

난청인의 주파수 선택도 둔화현상이 음질에 미치는 영향 평가

| |
|---------|
| 논 문 |
| 60-6-19 |

The Assessment on the Sound Quality of Reduced Frequency Selectivity of Hearing Impaired People

안 홍 섭* · 박 규 석* · 전 유 용** · 송 영 록*** · 이 상 민†
(Hong-Sub An · Gyu-Seok Park · Yu-Yong Jeon · Young-Rok Song · Sang-Min Lee)

Abstract - The reduced frequency selectivity is a typical phenomenon of sensorineural hearing loss. In this paper, we compared two modeling methods for reduced frequency selectivity of hearing impaired people. The two models of reduced frequency selectivity were made using LPC(linear prediction coding) algorithm and bandwidth control algorithm based on ERB(equivalent rectangular bandwidth) of auditory filter, respectively. To compare the effectiveness of two models, we compared the result of PESQ (perceptual evaluation of speech quality) and LLR(log likelihood ratio) using 36 Korean words of two syllables. To verify the effect on noise condition, we mixed white and babble noise with 0dB and -3dB SNR to speech words. As the result, it is confirmed that the PESQ score of bandwidth control algorithm is higher than the score of LPC algorithm, on the other hands, and the LLR score of LPC algorithm is lower than the score of bandwidth control algorithm. It means that both non-linearity and widened auditory filter characteristics caused by reduced frequency selectivity could be more reflected in bandwidth control algorithm than in LPC algorithm.

Key Words : Reduced frequency selectivity, Widened auditory filter, ERB, LPC

1. 서 론

난청을 가진 사람들은 청력 손실로 인해 의사소통에 많은 어려움을 느낀다. 이러한 난청은 크게 두 가지로 분류되는데, 하나는 외이, 고막, 중이 등 소리를 전달해주는 기관의 장애나 성능저하로 인하여 음파의 전달이 정상적으로 이루어지지 않는 상태로 발생하는 전음성 난청이고, 또 다른 하나는 달팽이관내의 소리 감지 기능에 이상이 생기거나 소리 자극을 뇌로 전달하는 청신경 또는 중추신경계의 이상으로 발생하는 감각신경성 난청이 있다. 이러한 난청의 치료 또는 보상에 있어서 대부분의 전음성 난청은 수술적 처치에 의해 극복이 가능한 상태이나 감각신경성 난청의 극복은 아직까지 많은 과제들이 남아 있다.

감각 신경성 난청은 청각의 문턱 값(threshold of hearing level)이 높고, 좁은 역동범위(dynamic range)를 가지며, 주파수 선택도(frequency selectivity)가 둔화된다는 특징을 가지고 있다[1]. 이러한 감각 신경성 난청의 해결 방법으로 제안되는 것 중 하나가 보청기의 착용이다. 보청기는 난청인

의 청력 손실을 보상해주도록 고안된 의료기기이다. 하지만 실제 보청기 착용자는 현재 보청기의 청력 보상기능에 아직 만족하지 못하고 있다. 이는 보청기가 난청인의 청력 보상을 하도록 고안되었지만, 난청인이 어떻게 소리를 듣는지는 모르는 상태에서 소리보상을 해주려고 하기 때문이다. 따라서 난청인이 어떻게 소리를 듣는지를 확인할 필요가 있는데, 우리가 주목한 점은 주파수 선택도 둔화현상이 난청인이 듣는 소리에 어떠한 영향을 미치는지에 관한 것이다. 특히 노인성 난청의 경우 둔화된 주파수 선택도가 음성신호를 얼마나 왜곡하는가를 평가해 보고자 한다.

본 논문에서는 주파수 선택도의 둔화 현상이 음질에 미치는 영향을 알아보기 위한 방법으로 두 가지 알고리즘을 살펴보고 비교 하였다. 두 알고리즘 모두 난청인의 청각필터가 정상인의 청각필터에 비하여 넓어진다는 현상에 착안하였으나 그 표현 방식에 있어서 한 가지는 선형 예측 부호화(linear prediction coding, LPC)를 이용해 스펙트럼상의 뭉개짐을 구현하여 선형적으로 주파수 선택도 둔화를 표현하였고[2], 다른 하나는 청각필터(auditory filter)를 구현한 후 이것의 ERB(equivalent rectangular bandwidth)를 넓히는 방식으로 비선형적으로 주파수 선택도 둔화를 표현한 것이다[3]. 이 두 가지 알고리즘이 음성신호에 주는 영향을 살펴보기 위하여 36개의 한국어 2음절 단어를 알고리즘에 적용하여 신호처리를 한 뒤 객관적 어음 인지도 평가 방식인 PESQ(perceptual evaluation of speech quality)와 LLR(log likelihood ratio)로 음성의 음질과 왜곡의 정도를 비교 분석하였다.

* 비 회 원 : 인하대학교 전자공학과 석사과정
** 비 회 원 : 인하대학교 전자공학과 박사과정
*** 비 회 원 : 인하대학교 정보전자공동연구소 연구교수
† 교신저자, 정회원 : 인하대학교, 전자공학과 부교수,
인하대학교 정보전자 공동연구소

