

인터랙티브 가상 스튜디오와 몰입형 시청자 환경 (Interactive Virtual Studio & Immersive Viewer Environment)

김래현, *박문호, 고희동, *변혜란
한국과학기술연구원, dochi_ko@kist.re.kr
*연세 대학교, mhpark@kist.re.kr, hrbyun@csai.yonsie.ac.kr

논문 초록

In this paper, we introduce a novel virtual studio environment where a broadcaster in the virtual set interacts with tele-viewers as if they are sharing the same environment as participants. A tele-viewer participates physically in the virtual studio environment by a dummy-head equipped with video "eyes" and microphone "ears" physically located in the studio. The dummy head as a surrogate of the tele-viewer follows the tele-viewer's head movements and views and hears through the dummy head like a tele-operated robot. By introducing the tele-presence technology in the virtual studio setting, the broadcaster can not only interact with the virtual set elements like the regular virtual studio environment but also share the physical studio with the surrogates of the tele-viewers as participants. The tele-viewer may see the real broadcaster in the virtual set environment and other participants as avatars in place of their respective dummy heads. With an immersive display like HMD, the tele-viewer may look around the studio and interact with other avatars. The new interactive virtual studio with the immersive viewer environment may be applied to immersive tele-conferencing, tele-teaching, and interactive TV program productions.

1. 개요

현재의 TV 방송 시스템에서 시청자(tele-viewer)는 방송국이 제공하는 정보를 수동적으로 받기만 하는 형태가 일반적이고 방송자(broadcaster)와 인터랙션이 있다고 하더라도 아주 미미한 정도이다. 현재의 방송 서비스에서 시청자와의 인터랙션은 편지나 PC 통신 등의 Off-line 방식이 대부분이고 on-line 인터랙션을 제공하더라도 음성을 통한 전화대화 정도에 머무르고 있다.

하지만 미래의 TV 방송 서비스는 시청자들이 적극적으로 참여하는 방송이 될 것이고 이를 위한 적절한 방송 시스템과 시청자 환경이 요구된다.

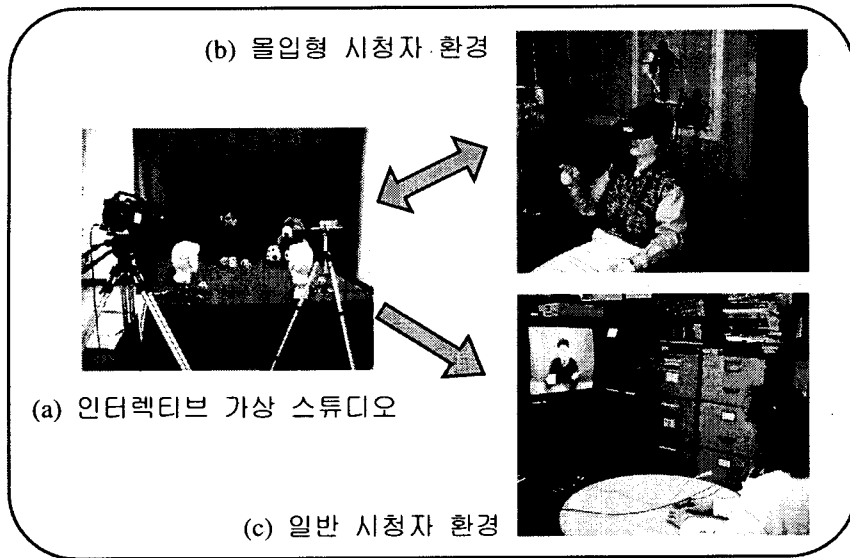
시청자들이 방송에 능동적으로 참여할 수 있는 인터랙티브 TV 방송의 어려움은 기본적으로 방송주체와 시청자들이 지역적으로 떨어져 있어서 이들간에 실시간으로 같은 물리적 공간을 공유할 수 없고 이를 효과적으로 해결할 수 있는 방송 환경이 없다는 데 있다.

본 논문에서는 이러한 문제를 극복할 수 하고 방송자와 시청자들간의 물리적인 공간의 공유와 자연스런 인터페이스를 제공할 수 있는 새로운 방송 환경을 제시하고자 한다.

