A ROBUST RECOGNITION METHOD FOR PARTIALLY OCCLUDED/DESTROYED OBJECTS

M. Masudur Rahman, Seiji Ishikawa

Department of Mechanical and Control Engineering
Kyushu Institute of Technology
Sensuicho 1-1, Tobata, Kitakyushu 804-8550, Japan
rahman@ss10.cntl.kyutech.ac.jp

ABSTRACT

This paper presents a new technique referred as a '*mean eigen window'* method for recognizing partially data-loss objects due to occlusion and/or destruction. We have proposed to store similar poses, that may include disturbed shapes, of an object in a particular window referred to as the '*eigen window'* and, finally, mean of poses of each window is taken into consideration in order to obtain a generalized eigen window called the '*mean eigen window*'. This mean eigen window is further used for recognizing an unfamiliar pose, including partially occluded or destroyed shapes, and the object type itself. We have applied the proposed approach to various data-loss environments and the method has successfully performed the recognition of an object with up to 20% of occlusion and/or destruction.

Index Terms: Object recognition, appearance-based method, eigen window, occlusion, PCA algorithm, computer vision.

1. INTRODUCTION

Object recognition is a promising area of research in a computer vision field and it has various industrial and military applications such as object picking, automatic target recognition, and surveillance and monitoring, etc.

The main difficulties for such tasks include: real-time performance, difficulty in segmentation, tracking object's poses in data-loss environments, and difficulty in obtaining appropriate models of the objects. Recently, visual learning methods based on eigenspace analysis [2-5, 7-9] have shown the potential to solve some of these problems. These methods learn object models from a series of pose images taken in the same environment as in the recognition mode. Thus, these methods overcome the difficulty related to object tracking and modeling. Furthermore, since such methods store an object model as a vector in a low dimensional feature space and recognize objects by comparison of the model and image vectors, recognition speed is very high and it can achieve real-time performance.

The eigen window method was initially proposed by K. Ohba and K. Ikeuchi [2] for stable verification of partially occluded objects, whereas the conventional eigenspace method was initially proposed by H. Murase [1]. Ohba and Ikeuchi proposed to collect various parts (particularly edges) of objects and put them into a particular window called the 'eigen window' and, then, the best matching among the object's parts is the recognized object or pose. However, partial segmentation contradicts with the concept of conventional eigenspace technique [1] and it makes complicated to collect similar parts of objects in a particular window. Besides, we cannot handle partially or largely occluded or destroyed objects using a conventional eigenspace method [1].

In order to employ an appearance-based eigenspace method for recognizing partially data-loss objects, we propose to collect various similar poses/image views from the partially occluded/destroyed object's shapes into an individual window, referred to as an "eigen window". Therefore, respective sets of similar images create various eigen windows. We, then, calculate a mean of each eigen window with respect to the collected/obtained poses, referred to as a "mean eigen window". This mean eigen window represents various object's poses in a generalized form. **Figure 1** shows an eigenspace with highlighting an eigen window. An eigen window having four similar poses is indexed of a particular object.

In Section 2, we review the eigenspace method, discuss the limitations of the eigenspace method, and explain how to overcome these limitations using the proposed mean eigen window method. Section 3 shows some of the experimental results and evaluates the performance. A concluding remark is placed in Section 4.

2. EIGEN WINDOW METHOD

First, we review the eigenspace technique [1] and discuss the limitations of the technique under apperance-change due to occlusion and shape-destruction. The eigen

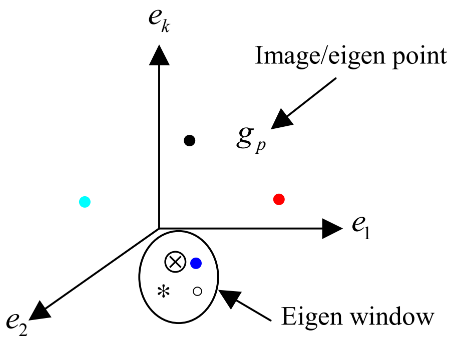

Figure 1. Demonstration of generating an eigen window.

window method [2] is also discussed and it's contradiction with the basic concept of traditional eigenspace technique is identified. Then, a mean eigen window method is proposed. This method is designed to overcome the preceding problems, which follows the basic eigenspace analysis [1] with simplifying the eigen window method [2].

