3D Reconstruction of Scale-Invariant Features for Mobile Robot localization

Dong-Fan Shen, Jong-Shill Lee†, Se-Kee Kil, Je-Goon Ryu††, Eung-Hyuk Lee††, Seung-Hong Hong

Dept. of Electronic Eng., Inha University, Korea
† Dept. of Biomedical Eng., Hanyang University, Korea
†† Dept. of Electronic Eng., Korea Polytechnic University, Korea

Summary
A key component of autonomous navigation of intelligent home robot is localization and map building with recognized features from the environment. To validate this, accurate measurement of relative location between robot and features is essential.
In this paper, we proposed relative localization algorithm based on 3D reconstruction of scale invariant features of two images which are captured from two parallel cameras. We captured two images from parallel cameras which are attached in front of robot and detect scale invariant features in each image using SIFT(scale invariant feature transform). Then, we performed matching for the two images' feature points and got the relative location using 3D reconstruction for the matched points.
Stereo camera needs high precision of two camera's extrinsic and matching pixels in two camera image. Because we used two cameras which are different from stereo camera and scale invariant feature point and it was easy to setup the extrinsic parameter. Furthermore, 3D reconstruction does not need any other sensor. And the results can be simultaneously used by obstacle avoidance, map building and localization. We set 20cm the distance between two cameras and capture the 3 frames per second. The experimental results show ±6cm maximum error in the range of less than 2m and ±15cm maximum error in the range of between 2m and 4m.
Key words:
DoG, SIFT, 3D reconstruction, Localization

Introduction

For a long time, diverse methods to allow robot to decide its location has been introduced around the world, but high-reliability technologies that can be actually used at home and in office environment is still at its early stage. Also, even though research is ongoing to increase reliability and to complement shortcomings by fusing environment recognition sensor including ultrasonic sensors, laser, and images, systems are complex and additional processes for sensor fusion has a lot of problems [1][2].
Compared to other sensors, image sensors retain much more environmental information. Diverse information can be extracted depending on image processing techniques, so self-localization techniques using only image sensors have recently been introduced. This kind of research includes a method using artificially manufactured land marks [3] and a method using natural land marks. When an artificial land mark is used, it is effective and the realization is easy; but, it is impossible to apply it when the environment changes. Hence the method to use natural land marks has been suggested [4].
Common research uses a method to detect a land mark by attaching a camera to a robot, which allows it to look ahead; and, many studies are ongoing, such as the latest research, vSLAM of the Evolution Robotics company [5]-[7]. Yet since the image distance information from this is obscure, relative distance, which is provided by an image and an encoder on two fixed locations, should be used in order to obtain three-dimensional location information of a land mark [8]-[11]. Therefore, an exact measurement of robot's movement is required to estimate a particular land mark and robot's relative location. However, when the measurements of robot's distance and position are encoded, it has a shortcoming that it is difficult to get an exact distance because errors are too big.
Wnuk[20] proposed an algorism to estimate locations and to build a map by using an encoder and two sets of image data, while following walls using a monocular vision system. In the case of wall following, it is possible to minimize the impact of position error in the course of the robot's movement because the angle between the robot's direction and the central axis of camera is 90° C. However, in this kind of system, one cannot use the vision system only for building a map and measuring locations to detect front obstacles.
In this paper, binocular camera system was suggested to resolve difficulties of monocular camera system which requires the location shifting of a mobile robot and an accurate measurement of shifted locations in order to detect ambient land marks. The existing binocular camera system constructs three dimensions (3D) through accurately setting up extrinsic parameters of two cameras and matching pixels. Thus, it requires a lot of calculations and it is difficult to install cameras. Therefore, in this study, we installed two

Manuscript received March 25, 2006.
Manuscript revised March 30 , 2006.

USB cameras on the upper part of a robot with a 20 cm distance between the two, and then we detected scale-invariant features from the two images using the Difference of Gaussian (DoG) and the distance to the robot by matching the two images' features.

2. Detecting Scale-Invariant Features and Matching

In order to estimate strong locations in a dynamic environment and to build a map, characteristics of natural land marks should be detected to distinguish locations in acquired images. Natural land marks commonly used indoor are doors, windows, and lightings on the ceiling. It can be difficult to recognize these natural land marks because of partial blocking of land marks, lighting changes, changes of the distance to a robot, and changes of visual points. Therefore, it is necessary to detect strong features which are invariant even when many changes as the aforementioned ones occur in images, and the Harris-Laplace method and the Difference of Gaussian (DoG) method are mainly used. The former based on the Harris corner detection has strong points in terms of changes of lighting and of visual points. However, the Laplacian Filter needs to be applied to Gaussian-Filtered images to obtain extreme values in scale space, so there are too many calculations. In contrast, the latter uses difference image of Gaussian-Filtered images, so there are only a few calculations [12],[13]. Therefore, in this paper, the Difference of Gaussian was used to extract scale-invariant features.

