コンピュータビジョンと
イメージメディア 102－5
(1996. 11. 21)

エッジ画像のパラメトリック固有空間を用いた物体の姿勢推定

金 珍祐† 岡崎耕三† 佐藤嘉伸‡ 田村進一‡

† 福井大学工学部
910 福井市文京3－9－1
‡ 大阪大学医学部
565 吹田市山田丘2－2

本稿では,2 次元照合による 3 次元物体の姿勢推定を行なうために, 多方向から撮影した画像のエッジをガウス関数でボカしたエッジ画像を用いる. これにより各画像間の相関値は増大し,$K-L$ 変換を用いると効率よくデータ圧縮される. $K-L$ 変換によりこれらの画像は固有ベクトルから構成される部分空間 (固有空間) 上の多様体に投影され, 寄与率の高い, 少ない個数のパラメトリック固有値で表示可能になる. それ故オクルージョンや位置ずれ, 拡大・縮小, 照明の変化を受けた場合でも入力画像をこの固有空間に投影し, その点に最も近い多様体上での点の位置と付随するパラメータを検出することにより, その物体の姿勢が推定される. エッジ抽出等の前処理を行なわず, 原画像をそのまま用いる法との比較を通して, 提案手法の有効性, ロバスト性を示す.
和文キーワード：姿勢推定, パラメトリック固有空間,K-L 変換, エッジ画像.

Pose estimation of an image using parametric eigen spaces of edge images

Jin Woo KIM† Kozo OKAZAKI† Yoshinobu SATO‡ Shinichi TAMURA‡

† Faculty of Engineering,Fukui Univerisity
Bunkyo 3-9-1,Fukui,910
‡ Osaka University Medical School
2-2,Yamadaoka,Suita,Osaka,565

Abstract

We propose a method to estimate pose of a three-dimensional object from it is two dimensional image based on a parametric eigen space method .In this method we use a Gaussian blurred edge image as the input image instead of the original image itself. There we compress the in put images using K-L transformation,and show better pose estimation can be achieved than the method which use the original images as it is inputs. The method has better efficiency in the case of with occlusion ,size and intensity changes of the objects,etc.
keywords:pose estimation,parametric eigen space ,K-L transform,edge image.

1. Introduction

Recognition of a 3 dimensional(3D) object from its 2 dimensional(2D) images, and the estimation of the spacial pose is an important research field having a lot of applications in different areas such as binpicking[1],part manufacturing , and camera control[2]. One of the methods to find an specified object in an image is template matching. In this method,first a template which has the same shape with the specific object is made. Then by moving the template over the image, a place with the highest correlation with the template is selected as the position of the specific object. Although this method is not efficient in some cases such that have a large noise in the images,occlusion,illumination variations,and pose changes. One way to cope with this problem is to memorize all of the 2-D appearance images of the 3-D object and then compare the input image with them.This method practically isn't possible as it needs a huge volume of computer memory. So far many algorithms have been proposed for pose estimation using eigen space which has advantage of information compression between similar objects. References [3] and [4] have proposed algorithems for pose estimation of face images. An individual person's discrimination by eigen space has also been proposed [5]. Here the input images are face images which are taken from different directions. Murase et al[6][7] proposed a method of pose estimation as well as object classification from the nearest trajectory in an eigen space composed by various kinds of object and lighting conditions. However,since they use the original image itself as the input image. Therefore,when there is an occlusion in the object,the model and the object don't coincide well,and there is a possibility of becoming not able to classify and estimate parameter (pose in this case). In this paper,we propose a method of composing the eigen space from Gaussian blurred edge images,and demonstrate that this method is robust for occlusion.

2. Making image sequence

We made 72 image frames for experiments as follows; Putting the object in a turn table and rotating it each time by 5 degrees,we made the image sequence. The images were taken by a CCD camera and were saved in an Excel frame memory connected to a computer. The light source was fixed at one place. Fig.1(a) shows 8 frames of original images. Using Sobel filter the edge images were made and then blurred by a Gaussian filter (σ = 2.0). We name the resulted images **as edge** images for short. Both the original and edge images were normalized as shown in the following section. The size of the edge images are to 40×40 pixels.

