PERFORMANCE ANALYSIS OF DUAL CHANNEL SPEECH ENHANCEMENT METHOD WITH DCT AND HADAMARD TRANSFORMS USING NORMALIZED LATTICE RECURSIVE LEAST SQUARE ALGORITHM

Dr. D. Deepa[1], Dr. M. G. Sumithra[2], Dr. C. Poongodi[3]

[123]Professors, Bannari Amman Institute of Technology, (India)

ABSTRACT

Enhancement of noisy speech signal plays an important role in many speech processing techniques. These techniques will vary based on the number of channels available in the application. This proposed approach is for dual channel applications where one channel has input noisy speech signal and second channel has reference speech signal. In addition to the normal processing, in the proposed method the input signals are preprocessed with Discrete Cosine Transform and Hadamard Transform to segregate the speech samples based on its frequency and to make the process more convergent. Among various adaptive algorithms like LMS, RLS etc., the proposed Normalized Lattice Recursive Least Square (NLRLS) adaptive algorithm gives better performance interms of convergence rate, minimum mean square error, Improved SNR. The performance analysis is done through various subjective and objective measures

Keywords: Speech enhancement, DCT, Hadamard Transform, Adaptive filtering, Normalized Lattice Recursive Least Square algorithm, Signal to Noise ratio, Mean Square Error.

I. INTRODUCTION

Speech processing applications includes compression, enhancement or recognition of speech signals. Processing of speech in digital domain is easy to store and retrieve. This contributes more towards the development of efficient speech processing techniques to face the issues created during speech processing. The objective of enhancing speech is to improve the quality of speech by reducing the interference and noise included during speech processing. This is important in a variety of contexts, like environments with interfering background noise in speech recognition systems, hands free environment for cars, hearing aid devices etc. The other important use of speech enhancement is to improve the perceptual quality of speech in order to reduce listener's fatigue condition.

Based on the number of channels available the speech enhancement technique will vary. In single channel system reference signal will not be available. This issue can be addressed in dual channel system. In dual channel system one channel has the noisy speech signal that is to be processed and the other channel consists of the reference clear speech signal. Adaptive filter is used for this type of noise cancellation. The main focus of this work is to reduce the additive noise present in the input noisy speech signal and restore to its original form. This proposed algorithm can be used in real time applications like speech recognition, if the input is noisy speech the system cannot recognize the signal due to clean desired signal, and also convergence rate is very important because as it is real time application. In order to have fast convergence Normalized Lattice Recursive Least Square algorithm is used and the performance is compared with conventional Recursive Least Square algorithm.

To analyze the performance of the proposed algorithm the following objective quality measures and considered. Objective speech quality measures analyzed (Yi Hu and Philipos C. Loizou ,2008) in this work are given below:

Signal to Noise Ratio (SNR): SNR is the ratio of signal power to noise power expressed in decibels (dB)

Mean Square Error (MSE): Mean Square Error (MSE) metric is frequently used in signal processing and is defined as the average value of the square of difference between clean and enhanced signal.

In addition to the above measures the output interms of time domain and frequency domain plots are obtained for the proposed algorithm

II. LITERATURE SURVEY

The following are some the survey findings related to the proposed speech enhancement algorithm.

Rankovic (1998) proposed adaptive linear filtering which improves effective speech to noise ratios by attenuating spectral regions with intense noise components to reduce the noise spread of masking onto speech in neighboring regions. This mechanism was examined in static listening conditions for seven individuals with sensorineural hearing loss.

An adaptive subband noise removal algorithm is proposed by Shields and Campbell (2001). It performs binaural preprocessing of speech signals for a hearing aid based on subband approach. It uses the Least Mean Square (LMS) algorithm in frequency limited subbands, and provides better performance due to the separation of signal based on its frequency the elimination of high frequency noise will be better in adaptive algorithm. The results show that there was some distortion and considerable amount of noise present in the output signal, which will lead to reduced intelligibility and create listener fatigue.

Sunitha and Udayashankara (2005) proposed speech enhancement methods, in the first method the noisy speech signal is transformed using DCT and processed using adaptive algorithm, in another method the noisy speech is transformed using DFT and processed using adaptive algorithm. These method has poor intelligibility.

