Abstract

Text to speech synthesis (TTS) is the production of artificial speech by a machine for the given text as input. The speech synthesis can be achieved by concatenation and Hidden Markov Model techniques. The voice synthesized by these techniques should be evaluated for quality. The study extends towards the comparative analysis for quality of speech synthesis using hidden markov model and unit selection approach. The quality of synthesized speech is analyzed for subjective measurement using mean opinion score and objective measurement based on mean square score and peak signal-to-noise ratio (PSNR). The quality is also accessed by Mel-frequency cepstral coefficient features for synthesized speech. The experimental analysis shows that unit selection method results in better synthesized voice than hidden markov model.

References

9. ITU-T P.830, Subjective performance assessment of telephone-band and wideband digital codecs
12. SR Quackenbush, TP Barnwell, MA Clements, Objective Measures of Speech Quality(Prentice-Hall, New York, NY, USA, 1988)
19. Sangramsing Kayte, Monica Mundada "Study of Marathi Phones for Synthesis of
Index Terms

Computer Science  Signal Processing

Keywords

Keyword TTS, MOS, HMM, Unit Selection, Mean, Variance.