

# 사용자 상호작용 서비스를 위한 실감방송 기술

## Realistic Broadcasting Technologies for User Interaction Services

□ 호요성, 김성열 / 광주과학기술원 실감방송연구센터

### 1. 서론

디지털 전송 기술이 비약적으로 발전하면서 가까운 미래에 우리는 언제 어디서나 다차원 실감미디어(multi-modal immersive media)를 접할 수 있을 것이다. 다차원 실감미디어란 단순히 1차원인 소리와 2차원 영상 정보가 아닌 인간의 오감을 표현하는 요소 정보를 의미한다. 최근에는 자유시점 TV(free-viewpoint TV, FTV)와 관련하여 3차원 비디오에 대한 관심이 MPEG 표준화 그룹을 중심으로 높아지고 있다[1]. 이와 더불어, 3차원 TV[2] 및 실감방송[3]은 사용자 상호작용 서비스를 통해서 시청자에게 보다 생생한 방송 콘텐츠를 제공하는 차세대 방송으로 주목을 받고 있다.

실감방송은 시청자가 화면을 보는 시점에 따라 변하는 3차원 장면을 생생하게 재현할 뿐 아니라, 시청자에게 다시점 영상 정보와 연계된 3차원 소리 정보를 재생하여 현장에 있는 것과 같은 착각을 일으키게 한다. 또한, 실감방송은 시각과 청각을 넘어 인간의 오감을 통해서 방송을 보고 듣고 느낄 수 있는 방송 서비스를 제공할 것으로 기대된다. 2005년도에 광주과학기술원 실감방송연구센터[4]는 깊이 카메라 기술을 기반으로 홈쇼핑 실감방송 콘텐츠를 제작했다[5]. 일반적인 홈쇼핑 방송 콘텐츠는 시청자가 상품과 상품정보를 수동적으로 보고 결정하는데 반해, 제작된 홈쇼핑 실감방송 콘텐츠는 시청자가 상품을 3차원으로 보고 직접 만지거나 조작하여 상품의 특성을 쉽게 파악한 후, 상품을 구매하는 시

※ 본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 지식경제부 및 정보통신연구진흥원의 대학IT연구센터(ITRC)의 지원에 의한 것입니다.

나리으로 구성되었다.

홈쇼핑 방송을 테마로 제작된 이 실감방송 콘텐츠는 다차원 실감미디어 중 컴퓨터 그래픽스 모델, 깊이 정보를 포함한 3차원 비디오, 촉감 정보를 사용하여 제작되었다. 쇼핑 호스트는 깊이 카메라로부터 획득한 깊이 정보를 포함한 3차원 비디오를 이용하여 모델링되었고, 홈쇼핑 무대 배경과 홈쇼핑 상품은 컴퓨터 그래픽스 도구를 이용하여 생성한 컴퓨터 그래픽스 정보로 표현하였다. 특히, 컴퓨터 그래픽스 모델에 촉감 정보를 포함시켜 시청자가 상품의 재질을 체험할 수 있었다. <그림 1>은 2005년도에 제작된 홈쇼핑 실감방송 콘텐츠를 보여준다.

실감방송 기술과 관련해서 이미 유럽에서는 2002년 3월부터 2년 동안 ATTEST 프로젝트를 통해 3차원 TV에 관한 기초 기술을 연구했으며[6], 2004년 9월부터는 20여 개의 대학과 연구소가 컨소시엄을 구성하여 3DTV 프로젝트를 수행하고 있다[7]. 또한, 일본은 3차원 컨소시엄과 초입장감 커뮤니케이션 포럼(ultra-realistic communication forum)을 결성하여 3차원 TV 및 실감방송과 관련된 3차원 정보 처리기술을 통합하여 연구 개발하고 있다[8]. 미국에서는 NASA가 3차원 영상과 관련된 연구를 수행하고 있으며, Microsoft사는 다시점 영상 기반의 모델링 및 렌더링 기법을 개발하고 있다. 실감방송

과 관련하여 국내의 경우, 한국전자통신연구원에서는 2002년 월드컵에서 개발된 디지털 TV 시스템을 이용하여 3차원 TV 시범 방송 서비스를 제공했다. 또한, 한국과학기술연구원에서는 3차원 디스플레이 장치를 연구하고 있으며, 삼성전자와 LG전자에서도 3차원 카메라, 무안경식 스테레오 LCD 모니터, 다시점 입체 모니터를 연구 개발하고 있다.

본 논문은 사용자 상호작용 서비스를 지원하는 실감방송 기술을 소개하려고 한다. 본 논문의 구성은 다음과 같다. 2장에서는 다차원 실감미디어와 사용자 상호작용의 개념 및 종류에 대해 살펴본다. 3장에서는 다차원 실감미디어를 이용하여 사용자 상호작용을 지원하는 실감방송 콘텐츠 제작과정을 설명하고, 4장에서는 본 논문의 결론을 맺는다.

## II. 다차원 실감미디어와 사용자 상호작용

기존의 미디어는 단순히 1차원인 소리와 2차원 영상을 표현하는 데 중점을 두었다. 이러한 1차원 소리와 2차원 영상 정보만으로는 생생한 현장감을 전달하기가 어렵다. 그러나 최근에 컴퓨터 기술이 급속히 발전하여, 시청자에게 몰입감을 줄 수 있는 다차원



<그림 1> 실감방송 콘텐츠

실감미디어의 표현 및 처리가 가능하게 되었다. 여기서, 다차원 실감미디어란 공간과 시간의 제약을 극복하는 다양한 형태의 요소정보로 인간의 오감을 통해서 보고 듣고 느낄 수 있는 정보를 의미한다. 다차원 실감미디어의 구체적인 예로는, 컴퓨터 그래픽스 모델 (computer graphic model), 다시점 비디오 (multi-view video), 다채널 오디오 (multi-channel audio), 촉각 정보 (haptic data) 등을 들 수 있다.