E-mail : sanglee@inha.ac.kr

접수일자 : 2011년 5월 11일

최종완료 : 2011년 5월 20일

2. 이론적 배경

2.1 주파수 선택도 둔화현상

음조(pitch)는 소리의 음악적인 요소, 즉 소리의 높낮이에 대한 특성을 말하는 것으로 순음의 경우에는 주파수에 관여하게 되고 주기적인 복합음의 경우에는 그 음의 기본 주파수(fundamental frequency)에 관여된다고 한다. 주파수 선택도는 서로 다른 주파수를 가진 두 음을 차례대로 듣고 어떤 소리가 높은 음을 가졌는지 구분 할 수 있는 정도 또는 자극 음을 주었을 때 시간에 따른 주파수의 변화를 판단할 수 있는 정도를 말한다. 구별이 가능한 최소 크기의 변화를 역치 분해능(difference limen, DL)이라고 하는데 주파수 선택도를 알아보기 위한 것을 주파수 상의 역치 분해능(difference limen for frequency, DLF)이라고 한다. 정상적인 청각 능력을 가진 경우 주요 대화 영역인 0.5~2kHz에서는 0.2%~0.3%의 주파수 차이가 있어도 주파수 구별이 가능하다. 그러나 달팽이관 손상을 받은 난청인들을 대상으로 DLF를 측정한 여러 연구의 전반적인 결론은 결과에 대한 개인 간의 차이가 있고 청력역치와의 관계는 뚜렷하지 않으나 주파수 선택도가 떨어진다고 한다. 이는 달팽이관의 손상이 있게 되면 청각 필터는 훨씬 넓어지며 위상 고정에 대한 정확성이 떨어진 것과 깊은 관계가 있을 것이다. 음조는 언어를 이해하는데 매우 중요할 뿐 아니라 나이, 성별, 감정 상태 등 언어 외적인 정보를 얻는데도 아주 중요한데 달팽이관 손상이 있는 경우 음조 감지 능력이 떨어져서 언어를 이해하는 데에 지장을 주게 된다.

정상인과 달팽이관 손상을 받은 난청인 사이의 주파수 선택도를 비교하는 것은 청각 필터의 모양에 어떠한 영향을 주는 가를 이용하면 쉽게 이해된다. 달팽이관 손상이 있는 난청인에게서 청각 필터의 모양을 구해보면 그림 1의 (b)와 같이 전반적으로 필터의 대역이 넓어지고 청력의 손실이 더 많은 난청인에게서 더 넓은 대역을 보이는데 특히 저주파 영역으로 대역이 훨씬 넓어져서 주위에 주로 잡음을 이루고 있는 낮은 음에 의해서 훨씬 더 잘 차폐 될 것을 시사하고 있다[4]. 또한 음색의 구분도 주파수 선택도에 의존한다고 알려져 있어 달팽이관 손상이 있는 경우 악기의 구분이나 모음 등의 구분이 쉽지 않다.

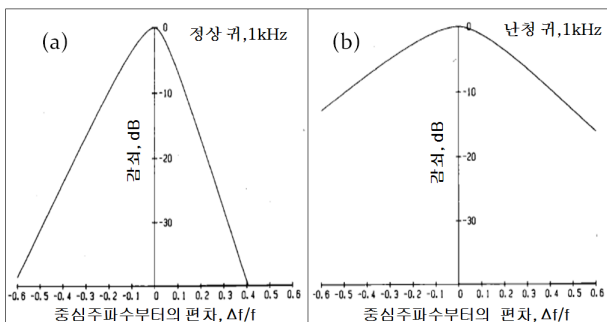


그림 1 중심 주파수 1kHz의 청각 필터: (a) 정상인, (b) 난청인
 Fig. 1 Auditory filter shapes of a center frequency of 1kHz for the normal (a) and impaired (b) ears

2.2 청각 필터(auditory filter)와 ERB(equivalent rectangular bandwidth)

필터는 청각학과 심리음향학의 많은 측면에서 사용되는 개념이다. 일반적으로 청각 필터는 겹쳐진 대역 통과 필터(band-pass filter)들의 집합으로 표현되는데 이것은 달팽이관의 위치에 따라 담당하는 음의 주파수 영역이 다를 것을 의미하는 토노토픽(tonotopic)의 개념을 가져오는데 사용되기도 한다[5]. 또한, 청각 필터는 기저막(basilar membrane)의 위치와 달팽이관의 주파수 선택도를 결정하는 것과 관련이 있다. 청각 필터는 비선형적이고 레벨에 종속적이며 대역폭은 기저막의 높은 주파수에 해당하는 영역에서 낮은 주파수에 해당하는 영역으로 갈수록 감소한다[6][7]. 따라서 청각 필터를 설계하고 그 개념을 사용 할 때는 청각 필터의 비선형적 특징을 반영해야 한다. 청각 필터에 관련된 또 다른 개념으로는 ERB가 있다. 이는 청각 필터와 중심주파수 그리고 임계대역폭(critical bandwidth)과 관련된 식으로 나타내어지는데 아래의 식 (2)에 이를 나타내었다. 청각 필터의 모양을 결정짓는 함수의 대표적인 방법은 아래 식 (1)에 나타낸 roex(p) 필터를 사용한 방법이 있다.

$$W(g) = (1 + pg)e^{-pg} \quad (1)$$

여기서 g 는 필터의 중심주파수인 f_c (center frequency)로부터의 주파수 차를 중심 주파수로 정규화한 것이며, p 는 필터의 경사를 제어하는 파라미터이다. W 는 필터의 모양을 표현하는 함수이며 p 값에 의하여 필터의 경사가 제어되기 때문에 중심주파수를 경계로 양쪽의 모양이 다른 비대칭 필터를 구현하기 위해서는 p 값을 중심주파수를 경계로 하여 아래쪽은 p_l 로 위쪽은 p_u 값으로 나누어 각각 다르게 설정하여 주면 된다. 여기서 청각 필터의 ERB는 $2f_c/p_l + 2f_c/p_u$ 로 표현 될 수 있고 Moore의 연구[8]에 따른 다음의 식 (2)와 같이 표현 할 수도 있다.

$$ERB = 24.7(0.00437f_c + 1) \quad (2)$$

여기서 f_c 와 ERB의 단위는 헤르츠(hertz)이다.

아래 그림 2는 위의 식 (1)을 이용하여 만든 중심주파수 (f_c)가 1kHz인 청각 필터의 모습을 나타낸다. 그림 2의 (a)는 p_l 과 p_u 값이 같은 대칭필터를 나타내며 p_l , p_u 값의 변화에 따라 필터의 기울기가 달라지는 모습을 볼 수 있다. 그림 2의 (b)는 p_l 과 p_u 값이 서로 다른 비대칭필터를 나타낸다.

