*System for authentication using Voice Biometric*

Saurabh Pandey

*Computer Engineering department Thakur college of Engineering and technology*

Mumbai, India

[Sp157311@gmail.com](mailto:       Sp157311@gmail.com)

Tarun Pandey

*Computer Engineering department Thakur college of Engineering and technology*

Mumbai, India

[tarunpandey1998@gmail.com](mailto:tarunpandey1998@gmail.com)

Smritisingh

*Computer Engineering department Thakur college of Engineering and*

*technology*

Mumbai, India [smrititanwar07@gmail.com](mailto:%20%20%20%20%20%20smrititanwar07@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 OFUSE

growth due to poor irrigation or low initial nitrogen content. Birds may fear a scarecrow for a time, but they will soon realizeitisnotaliveandreturninnumbers–meaningtheonly solutionisalethalone.Largeranimalslikebaboonsanddeer can only be kept away with tall often electrified fences with empty cans to make a noise. In the midst of all these preventative measures, farmers may forget about a sectionof an orchard or field – the consequences of which will only be apparent when it is too late to remedy. Crop diseases can be devastatingandclassifiedasfungal,bacterialorviral.Drones equippedwithInfraredcamerascanseeinsideplants,givinga clearimageoftheconditionthereof.Ifafarmercandetectan infectionbeforeitspreads,preventativemeasurescanbetaken

- like removing the plant -before the infection spreads to neighboring plants. Drones can be equipped with a multispectral camera that can detect the water content underground,whichcanallowafarmertodetermineifacrop row is parched or over hydrated. Farmers can adjust their irrigation until even water table is created, eliminating water as a source of crop distress. If multispectral camera is not available, the drone can still take overhead pictures of the fields and highlight areas where plants seem discolored or smaller than the rest of the field. This technique can also be used to detect underground leaks in existing irrigation systems. The power of wind is often underestimated and can interferewithdewfall,leadingtoirregularwaterdeliveryfor the field. Heavy winds can reduce pollination, making plants shed and produce lessfruit.

1. *Crime and Damage reports*

Thedronecanaidtheauthoritiesinidentifyingthevehicle used by the criminals, speeding up the investigation. Traditionally, some farmers have employed guards to patrol entrances and along fences, but this is expensive and, due to the existence of boredom and bribery, not alwayseffective.

1. *Forestry*

Open Forest can use drone-based forest and landscape mapping to provide a new perspective for valuation, monitoring and research. Hundreds of pictures taken by drones are stitched together to large and high resolution orthomaps. These orthomaps can then be integrated into GIS systems and used for analysis, planning and management. Nova drone uses drone technology to improve forest management and operational planning, including the monitoring of illegal activities and encroachment.

**978-1-7281-2098-0/19/$31.00 ©2019 IEEE**

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 usedas 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,varianceV,weight W and log probability.

B.Applications:

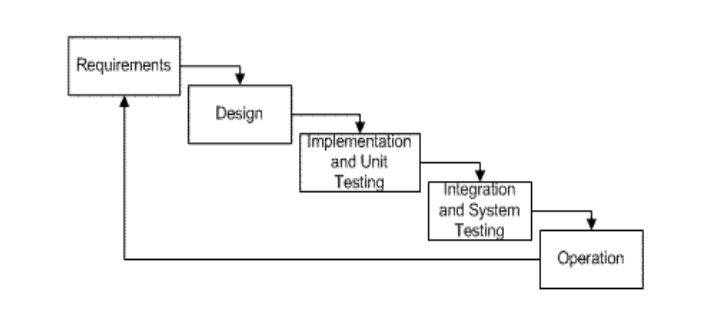
* Entertainment
* Healthcare
* Banking
* Education
* Independent software vendors(ISV)

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

We sincerely thank our guide Ms.Veena Kulkarni for her guidance and constant support and also for the stick to our backs. We also thank the project coordinators for arranging the necessary facilities to carry out the project work. We thank the HOD, Dr. SheetalRathi, Dean Academic, Dr. R.R Sedamkar, the Principal, Dr. B. K. Mishra and the college management for their support.

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 database , http://www.UK.Research.Att.com/Facedatabase.Html(visited on date).