컴퓨터 그래픽스 모델은 3차원 스캐너나 그래픽스 소프트웨어 도구를 이용하여 획득한다. 일반적으로 컴퓨터 그래픽스 모델은 기하학, 연결성, 광학적 정보로 이루어진 3차원 메쉬 모델로 표현된다 [9]. 3차원 메쉬 모델의 기하학 정보는 3차원 공간에서  $x, y, z$  좌표축에 대한 꼭지점(vertex)의 위치 정보이며, 연결성 정보는 3차원 모델의 표면을 정의하는 정보로서 꼭지점과 꼭지점 사이의 인접 관계를 나타낸다. 마지막으로 광학적 정보는 꼭지점의 색상, 법선 벡터, 텍스처 좌표 정보이다.

다시점 비디오는 여러 시점에서 다수의 카메라로 한 장면을 촬영한 다중 동영상의 집합이다 [10]. 다시점 비디오는 시청자에게 원하는 시점의 영상을 제공할 수 있는 다차원 실감미디어이다. 최근에 자유시점 TV 서비스를 위해 깊이 정보를 포함한 다시점 비디오 (multi-view video with depth information, MVD)에 대한 관심이 높아지고 있다 [11]. 다시점 비디오의 응용 분야는 자유시점 TV, 3차원 방송, 감시 카메라 영상, 파노라믹 영상, 홈 엔터테인먼트 등이 있다.

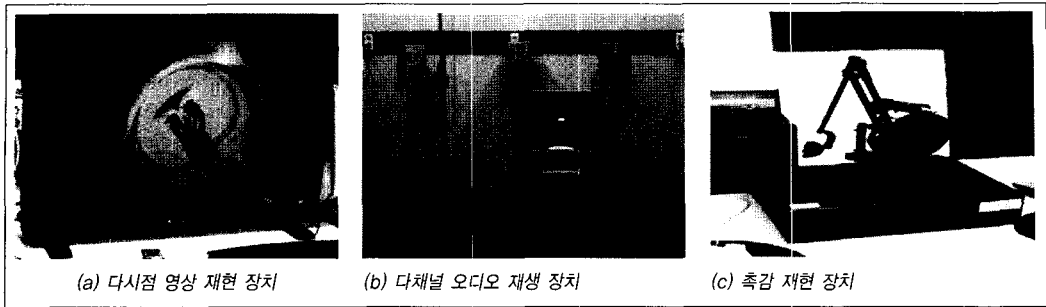
다채널 오디오는 다수의 마이크를 이용하여 여러 위치의 소리를 녹음하여 생성한 다중 오디오이다. 일반적으로 다채널 오디오는 머리 전달함수 (head related transfer function) 예측을 통해 소리의 방향과 거리를 유추하여 방향성을 지니는 소

리 정보를 제공한다. 최근에는 다채널 오디오와 다시점 비디오를 접목하여 시점에 따라 소리 정보를 제공하는 기술도 개발하고 있다 [12].

촉각 정보는 인간이 근육과 피부를 통해 느끼는 감촉에 관한 정보이다. 다시점 비디오와 다채널 오디오에 촉각 정보를 결합하면 눈으로 3차원 영상을 보고 귀로 3차원 소리를 들을 수 있을 뿐 만 아니라, 손을 이용하여 방송 콘텐츠를 만져볼 수 있어 몰입감을 극대화할 수 있다. 일반적으로 컴퓨터 그래픽스 모델의 촉각 정보는 범프맵으로 표현한다. 최근에는 깊이 정보를 촉각 정보로 간주하여 움직이는 물체를 만져볼 수 있는 촉각 시스템도 소개되었다 [13].

다차원 실감미디어를 이용한 사용자 상호작용 기술은 차세대 방송에서 시청자에게 현실감을 높여주는 중요한 요소이다. 사용자 상호작용을 통해서 시청자가 방송 콘텐츠에게 영향을 미치지거나, 방송 콘텐츠가 포함하고 있는 시나리오에 따라 시청자들이 반응하게 된다. 사용자 상호작용은 간접 상호작용 (indirect interaction)과 직접 상호작용 (direct interaction)으로 나눌 수 있다. 간접 상호작용은 방송 콘텐츠 자체가 아닌 미리 정해놓은 방송 콘텐츠의 시나리오에 따라 시청자가 반응하게 하는 것이다. 시청자가 보고자하는 방송 채널을 선택하는 것은 간접 상호작용의 가장 일반적인 예이다. 또한, 시청자가 현재 방송 콘텐츠를 전송하는 방송 채널 외의 다른 방송 채널로 전송되어진 추가적인 정보에 접근하거나 방송국에서 미리 정해놓은 인터넷 사이트에 접속하여 방송내용에 대한 소감을 적는 것도 간접 상호작용의 예이다.

직접 상호작용은 방송 콘텐츠에 따라 시청자가 직접 반응하게 하는 것이다. 직접 상호작용은 시청자의 오감을 자극하여 방송 콘텐츠로부터 실제와 같은 느낌을 전달해 주는 것이다. 직접 상호작용의 예로 자유시점 비디오 서비스, 다채널 오디오 서비스, 촉각



(그림 2) 직접 상호작용 장치

정보 서비스 등이 있다. 자유시점 비디오 서비스는 시청자가 보는 시점에 따라 해당 장면을 재현하여 시청자가 실제의 3차원 장면을 보는 듯한 착각을 하게 한다. 다채널 오디오 서비스는 음원의 위치를 여러 채널로 나누어서 시청자가 실제와 같은 오디오를 청취하게 한다. 촉각 정보 서비스는 미래형 방송에서 방송 콘텐츠를 직접 만지는 상호작용으로 시청자의 몰입감을 증대시키는데 중요한 역할을 할 것이다. <그림 2>는 직접 상호작용을 지원하는 장치를 보여준다.

