

잡음 형상화에 의한 오디오 워터마크 설계

이진걸[†]

요약

심리음향모델을 기반으로 주파수영역에서 잡음을 형상화하여 오디오 신호에 부가하였을 때 부가된 잡음이 지각되지 않는 방법을 제안하였다. 신호의 마스킹 문턱값으로부터 지각되지 않는 잡음의 준위를 구하는 것은 심리음향모델에서 확산함수와 관련된 디컨버루션을 수반하는데 난제(ill-conditioned problem)로 알려져 있다. 본 논문에서는 최적화 기법을 적용하여 잡음의 여기준위를 신호의 마스킹 준위에 일치시킴으로써 신호에 부가된 잡음이 청각적으로 지각되지 않는 한도 내에서 최대한의 잡음준위가 되도록 형상화하는 방법을 제시하고 실험적으로 그 타당성을 증명하였다.

Design of Audio Watermarks by Noise Shaping

Jingeol Lee[†]

ABSTRACT

A psychoacoustic model based noise shaping method is proposed. The method shapes the noise in the frequency domain such that its presence with a host signal will not be perceptually noticeable. The derivation of imperceptible noise levels from the masking thresholds of the signal involves deconvolution associated with the spreading function in the psychoacoustic model. It has been known as an ill-conditioned problem. In this paper, a constrained optimization is applied such that the noise excitation level conforms to the masking thresholds of the signal. Thus, the noises embedded in the signal will not be perceived by human ear, and its performance is demonstrated experimentally.

Key words: Psychoacoustics(심리음향학), Masking(마스킹), Noise Shaping(잡음 형상화)

1. 서론

잡음형상화 기법은 다양한 형태로 사용되어 왔다. 예를 들어 코드 여기선형예측(CELP) 코딩에서 지각가중필터(perceptual weighting filter)는 포먼트(formant)에 해당하는 주파수 영역보다 지각적으로 더 예민한 포먼트 널(null)의 주파수 영역에 있는 잡음 성분이 더 감쇄되도록 원래 신호와 합성신호의 간의 차(잡음)를 형상화한다[1]. 다른 예는 오디오 코딩에서 심리음향모델 및 TNS(temporal noise shaping)

를 이용하여 양자화 잡음이 지각되지 않도록 형상화된다[2]. 또한, 다른 예는 지각할 수 있는 왜곡이 최소화되도록 오디오 신호에 저작권정보와 같은 정보를 부가하는 오디오 워터마킹에서 찾을 수 있다[3].

이러한 응용들은 청각계의 마스킹 성질을 이용한다. 마스킹 성질은 순수한 정현파와 같은 강한 신호에 의해 양자화 잡음과 같은 약한 신호가 지각되지 않는 심리음향적인 현상이다. 마스킹하는 신호를 마스커(masker)라고 하고 마스킹되는 신호를 타겟(target)이라고 한다. 마스킹되기 위해서는 타겟의 여기준위(excitation level)가 청각영역에서 마스킹의 마스킹 문턱값(masking threshold)보다 낮아야 한다[4,5]. 신호의 마스킹 문턱값으로부터 지각되지 않는 잡음의 준위를 구하는 것은 심리음향모델에서 확산함수(spreading function)와 관련된 디컨버루션

※ 교신저자(Corresponding Author): 이진걸, 주소: 대전광역시 서구 도마2동 439-6(302-735), 전화: 042)520-5707, FAX: 042)520-5773, E-mail: jingeol@mail.paichai.ac.kr
접수일: 2005년 4월 6일, 완료일: 2005년 7월 7일

[†] 정회원, 배재대학교 전자공학과 부교수

을 수반하는데 난제(ill-conditioned problem)로 알려져 있다.

심리음향 모델을 기반으로 하는 오디오 잡음형상화는 신호의 마스크 문턱값을 필터 계수로 하여 입력 잡음을 형상화시키는 방법이 있으나 이는 형상화된 잡음의 여기준위를 고려하지 않기 때문에 심리음향학의 관점에서 정확한 방법이 아니며 또한, 필터의 이득설정을 명확히 제시하지 않는다[3]. 본 논문에서는 최적화기법을 이용하여 잡음의 여기준위를 신호의 마스크 문턱치에 일치시킴으로써 신호에 부가된 형상화된 잡음이 지각되지 않는 한도 내에서 최대한의 잡음준위가 되도록 잡음을 형상화시키는 방법을 제안하였다. 심리음향학에서 타겟의 여기준위가 청각영역에서 마스크의 마스크 문턱값보다 낮으면 타겟이 마스크되는 것은 잘 알려져 있다[5-7].

