



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년06월19일
(11) 등록번호 10-1276605
(24) 등록일자 2013년06월13일

(51) 국제특허분류(Int. Cl.)
H04N 13/00 (2006.01) H04N 21/83 (2011.01)
(21) 출원번호 10-2008-0057862
(22) 출원일자 2008년06월19일
심사청구일자 2011년06월08일
(65) 공개번호 10-2008-0112149
(43) 공개일자 2008년12월24일
(30) 우선권주장
1020070059709 2007년06월19일 대한민국(KR)
(뒷면에 계속)
(56) 선행기술조사문헌
KR100716142 B1*
KR1020070061227 A*
KR1020030042090 A
*는 심사관에 의하여 인용된 문헌
기술이전 희망 : 기술양도, 실시권허여, 기술지도

(73) 특허권자
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
(72) 발명자
윤국진
대전광역시 유성구 송림로 13, 송림마을 호반리젠
시빌 106동 1504호 (하기동)
조숙희
대전광역시 유성구 신성로 74-1, 301호 (신성동)
(뒷면에 계속)
(74) 대리인
특허법인 신성

전체 청구항 수 : 총 11 항

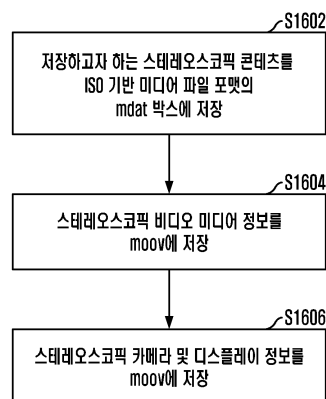
심사관 : 김기호

(54) 발명의 명칭 스테레오스코픽 데이터의 저장 및 재생을 위한 메타데이터구조 및 이를 이용하는 스테레오스코픽 콘텐츠 파일의 저장방법

(57) 요약

본 발명은 스테레오스코픽 데이터의 저장 및 재생을 위한 메타데이터 구조 및 이를 이용하는 스테레오스코픽 콘텐츠 파일의 저장 방법에 관한 것이다. 본 발명에 의한 저장 방법은 상기 스테레오스코픽 콘텐츠의 구간 정보에 관한 스테레오스코픽 비디오 미디어 정보를 상기 스테레오스코픽 콘텐츠를 재생하기 위하여 요구되는 메타데이터 정보로서 저장하는 단계와, 상기 저장하고자 하는 스테레오스코픽 콘텐츠의 부호화 스트림을 저장하는 단계를 포함한다. 스테레오스코픽 데이터는 모노스코픽(2차원) 및 스테레오스코픽(3차원) 데이터가 혼용으로 구성되는 경우 또는 스테레오스코픽 데이터로만 구성되는 경우 모두를 의미한다.

대표도 - 도16



(72) 발명자

허남호

대전광역시 유성구 은구비남로 34, 열매마을아파트
801동 1001호 (노은동)

김진웅

대전광역시 유성구 엑스포로 448, 305동 1603호 (전민동, 엑스포아파트)

이수인

대전광역시 서구 둔산로 155, 크로바아파트 106동 606호 (둔산동)

(30) 우선권주장

1020070061362	2007년06월22일	대한민국(KR)
1020070096267	2007년09월21일	대한민국(KR)
1020070098665	2007년10월01일	대한민국(KR)

이 발명을 지원한 국가연구개발사업

과제고유번호	2005-S-403-02
부처명	정보통신부 및 정보통신연구진흥원
연구사업명	IT차세대핵심기술개발
연구과제명	지능형 통합정보 방송(SmaR TV) 기술개발
주관기관	한국전자통신연구원
연구기간	2006-01-01 ~ 2006-12-31

특허청구의 범위

청구항 1

스테레오스코픽 콘텐츠의 구간 정보를 포함하는 스테레오스코픽 비디오 미디어 정보를 입력받는 단계; 및
 상기 스테레오스코픽 비디오 미디어 정보를 저장하기 위한 비디오 미디어 정보 영역을 포함하는 파일 구조에 상기 스테레오스코픽 비디오 미디어 정보를 저장하는 단계를 포함하며,
 상기 스테레오스코픽 비디오 미디어 정보는
 스테레오스코픽 구간과 모노스코픽 구간간의 변경 구간 수를 표시하는 정보를 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 2

제1항에 있어서,
 상기 파일 구조는
 상기 스테레오스코픽 콘텐츠를 저장하기 위한 콘텐츠 영역을 더 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 3

제 1 항에 있어서,
 상기 스테레오스코픽 콘텐츠에 대한 스테레오스코픽 카메라 및 디스플레이 정보를 입력받는 단계; 및
 상기 스테레오스코픽 카메라 및 디스플레이 정보를 저장하기 위한 카메라 및 디스플레이 정보 영역을 더 포함하는 상기 파일 구조에 상기 스테레오스코픽 카메라 및 디스플레이 정보를 저장하는 단계를 더 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 4

제1항에 있어서,
 상기 스테레오스코픽 비디오 미디어 정보는
 상기 스테레오스코픽 콘텐츠에 대한 스테레오스코픽 영상 구성 정보를 더 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 5

제4항에 있어서,
 상기 스테레오스코픽 영상 구성 정보는
 상기 스테레오스코픽 콘텐츠를 구성하는 영상의 구성방법에 따라 사이드 바이 사이드(side by side), 프레임 시퀀셜(frame sequential) 및 좌우영상 시퀀스(left/right view sequence)를 포함하는 타입으로 분류된 정보를 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 6

삭제

청구항 7

삭제

청구항 8

제1항에 있어서,
 상기 스테레오스코픽 비디오 미디어 정보는

연속하는 모노스코픽 또는 스테레오스코픽 콘텐츠의 시작 샘플들-상기 샘플은 비디오, 시간 연속적인 일련의 비디오 프레임들의 각각의 프레임(an individual frame of video, a time-contiguous series of video frames)을 의미함-의 수를 카운트하는 정보를 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 9

제1항에 있어서,

상기 스테레오스코픽 비디오 미디어 정보는

현재의 샘플이 스테레오스코픽인지 여부를 표시하는 정보를 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 10

제 3 항에 있어서,

상기 스테레오스코픽 카메라 및 디스플레이 정보는

디스플레이 파라미터 존재 여부에 대한 정보, 카메라 파라메타 존재 여부에 대한 정보, 카메라 이동 정보, 렌즈로부터 이미지 면까지의 거리 정보, 베이스라인으로부터 컨버전스 포인트까지의 거리 정보 또는 카메라의 배열 정보 중 어느 하나 이상을 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 11

제 3 항에 있어서,

상기 스테레오스코픽 카메라 및 디스플레이 정보는

좌우 이미지 사이의 최대 디스패리티 또는 최소 디스패리티 정보를 포함하는, 스테레오스코픽 콘텐츠의 저장 방법.

청구항 12

제 1 항에 있어서,

상기 파일 구조는

ISO 베이스 미디어 파일 포맷에 기초한 구조인, 스테레오스코픽 콘텐츠 저장 방법.

청구항 13

제 3 항에 있어서,

상기 파일 구조는

ISO 베이스 미디어 파일 포맷에 기초한 구조이며,

상기 스테레오스코픽 카메라 및 디스플레이 정보는

스테레오스코픽 구간의 오프셋 및 길이 정보를 저장하는 아이템 로케이션 박스(item location box)를 참조하기 위한 상기 스테레오스코픽 구간의 식별 정보를 포함하는, 스테레오스코픽 콘텐츠 저장 방법.

청구항 14

삭제

청구항 15

삭제

명세서

발명의 상세한 설명

기술 분야

- [0001] 본 발명은 스테레오스코픽 데이터의 저장 및 재생을 위한 메타데이터 구조 및 이를 이용하는 스테레오스코픽 콘텐츠 파일 저장 방법에 관한 것으로서, 특히 2차원 및 3차원 콘텐츠가 임의적으로 구성되어 제공되는 경우 요구되는 메타데이터(제어 정보) 및 이를 이용하는 스테레오스코픽 콘텐츠 파일의 저장 방법에 관한 것이다.
- [0002] 본 발명은 정보통신부의 IT차세대핵심기술개발사업의 일환으로 수행한 연구로부터 도출된 것이다[과제관리번호 : 2005-S-403-02, 과제명 : 지능형 통합정보 방송(SmarTV) 기술개발(Development of Super-intelligent Multimedia Anytime-anywhere Realistic TV(SmarTV) Technology)].

배경 기술

- [0003] 스테레오스코픽 콘텐츠를 이용한 응용 서비스 및 관련 기기에 대한 시장이 국내외에서 휴대폰, 디지털카메라, DVD, PDP 등을 중심으로 형성되고 있다. 이와 관련해 스테레오스코픽 콘텐츠의 획득, 저장 및 재생을 위해 부가적으로 요구되는 시스템 정보 또는 제어 정보(즉, 메타데이터)와, 이를 포함하는 파일 포맷에 대한 표준 제정이 요구되고 있다.
- [0004] 한국공개특허 제2006-0056070호(발명의 명칭: 엠팩-4 객체기술자 정보 및 구조를 이용한 3차원 동영상처리 장치 및 그 방법, 이하 “특허문헌 1”이라고 함)는 MPEG-4 객체기술자(Object Descriptor) 구조를 토대로 3차원 동영상 종류, 다양한 디스플레이 방식 및 시점수 등과 같은 새롭게 요구되는 정보를 추가한 3차원 동영상 객체기술자를 제공한다. 또한, 한국특허출원 제2006-0100258호(발명의 명칭: 스테레오스코픽 영상 데이터의 전송 방법, 이하 “특허문헌 2”라고 함)는 스테레오스코픽 콘텐츠에 대하여 부호화된 영상의 데이터에 대한 파일 포맷으로서, 부호화된 스테레오스코픽 영상 정보를 포함하는 영상 데이터부, 스테레오스코픽 영상 정보를 디코딩하고 재생하기 위한 메타데이터를 포함하는 헤더부를 구비하는 파일 포맷을 제공한다.
- [0005] 그러나 특허문헌 1 및 2는 2차원 콘텐츠 및 3차원 콘텐츠가 임의적으로 구성되어 제공되는 경우(즉, 2차원 콘텐츠 및 3차원 콘텐츠가 혼용되는 경우), 각 콘텐츠를 구분하는 방법, 각각의 3차원 콘텐츠가 서로 다른 스테레오스코픽 카메라 및 디스플레이 정보를 가질 경우 이를 지원하기 위한 스테레오스코픽 카메라 및 디스플레이 정보를 제공하는 방법, 2개의 부호화 스트림으로 구성되는 3차원 콘텐츠 경우 스테레오스코픽 트랙 참조 정보 등을 제공하지 못한다.

발명의 내용

해결 하고자하는 과제

- [0006] 본 발명은 이러한 문제점을 해결하기 위하여 제안된 것으로, 스테레오스코픽 콘텐츠의 저장 방법을 제공하는데 일 목적이 있다.
- [0007] 또한, 본 발명은 2차원(또는 2D) 콘텐츠 및 3차원 콘텐츠가 혼용되어 제공되는 경우 이를 다양한 3차원(또는 3D) 단말 환경에서 다운로드하여 재생할 때 2D 콘텐츠 및 3D 콘텐츠의 구분하고, 자동으로 패럴랙스 베리어를 온(On)/오프(Off)하여 2D 또는 3D로 디스플레이 할 수 있는 스테레오스코픽 콘텐츠의 저장 방법을 제공하는데 다른 목적이 있다.
- [0008] 또한, 본 발명은 파일에 저장되는 스테레오스코픽 콘텐츠가 구간별로 서로 다른 스테레오스코픽 카메라 및 디스플레이 정보를 가지는 경우, 이를 위한 스테레오스코픽 카메라/디스플레이 정보를 저장하는 방법을 제공하는데 또 다른 목적이 있다.
- [0009] 또한, 본 발명은 스테레오스코픽 콘텐츠의 기본 부호화 스트림이 2개인 경우 주요(메인)/부가 트랙을 구분하고, 기존 2D 단말과의 호환성을 유지하면서 각 트랙에 포함되어 사용될 수 있는 스테레오스코픽 카메라/디스플레이 정보의 중복성을 제거할 수 있는 스테레오스코픽 콘텐츠의 저장 방법을 제공하는데 또 다른 목적이 있다.
- [0010] 본 발명의 다른 목적 및 장점들은 종래의 ISO 기반 미디어 파일 포맷을 기반으로 하는 본 발명의 실시예에 의해 보다 분명하게 이해될 것이다. 또한, 본 발명의 목적 및 장점들은 특허청구범위에 나타난 수단 및 그 조합에 의해 실현될 수 있음을 쉽게 알 수 있을 것이다.

과제 해결수단

- [0011] 상기 목적을 달성하기 위해 제안된 본 발명은 스테레오스코픽 콘텐츠를 저장하는 방법에 있어서, 상기 스테레오스코픽 콘텐츠의 구간 정보에 관한 스테레오스코픽 비디오 미디어 정보를 상기 스테레오스코픽 콘텐츠를 재생하기 위하여 요구되는 메타데이터 정보로서 저장하는 단계와, 상기 저장하고자 하는 스테레오스코픽 콘텐츠의 부호화 스트림을 저장하는 단계를 포함하는 것을 일 특징으로 한다.
- [0012] 또한, 본 발명은 스테레오스코픽 콘텐츠를 저장하는 파일 포맷 구조에 있어서, 상기 스테레오스코픽 콘텐츠의 부호화 스트림을 저장하는 미디어 데이터 박스와, 상기 스테레오스코픽 콘텐츠의 구간 정보에 관한 스테레오스코픽 비디오 미디어 정보를 저장하는 박스를 포함하는 것을 다른 특징으로 한다.
- [0013] 상술한 목적, 특징 및 장점은 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이며, 그에 따라 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에 그 상세한 설명을 생략하기로 한다.
- [0014] 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 일실시예를 상세히 설명하기로 하며 설명을 도움을 위하여 종래의 ISO 기반 미디어 파일 포맷 기반으로 스테레오스코픽 콘텐츠를 저장하는 방법에 대하여 기술한다.

효과

- [0015] 전술한 바와 같은 본 발명에 의하면, 2차원 콘텐츠 및 3차원 콘텐츠가 임의적으로 구성되어 제공되는 스테레오스코픽 콘텐츠를 용이하게 저장하고 제공할 수 있으며, 2D 콘텐츠 및 3D 콘텐츠 구분 정보를 통하여 패럴랙스 베리어를 자동으로 2D 또는 3D 디스플레이로 변환하여 사용자는 스테레오스코픽 콘텐츠를 용이하게 감상할 수 있다.

