix} \frac{\partial h_i^T}{\partial y_i} \sigma_{IN_i}^{-2}, \tag{10}$$

where $\sigma^2_{x_v x_v}$, $\sigma^2_{x_v y_i}$ and $\sigma^2_{y_i y_i}$ are 3×3 blocks of the current state covariance matrix Σ. $\sigma^2_{IN_i}$ is the scalar innovation variance of y_i.

The updated system state and its covariance can be computed as

$$\begin{aligned} p(k+1|k+1) &= p(k+1|k) + K \cdot v_i(k+1), \\ \sigma^2(k+1|k+1) &= \sigma^2(k+1|k) - K \cdot \sigma^2_{IN_i}(k+1) \cdot K^T. \end{aligned} \tag{11}$$

This update is carried out sequentially for each innovation of the measurement.

Registration and deletion of landmarks When new three dimensional SIFT feature points are found on a recognized object using 3D OFM, absolute coordinates of those are computed using $y_n(x_v, h_n)$ of the measurement model. The computed absolute coordinates are added into the system state vector. We can also delete landmarks by deleting all rows and columns related to target landmarks from the covariance matrix.

3 Semantic-map-based Localization

The kidnapped robot problem refers the case where robot is lifted and manually repositioned in a different location in the environment, and has to relocate itself based on new sensor evidence. Our goal is to solve the kidnapped robot problems to locate robot position by only stereo vision information together with our semantic-map.

The relationship between absolute coordinate and relative robot coordinate system as shown in Fig. 2 is given as

$$\begin{bmatrix} x_{abs} \\ z_{abs} \end{bmatrix} = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} x_{rel} \\ z_{rel} \end{bmatrix} + \begin{bmatrix} x_t \\ z_t \end{bmatrix}, \tag{12}$$

where θ is the angle between relative and absolute coordinate and $[x_t \; z_t]^T$ is the robot position in absolute coordinate. Subtraction with two matched 3D SIFT points, we can obtain the relation as

$$\begin{bmatrix} \Delta x_{abs} \\ \Delta z_{abs} \end{bmatrix} = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} \Delta x_{rel} \\ \Delta z_{rel} \end{bmatrix}, \tag{13}$$

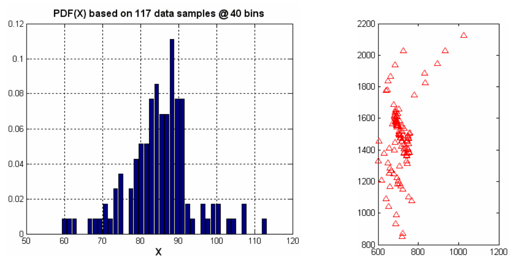
where $[\Delta x_{abs}\ \Delta z_{abs}]^T$ and $[\Delta x_{rel}\ \Delta z_{rel}]^T$ are the vectors of difference between matched points in absolute coordinate system, and in relative coordinate system, respectively. So, we can obtain the angle θ between the absolute coordinate system and the relative coordinate system by using two matched 3D SIFT points. The θ can be obtained by

$$cos\theta = \frac{[\Delta x_{abs}\ \Delta z_{abs}][\Delta x_{rel}\ \Delta z_{rel}]^T}{\|[\Delta x_{abs}\ \Delta z_{abs}]^T\| \cdot \|[\Delta x_{rel}\ \Delta z_{rel}]^T\|}. \tag{14}$$

And the robot location is computed by

$$\begin{bmatrix} x_t \\ z_t \end{bmatrix} = \begin{bmatrix} x_{abs} \\ z_{abs} \end{bmatrix} - \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} x_{rel} \\ z_{rel} \end{bmatrix}. \tag{15}$$

Ideally, if absolute and relative vectors of two pairs are correct, the robot can locate itself accurately. However, the estimated position has uncertainties because of errors in measurements and map building.



(a) the probability density function of estimated θ (b) the estimated location

Fig. 3. The statistics of the robot localization.

