

Abstract

Gender-age classification based on voice is very useful in the introduction of speech and the emotion recognition. It has also been implemented in face recognition, video summarization, determining different permission levels for different age groups, and others. Grouping different ages are divided into three groups: child, young, middle, and senior based on a certain age range. This research focuses on the gender-age classification based on the speaker's voice using combined Gaussian Mixture Model and Hidden Markov Model (GMM-HMM). Firstly, a feature vector is built using Mel-Frequency Cepstrum Coefficient (MFCC). Secondly, a training process is performed to produce an acoustic model for all speakers (male and female from various ages) in the training database. Finally, HMM is applied to detect both gender and age group. In this research, the database of speech is taken from Common Voice website, which contains many blog posts, old books, films, and other public speeches. The experimental result shows that the developed GMM-HMM model is capable of classifying age-gender with accuracy of up to 96.4%. This model can be improved by tuning the more precise parameters and using larger datasets.

Keywords: Classification, Mel-Frequency Cepstrum Coefficient, Acoustic Models, Gaussian Mixture Model, Hidden Markov Model