A METHODOLOGY FOR IMPROVING RECOGNITION RATE OF LINEAR DISCRIMINANT ANALYSIS IN VIDEO-BASED FACE RECOGNITION USING SUPPORT VECTOR MACHINES

Sreekar Krishna and Sethuraman Panchanathan

Center for Cognitive Ubiquitous Computing
Arizona Sate University
ASU Main, Tempe, AZ - 85287

ABSTRACT

This paper proposes a two-step methodology for improving the discriminatory power of Linear Discriminant Analysis (LDA) for video-based human face recognition. Results indicate that, under real-world video capture conditions, face images extracted from a video sequence have enough 3D rotations, illumination changes and background variations to reduce the discriminatory power of an LDA classifier. The proposed method involves deriving an LDA subspace from carefully selected subsets of face images that fall within a narrow range of pose angles, and then growing the classification regions in the LDA subspace using face images with a wider range of pose angle changes, illumination changes, and background variations. Polynomial Support Vector Machines (SVM) are shown to provide better recognition rates by defining the boundaries between clusters that represent the faces of different subjects. Results show that there is an improvement in the recognition rate when the LDA subspace is derived with this methodology than when it is derived with a set of face images with a widely divergent set pose angles, illumination variations, and backgrounds.

1. INTRODUCTION

Recently, with the availability of more powerful computers, real time processing of every frame in a video sequence has become possible. Taking advantage of this, face recognition researchers have started processing video sequences in an attempt to provide more reliable face recognition algorithms [1]. A few seconds of video contain considerably more information than a single face image. Within a single video sequence, it is possible to see multiple views of a face from various pose angles, and with a variety of illuminations, in addition to changes in the background. Even when the camera is stationary, 3D rotations of the face cause pose angle and illumination changes. Although these changes seem subtle to humans, they significantly affect the performance of face recognition algorithms.

This paper demonstrates how the learning characteristics of LDA changes for video-based face recognition due to the variations in pose, illumination and background, and substantiates the necessity for the new methodology. Our choice of LDA is based on a recent study [2], where the authors have shown that LDA-based face recognition is more robust to pose and illumination changes than three other widely used face recognition algorithms (Principle Component Analysis, Bayesian Intra-personal Classifier and Hidden Markov Model).

The rest of the paper is organized as follows: Section 2 reviews the work that has gone into using LDA and SVM with face recognition. Section 3 details the data set used for this study. Section 4 describes the proposed methodology for improving video-based face recognition rates. Section 5 presents the results, and Section 6 discusses those results. Section 7 presents conclusions and discusses future work.

sreekar.krishna@asu.edu; Ph: 480 326 6334

2. RELATED WORK

LDA was introduced to the domain of face recognition with the seminal paper by Etemad et al. [3] (Extension of this work includes [4] [5]). Here the authors generate a face class by adding Gaussian noise, and by taking the mirror image of the existing face images. Though this provided a means for generalizing face recognition around a limited set of face images, the resulting LDA subspace is not based on the types of variability that are encountered in video-sequence-based face recognition. Jonsson et al. [6] also reported on the performance of SVM for face recognition when used directly on the image set, and when used on an LDA-derived subspace, and concluded that there was no improvement in the recognition rate by using SVM. However, that work was also based on an artificial method of face class representation. In fact, most of the face recognition algorithms based on SVM [7] [8] [9] are based on standard face databases such as FERET, XM2VTS, the CMU PIE Database, AT&T, Oulu Physics Database, Yale Face Database, Yale B Database and MIT Database. None of these standard face databases include faces with the subtle variations in pose and illumination changes that are typical of face images extracted from video sequences. We have taken into consideration, in our data set, this important but ignored facet of the problems associated with face recognition and we show that our methodology performs much better than traditional LDA for video-based face recognition.

3. DATA SET

The proposed methodology for enhancing video-based face recognition was tested with images extracted from video sequences captured from 10 subjects. Each video sequence was captured while subjects looked into a camera, as both the camera and the subject moved. (This movement provided variations in background, pose and illumination - the latter of which were produced due to shadowing effects on the face). These video sequences were collected in an office environment, without any special lighting conditions.