발명의 실시를 위한 구체적인 내용

- [0016] 도 1은 스테레오스코픽 콘텐츠를 제공하기 위해 구성할 수 있는 다양한 콘텐츠 구성 형태를 나타낸다. 도 1의 (a)는 기본 부호화 스트림(elementary stream: ES)이 1개인 경우(즉, ES=1 경우)이고, 도 1의 (b)는 ES=2 경우이다. 단일 포맷(single format)은 스테레오스코픽 영상 구성 방식이 동일하며 하나의 카메라 파라미터 및 하나의 디스플레이 정보만을 가지는 포맷이다. 다중 포맷(multiple format)은 스테레오스코픽 영상 구성 방식이 다르거나, 또는 스테레오스코픽영상 구성 방식은 같으나 카메라 파라미터가 다른 경우, 또는 다수개의 디스플레이 정보 또는 기타 정보를 가지는 포맷이다.
- [0017] 스테레오스코픽 콘텐츠의 타입에는 i) 양안식 3차원 동영상 서비스와 같은 스테레오스코픽 비디오 콘텐츠(Stereoscopic Video Content), ii) 양안식 3차원 정지영상 서비스(예: 슬라이드 쇼), 2D (모노)비디오와 3D 데이터 서비스(특정 화면 또는 부분)의 결합과 같은 스테레오스코픽 이미지 콘텐츠(Stereoscopic Image Content), iii) 2D(모노)비디오와 3D 비디오(특정 화면 또는 부분)의 결합과 같은 스테레오스코픽 혼용 콘텐츠(Monoscopic and Stereoscopic Mixed Content)가 있다.
- [0018] 도 2 및 도 3은 본 발명에 의한 스테레오스코픽 비디오 어플리케이션 포맷(Stereoscopic video application format: SS-VAF)의 기본 구조도이다.
- [0019] 도 2는 MPEG-4 시스템 정보가 있는 파일 포맷 구조이고, 도 3은 MPEG-4 시스템 정보가 없는 파일 포맷 구조이다. 도시된 바와 같이, SS-VAF는 'ftyp', 'moov', 'mdat' 박스를 포함한다. 이하, SS-VAF의 박스들(boxes)의 신택스(Syntax)와 세만틱스(Semantics)를 설명한다. 본 발명에 따른 박스는 SS-VAF의 구조에 포함되며, 그 위치는 형태에 따라 달라질 수 있으며, 박스에 포함되는 정보는 별도로 사용될 수 있다.
- [0020] 1. 'scty'(Stereoscopic Content Type)

[0021] 'scty'는 기본적으로 구성되는 콘텐츠의 타입을 나타낸 것으로서, 모노스코픽 콘텐츠(일반 2D)와 스테레오스코픽 콘텐츠를 구분한다. 표 1은 'scty'에 대한 선택스를 나타낸다. 또한 'scty'를 구성하는 'Stereoscopic_Content_type'은 상기 'ftyp' 및 다른 기존의 박스에 포함되어 사용될 수 있다.

표 1

[0022]

```
Aligned(8) class StereoscopicContentTypeBox extend Box('sfty') {
    unsigned int(2)    Stereoscopic_Content_Type;
    unsigned int(6)    reserved;
}
```

[0023] 표 1에서 Stereoscopic_Content_Type는 스테레오스코픽 콘텐츠 타입(Stereoscopic content type)을 나타내며, 표 2와 같은 의미를 갖는다.

표 2

[0024]

Value	Description
00	Stereoscopic Content(3D)
01	Mono/stereoscopic Mixed Contents(2D+3D)
10~11	reserved

[0025] 2. 'sovf'(Stereoscopic Object Visual Format)

[0026] 'sovf'는 스테레오스코픽 콘텐츠의 영상 구성 포맷(또는 비주얼 포맷)을 나타낸다. 표 3은 'sovf'의 선택스를 나타낸다. 또한, 'sovf'내 포함된 'Stereoscoic_Object_VisualFormat'은 기존의 다른 박스 또는 스테레오스코픽 콘텐츠를 저장할 위하여 새롭게 정의된 박스 내에 포함될 수 있다.

표 3

[0027]

```
Aligned(8) class StereoscopicObjectDataVisualFormat extend Box('sovf') {
    unsigned int(4)    Stereoscopic_Object_VisualFormat;
    unsigned int(4)    reserved;
}
```

[0028] 표 3에서 'Stereoscopic_Object_VisualFormat'은 스테레오스코픽 콘텐츠의 영상 구성 정보를 나타내며, 표 4와 같은 실시예를 갖는다.

표 4

[0029]

Value	Description
0000	Side by Side
0001	Top-Down
0010	Field Sequential based on horizontal
0011	Field Sequential based on vertical
0100	Frame Sequential
0101-0111	Reserved
1000	Main + additional (full size)
1001	Main + additional (half of vertical)
1010	Main + additional (half of horizontal)
1011	Main + additional (half of vertical/horizontal)
1100	Main + depth Map

[0030] 'full size'는 부가 영상의 사이즈가 주요 영상의 사이즈와 동일한 것을 의미하며, 'half of vertical'은 부가 영상의 사이즈가 주요 영상의 세로 반절 사이즈임을 의미하고, 'half of horizontal'은 부가 영상의 사이즈가 주요 영상의 가로 반절 사이즈임을 의미하며, 'half of vertical/horizontal'은 부가 영상의 사이즈가 주요 영상의 가로/세로 반절 사이즈임을 의미한다.

[0031] 3. 'ssci'(Stereoscopic Content Information)

[0032] 'ssci'는 스테레오스코픽 콘텐츠가 포함하는 최소/최대 깊이(depth) 또는 디스패리티(disparity) 정보로서, 이를 통하여 3D 단말은 3D 디스플레이에 적합한 입체 영상을 재생할 수 있다. 표 5는 'ssci'의 신택스를 나타낸다. 또한, 'ssci'내 포함된 최소/최대 깊이 정보는 기존의 다른 박스 또는 스테레오스코픽 콘텐츠를 저장할 위하여 새롭게 정의된 박스 내에 포함될 수 있다. 표 5에서 Max_of_depth(disparity)는 최대 깊이/디스패리티 정보(픽셀단위)를 나타내고, Min_of_depth(disparity)는 최소 깊이/디스패리티 정보(픽셀단위)를 나타낸다.

표 5

[0033]

```
Aligned(8) class StereoscopicObjectContentInformation extend Box('ssci') {
    unsigned int(32) Max_of_depth or disparity;
    unsigned int(32) Min_of_depth or disparity;
}
```

[0034] 4. 'scpi'(Stereoscopic Camera Parameter Information)

[0035] 'scpi'는 스테레오스코픽 카메라 또는 저작도구에 의하여 획득/생성된 스테레오스코픽 콘텐츠의 카메라 파라미터 정보를 나타낸다. 표 6은 'scpi'의 신택스를 나타낸다. 또한, 'scpi'내 포함된 각 필드는 기존의 다른 박스 또는 스테레오스코픽 콘텐츠를 저장할 위하여 새롭게 정의된 박스 내에 포함될 수 있다.

표 6

[0036]

```
Aligned(8) class StereoscopicCameraParameterInformation extend Box('scpi') {
    unsigned int(32) Baseline;
    unsigned int(32) Focal_Length;
    unsigned int(32) Convergence_point_distnace;
    unsigned int(1) StereoscopicCamera_setting;
    unsigned int(7) Reserved;
}
```

[0037] 표 6에서 'Baseline'는 좌우 카메라 사이의 거리를 나타내고, 'Focal_Length'는 카메라 중심에서 이미지 면(image plane)(CCD 센서)까지의 거리를 나타내며, 'Convergence_point_distance'는 베이스라인(baseline)으로부터 수렴점(convergence point)까지의 거리를 나타낸다. 여기서, 수렴점은 좌우 카메라의 시선이 교차하는 지점을 가리킨다. 그리고 'StereoscopicCamera_setting'는 스테레오스코픽 촬영/데이터의 카메라 배치(camera arrangement)를 나타내며 표 7과 같은 의미를 갖는다.

표 7

[0038]

Value	Description
0	Parallel arrangement
1	Cross arrangement

[0039] 5. 'iods' (Object Descriptor Box)

[0040] 'iods'는 장면(Scene) 정보(BIFS or LAsEr)가 있을 경우, BIFS 스트림 및 OD 스트림 위치를 나타내기 위한 IOD(Initial Object Descriptor)를 표현하기 위한 정보를 나타낸다. 표 8은 'iods'의 선택스를 나타낸다.

표 8

[0041]

```
Aligned(8) class ObjectDescriptoratBox extend Box('iods') {
    ObjectDescriptor OD;
}
```

[0042] 6. 'soet' (Stereoscopic One es Type Box)

[0043] 'soet'는 인코더로부터 출력된 ES가 1개인 경우를 나타낸다. 표 9는 'soet'의 선택스를 나타낸다.

표 9

[0044]

```
Aligned(8) class StereoscopicOnesTypeBox extend Box('soet') {
    unsigned int(3) Stereoscopic_OneES_Type:
    unsigned int(5) Reserved;
}
```

[0045] 표 9에서 'Stereoscopic_OneES_Type'은 ES가 1개로 구성된 스테레오스코픽 데이터의 영상 구성 포맷의 실시예를 나타내며, 표 10과 같은 의미를 갖는다.

표 10

[0046]

Value	Description
000	Side by Side
001	Top-Down
010	Field Sequential based on horizontal
011	Field Sequential based on vertical
100	Frame Sequential
101	Mono/stereoscopic Mixed data
110	기준영상/부가데이터 기반으로 구성된 스테레오스코픽 데이터
111	Reserved

[0047] 7. 'stet' (StereoScopic Two es Type Box)

[0048] 'stet'는 인코더로부터 출력되는 ES가 2개일 때 각 ES의 타입을 나타낸다. 표 11은 'stet'의 선택스를 나타낸다.

표 11

[0049]

```
Aligned(8) class StereoscopicTwoesTypeBox extend Box('stet') {
    unsigned int(2) Stereoscopic_TwoES_Type:
}
```

[0050] 표 11에서 'Stereoscopic_TwoES_Type'는 2개의 ES로 구성된 스테레오스코픽 데이터의 영상 구성 포맷의 실시예를 나타내며, 표 12과 같은 의미를 갖는다.

표 12

Value	Description
000	기준영상
001	부가데이터(기준영상 기준 동일 사이즈 영상)
010	부가데이터(기준영상 기준 가로반절 사이즈 영상)
011	부가데이터(기준영상 기준 세로반절 사이즈 영상)
100	부가데이터(기준영상 기준 가로/세로반절 사이즈 영상)
101	부가데이터(Depth map)
110	부가데이터(Disparity map)
111	Reserved

8. 'sstt' (StereoScopic Time Table Box)

[0052] 삭제

[0053] 'sstt'는 모노스코픽 콘텐츠 및 스테레오스코픽 콘텐츠가 혼용되어 사용될 경우, 장면 내에서 모노스코픽 콘텐츠 및 스테레오스코픽 콘텐츠의 구분 시작 및 종료 정보를 나타낸다. 'sstt'의 제1 실시예에 대한 선택스가 표 13에 표시된다. 또한, 아래에 설명되는 'sstt'내 포함된 각 필드는 기존의 다른 박스 또는 스테레오스코픽 콘텐츠를 저장할 위하여 새롭게 정의된 박스 내에 포함될 수 있다

표 13

[0054]

```

Aligned(8) class StereoscopicTimeTable Box extend Box('sstt') {
    int i;
    unsigned int(8) Mono/stereoscopic_Scene_count;
    for(i=0; i<=Mono/stereoscopic_Scene_count; i++)
    {
        unsigned int(4) Mono/stereoscopic_identifier;
        unsigned int(4) Reserved;
        unsigned int(32) Start_Time;
    }
}

```

[0055] 표 13에서 'Mono/stereoscopic_Scene_count'는 2차원 콘텐츠 및 3차원 콘텐츠가 혼용되어 구성되었을 때 Mono/stereoscopic 장면 변화의 수를 나타낸다. 즉, 2D->3D->2D로 구성되는 스테레오스코픽 콘텐츠의 경우에 Mono/stereoscopic_Scene_count=2로 설정된다. 또한, 2차원 콘텐츠 및 3차원 콘텐츠가 혼용되지 않고 3차원 콘텐츠로만 구성되는 경우에 Mono/stereoscopic_Scene_count=1로 설정된다. 이러한 정보는 3차원 단말에서 2D/3D 디스플레이 자동 변환에 사용될 수 있다.

[0056] 'Mono/stereoscopic_identifier'는 시간별 콘텐츠 타입을 나타내고, 표 14와 같은 의미를 갖는다. 또한, 'Mono/stereoscopic_identifier'는 2차원 콘텐츠 또는 3차원 콘텐츠를 구분하는데 사용될 수 있다. 예를 들어, 'Mono/stereoscopic_identifier'는 1 비트가 할당되어 '0'이면 2D, '1'이면 3D로 콘텐츠를 구분하는데 사용될 수 있다. 'Start_Time'는 시간별 콘텐츠 타입 시작시간을 나타낸다.

표 14

[0057]

Value	Description
0000	Monoscopic
0001	Side by Side
0010	Top-Down
0011	Field Sequential based on horizontal
0100	Field Sequential based on vertical
0101	Frame Sequential
0110-0111	Reserved
1000	Main + additional (full size)
1001	Main + additional (half of vertical)
1010	Main + additional (half of horizontal)
1011	Main + additional (half of vertical/horizontal)
1100	Main + Depth Map
1101	Main + Disparity Map

[0058] 표 15는 'sstt'의 제2 실시예에 대한 신택스를 표시하고 있다. 표 15에서 'Start_Time'은 스테레오스코픽 콘텐츠의 시작 시간을 나타내며, 'End_Time'은 스테레오스코픽 콘텐츠의 종료 시간을 나타낸다.

표 15

[0059]

```
Aligned(8) class StereoscopicTimeTable Box extend Box('sstt') {
    int i;
    unsigned int(8) Stereoscopic_Scene_count;
    for(i=0;i<=Stereoscopic_Scene_count;i++)
    {
        unsigned int(32) Start_Time;
        unsigned int(32) End_Time;
    }
}
```

표 16은 'sstt'의 제3 실시예에 대한 신택스를 표시하고 있다. 'Start_Sample_number'는 시간별 Mono/Stereoscopic 콘텐츠 시작 샘플의 번호 또는 샘플의 개수를 나타낸다. 즉, 샘플의 개수는 Monoscopic 또는 Stereoscopic에 해당하는 전체 샘플의 개수를 의미한다. 여기서, 샘플(sample)은 비디오, 시간 연속적인 일련의 비디오 프레임들의 각각의 프레임(an individual frame of video, a time-contiguous series of video frames)을 의미한다.

[0060] 삭제

표 16

[0061]

```
Aligned(8) class StereoscopicTimeTable Box extend Box('sstt') {
    int i;
    unsigned int(16) Mono/stereoscopic_Scene_count;
    for(i=0;i<=Mono/stereoscopic_Scene_count;i++)
    {
        unsigned int(4) Mono/stereoscopic_identifier;
        unsigned int(32) Start_Sample_number;
    }
}
```

[0062] 표 17은 'sstt'의 제3 실시예에 대한 선택스를 표시하고 있다.

