Research Paper Engineering

Implementation Of Advanced Speech Enhancement System On Tms320c6713 Dsk System.

*N B Umashankar ** Anand Jatti*** Dr. S.C. Prasanakumar

*#216, 1st cross , 2nd Phase , 4th Block ,Banashankri 3rd Stage,Bangalore,Karnataka

**Associate Professor, Department of Instrumentation Technology , R.V. College of Engineering, Mysore Road, Bangalore,Karnataka

***Professor & Head,Department of Instrumentation Technology ,R.V. College of Engineering, Mysore Road,Bangalore, Karnataka

ABSTRACT

In modern handsfree speech communication environments often occurs a situation that the speech signal is superimposed by background noises which are sometimes captured at a higher level than the speech signal of interest. In order to improve the intelligibility and reduce the listeners stress, a noise reduction procedure called speech enhancement is applied which increase the signal to noise ratio.

This work employs implementation of some algorithms like MMSE-STA and MMSE-LSA in Matlab and compared with general spectral subtraction method. It was found that MMSE-LSA method yielded the best SNR and hence was chosen for real time implementation.

A TIA/EIA/IS127 Compliant Speech Enhancement System based on MMSE-LSA algorithm was implemented using TMS320C6713 DSK system. The improvement in the SNR was found to be 20db. The developed algorithm could be used in real time application like mobile phones, hearing aids and other speech activated devices.

Keywords :Speech Enhancement, Spectral subtraction, Log Spectral Amplitude

1 INTRODUCTION

The performance of speech coding and recognition system that operate in noisy environments (such as moving vehicles, crowded areas etc.) .Therefore, there has been a hot research topic on speech enhancement system since 1970s tell today. Speech enhancement is motivated by the need to improve the performance of voice communications systems in noisy environment. Applications range from front-ends for speech recognition systems condition to enhancement of telecommunications in aviation, military teleconferencing and cellular conditions. The goal is either to improve the perceived quality of the speech or to increase its intelligibility. The term intelligibility refers to our ability recognize the content of speech and quality refers to the aspect of the speech that determines the ease with which one can understand the speech.

The problem of enhancing speech degraded by background noise, when noisy speech alone is available, has been a research topic that has received great deal of attention over the past few decades. As mentioned before, there is only a single

Microphone available, is one of the most difficult situations in speech enhancement, since no reference signal of the noise is available, and the clean speech cannot be processed prior to being affected by the noise. The performance of single channel systems is usually limited because they tend to improve the quality of the noisy signal at the expense of some intelligibility loss. Therefore, there is a tradeoff between quality and intelligibility.

Applications of speech coding and speech recognition have been exploding these days. Many civilian and military working environments rely upon the human speech and hearing as the main method for person-to-person communication, often coded and transmitted via an auxiliary link, e.g. radio or wire. The human speech and hearing organ is inherently sensitive to interfering noise. Enduring interfering noise decreases speech intelligibility and makes speech communication troublesome, eventually causing human mistakes or damage to the hearing. Vehicles and tools used by personnel at a working site could emit interfering noise and thus negatively affect the communication and personal safety in the working environment. Hence, algorithms and implementations for speech enhancement are required. The speech enhancement implementations should preferably be robust to the rough environment in which they are intended. Moreover, versatility and flexibility are key features for speech enhancement devices, e.g. the ability to adapt to changing environment and to fit into a variety of applications. Battery is the predominant energy source for mobile solutions; hence optimized battery life time is necessary [1]-[2].

Enhancing of speech degraded by noise, or noise reduction, is the most important field of speech enhancement, and used for many applications such as mobile phones, VoIP, teleconferencing systems, speech recognition, and hearing aids.

