

Face Recognition with Liveness Detection using Eye and Mouth Movement

Avinash Kumar Singh, Piyush Joshi, G. C. Nandi

Robotics and Artificial Intelligence Laboratory

Indian institute of Information Technology, Allahabad - 211012, U.P, india

{avinashkumarsingh1986, piyushjoshi3839}@gmail.com, gcnandi@iiita.ac.in

Abstract— The recent literature on face recognition technology discusses the issue of face spoofing which can bypass the authentication system by placing a photo/video/mask of the enrolled person in front of the camera. This problem could be minimized by detecting the liveness of the person. Therefore, in this paper, we propose a robust liveness detection scheme based on challenge and response method. The liveness module is added as extra layer of security before the face recognition module. The liveness module utilizes face macro features, especially eye and mouth movements in order to generate random challenges and observing the user's response on account of this. The reliability of liveness module is tested by placing different types of spoofing attacks with various means, like using photograph, videos, etc. In all, five types of attacks have been taken care of and prevented by our system. Experimental results show that system is able to detect the liveness when subjected to all these attacks except the eye & mouth imposter attack. This attack is able to bypass the liveness test but it creates massive changes in face structure. Therefore resultant unrecognized or misclassified by the face recognition module. An experimental test conducted on 65 persons on university of Essex face database confirms that removal of eye and nose components results 75% misclassification.

Keywords— Face Recognition, Liveness Detection, Face Spoofing, Face Macro Features Movement, HAAR Classifier, Principal Component Analysis.

I. INTRODUCTION

Biometric authentication is a way to authenticate persons that usually a human does in his life. Every human being has some physiological or behavioural characteristics like (face, fingerprint, voice, gait, etc.), which make them unique, thereby differentiating them from others. Among all these biometric traits, we restrict ourselves to face. The history of face recognition technology started from 1960s. From the last 50 years, face recognition technology has experienced a rapid growth in itself [10]. Various methods have been proposed by various researchers so far, to recognize the person even in the bad illumination, different facial expressions, orientations, and even in partial occlusions [1]. The latest survey [2] shows that face recognition technology is the second most used biometric technology used by the market/users. Face recognition is now used in different areas like access controls, human robot interaction, surveillance, etc. Recently several researchers

quantified the integrity of the face recognition system and found that these systems are also vulnerable to different kinds of attacks. Most frequently used attack is the spoofing attack [7][8][6]. Face spoofing is an attack where attacker tries to bypass the face recognition system by placing a photo/video/mask of the enrolled persons in front of the camera. The problem lies within the working principle of Face Recognition. The principal of face recognition doesn't care about who is submitting the credentials. It only concerns about the person is enrolled or not. Hence whatever effort we will do for making our classification good and effort in using good resolution camera, nothing will help. Besides this, it helps attacker to perform their attack more accurately. This problem leads to the question where one can think about the significance of the accuracy and efficiency of the system, when the reliability is not assured. Liveness detection of the user could be a way to deal with this problem.

Therefore Researchers observed the need of security mechanism and proposed various ways to deal with this problem. On the basis of literature we have grouped possible solution in three main categories (1) liveness detection by using challenge and response method (2) liveness detection by utilizing face texture (image quality), and (3) liveness detection by combining two or more biometrics (multi-modal approach). In challenge and response method system throws some challenge in terms of eyes and mouth movement which can only be performed by real user not by photo, and analyses there response in account of the given challenges. In this regard most of the researchers [3][4][15] have utilized eye blinking, while in type-II researcher's exploits texture information (smoothness/roughness, edges, etc.) to distinguish between real and imposter [5][17]. Multimodal approach mostly uses speech and face as the combination to deal with this attack [18].

These spoofing techniques are going more complex day by day right from simple photograph to painted contact lenses and polymeric face, fingers. Hence a list of modern approaches to deal with all these circumstances is mentioned by [6][7]. He has suggested, in their report that for liveness detection, maximum utilization of face macro features is the best way; hence in this paper we utilized both the eye and mouth movement to detect the liveness with the constraint of challenge and response. Previous techniques discussed in the literature are mostly based on the eye blinking, which can be

