

<http://dx.doi.org/10.17703/JCCT.2018.4.4.309>

JCCT 2018-2-39

## 3D 가상현실기반의 발표훈련시스템

### Presentation Training System based on 3D Virtual Reality

정영기\*

Young-Kee Jung\*

본 연구에서는 실제와 같은 가상발표환경을 구현하여 실전에서 자신감 있게 발표할 수 있도록 도와주기 위한 3D 가상현실기반의 발표훈련시스템을 제안한다. 제안시스템은 발표자의 음성과 행동을 실시간으로 분석하여 가상공간의 청중들에게 반영되게 함으로서 사실감 있고 몰입도 높은 발표 및 면접 환경을 제공하였다. 발표자는 6DOF Tracking이 되는 HMD와 VR Controller를 착용하고 Kinect를 이용하여 가상공간 속에서의 시점 변화 및 인터랙션을 줄 수 있으며 가상공간은 사용자가 설정한 다양한 환경으로 변경이 가능하도록 하였다. 발표자는 가상공간 속에 별도로 제공되는 뷰에 표시된 프리젠테이션 파일 및 스크립트를 보며 내용 숙지 및 발표 숙달 연습을 하게 된다.

**주요어** : 발표훈련, 가상현실, 실감미디어, 동작분석

**Abstract** In this study, we propose a 3D virtual reality based presentation training system to help implement the virtual presentation environment, such as the real world, to present it confidently in the real world. The proposed system provided a realistic and highly engaging presentation and interview environment by analyzing the speakers' voice and behavior in real time to be reflected in the audience of the virtual space. Using HMD and VR Controller that become 6DOF Tracking, the presenter can change the timing and interaction of the virtual space using Kinect, and the virtual space can be changed to various settings set by the user. The presenter will look at presentation files and scripts displayed in separate views within the virtual space to understand the content and master the presentation.

**Key words** : Presentation Training, Virtual Reality, Realistic Media, Gesture Analysis

#### 1. 서 론

현대사회에서는 강사나 연설자와 같이 사람들 앞에서 연설하는 것을 전문으로 하는 사람들 외에도, 면접이나 회사 내에서의 업무 발표와 같이, 남들 앞에서 자신의 의견을 발표해야 하는 경우가 종종 있다. 그러나 종래에는 실제와 같은 발표상황을 구현하여 발표준비

를 도와주는 방법이 없었기 때문에 발표를 준비할 때 제 3자와의 관계를 고려하여 발표를 준비할 수 없었으며, 이에 따라 일반인들은 회의발표에 두려움을 느끼거나 완벽한 발표를 준비할 수 없었던 단점이 있었다. 유사 사례로 면접 체험기가 있으나 DID에 2D로 제작된 면접관이 녹음된 육성 질문을 재생하여 체험자가 해당 질문에 대응하는 시스템이고 실감미디어로 구축되지

\*정회원, 호남대학교 미래자동차공학부  
접수일: 2018년 8월 6일, 수정완료일: 2018년 9월 15일  
게재확정일: 2018년 9월 28일

Received: August 6, 2018 / Revised: September 15, 2018

Accepted: September 28, 2018

\*Corresponding Author: ykjung@honam.ac.kr

Dept. of Automotive Engineering, Honam University

않아 현실감 및 몰입감이 현저히 낮다.

따라서 발표자의 시선, 음성, 동작을 분석하여 피드백을 줌으로써 실전에서 체험자가 가상현실 콘텐츠를 체험하는 동안 실제 공간에 있는 느낌을 느낀다거나 체험자 스스로 콘텐츠와의 상호작용을 하는 동안 체험자의 호기심이 유발되고, 콘텐츠를 체험하는 자체에서 흥미를 느끼게 할 수 있는 가상현실훈련시스템에 대한 필요성이 대두되고 있다[1].

