Ear Recognition by Using Self Organizing Feature Map

Suad K. Mohammad
Department of Electrical Engineering, University of Technology/ Baghdad
Email: - rawankh2006@yahoo.com

Received on: 22/5/2012 & Accepted on: 7/3/2013

ABSTRACT

A wide variety of systems requires reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services. The purpose of such schemes is to ensure that the rendered services are accessed only by a legitimate user and no one else.

The aim of the work presented within this paper is to develop an optimum image compression system using haar wavelet transform and a neural network. In this paper we have developed and illustrated a recognition system for human ears using a Kohonen self-organizing map (SOM) or Self-Organizing Feature Map (SOFM) based retrieval system. SOM has good feature extracting property due to its topological ordering. The ear Analytics results for the 4 images of database reflect that the ear recognition using one of the neural network algorithms SOM for 4 persons. MATLAB programs were used to complete this work.

Keywords: Image Compression, Two-Dimensional Wavelet Packet Analysis, Haar Wavelet, Vector Quantization, Self-Organizing Feature Map (SOFM), Neural Network and pattern recognition.

.

تمييز الاذن باستخدام ميزة شبكة التنظيم الذاتي

الخلاصة

مجموعة واسعة من منظومات تقديم الخدمات تتطلب مخططات تميز تاشخصية معول عليها (موثوق بها) لتأكيد أو تحديد هوية الفرد طالب الخدمة. الغرض من هذه المخططات هو ضمان أن يتم الوصول إلى الخدمات المقدمة من قبل مستخدم فقط مشروعة وليس لأحد غيرهم. الهدف من العمل المقدم في هذا البحث هو وضع ضغط الصور المثلى النظام باستخدام المويجات تحويل هار والشبكة العصبية. في هذا البحث تم تطوير نظام التعرف على آذن الإنسان باستخدام ميزة شبكة ذاتية التنظيم (SOM) لانها لديها ترتيب طوبولوجي القائم على نظام استرجاع (SOFM) . وقد تم استخدام برامج (MATLAB) لاتمام هذا العمل. نتائج تحليلات الاذن لاربع صور من قاعدة البيانات تبين ان الشبكة العصبية SOM لاربع اشخاص.

INTRODUCTION

Kohonen Self-Organizing Maps (or just Self-Organizing Maps, or SOMs) is a type of neural network. Self- Organizing Maps are aptly named. “Self-Organizing” is because no supervision is required. SOMs learn on their own through unsupervised competitive learning. “Maps” is because they attempt to map their weights to conform to the given input data. The nodes in a SOM network attempt to become like the inputs presented to them. In this sense, this is how they learn. They can also be called “Feature Maps”, as in Self-Organizing Feature Maps. Retaining principle 'features' of the input data is a fundamental principle of SOMs, and one of the things that makes them so valuable. Specifically, the topological relationships between input data are preserved when mapped to a SOM network. Classification is the first application of the SOMs; it is not very practical on its own. One of these useful applications is described in the fourth section, Image Classification [1].

The purpose of this application is to show a useful implementation of the SOM algorithm. Given some database of pictures, this SOM will organize them based on content [2].

While the histogram data vector could be used on its own, it is better to have as many weight values as possible. This distinguishes images from each other even more, which is useful for the SOM during learning. The other weight vector used in this application is the area data vector.

EAR ANALYSIS

Ears have gained attention in biometrics due to the robustness of the ear shape. The shape does not change due to emotion as the face does, and the ear is relatively constant over most of a person’s life. Due to its semi-rigid shape and robustness against change over time, the ear has become an increasingly popular biometric feature. It has been shown that combining individual biometric methods into multi-biometric systems improves recognition. Ear Recognition has been identified as one of the attracting research areas and it has drawn the attention of many researchers due to its varying applications such as security systems, medical systems, entertainment, etc.

Ear recognition has the benefit of being a passive, non-intrusive system for verifying personal identity. Many supervised and unsupervised learning techniques have been reported for ear recognition. Various algorithms for ear recognition have been used which can be broadly divided into two approaches, namely, structure-based (Appearance based) and statistics-based (Feature based). Three different techniques - PCA, ICA & SOM, have been used for ear recognition [2]. Principal Component Analysis (PCA) is derived from Karhunen- Loeve's transformation. Given an s-dimensional vector representation of each ear in a training set of images, PCA tends to find a t-dimensional subspace whose basis vectors correspond to the maximum variance direction in the original image space [3]. This new subspace is normally lowered dimensional. If the image elements are considered as random variables, the PCA basis vectors are defined as eigenvectors of the scatter matrix. Independent Component Analysis (ICA) minimizes both second-order and higher-order dependencies in the input data and attempts to find the basis along which the data (when projected onto them) are statistically independent [4].