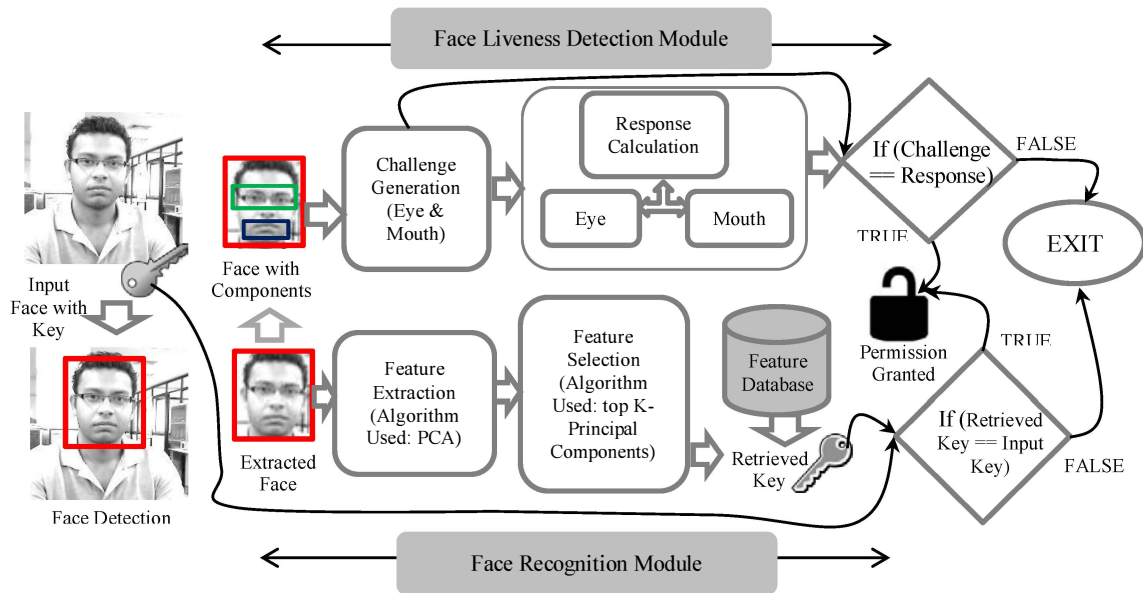


Figure 1: Face Recognition with Liveness Detection

easily forged by video imposter attack. Therefore on the basis of existing literature, we are proposing here a challenge and response technique in such a way that imposter will not be able to forge it briefly discussed in next section (challenge generation). This adds more difficulty to bypass the system, and if attacker tries to forge, it will result in wrong verification, resulting in authentication failure.

For testing our hypothesis we have designed five types of attacks by using (1) Photo Imposter Attack, (2) Eye imposter Attack, (3) Mouth Imposter Attack, (4) Eye and Mouth Imposter Attack, (5) Video Imposter Attack. The system successfully prevented attacks in all these conditions without any False Rejection Rate (FRR).

The rest of the paper is structured as follows: Section 2 describes proposed framework. Section 3 shows experimental setup that we have used in our approach followed by Section 4 that coherently states the results obtained and their discussion. In Section 5, we conclude the paper with its contribution towards the face biometrics and its future prospects.

II. PROPOSED FRAMEWORK

The work flow of the proposed framework is presented in figure 1. Here we have divided the approach into two parts (a) liveness detection and (b) face recognition. First we are testing for liveness of the user, and if person is live then the system will recognize the identity else not. Human faces are dynamic in nature and come in many forms and colours. Hence detection of human faces among different objects is even more problematic than recognizing any other object. Here, for detecting the faces and facial features, we are using Haar Classifier devised by Viola and Jones [9].

Challenge Generation:

Challenges are generated for testing whether the person is live or not. We assumed that if the person is live then he/she can move their face and facial feature too, but it is not true of all cases. This assumption will fail when attacker plays any recorded video of the genuine user, hence here we are generating the challenges in such a way that only the person who is live, can only response to those. Challenges are presented in terms of their eye and mouth movement (openness/closeness) in a sequence. These sequences are generated at random, so that no one can make a prior estimation about the challenges. The designed system is able to calculate the movement by measuring the teeth HSV value (Hue Saturation Value). Both of these challenges are generated at random, and presented at the same time (dependent of each other). This dependency is also due to the fact that attacker will not be able to forge it, even if he tries to forge it will lead to misclassification.

