Multi Subspace Analysis with Supervised Separable Clusters for Classification of Smooth Nonlinear Manifolds

Hari C.V.*, Praveen Sankaran

Department of Electronics & Communication Engineering
National Institute of Technology, Calicut, Kerala, India

Abstract

The classification of smooth varying data that lie close to each other becomes a hard problem since there is very small class separation between them. Availability of close and adjacent poses result in an assumption of smooth manifold. Distinguishing between poses that lie close to each other becomes a complex problem. This paper proposes classification of smooth varying manifolds based on multiple subspaces with discriminant analysis. We divide the smooth manifold into multiple disjointed, locally linear, separable clusters. The problem of identifying which cluster to use, is solved by dividing the entire process into two steps. First step, or projection using the entire smooth manifold, identifies a rough region of interest. Second step, or second projection uses trained cluster(s) from this neighbourhood to identify closest pose.

Keywords: Smooth Manifolds, Discriminant Analysis, Pose Classification, Multiple Subspaces.

1. Introduction

Generally face pose manifolds are smooth varying provided there is no or small variation in other parameters like lighting. The data points of multiple adjacent classes may appear overlapped or may be too close to each other in such a case. The probability for misclassification using normal discriminant analysis is high due to this overlapping. Multiple subspaces based discriminant analysis method proposed in this work, projects this overlapped data to other subspaces aiming to provide better class separation and hence better overall accuracy. The objectives of this paper are, to develop a generalized two layer framework on Linear Discriminant Analysis (LDA), and to prove the robustness of the approach that may then be extended on to other feature based methods.

*Corresponding author
Email address: haricv@gmail.com (Hari C.V.)

Chutorian and Trivedi [1] gives a survey on different head pose identification methods, where they discuss the implicit difficulties in head pose estimation and present an organized survey describing the evolution of the field. They classified manifold embedding methods into linear subspaces, kernelized subspaces and non linear subspaces. Manifold embedding methods follow low dimensional manifolds modelled from continuous variations in head pose. New images can be embedded into this manifold and then template matching or any other statistical analysis can be used for pose estimation.

Sherrah *et.al* [2] found that the most prominent linear dimensionality reduction method, Principal Component Analysis (PCA) yield an identity invariant representation and the similarities can be vigorously calculated. Also, PCA [3] manifolds can be used for pose estimation by projecting the image into the PCA subspace. LDA explicitly attempts to model the difference between the classes of data and is proper for pattern classification if the number of training samples of each class are large. LDA maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability [4]. Srinivasan and Boyer [5] demonstrated the use of view based eigenspaces for head pose estimation from head and shoulders video sequences. The kernalized translations of PCA and LDA are also good for classification [6] and in modelling multi-view faces [7]. The non linear dimensionality reduction techniques include methods like Isometric feature mapping (Isomap) [8], Locally Linear Embedding (LLE) [9] and Laplacian Eigenmaps [10] can also be used for head pose estimation. While these techniques capture the geometry of the data points in the high-dimensional space, the disadvantage of this family of manifold learning techniques is the unavailability of a projection matrix to embed out-of-sample data points after the training phase [11]. Balasubramanian *et.al* [11] suggested biased manifold embedding which uses the pose angle information for the formulation of biased neighbourhood of each point in embedding and justified it with Isomap, LLE and Laplacian eigenmaps.

Foytik and Asari [12] proposed a two layer framework for head pose estimation based on LDA and Canonical Correlation Analysis (CCA). In this paper, in the first layer, the supervised linear method - LDA is used for coarse pose estimation. In the second layer, fine pose estimation is performed using region dependent pose regressive transforms.

Wang and Tang [13] proposed random subspaces for the projection in discriminant analysis. Here, small number of training images are randomly selected and this can be used for making random subspaces. Finally, fusion methodologies are used for combining these random subspaces. Uray *et.al* [14] proposed multiple subspace projections based on the linear separability of classes [15].

