

## Voice Based Database Management System Using DSP Processor

Ms. Diksha Warikoo  
UG Student

Dept. of E&TC Engineering,  
PCCOE, Pune, Maharashtra, India  
dikshawarikoo@gmail.com

Ms. Sujata Wele  
UG Student

Dept. of E&TC Engineering,  
PCCOE, Pune, Maharashtra, India  
sujatawele@gmail.com

Mrs. Rajini.P.K.  
Asst. Prof.

Dept. of E&TC Engineering,  
PCCOE, Pune, Maharashtra, India  
rajani\_ranjith@yahoo.co.in

**Abstract** — Security is provided to customers through PIN/ID protection, to secure their data and information using password. This method require user to authenticate them by entering password. There are cases of fraud and theft when people can easily know the password. But there is a existing technology known as Bio-metric Identification System. It uses an individual's bio-metric characteristics, that is unique and therefore can be used to authenticate the user authority access. This invention is an implementation of speaker verification for attendance monitoring system using DSP Processor. First, speech signal will go to pre-processing phase, where it will remove the background noise then, speech signal's features will be extracted using Mel Frequency Cepstral Coefficients (MFCC) method. Then, using Hamming Window, the features will be matched with the reference speech in database. The speaker is identified by comparing the speech signal from the tested speaker. The main focus of this invention is speaker verification, which is compared between speech signal from unknown speaker to a database of known speaker using utterances. The speaker identification is used in this invention for creating database and identifying the students for maintaining attendance record. LCD Display is interfaced with Processor to show the presence of the students for a particular subject. So this can be used for monitoring the attendance of students. Also defaulter student's list is find out according to the criteria, and it is maintained in MS Excel sheet. Future scope for this can be, informing the monthly attendance to their parents through a text message on their mobile phones using a GSM Module interfaced to processor.

**Keywords**- Discrete Fourier Transform, GSM, Mel Frequency Cepstral Coefficient.

\*\*\*\*\*

### I. INTRODUCTION

The inherent superiority of speech over other modes of communication and its seamless integration with other services makes it the favorite choice for providing easy -to-use, efficient and affordable linkage medium between man and machine.

In this system RFID reader card is provided to respective subject teacher. By swapping card teacher will start their Monitoring for specific subject by checking card number on LCD screen. After this; teacher is ready to start Monitoring. Database in the form of voice is saved in .wave file. Attendance will take for six days of week. If speaker's samples are matched with his or her database then attendance will be done automatically in excel sheet. Linkage of excel sheet with Matlab program makes automatic monitoring

Excel sheet contains Roll no. of student. Monitoring is done for six days and weekly defaulter is decided on the days. If student is present for lecture less than four days then he or she will be considered as a defaulter for a week.

Verification of speaker is the process of deciding whether or not an unknown speech specimen was spoken by the individual speaker whose identity was claimed.

Growth of speaker recognition system during the last six decades is as follows: In 1950's Identifying Spoken digit 0 to 9, Later System able to recognize vowels. In 1960's Use of Spectral Resonance as well as Use of filter bank and logic circuits are developed. Replacement of filter banks by formant analysis is developed. In 1970's isolated word and discrete utterance recognition is developed. Pattern recognition idea also developed in this decade. 1980's Recognition spoken strings, various matching algorithms were develop i.e. HMM, Polynomial coefficients and Artificial Neural Networks (ANN) approach was developed. 1990's Error minimization concepts are developed at the same time description training and kernel-based methods was widen. 2000's Speech-to-Text

(Automatic Transcription technology). Spontaneous speech recognition and Various other techniques were developed and improvements over some earlier techniques were proposed.

A systematic study of various spectral features can be found in (Kinnunen.2004) Out of all existing feature, MFCC is the most popular and has become standard in speaker verification system.

