논문 2012-07-02

스마트 TV 환경을 위한 4D 미디어 방송 서비스 시스템의 설계 및 구현

(Design and Implementation of the 4D-Media
Broadcasting Service System in a Smart TV Environment)

윤재관*, 김민기, 장종현

(Jae-Kwan Yun, Min-Gi Kim, Jong-Hyun Jang)

Abstract : Recently, as the development of 3D contents service, production of 4D media that user can feel physical effects by the scene of movie has been increased. This 4D media uses the SMMD (Single Media Multiple Devices) method to give users enhanced effects with multiple devices playing rather than the SMSD (Single Media Single Device) method with single device playing in a smart TV environment. To give more sensible effects to users, from the beginning of making one media, we produced 4D media that consists of 3D contents and the SEM (Sensory Effect Metadata) to define effects of the scene. 4D media is transmitted to a HS (Home Server) to represent 3D contents in 3DTV and multiple devices with synchronized way. In this paper, we suggested the concept and overall 4D media broadcasting service system architecture, aggregator time acquisition, media time scheduling, real-sense device scheduling, evaluation results, user interfaces and implemented devices.

Keywords : 4D media, SEM, SMMD, SMSD, Multiple device synchronization.

1. 서론

최근 들어 아바타 3D를 선두로 하여 다양한 3D 콘텐츠가 제작되고 있다. 이에 따라 점차적으로 사용자의 오감을 자극할 수 있는 4D 미디어에 대한 요구가 증가하고 있다. 또한, 스마트 TV의 보급에 따라 극장, 체험관 등에서만 활용되던 4D 미디어를 가정에서 사용될 수 있는 방법에 대한 연구가 점차 진행되고 있다 [1].

이러한 4D 미디어를 처리하기 위해서는 하나의 미디어를 하나의 디바이스에서 처리하는 기존의 SMSD (Single Media Single Device) 방식을 벗어

나 하나의 미디어를 복수개의 디바이스를 이용하여 처리하는 SMMD (Single Media Multiple Devices) 기법을 사용하여야 한다 [2-4]. SMMD 기법을 사용하면 4D 방송 서비스와 같이 영상의 효과를 사용자 주변의 다양한 디바이스를 이용하여 실감 있게 제공하는 것이 가능하다. 이러한 SMMD 기법은 다양한 디바이스의 활용이 가능한 스마트 TV 환경에서 보다 효과적으로 사용될 수 있다 [5, 6].

본 논문에서는 스마트 TV 환경에서 SMMD 기법을 이용하여 사용자에게 실감 방송을 서비스 할 수 있는 4D 미디어를 위한 방송 서비스 시스템에 대하여 설명한다. 본 논문에서의 4D 방송 서비스 시스템은 야외나 실내에서 3D 카메라를 이용하여 영상을 촬영하면서 환경 정보를 수집할 수 있는 Aggregator를 이용하여 센서 정보도 같이 수집하여 영상을 만들고, 이를 가정 내의 홈서버로 전달하여 다양한 실감 효과 재현 디바이스를 동시에 재현함으로써 현장감 있는 영상을 재현할 수 있다. 이를 위하여 영상과 실감 효과를 동기화하여 인코

* 교신저자(Corresponding Author)

논문접수 : 2011. 07. 11., 수정일 : 2011. 08. 11.,
채택확정 : 2011. 09. 14.

윤재관, 김민기, 장종현 : 한국전자통신연구원

※ 본 연구는 지식경제부의 산업융합원천기술개발 사업의 일환으로 수행하였음. [10039171, 스마트 홈 환경에서 실감체험을 위한 융합형 미디어 서비스 플랫폼 기술 개발]"

당하고 이를 분석하며, 영상을 홈서버에 연결된 3DTV 및 실감 재현 디바이스를 이용하여 재현하는 방법에 대하여 설명한다. 본 논문은 1장의 서론에 이어 2장에서는 관련 연구로써 4D 미디어 서비스 시스템의 동향, 기존 방송 서비스 시스템과의 차이점에 대하여 설명한다. 그리고, 3장에서는 전체 방송 서비스 시스템의 설계, 4장에서는 시스템의 구현에 대하여 설명한 후 5장에서 결론을 맺는다.

