

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2023-85253  
(P2023-85253A)

(43)公開日 令和5年6月20日(2023.6.20)

(51)国際特許分類	F I		
H 0 4 N 21/434 (2011.01)	H 0 4 N	21/434	
H 0 4 N 21/438 (2011.01)	H 0 4 N	21/438	
G 1 0 L 19/008 (2013.01)	G 1 0 L	19/008	2 0 0
H 0 4 H 20/28 (2008.01)	G 1 0 L	19/008	1 0 0
H 0 4 H 20/95 (2008.01)	H 0 4 H	20/28	
審査請求 有		請求項の数	8 O L (全27頁) 最終頁に続く

(21)出願番号	特願2023-30769(P2023-30769)	(71)出願人	000002185
(22)出願日	令和5年3月1日(2023.3.1)		ソニーグループ株式会社
(62)分割の表示	特願2021-110252(P2021-110252)		東京都港区港南1丁目7番1号
	)の分割	(74)代理人	100093241
原出願日	平成27年8月31日(2015.8.31)		弁理士 宮田 正昭
(31)優先権主張番号	特願2014-180592(P2014-180592)	(74)代理人	100101801
(32)優先日	平成26年9月4日(2014.9.4)		弁理士 山田 英治
(33)優先権主張国・地域又は機関	日本国(JP)	(74)代理人	100095496
			弁理士 佐々木 榮二
		(74)代理人	100086531
			弁理士 澤田 俊夫
		(74)代理人	110000763
			弁理士法人大同特許事務所
		(72)発明者	塚越 郁夫
			東京都港区港南1丁目7番1号 ソニー
			最終頁に続く

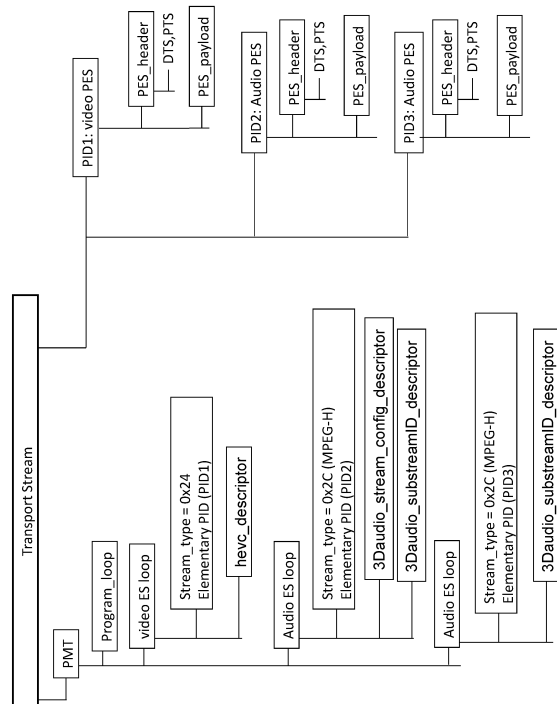
(54)【発明の名称】 受信装置および受信方法

(57)【要約】

【課題】複数種類のオーディオデータを送信する場合において受信側の処理負荷を軽減する。

【解決手段】複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを送信する。例えば、複数のグループの符号化データには、チャンネル符号化データおよびオブジェクト符号化データのいずれかまたは双方が含まれる。コンテナのレイヤに、複数のグループの符号化データのそれぞれの属性を示す属性情報を挿入する。例えば、コンテナのレイヤに、複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すストリーム対応関係情報をさらに挿入する。

【選択図】図13



**【特許請求の範囲】****【請求項 1】**

複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを受信する受信部を備え、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されており、

上記所定数のオーディオストリームから上記属性情報に基づいて取り出された所定の属性を持つグループの符号化データに対してデコード処理をしてスピーカシステムの各スピーカを駆動するためのオーディオデータを得るオーディオデコード部をさらに備える

受信装置。

10

**【請求項 2】**

上記所定の属性を持つグループの符号化データには、チャンネル符号化データおよびオブジェクト符号化データのいずれかまたは双方が含まれる

請求項 1 に記載の受信装置。

**【請求項 3】**

上記オーディオデコード部は、

上記チャンネル符号化データをデコードするときは、上記スピーカシステムのスピーカ構成へのダウンミックスまたはアップミックスの処理を行って上記各スピーカを駆動するためのオーディオデータを得る

請求項 2 に記載の受信装置。

20

**【請求項 4】**

上記オーディオデコード部は、

上記オブジェクト符号化データをデコードするときは、オブジェクト情報に基づき上記スピーカシステムのスピーカ構成に対応したスピーカレンダリング処理を行って上記各スピーカを駆動するためのオーディオデータを得る

請求項 2 に記載の受信装置。

**【請求項 5】**

上記オブジェクト情報は、上記オブジェクト符号化データのメタデータである

請求項 4 に記載の受信装置。

**【請求項 6】**

上記メタデータは、エクステンションエレメント (Ext\_element) として含まれる

請求項 5 に記載の受信装置。

30

**【請求項 7】**

上記所定数のオーディオストリームから、上記属性情報と、スピーカ構成およびユーザ選択情報に基づいて、上記所定の属性を持つグループの符号化データを取り出す符号化データ取り出し部をさらに備える

請求項 1 に記載の受信装置。

**【請求項 8】**

複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを受信する受信ステップを有し、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されており、

上記所定数のオーディオストリームから上記属性情報に基づいて取り出された所定の属性を持つグループの符号化データに対してデコード処理をしてスピーカシステムの各スピーカを駆動するためのオーディオデータを得るオーディオデコードステップをさらに有する

受信方法。

40

**【発明の詳細な説明】****【技術分野】****【0001】**

50

本技術は、受信装置および受信方法に関する。

【背景技術】

【0002】

従来、立体(3D)音響技術として、符号化サンプルデータをメタデータに基づいて任意の位置に存在するスピーカにマッピングさせてレンダリングする技術が提案されている(例えば、特許文献1参照)。

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特表2014-520491号公報

10

【発明の概要】

【発明が解決しようとする課題】

【0004】

5.1チャンネル、7.1チャンネルなどのチャンネル符号化データと共に、符号化サンプルデータおよびメタデータからなるオブジェクト符号化データを送信し、受信側において臨場感を高めた音響再生を可能とすることが考えられる。

【0005】

本技術の目的は、複数種類のオーディオデータを送信する場合にあって受信側の処理負荷を軽減することにある。

【課題を解決するための手段】

20

【0006】

本技術の概念は、

複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを送信する送信部と、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報を挿入する情報挿入部とを備える

送信装置にある。

【0007】

本技術において、送信部により、複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナが送信される。例えば、複数のグループの符号化データには、チャンネル符号化データおよびオブジェクト符号化データのいずれかまたは双方が含まれる、ようにされてもよい。

30

【0008】

情報挿入部により、コンテナのレイヤに、複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入される。例えば、コンテナは、デジタル放送規格で採用されているトランスポートストリーム(MPEG-2 TS)であってもよい。また、例えば、コンテナは、インターネットの配信などで用いられるMP4、あるいはそれ以外のフォーマットのコンテナであってもよい。

【0009】

このように本技術においては、コンテナのレイヤに、所定数のオーディオストリームに含まれる複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入される。そのため、受信側では、複数のグループの符号化データのそれぞれの属性を当該符号化データのデコード前に容易に認識でき、必要なグループの符号化データのみを選択的にデコードして用いることができ、処理負荷を軽減することが可能となる。

40

【0010】

なお、本技術において、例えば、情報挿入部は、コンテナのレイヤに、複数のグループの符号化データがそれぞれどのオーディオストリームに含まれるかを示すストリーム対応関係情報をさらに挿入する、ようにされてもよい。この場合、例えば、コンテナは、MPEG2-TSであり、情報挿入部は、属性情報およびストリーム対応関係情報を、プログラムマップテーブルの配下に存在する上記所定数のオーディオストリームのいずれか1つ

50

のオーディオストリームに対応したオーディオエレメンタリストリームループ内に挿入する、ようにされてもよい。このようにコンテナのレイヤにストリーム対応関係情報が挿入されることで、受信側では、必要なグループの符号化データが含まれるオーディオストリームを容易に認識でき、処理負荷を軽減することが可能となる。

**【0011】**

例えば、ストリーム対応関係情報は、複数のグループの符号化データのそれぞれを識別するグループ識別子と所定数のオーディオストリームのそれぞれのストリームを識別するストリーム識別子との対応関係を示す情報である、ようにされてもよい。この場合、例えば、情報挿入部は、コンテナのレイヤに、所定数のオーディオストリームのそれぞれのストリーム識別子を示すストリーム識別子情報をさらに挿入する、ようにされてもよい。例えば、コンテナは、MPEG2-TSであり、情報挿入部は、ストリーム識別子情報を、プログラムマップテーブルの配下に存在する所定数のオーディオストリームのそれぞれに対応したオーディオエレメンタリストリームループ内に挿入する、ようにされてもよい。

10

**【0012】**

また、例えば、ストリーム対応関係情報は、複数のグループの符号化データのそれぞれを識別するグループ識別子と所定数のオーディオストリームのそれぞれをパケット化する際に付されるパケット識別子との対応関係を示す情報である、ようにされてもよい。また、例えば、ストリーム対応関係情報は、複数のグループの符号化データのそれぞれを識別するグループ識別子と所定数のオーディオストリームのそれぞれのストリームタイプを示すタイプ情報との対応関係を示す情報である、ようにされてもよい。

20

**【0013】**

また、本技術の他の概念は、

複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを受信する受信部を備え、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されており、

上記受信されたコンテナが有する上記所定数のオーディオストリームを、上記属性情報に基づいて処理する処理部をさらに備える

受信装置にある。

**【0014】**

本技術において、受信部により、複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナが受信される。例えば、複数のグループの符号化データには、チャンネル符号化データおよびオブジェクト符号化データのいずれかまたは双方が含まれる、ようにされてもよい。コンテナのレイヤに、複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されている。処理部により、受信されたコンテナが有する所定数のオーディオストリームが、その属性情報に基づいて処理される。

30

**【0015】**

このように本技術においては、コンテナのレイヤに挿入されている複数のグループの符号化データのそれぞれの属性を示す属性情報に基づいて、受信されたコンテナが有する所定数のオーディオストリームの処理が行われる。そのため、必要なグループの符号化データのみを選択的に復号化して用いることができ、処理負荷を軽減することが可能となる。

