Fast Factored DCT-LMS Speech Enhancement for Performance Enhancement of Digital Hearing Aid

Sunitha. S.L., and V. Udayashankara

Abstract—Background noise is particularly damaging to speech intelligibility for people with hearing loss especially for sensorineural loss patients. Several investigations on speech intelligibility have demonstrated sensorineural loss patients need 5-15 dB higher SNR than the normal hearing subjects. This paper describes Discrete Cosine Transform Power Normalized Least Mean Square algorithm to improve the SNR and to reduce the convergence rate of the LMS for Sensory neural loss patients. Since it requires only real arithmetic, it establishes the faster convergence rate as compare to time domain LMS and also this transformation improves the eigenvalue distribution of the input autocorrelation matrix of the LMS filter. The DCT has good ortho-normal, separable, and energy compaction property. Although the DCT does not separate frequencies, it is a powerful signal decorrelator. It is a real valued function and thus can be effectively used in real-time operation. The advantages of DCT-LMS as compared to standard LMS algorithm are shown via SNR and eigenvalue ratio computations. . Exploiting the symmetry of the basis functions, the DCT transform matrix [A$_N$] can be factored into a series of ±1 butterflies and rotation angles. This factorization results in one of the fastest DCT implementation. There are different ways to obtain factorizations. This work uses the fast factored DCT algorithm developed by Chen and company. The computer simulations results show superior convergence characteristics of the proposed algorithm by improving the SNR at least 10 dB for input SNR less than and equal to 0 dB, faster convergence speed and better time and frequency characteristics.

Keywords—Hearing Impairment, DCT Adaptive filter, Sensorineural loss patients, Convergence rate.

I. INTRODUCTION

HEARING impairment is the preamble chronic disability, affecting people in the world. Many people have great difficulty in understanding speech with background noise. This especially true for a large number of elderly peoples and sensorineural impaired persons. Hearing loss or deafness can be broadly classified into 2 types

Conductive loss: This is associated with a defect of the middle ear or eardrum (conductive mechanism).

This type of hearing disability can be measured by audiograms and is considered as a mild disability. Because, it attenuates the incoming acoustic signal without introducing any significant distortion. So the intelligibility of the signal can be easily resorted by amplification. Sensorineural loss: This is a broad class of hearing impairments its origin is in the cochlea or auditory nervous system. Sensorineural loss disorders are difficulty to remedy. This type of defects may be due to congenital or hereditary factors, disease, tumors, old age, long-term exposure to industrial noise, acoustic trauma or the action of toxic agents etc.

The sensorineural loss patient's experiences difficulty in making fine distinction between speech sounds, particularly those having a predominance of high frequency Energy [5], [16]. He may hear the speaker's voice easily, but be unable to distinguish, for example, between the words 'fat' and 'sat' [7], [9]. Two features of sensorineural impairment particularly detrimental to the perception of speech are high tone loss and compression of the dynamic range of the ear. A high tone loss is analogous to low pass filtering. Amplification of the high tones may improve intelligibility, but in these circumstances dynamic range of the ear is a handicap [13], [14]. Because dynamic range of the impaired ear may not be sufficient to accommodate the range of intensities in speech signals. So, the stronger components of speech are perceived at a level, which is uncomfortably loud, while the weaker components are not heard at all [10], [11], [16].

Several investigations on speech intelligibility have demonstrated that subjects with sensorineural loss patients need 5 to 15db higher SNR than the normal hearing subjects. While most of the defects in transmission chain up to cochlea can now days be successfully rehabilitated by means of surgery. The great majority of the remaining inoperable cases are sensorineural hearing impaired patients [5], [16].

Digital technology has made an important contribution in the field of audio logy. Digital signal processing methods offer great potential for designing a hearing aid but, today's Digital Hearing Aid are not up to the expectation for sensorineural loss patients. Hearing-impaired patients applying for hearing aid reveal that more than 50% are due to sensorineural loss. So for only Adaptive filtering methods are suggested in the literature for the minimization of noise from the speech signal for sensorineural loss patients [8].

