

스테레오스코픽 비디오 응용 포맷 기술 개요

□ 이장원, 김규현, 서덕영, 박광훈 / 경희대학교 전자정보대학

요약

MPEG에서는 빠르게 급변하는 시장의 요구를 충족시키기 위해, 기존의 긴 시간과 요구되는 표준화 방식과 달리, 서로 다른 기존 규격들을 특정 응용 서비스의 목적에 맞도록 조합해 하나의 표준으로 규격화하여 'MPEG-A (MPEG-Applications: ISO/IEC 23000) 멀티미디어 응용 표준'이라는 명칭으로 표준화 활동을 활발히 진행하고 있다. 본 논문에서는 2007년 10월 중국 신천에서 개최된 제 82차 MPEG회의에서 작업 표준안 (Working Draft)으로 채택되어, 2008년 10월 부산에서 개최된 86차 MPEG회의에서 최종 국제표준 초안 (final draft international standard)으로 승인된 스테레오스코픽 비디오 응용 포맷(Stereoscopic Video Application Format) 기술을 소개하고자 한다. 스테레오스코픽 비디오 응용 포맷은 현재 시장에 등장하여 점차 수요가 증가하고 있는 스테레오스코픽 카메라, 디스플레이, 모바일 단말 등의 장치를 이용한 스테레오스코픽 영상의 획득과, 출력 서비스를 위한 요구사항 및 저장규격을 정의한다. 이러한 기술을 활용하여 기존의 특정 장치에서만 획득 및 출력가능했던 폐쇄적 서비스에서 벗어나, 스테레오스코픽 콘텐츠의 저장 및 사용시간 교환, 방송,

리치미디어 서비스에서의 활용 등과 같이 3차원 입체영상 콘텐츠의 활용도를 높일 수 있을 것으로 기대된다.

1. 서론

최근 멀티미디어 기술의 급격한 발전으로 디지털 카메라, 디스플레이, 모바일 단말 등 디지털 멀티미디어 장비들의 기술 지원 사양이 높아짐에 따라 다양한 멀티미디어 콘텐츠가 등장하고, 여기에 대한 소비자의 요구 수준도 높아지고 있다. 소비자는 보다 높은 수준의, 보다 새로운 방식의 멀티미디어 서비스를 요구하고 있으며, 사업자들은 이에 발맞추기 위해 보다 고사양의, 보다 다양한 멀티미디어 서비스 기술을 경쟁적으로 개발하고 있다. 스테레오스코픽 영상은 이러한 시장의 요구에 부합하여 등장한

* 교신저자 : 김규현(경희대학교 전자정보대학)

기술로서, 안경식 또는 무안경식으로 두 개의 카메라로부터 획득된 좌영상과 우영상을 동시에 하나의 디스플레이에 출력함으로써 사용자에게 입체감 있는 영상을 제공하는 방식이다.

현재 스테레오스코픽 카메라 및 디스플레이 장치를 장착하여, 콘텐츠의 획득 및 출력이 가능한 모바일 단말이 상용화 되어있다. 그러나 스테레오스코픽 콘텐츠의 저장을 위한 공통의 표준 규격의 부재로 인해 스테레오스코픽 콘텐츠의 활용도는 하나의 장비에서 획득된 영상을 해당 장비를 통해 볼 수 있는 형태로 제한되어 있었다. 따라서 스테레오스코픽 콘텐츠의 저장 및 사용자간 교환이나 방송 서비스를 지원하기 위한 스테레오스코픽 영상 서비스에 대한 공통된 파일 포맷 관련 표준 기술이 요구된다.

본 논문에서는 스테레오스코픽 콘텐츠의 저장에 대한 국제 표준규격을 제안하고 그 활용도를 높이기 위한 ‘스테레오스코픽 비디오 응용 포맷 (Stereoscopic Video Application Format)’ 기술의 전반에 대하여 소개하고자 한다. 스테레오스코픽 비디오 응용 포맷은 지난 2007년 10월 중국 신천에서 개최된 제 82차 MPEG 회의에서 작업 표준안 (Working Draft)으로 채택되었으며 2008년 10월 부산에서 개최된 86차 MPEG회의에서 최종 국제 표준 초안 (final draft international standard) 으로 승인되었다 [1].

