

ABSTRACT

Nowadays, the use of audio as the application of real-time media is desperately needed, especially in media applications using the differentiation process on audio data, such as content based encoding and audio compression and equalization between speech and music automatically. Therefore, it is required an efficient algorithm to segment the audio signal into speech signal or music signal. In this final project, we use an approach for detecting music boundaries and classify speech / music by using an algorithm called segmental continuous dynamic programming or shortened by Segmental CDP.

Segmental CDP algorithm can be used to identify the location of each piece of music and their limits based on various similarity segment and location information.

Feature extraction in the time domain is given two options, namely: ZCR (Zero Crossing Rate) and Bit Energy, while the MFCC feature extraction is in the frequency domain. Separation in mixed signals successfully performed using the threshold of the traits. The audio signal is categorized as a speech signal if the value of moving average energy bit \leq maximum value of the moving average bit energy speech, a moving average value of ZCR \geq minimum value of ZCR speech moving average, moving average MFCC and value \leq maximum value of moving average MFCC speech. The audio signal is categorized as a music signal if, the value of moving average energy bit \geq minimum value of moving average energy bit music, moving average ZCR value \leq the maximum value of moving average ZCR music, and moving average MFCC value \leq minimum value of moving average MFCC music.

Key words: segmental CDP, speech, music, segmentation, classification