An Intersecting Cortical Model Based Framework for Human Face Recognition

Ahmed G. Mahgoub[1,2,3], Amira A. Ebeid[1], Hossam-El-Deen M. Abdel-Baky[1], and El-Sayed A. El-Badawy[1,4]

[1]*Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Alexandria 21544, Egypt.*
[2]*Département de Génie Électrique et Informatique, Université Laval, Québec, QC, G1K 7P4, Canada.*
[3]*Research & Development Department, Palamar-Misr, Giza, Egypt.*
[4]*Alexandria Higher Institute of Engineering & Technology, Alexandria 21311, Egypt.*

ABSTRACT

This paper introduces a novel method for human face recognition based on a simplified approach for the Pulse Coupled Neural Network (PCNN) Algorithm. The face image is introduced to the Intersecting Cortical Model (ICM) to be iterated 200 times, and then the time signals for the faces are compared to make a decision. Experimental results for human face recognition confirm that the proposed method lends itself to higher classification accuracy relative to existing techniques.

Keywords: Correlation coefficient, face recognition, ICM, minimum-distance classifiers, PCNN.

1. INTRODUCTION

Face recognition may seem an easy task for humans, and yet computerized face recognition system still cannot achieve a completely reliable performance. The difficulties arise due to large variation in facial appearance, head size, orientation and change in environment conditions [1]. Automatic human face recognition has received substantial attention from researchers in biometrics, pattern recognition, and computer vision communities. The machine learning and computer graphics communities are also increasingly involved in face recognition. There is a large number of commercial, securities, and forensic applications that require the use of face recognition technologies [2].

Two major approaches are used for machine identification of human faces; geometrical local feature based methods, and holistic template matching based systems. Also, combinations of these two methods, namely hybrid methods, are used [3].

2. THE PULSE COUPLED NEURAL NETWORK

The PCNN is to a very large extent based on the Eckhorn model shown in Fig. 1 except for a few minor modifications. The early experiments demonstrated that the PCNN could process images such that the PCNN output was quite similar for images that were shifted, rotated, and scaled [4].

There are three parts to the model neuron: the dendritic

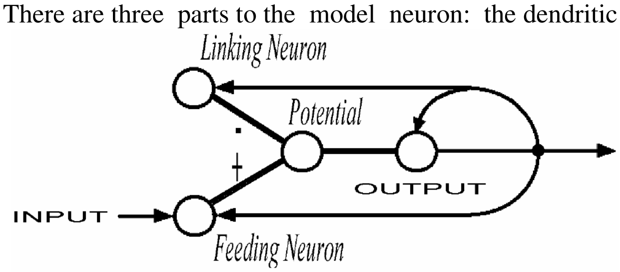

Fig. 1. The Eckhorn-type neuron.

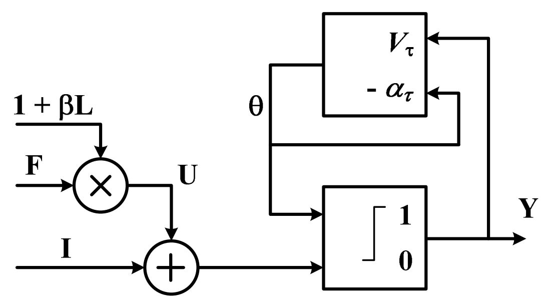

Fig. 2. The basic PCNN neuron.

tree, the linking modulation, and the pulse generator. The dendritic tree is divided into two principal branches in order to make two distinct inputs to the linking part. They are the primary input, termed the feeding input *F*, and the secondary input, termed the linking input *L*. The feeding receives an external stimulus as well as local stimulus, while the linking receives only the local stimulus. The feeding and the linking are combined in a second-order fashion to create the membrane voltage, *U*, that is then compared to a local threshold, θ. The basic simplified structure of the pulse-coupled Neuron is shown in the Fig. 2 [16]. Each input is a weighted sum from the synaptic connections on its dendritic branch. The synapses themselves are modeled as leaky integrators. An electrical version of a leaky integrator is a capacitor and a resistor in parallel, charged by a brief voltage pulse and decaying exponentially. Likewise, when a synapse receives a pulse, it is charged and its output amplitude rises steeply. The amount of rise depends on the amplitude-gain factor assigned to the synapses. It then decays exponentially