표 17

[0063]

```
Aligned(8) class StereoscopicTimeTable Box extend Box('sstt') {
    int i;
    unsigned int(8) Stereoscopic_Scene_count;
    for(i=0;i<=Stereoscopic_Scene_count;i++)
    {
        unsinged int(4) Stereoscopic_compositiontype;
        unsigned int(32) Start_sample_number;
        unsigned int(32) End_sample_number;
    }
}
```

[0064] 표 17에서 'Stereoscopic_compositiontype'는 다양한 스테레오스코픽 구성 타입으로 구성되는 경우 시간별 콘텐츠 타입을 나타내고, 표 18과 같은 의미를 갖는다. 'End_Sample_number'는 시간별 스테레오스코픽 콘텐츠 종료 샘플 번호 또는 샘플의 개수를 나타낸다.

표 18

[0065]

Value	Description
0000	Side by Side
0001	Top-Down
0010	Field Sequential based on horizontal
0011	Field Sequential based on vertical
0100	Frame Sequential
0101-0111	Reserved
1000	Main + additional (full size)
1001	Main + additional (half of vertical)
1010	Main + additional (half of horizontal)
1011	Main + additional (half of vertical/horizontal)
1100	Main + Depth Map
1101	Main + Disparity Map

[0066] 9. 'sesn' (Stereoscopic ES Num Box)

[0067] 'sesn'는 인코더로부터 출력되는 기본 부호화 스트림의 개수를 나타낸다. 표 19는 'sesn'에 대한 선택스를 나타낸다. 표 19에서 'stereoscopic_ESNum'는 스테레오스코픽 콘텐츠에 대하여 인코딩으로부터 출력되는 ES의 개수를 나타낸다.

표 19

[0068]

```
Aligned(8) class StereoscopicESNumBox extend Box('seen') {
    unsigned int(16) stereoscopic_ESNum;
}
```

[0069] 10. 'tref' (Track reference box)

[0070] 'tref'는 하나의 트랙에서 다른 트랙을 참조할 수 있도록 정보를 제공하는 ISO 기반 파일 포맷에 정의된 박스이다. 'tref'는 'trak'(Track Box)에 포함된다. 표 20은 'tref'의 일 실시예에 대한 신택스를 표시하고 있다. 여기서, 'track_ID'는 참조되는 트랙의 식별부호를 나타내며, 'reference_type'은 표 21과 같은 의미를 갖는다.

표 20

[0071]

```
aligned(8) class TrackReferenceBox extends Box ('tref') {}
aligned(8) class TrackReferenceTypeBox (unsigned int(32) reference_type) extends
Box(reference_type) {
    unsigned int(32) track_IDs[];
}
```

표 21

[0072]	hint	the referenced track(s) contain the original media for this hint track.
	cdsc	this track describes the referenced track.
	svdp	this track has a dependency for the referenced track as its primary view track, and it can contain the stereoscopic related 'meta' information for the referenced track.

[0073] 기본 부호화 스트림(Elementary Stream: ES)이 2개인 스테레오스코픽 비디오는 2개의 트랙이 존재하며, 종래의 LASEr와 같은 장면기술로서는 LASEr 내 비디오 관련 2개의 노드와 연결되어야 한다. 즉, 종래의 기술로는 ES가 2개인 스테레오스코픽 비디오는 2개의 객체로 인식되는 것이다. 하지만, 스테레오스코픽 비디오는 단말에서 최종적으로 하나의 3D 비디오 포맷으로 변환되어 재생되므로 하나의 객체로 인식되어야 한다. 즉, 스테레오스코픽 비디오는 2개의 트랙을 이용하여 구성되나, 이를 장면 재생하기 위해서는 하나의 3D 비디오 포맷으로 변환되어 재생되므로 하나의 노드만을 이용하여 연결되어야 한다. 이에 기본 부호화 스트림이 2개인 경우 두개의 트랙의 관계성을 나타내는 정보가 요구되며 이는 표 22와 같이 'tref' 내에 'svdp'에 스테레오스코픽 트랙 참조 정보를 정의하여 사용한다. LASEr가 사용되지 않더라도 ES가 2개인 스테레오스코픽 콘텐츠를 저장하기 위해서는 'svdp'와 같은 스테레오스코픽 트랙 참조 정보를 사용해야 한다.

[0074] 이 박스를 이용하여, 스테레오스코픽 콘텐츠를 구성함에 있어 ES가 2개인 경우에는, 2개의 트랙이 하나의 주요 트랙(main track)과 하나의 부가 트랙(additional track)으로 구분될 수 있다. 또한, 부가 트랙 내의 'tref' 박스를 이용하여 주요 트랙을 참조하고 관련 하나의 트랙에만 스테레오스코픽 비디오 관련 스테레오스코픽 카메라 및 디스플레이 정보를 저장함으로써 저장 정보의 중복성이 제거될 수 있다. 또한, 2개의 트랙을 가지더라도 LASEr의 비디오 노드의 1개에 연결될 수 있는 장점을 가진다.

[0075] 스테레오스코픽 콘텐츠가 다양한 스테레오스코픽 영상 구성 포맷 및 카메라 파라메타를 지원하는 경우 이를 위하여 본 발명에서 초기 스테레오스코픽 헤더(initial stereoscopic header: 'ishd')의 구조를 제안한다. 이하에서 실시예로서 'ishd'에 포함되는 정보는 별도로 분리되어 사용될 수 있다.

[0076] 스테레오스코픽 스트림 내 다양한 스테레오스코픽 포맷 및 카메라 파라메타가 존재하는 경우, 각 스테레오스코픽 스트림, 시작 및 길이의 정보는 'iloc'를 통하여 구분되며 각 구간(fragment)에 item_ID가 할당된다. 이에, 'ishd'는 item_ID를 기반으로 각각의 스테레오스코픽 포맷 또는 카메라 파라메타 정보를 제공해야 한다. 여기서, 아이템(item)은 하나의 시퀀스로 스테레오스코픽 구간과 모노스코픽 구간이 함께 존재할 때 하나의 스테레오스코픽 구간을 의미한다.

[0077] 또한, 스테레오스코픽 스트림 내에 3개의 스테레오스코픽 구간이 있을 경우, 각 구간에 적용된 'ishd' 정보가 다른 경우는 item_ID로 구분되어 각각의 정보를 기술한다. 하지만 각 구간에 적용된 'ishd' 정보가 같을 경우는 2, 3번째에 해당하는 스테레오스코픽 구간이 1번째에 해당하는 'ishd' 정보와 같다는 필드가 포함되어야 한다. 이러한 구조는 'ishd' 내에 포함되는 정보의 중복성을 제거할 수 장점이 있다.

[0078] 도 4는 본 발명에 따른 SS-VAF의 일 실시예의 구조도로서, MPEG-4 시스템 정보가 있는 경우이다. 도 5는 본 발명에 따른 SS-VAF의 다른 실시예의 구조도로서 MPEG-4 시스템 정보가 없는 경우이다.

- [0079] 한편, 모노스코픽 및 스테레오스코픽 콘텐츠가 혼용되어 있을 경우, 모노스코픽 또는 스테레오스코픽 콘텐츠가 언제 시작하고 끝나야 하는지에 대한 구분정보가 필요하다. 이는 모노/스테레오스코픽 콘텐츠 시작/종료 정보 또는 샘플의 2D/3D 구분 및 포함되는 Sample(AU)의 개수에 따라 구분될 수 있다.
- [0080] 도 6 내지 도 8은 도 1과 같은 다양한 콘텐츠 구성 형태를 지원하기 위한 저장 포맷의 개념적인 구조를 제공한다. 기본적으로 저장 포맷의 구조는 크게 'ftyp', 'moov' 및 'mdat'로 구성된다. 'ftyp'는 파일 타입을 정의하는 박스로서 스테레오스코픽 또는 모노스코픽/스테레오스코픽 혼용 콘텐츠 파일임을 나타내는 필드를 추가하여 3차원 콘텐츠 파일임을 나타낸다. 'moov'은 미디어 데이터를 재생하기 위한 모든 시스템(메타) 정보를 포함하며, 'mdat'은 실제 미디어 데이터를 포함한다. 이러한 구조를 바탕으로 스테레오스코픽 콘텐츠에 대한 새로운 부가정보가 요구되며, 이러한 부가정보의 위치에 따라 저장 포맷 구조는 달라진다.
- [0081] 도 6은 3차원 콘텐츠를 구성하는 소스의 개수 및 새로운 부가정보를 포함하는 'ishd'를 'moov'에 포함시킨 저장 포맷 구조를 나타낸다. 이때 3차원 콘텐츠에 대하여 소스가 1개인 경우(도 6(a))는 한 프레임 내에 좌우영상의 정보가 포함된 형태(예시: side by side 등)를 의미하며, 소스가 2개인 경우(도 6(b))는 좌우 영상의 정보가 각각 하나의 프레임으로 포함된 형태를 의미한다. 한편, 'moov' 내의 트랙의 개수는 포함된 미디어 데이터의 개수에 따라 달라진다. 'moov'내의 트랙은 'mdat'에 포함된 각각의 미디어 데이터를 재생하기 위한 모든 시스템(메타)정보를 포함한다.
- [0082] 이러한 기본적인 구조 아래 새로운 부가정보 및 이를 지원하기 위한 구조가 반드시 요구된다. 이에 본 발명은 'ishd' 새롭게 정의하고 이를 'moov'내 트랙 이전에 포함시킨다. 이러한 'ishd'의 위치는 'moov'내 또는 저장 포맷 내에서 달라질 수 있다.
- [0083] 도 7은 앞에서 정의된 'ishd' 정보를 'mdat'에 포함시킨 저장 포맷 구조를 나타낸다. 도 7(a)는 하나의 소스로 구성된 3차원 콘텐츠의 경우이고, 도 7(b)는 두개의 소스로 구성된 3차원 콘텐츠의 경우이다. 이는 기존의 'moov' 구조를 그대로 유지한 채 저장 포맷 구조 내 'ishd' 스트림이 'mdat'내 포함되어 있다는 정보만을 추가함으로써 가능하다.
- [0084] 도 8은 앞에서 정의된 'ishd' 정보를 'meta'에 포함시킨 저장 포맷 구조를 나타낸다. 도 8(a)는 하나의 소스로 구성된 3차원 콘텐츠의 경우이고, 도 8(b)는 두개의 소스로 구성된 3차원 콘텐츠의 경우이다.
- [0085] 표 22는 'ishd' 정보가 'mdat'에 포함되어 있다는 정보를 알려주기 위한 구조를 나타내며, 이러한 구조는 기존의 'stsd(sample description box)'내에 포함된다.

표 22

- [0086]

```
//ishd Sequences
Class ishdSampleEntry(codingname) extends SampleEntry ('ishd') { }
```
- [0087] 도 9는 도 6 내지 8의 구조를 바탕으로 장면기술자 정보가 포함되는 경우의 저장 포맷 구조를 나타낸다. 장면기술자는 다양한 멀티미디어의 자유로운 장면 구성과 사용자와의 인터랙션을 위한 것으로서 본 발명에서는 레이저(LASER)를 사용한다.
- [0088] 도 9의 (a)는 'moov' 내 장면기술자 스트림을 저장하기 위해 별도의 박스를 추가하는 경우를 나타낸다. 도 9의 (b)는 'mdat' 내 장면기술자 스트림을 추가하고, 'moov'내 이를 알려주기 위한 트랙을 별도로 추가하는 경우로서, 기존의 'stsd' 내에 장면기술자 스트림에 대한 정보를 추가한다. 즉, 트랙 내의 'stsd'를 검색하여 이 트랙이 어떤 정보(장면기술자/비디오/오디오)를 나타낸 것인가를 분석하고, 이를 토대로 'mdat'에 저장된 정보를 이용하여 디코딩하는 방법이다. 도 9의 (c)는 앞에서 정의된 장면기술자 정보를 'meta'에 포함시킨 저장 포맷 구조를 나타낸다.
- [0089] 표 23 내지 표 25는 도 1의 3차원 콘텐츠 구성 형태를 모두 지원하는 'ishd' 구조의 실시예들을 나타낸다.

표 23

[0090]

```
Class ishd {

    unsigned int(16) num_MonoStereo_scene;
    if(num_MonoStereo_scene){
        for(i=0;i<num_MonoStereo_scene;i++)

            unsigned int(16) start_sample_index;
            unsigned int(3) Composition_type;
            unsigned int(1) numofES;
            unsigned int(1) LR_first;
            unsigned int(3) Reserved;
        }
    } else

    unsigned int(3) Composition_type;
    unsigned int(1) numofES;
    unsigned int(1) LR_first;
    stereoscopicCameraInfo[0...1];
    stereoscopicContentsInfo[0...1];
}
```

표 24

[0091]

```
Class ishd {

    unsigned int(16) num_MonoStereo_scene;
    if(num_MonoStereo_scene){
        for(i=0;i<num_MonoStereo_scene;i++)

            unsigned int(16) start_sample_index;
            unsigned int(3) Composition_type;
            unsigned int(1) numofES;
            unsigned int(1) LR_first;
            unsigned int(3) Reserved;
        }
    } else

    unsigned int(3) Composition_type;
    unsigned int(1) numofES;
    unsigned int(1) LR_first;
}
```

표 25

[0092]

```

Class ishd {

    unsigned int(16) num_MonoStereo_scene;
    if(num_MonoStereo_scene){
        for(i=0;i<num_MonoStereo_scene;i++)

            unsigned int(16) start_sample_index;
            unsigned int(3) Composition_type;
            unsigned int(1) numofES;
            unsigned int(1) LR_first;
            stereoscopicCameraInfo[0...1];
            stereoscopicContentsInfo[0...1];
            unsigned int(3) Reserved;
        }
    } else

        unsigned int(3) Composition_type;
        unsigned int(1) numofES;
        unsigned int(1) LR_first;
        stereoscopicCameraInfo[0...1];
        stereoscopicContentsInfo[0...1];
    }
}

```

[0093] 표 23 내지 표 25에서 'num_MonoStereo_scene'은 2차원 콘텐츠 및 3차원 콘텐츠가 혼용되어 구성된 경우 장면의 개수를 나타내거나, 또한 다양한 3차원 콘텐츠가 혼용되어 구성된 경우 장면의 개수를 나타낸다. 예를 들어, 3차원 콘텐츠가 (2D)(3D)(2D)로 구성된 경우 Num_MonoStereo_scene=3이 된다. 또한 3차원 콘텐츠가 (side by side)(field sequential) 방식으로 구성된 경우에는 num_MonoStereo_scene=2가 된다. 이때 단일 포맷의 3차원 콘텐츠로만 이루어질 경우 num_MonoStereo_scene=1이 된다.

[0094] 'Start_sample_index'는 각 콘텐츠의 시작 샘플(즉, 일반적인 프레임) 번호 또는 각 콘텐츠 타입에 따른 포함되는 샘플의 개수로 사용될 수 있다. 'numofES'는 'mdat'에 포함된 비디오 부호화 스트림의 개수를 나타낸다.

