
Appearance and shape based face recognition using Backpropagation learning neural network algorithm with different lighting variations

Md. Rabiul Islam^{1,*}, Rizoan Toufiq¹, Md. Abdus Sobhan²

¹Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh

²School of Engineering & Computer Science, Independent University, Dhaka, Bangladesh

Email address:

rabiul_cse@yahoo.com (M. R. Islam), rizoantoufiq@yahoo.com (R. Toufiq), sobhan30@gmail.com (M. A. Sobhan)

To cite this article:

Md. Rabiul Islam, Rizoan Toufiq, Md. Abdus Sobhan. Appearance and Shape Based Face Recognition Using Backpropagation Learning Neural Network Algorithm with Different Lighting Variations. *Science Journal of Circuits, Systems and Signal Processing*. Vol. 2, No. 4, 2013, pp. 93-99. doi: 10.11648/j.cssp.20130204.11

Abstract: This paper presents an approach of face recognition system using Backpropagation learning neural network algorithm introducing appearance and shape based facial features to enhance the efficiency with different lighting variations. To extract the appearance and shape based facial feature, Active Appearance Model (AMM) has been applied. Appearance based facial feature is useful when the lighting condition is uniform. On the other hand when the environmental lighting condition is different, shape based facial features can perform better in comparison with appearance based feature because shape based structure is not changed with lighting variations. In this work, both appearance and shape based facial features are combined to enhance the recognition efficiency for various light variant system. For dimensionality reduction of appearance and shape based facial features, Principal Component Analysis (PCA) method has been used. Finally, error Backpropagation learning feed forward neural network algorithm has been used to classify the facial features. To measure the performance of the proposed appearance and shape based facial recognition system, VALID database has been used where each face has been captured with four different lighting variations. Experiments have been performed with Appearance-Only, Shape-Only and combined Appearance-Shape based feature and performance of the proposed system shows the superiority of the face recognition system.

Keywords: Appearance and Shape Based Facial Features, Face Recognition with Different Lighting Variations, Principal Component Analysis, Backpropagation Learning Neural Network, Active Appearance Model

1. Introduction

Biometry is currently a very active area of research spanning several sub-disciplines such as image processing, pattern recognition and computer vision. The main goal of biometry is to build systems that can identify people from some observable characteristics such as their face, fingerprints, iris, etc. Facial recognition is the identification of humans by the unique characteristics of their faces [1]. Face recognition based personal identification results in a highly accurate and secured system that has been widely applied in various practical fields [2, 3, 4].

Several excellent survey papers on face recognition techniques are available with a wide variety of methods [5, 6, 7, 8] that covers early face recognition approaches. While humans quickly and easily recognize faces under variable situations or

even after several years of separation, the problem of machine face recognition is still a highly challenging task in pattern recognition and computer vision [9, 10].

Different techniques can be used to track and process faces [11] such as neural networks [12, 13], eigenfaces [14], Markov chains [15] etc. Most of the face recognition research uses the public 2D face databases as the input pattern [16], with a recognition performance that is often sensitive to pose and lighting conditions. One way to override these limitations is to combine modalities: color, depth, 3D facial surface etc. [16, 17, 18, 19, 20, 21]. Most 3D acquisition systems use professional devices such as a traveling camera or a 3D scanner [19, 20]. Typically, these systems require that the subject remain immobile during several seconds in order to obtain a 3D scan and therefore may not be appropriate for some applications such as human

expression categorization using movement estimation. Moreover, many applications in the field of human face recognition such as human-computer interfaces, model-based video coding and security control [22, 23] need to be high-speed and real-time, for example, passing through customs quickly while ensuring security. Furthermore, the cost of systems based on sophisticated 3D scanners can easily make such an approach prohibitive for routine applications [1].

