IJCSN International Journal of Computer Science and Network, Volume 4, Issue 6, December 2015
ISSN (Online) : 2277-5420 www.IJCSN.org
Impact Factor: 0.417

Design of an Intelligent Speaker Recognition System using Mel Frequency Cepstrum Coefficients and Vector Quantization for Biometric Authentication

[1] Sreelakshmi V, [2] Dr. Gnana Sheela K.

[1] M.Tech Student, Department of ECE, TocH Institute of Science and Technology (CUSAT), Arakkunnam, Ernakulum 682313

[2] Professor, Department of ECE, TocH Institute of Science and Technology (CUSAT), Arakkunnam, Ernakulum 682313

Abstract - This paper gives an overview of automatic speaker recognition technology for biometric authentication. A person can be identified by various characteristics such as signature, fingerprints, voice, facial features, etc. This type of authentication methods is known as biometric person authentication. Speaker recognition refers to the process of automatically recognizing who is speaking on the basis of individual information included in speech waves. For a reliable and high accuracy of speech recognition, simple and efficient representation methods are required. In this paper, coefficients are extracted from incoming speech signal using MFCC and it represent trained vector of the speaker. Vector Quantization is the technique used for identification. To identify the speaker, the Euclidean distance between the acoustic vector of test input signal and the mapped codebook is calculated. The trained vector that produces the smallest Euclidean distance is identified as speaker.

Keywords - *Mel Frequency Cepstrum Coefficients (MFCC), Fast Fourier Transform (FFT), Mel Filter Bank; Windowing techniques, Euclidean Distance, Vector Quantization (VQ).*

1. Introduction

Nowadays the interest to different systems of biometric identification among users of computer systems grows up. Spheres of use of technologies of identification are not bounded. Government and private organizations are interested in technologies of face recognition as it allows increasing the level of protection of secret and confidential information. Companies that deal in the sphere of information technologies are interested in technologies of fingerprints, face, voice, iris recognition in order to prevent penetration of outside people to their net. Many famous people point out the increase popularity of biometric systems. Preferences are given to different methods of biometrics. For example, the president of Microsoft Bill Gates stated: “Biometric technologies, those that use voice, will be one of the most important IT innovations of the next several years [38]. Even most of the people come up with such a problem of losing the passwords, for-getting them or even worth, when the passports are stolen. People who agitate for using biometrics point out that those problems with password will not be urgent with the use of biometrics. Bill Gates said at the 2007 RSA Conference: “Passwords are not only weak; passwords have a huge problem. If you get more and more of them, the worse it is. Passwords are a headache for everyone, whether at home or the office, on your PC or your cell phone [39].

1.1 Biometric Authentication

The word “biometrics” came from Greek and it is divided it into two roots: “bio” means life and “metrics” to measure. Biometrical authentication or just biometrics is the process of making sure that the person is who he claims to be. Authentication of identity of the user can be done in three ways: 1) something that person knows (password), 2) something the person has (key, special card), 3) something the person is (fingerprints, footprint). Biometrics is based on anatomic uniqueness of a person and as follow it can be used for biometric identification of a person. Unique characteristics can be used to prevent unauthorized access to the system with the help of automated method of biometric control which, by checking unique physiological features or behavior characteristics identifies the person.

1.2 Speaker Recognition

Speaker recognition is the process of automatically recognizing who is speaking on the basis of individual information included in speech waves. This technique makes it possible to use the speaker's voice to verify their identity and control access to services such as voice dialing, banking by telephone, telephone shopping, database access services, information services, voice mail, security control for confidential information areas, and remote access to computers. Speaker recognition can be classified into identification and verification. Speaker identification is the process of determining which registered speaker provides a given utterance. Speaker verification, on the other hand, is the process of accepting or rejecting the identity claim of a speaker. All speaker recognition systems contain two main modules: feature extraction and feature matching. Feature extraction is the process that extracts a small amount of data from the voice signal that can later be used to represent each speaker. Feature matching involves the actual procedure to identify the unknown speaker by comparing extracted features from his/her voice input with the ones from a set of known speakers.

