International Journal of Engineering Research and General Science Volume 4, Issue 1, January-February, 2016
ISSN 2091-2730

A Survey on Different Algorithms for Automatic Speaker Recognition Systems

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

[1]M.Tech Student, [2] Professor, Department of ECE, TocH Institute of Science and Technology (CUSAT), Arakkunnam, Ernakulum 682313

Email: sreelakshmivk1991@gmail.com

Abstract- In this paper, a literature survey on Different algorithms used for Automatic Speaker Recognition Systems has been done. 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. Speech is a complicated signal produced as a result of several transformations occurring at several different levels: semantic, linguistic, articulatory, and acoustic. Differences in these transformations are reflected in the differences in the acoustic properties of the speech signal. An overview of different algorithms for Feature Extraction and Feature Matching techniques for Speaker recognition systems is presented in this paper.

Keywords- Feature Extraction, Feature matching, Mel Frequency Cepstrum Coefficients (MFCC), Linear Prediction Cepstrum Coefficients (LPCC), Gaussian Mixture Model (GMM), Euclidean Distance, Vector Quantization (VQ), Dynamic Time Warping (DTW), Neural Networks.

INTRODUCTION

The speech signal contains many levels of information. Primarily a message is conveyed via the spoken words. At other levels, speech conveys the information about the language being spoken, the emotion, gender, and the identity of the speaker. The automatic recognition of speaker and speech recognition are very closely related. While speech recognition sets its goals at recognizing the spoken words in speech, the aim of automatic speaker recognition is to identity the speaker by extraction, characterization and recognition of the information contained in the speech signal. The applications of speaker recognition technology are quite varied and continually growing. This technique makes it possible to use the speaker's voice for verification of their identity and thereafter enable the control access to services such as voice dialling and voice mail, tele-banking, telephone shopping, database access related services, information services, security control for confidential information areas, forensic applications, and remote access to computers. Speaker recognition technology is expected to create a host of new services that will make our daily lives more convenient.

Speaker recognition is a commonly used biometric today in most of the commercialization that has taken place for control of access to information services or user accounts on computers. Speaker recognition offers the ability to replace or augment the personal identification numbers and passwords with something that cannot be stolen or lost. There are two main factors, that make speaker recognition a compelling biometric; (1) Speech is natural signal to produce that is not considered threatening by the users to provide, and (2) the telephone system provides a familiar network of sensors for obtaining and delivering the speech signal.

Speaker Recognition

Speaker recognition is a biometric system which performs the computing task of validating a user's claimed identity using the characteristic features extracted from their speech samples. Speaker identification is one of the two integral parts of a speaker recognition system with speaker verification being the other one. On a brief note, speaker verification performs a binary decision which consists of determining whether the person speaking is the same person he/she claims to be or to put it in other words verifying their identity. Speaker identification on the other hand does the job of matching (comparing) the voice of the speaker (known or unknown) with a database of reference templates in an attempt to identify the speaker.

