IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 1, January 2011
ISSN (Online): 1694-0814
www.IJCSI.org

Fusion Of Facial Parts And Lip For Recognition Using Modular Neural Network

Anupam Tarsauliya[1], Shoureya Kant[2], Saurabh Tripathi[3], RituTiwari[4] and Anupam Shukla[5]

**[1]Department of Information Technology, IIITM
Gwalior: 474010, India**

**[2]Department of Information Technology, IIITM
Gwalior: 474010, India**

**[3]Department of Information Technology, IIITM
Gwalior: 474010, India**

**[4]Department of Information Technology, IIITM
Gwalior: 474010, India**

**[5]Department of Information Technology, IIITM
Gwalior: 474010, India**

Abstract

Face and Lip recognition has been benchmark problems in the field of biometrics and image processing. Various Artificial neural networks have been used for recognition purpose. This paper attempts at improving the recognition result for individuals by fusion of facial parts and lip using modular artificial neural network which employs parallel local experts with combinatory recognition techniques. Principal Component Analysis (PCA) and Regularized-Linear Discriminant Analysis (R-LDA) algorithm are used to extract low dimensional feature vector of images to drive neural networks effectively. Backpropagation Neural Network (BPNN) and Radial Basis Function Neural Network (RBFNN) are used as training algorithm for the database. Grimace database is used in this paper for carrying out the proposed methodology. Each facial image is divided into three sub image and a lip image. The three facial parts and the lip part are trained and tested individually. The fusion technique is applied using modular neural network by grouping sub images and the lip image in different network modules. Separate results obtained from each module are integrated to get the final result from the methodology used. This result set is compared with the result set obtained by training the sub images and the lip image individually. From the empirical and results finding it can be seen that the proposed methodology performs out better result.

Keywords: *Face, Lip Recognition, PCA, R-LDA, BPANN, RBFNN, Modular*

1. Introduction

Biometrics parameters like face, lips, iris and other various parameters are being used for individual identification and verification, recent works in field has enabled similar recognition for various states or expressions automatically. A major technical advancement in past few years has propelled face recognition technology into spotlight [1]. Face recognition is natural and passive and hence has clear advantage over other biometrics like fingerprints, iris etc. One of the major challenges involved in face recognition technique is the handling of various postures. Appearance based approaches for face recognition are being successfully developed and tested now a day [2]. These approaches are based on pixel intensity or intensity derived features. The performance of these methods depends upon types of training phase data, variations in pose, lighting and expressions [3].

Lip is one of the most important benchmark and advance parameter used for individual recognition of its varying postures [4]. Both Face and Lip can be represented as an image of size p x q pixels and further can be represented by a vector in p.q dimension space. In practical applications face and lip as parameters for identifying individuals outperform other biometrics [5]. Need for a automatic lip reading system is in spotlight because of its complication that stands for various moods and postures for facial expression.

Using Lip as modality for human identification has many advantages such as Lips biometrics is passive biometric i.e. individual interaction is needed. Images may be acquired from the distance without the knowledge of examined person. Lips biometrics is anatomical i.e. better results are expected than in behavioral biometrics. They are usually visible, may be implemented in hybrid face or voice recognition system [6].

Often a single biometric feature fails to provide sufficient evidences for verifying the identity of an individual. By fusion of multiple modalities pertaining to the field of biometric, the performance reliability of identification system can be improved. Due to its promising application, the fusion of multimodal biometric aspect is drawing more attention in recent years [7]. Face and lip multimodal biometrics are advantageous due to the use of non encroaching low cost image acquisition. Fusion of multimodal systems makes it difficult for intruders to trespass multiple biometric traits simultaneously.

Usually facial and lip images have their own representation of basis vectors of a high dimensional face vector space [8]. This large dimension is reduced for projecting the face vector to the basis vector by using various techniques such as Principal Component Analysis and Regularized- Linear Discriminant Analysis by approximating the original data with lower dimensional feature vector. PCA provides an effective technique for dimensionality reduction which involves computation of eigen values and eigen vectors by using the co-variance matrix of original input data vector [9]. The orthonormal vectors are computed from it which are the basis of the computed data. R-LDA reduces high variance of the eigen value estimates of the within class scatter matrix at expense of potentially increased bias [10].

