Al-khwarizmi
Engineering
Journal

Al-Khwarizmi Engineering Journal, Vol.1, No.2,pp 14-21, (2005)

Human Face Recognition Using Wavelet Network

Dr. Tarik Zeyad
Electrical Engineering Department/ College of Engineering/ University of Baghdad

(Received 6 March 2005; accepted 2 October 2005)

Abstract:-

This paper presents a study of wavelet self-organizing maps (WSOM) for face recognition. The WSOM is a feed forward network that estimates optimized wavelet based for the discrete wavelet transform (DWT) on the basis of the distribution of the input data, where wavelet basis transforms are used as activation function.

Keywords: Discrete Wavelet transform, WSOM, back propagation.

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 the environment conditions. Such difficulties make face recognition on 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 applications such as film processing, law enforcement, person identification, access control systems, etc [1].

A complete 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.

In order to design a high accuracy recognition system, the choice of the feature extractor is very crucial. Two main approaches to feature extraction have been extensively used in conventional techniques. The first one is based on extracting facial features that are local structure of the face images, for example, the shapes of eyes, nose and mouth. The structured based approaches deal with local information instead of global information. Therefore, they are not affected by irrelevant information in an image. It has been shown that the structured based approaches by explicit modeling of facial features have been troubled by the unpredictability of face appearance and environmental conditions. The second one is based on statistical approaches when features are extracted from the whole image and therefore use global information instead of local information. Since the global data of an image are used to determine the feature elements, data that are irrelevant to facial portion such as hair, shoulders and background may contribute to creation of erroneous feature vectors that can affect the recognition results.

2.Neural Network

A neural network is a set of processing elements, analogous to neurons in brain, connected in a multi-layer fashion to perform processing of input data vectors. They were originally described in physiological research on the brain, and were later implemented to solve classification problems.

Typically, a neural network consists of neurons connected in several layers: one input layer, one output layer and one or more "hidden" layers. At each layer, the neuron forms a weighted sum of all of the outputs of the previous layer and transforms the sum through a nonlinear function called a "squashing" function (often a "sigmoidal" function such a tanh{}) since it tens to compress the output data to avoid truncation. In a classification application, the outputs of the final layer are compared and the class identified with the largest output assigned [1].

The knowledge in the neural net is contained in the weights used to create the sums at the various neurons. These weights are calculated by training process labeled data. The most common method is "back-Propagation" in which the weights are iteratively adjusted based on their contribution to the output vector (computed using its partial derivatives) until all input vectors produce the desired outputs.

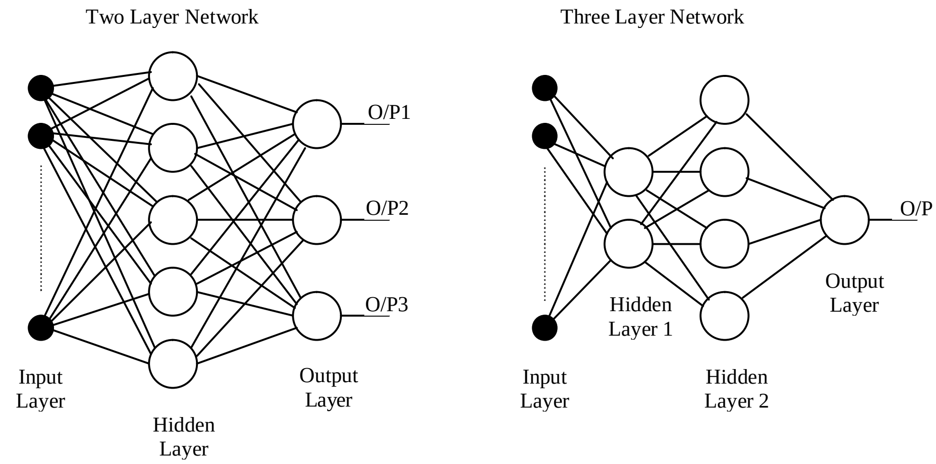

Figure (1): Neural Network with one and two hidden network.

