

## Efficient Sound Control Method in Virtual Environments Using Raytracing Based Diffraction

Jong-Hyun Kim\*, Jong-In Choi\*\*

\*Professor, School of Software Application, Kangnam University, Yongin, Korea

\*\*Professor, Dept. of Digital Media, Seoul Women's University, Seoul, Korea

### [Abstract]

In this paper, we propose diffraction-based sound control method to improve sound immersion in a virtual environment. The proposed technique can express the wave and flow of sound in a physical environment and a pattern similar to diffraction in real-time. Our approach determines whether there is an obstacle from the location of the sound source and then calculates the position of the new sound reflected and diffracted by the obstacle. Based on ray tracing, it determines whether or not it collides with an obstacle, and predicts the sound level of the agent behind the obstacle by using the vector reflected and refraction by the collision. In this process, the sound attenuation according to the distance/material is modeled by attenuating the size of the sound according to the number of reflected/refracted rays. As a result, the diffraction pattern expressed in the physics-based approach was expressed in real time, and it shows that the diffraction pattern also changes as the position of the obstacle is changed, thereby showing the result of naturally spreading the size of the sound. The proposed method restores the diffusion and diffraction characteristics of sound expressed in real life almost similarly.

▶ **Key words:** Sound propagation, Refraction vector, Sound flow, Sound diffraction, Sound wave

### [요 약]

본 논문에서는 가상환경에서 사운드 몰입을 개선시키기 위한 회절 기반의 사운드 제어 기법을 제안한다. 제안하는 기법은 물리적 환경에서 사운드의 파동과 흐름 그리고 회절과 유사한 패턴을 실시간으로 표현할 수 있다. 우리의 접근 방식은 사운드 근원지로부터 장애물이 있는지 판단을 한 뒤, 장애물로 인해 굴절과 회절된 새로운 사운드의 위치를 계산한다. 레이트레이싱 기반으로 장애물과의 충돌 여부를 판단하고, 충돌에 의해 반사와 굴절된 벡터를 이용하여 장애물 너머에 있는 에이전트에게 들리는 사운드의 크기를 예측한다. 이 과정에서 반사와 굴절된 레이의 개수에 따라 사운드의 크기를 감쇠시켜 거리와 재질에 따른 사운드 감쇠를 모델링한다. 결과적으로 물리 기반 접근법에서 표현되는 회절 패턴을 실시간으로 표현했으며, 장애물의 위치가 변경됨에 따라 회절 패턴도 변경됨을 보여주고, 이에 따라 사운드의 크기가 자연스럽게 확산되는 결과를 보여준다. 제안하는 방법은 현실에서 표현되는 소리의 확산과 회절 특징을 거의 유사하게 복원해냈다.

▶ **주제어:** 소리 전파, 굴절 벡터, 소리의 흐름, 소리의 회절, 소리의 파동

- 
- First Author: Jong-Hyun Kim, Corresponding Author: Jong-In Choi
  - \*Jong-Hyun Kim (jonghyunkim@kangnam.ac.kr), School of Software Application, Kangnam University
  - \*\*Jong-In Choi (funtech@swu.ac.kr), Dept. of Digital Media, Seoul Women's University
  - Received: 2022. 08. 22, Revised: 2022. 10. 31, Accepted: 2022. 11. 14.

## I. Introduction

최근에 가상공간을 실제 현실처럼 표현하는 물리 기반 환경이 중요하게 여겨지면서 다양한 신체적 감각을 기반으로 몰입을 올리려는 시도가 있었다[1]. 그 중에서 제일 많은 비중을 차지하는 감각은 대부분은 시각적인 요소이다. 이를 뒷받침 하듯이 게임이나 가상현실(Virtual Reality, VR) 콘텐츠에서는 시각적인 요소를 극대화하기 위한 다양한 기술들이 개발되고 있다. 최근에는 HMD(Head-mounted display)와 Non-HMD 사용자의 몰입을 개선하려는 연구들도 제안되었다[2]. 하지만, 시각적인 요소만으로 가상공간에서 느껴지는 현실적인 감각을 모두 표현하는 데는 한계가 있다.

