

論文94-31B-7-5

재귀적 지연추정기를 갖는 적응잡음제거 기법을 이용한 음성개선 (Speech Enhancement Using the Adaptive Noise Canceling Technique with a Recursive Time Delay Estimator)

姜海東*, 裴建星**

(Hae Dong Gahng and Keun Sung Bae)

要約

배경 잡음이 있는 음성신호에서 잡음 성분을 제거하기 위하여 재귀적 지연추정기를 갖는 단일입력 적응잡음제거 시스템을 제안하였다. 적응여파기의 기준 신호를 구하기 위하여, 기존의 방법에서는 입력 음성신호에서 프레임 단위로 분석된 피치 정보를 이용하지만 제안된 방법에서는 재귀적인 방법으로 매 샘플마다 추정된 지연 정보를 이용한다. 재귀적 지연추정기는 일반적인 자기상관함수와 평균절대차함수 등으로부터 이들의 재귀식을 유도하였다. 여파기 계수의 적응 조절을 위하여 정규화된 최소평균자승(NLMS) 및 재귀적 최소자승(RLS) 알고리즘 등이 사용되었다. 잡음에 의해 왜곡된 음성에 대한 음성개선 실험 결과, 제안된 방법으로 처리된 음성신호의 출력 SNR이 향상되고 켈스트럼 거리가 단축되며 음질이 기존 방법에 의한 결과보다 개선됨을 보였다.

Abstract

A single channel adaptive noise canceling (ANC) technique with a recursive time delay estimator (RTDE) is presented for removing effects of additive noise on the speech signal. While the conventional method makes a reference signal for the adaptive filter using the pitch estimated on a frame basis from the input speech, the proposed method makes the reference signal using the delay estimated recursively on a sample-by-sample basis. As the RTDEs, the recursion formulae of autocorrelation function (ACF) and average magnitude difference function (AMDF) are derived. The normalized least mean square (NLMS) and recursive least square (RLS) algorithms are applied for adaptation of filter coefficients. Experimental results with noisy speech demonstrate that the proposed method improves the perceived speech quality as well as the signal-to-noise ratio and cepstral distance when compared with the conventional method.

1. 서론

*正會員, 昌信專門大學 電子通信科
(Dept. of Elec. communication, Changshin
Junior College)

** 正會員, 慶北大學校 電子工學科
(Dept. of Elec. Kyungpook Nat'l Univ.)
接受日字 : 1993年 11月 11日

종래의 전화통신 뿐만아니라 사람과 사람 사이의 정보교환에 컴퓨터등이 개입되는 통신 형태가 상당한 비중을 차지하게 된 현대의 통신에서는 사람과 컴퓨터 사이의 정보 전달을 원활하게 하는 man-machine interface의 중요성이 대두되었으며,

man-machine interface 기술로서 컴퓨터를 이용한 음성신호의 분석 및 합성, 음성인식 등에 대한 연구가 꾸준히 진행되어 왔다.^[14] 대부분의 음성신호처리 시스템들이 주변 잡음이 없는 양질의 음성을 대상으로 연구가 진행되어 왔다. 이러한 여건의 음성신호처리 시스템의 성능은 잡음을 무시할 수 없는 실제 환경에서 급격히 나빠진다. 따라서 음성신호처리 시스템의 실용화를 위해서 음성개선, 즉, 수신된 음성신호에서 잡음을 경감시키는 기술에 대한 연구 및 개발이 지속적으로 이루어져 왔다.^[15,16]

LMS 알고리즘이 1960년대에 제안된 이래, 적응신호처리 기법은 컴퓨터의 신호처리 능력 및 VLSI 회로 설계 기술의 향상에 힘입어, 디지털 신호처리의 여러 분야에서 성공적으로 응용되고 있는데, 그 대표적인 것이 적응잡음제거기이다.^[15,17] 적응잡음제거 기법은, 사용되는 여파기의 계수를 자동으로 조절하는 능력을 가지고 있으며 원 신호에 포함된 잡음신호의 상세한 통계적 특성을 미리 알고 있지 않아도 된다는 장점이 있으나, 일반적으로 두가지 입력 신호를 필요로 한다. 즉, 입력 신호 내의 잡음과 높은 상관관계를 갖는 잡음을 따로 얻을 수 있어야 하는데 이 잡음신호를 기준 입력(reference input)이라고 부른다. 입력 신호와 높은 상관관계를 갖는 기준 신호를 얻기 위해서는 일반적으로 2차의 센서를 필요로 한다.

