신호 대 마스킹 비 개선을 통한 다채널 보청 알고리즘

A Novel Multi-Channel Hearing Aid Algorithm with SMR(signal-to-masking ratio) Improvement

김 헌 중*, 홍 민 철*, 차 형 태*

(Hun-Joong Kim*, Min-Cheol Hong*, Hyung-Tai Cha*)

*숭실 대학교 전자공학과

(접수일자: 2000년 10월 11일; 수정일자: 2000년 11월 10일; 채택일자: 2000년 11월 10일)

본 논문에서는 심리음향 (psychoacoustics)을 이용한 다채널 동적 음향 압축 알고리즘을 통해 감음신경성 난청 (sensorinural hearing loss)을 보정하여 정상적인 지각 환경을 제공하는 알고리즘을 제안한다. 제안된 알고리즘은 라우드니스 조정 함수 (loudness scaling function)를 통해 난청자에게 적합한 라우드니스 레벨을 보상 시켜 주고, 난청자가 인식하게 되는 신호에 대한 마스킹 특성을 해석하여, 지각 영역 (perceptual domain)에서 각각의 주파수 성분에 대한 음압 레벨과 마스킹 임계치 (masking threshold) 사이의 거리로 정의되어지는 SMR (signal-to-masking ratio)을 통해 주파수 대비 (spectral contrast)를 복원 시켜 줌으로써 정상적인 지각 환경을 제공하는 것이다.

핵심용어 : 감음 신경성 난청, 보청 알고리즘, 심리음향, 지각 신호 해석, 다채널 음향 압축, 라우드니스 보상

투고분야 : 음악음향 및 음향심리 분야 (8.7)

In this paper, we propose a novel hearing aid algorithm for sensorinural hearing loss restoration with multi-channel(band) dynamic range compression and psychoacoustics. In this way, we can present a normal perception condition to the impaired listener. The proposed algorithm make loudness scaling function achieve proper loudness level, and analysis masking property for the signal will be perceived to impaired listener, and then, restore normal spectral contrast using SMR(signal-to-masking ratio) defined by distance between the level of each frequency and masking threshold.

***Key words**: ensorinural hearing loss, Hearing aid algorithm, Psychoacoustics, Perceptual signal analysis, Multichannel dynamic range compression, Loudness restoration*

***Subject classification**: Musical acoustics and psychoacoustics* (8.7)

I. 서 론

감음신경성 난청 (sensorineural hearing loss)의 경우 특정 주파수에 대한 sensitivity change를 포함하고 있으므로 각각의 주파수 대역에 따라 불균일하게 가청 임계치가 상승하는 특성을 나타내는 청력 손실을 보상 시켜 주어야만 하는데, 다채널 (다중 밴드) 동적 음향 압축형 보청 장치 (multi-channel band) compression hearing instruments)는 난청자의 순음 청력도 (audio profile)에 따른 주파수 대역에 대해 라우드니스 특성을 보상시켜 줌으로써 개개인의 청력 손실정도에 훨씬 더 정확한 보정과 좋은 음성 인식 능력을 제공한다[1][2][3]. 그러나 이와 같이 많은 수의 채널 또는 대역의 정보를 이용하는 시스템의 경우, 채널 (대역) 수의 변화에 따라 음성 변별력에 다양한 특성을 나타내며, 각각의 주파수 대역에 대해 주파수 세기 변화 (frequency intensity variation)에 민감하게 반응하므로 주파수 대비 감소 (spectral contrast reduction) 문제가 발생하게 되어 청자의 주파수 selectivity를 감소시켜 오히려 음성

책임저자: 김헌중 (hjkim@engineer.soongsil.ac.kr)
156-743 서울시 동작구 상도동 1-1 숭실대학교 전자공학과
(전화: 826-9063; 팩스: 820-0711)

의 이해에 영향을 주는 결과를 초래한다[2][4]. 이것은 주변 주파수 성분들에 의한 마스킹 효과로 인해 주파수 변별력 (frequency resolving capacity)이 떨어진 것으로, 대부분의 경우 저주파 성분이 고주파 성분을 마스크할 때 발생한다[5][6]. 이러한 현상은 배경 잡음이 존재할 경우 더욱 심하게 나타난다[3]. 이러한 문제점을 보상하기 위해 음성 신호를 임계대역 (critical band)별로 분리하여 양쪽 귀에 나누어 제공함으로써 지각적인 측면에서 마스킹 효과를 줄이거나[6], 라우드니스 조정 과정 (loudness scaling procedure) 후에 저주파 영역에 대해 이득을 낮추거나, 낮은 레벨을 갖는 고주파 영역에 이득을 증가시키는 알고리즘을 적용하였다[5]. 또 주파수 대비 감소 문제를 해결하기 위해, 주파수 영역에서 신호의 최대값을 증가시키고 최소값을 감소시키는 방법으로 주파수 대비 (spectral contrast)를 개선시킬 수도 있겠으나[4], 보다 근본적인 문제 해결을 위해 심리음향적 요소 (psychoacoustical parameters)들을 이용한 알고리즘의 개발이 요구되어 진다.

본 논문에서는 이러한 난청자가가 느끼는 인지 신호 (perceived signal)에 대한 적절한 해석과 그에 따른 마스킹 특성 분석을 기반으로 SMR (signal-to-masking ratio)를 이용하여 감소된 주파수 대비를 복원함으로써 라우드니스 보정된 신호가 난청자에게 정상인과 같이 인식되어지도록 하는 알고리즘을 연구하였다.