Response Calculation:

Responses are calculated by counting the movements (eye and mouth movement). Eye openness and closeness is calculated by searching the eye in the region where eye should exist shown in figure 2, if found we assume that eye is open else it is close. Similarly mouth openness and closeness is calculated by searching teeth (symbolized by HSV value of the teeth) in the mouth region where it likely to exist, if noted mouth is open else close. Number of challenges thrown by the system acted as the threshold and if the sum of the responses is equal to the threshold, then system will recognize the person as live else not.

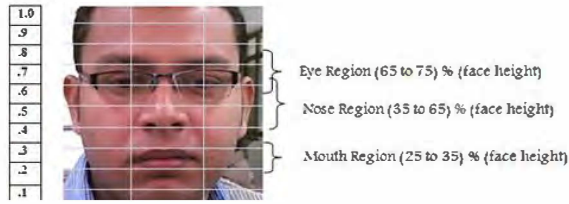


Figure 2: Human Face Fragmentation

We have used human face convention here, in which we have fragmented human face into different regions shown in figure 2. Eye will be found in eye region i.e 65% to 75% of the face height, Nose and Mouth will found in their respective regions (nose region 35% to 65% of face height, mouth region 25% to 35% of face height). The idea of fragmenting the face in this way is it reduces the search complexity. Eye, nose and mouth all are searched in their respective region.

Face Recognition:

Principal Component Analysis (PCA) [11][12][19] is good tool for dimensionality reduction as well as feature extraction. We used PCA for generating the Eigen faces [14][16], and Euclidean distance for recognition the person based on the distance between the test and train images. The minimum distance leads to the matched image.

Principal Component Analysis (PCA) at a glance:

Face can be represented as a matrix $I(x, y)$ having $m*n$ dimension, where $x \in m$, and $y \in n$. For the simplicity, we can also use this as a row or column vector like $I'(z)$ having $mn*1$ dimension, where $z \in mn$. In face image each face pixel is a feature and hence we have mn features, which is very hard to process due to time and space complexity of the program. Therefore in PCA we aim to find out what are those important features on behalf of which we can precede for the classification. Specifically we aim to find those features which have variance maximum [12]. In other words we find out those k features, such that $k \leq mn$ which are useful for representing others (linear combination). These k features are basically the principal components (Eigen value and Eigen vector). For finding out the best features several steps has been carried out discussed below

Steps carried out in Training Process:

Let we have P persons and every person has Q images, then we have total $\Delta = P*Q$ image, they have stored in a matrix called training database (T_db) having $mn*\Delta$ dimension.

- a) **Mean Calculation:** Mean is calculated by summing all the values (of observation) along each feature and divided by number of observation. Here mean $(\mu_z)_{mn*1}$ is

$$\text{calculated as: } (\mu_z)_{mn*1} = \frac{\sum_{z=1}^{mn} \sum_{i=1}^{\Delta} T_db(z,i)}{\Delta}$$

- b) **Zero Mean:** For aligning mean as zero, we are subtracting each face to its mean face described by (σ) .

$$(\sigma_z)_{mn*\Delta} = \sum_{z=1}^{mn} \sum_{i=1}^{\Delta} T_db(z,i) - \mu_z$$

- c) **Co-Variance Calculation:** By calculating the co-variance we aim to find out how each feature relates to other. If they are highly co-related to other, it means they can be expressed in terms of others, if negatively co-related means each one has its own importance, they cannot be expressed in the linear combination of others.

$$(\Sigma_z)_{mn*mn} = \sum_{z=1}^{mn} \sum_{y=1}^{mn} \sum_{i=1}^{\Delta} (T_db(z,i) - \mu_z) * (T_db(y,i) - \mu_y)'$$

From the above equation we can say that co-variance can be calculated as $(\text{Variance}) * (\text{Variance})^t$, resultant $mn*mn$ dimension which is hard to compute as well as process. Hence according to Turk et al [14] we used surrogate co-variance matrix here, i.e $(\Sigma)_{\Delta*\Delta} = (\sigma)^t * (\sigma)$.