적응잡음제거 시스템은 입력 센서의 수에 따라 단일입력 시스템과 다중입력 시스템으로 구분할 수 있다. 단일입력 적응잡음제거 시스템은 단지 하나의 센서를 통해서 입력 신호를 얻게 되므로 배경잡음이 stationary 하다는 가정하에 신호가 없는 구간에서 잡음의 통계적 특성을 분석하는 등의 방법을 이용한다. 대부분의 음성신호처리 시스템에서는 입력 신호가 한 개의 센서(마이크)를 통해서 얻어지므로 별도의 2차 센서를 통해 배경잡음과 높은 상관관계를 갖는 기준 신호를 얻지 않으면 적응잡음제거 기술을 음성개선에 적용하기가 어렵다. Sambur는 대부분의 음성신호처리 시스템에서와 같이 적응여파기의 기준 신호를 구하기 어려운 환경에서도 음성신호의 준주기적 특성을 이용하여 적응잡음제거가 가능한 단일입력 적응잡음제거 기법을 제안하였다.^[18] 그런데, 그의 방법에서는 기준 신호를 얻기 위한 지연 정보인 피치의 추정치를 입력 신호의 분석 구간 단위로 구하면서 적응여파기의 계수는 매 샘플마다 구하게 되므로 분석 구간 사이에 불연속성을 초래할 수 있으며, 또한 준주기적 특성을 갖지 않는 무성음부의 처리 방법등도 바람직스럽지 못했다.

본 논문에서는 기존 방법에서의 문제점인 분석 구

간 사이의 불연속성과 유/무성음 천이 구간에서 나타나는 성능저하 등을 해결하기 위해서 매 샘플시간마다 기준신호를 구할 수 있는 재귀적 지연추정기를 제안한다. 지연 정보의 추정 계산량을 줄이기 위하여 재귀식을 사용하고, 제안된 재귀적 지연추정기와 정규화된 최소평균자승(NLMS: normalized LMS) 및 재귀적 최소자승(RLS) 적응알고리즘 등을 사용하는 적응여파기로 단일입력 적응잡음제거 시스템을 구성하여 음성개선 실험을 하며, 기존 방법의 성능과 비교 검토한다. 본 논문의 구성은 다음과 같다. II장에서 단일입력 적응잡음제거 기법의 원리와 이 논문에서 제안된 재귀적 지연추정기에 대해 설명한다. 가우시안 백색잡음이 내재된 음성신호를 제안된 시스템 및 기존 적응잡음제거 시스템에 적용하여 개선된 음성신호의 결과를 III장에서 비교 검토하고, IV장에서 결론을 맺는다.

II. 재귀적 지연추정기

그림 1은 음성신호의 준주기적 특성을 이용하여 기준 신호를 얻는 단일입력 적응잡음제거 방법을 보인 것이다.^[7,8]

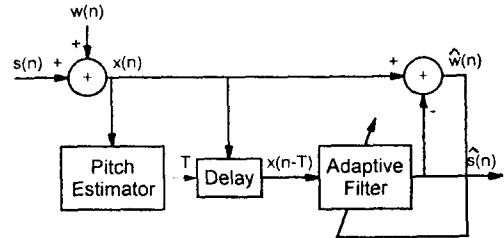


그림 1. 유성음 구간의 준주기적 특성을 이용하는 단일입력 적응잡음제거 시스템^[7,8]

Fig. 1. Single channel adaptive noise canceling system using the quasi-periodicity of the voiced sound.^[16,24]

음성신호 $s(n)$ 에 백색잡음 $w(n)$ 이 섞인 입력 신호 $x(n)$ 은 유성음부에서 준주기적 특성을 갖게 되므로 한 두 피치정도 지연된 신호 $x(n-T)$ 는 음성신호 성분 $s(n)$ 과는 높은 상관 관계를 가지면서 잡음 성분 $w(n)$ 과는 아주 낮은 상관 관계를 갖게 된다. 이때, 적응여파기의 출력 $\hat{s}(n)$ 은 예측오차신호 $\hat{w}(n)$ 의 에너지를 최소화함으로써 입력 신호 내의 음성신호 성분 $s(n)$ 과 최소자승오차를 갖는 대응신호(best least square fit)로 얻어지게 된다.^[8]

Sambur의 방법에서는 기준 신호를 얻기 위한 지

연 정보인 피치의 추정치를 음성의 분석 구간 단위로 구하면서 적응여파기의 계수는 음성샘플 단위로 update 하였다. 이로 인해 구해진 지연 정보에 분석 구간 사이의 불연속성을 초래하게 되는데, 본 연구에서는 이러한 문제를 해결하기 위해 매 샘플시간마다 지연 정보를 구할 수 있도록, 자기상관함수(ACF)와 평균절대차함수(AMDF) 등으로부터 이들의 재귀식을 유도하였다. ACF 와 AMDF 는 식(1), (2)와 같이 정의된다.

