(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2015 年 1 月 22 日(22.01.2015)

WIPO | PCT

(10) 国際公開番号
WO 2015/008613 A1

(51) 国際特許分類:
H04N 21/84 (2011.01) *H04N 21/439* (2011.01)
H04N 5/765 (2006.01) *H04N 21/4728* (2011.01)
H04N 5/91 (2006.01) *H04N 21/845* (2011.01)

(21) 国際出願番号: PCT/JP2014/067508

(22) 国際出願日: 2014 年 7 月 1 日(01.07.2014)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2013-151020 2013 年 7 月 19 日(19.07.2013) JP
PCT/JP2014/063752 2014 年 5 月 23 日(23.05.2014) JP
PCT/JP2014/066321 2014 年 6 月 19 日(19.06.2014) JP

(71) 出願人: ソニー株式会社(SONY CORPORATION) [JP/JP]; 〒1080075 東京都港区港南１丁目７番１号 Tokyo (JP).

(72) 発明者: 服部 しのぶ(HATTORI Shinobu); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 平林 光浩(HIRABAYASHI Mitsuhiro); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 中神 央二(NAKAGAMI Ohji); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 知念 徹(CHINEN Toru); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 史 潤宇(SHI Runyu); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 辻 実(TSUJI Minoru); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 山本 優樹(YAMAMOTO Yuki); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP).

(74) 代理人: 西川 孝, 外(NISHIKAWA Takashi et al.); 〒1600023 東京都新宿区西新宿７丁目５番２５号 西新宿木村屋ビルディング９階 Tokyo (JP).

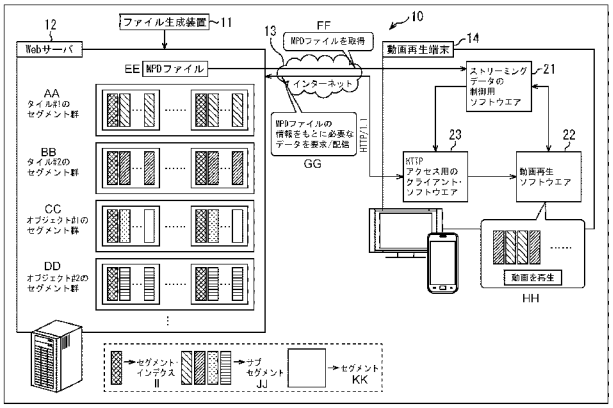
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN,

[続葉有]

(54) **Title**: INFORMATION PROCESSING DEVICE AND INFORMATION PROCESSING METHOD

(54) 発明の名称 : 情報処理装置および情報処理方法

[図1]



11... FILE GENERATION DEVICE
12... WEB SERVER
13... INTERNET
14... VIDEO REPLAY TERMINAL
21... CONTROL SOFTWARE FOR STREAMING DATA
22... VIDEO REPLAY SOFTWARE
23... CLIENT SOFTWARE FOR HTTP ACCESS
AA... TILE # 1 SEGMENT GROUP
BB... TILE #2 SEGMENT GROUP
CC... OBJECT #1 SEGMENT GROUP
DD... OBJECT #2 SEGMENT GROUP
EE... MPD FILE
FF... GET MPD FILE
GG... REQUEST/DISTRIBUTE REQUIRED DATA ON BASIS OF MPD FILE INFORMATION
HH... REPLAY VIDEO
II... SEGMENT INDEX
JJ... SUB-SEGMENT
KK... SEGMENT

(57) **Abstract**: The present invention pertains to an information processing device and an information processing method whereby the acquisition position of audio data in an image can be recognized. A web server transmits image frame size information indicating the image frame size of image data and audio position information indicating the acquisition position of audio data. The present invention can be applied to information processing systems, etc., including file generation devices, web servers, and video replay terminals, that perform tiled streaming using a method conforming to MPEG-DASH (Moving Picture Experts Group phase - Dynamic Adaptive Streaming over HTTP).

(57) 要約: 本開示は、画像上の音声データの取得位置を認識することができるようにする情報処理装置および情報処理方法に関する。 Web サーバは、画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを伝送する。本開示は、例えば、MPEG-DASH （Moving Picture Experts Group phase - Dynamic Adaptive Streaming over HTTP）に準ずる方式でタイルドストリーミングを行う、ファイル生成装置、Ｗｅｂサーバ、および動画再生端末を含む情報処理システム等に適用することができる。

WO 2015/008613 A1

IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告（条約第 21 条(3)）

明　細　書

発明の名称 ： 情報処理装置および情報処理方法

技術分野

[0001] 本開示は、情報処理装置および情報処理方法に関し、特に、画像上の音声データの取得位置を認識することができるようにした情報処理装置および情報処理方法に関する。

背景技術

[0002] 近年、インターネット上のストリーミングサービスの主流がOTT-V（Over The Top Video）となっている。この基盤技術として普及し始めているのがMPEG-DASH（Moving Picture Experts Group phase - Dynamic Adaptive Streaming over HTTP）である（例えば、非特許文献１参照）。

[0003] MPEG-DASHでは、配信サーバが１本の動画コンテンツ用に画面サイズと符号化速度が異なる動画データ群を用意し、再生端末が伝送路の状況に応じて最適な画面サイズと符号化速度の動画データ群を要求することにより、適応型のストリーミング配信が実現される。

先行技術文献

非特許文献

[0004] 非特許文献1：MPEG-DASH(Dynamic Adaptive Streaming over HTTP) (URL:http://mpeg.chiariglione.org/standards/mpeg-dash/media-presentation-description-and-segment-formats/text-isoiec-23009-12012-dam-1)

発明の概要

発明が解決しようとする課題

[0005] しかしながら、再生端末が、動画コンテンツ上の音声データの取得位置を認識することは考えられていない。

[0006] 本開示は、このような状況に鑑みてなされたものであり、画像上の音声データの取得位置を認識することができるようにするものである。

課題を解決するための手段

[0007] 本開示の第１の側面の情報処理装置は、画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを伝送する伝送部を備える情報処理装置である。

[0008] 本開示の第１の側面の情報処理方法は、本開示の第１の側面の情報処理装置に対応する。

[0009] 本開示の第１の側面においては、画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とが伝送される。

[0010] 本開示の第２の側面の情報処理装置は、画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを受け取る受け取り部と、前記受け取り部により受け取られた前記画像データの前記画枠サイズ情報と前記音声位置情報に基づいて、前記画像データに対応する画像上の前記音声データの取得位置を決定する位置決定部とを備える情報処理装置である。

[0011] 本開示の第２の側面の情報処理方法は、本開示の第２の側面の情報処理装置に対応する。

[0012] 本開示の第２の側面においては、画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とが受け取られ、受け取られた前記画像データの前記画枠サイズ情報と前記音声位置情報に基づいて、前記画像データに対応する画像上の前記音声データの取得位置が決定される。

[0013] なお、第１および第２の側面の情報処理装置は、コンピュータにプログラムを実行させることにより実現することができる。

[0014] また、第１および第２の側面の情報処理装置を実現するために、コンピュータに実行させるプログラムは、伝送媒体を介して伝送することにより、又は、記録媒体に記録して、提供することができる。

発明の効果

[0015] 本開示の第１の側面によれば、画像上の音声データの取得位置を認識させるための情報を伝送することができる。

[0016] また、本開示の第２の側面によれば、画像上の音声データの取得位置を認識することができる。

図面の簡単な説明

[0017] [図1]本開示を適用した情報処理システムの第１実施の形態の概要を説明する図である。

[図2]タイルの例を示す図である。

[図3]オブジェクトを説明する図である。

[図4]オブジェクト位置情報を説明する図である。

[図5]画枠サイズ情報を説明する図である。

[図6]MPDファイルの構造を示す図である。

[図7]「Period」、「Representation」、および「Segment」の関係を示す図である。

[図8]MPDファイルの階層構造を示す図である。

[図9]MPDファイルの構造と時間軸との関係を示す図である。

[図10]MPDファイルの記述例を説明する図である。

[図11]ファイル生成装置の構成例を示すブロック図である。

[図12]ファイル生成装置のファイル生成処理を説明するフローチャートである。

[図13]ストリーミング再生部の構成例を示すブロック図である。

[図14]ストリーミング再生部のストリーミング再生処理を説明するフローチャートである。

[図15]MPDファイルの他の記述例を説明する図である。

[図16]MPDファイルのさらに他の記述例を説明する図である。

[図17]オーディオストリームの配置例を示す図である。

[図18]gsixの記述例を示す図である。

[図19]Sample group entryとオブジェクトＩＤの対応関係を示す情報の例を示す図である。

[図20]AudioObjectSampleGroupEntryの記述例を示す図である。

[図21]type assignment boxの記述例を示す図である。

[図22]本開示を適用した情報処理システムの第２実施の形態の概要を説明する図である。

[図23]本開示を適用した情報処理システムのストリーミング再生部の構成例を示すブロック図である。

[図24]オブジェクトの位置の決定方法を説明する図である。

[図25]オブジェクトの位置の決定方法を説明する図である。

[図26]オブジェクトの位置の決定方法を説明する図である。

[図27]水平角度θ_{Ai}と水平角度θ_{Ai}'の関係を表す図である。

[図28]垂直角度γ_{v1}'と垂直角度γ_{Ai}'の関係を表す図である。

[図29]図２３のストリーミング再生部のストリーミング再生処理を説明するフローチャートである。

[図30]図２９の位置決定処理の詳細を説明するフローチャートである。

[図31]図３０の水平角度θ_{Ai}'推定処理の詳細を説明するフローチャートである。

[図32]本開示を適用した情報処理システムの第３実施の形態におけるオブジェクトの位置の決定方法を説明する図である。

[図33]本開示を適用した情報処理システムの第３実施の形態におけるオブジェクトの位置の決定方法を説明する図である。

[図34]本開示を適用した情報処理システムの第３実施の形態におけるオブジェクトの位置の決定方法を説明する図である。

[図35]本開示を適用した情報処理システムの第３実施の形態におけるオブジェクトの位置の決定方法を説明する図である。

[図36]本開示を適用した情報処理システムの第３実施の形態における位置決定処理の詳細を説明するフローチャートである。

[図37]コンピュータのハードウエアの構成例を示すブロック図である。

発明を実施するための形態

[0018] ＜第１実施の形態＞

（情報処理システムの第１実施の形態の概要）

図１は、本開示を適用した情報処理システムの第１実施の形態の概要を説明する図である。

[0019] 図１の情報処理システム１０は、ファイル生成装置１１と接続されるWebサーバ１２と動画再生端末１４が、インターネット１３を介して接続されることにより構成される。

[0020] 情報処理システム１０では、MPEG-DASHに準ずる方式で、Webサーバ１２が動画コンテンツの画像データをタイル単位で動画再生端末１４に配信（タイルドストリーミング）する。

[0021] 具体的には、ファイル生成装置１１は、動画コンテンツの画像データを取得し、タイル単位で符号化してビデオストリームを生成する。ファイル生成装置１１は、各タイルのビデオストリームを、セグメントと呼ばれる数秒から10秒程度の時間単位ごとにファイル化する。ファイル生成装置１１は、その結果得られる各タイルの画像ファイルをWebサーバ１２にアップロードする。

[0022] また、ファイル生成装置１１は、動画コンテンツの音声データをオブジェクト（詳細は後述する）ごとに取得し、オブジェクト単位で符号化してオーディオストリームを生成する。ファイル生成装置１１は、各オブジェクトのオーディオストリームをセグメント単位でファイル化し、その結果得られる各オブジェクトの音声ファイルをWebサーバ１２にアップロードする。

[0023] なお、オブジェクトとは、音源であり、各オブジェクトの音声データは、そのオブジェクトに取り付けられたマイクロフォンなどにより取得される。オブジェクトは、固定されたマイクスタンドなどの物体であってもよいし、人物などの動体であってもよい。

[0024] ファイル生成装置１１は、各オブジェクトの位置（音声データの取得位置）を表すオブジェクト位置情報（音声位置情報）、オブジェクトに固有のＩＤであるオブジェクトＩＤ等を含むオーディオメタデータを符号化する。ファイル生成装置１１は、オーディオメタデータの符号化データをセグメント

単位でファイル化し、その結果得られる音声メタファイルをWebサーバ１２にアップロードする。

[0025] また、ファイル生成装置１１は、動画コンテンツの画像の画枠サイズを表す画枠サイズ情報、各タイルの画像上の位置を表すタイル位置情報等を含む、画像ファイルや音声ファイルを管理するMPD（Media Presentation Description）ファイル（制御情報）を生成する。ファイル生成装置１１は、MPDファイルをWebサーバ１２にアップロードする。

[0026] Webサーバ１２は、ファイル生成装置１１からアップロードされた画像ファイル、音声ファイル、音声メタファイル、およびMPDファイルを格納する。

[0027] 図１の例では、Webサーバ１２には、タイルＩＤ「１」のタイルの複数のセグメントの画像ファイルからなるセグメント群と、タイルＩＤ「２」のタイルの複数のセグメントの画像ファイルからなるセグメント群とが格納されている。また、Webサーバ１２には、オブジェクトＩＤ「１」のオブジェクトの複数のセグメントの音声ファイルからなるセグメント群と、オブジェクトＩＤ「２」のオブジェクトの複数のセグメントの音声ファイルからなるセグメント群とが格納されている。図示は省略するが、音声メタファイルのセグメント群も同様に格納されている。

[0028] なお、以下では、タイルＩＤがｉであるタイルをタイル＃ｉといい、オブジェクトＩＤがｉのオブジェクトをオブジェクト＃ｉという。

[0029] Webサーバ１２は、伝送部として機能し、動画再生端末１４からの要求に応じて、格納している画像ファイル、音声ファイル、音声メタファイル、MPDファイル等を動画再生端末１４に送信する。

[0030] 動画再生端末１４は、ストリーミングデータの制御用ソフトウエア（以下、制御用ソフトウエアという）２１、動画再生ソフトウエア２２、HTTP（HyperText Transfer Protocol）アクセス用のクライアント・ソフトウエア(以下、アクセス用ソフトウエアという)２３などを実行する。

[0031] 制御用ソフトウエア２１は、Webサーバ１２からストリーミングするデータを制御するソフトウエアである。具体的には、制御用ソフトウエア２１は、

動画再生端末１４にWebサーバ１２からMPDファイルを取得させる。

[0032] また、制御用ソフトウエア２１は、動画再生ソフトウエア２２から指示される動画コンテンツの画像内の表示する領域である表示領域と、MPDファイルに含まれるタイル位置情報とに基づいて、表示領域内のタイルを特定する。そして、制御用ソフトウエア２１は、そのタイルの画像ファイルの送信要求をアクセス用ソフトウエア２３に指令する。

[0033] また、制御用ソフトウエア２１は、アクセス用ソフトウエア２３に音声メタファイルの送信要求を指令する。そして、制御用ソフトウエア２１は、表示領域、MPDファイルに含まれる画枠サイズ情報、および音声メタファイルに含まれるオブジェクト位置情報に基づいて、表示領域内の画像に対応するオブジェクトを特定する。そして、制御用ソフトウエア２１は、そのオブジェクトの音声ファイルの送信要求をアクセス用ソフトウエア２３に指令する。

[0034] 動画再生ソフトウエア２２は、Webサーバ１２から取得された画像ファイルと音声ファイルを再生するソフトウエアである。具体的には、動画再生ソフトウエア２２は、ユーザにより表示領域が指定されると、その表示領域を制御用ソフトウエア２１に指示する。また、動画再生ソフトウエア２２は、その指示に応じてWebサーバ１２から取得された画像ファイルと音声ファイルを復号し、合成して出力する。

[0035] アクセス用ソフトウエア２３は、HTTPを用いたインターネット１３を介したWebサーバ１２との通信を制御するソフトウエアである。具体的には、アクセス用ソフトウエア２３は、制御用ソフトウエア２１の指令に応じて、画像ファイル、音声ファイル、および音声メタファイルの送信要求を動画再生端末１４に送信させる。また、アクセス用ソフトウエア２３は、その送信要求に応じて、Webサーバ１２から送信されてくる画像ファイル、音声ファイル、および音声メタファイルを動画再生端末１４に受信させる。

