ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

Structure Preserving Embedding Based Face Recognition

Nishana S S, Subu Surendran
Dept. Computer Science & Engineering, SCT College of Engineering, Trivandrum,India
Associate Professor, Dept. Computer Science & Engineering, SCT College of Engineering, Trivandrum,India

Abstract— ***Face Recognition remains a challenging factor due to its high variability in expression, head rotation, aging ,etc. This paper proposes a technique called Structure Preserving Embedding (SPE) for face recognition. Structure Preserving Embedding is a non-linear dimensionality reduction technique which is effective in recognizing the intrinsic dimension of the data . Most machine learning and data mining techniques may not be effective for high-dimensional data. To handle this data efficiently, its dimensionality should be reduced. Dimension reduction is defined as the process of mapping high-dimensional data to a lower dimensional vector space. The SPE method performs dimensionality reduction on data for learning and classification purposes. The low dimensional data are then regressed using Support Vector Regression (SVR). SVR is used to learn a model that can predict the low dimensional embedding of the new datasets. Then a classification model using Support Vector Machine is generated to detect a face.***

Index Terms— ***Dimensionality Reduction, Semidefinite programming, Regression , Face recognition.***

I. INTRODUCTION

Face Recognition is a computer technology that automatically detects face and ignores everything else. Everyday actions are increasingly being handled electronically. This growth in electronic transactions results in great demand for fast and accurate user identification and authentication. Humans can identify each other by different ways. Likewise different technologies are available for the machines to identify a human. One of the most important person identification techniques is password known as Person Identification Number. The problem with this method is that, there is the chance of somebody to forget, loose, or even stolen the password by somebody else. For high security, this type of verification technique does not work properly. In such cases face recognition works efficiently.

Many different approaches are there to recognize a face. They are 1) Knowledge-based methods which uses a set of human coded rules to detect a face. 2) Feature invariant approaches which extract specific facial features such as eyes, nose, etc. 3) Template matching in which several standard patterns stored to describe the face as a whole or the facial features separately and 4) Appearance-based methods uses set of training images from which it learns the models that capture the representative variability of faces. Face detection is the process that discriminate face and non-face. The most important problem in face recognition is the 'curse of dimensionality' problem. High dimensional data leads to over fitting problems and increases computational complexity. This curse of dimensionality problem is reduced by dimensionality reduction techniques. So to efficiently detect faces in images, dimensionality reduction is an important and necessary operation. Dimensionality reduction is the process of mapping high-dimensional data to a lower-dimensional vector space. Reduction is done in a way so as to preserve all important features that is necessary for higher-level decision-making. For face detection and recognition, classical dimensionality reduction methods uses various techniques such as Eigen faces , Principal Component Analysis (PCA) [1], Independent Component Analysis [3], and Linear Discriminate Analysis [4]. But these methods fail to reveal the intrinsic dimension of a given dataset. Also they are inaccurate in detecting faces that exhibit variations in head pose, facial expression or illumination.

Here a new approach based on the graph embedding dimensionality reduction technique called Structure Preserving Embedding (SPE)[10] is used for face recognition. SPE is a non-linear dimensionality Reduction technique which maps high-dimensional observation data, that lies on a non-linear manifold, onto a single global coordinate system of lower dimensionality that preserves the global topological properties of the input data. Topology is preserved through a connectivity algorithm. SPE succeeds in recovering the underlying manifolds, whereas linear embedding methods, such as PCA[1] or Multi-Dimensional Scaling (MDS)[2], would map faraway data points to nearby points in the plane, creating distortions both in the local and global geometry. Many nonlinear dimensionality reduction algorithms, such as Locally Linear Embedding (LLE)[7], Maximum variance Unfolding (MVU)[5], and Minimum Volume Embedding (MVE)[11] begin by finding a sparse connectivity matrix A that describes local pair wise distances. Preserving distances does not explicitly preserve the structure of a data. LLE, MVU, and MVE, produce embeddings whose resulting connectivity no longer matches the inherent connectivity of the data. Here the distance between nodes that are connected are preserved but unconnected nodes are free to move apart. Structure Preserving Embedding (SPE) uses structure preserving constraints to solve this problem.