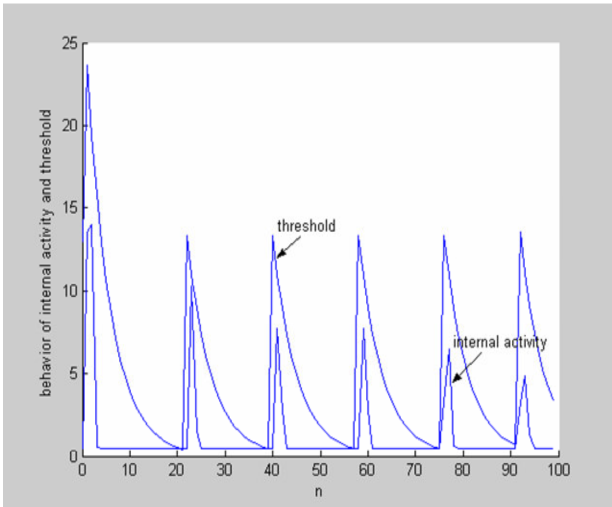

Fig. 3. The behavior of the internal activity and the threshold of the neuron.

Figure 3 shows how the internal activity rises until it becomes larger than the threshold value, then the neuron fires and the threshold sharply increases then begins its decay until once again the internal activity becomes larger than the threshold.

3. THE INTERSECTING CORTICAL MODEL

The intersecting cortical model (ICM), which was especially designed for image processing, is a model based on neural network techniques. The ICM was derived from several visual models especially Eckhorn's model. It is a simplified PCNN model introduced by J. Kinser [5] where there are no linking neurons. Since in the ICM there is only one time consuming calculation, the ICM is faster than the PCNN and hence it may be preferable to be used in many cases.

The state oscillators of all the neurons are represented by a 2D array F (the internal neurons states; initially $F_{ij}=0, \text{ for } \forall\, i,j$) and the threshold oscillators of all the neurons by a 2D array θ (initially $\theta_{ij}=0, \text{ for } \forall\, i,j$). Thus the neuron has state $F_{i,j}$ and threshold $\theta_{i,j}$. They are computed from,

$$F_{ij}[n+1]=fF_{ij}[n]+S_{ij}+\sum_{kl}W_{ijkl}Y_{kl}[n] \tag{1}$$

$$\theta_{ij}[n+1]=g\,\theta_{ij}[n]+hY_{ij}[n+1] \tag{2}$$

where S_{ij} is the stimulus (the input image, scaled so that the largest pixel value is 1.0), Y_{ij} is the firing state of the neuron (Y is the output image), f, g, and h are scalars (examples of their values are 0.9, 0.8, and 20.0), W_{ijkl} is the connection function through which the neurons communicate and n = 1, 2, …, N is the iteration number.

according to its time constant. These postsynaptic signals are summed to form the total signal out of that branch of the dendritic tree [17].

The scalars f and g are decay constants and thus less than

$$Y_{ij}[n+1]=\begin{cases}1, & \text{if } F_{ij}[n+1]>\theta_{ij}[n]\\ 0, & \text{otherwise}\end{cases} \tag{3}$$

The output of the ICM is the binary images Y[n] obtained after a number of n neural pulse iterations. Hence the images are called pulse images [6].

4. FACE RECOGNITON USING ICM-200

ICM-200 means using the ICM algorithm with 200 iterations for each face. To check the utility of our proposed algorithm experimental studies are carried out on the ORL database images of Cambridge[15]. Images from this database have been used to evaluate the performance of the proposed method. None of the 10 samples are identical to each other. They vary in position, rotation, scale and expression. In this database each person has changed his face expression in each of 10 samples (open/close eye, smiling/not smiling). For some individuals, the images were taken at different times; varying facial details (glasses/no glasses).

Eleven persons are introduced to the system. The persons are named as Fxx_yy where xx is the person and yy is the number of face of this person, example: F13_8 means face 8 for person number 13, and so on.

Matching Using Correlation Coefficients

In its simplest form, the correlation between *f(x)* and *w(x)*with sizes M and J, respectively, is

$$c(x)=\sum_{s} f(s)w(x) \quad \text{for } x = 0,\ 1,\ 2,\ \ldots,\ M-1 \tag{4}$$

This correlation function has the disadvantage of being sensitive to changes in amplitude of f and w. For example, doubling all values of f doubles the value *c(x).* An approach used to overcome this difficulty is to perform matching via the correlation coefficient, which is defined as

$$\gamma(x)=\frac{\sum_{s}[f(s)-\bar{f}(s)][w(x+s)-\bar{w}]}{\left\{\sum_{s}[f(s)-\bar{f}(s)]^{2}\sum_{s}[w(x+s)-\bar{w}]^{2}\right\}^{\frac{1}{2}}} \tag{5}$$

where x = 0, 1, 2, …, M - 1, $\bar{w}$ is the average value of w, and $\bar{f}$ is the average value of f [7].

