

A Survey on Voice Authentication Embedded Solution for Secured Access Control

Rohit S. Waykole

Dept. of E&TC, MIT College of Engineering, Pune, India

ABSTRACT: Biometrics is becoming more and more popular to identify people and authenticate them for access to secure areas and systems. This paper surveys the security systems based on the voice biometrics. It describes the methods and algorithms used in the existing systems. Accuracy and the speed are two of the biggest factors taken into consideration while implementing the algorithms, so this paper summarizes the current efforts to overcome these problems and provide better accuracy and speed.

KEYWORDS: biometrics, speakers, voice authentication

I. INTRODUCTION

A biometric is a physical trademark, a measure of a biological trait such as a unique mark. Biometrics has an extremely helpful application in security; it can be used to validate a person's identity and control access to a restricted area or electronic framework, in view of the fact that sure of these physical qualities can be utilized to remarkably recognize people. All security frameworks that utilization client based approval oblige clients to be precisely distinguished to guarantee that the right get to benefits are conceded. Biometrics as an confirmation instrument is intense in light of the fact that not at all like different strategies as of now used to confirm individuals, for example, passwords or get to control identifications, it can't be effectively taken away, lost, falsified, or overlooked. There are several categories of biometrics: fingerprints, hand geometry, retina, iris, face, handwriting, and voice

It is conceivable to verify a user through three diverse methodologies: something that he knows, similar to a secret key i.e. password or a PIN, something that he has, similar to a key, or something that he is, biometric characteristics. The biometric frameworks are more basic on the grounds that since the client does not need to recollect the watchword or to fear losing it, "yet they are not secret. You leave your fingerprints on all that you touch, and your iris examples can be watched anyplace you look." [1]. The improvement of the biometric frameworks is firmly coupled to the IT advancements, and this is the motivation behind why today they are extremely utilized. There are countless strategies: unique mark, iris, signature, walk, hand geometry, voice, retinal example, and so forth. It is conceivable to recognize these qualities into principle fields: -

- **Physiological:** unique mark, iris, hand, confront;
- **Behavioral:** voice, signature, walk.

As per their disparate quality they can be utilized as a part of various situations. In this paper, we will examine the biometric authentication using voice, specifying first of all how it works, the problems connected to its usage (like legal and privacy issues) and ultimately, the attack risks that such a system may suffer.

II. MOTIVATION

Today, security is very much essential in all kind of activities. Illegal activities are happening in every place today. So government and corporate sections are concentrating mainly on the security levels with their every invention. This will bring privacy all over the world. So in a thought of bringing privacy through security level, this project has been developed Voice authentication has various advantages. The cost of execution is low in light of the fact that there is no unique equipment required. A simple telephone or microphone is all that a user needs to authenticate using her voice. Different techniques for biometric confirmation like fingerprinting and retinal outputs require uncommon gadgets. Voice verification is easy to use and effectively acknowledged by clients. It is very characteristic to talk. It is not as normal to put an eye up to peruser. The idea of distinguishing individuals by voices is additionally very regular. Each time somebody answers a phone call, the regular nature is to attempt to distinguish the guest by his voice. Maybe most essential to the eventual fate of voice biometrics is that it is the just biometric that permits clients to confirm remotely. Permitting a client to call a telephone number and confirm with her bank vocally to play out an exchange is significantly less demanding than requesting that the client go to the bank face to face and confirm by means of unique mark. It rushes to enlist in a voice confirmation framework. The client is asked to talk a specific arrangement of words or states, or to represent a specific period of time. From that example, a computerized portrayal of the voice, called a

voiceprint, is made. A decent voiceprint is between 2-8 seconds of speech. Validating a client is proficient by looking at the voiceprint that was made at enlistment to an example given when the client needs to enter the limited territory or framework.

