

◆특집◆ Intelligent HCI를 위한 TSI 기술

가상현실에서의 음향효과 구현

이종민\*, 황요하\*

Sound Rendering Techniques for Virtual reality

Jong Min Lee\* and Yoha Hwang\*

Key Words : Virtual reality (가상현실), Sound rendering (음향효과), Dental simulator (치과 시뮬레이터)

1. 서론

음향효과는 가상현실 세계를 구현하여 인간과 상호작용을 가능하게 하는데 있어서 매우 중요한 요소로서 인간-컴퓨터의 상호작용(human-computer interaction) 시스템에서 양적인 정보뿐만 아니라 존재감, 사실감, 질감 등을 전달하는 아주 민감한 역할을 한다. 폴리음향(Foley sound)이라고 불리는 음향효과는 애니메이션, 영화, 게임 산업에서 넓게 사용되고 있으며, 일반적으로 음향효과를 내는 일은 노동 집약적이고 시간이 오래 걸리는 과정이기 때문에 소수의 음향 전문가들에 의해서 그들만의 경험적 지식에 의하여 만들어지고 있다. 그러나 가상현실 시스템에서는 이벤트(event)에 따라 이미 지정된 소리를 발생시키는 것이 아니라 인간의 움직임이나 반응에 따라 소리가 나도록 하여 사실감을 높여주는 것이 가능해야 한다. 따라서 인간-컴퓨터의 상호작용에 따라 실시간으로 컴퓨터가 음향효과를 만들어 내는 분야에 대한 연구는 가상환경 시스템에 반드시 필요하며, 2000년대 초반부터 어느 정도 의미 있는 결과물이 도출되고 있다.

소리는 객체 자체의 진동이나 객체 표면의 진동이 공기의 압력을 변화시켜서 발생하게 된다.

일반적으로 가상현실 시스템에서의 음향효과는 두 객체가 부딪치는 소리(충격음), 두 객체 표면이 접촉한 상태에서 구르거나 미끄러지는 소리(마찰음), 작동기(actuator)인 객체가 작동할 때의 소리(작동음)로 구분된다. 현재 작동음은 단순히 실제 작동기의 소리를 녹음하여 이벤트 발생 시 재생하는 수준에 있으며, 충격음의 경우에는 많은 연구가 진행되어 두 객체 사이의 충격음 모델을 가지고서 두 객체의 접촉상황에 따라 아주 실제적인 소리를 낼 수 있는 데까지 연구가 진행되었다. 또한 마찰음은 두 객체 표면의 텍스처(texture)에 따라 접촉력을 계산하여 발생시키고 있다.

충격음과 마찰음에 대한 기존의 연구 결과를 살펴보면 다음과 같다. 두 객체 사이에 접촉이 발생하면, 충격에너지가 각 객체로 전달되어 표면에 진동을 야기한다. 이러한 진동의 주파수, 크기, 감쇠는 어느 정도 각 객체의 형태와 재질에 따라서 결정된다<sup>1</sup>. 이러한 표면 진동이 주위의 공기 압력에 요동을 가하게 되어 소리로 감지된다. 이와 같은 “충격음(contact sound)”은 수음자(listener)에게 객체의 형태나 크기, 표면 감촉뿐만 아니라 접촉 위치, 접촉력, 재질 등과 같은 많은 정보를 주게 된다. 인간은 이런 소리를 이용하여 객체 사이에 발생한 사건을 인지, 식별할 수 있는 단서로 사용할 능력을 지니고 있다<sup>2,3</sup>. 이와 같이 실생활에서 두 객체가 상호작용하는데 소리가 중요한 만큼, 성공적인 가상환경이나 시뮬레이션(simulation)에서 음향효과가 필수적인 성분이라는 것은 명확하다.

가상환경에 음향효과를 포함시키는 방법은 물

\* 한국과학기술연구원 트라이볼로지연구소

Tel. 02-958-5669, Fax. 02-958-5659

Email [ljm2293@kist.re.kr](mailto:ljm2293@kist.re.kr)

진동 소음 신호의 분석, 음향효과 등의 분야에 관심을 두고 연구활동을 하고 있다.

