



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0059915  
(43) 공개일자 2017년05월31일

- (51) 국제특허분류(Int. Cl.)  
*H04L 29/08* (2006.01) *G10L 19/16* (2013.01)  
*H04L 29/06* (2006.01) *H04N 21/235* (2011.01)  
*H04N 21/435* (2011.01) *H04N 21/81* (2011.01)  
*H04N 21/845* (2011.01) *H04N 21/854* (2011.01)  
*H04N 21/858* (2011.01)
- (52) CPC특허분류  
*H04L 67/2804* (2013.01)  
*G10L 19/167* (2013.01)
- (21) 출원번호 10-2016-7011630  
 (22) 출원일자(국제) 2015년09월07일  
 심사청구일자 없음  
 (85) 번역문제출일자 2016년05월02일  
 (86) 국제출원번호 PCT/JP2015/075313  
 (87) 국제공개번호 WO 2016/039285  
 국제공개일자 2016년03월17일
- (30) 우선권주장  
 JP-P-2014-186155 2014년09월12일 일본(JP)
- (71) 출원인  
 소니 주식회사  
 일본국 도쿄도 미나토쿠 코난 1-7-1
- (72) 발명자  
 츠카고시 이쿠오  
 일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니  
 주식회사 내
- (74) 대리인  
 장수길, 이중희

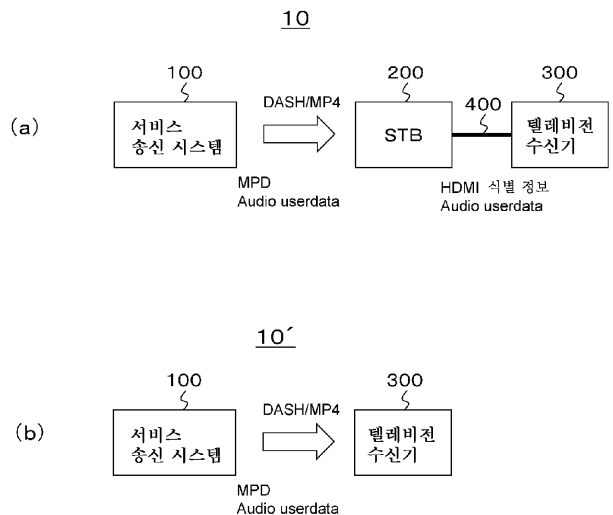
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 송신 장치, 송신 방법, 수신 장치 및 수신 방법

(57) 요약

오디오 스트림에 메타데이터가 삽입되어 있음을 수신측에서 용이하게 인식 가능하게 한다. 메타데이터가 삽입된 오디오 스트림을 수신 장치에서 취득하기 위한 메타 정보를 갖는 메타파일을 송신한다. 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를, 메타파일에 삽입한다. 수신측에서는, 오디오 스트림에 메타데이터가 삽입되어 있음을, 메타파일에 삽입되고 있는 식별 정보로부터 용이하게 인식할 수 있다.

대표도 - 도3



(52) CPC특허분류

*H04L 65/602* (2013.01)

*H04L 65/608* (2013.01)

*H04N 21/235* (2013.01)

*H04N 21/435* (2013.01)

*H04N 21/8113* (2013.01)

*H04N 21/8456* (2013.01)

*H04N 21/85406* (2013.01)

*H04N 21/8586* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

메타데이터가 삽입된 오디오 스트림을 수신 장치에서 취득하기 위한 메타 정보를 갖는 메타파일을 송신하는 송신부와,

상기 오디오 스트림에 상기 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 상기 메타파일에 삽입하는 정보 삽입부를 구비하는 송신 장치.

#### 청구항 2

제1항에 있어서,

상기 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보인 송신 장치.

#### 청구항 3

제2항에 있어서,

상기 메타데이터는, URI 정보를 나타내는 문자 코드인 송신 장치.

#### 청구항 4

제1항에 있어서,

상기 메타파일은, MPD 파일인 송신 장치.

#### 청구항 5

제4항에 있어서,

상기 정보 삽입부는,

“Supplementary Descriptor” 를 사용하여, 상기 메타파일에 상기 식별 정보를 삽입하는 송신 장치.

#### 청구항 6

제1항에 있어서,

상기 송신부는, 상기 메타파일을 RF 전송로 또는 통신 네트워크 전송로를 통해서 송신하는 송신 장치.

#### 청구항 7

제1항에 있어서,

상기 송신부는, 상기 메타데이터가 삽입된 오디오 스트림을 포함하는 소정 포맷의 컨테이너를 더 송신하는 송신 장치.

#### 청구항 8

제7항에 있어서,

상기 컨테이너는, MP4인 송신 장치.

#### 청구항 9

송신부에 의해, 메타데이터가 삽입된 오디오 스트림을 수신 장치에서 취득하기 위한 메타 정보를 갖는 메타파일을 송신하는 송신 스텝과,

상기 오디오 스트림에 상기 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 상기 메타파일에 삽입하는 정

보 삽입 스텝을 갖는 송신 방법.

**청구항 10**

메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신부를 구비하고,

상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고,

상기 오디오 스트림을, 상기 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 소정의 전송로를 통해서, 외부 기기에 송신하는 송신부를 더 구비하는 수신 장치.

**청구항 11**

제10항에 있어서,

기 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보인 장치.

**청구항 12**

제10항에 있어서,

상기 메타파일은, MPD 파일이며,

상기 메타파일에는, “Supplementary Descriptor”에 의해, 상기 식별 정보가 삽입되어 있는 수신 장치.

**청구항 13**

제10항에 있어서,

상기 송신부는, 화상 데이터의 블랭킹 기간에 상기 오디오 스트림 및 상기 식별 정보를 삽입하고, 그 화상 데이터를 상기 외부 기기에 송신함으로써, 상기 오디오 스트림 및 상기 식별 정보를 상기 외부 기기에 송신하는 수신 장치.

**청구항 14**

제10항에 있어서,

상기 소정의 전송로는, HDMI 케이블인 수신 장치.

**청구항 15**

수신부에 의해, 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신 스텝을 갖고,

상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고,

상기 오디오 스트림을, 상기 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 소정의 전송로를 통해서, 외부 기기에 송신하는 송신 스텝을 더 갖는 수신 방법.

**청구항 16**

메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신부를 구비하고,

상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고,

상기 식별 정보에 기초하여, 상기 오디오 스트림을 디코딩해서 상기 메타데이터를 추출하는 메타데이터 추출부와,

상기 메타데이터를 사용한 처리를 행하는 처리부를 더 구비하는 수신 장치.

**청구항 17**

제16항에 있어서,

상기 메타파일은, MPD 파일이며, 상기 메타파일에는, “Supplementary Descriptor”에 의해, 상기 식별 정보가 삽입되어 있는 수신 장치.

**청구항 18**

제16항에 있어서,

상기 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보이며,

상기 처리부는, 상기 네트워크 액세스 정보에 기초하여, 네트워크상의 소정의 서버에 액세스하는 수신 장치.

**청구항 19**

수신부에 의해, 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신 스텝을 갖고,

상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고,

상기 식별 정보에 기초하여, 상기 오디오 스트림을 디코딩해서 상기 메타데이터를 추출하는 메타데이터 추출 스텝과,

상기 메타데이터를 사용한 처리를 행하는 처리 스텝을 더 갖는 수신 방법.

**청구항 20**

네트워크 액세스 정보를 포함하는 메타데이터가 삽입된 오디오 스트림을 생성하는 스트림 생성부와,

상기 오디오 스트림을 갖는 소정 포맷의 컨테이너를 송신하는 송신부를 구비하는 송신 장치.

**발명의 설명**

**기술 분야**

[0001] 본 기술은, 송신 장치, 송신 방법, 수신 장치 및 수신 방법에 관한 것으로, 특히 오디오 스트림에 메타데이터를 삽입해서 송신하는 기술에 관련된 송신 장치 등에 관한 것이다.

**배경 기술**

[0002] 종래, 메타데이터를 오디오 스트림에 삽입해서 송신하는 것이 제안되어 있다(예를 들어, 특허문헌 1 참조).

**선행기술문헌**

**특허문헌**

[0003] (특허문헌 0001) 일본 특허 공개 제2012-010311호 공보

**발명의 내용**

**해결하려는 과제**

[0004] 메타데이터는 오디오 스트림의 예를 들어 유저 데이터 영역에 정의된다. 그러나, 모든 오디오 스트림에 메타데이터가 삽입되는 것은 아니다.

[0005] 본 기술의 목적은, 오디오 스트림에 메타데이터가 삽입되어 있음을 수신측에서 용이하게 인식 가능하게 해서 처리의 편의를 도모하는 데 있다.

**과제의 해결 수단**

- [0006] 본 기술의 개념은, 메타데이터가 삽입된 오디오 스트림을 수신 장치에서 취득하기 위한 메타 정보를 갖는 메타 파일을 송신하는 송신부와, 상기 오디오 스트림에 상기 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 상기 메타파일에 삽입하는 정보 삽입부를 구비하는 송신 장치에 있다.
- [0007] 본 기술에 있어서, 송신부에 의해, 메타데이터가 삽입된 오디오 스트림을 수신 장치에서 취득하기 위한 메타 정보를 갖는 메타파일이 송신된다. 예를 들어, 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보여도 된다. 이 경우, 예를 들어 메타데이터는, URI 정보를 나타내는 문자 코드여도 된다.
- [0008] 또한, 예를 들어 송신부는, 메타파일을 RF 전송로 또는 통신 네트워크 전송로를 통해서 송신하게 되어도 된다. 또한, 예를 들어 송신부는, 메타데이터가 삽입된 오디오 스트림을 포함하는 소정 포맷의 컨테이너를 더 송신하게 되어도 된다. 이 경우, 예를 들어 컨테이너는, MP4(ISO/IEC 14496-14:2003)이어도 된다.
- [0009] 정보 삽입부에 의해, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보가 메타파일에 삽입된다. 예를 들어, 메타파일은, MPD(Media Presentation Description) 파일이어도 된다. 이 경우, 예를 들어 정보 삽입부는, “Supplementary Descriptor” 를 사용하여, 메타파일에 식별 정보를 삽입하게 되어도 된다.
- [0010] 이렇게 본 기술에 있어서는, 메타데이터가 삽입된 오디오 스트림을 수신 장치에서 취득하기 위한 메타 정보를 갖는 메타파일에, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보가 삽입된다. 그로 인해, 수신측에서는, 오디오 스트림에 메타데이터가 삽입되어 있음을 용이하게 인식할 수 있다. 그리고, 예를 들어, 이 인식에 기초하여 오디오 스트림에 삽입되어 있는 메타데이터의 추출 처리를 행하는 것도 가능하여, 낭비 없이 확실하게 메타데이터의 취득이 가능하게 된다.
- [0011] 또한, 본 기술의 다른 개념은, 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신부를 구비하고, 상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고, 상기 오디오 스트림을, 상기 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 소정의 전송로를 통해서, 외부 기기에 송신하는 송신부를 더 구비하는 수신 장치에 있다.
- [0012] 본 기술에 있어서, 수신부에 의해, 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일이 수신된다. 예를 들어, 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보여도 된다. 메타파일에는, 오디오 스트림에 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있다.
- [0013] 예를 들어, 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보여도 된다. 또한, 예를 들어 메타파일은 MPD 파일이며, 이 메타파일에는, “Supplementary Descriptor” 에 의해, 식별 정보가 삽입되어 있어도 된다.
- [0014] 송신부에 의해, 오디오 스트림은, 이 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 소정의 전송로를 통해서, 외부 기기에 송신된다. 예를 들어, 송신부는, 화상 데이터의 블랭킹 기간에 오디오 스트림 및 식별 정보를 삽입하고, 이 화상 데이터를 외부 기기에 송신함으로써, 오디오 스트림 및 식별 정보를 외부 기기에 송신하게 되어도 된다. 또한, 예를 들어 소정의 전송로는, HDMI 케이블이어도 된다.
- [0015] 이렇게 본 기술에 있어서는, 메타데이터가 삽입된 오디오 스트림이, 이 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께 외부 기기에 송신된다. 그로 인해, 외부 기기측에서는, 오디오 스트림에 메타데이터가 삽입되어 있음을 용이하게 인식할 수 있다. 그리고, 예를 들어, 이 인식에 기초하여 오디오 스트림에 삽입되어 있는 메타데이터의 추출 처리를 행하는 것도 가능하여, 낭비 없이 확실하게 메타데이터의 취득이 가능하게 된다.
- [0016] 또한, 본 기술의 다른 개념은, 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신부를 구비하고, 상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고, 상기 식별 정보에 기초하여, 상기 오디오 스트림을 디코딩해서 상기 메타데이터를 추출하는 메타데이터 추출부와, 상기 메타데이터를 사용한 처리를 행하는 처리부를 더 구비하는 수신 장치에 있다.
- [0017] 본 기술에 있어서, 수신부에 의해, 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일이 수신된다. 메타파일에는, 오디오 스트림에 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어

있다. 예를 들어, 메타파일은 MPD 파일이며, 이 메타파일에는, “Supplementary Descriptor”에 의해, 식별 정보가 삽입되어 있어도 된다.

[0018] 메타데이터 추출부에 의해, 식별 정보에 기초하여, 오디오 스트림이 디코딩되어 메타데이터가 추출된다. 그리고, 처리부에 의해, 이 메타데이터를 사용한 처리가 행하여진다. 예를 들어, 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보이며, 처리부는, 네트워크 액세스 정보에 기초하여, 네트워크상의 소정의 서버에 액세스하게 되어도 된다.

[0019] 이렇게 본 기술에 있어서는, 메타파일에 삽입되고 있는, 오디오 스트림에 메타데이터의 삽입이 있음을 나타내는 식별 정보에 기초하여, 오디오 스트림으로부터 메타데이터가 추출되어 처리에 사용된다. 그로 인해, 오디오 스트림에 삽입되어 있는 메타데이터를 낭비 없이 확실하게 취득할 수 있어, 메타데이터를 사용한 처리를 적절하게 실행할 수 있다.

[0020] 또한, 본 기술의 또 다른 개념은, 네트워크 액세스 정보를 포함하는 메타데이터가 삽입된 오디오 스트림을 생성하는 스트림 생성부와, 상기 오디오 스트림을 갖는 소정 포맷의 컨테이너를 송신하는 송신부를 구비하는 송신 장치에 있다.

[0021] 본 기술에 있어서, 스트림 생성부에 의해, 네트워크 액세스 정보를 포함하는 메타데이터가 삽입된 오디오 스트림이 생성된다. 예를 들어, 오디오 스트림은, 오디오 데이터에 대하여, AAC, AC3, AC4, MPEGH(3D 오디오) 등의 부호화가 실시되어 생성되고, 그 유저 데이터 영역에 메타데이터가 매립된다.

[0022] 송신부에 의해, 그 오디오 스트림을 갖는 소정 포맷의 컨테이너가 송신된다. 여기서, 소정 포맷의 컨테이너란, 예를 들어 MP4, MPEG2-TS 등이다. 예를 들어, 메타데이터는, URI 정보를 나타내는 문자 코드여도 된다.

[0023] 이렇게 본 기술에 있어서는, 네트워크 액세스 정보를 포함하는 메타데이터가, 오디오 스트림에 매립되어 송신된다. 그 때문에, 예를 들어 방송국, 배신 서버 등으로부터, 네트워크 액세스 정보를, 오디오 스트림을 컨테이너로 해서 간단하게 송신할 수 있어, 수신측에서의 이용에 제공하는 것이 가능하게 된다.

**발명의 효과**

[0024] 본 기술에 의하면, 오디오 스트림에 메타데이터가 삽입되어 있음을 수신측에서 용이하게 인식 가능하게 된다. 또한, 본 명세서에 기재된 효과는 어디까지나 예시이며 한정되는 것이 아니고, 또한 부가적인 효과가 있어도 된다.

**도면의 간단한 설명**

- [0025] 도 1은 MPEG-DASH 베이스의 스트림 배신 시스템의 구성예를 도시하는 블록도.
- 도 2는 MPD 파일에 계층적으로 배치되어 있는 각 구조체의 관계의 일례를 도시하는 도면.
- 도 3은 실시 형태로서의 송수신 시스템의 구성예를 도시하는 블록도.
- 도 4는 MPD 파일 기술예를 도시하는 도면.
- 도 5는 「SupplementaryDescriptor」에 의한 「schemeIdUri」의 정의예를 도시하는 도면.
- 도 6은 트랜스포트 스트림 있어서의 비디오, 오디오의 액세스 유닛 배치예와, 오디오 스트림에 대한 메타데이터의 삽입 빈도를 설명하기 위한 도면.
- 도 7은 「<baseURL>」로 나타나는 로케이션처의 미디어 파일 실체에 대해서 설명하기 위한 도면.
- 도 8은 서비스 송신 시스템이 구비하는 DASH/MP4 생성부의 구성예를 도시하는 블록도.
- 도 9는 AAC의 오디오 프레임의 구조를 도시하는 도면.
- 도 10은 압축 포맷이 AAC인 경우에 메타데이터 MD가 삽입되는 「DSE(data stream element)」의 구성을 도시하는 도면.
- 도 11은 「metadata()」의 구성 및 그 구성의 주요한 정보의 내용을 도시하는 도면.
- 도 12는 「SDO\_payload()」의 구성을 도시하는 도면.
- 도 13은 커맨드 ID(cmdID)의 값의 의미를 도시하는 도면.

- 도 14는 AC3의 프레임(AC3 Synchronization Frame)의 구조를 도시하는 도면.
- 도 15는 AC3의 옥실러리 데이터(Auxiliary Data)의 구성을 도시하는 도면.
- 도 16은 AC4의 심플 트랜스포트(Simple Transport)의 레이어의 구조를 도시하는 도면.
- 도 17은 TOC(ac4\_toc()) 및 서브 스트림(ac4\_substream\_data())의 개략적인 구성을 도시하는 도면.
- 도 18은 TOC(ac4\_toc()) 중에 존재하는 「umd\_info()」의 구성을 도시하는 도면.
- 도 19는 서브 스트림(ac4\_substream\_data()) 중에 존재하는 「umd\_payloads\_substream()」의 구성을 도시하는 도면.
- 도 20은 MPEGH(3D 오디오)의 전송 데이터에 있어서의 오디오 프레임(1024 샘플)의 구조를 도시하는 도면.
- 도 21은 “Config”에 포함되는 각 “Frame”의 구성 정보(config)와, 각 “Frame”과의 대응이 유지되는 것을 설명하기 위한 도면.
- 도 22는 익스텐션 엘리먼트(Ext\_element)의 타입(ExElementType)과 그 값(Value)과의 대응 관계를 도시하는 도면.
- 도 23은 「userdataConfig()」의 구성을 도시하는 도면.
- 도 24는 「userdata()」의 구성을 도시하는 도면.
- 도 25는 송수신 시스템을 구성하는 셋톱 박스의 구성예를 도시하는 블록도.
- 도 26은 데이터 아일랜드 구간에 배치되는, 오디오·인포 프레임·패킷의 구조예를 도시하는 도면.
- 도 27은 송수신 시스템을 구성하는 텔레비전 수신기의 구성예를 도시하는 블록도.
- 도 28은 셋톱 박스의 HDMI 송신부와 나타내는 텔레비전 수신기의 HDMI 수신부의 구성예를 도시하는 블록도.
- 도 29는 TMDS 채널을 통해 화상 데이터가 전송되는 경우의 각종 전송 데이터의 구간을 도시하는 도면.
- 도 30은 텔레비전 수신기에 있어서의 메타데이터를 사용한 처리의 구체예를 설명하기 위한 도면.
- 도 31은 텔레비전 수신기에서 메타데이터에 기초하여 넷 서비스에 액세스하는 경우의 화면 표시의 천이예를 도시하는 도면.
- 도 32는 실시 형태에 있어서의 텔레비전 수신기에 있어서의 음성 출력계의 구성을 도시하는 블록도.
- 도 33은 텔레비전 수신기에 있어서의 음성 출력계의 다른 구성예를 도시하는 블록도.
- 도 34는 송수신 시스템의 다른 구성예를 도시하는 블록도.
- 도 35는 서비스 송신 시스템이 구비하는 TS 생성부의 구성예를 도시하는 블록도.
- 도 36은 오디오·유저 데이터·디스크립터의 구조예를 도시하는 도면.
- 도 37은 오디오·유저 데이터·디스크립터의 구조예에 있어서의 주요한 정보의 내용을 도시하는 도면.
- 도 38은 트랜스포트 스트림의 구성예를 도시하는 도면.
- 도 39는 송수신 시스템을 구성하는 셋톱 박스의 구성예를 도시하는 블록도.
- 도 40은 송수신 시스템을 구성하는 텔레비전 수신기의 구성예를 도시하는 블록도.

