MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

13—12

Image-based Control of Robot and Target Object Motions by Eigen Space Method

Isao Takahashi Koichiro Deguchi *
Faculty of Engineering
University of Tokyo

Abstract

This paper deals with the image-based control to move a robot-arm mounting a camera to the position where the given goal image was taken and, at the same time, target objects to the same positions as they were in the goal image. Their motions are determined to reduce the difference between the goal image and the image taken at their current positions.

We use the "eigen-space method" to obtain a compact representation of the image. To relate the image difference to the robot-arm and target object motions, we re-arrange the eigen space so that the representation linearly depends on the robot-arm and target object motions.

We describe this approach in detail and experiments using real robots and images.

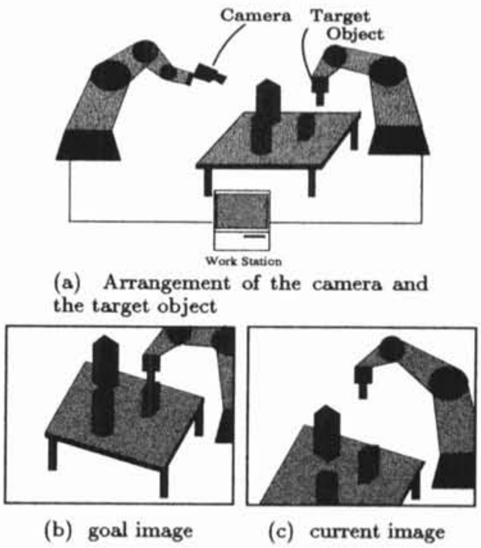
1 Introduction

It is an attractive application of machine vision to control the motion of a robot-arm mounting a camera based on the difference between a given goal image and an image taken at a current position of the robot-arm. Many methods have been proposed, but they are to control a robot-arm in a static scene where only the robot-arm position causes the change of the image. This enables the estimation of the robot-arm position relative to the goal position by observing some fixed objects by the camera. So the geometrical features (such as points, lines, and circles) of the observed objects are commonly used to relate the difference between images to the difference of robot-arm position from the goal[1][2].

In this paper, we deal with more complex situation; for a given goal image, our objective is to move a robot-arm mounting a camera to the position where a given goal image was taken and, at the same time, to move some target objects to the same positions as they were in the goal image. Their motions are determined to reduce the difference between the goal image and the current image taken by the camera step by step, and to let both the camera on the robot-arm and the target objects reach their respective goal positions.

*Address: 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656 Japan. E-mail: {isao,deguchi}@meip7.t.u-tokyo.ac.jp

A typical example is the case where a target object is picked up by a robot-arm and moved to a certain position, and its motion control is carried out based on the image taken by the camera mounted on another mobile robot-arm. For this purpose, also the camera position should be controlled to get a good observation (Figure 1).



(a) Arrangement of the camera and the target object

(b) goal image (c) current image

Figure 1: The target object is moved to a certain position and, at the same time, the camera is moved to get a good observation.

In these situation, it is not enough to observe some fixed objects by the camera. Moreover, what kind of image information must be utilized is not clear. It is desired to use image features which linearly depend on the motions of the robot-arm and target objects. An image composed of N pixels is represented as a point in N dimensional space. Though the number of pixels ($= N$) is usually very large, images obtained in a scene distributes within a small sub-space of the space. This sub-space is efficiently arranged by the "eigen-space method" which is often used in object recognition[3]. An image is represented as a linear combination of small number of "eigen images". They are given as eigen vectors of the covariance matrix of sample images. The subspace spanned by the eigen images is called "eigen-space". The coefficients of the linear combination (we call coefficient vector) can be interpreted as a compact representation of the image, and we use it to estimate the difference of the positions to their goal from images.

The relation between the coefficient vector and the positions of the robot-arm and the target objects is not clear when the eigen space is spanned by the eigen vectors, because the eigen vectors are obtained without the knowledge where the sample images are taken. To establish simple and clear relation between the coefficient vectors and the positions of the robot-arm and target objects, we rearrange the eigen space so that the coefficient vector linearly depends on the change of the robot-arm and target object positions.

