JOURNAL OF KOREA MULTIMEDIA SOCIETY, VOL. 7, NO. 12, DECEMBER 2004(pp. 1745-1753)

Pose Estimation of 3D Object by Parametric Eigen Space Method Using Blurred Edge Images

Jin-Woo Kim†

ABSTRACT

A method of estimating the pose of a three-dimensional object from a set of two-dimensioal images based on parametric eigenspace method is proposed. A Gaussian blurred edge image is used as an input image instead of the original image itself as has been used previously. The set of input images is compressed using K-L transformation. By comparing the estimation errors for the original, blurred original, edge, and blurred edge images, we show that blurring with the Gaussian function and the use of edge images enhance the data compression ratio and decrease the resulting from smoothing the trajectory in the parametric eigenspace, thereby allowing better pose estimation to be achieved than that obtainable using the original images as it is. The proposed method is shown to have improved efficiency, especially in cases with occlusion, position shift, and illumination variation. The results of the pose angle estimation show that the blurred edge image has the mean absolute errors of the pose angle in the measure of 4.09 degrees less for occlusion and 3.827 degrees less for position shift than that of the original image.

Keywords : Pose estimation, K-L transformation, parametric space

1. INTRODUCTION

Recognition of three-dimensional (3D) objects from two-dimensional (2D) images and estimating their spacial pose important research fields with a wide range of applications in a variety of areas, including bin picking, parts manufacturing[1], and camera control[2-4]. The representative method of finding a specific object in an image is template matching. In this method, a template with the same shape as the specific object is first made. The template is then moved over the image, and the location that has the highest correlation with the template is selected as the position of the specific object. However, this method is mot efficient in some cases-such as with images that have a large amount of noise, occlusion, illumination variations, and/or pose changes. One way to deal with such cases is to memorize all the 2D appearance images of the 3D object and then compare the input image with them, but this is often impractical due to the huge amount of computer memory required. Many algorithms that have been proposed for pose estimation make use of eigenspace[5,6], which gives the advantage of information compression between similar objects. Such algorithms have, for instance, been reported for pose estimation of facial images[7-11], as well as for the identification of individuals[12]. In these cases, the input images are facial images taken from different directions. Murase and Nayer[13-15] have proposed methods of pose estimation, as well as object classification, that employs and an illumination planning to be most distinguishable in appearance from each other using the nearest trajectory in an eigenspace using various kinds of objects and lighting conditions. However, they still use the original image as the input image. Therefore, when there is an occlusion, the model and the object do not coincide well, and

※ Corresponding Author : Jin-woo Kim, Address : (608-736) 314-79 Daeyeon-Dong, Nam-Gu, Busan, Korea, TEL : +82-51-607-5153, FAX : +82-51-625-1402
E-mail : jinwoo@ks.ac.kr
Receipt date : May 17, 2004, Approval date : Oct. 6, 2004
† Dept. of Multimedia Eng., Kyungsung Univ.

as a consequence it may not be possible to classify and estimate the parameter (the pose in this case). An algorithm that uses an edge image instead of the original image may be make effective in pose estimation because it emphasizes the edges, which are features of an object, while reducing the amount of redundant information. Here, the eigenspace needs to enhance the contribution ratio. This is made possible by blurring the edge image, where the correlation value between neighboring points becomes higher.

In this paper, we propose a method of composing the eigenspace from Gaussian blurred edge images in order to be able to estimate stably the pose of an object even when occlusion, shift of position, and/or change of illumination occur. We show that blurring by means of the Gaussian function and edge extraction enhance the data compression ratio and decrease the error due to smoothing. We also examine the estimation errors arising from a large positional shift, as well from expansion or reduction in object size, versus the number of eigenspace dimensions, and show the effectiveness of using blurred edge images in dealing with such disturbances. Our results demonstrate that the proposed method of edge blurring is effective for accurately estimating three-dimensional poses using sets of two-dimensional images.