Individuals are often recognized by their ears and the recent computing technology has now resulted similar recognition automatically. Initial ear recognition algorithms used simple geometric models, but recognition process has matured into a science of sophisticated mathematical representation and matching process.

Recognition algorithms can be divided into two main approaches,

1-Geometric, this looks at distinguishing features.

2-Photometric, this is a statistical approach that distills an image into values and comparing the values with templates to eliminate variances. The photometric approach is used for ear analysis.

TWO-DIMENSIONAL WAVELET PACKET ANALYSIS

In this section, the Wavelet Packet 2-D tool employs to analyze and compress an image of an ear. This is a real-world problem because maintains a large database of ear images is very difficult and the cost of storing all this data runs to hundreds of millions of dollars.

The technique involves a two-dimensional DWT, uniform scalar quantization (a process that truncates, or quantizes, the precision of the floating-point DWT output) and Huffman entropy coding (i.e., encoding the quantized DWT output with a minimal number of bits).

By turning to wavelets, in this application, wavelet compression is better than the more traditional Joint Photographic Experts Group (JPEG) compression, as it avoids small square artifacts and is particularly well suited to detect discontinuities (lines) in the ear image.

Note that the international standard JPEG 2000 will include the wavelets as a part of the compression and quantization process. This point out is the present strength of the wavelets.

IMAGE COMPRESSION

Image compression is a fast paced and dynamically changing field with many different varieties of compression methods available. Images contain large amount of data hidden in them, which is highly correlated.

A common characteristic of most images is that the neighboring pixels are correlated and therefore contain redundant information.

The investigation and design of computationally efficient and effective software algorithms for lossy image set of images. In wavelet image compression, parts of an image is described with reference to other parts of the same image and by doing so, the redundancy of piecewise self-similarity is exploited. There are a number of problems to be solved in image compression to make the process viable and more efficient. A lot of work has been done in the area of wavelet based lossy image compression. However, very little work has been done in lossless image compression using wavelets to improve image quality. So the proposed methodology of this paper is to achieve high compression ratio in images using 2D-Haar Wavelet Transform by applying different compression ratios are applied to the wavelet coefficients thresholds for the wavelet coefficients. That is, different compression ratios are applied to the wavelet coefficients belonging in the different regions of interest, in which either each wavelet domain band of the transformed image [5].

The proposed method suggests that a trained neural network can learn the non-linear relationship between the intensity (pixel values) of an image and its optimum compression ratio. Based on our hypothesis, a trained neural network could recognize the optimum haar compression ratio of an image upon its presentation to the neural network.

The development and implementation of this image compression system uses images of various objects, contrasts and intensities.

IMAGE DATABASE

The development and implementation of the proposed optimum image compression system uses four images from our database that have different objects, brightness and contrast. Haar compression has been applied to the four images shown in Figure (1).

The optimum Haar compression ratios for the four images were determined using the optimum compression criteria based on visual inspection of the compressed images. The image database is then organized into four sets. Training Image Set contains four images which are used for training the neural networks (SOM) within image compression system. The four images after Haar compression are shown in Figure (2), Figure (3), Figure (4) and Figure (5).

These four input data sets (images) will be used to train the SOM, and then the same sets will clustered and recognized with SOM. These 4 data set are shown in Tables (1, 2, 3 and 4) for images 1, 2, 3 and 4 with respect. Testing Image Set contains image which is used to test and evaluate the efficiency of the trained neural networks. This testing image is shown in Figure (10).

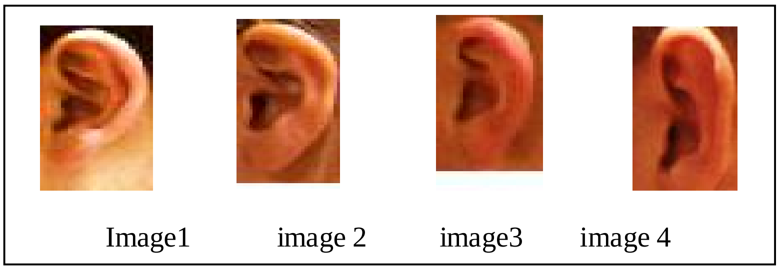

Figure (1) the origin pictures of ears.