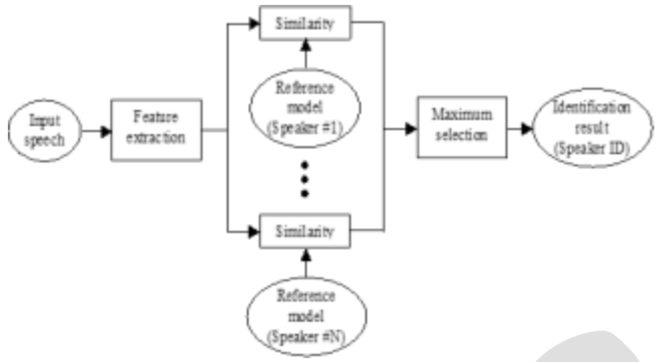

Figure1. Speaker Identification System

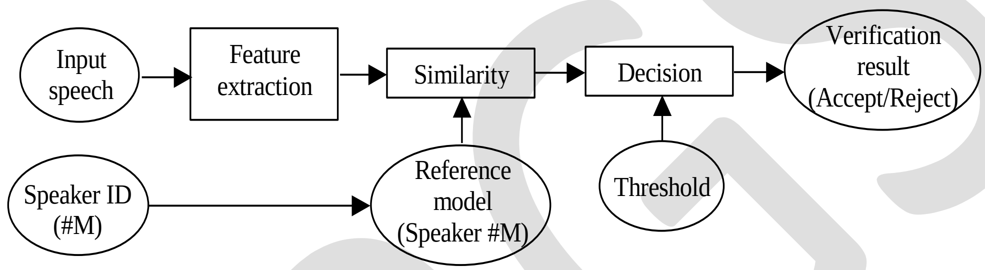

Figure2. Speaker Verification System

LITERATURE STUDY

During the past four decades, a large number of speech processing techniques have been proposed and implemented, and a number of significant advances have been witted in this field during the last one to two decades, which are spurred by the high speed developing algorithms, computational architectures and hardware. Speech recognition refers to the ability of a machine or program to recognize or identify spoken words and carry out voice. The spoken words are digitized into sequence of numbers, and matched against coded dictionaries so as to identify the words. Speech recognition systems are normally classified as to following aspects:

- Whether system requires users to train it so as to recognize users' speech patterns;
- Whether system is able to recognize continuous speech or discrete words;
- Whether system is able to recognize small vocabulary or large one.

A number of speech recognition systems are already available on the market now. The best can recognize thousands of words. Some are speaker-dependent, others are discrete speech systems. With the development of this field speech recognition systems are entering the mainstream, and are being used as an alternative to keyboards.

However nowadays more and more attention has been paid on speaker recognition field. 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.

Related Work

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. Those institutions include AT&T and its derivatives (Bolt, Beranek, and Newman); the Dalle Molle Institute for Perceptual Artificial Intelligence (Switzerland); MIT Lincoln Labs; National Tsing Hua University (Taiwan); Nippon Telegraph and Telephone (Japan); Rutgers University and Texas Instruments (TI). Sandia National Laboratories, National Institute of Standards and Technology, the National Security Agency etc. have conducted evaluations of speaker-recognition systems. It is to be noted that it is difficult to make reasonable comparison between the text-dependent approaches and the usually more difficult text-independent approaches. Text-independent approaches including Gish's segmental Gaussian model, Reynolds's Gaussian Mixture Model, need to deal with unique problems (e.g. sounds and articulations present in the test material but not in training).

It's difficult also to compare the binary choice verification task and the usually more difficult multiple-choice identification task. General trends depict accuracy improvements over time with larger tests i.e. enabled by larger data bases, thus enhancing confidence in performance measurements. These speaker recognition systems need to be used in combination with other authenticators (for e.g. smart cards) is case of high-security applications. The performance of current speaker-recognition systems, however, makes them ideal for a number of practical applications. There exist several commercial ASV systems, including those from Lernout & Hauspie, T-NETIX, Veritel, Voice Control Systems and many others. Perhaps the largest scale deployment of any biometric system to date is Sprint's Voice FONCARD. Speaker-verification applications include access control, telephone credit cards and a lot others. Automatic speaker-recognition systems could help a great deal in reducing crime over fraudulent transactions substantially. However it is imperative to understand the errors made by these ASV systems keeping note of the fact that these systems have gained widespread use across the world. They experience two kinds of errors:- Type I error (False Acceptance of an invalid user (FA)) and Type II error (False Rejection of a valid user(FR)). It uses a pair of subjects: an impostor and a target, to make a false acceptance error. These errors are the ultimate cause of concern in high-security speaker-verification applications.

Yoseph Linde et al (1980) developed an efficient and intuitive algorithm for the design of good block or vector quantizers with 'quite general distortion measures for use on either known probabilistic source descriptions or on a long training sequence of data. An approach of Lloyd algorithm is the basis for this paper which is not a variational technique, and involves no differentiation. Therefore it works well even when the distribution has discrete components, as is the case when a sample distribution obtained from a training sequence is used. The algorithm produces a quantizer meeting necessary but not sufficient conditions for optimality as with the common variational techniques. Both approaches assure atleast local optimality. Algorithms for both known distribution and unknown distribution are described [1].

In 1981, E. H. Wrench Jr. described the design and implementation of a real time speaker recognition system performing text independent, closed set speaker recognition with up to 30 talkers in real time. The reference speech used to characterize these 30 talkers can be extracted from 10 seconds of speech from each talker, and the actual recognition performed with less than one minute of speech from the unknown talker. Two algorithms developed by Markel and Pfeifer were investigated in this paper and the best method is implemented in a laboratory demonstration system [2].Izuan Hafez Ninggal et al (2006) presented a literature survey and discussion on the fundamentals of speaker features in speaker recognition, popular techniques used for feature extraction and their performance evaluation [4].