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

$$AMDF_n(k) = \frac{1}{N+1} \sum_{i=n-N}^n |x(i) - x(i-k)|, 0 \leq k \leq N \quad (2)$$

$ACF_n(k)$ 와 $AMDF_n(k)$ 는 각각 입력 신호 $x(n)$ 에 대해 k 번째 ACF와 AMDF의 계수이며, $N+1$ 은 분석 구간의 길이이다. 식(1), (2)의 값에서 최적의 주기 추정치 T 는 식(3), (4)와 같이 주어진다. 이때, t_U , t_L 은 각각 주기의 탐색을 위한 상한, 하한치이다.

$$T = \arg \text{Max}[ACF_n(k)] \quad (3)$$

$$t_L \leq k \leq t_U$$

$$T = \arg \text{Min}[AMDF_n(k)] \quad (4)$$

$$t_L \leq k \leq t_U$$

식 (1)을 재귀적인 수식의 형태가 되도록 수정하면 식(5)와 같이 쓸 수 있다.^[9]

$$RACF_n(k) = RACF_{n-1}(k) + \frac{1}{N+1} x(n)x(n-k) - \frac{1}{N+1} |x(n-N-1) - x(n-N-1-k)|, \quad (5)$$

같은 방법으로, 식 (2)를 재귀적인 수식의 형태가 되도록 수정하면 식(6)과 같이 쓸 수 있다.^[9]

$$RAMDF_n(k) = RAMDF_{n-1}(k) + \frac{1}{N+1} |x(n) - x(n-k)| - \frac{1}{N+1} |x(n-N-1) - x(n-N-1-k)|, \quad (6)$$

식 (5), (6)에서 알 수 있듯이 입력 신호 $x(n)$ 에 대한 RACF와 RAMDF는 입력신호 $x(n)$ 자신과 이 입력 신호와 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 재귀적으로 구할 수 있게 된다.^[9]

그림 2는 본 논문에서 제안된 재귀적 지연추정기의 블록도이다. 광대역 잡음이 음성신호에 더해진 경우, 대체로 음성신호의 고주파 대역에 대한 잡음의 영향에 비해 저주파 대역에 대한 잡음의 영향이 적다고 볼 수 있다. 따라서 잡음의 영향이 비교적 적으면서 지연 정보도 손상되지 않고 그대로 유지될 수 있는 저주파 대역의 입력 신호만으로 지연 정보를 추정하기 위하여 차단 주파수가 1 kHz인 저역 통과 여파기를 사용하였다.^[10]

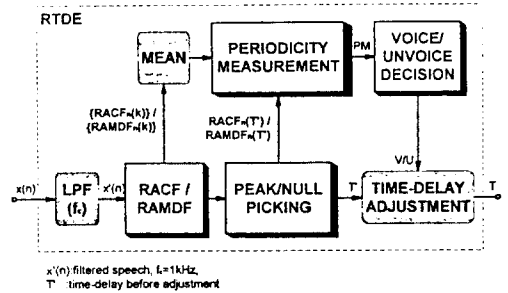


그림 2. 재귀적 지연추정기(RTDE)의 블록도
Fig. 2. Blockdiagram of the RTDE.

그림 3과 그림 4는 각각 그림 2의 RACF/RAMDF 로 표기된 블록의 세부 구조를 나타낸다.

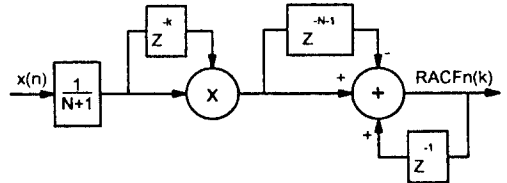


그림 3. 재귀적 자기상관함수의 구조
Fig. 3. Structure of the recursive ACF.

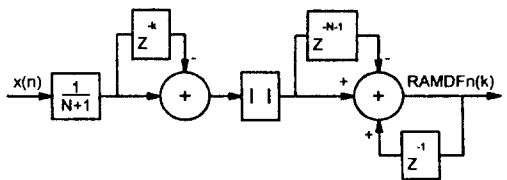


그림 4. 재귀적 평균절대차함수의 구조
Fig. 4. Structure of the recursive AMDF.

재귀적 지연추정기는 적응여파기의 기준 신호를 얻기 위한 지연 정보를 식(5)와 식(6)에 대해 각각 식

(3), (4)를 적용하여 재귀적 자기상관함수의 최대값 T' 과 재귀적 평균절대차함수의 최소값 T' 을 얻을 수 있다. 재귀적 자기상관함수의 최대값 혹은 재귀적 평균절대차함수의 최소값은 최종 지연 정보가 될 수 있는 후보값이 될 뿐만아니라 모든 k 에 대한 식(5)와 식(6)의 평균값과 함께 입력 신호 $x(n)$ 의 유성음/무성음의 판별에 이용되는 주기 측도를 구하는 데에도 이용된다.