In this paper, we attempt at presenting a novel fusion strategy for personal identification using facial parts and lip biometrics. The proposed paper shows that integration of face parts and lip biometrics can outperform single biometric indicators. We present a architecture based on modular neural network for integrating facial sub parts and lip based on PCA and R-LDA.

2. Methodology

We performed face recognition task on Grimace database (shown in Fig. 1), which contains 360 colored face images of 18 individuals forming 18 classes while there are 20 images present for each subject. Database images vary in expression & position. Figure 1 shows examples from Grimace database.

Fig 1. Grimace database of 18 individuals

We divide each individual image of 180 x 200 pixels into three subimages and a lip image. The three subimages which are left half, right half and lower half have dimension of 90 x 130 pixels, 90 x 130 pixels and 180 x 60 respectively. Lip image which corresponds to the lip section of image face has dimension of 70 x 32 pixels. Then the colored images are converted into gray scale & processed using histogram equalization. Figure 2 below shows the partitioning of the facial image.

Fig 2. Sub Images and Lip Image used

We then process the image database using Principal Component Analysis and Regularized-Linear Discriminant Analysis for normalization and dimensionality reduction i.e. we extract low dimensional feature vector of the images.

After the pre-processing stage Modular Neural Network (MNN) is used for the classification purpose. Backpropagation and radial basis algorithms are used for training the neural network on training set obtained from processed image database. 70% of each sub image including the lip image is used for training the ANN and 30% are used for testing purpose. Figure 3 below shows the training and testing methodology.

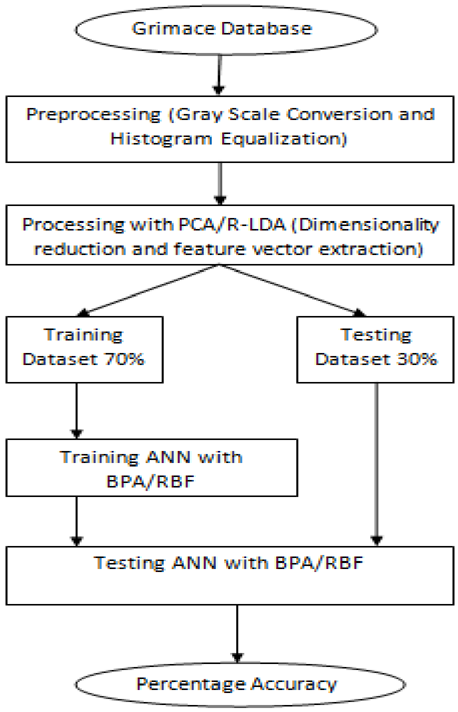

Fig 3. Flow Diagram for Methodology

The architecture consists of 3 sub-images and a lip image forming the reason to create 4 major sub-image modules. Each of these modules consists of 2 network modules, which are made to implement different combinations of principal component analysis, regularized linear discriminant analysis along with back propagation neural network and radial basis function neural networks. After learning the type 1 network module of each sub-images are integrated using in dedicated integration unit and type 2 network modules are integrated in integration unit dedicated for them. Both of these units are further integrated in module integration unit. The architecture is shown in figure 4.

The facial input image is sub divided into three sub facial image modules represented as Sub1, Sub2, Sub3 and a lip module as Lip respectively. Further each of these modules is divided into two separate network modules as N1 and N2 respectively. Sub1, Sub2, Sub3 are the upper left, upper right and lower half parts of the facial image respectively. N1 and N2 module of Sub1 is processed and trained using PCA with BPNN and R-LDA with RBFNN. N1 and N2 module of Sub2 is processed and trained using PCA with RBFNN and R-LDA with BPNN. N1 and N2 module of Sub3 is processed and trained using R-LDA with RBFNN and PCA with BPNN. N1 and N2 module of Lip is processed and trained using R-LDA with BPNN and PCA with RBFNN respectively. Network Module 1 – N1 of each of the sub image module and the lip module is integrated using various integration techniques stated below to form integrated result module M1. Similarly, Network Module 2 – N2 of each of the sub image module and the lip module is integrated using various integration techniques stated below to form integrated result module M2. M1 and M2 thus obtained are integrated using the same integration techniques to yield the final result.