현재의 방송 시스템은 제한된 형태의 간접 상호작용만 지원한다. 다시 말하면 지금까지의 상호작용 방송 서비스는 시청자가 방송 콘텐츠에 관심을 갖도록 하거나 시청자의 의견이 간접적으로 방송 콘텐츠에 영향을 미칠 수 있게 하는 정도이다. 따라서 현재의 상호작용형 방송이 실감나는 방송이라 하기엔 여전히 부족한 면이 많다. 게다가 우리는 지금까지 실제 방송에 이용할 수 있는 사용자 상호작용이 어떠한 것이 있는지 또 어떻게 사용자 상호작용 서비스를 실현할 수 있는지에 대한 이해와 연구도 부족한 상태이다. 본 논문은 사용자 상호작용 서비스를 지원하는 2차 홈쇼핑 실감방송 콘텐츠에 대한 제작기술을 소개하여, 실감방송 기술과 상호작용에 대한 이해를 돕고자 한다.

### III. 실감방송 콘텐츠 제작

실감방송 콘텐츠는 여러 가지 다차원 실감미디어를 결합하여 최종적으로 생성된 방송 콘텐츠를 의미한다. 실감방송 콘텐츠는 전체 3차원 콘텐츠(full 3D content)로서 다양한 직접 상호작용을 지원한다. 광주과학기술원 실감방송연구센터에서는 다차원 실감미디어를 이용하여 2차 실감방송 콘텐츠를 제작하였다. 홈쇼핑 방송을 주제로 다시 만들어진 2차 실감방송 콘텐츠는 2007년 ITRC 포럼에 전시되었다. 2차 홈쇼핑 실감방송 콘텐츠도 2005년도에 제작된 1차 홈쇼핑 실감방송 콘텐츠와 마찬가지로 자유시점 비디오, 다채널 오디오, 촉각 정보 서비스를 제공한다. 홈쇼핑 실감방송 콘텐츠를 제작하기 위해서 다차원 실감미디어 중 컴퓨터 그래픽스 모델, 깊이 비디오, 다채널 오디오, 촉각 정보를 사용하였다. 여기서, 깊이 비디오는 깊이 및 색상 정보를 표현하기 위해 2차원 비디오와 그에 대응되는 깊이 맵을 의미한다. 제작된 2차 홈쇼핑 실감방송 콘텐츠는 1차 홈쇼핑 실감방송 콘텐츠와 다음과 같은 차이점을 갖는다.

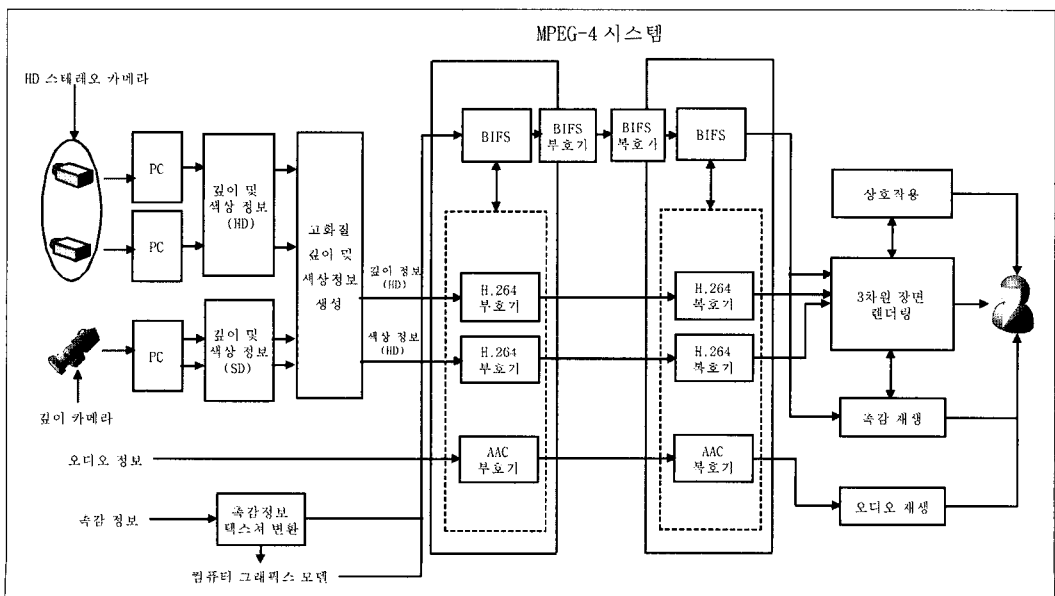
- 1) 1차 홈쇼핑 실감방송 콘텐츠는 깊이 카메라 기

술을 기반으로 제작한 표준 화질 (standard definition, SD)의 방송 콘텐츠인 반면, 2차 흡소핑 실감방송 콘텐츠는 복합형 카메라 시스템 (hybrid camera system) 을 이용하여 제작한 고 화질 (high definition, HD)의 방송 콘텐츠이다.