첫번째로 방송자와 시청자들 간에 물리적인 공간을 공유할 수 있는 Interactive Virtual Studio(IVS)을 소개한다. 이 IVS 안에는 시청자들은 자신의 대리인(surrogate) 역할을 하는 더미헤드(dummy head)로 참여하는데 방송자는 이 더미헤드를 통해 지역적으로 떨어진 시청자들과 물리적인 인터랙션을 할 수 있다. 그리고 virtual studio[1]기술을 사용하여 blue screen 앞의 방송자를 virtual set 에 존재하도록 하였다(그림 1-a).

두번째 방송 환경은 시청자측의 Immersive Tele-viewer Environment(ITE)이다. 기존의 시청자 환경은 수동적이고 2차원적인 인터페이스를 가지고 있다. 이에 반해 ITE는 HMD을 쓴 시청자가 가상 환경 안에서 방송자를 포함한 다른 참여자들과 가상 공간을 공유하고 서로 인터랙션할 수 있는 몰입적인 환경을 제공한다. 또한 스튜디오 안의 방송자의 real video image와 3차원 real audio 등으로 가상환경을 확장(augment)[2]하여 가상 환경의 실제감(reality)을 높였다(그림 1-b).

마지막으로 일반 시청자 환경이다. 기존의 시청자 환경과 마찬가지로 제 3자의 입장에서 인터랙션이 없이 수동적으로 방송국에서 보내준 화면만을 본다. 다른 점이 있다면 virtual set에서 실시간으로 참여하는 시청자들의 모습을 볼 수 있다는 것이다(그림 1-c).



[그림 1] 인터랙티브 방송 환경

본 연구에서는 다음과 같은 몇가지 새로운 시도를 하였다.

- 원격 참여자간의 공유하는 환경을 물리적인 환경과 가상 환경을 동시에 유지
- 원격 참여자간의 물리적인 인터랙션을 제공
- 가상 환경을 실제 세계의 정보로 확장할 때 비디오만을 사용하지 않고 모든 실제 세계의 정보를 이용한다. 즉, 확장(augmentation)의 일반화를 시도하였다.

앞으로 2 장에서는 관련된 연구를 알아보고 IVS 와 ITE 의 관계를 3 장에서 기술하고자 한다. IVS 와 ITE 에 대한 자세한 특징은 각각 4 장, 5 장에서 논하겠다. 6 장에서는 본 연구의 결과에 대한 예를 보여주고 7 장에서 결론을 내리고자 한다.

2. 관련 연구

이 장에서는 원격지의 참여자들이 서로 만나서 인터랙션 할 수 있는 환경을 제시하는 몇가지 시스템들을 알아보려고 한다.

먼저 인터넷 포이어(Internet Foyer)[3]는 월드 와이드 웹 홈 페이지들의 항해가능하고 참여 할 수 있는 3차원 시각화를 할 수 있는 협업 가상 환경(Callaborative Virtual Environment) 기술을 사용하였다. 인터넷 포이어에서의 시각화는 실제 빌딩내의 물리적 포이어(physical foyer)의 벽 위에 투사되어 이루어 진다. 반면 동시에 물리적 포이어로부터 캡처된 실시간 비디오 이미지는 가상 세계의 시각화를 위해 비디오 매핑된다. 이는 가상과 실제의 포이어의 방문자들이 3차원 웹상에서 서로의 비디오 이미지를 통해 상호 인식하고 대화한다.

다음으로 몰입환경의 원격회의 환경의 원형으로써 텔레포트(TELEPORT)[4]는 비록 지역적으로 떨어져 있지만 소그룹의 사람간의 가상적으로 얼굴을 마주 보는 만남을 지원한다. 이 시스템의 특징은 벽 크기의 디스플레이, 뷰어의 시점 추적, 그리고 참여자등과 배경의 분리등이고 이는 컴퓨터 비전과 컴퓨터가 만든 배경과 라이브 비디오의 실시간 합성등을 이용하였다. 이런 특징들은 실제 미팅 방을 가상 미팅 방으로 연장함으로써 다른 지역의 참여자간의 공간적 일치(spatial coherence)을 지원하는 자국없이 결합된 환경을 제공한다. 또한 TELEPORT 는 참여자들이 다른 참여자의 반응, 몸짓, 그리고 응시등을 지각할 수 있도록 해 준다.

