

방송과 증강현실의 접목을 위한 서비스 시나리오와 증강방송 메타데이터 설계

최범석, 김순철, 정영호, *이원돈

한국전자통신연구원

bschoi@etri.re.kr, choulsim@etri.re.kr, yhcheong@etri.re.kr, *wlee@cnu.ac.kr

Use Cases and Metadata Design for Integration of Broadcasting Service with Augmented Reality

Bumsuk Choi Soonchoul Kim Youngho Jeong, *Wondon Lee

Electronics & Telecommunications Research Institute, *Chungnam National Univ.

요 약

본 논문에서는 최근 모바일 디바이스에서 제공되고 있는 증강현실 서비스를 방송환경에 접목할 수 있는 서비스 시나리오에 대한 소개와 이를 실현하기 위한 증강방송 메타데이터에 대하여 소개한다. 방송환경이 모바일 환경과 다르기 때문에 완벽한 증강 서비스를 제공하는데 있어서 한계가 있다. 한편 TV의 대 화면과 양질의 TV 프로그램, 스마트 TV로의 발전에 따른 웹 환경 지원과 모션/음성 인식 인터페이스, 그리고 스마트 TV 애플리케이션의 등장은 증강방송의 가능성을 한층 높이고 있다. 이를 가능하게 하기 위하여 증강방송 메타데이터를 설계하였으며, 증강방송 메타데이터는 증강영역 및 시간, 증강 콘텐츠, 정합 정보 등을 포함한다. 증강방송 메타데이터를 방송환경에 적합하게 전송하기 위하여 스트리밍 서비스에 적합하도록 메타데이터 구조를 정의하였다.

1. 서론

최근 증강현실 기술은 모바일 또는 태블릿 PC에서 활발하게 개발되고 있다. 증강현실 분야도 광고, 오락, 교육, 여행 등 광범위해지고 있으며, 이에 대한 활용성도 급격하게 증가되고 있는 추세이다. 증강현실 서비스는 간단하게 설명하자면 실 세계를 사용자가 카메라를 통하여 비춰보면서 증강 콘텐츠를 오버레이 형태로 제공받는 서비스이다. 이 분야에 대한 전문가들은 증강현실(AR)을 가상현실(VR)과는 다른 개념으로 보고 있다[1]. 가장 큰 차이점은 가상현실에서는 실 세계와 인터랙션이 존재하지 않지만 증강현실은 존재하며 심지어 실 세계에 영향을 미치는 점이라고 설명한다. 모바일 환경에서 증강현실 앱들이 인기를 끄는 이유는 터치스크린과 같은 사용자 인터페이스와 방향, 가속도, GPS 센서 등을 제공하기 때문이다. 다음 그림은 증강현실 서비스의 대표적인 예로서 마커(marker) 위에 증강 콘텐츠를 오버레이 하여 디스플레이 하는 방식을 보여준다.



그림 1 - AR 서비스 예제

한편 방송 시청환경은 모바일 환경과는 다르게 방송 서버에서 일방적으로 미디어를 전송하는 방식으로 사용자가 자신의 의지대로 카메라의 방향을 바꿀 수 없다. 또한 모바일 디바이스에는 기본 장착되고 있는 다양한 센서들이 TV 단말에는 제공되지 않는다. 다시 말하면 증강현실 서비스를 제공하기 위하여 기본적으로 갖추어야 할 요소들이 방송환경에는 빠져있다[2]. 그럼에도 불구하고 TV에서도 부분적으로 증강방송을 제공하고 있다. 그림 2와 같이 스포츠 경기 방송에서 실물의 그라운드나 전광판에 광고물을 오버레이 하는 증강광고가 그것이다. 그러나 이러한 증강방송 서비스는 방송사에서 일방적으로 고정된 위치에 미리 정의한 그래픽 광고 콘텐츠를 오버레이 한 영상을 송출하는 방식이므로 진정한 증강방송이라고 보기 어렵다.