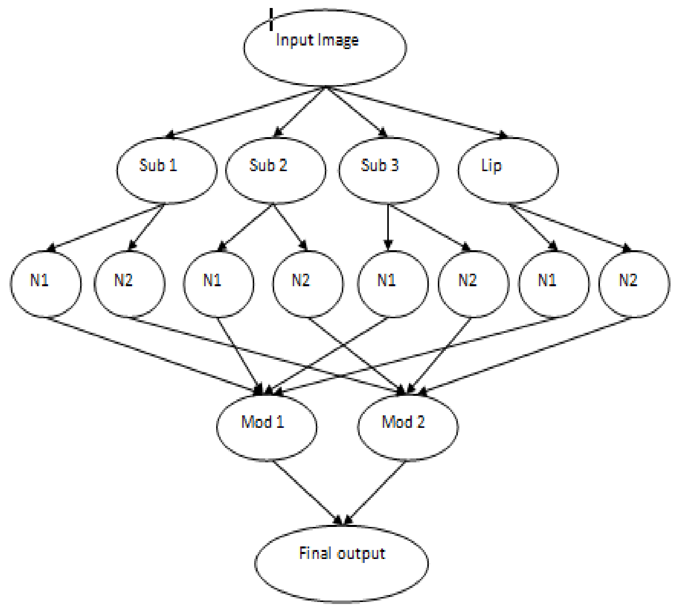

Fig 4. Used Modular Neural Network Architecture

Following strategies are used for integration: Probabilistic Sum Integration, Product Integration, Max integration, Min Integration and Polling Integration.

2.1 Back Propagation Algorithm

Backpropagation is the generalization of the Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way as defined by you [11]. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities.

Standard backpropagation is a gradient descent algorithm, as is the Widrow-Hoff learning rule, in which the network weights are moved along the negative of the gradient of the performance function. The term backpropagation refers to the manner in which the gradient is computed for nonlinear multilayer networks. There are a number of variations on the basic algorithm that are based on other standard optimization techniques, such as conjugate gradient and Newton methods [12].

2.2 Radial Basis Network

Radial basis function neural networks (RBFNNs) share features of the back propagation neural networks (BPNNs) for pattern recognition. They are being extensively used for on- and off-linear adaptive modeling and control applications. RBFNNs store information locally whereas the conventional BPNNs store the information globally.

RBF nets belong to the group of kernel function nets that utilize simple kernel functions, distributed in different neighborhoods of the input space, whose responses are essentially local in nature. The architecture consists of one hidden and one output layer. This shallow architecture has great advantage in terms of computing speed compared to multiple hidden layer nets.

The function newrb iteratively creates a radial basis network one neuron at a time. Neurons are added to the network until the sum-squared error falls beneath an error goal or a maximum number of neurons has been reached. The function newrb takes matrices of input and target vectors P and T, and design parameters goal and spread, and returns the desired network [13].

2.3 Principal Component Analysis

Principal component analysis (PCA) is a classical statistical method. It is a variable reduction procedure and it is appropriate to use, when obtained a number of observed variable s and wish to develop a smaller number of artificial variable. Principal component analysis (PCA) is a mathematical method that transforms a number of correlated variables into a number of uncorrelated variables called principal components. Hence it is used to approximate the original data with lower dimensional feature vectors.

Principal Component analysis is one of the most successful techniques used in image recognition. The principal component can be used for prediction, redundancy removal, feature extraction, data compression [14].

Let X= $(x^1, x^2, \ldots, x^i, \ldots, x^n)$ represents the $n \times N$ data matrix, where each xi is a lip vector of dimension *n*, concatenated from a $p \times q$ *lip* image.

$$Y = W^T q \qquad (1)$$

Where Y is the $m \times N$ feature vector matrix, m is the dimension of the feature vector and transformation matrix W is an $n \times m$ transformation matrix whose columns are the eigenvectors corresponding to the m largest eigen values.

The principle component $w1$ of the data set X is as follows:

$$W^i = \underset{\|w\|=1}{\arg\max}\ \mathrm{var}\{W^T x\} \qquad (2)$$

2.4 Regularized - Linear Discriminant Analysis

The R-LDA method presented here is based on a novel regularized Fisher's discriminant criterion, which is particularly robust against the SSS problem compared to the traditional one used in LDA. The purpose of regularization is to reduce the high variance related to the eigen value estimates of the within-class scatter matrix at the expense of potentially increased bias. The trade-off between the variance and the bias, depending on the severity of the SSS problem, is controlled by the strength of regularization [15].

