International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 5, Sept-Oct-2020, ISSN (Online): 2395-566X

A Review on Audio Video Based Person Authentication System Using Multi SVM

M.Tech.Scholar Ranjana Patel ,Krishna Kumar Vishwakarma
Dept. of Electronics & Communication Engineering
ranjanapatel332@gmail.com , kri_k_2006@yahoo.co.in
Rewa Institute of Technology
Rewa M.P.,India

***Abstract*- Object detection and tracking is often the first step in applications such as video surveillance. We propose a general moving object detection and tracking based on vision system using image difference algorithm. Then the speech of the person is recognized to get the feedback from the corresponding person. This process focuses on detection of human in a scene and then speech signal processing was done..We proposed new technique for human identification using fusion of both face and speech which can substantially improve the rate of recognition as compared to the single biometric identification for security system development. Our system using Viola Jones Algorithm for face detection. The proposed system uses Local Binary Pattern (LBP) as feature extraction techniques which calculate the local features. The Mel-Frequency cestrum coefficients feature extraction techniques are used for speech recognition in our project. The Extracted features given as input multi-SVM classifier to used to recognize the person and then display the result. This new system can be applied in various different fields such as identity verification and other potential commercial applications.**

***Keywords*-Multi SVM, LBP, Audio, Video, Feature Extraction.**

I.INTRODUCTION

Automatic person authentication has received an increasing interest for security applications in the last two decades. The objective of a person authentication system is to accept or reject the identity claim of a person using one or more physiological or behavioral characteristics associated with the person. Person authentication systems make use of one or more biometrics such as speech, face, fingerprint, signature, iris and hand geometry to accept or reject the identity claim of an individual.

The reason for the increased attention is that speech and face biometrics provides convenient and natural form of input. Government agencies are investing a considerable amount of resources into improving security systems as a result of recent terrorist events that dangerously exposed flaws and weaknesses in today's safety mechanisms. Badge or password-based authentication procedures are too easy to hack. Biometrics represents a valid alternative but it suffers drawbacks as well. Iris scanning, for example, is very reliable but too intrusive; fingerprints are socially accepted, but not applicable to non-consentient people.

On the other hand, person authentication represents a good compromise between what's socially acceptable and what's reliable, even when operating under controlled conditions. Human beings can remember hundreds or even thousands of faces in their whole 1 life and can easily identify a familiar face in different perspective variations, such as illumination variations, age variations, and pose variations. There is an increasing interest in biometric person authentication for commercial, security, surveillance and other applications, but authentication based on only one modality is unlikely to achieve acceptable performance for practical deployment. A potential way to overcome this is to combine information from more than one modality, and several important studies have confirmed this potential [5][6]. Person authentication systems can be used to prevent unauthorized access to automatic teller machines (ATMs), desktop PCs, workstations and in applications such as electronic commerce.

Currently, low cost methods such as passwords, personal identification numbers and magnetic cards are widely used for these applications. Traditionally, commercial applications have used personal identification numbers (PIN) and passwords, government applications have used identification cards, badges and forensic applications have relied on human experts to match biometric features. Person recognition system using speech, face and mouth modalities can be deployed in all the above application areas. In this thesis, we consider the particular situation of audio-video person authentication based on three sources of information: an audio signal (speech) and a video signal (face and mouth)

1. Face Recognition- by Humans Face recognition can be categorized into face authentication, face identification

and watch list [10]. The objectives of the three categories are listed below:

2. Face authentication: An individual who desires to be recognized claims an identity, usually through a personal identification number, a user name, or a smart card. The system conducts a one-to-one comparison to determine whether the claim is accepted or not. It is similar to "Does the face belong to a specific person? (or) Am I the identity I claim to be?".

3. Face identification: The system conducts a one-to-many comparison to establish an individual's identity without the subject to claim an identity. It is similar to "Whose face is this? (or) Who am I (or) What is my identity?".