III. PROPOSED SPEECH ENHANCEMENT ALGORITHM USING NORMALIZED LATTICE RECURSIVE LEAST SQUARE ALGORITHM

In this work, speech enhancement algorithm is proposed with the use of adaptive filter using Normalized Lattice Recursive least square (NLRLS) algorithm. Adaptive filter (Dessouky. M. I , 2008) processes two input signals. One of the input signals is the noisy speech to be enhanced and second input is the reference signal. In this work input noisy speech and reference signal are transformed from time domain to frequency domain using Discrete Cosine Transform and Hadamard Transform. This is to separate high frequency noise from noisy speech signal for speech enhancement process since analysis of signal is efficient in frequency domain rather in time domain.

Speech enhancement using NLRLS algorithm with DCT and Hadamard transformation of input signals as a preprocessing technique has four stages. First the reference speech and input noisy speech signals are transformed using Discrete Cosine Transform (DCT). The resultant signal is then processed using Hadamard transform, which frequency alignment is performed for the DCT samples; this process is useful for reducing high frequency noise with adaptive algorithm. Next the dual transformed signal is normalized using power normalization technique and finally these signals are applied to an adaptive filter where NLMS algorithm is used for adaptation. Due to arrangement of samples based on its frequency the enhancement is done promptly using NLRLS algorithm.

Step 1: Discrete Cosine Transform

Generally transforms will convert time domain signal to frequency domain signal. Discrete Cosine Transform converts the input time domain speech signal into a frequency domain signal by representing it as coefficients. It consists of real valued components. With smaller amount of coefficients this will give better approximation. It has strong energy compaction property which means that the signal information will be available in few low frequency components.

Step 2: Hadamard Transform

In the second stage, Hadamard transformation is applied to the cosine transformed noisy speech signal and reference signal. The Hadamard transform H_m is a $2^m \times 2^m$ matrix, the Hadamard matrix scaled by a normalization factor, which transforms 2^m real numbers x(n) into 2^m real numbers X(k). Recursively, the 1×1 Hadamard transform H_0 is defined by the identity $H_0 = 1$, and then H_m is defined for $m > 0$

$$H_m = \frac{1}{\sqrt{2}}\begin{pmatrix} H_{m-1} & H_{m-1} \\ H_{m-1} & -H_{m-1} \end{pmatrix}$$

The above equation specifies the Hadamard transform of order 'm', where the $1/\sqrt{2}$ is a normalization value. Other than the normalization factor, the Hadamard matrices are made up entirely of 1 and −1. The same Hadamard Transform equation is used during inverse process after adaptation.

Step 3: Power Normalization

The dual transformed signal is then normalized by the square root of its power $p_k(i)$. The powers are estimated based on the sliding rectangular window.

The power normalized signal $v_k(i)$ is given in equation

IJARSE
ISSN: 2319-8354

$$v_k(i) = \frac{u_k(i)}{\sqrt{p_k(i) + \xi}}$$

where $p_k(i) = \beta\, p_{k-1}(i) + (1-\beta)\, u^2{}_k(i)$,

β is the normalization constant between 0 to 1, the values are chosen based on the performance of the system and $p_{k-1}(i)$ is the power value of the previous sample. The small constant ξ is introduced to avoid numerical instabilities when $p_k(i)$ is close to zero.

Hadamard transformation followed by a power normalization stage causes the adaptive filter inputs to speed up the convergence of the adaptive weights. The output vector after power normalization is specified in equation.

$$v_k(i) = [\, v_k(0), v_k(1), \ldots .\ \ldots v_k(i-1)]^T$$

Further the dual transformed, power normalized speech samples are applied as an inputs for the adaptive filter using NLMS algorithm.

