HUMAN FACE RECOGNITION USING GABOR FILTER AND SELF ORGANIZING MAP NEURAL NETWORK

Dr. Tarik Zeyad **Mohamed Fadhel**
Electrical Engineering Department, College of Engineering,
University of Baghdad

ABSTRACT

This work implements the face recognition system based on two stages, the first stage is feature extraction stage and the second stage is the classification stage. The feature extraction stage consists of Self-Organizing Maps in a hierarchical format in conjunction with Gabor Filters and local image sampling.

The next stage is the classification stage, and consists of self-organizing map neural network; the goal of this stage is to find the similar image to the input image. The proposal method algorithm implemented by using C++ packages, this work is successful classifier for a face database consist of 20 people with six images for each person.

الخلاصة

يستخدم هذا البحث في تمييز الأوجه مرحلتين، الأولى أستخراج مواصفات هذه الأجسام والثانية تصنيف هذه الأوجه. تتم مرحلة أستخراج المواصفات بأستخدام الشبكات العصبية نوع (Self Organizing Map) في بنـــاء هرمي مع فلتر من نوع (Gabor) . الخطوة التالية هو أستخدام الشبكات العصبية والفلتر نوع (Gabor) في تمييز الأوجه وذلك بأستخدام لغة البرمجة C++ ونماذج صور لعشرين شخص مع ستة صور لكل شخص.

KEY WORDS

Gabor filter, SOM (Self Organizing Map).

INTRODUCTION

Face recognition is a very active topic of research work at the current time, for both technical and commercial reasons. Advancements in hardware and software technologies, as well as positive results from algorithm development efforts, have resulted in the increased feasibility of automatically identifying individuals from images of their faces. At the same time, applications in law enforcement, security and commercial identification have created a very real demand for accurate face-based identification of persons.

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. Such difficulties make face recognition one of the fundamental problems in pattern analysis. In recent years there has been a growing interest in machine recognition of faces due to potential commercial application such as film processing, law enforcement, person identification, access control systems, etc [Brunelli, et al., 1993].

A complete conventional human face recognition system should include three stages. The first stage involves detecting the location of face in arbitrary images. The second stage requires extraction of pertinent features from the localized image obtained in the first stage. Finally the third stage involves classification of facial images based on the derived feature vector obtained in the previous stage [Haddania, 1997].
In the field of pattern recognition, the combination of an ensemble of classifiers has been proposed to achieve image classification systems with higher performance in comparison with the best performance achievable employing a single classifier [Haddania, 1997].
A number of image classification systems based on the combination of outputs of different classifier systems have been proposed. Different structures for combining classifier systems can be grouped in three configurations. In the first group, the classifier systems are connected in cascade to create pipeline structure. In the second one, the classifier systems are used in parallel and their outputs are combined named parallel structure. Finally the hybrid structure is a combination of the pipeline and parallel structures. [Ho, et al., 1994].
Finally in this work self-organizing map (SOM) neural network is used as the classifier. Recently SOM neural networks have been found to be very attractive for many engineering problems. An important property of SOM neural networks is that they form a unifying link among many different research fields such as function approximation, regularization, noisy interpolation and pattern recognition.

GABOR WAVELETS

Since the discovery of crystalline organization of the primary visual cortex in mammalian brains by Hubel and Wiesel [Marcelja, 1980], an enormous amount of experimental and theoretical research has greatly advanced our understanding of this area and the response properties of its cells. On the theoretical side, an important insight has been advanced by Marcelja, [Marcelja, 1980] and Daugman, that simple cells in the visual cortex can be modeled by Gabor functions. The Gabor functions proposed by Daugman [Kepeneki, 2001], are local spatial band pass filters that achieve the theoretical limit for conjoint resolution of information in the 2D spatial and 2D Fourier domains.
Dennis Gabor first proposed Gabor functions as tools for signal detection in noise. Gabor filter showed that there exists a "quantum principle" for information; the conjoint time-frequency domain for 1- D signals must necessarily be quantized so that no signal or filter can occupy less than certain minimal area in it. However, there is a trade off between time resolution and frequency resolution. Gabor discovered that Gaussian modulated complex exponentials provide the best trade off. For such a case, the original Gabor elementary functions are generated with a fixed Gaussian, while the frequency of the modulating wave varies [Gabor, 1946].
Gabor filters, rediscovered and generalized to 2D, are now being used extensively in various computer vision applications. Daugman [Kepeneki, 2001] generalized Gabor function to the following 2D form in order to model the receptive fields of the orientation selective simple cells:

$$\Psi_i(\vec{x}) = \frac{\left\|\vec{k}_i\right\|^2}{\sigma^2} e^{-\frac{\|\vec{k}_i\|^2 \|\vec{x}\|^2}{2\sigma^2}} \left[e^{j\vec{k}_i \vec{x}} - e^{-\frac{\sigma^2}{2}} \right] \quad (1)$$

Each Ψ_i is a plane wave characterized by the vector $\vec{k}_i$ enveloped by a Gaussian function, where σ is the standard deviation of this Gaussian. The center frequency of i^{th} filter given by the characteristic wave vector is shown below [Kepeneki, 2001]:

$$\vec{k}_i = \begin{pmatrix} k_{ix} \\ k_{iy} \end{pmatrix} = \begin{pmatrix} k_v \cos \theta_\mu \\ k_v \sin \theta_\mu \end{pmatrix} \quad (2)$$

Having a scale and orientation given by (k_v, θ_μ), the first term in the brackets (1) determines the oscillatory part of the kernel, and the second term compensates for the DC value of the kernel. Subtracting the DC response, Gabor filters becomes insensitive to the overall level of illumination [Kepeneki, 2001].

Since the Gabor wavelet transform is introduced to computer vision area, one of the most important application areas for 2D Gabor wavelet representation is face recognition. In the U.S. Government activity (FERET program) to find the best face recognition system, a system based on Gabor wavelet representation of the face image is performed among other systems on several tests. Although the recognition performance of this system shows qualitative similarities to that of humans by now means, it still leaves plenty of room for improvement. [Yang, et al., 2000].

Gabor filter can capture salient visual properties such as spatial localization, orientation selectivity, and spatial frequency characteristics. Considering these excellent capacities and its great success in face recognition, Gabor features are chosen to represent the face image. Gabor filters are defined as follows:

$$\Psi_{u,v}(\bar{x}) = \frac{\left\|\vec{k}_{u,v}\right\|^2}{\sigma^2} e^{\frac{\|\bar{k}_i\|^2 \|z\|^2}{2\sigma^2}} \left[e^{j\bar{k}_{u,v}z} - e^{\frac{\sigma^2}{2}} \right] \qquad (3)$$

where $k_{u,v} = k_v e^{j\phi_u}$ and $k_v = \dfrac{k_{max}}{f^v}$ gives the frequency

and $\phi_u = \dfrac{u\pi}{8}, \phi_u \in (0, \pi)$ gives the orientation and z=(x,y).

$e^{jk_{u,v}z}$ is the oscillatory wave function, whose real part and imaginary part are cosine function and sinusoid function respectively. In equation (2), v controls the scale of Gabor filter, which mainly determines the center of the Gabor filter in the frequency domain; u controls the orientation of Gabor filters [Gabor, 1946].

The Gabor filters are used with the following parameters: five scales $v \varepsilon \{ 0, 1, 2, 3, 4, 5 \}$ and eight orientations $u \varepsilon \{ 0, 1, 2, 3, 4, 5, 6, 7 \}$ with $\sigma = 2\pi$, $k_{max} = \pi / 2$, and $f = \sqrt{2}$ [Perkins, 1975].

THE PROPOSED FACE RECOGNITION METHOD

The objective of this section is to put a block diagram which stands for the suggested model, convert each block into a suitable mathematical model , and finally define model parameters in order to improve its qualification , **Fig. (1)** shows the proposed face recognition model.

