다시점 비디오를 위한 시스템 디코더 모델 설계와 구현

이준철[O*] 최기한** 이정원** 김승호**
[O*]부천대학 인터넷과, *경북대학교 컴퓨터공학과
jclee[O*]@bc.ac.kr, {khchoi, jwlee}@mmlab.knu.ac.kr, shkim@knu.ac.kr

Design and implementation of System Decoder Model for Multi-view Video

J.C Lee[O*] K.H Choi** J.W Lee** S.H Kim**
*Dept. of Internet Information Science, BuCheon College
** Dept. of Computer Engineering, Kyungpook National University

요 약

3차원 비디오는 차세대 정보통신 서비스 분야의 핵심 기술로써 사용자에게 양질의 서비스를 제공하기 위한 미래 기술이며, 여러 개의 시점으로 구성되는 다시점 비디오는 고화질의 개념을 뛰어 넘어 더욱 사실감 넘치는 영상을 제공하며 사용자에게 미디어에 대한 실제감과 몰입감을 느끼게 한다. 본 논문에서는 입력 영상인 MP4파일 구조를 분석하고, MPEG을 기반으로 하는 다시점 비디오를 구현하는 방법을 제안한다.

1. 서 론

인간의 눈은 가로 방향으로 약 60mm 떨어져서 존재하는데, 이로 인해 나타나게 되는 양안시차(binocular disparity)는 입체감의 가장 중요한 요인이라 할 수 있다. 즉, 좌우의 눈은 각각 서로 다른 2차원 화상을 보게 되고, 이 두 영상이 망막을 통해 뇌로 전달되며, 뇌는 이를 정확히 서로 융합하여 본래 3차원 입체 영상의 깊이감과 실제감을 재생하는 것이다.

인간이 실제 눈으로 얻는 정보는 입체영상이므로 일상적으로 보고 있는 자연계의 정경에 보다 가깝고 자연스러운 영상시스템을 추구하는 인간의 욕망은 보다 많은 정보를 요구하고 초고속 정보통신망을 근간으로 구축될 정보고속화도로(Information Super Highway) 위에서 실현될 서비스들은 현재의 전화와 같이 단순히 " 듣고 말하는" 서비스로부터 문자, 음성, 영상을 고속 처리 하는 디지털 단말을 중심으로 한 " 보고 듣는" 멀티미디어형 서비스로 발전할 것이다. 궁극적으로는 " 시공간을 초월하여 실감 있고, 입체적으로 보고 느끼고 즐기는" 초 공간형 실감 3차원 비디오 서비스로 발전 할 것으로 예상된다.

3차원 비디오 기술은 고속 대용량의 정보통신 인프라 구축으로 21세기 전반에 걸쳐 새로운 영상 공간을 제공하는 서비스로 정보통신, 방송, 의료, 군사, 게임, 가상현실 등 분야에 이용된다. 이런 분야에서 요구되는 차세대 실감 3차원 입체 멀티미디어 정보통신의 핵심 기반 기술로, 방송기술의 발전 측면에서 칼라TV보다 선명한 화질을 제공하는 HDTV로 발전하고 있다. 미래에는 입체감과 현장감을 사용자에게 제공할 수 있는 3DTV가 서비스 될 것이며, 3차원 렌더링 하드웨어와 디스플레이 기술의 발전으로 2010년에 등장할 것이다[2].

본 논문에서는 입체감과 사실감 넘치는 영상을 제공하고, 이를 통해서 사용자에게 현장감과 몰입감을 느끼게 할 수 있는 다시점 비디오 서비스를 MPEG-4 시스템을 기반으로 하여 구현한다. 논문의 구성으로 2장에서는 MPEG-4 시스템과 MP4파일 구조를 분석하고, 3장에서는 다시점 비디오 서비스를 위한 SDM(System Decoder Model)구조를 제안하고 이를 배경으로 구현한다. 4장에서는 구현한 결과를 설명하고, 5장에서 결론을 내린다.

