Recognition Techniques and System Classification

Arvind Gupta [*]

Govt. M.V.M College Bhopal

Praveen Shrivastava **

Sadhu vaswani College Bhopal

1. Abstract

The voice is most primary mode of Communication among of human being. The communication among human computer interaction is called human computer interface. Voice potential of being important of interaction with computer .This paper gives an overview of major technological perspective and appreciation of the fundamental progress of recognition and also gives overview technique developed in each stage of recognition. This paper helps in choosing the technique along with their relative merits & demerits. A comparative study of different technique is done as per stages. This paper is concludes with the decision on feature direction for developing technique in human computer interface system using Hindi Language.

2. Introduction

Automatic speech recognition has been studied for many years, having as a goal the man – computer dialog. Nowadays, the communication with the computer can be done using a speech recognition and voice synthesis system. Speech recognition is a process of automatic drawing and description of linguistic information contained in the voice and it can be done using computers. speech recognition can be extended to the person, making the speaker recognition. The speech recognition is performed comparing the input voice signal with signals stored. Various parameters are extracted from the voice signal and the comparison is done based on various mathematical methods. There are two main types of recognition, namely isolated word recognition and continuous speech recognition. Also, to achieve advanced recognition systems you need to know if you use more speakers or speaker is always the same person. Independent speaker recognition systems are more complex because they must work adaptive, changing parameters with the change of the speaker. This paper aims to address technically and theoretically an effective example of application for registration in real-time voice processing and examination. Thus, it gives a series of theoretical and practical views on how the audio signal processing with description of possible applications can be developed using digital and processing systems.

3. SPEECH RECOGNITION TECHNIQUES

The goal of speech recognition is for a machine to be able to "hear," understand," and "act upon" spoken information. The earliest speech recognition systems were first attempted in the early 1950s at Bell Laboratories, Davis, Biddulph and Balashek developed an isolated digit Recognition system for a single speaker [1]. The goal of automatic speaker reorganization is to analyze, extract characterize and recognize information about the speaker identity. The speaker reorganization system may be viewed
as working in a four stages
1. Analysis
2. Feature extraction
3. Modeling
4. Testing

4 Speaker recognition and identification:-

Speaker recognition is basically divided into two parts: recognition and identification. This is a way to automatically identify who is the speaker on the basis of individual information included in speech. The main goal of this project is to identify the speaker from a list with reference speaker models. The algorithm has to compare a voice signal from an unknown speaker with a database consisting of known speakers. The system that has been previously trained with a number of speakers can recognize the unknown speaker In the figure below is presented the fundamental process of speaker identification. In most applications, the voice is used to confirm the identity of a speaker
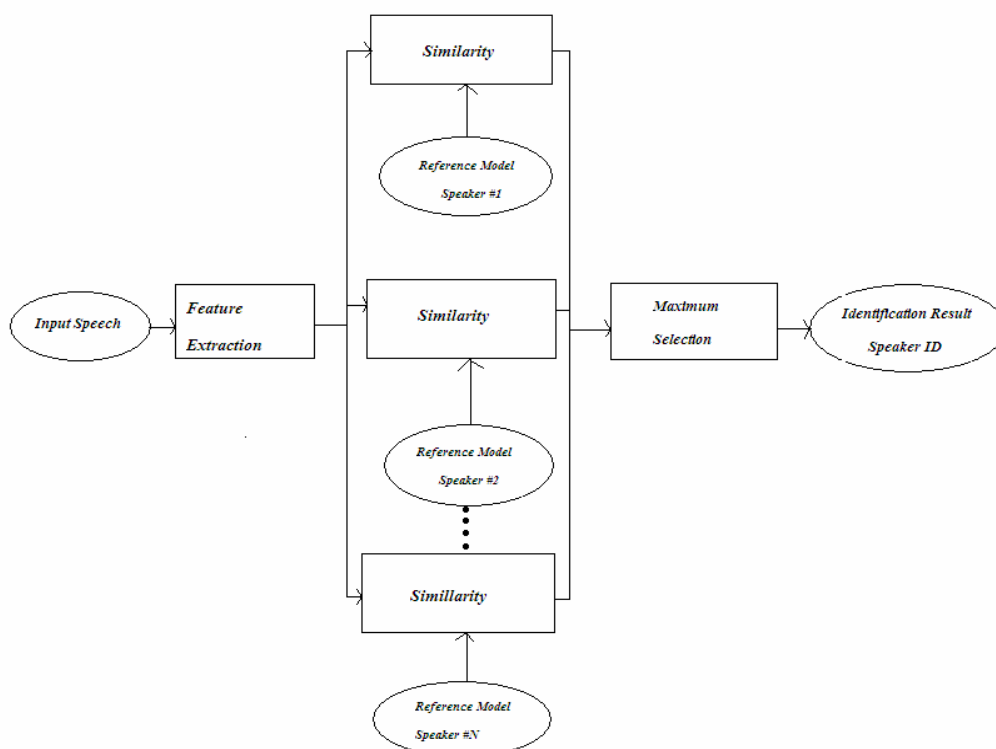
Figure 1.The fundamental process of speaker identification

4.1 Speech analysis technique
Speech data contain different type of information that shows a speaker identity. This includes speaker specific information due to vocal tract, excitation source and behavior feature. The information about the behavior feature also embedded in signal and that can be used for speaker recognition. The speech analysis stage deals with stage with suitable frame size for segmenting speech signal for further analysis and extracting [7]. The speech analysis technique done with following three techniques.

