Simulation Analysis of Multimodal Biometric Recognition using Artificial Neural Network Algorithm

Satish S. Babar[1], Sanjay S. Badhe[2], Sachin D. Indalkar[3]
Department of Electronics and communication Engineering
[1, 2. 3] Dr. D. Y. Patil College of Engineering, Ambi, Pune.

Abstract-*Personnel identification and verification systems are emerging fields in recent years. Biometric identification is considered as best method for offering security, to avoid unauthorized access. Multimodal biometrics is used now a day because the previously used unimodal biometric system has various drawbacks such as unacceptable error rates, noisy data and variations. So, a reliable and better recognition system requires robust multimodal biometric systems. Here, a new multimodal biometric system is proposed for recognition purpose called Artificial Neural Network (ANN). Face and Ear are the two traits used here. The benefit of using ANN is that it has got that mush tolerance level that even if captured image is little bit vary from its original place or position, ANN will not through an error. Paper presents the simulation results of the human identification using ANN. Error reduction is done by using haar filters. We must train the system to identify or match the face and ear. Based on the input image comparison will be done and system will produce the results accordingly.*

Keywords-Artificial Neural Networks; feature extraction; image processing; training;

I. INTRODUCTION

Authentication is first process we all have to do when we are visiting any organization. There are numerous types of authentication available, but here we will talk about biometric authentication. Because biometric identification is an effective method for human identification. Also, it is providing reliable personnel recognition systems to either confirm or reject the identity of an individual [1].

Conventional authentication systems based on passwords or smart cards has so many difficulties in remembering the passwords or in case forgetting smart cards but in case with biometric systems person don't have to carry anything with him or her as face and ear are the traits we are using for identification. In unimodal identification system, then also there will be some problems in the system such as non-informality, noisy data, limited degree of freedom, unacceptable error rates, spoof attacks. Suppose in case of finger print identification we have seen than some time fingerprint scanner doesn't recognize the person just because there is small cut at the finger. So just because of small issue systems fails to identify the person. Or we can take case of face recognition also. In our mobile phone, if we set unlock option as a face recognition then only by looking at the front camera our mobile phone will unlocks. Here also distance matters. The distance from camera to our face is fixed. Or may varies a little bit. If we keep our face too far from the camera, then camera won't be recognizing the person. So these kind of problems are there in unimodal recognition system. To overcome this issue, instead of unimodal identification system we can use multimodal identification system.

Multimodal identification system is more accurate compared to unimodal because in multimodal we can have multiple evidences of biometrics [2]. For example, suppose my face and ear are saved in a system, and if system has to identify me, then even if system fails to recognize my face and if it is recognizing my ear then also depending on the program, system can identify me. So chances of failure the less as compare to unimodal identification system. The important concept of multimodal biometrics system is that it integrates multiple sources of information obtained from different biometric cues [3], [4]. We can take either human behavior or physiological biometrics for identification purpose. Now, in human behavior let's say voice is the thing which we take for identification. If by any chance human voice changes then there will be a problem in identification of that person. But physiological biometrics has many advantages over this kind of problems. Human face or ear rarely changes or even we can say it does not change at all for many years. So obviously, instead of carrying smart card or remembering passwords for many years will not be feasible.

The use of ear in human identification system has more advantages such as, it is stable and the features are fixed and unchangeable [1]. Similarly, the face is also a passive identification method with some drawbacks like illumination pose change in unimodal biometric recognition system [5].

Even by using multiple traits for the project or system complexity increases but it give the more accurate results. In case of ANN, same program it has to run for two times because here we have both the traits as photo only and that too also of face and ear. So obviously system has to save the database first for both the objects that is ear and face. Then it will compare.

