Proceedings of the Fifth IASTED International Conference
Signal Processing, Pattern Recognition, and Applications
February 13–15, 2008, Innsbruck, Austria
ISBN Hardcopy: 978-0-88986-717-8 / CD: 978-0-88986-718-5

EMBEDDING VIEW-DEPENDENT COVARIANCE MATRIX IN OBJECT MANIFOLD FOR ROBUST RECOGNITION

Lina, Tomokazu Takahashi, Ichiro Ide, Hiroshi Murase
Graduate School of Information Science,
Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8601, Aichi, Japan
{lina, ttakahashi}@murase.m.is.nagoya-u.ac.jp, {ide, murase}@is.nagoya-u.ac.jp

ABSTRACT
Variations in camera-captured images usually occur naturally. For example, the appearance of an object usually differs for every pose and degradation effect might occur during the capturing process. While we could use a simple manifold to represent the variability of pose, relying on the simple manifold technique to deal with both pose and degradation problems is not possible, since a simple manifold does not take into account the information of sample distributions in feature space. In this paper, we propose a technique which embeds view-dependent covariance matrix in object manifold to develop a robust 3D object recognition system. Here, the view-dependent covariance matrices were obtained in an efficient way by interpolating eigenvectors and eigenvalues along the manifold. Experiment results showed that our developed 3D object recognition system could accurately recognize 3D objects even from images which are influenced by geometric distortions and quality degradation effects.

KEY WORDS
Object recognition, appearance manifold, covariance matrix, eigenspace

1. Introduction

Recognizing a 3D object from its 2D images raises many challenges. For more than a decade, the main issue for appearance-based approach is to solve the pose problem. It is not surprising that the performances of object recognition systems drop significantly when large pose variations are present in input images [1]. Therefore, various attempts have been made to handle this issue.

Earlier methods focused on constructing invariant features [2], synthesizing a prototypical frontal view [3], or classify pose problems [4]. Following the popular eigenface approach which was proposed by Turk and Pentland [5], many techniques have been extended to explore recognition in eigenspace domain. Recently, researchers improved the previous eigen model with the use of appearance manifold in eigenspace in order to achieve pose-invariant recognition. Addressing various problems, many types of appearance manifolds have been developed, such as the simple appearance manifold in [6-7] which could handle pose and illumination variations, the appearance manifold with probabilistic techniques [8-9] for handling various facial changes, the layer-transparent manifold in [10] for recognizing occluded objects, etc.

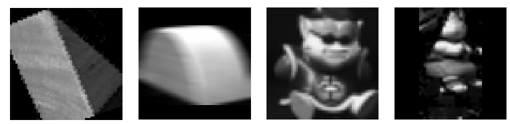

Figure 1. Samples of 3D objects with geometric distortions (translation, rotation) and quality degradations (motion blur)

We found that the work of Murase and Nayar in [7] is more favorable, due to its simplicity and applicability to more general pose variation problems. Thus, in this paper, we put our focus on their work. However, the disadvantage of their Parametric Eigenspace approach is the model only works well when the input images have no degraded effects. Unfortunately, this assumption is not realistic in real-world applications. Some degradation effects usually occur and contaminate the original images during the capturing process or segmentation process. Thus, relying on a simple manifold to handle this problem is not sufficient. Fig. 1 shows some image samples of 3D objects with some geometric distortion and quality degradation effects.

We have showed that to overcome this issue, constructing an appearance manifold with embedded covariance matrix is very useful. By using this appearance manifold model, the robustness of the system will be increased, since the manifold could capture the pose changes and the embedded covariance matrix could define the sample distribution information of every pose along the manifold. Moreover, since the appearance of an object in the captured image is different for every different pose, the covariance matrix value is also different for every pose. Thus, it is necessary to construct covariance matrix which is view-dependent.

Considering the models previously proposed by us in [11], a major limitation still exists, such as the need to critically control the correspondence between learning points in the interpolation process for manifold construction. When a huge number of learning points exists in the system, the controlling process becomes costly and time consuming.

In this paper, we propose the View-dependent Covariance matrix by Eigenvector Interpolation (VCEI) method where every mean vector and covariance matrix has different value for each training pose. The advantage of our proposed method is that it is not necessary to perform controlling process on the training images. In order to cover the untrained poses, we construct the appearance manifold by interpolating only the eigenvectors and eigenvalues of two consecutive trained poses. Thus, the appearance manifold will be noise-invariant and efficient.