마지막으로, 기존의 분산가상환경(Distributed Virtual Environment)[5,6]에 방송 프로그램이 콘텐츠(contents)로써 결합된 Inhabited TV[7]를 알아보자. 이 시스템은 분산가상환경을 이용하여 시청자들이 가상 공간에 아바타로 참여하여 채팅과 인터랙션을 지원한다. 또한 방송에서 사용되는 프로그램을 적용하여 방송에 참여할 수 있게 하였다. 실제 구현한 예인 "The Mirror"에서는 6개의 다자 참여형 가상 세계를 구축하여 시청자들이 이 가상 공간을 공유하면서 방송에 참여할 수 있게 하였다.

3. IVS 와 ITE 의 관계

3.1 물리적 환경(IVS) 그리고 가상적 환경(ITE)

우리는 원격지의 참여자들이 공유하는 공간을 하나의 환경으로 유지 하지 않고 물리적인 환경과 가상 환경을 동시에 유지하면서 참여자의 여건에 맞는 환경에서 자연스런 인터페이스를 제공하는 연구를 하였다[8].

본 연구에서도 방송자가 시청자들과 인터랙션하는 환경과 시청자들의 참여 환경은 다르다.

방송자는 방송을 진행하고 방송 내용을 전달하는 역할을 하므로 실제의 모습과 행동이 카메라를 통해서 시청자에게 전달되어야 하므로 방송자는 물리적인 환경에서 존재해야 하고 시청자들과 인터랙션 해야 한다. 반면에 시청자들에게 방송자와 다른 시청자들과 같은 공간에 존재하고 직접 방송에 참여하는 것 같은 몰입적인 환경을 제공하기 위해서는 가상환경이 필요하다.

따라서 IVS의 모든 참여자는 물리적 환경(physical environment)에서 존재하고(그림 2-a) ITE에 참여자는 가상 환경(virtual environment)에서 참여(그림 2-b)하게 된다.

이처럼 참여자의 환경이 일관적으로 유지됨으로써 직관적이고 자연스런 인터페이스를 제공하게 된다. IVS에서는 방송자가 참여자들을 물리적으로 인식하고 직접 만나는 것 같이 자연스럽게 인터랙션할 수 있다. 시청자들도 3차원 가상 환경에서 몰입감을 느낄 수 있을 것이다.



(a) Interactive Virtual Studio



(b) Immersive Tele-viewer Environment

[그림 2] 두 환경에서의 참여자들의 물리적인 관계

3.2 참여자와 대리자 모델

원격지에 존재하는 다수의 참여자들이 실시간으로 정보를 공유하고 대화를 할 수 있는 환경은 CSCW에서 많이 연구되어 왔다. CSCW의 2차원적인 인터페이스를 넘어서 직관적인 환경을 제공하는 DVE는 참여자간에 3차원 가상 환경을 공유하고 서로를 상호 인식(mutual awareness)하고 인터랙션할 수 있게 하였다. DVE에서는 각 참여자는 자신의 대리자로서 아바타(avatar)라고 하는 virtual human으로 표현된다. 이처럼 지역적으로 떨어진 참여자들 간의 자연스런 인터페이스를 위해서는 원격지의 참여자를 대신하여 그 사람의 형태와 위치를 보여주고 말과 행동 등을 흉내낼 수 있는 대리자가 필요하다.

인터랙티브 TV 방송에서도 방송자와 시청자들이 지역적으로 떨어져 있는 상황에서 로컬 참여자의 환경에 맞는 원격지의 참여자들의 적절한 대리자가 필요하다. 그림 3에서 본 연구의 인터랙티브 TV 방송 환경에서의 참여자와 대리자의 관계를 보여준다.

방송자는 물리적인 스튜디오(IVS)에서 실제로 참여하지만 원격지의 시청자 환경(ITE)에서는 비디오 이미지로써 참여하게 된다. 따라서 ITE의 시청자들은 방송자의 실제의 표정이나 제스처 등을 볼 수 있다.

반면에 시청자쪽에서는 두 가지 형태의 대리자가 필요하다. ITE에서 참여자는 기존의 DVE와 같이 가상 공간에서 아바타로써 존재한다. HMD를 쓰고 있는 시청자가 다른 참여자들의 아바타들과 가상 공간을 공유하고 서로 인터랙션을 할 수 있는 몰입적인 환경을 제공할 수 있다. IVS에서 시청자의 대리자인 더미헤드는 ITE의 시청자의 관점과 말 등을 흉내낼 수 있다. 만약 시청자가 HMD 등의 몰입적인 디스플레이 장치를 사용하지 않고는 시청자의 머리 움직임을 더미헤드에 반영할 수 없다.

