**An Efficient Speech Recognition Algorithm for Small Intelligent**

***Abstract***—The speech recognition technology makes it possible for people to communicate with intelligent electronic devices. However, existing speech recognition algorithms are overly complex for small intelligent electronic devices (e.g., mini speakers, intelligent toys, intelligent remote controls, etc.). For this, an efficient speech recognition algorithm is proposed. Firstly, the *Mel-scale Frequency Cepstral Coefficients* (MFCC) is applied to extract features of voices. Secondly, the *Support Vector Machines* (SVM) is used to train speech classification models.

*Index Terms*—MFCC, SVM, Speech Recognition

# I. INTRODUCTION

Because these small intelligent electronic devices are cost controlled so that they can not accept a large computation load and additional communication chips. Therefore, it is urgent to design an efficient speech recognition algorithm for small intelligent electronic devices.

Erlin et al. [1] analyzed discriminative features of voices in detail, including loudness and tonal harmony, and designed the audio classifier according to the *nearest neighbor* (NN) criterion. Guo and Li [2] designed a SVM based multi-stage audio classifier. Lu et al. [3] proposed a hidden Markov model based audio classification method. Based on the work of Guo and Li [2], Lin used the wavelet transform to extract sub-band energy and pitch frequency [4] as features of voices.

In this paper, an efficient speech recognition algorithm is proposed. It uses *Mel-scale Frequency Cepstral Coefficients* (MFCC) for representing voices and applies *Support Vector Machines* (SVM) to classify speeches.

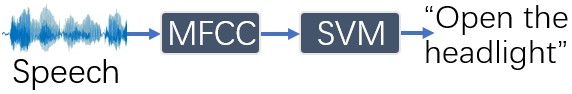


Fig. 1. The flow chart of the proposed method.

# II. METHOD

The proposed approach consists of two steps(i.e., MFCC and SVM), which are introduced as follows.

## A. MFCC

Due to MFCC [5] having an anti-noise ability, it is applied to extract features of speeches in this paper.

Because of the differences in speed and content of speech commands, the data lengths of speech commands are different, which is harmful to the following SVM.

For this, during the MFCC step, frame energy is descending sorted, and only top-128 frames are selected to construct features of speeches, making each speech holds the same dimensional feature vector.

## B. SVM

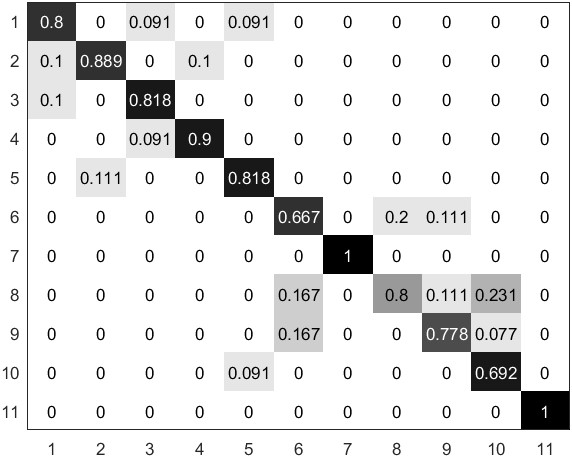
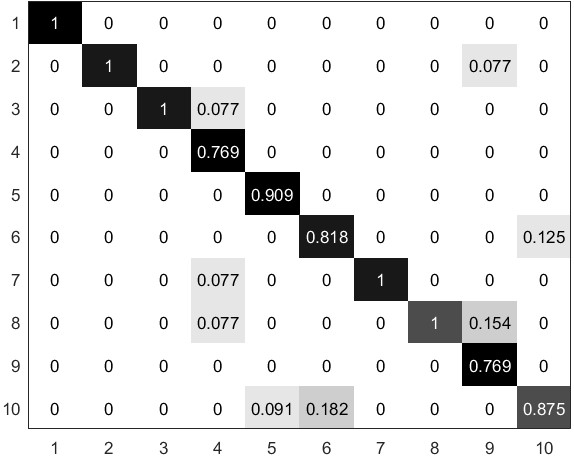
In practice, the interaction between human and small intelligent electronic devices does not require complex speech commands. The kinds of speech commands are limited to a specific electronic product. Therefore, we transform the speech recognition task as a classification task that aims to predict the class label of an input speech file. Considering that the SVM [6] has high efficiency and accuracy, it is applied to finish the speech classification task for small intelligent electronic devices in this paper.

# III. EXPERIMENT

## A. Database

In this paper, a new speech database is collected to validate the proposed method

# IV. CONCLUSION



### (a) (b)