International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

An Agent Based Multimodal Biometric Method for Secure Systems

Nidhi Srivastava

Abstract: ***A security system based on biometrics of a person depends on the features of a person. These can be physical features or behavioral features. User authentication is necessary for security control system and intelligent human computer interface, so that only the real user accesses it. We present a method for the authentication of person with help of multimodal data which includes face and speech. Agent based system has been used for fusion of these data which gives us a more reliable, efficient and secure biometric system. A method for face detection for recognition of faces has been given. For speech recognition Artificial Neural Network has been used. Both these combined together identify the speaker. This multimodal biometric method will give highly reliable secure system.***

Keywords**:** ***ANN, Audio-Visual Fusion, Biometric, Face Detection, Speech Recognition.***

I. INTRODUCTION

Biometrics is the science where digital technology is used to recognize the identity of individuals based on behavioral or physiological characteristics. Recently the need for safe and secure systems has grown tremendously. The first essential step for security systems is to authenticate the person who is to use the system. The earlier traditional methods of security like use of smart cards and keys, use of password and pins, etc. have become obsolete and also lack security. These can easily be stolen by anyone. A security system based on the physiological characteristics of a person rather than based on the above methods can significantly reduce the frauds and increase the security aspect [1,2]. The different physiological characteristics of a person which can be reaped for security systems include modalities like speech, iris, retinal scan, face, fingerprint, hand geometry, etc. The most favorable, technically worked upon modalities include speech, recognition of face, matching of fingerprint [1,2]. Although biometric systems using single modality give good performance, but so as to increase the performance and robustness of the systems based on biometric more than one modality can be used. A multimodal system will give really high performances. Use of modalities which are dissimilar in nature can provide independent and complementary information, thus improving difficulties typically associated with use of one modality [3]. The design of any system for recognition of a person, use of more than one modality is a difficult task and fusion of these modalities is even more difficult. We have proposed an agent fusion approach in which face and speech are fused together to identify the person. Section 1 gives the introduction. Section 2 gives the proposed method for user authentication. Section 3 gives the implementation and result. It shows face recognition, speech recognition and the fusion of two together with the help of agents. Finally Section 4 gives the conclusion.

II. PROPOSED MULTIMODAL BIOMETRIC METHOD

A method has been given in the paper which utilizes the face and the speech of a person so as to identify the real person. In my proposed method, fig. 1, firstly face detection of the person is done, followed by the speech recognition. Through this we identify the person. This information is finally passed to an agent which fuses the above two data and recognizes the speaker. In this we have used two biometric data - face and speech. Through this the person identification becomes easy and it is precise and quick as compared to the two data taken individually.

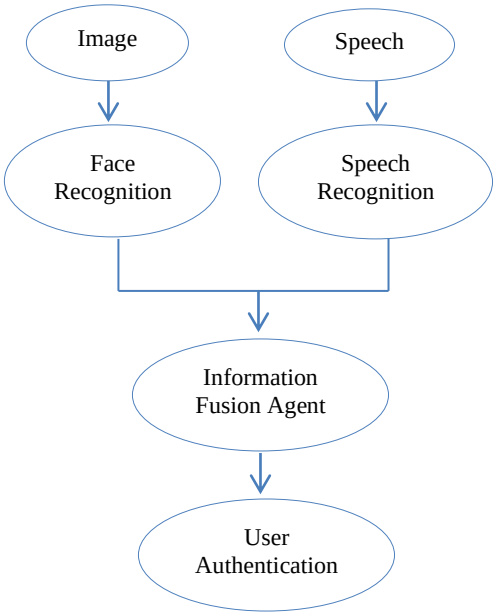

Fig. 1: Multimodal Biometric Method

III. IMPLEMENTATION & RESULT

A. Face recognition:

For identification of the person and the recognition, face detection plays an important role. In this the face of the person is detected from the given image.

Manuscript published on 30 August 2019.
***Correspondence Author(s)**
Dr. Nidhi Srivastava, Asst. Professor, Amity Institute of Information Technology, Amity University, Uttar Pradesh, Lucknow Campus, Lucknow, India.

For humans, detection of the face is very insignificant but same is a challenge for the machines. There are many challenges in detection of the face like presence of glasses, beard, mustaches, face posture, expression of face, face occlusion, etc. Also the shape and size of the face and skin color varies from person to person [4].