- 2) 1차 흡소핑 실감방송 콘텐츠는 자체적으로 개발한 장면 기술자를 이용하여 3차원 장면을 구성한 반면, 2차 흡소핑 실감방송 콘텐츠는 MPEG-4 이진 장면 기술자 (binary format for scene, BIFS)를 이용하여 3차원 장면을 기술하였다. 특히, 1차 흡소핑 실감방송 콘텐츠는 네트워크를 통한 스트리밍 서비스가 불가능 하였지만, 2차 흡소핑 실감방송 콘텐츠는 MPEG-4 멀티미디어 시스템을 기반으로 제작 되었기 때문에 스트리밍 서비스가 가능하다.
- 3) 2차 흡소핑 실감방송 콘텐츠는 깊이 카메라의 문제점을 세분화하고 이를 해결하였다. 1차 흡

소핑 실감방송 콘텐츠는 깊이 카메라에서 획득한 깊이맵의 광학적 잡음을 제거하는 데 초점을 두었지만, 2차 흡소핑 실감방송 콘텐츠는 광학적 잡음 제거 뿐 만 아니라, 경계 정제와 손실 영역 복원도 다루었다.

<그림 3>은 복합형 카메라 시스템을 이용하여 2차 흡소핑 실감방송 콘텐츠를 제작한 흐름도를 보여준다. 서버에서는 깊이 카메라와 고화질 스테레오 카메라를 결합하여 복합형 카메라를 구성한다. 구성된 복합형 카메라를 이용하여 고화질 깊이 및 색상 정보를 획득하고, 이를 각각 H.264 부호화기로 압축하여 두 개의 비디오 스트림을 생성한다. 한편, 오디오 녹음 장치를 이용하여 다채널 오디오 정보를 획득하고, 이를 AAC 부호화기로 압축하여 오디오 스트림을 생성한다. 또한, 촉감 정보를 범프맵으로 변환하고 컴퓨터 그래픽스 모델에 변환된 범



<그림 3> 흡소핑 실감방송 콘텐츠 제작 흐름도

프맵을 적용한다. 컴퓨터 그래픽스 모델 정보는 MPEG-4 BIFS에 의해 표현되고, BIPS 정보는 BIFS 부호화기를 이용하여 부호화된다.

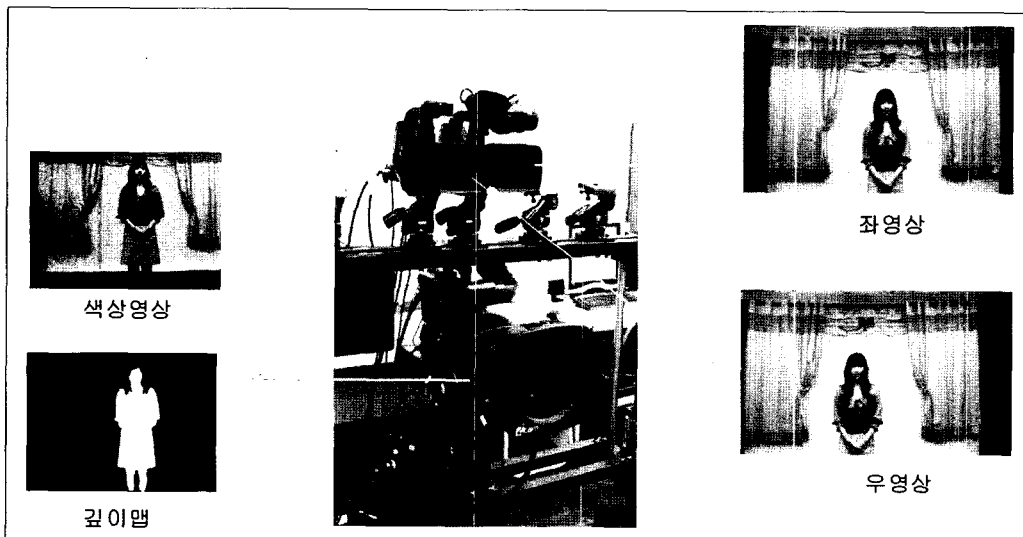
클라이언트에서는 BIFS 복호화기를 이용하여 BIFS 정보를 추출한다. BIPS 정보로부터 깊이 및 색상 정보의 스트림과 오디오 스트림의 위치를 파악하고 H.264 복호화기와 AAC 복호화기를 이용하여 각각 독립적으로 복호화한다. 또한, BIFS 정보로부터 컴퓨터 그래픽스 모델 정보와 촉감 정보를 추출한다. 그런 다음, 3차원 장면 재현기는 깊이 비디오 정보, 컴퓨터 그래픽스 모델 정보, 촉감 정보를 이용하여 3차원 장면을 렌더링하고, 오디오 재생기는 복호화된 오디오 신호를 재생한다. 3차원 장면은 깊이 비디오의 매 프레임마다 갱신된다.

### 1. 복합형 카메라를 이용한 동적 모델 생성

홈쇼핑 실감방송 콘텐츠의 동적 모델 정보인 쇼

핑호스트는 복합형 카메라 시스템으로부터 획득한 고화질 깊이 및 색상 정보를 이용하여 3차원 모델링된다[14]. 복합형 카메라 시스템은 두 개의 HD 카메라로 구성된 스테레오 카메라와 SD 깊이 카메라를 결합하여 구성된다. 각 카메라는 비디오 저장 보드가 갖추어진 PC와 연결되고, SD와 HD 동기화 신호를 동시에 발생시키는 동기화 신호 재생기가 모든 카메라에 연결되어 동기화 신호를 연속적으로 보낸다. 복합형 카메라 시스템은 매 프레임마다 스테레오 카메라에서 획득한 HD 좌우 영상과 깊이 카메라로부터 획득한 SD 색상 영상 및 깊이맵을 제공한다. <그림 4>는 복합형 카메라 시스템과 그 출력 영상을 보여준다.

