크로스토크 제거를 위한 바크스케일 등가 보상 필터를 이용한 음질 향상에 관한 연구

김 학 진† · 김 순 협††

요 약

본 논문에서는 자유음장 모델을 기초로 5.1채널 스피커 시스템의 보상함수를 해석하고, 이를 심리음향에 기초한 바크 스케일(bark scale) 등가 보상 밴드패스 필터에 적용하여 주파수 제한적인 음질 보상에 대한 실험을 통해 5.1채널 스피커 시스템에서 음향 재현시 나타나는 크로스토크를 제거함으로써 보다 입체적인 음을 재현하는 방법을 제안하였다. 정면 채널과 서라운드 채널의 음원을 기존의 2채널 및 5.1채널 시스템에서 제시하는 음원과 비교를 위하여 음질 평가와 함께 청취자를 대상으로 주관적인 평가를 실시하였다. 실험 결과 ±30°의 정면 스피커를 배치한 돌비 표준 스피커 배치에서 38dB 이상의 분리도를 얻었으며, 전문가 그룹의 음의 확산감에 대한 주관 평가에서는 5점 평가를 기준으로 0.4~0.5점, 음의 깊이감 평가에서는 0.1~0.2점이 향상되었다.

A Study on the Sound Quality Improvement Using the Equal Compensation Filter in Bark-scale for the Cross-talk Cancellation

Hack-Jin Kim† · Soon-Hyub Kim††

ABSTRACT

This paper deals a method to deliver more realistic sound by cancelling the cross-talk which is inherent to the 5.1 channel speaker system. The acoustical model for cross-talk cancellation is the free field model. This model minimizes distortion of sound. I used the bark scale sound quality compensation which based on psycho-acoustic. For the surround channels, band-limited sound quality compensation is performed in the frequency domain. I also performed the sound quality assessment test on the traditional 2 channel stereo and 5.1 channel system. This test is performed in the test chamber which satisfies the ITU-R specifications. I uses the IACC(Inter-Aural Cross-Correlation) to determine the preferences of the amateur and the golden ear experts to asses the trans-aural filter. According to the result from the proposed method, I got more the 38dB separation rates with the Dolby standard speaker array. The results on the diffusion by the subjective test with the experts shows 0.4~0.5 point increased then before.

키워드 : 크로스토크 제거(Cross-talk Cancellation), 입체음향(Spatial Sound), 트랜스오럴 필터(Transaural Filter)

1. 서 론

본 논문은 음향재현에 있어 가장 효율적이고 대중화된 방법인 5.1채널 시스템에서 청취자의 정면에 위치한 좌우측 및 중앙 채널 간의 크로스토크와 좌측후면(Ls), 우측후면(Rs)의 서라운드 음에 있어서의 크로스토크를 제거하여 입체적인 음을 재현하는 방법을 제안하였다. 크로스토크 제거에 대한 연구는 주로 2채널 또는 3채널의 자유음장 모델을 이용하고 스피커와 청취자 간의 전달함수를 통해 크로스토크를 제거하는 연구가 진행되었다[1-3].

자유음장 모델은 음질의 왜곡을 최소화할 수 있는 기법이다. 5.1채널의 경우 일반적인 스테레오 음원과는 다른 재생 포맷을 가지고 있다. 이 경우 정면과 서라운드의 역할이 각기 다르다. 따라서 본 논문에서는 5.1채널에서 각 채널간의 크로스토크 제거를 위해 자유음장 모델에 바크 스케일 등가 보상필터를 추가하여 크로스토크를 제거하고자 하였다. 또한 각 스피커 간에 공급되는 소리의 주파수 범위와 돌비 서라운드 프로로직(Dolby Surround Pro-logic) 등의 저작물간의 호환성을 위해 주파수 제한 필터를 조합, 구성하여 각 재생물(시료)에 대해 실험하였다[4, 5]. 재생시 2채널에 대해 정면채널의 효과에 대해 크로스토크 제거시와 그렇지 않을 때, 그리고 기존의 5.1채널 재생과 비교하여 정면 및 서라운드음에 대해 평가하였다. 일련의 실험들은 ITU-R 권고안의 기준을 만족하는 청취실에서, 그리고 5등

※ 본 연구는 2003년도 광운대학교 교내학술연구비 지원에 의해 수행되었습니다.
† 정 회 원 : 명지전문대학 겸임교수
†† 정 회 원 : 광운대학교 정보과학기술대학 원장
논문접수 : 2003년 9월 8일, 심사완료 : 2004년 3월 30일

급의 선호도로 평가하는 평가방법을 이용하여 평가하였다 [6, 7].

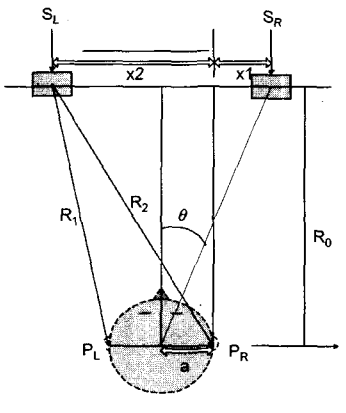
2장에서는 자유음장 모델을 이용한 전달함수의 해석에 대해서, 3장에서는 5.1채널 스피커 시스템에서 바크 스케일 등가 보상필터를 이용한 크로스토크 제거 방법에 대해서, 4장에서는 구현하고자 하는 5.1채널에서의 크로스토크 제거를 위한 트랜스오럴 필터 구현 프로그램에 대해서, 5장에서는 ITU-R 권고안에 따른 실험 및 평가에 대해서, 그리고 마지막 6장에서는 결론을 맺는다.