4. Watch list: Given a test face image and a set of reference face images, compare all the reference images with the test image, and rank them by the similarity score. If one or more similarity scores are greater than a threshold, select the top ranked reference image. It is similar to "Are you looking for me?

5. Audio-Video based Person -Authentication Audio-video based person authentication is a multimodal person authentication which combines speech and face modalities. The person authentication system, based on single modality such as speech or face, has limitations in both security and robustness. For example, it is possible to collect face images of the person to circumvent a face authentication system. Similarly, the prerecorded speech of a person can be used in the speaker verification system. The joint use of audio and video makes it difficult for an intruder to simultaneously spoof the multiple modalities. The face-based person authentication system is sensitive to variation in lighting conditions. Similarly, the speech-based system may fail when the acoustic environment is noisy. By combining the evidence from these modalities, it is possible to improve the performance and robustness of the system.

6. Issues in Person Authentication- Automatic person authentication appears to be difficult, while it is done effortlessly by human beings. The main reason for this difficulty is that it is difficult to articulate the mechanism human's use. Person authentication by a machine involves the following stages: face detection, feature extraction, modeling and decision logic.

7. Face Detection- Face detection aims to determine the position of a single face in an image. Most of the person authentication methods assume that the location of face in an image is known. Similarly, face tracking algorithms often assume that the initial location of the face is known [1]. Segmentation of a face region from a still image or video is the first step in an automatic person authentication system. The variability in scale or size of the face, orientation, facial expression and lighting conditions complicates the face localization task.

8. Feature Extraction -The primary and important task in person authentication is to extract features representing the person-specific information in the speech signal and face image. It is known that human beings use high level features such as style of speech and verbal mannerisms (for example, use of particular words and idioms) to recognize speakers from their voice. Intuitively, it is clear that these features contain important speaker-specific information. Difficulty arises due to limitations of the existing feature extraction techniques [20]. Current speaker recognition systems use segmental features (which characterize vertical axis (head to toe axis). The face in the image may not be upright and frontal. The images of a face vary due to the relative camera-face pose (frontal, 45 degree, profile, upside down), and some facial features such as an eye or the nose may become partially or wholly occluded.

9. Facial expression- The appearance of the face is directly affected by facial expressions such as smile, fear, surprise and emotion. Smile causes the largest variation in the appearance among all the facial expressions.

10. Illumination: Ambient lighting changes greatly within and between days and among indoor and outdoor environments. The pixel intensity values of the face image are directly proportional to the radiance of the light emitted from the face, which in turn depends on the lighting. The same face image appears different due to a change in lighting. The changes induced by lighting are often larger than the differences between individual faces. If a spotlight is used for capturing image or video, then the changes in the face appearance will be minimal. The spotlight can be used for person authentication applications with the cooperation of the subject. Unfortunately, in the surveillance type of application (person identification), the subject has to be identified under different lighting conditions, without any cooperation from the subject.

11. Aging: The person authentication system should be designed to operate over a longer period of time. The changes in the facial appearance of an individual over a period of time must be addressed.

12. Background noise: If the acoustic environment is not controlled, then the noise due to other sound sources must be removed from the speech signal.

13. Occlusion: Faces may be partially occluded by other objects. In an image with a group of people, some faces may partially occlude other faces.

II. LITERATURE SURVEY

This survey is perhaps the most effective representative and comprehensive because of the analysis of these

algorithms under a common statistical decision framework and the evaluation on a common database with more than hundred different subjects. The experimental results of this survey indicate that the Elastic Graph Matching (EGM) outperforms other techniques. Face detection can be regarded as a specific case of object-class detection. In object-class detection, the task is to find the locations and sizes of all objects in an image that belong to a given class. Examples include upper torsos, pedestrians, and cars. Face-detection algorithms focus on the detection of frontal human faces. It is analogous to image detection in which the image of a person is matched bit by bit.

