

# 증강현실에서 객체와 오디오의 상호작용

조현욱\* · 이종근\* · 이승진\* · 이종혁\*

\*경성대학교

## Interaction between Object and Audio in Augmented Reality

Hyun-wook Cho\* · Chong-geun Lee\* · Seung-jin Lee\* · Jong-hyeok Lee\*

\*Kyungsung University

E-mail : jhlee@star.ks.ac.kr

### 요 약

최근 멀티미디어 기술의 발달, 특히 음향 기술의 급격한 발달과 더불어 고품질 오디오에 대한 요구와 함께보다 현실감 있는 오디오를 재생하기 위한 실감 오디오기술 개발이 요구되고 있다. 이러한 요구를 만족시키기 위해 사용자의 가상현실 및 증강현실에서 실감나는 오디오 효과를 제공해 줄 수 있는 3차원 오디오에 대한 연구가 활발히 진행되고 있다.

본 논문에서는 증강현실에서 좀 더 나은 오디오 기술을 적용하여 실감나는 오디오 효과를 제공해 줄 수 있는 방법을 연구하고자 하였다. 연구한 내용은 가상세계와 실제세계의 현실감을 제공하기 위하여 마커 위에 띄워진 3D 모델의 움직임에 따라서 움직임에 맞는 사운드, 즉, 거리, 각도 등의 변화에 따른 사운드의 크기 및 피치 변화를 줄 수 있도록 하였다.

### 키워드

증강현실(Augmented Reality: AR), 증강학습, Audio Effect, Goblin XNA

## I. 서 론

새로운 디지털콘텐츠에 대한 관심은 정보통신 기술이 발전함에 따라서 항상 존재하였고, 특히, 이를 교육 혹은 학습과 연결시키려는 노력이 끊임없이 지속되어 왔다. '디지털 콘텐츠를 활용하면 교육력을 향상시켜 줄 것인가'라는 근원적인 질문이 최근에 들어와서 컴퓨터공학이 발달하면서 '공학이 교육의 모습을 어떻게 바꾸어 놓을 것인가'하는 것으로 보다 확장하게 되었다.[1] 그러나 아직도 교육용 디지털콘텐츠 연구자들의 주요 관심은 디지털 콘텐츠의 효과성 측면에 있다 할 수 있다. 이러한 관심에도 불구하고 디지털 콘텐츠가 단순히 다양한 정보양식의 제공을 통해 흥미나 관심을 높여 줄 것이라는 기대 외에 어떤 이유에서 구체적으로 학습 효과를 향상시켜 주는 지 또는 이 디지털 콘텐츠의 어떤 특성이 학습활동과 어떻게 관련되어 있는지에 대한 연구를 찾기 쉽지 않다.[2]. 학생들은 새로운 정보통신 기술이 적용된 교육용 콘텐츠에 몰입(follow) 정도가 높은 신기효과(novelty effect)를 가지고 있다. 유비쿼터스 학습 환경에서 적용될 모바일러닝(m-learning), 가상현실러닝(v-learning), 로봇러닝(r-learning) 증강현실(Augmented Reality: AR) 콘텐츠 등 신기술에 대한 기대가 크다.[3]

증강현실은 실제 현실세계에서 맥락성을 유지

하며 3차원의 가상객체를 통한 증강된 정보를 학습자에게 제공한다. 또한 기존 데스크탑 PC를 통해 지배적으로 활용되어 오던 그래픽 인터페이스 방식이 아닌 구체적인 실제 세계의 사물을 가지고 가상객체를 조작하는 실물형 인터페이스(Tangible User Interface: TUI)를 제공한다. 이러한 매체의 특성으로 말미암아 증강현실은 체험에 의한 학습(learning by doing)과 실제적인 학습(authentic learning)을 가능하게 함으로써 학습에서의 현존감(presence)과 몰입을 높여 학습효과를 극대화하는 결과를 얻을 수 있다.[4]

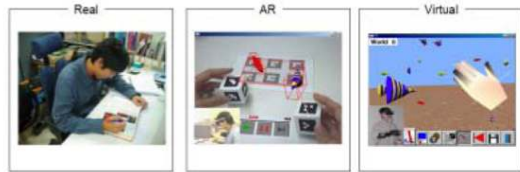
증강현실을 구현하는 기술 중 영상매칭, 마커 인식, 위치정보 등에 관한 내용은 활발히 연구되고 있으나 오디오 부분은 공간감각을 느끼게 하는 연구와[5] 사용자 인터페이스로 햅틱(haptic)과 같이 사용하는 오디오로[6] 상호작용 효과를 증대하도록 하였다.

