DBpia

論文 95-5-5-18

단층 인식자를 이용한 배경 잡음 제거에 관한 연구

正會員 田 先 燾*, 姜 哲 豪*

A Study on Background Noise Cancellation Using Single-layer Perceptron

Sun-Do June*, Chul-Ho Kang* Regular Members

要 約

본 연구는 활성화 함수가 있는 단층 인식자를 이용한 배경 잡음 제거 에 관한 연구이다. 이 단층 인식자는 시그모이드 함수를 갖는 활성화 함수 이고 단층의 네트워크로 구성되며 정규화된 최소 자승 평균(NLMS) 알고리즘으로 학습된다.

특히, 활성화 함수를 통계적 방법으로 분석하였고, 확률적 배경 잡음 의 기대치와 근사화시켰다. 컴퓨터 실험은 배경 잡음으로 열화된 음성에서 원 음성만을 추출하는 실험을 하였다. 이 배경 잡음은 자동차의 엔진 잡음과 여러 사람들이 말하는 혼합 음성을 사용하였다.

비교 평가치는 신호대 잡음비(SNR), 자승근 오차(RSE), 시간 영역과 주파수 영역에서의 스펙트로그램과 음성을 직접 들어보는 주관적 평가로 설정하였다. 실험 결과 제안한 단층 인식자는 정규화된 NLMS 적응 필터 보다 성능이 우수함을 알 수 있었다.

ABSTRACT

In this paper, single-layer Perceptron with activation function is applied to the background noise cancellation. This single-layer Perceptron system employs normalized least mean square(NLMS) algorithm, sigmoidal activation function, and one-layer network.

Especially, sigmoidal funtion is analyzed by statistical method, and approximated to the probabilistic background noise expectation. Through computer simulation the original speech has been extracted from the mixed signal with background noise. This background noise is mobile's engine sound and composite speech sound of many speakers.

* 광운대학교 전자통신공학과
Dept. of Electronic Comm. Eng., Kwangwoon Univ.
論文番號 : 94365-1222
接受日字 : 1994年 12月 22日

www.dbpia.co.kr

The criteria for performance evaluation is signal-to-noise ratio(SNR), root square error(RSE), spectrogram in the time and frequency domain, and hearing test as a subjective evaluation. Experimental results show that the performance of the proposed single-layer Perceptron is better than that of the adaptive filter with NLMS.

제 1 장 서 론

잡음 제거에 관한 연구는 주위 환경의 소음을 제거하고 잡음이 제거된 상태의 환경을 유지하는 능동 잡음 제거(active noise cancellation)에 관한 연구와 시스템에서 원하는 신호를 처리하기 위해 배경 잡음을 제거하는 연구로 나눌 수 있다. 통신이나 음성 정보 시스템에서 정보 신호나 음성이 잡음으로 열화된 상태에서도 잡음을 제거하고 원하는 신호만을 얻은 후 시스템에서 처리를 하여야 하는데, 결국 어떤 시스템이든지 배경 잡음 제거는 최적의 시스템을 유지하기 위한 선행 과정이 되어야 한다.

특히, 음성 정보 시스템에서는 음성 자체가 잡음(시스템에서 처리하고자 하는 음성 이외에 시스템에 입력되는 음성 내지 소음)과 상관 관계가 존재하기 때문에 원하는 음성만을 추출하기란 어려운 문제이다. 이런 이유로 배경 잡음에서 음성만을 추출해내는 연구는 현재에도 여러 학교 및 연구 단체에서 수행되고 있다. 그 예로서 MIT, Stanford, Tel-Aviv, Purdue 등 여러 대학과 Bell, Naval 연구소 등지에서 연구되는 과정이고 국내에서도 각 대학에서 연구 발표되는 추세에 있다.[8]

잡음 제거에 사용되는 방법은 크게 주파수 차감법(spectral subtraction), Wiener 필터링(Wiener filtering), 적응 필터링(adaptive filtering)으로 구분할 수 있다.[2] 주파수 차감법은 그 구현에 있어 계산량이 많아 실시간 처리에 적합치 않다. Wiener 필터링 방법 역시 계산량이 많으며 음성과 잡음이 상관 관계가 있을시에는 잡음만이 아니라 음성도 같이 제거되는 문제점이 있다. 적응 필터의 경우 least mean square(LMS) 알고리즘에 의해 실시간 처리가 가능하지만 Wiener 계수를 찾아가기까지 수렴 시간이 문제가 되며 이 방법 역시 음성과 잡음의 상관 관계가 존재할 경우에 Wiener 필터링 방법과 같은 문제점이 있다. 결국 현재까지 잡음 제거 방식 중에서 가장 최적이라고 선정되어 있는 것은 없다.

본 연구는 음성처리의 효율성을 위한 배경 잡음 제거를 시그모이드(sigmoid) 활성화 함수(activation function)의 단층 인식자(single-layer Perceptron)를 이용하였다. 제안한 인식자는 잡음 제거의 실시간 구현을 위한 LMS 방법으로 음성과 잡음이 상관 관계가 있을시에도 기존의 방법에서 있었던 문제점을 극복하려는 방법이다. 이러한 방법은 기존의 적응 필터[1][2]와는 달리 비선형 활성화 함수를 이용하는 방법이며 비선형 필터[10]에서 처럼 신경망을 선형 필터 전단에 구성하여 잡음 제거를 하는 기존의 방식[6][7]과도 다른 방법이다. 특히, 적응 필터에서 배경 잡음의 확률적 기대치를 적응 등화기에 주로 이용되는 Bussgang 알고리즘[3]과 비선형적인 estimation의 이론[4][5]을 변형 분석하여 인식자의 활성화 함수와 근사시켰다. 그리고 실제의 음성과 배경 잡음으로 모의 실험하였고, 제안한 인식자에서 잡음 제거된 음성을 기존의 적응 필터에서 잡음 제거한 음성과 시간 및 주파수 영역에서 비교했으며 SNR(signal-to-noise ratio) 및 RSE(root square error)로서 그 성능을 평가하여 그 우수성을 증명하였다.