[0095] 'Composition_type'은 2차원 콘텐츠 및 3차원 콘텐츠에 대하여 각각의 콘텐츠 포맷을 구별하기 위한 정보를 나타낸다. 이러한 'Start_sample_index' 및 'Composition_type' 정보는 2D/3D 디스플레이 모드를 지원하는 다양한 3차원 단말기에서 자동 디스플레이 온(on)/오프(off)를 하기 위한 기본 정보로서 활용될 수 있다. 'Composition_type'은 표 26과 같은 의미를 갖는다.

표 26

[0096]

Value		Description
3D	000	Side by Side
	001	Top-down
	010	Field sequential
	011	Frame sequential
	100	Vertical line interleaved format
	101	Left/right image
2D	110	Monoscopic left image
	111	Monoscopic right image

'LR_first'는 좌우영상 중 우선시 되는 영상을 나타낸 것으로서, 왼쪽 이미지와 오른쪽 이미지 중에서 먼저 인코딩되는 이미지를 표시한다.

[0097] 삭제

[0098] 'stereoscopicCameraInfo'는 3차원 콘텐츠에 대한 카메라 파라미터 정보를 나타낸 것으로서, 표 27은 'stereoscopicCameraInfo' 구성 형태의 일 실시예를 나타내며, 본 발명에서 제시한 카메라 파라미터 정보는 기존의 다른 박스 내 또는 새로게 정의된 박스에 포함될 수 있다. 표 27에서, 'StereoscopicCamera_setting'은 3차원 콘텐츠를 제작 또는 촬영할 때의 카메라의 배열 형태를 나타낸 것으로 "평행식(parallal)" 또는 "교차식(cross)"를 나타낸다. 'Baseline'은 스테레오스코픽 카메라 사이의 거리를 나타내고, 'Focal_Length'는 렌즈로부터 이미지 면(image plane)까지의 거리를 나타낸다. 또한, 'ConvergencePoint_distance'는 좌우 카메라를 연결하는 베이스라인(baseline)으로부터 수렴점(convergence point)까지의 거리를 나타낸다.

표 27

[0099]

```
StereoscopicCameraInfo {
    unsigned int(1) StereoscopicCamera_setting;
    unsigned int(7) Reserved=1111;
    unsigned int(16) Baseline;
    unsigned int(16) Focal_Length;
    unsigned int(16) ConvergencePoint_distance;
}
```

[0100] 'StereoscopicContentsInfo'는 3차원 콘텐츠에 대하여 디스플레이를 위한 최소한의 정보를 나타낸다. 표 28은 'StereoscopicContentsInfo' 구성 형태의 일 실시예를 나타내며, 포함되는 정보는 기존의 다른 박스 내 또는 새롭게 정의된 박스 내에 포함될 수 있다. 'Max_disparity'는 3차원 콘텐츠의 최대 디스패리티 크기를 나타내고, 'Min_disparity'는 3차원 콘텐츠의 최소 디스패리티 크기를 나타낸다.

표 28

[0101]

```
StereoscopicContentsinfo {
    unsigned int(16) Max_disparity;
    unsigned int(16) Min_disparity;
}
```

[0102] 'StereoscopicCameraInfo' 및 'StereoscopicContentsInfo' 내 정보들은 별도의 기술(예: MPEG-7 메타데이터)로 표현되어 저장될 수 있다.

[0103] 도 10은 ES=1일 때 SS-VAF의 일 실시예의 구성도이다.

[0104] 'ftyp'는 스테레오스코픽 콘텐츠 포함 여부 판별로서, 부호화 스트림이 전체 3D인 경우, 부호화 스트림이 2D/3D 혼용의 스트림으로 구성된 경우를 모두 스테레오스코픽 콘텐츠로 한다.

[0105] 스테레오스코픽 콘텐츠가 2D/3D 혼용의 스트림으로 구성된 경우, 스트림 상에서 시작 및 길이 정보를 알아야 한다. 이를 위해 ISO 기반 파일 포맷(14496-12)의 기존 박스인 'iloc'를 이용한다. 'iloc'(item location box)는 'meta' 박스에 포함된다. 'iloc'는 스테레오스코픽 콘텐츠의 경우, 저장 파일 내의 스테레오스코픽 구간의 위치를 제공한다.

[0106] 또한, 2D 또는 3D 스트림의 구분 등 관련 정보는 'ishd' 박스를 통하여 획득한다. 2D/3D 혼용 스트림의 경우 3D 스트림이 다수 개라도, 각각 동일한 정보라면(즉, 단일 포맷이라면) 한번의 'ishd' 정보만을 참조하여 3D 스트림의 관련 정보를 획득한다.

[0107] 단일 포맷의 3D 스트림으로만 구성된 경우는 'iloc'의 사용 없이 'ishd'만 사용하여 표현할 수 있다. 또한 ES=1

일 때 3D 스트림이 다중 포맷으로 구성된 경우는 'iloc'를 이용하여 각 포맷에 해당되는 오프셋(offset)/길이(length)값을 알고 각 포맷 정보는 'ishd' 박스를 통하여 획득한다. 이 경우 'ishd'는 다중 포맷에 해당되는 정보를 가지고 있다.

[0108] 도 11은 ES=2일 때 SS-VAF의 일 실시예의 구성도이다. ES=2인 경우, 좌우 스트림 정보는 각각의 'trak'에 포함되며 이러한 경우 전술한 바와 같이 3차원 재생 시 특정된 하나의 포맷으로 변환되어 디스플레이 되므로, 2개의 트랙으로 구성되지만 하나의 객체로 인식되기 위해 좌우 스트림은 서로 연관관계를 나타낼 필요가 있다. 예를 들어, 좌영상이 메인영상이고 우영상이 부가영상이라면, 우영상 스트림 정보를 포함하는 'trak'이 좌영상 스트림 정보를 포함하는 'trak'과의 관계성을 표현하면 이를 토대로 'ishd' 내용 중에서 중복된 것을 제거할 수 있다. 우영상 스트림 정보에 포함되는 'ishd' 중 카메라 파라미터 및 디스플레이 정보가 좌영상 스트림 정보에 포함되는 'ishd'와 동일하다면 별도로 기술하지 않고 좌영상 스트림 정보에 포함되는 'ishd' 내용을 따를 수 있다. 이러한 관계를 표현하기 위해서 본 발명에서는 'tref' 내의 'svdp'를 제안한다.

[0109] 좌우 스트림으로 저장된 3D 스트림 상에서 다중 포맷의 3D 스트림으로 구성된 경우는 각 포맷에 해당되는 3D 스트림의 구분, 시작 및 길이 정보를 알아야 한다. 이는 ISO 기반 파일 포맷 (14496-12)의 기존 박스인 'iloc'를 통하여 시작 및 길이 정보를 획득하고, 또한, 다중 포맷의 3D 스트림의 구분 등 관련 정보는 'ishd' 박스를 통하여 획득한다. 표 29는 단일 포맷의 'ishd'의 선택스 실시예를 나타낸다.

표 29

[0110] `Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)`

```

unsigned int(8) Stereoscopic_Composition_Type;
unsigned int(1) LR_first;
unsigned int(1) Is_camParams;
unsigned int(1) Is_disInfo;
unsigned int(5) Reserved;

//all the following are optional fields
if(Is_CamParams){
    unsigned int(32) Baseline;
    unsigned int(32) focallength;
    unsigned int(32) convergence_distance;
    unsigned int(1) Is_camera_cross
}

if(Is_camera_cross){
    unsigned int(32) rotation[];
    unsigned int(32) translation[];
    unsigned int(7) reserved;
}

if(Is_disInfo){
    Int(16) MinofDisparity;
    Int(16) MaxofDisparity;
}

```

[0111] 'Is_camParams'는 카메라 파라미터의 존재 여부 판단, 'Is_disInfo'는 스테레오스코픽 콘텐츠 디스플레이 정보 여부 판단, 'Baseline'는 좌우 카메라의 거리, 'focallength'는 렌즈에서 이미지 면(필름)까지의 거리, 'convergence_distance'는 좌우 카메라를 연결하는 베이스라인(baseline)의 중심으로부터 좌우 카메라의 시선이 교차하는 수렴점(convergence point)까지의 거리를 나타낸다. 'convergence_distance'는 평행축 카메라의 경우 무한대의 값을 가지며, 이를 표시하기 위해 모든 비트를 1로 한다.

- [0112] 또한, 'Is_camera_cross'는 '1'일 때 교차축 카메라, '0'일 때 평행축 카메라를 나타내고, 'rotation'은 객체까지의 카메라 포지션 각도(camera position angle to an object)를 나타낸다. 'translation'은 스테레오 카메라의 이동의 존재(모든 비트가 0인 경우 스테레오 카메라의 움직임이 없음), 'MinofDisparity'는 좌우 이미지의 최소 디스패리티 크기, 'MaxofDisparity'는 좌우 이미지의 최대 디스패리티 크기를 나타낸다.
- [0113] 표 30은 다중 포맷의 'ishd'의 선택스 실시예를 나타낸다. 'Item_count'는 다중 포맷의 경우 포맷의 정보를 기술하는 개수를 나타내고, 'Item_ID'는 각 포맷의 정수 이름(integer name)을 나타내며, 스테레오스코픽 콘텐츠 내 다양한 스테레오스코픽 포맷에 대한 구분을 위하여 사용된다. 이때 'Item_ID'는 'iloc'의 item_ID와 연동되어 사용된다.

표 30

[0114]

```

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

unsigned int(16)      item_count;
for (i=0; i<item_count; i++) {
    unsigned int(16) item_ID;
    unsigned int(8)    Stereoscopic_Composition_Type;
    unsigned int(1)    LR_first;
    unsigned int(1)    Is_camParams;
    unsigned int(1)    Is_disInfo;
    unsigned int(5)    Reserved;

    //all the following are optional fields
    if(Is_CamParams){
        unsigned int(32) baseline;
        unsigned int(32) focallength;
        unsigned int(32) convergence_distance;
        unsigned int(1)  Is_camera_cross
    }

    if(Is_camera_cross) {
        unsigned int(32) rotation[];
        unsigned int(32) translation[];
        unsigned int(7) reserved;
    }

    if(Is_disInfo){
        Int(16)  MinofDisparity;
        Int(16)  MaxofDisparity;
    }

    //other additional information
    if(other_flag){}
}

```

- [0115] 이하, AVC (Advanced Video Coding) SEI (Supplemental enhancement information)에 대하여 설명한다. SEI는 디코딩, 디스플레이 등에 연관된 메시지 정보를 포함하는 'stereo video information SEI' 정보를 포함하고 있으며, SEI 메시지는 AVC 스트림 내에 전송된다.
- [0116] 도 12는 NAL 유닛을 포함하는 단일 비디오 ES(single video elementary stream containing NAL units)의 흐름도로서, 도 12(a)는 AVC내 SEI의 기본 구조 중 'stereo video information SEI', 'reserved_sei_message'가 포함된 부분을 나타내고, 도 12(b)는 AVC 스트림내 SEI의 위치를 나타낸다. 표 31은 'Stereo video information

SEI' 메시지를 나타낸다.

표 31

[0117]

stereo_video_info(payloadSize){	C	Descriptor
field_views_flag	5	u(1)
if(field_views_flag)		
top_field_is_left_view_flag	5	u(1)
else {		
current_frame_is_left_view_flag	5	u(1)
next_frame_is_second_view_flag	5	u(1)
}		
left_view_self_contained_flag	5	u(1)
right_view_self_contained_flag	5	u(1)
}		

[0118]

'Field_views_flag'는 필드 기반 스테레오스코픽 스트림 존재 여부 판단을 나타낸다. 'Top_field_is_left_view_flag'는 '1'인 경우 vertical interlaced 포맷으로 구성된 스테레오스코픽 콘텐츠(left view 우선), '0'인 경우 vertical line interleaved 포맷으로 구성된 스테레오스코픽 콘텐츠(right view 우선)을 나타낸다. 'Current_frame_is_left_view_flag'는 '1'인 경우 현재 프레임이 left view를 나타내며 '0'인 경우 현재 프레임이 right view를 나타낸다. 'Next_frame_is_second_view_flag'는 '1'인 경우 현재 프레임과 다음 프레임으로 스테레오스코픽 영상을 구성하는 것을, '0'인 경우 현재 프레임과 이전 프레임으로 스테레오스코픽 영상을 구성하는 것을 나타낸다. 'Left_view_self_contained_flag'는 '1'인 경우 right view 상관성 없이 독립 스트림으로 코딩한 것을, '0'인 경우 right view의 상관성을 토대로 스트림을 코딩한 것을 나타낸다. 'Right_view_self_contained_flag'는 '1'인 경우 left view 상관성 없이 독립 스트림으로 코딩한 것을, '0'인 경우 left view의 상관성을 토대로 스트림을 코딩한 것을 나타낸다.

[0119]

'stereo video information SEI' 정보는 'stereoscopic_composition_type' 중에서 표 32의 포맷을 포함하고 있다. 단, 'side by side' 포맷은 지원하지 않는다.

표 32

[0120]

1	Vertical interleaved format
2	Frame sequential format
3	Field sequential format
4	Monoscopic left image
5	Monoscopic right image

[0121]

이하, 기존의 AVC SEI 정보를 활용한 스테레오스코픽 콘텐츠를 서비스하기 위한 방법 및 저장 포맷을 제안한다. 이는 스트레오스코픽 콘텐츠가 AVC를 통하여 부호화가 된 경우만 가능하다.

[0122]

기존의 'stereo video information SEI'를 토대로 각 스테레오스코픽 스트림에 요구되는 카메라 파라미터 및 디스플레이 정보를 추가하는 형태로서, 'reserved_sei_message'를 이용한 SS-VAF을 제안한다. 또한, 기존의 'stereo video information SEI'를 표 33과 같이 확장하여 사용할 수 있다. 'Side_by_side_flag'는 '1'인 경우 left view가 우선되어 한 프레임 내 좌우영상으로 구성되며, '0'인 경우 right view가 우선되어 한 프레임 내 좌우영상으로 구성된다. 표 33에서 "C"는 선택스의 카테고리를 의미하고, "u(1)"은 1 비트를 이용하는 "unsigned integer"를 의미한다.

표 33

[0123]

stereo_video_info(payloadSize){	C	Descriptor
field_views_flag	5	u(1)
if(field_views_flag)		
top_field_is_left_view_flag	5	u(1)
else {		
current_frame_is_left_view_flag	5	u(1)
next_frame_is_second_view_flag	5	u(1)
}		
else {		
left_view_self_contained_flag	5	u(1)
right_view_self_contained_flag	5	u(1)
}		
side_by_side_flag	5	u(1)
}		

[0124]

표 34는 AVC의 SEI 정보 중 'reserved_sei_message(payloadSize)'를 이용하여 스테레오스코픽 카메라 정보를 정의한다. 이때 포함되는 정보 외 다른 카메라 정보가 추가될 수 있으며 추가되는 정보는 별도로 사용될 수 있다. 이를 토대로 스테레오스코픽 콘텐츠 스트림에 대한 카메라 파라미터 정보를 획득한다.

