World of Computer Science and Information Technology Journal (WCSIT)
ISSN: 2221-0741
Vol. 2, No. 1, 1-7, 2012

A Review on Speech Recognition Challenges and Approaches

Vimala.C
Project Fellow, Department of Computer Science
Avinashilingam Institute for Home Science and Higher Education and Women
Coimbatore, India

Dr.V.Radha
Associate Professor, Department of Computer Science
Avinashilingam Institute for Home Science and Higher Education and Women
Coimbatore, India

Abstract— Speech technology and systems in human computer interaction have witnessed a stable and remarkable advancement over the last two decades. Today, speech technologies are commercially available for an unlimited but interesting range of tasks. These technologies enable machines to respond correctly and reliably to human voices, and provide useful and valuable services. Recent research concentrates on developing systems that would be much more robust against variability in environment, speaker and language. Hence today's researches mainly focus on ASR systems with a large vocabulary that support speaker independent operation with continuous speech in different languages. This paper gives an overview of the speech recognition system and its recent progress. The primary objective of this paper is to compare and summarize some of the well known methods used in various stages of speech recognition system.

Keywords- Speech Recognition; Feature Extraction; MFCC; LPC; Hidden Markov Model; Neural Network; Dynamic Time Warping.

I. Introduction

Speech is the most basic, common and efficient form of communication method for people to interact with each other. People are comfortable with speech therefore persons would also like to interact with computers via speech, rather than using primitive interfaces such as keyboards and pointing devices. This can be accomplished by developing an Automatic Speech Recognition (ASR) system which allows a computer to identify the words that a person speaks into a microphone or telephone and convert it into written text. As a result it has the potential of being an important mode of interaction between human and computers [1]. Although any task that involves interfacing with a computer can potentially use ASR. The ASR system would support many valuable applications like dictation, command and control, embedded applications, telephone directory assistance, spoken database querying, medical applications, office dictation devices, and automatic voice translation into foreign languages etc.

In the current Indian context, these machine-oriented interfaces restrict the computer usage to miniature fraction of the people, who are both computer literate and conversant with written English. Communication among human beings is dominated by spoken language. Therefore, it is natural for people to expect speech interfaces with computers which can speak and recognize speech in native language. It will enable even a common man to reap the benefit of information technology. India has a linguistically rich area which has 18 constitutional languages, which are written in 10 different scripts [2]. Hence there is a special need for the ASR system to be developed in their native language. This paper provides an overview of speech recognition system and the review of techniques available at various stages of speech recognition.

The paper is organized as follows. Section 2 presents the classification of speech recognition systems and section 3 explains about the growth of ASR systems. Section 4 explains about the overview of ASR system. Section 5 gives details about the speech feature extraction techniques. Section 6 investigates the speech recognition approaches, section 7 deals with the performance evaluation measures available for ASR system. Finally, the conclusion is summarized in section 8 with future work.

II. CLASSIFICATION OF SPEECH RECOGNITION SYSTEMS

Speech recognition systems can be separated in several different classes by describing the type of speech utterance, type of speaker model, type of channel and the type of vocabulary that they have the ability to recognize. Speech recognition is becoming more complex and a challenging task because of this variability in the signal. These challenges are briefly explained below.

A.. *Types of Speech Utterance*

An utterance is the vocalization (speaking) of a word or words that represent a single meaning to the computer. Utterances can be a single word, a few words, a sentence, or even multiple sentences. The types of speech utterance are:

1) *Isolated Words*

Isolated word recognizers usually require each utterance to have quiet on both sides of the sample window. It doesn't mean that it accepts single words, but does require a single utterance at a time. This is fine for situations where the user is required to give only one word responses or commands, but is very unnatural for multiple word inputs. It is comparatively simple and easiest to implement because word boundaries are obvious and the words tend to be clearly pronounced which is the major advantage of this type. The disadvantage of this type is choosing different boundaries affects the results.

2) *Connected Words*

Connected word systems (or more correctly 'connected utterances') are similar to isolated words, but allow separate utterances to be 'run-together' with a minimal pause between them.