3. 알고리즘 구현

3.1 선형 예측 부호화(LPC)를 이용한 주파수 선택도 둔화현상 표현

앞에서 말한 주파수 선택도 둔화현상이 난청인에게 어떠한 소리로 들리게 되는 지를 알아보기 위해 스펙트럼상의 무게 중심을 LPC로 나타낸 연구가 있었다[2]. LPC는 현재의 음성 신호 샘플 값은 과거의 음성출력 샘플 값과의 선형 결합으

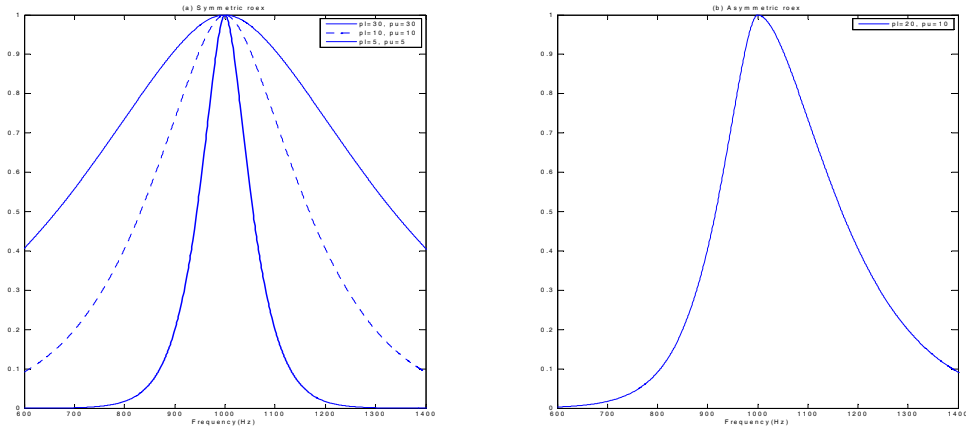


그림 2 roex(p)필터를 이용하여 만든 중심주파수가 1kHz인 청각 필터: (a) p, p_u값의 변화에 따른 대칭 필터의 변화모습, (b) 비대칭 필터

Fig. 2 Auditory filter shape of center frequency of 1kHz for the symmetric (a) and asymmetric roex filter

로 근사할 수 있다는 가정에서 시작된 분석법이다. LPC는 음성을 코딩하는 작업은 시간 영역에서 이루어지지만 음성의 분석 및 합성은 주파수 영역에서 이루어진다. LPC는 주파수 영역에서 코딩해 주면서 포먼트 주파수, 진폭 등의 파라미터 값을 추출해 낼 수 있다. LPC 계수를 계산하는 방법으로 가장 일반적으로 사용하는 방법은 자기상관법 (autocorrelation method)이다. 이 연구에서 사용한 함수는 자기상관법 중 가장 일반적인 방법인 Levinson-Durbin의 알고리즘[9]을 사용하여 계수를 계산하였다.

스펙트럼상의 뭉개짐은 그림 3의 실선에서 확인할 수 있듯이 주파수의 세밀한 성분들의 크기가 뭉개진다. 그림 3에서 LPC 차수에 따라 스펙트럼상의 뭉개짐 되는 정도가 다른 것을 볼 수 있다. 그림 3(a), (b)의 점선은 고속 푸리에 변환(FFT)시킨 음성 파형에 대하여 LPC를 적용하지 않은 음성의 주파수 스펙트럼으로 주파수 성분의 자세한 변화가 나타나고 그림 3(a)의 실선은 LPC 차수가 20차로 처리한 음성의 스펙트럼으로 주파수 성분의 대략적인 변화만 나타난다. 그림 3(b)의 실선은 12차로 처리한 음성의 스펙트럼이다.

3.2 청각 필터와 ERB를 이용한 주파수 선택도 둔화 현상 표현

앞에서 설명한 LPC를 이용한 주파수 선택도 둔화 현상 표현 방법이 수학적인 모델링 기법이라면 여기에서 설명할 방법은 청각 생리학을 바탕으로 한 수학적 모델링 기법이라고 할 수 있다. 난청인의 주파수 선택도 둔화 현상을 표현하기 위해서 먼저 청각 필터의 모양을 모델링 할 필요가 있는데 본 논문에서는 앞에서 설명한 roex(p) 필터를 이용하여 청각 필터의 모양을 모델링 하였다. 그 후 주파수 선택도 둔화 현상을 표현하는 방법으로 난청인의 넓어진 청각 필터를 모델링 하기 위해 ERB값을 키우는 방법을 제시한다. ERB값이 커지면 roex(p) 필터의 p파라미터 값이 작아지게 되고 이는 결과적으로 난청인의 넓어진 청각 필터를 나타내게 된다.

그림 4는 넓어진 청각 필터의 특성을 반영할 수 있는 알고리즘의 블록도이다. 먼저, 16kHz로 샘플링된 음성 신호 128개를 하나의 프레임으로 나누고 해밍윈도우를 통과 시킨 후 신호의 앞과 뒤에 각각 64개의 0을 삽입하여 제로 패딩의 과정을 행한다. 그 후 고속 푸리에 변환(FFT)을 통하여