2. Linear Discriminant Analysis

PCA and LDA are two regularly used linear techniques for data classification and dimensionality reduction. The difference between the classes of data is obviously explained in LDA and it is not treated in PCA. The main objective of LDA is to perform dimensionality reduction by sustaining the class discriminatory informations as much as possible. LDA tries to find the directions along the classes that are well separated by taking into consideration of scatter within and between classes. LDA maximizes the ratio of between class variance to the within class variance in any particular data set thereby guaranteeing maximal separability. LDA does not change the location after transformation but only tries to provide more class separability and draw a decision region between the given classes [4]. LDA is proper for pattern classification if the number of training samples of each class are large.

The between class scatter matrix is defined as

$$S_B = \sum_{i=1}^{c} N_i(\mu_i - \mu)(\mu_i - \mu)^T \tag{1}$$

and the within class scatter matrix is defined as

$$S_W = \sum_{i=1}^{c} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T \tag{2}$$

where μ_i is mean image of class X_i and N_i is the number of samples in class X_i and c is the number of classes. If S_w is non singular, the optimal projection W_{opt} is chosen as the matrix with orthonormal columns which maximizes the ratio of the determinant of the between class scatter matrix of the projected samples to the determinant of the within class scatter matrix of the projected samples, where

$$W_{opt} = arg\ max \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1\ w_2\ \ldots\ w_m] \tag{3}$$

where $\{w_i \mid i = 1, 2, \ldots m\}$ is the set of generalized eigen vectors of S_B and S_W corresponding to the m largest eigen values $\{\lambda_i \mid i = 1, 2, \ldots m\}$, i.e.,

$$S_B w_i = \lambda_i S_W w_i, \quad i = 1, 2, \ldots m \tag{4}$$

The maximum value of m is $c - 1$ [16].

3. Multi Subspace Discriminant Analysis

The classification based on multi subspace discriminant analysis consists of two steps. Initially, the entire manifold is used for the projection for identifying a rough area/class. Then, another subspace(s) is/are used for next level of classification. This paper proposes two types of multi subspace discriminant analysis. The first one is based on classification by considering a predefined subspace by considering the overlapping images and the second one is the classification by predefined subspaces based on the neighbourhoods.

Figure 1 shows the projection of two overlapping classes in top two dimensions of the LDA subspace. The non linearity of the images can be observed

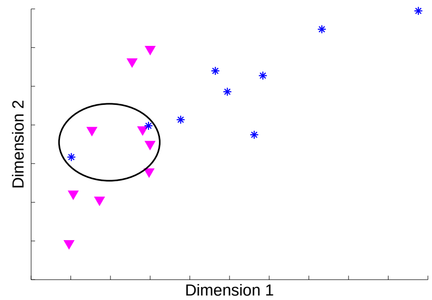

Figure 1: Projection of Two Classes.

from this projection and the points inside the circle may create an uncertainty in the classification. This images are projecting in to another subspace and in this space this points are well separated. Figure 2 shows the subspace projection in top two dimensions of LDA of this images and the images are well separated. This plane is giving good accuracy in classification than the first projection space for this points.

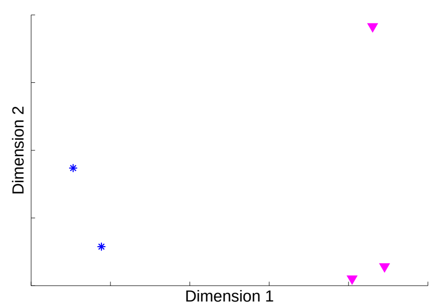

Figure 2: Subspace Projection of Overlapping Images.

3.1. Multi Subspace Discriminant Analysis by Considering Overlapping Images

The final and initial data points of adjacent classes of smooth varying manifold may be overlapped and reduces the classification accuracy and this overlapping area is considered as an area of uncertainty.

