Abstract

Speech is the output of a time varying excitation excited by a time varying system. It generates pulses with fundamental frequency F0. This time varying impulse trained as one of the features, characterized by fundamental frequency F0 and its formant frequencies. These features vary from one speaker to another speaker and from gender to gender also. In this paper the effect of gender on improving speech recognition is considered. Variation in F0 and formant frequencies is the main features that characterize variation in a speaker. The variation becomes very less within speaker, medium within the same gender and very high among different genders. This variation in information can be exploited to recognize gender type and to improve performance of speech recognition system through modeling separate models based on gender type information. Five sentences are selected for training. Each of the sentences are spoken and recorded by 20 female’s speakers and 20 male speakers. The speech corpus wills be
preprocessed to identify the voiced and unvoiced region. The voiced region is the only region which carries information about F0. From each voiced segment, F0 and the first three formant frequencies and also MFCC features are computed. Each forms the feature space labeled with the speaker identification: i.e., male or female. This information is used to parameterize the model for male and female. K-means algorithm is used during training as well as testing. Testing is conducted in two ways: speaker dependent testing and speaker independent testing. SPHINX-III software by Carnegie Mellon University has been used to measure the accuracy of speech recognition of data taking into account the case of gender separation which has been used in this research.

References

1. Breazeal, C. and Aryanda, L. (2000), 'Recognition of affective communicative intent in robot-directed speech,' in 'Proceedings of Humanoids2000'.
2. http://www.ece.auckland.ac.nz/p4p_2005/archive/reports2003/pdfs/p60_hlai015.pdf.
3. S. Davis and P. Mermelstein. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Transactions on Acoustics Speech and Signal Processing, 28:357–366, Aug1980.
4. Parris, E.S., Carey, M.I., Language Independent Gender Identification, Proceedings of IEEEICASSP, pp685-688, 1996.
5. Linde, Y., A. Buzo, and R.M. Gray, "An Algorithm for Vector Quantizer Design," IEEE Trans. on Communication, 1980, COM-28(1), pp.84-95.
6. Hartigan, J.A., Clustering Algorithm, 1975, New York, J. Wiley.
7. Gersho, A., "On the Structure of Vector Quantization," IEEE Trans. On Information Theory, 1982, IT-28, pp. 256-261.
8. Richard P. Lappmann, Speech recognition by Machines and Humans, SPEECH Comm., pp. 1-15, 1997.
9. Santosh K. Gaikwad, Bharti W. Gawali and Pravin Yannawar, A Review on speech recognition technique, International Journal of Computer Application, Vol. 10(3), pp. 16-24, 2010.
10. M. Prabha, P. Viveka and Bharathasreeja, Advanced gender recognition system using speech signal, IJSET, Vol.6(4), pp. 118-120, 2016.
11. Chetana Prakash and Suryakanth V Gangasetty, Fourier- Bessel based cepstral coefficient features for text-independent speaker identification, IICA, pp.913-930, 2-11.
12. MusaedAlhussein, Zalfiqr Ali, Muhammad Imran and Wadood Abdul, Automatic gender detection based on characteristics of vocal folds for mobile healthcare system, Hindawi, pp. 1-12, 2016.
13. Suma Swamy and K. V Ramakrishnan, An efficient speech recognition system, CSEIJ, Vol3(4), pp. 21-27, 2013.
14. Preeti Saini and ParneetKaur, Automatic speech recognition: A review, IJETT, Vol (2), pp. 132-136, 2013.
15. Bhupinder Singh, Neha Kapur and Puneet Kaur, Speech recognition with Hidden Markov model: A review, IJARCSSE, Vol. 2(3), pp. 400-403, 2012.
16. M.A Anusuya and S.K Katti, Speech recognition by Machine: A review, IJCSIS, Vol6(3), pp. 181-205, 2009.
Effect of Gender on Improving Speech Recognition System

Index Terms

Computer Science

Pattern Recognition

Keywords

Speech recognition (SR), Linear Prediction Coding (LPC), Accent, Speakers.