Abstract

In this paper we propose an unsupervised approach to speaker segmentation using autoassociative neural network (AANN). Speaker segmentation aims at finding speaker change points in a speech signal which is an important preprocessing step to audio indexing, spoken document retrieval and multi speaker diarization. The method extracts the speaker specific information from the Mel frequency cepstral coefficients (MFCC). The speaker change points are detected using the distribution capturing ability of the AANN model. Experiments are carried out on different audio databases, and the method is capable of detecting speaker changes with short duration of speech in an unsupervised manner.

Reference
Unsupervised Speaker Segmentation using Autoassociative Neural Network

integrated approaches in broadcast news speaker diarization. Computer, Speech and Language. 20, 303-330.


Index Terms

Computer Science Speech Processing

Key words

Audio indexing

Speaker segmentation

Mel frequency cepstral coefficients

Autoassociative neural networks