

입체음향 생성에 있어서 자연스러운 이동음 효과의 구현

(Implementation of Smooth Moving Sound Effect in 3D Sound Generation)

명 현[†] 김기홍[†] 김기호[†]
(Hyun Myung) (Ki Hong Kim) (Ki Ho Kim)

김용완[†] 김현빈^{**} 김풍민^{***}
(Young Wan Kim) (Hyun Bin Kim) (Poong Min Kim)

요약 최근 PC의 성능이 향상되고 디지털 신호처리 기술이 발달함에 따라 복잡한 계산을 필요로 하는 입체음향 처리를 PC상에서도 구현하는 것이 가능하게 되면서, 멀티미디어 분야에서 입체음향에 대한 관심이 고조되고 있다. 특히 2채널 방식의 입체음향 기술은 공간활용성과 경제성에서 우수하여 많은 연구가 진행되고 있다. 2채널 재생 방식에 있어서 위치음 효과는 비교적 단순한 반면에, 이동음 효과는 특정 점에서만 측정된 머리전달함수를 이용하여 이동하는 음을 생성해야 하기 때문에 해결해야 할 문제점이 많다. 본 논문에서는 입체음향을 2채널로 재생할 때 이동음 효과가 부드럽게 생성되도록 하는 방법에 대해 다루고자 한다. 특히 부드럽게 궤적을 생성하는 방법, 그리고 이동음의 출력이 이산적인 공간에서 부드럽게 연결되도록 하는 방법에 대해서 제안하고 실제 테스트를 통해서 그 유용성을 보여준다.

Abstract As it became possible to generate 3D sound on a PC environment due to the advances in computing performance and digital signal processing technology, 3D sound technology gains its focus in the multimedia area. Specifically, a two-channel based 3D sound technology is being studied by many researchers because of its space efficiency and economical structure. While the positional sound effect is simple in its implementation, the moving sound effect has many problems to be resolved as there are only discrete measured point of HRTF database. In this paper, we propose the method of generating smooth moving sound in a two-channel based 3D sound technique with respect to generating smooth trajectory, and the interpolation method of discrete measured HRTF data. We perform the tests in the PC environment and prove the utility of the proposed method.

1. 서론

입체음향은 음원(sound source)이 발생한 공간에 직접 위치하지 않은 청취자가 재생된 음향을 들었을 때에 음향으로부터 방향감, 거리감 및 공간감 등과 같은 공간

적 단서를 지각할 수 있는 음향을 말하며[1], 단순한 모노(mono)음이나 스테레오(stereo)음의 신호에 공간적 단서를 추가하여 입체음향신호로 변환하는 기술을 입체음향 생성기술이라 한다. 입체음향 생성 방식은 여러 대의 스피커를 이용하는 서라운드(surround) 타입의 멀티채널 방식과 2대의 스피커나 헤드폰을 이용하는 바이노럴(binaural) 타입의 2채널 방식으로 구분할 수 있다. 멀티채널 방식의 경우 핵심기술은 인코딩/디코딩(encoding/decoding)에 관한 기술인 반면, 2채널 방식은 입체음상정위(sound image localization), 가상음장 제어(virtual sound field control), 크로스톡 제거(crosstalk cancellation) 기술 등이 핵심기술이라 할 수 있다.

[†] 비 례 원 : 한국전자통신연구원 가상현실연구개발센터 연구원
myung@etri.re.kr
kumgh@etri.re.kr
khkim@etri.re.kr
ywkim@etri.re.kr

^{**} 경 회 원 : 한국전자통신연구원 가상현실연구개발센터 연구원
hbkim@etri.re.kr

^{***} 비 례 원 : (주)이미시스 대표이사
poong@etri.re.kr
논문접수 : 1999년 10월 26일
심사완료 : 2001년 7월 14일

입체음향관련 기술은 주로 미국, 일본 및 유럽과 같은 선진국을 중심으로 발달된 반면, 국내 기술은 이들 선진국에 비해 기술 수준이 아직 요원한 상태이다. 그러나, 최근에 연구소와 일부 대학에서 입체음향 관련 기술을 연구 중에 있으며, 입체음향을 지원하는 사운드카드나 기존 스테레오방식에서 입체음향을 지원하는 신제품들이 개발되고 있다.

입체음향을 구현하기 위해서는 공간상의 특정 위치에 음원을 위치시키는 입체음상정위 기술이 필수적인데, 이의 구현은 위치음과 이동음으로 나누어서 생각할 수 있다. 3차원 공간상의 특정 위치에서 음원이 들려지는 것과 같은 효과를 내는 것을 위치음 효과라고 하고, 3차원 상에서 궤적을 따라 음원이 움직이는 것처럼 들려지는 것을 이동음 효과라고 한다.