3.Introduction to Wavelet Transforms

Transforms are mathematical analysis in different spaces. These transforms can be applied to signals or images to get another aspect of information in the different domain. There are many kinds of transformation methods such as Fourier transform, Laplace, Hough, Z, Hilbert, ..etc. Wavelet transform is a mathematical function that decomposes in terms of time domain. We can analyze each component with a resolution method to its scale.

Why Wavelet Transform

Fourier transform can change data from time domain to frequency domain. Mathematical expressions for continuous and discrete Fourier transforms are described as follows:

$$F(\omega) = \int_R f(t)e^{-j\omega j}\,dt \qquad \text{(1-D continuous FT)}$$

$$F(u,v) = \iint_{R^2} f(x,y)e^{-j(ux+vy)}dxdy \qquad \text{(2-D continuous FT)}$$

$$F(\omega) = \sum_{n\in Z} f(n)e^{-j\omega n} \qquad \text{(1-D discrete FT)}$$

$$F(u,v) = \sum_{n\in Z}\sum_{m\in Z} f(n,m)e^{-j\omega(n+m)} \qquad \text{(2-D discrete FT)}$$

According to the above equations, one can only analyze either time domain data or frequency data at a time when using Fourier transform. However sometimes one needs to analyze time and frequency domain at the same time for non-stationary data (most of natural images and signals). For example, the abrupt change in the ECG frequency signal, which cannot be, analyzed adequately by the Fourier transform.

As a result, another approach to complement the original Fourier transform was suggested, called the short time Fourier transform (STFT) or windowed Fourier transform (WFT). For one-dimensional continuous transform, the following is the STFT mathematical expression.

It can be seen from the above equation that it is a windowed function (transform). However, one does not know the exact time information. He can only know the time interval. So, the wider window produces poorer frequency and better time information. On the other hand, the narrower windows offer better frequency and poorer time resolution. To solve this problem one can use the Wavelet transform.

Wavelet Self Organizing Maps

Self-organizing maps are a form of unsupervised learning. They are modeled on the fact that similar data is spatially organized in brain. In SOM n-dimensional, pattern space is mapped on to a single or two-dimensional output space. It can also be viewed as a nonlinear projection of the multi-dimensional input space onto a two- dimensional output space. Wavelet self-organizing maps are feed forward networks having four layers as shown in figure (2).

OUTPUT	B_1………...B_i………..B_m
WAVELET	O_1…………O_i……….O_m
SOM	Y_1…………Y_i………..Y_m
INPUT	A_1………….A_i………..A_m

Figure (2) The WSOM network.

The first layer is the input layer. The second layer is the SOM competitive learning algorithm [1], which quantizes and maps the input A to an N-node grid. The SOM layers maps on to the wavelet layer through an N×N matrix D. Every node in the wavelet layer has a discrete wavelet function associated with it. Each of these functions is encoded by the elements of the matrix D. Using the inverse wavelet transform the elements of the D matrix can be calculated. The activation of each wavelet unit plots a piecewise constant function at a particular scale, as the SOM nodes are successively activated. In case of a one-dimensional input space, the number of wavelet units is the sum of the geometric progression $1+1+2+..+2^{r-1} = 2^r = N$[1]. An approximation for the basis for the L^2 functions of the input space is given both by the SOM layer and the Wavelet layer. Now as the data is preprocessed by the SOM layer. The wavelet used in this approximation gets adapted to the distribution of the training data. The wavelet bases, so formed by the WSOM layer, retain many of the properties of the mother wavelet, including orthogonality, as the wavelets are defined in the grid co-ordinate of the SOM layer instead of the co-ordinate of the input space. Conventional ways of computing wavelet coefficients require that the input signal be stored in memory, and the wavelet coefficients, by looking at one observation at a time and computing the coefficients using the data rule. The number of coefficients of the WSOM layer is lesser than the number of inputs. In this case, the WSOM layer uses low-dimension wavelets to form bases for high dimensional input space. The WSOM systems have a few advantages over the SOM, though they have the same root mean square error. The representation of a function using WSOM requires lesser number of non-zero weights, and it supports methods for reducing the noise and recovering the signal.