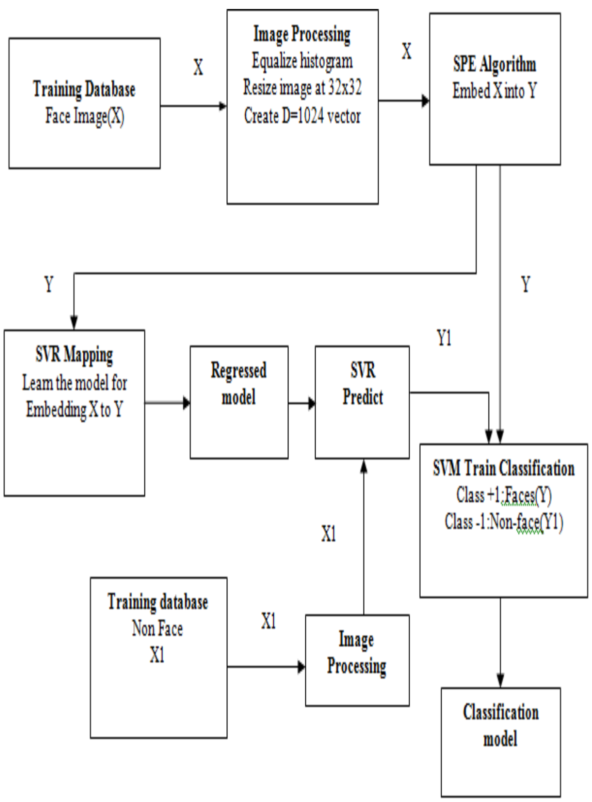

Fig 1 Face Recognition System.

A. Overview

This paper proposes an appearance –based method to detect faces in gray scale images subjected to a variety of condition such as expression, occlusion, aging and head rotation . Here a new technique called Structure Preserving Embedding (SPE) is used for face detection. SPE is a nonlinear dimensionality reduction technique which is used to extract features essential for face detection by reducing the image in high dimensional D space to lower dimension space d.

Once the facial data are transformed into a lower-dimensional space d. Support Vector Regression (SVR) is used to define a mapping from the input to the output space for these data. Thus, SVR computes the location of a point in d-space, given its location in the input dimension. A Support Vector Machine is then trained to classify the input as face or non-face. Once the detection is completed verification process is done. A matching process is then done to check whether the person is authenticated or not. Fig 1 gives the overall process in face recognition system.

Rest of the paper is organized as follows. Section 2 describes the Structure Preserving Embedding Algorithm. Section3 describes the Statistical Learning Methods which discuss about Support Vector Machine which is used to classify face and non-face, Support Vector Regression which is used to map new input to the low dimensional space. Section 4 provides experiments which show how the face recognition is performed and conclude in section 5.

II. STRUCTURE PRESERVING EMBEDDING

The SPE transformation algorithm is founded on the basis of simple geometric intuitions, where the input data is composed of N points Xi, $Xi \in R^D$, $i \in [1,N]$, each of dimensionality D, which were obtained by an underlying manifold. As an output, it produces N points Yi, $Yi \in R^d$, $i \in [1, N]$ where $d<<D$. It is a nonlinear dimensionality reduction technique that maps high dimensional data to low dimensional space. It's a graph embedding algorithm for embedding graphs in Euclidean space so that the embedding is low dimensional and preserves the global topological properties of the input graph. This algorithm can be effectively utilized to extract features necessary for face recognition. To preserve the topology, a connectivity algorithm is proposed which enforces certain linear constraints on the learned kernel matrix K. The K-nearest neighbor algorithm (KNN) greedily connects each node to the K neighbors to which the node has shortest distance, where K is an input parameter. The distance between a pair of points (i , j) with respect to a given positive semidefnite kernel matrix K, is Dij = Kii + Kjj - 2Kij .The matrix D constructed from elements Dij is a linear function of K. For each node, the distances to all other nodes to which it is not connected must be larger than the distance to the furthest connected neighbour of that node. The linear constraints applied to the kernel matrix, while using nearest neighbour algorithm is:

$$D_{ij} > (1 - A_{ij}) \max_m(A_{im}D_{im}) \qquad (2.1)$$

Where A is the connectivity matrix. For each node the connected distances are less than the unconnected distances , a greedy algorithm will find the exact same neighbours for that node, and thus the connectivity computed from K is exactly A.

