SPEECH ENHANCEMENT BY FREQUENCY-WEIGHTED BLOCK LMS ALGORITHM

D. H. Cho and C. K. Un

Communications Research Laboratory
Department of Electrical Engineering
Korea Advanced Institute of Science and Technology
P.O. Box 150, Chongyangni
Seoul, Korea

Abstract

In this paper, enhancement of speech corrupted by additive white or colored noise is stuided. The unconstrained frequency-domain block least-mean-square (UFBLMS) adaptation algorithm and its frequency-weighted version are newly applied to speech enhancement. For enhancement of speech degraded by white noise, the performance of the UFBLMS algorithm is superior to the spectral subtraction method or Wiener filtering technique by more than 3 dB in segmented frequency-weighted signal-to-noise ratio ($FWSNR_{SEG}$) when SNR of speech is in the range of 0 to 10 dB.

As for enhancement of noisy speech corrupted by colored noise, the UFBLMS algorithm is superior to that of the spectral subtraction method by about 3 to 5 dB in $FWSNR_{SEG}$. Also, it yields better performance by about 2 dB in FWSNR and $FWSNR_{SEG}$ than that of the time-domain least-mean-square (TLMS) adaptive prediction filter (APF).

In view of the computational complexity and performance improvement in speech quality and intelligibility, the frequency-weighted UFBLMS algorithm appears to yield the best performance among various algorithms in enhancing noisy speech corrupted by white or colored noise.

I. INTRODUCTION

The enhancement of noisy speech has been investigated by many researchers. As a result, various enhancement methods have been suggested and developed. Lim and Oppenheim surveyed the previous studies on enhancement of speech degraded by additive noise [1]. Speech enhancement may be done by one of the following three approaches. The first approach is to exploit certain perceptual aspects of speech. By high-pass filtering fricative sound and inserting short pauses before plosive sound, significant improvement in intelligibility has been obtained [2]. Also, the short time spectral magnitude has been used to enhance noisy speech [3]. The second approach exploits the fact that voiced portion has quasi-periodicity. Since the energy of a periodic signal is concentrated in the repetitive frequency bands and the interfering signal has in general energy over the entire frequency bands, comb filtering can reduce noise, while preserving the desired signal [4]. The third approach to speech enhancement is to exploit a speech production model. The parameters of the speech model are estimated, and then enhanced speech is generated by the synthesis system based on the same speech model or with the estimated speech model parameters [5]. In addition, to reduce narrowband noise, the use of a time-domain filter has been investigated. It may be formed from the inverse transform of the inverse of the estimated noise spectrum. This filter can be implemented using a time-domain noise suppression filter that is adapted segmentally based on samples of background noise [6].

In this paper, a new enhancement technique using the unconstrained frequency-domain block least-mean-square (UFBLMS) algorithm is proposed for speech corrupted by white or colored noise. The performance of the newly proposed method will be compared to those of existing enhancement algorithms. To test the effectiveness of each enhancement algorithm, we use objective measures such as frequency-weighted signal-to-noise ratio (FWSNR) and segmented FWSNR ($FWSNR_{SEG}$) that are closely correlated with perception [7]. Following this introduction, in Section II the new UFBLMS algorithm for enhancement of noisy speech is introduced. In this section, we will also discuss how one can apply this algorithm to enhance noisy speech corrupted by white or colored noise. In Section III, computer simulation is done to investigate the performance of various algorithms for enhancement of noisy speech. The performance improvements resulting from the use of various enhancement techniques are compared by objective quality measures. Finally, in Section IV, conclusions are made.

