A Hardware/Software Co-design model for Face Recognition using Cognimem Neural Network chip

Santu Sardar, Gaurav Tewari, K. A. Babu
Scientist, DRDO
Hyderabad, India
{santusardar, gaurav, anandbabu}@anurag.drdo.in

***Abstract*— Automated Face recognition is a technique employed in wide-range of practical applications, which include access control, identification systems, surveillance and law enforcement applications to name a few, and future improvements promise to spread the use of face recognition further still. Radial Basis Function Networks (RBFN) have proven effective approach for face recognition. Software implementations fail to capture the inherent parallelism of RBFN and incur long training time. Although, hardware implementations can speed up the training process, they may lead to inflexible solution. The main challenges of Face Recognition today are broad lightning variations, handling rotation in depth, together with personal appearance changes. A highly accurate face recognition system requires a number of complex sub-operations to be performed. To balance the flexibility of the involved sub-modules and to achieve high accuracy in face recognition, we propose an embedded computing system, consisting of a processor and dedicated fully parallelized Cognimem Neural Network chip based board. We will also identify the optimized algorithm for each of the involved sub-operations. Results obtained after testing our proposed system, with standard databases, show promising performances in terms of Recognition accuracy, False acceptance rate (FAR), False rejection rate (FRR), training time and testing time.**

Keywords- Radial basis function Networks (RBFN), Cognimem Neural Network chip, accuracy, FAR, FRR.

I. INTRODUCTION

In the recent years, the user authentication is becoming increasingly popular due to the security control requirement in identity authentication, access control, and surveillance, etc. Face recognition, among other biological authentication, is an amicable alternative as authentication can be done without hampering user activities and is economic with low-cost cameras and computers.

Over the past 20 years, extensive research works on various aspects of face recognition by human and machines have been conducted by psychophysicists, neuroscientist and engineering scientists[1, 2, 3, 4]. Automatic face recognition by computer can be divided into two main approaches i.e. content-based and face-based [1, 2]. In content-based approach, recognition depends on the relationship between human facial features such as eyes, mouth, nose, profile silhouettes and face boundary [5, 6, 7]. Every human face has similar facial features; a small deviation in the extraction may introduce a large classification error. Face-based approach attempts to capture and define the face as a whole [7, 8, 9 and 10]. The face is taken as a two-dimensional pattern of intensity variation. In this approach, face is matched with its underlying statistical regularities. Principal Component Analysis (PCA) [8, 9, 11] is proved to be an effective face-based approach.

Sirovich and Kirby first proposed using Karhunen- Loeve (KL) transform to represent human faces [11]. They proposed the idea of eigenfaces. Turk and Pentland developed a face recognition system using PCA [12]. But common PCA-based methods suffer from two limitations, namely, poor discriminatory power and large computational load. Swets and Weng [13] also observed this drawback of PCA approach and improved the discriminability of PCA by adding Linear Discriminant Analysis (LDA). But, for precise result, large number of samples for each class is required.

On the other hand, O'Toole et al. [14] showed that the eigenvectors with large eigenvalues are not the best for distinguishing face images. However, O'Toole et al. have not addressed much on the selection methods of eigenvectors.

Another problem in PCA method is the high computational load to find eigenvectors. To reduce the number of computations, we applied PCA on one of the sub-band of wavelet transformed image, for feature extraction.

The result in [12] show that 3-level wavelet has a good performance in face recognition applications. This method works on lower resolution, 16 x 16, instead of the original image resolution of 128 x 128. Therefore, this method reduces the computational complexity significantly when the resolution of training image is larger than 16 x 16, which is expected to be the case for a number of real-world applications. Moreover, experimental results demonstrate that, applying PCA on DWT sub-image gives better recognition accuracy and discriminatory power than applying PCA on the original image.

RBFN can be used for a wide range of application primarily because it can approximate any regular pattern and its training is faster than that of a multi-layer perceptron network. This faster learning speed comes from the fact that RBFN has just two layers of weights and each layer can be determined sequentially. But, software implementation of RBFN based ANN classifier has the problem of long training and testing time which will depend on the corresponding inputs. So, feature vectors are classified using Cognimem ANN chip based board which has 1024 Neurons operating in parallel. In this board, the recognition time is independent of the number of neurons that is being used. Results show an

978-1-61284-861-7/11/$26.00 ©2011 IEEE

excellent performance improvement because of the use of Cognimem ANN chip. Face detection & pre-processing methods are proposed to improve the performances further as is evident from the comparative results. The Data-flow diagram of our implementation is shown in fig. 1.

