An Intelligent Security System for a Bank Based on Amplitude of Blink Gesture Recognition

[1]Raed Awadh Bakunah, [2,3]Dr. Saeed Mohammed Baneamoon

[1]MSc Scholar, [2,3] Associate Professor

[1]Department of Information Technology, [2]Department of Computer Engineering, [3]Department of Computer Science

[1]Faculty of Engineering and Information Technology, Al-Rayan University, Mukalla, Yemen,

[2] College of Engineering & Petroleum, Hadhramout University, Mukalla, Yemen,

[3] College of Computers & Information Technology, Hadhramout University, Mukalla, Yemen.

Abstract : This paper proposes a human-security system interaction based on amplitude gesture recognition for intelligent bank security. In order to recognizing the amplitude gesture a different processes are implemented in the amplitude gesture recognition analyzing, these processes are face and facial landmark detection, Eye aspect ratio (EAR) and amplitude and duration of blink. The using web camera and successfully implemented a prototype proposed system is tested for the gesture of the amplitude blink with tow eyes successful. The proposed system is evaluated by a number of experiments for security in bank from robbery which is having a detection accuracy rate more efficient compared with other related approaches.

IndexTerms - **Human–computer interaction, Security system, Gesture recognition and Facial Landmarks.**

I. INTRODUCTION

Bank is one of the most important safety units, which must represents a top level of security. The bank become the target of criminals, because it has increased number of branches, which have expensive jewelry, important documents and a cash. In last years, there are increased numbers of robberies and crimes of banks which, occurred in cities during daylight hours when the employees were subjected to force, violence or gunpoint [1]. So there is strong need to develop technique to enable the employee to activate the alarm system, and help to inform the local security center or nearest police station to rapid intervention in order to abort the crime in timely manner, easy capture of suspects and keeps expensive contents, life of all employees and customers in safe without the knowledge, feeling and attention by suspects, attackers and accomplices.

Therefore, this paper proposes a new human-security interaction system for intelligent bank security, which based on amplitude of blink gesture detection technique in order to recognize special amplitude gesture of the employee.

II. RELATED WORK

The crimes of banks are increasing day by day. The developing of security systems have become the major concern to the entire world.

This section describes different technologies and techniques that have been used in security banks' systems. Some studies based only on electronic and embedded system technology that connected to different sensors as in [2] and [3], and other studies based on computer vision and image processing technologies that are used deferent methods and techniques in its systems. So some of this systems used human behavior recognition technique which focus on the behavior of attackers and robberies as in [4]. Others are depending on image processing technique to detect the changing in background as in [5] and [6]. Also some systems rely on image recognition and image matching processing techniques, which based on the comparison between two images, to see if there are differences in images as in [7]. In addition, there are some systems that used image matching processing technique integrated with levels of authentication and identification techniques such as RFID, biometric, face recognition, voice recognition and pattern recognition as in [8], [9], [10]. However most systems of this studies rely on communication technology to inform about crimes, and may be utilized in banks and financial institution for giving protection to highly-priced possessions in locker rooms and ATM machines, but the techniques are not depending on employees to detect the robberies, when they are feeling and watching the start of crime specially if the bank's system fall to recognizing on the crime. In this state the employee stay arms folded so he cannot do anything especially at gunpoint.

The studies in [11] and [12] are closed to our work, but they based on hand gesture technique which is easily to identify and observe by suspicious persons and accomplices that leads to exposing employees, customers and money to risk and escape the criminals in timely manner. Moreover this technique is ineffective in state when the hands are tied.

However from all previous studies, it is clear that there is an intense interest towards direction of robust and safety security systems with good performance..

III. BACKGROUND

3.1 Computer Vision and Image processing

An image is a two-dimensional matrix of pixels. Every image consists of a set of pixels. Pixels are the raw building blocks of an image. A pixel as the "color" or the "intensity" of light that appears in a given place in image. A color model (channel) is a way of describing colors. These are usually represented as tuples of numbers, typically as three or four color components. An RGB model is used most often, whereby each color is encoded as a composition of a red, green and blue channel [13].

Image processing deals with extracting low-level information from images by treating them as signals. In fact, image processing is often regarding as a subfield of signal processing and uses a large amount of the same techniques .This kind of processing can be used to detect edges or perform blurring [13] [14].