2. 관련 연구

다시점 3차원 TV는 최종 목표인 오감다차원 TV를 구현하기 위한 중간 과정에 있으며, 이를 위해 디스플레이, 촬영, 전송, 신호처리 등 다양한 기술이 요구된다. 다시점 3차원 동영상의 압축 및 전송 표준인 MPEG-4가 가장 효과적인 방법으로 고려 되고, 현재 MPEG-4에서 다시점 비디오 기술은 새로운 기술이라고 할 수 있다. 그리고, 이 기술은 3차원 AV구현의 기반 기술 중 하나가 될 것이다. 본 장에서는 ISO/IEC 14496의 Part 1인 MPEG-4 시스템과 다시점 비디오를 구현하기 위해 실험영상이 필요하며,

로서, 이에 대한 연구가 진행 되고 있다[1]. 이러한 일례

2.1 MPEG-4 시스템

MPEG-4는 내용 기반의 대화형 기능과 객체 단위의 데이터 처리 기능을 가지고 있다. 따라서, 영상내의 의미 있는 내용물들을 각각 다른 객체로 부호화하고, 이를 혼합하여 하나의 화면을 구성, 추후 원하는 객체만을 추출하여 재사용하거나, 객체들을 이용하여 전체 데이터를 재구성할 수 있다.

MPEG-4 시스템 표준에 규정된 시청각 표현 정보는 인코딩된 비주얼/오디오 정보와 해당하는 화면구성 정보에 상호 연관된 비주얼/오디오 화면을 생성하는 방법이다. 이런 기본동작이 그림 1을 통해 설명 한다.

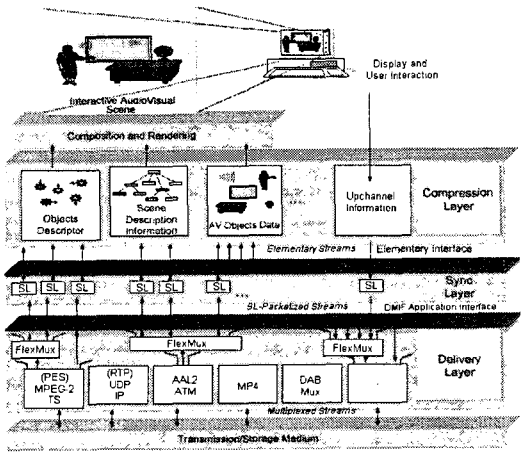


그림 1. MPEG-4 시스템 터미널 구조

Transmission/Strorage에서 데이터를 읽어 들이면 데이터에 맞는 프로토콜로 선택된다. 이는, MPEG-4에서 데이터에 맞는 정의된 프로토콜이 없기 때문에 데이터에 맞는 프로토콜을 선택하게 된다. 선택된 프로토콜은 FlexMux를 통해서 동기계층으로 전달된다. 전송계층과 동기계층 사이에 ISO/IEC 14496-6에서 제안된 DMIF 어플리케이션 인터페이스를 통해서 SL-Packetized 스트림들이 동기계층으로 전달된다. 동기계층을 지남으로써 ES(Elementary Streams)라 한다. 압축계층에서는 ES들이 OD(Object Description)인지 Scene을 구성하는 BIFS인지 혹은, AV Object데이터들인지에 따라 부호화된 포맷으로 데이터를 복원한다. 복원된 정보는 디바이스의 컴포지션, 랜더링을 위해 이용된다. 압축계층에서 복원된 정보를 이용하여 각각의 OD정보, 화면정보들을 컴포지션하고 랜더링과정을 통해서 장치에 재생되는 과정을 거치게 된다. 재생된 후 사용자 상호작용을 받게 되면 업 채널 정보를 압축계층으로 보내고 다시 동기계층, 전송계층을 통해 MP4파일 구조를 분석한다.