[0036] （タイルの例）

図２は、タイルの例を示す図である。

[0037] 図２に示すように、動画コンテンツの画像は、複数のタイルに分割され、

各タイルに１から順にタイルＩＤが付与される。図２の例では、動画コンテンツの画像は、４つのタイル＃１乃至タイル＃４に分割されている。

[0038] （オブジェクトの説明）

図３は、オブジェクトを説明する図である。

[0039] 図３の例では、動画コンテンツの音声として、画像内の８つのオブジェクトの音声が取得されており、各オブジェクトには１から順にオブジェクトＩＤが付与されている。オブジェクト＃１乃至オブジェクト＃５は、それぞれ、動体であり、オブジェクト＃６乃至オブジェクト＃８は、固定された物体である。また、図３の例では、動画コンテンツの画像は、５（縦）×７（横）のタイルに分割されている。

[0040] この場合、図３に示すように、ユーザにより２（縦）×３（横）個のタイルからなる表示領域３１が指定されると、表示領域３１には、オブジェクト＃１、オブジェクト＃２、およびオブジェクト＃６のみが含まれる。従って、動画再生端末１４は、例えば、オブジェクト＃１、オブジェクト＃２、およびオブジェクト＃６の音声ファイルのみをWebサーバ１２から取得し、再生する。

[0041] 表示領域３１内のオブジェクトは、以下に説明するように、画枠サイズ情報およびオブジェクト位置情報に基づいて特定することができる。

[0042] （オブジェクト位置情報の説明）

図４は、オブジェクト位置情報を説明する図である。

[0043] 図４に示すように、オブジェクト位置情報は、オブジェクト４０の水平角度θ_A（$-180° \leqq \theta_A \leqq 180°$），垂直角度$\gamma_A$（$-90° \leqq \gamma_A \leqq 90°$）、および距離$r_A$（$0 < r_A$）からなる。水平角度$\theta_A$は、例えば、画像の中心の撮影位置を原点（基点）Ｏとし、画像の水平方向をＸ方向、垂直方向をＹ方向、XY平面に垂直な奥行き方向をＺ方向としたときの、オブジェクト４０と原点Ｏを結ぶ直線とYZ平面との水平方向の角度である。垂直角度γ_Aは、オブジェクト４０と原点Ｏを結ぶ直線とXZ平面との垂直方向の角度であり、距離r_Aは、オブジェクト４０と原点Ｏの距離である。

[0044] なお、本明細書では、左回転および上回転の角度を正の角度とし、右回転および下回転の角度を負の角度とする。

[0045] （画枠サイズ情報の説明）

図５は、画枠サイズ情報を説明する図である。

[0046] 図５に示すように、画枠サイズ情報は、画枠の左端の水平角度θ_{v1}、右端の水平角度θ_{v2}、画枠の上端の垂直角度γ_{v1}、下端の垂直角度γ_{v2}、および距離r_vにより構成される。

[0047] 水平角度θ_{v1}は、例えば、画像の中心の撮影位置を原点Ｏとし、画像の水平方向をＸ方向、垂直方向をＹ方向、XY平面に垂直な奥行き方向をＺ方向としたときの、画枠の左端と原点Ｏを結ぶ直線とYZ平面との水平方向の角度である。水平角度θ_{v2}は、画枠の右端と原点Ｏを結ぶ直線とYZ平面との水平方向の角度である。従って、水平角度θ_{v1}と水平角度θ_{v2}とを合わせた角度は、水平方向の画角である。

[0048] 垂直角度γ_{V1}，γ_{v2}は、それぞれ、画枠の上端、下端と原点Ｏを結ぶ直線とXZ平面との角度であり、垂直角度γ_{V1}と垂直角度γ_{v2}とを合わせた角度は、垂直方向の画角である。距離r_vは、原点Ｏと画像の平面との距離である。

[0049] 以上のように、オブジェクト位置情報と画枠サイズ情報は、それぞれ、オブジェクト４０、画枠と、原点Ｏとの位置関係を表している。従って、オブジェクト位置情報と画枠サイズ情報に基づいて、各オブジェクトの画像上の位置を検出（認識）することができる。その結果、表示領域３１内のオブジェクトを特定することができる。

[0050] (MPDファイルの構造の説明)

図６は、MPDファイルの構造を示す図である。

[0051] MPDファイルの解析（パース）においては、動画再生端末１４は、MPDファイル（図６のMedia Presentation）の「Period」に含まれる「Representation」の属性から最適なものを選択する。

[0052] 動画再生端末１４は、選択した「Representation」の先頭の「Initialization Segment」のURL（Uniform Resource Locator）等を参照してファイルを

取得し、処理する。続いて、動画再生端末１４は、後続の「Media Segment」のURL等を参照してファイルを取得し、再生する。

[0053] なお、MPDファイルにおける、「Period」、「Representation」、および「Segment」の関係は、図７のようになる。つまり、１つの動画コンテンツは、「Period」により、セグメントより長い時間単位で管理することができ、各「Period」において、「Segment」によりセグメント単位で管理することができる。また、各「Period」において、「Representation」により、動画コンテンツをストリームの属性単位で管理することができる。

[0054] したがって、MPDファイルは、「Period」以下において、図８に示す階層構造を有する。また、このMPDファイルの構造を時間軸上に並べると図９の例のようになる。図９から明らかなように、同一のセグメントに対して複数の「Representation」が存在している。動画再生端末１４は、これらのうちのいずれかを適応的に選択することにより、ユーザにより選択された表示領域の画像ファイルと音声ファイルを取得し、再生することができる。

[0055] （MPDファイルの記述の説明）

図１０は、MPDファイルの記述を説明する図である。

[0056] 上述したように、情報処理システム１０では、動画再生端末１４において表示領域内のオブジェクトを特定可能にするため、MPDファイルに画枠サイズ情報が含まれる。この画枠サイズ情報は、図１０に示すように、Viewpoint DescriptorType elementを活用して新たに画枠サイズ情報（Viewing Angle）を定義するためのScheme(urn:mpeg:DASH:viewingAngle:2013)を拡張することにより、音声用の「Adaptation Set」と画像用の「Adaptation Set」に配置される。画枠サイズ情報は、画像用の「Adaptation Set」にのみ配置されるようにしてもよい。

[0057] また、MPDファイルの音声用の「Adaptation Set」には、音声メタファイル用の「Representation」が記述され、その「Representation」の「Segment」には、音声メタファイル（audionmetadata.mp4）を特定する情報としてのURL等が記述される。このとき、Role elementを活用して、「Segment」で指定さ

れるファイルが音声メタファイルであること（objectaudiometadata）が記述される。

[0058] MPDファイルの音声用の「Adaptation Set」にはまた、各オブジェクトの音声ファイル用の「Representation」が記述され、その「Representation」の「Segment」には、各オブジェクトの音声ファイル（audioObje1.mp4, audioObje5.mp4）を特定する情報としてのURL等が記述される。このとき、Viewpointを拡張することにより、音声ファイルに対応するオブジェクトのオブジェクトＩＤ（1,5）も記述される。

[0059] なお、図示は省略するが、タイル位置情報は、画像用の「Adaptation Set」に配置される。

[0060] （ファイル生成装置の構成例）

図１１は、図１のファイル生成装置１１の構成例を示すブロック図である。

[0061] 図１１のファイル生成装置１１は、画面分割処理部５１、画像符号化処理部５２、画像ファイル生成部５３、画像情報生成部５４、音声符号化処理部５５、音声ファイル生成部５６、MPD生成部５７、およびサーバアップロード処理部５８により構成される。

[0062] ファイル生成装置１１の画面分割処理部５１は、外部から入力される動画コンテンツの画像データをタイル単位に分割する。画面分割処理部５１は、タイル位置情報を画像情報生成部５４に供給する。また、画面分割処理部５１は、タイル単位の画像データを画像符号化処理部５２に供給する。

[0063] 画像符号化処理部５２は、画面分割処理部５１から供給されるタイル単位の画像データをタイルごとに符号化し、ビデオストリームを生成する。画像符号化処理部５２は、各タイルのビデオストリームを画像ファイル生成部５３に供給する。

[0064] 画像ファイル生成部５３は、画像符号化処理部５２から供給される各タイルのビデオストリームをセグメント単位でファイル化し、その結果得られる各タイルの画像ファイルをMPD生成部５７に供給する。

[0065] 画像情報生成部５４は、画面分割処理部５１から供給されるタイル位置情報と外部から入力される画枠サイズ情報とを画像情報としてMPD生成部５７に供給する。

[0066] 音声符号化処理部５５は、外部から入力される動画コンテンツのオブジェクト単位の音声データをオブジェクトごとに符号化し、オーディオストリームを生成する。また、音声符号化処理部５５は、外部から入力される各オブジェクトのオブジェクト位置情報、オブジェクトＩＤ等を含むオーディオメタデータを符号化し、符号化データを生成する。音声符号化処理部５５は、各オブジェクトのオーディオストリームとオーディオメタデータの符号化データとを音声ファイル生成部５６に供給する。

[0067] 音声ファイル生成部５６は、音声ファイル生成部として機能し、音声符号化処理部５５から供給される各オブジェクトのオーディオストリームをセグメント単位でファイル化し、その結果得られる各オブジェクトの音声ファイルをMPD生成部５７に供給する。

[0068] また、音声ファイル生成部５６は、メタファイル生成部として機能し、音声符号化処理部５５から供給されるオーディオメタデータの符号化データをセグメント単位でファイル化し、その結果得られる音声メタファイルをMPD生成部５７に供給する。

[0069] MPD生成部５７は、画像ファイル生成部５３から供給される各タイルの画像ファイルを格納するWebサーバ１２のURL等を決定する。また、MPD生成部５７は、音声ファイル生成部５６から供給される各タイルの音声ファイルと音声メタファイルを格納するWebサーバ１２のURL等を決定する。

[0070] MPD生成部５７は、画像情報生成部５４から供給される画像情報をMPDファイルの画像用の「AdaptationSet」に配置する。また、MPD生成部５７は、画像情報のうちの画枠サイズ情報をMPDファイルの音声用の「AdaptationSet」に配置する。MPD生成部５７は、各タイルの画像ファイルのURL等を、そのタイルの画像ファイル用の「Representation」の「Segment」に配置する。

[0071] MPD生成部５７は、各オブジェクトの画像ファイルのURL等を、そのオブジ

ェクトの音声ファイル用の「Representation」の「Segment」に配置する。また、MPD生成部５７は、情報生成部として機能し、音声メタファイルを特定する情報としてのURL等を音声メタファイル用の「Representation」の「Segment」に配置する。MPD生成部５７は、以上のようにして各種の情報が配置されたMPDファイル、並びに、画像ファイル、音声ファイル、および音声メタファイルをサーバアップロード処理部５８に供給する。

[0072] サーバアップロード処理部５８は、MPD生成部５７から供給される各タイルの画像ファイル、各オブジェクトの音声ファイル、音声メタファイル、およびMPDファイルを、Webサーバ１２にアップロードする。

[0073] （ファイル生成装置の処理の説明）

図１２は、図１１のファイル生成装置１１のファイル生成処理を説明するフローチャートである。

[0074] 図１２のステップＳ１１において、ファイル生成装置１１の画面分割処理部５１は、外部から入力される動画コンテンツの画像データをタイル単位に分割する。画面分割処理部５１は、タイル位置情報を画像情報生成部５４に供給する。また、画面分割処理部５１は、タイル単位の画像データを画像符号化処理部５２に供給する。

[0075] ステップＳ１２において、画像符号化処理部５２は、画面分割処理部５１から供給されるタイル単位の画像データをタイルごとに符号化し、各タイルのビデオストリームを生成する。画像符号化処理部５２は、各タイルのビデオストリームを画像ファイル生成部５３に供給する。

[0076] ステップＳ１３において、画像ファイル生成部５３は、画像符号化処理部５２から供給される各タイルのビデオストリームをセグメント単位でファイル化し、各タイルの画像ファイルを生成する。画像ファイル生成部５３は、各タイルの画像ファイルをMPD生成部５７に供給する。

[0077] ステップＳ１４において、画像情報生成部５４は、外部から画枠サイズ情報を取得する。ステップＳ１５において、画像情報生成部５４は、画面分割処理部５１から供給されるタイル位置情報と画枠サイズ情報とを含む画像情

報を生成し、MPD生成部５７に供給する。

[0078] ステップＳ１６において、音声符号化処理部５５は、外部から入力される動画コンテンツのオブジェクト単位の音声データをオブジェクトごとに符号化し、各オブジェクトのオーディオストリームを生成する。また、音声符号化処理部５５は、外部から入力される各オブジェクトのオブジェクト位置情報、オブジェクトＩＤ等を含むオーディオメタデータを符号化し、符号化データを生成する。音声符号化処理部５５は、各オブジェクトのオーディオストリームとオーディオメタデータの符号化データとを音声ファイル生成部５６に供給する。

[0079] ステップＳ１７において、音声ファイル生成部５６は、音声符号化処理部５５から供給される各オブジェクトのオーディオストリームをセグメント単位でファイル化し、各オブジェクトの音声ファイルを生成する。また、音声ファイル生成部５６は、音声符号化処理部５５から供給されるオーディオメタデータの符号化データをセグメント単位でファイル化し、音声メタファイルを生成する。音声ファイル生成部５６は、各オブジェクトの音声ファイルと音声メタファイルをMPD生成部５７に供給する。

[0080] ステップＳ１８において、MPD生成部５７は、画像情報生成部５４から供給される画像情報、各ファイルのURL等を含むMPDファイルを生成する。MPD生成部５７は、MPDファイル、各タイルの画像ファイル、各オブジェクトの音声ファイル、および音声メタファイルをサーバアップロード処理部５８に供給する。

[0081] ステップＳ１９において、サーバアップロード処理部５８は、MPD生成部５７から供給される各タイルの画像ファイル、各オブジェクトの音声ファイル、音声メタファイル、およびMPDファイルを、Webサーバ１２にアップロードする。そして、処理は終了する。

[0082] （動画再生端末の機能的構成例）

図１３は、図１の動画再生端末１４が制御用ソフトウエア２１、動画再生ソフトウエア２２、およびアクセス用ソフトウエア２３を実行することによ

り実現されるストリーミング再生部の構成例を示すブロック図である。

[0083] 図１３のストリーミング再生部９０は、MPD取得部９１、MPD処理部９２、メタファイル取得部９３、音声選択部９４、音声ファイル取得部９５、音声復号処理部９６、音声合成処理部９７、画像選択部９８、画像ファイル取得部９９、画像復号処理部１００、および画像合成処理部１０１により構成される。

[0084] ストリーミング再生部９０のMPD取得部９１は、受け取り部として機能し、Webサーバ１２からMPDファイルを取得し、MPD処理部９２に供給する。

[0085] MPD処理部９２は、MPD取得部９１から供給されるMPDファイルから音声メタファイル用の「Segment」に記述されるURL等の情報を抽出し、メタファイル取得部９３に供給する。また、MPD処理部９２は、MPDファイルから画像用の「AdaptationSet」に記述される画枠サイズ情報を抽出し、音声選択部９４に供給する。MPD処理部９２は、MPDファイルから、音声選択部９４から要求されるオブジェクトの音声ファイル用の「Segment」に記述されるURL等の情報を抽出し、音声選択部９４に供給する。

[0086] MPD処理部９２は、MPDファイルから画像用の「AdaptationSet」に記述されるタイル位置情報を抽出し、画像選択部９８に供給する。MPD処理部９２は、MPDファイルから、画像選択部９８から要求されるタイルの画像ファイル用の「Segment」に記述されるURL等の情報を抽出し、画像選択部９８に供給する。

[0087] メタファイル取得部９３は、MPD処理部９２から供給されるURL等の情報に基づいて、そのURLで特定される音声メタファイルをWebサーバ１２に要求し、取得する。メタファイル取得部９３は、音声メタファイルに含まれるオブジェクト位置情報を音声選択部９４に供給する。

[0088] 音声選択部９４は、位置決定部として機能し、MPD処理部９２から供給される画枠サイズ情報とメタファイル取得部９３から供給されるオブジェクト位置情報とに基づいて、各オブジェクトの画像上の位置を算出する。音声選択部９４は、各オブジェクトの画像上の位置に基づいて、ユーザにより指定さ

