



Face Recognition utilising Elman neural networks

Robert Chrisman

Computer Science Department, the University of Winnipeg

Email: Dr.chrisman@ mamu.edu.in

Abstract

Processing photos and recognizing them by facial architectural features lies among the basic applications of the neural network systems. This study aims investigate face recognition through utilising Elman Neural Networks. In this paper, the remittance waves (Wavelet) are used in order to withdraw more accurate details of the image, and then features were extracted based on the seven resolution and the four statistical properties (location measurement mean, standard deviation and skewness and kurtosis), which address the problems of image capture by the surveillance cameras. This study found that the recognition percentage reached 92%. This study indicated the ability of the seven resolutions and the four statistical features (measuring the middle position, measuring the standard deviation, measuring the inclination and measuring the inclination coefficient) in providing fixed characteristics that can be used as image features for the purpose of recognizing the wanted people. The study also concluded that the success of the neural network in recognizing the wanted due to its ability of handling ongoing data, and that the merging between the engineering features (the seven resolutions and for statistical features) gives good results in recognizing the wanted people.

Keywords: Face recognition, Elman Neural Networks, facial architectural, image capture.



1- Introduction:

Processing photos and recognizing them by facial architectural features lies among the basic applications of the neural network systems, especially in: Extracting distinguishing features and describing the photo, and in categorizing the photos or the extracted features (Brimblecombe, 2002). This system is designed to achieve the state security by recognizing the wanted photo. It depends on the seven resolutions and the four statistical features to extract the characteristics that solve the problems in taking photos.

Artificial neural networks have been used as a tool to complete the recognizing process and to add accuracy and efficiency in recognizing the needed facial photo; this is because the neural networks are considered as a strong tool in pattern distinction even if the information about these patterns is so few (Sung & Poggio, 1994). The neural network technique is used in the algorithms of analyzing the photo distinguishing features; as the neural network consists of a set of neurons which are connected to each others. Each one of them can process a group of signals applied on its entries and turns them into a group output signals of specific characteristics. The links between the neurons, which are labeled by the processes of weight, performs the central role in these neurons' functions as well. In this case, the neural networks have good generalization ability because they can distribute the learning experience resulting from each neuron to all the elements of the network (Stefan et al., 2004). As a matter of fact there are many kinds of neural networks. However, in this system, Elman neural network is used; and there have been good and guaranteed outcomes in the system operational use.

2- Previous Studies:

In 1987, Kirby and Sirovich (1987) conducted a study that depends on applying the principle of analyzing face elements and the standardized techniques of linear algebra in facial recognition; a thing which was considered as an essential leap later. Continues researches have been developed after that, and the idea of utilizing mathematical methodologies has emerged.



In 1997, Wiskott et al. (1997) started depicting the face as a graph whose ganglions are determined by certain points like the eyes, the nose and so on. These ganglions contain some of the wavelet transformation processes, while the edges represent the space between these points. After that, all the graphs of the face needed to be recognized are compared with graphs of the faces stored in the data base. In the beginning, the neural network methodologies depended on Kohonen Associative Map and used a small number of facial images.

Then an accurate recognition even in the case of cluttered and distorted input images was announced. This ability was also shown by using Optical Hardware. Karl and Kim (2006) used photo processing techniques to quickly identify the unique points that clearly appear on the face, and then to find the spaces and the angles in between these points to generate a significant signature of the face in the photo.

Lekshmi & Sasikumar (2009) identified that facial distinction process has been performed using wavelet analysis and the support Vector Machine (SVM). In other hand, Esbati & Shirazi (2011) clarified that facial recognition after a classification process of the data base which contains the photos of the faces depending on PCA and KPDA using SVM and Elman neural network.