0-7803-9332-5/05/$20.00 ©2005 IEEE

Subtle changes in pose angle were captured by asking the subjects to rotate their head slightly in arbitrary directions. The video sequences were collected over a time period of 1 minute at a capture rate of 10 frames per second, thus producing 600 images per subject. Fig.1 shows a sample set of 30 images extracted from one of these video sequences.

Fig. 1. Sample set of face images extracted from a video sequence

4. METHODOLOGY

4.1. Step 1: Generating the LDA face subspace

LDA is a subspace analysis method which projects high dimensional data into a lower dimensional space, the lower dimensions chosen such that data belonging to different classes are maximally separated. The LDA space is constructed from dominant eigen vectors of a separability matrix, S, which is the ratio of average between-class covariance and average within-class covariance. These eigen vectors represent the LDA space where the projected face images of different classes are best separated. The first step in the proposed method involves building such an LDA subspace from a compact set of images from each subject in the data set. From the 600 images of each subject, 10 frontal images were selected for each of the 10 subjects. (Fig.2 shows the set of 10 frontal images for one subject.) The resulting training set for the 10 subjects contained 100 frontal images. A high-dimensional feature vector was then generated for each of these frontal images by unwrapping them in a column-wise manner. These vectors were used to derive an N-dimensional LDA subspace that preserved the clustering (For our experiment, we found that a 6 dimensional LDA space was sufficient). Unfortunately, it is not possible to present this N-dimensional configuration in this paper, so Fig.3 shows a 2D LDA-based cluster plot for 3 subjects ($3 * 10$ frontal faces = 30 face images). To clearly define boundaries between the three clusters, Support Vector Machines were individually trained for each subject's 10 frontal images. The resulting boundaries are depicted in the Fig.3.

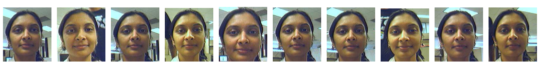

Fig. 2. 10 frontal images hand picked from the data set of a single subject

Fig.4 shows an LDA subspace generated for the same 3 subjects as in Fig.3, but derived from a subset of 10 randomly selected images per subject. There is no clear separation between the LDA face class of the 3 subjects. Contrast this with Fig.3, where the LDA subspace was derived from 10 frontal images of each subject, and where the separation between the clusters is clearly seen. In essence, training an LDA with frontal images alone provides

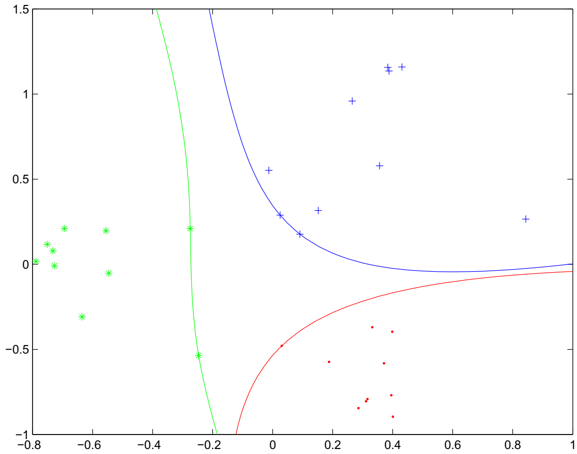

Fig. 3. LDA Subspace for 3 subjects trained with 10 frontal images per subject

better discrimination between classes. But deriving an LDA subspace from only frontal images does not produce a classification that reliably recognizes non-frontal face images, or images with other types of illumination, because the training data does not contain enough variations in these parameters. To allow the classifier to better distinguish between face images containing these types of variations, the regions of the LDA subspace representing each subject must be expanded around the frontal images as explained in Section 4.2.