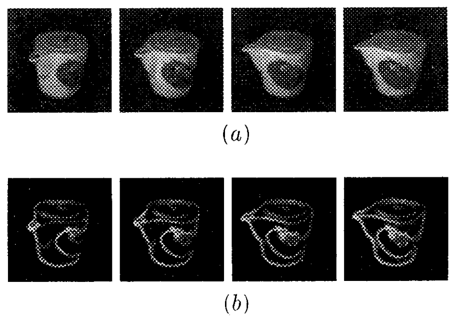


Fig.1 Examples of original(a) and edge image(b) sets obtained by rotating an object.

3. Eigen vector calculation and contribution rate

We make eigen vectors $\hat{\mathbf{x}}_m(N^2 \times 1)$ of the m - th input images by raster scanning of edge images($\mathbf{F}_m : N \times N$).

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

To remove the effects of intensity differences of input images(due to object position and sensor condition),we romalize the vectors.

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

Fig.1(b) shows the edge images. As these images are the blurred results of edge enhanced images,the correlation between two consecutive images in this sequence is high[8].We calculate the average A of images using equation(2);

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

where M is the number of images. Then the image covariance matrix $\mathbf{C}$ is made as below.

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

The C matrix can be expressed as below.

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

The eigen value is calculated using equation(4).

$$\mathbf{C}e_i = \lambda_i e_i \tag{4}$$

We arrange the eigen values in descending order and obtain the corresponding eigen vectors $(\mathbf{e}_1, \cdots, \mathbf{e}_k)$. These vectors contain the feature of edge images. Using these vectors as the basic vectors,it is possible to compose the eigen space. The trajectory of high correlated images in the eigen space will be compact and smooth. Fig.2 shows examples of the eigen vector (image). Contribution ratio Con of the principal components is expressed by

$$Con = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{K} \lambda_i} \tag{5}$$

where K and k indicate the number of dimensions of eigen space and that of subspace,respectively. The contribution ratio v.s number of dimensions for both images are shown in Fig.3. From these figures we can clearly see the higher compressibility of the proposed method (especially for lower dimensions)than the method using original images as its input.

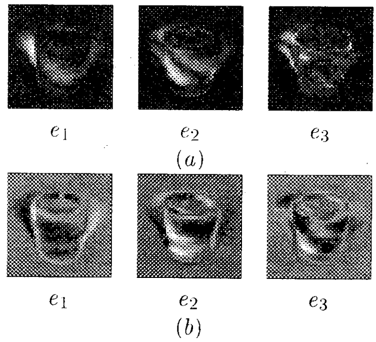


Fig.2 Eigenvectors for the object shown in Fig.

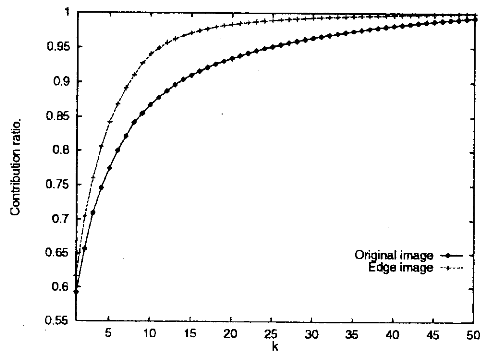


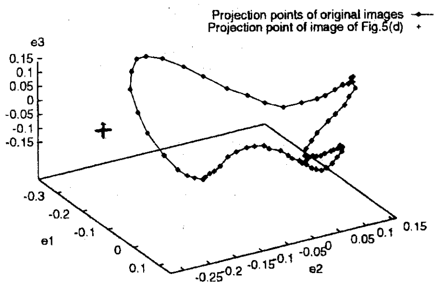
Fig.3 Cumulative contribution ratios for original and edge images.