리적인 모델로부터 소리를 합성하는 것이다<sup>4</sup>. 접촉력과 접촉위치가 주어진다면 객체의 음향 모델로부터 직접 계산하여 적절한 충격음을 합성해낼 수 있다. 여기서 음향 모델은 객체의 충격음을 시뮬레이션하는데 사용할 수 있는 수학적 표현을 의미한다. 이것은 기록한 소리를 단순히 재생시키거나 접촉력이나 접촉위치에 따라서 변조시키는 방법과는 아주 다른 방법이다. 모델에 기초한 방법은 적절한 접촉력 프로파일(profile)로 사용함으로써 어떠한 상호작용(긋기, 부딪치기, 구르기 등)에 대해서도 소리를 발생시킬 수 있다는 장점이 있다.

충격음이 지니고 있는 유용한 정보는 Fig. 1 과 같이 물질, 상호작용, 외형 등 세가지 범주로 나눌 수 있다<sup>5</sup>. 물질은 다시 복원력, 밀도, 감쇠, 내부구조로 나뉘고, 상호작용은 형태 및 작용력, 외형은 형상, 크기, 공진공간 등으로 구분된다. 이와 같은 정보는 충격음을 다루는 많은 적용분야에 유용하게 사용되고 있다. 예를 들면, 원격조정 시스템에 유용한 피드백(feedback)이 되며, 가상환경 시스템에 사실감을 증가시킨다.

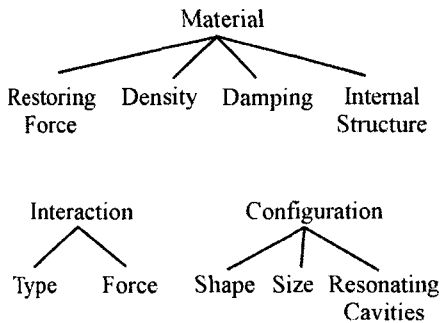


Fig. 1 Attributes of contact sounds<sup>5</sup>

가상환경에서 작동음(operative sound)의 모델링에 대한 연구는 현재까지 미미하며 본 논문에서는 AR(auto-regressive) 모델을 이용하여 작동음을 모델링하여 상황에 맞도록 재생하여 보았다.

## 2. 음향 모델

### 2.1 관련 연구 내용

음향 모델링 분야에서 많은 연구가 수행되어 왔으며, 한 예로서 악기의 음향 모델에 대한 많은

연구로 다양한 악기의 물리적인 모델링이 가능하게 되었다. Chigne 와 Doutaut 는 일차원 Euler-Bernoulli 방정식에 감쇠항과 복원력을 추가하여 나무 실로폰 막대의 음향을 모델링하였다<sup>6</sup>. 이 방정식은 경험적으로 구한 감쇠인자를 사용하여 실로폰 막대의 기하학적인 형상으로부터 유도되었다. Cook 과 Trueman 은 경험적 데이터로부터 현악기의 충격응답(impulse response)을 직접 모델링하기 위하여 주성분해석(principal component analysis)과 무한충격응답(Infinite Impulse Response, IIR) 필터 추정(estimation)기법, 왜곡선형단정(warped linear predication) 기법 등을 사용하였다<sup>7</sup>. Gaver 는 악기가 아닌 금속막대와 나무막대를 각 주파수 모드(mode)에서 지수 감쇠항을 지닌 파동방정식으로 모델링하였다<sup>5</sup>. 기본 모드와 초기 진폭, 감쇠인자는 경험적으로 결정하였고, 부차 모드는 기본 모드와의 비율로 계산하였다. van den Doel 은 물리적 근거에 의한 객체 충격음 모델을 제시하였는데, 일상적인 물체의 충격음을 효과적으로 표현할 수 있으며 실시간으로 음향을 합성하기에 알맞은 방법으로 이를 모달 공명 모델(modal resonance model)이라고 한다<sup>14</sup>. 이와 같은 연구를 바탕으로 Pai 등은 물체를 두드리거나 소리와 표면 거칠기를 고려한 긋는 소리 구르는 소리를 컴퓨터 가상 현실 공간에서 구현하였다<sup>8,9,10</sup>.