2. MAKING AN IMAGE SEQUENCE

The 72 image frames used in the experiments were prepared as follows. An image sequence was obtained by placing an object (a teapot) on a turntable and rotating it. At rotating intervals 5 degrees, images were taken by a CCD camera and saved in a frame memory connected to a computer. The light source was fixed at one place. Using a Sobel filter, edge images were made and then blurred by a Gaussian filter (standard deviation, σ = 2.0). The resultant images are hereafter referred to as blurred edge images. Both the original and blurred edge images were normalized as described in the following section. All the images used had a uniform size 64X64 pixels. Fig. 1 shows examples of images obtained by rotating the object: (a) original images, (b) blurred original images, (c) edge images, and (d) blurred edge images.

3. EIGENVECTOR CALCULATION[16]

First, raster vectors $\hat{\mathbf{x}}_m (N^2 \times 1)$ are obtained by raster scanning of the m-th blurred edge images $(\mathrm{F}_m = [\hat{\mathbf{f}}_{ij}] : N \times N ; m = 1,2,3,\cdots,M)$ as follows, where M is the number of images:

$$\hat{\mathbf{x}}_m = [\hat{\mathbf{f}}_{11}, \hat{\mathbf{f}}_{12}, \cdots, \hat{\mathbf{f}}_{NN}]^T$$

To remove the effects of intensity differences among input images (due to object position and sensor condition), the raster vectors are normalized:

$$\mathbf{x}_m = \frac{\hat{\mathbf{x}}_m}{\|\hat{\mathbf{x}}_m\|} \tag{1}$$

Examples of blurred edge images are shown in Fig. 1(d). Since these images are the blurred results of edge enhanced images, the correlation between two consecutive images in this sequence is high[17]. The average of the images is determined using the equation

$$\mathbf{A} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{x}_m \tag{2}$$

The image convariance matrix $\mathbf{C}$ is then calculated as follows:

$$\mathbf{C} = \frac{1}{M} \sum_{m=1}^{M} (\mathbf{x}_m - \mathbf{A})(\mathbf{x}_m - \mathbf{A})^T \tag{3}$$

The $\mathbf{C}$ matrix can be expressed thus:

$$\begin{bmatrix} C(0,0) & \cdots & C(0, N^2-1) \\ C(1,0) & \cdots & C(1, N^2-1) \\ \cdots & \cdots & \cdots \\ C(N^2-1,0) & \cdots & C(N^2-1, N^2-1) \end{bmatrix}$$

Finally, the eigenvalues are calculated using the equation

$$\mathbf{C}\mathbf{e}_i = \lambda_i \mathbf{e}_i \qquad (4)$$
$$(i = 1, 2, 3, \cdots, N \times N)$$

The eigenvalues are arranged in descending order and the corresponding base eigenvectors ($\mathbf{e}_1, \cdots, \mathbf{e}_k$) are obtained. Using these vectors, it is possible to compose an eigenspace that can efficiently compress the features of the images. The trajectory of highly correlated images in the eigenspace will be compact and smooth. Fig. 2 shows examples of eigenvectors (images), in which (a)~(d) correspond to frames (a)~(d) in Fig. 1.

4. PROJECTION TO PARAMETRIC EIGENSPACE

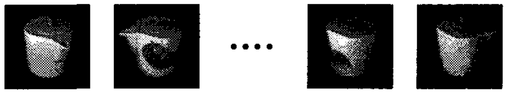
Subtracting the average images $\mathbf{A}$ from each image, the following equation gives the corresponding projection points $\mathbf{g}_{(p)}$, where p is a parameter describing the pose:

$$\mathbf{g}_{(p)} = [\ \mathbf{e}_1,\ \mathbf{e}_2, \cdots,\ \mathbf{e}_k\ ]^T(\ \mathbf{x}_m - \mathbf{A}\ ) \qquad (5)$$

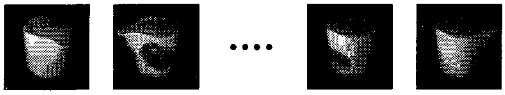
Generally, when the difference between two poses of an object is small, the images are highly correlated and their projected positions in the eigenspace are also close to each other. Fig. 3(•) shows the parametric eigensubspace trajectories (dictionary) for the object in Fig. 1 composed of only 3 eigenvectors in descending order. The sequences of the projected points are depicted as continuous smooth curves using B-spline interpolation. An unknown pose of an object can ascertained by projecting the image to the eigensubspace and estimating its pose as the nearest point on the trajectory in the manner detailed below.