II. 다채널 동적 음향 압축 알고리즘

2.1. 다채널 동적 음향 압축 알고리즘의 특징

다채널 동적 음향 압축 알고리즘은 가청 임계치 상승(threshold elevation), 라우드니스 조정 함수 (loudness scaling function), 라우드니스 불쾌 역치 (loudness discomfort levels)등을 각각의 분리된 대역을 통해 적용할 수 있으므로 개개인의 청력 손실 정도에 훨씬 더 정확한 보정을 제공할 수 있고, 많은 수의 채널을 적용함으로써 선형 또는 단일 채널 음향 압축형 보청 알고리즘과 같이 적은 수의 채널을 적용할 경우 한 채널 내에서 가청 영역이 크게 변화하여 발생할 수 있는 부정적인 효과를 최소화할 수 있어 훨씬 좋은 음성 인식률을 제공한다. 또한 좁은 대역내에 신호의 중요 주파수 성분이나 잡음 성분들을 한정시킬 수 있어, 낮은 S/N비를 갖는 대역의 경우 신호 성분만 존재할 때 이루어 졌던 것 보다 잡음이 첨가된 신호 성분들을 적게 증폭함으로써, 넓은 대역을 적용할 경우 발생할 수 있는 에너지가 큰 잡음 성분들에 의해 신호 성분들이 마스킹되어 버리는 것을 제어할 수 있게 되어, 다양한 음향 환경에서 긍정적인 결과를 얻을 수 있다. 그러나 이와 같이 많은 수의 채널을 갖는 시스템의 경우 주파수 대역에 걸쳐 주파수 세기 변이 (frequency intensity variation)에 민감하게 반응하므로, 주파수 대비 감소 문제가 발생하게 되어 청자의 주파수 selectivity를 감소시켜 오히려 음성 이해도를 감소시키는 결과를 초래하게 된다[1][2][4][7].

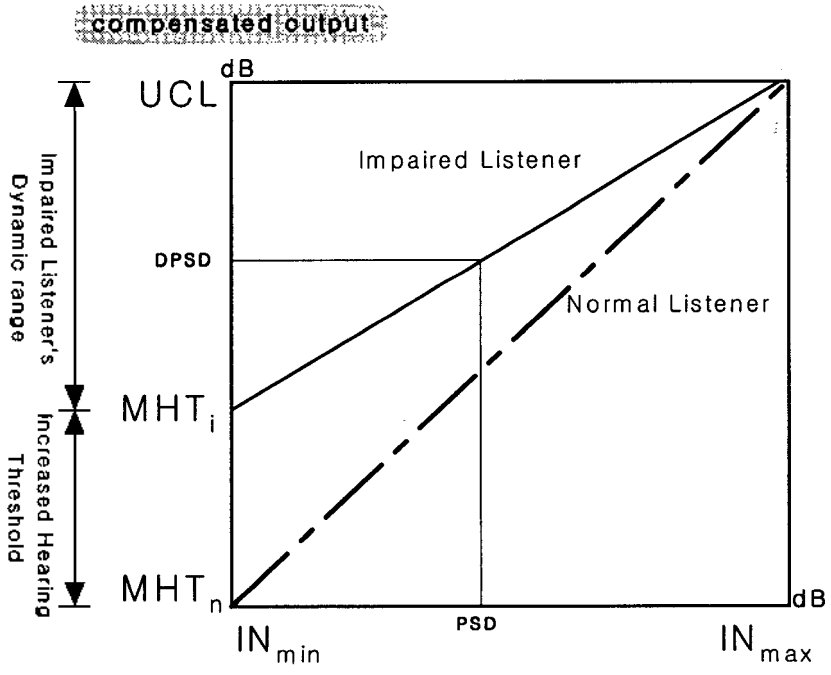

그림 1. 동적 가청 영역 보정 함수
Fig. 1. Dynamic range scaling function.

2.2. 가청 영역 보정

라우드니스 조정 (loudness scaling)의 목적은 MHT(minimum hearing threshold)가 상승한 난청자에게 지각적인 측면에서 정상적인 라우드니스 레벨을 복원시켜 주기 위함이다. 이때 이러한 라우드니스 조정 함수의 정의는 보청 프로세서의 특성에 의해 결정되어 지게 된다. 또한 라우드니스 조정 함수는 음성 신호와 일반 적인 환경 음향의 비선형 (nonlinear)적이고 시변 (time-varying)적인 특성을 고려하여 각각의 주파수 대역에 대해, 각기 다른 입력 레벨에 대해서 이득값을 조정하여 보정할 수 있도록 기울기를 변화하여 모든 입력 레벨에 대해서 지각적인 측면에서 적응적으로 라우드니스를 복원할 수 있어야 한다[5][7][8].

임의의 채널에 대한 동적 가청 영역 압축 (dynamic range compression)은 다음 그림 1에서와 같이 입력 신호의 에너지를 난청자의 감소된 가청 영역으로 전이 시키게 된다.

여기에서 라우드니스 조정 과정을 위한 파라미터인 입력 신호 에너지 범위 (IN_{min} ~ IN_{max})는 동적 가청 영역 보정을 위한 프로세서의 신호표현을 위한 bit resolution을 통해 결정되어 질 수 있으며[9], MHT_i은 정상 청력자들을 통한 통계적인 측정치를 통해 결정되어진 값 또는 ATH(Absolute threshold)로 표현될 수 있고, UCL (uncomfortable level)과 MHT_i는 난청자의 각각의 주파수 대역에 대한 순음 청력 측정을 통한 청력도 (audio profile)를 통해서 구해지게 된다. 또한 이때 보청 프로세서의 특성에 따라 MCL (most comfortable level)을 적용하거나, logarithmic 함수를 이용하기도 하기도 한다[1][5][8].

III. 제안된 심리음향 요소들을 통한 라우드니스 보정 알고리즘

다음 그림 2는 난청자의 지각적인 측면에서의 신호 해석과 마스킹 특성 개선을 통한 주파수 대비 회복을 위한 제안된 알고리즘에 대한 전체 구성 도이다.

이러한 심리음향을 이용한 라우드니스 보정과 주파수 대비 (spectral contrast) 개선을 통한 정상적인 지각 환경 복원 시스템의 구성은 다음과 같다.