이러한 연구와 구현은 그 활용에 있어 음성 및 화자인식 시스템이나 휴대용 무선 전화기 등에 이용되리라 기대된다.

제 2 장 기존의 적응 잡음 제거 방법

2.1. 잡음 제거를 위한 환경

그림 1은 배경 잡음 제거를 위한 환경이다. 이때 주 마이크(primary sensor)에서는 배경 잡음으로 열화된 음성 신호가 감지된다. 기준 마이크(reference sensor)에서는 배경 잡음이 감지된다. 여기서 음성 신호를 열화시키는 잡음은 기준 마이크에서 감지되는 잡음에서

www.dbpia.co.kr

미지의 채널을 통과한 신호가 된다.

2.2 적응 필터링 방법

적응 필터는 에러 신호를 최소화하는 방향으로 weight를 학습시켜 가는 LMS 알고리즘으로서 잡음 제거 및 반향기, 등화기 등 다양한 응용이 가능하다[1]. 잡음 제거에 응용할 경우 그 모식도는 그림 2와 같다.

이 때 s(n)은 음성 신호이고 x(n)은 배경 잡음 신호이며 h(n)은 두 마이크 사이의 채널이다. 이 때 배경 잡음 x(n)은 LMS 알고리즘으로 학습되는 weight w(n)과 내적되어 잡음을 추정한 후 잡음을 제거 시킨 음성 신호 e(n)을 얻는다. 여기서 weight의 학습은 LMS 알고리즘으로 다음과 같다.

$$\omega_i(n+1)=\omega_i(n)+\mu e(n)x(n-i) \quad (1)$$

여기서 μ는 수렴 계수로서 $0<\mu<\frac{2}{\sum_{i=0}^{N} x^2(n-i)}$ 의 범위에 있는 상수값을 선택하여야 안정하다. 또한 너무 적은 값을 선택할 경우 수렴이 늦어지게 된다. 이 때 N은 필터 차수이다.

제 3 장 제안한 단층 인식자

3.1. 잡음의 기대치와 시그모이드 함수와의 관계

두 개의 마이크의 구조를 갖는 적응 필터에서 필터의 출력값은 다음과 같다.

$$y=\sum_{i=0}^{N} w_i(n)x(n-i) = W^T X \quad (2)$$

이때, Wiener 계수인 최적 weight를 $\bar{\omega}(n)$라고 하며 이때의 weight는 $N+1$개의 차원을 갖는다. 그리고 최적 weight의 행렬을 $\overline{W}$로서 표시하고, X는 필터에 시간지연된 입력값들의 행렬이다. 이때 (2)식은 (3)식과 같이 쓸 수 있다.

$$\begin{aligned} y &= \overline{W}^T X + W^T X - \overline{W}^T X \\ &= \overline{W}^T X + (W-\overline{W})^T X = \overline{W}^T X + v \end{aligned} \quad (3)$$

여기서 v는 적응 필터가 Wiener 계수를 찾아가지 못한 정도가 된다.

이때, 주 마이크에 감지되는 간섭된 잡음 x_t의 필터 출력에서의 기대치는 다음과 같다.

$$\overline{x_t} = E(x_t|y) = \int_{-\infty}^{\infty} x_t\, p(x_t|y)dx_t \quad (4)$$

을 서로 독립이라고 보면, 간섭된 잡음의 기대값은 부록 1에 의해 다음과 같다.

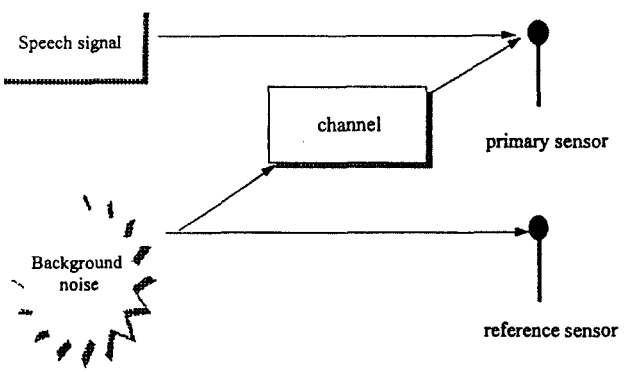

그림 1. 배경 잡음 제거를 위한 환경
Fig. 1. Environment for background noise cancellation

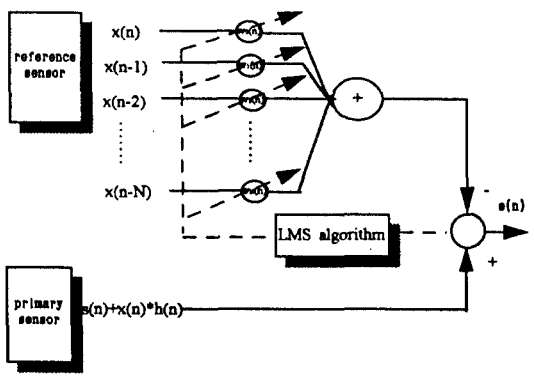

그림 2. 잡음 제거에 응용된 적응 필터
Fig. 2. Noise cancellation using adaptive filter

www.dbpia.co.kr

$$E(x_t|y) = \frac{\int_{-\infty}^{\infty} x_t \exp(-\frac{1}{2\delta_v^2}(y^2-2x_t\overline{W}^TB + x_t^2|\overline{W}|^2|B|^{2+}\frac{x_t^2\delta_v^2}{\delta_{x_t}^2}))dx_t}{\int_{-\infty}^{\infty} \exp(-\frac{1}{2\delta_v^2}(y^2-2x_t\overline{W}^TB + x_t^2|\overline{W}|^2|B|^{2+}\frac{x_t^2\delta_v^2}{\delta_{x_t}^2}))dx_t} \quad (5)$$

식(5)에서 기준 마이크와 주 마이크 사이의 미지의 채널을 행렬 A라고 할 때 $X^TA = x_t$이 된다. 이때, $X=Bx_t$의 관계가 있는 A의 역 형태로서 행렬 B가 존재한다. 또, A와 W의 차원을 같다고 가정한다면, B의 차원도 같게 된다. 그리고 x_t, v를 가우시안 분포라고 가정하고, δ_v, δ_v는 x_t, v의 편차 이다.