실제로 시각적인 요소에 의존하지 못하는 상황이나, 보아야 할 것이 보이지 않는 상황, 혹은 시야가 차단되거나 시야를 인지 못할 상황이 오면 청각적인 요소는 매우 중요하게 된다. 어두컴컴한 환경에서 소리가 임의의 순서로 재생되면 사람은 그에 따라 반응을 하고, 고개를 돌리게 된다. 또한, 장애물로 인해 시야가 가려져도 사운드의 회절로 인해 장애물 너머에 있는 사람은 그 소리를 듣게 된다 [3,4]. 하지만, 많은 VR 콘텐츠는 이러한 청각적인 특징을 고려하지 못하고 있으며, 특히 물체와의 상호작용에 따른 소리의 변화와 현실과 유사한 물리적인 소리의 특징이 반영되고 있지 않다. 만약에 이 소리의 부재가 지속되면 시각적인 요소가 뛰어날지라도 사용자가 소리에서 나타나는 부자연스러움 때문에 몰입이 저하되고 인지의 부조화를 느낄 수 있다. 반대로 청각적인 요소가 잘 표현된다면 시각적인 요소에 더해지면서 몰입이 개선될 수 있고, 청각적인 요소만으로도 다양한 정보들을 사용자에게 제공함으로써 새로운 방식으로 콘텐츠를 즐기게 할 수 있다.

최근에 물리 기반 사운드 합성기법을 이용하여 애니메이션과 VR에 활용하려는 시도가 꾸준히 발표되고 있다[5]. 일반적으로 VR에서 활용되는 오디오는 스테레오 헤드폰 환경을 기반으로 제공되기 때문에 다채널 오디오 신호로는 현장감 있는 사운드를 제공하지 못하므로 장면과 장면 사이에 갭이 존재한다. 이 문제를 완화하고 사실적인 오디오 사운드를 제공하기 위해 바이노럴(Binaural) 렌더링 기법을 이용한 연구들도 있다[6]. 그러나 바이노럴 렌더링의 선형 컨볼루션은 계산 복잡도가 높기 때문에 실시간으로 시행하기 어려운 접근법이다. 최근에 Kim 등은 주파수 영역에서 바이노럴 렌더링을 구현하고, 저주파 영역과 고주파 영역인 두 가지 주파수 영역에서 별도로 바이노럴 렌더링을 수행하였다[7]. 이러한 접근법에서 활용하는 스테레

오 헤드폰 기반의 VR 서비스에서는 사용자의 위치(방위각, 거리)의 변화에 관계없이 항상 일정한 소리 장면만을 가지는 오디오 신호를 사용한다. 결과적으로 이러한 접근법은 VR 서비스에서 사운드와 장면 사이에 또 다른 갭을 만들어 낸다. 이를 해결하기 위해 사용자의 방위각 변화에 대한 바이노럴 사운드의 SCC(Sound scene control) 기법과 사용자의 거리 변화에 대한 사운드 레벨 제어 기법이 제안되었다[8]. 위에서 언급한 방법들은 대부분 스테레오 사운드에만 집중하였으며, 사운드의 반사와 굴절, 그리고 회절에 의한 표현과 이를 통해 사운드의 크기를 제어할 수 있는 기법은 아니다. 이러한 제한적 표현은 가상환경에서 몰입을 저하시키는 요인이 된다.

앞에서 언급한 접근법들은 계산량이 크기 때문에 혼합 현실 및 실시간 애플리케이션에 활용이 어려우며, 높낮이가 다양한 음원 데이터에 대한 사운드 반사, 굴절 및 회절에 대한 특징을 다루지 않았기 때문에 특정 장면에 대해서만 제한적으로 활용이 가능하다. 이러한 제한점을 해결하기 위한 본 논문의 기술적 기여도는 다음과 같다 :

- 장애물에 의해 표현되는 회절 패턴을 근사화된 사운드 회절 위치를 통해 효율적으로 계산하는 방법
- 소리의 반사 및 회절에 따라 소리의 크기를 제어하는 방법
- 소리의 스타일을 다양한 음원에 전달하는 방법

### 1. Related work

증강현실(Augmented Reality, AR)은 현실 세계에 가상의 물체를 실시간으로 중첩시켜 현실감을 향상시키는 방법이다. AR을 체험하는 사용자는 주로 현실을 있는 환경을 그대로 보여주기 때문에 걷기 등의 동작을 취하기 쉽다. 하지만, 실제 환경에 있는 정보의 일부를 가상의 정보로 대체하면 사용자는 인지 및 행동을 결정하는데 있어서 어려움을 느낀다[9]. VR을 체험하는 사용자도 장애물과의 충돌을 피하기 위해서 실제 물체 주위에 경계 상자를 표시하고, 카메라를 사용하여 가상환경의 장애물 위치에 실제 장애물 이미지를 중첩하여 표현한다[10]. 하지만, 대부분 장애물은 시각적인 요소만이 아닌, 청각적인 요소로도 표현되기 때문에 청각을 제외하고는 온전한 장애물 인식은 불가능하다. 예를 들어, 보이지 않는 공간에서 사운드의 반사 및 회절로 인해 소리가 들리기 때문이다.