2. 자유음장모델을 이용한 전달함수 해석

본 논문에서도 고전적 자유음장을 선정하여 트랜스오럴 필터를 구현코자 하였으며, 기존 Hamada의 연구와는 달리 특정 스피커 위치에 독립적으로 동작하는 방안을 강구하였다.

Hamada 등 일부 학자들은 측정 머리전달함수를 이용한 크로스토크(cross-talk) 제거는 음색 왜곡과 음질 저하를 가져온다는 연구와 함께 이러한 부작용을 최소화하기 위해 고전적인 자유음장 모델을 선정하여 스테레오 다이폴(stereo dipole)을 연구 제안하였다[8-10]. 그러나 스테레오 다이폴은 특정위치의 특수한 스피커 배치에서만 재생하는 한계를 갖고 있다.

여기서 스피커로부터 방사되는 음원을 점음원으로 간주한다. 청취자는 항상 자신의 상체, 머리, 외이 등의 신체 구조에 따른 자신의 음향학적 음압 전달 특성에 따라 들리는 음을 가장 자연스러운 음으로 인식한다. 자유음장 모델은 양 귀 주위의 재현에 초점을 맞추고 있으며, 고막에서의 음압 재현은 아니다.



(그림 1) 청취자와 스피커에 대한 구성

(그림 1)은 스피커와 청취자의 기하학적 구성도이다. R_0는 음원에서 머리 중심까지의 거리이고, R_1과 R_2는 음원에 근접한 귀까지의 경로와 먼 쪽 귀와의 경로의 거리이다. θ는 머리의 중심점과 스피커와의 각도, 그리고 a는 청취자 머리의 반경이다. 여기서 직접경로 R_1과 교차경로 R_2의 계산은 다음과 같다.

$$R_1=\sqrt{(R_0\cdot\tan\theta-a)^2+R_0^2}$$
$$R_2=\sqrt{(R_0\cdot\tan\theta+a)^2+R_0^2} \quad (1)$$

근접 경로와 교차경로와의 음압비 g_c와 두 경로차이에 의한 음압 도달시간 τ_c는 다음과 같다.

$$g_c=C(R_2;f)/C(R_1;f)=R_1/R_2$$
$$\tau_c=(R_2-R_1)/C_0 \quad (2)$$

여기서 음압비 g_c는 항상 1 보다 작다.

3. 5.1채널 스피커 시스템에서의 크로스토크 제거 방법

3.1 트랜스오럴 필터의 적용

(그림 1)과 같이 구성된 좌·우측 스피커의 구동 음원 신호 $\{S_L, S_R\}$이라 할 때, 음압 전달함수로부터 양 귀의 직접 경로 음압은

$$P_{LL}(R_1;f)=S_L(f)\cdot e^{-jkR_1}/R_1$$
$$P_{RR}(R_1;f)=S_R(f)\cdot e^{-jkR_1}/R_1 \quad (3)$$

좌우 교차 성분인 크로스 토크 경로 음압은

$$P_{RL}(R_2;f)=S_L(f)\cdot e^{-jkR_2}/R_2$$
$$P_{LR}(R_2;f)=S_R(f)\cdot e^{-jkR_2}/R_2 \quad (4)$$

로 각각 기술된다. 따라서, 좌우 귀 음압 $\{P_L, P_R\}$은 식 (3)과 식 (4)의 음압 성분들의 합으로 기술된다.

$$\begin{bmatrix}P_L(f)\\P_R(f)\end{bmatrix}=\begin{bmatrix}\frac{e^{-jkR_1}}{R_1} & \frac{e^{-jkR_2}}{R_2}\\ \frac{e^{-jkR_2}}{R_2} & \frac{e^{-jkR_1}}{R_1}\end{bmatrix}\begin{bmatrix}S_L(f)\\S_R(f)\end{bmatrix} \quad (5)$$

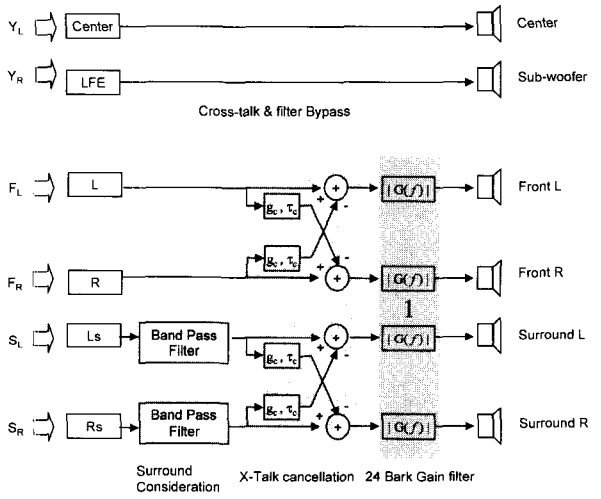
식 (5)는 좌, 우측 음원을 두 스피커에 구동할 때 양 귀 주변의 음압에 대한 모델이다. 이때 2채널 스피커 구동 신호는 다음과 같다.