2채널 재생 방식에 있어서 위치음 효과는 비교적 단순한 반면에, 이동음 효과는 구현에 있어서 고려해야 할 점들이 많다. 위치음 효과는 단순히 머리전달함수(Head Related Transfer Function: HRTF)와의 컨볼루션을 통해서 구할 수 있지만 이동음 효과는 특정 점에서만 측정된 머리전달함수를 이용하여 이동하는 음을 생성해야 하기 때문이다.

이동음 효과 생성을 위한 기존의 방법은 주로 HRTF의 주파수 영역에서의 보간을 이용하여 왔으나[2, pp.164-169][3-7], 가청의 클릭음을 발생시킨다는 문제점이 있었다. 또한 시간 영역에서 HRTF를 보간하게 되면 주파수 상에서는 전혀 다른 보간 특성을 내게 되므로 이러한 단점을 보완하여 측정된 HRTF가 아닌 모델로 합성된 HRTF를 이용해서 모델의 매개변수인 ITD(Interaural Time Difference)를 보간하는 방법도 이용되어 왔다.[2][8] 이 방법은 부드러운 보간 특성을 나타내기는 하지만 HRTF를 실측한 것이 아니기 때문에 약간 현실감이 떨어진다고 할 수 있다.

본 논문에서는 이러한 단점들을 보완하여 입체음향을 2채널로 재생할 때 이동음 효과가 부드럽게 생성되도록 하는 방법에 대해 다루고자 한다.

본 논문의 2절에서는 2채널 방식의 입체음상 정위 기술에 대해서 서술하고, 3절에서는 자연스러운 이동음 효과 구현 방법에 대해서 서술한다. 마지막으로 4절에서는 테스트를 통한 실험 결과에 대해서 토론하고 결론을 맺는다.

2. 2채널 방식의 입체음향 기술

음원이 발생한 공간 내에 있는 청취자의 양쪽 귀에 마이크로폰을 각각 설치하여 녹음한 신호를 바이노럴 신호

라 하며, 이 신호를 헤드폰으로 재생할 경우, 현장에서 직접 듣는 것과 같은 음상을 지각할 수 있다. 음원과 음상은 공간적 특성이 반드시 일치하지 않으며, 음원과 음상이 일치할수록 좋은 음질의 입체음향이 구현되었다고 할 수 있다. 바이노럴 신호에는 음원의 위치, 방향뿐만 아니라 음원을 둘러싸고 있는 공간, 즉 음장과 관련된 공간적 단서들이 포함되어 있다. 음장이란 음향학적 용어로 가청주파수의 음파가 존재하는 공간을 말한다. 2채널에 의한 방식은 인간이 두 개의 귀로 음향을 지각하는 특성을 이용하여, 음상정위와 음장제어에 의해 생성된 입체음향은 2채널에 의해서 재생하는 방식을 말한다.

바이노럴 타입의 2채널 입체음향을 생성하는 방식으로는 녹음과 필터링에 의한 방법이 있다. 녹음에 의한 방법은 청취자의 양쪽 귀에 장착한 두 개의 마이크로폰을 이용하여 현장음을 녹음하고 이를 재생하는 방식이다. 모든 사람이 자신의 바이노럴 신호를 녹음할 수 없기 때문에, 현재는 주로 표준치의 머리 모형(가진 더미헤드(dummy head))에 장착한 마이크로폰을 통하여 바이노럴 신호를 녹음하고 이를 일반 청취자에게 들려주는 방식을 이용한다[2].

필터링에 의한 방법은 원음을 변형하여 입체음을 생성하는 방식으로, 주로 머리전달함수가 필터로서 이용된다. 머리전달함수는 무향실 내에서, 더미헤드를 중심으로 구의 형태로 여러 각도에 배치한 스피커로부터 백색 잡음과 같은 임펄스 신호를 방사시켜, 더미헤드의 양쪽 귀에 장착한 마이크로폰으로 측정된 임펄스 응답을 푸리에(Fourier) 변환한 것을 말한다. 이 머리전달함수와 원음을 컨볼루션(convolution)하면 원하는 공간상의 위치에 음상을 정위시킬 수 있다.

2.1 머리전달함수

청취자가 음원에 대한 공간적 단서를 지각할 수 있는 것은 머리전달계의 고유 특성에 의해서 두 귀에 입사한 두 신호 간에 차이가 발생하기 때문이다. 이 두 신호차에 대한 특성은 머리전달함수에 내포되어 있어, 이를 이용하면 입체화되지 않은 단순음에 공간적 정보가 부가된 바이노럴 타입의 입체음을 생성할 수 있다. 머리전달함수는 입사하는 각도에 따라 달라지기 때문에, 여러 위치에서 임펄스에 대한 머리전달함수를 측정하고, 이를 데이터베이스(DB)로 구축하는 것이 필요하다. 이 DB로부터 원하는 위치에 해당하는 머리전달함수를 선택하여 단순음과 컨볼루션을 수행하면, 해당 위치에 음상을 정위시킬 수 있다.