**발명을 실시하기 위한 구체적인 내용**

- [0026] 이하, 발명을 실시하기 위한 형태(이하, 「실시 형태」라고 함)에 대해서 설명한다. 또한, 설명을 이하의 순서로 행한다.
- [0027] 1. 실시 형태
- [0028] 2. 변형예
- [0029] <1. 실시 형태>



- [0030] [MPEG-DASH 베이스의 스트림 배신 시스템의 개요]
- [0031] 처음에, 본 기술을 적용할 수 있는 MPEG-DASH 베이스의 스트림 배신 시스템의 개요를 설명한다.
- [0032] 도 1의 (a)는 MPEG-DASH 베이스의 스트림 배신 시스템(30A)의 구성예를 나타내고 있다. 이 구성예에서는, 미디어 스트림과 MPD 파일이, 통신 네트워크 전송로를 통해서 송신된다. 이 스트림 배신 시스템(30A)은, DASH 스트림 파일 서버(31) 및 DASH MPD 서버(32)에, N개의 수신 시스템(33-1, 33-2, ..., 33-N)이, CDN(Content Delivery Network)(34)을 통해서 접속된 구성으로 되어 있다.
- [0033] DASH 스트림 파일 서버(31)는, 소정의 콘텐츠의 미디어 데이터(비디오 데이터, 오디오 데이터, 자막 데이터 등)에 기초하여, DASH 사양의 스트림 세그먼트(이하, 적절히, 「DASH 세그먼트」라고 말함)를 생성하고, 수신 시스템으로부터의 HTTP 요구에 따라서 세그먼트를 송출한다. 이 DASH 스트림 파일 서버(31)는 스트리밍 전용의 서버여도 좋고, 또한 웹(Web) 서버로 겸용되는 경우도 있다.
- [0034] 또한, DASH 스트림 파일 서버(31)는, 수신 시스템(33(33-1, 33-2, ..., 33-N))으로부터 CDN(34)을 통해서 보내져 오는 소정 스트림의 세그먼트 요구에 대응하여, 그 스트림의 세그먼트를, CDN(34)를 통해서, 요구원의 수신기에 송신한다. 이 경우, 수신 시스템(33)은, MPD(Media Presentation Description) 파일에 기재되어 있는 레이트의 값을 참조하여, 클라이언트가 놓여 있는 네트워크 환경의 상태에 따라, 최적의 레이트의 스트림을 선택해서 요구를 행한다.
- [0035] DASH MPD 서버(32)는, DASH 스트림 파일 서버(31)에 있어서 생성되는 DASH 세그먼트를 취득하기 위한 MPD 파일을 생성하는 서버이다. 콘텐츠 매니지먼트 서버(도시하지 않음)로부터의 콘텐츠 메타 데이터와, DASH 스트림 파일 서버(31)에 있어서 생성된 세그먼트의 어드레스(url)를 바탕으로, MPD 파일을 생성한다. 또한, DASH 스트림 파일 서버(31)와 DASH MPD 서버(32)는 물리적으로 동일한 것이어도 된다.
- [0036] MPD의 포맷에서는, 비디오나 오디오 등의 각각의 스트림마다 리프리젠테이션(Representation)이라고 하는 요소를 이용하여, 각각의 속성이 기술된다. 예를 들어, MPD 파일에는, 레이트가 상이한 복수의 비디오 데이터 스트림마다, 리프리젠테이션을 나누어서 각각의 레이트가 기술된다. 수신 시스템(33)에서는, 그 레이트의 값을 참고로 해서, 상술한 바와 같이, 수신 시스템(33)이 놓여 있는 네트워크 환경의 상태에 따라, 최적의 스트림을 선택할 수 있다.
- [0037] 도 1의 (b)는, MPEG-DASH 베이스의 스트림 배신 시스템(30B)의 구성예를 나타내고 있다. 이 구성예에서는, 미디어 스트림과 MPD 파일이, RF 전송로를 통해서 송신된다. 이 스트림 배신 시스템(30B)은, DASH 스트림 파일 서버(31) 및 DASH MPD 서버(32)가 접속된 방송 송출 시스템(36)과, M개의 수신 시스템(35-1, 35-2, ..., 35-M)으로 구성되어 있다.
- [0038] 이 스트림 배신 시스템(30B)의 경우, 방송 송출 시스템(36)은, DASH 스트림 파일 서버(31)에서 생성되는 DASH 사양의 스트림 세그먼트(DASH 세그먼트) 및 DASH MPD 서버(32)에서 생성되는 MPD 파일을, 방송파에 실어서 송신한다.
- [0039] 도 2는, MPD 파일에 계층적으로 배치되어 있는 각 구조체의 관계의 일례를 나타내고 있다. 도 2의 (a)에 도시하는 바와 같이, MPD 파일 전체로서의 미디어·프리젠테이션(Media Presentation)에는, 시간 간격으로 구획된 복수의 피리어드(Period)가 존재한다. 예를 들어, 최초의 피리어드는 스타트가 0초부터, 다음 피리어드는 스타트가 100초부터, 등으로 되어 있다.
- [0040] 도 2의 (b)에 도시하는 바와 같이, 피리어드에는, 복수의 리프리젠테이션(Representation)이 존재한다. 이 복수의 리프리젠테이션에는, 어댑테이션 세트(AdaptationSet)로 그룹핑되는, 스트림 속성, 예를 들어 레이트가 상이한 동일 내용의 미디어 스트림에 관한 리프리젠테이션군이 존재한다.
- [0041] 도 2의 (c)에 도시하는 바와 같이, 리프리젠테이션에는, 세그먼트 인포(SegmentInfo)가 포함되어 있다. 이 세그먼트 인포에는, 도 2의 (d)에 도시하는 바와 같이, 이니셜라이제이션 세그먼트(Initialization Segment)와, 피리어드를 더 미세하게 구획한 세그먼트(Segment)마다의 정보가 기술되는 복수의 미디어 세그먼트(Media Segment)가 존재한다. 미디어 세그먼트에는, 비디오나 오디오 등의 세그먼트 데이터를 실제로 취득하기 위한 어드레스(url)의 정보 등이 존재한다.
- [0042] 또한, 어댑테이션 세트로 그룹핑되어 있는 복수의 리프리젠테이션 사이에서는, 스트림의 스위칭을 자유롭게 행할 수 있다. 이에 의해, 수신 시스템이 놓여 있는 네트워크 환경의 상태에 따라, 최적의 레이트의 스트림을 선

택할 수 있어, 도중에 끊김이 없는 배신이 가능하게 된다.

- [0043] [송수신 시스템의 구성]
- [0044] 도 3은, 실시 형태로서의 송수신 시스템의 구성예를 나타내고 있다. 도 3의 (a)의 송수신 시스템(10)은 서비스 송신 시스템(100)과, 셋톱 박스(STB)(200)과, 텔레비전 수신기(TV)(300)를 갖고 있다. 셋톱 박스(200)와 텔레비전 수신기(300)는, HDMI(High Definition Multimedia Interface) 케이블(400)을 통해서 접속되어 있다. 또한, 「HDMI」는 등록 상표이다.
- [0045] 이 송수신 시스템(10)에 있어서, 서비스 송신 시스템(100)은, 상술한 도 1의 (a)에 도시하는 스트림 배신 시스템(30A)의 DASH 스트림 파일 서버(31) 및 DASH MPD 서버(32)에 대응한다. 또한, 이 송수신 시스템(10)에 있어서, 서비스 송신 시스템(100)은, 상술한 도 1의 (b)에 도시하는 스트림 배신 시스템(30B)의 DASH 스트림 파일 서버(31), DASH MPD 서버(32) 및 방송 송출 시스템(36)에 대응한다.
- [0046] 이 송수신 시스템(10)에 있어서, 셋톱 박스(STB)(200) 및 텔레비전 수신기(TV)(300)는, 상술한 도 1의 (a)에 도시하는 스트림 배신 시스템(30A)의 수신 시스템(33(33-1, 33-2, ..., 33-N))에 대응한다. 또한, 이 송수신 시스템(10)에 있어서, 셋톱 박스(STB)(200) 및 텔레비전 수신기(TV)(300)는, 상술한 도 1의 (b)에 도시하는 스트림 배신 시스템(30B)의 수신 시스템(35(35-1, 35-2, ..., 35-M))에 대응한다.
- [0047] 또한, 도 3의 (b)의 송수신 시스템(10')은, 서비스 송신 시스템(100)과, 텔레비전 수신기(TV)(300)를 갖고 있다. 이 송수신 시스템(10')에 있어서, 서비스 송신 시스템(100)은, 상술한 도 1의 (a)에 도시하는 스트림 배신 시스템(30A)의 DASH 스트림 파일 서버(31) 및 DASH MPD 서버(32)에 대응한다. 또한, 이 송수신 시스템(10')에 있어서, 서비스 송신 시스템(100)은, 상술한 도 1의 (b)에 도시하는 스트림 배신 시스템(30B)의 DASH 스트림 파일 서버(31), DASH MPD 서버(32) 및 방송 송출 시스템(36)에 대응한다.
- [0048] 이 송수신 시스템(10')에 있어서, 텔레비전 수신기(TV)(300)는, 상술한 도 1의 (a)에 도시하는 스트림 배신 시스템(30A)의 수신 시스템(33(33-1, 33-2, ..., 33-N))에 대응한다. 또한, 이 송수신 시스템(10')에 있어서, 텔레비전 수신기(TV)(300)는, 상술한 도 1의 (b)에 도시하는 스트림 배신 시스템(30B)의 수신 시스템(35(35-1, 35-2, ..., 35-M))에 대응한다.
- [0049] 서비스 송신 시스템(100)은, DASH/MP4, 즉 메타파일로서의 MPD 파일과, 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4를, RF 전송로 또는 통신 네트워크 전송로를 통해서 송신한다. 서비스 송신 시스템(100)은, 오디오 스트림에 메타데이터를 삽입한다. 이 메타데이터로서는, 예를 들어 소정의 넷 서비스에 접속하기 위한 액세스 정보, 소정의 콘텐츠 정보 등이 생각된다. 이 실시 형태에서는, 소정의 넷 서비스에 접속하기 위한 액세스 정보가 삽입된다.
- [0050] 서비스 송신 시스템(100)은, MPD 파일에, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 삽입한다. 서비스 송신 시스템(100)은, 예를 들어 “Supplementary Descriptor” 를 사용하여, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 삽입한다.
- [0051] 도 4는, MPD 파일 기술예를 나타내고 있다. 「<AdaptationSet mimeType=“audio/mp4” group=“1” >」이라는 기술에 의해, 오디오 스트림에 대한 어댑테이션 세트(AdaptationSet)가 존재하고, 그 오디오 스트림은 MP4 파일 구조로 공급되고, 그룹1이 할당되어 있음이 나타나 있다.
- [0052] 또한, 「<SupplementaryDescriptor schemeIdUri=“urn:brdcst:AudioMetaContained” value=“true” />」라는 기술에 의해, 오디오 스트림에 메타데이터가 삽입되어 있음이 나타나 있다. 「SupplementaryDescriptor」에 의해, 「schemeIdUri」를, 종래의 규격에 있어서의 기존의 정의와는 달리, 방송 그 밖의 애플리케이션으로서 신규 정의가 가능하다. 도 5에 도시하는 바와 같이, 「schemeIdUri=“urn:brdcst:Audiometacontained”」는, 오디오 메타 정보가 포함됨, 즉 오디오 스트림에 메타데이터가 삽입되어 있음을 나타낸다. 예를 들어, 「value」가 “true” 일 때는, 오디오 메타 정보가 포함됨을 나타낸다. 「value」가 “false” 일 때는, 오디오 메타 정보가 포함되지 않음을 나타낸다.
- [0053] 또한, 「<SupplementaryDescriptor schemeIdUri=“urn:brdcst:codecType” value=“mpegH” />」라는 기술에 의해, 오디오 스트림의 코덱이 MPEGH(3D 오디오)임이 나타나 있다. 도 5에 도시하는 바와 같이, 「schemeIdUri=“urn:brdcst:codecType”」는, 코덱의 종류를 나타낸다. 예를 들어, 「value」는 “mpegH”, “AAC”, “AC3”, “AC4” 등으로 된다.
- [0054] 또한, 「<SupplementaryDescriptor schemeIdUri=“urn:brdcst:coordinatedControl” value=“true” />」라는

기술에 의해, 넷 접속에 필요한 정보가 복수의 미디어 스트림 사이에서 강조해서 공급되는 것이 나타나 있다. 도 5에 도시하는 바와 같이, 「schemeIdUri= “urn:brdcst:coordinatedControl”」은, 넷 접속에 필요한 정보가 복수의 미디어 스트림 사이에서 협조해서 공급됨을 나타낸다. 예를 들어, 「value」가 “true” 일 때는, 넷 접속 정보가 다른 어댑테이션 세트의 스트림과 협조해서 공급됨을 나타낸다. 「value」가 “false” 일 때는, 넷 접속 정보가 본어댑테이션 세트의 스트림만으로 공급됨을 나타낸다.

[0055] 또한, 「<SupplementaryDescriptor schemeIdUri= “urn:brdcst:type” value= “netlink” />」의 기술에 의해, 메타에 의한 서비스의 종류가 넷 접속인 것이 나타나 있다. 도 5에 도시하는 바와 같이, 「schemeIdUri= “urn:brdcst:type”」으로 메타에 의한 서비스의 종류를 나타낸다. 예를 들어, 「value」가 “netlink” 일 때는, 메타에 의한 서비스의 종류가 넷 접속임을 나타낸다.

[0056] 또한, 「<SupplementaryDescriptor schemeIdUri= “urn:brdcst:metaInsertionFrequency” value= “1” />」이라는 기술에 의해, 액세스 유닛 단위로 메타 정보가 공급되는 것이 나타나 있다. 도 5에 도시하는 바와 같이, 「schemeIdUri= “urn:brdcst:metaInsertionFrequency”」는, 액세스 유닛 단위로 메타 정보가 공급되는 빈도를 나타낸다. 예를 들어, 「value」가 “1” 일 때는, 1개의 유저 데이터 엔트리(entry)가 1액세스 유닛에 발생하는 것을 나타낸다. 「value」가 “2” 일 때는, 복수의 유저 데이터 엔트리가 1액세스 유닛에 발생하는 것을 나타낸다. 「value」가 “3” 일 때는, 1 이상의 유저 데이터 엔트리가, 랜덤 액세스 포인트로 구획되는 기간 중에 발생하는 것을 나타낸다.

[0057] 도 6의 (a)는, MP4로 컨테이너되는 비디오, 오디오의 액세스 유닛의 배치예를 나타내고 있다. 「VAU」는, 비디오 액세스 유닛을 나타내고 있다. 「AAU」는, 오디오 액세스 유닛을 나타내고 있다. 도 6의 (b)는, 「frequency\_type =1」인 경우이며, 각 오디오 액세스 유닛에 하나의 유저 데이터 엔트리(메타데이터)가 삽입되어 있음을 나타내고 있다.

[0058] 도 6의 (c)는, 「frequency\_type =2」인 경우이며, 하나의 오디오 액세스 유닛에 복수의 유저 데이터(메타데이터)가 삽입되어 있음을 나타내고 있다. 도 6의 (d)는, 「frequency\_type =3」인 경우이며, 랜덤 액세스 포인트를 포함하는 그룹마다, 그 선두의 오디오 액세스 유닛에 적어도 하나의 유저 데이터(메타데이터)가 삽입되어 있음을 나타내고 있다.

[0059] 또한, 도 4로 되돌아가서, 「<Representation id= “11” bandwidth= “128000” >」이라는 기술에 의해, 「Representation id= “11”」로서, 비트 레이트가 128kbps인 오디오 스트림의 존재가 나타나 있다. 그리고, 「<baseURL>audio/jp/128.mp4 </BaseURL>」이라는 기술에 의해, 그 오디오 스트림의 로케이션처가, 「audio/jp/128.mp4」로서 나타나 있다.

[0060] 또한, 「<AdaptationSet mimeType= “video/mp4” group= “2” >」의 기술에 의해, 비디오 스트림에 대한 어댑테이션 세트(AdaptationSet)가 존재하고, 그 비디오 스트림은 MP4 파일 구조로 공급되고, 그룹2가 할당되어 있음이 나타나 있다.

[0061] 또한, 「<SupplementaryDescriptor schemeIdUri= “urn:brdcst:VideoMetaContained” value= “true” />」의 기술에 의해, 비디오 스트림에 메타데이터가 삽입되어 있음이 나타나 있다. 도 5에 도시하는 바와 같이, 「schemeIdUri= “urn:brdcst:VideoMetaContained”」는, 비디오 메타 정보가 포함되는 것, 즉 비디오 스트림에 메타데이터가 삽입되어 있음을 나타낸다. 예를 들어, 「value」가 “true” 일 때는, 비디오 메타 정보가 포함됨을 나타낸다. 「value」가 “false” 일 때는, 비디오 메타 정보가 포함되지 않음을 나타낸다.

[0062] 또한, 「<SupplementaryDescriptor schemeIdUri= “urn:brdcst:codecType” value= “hevc” />」라는 기술에 의해, 비디오 스트림의 코덱이 HEVC임이 나타나 있다. 또한, 「<SupplementaryDescriptor schemeIdUri= “urn:brdcst:coordinatedControl” value= “true” />」라는 기술에 의해, 넷 접속에 필요한 정보가 복수의 미디어 스트림 사이에서 강조해서 공급되는 것이 나타나 있다.

[0063] 또한, 「<SupplementaryDescriptor schemeIdUri= “urn:brdcst:type” value= “netlink” />」라는 기술에 의해, 메타에 의한 서비스의 종류가 넷 접속임이 나타나 있다. 또한, 「<SupplementaryDescriptor schemeIdUri= “urn:brdcst:metaInsertionFrequency” value= “1” />」의 기술에 의해, 액세스 유닛 단위로 메타 정보가 공급되는 것이 나타나 있다.

[0064] 또한, 「<Representation id= “21” bandwidth= “20000000” >」의 기술에 의해, 「Representation id= “21”」로서, 비트 레이트가 20Mbps인 비디오 스트림의 존재가 나타나 있다. 그리고, 「<baseURL>video/jp/20000000.mp4 </BaseURL>」의 기술에 의해, 그 비디오 스트림의 로케이션처가, 「

video/jp/20000000.mp4」로서 나타나 있다.

- [0065] 여기서, 「<baseURL>」로 나타나는 로케이션처의 미디어 파일 실체에 대해서 설명한다. 는 프래그먼트 MP4(Non-Fragmented MP4)의 경우, 예를 들어 도 7의 (a)에 도시하는 바와 같이, “url 1”로서 정의되는 경우가 있다. 이 경우, 처음에 파일 타입이 기술되는 “ftyp” 박스가 배치된다. 이 “ftyp” 박스에 의해, 단편화되지 않은 MP4 파일임이 나타난다. 계속해서, “moov” 박스 및 “mdat” 박스가 배치된다. “moov” 박스에는, 모든 메타데이터, 예를 들어 각 트랙의 헤더 정보나 콘텐츠 내용의 메타 기술, 시각 정보 등이 포함된다. “mdat” 박스에는, 미디어 데이터 본체가 들어간다.
- [0066] 또한, 프래그먼트 MP4(Fragmented MP4)의 경우, 예를 들어 도 7의 (b)에 도시하는 바와 같이, “url 2”로서 정의되는 경우가 있다. 이 경우, 처음에 세그먼트 타입이 기술되는 “styp” 박스가 배치된다. 계속해서, 세그먼트 인덱스가 기술되는 “sidx” 박스가 배치된다. 그것에 이어서, 소정 개수의 무비 프래그먼트(Movie Fragment)가 배치된다. 여기서, 무비 프래그먼트는, 제어 정보가 들어가는 “moof” 박스와, 미디어 데이터 본체가 들어가는 “mdat” 박스로부터 구성된다. 하나의 무비 프래그먼트의 “mdat” 박스에는, 전송 미디어가 단편화되어 얻어진 단편이 들어가므로, “moof” 박스에 들어가는 제어 정보는 그 단편에 관한 제어 정보가 된다.
- [0067] 또한, 상술한 “url 1”과 “url 2”의 조합도 생각된다. 이 경우, 예를 들어 “url 1”이 이니셜라이제이션 세그먼트(initialization segment)로서, “url 1”과 “url 2”를 1 서비스의 MP4로 하는 것도 가능하다. 또는, “url 1”과 “url 2”를 하나로 통합하여, 도 7의 (c)에 도시하는 바와 같이, “url 3”으로서 정의하는 것도 가능하다.
- [0068] 셋톱 박스(200)는, 서비스 송신 시스템(100)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 DASH/MP4, 즉 메타파일로서의 MPD 파일과, 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4를 수신한다. MP4에 포함되는 오디오 스트림에는, 메타데이터로서, 소정의 넷 서비스에 접속하기 위한 액세스 정보가 삽입되어 있다. 또한, MPD 파일에는, “Supplementary Descriptor”에 의해, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보가 삽입되어 있다.
- [0069] 셋톱 박스(200)는, 오디오 스트림을, 이 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, HDMI 케이블(400)을 통해서, 텔레비전 수신기(300)에 송신한다.
- [0070] 여기서, 셋톱 박스(200)는, 비디오 스트림을 디코딩해서 얻어진 화상 데이터의 블랭킹 기간에 오디오 스트림 및 식별 정보를 삽입하고, 이 화상 데이터를 텔레비전 수신기(300)에 송신함으로써, 오디오 스트림 및 식별 정보를 텔레비전 수신기(300)에 송신한다. 셋톱 박스(200)는, 이 식별 정보를, 예를 들어 오디오·인포 프레임 패킷(Audio InfoFrame packet)에 삽입한다.
- [0071] 텔레비전 수신기(300)는, 도 3의 (a)에 도시하는 송수신 시스템(10)에 있어서는, 셋톱 박스(200)로부터, HDMI 케이블(400)을 통해서, 오디오 스트림을, 이 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 수신한다. 즉, 텔레비전 수신기(300)는, 셋톱 박스(200)로부터, 오디오 스트림 및 식별 정보가 블랭킹 기간에 삽입되어 있는 화상 데이터를 수신한다.
- [0072] 그리고, 텔레비전 수신기(300)는, 식별 정보에 기초하여, 오디오 스트림을 디코딩해서 메타데이터를 추출하고, 이 메타데이터를 사용한 처리를 한다. 이 경우, 텔레비전 수신기(300)는, 메타데이터로서의 소정의 넷 서비스 정보에 기초하여, 네트워크상의 소정의 서버에 액세스한다.
- [0073] 또한, 텔레비전 수신기(300)는, 도 3의 (b)에 도시하는 송수신 시스템(10')에 있어서는, 서비스 송신 시스템(100)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 DASH/MP4, 즉 메타파일로서의 MPD 파일과, 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4를 수신한다. MP4에 포함되는 오디오 스트림에는, 메타데이터로서, 소정의 넷 서비스에 접속하기 위한 액세스 정보가 삽입되어 있다. 또한, MPD 파일에는, “Supplementary Descriptor”에 의해, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보가 삽입되어 있다.
- [0074] 그리고, 텔레비전 수신기(300)는, 식별 정보에 기초하여, 오디오 스트림을 디코딩해서 메타데이터를 추출하고, 이 메타데이터를 사용한 처리를 한다. 이 경우, 텔레비전 수신기(300)는 메타데이터로서의 소정의 넷 서비스 정보에 기초하여, 네트워크상의 소정의 서버에 액세스한다.
- [0075] [서비스 송신 시스템의 DASH/MP4 생성부]
- [0076] 도 8은, 서비스 송신 시스템(100)이 구비하는 DASH/MP4 생성부(110)의 구성예를 나타내고 있다. 이 DASH/MP4

생성부(110)는, 제어부(111)와, 비디오 인코더(112)와, 오디오 인코더(113)와, DASH/MP4 포맷터(114)를 갖고 있다.