The following algorithm implemented the WSOM [1].

Variables:

A? (A_1……. A_i……A_M) is the input vector.

Y? (Y_1…….Y_i ……Y_M) is the vector of activity of the SOM layer.

O? (O_1……O_i…….O_M) is the vector of activity of the wavelet layer.

B? (B_1…….B_i…….B_M) is the vector of neural output.

W_j? (W_{1j}…..W_{ij}……W_{mj}) is the vector of weights from the input layer to the j^{th} node of the SOM layer.

d_s? (d_{1s}……d_{ij}…….d_{Mj}) is the non-adaptive weights from the SOM layer to the s^{th} node of the wavelet layer.

C_k? (c_{1k}……c_{ik}…….c_{Mk}) is the vector of the weights from the wavelet later to the k^{th} node of the output layer.

X_j? is the position of the j^{th} unit in the SOM layer on an integer-valued grid.

Parameters:

a: is the learning rate for weights c_{sk}.

β: is the learning rate for the weights W_{ij}.

s: is the neighborhood size used in the SOM algorithm.

j: is the index of the wining unit at the SOM layer.

h_{ij} decrease the complexity with the distance to the j^{th} unit in the SOM layer.

W- and W+ are the minimum and the maximum of initial weights W_{ij}.

β_0 is the initial value for β.

β_1 is the final value for β.

s_0 is the initial value for s.

s_1 is the final value for s.

t_1 is the required training sets to decrease β and s from β_0 and s_0 to β_1 and s_1 respectively.
n: is the total training set input number.

Algorithm

1-Set t=1, c_{sk}'s=0 and evenly distribute the weights W_{ij}'s in [W-,W+].
2-Decrease β:

$$\beta = \begin{cases} \beta_0 \left(\beta_0 / \beta_1 \right)^{\frac{t-1}{t_1-1}} & if\, t < t_1 \\ \beta_1 & if\, t = t_1 \end{cases}$$

3-Decrease s:

$$s = \begin{cases} s_0 \left(s_0 / s_1 \right)^{\frac{t-1}{t_1-1}} & if\, t < t_1 \\ s_1 & if\, t = t_1 \end{cases}$$

4-Obtain the j^{th} input vector A, and the output vector B.
5-find j=arg $\min_j \| A-W_j \|$.
6-Calculate the activity of the SOM layer: y_j=1 and y_j=0.
7-Calculate the activity of the wavelet layer: F=DY=d_{js}.
8-Compute the output $\hat{\beta}_k = \sum_{s=1}^{n} C_{sk} O_s$.
9-Adjust c_{sk} according to $\Delta C_{sk} = \alpha O_s \left(\beta_k - \hat{\beta}_k \right)$.
10-Set $h_{Jj} = exp\left(\left\| x_j - X_J \right\|^2 / s^2 \right)$.
11-Adjust W_{ik} according to : $\Delta W_{ik} = \beta_{ik} h_{Jj} (A_{ik} - W_{ik})$.
12-If t=n, then stop, else go to step 1 with (t=t+1).

4. Wavelet Neural Networks

A wavelet neural network (WNN) shown in figure (3), is a network layer with two layers whose output nodes form a linear combination of wavelet basis functions that are calculated in the hidden layer of the network. The basis used in WNN has been given the name "wavelons" [3]. These wavelons produce a localized response for an input impulse. That is, they produce a non-zero output when the input lies within a small area of the input space.

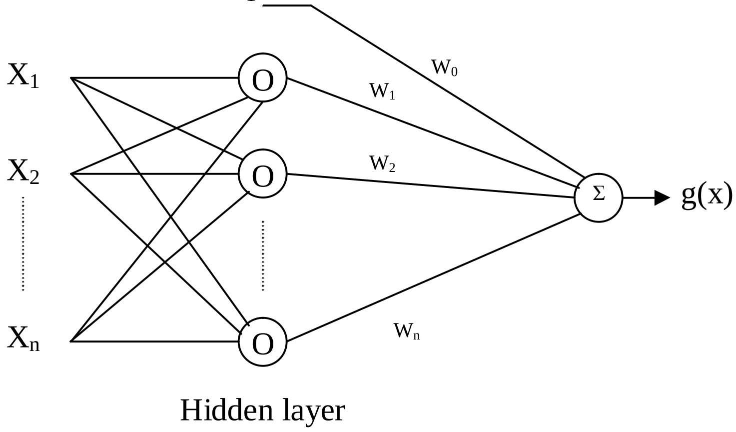

Figure (3): The proposed Wavelet Network.