$$S_L(f)=|G(f)|\cdot[P_L(f)-P_R(f)\cdot g_c e^{-j\omega\tau_c}]$$
$$S_R(f)=|G(f)|\cdot[P_R(f)-P_L(f)\cdot g_c e^{-j\omega\tau_c}] \quad (6)$$

크로스토크 제거 후 보상해 주는 필터 함수 $G(f)$는 다음과 같은 복소수 함수이다[2, 3].

$$G(f;R_0,\theta)=\frac{g_c\cdot e^{+j\omega\tau_c}}{1-g_c^2\cdot e^{-j2\omega\tau_c}} \quad (7)$$

따라서, 식 (7)의 $|G(f)|$는 좌우 스피커에 동일하게 인가하는 이득(gain) 값이다.우측 귀의 음압 $P_R=0$이라 가정하면, 좌측 스피커는 P_L에 비례하는 성분만을 구동하고 우측 스피커는 $(-g_c \cdot e^{-j\omega\tau})\times P_L$에 비례하는 신호를 구동하게 된다. 이때, 우측 구동 신호 $(-g_c \cdot e^{-j\omega\tau})\times P_L$는 좌측 구동 스피커 구동신호 P_L에 좌우측 귀의 상대적 음압비(Relative ILD) 만큼의 이득율 g_c ($g_c<1$) 적용과 양 귀 도달 시간차이 ITD) 만큼 시간지연인 τ_c 적용, 그리고 위상 180°변환을 각각 적용하는 과정을 의미한다. 이러한 세 가지 과정(ILD, ITD, 위상반전)을 거쳐 얻어진 우측 신호와 좌측 신호를 동시에 스피커에 인가할 때, 스피커의 교차 전달 경로에 의한 교차 전달 음압 성분이 제거된다. 이러한 연산 과정을 도식화하면 (그림 2)의 교차성분 제거부와 같이 표시된다.



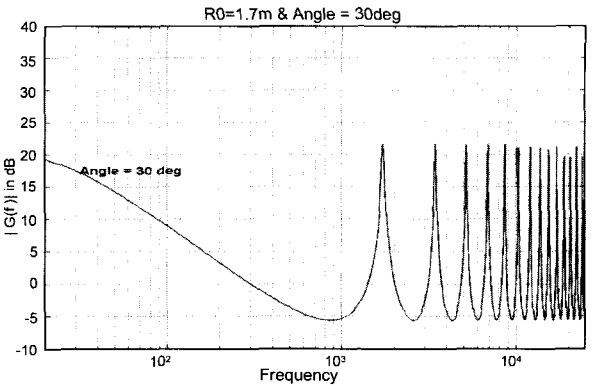
(그림 2) 5.1채널 트랜스오럴 필터링 과정

정면채널과 서라운드 채널에 인가되는 신호의 정보는 좌·우 채널에 동일하게 인가하게 되기 때문에, 청각 기관의 인지에는 영향을 미치지 않는다. (그림 1)에서 보인 바와 같이 청취자와 스피커의 거리 및 각도 $\{R_0, \theta\}$에 따라 좌우 채널의 이득함수 $|G(f)|$는 결정된다.

3.2 트랜스오럴 필터의 구현과정

좌우 스피커의 이득(Gain)함수인 $|G(f)|$를 구현하는 방법을 알아보면, 아래 (그림 3)은 (그림 1)에서 주어진 조건 음원과의 거리 R_0는 $1.7m$이고, 청취자와 스피커의 각도 θ는 30° 에 대한 이득함수 $|G(f)|$을 각각 보이고 있다.

이득함수의 최대값은 $|G|_{\max}=g_c/(1-g_c^2)$이고 최소값은 $|G|_{\min}=g_c/(1+g_c^2)$이다. (그림 3)에 보인 경우와 같이 거리 $R_0=1.7m$이고 스피커의 각도 30° 인 경우 $g_c(30°)=0.9577$이다.



(그림 3) 트랜스오럴 필터 이득 함수

청취자와 스피커의 각도가 30°인 경우 $|G|_{\max}=21.3dB$, $|G|_{\min}=-6dB$로 각각 확인되었다.

(그림 3)에서처럼 저주파에서 이득보상이 필요하며, 이것의 적절한 보상이 필요함을 알 수 있다. 트랜스오럴 필터의 구현에서 직면하는 가장 기본적 문제점은 (그림 3)에 보인 이득함수의 효과적 보상에 있다.

기존 트랜스오럴 필터 기법들은 위의 문제에 대하여 효과적인 해결 방안을 제시하지 못하고 있는 실정이다. (그림 3)에서 보인 바와 같이 일반적으로 이용되고 있는 청취자와 스피커의 각도가 30°인 경우, 20kHz까지의 가청 범위에서 11개의 피크를 갖는 이득함수 $|G(f)|$의 주파수 특성은 디지털 필터로 구현할 때 음색 왜곡과 음질 저하를 야기할 수 있는 인자임을 알 수 있다[2, 3].