40

**【0016】**

なお、本技術において、例えば、コンテナのレイヤには、複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すストリーム対応関係情報がさらに挿入されており、処理部は、属性情報の他に、ストリーム対応関係情報に基づいて所定数のオーディオストリームを処理する、ようにされてもよい。この場合、必要なグループの符号化データが含まれるオーディオストリームを容易に認識でき、処理負荷を軽減することが可能となる。

**【0017】**

50

また、本技術において、例えば、処理部は、属性情報およびストリーム対応関係情報に基づいて、スピーカ構成およびユーザ選択情報に適合する属性を持つグループの符号化データを含むオーディオストリームに対して選択的にデコード処理を施す、ようにされてもよい。

**【0018】**

また、本技術のさらに他の概念は、

複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを受信する受信部を備え、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されており、

上記受信されたコンテナが有する上記所定数のオーディオストリームから上記属性情報に基づいて所定のグループの符号化データを選択的に取得し、該所定のグループの符号化データを含むオーディオストリームを再構成する処理部と、

上記処理部で再構成されたオーディオストリームを外部機器に送信するストリーム送信部とをさらに備える

受信装置にある。

**【0019】**

本技術において、受信部により、複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナが受信される。コンテナのレイヤに、複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されている。処理部により、所定数のオーディオストリームから属性情報に基づいて所定のグループの符号化データが選択的に取得され、この所定のグループの符号化データを含むオーディオストリームが再構成される。そして、ストリーム送信部により、再構成されたオーディオストリームが外部機器に送信される。

**【0020】**

このように本技術においては、コンテナのレイヤに挿入されている複数のグループの符号化データのそれぞれの属性を示す属性情報に基づいて、所定数のオーディオストリームから所定のグループの符号化データが選択的に取得されて、外部機器に送信すべきオーディオストリームが再構成される。必要なグループの符号化データの取得を容易に行うことができ、処理負荷を軽減することが可能となる。

**【0021】**

なお、本技術において、例えば、コンテナのレイヤには、複数のグループの符号化データがそれぞれどのオーディオストリームに含まれるかを示すストリーム対応関係情報をさらに挿入されており、処理部は、属性情報の他に、ストリーム対応関係情報に基づいて、所定数のオーディオストリームから所定のグループの符号化データを選択的に取得する、ようにされてもよい。この場合、所定グループの符号化データが含まれるオーディオストリームを容易に認識でき、処理負荷を軽減することが可能となる。

**【発明の効果】****【0022】**

本技術によれば、複数種類のオーディオデータを送信する場合にあって受信側の処理負荷を軽減することが可能となる。なお、本明細書に記載された効果はあくまで例示であって限定されるものではなく、また付加的な効果があってもよい。

**【図面の簡単な説明】****【0023】**

【図1】実施の形態としての送受信システムの構成例を示すブロック図である。

【図2】3Dオーディオの伝送データにおけるオーディオフレーム(1024サンプル)の構造を示す図である。

【図3】3Dオーディオの伝送データの構成例を示す図である。

【図4】3Dオーディオの伝送データを1ストリームで送信する場合および複数ストリームで送信する場合におけるオーディオフレームの構成例を概略的に示す図である。

10

20

30

40

50

【図 5】3Dオーディオの伝送データの構成例において、3ストリームで送信する場合のグループ分割例を示す図である。

【図 6】グループ分割例（3分割）におけるグループとサブストリームの対応関係などを示す図である。

【図 7】3Dオーディオの伝送データの構成例において、2ストリームで送信する場合のグループ分割例を示す図である。

【図 8】グループ分割例（2分割）におけるグループとサブストリームの対応関係などを示す図である。

【図 9】サービス送信機が備えるストリーム生成部の構成例を示すブロック図である。

【図 10】3Dオーディオ・ストリーム・コンフィグ・デスク립タの構造例を示す図である。 10

【図 11】3Dオーディオ・ストリーム・コンフィグ・デスク립タの構造例における主要な情報の内容を示す図である。

【図 12】3Dオーディオ・サブストリームID・デスク립タの構造例と、その構造例における主要な情報の内容を示す図である。

【図 13】トランスポートストリームの構成例を示す図である。

【図 14】サービス受信機の構成例を示すブロック図である。

【図 15】サービス受信機におけるCPUのオーディオデコード制御処理の一例を示すフローチャートである。

【図 16】サービス受信機の他の構成例を示すブロック図である。 20

【発明を実施するための形態】

【0024】

以下、発明を実施するための形態（以下、「実施の形態」とする）について説明する。なお、説明を以下の順序で行う。

1. 実施の形態

2. 変形例

【0025】

< 1. 実施の形態 >

[送受信システムの構成例]

図 1 は、実施の形態としての送受信システム 10 の構成例を示している。この送受信システム 10 は、サービス送信機 100 とサービス受信機 200 により構成されている。サービス送信機 100 は、トランスポートストリーム TS を、放送波あるいはネットの packets に載せて送信する。このトランスポートストリーム TS は、ビデオストリームと、複数のグループの符号化データを含む所定数のオーディオストリームを有している。 30

【0026】

図 2 は、この実施の形態で取り扱う 3Dオーディオの伝送データにおけるオーディオフレーム（1024 サンプル）の構造を示している。このオーディオフレームは、複数の MPEGオーディオストリームパケット（mpeg Audio Stream Packet）からなっている。各 MPEGオーディオストリームパケットは、ヘッダ（Header）とペイロード（Payload）により構成されている。 40

【0027】

ヘッダは、パケットタイプ（Packet Type）、パケットラベル（Packet Label）、パケット長（Packet Length）などの情報を持つ。ペイロードには、ヘッダのパケットタイプで定義された情報が配置される。このペイロード情報には、同期スタートコードに相当する“SYNC”情報と、3Dオーディオの伝送データの実際のデータである“Frame”情報と、この“Frame”情報の構成を示す“Config”情報が存在する。

【0028】

“Frame”情報には、3Dオーディオの伝送データを構成するチャンネル符号化データとオブジェクト符号化データが含まれる。ここで、チャンネル符号化データは、SCE(S 50

ingle Channel Element)、CPE(Channel Pair Element)、LFE(Low Frequency Element)などの符号化サンプルデータで構成される。また、オブジェクト符号化データは、SCE(Single Channel Element)の符号化サンプルデータと、それを任意の位置に存在するスピーカにマッピングさせてレンダリングするためのメタデータにより構成される。このメタデータは、エクステンションエレメント(Ext\_element)として含まれる。

【0029】

図3は、3Dオーディオの伝送データの構成例を示している。この例では、1つのチャンネル符号化データと2つのオブジェクト符号化データとからなっている。1つのチャンネル符号化データは、5.1チャンネルのチャンネル符号化データ(CD)であり、SCE1, CPE1.1, CPE1.2, LFE1の各符号化サンプルデータからなっている。

10

【0030】

2つのオブジェクト符号化データは、イマーシブオーディオオブジェクト(IAO: Immersive audio object)およびスピーチダイアログオブジェクト(SDO: Speech Dialog object)の符号化データである。イマーシブオーディオオブジェクト符号化データは、イマーシブサウンドのためのオブジェクト符号化データであり、符号化サンプルデータSCE2と、それを任意の位置に存在するスピーカにマッピングさせてレンダリングするためのメタデータEXE\_\_E1(Object metadata)2とからなっている。

【0031】

スピーチダイアログオブジェクト符号化データは、スピーチランゲージのためのオブジェクト符号化データである。この例では、第1、第2の言語のそれぞれに対応したスピーチダイアログオブジェクト符号化データが存在する。第1の言語に対応したスピーチダイアログオブジェクト符号化データは、符号化サンプルデータSCE3と、それを任意の位置に存在するスピーカにマッピングさせてレンダリングするためのメタデータEXE\_\_E1(Object metadata)3とからなっている。また、第2の言語に対応したスピーチダイアログオブジェクト符号化データは、符号化サンプルデータSCE4と、それを任意の位置に存在するスピーカにマッピングさせてレンダリングするためのメタデータEXE\_\_E1(Object metadata)4とからなっている。

20

【0032】

符号化データは、種類別にグループ(Group)という概念で区別される。図示の例では、5.1チャンネルの符号化チャンネルデータはグループ1とされ、イマーシブオーディオオブジェクト符号化データはグループ2とされ、第1の言語に係るスピーチダイアログオブジェクト符号化データはグループ3とされ、第2の言語に係るスピーチダイアログオブジェクト符号化データはグループ4とされている。

30

【0033】

また、受信側においてグループ間で選択できるものはスイッチグループ(SW Group)に登録されて符号化される。また、グループを束ねてプリセットグループ(preset Group)とされ、ユースケースに応じた再生が可能とされる。図示の例では、グループ1、グループ2およびグループ3が束ねられてプリセットグループ1とされ、グループ1、グループ2およびグループ4が束ねられてプリセットグループ2とされている。

40

【0034】

図1に戻って、サービス送信機100は、上述したように複数のグループの符号化データを含む3Dオーディオの伝送データを、1ストリーム、あるいは複数ストリーム(Multiple stream)で送信する。

【0035】

図4(a)は、図3の3Dオーディオの伝送データの構成例において、1ストリームで送信する場合におけるオーディオフィレームの構成例を概略的に示している。この場合、この1ストリームに、“SYNC”情報と“Config”情報と共に、チャンネル符号化データ(CD)、イマーシブオーディオオブジェクト符号化データ(IAO)、スピーチダイアログオブジェクト符号化データ(SDO)が含まれる。

50

## 【 0 0 3 6 】

図 4 ( b ) は、図 3 の 3 D オーディオの伝送データの構成例において、複数ストリーム ( 各ストリームを、適宜、「サブストリーム」と呼ぶこととする )、ここでは 3 ストリームで送信する場合におけるオーディオフレームの構成例を概略的に示している。この場合、サブストリーム 1 に、“ S Y N C ” 情報と “ C o n f i g ” 情報と共に、チャンネル符号化データ ( C D ) が含まれる。また、サブストリーム 2 に、“ S Y N C ” 情報と “ C o n f i g ” 情報と共に、イマーシブオーディオオブジェクト符号化データ ( I A O ) が含まれる。さらに、サブストリーム 3 に、“ S Y N C ” 情報と “ C o n f i g ” 情報と共に、スピーチダイアログオブジェクト符号化データ ( S D O ) が含まれる。