5. POSE ESTIMATION OF AN OBJECT

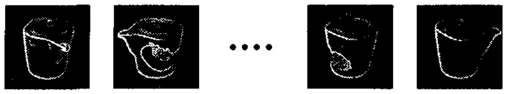
We assume here that the size and brightness of



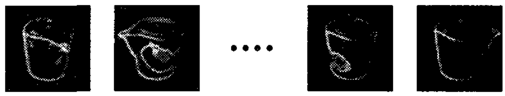
(a) Original images



(b) Blurred original images



(c) Edge images

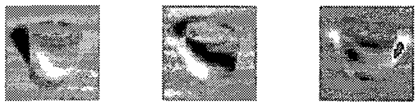


(d) Blurred edge images

Fig. 1. Set of images obtained by rotating an object(teapot).

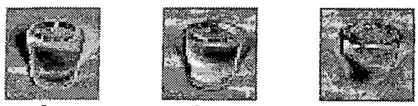
$\mathbf{e}_1$ $\mathbf{e}_2$ $\mathbf{e}_3$

(a) Original images



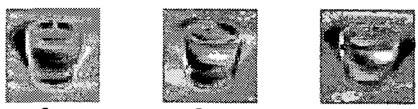
$\mathbf{e}_1$ $\mathbf{e}_2$ $\mathbf{e}_3$

(b) Blurred original images



$\mathbf{e}_1$ $\mathbf{e}_2$ $\mathbf{e}_3$

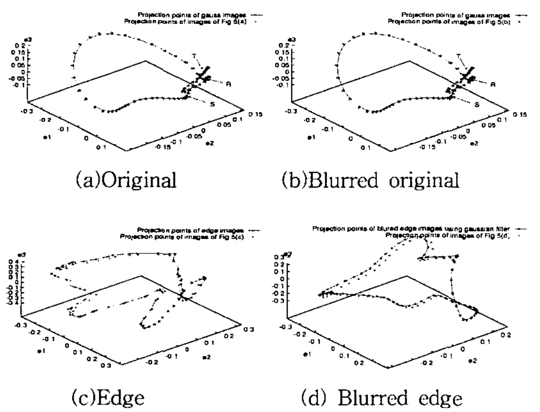
(c) Edge images



$\mathbf{e}_1$ $\mathbf{e}_2$ $\mathbf{e}_3$

(d) Blurred edge images

Fig. 2. Eigenvector images of the images shown in Fig. 1.



(a)Original (b)Blurred original

(c)Edge (d) Blurred edge

Fig. 3. Trajectories for the object depicted in Fig. 1 in parametric eigensubspace and projected points of occluded images.

an input image are normalized by preprocessing. The vector of a normalized input image is defined as $\mathbf{y}$ and is projected as the point $\mathbf{z}$ in the eigensubspace by the equation

$$\mathbf{z} = [\ \mathbf{e}_1,\ \mathbf{e}_2, \cdots,\ \mathbf{e}_k]^T(\ \mathbf{y} - \mathbf{A}) \tag{6}$$

The pose estimation value $\hat{p}_\theta$, pose estimation error ε_θ , relative distance error ε_d , and distance error $\hat{d}$ of the object are defined as follows:

$$\hat{p}_\theta = arg\min \|\ \mathbf{z} - \ \mathbf{g}_{(p)}\| \tag{7}$$

$$\varepsilon_\theta = p_\theta - \hat{p}_\theta \tag{8}$$

$$\varepsilon_d = \frac{\hat{d}}{\|\ \mathbf{z}\|} \tag{9}$$

$$\hat{d} = \min \|\ \mathbf{z} - \ \mathbf{g}_{(p)}\| \tag{10}$$

Here, p_θ is the correct pose value.