표 34

[0125]

stereo_camera_info(payloadSize) {	C	Descriptor
if(Is_CamParams){		
baseline;	5	U(32)
focallength;	5	U(32)
convergence_distance;	5	U(32)
if(Is_camera_cross) {		
rotation[];	5	U(32)
translation[];	5	U(32)
}		
}		

[0126]

표 35는 AVC의 SEI 정보 중 'reserved_sei_message(payloadSize)'를 이용하여 스테레오스코픽 콘텐츠를 디스플레이 하기 위하여 요구되는 정보를 정의한다. 이를 토대로 스테레오스코픽 콘텐츠 디스플레이 값을 추출한다.

표 35

[0127]

stereo_display_info(payloadSize) {	C	Descriptor
if(Is_disInfo){		
MinofDisparity;	5	U(16)
MaxofDisparity;	5	U(16)
}		

[0128]

다만, 상기 내용을 토대로, 이를 하나의 SEI_message로 통합하여 스테레오 카메라 및 디스플레이 정보를 제공할 수 있다.

[0129]

도 13은 Stereo video information 및 Reserved SEI를 이용한 SS-VAF을 나타낸다. 도시된 어플리케이션 포맷에서 LAsE 포함은 선택적이다.

- [0130] 스테레오스코픽 콘텐츠가 2D/3D 혼용의 스트림으로 구성된 경우 'iloc'를 통하여 구분된 3D 스트림의 구간 동안은 AVC 스트림 SEI message 내 정의된 3D 스트림 정보('stereo video information SEI', 'stereo camera information SEI', 'stereo display information SEI')를 획득한다. 단일 포맷의 3D 스트림으로만 구성된 경우는 'iloc'의 사용 없이 AVC 스트림 SEI message 내 정의된 3D 스트림 정보('stereo video information SEI', 'stereo camera information SEI', 'stereo display information SEI')를 사용하여 표현할 수 있다.
- [0131] 도 14는 ES=2일 때 단일 포맷의 스테레오스코픽 스트림만으로 구성된 경우의 SS-VAF의 구성도이다. ES=2인 경우, 좌우 스트림 정보는 각각의 'trak'에 포함된다. 이때, 좌우 스트림은 서로 연관관계를 나타낼 필요가 있다. 예를 들어, 좌영상이 메인영상이고 우영상이 부가영상이라면, 우영상 스트림 정보를 포함하는 'trak'이 좌영상 스트림 정보를 포함하는 'trak'과의 관계성을 표현하면 이를 토대로 'ishd' 내용 중 중복된 것은 제거할 수 있다. 이러한 관계성은 ISO 기반 파일 포맷에 포함된 'tref'를 이용한다. 이는 재생에 필요한 모든 'trak'_ID를 기술할 수 있어 우영상(부가영상) 스트림에 'trak'내 'tref'를 포함하여 기술한다.
- [0132] 표 36은 다양한 스테레오스코픽 구성 포맷 및 카메라 파라미터를 지원하는 경우의 'ishd'의 선택스 실시예를 나타낸다. 'item_ID'는 다음 정보가 정의되는 아이템의 ID를 나타내며, 1 이상의 값을 갖는다. 'current_indicator'는 '1'인 경우 다음에 기술되는 정보의 유효성을 나타내고, '0'인 경우 다음에 기술되는 정보가 앞서 기술되는 정보와 동일함을 의미한다. 단, Item_ID=1일 경우는 이후 기술되는 정보가 없음을 의미한다. 즉, 단말은 'Current_indicator'를 보고 다음에 기술되는 정보의 유효성을 판단한 다음에 0일 경우는 앞서 기술된 'ishd' 정보와 동일함을 인식하게 된다.
- [0133] 또한, 'LR_first'는 좌영상 및 우영상의 기준 위치 선정, 'Is_camParams'는 카메라 파라미터의 존재 여부 판단, 'Is_displaySafeInfo'는 스테레오스코픽 콘텐츠 디스플레이 정보 여부 판단을 나타낸다. 또한 'Baseline'는 좌우 카메라의 거리, 'focallength'는 CCD에서 이미지 플레인(필름)까지의 거리를 나타낸다. 또한 'Is_camera_cross'는 '1'인 경우 교차축 카메라, '0'인 경우 평행축 카메라를 나타낸다.

표 36

[0134]

```

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

unsigned int(16) item_ID;
unsigned int(1)  current_indicator;
unsigned int(8)  Stereoscopic_Composition_Type;
unsigned int(1)  LR_first;
unsigned int(1)  Is_camParams;
unsigned int(1)  Is_displaySafeInfo;
unsigned int(4)  Reserved;

//all the following are optional fields
if(Is_CamParams){
    unsigned int(32) baseline;
    unsigned int(32)   focallength;
    unsigned int(1)  Is_camera_cross
    if(Is_camera_cross) {
        unsigned int(32)   convergence_distance;
        unsigned int(32)   rotation[];
        unsigned int(32) translation[];
        unsigned int(7)  reserved;
    }
}
if(Is_displaySafeInfo){
    unsigned int(16) ViewingDisatance;
    int(16)          VerticalDisparity;
    int(16)          MinofDisparity;
    int(16)          MaxofDisparity;
}
//other additional information
if(other_flag){}

aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0){
    unsigned int(16)   item_count;
    ishdEntry[item_count] ishds;
}

```

또한, 'convergence_distance'는 베이스라인의 중심으로부터 수렴점까지의 거리(평행축 카메라의 경우 무한대의 값을 가짐, 이에 모든 비트가 1인 경우 무한대의 거리를 표시), 'rotation'은 객체에 대한 카메라 포지션 각도, 'translation'은 스테레오 카메라의 이동이 있을 경우 존재(모든 비트가 0인 경우 스테레오카메라 움직임이 없음)를 나타낸다. 또한 'VerticalDisparity'는 좌우 이미지의 수직 디스패리티 크기, 'MinofDisparity'는 좌우 이미지의 최소 수평 디스패리티 크기, 'MaxofDisparity'는 좌우 이미지의 최대 디스패리티 크기, 'item_count'는 다음 어레이 내의 엔트리의 개수를 나타낸다.

[0135] 삭제

[0136] 표 37은 다양한 카메라 파라미터를 지원하는 경우의 'ishd'의 제1 실시예의 신택스를 나타낸다. 여기서, 스테레오스코픽 구성 포맷은 동일하다고 가정하며, 바로 전의 'ishd' 정보가 참조된다. 또한 'ishd'에 포함되는 카메라 파라미터 및 디스플레이 정보는 별도의 박스로 분리하여 저장될 수 있다.

표 37

[0137]

```

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)
unsigned int(16) item_ID;
unsigned int(8) Stereoscopic_Composition_Type;
unsigned int(1) LR_first;
unsigned int(1) current_indicator;
unsigned int(1) Is_camParams;
unsigned int(1) Is_displaySafeInfo;
unsigned int(4) Reserved;

//all the following are optional fields
if(Is_CamParams){
    unsigned int(32) baseline;
    unsigned int(32) focallength;
    unsigned int(1) Is_camera_cross
    if(Is_camera_cross){
        unsigned int(32) convergence_distance;
        unsigned int(32) rotation[];
        unsigned int(32) translation[];
    }
    unsigned int(7) reserved;
}
if(Is_displaySafeInfo){
    unsigned int(16) ViewingDistance;
    int(16) VerticalDisparity;
    int(16) MinofDisparity;
    int(16) MaxofDisparity;
}
//other additional information
if(other_flag){}

aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0){
    unsigned int(16) item_count;
    ishdEntry[item_count] ishds;
}

```

[0138]

표 38은 다양한 카메라 파라미터를 지원하는 'ishd'의 제2 실시예의 신택스를 나타낸다. 스테레오스코픽 구성 포맷은 동일하다고 가정하며, 바로 전의 'ishd'가 참조된다.

표 38

[0139]

```

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

unsigned int(8) Stereoscopic_Composition_Type;
unsigned int(1) LR_first;
unsigned int(7) Reserved;
unsigned int(16) item_count;
for(i=0; i<item_count;i++) {
    unsigned int(16) item_ID;
    unsigned int(1) current_indicator;
    unsigned int(1) Is_camParams;
    unsigned int(1) Is_displaySafeInfo;
    unsigned int(5) Reserved;

    //all the following are optional fields
    if(Is_CamParams){
        unsigned int(32) baseline;
        unsigned int(32) focallength;
        unsigned int(1) Is_camera_cross
        unsigned int(7) reserved;
        if(Is_camera_cross){
            unsigned int(32) convergence_distance;
            unsigned int(32) rotation[];
            unsigned int(32) translation[];
        }
    }
    if(Is_displaySafeInfo){
        unsigned int(16) ViewingDisatance;
        int(16) VerticalDisparity;
        int(16) MinofDisparity;
        int(16) MaxofDisparity;
    }
    //other additional information
    if(other_flag){}
}

```

[0140] 표 39는 다양한 카메라 파라미터를 지원하는 경우의 'ishd'의 제3 실시예의 선택스를 나타낸다. 스테레오스코픽 구성 포맷은 동일하다고 가정하며, 'cameParams' 및 'displaysafeInfo'가 같은 임의의 Item_ID를 참조한다.

[0141] 표 39에서, 'Is_ref'는 '0'인 경우 참조되는 카메라 파라미터 및 디스플레이 정보가 없음을 표시하고, '1'인 경우 참조되는 Item_ID가 있음을 표시한다. 'current_refIndex'는 참조되는 Item_ID를 나타낸다.

표 39

[0142]

```

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

unsigned int(8) Stereoscopic_Composition_Type;
unsigned int(1) LR_first;
unsigned int(7) Reserved;
unsigned int(16)      item_count;
for(i=0; i<item_count;i++) {
    unsigned int(16) item_ID;
    unsigned int(1)   Is_camParams;
    unsigned int(1)   Is_displaySafeInfo;
    unsigned int(1)   Is_ref;
    unsigned int(5)   Reserved;

    //all the following are optional fields
    If(Is_ref){
        unsigned int(16)      current_refIndex;
    }
    else{
        if(Is_CamParams){
            unsigned int(32) baseline;
            unsigned int(32) focallength;
            unsigned int(1)   Is_camera_cross
            unsigned int(7)   reserved;
        }
        if(Is_camera_cross) {
            unsigned int(32)      convergence_distance;
            unsigned int(32)      rotation[];
            unsigned int(32) translation[];
        }
        if(Is_displaySafeInfo){
            unsigned int(16) ViewingDisatance;
            int(16)      VerticalDisparity;
            int(16)      MinofDisparity;
            int(16)      MaxofDisparity;
        }
    }
}
//other additional information
if(other_flag){}
}

```

[0143]

표 40은 다양한 카메라 파라미터를 지원하는 경우의 'ishd'의 제4 실시예의 신택스를 나타낸다. 스테레오스코픽 구성 포맷은 동일하다고 가정하며, 'cameParams' 및 'displaysafeInfo'가 서로 다른 임의의 Item_ID를 참조한다.

표 40

[0144]

```

Aligned(8) class Initial Stereoscopic Header box extend FullBox('ishd', version=0,0)

unsigned int(8) Stereoscopic_Composition_Type;
unsigned int(1) LR_first;
unsigned int(7) Reserved;
unsigned int(16) item_count;
for(i=0; i<item_count;i++) {
    unsigned int(16) item_ID;
    unsigned int(1) Is_camParams;
    unsigned int(1) Is_displaySafeInfo;
    unsigned int(1) Is_camParamsref;
    unsigned int(1) Is_displaySafeInforef;
    unsigned int(4) Reserved;

    //all the following are optional fields
    If(Is_ camParamsref || Is_displaySafeInforef){
        unsigned int(16) current_camrefIndex;
        unsigned int(16) current_displayrefIndex;
    }
    else{
        if(Is_CamParams) {
            unsigned int(32) baseline;
            unsigned int(32) focallength;
            unsigned int(1) Is_camera_cross
            unsigned int(7) reserved;
            if(Is_camera_cross) {
                unsigned int(32) convergence_distance;
                unsigned int(32) rotation[];
                unsgiend int(32) translation[];
            }
            if(Is_displaySafeInfo){
                unsigned int(16) ViewingDisatance;
                int(16) VerticalDisparity;
                int(16) MinofDisparity;
                int(16) MaxofDisparity;
            }
        }
    }
    //other additional information
    if(other_flag){}
}

```

[0145]

표 40에서, 'Is_camParamsref'는 '0'인 경우 참조되는 카메라 파라미터 정보가 없음을 나타내고, '1'인 경우 참조되는 Item_ID가 있음을 나타낸다. 'Is_displaySafeInforef'는 '0'인 경우 참조되는 디스플레이 safe 정보가 없음을 나타내고, '1'인 경우 참조되는 Item_ID가 있음을 나타낸다. 'current_cameraIndex'는 참조되는 Item_ID를 나타내고, 'Current_displayIndex'는 참조되는 Item_ID를 나타낸다.

[0146]

'ishd'는 스테레오스코픽 비디오 미디어 정보를 기록하는 'svmi'(Stereoscopic Video Media Information box)와 카메라 파라미터 및 디스플레이 정보를 기록하는 'scdi'(Stereoscopic Camera and Display Information box)로 분리될 수 있다. 'svmi'는 필수적(mandatory)이나, 'scdi'는 필수적이 아니므로 'ishd'를 'svmi'와 'scdi'로 분리하는 것은 불필요한 정보를 제거하는데 실익이 있다.

- [0147] 'svmi'는 스테레오스코픽 비주얼 타입과 구간 정보를 제공한다. 구체적으로, 스테레오스코픽 비디오 미디어 정보는 스테레오스코픽 영상 구성 타입 정보와, 왼쪽 이미지와 오른쪽 이미지 중에서 먼저 인코딩되는 이미지를 표시하는 정보와, 스테레오스코픽 콘텐츠의 기본 부호화 스트림이 스테레오스코픽 구간에서 모노스코픽 구간로 변경되거나 모노스코픽 구간에서 스테레오스코픽 구간로 변경될 때 구간의 수를 표시하는 정보와, 연속하는 샘플들의 수 또는 시작 샘플 번호를 카운트하는 정보와, 현재의 샘플이 스테레오스코픽인지 여부를 표시하는 정보 등을 포함한다.
- [0148] 'scdi'는 카메라 파라메타 존재 여부, 좌우 카메라 사이의 거리, 좌우 카메라의 배열, 메인 뷰 카메라로부터 서브 뷰 카메라까지의 상대 각도, 좌우 이미지 사이의 최대 디스패리티와 최소 디스패리티 등을 포함한다. 표 41은 'scdi'의 신택스의 일 실시예이다.

