

논문 2006-43SP-5-10

잡음 환경에서 압신을 이용한 인공 와우 환자의 언어 인지 향상 시뮬레이션 연구

(A simulation study of speech perception enhancement for cochlear
implant patients using companding in noisy environment)

이 영 우*, 지 윤 상*, 이 종 실**, 김 인 영**, 김 선 일**, 홍 성 화***, 이 상 민****

(Young Woo Lee, Yoon Sang Ji, Jong Shill Lee, In Young Kim, Sun I. Kim,
Sung Hwa Hong, and Sang min Lee)

요 약

본 연구에서 인공 와우 환자의 잡음 상황에서 음성 신호 강조와 잡음 제거를 위한 전 처리로서 companding strategy를 적용하고 이를 평가하였다. Companding은 인간의 청각 특성인 two tone suppression에 기반하며 이는 음성 스펙트럼 피크를 강화하고 배경 잡음을 감소시킨다. 하지만 companding은 잡음 제거와 스펙트럼 피크의 강화에 효과적인 반면, 제한된 채널의 수와 비선형 블록으로 인한 음성 정보 손실의 교환 특성을 가진다. 따라서 본 연구에서는 잡음 제거와 음성 정보 손실의 정도가 상대적인 두 companding 구조를 설계하여 개인마다 잡음 상황에서 언어 인지 특성차이에 따른 적절한 필터 बैं크를 도출하였으며, 낮은 신호 대 잡음 비 환경에서 인공 와우 환자의 언어 인지 향상을 위한 방법을 제시하였다. 제안된 알고리즘은 잡음 밴드 시뮬레이션을 이용하여 정상인 5명에게 평가되었다. 모든 피실험자에게서 효과적인 언어 인지의 향상이 관측되었고, 각 피실험자가 선호하는 필터 बैं크는 다르게 나타났다.

Abstract

In this study, we evaluated the performance of a companding strategy as a preprocessing for speech enhancement and noise reduction. The proposed algorithm is based on two tone suppression that is human's hearing characteristics. This algorithm enhances spectral peak of speech signal and reduces background noise, however it has tradeoff characteristics between speech distortion and noise reduction due to limited channel number and nonlinear block. Therefore, we designed two different companding structures that have relative characteristics of noise reduction and speech distortion and found suitable companding structures by difference of individual speech perception ability in noise environment. Thus we proposed speech perception enhancement of cochlear implant user in noise environment with low SNR. The performance of the proposed algorithm was evaluated through 5 normal hearing listeners using noise band simulation. Improvement of speech perception was observed for all subjects and each subject preferred the different type of companding structure.

Keywords: companding, noise reduction, cochlear implant, speech perception, noise band simulation

* 학생회원, ** 정회원, 한양대학교 의공학교실

(Dept. of Biomedical Engineering, Hanyang University)

*** 정회원, 성균관대학교 의과대학 이비인후과

(Dept. of Otolaryngology-Head and Neck Surgery, Sungkyunkwan University of Medicine)

**** 정회원, 인하대학교 전자전기공학부

(School of Electrical Engineering, Inha University)

※ 이 논문은 2005년도 전북대학교 지원 연구비와 보건복지부 보건의료기술개발사업 지원 (고유번호 : 02-PJ3-PG6-EV10-0001)에 의하여 이루어졌음.

접수일자: 2006년1월4일, 수정완료일: 2006년8월7일

I. 서 론

와우각내의 청각유모세포의 손상으로 인한 청각 장애의 경우 외부로부터의 음향을 증폭시켜 전달하더라도 청각세포의 변형이 일어나지 않아 청각신경에 활동전위가 발생하지 않으므로 음향에 관련된 정보가 뇌로 전달되지 않는다. 이런 경우 기존의 보청기에 의한 방법으로는 청각 기능의 보정이 불가능하며 청신경을 전기적으로 자극함으로써 소리의 인지를 가능하게 하는 인공 와우 장치를 적용해야 한다. 인공 와우 장치는 전극과 유도 코일로 이루어진 이식기와 어음처리기, 전송기로 이루어진 체외기로 구성되어 있다. 마이크로폰을 통해서 얻어진 음성 전기신호는 음성 신호처리를 통해서 그 성분이 분석되고, 특징들이 추출 변형되어서 와우내의 고실계에 위치한 수용기에 무선 전송방식으로 전달된다. 이 신호가 다시 복호화되어 와우내로 삽입된 전극을 통해 전기 신호가 각 주파수 성분에 따라서 정해진 와우 신경의 나선 신경절세포를 자극하게 되면 뇌로 전달된 신경 전달 신호를 소리로 인지하게 된다^[1]. 전극 삽입 수술을 받은 환자의 언어인지의 성능 개선을 위해 이식기의 개발과 음성 신호로부터 전극에 가해지는 전기 자극의 크기를 추출해내는 신호처리부의 신호처리 방법의 개선이 지속적으로 이루어지고 있다. 현재 인공 와우장치에서 가장 많이 사용되는 신호처리 방식인 CIS (continuous interleaved sampling) 방식에서는 선형대역필터뱅크를 통과한 입력신호가 포락선 검출과정을 거친 후 적절한 자극 반복률을 가진 쌍 위상 펄스열의 크기를 변조하게 된다. 잡음이 없는 상황에서는 CIS 방식은 인공와우장치 이식 환자들을 대상으로 한 실험에서 우수한 성능을 보이지만, 잡음 하에서는 언어 인지에 불편을 호소한다. 대부분의 인공 와우 시스템의 strategy는 각 필터 밴드의 에너지나 크기에 기반 하여 자극 채널이 결정되므로 낮은 신호 대 잡음 비 환경에서는 인공 와우 환자의 언어 인지력이 떨어지게 된다. 다양한 잡음 제거 알고리즘이 인공 와우 환자를 위해 제안되어 왔으며, 이런 알고리즘의 대부분은 2개 또는 그 이상의 마이크로폰을 이용한다는 가정에 기반 한다. Van Hoesel과 Clark는 4명의 Nucleus-22 환자에게 2개의 마이크로폰으로부터 들어오는 신호를 환자의 90°로부터 오는 잡음을 제거하기 위해 adaptive beam-forming 기술을 적용했다^[2]. 그러나 인공 와우에 두 번째 마이크로폰을 추가하는 것은 헤드셋을 착용하지 않고서는 어려우며, 양귀 인공 와우가 그 수단으로 제공