れる表示領域内のオブジェクトを選択する。音声選択部９４は、選択されたオブジェクトの音声ファイルのURL等の情報をMPD処理部９２に要求する。音声選択部９４は、その要求に応じてMPD処理部９２から供給されるURL等の情報を音声ファイル取得部９５に供給する。

[0089] 音声ファイル取得部９５は、受け取り部として機能し、音声選択部９４から供給されるURL等の情報に基づいて、そのURLで特定されるオブジェクト単位の音声ファイルをWebサーバ１２に要求し、取得する。音声ファイル取得部９５は、取得されたオブジェクト単位の音声ファイルを音声復号処理部９６に供給する。

[0090] 音声復号処理部９６は、音声ファイル取得部９５から供給されるオブジェクト単位の音声ファイルに含まれるオーディオストリームを復号し、オブジェクト単位の音声データを生成する。音声復号処理部９６は、そのオブジェクト単位の音声データを音声合成処理部９７に供給する。

[0091] 音声合成処理部９７は、音声復号処理部９６から供給されるオブジェクト単位の音声データを合成し、出力する。

[0092] 画像選択部９８は、MPD処理部９２から供給されるタイル位置情報に基づいて、ユーザにより指定される表示領域内のタイルを選択する。画像選択部９８は、選択されたタイルの画像ファイルのURL等の情報をMPD処理部９２に要求する。画像選択部９８は、その要求に応じてMPD処理部９２から供給されるURL等の情報を画像ファイル取得部９９に供給する。

[0093] 画像ファイル取得部９９は、画像選択部９８から供給されるURL等の情報に基づいて、そのURLで特定されるタイル単位の画像ファイルをWebサーバ１２に要求し、取得する。画像ファイル取得部９９は、取得されたタイル単位の画像ファイルを画像復号処理部１００に供給する。

[0094] 画像復号処理部１００は、画像ファイル取得部９９から供給されるタイル単位の画像ファイルに含まれるビデオストリームを復号し、タイル単位の画像データを生成する。画像復号処理部１００は、そのタイル単位の画像データを画像合成処理部１０１に供給する。

[0095] 画像合成処理部１０１は、画像復号処理部１００から供給されるタイル単位の画像データを合成し、出力する。

[0096] （動画再生端末の処理の説明）

図１４は、動画再生端末１４のストリーミング再生部９０（図１３）のストリーミング再生処理を説明するフローチャートである。

[0097] 図１４のステップＳ３１において、ストリーミング再生部９０のMPD取得部９１は、WebサーバＩ２からMPDファイルを取得し、MPD処理部９２に供給する。

[0098] ステップＳ３２において、MPD処理部９２は、MPD取得部９１から供給されるMPDファイルから、画像用の「AdaptationSet」に記述される画枠サイズ情報とタイル位置情報を取得する。MPD処理部９２は、画枠サイズ情報を音声選択部９４に供給し、タイル位置情報を画像選択部９８に供給する。また、MPD処理部９２は、音声メタファイル用の「Segment」に記述されるURL等の情報を抽出し、メタファイル取得部９３に供給する。

[0099] ステップＳ３３において、メタファイル取得部９３は、MPD処理部９２から供給されるURL等の情報に基づいて、そのURLで特定される音声メタファイルをWebサーバ１２に要求し、取得する。メタファイル取得部９３は、音声メタファイルに含まれるオブジェクト位置情報を音声選択部９４に供給する。

[0100] ステップＳ３４において、音声選択部９４は、MPD処理部９２から供給される画枠サイズ情報とメタファイル取得部９３から供給されるオブジェクト位置情報とに基づいて、ユーザにより指定される表示領域内のオブジェクトを選択する。音声選択部９４は、選択されたオブジェクトの音声ファイルのURL等の情報をMPD処理部９２に要求する。

[0101] MPD処理部９２は、MPDファイルから、音声選択部９４から要求されるオブジェクトの音声ファイル用の「Segment」に記述されるURL等の情報を抽出し、音声選択部９４に供給する。音声選択部９４は、MPD処理部９２から供給されるURL等の情報を音声ファイル取得部９５に供給する。

[0102] ステップＳ３５において、音声ファイル取得部９５は、音声選択部９４か

ら供給されるURL等の情報に基づいて、そのURLで特定される、選択されたオブジェクトの音声ファイルをWebサーバ１２に要求し、取得する。音声ファイル取得部９５は、取得されたオブジェクト単位の音声ファイルを音声復号処理部９６に供給する。

[0103] ステップＳ３６において、画像選択部９８は、MPD処理部９２から供給されるタイル位置情報に基づいて、ユーザにより指定される表示領域内のタイルを選択する。画像選択部９８は、選択されたタイルの画像ファイルのURL等の情報をMPD処理部９２に要求する。

[0104] MPD処理部９２は、MPDファイルから、画像選択部９８から要求されるオブジェクトの画像ファイル用の「Segment」に記述されるURL等の情報を抽出し、画像選択部９８に供給する。画像選択部９８は、MPD処理部９２から供給されるURL等の情報を画像ファイル取得部９９に供給する。

[0105] ステップＳ３７において、画像ファイル取得部９９は、画像選択部９８から供給されるURL等の情報に基づいて、そのURLで特定される、選択されたタイルの画像ファイルをWebサーバ１２に要求し、取得する。画像ファイル取得部９９は、取得されたタイル単位の画像ファイルを画像復号処理部１００に供給する。

[0106] ステップＳ３８において、音声復号処理部９６は、音声ファイル取得部９５から供給されるオブジェクト単位の音声ファイルに含まれるオーディオストリームを復号し、オブジェクト単位の音声データを生成する。音声復号処理部９６は、そのオブジェクト単位の音声データを音声合成処理部９７に供給する。

[0107] ステップＳ３９において、画像復号処理部１００は、画像ファイル取得部９９から供給されるタイル単位の画像ファイルに含まれるビデオストリームを復号し、タイル単位の画像データを生成する。画像復号処理部１００は、そのタイル単位の画像データを画像合成処理部１０１に供給する。

[0108] ステップＳ４０において、音声合成処理部９７は、音声復号処理部９６から供給されるオブジェクト単位の音声データを合成し、出力する。ステップ

Ｓ４１において、画像合成処理部１０１は、画像復号処理部１００から供給されるタイル単位の画像データを合成し、出力する。そして、処理は終了する。

[0109] 以上のように、Webサーバ１２は画枠サイズ情報とオブジェクト位置情報を伝送する。従って、動画再生端末１４は、例えば、表示領域内のオブジェクトを特定し、そのオブジェクトの音声ファイルを、表示領域内の画像に対応する音声ファイルとして選択的に取得することができる。このように、動画再生端末１４は必要な音声ファイルのみを取得することができるので、伝送効率は向上する。

[0110] なお、図１５に示すように、MPDファイルの画像用の「AdaptationSet」には、その画像と同時に再生したい音声に対応するオブジェクトを特定する情報としてのオブジェクトＩＤ（オブジェクト特定情報）が、記述されるようにしてもよい。この記述は、例えば、ViewpointのDescriptorType elementを活用して新たにオブジェクトＩＤ情報（audioObj）を定義するためのScheme(urn:mpeg:DASH:audioObj:2013)を拡張することにより、行われる。この場合、動画再生端末１４は、画像用の「AdaptationSet」に記述されたオブジェクトＩＤに対応するオブジェクトの音声ファイルを選択し、取得して再生する。

[0111] また、オブジェクト単位で音声ファイルが生成されるのではなく、全てのオブジェクトの符号化データが１本のオーディオストリームに多重化され、１つの音声ファイルが生成されるようにしてもよい。

[0112] この場合、図１６に示すように、MPDファイルの音声用の「AdaptationSet」には、音声ファイル用の「Representation」が１つ設けられ、「Segment」に、全てのオブジェクトの符号化データを含む音声ファイル（audioObje.mp4）のURL等が記述される。このとき、Viewpointを拡張することにより、音声ファイルに対応する全てのオブジェクトのオブジェクトＩＤ(1,2,3,4,5)も記述される。

[0113] また、この場合、図１７に示すように、MPDファイルの「Media Segment」

を参照して取得される音声ファイル（以下、適宜、音声メディアファイルともいう）のmdat boxには、サブサンプルとして、各オブジェクトの符号化データ（Audio object）が配置される。

[0114] 具体的には、音声メディアファイルには、セグメントより短い任意の時間であるサブセグメント単位でデータが配置される。このサブセグメント単位のデータの位置は、sidx boxにより指定される。また、サブセグメント単位のデータは、moof boxとmdat boxにより構成される。mdat boxは、複数のサンプルにより構成されるが、そのサンプルの各サブサンプルとして、各オブジェクトの符号化データが配置される。

[0115] また、音声メディアファイルのsidx boxの後には、サンプルの情報を記述するgsix boxが配置される。このように、サンプルの情報を記述するgsix boxが、moof boxとは別に設けられるので、動画再生端末１４は、サンプルの情報を素早く取得することができる。

[0116] gsix boxには、図１８に示すように、このgsix boxで管理する１以上のサンプルまたはサブサンプルからなるSample group entryの種別を表すgrouping_typeが記述される。例えば、Sample group entryが、オブジェクト単位の符号化データのサブサンプルである場合、図１７に示すように、Sample group entryの種別は「obja」である。音声メディアファイルには、複数のgrouping_typeのgsix boxが配置される。

[0117] また、図１８に示すように、gsix boxには、各Sample group entryのインデックス（entry_index）と、音声メディアファイル内の位置を表すデータ位置情報としてのバイトレンジ（range_size）とが記述される。なお、インデックス（entry_index）が０である場合、対応するバイトレンジは、moof boxのバイトレンジ（図１７の例ではa1）を表している。

[0118] MPDファイルの「Initialization Segment」を参照して取得される音声ファイル（以下、適宜、音声初期化ファイルともいう）には、各Sample group entryが、どのオブジェクトの符号化データのサブサンプルであるかを表す情報が記述される。

[0119] 具体的には、図１９に示すように、音声初期化ファイルのsbtl boxのsample group description box(sgpd)のAudioObjectSampleGroupEntryと対応付けられる、mvex boxのtype assignment box（typa）を用いて、その情報が表される。

[0120] 即ち、図２０のＡに示すように、AudioObjectSampleGroupEntryには、サンプルに含まれる符号化データに対応するオブジェクトＩＤ（audio_object_id）が１つずつ記述される。例えば、図２０のＢに示すように、４つのAudioObjectSampleGroupEntryのそれぞれに、オブジェクトＩＤとして、1,2,3,4が記述される。

[0121] 一方、図２１に示すように、type assignment boxには、AudioObjectSampleGroupEntryごとに、そのAudioObjectSampleGroupEntryに対応するSample group entryのパラメータ（grouping_type_parameter）として、インデックスが記述される。

[0122] 以上のように音声メディアファイルと音声初期化ファイルが構成される。従って、動画再生端末１４が、表示領域内のオブジェクトとして選択されたオブジェクトの符号化データを取得する際、音声初期化ファイルのstbl boxから、選択されたオブジェクトのオブジェクトＩＤが記述されたAudioObjectSampleGroupEntryが検索される。そして、mvex boxから、検索されたAudioObjectSampleGroupEntryに対応するSample group entryのインデックスが読み出される。そして、音声ファイルのsidxからサブセグメント単位のデータの位置が読み出され、gsixから、読み出されたインデックスのSample group entryのバイトレンジが読み出される。そして、サブセグメント単位のデータの位置とバイトレンジとに基づいて、mdat内に配置されている符号化データが取得される。これにより、選択されたオブジェクトの符号化データが取得される。

[0123] なお、上述した説明では、Sample group entryのインデックスとAudioObjectSampleGroupEntryのオブジェクトＩＤが、mvex boxを介して対応付けられたが、直接対応付けられるようにしてもよい。この場合、Sample group entr

yのインデックスが、AudioObjectSampleGroupEntryに記述される。

[0124] また、音声ファイルが複数のtrackから構成される場合、sgpdをmvexに格納すると、track間でsgpdを共用できる。

[0125] ＜第２実施の形態＞

（第２実施の形態の概要）

図２２は、本開示を適用した情報処理システムの第２実施の形態の概要を説明する図である。

[0126] なお、図２２において、図３と同一のものには、同一の符号を付してある。

[0127] 図２２の例では、図３の場合と同様に、動画コンテンツの画像は、５（縦）×７（横）のタイルに分割されており、動画コンテンツの音声として、オブジェクト＃１乃至＃８の音声が取得されている。

[0128] この場合に、ユーザにより、２（縦）×３（横）個のタイルからなる表示領域３１が指定されると、第２実施の形態では、図２２に示すように、その表示領域３１が、動画コンテンツの画像のサイズと同一のサイズに変換(拡大)され、表示画像１１１とされる。そして、表示画像１１１におけるオブジェクト＃１乃至＃８の位置に基づいて、オブジェクト＃１乃至＃８の音声が合成され、表示画像１１１とともに出力される。即ち、表示領域３１内のオブジェクト＃１、オブジェクト＃２、およびオブジェクト＃６だけでなく、表示領域３１外のオブジェクト＃３乃至＃５、オブジェクト＃７、およびオブジェクト＃８の音声も出力される。

[0129] （ストリーミング再生部の構成例）

本開示を適用した情報処理システムの第２実施の形態の構成は、ストリーミング再生部の構成を除いて、図１の情報処理システム１０の構成と同一であるので、以下では、ストリーミング再生部についてのみ説明する。

[0130] 図２３は、本開示を適用した情報処理システムのストリーミング再生部の構成例を示すブロック図である。

[0131] 図２３に示す構成のうち、図１３の構成と同じ構成には同じ符号を付して

ある。重複する説明については適宜省略する。

[0132] 図２３のストリーミング再生部１２０の構成は、MPD処理部９２、音声合成処理部９７、画像合成処理部１０１の代わりに、MPD処理部１２１、音声合成処理部１２３、画像合成処理部１２４が設けられる点、および、位置決定部１２２が新たに設けられる点が図１３のストリーミング再生部９０の構成と異なる。

[0133] ストリーミング再生部１２０のMPD処理部１２１は、MPD取得部９１から供給されるMPDファイルから音声メタファイル用の「Segment」に記述されるURL等の情報を抽出し、メタファイル取得部９３に供給する。また、MPD処理部１２１は、MPDファイルから画像用の「AdaptationSet」に記述される動画コンテンツの画像の画枠サイズ情報（以下、コンテンツ画枠サイズ情報という）を抽出し、位置決定部１２２に供給する。MPD処理部１２１は、MPDファイルから全てのオブジェクトの音声ファイル用の「Segment」に記述されるURL等の情報を抽出し、音声ファイル取得部９５に供給する。

[0134] MPD処理部１２１は、MPDファイルから画像用の「AdaptationSet」に記述されるタイル位置情報を抽出し、画像選択部９８に供給する。MPD処理部１２１は、MPDファイルから、画像選択部９８から要求されるタイルの画像ファイル用の「Segment」に記述されるURL等の情報を抽出し、画像選択部９８に供給する。

[0135] 位置決定部１２２は、メタファイル取得部９３により取得された音声メタファイルに含まれるオブジェクト位置情報と、MPD処理部１２１から供給されるコンテンツ画枠サイズ情報を取得する。また、位置決定部１２２は、ユーザにより指定される表示領域の画枠サイズ情報である表示領域画枠サイズ情報を取得する。位置決定部１２２は、オブジェクト位置情報、コンテンツ画枠サイズ情報、および表示領域画枠サイズ情報に基づいて、表示画像におけるオブジェクトの位置を決定（認識）する。位置決定部１２２は、決定されたオブジェクトの位置を音声合成処理部１２３に供給する。

[0136] 音声合成処理部１２３は、位置決定部１２２から供給されるオブジェクト

の位置に基づいて、音声復号処理部９６から供給されるオブジェクト単位の音声データを合成する。具体的には、音声合成処理部１２３は、オブジェクトの位置と音声を出力する各スピーカの位置とに基づいて、オブジェクトごとに、各スピーカに割り当てる音声データを決定する。そして、音声合成処理部１２３は、スピーカごとに、各オブジェクトの音声データを合成し、各スピーカの音声データとして出力する。オブジェクトの位置に基づいて各オブジェクトの音声データを合成する方法の詳細は、例えば、Ville Pulkki,“Virtual Sound Source Positioning Using Vector Base Amplitude Panning”, Journal of AES, vol.45, no.6, pp.456-466, 1997に記載されている。