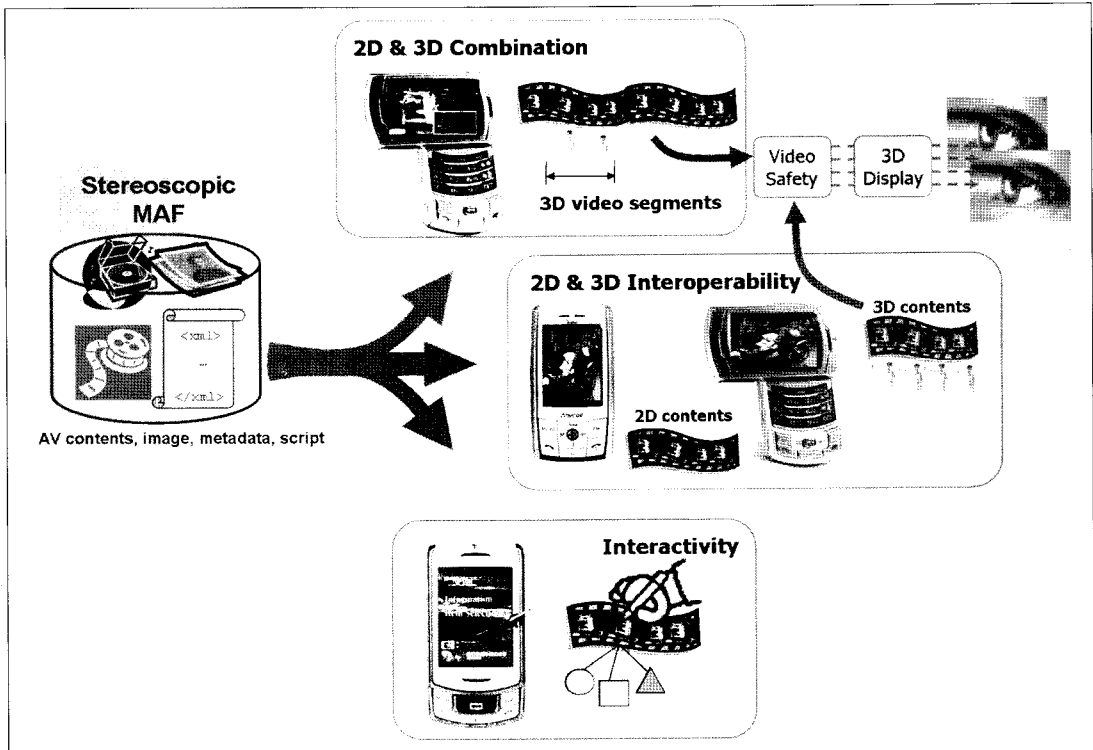
본 논문은 스테레오스코픽 비디오 응용 포맷 기술의 전반을 다음과 같은 구성으로 소개한다. 첫째로 본 논문에서 제안하는 저장방법이 포함해야 하는 기술의 범위 및 응용 시나리오와 이에 따른 요구사항을 명시하고 둘째로 스테레오스코픽 영상 구성 방식 (Composition Type)과 파일 구조 등 기술 세부사항에 대해 설명하며, 마지막으로 결론 및 향후 연구에 대해 기술한다.

II. 응용 시나리오 및 요구사항

위에서 언급한 바와 같이 현재 시장에 등장한 스테레오스코픽 카메라와 디스플레이가 장착된 모바일 단말은 스테레오스코픽 동영상 및 정지영상의 획득과 출력이 가능하다. 이처럼 현재 시장에서 사용되는 단말을 포함한 장비들의 기능을 기반으로 향후 활용될 응용 시나리오는 <그림 1>에서 제시한 바와 같이 예상되어질 수 있다 [2].

<그림 1>에서 제시한 바와 같이 첫째로 고려되어질 수 있는 응용 시나리오는 2D와 3D가 결합 (2D & 3D Combination)된 콘텐츠의 서비스이다. 즉, 하나의 멀티미디어 콘텐츠 전체가 스테레오스코픽 콘텐츠로 구성될 수 있으며, 또한 스테레오스코픽 콘텐츠와 모노스코픽 콘텐츠가 결합하여 하나의 멀티미디어 콘텐츠로도 구성되어질 수 있다. 예를 들어 한 화면에서 공간적으로 특정 부분의 이미지를 스테레오스코픽 영상으로 구성하여 눈에 띄는 효과를 줄 수 있고, 비보이 퍼포먼스나 광고 등과 같은 입체감의 효과를 강조하고 싶은 특정한 시간영역의 영상만을 스테레오스코픽 영상으로 구성할 수도 있다. 또한 스테레오스코픽 영상은 시점이 다른 두 가지 영상 데이터를 출력하기 때문에 장시간 보게 될 경우 모노스코픽 영상에 비해 쉽게 눈에 피로감을 줄 수 있어서, 2D와 3D의 시간적 혼합구성은 시각적 안전성 (Video Safety) 문제와도 연관된다. 따라서 스테레오스코픽 비디오 응용 포맷은 2D와 3D 콘텐츠를 서공간적으로 결합 (2D & 3D Combination)하여 하나의 콘텐츠로 기술할 수 있는 기능을 지원해야 한다.