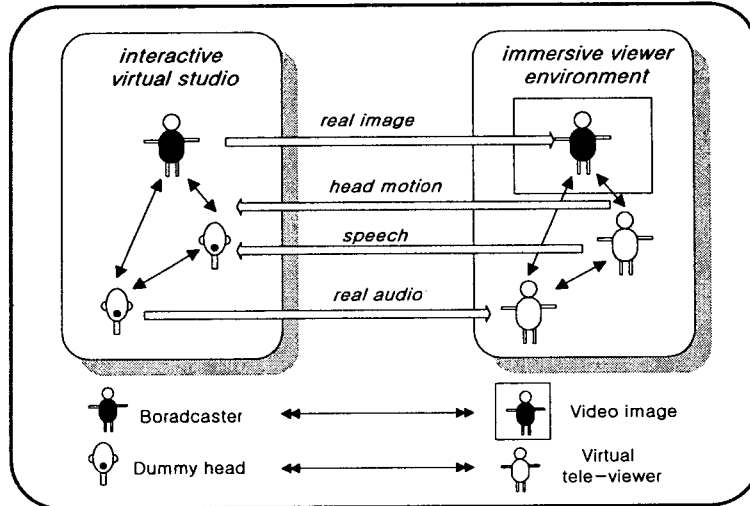


그림 3. IVS 와 ITE 와의 관계

4. 인터랙티브 가상 스튜디오(Interactive Virtual Studio, IVS)

4.1 참여자들 간의 물리적 인터랙션 제공

지역적으로 떨어진 참여자들이 정보를 교환하기 위해서는 네트워크 상에서 가상적으로 만난다. 3차원 가상 환경을 통해서 참여자들에게 좀더 직관적인 인터페이스를 제공하는 DVE에서도 참여자들 간의 물리적인 인터랙션을 제공하지는 못한다.

본 연구의 방송환경은 원격지의 참여자들에게 간접적으로 물리적 인터랙션을 제공함으로써 좀더 자연스런 인터페이스를 구현하였다.

방송자는 시청자의 대리자인 더미헤드를 통해서 원격지에 있는 시청자들과 물리적인 공간을 공유할 수 있고 마치 시청자와 직접 대화하는 방식으로 더미 헤드와 대화를 할 수 있다. 이를 통해 방송자는 자연스럽게 시선을 둘 수 있다.

이처럼 방송자와 시청자들이 물리적 인터랙션을 하기 위해서는 두 환경의 참여자의 물리적인 위치나 방향등이 일치해야 한다. 이것에 대한 설명은 5장에서 하기로 한다. 그림 3에서 보면 ITE의 방송자의 비디오 이미지와 아바타들의 위치와 방향은 IVS에서 위치한 방송자와 더미헤드의 물리적인 배치와 일치하도록 됨을 알 수 있다.

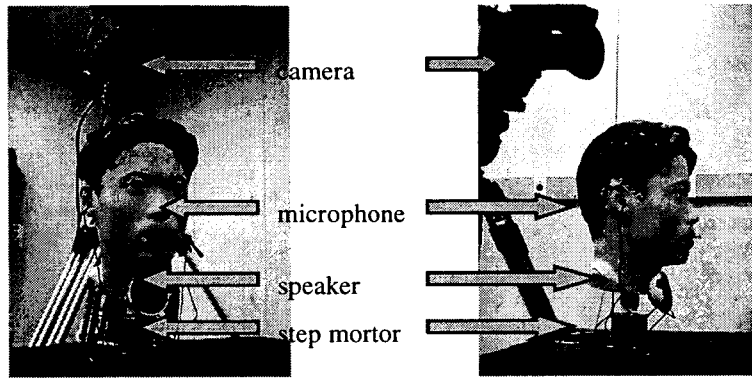
4.2 원격 존재자 로써 더미헤드(dummy head)

IVS의 더미 헤드는 시청자의 원격지의 물리적 대리자로서 역할을 수행하기 때문에 원격존재(tele-presence)기술의 원격 존재자(tele-operator)에 해당된다.