되지만 그 효과는 아직 연구되어 지고 있다. 대체 방안으로, 한쪽에 멀티 마이크로폰 기술이 사용될 수 있고, 그런 기술들은 현재 상업적으로 사용되고 있다. 그러나 일반적으로 멀티 마이크로폰 입력에 기반한 알고리즘보다 단일 마이크로폰 잡음 제거 알고리즘이 외관적인 문제로 인해 더 선호되어, 여러 가지 단일 마이크로폰 잡음 제거 방법이 인공 와우에 적용되어 왔다. 그 중 일부는 특징 추출 방법 (F0/F1/F2, MPEAK strategy)에 기반한 인공 와우 처리 장치에 적용되었으며, 이 방식으로 Weiss는 잡음 제거 알고리즘으로 전 처리된 신호는 포먼트 추출에 오차가 감소한다는 것을 검증했다^[3]. 그러나 최근의 음성 프로세서는 특징 추출 방법에 기반하는 게 아니라 보코더 타입의 방법에 기반 한다. Yang과 Fu는 보코더 방식의 인공 와우 프로세서를 이용해 스펙트럼 차분 알고리즘을 평가했으며, 스펙트럼 차분 알고리즘을 적용시킨 모든 피실험자에게서 효과적인 문장 인지의 향상이 관측되었다^[4]. 이와 같이, 다양한 연구가 잡음 제거 알고리즘에 의한 잡음 음성 신호를 전 처리하고, 인공 와우 환자에게 강조된 신호를 전달함에 의한 이득을 연구하기 위해 수행되어져 왔다. 본 연구에서 우리는 잡음 상황에서 신호 강조와 잡음 제거를 위한 전 처리로서 companding strategy를 적용하고 이를 평가하였다. 인간의 청각 특성 중 하나인 two tone suppression은 인간의 귀가 잡음 상황에서 뛰어난 성능을 보이는 이유 중의 하나이다. Two tone suppression이란 하나의 tone이 인접한 주파수 tone과 함께 존재할 때, 그 tone에 따른 고유 반응이 감소하는 특성이다. Turricchia 등은 이 특성을 모델링한 압신(companding, 이하 companding)을 제시하였으며, companding은 와우 동작 특성을 반영하기 위해 압축(compression)과 신장(expansion)의 비선형적인 기능을 신호처리 과정에 적용한다. 이는 스펙트럼 피크를 강화하고 배경 잡음을 감소시킨다^[5]. 하지만 companding strategy는 잡음 제거와 스펙트럼 피크의 강화에 효과적인 반면, 제한된 채널의 수와 비선형 블록으로 인한 음성 정보 손실의 특성을 가진다. 따라서 본 연구에서는 잡음 제거와 음성 정보 손실의 정도가 상대적인 두 companding 구조를 설계하여 개인마다 잡음 상황에서 언어 인지 능력에 적절한 필터 뱅크를 도출하여 낮은 신호 대 잡음 비 환경에서 인공 와우 환자의 언어 인지 향상을 위한 방법을 제시하였다. 평가 방법은 잡음 밴드 시뮬레이션을 이용한 정상인을 통한 임상 실험이 시행되었다.

II. 본 론

1. Compinging strategy 설계

Compinging의 구조는 그림 1과 같이 전처리 필터, 압축 블록, 후처리 필터, 팽창 블록으로 구성되어 있다. 각 채널의 전처리 필터뱅크와 후처리 필터뱅크는 같은 중심 주파수를 가진다. 인공 와우 시스템을 모델링하기 위해 총 22개의 채널을 구성하였으며, 각 채널의 중심 주파수는 Liberman이 제시한 고양이에서의 주파수 맵을 이용하여 인간의 와우에 적절한 주파수 맵으로 변환시킨 후 채널수에 따라 기저부로부터의 거리를 주파수로 변환시켰다^[6]. $d(i)$ 는 기저막의 기저부로부터 해당 채널의 백분율 거리이며, $f(i)$ 는 거리에 따라 변환된 주파수이다. 이렇게 계산된 주파수는 고양이의 최대 주파수 CF_{max} 와 인간의 최대 주파수 HF_{max} 에 의해 와우의 중심 주파수가 결정된다.

$$f(i) = 456 * 10^{0.021 * d(i)} - 364.8 \quad (1)$$

$$Fc(i) = f(i) * HF_{max} / CF_{max}$$

그림 2는 compinging의 단일 채널의 상세 구조를 보여준다. 전처리 필터 (F)는 광 대역 필터로 구성되었으며, 후처리 필터 (G)는 협 대역 필터로 구성되었다. 압축과 팽창은 포락선 검출기(ED), 비선형 블록, 곱셈기를 가지고 수행된다. 포락선 검출의 시상수는 각 채널의 중심 주파수에 의해 조절된다. 해당 채널의 광 대역 필터를 통과한 신호 성분들은 포락선 검출기를 통하여, 강한 신호 성분에 기반 하여 결정된 압축비에 의해 압축되게 된다. 압축된 신호 성분들은 해당 채널의 협 대역 필터를 통과하게 되고, 대역외의 신호 성분이 제거된 협 대역 내의 신호들 중 강한 신호 성분에 기반 하여 결정된 팽창비에 의해 팽창되게 된다. 즉 강한 스펙트럼 에너지를 가지는 채널은 강조시키고 약한 채널은 감소시킴으로써, two-tone suppression의 역할을 하여 해당 채널 주위의 잡음을 감소시키고 스펙트럼 피크가 강화된다.