- [0077] 제어부(111)는, CPU(111a)를 구비하고 있고, DASH/MP4 생성부(110)의 각 부를 제어한다. 비디오 인코더(112)는, 화상 데이터 SV에 대하여 MPEG2, H.264/AVC, H.265/HEVC 등의 부호화를 실시하여, 비디오 스트림(비디오 엘리먼트리 스트림)을 생성한다. 화상 데이터 SV는, 예를 들어 HDD 등의 기록 매체로부터 재생된 화상 데이터, 또는 비디오 카메라에서 얻어진 라이브 화상 데이터 등이다.
- [0078] 오디오 인코더(113)는, 음성 데이터 SA에 대하여, AAC, AC3, AC4, MPEGH(3D 오디오) 등의 압축 포맷에 의한 부호화를 실시하여, 오디오 스트림(오디오 엘리먼트리 스트림)을 생성한다. 음성 데이터 SA는, 상술한 화상 데이터 SV에 대응한 음성 데이터이며, HDD 등의 기록 매체로부터 재생된 음성 데이터, 또는 마이크로폰에 의해 얻어진 라이브 음성 데이터 등이다.
- [0079] 오디오 인코더(113)는, 오디오 부호화 블록부(113a) 및 오디오 프레임부(113b)를 갖고 있다. 오디오 부호화 블록부(113a)에서 부호화 블록이 생성되고, 오디오 프레임부(113b)에서 프레임이 행하여진다. 이 경우, 압축 포맷에 의해, 부호화 블록이 상이함과 함께, 프레임도 상이하다.
- [0080] 오디오 인코더(113)는, 제어부(111)에 의한 제어 하에서, 오디오 스트림에, 메타데이터 MD를 삽입한다. 이 실시 형태에 있어서, 메타데이터 MD는, 소정의 넷 서비스에 접속하기 위한 액세스 정보이다. 여기서, 소정의 넷 서비스로서는, 뮤직 넷 서비스나 오디오 비디오 넷 서비스 등의 모든 서비스가 대상이 된다. 여기서, 메타데이터 MD는, 오디오 스트림의 유저 데이터 영역에 매립된다.
- [0081] DASH/MP4 포맷터(114)는, 비디오 인코더(112)로부터 출력되는 비디오 스트림 및 오디오 인코더(113)로부터 출력되는 오디오 스트림에 기초하여, 콘텐츠인 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4를 생성한다. 또한, 콘텐츠 메타데이터나 세그먼트 URL 정보 등을 이용해서 MPD 파일을 생성한다. 여기서, MPD 파일에는, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보 등이 삽입된다(도 4 참조).
- [0082] 도 8에 도시하는 DASH/MP4 생성부(110)의 동작을 간단하게 설명한다. 화상 데이터 SV는 비디오 인코더(112)에 공급된다. 이 비디오 인코더(112)에서는, 그 화상 데이터 SV에 대하여 H.264/AVC, H.265/HEVC 등의 부호화가 실시되고, 부호화 비디오 데이터를 포함하는 비디오 스트림이 생성된다.
- [0083] 또한, 음성 데이터 SA는, 오디오 인코더(113)에 공급된다. 이 오디오 인코더(113)에서는, 그 음성 데이터 SA에 대하여, AAC, AC3, AC4, MPEGH(3D 오디오) 등의 부호화가 실시되고, 오디오 스트림이 생성된다.
- [0084] 이때, 제어부(111)로부터 오디오 인코더(113)에, 메타데이터 MD가 공급됨과 함께, 이 메타데이터 MD를 유저 데이터 영역에 매립하기 위한 사이즈 정보가 공급된다. 그리고, 오디오 인코더(113)에서는, 오디오 스트림의 유저 데이터 영역에, 메타데이터 MD를 매립하는 일이 행하여진다.
- [0085] 비디오 인코더(112)에서 생성된 비디오 스트림은, DASH/MP4 포맷터(114)에 공급된다. 또한, 오디오 인코더(113)에서 생성된, 유저 데이터 영역에 메타데이터 MD가 매립된 오디오 스트림은, DASH/MP4 포맷터(114)에 공급된다. 그리고, 이 DASH/MP4 포맷터(114)에서는, 콘텐츠인 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4가 생성된다. 또한, 이 DASH/MP4 포맷터(114)에서는, 콘텐츠 메타데이터나 세그먼트 URL 정보 등이 이용되어 MPD 파일이 생성된다. 이때, MPD 파일에는, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보 등이 삽입된다.
- [0086] [각 압축 포맷에 있어서의 메타데이터 MD의 삽입의 상세]
- [0087] 「AAC의 경우」
- [0088] 처음에, 압축 포맷이 AAC(Advanced Audio Coding)인 경우에 대해서 설명한다. 도 9는, AAC의 오디오 프레임의 구조를 나타내고 있다. 이 오디오 프레임은, 복수의 엘리먼트로 이루어져 있다. 각 엘리먼트(element)의 선두에는, 「id\_syn\_ele」의 3비트의 식별자(ID)가 존재하고, 엘리먼트 내용이 식별 가능하게 되어 있다.
- [0089] 「id\_syn\_ele」가 “0x4” 일 때, 유저 데이터를 둘 수 있는 엘리먼트인 DSE(Data Stream Element)임이 나타난다. 압축 포맷이 AAC인 경우, 이 DSE에 메타데이터 MD가 삽입된다. 도 10은, DSE(Data Stream Element())의 구성(Syntax)을 나타내고 있다.
- [0090] 「element\_instance\_tag」의 4비트 필드는, DSE 중의 데이터 중별을 나타내지만, DSE를 통일한 유저 데이터로서 이용하는 경우에는, 이 값을 “0”으로 해도 된다. 「Data\_byte\_align\_flag」는, “1”로 되고, DSE 전체가 바

이트 얼라인되도록 한다. 「count」 또는 그 추가 바이트수를 의미하는 「esc\_count」는, 유저 데이터의 사이즈에 따라 적절히 값이 결정된다. 「data\_stream\_byte」의 필드에, 「metadata()」가 삽입된다.

[0091] 도 11의 (a)는 「metadata()」의 구성(Syntax)을 나타내고, 도 11의 (b)는 그 구성에 있어서의 주요한 정보의 내용(semantics)을 나타내고 있다. 「userdata\_identifier」의 32비트 필드는, 미리 정의된 배열의 값이 세트됨으로써 오디오·유저 데이터임을 나타낸다. 「userdata\_identifier」가 “AAAA”로 유저 데이터를 나타내는 경우, 「metadata\_type」의 8비트 필드가 존재한다. 이 필드는, 메타데이터의 타입을 나타낸다. 예를 들어, “0x08”은, 소정의 넷 서비스에 접속하기 위한 액세스 정보이며, ATSC의 「SDO\_payload()」에 포함됨을 나타낸다. “0x08”일 때, 「SDO\_payload()」가 존재한다. 또한, 여기서는 “ATSC”로 했지만, 다른 규격화 단체에서 사용하는 것도 가능하다.

[0092] 도 12는, 「SDO\_payload()」의 구성(Syntax)을 나타내고 있다. 커맨드 ID(cmdID)가 “0x05”보다 작을 때, 「URI\_character」의 필드가 존재한다. 이 필드에, 소정의 넷 서비스에 접속하기 위한 URI 정보를 나타내는 문자코드가 삽입된다. 도 13은, 커맨드 ID(cmdID)의 값의 의미를 나타내고 있다. 또한, 이 「SDO\_payload()」는, ATSC(Advanced Television Systems Committee standards)로 규격화되어 있다.

[0093] 「AC3의 경우」

[0094] 이어서, 압축 포맷이 AC3인 경우에 대해서 설명한다. 도 14는, AC3의 프레임(AC3 Synchronization Frame)의 구조를 나타내고 있다. 「Audblock 5」의 “mantissa data”와, 「AUX」와, 「CRC」와의 합계 사이즈가 전체의 3/8을 초과하지 않도록, 음성 데이터 SA가 인코딩된다. 압축 포맷이 AC3인 경우, 「AUX」의 에리어에 메타데이터 MD가 삽입된다. 도 15는, AC3의 옥실러리 데이터(Auxiliary Data)의 구성(syntax)을 나타내고 있다.

[0095] 「auxdatae」가 “1”일 때, 「aux data」가 인에이블되고, 「auxdata1」의 14비트(비트 단위)로 나타나는 사이즈의 데이터가, 「auxbits」중에 정의된다. 그 때의 「auxbits」의 사이즈는 「nauxbits」에 기재된다. 본 기술에 있어서는, 「auxbits」의 필드가 「metadata()」라고 정의된다. 즉, 이 「auxbits」의 필드에, 상술한 도 11의 (a)에 도시하는 「metadata()」가 삽입되고, 그 「data\_byte」의 필드에, 도 11의 (a)에 도시하는 선택스 구조에 따라, 소정의 넷 서비스에 접속하기 위한 액세스 정보를 갖는, ATSC의 「SDO\_payload()」(도 12 참조)가 놓인다.

[0096] 「AC4의 경우」

[0097] 이어서, 압축 포맷이 AC4인 경우에 대해서 설명한다. 이 AC4는, AC3의 차세대 오디오 부호화 포맷의 하나로 되어 있다. 도 16의 (a)는, AC4의 심플 트랜스포트(Simple Transport)의 레이어의 구조를 나타내고 있다. 싱크워드(syncWord)의 필드와, 프레임 길이(frame Length)의 필드와, 부호화 데이터의 필드로서의 「RawAc4Frame」의 필드와, CRC 필드가 존재한다. 「RawAc4Frame」의 필드에는, 도 16의 (b)에 도시하는 바와 같이, 선두에 TOC(Table Of Content)의 필드가 존재하고, 그 후의 소정수의 서브 스트림(Substream)의 필드가 존재한다.

[0098] 도 17의 (b)에 도시하는 바와 같이, 서브 스트림(ac4\_substream\_data()) 안에는, 메타데이터 영역(metadata)이 존재하고, 그 안에 「umd\_payloads\_substream()」의 필드가 설치된다. 이 「umd\_payloads\_substream()」의 필드에, 소정의 넷 서비스에 접속하기 위한 액세스 정보를 갖는, ATSC의 「SDO\_payload()」(도 12 참조)가 놓인다.

[0099] 또한, 도 17의 (a)에 도시하는 바와 같이, TOC(ac4\_toc()) 안에는 「ac4\_presentation\_info()」의 필드가 존재하고, 또한 그 안에 「umd\_info()」의 필드가 존재하고, 그 안에 상술한 「umd\_payloads\_substream()」의 필드에 메타데이터의 삽입이 있음이 나타난다.

[0100] 도 18은, 「umd\_info()」의 구성(syntax)을 나타내고 있다. 「umd\_version」의 필드는, 버전 번호를 나타낸다. 「substream\_index」의 필드는, 인덱스값을 나타낸다. 버전 번호와 인덱스값이 있는 값의 조합이, 「umd\_payloads\_substream()」의 필드에 메타데이터의 삽입이 있음을 나타내는 것으로서 정의된다.

[0101] 도 19는, 「umd\_payloads\_substream()」의 구성(syntax)을 나타내고 있다. 「umd\_payload\_id」의 5비트 필드는, “0” 이외의 값으로 된다. 「umd\_userdata\_identifier」의 32비트 필드는, 미리 정의된 배열의 값이 세트됨으로써 오디오·유저 데이터임을 나타낸다. 「umd\_payload\_size」의 16비트 필드는, 이후의 바이트수를 나타낸다. 「umd\_userdata\_identifier」가 “AAAA”로 유저 데이터를 나타내는 경우, 「umd\_metadata\_type」의 8비트 필드가 존재한다. 이 필드는, 메타데이터의 타입을 나타낸다. 예를 들어, “0x08”은, 소정의 넷 서비스에 접속하기 위한 액세스 정보이며, ATSC의 「SDO\_payload()」에 포함됨을 나타낸다. “0x08”일 때, 「

SDO\_payload()」(도 12 참조)가 존재한다.

- [0102] 「MPEGH의 경우」
- [0103] 이어서, 압축 포맷이 MPEGH(3D 오디오)인 경우에 대해서 설명한다. 도 20은, MPEGH(3D 오디오)의 전송 데이터에 있어서의 오디오 프레임(1024 샘플)의 구조를 나타내고 있다. 이 오디오 프레임은, 복수의 MPEG 오디오 스트림 패킷(mpeg Audio Stream Packet)으로 이루어져 있다. 각 MPEG 오디오 스트림 패킷은, 헤더(Header)와 페이로드(Payload)에 의해 구성되어 있다.
- [0104] 헤더는, 패킷 타입(Packet Type), 패킷 라벨(Packet Label), 패킷 길이(Packet Length) 등의 정보를 갖는다. 페이로드에는, 헤더의 패킷 타입으로 정의된 정보가 배치된다. 이 페이로드 정보에는, 동기 스타트 코드에 상응하는 “SYNC”와, 3D 오디오의 전송 데이터의 실제 데이터인 “Frame”과, 이 “Frame”의 구성을 나타내는 “Config”가 존재한다.
- [0105] “Frame”에는, 3D 오디오의 전송 데이터를 구성하는 채널 부호화 데이터와 오브젝트 부호화 데이터가 포함된다. 여기서, 채널 부호화 데이터는, SCE(Single Channel Element), CPE(Channel Pair Element), LFE(Low Frequency Element) 등의 부호화 샘플 데이터로 구성된다. 또한, 오브젝트 부호화 데이터는, SCE(Single Channel Element)의 부호화 샘플 데이터와, 그것을 임의의 위치에 존재하는 스피커에 매핑시켜서 렌더링하기 위한 메타데이터에 의해 구성된다. 이 메타데이터는, 익스텐션 엘리먼트(Ext\_element)로서 포함된다.
- [0106] 여기서, “Config”에 포함되는 각 “Frame”의 구성 정보(config)와, 각 “Frame”과의 대응은, 이하와 같이 해서 유지된다. 즉, 도 21에 도시하는 바와 같이, “Config”에 각 “Frame”의 구성 정보(config)가 ID(elemIdx) 등록되지만, 각 “Frame”은 ID 등록된 순서로 전송된다. 또한, 패킷 라벨(PL)의 값은, “Config”와 그것에 대응한 각 “Frame”에서 동일한 값으로 된다.
- [0107] 도 20으로 되돌아가서, 이 실시 형태에서는, 익스텐션 엘리먼트(Ext\_element)로서 유저 데이터(userdata)가 포함되는 엘리먼트(Ext\_userdata)를 새롭게 정의한다. 이에 수반하여 “Config”에, 그 엘리먼트(Ext\_userdata)의 구성 정보(userdataConfig)를 새롭게 정의한다.
- [0108] 도 22는, 익스텐션 엘리먼트(Ext\_element)의 타입(ExElementType)과, 그 값(Value)과의 대응 관계를 나타내고 있다. 현 상황은, 0 내지 7이 정해져 있다. 128 이후는 MPEG 이외까지 확장 가능하므로, 예를 들어 128을 새롭게, “ID\_EXT\_ELE\_userdata” 타입의 값으로서 정의한다.
- [0109] 도 23은, 「userdataConfig()」의 구성(syntax)을 나타내고 있다. 「userdata\_identifier」의 32비트 필드는, 미리 정의된 배열의 값이 세트됨으로써 오디오·유저 데이터임을 나타낸다. 「userdata\_frameLength」의 16비트 필드는, 「audio\_userdata()」의 바이트수를 나타낸다. 도 24는, 「audio\_userdata()」의 구성(syntax)을 나타내고 있다. 「userdataConfig()」의 「userdata\_identifier」가 “AAAA”로 유저 데이터를 나타내는 경우, 「metadataType」의 8비트 필드가 존재한다. 이 필드는, 메타데이터의 타입을 나타낸다. 예를 들어, “0x08”은, 소정의 넷 서비스에 접속하기 위한 액세스 정보이며, ATSC의 「SDO\_payload()」에 포함됨을 나타낸다. “0x08”일 때, 「SDO\_payload()」(도 12 참조)가 존재한다.
- [0110] [셋톱 박스의 구성예]
- [0111] 도 25는, 셋톱 박스(200)의 구성예를 나타내고 있다. 이 셋톱 박스(200)는, 수신부(204)와, DASH/MP4 해석부(205)와, 비디오 디코더(206)와, 오디오 프레이밍부(207)와, HDMI 송신부(208)와, HDMI 단자(209)를 갖고 있다. 또한, 셋톱 박스(200)는, CPU(211)과, 플래시 ROM(212)과, DRAM(213)과, 내부 버스(214)와, 리모콘 수신부(215)와, 리모콘 송신기(216)를 갖고 있다.
- [0112] CPU(211)는, 셋톱 박스(200)의 각 부의 동작을 제어한다. 플래시 ROM(212)은, 제어 소프트웨어의 저장 및 데이터의 보관을 행한다. DRAM(213)은, CPU(211)의 워크메모리를 구성한다. CPU(211)는, 플래시 ROM(212)으로부터 판독한 소프트웨어나 데이터를 DRAM(213) 상에 전개해서 소프트웨어를 기동시켜, 셋톱 박스(200)의 각 부를 제어한다.
- [0113] 리모콘 수신부(215)는, 리모콘 송신기(216)로부터 송신된 리모콘 컨트롤 신호(리모콘 코드)를 수신하여, CPU(211)에 공급한다. CPU(211)는, 이 리모콘 코드에 기초하여, 셋톱 박스(200)의 각 부를 제어한다. CPU(211), 플래시 ROM(212) 및 DRAM(213)은 내부 버스(214)에 접속되어 있다.
- [0114] 수신부(204)는, 서비스 송신 시스템(100)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는

DASH/MP4, 즉 메타파일로서의 MPD 파일과, 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4를 수신한다. MP4에 포함되는 오디오 스트림에는, 메타데이터로서, 소정의 넷 서비스에 접속하기 위한 액세스 정보가 삽입되어 있다. 또한, MPD 파일에는, “Supplementary Descriptor”에 의해, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보가 삽입되어 있다.

- [0115] DASH/MP4 해석부(205)는, 수신부(204)에서 수신된 MPD 파일 및 MP4를 해석한다. DASH/MP4 해석부(205)는, MPD 파일에 포함되는 MPD 정보를 추출하여, CPU(211)에 보낸다. 여기서, 이 MPD 정보에는, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보 등도 포함된다. CPU(211)는, 이 MPD 정보에 기초하여, 비디오, 오디오의 스트림 취득 처리를 제어한다. 또한, DASH/MP4 해석부(205)는, MP4로부터 메타데이터, 예를 들어 각 트랙의 헤더 정보나 콘텐츠 내용의 메타 기술, 시각 정보 등을 추출하여, CPU(211)에 보낸다.
- [0116] DASH/MP4 해석부(205)는 MP4로부터 비디오 스트림을 추출하여, 비디오 디코더(206)에 보낸다. 비디오 디코더(206)는 비디오 스트림에 복호화 처리를 행해서 비압축 화상 데이터를 얻는다. 또한, DASH/MP4 해석부(205)는, MP4로부터 오디오 스트림을 추출하여, 오디오 프레이밍부(207)에 보낸다. 오디오 프레이밍부(207)는, 오디오 스트림에 대하여 프레이밍을 행한다.
- [0117] HDMI 송신부(208)는 HDMI에 준거한 통신에 의해, 비디오 디코더(206)에서 얻어진 비압축 화상 데이터와, 오디오 프레이밍부(207)에서 프레이밍된 후의 오디오 스트림을, HDMI 단자(209)로부터 송출한다. HDMI 송신부(208)는, HDMI의 TMDS 채널을 통해 송신하기 위해서, 화상 데이터 및 오디오 스트림을 패키징하고, HDMI 단자(209)에 출력한다.
- [0118] HDMI 송신부(208)는, CPU(211)의 제어 하에서, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 삽입한다. HDMI 송신부(208)는, 오디오 스트림 및 식별 정보를 화상 데이터의 블랭킹 기간에 삽입한다. 이 HDMI 송신부(209)의 상세는 후술한다.
- [0119] 이 실시 형태에 있어서, HDMI 송신부(208)는, 화상 데이터의 블랭킹 기간에 배치되는 오디오·인포 프레임·패킷(Audio InfoFrame packet)에 식별 정보를 삽입한다. 이 오디오·인포 프레임·패킷은, 데이터 아일랜드 구간 에 배치된다.
- [0120] 도 26은, 오디오·인포 프레임·패킷의 구조예를 나타내고 있다. HDMI에서는, 이 오디오·인포 프레임·패킷에 의해, 음성에 관한 부대 정보를 소스 기기로부터 싱크 기기로 전송 가능하게 되어 있다.
- [0121] 제0 바이트에 데이터 패킷의 종류를 나타내는 「Packet Type」이 정의되어 있고, 오디오·인포 프레임·패킷은 「0x84」로 되어 있다. 제1 바이트에 패킷 데이터 정의의 버전 정보를 기술한다. 제2 바이트에, 패킷 길이를 나타내는 정보를 기술한다. 이 실시 형태에서는, 제5 바이트의 제5 비트에, 「userdata\_presence\_flag」의 1비트 플래그 정보가 정의된다. 플래그 정보가 “1”일 때, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타낸다.
- [0122] 또한, 플래그 정보가 “1”일 때, 제9 바이트에 여러 가지 정보가 정의된다. 제7 비트 내지 제5 비트는 「metadata\_type」의 필드로 되고, 제4 비트는 「coordinated\_control\_flag」의 필드로 되고, 제2 비트 내지 제0 비트는 「frequency\_type」의 필드로 된다. 상세한 설명은 생략하지만, 이들 각 필드는, 도 4에 도시하는 MPD 파일에 부가되어 있는 각 정보와 동일한 정보를 나타낸다.
- [0123] 셋톱 박스(200)의 동작을 간단하게 설명한다. 수신부(204)에서는, 서비스 송신 시스템(100)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 DASH/MP4, 즉 메타파일로서의 MPD 파일과, 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4가 수신된다. 이렇게 수신되는 MPD 파일 및 MP4는, DASH/MP4 해석부(205)에 공급된다.
- [0124] DASH/MP4 해석부(205)에서는, MPD 파일 및 MP4가 해석된다. 그리고, DASH/MP4 해석부(205)에서는, MPD 파일에 포함되는 MPD 정보가 추출되어, CPU(211)에 보내진다. 여기서, 이 MPD 정보에는, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보도 포함된다. 또한, DASH/MP4 해석부(205)에서는, MP4로부터 메타데이터, 예를 들어 각 트랙의 헤더 정보나 콘텐츠 내용의 메타 기술, 시각 정보 등이 추출되어, CPU(211)에 보내진다.
- [0125] 또한, DASH/MP4 해석부(205)에서는, MP4로부터 비디오 스트림이 추출되어, 비디오 디코더(206)에 보내진다. 비디오 디코더(206)에서는, 비디오 스트림에 복호화 처리가 실시되어 비압축 화상 데이터가 얻어진다. 이 화상 데이터는, HDMI 송신부(208)에 공급된다. 또한, DASH/MP4 해석부(205)에서는, MP4로부터 오디오 스트림이 추출



된다. 이 오디오 스트림은 오디오 프레이밍부(207)에서 프레이밍된 후에, HDMI 송신부(208)에 공급된다. 그리고, HDMI 송신부(208)에서는, 화상 데이터 및 오디오 스트림이 패키징되어, HDMI 단자(209)로부터 HDMI 케이블(400)에 송출된다.