## 【 0 0 3 7 】

10

図 5 は、図 3 の 3 D オーディオの伝送データの構成例において、3 ストリームで送信する場合のグループ分割例を示している。この場合、サブストリーム 1 には、グループ 1 として区別されるチャンネル符号化データ ( C D ) が含まれる。また、サブストリーム 2 には、グループ 2 として区別されるイマーシブオーディオオブジェクト符号化データ ( I A O ) が含まれる。また、サブストリーム 3 には、グループ 3 として区別される第 1 の言語のスピーチダイアログオブジェクト符号化データ ( S D O ) と、グループ 4 として区別される第 2 の言語のスピーチダイアログオブジェクト符号化データ ( S D O ) が含まれる。

## 【 0 0 3 8 】

図 6 は、図 5 のグループ分割例 ( 3 分割 ) におけるグループとサブストリームの対応関係などを示している。ここで、グループ ID ( group ID ) は、グループを識別するための識別子である。アトリビュート ( attribute ) は、各グループの符号化データの属性を示している。スイッチグループ ID ( switch Group ID ) は、スイッチンググループを識別するための識別子である。プリセットグループ ID ( preset Group ID ) は、プリセットグループを識別するための識別子である。サブストリーム ID ( sub Stream ID ) は、サブストリームを識別するための識別子である。

20

## 【 0 0 3 9 】

図示の対応関係は、グループ 1 に属する符号化データは、チャンネル符号化データであって、スイッチグループを構成しておらず、サブストリーム 1 に含まれている、ことを示している。また、図示の対応関係は、グループ 2 に属する符号化データは、イマーシブサウンドのためのオブジェクト符号化データ ( イマーシブオーディオオブジェクト符号化データ ) であって、スイッチグループを構成しておらず、サブストリーム 2 に含まれている、ことを示している。

30

## 【 0 0 4 0 】

また、図示の対応関係は、グループ 3 に属する符号化データは、第 1 の言語のスピーチランゲージのためのオブジェクト符号化データ ( スピーチダイアログオブジェクト符号化データ ) であって、スイッチグループ 1 を構成しており、サブストリーム 3 に含まれている、ことを示している。また、図示の対応関係は、グループ 4 に属する符号化データは、第 2 の言語のスピーチランゲージのためのオブジェクト符号化データ ( スピーチダイアログオブジェクト符号化データ ) であって、スイッチグループ 1 を構成しており、サブストリーム 3 に含まれている、ことを示している。

40

## 【 0 0 4 1 】

また、図示の対応関係は、プリセットグループ 1 は、グループ 1、グループ 2 およびグループ 3 を含む、ことが示されている。さらに、図示の対応関係は、プリセットグループ 2 は、グループ 1、グループ 2 およびグループ 4 を含む、ことが示されている。

## 【 0 0 4 2 】

図 7 は、図 3 の 3 D オーディオの伝送データの構成例において、2 ストリームで送信する場合のグループ分割例を示している。この場合、サブストリーム 1 には、グループ 1 として区別されるチャンネル符号化データ ( C D ) と、グループ 2 として区別されるイマーシブオーディオオブジェクト符号化データ ( I A O ) が含まれる。また、サブストリーム 2 には、グループ 3 として区別される第 1 の言語のスピーチダイアログオブジェクト符号化

50

データ ( S D O ) と、グループ 4 として区別される第 2 の言語のスピーチダイアログオブジェクト符号化データ ( S D O ) が含まれる。

【 0 0 4 3 】

図 8 は、図 7 のグループ分割例 ( 2 分割 ) におけるグループとサブストリームの対応関係などを示している。図示の対応関係は、グループ 1 に属する符号化データは、チャンネル符号化データであって、スイッチグループを構成しておらず、サブストリーム 1 に含まれている、ことを示している。また、図示の対応関係は、グループ 2 に属する符号化データは、イマーシブサウンドのためのオブジェクト符号化データ ( イマーシブオーディオオブジェクト符号化データ ) であって、スイッチグループを構成しておらず、サブストリーム 1 に含まれている、ことを示している。

10

【 0 0 4 4 】

また、図示の対応関係は、グループ 3 に属する符号化データは、第 1 の言語のスピーチランゲージのためのオブジェクト符号化データ ( スピーチダイアログオブジェクト符号化データ ) であって、スイッチグループ 1 を構成しており、サブストリーム 2 に含まれている、ことを示している。また、図示の対応関係は、グループ 4 に属する符号化データは、第 2 の言語のスピーチランゲージのためのオブジェクト符号化データ ( スピーチダイアログオブジェクト符号化データ ) であって、スイッチグループ 1 を構成しており、サブストリーム 2 に含まれている、ことを示している。

【 0 0 4 5 】

また、図示の対応関係は、プリセットグループ 1 は、グループ 1、グループ 2 およびグループ 3 を含む、ことが示されている。さらに、図示の対応関係は、プリセットグループ 2 は、グループ 1、グループ 2 およびグループ 4 を含む、ことが示されている。

20

【 0 0 4 6 】

図 1 に戻って、サービス送信機 1 0 0 は、コンテナのレイヤに、3 D オーディオの伝送データに含まれる複数のグループの符号化データのそれぞれの属性を示す属性情報を挿入する。また、サービス送信機 1 0 0 は、コンテナのレイヤに、この複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すストリーム対応関係情報を挿入する。この実施の形態において、このストリーム対応関係情報は、例えば、グループ ID とストリーム識別子との対応関係を示す情報とされる。

【 0 0 4 7 】

サービス送信機 1 0 0 は、これらの属性情報およびストリーム対応関係情報を、例えば、プログラムマップテーブル ( P M T : Program Map Table ) の配下に存在する所定数のオーディオストリームのいずれか 1 つのオーディオストリーム、例えば最も基本的なストリームに対応したオーディオエレメンタリストリームループ内にデスクリプタとして挿入する。

30

【 0 0 4 8 】

また、サービス送信機 1 0 0 は、コンテナのレイヤに、所定数のオーディオストリームのそれぞれのストリーム識別子を示すストリーム識別子情報情報を挿入する。サービス送信機 1 0 0 は、このストリーム識別子情報を、例えば、プログラムマップテーブル ( P M T : Program Map Table ) の配下に存在する所定数のオーディオストリームのそれぞれに対応したオーディオエレメンタリストリームループ内にデスクリプタとして挿入する。

40

【 0 0 4 9 】

サービス受信機 2 0 0 は、サービス送信機 1 0 0 から放送波あるいはネットのパケットに載せて送られてくるトランスポートストリーム T S を受信する。このトランスポートストリーム T S は、上述したように、ビデオストリームの他に、3 D オーディオの伝送データを構成する複数のグループの符号化データを含む所定数のオーディオストリームを有している。そして、コンテナのレイヤに、3 D オーディオの伝送データに含まれる複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されていると共に、この複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すス

50

トリーム対応関係情報が挿入されている。

【 0 0 5 0 】

サービス受信機 2 0 0 は、属性情報およびストリーム対応関係情報に基づいて、スピーカ構成およびユーザ選択情報に適合する属性を持つグループの符号化データを含むオーディオストリームに対して選択的にデコード処理を行って、3 Dオーディオのオーディオ出力を得る。

【 0 0 5 1 】

[ サービス送信機のストリーム生成部 ]

図 9 は、サービス送信機 1 0 0 が備えるストリーム生成部 1 1 0 の構成例を示している。このストリーム生成部 1 1 0 は、ビデオエンコーダ 1 1 2 と、オーディオエンコーダ 1 1 3 と、マルチプレクサ 1 1 4 を有している。ここでは、オーディオの伝送データが、図 3 に示すように、1 つの符号化チャンネルデータと 2 つのオブジェクト符号化データとからなる例とする。

10

【 0 0 5 2 】

ビデオエンコーダ 1 1 2 は、ビデオデータ S V を入力し、このビデオデータ S V に対して符号化を施し、ビデオストリーム（ビデオエレメンタリストリーム）を生成する。オーディオエンコーダ 1 1 3 は、オーディオデータ S A として、チャンネルデータと共に、イマーシブオーディオおよびスピーチダイアログのオブジェクトデータを入力する。

【 0 0 5 3 】

オーディオエンコーダ 1 1 3 は、オーディオデータ S A に対して符号化を施し、3 Dオーディオの伝送データを得る。この 3 Dオーディオの伝送データには、図 3 に示すように、チャンネル符号化データ（C D）と、イマーシブオーディオオブジェクト符号化データ（I A O）と、スピーチダイアログオブジェクト符号化データ（S D O）が含まれる。そして、オーディオエンコーダ 1 1 3 は、複数のグループ、ここでは 4 つのグループの符号化データ含む 1 つまたは複数のオーディオストリーム（オーディオエレメンタリストリーム）を生成する（図 4（a）,（b）参照）。

20

【 0 0 5 4 】

マルチプレクサ 1 1 4 は、ビデオエンコーダ 1 1 2 から出力されるビデオストリームおよびオーディオエンコーダ 1 1 3 から出力される所定数のオーディオストリームを、それぞれ、P E S パケット化し、さらにトランスポートパケット化して多重し、多重化ストリームとしてのトランスポートストリーム T S を得る。

30

【 0 0 5 5 】

また、マルチプレクサ 1 1 4 は、プログラムマップテーブル（P M T）の配下に、複数のグループの符号化データのそれぞれの属性を示す属性情報と、複数のグループの符号化データがそれぞれどのオーディオストリームに含まれるかを示すストリーム対応関係情報を挿入する。マルチプレクサ 1 1 4 は、これらの情報を、例えば最も基本的なストリームに対応したオーディオ・エレメンタリストリームループ内に、3 Dオーディオ・ストリーム・コンフィグ・デスクリプタ（3 D audio\_stream\_config\_descriptor）を用いて挿入する。このデスクリプタの詳細については後述する。