Image matches with the image stores in database. Any facial feature changes in the database will invalidate the matching process. A reliable face-detection approach based on the genetic algorithm and the Eigen-face technique: Firstly, the possible human eye regions are detected by testing all the valley regions in the gray-level image. Then the genetic algorithm is used to generate all the possible face regions which include the eyebrows, the iris, the nostril and the mouth corners. Each possible face candidate is normalized to reduce both the lightning effect, which is caused by uneven illumination; and the shirring effect, which is due to head movement. The fitness value of each candidate is measured based on its projection on the eigen-faces. After a number of iterations, all the face candidates with a high fitness value are selected for further verification.

Saswati Debnath et.al (2019) Biometrics is an emerging technology in this era, which has been widely used in many application such as secured access to a computer and any other system, criminal identification, person authentication, etc. Face recognition is a biometric method of identifying a person by comparing the data with the stored information of that person. But the recognition of face can be affected by illumination variation, facial expression and other issues. Therefore, the authentication using only face image might be difficult for the system. To overcome this challenge, biometric authentications have to rely on more than one method. In this paper, we consider audio-video person authentication based on two sources of information. Single modality evidence has limitations in both security and robustness.

Therefore, here, audio recognition is added with visual recognition. Facial features, as well as speech is extracted separately from audio-visual data and integrate both the modality for secure user authentication. Mel Frequency Cepstral Coefficients (MFCC) is used as the speech feature. Viola-Jones and Scale-Invariant Feature Transform (SIFT) algorithm are used for visual feature extraction. After the extraction of audio and visual features, the feature selection method is employed. The two-phase algorithm consisting of statistical analysis, Analysis of Variance (ANOVA) with Incremental Feature Selection (IFS) is proposed to select significant features from audio-visual data. The Audio and Video processing is done in two separate phases using machine learning algorithms. The results of both the modalities are then combined at the decision level based on majority voting. It has been observed that multiple modalities of both audio-visual information give immensely good results compared to a standalone single modality.[1]

Aakarsh Malhotra et.al (2018) in many surveillance applications, the cameras are placed at overhead heights for human identification. In such real-world scenarios, the person of interest might be walking away from the camera and the only information available is "image of the person's head". In this research, we investigate the usage of head images for person recognition and propose it as a soft-biometric modality.

With its viability for human recognition, application of head images can also be extended with other face recognition algorithms for surveillance. We propose a head image database pertaining to 103 subjects with more than 600 images. In addition to the database, we propose a framework for head image-based person verification. As a pre-processing stage, the framework includes evaluation of two segmentation algorithms. We also perform benchmarking evaluations of various texture, key-point, and learning-based representation algorithms and establish the baseline results. The experiments suggest that head images can be effectively used to ascertain human identity and the availability of this database could pave further research in this field.[2]

Mahesh R. Pawar et.al (2018) Toady due to rapid increase in vehicles, there is an exponential increase in crime and accidents hence it has become challenge for governments to limit such crimes especially from professional thieves. This paper proposes designing and development of anti-theft as well as driver surveillance embedded system that uses biometric authentication to access the vehicle. This system contains camera which take the image of a person trying to get access of vehicle and compare with authorized person's image and then allowing or denying access. In the case of denial of vehicle access or even if there is an accident occurs, camera will capture the images and email it to the owner or authorizer. This will help to catch thieves, also allows the surveillance of driver and also the inner part of vehicle. The recent work on proposed embedded system is written in this paper. The system is designed and developed using raspberry pi, high resolution camera, vibration sensor and open source software.[3]

Nayansukh Patil et.al (2018) In recent years, all the activities of the countries over the world is carried out Digitally and all the information or data is shared over the network increasing the speed and efficiency of data, but

this transformation of data over the digital network has threat of security i.e. losing the data of the users by the third party unauthorized persons or attackers, cyber-crime has taking consistent efforts to improve the security over the network as all the scams now a days are carried digitally as the data transformation includes money transfer, online shopping, confidential data, social feeds, etc. As to maintain the security a unique identification value or term called password is given to every user and is asked to keep it secret, but the attacker still steals the password using various techniques so to avoid these threat we are using Honey words which will be generated by existing user password and if the attackers enter the password from the honeypot alarm is raised over administrator side, also we maintaining the IP and location tracking of the user and proposing a new technique called video click based captcha scheme to authenticate between humans and robots/bots overcoming the problems of graphical password scheme captcha. Thus, this whole architecture protects and secures the data and application over the online network reducing the threats against the unauthorized users.[4]