As shown in Fig. 3, the angle distribution indicates Gaussian-like distribution. In such a distribution, this function has a peak at the mean value. Thus, the mean value of θ is determined as the direction of robot in the absolute coordinate system. The process of robot localization is summarized as follows;

1. Eliminate outliers of θ : eliminates the outside of standard deviation
2. Compare the magnitudes of absolute and relative vectors :
   If $\|[\Delta x_{abs}\ \Delta z_{abs}]^T\|/\|[\Delta x_{rel}\ \Delta z_{rel}]^T\| > 1.5$
   or $\|[\Delta x_{abs}\ \Delta z_{abs}]^T\|/\|[\Delta x_{abs}\ \Delta z_{abs}]^T\| < 0.66$, the robot position estimated by Eq. (14) is discarded.
3. Determine the angle of the robot : the averaged value of θ is adopted as the estimated angle of the robot in the absolute coordinate system.
4. Determine robot position with the θ : averaged vector of the robot location is chosen as the estimated location of the robot.

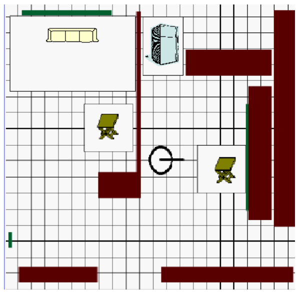
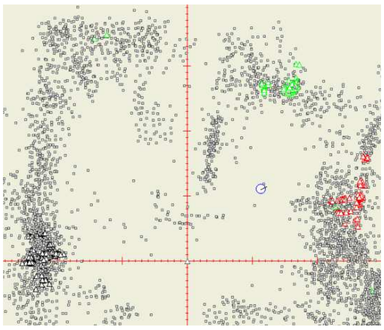


Fig. 4. the real scene of the indoor environment in which the map-building is performed.

Fig. 5. The topological map of the experimental space.



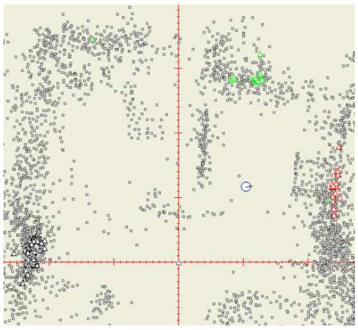


Fig. 6. The map building by only odometer.

Fig. 7. The map building by odometer and 3D SIFT.

4 Experimental Results

4.1 Performance Evaluation for Map-building and Localization

To determine how much the map-building is accurate, we solve the kidnapped robot problem at a given position of the robot. By comparing the solution of kidnap problem with the given position, we can measure the accuracy of the map.

Figure 4 shows the scene of the real indoor environment in which the map-building is performed. The circle in Fig. 5 represents the actual position of robot in the absolute coordinate system. And, robot is made to find its location by solving the kidnapped robot problems with one of maps in Fig. 6 to Fig. 9 as well as 3D SIFT features, not using odometer information. The resulting estimated position in each map is marked with circle as shown in Fig. 6 to Fig. 9.

Figure 6 shows the map which is built by relative localization by odometer. we abbreviate this method as ODO-LOC. Due to noisy odometer information, we obtain the distorted map as in Fig. 6.

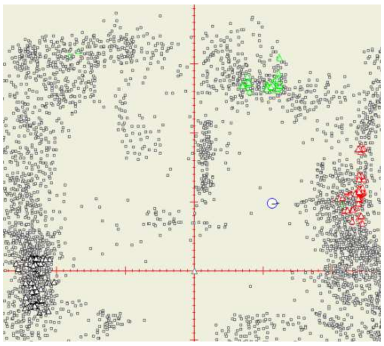


Fig. 8. The map building by odometer and 3D Harris corners.

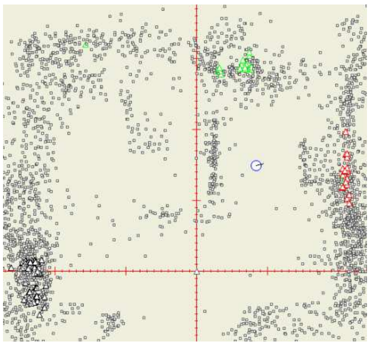


Fig. 9. The map building by 3D Harris corners and 3D SIFT.