본 논문의 구성은 2장에 심리음향모델을 요약하였으며, 3장에서는 최적화기법에 의한 잡음형상화 방법을 제안하였다. 4장에서는 제안된 방법의 적용성을 오디오 신호를 사용한 실험을 통하여 보였으며 결론은 5장에 서술하였다.

2. 심리음향 모델

2.1 마스크

마스크는 인간청각계의 지각적인 성질로서 강한 신호가 주파수, 시간적으로 가까이에 위치한 약한 신호를 지각하지 못하게 한다. 이들을 각각 동시(simultaneous), 비동시(nonsimultaneous) 마스크라고 한다[8]. 동시 마스크는 마스크와 타겟의 주파수, 전력, 톤 유사도(tone-like), 잡음 유사도(noise-like)에 영향을 받는다. 비동시 마스크는 전 마스크(pre-masking)과 후 마스크(post-masking)이 있다. 전 마스크는 강한 신호 이전에 5~20msec 동안 존재하는 약한 신호가 지각되지 않는 성질이다. 후 마스크는 강한 신호 이후에 50~200msec 동안 존재하는 약한 신호가 지각되지 않는 성질이다. 후 마스크와 관련된 마스크 문턱값은 지수적으로 감쇄하는 함수로 모델링된다[6,9]. 본 논문에서는 MPEG 오디오 심리음향 모델 1(계층 2)에 채택된 동시 마스크 모델을 사용하였다.

인간청각계는 주파수에 의존하는 제한된 분해능을 가진다. 이것은 500Hz이하의 주파수에서는 100

Hz, 500 Hz 이상의 주파수에서는 1/3 옥타브(octave)의 임계(critical) 대역폭으로 표시될 수 있다. 이 임계대역은 임계대역을, Z 라고 불리는 주파수의 측정으로도 사용될 수 있으며 단위는 바크(Bark)이다. 표 1은 임계대역의 컷오프(cutoff) 주파수를 나타내었다[4,8].

표 1. 임계대역 주파수 [4]

Rate (Barks)	Δf (Hz)	f_l (Hz)	f_u (Hz)
1	-	-	100
2	100	100	200
3	100	200	300
4	100	300	400
5	110	400	510
6	120	510	630
7	140	630	770
8	150	770	920
9	160	920	1080
10	190	1080	1270
11	210	1270	1480
12	240	1480	1720
13	280	1720	2000
14	320	2000	2320
15	380	2320	2700
16	450	2700	3150
17	550	3150	3700
18	700	3700	4400
19	900	4400	5300
20	1100	5300	6400
21	1300	6400	7700
22	1800	7700	9500
23	2500	9500	12000
24	3500	12000	15500
25	6550	15500	19500
26	6550	19500	24600

2.2 MPEG 오디오의 심리음향 모델1

MPEG 오디오의 심리음향 모델1은 32개 부대역(subband)의 신호대 마스크비를 제공하여 각 부대역에 대한 비트 할당에 사용된다. 본 논문은 MPEG 오디오의 동시 마스크 모델을 사용한 잡음형상화를 논하므로 마스크(MPEG오디오 심리음향 모델 1, 계층 2, 샘플링 주파수 44.1KHz)에 관련된 내용을 요약하였다[2].

(1) 시간-주파수 변환을 위한 FFT 계산

입력신호 $x(n)$ 에 한(Hann) 윈도우 $h(n)$ 를 적용한

후 1024point FFT를 취한다. 신호의 전력밀도 스펙트럼 $P(k)$ 은 다음과 같이 계산된다.

$$P(k) = 10 \log \left| \frac{1}{N} \sum_{n=0}^{N-1} h(n)x(n)e^{-j(2\pi/N)nk} \right|^2 \text{ dB}, \quad 0 \leq k \leq N/2 \quad (1)$$

여기서 $N=1024$ 이다.

(2) 절대 문턱값(absolute threshold)의 결정
주파수 영역에서 정의된 음압 준위인 절대 문턱값

$LT_q(k)$ 이 모델에 주어져 있다[2]. 인간은 특정 주파수의 음압이 이 문턱값을 넘게 되면 소리를 지각하게 된다.