4. Projection to parametric eigen space

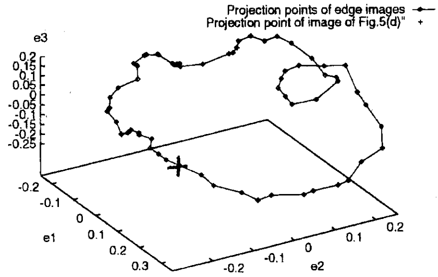
Subtracting the average image $\mathbf{A}$ from each image,the equation below gives the corresponding projection points $\mathbf{g}_{(p)}$ where p is a parameter describing the pose.

$$\mathbf{g}_{(p)} = [\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_k]^T(\mathbf{x}_m - \mathbf{A}) \tag{6}$$

Generally when the difference between two poses of an object is small, the images are highly correlated and their projected positions in the eigen space are also close to each other. Fig.4 shows the parametric eigenspace trajectory of the object of Fig.1 using only 3 eigen vectors which have the highest contribution rate. The projected point



(a)



(b)

Fig.4 Parametric eigenspace trajectory of the object shown in Fig.1.(a) original input image, (b) edge image.

sequence are expressed by a continuously smooth curve using interpolation. The interpolation is carried out by B-spline. An unknown pose of an object is estimated as follows(see detail in**5**); project the image to the eigen space and estimate its pose as the nearest point on the trajectory.

5. Pose estimation of object

First,we normalize the size and brightness of the input image. The vector of a normalized input image is defined asy and is projected as the point **z** in the sub eigenspace by equation(7).

$$\mathbf{z} = [\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_k]^T(\mathbf{y} - A) \tag{7}$$

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

$$\hat{p}_\theta = arg\min_p \|\mathbf{z} - \mathbf{g}_{(p)}\| \tag{8}$$

$$\varepsilon_e = p_\theta - \hat{p}_\theta \tag{9}$$

$$\varepsilon_d = \frac{\hat{d}}{\|\mathbf{z}\|} \tag{10}$$

$$\hat{d} = \min_p \|\mathbf{z} - \mathbf{g}_{(p)}\| \tag{11}$$

here the p_θ is a correct pose value.

6. Experiment of pose estimation

We project an original input image and its edge image on the subeigen space. Then we examine how the minimum distance between the projected points and the trajectories of the images vary according to occlusion, position shift,expansion,reduction and variation of the illumination and compare the results. Murase et al proposed an efficient algorithms of calculating the eigen values with such bulky data[9][10]. We transform the data to by Hessenberg matrix by Householder algorithm and calculate the eigen values. The computer we used is workstation Sakura of Data Processing Center,Kyoto Univ.

6.1 Occlusion

Fig.5 shows a frame of the original image sequence and its edge one, where some different parts of the object are occluded. The profile of projection for pose estimation onto the eigen subspace are shown in Fig.4(a), (b),respectively. The distance from the true point to the projected point for edge image is shorter than that of original input image. The relative distance error v.s the number of dimensions for the case of small occlusion (20 %, as shown in Fig.5(d)) is shown in Fig.6. According to this figure in the number of dimensions three, the distance errors are approximately 5 % and 0.2 % using original image and edge image as input, respectively. Changing the number of dimension,this error is almost constant in the case of our method,while its variation

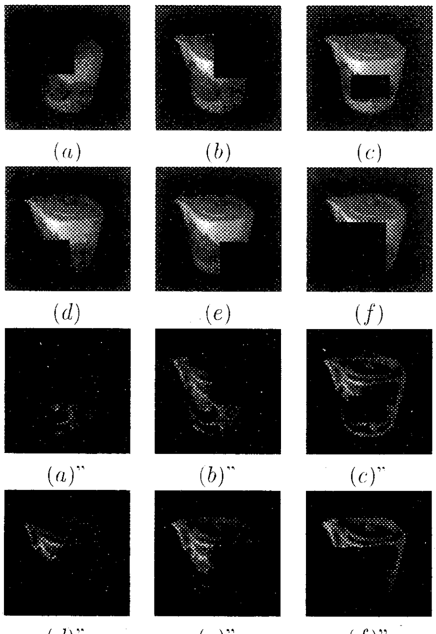


Fig.5 Examples of pose estimation experiments in case of with occlusion for original(without mark) and edge images(with mark["]).