MIT Media Lab.에서는 KEMAR 더미헤드를 사용하여 710지점에서 측정된 머리전달함수 DB를 인터넷에서

무료로 공개하고 있다[3]. 본 논문에서는 Neumann 더미헤드를 이용하여 직접 측정된 HRTF DB[4, 5]를 사용하였다.

2.2 입체음상 정위

청취자가 지각한 음상에 대한 위치를 파악하는 것을 음상정위라 하고, 음상을 공간 상의 특정 장소에 위치시키는 기술을 음상정위기술이라 한다. 이 기술을 이용하면 고정된 특정 위치에서 소리가 지각되는 “위치음” 효과와 소리가 한 위치에서 다른 위치로 움직이는 “이동음” 효과를 생성할 수 있다. 위치음 생성은 해당 위치의 머리전달함수를 단순음과 컨볼루션 연산을 행함으로써 얻을 수 있고, 이동음 생성은 소리가 이동하는 궤적 상에 해당하는 연속적인 머리전달함수들을 단순음과 컨볼루션을 수행함으로써 얻을 수 있다.

이동음의 경우, 머리전달함수를 이용하는 방식 외에 도플러(Doppler)효과를 이용하는 방식이 있는데, 이는 음원이 가까워지면 주파수가 높은 쪽으로 이동하고, 멀어지면 주파수가 낮은 쪽으로 이동하는 현상을 이용하는 것으로, 소리가 멀어지고 가까워지는 이동음 효과를 낼 수 있다.

일반적으로 음상정위의 구현은, 공간 상의 음원이 청취자의 두 귀에 이르기까지의 전달함수를 나타내는 머리전달함수에 의한 바이노럴 필터링 처리와 음원이 청취자로부터 떨어진 거리에 따른 소리세기(intensity) 변화를 시뮬레이션 하는 이득(gain) 조절 과정의 구현으로 구성된다. 그림 1은 모노 음원에 대한 2채널 입체음상정위(binaural spatialization) 과정을 나타내는 블록도이다. 그림 1에서 $x[n]$ 은 원음을 나타내고, (θ, ϕ, r) 은 각각 머리 중심에서부터 원음까지의 방위각(azimuth), 고도각(elevation), 거리를 나타낸다. 또한 $H_L(\theta, \phi)$, $H_R(\theta, \phi)$ 는 왼쪽, 오른쪽 귀의 방위각, 고도각에 따른 HRTF를 나타내며, $G_L(r)$, $G_R(r)$ 은 거리에 따른 이득을 나타낸다. HRTF와의 컨볼루션과 이득 조절을 거쳐서 왼쪽, 오른쪽 출력인 $x_L[n]$, $x_R[n]$ 이 나오게 된다.

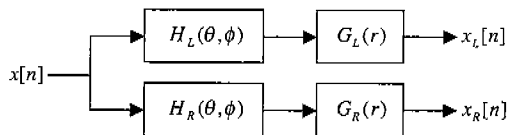


그림 1 Binaural Spatializer의 구성

입력 모노 신호에 대한 바이노럴 HRTF 필터링에 있어서 음상정위 위치를 지정하는 가상음원의 위치 정보

가 필요하며 가상음원의 청취자의 수평면(horizontal plane)과 정중면(median plane)에 대한 방위각과 고도각을 산정하여 이에 매핑되는 좌우 HRTF 쌍을 인덱싱하게 된다.

3. 자연스러운 이동음 효과의 구현

3.1 이동음 효과의 구현

2절에서 설명한 2채널 위치음 음상정위 모듈을 기반으로 하여 고정된 HRTF가 아닌 가변적인 HRTF와 가변적인 이득 제어를 통해서 한 위치에서 다른 위치로의 이동 음상을 구현할 수 있다. 이를 위해 움직이는 가상음원의 이동 궤적을 정의하고 궤적 상의 각 위치에 대응하여 변하는 HRTF를 연속적으로 매핑하고 필터링하는 것이 필요하다. [4-7]에 소개된 시스템 구현에서는 가상음원의 이동 궤적을 선분 세그먼트를 기반으로 구성하고 HRTF를 시간 영역에서 보간하였다. 기존에 사용하던 HRTF는 측정 공간을 메쉬(mesh) 형태로 분할한 특정 위치에서 측정된 불연속 HRTF이기 때문에 불연속 구간에서의 HRTF는 측정된 주위의 HRTF를 이용하여 보간 기법으로 재생하였던 것이다. 즉, 각 선분을 일정 개수의 세그먼트로 분할하고 각 분할된 세그먼트에 매핑되는 입력 신호를 보간된 HRTF를 이용하여 음상정위하는 과정을 연속적으로 수행함으로써 가상음원의 이동 속도에 따른 음상 이동을 구현하였다. 이때, 입력 오디오 신호의 처음 시작 부분이 가상음원의 이동 출발 위치에 대응하고 입력신호의 끝 부분은 가상음원의 이동 종료 위치에 대응시킴으로써, 오디오 파일의 소리가 궤적의 출발점에서 시작하여 가상음원의 속도에 따라 이동하다가 궤적의 끝점에서 마치도록 하는 방식을 택하였다. 즉, 오디오 파일이 어느 가수의 노래라면 음상정위 처리 결과는 가수가 청취자 주위를 지정된 일정 궤적에 따라 움직이면서 노래를 부르는 것과 같은 입체음향 오디오 파일이 되는 것이다.