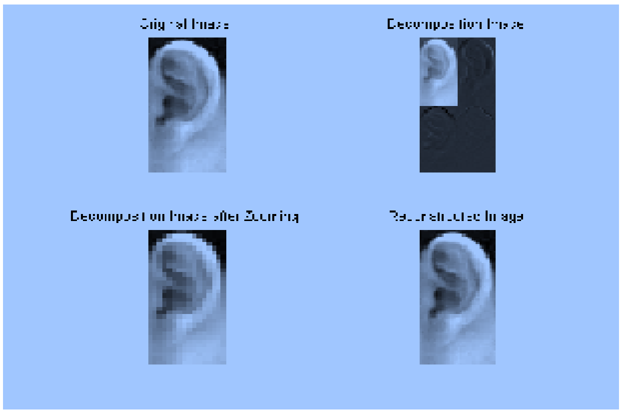

Figure (2) Image 1 after Haar compression.

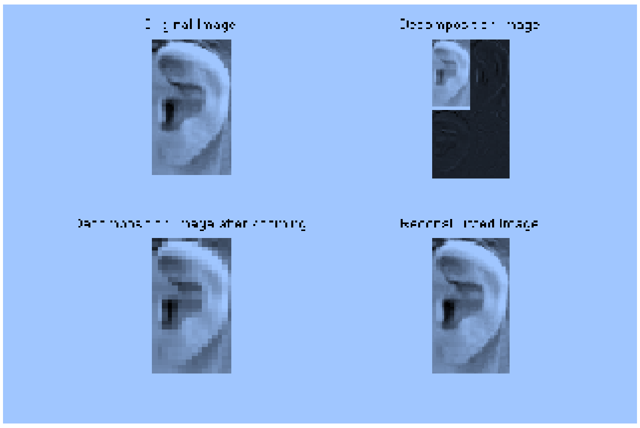

Figure (3) Image 2 after Haar compression.

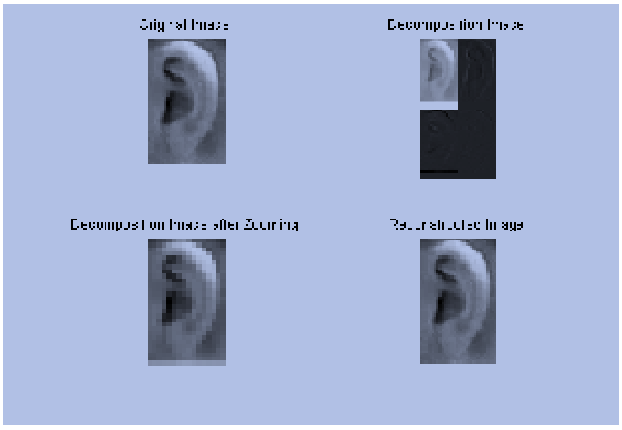

Figure (4) Image 3 after Haar compression.

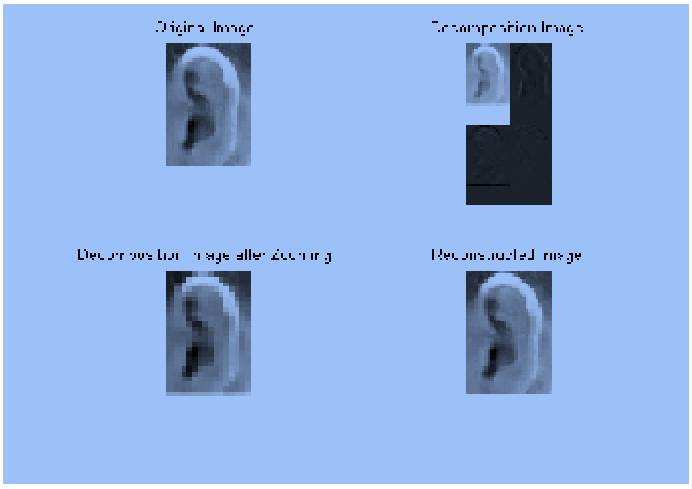

Figure (5) Image 4 after Haar compression.