Step 4: Adaptive Filtering

The key component in designing the adaptive filter is of different algorithms used for weight updation. In this work NLRLS algorithm is used for adaptation. The **Lattice Recursive Least Squares** adaptive filter is related to the standard RLS except that it requires fewer arithmetic operations (order N). It offers additional advantages over conventional LMS algorithms such as faster convergence rates, modular structure, and insensitivity to variations in eigen value spread of the input correlation matrix. Specifically the algorithms based on the lattice realization are very attractive because of their modular implementation and require a reduced number of arithmetic operations. As a consequence, the lattice recursive least-squares (LRLS) algorithms are considered fast implementations of the RLS problem. The LRLS algorithms are derived by solving the forward and backward linear prediction problems simultaneously. The lattice-based formulation provides the prediction and the general adaptive filter (joint-process estimation) solutions of all intermediate orders from 1 to N simultaneously. Consequently, the order of the adaptive filter can be increased or decreased without affecting the lower order solutions. This property allows the user to activate or deactivate sections of the lattice realization in real time according to performance requirements. Unlike the RLS algorithm, which requires only time-recursive equations, the lattice RLS algorithms use time-update and order-update equations.

A key feature of the LRLS algorithms is that the prediction process discloses the properties of the input signal. The internal signals of the prediction part retain in a sense non redundant information of the input signal that can be utilized in a decoupled form in the following processing. This mechanism is inherently built in the lattice algorithm derivations. The performance of the LRLS algorithms when implemented with infinite-precision arithmetic is identical to that of any other RLS algorithm. However, in finite-precision implementation each algorithm will perform differently.

The normalized form of the LRLS has fewer recursions and variables. It can be calculated by applying normalization to the internal variables of the algorithm which will keep their magnitude bounded by one.

Output of the NLRLS filter is then inverse transformed with Hadamard and IDCT to get the time domain signal, where the transformed signals are in frequency domain.

IV. EXPERIMENTAL RESULTS

The experimental results of the proposed DCT-Hadamard-NLRLS algorithm are given in terms of time domain and frequency domain plots further the performance measures are analyzed (D. Deepa and Dr. A. Shanmugam (2009). For time domain representation amplitude versus time plot is taken. Further spectrogram is given to represent the time versus frequency relation for the proposed method.