All speaker recognition systems have to serve two distinguished phases. The first one is referred to the enrolment or training phase, while the second one is referred to as the operational or testing phase. In the training phase, each registered speaker has to provide samples of their speech so that the system can build or train a reference model for that speaker. In case of speaker verification systems, in addition, a speaker-specific threshold is also computed from the training samples. In the testing phase, the input speech is matched with stored reference models and a recognition decision is made.

Speaker recognition, which involves two applications: speaker identification and speaker verification, is the process of automatically recognizing who is speaking on the basis of individual information included in speech waves. This technique makes it possible to use the speaker's voice to verify their identity and control access to services such as voice dialing, banking by telephone, telephone shopping, database access services, information services, voice mail, security control for confidential information areas, and remote access to computers.

Speaker verification (SV) is the process of determining whether the speaker identity is who the person claims to be. Different terms which have the same definition as *SV* could be found in literature, such as voice verification, voice authentication, speaker/talker authentication, talker verification. It performs a one-to-one comparison (it is also called binary decision) between the features of an input voice and those of the claimed voice that is registered in the system. Figure.1 shows the basic structure of SV system (SVS). There are three main components: Front-end Processing, Speaker Modeling, and Pattern Matching. Front-end processing is used to highlight the relevant features and remove the irrelevant ones. After the first component, we will get the feature vectors of the speech signal. Pattern Matching between the claimed speaker model registered in the database and the imposter model will be performed then. If the match is above a certain threshold, the identity claim is verified. Using a high threshold, system gets high safety and prevents impostors to be accepted, but in the mean while it also takes the risk of rejecting the genuine person, and vice versa.

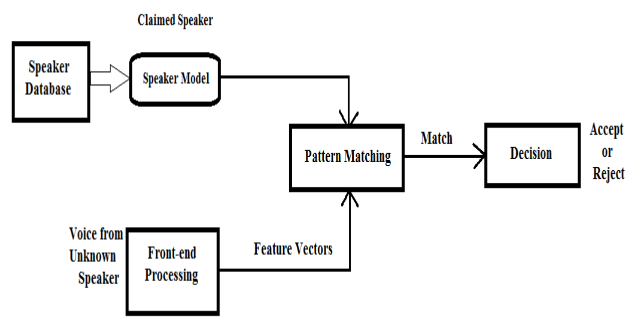

Fig 1. Basic structure of Speaker Verification

Speaker identification (SI) is the process of finding the identity of an unknown speaker by comparing his/her voice with voices of registered speakers in the database. Its a one-to-many comparison. The basic structure of SI system (SIS) is shown in Figure 2. The core components in SIS are the same as in SVS. In SIS, M speaker models are scored in parallel and the most-likely one is reported. In different situations, speaker recognition is often classified into closed-set recognition and open-set recognition. Just as their names suggest, the closed-set refers to the cases that the unknown voice must come from a set of known speakers; and the open-set means unknown voice may come from unregistered speakers, in which case "none of the above" option can be added to this identification system.

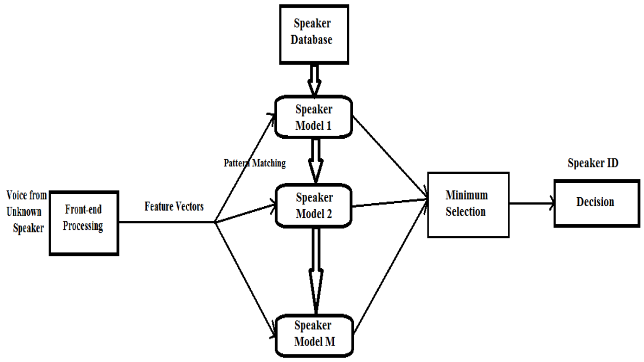

Fig 2. Basic Structure of Speaker Identification

2. Literature Survey

A considerable number of speaker-recognition activities are being carried out in industries, national laboratories and universities. Several enterprises and universities have carried out intense research activities in this domain and have come up with various generations of speaker-recognition systems.