【 0 0 5 6 】

また、マルチプレクサ 1 1 4 は、プログラムマップテーブル（P M T）の配下に、所定数のオーディオストリームのそれぞれのストリーム識別子を示すストリーム識別子情報を挿入する。マルチプレクサ 1 1 4 は、この情報を、所定数のオーディオストリームのそれぞれに対応したオーディオ・エレメンタリストリームループ内に、3 Dオーディオ・サブストリーム I D・デスクリプタ（3 D audio\_substreamID\_descriptor）を用いて挿入する。このデスクリプタの詳細については後述する。

40

【 0 0 5 7 】

図 9 に示すストリーム生成部 1 1 0 の動作を簡単に説明する。ビデオデータは、ビデオエンコーダ 1 1 2 に供給される。このビデオエンコーダ 1 1 2 では、ビデオデータ S V に対して符号化が施され、符号化ビデオデータを含むビデオストリームが生成される。この

50

ビデオストリームは、マルチプレクサ 1 1 4 に供給される。

【 0 0 5 8 】

オーディオデータ S A は、オーディオエンコーダ 1 1 3 に供給される。このオーディオデータ S A には、チャンネルデータと、イマーシブオーディオおよびスピーチダイアログのオブジェクトデータが含まれる。オーディオエンコーダ 1 1 3 では、オーディオデータ S A に対して符号化が施され、3 D オーディオの伝送データが得られる。

【 0 0 5 9 】

この 3 D オーディオの伝送データには、チャンネル符号化データ ( C D ) の他に、イマーシブオーディオオブジェクト符号化データ ( I A O ) と、スピーチダイアログオブジェクト符号化データ ( S D O ) が含まれる ( 図 3 参照 )。そして、このオーディオエンコーダ 1 1 3 では、4 つのグループの符号化データ含む 1 つまたは複数のオーディオストリームが生成される ( 図 4 ( a ) , ( b ) 参照 )。

10

【 0 0 6 0 】

ビデオエンコーダ 1 1 2 で生成されたビデオストリームは、マルチプレクサ 1 1 4 に供給される。また、オーディオエンコーダ 1 1 3 で生成されたオーディオストリームは、マルチプレクサ 1 1 4 に供給される。マルチプレクサ 1 1 4 では、各エンコーダから供給されるストリームが P E S パケット化され、さらにトランスポートパケット化されて多重され、多重化ストリームとしてのトランスポートストリーム T S が得られる。

【 0 0 6 1 】

また、マルチプレクサ 1 1 4 では、例えば最も基本的なストリームに対応したオーディオエレメンタリストリームグループ内に、3 D オーディオ・ストリーム・コンフィグ・デスクリプタが挿入される。このデスクリプタには、複数のグループの符号化データのそれぞれの属性を示す属性情報と、複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すストリーム対応関係情報が含まれている。

20

【 0 0 6 2 】

また、マルチプレクサ 1 1 4 では、所定数のオーディオストリームのそれぞれに対応したオーディオエレメンタリストリームグループ内に、3 D オーディオ・サブストリーム I D ・デスクリプタが挿入される。このデスクリプタには、所定数のオーディオストリームのそれぞれのストリーム識別子を示すストリーム識別子情報が含まれている。

【 0 0 6 3 】

[ 3 D オーディオ・ストリーム・コンフィグ・デスクリプタの詳細 ]

図 1 0 は、3 D オーディオ・ストリーム・コンフィグ・デスクリプタ ( 3 D audio\_stream\_config\_descriptor ) の構造例 ( Syntax ) を示している。また、図 1 1 は、その構造例における主要な情報の内容 ( Semantics ) を示している。

30

【 0 0 6 4 】

「 descriptor\_tag 」の 8 ビットフィールドは、デスクリプタタイプを示す。ここでは、3 D オーディオ・ストリーム・コンフィグ・デスクリプタであることを示す。「 descriptor\_length 」の 8 ビットフィールドは、デスクリプタの長さ ( サイズ ) を示し、デスクリプタの長さとして、以降のバイト数を示す。

【 0 0 6 5 】

「 NumOfGroups, N 」の 8 ビットフィールドは、グループの数を示す。「 NumOfPresetGroups, P 」の 8 ビットフィールドは、プリセットグループの数を示す。グループの数だけ、「 groupID 」の 8 ビットフィールド、「 attribute\_of\_groupID 」の 8 ビットフィールド、「 SwitchGroupID 」の 8 ビットフィールドおよび「 audio\_substreamID 」の 8 ビットフィールドが、繰り返えされる。

40

【 0 0 6 6 】

「 groupID 」のフィールドは、グループの識別子を示す。「 attribute\_of\_groupID 」のフィールドは、該当グループの符号化データの属性を示す。「 SwitchGroupID 」のフィールドは、該当グループがどのスイッチグループに属するかを示す識別子である。「 0 」は、いずれのスイッチグループにも属さないことを示す。「 0 」以外は、配属するス

50

ッチグループを示す。「audio\_substreamID」は、該当グループが含まれるオーディオ・サブストリームを示す識別子である。

【0067】

また、プリセットグループの数だけ、「presetGroupID」の8ビットフィールドおよび「NumOfGroups\_in\_preset, R」の8ビットフィールドが、繰り返される。「presetGroupID」のフィールドは、グループをプリセットした束を示す識別子である。「NumOfGroups\_in\_preset, R」のフィールドは、プリセットグループに属するグループの数を示す。そして、プリセットグループ毎に、それに属するグループの数だけ、「groupID」の8ビットフィールドが繰り返され、プリセットグループに属するグループが示される。本デスクリプタは、拡張デスクリプタの配下に配置されてもよい。

10

【0068】

[3Dオーディオ・サブストリームID・デスクリプタの詳細]

図12(a)は、3Dオーディオ・サブストリームID・デスクリプタ(3DAudio\_substreamID\_descriptor)の構造例(Syntax)を示している。また、図12(b)は、その構造例における主要な情報の内容(Semantics)を示している。

【0069】

「descriptor\_tag」の8ビットフィールドは、デスクリプタタイプを示す。ここでは、3Dオーディオ・サブストリームID・デスクリプタであることを示す。「descriptor\_length」の8ビットフィールドは、デスクリプタの長さ(サイズ)を示し、デスクリプタの長さとして、以降のバイト数を示す。「audio\_substreamID」の8ビット

20

フィールドは、オーディオ・サブストリームの識別子を示す。本デスクリプタは、拡張デスクリプタの配下に配置されてもよい。

【0070】

[トランスポートストリームTSの構成]

図13は、トランスポートストリームTSの構成例を示している。この構成例は、3Dオーディオの伝送データを2ストリームで送信する場合に対応している(図7参照)。この構成例では、PID1で識別されるビデオストリームのPESパケット「video PES」が存在する。また、この構成例では、PID2, PID3でそれぞれ識別される2つのオーディオストリーム(オーディオ・サブストリーム)のPESパケット「audio PES」が存在する。PESパケットは、PESヘッダ(PES\_header)とPESペイロード(PES\_payload)からなっている。PESヘッダには、DTS, PTSのタイムスタンプが挿入されている。多重化の際にPID2とPID3のタイムスタンプを合致させるなど、的確に付すことで両者の間の同期をシステム全体で確保することが可能である。

30

【0071】

ここで、PID2で識別されるオーディオストリームのPESパケット「audio PES」には、グループ1として区別されるチャンネル符号化データ(CD)とグループ2として区別されるイマーシブオーディオオブジェクト符号化データ(IAO)が含まれる。また、PID3で識別されるオーディオストリームのPESパケット「audio PES」には、グループ3として区別される第1の言語のスピーチダイアログオブジェクト符号化データ(SDO)と、グループ4として区別される第2の言語のスピーチダイアログオブジェクト符号化データ(SDO)が含まれる。

40

【0072】

また、トランスポートストリームTSには、PSI(Program Specific Information)として、PMT(Program Map Table)が含まれている。PSIは、トランスポートストリームに含まれる各エレメンタリストリームがどのプログラムに属しているかを記した情報である。PMTには、プログラム全体に関連する情報を記述するプログラム・ループ(Program loop)が存在する。

【0073】

また、PMTには、各エレメンタリストリームに関連した情報を持つエレメンタリストリームループが存在する。この構成例では、ビデオストリームに対応したビデオエレメン

50

タリストリームループ (video ES loop) が存在すると共に、2つのオーディオストリームに対応したオーディオエレメンタリストリームループ (audio ES loop) が存在する【0074】

ビデオエレメンタリストリームループ (video ES loop) には、ビデオストリームに対応して、ストリームタイプ、PID (パケット識別子) 等の情報が配置されると共に、そのビデオストリームに関連する情報を記述するデスクリプタも配置される。このビデオストリームの「Stream\_type」の値は「0x24」に設定され、PID情報は、上述したようにビデオストリームのPESパケット「video PES」に付与されるPID1を示すものとされる。デスクリプタの一つとして、HEVC デスクリプタが配置される。

【0075】

また、オーディオエレメンタリストリームループ (audio ES loop) には、オーディオストリームに対応して、ストリームタイプ、PID (パケット識別子) 等の情報が配置されると共に、そのオーディオストリームに関連する情報を記述するデスクリプタも配置される。このオーディオストリームの「Stream\_type」の値は「0x2C」に設定され、PID情報は、上述したようにオーディオストリームのPESパケット「audio PES」に付与されるPID2を示すものとされる。

【0076】

PID2で識別されるオーディオストリームに対応したオーディオエレメンタリストリームループ (audio ES loop) には、上述した3Dオーディオ・ストリーム・コンフィグ・デスクリプタおよび3Dオーディオ・サブストリームID・デスクリプタの双方が配置される。また、PID2で識別されるオーディオストリームに対応したオーディオエレメンタリストリームループ (audio ES loop) には、上述した3Dオーディオ・サブストリームID・デスクリプタのみが配置される。

【0077】

[ サービス受信機の構成例 ]

図14は、サービス受信機200の構成例を示している。このサービス受信機200は、受信部201と、デマルチプレクサ202と、ビデオデコーダ203と、映像処理回路204と、パネル駆動回路205と、表示パネル206を有している。また、このサービス受信機200は、多重化バッファ211-1~211-Nと、コンバイナ212と、3Dオーディオデコーダ213と、音声出力処理回路214と、スピーカシステム215を有している。また、このサービス受信機200は、CPU221と、フラッシュROM222と、DRAM223と、内部バス224と、リモコン受信部225と、リモコン送信機226を有している。

