



특집

스테레오스코픽 동영상 응용 파일 포맷 기술: Stereoscopic video application format

김규현 · 박광훈 · 서덕영 (경희대학교)

I. 서론

디지털 기술의 발전은 다양한 미디어 화질에서 커다란 변화를 가져 왔다. 특히, 방송에서는 기존의 SD급의 화질에서 HD급의 고화질 영상을 제공하게 되었으며, 이러한 디지털 기술의 발전은 단순히 미디어의 화질 뿐만 아니라 다양한 부가데이터 서비스를 제공하는 대화형 서비스를 가능하게 만들었다.

또한, 기존의 광대역폭을 중심으로 한 고화질의 미디어를 제공하는 디지털 기술은 더 나아가, 저대역폭의 통신망을 통해 다양한 방송 및 미디어 콘텐츠를 실시간으로 제공하는 서비스 기술로 발전되어 세계 최초로 국내에서는 DMB 서비스를 제공하게 되었다. 그러나, 이러한 고화질의 콘텐츠는 기존의 콘텐츠에 비해 사용자에게 보다 실감적인 서비스를 제공할 수는 있으나, 기본적으로 사람이 보는 사물이 갖고 있는 3차원 공간적인 특징을 제공하는 데에는 그 한계가 존재하였다.

이러한 2차원 중심의 콘텐츠에서 보다 실감적인 콘텐츠를 제공하기 위한 다양한 3차원 영상 처리 기술이 속속히 개발되어지고 있다. 이러한

3차원 실감 영상 콘텐츠중에서 가장 많이 사용되고 있는 콘텐츠는 2개의 카메라를 통해 얻어지는 스테레오스코픽 콘텐츠이다. 스테레오스코픽 콘텐츠는 기존의 2D를 기반으로 한 모노스코픽 영상에 비하여 깊이 정보를 제공함으로써 보다 다양한 실감 효과를 사용자에게 제공할 수 있으며, 기본적으로 스테레오스코픽 콘텐츠는 편광 필터를 사용한 안경이나 혹은 디스플레이어에 필터를 제공하는 무안경식 LCD 스크린을 통해 다양하게 활용되어지고 있다. 특히, 이러한 무안경식 스테레오스코픽 콘텐츠의 경우, 대화면의 스크린 보다는 작은 화면에서 보다 효과적으로 사용자에게 불편함없이 3차원 실감 효과를 제공할 수 있기에 작은 화면의 스크린을 장착한 다양한 디지털 멀티미디어 기기에서 많이 활용되어지고 있으며, 이러한 기기들은 2D/3D 기능을 동시에 제공하고 있어 모노스코픽 및 스테레오스코픽 콘텐츠를 사용자는 자유롭게 즐길 수 있게 되었다. 이미 손안의 멀티미디어 기기의 기능을 제공하고 있는 휴대폰과 디지털 카메라에서도 이러한 스테레오스코픽 콘텐츠를 촬영하고 디스플레이 할 수 있는 제품들이 시장에 소개되어 지고 있다. 그러나, 이러한 멀티미디어 기