Averaging the time signals for the faces

Suppose that each person class w_j is characterized by a mean vector m_j. That is, we use the mean time signal of each person as being representative of that class of time signals [8]:

$$m_j = \frac{1}{N_j}\sum_{x \in w_j} x \qquad j = 1,2,3,......W \qquad (6)$$

where N_j is the number of chosen time signals from class w_j, and the summation is taken over these time signals, x is a face time signal, and *W* is the number of pattern classes(persons). The time signals are added for the first 8 faces, and are averaged, and the last 2 faces are left as a 2

Fig. 4. The faces introduced to the system, with the last four faces for the same person [15].

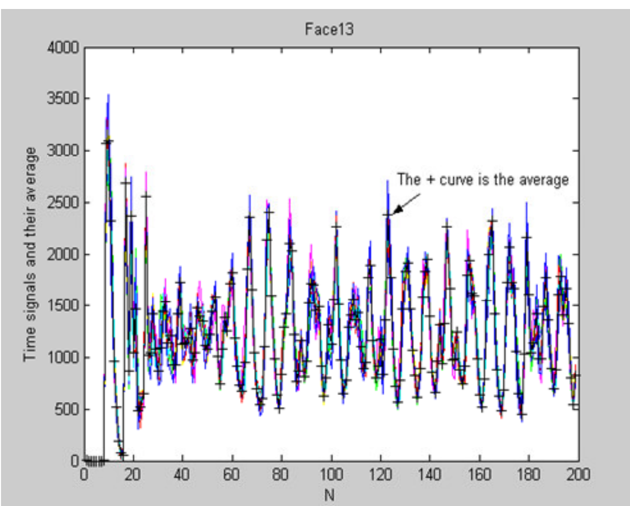

Fig. 5. Time signals for F13 and the average time signal shown with plus signs.

different inputs for the same person to the recognition algorithm. Figure 5 shows the time signals of the first 8 faces on the same graph, and the mean is plotted using the plus sign. To test the performance for a complete face recognition system based on the ICM-200 technique, the 10 faces for each person shown in the first column in Table I are correlated with the average of the persons shown in the first row of this table and the correlation coefficient is calculated.

The diagonal is the *minimum correlation coefficient* between the average and the 10 faces of the same person (shown bolded and underlined); while the other values in the table are the *maximum correlation coefficient* between the 10 faces for the person and the average of the person in the row. The minimum of the diagonal value must be greater than the maximum value in the column to identify the person.

Note that all averages are taken by averaging the sum of the first 8 faces, except the F16 person, where the average is the average of faces 1, 2, 3, 6, 7, 8, 9, 10 because faces 4 and 5 show completely instability with the average.

To detect a face, the input face introduced to the face recognition system will be correlated with all the averages faces stored in the system, and if the correlation coefficient exceeds the minimum value of the diagonal, this face will belong to that person. That is to say that, every average will not accept any face except the correlation coefficient exceeds the minimum of the diagonal.

The input faces are 10 faces for each person, and the number of persons is 11, so there are 110 different input faces. Each input face is correlated with all the averages to be detected; therefore we have 11 correlation processes for each face to be detected. The overall correlation processes of the system are 11 × 110 = 1210. The unrecognized faces are only 46 and the right ones are 1164.

Note that the bolded value that means an unrecognized face can include many unrecognized faces for the same person introduced to this average, and all these unrecognized faces are calculated when getting the overall performance. The overall performance of the system is 96.19%.

Matching Using Minimum-Distance Classifiers

One way to determine the class membership of an unknown pattern vector *x* is to assign it to the class of its closet prototype. Using the Euclidean distance as a measure of closeness reduces the problem to computing the distance measures:

$$D_j(x) = \|x - m_j\| \qquad j = 1,2,3,....W \tag{7}$$

x is then assigned to class w_j, if $D_j(x)$ is the smallest distance. That is, the smallest distance implies the best match in this formulation.

The Euclidean distance between two *n*-dimensional (row or column) vectors x and y is defined as the scalar [8]

$$\begin{aligned} d(x,y) &= \|x - y\| = \|y - x\| \\ &= [(x_1 - y_1)^2 + ... + (x_n - y_n)^2]^{1/2} \end{aligned} \tag{8}$$

Table II is similar to table I, but with using the minimum-distance classifiers as a technique for matching instead of the correlation coefficients.