둘째는 2D와 3D의 상호 운용성 (2D & 3D Interoperability) 이다. 현재 UCC (User Created Contents)와 같은 사용자들간의 콘텐츠 제작과 교환이 활발히 이루어지고 있는 것처럼 스테레오스코



〈그림 1〉 스테레오스코픽 콘텐츠의 응용 시나리오

픽 장비를 이용하여 획득된 콘텐츠를 사용자간에 교환하거나 방송 서비스를 하는 경우를 예상해 볼 수 있다. 스테레오스코픽 콘텐츠는 영상 사이즈나, 재생시간 등 여러 가지 기술 제원이 모노스코픽 콘텐츠와 다를 수 있기 때문에 기존의 모노스코픽 장비에서는 스테레오스코픽 콘텐츠의 재생이 불가능할 수 있다. 기존의 모노스코픽 장비에서 스테레오스코픽 콘텐츠를 재생, 혹은 스테레오스코픽 장비에서 스테레오스코픽 콘텐츠를 모노스코픽 영상으로 재생하는 경우를 고려하여 2D와 3D의 상호 운용성, 즉 스테레오스코픽 영상이라 할 지라도 기존의 모노스코픽 장비에서는 모노스코픽 방식으로 볼 수 있는 방안이 스테레오스코픽 비디오 응용 포맷 기술에서

고려되어야 한다.

셋째는 장면기술을 통한 양방향 대화형 서비스이다. 현재 스테레오스코픽 콘텐츠는 대화형 서비스에 이점이 있는 모바일 단말에서 널리 사용되고 있다. 장면기술을 이용한 대화형 기능 제공을 통해 사용자의 입력에 의한 2D, 3D 전환 등 다양한 리치미디어 서비스의 가능성을 제공할 수 있다.

현재 시장에서 사용되는 모바일 단말의 특성을 고려하여, 앞서 소개한 스테레오스코픽 콘텐츠의 응용 시나리오와 이에 따른 요구사항을 충족시키고자 스테레오스코픽 비디오 응용 포맷은 지원 컴포넌트로써 다음의 표 1과 같은 기술을 정의하였다.

스테레오스코픽 비디오 응용 포맷 기술은 ISO

<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

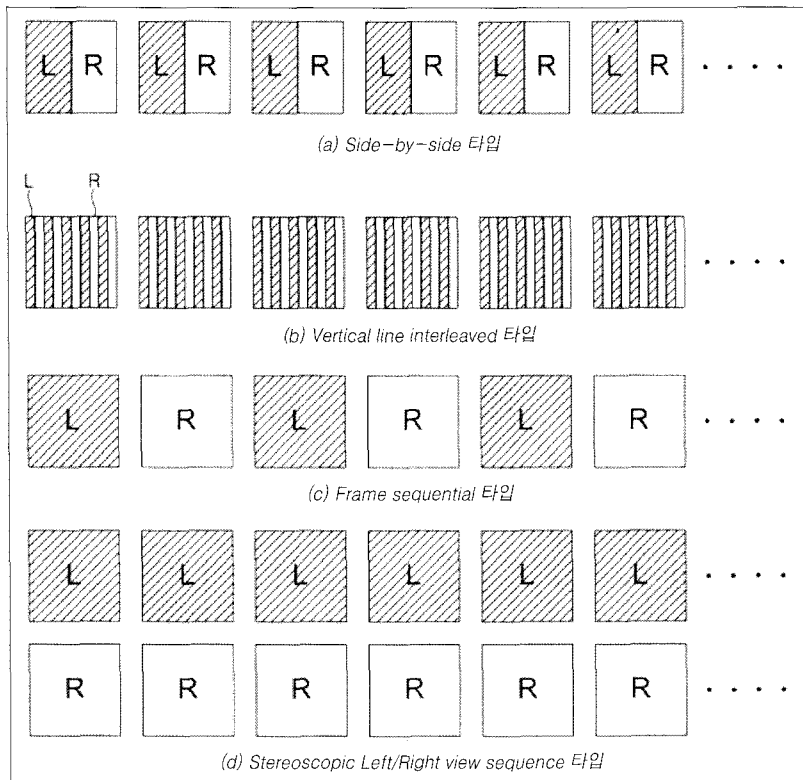
Simple Profile level 3, MPEG-4 Part10 Baseline Profile 1.3을 사용하고, 오디오로는 AAC-LC 및 AAC+, 보이스는 AMR과 EVRC, 정지영상은 JPEG, PNG를 사용한다. 또한, 앞서 언급한 대화형 기능 제공을 위해 장면기술로 LAsER (Lightweight Application Scene Representation) 를 사용한다 [4].