이때, $\delta_v^2 \ll \delta_{xt}^2$ 가 되게 하면 식 (5)에서 $\frac{\delta_v^2}{\delta_{x_t}^2} \simeq 0$ 이 된다.

단, (5)식은 다음 식을 만족한다.

$$2\pi\delta_x\delta_w p(y) = \int_{-\infty}^{\infty} \exp(-\frac{1}{2\delta_v^2}(y^2-2x_t\overline{W}^TB + x_t^2|\overline{W}|^2|B|^{2+}\frac{x_t^2\delta_v^2}{\delta_{x_t}^2}))dx_t$$

식(5)에서 x_t의 범위가 n_0에서 $-n_0$라고 가정하면, v의 범위는 $y-n_0\overline{W}^TB$에서 $y+n_0\overline{W}^TB$가 되고, 부록2에서 다음 식을 얻는다.

$$E(x_t|y) = \frac{y}{\overline{W}^TB} - \frac{\delta_v\,\delta_y}{(\overline{W}^TB)^2\sqrt{2\pi}\,\delta_{x_t}\exp(-\frac{y^2}{\delta_y^2})}\left[\exp(-\frac{(y-n_0\overline{W}^TB)^2}{2\delta_v^2}) - \exp(-\frac{(y+n_0\overline{W}^TB)^2}{2\delta_v^2})\right] \quad (6)$$

음성과 잡음이 상관 관계가 없다면 $\overline{W}^TB=1$ 이지만 음성과 잡음은 어느 정도 상관 관계를 가지고 있기 때문에 임의의 값을 갖게 된다.

결국 식(6)의 $\delta_v, \delta_{xt}, P(y), \overline{W}^TB$와 같은 파라미터를 적당한 값으로 주어졌을 때 식(7)에서 알맞는 λ_t(굴곡계수)과 F(활성화 함수의 최대값)을 변해줌에 따른 인식자의 시그모이드 활성화 함수 $f(y(n))$에 근사한다.

$$f(y(n)) = F\times(\frac{2}{1+\exp(-\lambda_t y(n))}-1) \quad (7)$$

그림 3은 식 (7)의 시그모이드 활성화 함수를 나타낸 것으로 λ_t을 0.01, 0.02, 0.035 로 선정하여 a, b, c 경우로 나누어서 표시된 그림이다.

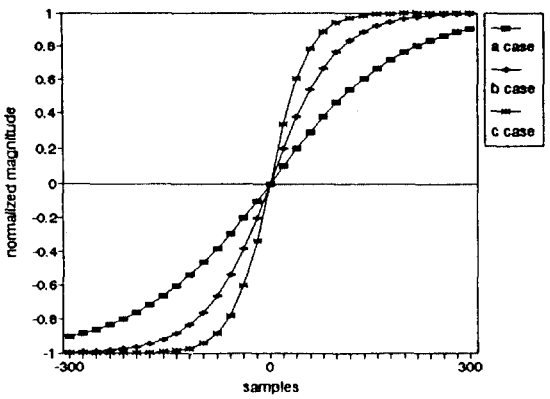

그림 3. 시그모이드 활성화 함수
Fig. 3. Sigmoid activation function

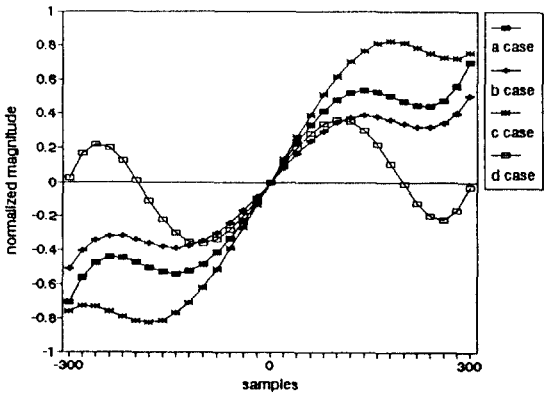

그림 4. 잡음 기대치
Fig. 4. Noise expectation

www.dbpia.co.kr

그림 4는 식 (6)의 각 파라미터에 따른 굴곡 정도를 보여 주는 그림이다.

a는 $\frac{1}{\overline{W}^T B}$ 이 720값을 가지며 $\frac{\delta_v \delta_y}{(\overline{W}^T B)^2 \sqrt{2\pi}\delta_{xt}}$ $n_0 \overline{W}^T B$이 290 값을 갖는다. 또, $\frac{1}{2\delta_v^2}$ 이 5×10^{-5}값을 갖고 $\frac{1}{\delta_y^2}$ 이 2×10^{-7}값을 갖는다.

b는 a와 모두 동일하고 단지 $\frac{1}{\overline{W}^T B}$ 만이 2.6값으로 변화된 형태이고 c는 a에서 $n_0 \overline{W}^T B$이 330값으로 변화된 형태이며 d는 a의 경우에서 $\frac{\delta_v \delta_y}{(\overline{W}^T B)^2 \sqrt{2\pi}\delta_{xt}}$ 이 1080값으로 변화된 형태이다.

결국 그림 3과 4에서 시그모이드 활성화 함수의 굴곡 계수와 잡음 기대치의 여러 파라미터 값에 의해 굴곡 정도의 유사성을 확인할 수 있다.

3.2. 단층 인식자의 구조

잡음의 기대치는 인식자의 활성화 함수와 근사되는 이유로 배경 잡음에 사용할 인식자의 구조는 다음과 같다.

그림. 5의 주 마이크에서 잡음으로 열화된 음성 신호가 감지되고 기준 마이크에서 잡음원이 감지된다. 이 때 e(n)은 인식자를 통해 잡음이 제거된 음성 신호가 된다.