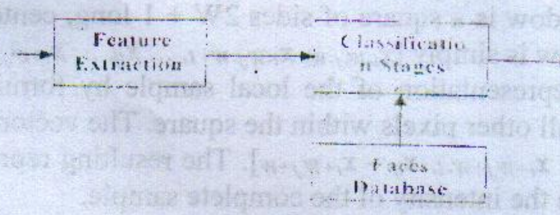
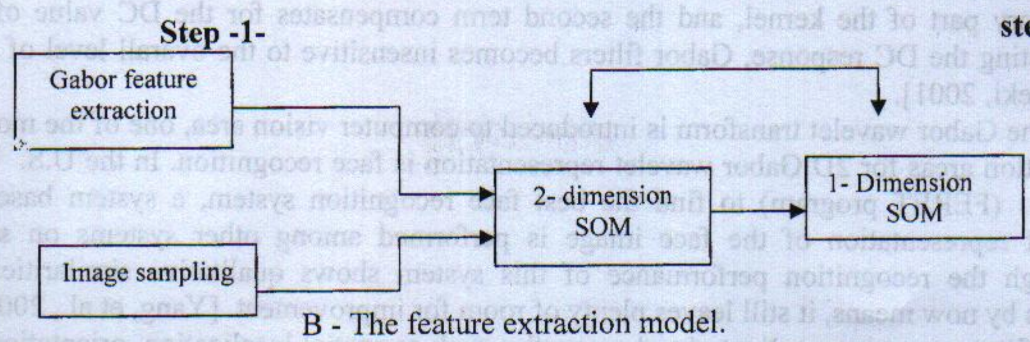
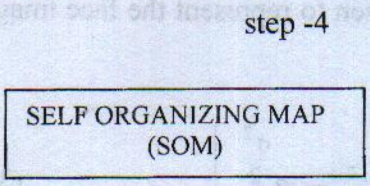


Figure (1) A- Proposed face recognition model.



B - The feature extraction model.



C - Classification model.

Gabor Feature Extraction

The Gabor filter has been successfully used in conjunction with the Multilayer Self-Organizing Maps (MSOM) as a feature extraction for face recognition. Gabor filters are orientation and frequency sensitive band pass filters that provide optimal resolution in both the space and spatial-frequency domains. For these reasons, they are suitable for extracting frequency dependent information like edge features from a small area. Gabor filters have also been used in other pattern recognition systems. The two-dimensional Gabor filter was described by equation (1).

The feature extraction system is a combination of the Gabor filter and three self-organizing maps (SOM). Each time the system is applied to an image the output is a set of 500 feature values, corresponding to 500 x and y position on the original images. The two 2-dimension SOMS are trained by the unsupervised method, the maps contain 10 by 10 nodes, with each node storing the eight values corresponding to a Gabor jet. The 1-dimension SOM has 256 nodes and each map has six values.

Local Image Sampling

Local Image Sampling is a used as an attempt to identify distinguishing features by presenting a small portion of the image to a system such as a Self-Organizing Map. This SOM job is then to identify the presented image portion as a useful feature. Thus by unsupervised learning the SOM will learn what different image portions represent. This is however not an ideal form of input into a one dimensional single layer SOM classifier as used in this work.

There are two methods for obtaining the Image Sample Vectors. These are described as follows:

1- The first method simply creates a vector from a local window on the image using the intensity values at each point in the window. Let x_{ij} be the intensity at the ith row and the j^{th} column of the given image. If the local window is a square of sides 2W + 1 long, centered on x_{ij}, then the vector associated with this window is simply $[x_{i-W,j-W}, x_{i-W,j-W+1}, \dots x_{ij}, \dots x_{i+W,j+W-1}, x_{i+W,j+W}]$.

2- The second method creates a representation of the local sample by forming a vector of the intensity of the centre pixel and all other pixels within the square. The vector is given by $[x_{ij} - x_{i-W,j-W}, x_{ij} - x_{i-W,j-W+1}, \dots x_{ij}, \dots x_{ij} - x_{i+W,j+W-1}, x_{ij} - x_{i+W,j+W}]$. The resulting representation becomes partially invariant to variations in the intensity of the complete sample.

Self Feature Clustering Self-organizing Maps

Self-Organizing Maps (SOM) were initially developed by Kohonen [Fausett, 1994]. The SOM can be arranged to perform classification and feature extraction, although in this section the SOM is only presented as a feature extraction method, in the next section the SOM will be re-examined as a classifier. The theoretical study of the SOM will be explained in the next step, as this knowledge is not essential when comparing feature extraction methods. The SOM is an ideal method to use in conjunction with another feature extraction such as the Gabor filter, or local image sampling as the feature values are topologically organized without using any prior knowledge.