Computer Vision is a subfield of computer science that deals with techniques for acquiring, processing, analyzing, and understanding digital images in order to extract numerical and symbolic information, in order to make decisions. This usually involves finding more high-level information than image processing, although image processing techniques are often used before applying computer vision techniques [15] [16].

3.2 Gesture

Many researchers had tried to define gestures but their actual meaning is still arbitrary. Gesture is defined as the motion of the body that is intended to communicate with other agents [17]. For a successful communication, a sender and a receiver must have the same set of information for a particular gesture. As per the context of the project, gesture is defined as an expressive movements of body parts, which have a particular messages, to be communicated precisely between a sender and a receiver.

A gesture is scientifically categorized into two distinctive categories: dynamic and static. A dynamic gesture is intended to change over a period of time whereas a static gesture is observed at the spurt of time. A waving hand means goodbye is an example of dynamic gesture and the stop sign is an example of static gesture [18] [19]. The possible location for origin of the human gestures or body gestures can be from any part of our body, but the most of the gesture-based studies that are carried out involve the gestures that have originated from the hand , face, head, eyes, eyelids and tongue [20] [21] [22] [19] [23] [24], because most normal people transfer the messages through this locations. To understand a full message, it is necessary to interpret all the static and dynamic gestures over a period of time, this process is called gesture recognition.

Gesture recognition is the process of recognizing and interpreting a stream continuous sequential gestures from the given set of input data. The main entire point of the gesture recognition, is to teach a machine to learn different gestures made by humans, and to interpret the same, which means to make the sensible conversations between a human and a machines possible.

With the rapid development of computer vision, image processing and video intelligent technologies, it's now possible to perceive gestures through vision techniques and integrate human gestures recognition techniques with video surveillance systems to achieve intelligent security systems. This integration is a great interest for recent computer vision researchers. It is helped in facilitate and better human control over human monitoring systems as in computer commands [19] [24], controlling machines as wheelchair as in [25] [22] and activity recognition [26], [27], [28].

However, a shift towards a user friendly environment has driven us to revisit to focus this area. The abstract from previous context that is to make computer understand human language and develop a user friendly environments as human computer interaction (HCI), which is a technology which lets computer and machine hear, see, and feel. At present, the field of human-computer interaction has entered the stage of multimodal interaction which making a computer understand the speech, facial expressions and human gestures as in [29].

IV. PROPOSED METHOD

The proposed method rely on the following important computer vision techniques to detect the amplitude blink of the gesture: Facial landmark, Eye aspect ratio (EAR) and Duration @ amplitude of blink. This method is fast, efficient, and easy to implement by using python with OpenCV and dlib libraries. The methodology of the proposed system include Face Detection, Facial Landmarks Detection and classification for amplitude and duration of blink.

The block diagram of complete system is shown in Figure1. Before detecting the facial landmark points of eyes, firstly localizing the face in stream video. Then detecting the amplitude blink gesture by measuring the current value of ear, and compares this value with predefined thresholds. After that measuring the duration and an amplitude of the current blink and compare them with predefined thresholds, in order to classifying and recognize the amplitude gesture.

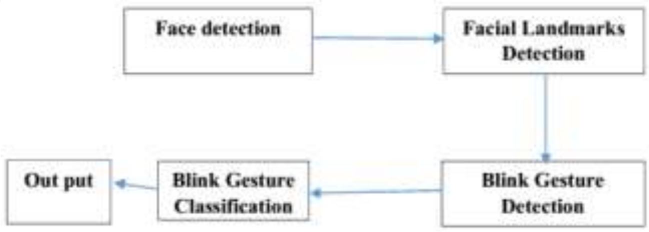

Figure 1. The proposed methodology for amplitude blink gesture detection.

4.1 Face and Facial Landmarks detection

The face plays an important role in visual communication. By looking at the face, human can automatically extract many nonverbal messages, such as humans' identity, intent, and emotion. In computer vision, face landmark consist of two tasks: face detection and facial landmarks [21]. The facial information gained through the facial landmarks locations (key points) can provide important information for human computer interaction, entertainment, security surveillance, and medical applications. These points are describing the unique location of a facial component. The facial landmarks detection is the process of localizing the key facial points on a face, including the eyes, eyebrows, nose, mouth, and jawline. Also the facial landmark detection algorithms aim to automatically identify the locations of the facial key landmark points on facial images or videos. It is usually assumed that face is already detected and given for most of the existing facial landmark detection algorithms. The detected face would provide the initial guess of the face location and face scale [30]. Finally, the idea in this paper is to work on 68 for all facial landmark points as in Figure 2.