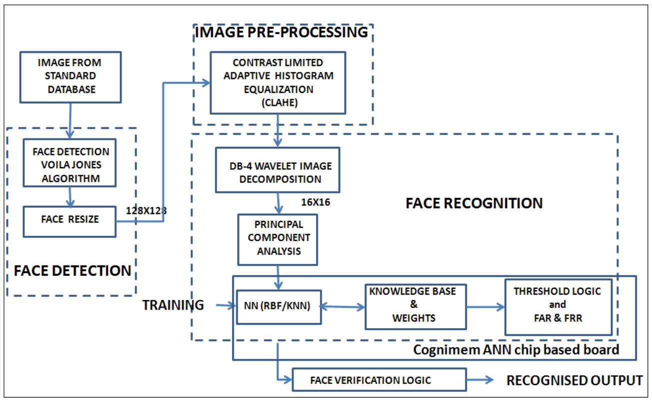

Figure 1. Data-flow diagram of ANN board based Face recognition system

This paper is organized as follows: Section II describes the face detection system and image pre-processing methods. Section III explains the application of DWT in face image processing. Section IV describes PCA, which is used for dimensionality reduction. Section V explains the features of the Cognimem ANN chip based board, which is used as a classifier. Performance results are given in section VI and the conclusions are made in section VII.

II. FACE DETECTION & IMAGE PRE-PROCESSING

The Images from standard database are first passed through the modified AdaBoost (Adaptive Boosting) framework for face detection. The detected face images are cropped and resized to 128x128 pixel size.

Before giving the data to the DWT module, histogram equalization or Contrast-limited adaptive histogram equalization (CLAHE) can be used for contrast enhancement of the face image. The histogram equalization approach is to design a transformation *T(.)* such that the gray values in the output is uniformly distributed in [0, 1]. This usually results in an enhanced image, with an increase in the dynamic range of pixel values. The CLAHE algorithm partitions the images into contextual regions and applies the histogram equalization to each one. This evens out the distribution of used grey values and thus makes hidden features of the image more visible. The full grey spectrum is used to express the image. CLAHE overcomes the limitations of standard histogram equalization.

Results are compared for face detection and recognition system performance without histogram equalization, with histogram equalization and with CLAHE. After analyzing the results, CLAHE is proposed in our face recognition system. Experiments performed for different sizes of the contextual regions reveal that dividing the image into 16 contextual regions (tiles) gives optimum performance for each of the test databases.

III. 2D-DWT ON IMAGES

One of the major issues when using wavelet transforms, is choosing a suitable wavelet filter. The Daubechies family of wavelet filters, developed by Ingrid Daubechies, is perhaps the most popular wavelet filters due to their many desirable characteristics. Daubechies wavelets have the property of having the maximum number of vanishing moments for a given order which makes them suitable for compression applications. Comparing with other member of Daubechies family, Db4 wavelet produces sharper edges and retains more detail, providing a closer resemblance to the original.

The DWT [15], which is based on sub-band coding [16], is found to yield a fast computation of Wavelet Transform [17]. The DWT is computed by successive low-pass and high-pass filtering of the discrete time-domain signal. At each level of decomposition, resolution of the image is reduced by half. To implement 2D-DWT, we have adopted the polyphase architecture. The Lifting [18] and Lattice [14] implementations require fewer computations than conventional polyphase implementation. However, the polyphase implementation can be made more efficient in case of long filters by incorporating techniques like Distributed Arithmetic (DA). Also, the lattice structure cannot be used for all linear phase filters and imposes restrictions on the length of the filters.

The poly phase decomposition of input sample is shown in fig. 2.

