System for authentication using Voice Biometric

Saurabh Pandey
Computer Engineering department Thakur college of Engineering and technology
Mumbai, India
Sp157311@gmail.com

Tarun Pandey
Computer Engineering department Thakur college of Engineering and technology
Mumbai, India
tarunpandey1998@gmail.com

Smriti singh
Computer Engineering department Thakur college of Engineering and technology
Mumbai, India
smrititanwar07@gmail.com

Abstract— **Traditionally the use of biometric devices has improved our ability to provide authenticated access to physical installations. With the rise in security breaches in recent years, many organizations believed that two-factor authentication would be the answer. Voice Biometric can be one of the factors of authentication.**

In this project, there will be three phases- Feature Extraction, Training and Matching. The features will be extracted by using MFCC (Mel Frequency Cepstral Co-efficient) technique. In the training phase the extracted features will be trained using the Gaussian Mixture Model (GMM). In the last phase, the algorithm on a dataset and the model will be used to predict the appropriate classes by finding the log likelihood of voice sample.

Keywords—
MFCC- *Mel-frequency Cepstral Coefficients*
GMM- *Gaussian Mixture Model*
EM- *Expectation Maximization*

I. INTRODUCTION

With recent advancement in technology voice recognition has become one of the efficient measure that is used to provide protection to human's computerized and electronic belongings. . It is one of the types of biometric that is used to identify and authenticate user on the basis of his/her voice. Voice recognition is divided into two types: text dependent and text independent. Text dependent recognition identifies user against a phrase while text independent recognition identifies the user irrespective of what he is saying. The success in both cases depends upon the various speaker characteristics which differentiate the one speaker from other. All voice recognition systems comprises of two modules feature extraction and feature matching. In feature extraction, data from the voice sample is extracted to represent the certain speaker and in feature matching the extracted features from the input voice sample is matched against a set known speaker.

II. EASE OF USE

- **Entertainment:** Voice recognition can be used to change TV or radio channels, open and close screens, and play movies. It can also help personalize customer experience. For instance, services such as Netflix and Hulu can be personalized by determining the age of the user through voice analysis, enabling them to access age-appropriate content.

- **Healthcare:** The global healthcare biometric market is expected to reach USD 14.5 billion by 2025, according to a recent report by Grand View Research, Inc. In an industry where data security is paramount, physicians can use voice biometrics to dictate and record patient's health conditions directly into the system and securely retrieve patient's personal history. This can significantly benefit patients who need to share medical records between various doctors. The system can also help dramatically reduce fraud for providers and payers by automating payment collection, and improve patient satisfaction by offering an additional payment option.

- **Banking:** Customers can use voice authentication to operate bank lockers. Banks, on the other hand, can leverage the system to enable highly secure and advanced voice-based payments. With fraud on the rise, credit card companies and banks such as Citibank and ANZ use voice biometrics to proactively identify fraudsters and authenticate callers at their call centers.

- **Education**: Educational institutions can use voice recognition to provide flexibility to students with visual disability, helping them take online exams using voice authentication.

- **Independent Software Vendors (ISV):** For ISVs, voice authentication can enable enterprise sign on mechanisms such as those based on Active Directory, enabling authentication uniformity across enterprise applications and strengthening compliance with accessibility standards.

III. How it works

The first step in any voice recognition system is for the user to give an input by speaking a word or a phrase into a microphone. Then an analog to digital converter converts the electrical signal into digitized form and stored it in the memory. The computer then attempts to determine the meaning of a voice sample by matching it with a template that has a known meaning. This is an analogy to the traditional inputs from a keyboard. The greatest hurdles in Speaker-independent speech recognition systems are articulations and variety of accents used by the people having different nationalities. Other factors that present a challenge to voice recognition technology are acoustical noise and variations in recording environment which are beyond speaker variability. The developed system is consisting of three processes:

- Features extracting
- Training
- Matching

In first process, the developed system will result as computed features of human voice. These features are voice features which are taken from the persons. These features are extracted by using MFCC (Mel Frequency Cepstral Coefficient) technique. MFCC is used as the acoustic features of human voice. It considers the human voice pitch in the form of frequencies and scale them on the Mel scale, these extracted feature are unique to others.In training process, the extracted features are trained using the Gaussian Mixture Modeling. Expectation Maximization (EM) algorithm is used to train the extracted features of human voice in system and then finally used to store in database.In the last phase, the model so constructed will be used in predicting the appropriate class using the validation dataset. We implement MFCC with GMM techniques in order to identify the speaker.

A. Algorithms

(a). MFCC Algorithm:

it is used as the acoustic features of human voice. It considers the human voice pitch in the form of frequencies and scale them on the Mel scale, these extracted feature are unique to others.

(b). GMM Algorithm:

In training process, the extracted features are trained using the Gaussian Mixture Modeling.it recognizes the speaker on the basis of log probability of voice vector and compares it to previously stored value.the log probability equal to the stored value provides access to the entire speaker.

(c). EM Algorithm:

It is used to estimate the maximum likelihood .and it returns the mean M,variance V,weight W and log probability.

B. Applications:

- Entertainment
- Healthcare
- Banking
- Education
- Independent software vendors(ISV)

C. Advantages:

- Increased Security
- Decreased Fraud
- Improved Customer Experience
- Reduced Costs
- High Reliability Rate
- Easy to use
- Minimally Invasive

D. Methodology:

The proposed system was designed using methodology of incremental model. Incremental Model is used for designing, integrating and testing the system. Incremental Model is chosen for this system because system can be developed and delivered in increments, accommodate changes that evolve with time, are easy to test and debug and easier to manage risks involved.
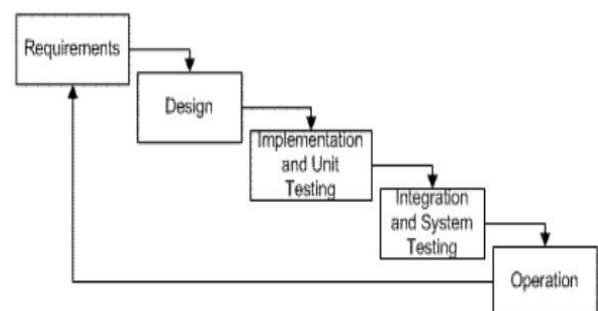
Acknowledgment

References

For a paper in a Proceedings/Transactions/Journal
[1]   J Woods and S O Neil,"Subband coding of Images", IEEE Trans on Acoustic speech signal processing, Vol 1, No.3, pp. 1278- 1288, 1986.

For a report in a conference record
[1]   S. T. Haiang and J.W. Woods, "Embedded image coding using zero blocks of subband/wavelet coefficients and context modeling", IEEE Int. Conf. on Circuits and Systems (ISCA2000), vol 3,pp.662-665, Los Angeles, CA, May 2000.

For an authored book
[1]   B. Klaus and P. Horn , Robot vision , Cambridge, MA, MIT Press, 1986.
[2]   Darwin, C.R. " The Expression of emotion in man and animals", John Murray publishers, 1st Edition, London, pp 8, 1978.

For a technical report
[1]   H. Freitag, "Design methodologies for LSI circuitry", IBM Tech, Place, Rep TR41736, pp 80-82, 1983.

For a Handbook
[1]   Transmission Systems for Communications, 3 rd ed., Winston- Salem, NC, 1985, pp 44-60.

For a website
[1]   URL   , http://www.UK.Research.Att.com/Facedatabase.Html(visited on date).