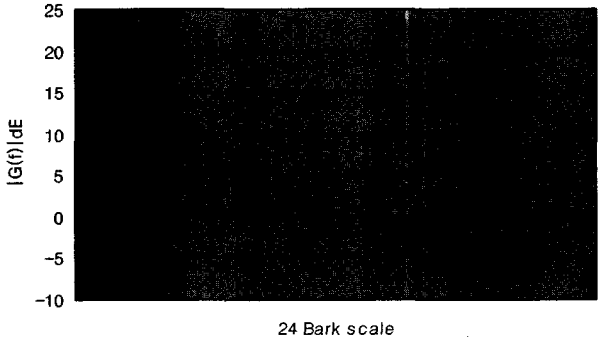
본 연구에서는 다수의 피크를 갖는 이득함수 $|G(f)|$의 주파수 특성을 만족하는 이득보상 기법을 해결하기 위하여, "마스킹(Masking)" 청각 특성[11, 12]을 이용한 유한 대역폭의 등가 이득율(Equivalent bandwidth gain)을 계산하여 24 밴드의 바크 스케일(bark scale)을 이용하여 이득 보상을 수행하는 방법을 채택하였다.

사람은 귀를 통하여 소리를 듣게 되는데, 귀는 외이(外耳), 중이(中耳), 그리고 내이(內耳)로 이루어져 있다. 특히, 내이에 해당하는 달팽이관에서 소리의 주파수 분석이 이루어진다. 달팽이관을 펼쳐 보면 하나의 긴 관으로 볼 수 있는데, 관의 각 위치에 따라 공진 주파수가 달라서 어떤 특정 위치점은 어떤 특정 주파수 대역을 담당하게 된다. 그런데, 사람이 소리를 감지하는 달팽이관의 주파수 스케일은 선형적이지 않고 지수 함수적인 특색을 가지고 있다. 즉, 사람의 소리 분석계를 필터 뱅크 시스템으로 볼 때, 각 대역은 중심 주파수가 높아질수록 대역폭이 커지는 형태를 나타낸다. 이런 주파수 스케일을 '바크 스케일'이라고 하며, 이것은 인간의 가청 주파수 대역을 24개의 대역으로 나눈다. 그리고 이 대역을 임계대역(Bark)이라고 부른다. <표 1>은 24개의 임계대역과 그에 해당되는 주파수 대역, 그리고 대역폭을 보여주고 있다[11, 12].

(그림 4)와 같이 이득함수 $|G(f)|$를 <표 1>에 따라서 각 주파수 대역별로 등가보상 필터를 구현하였다.

〈표 1〉 임계대역과 주파수 범위, 대역폭

Bark	Frequency Range(Hz)	Width	Bark	Frequency Range(Hz)	Width
1	0~ 100	100	13	1720~ 2000	280
2	100~ 200	100	14	2000~ 2320	320
3	200~ 300	100	15	2320~ 2700	380
4	300~ 400	100	16	2700~ 3150	450
5	400~ 510	110	17	3150~ 3700	550
6	510~ 630	120	18	3700~ 4400	700
7	630~ 770	140	19	4400~ 5300	900
8	770~ 920	150	20	5300~ 6400	1100
9	920~1080	160	21	6400~ 7700	1300
10	1080~1270	190	22	7700~ 9500	1800
11	1270~1480	210	23	9500~12000	2500
12	1480~1720	240	24	12000~15500	3500



(그림 4) 바크스케일 등가이득

특히, 30°인 경우 고주파 영역의 이득율은 이득함수 $|G(f)|$가 많은 최대 피크들 때문에 이득 보상량의 증가를 보인다. 그러나 2채널 스피커 시스템(스테레오 시스템)외에 이 방법을 그대로 현재의 모든 음향 시스템에 적용하는 것은 무리가 있다. 현재의 5.1채널 등의 시스템에서는 서브 우퍼의 사용 등 각 스피커간에 각기 다른 주파수 대역의 신호를 인가하여 각 스피커의 역할을 다르게 하고 있다.

4. 5.1채널용 크로스토크 생성프로그램 설계

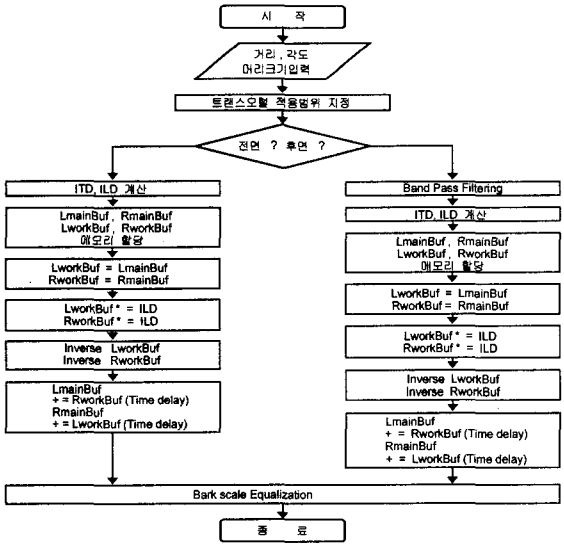
앞에서 설명한 트랜스오럴 필터링을 구현하기 위하여윈도우즈를 기반으로 하는 음원저작 툴을 제작하였다.