가상현실훈련시스템은 기술의 발전으로 인해 간단한 인터페이스만으로도 가상과 현실의 구분을 모호하게 만들 수 있는 실감나는 콘텐츠가 제공되고 있고 현실 공간 속에 가상의 콘텐츠를 불러들여 현실과 가상의 공간이 하나로 존재하게 되는 실감형 상호작용 가상현실로 진화하고 있다. 가상현실 기술 중 상호작용 기술은 실감 효과를 실시간으로 제시하는 가시화와 더불어 가상현실시스템을 구축하는 핵심기술 중 하나이고 가상현실훈련 시스템이 훈련자에게 실감 있는 교육과 훈련을 가능하게 한다[2][3][4][5].

상호작용기술로는 가상훈련 특성상 사용자의 몸, 손, 얼굴에 대한 인식/추적과, 사용자에게 피드백을 제공하기 위한 모션 센싱 기술인 Kinect, Myo, Leap Motion, RealSense가 있다. 이러한 모션 센싱 장치들은 가상현실에서 직관적이면서도 현실감 있는 훈련을 제공하는데 중요한 역할을 한다. Kinect는 적외선 패턴을 활용하여 사용자의 몸, 손 및 얼굴의 위치를 인식 및 추적한다. 사람의 동작을 각관절의 좌표값과의 변화로 사람의 동작의 의미를 파악하기 위해 Kinect 센서를 통해서 감지된 동작의 인식 및 분류에 대한 연구가 진행되었다 [6]. 또한 가상환경과의 상호작용 목적으로 사용자가 복잡한 wire와 센서를 부착해야만 하는 불편함을 해소하기 위한 마커리스(markerless) 방식의 인터페이스를 기반으로 한 상호작용 기법과 사용자의 동작추적 및 제스처 인식을 바탕으로 한 응용 연구가 활발히 진행되었다 [7][8].

본 연구에서는 발표자의 음성과 행동을 실시간으로 분석하여 가상공간의 청중들에게 반영되게 함으로서 사실감 있고 몰입도 높은 발표 및 면접 환경을 제공하는 발표면접훈련시스템을 제안하고자 한다. 제안시스템은 발표자의 목소리와 행동, 시선에 따라 가상의 청중들도 표정, 소리, 제스처 등의 피드백을 주어 체험자가 더욱 실감할 수 있도록 하였다. 또한 다양한 규모의 발

표 현장 체험이 가능하며 발표자가 입체 영상 및 Head Tracking 기술을 통해 가상공간에서 시선변화를 줄 수 있으며 실제와 같은 현장에서 발표를 하는듯한 몰입도 높은 경험을 제공하고자 한다.

## II. 발표 VR 훈련시스템

제안된 발표훈련시스템은 그림1과 같이 발표자가 6DOF Tracking이 되는 HMD와 VR Controller을 착용하고 Kinect를 이용하여 실제와 같은 가상공간 속에서의 시점 변화 및 인터랙션을 줄 수 있으며 가상공간은 사용자가 설정한 다양한 환경으로 변경이 가능하도록 설계하였다.



그림 1. 발표 VR 훈련시스템 개요  
Figure 1. Overview of Presentation VR Training System

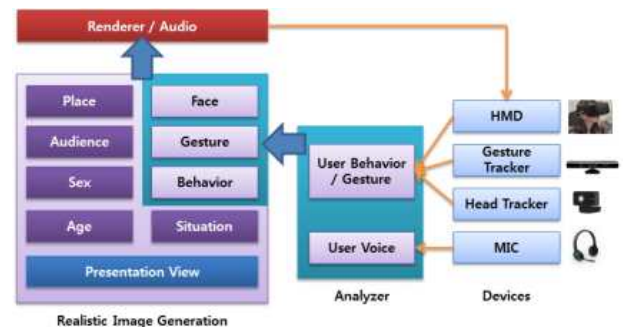


그림 2. 발표 VR 훈련시스템 구성  
Figure 2. Configuration of Presentation VR Training System

제안된 시스템은 크게 네 부분으로 구성된다. 첫째, 음성분석부는 발표자의 목소리가 청중과 면접관이 듣기에 적절한지 판단하기 위해 마이크로 들려오는 체험자의 목소리를 실시간으로 분석한다. 둘째, 발표자 행동 분석부는 발표자가 자연스러운 자세로 훈련에 임하는지를 판단하기 위해 손, 머리, 몸 전체의 움직임을 실시간으로 분석한다. 셋째, 발표자의 시선에 따라 가상청중의 실시간 피드백을 제공한다. 넷째, 가상청중의 다양한

제스처, 표정DB 및 발표현장과 같은 배경 및 소리를 생성한다.