사용자가 가상환경에서 소리를 어떻게 인지하고 어떤 효과를 주는지 고려하는 것은 사용자의 몰입을 높이는 데 없어서 안 될 요소이다. 사운드는 사용자에게 공간의 깊이를 전달할 뿐만 아니라 사물과의 상호작용을 가능하게 하

여 가상공간을 보다 몰입감 있게 만들어 공간 인식을 돕는다. 또한, 가상환경에서 소리의 움직임과 영상의 움직임을 동기화시키는 음원을 설정하면, 소리가 영상과 비동기식일 때보다 VR에 의한 멀미를 줄일 수 있다.

다양한 연구자들은 VR에서 청각 자극의 실현 가능성과 VR 경험에 미치는 영향에 대해 연구를 진행했다. Fujioka 등은 재질을 인지하는 분야에서 다감각 통합에 대한 연구를 진행했다[11]. 이 방법에서는 6가지 재질의 시각적 외관과 8가지 재질의 충격음을 결합한 시청각 실험을 수행하여 시각적 자극과 청각적 자극이 시각과 크게 연관이 있음을 보여주었다. 경우에 따라서 시각 정보가 모호한 경우 청각 정보가 재질을 결정하는 지배적인 요소가 된다. Kern 등은 사용자의 움직임과 동기화된 발자국의 유무가 사용자의 VR 경험에 어떤 영향을 미치는지에 대한 실험을 수행했으며, 발자국의 존재가 몰입감과 현실감을 향상시킨다는 것을 증명했다[12]. 이와 같이 시각 정보가 공간 인식의 대부분을 차지하는 가상환경에서 사용자는 청각 정보를 보조 수단으로 사용하면 가상환경을 올바르게 정확하게 인식할 수 있다.

이러한 소리를 활용하여 환경을 파악하고 장애물을 피하는 VR의 예는 시각 장애인을 위한 지팡이 훈련 체험이다. Lahav 등은 가상환경을 이용하여 시각 장애인을 위한 공간 인지 매핑과 방향을 습득할 수 있는 방법을 제시했다[13]. Inman 등은 시각 장애인을 위한 공간 인식 VR 훈련을 수행했으며, 소리가 장애물을 감지하고 회피하는데 사용할 수 있는 특정 공간 방향을 제공한다는 것을 증명했다[14]. 또한, 발자국 소리를 들으면 다양한 지형의 재질을 판단하고 이동 속도를 예측하는데 도움이 된다는 것을 실험을 통해 보여주었다. Siu 등은 시각 장애인이 복잡한 가상환경에서 보행하기 위한 소리와 진동기반 탐색 시스템을 제안했다[15]. 이 방법에서는 가상환경을 탐색하기 위한 5가지 재질을 선택하고(타일, 콘크리트, 금속, 나무, 카펫), 사용자가 지팡이 모양의 컨트롤러를 두들겨 소리와 진동을 수신하게 설계되었다. 소리는 시간적 속성을 가지고 있기 때문에 소리의 존재는 무언가가 일어나고 있다는 것을 알려주는 중요한 정보이다.

## II. The Proposed Scheme

### 1. Reflection and refraction of sound according to obstacles

본 논문에서는 장애물에 의해 표현되는 사운드의 반사와 굴절을 계산하기 위해 몇 가지 가정을 한다. 이 가정을

기반으로 장애물과 사용자(청중)는 다음과 같이 배치한다 (Fig. 1 참조) : 1) 지형은 평면과 같이 평평하고, 2) 소리에는 방향성이 있으며, 3) 소리 크기의 높낮이는 일정하다 (예 : 뚜~같은 반복되는 소리).

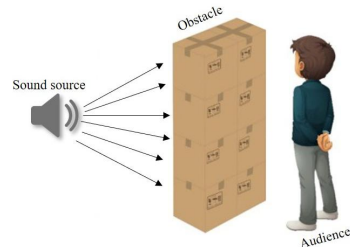


Fig. 1. Scene configuration.