6. POSE ESTIMATION EXPERIMENTS

In the original image Eigenspace, the error of pose estimation of the object increases according to the variants of occlusion, shift of position and lighting conditions, therefore it causes a problem for the estimation of an objects pose. This problem is very important for utilization of an outside environment. To solve this problem, firstly the edge image is used, instead of the original image, for this paper. The edge image emphasizes the edge of a subject object and reduces the amount of information of the other parts, so it is considered to make the estimation of pose influential for the case of occlusion. For the pose estimation of an object using Eigenspace method, the contribution ratio is needed to be enhanced in the low-dimension. However, we conduct the Gaussian function to raise the value of the correlation between neighboring images. Using the property of this correlation, we can expect the elevation of information compression ratio by K-L transformation ; this principle uses the fact that certain kinds of pose estimations can be done easily when the occlusion and shift of position of image occurs for pose estimation through parametric Eigenspace method.

To judge the effectiveness of our proposed method, we projected original, blurred original, edge, and blurred edge images to the eigen-subspace and then examineed how the minimum distance between the projected points and the trajectories of the images varied in the presence of occlusion, position shift, and illumination variations. The contribution ratio plotted against the number of dimensions for the original, blurred original, edge, and blurred edge images, are shown in Fig. 4 in descending order of eigenvalue. The plots clearly show that the highest degree of compressibility is achieved with the blurred edge image method, especially in the case of a small number of dimensions.

6.1 Occlusion

Fig. 5(a)~(d) shows the frame sequences of the original, blurred original, edge, and blurred edge images from Fig. 1 in which the top-right portion of the object is hidden. The occlusion window applied to the images is depicted in Fig. 5(e). We adopted the three prime eigenvectors from the maximun eigenvalue in descending order. The contribution ratios are (a)0.393460, (b)0.650857, (c)0.597741, and (d)0.72065, respectively. As mentioned earlies, Fig. 3(a)~(d) depicts the parametric eigensubspace trajectories of the dictionary (solid lines over) 360° obtained with images taken at intervals of 5° during rotation. The series of projected points for the occluded object (mark : ×) are also shown for each of the image types. The total number of the four kinds of occluded images was 288 (72×4), and the occluded area comprised 11.7% of the teapot image.

Fitst, in Fig. 3 we can see that the trajectories in the parametric eigensubspace of (a) and (b) resemble each other in that their shape is simple except for the right-hand side where the trajectories are folded. On the other hand, the trajectory in (c) is more complex but well spread, thereby separating each pose and reflecting the characteristics of the edge features. The trajectory in (d) is simplee than that in (c) because of edge blurring by the Gaussian function.

Next, the data from Fig. 3 are redrawn in Fig. 6 in order to clarify the estimation errors. In each frame of Fig. 6, the correct values of the pose are shown as a straight, diagonal dotted line and the estimated values by larger dots (•). Deviations from the diagonal line indicate errors. Table 1 (left column) shows the estimated pose angle average errors for the occlusion images. In the case of the original image, has some but peak limited in which the estimation errors valuess which occur mainly in the folded portion of the trajectory, the average error is relatively large. The blurred edge image has the lowest pose angle average estimation error, which is less than those of the edge, blurred original, and original images by 0.411°, 3.653°, and 4.09°, respectively.

6.2 Position Shifts

The object was shifted by 4.5% in an upward direction and the estimation errors were examined. The profiles of the pose angle estimations are shown in Fig. 7(a)~(d).

Estimations made using the original images show a large number of errors, which are reduced when the original images are blurred. On the other hand, even though the use of edge images gives large peak error values, the average of error is smaller than that of the original images. With blurred edge images, the errors are minimized.

At least ten expanding dimensions are considered to be required to obtain a contribution tatio of 0.7, which is needed for image spotting. Image spotting means recognizing the image of an object in a scene, which is one of the attractive applications of the pose estimation. The cumulative contribution ratios shown in Fig. 4 suggest that blurring may be effective for images spotting. In our experiment, thirteen dimensions were needed to obtain a contribution ratio of 0.7 using the original image (case (a) in Fig. 7), compared with only three dimensions using the proposed method (case (d) in Fig. 7).