3.1. 난청자의 신호 인지 (認知) 해석

입력 신호에 대한 난청자의 인지 신호 해석을 위한 난청자의 청력도는 임계 대역 (critical band)에 근거한 중심 주파수를 통해서 적용하고자 하는 채널 수를 적용한 난청자의 순음 청력 측정치를 이용하여 인터폴레이션을 취하여 추정하게 된다[8][10]. 다음 그림 3.는 여성음성의 모음 부분에 대해, 임계 대역에 근거한 중심 주파수를 갖는 8개의 대역을 통해 근사화된, 난청자의 청력도에 따른 입력 신호에 대한 난청자의 인지 신호를 나타낸다.

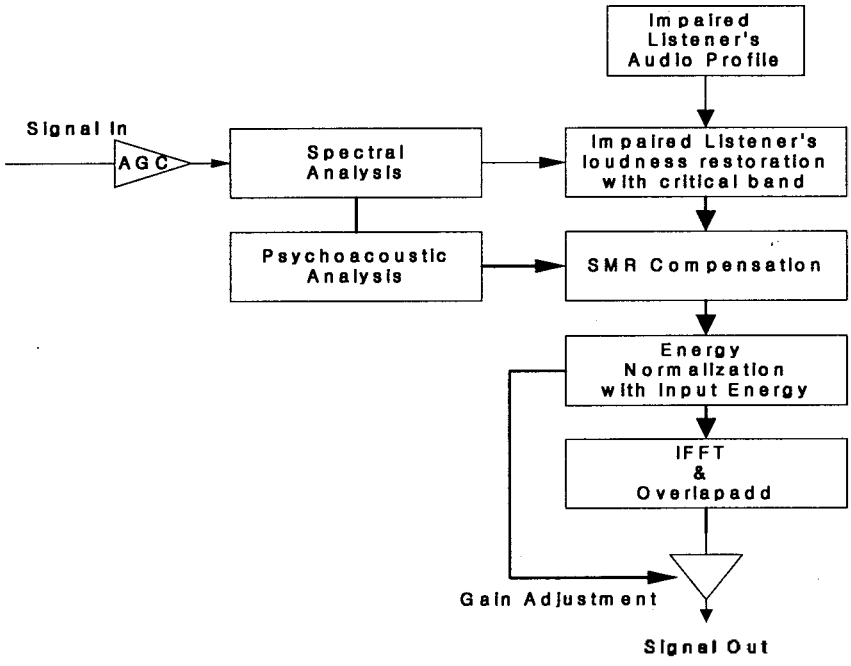

그림 2. 제안된 라우드니스 보정 알고리즘의 전체 구성도.
Fig. 2. A block diagram of the proposed loudness restoration algorithm.

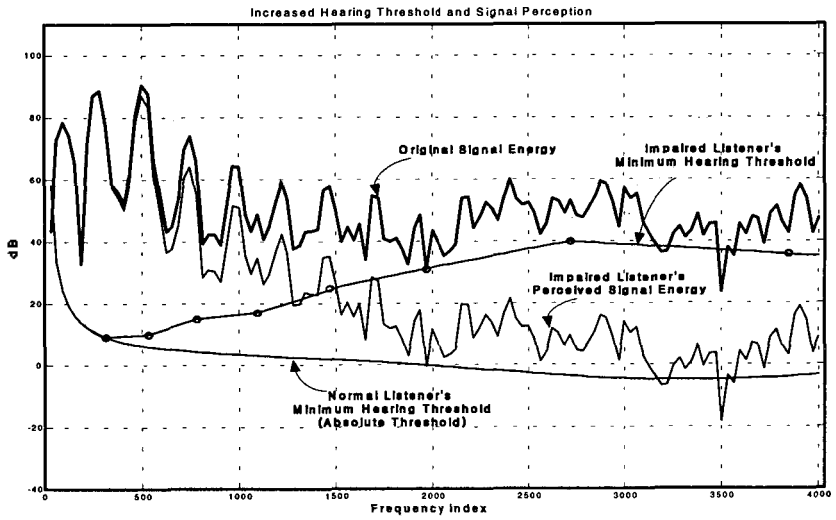

그림 3. 난청자의 청력도와 인지 신호 에너지
Fig. 3. Impaired listener's audio profile and perceived signal energy.

이 때 난청자의 MHT_i (minimum hearing threshold)는 정상인의 MHT_n에 비해 그 차이만큼 상승한 것으로 해석할 수 있다. 그러므로 입력 신호에 대한 난청자의 인지 신호 에너지는 정상인의 MHT_n에 비해 상승된 난청자의 가청 임계치 증가량만큼 감소시킴으로 해서 정상인의 마스킹 모델을 통해서 해석할 수 있다.

$$THD = MHT_i - MHT_n$$
$$IPSD = PSD - THD \quad \text{(in dB)} \tag{1}$$

단, 여기서 THD는 난청자와 정상인과의 최저 가청 한계치의 차이를, PSD는 입력 신호의 에너지 스펙트럼을, IPSD는 난청자의 입력 신호에 대한 인지 신호의 에너지 스펙트럼을 나타낸다.

3.2. 마스킹 특성 해석

사람의 기저막 (basilar membrane)에서의 주파수 분석을 나타내는 지각 영역에서의 마스킹 특성을 해석하기 위한 절차는 다음과 같다[11][12][13].

- Identification of Tonal and Noise Maskers.
- Decimation and Reorganization of Maskers.
- Calculation of Individual Masking Thresholds.
- Calculation of Global Masking Threshold.