Figure 2. Visual representation of 68 facial landmark points

However, one of aims in this paper is in detection of eyes components on videos by placing 6 (six) landmark points on each eye as shown in Figure 3.

Figure 3. Visual representation of 6 landmark points coordinates of each eye

4.2 Eye Aspect Ratio (EAR) Detection

The eye aspect ratio is simple calculation based on the ratio of distances between facial landmarks of the eyes. This technique was proposed by Soukupova and Cech's [31], [32]. It gives us a singular value, relating the distances between the vertical eye landmark points (p2, p3, p6, p5) to the distances between the horizontal landmark points (p1, p4) and it is computed by Eq. 1. The ear is mostly constant when eye is open and getting close to zero while closing an eye as shown in Figure 4.

$$EAR = \frac{\|p2 - p6\| + \|p3 - p5\|}{2\|p1 - p4\|} \qquad (1)$$

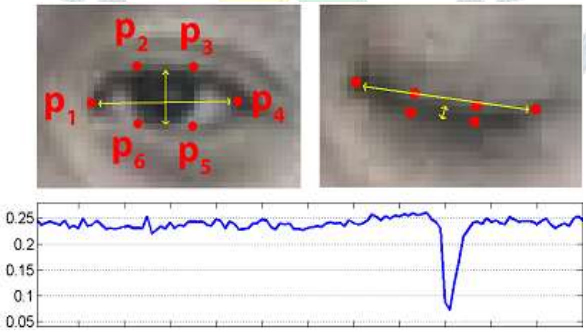

Figure 4. Open and closed eye and its landmark points and its ear value while open and closed state

4.3 Duration and Amplitude detection

The duration of blink is the duration of complete blink in a given time interval, another interesting blink parameters is the amplitude of blink, which describes the degree openness of eyelids of eye, these parameters differ between the people. Most current works based on duration parameter in its blink gestures interaction systems. Duration gestures which the changes are in duration time of blink, that mean the eyelids of eye stay closed in predefined period of time [33], [31].

In this paper is proposed new voluntary gestures based on amplitude parameter of the blink for the proposed security interaction system, which never used before as shown in Figure 5. The openness of eyelids are changing differently from normal, it increase in predefined period of time and decease in predefined period of time. In this type there is no need to completely close the eye.

The proposed gestures had specific and sufficient amplitude thresholds values, in order to avoid any trigger of a system by the spontaneous eye blink of human. In this type of gesture, no need to completely close the eyes.

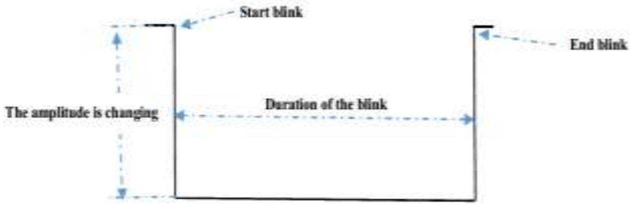

Figure 5. The proposed amplitude blink gesture

4.4 Experimental Proposed Method

In the system, the first section is localize the face by apply a pre-trained HOG + Linear SVM object detector algorithm in a given frame of a video stream, which is implemented in dlip library as shown in Figure 6.

Figure 6. Localized face

Then facial landmark detector is used which included in the dlib library to perform facial landmark detection of 68 key facial coordinates in order to localize the eyes landmarks as shown in Figure 6. This detector is an implementation of One Millisecond Face Alignment with an Ensemble of Regression Trees in [34].

Figure 7. Facial landmark points

Once we have the facial landmarks for both eyes as shown in Figure 8, we use Soukupova and Cech's method to measure EAR for each eye, which is ratio of distances between the six facial landmark points [31]. The ear value for each consecutive frame of both eyes is calculated and then the mean is taken and a thresholds for EAR are set in the code.

