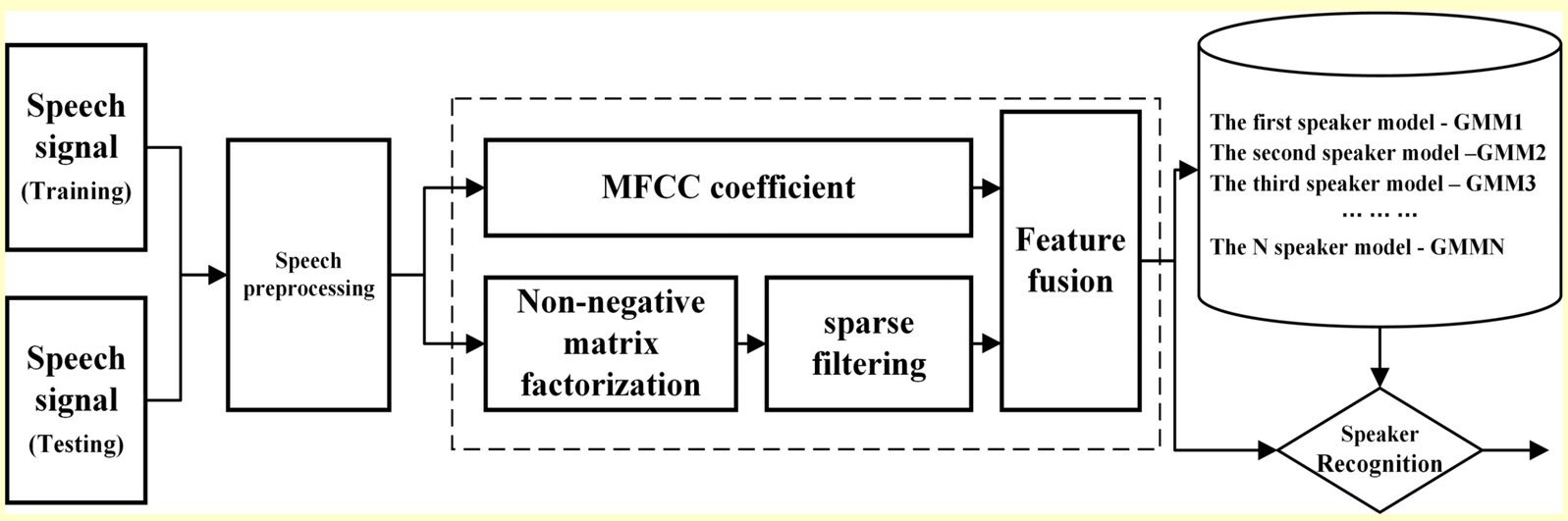
ICIEA 2022

16 - 19 Dec 22
Chengdu, China

ICIEA22-000060
A speaker recognition method based on GMM using NMF

Liming Huang, Dongbo Liu*, Yu Fang, Weibo Wang
School of Electrical and Electronic Information, Xihua University

A speaker recognition method based on GMM using non-negative matrix factorization