2.1 Eigenspace Technique

Let M be the number of the images in a training set of a particular object. Each image is converted into a column vector z_i of length N:

$$[z_1, z_2, \ldots, z_M] \tag{1}$$

By subtracting the average image of the all images, we obtain the training matrix,

$$Z = [z_1 - c, z_2 - c, \ldots, z_M - c] \tag{2}$$

where $\mathbf{c}$ is the average image, and the size of the matrix Z is $N \times M$. The sample covariance matrix $\boldsymbol{Q}$, $N \times N$, is obtained from

$$Q = ZZ^T \tag{3}$$

This sample covariance matrix provides a series of eigenvalues λ_i and eigenvectors e_i $(i = 1, 2, \ldots, N)$ where each corresponding eigenvalue and eigenvector pair satisfies:

$$\lambda_i e_i = Q e_i \tag{4}$$

That is, matrix Q can be decomposed into N orthonormal components, of which the eigenvalues are λ_i. Thus, each image set can be described by a set of eigenvectors with associated weight factors, i.e., eigenvalues. If the number of images M is much smaller than the number of pixels N, the implicit sample covariance matrix can be used instead of the sample covariance matrix Q to calculate the first k eigenvectors.

For the sake of memory efficiency, we will ignore small eigenvalues and their corresponding eigenvectors using a threshold value, T:

$$W_k = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \geq T \tag{5}$$

where k is sufficiently smaller than the original dimension N.

From this reduced set of eigenvectors, the matrix is constructed to project an image z_i (dimension N) into the eigenspace as an eigen point $\boldsymbol{g}_i$ (dimension k).

$$\boldsymbol{g}_i = \boldsymbol{E}^T (z_i - \boldsymbol{c}) \tag{6}$$

This eigenspace method can drastically reduce the dimension of the images (N) to the eigenspace dimension (k) while keeping several of the most effective features that summarize the original images.

2.2 Limitations of the Eigenspace Technique

The eigenspace representation, a collection of image poses or points in the eigenspace, is very sensitive to image conditions such as background noise, image shift, occlusion of objects, scaling of the image, and illumination-change. As an effort to reduce these disturbance effects in the eigenspace, we have seen various studies using the eigenspace technique [2-4, 9]. However, the literatures do not provide such convenient clues for avoiding occlusion and destruction of objects in a particular environment. There are a number of expected practical applications, e.g., robotic rescue for estimating a disaster or collecting goods from the debris, industrial application for monitoring inventories, car navigation for obstacle identification, etc. As described, Ohba and Ikeuchi [2] proposed an eigen window method for such applications. However, their proposal contradicts with the concept of basic eigenspace technique. Moreover, it is so complicated to collect similar parts (edges) of objects in a window and to use matching algorithm. Thus, we propose rather very simple technique that merges an eigenspace and an eigen window method.

2.3 Mean Eigen window

We have investigated that pattern of eigenspaces change with changing the object's shape due to disturbance effects. The rate of pattern changes depends on the disturbance appeared in the object's shapes. This paper is, as a primary step, focuses on a defined partial disturbance due to occlusion or destruction. To reduce the disturbance effects, we propose to apply a mean eigen window where similar disturbed or non-disturbed appearances/image views are collected in a particular window, called the *eigen window*, and mean of each eigen window is taken for obtaining a generalized form of the appearances/views called the *mean eigen window.* We refer to this method as the "*mean eigen window*" technique.

2.3.1 Training Eigen Windows

The training set of eigen windows is given as:

$$F_M^S = [F_1^1, \ldots, F_M^1; F_1^2, \ldots\ldots, F_M^2; F_1^s, \ldots, F_M^S] \tag{7}$$

where, the training set contains S objects and M poses/images.

Let us consider F_m^s, the collection of similar poses/shapes with respect to image m of object S. Each F_m^s has the form

$$F_m^s = [z_{m,1}^s, z_{m,2}^s, \ldots, z_{m,P}^s] \tag{8}$$

Here P is the number of similar poses or orientations in an eigen window. Then, a mean eigen window is defined as

$$\bar{\boldsymbol{z}}_m^s = \frac{1}{P}\sum_{i=1}^{P} \boldsymbol{z}_{mi}^s \tag{9}$$

Therefore, the training eigen windows of Eq. (7) reform as

$$\bar{F}_M^S \equiv [\bar{z}_1^1, \bar{z}_2^1 \ldots, \bar{z}_M^1; \bar{z}_1^2, \bar{z}_2^2 \ldots, \bar{z}_M^2; \bar{z}_1^S, \bar{z}_2^S, \ldots, \bar{z}_M^S] \tag{10}$$

If we calculate an eigenspace using this mean eigen window, Eq. (6) can be reformed as:

$$\boldsymbol{g}_m^s = \boldsymbol{E}^T (\bar{\boldsymbol{z}}_m^s - \mathbf{c}) \tag{11}$$

where $\boldsymbol{c}$ is the average eigen window across the all eigen windows.

