# 주파수 영역 블록 LMS 알고리즘을 이용한 잡음이 섞인 음성의 음질개선

# Enhancement of Noisy Speech by Frequency-Domain Block LMS Algorithm


* 조 동 호 (Cho, D.H.)

** 은 종 관 (Un, C.K.)



## 요 약

본 논문에서는 광대역 혹은 협대역 잡음이 섞인 음성의 음질을 향상시키기 위해서 빠른 수렴속도를 갖고 있는 UFBLMS 알고리즘을 음성처리에 응용한다. 광대역 잡음이 섞인 음성인 경우에는, 입력음성의 SNR이 0 dB에서 10 dB 사이 일때, UFBLMS 알고리즘의 성능이 spectral subtraction 방법이나 Wiener filtering 방법보다도 $FWSNR_{SEG}$ 척도로 약 3 dB 더 좋음을 알 수 있다.

협대역 잡음이 섞인 음성인 경우에는, UFBLMS 알고리즘의 성능이 spectral subtraction 방법보다 $FWSNR_{SEG}$ 척도로 약 3 dB에서 5 dB정도 더 좋다. 또한 UFBLMS 알고리즘은 TLMS 적응예측 방법보다 FWSNR 및 $FWSNR_{SEG}$ 척도로 약 2 dB정도 성능이 더 좋다.

여러 음질 향상 알고리즘의 계산상의 복잡도와 음질 향상도 및 인식도를 고려해보면 frequency weighting 기능을 갖고 있는 UFBLMS 알고리즘을 사용하는 것이 바람직함을 알 수 있다.

## ABSTRACT

In this paper, enhancement of speech corrupted by additive white or colored noise is studied. The unconstrained frequency-domain block least-mean-square (UFBLMS) adaptation algorithm and its frequency-weighted version are newly applied to speech enhancement, and their performances for enhancement are analyzed. 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.



* 과학기술원 통신공학연구실 선임연구원

** 과학기술원 전기 및 전자공학과 교수

## 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], [4]. In these algorithms, enhanced speech is obtained by combining the estimated spectral magnitude of speech with the short time spectral phase of the degraded speech.

The second approach exploits the fact that voiced portion has quasi-periodicity. The periodicity of a waveform is considered as harmonics with the fundamental frequency corresponding to the period of the time waveform in the frequency domain. 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 [5], [6]. Although this approach is applicable potentially to many different types of additive noise, it requires accurate pitch estimation which is difficult in the presence of noise.

The third approach to speech enhancement is to exploit a speech production model. In this case, speech is modeled by the response of an all-pole or pole-zero linear system representing the vocal tract driven by an excitation function. 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 [7], [8]. Although the performance of these systems has not been formally evaluated, some improvement in speech quality but little improvement in intelligibility have been reported. 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 [9].

The speech enhancement schemes discussed above have been applied to one input degraded by additive noise. If there are more than one input available for processing, further enhancement is possible. Each individual input may be processed separately using the speech enhancement system, and then the processed inputs may be appropriately combined. Examples are the adaptive noise cancelling algorithms in which correlation of noise in several inputs is exploited [10], [11]. Although these algorithms enhance noisy speech dramatically, and adapt well to changing noise statistics, they have some limitations in practical applications. For example, if the reference input contains some signal component, the noise canceller attempts to cancel the input signal as well as noise.

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 [12].

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, the relative gain by the use of the UFBLMS algorithm for enhancement of noisy speech is examined. In Section IV, 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. In addition, we consider the computational complexity of each enhancement algorithm. Finally, in Section V, conclusions are made.

## II. THE UFBLMS ALGORITHM WITH OR WITHOUT FREQUENCY WEIGHTING

Before discussing the frequency-weighted UFBLMS algorithm, we first consider a UFBLMS ADF. Realization of the UFBLMS ADF is shown in Fig. 1. Here, we assume that the impulse response of the UFBLMS ADF consists of M weights, and that the stationary input signal is processed block-by-block, each block having L data. In the following discussion, we use 'k' for the block index. Also, we use the notations $x_n$, $y_n$ and $d_n$ for the filter input, output and desired response, respectively. The UFBLMS ADF can be obtained by minimizing the frequency-domain block mean-squared error (FBMSE) defined as

$$\mathrm{FBMSE} = \varepsilon^{FB} \triangleq E[\mathbf{e}_k^* \mathbf{e}_k] \tag{1}$$

where the column vector $\mathbf{e}_k$ of N elements is the discrete Fourier transform (DFT) of the augmented time-domain error vector in the $k_{th}$ block. To obtain the error vector, the filter output is computed by linear convolution of the input and the time-domain weights. This linear convolution may be computed by the fast convolution approach using the N-point fast Fourier transform (FFT) and the overlap-save sectioning method. The length N for the FFT is $N=L+M-1+N_Z$, where $N_Z$ is the number of appended zeros.