### 1. 발표자 음성 분석

체험자의 목소리가 적절한 크기와 속도, 높낮이 등을 유지하고 있는지를 분석하기 위해 볼륨과 음역 주파수를 샘플링하여 그림3, 4와 같이 측정한다.



그림 3. 발표자의 음성 높이 측정 결과  
 Figure 3. Voice Height Measurement Results for Presenters



그림 4. 발표자의 음성 크기 측정 결과  
 Figure 4. Voice Size Measurement Results for Presenters



그림 5. 음성 분석후 가상청중 피드백  
 Figure 5. Virtual Audience Feedback after Voice Analysis

측정된 음성은 음량과 속도, 높낮이를 실시간으로 누적 분석되어 그림5와 같이 해당 반응의 가중치를 높이고 가상청중의 반응을 결정한다. 이와 같이 지속적으로 업데이트 되는 반응 가중치에서 임계값을 넘는 반응을 가상청중 및 면접관에게 적용시킨다. 음성분석에 따른 해당반응의 가중치를 높이는 기준은 그림6과 같이 판정한다. 체험자의 말이 자연스럽게 이어지지 않거나 너무

빠르고 목소리가 작은 경우 필터링 기준으로 음성의 높낮이가 10초 동안 90%의 점유율로 'D' 이하의 값을 가지는 것으로 정의되고 부정적인 반응의 가중치를 높게 된다.

가상청중은 표정, 행동, 소리로 발표자에게 피드백을 주어 보다 진지한 상황을 연출하게 되며 발표자는 상황에 대처하는 훈련과 발표 속도, 감정조절에 대해 훈련하게 된다. 또한 음성분석을 통해 1분 동안 해당 분석요인이 적용 되지 않을시 0.3의 가중치를 낮추어 너무 빈번한 피드백을 방지하도록 하였다.

발표자	종류	기준 값	시간	점유율/횟수	청중 반응
음성	음 높이	G 이하 (미)	10초	80	Loose
		D 이하 (도)	10초	90	Negative
	음성 공백	5초 이상	5초	100	Loose
		10초 이상	10초	70	Loose
	음성 크기	2 이하	10초	80	Loose
		0 이하	10초	80	Negative
	음성 속도	2.0 이상	20초	90	Distracted
	말끝 흐리기	0.5 이하	10초	80	Loose
	무의미	3회 이상	10초	3회 이상	Negative

그림 6. 음성분석에 따른 가상청중반응  
 Figure 6. Virtual Audience Response according to Voice Analysis

### 2. 발표자 행동 분석

발표자 행동 분석부는 입력된 행동의 변화량을 감지하여 청중 아바타가 취해야 할 피드백 테이블에 가중치를 계산하는 역할을 한다. 청중이 발표자에게 몰입할 수 있는 중요한 요소 중 하나로 발표자의 제스처가 있다. 자연스러운 제스처는 청중에게 발표자가 정보를 조금 더 시각적으로 이해하는데 많은 도움을 주게 된다.