3.3. 단층 인식자의 학습 방법

weight의 학습 알고리즘은 normalized least mean square(NLMS)로서 어떤 상태이든 안정하게 했으며 알고리즘의 에러 값은 원하는 신호에 활성화 함수를 통과한 출력값의 차가 되며 그 식은 다음과 같다.

$$w_i(n+1) = w_i(n) + \frac{\mu_m}{\sum_{i=0}^{N} x^2(n-i)}(d(n) - F\times(\frac{2}{1+\exp(-\lambda_D y(n))} - 1)) \quad (8)$$
$$x(n-i)$$

이때, 수렴 계수 μm의 범위는 $0<\mu<2$이다.

제 4 장 컴퓨터 시뮬레이션 결과 및 고찰

4.1. 컴퓨터 시뮬레이션 결과

"저는 신호처리 연구실의 전선도 입니다." 라는 문장을 입력되는 음성 신호로 하였고, 음성과의 상관 값이 65893.0인 소형 자동차 안에서 녹음한 엔진 잡음을 배경 잡음으로 설정(A 경우)한 경우와 음성과의 상관 값이 272312.5로서 A경우 보다 4배 이상의 상관 값을 갖는 시장에서 여러 사람들이 떠드는 혼합된 음성을 배경 잡음으로 사용한 경우(B 경우)에 대해 주 마이크(primary sensor)와 기준 마이크(reference sensor)사이의 다양한 채널 형태를 가정하여 실험하였다.

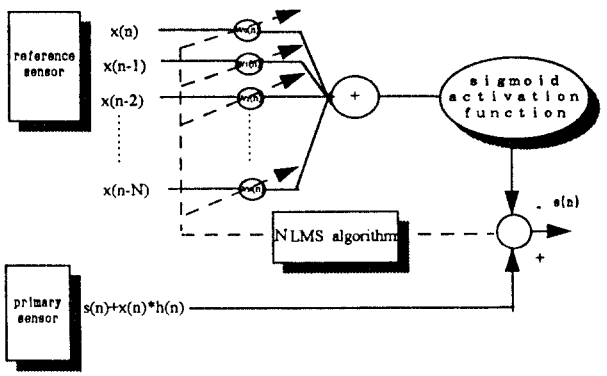

그림 5. 단층 인식자를 이용한 배경 잡음 제거
Fig. 5. Noise cancellation using single-layer Perceptron

www.dbpia.co.kr

실험은 NLMS의 인식자를 NLMS의 적응 필터와 비교하여 신호대 잡음비를 표1에서 표5까지 결과를 표시하였다. 이때 음성과 잡음 신호의 샘플링 주파수는 10 kHz로 하였다. 채널 형태는 기준 마이크에서 감지되는 배경잡음이 1/2배와 1/3배씩 감쇄되는 경우를 가정하여 실험 하였다. 그리고 저역 통과 필터(low-pass filter)와 대역 통과 필터(band-pass filter)의 특성을 갖는 채널과 미지의 채널을 가정한 경우도 실험하였으며 50차의 시간적인 지연을 갖는 채널도 가정하였다.

인식자의 시그모이드 활성화 함수 $f(y(n))$는 식(12)과 같다.

$$f(y(n)) = th \times Npk \times \left(\frac{2}{1+\exp(-\lambda y)} - 1\right) \quad (9)$$

여기서, $y(n)$은 입력과 weight들의 내적이고, 는 활성화 함수의 굴곡 정도를 보여준다. Npk는 기준 마이크로 감지되는 배경잡음의 최대값이고 th는 1에서 6정도의 값을 취함으로써 활성화 함수의 최대값을 결정했다.

weight의 학습 방법은 식(8)에 따르고 μdm은 1/N

표 1. 감쇄 특성을 갖는 채널에서의 SNR(A 경우)
Table 1. SNR for attenuation channel(A case)

채널 형태	입력	적응 필터	단층 인식자											
			th=1			th=1.5			th=2			th=3		
			λ = .0005	λ = .0007	λ = .0009	λ = .0005	λ = .0007	λ = .0009	λ = .0005	λ = .0007	λ = .0009	λ = .0005	λ = .0007	λ = .0009
Attenu-ation (1/2배)	-11.36	-5.10	3.70	3.81	3.59	3.79	3.37	2.89	3.45	2.81	2.26	2.66	1.92	1.37
	-6.25	-0.03	5.71	5.73	5.45	5.68	5.18	4.62	5.26	4.52	3.88	4.34	3.46	2.80
Attenu-ation (1/3배)	-8.75	-2.50	4.71	4.25	3.73	4.11	3.63	2.75	3.47	2.66	2.07	2.49	1.71	1.14
	-5.65	0.55	5.92	5.41	4.85	5.26	4.44	3.77	4.57	3.67	2.99	3.48	2.59	1.97

표 2. 필터 특성을 갖는 채널에서의 SNR(A 경우)
Table 2. SNR for filter channel (A case)

채널 형태	입력	적응 필터	단층 인식자											
			th=3			th=4			th=5			th=6		
			λ = .0005	λ = .0007	λ = .0009	λ = .0005	λ = .0007	λ = .0009	λ = .0005	λ = .0007	λ = .0009	λ = .0005	λ = .0007	λ = .0009
low pass	-5.23	2.75	7.38	7.53	7.29	7.55	7.23	6.71	7.39	6.77	6.11	7.09	6.29	5.57
band pass	-4.95	3.01	7.54	7.63	7.37	7.66	7.29	6.75	7.47	6.81	6.14	7.15	6.32	5.59
unkno-wn(10)	-6.39	3.49	8.06	7.94	7.55	8.05	7.50	6.87	7.73	6.93	6.20	7.32	6.39	5.62
unkno-wn(15)	-9.35	1.14	6.42	6.70	6.59	7.22	7.02	6.59	7.28	6.71	6.08	7.05	6.28	5.56
unkno-wn(20)	-11.41	-0.87	3.69	4.10	4.17	5.44	5.60	5.42	6.11	5.93	5.52	6.26	5.80	5.24

www.dbpia.co.kr

표 3. 감쇄 특성을 갖는 채널에서의 SNR(B 경우)
Table 3. SNR for attenuation channel (B case)