Table 2.1-Comparative Analysis of various techniques for Speaker Recognition

AUTHOR	YEAR	METHODOLOGY	REMARKS
Shanthi Therese S. et al.	2015	MFCC, VQ, Euclidean Distance, K-means Algorithm.	❖ Word level accuracy is 90.02%. ❖ Language recognition accuracy is 97.14%.
Milind U. Nemade et al.	2014	MFCC and GMM.	❖ Less time this is due to more Clock speed of CPU. ❖ More memory.
Mandeep Singh Walia	2014	MFCC using Discrete Fractional Fourier Transform and Vector Quantization.	❖ Modified Mel frequency cepstral coefficients using discrete fractional Fourier transform and vector quantization is used.
RiadhAjgou et al	2014	Auto Regressive(AR)-MFCC and Gaussian Mixture Model(GMM).	
Zhujianchen et al	2014	Linear Prediction Cepstrum Coefficients(LPCC), Mel Frequency Cepstrum Coefficients(MFCC) and Vector Quantization(VQ).	❖ Hybrid LPCC or MFCC feature extraction efficiency has a significant improvement.
Amit Kumar Singh et al.	2014	Mel Frequency Cepstrum Coefficients(MFCC), K-means, Vector Quantization(VQ).	❖ The failure rate of speaker recognition in first case was found to be was found to be 10% while in the second case as found to be 14%.
Liu Ting-ting et al.	2013	Fisher Criterion, Mel Frequency Cepstrum Coefficients(MFCC) and Vector Quantization(VQ).	❖ The efficiency of the algorithm is improved on the basis of high recognition rate, which is more than 80%.
Rishiraj Mukherjee et al.	2013	Gaussian Mixture Model(GMM) and Mel Frequency Cepstrum Coefficients(MFCC).	❖ At 10% FAR we achieved a 95% TAR (recognition ❖ accuracy) but for 5% FAR, the TAR is about 90%.
Shahzadi Farah et al.	2013	MFCC, LPC and VQ.	❖ SRS with MFCC and VQ showed better accuracy as compared to SRS with LPC and VQ.
Shivam Jain et al.	2013	MFCC, VQ, Fuzzy logic and Neural Networks.	❖ The results obtained from Neural Logic were found to be far more accurate than that of vector quantization method.
Nisha.V.S et al.	2013	Filter-bank Based Cepstral Parameters, LPC-based cepstral parameters, Channel Compensation Techniques, DTW, VQ, HMM, GMM.	❖ Stable features that remain insensitive to variation of speakers voice over time
Mahmoud I. Abdallaet al.	2013	MFCC, DWT, Neural Networks.	❖ DWT gives a recognition rate of 99.6 % versus 99.2% using mfccs only.

A.S.Bhalerao et al.	2013	MFCC and VQ.	❖ The identification rate is 90%.
Genevieve I. Sapijaszko et al.	2012	Real Cepstral Coefficients (RCC), MFCC, LPCC, PLPCC and VQ.	❖ Overall MFCC in a noise free. ❖ Recognition rate time(94%).
N.N.Lokhande et al.	2012	MFCC, VQ and Cepstral Mean Normalization.	❖ MFCC features gives more than 95%recognition performance. ❖ CMN features gives good performance over noisy data.
Dr. H B Kekre et al.	2012	MFCC and KMCG.	❖ KMCG is fast and simple algorithm. ❖ The proposed system gives moderate EER of 84%.
FatmazohraChelali et al.	2012	MFCC and VQLBG.	❖ Results show that vector quantization using a codebook of size 16 achieves good results compared to the system without quantization.
Amruta A. Malode et al.	2012	HMM, MFCC, VQ, Viterbi Decoding.	❖ Speaker Recognition system performance can be improved by VQ and HMM.
Jorge Martinez et al.	2012	MFCC, DFT, VQ	❖ The system achieved 100% of precision with a database of 10 speakers.
M.G.Sumithra et al.	2012	MFCC, MMFCC, BFCC, RPLP, LPCC, VQ.	❖ MFCC achieved an identification accuracy of 99.87% .
SupriyaTripathi et al.	2012	MFCC, VQ.	❖ The accuracy of VQ technique is seen to be higher than the MFCC approach.
M.HassanShirali-Shahreza et al.	2011	MFCC, VQ.	❖ Three feature selection choices are compared among which the variance normalization of MFCC is proposed to improve the accuracy.
DankoKomlen et al.	2011	MFCC, VQLBG and the *k*-NN classifier.	❖ The results obtained are encouraging, with an accuracy of more than 95%. In case of interference in the voice signal transmission. ❖ Accuracy ranges from 70% to 85%.
Tiwalade O. Majekodunmi et al.	2011	Fingerprint recognition, face recognition, iris recognition, speaker recognition.	❖ This paper presented a review on four biometric technologies.
Yuan Yujin et al.	2010	LPCC, MFCC, VQ, DTW.	❖ Combination of LPCC and MFCC improved the performance in aspects of the recognition rate and recognition time.
ZhiyiQu et al.	2010	MFCC, VQ Euclidean distance.	❖ The porno-audios are effectively detected using MFCC and VQ.
VibhaTiwari	2010	MFCC, VQ.	❖ MFCC is found to be the best technique.
LindasalwaMuda et al.	2010	MFCC and DTW.	❖ The two voice recognition Algorithms improves the voice recognition performance.
Li Shaomei et al.	2009	MFCC, VQ and Common Codebook.	❖ Save calculation and space resource while having better performance over current algorithm.
Ali Zulfiqar et al.	2009	MFCC and VQLBG.	❖ The accuracy is improved when number of vectors in VQ codebook is increased when sampling frequency is increased.
Izuan Hafez Ninggal et al.	2006	MFCC, PLP, PLAR.	❖ The identification accuracy with PLAR is 97.05% ❖ The successful rate obtained by using LAR is 94.76%.

