

SVD BASED ALGORITHMS FOR ROBUST FACE AND OBJECT RECOGNITION IN ROBOT VISION APPLICATIONS

S. Noushath

Department of Studies in Computer Science
University of Mysore
Mysore - 570 006, India
nawali_naushad@yahoo.co.in

Ashok Rao

Department of Electronics and Communications
Shri Jayachamarajendra College of Engineering
Mysore - 570 006, India
ashokrao.mys@gmail.com

G. Hemantha Kumar

Department of Studies in Computer Science
University of Mysore
Mysore - 570 006, India
hemanth_kumar@lycos.com

ABSTRACT

Contemporary society is highly networked and thus biometrics based surveillance has paramount importance for various security based reasons. The automatic, remote and robot vision based system are being deployed in a large way [11, 12]. The success of these schemes is highly dependent on robust algorithms for both face and object recognition. In this paper, we propose a very robust approach to face/object recognition based on Singular Value Decomposition (SVD). We first provide technical reasons to substantiate the claims made by T.Yuan et al [1], then we provide appropriate reasons for the robust behavior of SVD and finally, we corroborate the proposed concepts through extensive experiments on two standard databases which includes face and objects under both clean and noise conditions.

KEYWORDS

Face Recognition, Object Recognition, Singular Value Decomposition, Principal Component Analysis (PCA), Fisher Linear Discriminant, Analysis (FLD), Noise Analysis

1. INTRODUCTION

Among various biometric traits such as speech, fingerprint, handwriting, face recognition etc. there is tremendous attention on face recognition due to its technical challenges and potential applications. This area has roots which dates back to last four decades and is still a growing area in robot vision research [2]. It has many potential applications such as biometrics, surveillance systems, content based video retrieval systems, antique verification and face image transmissions. Nevertheless, there are still many challenging problems such as recognition under different noisy conditions, occlusions, illumination and/or pose variations.

There are two main approaches to face recognition. The first approach is feature based matching using the relationship between facial features such as eyes, nose and mouth [3, 4]. The second approach is template matching using the holistic features of the face images[4, 7, 8, 9].

Many successful methods have been proposed using holistic features of the faces and one popular category is based on Singular Value Decomposition (SVD) for face recognition [5, 6, 7, 8]. Hong [5] has proposed a SVD based recognition method which uses the singular values (SV) as the feature vectors.

The effectiveness of SVD has been tested in [5] and [6]. In [5], an error rate of 42.47% was

recorded which was thought to be caused by the statistical limitations of the small samples. Cheng [6] proposed a human face recognition method based on the statistical model of small sample size that also used the SV's as the face features, where a good discrimination ability was obtained with an accuracy rate of 100%. This accuracy is obtained on a moderate size face database and its performance on a larger database is unclear. In [7], a new face recognition method based on SVD was proposed by using reconstruction error on singular values of different face images and data fusion method. It was reported that the recognition accuracy of this method reaches 96% on the ORL face database. Furthermore, D.Zhang et.al. [8] have proposed a method based on SVD and conventional PCA [9] to improve the classification performance of the latter method under the condition when the face database contains only one image for training. This situation is called single sample per person problem which has evinced more importance in the current face recognition technology.

Recently a paper was proposed by T.Yuan et.al [1], in which, unlike the claims made by the aforementioned papers, they have claimed that the SVs contain little useful information for face recognition and most important information is encoded in the two orthogonal matrices of the SVD. They have ascertained their claims through simple demonstration and no supporting theory was provided to affirm their claims. Moreover, the reasons for the efficiency and robustness of SVD under various noise conditions (modelled by different distributions) are still not being proclaimed in the literature of face recognition community. In our work we address this issue by providing convincing reason.

The three essential contributions of our work are: (i) Providing technical reasoning to substantiate the claims made by T.Yuan et.al [1]. (ii) Providing appropriate reasons for the efficacy and robust behavior of orthogonal projections of SVD. (iii) Performance evaluation for comparing the results of SVD with the eigen vectors of the covariance matrix under both noisy as well the clean conditions.

The rest of the paper is organized as follows: In section 2, we briefly outline the algorithm of SVD. The experimental results and analysis are given in section 3. The conclusions are drawn in section 4.

2. SINGULAR VALUE DECOMPOSITION

The singular value decomposition (SVD) is one of the most important tools of numerical signal processing. It is employed in a variety of signal processing applications, such as spectrum analysis, filter design, system identification etc.

