Journal of Computational Linguistics
Volume 1, Issue 1, 2011, pp-01-04
Available online at: http://www.bioinfo.in/contents.php?id=46

Proximal SVM for Face Recognition

T. Hanumantha Reddy[a], K. Karibasappa[a], A. Damodaram[b]

[a]vijayanagar Engineering College, Bellary, Karnataka, thrbly@yahoo.com

[b]JNTU College of Engineering, Hyderabad, Andhra Pradesh, damodarama@jntu.ac.in

Abstract: In this paper, we have proposed a Proximal support Vector Machine based Face recognition using subspace analysis. The Proximal support vector Machine (SVM) aims at generating two parallel planes such that the each plane is closest to one of the two face classes to be classified and the two planes are as far apart as possible. The proximal support vector machine uses two parallel hyper planes for face classification while standard Support vector machine with Bayesian similarity is used for Hierarchical agglomerative clustering. The simple pair wise binary tree strategy is used to classify the multi class patterns using proximal Support Vector Machine. The Proximal support vector machine is computationally more efficient than the standard Support Vector Machine. The Experimental results confirm that our method outperforms some of the existing methods.

Key words: Face recognition, proximal support vector machine, Hierarchical Agglomerative clustering, Binary tree.

1 INTRODUCTION

Face recognition is an important application of pattern analysis due its many applications in real world. Even today the face recognition task is challenging because of the dynamism in the problem formulation. The different methods for face recognition are described in the paper [1]. The parameters which affect the accuracy and efficiency are size, expressions, varied lighting conditions, occlusions, and cosmetics. Nonetheless, developing an ideal face recognition system that replaces the human is impossible but the steady progress in the recognition research has made an attempt to develop a system that is very close to the human system.

In the image analysis literature, face/object recognition can be roughly divided into three cases.

(i) Appearance based approaches achieves visual representation based appearance properties of individual object or object categories [2].

(ii) Structure based approaches were introduced in the last 10 years to account for pictorial object representation and recognition.

(iii) Image parsing approaches were introduced in the last 5 years for object segmentation, detection and recognition. In this paper, we have proposed a computational efficient proximal support vector machine for face recognition. The rest of the paper is organized as follows. Section 2 describes the related work and background. Section 3 outlines the proposed approach. Section 4 shows the Experimental results. Conclusions are drawn in Section 5.

2 RELATED WORK AND BACKGROUND

The standard SVM is a statistical learning theory that can be used for pattern recognition task is described in the papers [3], [4]. The hierarchical component based support vector machines is presented for face detection and recognition [5]. The binary tree based standard SVM with Gaussian kernel is used for face recognition [6]. The one-against-all (OAA) Fuzzy support vector machine is used for multi class text classification [7].

The frame work of Eigen face recognition that learns and recognizes the human face in an unsupervised manner is proposed in the paper [8]. The combined adaptive clustering and multilevel subspace analysis using Bayesian-SVM method is used to enhance the face recognition performance [9]. The Proximal Support Vector Machine (PSVM) classifies the data points based on proximity to one of the two parallel planes that are pushed as far as possible contrast to the standard SVM that uses one hyper plane to separate the data.[10].

A relatively new technique to SVM classification is presented in the paper [11] where each of two dataset is proximal to one of two distinct planes that are not parallel to each other. The novel support vector multi class classifier is proposed that uses array of pair wise coupling classifier to improve the classification accuracy [12]. Bruce walter et al [13] have proposed an efficient locally-ordered algorithm for agglomerative clustering.