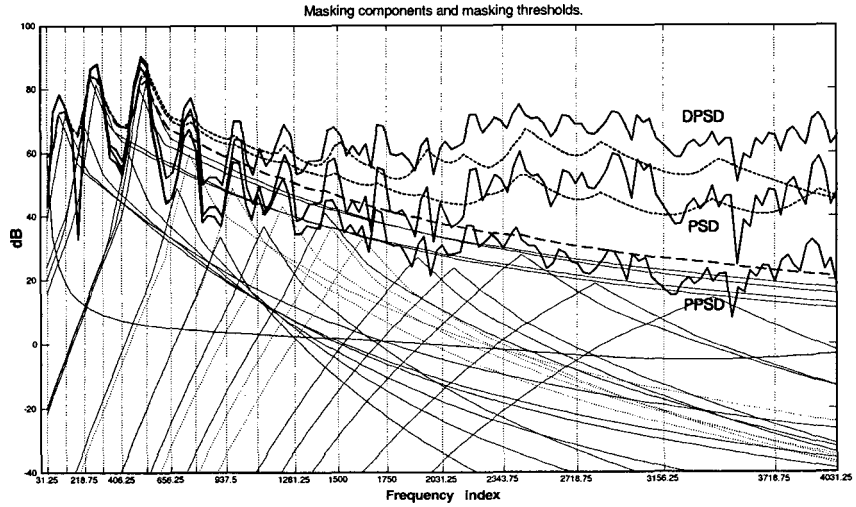

그림 4. 가청 영역 보정 처리된 신호
Fig. 4. The processed signals energy with dynamic range compensation.

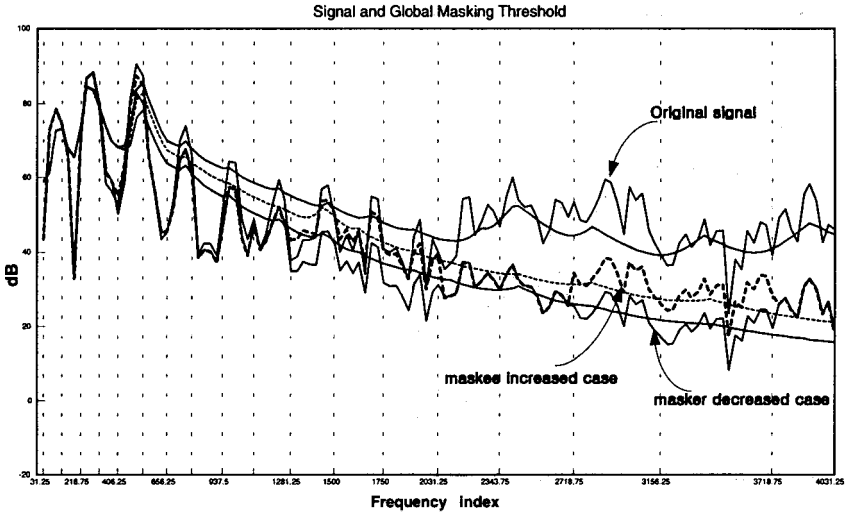

그림 5. Post-emphasis를 취한 경우에 대한 마스킹 임계치 비교
Fig. 5. Masking threshold comparison with the post-emphasized case.

그림 4는 청자에게 지각적으로 정상적인 라우드니스를 복원 시켜주기 위해 난청자의 청력도에 근거해서 각기 다른 주파수 대역에 대해서 입력 신호의 에너지 레벨에 따라, 지각적인 측면에서 적응적으로 라우드니스를 복원한 결과와 마스킹 임계치이다. 그림에서 PSD는 입력 신호 에너지를, DPSD는 실제 라우드니스 조정 과정을 거쳐 청자에게 전달되어지는 보정된 신호 에너지를, PPSD는 이러한 DPSD에 대한 청자가 느끼는 인지 신호의 에너지를 나타낸다.

$$PPSD = DPSD - PSD + IPSD \quad (\text{in } dB) \qquad (2)$$

이때 PSD와 PPSD의 차이는 각 주파수 대역에 대해 정상인과 난청자의 가청 영역이 그 차이만큼 전이된 것을 나타내고 있다. 또한 그림에서 살펴볼 수 있듯이, 각 주파수 대역에 따른 서로 다른 가청 범위의 영향으로 인해 청자에게 인식되어지는 신호의 에너지가 원신호의 에너지에 비해 변형된 주파수 파형으로 인한 주파수 대비 감소와 마스킹 특성을 갖게 되어 많은 부분의 주파수 성분들이 마스킹되어 음성 변별력에 제한을 가하고 있다.

이러한 변형된 주파수 파형으로 인한 주파수 대비 감소와 원치 않는 마스킹 영향에 의해 제약을 받고있는 음성 변별력 개선을 위해 다채널 동적 음향 압축 알고리즘의 적용 후에 주 (主) 마스커 (major masker)가 대부분 존재하게 되는 저주파 영역의 이득을 낮추어 주거나 낮은 레벨을 갖는 중간 또는 고주파 영역에 이득을 증가시키는 post-emphasis과정 등을 적용할 수 있다.

그림 5는 이와 같이 post-emphasis를 취한 경우에 대한 마스킹 임계치 비교이다. 이때 대부분의 주 (主) 마스커가 존재하게 되는 저주파 영역의 임계대역의 이득을 감소시킨 경우, 이득을 감소시킬 대역은 전역 마스킹 임계치 (global masking threshold)에 의해 마스킹되어 버린 대역에 대해서 영향을 많이 미치는 순음 성분과 잡음 성분의 마스커가 존재하는 임계 대역을 선택하고, 그러한 마스킹 곡선과 전역 마스킹 임계치의 차이값의 평균치만큼 감소시킨 것이다. 여기서는 Bark Index 5 ~ 6의 임계 대역에 대해 5dB 감소시킨 것이다. 또한 마스킹 임계치에 의해 마스킹 되어버린 마스키 (maskee)가 존재하는 임계 대역들에 대해서는 마스킹 임계치와 마스키의 차이의 평균치만큼 임계치 이상의 값을 갖도록 이득을 준 것이다. 이 경우 Bark Index 11 ~ 14의 임계 대역과 Bark Index 16 ~ 18의 임계 대역에 대해 각각 8.6 dB와 9.2 dB의 이득을 증가시킨 경우이다.