Our findings show that blurring is effective for the estimation of position shift. The mean absolute errors of the pose angle estimations for all the image types are given in Table 1. The estimation error for the blurred edge image decreased by 1.577°, 3.827°, and 0.987° respectively compared to those for the edge, original, and blurred original images. From these results, it is clear that proposed blurred edge image method is robust enough for the pose angle estimation of a shifted object.

6.3 Illumination Variation

In this section, we examine how the pose estimation distance error varies for the following cases (see Fig. 8(1)~(4)):

(i) The position of light source changes right or left direction.

(ii) The intensities of right and left light source change.

The conditions of illumination (light sources) are shown in Table 2.

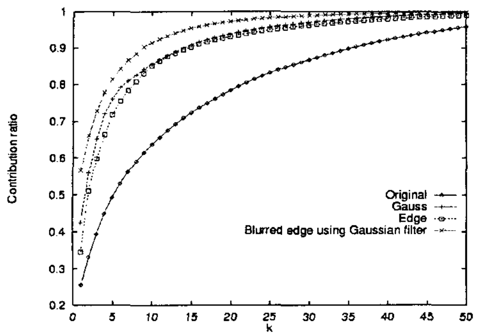
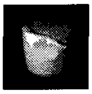


Fig. 4. Contribution ratios for four kinds of images.

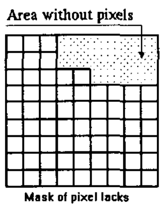
(a) Original image



(b) Blurred original image

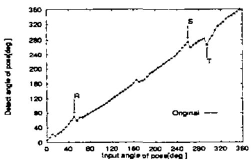
(c) Edge image

(d) Blurred edge image

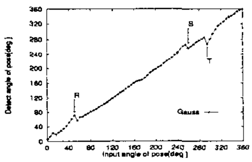


(e) Occlusion window

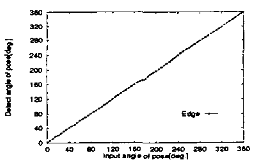
Fig. 5. Sample images from Fig.1 with the top-right portions occluded.



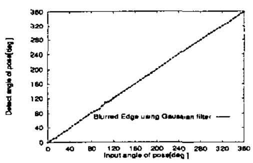
(a) Original



(b) Blurred original

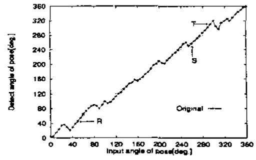


(c)Edge

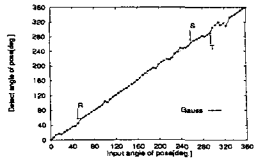


(d) Blurred edge

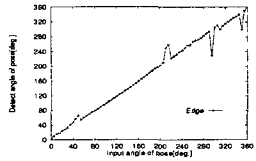
Fig. 6. Pose estimation profiles for occluded images in Fig. 5.



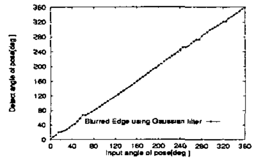
(a) Original



(b) Blurred original



(c) Edge



(d) Blurred edge

Fig. 7. Pose estimation profiles for a position shift.

(1) (2) (3) (4)

(a) Original images

(1) (2) (3) (4)

(b) Blurred edge images

Fig. 8. Images used for pose estimation in the case of light source variation

Table 1. Estimation accuracy for images containing occluded or shifted objects

Input images	Average error of pose estimation(degree)	
	Occlusion image	Position shift image
original image	4.770	5.715
Blurred original image	4.333	2.875
Edge image	1.091	3.465
Blurred edge image	0.680	1.888

Table 2. Illumination condition for Fig. 8.

case No.	Illumination	Average value of pot-area gray level
(1)	①	92.3
(2)	②	93.9
(3)	①+③	122.8
(4)	②+③	137.8

light source is located
① 25° right hand side
② 35° left hand side
③ front side from the object-toward camera line

The setting of the light sources:

- distance between the light source and the object is 40cm.
- elevation angle is 30° from the object.

In each case, the number of parametric eigen-subspace dimensions is three. Fig. 8 shows the original and blurred edge images used for the experiment in which the illuminations was changed.