본 연구에서는 움직일 수 있는 모델을 정하고 시스템이 마커를 인식했을 경우 모델을 띄운다. 키보드를 이용하여 모델을 움직였을 때 모델의 움직임에 따라 원래 오디오에 효과를 주어 출력하므로 실감나는 오디오를 들을 수 있게 한다.

## II. 관련연구

### 2.1 증강현실 정의 및 특징

증강현실은 가상현실(Virtual Reality)의 하나의 분야에 서 파생된 기술이다. 증강현실은 현실세계와 가상의 체험을 결합하는 기술을 의미한다. 즉, 실제 환경에 가상사물을 합성하여 원래의 환경에 존재하는 사물처럼 보이도록 하는 컴퓨터그래픽 기법이다.[7] 가상현실과 증강현실은 모두 가상환경에 바탕을 두고 있다. 그러나 일반적으로 가상현실은 사용자가 상상할 수 있는 공간과 사물을 컴퓨터로 구현하고 이것들과의 상호작용을 통하여 실제와 같이 몰입할 수 있는 가상의 세계를 의미한다. 증강현실 기술은 일반 가상현실 기술의 부류라고도 할 수 있으나, 실시간으로 현실세계와 가상객체가 동시에 혼합해서 제공되는 점에서 다르다고 할 수 있다. 또한 여러 가지 센서와 디스플레이 장치, 영상 합성기술 등을 필요로 하게 된다.[8] 증강현실과 가상현실의 관계를 나타내면 그림 1과 같으며, 왼쪽은 실제 환경을 나타내고 오른쪽은 컴퓨터에 의해 생성된 가상현실이다. 증강현실(AR)은 실제 환경과 가상 환경의 중간 단계이다.



실제환경 증강현실 가상현실  
그림 1.증강현실과 가상현실의 관계

### 2.2 오디오 및 오디오 효과 기술

모든 소리는 일정한 높이, 크기, 음색(音色)이라는 감각상의 요소를 지니며, 이것에 의해서 서로 다른 소리로 인식된다. 그 중에서 소리의 높이는 음파의 주파수, 크기는 음압, 음색은 파형(波形)이라는 물리적 요소에 대응한다. 음에는 아주 많은 물리적 특성이 있으며 음의 전달 속도(음속), 파장, 반사, 흡음, 회절, 굴절, 공진, 맥놀이, 도플러 효과 등이 있다. 도플러 효과는 음원이나 발음체가 소리를 내면서 사람의 귀에 접근해 오거나 멀어지면, 접근해 올 때는 음원의 음높이(주파수)보다 높고, 멀어질 때는 낮게 들리는데, 청취자가 움직일 때도 마찬가지이다. 이펙터는 전기 신호화한 음을 가공하여 원음과는 다른 음으로 변화시키는 일반적인 기기를 가리킨다. 이펙터의 분류는 크게 시간 제어계, 주파수 특성 제어계, 레벨 제어계, 디스토션계, 복합형이 있으며, 많이 사용되는 시간 제어계에는 에코(echo), 리버브(reverberation), 코러스(chorus), 플랜저(flanger), 페이저(phaser), 피치 시프터(pitch shifter) 등이 있다. 피치 시프터는 입력 신호의 높이(피치)를 변화시키는 이펙터로서, 피치 트랜스 포저라고도 한다. 보컬 음정의 분방함을 보정하는 것을 제외하고 피치를 놓친 음을 원음과 혼합해서 코러스 효과를 얻는 데도 사용한다.

## III. System 구현

### 3.1 구현 방법

System 구현을 위한 이미지처리와 증강현실 구현을 위한 개발 도구로는 높은 해상도의 모델 파일을 지원하고 보다 높은 AR을 위한 기술지원을 하는 Goblin XNA를 사용하였다.[9] 그리고 사운드 편집을 위한 Editing Tool로는 XNA Game Studio에서 제공되는 Microsoft Cross-Platform Audio Creation Tool(XACT) v3.0을 사용하였다.

#### 1) Goblin XNA의 전체구조

Goblin XNA의 전체적인 기본 노트구조는 그림 2와 같다.