To choose a unique Kernel from the admissible set in the convex hull generated by these linear constraints, an objective function which favours low dimensional embeddings is proposed.

The objective function $\max_{K \geq 0} tr(KA)$ subject to $tr(K) \leq 1$ recovers a low-rank version of spectral embedding. Also applying the linear constraint enforced by KNN, SPE preserves the global topology. This objective can be solved using Semi definite programming (SDP)[8] which is an optimization problem:

$$K' = \arg\max_{K \geq 0} tr(KA)$$

$$\text{subject to } D_{ij} > (1 - A_{ij}) \max_m(A_{im}D_{im}) \qquad (2.2)$$

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

To solve the optimization problem a software package called CSDP is used. The algorithm is a predictor corrector version of the primal dual barrier method [25] . CSDP is written in C for efficiency and portability. On systems with multiple processors and shared memory, CSDP can run in parallel. An optimized kernel K is obtained through this method.

Table1. Structure Preserving Embedding

Input	$(X_i)_{i=1}^{N}$
Step1	Find the affinity matrix using Linear kernel or RBF kernel
Step2	Find the connectivity matrix A through K Nearest Neighbour Algorithm. Initialize K=A.
Step3	Solve SDP K' = arg max $_{K \geq 0}$ $tr(KA)$ subject to $D_{ij} > (1 - A_{ij}) \max_m(A_{im}D_{im}$
Step4	Apply SVD to K and use the top eigenvectors as embedding coordinates

Once the low ranked optimized kernel is obtained, apply Singular Value Decomposition (SVD) to this kernel and use the top eigenvectors as embedding coordinates. Table 1 outlines the Structure Preserving Embedding Algorithm.

SVD theorem states that a matrix A can be represented as A= U S V^T ,where U is the Left singular vector and V is the Right singular vector . Calculating SVD consist of finding the eigen values and eigen vectors of AA^T and A^TA. The eigen vectors of AA^T makes up the column of U and that of A^TA makes the column V. The singular values in S are square root of eigen values from AA^T and A^TA.

III. EMBEDDING NEW DATA POINTS

SPE is an effective dimensionality reduction method. One of its limitations is the absence of a process for mapping between observations and the embedded space obtained from SPE in a particular dataset. Mapping is very necessary for dimensionality reduction in real-world problems, like face detection. Therefore, there must be some way to represent a new sample point within the manifold. For that purpose Support Vector Regression is used.

A. Support Vector Machine

A Support Vector Machine (SVM) is a technique that learns the decision surface through a process of discrimination, and has good generalization properties [26]. SVMs have been applied successfully to a number of applications, ranging from particle identification, face identification, text categorization, engine knock detection, bioinformatics and database marketing [29]. The SVM is based on *Structural Risk Minimization* theory [26, 30,31].

For given observations and interpretations, one finds the optimal approximation

$$f(x,a)=\alpha.\ g(x) + \beta \qquad (3.1)$$

where α represents the parameters of a learning machine, and g is a map from the original data space of X to a high dimensional feature space and β is the threshold [26]. If the interpretation only takes values +1 and -1, the learning problem is referred to as *Support Vector Classification*(SVC). Otherwise, if the domain contains continuous real values, it is *Support Vector Regression* (SVR).By introducing a kernel function

$$K(x ,y) = g(x).g(y) \qquad (3.2)$$

The SVC problem can be transformed to maximize

$$\begin{aligned} &W(a) = \sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i,j=1}^{l} a_i a_j y_i y_j K(x_i, x_j) \\ &Subject\ .to \sum_{i=1}^{l} a_i y_i = 0, 0 \leq a_i \leq C \\ &i = 1,2,.......l. \end{aligned} \qquad (3.3)$$