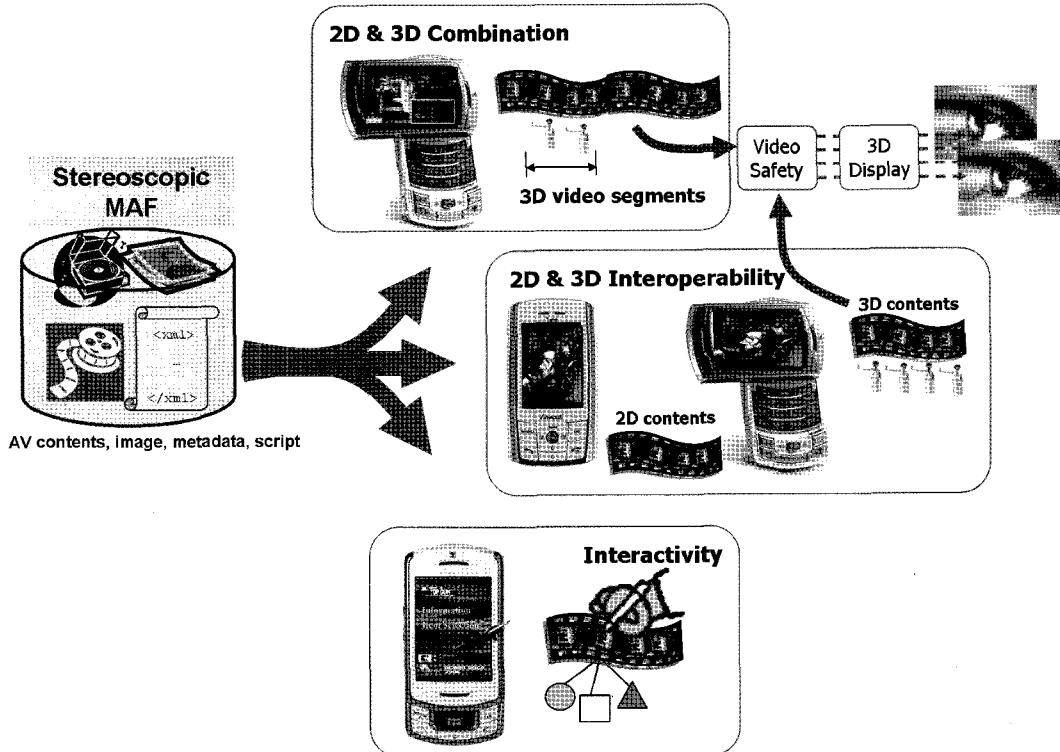
기를 통해 획득한 스테레오스코픽 콘텐츠는 세계적으로 공통으로 사용되는 저장 파일 포맷의 미비로 개인 멀티미디어 기기에서 획득한 스테레오스코픽 영상을 다른 기기에서는 활용될 수 없으며, 스테레오스코픽 동영상의 구조에 기반한 저장 파일 포맷이 아니기에 저장 효율에서도 그 한계를 갖고 있다. 따라서, 이러한 공통 파일 포맷 미비의 한계는 스테레오스코픽 콘텐츠의 확산을 가져오는데에 어려움으로 작용하고 있다^[14].

멀티미디어 콘텐츠의 압축 및 데이터 서비스 관련 기술의 국제표준을 제정하고 있는 MPEG에서는 이러한 파일 포맷의 필요성을 파악하여 지난 2007년 10월부터 스테레오스코픽 콘텐츠의 저장 파일 포맷에 대한 표준화 작

업을 시작하여, 2008년 10월에 최종국제표준안 (FDIS: Final Draft of International Standard)를 승인하였다. 본 고에서는 해당 표준이 대상으로 하는 어플리케이션 시나리오와 이를 위해 표준에서 제공하고자 하는 기능을 소개하고자 한다.

II. 스테레오스코픽 동영상 응용 파일 포맷을 통한 어플리케이션 시나리오

스테레오스코픽 카메라를 장착한 휴대폰과 같은 개인 휴대 단말은 현재 스테레오스코픽 정지 영상 및 동영상을 지원하고 있다. 이와같이 개인 휴대 단말을 통해 획득한 스테레오스코픽 영상



〈그림 1〉 스테레오스코픽 동영상 응용 파일 포맷 기능도

은 단순한 스테레오스코픽 영상 뿐만 아니라, 개인 휴대 단말에서 제공하는 다양한 기능들과 연계하여 새로운 서비스를 제공할 수 있도록 스테레오스코픽 파일 포맷을 정의하여야 한다. 즉, 기존의 2D 콘텐츠에서 제공하는 다양한 기능들이 스테레오스코픽 영상과도 함께 연계되어 제공될 수 있는 기능들을 하나의 파일 포맷에서 지원할 수 있어야 한다. 일례를 들어, <그림 1>에서 보는 바와 같이, 스테레오스코픽 영상은 기존의 2D 영상에 비해 사용자의 시각적인 피로도가 증가될 수 있기에, 기존의 2D 콘텐츠와 함께 구성되어, 사용자에게 시각적으로 강조를 하기 위한 용도를 활용되어 질 수 있으며, 또한, 스테레오스코픽 동상으로 구성된 콘텐츠의 경우, 스테레오스코픽 동영상을 재생하기 위한 단말이 아직은 널리 활용되어지고 있지 않고 있기에, 기존의 2D 단말에서 해당 콘텐츠를 재생하기 위하여 스테레오스코픽 영상을 2D영상으로 변환하여 재생할 수 있도록 파일 포맷에서는 그 기능을 정의하고 있어야 한다. 또한, 기존 2D 콘텐츠에서 활용하고 있는 다양한 대화형 기능을 통한 부가 서비스를 스테레오스코픽 영상에서도 지원하기 위한 기능도 스테레오스코픽 영상 파일 포맷에서는 함께 지원할 수 있도록 정의하여 한다 [16].