In 2009, Ali Zulfiqar et al presented a method in which feature vectors from speech are extracted by using Mel-Frequency Cepstral Coefficients and Vector Quantization (VQ) technique is implemented through Linde-Buzo-Gray algorithm. Two purposeful speech databases with added noise, recorded at sampling frequencies 8000 Hz and 11025 Hz, are used to check the accuracy of the developed speaker identification system in non-ideal conditions. An analysis is also provided by performing different experiments on the databases that number of vectors in VQ codebook and sampling frequency influence the identification accuracy significantly. Results show that MFCC based Speaker Identification system with VQ modeling technique has very good identification accuracy and therefore, it is robust against noise [5].

Li Shaomei et al (2009) proposed a new speaker recognition algorithm using acoustic feature distribution around common codebook to model speaker's characteristics. The common codebook is generated via the training data from all reference speakers, which is used to classify speech feature space, and the model of each reference speaker is described by the statistics of speaker's acoustic feature distribution around the common codebook. The method proposed in this paper can save both calculation time and space while having better performance over the current algorithm [6].

In 2010, Lindasalwa Muda et al. discussed two voice recognition algorithms which helps in improving the recognition performance of a speaker recognition system in this paper. The technique was able to authenticate the particular speaker based on the individual information that was included in the voice signal. The results show that these techniques could be used effectively for voice recognition purposes [7]. Vibha Tiwari (2010) discussed several feature extraction techniques and found out that Mel Frequency Cepstrum Coefficients is a well suited technique to describe the signal characteristics. The extracted speech features (MFCC's) of a speaker for designing a text dependent speaker recognition system are quantized to a number of centroids using vector quantization algorithm [8].

Zhiyi Qu et al (2010) presented a method applying MFCC and Vector Quantization algorithm for Pornographic Audio Detection. Firstly, MFCC of selected pornographic audios are extracted and then encoded into codebooks using VQ algorithm; Secondly, all of the codebooks obtained will be averaged to get an average codebook; Finally, the type of any newly input audio belonging to, either pornographic or non-pornographic, will be determined by measuring the Euclidean distance between the average codebook and its own codebook. Experiment results show that the algorithm can detect pornographic audios effectively [9].

Yuan Yujin et al (2010) developed a Linear Prediction Cepstrum Coefficient (LPCC) and Mel Frequency Cepstrum Coefficient (MFCC) are used as the features for text independent speaker recognition in this system. And the experiments compare the recognition rate of LPCC, MFCC or the combination of LPCC and MFCC through using Vector Quantization (VQ) and Dynamic Time Warping (DTW) to recognize the speaker's identity. Experimental results proved that the combination of LPCC and MFCC has a higher recognition rate [10].

In 2010, M.Hassan Shirali-Shahreza et al. described a comparison of the pre-processing techniques of MFCC features using the effect of these two processes on the accuracy of a Vector Quantization (VQ) speaker identification system. One of these pre-processing is creating delta and delta-delta coefficients and append them to MFCC to create feature vector and the other pre-processing is coefficients mean normalization. The results reported that removing the first coefficient can improve the accuracy and using the delta coefficients does not improve the accuracy and even decrease the accuracy [11]. Tiwalade O. Majekodunmi et al (2011) reviewed four biometric identification technologies (fingerprint, speaker recognition, face recognition and iris recognition) in which it discussed the mode of operation of each of the technologies and highlighted their advantages and disadvantages [12].

Danko Komlen et al (2011) described a system based on LBG vector quantization and the *k*-NN classifier. The features that were used for extraction are MFCC coefficients and energy of the sound signal. The developed system was evaluated on two sets of speakers. The results obtained showed an accuracy of more than 95%. The system was also evaluated for the case of interference in the voice signal transmission, and accuracy in this case ranges from 70% up to 85% [13]. Supriya Tripathi et al (2012) proposed the comparison of MFCC and Vector Quantization for Speaker Recognition. Feature vectors are extracted using Mel Frequency Cepstrum Coefficients (MFCC) and Vector Quantization is implemented through Linde Buzo Gray algorithm. These coefficients are used to identify an unknown speaker from a given set of speakers [14].