Mel Frequency Cepstral Coefficients (MFCCs) is a algorithm used in automatic speech and speaker recognition, introduced by Davis and Mermelstein in the 1980's, and have been popular ever since. Before introduction of MFCCs, Linear Prediction Coefficients (LPCs) and Linear Prediction Cepstral Coefficients (LPCCs) were the main feature type for speaker recognition.

### II. BLOCK DIAGRAM

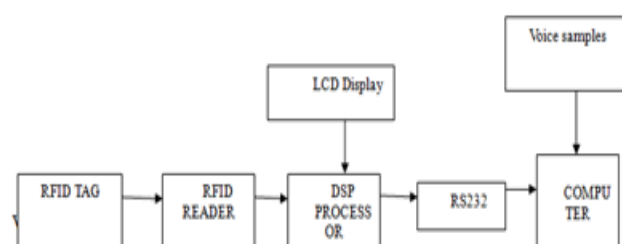


Figure 2.1: voice based database management system

#### Block Diagram Description:

1) Collection of data from students as voice samples: To implement this system, the database is collected. The speech has to be collected in high, low, pitched voices. This is a text dependent speaker verification; hence used utterances are Roll No 1, 2, 3....like this. Data is recorded

from speakers. Each speaker data set consists of 5 speech samples of varying lengths. This data set is split into three categories – training, tuning and testing. To record the samples a microphone is used. One factor that affected the results would be the distance between the microphone and the speaker's mouth.

#### 2) Processing of data collected from speaker:

In the this phase, the background model is created. The background model is basically a large pool of all sample data. The wave files are created, which are used in further phases. The wave file is a continuous signal, which must be broken down in discrete vectors. Each vector is about 15ms long; because we assume that in this duration the vector is stationary

#### 3) Extraction of features from earlier phase:

Feature extraction stage is the most important one in the entire process, since it is responsible for extracting relevant information like pitches, frequency from the speech frames. Common parameter used in speech recognition is Mel Frequency Cepstral Coefficients (MFCC) due to the following reasons:

- 1) The calculation of these parameter leads to a source-filter separation.
- 2) The parameters have an analytically tractable model.
- 3) Experience proves that these parameters work well in recognition applications.
- 4) Due to its significance, it will be described in two different subsections.
- 5) Another subsection will be devoted to dynamic features. They are the delta and acceleration coefficients, that mean to add the first or second derivative approximation.

#### 4) Verification of speaker:

The recently recorded data is used to find the speakers identity and the verification. This is used to retrieve the reference template from database The speaker then gives a samples of preferably a password. The samples are divided in to a number of frames. The features are evaluated for each frame. The feature vectors so obtained are compared with the reference patterns retrieved from the database for the speaker whose identity is verified

### III. HARDWARE IMPLEMENTATION

#### 1. RFID Module:

Specifications:

- Voltage – DC 5V
- Current - < 50mA
- Operating Frequency – 125Hz
- Read Distance – 10 cm
- Output – TTL (RS232)
- Antenna – Built – in

#### 2.DSP Processor:

Specification:

- Operating frequency 225MHz.
- 8MB of sync. DRAM.
- 512 KB non-volatile memory.
- 4 user LED and DIP switches.
- Power supply 5V.

The DSP Processor is used to display the Roll number of the student present on the LCD

### IV. SOFTWARE IMPLEMENTATION

In this system RFID reader card is provided to respective subject teacher. By swapping card teacher will start their Monitoring for specific subject by checking card number on LCD screen. After this; teacher is ready to start Monitoring. Database in the form of voice is save in .wave file. Attendance will take for six days of week. If speaker is mach with his or her database then attendance will be done automatically in excel sheet. Linkage of excel sheet with Matlab program makes automatic monitoring. Excel sheet contains Roll no. of student. Monitoring is done for six days and weekly defaulter is decided on the total days, If a student is present for lecture less than four days then he or she will be consider as a defaulter for a week.