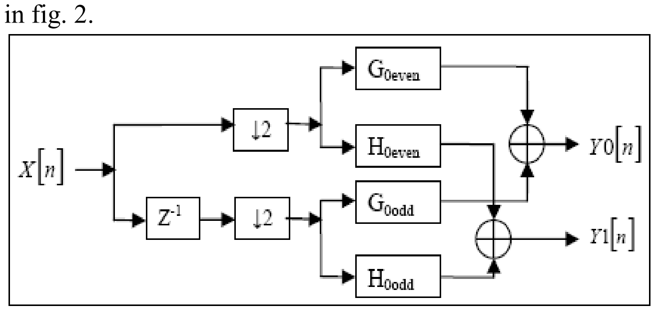

Figure 2. Poly Phase decomposition of Input Sample

In poly phase architecture, the input signal and filter coefficients are split into odd and even samples. The filters with G_{0even} and G_{0odd} are half as long as G_0, as they are obtained by splitting G_0. Since, the even and odd terms are filtered separately, by the even and odd coefficients of the filters, the filters can operate in parallel improving the efficiency.

IV. PRINCIPAL COMPONENT ANALYSIS

After we have the wavelet coefficients of the original face image, the PCA is applied for further data reduction. The idea of PCA is to transform original images into a corresponding eigenfaces [12]. Each eigenface represents certain features of the face, so one can reconstruct the original image from the training set by combining the eigenfaces.

Let $X=[X_1, X2,\ldots, XM]$ be the sample set of the face images, and the Φ_i denote $X_i - X_{mean}$, which the X_{mean} is average matrix. The covariance matrix C is calculated according to:

$$C = \frac{1}{M}\sum_{i=0}^{M} \Phi_i{}^*\Phi_i{}^t$$

According to the above equation, the eigenvectors (eigenfaces) $\boldsymbol{u}$ and the corresponding eigenvalues can be calculated. We only need to calculate the highest or dominant eigenvalues of the covariance matrix, because the higher the eigenvalue, the more characteristic features of a face does the particular eigenvector describe. After the image is transformed into its eigenface components, the eigen feature weight vectors are calculated as follows:

$$w_k = u_k^T (X_k - \overline{X}) \quad (k=1,2,\ldots\ldots.M)$$

These weight vectors are used as the input feature vectors for the Neural Network classifier.

V. COGNIMEM ANN CHIP BASED BOARD

Cognimem ANN chip based board performance is compatible to available off the shelf Neural Network Devices [19]. It has 1024 Neurons operating in Parallel & it can learn and recognize patterns of up to 256 bytes. The recognition time is independent of the number of neurons that are being used. It also has the provision to save and restore the acquired knowledge.

Fig. 3 shows the simplified functional blocks of the Neural Network chip. Feature vector as an input are being fed to all the Neurons and these 1024 Neurons operates in parallel. A Neuron Bus consolidates the results of all the Neurons. During training phase, the results of each of the neuron are examined by all other neurons to adjust the influence field. During recognition phase, the results of Neuron are read on the Neuron Bus. Search & Sort Circuit builds a table of results. Distance as calculated by each of the firing neurons is sorted in ascending order with corresponding category.

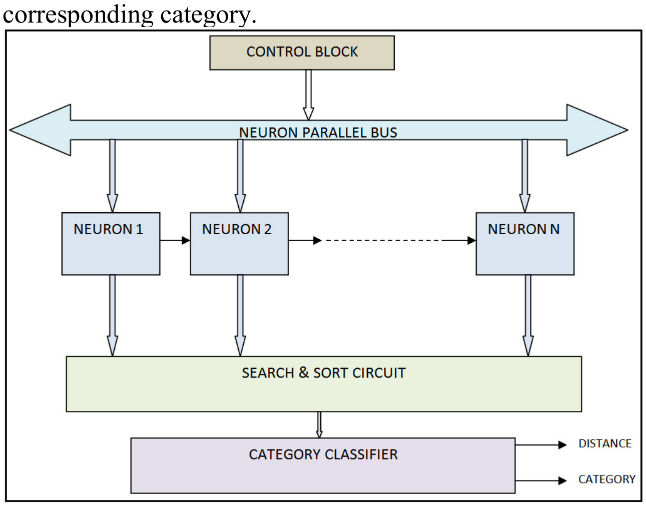

Figure 3. Functional block diagram of the Neural Network chip

Category Classifier declares the match as recognized, unrecognized and recognized with uncertainty depending on the selected classifier (RBF or k-nearest neighbor).