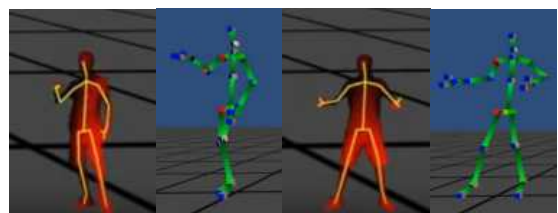


그림 7. 행동 분석 결과  
 Figure 7. Gesture Analysis Result

발표자의 행동을 분석하기 위해 Kinect v1을 사용하였으며 그림7은 체험자의 Depth Map과 실시간으로 추적되는 Skeletal Data를 3차원의 Bone 형태로 디스플레이한 모습이다. Kinect v1은 총 20개의 Skeletal 데이터를 추적할 수 있는데 제한된 장소에서의 발표나 면접시에는 상체의 움직임이 청중의 평가에 주효하게 반영되기 때문에 양손, 머리, 몸통 4개의 Bone 데이터의 6DOF 변화량을 0.5초마다 그림8과 같이 기록하였다.

Skeletal의 크기는 남녀 평균키인 1.7m로 설정하였다.



그림 8. 발표자의 손과 머리의 움직임 양 측정 결과  
Figure 8. Results of Measurement of the Amount of Movement of the Presenter's Hand and Head

측정된 움직임데이터는 실시간으로 누적 분석되어 그림9와 같이 해당 반응의 가중치를 높이고 가상청중의 반응을 결정한다.



그림 9. 행동 분석후 가상청중 피드백  
Figure 9. Virtual Audience Feedback after Geature Analysis

행동 분석에 따른 해당반응의 가중치를 높이는 기준은 그림 10과 같이 판정한다.

몸	신만	10초	10m 이상	몸을 너무 자주 흔들	
	경직	10초	1m 이하	몸이 경직되어 있음	
	뺨다리	5초	100	뺨다리를 한 채로 발표	
수	등	5초	100	등을 보인채로 발표	
	손 넣기	5초	100	손을 호주머니에 넣고 발표	
	얼굴 만지기	20초	3회 이상	얼굴/머리 부위를 자주 만짐	
	손 이동 양	20초	90	손을 계속 움직임	
	손 이동 속도	10초	10m 이상	손을 너무 빠르게 휘저음	
	뒷짐	10초	100	뒷짐 친 채로 발표	
	손 모으기	10초	100	손을 배꼽 아래에서 모은채로 발표	
	팔짱	10초	100	팔짱 낀 채로 발표	
	머리	한쪽 팔 잡기	10초	100	한쪽팔도 다른 팔을 잡기
		기울기	5초	100	머리를 기울인 채로 발표
시선	천정, 바닥	5초	100	단상의 자료 또는 스크린만 바라볼 때	
	발표자료, 좌우	8초	100	단상의 자료 또는 스크린만 바라볼 때	

그림 10. 행동분석에 따른 가상청중반응  
Figure 10. Virtual Audience Response according to Gesture Analysis

그림 11은 움직임이 없는 발표자에 따른 청중 아바타의 Loose 행동 동작 예를 보여주고 있다.

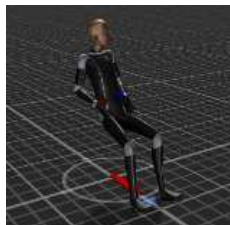


그림 11. 움직임이 없는 발표자에 따른 청중아바타의 Loose 행동  
Figure 11. Loose Behavior of the Audience Abata with no Movement

### 3. 발표자 시선 인식 피드백

청중의 입장에서 발표자의 시선을 자주 받는 것이 자신에 대한 관심과 몰입을 증대시킬 수 있는 요소로 작용된다. 따라서 이를 청중 아바타 피드백부에 적용할 수 있도록 청중 아바타는 체험자의 시선을 받는 회수와 시간을 측정한다.

청중 아바타의 시선 받는 시간과 빈도는 3D Picking 처리 방법을 통해 구현되는데 분할 스테레오 방식을 이용하기 때문에 Picking Ray가 2개가 되고 중심선에 일치하지 않기 때문에 양안 벡터의 중심부 벡터를 구하여 그림 12와 같이 3D Picking에 이용할 수 있도록 하였다.