II. THE UFBLMS ALGORITHM WITH OR WITHOUT FREQUENCY WEIGHTING

Let us consider a transversal adaptive digital filter (ADF) operated on the block-by-block basis. Prior to derivation of the frequency-weighted UFBLMS algorithm, we briefly review the UFBLMS ADF. Let M, L, and N be the number of filter weights, the block length, and the transform length of FFT, respectively. The UFBLMS ADF shown in Fig. 1 can be obtained by minimizing the frequency-domain block mean-squared error (BMSE). In the UFBLMS ADF, the frequency-domain error vector $\mathbf{e}_k$ in the k_{th} block is given by

$$\mathbf{e}_k = \mathbf{d}_k - \mathbf{P}_{0,L}(\mathbf{X}_k \mathbf{W}_k) \tag{1}$$

where $\mathbf{d}_k$ and $\mathbf{W}_k$ are the (Nx1) desired response and filter weight vectors, respectively, both in the frequency domain, and $\mathbf{X}_k$ is an (NxN) diagonal matrix whose diagonal elements are the transformed input data. In (1), the (NxN) matrix $\mathbf{P}_{0,L}$ realizes the sectioning procedures needed for computing the filter output, and is defined as

$$\mathbf{P}_{0,L} \triangleq \mathbf{F} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \tag{2}$$

where F is an (NxN) discrete Fourier transform matrix, $\mathbf{I}_L$ denotes an (LxL)

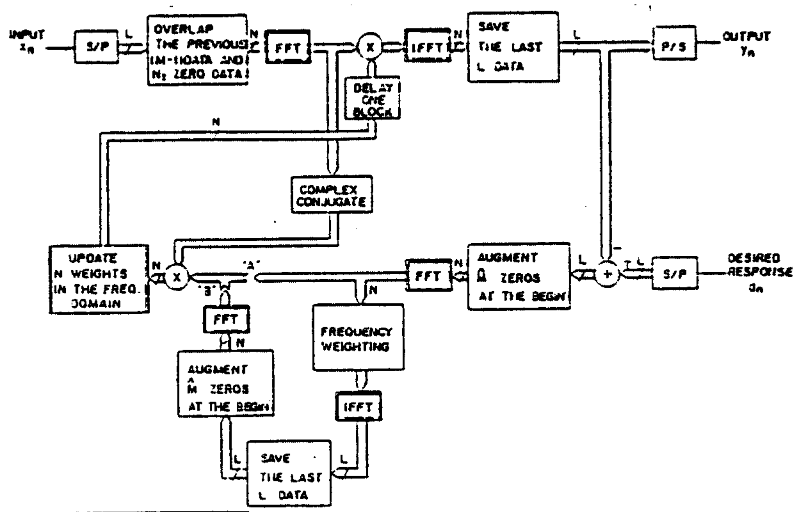

Fig. 1 Realization of UFBLMS and frequency weighted UFBLMS ADF's using FFT and overlap-save sectioning procedure ($N=L+M-1+N_z$, $\hat{L}\underline{\Delta}L-1+N_z$, $\hat{M}\underline{\Delta}M-1+N_z$, and $N_z \geq 0$).
[Note: UFBLMS ADF is realized with the position "A" connected and frequency-weighted UFBLMS ADF is realized with the position "B" connected. N_z is the number of zero data needed for augmenting the input data, thereby allowing to choose a suitable transform of length N. S/P = serial-to-parallel conversion and P/S = parallel-to-serial conversion.]

identity matrix, and $\mathbf{0}$ is a zero matrix.

As a performance criterion in adjusting the filter weights, we use the frequency-weighted block MSE ε^{FW} defined by [8]

$$\varepsilon^{FW} \underline{\Delta} E[\mathbf{e}_k^* \mathbf{H} \mathbf{e}_k] \tag{3}$$

where the asterisk and $E[\cdot]$ denote complex-conjugate transpose of a matrix and statistical expectation, respectively. In (3), $\mathbf{H}$ is an (NxN) diagonal matrix whose diagonal elements are of nonnegative values, and then magnitudes represent the relative significance of each frequency component. Following the same approach for the UFBLMS ADF [9], we can obtain from (1) and (3) a gradient of the frequency-weighted block MSE with respect to $\mathbf{W}_k$ as

$$\nabla\varepsilon^{FW}(\mathbf{W}_k) \underline{\Delta} \frac{\partial \varepsilon^{FW}}{\partial \mathbf{W}_k} = -2E[\mathbf{X}_k^* \mathbf{P}_{0,L} \mathbf{H} \mathbf{e}_k] \tag{4}$$