Fig. 4 shows the simplified architecture of a Neuron. Neuron signature block contains the signature of the pattern, for whose recognition, the neuron has been committed. Feature vector is weighed against the stored signature and distance is calculated in one of the two ways depending on the selected Norm, L1 or Lsup. L1 emphasizes the drift of the sum of all the components between the input vector and stored signature whereas Lsup emphasizes the largest drift of the same component between the input vector and stored signature. The selection of distance type is dependent on the end application. In our case, we are using L1 Norm. Distance comparison block compares the distance with Active Influence Field (AIF) of Neuron. If the distance is less than the AIF value, the feature is declared as "Recognized", else "Not recognized".

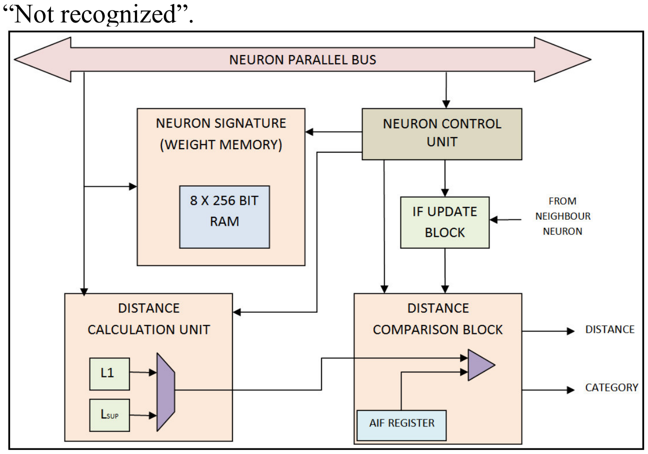

Figure 4. Simplified Architecture of a Neuron

VI. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed system, we used the following face image databases: Yale, cropped Yale [21], FERET [22] and ORL [23]. Images from the standard databases are first passed through the face detection module which crops the face image from the input image and then the cropped face is resized to 128x128 pixels. Cropped and resized face images are either pre-processed or directly fed as input to the DWT module. Two types of image pre-processing techniques are compared, namely histogram equalization and CLAHE. Db4 3-level DWT is performed over the pre-processed images and the HH sub-band of third level decomposition is passed as input to PCA module for dimensionality reduction. The feature vector for each image consists of 35 principal components corresponding to the highest 35 eigen values obtained from PCA module. The feature vector corresponding to each image is given as input to the Radial basis Function Network classifier of the ANN Board.

The Graph in Fig. 5 shows the relationship between recognition accuracy and number of principal components to be used as feature vector, obtained for different face databases.

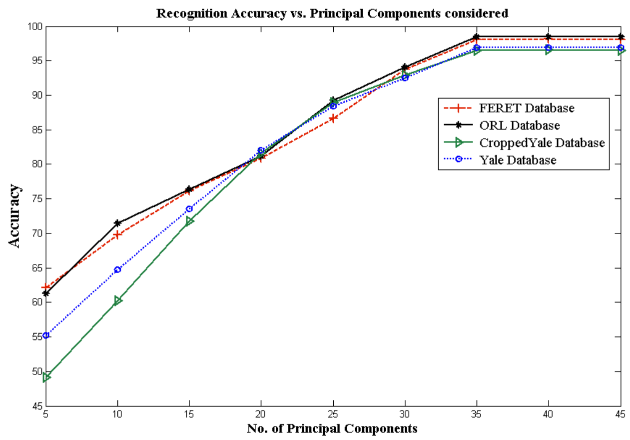

Figure 5. Recognition Accuracy vs. Number of Principal components

Analysis of the Graph shows that, taking more than 35 components doesn't have significant impact on the accuracy but increases the calculation overhead. On the other hand, reduction of the number of principal components reduces the recognition performance.

The two standard biometric measures to indicate the identifying power are False Rejection Rate (FRR) and False Acceptance Rate (FAR). FRR (Type I Error) and FAR (Type II Error) are inversely proportional measurements. FAR is fraction of the falsely accepted patterns divided by the number of all impostor patterns. The fraction of the number of rejected client (persons known by the system) patterns divided by the total number of client patterns is called False Rejection Rate (FRR).