- d) **Principal Component Calculation:** Principal components are the Eigen values and Eigen vectors those are computed on the basis of co-variance matrix, calculated in the previous stage.

$$\text{Then } (\sigma * \sigma^t) * \Omega = \lambda * \Omega \dots \dots \dots (1)$$

Where σ denotes Variance, λ denotes Eigen Values, and Ω denotes Eigen Vectors

Let post multiplied with σ to equation (1)

$$((\sigma * \sigma^t) * \Omega) * \sigma = (\lambda * \Omega) * \sigma$$

As we know multiplication is commutative hence we can write this as

$$(\sigma^t * \sigma) * (\sigma * \Omega) = \lambda * (\sigma * \Omega)$$

Where $(\sigma * \Omega)$ is the Eigen vector, since $(\sigma^t * \sigma)$ is having $\Delta*\Delta$ dimension, we will get the same dimension's Eigen values and Eigen vectors $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \dots, \lambda_{\Delta*\Delta})$ and $(\Omega_1, \Omega_2, \Omega_3, \Omega_4, \dots, \Omega_{\Delta*\Delta})$.

- e) **Selection of best Principal components (Eigen Vectors):** From the previous step we ended with $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \dots, \lambda_{\Delta*\Delta}$ Eigen values and $\Omega_1, \Omega_2, \Omega_3, \Omega_4, \dots, \Omega_{\Delta*\Delta}$ Eigen vector. We have to select those λ 's which have maximum values, because this shows the variance. Hence for selecting the best principal components (k), we have defined a threshold above which we have selected all principal components such that $(\lambda \geq TH)$.

- f) **Generating Eigenfaces:** Eigenfaces are generated by projecting the variance $(\sigma)_{mn*\Delta}$ over the best principal components.

$$(\text{Eigenfaces})_{mn*k} = (\sigma)_{mn*\Delta} * (\Omega_k)_{\Delta*k}$$

- g) **Generating Training Faces:** Each Training faces $(\Psi_1, \Psi_2, \Psi_3, \Psi_4, \dots, \Psi_{\Delta})$ is generates by projecting it over the variance.

$$(\Psi_i)_{k*1} = (\text{Eigenfaces})^t * \sigma(i)$$

where $i \in 1, 2, 3, 4, \dots, \Delta$, at the end we will get Training faces $\Psi_{k \times \Delta}$

Steps carried out in Testing Process:

Several steps has been carried out in Testing

1. Let we have a test face $(\alpha)_{mn \times 1}$
2. Calculate mean aligned face $(\beta)_{mn \times 1} = \sum_{z=1}^{mn} (\alpha_z - \mu_z)$
3. Calculate projected test face $(\gamma)_{k \times 1} = ((\text{Eigenfaces})^t)_{k \times mn} * (\beta)_{mn \times 1}$

Calculate the distance (Euclidean distance) between γ and Ψ , and whoever the minimum distance that will consider as the recognized identity.

As attacker tries to impersonate the identity, it will result in different classification (True or false Classification). All these results are stored in the file and if the face is live then whoever scored highest, their name will be fetched from the list. If the recognized name and the name presented by user is matched, permission will be granted to access the system else not. Here user name is treated as key which is assigned at the time of user enrolment. If the user is genuine (not an attacker), the presented name and the matched name will be same always.

III. EXPERIMENTAL SETUP

Experimental setup is divided into two units, (a) Attack generation and (b) Liveness detection and Recognition. In attack generation, we have shown all five types of spoofing attacks which we have used in this paper, and in prevention unit we have demonstrated how our system identify these attacks.

- a) **Attack Generation:** Five types of attacks listed below are used in this experiment shown in figure 3(a),(b),(c) and (d).

1. **Photo Imposter Attack:** A simple photograph of the genuine user is enough to bypass the authentication, if liveness measure has not been taken care of.
2. **Eye Imposter Attack:** If liveness measure is taken care by the user, like blinking or eye closity is detected by the system, it could be possible that by using this, attackers bypass this attack.
3. **Mouth Imposter Attack:** If only mouth movement is used in detecting the liveness then it could be possible that by using this kind of attack, attacker fools the system. This is same as Eye imposter, only difference is that attacker only removes eye region.
4. **Eye and Mouth Imposter Attack:** These kinds of attack where both the features are being forged by the attacker. This kind of attacks sometimes bypasses the liveness test.
5. **Video Imposter Attack:** Attacker can use any recorded video of the genuine user to forge the system; video is played reversely many times till system accepts this as a genuine trait.