쇼핑호스트의 3차원 깊이 정보를 추출하는 방법은 깊이맵 향상 단계, 쇼핑호스트 영역 추출 단계, 쇼핑호스트 영역 깊이맵 생성 단계로 나누어진다. 깊이맵 향상 단계는 깊이 카메라로부터 획득한 깊이맵에 포함된 광학적 잡음을 최소화하고, 손실된



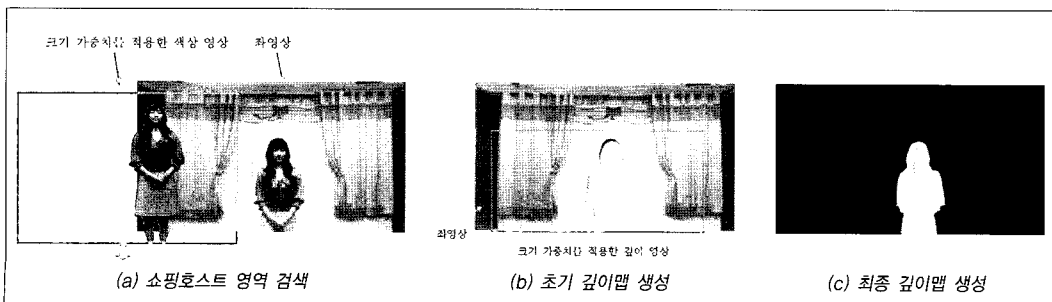
<그림 4> 복합형 카메라 시스템

쇼핑호스트 머리 영역의 깊이 정보를 복원한다[15]. 깊이맵의 광학적 잡음을 최소화하기 위해서, 우선 중간값 필터를 깊이맵에 적용하고 깊이맵을 다운샘플링한다. 그런 다음, 양측성 필터 (bilateral filter)를 다운샘플링된 깊이맵에 적용하고 선형 보간법을 이용하여 향상된 깊이맵을 생성한다. 깊이 카메라는 사람의 검은 머리카락과 같이 적외선을 흡수하는 영역에 대해 깊이 정보를 제공하지 못한다. 쇼핑호스트의 머리 영역은 깊이 카메라로부터 깊이 정보가 손실된 대표적인 영역이다. 쇼핑호스트의 머리 영역에 대한 깊이 정보를 복원하기 위해서, 쇼핑호스트의 얼굴 영역을 추적한 후, 얼굴 영역 주변에 다중 씨드(seed)를 할당한다. 씨드와 씨드 주변에 위치한 화소값을 머리색을 나타내는 임계치와 비교하여 머리 영역을 추적한다. 그런 다음, 얼굴 경계에 있는 깊이 정보를 고려한 비선형 보간법을 이용하여 머리 영역의 깊이 정보를 복원한다.

쇼핑호스트 영역 추출 단계는 스테레오 카메라로부터 획득한 HD 좌영상과 깊이 카메라로부터 획득한 SD 색상 영상으로부터 쇼핑호스트의 영역을 탐색하는 것이다. 쇼핑호스트 영역 추출 단계는 크기 가중치 선택 단계, 쇼핑호스트 영역 선택 단계, 크기 가중치 적용 단계로 나누어진다. 크기 가중치 선택 단계는 쇼

핑호스트를 복합형 카메라로 촬영하기 전에 미리 찍어둔 영상 패턴을 분석한다. 스테레오 카메라로 얻은 HD 영상 패턴과 깊이 카메라로 얻은 SD 영상 패턴의 셀 크기를 비교하여 크기 가중치를 계산한다. 쇼핑호스트 영역 선택 단계는 깊이 카메라로부터 획득한 깊이맵의 깊이 정보가 0이 아닌 색상 영역을 찾는다. 마지막으로 크기 가중치 적용 단계는 크기 가중치 선택 단계에서 계산된 크기 가중치를 SD 색상 영상과 SD 깊이맵의 쇼핑호스트 영역에 적용한다.

쇼핑호스트 영역 깊이맵 생성 단계는 크게 쇼핑호스트 영역 검색 단계와 쇼핑호스트 깊이맵 보완 단계로 나누어진다. 쇼핑호스트 영역 검색 단계는 <그림 5(a)>와 같이 크기 가중치가 적용된 쇼핑호스트 영상과 좌영상을 정합하여 좌영상에서의 쇼핑호스트 영역을 검색한다. 쇼핑호스트 영역 검색 단계에서 사용하는 정합 함수(matching function)로 최소 차분 절대치의 합(sum of absolute difference)을 이용한다. 쇼핑호스트 영역을 좌영상에서 추적한 후, 좌영상에 벗어난 영역은 제거된다. 쇼핑호스트 깊이맵 보완 단계는 스테레오 카메라로부터 획득한 좌영상과 우영상에 색상 분할에 기반한 스테레오 정합 알고리즘을 적용하여 좌영상의 깊이맵을 생성한다[16]. 또한, <그림 5(b)>와 같이, 쇼핑호스트 영역 검색 단계에서



<그림 5> 쇼핑호스트 영역 깊이맵 생성

추출한 쇼핑호스트 영역에 크기 가중치를 적용한 깊이맵을 덮어서 쇼핑호스트 초기 깊이맵을 생성한다. 그런 다음, 스테레오 매칭으로 얻어진 좌영상의 깊이맵과 초기 깊이맵의 경계부분을 정제하여 <그림 5(c)>와 같은 최종 쇼핑호스트 깊이맵을 생성한다.