Maheen Zulfiqar et.al (2019) Face is one of the most widely used biometrics for human identity authentication. Facial recognition has remained an interesting and active research area in the past several decades due to its ever growing applications in biometric authentication, content based data retrieval, video surveillance, access control and social media. Unlike other biometric systems, facial recognition based systems work independently without involving the individual, due to which it does not add unnecessary delay. Its ability of recognizing multiple persons at a time further adds to its speed. There are many face recognition methods based on traditional machine learning that are available in the literature. Improvements are being made with the constant developments in computer vision and machine learning.

However, most of the traditional methods lack robustness against varying illumination, facial expression, scale, occlusions and pose. With the advent of big data and graphical computing, deep learning has impressively advanced the traditional computer vision systems over the past decade. In this paper, we present a convolutional neural network based face recognition system which detects faces in an input image using Viola Jones face detector and automatically extracts facial features from detected faces using a pre-trained CNN for recognition. A large database of facial images of subjects is created, which is augmented in order to increase the number of images per subject and to incorporate different illumination and noise conditions for optimal training of the convolutional neural network.

Moreover, an optimal pertained CNN model along with a set of hyper parameters is experimentally selected for deep face recognition. Promising experimental results, with an overall accuracy of 98.76%, are obtained which depict the effectiveness of deep face recognition in automated biometric authentication systems.[5]

Mithun Dutta et.al (2018) In existing ATM system, most often only personal identification number (PIN) is used to verify authentic user which is not secured enough as it is very easy to copy. Sometimes thieves have a very strong way to steal account information, that's why bio-metric verification system can be a firm solution. The objective of this paper is to provide a more secured method using bio-metric features and message authentication technique. In our proposed method, PIN verification is combined with fingerprint recognition, to identify a customer during ATM transaction. Fingerprint is verified using efficient minutiae feature extraction algorithm. To assure the security while doing transaction through swipe machine, the client will confirm the transaction by an approval message through GSM technology. In both cases, location will be identified through GPS. If any illegitimate person tries to use the card it will automatically be blocked by the system and detail information will be sent to the customer through the message. Hence, the proposed method will provide more security by identifying and reducing the frauds.[6]

R. Divya et.al (2020) Bio-metric frameworks are getting to be progressively important, since they are more reliable and proficient for identity confirmation. One such biometric is gait. The pattern by which an individual walks is mentioned as gait. It's a locomotion that's achieved through the movement of a person's limb. Unlike several approaches gait is a behavioral biometric, that is taken into consideration for user authentication as it shows distinct patterns for every individual. Also, less obtrusion of user has made this biometric method to be more advantageous compared to others. During this survey we tend to concentrate on varied gait approaches, applications and various machine learning techniques which will be used for classification of gait features and its applications.[7]

Chandrakant P. Divate et.al (2018) The research work on, person identification using biometrics like, iris, face, fingerprint, blood and gait based are enormously performed over the period of time. Also the significant efforts have been going in the area of gender identification of a person. Using biometrics of a person atomization in classification of gender has gained massive importance in the area of research. Finding the person's gender using biometrics like fingerprints, iris, face, gait, etc. has been receiving a considerable attention in the past few decade in many areas like forensic, investigation, security applications, e-commerce authentication and automation systems [24]. This paper presents a brief review on the advanced research techniques in the area of gender classification. We have focused on a study of different biometrics a person, which can be utilized for

classifying a person as male or female and have presented a definite review of the present works. This paper will be helpful for the budding researchers to analysis and identify their ideas and interests about biometric that they may prefer in their research work of gender classification.[8]