The wavelet orthonormal basis is a set of functions formed by the scaling and translating the “mortlet” mother wavelets are used to decompose the input signal under analysis. This introduces an additional layer in neural network model. This layer is modified while the teaching process is in progress. The properties of wavelet transform emerging from a multi-scale decomposition of signals [4], allow the study of both stationary and non-stationary signals. On the other hand the neural network performs a stationary analysis of nonlinear as well as linear dependencies due to different possible structures and activation function. The neurons of the hidden layer in a neural network have wavelet activation functions of different resolutions. Wavelets are used as activation functions in locally responsive units. A wavelet network formed is formed on the basis of appropriate basis functions. Once created, it has the capability of the approximating any continuous nonlinear mapping to any high resolution. A simple wavelet neural network displays a much higher level of generalization and shorter computing time as compared to a three-layered feed forward neural network.

5. The Phases of Recognition Process

The recognition process contains two phases. The first phase is the twelve different faces were used during the learning phase of the recognition process. Each face provides the learning process with a twelve elements data vector used as the input of the learning WNN and this vector is used to calculate the weights of the wavelet network. The data vector represents the intensity of eight different points in the face plus the distance between the two eyes, the distance between the two ears, the distance between the end of the nose and the mouth and finally the total length of the face. These distances were all normalized according to the total number of pixels that represents the face. This normalization process is made in order to have a scale independent recognition process. The second phase of the recognition process was the test phase. In this phase the pictures of the twelve different faces used in the test with angle of rotation up to 40%. The recognition process gives correct results up to 30% rotation. Between 30% and 40% rotation percent only two of the pictures give correct results.

6. Discussion

In this paper the co-existence of neural networks and the wavelets is briefly discussed and it is seen how the neural network assist choosing the optimal shape of the wavelet. The recognition process gives good results with a good degree of face rotation but not more than 30% degree of face rotation.

7. References

1-Carpenter, G. A., “WSOM: Building Adaptive Wavelets with Self-Organizing Maps”, The 1998 IEEE International Joints Conference, Vol. 1, PP. 763-767.

2-Dongwook, C., “Face Recognition System Based on Wavelet Transform”, Department of Computer Science, Concordia University, 2003.

3-Po-Rong, C., Bao, F., “Nonlinear Computation Channel Equalization Using Wavelet Neural Networks”, The IEEE International conference on Computational Intelligence, Vol. (6), PP. 36o5-3610, 1994.

4-Kostka, P., Tkacz, E. J., Naeral, Z., and Malota, Z., “An Application of Wavelet Neural Network (WNN) for Heart Value Prostheses

Characteristics", IEEE, Engineering in Medicine and Biology Society, Vol. (4), PP. 2463-2465, 2000.

5-Vetterli, M. and Herely, C., "Wavelets and Filter Banks: Theory and Design", 2003. The IEEE International conference on Computational Intelligence, Vol. (6), PP. 2207-2232, 1994.

تمييز الوجه البشري باستخدام تحويل المويجة

د.طارق زياد

قسم الهندسة الكهربائية/كلية الهندسة
جامعة بغداد

الخلاصة:

تم في هذا البحث بناء منظومة لتمييز الأوجه باستعمال التحويل المويجي مع الشبكات العصبية من النوع ذات التنظيم الذاتي. هذا النوع من الشبكات تتعامل مع التحويل المويجي المتقطـــع كمصـــدر للبيانات التي يتم إدخالها الشبكة حيث يتم استعمال مكونات تحويل المويجة كدوال تحفيز فـــي داخـــل الشبكة.