Many studies and algorithms have been proposed for multimodal biometric fusion. The fusion of multimodal biometric system information can be performed at three different levels, namely, feature level, matching score level and decision level [6]. The LBP method used for feature extraction from face has various shortcomings, which leads to the requirement of a better feature extraction method so Mrs. PranaliPatil [7] present another technique called ULBP (Uniform Local Binary Pattern) which is applied for the face image feature extraction in multimodal biometric system. The application of ULBP can improve the overall accuracy of the system for better recognition. Also, the dimension of the extracted features can be reduced using ULBP. The histogram feature of the extracted ULBP pattern is plotted to serve as a useful tool for thresholding. But here also error rate and system process is not easy as compare to ANN. So here in this paper, we are trying to get the same results by another method called Artificial Neural Network Algorithm. The advantage of using ANN is that has got tolerance band. So even if captured image varies a bit, ANN will not through an error. This is not the case for other systems. In image processing, each image is defined by pixels. So if you compare image pixel by pixel, within a limit, system can produce the result that two images are same or not. Or for mathematics, if you compare two numbers, system straight away gives the result whether the two numbers are equal or not. But in case of ANN this is not the case.

II. LITERATURE SURVEY

Literature review presents several techniques for multimodal biometrics using the face and the ear images. Jitendra B. Jawale and Anjali S. Bhalchandra [8] have designed an ear based multiple geometrical feature extraction method. The important purpose of the design was to identify a person using ear biometrics. Generally, the human ear was a perfect source of data in many applications such as person identification. The ear biometric was a good solution, which satisfies the increasing need for security in various public places. The only reason behind this is that the ears were visible and in addition, it can be obtained easily even without the knowledge of the examined persons. Several advantages of the ear has enabled it to be best suitable for the ear biometric identification.

To satisfy the need for a better feature extraction technique, Dakshina Ranjan Kisku et al. [9] have proposed the Gaussian Mixture Model for face and ear biometrics. This model has made use of the Gabor wavelet filters to extract the facial and the ear features from the spatially enhanced face and ear biometrics. To create the measurement vectors of discrete random variables, this Gaussian mixture model was applied to the Gabor face and ear responses. In GMM, the estimation of density parameters was performed by using the expectation maximization algorithm. Through the dempster-shafer decision theory, the reduced feature sets of face and ear werefused together. This proposed scheme was validated and studied using two multimodal databases. The databases that were used are the IIT Kanpur databases and the virtual databases of the face and the ear images. Compared to the existing techniques, this work was proved to have better accuracy and significant improvements.

Fusion is an important process in multimodal biometrics. Ning Wang et al. [10] have proposed a complex fusion method at both the pixel level and the feature level. It was designed to overcome the various problems that have occurred in multimodal biometrics fusion. The designed method was applicable at the pixel level complex fusion. Based on the complex vectors the method was used to fuse the visible and the thermal face imagery. Here, the theoretical derivation at the pixel level complex fusion was extended to 2D based classification methods such as (2D) 2 PCA, (2D) 2 LDA and (2D) 2 FPCA. The accurate covariance matrix was evaluated by these methods and it reflected the differences inherited from separate sensors as well. The above methods were evaluated with the multimodal database, NVIE and it has been shown that the designed method was more efficient inidentification and verification.

Due to the difficulty in observing the facial expression by computer, PriyaMetri et al. [11] have designed a method for the recognition of emotions from the face, the hand and the body posture. Here, they have analyzed a way that enables the computer to be more aware of the user's emotional expressions. The method designed by them can recognize the emotions from facial expressions. The multimodal emotion recognition system was used by them that consist of two different models, one for facial expression recognition and the other for hand gesture recognition. Further, based on a third classifier, they have combined the results of the two different models. The experimental results have shown that the multimodal biometric system was capable of providing better results for recognition.

III. PROPOSED APPROACH

The whole and sole purpose of multimodal identification system is to improve the quality of identification or recognition over the conventional methods by comparing the results of multiple human traits. Here we are using face and ear as recognition objects to identify the individual using Artificial Neural Network Algorithm. Also, choosing a right modality is a challenging task in identification system. The proposed multimodal system is shown in Fig.1 the steps includes Image capturing, feature extraction, normalization, training, simulation, comparison and finally result. Image can be capture using camera with fixed distance and lightening condition and that can be given to the system. There are so many parameters which can be extracted from an image. Like color, texture, shape or posture. Posture is the feature here we will be extracted from the image. For the ANN training we need to normalize the image which can be done here. We have to train the ANN first. That means we have to tell the system first what data we are feeding. Each face and ear we have to mention to the system. After training, images can be saved in a database. Now at the time of simulation same process will be done till normalization. And after normalization is data is compared with our database to find out the identity. This is how we can achieve multimodal identification using ANN. The implementation is done using MATLAB and the performance of the algorithm will be evaluated.