[0137] 画像合成処理部１２４は、画像復号処理部１００から供給されるタイル単位の画像データを合成する。画像合成処理部１２４は、変換部として機能し、合成後の画像データに対応する画像のサイズを動画コンテンツのサイズに変換して、表示画像を生成する。画像合成処理部１２４は、表示画像を出力する。

[0138] （オブジェクトの位置の決定方法の説明）

図２４乃至図２６は、図２３の位置決定部１２２によるオブジェクトの位置の決定方法を説明する図である。

[0139] 表示画像１１１は、動画コンテンツ内の表示領域３１を抽出し、その表示領域３１のサイズを動画コンテンツのサイズに変換することにより生成される。従って、表示画像１１１は、図２４に示すように、表示領域３１の中心Cを表示画像１１１の中心C´に移動させた後、図２５に示すように、表示領域３１のサイズを動画コンテンツのサイズに変換したものと等価である。

[0140] よって、まず、位置決定部１２２は、以下の式（１）により、表示領域３１の中心０を表示画像１１１の中心０´に移動させるときの水平方向の移動量θ_{shift}を算出する。

[0141] [数1]

$$\theta_{shift} = \frac{\theta_{v1}' + \theta_{v2}' - \theta_{v1} - \theta_{v2}}{2} \quad \cdots (1)$$

[0142] 式（１）において、θ_{v1}'は、表示領域３１の表示領域画枠サイズ情報のうちの左端の水平角度であり、θ_{V2}'は、右端の水平角度である。また、θ_{v1}は、コンテンツ画枠サイズ情報のうちの左端の水平角度であり、θ_{v2}は、右端の水平角度である。

[0143] 次に、位置決定部１２２は、移動量θ_{shift}を用いて、以下の式（２）により、表示領域３１の中心Ｏを表示画像１１１の中心Ｏ′に移動させた後の表示領域３１の左端の水平角度$\theta_{v1\ shift}'$と右端の水平角度$\theta_{v2\ shift}'$を求める。

[0144] [数2]

$$\theta_{v1_shift}' = \mathrm{mod}(\theta_{v1}' + \theta_{shift} + 180°, 360°) - 180°$$

$$\theta_{v2_shift}' = \mathrm{mod}(\theta_{v2}' + \theta_{shift} + 180°, 360°) - 180° \quad \cdots (2)$$

[0145] 式（２）によれば、水平角度$\theta_{v1\ shift}'$と水平角度$\theta_{v2\ shift}'$は、-180°から180°までの範囲を超えないように求められる。

[0146] なお、上述したように、表示画像１１１は、表示領域３１の中心Ｏを表示画像１１１の中心Ｏ′に移動させた後、表示領域３１のサイズを動画コンテンツのサイズに変換したものと等価である。従って、水平角度θ_{V1}およびθ_{V2}については、以下の式（３）が成立する。

[0147] [数3]

$$\theta_{v1} = \frac{\theta_{v1} - \theta_{v2}}{\theta_{v1_shift}' - \theta_{v2_shift}'} * \theta_{v1_shift}'$$

$$\theta_{v2} = \frac{\theta_{v1} - \theta_{v2}}{\theta_{v1_shift}' - \theta_{v2_shift}'} * \theta_{v2_shift}' \quad \cdots (3)$$

[0148] 位置決定部１２２は、以上のようにして移動量θ_{shift}、水平角度$\theta_{v1\ shift}'$、および水平角度$\theta_{v2\ shift}'$を求めた後、表示画像１１１におけるオブジェクトの水平角度を求める。具体的には、位置決定部１２２は、以下の式（４）により、移動量θ_{shift}を用いて、表示領域３１の中心Cを表示画像１１１の中心C′に移動させた後のオブジェクト＃ｉの水平角度$\theta_{Ai\ shift}$を求める。

[0149]

[数4]

$$\theta_{Ai_shift} = \mathrm{mod}(\theta_{Ai}+\theta_{shift}+180°, 360°)-180° \quad \cdots (4)$$

[0150] 式（４）において、θ_{Ai}は、オブジェクト＃ｉのオブジェクト位置情報のうちの水平角度である。また、式（４）によれば、水平角度$\theta_{Ai\ shift}$は、-180°から180°までの範囲を超えないように求められる。

[0151] 次に、オブジェクト＃ｉが表示領域３１内に存在する場合、即ち$\theta_{v2\ shift}' < \theta_{Ai\ shift} < \theta_{v1\ shift}'$である場合、位置決定部１２２は、以下の式（５）により、表示画像１１１におけるオブジェクト＃ｉの水平角度θ_{Ai}'を求める。

[0152] [数5]

$$\theta_{Ai}' = \frac{\theta_{v1}-\theta_{v2}}{\theta_{v1_shift}'-\theta_{v2_shift}'} \cdot \left(\theta_{Ai_shift}-\frac{\theta_{v1}+\theta_{v2}}{2}\right) \quad \cdots (5)$$

[0153] 式（５）によれば、表示領域３１のサイズと表示画像１１１のサイズの比に応じて、表示画像１１１におけるオブジェクト＃ｉの位置と表示画像１１１の中心C´との距離を拡大することにより、水平角度θ_{Ai}'が求められる。

[0154] 一方、オブジェクト＃ｉが表示領域３１内に存在しない場合、即ち、$-180° \leqq \theta_{Ai\ shift} \leqq \theta_{v2\ shift}'$、または、$\theta_{v1\ shift}' \leqq \theta_{Ai\ shift} \leqq 180°$である場合、位置決定部１２２は、以下の式（６）により、表示画像１１１におけるオブジェクト＃ｉの水平角度θ_{Ai}'を求める。

[0155] [数6]

$$\theta_{Ai}' = \frac{\theta_{v2}+180°}{\theta_{v2_shift}'+180°} \cdot (\theta_{Ai_shift}+180°)-180°$$

$$(-180° \leq \theta_{Ai_shift} \leq \theta_{v2_shift}' \text{の場合})$$

$$\theta_{Ai}' = \frac{180°-\theta_{v1}}{180°-\theta_{v1_shift}'} \cdot (\theta_{Ai_shift}-180°)+180°$$

$$(\theta_{v1_shift}' \leq \theta_{Ai_shift} \leq 180° \text{の場合})$$

$$\cdots (6)$$

[0156] 式（６）によれば、図２６に示すように、オブジェクト＃ｉが表示領域３

１の右側の位置１５１に存在する（$-180° \leqq \theta_{Ai\ shift} \leqq \theta_{v2\ shift}'$）場合、角度Ｒ１と角度Ｒ２の比に応じて水平角度$\theta_{Ai\ shift}$を拡大することにより、水平角度$\theta_{Ai}'$が求められる。なお、角度Ｒ１は、表示画像１１１の右端から視聴者１５３の真後ろの位置１５４までの角度であり、角度Ｒ２は、中心の移動後の表示領域３１の右端から位置１５４までの角度である。

[0157] また、式（６）によれば、オブジェクト＃ｉが表示領域３１の左側の位置１５５に存在する（$\theta_{v1\ shift}' \leqq \theta_{Ai\ shift} \leqq 180°$）場合、角度Ｒ３と角度Ｒ４の比に応じて水平角度$\theta_{Ai\ shift}$を拡大することにより、水平角度$\theta_{Ai}'$が求められる。なお、角度Ｒ３は、表示画像１１１の左端から位置１５４までの角度であり、角度Ｒ４は、中心の移動後の表示領域３１の左端から位置１５４までの角度である。

[0158] また、位置決定部１２２は、水平角度θ_{Ai}と水平角度θ_{Ai}'に基づいて、表示画像１１１におけるオブジェクト＃ｉの垂直角度γ_{Ai}'を求める。具体的には、オブジェクト＃ｉが視聴者の前方にある場合、位置決定部１２２は、垂直角度γ_{Ai}'を水平角度θ_{Ai}'と同様に求める。

[0159] 一方、オブジェクト＃ｉが視聴者の後方にある場合、表示領域３１が拡大されると、水平角度θ_{Ai}が表すオブジェクト＃ｉの位置は視聴者の真後ろの方向に移動する。従って、垂直角度γ_{Ai}'が水平角度θ_{Ai}'と同様に求められると、オブジェクト＃ｉのオブジェクト位置情報のうちの垂直角度γ_{Ai}が表す位置は、視聴者の真上か真下の方向に移動する。しかしながら、オブジェクト＃ｉが視聴者の後方にある場合、オブジェクト＃ｉが水平角度θ_{Ai}の変更によって視聴者の前方に移動した場合を除いて、表示領域３１が拡大または縮小しても垂直角度γ_{Ai}'は変化させない、または、水平角度θ_{Ai}の変化方向と同一の方向に変化させることが望ましい。

[0160] よって、水平角度θ_{Ai}が表すオブジェクト＃ｉの位置と、水平角度θ_{Ai}'が表すオブジェクト＃ｉの位置が、両方とも視聴者の前方である場合、即ち$abs(\theta_{Ai}) < 90°$かつ$abs(\theta_{Ai}') < 90°$である場合、位置決定部１２２は、以下の式（７）により垂直角度γ_{Ai}'を求める。

[0161] [数7]

$$\gamma_{Ai}' = \begin{cases} \dfrac{\gamma_{v2}'+90^\circ}{\gamma_{v2}+90^\circ}\cdot(\gamma_{Ai}+90^\circ)-90^\circ & \text{for} \quad -90^\circ \leq \gamma_{Ai} < \gamma_{v2} \\ \dfrac{\gamma_{v1}'-\gamma_{v2}'}{\gamma_{v1}-\gamma_{v2}}\cdot\left(\gamma_{Ai}-\dfrac{\gamma_{v1}'+\gamma_{v2}'}{2}\right) & \text{for} \quad \gamma_{v2} \leq \gamma_{Ai} < \gamma_{v1} \\ \dfrac{90^\circ-\gamma_{v1}'}{90^\circ-\gamma_{v1}}\cdot(\gamma_{Ai}-\gamma_{v1})+\gamma_{v1}' & \text{for} \quad \gamma_{v1} \leq \gamma_{Ai} < 90^\circ \end{cases} \quad \cdots (7)$$

[0162] 式（７）において、γ_{v1}'は、表示領域３１の表示領域画枠サイズ情報のうちの上端の垂直角度であり、γ_{v2}'は、下端の垂直角度である。また、γ_{v1}は、コンテンツ画枠サイズ情報のうちの上端の垂直角度であり、γ_{v2}は、下端の垂直角度である。

[0163] また、水平角度θ_{Ai}が表すオブジェクト＃ｉの位置と、水平角度θ_{Ai}'が表すオブジェクト＃ｉの位置が、両方とも視聴者の後方である場合、即ちabs(θ_{Ai})≧90°かつabs(θ_{Ai}'≧90°である場合、位置決定部１２２は、以下の式（８）により垂直角度γ_{Ai}'を求める。即ち、垂直角度γ_{Ai}がそのまま垂直角度γ_{Ai}'にされる。

[0164] [数8]

$$\gamma_{Ai}' = \gamma_{Ai} \quad \cdots (8)$$

[0165] さらに、水平角度θ_{Ai}が表すオブジェクト＃ｉの位置と水平角度θ_{Ai}'が表すオブジェクト＃ｉの位置のいずれか一方が視聴者の前方であり、他方が後方である場合、即ちabs(θ_{Ai})＜90°かつabs(θ_{Ai}')≧90°、または、abs(θ_{Ai})≧90°かつabs(θ_{Ai}')＜90°である場合、位置決定部１２２は、以下のようにして垂直角度γ_{Ai}'を求める。

[0166] 即ち、この場合、オブジェクト＃ｉは視聴者の前方から後方へ、または後方から前方へ移動するため、単純にオブジェクト＃ｉが視聴者の前方または後方に位置すると決め付けることはできない。従って、この場合、オブジェ

クト＃ｉの水平角度が、視聴者の前方のオブジェクト＃ｉの位置を表す角度と、視聴者の左側を表す角度である90°または右側を表す角度である-90°との間を移動する段階と、90°または-90°と視聴者の後方のオブジェクト＃ｉの位置を表す角度との間を移動する段階に分けて考える必要がある。

[0167] 具体的には、位置決定部１２２は、オブジェクト＃ｉの水平角度が、90°または-90°と視聴者の後方のオブジェクト＃ｉの位置を表す角度との間を移動する段階の垂直角度γ_{Ai}の移動量を０に決定する。そして、位置決定部１２２は、オブジェクト＃ｉの水平角度が、視聴者の前方のオブジェクト＃ｉの位置を表す角度と、視聴者の左側を表す角度である90°または右側を表す角度である-90°との間を移動する段階の垂直角度γ_{Ai}の移動量を、式（７）と同様に求める。但し、式（９）に示すように、式（７）の垂直角度γ_{V2}'，γ_{V2}，γ_{V1}'，γ_{V1}は、式(10)で定義される垂直角度γ_{V2f}'，γ_{V2f}，γ_{V1f}'，γ_{V1f}に置換される。

[0168] [数9]

$$\gamma_{Ai}' = \begin{cases} \dfrac{\gamma_{v2f}'+90^\circ}{\gamma_{v2f}+90^\circ}\cdot(\gamma_{Ai}+90^\circ)-90^\circ & \text{for} \quad -90^\circ \le \gamma_{Ai} < \gamma_{v2f} \\ \dfrac{\gamma_{v1f}'-\gamma_{v2f}'}{\gamma_{v1f}-\gamma_{vf2}}\cdot\left(\gamma_{Ai}-\dfrac{\gamma_{v1f}'+\gamma_{v2f}'}{2}\right) & \text{for} \quad \gamma_{v2f} \le \gamma_{Ai} < \gamma_{v1f} \\ \dfrac{90^\circ-\gamma_{v1f}'}{90^\circ-\gamma_{v1f}}\cdot(\gamma_{Ai}-\gamma_{v1f})+\gamma_{v1f}' & \text{for} \quad \gamma_{v1f} \le \gamma_{Ai} < 90^\circ \end{cases} \quad \cdots (9)$$

[0169]

[数10]

$(\theta_{Ai}<90^\circ \ \& \ \theta_{Ai}'\geqq 90^\circ)$の場合:

$$\gamma_{v1f}=\gamma_{v1}$$
$$\gamma_{v2f}=\gamma_{v2}$$
$$\gamma_{v1f}'=F_{side}\cdot\gamma_{v1}$$
$$\gamma_{v2f}'=F_{side}\cdot\gamma_{v2}$$

$(\theta_{Ai}\geqq 90^\circ \ \& \ \theta_{Ai}'<90^\circ)$の場合:

$$\gamma_{v1f}=F_{side}\cdot\gamma_{v1}$$
$$\gamma_{v2f}=F_{side}\cdot\gamma_{v2}$$
$$\gamma_{v1f}'=\gamma_{v1}'$$
$$\gamma_{v2f}'=\gamma_{v2}' \qquad \cdots (10)$$

[0170] 式（１０）におけるF_{side}は、オブジェクト＃ｉの水平角度が90°または-90°である場合の表示領域３１の垂直角度と表示画像１１１の垂直角度との比を表すものであり、以下の式（１１）により求められる。

[0171] [数11]

$$F_{side}=\begin{cases}\dfrac{180^\circ-\dfrac{90^\circ}{180^\circ-\gamma_{Ai}}\cdot(180^\circ-\theta_{v1})}{\theta_{v1}}\cdot TransForm_Factor & for \quad \gamma_{Ai}\geq 0^\circ \\ \dfrac{\dfrac{90^\circ}{180^\circ+\gamma_{Ai}}\cdot(180^\circ+\theta_{v2})-180^\circ}{\theta_{v2}}\cdot TransForm_Factor & for \quad \gamma_{Ai}<0^\circ\end{cases} \qquad \cdots (11)$$

[0172] 式（１１）に示すように、F_{side}は、オブジェクト＃ｉの水平角度が90°または-90°である場合の表示領域３１の水平角度と表示画像１１１の水平角度の比にTransForm_Factorを乗じることで求められる。