E. H. Wrench	1981	Markel's technique and Pfeifer's technique. Linear Prediction Coder(LPC).	❖ Markel's technique performs better than the Pfeifer technique.
YosephLinde et al.	1980	Linear Predictive Coding(LPC), Block Quantizers.	❖ The results agreed with the PDP **11/34** run to within one percent.

3. Methodology

Speaker recognition is mainly divided into two categories**:** Speaker identification and Speaker verification. In speaker identification, which speaker has uttered the given speech is found out, whereas in speaker verification, the speaker who is claiming a particular identity is telling the truth or not is determined. The process comprised of three parts. First, preprocessing of the incoming speech sample is performed where we truncated the signal and performed thresholding on it. Then we extracted the features of speech signals using Mel frequency Cepstrum coefficients. These extracted features were then matched with a set of speakers using a Vector Quantization approach.

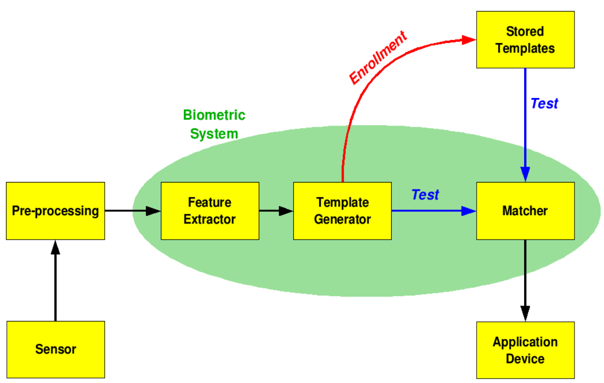

Fig 3. General Block Diagram of a Speaker Recognition System

3.1 Design Flow

Speaker recognition systems contain two main modules: feature extraction and feature matching. Feature extraction is the process that extracts a small amount of data from the voice signal that can later be used to represent each speaker. MFCC is used as the Feature Extraction technique in this project. Feature matching involves the actual procedure to identify the unknown speaker by comparing extracted features from his/her voice input with the ones from a set of known speakers. VQ with LBG algorithm is used for Feature Matching. Figure 3 illustrates the general block diagram of a speaker recognition system.

The speaker recognition system in this project have to serve two distinguished phases:

- Training Phase or Enrollment Session
  - ❖ A model of each speaker's voice is extracted and stored as template models.
  - ❖ Each registered speaker provides samples of their speech.
  - ❖ The voice is trained using MFCC and Vector Quantization.
  - ❖ The system builds or train a reference model for that speaker.

- Testing Phase or Operational Session
  - ❖ The voice of the speaker to be identified or verified is given as input speech.
  - ❖ The input speech is matched with the stored reference model.
  - ❖ The recognition decision depends upon the computed distance between the
  - ❖ reference template and the template derived from the input utterance.