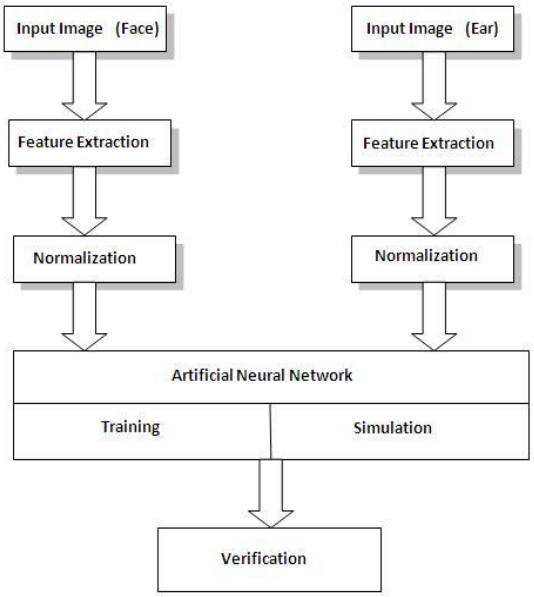

Fig.1: Proposed Multimodal System

A. Preprocessing

In the preprocessing stage of multimodal biometric system, the photos of face and ear has to be captured and keep in specific folder in the system as our database. Care should be taken at the time of capturing images that all the sample photos should be capture from specific distance only, in the same lightening conditions. They should be clear, not blurred. This is the only thing we have to do in the preprocessing, nothing else. As it is image can be given to the program for processing.

B. Training process for Face

Training process is same for both i.e. face and ear. It consists of following steps: Image acquisition, filtering image, neutralizing, segmentation, feature extractions, wavelet transform and database creation.

In the image acquisition, As images can be captured from same distance and in same lightening system, putting them into a database, we have to put them according to their respective labels. That means we know who is the person to be captured but system doesn't know. Here two things we have to mention, one is image itself and second one is its label. These two things we can concatenate and can be sending for further processing.

If any input is string array, then result is string array. If any input is cell array, and none are string arrays, then the result is cell array of character vectors. If all inputs are character array then result is character array only, for character array inputs, scrcat removes trailing ASCII white space characters: space, tab, vertical tab, newline, carriage return, and form feed. For cell and string array inputs, scrcat does not remove trailing white space.

In the filtering process, wavelet decomposition filter is used. It returns the wavelet decomposition of the matrix x at level N, using the wavelet named in character vector. Decomposition low-pass filter is used. Vector c is organized as a vector with A(N), H(N), V(N), D(N) where A, H, V and D are each a row vector. Each vector is the vector column-wise storage of a matrix. A contains the approximation coefficients, H contains the horizontal detail coefficients, V contains the vertical detail coefficients, D contains the diagonal details coefficients. In the neutralizing process, captured images are brought to the neutral level by wavelet decomposition filter. [Ea, Ed]=wenergy(C, L) returns Ea, which is the percentage of energy corresponding to the approximation and Ed, which is the vector containing the percentage of energy corresponding to the details. The minmax block outputs either the minimum or the maximum elements of the inputs. We can choose the function to apply by selecting one of the choices from the function parameter list. If the block has one input port, the input must be scalar or a vector. The block outputs a scalar equal to the minimum or maximum element of the input vector. If the block has multiple input ports, all nonscalar inputs must have the same dimensions.

In the segmentation process, image is separated from its background so that further processing will be on the image itself only rather than unnecessary background image or extra pixels.

In the feature extraction process, as shown in Fig. 2 image features are extracted and send it to wavelet transform filters.