This gives a separating function

$$f(x) = \sum_{i=1}^{l} y_i a_i K(x_i, x_j) + \beta \qquad (3.4)$$

On the other hand , SVR problem can be solved by maximizing

$$\begin{aligned} &W(a^*, a) = -\frac{1}{2}\sum_{i,j=1}^{l}(a_i^* - a_i)(a_j^* - a_j)y_i y_j \\ &K(x_i, x_j) - \varepsilon\sum_{i=1}^{l}(a_i^* + a_i) + y_i\sum_{i=1}^{l}(a_i^* - a_i) \\ &Subject\ .to \sum_{i=1}^{l}(a_i^* - a_i) = 0 \\ &0 \leq a_i^*, a_i \leq C \end{aligned} \qquad (3.5)$$

This gives the solution

$$f(x) = \sum_{i=1}^{l}(a_i^* - a_i)K(x, x_i) + \beta \qquad (3.6)$$

B. Mapping using Support Vector Regression (SVR)

An analytical method based upon non-linear regression is proposed by Wang et al. [19] . The objective of the proposal was to determine a backward analytical mapping to move back from the embedded space to the original space. What SPE need is a technique to determine how the new data points be embedded in the reduced space. The proposal by Wang et al. is easily adaptable to forward mapping problem. Let the input data is composed of *N* samples $Xi, Xi \in R^D$, $i \in [1,N]$, each of dimensionality *D.* The embedding of *Xi* by SPE is $Y_i \in R^d$ where d is he dimension of embedded space which is less than D. In SVR the input X is first mapped into a d dimensional feature space using nonlinear regression and then a linear model is constructed in the feature space. Each point Y_i in d dimensional space is regressed using

$$y_i = f_i(X) \qquad (3.7)$$

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

The following steps describe how $f_i(X)$ is generated.

Step1: Apply SPE to the training facial images X and find the low dimensional embedding co-ordinates Y as discussed in section 2.

Step2: This mapping is then regressed using Support Vector Regression.

$$y_i = f(X) = \sum_{j=1}^{M} \alpha_{ij} g_j(X) + \beta \quad (3.8)$$

where $g_j(X)$ is the kernel function and m is the number of training samples. For mapping new data points, the proper kernel function and its parameters must be determined.

Experiments shows that Radial Biasis Function (RBF) kernel was the best for regressing the mapping models. It is defined as:

$$g_j(X) = K(X,X_j) = exp(-\| X - X_j \|^2 / 2\rho^2) \quad (3.9)$$

Where X_j is the number of training samples and $K(X,X_j)$ is the kernel function.The proper form for $f_i(X)$ can be found using training sets with a sufficient number of facial images. In order to obtain useful information for function and parameter selection, a cost function must again be defined to evaluate regression performance. The entire database is divided into two different sets, one for training (s_1) which is used for estimating the optimal parameters, and one to evaluate the performance (s_2). The squared-error function is defined as:

$$E(F) = \sum_{j=1}^{N2} \| Y_j - F(X_j) \| \quad (3.10)$$

where $j=1,2,..N2$ and $N2$ is the number of samples in the test set. The optimal kernel parameters are determined using this cost function.

Step3: After obtaining the proper kernel function and parameters, α_{ij} and β in equation (3.8) are obtained by Support Vector Regression method. This gives the function $f_i(X)$ that determines the forward mapping from D space to low dimensional d space. This mapping function enables new images to be mapped into an embedded space using SVR.

IV. EXPERIMENTS AND RESULTS

This section describes the proposed face detection system based on SPE embeddings and Support Vector Machines and evaluates the performance of the method in terms of detection accuracy.

A. Experimental Methodology

First step to train the proposed face detector is to collect facial images from various face databases available on the internet. The facial database should reflect variation in head pose, expression, illumination and occlusion. The following subsection describes the various databases used for this work.

1. Face Databases

ORL Face Database [27]

The ORL Database of Faces , contains a set of face images taken between April 1992 and April 1994. There are ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open /closed eyes, smiling / not smiling) and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position. Fig 1 shows sample images from ORL face database.