In order to avoid using expensive and time intensive 3D acquisition devices and process, in this paper a combined 2D appearance and shape based facial feature based facial recognition system has been proposed with different lighting variations. Though the proposed approach has used a 2D method, it can perfectly cluster the visual features because when the appearance based feature is captured with noise (i.e. light variations) then shape based features can perform noise perfect. The system can also work when the shape

based feature is captured by noise highly. By combining this approach, the proposed face recognition technique can perform very well especially in various lighting environmental conditions.

2. Overview of the Proposed Face Recognition System

Fig. 1 shows the block diagram of the proposed face recognition system. At first facial images are taken using high quality digital camera. Then some facial image pre-processing such as noise removing technique to reduce the noise from the facial image, filtering technique to enhance the sharpness of the face and reduce the strong spikelike components, histogram equalization for reducing the number of gray levels in the image, edge detection to detect edges from the facial images etc have been applied.

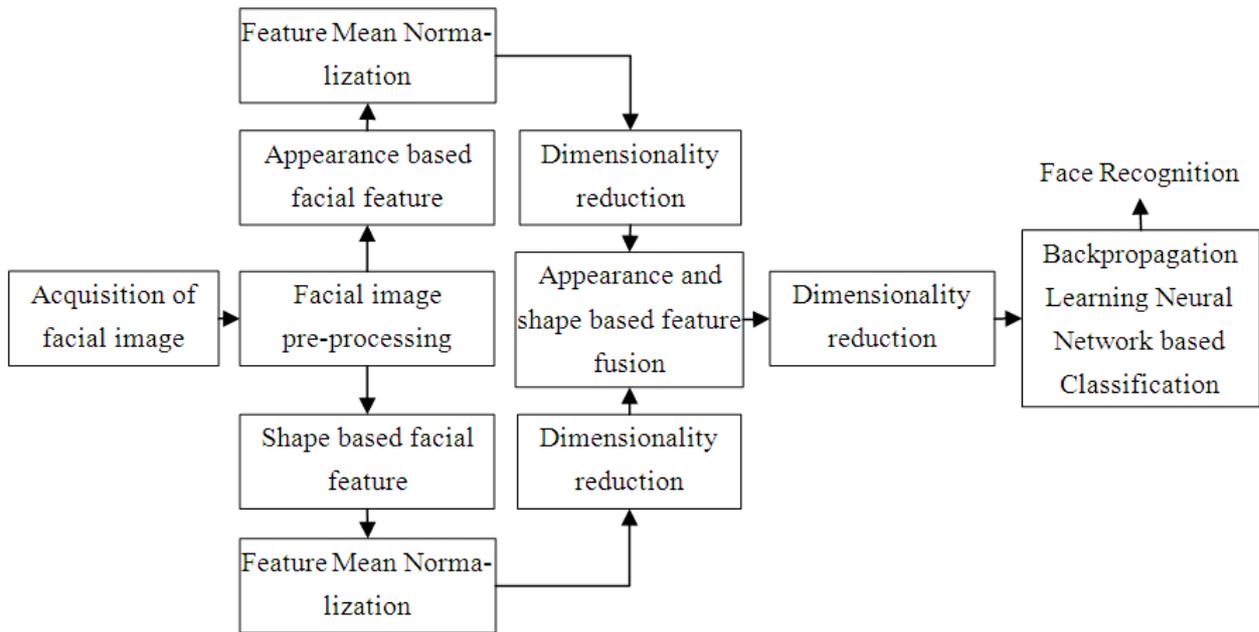


Figure 1. Paradigm of the proposed appearance and shape based face recognition system.

Active Appearance Model (AAM) has been applied to extract the appearance and shape based facial features. Since two different types of features are extracted from the facial images, if we want to combine those features then the dimension of the feature vector will increase a lot. As a result, Principal component analysis based dimensionality reduction technique has been used to reduced the dimensionality of the appearance and shape based feature vector separately. For combining two different types of feature into a single vector, feature normalization technique has been used. Finally, appearance and shape based feature fusion is performed and again use PCA for dimensionality reduction. To classify the facial features, Backpropagation learning neural network algorithm has been used for learning and recognition.