4.1 편집 프로그램

크로스토크 제거 기능을 정면 채널 및 서라운드 음원에 대해 처리할 수 있도록 하였다. 편집기능을 살펴보면 대용량 데이터 지원 포맷인 RIFF형인 Windows PCM 형태의 웨이브 데이타는 윈도우 메모리 보호를 위해 1024Kbyte 단위로 메모리에 로드된다. 이때 메모리는 수십 메가바이트에 해당하는 데이터의 효율적인 메모리 적재를 위해 가상 메모리를 사용하며 저장방식은 이중 연결리스트를 사용하였다. 또한 이 프로그램의 기본 틀은 MFC(Microsoft Foundation Class)와 WIN32 API를 이용하여 만들어졌으며 필터설계의 용이성을 위해 파일 처리부분과 메모리 처리부분은 특별히 설계되었다.

4.2 필터 프로그램 설계

원시 웨이브 데이타 파일에 대해 트랜스오럴 필터는 정면 채널과 서라운드 부분으로 나누어 크로스토크를 제거토록 하였다. 또한 다양한 환경에 적용 가능하며 고속연산을 할 수 있도록 하기 위해 청취자와 스피커간의 거리, 스피커의 각도, 청취자의 머리크기 등을 파라미터 값으로 받아 시간축 상에서 샘플단위로 연산을 수행하였다. 정면 채널에 대해서는 앞 절에서 설명한 방법에 따라 τ_c와 $|G(f)|$를 적용 후 바크 스케일에 따른 저주파 보상을 수행하였으며, 단음상 정위에 커다란 영향을 끼치지 않고 서브 우퍼에 공급되는 250Hz 이하의 주파수 범위에 대해서는 처리하지 않았다. 또한 서라운드 채널에 대해서는 5.1채널 포맷의 재생에서 서라운드 채널에 공급되는 주파수 범위에 따라 300Hz~7kHz대역의 대역제한 필터를 거친 후 크로스토크를 제거토록 하였다. 다음 (그림 5)는 크로스토크 제거과정을 나타내고 있다.



(그림 5) 트랜스오럴 필터 순서도

5. 실험 및 성능평가

5.1 실험방법 및 제시

음향 청취실험시 일반적으로 가정환경 및 보편적인 청취실을 기준을 만족한다. 이를 위해 ITU-R에서 정한 멀티채

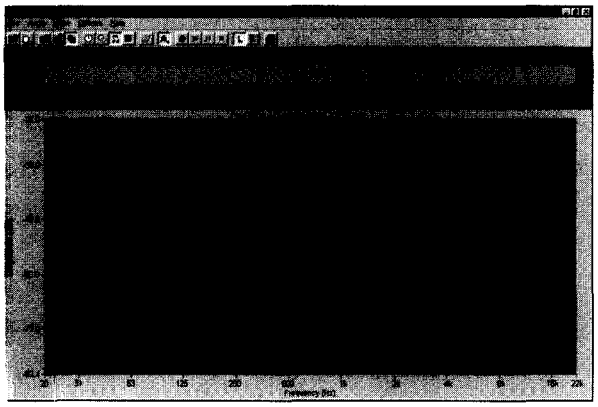
널 시스템의 평가 항목에 따라 맞추어 실험을 행하였다. 이는 (그림 7)과 같다[6, 7].

음악만을 재생하는 스피커의 경우 ±30°의 배치를 가진다.

본 논문에서는 멀티채널 스피커를 사용하여 정면 채널 및 서라운드 채널에 대하여 크로스토크를 제거시 청취자가 인지하는 질을 조사하는 것이 목적이다. 일반적으로 멀티채널 시스템에서 성능 평가시 다음의 항목에 대해 실시하고 있다[6, 7].

① Low-frequency Reproduction
② Listening & Viewing Position
③ Program material
④ Reproduction Level

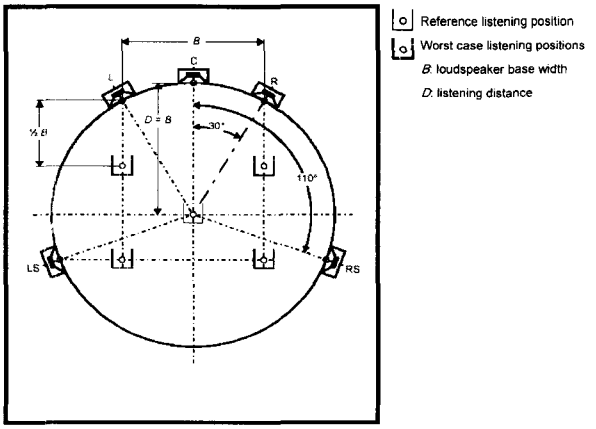
5.1.1 청취 실험실 음향 특성



(그림 6) 청취실의 음향 특성

5.1.2 청취 실험실 셋업

±30°의 두 개의 앞쪽 스피커는 ITU-R 권고안에 따라 1.7m의 높이, 주변 밝기는 스크린 주위에서 5 LUX정도의 조도를 유지하였다. 실험 전 재생물(시료)에 대해서는 비디오신호의 경우 미리 저장되었으며 오디오 신호는 ONKYO 5106



(그림 7) ITU-R 권고 청취실 레이아웃

디코더를 이용하여 WAV 파일로 저장되었다. 오디오만 있는 신호는 CD(44.1kHz 16 bit)를 이용하였다. 사용된 스피커는 평탄한 주파수 특성을 가진 2 way Tannoy series 5로 재생하였다.