Fig. 1에서 보듯이 장애물 판단을 위해 소리의 근원지 (Sound source)로부터 사운드가 들려야 하는 방향을 지정하여 레이를 쏘준다. 이때, 레이는 소리의 파동을 의미하며 레이의 생성 범위는 소리의 방향으로부터 반구 (Hemisphere) 방향으로 지정하였다. 본 논문에서는 사운드 근원지와 장애물, 그리고 사용자와의 거리를 가중치로 최대 20개의 무작위 추출(Random sampling)을 통해 레이를 생성하였다 (Fig. 2 참조).

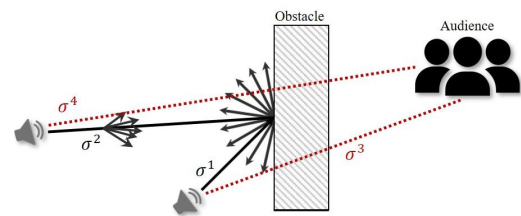


Fig. 2. Hemispherical ray direction considering the relationship between the sound source, obstacles and audience.

위 그림에서  $\sigma$ 는 사운드 근원지로부터의 거리 값이며, 본 논문에서는 이 값을 이용하여 에너지를 적분하기 위한 공간인 반구 형태의 레이 생성 영역을 제어한다.  $\sigma^{1,2}$ 는 사운드 근원지와 장애물과의 거리이며, 거리가 가까운  $\sigma^1$ 은 차폐될 영역이 가깝기 때문에 온전한 반구 형태로 적분 영역을 설정하고, 거리가 먼  $\sigma^2$ 는 상대적으로 적분 영역을 작게 설정한다. 거리가 먼 경우 상대적으로 차폐될 영역이 작기 때문에 온전한 반구 형태로 적분을 하지 않아도 충분한 사운드 에너지 계산이 가능하다. 만약에 장애물이 없다면 사운드 근원지와 사용자 간의 거리인  $\sigma^{3,4}$ 를 이용하여 적분 영역을 설정한다.

사운드 근원지로부터 발생된 레이가 장애물과 충돌을 하게 되면 충돌된 지점을 기준으로 레이를 굴절시킨다. 소

리는 빛과 마찬가지로 물체에 충돌되면 반사와 굴절 현상이 일어나기 때문이다. 이 같은 특징을 표현하고자 장애물과 충돌 이후의 굴절은 무작위로 굴절되며, 이 같은 결과가 의미하는 바는 소리 에너지의 굴절을 나타낸다. 최종적으로 장애물로 인한 소리 흐름의 변화는 굴절된 레이들로부터 대략적인 회절 소리를 근사하는데 사용된다. 여기서 근사화된 회절 소리는 원본 소리로부터 굴절된 소리이며, 이것을 기반으로 장애물 너머에 있는 청중에게 들리는 소리를 생성한다. 본 논문에서는 좀 더 정확하게 장애물의 기울기와 소리 방향의 연관성을 고려하기 위해, 장애물 표면의 법선벡터와 소리의 방향 벡터 사이에서 소리의 굴절 벡터를 생성하도록 하였다 (Fig. 3 참조). 장애물과의 충돌로 인해 굴절된 벡터들은 *ohm* 영역 내에서 굴절 형태로 변환되어 레이가 생성된다.

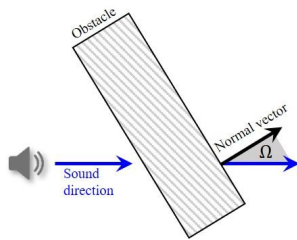


Fig. 3. Calculating range(*ohm*) of refracted vector.

## 2. Calculating diffraction of sound caused by obstacle

본 논문에서는 사운드의 회절을 계산하기 위해 앞에서 추가적으로 생성된 레이를 활용한다. 청중에게 들리는 근사화된 회절 소리의 크기는 장애물과 충돌된 횟수에 비례하며, 원본 소리의 크기를 기준으로 감쇠시켰다. 또한, 장애물이 소리의 근원지로부터 얼마나 가려졌는지에 따라 소리의 크기를 제어한다. 예를 들어, 장애물이 반만 가려진 경우, 장애물이 다 가렸을 경우, 장애물이 가려지지 않았을 경우마다 소리의 감쇠 정도가 달라야 한다. 본 논문에서는 이를 고려하기 위해 근사화된 회절 소리의 크기인  $\Gamma_{cp}$ 를 아래와 같이 계산한다 (수식 1 참조).

