Multilingual Person Identification

Praveen Singh Rathore[#1], Dr. Neeta Tripathi[*2]

 # ME Scholar,ET&T,SSCET(FET),SSGI,CSVTU,Bhilai,India
* Principal, S.S.I.T.M, SSGI,CSVTU, Bhilai, India

Abstract— **Speech conveys the word being spoken and information about the speakers. The speaker recognition is divided into two parts speaker identification and speaker verification. Present paper explores the idea to identify multi-lingual person by basic features. In the present approach the speech signals are recorded and basic features pitch and formant frequency has been calculated. The neural network approach is use for training of system. Here we consider two languages Hindi and Marathi including male and female. We base our approach on multilingual person identification using basic features.**

Keywords—**Pitch, formant frequency, multi-lingual identification and neural network.**

I. INTRODUCTION

Person Identification and verification is an actively growing area of research. Speaker recognition is one of the biometric methods to identify the person by features of the voice. The person identification is the task of determining an unknown speaker's identity. There are two levels of parameters present in speech signal. The low level parameter includes pitch, formant frequencies and the high level parameters include emotional level of speaker, speaking style etc. The person can be identified by their voices. Prosody features pitch and formant frequency are used here for the person identification. The speaker identification system for mono-lingual person cannot be use for the multilingual person identification. Now a day the people speaks more than one language. Country like India, multi-lingual person identification systems is required where person speaks more than one language. In our work we have considered two languages Hindi and Marathi including male and female. The training of the person identification system is done by neural network approach.

BASIC FEATURE

The speech feature extraction is a fundamental requirement of any speaker recognition system. The feature extraction means estimating a set of features from the speech signal that represent some speaker-specific information. Feature extraction means converting a speech signal to some type of parametric representation for analysis. There are several features we can extract from the speech signal. Here we are extracting the basic features that are pitch and first three formant frequency. There are several methods to calculate the pitch of voice. Here we are using PRAAT software for finding basic features.

II. DATABASE FOR SPEAKER IDENTIFICATION

A database has been prepared by recording sentences in two languages Hindi and Marathi. This database comprises of pitch and first three formant frequencies for the letter 'Cha' and 'Sha' in both languages. We have included 92 speakers.

FEATURE EXTRACTION

Proposed system is implemented to identify the identity of a multilingual person. Here we are considering two languages Hindi and Marathi. The first step of any speech or speaker recognition system is recording of the voice of speaker. To record the voice samples we have used microphone. The recording and segmentation is done by Gold wave software. The features pitch and first three formant frequencies are extracted through the PRAAT software. After extracting the different features the database has been prepared. This database has been statistically analysed and we found a set of rules. This set of rules can be implemented with the help of neural network approach. The system is mainly using feed forward neural networks with back propagation.

MULTILINGUAL SPEAKER IDENTIFICATION

A multi-lingual speaker identification system has been constructed. The basic building block needed is recording of speech in two languages, feature extraction, training through neural network and decision.
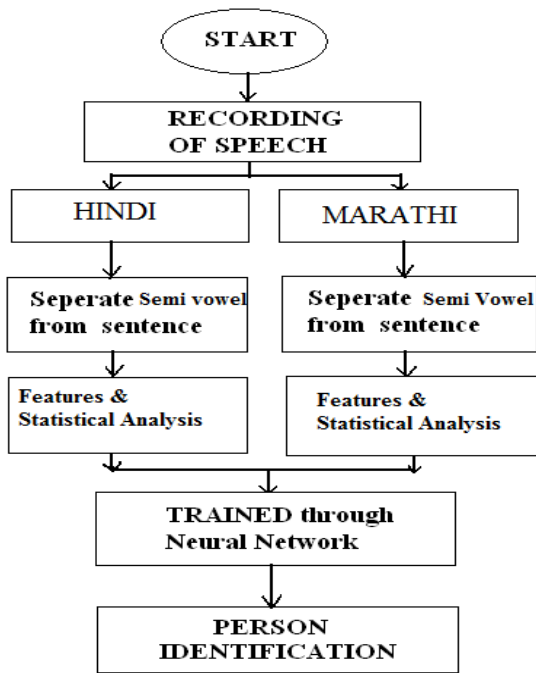
Figure 1 Multilingual Speaker Identification flow diagram

First of all the speakers voice (sentence) has been recorded in two languages which is having Cha and Sha letters in it. The letters Cha and Sha then extracted from the words. Then the basic features pitch and first three formant frequencies have been calculated. The statistical analysis has been done. These features are used to train a neural network.