Thus, using an instantaneously estimated gradient, we obtain from (4) a frequency-weighted UFBLMS weight adjustment algorithm as the following:

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \mathbf{X}_k^* \mathbf{P}_{0,L} \mathbf{H} \mathbf{e}_k \tag{5}$$

where μ is a convergence factor that controls the convergence behavior of the algorithm. In Fig. 1, a block diagram of the frequency-weighted UFBLMS ADF using the algorithm of (5) is shown together with that of the UFBLMS ADF. It is noted that, when $\mathbf{H}$ is an identity matrix, the frequency-weighted UFBLMS algorithm becomes identical to the UFBLMS algorithm since $\mathbf{P}_{0,L}\, \mathbf{e}_k = \mathbf{e}_k$. Also, it is noted that, when L is sufficiently larger than M, $\mathbf{P}_{0,L}$ can be approximated as an identity matrix. In that case, one can eliminate the FFT and inverse FFT operation that are needed just after the frequency weighting operation in the frequency-weighted UFBLMS ADF. In this work, we apply the UFBLMS and frequency-weighted UFBLMS algorithms discussed above to the enhancement of speech degraded by white or colored noise. Frequency weighting is done by using different convergence factors for each frequency component.

For noisy speech corrupted by white noise, enhanced speech is obtained by filtering the noisy speech through the UFBLMS or frequency-weighted UFBLMS ADF. This is shown in Fig. 2-(a). Note that, unlike the conventional algorithms such as the spectral subtraction method and the Wiener filtering method, the proposed enhancement algorithm requires no speech/silence discrimination. Hence, the computational complexity of the UFBLMS algorithm is simpler than those of the conventional enhancement algorithms.

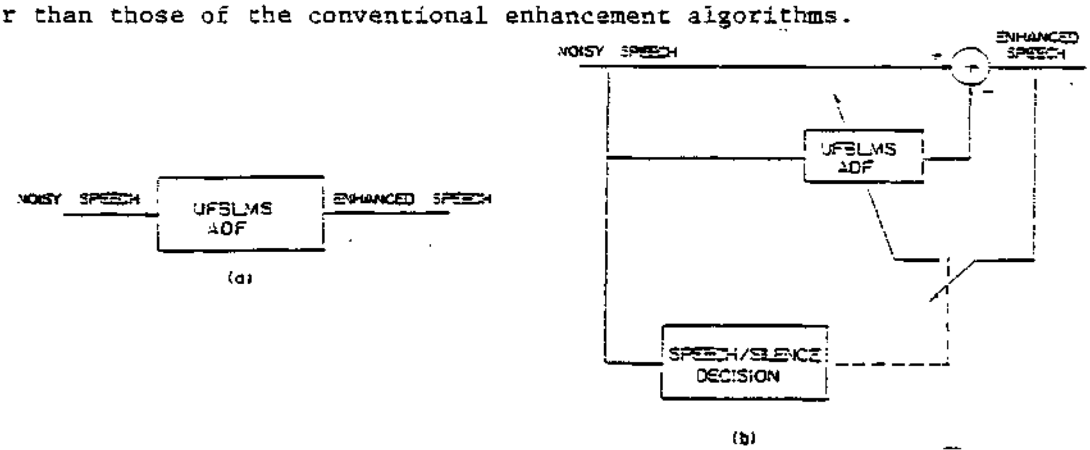

Fig. 2 Systems for enhancement of noisy speech by the UFBLMS algorithm.
(a) For speech corrupted by white noise
(b) For speech corrupted by colored noise

As for noisy speech corrupted by colored noise, we obtain enhanced speech by using a noise suppression filter based on the UFBLMS algorithm as shown in Fig. 2-(b). In this scheme, when noise in a silence interval is detected, the adaptive algorithm adjusts the weights of the noise suppression filter such that the error signal is minimized. When speech is detected, the adaptation algorithm is turned off, and the values of filter weights are kept at their current values. Adaptation resumes whenever speech activity no longer occurs.