【0078】

CPU221は、サービス受信機200の各部の動作を制御する。フラッシュROM222は、制御ソフトウェアの格納およびデータの保管を行う。DRAM223は、CPU221のワークエリアを構成する。CPU221は、フラッシュROM222から読み出したソフトウェアやデータをDRAM223上に展開してソフトウェアを起動させ、サービス受信機200の各部を制御する。

【0079】

リモコン受信部225は、リモコン送信機226から送信されたリモートコントロール信号 (リモコンコード) を受信し、CPU221に供給する。CPU221は、このリモコンコードに基づいて、サービス受信機200の各部を制御する。CPU221、フラッシュROM222およびDRAM223は、内部バス224に接続されている。

【0080】

受信部201は、サービス送信機100から放送波あるいはネットのパケットに載せて送られてくるトランスポートストリームTSを受信する。このトランスポートストリームTSは、ビデオストリームの他に、3Dオーディオの伝送データを構成する複数のグループの符号化データを含む所定数のオーディオストリームを有している。

【0081】

10

20

30

40

50

デマルチプレクサ 202 は、トランスポートストリーム TS からビデオストリームのパケットを抽出し、ビデオデコーダ 203 に送る。ビデオデコーダ 203 は、デマルチプレクサ 202 で抽出されたビデオのパケットからビデオストリームを再構成し、デコード処理を行って非圧縮のビデオデータを得る。

【0082】

映像処理回路 204 は、ビデオデコーダ 203 で得られたビデオデータに対してスケールリング処理、画質調整処理などを行って、表示用のビデオデータを得る。パネル駆動回路 205 は、映像処理回路 204 で得られる表示用の画像データに基づいて、表示パネル 206 を駆動する。表示パネル 206 は、例えば、LCD (Liquid Crystal Display)、有機 EL ディスプレイ (organic electroluminescence display) などによって構成されている。

10

【0083】

また、デマルチプレクサ 202 は、トランスポートストリーム TS から各種デスクリプタなどの情報を抽出し、CPU 221 に送る。各種デスクリプタには、上述した 3D オーディオ・ストリーム・コンフィグ・デスクリプタ (3DAudio\_stream\_config\_descriptor) や 3D オーディオ・サブストリーム ID・デスクリプタ (3DAudio\_substream ID\_descriptor) も含まれる (図 13 参照)。

【0084】

CPU 221 は、これらのデスクリプタに含まれている、各グループの符号化データの属性を示す属性情報、各グループがどのオーディオストリーム (サブストリーム) に含まれるかを示すストリーム関係情報などに基づいて、スピーカ構成および視聴者 (ユーザ) 選択情報に適合する属性を持つグループの符号化データが含まれるオーディオストリームを認識する。

20

【0085】

また、デマルチプレクサ 202 は、CPU 221 の制御のもと、トランスポートストリーム TS が有する所定数のオーディオストリームのうち、スピーカ構成および視聴者 (ユーザ) 選択情報に適合する属性を持つグループの符号化データを含む一つまたは複数のオーディオストリームのパケットを PID フィルタで選択的に取り出す。

【0086】

多重化バッファ 211-1 ~ 211-N は、それぞれ、デマルチプレクサ 202 で取り出される各オーディオストリームを取り込む。ここで、多重化バッファ 211-1 ~ 211-N の個数 N としては必要十分な個数とされるが、実際の動作では、デマルチプレクサ 202 で取り出されるオーディオストリームの数だけ用いられることになる。

30

【0087】

コンバイナ 212 は、多重化バッファ 211-1 ~ 211-N のうちデマルチプレクサ 202 で取り出される各オーディオストリームがそれぞれ取り込まれた多重化バッファから、オーディオフレーム毎に、オーディオストリームを読み出し、3D オーディオデコーダ 213 にスピーカ構成および視聴者 (ユーザ) 選択情報に適合する属性を持つグループの符号化データとして供給する。

【0088】

3D オーディオデコーダ 213 は、コンバイナ 212 から供給される符号化データに対してデコード処理を施し、スピーカシステム 215 の各スピーカを駆動するためのオーディオデータを得る。ここで、デコード処理される符号化データは、チャンネル符号化データのみを含む場合、オブジェクト符号化データのみを含む場合、さらにはチャンネル符号化データおよびオブジェクト符号化データの双方を含む場合の 3 通りが考えられる。

40

【0089】

3D オーディオデコーダ 213 は、チャンネル符号化データをデコードするときは、スピーカシステム 215 のスピーカ構成へのダウンミックスやアップミックスの処理を行って、各スピーカを駆動するためのオーディオデータを得る。また、3D オーディオデコーダ 213 は、オブジェクト符号化データをデコードするときは、オブジェクト情報 (メタデ

50

ータ)に基づきスピーカレンダリング(各スピーカへのミキシング割合)を計算し、その計算結果に応じて、オブジェクトのオーディオデータを、各スピーカを駆動するためのオーディオデータにミキシングする。

【0090】

音声出力処理回路214は、3Dオーディオデコーダ213で得られた各スピーカを駆動するためのオーディオデータに対して、D/A変換や増幅等の必要な処理を行って、スピーカシステム215に供給する。スピーカシステム215は、複数チャンネル、例えば2チャンネル、5.1チャンネル、7.1チャンネル、22.2チャンネルなどの複数のスピーカを備える。

【0091】

図14に示すサービス受信機200の動作を簡単に説明する。受信部201では、サービス送信機100から放送波あるいはネットの packets に載せて送られてくるトランスポートストリームTSが受信される。このトランスポートストリームTSは、ビデオストリームの他に、3Dオーディオの伝送データを構成する複数のグループの符号化データを含む所定数のオーディオストリームを有している。このトランスポートストリームTSは、デマルチプレクサ202に供給される。

【0092】

デマルチプレクサ202では、トランスポートストリームTSからビデオストリームの packets が抽出され、ビデオデコーダ203に供給される。ビデオデコーダ203では、デマルチプレクサ202で抽出されたビデオの packets からビデオストリームが再構成され、デコード処理が行われて、非圧縮のビデオデータが得られる。このビデオデータは、映像処理回路204に供給される。

【0093】

映像処理回路204では、ビデオデコーダ203で得られたビデオデータに対してスケール処理、画質調整処理などが行われて、表示用のビデオデータが得られる。この表示用のビデオデータはパネル駆動回路205に供給される。パネル駆動回路205では、表示用のビデオデータに基づいて、表示パネル206を駆動することが行われる。これにより、表示パネル206には、表示用のビデオデータに対応した画像が表示される。

【0094】

また、デマルチプレクサ202では、トランスポートストリームTSから各種デスクリプタなどの情報が抽出され、CPU221に送られる。各種デスクリプタには、3Dオーディオ・ストリーム・コンフィグ・デスクリプタや3Dオーディオ・サブストリームID・デスクリプタも含まれる。CPU221では、これらのデスクリプタに含まれている属性情報、ストリーム関係情報などに基づいて、スピーカ構成および視聴者(ユーザ)選択情報に適合する属性を持つグループの符号化データが含まれるオーディオストリーム(サブストリーム)が認識される。

【0095】

また、デマルチプレクサ202では、CPU221の制御のもと、トランスポートストリームTSが有する所定数のオーディオストリームのうち、スピーカ構成および視聴者選択情報に適合する属性を持つグループの符号化データを含む1つまたは複数のオーディオストリームの packets がPIDフィルタで選択的に取り出される。

【0096】

デマルチプレクサ202で取り出されたオーディオストリームは、多重化バッファ211-1~211-Nのうち対応する多重化バッファに取り込まれる。コンバイナ212では、オーディオストリームが取り込まれた各多重化バッファから、オーディオフレーム毎に、オーディオストリームが読み出され、3Dオーディオデコーダ213にスピーカ構成および視聴者選択情報に適合する属性を持つグループの符号化データとして供給される。

【0097】

3Dオーディオデコーダ213では、コンバイナ212から供給される符号化データに対してデコード処理が施され、スピーカシステム215の各スピーカを駆動するためのオ

10

20

30

40

50

オーディオデータが得られる。

【0098】

ここで、チャンネル符号化データがデコードされる時は、スピーカシステム215のスピーカ構成へのダウンミックスやアップミックスの処理が行われて、各スピーカを駆動するためのオーディオデータが得られる。また、オブジェクト符号化データがデコードされる時は、オブジェクト情報(メタデータ)に基づきスピーカレンダリング(各スピーカへのミキシング割合)が計算され、その計算結果に応じて、オブジェクトのオーディオデータが各スピーカを駆動するためのオーディオデータにミキシングされる。

【0099】

3Dオーディオデコーダ213で得られた各スピーカを駆動するためのオーディオデータは、音声出力処理回路214に供給される。この音声出力処理回路214では、各スピーカを駆動するためのオーディオデータに対して、D/A変換や増幅等の必要な処理が行われる。そして、処理後のオーディオデータはスピーカシステム215に供給される。これにより、スピーカシステム215からは表示パネル206の表示画像に対応した音響出力が得られる。

【0100】

図15は、図14に示すサービス受信機200におけるCPU221のオーディオデコード制御処理の一例を示している。CPU221は、ステップST1において、処理を開始する。そして、CPU221は、ステップST2において、受信機スピーカ構成、つまりスピーカシステム215のスピーカ構成を検知する。次に、CPU221は、ステップST3において、視聴者(ユーザ)によるオーディオ出力に関する選択情報を得る。

【0101】

次に、CPU221は、ステップST4において、3Dオーディオ・ストリーム・コンフィグ・デスク립タ(3Daudio\_stream\_config\_descriptor)の「groupID」、「attribute\_of\_GroupID」、「switchGroupID」、「presetGroupID」、「Audio\_substreamID」を読む。そして、CPU221は、ステップST5において、スピーカ構成、視聴者選択情報に適合する属性を持つグループが属するオーディオストリーム(サブストリーム)のサブストリームID(subStreamID)を認識する。

【0102】

次に、CPU221は、ステップST6において、認識されたサブストリームID(subStreamID)と、各オーディオストリーム(サブストリーム)の3Dオーディオ・サブストリームID・デスク립タ(3Daudio\_substreamID\_descriptor)のサブストリームID(subStreamID)と照合し、合致するものをPIDフィルタ(PID filter)で選択し、多重化バッファに取り込む。そして、CPU221は、ステップST7において、多重化バッファからオーディオフレーム毎に、オーディオストリーム(サブストリーム)を読み出し、3Dオーディオデコーダ213へ必要なグループの符号化データを供給する。