II. 4D 미디어 방송 서비스 시스템

본 장에서는 4D 미디어 방송 서비스 시스템의 동향과 기존 방송 서비스 시스템과의 차이점에 대하여 설명한다.

1. 4D 미디어 서비스 시스템의 동향

4D 미디어는 주로 미국의 유니버설 스튜디오에서 3D 영상에 실감 효과 디바이스와 소품들을 이용하여 재현하는 형태로 서비스를 제공하였다. 국내에서는 한국전자통신연구원(ETRI)에서 2007년 10월 중국 심천에서 열린 82차 MPEG 미팅을 시작으로 네덜란드의 필립스, 일본의 샤프, 오스트리아의 클라겐푸르트 대학과 더불어 MPEG-V에 표준화를 추진하여 ISO/IEC 23005-1~6번까지 국제표준이 제정되었다 [7]. 또한, 명지대에서 MPEG-V Part 3를 기반으로 사용자가 실감효과 메타데이터를 편리하게 저작할 수 있도록 4D 저작툴을 개발하였으며, 광주과학기술원(GIST)에서는 미디어와 햅틱 디바이스를 연동하여 미디어의 장면에 따라 햅틱 장갑에 효과 정보를 전달하여 촉감 효과를 제공하였다.

국외에서는 필립사에서 게임을 하면서 바람,

그림 1. 필립스의 amBX 게임킷
Fig. 1 amBX gamekit of phillips

진동, 조명 효과 등을 게임의 장면에 따라 연동하여 재현시켜주는 게임킷으로 amBX를 개발하였으며, 클라겐푸르트 대학에서는 필립스가 개발한 amBX 게임킷을 기반으로 미디어에 실감효과를 삽입하고 amBX를 통하여 재현하는 시스템을 개발하였다. 그림 1은 필립스의 amBX 게임킷을 보여준다.

2. 기존 방송 서비스 시스템과의 비교

기존의 방송 서비스 시스템은 카메라를 이용하여 영상을 촬영한 후 이를 MPEG-2 TS 스트림으로 실시간 스트리밍을 하거나 촬영한 영상을 편집하여 방송 서버를 이용하여 전송하는 방식을 사용하였다. 이러한 기존 방송 서비스 시스템을 활용하여 4D 방송 서비스를 하기 위해서는 효과를 삽입하기 위한 저작툴이 필요하고, 이를 송신하기 위한 방송 서버가 필요하다. 그리고, 저작툴을 이용하여 삽입된 실감 효과는 실제 환경과 무관하게 저작자의 추측에 의해서 생성되기 때문에 실제 효과를 표현하기 어렵다. 이를 해결하기 위하여 본 논문에서는 3DTV에서 활용할 수 있도록 3D 카메라를 이용하여 영상을 촬영한 후 이를 NVS에서 인코딩한다. 그리고, 다양한 센서가 내장된 Aggregator를 이용하여 녹화 당시의 환경정보를 같이 수집한 후 방송 서버인 4D Caster를 이용하여 MPEG-2 TS 스트림으로 전송한다. 마지막으로, 홈서버에서는 이를 수신하여 3D 영상은 3DTV에서 재현하고 3D 영상과 동기를 맞추어 실감 디바이스를 재현한다. 그림 2는 이에 대한 시스템 개념도를 보여준다.

III. 시스템 설계

본 장에서는 4D 미디어 방송 서비스 시스템의 구조, Aggregator의 시간 획득 방법, 미디어 시간 스케줄링에 대하여 설명한다.

1. 시스템의 구조

본 논문에서 제안하는 4D 미디어 방송 서비스 시스템은 스마트 TV 환경에서 3DTV와 하나 이상의 실감 디바이스를 이용하여 4D 미디어를 재현할 수 있다. 이를 위해 NVS(Network Video Server)를 이용하여 3D 영상(좌,우)을 수집한다. NVS에서 수집된 3D 영상은 방송 서버의 입력으로 송출되고, 국제 표준인 MPEG-4 over MPEG-2 TS 포맷을 사용하여 인코딩된다. 카메라를 이용하여 영

그림 2. 4D 미디어 방송 서비스 시스템 개념도
Fig. 2 System concept of the 4D media broadcasting service system