2.2 MP4 파일 구조

대화형 방송환경에서 MPEG-4 컨텐츠는 MP4파일 형식으로 저장된다. MP4파일은 IOD, OD, BIFS, 음성/영상 데이터를 ISO/IEC 13818-1 Amd.7 의 표준에 맞게 구성이되어 MPEG-2 TS 스트림으로 전송된다. 그림 2는 MP4 파일 구조를 나타낸 것이다.

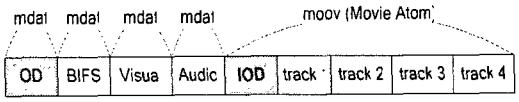


그림 2. MP4파일 구조

미디어 데이터를 AU (Access Unit) 단위로 가지는 미디어 데이터를 mdat(movie data atom)와 각 미디어 데이터의 메타정보를 가지는 moov (Movie Atom)로 구성된다. IOD는 track 1 과 track 2가 디코딩 되며, track 1은 BIFS(Binary Format for Scene)를 track 2는 OD를 참조하게 된다. BIFS는 장면을 생성하게 되고, OD를 통해서 참조하는 Audio 스트림 정보와 Visual 스트림 정보를 가져와서 장면에 뿌려 지게 된다.

3. SDM (System Decoder Model)

다시점 비디오에서 전송되는 각 시점 별 영상은 크게 두 가지 경우로 생각 할 수 있다. 그림 3과 같이 모든 시점들의 영상정보가 MVC(Multi View Codec)을 통해 하나의 ES에 포함되어 전송되는 경우와 그림 4와 같이 각 시점 별로 MVC를 통해 각각의 ES로 전송되는 두 가지의 경우로 생각 할 수 있다. 그림 3은 각 시점의 영상정보가 하나의 ES를 통해 전송되는 경우를 나타낸 그림이다.

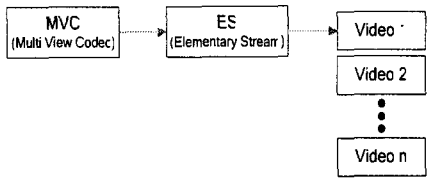


그림 3. 하나의 ES를 통해 전송되는 경우

그림 3은 MVC를 통해 여러 개의 시점의 영상정보가 하나의 ES로 통합되어 출력되는 경우이다. 각 시점간 영상의 중복성을 제거 하기 위해 기존의 3DAV 코덱의 Joint, Compatable 방식[3]으로 압축하여 하나의 ES를 생성하여 전송하는 방식이다. 전송되는 ES를 MVC에서 재구성하게 되므로, 기존의 MPEG-4의 신텍스 구조를 따르지

으로 보냄으로써 처리된다.
조는 사용자 상호작용에 따라 사용자가 원하는 시점을 선택해서 볼 수 없을 뿐 아니라 각 시점간의 상관관계에 대한 정보가 추가 되어야 한다.

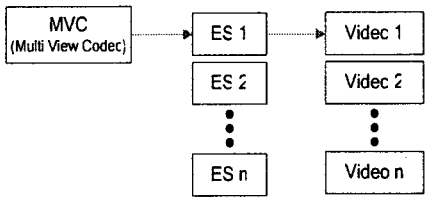


그림 4. 각각의 ES를 통해 전송되는 경우

그림 4와 같이 MVC에서 각 시점 별로 각각의 ES가 출력되는 경우, 각 영상 별로 MVC인코더를 사용하여 압축된 ES가 MUX를 통해 동기화되어 전송된다. 각 ES는 기존의 시스템과 호환성이 존재하고 각각의 ES로 전송되는 특징을 가진다. 그림 5는 각 시점의 영상정보가 MVC로 인코딩되어 각각의 ES로 전송될 경우를 나타낸 것 이다.

