Systems and Computers in Japan, Vol. 26, No. 8, 1995
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J77-D-II, No. 11, November 1994, pp. 2179-2187

Three-Dimensional Object Recognition from Appearance—Parametric Eigenspace Method

Hiroshi Murase, *Member*

NTT Basic Research Laboratories, Atsugi, Japan 243

Shree K. Nayar, *Nonmember*

Columbia University, New York, NY10027, USA

SUMMARY

Technology of recognizing a 3-dimensional object and finding its direction from a 2-dimensional image is important in practical applications such as classification of industrial products. A typical conventional method for this purpose uses the 3-dimensional structure of the object such as its edges and surface shapes. However, extraction of a 3-dimensional structure with a high accuracy, notably that of an arbitrary shaped object, is difficult.

This paper proposes a method of recognizing a 3-dimensional object by using a 2-dimensional collation. A 2-dimensional collation which requires no 3-dimensional feature has never seriously been examined, because it has been considered that the amounts of computation and a memory for learning the 2-dimensional image data (which are very complex due to the variations of viewing and lighting angles) are not acceptable. The proposed method can learn a 3-dimensional object as a set of 2-dimensional images by using a new parametric eigenspace approach with a small memory capacity. The proposed method can easily learn a 3-dimensional object from its 2-dimensional image, and can recognize the object and estimate its pose. This paper includes experimental comparisons between the proposed method and other 2-dimensional collation methods.

Key words: Object recognition; pose estimation; principal component analysis; eigenvector; manifold; visual learning.

1. Introduction

There have been many studies on technology of recognizing a 3-dimensional object from its 2-dimensional image with estimation of its direction (pose) [1, 2], since this technology has many applications such as classification of industrial products and monitoring of objects in various environments. There have been two basic methods of collation of a 3-dimensional object and its model: a method using a 3-dimensional structure, and a method using a 2-dimensional structure. Most conventional methods have been based on the former, because it has been considered that the appearance of a 3-dimensional object varies widely depending on the directions of illuminations and viewing angles. A proposal of a framework for 3-dimensional representations by Marr [3] has accelerated the trend of the 3-dimensional approach. Most recent methods use collation between an object and its memorized model, after the 3-dimensional structure of the object is reconstructed from its 2-dimensional image using shades, edges and movements. There have been some methods which use a 2-dimensional collation, e.g., a method using the positions of refraction points and

ISSN0882-1666/95/0008-0045

terminal points [5]. However, they use the collation between feature points, and not 2-dimensional patterns.

This paper describes "a parametric eigenspace method" which recognizes a 3-dimensional object by using a 2-dimensional collation of image-signal levels. This represents a 2-dimensional image, which changes continuously due to the changes of direction and illumination of a 3-dimensional object, by using a manifold on a subspace (eigenspace) consisting of the eigenvector of the image. In the learning stage, the manifold is constructed by calculating an eigenspace from the set of images of the object. In the recognition state, the category and pose of the object are estimated by projecting the input image onto a point in the eigenspace, and by detecting a position of the point which is nearest to the point on the manifold. In this method, an object can be learned automatically by feeding an example of the object. In the recognition stage, it is possible to carry out the recognition of a 3-dimensional object in an input and the detection of the pose of the object simultaneously.

The proposed method has a closed relationship with the subspace method of pattern recognition [6, 7]. Character recognition methods using subspace [8, 9], and the face recognition method using an eigenface [13, 14] are examples of applications of the eigenvector of pixels. These examples are aimed mainly at classification of patterns, and not detection of parameters such as the pose of an object or representation of a 3-dimensional object.

The main part of the paper describes the 3-dimensional collation, the 2-dimensional collation, the parametric eigenspace method, and experimental comparisons between the proposed method and other method.

2. Three- and Two-Dimensional Collations in Object Recognition

As already described, there have been two approaches in image recognition of a 3-dimensional object: the 3-dimensional approach, and the 2-dimensional approach. Let us compare their features.

The 3-dimensional collation method has a dimensional perfectness since it is basically 3-dimensional. A model seen from any direction can be represented by the same form in the 3-dimensional treatment. The amount of 3-dimensional descriptions of a model is less than that of a 2-dimensional description. At the stage of an object recognition in this approach, the collation is carried out after the 3-dimensional features have been reconstructed from its 2-dimensional image. However, this reconstruction process has never been perfect, and the reconstruction is not stable. Similar problems occur when the model of an input model is learned. Therefore, usually a model (such as a manually prepared CAD model) is employed for the recognition procedure.