식(7), (8)은, RACF와 RAMDF 각각에 대해, 입력 신호 $x(n)$ 이 유성음인가 무성음인가를 결정하는데 이용되는 주기 측도인 PM(periodicity measurement)을 구하는 식이다. 여기서 $E\{\}$ 는 기대값을 나타내며, C 는 잡음의 심한 정도에 따라 PM을 조절하는 상수이다.

$$PM = \frac{C \cdot RACF_n(T')}{E \{RACF_n(k)\}} \quad (7)$$

$$PM = C[E\{RAMDF_n(k)\} - RAMDF_n(T')] \quad (8)$$

식(7)의 오른쪽 식에서 분자의 값은 RACF 계수 중 가장 큰 값을 갖는 계수이므로, 분자가 분모보다 월등히 커진다는 것은 입력되는 n 번째 샘플이 유성음을 나타낸다. 식(8)의 경우, 두번째 항이 최소값을 가지는 RAMDF 계수이므로 첫번째보다 월등히 작아진다는 것은 입력되는 n 번째 샘플이 유성음을 나타낸다. 주기 측도 PM으로부터 입력 신호 $x(n)$ 이 유성음인가 무성음인가를 판단한 다음 유성음일 경우 지연 정보의 후보값으로 구한 T' 을 최종 지연 정보로 결정하고, 무성음인 경우 지연 정보를 1로 조정하여 결정하게 된다.

표 1은 피치를 구하는 일반식과 지연 정보를 구하는 재귀식의 특징들을 비교해 놓은 것이다. $N+1$ 이 분석 구간의 길이일 때, ACF의 경우, $N(N+1)/2$ 의 덧셈과 $(N+1)(N+2)/2$ 의 곱셈 과정이 필요하며, AMDF의 경우, $(N+1)$ 의 덧셈 과정이 필요하다. 더 우기, ACF와 AMDF의 경우, windowing과 pre-emphasis와 같은 전처리 과정이 부가적으로 필요하므로 실제로는 더 많은 덧셈과 곱셈 과정이 필요하게 된다. 반면 재귀적 지연추정기를 이용하여 $N+1$ 의 길이 갖는 분석 구간에서 지연 정보를 1회 구하고자 할 때 RACF의 경우, $2(N+1)$ 의 덧셈과 $2(N+1)$ 의 곱셈과정이 필요하고, RAMDF의 경우, $4(N+1)$ 의 덧셈 과정만이 필요하다. 그러므로, $N+1$ 개의 샘플에 대해 매 샘플마다 지연 정보를 추정하고자 할 경우 계산량은 $N+1$ 배 만큼 증가하게 된다.

표 1. 일반적 방법과 재귀식 방법의 비교
Table 1. Comparisson of the conventional and recursive methods.

	일반		재귀	
	ACF	AMDF	RACF	RAMDF
분석단위	프레임		샘플	
전처리과정	preemphasis windowing		없음	
덧셈	$N(N+1)/2$	$(N+1)^2$	$2(N+1)$	$4(N+1)$
곱셈	$(N+1)(N+2)/2$		$2(N+1)$	
내삽음성 (실험결과 참조)	약함		강함	
프레임간 불연속성	있음		없음	

지연 정보를 1회 구하는 데 필요한 계산량, ACF와 AMDF의 덧셈과 곱셈 계산과정의 경우, 이들의 일반식에서 overlap되는 부분만 고려하였으며, 전처리과정에 소요되는 계산과정은 포함되지 않았다. $N+1$ 은 분석 구간 길이.

음성개선을 위해 본 연구에서 제안된 재귀적 지연추정기를 이용한 적응잡음제거 시스템의 블록도는 그림 5와 같다.

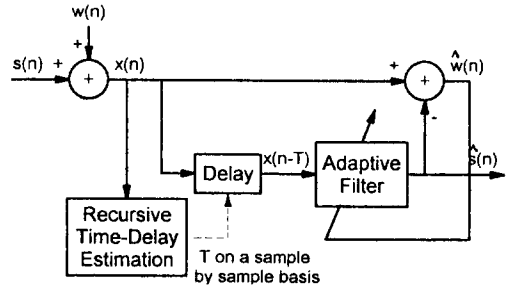


그림 5. 제안된 시스템의 구조

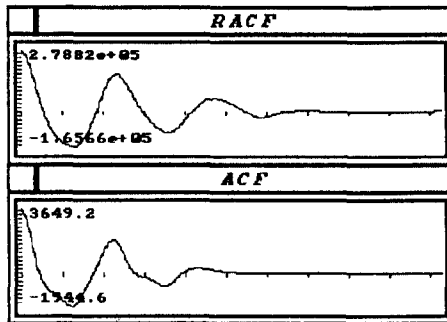
Fig. 5. Structure of the proposed system.