상을 촬영하는 동안, Aggregator는 온도, 조도, 습도, 가속도, 각속도, 가스량, 풍향, 위치 정보 등의 환경 정보를 수집한다. 수집된 환경정보는 SEM(Sensory Effect Metadata)[8] 형태로 방송 서버에 지속적으로 전달되고, 방송 서버에서는 촬영된 영상과 동기를 맞추어 MPEG-4 over MPEG-2 TS로 인코딩한다. 일반적으로 오디오/비디오 인코딩은 고해상도일수록 시간이 걸리는데, Aggregator에서 수집된 실감 효과는 초당 데이터 양은 많을 수 있으나 압축을 하지 않기 때문에 오디오/비디오의 인코딩 시간에 맞게 실감 효과의 데이터 시간을 맞추어 송출하여야 한다. 이를 위해서 미디어 시간 동기화가 필요하다.

MPEG-4 over MPEG-2 TS 인코딩이 완료되면 방송 서버는 홈서버로 스트림을 브로드캐스팅한다. 홈서버는 MPEG-2 TS 디먹스를 이용하여 TS 헤더를 제거하고, 디코딩된 트랙은 릴레이 프로그램을 이용하여 홈서버에 연결된 오디오/비디오 재현 디바이스로 전송된다. 재현 디바이스 내부의 MPEG-4 플레이어는 홈서버에 있는 동기화 모듈에 맞게 스트림을 플레이한다. 디코딩된 SEM은 MPEG-2 DEMUX/Relay를 통해 SEM 디코더로 전송된다. SEM 디코더에서는 SEM을 해석한 후 홈서버에 연결되어 있는 실감 재현 디바이스에 맞게 효과 속성과 효과 값을 제어 시간 및 제어 값으로 변경한다. 동기화 모듈은 3D 영상을 플레이하면서 발생하는 재생 시간에 맞는 제어 값을 찾아 실감 재현 디바이스를 제어한다. 그림 3은 제안된 4D 미디어 방송 서비스 시스템의 구조도를 보여준다.

2. Aggregator 시간 획득 방법

본 논문에서 제안하는 4D 미디어 방송 서비스

시스템에서는 센서를 이용하여 영상과 동시에 실감 효과를 획득하는 카메라를 사용한다. 이 과정에서 영상에 맞게 실감 효과를 재현하기 위해서는 영상을 송출하는 NVS의 시간과 Aggregator의 RTC(Run-Time Clock) 시간을 동일하게 맞춰야 한다. 대부분의 경우 카메라와 Aggregator가 분리되어 촬영에 활용되는데 이런 경우에는 절대 시간으로 두 시간을 설정하기 어렵기 때문에 상대 시간을 이용하여 두 디바이스 간의 시간차를 계산하게 된다. 이를 위해 Aggregator에서 방송 서버로 실감 효과를 보낼 때는 동기화를 위한 시간을 계산할 수 있는 RTC 시간을 기록하게 된다.

우선, 방송 서버에서 NVS 및 Aggregator의 시작을 알기 위하여 Aggregator로부터 시작(START) 명령을 받은 후에 촬영된 영상과 Aggregator의 실감 효과를 인코딩하게 된다. 실제 Aggregator가 카메라에 장착된 경우에는 카메라의 녹화 버튼을 누름과 동시에 방송 서버로 두 데이터를 모두 송출하면 되지만 그렇지 않은 경우 Aggregator가 기준이 되어 시작 명령을 방송 서버로 전달한다. 또한, Aggregator로부터 종료(STOP) 명령을 받으면 방송 서버는 인코딩을 종료한다. Aggregator의 RTC는 SEM내의 start 엘리먼트에 기록된다.