Figure 8. Landmark points of eyes

The second section is to determine when the blink is considered to be a gesture, to do that we first set tow threshold values of ear that must be set in code, higher one for open eye (HT) and lower one for closed eye (LT). After that we measure the real detected ear value and compare it with LT and HT threshold values, if the comparison is true, then we determine the amplitude and duration of the detected blink and compare it with predefined thresholds.

The duration is the number of sequence frames when eyes are in closed state or when value of ear = < LT value, so TFseq predefined threshold is the number of sequence frames must be set in the code. The amplitude of blink is openness of the eyes, and the amplitude gesture is the change in ear value from low to high or high to low in TFseq predefined period of time (number of frames) passing through LT and HT threshold values. Figure 9 shows the signal of complete amplitude gesture with LT, HT, TFseq and SLT thresholds, while SLT is the shafted low threshold when there is no need to close eyes completely.

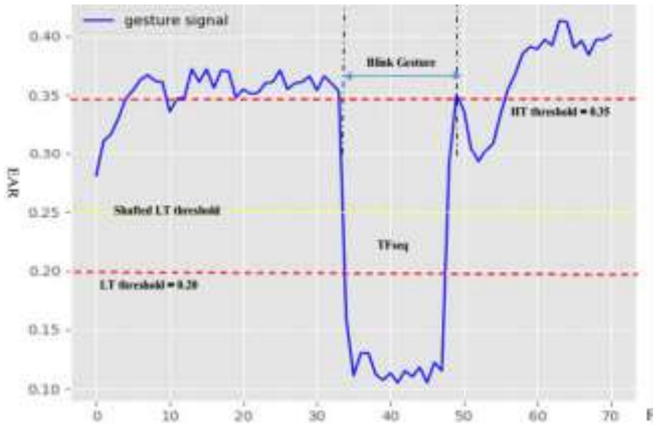

Figure 9. Signal of amplitude gesture and its thresholds

Once the comparison is true, the alarm is activated the system and sends SMS warning messages. This section is performed by a code on python language and the flowchart of the complete system is shown in Figure 10.

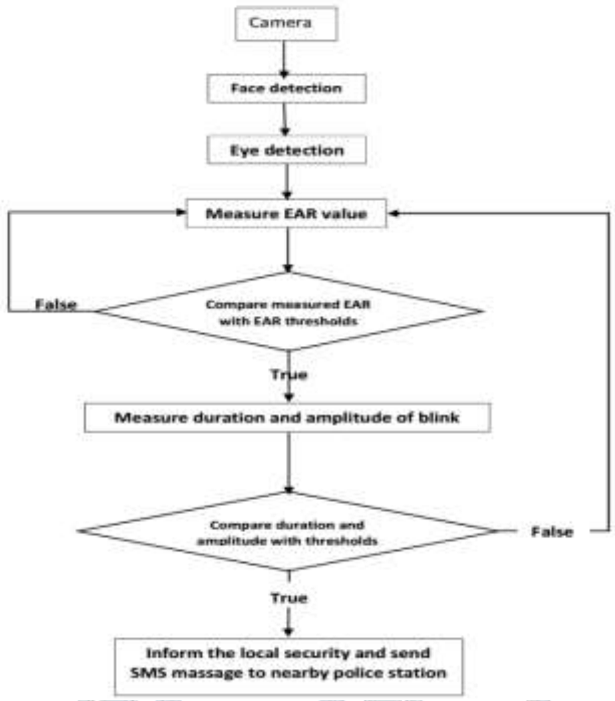

Figure 10. The experimental flowchat of complete system.

V. EXPERIMENTAL RESULT

The proposed system for Amplitude blink gesture detection was tested using Intel Core 2 Duo CPU at 2.13 GHz processor 4G RAM on the 30 fps sequences frames from the USB Logitech 270 Pro webcam and Jewa web cameras as shown in Figure 11. The size of the input images sequences were equal to 480, 640 pixels. Testing of the system took place in a room illuminated by 2 fluorescent lamps and daylight from tow windows.


Figure 11. Core 2 Duo laptop, HD web camera and medium resolution web camera

We did initial training, so we recorded videos for four participants who are perform the predefined gesture. The USB web camera was fixed at distance about 70 cm away from the participant's face. The participant must keep the distance to the camera as stable as possible. Each participant was asked to blink some times (normal blinks and amplitude gesture blinks, alternately). Finally we did experiments in Python, OpenCV, and dlib in video streams in orther to abtain results as in sections: A, B, C and D.