그러나 [4-7]에서 구현된 이동음 효과는 다음과 같은 문제점들을 가지고 있었다.

- 1) 가변 HRTF 필터링에 있어서 HRTF의 필터계수가 바뀌게 되는 경계 부분에서 출력 신호에 클릭(click)음 발생 문제: HRTF가 바뀌는 지점에서 출력 신호의 크기(amplitude envelop)가 급격히 변함으로써 생기게 된다.
- 2) 측정되지 않은 위치의 HRTF에 대한 보간 생성 문제: HRTF 데이터 베이스가 연속적이지 않고 유한한 점에서 측정되었기 때문에 임의 위치의 HRTF를 얻기 위해서는 보간이 필요하다.

3.2 궤적의 구현 방법

궤적을 구현할 때 기존에는 선형 보간(linear interpolation)을 이용하여 구현하였다. 즉, 정해진 시간 간격으로 공간상의 궤적점들을 샘플링하고 그 중간 값들은 직선으로 보간하여 얻어내었다. 정해진 시간 간격이 커질수록 이 방법은 부드러운 곡선 궤적에서 멀어져서 모서리가 있는 다각형의 궤적 모습을 가지게 된다. 이 때문에 HRTF DB가 부드럽게 이어지지 못하고 직선 궤적상의 HRTF DB로 원음과 컨볼루션이 수행되어 결과적으로 부드럽지 못한 이동음을 생성한다.

이러한 단점을 극복하고 부드러운 이동음 효과를 얻기 위해서는 샘플링된 점들 사이의 궤적 자체를 부드럽게 이어주는 것이 필요하다. 이를 위해서 보간법들 중에서 널리 쓰이는 큐빅 보간법(cubic spline)을 도입하였다. 큐빅 보간은 각 점들에서의 비분을 연속으로 만들어 주어 궤적이 부드러운 곡선을 유지하도록 해준다.

원음을 일정한 시간 간격마다 샘플링한 개수가 N 개이고, 일정한 시간 간격마다 얻어진 샘플의 인덱스를 t_i ($i=1, \dots, N$)라고 하자. 또한 k 번째 점에서의 3차원 좌표를 $P_k=(P_x(t_k), P_y(t_k), P_z(t_k))$ 라고 하자. 이때 t_j 점과 t_{j-1} 점 사이에 있는 t 점의 좌표 P 는 다음과 같이 보간할 수 있다[9].

$$P = AP_j + BP_{j+1} + CP_j'' + DP_{j-1}'' \quad (1)$$

$$A = \frac{t_{j+1}-t}{t_{j+1}-t_j}, \quad B = 1-A$$

$$C = \frac{1}{6}(A^3-A)(t_{j+1}-t_j)^2, \quad D = \frac{1}{6}(B^3-B)(t_{j+1}-t_j)^2 \quad (2)$$

$$\frac{(t_j-t_{j-1})}{6} P_{j-1}'' + \frac{(t_{j+1}-t_{j-1})}{3} P_j''$$

$$+ \frac{(t_{j+1}-t_j)}{6} P_{j+1}'' = \frac{P_{j+1}-P_j}{t_{j+1}-t_j} - \frac{P_j-P_{j-1}}{t_j-t_{j-1}} \quad (3)$$

먼저 (3)식을 이용하여 각 점에서의 2차 도함수인 P_j'' ($i=1, \dots, N$)을 구한 다음 그 결과와 (2)식을 이용하여 t 점의 좌표인 P 를 (1)식으로 구한다.