(3) 오디오 신호의 톤(tonal) 성분과 톤이 아닌(non-tonal) 성분 구분

마스킹 성분의 톤 유사도가 마스킹 문턱값에 영향을 미치기 때문에 톤 성분과 톤이 아닌 성분을 구분하는 것이 필요하다. 이 절차는 국부 최대치를 정하고 임계대역 안에서 국부 최대치를 제외한 성분들의 세기를 계산함으로써 이루어진다. 다음의 세가지 동작이 톤 성분과 톤이 아닌 성분의 음압을 구하기 위해 수행된다.

a) 국부 최대치의 표시

다음의 조건을 만족할 경우 주파수 성분 $P(k)$ 는 국부 최대치로 표시된다.

$$P(k) \geq P(k-1) \quad \text{and} \quad P(k) \geq P(k+1) \quad (2)$$

b) 톤 성분의 추출 및 음압의 계산

다음의 조건이 만족되면 국부 최대는 톤 성분이 된다.

$$P(k) - P(k+j) \geq 7 \text{ dB} \quad (3)$$

여기서 j 는 다음에 따라 선택된다.

$j = -2, 2$	$2 < k < 63$
$j = -3, -2, 2, 3$	$63 \leq k < 127$
$j = -6, \dots, -2, 2, \dots, 6$	$127 \leq k < 255$
$j = -12, \dots, -2, 2, \dots, 12$	$255 \leq k \leq 500$

$P(k)$ 가 톤 성분이 되면 톤 성분의 음압은 다음과 같이 계산된다.

$$P_m(k) = 10 \log \left\{ 10^{\frac{P(k-1)}{10}} + 10^{\frac{P(k)}{10}} + 10^{\frac{P(k+1)}{10}} \right\} \text{ dB} \quad (4)$$

조사된 주파수 영역내의 모든 주파수 성분은 $-\infty$ dB가 된다.

c) 톤이 아닌 성분의 추출 및 전력계산

26개 각각의 임계대역 내에서 톤 성분이 0이 되고 남은 주파수 성분의 전력을 합하여 그 임계대역에 해당하는 톤이 아닌 성분의 음압준위 $P_{nm}(k)$ 를 구한다.

(4) 관련된 마스커만 구하기 위한 마스커의 선별
이것은 전역 마스킹 문턱값의 계산에 사용되는 마스커의 수를 줄이기 위하여 사용되는 과정이다.

(5) 개별 마스킹 문턱값의 계산

수식 (1)에서와 같이 512개의 주파수영역 샘플 중, i 로 표시되는 일부 샘플에서만 전역 마스킹 문턱값 계산이 이루어진다. 이 샘플의 개수는 130개이다.

톤 성분과 톤이 아닌 성분의 개별 마스킹 문턱값은 다음과 같다:

$$\begin{aligned} LT_{tm}\{z(j), z(i)\} &= P_m\{z(j)\} + av_m\{z(j)\} + vf\{z(j), z(i)\} \text{ dB} \\ LT_{nm}\{z(j), z(i)\} &= P_m\{z(j)\} + av_{nm}\{z(j)\} + vf\{z(j), z(i)\} \text{ dB} \end{aligned} \quad (5)$$

여기서 아래 첨자 tm 은 톤 성분, nm 은 톤이 아닌 성분을 나타낸다. $LT\{z(j), z(i)\}$ 는 $z(j)$ 에 해당하는 주파수 샘플 j 에 있는 마스킹 성분에 의한 $z(i)$ 에 해당하는 주파수 샘플 i 에서의 개별 마스킹 문턱값이다. 전력 $P\{z(j)\}$ 는 $z(j)$ 에 해당하는 주파수 샘플 j 에서의 마스킹 성분의 음압준위이다. av 는 마스킹 인덱스(index)이고 vf 는 확산함수이다.

(6) 전역 마스킹 문턱값의 결정

주파수 샘플 i 에서 전역 마스킹 문턱값 $LT_g(i)$ 는 주파수 샘플 j 의 톤 성분과 톤이 아닌 성분 마스커에 해당하는 개별 마스킹 문턱값과 절대 문턱값 $LT_q(i)$ 를 합하여 구한다.

$$LT_g(i) = 10 \log \left(10^{LT_q(i)/10} + \sum_{j=1}^m 10^{LT_{tm}\{z(j), z(i)\}/10} + \sum_{j=1}^n 10^{LT_{nm}\{z(j), z(i)\}/10} \right) \quad (6)$$

여기서 m 은 톤 성분 마스커의 개수이고 n 은 톤이

아닌 성분 마스크의 개수이다.