A. Participant vs. EAR Experiment Results

The EAR values for 4 different people were recorded and the ear values for completely closed, normal open and maximum open of the eyes were measured. For accuracy the ear was calculated for several times and the mean value was recorded. Table 1shows the mean values. Using the values a graph was plotted as shown in Figure 12 and the thresholds of LT and HT were computed as in Eq. 2 and Eq. 3.

Table 1 The mean values of ear for complete closed, normal open and max openness of the eyes.

Participants	EAR average for eyes closed	EAR for normal eyes open	EAR Values of max eye openness
Participant 1 Raed	0.104	0.345	0.425
Participant 2 Aya	0.092	0.323	0.389
Participant 3 Saeed	0.072	0.313	0.382
Participant 4 Sami	0.103	0.294	0.358

$$LT = \frac{\text{Big ear value in closed state + small ear in normal open state}}{2} \quad (2)$$

$$HT = \frac{\text{Big ear in normal open state + small ear in max openness state}}{2} \quad (3)$$

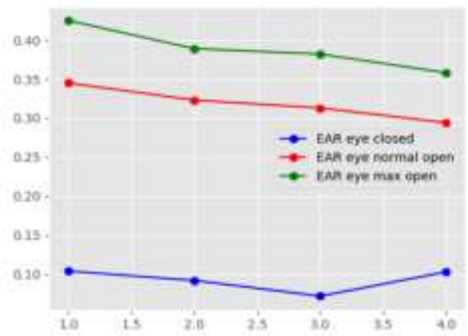

Figure 12. EAR of 4 participants for open, closed and max openness of the eyes.

B. Gestures vs. Conditions Experiment Results

Experiments were done with one participant at different luminous intensities, head orientation, wear glasses and occlusions. We noticed that the amplitude gesture stay detected at different luminous levels even in low lights as shown in Figure 13, in different views of face orientation as shown in Figure 14, in wearing glasses, in deferent expression and in some Occlusions as shown in Figure 15 (a) and (b).

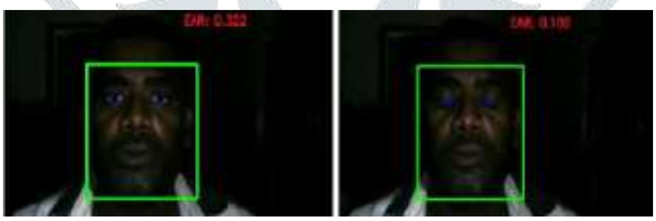

Figure 13. Gestures stays detected even in low lights.

Figure 14. Camera put it in different directions up, down, side views

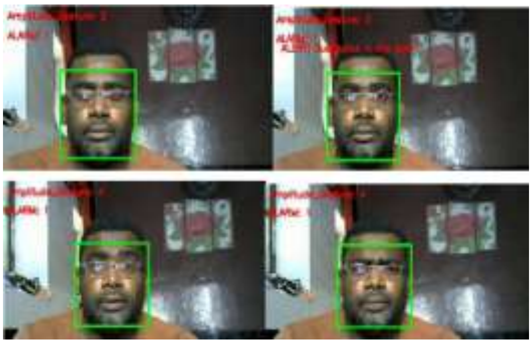

Figure 15 (a) Amplitude gesture in glasses state.

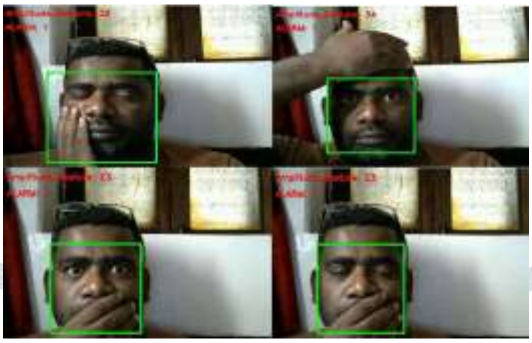

Figure 15 (b): Gesture was detected in deferent Occlusions

C. Real Time Capabilities Results

We first analyzed whether the system was able to process the video and recognize gestures in real-time. To do this, we measured the processing time for each frame and computed the mean frame rate.