Base Media File format (ISO/IEC 14496-12)을 기반으로 설계되었다 [3]. 본 기술은 모바일 단말에서의 사용을 주 목적으로 설계되었기 때문에, 표1에 나타난 바와 같이 비디오 코덱으로 MPEG-4 Part2

III. 기술 세부사항

1. 스테레오스코픽 콘텐츠 컴포지션 타입

스테레오스코픽 콘텐츠의 컴포지션 타입은 하나



<그림 2> 스테레오스코픽 콘텐츠 컴포지션 타입

혹은 두 개의 스트림에서 좌영상과 우영상 데이터의 위치 구성 방식을 말하며, 현재 시장에는 'side-by-side', 'above and below', 'pixel-by-pixel' 타입과 여러 종류의 interleaved, sequential 타입이 존재한다. 단일 표준에서 모든 종류의 컴포지션 타입을 정의하기는 어려우므로 스테레오스코픽 비디오 응용 포맷에서는 모바일 단말을 우선으로 널리 사용되는 타입들을 고려하여 'Side-by-side', 'Vertical line interleaved', 'Frame sequential', 'Stereoscopic Left/Right view sequence' 타입의 네 가지 컴포지션 타입을 정의한다.

Side-by-side 타입은 가장 널리 사용되는 스테레오스코픽 컴포지션 타입의 하나로 <그림 2(a)>와 같이 하나의 프레임에 좌우 반으로 나누어 좌영상과 우영상을 옆으로 위치시킨다. Side-by-side 타입은 영상의 해상도가 반으로 줄기는 하지만 기존의 비트율 그대로 영상을 압축할 수 있고, 기존 시스템의 변환 없이 플레이어에서 재생될 수 있다는 이점이 있다.

Vertical line interleaved 타입은 <그림 2(b)>와 같이 하나의 프레임에 수직방향으로 좌영상과 우영상을 하나의 픽셀 라인씩 번갈아 위치시키는 방식이다. Vertical line interleaved 타입은 모든 수직라인마다 불연속성이 존재하므로 압축 효율이 낮다. 그러나 모바일 디스플레이 장치에서 가장 많이 사용되는 패럴랙스 배리어 디스플레이 방식에 변환 없이 직접 적용될 수 있다는 이점이 있다.

Frame sequential 타입은 <그림 2(c)>와 같이 좌영상 프레임과 우영상 프레임을 번갈아 하나의 스트림을 구성한다. 모노스코픽 콘텐츠의 두 배의 비트율이 보장된다면 풀(full) 해상도의 스테레오스코픽 콘텐츠 재생이 가능하다.

Stereoscopic Left/Right view sequence 타입은

<그림 2(d)>와 같이 두 개의 독립된 좌우 영상의 스트림을 구성하는 방식이다. 독립된 두 개의 스트림이 직접 모노스코픽 콘텐츠로 활용될 수 있는 이점이 있으며 스테레오스코픽 콘텐츠 재생 시 각각의 프레임 쌍들에 대한 동기화가 필요하다.

2. 스테레오스코픽 비디오 응용 포맷 파일 구조

스테레오스코픽 비디오 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 크게 파일 타입을 정의하는 파일 타입 박스(ftyp), 비디오 등의 미디어 재생 관리를 위한 정보를 기술하는 무비 박스(moov), 실제 데이터를 포함하는 미디어 데이터 박스(mdac)로 구성되어 있다. 미디어 재생에 관한 세부적인 정보들은 무비 박스 안의 트랙 박스(trak)에 기술된다.

스테레오스코픽 비디오 응용 포맷 파일의 무비 박스에는 하나 혹은 여러 개의 스테레오스코픽 비디오 트랙 박스와 장면기술을 위한 LAsER 트랙 박스가 포함될 수 있다. 스테레오스코픽 비디오 트랙 박스는 미디어 정보 기술을 위한 미디어 박스(mdia)를 포함하며 트랙간 참조관계를 나타내는 트랙 레퍼런스 박스(tref)와 트랙 레벨의 메타데이터 박스(mdac)가 추가적으로 포함될 수 있다.