### 2.2 충격음의 모달 공명 모델

객체가 접촉으로 인하여 선형-탄성 거동을 보인다고 가정하면, 물체 표면의 진동은 다음 식(1)과 같은 식으로 표현된다.

$$\left( A - \frac{1}{c} \frac{\partial^2}{\partial t^2} \right) \cdot \mu(z, t) = F(z, t) \quad (1)$$

여기서,  $A$  는 자기수반(self-adjoint) 미분연산자,  $c$  는 매질 내의 음속 상수,  $t$  는 시간,  $z$  는 표면의 평형점,  $\mu$  는  $z$  로부터의 표면 편차,  $F$  는 접촉력이다. 외력이 없다면 식(1)의 해는 식(2)와 같다.

$$\mu(z, t) = \sum_{i=1}^{\infty} (a_i \sin(\omega_i ct) + b_i \cos(\omega_i ct)) \cdot \Psi_i(x) \quad (2)$$

여기서,  $a_i$  와  $b_i$  는 경계조건에 의하여 결정되며,  $\omega_i$  는 연산자  $A$  의 고유값(eigenvalue)와 관련

되어 있으며,  $\Psi_i(x)$  는 이에 대응하는 고유함수 (eigen-function)이다. 소리를 듣는 사람이 객체로부터 충분히 멀리 떨어져 있을 때, 식(2)의  $\mu$ 로부터 기인하는 음압(sound pressure)을 근사치로 구하면 식(3)과 같이 충격응답함수와 같다.

$$p(x,t) = \sum_{i=1}^{N_f} a_{x,i} \cdot e^{-d_{x,i}(t)} \cdot \sin(\omega_{x,i}t) \quad (3)$$

이 모델은  $N_f$  개의 주파수 모드 합으로 시간  $t$ 에서의 음압  $p$ 를 표현하고 있다. 여기서,  $x$ 는 충격력의 가해진 위치이다. 모델에서 각각의 모드  $i$ 는 주파수  $\omega_{x,i}$ , 초기진폭  $a_{x,i}$ , 지수감쇠인자  $d_{x,i}$ 로 이루어진다. 감쇠인자는 내부마찰에 기인한 물질감쇠를 모델링하는데, 이는 내부마찰 매개변수  $\phi$ 를 사용하여 식(4)와 같이 표현된다.

$$d_{x,i} = 2\omega_{x,i} \tan(\phi) \quad (4)$$

객체 표면 상의 각 위치  $x$ 에 대하여 매개변수  $\omega_{x,i}$ ,  $a_{x,i}$ ,  $d_{x,i}$ 는 모든  $N_f$ 개의 모드에 대하여 구해야만 된다. 음향모델이 충격응답 모델이기 때문에, 선형 작용력에 의해서 발생하는 충격음은 단순히 식(3)을 사용하여 Fig. 2와 같이 접촉력과 모델의 콘볼루션(convolution)으로 소리를 합성할 수 있다.

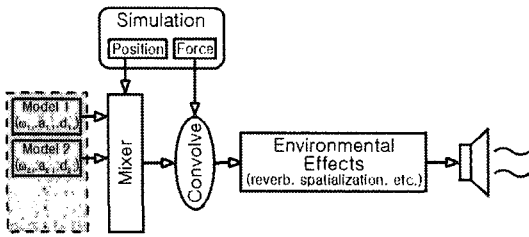


Fig. 2 Structure of convolution synthesizer for modal resonance model of contact sound<sup>5</sup>

이와는 달리 음향모델을 각 공진주파수에 대하여 IIR 필터로 변환하면 식(5)와 같은데, 이를 Fig. 3과 같이 필터뱅크(filter bank)를 구성하여 소리를 합성하는 방법이 있다<sup>11</sup>.

$$H_{x,i}(z) = \frac{1 - 2e^{-2d_{x,i}T} \cos(\omega_{x,i}T) + e^{-4d_{x,i}T}}{1 - 2e^{-2d_{x,i}T} \cos(\omega_{x,i}T)z^{-1} + e^{-4d_{x,i}T}z^{-2}} \quad (5)$$

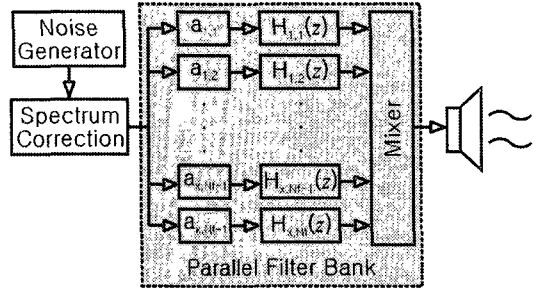


Fig. 3 Structure of filter bank synthesizer for modal resonance model of contact sound<sup>11</sup>