In section 2, we describe how to arrange the eigen-space. The control scheme using the obtained eigen-space is described in section 3. Several experiments using real robots and images are shown in section 4. And we conclude our proposition in section 5.

2 Arrangement of the eigen-space

This section describes how to obtain a coefficient vector of the image which linearly depends on the change of the robot-arm and target object positions. To accomplish this, first, we apply the eigen space method for sample images taken in our situation to obtain the coefficient vector of the image. This operation intends to reduce the dimensions of the image. Next, we rearrange the sub-space of the eigen space so that the rearranged coefficient vector will linearly depend on the change of the robot-arm and target object positions.

As the result of those operations, rearranged coefficient vectors form a almost flat surface with respect to the robot-arm and the target object positions in their product space. We determine the path to lead to the goal on the surface (Figure2).

We describe the algorithm to arrange the eigen space and rearrange the sub-space of the eigen space in the following subsections.

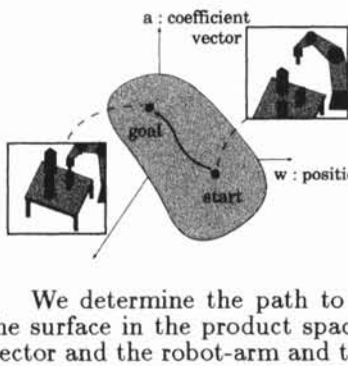


Figure 2: We determine the path to lead to the goal on the surface in the product space of the coefficient vector and the robot-arm and target object positions.

2.1 Eigen-space method

First, we take a number ($= K$) of sample images near the goal positions by shifting the camera and target objects. Each image is represented as N dimensional vector $\mathbf{I}_j (1 \leq j \leq K)$ composed of N pixel values. We assume here that their mean equals to zero for simplicity.

We define the image matrix P as

$$P = [\mathbf{I}_1, \mathbf{I}_2, \cdots, \mathbf{I}_K]. \tag{1}$$

Then, the covariance matrix Q of sample images is given as

$$Q = \frac{1}{K} P P^T. \tag{2}$$

We obtain eigen vectors $\mathbf{e}_i$ $(1 \leq i \leq M)$ of the matrix Q for its M largest eigen values as

$$\lambda_i \mathbf{e}_i = Q \mathbf{e}_i. \tag{3}$$

$$(\|\mathbf{e}_i\|^2 = 1, \lambda_1 \geq \cdots \geq \lambda_M) \tag{4}$$

Those eigen vectors are called "eigen images", and the space spanned by those eigen images is called eigen space. Then, an given image $\mathbf{I}$ is well represented as a linear combination of those eigen images as

$$\hat{\mathbf{I}} = \sum_{i=1}^{M} a_i \mathbf{e}_i. \tag{5}$$

Those eigen images are the best set of M orthonormal basis which minimize the mean square error($\frac{1}{K}\sum_{j=1}^{K} \|\mathrm{I_j} - \hat{\mathrm{I}}_\mathrm{j}\|^2$) for the given image set.

Each coefficient a_i for an image $\mathbf{I}$ is obtained by inner product of the image and the corresponding eigen image as

$$a_i = \mathbf{e}_i^T \mathbf{I}. \tag{6}$$

We call the M dimensional vector composed of those coefficients "coefficient vector" which is given as

$$\boldsymbol{a} = (a_1, \cdots, a_M)^T = E^T \mathbf{I} \tag{7}$$

where

$$E = [\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_M]. \tag{8}$$

2.2 Rearranging the eigen space

Now, we have obtained the eigen images to form a compact representation of the image as the coefficient vector. The next step is to rearrange the eigen images so that each component of the coefficient vector will linearly depend on the change of the robot-arm and target object positions.

Let us denote the robot-arm and target object positions with a vector $\boldsymbol{w}$. Then, possible combinations of a_i and $\boldsymbol{w}$ form a surface in a_i-$\boldsymbol{w}$ space. Denoting its surface normal vector with $\boldsymbol{l}_i$, $L^T (= [\boldsymbol{l}_1, \cdots, \boldsymbol{l}_M]^T)$ is interpreted as so-called the interaction matrix[1].