The theory of SVD states that any matrix A of size $m \times n$ can be factorized into a product of unitary matrices and a diagonal matrix, as the following:

$$A = U \Sigma V^T \quad (1)$$

where $U \in \mathfrak{R}^{m \times m}$ is unitary, $V \in \mathfrak{R}^{n \times n}$ is unitary, and $\Sigma \in \mathfrak{R}^{m \times n}$ has the form $\Sigma = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_p)$, where p is the minimum value of m or n . The diagonal elements of Σ are called the singular values of A and are usually ordered in descending manner.

We use the orthogonal matrix U as the projection vectors (instead of eigen vectors of covariance matrix) for feature extraction while implementing Eigenface [9] and Fisherface [10] methods. In all our discussions, we call these methods as PCA+SVD and FLD+SVD respectively. In contrast to eigenvector calculations involved in conventional algorithms, the SVD has several advantages: (i) computationally efficient and (ii) it is robust under noise conditions.

The SVD are known to be more robust than usual Eigen vectors of covariance matrix. This is because, the robustness are determined by the directional vectors rather than mere scalar quantity like magnitudes (Singular values stored in Σ). Since U and V matrices are inherently orthogonal in nature, these directions are encoded in U and V matrices. This is unlike the Eigenvectors which need not be orthogonal. Hence, a small perturbations like noise (in case of images or signals) have very little effect to disturb the orthogonal properties encoded in the U and V matrices. This we believe could be the main reason

for the robust behavior of the SVD. This also affirms the claims made by T.Yuan et al [1].

3. EXPERIMENTAL RESULTS

In this section, we conduct several tests to affirm the proposed claims. Robot vision applications encounter most frequently face and object recognition tasks generally under noisy environment. For this purpose we have used two standard dabases namely ORL face database¹ and COIL-20 object database². In all our experiments, the nearest neighbor classifier (Euclidean distance) is used for classification purpose and the number of projection vectors (k) is controlled by the following equation:

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} \geq \theta \quad (2)$$

Where λ_i are eigen values such that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \dots \geq \lambda_N$, N is the total number of eigen values, k is any scalar value and θ is a preset threshold. In most of our experiments, θ is set to 0.98 to select the value of k . We call this as *confidence interval* throughout the paper. Finally, all our experiments are carried out on a P4 3GHz CPU and 1GB RAM memory on Matlab 7 platform.

3.1. RESULTS ON ORL DATABASE

The ORL database contains 400 images corresponding to 40 different persons with each person having ten different views. The variations include changes in expression, facial details and slight variations in pose. Few example images are shown in Fig.1. On this database, we have evaluated the performance of algorithms under two test conditions: 1) Performance under clean or noise free conditions, i.e. when both sample size and pose are varied and 2) Rigorous experiments under different noisy conditions modeled by five various distributions: Uniform (for Salt-and-Pepper behavior), Gaussian, Weibull, Exponential and Beta.

First, an experiment was performed using first five samples per person for training and the remaining images are used for testing. So effectively, 200 ($40 \times 5 = 200$) images were constituted in both training and testing sets. Table 1 gives the comparison of PCA based methods on recognition accuracy³, corresponding dimension of the feature set and running time for 98% confidence interval. It is clear from the table that the SVD based PCA and FLD methods outperformed their conventional counterpart algorithms interms of recognition accuracy and running time costs.



Figure 1 Few Sample Images from ORL Database

Table 1 Comparing Methods for Five Training Samples (98% Confidence Interval)

Methods	Accuracy (%)	Dimension	Time (secs)
PCA [9]	92.00	151	25.77
PCA+SVD	94.00	187	17.89
FLD [10]	83.50	149	36.98
FLD+SVD	86.25	187	29.33

The issue of noise modeling is crucial to check the robust behavior of the algorithm under real time pattern recognition and computer vision tasks. In our work, we have modeled five noise environments using different continuous distributions and used them in their descretized version⁴

³ Recognition accuracy is defined as the ratio of correctly classified images to the total number of test samples

⁴ Matlab command `imnoise` is used to add salt-and pepper effect. Control parameters μ and σ are fixed to 0 and 0.5 respectively for Gaussian noise, the λ value is fixed to 1 while creating exponential noise, K and α are fixed to 0.5 and 2 to generate weibull noise and both α and β values are fixed to 0.5 to create the beta noise, in their respective distributions

¹ <http://www.uk.research.att.com/facedatabase.html>

² www1.cs.columbia.edu/CAVE/research/softlib/coil-20.html

We selected first image from each class and generated 10 noise images by varying the noise density from 0.1, 0.2, . . . , 1.0. Likewise, 50 noise images are created for each class corresponding to 5 different distributions. Since ORL contains 40 classes, 2000 noisy images are created, which we use as test samples. First five clean samples from each class of ORL are used as training samples. Table 2 presents the average recognition accuracy (average of ten different noise densities) obtained by various algorithm. Again it can be ascertained that the SVD based algorithms remained robust for all the noise conditions.