The 2-dimensional collation method recognizes, in principle a 3-dimensional object by using a 2-dimensional comparison of the input image and a set of 2-dimensional images stored. Therefore, extraction of 3-dimensional features are not required. This method has never been tested, since the method appeared to need a large amount of computation with a large memory for leaning models which contain wide variations due to illuminations and viewing angles. If these problems are solved, it would be possible to construct a versatile recognition system which can apply even to an arbitrary shaped 3-dimensional object.

The proposed method uses the 2-dimensional approach. A 3-dimensional object can be described as a set of 2-dimensional images by employing a parametric eigenspace method. Hence, the object can be learned from the examples of 2-dimensional image so that a 3-dimensional object can be recognized.

3. Learning Using Parametric Eigenspace Method

An image representing the appearance of an object varies widely depending on the position of its illumination and viewing angle. Figure 1 shows an example of a set of images of an object obtained by rotating it. The method of memorizing these images is the subject of the learning. An extraction of essential information of an image from its large set of 2-dimensional images is equivalent to a coding of the images. This is the base of representation of images in this approach, and this is called "parametric eigenspace method."

In the learning stage, a parametric eigenspace is constructed from a set of learning image samples. This process consists of two stages. In the first stage, a subspace (eigenspace) with eigenvector is constructed from the learning image samples. In the second stage, a series of learning images (which vary continuously) are projected on an eigenspace so that the series of the original images is represented by constructing the manifold. If there are more than two objects, the same number of manifolds as the objects are constructed. In

Fig. 1. An image set obtained by rotating an object.

the recognition process, the input image is projected on a point on the eigenspace, then the recognition of the object and the detection of its pose are carried out.

3.1. Normalization of images

First, the part of an object in an input is extracted. The extraction in this experiment is carried out by using a threshold or a difference in the image from the background. Then, the value of zero is substituted in the part of the image other than the object. The size of the object is normalized so that the object contacts a square without changing the ratio of the vertical to horizontal dimensions. By scanning this square image, a series of pixels is obtained. Then the original image is represented by a vector

$$\hat{x} = [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_N]^T$$

where $[\hat{x}_1, \hat{x}_2, ..., \hat{x}_N]$ is the obtained series of pixels, and N is the number of pixels.

To eliminate the influence of a variation of sensor, sensitivity, the brightness is normalized. Let the normalized image vector be x. The normalization is achieved by

$$x = \frac{\hat{x}}{\|\hat{x}\|}$$

The appearance of an object varies depending on its pose and illumination. Let us assume that an object rotates around a vertical axis, and the object is illuminated by a point light source which moves along a line

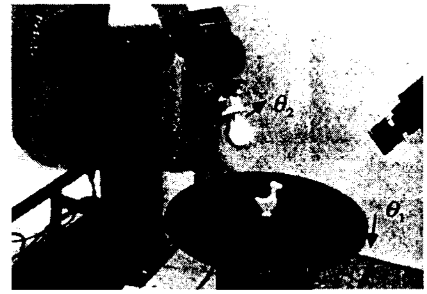


Fig. 2. Setup used for automatic acquisition of object image sets.

in addition to a general background light. This assumption is acceptable for a natural indoor scene such as the interior of a factory. To treat an object having an arbitrary pose, the parameters can be increased.

Let us consider a case where P kinds of objects are learned. Let a set of images of an object, which rotates in 360° under an illumination with varying direction, be

$$\{x_{1,1}^{(P)}, \cdots, x_{R,1}^{(P)}, x_{1,2}^{(P)}, \cdots, x_{R,L}^{(P)}\}$$

where R is the number of the increment in the rotation of the object, and L is the number of directions of the light source. Let us call this "image set of pth object." The image sets all of the objects are represented by

$$\begin{aligned} &x_{1,1}^{(1)}, \cdots, x_{R,1}^{(1)}, x_{1,2}^{(1)}, \cdots, x_{R,L}^{(1)}, \\ &x_{1,1}^{(2)}, \cdots, x_{R,1}^{(2)}, x_{1,2}^{(2)}, \cdots, x_{R,L}^{(2)}, \\ &\vdots \\ &x_{1,1}^{(P)}, \cdots, x_{R,1}^{(P)}, x_{1,2}^{(P)}, \cdots, x_{R,L}^{(P)}\} \end{aligned}$$

Let us call each image vector in this set "learning sample." These samples were obtained in this experiment by using a computer-controlled turntable and a direction-adjustable robot arm for the light source as shown in Fig. 2. By using this apparatus, the learning samples are obtained automatically.