III. 스테레오스코픽 동영상 응용 파일 포맷 요소 기능

II장에서 소개되어진 스테레오스코픽 동영상 응용 포맷을 통한 다양한 어플리케이션을 제공하기 위해 파일 포맷에서 정의하고 있는 주요 기능은 아래와 같다[1].

1. 모노스코픽/스테레오스코픽 혼합 콘텐츠 서비스 지원

제공되어지는 콘텐츠는 스테레오스코픽 동영상만으로 구성되어질 수도 있으나, 또한, 제공하고자 하는 콘텐츠의 효과를 극대화하기 위해 기존의 2D 기반의 모노스코픽 콘텐츠에 일부만 을 스테레오스코픽 동영상으로 제공할 수 있다. 일례로, 광고 콘텐츠의 경우, 사용자의 시각적 피로도를 고려하여 2D 콘텐츠로 주요 광고 시나리오를 구성하고, 해당 제품에 대한 시각적 효과를 통한 강조를 위해 특정 구간을 스테레오스코픽 콘텐츠로 구성할 수 있다. 따라서, 제공하는 스테레오스코픽 동영상 응용 포맷은 이러한 혼합 콘텐츠를 효과적으로 저장할 수 있어야 한다.

2. 모노스코픽/스테레오스코픽 콘텐츠간의 호환성 확보

휴대 단말기의 경우, 다양한 네트워크 기능을 제공하고 있으며, 이러한 네트워크 기능을 통해 사용자가 획득 또는 가공한 콘텐츠를 다른 사용자에게 전달 또는 공유할 수 있다. 이러한 네트워크 기능을 통해 획득 또는 가공한 스테레오스코픽 동영상을 다른 사용자에게 전달하였으나, 현재 시장에서는 스테레오스코픽 콘텐츠를 디스플레이할 수 있는 기능을 장착한 멀티미디어 기기 보다는 모노스코픽 디스플레이 기능을 갖고 있는 기기가 상대적으로 많이 존재하고 있기에, 전달받은 스테레오스코픽 동영상을 재생할 수 없는 현상이 발생할 수 있다. 이에, 표준에서 제정하고자 하는 스테레오스코픽 동영상 응용 포맷에서는 저장된 스테레오스코픽 동영상을 모노스코픽 디스플레이만을 장착한 멀티미디어 기기에

서는 모노스코픽으로 제공할 수 있는 기능을 제공하여 한다.

3. 스테레오스코픽 콘텐츠 기반 대화형 서비스

점차로 멀티미디어 서비스는 기존의 단방향성에서 벗어나 사용자의 요구사항을 전달하고 이에 따라 다양한 서비스를 제공하는 대화형 기능을 요구하고 있기에, 본 논문에서 소개하고 있는 스테레오스코픽 동영상 응용 포맷에서는 이러한 대화형 기능을 제공할 수 있어야 한다.