이와 같이 가청 영역 보정 후에 이러한 음성 인지도를 향상시키기 위한 post-emphasis를 적용하는 방법은 주 (主) 마스커의 영향에 의해 마스킹 되었던 주파수 성분들이 임계치 밖으로 나옴으로써, 약간의 개선점을 제공해 줄 수 있으나, 시변 (time-varying) 환경 하에서 정확한 이득의 수치를 조정하기 어렵고 오히려 변형된 주파수 파형으로 인해 원치 않는 부정적인 효과를 발생시킬 수가 있다. 또한 이득을 준 대역의 근접 영역에서는 이러한 대역에 의한 역방향 마스킹 (backward masking)과 순방향 마스킹 (forward masking)의 영향으로 마스킹 특성이 바뀌는 결과를 초래하고, 라우드니스 조정 과정을 통해 발생한 감소된 주파수 대비를 보상해주어야 하는 문제가 여전히 존재하게 된다.

3.3. SMR (signal-to-masking ratio) 보상

이와 같은 마스킹 특성 개선과 주파수 대비 감소

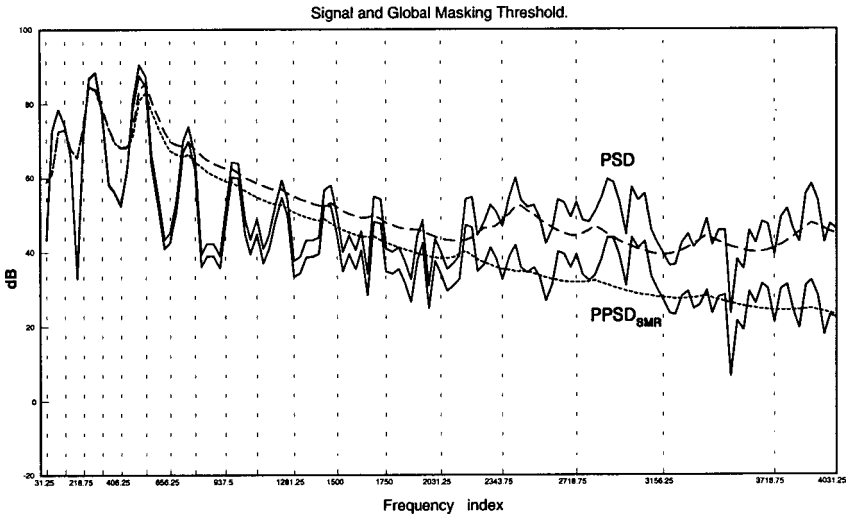

그림 6. SMR 보상된 신호와 원신호의 마스킹 임계치 비교
Fig. 6. Masking threshold comparison with the SMR compensated signal and the original signal.

문제점 등은 주파수 영역에서는 이것을 어떻게 보상할지 모호해지지만, 이것을 사람의 청각인지 특성을 이용한 지각 영역에서 각각의 주파수 성분에 대한 에너지 레벨과 마스킹 임계치 사이의 거리(distance)로 정의되어지는 SMR(signal-to-masking ratio)을 통해 개선시킬 수 있다. 다음 그림 6은 이러한 신호 대 마스킹 비를 이용하여 원신호의 마스킹 특성과 주파수 대비를 갖도록 라우드니스 보정된 신호에 대해 SMR 보상을 취한 신호 ($DPSD_{SMR}$)에 대한 인지 신호 ($PPSD_{SMR}$)와 입력 신호 (PSD)의 마스킹 임계치의 비교이다.

$$DPSD_{SMR} = DPSD + (SMR_o - SMR_l)$$
$$PPSD_{SMR} = PPSD + (SMR_o - SMR_l) \quad (\text{in } dB) \qquad (3)$$

where, SMR_o : original input signal's SMR
SMR_l : loudness compensated signal's SMR

그림에서 볼 수 있듯이 전체적으로 라우드니스 조정된 신호의 에너지 스펙트럼 파형이 원신호의 SMR에 맞게 재구성이 이루어져 원신호와 같은 마스킹 특성과 주파수 대비를 복원하여 원신호 같이 인식할 수 있게 되었다.

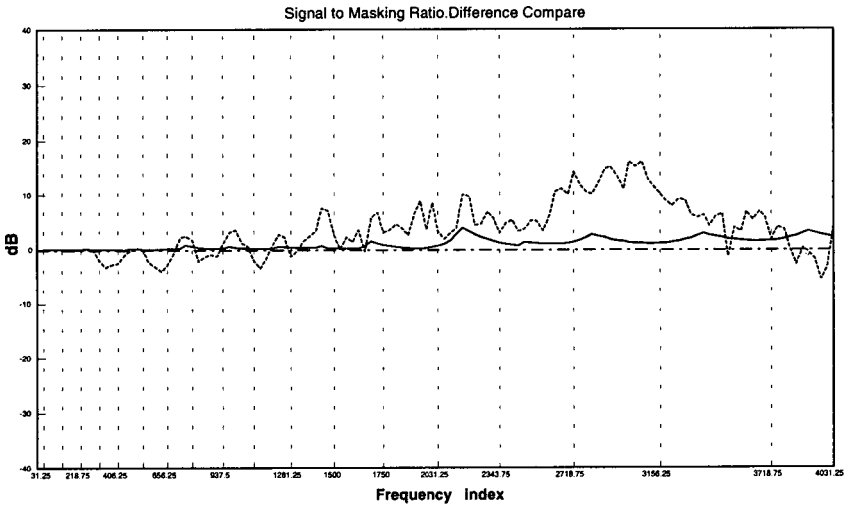

그림 7. SMR 보상된 신호와 라우드니스 보정된 신호에 대한 원신호와의 SMR 차등값 비교
Fig. 7. SMR difference comparison with the original signal's SMR for the SMR compensated signal and the loudness scaled signal.