Table 1. Averaged error between the robot actual position and the position estimated.

METHOD	X (mm)	Z (mm)	θ (degree)
ODO-LOC	277.17	880.19	19.47
ODO-SIFT-KAL	107.37	558.70	10.34
HARI-LOC	174.30	758.83	8.85
HARI-SIFT-KAL (proposed method)	127.01	218.51	8.87

Figure 7 shows the map which is built by extended Kalman filter using odometer and 3D SIFT features. We will call this method as ODO-SIFT-KAL. As shown in Fig. 7, the map appears to be more accurate when compared with the map built by ODO-LOC. This can be also observed from Table 1 by comparing solutions of kidnap problems for the ODO-LOC map and ODO-SIFT-KAL map.

Figure 8 shows the map which is built by relative localization by optical flows of 3D Harris corners. We abbreviate this method as HARI-LOC. Figure 9 shows the map-building by our proposed method in Sec. II, which will be called as HARI-SIFT-KAL. From Table 1, the map by HARI-SIFT-KAL is observed to be the most accurate map in this experiment.

5 Concluding Remarks

In this work, we have proposed autonomous semantic-map building combined with vision-based robot localization. We have used nonsymbolic SIFT-based object features with symbolic object names as landmarks and we have used vision-based relative localization as the process model of our EKF. Thus our method is able to be applied successfully to environments in which encoders are not available, such as humanoid robots.

In most previous methods, landmarks have been composed of points without semantics. Thus, an additional anchoring technique was often required for interaction. However, in our work, symbolic object names with their 3D feature location have been used as landmarks. Using such symbolic object names as landmarks is very useful when humans interact with the robot.

We also solve the kidnapped robot problem to determine how much the map-building is accurate. Note that the more is a map accurate, the more is correct the solution of kidnap problem. Consequently, in our proposed HARI-SIFT-KAL map-building, the solution of kidnap problem shows the most accurate result. Thus, it is concluded that the map by HARI-SIFT-KAL appears to be most accurate in the map-building.

Acknowledgement

This work is supported by the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Korea Ministry of Commerce, Industry and Energy. And, all correspondences of this work should be addressed to I. H. Suh.

References

1. A. J. Davison and W. Murray, "Simulatoneous Localization and Map-Building Using Active Vision," *IEEE Transaction on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 865–880, 2002.
2. A. J. Davision, "Real-Time Simultaneous Lacalisation and Mapping with a Single Camera," *Proceedings of Ninth IEEE International Conference on Computer Vision,* vol. 2, pp. 1403–1410, 2003.
3. M.W.M. G. Dissanayake, P. Newman, S. Clark, and H. F. Durrant-Whyte, "A Solution to the Simultaneous Localization and Map Building (SLAM) Problem, " *IEEE Transaction on Robotics and Automation,* vol. 17, pp. 229–241, 2001.
4. N. Doh, H. Choset, and W. K. Chung, "Accurate Relative Localization Using Odometry," *Proceedings of IEEE International Conference on Robotics and Automation,* " vol.2, pp. 1606–1612, 2003.
5. C. Harris and M. J. Stephens, "A Combined Corner and Edge Detector," *Alvey Vision Conference,* pp. 147–152, 1988.
6. D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision,* pp. 91–110, 2004.
7. B. Lucas and T. Kanade, "An Interative Image Registration Technique with an Application To Stereo Vision," *Proceedings of DARPA IU Workshop,* pp. 121–130, 1981.
8. P. S. Maybeck, *The Kalman Filter:An Introduction to Concepts,* Spinger-Verlag, 1990.
9. I. Shimshoni, "On Mobile Localization From Landmark Bearings," *IEEE Transactions on Robotics and Automation,* vol. 18, no. 6, pp. 971–976, 2002.
10. S. Se, D. G. Lowe, and J. Little, "Mobile Robot Localization And Mapping with Uncertainty using Scale-Invariant Visual Landmarks," *International Journal of Robotics Research,* vol. 21, pp. 735–758, 2002.