Chapter 9

Conclusions and Future Scopes

The work described in this thesis is focused on designing single channel real time speech enhancing system for the low SNR (0-5dB) range. By qualitative and quantitative analysis, the explanation in the thesis has shown that a practical single channel real time embedded speech enhancement technique can enhance signal quality. The hybrid approach suggested here can work in 0-5dB SNR range and can handle additive noise and convolutive distortion. Both the objective and subjective tests advocate the improvements compared to original algorithms in low SNR range (0-5dB) with the hybrid approach. The major difficulty in real time implementation is due to non-causal nature of RASTA algorithm. The sub-framing approach is used to solve this problem. This thesis has described the real time implementation of hybrid algorithm on PC as well as on DSK6713 through SIMULINK, Real Time Workshop and Target Support Package TC6. The profiling results are obtained and compared. For PC implementation the algorithm works fine and gives the real time enhanced speech output. But for DSK6713 implementation it is not the case. The enormous resources available on PC are responsible for the better performance. For DSK implementation as already indicated the main loop requires optimization as the execution can't be completed within base sample time. So it is concluded here that the hybrid algorithm is found suitable for real time embedded implementation in communication systems but requires optimization before final real time hardware implementation.

Table 9.1 shows the list of contemporaneous research work done in similar direction by other researchers. It is not possible to make exact comparisons as the results are reported using different database and performance measures. Also no one has reported about embedded or real time implementations. However, an attempt is made here to brief the most common results and can be compared with the hybrid approach. The hybrid approach offers comparatively appreciable results under different noisy and reverberant conditions.

<table>
<tr><th>Ref. no.</th><th>Brief of technique/principle</th><th colspan="4">Results reported</th></tr>
<tr><td rowspan="4">[1]</td><td rowspan="4">Multidimensional MMSE STSA estimators based on correlation between spectral components, an optimization parameter γ between 0 to 1 places lower and upper bounds.</td><td colspan="4">Advantageous at high SNR, PESQ under white noise:</td></tr>
<tr><td>γ</td><td>0</td><td>0.5</td><td>1</td></tr>
<tr><td>5 dB</td><td>1.3</td><td>1.24</td><td>1.21</td></tr>
<tr><td>10 dB</td><td>1.57</td><td>1.52</td><td>1.46</td></tr>
<tr><td>[2]</td><td>MMSE STSA85 for speech enhancement and independent component analysis (ICA) for noise estimation</td><td colspan="4">Real railway station is used to test the algorithm,
MOS: 4.2</td></tr>
</table>

[3]	β power MMSE STSA85 (β SA), β is power of the optimization cost function	β=-1 gives good compromise between noise reduction and speech distortion, MOS at 0 dB white noise: 2.7, PESQ under white noise:	
		0 dB	1.47
		5 dB	1.72
		10 dB	1.96
[4]	Maximum likelihood phase equivalence of speech and noise, Generalized Gamma distribution function for speech and noise spectral amplitudes	SSNR improvement compared to MMSE STSA85 under white noise:	
		0 dB	8.6%
		5 dB	7.4%
		10 dB	6.3%

Table 9.1 Comparison of results from the research papers published contemporaneous

Significant progress has been made in the development of single channel speech enhancement algorithms. Robust human-human communication with only two sensors and channels even in adverse conditions still haunts researchers in the field. Future work will explore some of the research directions pointed out in the thesis so far.

- The hybrid approach suggested here can be optimized by merging MMSE and RASTA algorithm together. Also the RASTA filters can be redesigned with better specifications.
- The algorithm is still unable to handle highly non-stationary noise. So such a scheme can be incorporated into hybrid algorithm.
- This thesis highlighted the importance of magnitude spectrum information for estimating the true clean speech magnitude in all algorithms. However, an attempt can be made at estimating phase and the complex spectral subtraction instead of magnitude spectral subtraction can be used.
- Instead of single channel approach a multichannel approach can also be investigated for viable real time implementation. In the present scenario the size and cost of microphone array limits the multichannel approach to challenge the single channel approach for speech enhancement. Multichannel speech enhancement algorithms are more robust for different noise conditions compared to single channel speech enhancement techniques [5]. With the advent of nano-technology and MEMS, the miniaturisation of

microphones is emerging quickly. This will overcome the said disadvantage of the multi channel techniques. In near future, many devices like mobile phones, laptops, PCs etc. can have the microphone array embedded into them. In fact, the research work in this direction has already begun [6]. The commercial noise canceller in mobile phones will be soon in market as reported in [7].

- The multi speaker separation problem can also be tackle by using array processing.
- Also the real time implementation suggested here is done using embedded target toolbox of MATLAB. So the developed assembly code may not be highly optimized. Further optimization of the assembly code can be done.
- With a faster DSP or media processor or even an FPGA implementation, a number of improvements can be made without the need for higher power algorithms.
- Also due to high complexity the algorithm can be implemented using soft computing techniques like fuzzy logic, neural network and genetic algorithms.

The investigation of these implications is a valuable topic for future research and might yield substantial improvements.