$$\Gamma_{cp} = (1 - \alpha h)^2 \quad (1)$$

여기서  $h$ 는 장애물에 충돌된 레이의 개수이며,  $\alpha$ 는 장애물 재질에 따른 소리의 감쇠 정도이고, 0에서 1사이의 값을 가진다. 근사화된 회절 소리의 위치는 아래와 같이 계산한다 (수식 2 참조).

$$\Gamma_{cp}^p = \frac{\sum_{k=0}^{\|V_h\|} V_h^k + \left( \sum_{k=0}^{\|V_{nh}\|} V_{nh}^k w \right)}{n} \quad (2)$$

여기서  $n$ 은 굴절 벡터의 개수,  $V_h$ 와  $V_{nh}$ 는 각각 장애물과 충돌된 굴절 벡터와 충돌되지 않은 벡터를 의미한다.  $w$ 는 장애물의 재질에 따른 굴절 정도를 의미한다. 이로써 장애물이 있을 시에 장애물로 인한 회절 소리와 감쇠, 그리고 소리 방향에 따른 굴절 위치를 계산함으로써 장애물 너머 생성되는 소리의 굴절 흐름을 예측할 수 있다. 근사화된 회절 소리 값 계산 후 장애물을 기준으로 사용자가 어디 있는지를 판별해서 장애물 너머에 위치해 있다면 회절 소리만을 이용하여 소리의 크기를 계산하고, 그렇지 않으면 원본 소리를 고려하여 소리의 크기를 계산한다.

## 3. Efficient style transfer of sound

일반적으로 스타일 전달은 컴퓨터비전에서 영상의 스타일을 다른 영상으로 복사하는 것을 의미한다. 본 논문에서는 음원의 스타일을 다른 음원으로 손쉽게 전달할 수 있는 방법에 대해 설명한다. 앞에서 설명한 소리의 강도 맵은 레이트레이싱을 통해서만 계산하며, 이렇게 나온 맵은 레이를 통해 얻어진 사운드 에너지와 유사하다. 물론, 앞에서 설명했듯이 렌더링분야에서 사용하는 일반적인 레이트레이싱은 이지만 (수식 3 참조)[21], 본 논문에서는 소리의 굴절 및 회절을 계산하기 위해 반구영역에 해당하는 적분 공간을 제어하고(수식 3에서  $H^2$ ), 장애물의 재질인 *ohm*에 따라 사운드 에너지가 달라진다.

$$L(x, \omega) = L_e(x, \omega) + \int_{H^2} f_r(x, \omega' \rightarrow \omega) L(x^*(x, \omega'), -\omega') \cos \theta' d\omega' \quad (3)$$

본 논문에서는 계산된 소리의 강도 맵으로부터 음원의 스타일을 효율적으로 표현하기 위해 다음과 같은 수식을 이용한다 (수식 4 참조).

$$\frac{\Gamma^{i,j\nu}}{\max(\Gamma^{i,j \in N\nu})} \quad (4)$$

여기서  $\Gamma^{i,j}$ 는 사운드의 굴절과 회절로 계산된 사운드의 강도 맵이며,  $\nu$ 는 새로운 소리에 적용시킬 입력 음원 파일이다. 예를 들어,  $\nu$ 가 캐논 음원 파일이라면, 이 값에 기존 맵에서 측정된 가중치 값을 곱해주며, 이 값들 중 가장 큰 값으로 나누어 정규화를 시킨다.

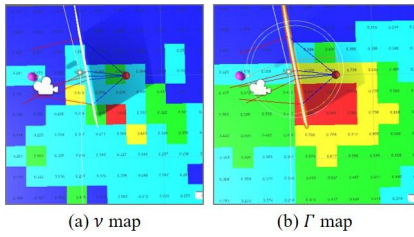


Fig. 4. Maps with different characteristics ( $\nu$  : Variations On The Canon).

Fig. 4에서 보듯이 두 개의 맵들은 서로 다른 특징을 가지므로 사실상 다른 소리이다. 본 논문에서는 이러한 특징을 효율적으로 전달시키기 위해 앞에서 수식 4를 통해 소리를 제어한다. 이 과정에서  $\nu$  맵은 장애물을 고려하지 않고, 샘플링 위치로부터  $\nu$  근원지까지의 유클리드 거리를 통해 사운드의 크기를 계산한 맵이다. 위 그림은 음의 높낮이가 일정한 소리뿐만 아니라, 캐논변주곡과 같이 임의의 소리 데이터에서도 사운드 제어가 가능하다는 결과를 보여주고 있는 그림이다. 그림 4a는 캐논변주곡이 거리에 따라 들리는 맵은 표현한 것이며, 그림 4b는 음의 높낮이가 일정한 소리에서 맵을 추출한 결과이다. 본 논문에서는 그림 4b에 대한 맵을 활용하여 복잡한 소리 데이터에서도 사운드가 제어될 수 있도록 하였다.