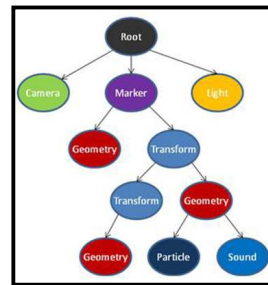
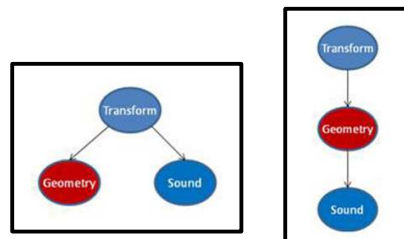


그림 2 Goblin XNA의 노트구조

우선 전체적인 구조는 Root → Marker → Transform → Geometry → Sound로 이루어져 있다. Geometry는 3D모델에 해당된다. Transform은 Translation, Rotation, Scale과 같은 변형을 정의한다. Marker는 Marker에 대한 정보가 포함되어 있다.

Sound 노트의 경우에는 Geometry와 Transform 노트에 어떻게 연결되는가에 따라서 다르다. 즉, 사운드가 Geometry Node의 transformation에 영향을 받을 수도 있으며, 받지 않을 수도 있다. 이를 그림 3에 나타내었다.



Not affected by Geometry Node(왼쪽) / Affected by Geometry Node(오른쪽)

그림 3 사운드 노트구조

#### 2) 사운드 파일 편집 과정

본 System은 다음과 같은 사운드 파일의 편집 과정을 통해 사운드의 변화가 구현되어진다. 우선 XNA Game Studio에 포함 되어 있는 Microsoft Cross-Platform Audio Creation Tool(XACT) v3.0

을 이용하여 새로운 Project를 생성한 뒤 왼쪽의 메뉴를 보면 여러 가지의 메뉴가 있다. 그 중에서 Wave Banks와 Sound Banks. 그리고 RPC Preset을 사용한다. 각각의 Banks를 생성한 다음 오른쪽 창에서 \*.Wav파일을 Wave Bank → Sound Bank(Sound Name) → Sound Bank(Cue Name) 순서로 Drag & Drop을 통하여 추가 한다. 이를 그림 4에 나타내었다.

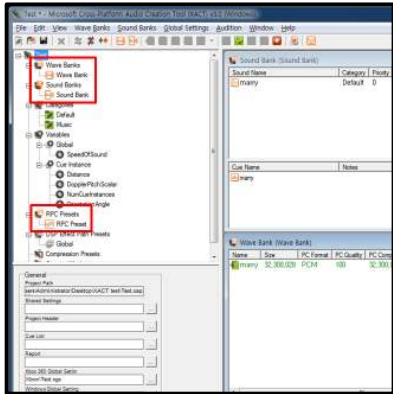


그림 4 XACT v3.0의 기본구조

그리고 RPC Preset메뉴로 들어가서 7가지의 Variable 및 5가지의 Parameter를 설정한 뒤 나타나는 그래프를 이용하여서 편집한다. 이를 그림 5에 나타내었으며, Variable과 Parameter의 종류는 표 1과 같다.

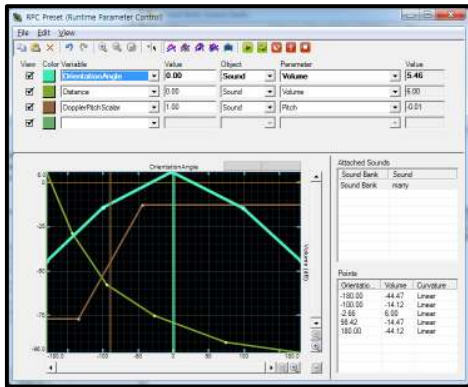


그림 5 RPC Preset 설정

표1. RPC Preset 항목

Variable	Parameter
Attack Time	Pitch
Distance	Volume
Doppler Pitch Scalar	Filter Frequency
Speed Of Sound	Filter Q Factor
Orientation Angle	Reverbe Send
Release Time	
NumCueInstances	