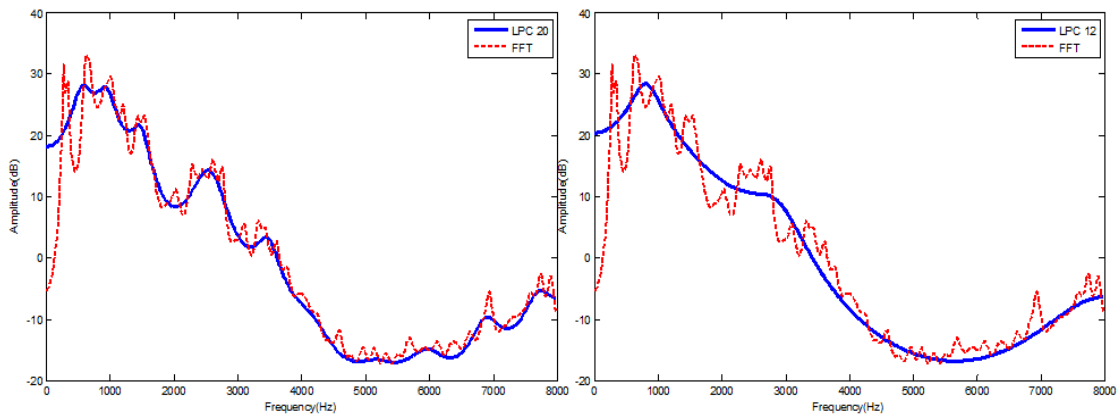


그림 3 각 LPC차수에 따른 음성 스펙트럼: (a) 20차, (b) 12차

Fig. 3 The power spectrum graph of speech according to LPC order for the 20 (a) and 12 (b) LPC order

주파수 영역으로 변환 후 파워와 위상으로 나눈다. 나뉜 파워에는 식(1)과 식(2)를 통하여 구한 청각필터의 특성을 모델링한 256-by-256 행렬을 곱한 후 원신호의 위상과 더한 후에 역 고속 푸리에 변환(IFFT)를 한다. 그 후 overlap-add를 통하여 한 개의 출력 프레임에 4개의 프레임이 겹치도록 연산을 수행한다.

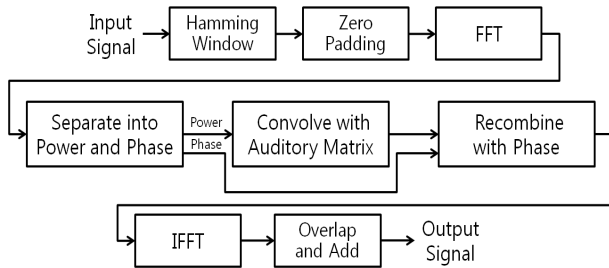


그림 4 청각필터의 특성을 음성에 적용하기 위한 알고리즘의 블록도

Fig. 4 The block diagram of algorithm for reflect the characteristic of auditory filter

그림 5에 이러한 과정을 거쳐서 출력된 음성의 스펙트럼을 나타내었다. 여기서 ERB*1은 정상 청각 필터를 나타내며 ERB*3, ERB*6은 난청인의 넓어진 청각필터를 나타낸다. 그림 5는 한국어 문장 “안녕하세요”를 왼쪽 위에서 각각

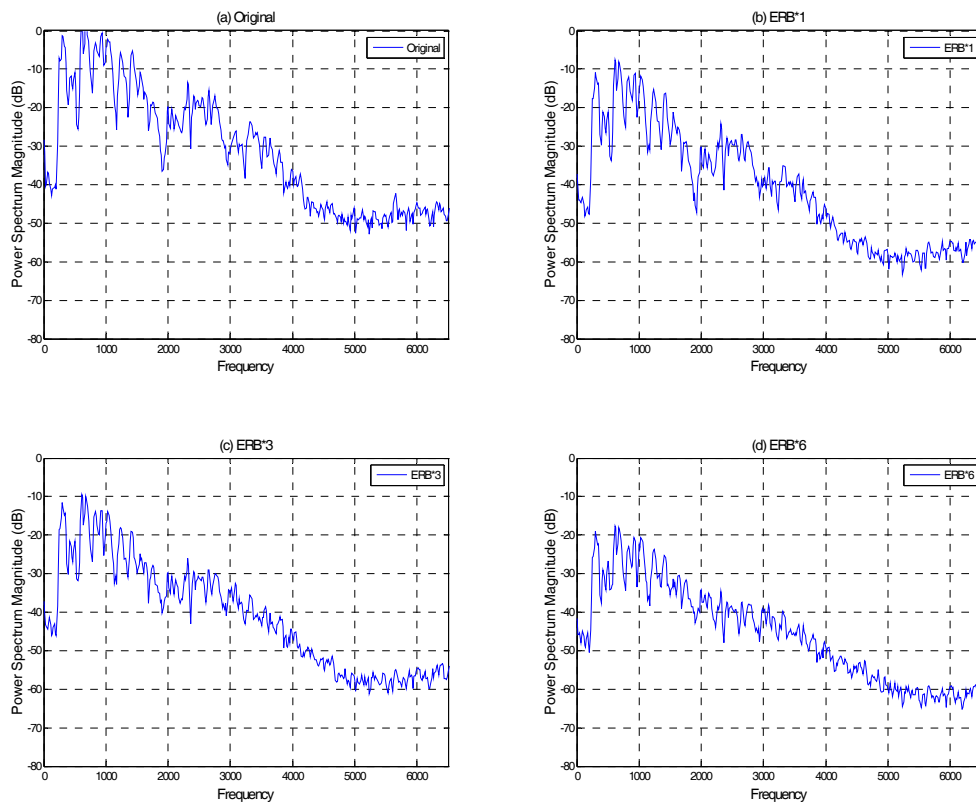


그림 5 원신호와 ERB*1, ERB*3, ERB*6 의 청각 필터를 통과시킨 후 출력된 음성의 스펙트럼

Fig. 5 The power spectrum graph for the original speech signal (a) and were passed auditory filters ERB*1 (b), ERB*3 (c) and ERB*6 (d)

원신호(a)와 ERB*1(b), ERB*3(c), ERB*6(d) 의 청각 필터를 통과시킨 후 출력된 음성의 스펙트럼을 나타낸 것이다.