그림 2 - 전통적인 증강방송 예제

한편 최근 TV의 발전도 놀랍다. 최근 TV의 대명사로 알려진 스마트 TV는 인터넷 연결을 필수화하고 있으며 스마트폰과 동일하게 다양한 스마트 TV 애플리케이션들을 제공하고 있다. 또한 스마트 리모콘, 음성인식, 그리고 모션인식도 기본적으로 제공하고 있기 때문에 기존 TV 단말에 비하여 사용자 인터페이스가 대단히 편리해 졌다. 또한 셋탑박스의 성능이 점차 높아지면서 증강방송 서비스에 필요한 복잡한 연산이 가능해지고 있다. 마지막으로 방송환경의 가장 큰 장점인 대 화면과 고화질의 콘텐츠로 말미암아 적절한 증강방송 서비스를 개발한다면, 향후 TV에서도 다양한 증강현실 서비스를 제공받을 수 있을 것이다.

본 논문의 구성은 다음과 같다. 2 절에서는 증강방송을 위한 서비스 시나리오에 대해 살펴본 후, 3 절에서는 이러한 서비스를 제공하기 위한 증강방송 메타데이터에 대하여 설명한다. 마지막으로 4 절에서는 본 논문에 대한 결론을 맺는다.

2. 증강방송 시나리오

본 절에서는 방송환경에 적용할 수 있는 증강방송 서비스 시나리오에 대하여 소개한다. 일부 서비스는 현재의 방송환경에서 바로 적용할 수 있는 반면, 일부 시나리오는 향후 TV 단말이 성능이나 기능이 추가된다면 적용 가능한 시나리오이다.

2.1. 증강영역 기반 서비스

증강영역이란 비디오 장면에서 증강 콘텐츠가 오버레이 될 특정 영역을 의미한다. 방송 프로그램의 특성상 방송 프로그램 장면의 아무 곳이나 증강 콘텐츠를 오버레이 할 수 없다. 따라서 방송사에서 미리 정의한 영역 안에서만 증강 콘텐츠를 오버레이 하도록 제한할 필요가 있다. 그림 3은 증강영역을 활용한 증강방송의 예제이다. 증강영역으로 사용될 장면의 일부분을 미리 정의하고 그 위에 올라갈 증강 콘텐츠를 지정하면 단말에서 해당영역에 지정된 증강 콘텐츠를 오버레이 하여 보여주는 방식이다. 증강영역이 장면의 고정된 위치에 지속적으로 표현될 수도 있지만, 프레임 마다 증강영역의 위치가 달라질 수도 있다. 이때 증강영역이 활성화되는 시간과 소멸되는 시간을 정의하고 동영상의 매 프레임 별 증강영역의 움직임도 역시 표현해 주어야 한다[3].