3. Facial Feature Extraction and Dimensionality Reduction

To extract the facial features from the face image, Active Appearance Model (AAM) has been used. An Active Appearance Model is an integrated statistical model which combines a model of shape variation with a model of the appearance variations in a shape-normalized frame. An AAM contains a statistical model of the shape and gray level appearance of the object of interest which can generalize to almost any valid example. Matching to an image involves finding model parameters which minimize the difference between the image and a synthesized model example, projected onto the image. The AAM is constructed based on

a training set of labeled images where landmark points are marked on each example face at key positions to outline the main features [24].

Fig. 2(b) shows an example face with landmark points where specific points are connected to distinguish among eyes, leaps, nose etc. Stephen and Fred [25] have developed an Active Shape Model (ASM) which has been used in this work to detect the facial features. After getting the connected landmark points, the binary image has been taken and treated as shape edge model of the face. The Region Of Interest (ROI) has been chosen according to ROI selection algorithm [26]. Lastly the background noise has been eliminated [27] and appearance based facial feature has been found. The procedure of ROI selection for two different facial images is shown in Fig. 2.

After extracting the appearance and shape features from the facial images, PCA has been used to reduce the dimensionality of the appearance and shape based feature vector. Finally, 110 features have been extracted from the PCA based dimensionality reduction method of combined appearance and shape based feature fusion which is shown in Fig. 4.

A face image $I(x,y)$ is a two dimensional N by N array of intensity values. An image may also be considered as a vector of dimension N^2 . Let the training set of face images be $\Gamma_1, \Gamma_2, \Gamma_3... \Gamma_M$. The average face of the set is defined by

$$\Psi = \frac{1}{M} \sum_i^M \Gamma_i \quad (1)$$

Each face differs from the average by the vector,

$$\Phi_i = \Gamma_i - \Psi \quad (2)$$

$$\lambda_k = \frac{1}{M} \sum_{n=1}^M (\mu_k^T \Phi_n)^2 \quad (3)$$

$$\mu_l^T \mu_k = \delta_{lk} = \begin{cases} 1, & \text{if } l = k \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

The vectors μ_k and scalars λ_k are the eigenvectors and eigenvalues respectively of the covariance matrix

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T \quad (5)$$

$$= AA^T$$

where $A = [\Phi_1, \Phi_2... \Phi_M]$. The matrix C , however, is N^2 by N^2 , and determining the N^2 eigenvectors and eigenvalues is an intractable task for typical image size. Fortunately, we can solve for the N^2 dimensional eigenvectors in this case by first solving for the eigenvectors of a M by M matrix and then taking appropriate linear combinations of the face images Φ_i . Consider the eigenvectors v_i of $A^T A$ such that

$$A^T A v_i = \mu_i v_i \quad (6)$$

$$AA^T A v_i = \mu_i A v_i \quad (7)$$

From which, we see that $A v_i$ are the eigenvectors of $C = AA^T$. Following this analysis, we construct the M by M matrix $L = A^T A$ where, $L_{mn} = \Phi_m^T \Phi_n$ and find the M eigenvectors, v_i , of L . These vectors determine linear combinations of the M training set face images to form the eigenfaces μ_i ,

$$\mu_l = \sum_{k=1}^M v_{lk} \Phi_k, \quad l = 1, 2, \dots, M \quad (8)$$

With this analysis, the calculations are greatly reduced from the order of the number of pixels in the images (N^2) to the order of the number of images in the training set (M) [28, 29, 30, 31]

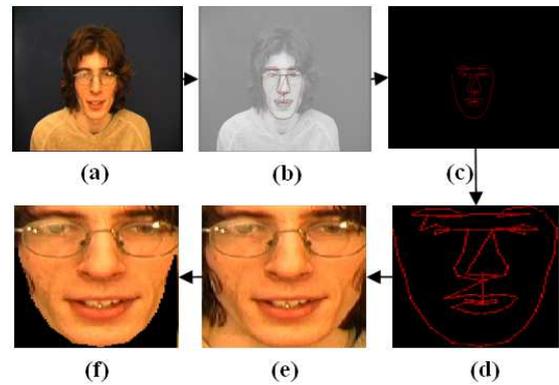


Figure 2. Facial image pre-processing for the proposed system (a) Original image (b) Output taken from Starns Active Appearance Model (c) Extracted facial edges (d) Shape based features (e) Region Of Interest (ROI) selection with background noise (f) Appearance based facial features.