The vision system of a mobile robot detects land marks at a fixed distance and point, so detecting features should have invariant features in scale space. Lindeberg[19] contended that only Gaussian Filter can be a Smoothing Filter to be used to detect scale-invariant features. Therefore, one image's scale space is expressed as formula (1). That is, input image $I(x,y)$, and the Gaussian Function, $G(x,y,\sigma)$, is shown as a convolution [14].

$$L_s(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \qquad (1)$$

The DoG method is a method to effectively detect the stable locations of feature points, by using extreme points in the difference of convoluted Gaussian functions, as well as images; and this can be obtained by using formula (2).

$$\begin{aligned} DoG(x,y,\sigma) &= (G(x,y,k\sigma) - G(x,y,k\sigma)) * I(x,y) \\ &= L(x,y,k\sigma) - L(x,y,\sigma) \end{aligned} \qquad (2)$$

Here a constant, k, usually uses a multiplier of $\sqrt{2}$ [5], and $G(x,y,\sigma)$ is defined as formula (3).

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \qquad (3)$$

As the 2D Gaussian function is separable, its convolution with the input image can be efficiently computed by applying two passes of the 1D Gaussian function in horizontal and vertical directions. This paper uses the 1-D DoG kernel with 7 sample points and $\sigma = \sqrt{2}$, which Lowe [14] suggested. Figure 1 shows the 1D DoG Kernel that was used.

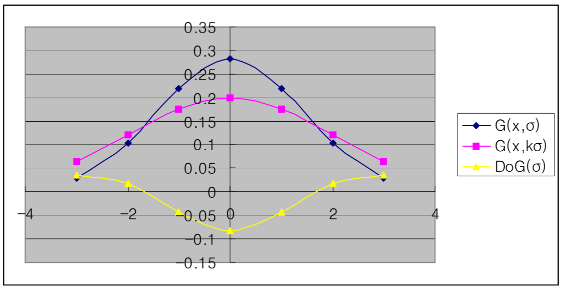


Fig. 1 1D DoG kernel.

When Input image data is like Fig 2, the process to form scale space by applying the above DoG kernel is shown in Table 1.

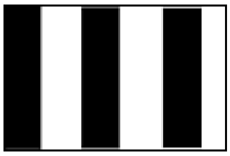
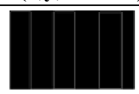
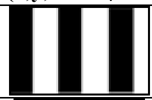
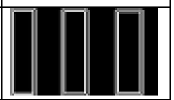
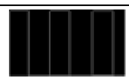
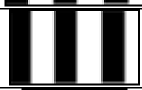
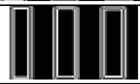
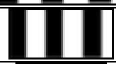
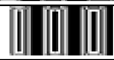
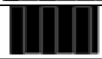
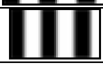
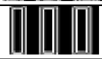


Fig. 2 Input image.

Table 1: Gaussian and DoG images at the second scale

Lev	Gaussian Images		DoG image
	L(x,y,Lev*kσ)	L(x,y,Lev* $\sqrt{2}$ kσ)	
1			
2			
3			
4			
· · ·	, · ·	· · ·	· · ·
8			

In Table 1, when $k = 1$, $\sigma = \sqrt{2}$, and 8 scale levels (Lev), Gaussian images and DoG image were shown at each level, which were created by applying Gaussian Smoothing Filter two times.

In created DoG images, with locations of local maximum and minimum points, scale invariant feature points are selected as candidates. Local maximum and minimum points were used to compare each pixel of DoG image at each level with total 26 pixels - 8 ambient pixels, 9 pixels corresponding to adjacent high level and 9 pixels to low level - and made decisions.