【0103】

次に、CPU221は、ステップST8において、オブジェクト符号化データをデコードするか否かを判断する。オブジェクト符号化データをデコードするとき、CPU221は、ステップST9において、オブジェクト情報(メタデータ)に基づき、アジマス(方位情報)とエレベーション(仰角情報)によりスピーカレンダリング(各スピーカへのミキシング割合)を計算する。その後、CPU221は、ステップST10に進む。なお、ステップST8において、オブジェクト符号化データをデコードしないとき、CPU221は、直ちに、ステップST10に進む。

【0104】

このステップST10において、CPU221は、チャンネル符号化データをデコードするか否かを判断する。チャンネル符号化データをデコードするとき、CPU221は、ステップST11において、スピーカシステム215のスピーカ構成へのダウンミックスやアップミックスの処理を行って、各スピーカを駆動するためのオーディオデータを得る。そ

10

20

30

40

50

の後に、CPU 221は、ステップST12に進む。なお、ステップST10において、オブジェクト符号化データをデコードしないとき、CPU 221は、直ちに、ステップST12に進む。

【0105】

このステップST12において、CPU 221は、オブジェクト符号化データをデコードするとき、ステップST9の計算結果に応じて、オブジェクトのオーディオデータを、各スピーカを駆動するためのオーディオデータにミキシングし、その後、ダイナミックレンジ制御を行う。その後、CPU 221はステップST13において、処理を終了する。なお、オブジェクト符号化データをデコードしないとき、CPU 221は、ステップST12をスキップする。

10

【0106】

上述したように、図1に示す送受信システム10において、サービス送信機100は、コンテナのレイヤに、所定数のオーディオストリームに含まれる複数のグループの符号化データのそれぞれの属性を示す属性情報を挿入する。そのため、受信側では、複数のグループの符号化データのそれぞれの属性を当該符号化データのデコード前に容易に認識でき、必要なグループの符号化データのみを選択的にデコードして用いることができ、処理負荷を軽減することが可能となる。

【0107】

また、図1に示す送受信システム10において、サービス送信機100は、コンテナのレイヤに、複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すストリーム対応関係情報を挿入する。そのため、受信側では、必要なグループの符号化データが含まれるオーディオストリームを容易に認識でき、処理負荷を軽減することが可能となる。

20

【0108】

< 2 . 変形例 >

なお、上述実施の形態において、サービス受信機200は、サービス送信機100から送信されてくる複数のオーディオストリーム（サブストリーム）から、スピーカ構成、視聴者選択情報に適合する属性を持つグループの符号化データが含まれるオーディオストリームを選択的に取り出し、デコード処理を行って所定数のスピーカ駆動用のオーディオデータを得る構成となっている。

30

【0109】

しかし、サービス受信機として、サービス送信機100から送信されてくる複数のオーディオストリーム（サブストリーム）から、スピーカ構成、視聴者選択情報に適合する属性を持つグループの符号化データを持つ1つまたは複数のオーディオストリームを選択的に取り出し、スピーカ構成、視聴者選択情報に適合する属性を持つグループの符号化データを持つオーディオストリームを再構成し、その再構成オーディオストリームを、構内ネットワーク接続されたデバイス（DLNA機器も含む）に配信することも考えられる。

【0110】

図16は、上述したように再構成オーディオストリームを構内ネットワーク接続されたデバイスに配信するサービス受信機200Aの構成例を示している。この図16において、図14と対応する部分には同一符号を付し、適宜、その詳細説明は省略する。

40

【0111】

デマルチプレクサ202では、CPU 221の制御のもと、トランスポートストリームTSが有する所定数のオーディオストリームのうち、スピーカ構成および視聴者選択情報に適合する属性を持つグループの符号化データを含む1つまたは複数のオーディオストリームのパケットがPIDフィルタで選択的に取り出される。

【0112】

デマルチプレクサ202で取り出されたオーディオストリームは、多重化バッファ211-1~211-Nのうち対応する多重化バッファに取り込まれる。コンバイナ212では、オーディオストリームが取り込まれた各多重化バッファから、オーディオフレーム毎に

50

、オーディオストリームが読み出され、ストリーム再構成部 231 に供給される。

【0113】

ストリーム再構成部 231 では、スピーカ構成、視聴者選択情報に適合する属性を持つ所定グループの符号化データが選択的に取得され、この所定グループの符号化データを持つオーディオストリームが再構成される。この再構成オーディオストリームは配信インタフェース 232 に供給される。そして、この配信インタフェース 232 から、構内ネットワーク接続されたデバイス 300 に配信（送信）される。

【0114】

この構内ネットワーク接続には、イーサネット接続、「WiFi」あるいは「Bluetooth」などのワイヤレス接続が含まれる。なお、「WiFi」、「Bluetooth」は、登録商標である。 10

【0115】

また、デバイス 300 には、サラウンドスピーカ、セカンドディスプレイ、ネットワーク端末に付属のオーディオ出力装置が含まれる。再構成オーディオストリームの配信を受けるデバイス 300 は、図 14 のサービス受信機 200 における 3D オーディオデコード 213 と同様のデコード処理を行って、所定数のスピーカを駆動するためのオーディオデータを得ることになる。

【0116】

また、サービス受信機としては、上述した再構成オーディオストリームを、「HDMI (High-Definition Multimedia Interface)」、「MHL (Mobile High definition 20 on Link)」、「DisplayPort」などのデジタルインタフェースで接続されたデバイスに送信する構成も考えられる。なお、「HDMI」、「MHL」は、登録商標である。

【0117】

また、上述実施の形態において、コンテナのレイヤに挿入されるストリーム対応関係情報は、グループ ID とサブストリーム ID との対応関係を示す情報であった。すなわち、グループとオーディオストリーム（サブストリーム）を関連付けるためにサブストリーム ID を用いたものである。しかし、グループとオーディオストリーム（サブストリーム）を関連付けるためにパケット識別子（PID: Packet ID）あるいはストリームタイプ（stream\_type）を用いることも考えられる。なお、ストリームタイプを用いる場合には、各オーディオストリーム（サブストリーム）のストリームタイプを変える必要がある。 30

【0118】

また、上述実施の形態においては、各グループの符号化データの属性情報を「attribute\_of\_groupID」のフィールドを設けて送信する例を示した（図 10 参照）。しかし、本技術は、送受信機間でグループ ID（GroupID）の値自体に特別な意味を定義することで、特定のグループ ID を認識すれば符号化データの種類（属性）が認識できるような方法も含むものである。この場合、グループ ID は、グループの識別子として機能する他に、そのグループの符号化データの属性情報として機能することとなり、「attribute\_of\_groupID」のフィールドは不要となる。 40

【0119】

また、上述実施の形態においては、複数のグループの符号化データに、チャンネル符号化データおよびオブジェクト符号化データの双方が含まれる例を示した（図 3 参照）。しかし、本技術は、複数のグループの符号化データに、チャンネル符号化データのみ、あるいはオブジェクト符号化データのみが含まれる場合にも同様に適用できる。

【0120】

また、上述実施の形態においては、コンテナがトランスポートストリーム（MPEG-2 TS）である例を示した。しかし、本技術は、MP4 やそれ以外のフォーマットのコンテナで配信されるシステムにも同様に適用できる。例えば、MPEG-DASH ベースのストリーム配信システム、あるいは、MMT (MPEG Media Transport) 構造伝送 50

ストリームを扱う送受信システムなどである。

【 0 1 2 1 】

なお、本技術は、以下のような構成もとることができる。

( 1 ) 複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを送信する送信部と、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報を挿入する情報挿入部とを備える

送信装置。

( 2 ) 上記情報挿入部は、

上記コンテナのレイヤに、上記複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すストリーム対応関係情報をさらに挿入する

前記( 1 )に記載の送信装置。

( 3 ) 上記ストリーム対応関係情報は、

上記複数のグループの符号化データのそれぞれを識別するグループ識別子と上記所定数のオーディオストリームのそれぞれを識別するストリーム識別子との対応関係を示す情報である

前記( 2 )に記載の送信装置。

( 4 ) 上記情報挿入部は、

上記コンテナのレイヤに、上記所定数のオーディオストリームのそれぞれのストリーム識別子を示すストリーム識別子情報をさらに挿入する

前記( 3 )に記載の送信装置。

( 5 ) 上記コンテナは、MPEG2-TSであり、

上記情報挿入部は、

上記ストリーム識別子情報を、プログラムマップテーブルの配下に存在する上記所定数のオーディオストリームのそれぞれに対応したオーディオエレメンタリストリームループ内に挿入する

前記( 4 )に記載の送信装置。

( 6 ) 上記ストリーム対応関係情報は、

上記複数のグループの符号化データのそれぞれを識別するグループ識別子と上記所定数のオーディオストリームのそれぞれをパケット化する際に付されるパケット識別子との対応関係を示す情報である

前記( 2 )に記載の送信装置。

( 7 ) 上記ストリーム対応関係情報は、

上記複数のグループの符号化データのそれぞれを識別するグループ識別子と上記所定数のオーディオストリームのそれぞれのストリームタイプを示すタイプ情報との対応関係を示す情報である

前記( 2 )に記載の送信装置。

( 8 ) 上記コンテナは、MPEG2-TSであり、

上記情報挿入部は、

上記属性情報および上記ストリーム対応関係情報を、プログラムマップテーブルの配下に存在する上記所定数のオーディオストリームのいずれか1つのオーディオストリームに対応したオーディオエレメンタリストリームループ内に挿入する

前記( 2 )から( 7 )のいずれかに記載の送信装置。

( 9 ) 上記複数のグループの符号化データには、チャンネル符号化データおよびオブジェクト符号化データのいずれかまたは双方が含まれる

前記( 1 )から( 8 )のいずれかに記載の送信装置。

( 10 ) 送信部より、複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを送信する送信ステップと、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報を挿入する情報挿入ステップとを有する

送信方法。

(11) 複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを受信する受信部を備え、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されており、

上記受信されたコンテナが有する上記所定数のオーディオストリームを、上記属性情報に基づいて処理する処理部をさらに備える

受信装置。

(12) 上記コンテナのレイヤには、上記複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すストリーム対応関係情報がさらに挿入されており、