식 (2), (3)와 같은 통과 함수를 가지는 대역통과 필

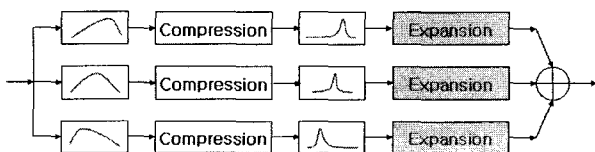


그림 1. Compinging의 블록도
Fig. 1. Block diagram of compinging.

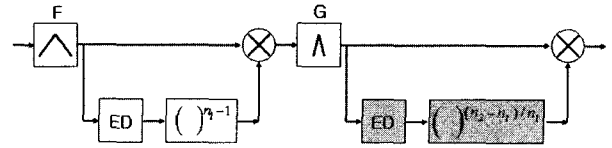


그림 2. Compinging 구조의 단일 채널
Fig. 2. A single channel of compinging structure.

터를 MATLAB에서 쌍 일차 변환 방법을 이용하여 디지털 영역에서 광 대역 필터(F)와 협 대역 필터(G)를 설계하였다. q_1 과 q_2 는 광, 협대역 필터의 밴드폭을 결정하는 계수이며, τ_i 는 각 채널의 중심주파수에 따른 포락선 검출기의 시상수이다.

$$F_i(s) = \left(\frac{2 \left(\frac{\tau_i}{q_1} \right) s}{\tau_i^2 s^2 + 2 \left(\frac{\tau_i}{q_1} \right) s + 1} \right)^2 \quad (2)$$

$$G_i(s) = \left(\frac{2 \left(\frac{\tau_i}{q_2} \right) s}{\tau_i^2 s^2 + 2 \left(\frac{\tau_i}{q_2} \right) s + 1} \right)^2 \quad (3)$$

본 연구에서는 q_1 은 1.5, q_2 는 4.5로 설정하였으며, 각 채널의 입력과 출력이 선형적으로 되게 하기 위해 n_1 은 0.3, n_2 는 1로 설정하여 압축 단계에서 신호가 압축되고, 팽창 단계에서 압축을 복원하게 하였다.

잡음 제거의 평가를 위해 신호 대 잡음비(SNR, 이하 SNR) 측정 알고리즘을 설계하였다. 하지만, 보통의 SNR로는 난청인을 대상으로 대역 전체를 자극하는 인공와우의 음성 강조를 평가하기 어렵다. 그러므로 고주파 잡음대 저주파 잡음 음성을 비교하는 정규화 잡음 음성 대 잡음 비율(XNR, 이후 XNR)이 필요하다. 화자에 따라 다르게 나타나기는 하지만, Yamauchi는 일반적으로 높은 주파수 영역은 잡음 스펙트럼 성분이 음성 스펙트럼 성분보다 많다는 성질을 이용하여 식 (4)과 같은 잡음 측정 방식을 제안하였다^[7].

$$\begin{aligned} NL_{Xhigh} &= \int_{10k}^{Fs/2} |X(\lambda, f)| df \quad (4) \\ &= \int_{10k}^{Fs/2} \{|S(\lambda, f)_{RIGHT}| + |N(\lambda, f)_{RIGHT}|\} df \\ &\approx \int_{10k}^{Fs/2} |N(\lambda, f)| df \\ &= NL_{Nhigh} \end{aligned}$$

NL_{Xhigh} 는 높은 주파수 영역에 잡음 음성 수준이고, NL_{Nhigh} 는 높은 주파수 영역에 잡음 수준이다. Com-panding은 강한 스펙트럼 에너지를 가지는 채널은 강조 시키고 약한 채널은 감소시키므로 음성이 거의 존재하지 않는 높은 주파수 영역에서는 잡음이 많이 감소하고 음성 정보가 많이 존재하는 낮은 주파수 영역에서는 음성 신호가 많이 강조되므로 제안된 방식을 이용하여 XNR을 알아낼 수 있다. 잡음 음성의 전력은 0~8kHz의 범위에서 측정하였고, 잡음 신호의 전력은 10~22.05kHz에서 측정하여 식 (5)처럼 그 비율을 XNR로 가정하였다.

$$XNR = \frac{NL_{Xlow}}{NL_{Nhigh}}, NL_{Xlow} = \int_0^{8k} |X(\lambda, f)| df \quad (5)$$

이 XNR을 정규화 하여, 잡음을 인가한 SNR과 설계된 방법에 의해 측정된 XNR이 일치하게 하기 위해 K-SPIN에서 추출한 단음절 3개를 이용하여 보정식 (6)을 구성하였다. XNR_N 은 정규화 된 XNR이다.

$$XNR_N = c - b * \ln\left(\frac{a}{(XNR - d)} - 1\right) [dB] \quad (6)$$

여기에서, a 는 36.34, c 는 8.6149, b 는 8.8927, d 는 -3.4795이다. 설계된 정규화 XNR 측정 전체 블록도는 그림 (3)과 같다.

단음절 10개에 20dB부터 -18dB의 SNR의 잡음을 적용한 후, 설계된 방법에 의해 XNR_N 을 측정하였으며 그림 (4)와 같이 알고리즘의 성능을 검증하였다.

