

논문 2010-47SP-4-15

# 선스펙트럼 쌍의 복원에 의한 잡음억제 기법

( Noise Suppression Method for Restoring Line Spectrum Pair )

최 재 승\*

( Jae Seung Choi )

## 요 약

본 논문에서는 시간지연신경회로망과 주파수영역의 파라미터를 가지는 선스펙트럼 쌍을 사용하여 정규화 방법에 기초한 잡음억제 시스템을 제안한다. 먼저, 시간지연신경회로망은 선형예측분석하여 구해진 잡음이 부가된 음성신호의 선스펙트럼 쌍을 학습시킨다. 그리고 제안한 시스템은 시간지연신경회로망을 학습시킴으로써 배경잡음에 의하여 열화된 잡음이 부가된 음성신호를 강조한다. 따라서 제안한 시간지연신경회로망은 학습에 의하여 잡음이 부가된 선스펙트럼 쌍의 값을 잡음이 부가되기 전의 값으로 복원하여 잡음을 억제한다. 제안한 시스템은 스펙트럼 왜곡율의 평가법을 사용하여 배경잡음에 의하여 열화된 음성신호에 대하여 효과적인 것을 확인한다.

## Abstract

This paper describes a noise suppression system based on a normalization method using a time-delay neural network and line spectrum pair having a parameter of frequency domain. First, a time-delay neural network is trained using line spectrum pair values of noisy speech signals obtained by linear prediction analysis. After trained the time-delay neural network, the proposed system enhances speech signals that are degraded by a background noise. Accordingly, the proposed time-delay neural network restores from the line spectrum pair values of noisy speech signals to the line spectrum pair values of clean speech signals. It is confirmed that this system is effective for speech signals degraded by a background noise, judging from spectral distortion measurement.

**Keywords :** Time-delay neural network, line spectrum, normalization method, spectral distortion measurement.

## I. 서 론

현대의 정보처리라는 그 대부분을 디지털 계산기가 계산 및 처리 등을 담당하고 있다. 이러한 이유 때문에 정보처리가 디지털계산기와 거의 동일하다는 생각이 적용되고 있지만 실제로는 그렇지 않다고 본다. 본래 정보처리라는 것은 생물이 환경 속에서 살아나가기 위하여 운영되고 있으며, 현재의 디지털계산기에 의한 정보처리는 이러한 것 중에서 극히 일부분을 기계에 대체한 것에 불과하다. 정보처리는 인간의 뇌의 구조를 기초로 하여 신경회로망 상의 네트워크 구조에 의한 정보처리

의 가능성을 탐구하려는 시도를 해왔다. 이러한 시도는 뇌와 같은 비선형이고 높은 병렬성을 가진 시스템에의 흥미, 자기학습, 자기조직되어 가는 시스템에의 흥미 등에 지지되어 하나의 주류로서 연구의 흐름을 만들어 왔다. 이러한 연구 중에서 도출되어진 것으로는 신경회로망(Neural Network; NN)<sup>[1]</sup>의 오차역전파학습법(Back-propagation algorithm)<sup>[2]</sup> 또는 이 응용 중의 하나인 시간적 변화를 인식하는 구조를 가진 시간지연신경회로망(Time-Delay Neural Network; TDNN)<sup>[3]</sup> 등이 있다. 본 논문에서는 이러한 TDNN 및 주파수영역의 파라미터를 가지는 선스펙트럼 쌍을 사용하여 잡음이 중첩된 음성신호로부터 음성신호의 부분을 강조하여 잡음을 억제하는 것을 목적으로 연구를 진행한다. 즉, 본 논문에서는 TDNN의 입력인 잡음이 중첩된 선스펙트

\* 정회원, 신라대학교 전자공학과  
(Department of Electronics Engineering,  
Silla University)

접수일자: 2010년1월4일, 수정완료일: 2010년6월15일

럼 쌍의 값을 2종류의 정규화 방법을 사용하여 학습시킨 후에 잡음이 억제되어 복구된 선스펙트럼 쌍의 값을 출력되게 된다. 본 논문에서는 제안한 잡음억제 기법을 평가하기 위하여 스펙트럼 왜곡율 및 신호대잡음비의 평가법을 사용하여 배경잡음에 의하여 열화된 음성신호에 대하여 효과적인 것을 실험으로 확인한다.