2 MINIMUM MEAN SQUARE ERROR –LOG SPECTRAL AMPLITUDE TECHNIQUE

The algorithm is based on the observation that for each frequency band Ω i the smallest value of .$R_{xx}(n,\Omega_i)$ (the power spectral density (PSD) estimate of a noisy speech signal) that is observed in a sufficiently large number of consecutive frames corresponds to the noise only. Consequently, by tracking these minima in a sliding window covering several frames, an estimate for the noise magnitude spectrum can be obtained. To get a reliable noise power estimation the frame size must be large enough to bridge speech activity. At first the PSD estimates of the noisy speech signal .$\hat{R}_{xx}(n, \Omega_i$ are to be calculated. An effective way is offered by an exponential decaying window (first order recursive averaging):

$$R_{xx}(n, \Omega_i)=\alpha_N\ R_{xx}(n-1, \Omega_i)+1-\alpha_N)\ X^2(n, \Omega_i)$$

where αN (αNє[0,1]) is a smoothing factor

The higher this factor is, the more stable and smooth the estimate will be. On the other hand the ability to track sudden changes will decrease. When equ. 5 is used for the estimation of the noise PSD a relatively low factor αN should be chosen (we achieved the best results with αN = 0.85).

The minimum noise power estimate Nmin ^ $2(n,\Omega i)$ of subband i is obtained by frame-wise Comparison of the actual smoothed signal power estimate $\hat{R}_{xx}(n,\Omega_i)$ and some preceding PSD values . $(n-1, \Omega_i)$ with r = 1, 2, 3, . . . , R-1, which are stored in a FIFO register. The introduction of that FIFO register is the special feature of the minimum statistics algorithm. The depth of the FIFO is given by R. (See Fig. 4 for the structure of the noise power estimation algorithm.)

If the actual subband power $X^2(n,\Omega i)2(n,\Omega i)$ is smaller than the estimated minimum noise power N2min (n,Ωi) the minimum noise power spectrum is updated immediately:

$$N^2_{min}(n,\Omega_i)=\min\{R_{xx}{}^2(n-r, \Omega_i)\}$$

$$1 \le r \le R-1$$

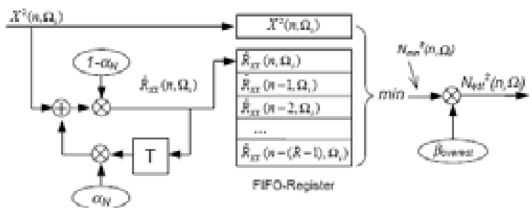

Figure 1 Structure of the sub band noise power estimation algorithm

Thus, in case of decreasing noise power, we achieve a fast update of the minimum power estimate. In case of increasing noise power the update of noise estimates is delayed by R samples.

3. TIA/EIA/IS-127 Compliant Speech Enhancement System

A TIA/EIA/IS127 Compliant Speech Enhancement System is pre-processing block in Enhanced Variable Rate Codec (EVRC) used to enhance the speech signal before encoding the speech signal .

The main components of the TIA/EIA/IS127 Compliant Speech Enhancement System are:

1. High Pass System
2. Adaptive Noise Suppression System

- High Pass System comprises 6th order Butterworth filter implemented using 3 sections of Biquad Filter.
- Noise Suppression System contains sub-systems given by:

Frequency Domain Conversion, Channel Energy Estimation, Channel SNR Estimate, Voice Metric Calculation, Spectral Deviation Estimator, Background Noise Update Decision, SNR Estimate Modification, Channel Gain Computation, Frequency Domain Filtering, Background Noise Estimate Update and Time Domain Signal Reconstruction.

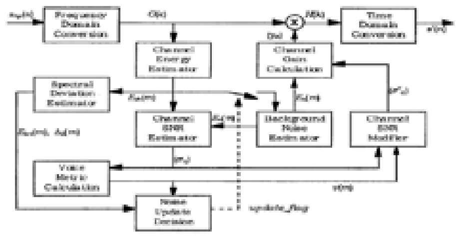

Figure.2 TIA/EIA/IS-127 Compliant Speech Enhancement System

4 Hardware Implementation of Noise Suppression System

➢ A MMSE-LSA based speech enhancement algorithm discussed in the previous section is written in C language in the Code Composer Studio environment.

➢ The TMS320C6713 DSK has a codec which has pins called LINEIN and LINEOUT. The input speech signal is given from PC through a wavosaur software where the signal is made to run continuously.

➢ Through a USB cable the code is flashed on to a TMS kit which is powered by USB cable or a AC adapter.

➢ The output the algorithm is made available at the LINEOUT of codec of the kit. The enhanced speech can be heard by connecting headphones to this pin. Also the corresponding waveform can be viewed on wavosaur software.

