Learning English Pronunciation using Hand-Held Devices
SM1-02
Supervised by Dr. Siu Man Hung
Group member:
Lau Yeuk Chung
Introduction

An improvement in the correct pronunciation of English words is hard to achieve through self-learning. Using speech recognition techniques, learning to pronounce English words correctly can become easier and more convenient. Students can learn anytime and anywhere using a computer.

At HKUST, a PC-based system called 'Plaser' (Pronunciation Learning via Automatic Speech Recognition)[1], a multi-media tool that uses automatic speech recognition technology for effective English pronunciation learning was developed.

The aim of this project is to develop a system with functionalities similar to those of the existing system 'Plaser' on a hand-held device (a windows CE based machines).

Features

Some features developed in this project are listed below.

1) User’s speech can be retrieved by the microphone of the hand-held device.
2) The application can point out which part of the user’s speech is different from standard and calculate a reference score for the user using the technique of speech recognition.
3) After processing, the reference score is visualized as color segments of the word and displayed to the user.
System Architecture

**Application on Windows CE**

- **Gaussian mixture model decoding results**
- **Result of forced alignment**

**The recognizer**

- **Wave to MFCC convert**
  - **Wave recorded**
  - **MFCC of the recorded wave**

- **Standard input**
  - **Speech data of user**
  - **Button input**
  - **Touch screen**

**The User interface**

- **Teaching materials and result**
- **Colors of alphabet of a word**

**Html viewer**

- **Garbage rejection and confidence measure**
  - **Pre trained hmm model**
Results

System performance

<table>
<thead>
<tr>
<th></th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>50.43%</td>
</tr>
<tr>
<td>Speed</td>
<td>33.25 ms per frame per phone</td>
</tr>
</tbody>
</table>

Snap shots of user interface

Results displayed through color segments of words

Conclusion

This project built a prototype program of using automatic speech recognition on teaching on hand-held devices. It has market value as there is only a few products with similar functionality on the market at the moment.