In order to correct the stability of lighting, candidate points found in the DoG images were used to obtain magnitude (m) and orientation (θ) by using formulas (4) and (5), in terms of local images ($L_{x,y}$) in which a radius of each feature point location is 9 pixels.

$$m = \sqrt{(L_{x+1,y} - L_{x-1,y})^2 + (L_{x,y+1} - L_{x,y-1})^2} \quad (4)$$

$$\theta = \tan^{-1}\left(\frac{L_{x,y+1} - L_{x,y-1}}{L_{x+1,y} - L_{x-1,y}}\right) \quad (5)$$

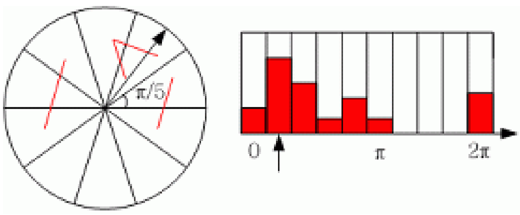


Fig. 3 An orientation histogram and feature point orientation.

Fig. 3 An orientation histogram and feature point orientation

As shown in the left one of Figure 3, azimuth angles of all sample points within the circle area surrounding feature points were obtained; and as in the right one, a histogram was drawn up by quantize an azimuth to 10 levels and an angle with the greatest value was set as a standard orientation of feature points. As the distribution on the histogram slowly responds to lighting changes, a standard orientation has a strong feature against lighting changes.

Feature points detected by the method to detect scale invariant feature points used the SIFT Descriptor of Lowe [15], [16]. A standard orientation was obtained by using the method to get the fore-mentioned azimuth within the area surrounding the feature points, and a new orientation was obtained by subtracting a standard orientation from the orientation of pixels in the surrounding area, which becomes a "canonical orientation" invariant to rotation transformation of images. Orientation Histogram made by quantizing a canonical orientation, which is invariant to rotation transformation, to 8 angles was used as feature vectors for feature points. In this paper, as a histogram quantized to 8 angles for the 4×4 area was used, each feature point becomes a feature vector of the 4×4×8=128 dimension. Therefore, the descriptor of the scale invariant feature points was composed of the 128 feature vectors and coordinates, scale, and standard orientations.

The process to match feature points on immovable images is done through comparison of feature vectors. In this paper, it was estimated if two feature points are the same points by comparing a sum of square of feature vector differences with a threshold value, which Lowe [14] suggested. If the threshold value is larger, then there are many matching points but there are points matched incorrectly; if the threshold value is smaller, then the accuracy of matching improves but the number of matched points decreases. Since the two cameras are installed in parallel, in addition to the comparison of feature vectors, the matching of two images can be done more accurately by judging if they have the same scale level and canonical orientation. However, the matching process with reference image can encounter image changes that occur due to various changes such as lighting change, scale change, rotation change, or a combination of these changes, so more stable feature points should be detected.

In this study, feature points that are sensitive to noise and have low contrast were first removed with the method suggested by Brown and Lowe [15]. In order to ensure the stability of feature points, removing the sample points with low contrast is not sufficient. The DoG actually gets extreme values at edge. Yet depending on edge, points with extreme value of a small DoG value are sensitive to little noise and become unstable points, and thus should be removed.

In the DoG image, extreme values that have small values at edge get a big principle curvature for a horizontal direction of edge, but a principle curvature with small values for the vertical direction. Principle curvatures can be obtained by using a 2×2 Hessian matrix (H_h) as Formula (6) at the location and the scale of a key point.

$$H_h = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (6)$$

Derivatives D_{xx}, D_{xy}, D_{yy}, can be obtained by using the difference around sample points in the DoG image. When the method by Harris and Stephens [17] is used, it can be obtained without getting an Eigenvalue, yet only with ratio as shown in Formula (7).

$$\frac{trace(H_h)}{\det(H_h)} < \frac{(r+1)^2}{r} \quad (7)$$

In order to remove feature points with the ratio between principle curvatures of more than 10, r was set as 10 and feature points with a value for the ratio between curvatures of more than 10 were removed. Through the removal of features points with low contrast and the improvement of edge response characteristics, the accuracy of matching scale-transformed images only by comparing feature vectors improved.

3. Detecting Distance of Feature Points and Localization

In order to detect a distance to feature points, a process to transform pixel coordinates of an image plane with a coordinate of the real world. Figure 4 shows the process of transforming coordinates with an ideal pin-hole camera model. M is a location of 3D real-world coordinate, and m is a location of image plane by ideal lens. The center point c on the camera coordinate system is an optical center, and the distance between c and p is the focal distance, f [18]. In this model, all points on extension lines of $\vec{Cm}$ have the same $m(u,v)$ on the image plane.