원신호와 ERB*1을 비교해 보면 전체적인 음압이 떨어진 것 이외에는 큰 차이점을 찾아 볼 수 없다. 하지만 원신호와 ERB*3, ERB*6을 비교해 보면 넓어진 청각 필터의 효과로 인하여 음성신호의 음조특성이 사라진 것을 볼 수 있으며 이는 청각 생리학을 바탕으로한 수학적 모델링 기법이 주파수 선택도 둔화 현상을 잘 나타내고 있음을 보여준다.

4. 객관적 음질평가 실험

4.1 객관적 음질평가 실험 방법

실험에 사용된 단어는 한국어 2음절 단어 36개이다. 각각의 단어들은 잡음이 없는 상황, 백색(white) 잡음이 섞인 상황, 배블(babble) 잡음이 섞인 상황으로 구별되며 이는 다시 각각 0dB 신호 대 잡음비인 경우와 -3dB 신호 대 잡음비인 경우로 나누어진다. 이렇게 원음과 노이즈가 섞인 음성을 12차수 LPC와 20차수 LPC 알고리즘을 적용한 것과 ERB*3, ERB*6 의 알고리즘을 적용한 것으로 나누어 실험에 사용하였다. 객관적 음질 평가 도구로는 음질을 나타내는 PESQ와 음성의 왜곡 정도를 평가하는 LLR을 사용하였다.

4.2 객관적 음질평가 실험 도구

4.2.1 PESQ(Perceptual Evaluation of Speech Quality)

PESQ는 음성 품질의 객관적 평가 도구로서 음성 품질의 자동 평가를 위해 개발되었다. 이는 ITU-T의 권고안 P.862로 표준화 되었으며 오늘날 전화기 제조업체, 네트워크 장비업체와 통신 사업자들의 객관적인 음성 품질 테스트에 사용되는 전 세계적인 산업 표준이다.

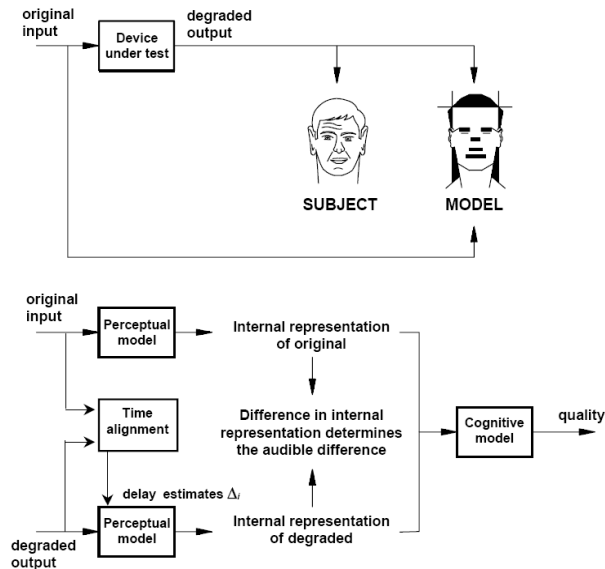


그림 6 PESQ의 개념을 나타내는 블록다이어그램
Fig. 6 The block diagram for the conception of PESQ

PESQ의 평가 결과는 주관적 평가결과인 MOS와 0.935의 높은 상관도를 보여 그 안정성이 이미 확인 되었다[10]. PESQ 점수는 최종적으로 계산된 교란 값과 평균비대칭 교란 값의 선형조합으로 구해지는데 그 범위는 -0.5에서 4.5 사이의 값을 가지며 원음과 유사 할수록 높은 점수가 나오게 된다. PESQ의 개념을 나타내는 블록다이어그램을 그림 6에 나타내었다. 또한 PESQ의 이름에서도 알 수 있듯이 인간의 소리 인지에 대한 요소도 포함되어있다. 그림 6에 PESQ의 개념에 대한 블록다이어그램을 나타내었다[11]. PESQ는 PESQ값을 구하려는 음과 그 음의 원음이 필요하다. 이 두 음은 레벨(level) 표준화 과정을 거치고 IRS 필터링을 한 후, 지연(delay) 보상 과정을 지나 음성 모델 과정을 거쳐서 최종 PESQ값이 나오게 된다.

4.2.2 LLR(Log Likelihood Ratio)

LLR은 음성의 왜곡을 객관적으로 측정하는 도구로서 음성신호간의 all-pole 모델의 차이를 계산한다. 음성신호는 p차의 all-pole 모델로 나타내어진다.

$$x(n) = \sum_{i=1}^p a_x(i)x(n-i) + G_x u(n) \quad (3)$$

여기서 $a_x(i)$ 는 all-pole필터의 계수이며, G_x 는 필터의 이득 그리고 $u(n)$ 은 백색잡음의 단위 분산이다. all-pole을 기초로 하는 측정은 LLR과 IS(Itakura-Saito)가 대표적이다.

$$LLR = \log_{10} \left(\frac{\overline{a_x} R_x \overline{a_x}^T}{a_x R_x a_x^T} \right) \quad (4)$$

여기서 a_x^T 는 깨끗한 음성신호의 LPC계수이며, $\overline{a_x}^T$ 는 향상된 음성신호의 LPC계수이며, R_x 는 깨끗한 음성신호의 자기상관 행렬이다[12].

5. 실험 결과

5.1 PESQ 평가 실험 결과

표 1은 백색(white) 잡음에서의 실험 결과의 평균 PESQ 수치이고 표 2는 배블(babble) 잡음에서의 실험결과와 평균 PESQ 수치이다. 그림 7은 2음절 한국어 단어 36개의 평균 PESQ 값을 나타낸 그림이다. 각각의 단어들은 잡음이 없는 상황, 0dB SNR인 상황, -3dB SNR인 상황으로 나누어지며 ERB*3, 20차수 LPC, ERB*6, 12차수 LPC 알고리즘을 통과 시킨 후 노이즈가 없는 상황에는 원음과 비교를 하고 노이즈 상황에서는 노이즈와 원음을 섞은 음과의 비교를 통해 평균 PESQ 값을 계산하였다. 그림 7의 (a)는 백색(white)잡음에서의 상황이며 (b)는 배블(babble)잡음에서의 상황이다.

