Isolated Word Recognition

Stu.no: 1652770 Name: 梁琛

Introduction

This is a course lab using the hmm algorithm to apply speech recognition. Some changes have been made to optimize the acurracy of the original code provided by the course lab2. Detailed description will be given in the lab report. In this README file, the following part show the usage steps to run this programme.

Usage

Note: Step 1, 2, 3 is for the user who wants to train another model. So if you don't want to add any samples or re-train the model, just skip these step.

P.S. *step means that this step can be skiped.

Training model

1. *Create Training Record Samples

   - Run `createRecords.m`

   - In this step, you are creating training samples, so input **'record'** when *'Input the type name:[record/test]'* shows int the command line.

   - Following the text tip, input the word that you are going to record (*like: 'hello'*)

     - By default, the recording times of the sample is **10**.

   - After seeing *'Start speaking: (times)'* in the command line window, you can start to say the word.

   - The recording lasts **2 seconds** and if the *'End of recording: (times)'* shows in the command line, the recording finishs.

   - After that, the records will be saved in **'/record'**. If you are not satisfied with the records due to some noise, you can re-input the same input and the new recording will cover the old one.

   In my model, there are 10 words: `{'你好', '谢谢', '再见', '晚安', '早安', '抱歉', '梦想', '奇迹', '疾病', '灾难'}`.

   Each word has 10 .wav files in '/record'.

2. *Generate The Training Data
   - Prepare the `.wav` files in '/record' following step 1.
   - Run `generateSamples.m`
   - Modify the variable `names = {'你好', '谢谢', '再见', '晚安', '早安', '抱歉', '梦想', '奇迹', '疾病', '灾难'}`
   - The `wav` files will be converted and saved in `training.mat` automatically for next step.

3. *Train The Model
   - Run `main.m`
   - The trained model will be saved in the **myhmm.mat** automatically.

Word Recognization

4. Test The Result With **GUI**

   P.S. If you have added some samples, please find the function `axes1_CreateFcn` and `g_interfaces_OpeningFcn`. And modify the variables `word` and `text`.
   - Run `g_interfaces.m`
   - Press button `开始录音` and `结束录音` to record your voice.
     - *'Start speaking.'* and *'End of recording.'* will be showed in the text box to be a instruction.
   - Press button `播放` to play the recording after the above step.
   - Press button `识别波形` to get the recognition result.
   - You can also add *.wav* file from the menu.

Test Accuracy

5. Create Testing Record Samples
   - This step is very similar to *Step 1*, just following step 1 with the input **'test'** instead of **'record'**.
   - The testing records will be saved in **'/test'** automatically.

5. Get The Accuracy
   - Run `getAccuracy.m`
   - It will grab the test *wav files* in the `test` subdirectory and use the model to print the results in command line.

Interface