상기에서 소개되어진 요구사항을 만족하는 스테레오스코픽 동영상 응용 포맷을 정의함에 있어 MPEG에서는 <표 1>에서 보여진 것과 같은 기존의 제정된 표준 기술을 사용하고자 하고 있다. 스테레오스코픽 동영상 응용 포맷은 기본적으로 ISO (International Standard Organisation)에서 정의한 기본 포맷인 ISO Base Media File Format과 호환성을 확보하고, 스테레오스코픽 동영상 및 오디오 압축/복원 기술로는 MPEG-4 part2 (SP@level 3), MPEG-4 part10 (BP@level 1.3) 및 AAC-LC, AAC+를 활용하였으며, 휴대폰에서 활용되어질 것을 고려하여 음성 압축/복원 기술로는 AMR, EVRC를 활용하였다. 또한, 다양

<표 1> 스테레오스코픽 비디오 응용 포맷 지원기술 및 사양

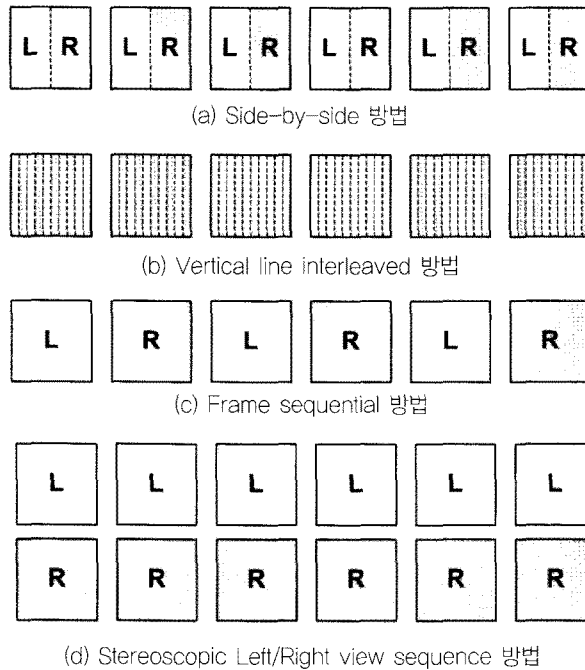
분류	규격 및 사양
파일포맷	ISO Base Media File Format
비디오	MPEG-4 Part2 (SP@level 3), MPEG-4 Part10 (BP@level 1.3)
오디오	AAC-LC, AAC+
보이스	AMR, EVRC
장면기술	MPEG-4 Part20 LASer
정지영상	JPEG, PNG

한 대화형 서비스 기술을 제공하기 위하여 LASer 기술을, 그리고, 다양한 부가데이터를 위하여 JPEG과 PNG 기술을 지원토록 구성하고 있다^{[13][5]}.

IV. 스테레오스코픽 동영상 구성 방법

스테레오스코픽 동영상은 기본적으로 좌우의 영상을 통해 입체감을 사용자에게 전달하기에, 좌우의 동영상을 하나의 스트림으로 구성하는 방법이 정의되어야 한다. 이러한 동영상 구성방법으로 'side-by-side', 'above and below', 'pixel-by-pixel' 타입과 여러 종류의 interleaved, sequential 타입이 존재한다. 단일 표준에서 모든 종류의 컴포지션 타입을 정의하기는 한계가 존재하기에 스테레오스코픽 동영상 응용 포맷에서는 모바일 단말을 우선으로 널리 사용되는 타입들을 고려하여 <그림 2>에서 나타난 바와 같이 'Side-by-side', 'Vertical line interleaved', 'Frame sequential', 'Stereoscopic Left/Right view sequence' 타입의 네 가지 구성방법을 정의하였다.

Side-by-side 방식은 가장 널리 사용되는 스테레오스코픽 구성방법으로 <그림 2> (a)에서 보이는 바와 같이 하나의 프레임을 좌우 반으로 나누어 좌영상과 우영상을 옆으로 위치시키는 방법으로 영상의 해상도가 반으로 주는 한계점이 존재하기는 하지만 기존의 비트율 그대로 영상을 압축할 수 있고, 기존 시스템의 변환 없이 플레이어에서 재생될 수 있다는 이점이 있다. Vertical line interleaved 구성방법은 <그림 2> (b)와 같이 하나의 프레임에 수직방향으로 좌영상과 우영상을 하나의 픽셀 라인씩 번갈아 위치시키는 방식으로서, 모든 수직라인마다 불연속성이 존재하므