3. 지각되지 않는 잡음준위의 유도

2.2절의 심리음향 모델에 따라 잡음의 여기준위는 잡음의 마스크 문턱값에서 마스크 인덱스와 절대 문턱값을 제거하여 구할 수 있다. 즉, 각각의 마스크의 여기준위는 마스크 준위보다 마스크 인덱스만큼 높다[7]. 따라서, 잡음의 여기준위는 (6)에서 마스크 인덱스와 절대 문턱값에 해당하는 항을 제거하여 구해진다.

$$ET_g(i) = 10 \log \left(\sum_{j=1}^m 10^{ET_m\{z(j), z(i)\}/10} + \sum_{j=1}^n 10^{ET_{nm}\{z(j), z(i)\}/10} \right) \quad (7)$$

여기서,

$$ET_m\{z(j), z(i)\} = X_m\{z(j)\} + vf\{z(j), z(i)\} \text{ dB}$$

$$ET_{nm}\{z(j), z(i)\} = X_{nm}\{z(j)\} + vf\{z(j), z(i)\} \text{ dB}$$

$ET_g(i)$ 는 전역 여기준위이고 $ET_m\{z(j), z(i)\}$ 와 $ET_{nm}\{z(j), z(i)\}$ 는 각각 톤 성분과 톤이 아닌 성분의 개별적인 여기준위이다.

잡음의 여기준위가 신호의 마스크 문턱값보다 낮으면 잡음은 지각되지 않는다. 이렇게 잡음을 형상화하기 위하여 각각의 임계대역에서 잡음의 음압을 선형 필터로 변경하는 방법이 제안되었다. 이 방법은 지각필터(Perceptual filter)와 같이 임계대역 별로 필터 계수값이 정해진다[9]. 2.2절에서와 같이 음압을 구하기 위해 각각의 임계대역 내의 주파수 성분의 전력이 더해지므로 필터 계수는 각각의 임계대역 내에서 일정한 값이 되어 형상화된 잡음은 다음과 같이 주어진다.

$$\hat{N}(k) = H(b)N(k), \quad k_{b_0} \leq k \leq k_{b_b}, \quad 0 \leq b \leq B-1 \quad (8)$$

여기서 k_{b_0} , k_{b_b} 는 각각 임계대역 b 의 하한, 상한경계(표 1)이고 B 는 임계대역의 개수이다. (8)은 각각의 임계대역의 주파수 범위 내에서 잡음의 주파수 성분 $N(k)$ 이 해당하는 필터 계수 $H(b)$ 에 의해 변경됨을 나타낸다. 임의의 주파수에서 잡음의 여기준위는 이웃한 톤 성분과 톤이 아닌 성분의 확산된 음압

준위를 합하여 구해지는 것을 고려하여 그 주파수에서의 잡음준위는 다음과 같이 이웃한 톤 성분과 톤이 아닌 성분의 크기를 조절하여 형상화될 수 있다.

$$\sum_{j=0}^{B-1} H^2\{b(j)\} 10^{ET\{b(j), z(i)\}/10} = 10^{ET_g(i)/10}, \quad ET_g(i) < LT_g(i)$$