표 1 백색(white) 잡음에서의 평균 PESQ

Table 1 Average PESQ score in the white noise

| | ERB*3 | 20LPC | ERB*6 | 12LPC |
|----------|--------|--------|--------|--------|
| No noise | 3.6382 | 2.7724 | 3.0494 | 2.7966 |
| 0dB SNR | 4.1452 | 3.6688 | 3.7204 | 3.6468 |
| -3dB SNR | 4.1938 | 3.8130 | 3.8138 | 3.7908 |

표 2 배블(babble) 잡음에서의 평균 PESQ

Table 2 Average PESQ score in the babble noise

| | ERB*3 | 20LPC | ERB*6 | 12LPC |
|----------|--------|--------|--------|--------|
| No noise | 3.6382 | 2.7724 | 3.0494 | 2.7966 |
| 0dB SNR | 4.1194 | 3.6432 | 3.5730 | 3.6222 |
| -3dB SNR | 4.1136 | 3.7498 | 3.6074 | 3.7062 |

본 연구에서는 난청인의 넓어진 청각필터를 표현하기 위한 방법으로 LPC와 청각필터의 ERB를 제시 하였는데 이를 비교하기 위해서 ERB*3, ERB*6에 각각 대응하는 LPC의 차수를 20개와 12개로 제시한다.

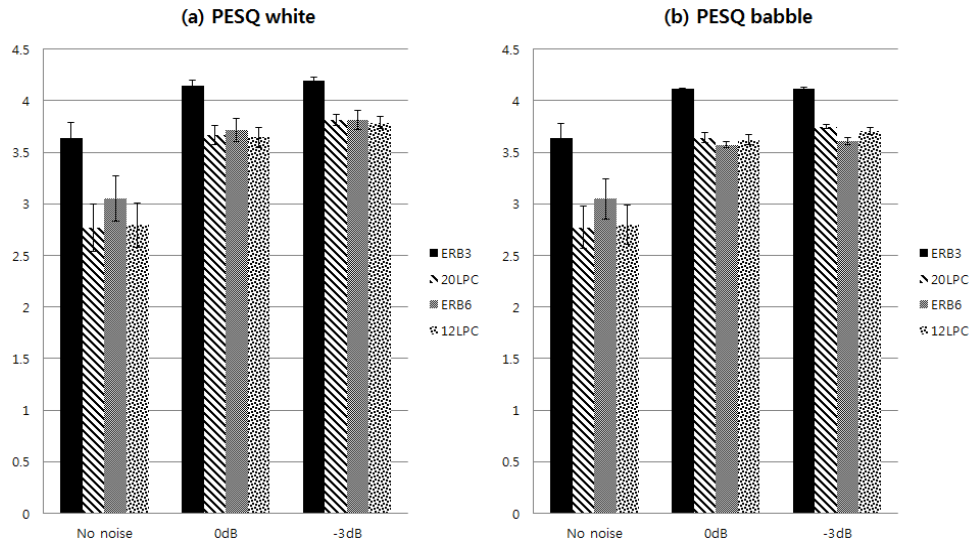


그림 7 PESQ 실험 결과: (a) 백색(white) 잡음, (b) 배블(babble) 잡음
 Fig. 7 PESQ experiment result for the white (a) and babble (b) noise

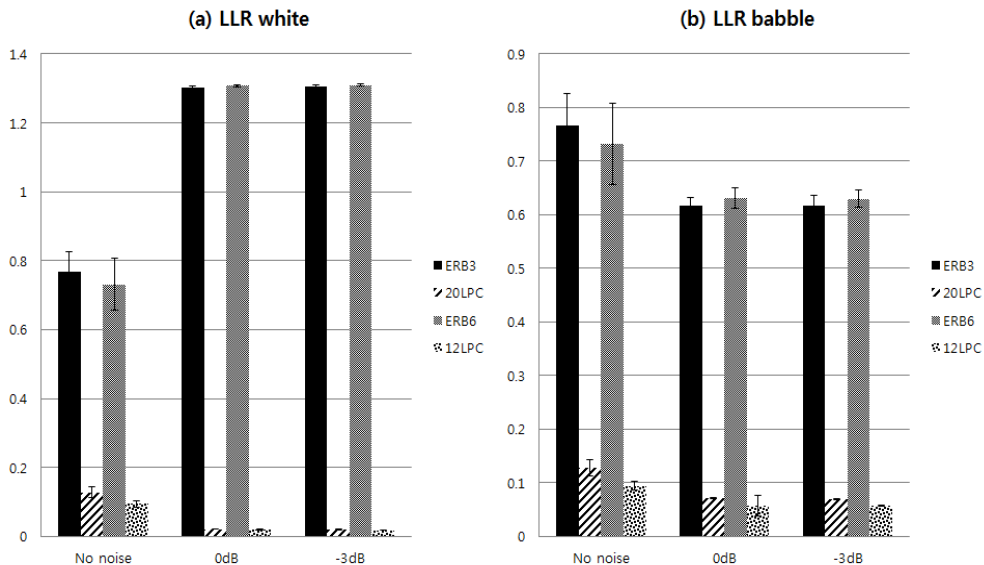


그림 8 LLR 실험 결과: (a) 백색(white) 잡음, (b) 배블(babble) 잡음
 Fig. 8 LLR experiment result for the white (a) and babble (b) noise