Table I. Correlation coefficient.

	Av9	Av11	Av12	Av13	Av14	Av15	Av16	Av17	Av18	Av19	Av20
F9	***0.9850***	0.9197	0.9749	0.9829	0.9695	0.9492	0.9155	0.9608	0.9597	0.9478	**0.9735**
F11	0.9206	***0.9981***	0.9478	0.9389	0.9809	0.8617	0.9000	0.9700	0.9451	0.9812	0.9426
F12	0.9823	0.9451	***0.9816***	0.9780	0.9646	0.9324	0.8341	0.9581	0.9624	0.9573	**0.9627**
F13	0.9828	0.9302	**0.9874**	***0.9830***	0.9682	0.9412	0.9109	0.9653	**0.9876**	0.9708	**0.9607**
F14	0.9749	0.9833	0.9758	0.9757	***0.9843***	0.9263	0.9113	**0.9778**	0.9692	0.9827	**0.9774**
F15	0.9495	0.8622	0.9186	0.9238	0.9140	***0.9665***	0.9040	0.8862	0.9065	0.9096	0.9528
F16	0.9325	0.9146	0.8806	0.9338	0.9309	0.9537	***0.9573***	0.9326	0.9162	0.9399	**0.9567**
F17	0.9576	0.9679	**0.9836**	0.9647	**0.9845**	0.9032	0.9329	***0.9714***	0.9642	0.9750	**0.9607**
F18	0.9519	0.9369	0.9756	**0.9900**	0.9639	0.9265	0.8949	0.9516	***0.9847***	0.9746	0.9485
F19	0.9503	0.9721	0.9593	0.9782	0.9851	0.9107	0.9367	**0.9767**	0.9784	***0.9901***	0.9519
F20	0.9829	0.9761	0.9655	0.9425	0.9741	**0.9761**	0.9543	0.9668	0.9471	0.9642	***0.9556***

Table II. Minimum-distance classifier.

	Av9	Av11	Av12	Av13	Av14	Av15	Av16	Av17	Av18	Av19	Av20
F9	***1.9712***	4.7737	2.6469	2.7082	2.8564	3.7353	5.1969	3.2128	3.6560	3.9565	2.6178
F11	4.5423	***0.5754***	4.0977	4.6452	2.2132	6.1252	5.3396	2.8201	4.3855	2.3992	3.7701
F12	2.5623	4.1849	***2.5011***	2.6371	2.8554	4.3401	6.8146	3.3277	3.5014	3.5845	3.3267
F13	2.4406	4.9767	1.9176	***1.1683***	3.3583	4.2892	5.2301	3.3703	1.9965	2.9990	3.7994
F14	2.9117	2.0713	2.7098	2.9500	***2.0498***	4.7275	5.0608	2.4455	3.5479	2.3273	2.3521
F15	3.7865	6.1404	4.9241	4.8384	4.8723	***3.2435***	5.0919	5.6146	5.4927	5.0828	3.6682
F16	**4.3740**	**4.6700**	**5.7930**	**4.4900**	**3.9300**	**3.5960**	***9.5390***	**4.2980**	**5.1830**	**4.1300**	**3.3323**
F17	3.3312	**2.9824**	**2.1480**	**3.2877**	**2.1034**	5.1528	4.3757	***3.3301***	3.4909	**2.6184**	**3.2673**
F18	4.1165	4.7827	2.8268	**1.8772**	3.7507	4.8175	5.9625	4.0592	***2.1948***	2.9302	4.3430
F19	3.6624	2.3139	3.4092	2.5509	2.0820	4.9938	4.2434	2.5443	2.9990	***1.7030***	3.6248
F20	**2.2242**	**2.6683**	**3.1123**	**2.5885**	**2.5885**	**2.6988**	5.2350	**2.9474**	4.2426	**3.1544**	***3.4139***

The diagonal is the *maximum distance* between the average and the 10 faces of the same person (shown bolded and underlined); while the other values in the table are the *minimum distance* between the 10 faces for the person and the average of the person in the row.

The minimum of the diagonal value must be lower than the maximum value in the column to identify the person. There are 147 missed faces using the minimum distance classifier algorithm, with a recognition rate of 87.85%.

Table III. Comparison with different methods.

Method	Recognition Rate (%)	Reference
Image LDA	78.84	[9]
2D PCA	77.87	[10]
Kernel Fisher	93.94	[11]
ICA-FX	96.36	[12]
Fisher Face	93.7	[13]
ICM-200 using MD classifier	**95.87**	This paper
ICM-200 using correlation coefficient	**96.19**	This paper

But with excluding the 2 faces for person F16 (face 4 and face5) that maintained the instability with the correlation coefficients, the unrecognized faces will be only 49 faces, and the recognition rate can be increased to 95.87%.