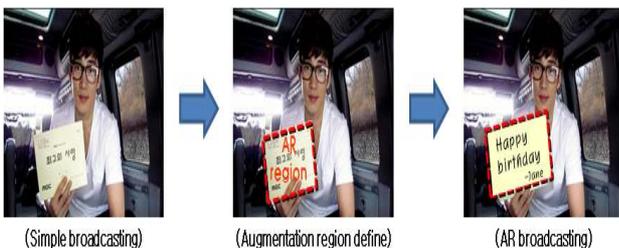


그림 3 - 증강영역 기반 서비스 예

2.2. 인터랙티브 증강방송 서비스

증강현실 서비스에서 빼놓을 수 없는 것이 사용자와의 인터랙션이다. 오버레이 된 증강 콘텐츠를 사용자가 직접 제어할 수 있다면 더욱 흥미로운 것이다. 그림 4의 예제처럼 TV 프로그램에 오버레이 된 인간의 장기를 나타낸 3D 모델을 사용자가 직접 돌려볼 수 있다면 교육의 효과가 더욱 증대될 것이다. 이를 위하여 다양한 방법이 가능하다. 최근 리모콘 안에 센서가 장착된 스마트 리모콘이 상용화 되면서 리모콘을 통하여 다양한 제어가 가능해졌다. 또한 모션인식기술까지 접목되어 리모콘 없이 제스처만으로 인터랙션이 가능하다. 기존의 증강현실 기술을 접목한다면 사용자가 마커이미지를 들고 있으면 TV에 장착된 카메라를 통하여 마커이미지의 위치를 추적하고 이를 증강 콘텐츠와 연동하므로 증강 콘텐츠를 제어할 수 있다.



그림 4 - 인터랙션 기반 증강 방송 서비스 예

2.3. 증강 서비스 제공자 선택

TV 채널을 사용자가 선택하듯이, 증강 콘텐츠를 사용자가 원하는 콘텐츠로 변경할 수 있다면 같은 TV 프로그램에도 다양한 종류의 증강 콘텐츠를 제공받을 수 있을 것이다. 모바일에서는 이미 비슷한 서비스를 제공하고 있다. “Layer Reality Browser” 라는 증강현실 앱(그림 5)은 다양한 타입의 증강 서비스가 올라갈 수 있는 플랫폼을 제공하고 있다. 예를 들어 위치기반 서비스 타입을 선택하면 증강된 거리영상에서 가까운 식당, 카페, 상점의 위치를 오버레이하여 보여준다. 마찬가지로 게임 타입을 선택하면 증강 게임을 즐길 수도 있다.

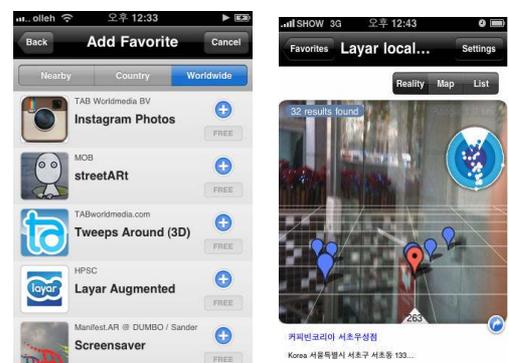


그림 5 - LayaR 앱의 서비스 타입 제공 예

방송국에서는 증강영역과 시간을 정의하고 TV 단말에서 사용자가 원하는 증강 콘텐츠를 선택할 수 있도록 하므로 증강방송에서도 동일한 서비스를 제공할 수 있다. 이러한 비즈니스 모델은 증강 앱 마다 별도의 증강 브라우저를

개발하는 방식에서 벗어나서 동일한 증강 브라우저 위에 다양한 증강 서비스 제공자가 자신의 콘텐츠 제공할 수 있다는 점에서 큰 장점이 있다.

2.4. 위치기반 증강방송 서비스

증강현실 서비스는 위치기반 서비스 앱에서부터 인기를 얻기 시작했다. 사용자가 증강 브라우저를 통하여 실 세계를 보면서 가고자 하는 장소의 위치를 직관적으로 알 수 있기 때문이다. TV 단말은 모바일 디바이스와 달리 위치기반 서비스를 하기 위한 어떠한 센서도 없으며, 사용자가 TV 방향을 바꾸어 보았자 방송 카메라의 위치가 바뀌지도 않는다. 그렇다면 방송환경에서 위치기반 증강방송 서비스를 적용할 방법은 무엇인가? 최근 방송 카메라에 GPS 센서가 부착된 장비가 개발되고 있다. GPS 센서가 없다면 방송 프로그램에 제작한 후에 장면에 대한 GPS 값을 메타데이터로 기술하여 전달할 수도 있다. 예를 든다면 골프 게임 방송을 할 때, 시청자는 골프 게임이 벌어지는 필드의 위치를 알고 싶어 할 수도 있고 골프 그린의 모양을 보고 싶어 할 수도 있다. 이 경우 증강 서비스 제공자는 방송 프로그램에 GPS 정보를 참고하여 사용자가 원하는 정보를 골프 방송에 오버레이 하여 제공할 수 있다.