III.LITERATURE SURVEY

In literature, the problem and the previous techniques of voice authentication are described

HairolNizam Mohd. Shah et.al.[1] introduced the voice recognition algorithm is produced by utilizing MFCC technique to separate the element of the voice flag. The reference voice is being put away in preparing stage and contrast and the voice in testing stage to coordinate the both outcomes. The framework is effectively perceive the validate client's voice and rejected all the others impostor's voice. The yield result is separated into two classes which are acknowledged and dismisses. In the event that acknowledged, the Arduino will actuate the magnet door to unlock. On the off chance that the yield is rejected, the Arduino will remain the magnet entryway as bolt and the bell will caution for 1 second.

Y. Taigman, et.al.[2] proposed the perfect face classifier would perceive confronts in precision that is just coordinated by people. The fundamental facedescriptor should be invariant to posture, light, expression, and image quality. It ought to likewise be general, in the feeling that it could be connected to different populaces with little adjustments, if any by any means. Furthermore, short descriptors are ideal, and if conceivable, inadequate elements. Unquestionably, fast calculation time is likewise a worry. They accept that this work, which departs from the recent trend of using more features and employing a more powerful metric learning technique, has addressed this challenge, closing the vast majority of this performance gap. The work illustrates that coupling a 3D show based arrangement with extensive limit feedforward models can adequately gain from numerous illustrations to beat the disadvantages and constraints of past strategies. The capacity to introduce a checked change in face acknowledgment bears witness to the capability of such coupling to end up distinctly huge in other vision spaces too.

Scheffer et.al.[3] proposed an extra preferred standpoint of biometric speaker recognition frameworks is their capacity of performing remote verification by utilization of telephones on the other hand versatile applications, without the need of utilizing specific gear, barring the amplifier of the telephone. Voice biometrics are thought to be an exceptionally convenient and natural method for identity recognition.

Weiwu Jiang et.al.[4] proposed the different acoustic features, speaker displaying systems, session-changeability lessening strategies, and VAD plans have been utilized for individual systems. This procedure has prompted to a huge execution pick up when the subsystems were melded. In particular, the combination framework lessens the EER by 42% and min DCF by 56% when contrasted and the best individual subsystems. It was additionally found that the recently proposed FSH subsystem is integral to JFA and performs altogether superior to anything JFA when its projection frameworks were prepared by the sort of discourse that matches the evaluation conditions.

D. P. Munteanu et.al.[5] introduced a programmed speech verification system was proposed in view of a nonstop discourse recognizer. In the preparing stage, are prepared two recognizers: one speaker dependent for the customer model and one speaker-free for the world model. Both recognizers have a similar structure what's more, number of parameters. In the check organize, for the input state, they are playing out a constrained arrangement Viterbi system. A standardized acoustic score is acquired and contrasted and an edge for acknowledgment choice. The trial comes about uncover that the strategy proposed here can acquire mistake rates under 1%.

A. Larcher, J. F. et.al.[6] introduced the new method is designed for embedded applications. It takes favorable circumstances of a GMM/UBM content autonomous approach and the HMM/Viterbi speech recognition control. What's more we propose to fortify the Transient Structure Information demonstrating by a synchronization issued from a video stream. The utilization of the transient data in a speaker acknowledgment framework permits to move forward execution especially when a moderately little amount of discourse information is accessible for preparing and test. Execution of our approach is proportional to the GMM/UBM gauge framework when not considering the phonetic substance (case of EER in KNOWN condition, GMM: 4.49, EBD: 4.56) though the proposed approach beats the GMM/UBM when impostors don't have the foggiest idea about the customer expression (EER in UNKNOWN condition, GMM: 0.87, EBD: 0.56). Moreover, the video-learned synchronization prompts to a pickup in all circumstances, when impostors know or not the client utterance, what's more, beats or is identical to the GMM/UBM benchmark in all circumstances.