Unlike the PCA method that extracts features to best represent face images; R-LDA method tries to find the sub space that best discriminates different face classes. By applying this method one can find the projection direction that on one hand maximizes the distance between the face images of different classes. On the other hand minimizes the distance between the face images of the same class.

2.5 Modular Neural Network

A modular architecture allows decomposition and assignment of task to several modules. Therefore, separate architecture can be developed to each solve a sub-task with the best possible architecture and individual modules or building blocks may be combined to form a comprehensive system. The module decomposes the problem into two or more subsystem that operates on inputs without communicating to each other. The input units are mediated by an integrating unit that is not permitted to feed information back to modules. The modular architecture combines two learning schemes ,supervised and competitive [16] .The supervised learning scheme is used to train the different module of networks and a gating netwok operates in a competitve modeto assign different patterns of the task to a module through a mechanism that acts as a mediator. Figure 5 given below illustrates the architecture of a typical MNN.

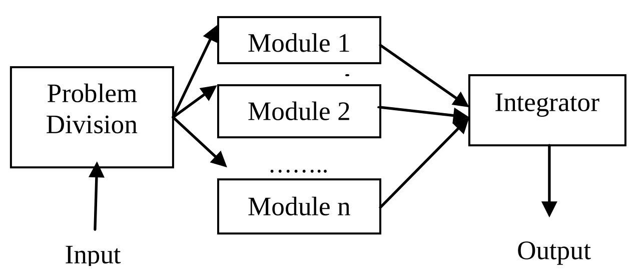

Fig 5. General Modular Neural Network Architecture

The integration of the result from various modules is performed using
Probabilistic Sum Integration, Product Integration, Max integration, Min Integration and Polling Integration techniques. On the basis of the outputs that the various modules or ANNs generate, the decision regarding the final

system output is made. Integration is the mechanism to combine the various outputs into a single output of the system. The integration again depends upon the design of the MNN and the manner in which the division of the problem has taken place.

The mechanism to compute the final output by these outputs is done by the integrator using a variety of methods.

Sum rule is applicable when a high level of noise and/or high ambiguity in the classification problem cause the posterior estimated by a classifier not to deviate much from the prior. Max rule approximates the mean by the maximum of the posteriors. Min rule approximates the mean by the minimum of the posteriors [17].

In polling [18], each of the modules gives as its output the class to which the input may belong as per its knowledge and methodology. The integrator gets the various classes that are potentially the output of the system by the various modules. The task of the integrator is to decide the system output as a means of voting between the various modules. Each module casts one vote in favor of the class which is its output. The votes for the various classes are collected and the class getting the largest vote is regarded as the final system output by the integrator. In case of a tie between the modules, any one of the classes may be randomly chosen out of the classes involved in a tie.

In probabilistic sum, every module gives as many outputs as there are classes in the system. Each output measures the probability of the occurrence of any class as the final output class. The probability lies in between 0 to 1. An output of 1 by any module for any class means that as per the recordings of the module that particular class is surely the class to which the input perfectly maps to. An output of 0 by any module for any class means that as per the recordings of the module that particular class is surely not the class to which the input perfectly maps to. Every module hence computes the probability for each of the classes in the system. This probability vector comprising of all the probabilities is passed to the integrator for computing the final output.

3. Results

Table 1 and Table 2 below depict the network parameters for BPNN and RBFNN respectively.

Table 1. Network Parameters for BPNN

Learning Rate	Momentum	Epochs	Goal
0.001	0.5	10,000	1.0e-3

Table 2. Network Parameters for RBFNN

Spread	Epochs	Goal
0.4	10,000	1.0e-3

Subimage 1 which is upper left of the face having dimension of 90 x 130 pixels has been taken as separate module and is further divided into two network modules as PCA with BPNN and PCA with RBFNN respectively. The individual network module are trained and tested with their corresponding ANN's respectively. The result thus obtained is shown in table1. It can be seen from the table that module trained with RBFNN gives better result than module trained with BPNN.

Table 3. Individual results for Sub Image 1

SubImage 1 Network Module	Feature Extraction Classification Techniques	Matching Score
Network Module 1	PCA and BPNN	159/180
Network Module 2	PCA and RBFNN	163/180

Subimage 2 which is upper right of the face having dimension of 90 x 130 pixels has been taken as separate module and is further divided into two network modules as R-LDA with BPNN and R-LDA with RBFNN respectively. The individual network module are trained and tested with their corresponding ANN's respectively. The result thus obtained is shown in table2. It can be seen from the table that module trained with BPNN gives better result than module trained with RBFNN.