3. 증강방송 메타데이터

증강방송 메타데이터란 방송 콘텐츠에 증강 콘텐츠를 오버레이하기 위해 필요한 정보 즉, 증강 콘텐츠가 표현되어야 할 영역이나 위치, 표현방법, 증강 콘텐츠의 타입, 증강 콘텐츠의 속성, 방송 콘텐츠 제작에 사용된 카메라 각종 센서 정보, 방송 콘텐츠와 증강 콘텐츠의 동기화를 위한 시간 정보 등을 포함하는 XML 기반의 메타데이터이다. 증강방송 메타데이터는 먼저 저작서버에서 방송 콘텐츠를 기반으로 사용자의 저작행위를 통하여 생성이 되며, 전송서버에서 방송 콘텐츠와 증강방송 메타데이터를 다중화하여 방송단말에 전달하게 된다. 방송단말은 증강방송 메타데이터를 방송 스트림에서 추출하고 이를 분석하고 방송 콘텐츠와 동기화 처리를 통하여 방송 콘텐츠에 증강 콘텐츠를 오버랩하여 표현한다. 증강방송 메타데이터의 구조에 대한 이해를 돕기 위하여 증강방송의 장면(Scene), 참조영역(Reference region), 증강영역(Augmented region), 증강객체(Augmented object)에 대한 개념을 보여주는 그림과 각 영역에 대한 의미는 아래와 같다[4].

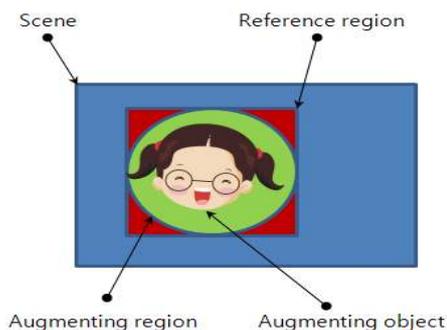


그림 6 - 증강방송 메타데이터 기본 개념

- Scene : 동영상의 장면
- Reference region : 증강영역의 3 차원 움직임을 추출하기 위한 영역. (단말에서 증강영역을 실시간으로 추적하기 위해서 레퍼런스 이미지로 사용할 영역)
- Augmented region : 증강 콘텐츠가 오버랩 되어 디스플레이 될 영역
- Augmented object : 증강 콘텐츠(객체)

증강방송은 기본적으로 전송 서비스 형태이므로 증강방송 메타데이터의 구조를 정의함에 있어서 메타데이터 전송에 용이하도록 구조를 정의할 필요가 있다. 방송단말에서 방송 콘텐츠와 증강 콘텐츠가 적절히 동기화되어 재현되기 위해서는 증강 콘텐츠가 표현되어야 할 타임정보가 매우 중요하다. 따라서 증강방송 메타데이터를 전송하는데 있어서 프레그먼트이션의 기준이 되어야 할 정보가 time stamp이며, 이러한 time stamp 를 포함하여 증강영역 또는 증강 콘텐츠의 업데이트 정보를 함께 instruction 단위로 묶어 전송한다. First instruction 에는 증강영역, 참조영역, 증강객체(콘텐츠) 등에 대한 모든 정보가 포함되지만, 이와 관련된 following instruction 에는 변화된 정보에 대한 내용만 포함시켜 전송하게 된다. First instruction 과 following instruction 들간의 관계는 instruction ID 를 레퍼런스하여 표현한다. 그림 7 은 증강방송 메타데이터 전송 개념을 나타낸다.