Sunitha S L is with JSS Research Foundation, SJCE Campus, Mysore-570006, Karnataka, India (e-mail: sunithanov27@yahoo.co.in , phone: 08194-430498).

V. Udayashankara is with JSS Research Foundation as Research Scientist, SJCE campus, Mysore-570006, (e-mail: udaya@sjce.ac.in, phone: 0821-2512568).

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:1, No:10, 2007

A. Adaptive Filtering Method

The least mean square algorithm was first introduced by Widrow and Hoff in 1959 is simple, robust and is one of the most widely used algorithm for adaptive filtering. LMS algorithm is very popular because of its simplicity and easy of computations. LMS algorithm is generally the best choice for many different applications [18], [19]. This method can be effectively applied to reduce the noise i.e. to improve the SNR for sensorineural loss patients [6], [12], [15]. Unfortunately, its convergence rate is highly dependent on the feedback coefficient μ and the input power to the adaptive filter [18], [19]. The mean square error of an adaptive filter trained with LMS decreases over time as a sum of exponentials whose time constants are inversely proportional to the eigenvalues of the autocorrelation matrix of the filter inputs. Therefore, small eigenvalues create slow convergence modes in the means square error function. Large on the other hand, put a limit on the maximum learning rate that can be chosen without encountering stability problems [1]-[3].

In this work we use Discrete Cosine Transform Power Normalized Least Mean Square algorithm to improve the SNR and to reduce the convergence rate of the LMS for sensorineural loss patients. DCT-LMS algorithm is suited for non-stationary inputs like speech signals and the convergence time is also less compare to direct LMS techniques and Discrete Fourier Transform adaptive algorithms [17]. The DCT is a technique that converts a spatial domain waveform into its constituent frequency components as represented by a set of coefficients. Typically the DCT coefficients produced have most of the block's energy in a few frequency domain elements DCT is orthonormal, separable, frequency basis much like a Fourier transform [18].

The DCT has a strong energy compaction property. Most of the signal information tends to be concentrated in a few low frequency components of the DCT. It is a close relative of DFT – a technique for converting a signal into elementary frequency components, and thus DCT can be computed with a Fast Fourier Transform. Unlike DFT, DCT is a real valued and provides a better approximation of a signal with fewer coefficients. The DCT is central to many kinds of signal processing but DCT is mainly used in image processing application. DCT is very popular and used extensively in current image compression algorithms [18]. For non-stationary signals like speech signals, the DCT provides good approximation of a signal with fewer coefficients [3], [4].

B. Fast DCT Algorithm Developed by Chen, Smith and Fralick

The relationship between a given N point sequence $x(n)$ and the DCT of $x(n)$, $X(k)$ can be described in a matrix form as follows

$$X = \sqrt{\frac{2}{N}} [A_N] x \tag{1}$$

Where $x = [x(0), x(1),..........x(n),.....x(N-1)]^T$ is the vector form of the given N point sequence $x(n)$

$$\{x(n), X = [X(0), X(1),.....X(k),....X(N-1)]^T$$

is the N-th order of the DCT matrix. When N is a power of 2, the DCT matrix $[A_N]$ can be factorized into a product of sparse matrices.

There are different ways to obtain sparse matrix factorizations, resulting in different fast DCT algorithms. This work uses the fast DCT algorithm developed Chen and company. Chen, Smith and Fralick developed a fast DCT algorithm based on the following sparse matrix factorizations of the DCT matrix:

$$[A_N] = [\overline{P_N}] \begin{pmatrix} [A_{N/2}] & 0 \\ 0 & \overline{\overline{R_{N/2}}} \end{pmatrix} [B_N] \tag{2}$$

Where $[\overline{P_N}]$ is a permutation matrix, which permutes the even rows in decreasing order in the lower half. The matrix $[B_N]$ is a butterfly matrix:

$$[B_N] = \begin{pmatrix} I_{N/2} & \bar{I}_{N/2} \\ \bar{I}_{N/2} & I_{N/2} \end{pmatrix} \tag{3}$$