When each surfaces is a flat plane, the interaction matrix becomes constant, and using the pseudo-inverse of this matrix, we estimate the positions relative to their goal by the difference between the goal image and the image taken at current robot-arm and target object positions (Figure3).

To achieve this strategy, we determine rearranged eigen images $\{\phi_1, \phi_2, \cdots, \phi_S\}$ $(S \leq M, \|\phi_i\|^2 = 1)$ so that the rearranged coefficient vector $\boldsymbol{a}'$ is well

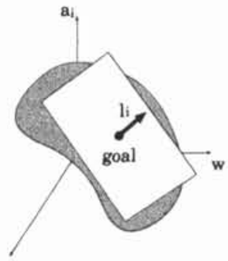


Figure 3: The normal vector of the surface in a_i-w space is interpreted as the column vector of the interaction matrix.

approximated by linear form of the robot-arm and target object positions weighted with the interaction matrix L^T. The rearranged coefficient vector for the image $\mathbf{I}_i$ is obtained as

$$\begin{array}{c} \boldsymbol{a}'_i = (a'_1, \cdots, a'_S)^T = E'^T \mathbf{I}_i. \\ (E' = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \cdots, \boldsymbol{\phi}_S]) \end{array} \quad (9)$$

If such the rearranged eigen images and the interaction matrix are obtained, we well approximate the coefficient vector as

$$\begin{array}{c} \hat{\boldsymbol{a}}'_i = L^T(\boldsymbol{w}_i - \boldsymbol{w}_G) + \boldsymbol{a}'_G. \\ (\mathrm{G} \ : \ \mathrm{goal}) \end{array} \quad (10)$$

They are obtained recursively. We determine ϕ_1 and $\boldsymbol{l}_1$ at first, then determine ϕ_2 and $\boldsymbol{l}_2$, and so on.

We represent the rearranged eigen image as a linear combination of the original eigen images obtained in section 2.1. To compose ϕ_1 as a linear combination of $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_M\}$, let us denote the weighting coefficients with M dimensional vector $\boldsymbol{d}_1$,

$$\begin{array}{c} \phi_1 = E\boldsymbol{d}_1. \\ (||\boldsymbol{d}_1||^2 = 1) \end{array} \quad (11)$$

Then, the mean square approximation error for the first component of the coefficient vector is given as

$$Err = \frac{1}{K}||\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}||^2 \quad (12)$$

where

$$\boldsymbol{\beta} = Y^T \phi_1 \quad (13)$$

$$\hat{\boldsymbol{\beta}} = W^T \boldsymbol{l}_1 \quad (14)$$

$$Y = [\mathbf{I}_1 - \mathbf{I}_G, \mathbf{I}_2 - \mathbf{I}_G, \cdots, \mathbf{I}_K - \mathbf{I}_G] \quad (15)$$

$$W = [\boldsymbol{w}_1 - \boldsymbol{w}_G, \boldsymbol{w}_2 - \boldsymbol{w}_G, \cdots, \boldsymbol{w}_K - \boldsymbol{w}_G]. \quad (16)$$

This error(12) is minimized when

$$\boldsymbol{l}_1 = (WW^T)^{-1} W \boldsymbol{\beta}. \quad (17)$$

Then, its minimum value is

$$\begin{array}{rcl} Err_{min} & = & \boldsymbol{d}_1^T B \boldsymbol{d}_1. \\ (B \equiv \frac{1}{K} E^T Y (I - W^T (WW^T)^{-1} W) Y^T E) \end{array} \quad (18)$$