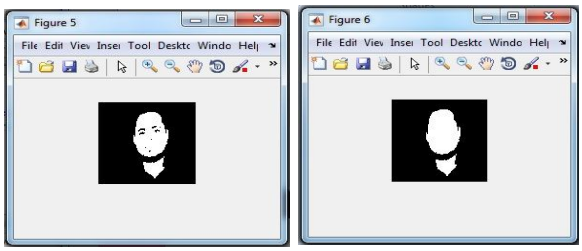

Fig.2 Feature extraction

In the wavelet transform and database creation process, Image and its label is stored in the database as shown in Fig.3 and final details of the process are as shown in the Fig.4 . The block expands any scalar inputs to have the same dimensions as the nonscalar inputs. The block outputs a signal having the same dimensions as the input. Each output element equals the minimum or maximum of the corresponding input elements. The MinMax block ignores any input value that is NaN, except when every input value is NaN. When all input values are NaN, the output is NaN, either as a scalar or the value of each output vector element.

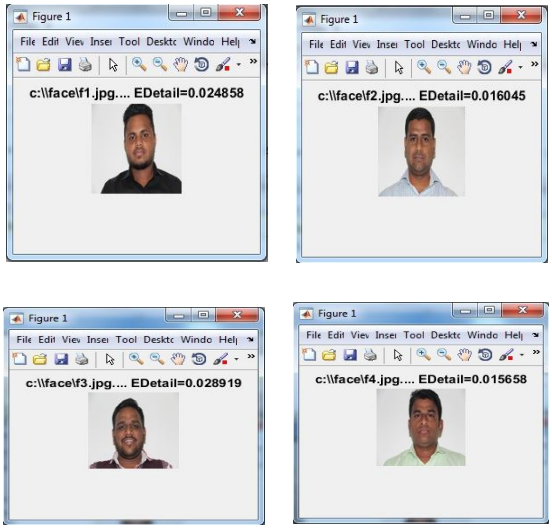

Fig.3 Sample face images as Database for ANN Algorithm

In the training process, traingdx is a network training function that updates weight and bias values according to gradient descent momentum and an adaptive learning rate. Net.trainFcn='traingdx' sets the network trainfcn property.

Training occurs according to traingdx training parameter shown here with their default values:

net.trainParam.show=1; Epochs between displays
net.trainParam.showCommandLine=0; generate command line output

net.trainParam.showWindow = 1; show training GUI
net.trainParam.epochs = 5000; Maximum number of epochs to train
net.trainParam.goal = 0.00001; Performance goal
net.trainParam.max_fail = 10; Maximum validation failures
net.trainParam.mem_reduc = 5; Factor to use for memory/speed trade off.
net.trainParam.min_grad =1e-10; Minimum performance gradient
net.trainParam.mu = 0.001; Initial Mu
net.trainParam.mu_dec = 0.1; Mu decrease factor
net.trainParam.mu_inc = 10; Mu increase factor
net.trainParam.mu_max = 1e10; Maximum Mu
net.trainParam.time = inf; Maximum time to train in seconds

The function traingdx combines adaptive learning rate with momentum training. It is invoked in the same way as traingda, except that it has the momentum coefficient mc as an additional training parameter. traingdx can train any network as long as its weight, net input, and transfer functions have derivative functions. Backpropagation is used to calculate derivatives of performance perf with respect to the weight and bias variables X. Each variable is adjusted according to gradient descent with momentum,

dX = mc*dXprev + lr*mc*dperf/dX

where dXprev is the previous change to the weight or bias. For each epoch, if performance decreases toward the goal, then the learning rate is increased by the factor lr_inc. If performance increases by more than the factor max_perf_inc, the learning rate is adjusted by the factor lr_dec and the change that increased the performance is not made.

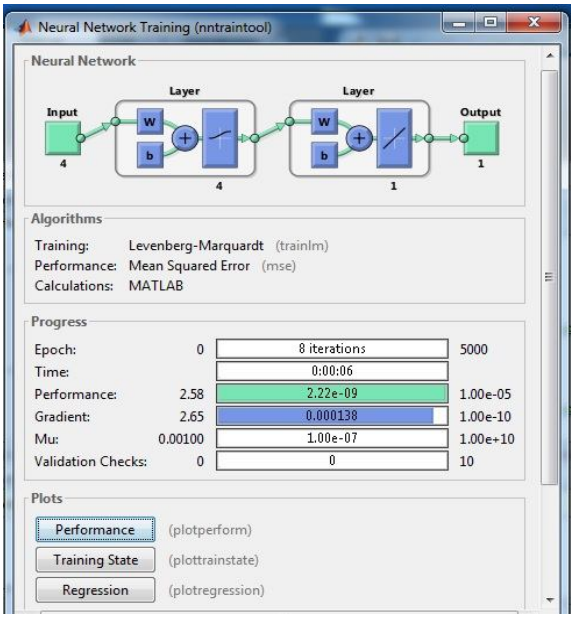

Fig.4 ANN Training details for Face

Training stops when any of these conditions occurs:

- The maximum number of epochs (repetitions) is reached.
- The maximum amount of time is exceeded.
- Performance is minimized to the goal.
- The performance gradient falls below min_grad.
- Validation performance has increased more than max_fail times since the last time it decreased (when using validation).

The echo command controls the display (or *echoing*) of statements in a function during their execution. Normally, statements in a function file are not displayed on the screen during execution. Command echoing is useful for debugging or for demonstrations, allowing the commands to be viewed as they execute.

Y = round(X) rounds each element of X to the nearest integer. In the case of a tie, where an element has a fractional part of exactly 0.5, the round function rounds away from zero to the integer with larger magnitude.

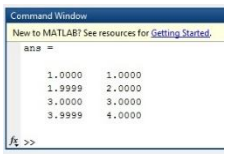

Fig.5 Rounding off

C. Training process for ear

As we can see, from the above face training process, images of the faces are captured, filtered, neutralized, segmented and finally through ANN training network they saved in a database for identification purpose. Now the same process has to be follow to create ear images database. Everything will be same except input image. And results are as follows:

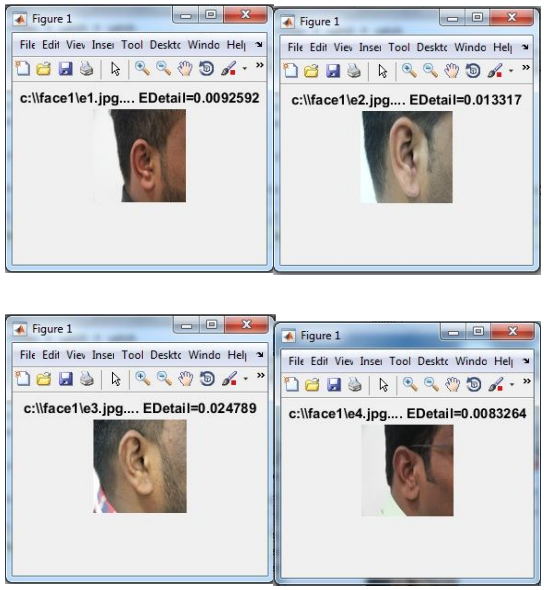

Fig.6 Sample ear images as Database for ANN Algorithm

Training process details are as follows :

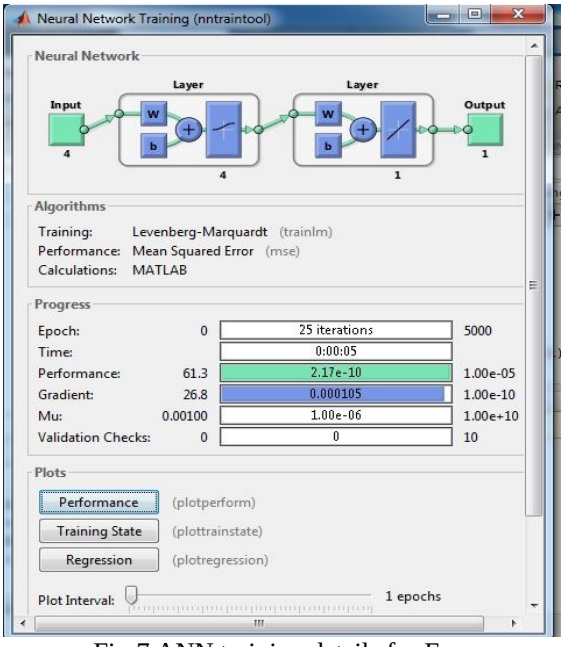

Fig.7 ANN training details for Ear

D. Simulation Process

Most of the simulation process is same as training process. In the simulation process also, we have to take input image, process it and finally through ANN algorithm it will

compare input image data with the data present in database. And depending on the closeness of the matching record, ANN will show the result. If data is closely match with the database then it will extract the name of the person which we had saved in the training process. If the data is not matching with the database then ANN will show accordingly that person is not identified because it is not there in the database.