Where $[I_{N/2}]$ is the identity matrix of $\frac{N}{2} X \frac{N}{2}$, and $[\bar{I}_{N/2}]$ is the opposite identity matrix of $\frac{N}{2} X \frac{N}{2}$. The matrix $[\overline{\overline{R_{N/2}}}]$ is derived from the matrix $[R_N]$ by reversing the orders of both the rows and columns of $[R_N]$. The (i,k) element $r_{i,k}$, of the matrix $[R_N]$ is given by

$$r_{i,k} = \cos \frac{(2i+1)(2k+1)\pi}{4N} \tag{4}$$

The signal flow graph for N=8 is as shown in Fig. 2.

In section 1, we briefly discussed about the sensorineural loss patients and brief review about the convergence rate of the LMS adaptive algorithm and about Fast DCT algorithm. Section 2, describes LMS filtering in DCT domain. Simulated results are discussed in section 3 and section 4 concludes the paper.

II. DCT-LMS

DCT-LMS is composed of three stages as shown in Fig. 1.

Stage 1: Transformation

The input to the filter is $x_k = [x_k, x_{k-1},........., x_{k-n+1}]^T$.

The transformation operation is done with the help of Fast factored DCT by Chen and Company

$$u_k(n) = T_n[x_k] \tag{5}$$

The transform outputs then form a vector

$$u_k(n) = [u_k(0), u_k(1),........u_k(n-1)]^T$$

Stage 2: Power Normalization

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:1, No:10, 2007

The transformed signal $u_k(i)$ is then normalized by the square root of their power $p_k(i)$.

Where $i = 0, 1, \ldots\ldots n-1$.

The powers $p_k(i)$ can be estimated by the following methods:

The powers $p_k(i)$ can be estimated by filtering the $u_k^2(i)$ with an exponentially decaying window of parameter $\beta \in (0,1)$.

The powers $p_k(i)$ can also be estimated based on a sliding rectangular window or with the help of an arbitrary weighting filter.

In this work, power normalization is as follows.

Power normalizing $T_n x_k$ transforms its elements

$$(T_n x_k)(i) \text{ into } \frac{(T_n x_k)(i)}{\sqrt{Power of (T_n x_k)(i)}} . \tag{6}$$

Where the power of $(T_n x_k)(i)$ can be found on the main diagonal of B_n.

Then the power-normalized signal is

$$v_k(i) = \frac{u_k(i)}{\sqrt{p_k(i) + \varepsilon}} \tag{7}$$

Where $p_k(i) = \beta p_{k-1}(i) + (1-\beta) u_k^2(i)$ \hfill (8)

for $i = 0, 1, \ldots\ldots n-1$. The small constant ε is introduced to avoid numerical instabilities when $p_k(i)$ is close to zero.

The signals $v_k(i)$ are equal to the discrete cosine transformed outputs $u_k(i)$, but the learning constant μ in LMS filtering is replaced by a diagonal matrix whose elements are proportional to the inverse of the powers $p_k(i)$. This type of LMS is referred to as power-normalized LMS. Discrete cosine transformation followed by a power normalization stage, causes the eigenvalues of the LMS filter inputs to cluster around one and speeds up the convergence of the adaptive weights.

The autocorrelation matrix after transformation and power normalization is thus

$$S_n \approx E(diag B_n)^{-1/2} B_n (diag B_n)^{-1/2} . \tag{9}$$

If T_n decorrelated x_k exactly, B_n would be diagonal, S_n would be an identity matrix I_n, and all the eigenvalues of S_n would be equal to one, but since practically the DCT is not a perfect decorrelator, this does not work out exactly [2]. But the power normalization makes the eigenvalues of the LMS filter inputs to cluster around one and speeds up the convergence of adaptive weights.