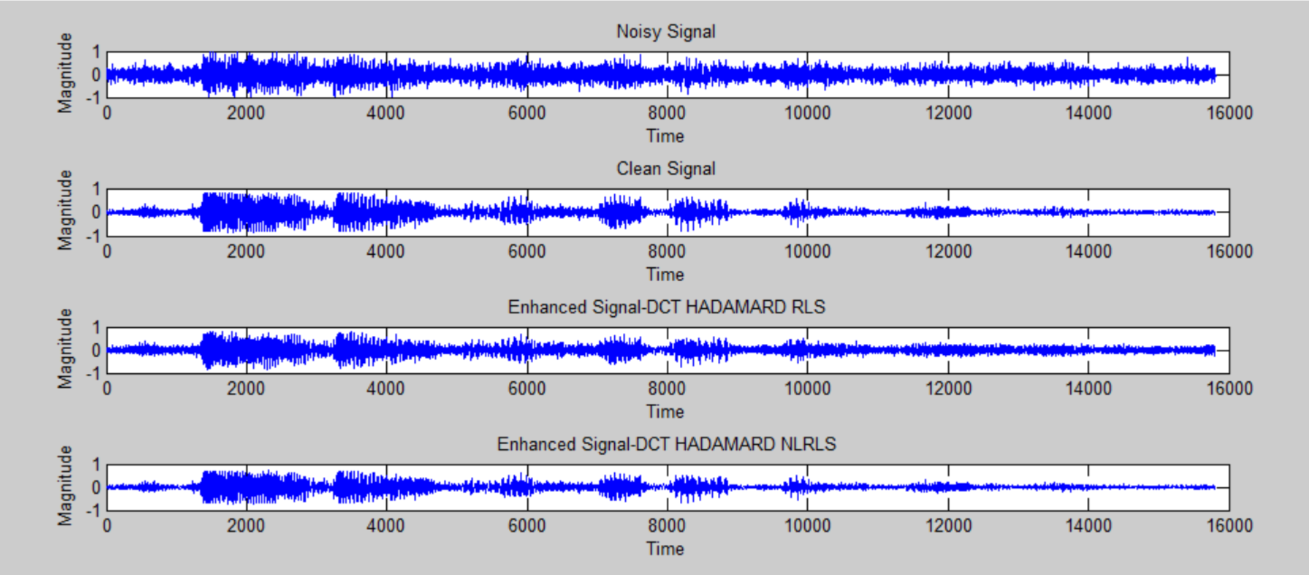

Figure 1 Time domain plot for Cellular 0 dB noisy signal

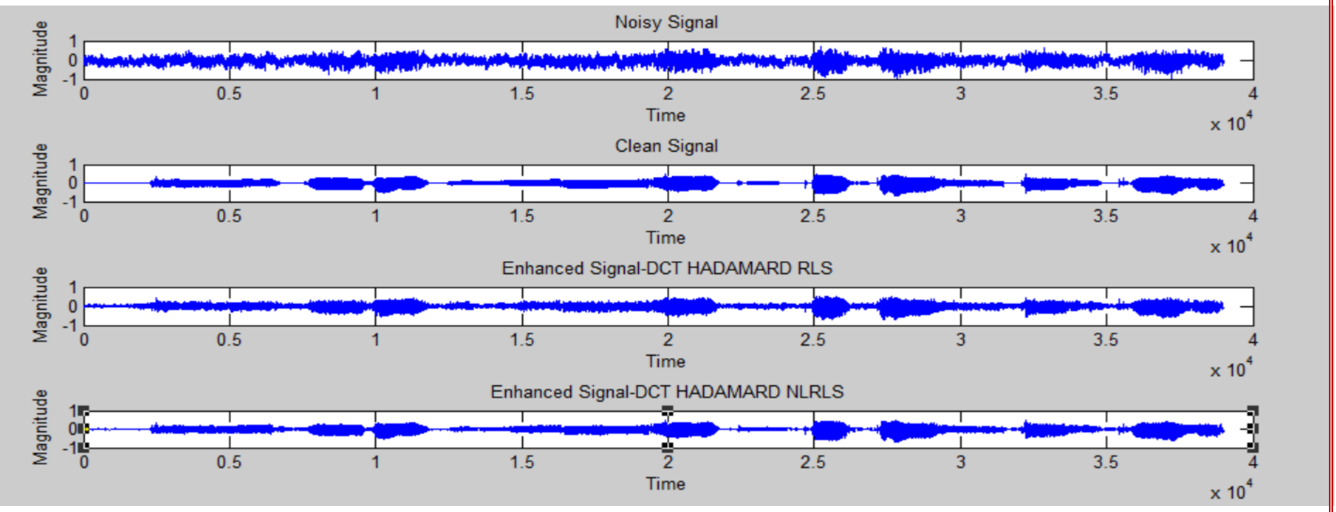

Figure 2 Time domain plot for Pink Stationary 0 dB noisy signal

Figure 1 and 2 shows the plots of the time domain results for Cellular 0 dB and pink stationary 0 dB noisy signals respectively of the proposed method. From the top, it is noisy signal, clean signal, signal enhanced by DCT- HADD-RLS algorithm and signal enhanced by proposed method respectively. In the proposed method the signal is closure to the clean signal.