3) *Continuous Speech*

Continuous speech recognizers allow users to speak almost naturally, while the computer determines the content. Basically, it's computer dictation. It includes a great deal of "co articulation", where adjacent words run together without pauses or any other apparent division between words. Continuous speech recognition systems are most difficult to create because they must utilize special methods to determine utterance boundaries. As vocabulary grows larger, confusability between different word sequences grows.

4) *Spontaneous Speech*

This type of speech is natural and not rehearsed. An ASR system with spontaneous speech should be able to handle a variety of natural speech features such as words being run together and even slight stutters. Spontaneous (unrehearsed) speech may include mispronunciations, false-starts, and non-words.

B. *Types of Speaker Model*

All speakers have their special voices, due to their unique physical body and personality. Speech recognition system is broadly classified into two main categories based on speaker models namely speaker dependent and speaker independent.

1) *Speaker dependent models*

Speaker dependent systems are designed for a specific speaker. They are generally more accurate for the particular speaker, but much less accurate for other speakers. These systems are usually easier to develop, cheaper and more accurate, but not as flexible as speaker adaptive or speaker independent systems.

2) *Speaker independent models*

Speaker independent systems are designed for variety of speakers. It recognizes the speech patterns of a large group of people. This system is most difficult to develop, most expensive and offers less accuracy than speaker dependent systems. However, they are more flexible.

C. *Types of Vocabulary*

The size of vocabulary of a speech recognition system affects the complexity, processing requirements and the accuracy of the system. Some applications only require a few words (e.g. numbers only), others require very large dictionaries (e.g. dictation machines). In ASR systems the types of vocabularies can be classified as follows.

- Small vocabulary - tens of words
- Medium vocabulary - hundreds of words
- Large vocabulary - thousands of words
- Very-large vocabulary - tens of thousands of words
- Out-of-Vocabulary- Mapping a word from the vocabulary into the unknown word

Apart from the above characteristics, the environment variability, channel variability, speaking style, sex, age, speed of speech also makes the ASR system more complex. But the efficient ASR systems must cope with the variability in the signal.

III. GROWTH OF ASR SYSTEMS

Building a speech recognition system becomes very much complex because of the criterion mentioned in the previous section. Even though speech recognition technology has advanced to the point where it is used by millions of individuals for using variety of applications. The research is now focusing on ASR systems that incorporate three features: large vocabularies, continuous speech capabilities, and speaker independence. Today, there are various systems which incorporate these combinations. However, with these numerous technological barriers in developing ASR system, still it has reached the highest growth. The milestone of ASR system is given in the following table 1.

TABLE 1. GROWTH OF ASR SYSTEM

Year	Progress of ASR System
1952	Digit Recognizer
1976	1000 word connected recognizer with constrained grammar
1980	1000 word LSM recognizer (separate words w/o grammar)
1988	Phonetic typewriter
1993	Read texts (WSJ news)
1998	Broadcast news, telephone conversations
1998	Speech retrieval from broadcast news
2002	Rich transcription of meetings, Very Large Vocabulary, Limited Tasks, Controlled Environment
2004	Finnish online dictation, almost unlimited vocabulary based on morphemes
2006	Machine translation of broadcast speech
2008	Very Large Vocabulary, Limited Tasks, Arbitrary Environment
2009	Quick adaptation of synthesized voice by speech recognition (in a project where TKK participates in)
2011	Unlimited Vocabulary, Unlimited Tasks, Many Languages, Multilingual Systems for Multimodal Speech Enabled Devices
Future Direction	Real time recognition with 100% accuracy, all words that are intelligibly spoken by any person, independent of vocabulary size, noise, speaker characteristics or accent.

IV. OVERVIEW OF AUTOMATIC SPEECH RECOGNITION (ASR) SYSTEM

The task of ASR is to take an acoustic waveform as an input and produce output as a string of words. Basically, the problem of speech recognition can be stated as follows. When given with acoustic observation $X = X_1,X_2 \ldots X_n$, the goal is to find out the corresponding word sequence $W = W_{1,}W_2 \ldots W_m$ that has the maximum posterior probability P(W|X) expressed using Bayes theorem as shown in equation (1). The following figure 1 shows the overview of ASR system.