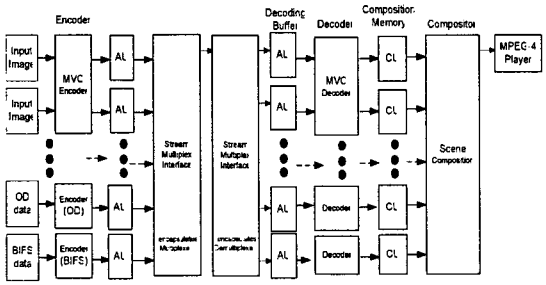
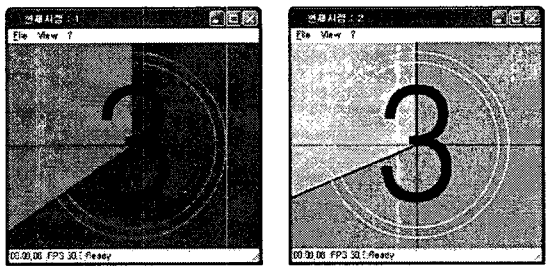


그림 5. SDM 설계 구조

수신측에서 DeMUX를 통해 각각의 디코딩 버퍼로 들어가서 MVC 디코더에서 디코딩되며, 컴포지션 메모리를 거쳐 컴포지터가 Scene과 OD를 조합하고 MPEG-4 플레이어를 통해 재생된다. 사용자 상호작용을 받으면 DeMUX부터 다시 위와 같은 과정을 거치게 된다. 아직, 다시점 비디오를 지원하기 위한 인코더 부분이 표준이 정의 되지 않았기 때문에, 그림 5의 인코더 부분은 MVC인코더를 거치면 각각의 AU가 나온다는 것을 가정하고 SDM을 설계한 것이다.

4. 실험 결과

다시점 비디오를 재생하기 위해서 n개의 시점을 고려하였고, 본 논문에서는 두 개의 시점만을 고려 하여 실험하고, 코덱에서 새로운 신텍스 구조가 구성된다. 이런 구하였다. 실험을 위해 GPAC0.4.0버전을 이용하였고, 다른 영상을 입력 받은 2대의 카메라를 이용하였으며, MP4BOX 툴로 실험영상을 생성하였다. 그림 6은 본 논문에서 제안한 SDM구조를 고려하여 재생한 결과화면이다.



(a) 시점 1 (b) 시점 2

그림 6. 실험 결과

사용자 상호작용은 키보드의 입력으로서 인식을 하며, 사용자가 원하는 시점을 선택하여 재생 할 수 있다. 현재는 두 개의 시점만을 고려 하였지만, 이를 바탕으로 더 많은 시점을 고려 할 수 있을 것이다.

5. 결론

사용자 상호작용에 대해 원하는 시점을 선택적으로 재생 할 수 있는 것을 실험 결과를 통해 확인 하였다. 미래에 사용자에게 입체감과 생동감을 느끼게 하기 위해서 다시점을 기반으로 한 인터렉티브 기능이 추가 될 것으로 기대 된다. 본 논문에서는 두 개의 시점까지 고려하였고, 인터렉티브 기능을 구현하였다. 차후에, 각각의 카메라에서 입력되는 영상의 상호관계와 n개의 시점을 고려하였을 때 시스템의 성능에 따라 사용자에게 자연스럽게 재생될 수 있는 방법이 연구 되어야 한다.

참고문헌

1. M. Price, J. M. Price, J. Chandaria, O. Grau, G.A. Thomas, D. Chatting, J. Thorne, G. Milnthorpe, P. Woodward, L. Bull, E-J. Ong, A. Hilton, J. Mitchelson, J. Starck, " Real-Time Production and Delivery of 3D Meida," BBC R&D White Paper, 2002.
2. Advanced Three-Dimensional Television System Technologies, http://www.iti.gr/db.php/en/projects/ATTEST.html
3. B. Haskell, A. Purl, A. N. Netravali, Digital Video: An Introduction to MPEG-2, New York: Chapman & Hall, 2003