b) **Liveness Detection and Recognition:** Recognition and liveness detection both are taking place simultaneously. Same frame is used for testing the liveness as well as recognition. The designed system first blocks the attacks in its primary phase by detecting the liveness. If the system ensures that the person is live, then only it verifies the respective identity. A user has to be first enrolled in the system with his key (i.e. name) and it will be stored with its face data. Whenever a person is recognized by the system, respective key will be drawn and stored in a “recognized identity key database”. At the time of authentication, system asks to submit the key and then throws random challenges to the user to test liveness. If the valid responses are given by the user, then only submitted key will be verified with the

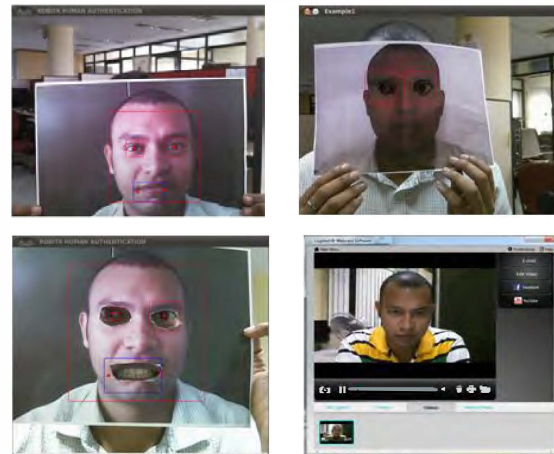



Figure 3(a): Top First, Image Imposter Attack. Figure 3(b): Top Second, Eye Imposter Attack. Figure 3(c): Bottom First, Eye and Mouth Imposter Attack. Figure 3(d): Bottom Second, Video Imposter Attack.

maximum occurred key drawn from the “recognized identity key database”. Examples of spoof detection and genuine user recognition are presented here to show the effectiveness of the approach. Here we have used 4 symbols to generate the challenges.

- (a) Eye Open  (b) Eye Close 
(c) Mouth Open  (d) Mouth Close 

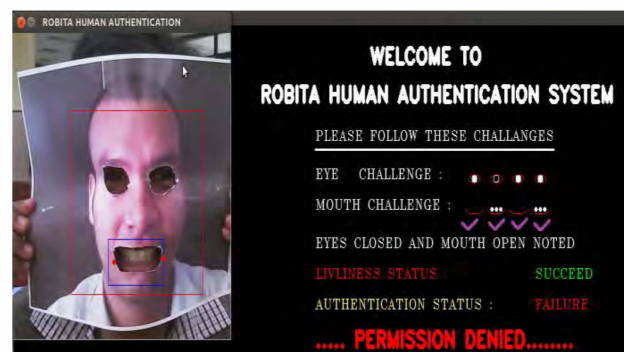


Figure 4: A Spoof Attack detected by the System

If the attacker has passed the liveness test, the system blocks it because of failure of its verification (Recognition module). To pass the liveness test, attacker has to forge both the eye and mouth region of the genuine user but it will result in major change in face and hence result in misclassification. The key presented by the attacker and that drawn from the “recognized identity key database” will not match shown in figure 4. Figure 5 shows the recognition of a genuine user, if the user follows instructions given by the system, it will pass the liveness test. The system will verify the user by matching its presented key and the key drawn from the “recognized identity key database”. Here, the person is already registered in the system, hence, the system is showing successful authentication.

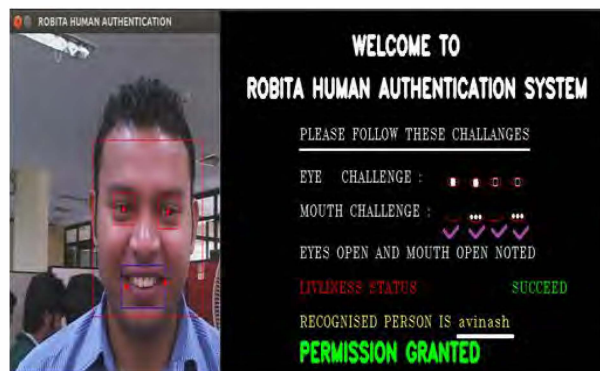


Figure 5: A Genuine User Recognized by the System

IV. RESULTS AND DISCUSSION

Results that we have got are summarized in table 1. It shows how the system behaves when attacks are placed. The report is very convincing. Only one attack bypasses the liveness test but it is further blocked by verification module. When the attacker forges both the eye and mouth, it will result in misclassification as shown in figure 4.