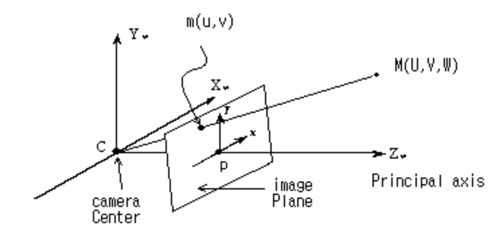


Fig. 4 Pin-hole camera model for geometrical analysis.

To represent this characteristic of reflection, the transformation of image coordinates and real-world coordinate can be defined by using the perspective coordinate system. By adding a scale vectors to, it is expressed as $\tilde{m} = [x, y, 1]^T$ $\tilde{M} = [X, Y, Z, T]^T$. The relation between perspective coordinates and real coordinates is U=X/T, V=Y/T, and W=W/T. A reflection equation using the perspective coordinate system is defined as Formula (8).

$$\tilde{m} = A\begin{bmatrix} R & T \\ 0_3 & 1 \end{bmatrix}\tilde{M}, \qquad A = \begin{bmatrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

In Formula (8), $[R,T]$ is the extrinsic parameter of the camera, expressing the rotation and moving transformation between real-world coordinate system and the camera coordinate system, and A indicates the intrinsic parameter of the camera.

In order to get the distance of feature points by using Formula (8), first the intrinsic parameters of two camera's A_l and A_r are obtained through offline camera calibration. For the convenience of calculation and installation, the two cameras are set in parallel, having constants, R and T. Parameters used in this paper are defined as in Formula (9). In Formula (9), d is the distance between two cameras.

$$R_l = R_r = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad T_l = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad T_r = \begin{bmatrix} d \\ 0 \\ 0 \end{bmatrix}$$
$$A_l = A_r = \begin{bmatrix} 278.7 & 0 & 155.2 \\ 0 & 272.5 & 118.4 \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$

Through matching feature points, extrinsic parameters for the same point, M, in real-world get the other two image coordinates, $\tilde{m}_l$ and $\tilde{m}_r$; so, by substituting Formula (8), the 1st linear simultaneous equation as Formula (10) can be obtained.

$$\begin{cases} \tilde{m}_l = \tilde{A}_l \begin{bmatrix} R_l & T_l \\ 0_3^T & 1 \end{bmatrix} \tilde{M} \\ \tilde{m}_r = \tilde{A}_r \begin{bmatrix} R_r & T_r \\ 0_3^T & 1 \end{bmatrix} \tilde{M} \end{cases} \tag{10}$$

As Formula (10) is 4 linear simultaneous equation, it can obtain 4 unknowns, the reflection coordinates , and real-world coordinates that has the center of camera as the origin, M(U, V, W), can be obtained with U= X/T, V= Y/T, and W= W/T. The actual localization process through detecting feature points and the distance to the robot is shown in Figure 5.

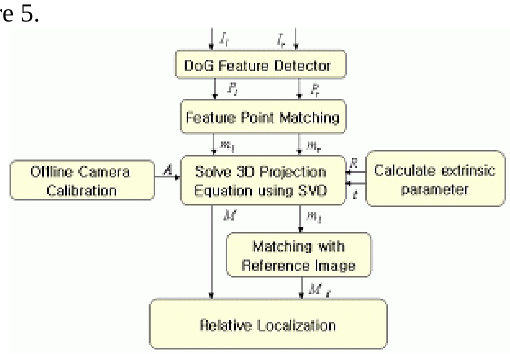


Fig. 5 Flowchart of relative localization.

For input images of two cameras I_l and I_r the DoG method was applied to each to get features P_l and P_r . Two perspective coordinates, m_l and m_l , for the same point are obtained through matching P_l and P_r; 4 linear simultaneous equation are formulated by using the camera's intrinsic parameter, A, and extrinsic parameters, R, and T, obtained from the pre-treatment process; and, a solution to the linear equation can be acquired by using the SVD method. And, detected feature points and the distance between cameras, M, are obtained by using the relation of perspective coordinates and real-world coordinates; and, it is judged whether there are any obstacles in the direction of movement, using the distribution of feature points' distances. For localization, first, reference image, feature points descriptor, and distance information between feature points and robot are calculated and saved. Next, relative localization is done by comparing the measured distance, M, and saved distance information of feature points detected through matching input image and saved reference image, M_d.