2.3.2 Matching Operation

Since we have obtained a mean eigen window which is similar to a simple eigenspace made from a set of image sequences, the system is prepared to accept the minimum description length principle that uses the L_1-norm. From an input unfamiliar/unknown image or pose including partially occluded or destroyed shapes, a sub-window image is obtained. The similarity between a training eigen window and an input eigen window is evaluated by calculating their distance in the eigenspace. The minimum distance

$$d_{M*}^{S^*} \equiv \min_{s,m} \{ \| \boldsymbol{g} - \boldsymbol{g}_m^s \| \} \tag{12}$$

is calculated to find the nearest learned point in the mean eigen window or eigenspace. For a certain threshold $\varepsilon(>0)$, if $d_{M*}^{S^*} < \varepsilon$ holds, we conclude that the unknown object of the unknown posture M' is similar to the one represented by the point $\boldsymbol{g}_{M*}^{S^*}$. It is noted that the unknown image point is denoted by $\boldsymbol{g}$.

3. EXPERIMENTAL RESULTS

In our study, we have considered that shapes of objects do not change ambiguously and object disturbance should not be more than 20% of the total shape of objects. Our study is limited to partially occluded or destroyed object's representation and recognition. The definition of occlusion refers that an object is placed with some other objects and some parts of the object cannot be viewed properly. Similarly, partial destruction refers that some portions of a particular object are lost by any means and a complete shape of the object is not available. In the experiment, we have taken a particular object with various disturbed and non-disturbed image situations that give us 5(=P) sets of image sequences. Here we suppose S=1.

A turntable is taken to obtain 18(=M) various poses via a digital camera in 20-degree rotation of each object's situation. Therefore, we have obtained a total of 90 training image samples, which are used for generating eigen windows.

Some of the considered object's situations such as non-disturbed situation, partially occluded and destroyed situation, etc., are shown in **Figure 2**. Figure 2 also shows some of image segmentation from their original video images. We have shown only the first pose of each image situation. This image segmentation with just an image

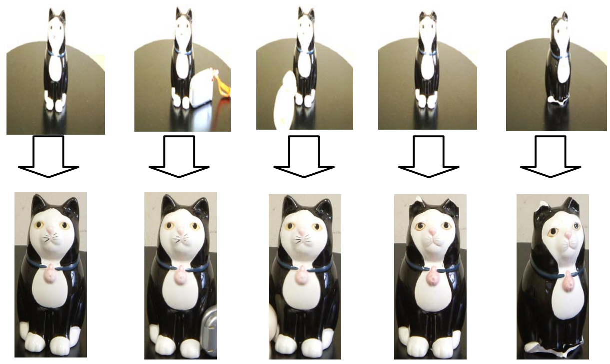

Figure 2. Some video image frames in the upper low and the images in the lower row used in the experiment.

portion eliminates major occlusion and we include the rest of disturbance, which cannot be removed by the segmentation in generating an eigenspace. It should be noted that we have not extracted or segmented any part of the objects even they are occluded. One may choose to extract the occluded parts for making eigenspace. However, it makes partial segmentation of the objects that contradicts to the conventional method and it is also computationally expensive. Once we have prepared the image sets, we generate eigen window by Eq. (7) and then make a mean eigen window by Eq. (9). Finally, the eigenspace is obtained by Eq. (11).

Therefore, we have obtained a mean eigen window which contains 18 ($=\overline{F}$) mean appearance-points (or mean window) in the eigenspace. **Figure 3** shows a magnified eigen windows created from 5 sets of training images. One may get confused that the indexed mean eigen widow are created by taking the mean of similar pose images but not from their discrete points. A mean eigen window is also placed in **Figure 4**. This mean eigen window has been created from the training samples, i.e., 5 sets of images and employed for the further recognition.

In case of recognition test, we have also taken the similar number of images (*P*=5 and *M*=18). It should be noted that these testing samples also include disturbed and non-disturbed object's shapes. The way of data-loss should not necessarily be same or similar between the training and testing images. We have projected each set of unknown orientations onto the mean eigen window to verify the performance of the proposed technique. **Table 1** shows experimental activities including the obtained recognition rates. We have obtained an average of 91.66% recognition rate. For reference, we have also calculated the recognition rates based on the original method [1] in order to compare the performance and we have obtained 53.66% of recognition rates.