그림 3. 4D 미디어 방송 서비스 시스템 구조도
Fig. 3 System architecture of the 4D media broadcasting service system

표 1에서 보는 것과 같이 type 엘리먼트는 실감 효과의 종류를 의미하고, value 엘리먼트는 센서로부터 수집된 값을 나타낸다. color 엘리먼트는 실제 재현 디바이스가 표현해야 할 색정보를 포함하고 timeScale은 밀리초 단위이며, start는 센싱된 시간을 나타낸다. Aggregator의 센싱 시간은 시스템의

설정에 따라 바뀔 수 있으며, 센싱 시간이 짧은 경우 데이터의 정확성이 보장되지만 전송되는 SEM

표 1. 조도 센서로부터 수집된 SEM

Table 1. SEM gathered from the illumination sensor

조도	<pre> <SensedInfoList> <SensedInfo type="LightSensorType" id="LS001" sensorIdRef="LSID001" activate="false" value="450.833344" color="#000000"> <TimeStamp type="ClockTickTimeType" timeScale="1000" pts="59" start="T09:26:07:884"/> </SensedInfo> </SensedInfoList> </pre>
	<pre> <SensedInfoList> <SensedInfo type="LightSensorType" id="LS001" sensorIdRef="LSID001" activate="true" value="33830.000000" color="#FFFFFF"> <TimeStamp type="ClockTickTimeType" timeScale="1000" pts="68" start="T09:20:05:884"/> </SensedInfo> </SensedInfoList> </pre>

데이터의 양이 많아지고, 센싱 시간이 긴 경우에는 전송되는 데이터의 양은 적어지지만 빠른 실감 효과 처리가 어렵다.

3. 미디어 시간 스케줄링

미디어 시간 스케줄링은 NVS에서 인코딩되는 3D 영상의 시간과 Aggregator로 먼저 전송되는 SEM의 start 엘리먼트의 시간을 비교하여 실제 송

출되는 TS 데이터의 인코딩 시간을 스케줄링한다. 즉, 촬영된 영상의 시간에 맞게 SEM을 MPEG-2 TS에 삽입한 후 브로드캐스팅하게 된다. 그림 4에서 보는 것과 같이 우선, MPEG-2 TS의 헤더 내에 있는 payload_unit_start_indicator 정보를 체크하여 1로 설정되어 있는 경우를 찾는다. 이 값이 1인 경우에는 PES 헤더가 존재하는 패킷이므로, MPEG-2 TS의 payload를 분석한다. 그림 5는 MPEG-2 TS 헤더의 구조를 보여준다.

그림 5. MPEG-2 TS 헤더 구조

Fig. 5 MPEG-2 TS header structure

본 논문에서는 MPEG-2 TS에 저장되어 있는 시간을 얻기 위하여 PTS (Presentation Time Stamp) 값을 얻은 후 이를 90KHz의 정밀도를 기준으로 마이크로 초로 변환한다. 변환된 마이크로 초 값과 Aggregator의 start 엘리먼트의 시간을 마

그림 4. 미디어 시간 동기화 방법

Fig. 4 media time synchronization method

이크로 초로 변환한 값을 비교하여 인코딩하는 순서를 결정한다. 현재는 3D 영상의 시간과 SEM의 시간을 20ms 이하로 보정을 하여 인코딩을 하는데 영상의 경우 딜레이를 줄 때 시청자가 거부감을 느낄 수 있으므로, 가능한 SEM의 인코딩 시간을 변경하여 영상과 시간을 맞춘다. 방송 서버에서 홈서버로 전송된 MPEG-2 TS 스트림은 디렉스에 의해서 분석된다. 촬영된 3D 영상은 MPEG-2 DEMUX/Relay 프로그램에 의해 홈서버에 연결된 SMP 8635 H/W 비디오/오디오 디코더(*Revision C, Production Chip MRUA library version:mrua_SMP_8634_2.8.2.0*)에서 플레이된다.

그림 4에서 MPEG-2 DEMUX/Relay는 방송 서버로부터 송출된 스트림에서 실감 효과에 대한 PID의 MPEG-2 TS를 분류하여 PES를 추출하고, PES에서 PID 0x103을 분석하여 SEM에 대한 텍스트를 추출해 낸다. 그리고, 릴레이에 설정된 IP에 따라 SEM을 처리할 수 있는 재현 모듈로 SEM을 송출하고 재현 모듈은 3D 영상의 시간에 맞게 실감 효과 디바이스를 제어한다.

IV. 시스템의 성능 평가 및 구현

본 장에서는 3D 영상 및 SEM 동기화 성능 평가를 수행하고, 사용자 인터페이스 및 구현된 디바이스에 대하여 설명한다.