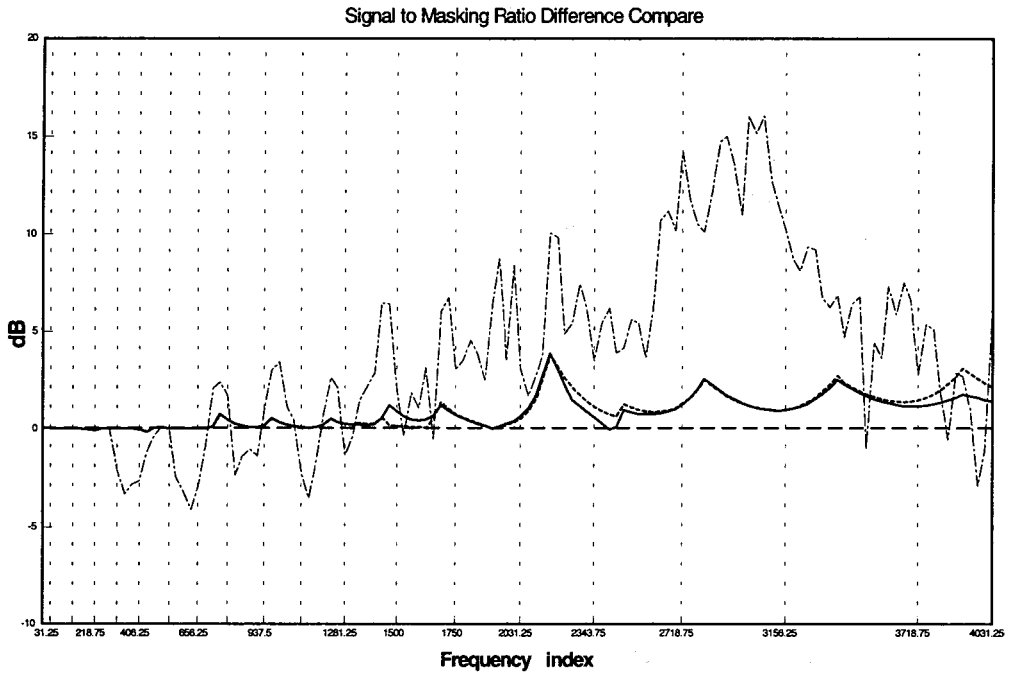

그림 8. 원신호의 순음 성분 인덱스와 잡음 성분 인덱스를 사용한 경우의 SMR 차등값
Fig. 8. SMR difference of the SMR compensated signal with original signal's tonality and non-tonality indices.

그림 7은 원신호와 SNR보상된 신호의 SNR 차이를 비교한 것 (실선)이다. 그림에서 볼 수 있듯이 라우드니스 보정된 신호에 대한 것 (점선)에 비해 차등 값의 전체적인 변이 편차가 낮고, 전 대역에 걸쳐 거의 일정한 값을 유지해 SMR 보상된 신호가 원신호와 같은 주파수 대비와 마스킹 특성을 유지할 수 있게 됨으로써 원신호와 같이 지각할 수 있게 된다. 그러나 본 예에서 볼 수 있듯이 마스킹 정도가 컸던 부분, 이 경우 고주파 영역으로 갈수록 SMR의 차이가 커지게 된다. 이것은 라우드니스 조정 함수를 통해 라우드니스 보정된 신호의 순음 성분과 잡음 성분들이 SMR비에 근거해서 마스킹 특성과 주파수 대비를 보상 시켜줌으로 해서 원신호와 같은 순음 성분과 잡음 성분 인덱스들을 복원함으로 해서 이러한 복원된 마스커들의 영향에 의해 전체적인 마스킹 임계치가 바뀌어지기 때문이다. 그러므로 SMR, 해석 시 원신호의 마스커 인덱스들을 사용함으로써 적절한 이득 보상이 이루어지도록 해야 한다.

그림 8은 SMR 보상 때에 원래의 마스커 인덱스들을 사용한 결과 (실선)로써, 그림에서 볼 수 있듯이 임계 대역 12, 14, 18에서 적절한 마스킹 특성을 회복함으로써 평균 SMR 차등값이 낮아지게 된다. 이러한 순음 성분과 잡음 성분의 인덱스에 따른 SMR 차등값의 차이는 그 대역의 가청 범위가 클수록, 고 주파영역으로 갈수록 다른 대역들의 마스킹 영향으로 인해 그 차이가 더욱 커지게 된다.

IV. 실험 결과

제안된 알고리즘의 성능 평가를 위한 실험은 난청 시뮬레이션 (hearing impairment simulation)을 통해 2년 이상 음성 또는 오디오 신호처리 분야에 종사한 정상 청력을 가진 전문인 10명을 대상으로 어음 명료도 검사 (speech discrimination test)를 통해 수행되었다. 이때 제안된 알고리즘의 평가를 위해 난청자를 통한 주관적인 평가가 수반되어져야 하나, 이러한 난청자를 통한 검증에는 난청자 개개인의 주관적인 요소들, 즉, 나이, 성별, 학력수준 등 많은 변수들이 작용하게 되어 객관적인 결과를 얻는데 많은 어려움과 시간이 요구되어 지게 된다. 이러한 제약점들의 극복을 위해 난청자의 상승된 가청 한계치를 이용하여 정상인을 통한 난청 시뮬레이션 환경을 통해 긍정적인 결과를 얻을 수 있다[14]. 본 논문에서는 난청자가 지각하게 되는 인지신호 해석을 통해 정상인을 이용한 난청 시뮬레이션 환경을 다음 그림 9와 같이 구성 하였다.