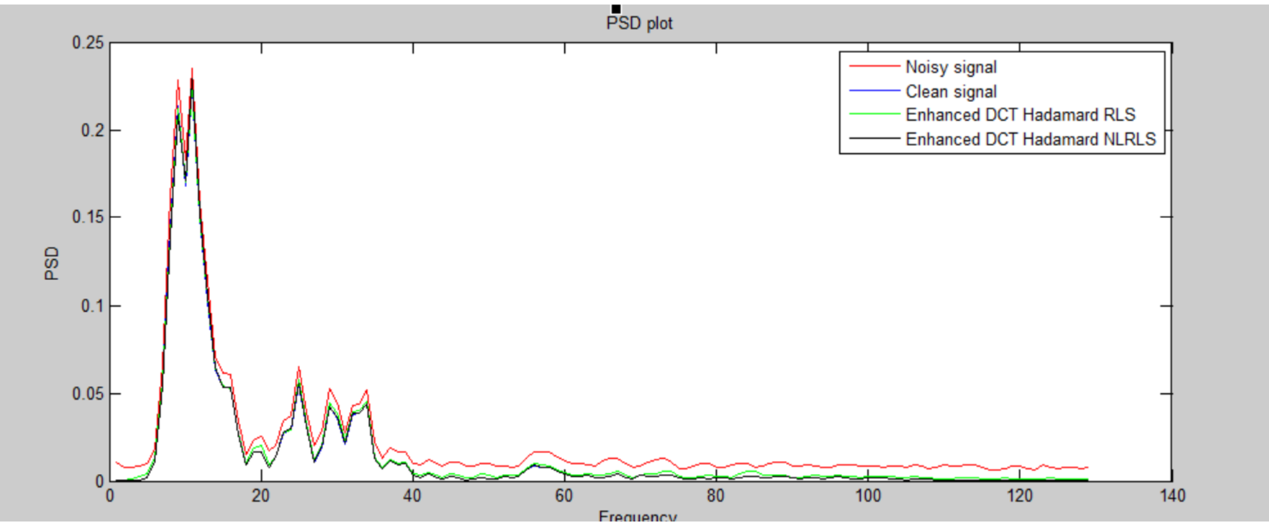

Figure 3 PSD plot for Pink Stationary 0 dB noisy signal

Figure 3 shows the Power spectral Density plot of the proposed method for Pink stationary 0db noisy signal. It shows that the clean and enhanced signals using proposed method are very close in their power spectrum

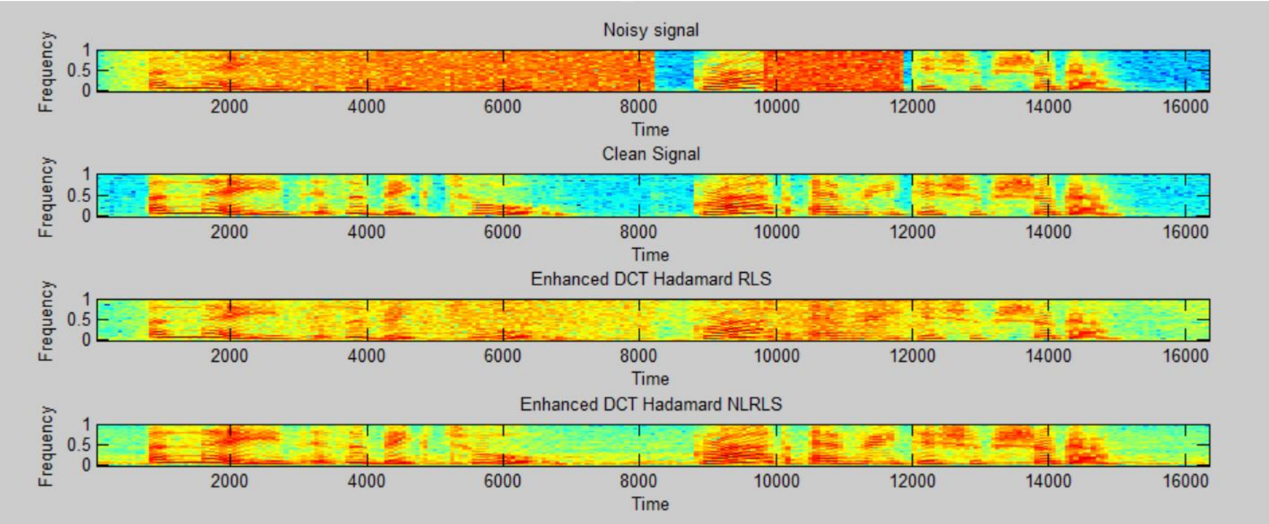

Figure 4 Spectrogram plot for Cellular 0dB noisy signal

Figure 4 shows the spectrogram of the proposed method and compared with the conventional method. The spectral components of the conventional method have some noise components with reduced speech level and in the proposed method it is similar to the clean signals spectrogram

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 10, October 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

The performances of the proposed dual channel speech enhancement algorithms are analyzed based on the objective and subjective quality measures.