미디어 박스에 포함되는 스테레오스코픽 비디오

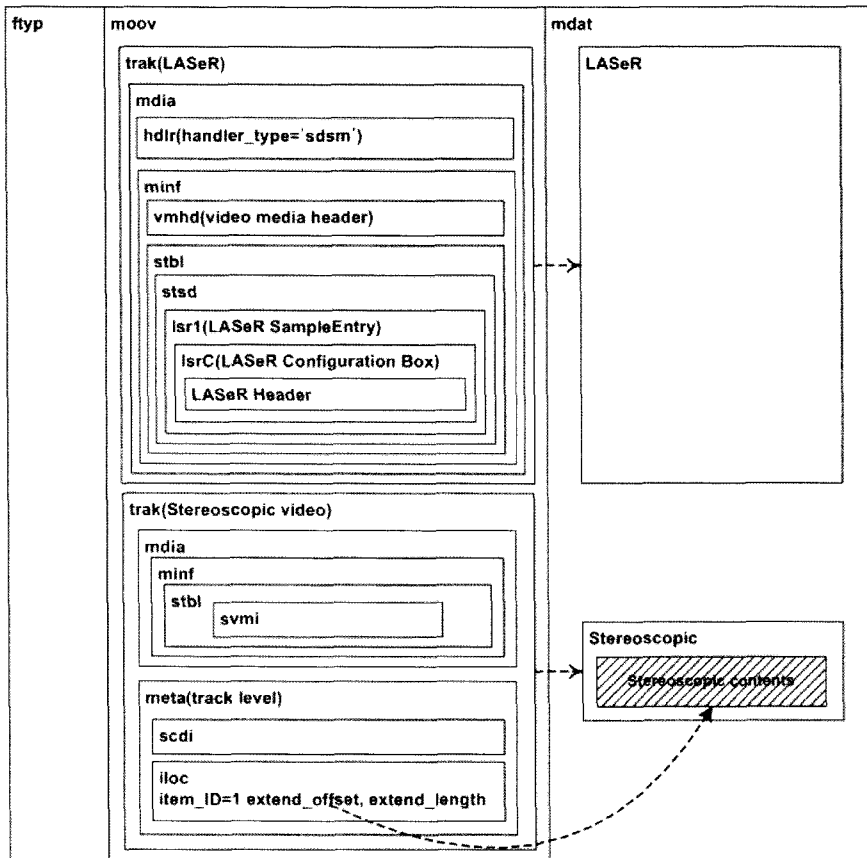
미디어 정보 박스는 스테레오스코픽 콘텐츠의 컴포지션 타입 등의 비주얼 타입과 스테레오, 모노스코픽 구간에 대한 정보를 기술한다.

스테레오스코픽 비디오 응용 포맷에서의 트랙 레퍼런스 박스는 <그림 2(d)>에서 나타내고 있는 Stereoscopic Left/Right view sequence 타입에서 두 개의 스테레오스코픽 비디오 트랙 사이의 참조 관계를 기술한다.

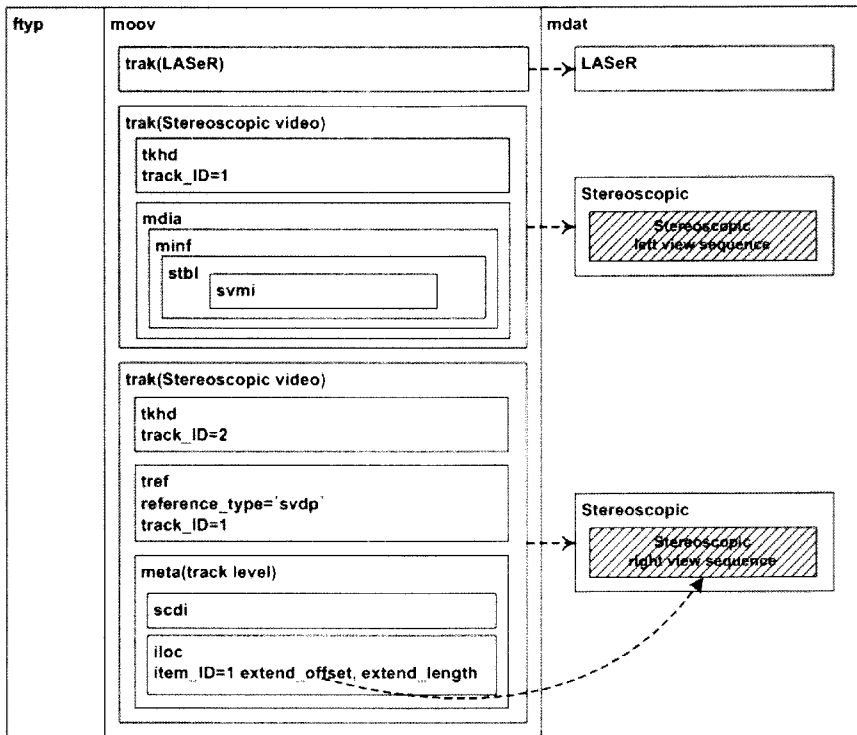
트랙레벨의 메타데이터 박스는 스테레오스코픽 카메라, 디스플레이 정보 박스와 아이템 로케이션

박스 (iloc)를 포함할 수 있다. 아이템 로케이션 박스는 extent_offset, extent_size 필드를 통해 파일 내의 일정 부분을 아이템으로 명명할 수 있으며, 스테레오스코픽 비디오 응용 포맷에서의 아이템 (Item)은 스테레오스코픽 콘텐츠의 구간을 지칭한다. 아이템 로케이션 박스를 통해 식별된 스테레오스코픽 구간은 카메라, 디스플레이 정보가 부여되어 재생될 수 있다.

