Abstract

Natural synthesis of speech needs to identify the minute variations in phoneme during reproduction, which is affected by many factors. One well-known problem with speech synthesis is the occurrence of audible discontinuities at phoneme boundaries, which lead to the unnaturalness of synthetic speech. This study basically focuses on introducing a novel method with low bit rate to improve the naturalness of synthetic speech.

The research presents a sinusoidal noise based mathematical method to reform the transition regions from one phoneme to another phoneme with lesser number of parameters. The speech information which are amplitude, phase and frequency were extracted using three different algorithms. They are Fast Fourier Transform (FFT) algorithm, Auto Regressive model (AR) with Linear Predictive Coding (LPC) algorithm and Auto Regressive Moving Average model (ARMA) with Steiglitz-McBride method. Polynomial coefficients were estimated to represent the speech information in lesser number of parameters. The results show that the synthesized output is highly correlated to the source signal in FFT method than AR model and ARMA model.
References


Index Terms

Computer Science
Signal Processing
Keywords

Fast Fourier Transform algorithm (FFT), Auto Regressive model (AR), Auto Regressive Moving Average model (ARMA), Speech Synthesis, Correlation Coefficient, Phoneme Transition