Neural networks were used later within a way to categorize people according to their gender, which automatically extracted a ray of features of 16 dimensions like eyebrows thickness, mouth and nose width, six chin radii, and so on (Ramo & Ramo, 2011). A Hyper Basis Function (Hyper BF) was used for each of the genders. The input images were measured according to scale and rotation using eyes' positions, which were pointed automatically. After comparing the outputs of the two Hyper Basis Functions, the hyper with the higher output determines the gender label in the tested image. In the actual categorizing experiments, only a partial set of the features ray of the 16 dimensions was used. The validation of some results of the automatic categorization was proved through using humans. Recognition procedures usually consist of three stages (Bouzalmat, et al., 2011):



- Reading and initial processing of the image (location alignment, segmentation, clutter removal and other enhances).
- Describing the image (extracting the distinguishing features of the image in the light of stored patterns like the surface curving patterns).
- Categorizing by using the neural network algorithms (where these features are analyzed and the final recognition decision is made).

Methods of categorization were enlarged according to the gender which led to a facial recognition using an expanded ray of features of 35 dimensioned and a Hyper BF for each person. The results of the categorization of the different groups from the synthetics (coefficients, centers, and Hyper BF metrics)) were attained. There was also some participation in explaining the caricature behavior for the Hyper BF by psychophysical studies.

3- The Seven Moments:

The seven moments were used as they are typical for a number of applications to perform the fixed distinction of a two dimensioned photo. I have also depended on the algebraic method using an non-linear equation for the values which represent the geometric moments (GM), which is a characteristic needed for fixing when interpreting the photo, modifying the size (zooming in and out) and rotation. This type of values is the perfect scale to recognize the Arabic and English letters and the digital photos. In this study, the (GM) technique and its set of fixed values were used due to its characteristics for being fixed as opposed to converting, rotation and scoring and its features for each form of the group. This set of constants is able to use dual or a real value depending on the type of the problem that is needed to be resolved (Al Hadidi, 2002).

The central means are known as (mpq) and as follows:



$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dx dy \quad \dots(11)$$

Where $p, q = 0, 1, 2, \dots$

X, Y are the two coordinates of the first photo unit respectively.

P, q represents the layer of the regular mean.

F(x, y): is the value of the contrast equation of the image and it is 1 or 0 in the two dimensioned photos.

X and Y represent the standard centers of the image in the direction of x and y.

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x,y) dx dy \quad \dots(12)$$

$$\bar{y} = \frac{m_{01}}{m_{00}}, \quad \bar{x} = \frac{m_{10}}{m_{00}}$$

$$m_{pq} = \sum_x \sum_y [(x - \bar{x})^p (y - \bar{y})^q] f(x,y)$$

$$\mu_{pq} = \frac{m_{pq}}{\mu_{00}} \quad \dots(14)$$

$$g = 1/2(p + q) + 1 \quad \dots(15)$$

$$p + q = 2, 3, \dots$$

Seven invariant means could be derivated from the second and third mean which are given as follows:

$$\phi_1 = \eta_{20} + \eta_{02} \quad \dots(16)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4 \eta_{11}^2$$

...(12)

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad \dots(17)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad \dots(18)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad \dots(19)$$

$$\phi_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad \dots(20)$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad \dots(21)$$



(Hu, 1962) proved that these means are invariant for translation, rotation and scaling.

4- Elman Neural Network:

The neural networks were used in this field because they are specialized in differentiating the different patterns in the performing processes. In this research, Elman neural network was used. This kind of networks consists at least of three layers of cells: the input layer, the hidden layer and the output layer. Yet, it also has a context layer which gives a feedback of no weight; as for the output, it moves from the hidden layer towards the output layer.

Elman's network keeps the values and give them back when the following performance is done to the network: these values are sent through the training and retrieving of the weights to the hidden layer, and because Elman's network has a kind of network with a feedback effect, meaning they have a link from the hidden layer backward to a private layer, this network depends on the (current inputs, previously registered position and the network's output). This means that they can specify accurately what the preserved position of the last process is. In other words, we will have an additional hidden layer which records the value resulted from the previous processes and which will be of much help in the training course (15) (17).