스테레오스코픽 비디오 응용 포맷의 파일 구조는 위에서 설명한 박스들을 기반으로 하여, 컴포지션



<그림 3> 단일 스테레오스코픽 트랙의 파일 포맷 구조



〈그림 4〉 Left/Right view sequence 타입의 파일 포맷 구조

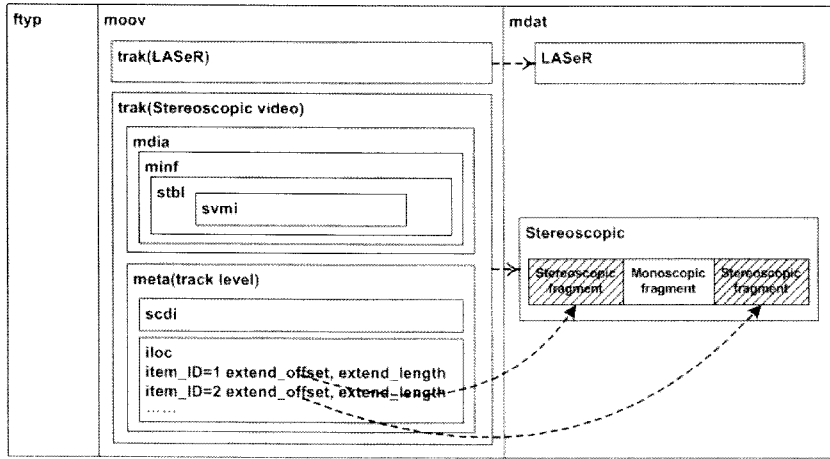
타입과 스테레오스코픽 트랙의 개수에 따라 다음과 같이 형성될 수 있다

첫째로 <그림 3>과 같이 전체 스테레오스코픽 콘텐츠가 단일 트랙으로 구성된 경우이다. 앞장에서 설명한 side-by-side, vertical line interleaved, frame sequential 타입이 단일 트랙으로 구성될 수 있으며, 아이템 로케이션 박스는 하나의 아이템으로 스테레오스코픽 비디오 데이터의 전구간을 지칭하며, 이 구간에 스테레오스코픽 카메라, 디스플레이 정보 박스의 카메라 파라미터들을 적용하여 렌더링 할 수 있게 한다.

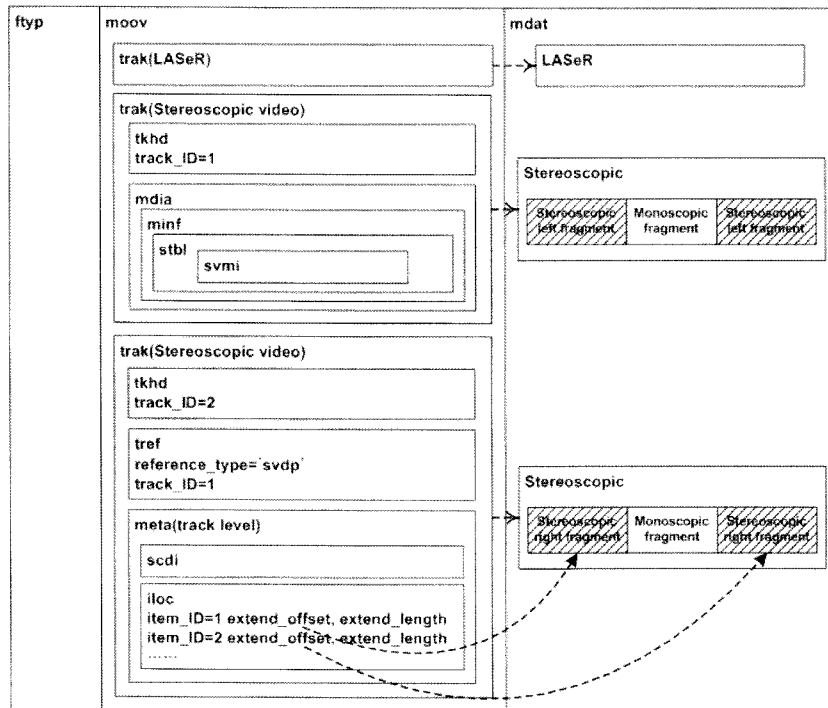
둘째로 <그림 4>와 같이 전체 스테레오스코픽 콘텐츠가 Left/Right view sequence 타입으로 구성된

경우이다. 스테레오스코픽 비디오 트랙은 좌영상, 우영상 각각 두 개로 구성되며, 트랙 레퍼런스 박스를 통해 두 개의 트랙이 연결된다. 이 경우 스테레오스코픽 콘텐츠 정보 기술을 위한 박스들의 내용이 중복될 수 있으므로, 스테레오스코픽 비디오 미디어 정보 박스와, 카메라 디스플레이 정보 박스, 아이템 로케이션 박스는 하나의 트랙에만 포함될 수도 있다.