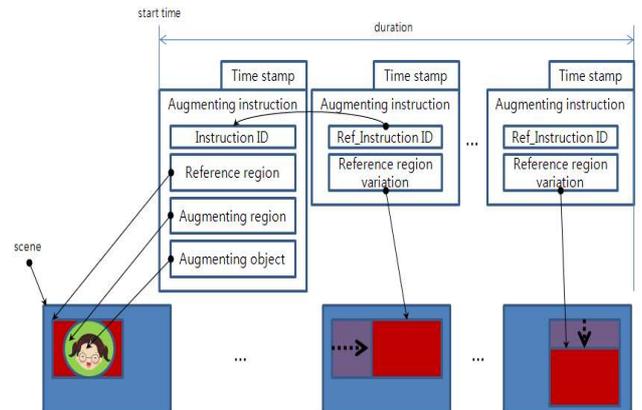


그림 7 - 증강방송 메타데이터 기본 개념

다음 그림은 증강방송 메타데이터의 상위 구조를 나타낸다.

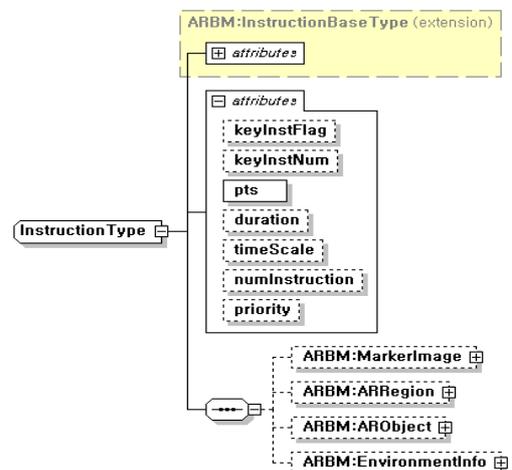


그림 8 - 증강방송 메타데이터 상위 구조

서브 엘리먼트 MarkerImage 는 방송 단말이 증강영역을 자동트래킹 하기 위한 레퍼런스 이미지를 포함한다. 또는 사용자가 마커이미지를 통하여 증강 콘텐츠를 제어하기 원할 경우 TV 단말에서 트래킹해야 할 마커이미지를 포함 할 수도 있다. 서브 엘리먼트 ARRegion 은 증강 콘텐츠가 오버레이 되어야 할 영역정보와 증강영역이 시간에 따라 움직일 경우 움직임에 대한 정보도 표현한다. 이러한 움직임은 보통 스케일, 로테이션, 위치이동의 3 가지 변수에 의하여 표현된다. 서브 엘리먼트 ARObjct 는 오버레이 될 증강 콘텐츠의 위치, 인코딩 정보, 포맷 등의 속성정보를 포함한다. 마지막으로 서브 엘리먼트 EnvironmentInfo 는 증강 콘텐츠의 자연스러운 정합에 필요한 요소, 예를 들어 광원에 대한 정보를 포함한다.

```
<Instruction id="Instruction 1" firstInstFlag="true" augNum="1" pts="90" duration="200">
  <ARRegion>
    <Coordinates x1="-0.181250" y1="0.636111" x2="0.482812" y2="0.636111" x3="0.482812" y3="0.636111">
    <SRT sx="0.006000" sy="0.006000" rx="0.000000" ry="0.000000" rz="0.000000" tx="0.000000" ty="0.000000" tz="0.000000">
  </ARRegion>
  <ARObjct clearFlag="false" service="entertain">
    <Remote>como.x</Remote>
  </ARObjct>
  <EnvironmentInfo>
    <Light type="0" color="#FFFFFF" intensity="">
      <Position px="0.00" py="0.00" pz="0.00">
    </Position>
    <Rotation wx="0.00" wy="0.00" vz="0.00">
    </Rotation>
  </Light>
  <Camera fov="0.00">
  </Camera>
</EnvironmentInfo>
</Instruction>
```