### III. Results

본 연구의 결과들을 만들기 위해 실험한 환경은 Intel Core i7-7700K CPU, 32GB RAM, Geforce GTX 1080Ti GPU가 탑재된 컴퓨터를 이용했다. 본 논문에서 생성된 회절 소리가 장애물을 잘 반영하여 현실에서 장애물 너머에 있는 소리의 흐름과 유사한 패턴을 만들어 냈는지 확인해 볼 필요가 있다. 청중의 역할을 하는 객체의 위치를 격자의 노드 위치라고 가정하고, 모든 노드에서 소리의 크기를 계산하였다. 이 과정에서 들리는 소리가 벽에 다가갈수록 커지고, 벽이 얼마나 가려졌는지에 따라 같은 위치에서도 벽에 많이 가려졌다면 작게, 벽이 조금만 가려졌다면 크게 들리는 것을 실험을 통해 확인하였다. 또한, 장애물의 기울기에 따라 회절 소리의 위치가 바뀌면서 왼쪽에서 크게 들릴지, 오른쪽에서 크게 들리는지 소리의 크기가 스테레오 사운드 형식처럼 반영되는 것을 확인했다. 다만, 이 부분에서 소리의 크기를 직관적으로 확인하기 위해 레인보우 형태의 크기 맵으로 시각화하여 결과를 분석했다. Fig. 5는 본 논문에서 제안한 방법을 통해 생성된 소리의 크기 맵이다. 작게 삽입된 그림은 소리를 계산할 때 사용한 장

면이며, 장애물을 나타내는 흰색 벽을 기준으로 오른쪽이 사운드 근원지이다. 앞에서 설명했듯이 사운드 근원지로부터 생성된 레이의 굴절로 인해 벽면에 의해 가려진 부분에서도 사운드의 크기가 시각화가 된 것을 볼 수 있다 (Fig. 5 참조). 본 논문에서는 장면을 구성할 때는 저해상도 격자로 공간을 이산화하고, 실제 사운드 샘플링은 해당 격자를 스무딩(Smoothing)하여 부드럽게 변화된 사운드의 크기 값을 이용하였다. 맵에서는 소리가 클수록 빨간색으로, 소리가 작을수록 파란색으로 표현했다.

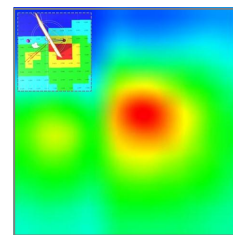


Fig. 5. Generated magnitude map of sound.

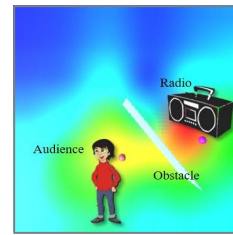


Fig. 6. Sound diffraction with our method.

장애물에 따른 소리의 굴절과 반사로 인해 나타나는 현상인 실제 회절과 본 논문에서 제안한 방법을 비교해보면 형태가 유사하게 나타나는 것을 볼 수 있다. Fig. 6에서 보면 사운드의 근원지인 라디오 부근에서 소리의 크기가 크고, 멀어질수록 사운드의 감쇠가 표현된 것을 보여주고 있다. 뿐만 아니라 추가적인 레이를 기반으로 사운드의 회절을 잘 표현하였다. Fig. 5와 비교했을 때 사운드의 회절과 굴절로 인한 사운드의 크기가 좀 더 명확하게 보이는 것을 확인할 수 있다 (Fig. 6 참조). 굴절로 인해 사운드의 크기라 노란색으로 표현되었을 뿐만 아니라, 회절로 인해 생성된 사운드가 크기가 장애물 형태를 고려하여 청록색으로 표현되었다.

Fig. 7은 장애물의 회전을 고려하여 소리의 크기와 회절의 표현을 확인하기 위한 결과이다. 각 장애물의 회전 각도에 따라 회절 소리가 변화하고, 이에 따라서 소리의 흐름도 이동되는 것을 확인할 수 있다.