## II. 선형예측 분석 및 선스펙트럼 쌍

음성신호의 표본값 사이에는 커다란 상관관계가 있으며 음성의 특징 추출을 위하여 이것을 이용한 예측부호화가 실시되어진다<sup>[4]</sup>. 이러한 예측은 어떤 시점  $n$ 에서의 표본값은 그 이전의 연속한  $p$ 개의 표본값의 선형하중값으로서 근사할 수 있으며, 또한 음성파형은 정상적으로 선형예측계수는 시간적으로 변화하지 않는다고 가정한다. 이러한 가정 하에서, 선형예측오차의 2승 평균값을 최소로 하도록 선형예측계수를 결정한다. 그러나 선형예측 계수는 다음과 같은 2가지의 결점을 가지고 있다고 알려져 있다. 첫 번째는 선형예측계수의 저비트 양자화에 의하여 합성회로의 스펙트럼 특성이 입력음성의 스펙트럼 특성으로부터 크게 차이가 나며 음질의 열화가 크게 된다. 두 번째로는 저비트 양자화에 의한 합성계가 가끔 불안정하게 되며 발진을 일으킨다. 이상의 결점을 개선하기 위하여 제안된 방법 중의 하나가 통계학에 기초를 둔 편자기상관계수(PARCOR 계수)<sup>[5]</sup>이다. 편자기상관 방식은 음성의 생성모델에 기초를 두고 있으며, 시간적으로 직접 내압의 적합성이 부족하고 왜곡이 커지는 결점을 가지고 있다. 따라서 이러한 결점을 보완하기 위하여 단일의 물리적인 의미를 가지고 있으며 주파수 영역의 파라미터인 선스펙트럼 쌍(Line Spectrum Pair; LSP)<sup>[6]</sup>이 고안되었다. 선스펙트럼 쌍의 계수의 주파수는 음성의 스펙트럼의 포르만트 주파수에 대응하는 것이며, 저 비트에 대한 양자화 특성과 저 프레임율에 대한 선형내압성에 뛰어나다.

## III. 실험 데이터 및 시간지연신경회로망

### 1. 실험 데이터 및 실험 평가법

본 실험에서 사용한 음성데이터는 일본인 남성화자에 의한 단어 "aioi(M1)", "hachioji(M2)", "bizen(M3)"과 영어숫자로 구성된 Aurora-2 데이터베이스<sup>[7]</sup>의 2종류를 사용하였으며, 모두 샘플링 주파수는 8 kHz이며

양자화 비트수는 12비트이다. 제안한 시스템은 Aurora-2 데이터베이스의 테스트 셋 A의 음성데이터에서 여성화자 "6751(F1)"를 사용하였다. 그리고 본 실험에 사용한 잡음은 백색잡음(white noise), 자동차잡음(car noise)이다. 여기에서 백색잡음은 컴퓨터에 의해서 작성된 가우스 잡음이며, 자동차잡음은 Aurora2 데이터베이스에 포함된 잡음이다.

본 실험에 있어서의 평가방법으로는 스펙트럼 왜곡율(Spectral Distortion; SD) 및 시간영역의 평가척도인 SNR(Signal-to-Noise Ratio)을 사용한다. SD는 입출력 신호의 대수 스펙트럼의 차를 구한 것으로 식 (1)과 같이 정의한다.

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{W} \int_0^w [S_x(w) - S_y(w)]^2 dw} \quad (1)$$

단,  $S_x(w)$ ,  $S_y(w)$ 는 입출력 신호의 대수 스펙트럼(dB)이며,  $N$ 은 측정구간의 프레임 수,  $W$ 는 신호의 대역폭이다. SNR의 척도를 식 (2)와 같이 정의한다.

$$SNR = 10 \log_{10} \frac{\sum_{i=1}^N X(i)^2}{\sum_{i=1}^N \{X(i) - Y(i)\}^2} \quad (2)$$

여기에서,  $X(i)$ ,  $Y(i)$ 는 각각 입력신호 및 출력신호의 표본값이며  $N$ 은 측정구간의 표본수를 나타낸다.