표 2 사용된 비디오/오디오 인코딩 포맷

Table 2. Video/audio encoding format used for experiment

VIDEO		AUDIO	
Codec	MPEG-2 TS(CBR)	Codec	MPEG-Layer 2
Width,Height	1920x1080	Bit per sample	16bit/sample
Frame rate	29.97	Channels	2 channels
Average bitrate	17,000KBits/sec	Average bitrate	384KBits/sec

1. 영상 및 SEM 동기화 성능 평가

본 장에서는 3D 카메라로부터 입력된 3D 영상과 Aggregator를 이용하여 바람, 조도, 온도에 대한 센서 데이터를 획득한 후 본 논문에서 제시한 스케줄링 방법으로 동기화 시험을 하였다. 우선 NVS에서는 좌우(L, R) 화면 분할(side-by-side) 방식으로 전송되는 3D 영상을 표 2의 같은 포맷으

로 인코딩하였다. 3D 영상은 좌우 영상을 합쳐 FULL-HD 사이즈로 인코딩하고 오디오는 스테레오 방식으로 인코딩하였다.

Aggregator로부터 수집된 센서 데이터는 실제 디바이스가 제어될 수 있는 이벤트가 발생한 데이터에 대하여 3D 영상과 SEM 데이터의 오차를 각 샘플마다 20회 이상 수집하여 평균을 낸 값으로 비교하였다. 그림 6에서 보는 바와 같이 +값으로 같 경우에는 Aggregator의 시간이 스트림에 비해 늦게 도착한 경우이고, -값으로 같 경우에는 Aggregator가 먼저 도착한 경우이다. 그림 6은 보정 전에는 800ms까지 차이가 나는 오차를 20ms 이하로 줄인 결과를 보여준다.

그림 6. 성능 평가 결과

Fig. 6 Results of the performance evaluation

2. 사용자 인터페이스 및 구현된 디바이스들

그림 7은 4D 미디어 방송 서비스 시스템의 사용자 인터페이스 화면을 보여준다. 본 사용자 인터페이스는 5개의 주된 기능이 있다. 첫번째 기능은 2D 혹은 3D 카메라로 녹화된 오디오/비디오를 송출할 NVS의 주소를 설정하는 기능이다. 방송 서버는 한번에 4개의 NVS 송출 서버를 선택할 수 있으며, UDP/IP를 이용하여 홈서버로 MPEG-4 over MPEG-2 TS 스트림을 유니캐스팅 혹은 브로드캐스팅할 수 있다. 두 번째 기능은 Aggregator로부터 수집된 SEM의 로그를 보여주는 기능이다. Aggregator로부터 A-Start 명령이 전송되면 방송 서버는 카메라 영상과 SEM을 인코딩하여 홈서버로 전송하고 A-Stop 명령이 전송되면 방송을 중단한다. 세 번째 기능은 실시간으로 실감 정보를 전송하는 기능이다. 환경 정보는 카메라 영상을 촬영하면서 자동으로 삽입되지만 만약 방송을 송출할 경우 실시간으로 SEM을 전송하고 싶을 경우 본 기능

을 사용하게 된다. SEM은 환경 정보 XML과 달리 영상과 시간을 맞추어 인코딩 되지 않고 버튼이 눌리는 즉시 실시간으로 SEM이 전송된다. 네 번째 기능은 SEM 전송 로그이다. 본 기능은 실시간으로 SEM이 전송되거나 환경 정보에 의한 SEM이 전송될 경우 로그가 표시된다. 다섯 번째 기능은 현재 MPEG-4 over MPEG-2 TS 스트림이 전송될 경우 차지하는 네트워크 밴드width를 실시간으로 보여준다. 그림 8은 본 논문에서 구현 및 사용된 Aggregator를 보여준다.