표 41

[0149]

```

unsigned int(16)  item_count;
for(i=0; i<item_count;i++) {
    unsigned int(16) item_ID;
    unsigned int(1)    Is_camParams;
    unsigned int(1)    Is_displaySafeInfo;
    unsigned int(1)    Is_ref;
    unsigned int(5)    Reserved;

    //all the following are optional fields
    If(Is_ref){
        unsigned int(16)    current_refIndex;
    }
    else{
        if(Is_CamParams){
            unsigned int(32) baseline;
            unsigned int(32)    focallength;
            unsigned int(1)    Is_camera_cross
            unsigned int(7)    reserved;
        }
        if(Is_camera_cross) {
            unsigned int(32)    convergence_distance;
            unsigned int(32)    rotation[];
            unsigned int(32) translation[];
        }
        if(Is_displaySafeInfo){
            unsigned int(16) ViewingDisatance;
            int(16)    VerticalDisparity;
            int(16)    MinofDisparity;
            int(16)    MaxofDisparity;
        }
    }
}

```

- [0150] 전술한 바와 같은 'tref'(Track reference box)를 통해 각 트랙에 포함되는 'scdi' 정보의 중복성을 제거할 수 있다. ES=2인 경우, 각 트랙의 'iloc'는 'scdi' 정보를 제공하기 위하여 스테레오스코픽 구간을 item_ID를 통하여 구분하는 역할을 수행하며, 이때 각 트랙의 스테레오스코픽 구간은 동일한 item_ID를 가지며 동일한 카메라 및 디스플레이 정보를 가진다. 'tref'의 'svdp'를 통해 기본 부호화 스트림은 주요/부가 트랙으로 구분되어지며 하나의 트랙에만 'iloc' 박스가 존재하더라도 3D 디스플레이를 할 때 스테레오스코픽 구간 별로 동기되어 재생이 가능하다.
- [0151] 또한 'tref'를 통해 각 트랙에 포함되는 동일한 스테레오스코픽 비디오 미디어 정보('svmi') 정보의 중복성을

제거할 수도 있다. 'ftyp'를 통해 스테레오스코픽 콘텐츠를 인식한 상황에서, 'tref'의 'svdp'를 통하여 주요/부가 트랙으로 구분되어지며 하나의 트랙에만 'svmi' 박스가 존재하면 다른 트랙은 자동적으로 인식될 수 있다. 'svmi' 박스는 필수적이므로 주요/부가 트랙 내 존재할 수도 있으며, 경우에 따라 주요 트랙 내에만 존재할 수 있다.

[0152] 도 15는 ES=2인 경우 SS-VAF의 일 실시예의 구조도이다. 도 15에 도시된 SS-VAF는 'svmi'와 'scdi' 박스를 구비한다.

[0153] 스테레오스코픽 콘텐츠를 구성함에 있어 ES가 2개인 경우, 2개의 트랙('trak')이 존재하며 이는 하나의 주요 트랙과 부가 트랙으로 구분할 수 있다. 이에 따라 부가 트랙 내 'tref'의 'svdp'를 이용하여 주요 트랙을 참조하고 관련 'scdi' 정보를 포함하는 정보를 제공한다. 이러한 구조는 각 트랙에 포함되는 동일한 'scdi' 정보의 중복성을 제거하는 장점을 가진다. 여기서, track_ID는 참조되는 트랙의 ID를 나타낸다. reference_type이 'svdp'이면 트랙이 참조 트랙에 대해 스테레오스코픽 카메라 및 디스플레이 정보를 포함함을 또한 나타낸다.

[0154] 사용자가 3D 디스플레이 모드로 시청하는 도중 2D 디스플레이 모드로 변경시 단말은 주요 트랙에 해당하는 영상만 디스플레이 함으로써 2D 디스플레이 모드를 만족한다. 한편, ES가 2개인 트랙에서 기본 부호화 스트림의 중간에 동일한 모노스코픽 데이터가 존재할 수 있다. 모노스코픽 데이터는 2개의 트랙에 같은 내용이 저장되는 형식으로 3D로 디스플레이 될 수 없는 데이터이다. 이러한 경우 단말에서는 2개의 트랙 중 어떤 트랙의 모노스코픽 데이터를 디스플레이 할지 결정을 해야 하는데 본 발명을 통하여 구분된 주요 트랙에 해당하는 영상만을 디스플레이 하도록 한다.

[0155] 또한, 스테레오스코픽 비디오의 ES가 2개인 경우 2개의 트랙이 존재하며 종래의 LASEr 등 장면기술자로서는 스테레오스코픽 비디오가 2개의 객체로 인식되어 LASEr 내 비디오 관련 노드의 2개와 연결되어야 한다. 하지만, 스테레오스코픽 비디오는 단말에서 최종적으로 하나의 3D 비디오 포맷으로 변환되어 재생되어야 하므로, LASEr에 하나의 객체로 인식되어야 한다. 즉, 스테레오스코픽 비디오는 2개의 트랙을 이용하여 구성되나, 이를 장면 재생하기 위해서는 하나의 3D 비디오 포맷으로 변환되어 재생되어야 하므로, 결국 하나의 노드만을 이용하며 연결되어야 한다. 이에 본 발명에서 제안한 'tref'내의 'svdp'를 이용하여 스테레오스코픽 비디오는 주요/부가 트랙으로 구분되며, LASEr 내 비디오 관련 노드는 주요 트랙에 해당되는 'trak'_ID 또는 미디어 스트림만을 링크한다.

[0156] 도 16은 본 발명에 따라 ISO 기반 미디어 파일 포맷으로 스테레오스코픽 콘텐츠를 저장하는 방법의 일 실시예의 흐름도이다. 먼저, 저장하고자 하는 스테레오스코픽 콘텐츠를 ISO 기반 미디어 파일 포맷의 'mdat' 박스에 저장한다(S1602). 다음에는 스테레오스코픽 콘텐츠에 대한 스테레오스코픽 비디오 미디어 정보와, 스테레오스코픽 카메라 및 디스플레이 정보를 스테레오스코픽 콘텐츠의 메타데이터로서 'moov'에 저장한다(S1604, S1606).

[0157] 도 17은 본 발명에 따라 ISO 기반 미디어 파일 포맷으로 스테레오스코픽 콘텐츠를 저장하는 방법의 다른 실시예의 흐름도이다. 먼저 저장하고자 하는 스테레오스코픽 콘텐츠를 ISO 기반 미디어 파일 포맷의 'mdat' 박스에 저장한다(S1702). 다음에는 스테레오스코픽 콘텐츠의 기본 부호화 스트림이 2개 이상인 경우, 기본 부호화 스트림들을 주요 트랙 또는 부가 트랙으로 구분하는 정보('svdp')를 ISO 기반 미디어 파일 포맷의 'tref' 박스에 저장한다(S1704). 다음에는 주요 트랙에만 링크되는 비디오 관련 노드를 구비하는 스테레오스코픽 콘텐츠에 대한 LASEr를 저장한다(S1706). 이때, LASEr에 링크되는 부분은 사용되지 않을 수 있으며 'tref'만 사용될 수 있다.

[0158] 도 18은 도 18은 본 발명에 의한 스테레오스코픽 콘텐츠 저장 방법의 또 다른 실시예의 흐름도이다. 여기에서는 전술한 바와 같은 AVC SEI를 이용하여 스테레오스코픽 콘텐츠를 저장한다. 먼저 AVC를 통해 부호화된 스테레오스코픽 콘텐츠를 저장한다(S1802). 다음에는 'reserved_sei_message'를 이용하여 상기 스테레오스코픽 콘텐츠의 각 스트림에 요구되는 카메라 파라미터 및 디스플레이 정보를 저장한다 (S1804).

[0159] 여기에서 'stereo video information SEI'는 스테레오스코픽 영상 구성이 'side by side' 타입임을 가리키는 정보를 추가로 포함한다. 카메라 파라미터는 좌우 카메라 사이의 거리, 좌우 동일 초점 거리(focal_length), 베이스라인으로부터 컨버전스 포인트까지의 거리, 좌우 카메라가 교차 배열하는 경우 좌우 카메라의 회전(rotation)과 이동(translation) 중 적어도 하나를 포함한다. 디스플레이 정보는 좌우 이미지 사이의 최대 디스패리티와 최소 디스패리티를 포함한다.

[0160] 상술한 바와 같은 본 발명의 방법은 프로그램으로 구현되어 컴퓨터로 읽을 수 있는 형태로 기록매체(씨디롬, 램, 롬, 플로피 디스크, 하드 디스크, 광자기 디스크 등)에 저장될 수 있다. 이러한 과정은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있으므로 더 이상 상세히 설명하지 않기로 한다.

[0161] 이상에서 설명한 본 발명은, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 있어 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하므로 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니다.

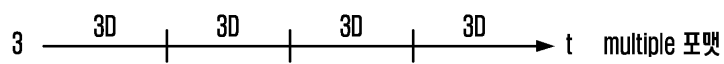
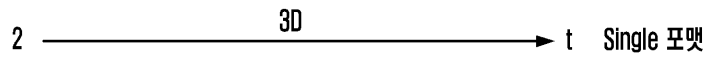
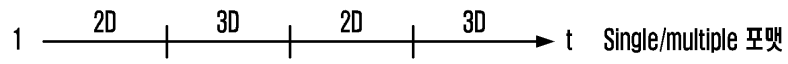
도면의 간단한 설명

- [0162] 도 1은 스테레오스코픽 콘텐츠를 제공하기 위해 구성할 수 있는 다양한 콘텐츠 구성 형태를 도시한다.
- [0163] 도 2는 본 발명에 의한 스테레오스코픽 비디오 어플리케이션 포맷을 위한 기본 파일 포맷의 구조도로서, MPEG-4 시스템 정보가 있는 경우이다.
- [0164] 도 3은 본 발명에 의한 스테레오스코픽 비디오 어플리케이션 포맷을 위한 기본 파일 포맷의 구조도로서, MPEG-4 시스템 정보가 없는 경우이다.
- [0165] 도 4는 본 발명에 의한 스테레오스코픽 비디오 어플리케이션 포맷의 제1 실시예의 구조도로서, MPEG-4 시스템 정보가 있는 경우이다.
- [0166] 도 5는 본 발명에 의한 스테레오스코픽 비디오 어플리케이션 포맷의 제2 실시예의 구조도로서, MPEG-4 시스템 정보가 없는 경우이다.
- [0167] 도 6은 본 발명에 의한 'ishd'를 'moov'에 포함시킨 저장 포맷의 구조도로서, (a)는 하나의 소스로 구성된 3차원 콘텐츠의 경우이고, (b)는 두개의 소스로 구성된 3차원 콘텐츠의 경우이다.
- [0168] 도 7은 본 발명에 의한 'ishd'를 'mdat'에 포함시킨 저장 포맷의 구조도로서, (a)는 하나의 소스로 구성된 3차원 콘텐츠의 경우이고, (b)는 두개의 소스로 구성된 3차원 콘텐츠의 경우이다.
- [0169] 도 8은 본 발명에 의한 'ishd'를 'meta'에 포함시킨 저장 포맷의 구조도로서, (a)는 하나의 소스로 구성된 3차원 콘텐츠의 경우이고, (b)는 두개의 소스로 구성된 3차원 콘텐츠의 경우이다.
- [0170] 도 9는 본 발명에 의한 'ishd'와 LASEr를 포함시킨 저장 포맷의 구조도로서, (a)는 'ishd'와 LASEr가 'moov'에 포함된 경우이고, (b)는 'ishd'는 'moov'에 포함되고 LASEr는 'mdat'에 포함된 경우이며, (c)는 'ishd'와 LASEr가 'meta'에 포함되는 경우이다.
- [0171] 도 10은 본 발명에 의한 'ishd'와 'iloc'가 포함된 SS-VAF의 구조도로서 ES=1인 경우이다.
- [0172] 도 11은 본 발명에 의한 'ishd'와 'iloc'가 포함된 SS-VAF의 구조도로서 ES=2인 경우이다.
- [0173] 도 12(a)는 AVC 내의 SEI의 기본 구조 중 'stereo video information SEI', 'reserved_sei_message'가 포함된 부분을 도시하고, 도 12(b)는 AVC 스트림내 SEI의 위치를 나타낸다.
- [0174] 도 13은 본 발명에 의한 'stereo video information SEI'와 'reserved_sei_message'를 이용한 SS-VAF의 구조도이다.
- [0175] 도 14은 본 발명에 의한 'tref', 'ishd', 'iloc' 박스들을 갖는 SS-VAF의 구조도로서, ES=2일 때 단일 포맷의 스테레오스코픽 스트림만으로 콘텐츠가 구성된 경우이다.
- [0176] 도 15은 본 발명에 의한 'tref', 'iloc', 'svmi', 'scdi' 박스들을 갖는 SS-VAF의 구조도이다.
- [0177] 도 16는 본 발명에 의한 스테레오스코픽 콘텐츠 저장 방법의 일 실시예의 흐름도이다.
- [0178] 도 17은 본 발명에 의한 스테레오스코픽 콘텐츠 저장 방법의 다른 실시예의 흐름도이다.
- [0179] 도 18은 본 발명에 의한 스테레오스코픽 콘텐츠 저장 방법의 또 다른 실시예의 흐름도이다.