TABLE 1. SUMMARIZED REPORT OF SYSTEM UNDERGOING ATTACK

Attack Module	Liveness Detection Module	Verification Module
Photo Imposter Attack	No	No
Eye Imposter Attack	No	No
Mouth Imposter Attack	No	No
Eye & Mouth Imposter Attack	Yes	No
Video Imposter Attack	No	No

a) Why Eye and Mouth Removal Effect the Misclassification more than any other:

For verifying this concept we have used two assumptions, which give supports to our hypothesis.

Support 1: Face recognition literature shows that, in face most important features are of eyes, nose and mouth. They are also named as T features. From the given literature, we can say that these T feature have major contribution to those m features. Therefore absence of anyone from these can affect the classification.

Support 2: The human perception also ensures that these features of the face make the person unique from others. So absence of any of these features could lead to misclassification.

For verifying these statements, we have used face images of University of Essex, UK (face94) [13] shown in figure 6, first and second block. In our experiment we removed eye, mouth and eye-mouth both region of the face described in figure 6, Third, fourth and fifth block and calculated the efficiency of the PCA on each. There are 65 people in the database, both Male and Female. Each person is having 15 images; we have used 9 images for training and 6 images for testing (60% training and 40% testing). For removing the eye and mouth region from the face we first detected these regions in the face using the haar classifier [9], and place an ellipse filled up with black colour on these regions. The black colour having bit 0, shows that this feature is absent from the face. The reason for choosing ellipse from the available shapes (circle, rectangle, etc.) is that it provides the optimal region covered by that feature.

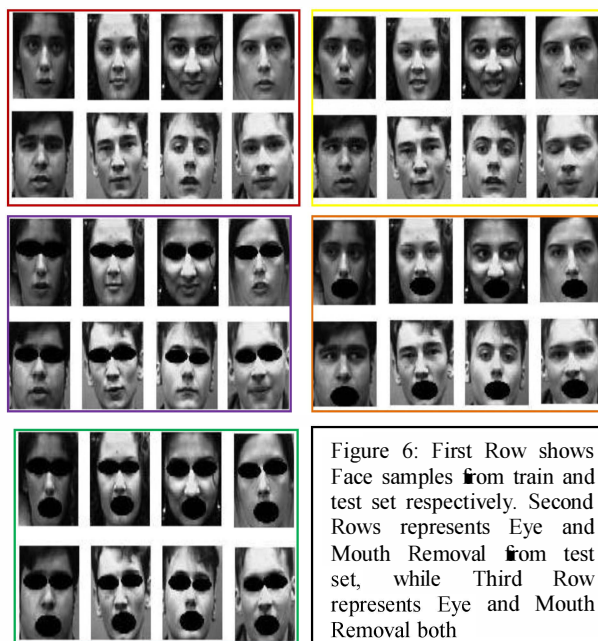


Figure 6: First Row shows Face samples from train and test set respectively. Second Rows represents Eye and Mouth Removal from test set, while Third Row represents Eye and Mouth Removal both

Misclassifications of these experiments are summarized in figure 7, which shows the misclassification ratio when attacks are performed. Misclassification shows that how many times system generated false key.

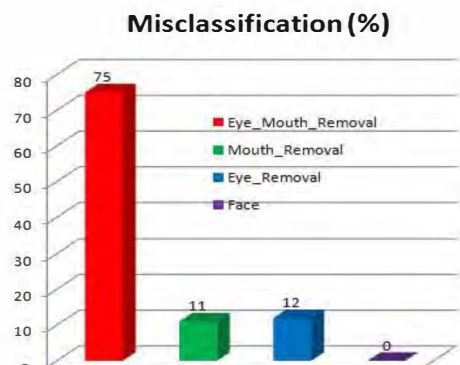


Figure 7: Misclassification rate of various imposter Attacks

V. CONCLUSIONS

Liveness detection is necessary because it ensures whether the person is physically present or not. As no two persons can have the same biometric traits, system will identify that a genuine user is presenting their biometrics. Experimentally it is proven that, to bypass the liveness test, attacker has to remove eye and mouth region of the photograph and has to place his eye and mouth instead. These attempts bring about drastic changes in the structure of genuine user's face, hence resulting in misclassification. Eye and mouth movements are identified by using the Haar classifier, Eye openness and closeness in a definite time interval shows the eye liveness while the teeth HSV value estimation shows the movement in mouth region. HSV value of the teeth is calculated by estimating the mouth's Region of Interest (ROI). We have used a challenge and response method here to test the liveness while Principal Component Analysis (PCA) has been used effectively for recognizing the person. PCA is a good candidate for dimensionality reduction; therefore we used this to prove our hypothesis. We have tested our approach on 40 people each is having 20 images. The system has shown a good accuracy ratio and has successfully identified all the trained persons.