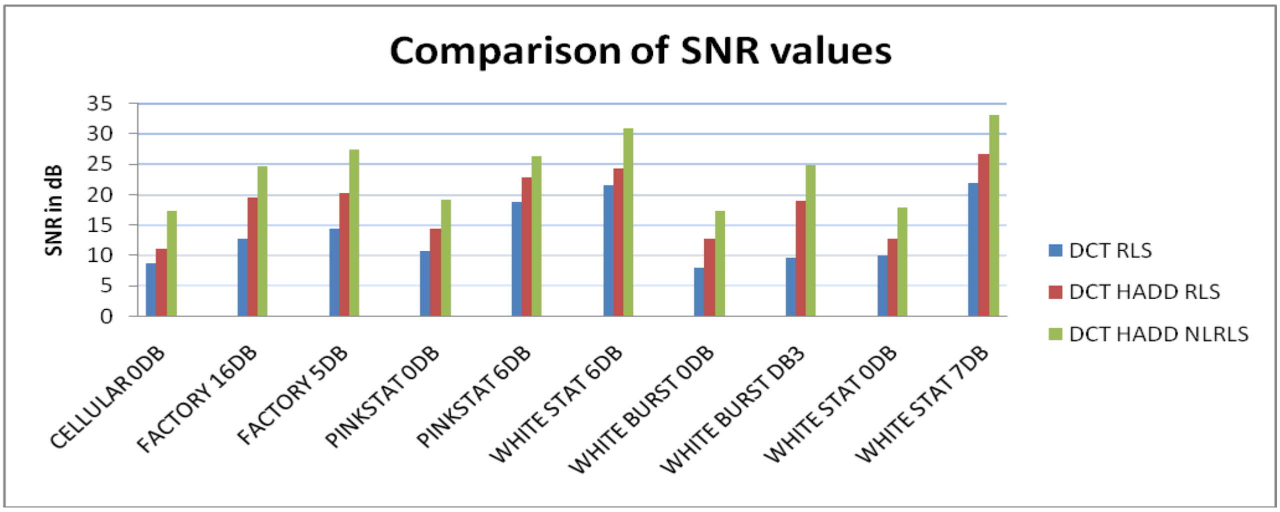

Figure 5 Comparison of SNR values

Figure 5 illustrates the SNR comparison of proposed method with DCT-RLS, DCT-HADD-RLS and DCT-HADD- NLRLS algorithms. From the results it is identified that SNR value of DCT-HADD-NLMS is increased compared to all other methods due to proposed normalized lattice algorithm and dual transform.

Table 1 gives the MSE values of the proposed dual channel speech enhancement algorithms for NOIZEUS database samples. It shows that the MSE values are reducing when the conventional adaptive algorithms are preprocessed with the Hadamard transform and further the LRLS algorithm is normalized.

Table 2 Comparison of mean square error values of proposed dual channel algorithms

Input Noisy Signal	Input SNR in dB	DCT RLS	DCT HADD RLS	DCT HADD NLRLS
CELLULAR 0DB	0	4.32E-03	3.20E-03	1.19E-03
FACTORY 16DB	16	6.42E-03	4.19E-03	3.30E-03
FACTORY 5DB	5	0.006513	0.00513	0.0029
PINKSTAT 0DB	0	4.26E-03	2.60E-03	1.70E-03
PINKSTAT 6DB	6	4.37E-03	3.70E-03	2.50E-03
WHITE STAT 6DB	6	5.32E-03	3.24E-03	1.90E-03
WHITE BURST 0DB	0	4.23E-03	2.30E-03	1.43E-03
WHITE BURST DB3	3	4.21E-03	2.12E-03	1.10E-03
WHITE STAT 0DB	0	5.24E-03	2.40E-03	1.47E-03
WHITE STAT 7DB	7	5.70E-03	2.57E-03	1.10E-03

V. CONCLUSION

In this work speech enhancement algorithm for dual channel environment is proposed and the performances of the adaptive filter using NLRLS algorithm is discussed. It is identified from the proposed method that adaptive filter operated in frequency domain input signals perform well than that of the input signals in time domain. Compared to the conventional Dual transformed RLS algorithm, SNR is improved better in the proposed method for different types of noisy environment. The mean square error also reduced in the proposed comparatively. With this performance analysis it is identified that the proposed algorithm can perform better than conventional RLS algorithm with faster convergence.