3.2. Calculation of eigenvectors

As shown by the example of image series in Fig. 1, two adjacent images have a very high correlation between each other. The image is compressed by using this correlation. In this experiment, a Karhunen-Loève expansion, which can compress an image most efficiently from the point of view of a mean square error, is employed to compress the image set. This method

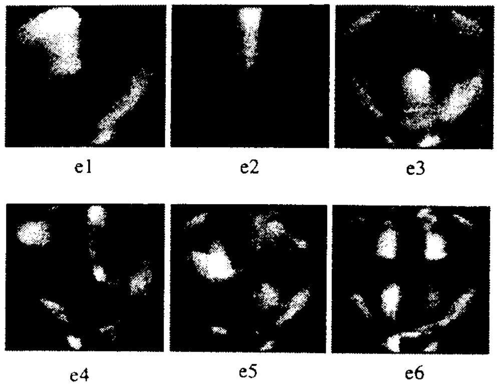


Fig. 3. Eigenvectors for the object shown in Fig. 1.

approximates the original image by using a subspace (eigenspace) spanned by the eigenvectors of the convariance matrix of the image set. Let us calculate the universal eigenspace (which is the eigenspace of all objects) and the eigenspace of object p.

An image representing all the learning images (the mean image) is given by

$$c = \frac{1}{RLP} \sum_{p=1}^{P} \sum_{r=1}^{R} \sum_{l=1}^{L} x_{r,l}^{(p)}$$

By subtracting the mean image from each image, a matrix

$$X \equiv [x_{1,1}^{(1)} - c, \cdots, x_{R,1}^{(1)} - c, \cdots, x_{R,L}^{(P)} - c]$$

is made. The convariance matrix of the image set is given by

$$Q \equiv XX^T$$

The eigenspace (e.g., k-dimensional) is given by solving

$$\lambda_i e_i = Q e_i$$

and by taking the eigenvectors (e_1 ... e_k) which correspond to k the largest eigenvalues ($\lambda_1 \geq \cdots \geq \lambda_k \geq \lambda_N$) as the base vector. Generally, it is difficult to calculate eigenvectors having a large dimension such as the convariance of an image. Note that this experiment treats 16,384 dimensions. If the number of pixels is small, it is possible to calculate the eigenvector by using, for example, the singular-value decomposition or the STA method [10]. The universal eigenspace is a space suitable of representing the set of all the objects, and is useful for identifying objects.

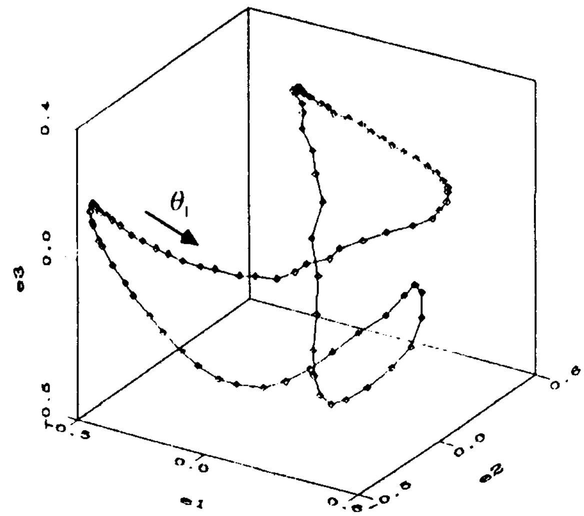


Fig. 4. Parametric eigenspace representation of the object shown in Fig. 1.