[0173] 式（１１）におけるTransForm_Factorは、垂直角度γ_{v1}、γ_{v2}と水平角度θ_{v1}

、θ_{v2}の比を表すファクタであり、以下の式（１２）で表される。このファクタTransForm_Factorを用いて、水平角度の変化量に対応する垂直角度の変化量を推定することができる。

[0174] [数12]

$$\mathrm{TransForm_Factor} = \frac{\gamma_{v1}' - \gamma_{v2}'}{\gamma_{v1} - \gamma_{v2}} \Big/ \frac{\theta_{v1}' - \theta_{v2}'}{\theta_{v1} - \theta_{v2}} \qquad \cdots (12)$$

[0175] 以上のように、オブジェクト＃ｉが視聴者の前方から後方へ移動する場合、即ちabs(θ_{Ai})＜90°かつabs(θ_{Ai}')≧90°である場合、オブジェクト＃ｉの垂直角度の変化量は、オブジェクト＃ｉの水平角度が水平角度θ_{Ai}から90°または-90°まで移動する場合の変化量となる。従って、式（７）のγ_{V1}'とγ_{V2}'は、abs(θ_{Ai}')＝90°である場合の画枠サイズ、即ち、F_{side}と垂直角度γ_{v1}とγ_{v2}の積となる。

[0176] また、オブジェクト＃ｉが視聴者の後方から前方へ移動する場合、即ちabs(θ_{Ai})≧90°かつabs(θ_{Ai}')＜90°である場合、オブジェクト＃ｉの垂直角度の変化量は、オブジェクト＃ｉの水平角度が90°または-90°から水平角度θ_{Ai}'まで移動する場合の変化量となる。従って、式（７）のγ_{v1}とγ_{v2}は、Fsideと垂直角度γ_{v1}とγ_{v2}の積となる。

[0177] ここでは、オブジェクト＃ｉが視聴者の後方にある場合の垂直角度は変化しないものするが、垂直角度は、水平角度と同じ方向に変化してもよい。即ち、水平角度が表す位置が視聴者の真後ろの方向に移動する場合には、垂直角度が表す位置も視聴者の真後ろの方向に移動し、水平角度が表す位置が視聴者の真正面の方向に移動する場合には、垂直角度が表す位置も視聴者の真正面の方向に移動するようにしてもよい。

[0178] また、位置決定部１２２は、表示画像１１１におけるオブジェクト＃ｉの距離r_{Ai}'を、オブジェクト＃ｉのオブジェクト位置情報のうちの距離r_{Ai}に決定する。位置決定部１２２は、以上のようにして求められるオブジェクト

＃ｉの水平角度θ_{Ai}'、垂直角度γ_{Ai}'、および距離r_{Ai}を、オブジェクト＃ｉの位置として、音声合成処理部１２３に供給する。

[0179] 図２７は、水平角度θ_{Ai}と水平角度θ_{Ai}'の関係を表す図である。

[0180] 図２７のグラフにおいて、横軸は、水平角度θ_{Ai}を表し、縦軸は、水平角度θ_{Ai}'を表す。

[0181] 図２７に示すように、$\theta_{V2}' < \theta_{Ai} < \theta_{V1}'$である場合、水平角度$\theta_{Ai}$は、移動量$\theta_{shift}$だけ移動され、拡大されることにより、水平角度$\theta_{Ai}'$になる。また、$-180° \leqq \theta_{Ai} \leqq \theta_{v2}'$、または、$\theta_{v1}' \leqq \theta_{Ai} \leqq 180°$である場合、水平角度$\theta_{Ai}$は、移動量$\theta_{shift}$だけ移動され、縮小されることにより、水平角度$\theta_{Ai}'$になる。

[0182] 図２８は、垂直角度γ_{v1}'と垂直角度γ_{Ai}'の関係を表す図である。

[0183] 図２８のグラフにおいて、横軸は、垂直角度γ_{v1}'を表し、縦軸は、垂直角度γ_{Ai}'を表す。

[0184] 図２８の例では、水平角度θ_{Ai}と垂直角度γ_{Ai}が45°である。また、垂直角度γ_{v1}, γ_{v2}は、水平角度θ_{v1}, θ_{v2}と同一の比率で変化する。即ち、TransForm_Factor＝１である。また、$\theta_{v1}=29°$、$\theta_{v2}=-29°$、$\gamma_{v1}=17.5°$、$\gamma_{v2}=-17.5$である。

[0185] この場合、図２８に示すように、$0° < \gamma_{v1}' < (F_{side}*\gamma_{v1})$である場合、垂直角度$\gamma_{Ai}'$は、表示画像１１１の垂直方向の画枠サイズが大きいほど、大きくなる。また、$(F_{side}*\gamma_{v1}) \leqq \gamma_{v1}' \leqq 90°$である場合、垂直角度$\gamma_{Ai}'$は、表示画像１１１の垂直方向の画枠サイズの大きさによらず、$\gamma_{v1}' = (F_{side}*\gamma_{v1})$のとき、即ちオブジェクト＃ｉの水平角度$\theta_{Ai}'$が90°のときの垂直角度$\gamma_{Ai}'$のままになる。

[0186] （ストリーミング再生部の処理の説明）

図２９は、図２３のストリーミング再生部１２０のストリーミング再生処理を説明するフローチャートである。

[0187] 図２９のステップＳ１３１において、ストリーミング再生部１２０のMPD取得部９１は、Webサーバ１２からMPDファイルを取得し、MPD処理部１２１に供

給する。

[0188] ステップＳ１３２において、MPD処理部１２１は、MPD取得部９１から供給されるMPDファイルから、画像用の「AdaptationSet」に記述されるコンテンツ画枠サイズ情報とタイル位置情報を取得する。MPD処理部１２１は、画枠サイズ情報を位置決定部１２２に供給し、タイル位置情報を画像選択部９８に供給する。また、MPD処理部１２１は、音声メタファイル用の「Segment」に記述されるURL等の情報を抽出し、メタファイル取得部９３に供給する。

[0189] ステップＳ１３３において、メタファイル取得部９３は、MPD処理部１２１から供給されるURL等の情報に基づいて、そのURLで特定される音声メタファイルをWebサーバ１２に要求し、取得する。メタファイル取得部９３は、音声メタファイルに含まれるオブジェクト位置情報を位置決定部１２２に供給する。

[0190] ステップＳ１３４において、位置決定部１２２は、オブジェクト位置情報、コンテンツ画枠サイズ情報、および表示領域画枠サイズ情報に基づいて、表示画像におけるオブジェクトの位置を決定する位置決定処理を行う。この位置決定処理の詳細は、後述する図３０を参照して説明する。

[0191] ステップＳ１３５において、MPD処理部１２１は、MPDファイルから全てのオブジェクトの音声ファイル用の「Segment」に記述されるURL等の情報を抽出し、音声ファイル取得部９５に供給する。

[0192] ステップＳ１３６において、音声ファイル取得部９５は、MPD処理部１２１から供給されるURL等の情報に基づいて、そのURLで特定される全てのオブジェクトの音声ファイルをWebサーバ１２に要求し、取得する。音声ファイル取得部９５は、取得されたオブジェクト単位の音声ファイルを音声復号処理部９６に供給する。

[0193] ステップＳ１３７乃至Ｓ１４０の処理は、図１４のステップＳ３６乃至Ｓ３９の処理と同様であるので、説明は省略する。

[0194] ステップＳ１４１において、音声合成処理部１２３は、位置決定部１２２から供給されるオブジェクトの位置に基づいて、音声復号処理部９６から供

給されるオブジェクト単位の音声データを合成し、出力する。

[0195] ステップＳ１４２において、画像合成処理部１２４は、画像復号処理部１００から供給されるタイル単位の画像データを合成する。

[0196] ステップＳ１４３において、画像合成処理部１２４は、合成後の画像データに対応する画像のサイズを動画コンテンツのサイズに変換し、表示画像を生成する。そして、画像合成処理部１２４は、表示画像を出力し、処理を終了する。

[0197] 図３０は、図２９のステップＳ１３４の位置決定処理の詳細を説明するフローチャートである。この位置決定処理は、例えば、オブジェクトごとに行われる。

[0198] 図３０のステップＳ１５１において、位置決定部１２２は、表示画像における水平角度θ_{Ai}'を推定する水平角度θ_{Ai}'推定処理を行う。この水平角度θ_{Ai}'推定処理の詳細は、後述する図３１を参照して説明する。

[0199] ステップＳ１５２において、位置決定部１２２は、表示画像における垂直角度γ_{Ai}'を推定する垂直角度γ_{Ai}'推定処理を行う。この垂直角度γ_{Ai}'推定処理の詳細は、水平方向が垂直方向に代わる点を除いて、ステップＳ１５１の水平角度θ_{Ai}'推定処理と同様であるので、詳細な説明は省略する。

[0200] ステップＳ１５３において、位置決定部１２２は、表示画像における距離r_{Ai}'を、メタファイル取得部９３から供給されるオブジェクト位置情報のうちの距離r_{Ai}に決定する。

[0201] ステップＳ１５４において、位置決定部１２２は、水平角度θ_{Ai}'、垂直角度γ_{Ai}'、および距離r_{Ai}を、オブジェクト＃ｉの位置として音声合成処理部１２３に出力する。そして、処理は、図２９のステップＳ１３４に戻り、ステップＳ１３５に進む。

[0202] 図３１は、図３０のステップＳ１５１の水平角度θ_{Ai}'推定処理の詳細を説明するフローチャートである。

[0203] 図３１のステップＳ１７１において、位置決定部１２２は、メタファイル取得部９３から供給されるオブジェクト位置情報のうちの水平角度θ_{Ai}を取得

する。

[0204] ステップＳ１７２において、位置決定部１２２は、MPD処理部１２１から供給されるコンテンツ画枠サイズ情報と、ユーザにより指定される表示領域画枠サイズ情報を取得する。

[0205] ステップS１７３において、位置決定部１２２は、コンテンツ画枠サイズ情報と表示領域画枠サイズ情報に基づいて、上述した式（１）により、移動量θ_{shift}を算出する。

[0206] ステップS１７４において、位置決定部１２２は、移動量θ_{shift}と表示領域画枠サイズを用いて、上述した式（２）により、水平角度$\theta_{v1\ shift}'$および$\theta_{v2\ shift}'$を算出する。

[0207] ステップＳ１７５において、位置決定部１２２は、水平角度θ_{Ai}と移動量θ_{shift}を用いて、上述した式（４）により、水平角度$\theta_{Ai\ shift}$を求める。

[0208] ステップＳ１７６において、位置決定部１２２は、オブジェクト＃ｉが表示領域３１内に存在する（オブジェクト＃ｉの水平角度が表示領域３１の両端の水平角度の間である）かどうか、即ち$\theta_{v2\ shift}' < \theta_{Ai\ shift} < \theta_{v1\ shift}'$であるかどうかを判定する。

[0209] ステップＳ１７６でオブジェクト＃ｉが表示領域３１内に存在すると判定された場合、即ち$\theta_{v2\ shift}' < \theta_{Ai\ shift} < \theta_{v1\ shift}'$である場合、処理はステップＳ１７７に進む。ステップＳ１７７において、位置決定部１２２は、コンテンツ画枠サイズ情報、水平角度$\theta_{v1\ shift}'$および$\theta_{v2\ shift}'$、並びに水平角度$\theta_{Ai\ shift}$に基づいて、上述した式（５）により水平角度θ_{Ai}'を算出する。

[0210] 一方、ステップＳ１７６でオブジェクト＃ｉが表示領域３１内に存在しないと判定された場合、即ち、$-180° \leqq \theta_{Ai\ shift} \leqq \theta_{v2\ shift}'$、または、$\theta_{v1\ shift}' \leqq \theta_{Ai\ shift} \leqq 180°$である場合、処理はステップＳ１７８に進む。ステップＳ１７８において、位置決定部１２２は、コンテンツ画枠サイズ情報、水平角度$\theta_{v1\ shift}'$または$\theta_{v2\ shift}'$、および水平角度$\theta_{Ai\ shift}$に基づいて、上述した式（６）により水平角度θ_{Ai}'を算出する。

[0211] ステップＳ１７７またはステップＳ１７８の処理後、処理は、図３０のス

テップＳ１５１に戻り、処理はステップＳ１５２に進む。

[0212] なお、第２実施の形態では、表示画像のサイズが動画コンテンツのサイズと同一であるようにしたが、異なっていてもよい。

[0213] また、第２実施の形態において、全てのオブジェクトの音声データが合成されて出力されるのではなく、一部のオブジェクト(例えば、表示領域内のオブジェクト、表示領域から所定の範囲内のオブジェクト等)の音声データのみが合成されて出力されるようにしてもよい。出力する音声データのオブジェクトの選択方法は、予め決定されていてもよいし、ユーザにより指定されるようにしてもよい。

[0214] ＜第３実施の形態＞

（表示画像におけるオブジェクトの位置の決定方法の説明）

本開示を適用した情報処理システムの第３実施の形態は、図２３の位置決定部１２２による表示画像におけるオブジェクトの位置の決定方法を除いて、第２実施の形態と同一であるため、以下では、図２３の位置決定部１２２による表示画像におけるオブジェクトの位置の決定方法についてのみ説明する。

[0215] 視聴者が動画コンテンツを試聴する場合、視聴者から画枠への距離が変化すると、視聴者から見る画枠の角度もそれに応じて変化する。例えば、視聴者が画枠に近づくと、画枠は大きく見えるようになり、画枠から遠くなると、画枠は小さく見えるようになる。従って、表示領域や画枠サイズの変化は、視聴者の移動によって表現することができる。そこで、第３実施の形態では、視聴者の移動によって表示領域のサイズ変更が表現される。

[0216] 図３２乃至図３５は、本開示を適用した情報処理システムの第３実施の形態における表示画像におけるオブジェクトの位置の決定方法を説明する図である。

[0217] 図３２と図３３は、動画コンテンツ内の表示領域３１のサイズを変更する前の状態を説明する図である。図３２は、視聴者の頭頂部から見た上面図であり、図３２は視聴者の右側から見た側面図である。

[0218] 図３４と図３５は、視聴者が移動し、視聴者の移動によって視聴者から見る表示領域３１のサイズが動画コンテンツのサイズと同じサイズになった場合の状態を説明する図である。図３４は視聴者の頭頂部から見た上面図であり、図３５は視聴者の右側から見た側面図である。

[0219] 図３２乃至図３５に示すように、表示領域３１の内部にあるオブジェクト＃１、並びに、表示領域３１の外部にあるオブジェクト＃２およびオブジェクト＃３と、視聴者との相対的位置関係は、視聴者の移動によって変化する。よって、位置決定部１２２は、表示領域３１のサイズを動画コンテンツのサイズに変更した場合のオブジェクト位置情報、即ち表示画像におけるオブジェクト位置情報を、視聴者の移動によって変更されたオブジェクトのオブジェクト位置情報を求めることで推定する。

[0220] 具体的には、まず、位置決定部１２２は、式（１３）により、移動前の表示領域３１の横幅$Screen_{Width}$、左端から中央までの幅$Screen_{LeftWidth}$、および上端から中央までの高さ$Screen_{TopHeight}$を求める。

[0221] [数13]

$$Screen_{LeftWidth}=r_v\cdot\tan(\theta_{v1}')$$

$$Screen_{TopHeight}=r_v\cdot\tan(\gamma_{v1}')$$

$$Screen_{Width}=r_v\cdot(\tan(\theta_{v1}')-\tan(\theta_{v2}')) \quad \cdots(13)$$

[0222] 式（１３）において、r_vは、視聴者から画枠までの距離であり、θ_{v1}'は、表示領域３１の表示領域画枠サイズ情報のうちの左端の水平角度であり、θ_{v2}'は、右端の水平角度であり、γ_{v1}'は、上端の垂直角度である。

[0223] 次に、位置決定部１２２は、式（１４）により、視聴者の移動によって視聴者から見る表示領域３１のサイズが動画コンテンツのサイズと同じサイズになった場合の、視聴者から画枠までの距離r_v'を推定する。