The remainder of this paper is organized as follows: we describe the construction process of manifold in Section 2. Section 3 presents the description of embedding process of view-dependent covariance matrix in an object manifold. Section 4 shows and discusses the performance evaluation in recognizing 3D objects. Finally, Section 5 presents our conclusion.

2. Manifold of Object

Generally, the appearance-based approaches deal with a set of learning images in various capturing conditions. Since these images are usually high-dimensional images, they could not be applied directly due to efficiency reasons. Here, PCA is used to efficiently represent a collection of images by reducing their dimensionality. PCA represents a linear transformation that maps the original *n*-dimensional space onto a *k*-dimensional feature subspace where normally *k*<<*n*.

Next, the first *k* eigenvectors will be used to project *S* learning samples of *P* objects with *H* poses. Thus, with $\mathbf{x}_s^{(p)}(\theta_h)$ the *s* sample image of object *p* with horizontal viewpoint θ_h and $\mathbf{e}_i$ are the eigenvectors, the new feature vectors $\mathbf{g}_s^{(p)}(\theta_h) \in \Re^k$ are defined by $\mathbf{g}_s^{(p)}(\theta_h) = [\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k]^T \left(\mathbf{x}_s^{(p)}(\theta_h) - \mathbf{c}\right)$ where $\mathbf{c}$ is the mean vector of sample images. These eigenvectors $\mathbf{e}_i$ were obtained by solving the eigen decomposition $\lambda_i \mathbf{e}_i = \mathbf{Q}\mathbf{e}_i$, where $\mathbf{Q}$ is the auto correlation matrix of the training set and λ_i is the eigenvalue associated with the eigenvector $\mathbf{e}_i$. Note that in this section, the eigenvectors and eigenvalues are used only to construct the eigenspace. Meanwhile, later in the next sections, the eigenvectors and eigenvalues are derived from the covariance matrix of each training-pose.

In the Simple Manifold (SM) method, after transforming learning images onto the eigenspace, the manifold of an object could be obtained by interpolating the mean vector of training images of one pose to its consecutive poses. Practically, we can simply apply an interpolation method to construct the manifold between training images in two-consecutive poses. Fig. 2 shows the illustration of the construction process of a simple manifold of an object in the feature space.

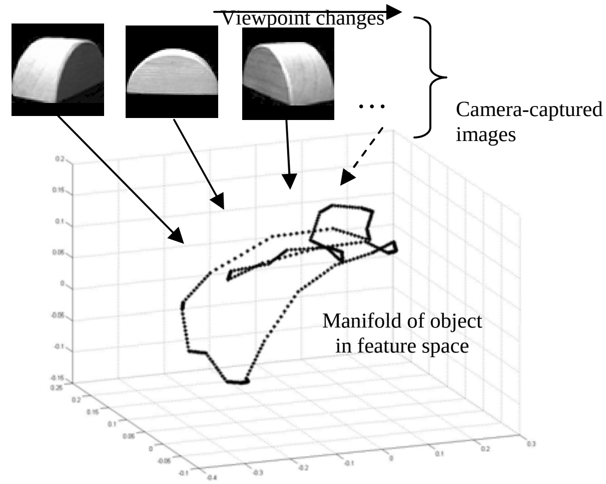

Figure 2. A simple manifold of an object in feature space

3. Embedding View-dependent Covariance Matrix in Object Manifold

This section describes the process of constructing the appearance manifold with embedded covariance matrix in eigenspace and the recognition process of input images using the Mahalanobis distance measurement.

The construction process of View-dependent Covariance matrix by Eigenvector Interpolation (VCEI) method consists of two stages of interpolation process: the interpolation of mean vectors and the interpolation of eigenvectors and eigenvalues. The mean vector is used to represent the center point of samples in each learning pose, while the eigenvectors and eigenvalues represent the distribution of samples in each pose. The interpolation process is useful to cover the information of untrained poses.

Basically, the interpolation process of the mean vector can be done by simply using one of the several existing interpolation algorithms. Meanwhile, the interpolation process of eigenvectors and eigenvalues are done based on high-dimensional rotation theory. As the eigenvectors and eigenvalues can be considered as axes directions and lengths of a hyper-ellipsoid in the eigenspace, we consider obtaining the covariance matrices of untrained poses by rotating the hyper-ellipsoids of two-consecutive trained poses. Fig. 3 shows the 2D illustration of the interpolation process of eigenvectors and eigenvalues in the feature space. Meanwhile, Fig. 4 shows the construction process of an appearance manifold with view-dependent embedded covariance matrix. Here, we