4. Backpropagation Learning Neural Network Based Facial Feature Classification

An ANN which is found to be useful in addressing problems requiring recognition of complex patterns and performing nontrivial mapping functions is the back-propagation network [32]. The network that has been used in this system is a feed-forward network with tan-sigmoid transfer functions in both the hidden layer and the output layer which is shown in Fig. 3.

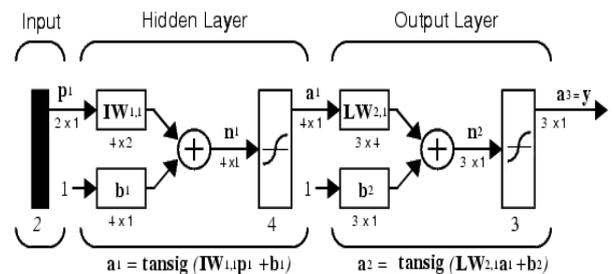


Figure 3. Architecture of Back-propagation neural network.

If the input vector is $I = [p_1, p_2, \dots, p_n]$, then the output of hidden layer has been calculated as following,

$$n = IW + b \tag{9}$$

$$a = f(n) = \frac{2}{(1 + e^{(-2*n)})^{-1}} \tag{10}$$

Where, W = weight vector and b = bias input. The error is calculated as the difference between the target output and the network actual output. The goal is to minimize the average of the sum of these errors.

$$mse = \frac{1}{M} \sum_{k=1}^M e(k)^2 = \frac{1}{M} \sum (t(k) - a(k))^2 \tag{11}$$

Here, mse means mean square error, $t(k)$ represents the target output and $a(k)$ represents the network output. The weights and bias values are updated based on the goal average error value.

In face recognition system, the final weights and bias values are calculated in training stage. In test phase, the output of the network has been calculated for the new input image and compared with the target output to select the class of the input facial image. The details of the proposed system architecture with Backpropagation learning neural network algorithm is shown in Fig. 4.

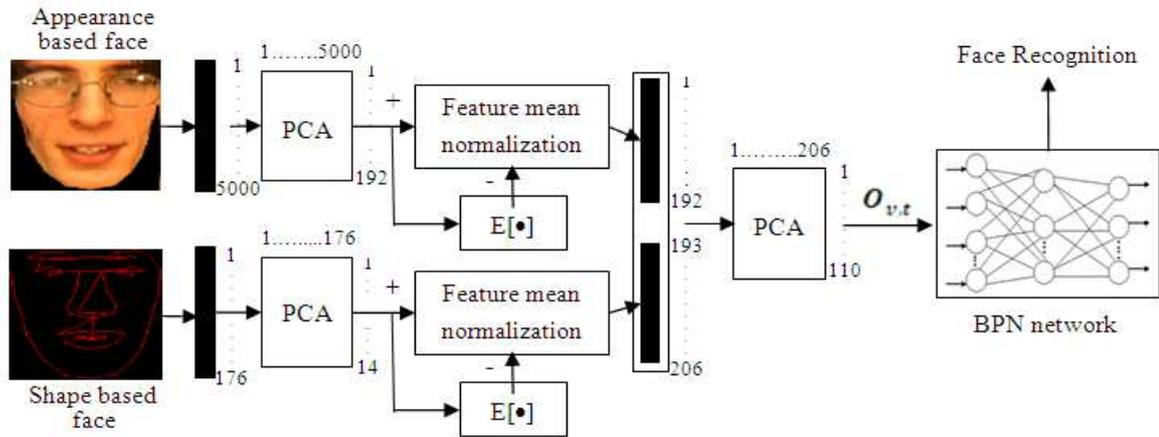


Figure 4. Architecture of the proposed system in detail.