While this minimum value must be small, the variance of the coefficient must be sufficiently large. So we take a ratio of the minimum value and the variance as

$$\frac{Err_{min}}{\sigma^2} = \frac{\boldsymbol{d}_1^T B \boldsymbol{d}_1}{\boldsymbol{d}_1^T \Lambda \boldsymbol{d}_1} \quad (19)$$

where

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_M). \quad (20)$$

and determine $\boldsymbol{d}_1$ to minimize this ratio. Such the vector $\boldsymbol{d}_1$ is obtained as the eigen vector which corresponds to the smallest eigen value of the matrix

$$\Lambda^{-1} B. \quad (21)$$

Substituting the obtained $\boldsymbol{d}_1$ into (11) and (17), we obtain the first rearranged eigen image ϕ_1 and the first column $\boldsymbol{l}_1$ of the interaction matrix. Second rearranged eigen images are obtained by determining $\boldsymbol{d}_2$ which minimizes (19) in condition that it is orthogonal to $\boldsymbol{d}_1$. Repeating these process, we compose the rearranged eigen images $\{\phi_1, \phi_2, \cdots, \phi_S\}$ and obtain the interaction matrix as

$$L^T = [\boldsymbol{l}_1, \boldsymbol{l}_2, \cdots, \boldsymbol{l}_S]^T. \quad (22)$$

The space spanned by the rearranged eigen images becomes the sub-space of the original eigen space.

3 Control Scheme

We use a simple feedback algorithm to move a robot-arm and target objects step by step to their goal positions, which is so-called "Look and Move" strategy. When the goal image and its coefficient vector $\mathbf{a}'_G$ are given, we start next processes and repeat until the same image as the given goal image is obtained.

1. Take an image $\mathbf{I}$ at a current positions of the robot-arm and the target object.
2. Obtain the coefficient vector $\mathbf{a}'$ of the image,
$$\mathbf{a}' = E'^T \mathbf{I}$$
3. Estimate the difference of the robot-arm and target object positions to their goal positions,
$$\Delta \mathbf{w} = L^{T+}(\mathbf{a}' - \mathbf{a}'_G)$$
(L^{T+} : Pseudo inverse of the interaction matrix)
4. Move by small amounts $k\Delta \boldsymbol{w}$ toward the estimated goal positions.
5. If the amount of motions is sufficiently small, the control finishes. Otherwise, go to 1 and repeat these steps.

4 Experiments

This section shows some experimental results. A robot-arm mounting a camera moved in a scene where a target object (grasped by another robot-arm) was also moving. Other objects in the scene were static.

The robot-arm mounting the camera and the target object were moved to the same positions where the given goal image had been taken, and where the target object had been in the goal image. In the experiments shown here, we took 61 images at different camera positions with the same target object position, and 60 images at the different target object positions with the same camera position. From these images, we obtained 20 eigen images. Then, we used them to obtain 8 rearranged eigen images. The goal image and rearranged eigen images are shown in Figure 4.

Figure 5 shows the relation between the coefficient vector and the robot-arm and target object positions. Two components of the rearranged coefficient vector are plotted with respect to two coordinates of the robot-arm and target object positions. This figure shows those relations form almost flat surfaces.

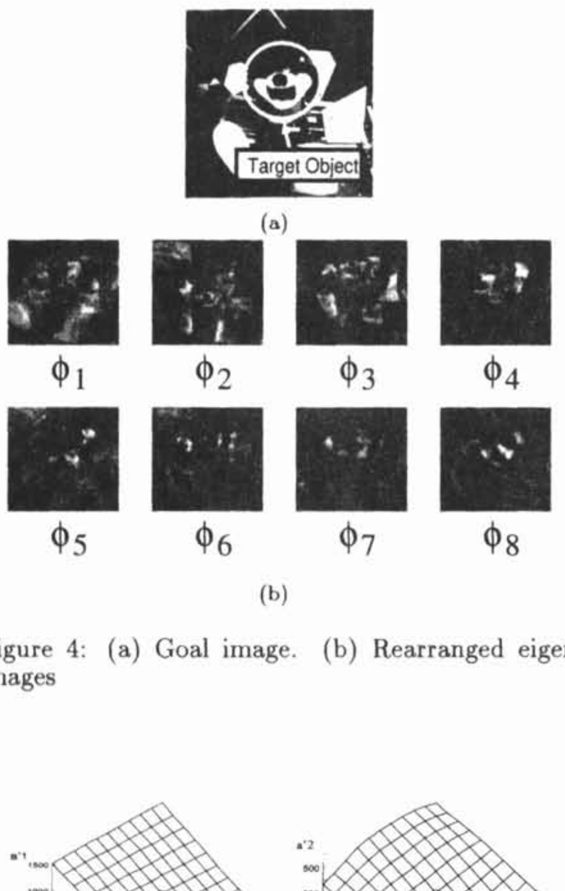
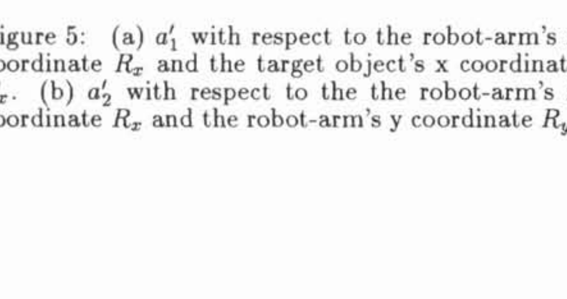