In the UFBLMS ADF, the frequency-domain error vector $\mathbf{e}_k$ in the k-th block is given by

$$\mathbf{e}_k = \mathbf{d}_k - \mathbf{P}_{0,L}(\mathbf{X}_k \mathbf{W}_k) \tag{2}$$

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 (2), the (NxN) matrix $\mathbf{P}_{0,L}$ realizes the sectioning procedure needed for computing the filter output, and is defined as

$$\mathbf{P}_{0,L} \triangleq \mathbf{F} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^{-1} \tag{3}$$

where F is an (NxN) discrete Fourier transform matrix, $\mathbf{I}_L$ denotes and (LxL) identity matrix, and $\mathbf{0}$ is a zero matrix.

As a performance criterion in adjusting the filter weights, we use the frequency-weighted block $MSE_\varepsilon{}^{fw}$ defined by [13]

$$\varepsilon^{FW} \triangleq E[\mathbf{e}_k^* \mathbf{H} \mathbf{e}_k] \tag{4}$$

where the asterisk and E[·] denote complex-conjugate transpose of a matrix and statistical expectation, respectively. In (4), H is an (NxN) diagonal matrix whose diagonal elements are of nonnegative values, and their magnitudes represent the relative significance of each frequency component. Following the same approach used for the UFBLMS ADF [14], we can have from (2) and (4) a gradient of the frequency-weighted block MSE with respect to $W_k$ as

$$\nabla\varepsilon^{FW}(W_k) \triangleq \frac{\partial \varepsilon^{FW}}{\partial W_k} = -2E[X_k^* P_{0,L}\, He_k] \qquad (5)$$

Thus, using an instantaneously estimated gradient, we obtain from (5) a frequency-weighted UFBLMS weight adjustment algorithm as the following:

$$W_{k+1} = W_k + \mu\, X_k^* P_{0,L}\, He_k \qquad (6)$$

where $\mu$ is a convergence factor that controls the convergence behavior of the algorithm. In Fig. 1, a block diagram of the frequency-

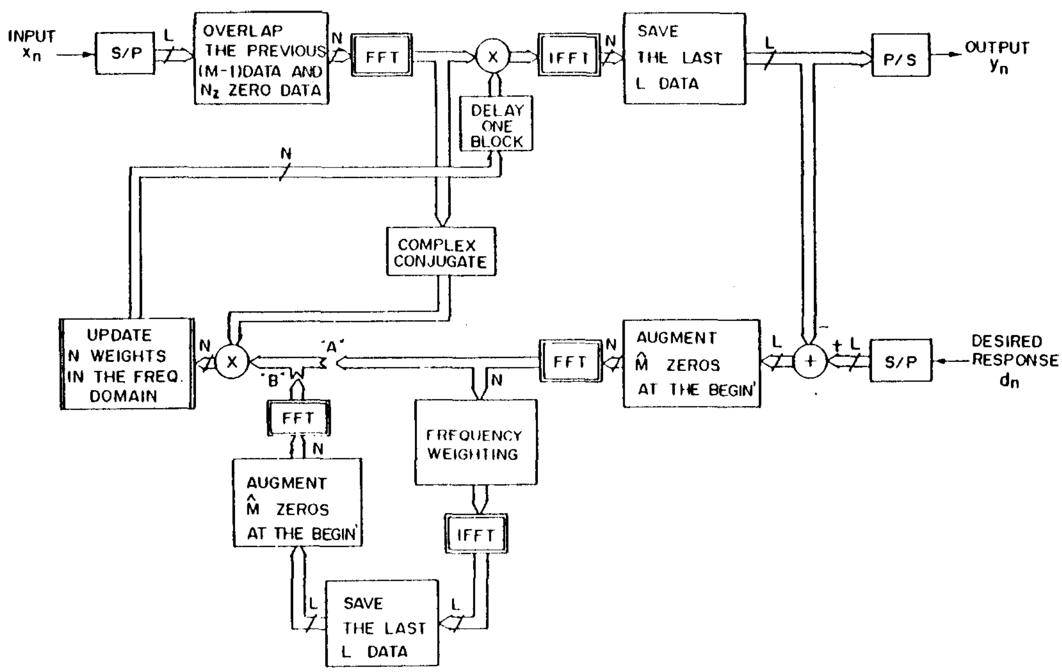


Fig. 1. Realization of UFBLMS and frequency-weighted UFBLMS ADF's using FFT and overlap-save sectioning procedure ($N=L+M-1+N_Z$, $\hat{L} \triangleq L-1+N_Z$, $\hat{M} \triangleq M-1+N_Z$, and $N_Z \geq 0$).

[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.]

weighted UFBLMS ADF using the algorithm of (6) is shown together with that of the UFBLMS ADF. It is noted that, when H is an identity matrix, the frequency-weighted UFBLMS algorithm becomes identical to the UFBLMS algorithm since $P_{0,L}\ e_k = e_k$. Also, it is noted that, when L is sufficiently larger than M, $P_{0,L}$ can be approximated as an identity matrix. In that case, one can eliminate the FFT and inverse FFT operations 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.

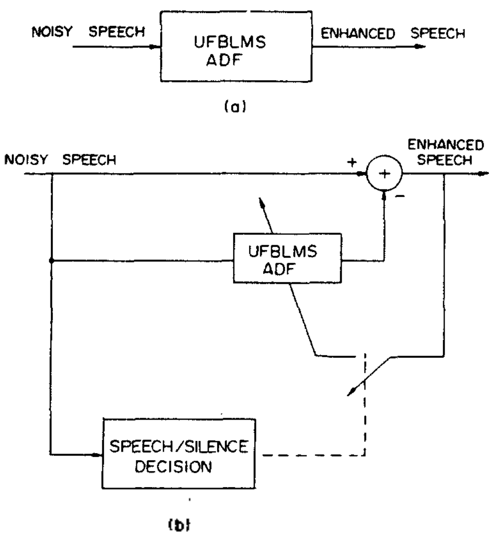


Fig. 2. Systems for enhancement of noisy speech by the UFBLMS algorithm.
(a) For speech corrupted by white noise
(b) For speech corrupted by colored noise

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. PERFORMANCE OF THE UFBLMS ENHANCEMENT ALGORITHM

We study first the performance improvement obtainable by using a UFBLMS ADF for noisy speech corrupted by white noise. The noisy signal degraded by additive white noise, $x_w(k)$, may be expressed as

$$x_w(k) \triangleq s(k) + n_w(k), \qquad k=1, 2, \cdots \tag{7}$$

where $s(k)$ and $n_w(k)$ are the $k_{th}$ samples of clean speech and white noise, respectively. Here, we assume that $s(k)$ and $n_w(k)$ are mutually uncorrelated.

We can estimate $s(k)$ from M observations of the noisy speech as [15]

$$\hat{s}(k) \simeq \sum_{\substack{\ell=-\frac{M}{2} \\ \ell \neq 0}}^{\frac{M}{2}} a_\ell\, x_w(k-\ell), \qquad k=1, 2, \cdots \tag{8}$$

where the coefficients $\{a_\ell\}$ are chosen to minimize the mean-squared interpolation error given by

$$\varepsilon(m) \triangleq E[x_w(k) - \sum_{\substack{\ell=-\frac{M}{2} \\ \ell \neq 0}}^{\frac{M}{2}} a_\ell x_w(k-\ell)]^2 \tag{9}$$

Using the orthogonality principle, one can show that the coefficients $\{a_\ell\}$ are the solution of the Wiener-Hopf equations

$$\sum_{\substack{\ell=-\frac{M}{2} \\ \ell \neq 0}}^{\frac{M}{2}} a_\ell \, \rho(k-\ell) = \rho(k) \quad k = \pm 1, \pm 2, \cdots, \pm\frac{M}{2} \tag{10}$$

where $\rho(k)$ is the convariance function of noisy speech $x_w(k)$. Once the coefficients $\{a_\ell\}$ are determined, we can get the estimate $\{s(k)\}$ from (8).

To investigate the performance gain resulting from the use of the UFBLMS enhancement algorithm discussed above, we define the improvement factor $\eta$ as

$$\eta \triangleq \frac{SNR_e}{SNR_{ne}} \tag{11}$$

Here, $SNR_{ne}$ is the SNR of speech prior to enhancement, and is given by

$$SNR_{ne} \triangleq \frac{\sum_{j=1}^{J} s^2(j)}{\sum_{j=1}^{J} n_w^2(j)} \tag{12}$$

where J is the number of total speech samples. And, $SNR_e$ is the SNR after enhancement, and is expressed as

$$SNR_e \triangleq \frac{\sum_{j=1}^{J} s^2(j)}{\sum_{j=1}^{J} \{s(j) - \sum_{\substack{\ell=-\frac{M}{2} \\ \ell \neq 0}}^{\frac{M}{2}} a_\ell s(j-\ell) - \sum_{\substack{\ell=-\frac{M}{2} \\ \ell \neq 0}}^{\frac{M}{2}} a_\ell n_w(j-\ell)\}^2} \tag{13}$$

From (11), (12) and (13), we can note that the improvement ratio $\eta$ becomes larger as the estimated speech approximates real speech more accurately.

Next, let us consider the performance gain by a UFBLMS noise suppression filter for speech corrupted by colored noise. The noisy speech corrupted by colored noise, $x_c(k)$, may be represented by

$$x_c(k) \triangleq s(k) + n_c(k), \quad k = 1, 2, \cdots \tag{14}$$

where $s(k)$ and $n_c(k)$ are the samples of clean speech and colored noise, respectively. Again we assume that $s(k)$ and $n_c(k)$ are mutually uncorrelated.

From M observation of noisy speech, we can estimate $n_c(k)$ as follows:

$$\hat{n}_c(k) \simeq \sum_{\substack{\ell=-\frac{M}{2} \\ \ell \neq 0}}^{\frac{M}{2}} a_\ell x_c(k-\ell) \quad k = 1, 2, \cdots \tag{15}$$

where the interpolation coefficients $\{a_\ell\}$ are chosen to minimize the mean-squared error given by

$$\varepsilon(m) \triangleq E[x_c(k) - \sum_{\substack{\ell=-\frac{M}{2} \\ \ell \neq 0}}^{\frac{M}{2}} a_\ell x_c(k-\ell)]^2 \tag{16}$$

where the coefficients $\{a_\ell\}$ are the solution of the Wiener-Hopf equations.

Once the coefficients are determined, we can obtain the estimate $\{n_c(k)\}$ from (15). Then, we obtain enhanced speech $s(k)$ from (14). Also, using $\{a_\ell\}$, we can form a transversal filter (see Fig. 2-(b)) for suppression of the colored noise as the following:

$$A(z) \triangleq 1 - \sum_{k=-\frac{M}{2}}^{\frac{M}{2}} a_k z^{-k} \tag{17}$$

We can note that the tap weights $\{H(\ell)\}_{\ell=-\frac{M}{2}}^{\frac{M}{2}}$ for the impulse response of the noise suppression filter are related to the interpolation coefficients $\{a_\ell\}$ by

$$H(0)=1,\ H(\ell)=-a_\ell,\quad \ell=\pm1,\pm2,\cdots,\pm\frac{M}{2}. \tag{18}$$

Thus, the enhanced speech is given by

$$\hat{s}(k)=\sum_{\ell=-\frac{M}{2}}^{\frac{M}{2}} H(\ell)s(k-\ell)+\sum_{\ell=-\frac{M}{2}}^{\frac{M}{2}} H(\ell)n_c(k-\ell),\quad k=1,2,\cdots \tag{19}$$

Consequently, from (11) and (19), the improvement factor $\eta$ is expressed as

$$\eta=\frac{\sum_{j=1}^{J} n_c^2(j)}{\sum_{j=1}^{J}\{s(j)-\sum_{\ell=-\frac{M}{2}}^{\frac{M}{2}} H(\ell)s(j-\ell)-\sum_{\ell=-\frac{M}{2}}^{\frac{M}{2}} H(\ell)n_c(j-\ell)\}^2} \tag{20}$$

From (20), it may be noted that the performance gain of the UFBLMS algorithm can be larger than that of the prediction filter using various LMS algorithms. The reason is due to the fact that, unlike the APF algorithm which uses only backward information, the UFBLMS algorithm is based on forward and backward prediction. In other words, the colored noise can be suppressed more by an interpolation filter than by a prediction filter.

## IV. 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 meausres. 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 prgoram. 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 can be improved significantly for noisy speech by the UFBLMS algorithm. Perhaps, the reason may be due to the fact that 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

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

| Enhancement Algorithm | Measure \ Input SNR | 0 dB | 5 dB | 10 dB |
|---|---|---|---|---|
| UFBLMS | FWSNR | 11.88 | 14.26 | 14.40 |
| | $FWSNR_{SEG}$ | 7.03 | 11.52 | 13.45 |
| Frequency-weighted UFBLMS | FWSNR | 11.79 | 14.30 | 14.40 |
| | $FWSNR_{SEG}$ | 6.87 | 11.38 | 13.45 |
| Spectral subtraction with Hamming window | FWSNR | 9.87 | 12.78 | 16.38 |
| | $FWSNR_{SEG}$ | 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.

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 1 that the UFBLMS algorithm is superior to the spectral subtraction method or Wiener filtering technique by more than 3 dB in $FWSNR_{SEG}$.

Fig. 3 shows waveforms of clean speech, 5 dB noisy speech and enhanced speech by various enhancement algorithms. We can note from this figure that the enhanced waveform by the UFBLMS algorithm is closer to the clean speech than those obtained by the other two enhancement techniques. Also, we can note that, unlike

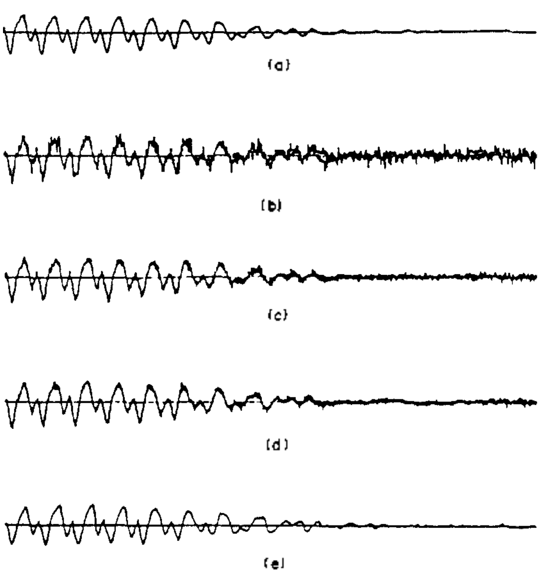


Fig. 3. Waveforms of clean, noisy and enhanced speech.
(a) Original clean speech
(b) 5 dB noisy speech
(c) Enhanced speech by the spectral subtraction method with Hamming window
(d) Enhanced speech by the Wiener filtering
(e) Enhanced speech by the UFBLMS algorithm

other existing methods, noise suppression by the UFLBMS algorithm is done very effectively.

In addition, Fib. 4 shows linear predictive coding (LPC) spectra of clean, noisy and en-

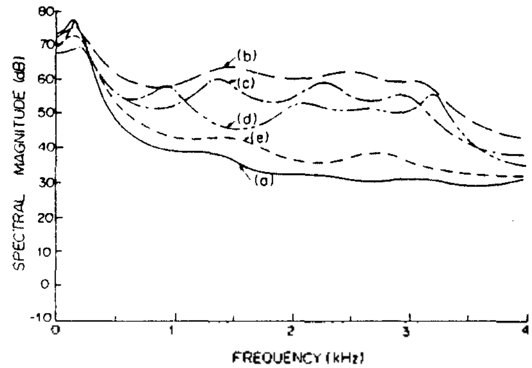


Fig. 4. Spectral envelopes of clean, noisy and enhanced speech.
(a) Original clean speech
(b) 5 dB noisy speech
(c) Enhanced speech by the spectral subtraction method with Hamming window
(d) Enhanced speech by the Wiener filtering
(e) Enhanced speech by the UFBLMS algorithm

hanced 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, highpass 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 [9], 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. 5 shows the average noise spectrum of which narrowband ridges correspond to the fundamental (1550 Hz) and first harmonic (3100 Hz) narrowband 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 [16].

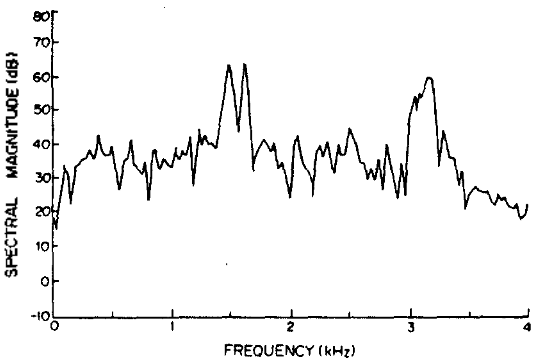


Fig. 5. Spectrum of colored noise.

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.

Table II. Performance 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 | 13.83 |
| | $FWSNR_{SEG}$ | 1.34 | 4.54 | 8.60 |
| Adaptive prediction filtering | FWSNR | 11.82 | 14.14 | 16.41 |
| | $FWSNR_{SEG}$ | 6.38 | 9.06 | 12.04 |

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.

Fig. 6 shows waveforms of clean speech, noisy speech and enhanced speech by various enhancement algorithms. It can be seen from this figure that the waveform enhanced by the UFBLMS algorithm is closer to clean speech in comparison with other enhancement techniques. Also, Fig. 7 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.

Finally, let us consider the computational

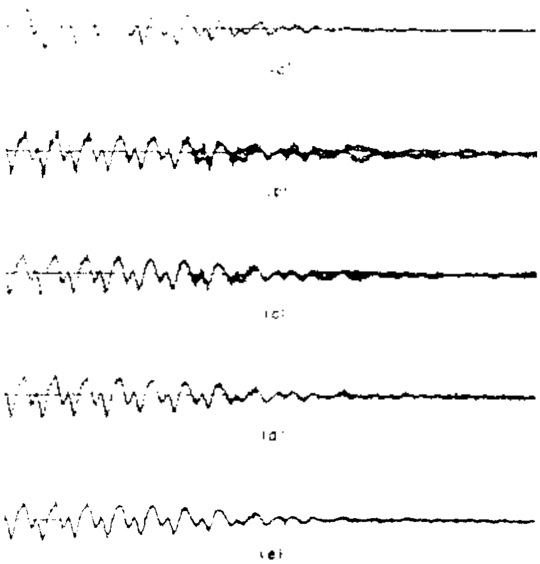

Fig. 6. Waveforms of clean, noisy and enhanced speech.
(a) Original clean speech
(b) 5 dB noisy speech degraded by colored noise
(c) Enhanced speech by the spectral subtraction method with Hamming window
(d) Enhanced speech by the TLMS adaptive prediction filter
(e) Enhanced speech by the UFBLMS algorithm ($\beta$=0.9)

complexity of each enhancement algorithm in view of application to a narrow-band speech compression system. For convenience, we assume that the speech enhancement subsystem is cascaded with a vocoder, and that noise characteristics such as power spectrum and prediction coefficients are known. Hence, we do not consider the computational complexity for obtaining noise characteristics.

Let N, M and L be the frame length, the number of prediction coefficients and the output data length, respectively. Then, we can obtain the computational complexities of various algorithms as shown in Tables III through VI. In these Tables we can see that the UFBLMS algorithm requires far less number of computations than the spectral subtraction method or the Wiener filtering technique. For enhancement of noisy speech corrupted by white noise, we can see that, unlike the existing

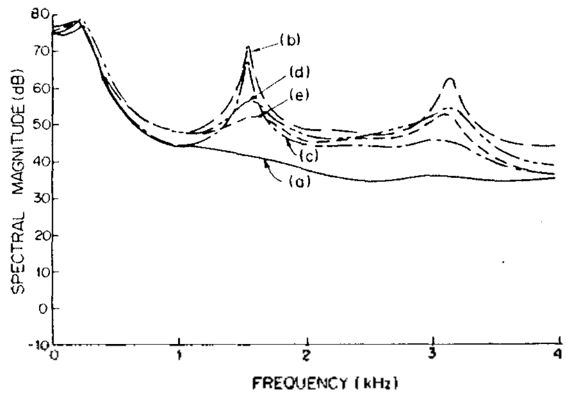


Fig. 7. Spectral envelopes of clean, noisy and enhanced speech.
(a) Original clean speech
(b) 5 dB noisy speech corrupted by colored noise
(c) Enhanced speech by the spectral subtraction method with Hamming window
(d) Enhanced speech by the TLMS adaptive prediction filter
(e) Enhanced speech by the UFBLMS algorithm ($\beta$=0.9)