Face detection plays a crucial role in certain applications like face recognition, face expression recognition and lip reading. Intelligent systems and security related systems find it important to detect the human faces from the given images or videos. For detection of human faces there are many different approaches like Template matching methods, Knowledge-based methods, Appearance-based methods and Feature invariant approaches

- Knowledge based methods: These methods make use of the rules which are made based on the knowledge of the human face. Easy, simple and straightforward rules are made by the researchers depending on the facial features like location of the two eyes and the distance between these two, the symmetry between nose, eyes and mouth. The face can be easily localized based on this method.
- Feature invariant approaches: As the name suggests this method emphasizes on the features of the face. Various approaches are used to detect the different features of the face and extract the mouth, nostrils, ear, and eyes from the given image. Statistical model is then built from the extracted features and relationship is created between them which verify the presence of the face. This includes detection through features of the face, texture of the skin, color of the skin and multiple features.
- Template matching methods: In this a prototype of the face is created of the standard front face which is manually predefined or parameterized by a function. This is called as the template. Correlation between the input face and the template is calculated for each of the features which help in deciding if a face exists or not.
- Appearance-based methods: In this some sample images are taken which are trained and based on these images templates are built. This method makes use of machine learning and statistical analysis tools to detect images that have face and images which do not have face. Some of the systems used in this are Neural Network, Information-Theoretical Approach, Eigen Face, Naïve Bayes Classifier, distribution based, Support Vector Machine, Hidden Markov Model, etc.[5,6]

In this approach for detection of the face we have made use of skin color which is a feature invariant approach. Various approaches exists which help in identifying the skin area and non-skin area like RGB, normalized RGB, YCbCr, HSV, YUV, YIQ, clustering, etc. There are many clustering algorithms like Hierarchical Clustering, K-Means Clustering, Fuzzy Clustering etc. which are used in various applications. Among different kinds of clustering algorithm, K-means is a very straightforward clustering algorithm. It is one of the simplest unsupervised learning algorithms and is extensively preferred because of its simplicity, practicality and effectiveness [7]. In this paper I tested my images with three different methods for classification of the skin color-RGB, YCbCr and K Means. We tested these methods on a database of thirty images taken under same illumination and lighting conditions. The experiment has been done on MATLAB 2015a using image processing toolbox. Skin color detection was done on the images. For detection of the human face more work is to be done on the obtained image. Binary image containing 1 and 0 is derived where skin region is shown with 1 and non-skin region is shown with 0. Morphological operations are then performed on this image. All regions which fail to have the required number of pixels needed for face are discarded. Then morphological opening is performed followed by filling the small hole regions. The face too contains hole in place of eyes and nostrils. These are filled so as to get a connected region. Then morphological closing is performed. Finally width to height ratio is calculated and if it is found to be within the specified value the region is declared to be face [11]. For detection of skin color each of the above mentioned method was individually tested on our database and finally used to detect the face. Each of the three color spaces used in the experiment are described below and also face detection is done. Out of thirty, face detection in three images is shown.RGB is the most commonly used color space and it matches the primary colors which are three in number - red, green and blue, respectively. From the database images, we have taken a general upper and lower range of the colors in which the skin colored pixels will lie. We took out roughly the red, green and blue value of the skin colored pixels. So, a threshold for the RGB values was taken.

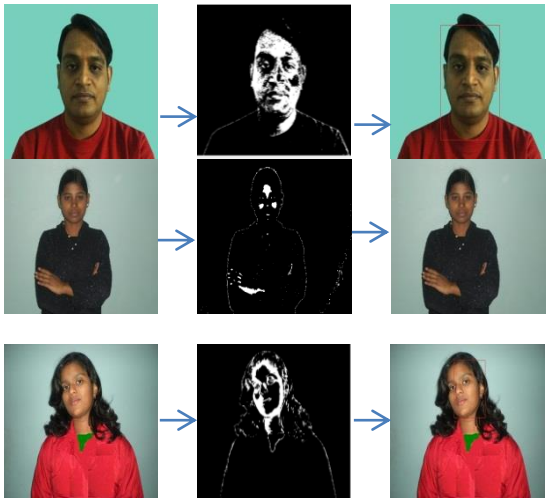

Fig. 2: The Original, Binary And Detected Face When RGB Is Used

For each of these colors a minimum and a maximum value was taken out. Further, each of the pixels in the image was checked if it lied between the given values of RGB. Fig. 2 shows face detection in three images. Face in first image was successfully detected. In second image the face could not be detected and in last image only half of the face was detected. The YCbCr belongs to the family of television transmission color spaces and has two parts-luminance and chrominance.