### 2. TDNN의 구조

음성신호 등을 입력으로 하여 신경회로망에 학습하는 경우, 신호의 시간변화가 중요한 요소가 되는 경우가 있다. 이러한 것을 고려한 신경회로망으로써 TDNN이 고안되었으며 본 논문에서는 TDNN을 사용한다. 본 실험에 사용한 TDNN은 그림 1과 같으며, 입력층은 10유닛에 9프레임, 중간층 제1층은 15유닛에 6프레임, 중간층 제2층은 10유닛에 3프레임, 출력층은 10유닛의 구조를 가진다.

그림 1에서 입력층의 가로축 방향은 시간을 나타내고 있으며, 시간축 방향의 최소 단위를 1프레임이라고 한다. 그림 1에서 입력층의 4프레임 분의 유닛을 중간층 제1층(중간층 1)의 1프레임의 유닛과 결합한다. 또한 중간층 제1층 및 제2층 사이에서는, 중간층 제1층의 4프레임 분이 중간층 제2층의 1프레임 분의 유닛에 결합하고 있다. 이것은 보다 장기적인 시간변화 패턴을 제2

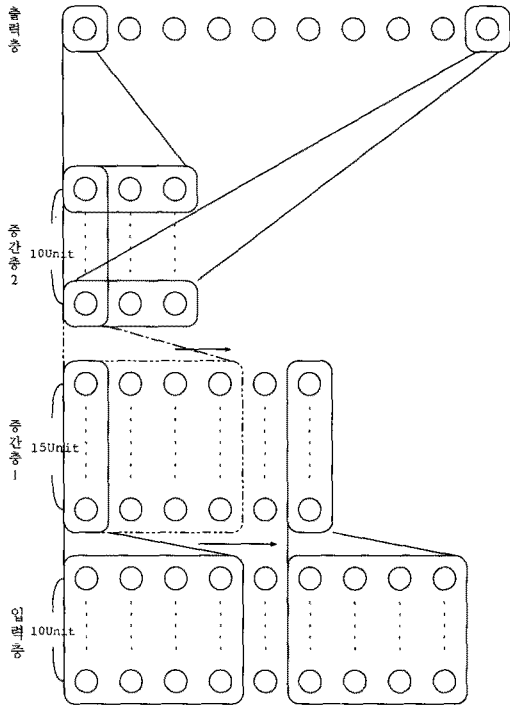


그림 1. 제안한 TDNN 시스템  
Fig. 1. The proposed TDNN system.

층의 유닛에 반영하기 위한 결합 구조이다. 더욱이 출력층은 중간층 제2층의 각 가로 1열이 각각의 출력층의 각 1유닛에 대응하여 접속되어 있다. 이러한 구조를 가지고 있기 때문에 신호의 시간변화 패턴이 반영 가능하도록 한 네트워크가 구축 가능하다.

3. TDNN의 학습법

음성신호라는 주기성이 강한 신호를 TDNN을 사용하여 처리하는 경우에 각 프레임 간에서는 정보의 차이가 적다고 생각되어 진다. 그래서 음성신호의 처리가 프레임의 차이에 의해서 변화되지 않게 하기 위해서, 입력층으로부터 중간층 1의 제1프레임에 접속하는 가중치와 중간층 1의 제2 프레임에 접속하는 가중치의 값은 동일한 값이 되도록 하며, 이 이후도 동일하게 적용한다. 이것을 실현하기 위해서는 다음과 같은 가중치 학습을 실시한다. (1) 중간층 1의 각 프레임에 입력층부터 접속된 가중치를 각각 별도의 신경회로망을 구축한다. 그리고 별도로 먼저 1회의 학습을 오차역전파학습법의 순서에 의해서 학습을 하여 각각의 가중치의 변화 분을 구한다. (2) 중간층 1의 각 프레임에 입력층부터 접속된 가중치의 각각의 변화 분을 각 대응하는 가중치의 변화 분에 대해서 전부 계산하여 평균치를 구한다. 이 후에

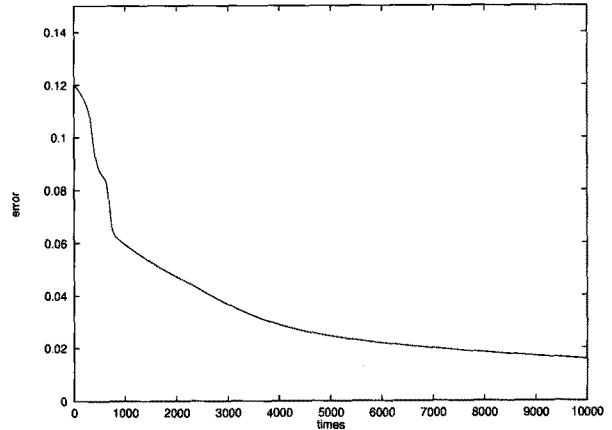


그림 2. 오차의 수축 곡선의 예  
Fig. 2. The example for error curves of training.