유성음부에서는, 재귀적 지연추정기를 이용하여 매 샘플마다 지연값을 추정하고 추정된 값만큼 지연된 잡음음성신호를 적응여파기의 기준 신호로 사용하여 매 샘플마다 잡음이 제거된다. 무성음부에서는, 한 샘플 지연된 입력 신호를 적응여파기의 기준 신호로 사용하여 매 샘플마다 잡음 성분이 적응적으로 제거된다. 제안된 방법에서는, 이와 같이, 매 샘플마다 기준 신호를 위한 지연 정보를 구하고 매 샘플마다 적응여파기로 처리함으로써 지연 정보를 구하는데 있어 분석 구간사이의 불연속성이 제거되며, 무성음부에서도 유성음부에서와 같이 입력 신호를 매 샘플마다 처리함으로써 유/무성음 천이 부분에서의 불연속성이 제거되고 유성음과 무성음의 처리 과정이 서로 일관성을 가지므로 무성음부의 처리가 용이해진다.

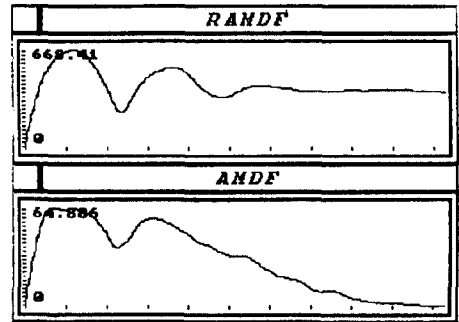
Ⅲ. 실험 및 결과

백색잡음이 내재된 음성신호를 재귀적 지연추정기를 사용하는 제안된 시스템과 기존 적응잡음제거 시스템⁽⁸⁾에 적용하여 개선된 음성신호의 결과를 비교 검토하였다. 컴퓨터 시뮬레이션으로 만들어진 백색가우시안 잡음을 한국어 문장으로 구성된 음성신호에

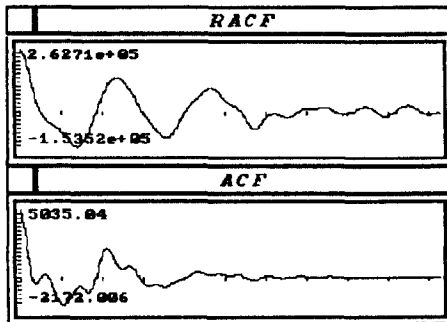
첨가하여 신호대 잡음비(SNR: signal-to-noise ratio)의 범위가 10 dB 에서 -10 dB가 되는 잡음음성을 얻었다. 적응여파기로는 정규화된 최소평균자승(NLMS) 알고리즘을 사용하는 횡단형 여파기와 재귀적 최소자승(RLS) 알고리즘을 사용하는 격자형 여파기를 사용하였다. 입력 SNR과 출력 SNR의 계산은 식(9), (10)과 같이 정의된 식들을 이용하였으며,



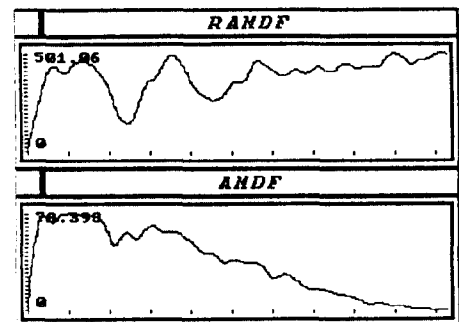
(a)clean speech



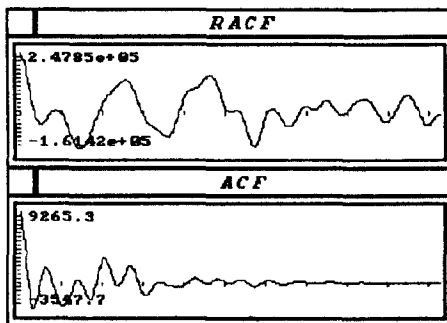
(a)clean speech



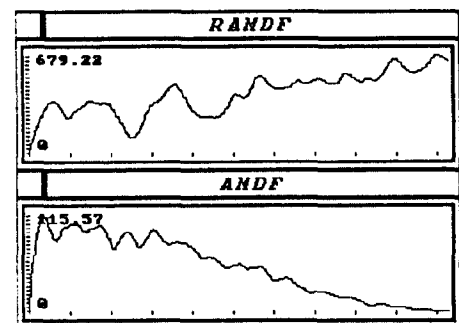
(b)noisy speech (input SNR = -5dB)



(b)noisy speech (input SNR = -5dB)



(c)noisy speech (input SNR = -10dB)



(c)noisy speech (input SNR = -10dB)

그림 6. RACF와 ACF의 비교
Fig. 6. Comparison of RACF and ACF.

그림 7. RAMDF와 AMDF의 비교
Fig. 7. Comparison of RAMDF and AMDF.