[0224] [数14]

$$r_v'=\frac{Screen_{Width}}{\tan(\theta_{v1})-\tan(\theta_{v2})} \quad \cdots(14)$$

[0225] 式（１４）において、θ_{v1}は、移動後の表示領域３１の表示領域画枠サイズ情報、即ち、コンテンツ画枠サイズ情報のうちの左端の水平角度であり、θ_{v2}は、右端の水平角度である。

[0226] 次に、位置決定部１２２は、式（１５）により、視聴者の移動によって視聴者から見る表示領域３１のサイズが動画コンテンツのサイズと同じサイズになった場合の、視聴者の前後方向の移動量$Length_{shift}$、左右方向の移動量$Width_{shif}$、および上下方向の移動量$Height_{shift}$を推定する。

[0227] [数15]

$$Length_{shift}=r_v-r_v'$$

$$Width_{shift}=Screen_{LeftWidth}-r_v'\cdot\tan(\theta_{v1})$$

$$Height_{shift}=Screen_{TopHeight}\cdot F_{Aspect_ratio}-r_v'\cdot\tan(\gamma_{v1}) \quad \cdots(15)$$

[0228] 式（１５）において、γ_{v1}は、移動後の表示領域３１の表示領域画枠サイズ情報、即ち、コンテンツ画枠サイズ情報のうちの上端の垂直角度である。また、$F_{Aspect\ ratio}$は、移動前と移動後の視聴者から見る表示領域３１の垂直方向の幅と水平方向の幅との比の変化量を表すファクタであり、以下の式（１６）で表される。このファクタ$F_{Aspect\ ratio}$を用いて、水平方向の幅の変化量に対応する垂直方向の幅の変化量を推定することができる。

[0229] [数16]

$$F_{Aspect_ratio}=\frac{(\tan(\gamma_{v1}')-\tan(\gamma_{v2}'))\cdot(\tan(\theta_{v1})-\tan(\theta_{v2}))}{(\tan(\theta_{v1}')-\tan(\theta_{v2}'))\cdot(\tan(\gamma_{v1})-\tan(\gamma_{v2}))} \quad \cdots(16)$$

[0230] 式（１６）において、γ_{v2}は、移動後の表示領域３１の表示領域画枠サイズ情報、即ち、コンテンツ画枠サイズ情報のうちの下端の垂直角度である。γ_{v2}'は、表示領域３１の表示領域画枠サイズ情報のうちの下端の垂直角度である。

[0231] 最後に、以上により求められた情報を用いて、位置決定部１２２は、表示画像におけるオブジェクトのオブジェクト位置情報を推定する。

[0232] 具体的には、オブジェクトが表示領域３１上に存在する場合、即ち、θ_{v2}' ≦θ_{Ai}≦θ_{v1}'かつγ_{v2}' ≦γ_{Ai}≦γ_{v1}'である場合（図３２乃至図３５の例では、オブジェクト＃１である場合）、位置決定部１２２は、以下の式（１７）により、表示画像におけるオブジェクト位置情報を推定する。

[0233] [数17]

$$\theta_{Ai}' = \arctan\left(\frac{r_V \cdot \tan(\theta_{Ai}) - Width_{shift}}{r_V'}\right)$$

$$\gamma_{Ai}' = \arctan\left(\frac{r_V \cdot \tan(\gamma_{Ai}) - Height_{shift}}{r_V'}\right)$$

$$r_{Ai}' = \frac{r_V'}{\cos(\theta_{Ai}') \cdot \cos(\gamma_{Ai}')} \quad \cdots (17)$$

[0234] 式（１７）において、θ_{Ai}'は、移動後のオブジェクト＃ｉの水平角度であり、θ_{Ai}は、移動前のオブジェクト＃ｉの水平角度である。γ_{A1}'は、移動後のオブジェクト＃ｉの垂直角度であり、垂直角度γ_{Ai}は、移動前のオブジェクト＃ｉの水平角度である。r_{Ai}'は、移動後のオブジェクト＃ｉの距離であり、r_{Ai}は、移動前のオブジェクト＃ｉの距離である。

[0235] 一方、オブジェクトが表示領域３１上に存在しない場合、即ち、θ_{v2}' ≦θ_{Ai}≦θ_{v1}'かつγ_{v2}' ≦γ_{Ai}≦γ_{v1}'ではない場合（図３２乃至図３５の例では、オブジェクト＃２とオブジェクト＃３である場合）、位置決定部１２２は、以下の式（１８）により、表示画像におけるオブジェクト位置情報を推定する。

[0236]

[数18]

$$\theta_{Ai}' = \arctan\left(\frac{r_{Ai}\cdot\sin(\theta_{Ai})\cdot\cos(\gamma_{Ai}) - Width_{shift}}{r_{Ai}\cdot\cos(\theta_{Ai})\cdot\cos(\gamma_{Ai}) - Length_{shift}}\right)$$

$$\gamma_{Ai}' = \arctan\left(\frac{r_{Ai}\cdot\sin(\gamma_{Ai}) - Height_{shift}}{\sqrt{(r_{Ai}\cdot\cos(\theta_{Ai})\cdot\cos(\gamma_{Ai}) - Length_{shift})^2 + (r_{Ai}\cdot\sin(\theta_{Ai})\cdot\cos(\gamma_{Ai}) - Width_{shift})^2}}\right)$$

$$r_{Ai}' = \sqrt{(r_{Ai}\cdot\cos(\theta_{Ai})\cdot\cos(\gamma_{Ai}) - Length_{shift})^2 + (r_{Ai}\cdot\sin(\theta_{Ai})\cdot\cos(\gamma_{Ai}) - Width_{shift})^2 + (r_{Ai}\cdot\tan(\gamma_{Ai}) - Height_{shift})^2}$$

・・・(18)

[0237] なお、式（１８）は、以下の式（１９）のように変形してもよい。

[0238] [数19]

$$x_{Ai}' = r_{Ai}\cdot\cos(\theta_{Ai})\cdot\cos(\gamma_{Ai}) - Length_{shift}$$

$$y_{Ai}' = r_{Ai}\cdot\sin(\theta_{Ai})\cdot\cos(\gamma_{Ai}) - Width_{shift}$$

$$z_{Ai}' = r_{Ai}\cdot\tan(\gamma_{Ai}) - Height_{shift}$$

$$\theta_{Ai}' = \arctan\left(\frac{y_{Ai}'}{x_{Ai}'}\right)$$

$$\gamma_{Ai}' = \arctan\left(\frac{z_{Ai}'}{\sqrt{x_{Ai}'^2 + y_{Ai}'^2}}\right)$$

$$r_{Ai}' = \sqrt{x_{Ai}'^2 + y_{Ai}'^2 + z_{Ai}'^2} \quad \text{・・・(19)}$$

[0239] 式（１９）において、x_{Ai}'、y_{v1}'、z_{Ai}'は、それぞれ、移動後のオブジェクト＃ｉのxyz座標系における前後方向（奥行き方向）の位置、左右方向の位置、上下方向の位置である。式（１９）は、式（１８）の変形で得られたものであり、式（１８）と式（１９）は、等価である。

[0240] ここでは、視聴者から画枠までの距離r_vと移動前のオブジェクト＃ｉの距離r_{Ai}を用いて、移動後のオブジェクト位置情報を推定したが、r_vとr_{Ai}のいずれかが与えられていない場合は、$r_v = r_{Ai}$と仮定しても良い。さらに、両方

が与えられていない場合は、$r_v = r_{Ai} = 1$と仮定しても良い。

[0241] なお、r_vとr_{Ai}の少なくとも一方が与えられていない場合には、視聴者の前後方向の移動量がr_vを超えないようにするために、式（１４）で求められたr_v'が$2*r_v$を超えないという制限を設けても良い。即ち、式（１４）で求められたr_v'を用いて、$r_v' = \min(2*r_v, r_v')$という式により、最終的なr_v'を求めるようにしてもよい。

[0242] （ストリーミング再生部の処理の説明）

本開示を適用した情報処理システムの第３実施の形態のストリーミング再生部によるストリーミング再生処理は、ステップＳ１３４の位置決定処理を除いて図２９のストリーミング再生処理と同一であるため、以下では、位置決定処理についてのみ説明する。

[0243] 図３６は、本開示を適用した情報処理システムの第３実施の形態における位置決定処理の詳細を説明するフローチャートである。

[0244] 図３６のステップＳ１８１において、位置決定部１２２は、移動前の表示領域３１の表示領域画枠サイズ情報とコンテンツ画枠サイズ情報を取得する。具体的には、位置決定部１２２は、移動前の表示領域３１の画枠の左端の水平角度θ_{v1}'および右端の水平角度θ_{v2}'、並びに、上端の垂直角度γ_{v1}'および下端の垂直角度γ_{v2}'を取得する。また、位置決定部１２２は、動画コンテンツの画枠の左端の水平角度θ_{v1}および右端の水平角度θ_{v2}、上端の垂直角度γ_{v1}および下端の垂直角度γ_{v2}'、並びに距離r_vを取得する。

[0245] ステップＳ１８２において、位置決定部１２２は、水平角度θ_{v1}'およびθ_{v2}'、垂直角度γ_{v1}'、並びに距離r_vを用いて、上述した式（１３）により、横幅$Screen_{Width}$、幅$Screen_{LeftWidth}$、および高さ$Screen_{TopHeight}$を求める。

[0246] ステップＳ１８３において、位置決定部１２２は、横幅$Screen_{Width}$並びに水平角度θ_{v1}およびθ_{v2}を用いて、上述した式（１４）により、移動後の距離r_v'を推定する。

[0247] ステップＳ１８４において、位置決定部１２２は、幅$Screen_{LeftWidth}$、高さ$Screen_{TopHeigh}$、距離r_vおよびr_v'、水平角度$\theta_{v1}, \theta_{v2}, \theta_{v1}'$、および$\theta_{v2}'$、並

びに垂直角度γ_{v1}, γ_{v2}, γ_{v1}'、およびγ_{v2}'を用いて、上述した式（１５）および（１６）により、視聴者の移動量$Length_{shift}$, $Width_{shift}$、および$Height_{shift}$を推定する。

[0248] ステップＳ１８５において、位置決定部１２２は、再生対象の最初のオブジェクト＃ｉを選択する。

[0249] ステップＳ１８６において、位置決定部１２２は、選択されたオブジェクト＃ｉの移動前のオブジェクト位置情報である水平角度θ_{Ai}、垂直角度γ_{Ai}、および距離r_{Ai}を取得する。

[0250] ステップＳ１８７において、位置決定部１２２は、オブジェクト＃ｉが表示領域３１上に存在するかどうか、即ち$\theta_{v2}' \leqq \theta_{Ai} \leqq \theta_{v1}'$かつ$\gamma_{v2}' \leqq \gamma_{Ai} \leqq \gamma_{v1}'$であるかどうかを判定する。

[0251] ステップＳ１８７でオブジェクト＃ｉが表示領域３１上に存在すると判定された場合、即ち$\theta_{v2}' \leqq \theta_{Ai} \leqq \theta_{v1}'$かつ$\gamma_{v2}' \leqq \gamma_{Ai} \leqq \gamma_{v1}'$である場合、処理はステップＳ１８８に進む。

[0252] ステップＳ１８８において、位置決定部１２２は、移動量$Width_{shif}$および$Height_{shift}$、距離r_vおよびr_v'、水平角度θ_{Ai}およびθ_{Ai}'、並びに垂直角度γ_{Ai}およびγ_{Ai}'を用いて、上述した式（１７）により、表示画像におけるオブジェクト＃ｉのオブジェクト位置情報を推定する。そして、処理はステップＳ１９０に進む。

[0253] 一方、ステップＳ１８７でオブジェクト＃ｉが表示領域３１上に存在しないと判定された場合、即ち、$\theta_{v2}' \leqq \theta_{Ai} \leqq \theta_{v1}'$かつ$\gamma_{v2}' \leqq \gamma_{Ai} \leqq \gamma_{v1}'$ではない場合、処理はステップＳ１８９に進む。

[0254] ステップＳ１８９において、位置決定部１２２は、移動量$Length_{shift}$, $Width_{shift}$、および$Height_{shift}$、距離r_{Ai}、水平角度θ_{Ai}およびθ_{Ai}'、並びに垂直角度γ_{Ai}およびγ_{Ai}'を用いて、上述した式（１８）により、表示画像におけるオブジェクト＃ｉのオブジェクト位置情報を推定する。そして、処理はステップＳ１９０に進む。

[0255] ステップＳ１９０において、位置決定部１２２は、再生対象のすべてのオ

ブジェクトのオブジェクト位置情報が推定されたかどうかを判定する。ステップS１９０ですべてのオブジェクトのオブジェクト位置情報が推定されていないと判定された場合、処理はステップＳ１９１に進む。

[0256] ステップＳ１９１において、位置決定部１２２は、再生対象のオブジェクトのうちの、現在選択されているオブジェクトの次のオブジェクトを選択し、処理をステップS１８６に戻す。

[0257] 一方、ステップＳ１９０ですべてのオブジェクトのオブジェクト位置情報が推定されたと判定された場合、処理は、Ｓ１９２に進む。

[0258] ステップＳ１９２において、位置決定部１２２は、表示画像における再生対象のすべてのオブジェクトのオブジェクト位置情報を音声合成処理部１２３に出力する。ステップＳ１９２の処理後、処理は、図２９のステップＳ１３４に戻り、ステップＳ１３５に進む。

[0259] なお、第３実施の形態では、表示画像のサイズが動画コンテンツのサイズと同一であるようにしたが、異なっていてもよい。

[0260] また、第３実施の形態において、再生対象の全てのオブジェクトの音声データが合成されて出力されるのではなく、一部のオブジェクト(例えば、表示領域内のオブジェクト、表示領域から所定の範囲内のオブジェクト等)の音声データのみが合成されて出力されるようにしてもよい。出力する音声データのオブジェクトの選択方法は、予め決定されていてもよいし、ユーザにより指定されるようにしてもよい。

[0261] ＜第４実施の形態＞

（本開示を適用したコンピュータの説明）

上述したWebサーバ１２の一連の処理は、ハードウエアにより実行することもできるし、ソフトウエアにより実行することもできる。一連の処理をソフトウエアにより実行する場合には、そのソフトウエアを構成するプログラムが、コンピュータにインストールされる。ここで、コンピュータには、専用のハードウエアに組み込まれているコンピュータや、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用の

パーソナルコンピュータなどが含まれる。

[0262] 図３７は、上述したWebサーバ１２の一連の処理をプログラムにより実行するコンピュータのハードウエアの構成例を示すブロック図である。

[0263] コンピュータにおいて、CPU (Central Processing Unit) ２０１, ROM (Read Only Memory) ２０２, RAM (Random Access Memory) ２０３は、バス２０４により相互に接続されている。

[0264] バス２０４には、さらに、入出力インタフェース２０５が接続されている。入出力インタフェース２０５には、入力部２０６、出力部２０７、記憶部２０８、通信部２０９、及びドライブ２１０が接続されている。

[0265] 入力部２０６は、キーボード、マウス、マイクロフォンなどよりなる。出力部２０７は、ディスプレイ、スピーカなどよりなる。記憶部２０８は、ハードディスクや不揮発性のメモリなどよりなる。通信部２０９は、ネットワークインタフェースなどよりなる。ドライブ２１０は、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリなどのリムーバブルメディア２１１を駆動する。

[0266] 以上のように構成されるコンピュータでは、CPU２０１が、例えば、記憶部２０８に記憶されているプログラムを、入出力インタフェース２０５及びバス２０４を介して、RAM２０３にロードして実行することにより、上述した一連の処理が行われる。

[0267] コンピュータ（CPU２０１）が実行するプログラムは、例えば、パッケージメディア等としてのリムーバブルメディア２１１に記録して提供することができる。また、プログラムは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提供することができる。

[0268] コンピュータでは、プログラムは、リムーバブルメディア２１１をドライブ２１０に装着することにより、入出力インタフェース２０５を介して、記憶部２０８にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部２０９で受信し、記憶部２０８にイン

ストールすることができる。その他、プログラムは、ROM２０２や記憶部２０８に、あらかじめインストールしておくことができる。

[0269] なお、コンピュータが実行するプログラムは、本明細書で説明する順序に沿って時系列に処理が行われるプログラムであっても良いし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであっても良い。