원격 존재 기술은 로컬에 있는 조작자가 원격지의 물리적 공간에 있는 대리자(tele-operator)를 통해 원격지에서 자신의 일을 할 수 있다. 이를 위해서는 tele-operator를 네트워크를 통해 조작할 수 있어야 하고 원격지의 물리적인 정보를 다시 조작자에게 피드백 시킨다.

더미 헤드는 시청자의 말과 머리 움직임을 통해 지역적으로 떨어진 스튜디오에서 물리적인 인터랙션을 하고 스튜디오로 부터 더미 헤드가 보는 영상과 오디오를 시청자에게 피드백 시켜준다.

더미헤드는 시청자의 말을 스튜디오에 전달하기 위한 스피커, 머리 움직임을 흉내내기 위한 스텝 모터가 장착되어 있다. 또한 시청자의 뷰에서 얻을 수 있는 영상과 오디오를 위해 카메라와, 양쪽 귀에 마이크로폰 등이 있다(그림 4).



[그림 4] Dummy head as tele-operator

5. 몰입형 시청자 환경(Immersive Tele-viewer Environment, ITE)

앞으로 ITE 가 갖고 있는 특징과 기능에 대해 알아보자

5.1 Augmentation by visual and aural queue :

우리가 살고 있는 실제 세계를 가상 공간에서 그대로 재현하는 것은 현재의 기술로써는 어렵다. 그래서 카메라에서 보는 실제 세계의 영상과 컴퓨터가 만들어 낸 화면을 합성하여 가상 공간을 확장(augmentation)하는 기술을 **Mixed Reality** 또는 **Augmented Virtuality**[9]라고 한다.

본 연구에서는 방송을 진행하는 방송자의 역할이 중요하므로 방송자의 얼굴 표정이나 제스처를 그대로 가져와서 가상 공간을 확장하였다. 또한 방송에 필요한 실제 소품들도 마찬가지로 방법으로 가상 공간에 들 수 있다(그림 4).

실제 비디오 영상을 이용한 가상 공간의 확장과 더불어 실제 오디오를 더미 헤드(Dummy Head)를 통해서 가상 공간으로 가져와서 시청자에게 좀더 몰입적인 환경을 제공하였다.

또한 실제 공간에서 참여자들 간의 물리적인 위치 관계도 가상 세계에 매핑 되어야 하고 테이블 같은 참여자와 관련된 물체는 가상 공간에서도 같은 위치와 크기로 존재해야 한다.

이처럼 실제 영상뿐만 아니라 실제 세계의 다양한 정보를 통해서 가상 공간을 확장하는 방법은 본 연구에서 최초로 시도되는 것이고 **Mixed Reality** 가 실제 세계의 영상만으로 확장되는 것이 아니고 실제 세계의 모든 정보를 이용할 수 있음을 보여주는 것이다.

다음 절에서는 실제 세계와 가상 세계간의 공간적인 일치를 구현하기 위한 방법을 소개한다. 본 연구에서는 실제 영상과 실제 오디오를 가지고 가상 공간을 확장하였기 때문에 이 두 가지 모두를 고려한 정합(registration)이 필요하다.

5.2 Visual registration

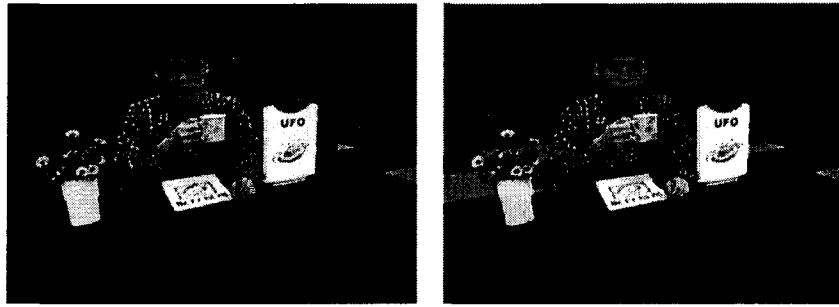
우리는 방송자의 실제 비디오 이미지를 가상 공간의 video wall 이라고 하는 panel 에 비디오 매핑을 함으로써 2D 이미지를 가상 공간의 virtual object 로 만들었다. 하지만, 비디오 이미지는 가상 공간과는 다른 실제 배경을 담고 있고 특정한 카메라 뷰로 고정되어 있다. 이 두 요소들은 공간적인 일치감을 방해하고 있고 결과적으로 방송자는 가상 공간의 TV 속에서 존재하는 것과 같이 된다.