무성음부와 목음부도 유성음부와 함께 출력 SNR 계산에 포함되었다.

$$\text{SNR} = 10 \log_{10} \frac{\sum_n s^2(n)}{\sum_n [s(n) - x(n)]^2} \quad (9)$$

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

여기서, $s(n)$ 은 잡음이 섞이지 않은 원래의 음성신호, $x(n)$ 은 잡음이 내재된 음성신호, 그리고 $\hat{s}(n)$ 은 적응잡음제거 시스템에 의해 처리된 음성신호이다. 음성개선 결과를 비교하기 위해 사용된 스펙트럼 거리는 식(11)과 같이 정의된 켈스트럼 거리를 이용하였다.

$$\text{cepstrum distance} = \sqrt{\sum_{i=1}^L (c_i - \hat{c}_i)^2} \quad (11)$$

여기서, c_i 는 켈스트럼 거리 비교에 기준이 되는 잡음이 섞이지 않은 음성신호 $s(n)$ 에서 구해진 LPC 켈스트럼 계수이고, \hat{c}_i 는 비교 대상이 되는 잡음음성 $x(n)$ 이나 처리된 음성신호 $\hat{s}(n)$ 에서 구해진 LPC 켈스트럼 계수이다.

그림 6과 7은 각각 제안된 재귀적 지연추정기(RACF, RAMDF)와 일반적인 ACF나 AMDF의 지연 정보 추정 성능에 미치는 잡음의 영향을 보여주기 위해 유성음 구간에서의 함수값을 나타낸 것이다. 잡음의 영향이 비교적 적은 저주파 대역의 입력 신호만으로 지연 정보를 추정하기 위하여 차단 주파수가 1 kHz인 3차 Cauer 저역 통과 여파기를 기존 방법과 제안된 방법 모두에서 사용하였다. 기존의 방법의 경우, Hamming 창함수와 preemphasis 과정이 피치 추정 이전의 전처리 과정으로 적용되었다. 그림 6과 7에서 알 수 있듯이, 입력 신호에 잡음의 영향이 커지면 커질수록 ACF나 AMDF는 왜곡이 많이 되고 그만큼 지연 정보를 추정하기도 어려워진다. 반면 RACF나 RAMDF는 잡음의 영향에 둔감하므로 비교적 안정되게 지연 정보를 추정할 수 있음을 볼 수 있다.

그림 8은 합성음 /아/에 입력 SNR이 0 dB가 되도록 잡음을 인가하여 NLMS-RAMDF 시스템에 의해 처리된 음성신호의 LPC 스펙트럼이다. 그림 8에서, 제안된 시스템에 의해 처리된 음성신호의 스펙트럼(short-dashed line)이 처리전의 잡음음성의 스펙트럼(long-dashed line)보다 원 음성의 스펙트럼

(solid line)에 훨씬 가까워짐을 보여 준다.

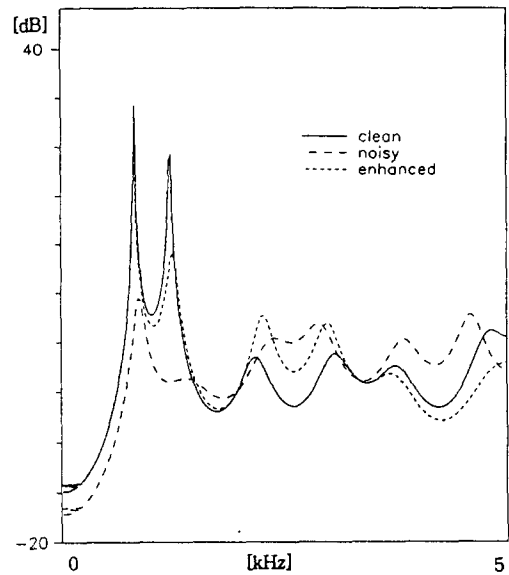


그림 8. NLMS-RAMDF 시스템에 의해 처리된 유성음 구간의 스펙트럼 (입력 SNR = 0 dB)

Fig. 8. LPC spectrum of a voiced sound processed by the NLMS-RAMDF system. (Input SNR = 0 dB)

그림 9는 10 kHz로 샘플링된 잡음이 없는 문장 "북풍은 힘껏 불었습니다"에 대해 재귀적 지연추정기를 이용하여 지연 정보를 구하는 과정에서 얻어지는 주기 측도, 유/무성음 정보 및 지연 정보 등의 contour를 나타낸 것이다. 재귀적 지연추정기가 한번에 고려하는 음성신호의 샘플 수는 201 샘플이고, 지연 정보를 구하기 위해 재귀적 AMDF 함수가 사용되었으며, 주기 측도로부터 유/무성음을 결정하기 위한 문턱값은 실험적으로 450으로 하였으며 목음 구간도 무성음 구간으로 분류하여 처리하였다. 그림 9에서 2 초이후의 지연 정보 contour에 double 피치가 검출되는 것을 볼 수 있다. Double 피치만큼 지연된 입력신호는 원래의 음성신호와 높은 상관관계를 갖는 기준신호가 되므로 double 피치도 지연 정보로 사용할 수 있다.