2. 필터 बैं크 설계

모든 청각 보조 장치들은 환자 개개인에 적합한 피팅 과정을 거친다. 이는 각 환자들의 언어 인지 특성, 잔존 청력 상태 등 다양한 원인으로, 개인마다 잡음 상황에서 언어 인지 능력에 차이가 생긴다. 그러므로 같은 신호대 잡음비라도 개인마다 선호하는 필터 बैं크 타입은 다를 것이다. 필터 बैं크 파라미터 변경에 따라 잡음 제거와 음성 정보 손실은 교환 특성을 지닌다. 식 (2)과 (3)에서 q_1 과 q_2 에 의한 대역 필터의 형태는 two-tone suppression의 주파수 윤곽을 결정한다. q_1 이 작아질수록 광 대역 필터의 밴드 폭은 커지게 되고, q_2 가 커질수록 협 대역 필터의 밴드 폭은 감소되어 더욱 큰 two-tone suppression이 발생되어 해당 채널 주위의 잡음 감소가 커지는 반면 음성 정보에 왜곡 또한 커진다. q_1 값이 커지고, q_2 값이 작아지는 경우 two-tone suppression의 효과가 미약하여 그 반대로 잡음 제거는 미비하

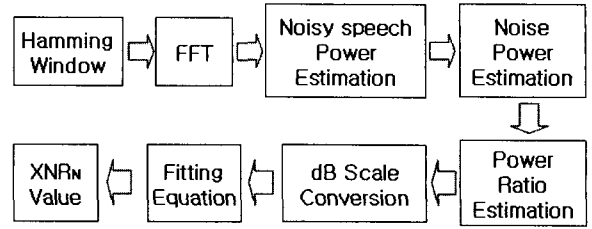


그림 3. 신호 대 잡음 비 측정 전체 블록도
Fig. 3. Total block diagram of XNR_N estimation.

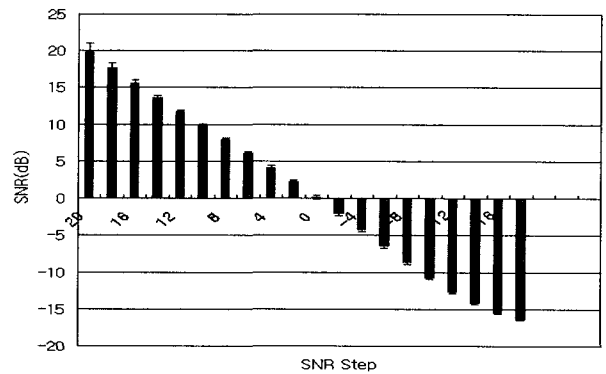


그림 4. Fitting된 XNR_N 의 평균, 표준편차
Fig. 4. Average and deviation of Fitting XNR_N .

지만, 음성 정보의 왜곡이 줄어든다. 본 연구에서는 q_1 과 q_2 를 조정하여 이런 특성들이 상대적인 두 종류의 필터 बैं크를 설계하고 이를 적용하여 개인 청력 특성에 적합한 필터 बैं크를 도출하고자 하였다. 잡음 제거의 평가는 앞서 사용한 정규화 XNR 측정 알고리즘을 사용하였고, 음성 정보 손실의 평가는 음성 명료도와 관계가 깊은 식 (7)의 스펙트럴 왜곡(Spectral distortion, SD)을 사용했다^[8]. W 는 신호의 대역폭, $S_s(f)$ 및 $S_y(f)$ 는 입력신호 $s(t)$ 와 출력신호 $y(t)$ 의 대수 스펙트럼이며, 식 (8), (9)과 같이 정의한다. $S(f)$ 와 $Y(f)$ 는 각각 주파수 f 에서의 입력 신호와 출력 신호의 스펙트럼이다.

$$SD = \sqrt{\frac{1}{W} \int_0^W |S_s(f) - S_y(f)|^2 df} (dB) \quad (7)$$

$$S_s(f) = 10 \log_{10} |S(f)|^2 \quad (8)$$

$$S_y(f) = 10 \log_{10} |Y(f)|^2 \quad (9)$$

3. 인공 와우 음성 변조

인공 와우 환자의 음성 변조는 인공 와우 장치에 의해 중앙 신경계에 전달되는 스펙트럼의 저하와 유사하

게 정상 청각 인에게 전달하기 위해 제안되었다. 최초에 Shannon^[9]과 Dorman^[10]이 두 가지 접근을 하였는데 Dorman은 입력 신호를 N 밴드로 대역통과 필터를 한 후 각 밴드의 포락선에 의해 필터의 중심 주파수를 가지는 N개의 사인곡선을 변조하였다. Shannon은 사인곡선 대신 잡음 밴드로 변조하였다. 두 방식을 사용한 정상 청각 인에서의 언어 인지 능력은 유사하지만 최근에는 잡음 밴드 변조 방식이 널리 채택되고 있으며, 인공 와우 음성 변조는 일반적으로 잡음 밴드 변조로 언급되고 있다. 잡음 밴드 시뮬레이션은 처리방식의 유사성 때문에 CIS strategy를 시뮬레이션하기 위해 사용된다. 최근에 Michael 또한 인공 와우 환자의 중심주파수 분별력에 관한 연구를 위해 정상 청각 인을 대상으로 잡음 밴드 변조를 이용하였다^[11]. 인공 와우 음성 변조를 사용한 정상 청각인의 언어 인지 능력은 인공 와우 환자의 최상의 언어 인지 능력과 유사하다. 그러므로 CIS 나 SPEAK (spectral peak), ACE (advanced combinational encoder) 등과 같은 포락선 추출에 기반한 strategy의 적용 시 가능한 성능의 상한계를 알 수 있다. 식 10에서 보듯이 입력신호를 $x(t)$ 라 하게 되면, 밴드 제한된 잡음 신호 $n(t)$ 를 Hilbert 변환을 통해 포락선을 추출함에 의해 fine structure 잡음을 얻을 수 있다. 이 포락선은 $x(t)$ 의 fine structure를 랜덤하게 변조한다. r 은 -0.5와 0.5사이에서 균일하게 분포된 랜덤 수이고, l 은 0과 1사이의 잡음 지수이다. 식 (10)에 의해 생성된 신호는 랜덤화에 의한 고주파 요소를 포함하므로 대역 통과 필터에 의해 제거한다. 대역 통과된 신호는 l 이 0일 때 원신호가 복원되고, l 이 1일 때 출력은 채널의 수에 따른 잡음 밴드 변조가 된다^[12]. 식 (10)에 의해 생성된 밴드 제한된 잡음 X_{FS} 는 식 (11)에서 각 채널별로 대역 통과 필터를 거치게 되고 동일한 대역 통과 필터를 거친 입력 신호를 통해 변조된 신호 y_i 를 생성한다.