Overlapping data points of a smooth varying manifold in the area of uncertainty are projected to another space, and this subspace can be used for the decision making of overlapping test data points in a classification problem. Initially, different classes are created in such a way that each class consists of a specific number of different pose angles of multiple people. All classes together

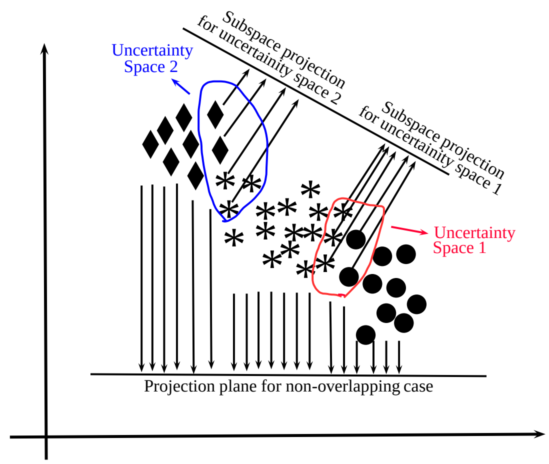

Figure 3: Multi Subspace LDA.

will cover the entire training dataset. These different classes are used for the training of the first layer projection features. Then, overlapping data points of adjacent classes forms the new classes for the second subspace. Consider ten images in a class for training. In this, 8^{th}, 9^{th} and 10^{th} of first class and 1^{st} and 2^{nd} of second class are taken as the overlapping data points. These data points form the first new class for the second subspace for this multi subspace discriminant analysis. Similarly, 8^{th}, 9^{th} and 10^{th} of second class and 1^{st} and 2^{nd} of third class form another class for the second subspace for this multi subspace discriminant analysis and so on. Then similar to the first case, LDA is applied on these new classes and projected to the resulting subspace. In this new plane, the new classes are well separated. This new plane is used for the projection of test images in the overlapping cases as shown in figure 3.

3.2. Multi Subspace Neighbourhood Discriminant Classifier

This approach also consists of two steps. Initially, similar to the previous case, the entire manifold is used in the first layer projection for identifying a rough area/class. In the second layer, data is projected onto a predefined subspace, formed using defined neighbourhoods, to identify the closest pose.

Initially, different classes are created in such a way that each class consists of a specific number of different pose angles of multiple people. All classes together will cover the entire training dataset. These different classes are used for the training of the first layer projection features. Test data/image is projected using these features and 'k' nearest neighbours are calculated. This calculation is based on the euclidean distance between the test image and training images.

For the second layer projection, different number of predefined projection spaces are created using subsets of the larger training set. These subsets are chosen such that, the entire dataset is covered, and each point may occur in multiple training subsets. After projecting the test data in the first layer, we look for the 'k' nearest neighbours of test data. We choose that predefined subspace, which is pointed to by the 'k' nearest neighbours.

Consider ten images in a class for training in the first layer and the value of 'k' is taken as 3 for the neighbourhood calculation. If the neighbours of a test image are 2^{nd}, 3^{rd} and 4^{th} poses, then the subspace containing 2^{nd}, 3^{rd} and 4^{th} pose images is used for the second layer projection. If a predefined subspace with these pose images are not existing, then the subspace with maximum number of images/poses of this neighbourhood is used for the second layer projection. For example, assume the predefined subspaces are $S_1 = [1, 2, 3, 4]$, $S_2 = [3, 4, 5, 6]$, $S_3 = [5, 6, 7, 8]$, $S_4 = [7, 8, 9, 10]$ and $S_5 = [9, 10, 1, 2]$, where S_1, S_2, S_3, S_4 and S_5 consists of the corresponding pose images of all classes. Then, if the neighbours of a test image are 2^{nd}, 3^{rd} and 4^{th} pose images, then S_1 subspace is used for the second layer projection because 2^{nd}, 3^{rd} and 4^{th} poses are in this subspace. If the neighbours are 4^{th}, 5^{th} and 9^{th} pose images, then S_2 subspace is used for the second layer projection because this subspace consists maximum number of poses.