실험은 음원이 저장된 PC의 중앙에 설치된 마이크를 통하여 크로스토크의 성능이 시험되었고, 주관평가 시에는 중앙에 의자를 설치하여 청취자와 스피커간의 거리가 1.7m를 유지하도록 하였다. 서브 우퍼(LFE X-over)의 경우는 SR용도의 18″ 제품을 사용하였다. LFE X-over는 주파수 20Hz~000Hz의 가변이 가능한 액티브형 dBX 320를 사용하였다.

5.1.3 평가용 자료

평가용 프로그램 자료는 클래식, 대중음악, 보컬 등을 포함한 것이 바람직하지만 EBU에 의해 제작된 SQM(Sound Quality Assessment Material)을 사용하기도 한다. 가정 환경에서의 평가를 위해서는 현재까지 특별히 제작된 것이 없으며 각 개발자들이 개별적으로 녹음하여 사용하고 있다[6, 7]. 본 논문에서의 실험은 44.1kHz 16bit로 녹음된 음원(Dry source) 2종과 5.1채널로 녹음된 음원 3종 총 5종의 음원을 이용하였다.

(1) 쥬라기 공원
첫 번째로 쥬라기 공원에 들어갔을때의 장면, 여기서 디아노사우르스 등의 동물 등장
(2) 쥬라기 공원
티렉스가 등장시의 장면, 비오는 소리등의 확산음장
(3) 라이언 일병구하기 - 오마하 해변 상륙장면
(4) 브리튼 벤자민 "Young persons"
(5) 레베카피존 "spanish harlem"

5.1.4 주관 평가

주관적인 평가는 입체적인 느낌(spatial impression)을 재현하는 것에 대하여 그것의 해상도와 질을 추정하는 것이 주목적으로 프로그램의 종류에 따라 다른 평가 항목을 가져야 한다. 시청각 프로그램을 위한 주제에 대한 입체적인 인상에 1dB 위치에서 5점의 간격을 가진 척도의 등급으로 평가를 실시한다. 0은 시청각 프로그램에 의해 생성되는 입체적인 인상이 입체적인 소리의 인상과 아무 관련이 없다는 것을 정의하며 5의 등급은 소리의 인상이 시각의 인상에 따라서 완전히 일치하는 것을 의미한다. 입체적인 소리에 대해 다음과 같은 인자를 포함한다.

1) 앞의 방향에서 넓이에 관한 입체적인 인상
2) 앞의 방향에서 깊이에 관한 입체적인 인상
3) 소리에 의해 청취자가 느끼는 음악같은 느낌의 정도

1)~3)의 인자에 대하여 평가자가 느끼는 성능에 대한 가중치는 평가하는 청취자에 의해 결정되도록 한다. 오디오만

재생되는 프로그램에 대해 '5'는 녹음이 된 공간을 완전하게 재현하는 것을 뜻하며, '0'은 재현되는 공간과 녹음하는 공간사이에 아무 상관이 없음을 의미한다. 음의 품질평가는 다음과 같은 ITU-DS.562.3의 권고를 따르며, 이 논문에서는 음악의 평가에 흔히 사용되는 다음의 5등급 척도를 사용하였다[6, 7].

〈표 2〉 5등급 선호도 척도

점 수	척 도
5	아 주 좋 다
4	좋 다
3	보 통
2	나 쁘 다
1	아주 나쁘다

5.1.5 자극의 제시

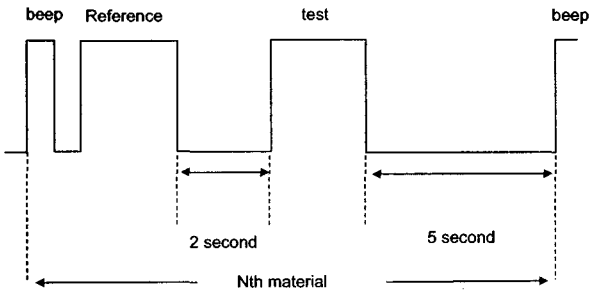
ITU-R 권고에서는 일대일 비교법에 의한 2개 음의 차이를 엄밀하게 비교할 경우에는 다음의 방법으로 제시할 것을 권고하고 있다. A가 원음이고 B가 평가음일 경우 A, B 모두 평가음일 경우가 있다. 어느 경우에도 제시 순서와 조합을 랜덤하게 해야 한다. 하나의 평가 세션(session)은 시간을 약 15분~20분 이하로 하며 휴식시간은 평가시간과 같은 정도로 하는 것으로 한다. 간단한 평가를 위해서 1회만 제시하여 평가하기도 한다. 제시방법은 다음의 <표 3>과 (그림 8)에서 보는 것과 같은 제시방법을 사용하여 평가하였다[6, 7].