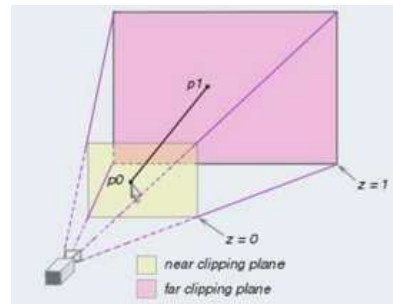


그림 12. 3D Picking 처리  
Figure 12. 3D Picking Process

그림13과 14는 청중 아바타가 체험자의 시선을 받는 경우와 받지않는 경우의 장면을 보여주고 있다.



그림 13. 발표자의 시선을 받는 청중 아바타  
Figure 13. A Presenter's Eye-Catching Audience Avatar



그림 14. 발표자의 시선을 받지 않는 청중 아바타  
Figure 14. A Presenter's Non Eye-Catching Audience Avatar

각 청중 아바타가 체험자의 시선을 받는 시간에 따라 청중 아바타 피드백부에 가중치를 더하는 기법을 사용하였는데 그림15에서는 A, C, D 청중을 제외한 나머지

지 청중은 Loose의 가중치를 높이게 되어 현 상황이 지속될 경우 결과적으로 B, E, F는 주변을 살피거나 옆 사람에게 대화를 하는 등의 행동을 취하게 된다.

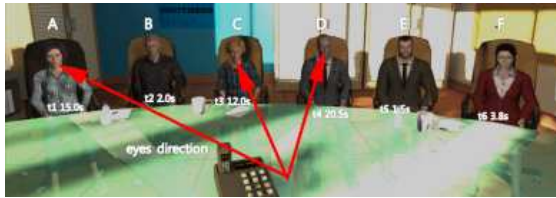


그림 15. 발표자의 시선이 머무르는 시간에 따른 청중 아바타의 행동

Figure 15. Audience Avatars' Behavior over the Time that the Presenter's Eyes Stay

#### 4. 실감영상 생성

실감영상 생성 부는 발표자가 보게 될 영상을 생성하는 부로 장소, 청중, 상황, 배경 등에 대한 일반적 3D 렌더링 처리와 청중 아바타의 피드백 표정 및 행동, 제스처를 적용할 피드백 엔진으로 구성된다. 그리고 PPT View는 Presentation File을 Image 또는 HTML로 컨버팅하여 화면 코너에 표시되어 발표자가 발표 중 내용을 확인 할 수 있도록 한다.

발표자의 화면에 표시되는 실감영상은 그림 16과 같으며 발표 장소, 청중, 소리, Presentation View가 표시된다.



그림 16. 실감공간 영상생성

Figure 16. Realistic Space Image Generation

발표자는 가상공간 속에 별도로 제공되는 뷰에 표시된 프리젠테이션 파일 및 스크립트를 보며 내용 숙지 및 발표 숙달 연습을 하게 된다.

청중 아바타는 발표자의 발표에 따라 다양한 행동과 함께 매칭되는 표정을 짓도록 하여 발표자가 더욱 몰입할 수 있는 환경을 제공해주도록 하였다. 청중 아바타가 눈썹, 눈꺼풀, 윗 입, 아랫 입, 턱, 입꼬리 등을 움직일 수 있도록 BlendSahpe이 적용되었다.

표정 애니메이션을 제작하는 방법은 영화에 사용되는 Optical 센서를 얼굴 표면에 여러개 부착하여 적외선 카메라로 이를 추적하는 마커 기반 모션캡처 방법과

마커 없이 PC카메라 또는 Depth 카메라를 이용하는 마커리스 모션캡처 방식이 있다. 본 연구에서는 마커 기반 모션캡처에 비해 애니메이션 획득 시간이 빠르고 후처리 가공이 거의 필요 없으며 저비용으로 구현할 수 있는 마커리스 모션캡처 방식으로 표정 애니메이션을 제작하였다. 그림17과 같이 액터의 얼굴 형태와 특징점을 최적화하기 위해 캘리브레이션 과정을 거치며 기본 표정 템플릿 데이터를 생성하여 액터와 흡사한 표정 애니메이션을 취득할 수 있도록 하였다.