III. SIMULATION RESULTS AND DISCUSSION

In this section we first investigate the performance of the proposed UFBLMS algorithm by simulation when noise is white, and compare it to those of the existing enhancement techniques, such as the spectral subtraction method [3] and the Wiener filtering method [1], by various objective measures. As the input to these systems, real speech bandlimited to 3.4 kHz and sampled at 8 kHz was used. To obtain noisy speech, we generated white Gaussian noise using a random number generation program. We then processed it by a low-pass filter whose 3 dB cutoff frequency was 3.4 kHz, and added the resulting noise to the clean speech.

Table I shows the results of speech enhancement by various enhancement algorithms. We can see from this table that the improvement resulting from the use of the UFBLMS algorithm is similar to that of the frequency-weighted UFBLMS algorithm. Also, it is noted that the performance could be improved significantly for noisy speech with the UFBLMS algorithm. Perhaps, the reason may be due to the fact that the UFBLMS algorithm is based on forward and backward prediction, and enhancement is done in the frequency-domain. Also, we can note from Table I that the improvement by various enhancement algorithms decreases as the SNR of input noisy speech becomes higher. The reasons are thought to be due to the nonstationary characteristics of speech and also because noise spectrum is estimated approximately. When frequency-weighted objective measures are used as performance criteria, we can see from Table I that the UFBLMS algorithm is superior to the spectral subtraction method or Wiener filtering technique by more than 3 dB in $FWSNR_{SEG}$.

Table I. Improvement resulting from enhancement algorithms for noisy speech corrupted by white noise.

Enhancement Algorithm	Input SNR / Measure	0 dB	5 dB	10 dB
UFBLMS	FWSNR	11.38	14.25	14.40
	$FWSNR_{SEG}$	7.03	11.52	13.45
Frequency-weighted UFBLMS	FWSNR	11.79	14.20	14.40
	$FWSNR_{SEG}$	6.87	11.28	13.45
Spectral subtraction with Hamming window	FWSNR	9.87	12.78	16.23
	$FWSNR_{SEG}$	3.51	6.43	10.11
Wiener filtering	FWSNR	10.56	13.32	16.68
	$FWSNR_{SEG}$	4.43	7.21	10.62

Note: Noisy speeches of 0, 5 and 10 dB in SNR correspond to those of 5.0, 8.35 and 12.69 dB in FWSNR, and -1.53, 1.81 and 6.15 dB in $FWSNR_{SEG}$, respectively.

Fig. 3 shows LPC spectra of clean, noisy and enhanced speech by various enhancement algorithms. It can be seen that the LPC spectra of the enhanced speech by the UFBLMS algorithm approximates the LPC spectral envelope of clean speech most closely among the three enhancement techniques, especially in the frequency range of 1 to 3 kHz. Also, it is worthwhile to mention that, with the UFBLMS algorithm, high-pass filtering may be combined with the enhancement algorithm to improve speech quality and intelligibility. That is, high-pass filtering can be done simultaneously with enhancement in the frequency domain. In this case, different convergence factors may be used for each frequency component.

Next, we investigate the performance of the UFBLMS enhancing algorithm when the noise is colored, and compare it to those of the existing enhancement techniques, such as the spectral subtraction method [3] and the adaptive prediction filtering (APF) method [6], by objective quality measures.

To obtain noisy speech corrupted by colored noise, we generated white Gaussian noise. Then, we processed it by a band-pass filter, and added the resulting noise to clean speech. Fig. 4 shows the average noise spectrum of which narrow-band ridges correspond to the fundamental (1550 Hz) and first harmonic (3100 Hz) narrow-band noise of the helicopter engine [3]. For speech/silence discrimination which is required as a part of the enhancement algorithm for the colored noise case, we used a speech/silence detection method based on spectral magnitude, power and autocorrelation of segmented speech [10].