실제로 쇠구슬로 쇠막대를 때렸을 때 충격음과 이 충격음을 위에서 설명한 모달 공진 모델로 모델링하여 합성하여 스피커를 통하여 출력된 가상 충격음을 스펙트로그램(spectrogram)으로 Fig. 4~5에 나타내었다. Fig. 4는 실제 충격음으로, 쇠막대의 구조와 관련된 여러 개의 공진주파수와 이에 수반된 잡음이 함께 나타나는 것을 알 수 있다. 주요 공진주파수 15 개를 선택하여 각 주파수 별로 앞서 설명한 모달 공명 모델을 구하여, 합성한 결과는 Fig. 5에 나타나 있는데, 두 그래프를 비교해보면 잡음을 제외하고는 아주 유사하다는 것을 알 수 있다. 이의 검증을 위하여 특별한 음악적 경험이 없는 피실험자들에게 실제 충격음과 합성된 충격음을 들려주고 구분하는 가를 실험하였는데, 대부분의 피실험자들이 두 소리의 차이를 구분하지 못하는 아주 훌륭한 결과를 얻었다<sup>11</sup>.

## 2.3 마찰음의 모델

### 2.3.1 굽는 소리와 미끄러지는 소리

두 객체가 표면을 접촉하여 미끄러지거나 한 객체가 다른 객체를 굽는 소리는 표면 거칠기 모델로 접촉력을 구하여 합성한다. 그러나 표면 거칠기는 가상현실 화면에서 사용하고 있는 표면 텍스처 모델을 사용할 수 없는데, 그 이유는 화면용 표면 텍스처가 음향 합성에 필요한 표면 텍스처보다 분해능이 나쁘기 때문이다.

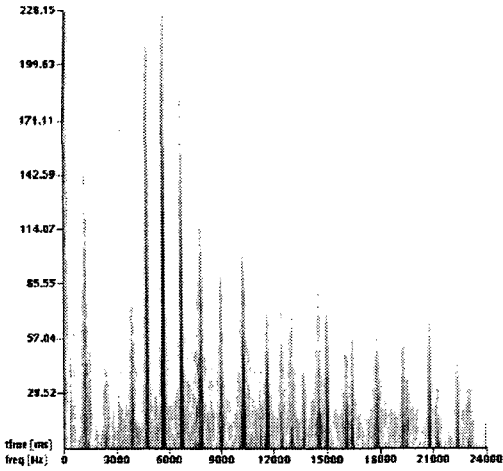


Fig. 4 Spectrogram of real sound<sup>11</sup>

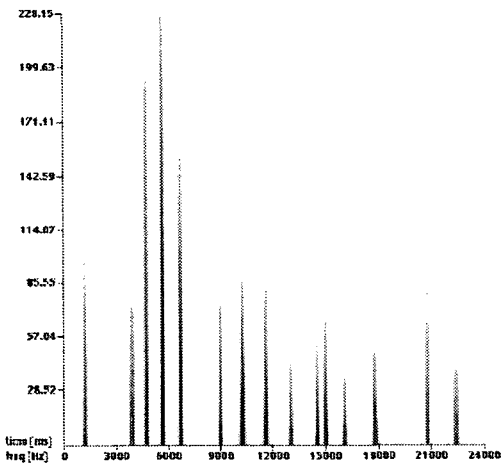


Fig. 5 Spectrogram of synthesized sound<sup>11</sup>

따라서 음향용 표면 텍스처를 따로 구하여야 하는데, 음향의 샘플링 주파수가  $f_s$  이고 두 표면이 미끄러지는 최대속도가  $v_m$  이라면  $v_m/f_s$  만큼의 분해능을 가진 표면 텍스처를 준비해야만 최대 접촉 속도에서 표면 프로파일을 음향 샘플링 속도로 샘플링 할 수 있기 때문이다. 이와 같은 조건이 만족되면 표면 거칠기와 접촉 속도, 접촉력으로 이루어진 긁는 힘(scraping force) 모델을 구할 수 있으며 이 긁는 힘을 모달 공명 모델에 입력하면 긁는 소리를 합성해 낼 수 있다.