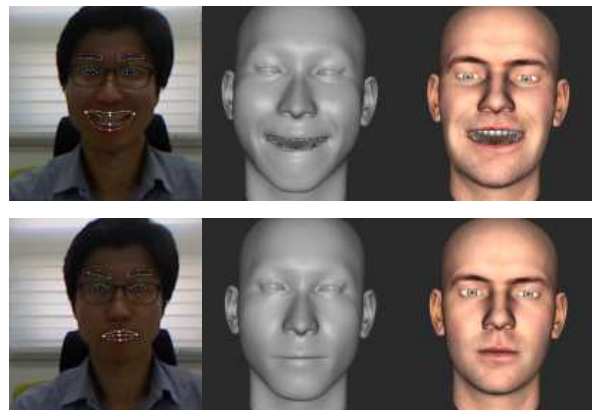


그림 17. 마커리스 표정 모션캡처

Figure 17. Markerless Facial Motion Capture

청중 아바타의 제스처 애니메이션은 청중 아바타가 주로 앉아 있는 행동을 취하기 때문에 그림18과 같이 마커 기반 모션캡처를 활용하였다. 모션캡처 후처리에는 Motion Builder를 이용하여 Spike나 Gap 등을 제거하여 자연스러운 애니메이션을 제작하였다.



그림 18. 마커기반 제스처 모션캡처

Figure 18. Marker based Gesture Motion Capture

소리는 실감공간상에 몰입할 수 있도록 하는 매우

중요한 요소로써 각 프리젠테이션 장소에 맞는 환경음과 사무기기 소리, 노이즈, 아바타 반응 소리 등을 녹음 제작하였다. 피드백 사운드는 그림19와 같이 크게 환경음과 청중음으로 구분되는데 환경음은 실감공간과 매치되는 각종 주변음과 사무기기음, 노이즈 등을 포함한다. 청중음은 돌발음, 긍정음, 지루함, 부정음으로 카테고리리를 나눠 제작하였다.

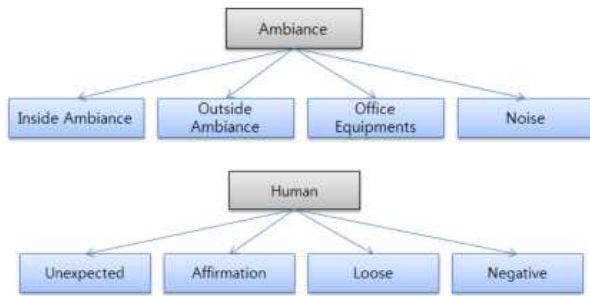


그림 19. 제작된 환경음과 청중음  
Figure 19. Environmental Sound and Audience Sound Produced

### III. 시스템 구현 및 실험

발표 면접 훈련 시스템은 그림 20과 같이 발표와 면접의 두 가지 훈련 시스템이 통합된 형태로써 각각 8개의 씬을 가진다.

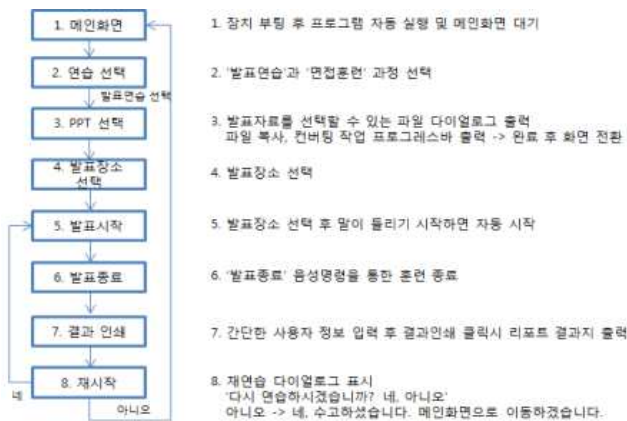


그림 20. 콘텐츠 시나리오  
Figure 20. Contents Senario

체험자는 부스내에 들어가서 발표 또는 면접 훈련을 선택한 후 HMD를 착용하면 가상의 실감공간 내에 청중과 면접관이 보이며 체험자는 선택한 시나리오에 맞게 발표 훈련을 하거나 면접관의 질문에 성실히 대답하도록 하였다. 체험자의 말, 행동, 시선에 따라 가상 청중과 면접관이 표정과 행동, 소리로써 다양한 감정 상

태를 피드백 하므로 당황하지 않고 끝까지 준비한 내용을 발표하는 훈련을 마치고 스크린에 인쇄 버튼을 터치하여 출력된 훈련 결과를 확인하도록 하였다.