$$X_{FS}(t) = |H(n(t))| * \cos(\angle(H(x(t))) + 2\pi lr) \quad (10)$$

$$y_i(t) = |x_i(t)| * \cos(\angle(X_{FS}(t))) \quad (11)$$

III. 실험 및 결과

지금까지 기술한 것과 같은 기본적인 구성조건 아래에서 companding strategy에 의한 잡음제거의 효과를 XNR_N 을 평가 기준으로 구하였으며, 다양한 필터 बैं크

를 적용시킨 companding strategy에 의한 잡음제거와 음성 정보 손실의 효과를 XNR_N 과 SD를 평가 기준으로 구하였다. 이렇게 전 처리된 음성 신호의 잡음 상황에서 언어 인지 향상 효과를 인공 와우 시뮬레이션에 의해 변환된 음성 신호의 인지 역치 값을 평가 기준으로 구하였다.

1. Companding strategy에 의한 신호 대 잡음비 향상

설계된 companding strategy를 검증하기 위해 K-SPIN (Korea speech perception in noise) 에서 추출한 /dol/이라는 단음절에(그림 5, 8) 2dB SNR의 백색 잡음

표 1. 잡음 제거에 의한 신호 대 잡음비 향상 값
Table 1. Enhanced XNR_N value due to noise reduction.

Input SNR(dB)	Enhanced XNR_N (dB)	Input SNR(dB)	Enhanced XNR_N (dB)
20	46.0827	0	9.2973
18	42.5062	-2	7.4292
16	32.4625	-4	5.5814
14	27.1207	-6	3.7412
12	23.2967	-8	1.9120
10	20.2350	-10	0.1153
8	17.6305	-12	-1.6062
6	15.3189	-14	-3.1328
4	13.2007	-16	-4.3654
2	11.2096	-18	-5.2571

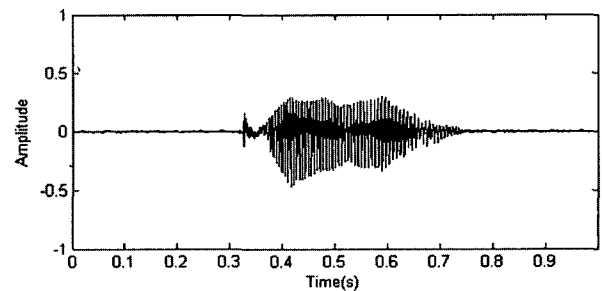


그림 5. /dol/ 원 신호 파형
Fig. 5. Waveform of original sound of /dol/.

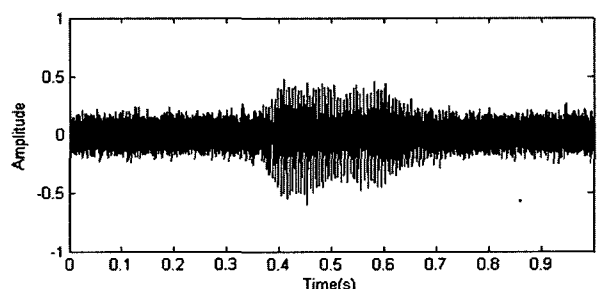


그림 6. 2dB의 신호 대 잡음 비 신호 파형
Fig. 6. Waveform of sound of 2dB SNR.

음을 첨가하여(그림 6) companding strategy 유무에 따른 결과를 시간 축 및 스펙트럼 상에서 비교하였다. Companding strategy 적용 결과, 시간 영역 분석을 통해 잡음이 많이 감소한 것을 볼 수 있으며(그림 7), 주

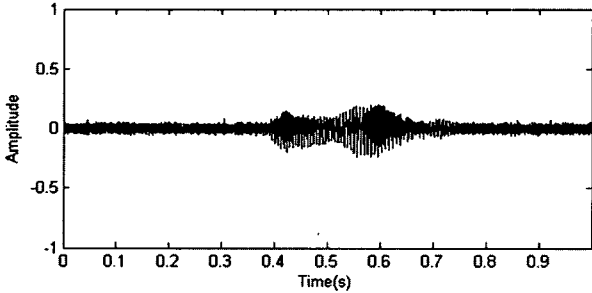


그림 7. Companding strategy 처리된 신호 파형
Fig. 7. Waveform of sound processed with companding strategy.

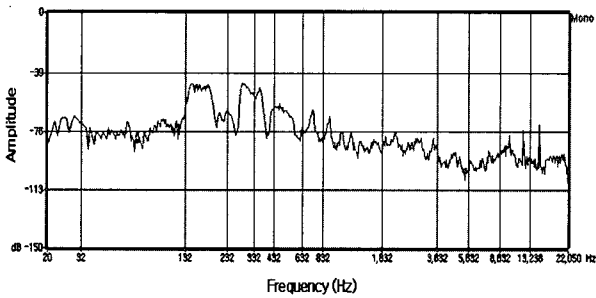


그림 8. 원 신호의 주파수 분석
Fig. 8. Frequency analysis of original sound.

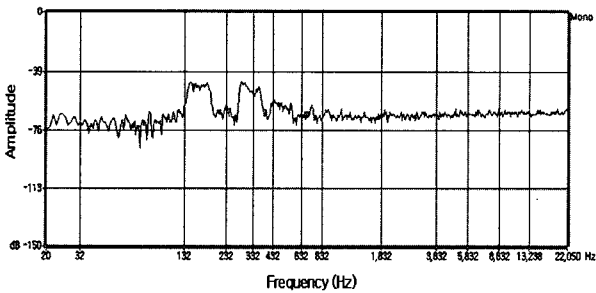


그림 9. 2dB 신호 대 잡음 비 신호의 주파수 분석
Fig. 9. Frequency analysis of sound of 2dB SNR.