채널 형태	입력	적응 필터	단층 인식자 th=1 λ=.0005	th=1 λ=.0007	th=1 λ=.0009	th=1.5 λ=.0005	th=1.5 λ=.0007	th=1.5 λ=.0009	th=2 λ=.0005	th=2 λ=.0007	th=2 λ=.0009	th=3 λ=.0005	th=3 λ=.0007	th=3 λ=.0009
Attenu-ation (1/2배)	-12.28	-6.36	3.86	3.99	3.75	3.96	3.49	2.95	3.58	2.86	2.22	2.68	1.81	1.13
	-17.39	-11.48	2.33	2.41	2.13	2.40	1.89	1.35	1.99	1.26	0.60	1.08	0.14	-0.78
Attenu-ation (1/3배)	-11.69	-5.77	4.16	3.69	3.15	3.55	2.76	2.11	2.88	2.00	1.33	1.82	0.89	0.14
	-14.78	-8.87	3.17	2.68	2.15	2.54	1.77	1.12	1.88	1.02	0.30	0.82	-0.21	-1.30

표 4. 필터 특성을 갖는 채널에서의 SNR(B 경우)
Table 4. SNR for filter channel (B case)

채널 형태	입력	적응 필터	단층 인식자 th=1 λ=.0005	th=1 λ=.0007	th=1 λ=.0009	th=41.5 λ=.0005	th=41.5 λ=.0007	th=41.5 λ=.0009	th=2 λ=.0005	th=2 λ=.0007	th=2 λ=.0009	th=3 λ=.0005	th=3 λ=.0007	th=3 λ=.0009
low pass	-12.74	-2.30	5.62	4.83	4.18	4.71	3.77	3.04	3.90	2.94	2.22	2.73	1.79	1.10
band pass	-12.47	-2.04	5.68	4.89	4.20	4.73	3.78	3.05	3.92	2.94	2.23	2.74	1.79	1.10
unkno-wn(10)	-1160	-1.60	-1.23	-1.11	-1.11	2.80	3.22	3.37	4.71	5.01	5.01	5.81	5.63	5.28
unkno-wn(15)	-13.96	-3.41	-7.32	-7.80	-8.17	-2.89	-2.92	-3.02	0.74	0.90	0.90	4.24	4.29	4.12
unkno-wn(20)	-15.62	-5.51	-10.8	-11.2	-11.5	-8.23	-8.58	-8.85	-4.62	-4.73	-4.87	0.59	0.73	0.69

표 5. 시간 지연 특성을 갖는 채널에서의 SNR (A,B 경우)
Table 5. SNR for delay channel (A,B case)

입 력	적응 필터	단층 인식자 th=1 λ=.0005	th=1 λ=.0007	th=1 λ=.0009	th=1.5 λ=.0005	th=1.5 λ=.0007	th=1.5 λ=.0009	th=2 λ=.0005	th=2 λ=.0007	th=2 λ=.0009	th=3 λ=.0005	th=3 λ=.0007	th=3 λ=.0009
(case A) -4.26	1.87	2.75	3.56	4.13	4.29	5.15	5.70	5.23	5.98	6.35	6.21	6.57	6.56
(case B) -10.85	-4.45	-1.38	-0.50	0.06	1.00	2.09	2.78	2.42	3.44	3.98	3.90	4.49	4.60

www.dbpia.co.kr

으로 선정 했다. 이때 인식자의 차수 N은 그 성능이 떨어지지 않는 상태에서 최소값인 150차수로 실험했다.

이때 입력의 SNR은 $10\log \sum_{i=0}^{M} \frac{s^2(n)}{x_i^2(n)}$

여기서 s(n)은 음성신호이며 x_i (n)는 채널을 통과한 잡음 신호로서 주 마이크에서 음성을 열화시킨 잡음이고 M은 처리된 음성의 전체 샘플치이며 출력의 SNR은

$$10\log \sum_{i=0}^{M} \frac{s^2(n)}{(x_i(n)-f(y(n)))^2} \quad (11)$$

여기서 f(y(n))은 잡음으로 추정되는 인식자의 출력이다.

그림 6부터 그림 9는 가정한 여러 채널 중에서 10차의 미지의 필터에서 여러 사람들의 떠드는 혼합 음성을 배경 잡음으로 선택한 경우(B 경우)에 대해 원 음성 신