The code we developed takes the noisy speech signal through the 'line in' terminal of the DSK and the processed output comes out through the speakers or headphones which are connected to the 'line out' terminal of the DSK

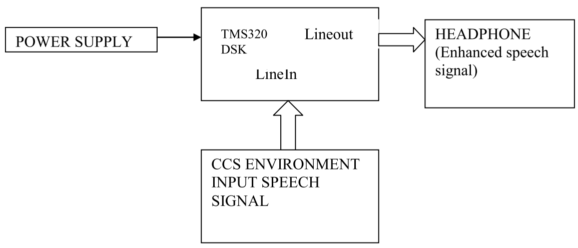

Figure 3 Block Diagram showing DSP Kit implementation of speech enhancement system

5 RESULTS

The signal to noise ratio obtained for different methods are tabulated as follows. We see that best SNR is obtained by MMSE-LSA method which is 20dB superior than the conventional spectral subtraction method.

Noise Removal Method	SNR Improvement (in dB)
General Spectral Subtraction Method (SS)	7.31
Parametric based Spectral Subtraction Method	25.15
Multiband Spectral Subtraction Method	27.29

MMSE_STSA	17.89
MMSE_LSA	22.74
ANS using the MMSE_LSA based method described in the EVRC standard	38.82

Table Comparison of various spectral techniques

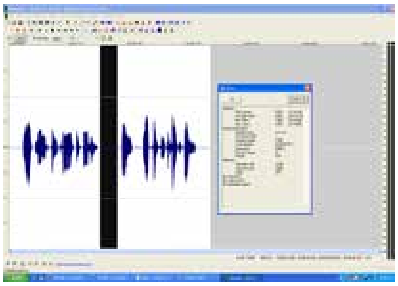

Figure 4 Speech Signal Enhanced by MMSE-LSA technique

REFERENCES

K. Wu and P. Chen, (2001) , Efficient speech enhancement using spectral subtraction for | car hands - free application, International Conference on Consumer Electronics, | , Vol. 2, 220-221. | [2] J. Lim and A. Oppenheim, (1979) ,Enhancement and bandwidth compression of noisy speech,Proc. IEEE, Dec., Vol 67, No. 12, 221-239 | [3] N. Virag, (1996),Speech enhancement based on masking properties of the human auditory system, Master thesis, Swiss Federal Institute of Technology, ,Vol 1, 796-799 . | [4] Anuradha R. Fukane, Shashikant L. Sahare,(2011), Role of Noise Estimation in Enhancement of Noisy Speech Signals for Hearing Aids, International Conference on Computational Intelligence and Communication Systems, Gwalior, India , 648 – 652. | [5] Mehdi Yektaeian, Rassul Amirfattahi,(2007),Comparison of Spectral Subtraction Methods used in Noise Suppression Algorithms , 6th International Conference on Information, Communications & Signal Processing, Singapore, Vol 15, Issue 6, 1–4. | | [6]Mansanori Kato , Akihiko Sugiyama,(2006),Noise suppression with high speech quality based on weighted noise estimation and MMSE STSA, Journal of Electronics and communication in Japan, Vol. 89, 43-53. | [7]Sunil D Kamath ,Philipos C Loizou ,(2006), A Multib and Spectral Subtraction Method for Enhancing Speech corrupted by colored noise, | IEEE International conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, Florida, USA, vol.4, IV-4164. | [8] Yang lu, (2011),Review of spectral subtraction method for speech enhancement technique, International Journal of Electronics and communication Technology, Vol 2 , Issue 4 , 189-194 , ISSN : 2230-7109. | [9]Boh Lim Sim, Yit Chow Tong, ,Joseph S. Chang, and Chin Tuan Tan, Member, IEEE,(1998), A Parametric Formulation of the Generalized Spectral Subtraction Method, IEEE Transactions On Speech And Audio Processing, Vol. 6, Issue. 4, 328-336. | [10] Zhi-Heng Lu, Huai-Zong Shao, and Tai-Liang Ju , (2009), Speech Enhancement Algorithm Based on MMSE Short Time Spectral Amplitude in Whispered Speech", Journal Of Electronic Science And Technology Of China, VOL. 7, NO. 2, 115-118. |