가상훈련장은 소회의실, 중회의실, 대강당을 구축하고 고품질의 영상을 얻기위해 Lighting과 Render to Texture, Transparent Material을 적용하여 실감공간을 구축하였다. 실시간 기반으로 프리젠테이션 현장을 입체 영상 (양안 뷰) 구현 및 100° 이상의 FOV를 확보하기 위해 왜곡 보정과 분할 방식의 스테레오가 필요하여 100°의 체험자 FOV 제공을 위한 어안 렌더링(fish eye rendering) 구현하였다. 그림 21은 왜곡 처리와 분할 스테레오 렌더링이 적용된 소회의실, 중회의실, 대강당에 해당하는 실감공간과 청중아바타의 구현 모습이다.

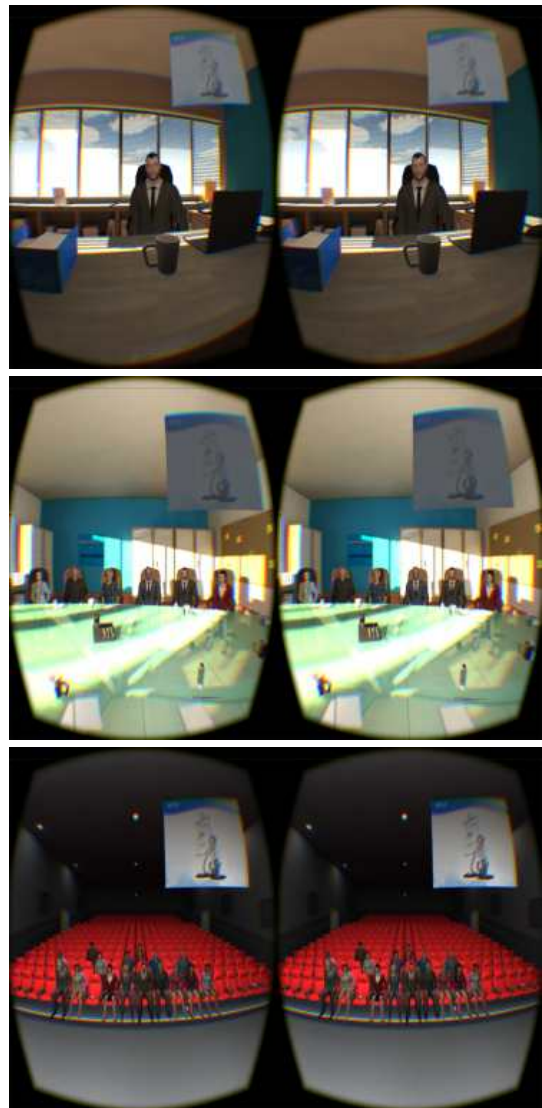


그림 21. 가상훈련장 및 청중아바타  
Figure 21. Virtual Training Center and Audience Avatar

발표훈련시스템은 그림 22와 같이 HMD를 착용하면 가상의 실감공간 내에 청중이 보이며 체험자는 선택한 시나리오에 맞게 발표 훈련을 하거나 면접관의 질문에 성실히 대답하게 된다. 체험자의 화면에 표시되는 영상은 그림23과 같으며 실감 영상에는 발표 장소, 청중, 소리, Presentation View가 표시된다. 가상공간상에 발표해야 하는 프리젠테이션 파일 내용 (PPT)이 별도 뷰로 제공되어 체험자가 뒤를 쳐다보지 않고 청중들의 반응을 보며 내용 숙달 및 발표 훈련이 가능하도록 하였다.