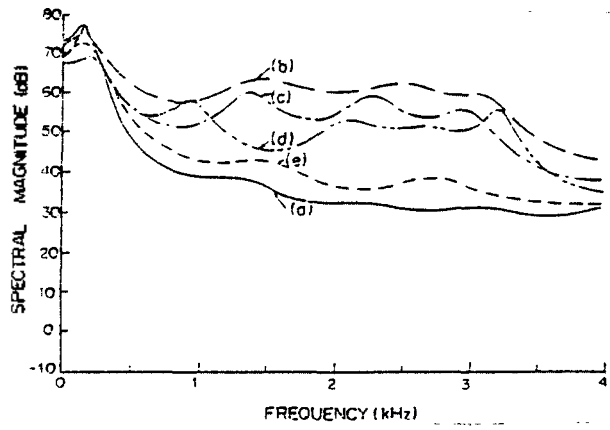

Fig. 3 Spectral envelopes of clean, noisy and enhanced speech.
(a) Original clean speech (b) 5 dB noisy speech
(c) Enhanced speech by the spectral subtraction method with Hamming window
(d) Enhanced speech by the Wiener filtering
(e) Enhanced speech by the UFBLMS algorithm

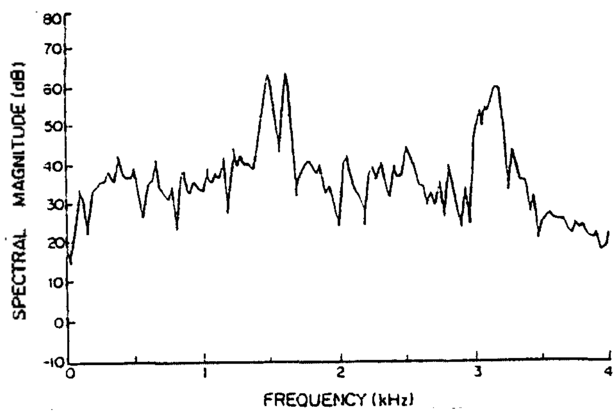

Fig. 4 Spectrum of colored noise.

Table II. Improvement resulting from enhancement algorithms for noisy speech corrupted by colored noise.

Enhancement Algorithm \ Measure \ Input SNR		0 dB	5 dB	10 dB
UFBLMS	FWSNR	14.27	16.42	18.77
	$FWSNR_{SEG}$	8.50	10.68	13.07
Spectral subtraction with Hamming window	FWSNR	6.49	9.71	13.33
	$FWSNR_{SEG}$	1.24	4.54	8.60
Adaptive prediction filtering	FWSNR	11.32	14.14	16.41
	$FWSNR_{SEG}$	6.38	9.06	12.34

Note: Noisy speeches of 0, 5 and 10 dB in SNR correspond to those of 3.74, 7.00 and 11.29 dB in FWSNR, and -2.04, 1.22 and 5.51 dB in $FWSNR_{SEG}$, respectively.

Table II shows the improvement that results from the use of various enhancement algorithms. When the FWSNR and $FWSNR_{SEG}$ measures are used as performance criteria, the improvement by the UFBLMS algorithm is more than 7 dB. Also, we have found that the performance of the UFBLMS algorithm is almost the same as that of the frequency-weighted UFBLMS algorithm. However, according to our intelligibility test, the frequency-weighted UFBLMS algorithm appears to be more effective for improvement of speech intelligibility. It is noted that the performance of the UFBLMS algorithm is superior to that of the spectral subtraction method by about 3 to 5 dB in FWSNR and $FWSNR_{SEG}$. Also, the performance of the UFBLMS algorithm is better by about 2 dB in FWSNR and $FWSNR_{SEG}$ than that of the APF algorithm. In addition, Fig. 5 shows LPC spectra of clean speech, noisy speech and enhanced speech by various enhancement algorithms. It is seen from this figure that the spectral envelope by the UFBLMS algorithm is closer to the spectral envelope of clean speech than other enhancement techniques.