- [0126] HDMI 송신부(208)에서는, CPU(211)의 제어 하에서, 화상 데이터의 블랭킹 기간에 배치되는 오디오·인포 프레임·패킷에, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보가 삽입된다. 이에 의해, 셋톱박스(200)로부터 HDMI 텔레비전 수신기(300)에, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 송신하는 일이 행하여진다.
- [0127] [텔레비전 수신기의 구성예]
- [0128] 도 27은, 텔레비전 수신기(300)의 구성예를 나타내고 있다. 이 텔레비전 수신기(300)는, 수신부(306)와, DASH/MP4 해석부(307)와, 비디오 디코더(308)와, 영상 처리 회로(309)와, 패널 구동 회로(310)와, 표시 패널(311)을 갖고 있다.
- [0129] 또한, 텔레비전 수신기(300)는, 오디오 디코더(312)와, 음성 처리 회로(313)와, 음성 증폭 회로(314)와, 스피커(315)와, HDMI 단자(316)와, HDMI 수신부(317)와, 통신 인터페이스(318)를 갖고 있다. 또한, 텔레비전 수신기(300)는, CPU(321)와, 플래시 ROM(322)과, DRAM(323)과, 내부 버스(324)와, 리모콘 수신부(325)와, 리모콘 송신기(326)를 갖고 있다.
- [0130] CPU(321)는, 텔레비전 수신기(300)의 각 부의 동작을 제어한다. 플래시 ROM(322)은, 제어 소프트웨어의 저장 및 데이터의 보관을 행한다. DRAM(323)은, CPU(321)의 워크메모리를 구성한다. CPU(321)는, 플래시 ROM(322)으로부터 판독한 소프트웨어나 데이터를 DRAM(323) 상에 전개해서 소프트웨어를 기동시켜, 텔레비전 수신기(300)의 각 부를 제어한다.
- [0131] 리모콘 수신부(325)는, 리모콘 송신기(326)로부터 송신된 리모콘 컨트롤 신호(리모콘 코드)를 수신하여, CPU(321)에 공급한다. CPU(321)는, 이 리모콘 코드에 기초하여, 텔레비전 수신기(300)의 각 부를 제어한다. CPU(321), 플래시 ROM(322) 및 DRAM(323)은, 내부 버스(324)에 접속되어 있다.
- [0132] 통신 인터페이스(318)는, CPU(321)의 제어 하에서, 인터넷 등의 네트워크 상에 존재하는 서버 사이에서 통신을 행한다. 이 통신 인터페이스(318)는, 내부 버스(324)에 접속되어 있다.
- [0133] 수신부(306)는, 서비스 송신 시스템(100)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 DASH/MP4, 즉 메타파일로서의 MPD 파일과, 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4를 수신한다. MP4에 포함되는 오디오 스트림에는, 메타데이터로서, 소정의 넷 서비스에 접속하기 위한 액세스 정보가 삽입되어 있다. 또한, MPD 파일에는, “Supplementary Descriptor”에 의해, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보 등이 삽입되어 있다.
- [0134] DASH/MP4 해석부(307)는, 수신부(306)에서 수신된 MPD 파일 및 MP4를 해석한다. DASH/MP4 해석부(307)은 MPD 파일에 포함되는 MPD 정보를 추출하여, CPU(321)에 보낸다. CPU(321)는, 이 MPD 정보에 기초하여, 비디오, 오디오의 스트림 취득 처리를 제어한다. 또한, DASH/MP4 해석부(307)는, MP4로부터 메타데이터, 예를 들어 각 트랙의 헤더 정보나 콘텐츠 내용의 메타 기술, 시각 정보 등을 추출하여, CPU(321)에 보낸다.
- [0135] DASH/MP4 해석부(307)는, MP4로부터 비디오 스트림을 추출하여, 비디오 디코더(308)에 보낸다. 비디오 디코더(308)는, 비디오 스트림에 복호화 처리를 행해서 비압축 화상 데이터를 얻는다. 또한, DASH/MP4 해석부(307)는, MP4로부터 오디오 스트림을 추출하여, 오디오 디코더(312)에 보낸다.
- [0136] HDMI 수신부(317)는 HDMI에 준거한 통신에 의해, HDMI 케이블(400)을 통해서 HDMI 단자(316)에 공급되는 화상 데이터 및 오디오 스트림을 수신한다. 또한, HDMI 수신부(317)는, 화상 데이터의 블랭킹 기간에 삽입되어 있는 여러 가지 제어 정보를 추출하여, CPU(321)에 송신한다. 여기서, 이 제어 정보에는, 오디오·인포 프레임·패킷에 삽입된, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보 등(도 26 참조)도 포함된다. 이 HDMI 수신부(317)의 상세는 후술한다.
- [0137] 영상 처리 회로(309)는, 비디오 디코더(308)에서 얻어진, 또는 HDMI 수신부(316)에서 얻어진 화상 데이터, 나아가서는, 통신 인터페이스(318)로 넷 상의 서버로부터 수신된 화상 데이터 등에 대하여 스케일링 처리, 합성 처리 등을 행하여, 표시용 화상 데이터를 얻는다.
- [0138] 패널 구동 회로(310)는, 영상 처리 회로(308)에서 얻어지는 표시용 화상 데이터에 기초하여, 표시 패널(311)을

구동한다. 표시 패널(311)은, 예를 들어 LCD(Liquid Crystal Display), 유기 EL 디스플레이(organic electroluminescence display) 등으로 구성되어 있다.

- [0139] 오디오 디코더(312)는, DASH/MP4 해석부(307)에서 추출된, 또는 HDMI 수신부(317)에서 얻어진 오디오 스트림에 대하여 복호화 처리를 행해서 비압축 음성 데이터를 얻는다. 또한, 오디오 디코더(312)는, CPU(321)의 제어 하에서, 오디오 스트림에 삽입되어 있는 메타데이터를 추출하여, CPU(321)에 보낸다. 이 실시 형태에 있어서, 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보이다(도 12 참조). CPU(321)는, 적절히, 텔레비전 수신기(300)의 각 부에 메타데이터를 사용한 처리를 행하게 한다.
- [0140] 또한, CPU(321)에는 DASH/MP4 해석부(307)로부터 MPD 정보가 공급된다. CPU(321)는, 이 MPD 정보에 포함되는 식별 정보에 의해, 오디오 스트림에 메타데이터가 삽입되어 있음을 미리 인식할 수 있어, 메타데이터의 추출이 행해지도록 오디오 디코더(312)를 제어할 수 있다. 음성 처리 회로(313)는, 오디오 디코더(312)에서 얻어진 음성 데이터에 대하여 D/A 변환 등의 필요한 처리를 행한다. 음성 증폭 회로(314)는, 음성 처리 회로(313)로부터 출력되는 음성 신호를 증폭해서 스피커(315)에 공급한다.
- [0141] 도 27에 도시하는 텔레비전 수신기(300)의 동작을 간단하게 설명한다. 수신부(306)에서는, 서비스 송신 시스템(100)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 DASH/MP4, 즉 메타파일로서의 MPD 파일과, 비디오나 오디오 등의 미디어 스트림(미디어 세그먼트)이 포함되는 MP4가 수신된다. 이렇게 수신되는 MPD 파일 및 MP4는, DASH/MP4 해석부(307)에 공급된다.
- [0142] DASH/MP4 해석부(307)에서는, MPD 파일 및 MP4가 해석된다. 그리고, DASH/MP4 해석부(307)에서는, MPD 파일에 포함되는 MPD 정보가 추출되어, CPU(321)에 보내진다. 여기서, 이 MPD 정보에는, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보도 포함된다. 또한, DASH/MP4 해석부(307)에서는, MP4로부터 메타데이터, 예를 들어 각 트랙의 헤더 정보나 콘텐츠 내용의 메타 기술, 시각 정보 등이 추출되어, CPU(321)에 보내진다.
- [0143] 또한, DASH/MP4 해석부(307)에서는, MP4로부터 비디오 스트림이 추출되어, 비디오 디코더(308)에 보내진다. 비디오 디코더(308)에서는, 비디오 스트림에 복호화 처리가 실시되어 비압축 화상 데이터가 얻어진다. 이 화상 데이터는, 영상 처리 회로(309)에 공급된다. 또한, DASH/MP4 해석부(307)에서는, MP4로부터 오디오 스트림이 추출된다. 이 오디오 스트림은, 오디오 디코더(312)에 공급된다.
- [0144] HDMI 수신부(317)에서는, HDMI에 준거한 통신에 의해, HDMI 케이블(400)을 통해서 HDMI 단자(316)에 공급되는 화상 데이터 및 오디오 스트림이 수신된다. 화상 데이터는, 영상 처리 회로(309)에 공급된다. 또한, 오디오 스트림은 오디오 디코더(312)에 공급된다.
- [0145] 또한, HDMI 수신부(317)에서는, 화상 데이터의 블랭킹 기간에 삽입되어 있는 여러 가지 제어 정보가 추출되어, CPU(321)에 보내진다. 이 제어 정보에는, 오디오·인포 프레임·패킷에 삽입된, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보 등도 포함된다. 그로 인해, CPU(321)는, 이 식별 정보에 기초하여, 오디오 디코더(312)의 동작을 제어하고, 오디오 스트림으로부터 메타데이터를 추출시키는 것이 가능하게 된다.
- [0146] 영상 처리 회로(309)에서는, 비디오 디코더(308)에서 얻어진, 또는 HDMI 수신부(317)에서 얻어진 화상 데이터, 나아가서는, 통신 인터페이스(318)로 넷 상의 서버로부터 수신된 화상 데이터 등에 대하여 스케일링 처리, 합성 처리 등이 실시되어, 표시용 화상 데이터가 얻어진다. 여기서, 텔레비전 방송 신호를 수신해서 처리하는 경우에는, 영상 처리 회로(309)에서는, 비디오 디코더(308)에서 얻어진 화상 데이터가 취급된다. 한편, 셋톱 박스(200)가 HDMI 인터페이스로 접속되는 경우에는, 영상 처리 회로(309)에서는, HDMI 수신부(317)에서 얻어진 화상 데이터가 취급된다.
- [0147] 영상 처리 회로(309)에서 얻어진 표시용 화상 데이터는 패널 구동 회로(310)에 공급된다. 패널 구동 회로(310)에서는, 표시용 화상 데이터에 기초하여, 표시 패널(311)을 구동하는 일이 행하여진다. 이에 의해, 표시 패널(311)에는, 표시용 화상 데이터에 대응한 화상이 표시된다.
- [0148] 오디오 디코더(312)에서는, DASH/MP4 해석부(307)에서 얻어진, 또는 HDMI 수신부(316)에서 얻어진 오디오 스트림에 대하여 복호화 처리가 행하여져서 비압축 음성 데이터가 얻어진다. 여기서, 텔레비전 방송 신호를 수신해서 처리하는 경우에는, 오디오 디코더(312)에서는, DASH/MP4 해석부(307)에서 얻어진 오디오 스트림이 취급된다. 한편, 셋톱 박스(200)가 HDMI 인터페이스로 접속되는 경우에는, 오디오 디코더(312)에서는, HDMI 수신부(317)에서 얻어진 오디오 스트림이 취급된다.

- [0149] 오디오 디코더(312)에서 얻어지는 음성 데이터는, 음성 처리 회로(313)에 공급된다. 음성 처리 회로(313)에서는, 음성 데이터에 대하여 D/A 변환 등의 필요한 처리가 실시된다. 이 음성 데이터는, 음성 증폭 회로(314)에서 증폭된 후에, 스피커(315)에 공급된다. 그로 인해, 스피커(315)로부터, 표시 패널(311)의 표시 화상에 대응한 음성이 출력된다.
- [0150] 또한, 오디오 디코더(312)에서는, 오디오 스트림에 삽입되어 있는 메타데이터가 추출된다. 예를 들어, 이 메타데이터의 추출 처리는, 상술한 바와 같이, CPU(321)가, 오디오 스트림에 메타데이터가 삽입되어 있음을 식별 정보에 기초하여 파악하여, 오디오 디코더(312)의 동작을 제어함으로써, 낭비 없이 확실하게 행하여진다.
- [0151] 이렇게 오디오 디코더(312)에서 추출되는 메타 데이터는, CPU(321)에 보내진다. 그리고, CPU(321)의 제어에 의해, 적절히, 텔레비전 수신기(300)의 각 부에서 메타데이터를 사용한 처리가 행하여진다. 예를 들어, 네트워크 상의 서버로부터 화상 데이터를 취득하고, 멀티 화면 표시가 행하여진다.
- [0152] [HDMI 송신부, HDMI 수신부의 구성예]
- [0153] 도 28은, 도 25에 도시하는 셋톱 박스(200)의 HDMI 송신부(HDMI 소스)(208)와, 도 27에 도시하는 텔레비전 수신기(300)의 HDMI 수신부(HDMI 싱크)(317)의 구성예를 나타내고 있다.
- [0154] HDMI 송신부(208)는, 유효 화상 구간(이하, 적절히, 액티브 비디오 구간이라고도 함)에 있어서, 비압축의 1화면분의 화상의 화소 데이터에 대응하는 차동 신호를, 복수의 채널을 통해, HDMI 수신부(317)에 일 방향으로 송신한다. 여기서, 유효 화상 구간은, 어떤 수직 동기 신호로부터 다음 수직 동기 신호까지의 구간으로부터, 수평 귀선 구간 및 수직 귀선 구간을 제외한 구간이다. 또한, HDMI 송신부(208)는, 수평 귀선 구간 또는 수직 귀선 구간에 있어서, 적어도 화상에 부수되는 음성 데이터나 제어 데이터, 그 밖의 보조 데이터 등에 대응하는 차동 신호를, 복수의 채널을 통해, HDMI 수신부(317)에 일 방향으로 송신한다.
- [0155] HDMI 송신부(208)와 HDMI 수신부(317)를 포함하는 HDMI 시스템의 전송 채널에는, 이하의 전송 채널이 있다. 즉, HDMI 송신부(208)로부터 HDMI 수신부(317)에 대하여 화소 데이터 및 음성 데이터를, 픽셀 클럭에 동기하여, 일 방향으로 시리얼 전송하기 위한 전송 채널로서의, 3개의 TMDS 채널#0 내지 #2가 있다. 또한, 픽셀 클럭을 전송하는 전송 채널로서의, TMDS 클럭 채널이 있다.
- [0156] HDMI 송신부(208)는, HDMI 트랜스미터(81)를 갖는다. 트랜스미터(81)는, 예를 들어 비압축 화상 화소 데이터를 대응하는 차동 신호로 변환하고, 복수의 채널인 3개의 TMDS 채널#0, #1, #2를 통해, HDMI 케이블(400)을 통해서 접속되어 있는 HDMI 수신부(317)에, 일 방향으로 시리얼 전송한다.
- [0157] 또한, 트랜스미터(81)는, 비압축 화상에 부수되는 음성 데이터, 나아가서는, 필요한 제어 데이터 그 밖의 보조 데이터 등을, 대응하는 차동 신호로 변환하고, 3개의 TMDS 채널#0, #1, #2로 HDMI 수신부(317)에, 일 방향으로 시리얼 전송한다.
- [0158] 또한, 트랜스미터(81)는, 3개의 TMDS 채널#0, #1, #2를 통해 송신하는 화소 데이터에 동기한 픽셀 클럭을, TMDS 클럭 채널을 통해, HDMI 케이블(400)을 통해서 접속되어 있는 HDMI 수신부(317)에 송신한다. 여기서, 1개의 TMDS 채널#i(i=0, 1, 2)에서는, 픽셀 클럭의 1클럭 사이에, 10비트의 화소 데이터가 송신된다.
- [0159] HDMI 수신부(317)는, 액티브 비디오 구간에 있어서, 복수의 채널을 통해, HDMI 송신부(208)부터 일 방향으로 송신되어 오는, 화소 데이터에 대응하는 차동 신호를 수신한다. 또한, 이 HDMI 수신부(317)는, 수평 귀선 구간 또는 수직 귀선 구간에 있어서, 복수의 채널을 통해, 송신부(208)부터 일 방향으로 송신되어 오는, 음성 데이터나 제어 데이터에 대응하는 차동 신호를 수신한다.
- [0160] 즉, HDMI 수신부(317)는 HDMI 리시버(82)를 갖는다. 이 HDMI 리시버(82)는, TMDS 채널#0, #1, #2를 통해, HDMI 송신부(208)부터 일 방향으로 송신되어 오는, 화소 데이터에 대응하는 차동 신호와, 음성 데이터나 제어 데이터에 대응하는 차동 신호를 수신한다. 이 경우, HDMI 송신부(208)로부터 TMDS 클럭 채널을 통해 송신되어 오는 픽셀 클럭에 동기해서 수신한다.
- [0161] HDMI 시스템의 전송 채널에는, 상술한 TMDS 채널#0 내지 #2 및 TMDS 클럭 채널 이외에, DDC(Display Data Channel)(83)나 CEC 라인(84)이라고 불리는 전송 채널이 있다. DDC(83)는, HDMI 케이블(400)에 포함되는 도시하지 않은 2개의 신호선으로 이루어진다. DDC(83)는, HDMI 송신부(208)가, HDMI 수신부(317)로부터, E-EDID(Enhanced Extended Display Identification Data)를 판독하기 위해서 사용된다.
- [0162] HDMI 수신부(317)는, HDMI 리시버(81) 이외에, 자신의 성능(Configuration/capability)에 관한 성능 정보인 E-

EDID를 기억하고 있는, EDID ROM(Read Only Memory)(85)을 갖고 있다. HDMI 송신부(208)는, 예를 들어 CPU(211)(도 20 참조)로부터의 요구에 따라, HDMI 케이블(400)을 통해서 접속되어 있는 HDMI 수신부(317)로부터, E-EDID를, DDC(83)를 통해서 판독한다.

- [0163] HDMI 송신부(208)는 판독한 E-EDID를 CPU(211)에 보낸다. CPU(211)는, 이 E-EDID를, 플래시 ROM(212) 또는 DRAM(213)에 저장한다.
- [0164] CEC 라인(84)은, HDMI 케이블(400)에 포함되는 도시하지 않은 1개의 신호선으로 이루어지고, HDMI 송신부(208)와 HDMI 수신부(317) 사이에서, 제어용 데이터의 쌍방향 통신을 행하기 위해서 사용된다. 이 CEC 라인(84)은, 제어 데이터 라인을 구성하고 있다.
- [0165] 또한, HDMI 케이블(400)에는, HPD(Hot Plug Detect)라고 불리는 핀에 접속되는 라인(HPD 라인)(86)이 포함되어 있다. 소스 기기는, 그 라인(86)을 이용하여, 싱크 기기의 접속을 검출할 수 있다. 또한, 이 HPD 라인(86)은, 쌍방향 통신로를 구성하는 HEAC- 라인으로서도 사용된다. 또한, HDMI 케이블(400)에는, 소스 기기로부터 싱크 기기에 전원을 공급하기 위해서 사용되는 전원 라인(87)이 포함되어 있다. 또한, HDMI 케이블(400)에는, 유틸리티 라인(88)이 포함되어 있다. 이 유틸리티 라인(88)은 쌍방향 통신로를 구성하는 HEAC+ 라인으로서도 사용된다.
- [0166] 도 29는, TMDS 채널#0, #1, #2에 있어서, 가로×세로가 1920 픽셀×1080 라인인 화상 데이터가 전송되는 경우의, 각종 전송 데이터의 구간을 나타내고 있다. HDMI의 3개의 TMDS 채널#0, #1, #2를 통해 전송 데이터가 전송되는 비디오 필드(Video Field)에는, 전송 데이터의 종류에 따라, 비디오 데이터 구간(17)(Video Data Period), 데이터 아일랜드 구간(18)(Data Island Period) 및 컨트롤 구간(19)(Control Period)의 3종류의 구간이 존재한다.
- [0167] 여기서, 비디오 필드 구간은, 어떤 수직 동기 신호의 상승 엣지(Active Edge)로부터 다음의 수직 동기 신호의 상승 엣지까지의 구간이며, 수평 귀선 기간(15)(Horizontal Blanking), 수직 귀선 기간(16)(Vertical Blanking) 및 비디오 필드 구간으로부터, 수평 귀선 기간 및 수직 귀선 기간을 제외한 구간인 유효 화소 구간(14)(Active Video)으로 나뉜다.
- [0168] 비디오 데이터 구간(17)은 유효 화소 구간(14)에 할당된다. 이 비디오 데이터 구간(17)에서는, 비압축의 1화면분의 화상 데이터를 구성하는 1920 픽셀(화소)×1080 라인분의 유효 화소(Active Pixel)의 데이터가 전송된다. 데이터 아일랜드 구간(18) 및 컨트롤 구간(19)은, 수평 귀선 기간(15) 및 수직 귀선 기간(16)에 할당된다. 이 데이터 아일랜드 구간(18) 및 컨트롤 구간(19)에서는, 보조 데이터(Auxiliary Data)가 전송된다.
- [0169] 즉, 데이터 아일랜드 구간(18)은, 수평 귀선 기간(15)과 수직 귀선 기간(16)의 일부분에 할당되어 있다. 이 데이터 아일랜드 구간(18)에서는, 보조 데이터 중, 제어에 관계되지 않는 데이터인, 예를 들어 음성 데이터의 패킷 등이 전송된다. 컨트롤 구간(19)은, 수평 귀선 기간(15)과 수직 귀선 기간(16)의 다른 부분에 할당되어 있다. 이 컨트롤 구간(19)에서는, 보조 데이터 중의, 제어에 관계되는 데이터인, 예를 들어 수직 동기 신호 및 수평 동기 신호, 제어 패킷 등이 전송된다.
- [0170] 이어서, 도 30을 참조하여, 텔레비전 수신기(300)에 있어서의 메타데이터를 사용한 처리의 구체예를 설명한다. 텔레비전 수신기(300)는, 예를 들어 메타데이터로서, 초기 서버 URL, 네트워크·서비스 식별 정보, 대상 파일명, 세션 개시·종료 커맨드, 미디어 기록·재생 커맨드 등을 취득한다. 또한, 상술에서는, 메타데이터가 소정의 넷 서비스에 접속하기 위한 액세스 정보라는 내용을 설명했지만, 여기서는 그 밖의 필요한 정보도 메타데이터에 포함되어 있는 것으로 한다.
- [0171] 네트워크 클라이언트인 텔레비전 수신기(300)는, 초기 서버 URL을 사용하여, 프라이머리 서버에 액세스한다. 그리고, 텔레비전 수신기(300)는, 프라이머리 서버로부터, 스트리밍 서버 URL, 대상 파일명, 파일의 타입을 나타내는 마임 타입, 미디어 재생시간 정보 등의 정보를 취득한다.
- [0172] 그리고, 텔레비전 수신기(300)는, 스트리밍 서버 URL을 사용하여, 스트리밍 서버에 액세스한다. 그리고, 텔레비전 수신기(300)는 대상 파일명을 지정한다. 여기서, 멀티캐스트에서의 서비스를 받는 경우에는, 네트워크 식별 정보와 서비스 식별 정보로 프로그램의 서비스를 특정한다.
- [0173] 그리고, 텔레비전 수신기(300)는, 세션 개시·종료 커맨드에 의해, 스트리밍 서버와의 사이의 세션을 개시하거나, 또는 그 세션을 종료한다. 또한, 텔레비전 수신기(300)는, 스트리밍 서버와의 사이의 세션 수행 중에, 미디어 기록·재생 커맨드를 사용하여, 스트리밍 서버로부터 미디어 데이터를 취득한다.