上記処理部は、

上記属性情報の他に、上記ストリーム対応関係情報に基づいて、上記所定数のオーディオストリームを処理する

前記(11)に記載の受信装置。

(13) 上記処理部は、

上記属性情報および上記ストリーム対応関係情報に基づいて、スピーカ構成およびユーザ選択情報に適合する属性を持つグループの符号化データを含むオーディオストリームに対して選択的にデコード処理を施す

前記(12)に記載の受信装置。

(14) 上記複数のグループの符号化データには、チャンネル符号化データおよびオブジェクト符号化データのいずれかまたは双方が含まれる

前記(11)から(13)のいずれかに記載の受信装置。

(15) 受信部により、複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを受信する受信ステップを有し、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されており、

上記受信されたコンテナが有する上記所定数のオーディオストリームを、上記属性情報に基づいて処理する処理ステップをさらに有する

受信方法。

(16) 複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを受信する受信部を備え、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す属性情報が挿入されており、

上記受信されたコンテナが有する上記所定数のオーディオストリームから上記属性情報に基づいて所定のグループの符号化データを選択的に取得し、該所定のグループの符号化データを含むオーディオストリームを再構成する処理部と、

上記処理部で再構成されたオーディオストリームを外部機器に送信するストリーム送信部とをさらに備える

受信装置。

(17) 上記コンテナのレイヤには、上記複数のグループの符号化データがそれぞれのオーディオストリームに含まれるかを示すストリーム対応関係情報がさらに挿入されており、

上記処理部は、

上記属性情報の他に、上記ストリーム対応関係情報に基づいて、上記所定数のオーディオストリームから上記所定のグループの符号化データを選択的に取得する

前記(16)に記載の受信装置。

(18) 受信部により、複数のグループの符号化データを含む所定数のオーディオストリームを有する所定フォーマットのコンテナを受信する受信ステップを有し、

上記コンテナのレイヤに、上記複数のグループの符号化データのそれぞれの属性を示す

10

20

30

40

50

属性情報が挿入されており、

上記受信されたコンテナが有する上記所定数のオーディオストリームから上記属性情報に基づいて所定のグループの符号化データを選択的に取得し、該所定のグループの符号化データを含むオーディオストリームを再構成する処理ステップと、

上記処理ステップで再構成されたオーディオストリームを外部機器に送信するストリーム送信ステップとをさらに有する

受信方法。

【 0 1 2 2 】

本技術の主な特徴は、コンテナのレイヤに、所定数のオーディオストリームに含まれる複数のグループの符号化データのそれぞれの属性を示す属性情報や複数のグループの符号化データがそれぞれどのオーディオストリームに含まれるかを示すストリーム対応関係情報を挿入することで、受信側の処理負荷を軽減可能としたことである（図 1 3 参照）。