$$= 10^{LT_g(i)/10}, \quad ET_g(i) \geq LT_g(i) \quad (9)$$

여기서 $ET\{b(j), z(i)\}$ 는 임계대역 j 에 있는 톤 성분과 톤이 아닌 성분의 개별적인 마스크로부터 주파수 i 에 확산된 음압준위이다. 필터계수는 (9)의 전력항과 함께 표시되도록 자승되었다. 주파수 i 에서 잡음의 여기준위가 신호의 마스크 문턱값보다 높은 경우, 여기준위가 마스크 문턱값에 일치되도록 조절되었다. 그렇지 않은 경우, 여기준위의 변화는 없다. (9)가 적용되는 주파수는 톤 성분과 톤이 아닌 성분이 있는 주파수로 하였다. 왜냐하면 톤 성분과 톤이 아닌 성분이 있는 주파수에서 잡음의 여기준위는 국부적으로 최대값이 되기 때문이다. 따라서, 이 주파수에서 잡음의 여기준위를 신호의 마스크 문턱값에 일치시키면 다른 주파수에서는 확산함수에 의해 잡음의 여기준위가 신호의 마스크 문턱값보다 낮아진다. (9)를 적용하면 $E \cdot h = m$ 형태의 선형 대수식이 된다. E 는 적용되는 주파수의 개수를 행의 수로 하고 임계대역의 개수를 열의 수로 하는 행렬이고 각 행의 원소들은 해당하는 임계대역에 있는 개별적인 마스크로부터 임의의 주파수로 확산된 음압준위이다. 벡터 h 는 필터계수의 자승 $H^2\{b(j)\}$ 으로 임계대역의 개수와 같은 개수의 원소를 가진다. 벡터 m 은 (9)에서와 같이 적용되는 주파수에서 상대적인 크기에 따라 여기준위이거나 마스크 문턱값으로 구성되고 주파수 개수와 같은 개수의 원소를 가진다. 미지수인 필터계수의 자승의 개수는 수식(주파수)의 개수보다 많거나 적을 수 있으며 SVD(singular value decomposition)를 이용한 방법에 따라 해를 구할 수 있다. 그러나, SVD를 이용한 방법은 잡음준위에 따라 필터계수의 자승으로 음의 값이 되는 경우가 있어 적절한 방법이 아님을 확인하였다. 이러한 문제를 극복하기 위해 잡음 형상화는 다음과 같이 최적화기법으로 수식화되었다.

$$\min_h \|E \cdot h - m\|_2, \quad 0 \leq h_0, h_1, \dots, h_{B-1} \leq 1 \quad (10)$$

(10)을 적용하면 형상화된 잡음의 여기준위는 일치시키고자 하는 마스킹 문턱값과 다소 다를 것으로 예상된다. 왜냐하면 확산함수는 개별적인 마스커의 음압에 따라 형태를 달리하기 때문이다. 높은 음압준위가 더 천천히 확산되기 때문에 형상화된 잡음의 준위는 일치시키고자 하는 마스킹 문턱값보다 다소 낮을 것으로 예상된다. 그러나, 다음 장에서 보일 일련의 실험으로부터 이러한 오차는 무시할 정도임을 확인하였다. (10)의 마스킹 문턱값에 일정한 값을 감함으로써 지각되지 않는 잡음의 준위에 여유(margin)를 줄 수 있다.

4. 실험결과

제안된 잡음형상화 방법은 실험을 통하여 형상화된 잡음의 여기준위가 신호의 마스킹 문턱치에 일치함과 쌍 비교(pair comparison)를 통해 원음과 제안한 방법에 의해 형상화된 잡음이 부가된 음이 청각적으로 구별되지 않음을 확인한다. 제안한 방법에 의해 형상화된 잡음의 여기준위가 신호의 마스킹 문턱치에 일치함을 확인하기 위하여 그림 1에서와 같은 피리연주 한 프레임(1024 샘플, 샘플링 주파수: 44.1KHz, 16비트 입력)의 신호에 대해, 크기의 범위가 -50에서 50(16비트 입력)인 유사 랜덤(pseudorandom) 입력 잡음을 형상화시켜 신호의 마스킹 문턱값과 입력 및 형상화된 잡음의 여기준위를 그림 2에 나타내었다. 그림 1에서 신호는 거의 주기적이므로 기본 주파수(약 400Hz)와 그것의 고조파 성분으로 구성되어 있