As can be seen from Table 2, the time for individual frame without any detection processing of face, landmarks and gesture is between 110ms, 142ms, and the mean frame rate was 5.9 Fps. If the system performs face and landmarks detection, the processing time up between 125ms, 200ms. So the mean frame rate drop to 5.3 Fps. While the baseline corresponds to displaying the video on the screen without any processing was closed to frame rate of the camera around 27 fps.

Table 2 Frame rate and processing time with / without face and landmarks recognition and also the baseline corresponds to displaying the video on the screen without any processing.

	Baseline	Without any face appearance	Detection
Mean frame rate	**27. 3 fps**	**5.9 fps**	**5.3 fps**
Processing time	**0 – 37.2ms**	**110-142 ms**	**125–200 ms**

In Table 3 the time of complete gesture for different number of frames was measured to find the most appropriate number of frames (TFseq) to set in the code. Using values in (Table III) a graph was plotted as shown in Figure 16. And also Figure 17 shows some samples of complete gesture.

Table 3 Gesture processing time for number of frames.

TFseq to set in code	Number of Frames for gesture	Processing time of Gesture
4	**6**	**1.257s**
5	**7**	**1.467**
6	**8**	**1.668s**
8	**9**	**1.879s**
12	**16**	**3.343s**
***13**	***15**	***3.137s**
***14**	***16**	***3.353s**
15	**17**	**3.553s**
16	**18**	**3.718s**
18	**20**	**4.186s**

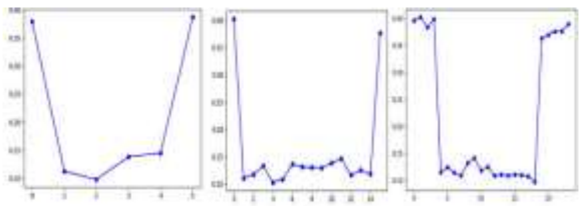

Figure 16. Samples of complete signal of amplitude gesture.

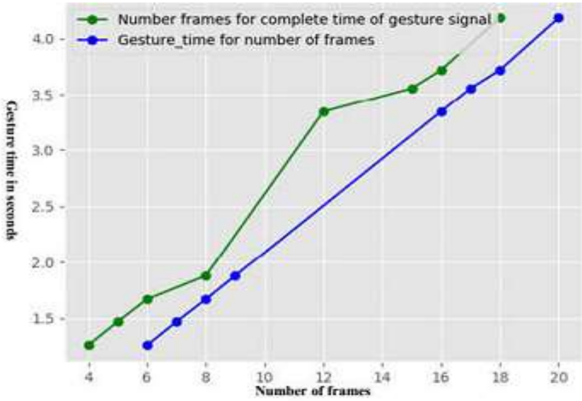

Figure 17. Gesture time for number of frames and number of frames for complete time of gesture signal.

D. Recognition Performance Experiment Results

The detection of the Amplitude blink gesture was done in real time with the camera speed of 30 fps. Two kinds of errors were identified: false detection (the system detected gesture when it was not present) and missed gesture (a present gesture that was not detected by the system). The possible decisions distribution of the gesture detector output is presented as shown in Table 4.

Table 4 Possible decisions of gesture output

Possible decision		Actual state	
		Gesture present	Gesture not present
Result of detection	Gesture detected	TP	FP
	No gesture detected	FN	TN

The correctly detected of the gesture are denoted as True Positives (TP), the false detections are denoted as False Positives (FP) and the missed gestures are denoted as False Negatives (FN). Based on these parameters, measure of the system performance was introduced the accuracy as shown in Eq. 4 [24].

$$\textbf{Accuracy} = \frac{\textbf{TP}}{\textbf{TP + FP + FN}} \quad (4)$$

This measure was used to assess the system robustness in detecting the Amplitude blink gestures as well as for the overall system performance. The tests of the system and low and good resolution of cameras and others conditions were also performed and the results are presented as in Tables 5, 6 and 7).

Table 5 Accuracy of the Amplitude blink gesture recognition system averaged over all four participants.

videos	GT	DAG	FN	FP	ACC
Participant1	19	18	1		94.7%
Participant2	3	3			100%
Participant3	6	5	1		83.3%
Participant4	10	8	2		80%
Averaged Accuracy					89.5

Table 6 The distribution of the system performance and accuracy, in deferent conditions states: direction views of camera, glasses and low lights.

