Abstract

In this paper, the performance of speaker modeling schemes such as vector quantization (VQ) and Gaussian mixture model (GMM) is compared for speaker identification. Along with the effective size of feature set, model based approaches are typically used as a solution for robustness issues of speaker recognition systems. Gaussian Mixture Model (GMM) is versatile parameter estimation approach whereas; Vector Quantization (VQ) is based on template modeling. Here, first, MFCC features are used to extract speaker specific speech features for text-independent speaker identification. MFCC features are then modeled using Vector Quantization (VQ) and Gaussian mixture model (GMM) and their performance is compared in the context of speaker identification. The average recognition rate achieved for MFCC with GMM is 99.2% and for MFCC with VQ is 98.4% on TIMIT database consisting of 64 speakers.

References

1. R. Togneri and D. Pullella, "An overview of speaker identification: Accuracy and
Comparison of Vector Quantization and Gaussian Mixture Model using Effective MFCC Features for Text-independent Speaker Identification


Index Terms

Computer Science  Signal Processing
Keywords

Speaker Identification, Text-independence, Feature Extraction, Vector Quantization, Gaussian Mixture Model