The eigenspace of object p is a space which is calculated using the image set of object p alone. Letting the convariance matrix be $Q^{(p)}$, the eigenspace of the pth object is given by solving

$$\lambda_i^{(p)} e_i^{(p)} = Q^{(p)} e_i^{(p)}$$

and by taking the eigenvector as the base. Since the eigenspace of object p is a space suitable for representing the object, the eigenspace used for estimating the pose of the object after its name, is identified. Figure 3 shows examples of the eigenvector processed from the images shown in Fig. 1.

3.3. Parametric eigenspace representation of external aspect of objects

Let us represent a 3-dimensional object, which varies continuously due to its pose and/or the position of a light source, on a manifold in a eigenspace. When a vector representing the remainder of a learning sample minus the mean image is projected onto an eigenspace, by using

$$g_{r,l}^{(p)} = [e_1, e_2, \cdots, e_k]^T (x_{r,l}^{(p)} - c)$$

a single image corresponds to a single point. Therefore, the projection of learning samples corresponding to a one turn (360°) form a series of 2-dimensional points. Because, when the pose of an object varies little, its images also change little, the correlation between each image is high, and image having a high correlation are projected closely on an eigenspace.

Figure 4 shows an example of a parametric representation of an object (shown in Fig. 1). This series of points is illustrated in a 3-dimensional space, although they are in fact a series of points in a multiple-dimensional space. This series of points is represented as a continuous line by interpolating between the points. A cubic spline [17] was used for interpolation in this experiment. Images of an object under a varying position of a light source can be treated in a similar way, i.e., a 2-dimensional manifold represented by the pose of an object and the position of a light source is constructed on an eigenspace. Let this manifold be $g^{(p)}(\theta_1, \theta_2)$, where θ_1 and θ_2 correspond to the parameters of the rotation and the position of the light source, respectively. This manifold contains the pose of the object and the position of the light source which exist in the learning sample, because these are interpolated. The same number of manifolds as the number categories of objects are constructed in the universal eigenspace.

The manifold of object p also is constructed in eigenspace of object p. The learning samples of object are projected on the eigenspace of object p by using

$$f_{r,l}^{(p)} = [e_1^{(p)}, e_2^{(p)}, \cdots, e_k^{(p)}]^T (x_{r,l}^{(p)} - c^{(p)})$$

so that manifolds are constructed with interpolation, where $c^{(p)}$ is the mean of learning sample of object p. The interpolation is represented by $f^{(p)}(\theta_1, \theta_2)$.

4. Recognition

First, the same process as in the learning stage is applied to the recognition process, i.e., the region of an object is extracted from the original image using subtraction. The size and brightness of the extracted region are normalized. Let the vector of the normalized image be y. This vector is projected to a point z universal eigenspace by using

$$z = [e_1, e_2, \cdots, e_k]^T (y - c)$$

where c is the aforementioned mean image. The recognition of an object is carried out by investigating the position of the projected point z on the manifolds (P categories). The distance between point z and the manifold $g^{(p)}(\theta_1, \theta_2)$ is given by

$$d_1^{(p)} = \min_{\theta_1, \theta_2} \| z - g^{(p)}(\theta_1, \theta_2) \|$$

The recognition of the object is achieved by finding p which minimizes this distance. In this experiment, p is

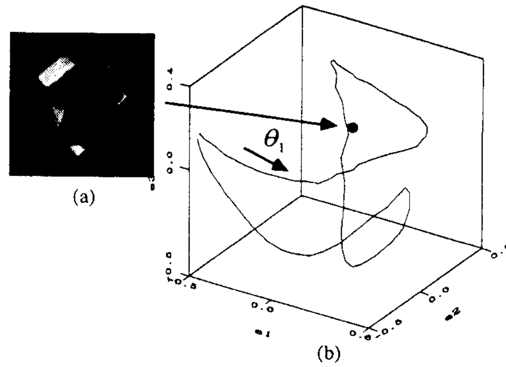


Fig. 5. (a) An input image; (b) the input image is mapped to a point in the eigenspace. The location of the object point on the curve determines the pose of the object.

obtained by searching a table which contains all the representative points on the manifold, since this process increases the processing speed.