이 평균치를 원래의 가중치에 부가한다. 중간층 1부터 상위층에서의 학습도 위에서 기술한 방법을 취한다. 그림 2는 TDNN을 사용한 오차 수축곡선으로서, 백색잡음을 음성신호 “aioi”(SD=10.12)에 부가하여 TDNN의 학습을 실시하였을 때의 오차의 변화를 나타낸다.

IV. 선스펙트럼 쌍의 복원에 의한 잡음억제

본장에서는 TDNN의 학습에 의하여 잡음이 중첩된 선스펙트럼 쌍의 값을 잡음이 혼입하기 전의 값으로 되돌려서 잡음을 억제하는 실험에 대하여 기술한다.

1. 실험의 원리

II장에서 기술한 방법으로 음성신호 표본값을 선형 예측분석하여 선스펙트럼 쌍을 구한다. 단, 본 실험에서는 분석 차수를 10차로 한다. 이렇게 함으로써 원래의 음성신호 표본값은 10개의 선스펙트럼 주파수의 값과 잔차신호에 의해서 완전히 복구가 가능하다. 분석대상의 음성신호 표본값에 잡음이 중첩된 경우, 물론 잡음은 선스펙트럼 쌍의 값 및 잔차신호의 양쪽에 영향을 미치지만, 본 실험에서는 잔차신호 쪽에는 조작을 가하지 않고 선스펙트럼 쌍의 값만을 잡음이 중첩되기 전의 값에 근접하게 하는 것을 목적으로 하고 있다. 따라서 이를 위하여, TDNN의 학습 시에는 입력신호로써 잡음이 중첩된 음성신호 표본값으로부터 구한 선스펙트럼 쌍의 값 및 교사신호로써 잡음이 없는 음성신호 표본값으로부터 구한 선스펙트럼 쌍의 값을 사용한다.

본 학습을 실시한 후의 TDNN에 잡음이 중첩된 입력음성신호로부터 구한 선스펙트럼 쌍의 값을 부여하여

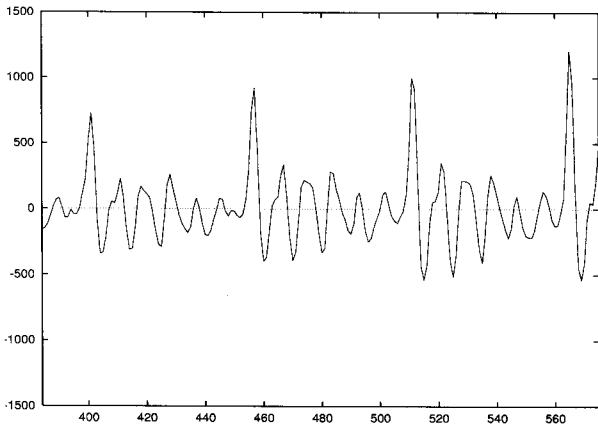


그림 3. 음성신호  
Fig. 3. Speech signal.

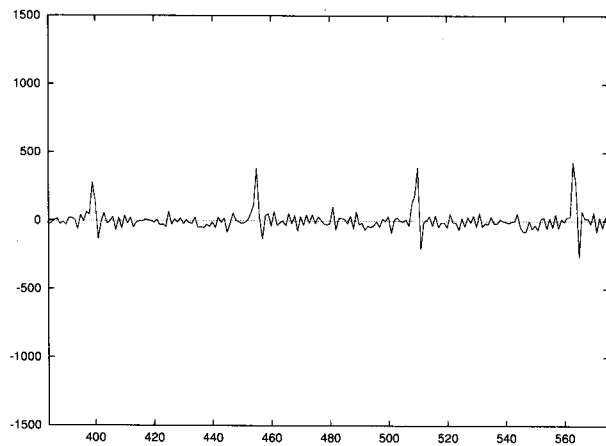


그림 4. 잔차신호  
Fig. 4. Residual signal.