The major drawbacks of Backpropagation learning neural network algorithm is the training time and local minima. Convergence time of the Backpropagation algorithm is inversely proportional to the error tolerance rate. In learning, effective use of error rate can decrease the convergence time. At first select the final error rate. Then converge the weights such that all the patterns overcome some of the percentage error of the total system (the error must be higher than the final error rates). Finally converge the system to the next lower error rate until cross the final targeted error. For example, if the error rate of the system is 0.001, first the converged error rate for all of the patterns is 0.009, then 0.005, 0.003 and finally 0.001. This process is known as SET-BPL [33].

100 speech utterances are trained in Backpropagation learning neural network and the effects of applying SET-BPL of the proposed system is shown in Fig. 5. Sometimes local minima problem has been occurred in Backpropagation learning neural network algorithm. As a result, some precautions such as addition of internal nodes, lowering the gain term etc has been considered to set the learning parameters. More addition of internal nodes and lowering the gain term can increase the convergence time. To overcome these learning difficulties, momentum term has been used to speed up the convergence process for this proposed appearance and shape based speaker identification system.

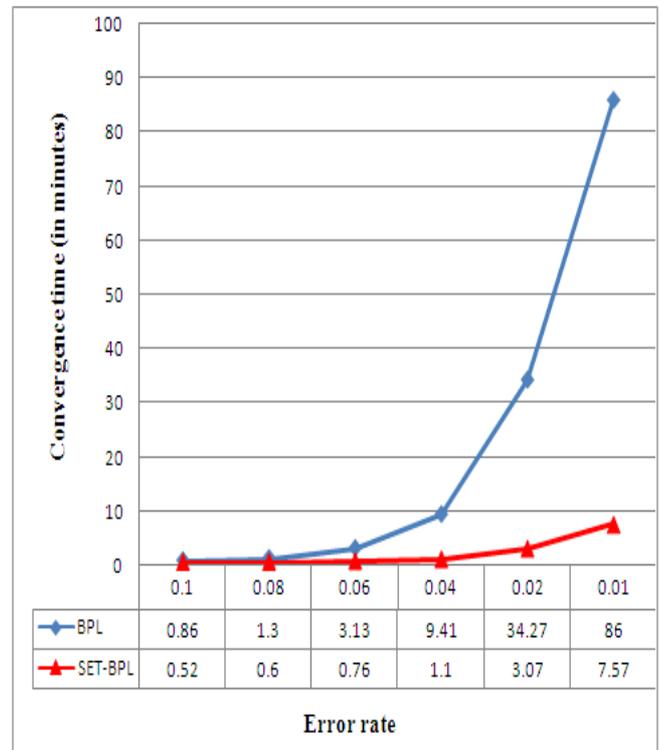


Figure 5. Effects of applying SET-BPL process of the proposed system.

5. Experimental Results and Performance Analysis

Performance has been analyzed according to various dimensions with VALID Audio-Visual database [34] to measure the performance of the proposed appearance and shape based face recognition system. Though VALID is a multimodal Audio-Visual database, it provides 106 subjects each with four office lighting conditions, gathered periodically over one month giving some temporal variation and one studio session with controlled lighting. The five sessions were recorded over a period of one month, allowing for variation, clothing, facial hair, hairstyle and cosmetic appearance of the subjects and also variation of the visual background. The database is designed to be realistic and challenging, hence the other four sessions were recorded in noisy real world scenarios with no control on illumination i.e. uncontrolled environment. One facial image that are captured in controlled lighting environment is used for learning and four other images with different lighting variations are used for testing. To populate the experimental results, 100 subjects has been used from this database.