Table 4. Individual results for Sub Image 2

SubImage 2 Network Module	Feature Extraction Classification Techniques	Matching Score
Network Module 1	R-LDA and RBFNN	160/180
Network Module 2	R-LDA and BPNN	163/180

Subimage 3 which is upper left of the face having dimension of 180 x 60 pixels has been taken as separate module and is further divided into two network modules as R-LDA with RBFNN and PCA with BPANN respectively. The individual network module are trained and tested with their corresponding ANN's respectively. The result thus obtained is shown in table3. It can be seen from the table that module trained with RBFNN gives better result than module trained with BPNN.

Table 5. Table Individual results for Sub Image 3

SubImage 3 Network Module	Feature Extraction Classification Techniques	Matching Score
Network Module 1	R-LDA and RBFNN	163/180
Network Module 2	PCA and BPNN	157/180

Lip Image which is lip part of the face having dimension of 70 x 32 pixels has been taken as separate module and is further divided into two network modules as R-LDA with BPNN and PCA with RBFNN respectively. The individual network module are trained and tested with their corresponding ANN's respectively. The result thus obtained is shown in table4. It can be seen from the table that module trained with RBFNN gives better result than module trained with BPNN.

Table 6. Individual results for Lip Image

Lip Image Network Module	Feature Extraction Classification Techniques	Matching Score
Network Module 1	R-LDA and BPNN	161/180
Network Module 2	PCA and RBFNN	165/180

Table 5 below shows the integrated results obtained from five various integrated techniques namely Sum, Max, Min, Product and Polling for network module 1 and network module 2 of sub images (1 – 3) and the lip image modules.

Table 7. Integrated results for network modules

Integration Technique	Network Module 1	Network Module 2
Probabilistic Sum Integration	164/180	166/180
Max Integration	163/180	165/180
Min Integration	162/180	166/180
Product Integration	167/180	165/180
Polling Integration	169/180	170/180

Table 6 below shows the final integrated result obtained from module 1 and module 2 using various integrated techniques namely Sum, Max, Min, Product and Polling for network module 1 and network module 2 of subimages (1 – 3) and the lip image modules.

Table 8. Final result obtained

Integration Technique	Final Network Module Output
Probabilistic Sum Integration	165/180
Max Integration	164/180
Min Integration	168/180
Product Integration	167/180
Polling Integration	172/180

From the table obtained we can see that the most optimized result is obtained from polling integration technique with recognition of 95.56%.

5. Conclusion

This paper attempts at bringing forth the advantage of fusion of multiple biometrics modalities over single biometric feature identification mechanism. The facial parts along with the lip image are fused together using modular neural network architecture and this system is then analyzed and compared with individual performance. The preprocessing stage involves the image preprocessing of gray scale conversion and histogram equalization, processing with PCA/R-LDA for dimensionality reduction and feature vector extraction. The preprocessing stage is followed by the training the classification mechanism using BPA/RBF algorithms. From the table 1-4, it can be seen that training of module with RBFNN gives better result for three of the image module which is also the case for the lip image. It is also noticed that lip image module gave better matching score than other three sub images taken for the experiment. Different integration techniques are used for integrating the results of the modules, integrated results are found to better than the individual scoring results obtained of a module; it can be seen from table 5 and table 6 that polling integration technique gives the best result in comparison to rest of the techniques used for the integration purpose. The best matching score obtained from the three facial sub image comes out to be 90.55% with RBFNN as training algorithm. Lip image shows the matching score of 91.67% using RBFNN as training algorithm. Results obtained from the proposed fusion techniques stands better than individual matching scores at 95.56% with polling as the integration technique followed by min integration at 93.34%.

6. Future scope

Further research work can be done in this field, by fusion of other biometric features. Neural network parameter can be optimized using evolutionary algorithms for getting better results. Other training algorithm in place of BPA and RBF may give better performance than used algorithms. Modular architecture i.e. the number of modules used for an image can be varied to get better result. Combinations of other different processing, training and integration technique may be used.