10

【 符号の説明 】

【 0 1 2 3 】

1 0 . . . 送受信システム

1 0 0 . . . サービス送信機

1 1 0 . . . ストリーム生成部

1 1 2 . . . ビデオエンコーダ

1 1 3 . . . オーディオエンコーダ

1 1 4 . . . マルチプレクサ

20

2 0 0 , 2 0 0 A . . . サービス受信機

2 0 1 . . . 受信部

2 0 2 . . . デマルチプレクサ

2 0 3 . . . ビデオデコーダ

2 0 4 . . . 映像処理回路

2 0 5 . . . パネル駆動回路

2 0 6 . . . 表示パネル

2 1 1 -1 ~ 2 1 1 -N . . . 多重化バッファ

2 1 2 . . . コンバイナ

2 1 3 . . . 3 D オーディオデコーダ

30

2 1 4 . . . 音声出力処理回路

2 1 5 . . . スピーカシステム

2 2 1 . . . C P U

2 2 2 . . . フラッシュ R O M

2 2 3 . . . D R A M

2 2 4 . . . 内部バス

2 2 5 . . . リモコン受信部

2 2 6 . . . リモコン送信機

2 3 1 . . . ストリーム再構成部

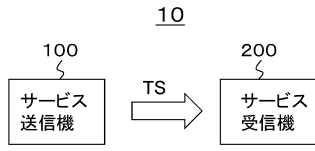
2 3 2 . . . 配信インタフェース

40

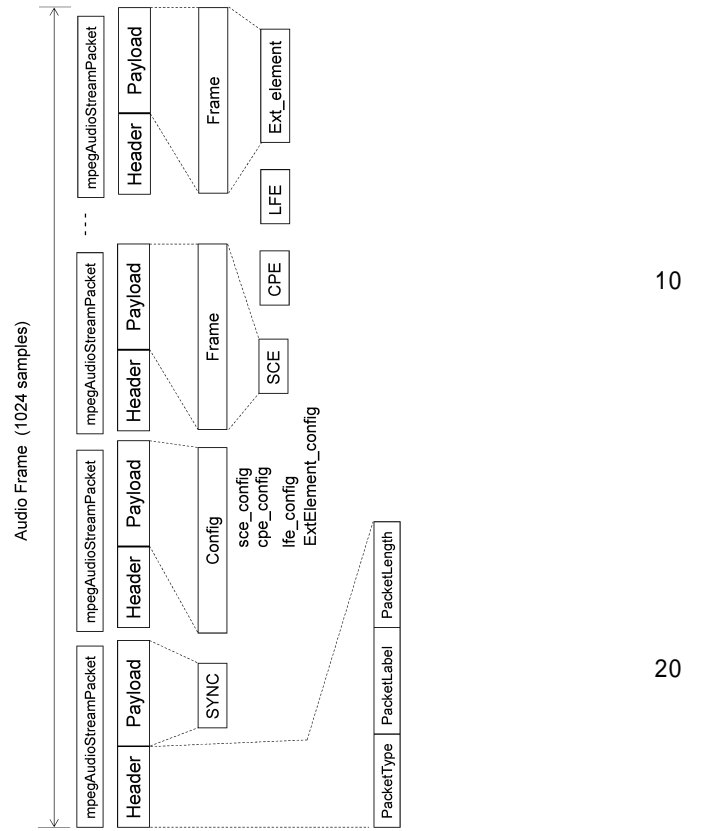
3 0 0 . . . デバイス

【 図 面 】

【 図 1 】



【 図 2 】



10

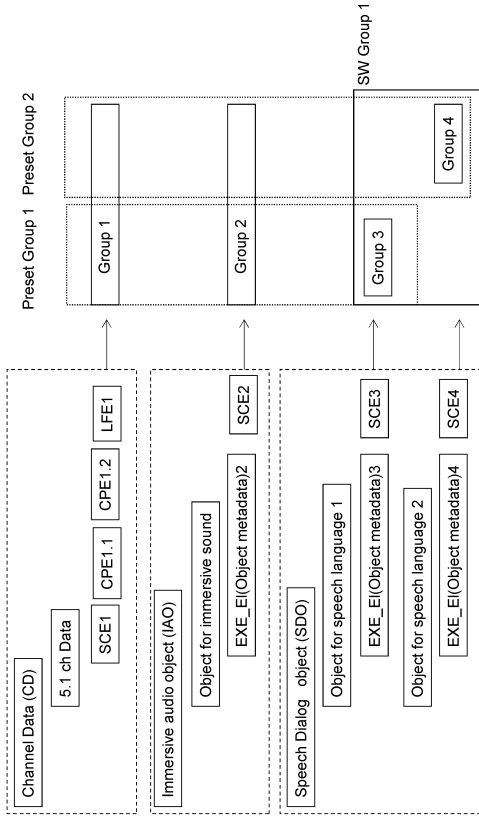
20

30

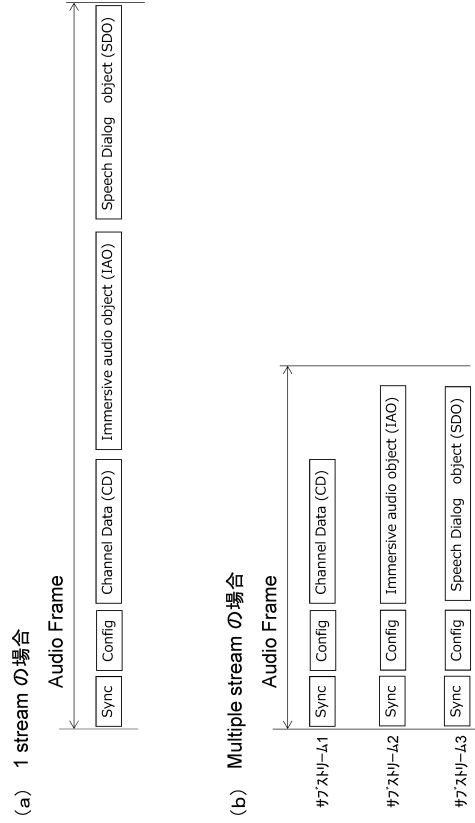
40

50

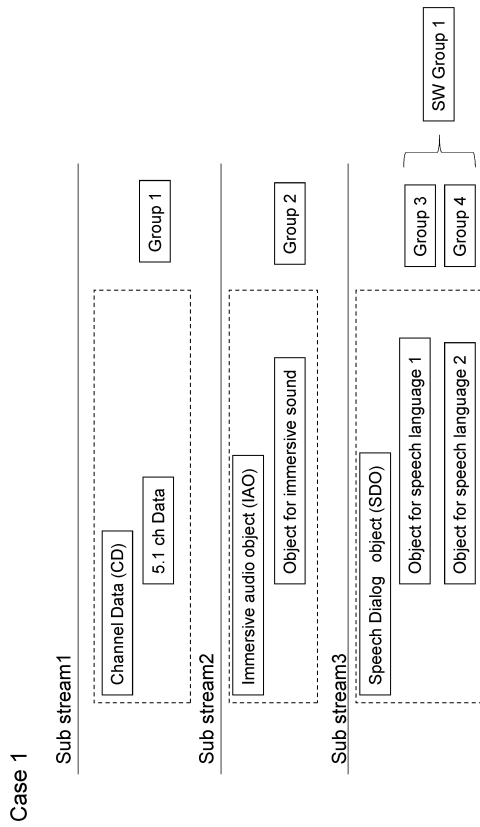
【 図 3 】



【 図 4 】



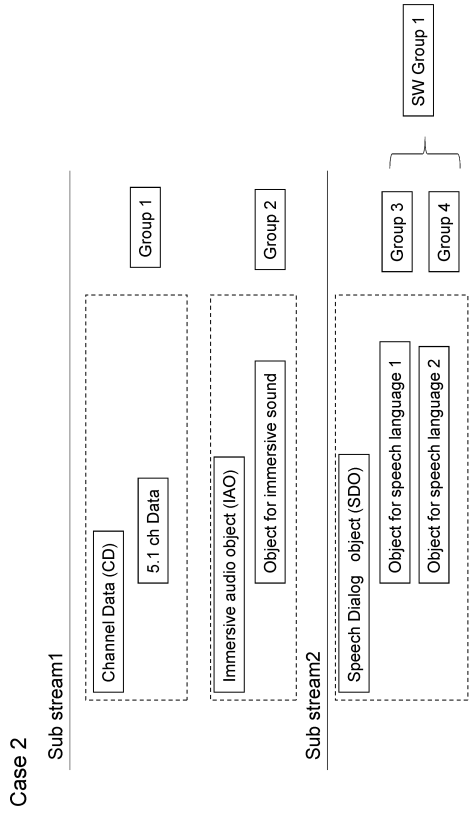
【 図 5 】



【 図 6 】

groupID	attribute	switchGroupID	presetGroupID [1]	subStreamID [2]
1	Channel data	0	1	1
2	Object sound	0	2	2
3	Object language1	1	3	3
4	Object language2	1	4	3

【 図 7 】



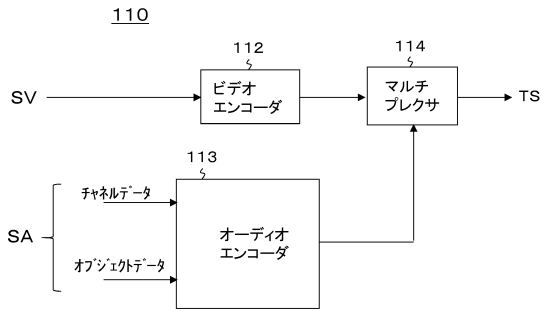
【 図 8 】

groupID	attribute	switchGroupID	presetGroupID [1]	subStreamID
1	Channel data	0	1	1
2	Object sound	0	2	1
3	Object language1	1	3	2
4	Object language2	1	4	2

10

20

【 図 9 】



【 図 10 】

3Daudio\_stream\_config\_descriptor

Syntax	No. of Bits	Format
3Daudio_stream_config_descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
NumOfGroups, N	8	uimsbf
NumOfPresetGroups, P	8	uimsbf
for (i = 0 ; i < N ; i++){		
groupID	8	uimsbf
attribute_of_groupID	8	uimsbf
SwitchGroupID	8	uimsbf
audio_substreamID	8	uimsbf
}		
for (j = 0 ; j < P ; j++){		
presetGroupID	8	uimsbf
NumOfGroups_in_preset, R	8	uimsbf
for (k = 0 ; k < R ; k++){		
groupID	8	uimsbf
}		
}		
}		

30

40

50

【 図 1 1 】

Semantics of 3Daudio\_stream\_config\_descriptor

NumOfGroups (8bits)  
Groupの数を示す。(1以上の値)

NumOfPresetGroups (8bits)  
PresetGroupの数を示す。(1以上の値)

groupID (8bits)  
groupの識別子を示す。(0以外の値)

attribute\_of\_groupID (8bits)  
該当groupの符号化データの属性を示す。(0以外の値)

switchGroupID (8bits)  
該当groupがどのswitchGroupに属するかを示す識別子。いずれのswitchGroupにも属さない。

audio\_substreamID (8bits)  
'0' 以外 配属するswitchGroupを示す。

presetsGroupID (8bits)  
オーディオ・サブストリームの識別子。

NumOfGroups\_in\_preset (8bits)  
Groupをプリセットした束を示す識別子。

NumOfGroups\_in\_preset (8bits)  
presetsGroupに属するGroupの数を示す。

【 図 1 2 】

3Daudio\_substreamID\_descriptor

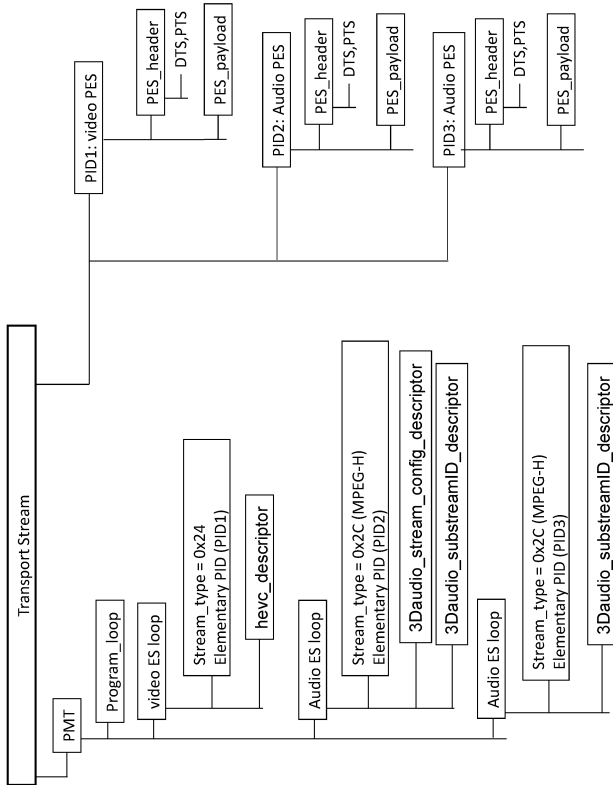
Syntax	No. of Bits	Format
3Daudio_substreamID_descriptor{		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
audio_substreamID	8	uimsbf
}		

(a)

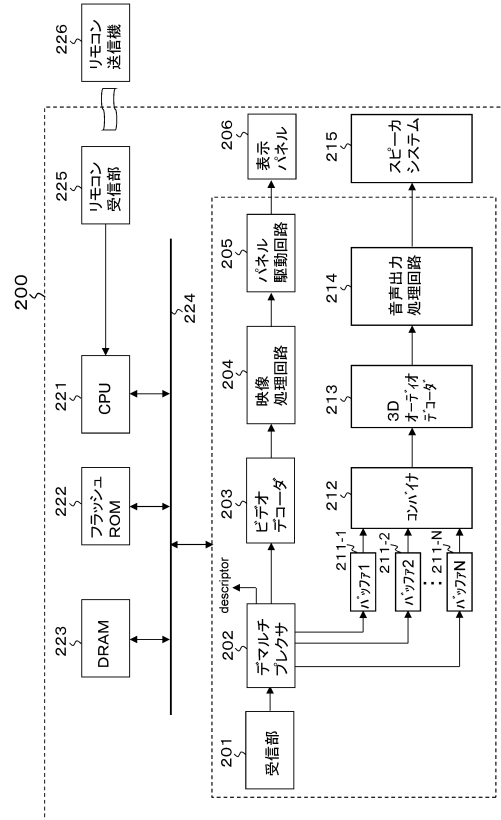
audio\_substreamID (8bits)  
オーディオ・サブストリームの識別子。

(b)

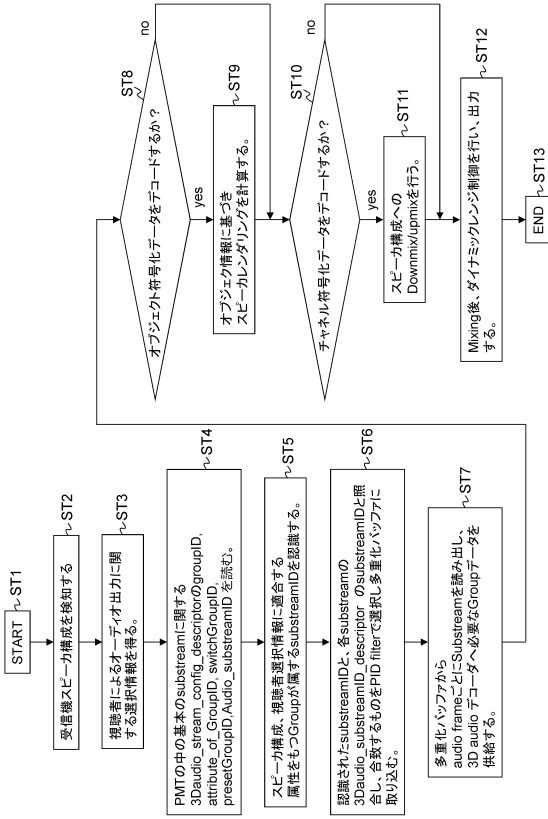
【 図 1 3 】



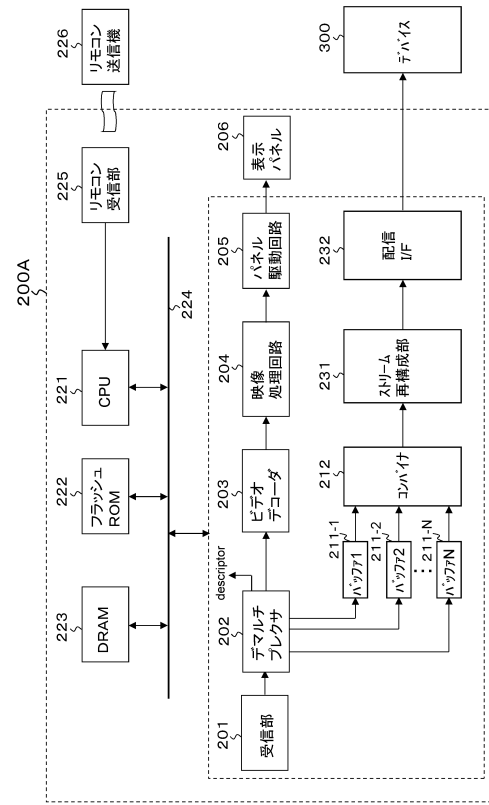
【 図 1 4 】



【 図 15 】



【 図 16 】



10

20

30

40

50

フロントページの続き

(51)国際特許分類

F I  
H 0 4 H 20/95

グループ株式会社内