기존 방법과의 비교를 위하여 그림 10에는 그림 9에 나타낸 문장에 대해 매 분석 구간마다 구한 유/무성음 정보 및 피치 정보의 contour를 나타내었다. 그림 10에서, 비선형 평활화 과정을 거쳤으므로 상관관계가 높은 지연 정보인 double 피치가 제거되었다.

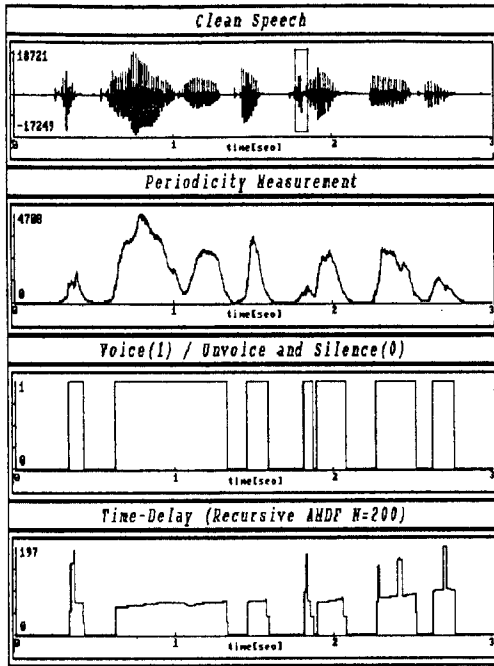


그림 9. 주기 측도 및 지연 정보 추정
Fig. 9. Periodicity measurement and estimated time-delay.

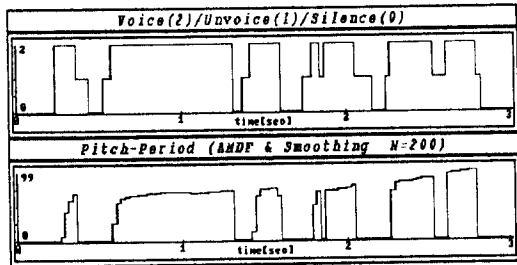


그림 10. 기존 방법에 의해 구해진 유/무성음 정보와 피치 정보
Fig. 10. Voice/Unvoice decision and estimated pitch using the conventional method.

잡음음성에 대해 일반적인 ACF와 AMDF를 이용해서 피치 정보 및 /유/무성음 정보를 추정할 경우 많은 오차가 발생되므로 본 연구에서는 기존 방법을 사용한 대부분의 음성개선 실험에서는 잡음이 섞이지 않은 원래의 음성신호에서 20ms의 분석 구간으로 매 10ms씩 이동시키면서 구한 피치 정보와 유/무성음 정보를 그대로 사용하였다.

표 2. 처리된 음성의 출력 SNR (dB)
Table 2. SNR of the processed speech. (in dB)

Input SNR	NLMS		RLS	
	RACF	RAMDF	RACF	RAMDF
-10.0	1.73	0.92	-0.79	-1.10
-5.0	4.29	3.78	3.02	2.87
0.0	6.79	6.68	6.58	6.62
5.0	8.49	8.45	9.13	9.13
10.0	9.67	9.63	10.76	10.79

표 3. 처리된 음성의 켈스트럼 거리
Table 3. Cepstral distance of the processed speech.

Input SNR[dB]	Noisy Speech	NLMS		RLS	
		RACF	RAMDF	RACF	RAMDF
-10	3.88	3.25	2.96	3.55	3.34
-5	3.57	2.74	2.70	3.12	3.01
0	3.20	2.62	2.52	2.65	2.63
5	2.79	2.63	2.61	2.48	2.43
10	2.35	2.53	2.48	2.22	2.16

표 2와 표 3은 제안된 방법에 의해 처리된 음성의 출력 SNR과, 켈스트럼 거리를 각각 입력 SNR에 대하여 나타낸 것이며, 그림 11은 표 2의 결과를 그래프로 나타낸 것이다.

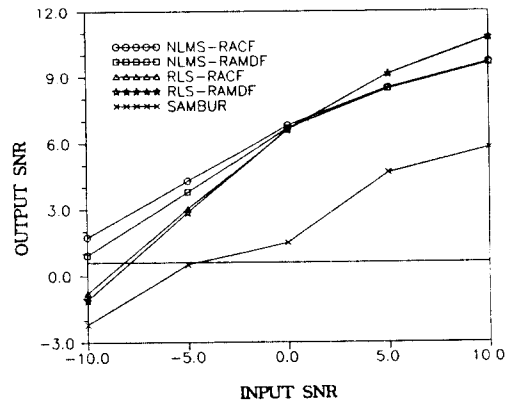


그림 11. 처리된 음성의 출력 SNR
Fig. 11. SNR of the processed speech signal.