그림 22. 프리젠테이션 훈련 중인 발표자 모습  
 Figure 22. Presenters in Presentation Training



그림 23. 발표자와 상호작용중인 가상청중의 모습  
 Figure 23. Virtual Audience Interacting with Presenters

발표자의 말, 행동, 시선에 따라 가상 청중과 면접관이 표정과 행동, 소리로써 다양한 감정 상태를 피드백하므로 당황하지 않고 끝까지 준비한 내용을 발표하도록 하였다. 그림24은 구현된 발표면접훈련시스템이다.



그림 24. 발표면접훈련시스템  
 Figure 24. Presentation Training System

## IV. 결 론

본 연구에서는 3차원 가상현실기술을 이용하여 사실감 있고 몰입도 높은 가상 발표훈련 시스템을 구현하여 현실세계에서의 실제 발표에 완벽하게 대처할 수 있게 하는 가상발표훈련시스템을 제안하였다.

특히 체험자의 콘텐츠 이용 몰입도와 훈련의 효과를 높이기 위해 체험자의 음성과 행동을 실시간으로 분석하여 가상공간의 청중들에게 반영되게 함으로서 사실감 있고 몰입도 높은 발표 및 면접 환경을 제공하였다. 청중 아바타는 체험자와 상호 작용을 하기 때문에 눈의 시선과 머리의 방향, 표정, 손가락, 다리의 떨림 등 청중 아바타가 움직일 수 있는 모든 요소는 체험자에게 중요한 피드백 정보가 될 수 있도록 하였다. 이를 위해 청중 아바타 표정 DB와 청중 아바타 제스처 DB를 구축하였으며 체험자의 행동과 음성 분석을 통해 실제와 같은 자연스러운 동작과 표정을 취할 수 있도록 하여 체험자 반응형 실감미디어를 구현하였다.

## References

- [1] Lee-kwon Choi, "A Study on the Development of Experienced Content Using Virtual Reality Technology," The magazine of KIICE, Vol. 16 No. 1, Apr. 2015, pp. 46-53.
- [2] Y.W. Kim, D.S. Jo, Y.H Kim, H.M. Kim, K.H. Kim, "Realistic Interaction Technology for Virtual Reality," Electronics and Telecommunications Trends, Vol. 27 No. 3, 2012, pp. 1225-6455.
- [3] Ji Hye Lee, "VR System Environment Technologies and User Input Elements," KSDC Journal, Vol. 24 No. 2, 2018, pp. 585-596.
- [4] Young-Dae Lee, Won-Sik Lee, Jeong-Jin Kang, "Development of Universal Sports Simulator Fusing 5 Senses ," JCCT, Vol. 1 No. 1, 2015, pp. 73-77.
- [5] Im-Chul Kang, Beom-Seok Kim, Gi-Taek Hur, Young-Hyuk Ko, "A Study on the Implementation of Multi-touch model using a Haptic Device in Virtual Reality ," JIIBC, Vol. 10 No. 4, 2010, pp. 83-90.
- [6] Byung-jun Cho, Ha-Young Jang, Byoung-Tak Zhang, "Motion Recognition and Classification using Kinect Sensor Data," KSC, Vol.39 No.28, 2012, pp. 318-320.

- [7] Yoon-Je Kim, Tack-Kyun Koh, Min-Ho Yoon, Tae-Young Kim, "Hand Gesture Recognition Method based on the MCSVM for Interaction with 3D Objects in Virtual Reality," KIPS Conference, Vol.24 No.2, 2017, pp. 1088-1091.
- [8] K.K. Kim, H.J. Kim, S.H. Cho, J.Y. Lee, "Gesture Recognition for Natural Human-Robot Interaction," Electronics and Telecommunications Trends, Vol. 20 No. 2, 2005, pp. 15-20.

※ 이 논문은 2016년도 호남대학교 학술연구  
비 지원을 받아 연구되었음.