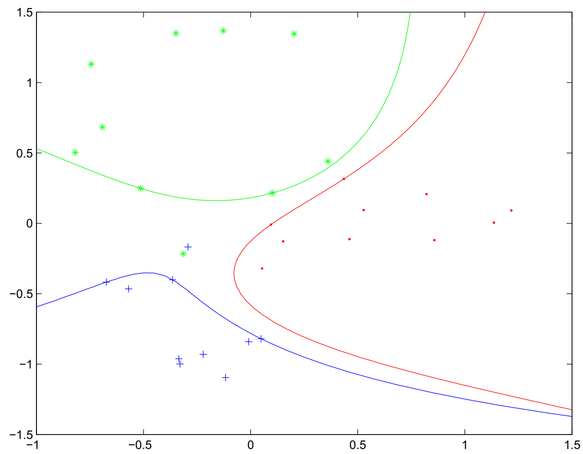

Fig. 4. LDA Subspace for 3 subjects trained with 10 randomly picked images per subject

4.2. Step 2: Expanding the face class in the LDA space

To expand the region for each subject, a set of face images can be randomly selected from that subject's video, a high-dimensional feature vector can be generated for each of these face images. These vectors can then be projected onto the lower dimensional LDA subspace, and associated with the corresponding frontal face clusters. Fig.5 shows the projection of 40 randomly selected images per subject on an LDA subspace that was derived from 10 frontal images per subject. It can be seen that the randomly selected images project into the LDA space near the original clusters, in contrast to Fig.6, which shows the projection of the same 40 images on an LDA subspace derived from 10 randomly selected images per subject.

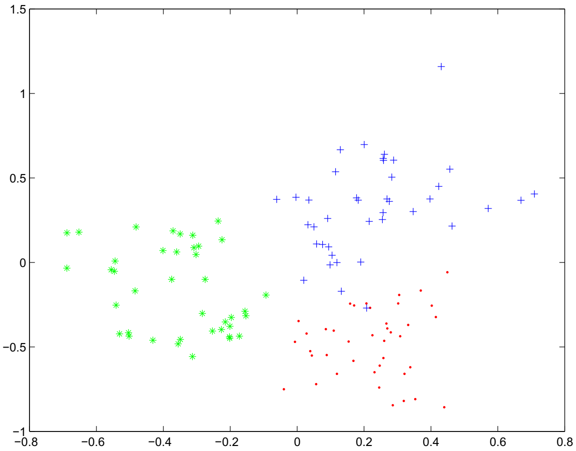

Fig. 5. LDA subspace trained with frontal images, with 40 random face images per subject projected into it

The expanded clusters need clearly defined boundaries, which can be provided by SVM. The kernel used for the SVM is a polynomial of degree 5. Fig.7 shows the resulting class boundaries formed with the 40 additional faces projected onto an LDA subspace derived from 10 frontal images per subject. For comparison, Fig.8 shows the SVM boundaries formed when the same 40 additional face images are projected onto an LDA subspace derived from 10 random images per subject.

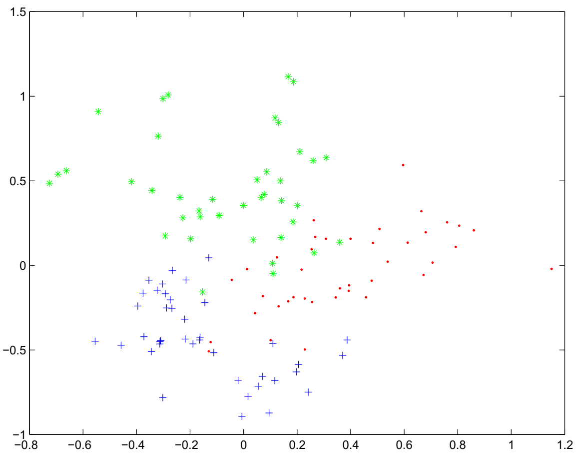

Fig. 6. LDA subspace trained on 10 random images per subject, with 40 random face images per subject projected into it

5. RESULTS

To test the recognition accuracy, test images were randomly selected, feature vectors were generated, and those feature vectors were then projected into the lower-dimensional LDA subspace to determine how close each was to the correct subject cluster. Two tests were conducted - one with 50 randomly selected images per subject, and the other with 100 randomly selected images per subject. Fig.9 shows the recognition rates when an MSE classifier was used with the subspace produced by (1) the proposed method of using frontal images, followed by random images, and (2) the traditional method of using frontal and random images indiscriminantly. Fig.10 shows the recognition rates with an SVM classifier on LDA subspace derived with (1) the proposed method of using frontal images followed by random images, and (2) the traditional method of using both the frontal and random images.