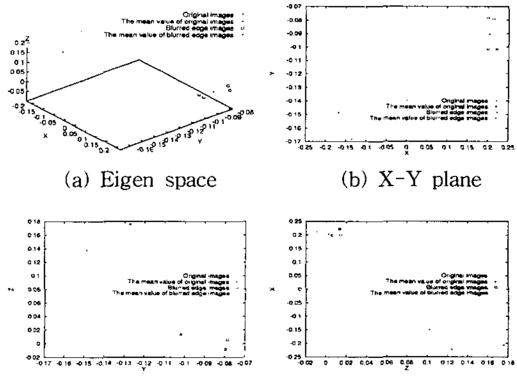
Fig. 9(a) shows the error profiles of the 3D eigensubspace for the images with illumination variation. Figures 9(b), (c), and (d) show the projected profiles of the dispersions for the illumination changes (original images, ◇; blurred edge images, □) onto the x-y, y-z, and z-x planes, respectively. The group at the left side is that of the original images and the group at the right side is that of the blurred edge images. Two groups are well separated. The mean projected points for the four case of original and blurred images are indicated by "+" and "x", respectively. It can be seen that the blurred edge image group is less dispersed (i.e, more stable) than the original image group. The distances from the x=0 plane along the x axis, which correspond to the first primary compoments, are at the same level for both images (Fig. 9(b)). The distances of the blurred edge images (Fig. 9(c), (d)) from the y=0 plane along the y axis and the z=0 plane along the z axis, which respectively correspond to the second and third primary components, are smaller than those of the original images. These results show that the processing of differentiation (edge detection) and blurring in the parametric eigenspace method produces a compact eigensubspace and is robust for changes in light intensity.

Table 3 shows the average absolute errors of pose angle estimation for illumination variations, from which we can see that the blurred edge image method is robust also for illumination variations.

6.4 Comparison with Other Methods

For the case of occlusion that frequently occurs in the pattern recognition, we conducted a comparison study using a correlation method and Eigenspace method[7].

In the comparison experiment, we used 36



(a) Eigen space (b) X-Y plane

(c) Y-Z plane (d) Z-X plane

Fig. 9. Projected points of original images (left-side group in (a)) and blurred edge images (right-side group) under varying illumination.

Table 3. Estimation accuracy for original and blurred edge images with variation of illumination.

Input images of Fig. 8	Average error of pose estimation(degree)	
	Original image	Blurred edge image
(1)	1.27	0.004
(2)	2.5	1.23
(3)	3.17	0.76
(4)	4.25	0.26

occlusion images that were produced by Fig. 5 and rotated by 10 degrees. There are the results of the comparison experiment between our proposed method and other existing methods, on the Table 4.

The method we proposed could gain relatively high accuracy, compared to the conventional methods, and we could confirm that it is possible to estimate the pose of objects.

7. CONCLUSIONS

The parametric eigenspace method proposed by Murase and coworkers allows the pose of a 3D object to be estimated by 2D object matching. In this case, the matching process between the huge amount of data can be simplified by K-L transformation because the images of a rotating object have a strong correlation, and the objcet can be

Table 4. The results of the comparsion experiment with other methods

	recognition ratio (in the case of occlusion)
correlation method	38.46%
Eigenspace method	40.28%
the proposed method	99.47%

expressed by a small number of eigenvectors. We confirmed that the redundancy of edge image information is low compared to that of the original image, on the basis of which we proposed a new method for pose estimation in parametric eigen-space using edge images blurred by Gaussian filtering instead of the original images. We have shown that blurring by the Gaussian function and edge extraction enhance the data compression ratio and decrease the occurrence of errors. We have also shown how blurring and edge extraction of an object are effective in decreasing the pose estimation error, and consequently the effectiveness of the blurred edge image.

The conclusive results are as follows:

(1) Blurring and edge extraction in K-L transformation enhance the compression ratio.

(2) Blurring and edge extraction are both robust with respect to occlusion and pose shift.

(3) A blurred edge image is the most accurate and robust type of image for pose estimation under circumstances of occlusion, pose shift, and/or illumination variation.