ACKNOWLEDGMENT

We would like to thank all the students, research scholars, and project associates of Robotics and Artificial Intelligence lab of Indian Institute of Information Technology Allahabad for donating their face biometric data, and their cooperation during several time testing over this project.

REFERENCES

[1] W. Zhao, R. Chellappa, A. Rosenfeld, P.J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys, 2003, pp. 399-458

[2] A.F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey", Pattern Recognition Letters, vol.28, issue 15, pp.1885-1906, Oct 2007

[3] Westeyn, T., Pesti, P., Park, K.-H. and Starner, T. (2005) Biometric identification using song-based blink patterns. *Proc. HCI Int'l '05*. Mahwah, NJ: Lawrence Erlbaum.

[4] H.-K. Jee, S.-U. Jung, and J.-H. Yoo, "Liveness detection for embedded face recognition system," in *World Academy of Science, Engineering and Technology Vol. 18*, 2006.

[5] Peixoto, B.; Michelassi, C.; Rocha, A., "Face liveness detection under bad illumination conditions," *Image Processing (ICIP), 2011 18th IEEE International Conference on*, vol., no., pp.3557-3560, 11-14 Sept. 2011.

[6] S. A. C. Schuckers, "Spoofing and Anti-Spoofing Measures," vol. 7, Information Security Tech. Rep., 2002.

[7] K. Nixon, V. Aimale, and R. Rowe, "Spoof detection schemes," in Handbook of Biometrics. Springer US, 2008.

[8] Biggio, B.; Akhtar, Z.; Fumera, G.; Marcialis, G.L.; Roli, F.; , "Security evaluation of biometric authentication systems under real spoofing attacks," *Biometrics, IET*, vol.1, no.1, pp.11-24, March 2012.

[9] Wilson, P. I. and Fernandez, J. (2006). Facial feature detection using haar classifiers. *Journal of Computing Sciences in Colleges*, 21:127-133.

[10] Jain, A., Kumar, A.: Biometrics of Next Generation: an Overview, Second Generation Biometrics. Springer, Heidelberg (2010).

[11] JOLLIFFE, I. T., 2002. Principal Component Analysis. Second ed. Springer Series in Statistics. New York: Springer-Verlag New York.

[12] HOTELLING, Harold, 1933. Analysis of a Complex of Statistical Variables into Principal Components, "Journal of Educational Psychology", 24(6 & 7), 417-441 & 498-520.

[13] Face Recognition Data, University of Essex, UK, Face 94, http://cswwww.essex.ac.uk/mv/all_faces/faces94.html.

[14] Turk, M.A.; Pentland, A.P.; "Face recognition using eigenfaces," *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91. IEEE Computer Society Conference on*, vol., no., pp.586-591, 3-6 Jun 1991.

[15] Gang Pan; Lin Sun; Zhaohui Wu; Shihong Lao; , "Eyeblick-based Anti-Spoofing in Face Recognition from a Generic Webcam," *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, vol., no., pp.1-8, 14-21 Oct. 2007.

[16] M. Kirby, L. Sirovich, Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 1, January 1990, pp. 103-108

[17] Xiaoyang Tan , Yi Li , Jun Liu , Lin Jiang, Face liveness detection from a single image with sparse low rank bilinear discriminative model, Proceedings of the 11th European conference on Computer vision: Part VI, September 05-11, 2010

[18] Girija Chetty , Michael Wagner, Audio-visual multimodal fusion for biometric person authentication and liveness verification, Proceedings of the 2005 NICTA-HCSNet Multimodal User Interaction Workshop, p.17-24, September 13, 2005, Sydney, Australia.

[19] Avinash Kumar Singh, G.C Nandi, "Face recognition using facial symmetry", proceedings of the 2nd ACM International Conference on Computational Science, Engineering and Information Technology (CCSEIT-2012), Coimbatore, pp. 550-554, October, 2012.