Nearest neighbour algorithm with euclidean distance measurement is used for classification. Distance between the feature vectors of training and testing samples after LDA projection is computed. The image in the database with shortest euclidean distance with the test image is the winner. Leave-one-out testing strategy is followed here. One sample from the database from each class is kept out as the test image and the rest is used for training. In each iteration a different sample is kept out as the test image. Experiment is iterated till all the images in the database is used up as a test case. The dimensionality of the feature subspace is varied and accuracy is plotted against it. Accuracy is taken as the ratio between the number of correctly classified images to the total number of examples.

4. Results and Discussions

The proposed multi subspace LDA method is tested and evaluated in a smooth varying face pose manifold. The CUbiC FacePix database [17, 18] is used for the experimental analysis. For training, we created 18 different bins. Each class consists of 10 different pose angles of same person. For first subspace, in the positive angles, the first class consists of pose angles from 1 to 10, second class consists of 11 to 20 and so on. Similarly, in the negative side, first class consists of -1 to -10, second class consists of -11 to -20 and so on. This classes are using for the first layer of projection.

For second subspace, in the case of multi subspace discriminant analysis by considering overlapping images, the most probable overlap images form different classes. In this, 8^{th}, 9^{th} and 10^{th} of one class and 1^{st} and 2^{nd} of nearest class form the different classes. Here, there are 18 classes for the normal projection and 17 classes for multi subspace projection in this experiment. In the case of neighbourhood based discriminant analysis, we created four different type of arrangements of data to create predefined subspaces.

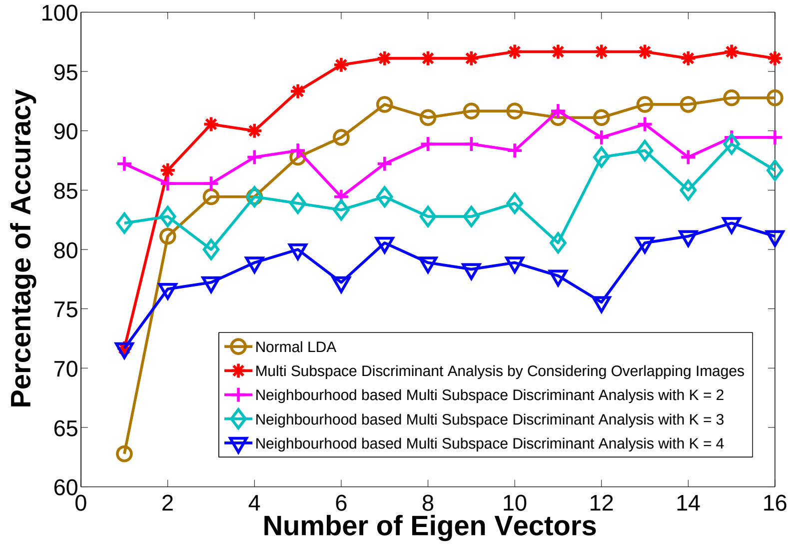

Figure 4: Accuracy Curve for First Multi Subspace Arrangement.

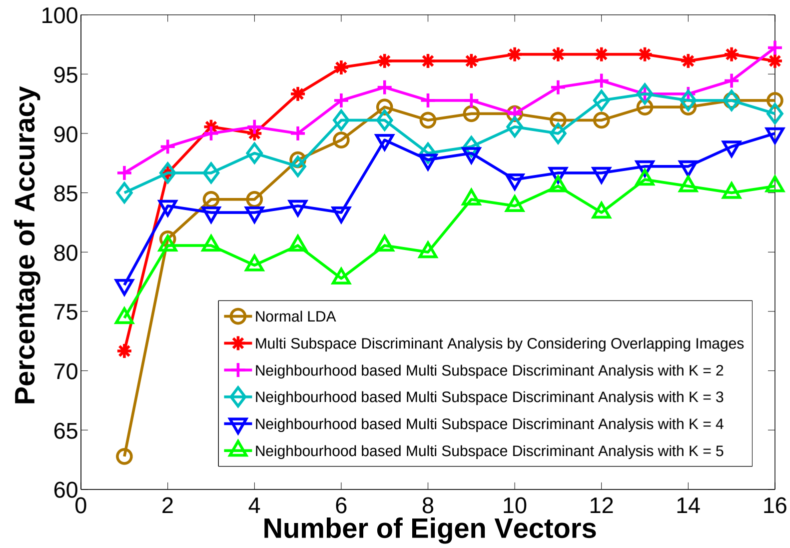

Figure 5: Accuracy Curve for Second Multi Subspace Arrangement.