현재 끝이 뾰족한 금속 탐침봉으로 금속표면을 긁는 소리나 금속 구슬이 금속 표면에서 미끄

러지는 소리와 같이 접촉 면적을 점으로 가정할 수 있는 경우에 한하여 이와 같은 방법으로 음향 효과를 구현하고 있다<sup>8,10</sup>. 그러나 넓은 면적을 접촉한 상태에서 미끄러지는 경우에는 두 객체의 표면 거칠기에 따라서 많은 접촉점들이 발생하며, 또 접촉점이 바뀌는 등 계산상 문제로 아직 구현되지 못하고 있다.

### 2.3.2 구르는 소리

굵는 힘처럼 구르는 힘(rolling force)은 접촉 표면의 불규칙성 때문에 발생하지만, 접촉점에서 접촉 표면 사이에 상대속도가 없다는 점이 다르기 때문에 구르는 소리와 굵는 소리는 다르다. 그러나 구르는 소리의 발생 구조에 대한 물리적 설명이 아직 완벽하지 않다. 더욱이 소리로 구름과 미끄러짐을 인지하는 방법에 대하여 정확하게 알고 있지 못하다. 구르는 소리의 주기성이 미끄러지는 소리와 구분하는데 아주 중요한 역할을 한다는 연구결과도 있으나<sup>12</sup>, 그 밖에 또 다른 측면이 있다는 것도 확실하다. 또 다른 추정은 구르는 객체는 구나 원기둥이고 구를 때 발생하는 충돌은 접촉면 앞쪽에서만 발생하기 때문에, 구르면서 발생하는 충돌이 부드럽게 된다는 가정이다. 이런 가정을 표면 프로파일의 저주파 성분이 구르는 소리에 큰 역할을 하게 만들기에, 굵는 소리 모델에 구르는 느낌을 줄 수 있는 조절 가능한 차단(cutoff) 주파수를 지닌 저 대역 통과(low-pass) 필터를 추가함으로써 구르는 소리 모델이 구현된다<sup>10</sup>. 그러나 이런 간단한 모델은 굵는 소리 모델보다는 설득력이 떨어진다. 구르는 소리를 분석해보면 미끄러지는 소리 보다는 더 강하게 모드와 관련되어 있다. 이는 순수한 모달 모델보다는 잡음(noise)로 구동된  $\gamma=2$  인 감마-톤(gamma-tone) 모델이 훨씬 더 구르는 소리를 만들어 내는 것을 미루어 보아도 알 수 있다<sup>13</sup>. 이 감마-톤 필터는 식(6)과 같은 공명(reson) 필터와 스펙트럼이 동일하다.

$$y(t) = (1 - R^2) \sin \theta \cdot x(t) + 2R \cos \theta \cdot y(t-1) - R^2 y(t-2) \quad (6)$$

$$R = 1 - d_r / 2,$$

$$\text{where, } \cos \theta = \frac{2R \cos(\omega_r)}{1 + R^2}$$

여기서,  $\omega_r$  과  $d_r$  은 공명(reson) 주파수와 감쇠비이다. 이런 공명필터의 스펙트럼은 실제 구르는

소리와 같이 객체의 공진 주파수 주변이 강조되어 있어서 많이 사용되고 있다.

### 2.4 작동음의 AR 모델

대부분의 기계는 회전운동이나 병진운동을 하거나 또는 회전, 병진운동을 동시에 하는 작동기를 가지고 있다. 이러한 기계가 작동하며 발생시키는 소리는 앞서 설명한 구조적 모드를 지닌 감쇠진동에 의한 소음보다는 기계의 작동주파수와 연관된 소음이 주성분을 이루고 그 외에 병진운동에 따른 마찰력에 의한 소음성분이 존재한다. 이와 같은 기계 작동음은 모달 공명 모델로는 합성될 수 없다.