이 실험에서 일대일 평가방법에 따라 미리 녹음된 소스를 A : 기존의 음, B : 크로스토크 제거 된 음을 랜덤 하게 피험자에게 제시하였다. 6명의 피험자에게 제시하는 음은 각각 제시 순서가 다른 파일로 미리 녹음하여 주관평가를 실시하였다.

〈표 3〉 신호의 제시순서

제 시 순 서	R-A	A-R	R-A를 랜덤 제시
평 가			5점의 평가

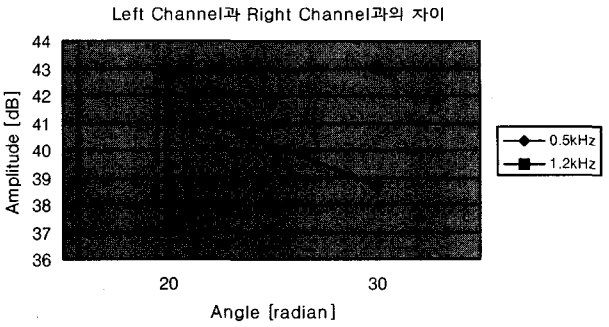
R : Reference 신호(original), A : Test 신호



(그림 8) 신호 제시의 순서

5.2 트랜스오럴 성능평가

크로스토크 제거기의 성능평가를 위해 주파수 제한적인 대역에 걸쳐서 성능을 측정하였다.



(그림 9) 0.5kHz Sine Signal과 1.2kHz Sine Signal의 스피커 20°와 30° 위치에 다른 성능 평가

(그림 9)에서는 20°와 30°에서 좌, 우측(Left & Right)의 값들을 Amplitude의 차이로 본 것이다. 결과에서 ±30°에서의 크로스토크 성능을 살펴보면 그 결과 38~42dB의 크로스토크 제거(cross-talk cancellation) 성능을 가지고 있고 1.2kHz에서 균형 있는 성능을 갖는 것으로 나타났다.

5.3 5.1채널에서 서라운드채널에 대한 주관평가

실험에 참여한 청취자는 총 10명이며 5명의 전문가(음향연구 3년 이상자)와 5명의 비전문가(일반인)가 2가지 항목이 있는 설문지에 대해 최고 5점 내에서 0.1간격의 점수를 부여토록 하였다[6, 7]. 총 5종의 음원에 대해 평가하였으며, 그 중 단일음원에 의한 평가결과와 음악을 이용한 평가 결과는 다음과 같다.

〈표 4〉 5.1채널(DTS)에서 서라운드채널 음원에 대한 크로스토크 제거 적용여부

①		일반 재생	트랜스오럴 재생	②		일반 재생	트랜스오럴 재생
전문가	청취자 1	3.8	4.2	전문가	청취자 1	4.2	4.2
	청취자 2	4.0	4.5		청취자 2	4.5	4.6
	청취자 3	3.6	4.1		청취자 3	4.1	4.1
	청취자 4	3.8	4.4		청취자 4	4.1	4.2
	청취자 5	3.8	4.3		청취자 5	4.0	4.3
	평 균	3.8	4.3		평 균	4.2	4.3
비전문가	청취자 6	4.1	4.2	비전문가	청취자 6	3.8	3.8
	청취자 7	4.0	4.1		청취자 7	4.1	4.1
	청취자 8	3.9	4.0		청취자 8	4.2	4.2
	청취자 9	3.8	4.2		청취자 9	4.1	4.3
	청취자10	4.0	4.1		청취자10	4.2	4.1
	평 균	3.9	4.1		평 균	4.1	4.1

〈표 5〉 5.1채널(Dolby ProLogic)에서 서라운드채널 음원에 대한 크로스토크 제거 적용여부

①		일반 재생	트랜스오럴 재생	②		일반 재생	트랜스오럴 재생
전문가	청취자 1	3.8	4.2	전문가	청취자 1	4.0	4.2
	청취자 2	4.1	4.4		청취자 2	4.2	4.2
	청취자 3	3.9	4.5		청취자 3	4.0	4.3
	청취자 4	4.1	4.4		청취자 4	4.2	4.4
	청취자 5	4.0	4.4		청취자 5	3.9	4.2
	평 균	4.0	4.4		평 균	4.1	4.3
비전문가	청취자 6	4.1	4.1	비전문가	청취자 6	4.1	4.2
	청취자 7	4.0	4.2		청취자 7	3.9	4.3
	청취자 8	3.9	3.9		청취자 8	4.2	4.0
	청취자 9	4.1	4.2		청취자 9	4.0	4.2
	청취자10	4.0	4.3		청취자10	3.8	4.3
	평 균	4.0	4.1		평 균	4.0	4.2

서라운드 채널에 대한 음상 확장감에 대한 선호도 평가 결과는 전문가 그룹이 DTS 부문에서 평균 3.8점에서 평균 4.3점으로 약 0.5점 향상되었으며, 돌비 프로로직 부문에서는 평균 0.4점이 향상된 결과를 보였다. 비전문가 그룹에 대한 선호도 평가에서는 DTS 부문이 평균 3.9점에서 평균 4.1점으로 약 0.2점 향상되었으며, 돌비 프로로직 부문에서는 평균 4.0점에서 평균 4.4점으로 평균 0.4점이 향상된 결과를 보였으며, 음상 깊이감에 대한 평가에서는 전문가 그룹이 DTS 부문에서 평균 4.1점에서 평균 4.2점으로 향상되었으며, 돌비 프로로직 부문에서는 평균 0.2점이 향상된 결과를 보였다. 비전문가 그룹의 선호도 평가에서는 DTS 부문이 변함없이 4.1점을 유지하고 있으며, 돌비 프로로직 부문에서는 평균 0.2점이 향상된 것으로 나타났다.