The output vector after power normalization is

$$v_k(n) = [v_k(0), v_k(1), \ldots\ldots v_k(n-1)]^T \tag{10}$$

Stage 3: LMS filtering

The resulting equal power signals $v_k(i)$ are applied as an input to an adaptive linear combiner whose weights $w_k(i)$ are adjusted using LMS algorithm described below. The weight vector is defined as

$$w_k(n) = [w_k(0), w_k(1), \ldots\ldots w_k(n-1)]^T \tag{11}$$

Then the filter output is given by

$$y_k(n) = w_k^T(n) v_k(n) \tag{12}$$

and the instantaneous output error is

$$e_k = d_k - \sum_{i=0}^{n-1} y_n(i) . \tag{13}$$

Where d_k is the desired signal.

This error is used to update the adaptive filter taps using a modified for of the LMS algorithm

$$w_{k+1}(i) = w_k(i) + \mu e_k v_k(i) \tag{14}$$

for $i = 0, 1, \ldots\ldots n-1$.

The parameters used in algorithm are:
Desired signal is the speech sentence in English. Number of samples=20000, β=0.45
μ=0.075 and filter order=32.

III. SIMULATED RESULTS

The algorithm works on the corrupted speech signals with different types of noise signals like cafeteria noise, low frequency noise, babble noise etc. in several Signals to Noise Ratios. The various parameters like β, μ, and filter order were changed and their influence has been checked. For different input Signal to Noise Ratios the output Signal to Noise Ratios and convergence ratios are calculated. Although the SNR improvement has a limited meaning in the speech processing, we used this figure to indicate an over-all score. A more meaningful quantity is the eigenvalue Spread that is calculated to find out how well the algorithm convergence to the optimum Wiener solution. We have found that both the parameters SNR and convergence ratio are strongly depending on the number of samples in the input signal, β, μ and filter order. As the number of samples in the input signal increases SNR decreases and convergence ratio increases. Fig. 3, 4, 5 and 6 shows the input signal, that is corrupted signal, desired signal and the filtered signal for different Signal to Noise Ratio's of the input signal.

The Table 1 shows the Signal to Noise ratios of the DCT adaptive filtered outputs for different Signal to Noise ratios of the input signals and Table 2 shows the computational complexity of DCT and Factored DCT.

IV. CONCLUSION

Table I shows that the noise has been successfully removed from the input noisy signal and the speech quality is also good. The SNR improvement of at least 10 dB is obtained for the input Signal to Noise Ratios less than and equal to 0dB, which is higher than the other transformation techniques like DFT and Wavelet transforms [17], [20]. We have already stated that the filtering technique depends on the number of

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:1, No:10, 2007

samples in the input signal, β, μ and filter order. In this work, we have tested for only few values and their influence has been checked. By testing with few more different values, it may be possible to get further improvements. The algorithm convergence time and stability depends upon the ratio of the largest to the smallest eigenvalues associated with the correlation matrix of the input sequence. As the eigenvalue spread of the input autocorrelation matrix increases, the convergence speed of LMS deteriotes. So in this case, we derived the eigenvalue distribution for the input auto correlation matrix after DCT and power normalization. This provides the good tracking capabilities in different noisy environments. Even in the case of DFT-LMS and DWT-LMS, the eigenvalue distribution of the input autocorrelation matrix is calculated after the Transformation and power normalization. But, it is unable to give good SNR improvement and the convergence ratio is also very high [17], [20]. Proposed algorithm is not comparable with direct least mean square algorithm in terms of convergence ratio, where the eigenvalue ratio is in terms of thousands [6], [18].

REFERENCES

[1] 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, July 1998

[2] Francosie Beaufays. " Trasnform domain adaptive filters: An analytical approach". *IEEE Transactions on Signal processing, vol.* 43, no. 2, Feb 1995

[3] Mohammed A Shamma. " Improving the speed and performance of adaptive equalizers via transform based adaptive filtering ". 2591 Ashurst Rd. University Heights, Ohio, 44118, USA.

[4] *Adaptive filter theory.* By Simon Haykin.