We have created GUI browse option in the program so that when we run the program, we can be able to select the image from the respective folder.

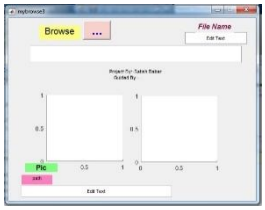

Fig.8 GUI to select image from Floder

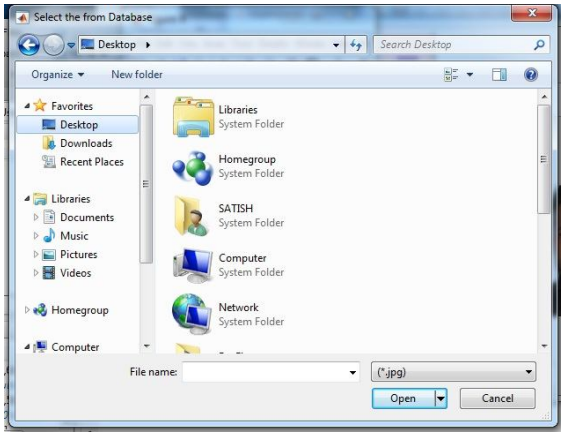

Fig.9 Fetching data from the Database

In the simulation process, we have to pick two images as we said we have two traits for comparison, the face and the ear. Through browse option we will pick the images of face and ear.

Here we will take two cases, one will be the correct data and second will be the wrong one. As we know if we put correct data in the system then the system will give correct output as shown in Fig. 11. And when we put wrong data then the system will give result accordingly as shown in Fig. 12. This is the best part of the ANN that we know the input and we know the output.

Once we select the image, ANN will perform all the preprocessing task to bring up the image parameters according to the training process. Through ANN algorithm it will match the data of the face and ear image with data present in the database and finally ANN will produce the result.

Fig.10 Image Detection

IV. RESULTS AND DISCUSSION

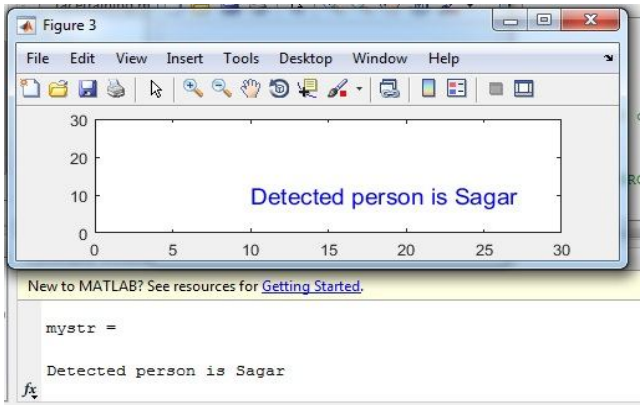

Fig. 11 Result1

So here you can see from the above results that without doing much editing and any other modifications in the input images ANN algorithm has given the final results. As ANN has its algorithm to verify the data provided, its accuracy is more than the any other identification system present. In the advantage, we have used two traits for further more accuracy. So even if one condition fails, system should not show false result, it will wait for second step verification. So here in ANN algorithm, possibility of errors we have minimized by providing two samples to the algorithm.

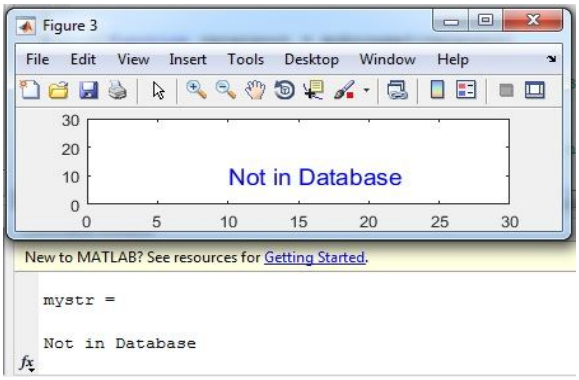

Fig.12 Result2

Artificial Neural Network well known for its excellence in the automatic weight reduction procedure which brings down the errors close to zero level. Following graph shows the reduction in the weights throughout the simulation process.