SELF-ORGANIZING MAPS (SOMS)

(SOM) is an unsupervised learning process that has the property of topology preservation. Unsupervised learning algorithms use patterns that are typically redundant raw data having no labels regarding their class membership. In this mode of learning, the network must discover for itself any possibly existing patterns, regularities, separating properties etc. While discovering this, the network undergoes change of its parameters, which is called self-organization. Unsupervised learning is sometimes called learning without teacher. We are using the neural network based unsupervised learning algorithm known as Self Organizing Map [6].

The self-organizing map, or SOM, introduced by Teuvo Kohonen is an unsupervised learning process which learns the distribution of a set of patterns without any class information.

As a neural unsupervised learning algorithm, Kohonen's Self-Organizing Maps (SOM) has been widely utilized in pattern recognition area. Using the SOM as a feature extraction method in ear recognition applications is a promising approach, because the learning is unsupervised, no pre-classified image data are needed at all. When high compressed representations of ear images or their parts are formed by the SOM, the final classification procedure can be fairly simple, needing only a moderate number of labeled training samples. The (SOM) is unlike most classification or clustering techniques in that it provides a topological ordering of the classes.

Similarity in input patterns is preserved in the output of the process. The topological preservation of the (SOM) process makes it especially useful in the classification of data which includes a large number of classes [7].

Structure of a SOM

The structure of a SOM is fairly simple, and is best understood with the use of an illustration such as Figure (6).

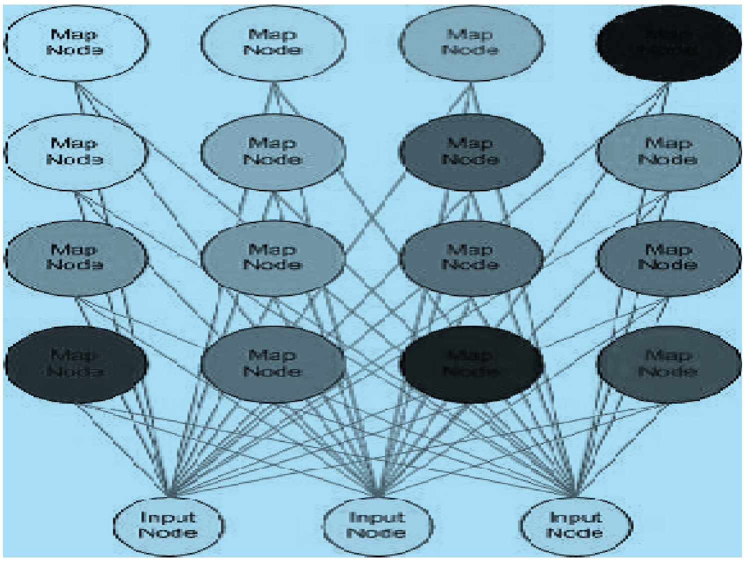

Figure (6) 4 x 4 SOM network.

Figure (6) shows a 4x4 SOM network (4 nodes down, 4 nodes across). It is easy to overlook this structure as being trivial, but there are a few key things to notice.

First, each map node is connected to each input node. For this small 4x4 node network, that is 4x4x3=48 connections. Secondly, notice that map nodes are not connected to each other. The nodes are organized in this manner, as a 2-D grid makes it easy to visualize the results. This representation is also useful when the SOM algorithm is used. In this configuration, each map node has a unique (i,j) coordinate. This makes it easy to reference a node in the network, and to calculate the distances between nodes. Because of the connections only to the input nodes, the map nodes are oblivious as to what values their neighbors have [8].

A map node will only update its' weights based on what the input vector tells it.

THE SOM ALGORITHM

The Self-Organizing Map algorithm can be broken up into 6 steps.

1) Each node's weights are initialized.

2) A vector is chosen at random from the set of training data and presented to the network.

3) Every node in the network is examined to calculate which ones' weights are most like the input vector. The winning node is commonly known as the *Best Matching Unit* (BMU).
4) The radius of the neighborhood of the BMU is calculated. This value starts large. Typically it is set to be the radius of the network, diminishing each time-step, (Equation 2a and 2b).
5) Any nodes found within the radius of the BMU, calculated in 4), are adjusted to make them more like the input vector (Equation 3a, 3b). The closer a node is to the BMU, the more its' weights are altered (Equation 3c).
6). Repeat 2) for N iterations.