3 PROPOSED APPROACH

The Architecture of the proposed system is shown in the Fig 1. The proposed approach uses method presented by Mangasarian [11] for face recognition. However, we have used Standard support vector machine for Hierarchical clustering. Firstly, the principal component analysis (PCA) or Karhunen-Loeve expansion is performed to extract the Eigen vectors and Eigen faces of the training images. These Eigen vectors are used by the Hierarchical SVM clustering to construct the clusters based on similarity. After clustering, c(c-1)/2 PSVMs are used to classify the incoming image using binary tree strategy. The projection coefficient of the incoming image goes through the binary tree from bottom to upwards until winner is declared. After several comparisons, a unique class label appears at the top of the tree and this unique label is the exact match of the incoming image. The rest of the section is organized as follows. Section 3.1 describes standard support vector machine, section 3.2 explains the generalized **E**igen value proximal support vector machine, section 3.3 depicts the Hierarchical agglomerative clustering, and section 3.4 explains the face recognition algorithm.

3.1 Support Vector Machine (SVM)

The main idea of SVM comes from the mapping of input space to high dimensional space and designing optimal hyper plane in terms of margin [3]. Let (x_1,y_1) (x_2,y_2) ….. (x_k,y_k) Є R^N and y_i Є { -1,+1} be the k^{th} training samples in the input space, where y_i indicates the class membership of x_i.

Let ф be non-linear mapping $x \rightarrow ф(x)$.

The optimal hyper plane is defined as below.

$$w_0 \Phi(x) + b_0 = 0 \quad (1)$$

It is proved that the vector w_0 is linear combination of weight vector which are vectors x_i that satisfy

$$y_i(w_0.\Phi(x_i) + b_o = 1 \quad (2)$$

$$w_0 = \sum_{\sup portvectors} y_i \alpha \Phi(x_i) \quad (3)$$

The linear decision function,

$$f(x) = sign(\sum_{\sup portvectors} y_i \alpha \phi(x_i).\phi(x) + b_o) \quad (4)$$

To make the data more separable, the dot product is replaced by a RBF Gaussian kernel.

$$k(x_i, x_j) = \phi(x_i).\phi(x) \quad (5)$$

The i^{th} row and j^{th} column entry in a RBF Gaussian kernel matrix is evaluated as,

$$k(x_i, x_j) = \exp(-\frac{\| x_i - x_j \|^2}{2\sigma^2}) \quad (6)$$

Where, $\sigma > 0$

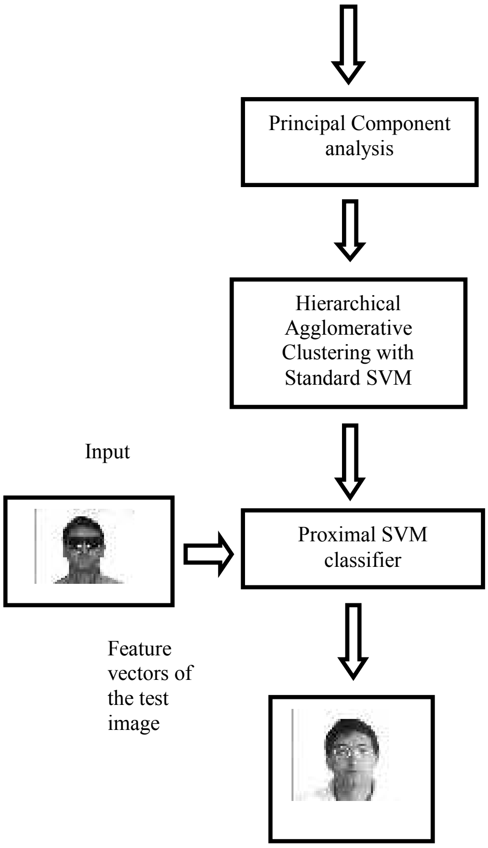

Fig 1 shows the architecture of the proposed system

3.2 Generalized Eigenvalue Proximal Support Vector Machine Classifier (GEPSVM)

In this section, we briefly describe the mathematical background of Proximal Support vector machine. To classify m points in the n-dimensional space R^n, represented by m_1 x n matrix A belongs to class 1 and the matrix m_2 x n B belongs to class 2, with $m_1 + m_2 = m$. For this problem of classification, two parallel planes are generated such that each plane is closest to one of two datasets to be classified by the machine. The classifying plane is mid way between these two parallel planes. Fig 2 shows the pictorial representation of data classification using proximal support vector machine.