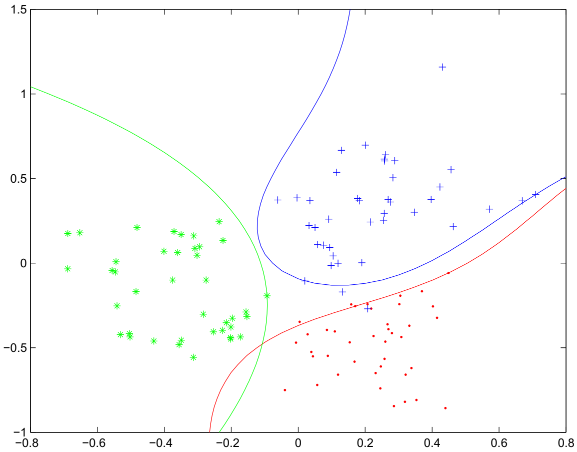

Fig. 7. SVM class boundaries for the new training method

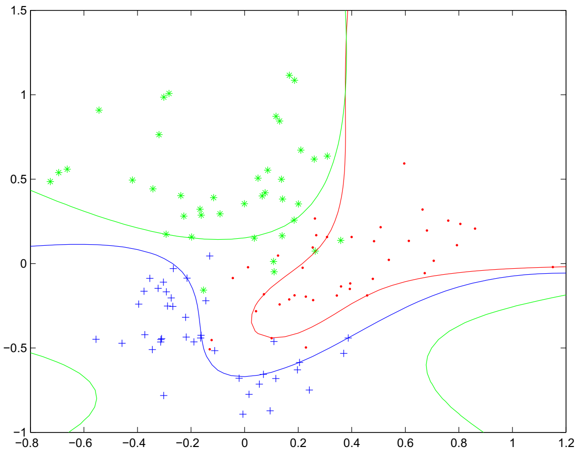

Fig. 8. SVM class boundaries for the regular LDA training procedure

6. DISCUSSION OF THE RESULTS

Visual inspection of Fig.9 and Fig.10 suggests that there is an improvement in the recognition rate when the LDA subspace is derived with the proposed method. It is also evident that there is an improvement in the recognition rate as the algorithm experiences more variations in the training images.

Fig.11 compares the performance of SVM and MSE classifiers for different size training image sets. (In both cases the LDA

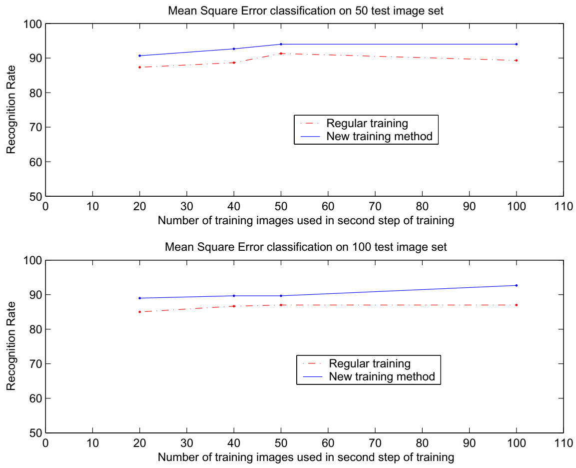

Fig. 9. Recognition rates with MSE classifier

subspace was derived using the proposed frontal-face method). It is evident that the SVM classifier performance is slightly better. When the LDA subspace was derived with only 20 images, SVM performed similar to MSE. An examination of the 20 frontal faces used in this case showed that there was little variation between the images, and thus the face clusters in the LDA subspace remained linearly separable.