When the location of a feature point on reference image is $M_r(X_r, Y_r, Z_r)$, and the location of a feature point on input image, which is matched with reference image, is $M_i(X_i, Y_i, Z_i)$, two feature points satisfy Formula (11). In indoor environment, a robot navigates on a plane ground, so it satisfies Formula (11).

$$\begin{bmatrix} X_r \\ Z_r \end{bmatrix} = \begin{bmatrix} X_i \\ Z_i \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} + \begin{bmatrix} T_x \\ T_z \end{bmatrix} \tag{11}$$

In Formula (11), T_x, T_y and θ are a current location and position of the robot corresponding to reference location respectively. Therefore, when every location has more than two matching points, a robot's moved location and position can be obtained by using Formula (11).

4. Experiments and Results

Figure 6 shows the experimental environment. Figure 6 - (left) shows a camera installed in Hanuri-RD robot of the Hanwool Robotics Company, and Figure 6 - (right) shows the whole system that receives input of cameral image from Centrino 1.6G IBM laptop through a USB port and controls the robot. Two Logitech USB cameras were used.

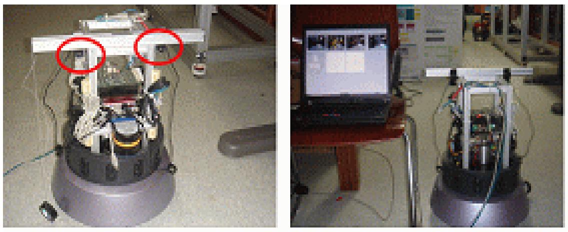


Fig. 6 Experimental environment.

For localization through a vision system, the detection of scale invariant features and the measurement of exact distance are the keys. Hence, an experiment in which the same objects at intervals of 30cm measure the distance while moving was conducted. Figure 7 shows input images from left-side and right-side cameras when a 70cm-wide board is at 90cm and 120cm. Figure 8 shows matched scale-invariant feature points on input images from left-side and right-side cameras and the matching relation on image plane. Figure 9 shows the 3D coordinates of matched feature points. The grid size of Figure 9 is 30x30cm.

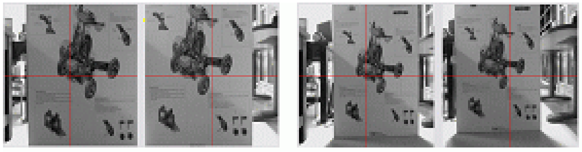


Fig. 7 Two camera input image. (left) 90cm. (right)120cm.

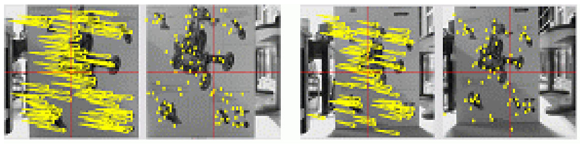


Fig. 8 Matched Feature points. (Left)90cm. (right)120cm.

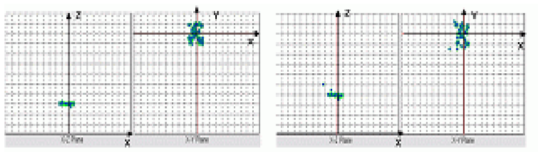


Fig. 9 Measured 3D location for features. (left)90cm. (right)120cm. (grid size is 30x30cm).

Table 2: Results of the measured distance

Real distance(cm)	Number of Matched points	Distance Error(cm)
60	97	±2
90	72	±2
120	55	±2
150	49	±3
180	32	±6
210	25	±15

Table 2 is the results of the above experiments conducted at 60cm - 210cm at intervals of 30cm. According to the table, when the real distance is less than 150 cm, the distance errors are within 10 cm, and as the distance increases, so does the error.

In order to observe the measurement accuracy for an object with more complex structure, two boards were arranged at angles of 90 degrees and of 120 degrees and the 3D coordinates of feature points were obtained from a fixed distance. Figure 10 is input image and Figure 11 and 12 shows matched feature points and the 3D coordinates of feature points. In Figure 12, one can see that the distribution of detected feature points in fact expresses the arrangement angles of board

Fig. 9 Input Image. (left) 120 degree. (right) 90 degree.

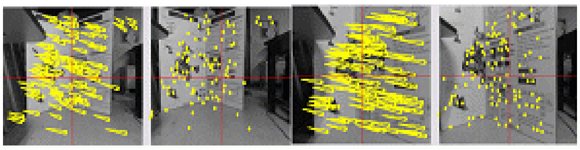


Fig. 10 Matched points. (left) 120 degree. (right) 90 degree.