REFERENCES

[1.] A. Rix, J. Beerends, M. Hollier, and A. Hekstra, “Perceptual evaluation of speech quality (PESQ)-A new method for speech quality assessment of telephone networks and codecs,” in Proceedings of IEEE International Conference on Acoustics, speech and signal processing vol. 2, pp. 749–752, 2001.

[2.] Ben Gold, Nelson Morgon, “Speech and Audio Signal Processing”, John Wiley & Sons, 2000.

[3.] James D. Jhonston, “Transform Coding of audio signals using perceptual Noise criteria”, IEEE Journal on selected areas in communications, vol. *6*, no. 2, 1988.

[4.] L.R.Rabiner and R.W.Schafer, “Digital Processing of Speech Signals”. Pearson Publications 2004.

[5.] Monson H. Hayes, “Statistical Digital signal processing and Modelling”, John Wiley & Sons, 2005.

[6.] Philipos C. Loizou, “Speech Enhancement Theory and Practice”, CRC Press, Taylor& Francis, 2007.

[7.] Rankovic.C.M., “Factors governing speech reception benefits of adaptive linear filtering for listeners with sensorineural loss”. Journal of the Acoustical Society of America, 103(2). pp. 1043-57, 1998.

[8.] Shanks. J.E., Wilson. R.H.Larson, Williams. D.,“ Speech recognition performance of patients with sensorineural hearing loss under unaided and aided conditions using linear and compression hearing aids”. Ear hear. 23(4), pp. 280-90, 2002.

[9.] Shaul Florian and Neil J Bershad. “ A Weighted Normalized Frequency Domain LMS Adaptive Algorithm”. IEEE Transactions on Acoustics speech and signal processing. Vol. 36, no. 7, 1998.

[10.] Simon Haykin, “Adaptive Filter Theory”, Pearson Education Asia, 4th Edition, 2002.

[11.] Sunitha. S.L., and V. Udayashankara. “Fast Factored DCT-LMS Speech Enhancement for Performance Enhancement of Digital Hearing Aid”. World Academy of Science, Engineering and Technology, 10, 2005.

[12.] Sunitha S L and Dr.V Udayashankara,“ DFT-LMS Speech Enhancement Technique for Sensorinueural loss Patients”. Journal of Bioinformatics India, Vol 3, Jan- March 2005.

[13.] Yi Hu and Philipos C. Loizou, “Evaluation of Objective Quality Measures for Speech Enhancement”, IEEE Transactions on Audio, Speech and Language Processing, Vol. 16.No 1 PP. 229-238, 2008.

[14.] D. Deepa and Dr. A. Shanmugam (2010), “Dual Channel Speech Enhancement Using Hadamard-LMS Algorithm With DCT Preprocessing Technique” in International Journal of Engineering Science and Technology (IJEST), Vol. 2 (09) PP 4424 to 4430.

[15.] D. Deepa and Dr. A. Shanmugam (2009), “Time And Frequency Domain Analysis Of Subband Spectral Subtraction Method Of Speech Enhancement Using Adaptive Noise Estimation Algorithm ” in International Journal of Engg. Research & Indu. Appls. (IJERIA)., ISSN 0974-1518, Vol.2, No. VII, PP 57-72 .

[16.] D. Deepa and Dr. A. Shanmugam "Enhancement of Noisy speech signal using Hadamard-RLS algorithm with DCT preprocessing technique for digital hearing aid" 2011 IEEE International Conference on Computational Intelligence and Computing Research (2011 ICCIC) December 2011, Cape Institute of Technology, Levengipuram, Kanyakumari and Published in IEEE Explore.

[17.] WEBSITES FOR SPEECH SAMPLES

a. http://www.cslu.ogi.edu/nsel/data/SpEAR_database.html

b. http://www.utdallas.edu/~loizou/speech/noizeus/