Retrieval Number: J92180881019/19©BEIESP
DOI: 10.35940/ijitee.J9218.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

The color is represented as weighted sum of RGB values known as luminance (Y) and chrominance. The difference between the luminance and B and R values, gives us the chrominance. The luminance information is stored in Y. The chrominance is divided into two parts, Cb and Cr. On subtracting the blue from a reference value we get Cr, and on subtracting Cr from red value we obtain Cb. YCbCr is preferred for detection of skin as it is simple and very easily separates the luminance and the chrominance component [8]. The value for Cb and Cr for skin is as given in [9]. The result is given in Fig. 3. The face in the first and second image was easily detected. But in third image the face is not detected.

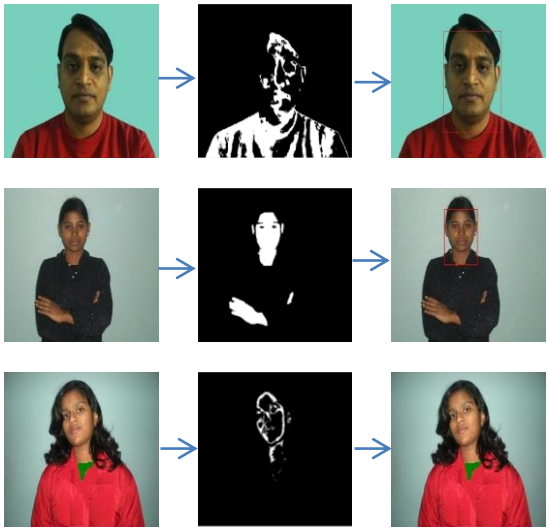

Fig. 3: The Original, Binary And Detected Face When Ycbcr Is Used

K-means clustering, depending on the attributes, classify the objects into different groups, k. In this the number of clusters is fixed and the given data is classified through this algorithm. We assume that there are k clusters.

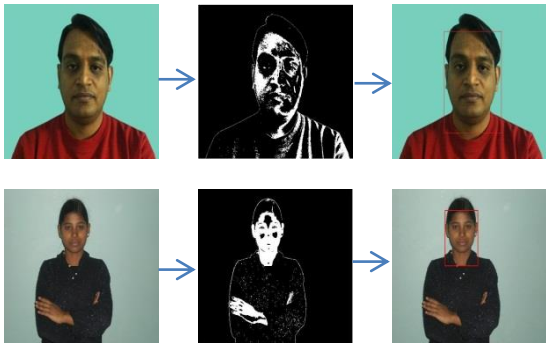

Fig. 4 The Face In Each Of The Image Could Be Easily Detected.

Fig. 4: The original, binary and detected face when K-Means is used

The procedure used in K means can be described as follows. Number of cluster, k is set and centroid coordinate is determined. Distance between the centroid and the object is calculated. Then the objects are grouped together based on the minimum distance. The above steps are repeated unless no object changes its course from one to the other group [7]. In K-means clustering, selection of K is critical. In this the value of K is taken as 5 and based on this value of k the segmentation of the color is done [10]. As can be seen from Experimental results show that in the images taken by us, for skin color segmentation and final detection of face K-means gives the best result. When RGB was used for skin color segmentation in the algorithm then of the total 30 images only 18 faces were correctly detected (Fig. 2). When YCbCr was used then 21 faces were correctly detected (Fig. 3) and when K means was used then 29 faces could be correctly identified (Fig. 4). The comparison between the three algorithms and implementation to detect face is shown in table 1. The table also shows the average time taken to detect face when different algorithms are used.

Table I: Results of our Experiment

Algorithm used	No. of Faces Detected	No. of Faces not Detected	Accuracy (Percentage)	Time Taken (sec)
RGB	18	12	60%	2.2
YCbCr	21	09	70%	3.5
K-Means	29	01	96.66%	4.2

It is very clear from the table that when K means is used the total time taken to detect face is far more as compared to that of RGB and YCbCr. But as far as accuracy is concerned, it is much high as compared to that of other two and so a compromise on time can be made. Once the face has been successfully detected it is matched with the existing database of faces using Neural Network. If the face matches, the person is identified else not.