Deferent conditions videos	GT	DAG	FN	FP	ACC
Down	10	7	3	1	64%
Side	10	8	2		80%
Middle view	40	39	1		97.5%
Glasses state	20	19	1		95%
Middle and low light	20	20			100%

Table 7 Low and high Accuracy of the system.

Measure	Accuracy
Good resolution	96.25%
Low resolution	89.5%

When GT is Total Amplitude Gesture (Grown True Gesture), DAG is detected Amplitude gesture and ACC is an accuracy. Table 8 shows a comparison of proposed system with Gupta and Kumar systems.

Table 8 The comparison of proposed system with Gupta and Kumar works.

Works	Used technique	Observe by suspicious and accomplices	Accuracy
Gupta [11]	**Hand gesture recognition**	**easily**	**95.7**
Kumar [12]	**Hand gesture recognition**	**easily**	**94.5**
Proposed system	**Amplitude blink gesture recognition**	**difficult**	**96.5**

VI. CONCLUSION

In this paper security interaction system based on the amplitude blink gesture recognition technique is proposed. The employee is pre-trained about how and what gesture to make in case of robbery in bank before joining at the bank. This proposed system combines facial landmark, eye aspect ratio and amplitude and duration computer vision techniques to recognize the amplitude gestures. The implemented real time amplitude- blink gesture recognition security system has an accuracy of more than 96% which, is better as compared with the existing results of related works in [11] and [12].

.

REFERENCES

[1] Royanews, "royanews," [Online]. Available: https://royanews.tv/news/171426.

[2] K. Archana, P. B. Reddy and A. Govardhan, "To enhance the security for ATM with the help of sensor and controllers," *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai, pp. 1012-1015, 2017.

[3] T. M. F. Elshafiey, "Design and Implementation of a Museum and Banks Security System Using Antenna as IR Proximity Sensor and PSoC Technology," *2011 IEEE Symposium on Wireless Technology and Applications (ISWTA)*, Langkawi, pp. 156-161, 2011.

[4] B. Georis, M. Maziere, F. Bremond and M. Thonnat, "A video interpretation platform applied to bank agency monitoring," *IEEE Intelligent Distributed Surveilliance Systems*, London, UK, pp. 46-50, 2004.

[5] Y. Wu, "Research on bank intelligent video image processing and monitoring control system based on OpenCV," *2009 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication*, Hong Kong, pp. 211-214, 2009.

[6] N. Khera and A. Verma, "Development of an intelligent system for bank security," *2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence)*, Noida, pp. 319-322, 2014.

[7] H. Yi, J. Liu and X. Wang, "Automatic Alarm System for Self-Service Bank Based on Image Comparation," *2011 International Symposium on Computer Science and Society*, Kota Kinabalu, pp. 48-50, 2011.

[8] A. Verma, "A Multi Layer Bank Security System," *2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE)*, Chennai, pp. 914-917, 2013.

[9] B. Joshua, J. Grace and A. M. George, "Open CV Pattern Based Smart Bank Security System With Theft & Identification Using Android," *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, Bangalore, pp. 254-257, 2016.

[10] P. Kumari, "High protection bank locker security system using live image and voice authentication," *International Journal of Scientific Research and Review / UGC Journal No.: 64650*, vol. Volume 07, no. Issue 03, pp. 2616-2619, March 2019.

[11] A. Gupta, Y. Kumar and S. Malhotra, "Banking security system using hand gesture recognition," *2015 International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE)*, Noida, pp. 243-246, 2015.

[12] S. Kumar, G. Yadav, H. P. Singh, S. Malhotra and A. Gupta, "SSVM Classifier and Hand Gesture Based Sign Language Recognition," *2018 International Conference on Intelligent Circuits and Systems (ICICS)*, Phagwara, pp. 456-460, 2018.

[13] R. Adrian, "Practical Python and OpenCV", 3rd Edition, pyimagesearch.com, 2016.

[14] Wikipedia, "Image processing," [Online]. Available: https://en.wikipedia.org/wiki/Digital_image_processing.