The ganglions in the hidden layer are performed by the TANSING referential equation which works on an area that is restricted in the terms of input by using the following equation (18):

$$a_1(k) = \text{tansig} (IW_{1,1}p + LW_{1,1}a_1(k-1) + b_1) \quad \dots\dots (22)$$

The ganglions in the output layer are extracted by the PURELIN linear equation which works on an area that is restricted in terms of output using the following equation (12):

$$a_2(k) = \text{purelin} (LW_{2,1}a_1(k) + b_2) \quad \dots\dots (23)$$

Where: P: the input matrix, IW1,1: the weight matrix of the hidden layer, a1, a2: the output matrix, b1, b2: the basis matrix, LW2,1: the weight matrix for the output layer, K: feedback rotation



PURELINE: is a neural transformation equation which works on the expense of the layers outputs from the inputs of the network.

The weights were prepared at random where the values are small and in the scale of [0, 1]. They have been adapted as initial weights for the network and then put in a data base.

5- Research Algorithm:

In this research, remittance waves (wavelet) was used to extract more accurate details, and the seven resolutions were used plus the four statistical features which solve a lot of the problems in taking the photo. Elman's network was used as a tool to complete the person recognition and to add more accuracy and efficiency in recognizing the needed face image. After getting good and guaranteed results in the function of the system, the values, which were attained through the extraction from the features of each photo that was taken, were stored in the database. In this system, two levels of recognition were used:

Level one: the user input the stored images and extracts the features (the seven resolutions, and the four statistical features) for each image and stores it in the database. The used network is trained on the attained results.

Level two: the user inputs the image he want to recognize and extract its features and get results without storing them in the database, and then input the results into the used neural network and compare the finale values with those stored in the database to show the final result of the recognition system.

5-1- Reading Photos: in this system, colorful images which consist of three layers (R,G,B) are read from the file of the database which is stored in a common path, and due to the big size of the photo the format used in the system is (BMP) or (BitMap). This kind doesn't have problems of losing data, where the image is a two or three dimensioned matrix; pictures of this size are usually so big and unidentified. This is when the size of the taken photos is changed to a particular size (128,128) so that the new size is specified to the face of the person. A certain



number of images for some people who are needed to be recognized were used in this system, where seven photos of each person were input.

5-2- Photo Configuration: after storing the photo in the database in its colored version, it is transformed into grey images. The reason is that the seven resolutions and the statistical features which are used to extract the characteristics of the photo work on this kind of photos; in addition to that, the format of the grey images makes it easy to deal with photos when applying the rules. Plus, in the case of grey colors, the photos are clear of the effects which are found on them if they are colorful, and hence we get clearer results. This procedure is performed during the functions of the initial processes except for the image cuts because the seven resolutions takes their values from the characteristics of the whole image, and in case the image is divided into several parts, the values of the seven resolutions will be taken from each of the parts. After the extraction of the features is done, normalization is done by the Min-max method.

5-3- Wavelets: A separated two dimensioned wavelet was used to convert a photo and the result was four sets of features; one of them is a copy of the original copy with a low resolution and is called rounding features, whereas the other three sets are copies of the original photo with a high resolution called detailed features; these are a horizontal set, a vertical set and a diametric set. Most of the information is found in the rounding features, and we can zip the photo by taking these resolutions only (Al Neamy,2006).The purpose of using wavelets is to minimize the photo size to a quarter of its original size and which contains more accurate details. We used the wavelet of the two dimensioned detached to transform the photo, and the result was for sets of factors: 1- a copy of the original image with a low resolution which is called (rounding coefficients A) that we will make use of in this research. 2- The other three sets are copies of the original image with high resolutions which are called (details coefficients: vertical set, horizontal set and diametric set). Most of the information are gathered in the rounding coefficients, and we can press the photo by taking these coefficients alone (Al Neamy, 2006) (AlBadrani, 2001).



The input of the wavelet was the grey gradation data of the image. As for the output of the wavelet, a part of the low resolution coefficient (rounding coefficient A) will be taken and from it the features could be extracted.