3) 콘텐츠의 제작 및 출력

출력될 콘텐츠의 모델파일들은 해당 프로젝트의 Content폴더와 프로젝트의 Content영역에도 추가한다. 그리고 모델을 로드하여서 크기설정, 위치설정, 모델이 향하는 방향 및 각도 설정 등을 하여 준다. 그리고 Sound 노드를 하나 생성하여서 이를 3D모델에 연결시키고 3D모델은 마커에 연결하여 준다. 이를 그림 6에 나타내었다.

```
private void CreateObject()
{
    MakingMarker(); // 마커로드하는 부분
    ModelLoader loader = new ModelLoader();

    /* Apple */
    /* Apple */
    /* Apple */
    Model appleModel = (Model)loader.Load("apple"); // 모델 .fbx이름
    ModelInfo apple = new ModelInfo(appleModel, "apple");

    GeometryNode appleNode = apple.nodeInfo();
    TransformNode appleTransNode = apple.transInfo(new Vector3(0, 0, 0),
        new Vector3(0.8f, 0.8f, 0.8f), new Vector3(0, 1, 0), 270);

    appleTransParentNode = new TransformNode();
    appleTransParentNode.Translation = Vector3.Zero;
    soundTransNode = new TransformNode();
    soundTransNode.Translation = Vector3.Zero;
    testSoundNode = new SoundNode();
    test2SoundNode = new SoundNode();

    markerModel.AddChildID(appleTransParentNode);
    appleTransParentNode.AddChildID(appleTransNode);
    appleTransNode.AddChildID(appleNode);
    appleNode.AddChildID(testSoundNode);
}
```

그림 6. 3D 모델의 로드 및 노드연결

모델을 움직이기 위해서는 움직여 주는 부분을 프로그램이 실행중일 때 계속 호출이 되는 Update()메소드에 넣어주어야 하며, 각 축을 의미하는 X, Y, Z키와 +, - 키의 합성으로 움직일 수 있도록 하였다. 변화된 값을 Transform 노드에 업데이트를 시켜줌으로써 모델이 움직이게 된다. 이를 그림 7에 나타내었다.

```
protected override void Update(GameTime gameTime)
{
    KeyboardState state = Keyboard.GetState();

    // X, Y, Z의 각 좌표에 해당하는 X, Y, Z키중 하나를 누른 상태에서
    // +, -키를 누르게 되면 좌표가 변경
    // float의 주기는 좌표에 물리값에 크게
    if (state.IsKeyDown(Keys.J))
    {
        if (state.IsKeyDown(Keys.Add))
        {
            transX += 0.3f;
        }
        else if (state.IsKeyDown(Keys.Subtract))
        {
            transX -= 0.3f;
        }
    }
    if (state.IsKeyDown(Keys.I))
    {
        if (state.IsKeyDown(Keys.Add))
        {
            transY += 0.3f;
        }
        else if (state.IsKeyDown(Keys.Subtract))
        {
            transY -= 0.3f;
        }
    }
    if (state.IsKeyDown(Keys.Z))
    {
        if (state.IsKeyDown(Keys.Add))
        {
            transZ += 0.3f;
        }
        else if (state.IsKeyDown(Keys.Subtract))
        {
            transZ -= 0.3f;
        }
    }

    // 키보드 입력에 의하여 변경된 값을 입력받아 좌표값을 업데이트
    appleTransParentNode.Translation = new Vector3(transX, transY, transZ);
    soundTransNode.Translation = new Vector3(transX, transY, transZ);
}
```

그림 7. 3D 모델 움직임

3.2 시스템 구현 결과

개인용 컴퓨터(Intel core 2 CPU, 1.86GHz, 3GB RAM)에서 운영체제는 MS Windows 7 Ultimate K SP1, AR 개발을 위한 프레임워크는 Goblin XNA 3.4, 프로그램 언어는 C#, Sound Editing Tool은 XACT v3.0을 이용하여 구현하였으며, 증강현실에 사용되는 3D모델은 “트럭”을 사용하였고, 사운드는 트럭이 움직일 때 나는 엔

진 소리를 사용하였다.

그림 8은 거리에 따른 크기의 변화를 나타낸 것이다. 트럭이 뒤쪽으로 갈수록 소리의 크기가 작아지는 것을 볼 수 있다.