출력을 구한다. 이 출력으로 구한 선스펙트럼 쌍의 값과 분석 시에 구한 잔차신호로부터 음성신호를 합성하여 잡음이 중첩되지 않은 음성신호와 비교하여 효과를 조사한다. 그림 3과 그림 4는 본 실험에서 구한 “aioi” 3프레임에 대한 음성 및 잔차신호를 각각 나타낸다.

2. TDNN의 정규화 방법 및 학습방법

본 실험에서는 TDNN에의 입력인 선스펙트럼 쌍의 값을 0부터 1까지의 범위로 정규화하며, 다음의 2가지 데이터 규화 방법을 사용한다. (A) 선스펙트럼 쌍의 값은 0부터  $\pi$ 의 범위 내이기 때문에, 값을  $\pi$ 로 나눈 값을 입력으로 한다. (B) 선스펙트럼 쌍의 값은 10차로 분석하였기 때문에 각 프레임에 대해서 10개이다. 각각 프레임에 대하여 평균값을 구하여 평균값과의 차에 0.5를 부가한 값을 입력으로 한다. 이와 같이 정규화한 각 프레임의 10개의 선스펙트럼 쌍의 값은 입력층의 각 유닛에 전달되어 다음과 같은 방법으로 학습이 실시된다.

(1) 각각 유닛의 가중치 및 문턱값을 임의적으로 0부터 1까지의 범위 내에서 초기화하여 다음 식의 학습계수  $\epsilon$  및 가속도계수  $\alpha$ 를 지정한다. 본 실험에서는  $\epsilon = \alpha = 0.3$ 으로 하였다.

$$\Delta w_{ij}^{k-1k}(t+1) = -\epsilon d_j^k o_i^{k-1} + \alpha \Delta w_{ij}^{k-1k}(t) \quad (3)$$

여기에서  $w_{ji}^{k-1k}$ 는  $k-1$ 층의 제 $i$ 유닛으로부터  $k$ 층의 제 $j$ 유닛에의 결합의 가중치를,  $d_j^k$ 는  $k$ 층의 제 $j$ 유닛의 회망출력을,  $o_i^k$ 는  $k$ 층의 제 $i$ 유닛의 출력을 나타낸다. (2) 입력층에 제1프레임으로부터 제9프레임의 잡음이 중첩된 음성신호의 표본값으로부터 구한 선스펙트럼 쌍의 값이 입력신호로써 부여된다. (3) 제3프레임의 잡음이 중첩되지 않은 음성신호 표본값으로부터 구한 선스펙트럼 쌍의 값이 교사신호로써 부여된다. (4) III.3절의 TDNN 학습방법으로 학습을 실시한다. (5) 입력층에 제2프레임으로부터 제10프레임의 선스펙트럼 쌍의 값을 입력신호로써 부여된다. (6) 제4프레임의 선스펙트럼 쌍의 값을 교사신호로써 부여된다. (7) 이하, 동등한 조작을 반복하여 모든 프레임에 대해서 학습이 종료된 시점에서 (2)에 되돌아간다. 단, 되돌아가는 횟수는 학습횟수로 포함된다. 따라서 학습 후의 네트워크에 잡음이 중첩된 선스펙트럼 쌍을 입력으로 부여하면, 출력으로써 선스펙트럼 쌍의 값이 구해지게 된다.

V. 실험 결과

본 장에서는 TDNN을 사용하여 잡음을 억제하여 음성을 강조하는 실험결과에 대해서 기술한다. 먼저, 학습 후의 입력으로써 부여하는 신호를 학습 시에 입력으로써 사용한 신호와 동일한 신호를 사용하여 출력을 구한

표 1. 학습 시의 입력데이터와 학습 후의 입력데이터가 동일한 경우의 결과  
Table 1. The results when the training and the estimating data were same.