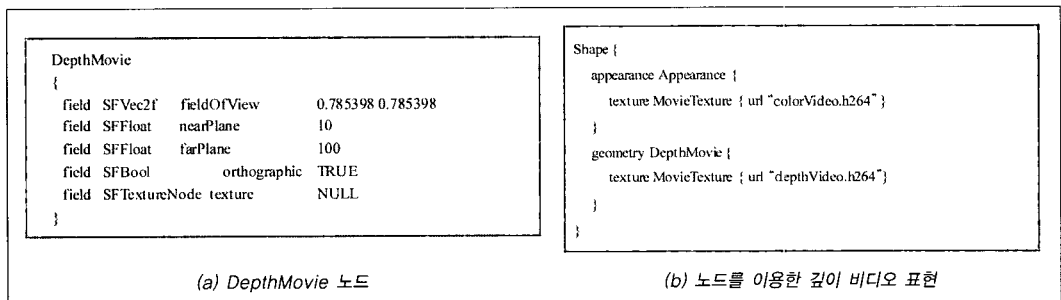
## 2. MPEG-4 멀티미디어 시스템

방송 콘텐츠를 클라이언트에 전송하기 위해 멀티미디어 시스템이 필요하다. MPEG-1과 MPEG-2와 같은 기존의 멀티미디어 시스템은 사용자 상호작용 기능에 초점을 맞추기 보다는 오디오 비디오 정보의 부호화와 오디오와 비디오 간의 동기화 문제에 초점을 맞추었다. 실감방송 콘텐츠를 효과적으로 표현하고 전송하기 위해, 본 논문에서는 유동적으로 사용자 상호작용 서비스를 제공할 수 있는 MPEG-4 멀티미디어 시스템을 소개한다. 기존의 멀티미디어 시스템과 MPEG-4 멀티미디어 시스템의 가장 큰 차이점은 객체 기반의 오디오 비디오 표현 모델 (object-based audio-visual representation model)이다[17]. MPEG-4 멀티미디어 시스템에서 객체 기반의 3차원 장면은 시공간적으로 관계가 있는 객체들의 조합이다. 결과적으

로, MPEG-4 멀티미디어 시스템은 깊이 및 색상 정보나 컴퓨터 그래픽스 모델과 같은 다양한 다차원 실감미디어를 결합할 수 있고, 시청자에게 직접 상호작용이 가능한 3차원 장면을 제공한다.

MPEG-4 멀티미디어 시스템은 3차원 장면을 기술하는 장면 기술자인 MPEG-4 BIFS를 제공한다. BIFS는 독립적인 객체들이 시공간적으로 어떻게 결합되는지 정의한다. 일반적으로 3차원 장면에 포함된 모든 객체는 BIFS의 Shape 노드(node)에 의해 기술되는데, Shape 노드는 외관 정보와 기하학 정보를 포함한다. 외관 정보는 MovieTexture 노드를 이용하여 2차원 비디오 정보로 표현된다. 그러나 불행히도 MPEG-4 멀티미디어 시스템은 외관 정보에 대한 기하학 정보 노드를 지원하지 않는다. 따라서, 본 논문에서는 깊이 비디오를 표현할 수 있는 새로운 노드인 DepthMovie 노드를 추가한다. <그림 6(a)>는 기하학 정보를 저장하는 새로운 DepthMovie 노드를 보여준다.

DepthMovie 노드의 상위 4개의 필드(field)는 카메라 매개변수를 저장하는 DepthImage 노드의 필드와 같다[18]. 마지막 필드인 texture 필드는 일반적인 2차원 비디오를 지칭하는 형태를 가지며, 깊이 비디오 정보를 MovieTexture 노드를 이용하여 정



<그림 6> MPEG-4 노드 정의



의하게 한다. <그림 6(b)>는 DepthMovie 노드와 MovieTexture 노드를 이용하여 깊이 비디오를 기술한 BIFS의 예를 보여준다.

H.264 부호화기에 의해 압축된 깊이 및 색상 정보 스트림과 BIFS 부호화기로 압축된 BIFS 스트림은 MPEG-4 멀티미디어 시스템이 제공하는 표현 형식에 따라 MP4 파일로 다중화된다. 생성된 MP4 파일은 Darwin 스트리밍 서버와 같은 전송 서버를 이용하여 IP 네트워크를 통해 전송되어 재생되거나 [19], CD나 하드디스크와 같은 국부적 저장 장치로부터 바로 읽어 재생된다. MPEG-4 멀티미디어 시스템을 구현하기 위해 본 논문은 GPAC 패키지를 이용하였다[20]. GPAC 패키지는 MPEG-4 멀티미디어 시스템을 쉽게 구현할 수 있도록 제공되는 참조 소프트웨어이다. GPAC 패키지는 Osmo라는 MPEG-4 재생기를 지원하는데, 생성된 MP4 파일을 재생하는데 이용하였다.