3.2 Feature Extraction

The purpose of this module is to convert the speech waveform, using digital signal processing (DSP) tools, to a set of features (at a considerably lower information rate) for further analysis. This is often referred to as the signal-processing front end. A wide range of possibilities exist for parametrically representing the speech signal for the speaker recognition task, such as LPC, MFCC and others. MFCC is perhaps the best known and most popular, and is described in this paper.

3.2.1 Frame Blocking

In this step, the continuous speech signal is blocked into frames of N samples, with adjacent frames being separated by M (M < N). The first frame consists of the first N samples. The second frame begins M samples after the first frame, and overlaps it by N – M samples. Similarly, the third frame begins 2M samples after the first frame (or M samples after the second frame) and overlaps it by N - 2M samples. This process continues until all the speech is accounted for within one or more frames [40]. The values for N and M are taken as N = 256 (which is equivalent to ~ 30 msec windowing and facilitate the fast radix-2 FFT) and M = 100.

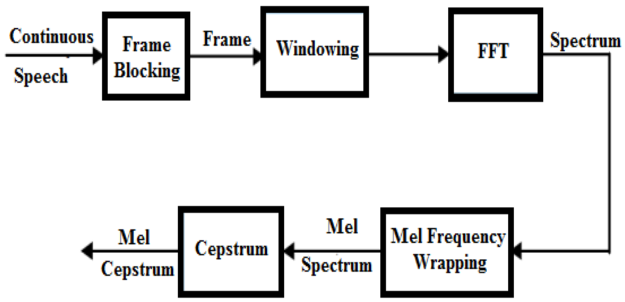

Fig 4. Block Diagram of MFCC Processor

3.2.2 Windowing

The next step in the processing is to window each individual frame so as to minimize the signal discontinuities at the beginning and end of each frame. Windowing is to reduce the effect of the spectral artifacts from framing process. If the window is defined as w(n), $0 \le n \le N-1$, where N is the number of samples in each frame, then the result of windowing is the signal

$$y_1(n) = x_1(n)\,w(n), \quad 0 \le n \le N-1 \qquad (1)$$

Typically the Hamming window is used, which has the form:

$$w(n)=0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \le n \le N-1 \qquad (2)$$

3.2.3 Fast Fourier Transform

The next processing step is the Fast Fourier Transform, which converts each frame of N samples from the time domain into the frequency domain. FFT reduces the computation time required to compute a discrete Fourier transform and improves the performance by a factor of 100 or more over direct evaluation of the DFT. FFT reduces the number of complex multiplications from N2 to (N/2)log2N and its speed improvement factor is N2 /(N/2) Log2N).

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi n/N}, \quad k=0,1,2,\ldots,N-1 \qquad (3)$$

3.2.4 Mel Frequency Wrapping

One approach to simulating the subjective spectrum is to use a filter bank, spaced uniformly on the mel scale. That filter bank has a triangular band-pass frequency response, and the spacing as well as the bandwidth is determined by a constant mel frequency interval. The modified spectrum of S(ω) thus consists of the output power of these filters when S(ω) is the input. The number of mel spectrum coefficients, K, is typically chosen as 20.

3.2.5 Cepstrum

In this final step, we convert the log mel spectrum back to time. The result is called the mel frequency cepstrum coefficients (MFCC). The cepstral representation of the speech spectrum provides a good representation of the local spectral properties of the signal for the given frame analysis. Because the mel spectrum coefficients (and so their logarithm) are real numbers, we can convert them to the time domain using the Discrete Cosine Transform (DCT).

3.3 Feature Matching

The state-of-the-art in feature matching techniques used in speaker recognition includes Dynamic Time Warping (DTW), Hidden Markov Modeling (HMM), and Vector Quantization (VQ). In this project, the VQ approach will be used, due to ease of implementation and high accuracy. VQ is a process of mapping vectors from a large vector space to a finite number of regions in that space. Each region is called a cluster and can be represented by its center called a codeword. The collection of all codewords is called a codebook.