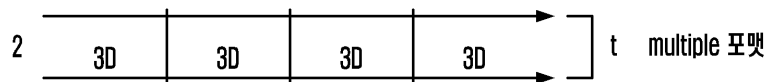
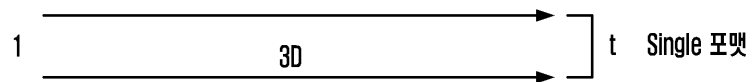
도면

도면1

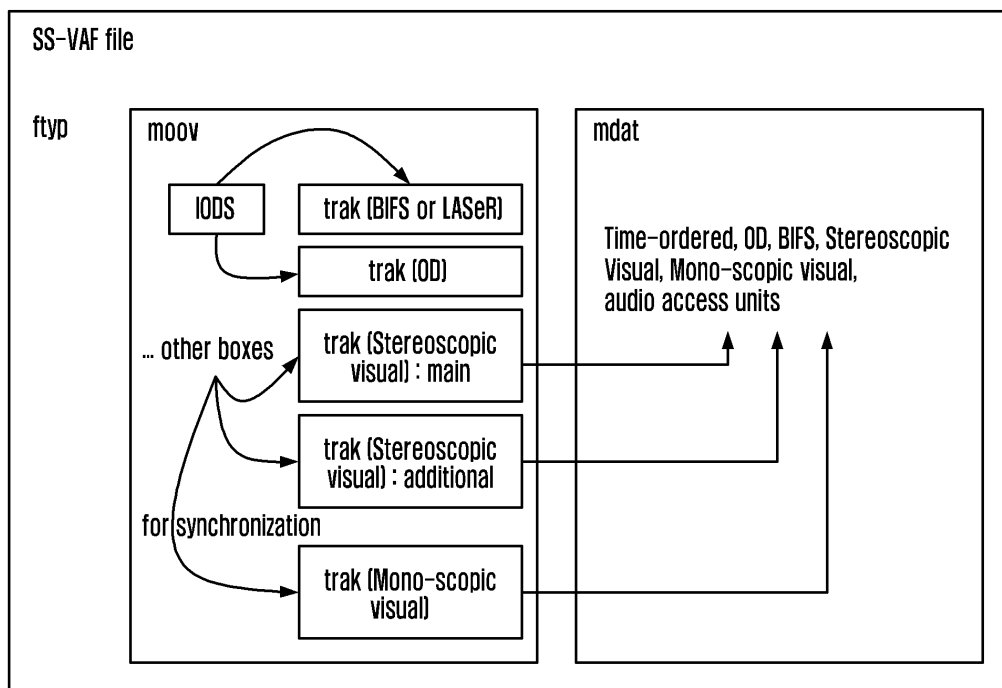
(a)



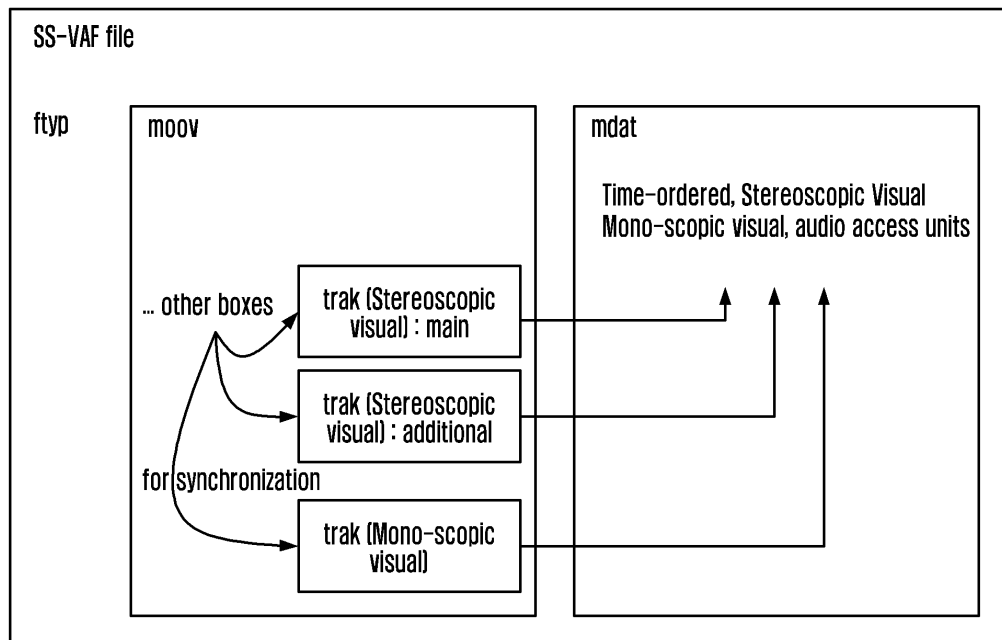
(b)



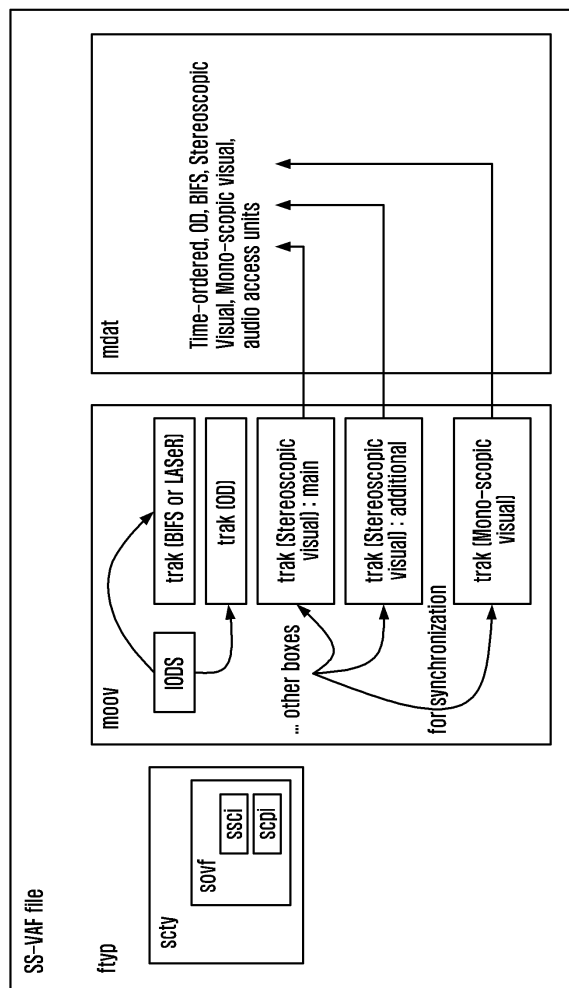
도면2



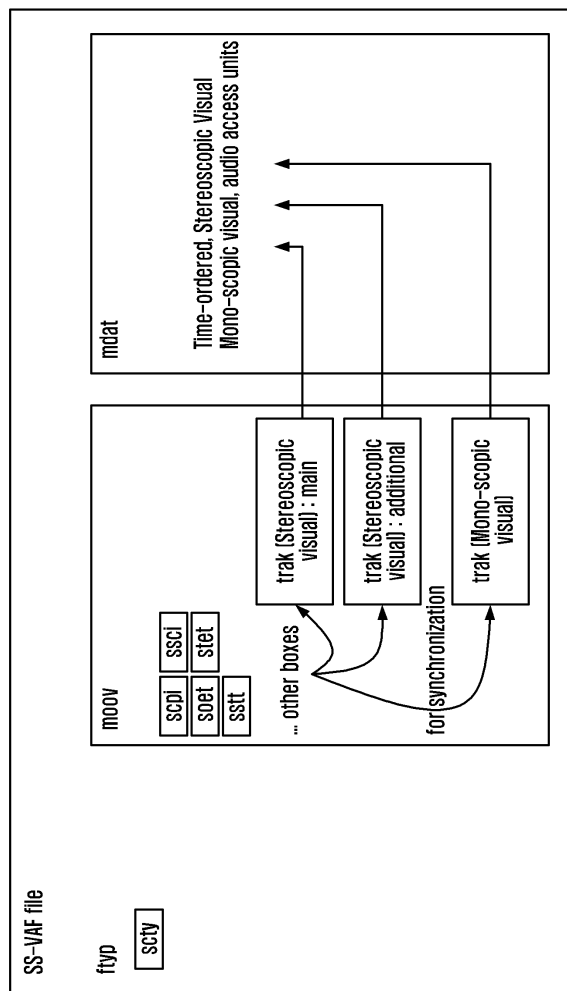
도면3



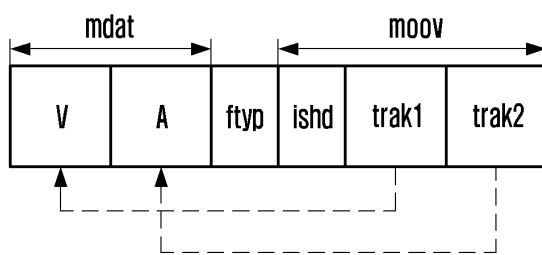
도면4



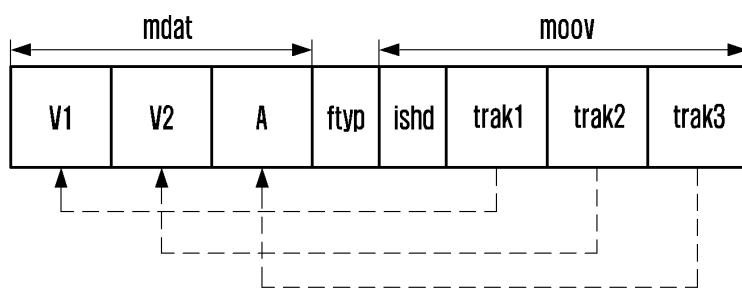
도면5



도면6

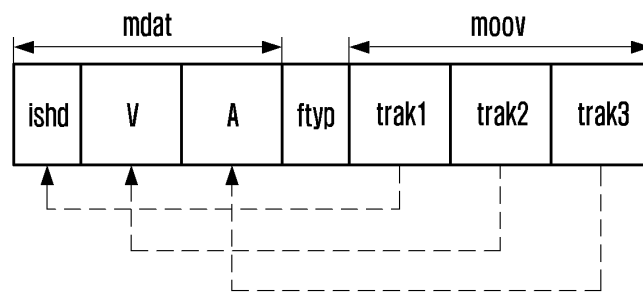


(a)

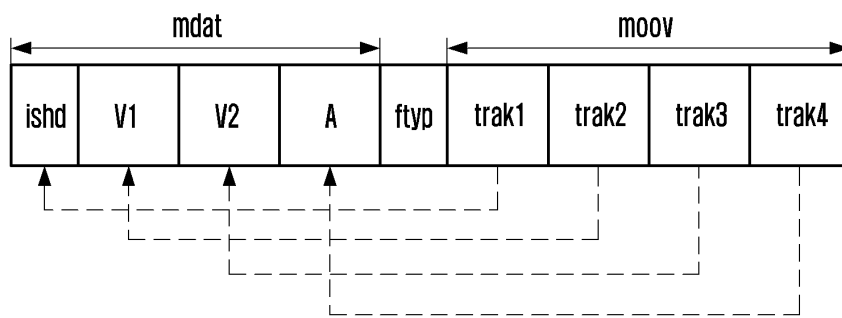


(b)

도면7

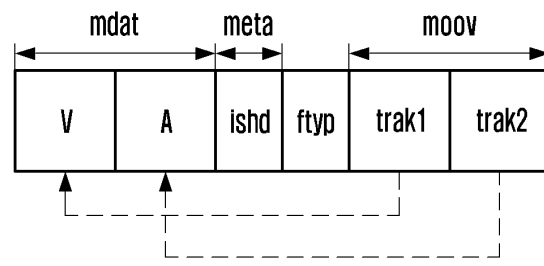


(a)

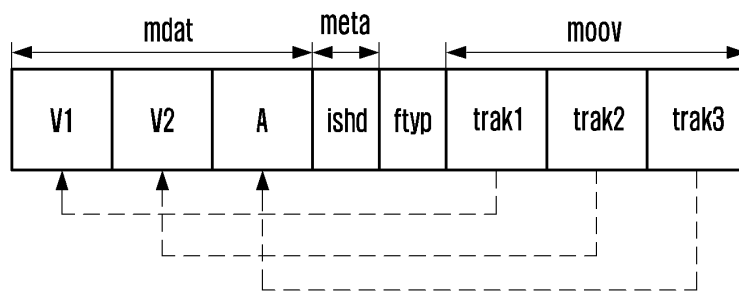


(b)

도면8

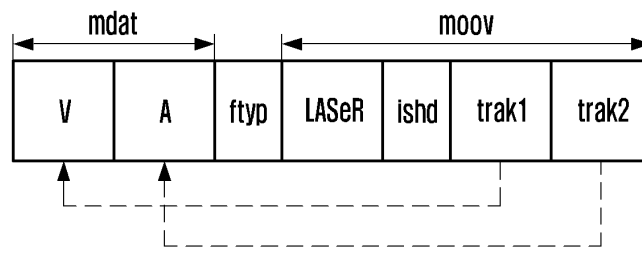


(a)

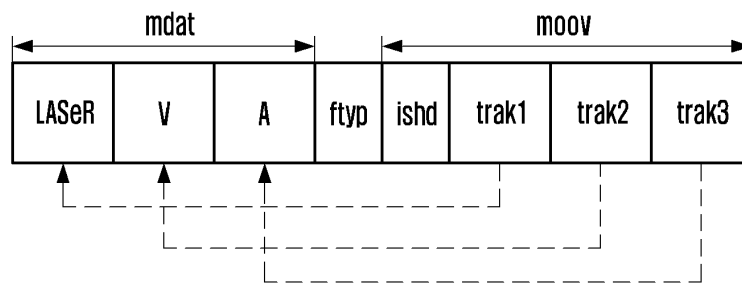


(b)

도면9

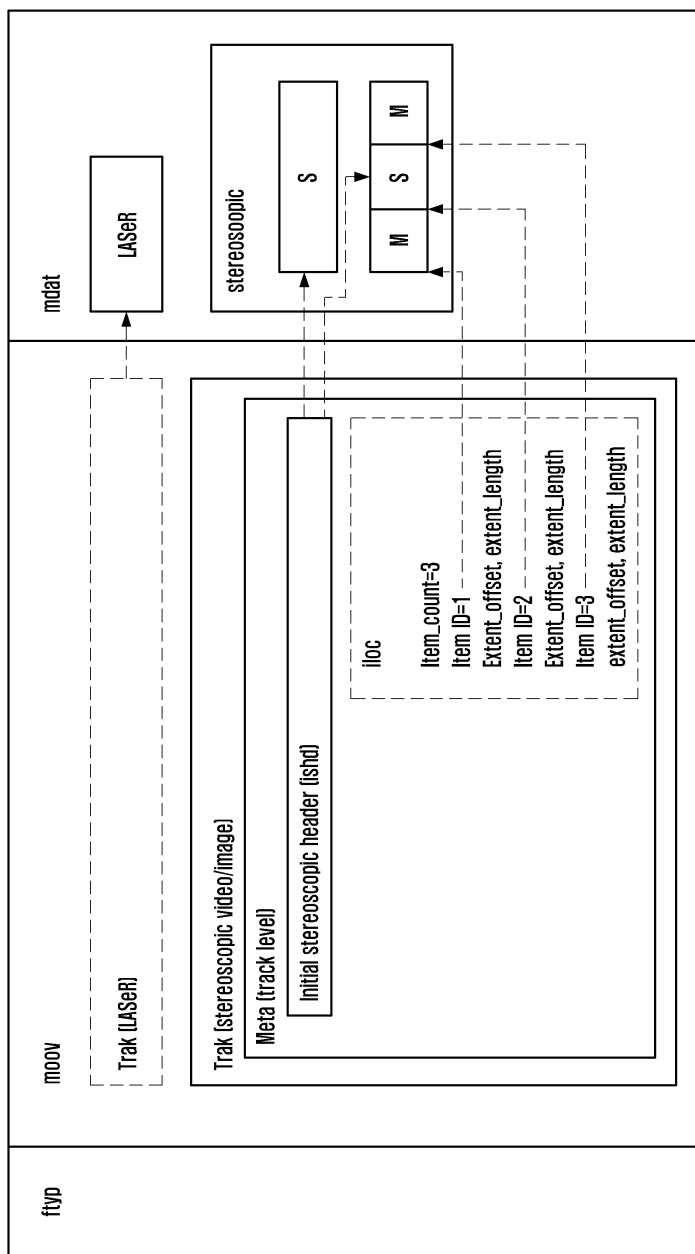


(a)