### 3. 사용자 상호작용 서비스 지원

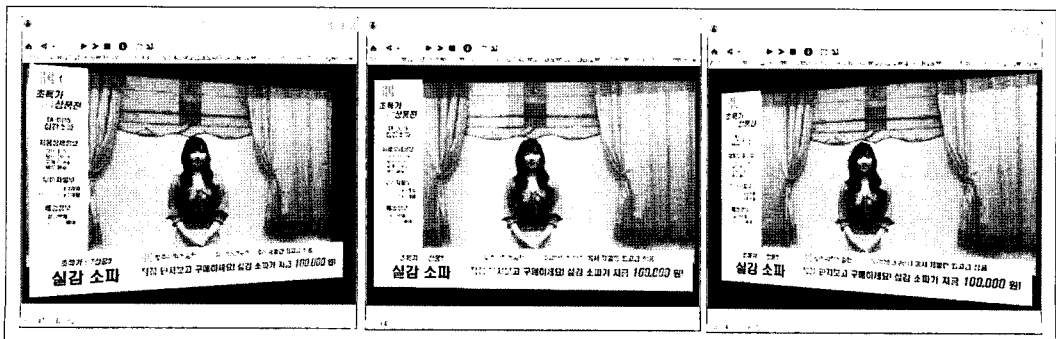
<그림 7>은 제작한 2차 홈쇼핑 실감방송 콘텐츠를 보여준다. 2차원 홈쇼핑 실감방송 콘텐츠는 스트리밍이 가능한 MP4 파일이고, 메쉬 삼각화 기법을



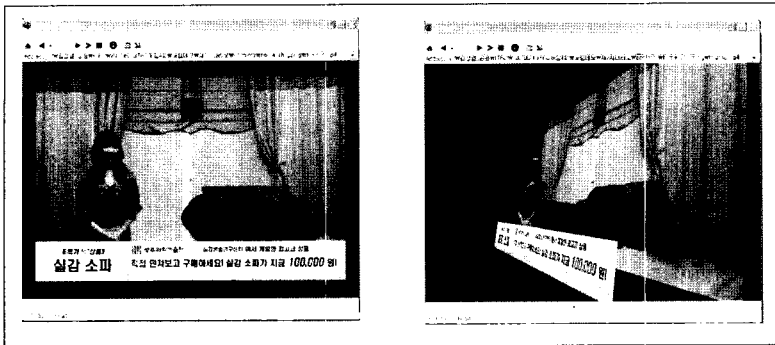
<그림 7> 홈쇼핑 실감방송 콘텐츠

이용하여 3차원 장면을 렌더링하였다[21]. 홈쇼핑 실감방송 콘텐츠는 복합형 카메라 시스템을 이용하여 깊이 및 색상 영상을 생성하여 제작하였기 때문에 고화질 콘텐츠이다. 또한, 자유시점 영상 생성, 컴퓨터 그래픽스 모델과의 합성, 촉감 제시와 같은 사용자 상호작용 기능을 제공한다.

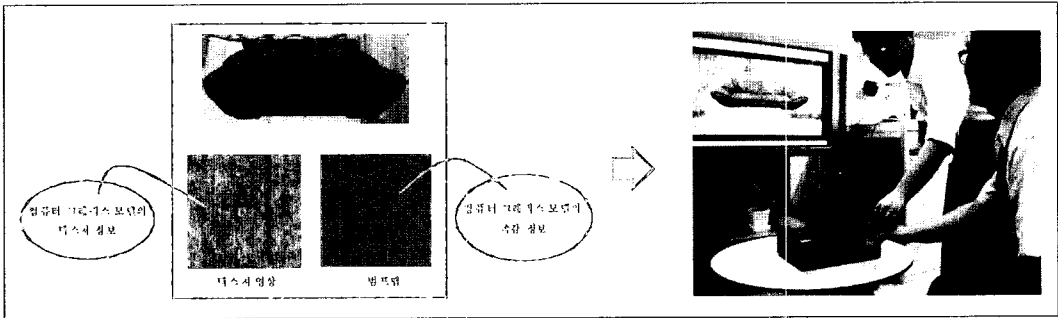
제작한 홈쇼핑 실감방송 콘텐츠는 쇼핑몰스트, 배경, 자막과 같이 3차원 장면에 포함된 모든 객체를 메쉬 삼각화 기술을 이용하여 3차원 모델링하였기 때문에, 시청자가 가상 카메라의 위치를 보고자



<그림 8> 자유시점 영상 서비스



〈그림 9〉 컴퓨터 그래픽스 모델과 합성



〈그림 10〉 촉각 제시

하는 시점으로 변경하여 원하는 시점의 영상을 즐길 수 있다. 〈그림 8〉은 자유시점 영상을 재현한 결과를 보여준다.

또한, 홈쇼핑 실감방송 콘텐츠는 MPEG-4 BIFS를 이용하여 3차원 장면을 표현하고 기술하였기 때문에, 컴퓨터 그래픽스 모델과 쉽게 합성된다. 〈그림 9〉에서 소파는 그래픽스 소프트웨어 도구로 제작한 것으로서, 깊이 및 색상 정보로부터 생성된 동적 모델인 쇼핑호스트와 쉽게 결합된다.

마지막으로 제작한 홈쇼핑 실감방송 콘텐츠는 촉각 제시 상호작용을 지원한다. 〈그림 10〉에서 보듯이, 컴퓨터 그래픽스 모델의 텍스처 정보와 촉각 정보를 영상으로 변환하여 컴퓨터 그래픽스 모델 정

보에 적용한다. 결과적으로, 촉각 정보가 포함된 컴퓨터 그래픽스 모델은 촉각 장치와 연동되어 손으로 컴퓨터 그래픽스 모델을 직접 만질 수 있다.

#### IV. 결론

지금까지 사용자 상호작용 서비스를 위한 실감방송 기술을 대략적으로 살펴보았다. 복합형 카메라 시스템로부터 3차원 비디오를 위한 실감미디어를 획득하였고, 획득한 실감미디어를 편집하고 제작하여 방송 콘텐츠를 생성하였다. 상호작용을 지원하는 3차원 방송 및 실감방송은 방송과 통신기술의 융

함으로 점점 늘어가는 시청자의 다양한 요구에 부응할 수 있는 미래형 방송 서비스이다. 다양한 실감 미디어를 효과적으로 획득, 편집, 압축, 전송 및 재현함으로써 시청자에게 보다 질 높은 방송 서비스

를 제공할 수 있으며, 차세대 방송과 관련된 요소 기술을 연구 개발하고 이에 대한 적극적인 국제 표준화 활동을 추진한다면 세계시장에서 우리나라가 유리한 고지를 선점할 수 있을 것이다.