After the name of object p is recognized, its pose is estimated by using the following procedure: the input image y is projected on the eigenspace of the object by using

$$z^{(p)} = [e_1^{(p)}, e_2^{(p)}, \cdots, e_k^{(p)}]^T (y - c^{(p)})$$

The detection of the pose of an object is equivalent to the detection of position of point $z^{(p)}$ on the manifold. For this purpose, θ_1 which minimizes

$$d_2^{(p)} = \min_{\theta_1, \theta_2} \| z^{(p)} - f^{(p)}(\theta_1, \theta_2) \|$$

is found. Figure 5 shows the relationship between the original image, a point in the eigenspace to which the point is projected, and θ which has the minimum distance. This example used a 1-dimensional manifold (a curved line), al-though an actual search is carried out in a 2-dimen-sional manifold (a curved surface).

5. Recognition Experiments

Four objects as shown in Fig. 6(a) were used for the experiments. A set of learning samples of each of the objects was made by using a computer-controlled turntable (Fig. 1). Each set has 90 poses taken at 4° intervals (360° in all). The light source is changed at 5 positions with intervals of 30° around the turntable. Therefore 4 × 90 × 5 - 1800 learning samples were

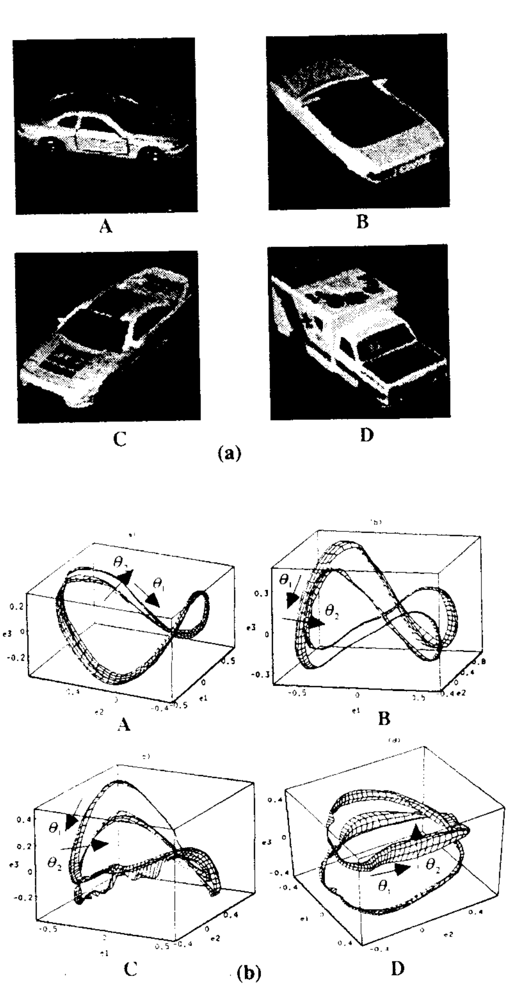


Fig. 6. (a) The objects in the experiments; (b) the parametric eigenspace for the objects in (a).

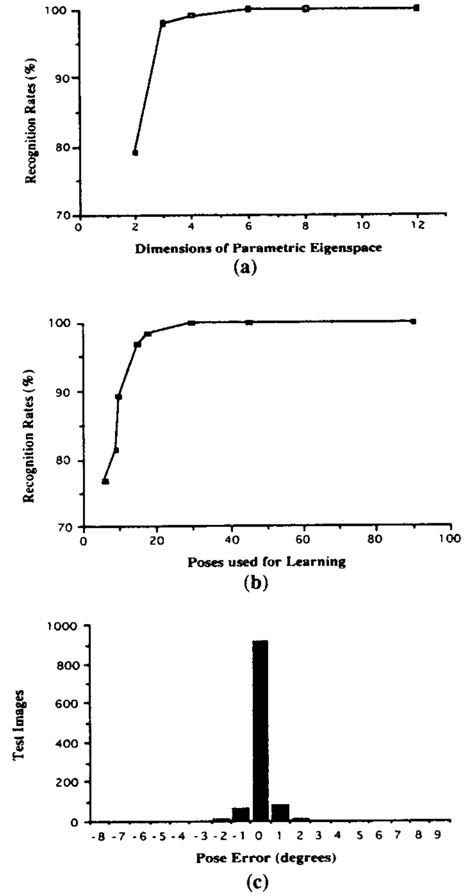


Fig. 7. (a) Recognition rate plotted as a function of the number of eigenspace dimensions; (b) recognition rate plotted as a function of the number of poses of each learning object; (c) histogram of the errors in the pose estimation

made for each set. Each image is normalized in 128 × 128 pixels, and stored in a real number of 4 bytes (118 Mbytes in all). Data for the recognition experiments were made by sampling 90 poses (at positions with a phase difference of 2° referring to the learning samples) and three positions of the light source (1800 images in all). In the learning stage, the process described in section 3 was applied, and the parametric eigenspace is calculated for each object. Figure 6(b) shows examples. If a single manifold is in an 8-dimensional eigenspace, the amount of the data is about 640 Kbytes for storing the eigenvectors, and 115 Kbytes for a table for the manifold.