Jorge Martinez et al (2012) proposed a new method for automatic speaker recognition method in which MFCCs are used to extract features from the voice signal and Vector Quantization to identify the speaker. The approximation to the human voice behaviour is good when using MFCC because the MFCC uses the Mel scale [15]. M. G. Sumithra et al (2012) discussed various feature extraction techniques for text independent speaker identification such as Mel-frequency cepstral coefficients(MFCC), Modified Mel-frequency cepstral coefficients(MMFCC), Bark frequency cepstral coefficients(BFCC), Revised Perceptual liner prediction (RPLP) and linear predictive coefficient cepstrum (LPCC) are implemented and the comparison is done based on performance and computation time. Vector quantization (VQ) codebook has been used for modeling speaker identity. MFCC obtained a very less false rejection rate for the optimum distance minimum value. Also it yields higher identification accuracy [16].

In 2012, Amruta A. Malode et al described the system is able to recognize the speaker by translating the speech waveform into a set of feature vectors using Mel Frequency Cepstral Coefficients (MFCC) technique. Vector Quantization (VQ) is used to make the same number of MFCC coefficients. This paper proposed the idea that the Speaker Recognition system performance can be improved by VQ and HMM. It can improve the accuracy of the system upto 99.99% by more training of data and denoising methods [17]. Fatma zohra Chelali et al (2012) developed a speaker-dependent Arabic phonemes recognition system using MFCC analysis and the VQLBG algorithm. The system is examined with and without vector quantization to analyze the effect of compression in an acoustic parameterization phase. Experimental results show that vector quantization using a codebook of size 16 achieves good results compared to the system without quantization for a majority of the phonemes studied [18]. In 2012, Dr. H. B. Kekre et al. implemented a method using a combination of Mel Frequency Cepstral Coefficients (MFCC) & Kekre's Median Codebook Generation Algorithm (KMCG) for automatic speaker recognition system. The system is implemented as a text-dependent system in which MFCC algorithm is used for feature extraction and codebook generation and feature matching is done using KMCG algorithm. It provides simplicity in implementation and achieves high level of accuracy [19].

N. N. Lokhande et al. (2012) introduced robust features of MFCC for speech recognition in noisy environments. In noisy environments, the performance of MFCC features is degraded. These features further normalized to get new set of features known as Cepstral Mean Normalized (CMN) features. The recognition rate in presence of additive noise is improved using normalized features. Performance of these noise robust features evaluated on own created English digit database using Vector Quantization technique [20].

Genevieve I. Sapijaszko et al (2012) presented a survey that compares and contrasts recent window frames algorithms such as Real Cepstral Coefficients (RCC), Mel Cepstral Coefficients (MFCC), Linear Predictive Cepstral Coefficients (LPCC), and Perceptual Linear Predictive Cepstral Coefficients (PLPCC). These feature extraction methods will be used in conjunction with a Vector Quantization (VQ) method and a Euclidean distance classifier to find the best recognition rate among the feature extraction features. When comparing recognition time, MFCC was faster than all other methods. Overall MFCC in a noise free environment was the best method in terms of recognition rate and recognition rate time [21]. In 2013, A. S. Bhalerao et al proposed an algorithm for implementing automatic speaker recognition system on TMS320C6713 DSP kit using MFCC for feature extraction of the incoming speech signal and Vector Quantization for classification and Euclidean distance between MFCC and trained vectors for speaker identification [22]. Mahmoud I. Abdalla et al (2013) introduced a new method for feature extraction is presented for speech recognition using a combination of discrete wavelet transform (DWT) and mel Frequency Cepstral Coefficients (MFCCs)**.** The objective of this method is to enhance the performance of the proposed method by introducing more features from the signal. The performance of the Wavelet-based Mel Frequency Cepstral Coefficients method is compared to mel Frequency Cepstral Coefficients based method for features extraction [23]. Nisha.V. S. et al (2013) presented a survey on feature extraction and feature matching or modelling techniques that are currently used for Speaker Recognition Systems [24].

In 2013, Shivam Jain et al developed a Text Dependent Automatic Speaker recognition system and it is simulated using MATLAB. The system is trained to store voice of the same person under various physiological conditions such as coughing, shouting, during chewing, mouth covered etc for a better computational efficiency. A dictionary is created to store the signature features of each user's voice. A neural networks is then trained using back propagation and accordingly weights are obtained to recognize voice in the testing phase. The efficiency of the proposed system is then compared to the system implemented using vector quantization [25].