정규화	입력 데이터	학습 시	학습 후	
		입력 SD(dB)	출력 SD(dB)	개선 SD(dB)
(A)	M1	10.12	6.84	3.28
	M2	12.20	8.47	3.73
	M3	13.46	7.44	6.02
(B)	M1	10.12	5.97	4.15
	M2	12.20	7.65	4.55
	M3	13.46	6.79	6.67

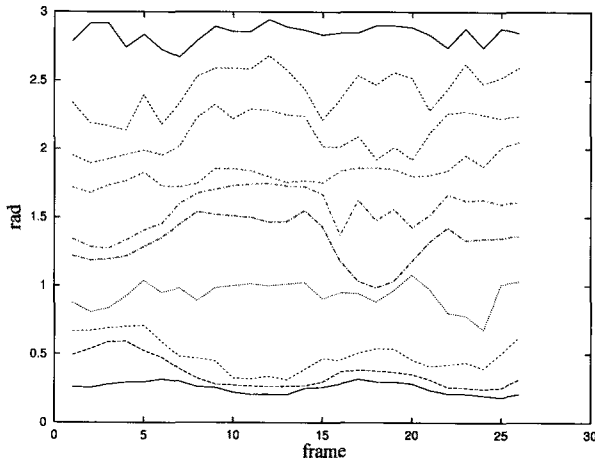


그림 5. 음성신호 "aioi"의 LSP 값  
Fig. 5. The values of LSP for speech signal "aioi".

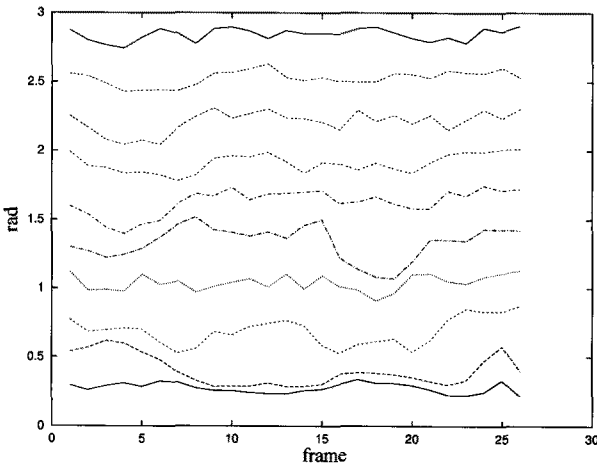


그림 6. SD=10.39의 음성신호 "aioi"의 LSP 값  
Fig. 6. The values of LSP for noisy speech signal "aioi" in case of SD=10.39.

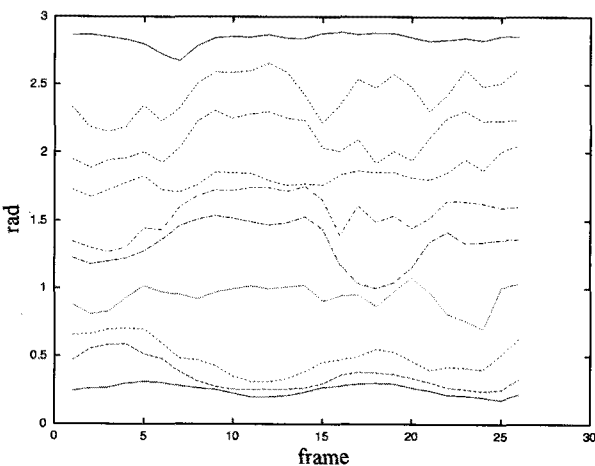


그림 7. 학습 후의 "aioi"의 LSP 값  
Fig. 7. The values of LSP for noisy speech signal "aioi" after training the TDNN.

경우에 대하여 실험을 실시하였다. 그 결과를 표 1에 나타낸다. 표 중의 'SD=10.39 aioi'는 원래의 음성신호 샘플 "aioi"에, 어떤 강도로 백색잡음을 부가한 결과, 이 SD의 값이 10.39(dB)로 된 음성신호의 샘플을 나타낸다. 표 안의 정규화방법 (A) 및 (B)는 IV.2 절에서 나타낸 정규화방법에 각각 대응한다. 또한 학습 시에 입력 데이터에 대한 교사데이터는 잡음이 중첩되지 않은 음성신호 샘플이다. 그리고 이후의 실험에서는 TDNN의 학습횟수를 모두 10,000회하여 실험을 실시하였다.