The recognition experiments were carried out by applying the process described in section 4. Figure 7(a) shows the relationship between the recognition rate and the dimension of the parametric eigenspace. This shows that the recognition rate saturates after 8 dimensions. It is possible in some cases to recognize an object which does not exist in the learning sample by using this method. Therefore, it is worth investigating the number of poses which is sufficient for the recognition. Figure 7(b) shows the recognition rate with a reduced number of poses. This shows that about 15 poses are sufficient to recognize an object as complex as these examples. Figure 7(c) shows the mean value. Another experiment has been carried out to find the effect of the error on the recognition accuracy when an object is extracted incorrectly from its background. The results show that this kind of small extraction error (less than 3 percent of the size of an object) has no significant effect on the recognition accuracy (see [18]).

6. Discussions

6.1. Comparison with other methods

Let us compare the proposed method, which is basically a 2-dimensional collation of a 3-dimensional object, with other similar methods. It is interesting to determine the effectiveness of conventional 2-dimensional image recognition methods when they are used for recognizing 3-dimensional objects. Three typical methods are selected for the comparison: (1) simple correlation method [14]; (2) the eigenface method used by Turk et al. [14]; and (3) the projection method for letter recognition used by Murase et al. [9]. The outline of each method is as follows.

(1) Simple correlation method

In the learning stage, the set of mean image vectors, $c_1^{(p)}$ for the learning is calculated for each object In the recognition stage, the correlation function $a_3^{(p)} = y^T c_1^{(p)}$ between the input image y and the image vector is calculated, and a p which makes the maximum correlation is regarded as the final result.

(2) Eigenface method

In the learning stage, an eigenspace is made by calculating the eigenvector of the convariance of a set of images, assuming an 8-dimensional space. The mean value $c_2^{(p)}$ is calculated for each object p by projecting the learning sample onto the space. In the recognition stage, the input image is projected onto the eigenspace, and the distance $d_4^{(p)} = |z - c_2^{(p)}|$ between the point and the mean value also is calculated. An object p which minimizes the distance is regarded as the final result.

(3) Subspace method

In the learning stage, the eigenvector $[e_1^{(p)}, e_2^{(p)}, \cdots, e_k^{(p)}]$ of the auto-correlation matrix of the set of images is calculated for each object p, and the eigenspace of each object p is made. In the recognition stage, the input image y is projected onto the eigenspace (assuming 8-dimensional) of each object. An object p which makes the projection energy

$$d_5^{(p)} = \sum_{k=1}^{K} (y^T e_k^{(p)})^2$$

maximum is regarded as the final result.

Table 1. Experimental comparison of methods

Method	1	2	3	Proposed
Recognition rate	68.6%	66.8	98.7	99.8
Detection of pose	Not possible	Not possible	Not possible	Possible

Eight objects, including 4 objects shown in Fig. 1, were used in the comparison experiments; 144 images containing 18 poses were used for the learning, and 720 images containing 90 poses were used for the recognition. The position of the light source is fixed throughout all the cases. Table 1 shows the recognition rate for each method. The table shows that methods (1) and (2) are effective for relatively simple 2-dimensional patterns, but they are less accurate for a complex object such as those shown in Figs. 1 and 6(a). Method (3) can recognize a 3-dimensional object with a reasonable accuracy, but cannot detect its pose. The proposed method recognizes any object with a higher accuracy than the other three methods, and can detect the pose of an object.