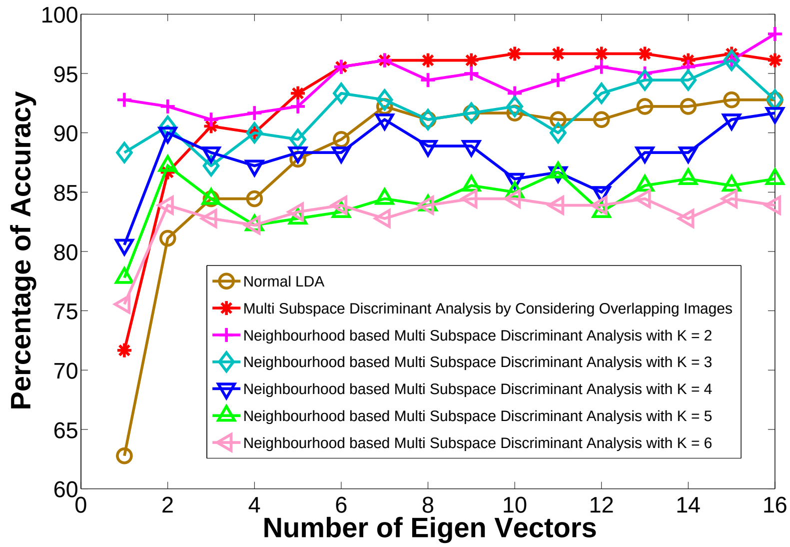

Figure 6: Accuracy Curve for Third Multi Subspace Arrangement.

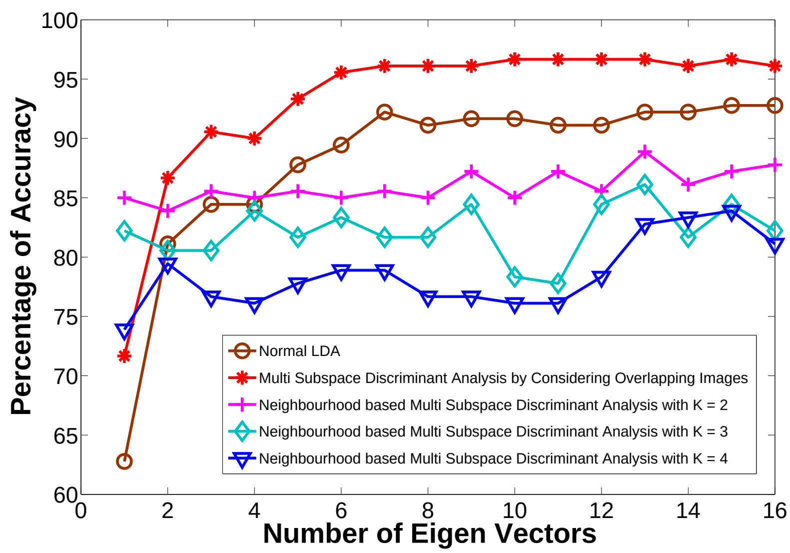

Figure 7: Accuracy Curve for Fourth Multi Subspace Arrangement.