6. 결 론

본 논문에서는 기존의 스테레오에 대하여 크로스토크 제거를 확장시켜 5.1채널을 채용한 시스템에서 크로스토크의 적용방법과 이에 따른 바크 스케일 등가 보상필터를 이용한 트랜스오럴 필터를 구현하였다. 구현된 필터를 이용하여 실험한 결과 스테레오 폭, 즉 확산감의 영향은 크로스토크의 영향을 제거하였을때 가장 양호하게 나타났으며, 확산감의 증가는 모든 위치에서 입체적 인상 재현의 질적 향상을 확인하였다. 또한 정면채널의 크로스토크보다는 서라운드채널에서 그 효과가 더 크다는 것을 확인할 수 있었다.

이 논문에서는 서브우퍼에 대한 영향에 대해서는 실험결과를 넣지 않았다. 서브우퍼는 음상정위 및 확산감에 영향을 끼치지 않는다. 입체적 인상의 질은 시청각, 청각 제시의 제작물에 대해 크로스토크 영향에 의한 확산감에 의하여 주관적 인지 넓이의 인상과 선형적인 관계가 있다.

제안된 트랜스오럴 필터는 38dB 이상의 채널 분리도를 보였다.주관 평가시 선택된 제작물(시료)은 입체정보의 질의 재현과 음장의 변화가 명백한 제작물을 선택하였으며, 유럽의 EUREKA project 주관 평가시 사용된 제작물을 참고하였다. 전문가 및 비전문가를 대상으로 실험한 결과 전문가계층에서는 0.4~0.5점 이상 확산감 향상을 얻은 것으로 평가하였다. 따라서 기존의 5.1채널 제작물을 본 논문에서 제안된 방법으로 제작 및 재생시 입체감을 향상시킬 수 있다.

참 고 문 헌

[1] 정완섭, 김순협 "음향환경 제시용 3D sound 음장 재현장치에 관한 연구", '97 한국음향학회 학술발표대회 논문집.

[2] 정완섭, "공간 음상정위를 위한 Transaural 필터 구현기법", 한국음향학회, 1999.

[3] G. Theile, "The New Sound Format 3/2-stereo," The 94th AES Convention, Berlin, preprint 3550, Mar., 1993.

[4] Michael A Gerzon and Geoffrey J Barton, 1992, Ambisonic Decoders for HDTV, presented at the 92nd Audio Engineering Society Convention, Vienna, 24-27 March. Preprint 3345.

[5] Dolby Laboratory Inc. http://www.dolby.com "Some guidlines for producing music in 5.1 channel surround".

[6] Bech "Calibration of Relative Level Difference of a Domestic Multichannel Sound Reproduction system," J Audio Eng, April, 1984.

[7] ITU-R BS775-1, "Multichannel stereophonic Sound system with & without Accompanying Picture".

[8] T. Takeuchi, P. A. Nelson and O. Kirkeby "Robustness of the Performance of the "Stereo Dipole" to Misalignment of Head Position," Univ. of Southhampton.

[9] O. Kirkeby, P. A. Nelson, H. Hamada "The "Srereo Dipole" - Binaural Sound Reproduction Using Two Closely Spaced loudspeakers," AES Preprint 4463(I6), 102nd AES Convention, Munich : Germany, 1997.

[10] O. Kirkeby, P. A. Nelson, H. Hamada "Local Sound field reproduction using two closely spaced loudspeakers," J. Acoust. Soc. Am., 104(4), pp.1973-1981, 1998

[11] 강성훈, "방송 음향", 기전연구사, 1997.

[12] E. Zwicker, H. Fastl, Psychoacoustics : Facts and Models, Springer-Verlag, Berlin, 1990.

김 학 진

e-mail : kimhj@shinbo.co.kr

1986년 건국대학교 전자계산학과(학사)

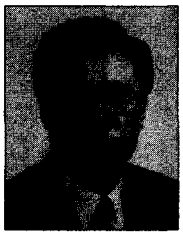
1997년 연세대학교 전자공학과(석사)

2004년 광운대학교 컴퓨터공학과 공학박사

1986년~현재 신용보증기금 근무

2000년~현재 명지전문대학 겸임교수

관심분야 : 음성신호처리, 음성인식, 자연어처리, 3-D 멀티미디어



김 순 협

e-mail : kimsh@daisy.gwu.ac.kr

1974년 울산대학교 전자공학과

1976년 연세대학교 석사

1983년 연세대학교 공학박사

1998년~1999년 한국음향학회 회장

1999년~2003년 광운대학교 사회교육원장

1979년~현재 광운대학교 컴퓨터공학과 교수

2000년~현재 한국음향학회 명예회장

2003년~현재 광운대학교 정보과학기술대학 원장

관심분야 : 음성신호처리, 음성인식, 자연어처리, 3-D 멀티미디어