B. Speech Recognition: The speech of the person is preprocessed to remove noise and then the features of the speech are extracted and classified and finally the speech is recognized. In our experiment MATLAB 2015a is used. Voices of the subjects whose faces are stored in the database are recorded and stored. MFCC is used for feature extraction and Back propagation algorithm of Neural Network is used for speech classification. Twelve coefficients of MFCC have been used.

C. The network is first created and then it is trained and finally simulated. Neural Network toolbox was used and the performance was evaluated using mean square error.

The square of the output is subtracted from the square of the target and the result is averaged which gives us the Mean Squared Error (MSE). Lower values of MSE are preferred and regarded as good [12]. Zero value means no error. Fig. 5 in our experiment shows that the network error touches almost zero. The Receiver operating characteristics (ROC) graph Fig. 6 is towards left and top. This gives good result and the network could definitely identify the speaker. This system achieved almost 95% accuracy.

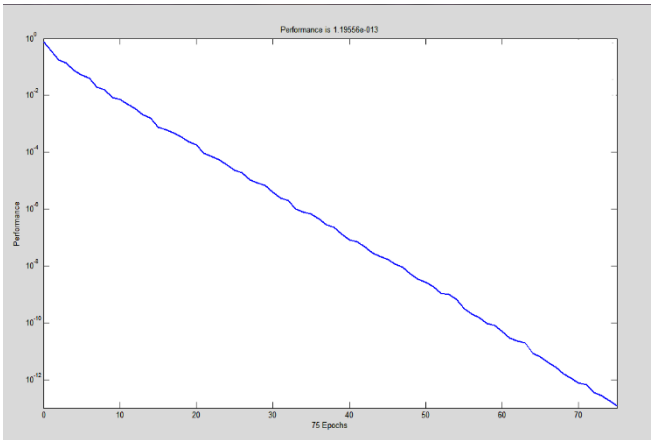

Fig. 5: MSE In The Training Phase With Feed Forward Back Propagation Network

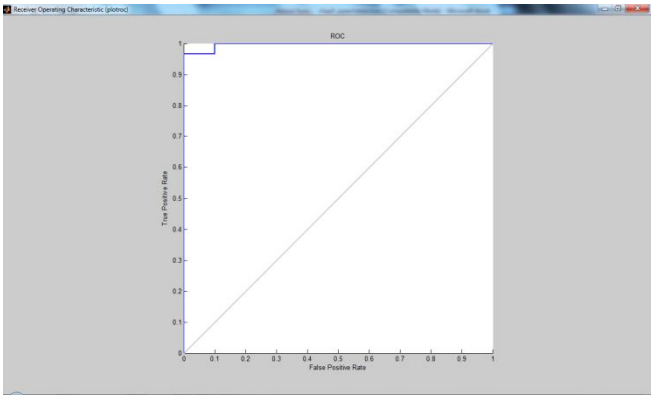

Fig. 6: Receiver Operating Characteristics (ROC) Graph

C. Information Fusion Agent: The information from various biometric modalities can be fused together through various techniques available. Agents have been used in this paper. An agent is a self-sufficient, autonomous entity. The agent comprehends the information which it receives from the surroundings and the given environment and based on this takes decision about what action it is to be done [13].

Agents are small programs designed to perform a specific task. These agents have the capability to learn and reason and communicate in the environment and act the way they are designed to behave [14]. The fusion of the modalities can be attained at different levels as presented in fig.7 [15]. This includes fusion at different levels like data, feature, score and decision. At data level the raw data is captured and fused.

In fusion at feature level, features of the various biometric attribute are extracted and combined to gather information. The score level fusion compares and shows the similarities between the gathered feature vector and the feature vector stored in template. At the decision level fusion, each of the traits feature is extracted and classified as accepted or rejected. In our case the decision level fusion system is appropriate to be used as we have got the decision from both the modalities, face and speech, and these can be fused together to give us the result. The agent fuses the data given by the two modalities and gives the desired information.