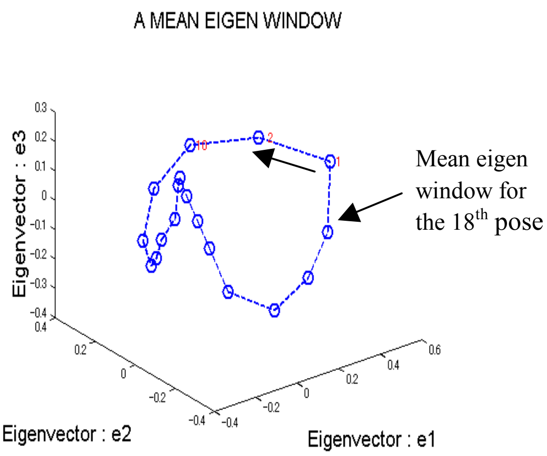

Figure 3. Magnified eigen window. Five different points sets obtained from the 5 sets of image situations.

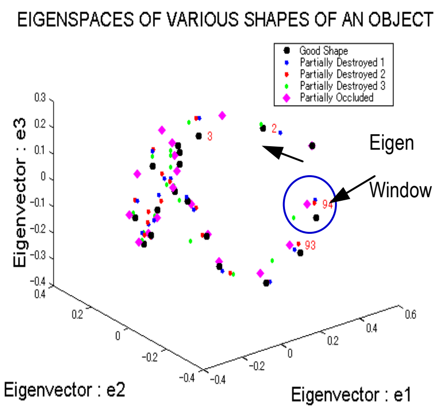

Figure 4. A mean eigen window used for further recognition purposes.

Table 1: Images used in the experiment and experimental results.

<table>
<tr><th>Samples</th><th>Non-disturbed shape</th><th>Occluded shape</th><th>Destroyed shape</th><th>Total pose</th><th>Recognition rates (Avg.)</th></tr>
<tr><td>Training</td><td>1</td><td>2</td><td>2</td><td>$18\times5=90$</td><td rowspan="2">91.66%</td></tr>
<tr><td>Testing</td><td>1</td><td>2</td><td>2</td><td>$18\times5=90$</td></tr>
</table>

4. CONCLUSIONS

This paper describes a robust recognition method, referred to as the *mean eigen window* method, to extend the standard eigenspace method and to simplify the eigen window technique which is able to recognize partially occluded/destroyed objects. The present study has overcome the limitations occurred in the eigen window technique and the eigenspace technique has been extended for tracking objects in the occluded or data-loss environments. However, the present investigation was limited to a particular object with its various data-loss environments. The proposed approach can be applicable to various industrial and military applications such as object picking, automatic target recognition, and surveillance and monitoring, etc.

The limitations of the *mean eigen window* method may be recognition under large occlusion and/or destruction and under various illumination conditions. Future work will concentrate on recognizing objects considering the unsolved issues.

5. REFERENCES

[1] Murase H., Nayar S. K.: "Visual Learning and Recognition of 3D Objects from Appearance," *Int'l J. Computer Vision,* Vol. 14, No. 1, pp. 5-24 (1995).

[2] Ohba K., Ikeuchi, K.: “Detectability, Uniqueless and Reliability of Eigen Windows for Stable Verifications of Partially Occluded Objects,” *IEEE Trans. on Pat. Anal. Machine Intell.*, Vol. 9, pp. 1043-1047 (1997).

[3] Leonardis A., Bischof H.: “Robust Recognition Using Eigenimages”, *Computer Vision and Image Understanding*, Vol.78, pp.99-118 (2000).

[4] Borotschnig, H. et. al.: “Appearance-Based Active Object Recognition”**,** *Image and Vision Computing,* Vol. 18, No. 9, pp. 715-727(2000).

[5] Turk, M. A., Pentland, A. "Face Recognition Using Eigenfaces," *Proc. CVPR 1991,* pp. 586-591, 1991.

[6] Pentland, P.A., Horowitz, B: "Recovery of Non-rigid Motion and Structure," *IEEE Trans. Pattern Analysis and Machine Intelligence,* Vol. 13, No. 7, pp. 730-742 (1991).

[7] Rahman, M. M., Ishikawa, S.: “Representing Human Motions by an Eigenspace”, *Journal of Biomedical Soft Computing and Human Sciences*, Vol. 9, No. 1, pp.27-33 (2003).

[8] Moghaddam, B., Pentland, A. P.: "Probabilistic Visual Learning for Object Detection," *Proc. 5th Int'l Conf. Computer Vision*, CD-ROM (1995).

[9] Ohba, K., Ikeuchi, K.: "Recognition of the Multi Specularity Objects Using the Eigen Window," *Proc. Int'l Conf. Pattern Recognition*, CD-ROM (1996).