International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

Speaker Recognition by Applying DWT Method with Different Speaker

Kanak Lata
M.Tech Scholar, AL-Falah University, Haryana, India
EMAIL: kanaklata27@gmail.com

Assit. Prof. Ravinder Kumar
al-falah university, haryana, india
email:**Ravi22aug@Rediffmail.Com**

Abstract

Speaker recognition is the computing task of validating a user's claimed identity using characteristics extracted from their voices. Voice -recognition is combination of the two where it uses learned aspects of a speaker's voice to determine what is being said - such a system cannot recognize speech from random speakers very accurately, but it can reach high accuracy for individual voices it has been trained with, which gives us various applications in day today life. The design of speech recognition system require careful attentions to the challenges or issue such as definition of various types of speech classes, speech representation, feature extraction techniques, database and performance evaluation. This paper gives an overview of major technological perspective and appreciation of the fundamental study of speaker recognition and gives overview technique developed in each stage of speech recognition and also summarize and compare different speech recognition systems and identify research topics and applications which are at the forefront of this exciting and challenging field.

Key Words: Speech Recognition; Speaker recognition; LPC; HMM; DTW; Voice Testing

Introduction

Speech being a natural mode of communication for humans can provide a convenient interface to control devices. Some of the speech recognition applications require speaker-dependent isolated word recognition. Current implementations of speech recognizers have been done for personal computers and digital signal processors. However, some applications, which require a low-cost portable speech interface, cannot use a personal computer or digital signal processor based implementation on account of cost, portability and scalability. For instance, the control of a wheelchair by spoken commands or a speech interface for computer.

Spoken language interfaces to computers is a topic that has lured and fascinated engineers and speech scientists alike for over five decades. For many, the ability to converse freely with a machine represents the ultimate challenge to our understanding of the production and perception processes involved in human speech communication. In addition to being a provocative topic, spoken language interfaces are fast becoming a necessity. In the near future, interactive networks will provide easy access to a wealth of information and services that will fundamentally affect how people work, play and conduct their daily affairs. Today, such networks are limited to people who can read and have access to computers---a relatively small part of the population even in the most developed countries. Advances in human language technology are needed for the average citizen to communicate with networks using natural communication skills using everyday devices, such as telephones and televisions. Without fundamental advances in user centered interfaces, a large portion of society will be prevented from participating in the age of information, resulting in further stratification of society and tragic loss in human potential.

Types of Speech

Speech can be classified into following categories:

***A.*Isolated words**

Isolated word recognizers accept single word at a time. These systems have "Listen/Not-Listen" states, where they require the speaker to wait between utterances. Isolated Utterance might be a better name for this class.

***B.*Connected words**

Connected word speech recognition is the system where the words are separated by pauses. Connected word speech recognition is a class of fluent speech strings where the set of strings is derived from small-to-moderate size vocabulary such as digit strings, spelled letter sequences, combination of alphanumeric. Like isolated word speech recognition, this set too has a property that the basic speech-recognition unit is the word/phrase to much extent.

***C.*Continuous speech**

Continuous speech recognition deals with the speech where words are connected together instead of being separated by pauses. As a result unknown boundary information about words, co-articulation, production of surrounding phonemes and rate of speech effect the performance of continuous speech recognition systems. Recognizers with continuous speech capabilities are some of the most difficult to create because they utilize special methods to determine utterance boundaries.

Principles of Speaker Recognition

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.

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.

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.

It’s a one-to-many comparison [3]. The basic structure of SI system (SIS) is shown in 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.