H. S. Mohammadi and R. Saeidi, et.al.[7] proposed a novel GMM structure called sorted GMM is presented which is profited from a fast scoring capacity while its memory prerequisite is quite recently hardly higher than customary

GMM. Additionally, the preparation and streamlining plans of GMMs for the proposed strategy are portrayed. The utilization of the proposed strategy in speaker confirmation system is laid out. The execution of a speaker confirmation framework in light of the new strategy was contrasted and other GMM based speaker confirmation frameworks tentatively. The after effects of tests demonstrate that the streamlined sorted GMM presents an attractive execution and it outflanks the as of now known SBM-SGMM which itself gives an adequate execution for quick scoring GMMs while its computational cost is not as much as that of the SBM-SGMM.

IV. PROPOSED SYSTEM

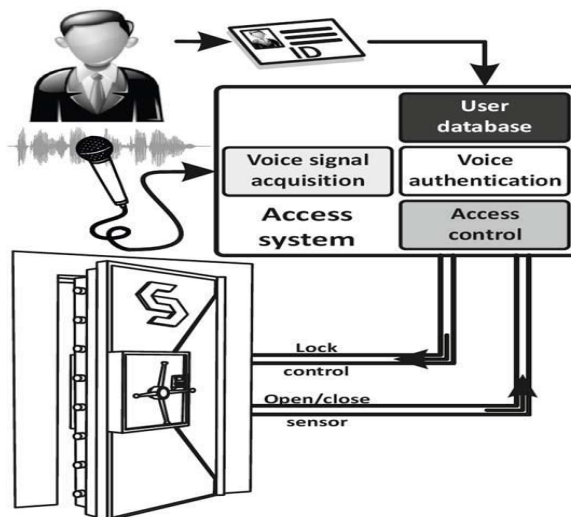


Fig: System Architecture

The system contains a few modules, the undertaking of which is to be up to standard. The most vital piece of the whole framework is the Access Control Module, which forms all of the information got by the framework. Alternate modules guarantee the outside correspondence, information enrollment and legitimate capacity of the User Database. It additionally takes into account productive handling of computerized signs on account of an inherent skimming point co-processor. The execution of the committed openly accessible continuous working framework, controls the work of the considerable number of peripherals, and additionally preparing the obtained biometric information, and permitting the correspondence with an outer gadget in a local area network.

V. CONCLUSION

The paper surveys the existing voice based security systems and the methods used for them and gives the idea of implementation of an automatic door lock, which uses voice biometrics for speaker authentication and access control.

REFERENCES

- [1] Hairol Nizam Mohd. Shah*, Mohd. Zamzuri Ab Rashid, Mohd. Fairus Abdollah, Muhammad Nizam Kamarudin, Chow Kok Lin and Zalina Kamis "Biometric Voice Recognition in Security System" *Indian Journal of Science and Technology*, Vol 7(2), 104–112, February 2014
- [2] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, USA, pp. 1891–1898, Jun. 2014.
- [3] N. Scheffer, L. Ferrer, M. Graciarena, S. Kajarekar, E. Shriberg, and A. Stolcke, "The SRI NIST 2010 speaker recognition evaluation system," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Prague, Czech Republic, pp. 5292–5295, May 2011.
- [4] Weiwu Jiang, Man-Wai Mak, Wei Rao and Helen Meng "The Hkcpu System For The Nist 2010 Speaker Recognition Evaluation 1978-1-4577-0539-7/112011 IEEE 5288 ICASSP 2011 pp. 5288-5291
- [5] D.-P. Munteanu and S.-A. Toma, "Automatic speaker verification experiments using HMM," in *Proc. IEEE 8th International Conference on Communications*, Bucharest, Romania, pp. 107–110, Jun. 2010.
- [6] A. Larcher, J.-F. Bonastre, and J. S. Mason, "Short utterance-based video aided speaker recognition," in *Proc. IEEE 10th Workshop on Multimedia Signal Processing*, Cairns, Australia, pp. 897–901, Oct. 2008.
- [7] H. S. Mohammadi and R. Saeidi, "Efficient implementation of GMM based speaker verification using sorted Gaussian mixture model," in *Proc. European Signal Processing Conf. EUSIPCO*, Florence, Italy, Sep. 2006