본 연구에서는 visual 측면에서 공간적인 일치를 이루기 위해 virtual studio 기술에서 사용하는 크로마키잉을 사용한 subject extraction 과 실제 카메라와 가상 카메라의 차이를 없애기 위한 camera calibration 을 구현하였다.

Subject extraction :

실제 비디오 영상 안의 사람과 가상 세트와 합성하기 위해서 크로마키잉 기법을 사용한다. 일반적으로 가상 스튜디오는 blue screen 으로 둘러 싸여 있고 그 안에 앵커가 위치하게 된다. 크로마키잉 기법은 카메라가 잡은 영상에서 blue 인 부분에만 가상 세트가 보이도록 함으로써 마치 앵커가 가상 세트에 존재하는 것 같은 효과를 준다.

이와 마찬가지로 ITE 의 가상 환경에서 참여자를 넣기 위해서는 IVS 에 blue screen 을 설치하였다(그림 2)



(a) video mapping without alpha-key (b) video mapping with alpha-key
 [그림 5] 알파 키 정보를 가질 때와 없을 때의 비디오 매핑의 비교

기존의 가상 스튜디오 시스템과 기본적인 원리는 같지만 큰 차이점이 있다. ITE에서는 비디오 영상과 가상 세트의 영상을 외부의 크로마 키어로 합성하는 것이 아니고 알파 키 정보를 갖고 있는 비디오 영상을 가상 공간의 비디오 월에 매핑을 하여 가상 공간내의 virtual object로 만들었다. 이로써 HMD을 쓴 시청자가 제한이 따르지만 자신의 뷰를 바꿀 수 있게 된다. 만약에 기존의 가상 스튜디오와 같이 외부에서 합성하려고 한다면 시청자가 보고자 하는 방향과 일치하도록 카메라가 동작하여야 한다. 그림 5는 비디오 월에 방송자의 영상이 매핑된 화면을 보여준다. 그림 5-a는 알파 키 정보가 없는 영상일 때를 보여주고 5-b에서는 알파 키 정보를 가진 영상을 매핑 했을 때를 보여준다.

Camera calibration :

공간적인 일치감을 주기 위해서는 subject extraction 과 더불어 필요한 작업이 camera calibration 이다. 이는 영상이 매핑된 비디오 월의 real subject 들이 가상 공간에서 공간적인 일치감을 얻기 위해서 비디오 월의 위치, 방향 그리고 크기 등이 적절하게 조정되어야 한다.

먼저 앞에서 설명했듯이 IVS 와 ITE 의 참여자들은 두 환경에서 같은 공간적인 배치를 하고 있기 때문에 실제 카메라의 위치와 가상 카메라의 위치도 일치시켜야 한다.

다음으로 비디오 월의 subject 들의 크기와 위치가 가상 공간의 virtual object 들과 공간적 일치감이 없기 때문에 이를 조절해 주어야 한다. 이를 위해 IVS 에서 크기와 위치를 알고 있는 landmark 을 두고 ITE에서는 landmark 와 같은 크기의 virtual ruler 을 landmark 와 같은 위치의 가상 공간에 배치한다(그림 6-a). 그리고 이 둘의 차이를 비교하여 virtual ruler 와 일치하도록 video wall 의 landmark 을 조절한다. 이 과정을 그림 6-b,c,e 에서 보여준다.



(a) (b) (c) (d)
 [그림 6] An example of camera calibration

5.3 Aural registration

공간적인 일치감은 visual factor 뿐만 아니라 aural factor 에서도 필요하다. 실제 공간에서 얻은 오디오가 3 차원 가상 공간에서 visual factor 와 충돌되면 이런 공간적인 일치감이 깨지게 된다.

Aural registration 을 구현하기 위해서는 camera calibration 에서 언급했듯이 참여자들의 물리적인 배치가 ITE 에서 그대로 일치해야 한다. 즉 두 환경이 같은 coordination 을 가져야 한다. 다음으로 시청자가 머리를 움직일 때 더미 헤드 가 이를 흉내냄으로써 시청자의 뷰에 따라 해당되는 실제 소리를 듣게 된다.