〈그림 2〉 스테레오스코픽 동영상 구성 방법

로 압축 효율이 낮으나, 모바일 디스플레이 장치에서 가장 많이 사용되는 패럴랙스 배리어 디스플레이 방식에 변환 없이 직접 적용될 수 있다는 이점이 있다. 또한, Frame sequential 구성방법은 <그림 2> (c)와 같이 좌영상 프레임과 우영상 프레임을 번갈아 하나의 스트림을 구성방법으로서, 모노스코픽 콘텐츠의 풀(full) 해상도의 스테레오스코픽 콘텐츠 재생이 가능하다. Stereoscopic Left/Right view sequence 구성방법은 <그림 2> (d)와 같이 독립된 두 개의 스트림이 직접 모노스코픽 콘텐츠로 활용될 수 있는 이점이 있으나, 스테레오스코픽 콘텐츠 재생 시 각각의 프레임 쌍들에 대한 동기화가 추가로 필요하다^[56].

V. 스테레오스코픽 동영상 파일 포맷 기본 구조

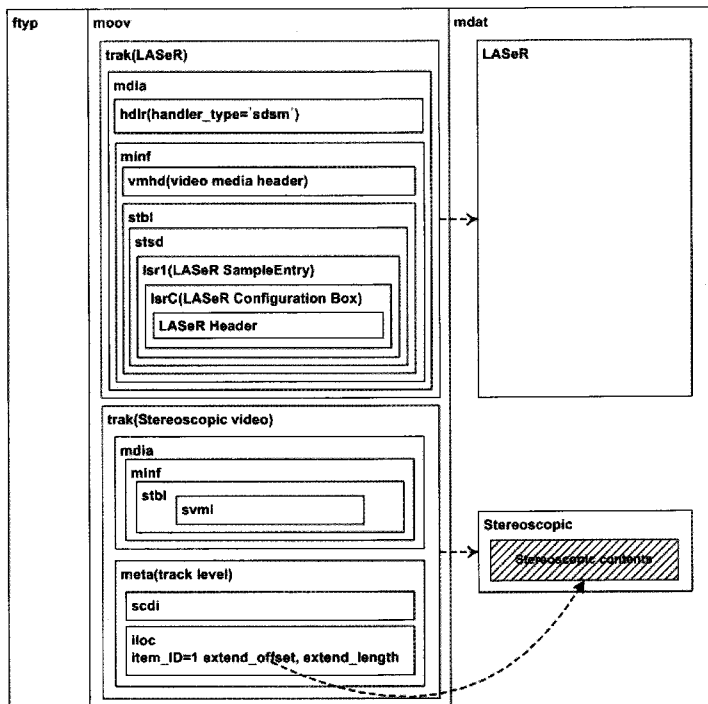
스테레오스코픽 동영상 응용 포맷의 파일 구조는 ISO Base Media File Format(ISO/IEC 14496-12)을 기반으로 설계되었으며 스테레오스코픽 비디오 미디어 정보 기술을 위한 스테레오스코픽 비디오 미디어 정보 박스(svmi), 카메라와 디스플레이 정보 기술을 위한 스테레오스코픽 카메라, 디스플레이 정보 박스(scdi)가 새롭게 추가되었다.

ISO Base Media File Format은 재생 관리를 위한 정보들이 박스(box) 형태의 구조체로 구성되어 있으며, 일반적으로 파일 타입을 정의하는 파일 타입 박스(ftyp), 비디오 등의 미디어 재생 관리를 위한 정보를 기술하는 무비 박스