Abdul Razaque et.al (2018) Advancement of electronic media not only improved our living standard, but created several amenities that makes work bit easier. On the other hand, media faces several challenges. Biometric systems are promising technology with accurate data and secure features. However, many existing Biometric systems based on a single authentication process have higher vulnerability than Multi-biometric systems. Multi-biometric systems accumulate evidence from more than one biometric trait in order to recognize a person. They provide higher recognition accuracy and larger population coverage. Multi-biometric systems store multiple biometric templates for each user, which results in increased risk to user privacy and system security. To secure the individual biometric template by using the fusion method to store the data. The fuzzy vault is proposed. By the end of this paper we would like to propose an idea to further improve the data security and also make this process take the less time.[9]

A.Muthu Kumar et.al (2019) Biometrics is a mode of popularity and identification affirmation that uses physical attributes of a particular person that are impossible or as a minimum hard to mask. Finger-Knuckle-Print and ear print are examined to be maximum reliable physiological bio metric methods. These days a Finger Knuckle Print authentication based totally bio metric system has been efficaciously applied in various sectors. These unique Finger Knuckle Print and ear print authentication offers in reveals its user image that matches with its very own template that is stored inside the database for verification. An ear print is a replica of the parts outside the ear which have impressed a particular surface [3].

Many variants of local binary patterns are widely used for feature extraction process due to their satisfactory performance, for recognition purpose. An effective method of extracting feature patterns for Finger Knuckle Print (FKP) and ear has been proposed in this project. After extracting features of ear and FKP separately, they are fused by feature level fusion process which is done registration phase. The fused results of FKP and ear have been compared with the stored fused values in authentication phase. If the level of matching is achieved, that particular person gets authenticated to access the things else the person will be impostor [4]. This system gives more accuracy in less time. Challenges: The accuracy of authentication is improved in less duration. The chances of intruding is reduced[10].

III. PROPOSED SYSTEM

In this process, we propose a face tracking algorithm with temporal-spatial information and trajectory of confidence. The whole process is divided into Video and Speech association. Trajectories with high confidence are associated with the detection result of the current frame during local association, whereas trajectories with low confidence are associated with the detection results of the current frame are not matched during global association. We determine the association results using a combined model of digital image processing and digital signal processing.

The major steps carried out are, Detection and recognition. Face detection is a computer technology being used in a variety of applications that identifies human faces in digital images. Face detection also refers to the psychological process by which humans locate and attend to faces in a visual scene. In face based systems this can be in the form of a change in the illumination direction and/or face pose variations. Multi-modal systems use more than one biometric at the same time. In this face detection process is implemented the viola jones is used to detect the face region and the detected region will extracted. In the feature extraction process, we can implement the LBP for pattern extraction in image, and MFCC, Energy features extraction in the speech signal. We implement the Multi SVM Classifier is used to recognize the person and then display the result.

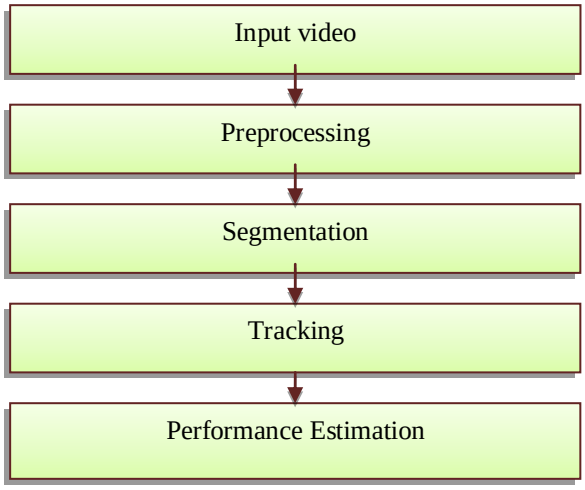

Fig1. proposed flow chart.