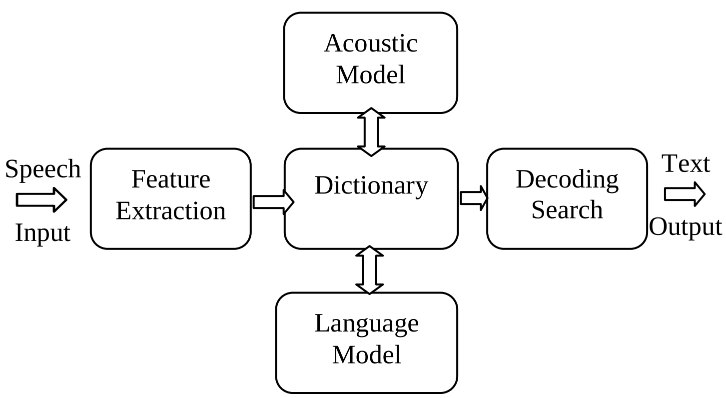

Figure 1.Overview of ASR system

$$W = \arg\max_{w} P(W/X) = \arg\max_{w} \frac{P(W)P(X/W)}{P(X)} \qquad (1)$$

Where P(W) is the probability of word W uttered and P(X|W) is the probability of acoustic observation of X when the word W is uttered.

In order to recognize speech, the system usually consists of two phases. They are called pre-processing and post-processing. Pre-processing involves feature extraction and the post-processing stage comprises of building a speech recognition engine. Speech recognition engine usually consists of knowledge about building an acoustic model, dictionary and grammar. Once all these details are given correctly, the recognition engine identifies the most likely match for the given input, and it returns the recognized word.

An essential task of developing any ASR system is to choose the suitable feature extraction technique and the recognition approach. The suitable feature extraction and recognition technique can produce good accuracy for the given application. Hence, these two major components are reviewed and compared based on its merits and demerits to find out the best technique for speech recognition system. The various types of feature extraction and speech recognition approaches are explained in the following section.

V. SPEECH FEATURE EXTRACTION TECHNIQUES

Feature Extraction is the most important part of speech recognition since it plays an important role to separate one speech from other. Because every speech has different individual characteristics embedded in utterances. These characteristics can be extracted from a wide range of feature extraction techniques proposed and successfully exploited for speech recognition task. But extracted feature should meet some criteria while dealing with the speech signal such as:

- Easy to measure extracted speech features
- It should not be susceptible to mimicry
- It should show little fluctuation from one speaking environment to another
- It should be stable over time
- It should occur frequently and naturally in speech

The most widely used feature extraction techniques are explained below.

A. *Linear Predictive Coding (LPC)*

One of the most powerful signal analysis techniques is the method of linear prediction. LPC [3][4] of speech has become the predominant technique for estimating the basic parameters of speech. It provides both an accurate estimate of the speech parameters and it is also an efficient computational model of speech. The basic idea behind LPC is that a speech sample can be approximated as a linear combination of past speech samples. Through minimizing the sum of squared differences (over a finite interval) between the actual speech samples and predicted values, a unique set of parameters or predictor coefficients can be determined. These coefficients form the basis for LPC of speech [10]. The analysis provides the capability for computing the linear prediction model of speech over time. The predictor coefficients are therefore transformed to a more robust set of parameters known as cepstral coefficients. The following figure 2 shows the steps involved in LPC feature extraction.

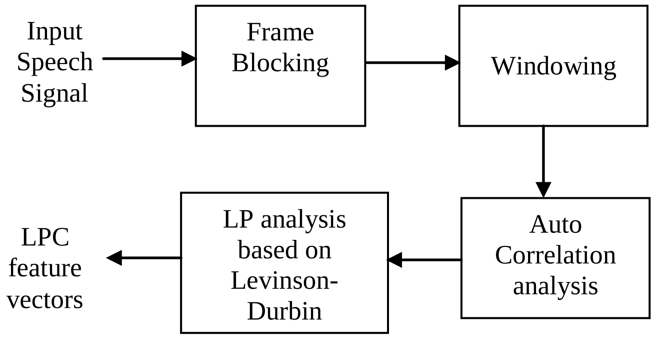

Figure 2. Steps involved in LPC Feature extraction

B. Mel Frequency Cepstral Coefficients (MFCC)

The MFCC [3] [4] is the most evident example of a feature set that is extensively used in speech recognition. As the frequency bands are positioned logarithmically in MFCC [6], it approximates the human system response more closely than any other system. Technique of computing MFCC is based on the short-term analysis, and thus from each frame a MFCC vector is computed. In order to extract the coefficients the speech sample is taken as the input and hamming window is applied to minimize the discontinuities of a signal. Then DFT will be used to generate the Mel filter bank. According to Mel frequency warping, the width of the triangular filters varies and so the log total energy in a critical band around the center frequency is included. After warping the numbers of coefficients are obtained. Finally the Inverse Discrete Fourier Transformer is used for the cepstral coefficients calculation [3] [4]. It transforms the log of the quefrench domain coefficients to the frequency domain where N is the length of the DFT. MFCC can be computed by using the formula (2).

$$Mel(f)= 2595*\log10(1+f/700) \qquad (2)$$

The following figure 3 shows the steps involved in MFCC feature extraction.

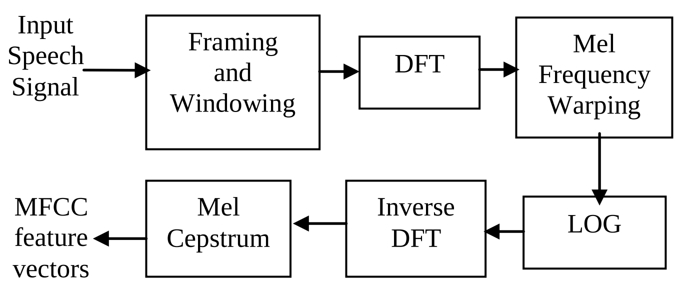

Figure 3. Steps involved in MFCC feature extraction

VI. SPEECH RECOGNITION APPROACHES

In the earlier years, dynamic programming techniques have been developed to solve the pattern-recognition problem [12]. Subsequent researches were based on Artificial Neural Network (ANN) techniques, in which the parallel computing found in biological neural systems is mimicked. More recently, stochastic modeling schemes have been incorporated to solve the speech recognition problem, such as the Hidden Markov Modeling (HMM) approach. At present, much of the recent researches on speech recognition involve recognizing continuous speech from a large vocabulary using HMMs, ANNs, or a hybrid form [12]. These techniques are briefly explained below.

A. Template-Based Approaches

Template based approaches to speech recognition have provided a family of techniques that have advanced the field considerably during the last two decades. The underlying idea of this approach is simple. It is a process of matching unknown speech is compared against a set of pre-recorded words (templates) in order to find the best match (Rabiner et al., 1979). This has the advantage of using perfectly accurate word models; but it also has the disadvantage that the pre-recorded templates are fixed, so variations in speech can only be modeled by using many templates per word, which eventually becomes impractical. Template preparation and matching become prohibitively expensive or impractical as vocabulary size increases beyond a few hundred words. This method was rather inefficient in terms of both required storage and processing power needed to perform the matching. Template matching was also heavily speaker dependent and continuous speech recognition was also impossible.

B. Knowledge-Based Approaches

The use of knowledge/rule based approach to speech recognition has been proposed by several researchers and applied to speech recognition (De Mori & Lam, 1986; Alikawa, 1986; Bulot & Nocera, 1989), speech understanding systems (De Mori and Kuhn, 1992). The “expert” knowledge about variations in speech is hand-coded into a system. It uses set of features from the speech, and then the training system generates set of production rules automatically from the samples. These rules are derived from the parameters that provide most information about a classification. The recognition is performed at the frame level, using an inference engine (Hom, 1991) to execute the decision tree and classify the firing of the rules. This has the advantage of explicitly modeling variations in speech; but unfortunately such expert knowledge is difficult to obtain and use successfully, so this approach was judged to be impractical, and automatic learning procedures were sought instead.