6.2. Application to motion images

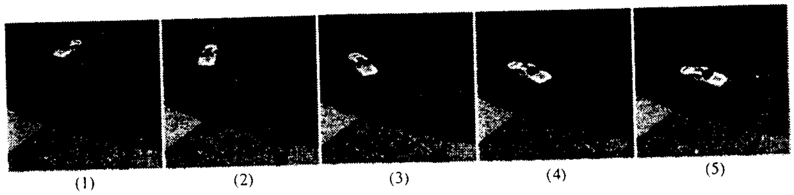
Figure 8 shows an example of application of the proposed method to a motion image. Figure 8(a) shows the sequence of the original input images containing an object. Figure 8(b) shows the object extracted from each of the input images by using the difference from its background. Figure 8(c) shows the learning samples in the closest pose. Figure 8(d) shows the pose the object computed.

6.3. Application to face images

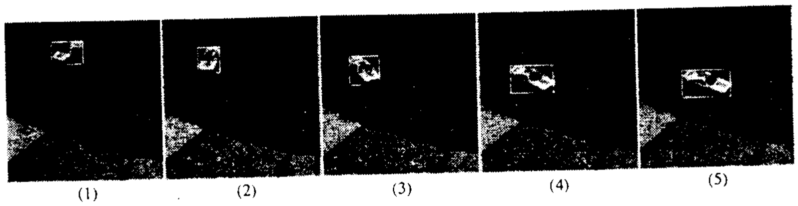
Figure 9 shows an application of the proposed method to face images. This experiment uses a set of learning samples with 18 directions. The method can recognize a face in an arbitrary direction, and can automatically detect its direction (pose).

6.4. Comparison of the proposed method with human 3-D object recognition

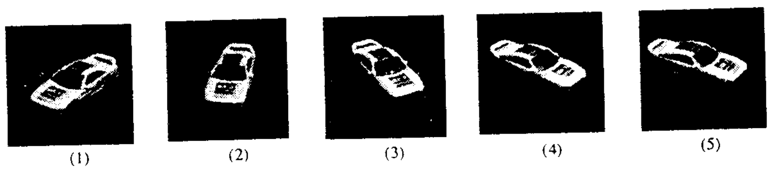
In psychology, whether the human uses a 2- or 3-dimensional collation in recognition of a 3-dimensional



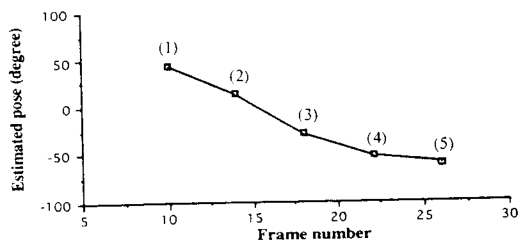
(a) Image sequence of a moving object.



(b) Segmentation of the moving object.



(c) Learning sample with closest pose.



(d) Computed pose of the moving objetct.

Fig. 8. Parametric eigenspace method applied to the image sequence of a moving car.

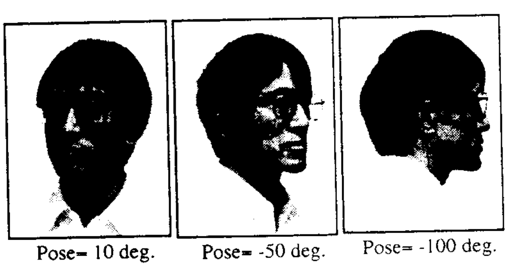


Fig. 9. Pose estimation for the faces.

object is an interesting problem. Edelman et al. have found in their psychological experiment on mental rotation [16], the following facts [17]: the human recognizes an object which is familiar to him/her in a 2-dimensional collation, although he/she recognizes an object which is not familiar to him/her in a 3-dimensional collation. This shows that the human recognizes an object which can be seen frequently in everyday life by using a simple 2-dimensional collation.

7. Conclusions

This paper describes a method which recognizes a 3-dimensional object in an arbitrary direction and detects its pose by using a 2-dimensional collation.

This paper also describes the parametric eigenspace method which represents a continuously varying series of image by using a manifold in an eigenvector space. The proposed method can memorize a 3-dimensional object as a set of 2-dimensional images with a small capacity of memory. As a result, it has become possible to learn a 3-dimensional object from its 2-dimensional image samples, to recognize the object by using a 2-dimensional collation, and to detect its pose, without extracting 3-dimensional structures such as edges or shapes which are difficult in conventional methods.