[5] V.Udayashankara, A.P.Shivaprasad., " Digital Hearing Aid A Review", *World congress on Medical physics & Biomedical Engineerin*g. Brejil, Aug. 1994. pp. 21-26.

[6] V.Udayashankara , A.P.Shivaprasad. " The application of volt era LMS Adaptive filtering to speech enhancement for the Hearing Impairment ". *4th Euro-speech conference on speech communication and Technology,* Mandrid, Spain. Sept.1995. pp. 18-21.

[7] Moore. B.C., Stainsby. T.H., Alcantara. J.I., Kuhnel. V., "The effect of speech intelligibility of varying compression time constants in a digital hearing aid". *International Journal on Audio logy.* 2004 Jul-Aug: 43(7), pp. 399-409.

[8] Chug. "Challenges and recent developments in Hearing Aids", Trends *Amplif.*2004:8(3). Pp. 83-124.

[9] 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. 2002 Aug: 23(4), pp. 280-90.

[10] Baer.T., Moore.B.C., Kulk.K., " Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies". Journal on Acoustic soc Am. 2002 Sept: 112(3 pt 1)., pp. 1133-44.

[11] Hornsby.B.W., Ricketts.T.A., "The effects of compression ratio, signal-to-noise ratio and level on speech recognition in normal-hearing listners". Journal on Acoustics soc Am. 2001 June: 109(6). Pp. 2964-73.

[12] Shields.P.W., Campbell.B.R., " Improvements in intelligibility of noisy reverberant speech using a binaural sub band adaptive noise-cancellation processing scheme". Journal on Acoustics Soc Am. 2001Dec: 110(6). pp. 3232-42.

[13] Wouters.J., Litiere.L., Van Wieringen.A., " Speech intelligibility in noisy environment with one and two microphone hearing aids". Audiology. 1999 Mar-Apr: 38(2). Pp 91-8.

[14] Rankovic.C.M., " Factors governing speech reception benefits of adaptive linear filtering for listeners with sensorineural loss". Jpurnal on Acoustics Soc Am. 1998 Feb: 103(2). Pp. 1043-57.

[15] Baer. T., Moore.B.C., Gatechouse. S., "Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality, and response times". Journal on Rehabilitation Res Dev. 1993:30(1). Pp.. 49-72.

[16] Dr. Harry Levitt, " Noise reduction in Hearing aids: An overview ". Journal of Rehabilitation Research and Development, Vol.38. No.1, Jan-2001.

[17] 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.

[18] Simon Haykin, Adaptive Filter Theory", Pearson Education Asia, 4th Edition, 2002.

[19] Bernard Widrow and Samuel D.Stearns. " Adaptive Signal Processing", *Pearson Education Asia*, 2002.

[20] Sunitha S.L and Dr.V. Udayashankara. " DWT-LMS Speech Enhancement Technique for Performance Enhancement of Digital Hearing Aid", *ICSCI*, Jan 2005. (Best paper of the session has been taken for this paper)
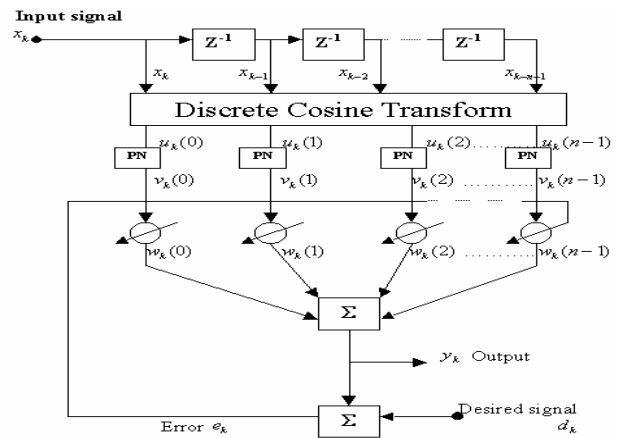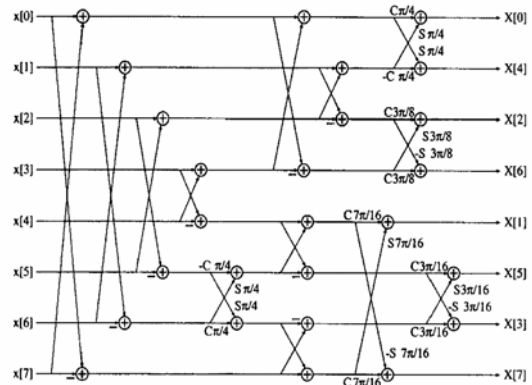
Fig. 1 Block Diagram of DCT-LMS algorithm