- [0174] 또한, 도 30의 예에서는, 프라이머리 서버와 스트리밍 서버가 별개로 존재하고 있다. 그러나, 이들 서버는, 일체적으로 구성되어 있어도 된다.
- [0175] 도 31은, 텔레비전 수신기(300)에서 메타데이터에 기초하여 넷 서비스에 액세스하는 경우의 화면 표시의 천이예를 나타내고 있다. 도 31의 (a)는 표시 패널(311)에 화상이 표시되지 않은 상태를 나타낸다. 도 31의 (b)는 방송 수신에 개시되고, 이 방송 수신에 관한 메인 콘텐츠가, 표시 패널(311)에, 전체 화면 표시되어 있는 상태를 나타낸다.
- [0176] 도 31의 (c)는, 메타데이터에 의한 서비스에의 액세스가 있고, 텔레비전 수신기(300)와 서버 사이의 세션이 개시된 상태를 나타낸다. 이 경우, 방송 수신에 관한 메인 콘텐츠가 전체 화면 표시로부터 부분 화면 표시로 된다.
- [0177] 도 31의 (d)는, 서버로부터의 미디어 재생이 행하여지고, 표시 패널(311)에, 메인 콘텐츠의 표시와 병행하여, 넷 서비스 콘텐츠(1)가 표시된 상태를 나타낸다. 그리고, 도 31의 (e)는, 서버로부터의 미디어 재생이 행하여지고, 표시 패널(311)에, 메인 콘텐츠의 표시와 병행하여, 넷 서비스 콘텐츠(1)의 표시와 함께, 넷 서비스 콘텐츠(2)가 메인 콘텐츠 표시 상에 중첩 표시된 상태를 나타낸다.
- [0178] 도 31의 (f)는, 넷으로부터의 서비스 콘텐츠의 재생이 종료하고, 텔레비전 수신기(300)와 서버 사이의 세션이 종료한 상태를 나타낸다. 이 경우, 표시 패널(311)에, 방송 수신에 관한 메인 콘텐츠가 전체 화면 표시되는 상태로 복귀된다.
- [0179] 또한, 도 27에 도시하는 텔레비전 수신기(300)는, 스피커(315)를 구비하고, 도 32에 도시하는 바와 같이, 오디오 디코더(312)에서 얻어진 음성 데이터가 음성 처리 회로(313) 및 음성 증폭 회로(314)를 통해서 스피커(315)에 공급되고, 이 스피커(315)로부터 음성이 출력되는 구성으로 되어 있다.
- [0180] 그러나, 도 33에 도시하는 바와 같이, 텔레비전 수신기(300)는 스피커를 구비하지 않고, DASH/MP4 해석부(307) 또는 HDMI 수신부(317)에서 얻어지는 오디오 스트림을 인터페이스부(331)로부터 외부 스피커 시스템(350)에 공급하는 구성도 생각된다. 인터페이스부(331)는, 예를 들어 HDMI(High-Definition Multimedia Interface), SPDIF(Sony Philips Digital Interface), MHL(Mobile High-definition Link) 등의 디지털 인터페이스이다.
- [0181] 이 경우, 오디오 스트림은, 외부 스피커 시스템(350)이 갖는 오디오 디코더(351a)에서 복호화 처리가 행하여지고, 이 외부 스피커 시스템(350)으로부터 음성이 출력된다. 또한, 텔레비전 수신기(300)가 스피커(315)를 구비하는 경우(도 32 참조)라도, 오디오 스트림을 인터페이스부(331)로부터 외부 스피커 시스템(350)에 공급하는 구성(도 33 참조)도 더 고려된다.
- [0182] 상술한 바와 같이, 도 3의 (a), (b)에 도시하는 송수신 시스템(10, 10')에 있어서, 서비스 송신 시스템(100)은, MPD 파일에, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 삽입한다. 그로 인해, 수신측(셋톱 박스(200), 텔레비전 수신기(300))에서는, 오디오 스트림에 메타데이터가 삽입되어 있음을 용이하게 인식할 수 있다.
- [0183] 또한, 도 3의 (a)에 도시하는 송수신 시스템(10)에 있어서, 셋톱 박스(200)는, 메타데이터가 삽입된 오디오 스트림을, 이 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 텔레비전 수신기(300)에, HDMI로 송신한다. 그로 인해, 텔레비전 수신기(300)에서는, 오디오 스트림에 메타데이터가 삽입되어 있음을 용이하게 인식할 수 있고, 이 인식에 기초하여 오디오 스트림에 삽입되어 있는 메타데이터의 추출 처리를 행함으로써, 낭비 없이 확실하게 메타데이터를 취득해서 이용할 수 있다.
- [0184] 또한, 도 3의 (b)에 도시하는 송수신 시스템(10')에 있어서, 텔레비전 수신기(300)는, MPD 파일에 삽입되어 있는 식별 정보에 기초하여, 오디오 스트림으로부터 메타데이터를 추출해서 처리에 사용한다. 그로 인해, 오디오 스트림에 삽입되어 있는 메타데이터를 낭비 없이 확실하게 취득할 수 있어, 메타데이터를 사용한 처리를 적절하게 실행할 수 있다.
- [0185] <2. 변형예>
- [0186] 또한, 상술한 실시 형태에 있어서, 송수신 시스템(10, 10')으로서, DASH/MP4를 취급하는 예를 나타냈지만, MPEG2-TS를 취급하는 예도 마찬가지로 생각된다.
- [0187] [송수신 시스템의 구성]
- [0188] 도 34는, MPEG2-TS를 취급하는 송수신 시스템의 구성예를 나타내고 있다. 도 34의 (a)의 송수신

시스템(10A)은, 서비스 송신 시스템(100A)과, 셋톱 박스(STB)(200A)와, 텔레비전 수신기(TV)(300A)를 갖고 있다. 셋톱 박스(200A)와 텔레비전 수신기(300A)는, HDMI(High Definition Multimedia Interface) 케이블(400)을 통해서 접속되어 있다. 도 3의 (b)의 송수신 시스템(10A')은, 서비스 송신 시스템(100A)과, 텔레비전 수신기(TV)(300A)를 갖고 있다.

- [0189] 서비스 송신 시스템(100A)은, MPEG2-TS의 트랜스포트 스트림 TS를, RF 전송로 또는 통신 네트워크 전송로를 통해서 송신한다. 서비스 송신 시스템(100A)은, 오디오 스트림에 메타데이터를 삽입한다. 이 메타 데이터로서는, 예를 들어 소정의 넷 서비스에 접속하기 위한 액세스 정보, 소정의 콘텐츠 정보 등이 생각된다. 여기서는, 상술한 실시 형태와 마찬가지로, 소정의 넷 서비스에 접속하기 위한 액세스 정보가 삽입되는 것으로 한다.
- [0190] 서비스 송신 시스템(100A)은 컨테이너의 레이어에, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 삽입한다. 서비스 송신 시스템(100A)은 이 식별 정보를, 예를 들어 프로그램 맵 테이블(PMT:Program Map Table)의 관리 하의 오디오·엘리멘터리 스트림·루프 내에 디스크립터로서 삽입한다.
- [0191] 셋톱 박스(200A)는, 서비스 송신 시스템(100A)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 트랜스포트 스트림 TS를 수신한다. 이 트랜스포트 스트림 TS에는, 비디오 스트림 및 오디오 스트림이 포함되어 있고, 오디오 스트림에는 메타데이터가 삽입되어 있다.
- [0192] 셋톱 박스(200A)는 오디오 스트림을, 이 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, HDMI 케이블(400)을 통해서, 텔레비전 수신기(300A)에 송신한다.
- [0193] 여기서, 셋톱 박스(200A)는 비디오 스트림을 디코딩해서 얻어진 화상 데이터의 블랭킹 기간에 오디오 스트림 및 식별 정보를 삽입하고, 이 화상 데이터를 텔레비전 수신기(300A)에 송신함으로써, 오디오 스트림 및 식별 정보를 텔레비전 수신기(300A)에 송신한다. 셋톱 박스(200A)는 이 식별 정보를, 예를 들어 오디오·인포 프레임·패킷(Audio InfoFrame packet)에 삽입한다(도 26 참조).
- [0194] 텔레비전 수신기(300A)는, 도 34의 (a)에 도시하는 송수신 시스템(10A)에 있어서는, 셋톱 박스(200A)로부터, HDMI 케이블(400)을 통해서, 오디오 스트림을, 이 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께 수신한다. 즉, 텔레비전 수신기(300A)는 셋톱 박스(200A)로부터, 오디오 스트림 및 식별 정보가 블랭킹 기간에 삽입되고 있는 화상 데이터를 수신한다.
- [0195] 그리고, 텔레비전 수신기(300A)는 식별 정보에 기초하여, 오디오 스트림을 디코딩해서 메타데이터를 추출하고, 이 메타데이터를 사용한 처리를 한다. 이 경우, 텔레비전 수신기(300A)는, 메타데이터로서의 소정의 넷 서비스 정보에 기초하여, 네트워크상의 소정의 서버에 액세스한다.
- [0196] 또한, 텔레비전 수신기(300A)는, 도 34의 (b)에 도시하는 송수신 시스템(10A')에 있어서는, 서비스 송신 시스템(100A)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 트랜스포트 스트림 TS를 수신한다. 이 트랜스포트 스트림 TS에 포함되는 오디오 스트림에는, 메타데이터로서, 소정의 넷 서비스에 접속하기 위한 액세스 정보가 삽입되고 있다. 또한, 컨테이너의 레이어에, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보가 삽입되어 있다.
- [0197] 그리고, 텔레비전 수신기(300A)는 식별 정보에 기초하여, 오디오 스트림을 디코딩해서 메타데이터를 추출하고, 이 메타데이터를 사용한 처리를 한다. 이 경우, 텔레비전 수신기(300A)는, 메타데이터로서의 소정의 넷 서비스 정보에 기초하여, 네트워크상의 소정의 서버에 액세스한다.
- [0198] [서비스 송신 시스템의 TS 생성부]
- [0199] 도 35는, 서비스 송신 시스템(100A)이 구비하는 TS 생성부(110A)의 구성예를 나타내고 있다. 이 도 35에 있어서, 도 8과 대응하는 부분에는, 동일 부호를 붙여서 나타내었다. 이 TS 생성부(110A)는, 제어부(111)와, 비디오 인코더(112)와, 오디오 인코더(113)와, TS 포맷터(114A)를 갖고 있다.
- [0200] 제어부(111)는 CPU(111a)를 구비하고 있고, TS 생성부(110A)의 각 부를 제어한다. 비디오 인코더(112)는, 화상 데이터 SV에 대하여, MPEG2, H.264/AVC, H.265/HEVC 등의 부호화를 실시하여, 비디오 스트림(비디오 엘리멘터리 스트림)을 생성한다. 화상 데이터 SV는, 예를 들어 HDD 등의 기록 매체로부터 재생된 화상 데이터, 또는 비디오 카메라에 의해 얻어진 라이브 화상 데이터 등이다.
- [0201] 오디오 인코더(113)는, 음성 데이터 SA에 대하여, AAC, AC3, AC4, MPEGH(3D 오디오) 등의 압축 포맷에 의한 부호화를 실시하여, 오디오 스트림(오디오 엘리멘터리 스트림)을 생성한다. 음성 데이터 SA는, 상술한 화상 데이

터 SV에 대응한 음성 데이터이며, HDD 등의 기록 매체로부터 재생된 음성 데이터, 또는 마이크로폰에 의해 얻어진 라이브 음성 데이터 등이다.

- [0202] 오디오 인코더(113)는, 오디오 부호화 블록부(113a) 및 오디오 프레임부(113b)를 갖고 있다. 오디오 부호화 블록부(113a)에서 부호화 블록이 생성되고, 오디오 프레임부(113b)에서 프레임이 행하여진다. 이 경우, 압축 포맷에 의해, 부호화 블록이 상이함과 함께, 프레임도 상이하다.
- [0203] 오디오 인코더(113)는, 제어부(111)에 의한 제어 하에서, 오디오 스트림에, 메타데이터 MD를 삽입한다. 이 메타데이터 MD로서는, 예를 들어 소정의 넷 서비스에 접속하기 위한 액세스 정보, 소정의 콘텐츠 정보 등이 생각된다. 여기서는, 상술한 실시 형태와 마찬가지로, 소정의 넷 서비스에 접속하기 위한 액세스 정보가 삽입되는 것으로 한다.
- [0204] 이 메타데이터 MD는, 오디오 스트림의 유저 데이터 영역에 삽입된다. 상세한 설명은 생략하지만, 각 압축 포맷에 있어서의 메타데이터 MD의 삽입은 상술한 실시 형태에 있어서의 DASH/MP4 생성부(110)에 있어서의 경우와 마찬가지로 행하여지고, 메타데이터 MD로서 「SDO\_payload()」이 삽입된다(도 8-도 24 참조).
- [0205] TS 포맷터(114A)는, 비디오 인코더(112)로부터 출력되는 비디오 스트림 및 오디오 인코더(113)로부터 출력되는 오디오 스트림을, PES 패킷화하고, 또한 트랜스포트 패킷화해서 다중하고, 다중화 스트림으로서의 트랜스포트 스트림 TS를 얻는다.
- [0206] 또한, TS 포맷터(114A)는, 프로그램 맵 테이블(PMT)의 관리 하에서, 오디오 스트림에 메타데이터 MD가 삽입되어 있음을 나타내는 식별 정보를 삽입한다. 이 식별 정보의 삽입에는, 오디오·유저 데이터·디스크립터(audio\_userdata\_descriptor)를 사용한다. 이 디스크립터의 상세에 대해서는 후술한다.
- [0207] 도 35에 도시하는 TS 생성부(110A)의 동작을 간단하게 설명한다. 화상 데이터 SV는 비디오 인코더(112)에 공급된다. 이 비디오 인코더(112)에서는, 그 화상 데이터 SV에 대하여 H.264/AVC, H.265/HEVC 등의 부호화가 실시되어, 부호화 비디오 데이터를 포함하는 비디오 스트림이 생성된다.
- [0208] 또한, 음성 데이터 SA는, 오디오 인코더(113)에 공급된다. 이 오디오 인코더(113)에서는, 그 음성 데이터 SA에 대하여, AAC, AC3, AC4, MPEGH(3D 오디오) 등의 부호화가 실시되어, 오디오 스트림이 생성된다.
- [0209] 이때, 제어부(111)로부터 오디오 인코더(113)에, 메타데이터 MD가 공급됨과 함께, 이 메타데이터 MD를 유저 데이터 영역에 매립하기 위한 사이즈 정보가 공급된다. 그리고, 오디오 인코더(113)에서는, 오디오 스트림의 유저 데이터 영역에, 메타데이터 MD를 매립하는 일이 행하여진다. 비디오 인코더(112)에서 생성된 비디오 스트림은, TS 포맷터(114A)에 공급된다. 또한, 오디오 인코더(113)에서 생성된, 유저 데이터 영역에 메타데이터 MD가 매립된 오디오 스트림은, TS 포맷터(114A)에 공급된다.
- [0210] 이 TS 포맷터(114A)에서는, 각 인코더로부터 공급되는 스트림이 패킷화되어 다중되고, 전송 데이터로서 트랜스포트 스트림 TS가 얻어진다. 또한, 이 TS 포맷터(114A)에서는, 프로그램 맵 테이블(PMT)의 관리 하에서, 오디오 스트림에 메타데이터 MD가 삽입되어 있음을 나타내는 식별 정보가 삽입된다.
- [0211] [오디오·유저 데이터·디스크립터의 상세]
- [0212] 도 36은, 오디오·유저 데이터·디스크립터(audio\_userdata\_descriptor)의 구조예(Syntax)를 나타내고 있다. 또한, 도 37은, 그 구조예에 있어서의 주요한 정보의 내용(Semantics)을 나타내고 있다.
- [0213] 「descriptor\_tag」의 8비트 필드는, 디스크립터 타입을 나타낸다. 여기서는, 오디오·유저 데이터·디스크립터임을 나타낸다. 「descriptor\_length」의 8비트 필드는, 디스크립터의 길이(사이즈)를 나타내고, 디스크립터의 길이로서, 이후의 바이트수를 나타낸다.
- [0214] 「audio\_codec\_type」의 8비트 필드는, 오디오의 부호화 방식(압축 포맷)을 나타낸다. 예를 들어, “1”은 「MPEGH」를 나타내고, “2”는 「AAC」를 나타내고, “3”은 「AC3」을 나타내고, “4”는 「AC4」를 나타낸다. 이 정보의 부가에 의해, 수신측에서는, 오디오 스트림에 있어서의 음성 데이터의 부호화 방식을 용이하게 파악할 수 있다.
- [0215] 「metadata\_type」의 3비트 필드는, 메타데이터의 타입을 나타낸다. 예를 들어, “1”은, 「userdata()」의 필드에, 소정의 넷 서비스에 접속하기 위한 액세스 정보를 갖는, ATSC의 「SDO\_payload()」가 놓이는 것을 나타낸다. 이 정보의 부가에 의해, 수신측에서는, 메타데이터의 타입, 즉 메타데이터가 어떤 메타데이터인지를 용이하게 파악할 수 있고, 예를 들어 취득할지 여부의 판단을 행하는 것도 가능하게 된다.

- [0216] 「coordinated\_control\_flag」의 1비트 플래그 정보는, 메타데이터가 오디오 스트림에만 삽입되어 있는지 여부를 나타낸다. 예를 들어, “1”은 다른 컴포넌트의 스트림에도 삽입되어 있음을 나타내고, “0”은 오디오 스트림에만 삽입되어 있음을 나타낸다. 이 정보의 부가에 의해, 수신측에서는, 메타데이터가 오디오 스트림에만 삽입되어 있는지 여부를 용이하게 파악할 수 있다.
- [0217] 「frequency\_type」의 3비트 필드는, 오디오 스트림에 대한 메타데이터의 삽입 빈도의 타입을 나타낸다. 예를 들어, “1”은 각 오디오 액세스 유닛에 하나의 유저 데이터(메타데이터)가 삽입되어 있음을 나타낸다. “2”는 하나의 오디오 액세스 유닛에 복수의 유저 데이터(메타데이터)가 삽입되어 있음을 나타낸다. 또한, “3”은 랜덤 액세스 포인트를 포함하는 그룹마다, 그 선두의 오디오 액세스 유닛에 적어도 하나의 유저 데이터(메타데이터)가 삽입되어 있음을 나타낸다. 이 정보의 부가에 의해, 수신측에서는, 오디오 스트림에 대한 메타데이터의 삽입 빈도를 용이하게 파악할 수 있다.
- [0218] [트랜스포트 스트림 TS의 구성]
- [0219] 도 38은, 트랜스포트 스트림 TS의 구성예를 나타내고 있다. 이 구성예에서는, PID1로 식별되는 비디오 스트림의 PES 패킷 「video PES」가 존재함과 함께, PID2로 식별되는 오디오 스트림의 PES 패킷 「audio PES」가 존재한다. PES 패킷은, PES 헤더(PES\_header)와 PES 페이로드(PES\_payload)로 이루어져 있다. PES 헤더에는, DTS, PTS의 타임 스탬프가 삽입되어 있다. 오디오 스트림의 PES 패킷의 PES 페이로드에는 메타데이터를 포함하는 유저 데이터 영역이 존재한다.
- [0220] 또한, 트랜스포트 스트림 TS에는, PSI(Program Specific Information)로서, PMT(Program Map Table)이 포함되어 있다. PSI는, 트랜스포트 스트림에 포함되는 각 엘리멘터리 스트림이 어느 프로그램에 속해 있는지를 기재한 정보이다. PMT에는, 프로그램 전체에 관련된 정보를 기술하는 프로그램 루프(Program loop)가 존재한다.
- [0221] 또한, PMT에는, 각 엘리멘터리 스트림에 관련된 정보를 갖는 엘리멘터리 스트림 루프가 존재한다. 이 구성예에서는, 비디오 스트림에 대응한 비디오 엘리멘터리 스트림 루프(video ES loop)가 존재함과 함께, 오디오 스트림에 대응한 오디오 엘리멘터리 스트림 루프(audio ES loop)가 존재한다.
- [0222] 비디오 엘리멘터리 스트림 루프(video ES loop)에는, 비디오 스트림에 대응하여, 스트림 타입, PID(패킷 식별자) 등의 정보가 배치됨과 함께, 그 비디오 스트림에 관련된 정보를 기술하는 디스크립터도 배치된다. 이 비디오 스트림의 「Stream\_type」의 값은 「0x24」로 설정되고, PID 정보는, 상술한 바와 같이 비디오 스트림의 PES 패킷 「video PES」에 부여되는 PID1을 나타내게 된다. 디스크립터의 하나로서, HEVC 디스크립터가 배치된다.
- [0223] 또한, 오디오 엘리멘터리 스트림 루프(audio ES loop)에는, 오디오 스트림에 대응하여, 스트림 타입, PID(패킷 식별자) 등의 정보가 배치됨과 함께, 그 오디오 스트림에 관련된 정보를 기술하는 디스크립터도 배치된다. 이 오디오 스트림의 「Stream\_type」의 값은 「0x11」로 설정되고, PID 정보는, 상술한 바와 같이 오디오 스트림의 PES 패킷 「audio PES」에 부여되는 PID2를 나타내게 된다. 디스크립터의 하나로서, 상술한 오디오 유저 데이터 디스크립터(audio\_userdata\_descriptor)가 배치된다.
- [0224] [셋톱 박스의 구성예]
- [0225] 도 39는, 셋톱 박스(200A)의 구성예를 나타내고 있다. 이 도 39에 있어서, 도 25와 대응하는 부분에는 동일 부호를 붙여서 나타내고 있다. 수신부(204A)는, 서비스 송신 시스템(100A)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 트랜스포트 스트림 TS를 수신한다.
- [0226] TS 해석부(205A)는, 트랜스포트 스트림 TS로부터 비디오 스트림의 패킷을 추출하여, 비디오 디코더(206)에 보낸다. 비디오 디코더(206)는, 디멀티플렉서(205)에서 추출된 비디오의 패킷으로부터 비디오 스트림을 재구성하고, 복호화 처리를 행해서 비압축 화상 데이터를 얻는다. 또한, TS 해석부(205A)는, 트랜스포트 스트림 TS로부터 오디오 스트림의 패킷을 추출하고, 오디오 스트림을 재구성한다. 오디오 프레이밍부(207)는, 이렇게 재구성된 오디오 스트림에 대하여 프레이밍을 행한다.
- [0227] 또한, TS 해석부(205A)로부터 전송되는 오디오 스트림을 오디오 프레이밍부(207)에 보내는 것과 병행하여, 도시하지 않은 오디오 디코더로 디코딩하여, 오디오 출력을 행하는 것도 가능하다.
- [0228] 또한, TS 해석부(205A)는, 트랜스포트 스트림 TS로부터 각종 디스크립터 등을 추출하여, CPU(211)에 송신한다. 여기서, 디스크립터에는, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보로서의 오디오 유저 데이터 디스크립터(도 36 참조)도 포함된다.