5-4- Extracting Features: extracting features is done after attaining a two dimensioned matrix (32, 32) which represents the image after transforming it from the grey form to the wavelet. Attaining these features is done on two axes:

Axis One: the seven resolutions; this has been depended on as a basis to extract features because these features remain fixed during rotation, maximizing, minimizing and transferring, and using them will make the extracted features both small and effective. It is from these seven resolutions and their name that a seven- valued number of the correct group as an outcome of extracting the photo's features. The rank of the *order* of the seven features is three.

Axis Two: the four statistical features; the values of these features vary according to the variety of the photo, the thing that makes it an advantage that can benefit in recognizing a person's image from the others'. With the help of these four features (1- measuring the middle position by applying the equation(6). 2- measuring the standard deviation by applying the equation(7). 3- measuring the inclination by applying the equation(8). 4- measuring the inclination coefficient by applying the equation(9)), we can complete the task of calculating the features of the image which imply in (11) a value as a result of each of the seven resolutions in addition to the four statistical features of the image which are stored in a database of the type Excel in a serial arrangement and is enhanced by the objectives data and is considered as an input of the used neural network. The statistical features were use actually to increase the characteristics.

5-5- Database: when applying the seven resolutions and the four statistical features on each image, the results are stored in the database which consists of a number of lines and 14 columns, where the sequence of the line refers to the name of the image whose features are stored vertically and which is followed by the next image. In this system, a sample of 16 persons' images is input- there could be any other set of numbers- in a serial arrangement and every person has got 7 images in the sequence that the first seven images are for the first



person plus their features taken from the extracted characteristics, followed significantly by the next person's seven images. This work can be done on a huge number of people. There are samples of the people who were under the testing trial.

After storing the extraction values of the features in the database, other values are to be stored; those are the objectives values that are added to the database as well. The objective contains a value for each person stored in the database and these values are stored in the Binary form. In this system, a sample of 16 persons' images was input- any number can be input- so as each person has got seven images where the first seven images are for the first person. This is why the value of the objective is fixed which (0001) is.

5-6- Construction and Training of the Network: in this research, Elman's neural network that was compatible with the system was used because of its accurate results and efficiency in recognizing the patterns. This network depends basically in its function on the database where it takes the values of both the features and the objectives out of it. The network receives two matrixes which will be configured to be ready to input into the network; each one of them is a singular matrix. The first one contains results attained in the process of the features extraction (the seven resolutions and the four features) and the size of the first matrix (11) is a value that consists of seven values which are the results of the seven resolutions and four values which are the results of the four basic statistical features. These values determine their position in the database from the first till the second column and are taken in a serial arrangement according to the sequence of the images in the database. The matrix is configured by rotating it from the horizontal to the vertical shape to get it in the network according to the input form that should be vector. As for the second matrix, it contains the stored objective in the database and the size of the second matrix is 3 values which represent the number of the objective value and is positioned in the database from column 12 to column 14.

The Input Layer: it is the first layer of Elman's network which contains the first matrix (features extraction results), and the number of the ganglions in this layer is 11 just like the number of the extracted features that are stored in the database which are 11 as well.



The Context Layer: it has the same number of ganglions as the input layer because the network has got a temporary feedback and which gives a weightless feedback.

The Output Layer: it consists of three ganglions according to the objective data.

After the stage of constructing the network, the network starts a training stage by applying the (22) (23) rules.

Then, the extracted features values which are stored in the database are taken. These values (extracting features) which are input into the network are for producing the perfect weight (training); where the network is trained on a percentage of error that is estimated by (0,001) and a number of rotations estimated by (1079) to get to the needed objective.

6- Recognizing the Wanted Person System:

When all the past stages are done, the stage of recognizing the wanted person starts, which is the final stage in this system where the image goes through the following stages: (reading the image of the wanted person if he is among the wanted or not, transforming the image into grey colors, moving to wavelet, extracting features (the seven resolutions and the four statistical features), recognition system, showing the final result of the program.

