HMM Adaptation for child speech synthesis

Avashna Govender, Febe de Wet, Jules-Raymond Tapamo

Abstract

Hidden Markov Model (HMM)-based synthesis in combination with speaker adaptation has proven to be an approach that is well-suited for child speech synthesis. This paper describes the development and evaluation of different HMM-based child speech synthesis systems. The aim is to determine the most suitable combination of initial model and speaker adaptation techniques to synthesize child speech. The results of the study indicate that gender-independent initial models perform better than gender-dependent initial models and Constrained Structural Maximum a Posteriori Linear Regression (CSMAPLR) followed by maximum a posteriori (MAP) is the speaker adaptation technique combination that yields the most natural and intelligible synthesized child speech.