Abstract

The class-specific automatic speech recognition systems construct an individual classifier for each class based on its own feature set, wherein the feature set for each class is selected such that it distinguishes that class from the other classes most accurately. Consequently, different feature set sequences must be fed into each of the classifiers, and the output of each of the classifiers must be combined to predict the actual class of the observation sequences. However, speech is continuous, and to be able to apply class-specific features, speech should be segmented and fed to the classifiers, which requires the identification of segmentation cues. This paper proposes a framework that jointly segments, and combines the output of the class-specific classifiers in the absence of any segmentation cues using a recursive formulation.

Index Terms

Computer Science

Signal Processing

Keywords

Class-specific feature set speech segmentation speech recognition