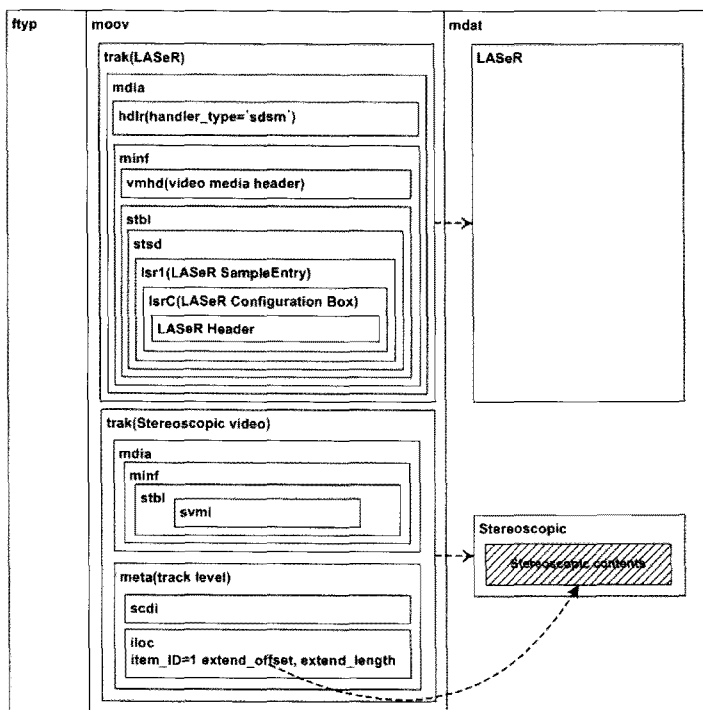
(moov), 실제 데이터를 포함하는 미디어 데이터 박스 (mdat)로 구성되어 있으며, 특히, 미디어 재생에 관한 세부적인 정보들은 무비 박스 안의 트랙 박스 (trak)에 기술된다^[3].

<그림 3>에서는 상기에서 소개한 ISO Base Media File Format을 기반으로 IV장 <그림 2>에서 정의한 스테레오스코픽 동영상중, 하나의 시퀀스로 스테레오스코픽 동영상을 - Side-by-side, Vertical line interleaved, Frame sequential 구성 방법- 대상으로 스테레오스코픽 동영상 파일 포맷을 정의한 것을 소개하고 있다. 또한, <그림 4>에서는 두개의 시퀀스로 스테레오스코픽 동영상을 구성하는 Left/Right view sequence 구성방법의 경우에 대한 파일 포맷 구조를 소개하고 있다. 스테레오스코픽 동영상 응용

포맷 파일의 무비 박스에는 하나 혹은 여러 개의 스테레오스코픽 비디오 트랙 박스와 장면기술을 위한 LAsER 트랙 박스가 포함 될 수 있다. 스테레오스코픽 비디오 트랙 박스는 미디어 정보 기술을 위한 미디어 박스 (mdia)를 포함하며 트랙 간 참조관계를 나타내는 트랙 레퍼런스 박스 (tref)와 트랙 레벨의 메타데이터 박스 (mdat)가 추가적으로 정의되어 있으며, 미디어 박스에 포함되는 스테레오스코픽 비디오 미디어 정보 박스는 스테레오스코픽 동영상 구성방법과 스테레오 및 모노스코픽 구간에 대한 정보를 기술하도록 정의하였다. 또한 두개의 시퀀스로 구성된 스테레오스코픽 동영상의 경우, 두개 트랙간의 상호 관련 정보를 tref박스에서 기술하도록 정의하였다. 또한, 트랙레벨의 메타데이터 박스에서는



<그림 3> 단일 스테레오스코픽 동영상 트랙의 파일 포맷 구조



〈그림 4〉 복수 스테레오스코픽 동영상 트랙의 파일 포맷 구조

스테레오스코픽 카메라, 디스플레이 정보 박스를 통해 카메라와 디스플레이 정보를 기술하도록 하였으며, 아이템 로케이션 박스 (iloc)의 extent_offset, extent_size 필드를 통해 스테레오 및 모노스코픽 동영상의 개별 구간을 지칭할 수 있도록 정의하였다 [5][6].

VI. 결론

다양한 멀티미디어 단말은 기존의 2D에서 발전하여 3D서비스를 제공하기 위한 기술개발이 진행되고 있다. 실사 동영상을 기반으로 한 3D 서비스를 제공하기 위한 기술로 스테레오스코픽

동영상 압축 기술이 개발되어지고 있으며, 또한, 이러한 스테레오스코픽 도영상을 보다 많은 사용자에게 공유되어지게 하기 위한 방법으로 MPEG에서는 스테레오스코픽 동영상 파일 포맷 저장방법이 정의되어졌다. 본 논문에서는 현재 MPEG에서 진행중인 스테레오스코픽 동영상을 효율적으로 저장하기 위한 ISO Base Media File Format 기반의 Stereoscopic Video Application Format 국제표준 규격을 소개하였다. 해당 규격에서는 단순한 스테레오스코픽 영동상 뿐만 아니라, 2D 동영상과 혼합된 스테레오스코픽 동영상 콘텐츠 및 다양한 대화형 서비스를 지원할 수 있도록 규격을 정의하고 있으며, 또한 기존의 2D 동영상만을 지원하는 단말에서 스테레오스코픽