[0270] また、動画再生端末１４のハードウエア構成は、図３７のコンピュータと同様の構成にすることができる。この場合、例えば、CPU２０１が、制御用ソフトウエア２１、動画再生ソフトウエア２２、およびアクセス用ソフトウエア２３を実行する。動画再生端末１４の処理は、ハードウエアにより実行することもできる。

[0271] 本明細書において、システムとは、複数の構成要素（装置、モジュール（部品）等）の集合を意味し、すべての構成要素が同一筐体中にあるか否かは問わない。したがって、別個の筐体に収納され、ネットワークを介して接続されている複数の装置、及び、１つの筐体の中に複数のモジュールが収納されている１つの装置は、いずれも、システムである。

[0272] なお、本開示の実施の形態は、上述した実施の形態に限定されるものではなく、本開示の要旨を逸脱しない範囲において種々の変更が可能である。

[0273] 例えば、ファイル生成装置１１は、タイル単位で画像ファイルを生成するのではなく、全てのタイルの符号化データを多重化してビデオストリームを生成し、１つの画像ファイルを生成するようにしてもよい。

[0274] また、表示領域３１に対応する音声は、表示領域３１内のオブジェクトの音声だけでなく、表示領域３１外の固定されたオブジェクトの音声も含むようにしてもよい。

[0275] さらに、音声データには、オブジェクト単位の音声データのほか、チャネルオーディオの音声データ、HOA (Higher－Order Ambisonics) オーディオの音声データ、SAOC(Spatial Audio Object Coding)の音声データ、音声データのメタデータ(scene情報、動的・静的metadata)などが含まれるようにしても

よい。この場合、例えば、サブサンプルとして、各オブジェクトの符号化データだけでなく、これらのデータの符号化データも配置される。

[0276] さらに、本開示は、以下のような構成もとることができる。

[0277] （１）

画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを伝送する伝送部

を備える情報処理装置。

（２）

前記画枠サイズ情報は、前記画像データの水平方向および垂直方向の画角、並びに、前記画角の基点と画像平面との距離を示す情報である

ように構成された

前記（１）に記載の情報処理装置。

（３）

前記音声位置情報は、前記音声データを取得した位置と基点とを結ぶ線の水平方向および垂直方向の角度、並びに、前記音声データを取得した位置と前記基点との距離を示す情報である

ように構成された

前記（１）または（２）に記載の情報処理装置。

（４）

前記音声位置情報を含む前記音声データのメタデータファイルを生成するメタファイル生成部と、

前記画枠サイズ情報と前記メタデータファイルを特定する情報を制御情報として生成する情報生成部と、

をさらに備え、

前記伝送部は、前記情報生成部により生成された前記制御情報と前記メタファイル生成部により生成された前記メタデータファイルを伝送する

ように構成された

前記（１）乃至（３）のいずれかに記載の情報処理装置。

（５）

前記伝送部は、前記画枠サイズ情報と前記音声位置情報に基づいて要求される、前記画像データに対応する画像のうちの表示する領域である表示領域に対応する前記音声データを伝送する

ように構成された

前記（１）乃至（４）のいずれかに記載の情報処理装置。

（６）

前記音声データは、オブジェクト単位の音声データであり、

前記伝送部は、前記表示領域に対応する前記オブジェクトの音声データを伝送する

ように構成された

前記（５）に記載の情報処理装置。

（７）

複数の前記オブジェクトの音声データと、各オブジェクトの前記音声データの音声ファイル内の位置を表すデータ位置情報とを含む前記音声ファイルを生成する音声ファイル生成部

をさらに備え、

前記伝送部は、前記音声ファイル生成部により生成された前記音声ファイル内の前記表示領域に対応する前記オブジェクトの音声データを伝送する

ように構成された

前記（６）に記載の情報処理装置。

（８）

複数の前記オブジェクトの音声データのそれぞれを含む音声ファイルを生成する音声ファイル生成部

をさらに備え、

前記伝送部は、各オブジェクトの前記音声ファイルを特定するファイル特定情報を伝送し、前記画枠サイズ情報、前記音声位置情報、および前記ファイル特定情報に基づいて要求される前記表示領域に対応する前記オブジェク

トの音声データを含む前記音声ファイルを伝送する

ように構成された

前記（６）に記載の情報処理装置。

（９）

情報処理装置が、

画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを伝送する伝送ステップ

を含む情報処理方法。

（１０）

画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを受け取る受け取り部と、

前記受け取り部により受け取られた前記画像データの前記画枠サイズ情報と前記音声位置情報に基づいて、前記画像データに対応する画像上の前記音声データの取得位置を決定する位置決定部と

を備える情報処理装置。

（１１）

前記画枠サイズ情報は、前記画像データの水平方向および垂直方向の画角、並びに、前記画角の基点と画像平面との距離を示す情報である

ように構成された

前記（１０）に記載の情報処理装置。

（１２）

前記音声位置情報は、前記音声データを取得した位置と基点とを結ぶ線の水平方向および垂直方向の角度、並びに、前記音声データを取得した位置と前記基点との距離を示す情報である

ように構成された

前記（１０）または（１１）に記載の情報処理装置。

（１３）

前記音声データは、オブジェクト単位の音声データであり、

前記位置決定部は、前記画枠サイズ情報と各オブジェクトの前記音声位置情報に基づいて、前記画像データに対応する画像上の各オブジェクトの前記音声データの取得位置を決定する

ように構成された

前記（１０）乃至（１２）のいずれかに記載の情報処理装置。

（１４）

前記位置決定部により決定された各オブジェクトの前記音声データの取得位置に基づいて、前記画像データに対応する画像のうちの表示する領域である表示領域に対応する前記オブジェクトの音声データを選択する音声選択部

をさらに備え、

前記受け取り部は、前記音声選択部により選択された前記音声データを受け取る

ように構成された

前記（１３）に記載の情報処理装置。

（１５）

前記受け取り部は、複数の前記オブジェクトの音声データと、各オブジェクトの前記音声データの音声ファイル内の位置を表すデータ位置情報とを含む前記音声ファイル内の、前記音声選択部により選択された前記オブジェクトの音声データを受け取る

ように構成された

前記（１４）に記載の情報処理装置。

（１６）

前記受け取り部は、複数の前記オブジェクトの音声データのそれぞれを含む音声ファイルのうちの、前記音声選択部により選択された前記オブジェクトの音声データを含む前記音声ファイルを受け取る

ように構成された

前記（１４）に記載の情報処理装置。

（１７）

前記位置決定部により決定された各オブジェクトの前記音声データの取得位置に基づいて、複数の前記オブジェクトの音声データを合成する音声合成処理部

をさらに備える

前記（１３）に記載の情報処理装置。

（１８）

前記画像データに対応する画像のうちの表示する領域である表示領域の画像データのサイズを、前記画像データのサイズに変換する変換部

をさらに備え、

前記位置決定部は、前記画像データの前記画枠サイズ情報、各オブジェクトの前記音声位置情報、および前記表示領域の画枠サイズ情報に基づいて、前記変換部によりサイズが変換された前記表示領域の画像データに対応する画像上の各オブジェクトの前記音声データの取得位置を決定する

ように構成された

前記（１７）に記載の情報処理装置。

（１９）

前記表示領域の画枠サイズ情報は、前記表示領域の水平方向および垂直方向の画角、並びに、前記画角の基点と画像平面との距離を示す情報である

ように構成された

前記（１８）に記載の情報処理装置。

（２０）

画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを受け取る受け取りステップと、

前記受け取りステップの処理により受け取られた前記画像データの前記画枠サイズ情報と前記音声位置情報に基づいて、前記画像データに対応する画像上の前記音声データの取得位置を決定する位置決定ステップと

を含む情報処理方法。

符号の説明

[0278] １１　ファイル生成装置，　１２　Webサーバ，　１４　動画再生端末，５６　音声ファイル生成部，　５７　MPD生成部，　９１　MPD取得部，　９４　音声選択部，　９５　音声ファイル取得部，　１２２　位置決定部，１２３　音声合成処理部，　１２４　画像合成処理部

請求の範囲

[請求項1] 画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを伝送する伝送部

を備える情報処理装置。

[請求項2] 前記画枠サイズ情報は、前記画像データの水平方向および垂直方向の画角、並びに、前記画角の基点と画像平面との距離を示す情報である

ように構成された

請求項１に記載の情報処理装置。

[請求項3] 前記音声位置情報は、前記音声データを取得した位置と基点とを結ぶ線の水平方向および垂直方向の角度、並びに、前記音声データを取得した位置と前記基点との距離を示す情報である

ように構成された

請求項１に記載の情報処理装置。

[請求項4] 前記音声位置情報を含む前記音声データのメタデータファイルを生成するメタファイル生成部と、

前記画枠サイズ情報と前記メタデータファイルを特定する情報を制御情報として生成する情報生成部と、

をさらに備え、

前記伝送部は、前記情報生成部により生成された前記制御情報と前記メタファイル生成部により生成された前記メタデータファイルを伝送する

ように構成された

請求項１に記載の情報処理装置。

[請求項5] 前記伝送部は、前記画枠サイズ情報と前記音声位置情報に基づいて要求される、前記画像データに対応する画像のうちの表示する領域である表示領域に対応する前記音声データを伝送する

ように構成された

請求項１に記載の情報処理装置。

[請求項6] 前記音声データは、オブジェクト単位の音声データであり、

前記伝送部は、前記表示領域に対応する前記オブジェクトの音声データを伝送する

ように構成された

請求項５に記載の情報処理装置。

[請求項7] 複数の前記オブジェクトの音声データと、各オブジェクトの前記音声データの音声ファイル内の位置を表すデータ位置情報とを含む前記音声ファイルを生成する音声ファイル生成部

をさらに備え、

前記伝送部は、前記音声ファイル生成部により生成された前記音声ファイル内の前記表示領域に対応する前記オブジェクトの音声データを伝送する

ように構成された

請求項６に記載の情報処理装置。

[請求項8] 複数の前記オブジェクトの音声データのそれぞれを含む音声ファイルを生成する音声ファイル生成部

をさらに備え、

前記伝送部は、各オブジェクトの前記音声ファイルを特定するファイル特定情報を伝送し、前記画枠サイズ情報、前記音声位置情報、および前記ファイル特定情報に基づいて要求される前記表示領域に対応する前記オブジェクトの音声データを含む前記音声ファイルを伝送する

ように構成された

請求項６に記載の情報処理装置。

[請求項9] 情報処理装置が、

画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを伝送する伝送ステップ

を含む情報処理方法。

[請求項10] 画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを受け取る受け取り部と、

前記受け取り部により受け取られた前記画像データの前記画枠サイズ情報と前記音声位置情報に基づいて、前記画像データに対応する画像上の前記音声データの取得位置を決定する位置決定部と

を備える情報処理装置。

[請求項11] 前記画枠サイズ情報は、前記画像データの水平方向および垂直方向の画角、並びに、前記画角の基点と画像平面との距離を示す情報である

ように構成された

請求項１０に記載の情報処理装置。

[請求項12] 前記音声位置情報は、前記音声データを取得した位置と基点とを結ぶ線の水平方向および垂直方向の角度、並びに、前記音声データを取得した位置と前記基点との距離を示す情報である

ように構成された

請求項１０に記載の情報処理装置。

[請求項13] 前記音声データは、オブジェクト単位の音声データであり、

前記位置決定部は、前記画枠サイズ情報と各オブジェクトの前記音声位置情報に基づいて、前記画像データに対応する画像上の各オブジェクトの前記音声データの取得位置を決定する

ように構成された

請求項１０に記載の情報処理装置。

[請求項14] 前記位置決定部により決定された各オブジェクトの前記音声データの取得位置に基づいて、前記画像データに対応する画像のうちの表示する領域である表示領域に対応する前記オブジェクトの音声データを選択する音声選択部

をさらに備え、

前記受け取り部は、前記音声選択部により選択された前記音声データを受け取る

ように構成された

請求項１３に記載の情報処理装置。

[請求項15] 前記受け取り部は、複数の前記オブジェクトの音声データと、各オブジェクトの前記音声データの音声ファイル内の位置を表すデータ位置情報とを含む前記音声ファイル内の、前記音声選択部により選択された前記オブジェクトの音声データを受け取る

ように構成された

請求項１４に記載の情報処理装置。

[請求項16] 前記受け取り部は、複数の前記オブジェクトの音声データのそれぞれを含む音声ファイルのうちの、前記音声選択部により選択された前記オブジェクトの音声データを含む前記音声ファイルを受け取る

ように構成された

請求項１４に記載の情報処理装置。

[請求項17] 前記位置決定部により決定された各オブジェクトの前記音声データの取得位置に基づいて、複数の前記オブジェクトの音声データを合成する音声合成処理部

をさらに備える

請求項１３に記載の情報処理装置。

[請求項18] 前記画像データに対応する画像のうちの表示する領域である表示領域の画像データのサイズを、前記画像データのサイズに変換する変換部

をさらに備え、

前記位置決定部は、前記画像データの前記画枠サイズ情報、各オブジェクトの前記音声位置情報、および前記表示領域の画枠サイズ情報に基づいて、前記変換部によりサイズが変換された前記表示領域の画像データに対応する画像上の各オブジェクトの前記音声データの取得

位置を決定する

ように構成された

請求項１７に記載の情報処理装置。

[請求項19] 前記表示領域の画枠サイズ情報は、前記表示領域の水平方向および垂直方向の画角、並びに、前記画角の基点と画像平面との距離を示す情報である

ように構成された

請求項１８に記載の情報処理装置。

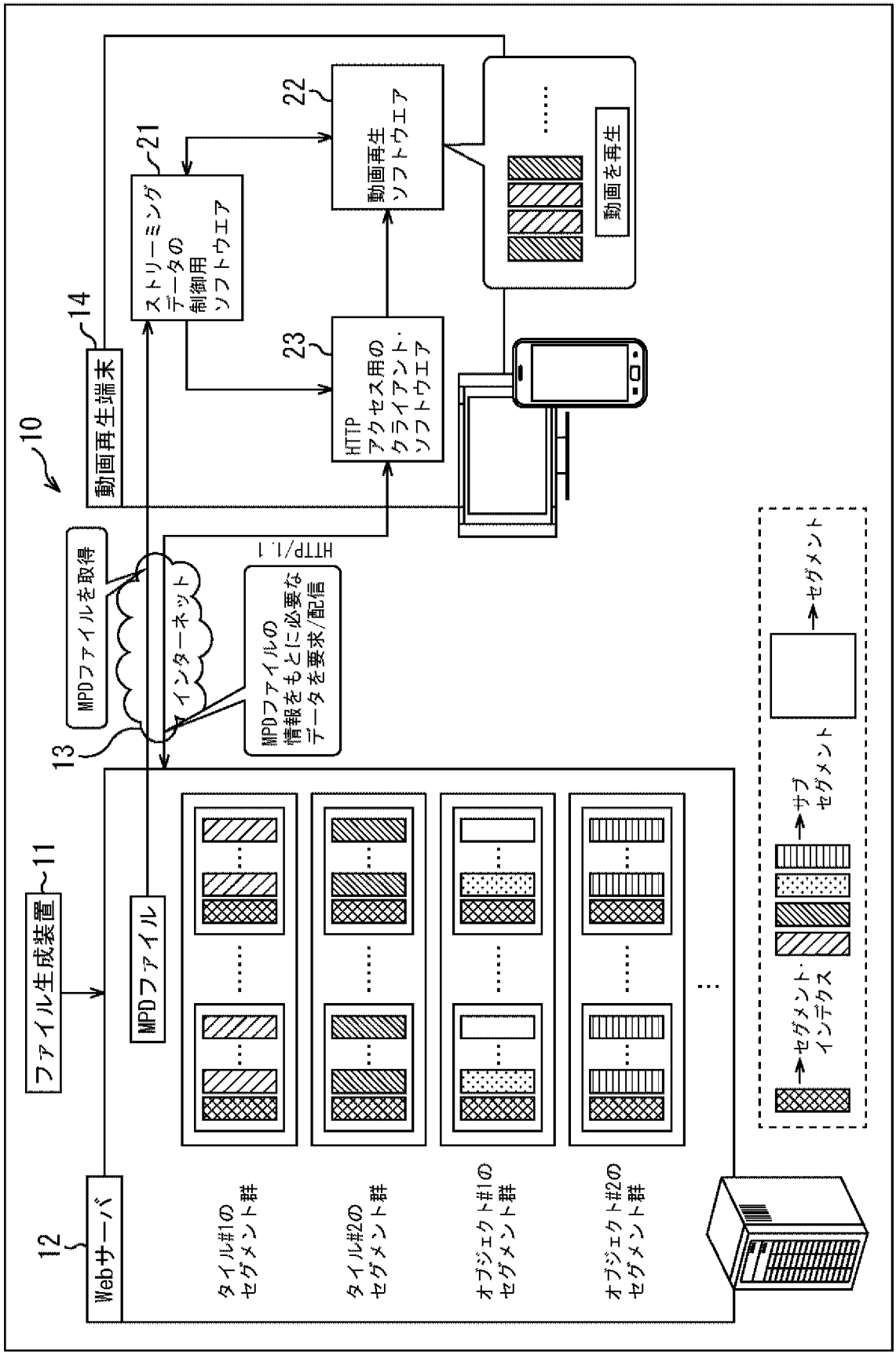
[請求項20] 画像データの画枠サイズを示す画枠サイズ情報と、音声データの取得位置を示す音声位置情報とを受け取る受け取りステップと、