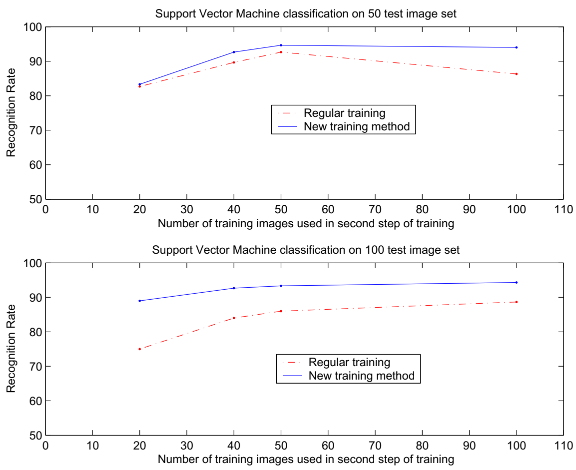

Fig. 10. Recognition rates with SVM classifier

7. CONCLUSION AND FUTURE WORK

This paper has presented a methodology to improve the recognition rate obtained from LDA-derived subspaces for video-based face recognition by using SVM. When the LDA subspace was derived from a relatively similar set of face images (all frontal) it provided better recognition performance than when the subspace was derived from a set of widely divergent (randomly selected) images. Work is in progress to develop algorithms that can learn the face class boundaries on the LDA subspace automatically as they experience varying images of a subject during the classification phase.

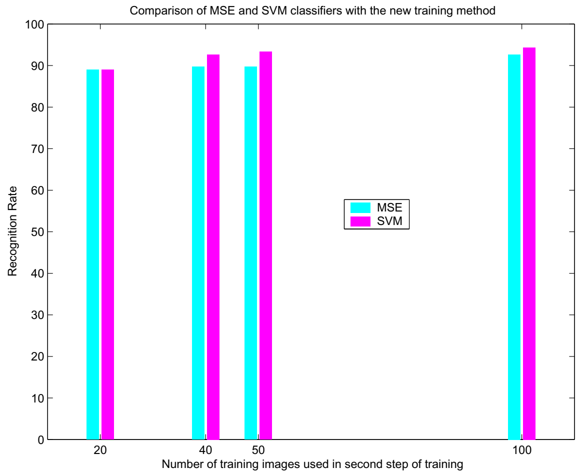

Fig. 11. Comparison of recognition rates with MSE and SVM classifiers

8. REFERENCES

[1] Price J.R.; Gee T.E.;, "Towards robust face recognition from video," *Applied Imagery Pattern Recognition Workshop, AIPR 2001 30th,*, pp. 94 – 100, 10-12 Oct. 2001.

[2] Little G.; Krishna S.; Black J.; Panchanathan S.;, "A methodology for evaluating robustness of face recognition algorithms with respect to variations in pose and illumination angle," *accepted in ICASSP 2005*, March 2005.

[3] Etemad K.; Chellappa R.;, "Discriminant analysis for recognition of human face images," *Journal of Optical Society of America*, vol. 14, pp. 1724–1733, 1997.

[4] Zhao W.; Chellappa R.; Phillips P.J.;, "Subspace linear discriminant analysis for face recognition," *IEEE Trans. on Image Processing*, 1999.

[5] Tae-Kyun Kim; Hyunwoo Kim; Wonjun Hwang; Seok Cheol Kee; Jong Ha Lee, "Component-based lda face," *British Machine Vision Conference 2002*, pp. 507–516, 2002.

[6] Jonsson K.; Kittler J.; Li Y.P.; Matas J.;, "Support vector machines for face authentication," *British Machine Vision Conference 1999*, pp. 543–553, 1999.

[7] Li Y.; Gong S.; Sherrah J.; Liddell H.;, "Support vector machine based multi-view face detection and recognition," *Journal of Image and Vision Computing*, vol. 22, pp. 413–427, 2004.

[8] Heisele B.; Ho P.; Poggio T.;, "Face recognition with support vector machines: Global versus component-based approach," *IEEE International Conference on Computer Vision*, pp. 688–694, 2001.

[9] Bicego M.; Lacono G.; Murino V.;, "Face recognition with multilevel b-splines and support vector machines," *Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*, pp. 17 – 24, 2003.