재귀적 지연추정기가 한번에 고려하는 음성신호의 범위는 10 kHz의 sampling rate에서 201 샘플로 하였으며, 모든 적응여파기의 차수는 14차이다. NLMS의 알고리즘의 수렴인자 μ 는 0.01, RLS 알고리즘의

가중치 인자는 0.99를 사용하였다. 표 2 및 3에서 알 수 있듯이, 입력 SNR이 10 dB인 경우에는, 적응처리 전과 후의 음질(SNR, 켈스트럼 거리)은 거의 변하지 않았다. 그러나, 잡음의 정도가 심하면 심할수록 음질은 더욱 많이 개선됨을 알 수 있다. 대체로, 입력 SNR이 높은 경우에는 RLS 시스템이 NLMS 시스템보다 약간 나은 결과를 보였지만 입력 SNR이 낮은 경우에는 NLMS 시스템이 RLS 시스템보다 약간 나은 결과를 보였다. 또한 제안된 방법이 Sambur의 방법에 비해 더 좋은 결과가 얻어짐을 볼 수 있다.

IV. 결 론

배경 잡음이 내재된 음성신호에서 잡음 성분을 제거하기 위하여 재귀적 지연추정기를 갖는 단일입력 적응잡음제거 시스템을 제안하였으며, 정규화된 최소평균자승(NLMS: normalized LMS) 및 재귀적 최소자승(RLS) 적응알고리즘을 사용하는 적응잡음제거 시스템을 구성하여 음성개선 실험을 하였다. 제안된 방법에서는 매 샘플시간마다 적절한 지연 정보를 적응잡음제거 시스템에 전달하므로 기준 신호를 얻기 위한 지연 정보를 구할 때 음성 분석기간 사이의 불연속성이 제거되었으며, 무성음의 처리도 용이하게 된다.

제안된 시스템을 적용한 음성개선 실험 결과, 출력 SNR이 향상되고 켈스트럼 거리도 단축되는 등 처리된 음성의 음질이 Sambur 방법에 의한 결과보다 더 개선되었다. 입력 SNR이 높은 경우에는 RLS 시스템이 NLMS 시스템보다 약간 나은 결과를 보였지만 입력 SNR이 낮은 경우에는 NLMS 시스템이 RLS 시스템보다 약간 나은 결과를 보였다. Informal한 청취시험에서도 대체적으로 Sambur 방법보다 제안된 방법에 의한 출력음성의 음질이 양호함을 보여 주었다.

* 이 연구는 1993년도 한국과학재단 연구비지원에 의한 결과임.

과제번호 : 931-0900-066-1

參 考 文 獻

- [1] L.W. Rabiner and R.W. Schafer, Digital Processing of Speech Signals, Prentice-Hall, New Jersey, 1978.
- [2] D. O'Shaughnessy, Speech Communication, Addison-Wesley, New York, 1990.
- [3] D.R. Reddy, "Speech recognition by machine: a review," Proc. IEEE, vol. 64, pp.502-531, April 1976.
- [4] J. Flanagan, C. Coker, L. Rabiner, R. Schafer, and N. Umeda, "Synthetic voice for computers," IEEE Spectrum vol.7, pp.22-45, 1970.
- [5] J.S. Lim, Speech Enhancement, Prentice-Hall, New Jersey, 1983.
- [6] Y. Ephraim, "Statistical-model-based speech enhancement systems," Proc. of IEEE, vol. 80, pp. 1526-1555, Oct. 1992.
- [7] J.R. Deller, Jr., J.G. Prokis, and J.H. L. Hansen, L.J. Griffiths, Discrete-Time Processing of Speech Signals, Macmillan Publishing Company, New Jersey, 1993.
- [8] M.R. Sambur, "Adaptive noise canceling for speech signals," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-26, pp.419-423, Oct. 1978.
- [9] 강해동, 배건성, "음성신호의 적응잡음제거에 관한 연구," 제10회 음성통신 및 신호처리 워크샵 논문집, 제 SCAS-10권 1호, pp.145-148, 1993년 8월
- [10] L.R. Rabiner, "On the use of autocorrelation analysis for pitch detection," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-25, pp. 24-33, Feb. 1977.

著者紹介



姜海東(正會員)

1961年 10月 9日生. 1987年 2月
경북대학교 전자공학과 졸업(공학
사). 1989年 2月 경북대학교 대학
원 전자공학과 졸업(공학석사).
1994年 2月 경북대학교 대학원 전
자공학과 졸업(공학박사). 주관심

분야는 음성신호처리, 적응신호처리 등임.



裴建星(正會員)

1953年 11月 9日生. 1977年 2月
서울대학교 전자공학과 졸업(BS).
1979年 2月 한국과학기술원 전기 및
전자공학과 졸업(MS). 1989年 5
月 미국 Univ. of Florida(Ph.
D). 주관심분야는 디지털 신호처

리, 음성신호처리, 디지털 통신 등임.