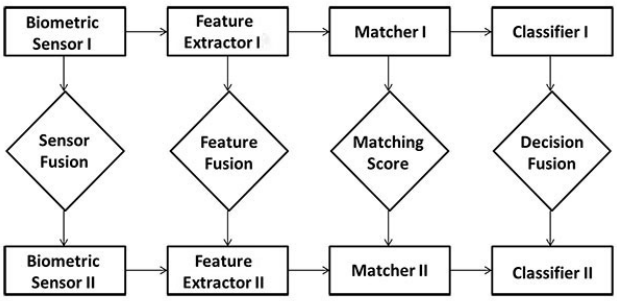

Fig. 7: Fusion levels in Multimodal Biometric Systems

IV. CONCLUSION

In this paper, we have given an agent based multimodal biometric method which can be easily used to authenticate the user and improve the security of a system. Firstly the face detection technique is given which can easily detect the front face of a person. This is a robust algorithm. This algorithm is capable of detecting skin region in the given image and also detects the face. This algorithm has been executed on MATLAB R2015a and it successfully detects face region and matches with the stored database. Further Neural Network Toolbox of MATLAB R2015a has been used for speech recognition and speaker identification. Finally, both the face and speech are combined together with the help of agents. This information fusion agent helps in the security of any system as only the person who is successfully recognized can use it. The fused multimodal biometric system is definitely more efficient and reliable as compared to the single biometric system.

REFERENCES

1. Rishiraj Mukherjee, Tanmoy Islam, Ravi Sankar. "Text dependent speaker recognition using shifted MFCC", Proceedings of IEEE Southeastcon, 2012
2. A. Kounoudes, V. Kekatos, S. Mavromoustakos. "Voice Biometric Authentication for Enhancing Internet Service Security", 2nd International Conference on Information & Communication Technologies, 2006
3. P.S. Aleksic, A.K. Katsaggelos. "Audio-Visual Biometrics", Proceedings of the IEEE, Vol. 94, No. 11, November 2006
4. Sanjay Kr. Singh, D. S. Chauhan, Mayank Vatsa, Richa Singh "A Robust Skin Color Based Face Detection Algorithm" Tamkang Journal of Science and Engineering, Vol. 6, No. 4, pp. 227-234, 2003
5. Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, "Detecting Faces In Images: A Survey", IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, pp.34-58, January 2002.
6. Ming-Hsuan Yang, Narendra Ahuja. "Face Detection and Gesture Recognition for Human- Computer Interaction", Springer Nature, 2001.
7. Sneha Silvia, Y. Vamsidhar, G. Sudhakar, "Colour Image Clustering using K-Means", International Journal of Computer Science and Technology, Vol. 2, SP 1, December 2011
8. Salimi, Fahimeh, Hadi Farzin, and Farzad Ebrahimi. "A novel skin detection method using Generalized Discriminant Analysis", International Conference on Computer Systems and Industrial Informatics, 2012.
9. Mohamed A. Berbar, Qatar, "Faces and Facial Features Detection in Color Images", Proceedings of the Geometric Modeling and Imaging—New Trends, IEEE, 2006.
10. Fang X., Wenquan Gu W., Huang C. "A Method of Skin Color Identification Based On Color Classification", International Conference on Computer Science and Network Technology (IEEE), pp 2355 – 2358, 24-26 December 2011

11. Nidhi Srivastava, Harsh Dev, S. Qamar Abbas "Framework for Face Detection", IJCA, Volume 58– No.17, November 2012.
12. Nidhi Srivastava, "Speech recognition using MFCC and Neural Network" Proceedings of National Conference on Challenges & Opportunities for Technological Innovation in India" 16th February 2013.
13. Maximilian Kruger, Achim Schafer, Andreas Tewes, Rolf P. Wurtz, "Communicating Agents Architecture with Applications in Multimodal Human Computer Interaction", Informatik 2004.
14. Elfriede I. Krauth, Jos van Hillegersberg, Steef L. van de Velde, "Agent-based Human-computer-interaction for Real-time Monitoring Systems in the Trucking Industry" IEEE Proceedings of the 40th Hawaii International Conference on System Sciences pp. 1-7, 2007.
15. Ashish Mishra "Multimodal Biometrics it is: Need for Future Systems", IJCA, Volume 3, No.4, June 2010 pp: 28-33

AUTHORS PROFILE

Dr. Nidhi Srivastava is currently working as Assistant professor in Amity University. Her research area is Human Computer Interaction, e-governance and Cloud Computing. She has about 14 years of teaching experience. She has published many research papers in various reputed national and international journals and conferences. She is in the editorial and reviewer board of many prestigious journals and conferences.

.