In this verification process voice signal is taken as input speech signal by microphone. This voice signals are stored in to database. Speaker verification contains two stages that is training phase and testing phase. In training phase system has to train with speaker voice and stored the voice as an input. In testing phase voice samples will match with stored ones and maximum score will show verification successfully. This is an overall concept behind this Speaker Verification.

Stages in the Development of Speaker Verification System. Speech data contains different types of information that conveys speaker identity. These include speaker specific information due to the vocal tract, excitation source and behavioral traits. The structure and dimension of vocal tract as well as the excitation source are unique for each speaker. This uniqueness is embedded into the speech signal during speech production and can be used for speaker verification. The possible verification errors are rejection and acceptance .In the case of rejection, a true speaker is rejected and in the case of false acceptance, an impostor gets accepted as the true speaker. The system usually computes the distance between the two templates and compares it with a threshold which is decided during the training phase. The selection of threshold is a crucial part.

MFCCs are commonly derived as follows:

1. Take the Fourier transform of a signal.
2. Mapping the powers of the spectrum obtained onto the Mel scale
3. Take the logs of the powers at each of the Mel frequencies.
4. Take the discrete cosine transform of the list of Mel log powers as a signal
5. The MFCCs are the amplitudes of the resulting spectrum.

## V. RESULTS

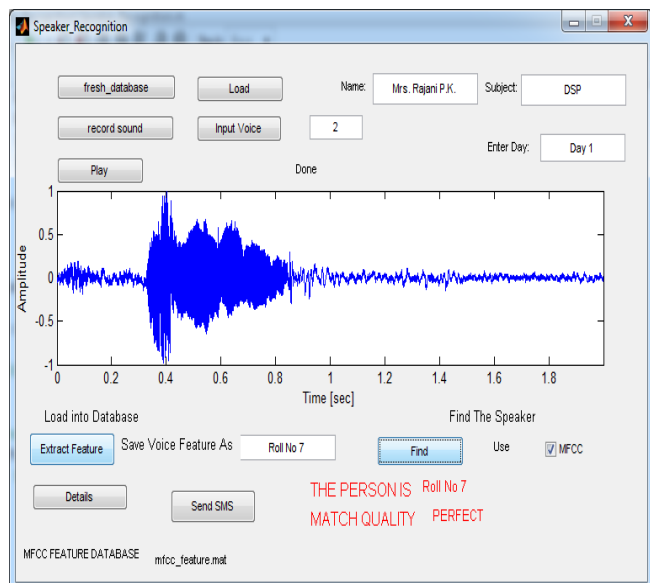


Fig: Speaker verification result

Start button :-This button is used to start the system and enable the hardware.

With the help of this button co-ordination of Hardware and Software will be done.

Fresh Database:- This button is used at the first time of data collection. For refreshing the database.

Record Sound:- By using this button user can able to record voice samples and saved without MFCC features.

Load:- This is used to reload saved sample in the form of sound wave on speaker recognition page. This saved sound wave further used for extract feature.

Extract Feature:- This button used for extract MFCC features and further saved in database for matching.

Input Voice:- This button used to collect sample for voice matching purpose.

Find:- This button used to matched input voice sample with database sample and show the results on the basis of defined threshold.

Play:- With the help of this button we can able to hear recorded voice sample.

Save voice feature as :- Voice features are saved as same name or as same extension as saved in time of recording.

Stop:- Stop button is used to stop the working of Monitoring system.

Send SMS:-This button is used to send SMS to defaulters as per defined contact details in attendance sheet.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Total	Contact Number
Roll No 1		1	1	1	1	1	4	7507772949
Roll No 2	1	1	1	1	1	1	5	9750769759
Roll No 3				1	1	1	2	5762620442
Roll No 4	1	1	1	1	1	1	5	8857800154
Roll No 5	1	1	1	1	1		4	8864811506
Roll No 6				1	1		2	9970391481
Roll No 7	1	1			1	1	4	954008811

Fig: Attendance Sheet in Excel

This is the default format of attendance sheet in excel, in this user has to enter Roll No details like Roll No and his contact details. System Verify Roll No in the excel sheet and as per verification result if matched it marked as 1 against that day. If person is not verified for that day, cell in attendance sheet will be display as BLANK and it indicated as absents or unavailable.