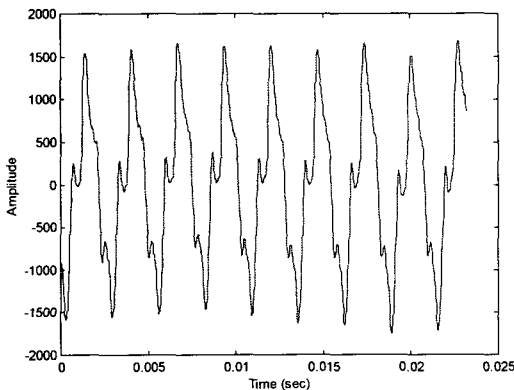


그림 1. 피리연주 파형

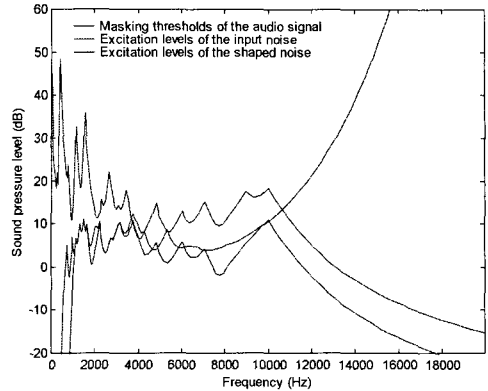


그림 2. 마스킹 문턱값과 여기준위

음을 알 수 있으며, 그림 2에서 신호의 마스킹 문턱값에서 국부 최대치는 기본 주파수와 그것의 고조파 성분에 해당하는 톤성분을 나타낸다. 그림 2에서 형상화된 잡음의 여기준위는 (8)에 의해 필터링된 잡음의 주파수 성분을 이용하여 계산하였다. 주파수 범위 4000~11,000 Hz에서 입력잡음의 여기준위는 신호의 마스킹 문턱값보다 높은 반면 같은 주파수 대역에서 형상화된 잡음의 여기준위의 국부최대치가 신호의 마스킹 문턱값에 거의 일치함을 알 수 있다. 이 일치에서 약간 벗어나는 오차는 개별적인 마스커의 음압에 따라 값을 달리하는 확산함수와 최적화기법에 기인한다. 이 실험을 통하여 제안한 방법으로 형상화된 잡음의 여기준위를 신호의 마스킹 문턱값에 일치시킬 수 있음을 보였다.

다음 실험은 유사 랜덤(pseudorandom) 입력잡음을 형상화시켜 전체 피리연주 신호에 부가함으로써 제안한 방법의 구현성을 보인다. 이 실험에서 분석 프레임은 이웃한 프레임과 50% 중첩되고 형상화된 잡음은 중첩 합(overlapped addition) 방법에 의해 구해진다. 중첩 합 방법에서 FFT 크기는 프레임의 크기와 같은 1024로 하였다. 그림 3은 입력 오디오 신호, 형상화된 잡음, 형상화된 잡음이 부가된 신호를 보인다.

그림 3에서 형상화된 잡음의 크기가 입력 오디오 신호의 크기에 따라 변화함을 보인다. 이것은 입력 오디오 신호의 크기에 따라 마스킹 문턱값이 변함에 따라 형상화된 잡음의 크기도 변하기 때문이다.

주관적인 음질시험을 통하여 형상화된 잡음이 부가된 음으로부터 왜곡이 지각되는지를 확인하기 위

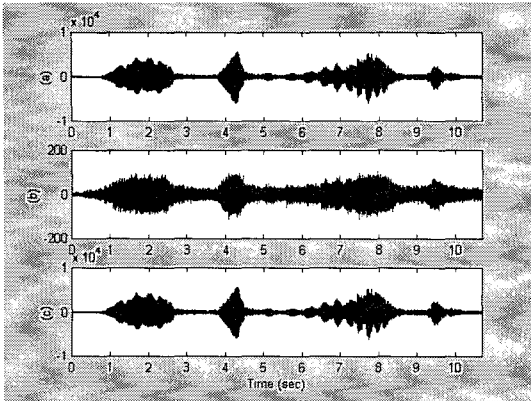


그림 3. 파형: (a) 입력 오디오 신호, (b) 형상화된 잡음, (c) 형상화된 잡음이 부가된 신호.

하여 쌍 비교실험을 하여 그 결과를 표 2에 나타내었다. 쌍 비교실험은 피리연주(10.8초)를 포함하여 심청가(12.1초), 비엔나 소년합창(11.4초), 베토벤 소나타 1번(9.8초), 차이코프스키 심포니 6번(8.3초)으로 구성된 특성이 다른 5개의 오디오 신호를 사용하였다. 쌍 비교실험에서 실험의 참가자들은 35명의 음악

을 전공하지 않는 대학생으로 구성하였으며 각각의 오디오 신호에 대해 원음과 형상화된 잡음이 부가된 음을 두 번 들려주고 음질이 우수한 것을 선택하도록 하였다. 이상의 쌍 비교실험은 원음과 형상화된 잡음이 부가된 음이 청각적으로 거의 구별되지 않음을 보여준다.