표 1의 결과로부터, 학습에 사용한 음성신호를 학습 후의 네트워크에 입력하여 출력을 구한 경우, 정규화방법 (A), (B)에 대해서 SD를 경감하는 효과가 있는 것을 알 수 있다. 학습에 의한 선스펙트럼 쌍의 값의 변화의 예를 그림 5, 6, 7에 나타낸다. 그림 5는 음성신호 "aioi"로부터 구한 선스펙트럼 쌍의 값을, 그림 6은 이 파형에 대한 백색잡음을 부가한 SD=10.39의 음성신호 "aioi"로부터 구한 선스펙트럼 쌍의 값을, 그림 7은 TDNN을 학습시킨 후의 "aioi"의 선 스펙트럼 쌍의 값을 각각 그래프로써 나타낸다. 여기에서, 1프레임은 192 표본, 표본 주파수는 8 kHz이다. 그림 6을 보면, 잡음이 중첩됨에 따라서 선스펙트럼 쌍의 값이 영향을 받는 것을 알 수 있다. 그림 7에서 알 수 있듯이, TDNN을 학습시킴으로써 학습 후의 "aioi"의 선스펙트럼 쌍의 값이 그림 5의 잡음을 중첩하기 전의 선스펙트럼 쌍의 값에 가깝게 되어가는 것을 볼 수 있다. 따라서 본 실험은 학

표 2. 학습 시의 입력데이터에 강도가 다른 잡음을 부가한 경우의 결과

Table 2. The results when training and estimating noisy intensity were different.

정규화	입력 데이터	학습 시 입력 SD(dB)	학습 후		
			SD (dB)		
			입력 SD	출력 SD	개선 SD
(A)	M1	10.12	12.24	7.71	4.53
			14.43	8.68	5.75
	M2	12.20	14.22	9.12	5.10
			15.94	10.24	5.70
	M3	13.46	16.30	10.35	5.95
			18.37	8.56	9.81
(B)	M1	10.12	12.24	7.64	4.60
			14.43	10.72	3.71
	M2	12.20	14.22	9.10	5.12
			15.94	10.81	5.13
	M3	13.46	16.30	9.29	7.01
			18.37	9.21	9.16

표 3. 학습 시의 입력데이터와는 다른 음성데이터를 사용한 경우의 결과

Table 3. The results when the training and the estimating speech data were different.

정규화	학습 시		학습 후			
	입력 데이터	입력 SD(dB)	입력 데이터	SD (dB)		
				입력 SD	출력 SD	개선 SD
(A)	M1	10.12	M2	14.22	10.71	3.51
			M2	16.10	11.10	5.00
			M3	16.30	9.87	6.43
			M3	18.55	11.06	7.49
			F1	16.39	10.43	5.96
			F1	18.65	12.62	6.03
(B)	M1	10.12	M2	14.22	12.14	2.08
			M2	16.10	13.88	2.22
			M3	16.30	11.97	4.33
			M3	18.55	12.39	6.16
			F1	16.39	12.46	3.93
			F1	18.65	12.87	5.78

표 4. 학습 시의 입력데이터와는 다른 음성데이터 및 잡음을 사용한 경우의 결과

Table 4. The results when the training and the estimating speech data and noise were different.

정규화	학습 시		학습 후			
	입력 데이터	입력 SD(dB)	입력 데이터	SNR (dB)		
				입력 SNR	출력 SNR	개선 SNR
(A)	M1	10.12	F1	14.57	17.09	2.52
			F1	9.41	12.00	2.59
(B)	M1	10.12	F1	14.57	16.30	1.73
			F1	9.41	11.88	2.47

습 후의 TDNN이 효과가 있는 것을 보여준다.

표 2는 백색잡음의 강도를 변경하여도 효과가 얻어지는가를 조사하기 위하여, 학습 시 및 학습 후의 입력 데이터는 동일하지만 학습 시에 입력으로써 사용한 데이터와는 잡음강도가 다른 음성신호를 학습 후의 네트워크에 입력하여 출력을 구한 결과를 나타낸다. 표 2의 결과로부터, 잡음강도를 변경하여도 SD를 경감하는 효과가 구해지는 것을 알 수 있다. 표 3은 TDNN의 일반성을 조사하기 위하여, 음성데이터에 백색잡음을 부가한 것을 학습 후의 네트워크에 입력하여 출력을 구한 결과이며, 학습 시에 사용한 입력데이터와 다른 음성데이터에 대해서도 SD 값을 경감하는 효과가 구해진다.