Shahzadi Farah et al (2013) implemented a speaker recognition system (SRS) using Mel-Frequency Cepstrum Coefficients (MFCC), Linear Prediction Coding (LPC) as feature extraction techniques and Vector Quantization (VQ) as speaker classification technique and investigated the effect of noise and pitch alteration on accuracy of the system. Speaker Recognition System with MFCC and VQ showed better accuracy as compared to Speaker Recognition System with LPC and VQ. The accuracy of speaker recognition decreases with increase of noise and the effect of pitch alteration resulted in lower classification accuracy [26]. Rishiraj Mukherjee et al (2013) introduce a novel method to recognize/identify speakers including a new set of features, the shifted MFCC which allowed inclusion of accent information in the recognition algorithm. The algorithm was evaluated using TIDIGIT dataset and the results showed on the average 10% improvement over the performance of previous works [27].

In 2013, Liu Ting-ting et al proposed a paper which mainly included the study of the text-independent speaker recognition. MATLAB software is used to realize the design of the system. Preprocessing of the speech signal is performed as the initial step. Then the features are extracted which involved differential MFCCs. Pattern match judgment is based on vector quantization (VQ) model. The optimum codebook is generated by LBG algorithm. The identification of the speaker is achieved by calculating the distortion between the reference models and the testing model. The decreased feature parameters is obtained through Fisher criterion, which helps reduce the space complexity. The efficiency of the algorithm is improved on the basis of high recognition rate, which is more than 80% [28].

Amit Kumar Singh et al (2014) presented a performance evaluation of MFCC technique when applied to K means clustering using two experiments. The speech features were directly matched in the first experiment and in the second case, a VQ codebook was created by clustering the training features of the speakers. The recognition rate is determined vitally by the choice of number of clusters. The failure rate of speaker recognition in first case was found to be 10% while in the second case was found to be 14%. A better idea regarding the choice of ideal number of clusters for a better recognition is provided in this paper [29]. In 2014, Zhujianchen et al. introduced a method in which the simulation results show that the hybrid LPCC or MFCC feature extraction has a significant improvement in efficiency based on the speaker recognition among the feature extraction, LPCC and MFCC coefficients for complementary advantage, its integration [30].

Riadh Ajgou et al (2014) derived a scheme to improve the performance of Remote Speaker Recognition System in noisy environment in which feature extraction framework is based on the well-known MFCC and autoregressive model (AR) features since MFCC is a very useful feature for speech processing in clean conditions but it deteriorates in the presence of noise. The use of AR-MFCC approach has provided significant improvements in identification rate accuracy when compared with MFCC in noisy environment. However, in terms of runtime, AR-MFCC requires more time to execute than MFCC [31]. Mandeep Singh Walia (2014) proposed a new method that uses modified Mel Frequency Cepstrum Coefficients (MFCC) using discrete fractional fourier transform for feature extraction and Vector Quantization for feature matching or modeling. Speakers identified with comparison between training and testing speech samples [32].

In 2014, Milind U. Nemade implemented a real time speech recognition system. MFCC is used for designing a text dependent speaker identification system. The DSP processor TMS320C6713 with Code Composer Studio (CCS) has been used for real time speech recognition in this paper. After feature extraction from recorded speech, each Euclidian Distance (ED) from all training vectors is calculated using Gaussian Mixture Model (GMM) as it gives better recognition for the speaker features. The command/voice having minimum ED is applied as similarity criteria [33].

Shanthi Therese S. et al (2015) presented a speaker based Language Independent Isolated Speech Recognition System. The most popular feature extraction technique Mel Frequency Cepstral Coefficients (MFCC) is used for training the system. Representative specific features are identified using K-Means algorithm. Euclidian distance function is used for calculating the Distortion measure. Pitch contour characteristics are used to identify the language specific features. Decision rules are formed to recognize language and speech of the given input [34].

CONCLUSION

Speaker recognition techniques alongside with facial image recognition, fingerprints and retina scan recognition represent some of the major biometric tools for identification of a person. Each of these techniques carries its advantages and drawbacks. The question to what degree each of these techniques provides unique person identification remains largely unanswered. If these methods can provide unique identification then, it is still not clear what kind of parametric representations contain information which is essential for the identification process, and for how long and under what conditions, this representation remains valid? As long as these questions are unanswered, there is a scope for research and improvements.