5. 결 론

오디오 신호에 부가된 잡음이 지각되지 않도록 심리음향 모델을 기반으로 하여 잡음을 형상화하는 방법을 제안하였다. 지각되지 않는 잡음의 준위를 최적화기법을 사용하여 형상화된 잡음의 여기준위를 오디오 신호의 마스킹 문턱값에 일치시켜 구하였다. 한 프레임의 오디오 신호를 이용하여 형상화된 잡음의 여기준위가 오디오 신호의 마스킹 문턱값에 일치됨을 보였으며 쌍 비교실험을 통하여 원음과 형상화된 잡음이 부가된 음이 청각적으로 거의 구별되지 않음을 보였다. 제안한 잡음형상화 방법으로 신호에 부가된 형상화된 잡음이 지각되지 않는 한도 내에서 최대

표 2. 쌍 비교

피리연주		
구분	선호 (%)	구분되지 않음(%)
원음	20	63
형상화된 잡음이 부가된 음	17	
심청가		
구분	선호 (%)	구분되지 않음(%)
원음	20	60
형상화된 잡음이 부가된 음	20	
비엔나 소년합창		
구분	선호 (%)	구분되지 않음(%)
원음	20	57
형상화된 잡음이 부가된 음	23	
베토벤 소나타 1번		
구분	선호 (%)	구분되지 않음(%)
원음	17	60
형상화된 잡음이 부가된 음	23	
차이코프스키 심포니 6번		
구분	선호 (%)	구분되지 않음(%)
원음	26	
형상화된 잡음이 부가된 음	20	54

한 크기의 잡음준위를 구할 수 있다. 한편, 오디오 워터마킹에서는 부가되는 워터마크의 비지각성 및 강인성이 요구되며 워터마크의 준위가 높을수록 강한 것으로 알려져 있다. 따라서, 제안한 방법은 부가되는 워터마크의 비지각성과 함께 예상되는 여러 가지 공격에 강인한 오디오 워터마킹에 적용이 기대된다.

참 고 문 헌

- [1] A. M. Kondoz, *Digital Speech Coding for Low Bit Rate Communications Systems*, John Wiley & Sons Ltd, New York, NY, 1994.
- [2] ISO/IEC 11172-3, "Information technology-Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s Part3: Audio."
- [3] Mitchell D. Swanson, Bin Zhu, Ahmed H. Tewfik, and Laurence Boney, "Robust Audio Watermarking using Perceptual Masking," *Signal Processing*, Vol. 66, pp. 337-355, 1998.
- [4] Raymond N. J. Veldhuis, "Bit Rates in Audio Source Coding," *IEEE Journal on Selected Area in Communications*, Vol. 10, No. 1, pp. 86-96, Jan. 1992.
- [5] M. R. Schroeder, B. S. Atal, and J. L. Hall, "Optimizing Digital Speech Coders by Exploiting Masking Properties of the Human Ear," *Journal of Acoustic Society of America*, Vol. 66, No. 6, pp. 1647-1652, Dec. 1979.
- [6] John G. Beerends and Jan A. Stemerdink, "A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation," *J. Audio Eng. Soc.*, Vol. 40, No. 12, pp. 963-978, Dec. 1992.
- [7] Eberhard Zwicker and U. Tilmann Zwicker, "Audio Engineering and Psychoacoustics: Matching Signals to the Final Receiver, the Human Auditory System," *J. Audio Eng. Soc.*, Vol. 39, No. 3, pp. 115-126, Mar. 1991.
- [8] Davis Pan, "A Tutorial on MPEG/Audio Compression," *IEEE Multimedia Magazine*, Vol. 2, No. 2, pp. 60-74, Jun. 1995.
- [9] Dionysis E. Tsoukalas, John Mourjopoulos, and George Kokkinakis, "Perceptual Filters for Audio Signal Enhancement," *J. Audio Eng. Soc.*, Vol. 45, No. 1/2, pp. 22-36, Jan./Feb. 1997.



이 진 길

1981년 2월 고려대학교 전자공학과 졸업(공학사)
 1985년 8월 고려대학교 대학원 전자공학과 졸업(공학석사)
 1994년 12월 미국 University of Florida 전기공학과 졸업(공학박사)

1982년 9월~1990년 12월 국방과학연구소, 연구원
 1995년 6월~1997년 2월 삼성전자주식회사, 수석연구원
 1997년 3월~현재 배재대학교 전자공학과, 부교수
 관심분야: 워터마킹, 음향학