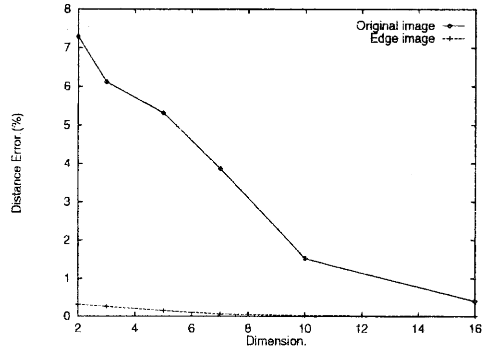


Fig.6 Distance errors v.s. dimensions for the case of small occlusion((d):20 %).

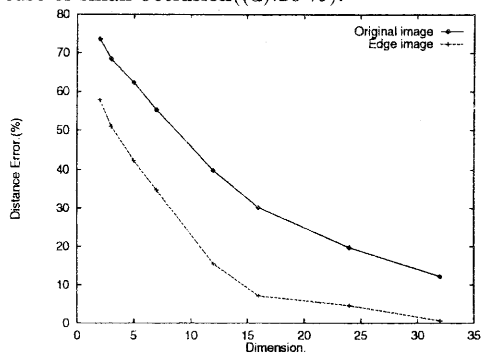


Fig.7 Distance errors of estimation v.s. dimensions for the case of large occlusion((f):60 %).

is large in the case of the other method using the original input image. The results are also almost similar to them for the (a), (b), (c), and (e) cases. Fig.7 shows the case (f) with large occlusion (60 %). The results of all cases are summarized in Table 1,which shows the higher efficiency of our proposed method. The pose estimation by edge image is less than that of original image.

6.2 Position shift,expantion,reduction and variation of illumination

(1)Position shift

The position of the object was shifted to right,left,up and down direction to examine those effects.The relation between the relative distance error ε_d and the number of dimensions for 30 % shift to left side is shown in Fig.8. The relative distance error of original and edge image are proximately 23 % and 2 % at the dimension number 2,respectincly. Those of edge image are definitly improved as the whole.

(2)Expantion and reduction

We examine effects of expantion and reduction of images by both 20 %. The results of the reduction are shown in Fig.9. The relative distance error of the original and edge image are approximately 9 % and 3 %,respectively,at the point of dimension number 2. The errors of edge image are improved even when increasing the number of dimension. The results of the expantion are shown in Fig.10. The relative distance errors of the original and edge image are 25 % and 4 % respectively at the point of dimension number 2. The error of the edge image are improved even when increasing the number of dimensions. The errors for expantion and reduction in edge image are small compared to those of original image on the whole.

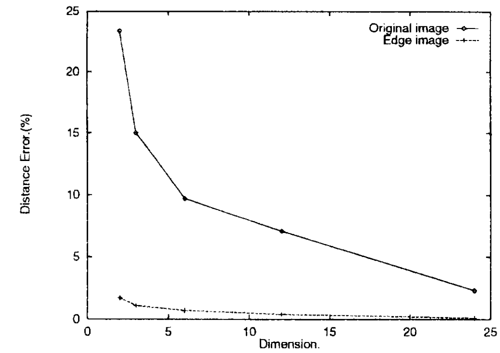


Fig.8 Distance errors of estimation v.s. dimensions for the case of position shift(30 %. in left direction).

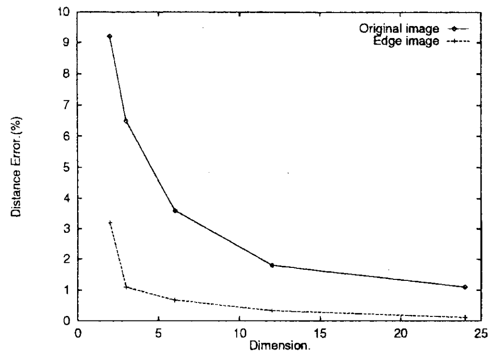


Fig.9 Distance errors v.s. dimensions for the case of size reduction(20 %).