셋째로 <그림 5>처럼 스테레오-모노스코픽 혼합 구성 콘텐츠가 단일트랙으로 구성된 경우를 생각해 볼 수 있다. 파일 포맷의 기본 구조는 <그림 3>과 유사하며 스테레오스코픽 데이터 구간 식별을 위한 아이템 로케이션 박스의 기술 내용이 달라진다. 아이템 로케이션 박스는 스테레오스코픽 구간의 개수만



<그림 5> 스테레오-모노스코픽 혼합 구성 콘텐츠의 단일 트랙 파일 포맷 구조



<그림 6> 스테레오-모노스코픽 혼합 구성 콘텐츠의 Left/Right view sequence 타입 파일 포맷 구조

클립 아이템을 명명하여 스테레오스코픽 데이터 구간의 위치를 기술한다. 이처럼 스테레오스코픽 데이터

구간을 기록함으로써 스테레오-모노스코픽 혼합 영상의 파일 저장이 가능해지며, 스테레오 구간마다

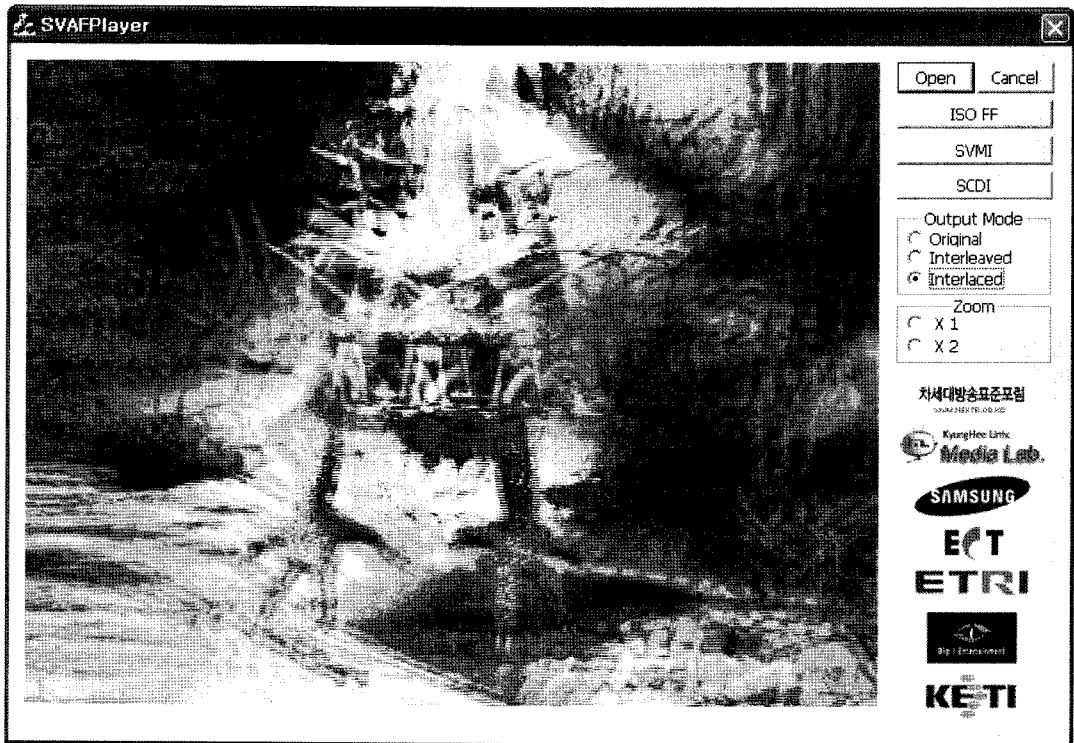
각기 다른 카메라 디스플레이 정보를 적용할 수 있게 된다.

마지막으로 <그림 6>와 같이 스테레오-모노스코픽 혼합구성 콘텐츠가 Left/Right view sequence 타입으로 구성된 경우이다. 파일 포맷의 기본 구조는 <그림 4>와 유사하며, <그림 5>의 경우와 같이 아이템 로케이션 박스가 여러 개의 스테레오스코픽 구간을 지칭한다.