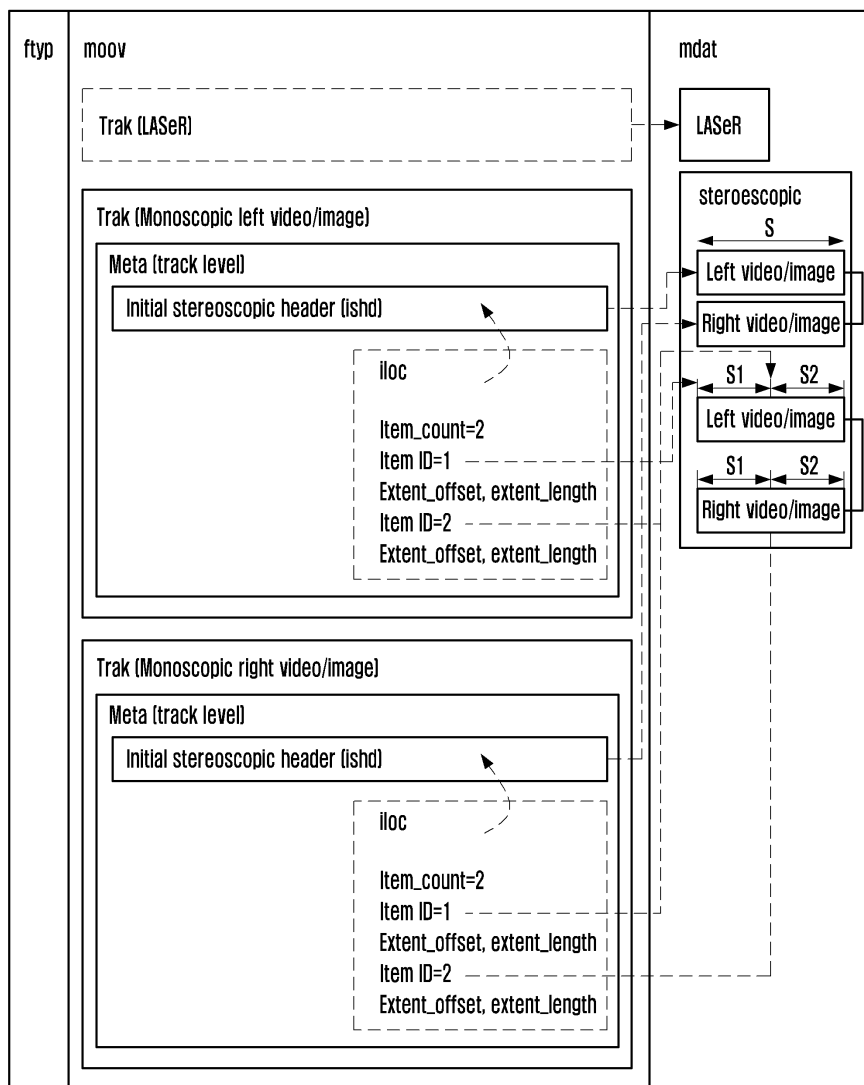


(b)

도면10



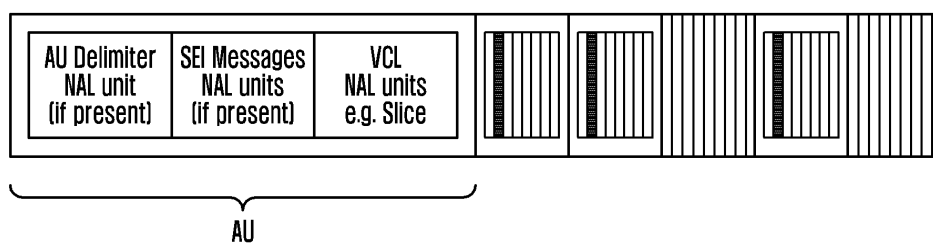
도면11



도면12

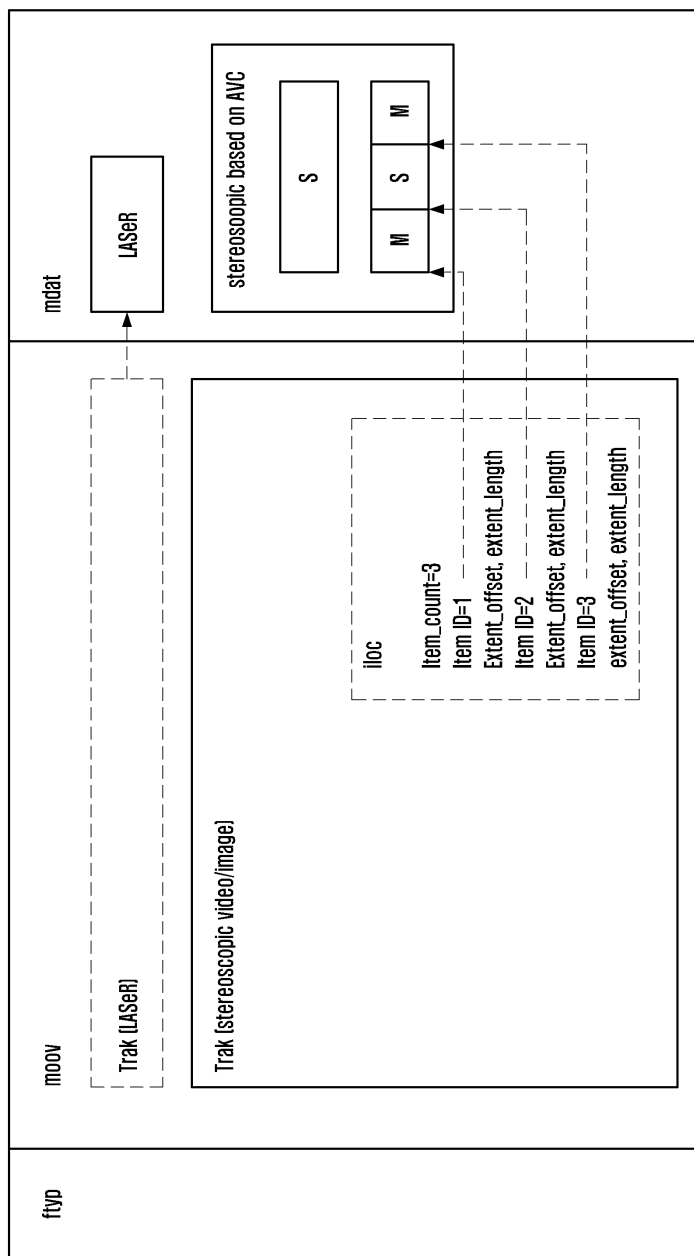
else if (payloadType == 21)		
stereo_video_info (payloadSize)	5	
else		
reserved_sei_message (payloadSize)	5	

(a)

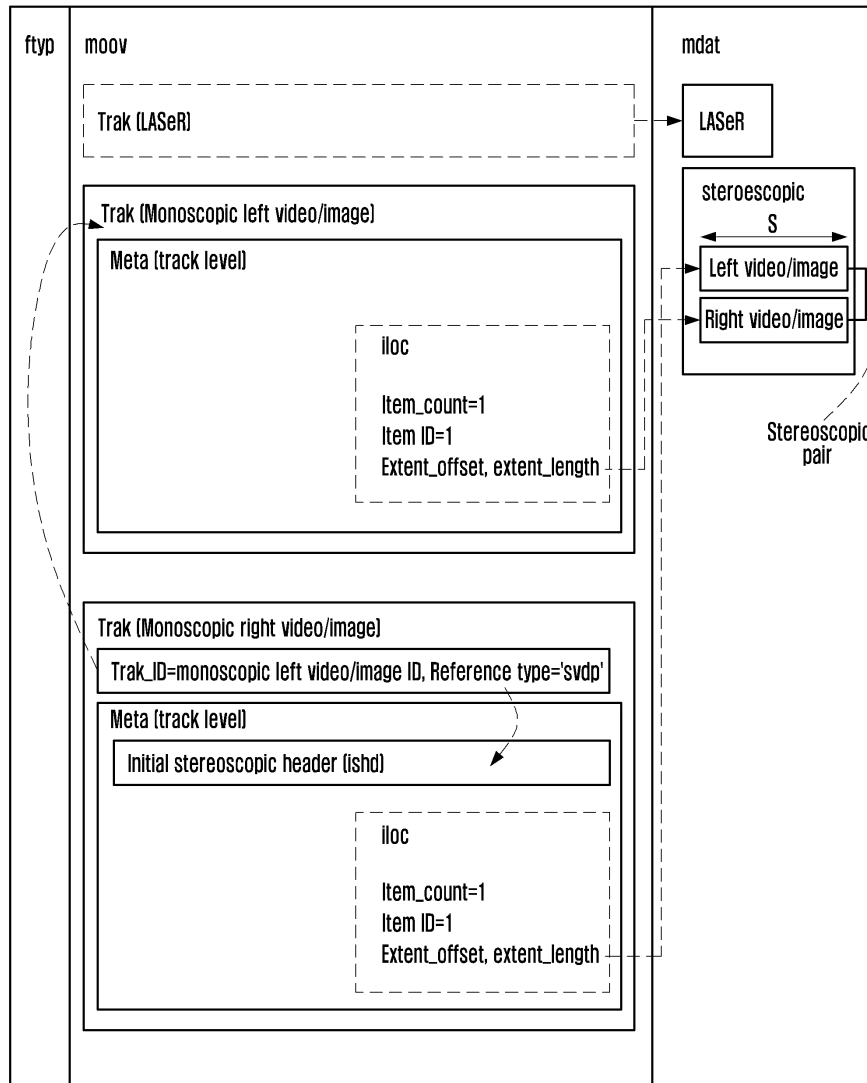


(b)

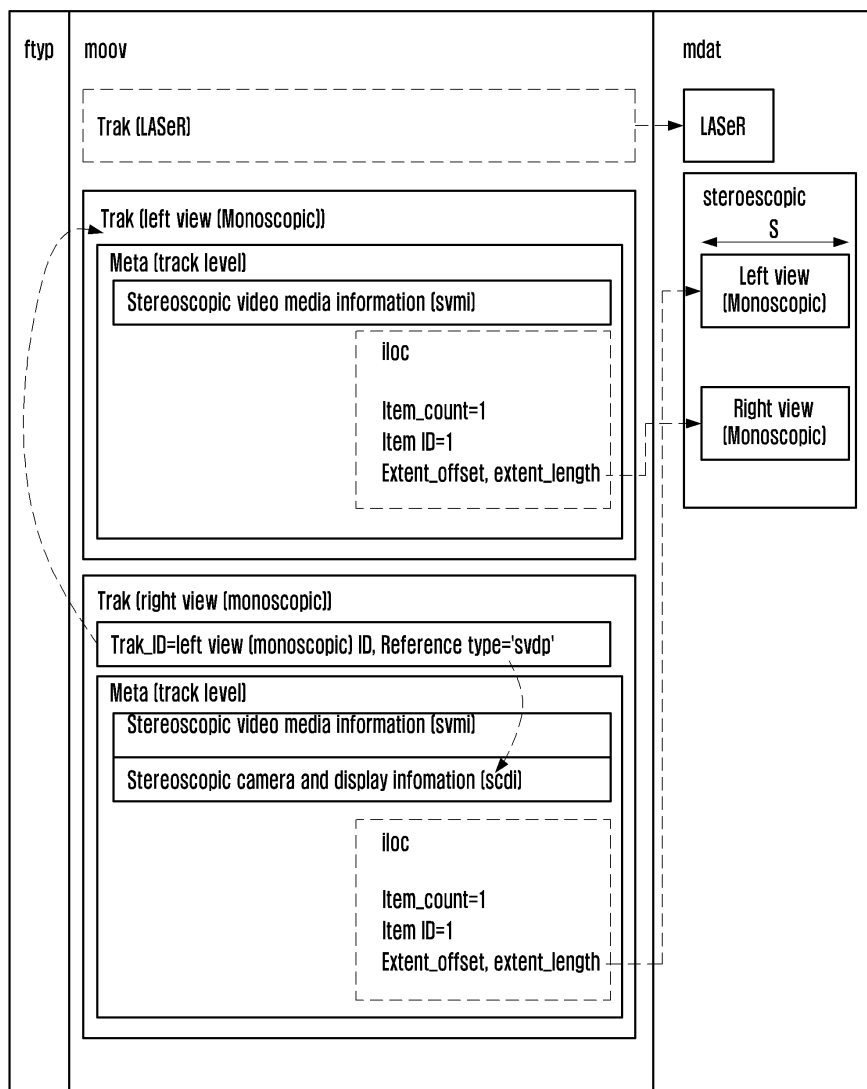
도면13



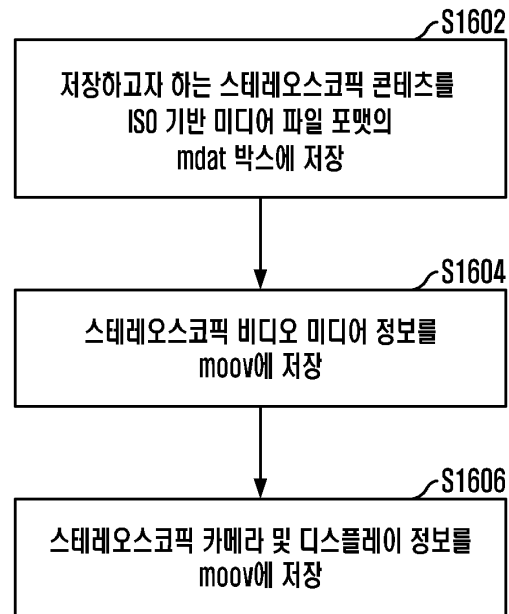
도면14



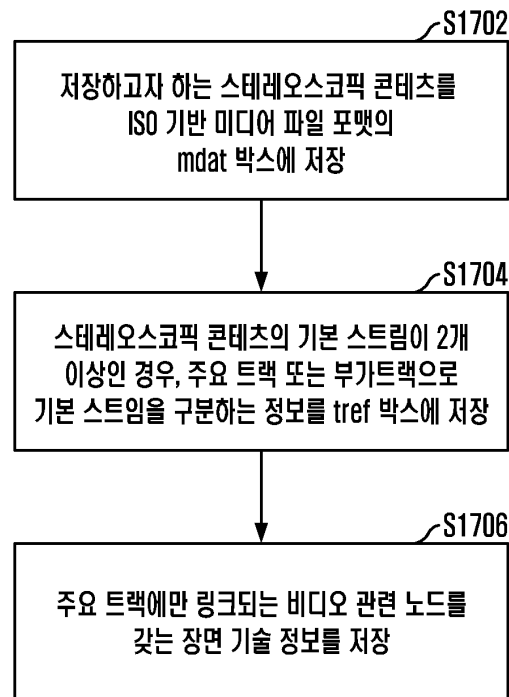
도면15



도면16



도면17



도면18