III. RESULT

In this paper, investigation has been made for two basic features pitch and formant frequency for male and female speakers in two languages Hindi and Marathi. The analysis is done for the letter CHA and SHA in two languages. The following observations are recorded.

Figure 2 indicates the percentage change in Pitch from Hindi to Marathi language. The pitch of speaker in Hindi language is less as compare to Marathi language while speaking the letter "CHA". Hence the percentage change in pitch for Hindi to Marathi language is negative.
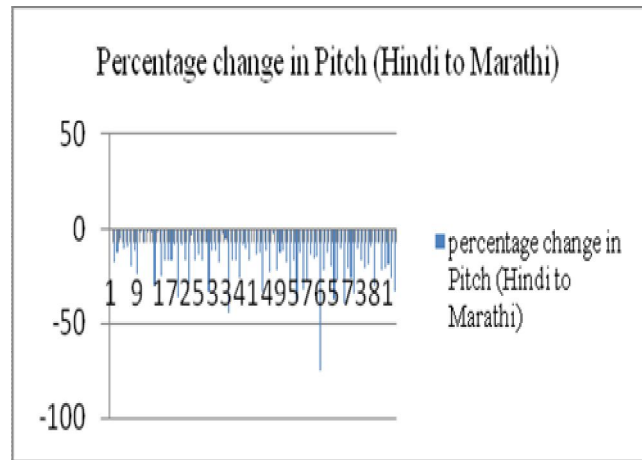

Figure 2 Percentage change in Pitch of speakers (Hindi to Marathi)

Figure 3 indicates the comparative study of pitch and formant frequency of speakers. It indicates that 16 male and 7 female speakers having negative pitch difference but positive formant frequency difference. It also shows that the speakers are having negative percentage changes of pitch and the positive percentage change for formant frequency (third formant frequency).
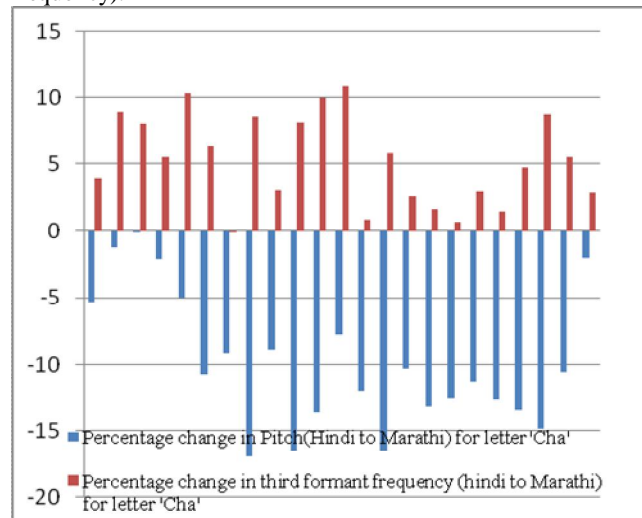

Figure3 Percentage variations in pitch and formant frequency of speakers for Hindi to Marathi

Figure 4 gives the percentage change for Pitch (CHA) is negative and percentage change for Pitch (SHA) is positive, and formant frequency percentage change is also positive for all the cases.
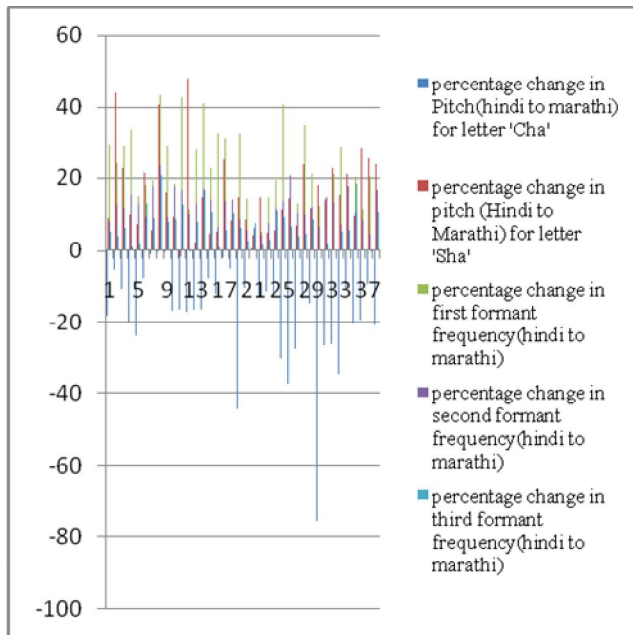
Figure 4 Pitch and formant frequency percentage change from Hindi to Marathi