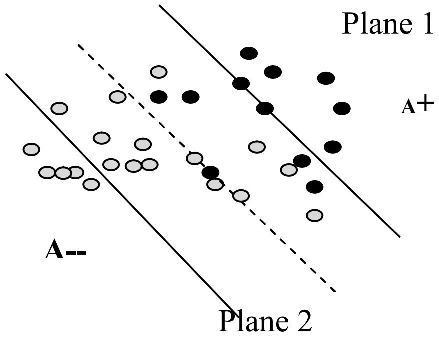

Fig 2 shows the proximal parallel planes

The two parallel planes classify the points based on their proximity: $x^1 w = \gamma + 1$ and $x^1 w = \gamma - 1$ that are determined by the quadratic function $w \in R^n, \gamma \in R$ for some v >0 as follows.

$$\min_{(w,\gamma) \in R^n} \frac{v}{2}(\| e - (Aw - e\gamma) \|^2 + \| e + (Bw - e\gamma) \|^2) + \frac{1}{2} \| \begin{bmatrix} w \\ \gamma \end{bmatrix} \|^2 \quad (7)$$

By the parallel condition on proximal planes and require each plane be closest to the one of face class and far away from second class. Thus two planes in n-dimensional points are formulated as follows.

$$x^1 w^1 - \gamma^1 = 0 \quad and \quad x^2 w^2 - \gamma^2 = 0 \quad (8)$$

To obtain the plane closest to the class 1 and far away from class 2, the following optimization formulation can be used.

$$\min_{(w,\gamma) \neq 0} \frac{\| Aw - e\gamma \|^2 / \| \begin{bmatrix} w \\ \gamma \end{bmatrix} \|^2}{\| Bw - e\gamma \|^2 / \| \begin{bmatrix} w \\ \gamma \end{bmatrix} \|^2} \quad (9)$$

3.3 Hierarchical clustering

The major problem with pair wise strategy **is** to train c(c-1)/2 proximal support vector machines for classification. The efficiency of the recognition depends on the parallel planes of the proximal Support vector machine. Therefore, it is reasonable to train the PSVM near the plane which is closest to the class. To achieve this and to increase the recognition accuracy, we first partition the gallery images into clusters with each cluster containing only similar images [9].

The basic process of Hierarchical clustering is defined as below.

Algorithm: Hierarchical agglomerative clustering
Begin
Step1: Initialize the one cluster per face class
Step 2: Find the nearest pair of clusters based on similarity measure, and then merge them into one cluster.
Step 3: Repeat step 2 until even number of clusters are obtained.
End //Hierarchical agglomerative clustering//

In the above algorithm, the important condition is to determine the stop rule of the clustering process. In our method the even number of cluster is used to stop the algorithm. The motivation to use even number of cluster is to use each cluster as leaf node of the binary tree. Another key design parameter is to select the similarity measure between the clusters. We have used Bayesian-SVM similarity measure between the face images.
To find the similarity between the face clusters, we evaluated the intra personal variation set $\{\Delta_I \mid \Delta_I \in \Omega_I\}$ and the extra personal variation set $\{\Delta_E \mid \Delta_E \in \Omega_E\}$. The Eigen value matrix Λ_I and Eigen vector matrix V_I of the intra personal and extra personal variations are computed.

$$\Delta'_I = \Lambda_I^{-1/2} V_I^T \Delta_I \qquad (10)$$

$$\Delta'_E = \Lambda_I^{-1/2} V_I^T \Delta_E \qquad (11)$$

For pair of images i and j in the database, we first compute the image difference $\Delta^{ij}{}_E$ and then project it in an Intra personal subspace.

$$\Delta^{ij}{}_E = \Lambda_I^{-1/2} V_I^T \Delta^{ij}{}_E \qquad (12)$$

The similarity measure is defined with the decision function of the standard support vector machine, as defined in the equation (4).

$$S_{ij} = f(\Delta^{ij}{}_E) \qquad (13)$$

The away from the hyper plane indicates that the image difference is closer to the intrapersonal variation.

3.4 Proximal SVM based Face Recognition Algorithm

The Proximal SVM based face recognition algorithm has two steps, as shown in the Fig 3. Initially, Projection coefficients are evaluated for the input face image. The projections coefficients are used as feature vectors of the probe image.
We have used bottom-up binary tree for face classification using PSVM. By comparison between each pair, one class is chosen as winner of the two current classes. The projection coefficient of the incoming image goes through the binary tree from bottom to upwards until winner is declared. After several comparisons, a unique class label appears at the top of the tree.

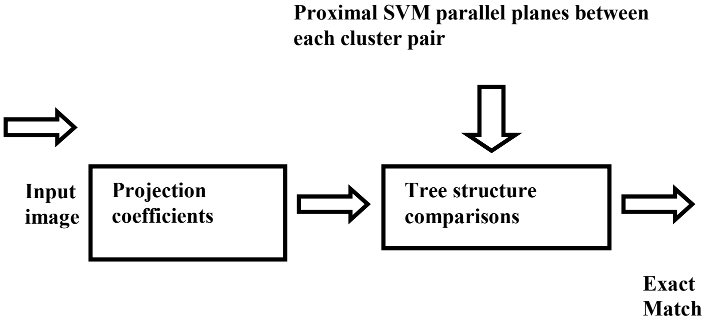

Fig 3 shows the flow line of face recognition

4 EXPERIMENTAL RESULTS

The proposed face recognition system is evaluated on two databases and the face image database details are shown in the Table 1. It also describes additional information such as number of images, color, and size. The performance of the face recognition algorithm is compared with other algorithms to show its success. In this study, we have used AR and Essex database.

Table 1 Face database used for face recognition

Name	Number	Color	Size
AR database [14]	400	Grey Scale	92 x 112
Essex database[15]	7900	RGB	180 x 200

4.1 Experimental study on AR database

The first experiment is performed on the AR face database, which contains 126 persons face images. There are 26 images for each person captured at different times. There are variations in expressions and facial details such as glasses or no glasses. All the images are frontal up-right with homogeneous background. In this study, we have used 1638 face samples, 13images per person for 126 people, as training set and the remaining 1638 face samples as testing set. During training, calculate the Eigen vectors and use the Hierarchical agglomerative clustering to reduce the number of clusters. Later, train the proximal support vector machine for face classification. In our experimental study, the elapsed time for training the training set is 93.54 seconds and elapsed time for training the test set is 101.04 seconds. The user is prompted to enter the integer number for input image to recognize it by using the system. The system achieves the error recognition rate of 1.42 %. The output of the proposed system on AR face database is shown in the Fig 4.

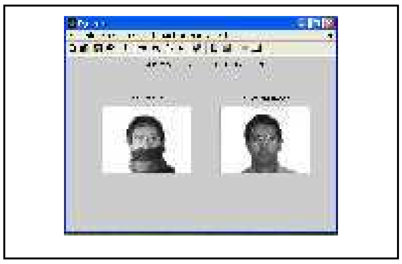

Fig 4 shows the output of the proposed system on AR database

Fig 5 shows the output of the proposed system on Essex database

4.2 Experimental study on Compound database

Finally, a subset of the compound dataset is used as the training set for computing the Eigen faces. It is composed of 2605 images: five images per person randomly chosen from AR face database, and Essex database. The other subset of 2605 images: five images per person randomly chosen from AR face database, and Essex database are used as test set. The output of the proposed system on Essex database is shown in the Fig 5.

In this study, the number of clusters are constructed after the Hierarchical agglomerative clustering and the number of clusters is c= 300. The number of proximal support vector machines trained for c clusters is 44850 (c(c-1)/2 pairs). It achieves the 1.60% of error recognition rate on compound database. Comparison of error recognition rates of different face recognition algorithms to the proposed system is shown in Table 2 and the comparison is pictorially depicted by using the chart diagram is as shown in the Fig 6.

Table 2 Brief comparison of error recognition rates

Method	Recognition error rate %
PCA	15.4
LDA	9.70
Bayesian (BAY)	6.70
One-vs-all Bayesian-SVM (OSVM)	4.00
HAC Bayesian SVM (HSVM)	3.70
Proximal Support vector machine (TSVM)	1.60

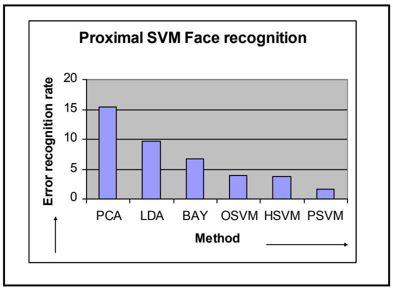

Fig 6 shows the Bar chart of Error recognition rate

5 Conclusions

We have presented a method for face recognition using proximal support vector machine using subspace analysis. The proximal support vector machine uses two parallel planes for binary classification. The strategy of the proximal SVM is to generate a plane that is closest to the one of the two face classes. The simple Hierarchical agglomerative clustering method is used to construct the face clusters based on Bayesian similarity measure. We have used pair wise PSVM method for face recognition that requires c(c-1)/2 pairs for classification. The proposed system uses standard support vector machine for clustering and proximal support vector machine for classification. We have performed experimental study on AR and Essex face databases. The experimental result gives the proof that the proposed system can outperform some of the existing methods.

References

[1] Zhao W et al, "Face Recognition- A literature survey", ACM Computing surveys, pp no 399- 458, 2003.
[2] Nayar S K et al, "Parametric Appearance Representation", early visual learning, Chapter 6, Oxford University press,1996.
[3] Burges C J C, "A Tutorial on Support vector machine for pattern recognition", Data mining and knowledge discovery, Vol 2, pp no 121-167, 1998.
[4] Mangasarian O L "Generalized Support Vector Machines", Advances in Large Margin Classifiers, pp no 135-146, Cambridge, 2000.
[5] Bernd Heisele et al, "A Component based Frame work for Face detection and identification", International Journal of Computer vision, Vol 74(2), pp no 167-181, 2006.
[6] Guodong Guo et al, "Support vector machines for Face Recognition", Special issue on ANN for image analysis and Computer Vision, Vol 19, pp no 631-638, 2001.
[7] Tai-Yuc Wang et al,"Fuzzy support vector machine for multi-class text categorization", Information processing an Management, Vol 43, pp no 914-929, 2007.
[8] Mathew A Turk et al "Face Recognition Using Eigenfaces", Cognitive Neuroscience, 3(1), pp no 71-86, 1991.
[9] Zhifeng et al, "Using SVM to enhance the performance of Bayesian Face Recognition", IEEE Transactions on Information Forensics and Security, Vol 2, No 2, pp no 174-180, June 2007.
[10] Glenn Fung, "Proximal Support Vector Machine Classifiers", Proceedings of 7th ICM KDD, pp no 77-86, 2001.
[11] Mangasarian Olvi L et al, "Multi Surface Proximal Support Vector machine classification via generalized Eigenvalues", IEEE PAMI Vol 28, no 1, pp no 69-74, 2006.
[12] Zeyu Li et al, "A Novel SVM Multi class classifier based on pariwise coupling", IEEE Conference on System, man and Cybernetics, Vol 7, pages: 5, 2002.
[13] Bruce Walter et al, "Fast Agglomerative Clustering for Rendering", IEEE Symposium on Interactive Ray Tracing, pp no 81-86, 2008.
[14] A. M Martinez. and R. Benavente, 1998"The AR face database," CVC Tech. Report #24.
[15] Essex database Available at:
htttp://cswww.essex.ac.uk/mv/allfaces/faces94.zip