먼저 잡음이 없는 상황에서 평균 PESQ값을 살펴보면 ERB의 경우 ERB*3의 값이 ERB*6의 값보다 높게 나왔음을 알 수 있다. 이는 청각 필터가 넓어질수록 음성의 품질이 떨어짐을 의미한다. 따라서 청각 필터의 ERB를 이용한 알고리즘은 난청인의 넓어진 청각 필터를 표현 할 수 있다고 말할 수 있다. 이는 잡음상황에서도 마찬가지이다. 두 알고리즘을 비교하는 관점에서 보면 LPC를 이용한 알고리즘의 경우 모든 상황에서 PESQ값이 ERB를 이용한 알고리즘보다 낮게 나온 것을 볼 수 있다. 이는 LPC알고리즘이 스펙트럼을 뭉칠 때 선형적인 특성을 따른 것에 기인한 결과로 보여진다. 앞에서 설명한 청각 필터의 특성상 청각필터는 비선형적인 특성을 가지고 있는데 ERB를 이용한 알고리즘의 경우 이러한 청각필터의 특성을 잘 따랐기 때문에 PESQ

값이 높게 나왔지만, LPC의 경우는 그렇지 않기 때문에 상대적으로 PESQ값이 낮게 나온 것이다. 이는 난청인의 주파수 선택도 둔화 현상을 표현하기 위해서는 청각필터와 그것의 ERB를 이용한 알고리즘이 더 나은 결과를 가져온다는 것을 의미하며 이는 아래의 LLR의 결과에서도 마찬가지로 나타난다.

5.2 LLR 평가 실험 결과

표 3은 백색(white) 잡음에서의 실험 결과의 평균 LLR 수치이고 표 4는 배블(babble) 잡음에서의 실험결과의 평균 LLR 수치이다. 그림 8은 2음절 한국어 단어 36개의 평균 LLR 값을 나타낸 그림이다. 각각의 단어들은 잡음이 없는

상황, 0dB SNR인 상황, -3dB SNR인 상황으로 나누어지며 ERB*3, 20차수 LPC, ERB*6, 12차수 LPC 알고리즘을 통과 시킨 후 노이즈가 없는 상황에는 원음과 비교를 하고 노이즈 상황에서는 노이즈와 원음을 섞은 음과의 비교를 통해 평균 LLR 값을 계산하였다. 그림 8의 (a)는 백색(white)잡음에서의 상황이며 (b)는 배블(babble)잡음에서의 상황이다.

표 3 백색(white) 잡음에서의 평균 LLR

Table 3 Average LLR score in the white noise

| | ERB3 | 20LPC | ERB6 | 12LPC |
|----------|--------|--------|--------|--------|
| No noise | 0.7668 | 0.1281 | 0.7313 | 0.0946 |
| 0dB | 1.3019 | 0.0219 | 1.3074 | 0.0195 |
| -3dB | 1.3049 | 0.0198 | 1.3090 | 0.0180 |

표 4 배블(babble) 잡음에서의 평균 LLR

Table 4 Average LLR score in the babble noise

| | ERB3 | 20LPC | ERB6 | 12LPC |
|----------|--------|--------|--------|--------|
| No noise | 0.7668 | 0.1281 | 0.7313 | 0.0946 |
| 0dB | 0.6160 | 0.0711 | 0.6309 | 0.0579 |
| -3dB | 0.6176 | 0.0702 | 0.6297 | 0.0573 |

LLR값은 0에 가까울수록 왜곡이 적음을 의미한다. 그림을 살펴보면 PESQ값과는 반대의 결과가 나온 것을 볼 수 있다. ERB를 이용한 알고리즘의 경우 높은 PESQ값 즉, 음질이 좋음에도 불구하고 원음과의 왜곡이 심하다는 결과를 보이고 있고 LPC의 경우 낮은 PESQ값에도 불구하고 원음과의 왜곡의 거의 없다는 결과를 보이고 있다. 이는 LLR이 단순히 선형적으로 원음과 비교음의 차이를 계산하기 때문에 생기는 결과이다. 반면에 PESQ의 경우 인간의 청각특성을 반영하여 원음과 비교음의 인지 차이를 계산하므로 LLR과는 반대의 결과가 나온 것이다. 따라서 LLR의 결과역시 난청인의 주파수 선택도 둔화현상을 표현하기 위해서는 비선형적인 청각필터의 특성을 이용한 ERB알고리즘이 더 나은 결과를 가져온다는 것을 의미한다.

6. 결 론

난청을 가진 사람은 사회생활에 여러 가지 불편을 겪는다. 이러한 난청의 종류 중에서도 전음성 난청보다는 감각신경성 난청이 더 중요한데 이것은 전음성 난청은 비교적 쉽게 치료가 되기 때문이다. 감각신경성 난청의 대표적인 현상으로 주파수 선택도 둔화 현상이 있는데 이것은 난청인의 어음인지에 있어서 매우 중요한 현상임에도 불구하고 그 연구는 미약한 실정이다. 따라서 본 논문에서는 난청인의 주파수 선택도 둔화현상이 음질에 미치는 객관적 영향을 알아보기 위하여 선형적 수학적 모델링 기법인 LPC를 이용한 알고리즘과 청각 생리학에 기초를 둔 비선형적 수학적 모델링 방법인 청각필터와 그것의 ERB를 이용하는 알고리즘을 비교하고 분석해 보았다. 두 가지 알고리즘을 비교하고 분석하기 위해서 객관적인 음성 평가 도구가 필요한데 본 논

문에서는 음성의 품질을 테스트 하는데 있어서 전 세계적인 표준인 PESQ와 음성의 왜곡을 평가할 수 있는 도구로서 LLR을 이용하였다. 실험에는 36개의 한국어 2음절 음성을 사용하였는데 이들 음성은 각각 잡음이 없는 상황과 잡음비가 0dB인 상황 그리고 잡음비가 -3dB인 상황을 준비하였고 사용된 잡음은 백색(white)잡음과 배블(babble)잡음이다.