3.3 HRTF 보간과 출력 보간

앞 절에서의 보간법에 의해 원음의 전체 샘플링 시간 내에 임의의 시간에서의 음원의 좌표를 찾아낼 수 있게 되었다. 그러나 우리가 필요로 하는 것은 모든 시간에서의 좌표가 아니라 HRTF와 원음을 컨볼루션할 때의 좌표만 필요하다. 컨볼루션 길이를 L 이라고 하고, IIRTF의 시간 영역 신호인 임펄스 응답 HRIR(Head Related Impulse Response)의 필터 길이를 F 라고 하면 $L \geq 2F-1$ 를 만족해야 한다. 이렇게 하여 Overlap-save 방법을 구현하게 되면 매 $(L-F)$ 샘플마다 새로운 HRTF

가 제공되어야 한다. 그러면 한번 컨볼루션할 때마다 $(L-F)$ 개의 샘플에 대해서 처리하게 되는데, 각 샘플마다 원음의 위치가 다르게 된다. 올바른 이동음 효과를 얻기 위해서는 $(L-F)$ 개의 샘플 내에서도 보간을 시켜주어야 하는 것이다. 이 때 기존의 방법인 [2, pp.164-169] [4-7]에서는 HRTF의 임펄스 응답, 즉 HRIR을 보간시키는 방법을 사용했다. 그러나 이렇게 할 경우 두 지점간의 HRIR 특성이 급격히 변하여 보간이 별로 의미가 없을 수 있다. 실제로 HRIR의 보간을 이용한 이동음은 가청의 콜택음을 부가적으로 생성하게 되어 매끄러운 음을 주지 못한다.[10]

또한 시간 영역에서 IIRIR을 보간하게 되면 주파수 상에서는 전혀 다른 보간 특성을 내게 되므로 이러한 단점을 보완하여 측정된 HRTF가 아닌 모델로 합성된 HRTF를 이용해서 모델의 매개변수인 ITD를 보간하는 방법도 이용되어 왔다.[2, pp.166-169] [8] 이 방법은 하나의 매개변수만을 보간하기 때문에 주파수 영역에서 부드러운 보간 특성을 나타내기는 하지만 HRTF를 실측한 것이 아니고 모델링한 것이기 때문에 약간 현실감이 떨어진다고 할 수 있다.

따라서 본 논문에서는 HRIR을 주파수 영역에서 보간하는 것보다는 각 지점에서 서로 다른 HRTF를 써서 출력음을 생성하고, 그 출력음을 시간 영역에서 매끄럽게 보간하는 방법을 사용하기로 한다. 현재 음원의 위치가 이전의 IIRIR인 h_{old} 와 새로운 위치의 IIRIR인 h_{new} 사이에 위치한다고 가정하자. 그때의 음원이 $x[n]$ 으로 표현된다면 출력 $y[n]$ 은 다음과 같이 두 출력음 h_{old} , h_{new} 의 보간으로서 얻어진다.

$$y_{new}[n] = x[n] * h_{new}[n]$$

$$y_{old}[n] = x[n] * h_{old}[n]$$

$$y[n] = y_{old}[n] \cdot c_1[n] + y_{new}[n] \cdot c_2[n] \quad (4)$$

이때 c_1, c_2 는 보간을 행하는 상수 계수들로서 부드러운 출력이 얻어지도록 설계되어야 한다.

3.4 출력 보간 함수의 비교

출력 보간 함수를 어떤 함수로 한 것이냐, 즉, 보간 계수인 c_1, c_2 를 어떻게 잡을 것이냐에 따라서 출력음이 달라지게 된다. [10]에서는 음의 크기(intensity)가 거리의 제곱에 반비례한다는 사실을 이용하고, 같은 거리에 있다는 가정하에 $c_1^2 + c_2^2 = 1$ 이 유지되도록 하였다. 그리고, 시작점과 끝점에서 fade-in, fade-out 효과를 주기 위해서 삼각함수를 이용하여 다음과 같이 계수를 구성하였다.

$$c_1[n] = \sin(\theta), \quad c_2[n] = \cos(\theta), \quad \theta = \frac{\pi}{2} \cdot \frac{n}{(L-F)} \quad (5)$$

그러나 HRIR이 h_{old} 에서 h_{new} 로 변할 때 거리가 일정하다는 가정은 항상 옳은 것은 아니므로 위의 계수 구성은 부드러운 음을 주지 못할 때도 있다. 이를 해결하기 위해서는 (4)식과 (5)식에서 거리에 대한 보정을 해 주어야 한다. 거리에 대한 보정을 포함한 새로운 계수 구성은 다음과 같다.

$$\begin{aligned} y_{new}[n] &= x[n] * h_{new}[n] \\ y_{old}[n] &= x[n] * h_{old}[n] \\ y[n] &= y_{old}[n] \cdot c_1[n] \cdot g_{old} + y_{new}[n] \cdot c_2[n] \cdot g_{new} \end{aligned} \quad (6)$$

$$c_1[n] = \sin(\theta), \quad c_2[n] = 1 - c_1[n], \quad \theta = \frac{\pi}{2} \cdot \frac{n}{(L-F)} \quad (7)$$

(6)식에서 g_{old} , g_{new} 는 이전 위치 및 새로운 위치의 거리에 따른 이득을 나타낸다. 또한 (7)식과 같은 사인 함수를 사용하지 않고 선형 함수로 다음과 같이 구성할 수도 있다.

$$c_1[n] = \frac{n}{(L-F)}, \quad c_2[n] = 1 - c_1[n] \quad (8)$$

위와 같이 출력의 보간은 사용자의 기호에 따라 선형 또는 비선형으로 다르게 할 수 있다.

