Abstract

Emotion recognition from speech has emerged as an important research area in the recent past. The purpose of speech emotion recognition system is to automatically classify speaker’s utterances into seven emotional states including anger, boredom, disgust, fear, happiness, sadness and neutral. The speech samples are from Berlin emotional database and the features extracted from these utterances are Teager-based delta-spectral cepstral coefficients (T-DSCC) which are shown to perform better than MFCC. Dynamic Time Warping (DTW) and its variant Improved Features for DTW (IFDTW) is used as a classifier to classify different emotional states. Unlike in conventional DTW, we do not use the minimum distance for classification. Rather, the median distance criterion is employed for improved emotion
classification. The proposed emotion recognition system gives an overall classification accuracy of 97.52%.

References

- T. Pao, Y. Chen, J. Yeh, and P. Li, "Mandarin emotional speech recognition based on SVM and NN," Proc. 18th Intl. Conf. on Pattern Recognition (ICPRapos;06), vol. 1, pp. 1096-1100, Sep 2006
- S. Salvador, and P. Chan, "FastDTW: toward accurate dynamic time warping in linear time and space," Proc. 3rd KDD Workshop on Mining Temporal and Sequential Data, pp. 70-80, Aug 2004

Index Terms

Computer Science
Pattern Recognition

Keywords

Emotion Recognition  Mfcc  Dscc  Teager Energy Operator  Dynamic Time Warping