The equations utilized by the algorithm are as follows [9]:

Equation 1 – Calculate the BMU.

$$DistFromInput^2 = \Sigma\ Ii - Wi\ i = 0 \quad \ldots(1)$$

$$\sum_{i=0}^{i=n} (Ii - Wi\ )2 \quad \ldots(2)$$

I = current input vector
W = node's weight vector
n = number of weights

Equation (2a) – Radius of the neighborhood.

$$\sigma(t) = \sigma_0\, e(-t/\lambda) \quad \ldots(3)$$

t = current iteration
λ = time constant (Equation 2b)
σ_0 = radius of the map

Equation (2b) – Time constant

$$\lambda = num.\ Iterations\ /map\ Radius \quad \ldots(4)$$

Equation (3a) – New weight of a node.

$$W(t+1) = W(t) + \Theta(t)L(t)(\,I(t) - W(t)) \quad \ldots(5)$$

Equation (3b) – Learning rate.

$$L(t) = L0e^{(-t/\lambda)} \quad \ldots(6)$$

Equation (3c) – Distance from BMU.

$$\Theta(t) = e\left(-dist.\ From\ BMU^{2} / (2\sigma^{2}(t))\right) \quad \ldots(7)$$

THE PROPOSED DESIGN

The basic block diagram of ear recognition for the project is as shown in figure (7).

A. Training

Data base of ears → SOM for training

B. Mapping

Trained SOM → Recognized ear

Figure (7) The Ear Recognition.

TRAINING THE NEURAL

Data Base of Ears: There are four different images of each of 4 distinct persons.

When the input data is too large to be processed and it is suspected to be notoriously redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features (also named features vector).

Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. In our design the four data set (16x32) is taken for the ear's pictures that shown in Figure (1).

Training of SOM: The recognition process involves input and target vector. Training the SOM requires no target vector. A SOM learns to classify the training data without any external supervision which results in the formation of different Clusters or Classes.

There are various cluster topologies such as grid, hex and random. In this paper the hex topology is used to cover maximum area of the neurons under training [10].

MAPPING

Trained SOM: The mapping of input image is performed with the trained clusters of the database. The input matrix of our design will be (16x128) that shown in Table (5), after the training process of SOM the recognition process is obtained and the SOM classified the input data set into four classes, the output matrix that obtained after the training and recognition processes will be (4x128) as in Table (6).The SOM design shown in figure (8).After the training process the Mean Square Error (MSE) is shown in figure (9) with performance 0.0037318 at epoch 16.

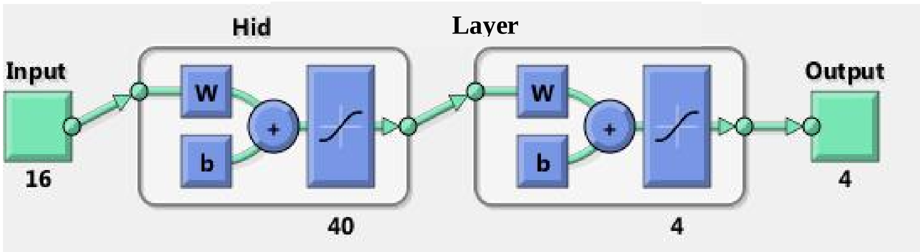

Figure (8) The SOM Design.