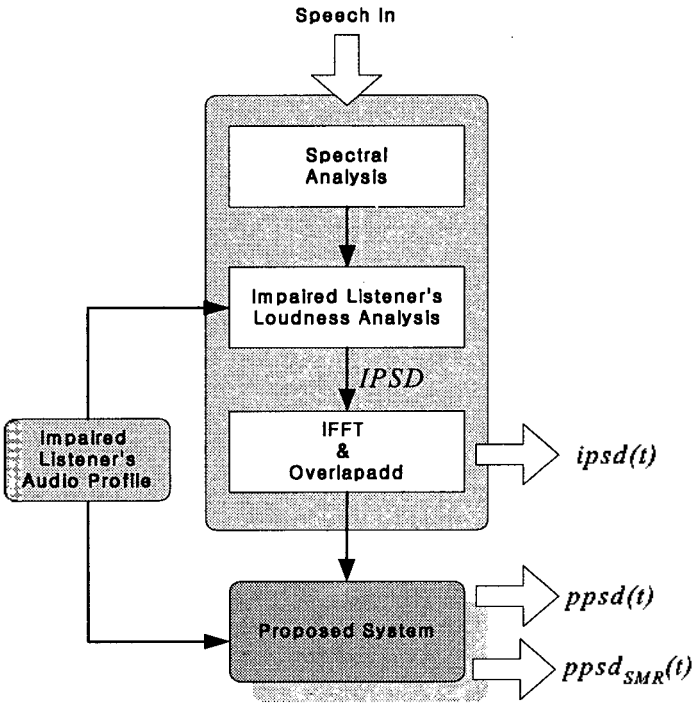

그림 9. 난청 시뮬레이션을 위한 구성도
Fig. 9. A block diagram for the hearing impairment simulation.

실험은 20대에서 40대의 정상인 남녀의 단음절 어음 (monosyllabic word)과 단문장을 통해 난청자가 각각의 상황에서 인식하게 되는 신호인 *ipsd*(청각 장애인의 인지 신호), *ppsd*(라우드니스 보정된 신호에 대한 인지 신호), $ppsd_{SMR}$ (SMR 보상된 신호에 대한 인지 신호)를 통해 실험 참가자의 어음 변별력을 측정 하였다. 이때 이러한 인지 신호 해석을 위한 주파수 분석은 256-point hanning window와 FFT를 통해 이루어 졌으며, overlap-add 방식을 적용하여, 8kHz로 샘플링 하고 16bit resolution을 갖는 입력 데이터를 128 샘플마다 16ms 분량으로 처리하게 된다. 이때 난청 보정을 위해 이득 조정된 신호의 경우 전체적인 평균 에너지가 상승하게 되어 IFFT를 통해 시간 영역으로 변환할 경우 시스템의 신호 표현을 위한 bit resolution의 제한으로 인해 시간 영역의 신호에 대해 클리핑 (clipping)이 발생하게 된다. 이러한 클리핑으로 인한 intermodulation의 발생과 같은 원치 않는 부산물등에 의한 시간영역에서의 마스킹 영향을 피하기 위해[9], 입력 신호의 최대값은 16bit resolution으로 표현될 경우 최대값인 96dB를 갖도록 하고, 출력 값은 입력 신호의 평균 에너지를 통해 정규화하게 된다. 또한 각 피실험자에게 제공되어지는 신호의 레벨은 각각의 청자의 MCL (most comfortable level)에 맞추어 조정하여 제공하였다.

다음 표 1은 실험에 적용된 난청자의 청력도이다.

표 1. 적용된 난청자의 청력도와 주파수 대역
Table 1. The applied impaired listener's audio profile and frequency bands.

Bark Index	Frequency Range	Band Index	Center Freq.	MHTI
1	0 ~ 93.75	....	....	
2	93.75 ~ 187.5			
3	187.5 ~ 281.25	1	281.25	9.2
4	281.25 ~ 406.25			
5	406.25 ~ 500.0	2	500	10
6	500.0 ~ 625.0			
7	625.0 ~ 750.0	3	750	15
8	750.0 ~ 906.25			
9	906.25 ~ 1062.5	4	1062.5	17
10	1062.5 ~1250.0			
11	1250.0 ~ 1437.5	5	1437.5	25
12	1437.5 ~ 1687.5			
13	1687.5 ~ 1937.5	6	1937.5	32
14	1937.5 ~ 2281.25			
15	2281.25 ~ 2687.5	7	2687.5	40
16	2687.5 ~ 3187.5			
17	3187.5 ~ 3812.5	8	3812.5	35
18	3812.5 ~ 4000.0			

SDT (speech discrimination test) 점수는 피실험자의 응답에 대한 정답율을 통해 얻어지게 된다. 다음 표 2는 실험에 적용된 유음어 단음절 어음 리스트

이다. 이때 각각의 어음은 여성과 남성의 목소리로 녹음되어 무작위로 쌍을 이루며, 각각의 어음 쌍은 피실험자에 따라 무작위 순으로 제공되어 지게 된다.

표 2. 어음 변별력 실험에 사용된 유음어 어음 쌍
Table 2. A word list for the speech discrimination test.

힘 - 합	논 - 눈	맛 - 멋
솔 - 술	국 - 독	솜 - 숀
닭 - 갓	감 - 땀	향 - 흉
법 - 벌	짐 - 김	통 - 뽕
삼 - 섬	싹 - 쑥	독 - 뚝
복 - 목	갈 - 칼	낮 - 높
망 - 명	알 - 일	새 - 재
꽃 - 굿	금 - 곰	발 - 볼
실 - 살	붓 - 북	풀 - 볼
왕 - 용	굴 - 글	쥐 - 죄

표 3. 어음 변별력 실험 결과
Table 3. The speech discrimination test result.

Subject	sex	age	$l_{psd}(t)$ SDT score	$l_{ppd}(t)$ SDT score	$p_{psdsum}(t)$ SDT score
1	M	42	60	76	92
2	M	41	52	73	87
3	M	30	56	76	90
4	M	30	54	74	86
5	M	30	48	67	84
6	M	27	50	66	88
7	M	26	53	72	90
8	F	24	57	76	89
9	F	25	58	76	88
10	F	22	60	78	89

표 3의 실험 결과에서 볼 수 있듯이, 다채널 음향 압축형 알고리즘을 적용하여 난청자에게 적합한 라우드니스 레벨을 보상 시켜준 후에, SMR을 통한 주파수 대비보상을 통해 어음 변별력에 많은 개선을 보이는 것을 알 수 있다. 실험 결과 다채널을 이용한 라우드니스 보정 결과에 비해 평균적으로 14.9%가 개선되었으며, 이것은 난청자가 이러한 라우드니스 보정을 통해 평균적으로 16.7%의 어음 변별력이 개선되어짐을 감안해 볼 때, 결론적으로 16.7+14.9=31.6%의 개선 효과를 볼 수 있게 됨을 나타내고 있다.