Table 2 Performance Comparison for Different Noise Conditions for 98% Confidence Intervals

Noise Conditions	Methods			
	PCA	PCA + SVD	FLD	FLD + SVD
Gaussian	96.75	98.50	89.25	90.00
Salt-and-Pepper	63.75	64.75	58.50	60.75
Exponential	79.00	82.50	75.75	79.75
Weibull	67.00	73.50	64.25	71.00
Beta	100	100	98.50	99.00

3.2. RESULTS ON COIL-20 DATABASE

This database contains 1440 grayscale image of 20 objects. The objects are placed on a motorized turntable against a black background. The turn table was rotated through 360 degrees to vary object pose with respect to fixed camera. Images of the objects were taken at pose interval of 5 degrees, which correspond to 72 images per object. The sample views of twenty objects considered in this database are shown in Fig.2

We conducted an experiment on this database by varying the number of training samples for 98% confidence interval. This result is shown in Table 3. It can be seen that the SVD based algorithms outperformed their counterparts. These results endorse the applicability of the SVD based algorithms for object recognition.



Figure 2 Few Sample Images from COIL-20 Database

Table 3 Recognition Accuracy (%) for Varying Number of Training Samples (98% Confidence Intervals)

Noise Conditions	Number of Training Samples			
	18	24	36	48
PCA	75.97	82.70	90.48	93.33
PCA+SVD	77.15	84.51	91.52	94.44
FLD	67.63	72.77	81.73	88.75
FLD+SVD	72.91	78.81	87.01	91.94

4. CONCLUSION

In this paper we have proposed SVD based algorithms for robust face/object recognition. We have ascertained the efficacy of the SVD based algorithms for both face and object recognition which are useful in later robot vision related tasks. We have conducted experiments under both clean and noise conditions. For later case, we have modeled five different noise conditions using continuous distributions and used them in their discretized version. Overall, we observed SVD based algorithm performing significantly better than conventional algorithms which use covariance matrix construction and subsequent Eigen value decompositions.

5. REFERENCES

- [1] T. Yuan., T.Tan, Y.Wang and F.Fang, 2003, Do singular values contain adequate information for face recognition?, *Pattern Recognition*, Vol.36, 649-655.
- [2] R.Chellappa, C.L.Wilson, S.Sirohey, 1995, Human and machine recognition of faces: a survey, *Proceeding of IEEE*, Vol.38, No.5, 705-740.
- [3] A.Samal, P.A.Iyengar, 1992, Automatic recognition and analysis of human faces and facial expressions: a survey, *Pattern Recognition*, vol.25, No.1, 65-77.
- [4] R.Brunelli, T.Poggio, 1993, Face Recognition: feature versus templates, *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol.15, No.10, 1042-1052.
- [5] Z.Hong, 1991, Algebraic feature extraction of image for recognition, *Pattern Recognition*, Vol.24, 211-219.
- [6] Yong-Qing Chen, 1991, Human face recognition method based on the statistical model of small sample size, *SPIE Proceedings of the Intelligent Robots and Computer Vision*, Vol.16, 85-95.
- [7] Wang Yun-Hong, Tan Tie-Niu, Zhu Yong, 2000, Face identification based on singular value decomposition and data fusion, *Chinese Journal of Computers*, Vol.23, No.6(in chinese).
- [8] Peter N. Belhumer, J.P.Hespanha and David J.Kriegman, 1997, Eigenfaces vs Fisherfaces: Recognition using class specific linear projection, *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, 711-720.
- [9] D.Zhang, S.Chen and Z.-H.Zhou, 2005, A new face recognition method based on SVD perturbation for single example image per person, *Applied Mathematics and Computation*, Vol.163, No.2, 895-907.
- [10] M.Turk and A.Pentland, 1991, Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, Vol.3, No.1, 71-86.
- [11] Mark Fiala and Anup Basu, 2004, Robot navigation using panoramic tracking, *Pattern Recognition*, Vol.37, No.11, 2195-2215.
- [12] Ilkka Autio and Tapio Elomaa, 2003, Flexible view recognition for indoor navigation based on Gabor filters and support vector machines, *Pattern Recognition*, Vol.36, No.12, 2769-2779.