3. 스테레오스코픽 비디오 응용 포맷 파일 재생

앞서 설명한 스테레오스코픽 응용 포맷의 파일구

조로 파일을 구성하여 스테레오스코픽 콘텐츠를 재생해 볼 수 있다. 플레이어는 스테레오스코픽 응용 포맷 파일을 파싱 (parsing)하여 스테레오스코픽 비디오 미디어 정보 박스 (svmi)를 통해 스테레오스코픽 컴포지션 타입과 스테레오-모노 구간 정보를 식별하고 스테레오스코픽 카메라 디스플레이 박스 (scdi)의 내용을 바탕으로 카메라, 디스플레이 파라미터를 활용하여 스테레오스코픽 영상을 디스플레이 장비의 특성에 맞게 좌, 우 영상을 구성하여 재생한다. <그림 7>은 스테레오스코픽 응용 포맷 파일을 좌영상과 우영상을 수평라인으로 교대 배열한 인터레이스 (interlace) 방식으로 영상을 재생하는 모습의 예이다.



<그림 7> 스테레오스코픽 비디오 응용 포맷 파일 재생의 예

IV. 결론

멀티미디어 기술이 발전함에 따라 스테레오스코픽 카메라, 디스플레이 등의 기기들이 등장하였고 이에 따라 소비자는 실감형 입체영상을 쉽게 접할 수 있게 되었다. 하지만 현재의 스테레오스코픽 콘텐츠는 단말마다 각기 다른 방식으로 획득, 재생되어 사용자간 콘텐츠 교환이나 방송 서비스를 하기에는 한계가 있다. 이와 관련하여 스테레오스코픽 콘텐츠의 활용도를 높이기 위한 저장 포맷관련 표준 기술 규격이 요구되어 왔다.

본 논문에서는 이러한 요구에 부합하기 위하여, 스테레오스코픽 콘텐츠의 규격화된 표준 기술 제정을 목적으로 한, 스테레오스코픽 비디오 응용 포맷에 대한 응용 시나리오 및 요구사항, 기술 세부사항 등 전반에 대해 소개하였다.

향후 모바일 단말에서의 사용을 주 목적으로 고안된 이 기술을 기반으로, 카메라 파라미터 및 디스플레이를 위한 파라미터들을 보완하여, 고해상도 출력 장치를 위한 스테레오스코픽 영상 저장방법으로의 확장 및 보완 연구가 필요하다.

참고 문헌

- [1] "Text of ISO/IEC FDIS 23000-11 for Stereoscopic Video Application Format", ISO/IEC JTC 1/SC 29/WG 11/MPEG2008/N10283, Korea, Busan, Oct., 2008
- [2] "MAFs Overview", ISO/IEC JTC 1/SC 29/WG 11/MPEG2008/N10233, Korea, Busan, Oct. 2008
- [3] "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
- [4] "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

필자 소개



이장원

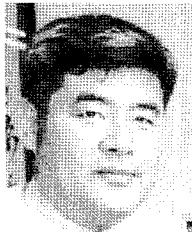
- 2007년 2월 : 경희대학교 전자공학과 졸업(공학사)
- 2007년 3월 ~ 현재 : 경희대학교 전자전파공학과 석박사과정
- 주관심 분야 : 멀티미디어 서비스, 디지털 방송 분야

필자 소개



김규현

- 1989년 2월 : 한양대학교 전자공학과 공학사
- 1992년 9월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학석사
- 1996년 7월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학박사
- 1996년 ~ 1997년 : 영국 University of Sheffield, Research Fellow
- 1997년 ~ 2006 : 한국전자통신연구원 대화형미디어연구팀장
- 2006년 ~ 현재 : 경희대학교 전자정보대학 부교수
- 주관심분야 : 영상처리, 멀티미디어통신, 디지털 대화형 방송



서덕영

- 1980년 2월 : 서울대 핵공학과 학사
- 1986년 2월 : 미국 Georgia Tech. 핵공학 석사
- 1990년 6월 : 미국 Georgia Tech. 전기및컴퓨터공학과 박사
- 1990년 ~ 1992년 : 상공부 생산기술 연구원, HDTV 연구개발단 선임연구원
- 1992년 3월 ~ 현재 : 경희대학교 전자정보학부 교수
- 주관심 분야 : networked video



박광훈

- 1985년 2월 : 연세대학교 전자공학과 학사
- 1987년 7월 : 연세대학교 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 멀티미디어연구소 책임연구원
- 1997년 3월 ~ 2001년 2월 : 연세대학교 문리대학 전산학과 부교수
- 2001년 3월 ~ 현재 : 경희대학교 컴퓨터공학과 교수
- 주관심분야 : 멀티미디어, 비디오 신호처리, 패턴인식, 영상처리, 계산지능