본 논문에서는 작동음 모델링 방법의 하나로 제시된 AR 모델을 이용한 자동차 소음 합성방법<sup>14</sup> 과 같은 방법을 이용하여 휴대용 전기 연마기의 공회전 소리, 작업소리를 각각 AR 모델링하여 치아 가공 시뮬레이터를 위한 작업소리를 합성하였다. 치과 시뮬레이터는 햅틱(haptic) 장치와 별도의 컴퓨터로 구성되었으며, 음향효과를 위한 컴퓨터는 DSP(digital signal processor) 시스템을 외장하고 있다. 시뮬레이터 컴퓨터와 TCP/IP 로 통신을 통하여 작동여부, 접촉여부, 접촉면적에 대한 정보들을 얻어서, 현재 작업상태에 따라 실시간으로 DSP 시스템에 구현되어 있는 음향모델을 바꾼다. 선택된 음향모델에 따라 DSP 시스템이 실시간으로 발생시키는 음향신호는 스피커를 구동하여 소리로 바뀐다.

휴대용 전기 연마기가 공회전할 때 소리와 작업할 때 소리를 각각 식(7)과 같은 AR 모델로 모델링하여 치아를 가공하는 소리를 합성하였다.

$$y(t) = a_1 \cdot y(t-1) + \dots + a_M \cdot y(t-M) + e(t) \quad (7)$$

여기서,  $M$  은 AR 모델 차수(order)이고,  $a_1, \dots, a_M$  는 AR 계수,  $e(t)$  는 가우스 잡음(Gaussian noise)이다. 최종적으로 AR(300)모델이 사용되었으며 작업소리는 접촉 면적에 따라 입력 잡음의 크기를 연동시켜 좀더 현실감을 부여하였다. 실제 소리와 합성된 소리의 스펙트럼이 Fig. 6~7 에 나타나 있다. Fig. 6 은 공회전할 때의 경우이고, Fig. 7 은 작업할 때의 소리이며, 실선은 실제 소리이고, 흐린 점선은 합성된 소리이다. 그림에서 보는 바와 같이 실제소리와 합성된 소리의 특성은 매우

비슷하다. 단, 작업할 때의 합성된 소리가 2,500Hz 이상의 고주파 영역에서 파워가 약간 떨어지는 경향을 보여주고 있다.

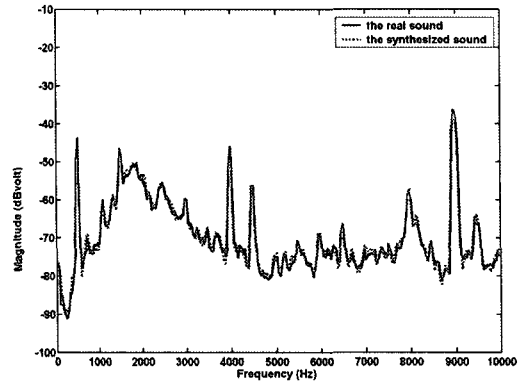


Fig. 6 Sound spectrum of idling condition

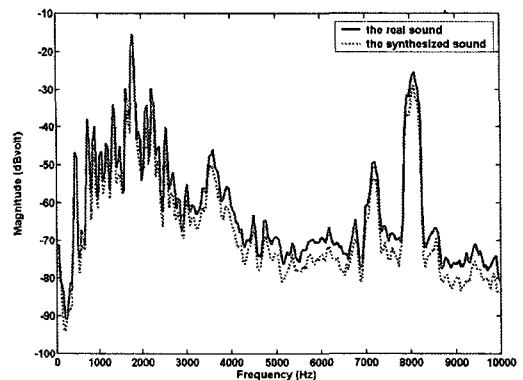


Fig. 7 Sound spectrum of drilling condition

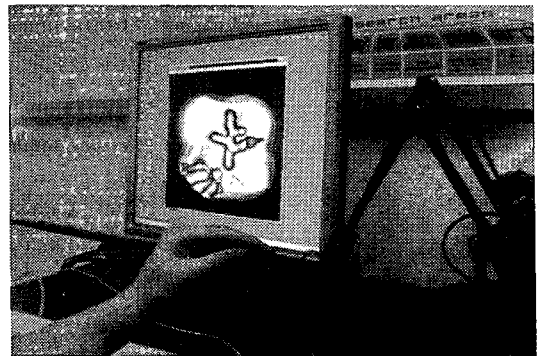


Fig. 8 Haptic rendering system for dental simulation

이와 같은 결과를 Fig. 8 과 같은 햅틱 랜더링

(haptic rendering) 시스템과 연결하여 음향효과를 부여함으로써 사용자에게 치아를 가공하고 있는 느낌을 훨씬 향상시키는 결과를 얻었다.