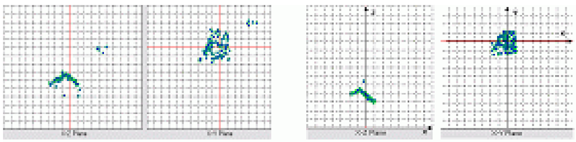


Fig. 11 Measured 3D location of feature points. (left) 120 degree. (right) 90 degree. (grid size 30x30cm)

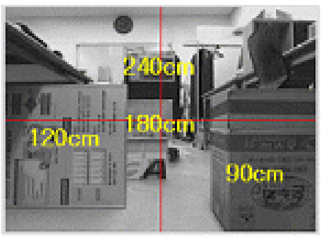


Fig.12 Input image.

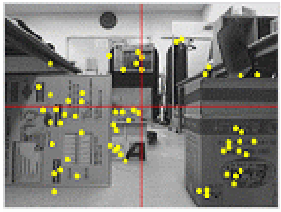


Fig. 13 Matched points.

In the robot navigating environment, input images of the camera system reflect various objects placed at fixed distances and fixed angles. Therefore, the important requirements to search for the path and to avoid the obstacles are to distinguish objects placed at fixed distances and directions in input images and to accurately measure each distance. Hence, we conducted an experiment with installing various obstacles at a fixed distance in the direction of the robot's navigation. Figure 12 shows input images and the distance information of obstacles. Figure 13 shows matched figure points on left-side and right-side input images; and, Figure 14 shows the distribution of matched feature points on the x-z plane. As shown in the figure, front obstacles at 90cm, 120cm, and 180cm were accurately measured.

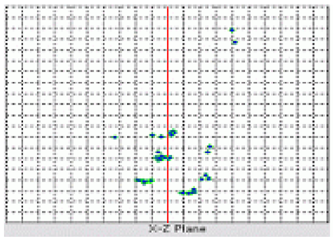


Fig.14 Distance of feature points.

After saving the reference images and the detected feature points for localization experiment, we got input images of a fixed location, which included the saved feature points and experimented to match with reference images. Figure 15 shows the saved reference images, feature points, and the 3D coordinate. Figure 16, 17, and 18 shows images which was input after the robot moved, feature points matched with reference images, and the 3D coordinate of feature points. Since coordinate values of feature points detected at the fixed locations are the values of the coordinate system centered on the robot, the 3D coordinate of feature points in reference images and the 3D coordinate of corresponding feature points in input images can be used, through simple coordinate transformation, to estimate relative locations corresponding to reference locations.

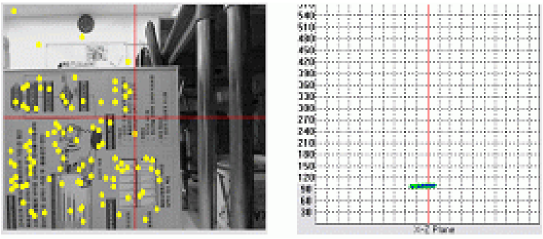


Fig. 15 Reference image feature points and measured distance.

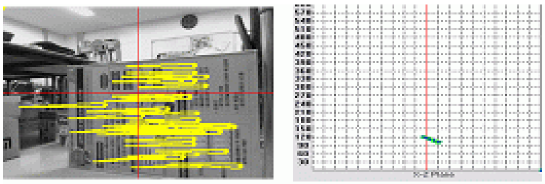


Fig. 16 Left rotated input image, matched feature points with reference image.

Fig. 17 Right rotated image and matched feature points with reference image.

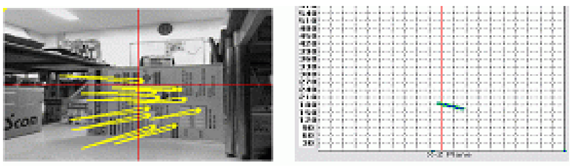


Fig. 18 Rotation, moved image and matched feature points with reference image.

Table 3 is the results of relative localization by applying the above results to Formula (11).