Associated with each cluster of vectors is a representative codeword. Each codeword resides in its own Voronoi region. Given an input vector, the codeword that is chosen to represent it is the one in the same Voronoi region. The representative codeword is determined to be the closest in Euclidean distance from the input vector. The Euclidean distance is defined by:

$$d(x, y_i) = \sqrt{\sum_{j=1}^{k} (x_j - y_{ij})^2} \qquad (4)$$

where x_j is the j^{th} component of the input vector, and y_{ij} is the j^{th} is component of the codeword y_i.

4. Simulation Results

Simulations were done in MATLAB. MATLAB is a high-level language and interactive environment for numerical computation, visualization, and programming. Using MATLAB, one can analyze data, develop algorithms, and create models and applications. Simulation is completed in two phases: Training Phase and Testing Phase. MATLAB program for both Speaker Identification and Speaker Verification had been simulated and the results are analysed.

5. Conclusion

The automatic speaker recognition system consists of 2 phases: enrollment and testing phase. In the enrollment phase, a database of 8 speakers were created and stored in as a reference. The set of speaker's voice samples are trained using MFCC and Vector Quantization. These feature vectors are stored as reference models. In the testing phase, the unknown speaker's identity is matched against the reference models and the recognition is made. Speaker identification and verification are simulated and the results are verified. Simulation is completed using MATLAB 2013a. Speaker recognitions accuracy of 100% was obtained for a set of 8 pre-recorded speakers. The system is also trained using a set of 9 real time recorded speakers and obtained an approximate accuracy of 80% in this case. Implementation phase can be completed using Digital Signal Processor which includes the real time speaker recognition for Biometric Authentication.

References

[1] Yoseph Linde, Andres Buzo and Robert M. Gray, "An Algorithm for Vector Quantizer Design", *IEEE Transactions on Communications*, Vol.28, No.1, pp.84-95, January 1980.

[2] E. H. Wrench, "A Realtime Implementation of a Text Independent Speaker Recognition System", *IEEE International Conference on Acoustics, Speech, and Signal Processing,* Vol.6, pp.193-196, April 1981.

[3] Jialong He, Li Liu, and Gunther Palm, "A New Codebook Training Algorithm for VQ-based Speaker Recognition", *In proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol.2, pp.1091-1094, April 1997.

[3] Izuan Hafez Ninggal & Abdul Manan Ahmad, "The Fundamental of Feature Extraction in Speaker Recognition : A Review", *In proc. of the Postgraduate Annual Research Seminar*, pp.435-439, 2006.

[4] Ali Zulfiqar, Aslam Muhammad and Martinez Enriquez A. M., "A Speaker Identification System using MFCC Features with VQ Technique", *In proc. of Third International Symposium on Intelligent Information Technology Application*, Vol.3, pp.115-118, November 2009.

[5] Li Shaomei, Guo Yunfei, Wei Hongquan "Speaker Recognition via Statistics of Acoustic Feature Distribution", *In proc of International Conference on Multimedia Information Networking and Security*, Vol.2, pp.190-192, 2009.

[6] Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi, "Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques", *In proc of Journal of Computing*, Vol.2, No.3, pp.138-143, March 2010.

[7] Vibha Tiwari, "MFCC and its applications in speaker recognition", *In proc of International Journal on Emerging Technologies*, Vol.1, pp.19-22, February 2010.

[8] Zhiyi Qu, Jing Yu, and Qiang Niu, "Pornographic Audios Detection Using MFCC Features and Vector Quantization", *In proc of IEEE International Conference on Computational and Information Sciences*, pp.924-927, December 2010.

[9] Yuan Yujin, Zhao Peihua and Zhou Qun, "Research of Speaker Recognition Based on Combination of LPCC and MFCC", *In proc of IEEE International Conference on Intelligent Computing and Intelligent Systems*, Vol.3, pp.765-767, October 2010.

[10] M.Hassan Shirali Shahreza and Sajad Shirali Shahreza, "Effect of MFCC Normalization on Vector Quantization Based Speaker Identification", *In proc of IEEE International Symposium on Signal Processing and Information Technology(ISSPIT),* pp.250-253, December 2010.

[11] Tiwalade O. Majekodunmi and Francis E. Idachaba, "A Review of the Fingerprint, Speaker Recognition, Face Recognition and Iris Recognition Based Biometric Identification Technologies", *In proc of the World Congress on Engineering*, Vol.2, pp.1681-1687, July 2011.