To evaluate the performance of the proposed face recognition system, Appearance based, Shape based and combined Appearance-Shape based facial features are tested. Finally, the results are compared and best result has been taken for the proposed appearance and shape based facial recognition system. Receiver Operating Characteristics (ROC) curve is populated where a trade off is made between security and user friendness which is shown in Fig. 5.

From the ROC curve, it is observed that the performance of the Appearance-Shape based feature can perform better than Appearance based only and Shape based only features. For example, at a FRR = 40%, the Appearance based, Shape based and Appearance-Shape based FAR are 28%, 23% and 16% respectively. Therefore, this scenario represents that combined Appearance-Shape based feature can perform better than individual feature such as Appearance based only or Shape based only feature.

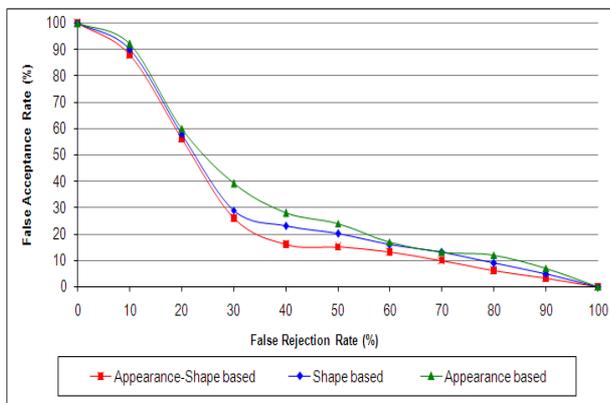


Figure 7. Performance comparison among Appearance based, Shape based and Appearance-Shape based facial recognition system.

Since the system performance has been degraded according to various noises, filtering techniques have been adapted to increase the efficiency of the proposed face recognition system. Filtering is used for modifying or enhancing an image, to emphasize certain features or remove other features from an image. Wiener filtering technique has been used to remove or reduce white Gaussian noise from the facial image. Wiener filter can be used adaptively to an image where the variance is large, wiener filter performs little smoothing and where the variance is small, wiener filter performs more smoothing. Wiener filtering technique is more selective than a comparable linear filter, preserving edges and other high-frequency parts of an image.

Based on wiener filtering technique, experiment has been performed on the basis of noisy facial image according to Appearance based, Shape based and Appearance-Shape based feature where Backpropagation learning neural network has been used to measure the accuracy of the noise treatment system. Fig. 6 shows the result after applying wiener filtering technique. From the experiments, it is clearly visible that when the white Gaussian noises increase with different lighting variations, the appearance based feature affects a lot and appearance based feature performance degrades huge amount. The lighting variations can not affect the shape based feature. As a result, the shape based feature retain the recognition accuracy which will further increase combined appearance and shape based facial recognition performance.

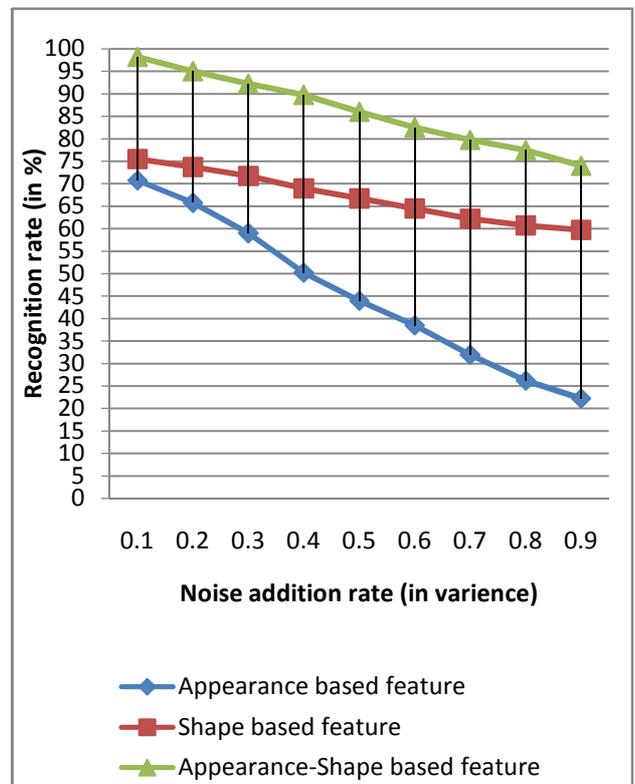


Figure 6. Proposed face recognition rate according to wiener filtering technique with different noise addition rate.