Table 3: Results of relative localization

	Matched Features	Relative Location	Angle	Distance Error (cm)	Angle Error (Degree)
Fig.16	28	(-18, -10)	-27.9°	2.24	1.1
Fig.17	31	(5.2, 6.4)	10.3°	1.5	1.7
Fig.18	12	(13.4, 5.3)	-24.8°	8.9	3.1

After we conducted above experiments from various angles and locations, we found that: the result of relative localization within 2m and -30 ~ 30 was location error was on average ±11cm and the mean error of position was ±3. In the range beyond this, there were few matched feature points and too much matching error.
The processing speed of the proposed method was total 330ms - 200ms to capture two frames of 320 x 240 gray image input from left- and right-side cameras, 60ms to detect feature points from the two images, 20ms per 100 feature points for matching left and right images and detecting the distance, and 30ms to match with reference image. Table 4 shows the comparison of the proposed method and the method using vSLAM [4]. According to this, the proposed method measures the distance more accurately than vSLAM, does not depend on the information of an encoder, and is able to extract the distance information of ambient objects even at idle state. However, compared to vSLAM that uses only one camera, this method using two cameras requires more resource. Therefore, even though a system that uses a monocular camera like vSLAM is inexpensive but since it depends on the information from an encoder, it is difficult to expect a high accuracy of distance measurement and localization.

Table 4: Comparison with vSLAM

	Proposed Method	vSLAM
Distance Accuracy	±10cm(Maximum)	±20cm(median)
No. of Camera	2	1
Encoder Information	Not used	Used
Resource	Much	Little

5. Conclusions and Future Research

In this study, the DoG method was applied to two images captured from two general USB camera attached in parallel to a mobile robot; scale invariant features were detected; 3D coordinates of features were obtained through matching; and relative localization was performed through matching with reference image. Since the method proposed in this study uses two cameras, it demands more resource than a system that uses one camera, but it can carry out an accurate localization without any assistance from foreign sensors like an encoder, and it is capable of avoiding any obstacles in front. Also, it is easy to install compared to the existing binocular cameras, and due to its fewer amount of calculations, it can be utilized for a robot's intelligent

navigation within complex and dynamic environments like an indoor one
For global localization in future, building a map database based on feature points is necessary. Also, it is required to develop an effective way to find matched feature points from numerous reference images and feature points stored in the map database.

Acknowledgments

This study was supported by a grant of the project "Core Technology Development Program", Misnistry of Commerce, Industry and Energy(MOCIE), Republic of Korea.

References

[1] J. Borenstein, H.R. Everett and L. Feng, Where am I? - Sensors and Methods for Mobile Robot Positioning, university of Michigan, 1996.

[2] A.J. Davison and D.W. Murry, "Simultaneous Localization and Map-Building Using Active Vision," IEEE Trans. on PAMI, Vol.24, No.7, July, 2002.

[3] 권인소, 장기정, “인공표식과 자연표식을 결합한 강인한 자기위치 추정”, 한국정밀공학회, 춘계학술대회논문집, 2003.

[4] Niklas Karlsson, et al., "Core Technologies for Service Robotics," Proc. of Int. Conf. on IROS2004, 2004.

[5] D. Lowe, "Local feature view clustering for 3D object recognition," Proc. of the 2001 IEEE/RSJ, Int. Conf. on Computer Vision and Pattern Recognition, Kauai, Hawaii, USA, December 2001.

[6] S. Se, D. Lowe and J. Little, "Local and global localization for mobile robots using visual landmarks," Proc. of the 2001 IEEE/RSJ, Int. Conf. on Intelligent Robots and Systems, Maui, Hawaii, USA, October 2001.

[7] J. Wolf, W. Burgard and H. Burkhardt, "Robust vision-based localization for mobile robots using an image retrievel system based on invariant features," Proceedings of the 2002 IEEE Int. Conf. on Robotics and Automation, Washington, DC, USA, May, 2002.

[8] D. Fox, W. Burgard and S. Thrun, "Markov localization for mobile robots in dynamic environments," Journal of Artificial Intelligence Research 11, pp. 391-427, 1999.

[9] J.S. Gutmann and K. Konolige, "Incremental Mapping of Large Cyclic Environments," Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), San Francisco, CA, 2000.

[10] S.I. Roumeliotis and G.A. Bekey, "Bayesian estimation and Kalman filtering : A unified framework for mobile robot localization," Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 2985-2992, San Francisco, CA, 2000.

[11] S. Thrun, Mapping : A Survey, Technical Report, CMU-CS-02-111, Carnegie Mellon University, Pittsburgh, PA, USA, February, 2000.

[12] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," IEEE Conference on Computer Vision and Pattern Recognition, pp.257-263, June 2003.

[13] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," in Proceedings of the 8th International Conference on Computer Vision, pp.525-531, 2001.