In this paper, only two parameters (a single rotation of an object and the positions of a light source) are considered. We are planning to expand our method in the future to cases where more parameters are involved.

Acknowledgment. The authors wish to express their thanks to Prof. T. Poggio (MIT) and Dr. D. Weinshall (IBM Washington Laboratory) for their important advice, to Dr. Tatsuya Kimura (Director, NTT Basic Research Laboratories), Dr. Ryokei Nakatsu (Head of Science Department, NTT BRL), and Dr. Seichiro Naito (Leader, NTT BRL) for their encouragement.

REFERENCES

1. R. T. Chin and C. R. Dyer. Model-based recognition in robot vision. ACM Computing Surveys, **18**, 1, pp. 67-108 (March 1986).
2. P. J. Besl and R. C. Jain. Three-dimensional object recognition. ACM Computing Surveys, **17**, 1, pp. 75-145 (1985).
3. D. Marr. Vision. W. H. Freeman, New York (1982).
4. S. Ullman and R. Basri. Recognition by linear combination of models. IEEE Trans. on Pattern Analysis and Machine Intelligence, **13**, 10, pp. 992-1006 (Oct. 1991).
5. T. Poggio and S. Edelman. A network that learns to recognize three-dimensional objects. Nature, **343**, pp. 263-266 (1990).
6. K. Fukunaga. Introduction to Statistical Pattern Recognition. Academic Press, London (1990).
7. E. Oja. Subspace Methods of Pattern Recognition. Research Studies Press, Hertfordshire (1983).
8. T. Iijima. Letter recognition system. ASPET/71. Journal of Institute of TV, Japan, **27**, 3, pp. 157-164 (1973).
9. H. Murase, F. Kimura, M. Yoshimura and Y. Miyake. An improvement of the auto-correlation hand-printed matrix Hiragana. Trans. I.E.I.C.E., Japan, **J64-D**, 3, pp. 276-283 (1981).
10. H. Murase and M. Lindenbaum. Spatial temporal adaptive method for partial eigenstructure decomposition of large images. NTT Technical Report No. 6527 (March 1992) (IEEE Trans. on Image Processing, to appear).
11. H. Murase and S. K. Nayar. Learning object models from appearance. AAAT-93, American Association for Artificial Intelligence, pp. 836-843 (July 1993).
12. H. Murase and S. K. Nayar. Learning and recognition of 3D objects from appearance. IEEE Qualitative Vision Workshop, pp. 39-50, New York (June 1993).
13. L. Sirovich and M. Kirby. Low dimensional procedure for the characterization of human faces. Journal of Optical Society of America, **4**, 3, pp. 519-524 (1987).
14. M. A. Turk and A. P. Pentland. Face Recognition Using Eigenfaces. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 586-591 (June 1991).
15. W. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling. Numerical Recipes in C. Cambridge University Press, Cambridge (1988).
16. M. Tarr and S. Pinker. Mental rotation and orientation-dependence in shape recognition. Cognitive Psychology, **21**, pp. 233-282 (1989).
17. S. Edelman and D. Weinshall. A self-organizing multiple-view representation of 3D objects. Biological Cybernetics, **64**, pp. 209-219 (1991).
18. H. Murase and S. K. Nayar. Illumination planning for object recognition in structured environments. Proc. of IEEE Conference on Computer Vision and Pattern Recognition (June 1994).

AUTHORS (from left to right)

Hiroshi Kimura received his B.E. and M.S. degrees in 1978 and 1980, respectively, from Electronics Department, Nagoya University, and his Dr. of Eng. degree from there in 1987. He joined NTT in 1980. He has engaged in research on character and figure recognitions, and computer vision. He has been a guest researcher at Computer Department of Columbia University (USA) for one year in 1992. Presently, he is a Distinguished Research Scientist at NTT Basic Research Laboratories. He was awarded a Young Engineer Award by IECE Japan in 1985; and the Best Paper Award of CVPR (IEEE) in 1994. He is a member of IEEE; Institute of Information Processing, Japan; and AVIRG.

Shree K. Nayar received his Ph.D. from Carnegie-Mellon University in 1990. He has been an Associate Professor at Columbia University since 1991. He was awarded the Marr Prize from ICCV Conference in 1990. He has engaged in research on computer vision, notably physics-based vision, and robot vision.