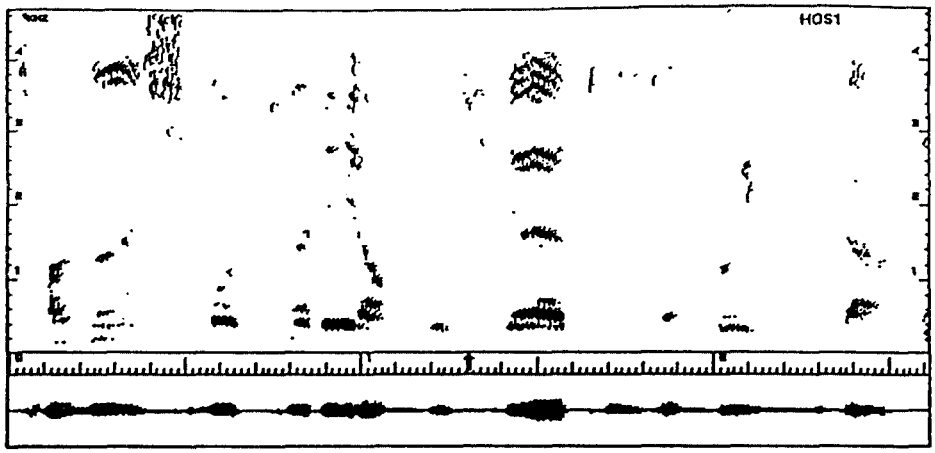

그림 6. 원 음성 신호
Fig. 6. Original speech

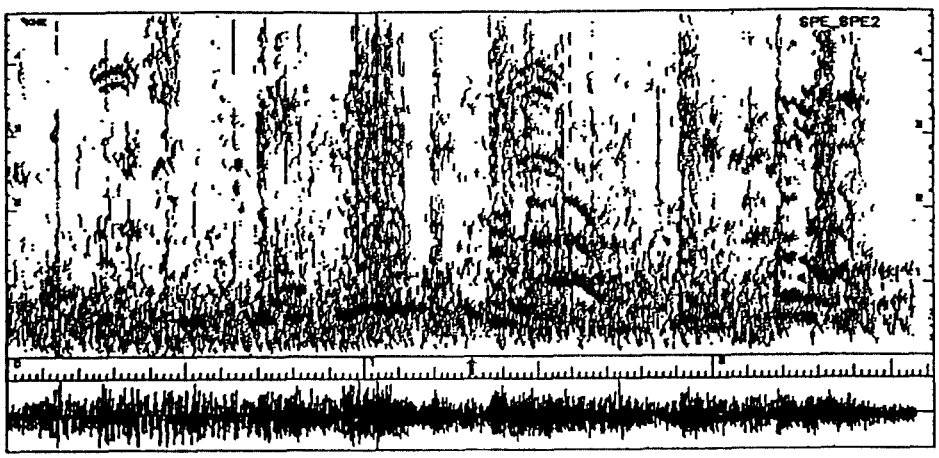

그림 7. 잡음으로 열화된 음성 신호(B경우)
Fig. 7. Speech corrupted by noise (B case)

www.dbpia.co.kr

배경 잡음으로 선택한 경우(B 경우)에 대해 원 음성 신호와 잡음으로 혼합된 음성 신호 그리고 적응 필터와 단층 인식자로 추출한 음성 신호를 스펙트로그램과 시간 영역에서의 파형으로 표시한 그림이다.

그림 10와 11는 채널을 10차의 미지의 필터로 가정한 경우에 적응 필터와 단층 인식자의 RSE(root square error)의 정규화 된 형태를 나타낸 것이다. 이때 RSE의 식은 다음과 같다.

$$RSE = \sqrt{(s(n)-e(n))^2} \quad (12)$$

단, e(n)은 적응 필터나 인식자에서 추출한 음성 신호가 된다.

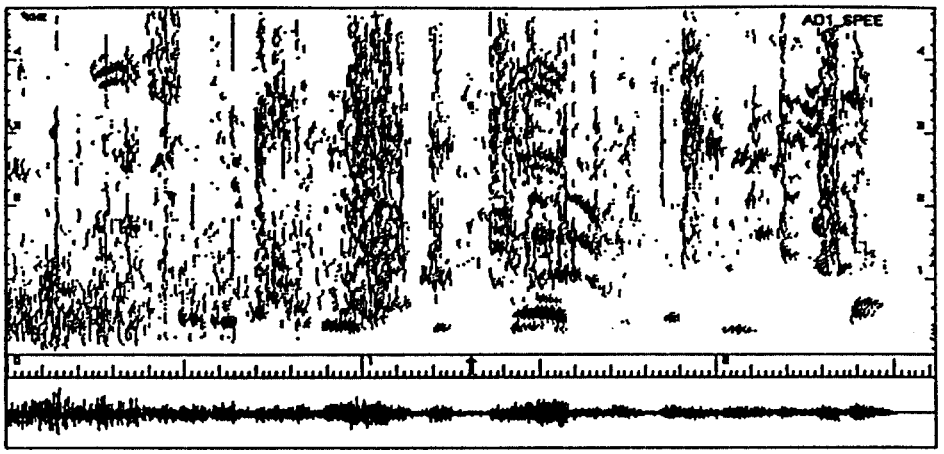

그림 8. 적응 필터로 얻은 음성 신호(B 경우)
Fig. 8. Speech enhanced by adaptive filter (B case)

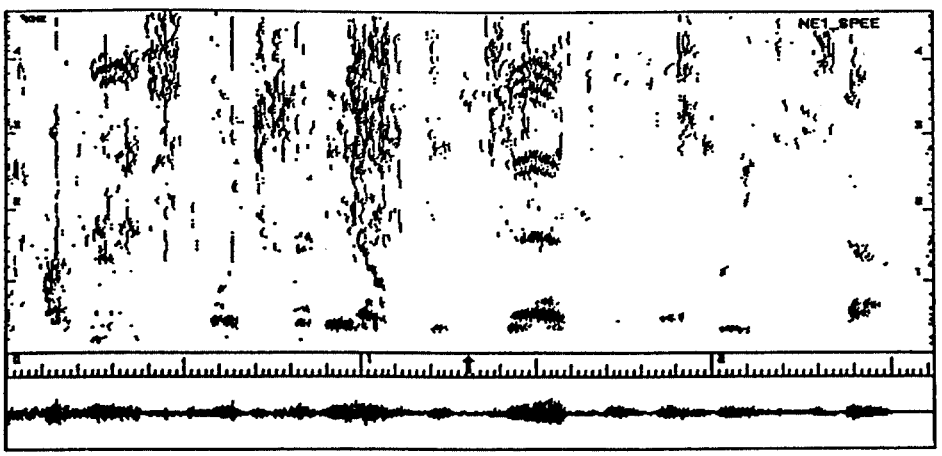

그림 9. 단층 인식자로 얻은 음성 신호 (B 경우)
Fig. 9. Speech enhanced by single-layer Perceptron (B case)

www.dbpia.co.kr

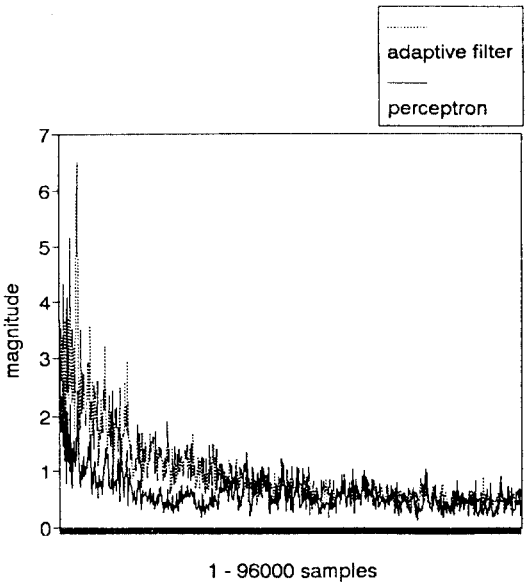

그림 10. 적응 필터와 단층 인식자의 RSE (A 경우)
Fig. 10. RSE for adaptive filter and single-layer perceptron (A case)