- [0229] 상세한 설명은 생략하지만, 이 도 39에 도시하는 셋톱 박스(200A) 이외는, 도 25에 도시하는 셋톱 박스(200)와 마찬가지로 구성되고, 마찬가지로의 동작을 한다.
- [0230] [텔레비전 수신기의 구성예]
- [0231] 도 40은 텔레비전 수신기(300A)의 구성예를 나타내고 있다. 이 도 40에 있어서, 도 27과 대응하는 부분에는 동일 부호를 붙여서 나타내고 있다. 수신부(306A)는, 서비스 송신 시스템(100A)으로부터, RF 전송로 또는 통신 네트워크 전송로를 통해서 보내져 오는 트랜스포트 스트림 TS를 수신한다.
- [0232] TS 해석부(307A)는, 트랜스포트 스트림 TS로부터 비디오 스트림의 패킷을 추출하여, 비디오 디코더(308)에 보낸다. 비디오 디코더(308)는, 디멀티플렉서(205)에서 추출된 비디오의 패킷으로부터 비디오 스트림을 재구성하고, 복호화 처리를 행해서 비압축 화상 데이터를 얻는다. 또한, TS 해석부(307A)는, 트랜스포트 스트림 TS로부터 오디오 스트림의 패킷을 추출하고, 오디오 스트림을 재구성한다.
- [0233] 또한, TS 해석부(307A)는, 트랜스포트 스트림 TS로부터 오디오 스트림의 패킷을 추출하고, 오디오 스트림을 재구성한다. 또한, TS 해석부(307A)는, 트랜스포트 스트림 TS로부터 각종 디스크립터 등을 추출하여, CPU(321)에 송신한다. 여기서, 이 디스크립터에는, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보로서의 오디오·유저 데이터·디스크립터(도 36 참조)도 포함된다. 상세한 설명은 생략하지만, 이 도 40에 도시하는 텔레비전 수신기(300A) 이외는, 도 27에 도시하는 텔레비전 수신기(300)와 마찬가지로 구성되고, 마찬가지로의 동작을 한다.
- [0234] 상술한 바와 같이, 도 34의 (a), (b)에 도시하는 화상 표시 시스템(10A, 10A')에 있어서, 서비스 송신 시스템(100A)은, 오디오 스트림에 메타데이터를 삽입함과 함께, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 컨테이너의 레이어에 삽입한다. 그로 인해, 수신측(셋톱 박스(200A), 텔레비전 수신기(300A))에서는, 오디오 스트림에 메타데이터가 삽입되어 있음을 용이하게 인식할 수 있다.
- [0235] 또한, 도 34의 (a)에 도시하는 화상 표시 시스템(10A)에 있어서, 셋톱 박스(200A)는 메타데이터가 삽입된 오디오 스트림을, 이 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 텔레비전 수신기(300A)에, HDMI에 의해 송신한다. 그로 인해, 텔레비전 수신기(300A)에서는, 오디오 스트림에 메타데이터가 삽입되어 있음을 용이하게 인식할 수 있고, 이 인식에 기초하여 오디오 스트림에 삽입되어 있는 메타데이터의 추출 처리를 행함으로써, 낭비 없이 확실하게 메타데이터를 취득해서 이용할 수 있다.
- [0236] 또한, 도 34의 (b)에 도시하는 화상 표시 시스템(10A')에 있어서, 텔레비전 수신기(300A)는, 오디오 스트림과 함께 수신되는 식별 정보에 기초해서 이 오디오 스트림으로부터 메타데이터를 추출해서 처리에 사용한다. 그로 인해, 오디오 스트림에 삽입되어 있는 메타데이터를 낭비 없이 확실하게 취득할 수 있어, 메타데이터를 사용한 처리를 적절하게 실행할 수 있다.
- [0237] 또한, 상술한 실시 형태에 있어서, 셋톱 박스(200)는 텔레비전 수신기(300)에 화상 데이터, 오디오 스트림을 송신하는 구성으로 되어 있다. 그러나, 텔레비전 수신기(300) 대신에 모니터 장치 또는 프로젝터 등에 송신하는 구성도 생각된다. 또한, 셋톱 박스(200) 대신에, 수신 기능을 구비한 레코더, 퍼스널 컴퓨터 등인 구성도 생각된다.
- [0238] 또한, 상술한 실시 형태에 있어서, 셋톱 박스(200)와 텔레비전 수신기(300)는 HDMI 케이블(400)로 접속되어 이루어지는 것이다. 그러나, 이들 사이가, HDMI와 마찬가지로의 디지털 인터페이스로 유선 접속되는 경우, 나아가서는, 무선에 의해 접속되는 경우에도, 본 발명을 마찬가지로 적용할 수 있음은 물론이다.
- [0239] 또한, 본 기술은, 이하와 같은 구성도 취할 수 있다.
- [0240] (1) 메타데이터가 삽입된 오디오 스트림을 수신 장치에서 취득하기 위한 메타 정보를 갖는 메타파일을 송신하는 송신부와, 상기 오디오 스트림에 상기 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 상기 메타파일에 삽입하는 정보 삽입부를 구비한다 송신 장치.
- [0241] (2) 상기 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보인 상기 (1)에 기재된 송신 장치.
- [0242] (3) 상기 메타데이터는, URI 정보를 나타내는 문자 코드인 상기 (2)에 기재된 송신 장치.
- [0243] (4) 상기 메타파일은, MPD 파일인 상기 (1) 내지 (3) 중 어느 한 항에 기재된 송신 장치.
- [0244] (5) 상기 정보 삽입부는, “Supplementary Descriptor” 를 사용하여, 상기 메타파일에 상기 식별 정보를 삽입하

는 상기 (4)에 기재된 송신 장치.

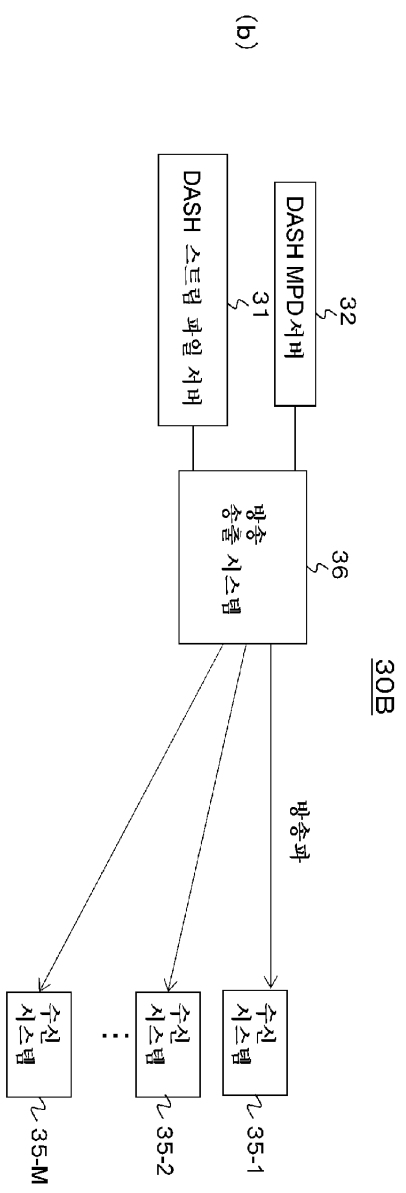
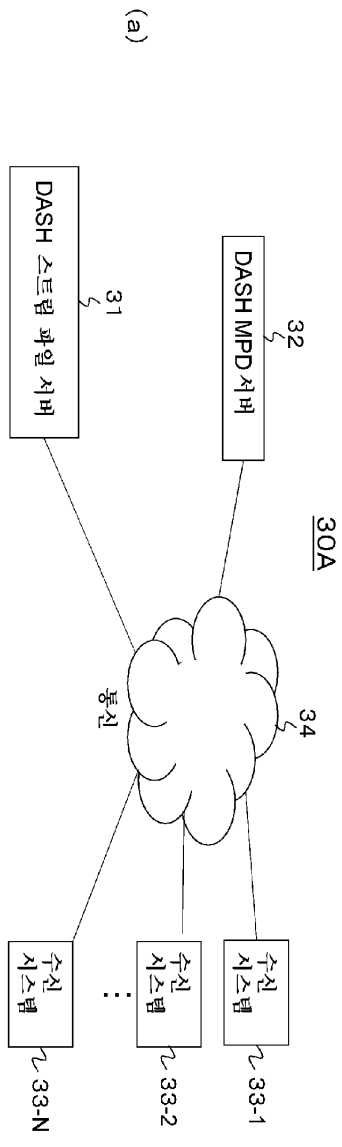
- [0245] (6) 상기 송신부는, 상기 메타파일을 RF 전송로 또는 통신 네트워크 전송로를 통해서 송신하는 상기 (1) 내지 (5) 중 어느 한 항에 기재된 송신 장치.
- [0246] (7) 상기 송신부는, 상기 메타데이터가 삽입된 오디오 스트림을 포함하는 소정 포맷의 컨테이너를 더 송신하는 상기 (1) 내지 (6) 중 어느 한 항에 기재된 송신 장치.
- [0247] (8) 상기 컨테이너는, MP4인 상기 (7)에 기재된 송신 장치.
- [0248] (9) 송신부에 의해, 메타데이터가 삽입된 오디오 스트림을 수신 장치에서 취득하기 위한 메타 정보를 갖는 메타파일을 송신하는 송신 스텝과, 상기 오디오 스트림에 상기 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 상기 메타파일에 삽입하는 정보 삽입 스텝을 갖는 송신 방법.
- [0249] (10) 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신부를 구비하고, 상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고, 상기 오디오 스트림을, 상기 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 소정의 전송로를 통해서, 외부 기기에 송신하는 송신부를 더 구비하는 수신 장치.
- [0250] (11) 상기 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보인 상기 (10)에 기재된 수신 장치.
- [0251] (12) 상기 메타파일은, MPD 파일이며, 상기 메타파일에는, “Supplementary Descriptor”에 의해, 상기 식별 정보가 삽입되어 있는 상기 (10) 또는 (11)에 기재된 수신 장치.
- [0252] (13) 상기 송신부는, 화상 데이터의 블랭킹 기간에 상기 오디오 스트림 및 상기 식별 정보를 삽입하고, 그 화상 데이터를 상기 외부 기기에 송신함으로써, 상기 오디오 스트림 및 상기 식별 정보를 상기 외부 기기에 송신하는 상기 (10) 내지 (12) 중 어느 한 항에 기재된 수신 장치.
- [0253] (14) 상기 소정의 전송로는, HDMI 케이블인 상기 (10) 내지 (13) 중 어느 한 항에 기재된 수신 장치.
- [0254] (15) 수신부에 의해, 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신 스텝을 갖고, 상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고, 상기 오디오 스트림을, 상기 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보와 함께, 소정의 전송로를 통해서, 외부 기기에 송신하는 송신 스텝을 더 갖는 수신 방법.
- [0255] (16) 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신부를 구비하고, 상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고, 상기 식별 정보에 기초하여, 상기 오디오 스트림을 디코딩해서 상기 메타데이터를 추출하는 메타데이터 추출부와, 상기 메타데이터를 사용한 처리를 행하는 처리부를 더 구비하는 수신 장치.
- [0256] (17) 상기 메타파일은, MPD 파일이며, 상기 메타파일에는, “Supplementary Descriptor”에 의해, 상기 식별 정보가 삽입되어 있는 상기 (16)에 기재된 수신 장치.
- [0257] (18) 상기 메타데이터는, 소정의 넷 서비스에 접속하기 위한 액세스 정보이며, 상기 처리부는, 상기 네트워크 액세스 정보에 기초하여, 네트워크상의 소정의 서버에 액세스하는 상기 (16) 또는 (17)에 기재된 수신 장치.
- [0258] (19) 수신부에 의해, 메타데이터가 삽입된 오디오 스트림을 취득하기 위한 메타 정보를 갖는 메타파일을 수신하는 수신 스텝을 갖고, 상기 메타파일에는, 상기 오디오 스트림에 상기 메타데이터의 삽입이 있음을 나타내는 식별 정보가 삽입되어 있고, 상기 식별 정보에 기초하여, 상기 오디오 스트림을 디코딩해서 상기 메타데이터를 추출하는 메타데이터 추출 스텝과, 상기 메타데이터를 사용한 처리를 행하는 처리 스텝을 더 갖는 수신 방법.
- [0259] (20) 네트워크 액세스 정보를 포함하는 메타데이터가 삽입된 오디오 스트림을 생성하는 스트림 생성부와, 상기 오디오 스트림을 갖는 소정 포맷의 컨테이너를 송신하는 송신부를 구비하는 송신 장치.
- [0260] 본 기술의 주된 특징은, DASH/MP4의 패킷에서 오디오 스트림에 메타데이터를 삽입하는 경우에, 오디오 스트림에 메타데이터가 삽입되어 있음을 나타내는 식별 정보를 MPD 파일에 삽입함으로써, 수신측에 있어서, 오디오 스트림에 메타데이터가 삽입되어 있음을 용이하게 인식 가능하게 한 것이다(도 3, 도 4 참조).

**부호의 설명**

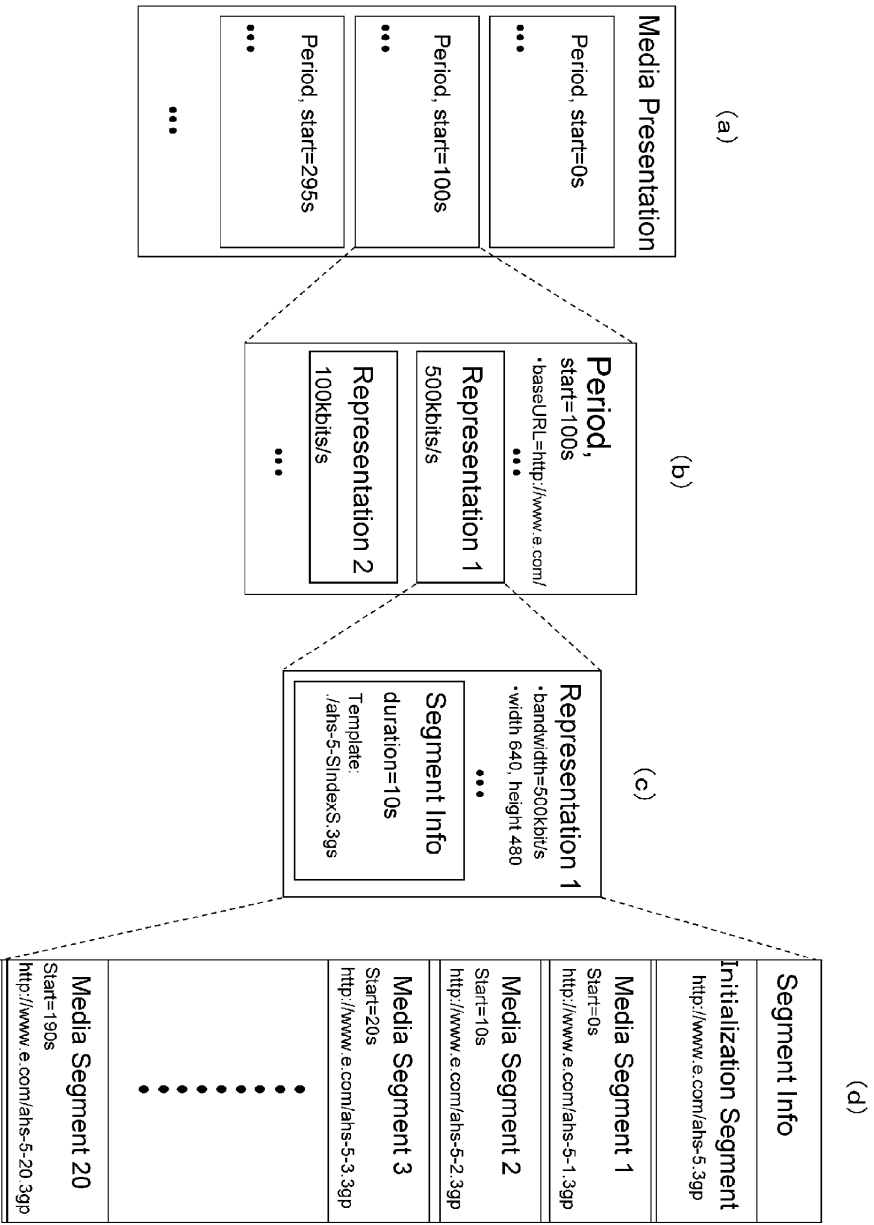
- [0261] 10, 10', 10A, 10A' : 송수신 시스템

- 14 : 유효 화소 구간
- 15 : 수평 귀선 기간
- 16 : 수직 귀선 기간
- 17 : 비디오 데이터 구간
- 18 : 데이터 아일랜드 구간
- 19 : 컨트롤 구간
- 30A, 30B : MPEG-DASH 베이스의 스트림 배신 시스템
- 31 : DASH 스트림 파일 서버
- 32 : DASH MPD 서버
- 33, 33-1 내지 33-N : 수신 시스템
- 34 : CDN
- 35, 35-1 내지 35-M : 수신 시스템
- 36 : 방송 송출 시스템
- 81 : HDMI 트랜스미터
- 82 : HDMI 리시버
- 83 : DDC
- 84 : CEC 라인
- 85 : EDID ROM
- 100, 100A : 서비스 송신 시스템
- 110 : DASH/MP4 생성부
- 110A : TS 생성부
- 111 : 제어부
- 111a : CPU
- 112 : 비디오 인코더
- 113 : 오디오 인코더
- 113a : 오디오 부호화 블록부
- 113b : 오디오 프레임부
- 114 : DASH/MP4 포맷터
- 114A : TS 포맷터
- 200, 200A : 셋톱 박스(STB)
- 204, 204A : 수신부
- 205 : DASH/MP4 해석부
- 205A : TS 해석부
- 206 : 비디오 디코더
- 207 : 오디오 프레임부
- 208 : HDMI 송신부

- 209 : HDMI 단자
- 211 : CPU
- 212 : 플래시 ROM
- 213 : DRAM
- 214 : 내부 버스
- 215 : 리모콘 수신부
- 216 : 리모콘 송신기
- 300, 300A : 텔레비전 수신기
- 306, 306A : 수신부
- 307 : DASH/MP4 해석부
- 307A : TS 해석부
- 308 : 비디오 디코더
- 309 : 영상 처리 회로
- 310 : 패널 구동 회로
- 311 : 표시 패널
- 312 : 오디오 디코더
- 313 : 음성 처리 회로
- 314 : 음성 증폭 회로
- 315 : 스피커
- 316 : HDMI 단자
- 317 : HDMI 수신부
- 318 : 통신 인터페이스
- 321 : CPU
- 322 : 플래시 ROM
- 323 : DRAM
- 324 : 내부 버스
- 325 : 리모콘 수신부
- 326 : 리모콘 송신기
- 350 : 외부 스피커 시스템
- 400 : HDMI 케이블

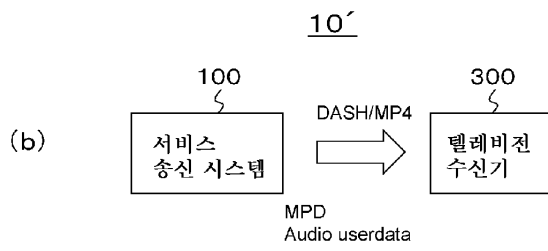
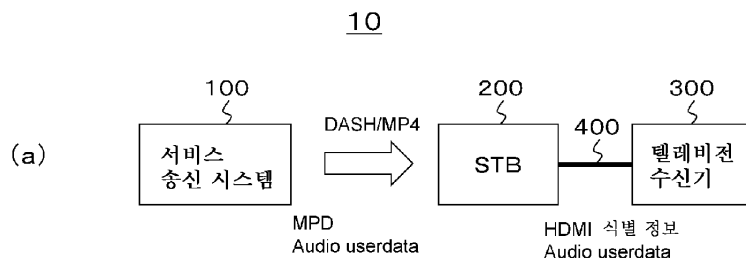


도면  
도면1



도면2

도면3



MPD 파일 기술예

```

<MPD>
<Period>
  <AdaptationSet mimeType="audio/mp4" group="1">
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:codecType" value="mpegh"/>
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:coordinatedControl" value="true"/>
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:AudioMetaContained" value="true"/>
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:netlink:metaInsertionFrequency" value="1"/>
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:type" value="netlink"/>
    <Representation id="1" bandwidth="128000">
      <baseURL>audio/jp/128.mp4</BaseURL>
    </Representation>
  </AdaptationSet>
  <AdaptationSet mimeType="video/mp4" group="2">
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:codecType" value="hevc"/>
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:coordinatedControl" value="true"/>
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:VideoMetaContained" value="true"/>
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:metaInsertionFrequency" value="1"/>
    <SupplementaryDescriptor schemeIdUri="urn:brdcs:type" value="netlink"/>
    <Representation id="21" bandwidth="20000000">
      <baseURL>video/jp/20000000.mp4</BaseURL>
    </Representation>
  </AdaptationSet>
</Period>
</MPD>

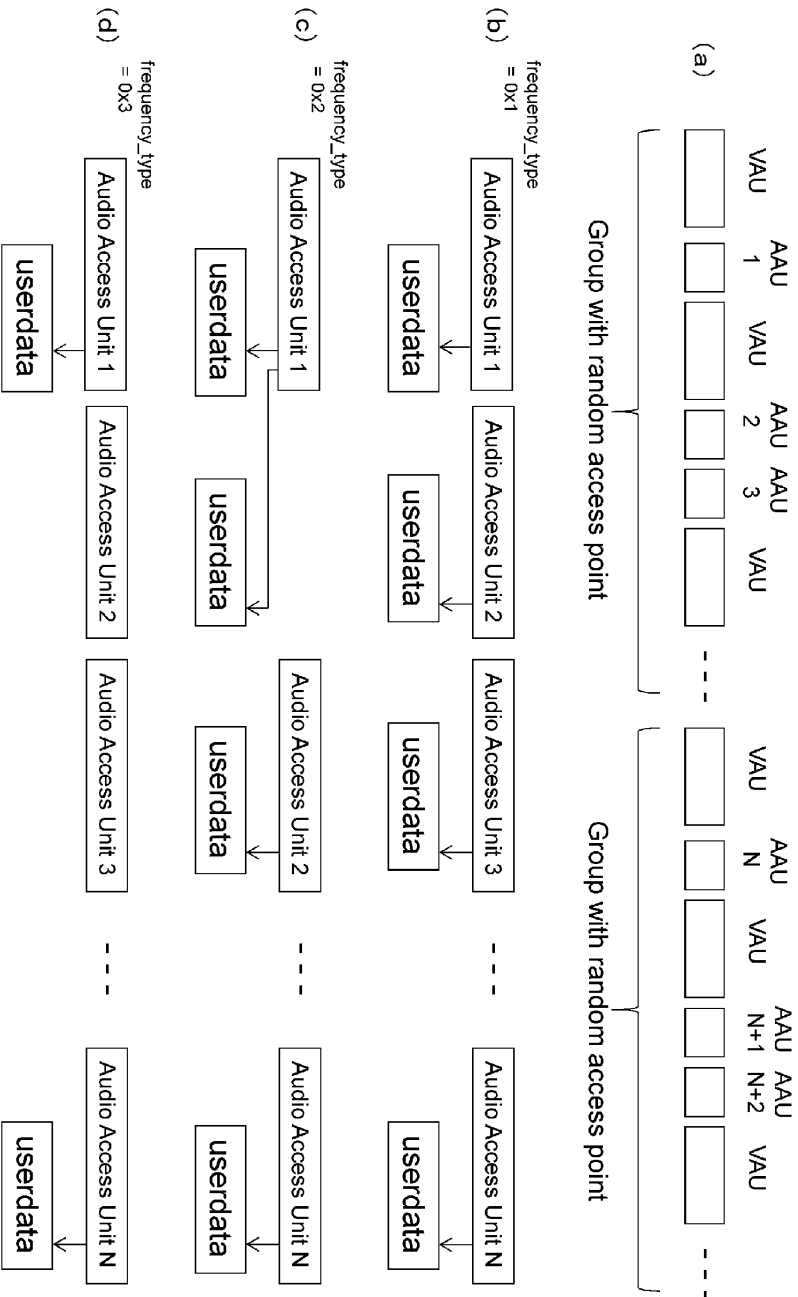
```



SupplementaryDescriptor에 의해, schemeIdUri를 이하와 같이 정의한다.