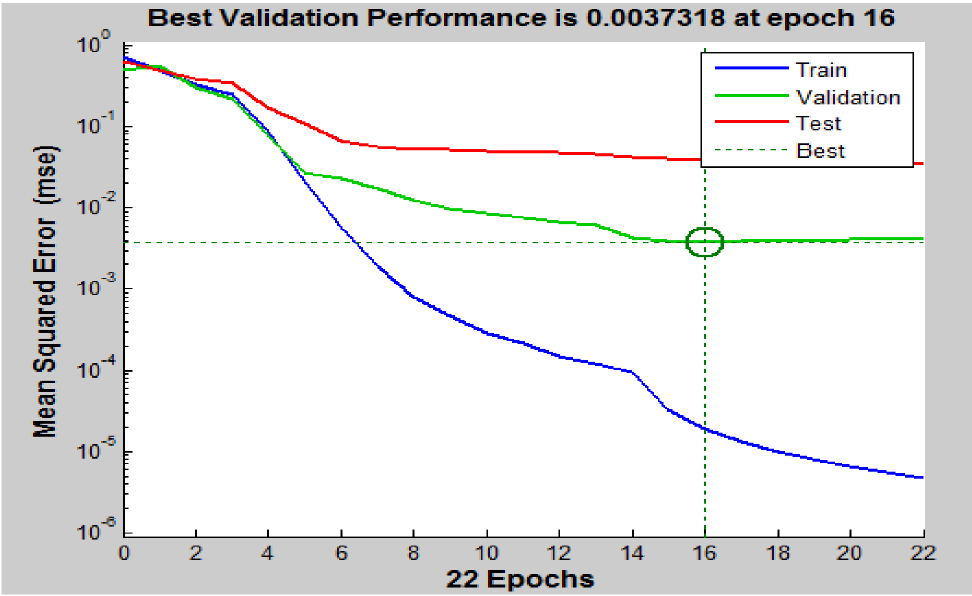

Figure (9) The Mean Square Error (MSE).

Recognized Ear: The best match determined using Euclidean distance formula is the output of this process. The minimum distance between input image and classifiers or clusters is the actual recognized ear image.

Ear recognition is one of the few biometric methods that possess the merits of both high accuracy and low intrusiveness. Pattern recognition is one of the important steps in Image Processing.

Given a pattern, the first step in the pattern recognition is to select a set of features or attributes from the universe of available features that will be used to classify the pattern.

Next, the original pattern must be transformed into a representation that can be easily manipulated programmatically; in this research we use wavelet image compression. . After the data are processed to remove noise, features in the data that are defined as relevant to pattern matching are searched for. In the classification stage, data are classified based on measurements of similarity with other patterns. The pattern recognition process ends when a label is assigned to the data, based on its membership in a class.

In ear recognition system, we have database of images stored in the system. Whenever we get a new image as in figure (10), it is compared with the database of images already stored in the system. SOM operates in two modes: training and mapping.

1-Training builds the map using input examples. It is a competitive process, also called vector quantization.

2-Mapping automatically classifies a new input vector.

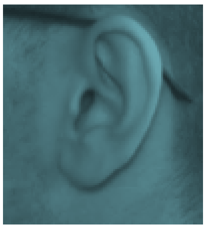

Figure (10) The new image for testing.

The program for our ear Analytics project consists of following steps:

Step 1 – start.
Step 2 – Initialize the Map for Clustering.
Step 3 – Set t = 0 and Repeat the following steps until t < e.
Where t is the iteration rate and e is the error rate.
Step 4 – Get the Best Matching Unit.
Step 5 – Scale Neighbors
Step 6 – Increase t by small amount.
Step 7 – end.

After executing this program the MSE is obtained shown in Figure (9) and the Gradient is shown in Figure (11) also the Error Histogram and the Training Confusion Matrix are shown in Figures (12) and Figure (13).

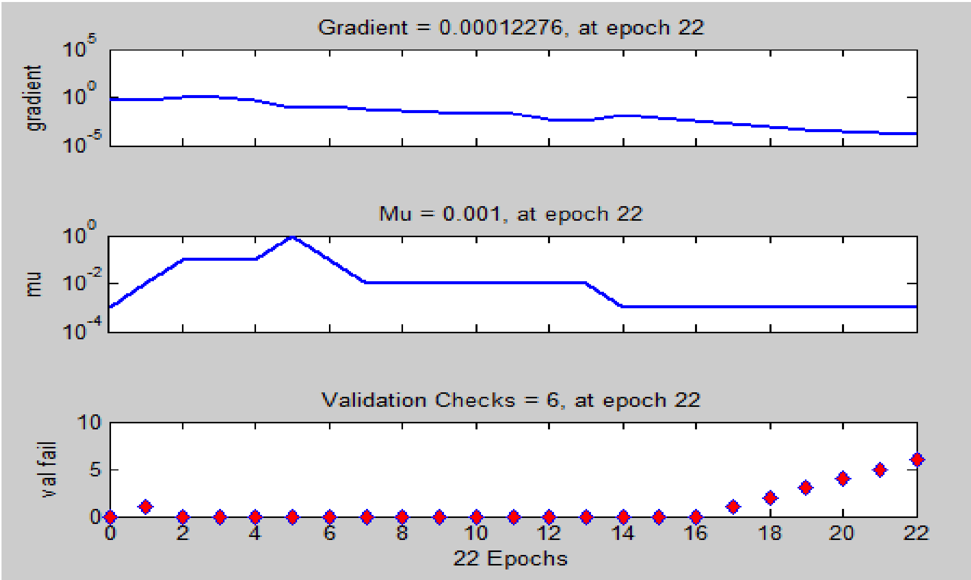

Figure (11) The Gradient.

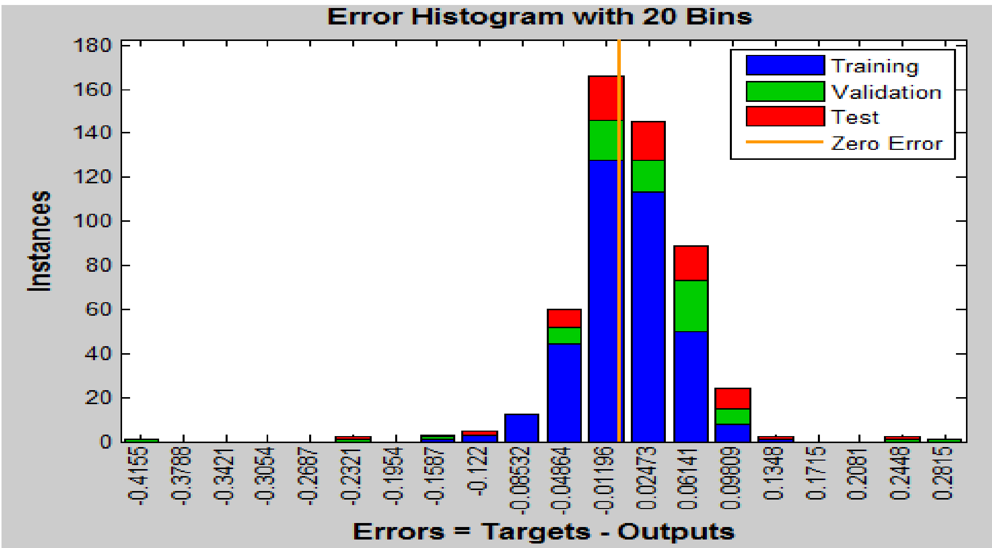

Figure (12) The Error Histogram.

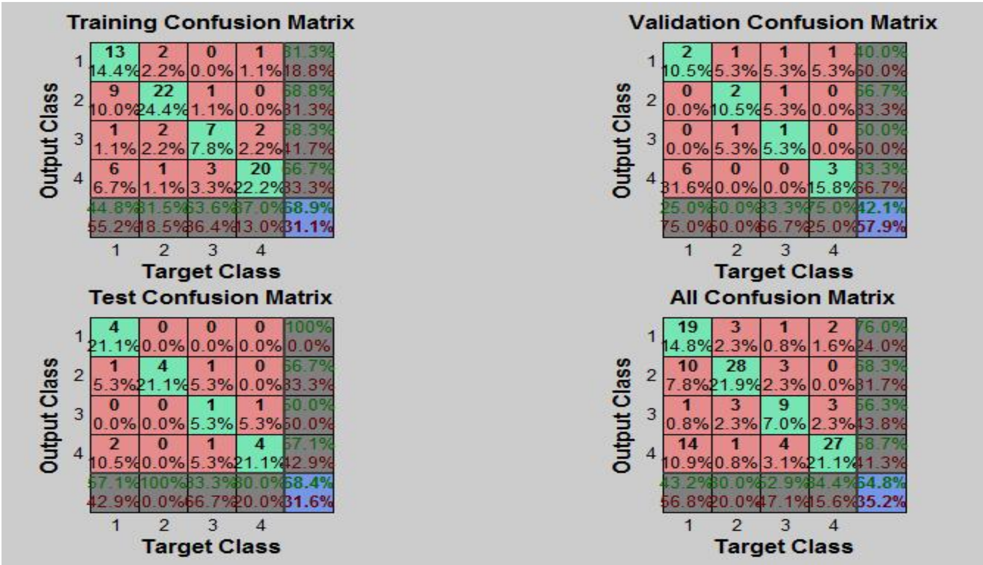

Figure (13) The Training Confusion Matrix.