6. Conclusions and Observations

Experimental results shows that by using shape based facial feature with appearance based feature, the recognition performance increases a significant amount. Especially when noise level increases gradually the recognition rate of appearance based facial feature gradually decreases more than the shape based feature. With different lighting variations when appearance based features are affected largely then shape based feature can perform better. In this way by combining appearance and shape feature, the recognition performance can be increased in a satisfied level even the lighting variations can create noise heavily. So, the proposed combined Appearance-Shape based system can be useful in noise robust face recognition system especially with different lighting variations. This system is also useful when the structure of faces are changed by affecting different environmental noises. In this case, shape based performance decreases but on that time appearance based performance retain the proposed system accuracy. This system accuracy can only degrades when both appearance and shape based features are affected by environment noises. Some methods can be introduced to protect this situation and it will be the further work of this system.

References

- [1] Dominique Ginhac, Fan Yang, Xiaojuan Liu, Jianwu Dang, and Michel Paindavoine, "Robust Face Recognition System Based on a Multi-Views Face Database," Recent Advances in Face Recognition, 1st ed., In-The, 2008, pp. 27-38.
- [2] S. Akamatsu, "The research trend of face recognition by computer," The Journal of IEICE, vol. 80, no. 3, pp. 257-266, 1997.
- [3] O. Hasegawa, S. Morisima, and M. Kameko, "Information processing of face," IEICE Trans. D-II, vol. 80, no. 8, pp. 2047-2065, 1997.
- [4] M. Yachida, Robot vision, Syokoudou Publishing, 1993.
- [5] A. Samal, and P.A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: a survey," Patt. Recogn. 25 (1), pp. 65-77, 1992.
- [6] D. Valentin, H. Abdi, and A.J. O>Toole, "G.W. Cottrell, Connectionist models of face processing: a survey," Patt. Recogn. 27 (9), pp. 1209-1230, 1994.
- [7] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey," Proc. IEEE 83 (5), pp. 705-740, 1995.
- [8] J. Zhang, Y. Yan, and M. Lades, "Face recognition: eigenface, elastic matching, and neural nets," Proc. IEEE 85 (9), pp. 1423-1435, 1997.
- [9] I. Craw, N. Costen, T. Kato and S. Akamatsu, "How should we represent faces for automatic recognition?," IEEE Trans. Patt. Anal. Mach. Intell. 21 (8), pp. 725-736, 1999.
- [10] A. M. Burton, V. Bruce, P.J.B. Hancock, "From pixels to people: a model of familiar face recognition," Cognitive Sci. 23, pp. 1-31, 1999.
- [11] Yang M., Kriegman, D., and Ahuja, N., "Detecting faces in images: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 1, pp. 34-58, 2001.
- [12] Féraud R., Bernier O. J., Viallet, J., and Collobert, M., "A fast and accurate face detector based on neural networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 1, pp. 42-53, 2001.
- [13] Rowley H., Baluja S., and Kanade T., "Neural network-based face detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, 1998.
- [14] Turk M., and Pentland A., "Eigenfaces for recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp 71-86, 1991.
- [15] Slimane M., Brouard T., Venturini G., and Asselin de Beauville J. P., "Unsupervised learning of pictures by genetic hybridization of hidden Markov chain," Signal Processing, Vol. 16, No. 6, pp. 461-475, 1999.
- [16] Phillips P., Grother P., Micheals R., Blackburn D., Tabassi E., and Bone J., "Face recognition vendor test 2002," Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG), pp. 44, 2003.
- [17] Tsalakanidou F., Tzovaras D., and Strintzis M., "Use of depth and colour eigenfaces for face recognition," Pattern Recognition Letters, vol. 24, no. 9-10, pp. 1427-1435, 2003.
- [18] Beumier C., and Acheroy M., "Face verification from 3D and grey level clues," Pattern Recognition Letters, vol. 22, no. 12, pp. 1321-1329, 2001.
- [19] Hehser C., Srivastava A, and Erlebacher G., "A novel technique for face recognition using range imaging," Proceedings of 7th International Symposium On Signal Processing and Its Applications (ISSPA), pp. 201-204, 2003.
- [20] Lu X., Coalbry D. and Jain A., "Three-dimensional model based face recognition," Proceedings of the 17th International Conference on Pattern Recognition (ICPR), vol. 1, pp. 362-366, 2004.
- [21] Bowyer K., Chang K., and Flynn P., "A survey of 3D and multi-modal 3D+2D face recognition," Proceedings of the 16th International Conference on Pattern Recognition, pp. 358-361, 2002.
- [22] Kobayashi K., "Mobile terminals and devices technology for the 21st century," NEC Research Development, vol. 42, no. 1, pp. 15-24, 2001.
- [23] Yeh Y., and Lee C., "Cost effective VLSI architectures and buffer size optimization for full search block matching algorithms," IEEE Transactions on VLSI Systems, vol. 7, no. 3, pp. 345-358, 1999.
- [24] Xiaoguang Lu, "Image Analysis for Face Recognition," Personal notes, Dept. of Computer Science & Engineering, Michigan State University, East Lansing, MI, May 2003.
- [25] Stephen Milborrow, and Fred Nicolls, "Locating Facial Features with an Extended Active Shape Model," MSc dissertation, Faculty of Engineering, University of Cape Town, South Africa, Nov. 2007.