In Table III we compare the performance of the proposed system with some known results published in literature; it can be concluded from this table that using ICM-200 gives promising results.

5. CONCLUSION AND FUTURE WORK

A novel method for face recognition using the time signals of the ICM is presented in this paper. The correlation coefficient and the minimum distance classifiers are used for making the decision. The proposed systems show that the rotation, translation, and varying the facial details, leave the face recognition performance relatively unaffected. The recognition rates for the proposed systems are over 95%.

For installing a complete system the time signals of the average are saved, but the time signal for each introduced face must be generated each time the person is introduced to the system, i.e. each introduced face must be iterated to generate its time signal. Although the ICM is considered a fast algorithm compared to the PCNN model, but it still consumes time for generating the time signal for the face because of the 200 iterations it needs to perform to generate a single time signal; this time will be increased when increasing the number of iteration. So further research may be directed towards the implementation of this system on a field programmable gate array (FPGA) technology which will give good results in saving the time of iterations as shown in [14].

6. REFERENCES

[1] J. Daugman, “Face Detection: A Survey,” Computer Vision and Image Understanding, Vol. 83, No. 3, pp. 236-274, Sept. 2001.

[2] Xiaguang Lu, “3D Face Recognition Across Pose and Expression,” Doctor of Philosophy Thesis, Dept. Computer Science and Engineering, Michigan State University, USA, 2006.

[3] Ahmet Bahtiyar Gul, “Holistic Face Recognition by Dimension Reduction,” M.Sc. Thesis, Dept. of Electrical and Electronics Engineering, Middle East Technical University, Iran, 2003.

[4] J. M. Kinser, T. Lindblad, “Image Processing Using Pulse-Coupled Neural Networks”, Springer-Verlag London Limited, ISBN 3-540-76264-7, 1998.

[5] J. M. Kinser, “A simplified Pulse-coupled Neural Network, in: .K. Rogers, D. W. Ruck (Eds.),” Applications and Science of Artificial Neural Networks II, Proceedings of SPIE, Vol.2760,No.3,1996,pp .563 –567.

[6] Ulf Ekblad, Jason M. Kinser, Jenny Atmer, and Nils Zetterlund “The Intersecting Cortical Model in Image Processing,” Nuclear Instruments and Methods in Physics Research A 525, 392-396, 2004.

[7] Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing,* Second edition, Prentice-Hall, Inc., 2002.

[8] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, *Digital Image Processing Using Matlab*, Prentice-Hall, Inc., 2004.

[9] K.Liu,Y.Q.Cheng, J.Y.Yang, “Algebraic Feature Extraction for Image Recognition Based on an Optimal Discrimination Criterion,” Pattern Recognition 26 (6)(1993)903 –911.

[10] J.Yang, J.Y.Yang, “From Image Vector to Matrix: A Straightforward Image Projection Technique —IMPCA vs. PCA,” Pattern Recognition 35 (9)(2002)pp.1997 –1999.

[11] M. H. Yang, “Face Recognition Using Kernel Methods”, Advances of Neural Information Processing Systems,vol.14, MIT Press, 2002, pp.215 –220.

[12] N. Kwak, C. Choi, N. Ahuja, “Face Recognition Using Feature Extraction Based on Independent Component Analysis,” Proceedings of International Conference on Image Processing, vol.2, 2002, pp.337 –340.

[13] Y. Gao, M. K. Leung, “Face Recognition Using Line Edge Map,” IEEE Trans. Pattern Anal. Mach. Intell. 24 (6)(2002)764 –779.

[14] V. Beèanovic, “Signal and Image Processing in Engineering Physics”, Ph.D. Thesis, Royal Institute of Technology (KTH), Stockholm, 2000.

[15] ORL data base.

[16] Alexei N. Skourikhine, Lakshman Prasad, Bernd R. Schlei, “Neural Network for Image Segmentation,” Proc. of Conference on Applications and Science of Neural Networks, Fuzzy Systems and Evolutionary Computation, part of 45th SPIE's International Symposium on Optical Science and Technology, San Diego, Calif., July 31–August 4, 2000.

[17] John L. Johnson, “Pulse-Coupled Neural Nets: Translation, Rotation, Scale, Distortion, and Intensity Signal Invariance for Images,” Applied Optics, vol. 33, No. 26,Sep., 1994.