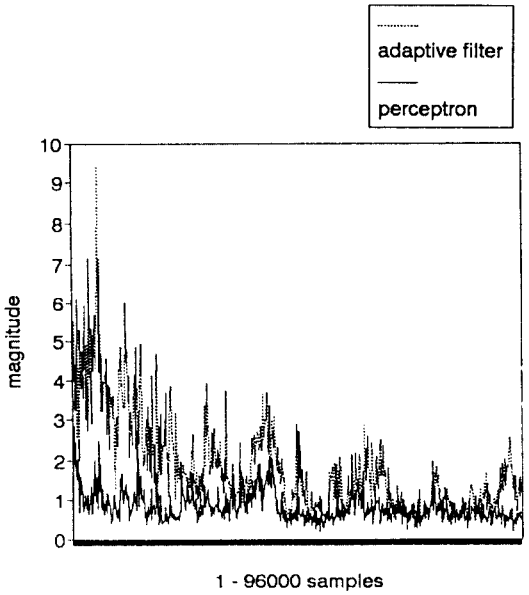

그림 11. 적응 필터와 단층 인식자의 RSE (B 경우)
Fig. 11. RSE for adaptive filter and single-layer perceptron (B case)

4.2. 고찰

기존의 방법인 적응 필터와 제안한 단층 인식자를 비교했을 때 활성화 함수의 파라미터 th와 λ를 조절함에 따라 SNR비의 개선을 보았다. 이것은 시그모이드 활성화 함수가 주 마이크에서 음성을 열화시킨 잡음의 기대치 역할을 하고 있음을 알 수 있다.

또한 스펙트로그램에서 적응 필터와 비교해 볼 때 단층 인식자에서 얻은 음성은 잡음의 전력 밀도가 더 많이 제거됨을 볼 수 있다. RSE에서는 초기의 RSE 보다 시간의 흐름에 따라 그 값의 감소를 보여주는데 이것은 수렴을 의미하며, 실험 결과는 단층 인식자가 적응 필터보다 빠른 수렴 상태에 도달하고 있음을 보여준다.

결국 실험 결과에서 볼 때, 단층 인식자에서 시그모이드 활성화 함수가 적응 필터에서 그 수렴을 빠르게 하는 역할과 3.1 단원에서 분석한 잡음의 기대치 역할을 수행함을 보여 주는 것이다.

제 5 장 결론

본 연구에서는 배경 잡음으로 열화된 음성으로부터 잡음을 제거하기 위해 시그모이드 활성화 함수가 있는 단층 인식자를 이용하였으며 기존의 적응 필터 방식과 성능을 비교하였다.

실험 결과는 시그모이드 활성화 함수의 굴곡 계수와 최대치를 적당히 조절함으로서 SNR은 NLMS 알고리즘의 적응 필터보다 2에서 7dB 정도 개선이 있었고, 잡음으로 열화된 음성 잡음보다 8에서 17dB 이상까지 SNR이 개선되었다. 그리고 사용된 두 마이크 사이의 채널을 여러 형태로 가정하여 실험하였는데 이러한 가정에 의한 실험은 실질적인 활용에 있어 어떤 상황에서도 배경 잡음의 제거를 보여주는 것이다.

또한 시간 영역에서의 파형 및 주파수 영역에서의 밀도 형태를 보았을 때 NLMS 알고리즘의 적응 필터보다

www.dbpia.co.kr

단층 인식자의 성능이 더 좋음을 알 수 있었고, RSE는 적응 필터보다 단층 인식자가 더 작은 값을 유지함으로 수렴의 향상을 알 수 있었다.

특히 A/D 변환과 D/A 변환이 내장된 DSP board를 사용하여 잡음 제거된 음성을 들었을 때 적응 필터를 통해 추출된 음성보다 단층 인식자를 통해 추출된 음성이 선명함을 알 수 있었다. 결국, 단층 인식자의 시그모이드 활성화 함수는 음성과 상관 관계가 있는 잡음의 기대치로서의 역할을 하고 있음을 알 수 있다.

그러나, 단층 인식자의 활성화 함수는 굴곡 계수와 최대치를 선정해 주어야 하는 문제가 있다. 본 연구에서는 가장 최적의 상태를 얻기 위해 여러 가지 경우를 실험하였는데 잡음의 형태와 음성과의 상관 관계로부터 최적의 굴곡 계수와 최대치를 자동적으로 찾는 방법은 앞으로의 연구 과제이다.

부록 1

$$E(x_t|y) = \int_{-\infty}^{\infty} \frac{x_t\, p(y|x_t)p(x_t)}{p(y)}\, dx_t$$

$$= \int_{-\infty}^{\infty} \frac{x_t\, p(v)p(x_t)}{p(y)}\, dx_t$$

$$= \frac{\int_{-\infty}^{\infty} x_t\, p(v)p(x_t)dx_t}{\int_{-\infty}^{\infty} p(y|x_t)p(x_t)dx_t}$$

$$= \frac{\int_{-\infty}^{\infty} x_t\, p(v)p(x_t)dx_t}{\int_{-\infty}^{\infty} p(v)p(x_t)dx_t}$$

$$= \frac{\int_{-\infty}^{\infty} \frac{x_t}{\delta_v\sqrt{2\pi}} \exp(-\frac{1}{2\,\delta_v^{\,2}}(y-x_t\,\overline{W}^T B))\, \frac{1}{\delta_{x_t}\sqrt{2\pi}} \exp(-\frac{1}{2\,\delta_{x_t}^{\,2}}\, x_t^{\,2})\, dx_t}{\int_{-\infty}^{\infty} \frac{1}{\delta_v\sqrt{2\pi}} \exp(-\frac{1}{2\,\delta_v^{\,2}}(y-x_t\,\overline{W}^T B))\, \frac{1}{\delta_{x_t}\sqrt{2\pi}} \exp(-\frac{1}{2\,\delta_{x_t}^{\,2}}\, x_t^{\,2})\, dx_t}$$