[15] Wikipedia, "computer vision," 2017. [Online]. Available: https://en.wikipedia.org/wiki/Computer_vision.

[16] D. A. Rosebrock, Deep Learning for Computer Vision with Python, 1st Edition, PyImageSearch.com, 2017.

[17] S. MEENA, "A Study on Hand Gesture Recognition Technique," National Institute Of Technology, Roukela, INDIA, 2011.

[18] V. A. Shanthakumar, C. Peng, J. Hansberger, L. Cao, S.C. Meacham and V.R. Blakely, "Design and evaluation of a hand gesture recognition approach for real-time interactions, " *Multimed Tools Appl* **79,** pp.17707–17730, Feb. 2020.

[19] S. Hussain, R. Saxena, X. Han, J. A. Khan and H. Shin, "Hand gesture recognition using deep learning," *2017 International SoC Design Conference (ISOCC)*, Seoul, pp. 48-49, 2017.

[20] K. Wang, R. Zhao and Q. Ji, "Human Computer Interaction with Head Pose, Eye Gaze and Body Gestures," *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, Xi'an, pp. 789-789, 2018.

[21] C. Pradhan, D. Banerjee, N. Nandy and U. Biswas, "Generating Digital Signature using Facial Landmlark Detection," *2019 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, pp. 0180-0184, 2019.

[22] A. Gupta, G. Chitranshi, V. Sharma and H. Singh, "Controlling devices using facial expression identification and recognition for assisting physically disabled," *2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*, Noida, 2017, pp. 574-576, 2017.

[23] V. Vaitukaitis and A. Bulling, "Eye Gesture Recognition on Portable Devices," *UbiComp '12: Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 711-714, Sept. 2012.

[24] A. K. ´lak, "Eye-blink detection system for human–computer interaction," *Universal Access in the Information Society 11(4) Spirnger*, pp. 1-11, 2011.

[25] H M Ravindu T Bandara,K S Priyanayana,A G Buddhika P Jayasekara,D P Chandima,R A R C Gopura, " An Intelligent Gesture Classification Model for Domestic Wheelchair Navigation with Gesture Variance Compensation," *Applied Bionics and Biomechanics, Hindawi*, pp. 1-11, Jan. 2020.

[26] A. Karakostas, I. Lazarou, G. Meditskos, T. G. Stavropoulos, I. Kompatsiaris and M. Tsolaki, "Sensor-based in-home monitoring of people with dementia using remote web technologies," *2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL)*, Thessaloniki, pp. 353-357,, 2015.

[27] H. W. Guesgen, "Human Behavior Recognition Technologies: Intelligent Applications for Monitoring and Security", New Zealand: Massey University, 2013.

[28] C. Fauzi, S. Sulistyo and Widyawan, "A survey of group activity recognition in smart building," *2018 International Conference on Signals and Systems (ICSigSys)*, *IEEE*, Bali, pp. 13-19, 2018.

[29] R. Zhao, K. Wang, R. Divekar, R., Rouhani, H. Su, Q. Ji, "An immersive system with multi-modal human-computer interaction," *Automatic Face & Gesture Recognition (FG 2018) 2018 13th IEEE International Conference on.*, pp. 517-524, 2018.

[30] Y. Wu, "Facial Landmark Detection: A Literature Survey," *IJCV, Springer*, pp. 1-28, 2018.

[31] T. Soukupov and J. Cech, " Real-Time Eye Blink Detection using Facial Landmarks," *21st Computer Vision Winter WorkshopLukaˇCehovin, Rok Mandeljc, VitomirˇStruc (eds.)Rimske Toplice, Slovenia, February* 2016.

[32] T. Soukupov´a, " Eye-Blink Detection Using Facial Landmarks," Center for Machine Perception, Technical University Technick´a 2, 16627 Prague 6, Czech Republic, Prague, May 2016.

[33] Z. A. Haq and Z. Hasan, "Eye-blink rate detection for fatigue determination", *1st India International Conference on Information Processing (IICIP)*, pp. 1-5, 2016..

[34] Kazemi, "One Millisecond Face Alignment with an Ensemble of Regression Trees," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1867-1874, 2014..

[35] S. Hussain, "Hand Gesture Recognition Using Deep Learning," *International SoC Design Conference (ISOCC), IEEE*, pp. 48-49, Nov. 2017.