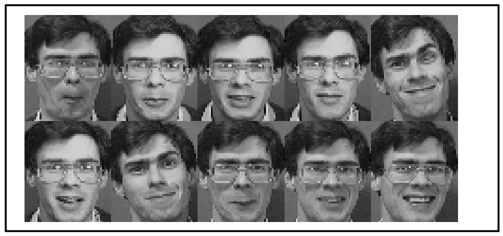

Fig 1 sample images from ORL face database

UMIST Database [28]

The UMIST Face Database consists of 564 images of 20 people. Each covers a range of poses from profile to frontal views. Subjects cover a range of race/sex/appearance. This database offers facial images with large variations in head rotation. Fig 2 shows the sample images from UMIST database..

Fig 2 sample images from UMIST database

2. System Description

The facial images obtained from the above described sources should be cropped and divided into two databases, DB1 and DB2. DB1 reflect facial images with illumination, expression and occlusion features. DB2 consists of facial images with varying head rotation. A preprocessing stage was first applied to normalize and resize all images. Then all the training images in the D dimensional space should be transformed to lower dimensional space d using the SPE algorithm. This mapping was then regressed using Support Vector Regression. Then a classifier should be trained to distinguish between face and non-face. These steps are repeated for the two databases DB1 and DB2. Thus two regression and classification models are created that should embed and classify new images.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

Step1: Initial step is a pre-processing stage. Here the images need to normalize and resize. Illumination compensation was performed using histogram equalization.

Step2: Then all the training input images X are embedded to low dimensional space using SPE. This created a set of embedded points (*Y*) in *d*-space.

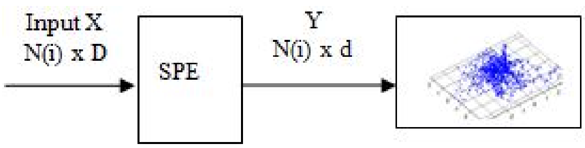

Fig 3 Low dimensional embedding

Step3: Then this embedding is regressed using SVR. SVR creates a model which can predict the low dimensional representation of the newly introduced images. This model is used to find the lower dimensional representation of the non-face images. The non-face images used for training are taken from the Viola-Jones database.

Step4: Train a classifier to distinguish between face and non-face.

Step5: Final step is the verification process. Face detector is mostly used for security purpose to identify authenticated persons. If a face is detected , then it should determine whether the person is authenticated or not. For that a matching process is done with the detected face and the already trained faces. If a matching is found, the person is authenticated and his information is displayed. Fig 4 gives the details of how the system works.

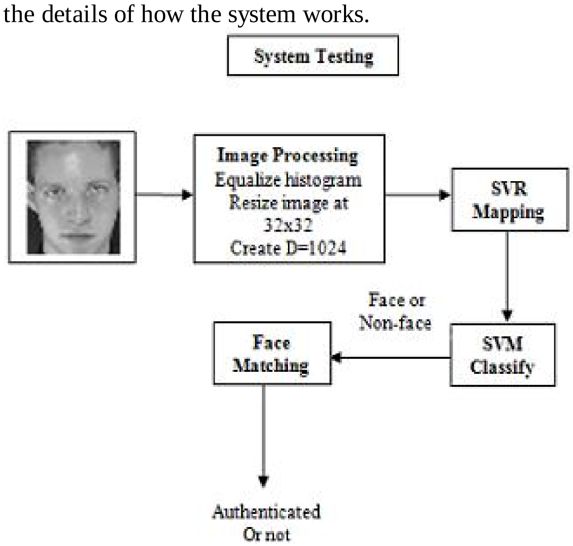

Fig 4 System Testing

ORL Face database

SPE based face recognition technique is tested against the CBCL faces database. It gives a Detection rate about 96% with a lower dimension *d*=10. An analysis is done against the Principal Component Analysis (PCA) based face recognition and found that the proposed SPE based face detection technique gives much better result than PCA. The various detection rates for different dimension for the two techniques are shown in table 2.

Table 2 Detection Rate for various dimensions for SPE and PCA Face Detection for DB1

Dimension(d)	1	2	6	9	10
Detection Rate (SPE)	53%	72%	76%	92%	96%
Detection Rate (PCA)	32%	43%	52%	75%	82%

UMIST Database

For the database2 (DB2), the head pose database, the proposed SPE method gives a detection rate of 94% with a lower dimension *d*=25. The SPE based face recognition is compared with PCA method. The obtained result is shown in table 3.