Fig. 8은 본 논문에서 제안한 기법을 통해 캐논 변주곡의 사운드를 제어한 결과이다. Fig. 8a는 캐논변주곡의 사운드 크기이며, 그림에서 보듯이 주변 환경을 고려하지 않

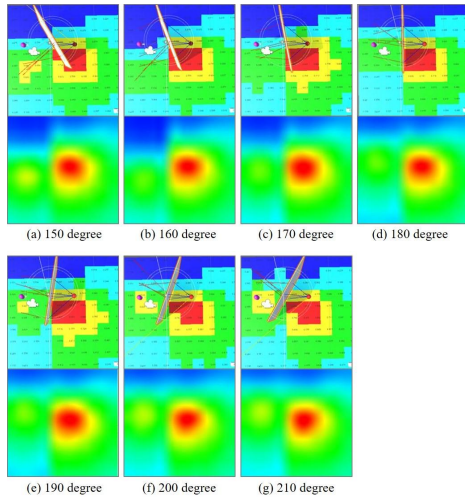


Fig. 7. Diffraction map that changes with rotating obstacles.

고 사운드 근원지와 거리에 따라 소리 크기가 제어되는 결과를 보인다. 그에 비해 본 논문에서 제시한 음원 스타일 전달 방식은 캐논 변주곡의 스타일은 유지할 한 채, 주변 장애물과의 관계에 따라 사운드의 굴절 및 회절이 잘 표현되었다. 이 방식은 기존의 인공지능이나 컴퓨터 비전에서 사용하는 영상처리 기반의 스타일 전달과는 정확도 측면에서 다소 떨어지지만, 임의의 음원파일에 대해서 실시간으로 손쉽게 소리를 제어할 수 있다는 점에서 큰 장점을 가진다.

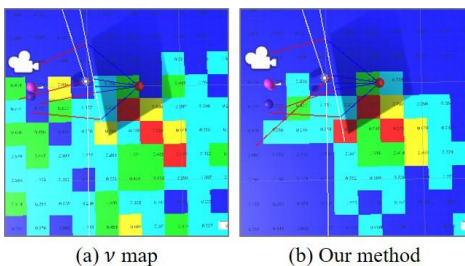


Fig. 8. Result of controlling the sound of the Canon Variations.

#### IV. Conclusions

본 논문에서는 소리의 근원지로부터 장애물 여부를 판단하여 장애물과 사운드 레이가 충돌할 경우 회절 소리를 생성하여 소리의 감쇠, 소리의 위치를 근사할 수 있는 방법을 제시했다. 또한, 회절 소리의 위치를 계산하기 위해 소리의 굴절 벡터를 활용하는 방법을 제시하였고, 임의의 사운드 파일을 손쉽게 제어할 수 있는 효율적인 스타일 전달 기법을 제시하였다. 이를 통해 실제 소리 현상인 회절

과 같은 소리의 파동 및 흐름을 효율적으로 모델링했으며, 소리의 스타일 전달을 통해 임의의 사운드에 대해서도 소리의 회절 및 굴절을 실시간으로 표현할 수 있었다. 기존 기법들은 전처리 시간이 많이 필요하지만, 본 논문에서는 간단한 레이 기반으로 사운드의 회절을 추론하기 때문에 효율적으로 처리가 가능하다. 뿐만 아니라, 높낮이가 같은 음원으로 부터 복잡한 사운드 데이터의 크기를 간단하게 제어할 수 있기 때문에 실시간 애플리케이션에서 효율적으로 활용할 수 있다.

그럼에도 불구하고 몇 가지 한계점이 존재한다. 지형의 높낮이를 고려하지 못하고 평평한 지면이라는 전제로 사운드의 크기를 계산했기 때문에, 향후 복잡한 지역에서 사운드를 제어할 수 있도록 알고리즘을 확장해야 한다. 또한, 장애물의 재질을 간단한 변수로 제어하기 때문에 좀 더 다양한 재질을 통한 사운드를 표현하기에는 한계가 있다. 향후 이러한 문제를 완화시킬 수 있도록 알고리즘을 확장할 계획이며, 나아가 소리의 파동과 흐름을 평면 형태가 아닌 3차원 공간상에서도 공기와 바람으로 인해 효율적으로 전파될 수 있도록 확장할 계획이다.

#### ACKNOWLEDGEMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2022R1F1A1063180).

#### REFERENCES

- [1] Wang, Yong, Uma Jayaram, Sankar Jayaram, and Kevin Lyons. "Physically based modeling in virtual assembly." In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, pp. 295-305, 2001. DOI:10.20870/IJVR.2001.5.1.2679.
- [2] Gugenheimer, Jan, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. "Sharevr: Enabling co-located experiences for virtual reality between hmd and non-hmd users." In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 4021-4033, 2017. DOI: 10.1145/3025453.3025683.
- [3] Van Den Doel, Kees, Paul G. Kry, and Dinesh K. Pai. "FoleyAutomatic: physically-based sound effects for interactive simulation and animation." In Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp.

- 537-544, 2001. DOI: 10.1145/383259.383322.
- [4] De Poli, G. and Rocchesso, D., 1998. "Physically based sound modelling", *Organised Sound*, Vol. 3, No. 1, pp.61-76, 2001. DOI: 10.1017/S1355771898009182.
- [5] Schröder, Dirk. "Physically based real-time auralization of interactive virtual environments". Vol. 11. Logos Verlag Berlin GmbH, 2011.
- [6] Breebaart, Jeroen, Jürgen Herre, Craig Jin, Kristofer Kjörling, Jeroen Koppens, Jan Plogsties, and Lars Villemoes. "Multi-channel goes mobile: MPEG Surround binaural rendering." In *Audio Engineering Society Conference: 29th International Conference: Audio for Mobile and Handheld Devices*. Audio Engineering Society, 2006.
- [7] Kim, Kwangki. "A study on complexity reduction of binaural decoding in multi-channel audio coding for realistic audio service.", 2015. DOI: 10.12988/ces.2016.512315.
- [8] Kim, Kwangki. "Sound scene control of multi-channel audio signals for realistic audio service in wired/wireless network." *International Journal of Multimedia and Ubiquitous Engineering*, Vol. 9, No. 2, pp. 253-258, 2014. DOI: 10.14257/ijmue.2014.9.2.25.
- [9] Nagao, Katashi, Menglong Yang, and Yusuke Miyakawa. "Building-scale virtual reality: reconstruction and modification of building interior extends real world." *International Journal of Multimedia Data Engineering and Management*, Vol. 10, No. pp. 1-21, 2019.
- [10] Połap, Dawid, Karolina Kęsik, Kamil Książek, and Marcin Woźniak. "Obstacle detection as a safety alert in augmented reality models by the use of deep learning techniques." *Sensors*, Vol. 17, No. 12, 2017. DOI: 10.3390/s17122803.
- [11] Fujisaki, Waka, Naokazu Goda, Isamu Motoyoshi, Hidehiko Komatsu, and Shin'ya Nishida. "Audiovisual integration in the human perception of materials." *Journal of Vision*, Vol. 14, No. 4, pp. 12-12, 2014. DOI: 10.1167/14.4.12.
- [12] Kern, Angelika C., and Wolfgang Ellermeier. "Audio in VR: effects of a soundscape and movement-triggered step sounds on presence." *Frontiers in Robotics and AI*, Vol. 7, pp. 20, 2020. DOI: 10.3389/frobt.2020.000.
- [13] Lahav, Orly, and David Mioduser. "Multisensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping—a case study." In *EdMedia+ Innovate Learning*, pp. 1046-1051. Association for the Advancement of Computing in Education, 2001.
- [14] Inman, Dean P., Ken Loge, and Aaron Cram. "Teaching orientation and mobility skills to blind children using computer generated 3D sound environments." *Georgia Institute of Technology*, 2000.
- [15] Siu, Alexa F., Mike Sinclair, Robert Kovacs, Eyal Ofek, Christian Holz, and Edward Cutrell. "Virtual reality without vision: A

haptic and auditory white cane to navigate complex virtual worlds." In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pp. 1-13. 2020. DOI: 10.1145/3313831.3376353.

## Authors



Jong-Hyun Kim received the B.A. degree in the department of digital contents at Sejong University in 2008. He received M.S. and Ph.D. degrees in the department of computer science and engineering at Korea University,

in 2010 and 2016. Prof. Kim is an associate professor in the school of software application in Kangnam University. His current research interests include fluid animation and virtual reality.



Jong-In Choi received the Ph.D. in 2016 from the Department of Computer Science and Engineering from Korea University. He joined Nexon Korea as a lead client programmer. And he also has worked at

NCSOFT Korea R&D Center and AI Center. Now he is a professor at the department of digital media design and application in Seoul Women's University.