- [26] R. Herpers, G. Verghese, K. Derpains, and R. McCready, "Detection and tracking of face in real environments," IEEE Int. Workshop on Recognition, Analysis and Tracking of Face and Gesture in Real- Time Systems, Greece, pp. 96-104, 1999.
- [27] Ruba Kathavarayan, and Murugesan Karuppasamy, "Preserving Global and Local Features for Robust Face Recognition under Various Noisy Environments," Int. Journal of Image Processing (IJIP), vol. 3, no. 6, pp. 328-340, 2010.
- [28] Matthew Turk and Alex Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, vol. 3, no. 1, Massachusetts Institute of Technology, 1991.
- [29] Matthew Turk and Alex Pentland, "Face Recognition Using Eigenfaces, Vision And Modeling Group", The Media Laboratory, Massachusetts Institute of Technology, 1991.
- [30] Jon Shlens, "A Tutorial on Principal Component Analysis, Derivation, Discussion and Singular Value Decomposition", Version 1, 25 March, 2003.
- [31] Alok Sharma, Kuldip K. Paliwal, and Godfrey C. Onwubolu, "Splitting Technique Initialization in Local PCA", Journal of Computer Science 2(1), pp.53-58, 2006.
- [32] James A. Freeman, and David M. Skapura, Neural Networks, Algorithms, applications and Programming Techniques, Addison-Wesley Pub (Sd), June, 1991.
- [33] Md. Rabiul Islam, and M. Abdus Sobhan, "Improving the Convergence of Backpropagation Learning Neural Networks Based Bangla Speaker Identification System for Various Weight Update Frequencies, Momentum Term and Error Rate", International Conference on Computer Processing of Bangla (ICCPB 2006), Independent University, Bangladesh, pp. 27-34, 2006.
- [34] Niall A. Fox, Brian A. O'Mullane, and Richard B. Reilly, "VALID: A New Practical Audio-Visual Database, and Comparative Results," Lecture Notes in Computer Science, Audio- and Video-Based Biometric Person Authentication, vol. 3546, pp. 201-243, 2005.