$$= \frac{\int_{-\infty}^{\infty} x_t \exp(-\frac{1}{2\,\delta_v^{\,2}}(y^2 - 2x_t\,\overline{W}^T B + x_t^{\,2}|\overline{W}|^2|B|^{2+} \frac{x_t^{\,2}\,\delta_v^{\,2}}{\delta_{x_t}^{\,2}}))dx_t}{\int_{-\infty}^{\infty} \exp(-\frac{1}{2\,\delta_v^{\,2}}(y^2 - 2x_t\,\overline{W}^T B + x_t^{\,2}|\overline{W}|^2|B|^{2+} \frac{x_t^{\,2}\,\delta_v^{\,2}}{\delta_{x_t}^{\,2}}))dx_t}$$

부록 2

$$E(x_t|y) = \frac{\int_{-\infty}^{\infty} (\frac{y-v}{\overline{W}^T B}) \exp(-\frac{1}{2\delta_v^{\,2}}(y - x_t\,\overline{W}^T B)^2)dx_t}{\int_{-\infty}^{\infty} \exp(-\frac{1}{2\delta_v^{\,2}}(y-x_t\,\overline{W}^T B)^2)dx_t}$$

$$= \frac{y}{\overline{W}^T B} + \frac{1}{(\overline{W}^T B)^2} \frac{\int_{\infty}^{-\infty} v\, \exp(-\frac{v^2}{2\,\delta_v^{\,2}})dv}{2\,\pi\,\delta_v\delta_{x_t}\, p(y)}$$

$$= \frac{y}{\overline{W}^T B} - \frac{1}{(\overline{W}^T B)^2} \frac{\int_{y-n_0\overline{W}^T B}^{y+n_0\overline{W}^T B} v\, \exp(-\frac{v^2}{2\,\delta_v^{\,2}})dv}{2\,\pi\,\delta_v\delta_{x_t}\, p(y)}$$

$$= \frac{y}{\overline{W}^T B} - \delta_v \frac{\exp(-\frac{(y-n_0\,\overline{W}^T B)^2}{2\,\delta_v^{\,2}}) - \exp(-\frac{(y+n_0\,\overline{W}^T B)^2}{2\,\delta_v^{\,2}})}{(\overline{W}^T B)^2\, 2\,\pi\,\delta_{x_t}\, p(y)}$$

$$= \frac{y}{\overline{W}^T B} - \delta_v\,\delta_y \frac{\exp(-\frac{(y-n_0\,\overline{W}^T B)^2}{2\,\delta_v^{\,2}}) - \exp(-\frac{(y+n_0\,\overline{W}^T B)^2}{2\,\delta_v^{\,2}})}{(\overline{W}^T B)^2\sqrt{2\,\pi}\,\delta_{x_t}\exp(-\frac{y^2}{\delta_y^{\,2}})}$$

www.dbpia.co.kr

참고 문헌

1. B. Widrow, Samuel D. Stearn, "Adaptive Signal Processing," Prentice-Hall, 1985.
2. John R. Deller, John G. Proakis, and John H. L. Hansen, "Discrete-Time Processing of Speech Signals," Macmillan Publishing Company, 1993.
3. Simon Haykin, "Adaptive Filter Theory," Prentice-Hall, 1991.
4. James L. Melsa, David L. Cohn, "Decision and Estimation Theory," McGraw-Hill, 1978.
5. Louis L. Scharf, "Statistical Signal Processing," Addision-Wesley Publishing Company, 1991.
6. W. G. Knecht, "Nonlinear Noise Filtering and Beamforming Using the Perceptron and its Volterra Approximation," *IEEE Transaction on Speech and Audio Processing*, pp.52-62, January, 1994.
7. T. Troudet, W. Merrill, "A Novel Approach to Noise-Filtering Based on a Gain-Scheduling Neural Network Achitecture," *World Congress on Neural Network*, pp.38-45, June ,1994.
8. Ehud Weinstein, Meir Feder, and Alan V. Oppenheim, "Multi-Channel Signal Separation by Decorrelation," *IEEE Transaction on Speech and Audio Processing*, Vol. 1, pp.405-413, October, 1993.
9. Meir Feder, Alan V. Oppenheim, and Ehud Weinstein, "Maximum Likelihood Noise Cancellation Using the EM Algorithm," *IEEE Transaction on Acoustics Speech and Signal Processing*, Vol. 37, pp.204-216, February, 1989.
10. John C. Stapleton, Steven C. Bass, "Adaptive Noise Cancellation for Class of Nonlinear, Dynamic Reference Channels," *IEEE Transaction on Circuit and Systems*, Vol. CAS-32, No. 2, pp. 143-150, February, 1985.

田 先 燾(Sun Do June) 정회원

1967년 5월 9일생
1993년 2월 : 광운대학교 전자통신공학과 졸업(공학사)
1995년 2월 : 광운대학교 대학원 전자통신공학과 졸업(공학석사)
1995년 3월~현재 : 광운대학교 대학원 전자통신공학과 박사과정
※주관심 분야 : 적응 신호 처리, Neural Networks

姜 哲 豪(Chul Ho Kang) 정회원

1952년 2월 5일생
1975년 2월 : 한양대학교 전자공학과 졸업(공학사)
1979년 2월 : 서울대학교 대학원 전자공학과 졸업(공학석사)
1977년~1981년 : 국방과학연구소(연구원)
1988년 2월 : 서울대학교 대학원 전자공학과 졸업(공학박사)
1991년 1월~1992년 1월 : 미국 일리노니 대학 객원교수
1983년~현재 : 광운대학교 전자통신공학과 교수
※주관심 분야 : 적응 신호 처리, 음성 인식, Neural Networks 응용

www.dbpia.co.kr