(3)variation of illumination

Experiments of changing illumination are carried out for 4 cases totaly ;two cases for one light source and two cases for two light sources where the light sources are moved to right and left sides largely. We examined how the variance of each eigen values distributes in the pose estimation. Original and edge images used for the experiment are show in Fig.11. Pose estimation errors for images of Fig.12(a) are profiled on the 3D subeigen space.Figures(b),(c) and (d) are projections onto x-y,y-z and z-x plane, respectively. The group of coordinates origin side and right side are those of original and edge images,respectively.Both of the

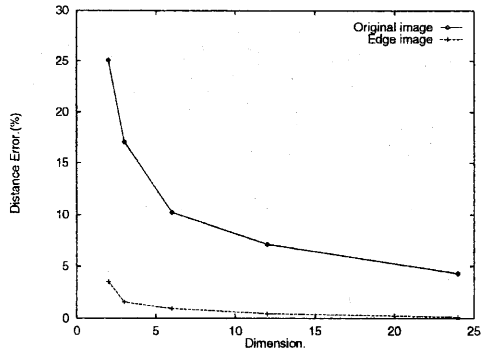


Fig.10 Distance errors v.s. dimensions for the case of expansion(20 %).

estimation errors are small compared to those of the position shift, expantion and reduction.The errors of edge images are improved 6-8 %

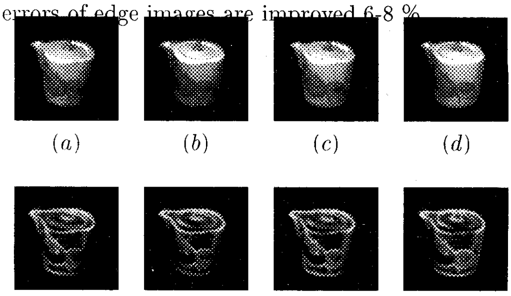


Fig.11 Distance errors of pose estimation v.s dimensions in the case of light source variation for input original(without mrk) and edge images(with mark["]).(a),(b),(a)",(b)":light source is one,(c),(d),(c)",(d)":light source is two.

(4)Distance error and pose estimation error

Calculation results of the relative distance error$\varepsilon_d$and pose estimation errorε_θ for the case of (1) $\sim$ (3) are shown in Table 2.

7. Conclusive remarks

The parametric eigen space method proposed by Murase it et estimates a pose of 3D object by 2D object matching.Matching process is simplifide by

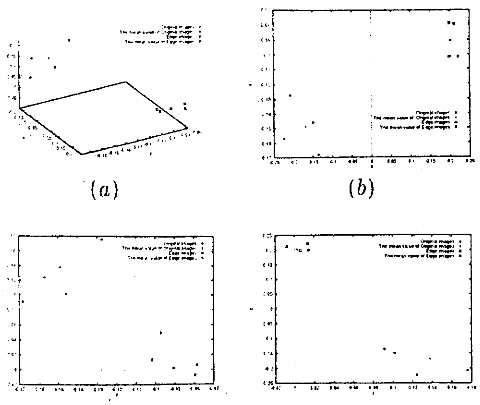


Fig.12 Estimated points of input original and edge images in 3D parametric space for the case of light source variation .(a) is 3D eigen space,(b),(c) and (d) are profile of projections on x-y,y-z and z-x plane,respectively.Left side points of each figure are projection of original image, and right side points of each figure are projection of egde image.

K-L transform because the images of rotating object have strong correlation. and the object can be expressed by small number of eigenvalues. We perceived the features that the redundancies of information of the edge image are decreased compared to that of original image. Thus,we have proposed a new method of pose estimation in parametric eigen space using the edge images instead of the original input image. We have examined that how the occlusion,expantion,reduction and changing illumination of an object to pose estimation and the number of the dimensions have affected the distance errors. We demonstrated that the proposed method is robust compared to the convention method using original image.