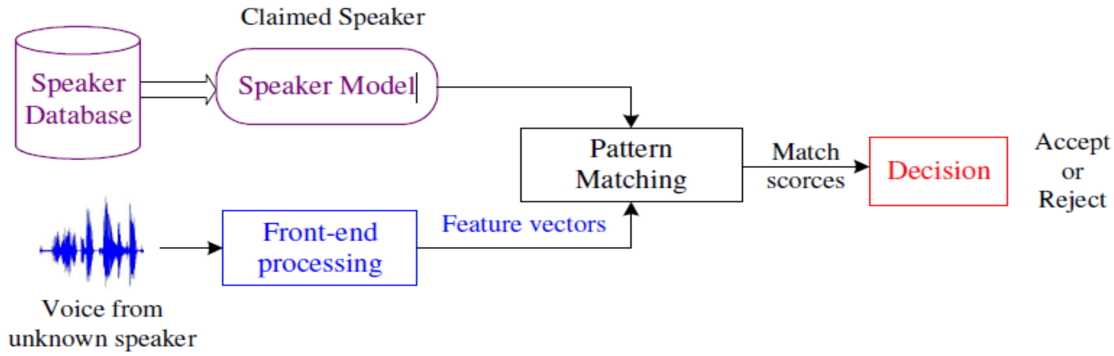

Figure 1: Basic structure of Speaker Verification

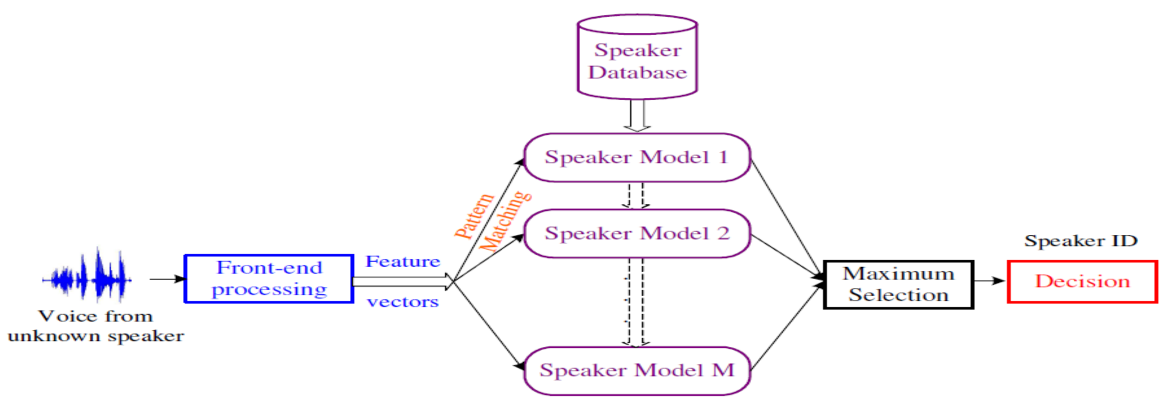

Figure 2: Basic structure of Speaker Identification

Speaker recognition is usually a general name referring to two different subtasks: speaker identification (SI) and speaker verification (SV).

Our research primly concentrates on the identification task. The aim in SI is to recognize the unknown speaker from a set of known speakers (closed-set SI).

Referring to the Figure 1 below, we can see that a speaker recognition system is composed of the following modules:

Front-end processing - the "signal processing" part, which converts the sampled speech signal into set of feature vectors, which characterize the properties of speech that can separate different speakers. Front-end processing is performed both in training- and recognition phases.

Speaker modelling - this part performs a reduction of feature data by modelling the distributions of the feature vectors.

Speaker database - the speaker models are stored here.

Decision logic - makes the final decision about the identity of the speaker by comparing unknown feature vectors to all models in the database and selecting the best matching model.

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

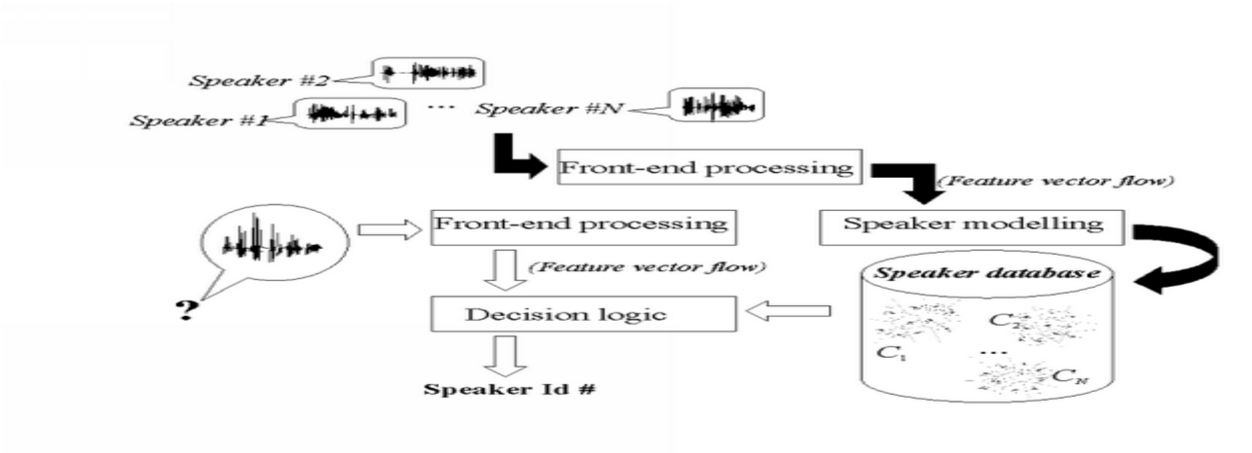

Figure 3: Schematic diagram of the closed-set speaker identification system

THE GOAL OF OUR RESEARCH:

Algorithm development:

*Feature extraction

*Speaker modeling using clustering and VQ

*Discriminative matching

*Real-time issues, optimization Applying speaker recognition technology in real world.