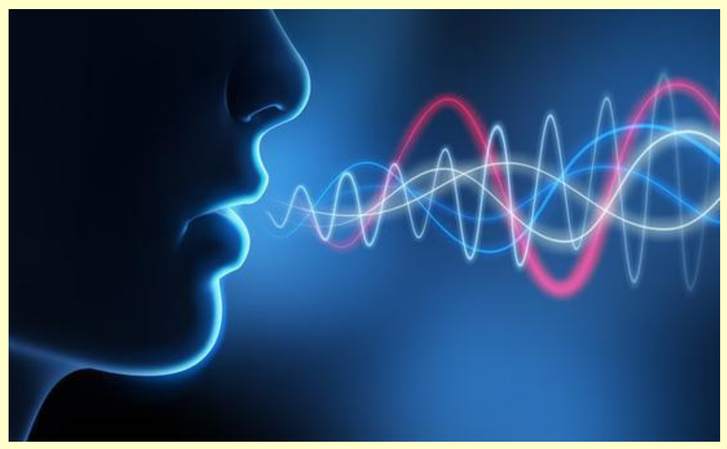


Fig.1 Overall framework of speaker recognition

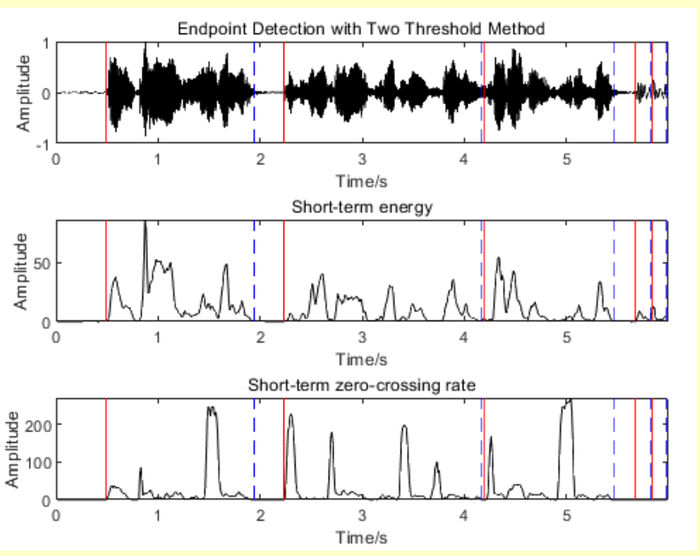


Fig.2 Endpoint Detection with Two Threshold Method

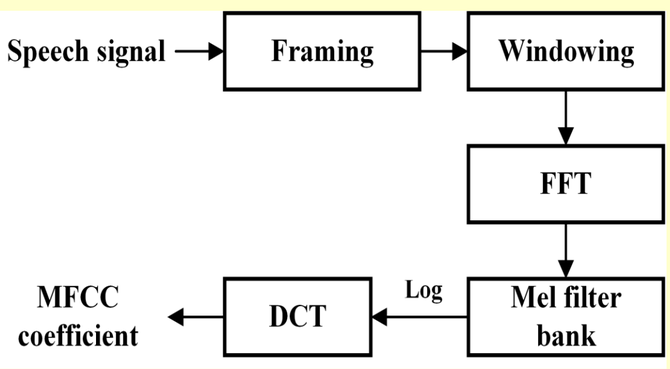


Fig.3 MFCC feature extraction

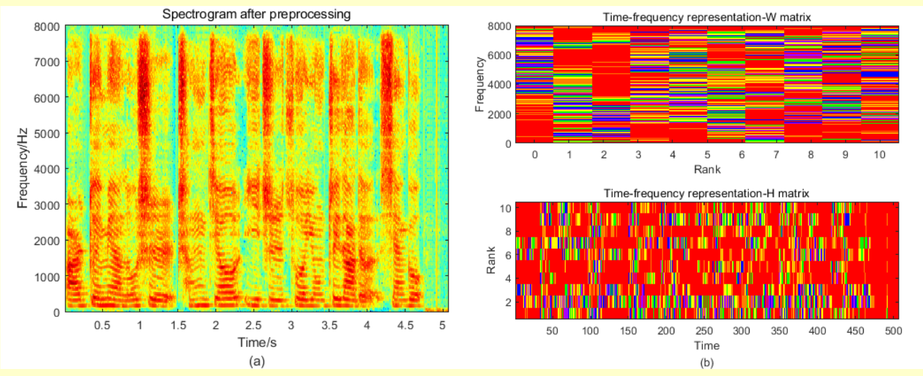


Fig.4 Non-negative matrix factorization spectrogram

The silent segment of the speech signal is cut by double-threshold endpoint detection. After preprocessing, the MFCC features and the spectrogram were obtained respectively. The W feature matrix is obtained from the spectrogram through non-negative matrix factorization using Euclidean distance as a loss function, and after sparse filtering, it is fused with MFCC features.

Experimental results show that the general method performs well in a calm environment, but is more sensitive to environmental changes; the improved method in this paper provides relatively more feature information, so that it can maintain relatively good performance when the signal-to-noise ratio decreases.

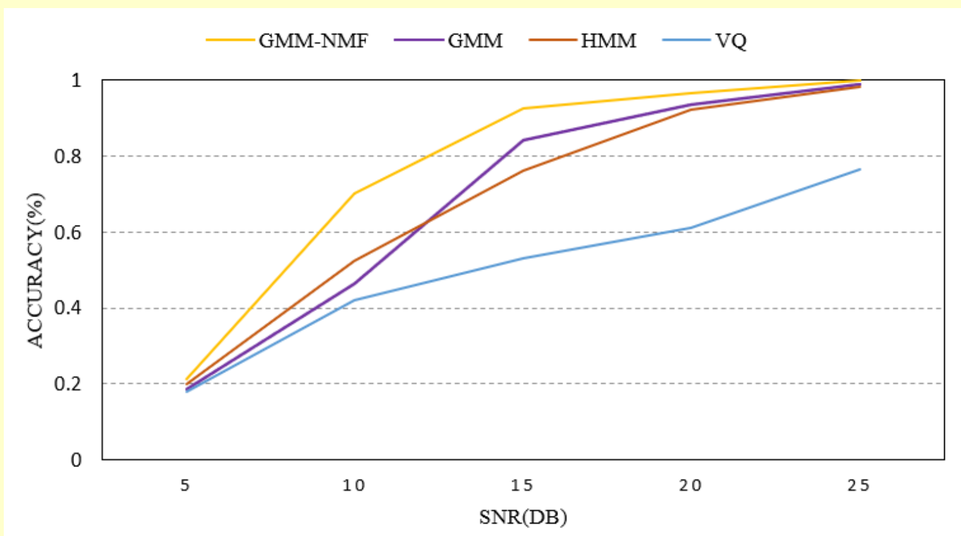


Fig.5 The recognition accuracy of different SNR under Factory1 noise

Algorithm	Global Accuracy (%)	15 dB to 25 dB Average Accuracy (%)
GMM (MFCC+NMF)	76.0	96.3
GMM (MFCC)	68.4	92.3
HMM	68.0	89.2
VQ	51.7	66.2

Summary

This paper applies non-negative matrix factorization to feature extraction from spectrograms and fuse it with MFCC for speaker recognition. Compared with not using fused features, the accuracy of this method is improved in noisy environments. The reason is that the inclusion of W matrix information provides more information for the GMM model that uses only MFCC as features at low signal-to-noise ratio.