1. Modules

- Input
- Preprocessing
- Image Resize
- Noise removal of signal.
- Face detection & extraction
- Feature extraction

- Classification

IV. CONCLUSION

In this Paper, methods are proposed for person authentication using speech, face and visual speech modalities. The performance of text independent person authentication system will be evaluated for by dataset. The performance of text dependent person authentication system is evaluated for real time video using a camera with a resolution of 160 × 120. In this work, we used the method proposed in for determining the face region in the video. The eye region has low luminance, low red chrominance and high blue chrominance when compared to the forehead region of the face. Using these facts, the face region is processed to determine the locations of the eyes. The facial features are extracted relative to the locations of the eyes for each frame in the video. The LBP operation is used for facial feature extraction. The static nature of the visual speech features is extracted relative to the locations of the eyes and mouth using Feature Extraction Method And finally person authentication will be classified by multi SVM Classifier ,After Classification estimated parameters will be calculated in terms of accuracy.

REFERENCES

[1] Saswati Debnath; PinkiRoyMulti-modal authentication system based on audio-visual dataTENCON 2019 - 2019 IEEE Region 10 Conference (TENCON) Year: 2019 ISBN: 978-1-7281-1895-6 DOI: 10.1109/IEEEKochi, India, India

[2] Aakarsh Malhotra; Richa Singh; MayankVatsa; Vishal M. PatelPerson Authentication Using Head Images2018 IEEE Winter Conference on Applications of Computer Vision (WACV)Year: 2018 ISBN: 978-1-5386-4886-5 DOI: 10.1109/IEEELake Tahoe, NV, USA

[3] Mahesh R. Pawar; Imdad Rizvi IoT Based Embedded System for Vehicle Security and Driver Surveillance 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)Year: 2018 ISBN: 978-1-5386-1974-2 DOI: 10.1109/IEEECoimbatore, India

[4] NayansukhPatil ; Rachana PatilAchieving Flatness: with Video Captcha, Location Tracking, Selecting the Honeywords2018 International Conference on Smart City and Emerging Technology (ICSCET)Year: 2018 ISBN: 978-1-5386-1185-2 DOI: 10.1109/IEEEMumbai, India

[5] MaheenZulfiqar ; Fatima Syed ; Muhammad Jaleed Khan ; KhurramKhurshidDeep Face Recognition for Biometric Authentication2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)Year: 2019

[6] MithunDutta ;KangkhitaKaem Psyche ; Tania Khatun ; Md. Ashiqul Islam ; Md. AzijulIslamATM Card Security Using Bio-Metric and Message Authentication Technology2018 IEEE International Conference on Computer and Communication Engineering Technology (CCET)Year: 2018 ISBN: 978-1-5386-7437-6 DOI: 10.1109/IEEEBeijing, China

[7] R. Divya ; Raja LavanyaA Systematic Review on Gait Based Authentication System2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)Year: 2020 ISBN: 978-1-7281-5197-7 DOI: 10.1109/IEEECoimbatore, India, India

[8] Chandrakant P. Divate ; Syed Zakir Ali Study of Different Bio-Metric Based Gender Classification Systems2018 International Conference on Inventive Research in Computing Applications (ICIRCA)Year: 2018 ISBN: 978-1-5386-2456-2DOI: 10.1109/IEEECoimbatore, India

[9] Abdul Razaque ;PrudhviSagarSreeramoju ; Fathi H. Amsaad ; Chaitanya Kumar Nerella ; MusbahAbdulgader ; HarshaSaranuMulti-biometric system using Fuzzy Vault2016 IEEE International Conference on Electro Information Technology (EIT)Year: 2016 ISBN: 978-1-4673-9985-2 DOI: 10.1109/IEEEGrand Forks, ND, USA

[10] A.Muthu Kumar ; A. Chandralekha ; Y. Himaja ; S. MounikaSaiLocal Binary Pattern based Multimodal Biometric Recognition using Ear and FKP with Feature Level Fusion 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS) Year: 2019DOI: 10.1109/IEEETamilnadu, India.