### 3. 결론

음향효과는 가상현실 시스템에서 정보의 전달, 존재감의 향상, 사실감, 질감의 전달과 같은 매우 중요한 역할을 한다. 가상현실에서 인간과 컴퓨터의 상호작용을 위한 음향 효과는 크게 충격음, 마찰음, 작동음으로 구분되며, 이들은 수학적인 모델로 표현되어야 한다. 본 논문에서는 각각의 모델링을 위한 기존의 최신 기술을 설명하였고, 이 중에서 가장 연구가 미진한 작동음 모델링 방법을 AR 모델을 기준으로 설명하고, 이를 치과 시뮬레이터에 초기 적용을 하였다. 치아 가공을 위한 연마기의 공회전과 작업 소리를 AR(300)모델로 구현하고, 작업 시에는 접촉 면적의 크기와 연동되게 소리의 크기를 조절하여 적용한 결과, 햅틱장치의 힘-피드백(force-feedback)과 화면상의 변형만을 구현하는 방식에 훨씬 개선된 사실감을 부여 할 수 있었다.

### 후 기

본 논문은 한국과학기술연구원의 기관고유연구사업인 “Tangible Space Initiative 기술 개발” 연구과제의 연구결과임을 밝힌다.

### 참고문헌

- van den Doel, Kees and Pai, D. K., "The Sound of Physical Shapes," Presence, Vol. 7, pp. 382 - 395, 1998.
- Gaver, W. W., "Synthesizing Auditory Icons," Proceedings of the ACM INTERCHI '93, pp. 228 - 235, 1993.
- Klatzky, R., Pai, D. K. and Krotkov, E., "Perception of Material from Contact Sound," Presence, Vol. 9, pp. 399 - 410, 2000.
- van den Doel, K., Sound Synthesis for Virtual Reality and Computer Games, Ph.D. thesis, University of British Columbia, 1999.
- Gaver, W. W., Everyday Listening and Auditory Icons, Ph.D. thesis, University of California in San Diego, 1998.
- Chaigne, A. and Doutaut, V., "Numerical Simulations of Xylophones," Journal of Acoustical Society of America, Vol. 101, pp. 539 - 557.
- Cook, P. R. and Trueman, D., "A Database of Measured Musical Instrument Body Radiation Impulse Responses, and Computer Applications for Exploring and Utilizing the Measured Filter Functions," International Symposium of Musical Acoustics, Acoustical Society of America, Woodbury, 1998.
- Pai, D. K., van den Doel, K., James, D. L., Lloyd, J. E. Lichmond, J. L. and Yau, S. H., "Scanning Physical Interaction Behavior of 3D Objects," SIGGRAPH 2001, Computer Graphics Proceedings, Los Angeles, 2001.
- Richmond, J. L. and Pai, D. K., "Active Measurement of Contact Sound," Proceedings of the 2000 IEEE International Conference on Robotics & Automation, San Francisco, pp. 2146 - 2152, 2000.
- van den Doel, K., Kry, P. G. and Pai, D. K., "Foley Automatic: Physically-based Sound Effects for Interactive Simulation and Animation," SIGGRAPH 2001, Computer Graphics Proceedings, Los Angeles, 2001.
- Munteanu, A. D., Guggiana, V., Schauer, H., Motavalli, M. and Rauterberg, M., "Automatic Impact Sound Generation for Using in Non-visual Interfaces," The First Annual ACM Conference on Assistive Technologies-ASSETS'94, New York, pp. 100 - 106, 1994.
- Houben, M. M. J., Hermes, D. J. and Kohlraush, A., "Auditory Perception of the Size and Velocity of Rolling Balls," IPO Annual Progress Report, Vol. 34, 1999.
- Hermes, D. J., "Synthesis of the Sounds Produced by Rolling Balls," Internal IPO report, No. 1226, IPO, Center for User-System Interaction, Eindhoven, The Netherlands, 2000.
- Cho, J. H., Lee, J. M. and Hwang, Y., "A Study on the Reproduction of Acoustic Characteristics of a Car's Exhaust Noise Using Digital Filtering Technique," Transaction of the Korean Society of Automotive Engineers, Vol. 1, No. 3, pp. 55 - 62, 1993.