SIMULATION AND DISCUSSION

Step 1: Run the speaker recognition code in Matlab.

Step 2: After run the code it shows Speaker identification system. (See figure 4)

Step 3: Now select One-to-One Speaker Recognition Verification

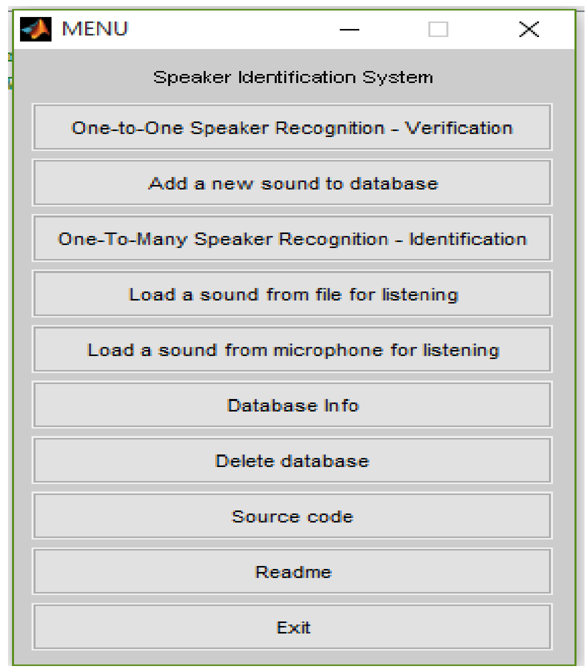

Figure 4: Speaker identification system

Step 4: after selection we have two option to select audio from Disk and microphone.

Step 5: If I select audio from Disk then I take an audio and after that I have same option to select audio again if we select audio from Disk then it shows recognized because I select both audio from same speaker1.See figure 5

International Journal of Research
IJR
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

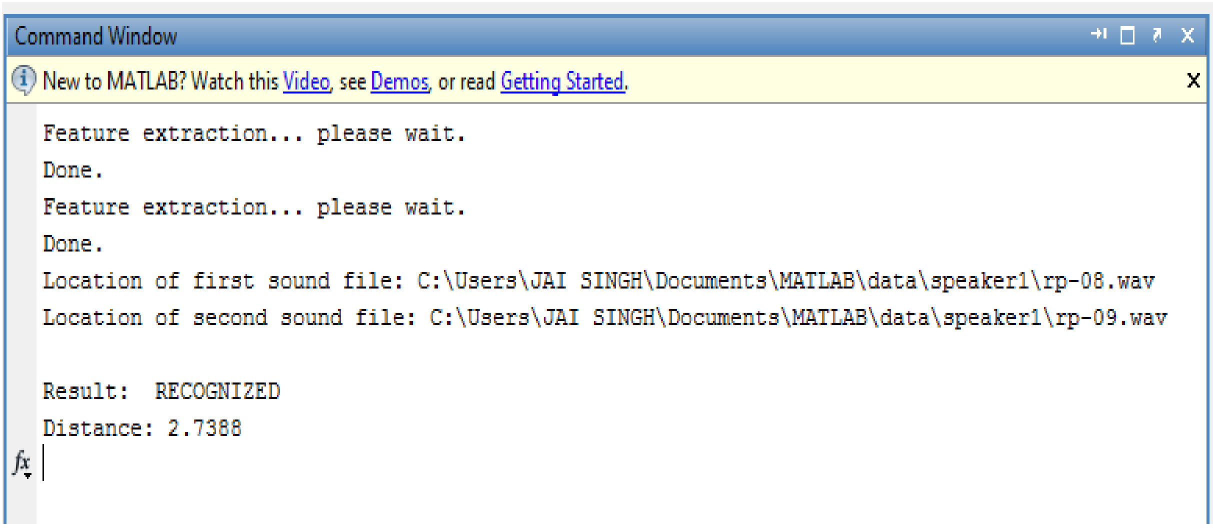

Figure 5: Speaker Recognition Verification when recognized

If I select audio from different speaker then it shows UNRECOGNIZED. (See figure 6)