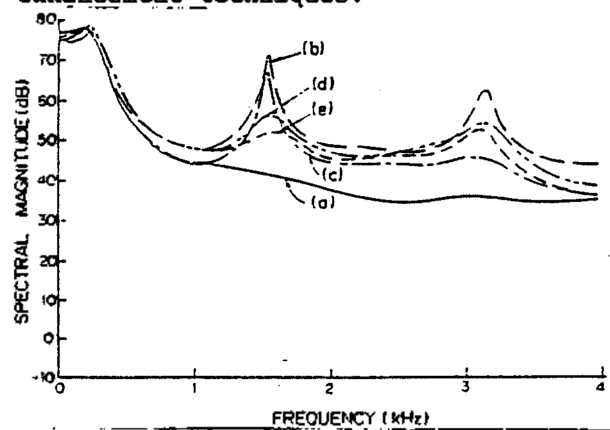

Fig. 5 Spectral envelopes of clean, noisy and enhanced speech.

(a) Original clean speech
(b) 5 dB noisy speech corrupted by colored noise
(c) Enhanced speech by the spectral subtraction method with Hamming window
(d) Enhanced speech by the TLMS adaptive prediction filter
(e) Enhanced speech by the UFBLMS algorithm (β=0.9)

V. CONCLUSIONS

In this paper, a new technique using the UFBLMS algorithm has been proposed to enhance noisy speech degraded by white or colored noise, and evaluated by various objective measures. According to the simulation results for speech corrupted by white noise, the UFBLMS algorithm is superior to the spectral subtraction method or Wiener filtering technique by more than 3 dB in $FWSNR_{SEG}$ when the SNR of speech is in the range of 0 to 10 dB. In general, the improvement decreases as the SNR of input speech becomes higher. With the UFBLMS algorithm, high-pass filtering may be combined with the enhancement algorithm to improve the speech quality and intelligibility further. For degraded speech by colored noise, the performance of the UFBLMS algorithm is superior to that of the spectral subtraction method by about 3 to 5 dB in FWSNR and $FWSNR_{SEG}$. Also, the performance of the UFBLMS algorithm is better by about 2 dB in FWSNR and $FWSNR_{SEG}$ than that of the TLMS APF algorithm. In view of the computational complexity and improvement in speech quality and intelligibility, it can be concluded that the frequency-weighted UFBLMS algorithm yields the best results among various algorithms so far proposed in enhancing noisy speech corrupted by white or colored noise.

REFERENCES

1. J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," Proc. IEEE, vol. 67, pp. 1586-1604, Dec. 1979.
2. H. Drucker, "Speech processing in a high ambient noise environment," IEEE Trans. Audio Electroacoust., vol. AU-16, pp. 165-169, June 1968.
3. S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-27, pp. 113-120, Apr. 1979.
4. R. H. Frazier et al., "Enhancement of speech by adaptive filtering," in Proc. 1976 IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 251-253, Apr. 1976.
5. J. S. Lim and A. V. Oppenheim, "All-pole modeling of degraded speech," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-26, pp. 197-210, June 1978.
6. L. Hoy et al., "Noise suppression methods for speech applications," in Proc. 1983 IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 1133-1136, Apr. 1983.
7. S. G. Wilson and S. Husain, "Adaptive tree encoding of speech at 8000 bits/s with a frequency weighted error criterion," IEEE Trans. Commun., vol. COM-27, pp. 165-170, Jan. 1979.
8. J. C. Lee, C. K. Un and D. H. Cho, "A frequency-weighted block LMS algorithm and its application to speech processing," Proc. IEEE, vol. 73, pp. 1173-1174, June 1985.
9. J. C. Lee and C. K. Un, "Performance of transform-domain LMS adaptive digital filters," submitted to IEEE Trans. Acoust., Speech, Signal Processing.
10. C. K. Un and H. S. Kim, "Speech/silence discrimination of noisy speech," in Proc. of Japan-Korea Joint Symposium on Acoustics, pp. 363-366, Nov. 1983.