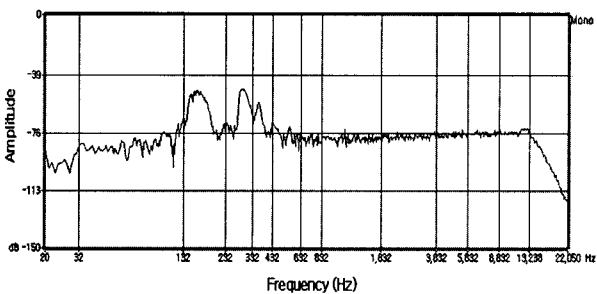


그림 10. Companding strategy 처리된 신호 주파수 분석
Fig. 10. Frequency analysis of sound processed with companding strategy.

파수 영역 분석을 통해 잡음 인가된 스펙트럼보다(그림 9) 스펙트럼 피크 값이 더 강조된 것을 알 수 있다(그림 10). 각 SNR에서 잡음 제거로 인한 XNR_N 의 향상은 표 1과 같다.

2. 필터 뱅크 설계에 따른 음성 손실과 신호 대 잡음비 향상의 교환

그림 11은 q_1 값을 1로, q_2 값을 6으로 하여 잡음 제거에 비중을 둔 고역제 필터 뱅크이며, 그림 13은 q_1 값을 2.8로, q_2 값을 4.5로 하여 음성 정보 비 손실에

표 2. 필터 뱅크 설계에 따른 XNR_N 향상 값과 SD 값
Table 2. Enhanced XNR_N and SD due to filterbank design.

Filterbank ($q_1 : q_2$)	SD (dB)	Enhanced XNR_N (dB)
1 : 6	1.3613	13.6317
1.5 : 5.5	1.3405	12.5408
2 : 5	1.3263	11.6631
2.8 : 4.5	1.3135	11.0325

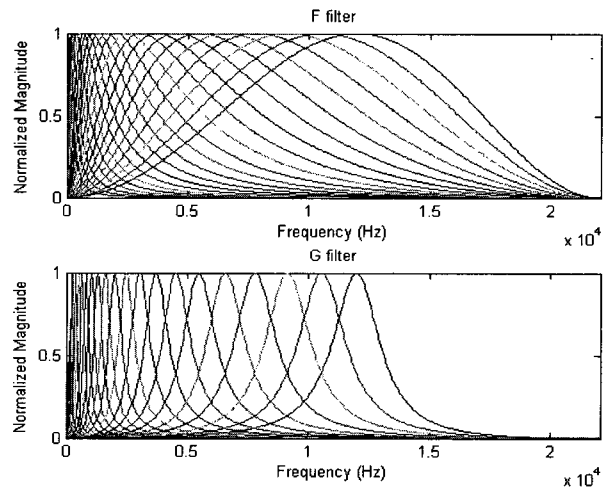


그림 11. 고역제 필터 뱅크 설계 ($q_1=1, q_2=6$)
Fig. 11. Filterbank of high suppression ($q_1=1, q_2=6$).

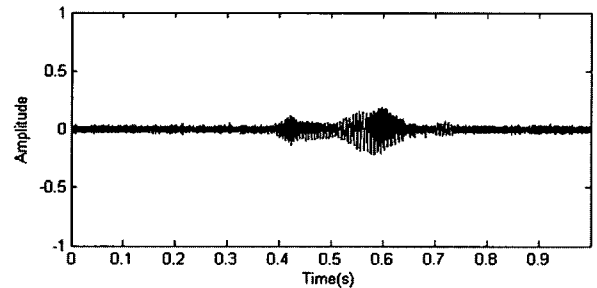


그림 12. 고역제 필터 뱅크에 의한 companding strategy 처리된 파형
Fig. 12. Waveform of sound processed with filterbank of high suppression.

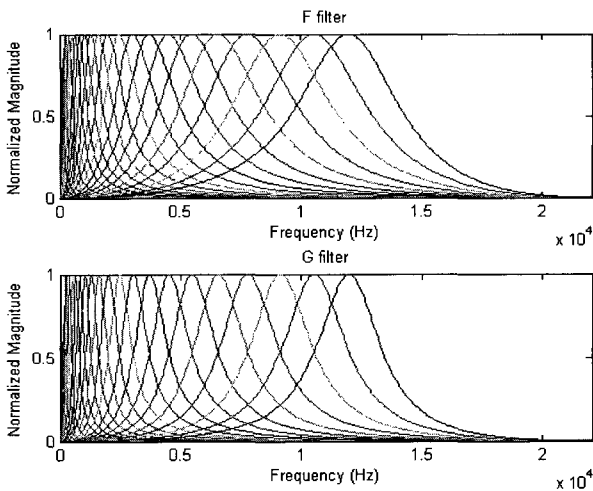


그림 13. 저 억제 필터뱅크 설계 ($q_1=2.8, q_2=4.5$)

Fig. 13. Filterbank of low suppression ($q_1=2.8, q_2=4.5$).

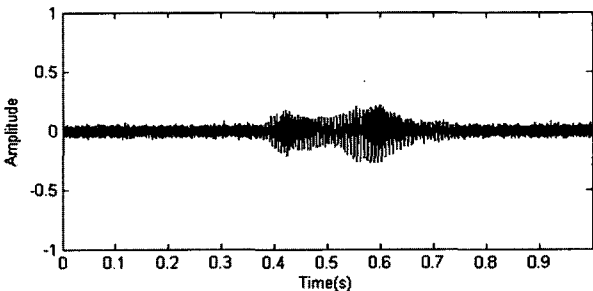


그림 14. 저 억제 필터뱅크에 의한 companding strategy 처리된 파형

Fig. 14. Waveform of sound processed with filterbank of low suppression.

비중을 둔 저 억제 필터이다. 앞서 사용한 단음절에 2dB의 백색 잡음을 첨가하여 두 필터뱅크를 적용시킨 결과, 잡음 제거와 음성 정보 손실이 교환되었음을 볼 수 있다(그림 12, 14). 다양한 필터뱅크에 따른 잡음 제거와 음성 정보 손실의 교환 정도는 표 2와 같다.