<p>schemeIdUri="um:brdct:codecType"으로, 코덱의 종류를 나타낸다. value: "mp3", "AAC", "AC3", "AC4"라고 기술</p>	
<p>schemeIdUri="um:brdct:coordinatedControl"로, 넷 접속에 필요한 정보가 복수의 미디어 스트림 사이에서 협조해서 공급됨을 나타낸다. value: "true" (예를 들어, 넷 접속 정보가 다른 Adaptation Set의 스트림과 협조해서 공급된다. 넷 접속 정보가 audio stream으로도, video stream으로도 공급됨) value: "false" 넷 접속 정보가 본 Adaptation Set의 스트림만으로 공급된다.</p>	
<p>schemeIdUri="um:brdct:audioMetaContained"로, 오디오 메타 정보가 포함되는 것을 나타낸다. value: "true" 오디오 메타 정보가 포함된다. value: "false" 오디오 메타 정보가 포함되지 않는다.</p>	
<p>schemeIdUri="um:brdct:metainserionFrequency"로, 액세스 유닛 단위로 메타 정보가 공급되는 빈도를 나타낸다. value: "1" 하나의 유저 데이터 entry가 1액세스 유닛에 발생한다. value: "2" 복수의 유저 데이터 entry가 1액세스 유닛에 발생한다. value: "3" 하나 이상의 유저 데이터 entry가, 랜덤 액세스 포인트로 구획되는 기간 중에 발생한다.</p>	
<p>schemeIdUri="um:brdct:videoMetaContained"로, 비디오 메타 정보가 포함되는 것을 나타낸다. value: "true" 비디오 메타 정보가 포함된다. value: "false" 비디오 메타 정보가 포함되지 않는다.</p>	
<p>schemeIdUri="um:brdct:type"로, 메타에 의한 서비스의 종류를 나타낸다. value: "neilink" 메타에 의한 서비스의 종류가, 넷 접속이다.</p>	

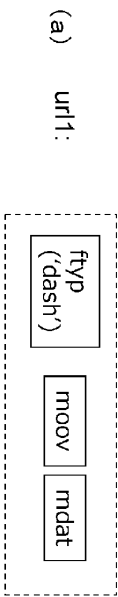
도면5



도면6

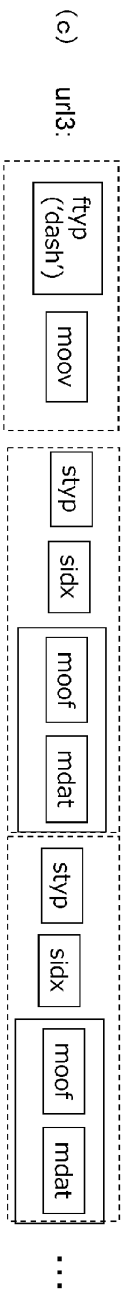
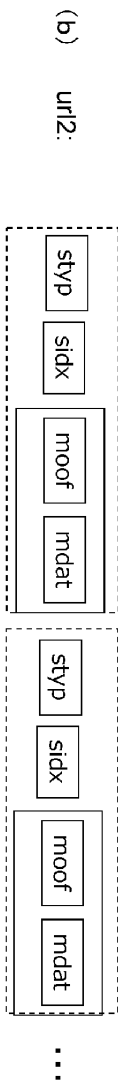
<baseURL>로 나타나는 로케이션처의 미디어 파일 실체

Non-Fragmented MP4의 경우



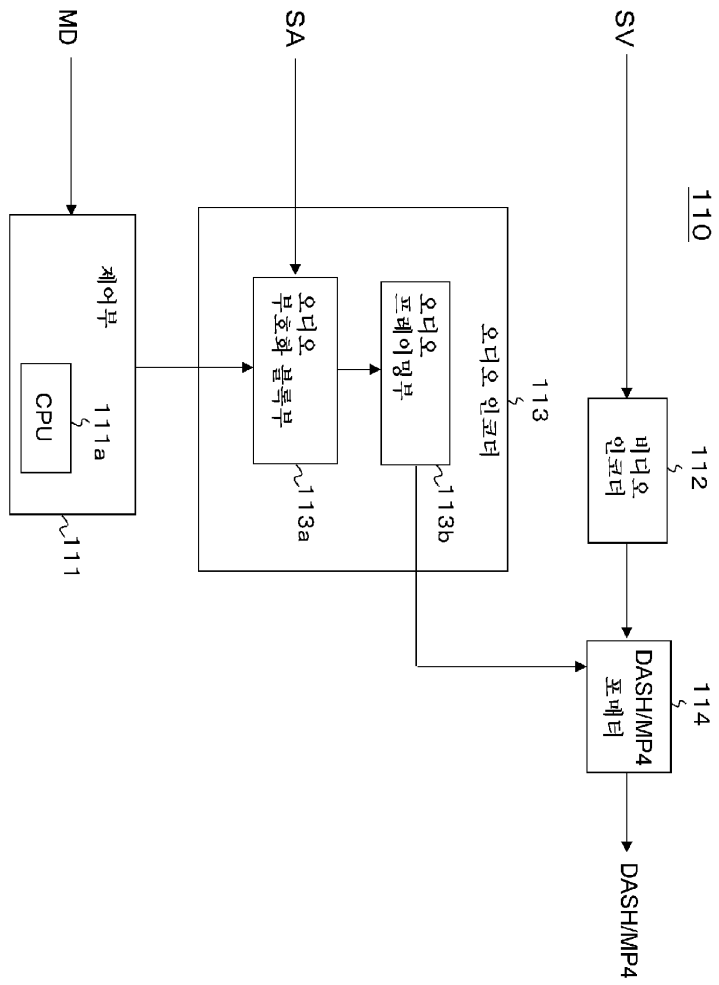
styp: segment type box  
 moov: movie header  
 ftyp: file type box  
 moof: movie fragment  
 mdat: media data container  
 sidx: segment index box

Fragmented MP4의 경우



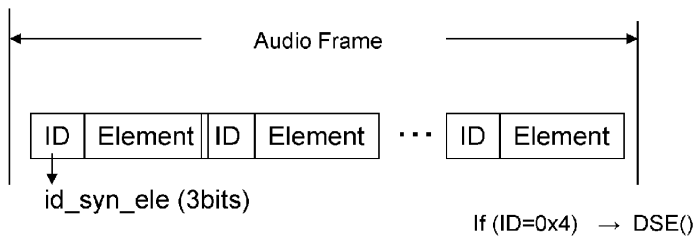
도면7

도면8



도면9

AAC의 Audio Frame 구조



도면10

Syntax of data\_stream\_element

Syntax	No. of bits	Mnemonic
data_stream_element ()		
[		
element_instance_tag;	4	uimsbf
data_byte_align_flag;	1	uimsbf
cnt=count	8	uimsbf
if (cn t==255) {		
Cnt + = esc_count;	8	uimsbf
}		
if (data_byte_align_flag) {		
byte_alignment ();		
}		
for (l = 0, i<cnt, i++) {		
data_stream_byte [element_instance_tag] [i];	8	uimsbf
}		
}		

(a)

Syntax	No. of Bits	Format
metadata {}		
userdata {}	32	uimssf
if (userdata_identifier == "AAAA" ) {		
metadata_type	8	uimssf
if (metadata_type == "0x08") {		
SDO_payload()		
}else{		
reserved		
}		
}		
}		

(b)

<b>userdata_identifier (32bits)</b>	Shall be set to the value of predefined sequence for audio_userdata.
<b>metadata_type (8bits)</b>	Describes type of metadata. '0x08' Unconditioned character to describe URL access, which specifies to contain ATSC SDO_SDO_payload().

도면11

도면12

Syntax of SDO\_payload()

Syntax	No. of Bits	Format
SDO_payload() {		
if (cmdID<0x05) {		
for (k=0; k<L-1; k++) {		
URL_character	8	uimsbf
}		
} else {		
reserved	var	
}		
}		

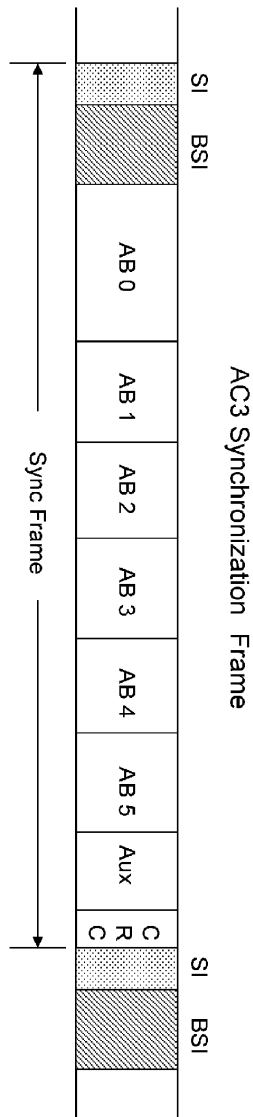
Syntax/semantics of SDO\_payload is compliant to ATSC A/105 Candidate Standard.

도면13

cmdID value	Meaning
0x00	Interactive services Trigger – TDO model
0x01	Interactive services Trigger – Direct Execution model
0x02	Location of PDI Table
0x03	Location of Usage Reporting Data Server
0x04	Base URL for Internet delivery of signaling and announcements
0x05-0x1F	Reserved for future ATSC use



도면14

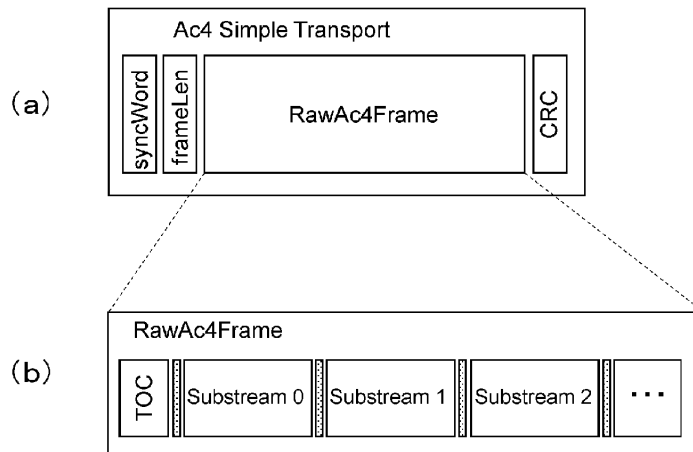


도면15

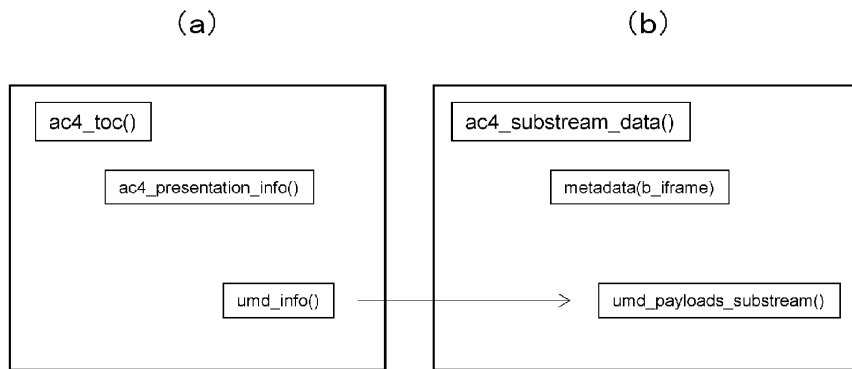
AC3의 AUXILIARY DATA의 구성

Syntax	Word Size
auxdata()	
{	
auxbits	n auxbits
if(auxdatae)	
{	
auxdatal	14
}	
auxdatae	1
}	
/* end of auxdata */	

도면16



도면17



도면18

Syntax of umd\_info()

Syntax	No. of bits	Mnemonic
umd_info(){		
umd_version	2	umrsbf
k_id		
substream_index	2(+2)	umrsbf
:		
:		
}		

Substream\_index 인덱스값

Syntax	No. of Bits	Format
umnd_payload_substream() {		
umnd_payload_id	5	uimsbf
umnd_userdata_identifier	32	uimsbf
umnd_payload_size	16	uimsbf
if (umnd_userdata_identifier == "AAAA") {		
umnd_metadata_type	8	uimsbf
if (umnd_metadata_type == "0x08") {		
SDO_payload()		
}else{		
reserved		
}		
}		
}		

umnd\_userdata\_identifier (32bits) Shall be set to the value of predefined sequence for audio userdata.

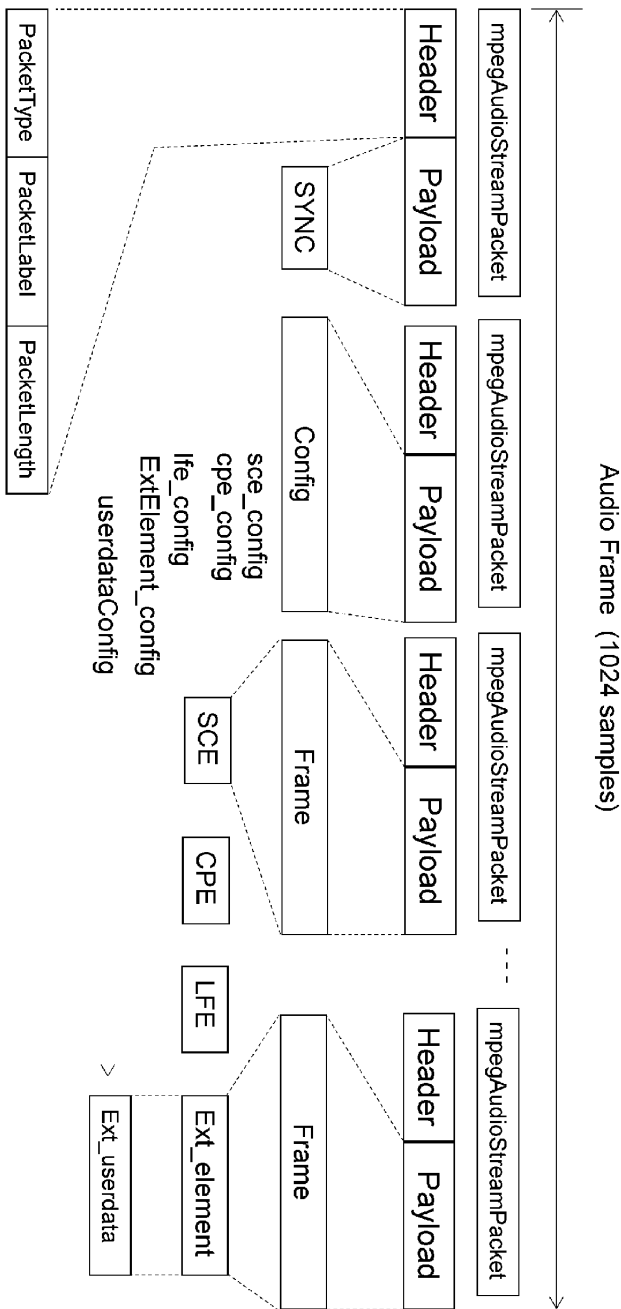
umnd\_payload\_size (16bits) Describes the total number of bytes following this element.

umnd\_metadata\_type (8bits) Describes type of metadata.

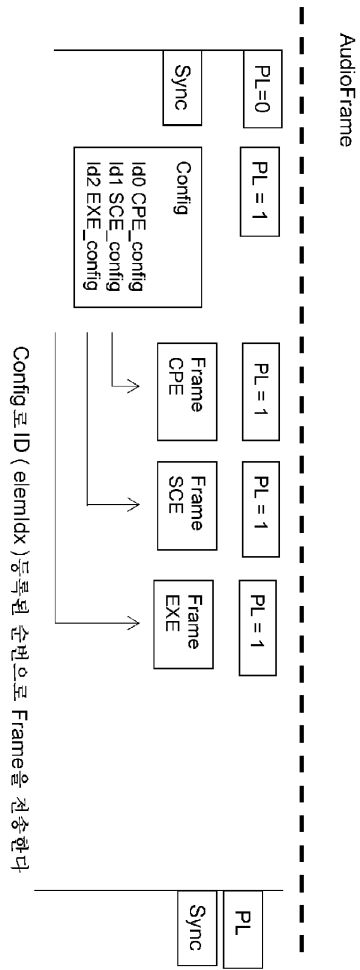
'0x08' Unconditioned character to describe URL access, which specifies to contain ATSC SDO\_payload().

도면19

도면20



도면21



도면22

usacExtElementType	Value
ID_EXT_ELE_FILL	0
ID_EXT_ELE_MPEGS	1
ID_EXT_ELE_SAOC	2
ID_EXT_ELE_AUDIOPREROLL	3
ID_EXT_ELE_UNI_DRC	4
ID_EXT_ELE_OBJ_METADATA	5
ID_EXT_ELE_SAOC_3D	6
ID_EXT_ELE_HQA	7
/* reserved for ISO use */	8-127
/* reserved for use outside of ISO scope */	128 and higher
ID_EXT_ELE_userdata	128

Syntax of userdataConfig()

Syntax	No. of bits	Mnemonic
userdataConfig()		
{		
userdata_identifier	32	uimnsbf
userdata_framelength	16	uimnsbf
}		

(a)

(b)

<p>userdata_identifier (32bits)    Shall be set to the value of predefined sequence for audio userdata.                  userdata_framelength (16bits)    Describes the total number of bytes in audio_userdata().</p>
--

도면23



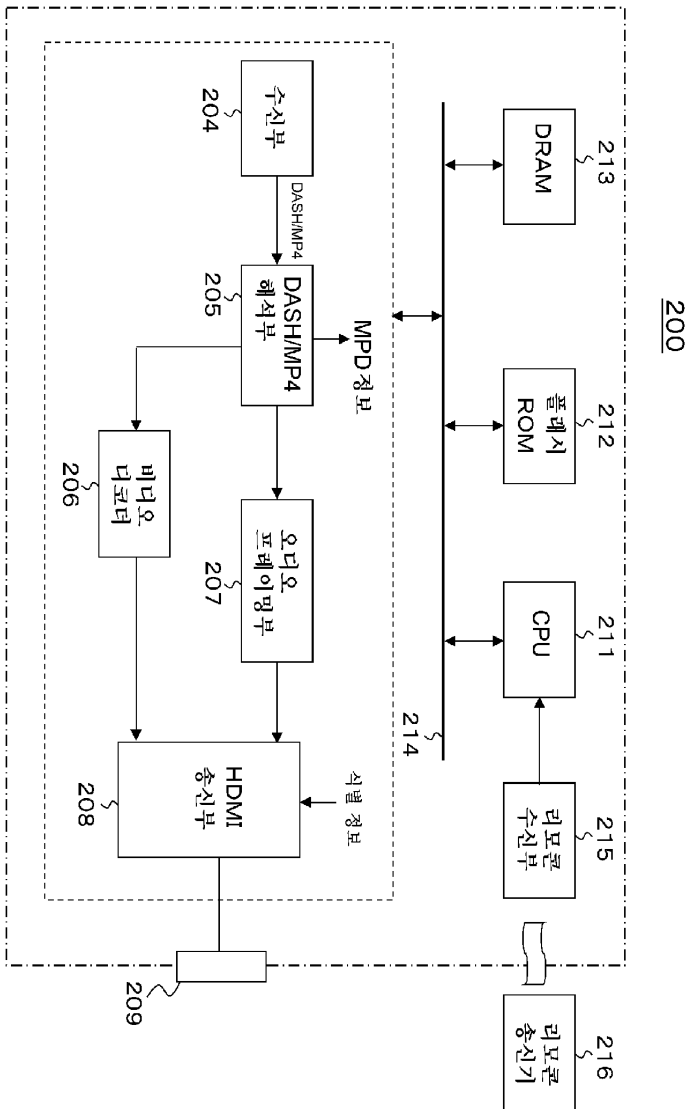
Syntax of audio\_userdata()

Syntax	No. of bits	Mnemonic
audio_userdata()		
{		
if (userdata_identifier == "AAAA") {		
metadataType	8	uimrbf
if (metadataType == "0x08") {		
SDO_payload()		
}else{		
reserved		
}		
}		

도면24

metadataType (8bits)      Describes type of metadata.  
 '0x08'      Unconditioned character to describe URL access,  
 which specifies to contain ATSC SDO\_payload().

도면25

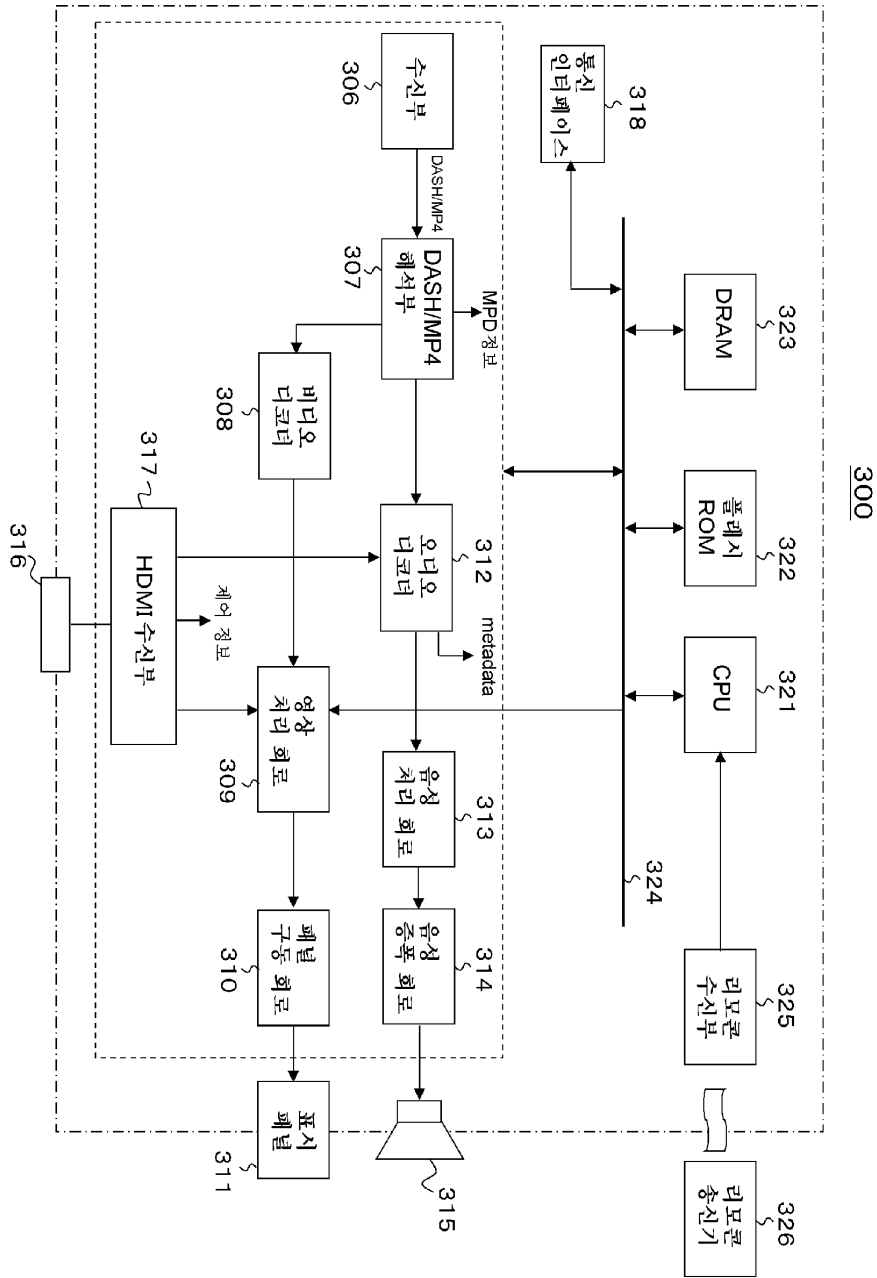


도면26

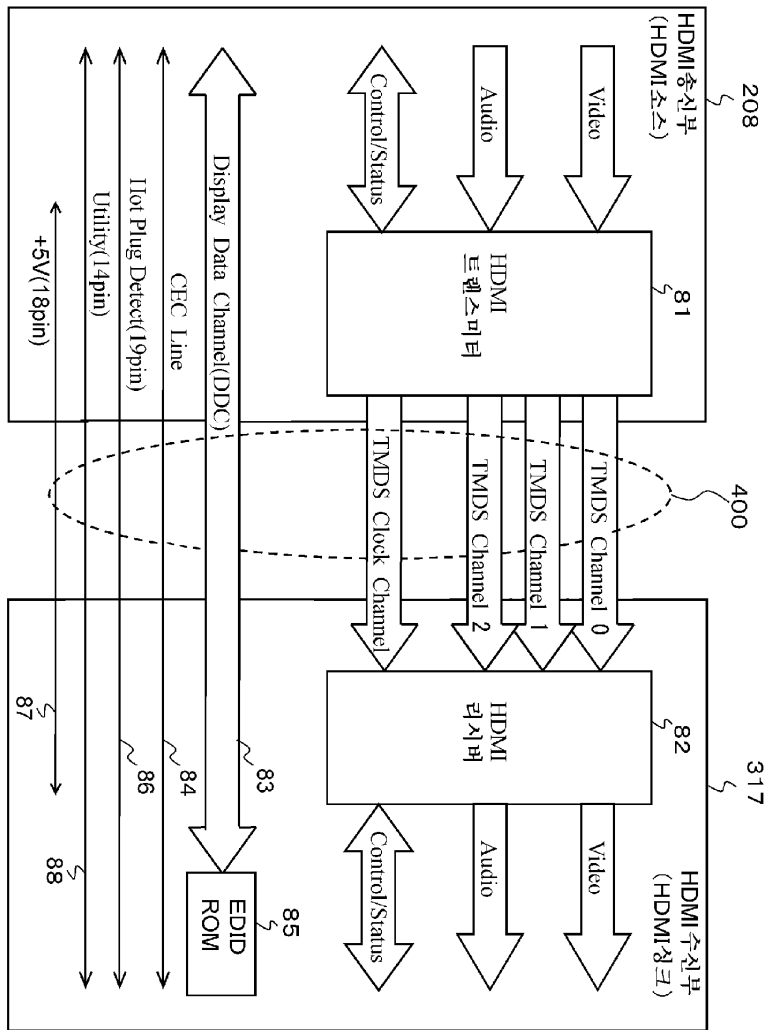
Audio InfoFrame packet의 구조에

Byte#	7	6	5	4	3	2	1	0
0	Packet Type = 0x84							
1	Version							
2	0							
3	Checksum							
4	CT3	CT2	CT1	CT0	Rsvd(0)	CC2	CC1	CC0
5	Reserved (0)	Reserved (0)	Userdata presence_flag	SF2	SF1	SF0	SS1	SS0
6	Format depends on coding type							
7	CA7	CA6	CA5	CA4	CA3	CA2	CA1	CA0
8	DM_INH	LSV3	LSV2	LSV1	LSV0	Rsvd (0)	LFPBL1	LFPBL0
9	metadata_type		coordinated control_flag		Rsvd (0)	frequency_type		
10	Reserved (0)							
11	Reserved (0)							
12	Reserved (0)							
13	Reserved (0)							
14...30	NA Reserved (0)							