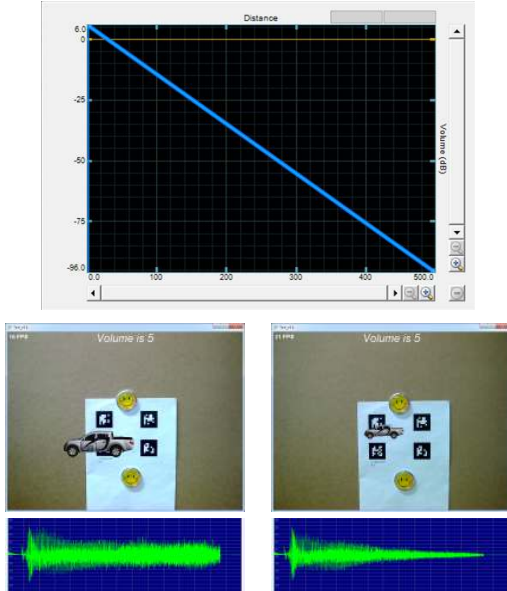


그림 8. 거리에 따른 크기 변화

그림 9는 Pitch의 변화를 회전각의 변화에 따라서 나타낸 것이다. Pitch란 음의 고저를 의미한다. 회전각이 90을 기준으로 Pitch값의 변화는 트럭이 오른쪽으로 가면 작아지고 왼쪽으로 가면 커지는 것을 볼 수 있다.

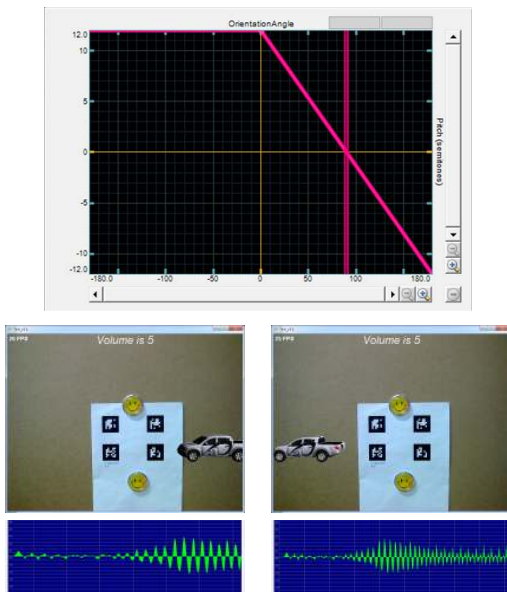


그림 9. 회전각에 따른 Pitch 변화

#### IV. 결 론

정보통신의 급속한 발전과 새로운 디지털 콘텐츠에 대한 관심이 높아지고 있는 환경 속에서 증강현실 기술은 단순한 학습 콘텐츠의 단점을 보완하여 가상으로 체험하며 학습효과를 극대화시킬 수 있는 첨단 기술이다.

본 연구에서는 움직일 수 있는 모델을 정하고 시스템이 마커를 인식했을 경우 모델을 띄운다. 키보드를 이용하여 모델을 움직였을 때 모델의 움직임에 따라 원래 오디오에 효과를 주어 출력하므로 실감나는 오디오를 들을 수 있게 하였다.

오디오의 효과는 모델이 움직였을 때 움직인 거리에 따른 소리의 크기를 변화하도록 하였으며, 모델이 회전하도록 했을 경우 회전각에 따른 음정 즉 피치의 변화를 가져오도록 하여 증강현실에서 현존감을 극대화하는 결과를 얻을 수 있도록 하였다.

향후 연구로는 도플러 효과와 여러 가지 오디오 효과를 동시에 적용한 학습콘텐츠를 만들고 이들의 효과를 테스트 해보는 연구가 필요하다.

#### 참고문헌

- [1] Banathy, B. H., Systems design of education, NJ: Educational Technology Publications, 1991.
- [2] Crook, C., Computers and the collaboration experience of learning, NY: Routledge, 1994.
- [3] 장상현, 계보경, u-러닝 환경에서의 에듀테인먼트 개발 및 적용, 정보과학회지, 제4권, 제호, 2006.
- [4] 계보경, 증강현실 기반학습에서 매체특성과 현존감 학습몰입 학습효과의 관계 규명, 이화여자대학교, 박사학위 청구논문, 2007.
- [5] Lindeman, R.W., et al., "An Empirical Study of Hear-Through Augmented Reality: Using Bone Conduction to Deliver Spatialized Audio," VR Conf., pp. 35-42, 2008.
- [6] Henrysson, A., et al., "Face to Face Collaborative AR on Mobile Phones," Proc. Fourth IEEE and ACM International Symposium, pp.80-89, 2005.
- [7] 방준성, 최은주, 증강현실 국내외 기술동향과 발전전망, 한국과학기술정보연구원.
- [8] 장상현, 계보경, 증강현실 콘텐츠의 교육적 적용, 한국교육학술정보원.
- [9] <http://goblinxna.codeplex.com>