The images used were 256 x 256 pixels inputs to the feature extraction stage. The output of the feature extraction stage is referred to as a feature image. The feature image has reduced dimensions compared to the original image and is used as the input to the classification stage.

The 2-dimension SOMs are trained by the unsupervised method, which is explained in the next step, the maps contain 10 by 10 nodes, with each node storing the eight values corresponding to a Gabor jet. The 1-dimension SOM has 256 nodes and each map has six values.

The main advantages of this system are the reduced training time; data compression rate, distortion tolerant feature extraction, and the ability to train unsupervised the feature extraction stage. Further analysis and expansion in terms of the face database size is required so that more quantitative results can be obtained to further determine the potential of SOM-based feature extraction and model-free networks for face recognition.

CLASSIFICATION

The Self-organizing Map as a Classifier

The Self-Organizing Map (SOM) as introduced by Teuvo Kohonen is one of the most widely used neural network structures in pattern recognition. The SOM has also been applied to various other applications such as speech analysis, process control systems, robotics and telecommunications [Fausett 1994].

Self-Organizing Maps are generally less computationally expensive and perform better with large complex data compared to the traditional neural networks. The learning and topological arrangement of the Self-Organizing Map has been observed to be similar to the behavior observed in higher animals' primary visual cortex. Thus going back to the original motivation of neural networks: to model the operation of the human brain.

Algorithm

1- [Initialize weights] : Initialize weights wij ($1 < i < N$) to small random values. Set the initial radius of the neighborhood around node j to Nj(t).

2-[Present input]: The vector x0, x1, …, xN-1, where xi is the input to node j at time t.

3- [Calculate distance dj] : Calculate distance between the input and each node j, given by

Dj = ΣN [xi(t) – wij(t)]2 .

4- [Determine dj*]: Determine dj* which is the minimum value of dj.

5-[Update weights] Update weights for node j* and its neighbors defined by Nj*(t). New weights are:

wij(t+1) = $w_{ij}(t) + a(t)\,[x_i(t) - w_{ij}(t)]$

for j in $N_{j^*}(t)$ where $a(t)$ is a learning rate and both $a(t)$ and $N_{j^*}(t)$ decrease in t.

6- [Repeat]: Repeat by going to step 2 if the stopping condition, (t = b) is not satisfied, where b represents the total number of iterations.

The time dependent values of $a(t)$ and $Nj(t)$ are calculated by the following

RESULTS OF TESTING PROPOSED ALGORITHM

In this section, the properties of the suggested face recognition system are discussed using (20) different persons pictures. Each person has (6) different images. Each of these images is of 256 × 256 pixel.

For the training of the feature extraction methods, Gabor and Local Sampling, the number of reported iterations is constant. This is because for every time an image is used to train the three SOMS. The constant training time for all the input images is required; each SOM is actually presented with 500 sets of input data and is required to update the SOMs 500 times for each image. The two 2-D SOMs in the feature extraction contain (10x10) nodes and each node store eight values corresponding to Gabor jets and the one dimensional SOM has 256 nodes and stores six values.

In local image sampling training, the number of features that is used is larger than for the Gabor feature, the two 2-dimension SOM contains (10×10) node and one dimension contains 256 nodes. .

The following figure shows the output of the programs, **Fig. (2)** shows the output when the (learning rate is 0.2 and 0.3 threshold), the output is different image for a different person, this is due to threshold level where it represents the main factor in SOM neural network.

This state of programs uses 7 exemplars and expands later to more exemplars. Considering the capacity of the neural, the number of the exemplars for than the capacity of SOM neural network will take the whole system into collapsing.

Fig. (2) The program output with the parameters (2.1) learning rate, 7 exemplars, 0.3 thresholds.

The threshold function is an important parameter in the performance of neural network, and the decrease in the threshold value will diverge the system output from the desired output, the activation function that is used here is Sigmoid function.