7. Acknowledgement

We greatly acknowledge and admire the co-operation extended to us by Mr. Rahul Kala, Ph.d. Scholar, School of Cybernatics, University Of Reading. His sincere efforts and guidance paved the way towards the completion of this paper.

8. References

[1] Szewczyk, R.; "New Features Extraction Method for People Recognition on the Basis of the Iris Pattern," Mixed Design of Integrated Circuits and Systems, 2007. MIXDES'07. 14th International Conference on , vol., no., pp.645-650, 21-23 June 2007

[2] Kresimir Delac, Mislav Grgic and Sonja Grgic, "Generalization Abilities of Appearance-Based Subspace Face Recognition Algorithms", 12th Int. Workshop on Systems, Signals & Image Processing, 22-24 September 2005.

[3] Sheetal Chaudhary, Rajender Nath, "Hybrid Approach for Template Protection in Face Recognition System", Global Journal of Computer Science and Technology, P a g e 34 Vol. 10 Issue 5 Ver. 1.0 July 2010

[4] Mok, L.L.; Lau, W.H.; Leung, S.H.; Wang, S.L.; Yan, H.; , "Person authentication using ASM based lip shape and intensity information," Image Processing, 2004. ICIP '04. 2004 International Conference on , vol.1, no., pp. 561- 564 Vol. 1, 24-27 Oct. 2004

[5] Lievin, M.; Luthon, F.; , "A hierarchical segmentation algorithm for face analysis. Application to lipreading," Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on , vol.2, no., pp.1085-1088 vol.2, 2000

[6] Michał Choras, “The lip as a biometric”, Theoretical Advances, Pattern Anal Applic (2010) 13:105–112 8 January 2009

[7] Nageshkumar.M, Mahesh.PK and M.N. Shanmukha Swamy,” An Efficient Secure Multimodal Biometric Fusion Using Palmprint and Face Image,” IJCSI International Journal of Computer Science Issues, Vol. 2, 2009

[8] Nor'aini , A.J.; Raveendran, P.; Selvanathan, N. “A comparative analysis of feature extraction methods for face recognition system” Sensors and the International Conference on new Techniques in Pharmaceutical and Biomedical Research, 2005 Asian Conference on 5-7 Sept. 2005.

[9] M.H. Tan, J.K. Hammond.” A non-parametric approach for linear system identification using principal component analysis”, Mechanical Systems and Signal Processing 21 (2007) 1576–1600, 7 September 2006

[10] Juwei Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, "Face Recognition Using LDA Based Algorithms", IEEE Transactions on Neural Networks, Vol. 14, No.1, Page: 195-200, January 2003.

[11] K.Rama Linga Reddy, G.R Babu, Lal Kishore, M.Maanasa,” Multiscale Feature And Single Neural Network Based Face Recognition”, Journal of Theoretical and Applied Information Technology,2005

[12] Saeid Iranmanesh, M. Amin Mahdavi,” A Differential Adaptive Learning Rate Method for Back-Propagation Neural Networks”, World Academy of Science, Engineering and Technology 50,2009.

[13] Ch. Satyananda Reddy, P. Sankara Rao, KVSVN Raju, V. Valli Kumari,” A New Approach For Estimating Software Effort Using RBFN Network”, IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.7, July 2008

[14] Khoukhi, A.; Ahmed, S.F.; , "Fuzzy LDA for face recognition with GA based optimization," Fuzzy Information Processing Society (NAFIPS), 2010 Annual Meeting of the North American , vol., no., pp.1-6, 12-14 July 2010

[15] Salehi, N.B.; Kasaei, S.; Alizadeh, S.; , "Face Recognition Using Boosted Regularized Linear Discriminant Analysis," Computer Modeling and Simulation, 2010. ICCMS '10. Second International Conference on , vol.2, no., pp.89-93, 22-24 Jan. 2010

[16] Sarosh I. Khan , Stephen G. Ritchie,” Statistical and neural classifiers to detect traffic operational problems on urban arterials”, Transportation Research Part C 6 (1998) 291-314,1998

[17] Shakhnarovich, G.; Darrell, T.; , "On probabilistic combination of face and gait cues for identification," Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on , vol., no., pp.169-174, 20-21 May 2002

[18] Rahul Kala, Ritu Tiwari, Anupam Shukla, “Real Life Applications of Soft Computing”, CRC Press, May 21, 2010