4. 테스트 결과

테스트시의 HRTF 필터링에는 무향실에서 Neumann KU-100 더미헤드폰을 사용하여 직접 측정된 HRTF 데이터[4]를 사용하였으며 각 방향에 대한 HRTF 필터는 256개의 임펄스 응답 데이터를 FIR 필터 계수로 사용하였다. 또한, 단구간 푸리에 변환을 기반으로 한 Overlap-Save 방식 필터링[11]을 구현하였고 FFT 기반 주파수 영역 컨볼루션 알고리즘을 사용하였다.

$L=1024$, $F=256$, $g=1/r$ (r 은 머리 중심으로부터 음원까지의 거리)을 이용하여 샘플 wave 파일에 대해 이동음을 고찰하였다. 그림 2는 각각의 보간법을 이용하여 생성된 궤적을 나타낸다. 그림에서 점으로 나타난 것은 GUI를 통해 입력한 궤적점들을 나타낸다. 점들을 샘플링하는 시간 간격이 클수록 들성 들성하게 나타난다. 직선으로 연결된 것이 선형 보간에 의해 생성된 궤적이고, 곡선으로 연결된 것이 큐빅 보간에 의해 생성된 궤적이다. 큐빅 보간이 훨씬 자연스런 궤적을 제공하며 실제 HRTF DB의 인덱스가 자연스럽게 연결되도록 해준다. 청취 테스트 결과 선형 보간법 보다는 큐빅 보간법이 훨씬 자연스러운 이동음을 들려주었다.

그림 3은 모노인 원음 파형과 HRTF 보간을 이용하였을 때의 이동음 파형, 그리고 선형 출력 보간을 이용하였을 때의 이동음 파형을 보여준다. 전체 파형 중에서 일부 구간만을 동일하게 샘플링하여 보여준 것인데,

HRTF 보간을 이용하였을 경우 그림에서 점선 원으로 표시된 것처럼 클릭음이 발생하였으나 출력 보간을 이용하였을 경우에는 클릭음이 발생하지 않았다.

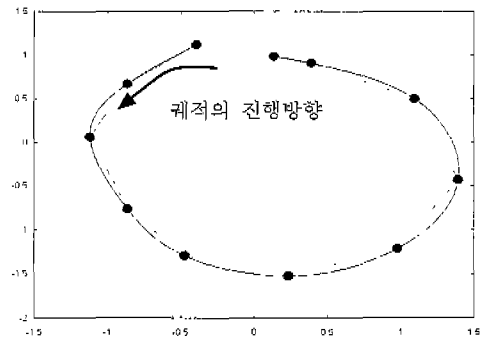


그림 2 점으로 표시된 것은 GUI를 통해 입력한 궤적. 직선으로 표시된 것은 선형 보간을 이용한 궤적. 곡선으로 표시된 것은 큐빅 보간을 이용한 궤적

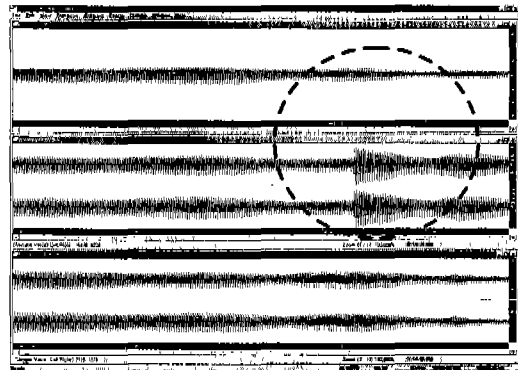


그림 3 (상) 원음 파형 (중) HRTF 보간 이동음 파형. 원음으로 표시된 부분에서 클릭음이 발생할 (하) 선형 출력 보간 이동음 파형

그림 4는 출력 보간법 중에서도 보간 계수를 구성하는 방법에 따른 영향을 살펴보기 위해서 실험을 한 결과이다. 제안된 출력 보간법으로 하였을 때는 선형으로 하는 것과 sine 함수 방식으로 하는 것이 거의 동일한 출력 효과를 보여 주었으나, [10]에서 제안된 sine 함수 방식으로 보간하였을 때에는 정현파형의 매개변수의 영향에 의해 저주파수의 정현파형이 발생함을 관찰할 수 있었다. 제안된 sine 함수 방식은 거리에 대한 보정을 따로 하였으나, [10]의 방법은 거리에 대한 보정이 없이 동일한 거리로 가정하여 계산하였기 때문에 이러한 정

현파형의 효과가 크게 나타나리라 생각된다. 그림 4(하) 파형을 더욱 자세히 그려보면 그림 5처럼 보이는데, 이를 청취하게 되면 이동음 효과는 느껴지지만 정현파형의 저주파수 대역음이 느껴짐을 알 수 있다.