First arrangement consists of five different spaces and they are $S_1 = [1, 2, 3, 4]$, $S_2 = [3, 4, 5, 6]$, $S_3 = [5, 6, 7, 8]$, $S_4 = [7, 8, 9, 10]$ & $S_5 = [9, 10, 1, 2]$. Second arrangement consists of three different spaces and they are $S_1 = [1, 2, 3, 4, 5]$, $S_2 = [5, 6, 7, 8, 9]$ & $S_3 = [8, 9, 10, 1, 2]$. Third one is also created with three spaces and they are $S_1 = [1, 2, 3, 4, 5, 6]$, $S_2 = [5, 6, 7, 8, 9, 10]$ & $S_3 = [8, 9, 10, 1, 2, 3]$. Last arrangement has four different spaces and they are $S_1 = [3, 4, 5, 6]$, $S_2 = [5, 6, 7, 8]$, $S_3 = [7, 8, 9, 10]$ & $S_4 = 9, 10, 1, 2]$. Here, $S_1 = [1, 2, 3, 4]$ from class 1 will form class 1 in the second layer. Similarly class 2 in second layer will be defined by the same set from class 2 of original data set.

The various results are as shown in figure 4, 5, 6 and 7 for first, second, third and fourth multi subspace arrangements respectively. The accuracy of multi subspace neighbourhood discriminant classifier is tested for different $'k'$ values. In all cases, multi subspace discriminant analysis by considering overlapping images method performs better than neighbourhood based multi subspace discriminant classifier.

5. Conclusion

This paper proposed two different types of multi subspace discriminant analysis for classification of smooth varying manifolds. The area of uncertainty in classification of a smooth manifold is projected to another subspace to obtain better class separability. This subspace is used for the classification of data points. The experiments performed on a standard face pose database demonstrates the robustness of our method. The approach is scalable to other projection/embedding technique to yield better accuracies.

References

[1] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation in computer vision: A survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 4, pp. 607–626, 2009.

[2] J. Sherrah, S. Gong, and E.-J. Ong, "Face distributions in similarity space under varying head pose," *Image and Vision Computing*, vol. 19, no. 12, pp. 807–819, 2001.

[3] S. J. McKenna and S. Gong, "Real-time face pose estimation," *Real-Time Imaging*, vol. 4, no. 5, pp. 333–347, 1998.

[4] S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis-a brief tutorial," *Institute for Signal and information Processing*, 1998.

[5] S. Srinivasan and K. L. Boyer, "Head pose estimation using view based eigenspaces," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, vol. 4. IEEE, 2002, pp. 302–305.

[6] J. Wu and M. M. Trivedi, "A two-stage head pose estimation framework and evaluation," *Pattern Recognition*, vol. 41, no. 3, pp. 1138–1158, 2008.

[7] L. Chen, L. Zhang, Y. Hu, M. Li, and H. Zhang, "Head pose estimation using fisher manifold learning." in *AMFG*, 2003, pp. 203–207.

[8] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[9] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[10] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[11] V. N. Balasubramanian, J. Ye, and S. Panchanathan, "Biased manifold embedding: A framework for person-independent head pose estimation," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–7.

[12] J. Foytik and V. K. Asari, "A two-layer framework for piecewise linear manifold-based head pose estimation," *International Journal of Computer Vision*, vol. 101, no. 2, pp. 270–287, 2013.

[13] X. Wang and X. Tang, "Using random subspace to combine multiple features for face recognition," in *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*. IEEE, 2004, pp. 284–289.

[14] M. Uray, P. M. Roth, and H. Bischof, "Efficient classification for large-scale problems by multiple lda subspaces." in *VISAPP (1)*, 2009, pp. 299–306.

[15] A. M. Martinez and M. Zhu, "Where are linear feature extraction methods applicable?" *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 12, pp. 1934–1944, 2005.

[16] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 711–720, 1997.

[17] J. A. Black Jr, M. Gargesha, K. Kahol, P. Kuchi, and S. Panchanathan, "A framework for performance evaluation of face recognition algorithms," in *Proc. SPIE*, vol. 4862, 2002, p. 164.

[18] D. Little, S. Krishna, J. Black, and S. Panchanathan, "A methodology for evaluating robustness of face recognition algorithms with respect to variations in pose angle and illumination angle," in *Proceedings of the IEEE international conference on acoustics, speech and signal processing*, 2005.