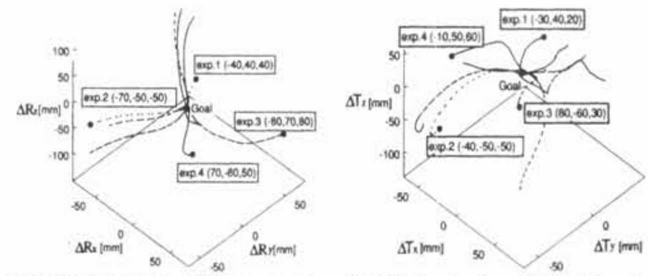
Figure 4: (a) Goal image. (b) Rearranged eigen images



(a) (b)

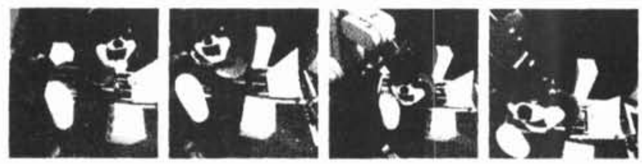
Figure 5: (a) a'_1 with respect to the robot-arm's x coordinate R_x and the target object's x coordinate T_x. (b) a'_2 with respect to the the robot-arm's x coordinate R_x and the robot-arm's y coordinate R_y.

Using these rearranged eigen images and the interaction matrix, we controlled the robot-arm and target object motions with the algorithm described in section 3. Their motions were converged to the goal from wide range of initial positions. Figure 6 shows the trajectories of the robot-arm and the target object starting at several initial positions.



(a) Trajectories of the robot-arm motion.

(b) Trajectories of the target object motion.



(c) exp.1 (d) exp.2 (e) exp.3 (f) exp.4

Figure 6: Experimental results starting at several initial positions. (c) ~ (f) are the images taken at the initial positions.

5 Conclusion

We have proposed a method which controls a robot-arm mounting a camera and, at the same time, objects motions in the scene. We achieve, by a single camera, to control of its self-motion and target object motion observed by the camera. Our method need not employ complex geometrical relations between the robot-arm and objects in the scene. So, more complex situation, for example, with multiple target objects, can be treated as the same manner. The case where there are two moving objects in the scene has already examined with good results. We are going to examine with multiple cameras as a further work.

References

[1] B. Espiau, F. Chaumette, and P. Rives, "A New Approach to Visual Servoing in Robotics," IEEE Trans. on Robotics and Automation, Vol.8, No.3, pp.313-326, 1992.

[2] J.T.Feddema, C.S.George Lee and O.R Mitchell, "Feature-based Visual Servoing of Robotic Systems," In K. Hashimoto, editor, Visual Servoing, World Scientific Series in Robotics and Automated Systems, Vol.7, pp.105-138, 1993.

[3] H.Murase and S.K.Nayer, "Learning and Recognition of 3D Objects from Appearance," Proc. of IEEE Workshop on Qualitative Vision, pp.39-50, 1993.