Fig. 2 Signal flow graph of fast DCT for N=8

TABLE I
OUTPUT SNR FOR ZERO DB INPUT SNR

Input SNR in dB	Output SNR in dB	Eigenvalue Ratio
-10	10.2	5.5
0	10.0	6.09
+5	11.24	5.44
+10	13.20	5.6

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:1, No:10, 2007

TABLE II
NUMBER OF MULTIPLICATIONS AND ADDITIONS REQUIRED FOR DIFFERENT DCT ALGORITHMS

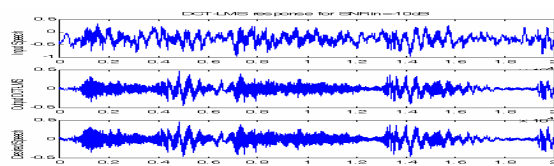| Type of transformation	No. of Multiplications	No. of Additions
DCT	64	56
Fast Factored DCT	13	29

Fig. 3 DCT response for input SNR= -10dB
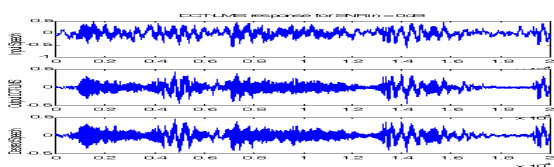
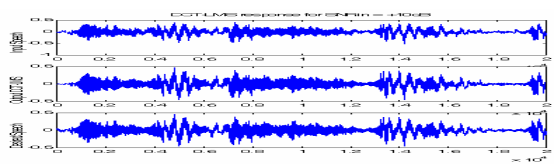Fig. 4 DCT response for input SNR= 0dB
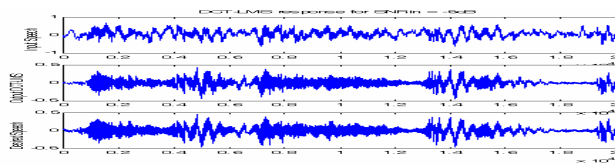

Fig. 5 DCT response for input SNR= +10dB

Fig. 6 DCT response for input SNR= -5dB

Sunitha S.L received the B.E. degree in Instrumentation Technology from Malnad College of Engineering, Hassan University of Mysore, India. M Tech. degree in Biomedical Engineering from Sri Jayachamarajendra College of Engineering, Mysore from Visveswaraiah Technological University, Karnataka. India.Currently she is a research scholar at Sri Jayachamarajendra College of Engineering, India and is teaching at SJMIT, Chitradurga. Her research interests include Digital Signal Processing, Speech processing, Neural Networks and Wavelet transforms.

V. Udayashankara received the B.E. degree in Electronics and Communication Engineering from Sri Jayachamarajendra College of Engineering, university of Mysore, India, in 1983, M.E. degree in Electrical Communication Engineering from the Indian Institute of Science, India, in 1988 and Ph.D. from Indian Institute of Science, India, in 1996.
Currently he is a Professor with the department of Instrumentation Technology, at Sri Jayachamarajendra College of Engineering, Visveswaraiah Technological University, India. His research interests include Rehabilitation Engineering, Digital Signal Processing, Speech recognition, Speech enhancement and EEG analysis. He has authored more than 50 publications in Journals, National and International Conferences in these areas.