이외에도 여러 가지 음원 파형에 대해서 테스트를 해 본 결과 큐빅 보간을 이용하여 궤적을 생성하고, 또한 이를 HRTF 보간이 아닌 제안된 출력 보간법을 이용하였을 때 더욱 부드러운 이동음이 생성됨을 알 수 있었다.

출력의 보간을 선형으로 할 것이냐 사인 함수로 할 것이냐는 사용자의 기호에 따라 다르지만 실제로 다양한 음원으로 테스트 한 결과 그다지 많은 차이를 느끼지 못하였다.

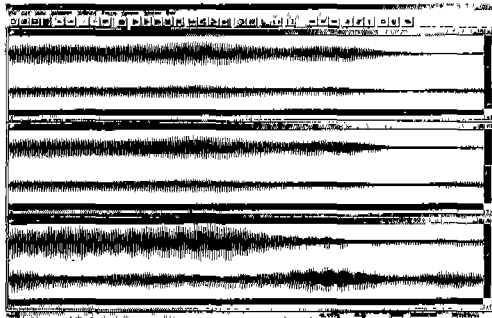


그림 4 각 보간법의 비교 (상) 제안된 선형 보간에 의한 출력 (중) 제안된 sin 함수 보간에 의한 출력 (하) [10]의 sin 함수 보간에 의한 출력

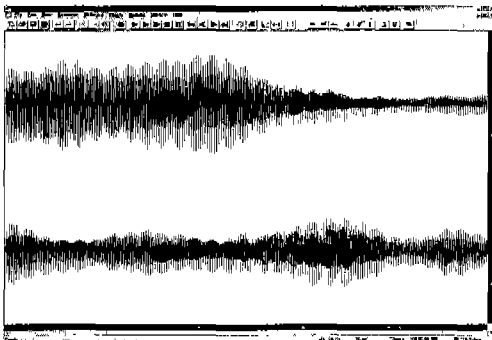


그림 5 그림 4(하)의 방법으로 하였을 때 더욱 확대한 파형. 사인 파형에 의한 저주파 잡음이 동반됨

5. 결론

본 논문에서는 바이노럴 타입의 2채널 입체음향을 생성하는 방식에 있어서 가장 중요한 효과라고 할 수 있

는 음상정위 부분 중에서 음원이 3차원 공간상의 궤적을 따라 이동하는 듯한 효과를 주는 이동음 효과의 개선에 중점을 두었다. 궤적을 생성하는 방법에 있어서는 선형 보간법과 큐빅 보간법을 비교하여 테스트한 결과 큐빅 보간법이 더욱 자연스러운 음을 생성함을 알 수 있었다. 생성된 궤적 상에서 이산화된 HRTF 데이터베이스를 이용하여 보간하는 방법에 있어서는 기존에 HRTF를 주파수 영역에서 보간하던 방법 대신에 출력음을 선형 또는 비선형으로 시간 영역에서 보간하는 방법이 보다 자연스러운 음을 생성하였다. 이와 같이 본 논문에서 제안된 이동음 효과 방법을 이용할 경우 3차원 공간상의 임의의 위치에서 임의의 다른 위치로 음원이 이동하는 효과를 자연스럽게 생성할 수 있다.

참고 문헌

- [1] 이희중, 서상원, 이명진, 김용완, "입체음향 기술", 소프트웨어 기술동향, vol.2, no.2, pp.54-80, 1998.
- [2] Durand R. Begault, *3-D Sound for Virtual Reality and Multimedia*, New York, Academic Press Inc., 1994.
- [3] <http://sound.media.mit.edu>
- [4] 서상원, 김재우, 이명진, 김용완, "Neumann 더미헤드를 사용한 머리전달함수(HRTF)의 측정", 한국감성과학회 춘계학술발표회, 1998.
- [5] Hyun Myung, Ki Hong Kim, Yong Wan Kim, Poong Min Kim, SangWoon Suh, Hyun Suk Kim, Jae Woo Kim, and Hyun Bin Kim, "The Development of 3D Sound Signal Editor "SoriWave" for Multimedia Contents," *Audio Engineering Society 106th Convention*, 4946(P4), Munich, Germany, May, 1999.
- [6] 이명진, 이희중, 김현빈, "게임 S/W 개발용 입체음향 저작도구 개발에 관하여", 한국음향학회 학술대회 논문집, 제16권 제1(s)호, pp.197-200, 1997.
- [7] 서상원, 김기홍, "입체음향 생성 저작도구 SoriWave를 위한 음상정위 인터페이스의 개발", HCI'99 학술대회지, pp.515-519, 1999.
- [8] E. M. Wenzel and S. H. Foster, "Perceptual consequences of interpolating head-related transfer functions during spatial synthesis," in *Proc. of the ASSP (IEEE Workshop on Applications of Signal Processing to Audio and Acoustics)*, New York, IEEE Press, 1993.
- [9] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Second edition, Cambridge University Press, pp.113-116, 1995.
- [10] An-Nan Suen, Jhing-Fa Wang, and Jia-Ching