For evaluating our implementation, we have used training time, testing time, recognition accuracy, FAR & FRR as Performance evaluation metric.

The experiments were performed by taking four different face databases. For each database results are obtained for different pre-processing methods. First we obtained the results without any pre-processing. Then we applied Histogram Equalization to the Database. Lastly, we applied CLAHE to the input images. The image was divided into 16 contextual regions along each axis. Table I-IV compares the results for above three cases.

Table I compares the result for Yale Face Database. Eighty images, each with resolution 64x64 were taken. Similarly Table II compares the results for Cropped Yale Face Database containing 300 images, each having resolution 128x128. Table III contains the comparative results for ORL Face Database. Total 400 images were taken each having resolution 128x128. Table IV shows the results obtained by using 216 images of FERET Face Database, each having resolution 256x384.

Comparison of our implementation with other published approaches is given in Table 5. We have also included the results obtained with different pre-processing methods for comparison purpose.

TABLE I. RESULT COMPARISON FOR VARIOUS PRE-PROCESSING METHODS (YALE DATABASE)

Image Pre-processing	Training time (ms)	Testing time (ms)	Accuracy (%)	FAR	FRR
None	538.28	9.5	89.45	0.09	0.40
Histogram Equalization	576.18	9.7	90.27	0.06	0.15
CLAHE	871.63	15	96.995	0	0.08

TABLE II. RESULT COMPARISON FOR VARIOUS PRE-PROCESSING METHODS (CROPPED YALE)

Image Pre-processing	Training time (ms)	Testing time (ms)	Accuracy (%)	FAR	FRR
None	4105	16.7	89.00	0.11	0.49
Histogram Equalization	4139.5	17.25	91.51	0.07	0.23
CLAHE	6022	22.35	96.55	0	0.06

TABLE III. RESULT COMPARISON FOR VARIOUS PRE-PROCESSING METHODS (ORL DATABASE)

Image Pre-processing	Training time (ms)	Testing time (ms)	Accuracy (%)	FAR	FRR
None	3124	16.4	92.00	0.1	0.26
Histogram Equalization	3285	16.49	95.70	0.1	0.19
CLAHE	4484	21.48	98.48	0	0.05

TABLE IV. RESULT COMPARISON FOR VARIOUS PRE-PROCESSING METHODS (FERET DATABASE)

Image Pre-processing	Training time (ms)	Testing time (ms)	Accuracy (%)	FAR	FRR
None	3124	16.4	87.90	0.15	0.37
Histogram Equalization	3285	16.49	92.30	0.09	0.22
CLAHE	4484	21.48	98.15	0.01	0.06

TABLE V. PERFORMANCE COMPARISON WITH OTHER FRS IMPLEMENTATIONS

Feature Extraction method	Classification method	Accuracy (%)			
		YaleFaces	CroppedYale	ORL	Feret
Proposed Implementation					
DB4 DWT+PCA	RBF ANN (Cognimem Board)	89.45	89.00	92.00	87.90
Histogram Equalization+DB4 DWT+ PCA	RBF ANN (Cognimem Board)	90.27	91.51	95.70	92.30
CLAHE (Tiles: 16) + DB4 DWT+PCA	RBF ANN (Cognimem Board)	96.995	96.55	98.48	98.15
Existing Implementations					
Wavelet Transform [20]	MLP	89.45	---	94.25	---
PCA [20]	MLP	81.27	---	90.00	---
DCT [20]	MLP	---	---	97.45	---
Wavelet Transform [20]	NFL (18)	---	---	95.40	---
LDA [20]	MLP	89.24	---	93.90	---
DB4+PCA [20]	MLP	90.35	---	97.68	---

VII. CONCLUSION

A hardware/software co-design model using wavelet based PCA method and ANN chip is developed so as to overcome the limitations of the original PCA method and software implementation of ANN classifier respectively. We have utilized a Cognimem neural network chip based board in order to carry out the classification of faces. We adopted RBF classifier in the board which is fed by the reduced input feature vectors generated by the C++ based system consisting of face detection, resizing, preprocessing, 2D-DWT and PCA methods. Our co-design approach has yielded better recognition accuracy than the existing methods. By using the proposed algorithms for image pre-processing (CLAHE) and feature extraction (Db4 DWT & PCA), we achieved significant performance improvement. Further, the training and testing time are reduced largely due to the inherent parallelism of the Cognimem ANN chip based board.