[12] Danko Komlen, Tomislav Lombarovic, Mario Ogrizek Tomas, Denis Petek and Andrej Petkovic, "Text Independent Speaker Recognition Using LBG Vector Quantization", *In proc of the 34th International Convention MIPRO*, pp.1652-1657, May 2011.

[13] Supriya Tripathi and Smriti Bhatnagar, "Speaker Recognition", *In proc of Third International Conference on Computer and Communication Technology*, pp.283-287, November 2012.

[14] Jorge Martinez, Hector Perez, Enrique Escamilla and Masahisa Mabo Suzuki, "Speaker recognition using Mel Frequency Cepstral Coefficients (MFCC) and Vector Quantization (VQ) Techniques", *In proc of IEEE International Conference on Electrical Communications and Computers*, pp.248-251, February 2012.

[15] M. G. Sumithra and A. K. Devika, "A Study on Feature Extraction Techniques for Text Independent

Speaker Identification", *In proc of IEEE International Conference on Computer Communication and Informatics (ICCCI),* pp.1-5, January 2012.

[16] Amruta A. Malode and Shashikant L. Sahare, "An Improved Speaker Recognition by Using VQ & HMM", *In proc of IEEE Third International Conference on Sustainable Energy and Intelligent System (SEISCON),* pp.377-383, December 2012.

[17] Fatma zohra Chelali and Amar Djeradi, "MFCC and vector quantization for Arabic fricatives Speech/Speaker recognition", *In proc of IEEE International Conference on Multimedia Computing and Systems*, pp.284-289, May 2012.

[18] Dr. H B Kekre, Dr. V A Bharadi, A R Sawant, Onkar Kadam, Pushkar Lanke and Rohit Lodhiya, "Speaker Recognition using Vector Quantization by MFCC and KMCG Clustering Algorithm", *In proc of IEEE International Conference on Communication, Information & Computing Technology (ICCICT)*, pp.1-5, October 2012.

[19] N. N. Lokhande, N. S. Nehe and P. S. Vikhe, "MFCC Based Robust Features for English Word Recognition", *In proc of Annual IEEE India Conference (INDICON),* pp.798-801, December 2012.

[20] Genevieve I. Sapijaszko and Wasfy B. Mikhael, "An Overview of Recent Window Based Feature Extraction Algorithms for Speaker Recognition", *In proc of IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS),* pp.880-883, August 2012.

[21] A. S. Bhalerao and V. B. Malode, "Implementation of Automatic Speaker Recognition on TMS320C6713 Using MFCC", *In proc of IEEE International Conference on Computer Communication and Informatics (ICCCI),* pp.1-4, January 2013.

[22] Mahmoud I. Abdalla, Haitham M. Abobakr and Tamer S. Gaafar, "DWT and MFCCs based Feature Extraction Methods for Isolated Word Recognition", *In proc of International Journal of Computer Applications*, Vol.69, No.20, pp.21-26, May 2013.

[23] Nisha.V. S. and M. Jayasheela, "Survey on Feature Extraction and Matching Techniques for Speaker Recognition Systems", *In proc of International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE),* Vol.2, No.3, pp.357-362, March 2013.

[24] Shivam Jain, Preeti Jha and Suresh R., "Design and Implementation of an Automatic Speaker Recognition System using neural and fuzzy logic in Matlab", *In proc of IEEE International Conference on Signal Processing and Communication (ICSC),* pp.319-324, December 2013.

[25] Shahzadi Farah and Azra Shamim, "Speaker Recognition System Using Mel-Frequency Cepstrum Coefficients, Linear Prediction Coding and Vector Quantization", *In proc of IEEE 3rd International Conference on Computer, Control and Communication (IC4),* pp.1-5, September 2013.

[26] Rishiraj Mukherjee, Tanmoy Islam, and Ravi Sankar, "Text Dependent Speaker Recognition Using Shifted MFCC", *In proc of IEEE Southeastcon*, pp.1-4, March 2013.

[27] Liu Ting-ting and Guan Sheng-xiao, "On Text-independent Speaker Recognition via Improved Vector Quantization Method", *In proc of the 32nd Chinese Control Conference,* pp.3912-3916, July 2013.