Table 1 Comparisons of estimation accuracy for original and edge image with occlusion. [anti-clockwise]

	The relative distance error ε_d(%)		The pose estimation error ε_θ(degree)	
	Original image	Edge image	Original image	Edge image
(a)25(%)	7.7	0.4	2.8	-0.06
(b)25(%)	7.4	0.6	2.1	0.65
(c)15(%)	5.1	0.8	-2.7	-0.81
(d)20(%)	6.1	0.2	2.4	0.07
(e)20(%)	6.2	0.5	2.6	0.13
(f)60(%)	68.6	51.4	-75.1	37.3

Table 2.

[anti-clockwise: +]

		Therelative distance ε_d(%)		The pose estimation ε_θ(degree)	
		Original image	Edge image	Original image	Edge image
Position shift (30 %)	left	9.8	1.1	-35.4	0.9
	right	10.8	1.6	38.7	-0.64
	up	10.5	1.2	-40.3	1.25
	down	6.5	1.4	32.5	-1.37
expantion(20 %)		6.3	0.6	8.7	0.07
reduction(20 %)		16.2	1.4	85.4	1.4
variation of illumination	(a)	5.7	0.2	-1.27	0.04
	(b)	7.8	2.4	2.5	1.23
	(c)	8.3	4.7	-3.17	0.76
	(d)	8.9	3.1	4.25	-0.26

References

[1] K.Ikeguchi,S.Nagata,E.Horn,and K.Nishihara., " Determining Gripper Configuration in Bin Picking Tasks Using Photometric Stereo System and PRISM Stereo System, " The Institute of Electronics and Communication Engineers, vol.J68-D,no.4,pp.546-553,April 1985.

[2] T.Noguchi,and K.Deguchi., " Visual Servoing without Feature Corresponderces Using Eigenspace Method, "Transactions of the Society of Instrument and Control Engineers, vol.32,No.10,pp.1439-1446,1996.

[3] Turk M.A. and Pentland A.P., "Face Recognition Using Eigenfaces, "Proc.of IEEE Conference on Computer Vision and Pattern Recognition,pp.586-591,June 1991.

[4] Y.Sakurai, H.Iwase, H.Takemura., N.Yokoya, and T.Kato, " Face Recognition by Eigen-space Method Using Cylindrical Range Data, "Techinal Report Of IEICE,PRU95-193,pp.23-28,1996-01.

[5] A.Nagata,T.Okazaki,S.C.Chang,and H.Harashima., " Basis Generation and Description of Facial Images Using Principal-Component Analysis, " The Institute of Electronics and Communication Engineers, vol.J79-D-II,no.7,pp.1230-1235,July 1996.

[6] H.Murase and Shree K.Nayar., " 3D Object Recognition from Appearance — Parametric Eigenspace Method —, " The Institute of Electronics and Communication Engineers, " The Institute of Electronics and Communication Engineers, vol.J77-D-II,no.11,pp.2179-2187,Dec 1994.

[7] H.Murase and Shree K.Nayar., "Image Spotting of 3D Object Using Multi-Resolution and Eigenspace Representation, " Transactions of Information Processing Society of Japan, vol.36,no.10,pp.2234-2243,Oct.1995.

[8] S.Tomita, S.Noguchi, and J.Oizumi., " Theory of Feature Extraction for Patterns by the Karhunen-Loeve Orthogonal System, " The Institute of Electronics and Communication Engineers, vol.53-C,no.12,pp.897-903,Dec.1970.

[9] Roseborough J.B. and Murase H., " Partial eigenvalue decomposition for large image sets using run-length encoding, " Pattern Recognition,vol.28,pp.421-430,1995.

[10] Murase H. and Lindenbaum M., " Partial Eigenvalue Decomposition of Large Images Using Spatial Temporarl Adaptive Method,",IEEE Trans. on Image processing,Vol.4,No.5,May 1995.