REFERENCES

[1] R. Chellappa, C. L. Wilson and S. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of the IEEE*, May 1995, Vol. 83, No. 5, 705-740.

[2] G. Chow and X. Li, "Towards a system for automatic facial feature detection," *Pattern Recognition*, 1993, Vol. 26, No. 12, 1739-1755.

[3] Pawan Sinha, Benjamin Balas, Yuri Ostrovsky and Richard Russell, "Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About, " *Proceedings of the IEEE*, 2006, Vol. 94, No. 11.

[4] Zhang Yankun and Liu Chongqing, "Efficient face recognition method based on DCT and LDA," *Journal of Systems Engineering and Electronics*,2004,15(2):211-216.

[5] F. Goudail, E. Lange, T. Iwamoto, K. Kyuma and N. Otsu, "Face recognition system using local autocorrelations and multiscale integration," *IEEE Trans. PAMI*, 1996, Vol. 18, No. 10, 1024-1028.

[6] K. M. Lam and H. Yan, "Locating and extracting the eye in human face images", *Pattern Recognition*, 1996, Vol. 29, No.5 771-779.

[7] D. Valentin, H. Abdi, A. J. O'Toole and G. W. Cottrell, "Connectionist models of face processing: A Survey," *Pattern Recognition*, 1994, Vol. 27, 1209-1230.

[8] A. L. Yuille, P. W. Hallinan and D. S. Cohen," Feature extraction from faces using deformable templates," *Int. J. of Computer Vision*, 1992, Vol. 8, No. 2, 99-111.

[9] M. Kirby and L. Sirovich," Application of the Karhunen- Loeve procedure for the characterization of human faces" *IEEE Trans. PAMI.*, 1990, Vol. 12, 103-108.

[10] Bahadir K. Gunturk, Aziz U. Batur, Yucel Altunbasak, Monson H. Hayes and Russell M. Merscreau, "Eigenface-Domain Super-Resolution for Face Recognition," *IEEE. Trans. on Image Processing*, 2003, Vol. 12, No. 5, 597-606.

[11] M. V. Wickerhauser, Large-rank "approximate component analysis with wavelets for signal feature discrimination and the inversion of complicated maps," *J. Chemical Information and Computer Sciences*, 1994, Vol. 34, No. 5, 1036-1046.

[12] D. L. Swets and J. J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Trans. PAMI*, 1996, Vol. 18, No. 8, 831-836.

[13] K.Fukunaga, "introduction to Statistical Pattern Recognition," Academic press.

[14] Juyoung Kim and Taegeun Park, "High Performance VLSI Architecture of 2D Discrete Wavelet Transform with Scalable

Lattice Structure" World Academy of Science, Engineering and Technology 54-2009.

[15] J.T. Chien, "Discriminant Waveletfaces and Nearest Feature Classifiers for Face Recognition", *PAMI.* , 2002.

[16] M. Alam, C.A. Rahman, W. Badawy, G. Jullien, "Efficient Distributed Arithmetic Based DWT Architecture for Multimedia Applications", Proceedings of the 3rd IEEE International Workshop on System-on-Chip for Real-Time Applications, pages 333 -336, June 2003.

[17] F. Marino, "Efficient high-speed/low-power pipelined architecture for the direct 2-D discrete wavelet transform," IEEE Trans. Circuits Systems II, Analog Digit. Process., vol.47, no.12, pp.1476-1491, 2000 .

[18] P. Wu and L.-G. Chen, "An efficient architecture for two-dimensional discrete wavelet transform," IEEE Trans. Circuits Systems, Video Technol., vol.11, no.4, pp.536-545, 2001.

[19] http://www.general-vision.com/product_CM1K.htm.

[20] Masoud Mazloom and Saeed Ayat, "Combinational Method for Face Recognition: Wavelet, PCA and ANN," IEEE Computer Society, 2008.

[21] Yale Face Database: http://cvc.yale.edu/projects/yalefaces/yalefaces.html.

[22] FERET Face Database: http://faces.nist.gov/colorferet

[23] ORL Face Database: http://www.orl.co.uk/facedatabase.html.