C. Neural Network-Based Approaches

Another approach in acoustic modeling is the use of neural networks. They are capable of solving much more complicated recognition tasks, but do not scale as excellent as Hidden Markov Model (HMM) when it comes to large vocabularies. Rather than being used in general-purpose speech recognition applications they can handle low quality, noisy data and speaker independence [7] [11]. Such systems can achieve greater accuracy than HMM based systems, as long as there is training data and the vocabulary is limited. A more general approach using neural networks is phoneme recognition. This is an active field of research, but generally the results are better than HMMs [7] [9]. There are also NN-HMM hybrid systems that use the neural network part for phoneme recognition and the HMM part for language modeling.

D. Dynamic Time Warping (DTW)-Based Approaches

Dynamic Time Warping is an algorithm for measuring similarity between two sequences which may vary in time or speed [8]. A well known application has been ASR, to cope with different speaking speeds. In general, it is a method that allows a computer to find an optimal match between two given sequences (e.g. time series) with certain restrictions, i.e. the sequences are "warped" non-linearly to match each other. This sequence alignment method is often used in the context of HMM. In general, DTW is a method that allows a computer to find an optimal match between two given sequences (e.g. time series) with certain restrictions. This technique is quite efficient for isolated word recognition and can be modified to recognize connected word also [8].

E. Statistical-Based Approaches

In this approach, variations in speech are modeled statistically (e.g., HMM), using automatic learning procedures. This approach represents the current state of the art. Modern general-purpose speech recognition systems are based on statistical acoustic and language models. Effective acoustic and language models for ASR in unrestricted domain require large amount of acoustic and linguistic data for parameter estimation. Processing of large amounts of training data is a key element in the development of an effective ASR technology nowadays. The main disadvantage of statistical models is that they must make a priori modeling assumptions, which are liable to be inaccurate, handicapping the system's performance.

Hidden Markov Model (HMM)-Based Speech Recognition

The reason why HMMs are popular is because they can be trained automatically and are simple and computationally feasible to use [2] [5]. HMMs to represent complete words can be easily constructed (using the pronunciation dictionary) from phone HMMs and word sequence probabilities added and complete network searched for best path corresponding to the optimal word sequence. HMMs are simple networks that can generate speech (sequences of cepstral vectors) using a number of states for each model and modeling the short-term spectra associated with each state with, usually, mixtures of multivariate Gaussian distributions (the state output distributions). The parameters of the model are the state transition probabilities and the means, variances and mixture weights that characterize the state output distributions. Each word, or each phoneme, will have a different output distribution; a HMM for a sequence of words or phonemes is made by concatenating the individual trained HMM [12] for the separate words and phonemes.

Current HMM-based large vocabulary speech recognition systems are often trained on hundreds of hours of acoustic data. The word sequence and a pronunciation dictionary and the HMM [6] [12] training process can automatically determine word and phone boundary information during training. This means that it is relatively straightforward to use large training corpora. It is the major advantage of HMM which will extremely reduce the time and complexity of recognition process for training large vocabulary.

VII. PERFORMANCE EVALUATION OF ASR TECHNIQUES

The performance of a speech recognition system is measurable. Perhaps the most widely used measurement is accuracy and speed. Accuracy is measured with the Word Error Rate (WER), whereas speed is measured with the real time factor. WER can be computed by the equation (3)

$$WER = \frac{S + D + I}{N} \qquad (3)$$

Where S is the number of substitutions, D is the number of the deletions, I is the number of the insertions and N is the number of words in the reference.

The speed of a speech recognition system is commonly measured in terms of Real Time Factor (RTF). It takes time *P* to process an input of duration *I*. It is defined by the formula (4)

$$RTF = \frac{P}{I} \qquad (4)$$

The comparison of the various speech recognition research based on the dataset, feature vectors, and speech recognition technique adopted for the particular language are given in the table 2.

TABLE 2. COMPARISION OF VARIOUS SPEECH RECOGNITION APPLICATIONS BASED ON DATASET, FEATURE EXTRACTION AND RECOGNITION APPROACH

Author	Year	Research Work	Nature of the Data	Feature Extraction Technique	Recognition Technique	Language	Accuracy
Meysam Mohamad pour, Fardad Farokhi		Spoken digit recognition	Isolated Digit	Discrete Wavelet Transform (DWT)	Multilayer Perceptron + UTA algorithm	English	98%

Ghulam Muhammad, Yousef A. Alotaibi, and Mohammad Nurul Huda	2009	Automatic Speech Recognition for Bangia Digits	Small vocabulary Speaker independent Isolated digit	Mel-Frequency Cepstral Coefficients (MFCCs)	Hidden Markov Model (HMM)	Bangia	more than 95% for digits (0-5) and less than 90% for digits (6-9)
Corneliu Octavian Dumitru, Inge Gavat	2006	A Comparative Study of Feature Extraction Methods Applied to Continuous Speech Recognition in Romanian Language	Large vocabulary Speaker independent Continuous speech	PLP, MFCC, LPC	Hidden Markov Models (HMM)	Romanian	MFCC-90,41%, LPC-63,55%. and PLP 75,78%
Douglas O'shaughnessy		Interacting With Computers by Voice: Automatic Speech Recognition and Synthesis		LPC	HMM	English	Good acuuracy
Sid-Ahmed Selouani, Yousef Ajami Alotaibi	2003	Investigating Automatic Recognition of Non-Native Arabic Speech	Large vocabulary Speaker independent Phonetic/word	MFCC	HMM	Arabic	New words makes less accuracy for non-native speakers
Vimal Krishnan V.R Athulya Jayakumar Babu Anto.P	2008	Speech Recognition of Isolated Malayalam Words Using Wavelet Features and Artificial Neural Network	Small vocabulary Speaker independent Isolated word	Discrete Wavelet Transform	Artificial Neural Network (ANN)	Malayalam	89%
Zhao Lishuang , Han Zhiyan	2010	Speech Recognition System Based on Integrating feature and HMM	Large vocabulary Speaker independent vowels	MFCC	Genetic Algorithm + HMM	Chinese	effective and high speed and accuracy
Bassam A. Q. Al-Qatab , Raja N. Ainon		Arabic Speech Recognition Using Hidden Markov Model Toolkit (HTK)		MFCC	HMM	Arabic	97.99%
Javed Ashraf , Dr Naveed Iqbal, Naveed Sarfraz Khattak, Ather Mohsin Zaidi		Speaker Independent Urdu Speech Recognition Using HMM	Small vocabulary Speaker independent Isolated word	MFCC	Hidden Markov Model	Urdu	Little variation in WER for new speakers
N.Uma Maheswari, A.P.Kabilan, R.Venkatesh		A Hybrid model of Neural Network Approach for Speaker independent Word Recognition		LPC	Hybrid model of Radial Basis Function and the Pattern Matching method	English	91%
Raji Sukumar.A Firoz Shah.A Babu Anto.P	2010	Isolated question words Recognition from speech queries by Using artificial neural networks	Medium vocabulary Speaker dependent Isolated word	DWT	ANN	Malayalam	80%

R. Thangarajan, A.M. Natarajan and M. Selvam		Phoneme Based Approach in Medium Vocabulary Continuous Speech Recognition in Tamil language	Medium vocabulary Speaker independent Continuous Speech	MFCC	Hidden Markov Model (HMM)	Tamil	good word accuracy for trained and test sentences read by trained and new speakers
A.Rathinavelu, G.Anupriya, A.S.Muthananthamurugavel	2007	Speech Recognition Model for Tamil Stops	Small vocabulary Speaker independent phonems	first five formant values	Feed forward neural networks	Tamil	81%
M. Chandrasekar, and M.Ponnavaikko	2008	Tamil speech recognition: a complete model	Medium vocabulary Speaker dependent Isolated Speech	MFCC	Back Propagation Network	Tamil	80.95%
A.P.Henry Charles1 & G.Devaraj2	2004	Alaigal-A Tamil Speech Recognition	Speaker independent Continuous speech	MFCC	HMM	Tamil	Offers High Performance

VIII. CONCLUSION

Speech recognition has been in development for more than 50 years, and has been entertained as an alternative access method for individuals with disabilities for almost as long. In this paper, the fundamentals of speech recognition are discussed and its recent progress is investigated. The various approaches available for developing an ASR system are clearly explained with its merits and demerits. The performance of the ASR system based on the adopted feature extraction technique and the speech recognition approach for the particular language is compared in this paper. In recent years, the need for speech recognition research based on large vocabulary speaker independent continuous speech has highly increased. Based on the review, the potent advantage of HMM approach along with MFCC features is more suitable for these requirements and offers good recognition result. These techniques will enable us to create increasingly powerful systems, deployable on a worldwide basis in future.

REFERENCES

[1] Bassam A. Q. Al-Qatab , Raja N. Ainon, “Arabic Speech Recognition Using Hidden Markov Model Toolkit(HTK)”, 978-1-4244-6716-711 0/$26.00 ©2010 IEEE.

[2] M. Chandrasekar, M. Ponnavaikko, “Tamil speech recognition: a complete model”, Electronic Journal «Technical Acoustics» 2008, 20.

[3] Corneliu Octavian DUMITRU, Inge GAVAT, “A Comparative Study of Feature Extraction Methods Applied to Continuous Speech Recognition in Romanian Language”, 48th International Symposium ELMAR-2006, 07-09 June 2006, Zadar, Croatia.

[4] DOUGLAS O’SHAUGHNESSY, “Interacting With Computers by Voice: Automatic Speech Recognition and Synthesis”, Proceedings of the IEEE, VOL. 91, NO. 9, September 2003, 0018-9219/03$17.00 © 2003 IEEE.

[5] Ghulam Muhammad, Yousef A. Alotaibi, and Mohammad Nurul Huda , “Automatic Speech Recognition for Bangia Digits”, Proceedings of 2009 12th International Conference on Computer and Information Technology (ICCIT2009) 21-23 December, 2009, Dhaka, Bangladesh, 978-1-4244-6284-1/09/$26.00 ©2009 IEEE.

[6] A.P.Henry Charles & G.Devaraj, “Alaigal-A Tamil Speech Recognition”, Tamil Internet 2004, Singapore.

[7] Meysam Mohamad pour, Fardad Farokhi, “An Advanced Method for Speech Recognition”, World Academy of Science, Engineering and Technology 49 2009.

[8] Santosh K.Gaikwad, Bharti W.Gawali and Pravin Yannawar, “A Review on Speech Recognition Technique”, International Journal of Computer Applications (0975 – 8887) Volume 10– No.3, November 2010.

[9] Raji Sukumar.A, Firoz Shah.A and Babu Anto.P, “Isolated question words recognition from speech queries by Using artificial neural networks”, 2010 Second International conference on Computing, Communication and Networking Technologies, 978-1-4244-6589-7/10/$26.00 ©2010 IEEE.

[10] N.Uma Maheswari, A.P.Kabilan, R.Venkatesh, “A Hybrid model of Neural Network Approach for Speaker independent Word Recognition”, International Journal of Computer Theory and Engineering, Vol.2, No.6, December, 2010 1793-8201.

[11] Vimal Krishnan V. R, Athulya Jayakumar and Babu Anto.P, “Speech Recognition of Isolated Malayalam Words Using Wavelet Features and Artificial Neural Network”, 4th IEEE International Symposium on Electronic Design, Test & Applications, 0-7695-3110-5/08 $25.00 © 2008 IEEE

[12] Zhao Lishuang , Han Zhiyan, “Speech Recognition System Based on Integrating feature and HMM”, 2010 International Conference on Measuring Technology and Mechatronics Automation, 978-0-7695-3962-1/10 $26.00 © 2010 IEEE.

AUTHORS PROFILE

Ms. Vimala.C, currently doing Ph.D and working as a Project Fellow for the UGC Major Research Project in the Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women. She has more than 2 years of teaching experience and 1 year of research experience. Her area of specialization includes Speech Recognition and Synthesis. She has 10 publications at National and International level conferences and journals. Email id: vimalac.au@gmail.com

Dr. V. Radha, is the Associate Professor of Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, TamilNadu, India. She has more than 21 years of teaching experience and 7 years of Research Experience. Her Area of Specialization includes Image Processing, Optimization Techniques, Voice Recognition and Synthesis, Speech and signal processing and RDBMS. She has more than 45 Publications at national and International level journals and conferences. She has one Major Research Project funded by UGC. Email id: radhasrimail@gmail.com.