First: the image is input in its colorful form, and then it is transformed into the grey form.

Second: the remittance waves (wavelet) are applied on the image where the most accurate part is to be attained.

Third: it goes into the stage of extracting the features and stored in a single matrix without passing through to the database.

Fourth: this matrix goes into the network to get the final results.

Fifth: these results are compared with the results that are already stored in the database, and either they are identical or close to them with an error percentage of (0,001). The system gives a feedback of these results which refer to a certain person.



7- Discussing the Results:

After fulfilling the suggested system and comparing it with the ideas applied by previous researchers, this study found that the recognition percentage reached 92% and it was calculated in the following way: Recognition percentage= the number of people the system was trained on/the number of people recognized .Some researchers used the linear algebra techniques in past researches. In these techniques, there is a flaw when maximizing or minimizing the image or when rotating the head.

All these flaws have been overcome in the seven resolutions which are not affected by rotating or motion, but they depend on fixed characteristics which are extracted from the features of the face; the thing that distinguishes this research from the others.

8- Conclusions:

This study indicated the ability of the seven resolutions and the four statistical features (measuring the middle position, measuring the standard deviation, measuring the inclination and measuring the inclination coefficient) in providing fixed characteristics that can be used as image features for the purpose of recognizing the wanted people. Wavelet was used to get a picture of more accurate features and details. The study also concluded that the success of the neural network in recognizing the wanted due to its ability of handling ongoing data, and that the merging between the engineering features (the seven resolutions and for statistical features) gives good results in recognizing the wanted people.

Resources:

- Al Hadidi, Ibrahim Ahmad Saleh, (2002), pressing digital images using wavelet and directional quantum, a master research, college of mathematics and computer sciences, Al Mousl University.



- Anissa Bouzalmat, et al (2011) “Face Detection And Recognition Using Back Propagation Neural Network And Fourier Gabor Filters”, Signal & Image Processing: An International Journal (SIPIJ) Vol. 2, No. 3, September 2011, DOI : 10.5121/sipij.2011.2302 15.
- Brimblecombe , Phil (2002) “Face Detection using Neural Networks”, H615 – Meng Electronic Engineering, School of Electronics and Physical Sciences, URN: 1046063. Engineering.
- Fawziya M. Ramo, Ramadan M. Ramo, (2011): "Face Detection using Neural Networks", College of computers & Mathematic Scientific, Computer Science Department, University of Mosul, Iraq Journal Of Statistical(Iraq. J. S. S.).
- Hossein Esbati and Jalil Shirazi,(2011), "Face Recognition with PCA and KPCA using Elman Neural Network and SVM", World Academy of Science, Engineering and Technology.
- Karl B. J. Axnick, and Kim C. Ng, (2006), "Fast Face Recognition", Intelligent Robotics Research Centre (IRRC), <http://pixel.otago.ac.nz/ipapers/60.pdf>.
- L. Sirovich and M. Kirby, "A Low-Dimensional Procedure for the Characterization of Human Faces," J. Optical Soc. Am. A, 1987, Vol. 4, No.3, 519-524.
- L. Wiskott, J.M. Fellous, N. Kruger, and C. von der Malsburg, “Face recognition by elastic bunch graph matching,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 775–779, 1997.
- Praseeda Lekshmi., Dr.M.Sasikumar,(2009), "Analysis of Facial Expression using Gabor and SVM", International Journal of Recent Trends.
- Stefan W., Christian I. & Uwe H (2004) “Evolutionary Optimization of Neural Networks for Face Detection”, Proceedings of the 12th European Symposium on Artificial Neural Networks, Evere, Belgium: d-side publications.
- Sung, KahKay & Poggio, Tomaso (1994) Example Based Learning For View Based Human Face Detection, Massachusetts Institute of Technology Artificial Intelligence



Laboratory and Center For Biological And Computational Learning, Memo 1521,
CBCL Paper 112, MIT, and December.