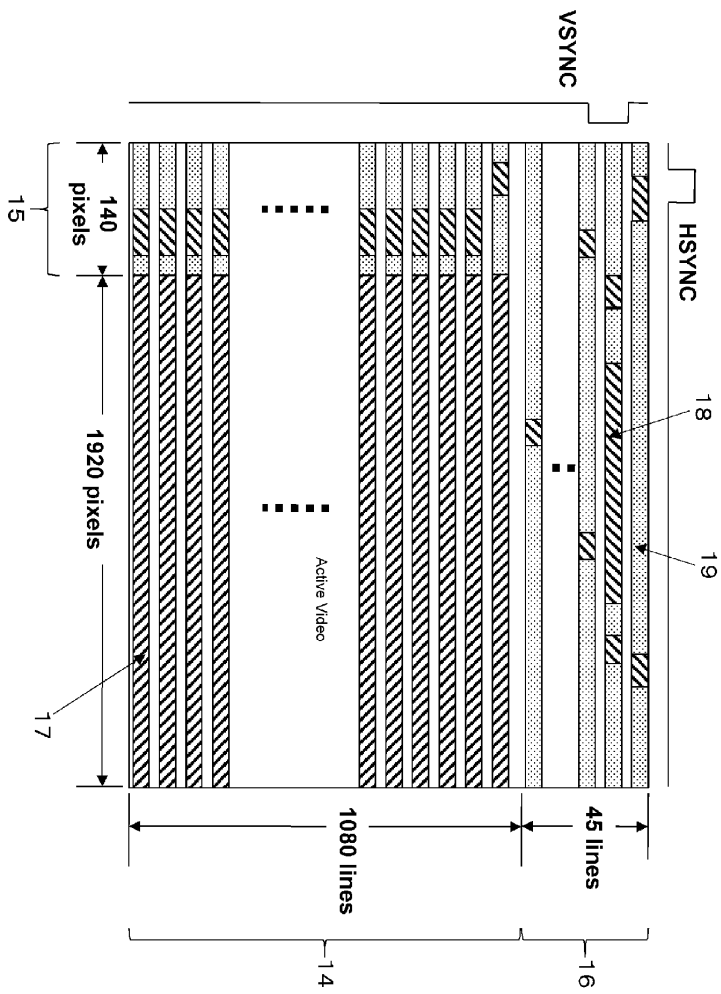
도면27



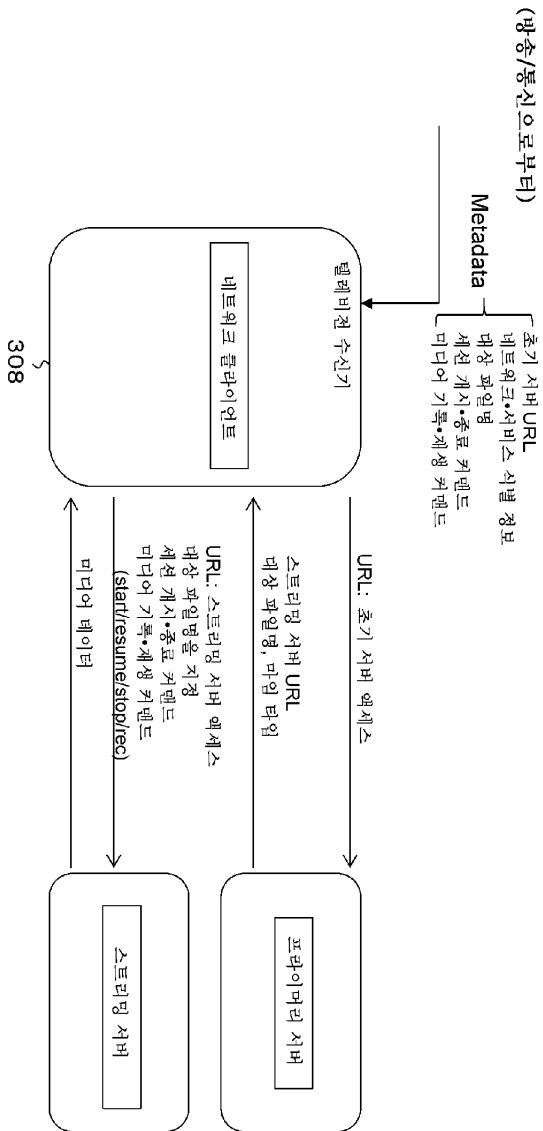
도면28



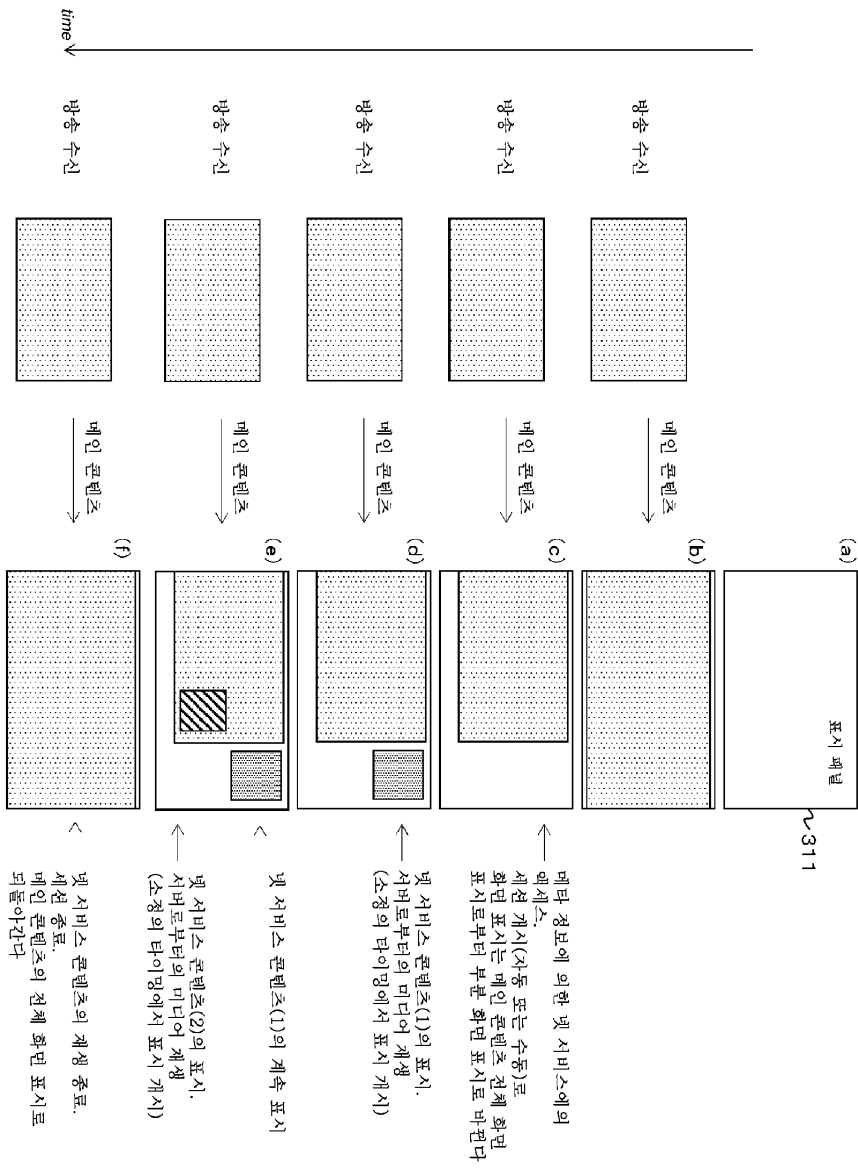
도면29



도면30

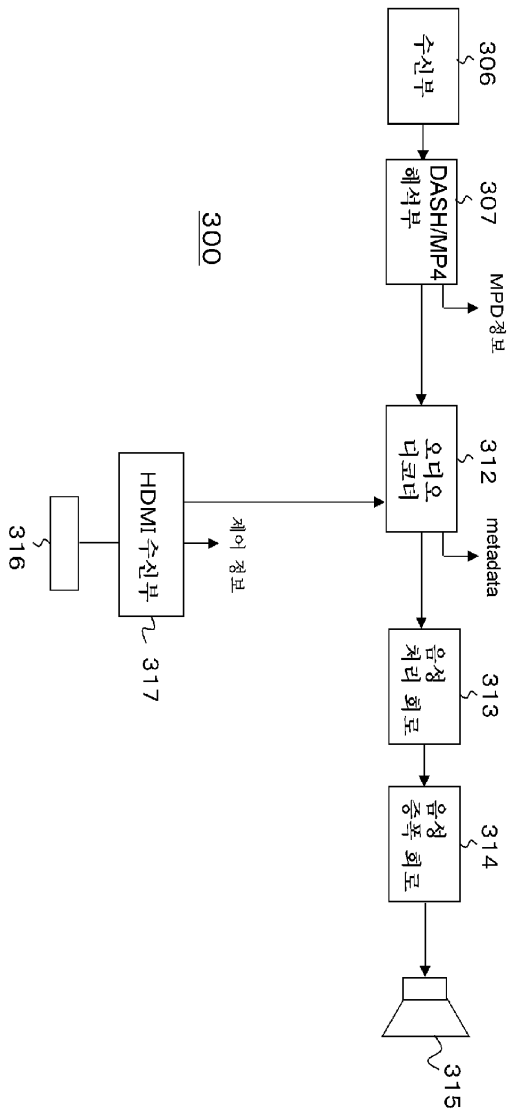


도면31

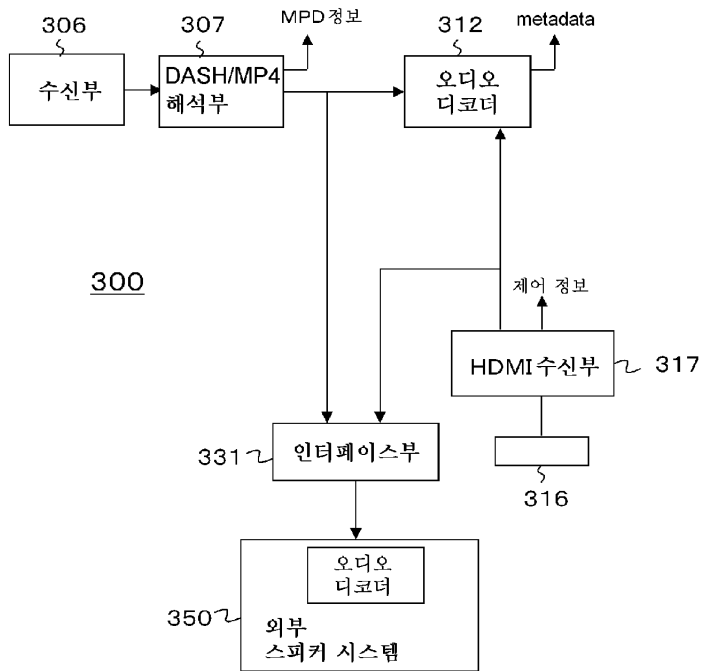




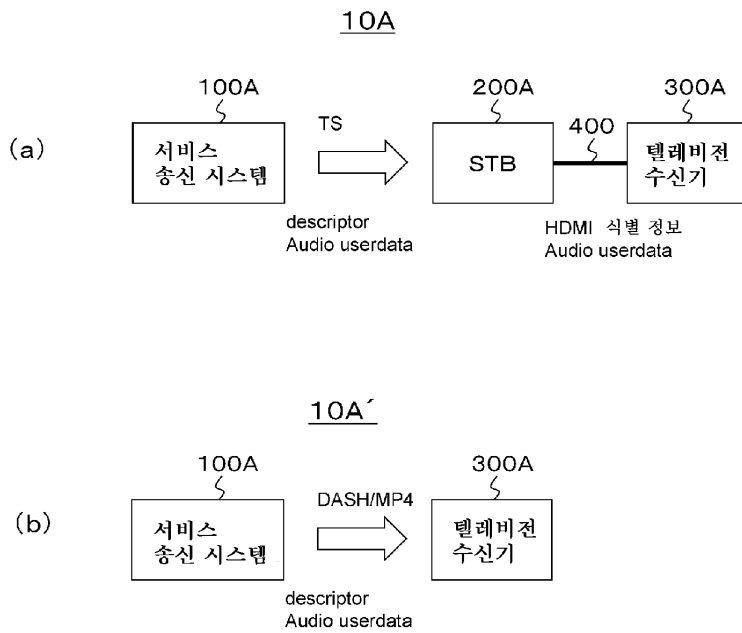
도면32



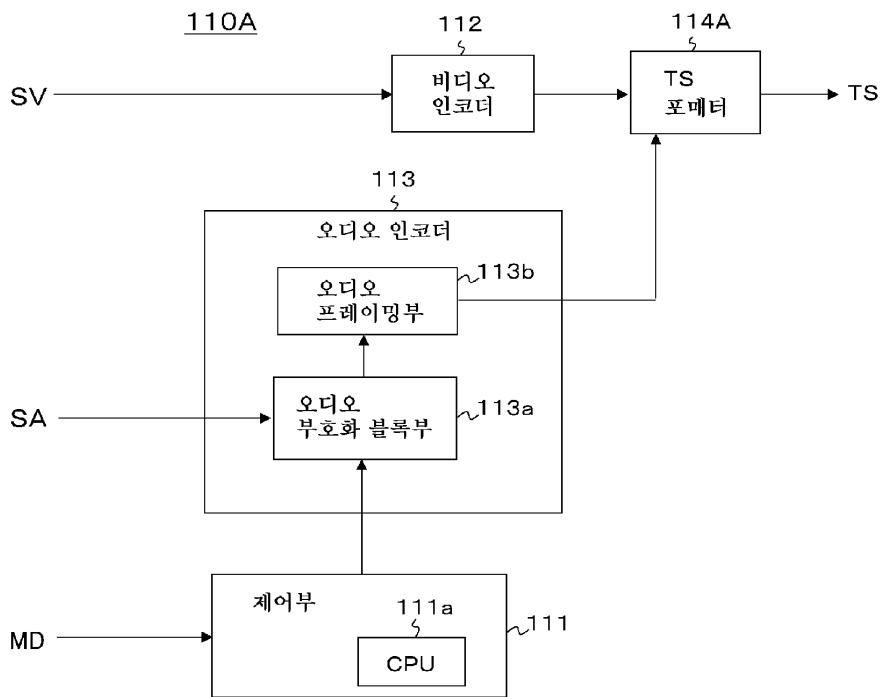
도면33



도면34



도면35



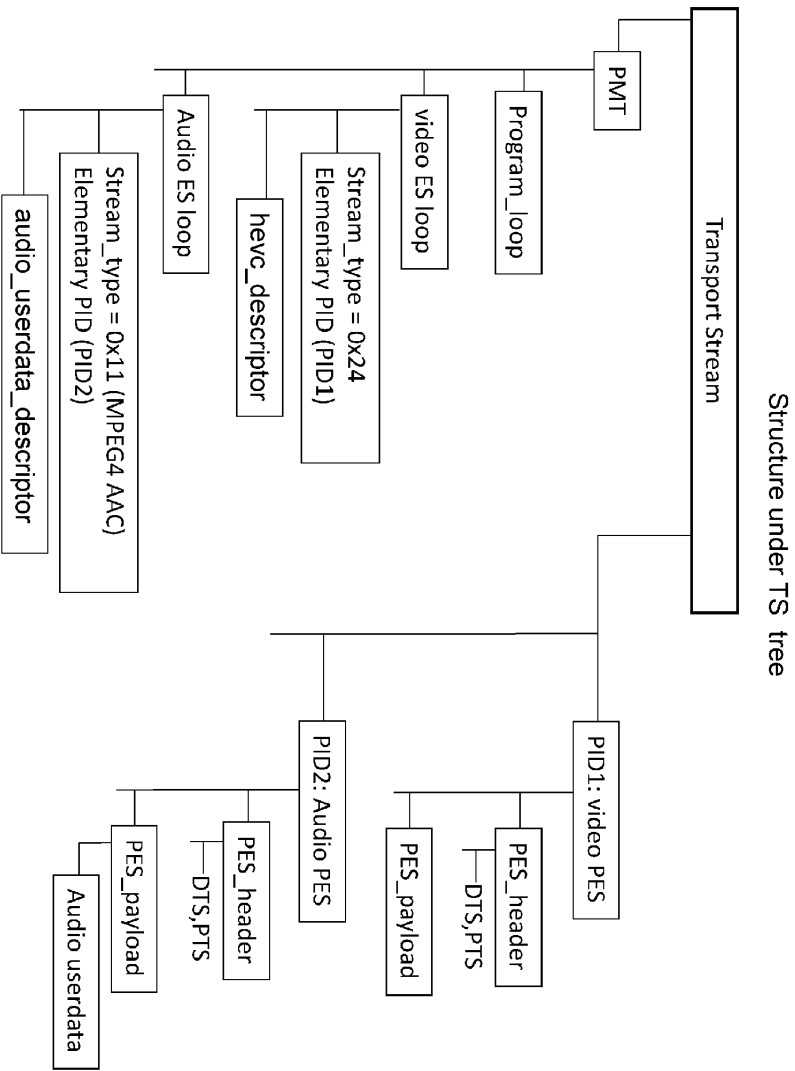
도면36

audio\_userdata\_descriptor

Syntax	No. of Bits	Format
audio_userdata_descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
audio_codec_type	8	uimsbf
metadata_type	3	uimsbf
coordinated_control_flag	1	bslbf
reserved	1	
frequency_type	3	uimsbf
}		

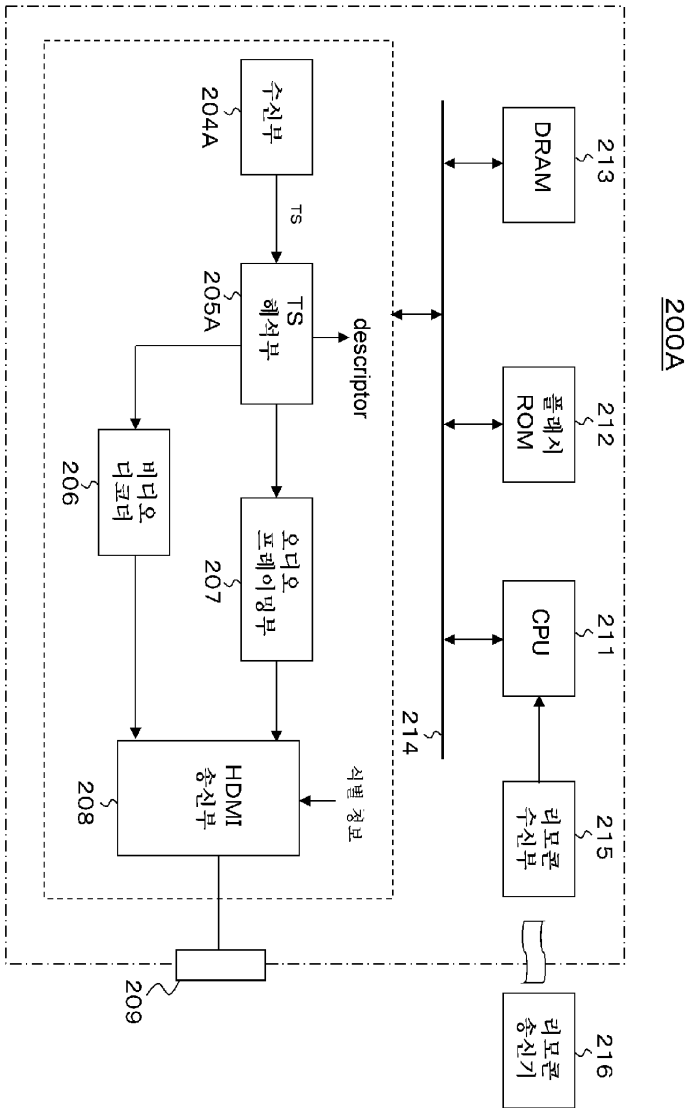
도면37

Semantics of descriptor	
<b>audio_codec_type</b> (8bits)	describes the type of audio codec.
0x1	MPEGH
0x2	AAC
0x3	AC3
0x4	AC4
others	reserved
<b>metadata_type</b> (3bits)	describes the type of information carried in userdata.
0x1	Unconditioned character to describe URL access, which specifies to contain ATSC SDO_payload().
others	reserved
<b>coordinated_control_flag</b> (1bit)	describes whether the information for control is carried in multiple components.
'1'	the control information is provided by the other component stream as well. (such as video stream)
'0'	the control information is provided by audio stream only.
<b>frequency_type</b> (3bits)	describes the frequency of userdata insertion among the access units.
0x1	exactly one userdata is inserted every audio access unit
0x2	more than one userdata is inserted in an audio access unit
0x3	at least one userdata is inserted at the top of every random access point
others	reserved



도면38

도면39



도면40