표 4는 학습 시에 입력으로써 사용한 음성데이터와

는 다른 음성데이터를 사용한 경우에 대하여, 특히 자동차잡음을 학습 후에 부가하였을 때의 SNR에 의한 잡음제거에 대한 실험이다. 표 4의 결과로부터, 학습 시에 사용한 입력데이터와 다른 음성 및 잡음을 사용하였을 경우에 SNR이 최대 2.59 dB 정도 향상되는 것을 확인할 수 있었다.

지금까지 실시한 실험의 결과로부터, 선스펙트럼 쌍의 값을 TDNN에 의하여 복원함으로써 잡음이 많이 중첩된 음성에 대해서도 SD 및 SNR의 값을 경감하는 것을 알 수 있었다. 실제로 TDNN의 출력으로써 구해진 선스펙트럼 쌍의 값으로부터 복원한 음성을 귀로 청취해보았을 때 상당히 백색잡음 및 자동차잡음이 경감된 것을 알 수 있었다.

## VI. 결 론

본 논문에서는 잡음억제 및 음성강조의 방법으로써 TDNN에 의한 수법을 제안하여 실험을 실시하였다. 특히 선 스펙트럼 쌍의 복원의 수법을 사용하여 잡음이 많이 부가된 음성데이터를 사용한 경우의 SD 및 SNR의 개선량이 향상되었으며, 학습에 걸린 시간이 적기 때문에 효율이 좋다고 생각되어 진다.

이상의 결과 및 각 실험에 대한 고찰을 기본으로 하여 향후의 과제로는 다음과 같다. (1) 실험에 사용하는 음성데이터를 다양하게 할 필요가 있다. (2) 선형예측계수의 복원에 의한 수법에 있어서 잔차신호에도 어떤 조작을 부가한다. (3) 음질 평가 방법으로 MOS 혹은 PESQ 등으로 평가할 필요가 있다. 이상으로 선 스펙트럼 쌍의 복원의 실험을 통하여 시간변화를 고려한 TDNN을 사용하여 본 수법이 SD 및 SNR 값이 개선가능하다는 결과로부터 본 논문이 음성강화에 도움이 될 것으로 생각된다.

## 참 고 문 헌

- [1] K. Daqrouq, I. N. Abu-Isbeih, M. Alfauri, "Speech signal enhancement using neural network and wavelet transform", 2009. 6th International Multi-Conference on Systems, Signals and Devices, pp. 1-6, 2009.
- [2] Ooyen A. V. and Nienhuis B. "Improving the convergence of the back-propagation algorithm," Neural Networks 5, 3, pp. 465-471, 1992.

- [3] J. B. Hampshire, A.H. Waibel, "A novel objective function for improved phoneme recognition using time delay neural networks", IEEE Transactions on Neural Networks, Vol. 1, No. 2, pp. 216-228, 1990.
- [4] 최재승, "음성신호의 선형예측계수에 의한 잡음량의 인식", 대한전자공학회 논문지 제46권 SP편 제2호, pp. 120-126, 2009년.
- [5] 이용준, 김형래, "한국어 합성을 위한 PARCOR 계수의 부호화와 압축에 관한 연구", 대한전자공학회 추계종합학술대회 논문집 제8권 2호, pp. 155-157, 1985.
- [6] 강정원, 정대권, 김성수, 정재호, 우홍채, 이인성, "음성신호의 Line Spectrum Pair (LSP) 주파수의 2단 벡터양자화", 대한전자공학회 신호처리합동학술대회 논문집 제7권 1호, pp. 288-291, 1994.
- [7] H. Hirsch and D. Pearce, "The AURORA experimental framework for the performance evaluations of speech recognition systems under noisy conditions", in Proc. ISCA ITRW ASR2000 on Automatic Speech Recognition: Challenges for the Next Millennium, Paris, France, 2000.

---

 저 자 소 개
 

---



최 재 승(정회원)  
 1989년 조선대 전자공학과 학사  
 1995년 오사카시립대학 정보통신  
 공학과 석사  
 1999년 오사카시립대학 정보통신  
 공학과 박사  
 2000년~2001년 일본 마쯔시타  
 전기산업주식회사  
 AVC사 연구원

2002년~2007년 경북대디지털기술연구소  
 책임연구원

2007년~현재 신라대학교 전자공학과 교수  
 <주관심분야: 음성신호처리, 음성인식, 신경회로  
 망, 청각모델 등>