## VI. APPLICATIONS

1. This project enables student monitoring in schools and colleges. Reducing manually keeping the record of attendance of each student.

2. In places like private conferences where number of visitors need to be monitored such systems can be used.

## VII. CONCLUSION AND FUTURE SCOPE

The speaker verification systems yield good performance, if speech data is noise free. Practically, the speech data is noisy due to environment and channel conditions and hence the degradation in performance. Product of System would enroll users, verify their voice-print, and provided a GUI interface for users to do so. Different speakers required different threshold levels. This shows that a global and static threshold could not be defined. By performing speaker verification, this invention can achieve accuracy but this system is not robust in noisy environment. For noise free environment it gives better accuracy but accuracy decrease as noise increases

This project can be used to send SMS note to Parents as Future Scope. System can be used as online which helps Parents to see detailing of their adolescent. In places where security is of great importance this project can be implemented.

---

REFERENCES

- [1] Md Jahangir Alama, Tomi Kinnunen, Patrick Kenny, Pierre Ouellet, Douglas O'Shaughnessy, "Multitaper MFCC and PLP features for speaker verification using I-vectors". *Speech Communication* 55(2):237-251. 2013
- [2] Md Sahidullah, Goutam Saha "A Novel Windowing Technique for Efficient Computation of MFCC for Speaker Recognition", *IEEE*, 2013
- [3] Anand Vardhan Bhalla, Shailesh Khaparkar, Mudit Ratan Bhalla. Gyan Ganga Collage of Technology, "Performance Improvement of Speaker Recognition System". *IJARCSSE* Volume 2, Issue 3, March 2013.
- [4] M. Pathak and B. Raj, "Privacy preserving speaker verification using adapted GMMs," in *Proc. Interspeech*, 2011.
- [5] Amin Fazel and Shantanu Chakrabarty, "An Overview of Statistical Pattern Recognition Techniques for Speaker Verification" - *IEEE Circuits and Systems Magazine*, 11(2), 2011, pp. 62-81.
- [6] N. Dehak, P. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-end factor analysis for speaker verification". *IEEE Trans. on Audio, Speech and Language Processing*, 19(4), 788-798. 2011.
- [7] Alam, J., CRIM, Montreal, QC, Canada, Kinnunen, T., Kenny, P., Ouellet, P.. Multi-taper MFCC features for speaker verification using I-vectors. *Automatic Speech Recognition and Understanding (ASRU)*, 547 - 552. 2011
- [8] M. Senoussaoui, P. Kenny, N. Brummer, E. de Villiers, and P. Dumouchel, "Mixture of PLDA models in I-vector space for gender independent speaker recognition". *Proc. of Interspeech*, Florence, Italy, 25-28. 2011
- [9] N. Dehak, P. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-end factor analysis for speaker verification". *IEEE Trans. on Audio, Speech and Language Processing*, 19(4), 788-798. 2011.
- [10] G. R. Doddington. "Speaker Recognition - Identifying People by Their Voices", in *Proceedings of IEEE*, Vol. 73, No. 11, 1986, pp. 1651-1644.
- [11] G. R. Doddington, "Speaker recognition—Identifying people by their voices," *Proc. IEEE*, vol. 73, pp. 1651–1664, Nov. 1985.

**Websites:**

- [01] <http://www.mathworks.com>
- [02] <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html>

**Books:**

- [1] Shaila D. Apte, "Speech Processing and Recognition", 1st edition.
- [2] Jacob Benesty, Jingdong Ching, "Noise Reduction in Speech Processing", 1st edition.