[14] G. Lowe, "Distinctive image features from scale-invariant keypoints," preprint, submitted to International Journal of Computer Vision, 2004.

[15] M. Brown and G. Lowe, "Invariant features from interest point groups," In British Machine Vision Conference, Cardiff, Wales, pp.656-665, 2002.

[16] G. Lowe, "Object recognition from local scale-invariant features," International Conference on Computer Vision, Corfu, Greece, pp.1150-1157, 1999.

[17] C. Harris and M. Stephens, "A combined corner and edge detector," In 4th Alvey Vision, Conference, Manchester, UK, pp. 147-151, 1988.

[18] Zhengyou Zhang, "Flexible Camera Calibration By Viewing a Plane From Unknown Orientations," The Proceedings of the 7th IEEE International Conference on Computer Vision, Vol. 1, pp. 666-673, 1999.

[19] Lindeberg, Tony, "Scale-space theory: A basic tool for analysing structures at different scales," Journal of Applied Statistics, 21, 2(1994), pp.224-270.

[20] Kamil Wnuk, Faith Dang, and Zachary Dodds, "Dense 3D Mapping with Monocular Vision," Int. Conf. on Autonomous Robots and Agents, pp.212-217, 2004.

Dong-Fan Shen received the B.S. degree in Computer Engineering fromYanBian University, JiLin, China, in 1998. and the M.S. degree in Electronic Engineering from School of Electronic Engineering from Inha University, Incheon, Korea in 2001. He has been received the Ph.D. course in Electronic Engineering at the School of Electronic Engineering from Inha University, Incheon, Korea. His main research interests are in the areas of robot vision, embedded system.

Jong-Shill Lee received the B.S., M.S. and Ph.D degrees in Electronic Engineering from Inha University in 1995, 1997 and 2005, respectively. From 2001 to 2005, he was a lecturer in the Department of Electronic Engineering, Korea Polytechinc University. He has taught courses in digital signal processing and digital image processing. He is currently a Research Professor with the Department of Biomedical Engineering, University of Hanyang. His research activities are mainly in the areas of digital signal processing, bio signal processing, rehabilitation system, robot vision system.

Se-Kee Kil received the B.S. degree in Electronic Engineering from Inha University, Incheon, Korea, in 1998. and the M.S. degree in Electronic Engineering from School of Electronic Engineering from Inha University, Incheon, Korea in 2000. He has been received the Ph.D. course in Electronic Engineering at the School of Electronic Engineering from Inha University, Incheon, Korea from 2000 His main research interests are in the areas of biomedical signal processing, healthcare system and various industrial applications.

Je-Goon Ryu received the B.S. degree in Electronic Engineering from Inha University, Incheon, Korea, in 1999, and the M.S. degree in Electronic Engineering from Inha University, Incheon, Korea, in 2004, respectively. From 1999 to 2001, he was a junior researcher at R&D of MECA Information Communication Co. Ltd. Since 2004, he has been a junior researcher at Intelligence Healthcare system Research Center, Korea Polytechnic University. His main research interests are in the areas of service robot control, mobile healthcare system, robot vision, embedded system, and various industrial applications.

Eung-Hyuk Lee received the B.S. degree in Electrionics Engineering from Inha University, Inchon, Korea, in 1985, and the M.S. degree and the Ph.D. degree in Electronic Engineering from School of Electronic Engineering from Inha University, Inchon, Korea, in 1987 and 1987, respectively. From 1987 to 1992, he was a researcher at Industrial Robot Lab. of Daewoo Heavy Industry Co. Ltd. From 1995 to 2000, he was a assistive professor at Dept. of Computer Engineering in Konyang University. Since 2000, he has been with the Department of Electronics Engineering at Korea Polytechnic University. His main research interests are in the areas of service robot control, mobile healthcare system, image processing and various industrial applications.

Seung-Hong Hong received the B.S. degree in Electrical Engineering from Inha University in 1963. He received the M.S. degree in Electrical Engineering from Inha University in 1966, and the Ph.D. degree at the School of Biomedical Engineering from Tokyo University, Japan in 1975, respectively. He was a director and chairman at the Institute of Electronic Engineer of Korea from 1981 to 1994. He was a director, vice chairman, and chairman at the IEEE Korea Section from 1983 to 1997. He has been with the Department of Electronics Engineering at Inha University. His main research interests are in the areas of bio-signal processing, rehabilitation engineering, medical image processing.