Table 3 Detection Rate for various dimensions for SPE and PCA Face Detection for DB2

Dimension(d)	1	5	10	15	25
Detection Rate (SPE)	55%	72%	75%	86%	94%
Detection Rate (PCA)	34%	40%	48%	53%	69%

Receiver Operator Curve (ROC)

The system can exhibit two types of errors. One is if a non-face image is given but the system recognizes it as face. This is called False Positive (FP). The other is that a face is given but the system identifies it as non-face which is called False Negative. It has two correct answers, if a face is accepted it is called True Positive(TP) and if a non-face is rejected, it is True Negative (TN). So a system performance is measured by using False Positive Rate (FPR), False Negative Rate (FNR), True Positive Rate (TPR) or True Negative Rate (TNR).

False Positive Rate (FPR) = FP/ (FP + TN)
False Negative Rate (FNR) = FN/ (FN + TP)
True Positive Rate (TPR) = TP/ (TP + FN)
True Negative Rate (TNR) =TN/ (TN + FP)

Receiver Operator Curve is a graphical plot which illustrates the performance of a binary classifier. It is created by plotting the True Positive Rate (TPR) vs. False Positive Rate (FPR). The ROC curve for the proposed SPE face

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

recognition technique and the PCA for the two databases DB1 and DB2 are shown in fig 5 and 6.

In this section the proposed technique is tested against different face and non-face images and a comparison is done against the PCA algorithm. The results shows that the proposed SPE face detection method outperforms other face detection methods.

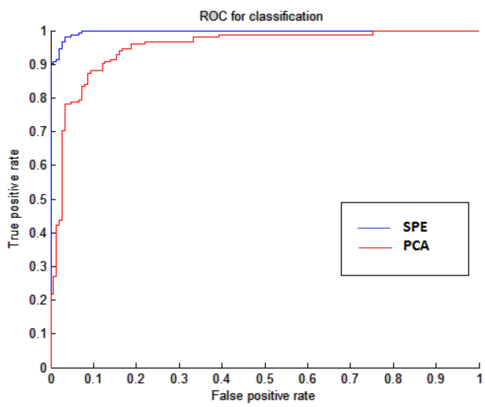

Fig 5 ROC for SPE(——) and PCA(——) algorithm for DB1

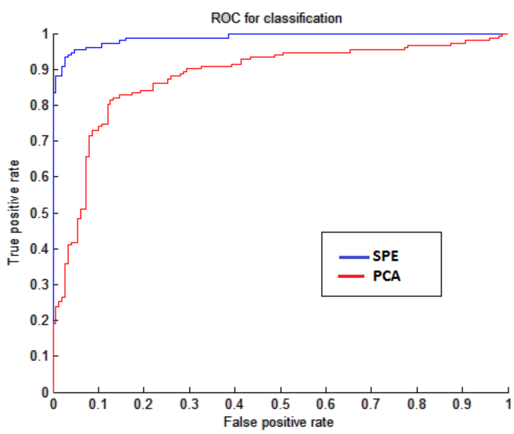

Fig 6 ROC for SPE(——) and PCA(——) algorithm for DB2

V. CONCLUSION

Face detection is a challenging factor because of its variability in expression, head pose, occlusion and aging. Several techniques are there to recognize a face. The most important factor that is to be considered while developing an application program is the dimensionality of the data. High dimensional data leads to over fitting problems and increases computational complexity. So the dimensionality should be reduced for the efficient processing of the data's. Also it reduces the storage space needed. This paper proposes an appearance based method which uses dimensionality reduction technique called Structure Preserving Embedding to detect faces. This new approach uses Semi Definite Programming to obtain an optimized Kernel which provides good result. From the result obtained by testing the proposed technique against different datasets, it is clear that SPE is successful in recognizing faces in varying conditions and provides good results compared to other techniques.

REFERENCES

[1] I.T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, New-York, 1986.

[2] T. Cox and M. Cox. Multidimensional scaling. Chapman & Hall, London, UK, 1994.

[3] Aapo Hyvärinen and Erkki Oja, "Independent Component Analysis: Algorithms and Applications",2000.