### 참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11 N8944, "Preliminary FTV Model and Requirements," 2007.
- [2] C. Fehn, R. Barre, and S. Pastoor, "Interactive 3-D TV - Concepts and Key Technologies," Proceedings of the IEEE, vol. 94, no. 3, pp. 524-538, 2006.
- [3] 호요성, 김성열, "다차원 실감미디어와 실감방송 기술의 이해," 방송과 기술, vol. 108, pp. 90-97, 2004.
- [4] 광주과학기술원 실감방송연구센터, <http://rbrc.gist.ac.kr/>
- [5] J. Cha, S.M. Kim, S.Y. Kim, S.U. Yoon, I. Oakley, J. Ryu, K.H. Lee, W. Woo, and Y.S. Ho, "Client System for Realistic Broadcasting: A First Prototype," Lecture Notes in Computer Science, vol. 3768, pp. 176-186, 2005.
- [6] A. Redert, M. Beeck, C. Fehn, W. IJsselsteijn, M. Pollefeys, L. Gool, E. Ofek, I. Sexton, and P. Surman, "ATTEST: Advanced Three-Dimensional Television System Technologies," Proc. of International Symposium on 3D Data Processing, pp. 313-319, 2002.
- [7] Integrated Three-Dimensional Television-Capture, Transmission, and Display, <http://3dtv.zcu.cz/>.
- [8] Ultra-realistic Communications Forum, <http://www.scot.or.jp/urcf/>.
- [9] G. Taubin and J. Rossignac, "Geometry Compression through Topological Surgery," Proc. of SIGGRAPH, pp. 84-115, 1998.
- [10] P. Markle, A. Smolic, K. Muller, T. Wiegand, "Efficient Predictin Structures for Multiview Video Coding," IEEE Trans. on Circuits and Systems for Video Technology, vol. 17, no. 11, pp. 1461-1473, 2007.
- [11] ISO/IEC JTC1/SC29/WG11 M15090, "Improvement of Depth Map Estimation and View Synthesis," 2008.
- [12] 김홍국, "Audio Processing for Multiview Systems," DSP 워크샵, 2007.
- [13] J. Cha, S.Y. Kim, Y.S. Ho, and J. Ryu, "3D Video Player System with Haptic Interaction Based on Depth Image-Based Representation," IEEE Trans. on Consumer Electronics, vol. 52, no. 2, pp. 477-484, 2006.
- [14] 김성열, 이은경, 호요성, "깊이 카메라와 스테레오 카메라를 이용한 관심영역 향상 깊이맵 생성 방법," 영상처리 및 이해에 관한 워크샵, pp. 214, 2008.
- [15] J.H. Cho, I.Y. Chang, S.M Kim, and K.H. Lee, "Depth Image Processing Technique for Representing Human Actors in 3DTV Using Single Depth Camera," Proc. of 3DTV conference, paper no. 15. 2007.
- [16] C. Zitnick, S. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality Video View Interpolation Using a Layered Representation," Proc. of SIGGRAPH, pp. 600-608, 2004.
- [17] F. Pereira, "MPEG-4: Why, What, How, and When," Signal Processing Image Communication, vol. 15, pp. 1071-1085, 1995.
- [18] L. Levkovich-Maslyuk, A. Ignatenko, A. Zhirkov, A. Konushin, I. Park, M. Han, and Y. Bayakovski, "Depth Image-based Representation and Compression for Static and Animated 3-D Objects," IEEE Trans. on Circuits and Systems for Video Technology, vol. 14, no. 7, pp. 1032-1045, 2004.
- [19] Darwin Streaming Server, <http://developer.apple.com/opensource/server/streaming/>.
- [20] GPAC project, <http://gpac.sourceforge.net/>.
- [21] S.Y. Kim and Y.S. Ho, "Hierarchical Decomposition of Depth Map Sequences for Representation of Three-dimensional Dynamic Scenes," IEICE Trans. on Information and Systems, vol. E90-D, no. 11, pp. 1813-1820, 2007.

필자소개



호요성

- 1981년 2월 : 서울대학교 전자공학과 (학사)
- 1983년 2월 : 서울대학교 전자공학과 (석사)
- 1983년 3월 ~ 1995년 9월 : 한국전자통신연구원 선임연구원
- 1989년 12월 : University of California, Santa Barbara Department of Electrical and Computer Engineering (박사)
- 1990년 1월 ~ 1993년 5월 : 미국 Philips 연구소 Senior Research Member
- 1995년 9월 ~ 현재 : 광주과학기술원 정보통신공학과 교수
- 주관심분야 : 디지털 신호처리, 영상신호처리 및 압축, 멀티미디어 시스템, 디지털 TV와 고선명 TV, MPEG 표준, 3차원 TV, 실감방송



김성열

- 2001년 2월 : 강원대학교 정보통신공학과 (학사)
- 2003년 2월 : 광주과학기술원 정보통신공학과 (석사)
- 2003년 2월 ~ 현재 : 광주과학기술원 정보통신공학과 박사과정
- 2007년 12월 ~ 현재 : 한국전자통신연구원 위촉연구원
- 주관심분야 : 3차원 메쉬 모델링 및 압축, 디지털 신호처리, 멀티미디어 시스템, 3차원 TV, 실감방송