적용 알고리즘의 경우 LPC알고리즘은 차수가 12인 알고리즘과 20인 알고리즘을 사용하였고 청각 필터의 ERB를 사용한 알고리즘은 ERB의 3배와 ERB의 6배를 적용시키는 알고리즘을 사용하여 음성 신호 처리를 하였다. 처리된 음성 신호는 PESQ와 LLR을 통해서 비교를 하였다. 그 결과 PESQ에서 비선형적인 청각필터의 특성을 이용한 ERB 알고리즘의 경우 난청인의 넓어진 청각 필터를 표현하기에 적합한 결과가 나왔지만 LPC의 경우 선형적으로 스펙트럼을 뭉개는 알고리즘의 한계로 인하여 난청인의 넓어진 청각 필터효과를 제대로 표현하지 못함이 밝혀졌다. 또한 LLR의 경우 있어서도 LLR의 선형적인 계산 특성상 낮은 PESQ값을 보임에도 불구하고 왜곡이 적게 나오는 LPC알고리즘이 청각필터의 ERB를 이용한 알고리즘보다 더 좋지 않음을 알 수 있었다. 따라서 차후에 난청인의 난청 현상을 모델링하는데 있어서 감각신경성 난청의 대표적인 현상인 넓어진 청각 필터를 표현하기 위해서는 LPC알고리즘을 사용한 방법보다는 청각 필터의 ERB를 이용한 알고리즘을 사용하는 것이 난청인의 난청 현상을 더 잘 모델링 할 수 있을 것이다. 나아가 본 논문의 결과를 이용하여 후후 난청인의 난청현상을 모델링하는 난청시뮬레이터를 개발 한다면 보청기 성능 평가 등 난청 연구에 큰 도움이 될 수 있으리라 사료된다.

감사의 글

본 연구는 지식경제부 바이오 의료기기 전략기술 개발사업의 지원(과제번호: 10031764)에 의하여 이루어졌음.

참 고 문 헌

- [1] Brian C. J. Moore and Brian R. Glasberg, "Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech," *Acoustical Society of America*, vol. 94, no. 4, 1993, pp. 2050-2062.
- [2] 주상익, 강현덕, 송영록 and 이상민, "난청인의 주파수 선택도와 비대칭적 청각 필터를 고려한 난청 시뮬레이터 개발에 관한 연구", *전기학회논문지*, vol. 59, no. 4, pp. 831-840, 2010.
- [3] Thomas Baer and Brian C. J. Moore, "Effects of spectral smearing on the intelligibility of sentences in noise," *Acoustical Society of America*, vol.94, pp. 1229-1241, 1993.
- [4] Glasberg, B.R., and Brian C. J. Moore, "Auditory filter shapes in subjects with unilateral and bilateral cochlea impairments". *Acoust Society of America*,

vol. 79, no. 4, pp. 1020-33, 1986.

- [5] R. Munkong, B.-H. Juang, "Auditory perception and cognition", IEEE Signal Processing Magazine, vol. 25, no. 3, pp. 98-117, 2008.
- [6] Gelfand, S. A. "Hearing: an introduction to psychological and physiological acoustics", (4th ed.). New York, Marcel Dekker, ISBN 0-585-26606-9, 2004.
- [7] Moore, B. C. J. "Cochlear hearing loss", London, Whurr Publishers Ltd, ISBN 0-585-12256-3, 1998.
- [8] Glasberg, B.R., and Brian C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," Hearing Research., vol.47, pp. 103-118 1990.
- [9] A. M. Kondo, "Digital Speech, coding for low bit rate communication system Second Edition", West Sussex, John Wiley & sons Ltd, ISBN 0-470-87007-9(HB), 2004.
- [10] ITU-T, "Perceptual Evaluation of Speech Quality(PESQ), an Objective Method for End-to-end Speech Quality Assessment of Narrowband Telephone Network and Speech Codecs", ITU-T Rec. P.862 Feb. 2001.
- [11] Beerends, J. G., Hekstra, A. P., Rix, A. W. and Hollier, M. P. "Perceptual Evaluation of Speech Quality (PESQ), the new ITU standard for end-to-end speech quality assessment, part II - psychoacoustic model", Journal of the Audio Engineering Society, vol. 50, no. 10, pp. 765-778, October 2002.
- [12] Philipos C. Loizou, "Speech Enhancement", CRC Press, ISBN 978-0-8493-5032-0, 2007.

저 자 소 개



안 홍 섭 (安 洪 燮)

2010년 인하대학교 전자공학과 졸업,
2010년~현재 인하대학교 전자공학과 석사과정
Tel : 070-8258-1549
Fax : 032-868-3654



박 규 석 (朴 奎 錫)

2010년 인하대학교 전자공학과 졸업,
2010년~현재 인하대학교 전자공학과 석사과정
Tel : 070-8258-1549
Fax : 032-868-3654



전 유 용 (全 庾 鏞)

2006년 전북대학교 생체정보공학부 졸업,
2009년 인하대학교 전자공학과 석사 학위,
2009년~현재 인하대학교 전자공학과 박사과정
Tel : 070-8258-1549
Fax : 032-868-3654



송 영 록

2001년 인천대학교 정보통신공학부 학부 졸업,
2003년 인천대학교 정보통신공학과 석사 졸업,
2009년 인천대학교 정보통신공학과 박사 졸업,
2009년~현재 인하대학교 정보전자공동연구소 연구교수
Tel : 070-8258-1549
Fax : 032-868-3654



이 상 민 (李 相 旻)

1987년 인하대학교 전자공학과 학사 졸업,
1989년 인하대학교 전자공학과 석사 졸업,
2000년 인하대학교 전자공학과 박사 졸업.
1989년 1월~1994년 7월 LG이노텍 선임 연구원, 1995년 1월~2002년 3월 삼성종합 기술원 책임 연구원, 2002년 4월~2005년 2월 한양대학교 의공학교실 연구교수, 2005년 3월~2006년 8월 전북대학교 생체정보공학부 조교수, 2006년 9월~현재 인하대학교 전자전기공학부 부교수
Tel : 032-860-7420
Fax : 032-868-3654
E-mail : sanglee@inha.ac.kr