ABSTRACT

Concatenation synthesizer is a synthesizer that is capable of producing a signal through the transcription of speech automatically grapheme-to-phoneme to sentence spoken. Research concatenation synthesizer in various language versions are being developed, and in particular languages has achieved satisfactory results. However, some problems in concatenation synthesizer is still not completely solved. The approach of each of these issues will always refer to the concatenation synthesizer achievement that meets the quality intelligibility (sounds obvious) and naturalness (naturalness).

In this final project, has realized a new model of feature extraction for comparison of previous studies and also have realized a simple method for comparison with the test data HNN in execution. With this feature extraction of the average value of RMSE synthesis signal to the original signal to 15 sentences in the test increased from 0.349273 to 0.1919, but for the average correlation coefficient of -0.0060, which means the signal Sintesi contrary to the original signal. As for the simpler method is faster 1.63086 1.97588 dibandngkan use of HNN. While the MOS in the assessment, to integibillity increased from 2.666667 to 2.815385, which means more comprehensible voice clarity. But for fludity and naturalness is decreasing compared to previous studies.

Design and simulation modules feature extraction and the simple method, could be the basis dala improve the quality of synthesized speech.

Keywords: Hopfield neural network, generator prosodic, concatenation synthesizer