동영상중 하나의 시점만을 제공하여 해당 콘텐츠를 소비할 수 있도록 규격을 정의하고 있다. 본 논문에서 소개한 스테레오스코픽 동영상 파일 포맷을 통해 스테레오스코픽 동영상이 보다 손쉽게 사용자들간에 공유되어짐으로 다양한 스테레오스코픽 동영상 관련 산업이 발전할 수 있을 것으로 기대한다.

참고문헌

- [1] "MAFs Overview", ISO/IEC JTC 1/SC 29/WG 11/MPEG2007/N9164, Lausanne, Switzerland, July 2007
- [2] "Request for Subdivision of ISO/IEC 23000", ISO/IEC JTC 1/SC 29/WG 11/MPEG2007/N9413, Shenzhen, China, Oct., 2007
- [3] "Information technology ? Coding of audio-visual objects ? Part 12:ISO base media file format", ISO/IEC JTC 1/SC 29/WG 11/MPEG2008/N9678, Antalya, Turkey, Jan., 2008
- [4] "Information technology ? Coding of audio-visual objects ? Part 20: Lightweight Application Scene Representation (LAsER) and Simple Aggregation Format (SAF)", ISO/IEC JTC 1/SC 29/WG 11/MPEG2005/N7588, Nice, France, Oct., 2005
- [5] "Text of ISO/IEC CD 23000-11 for Stereoscopic Video Application Format", ISO/IEC JTC 1/SC 29/WG 11/MPEG2008/N10283, Busan, Korea, Oct., 2008
- [6] 이장원, 김규현, 서덕영, 박광훈, 이운진, "ISO Base media file format 기반 스테레오스코픽 영상 저장방법", 한국통신학회 하계 학술대회, vol. 30., pp.652-655, 2008

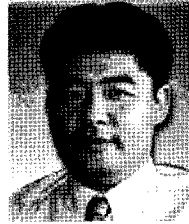
저자소개



김규현

1989년 2월 한양대학교 전자공학과 공학사
 1996년 7월 영국 University of Newcastle
 upon Tyne 전기전자공학과 공학박사
 1996년 8월 ~ 1997년 7월 영국 University of
 Sheffield, Reserach Fellow
 1997년 9월 ~ 2006년 2월 한국전자통신연구원
 대화형미디어연구팀장
 2006년 ~ 현재 경희대학교 전자정보대학 부교수
 주관심분야 : 영상처리, 멀티미디어, 디지털 대화형
 방송

저자소개



서덕영

1980년 2월 서울대학교 원자핵공학과 학사
 1985년 2월 Georgia Tech 핵공학 석사
 1990년 2월 Georgia Tech 전자공학 박사
 1990년 9월 ~ 1992년 3월 상공부 생산기술구연
 원 HDTV 연구개발단 선임연구원
 2002년 2월 ~ 2003년 2월 미국 North
 Carolina State Univ. 방문교수
 1992년 3월 ~ 현재 경희대학교 전자정보대학 교수
 주관심분야 : Networked video



박광훈

1985년 2월 연대대학교 전자공학과 학사
 1987년 2월 연대대학교 전자공학과 석사
 1991년 1월 Case Western Reserve
 University, Dept. of EEAP 석사
 1995년 1월 Case Western Reserve
 University, Dept. of EEAP 박사
 1995년 3월 ~ 1997년 2월 현대전자 멀티미디어
 연구소 책임연구
 1997년 3월 ~ 2001년 2월 연세대학교 문리대학
 전산학과 부교수
 2001년 3월 ~ 현재 경희대학교 전자정보대학 교수
 주관심분야 : 멀티미디어, 비디오 신호처리, 패턴인
 식, 영상처리, 계산지능