Voice-Enabled PDA

FP1-04

Supervisor: Professor Pascale Fung

Students: Kwok Tsz Fai
          Li Tik Ho
System Architecture

![Diagram of System Architecture](image)

**Fig. 1** Training Process

![Diagram of Training Process](image)

**Fig. 2** Recognition Process

![Diagram of Recognition Process](image)
Overview

In the last few years, a kind of computer which is called as Personal Digital Assistant (PDA) became popular because of its mobility. This handheld device that combines computing, fax sender, web browser and personal organizer features did actually act as a good assistant for user.

However, PDA does not help the one who has difficulty in using hand-type or taping on the screen. There is a more efficient way to communicate with the handheld device. Voice can act as commands to control the PDA. Even a user with little knowledge on computer can use the voice-controlled device easily.

We aimed at producing a PDA program that can recognize commands of spoken words in our project. A system would be implemented to help user to find the phone book or address book in PDA. It is a speaker-independent system, which does not require the user to record his or her voice beforehand.

Software Specifications:

• Running in PalmOS 5.
• Users are allowed to use voice speech to find a name in the name list.
• Design for Cantonese Speaker.
• Capable to update the Chinese name list dynamically.
• Capable to bind 50 names in name list.
• Speaker Independent, consistent result will be shown for different users.
• Isolated words would be recognized.
Testing Results:

<table>
<thead>
<tr>
<th>Number of words to be recognized</th>
<th>Accuracy (%)</th>
<th>Run time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>100</td>
<td>34-36</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>64-68</td>
</tr>
<tr>
<td>30</td>
<td>100</td>
<td>90-94</td>
</tr>
<tr>
<td>40</td>
<td>100</td>
<td>121-125</td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>155-165</td>
</tr>
</tbody>
</table>

The run time and number of word phase relationship:

Run time = (number of words + 1) X 3s

Conclusion:

1. If there is more name in grammar book, the processing time increases.
2. If the user speaks slowly, the processing time increases.
3. Although the algorithm’s speed could be improved by decreasing the number of mix-
tures, the accuracy would decrease too.
4. Speaker independent, the system can work for different users without previous re-
cording.
5. The name list can be up to a size of 50, the accuracy is also high.
6. It works for Cantonese speaker.
7. It is a Palm OS applicable program.
8. The name list can be updated for different name combinations.