The proposed method was demonstrated to possess both good accuracy and robustness compared with conventional methods that use original images.

8. REFERENCES

[1] K. Ikequchi, S. Nagata, E. Horn, and K. Nishihara, "Determining gripper configuration in bin picking tasks using photometric stereo system and PRISM stereo system"(in Japanese), IEICE Trans. D J68-D, pp. 546-553, 1985.

[2] T. Noguchi, and K. Deguchi, "Visual servoing without feature correspondences using eigen-space method"(in Japanese), Trans. Society of Instrument and Control Engineers 32, pp. 1439-1446, 1996.

[3] J. Qiang, "3D Face pose estimation and tracking from a monocular camera," Image and Vision Computing, 20, pp. 499-511, 2002.

[4] E. J. Gonzalez-Galvan, S. R. Cruz-Ramirez, et. al, "An efficient multi-camera, multi-target scheme for the three-dimensional control of robots using uncalibrated vision," Robotics and Computer-Integrated Manufacturing, 19, pp. 387-400, 2003.

[5] E. Oja, "Subspace methods of Pattern Recognition." Hertfordshire: Research Studies, 1983.

[6] A. Leonardis, H. Bischof, and J. Maver, "Multiple eigenspaces," Pattern Recognition, 35, pp. 2613-2627, 2002.

[7] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 586-591, 1991.

[8] Y. Sakurai, H. Iwase, H. Takemura, N. Yokoya, and T. Kato, "Face recognition by eigen-space method using cylindrical range data." (in japanese), IEICE Technical Report PRU95-193, pp. 23-28, 1996.

[9] Yongping Li, "Linear Discriminant Analysis and its application to Face Identification", Ph. D. Thesis Paper, University of Surrey, September 2000.

[10] S. Chen, J. Liu, and Z. H. Zhou, "Making FLDA applicable to face recognition with one sample per person," Pattern Recognition, 37, pp. 1553-1555, 2004.

[11] A. Lanitis, T. F. Cootes, and C. J. Taylor, "Toward automatic simulation of aging effects on face images," IEEE Trans on PAMI, 24,

pp. 0442-0455, 2002.

[12] A. Nagata, T. Okazaki, S. C. Ching, and H. Harashima, "Basis generation and description of facial images using principal-compontent analysis." (in Japanese), IEICE Trans. D-II J79-D-II, pp. 1230-1235, 1996.

[13] H. Murase and S. K. Nayer, "3D object recongnition from appearance-parametric eigenspace method-." (in Japanese), IEICE Trans. D-II J77-D-II, pp. 2179-2187, 1994.

[14] H. Murase and S. K. Nayer, "Image spotting of 3D object using multi-resolution and eigen-space representation," Trans. Information Processing Society of Japan 36, pp. 2234-2243, 1995.

[15] H. Murase and S. K. Nayar, "Illumination planning for object recognition using parametric eigenspaces," IEEE Trans. on Pattern Analysis and Machine Intelligence 16, pp. 1219-1227, 1994.

[16] R. C. Gonzalez, and R. E. Woods, "Digital image processing," Addison Wesley, 1993.

[17] S. Tomita, S. Nuchi, and J. Oizumi, "Theory of feature extraction for patterns by the Karhunen-Loeve orthogonal system." (in Japanese), Trans. Institute of Electronics and Communication Engineers 53-C, pp. 897-903, 1970.

Jin-Woo Kim

He received the B.S degree in Electrical Engineering from Myongji University in 1992 and the M.S. and Ph.D. degrees in Electronic Engineering and System desigen Engineering from Fukui University, Fukui, Japan, in 1996 and 1999, respectively. From 1998 to 1999, he was a researcher at Fukui University, Fukui, Japan. From 2000 to 2003, he was a contract Professor in the Department of Information communication and Computer Engineering at Hanbat national University, Daejeon, Korea. Since 2003 he has been with the Department of Multimedia Engineering at Kyungsung University, Busan, Korea, where he is currently a full-time lecturer. His research interests include image processing, pattern recognition, and medical imaging technology.