Abstract

In this paper an efficient implementation of speech to text converter for mobile application is presented. The prime motive of this work is to formulate a system which would give optimum performance in terms of complexity, accuracy, delay and memory requirements for mobile environment. The speech to text converter consists of two stages namely front-end analysis and pattern recognition. The proposed method uses effective methods for voice activity detection in preprocessing, feature extraction and recognizer. The energy of high frequency part is separately considered as zero crossing rate to differentiate noise from speech. RASTAPLP feature extraction method is used in which RASTA filter suppresses the spectral components that change more slowly or quickly than the typical range of change of speech thus avoiding unnecessary information in the extracted features. In the proposed system Generalized Regression Neural Network is used as recognizer in which syllable level recognition is used that reduces memory requirement and complexity for mobile application. Thus a small database containing all possible syllable pronunciation of the user is sufficient to give recognition accuracy closer to 100%. Reduction in 50% with respect to delay and memory requirement is proved in the proposed system. Thus the proposed technique entertains
realization of real time speaker dependant applications like mobile phones, PDAs etc.

References

Index Terms

Computer Science  Mobile Applications

Keywords

Preprocessing  Voice activity detection  RASTAPLP  Neural network  syllable based recognition