4.1.1 Segmentation analysis
In this case speech is analyzed using the frame size and shift in the range of 10-30 ms to extract speaker information. Studied made in used segmented analysis to extract vocal tract information of speaker recognition.

4.1.2 Sub segmental analysis
Speech analyzed using the frame size and shift in range is known as Sub segmental analysis. This technique is used to mainly analyze and extract the characteristic of the excitation state [8].

4.1.3 Performance of System
The performance of speaker recognisation system depends on the technique employed in the various stages of speaker recognisation system. The state of art of speaker recognisation system mainly used segmental analysis, Mel frequency Spectral coefficients (MFFCs), Gaussian mixture model (GMM) and feature extraction, modeling and testing stage. There are practical issues in the speaker recognisation field other technique may also have to be used for resulting a good speaker recognisation performance some of practical issues

4.1.4.1. A universal background Model (UBM) is a model used in a speaker verification system to represent general person independent the feature characteristics to be compared against a model of person specific feature characteristics when making an accept or reject decision [2].

4.3 Modeling Technique
The objective of modeling technique is to generate speaker models using speaker specific feature vector. The speaker modeling technique divided into two classification speaker recognition and speaker identification. The speaker identification technique automatically identify who is speaking on basis of individual information integrated in speech signal The speaker reorganization is also divided into two parts that means speaker dependant and speaker independent. In the speaker independent mode of the speech reorganization the computer

should ignore the speaker specific characteristics of the speech signal and extract the intended message .on the other hand in case of speaker reorganization machine should extract speaker characteristics in the acoustic signal [13]. The main aim of speaker identification is comparing a speech signal from an unknown speaker to a database of known speaker .The system can recognize the speaker, which has been trained with a number of speakers. Speaker recognition can also be divide into two methods, text-dependent and text independent methods. In text dependent method the speaker say   key words or sentences having the same text for both training and recognition trials. Whereas text independent does not rely on a specific texts being spoken [14]. Following are the modeling which can be used in speech recognition process:

6. REFERENCES

[1]   R.Klevansand R.Rodman, "Voice Recognition, Artech House, Boston, London 1997.

[2]   Samudravijaya K. Speech and Speaker recognition tutorial TIFR Mumbai 400005

[3]   Kevin Brady, Michael Brandstein, Thomas Quatieri, Bob Dunn "An Evaluation Of Audio-Visual person Recognition on the XM2VTS corpus using the Lausanne protocol" MIT Lincoln Laboratory, 244 Wood St., Lexington MA

[4]   W. M. Campbell_, D. E. Sturim W. Shen D. A. Reynolds, J. Navr´atily "The MIT- LL/IBM Speaker recognition System using High performance reduced Complexity recognition" MIT Lincoln Laboratory IBM 2006.

[5]   Zahi N.Karam,William M.Campbell "A new Kernel for SVM MIIR based Speaker recognition "MIT Lincoln Laboratory, Lexington, MA, USA. [6] Asghar .Taheri ,Mohammad Reza Trihi et.al,Fuzzy Hidden Markov Models for speech recognition on based FEM Algorithm, Transaction on engineering Computing and Technology V4 February 2005,IISN,1305-5313.

[7]   GIN-DER WU AND YING LEI " A Register Array based Low power FFT Processor for speech recognition" Department of Electrical engineering national Chi Nan university Puli ,545 Taiwan [8] Nicolás Morales1, John H. L. Hansen2 and Doorstep T. Toledano1 "MFCC Compensation for improved recognition filtered and band limited speech" Center for Spoken Language Research, University of Colorado at Boulder, Boulder (CO), USA

[9]   M.A.Anusuya , S.K.Katti "Speech Recognition by Machine: A Review" International journal of computer science and Information Security 2009.

[10]   Goutam Saha, Ulla S. Yadhunandan " Modifield Mel- Frequency Cepstral coefficient Department of Electronics and Electrical communication Engineering India Institute of Technology ,Kharagpur Kharagpur-721302 West Bengal,India. .

[11]   Kenneth Thomas Schutte "Parts-based Models and Local Features for Automatic Speech Recognition" B.S., University of Illinois at Urbana-Champaign (2001) S.M., Massachusetts Institute of Technology (2003).

[12]   Zaidi Razak, Noor Jamaliah Ibrahim, Emran Mohd Tamil, Mohd Yamani Idna Idris "Quarnic Verse recitation feature extraction using Mel-Frequency Cepstral Coefficient(MFCC)" Department of Al-Quran & Al-Hadith, AcademyOf Islamic Studies, University of Malaya .

[13]   Samudravijay K "Speech and Speaker recognition report" source: Viewed on 23 Feb. 2010.

[14]   Sannella, M Speaker recognition Project Report report" From

[15]   IBM (2010) online IBM Research Source:-http://www.research.ibm.com/Viewed 12 Jan 2010.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/   All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Academic conference: http://www.iiste.org/conference/upcoming-conferences-call-for-paper/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar