

음성신호의 단일입력 적응잡음제거

A Single Channel Adaptive
Noise Cancellation for Speech Signals

강 해 동*, 배 건 성**
(Hae Dong Gahng, Keun Sung Bae)

요 약

음성신호에 내재한 배경잡음을 제거하는 단일입력 적응잡음제거 시스템을 구성하였다. 기존 방법에서는 프레임 단위로 분석된 음성신호의 피치 정보를 이용하여 적응여파기의 기준신호를 얻는데 비해 제안된 방법에서는 매 샘플마다 지연 정보를 추정하여 기준신호를 만든다. 입력되는 음성신호로부터 매 샘플시간마다 지연 정보를 구하기 위하여 일반적인 자기상관 함수와 평균절대차 함수로부터 재귀적 자기상관함수와 재귀적 평균절대차함수를 유도하였다. 정규화된 최소평균자승(NLMS) 적응알고리즘을 사용하는 단일입력 잡음제거 시스템에 제안된 지연추정 방법을 적용하여 백색 가우시안 잡음에 왜곡된 음성에 대해 음성개선 실험을 하였으며, 기존 방법과의 성능비교 실험을 하였다. 제안된 방법에 의한 음성개선이 기존 방법보다 음질 및 SNR 면에서 더 좋은 결과를 보였다.

ABSTRACT

A single channel adaptive noise canceling (ANC) technique is presented for removing effects of additive noise on the speech signal. The conventional method obtains a reference signal using the pitch estimated on a frame basis from the input speech. The proposed method, however, gets the reference signal using the delay estimated recursively on a sample by sample basis. To estimate the delay, we derive recursion formula of autocorrelation function and average magnitude difference function. The performance of the proposed method is evaluated for the speech signals distorted by the additive white Gaussian noise. Experimental results with normalized least mean square (NLMS) adaptive algorithm demonstrate that the proposed method improves the perceived speech quality quite well besides the signal-to-noise ratio.

*창신전문대학교 전자통신과

*Department of Electronic Communication, Changshin Junior College

**경북대학교 전자공학과

**Department of Electronics, Kyungpook National University

접수일자: 1993년 9월 20일

I. 서 론

정보화 사회에서는 사람간의 기존 음성통신 외에 컴퓨터 등이 개입되는 사람과 기계간의 음성통신 형태가 상당한 비중을 지니게 된다. 음성통신 분야에

사용되는 기본적인 기술에 관련된 연구의 대부분은 주로 주변잡음이 없거나 무시할 수 있을 정도의 양질의 음성을 대상으로 연구가 진행되어 왔는데 잡음 환경의 영향을 고려하지 않은 음성신호처리 시스템의 성능은 실지 환경에서 급격히 감소하게 된다. 또한 잡음이 있는 환경하에서는 정확한 음성정보의 전달에 어려움이 따르므로 청취성(listenability)도 떨어진다. 따라서 수신된 음성신호에서 잡음을 경감시키는 음성개선(speech enhancement) 기술에 대한 연구 및 개발이 꾸준히 진행되고 있다[1-15].

음성개선을 위해 제안된 기법은 크게 3가지 범주로 나눌 수 있다. 첫번째 범주에 속하는 기술은 짧은 구간의 스펙트럼 추정(short-time spectral estimation)에 기초를 둔 방법으로, 스펙트럼 감산법(spectral subtraction)[4, 5]과 Wiener 여파기를 이용한 잡음 제거 방법[6, 7] 등이 이에 속한다. 이들 방법은 추정된 잡음의 스펙트럼을 입력 잡음신호에서 빼거나 깨끗한 음성의 스펙트럼을 추정하고 이를 noncausal Wiener filter에 적용해 잡음이 제거된 음성을 재합성하는 방법으로서, 미리 잡음이나 원하는 음성의 통계적 특성이 잘 알려져 있을 때 적합하다. 두번째 범주에 속하는 기술은 음성신호의 통계적 모델에 기초하여 hidden Markov model(HMM)이나 expectation and maximization(EM) 등의 기술을 음성개선에 응용하는 방법이다[8-10]. 음성신호가, 통계적으로 볼 때, 일련의 독립적인 부음원(subsource)으로 구성되며, 이들 각각의 부음원이 가장 가까운 음을 생성하는 특정한 상태(state)를 대표한다고 가정하고, 음성신호를 HMM 등으로 모델링하는 방법이다. 이 방법은 잡음의 영향이 비교적 크지 않은 입력 SNR이 10에서 30dB 사이의 경우에 좋은 효과를 볼 수 있으나 [8] 음성신호와 잡음신호의 표준 모델을 미리 구해 두어야 한다. 세번째 범주의 기술은 음성신호의 준주기(quasi-periodic)적 특성을 이용하는 방법으로 comb filtering[11]과 적응여파기를 이용한 잡음제거 기술 [12, 13] 등이다. 다른 범주의 방법들과는 달리, 적응잡음제거 기술은 음성신호에 포함된 잡음신호의 상세한 통계적 특성을 미리 알고 있지 않아도 되는 장점이 있다. 적응잡음제거 방법은 입력 음성신호에 내재된 잡음과 높은 상관관계를 갖는 별도의 잡음신호를 필요로 하는데 이 잡음신호를 기준입력(reference input)이라고 부른다.

음성신호에 대한 적응잡음제거 시스템은 입력 센

서의 수에 따라 단일입력 시스템과 이중 (혹은 다중) 입력 시스템으로 구분할 수 있다. 단일입력 음성개선 시스템은 단지 하나의 마이크를 통해서 입력신호를 얻게 되므로 배경잡음이 stationary 하다는 가정하에 목음 구간에서의 잡음의 통계적 특성을 분석하는 방법 등이 이용된다. 하지만 잡음이 stationary 하다고 볼 수 없을 뿐 아니라 목음 구간만으로 잡음의 특성을 분석하기에는 데이터 양이 불충분하며, 잡음이 내재된 음성신호에 목음 구간을 정확하게 분류해 내는데 많은 어려움이 있다.

Sambur[12]는 별도의 센서를 사용하지 않는 경우에도 음성신호의 적응잡음제거가 가능한 단일입력 적응잡음제거 기법을 제안하였다. 음성신호의 유성음부가 준주기적 특성을 가지므로 잡음이 내재된 입력신호를 한 두 피치(pitch)정도 지연시킴으로써 입력 음성신호와 높은 상관관계를 갖는 신호를 얻을 수 있는데, 그는 이 신호를 적응여파기의 기준신호로 사용하였다. 지연 값을 구하기 위해 입력신호를 음성신호의 통계적 특성이 변하지 않을 정도의 구간으로 나누고 평균절대차함수(AMDF: average magnitude difference function)[18]와 비선형 평활화과정[19]을 이용하여 분석구간별로 피치를 구하였으며 매 분석구간마다 추정된 피치만큼 지연된 입력신호를 얻어 최소평균자승(LMS: least mean square) 알고리즘을 사용하는 TDL(tapped delayed line) 구조의 적응여파기의 기준신호로 사용하여 내재된 잡음성분을 제거하고자 하였다. 그런데 기존 방법에서는 피치를 분석구간 단위로 구하며 분석구간 내에서 동일한 지연값을 사용하게 되므로 입력되는 때 음성신호 샘플에 대해 가장 높은 상관관계를 갖도록 지연된 기준신호를 얻을 수 없다. 또한, 잡음의 영향으로 인해 잘못 추정된 피치가 적용될 경우 그 영향이 분석구간 전체에 미치게 된다. 특히, 준주기적인 특성을 갖지 않는 무성음 구간에 대해서는 입력신호가 시스템을 그냥 통과하게 하거나 직전의 유성음 구간 처리때 얻어진 적응여파기의 계수를 update 시키지 않고 그대로 적용함으로써 천이 구간에서 청취피로(listening fatigue)를 초래하는 음질의 불연속성이 존재하게 되며 무성음 구간부터 시작되는 문장에는 처음부터 적응여파기의 계수를 update 시킬 수 없는 점 등 무성음 구간의 처리 방법이 바람직스럽지 못하다.

본 논문에서는, 분석구간 사이의 불연속성, 무성음 구간의 바람직하지 않은 처리 방법 등 기존 방법에서

의 문제점들을 해결하기 위하여 매 샘플시간마다 입력신호와 높은 상관관계를 갖는 지연 정보를 구할 수 있는 방법을 제안하고, 이를 이용하여 정규화된 최소평균자승(NLMS: normalized LMS) 알고리즘을 사용하는 적응여파기로 단일입력 적응잡음제거 시스템을 구성하여 음성개선 실험을 하였으며 기존 방법과의 성능 비교 실험을 수행하였다. 본 논문의 구성은 다음과 같다. II장에서 단일입력 적응잡음제거 기법의 원리에 대해 설명하고, 본 논문에서 제안된 지연 추정 방법 및 이를 이용한 단일입력 적응잡음제거 시스템에 대해 설명한다. III장에서는 가우시안 백색잡음이 내재된 음성신호를 제안된 시스템 및 기존 적응잡음제거 시스템에 적용하여 음성개선된 결과를 비교 검토하고, IV장에서 결론을 맺는다.

II. 단일입력 적응잡음제거

대부분의 음성신호처리 시스템에서는 입력신호가 1개의 센서(마이크)를 통해서 얻어지므로 별도의 2차 센서를 통해 배경잡음과 높은 상관관계를 갖는 기준신호를 얻지 않는 한 적응잡음제거 기법을 음성개선에 적용하기가 어렵다. Sambur는 대부분의 음성신호처리 시스템에서와 같이 적응여파기의 기준신호를 구하기 어려운 환경에서도 음성신호의 적응잡음제거가 가능한 단일입력 적응잡음제거 방법을 제안하였다. 그림 1은 음성신호의 유성음 구간의 준주기적 특성을 이용하여 기준신호를 얻는 단일입력 적응잡음제거 방법을 보인 것이다[12, 20]. 음성신호 $s(n)$ 에 백색잡음 $w(n)$ 이 섞인 입력신호 $x(n)$ 은 유성음 구간에서 준주기적 특성을 갖게 되므로 한 두 피치정도 지연된 신호 $x(n-T)$ 는 음성신호 성분 $s(n)$ 과는 높은 상관관계를 가지면서 잡음성분 $w(n)$ 과는 아주 낮은 상관관계를 갖게 된다. 이때, 적응여파기의 출력 $\hat{s}(n)$ 는 전체 시스템의 오차신호 $\hat{w}(n)$ 의 에너지를 최소화함으로써 입력신호 내의 음성신호 성분 $s(n)$ 의 최소자승오차를 갖는 대응신호(best least square fit)로 얻어지게 되는데 그 과정은 식(1)-(3)과 같다.

$$\begin{aligned} E[\hat{w}^2(n)] &= E[(s(n) + w(n) - \hat{s}(n))^2] \\ &= E[w^2(n) + (s(n) - \hat{s}(n))^2 + 2w(n)(s(n) - \hat{s}(n))] \end{aligned} \quad (1)$$

$$E[\hat{w}^2(n)] = E[w^2(n)] + E[(s(n) - \hat{s}(n))^2] \quad (2)$$

$$\min E[\hat{w}^2(n)] = E[w^2(n)] + \min E[(s(n) - \hat{s}(n))^2] \quad (3)$$

여기서 $E[\cdot]$ 는 기대값(expected value)을 나타낸다. 음성신호 성분 $s(n)$ 과 잡음성분 $w(n)$ 은 상관관계가 없으므로 식(1)은 식(2)로 표시될 수 있다. 식(2)에서, $E[w^2(n)]$ 은 입력신호 내의 잡음성분의 에너지로 고정된 값을 가지므로 추정된 잡음성분의 에너지 $E[\hat{w}^2(n)]$ 이 최소화될 때 $E[(s(n) - \hat{s}(n))^2]$ 도 최소화 되어 식(3)과 같이 표시된다. 따라서, 적응여파기의 출력 $\hat{s}(n)$ 은 입력신호 내의 음성신호 성분 $s(n)$ 에 최소자승오차를 갖는 추정값이 된다.

유한임펄스응답(FIR) 형태로 구성된 적응여파기의 출력 $\hat{s}(n)$ 은 식(4)와 같이 표현된다.

$$\hat{s}(n) = \sum_{i=0}^L b_i x(n-i-T) \quad (4)$$

여기서 $x(n)$ 은 잡음이 포함된 음성신호, T 는 기준신호를 만들기 위한 지연 정보, 그리고 b_i , $0 \leq i \leq L$, FIR 여파기의 계수들이다. Sambur는 지연 정보를 구하기 위한 피치 추정시 수정된 AMDF와 비선형 평활화 과정을 입력신호의 매 분석구간마다 적용하였다[18, 19]. 여파기의 계수, b_i 는 유성음 구간에서는 Widrow와 Hoff가 제안한 LMS 알고리즘[16]에 의하여 매 샘플시간마다 update하였으며, 무성음 구간의 경우에는 입력신호가 시스템을 그냥 통과하게 하여 잡음을 제거하지 않거나 직선의 유성음 구간 처리 때 얻어진 적응여파기의 계수를 update시키지 않고 무성음에 그대로 적용하였다.

LMS 알고리즘에서, 시간 n 일때의 여파기 계수 벡터 $[b_0 \ b_1 \ b_2 \ \dots \ b_L]$ 을 B_n 이라 하고 전체 시스템의 오차신호와 입력신호 벡터를 각각 식(5), (6)으로 표시하면,

$$\hat{w}(n) = x(n) - \hat{s}(n), \quad (5)$$

$$X_n = [x(n) \ x(n-1) \ \dots \ x(n-L)] \quad (6)$$

시간 $n+1$ 일 때의 여파기 계수는

$$B_{n+1} = B_n + 2\mu \hat{w}(n) X_n^T \quad (7)$$

로 update 된다. 여기서 μ 는 수렴속도와 안정성을 조

절하는 인자이며 $0 \leq \mu \leq 1/\lambda_{max}$ 의 값을 가져야 한다. 이때 λ_{max} 는 입력신호의 자기상관행렬의 최대 고유치 (eigenvalue)이다.

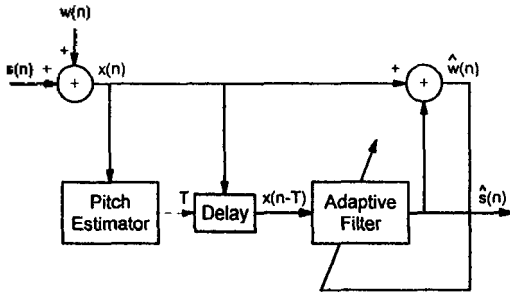


그림 1. 유성음 구간의 준주기적 특성을 이용하는 단일입력 적응잡음제거 시스템[12]

Fig 1. Single channel adaptive noise canceling system using the quasi-periodicity of the voiced sound [12]

Sambur의 방법에서는 기준신호를 얻기 위한 지연 정보인 피치의 추정치를 음성의 분석구간 단위로 구하면서 적응여파기의 계수는 음성샘플 단위로 update 하였다. 따라서 분석구간 내에서 동일한 지연값을 사용하게 되므로 입력되는 매 음성신호 샘플에 대해 가장 높은 상관관계를 갖도록 지연된 기준신호를 얻을 수 없다. 또한, 잡음의 영향으로 인해 잘못 추정된 피치가 적용될 경우 그 영향이 분석구간 전체에 미치게 된다. 본 연구에서는 이러한 문제를 해결하기 위해 매 샘플시간마다 지연 정보를 구할 수 있도록, 일반적인 자기상관함수와 평균절대차함수 등으로부터 재귀식을 유도하였다. 일반적인 ACF와 AMDF는 식 (8), (9)와 같이 정의 된다.

$$ACF_n(k) = \frac{1}{N+1} \sum_{i=n-N}^n x(i)x(i-k), 0 \leq k \leq N \quad (8)$$

$$AMDF_n(k) = \frac{1}{N+1} \sum_{i=n-N}^n |x(i)-x(i-k)|, 0 \leq k \leq N \quad (9)$$

$ACF_n(k)$ 와 $AMDF_n(k)$ 는 각각 입력신호 $x(n)$ 에 대해 k 번째 ACF와 AMDF의 계수이며 $N+1$ 은 분석구간의 길이이다. 단일입력 적응잡음제거 시스템의 기준신호를 얻기 위해 필요한 지연 정보인 피치 정보는 식

(10), (11)과 같이 구할 수 있다.

$$T = \arg\text{Max}_{t_L \leq k \leq t_U} [ACF_n(k)], \quad \text{for ACF} \quad (10)$$

$$T = \arg\text{Min}_{t_L \leq k \leq t_U} [AMDF_n(k)], \quad \text{for AMDF} \quad (11)$$

여기서 t_L, t_U 은 각각 주기의 상한, 하한치이다.

매 샘플시간마다 지연 정보를 구할 수 있도록 일반적인 자기상관함수나 평균절대차함수 등을 재귀적인 수식으로 유도하는 과정은 식(12), (13)과 같다. 식 (8)로부터 재귀적 자기상관함수는 식(12)와 같이 주어진다.

$$\begin{aligned} ACF_n(k) &= \frac{1}{N+1} \sum_{i=n-N}^{n-1} x(i)x(i-k) \\ &+ \frac{1}{N+1} [x(n)x(n-k) - x(n-N)x(n-N-k)] \\ &= ACF_{n-1}(k) + \frac{1}{N+1} x(n)x(n-k) \\ &- \frac{1}{N+1} x(n-N)x(n-N-k), 0 \leq k \leq N \quad (12) \end{aligned}$$

같은 방법으로, 식(9)로부터 재귀적 평균절대차함수는 식(13)과 같이 주어진다.

$$\begin{aligned} AMDF_n(k) &= \frac{1}{N+1} \sum_{i=n-N}^{n-1} |x(i)x(i-k)| \\ &+ \frac{1}{N+1} [|x(n)-x(n-k)| - |x(n-N)-x(n-N-k)|] \\ &= AMDF_{n-1}(k) + \frac{1}{N+1} |x(n)-x(n-k)| \\ &- \frac{1}{N+1} |x(n-N)-x(n-N-k)|, 0 \leq k \leq N \quad (13) \end{aligned}$$

식 (12), (13)에서 알 수 있듯이 입력신호 $x(n)$ 에 대한 재귀적 ACF와 AMDF는 입력신호 $x(n)$ 자신과 k 샘플만큼 이전의 신호 $x(n-k)$, 그리고 이들 신호와 각각 $N+1$ 샘플만큼 이전의 신호들 $x(n-N-1), x(n-N-1-k)$ 만으로 매 샘플시간마다 구할 수 있게 된다[21]. 매 샘플단위로 지연 정보를 추정할 때 적응여파기의 기준신호를 얻기위한 지연 정보는 식 (12), (13)에 대해 각각 식(10), (11)을 적용하여 같

은 방법으로 얻어지며 그 결과를 이용하여 유성음/무성음의 판별도 가능하게 된다.

음성개선을 위해 본 연구에서 제안된 지연추정 방법을 이용한 적응잡음제거 시스템의 블록도는 그림 2와 같다. 유성음 구간에서는, 매 샘플단위로 지연정보를 추정하고 추정된 값만큼 지연된 잡음음성신호를 적응여파기의 기준신호로 사용하여 매 샘플마다 잡음이 제거된다. 무성음 구간에서는, 한 샘플 지연된 입력신호를 적응여파기의 기준신호로 사용하여 매 샘플마다 잡음 성분이 적응적으로 제거된다. 제안된 방법에서는, 이와 같이, 매 샘플마다 기준신호를 위한 지연 정보를 구하고 매 샘플마다 적응여파기로 처리함으로써 매 샘플마다 입력신호 내의 음성신호와 가장 상관관계가 높은 기준신호를 얻을 수 있고 분석구간 사이의 불연속성이 제거된다. 무성음 구간에서도 유성음 구간에서와 같이 입력신호를 매 샘플마다 처리함으로써 유/무성음 천이 구간에서의 불연속성이 제거되고 유성음과 무성음의 처리 과정이 서로 일관성을 가지므로 무성음 구간의 처리가 용이해진다. 광대역 잡음(broadband noise)이 입력신호에 더해진 경우, 일반적으로 저주파 대역의 잡음신호 에너지가 음성의 에너지 보다 적으므로 고주파 대역에 대한 잡음의 영향에 비해 저주파 대역에 대한 잡음의 영향이 적다고 볼 수 있다. 따라서, 잡음의 영향이 비교적 적은 저주파 대역의 입력신호만으로 지연 정보를 추정하기 위하여 차단주파수가 1kHz인 저역통과여파기를 지연추정 과정 직전에 사용하였다.

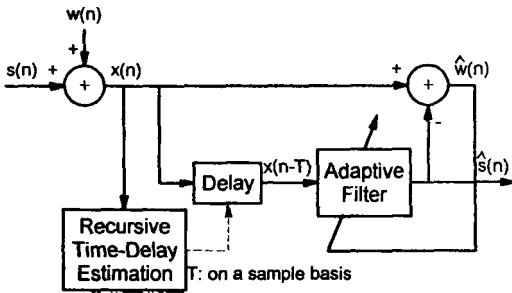


그림 2. 제안된 시스템의 구조
Fig 2. Structure of the proposed system

III. 실험 및 결과

본 장에서는 백색잡음이 내재된 음성신호를 매 샘플

마다 지연 정보를 추정하는 제안된 시스템과 기존 적응잡음제거 시스템[12]에 적용하여 음성개선된 결과를 비교 검토하였다. 컴퓨터 시뮬레이션으로 만들어진 백색 가우시안 잡음(AWGN: additive white Gaussian noise)이 신호대 잡음비(SNR: signal-to-noise ratio)의 범위가 10dB에서 -10dB가 되도록 한국어 문장으로 구성된 음성신호에 첨가하여 잡음 음성을 얻었다. 적응여파기로는 정규화된 최소평균자승(NLMS) 알고리즘을 사용하는 횡단형 여파기를 사용하였다. 음성개선 실험결과를 비교 검토하기 위해 사용된 입력 SNR과 출력 SNR의 계산은 주어진 음성신호 전체에 대해 식(25), (26)을 이용하여 계산되는데, 무성음 구간 및 묵음 구간도 유성음 구간과 함께 출력 SNR 계산에 포함되었다.

$$\text{입력 SNR} = 10 \log_{10} \frac{\sum_n s^2(n)}{\sum_n [s^2(n) - x(n)]^2} \quad (25)$$

$$\text{출력 SNR} = 10 \log_{10} \frac{\sum_n s^2(n)}{\sum_n [s(n) - \hat{s}(n)]^2} \quad (26)$$

실험에 사용된 적응잡음제거 시스템은 사용된 여파기의 종류와 지연 정보추정 방법에 따라 NLMS-ACF, NLMS-AMDF로 나누어진다.

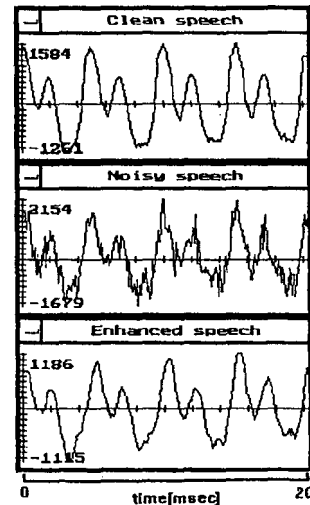


그림 3. NLMS-AMDF 시스템에 의해 처리된 유성음 구간의 음성신호(입력 SNR = 0dB)
Fig 3. A voiced sound processed by NLMS-AMDF system(Input SNR = 0dB)

그림 3은 입력 SNR이 0dB 일때의 NLMS-AMDF 시스템에 의해 처리된 유성음 구간의 음성신호를 보인 것이며, 그림 4는 그림 3에 주어진 유성음 구간에 대응하는 LPC 스펙트럼을 나타낸 것인데, 제안된 시스템에 의해 처리된 음성의 스펙트럼(short-dashed line)이 처리전의 잡음음성의 스펙트럼(long-dashed line)보다 원래 음성신호의 스펙트럼(solid line)에 훨씬 가까와짐을 보여 준다.

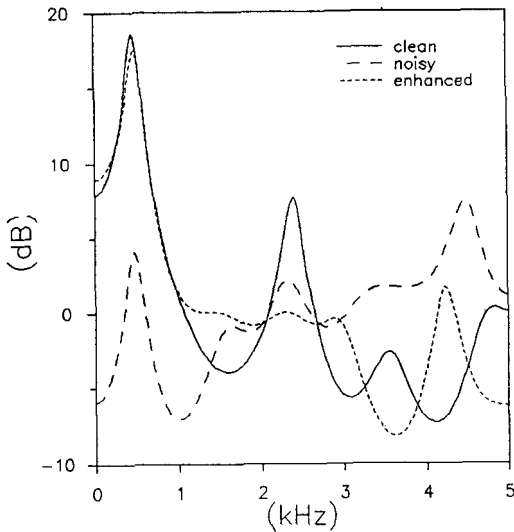


그림 4. NLMS-AMDF 시스템에 의해 처리된 유성음 구간의 스펙트럼(입력 SNR = 0dB)

Fig. 4. LPC spectrum of a voiced sound processed by the NLMS-AMDF system(Input SNR = 0dB)

표 1과 표 2는 여러개의 한국어 문장에 대해 제안된 방법에 의해 처리된 음성신호의 출력 SNR과, cepstrum 거리를 각각 입력 SNR에 대하여 나타낸 것이며, 그림 5는 표 1의 출력 SNR을 기존의 Sambur 방법에 의해 처리된 결과와 비교하여 그래프로 그린 것이다. 매 샘플단위로 지연 정보를 추정할 때 한번에 고려하는 음성신호의 범위는 10kHz의 sampling rate에서 201 샘플로 하였으며 적응여파기의 차수는 14차로 하였다. NLMS의 알고리즘의 수렴인자 μ 는 0.01을 사용하였다. 표 1과 표 2에서, 입력 SNR이 10dB인 경우에는 적응처리 전과 후의 음질(SNR, cepstrum 거리)은 거의 개선되지 않음을 볼 수 있다. 그러나, 잡음의 정도가 심하면 심할수록 음질은 더욱 많

이 개선됨을 볼 수 있는데, 입력 SNR이 0dB 일 때 NLMS-ACF 시스템으로 처리된 경우에는 6.8dB 정도의 출력 SNR이 개선됨을 보였다.

표 1. 처리된 음성의 출력 SNR(dB)

Table 1. SNR of the processed speech signal (in dB)

Input SNR[dB]	NLMS	
	RACF	RAMDF
-10.0	1.732353	0.919702
-5.0	4.287469	3.783259
0.0	6.794404	6.676066
5.0	8.492607	8.447666
10.0	9.673522	9.626303

표 2. 처리된 음성의 cepstrum 거리

Table 2. Cepstral distance of the processed speech signal

Input SNR[dB]	Cepstral Distance		
	Noisy Speech	NLMS	
		ACF	AMDF
-10	3.875357	3.253632	2.955949
-5	3.573388	2.741590	2.702813
0	3.200887	2.619581	2.519471
5	2.794361	2.631146	2.610891
10	2.346024	2.531770	2.482355

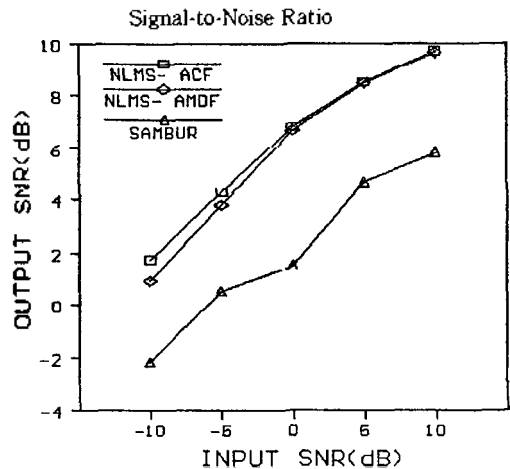


그림 5. 처리된 음성의 출력 SNR

Fig 5. SNR of the processed speech signal

그림 6은 한국어 문장 “북풍은 힘껏 불었습니다”에 대한 원음성 및 입력 SNR이 0dB인 잡음 음성신호에 대해 기존 방법으로 처리한 결과를 보인 것이다. 그림 6에서 화살표로 표시된 부분과 같이 기존 방법에서는 분석구간 단위로 피치 정보를 구할 뿐만아니라 무성음 구간의 경우 잡음을 제거하지 않고 시스템을

그냥 통과시키므로 유/무성음 천이 구간 및 무성음 구간에서 청취 피로를 유발하는 음성신호의 불연속성을 볼 수 있다. 그림 7은 입력 SNR이 5dB인 한국어 문장 “북풍은 힘껏 불었습니다”를 NLMS-AMDF 시스템으로 잡음제거한 결과를 나타낸 것인데, 그림 6에서와 같은 유/무성음 천이 구간에서의 불연속성

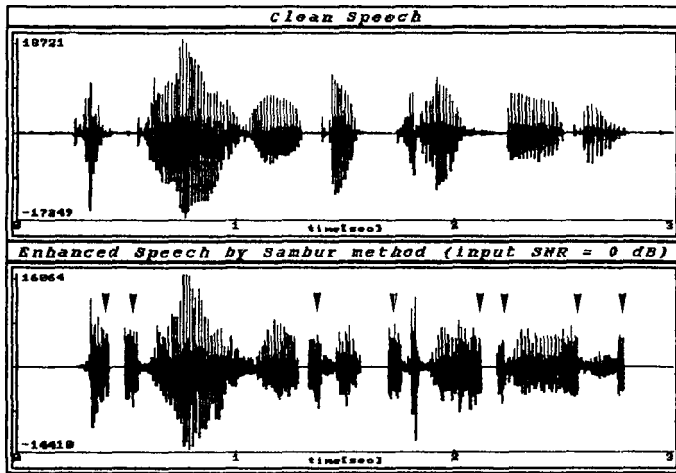


그림 6. 기존 방법으로 처리시 발생하는 천이 구간에서의 음의 불연속성
 Fig. 6. Discontinuity of speech sound in the transition region processed by the conventional method

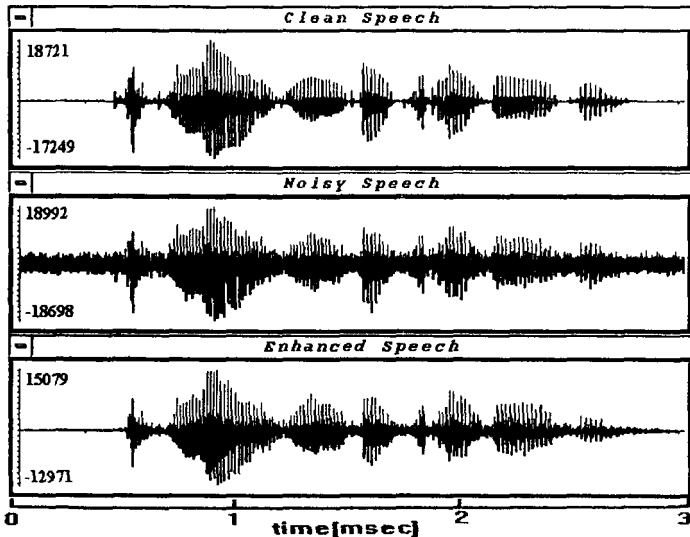


그림 7. NLMS-AMDF 시스템에 의해 처리된 한국어 문장의 음성신호(입력 SNR = 5dB)
 Fig. 7. Processed speech signal of a Korean sentence processed by NLMS-AMDF system(Input SNR = 5dB)

이 제거 되었음을 볼 수 있다. 잡음 성분이 많이 제거 되었음을 볼 수 있으며, Informal한 청취시험에서도 기존 Sambur 방법에 의해 처리된 음성신호에 비해 출력 음성신호의 잡음 제거 성능이 향상되었음을 보여 주었다.

VI. 결 론

본 논문에서는 기존 음성개선 방법에서의 문제점인 분석구간 사이의 불연속성과 유/무성음 천이 구간이나 무성음 구간에서 나타나는 성능 저하 등을 해결하기 위해서 매 샘플시간마다 지연 정보를 추정할 수 있는 단일입력 적응잡음제거 시스템을 제안하고, 이를 이용하여 정규화된 최소평균자승(NLMS : normalized LMS) 알고리즘을 사용하는 적응여파기를 구성하여 음성개선 실험을 하였으며 기존 방법과의 성능 비교 실험을 수행하였다. 제안된 방법에서는, 매 샘플시간마다 적절한 지연 정보를 적응잡음제거 시스템에 전달하므로, 지연정보를 구하는데 있어, 음성 분석구간 및 유/무성 천이 구간에서의 불연속성 등 기존 방법의 문제점을 제거하였으며, 무성음 구간의 처리도 용이하게 되었다. 제안된 시스템을 적용한 음성개선 실험 결과, 출력 SNR이 향상되고 왜스트럴 거리도 단축되는 등 기존 방법에 의한 결과보다 많이 개선됨을 보였다. Informal한 청취시험에서도 기존 Sambur 방법에 의해 처리된 음성신호에 비해 출력 음성신호의 잡음제거 성능이 향상되었음을 보여 주었다. 앞으로, 실제 음성(컴퓨터로 시뮬레이션된 잡음이 아닌 실제 바깥 환경의 잡음이 내재된 음성)에 대한 실험을 통하여, 제안된 시스템에 대한 연구가 계속 수행되어야 할 것이다.

참 고 문 헌

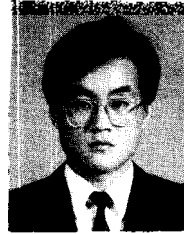
1. J.S. Lim, *Speech Enhancement*, Prentice-Hall, 1983.
2. J.S. Lim and A.V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proc. IEEE*, vol.67, no.12, Dec.1979.
3. D. O'Shaughnessy, "Enhancing speech degraded by additive noise or interfering speakers," *IEEE Communication Magazine*, pp.46-52, Feb. 1989.
4. S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp.113-120,

- Apr. 1979.
5. A. Schaub and P. Schaub, "Spectral sharpening for speech enhancement/noise reduction," *Proc. of Int. Conf. on Acoust., Speech, signal Processing*, vol.2, pp. 993-996, May 1991.
6. 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.
7. J. Hansen and M. Clements, "Constrained iterative speech enhancement with to speech recognition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-39, no.4, pp.21-27, Apr. 1989.
8. Y. Ephraim, "Statistical-model-based speech enhancement systems," *Proc. of IEEE*, vol.80, no.10, pp. 1526-1555, Oct. 1992.
9. Y. Ephraim, D. Malah, and B.H. Juang, "On the application of hidden Markov models for enhancing noisy speech," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, no.12, pp.1846-1989, Dec. 1989.
10. M. Feder, A.V. Oppenheim, and E. Weinstein, "Maximum likelihood noise cancellation using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, no.2, pp.204-216, Feb. 1989.
11. J.S. Lim, A.V. Oppenheim, and L.D. Braida, "Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, no.4, pp.354-358, Apr. 1991.
12. M.R. Sambur, "Adaptive noise canceling for speech signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp.419-423, Oct. 1978.
13. 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, no.6, pp. 752-753, Dec. 1989.
14. H.L. Nquyen Thi, C. Jutten, and J. Caelen, "Speech enhancement : analysis and comparison of methods on various real situations," *Signal Processing IV, Theorems and Applications*, vol.1, pp. 303-306, 1992.
15. E. Masgrau, J.A. Rodriguez-Fonollosa, and A. Ardanuy, "Enhancement of speech by using higher-order spectral modelling," *Signal Processing*

IV. Theorems and Applications, vol.1, pp.307-310, 1992.

16. B.Widrow et al., "Adaptive noise canceling: principles and applications," *Proc. of IEEE*, vol.63, pp. 1692-1716, Dec. 1975.
17. L.R. Rabiner, "On the use of autocorrelation analysis for pitch detection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, no.1, pp. 24-33, Feb. 1977.
18. M.J. Ross, H.L. Shaffer, A. Cohen, R. Freudberg, and H.J. Manley, "Average magnitude difference function pitch extractor," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp.353-362, Oct. 1974.
19. L. Rabiner, M.Sambur, and C. Schmidt, "Application of a nonlinear smoothing algorithm to speech processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp.552-557, Dec. 1975.
20. J.R. Deller, Jr., J.G. Prokis, and J.H.L. Hansen, L.J. Griffiths, *Discrete-Time Processing of Speech Signals*, Macmillan Publishing Company, 1993.
21. 강해동, 배건성, "음성신호의 적응잡음제거에 관한 연구," 제10회 음성통신 및 신호처리 워크샵 논문집, 제 SCAS-10권 1호, pp.145-148, 8월 1993년.

▲姜海東(Hae Dong Gahng) 1961년 10월 9일생



1980년 3월~1987년 2월: 경북
대학교 전자공학과
(공학사)

1987년 3월~1989년 2월: 경북
대학교 전자공학과
(공학석사)

1989년 3월~1994년 2월: 경북
대학교 전자공학과
(공학박사)

※주관심분야: 음성신호처리, 적응신호처리 등

▲裨建星(Keun Sung Bae) 1953년 11월 9일생



1973년 3월~1977년 2월: 서울
대학교 전자공학과
(BS)

1977년 3월~1979년 2월: 한국
과학원 전기 및 전
자공학과(MS)

1984년 8월~1989년 5월: 미국
Univ. of Florida
(Ph.D)

1979년 3월~현재: 경북대학교 전자공학과 부교수

※주관심분야: 디지털 신호처리, 음성신호처리, 디
지탈통신 등