[28] Amit Kumar Singh, Rohit Singh, Ashutosh Dwivedi, "Mel Frequency Cepstral Coefficients Based Text Independent Automatic Speaker Recognition Using Matlab", *In proc of International Conference on Reliability, Optimization and Information Technology (ICROIT),* pp.524-527, February 2014.

[29] Zhu Jianchen and Liu Zengli, "Analysis of Hybrid Feature Research Based on Extraction LPCC and MFCC", *In proc of Tenth International Conference on Computational Intelligence and Security,* pp.732-735, Nov.15-16,2014.

[30] Riadh Ajgou, SalimSbaa, Said Ghendir, Ali Chamsa and A. Taleb-Ahmed, "Robust Remote Speaker Recognition System Based on AR-MFCC features and Efficient Speech activity detection Algorithm", *In proc of 11th International Symposium on Wireless Communications Systems (ISWCS),* pp.722-727, August 2014.

[31] Mandeep Singh Walia," Discrete Fractional Fourier Transform and Vector Quantization Based Speaker Identification System", *In proc of Fourth International Conference on Advanced Computing & Communication Technologies*, pp.459-463, February 2014.

[32] Milind U Nemade and Satish K shah, "Real Time Speech Recognition Using DSK TMS320C6713", *In proc of International Journal of Advanced Research in Computer Science and Software Engineering(IJARCSSE),* Vol.4, No.1, pp.461-469, January 2014.

[33] Shanthi Therese S., Chelpa Lingam, "Speaker based Language Independent Isolated Speech Recognition System", International Conference on Communication, Information & Computing Technology (ICCICT), pp. 1-7,January 2015.

[34] Campbell, J.P., Jr.; "Speaker recognition: a tutorial" *In proc of the IEEE,* Vol.85, No.9, pp 1437 – 1462, Sept. 1997.

[35] Sadaoki Furui, "An overview of Speaker Recognition Technology", *In proc of ECSA workshop on Automatic Speaker Recognition, Identification and Verification*, pp.1-9, 1994.

[36] Aleksandra Babich, "Biometric Authentication. Types of biometric identifiers", *Bachelor's Thesis Degree Programme in Business Information Technology 2012.*

[37] Sujatha K., Nageswara Rao P.V., Rao A.A., Prasad K.R. and Deepthi M.S.B., "Biometric Identity Verification Using Automatic Speaker Recognition", *In proc of IEEE International Conference on*

Electrical, Electronics Signals, Communication and Optimization (EESCO), pp.1-5, January 2015.

[38] http://www.usfst.com/article/Demystifying-Voice-Biometrics--The-Future-of-Security-is-Available-Today/ cited 21.04.2012.

[39] http://www.technewsworld.com/story/59728.html cited 13.04.2012.

[40] John G. Proakis and Dimitris G. Manolakis, “Digital Signal Processing”, New Delhi: Prentice Hall of India. 2002.

[41] Seddik, H.; Rahmouni, A.; Sayadi, M.; “Text independent speaker recognition using the Mel frequency cepstral coefficients and a neural network classifier”, *In proc of First International Symposium on Control, Communications and Signal Processing, Proceedings of IEEE,* pp 631 – 634, 2004.

[42] Y. Linde, A. Buzo & R. Gray, “An algorithm for vector quantizer design”, *In proc of IEEE Transactions on Communications*, Vol. 28, No.1, pp.84-95, Jan 1980.

[43] Moureaux, J.M., Gauthier P, Barlaud, M and Bellemain P.”Vector quantization of raw SAR data”, *In proc of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol.5, pp 189 -192, April 1994.

AUTHOR PROFILE

Sreelakshmi V. completed her B.Tech in Electronics & Communication Engineering under Mahatma Gandhi University. Currently she is pursuing M.Tech in Electronics with specialization in VLSI and Embedded System under Cochin University of Science and Technology (CUSAT).

Dr. Gnana Sheela K received her Ph D in Electronics & S Communication from Anna University, Chennai. She is working as a Professor, Department of ECE at TocH Institute of Science and Technology. She has published 20 international journal papers. She is a life member of ISTE.