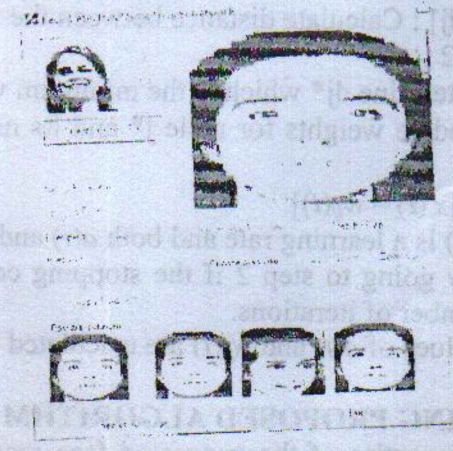


Fig. (3) The program output with parameter (0.2) learning rate, 40 exemplars, 0.5 threshold.

Fig. (3) shows the program output with the same learning rate and number of exemplar for the previous figure but with the decrease in the threshold value to (0.5), in this state the face recognition system output still stable for the given number of exemplars. The image size used in this program is (256×256) and the format of the images is (BMP).

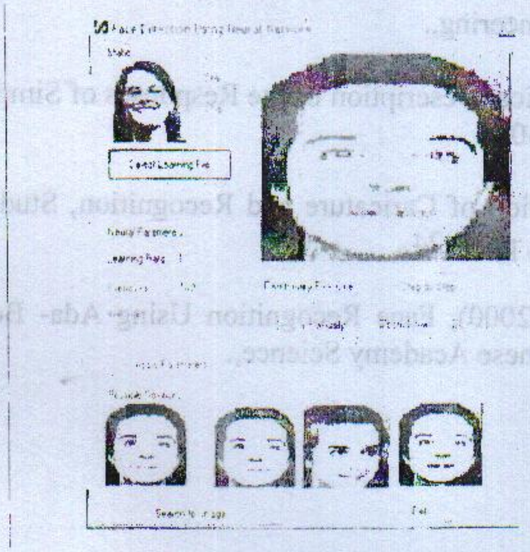


Fig. (4) The program output with parameter (0.1) learning rate, 140 exemplars, 0.9 threshold.

Finally the performance of the face recognition system will be improved with the use of (0.9) threshold value, (0.1) learning rate and (140) number of exemplars. **Fig. (4)** shows this state of program.

CONCLUSION

This work presents a human face recognition system based on the extraction of features

1- The system of face recognition depends on three parameters, learning rate, number of exemplars and threshold value.
2- The increase in the threshold value will converge the input image to the desired image and the decrease in the learning rate will lead to better tuning of the main variable in the network (weight). The increase in the number of the exemplars will increase the knowledge base of the network and the system will converge better.
3- In the comparison of the results with the results of related works, the time that the program takes for training is smaller, and the number of feature that are extracted for the classification stages is larger.

REFERENCES

R. Brunelli, T. Poggio. (1993), Face Recognition: Features versus Temples. IEEE Transactions on Pattern Analysis and Machine Intelligence , 11(6):1042– 1052,.

L. Fausett, (1994), Fundamentals of Neural Networks, Prentice Hall International, Inc.,.

D. Gabor, (1946), Theory of Communication, J. IEE, vol. 93, , pp. 429-459, .

Haddadnia, (1997), N-Feature Neural Network Human Face Recognition, Electrical and Computer Engineering Department, University of Windsor,.

T. K. Ho, J. J. Hull and S. N. Srihari, (1994), Decision Combination in Multiple Classifier Systems, IEEE Trans. On Patt. Anal. and Mach. Intel., Vol. 16, No. 1, pp. 66-75,Jan.

B. Kepenekci, (2001), Face Recognition Using Gabor Wavelet Transform', A Thesis to the Graduate School of Natural Sciences of the Middle East Technical University, the Department of Electrical and Electronics Engineering,.

S. Marcelja, (1980), Mathematical Description of the Responses of Simple Cortical Cells, J. Optical Soc. Am., Vol. 70, pp. 1297-1300,.

D. Perkins, (1975), A definition of Caricature and Recognition, Studies in the Anthropology of Visual Communication, Vol. 2, pp. 1-24,.

P. Yang, S. Shan, W. Gao, (2000), Face Recognition Using Ada- Boosted Feature, Institute of Computing Technology of Chinese Academy Science,.