Wang, "VLSI Implementation of 3-D Sound Generator," *IEEE Trans. on Consumer Electronics*, vol.43, no.3, pp.679-688, Aug. 1977.

[11] P. M. Embree, *C Algorithms for Real Time DSP*, Prentice Hall Inc., 1995.



명 현

1992년 2월 KAIST 전기 및 전자공학파 학사 졸. 1994년 2월 KAIST 전기 및 전자공학과 석사 졸. 1998년 9월 KAIST 전기 및 전자공학과 박사 졸. 1996년 SEAL(Simulated Evolution and Learning)'96 국제학회 사무총장. 1998년 9월 ~ 현재 한국전자통신연구원 가상현실연구개발센터 선임연구원. 현 한국감성과학회, 한국생물정보학회 정회원. 관심분야는 가상현실, 생체신호 HCI, 최적화, 진화연산, 인공지능, 바이오 인포매틱스, 입체음향



김 기 홍

1994년 2월 경북대학교 전자공학과 학사 졸. 1996년 2월 경북대학원 전자공학과 석사 졸. 1995년 12월 ~ 1998년 5월 시스템공학연구소 (SERI). 1998년 5월 ~ 현재 한국전자통신연구원 가상현실연구개발센터 선임연구원. 관심분야는 음성신호처리, 생체신호처리, HCI



김 기 호

1990년 광운대학교 컴퓨터공학과 석사 졸. 2000년 광운대학교 컴퓨터공학과 박사 졸. 1984년 ~ 현재 한국전자통신연구원 가상현실연구개발센터 선임연구원. 관심분야는 의료가상현실, 입체영상, 입체음향, 실물모델링, 멀티미디어, 멀티미디어 통신 및 응용, 네트워크 보안, 네트워크 게임



김 용 완

1992년 3월 ~ 1996년 2월 인하대학교 전자공학과 학사 졸. 1996년 3월 ~ 1998년 2월 광주과학기술원 정보통신공학과 석사 졸. 1998년 2월 ~ 1998년 8월 시스템공학연구소(SERI). 1998년 8월 ~ 현재 한국전자통신연구원 가상현실연구개발센터. 관심분야는 가상현실, 온라인 게임 엔진, 입체음향, 영상처리, 영상압축, HCI



김 현 빈

1978년 3월 ~ 1985년 2월 중앙대학교 정경대학 응용통계학과 학사 졸. 1986년 3월 ~ 1988년 8월 중앙대학교 대학원 응용통계학과 석사 졸. 1993년 4월 ~ 1996년 3월 Okayama 대학 대학원 자연과학연구과 박사 졸. 1991년 3 ~ 1993년 3월 Nagoya 대학 대학원 공학연구과 연구원. 1984년 11월 ~ 1998년 6월 시스템공학연구소(SERI) 책임연구원. 1998년 7월 ~ 현재 한국전자통신연구원 가상현실연구개발센터 책임연구원, 팀장. 1999년 12월 멀티미디어 기술사 자격증. 2000년 12월 정보통신부 장관상 수상. 2001년 7월 ~ 12월 국가과학기술자문회의 전문위원. 현 한국감성과학회 이사. 현 한국음향학회 편집위원회 위원. 관심분야는 음향신호처리(입체음향), 음성정보처리(인식, 합성), 생체신호처리(Biometric, Brain Wave), 감성정보처리, 인공지능, 휴먼인터페이스



김 봉 민

1985년 2월 단국대학교 전산통계학과 학사 졸. 1993년 2월 고려대학교 전산학과 석사 졸. 1985년 2월 ~ 2000년 5월 한국전자통신연구원 가상현실연구센터 책임연구원. 1992년 6월 우수연구사례전시 표창장 (과학기술부). 2000년 12월 정보통신부 장관상. 2000년 5월 ~ 현재 (주) 이머시스 대표이사. 현 한국멀티미디어학회 이사. 현 정보과학회 정회원, 한국특수교육학회 정회원, 정보처리학회 운영위원. 관심분야는 입체음향, 멀티미디어, 인터넷 방송