
In recent years wavelet packet (WP) transform has been used as an important speech representation tool. WP based acoustic features have found to be more effective than the short time Fourier transform (STFT) based features to capture the information of unvoiced phoneme in continuous speech. But wavelet features fail to carry the same usefulness to represent the voiced phonemes such as vowels, nasals. This paper proposes a new WP sub-band based features by taking care of harmonic information of voiced speech signal. It has been noticed that most of the voiced energy of the speech signal lies in between 250Hz-2000Hz. Thus the proposed technique emphasizes the individual sub-band harmonic energy upto 2 kHz. The speech signal is decomposed into 16 wavelet sub-bands and Harmonic energy features (HEF) are combined with wavelet packet cepstral features (WPCC). More in IET Signal Processing Digital Library (DOI: 10.1049/iet-spr.2014.0282)
Details

Admissible Wavelet Packet Sub-band based Harmonic Energy Features for Hindi Phoneme Recognition. Astik Biswas, P.K.Sahu, Anirban Bhowmick, Mahesh Chandra

Dept. of Electrical Engineering

In recent years wavelet packet (WP) transform has been used as an important speech representation tool. WP based acoustic features have found to be more effective than the short time Fourier transform (STFT) based features to capture the information of unvoiced phoneme in continuous speech. But wavelet features fail to carry the same usefulness to represent the voiced phonemes such as vowels, nasals. This paper proposes a new WP sub-band based features by taking care of harmonic information of voiced speech signal. It has been noticed that most of the voiced energy of the speech signal lies in between 250Hz-2000Hz. Thus the proposed technique emphasizes the individual sub-band harmonic energy upto 2 kHz. The speech signal is decomposed into 16 wavelet sub-bands and Harmonic energy features (HEF) are combined with wavelet packet cepstral features (WPCC) to enhance the performance of voiced phoneme recognizer. A standard phonetically balanced Hindi database is taken to analyze the performance of the proposed feature set. The noisy phoneme recognition task is also carried out to study the robustness. Significant improvement is obtained with the proposed feature set in voiced phoneme recognition over WPCC and conventional MFCC. More in IET Signal Processing Digital Library (DOI: 10.1049/iet-spr.2014.0282)

Link:-
http://digital-library.theiet.org/content/journals/10.1049/iet-spr.2014.0282