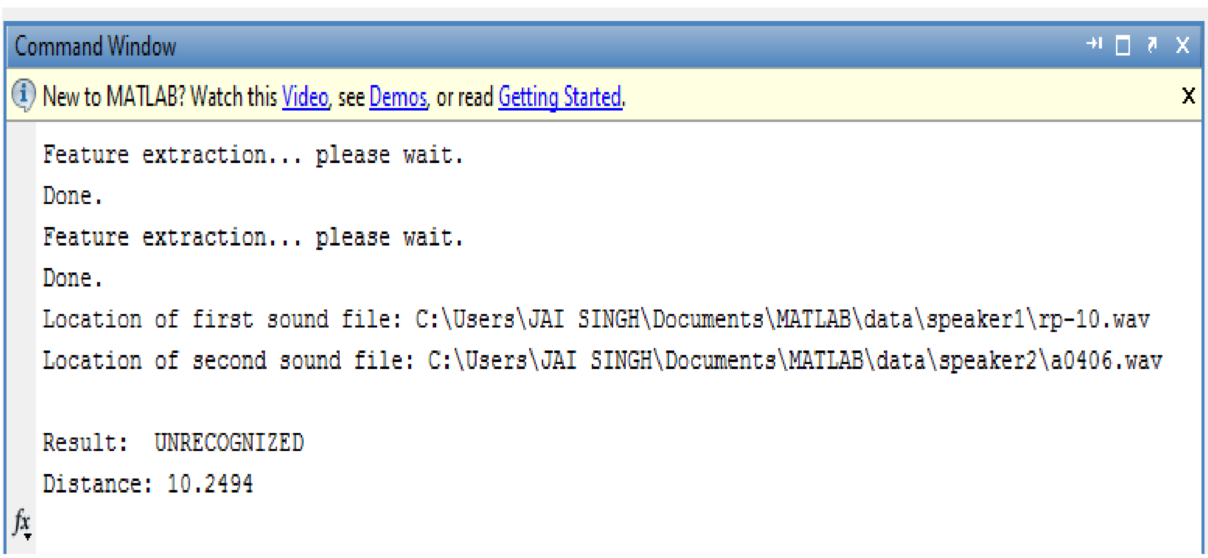

Figure 6: Speaker Recognition Verification when unrecognized

Conclusion

In this work, we studied and analyzed different techniques for speaker identification. In the first part, we started from the identification background, which is based on the digital signal theory and modeling of the speaker vocal tract. Then we discussed various techniques for reducing amount of test dataor feature extraction. Further, we studied most popular speaker modelingmethods, which are commonly used in the speaker identification. In the second part, we studied techniques, discussed in the previous part, from the real-time systems point of view. We proposed different optimization approaches to the speaker identification. However, we discussed only methods related to the speaker identification area, and left out from discussion general optimization methods.

We proposed a speaker pruning as a novel approach to reducing amount of distance calculations in the matching step. This method is heuristic, and therefore, improves identification speed at the cost of increasing of the probability of incorrect identification. We proposed two variations of the pruning algorithm and made approximate time complexity analysis for this methods and concluded that it significantly improves matching step. Finally, we studied speaker pruning empirically and found out that theory analysis was correct and it really improves identification speed. We also compared different parameter combinations for both variants of speaker pruning.

By seeing Distance different in both different speaker we can know the difference of speaker recognition very well.

References:

[1.] Bassam A. Q. Al-Qatab, Raja N. Ainon, "Arabic Speech Recognition Using Hidden Markov Model Toolkit(HTK)", 978-1- 4244-6716-7110/$26.00©2010 IEEE..

[2.] Dat Tat Tran, "Fuzzy Approaches to Speech and Speaker Recognition", a thesis submitted for the degree of PhD of the University of Canberra.

[3.] R.Klevansand, R.Rodman, "voice Recognition, Artech House, Boston, London 1997.

[4.] B. S. Atal, "Automatic Recognition of Speakers from their Voices", Proceedings of the IEEE, vol 64, 1976, pp 460 – 475..

[5.] R.Stapert, J. Mason, "A Segmental Mixture Model for Speaker Recognition",Eurospeech 2001 – Scandinavia

[6.] J. M.Naik, "Speaker Verification: A Tutorial", IEEE Communications Magazine, January 1990, pp.42-48.

[7.] Barry, P. Hedelin and J. Skoglund, "Vector quantization based on Gaussian mixturemodels", IEEE Transactions on Speech and Audio Processing, Vol. 8, No 4, July2000, pp. 385-401.

[8.] M. C. Huggins, J. J. Grieco, "Confidence Metrics for Speaker Identification",ICSLIP 2002, Denver, pp. 1381-1384