In this survey paper, a study on different feature extraction and feature matching tecniques for speaker recognition systems have been presented. The LPC features were very popular in the early speaker-identification and speaker verification systems. However, comparison of two LPC feature vectors requires the use of computationally expensive similarity measures such as the Itakura-Saito distance and hence LPC features are unsuitable for use in real-time systems. MFCCs are based on the known variation of the human ear’s critical bandwidths with frequency, filters spaced linearly at low frequencies and logarithmically at high frequencies have been used to capture the phonetically important characteristics of speech. Pattern matching is often based on Hidden Markov Models (HMMs), a statistical model which takes into account the underlying variations and temporal changes of the accoustic pattern and other models include Vector Quantization and Dynamic Time Warping is used, this algorithm measures the similarity in between two sequences that vary in speed or time, even if this variation is non-linear such as when the speaking speed changes during the sequence.

REFERENCES:

[1] Yoseph Linde, Andres Buzo and Robert M. Gray, “An Algorithm for Vector Quantizer Design”, IEEE Transactions on Communications, vol. com-28, no. 1,pp.84-95, January 1980.

[2] E. H. Wrench, Jr., “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”, IEEE International Conference on Acoustics, Speech, and Signal Processing, vol.2, pp.1091-1094, April 1997.

[4] Izuan Hafez Ninggal & Abdul Manan Ahmad, “The Fundamental of Feature Extraction in Speaker Recognition : A Review”, Proceedings of the Postgraduate Annual Research Seminar, pp.435-439, 2006.

[5] Ali Zulfiqar, Aslam Muhammad and Martinez Enriquez A. M., “A Speaker Identification System using MFCC Features with VQ Technique”, Third International Symposium on Intelligent Information Technology Application,vol.3, pp.115-118, November 2009.

[6] Li Shaomei, Guo Yunfei, Wei Hongquan “Speaker Recognition via Statistics of Acoustic Feature Distribution”, International Conference on Multimedia Information Networking and Security, vol.2, pp.190-192, 2009.

[7] Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi, “Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques”, Journal of Computing, vol.2, issue.3, pp.138-143, March 2010.

[8] Vibha Tiwari, “MFCC and its applications in speaker recognition”, International Journal on Emerging Technologies, vol.1, issue.,pp.19-22, February 2010.

[9] Zhiyi Qu, Jing Yu, and Qiang Niu, “Pornographic Audios Detection Using MFCC Features and Vector Quantization”, IEEE International Conference on Computational and Information Sciences, pp.924-927, December 2010.

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

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

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

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

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

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

[16] M. G. Sumithra and A. K. Devika, "A Study on Feature Extraction Techniques for Text Independent Speaker Identification", IEEE International Conference on Computer Communication and Informatics (ICCCI), pp.1-5, January 2012.

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

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

[19] 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", IEEE International Conference on Communication, Information & Computing Technology (ICCICT), pp.1-5, October 2012.

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

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

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

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

[24] Nisha.V. S. and M. Jayasheela, "Survey on Feature Extraction and Matching Techniques for Speaker Recognition Systems", International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE),vol.2, issue 3, pp.357-362, March 2013.

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

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

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

[28] Liu Ting-ting and Guan Sheng-xiao, “On Text-independent Speaker Recognition via Improved Vector Quantization Method”, Proceedings of the 32nd Chinese Control Conference, pp.3912-3916, July 2013.

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

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

[31] 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”, 11th International Symposium on Wireless Communications Systems (ISWCS), pp.722-727, August 2014.

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

[33] Milind U Nemade and Satish K shah, “Real Time Speech Recognition Using DSK TMS320C6713”,International Journal of Advanced Research in Computer Science and Software Engineering(IJARCSSE); vol.4,no.1,pp.461-469, January 2014.

[34] 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.

[35] J. P. Campbell, “Speaker recognition: A tutorial,” Proc. IEEE, vol. 85, no. 9, pp. 1437–1462, Sep. 1997.

[36] Sadaoki Furui, “An overview of Speaker Recognition Technology”, ECSA workshop on Automatic Speaker Recognition, Identification and Verification, pp.1-9, 1994.

[37] Aleksandra Babich, “Biometric Authentication. Types of biometric identifiers”, Bachelor’s Thesis Degree Programme in Business Information Technology 2012.

[38] Sujatha K., Nageswara Rao P.V., Rao A.A., Prasad K.R. and Deepthi M.S.B., “Biometric Identity Verification Using Automatic Speaker Recognition”, IEEE International Conference on Electrical, Electronics Signals, Communication and Optimization (EESCO), pp.1-5, January 2015.

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

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