V. 결 론

본 논문에서는 심리 음향적 요소들을 통한 정상적인 지각 환경을 제공하는 알고리즘을 제안한다. 제안된 알고리즘은 라우드니스 조정 함수를 통해 난청자에게 적합한 라우드니스 레벨을 보상 시켜주고, 청자가 인식하게 되는 신호에 대한 마스킹 특성를 해석하고, 지각 영역에서 각각의 주파수 성분에 대한 에너지 레벨과 마스킹 임계치 사이의 거리로 정의되어지는 신호 대 마스킹 비 (SMR)를 통해 정상적인 주파수 대비를 회복시켜주는 것이다. 정상인을 통한 유음어 어음 쌍을 이용한 어음 변별력 실험 결과, 다채널 음향 압축형 알고리즘을 이용한 라우드니스 보정의 경우보다 약 14.9%의 개선점을 나타내었다.

이와 같이 SMR을 통한 보정 알고리즘을 적용할 경우, 채널 수의 증가에 따른 원치 않는 부정적인 효과를 최소화 하여 좁은 대역을 통해 정확한 보정을 제공할 있으며, 정상적인 마스킹 특성의 회복으로 인해 다양한 음향 환경 및 오디오 환경에서도 정상적인 지각 특성을 회복할 수 있다. 그러나 현재 마스킹 특성 해석을 위한 알고리즘은 너무 복잡하므로, 보청 장치에 적합한 효율적인 알고리즘에 대한 연구가 계속 진행되어야 겠다.

참 고 문 헌

1. Loizou, P. "Introduction to cochlear implants, " IEEE Engineering in Medicine and Biology Magazine, Vol. 18, No.1, pp. 32-42, 1999.
2. E. William Yund, Krista M. Buckles, "Multichannel compression hearing aids : Effect of numberof channels on speech discrimination," J. Acoust. Soc. Amer., Vol. 97, No.2, May. 1995.
3. Mark Kahrs, Karlheinz Brandenburg, *Application of digital signal processing to audio and acoustics*, Kluwer Academic Publishers, 1998.
4. O. Poroy and P. Loizou, "Minimum spectral contrast for vowel identification in CIS-type processors," Conference on Implantable Auditory Prostheses, Asilomar, Monterey, CA. 1999.
5. Holube I, Wesselkamp M, "Application of psychoacoustics in hearing aid development," Acustica, Vol. 85, No.5, pp. 642-645, Sep. 1999.
6. D.Chaudhari, P.Pandey, "Dichotic presentation of speech signal with critical band filtering for improving speech perception," Proc. IEEE ICASSP'98, Washington, USA, May. 1998.
7. Paul H. Stypulkowski, "Fitting Strategies for Multiple-Memory Programmable Hearing Instruments," American Journal of Audiology, Vol. 2, No.2, July. 1993.
8. Young-cheol Park, Hae-su Lee, Yong-chun Kim, Young-beom Jang, Sang-il Park, "High performance digital hearing aid processor with psychoacoustic loudness correction," Proc. IEEE ICCE'97, New York, USA, Jun 1997.
9. Jeremy Agnew, "Audible circuit noise in hearing aid

amplifiers," J. Acoust. Soc. Amer., Vol. 102, No.5, Nov. 1997.
10. N.Black, M.Lydon, N.Waterman, M.Powderley, "Programmable Audiogram Matching Using a Frequency Sampling Filter Implemented On the texas tms320c30," Proc. IEEE ICASSP'92, California, USA, Mar. 1992.
11. Ted Paiter, "A Review of Algorithms for Perceptual Coding of Digital Audio Signals," Proc. IEEE, 13th Conference on Digital Signal Processing., Vol. 1, pp. 179-208, Jun. 1997.
12. Bosse Lincoln, "An Experimental High Fidelity Perceptual Audio Coder Project in MUS420 Win 97," http://ccrma-www.stanford.edu/~bosse/proj/proj.html
13. ISO/IEC JTC1/SC29/WG11 MPEG, IS11172-3 "Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s, Part 3: Audio," 1992.
14. Dong Wook Kim, Young Cheol Park, Won Ky Kim, Seon Joon Park, Won Doh, Seung Woo Shin, Dae Hee Youn, "Simulation of Hearing Impairment with Sensorineural Hearing Loss," Proc. IEEE EMBS, Chicago, USA, Oct. 1997.

▲ 김 헌 중 (Hun-Joong Kim)

1997년 : 관동대학교 전자 통신공학과 (공학사)
1999년 : 숭실대학교 전자공학과 (공학석사)
1999년~현재 : 숭실대학교 전자공학과 박사과정 재학
※ 주관심분야 : 오디오 및 음성신호 처리, 통신 신호처리, ASIC 설계

▲ 홍 민 철 (Min-Cheol Hong)

1988년 : 연세대학교 전 자공학과 졸업(공학사)
1990년 : 연세대학교 전자공학과 졸업(공학석사)
1990~91년 : LG 정보통신 연구소 연구원
1997년 : Northwestern University(공학박사)
1997~98년 : Northwestern University Post Doctoral Research Fellow
1998~00년 : LG 전자 연구소 선임 연구원
2000~현재 : 숭실대학교 전임강사
※주관심분야 : Image Restoration, Non-linear Video Processing/Filtering Blind Deconvolution, Video Compression

▲ 차 형 태 (Hyung-Tai Cha)

1985년 : 숭실대(공학사)
1988년 : The University of Pittsburgh (공학석사)
1993년 : The University of Pittsburgh (공학박사)
1993~96년 : 삼성전자 신호처리 연구소 선임연구원
1996~98 : 숭실대학교 전임 강사
1998~현재 : 숭실대학교 조교수
※주관심분야 : Audio / Video Coding, Morphology, ASIC 설계