그림 7. 4D 방송 서비스 시스템의 사용자 인터페이스

Fig. 7 User interfaces for the 4D media broadcasting service

그림 8 구현에 사용된 Aggregator

Fig. 8 Aggregator for implementation

V. 결 론

최근 들어, 실감 방송, 실감 뉴스, 실감 교육 등과 같이 사용자에게 실감 효과를 줄 수 있는 방송 매체에 대한 요구가 증가하고 있다. 본 논문에서는 스마트 TV를 사용하는 스마트 홈 환경에서 극장에서와 같은 4D와 같은 효과를 재현할 수 있는 4D 미디어 방송 서비스 시스템에 대하여 설계 및 구현하였다. 본 논문에서는 하나 이상의 카메라를 이용하여 촬영한 영상과 센서를 통해 수집된 환경 정보

를 하나의 MPEG-4 over MPEG-2 TS 스트림으로 인코딩하고 이를 홈서버로 전송한 후 스마트 TV, 노트북, 스마트 폰과 같은 복수개의 플레이어와 복수개의 실감 재현 디바이스를 이용하여 재현하는 방법에 대하여 설명하였다. 향후에는 n-Screen 적용을 위한 비디오 압축 기법 연구가 필요하다.

참고문헌

- [1] J.S. Han, J.K. Yun, J.H. Jang, K.R. Park, "User-Friendly Home Automation based on 3D Virtual World," IEEE Transactions on Consumer Electronics, Vol. 56, No. 3, pp. 1843-1847, 2010.
- [2] S.W. Park, N.S. Park, J.T. Kim, E.H. Paik, "Provision of the Expressive Multisensory Adaptation Platform for Heterogeneous Multimedia Devices in the Ubiquitous Home," IEEE Transactions on Consumer Electronics, Vol. 55, No. 1, pp.126-131, 2009.
- [3] F.B. Segui, J.C.G. Cebollada, J.L. Mauri, "Multiple Group and inter-stream synchronization techniques: A comparative study," Information Systems, Vol 34, No 1, pp.108-131, 2009.
- [4] J.K. Yun, J.H. Jang, K.R. Park, D.W. Han, "Real-Sense Representation Technology Using Multiple Devices Synchronization," Proceedings on 7th IFIPWG International Workshop, pp.343-353, 2009.
- [5] K.R. Yoon, B.S. Choi, E.S. Lee, T.B. Lim, "4-D broadcasting with MPEG-V," Proceedings on Multimedia Signal Processing (MMSP), pp.257-262, 2010.
- [6] J.K. Yun, J.H. Jang, K.R. Park, "Real-sense media synchronization technology based on the SMMD," Proceedings on Digest of Technical Papers International Conference on Consumer Electronics (ICCE), pp.71-72, 2010.
- [7] C. Timmerer, S. Hasegawa, (eds.) "Working Draft of ISO/IEC 23005 Sensory Information," SO/IEC JTC 1/SC 29/WG 11/N10475, Lausanne, Switzerland, 2009.
- [8] M. Walzl, C. Timmerer, H. Hellwagner, "A Test-Bed for Quality of Multimedia Experience Evaluation of Sensory Effects," Proceedings on the

First International Workshop on Quality of
Multimedia Experience (QoMEX 2009), pp.145-150,
2009.

저 자 소 개

윤재관

1997년 건국대 전자계산학과 학사. 1999년 건국대학교 컴퓨터공학과 석사. 2003년 건국대학교 컴퓨터공학과 박사. 현재, 한국전자통신연구원 그린컴퓨팅연구부 실감플랫폼연구팀 선임연구원.

관심분야: Real-Time OS, 임베디드 S/W, 실시간 데이터베이스 시스템, 실감 방송 서비스, 실감 미디어 서비스, 실감형 인터랙션 서비스.

Email: jkyun@etri.re.kr

장종현

2000년 충남대 석사. 2004년 Hufs 전자계산학과 박사. 현재, 한국전자통신연구원 그린컴퓨팅연구부 실감플랫폼연구팀 책임연구원(팀장).

관심분야: 실시간 미들웨어, 홈네트워킹 시스템, 실감 미디어 서비스 시스템.

Email: jangjh@etri.re.kr

김민기

2009년 목원대 컴퓨터공학과 학사. 현재, 한국전자통신연구원 그린컴퓨팅연구부 감성융합미들웨어연구팀 연구원.

관심분야: 홈네트워킹 시스템, 상호연동 미들웨어, 실감 미디어 서비스 시스템.

Email: minki1982@etri.re.kr