3. 인공 와우 환자 시뮬레이션 임상 실험

본 연구에서는 인공 와우 시스템을 가정하기 위해, 잡음 지수는 1로 설정하여 신호의 포락선만 전달하였으며, 채널은 6채널로 설정하였다. 앞서 사용한 /돌/이라는 단음절을 설정된 채널수와 잡음 지수를 이용하여 시뮬레이션 하였으며, 변환된 신호의 시간 영역에서의 결과는 그림 15와 같다.

실험은 Training, 개인의 신호 대 잡음 비 역치 값 측정, 알고리즘 적용 후 개인의 신호 대 잡음 비 역치 값 측정으로 진행하였다. 실험에 사용된 12개의 단어는 K-SPIN에서 스펙트럼의 유사성을 기준으로 추출하였으며, 잡음은 백색 잡음을 이용하였다. 실험 대상은

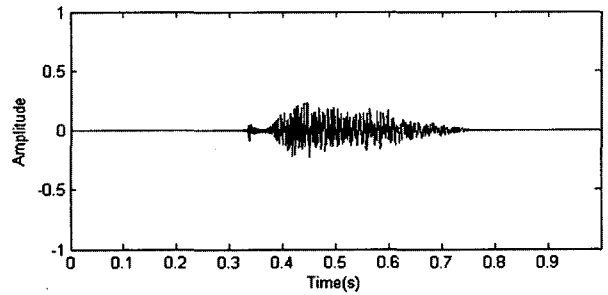


그림 15. 6 채널, 잡음지수 1로 변조된 인공 와우 음성 시뮬레이션 파형

Fig. 15. Waveform of simulated sound using 6 channel and noise factor 1.

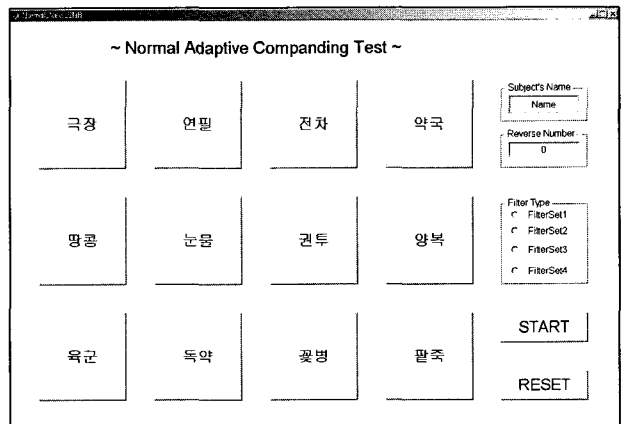


그림 16. 신호 대 잡음비 역치 값 측정 GUI

Fig. 16. GUI of SNR Threshold estimation.

26~30살의 남성 5명을 대상으로 하였다. 먼저, 인공 와우 음성 시뮬레이션에 의해 변조된 소리와 실험 환경에 실험자가 익숙해지기 위한 반복 훈련을 실시하였다. 반복 훈련은 변조된 음성 데이터를 랜덤으로 실험자에게 들려주었으며, 24번을 연속으로 맞춘 후에 훈련이 종료하게 된다. 잡음 상황에서 개인의 SNR 역치 값을 구하기 위해 원 음성 데이터에 백색 잡음을 더해주고 이를 시뮬레이션 변환하여 실험자에게 들려주었다. 역치 값 측정 방식은 two-alternative force-choice 패러다임을 사용하여 측정되었다. 1-up 1-down 적응 방법은 50% 인지율을 추정하기 위해 사용되었다^[13]. 실험의 시작에서 최초 SNR은 20dB로 설정되었고, 실험자가 신호를 맞추면 SNR이 2dB 감소하게 되고, 틀리면 SNR이 2dB 증가하게 된다.

증감소가 교차하는 횟수(reverse number)가 14회일 때, 실험이 끝나게 되고, 마지막 10개의 reverse number의 SNR 값을 평균하여 측정하였다. 마지막으로, companding strategy를 적용 후 잡음 상황에서 개인의 SNR 역치 값을 구하기 위해 2가지 설계된 필터뱅크에 대해 실험하였으며, 그 외 실험 프로토콜은 앞 실험과

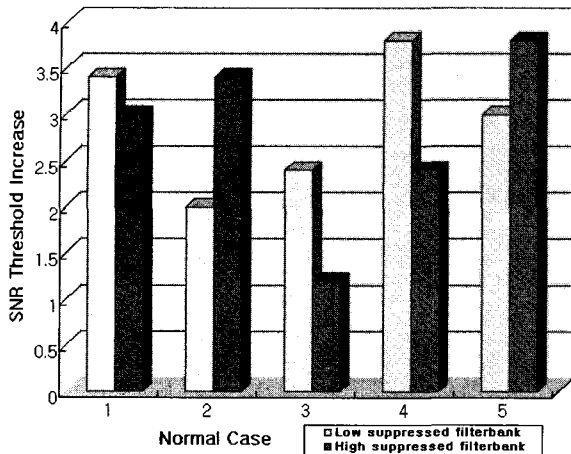


그림 17. 5명의 신호 대 잡음비 역치 값 향상 값
Fig. 17. Enhanced SNR threshold from 5 subjects.

동일하다. 그림 16은 MATLAB 환경에서 개발된 실험 시 사용된 프로그램이다. 본 실험 결과 그림 17과 같이 모든 실험자들이 각 개인의 SNR 역치 값보다 더 낮은 SNR에서 언어 인지 향상을 보였다.

필터 बैं크 설계에 따른 개인 간의 언어 인지 특성으로 인해 개인마다 적절한 필터 बैं크가 다르게 나타났으며, 고역제 필터 बैं크에서는 평균 2.92dB의, 저역제 필터 बैं크에서는 평균 2.76dB의 SNR 역치 값이 증가했다.