Table III. Number of multiplications required for the UFBLMS algorithm enhancing noisy speech corrupted by white noise (N=252, M=10 and L=243).

| Operation | Multiplications | |
|---|---|---|
| UFBLMS DFT | 3 N point DFT | 2544 |
| Computation of output in the frequency domain | $4 \times (\frac{N}{2})$ | 504 |
| Computation of estimated gradient in the frequency domain | $4 \times (\frac{N}{2})$ | 504 |
| Adjustment of weights | $\frac{N}{2}$ | 126 |
| Total | | 3678 |

Note: For DFT, we use the N point DFT by the nested algorithm [17].

Table IV. Number of multiplications required for the spectral subtraction method. (N=252)

| Operation | Multiplications | |
|---|---|---|
| DFT of noisy speech | N point DFT | 848 |
| Square of noisy speech spectrum | 2N | 504 |
| Square root of enhanced spectrum (Square root $\simeq$ 20 multiplications) | 10N | 2520 |
| Phase information (Division $\cong$ 10 multiplications) | 5N | 1260 |
| Enhanced speech in the frequency domain | N | 252 |
| Inverse DFT of enhanced speech in the frequency domain | N point DFT | 848 |
| Total | | 6232 |

Table V. Number of multiplications required for Wiener filtering based on all-pole modeling. (N=252, M=10)

| Operation | Multiplications | |
|---|---|---|
| Autocorrelation of noisy speech | (M+1) N | 2772 |
| Calculation of prediction coefficients | $M^2$ | 100 |
| Spectral envelope of noisy speech | N point DFT | 848 |
| Optimum filter design (Division $\cong$ 10 multiplications) | 5N | 1260 |
| Enhanced speech spectrum | N | 252 |
| Enhanced speech waveform | N point DFT | 848 |
| Total | | 6080 |

Table VI. Number of multiplications required for noise suppression filter using TLMS or UFBLMS algorithm (M=10, L=243).

| Operation | Multiplications | |
|---|---|---|
| Enhanced speech | L x M | 2430 |
| Total | | 2430 |

enhancement algorithms, the UFBLMS algorithm does not require speech/silence detection for which the amount of computations can be fairly large. Moreover, high-pass filtering can be combined into the frequency-weighted UFBLMS algorithm to enhance noisy speech corrupted by white or colored noise. Hence, in view of the computational complexity and improvement in speech quality and intelligibility, it is preferred to use the frequency-weighted UFBLMS algorithm for enhancement of noisy speech.

## 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. Also, the improvement by the UFBLMS algorithm for noisy speech has been considered analytically.

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 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. C.K. Un and K.Y. Choi, "Improving LPC analysis of noisy speech by autocorrelation subtraction method," in Proc. 1981 IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 1082-1085, Mar. 1981.
5. 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.
6. J.S. Lim, A.V. Oppenheim, and L.D. Braida, "Evaluation of adaptive comb filtering method enhancing speech degraded by white noise addition," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-26, pp. 354-358, Aug. 1978.
7. 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.
8. H. Kobatake, J. Inari, and S.I. Kakuta, "Linear predictive coding of speech signals in a high ambient noise environment," in Proc. 1978 IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 472-475, June 1978.
9. 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.
10. B. Widrow et al., "Adaptive noise cancelling: principles and applications," Proc. IEEE, vol. 63, pp. 1692-1716, Dec. 1975.
11. S.F. Boll and D.C. Pulsipher, "Suppression of acoustic noise in speech using two microphone adaptive noise cancellation," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-28, pp. 752-753, Dec. 1980.
12. 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.
13. 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. 1137-1138, June 1985.
14. J.C. Lee and C.K. Un, "Performance of transform-domain LMS adaptive digital filters," submitted to IEEE Trans. Acoust., Speech, Signal Processing.
15. E. Masry, "Closed-form analytical results for the rejection of narrow-band interference in PN spread-spectrum systems — part II: linear interpolation filters," IEEE Trans. Commun., vol. COM-33, pp. 10-19, Jan. 1985.
16. 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.

17. D.P. Kolba and T.W. Parks, "A prime factor FFT algorithm using high speed convolution," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-25, pp. 281-294, Aug. 1977.