[4] S. Balakrishnama, A. Ganapathiraju," Linear Discriminate Analysis - A Brief Tutorial".

[5] Kilian Q. Weinberger and Lawrence K. Saul," An Introduction to Nonlinear Dimensionality Reduction by Maximum Variance Unfolding". *American Association for Artificial Intelligence,2006.*

[6] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Computation, 15(6), 2002.

[7] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. Science, 290(5500), 2000.

[8] L. Vandenberghe and S. Boyd. Semidefinite programming. SIAM Review, 38(1):49–95, 1996.

[9] K. Q. Weinberger, F. Sha, and L. K. Saul. Learning a kernel matrix for nonlinear dimensionality reduction. In Proceedings of the Twenty First International Conference on Machine Learning (ICML-04), pages 839–846, Banff, Canada, 2004.

[10] Blake Shaw &Tony Jebara,(2009).Structure Preserving Embedding Proceedings of the 26 th International Conference on Machine Learning.

[11] Shaw, B., & Jebara, T. (2007). Minimum volume embedding. Proc. of the 11th International Conference on Articial Intelligence and Statistics.

[12] Burer, S., & Monteiro, R. D. C. (2003). A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Mathematical Programming (series B), 95(2), 329{357}.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013
[13] M. Bartlett and T.J. Sejnowski, "Independent components of face images: A representation for face recognition", *Proc. of the 4th Annual Joint Symposium on Neural Computation*, 1997.

[14] Y. Li, S. Gong, and H. Liddell, "Support vector regression and classification based multi-view face detection and recognition", *Proc. Int'l Workshop on Automatic Face- and Gesture- Recognition*, 2000.

[15] Y. Li, S. Gong, J. Sherrah, and H. Liddell, "Multi-view Face Detection Using Support Vector Machines and Eigenspace Modelling", *Proc. Int'l Conf. on Knowledge-based Intelligent Eng. Sys. and Allied Tech.*, pp. 241-245, 2000

[16] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: An Application to Face Detection", *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, pp. 130-136, 1997.

[17] A. Yuille, P. Hallinan, and D. Cohen, "Feature Extraction from Faces Using Deformable Templates", *Int'l J Computer Vision*, 8(2):99-111, 1992.

[18] M.-H. Yang, D. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(1):34-58, 2002.

[19] J. Wang, Z. Changshui, and K. Zhongbao, "An Analytical Mapping for LLE and Its Application in Multi-Pose Face Synthesis", *14th British Machine Vision Conf.*, 2003.

[20] E. Vazquez and E. Walter, "Multi-output support vector regression", *13th IFAC Symposium on System Identification*, pp.1820-1825, 2003

[21] M.A. Turk and A.P. Pentland, "Face Recognition using Eigenfaces", *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, pp. 586-591, 1991.

[22] K.-K. Sung and T. Poggio, "Example-Based Learning for View- Based Human Face Detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(1):39-51, 1998.

[23] B. Scholkopf, "Statistic al Learning and Kernel Methods", *Technical Report MSRTR- 2000-23*, Microsoft Research Limited, 2000.

[24] Y. Li, S. Gong, and H. Liddell, "Support vector regression and classification based multi-view face detection and recognition", *Proc. Int'l Workshop on Automatic Face- and Gesture- Recognition*, 2000.

[25] Anders Forsgren and Phililp E. Gill, "Primal-Dual Interior Methods For NonConvex Non linear Programming" *Society for Industrial and Applied Mathematics*, Vol. 8, No. 4, pp. 1132{1152}, November 1998.

[26] V. Vapnik, *The nature of statistical learning theory*, 2nd Edition, Springer-Verlag,1997.

[27]http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces.tar.Z

[28]http://www.sheffield.ac.uk/eee/research/iel/research/face

[29] C. Bennett and C. Campbell, "Support Vector Machines: Hype or Hallelujah?", *SIGKDD Explorations*, 2(2):1-13, 2000.

[30] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, pages 1–47, 1998.

[31] M. Hearst, B. Scholkopf, S. Dumais, E. Osuna, and J. Platt. Trends and controversies –support vector machines. *IEEE Intelligent Systems*, 1998.

.