그림 9 – First instruction 샘플

그림 9 는 증강방송 메타데이터의 예를 보여준다. 본 예제는 증강영역에 대한 first instruction 을 나타내며, 증강 콘텐츠 (파일명: como.x)를 프리젠테이션 타임(pts) 90 에서 시작하여 200 타임까지 지속하라는 의미를 나타낸다. 증강영역은 4 점을 연결한 사각형으로 표현하고 있으며, 초기 증강영역의 위치, 스케일, 회전 정보를 (tx, ty), (sx, sy), 그리고 (rx, ry, rz) 형태로 표현하고 있다. 만일 증강영역이 프리젠테이션 시간동안 움직이지 않는다면 first instruction 만으로 충분하지만, 증강영역이 움직인다면 이러한 움직임이 있을 때마다 following instruction 들을 추가해 주어야 한다. 다만 following instruction 에는 증강영역의 움직임 정보만 표현해 주면 된다. First instruction 과 following instruction 사이의 관계는 augNum 에 의하여 표현된다. 그림 10 은 following instruction 의 예제를 보여준다.

```
<Instruction id="Instruction 2" firstInstFlag="false" pts="100" augNum="1">
  <ARRegion>
    <SRT tx="0.300000" ty="0.036111"/>
  </ARRegion>
</Instruction>
```

그림 10 – Following instruction 샘플

4. 결론

방송 환경은 모바일 환경과 달리 진정한 증강현실 서비스를 구현하기에 한계가 있다. 그러나 TV 단말이 갖고 있는 대 화면의 장점과 최근 스마트 TV 의 특징을 적절하게 조합할 수 있다면 시청자에게 매력적인 증강방송을 제공할 수 있다. 본 논문에서는 이러한 관점에서 방송환경에 접목할 수

있는 증강방송 서비스 시나리오를 제안하였다. 일부 시나리오는 현재의 방송환경에 바로 접목이 가능하지만 몇몇 다른 시나리오는 향후 방송단말의 성능이나 기능이 업그레이드 된다면 적용할 수 있을 것이다. 본 논문에서는 증강방송 서비스를 제공하기 위한 증강방송 메타데이터의 구조와 샘플 메타데이터를 통하여 설명하였다.

향후 연구로서, 이러한 증강방송 메타데이터를 사용자가 편리하게 제작할 수 있도록 GUI 기반의 저작툴을 개발할 예정이며, 증강영역의 움직임을 효과적으로 추출하기 위한 레퍼런스 이미지 기반 자동추출 알고리즘을 저작툴에 적용할 예정이다. 마지막으로 백그라운드 동영상과 증강 콘텐츠를 자연스럽게 정합하기 위한 환경정보를 정의하고 이를 단말에서 이용하여 디스플레이 하는 방법에 대한 연구를 진행할 예정이다.

5. 감사의 글

본 논문의 연구결과는 한국방송통신부(KCC)에서 지원하는 "Beyond Smart TV 기술 개발"(11921-03001) 과제로부터 산출되었습니다.

6. 참조문헌

- [1] Tobias Daniel Kammann, "Interactive Augmented Reality in Digital Broadcasting Environments" Diploma Thesis, November, 2005.
- [2] digitalfernsehen.deInteraktivität: Japaner setzen auf Open TV
http://www.digitalfernsehen.de/news/news_46594.html, Sept., 2005.
- [3] ISO/IEC JTC1/SC29/WG11, m22476, B. S. Choi, Young Ho Jeong, Geneva, Swiss, November 2011.
- [4] Layar application product description;
<http://www.layar.com>, May, 2012.
- [5] ISO/IEC JTC1/SC29/WG11, m23635, Junghak Kim, Bum-suk Choi, San jose, USA, February 2012.
- [6] Layar application product description;
<http://www.layar.com>, May, 2012.
- [7] ISO/IEC JTC1/SC29/WG11, m23635, Junghak Kim, Bum-suk Choi, San jose, USA, February 2012.
- [8] ISO/IEC JTC1/SC29/WG11, m21494, B. S. Choi, Young Ho Jeong, Geneva, Swiss, November 2011.