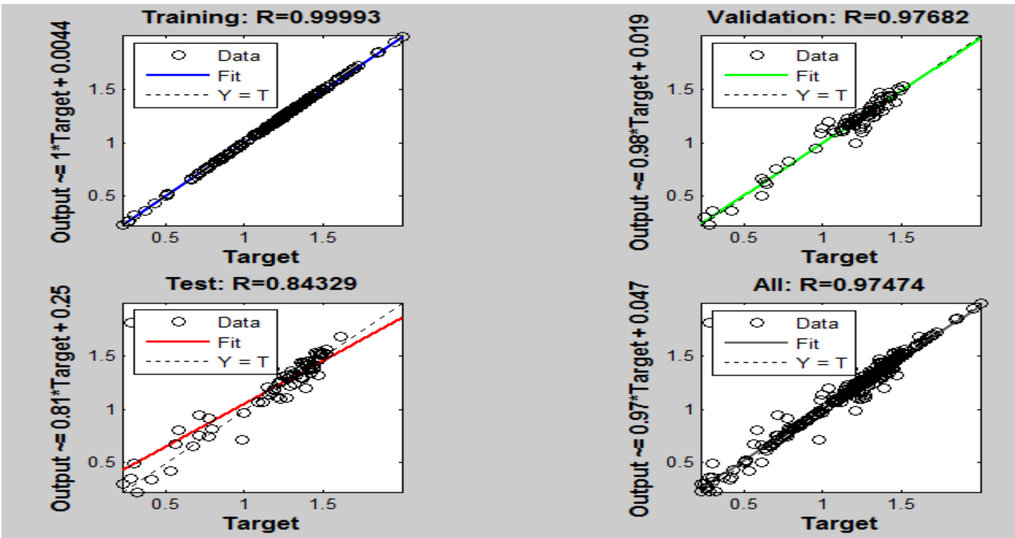

Figure (14) The Training Regression.

CONCLUSIONS

A self-organizing map (SOM) based retrieval system is proposed for performing ear matching in large database. The proposed system provides a small subset of ear that are most similar to a given query ear, from which user can easily verify the matched images. The architecture of the proposed system consists of two major parts. First, the system provides a generalized integration of multiple feature-sets using multiple self-organizing maps. Second, a SOM is trained to organize all the ear images in a database through using the compressed feature vector. Using the organized map, similar ear to a query can be efficiently identified. SOM is statistic-based Ear Recognition algorithm.

In this paper an improved SOM method is proposed. The recognition is obtained for 4 persons' 8 images of database, where the training is done on four images only and tested on remaining images. Thus, the proposed method is an efficient ear recognition process.

REFERENCES

[1]. Misiti, M. Y. Oppenheim and G. Poggi, "Wavelet Toolbox User's Guide", Version 2.1.The Mathworks, Inc. 2000.

[2]. Chang, S. GB Yu and M Vetterli, "Adaptive Wavelet Thresholding for image Denoising and Compression" IEEE Transactions on Image Processing, Vol. 9, No. 9, September, 2000.

[3].Kumar, C.S. Rai and Shakti, "Dinesh Face Recognition using Self-Organizing Map and Principal Component Analysis", IEEE, pages 0-7803-9422-4/05, 2005.

[4].Ni'matus Sholihah, "Ear Recognition using Hybrid Features and Multilayer Neural Networks", Department of computer graphics and multimedia, University Technology Malaysia, 2011.

[5]. Bhanu and H. Chen, B. "Human ear recognition in 3D", In Workshop on Multimodal User Authentication, pages 91–98, 2003.

[6]. Yan and K.W. Bowyer, P. "Ear Biometrics Using 2D and 3D Images", Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshop Advanced 3D Imaging for Safety and Security, June, 2005.

[7].Yin, Razdan and Corner, "EARS: A System for Geometric and Anthropometric Evaluation of Human Body Scans", Computer-Aided Design and Applications, Proceedings of CAD Conference, university Wisconsin, New York, vol. 6, pages 431-445, (2009).

[8]. Pun and Y. Moon, K. "Recent Advances in Ear Biometrics", In Proceedings of the Sixth International Conference on Automatic Face and Gesture Recognition, pages 164 – 169, May 2004.

[9]. Jain, R. Bolle, . and S. Pankanti, "Personal Identification in Networked Society", Kluwer Academic Publishers, Boston, 2002.

[10]. Victor and K. Chang, AB. "Comparison and combination of ear and face images in appearance- based biometrics", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25 no. 9, pages 1160-1165, September, 2003.