The basic features can thus be use to train the neural network. Here we have trained two separate neural networks. The software used here are MATLAB, Goldwave and PRAAT. After training the network we get the results as identity of speakers, language and gender. The result The network trained is for close set of data of 92 speakers including male and female.

IV. CONCLUSION

We have developed a system which combines the pitch and formant frequency to identify multilingual speaker. In our system two neural networks have been trained. We proposed an approach for multilingual speaker identification to extract the prosody features of speech signal in two languages Hindi and Marathi, and then all the related features are used to train the system with the help of neural network. This work includes the letter 'Cha' and 'Sha'. We recorded 92 speaker voices including male and female. The development process includes feature extraction; training through feed forward neural network approach, and the results is the speaker's identity, gender and language.

REFERENCES

[1]   B. S. Atal, "Automatic Recognition of Speakers from Their Voices," Proceedings of the IEEE vol. 64, pp. 460-475, 1976.

[2]   G. R. Doddington, "Speaker Recognition-Identifying People by their Voices," Proceedings of the IEEE, vol. 73, No. 11, pp 1651-1664, 1985.

[3]   J. Campbell, "Speaker recognition: A tutorial", Proceedings of the IEEE, vol.85, pp 1437–1462, 1997.

[4]   Lawrence Rabiner and Biing-Hwang Juang, "Fundamental of Speech Recognition", Prentice-Hall, Englewood Cliffs, N.J., 1993.

[5]   U.Shrawankar, V.Thakre, "Feature extraction for a speech recognition system in noisy environment: A study", ICCEA, vol.1, pp 358-361,March 2010.

[6]   Utpal Bhattacharjee and Kshirod Sarmah, "Speaker verification using acoustic and prosodic features Advance Computing International Journal (ACIJ), Vol.4, No.1, January 2013.

[7]   E.E. Shriberg, "Higher Level Features in Speaker Recognition", Springer, vol 43. Computer science artificial intelligence, pp 241-259, 2007.

[8]   Herve Bourlard, John Dines, Mathew Magimai, Doss, "Current trends in multilingual speech processing", Sadhana, vol. 36, Part 5, pp. 885-915,October 2011.

[9]   Gurpreet kaur,Harjeet kour, "Multi Lingual Speaker Identification on Foreign Languages Using Artificial Neural Network with Clustering". International general of Computer Science and Software Engineering, volume 3, Issue 5, May 2013.

[10] Suzan A. Mahmood, Loay E. George, "Speaker Identification using back propagation neural network". (JZS) Journal of Zankoy Sulaimani, vol.10(1), pp 61-66, December 2008.

[11] Sanjay Decate, Anupam Shukla,"Speaker Recognition Based on Multilingual Speech Features using Neural Network Models",Oriental COCOSDA,2010.

[12] Lawrence R. Rabiner, Michael J. cheng, Aaron Rosenberg, Carol A. Mcgonegal, "A comparative performance study of several pitch detection algorithms" IEEE transactions on acoustics, speech, and signal processing, vol. assp-24,no. 5, pp399-417,October 1976.

[13] Biljana Prica and Sinisa Ili, "Recognition of Vowels in Continuous Speech by Using Formants,ELEC. ENERG. vol. 23, no. 3, pp379-393, December 2010.

[14] Hui Lin, Jui ting Huang, Yun hsuan Sung, "Recognition of multilingual speech in mobile applications".Acoustic speech and signal processing(ICASSP),pp 4881-4884,2012.