CHAPTER 8
CONCLUSION AND FUTURE ENHANCEMENT

8.1 Conclusion

It remains to be seen whether or not speech recognition captures the imagination of the general population. Certainly it appears most likely that people will confront this technology, initially through their telephone rather than with their computer. Continuous speech (i.e., without pauses) and speaker independence (i.e., no need for training) are both highly desired prizes in the world of speech recognition. Given the dazzling pace of new technology, these developments cannot be far away.

Speech recognition emerged as a computer access option at the beginning of this decade. Although its potential was recognised, indeed anticipated, the early systems were thwarted by high cost, slow processors and expensive computer memory. The 1990s have witnessed dramatic developments in computer technology, and these changes have been fortuitous for speech recognition. Processors have become faster, memory capacity has increased enormously and the cost of computer systems has diminished. More work remains to be done on the problem of real-time implementation. The challenge is to process these recordings and extract the knowledge in a form amenable to further analysis and reuse.

The major contributions of this research work could be summarized as follows:

1. Specified the Importance of Acoustics in speech recognition system.
2. A new method is proposed to filter unambiguous speech signal using Input signal Processing Model for a practical domain and integrated acoustic-phonetic
knowledge in statistical speech recognition using Sphinx-4 open source speech engine. It was tested with Railway help desk system.

3 Technique and design strategies used to define better word model (sub phonetics) in speech recognition to reduce word error rate and for performance improvement.

4. Implemented HMM with Viterbi, forward-backward Algorithms for parallel rate specific acoustic model for fast and slow speech in Speech recognition. The current approach used identical pronunciations, but different phone units have taken to model fast speech versus slow speech.

5. Designed and implemented a standard usability procedure for providing Speech Procedures and also error recognition and correction strategies.

6. Implemented a new algorithm to identify and verify the speaker identity and explained about different types of Speaker identification procedures and their performance comparisons.

The most important issue is how to make the speech recognition systems robust against acoustic and linguistic variation in continuous speech. In this context, a paradigm shift from speech recognition to understanding, where underlying messages of the speaker, that is, meaning/context that the speaker intended to convey, are extracted, instead of transcribing all the spoken words, will be indispensable. To reach such a goal, the research need to have an efficient way of representing, storing, retrieving, and utilizing world knowledge.

8.2 Future Enhancement

Even though the thesis have taken a big step forward in being able to transfer speech recognition systems from the laboratory to the real world, the field of environment independent recognition is still in its infancy. This section, describe future research that appears to be worthwhile.
Incorporating spelling information should be done to improve the senonic baseform in order to take advantage of the adaptation capability. Another improvement would be to incorporate senone bigram models or other higher level models in the state quantization procedure.

To make the system more robust to the noisy environment, especially for low SNR environment, some new modifications to the spectral subtraction Method may be researched for low-cost applications. One of these methods may works to use the statistical way to find a distribution which can separate the noise and speech from each other. It is necessary to investigate the behavior of many different microphones and environments so that more general conclusions can be drawn. It would also be desirable to test the algorithms with telephone speech.

The absence of a standard database for evaluation of algorithms is definitely not benefiting in the field of environment-independent recognition. The continuous speech recognition community has benefited from the existence of common tasks and databases, so that direct comparison between algorithms is more straight forward. In future similar efforts can be directed toward the development of standard databases for environment-independent recognition.