6 구현 예 : TV 게임 쇼

우리는 이 새로운 방송 환경을 실험하기 위해 한 명의 방송 진행자와 두 명의 시청자가 참여하는 게임 쇼를 구현하였다.

시나리오는 방송 진행자가 미리 준비된 단어를 설명하면 두 명의 참여자가 그 단어를 맞춘다. 시청자가 단어를 맞추면 방송 진행자는 그 단어를 나타내는 참여자들에게 보여주고 그림을 보여 주고 다음 문제를 내게 된다. 이때 문제를 많이 맞추는 시청자가 우승자가 되고 경품을 타게 된다.



(a) 오른쪽 시청자

(b) 왼쪽 시청자

(c) 오른쪽 시청자

[그림 7] 게임 쇼의 장면

그림 7은 실제 게임 쇼를 진행할 때 HMD을 쓴 시청자들이 보는 장면이다. 그림 7-a는 탁자의 왼쪽에 앉아 있는 시청자가 문제를 답안을 들고 있는 방송 진행자를 보는 장면이고 7-b는 탁자 오른쪽에 앉아 있는 시청자가 보는 화면이다. 마지막으로 그림 7-c는 오른쪽에 위치한 시청자가 왼쪽 시청자를 보는 장면이다.

7. 결론

미래의 방송 서비스는 시청자들이 직접 참여할 수 있는 환경이 될 것이다. 지금까지 이러한 인터랙트 TV 방송을 위한 두 가지의 새로운 환경을 제시하였다.

하나는 원격지의 시청자들과 물리적 공간을 공유하고 인터랙션 할 수 있는 Interactive Virtual Studio 이다. 시청자는 이 환경에서 자신의 물리적 대리자인 더미 헤드를 통해서 참여하고 방송자는 이 더미헤드를 통해 원격의 시청자와 인터랙션한다.

또 하나의 방송 환경은 시청자가 HMD을 쓰고 몰입적인 환경에서 방송에 참여할 수 있는 Immersive Tele-viewer Environment 이다. 이 환경에서는 모든 참여자 간의 독립적인 관점에서 가상 공간을 공유할 수 있고 더미헤드를 통해 얻은 비디오 영상과 오디오로 가상 공간의 현실감을 높였다. 이 두 환경이 서로 쌍을 이루어서 각 참여자의 환경에 맞는 자연스런 인터랙티브 환경을 제공하고 있다.

또한 본 연구에 제시한 인터랙티브 방송 환경을 게임 쇼를 통해 가능성을 검증하였다.

이러한 인터랙티브 TV 방송 환경은 TV 방송 뿐만 아니라 앞으로 원격 만남, 원격 교육 등에서도 활용 할 수 있을 것이다.

참고 문헌

- [1] Heedong Ko, Hyunsuk Kim, Laehyun Kim, Jaehong Ahn, Kyungdong Park, "Development of a virtual studio system for live broadcasting of election results : VdreamSet", In Proceedings of IWAIT '99 in Taiwan
- [2] Augmented Reality, <http://www.cs.rit.edu/~jrv/research/ar/>
- [3] The Internet Foyer, <http://www.crg.cs.nott/ research/projects/Foyer/>
- [4] Christian J. Breiteneder, Simon J. Gibbs, Costas Arapis, "TELEPORT - An Augmented Reality Teleconferencing Environment", In Proceedings of 3th Eurographics Workshop on VE,
- [5] Laehyun Kim, Heedong Ko, Hyun suk Kim, "Hierarchical World Modeling using an object-oriented methodology for Distributed Virtual Environment", In Proceedings of 98'CG workshop in Taiwan
- [6] Rrrnie Roehl, Distributed Virtual Reality - An Overview, <http://sune.uwaterloo.ca/~broehl/distrib.html>
- [7] Inhabited Television, <http://www.crg.cs.nott research/applications/itv/>
- [8] Moonho Park, Laehyun Kim, Heedong Ko, and Hyeran Byun, "A Mixed Environment for Tele-meeting between Real human and Virtual human", To appear in IEEE/RSJ IROS '99
- [9] Milgrim, P. and Kishino, F., A Taxonomy of Mixed Reality visual Display, IEICE Transactions on Information Systems, Vol E77-D, No. 12, 1994