Automated Telephone Number Retrieval using Speech Recognition

Author: I Mun, Chia
Supervisor: Associate Professor E. Ambikairajah

1. Abstract
This thesis involves the implementation of an automated telephone number retrieval system using a speaker dependant, isolated-word, speech recognition system in MATLAB®. Several schemes of improving recognition accuracy and robustness were incorporated, with each displaying varying levels of effectiveness.

2. Introduction
Automatic Speech Recognition (ASR) is the process whereby a computer correctly identifies spoken words. With the proliferation of electronic devices and the demand for them to be smaller and more mobile, ASR is emerging as an attractive alternative to conventional user interfaces. ASR is natural, flexible, efficient and relatively independent of device size, and there is an increasing need for ASR systems that are accurate, efficient and robust.

3. Speech Recognition System Block Diagram

4. Implementation and Testing
The system was trained using a set 36 names and tested using varying levels of additive noise and with different schemes of improving accuracy and robustness. The schemes used included:
- Dynamic MFCC features
- Cepstral Mean Normalisation
- Temporal Masking
- Wiener Filtering

5. Results
Recognition accuracy was good in ‘clean’ conditions but degraded quickly with increasing levels of noise. The different schemes used helped to improve performance.

6. Conclusion
Automatic speech recognition is a technology that has much potential, unfortunately it is often hampered by a decrease in accuracy when performing under non-ideal conditions. E.g. in the presence of noise. Of all the different schemes investigated that helped improve robustness, most were only mildly effective with the exception of Temporal Masking and Wiener Filtering. It is recommended that further research on both schemes and their benefits to speech recognition be carried out.