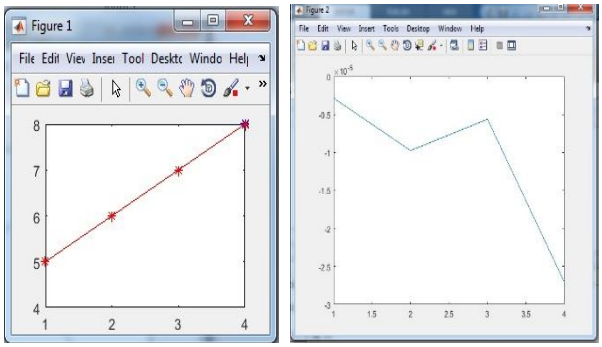

Fig.13 Error rate reduction

V. CONCLUSION

In this paper, an Artificial Neural Network algorithm is being proposed for human identification using face and ear images. This proposed algorithm has provided the better results than the existing methods and it can be applied for human identification system. ANN is well known for its automatic weight reduction procedure which reduces the errors and brings down the errors close to the zero level which increases the accuracy of the system. The designed method have proved that it is capable of improving the performance of the multimodal biometric system. The evaluation metrics have revealed that the designed method provides better performance and improved accuracy rate.

REFERENCES

[1] Ramesh Kumar. P and K.Nageswara Rao, "Pattern Extraction Methods for Ear Biometrics – A Survey", World Congress on Nature and Biological inspired computing, pp.1657-1660, 2009.

[2] Steven Cadavid, Mohammad H. Mahoor and Mohamed Abdel- Mottaleb, "Multi-modal Biometric Modeling and Recognition of the Human Face and Ear", International Workshop on safety security and Rescue Robotics, pp.1 – 6, 2009.

[3] A. K. Jain and A. K. Ross, "Multibiometric systems," Communications of the ACM, Vol. 47, No.1, pp. 34–40, 2004.

[4] A. Rattani, D. R. Kisku, M. Bicego, and M. Tistarelli, "Robust feature-level multibiometric classification," Proceedings of the Biometric Consortium Conference – A special issue in Biometrics, pp. 1–6, 2006.

[5] Wei Jin, Bin Li and Ming Yu, "Feature Extraction Based on Equalized ULBP for Face Recognition", International Conference on computer science and electronics engineering, Vo.2, pp. 532 – 536, 2012.

[6] Yeong Gon Kim, Kwang Yong Shin, EuiChul Lee and Kang Ryoung Park, "Multimodal Biometric System Based on the Recognition of Face and Both Irises",International Journal of Advanced Robotic Systems, Vol.9, No.65, 2012.

[7] Mrs. PranaliPatil and Prof. Dr.D.S.Bormane "Multilevel security using score level fusion of face and ear based biometric modulities" International Journal of Engineering Research & Technology(IJERT) ISSN:2278-0181 vol.3 Issue 5, May 2014.

[8] Jitendra B. Jawale and Anjali S. Bhalchandra, "The Human Identification System Using Multiple Geometrical Feature Extraction of Ear –An Innovative Approach"International Journal of Emerging Technology and Advanced Engineering, Vol.2, No.3, 2012.

[9] Dakshina Ranjan KISKU, Phalguni GUPTA, Hunny MEHROTRA and JamunaKanta SING, "Multimodal Belief Fusion for Face and Ear Biometrics",Intelligent Information Management, Vol.1,pp. 166-171,2009.

[10] Ning Wang, Qiong Li, Ahmed A. Abd El-Latif, Jialiang Peng and XiamuNiu, "Multibiometric Complex Fusion for Visible and Thermal Face Images", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 6, No. 3, 2013.

[11] PriyaMetri, Jayshree Ghorpade and and Ayesha Butalia, "Facial Emotion Recognition Using Context Based Multimodal Approach", International Journal of Interactive Multimedia and Artificial Intelligence, Vol.1, No. 4, pp. 12-15, 2012.