IV. 결 론

본 연구를 정상 청각 인을 대상으로 적용하였을 때, companding strategy가 잡음 상황에서 언어 인지 향상에 도움이 됨을 알 수 있었다. 현재 인공 와우 환자에게 적용되는 strategy는 각 필터 밴드의 에너지나 크기에 기반 하여 자극 채널을 결정하므로 0dB 이하의 신호 대 잡음비에서는 언어 인지 능력이 현저하게 감소한다. 제안된 알고리즘은 각 채널의 중심주파수 스펙트럼 피크를 강조시키므로 인공 와우 환자의 잡음 제거 및 음성 강조에 큰 효과가 있을 것으로 기대된다. 인공 와우 시스템을 환자에게 시술 후 사용 시에 일반적으로 각 개인의 불편함을 느끼지 않는 가장 큰 크기의 C-level, 인지 가능한 크기의 역치값인 T-level을 측정하여 시스템 환경에 적용한다. 기존 방식에 추가하여 본 연구에서 제시한 각 개인의 청력 특성에 맞는 적절한 필터 बैं크 정보를 적용한다면 잡음 상황에서 인공 와우 환자의 언어 인지 향상에 큰 도움이 될 것이다.

참 고 문 헌

- [1] P.C.Loizou, "Introduction to Cochlear Implants" IEEE Engineering in Medicine and Biology, pp. 32-42, 1999.
- [2] van Hoesel, R.J.M., Clark, G.M., "Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients" The Journal of the Acoustical Society of America, vol. 97, no. 4, pp.2498-, 1995.
- [3] Weiss, M.R., "Effects of noise and noise reduction processing on the operation of the Nucleus-22 cochlear implant processor" Journal of rehabilitation research and development, vol. 30, no. 1, pp. 117-, 1993.
- [4] Yang, L., Fu, Q., "Spectral subtraction-based speech enhancement for cochlear implant patients in background noise" The Journal of the Acoustical Society of America, vol. 117, no. 3, pp. 1001-1004, 2005.
- [5] L. Tricchia and R. Sarpeshkar, "A bio-inspired companding strategy for spectral enhancement" IEEE transactions on speech and audio processing, vol. 13, no. 2, pp. 243-253, 2005.
- [6] M. Charles Liberman, "The cochlear frequency map for the cat : Labeling auditory-nerve fibers of known characteristic frequency" The Journal of the Acoustical Society of America, vol. 72, no. 5, pp. 1441-1449, 1982.
- [7] Junpei Yamauchi, Tetsuya Shimamura, "Noise estimation using high frequency regions for speech enhancement in low SNR environments" Speech Coding, IEEE Workshop Proceeding 2002, pp. 59-61, 2002.
- [8] K. Itoh, et al. "A study of Objective Quality Measures for Digital Speech Waveform Coding Systems", The Institute of Electronics, Information and Communication Engineers(IEICE), Vol. J 66-A, No. 3, pp. 274-281, 1983.
- [9] R.V.Shannon, F.G.Zeng, V.Kamath, J.Wygonski, M.Ekelid. "Speech recognition with primarily temporal cues", Science, 270, pp. 303-304, 1995.
- [10] M.F.Dorman, P.C.Loizou, D.Rainey. "Speech understanding as a function of the number of channels of stimulation for processors using sine-wave and noise-band outputs", The Journal of the Acoustical Society of America, 102, pp. 2403-2411, 1997.
- [11] Michael K.Qin and Andrew J.Oxenham, "Effects of Envelope-Vocoder Processing on F0 Discrimination and Concurrent-Vowel Identification" Ear & Hearing, vol. 26, pp. 451-460, 2005.

- [12] J. T Rubinstein, C. Turner, "A Novel acoustic simulation of cochlear implant hearing : Effects of Temporal Fine Structure" Nerual Engineering, Conference Proceeding. First international IEEE EMBS Conference on, pp.142-145, 2003.
- [13] Levitt, H. "Transformed Up-Down Methods in Psychoacoustics" The Journal of the Acoustical Society of America, vol. 49, no. 2B, pp. 467-477, 1970.

— 저 자 소 개 —



이 영 우(학생회원)
 2004년 인제대학교 의용공학과
 학사 졸업.
 2006년 한양대학교 의용생체공학
 과 석사과정.
 <주관심분야 : 생체 신호 처리, 청
 각 보철 장치>



지 윤 상(학생회원)
 2003년 고려대학교 제어계측
 공학과 학사 졸업
 2005년 한양대학교 의용생체
 공학과 석사 졸업
 2005년 한양대학교 의용생체
 공학과 박사 과정

<주관심분야 : 음성 신호처리, 인공와우 부호화 >



이 종 실(정회원)
 1995년 인하대학교 전자공학과
 학사 졸업.
 1997년 인하대학교 전자공학과
 석사 졸업.
 2005년 인하대학교 전자공학과
 박사 졸업.

2001년~2005년 한국산업기술대학교 전자공학과
 겸임교수
 2005년~한양대학교 의공학교실 연구교수
 <주관심분야 : 생체신호처리, 무선센서네트워크>



홍 성 화(정회원)
 1984년 서울대학교 의과대학
 이비인후과 학사 졸업.
 1991년 서울대학교 의과대학
 이비인후과 석사 졸업.
 1996년 서울대학교 의과대학
 이비인후과 박사 졸업.

1993년 충북대학교 의과대학 이비인후과 조교수
 2001년~성균관대학교 의과대학 이비인후과학
 부교수
 2001년~삼성 서울 병원 이비인후과 전문의
 <주관심분야 : 청각학, 전기생리학>

김 인 영(정회원)
 대한전자공학회논문지 제 42권 SC편 제 3호 참조

김 선 일(정회원)
 대한전자공학회논문지 제 42권 SC편 제 3호 참조

이 상 민(정회원)-교신저자
 대한전자공학회논문지 제 42권 SC편 제 3호 참조