前記受け取りステップの処理により受け取られた前記画像データの前記画枠サイズ情報と前記音声位置情報に基づいて、前記画像データに対応する画像上の前記音声データの取得位置を決定する位置決定ステップと

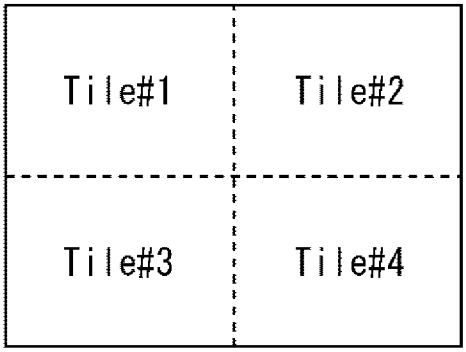
を含む情報処理方法。

[図1]

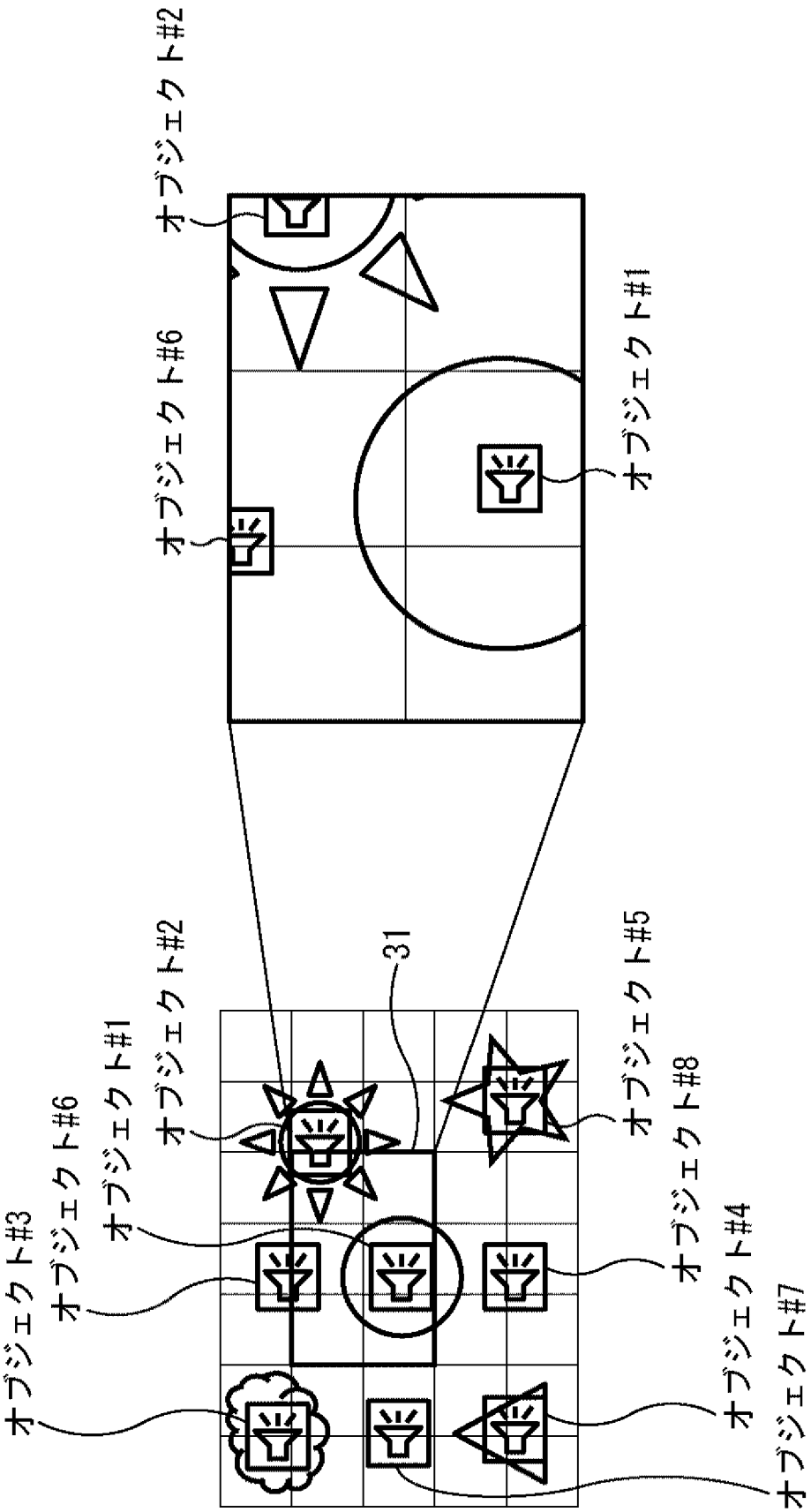
図1


10
動画再生端末
14
ストリーミング
データの
制御用
ソフトウエア
21
22
動画再生
ソフトウエア
23
HTTP
アクセス用の
クライアント・
ソフトウエア
動画を再生
HTTP/1.1
MPDファイルを取得
インターネット
MPDファイルの
情報をもとに必要な
データを要求/配信
13
ファイル生成装置
11
MPDファイル
12
Webサーバ
タイル#1の
セグメント群
タイル#2の
セグメント群
オブジェクト#1の
セグメント群
オブジェクト#2の
セグメント群
セグメント
サブ
セグメント
セグメント・
インデクス

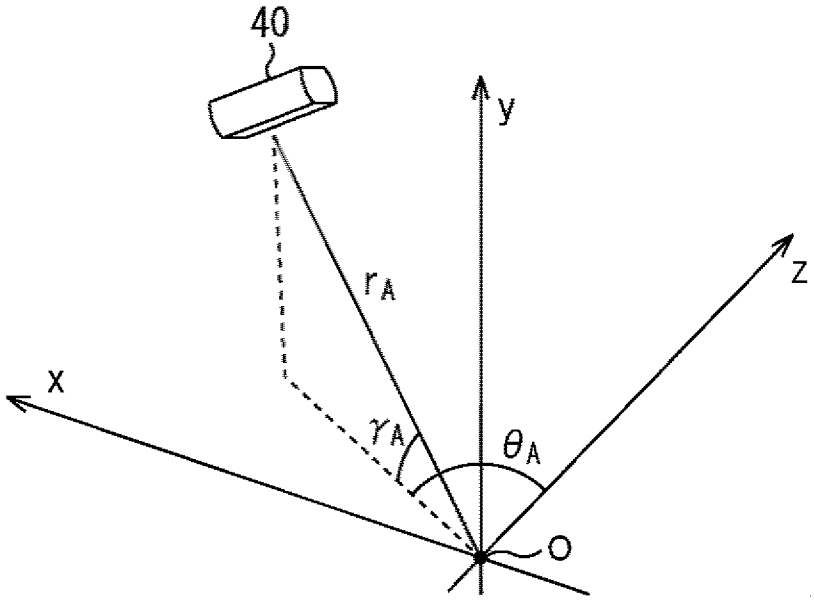
[図2]
図2



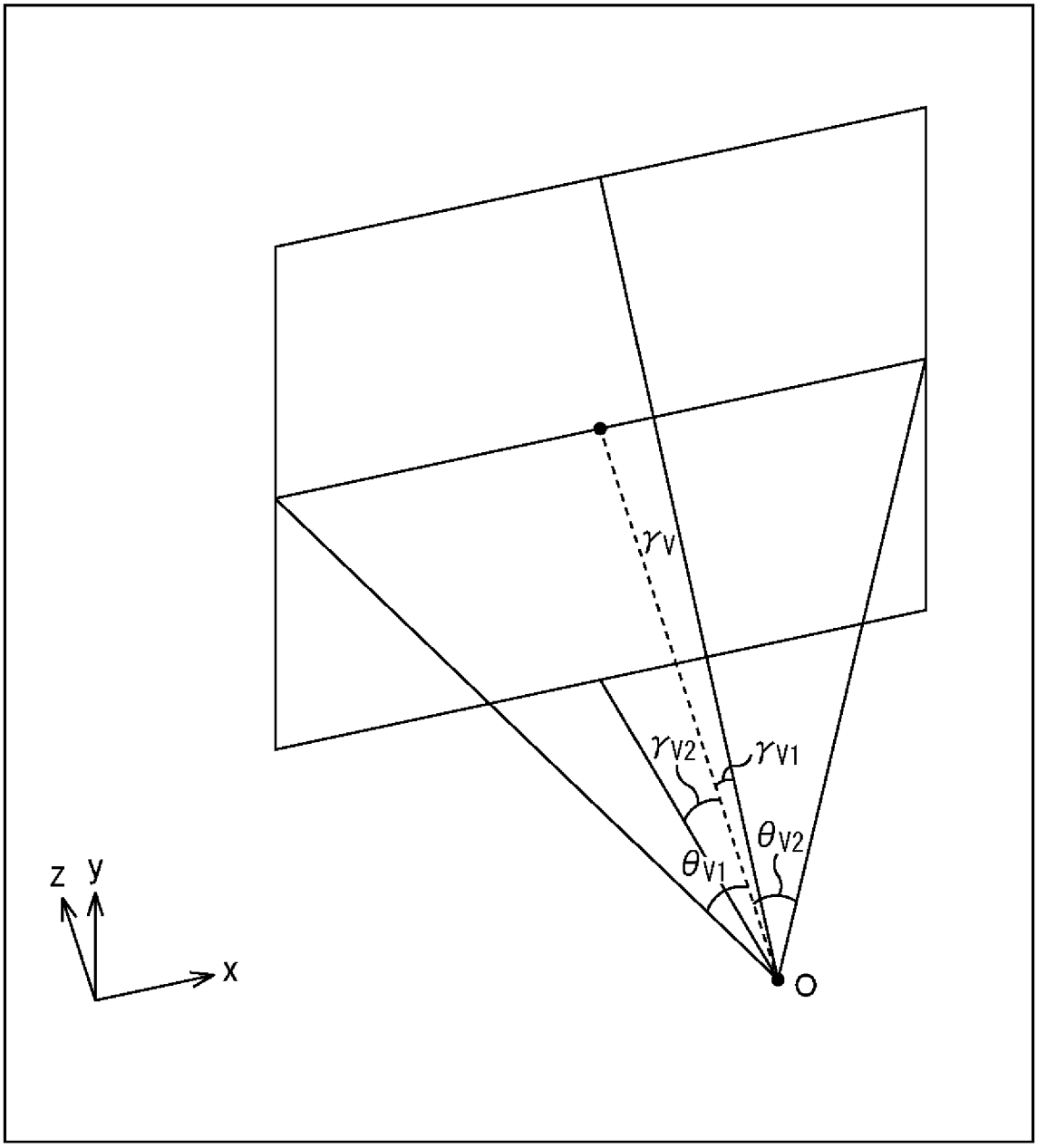
[図3]
図3


オブジェクト#2
オブジェクト#6
オブジェクト#1
31
オブジェクト#2
オブジェクト#1
オブジェクト#6
オブジェクト#3
オブジェクト#5
オブジェクト#8
オブジェクト#4
オブジェクト#7

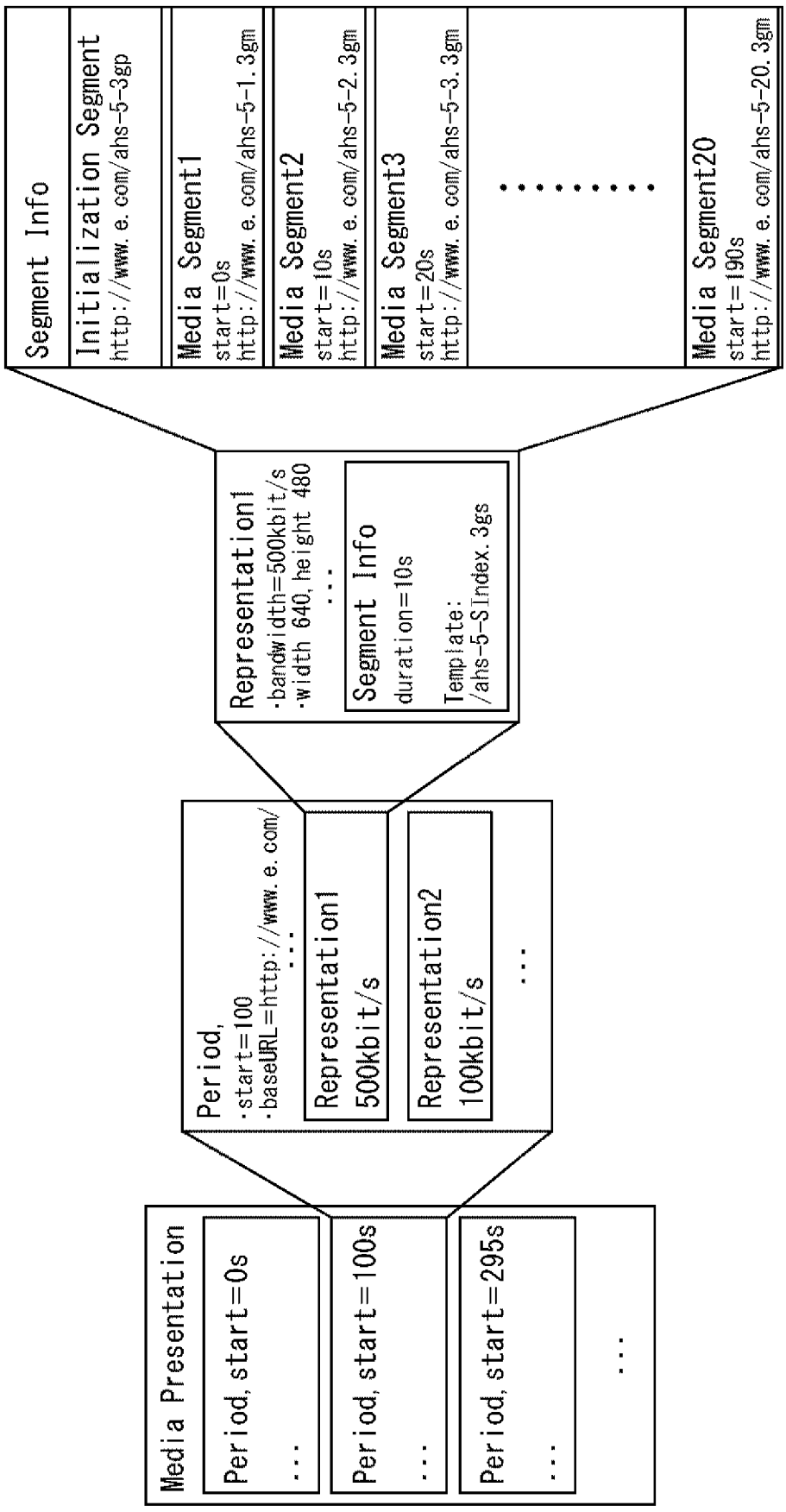
[図4]
図4



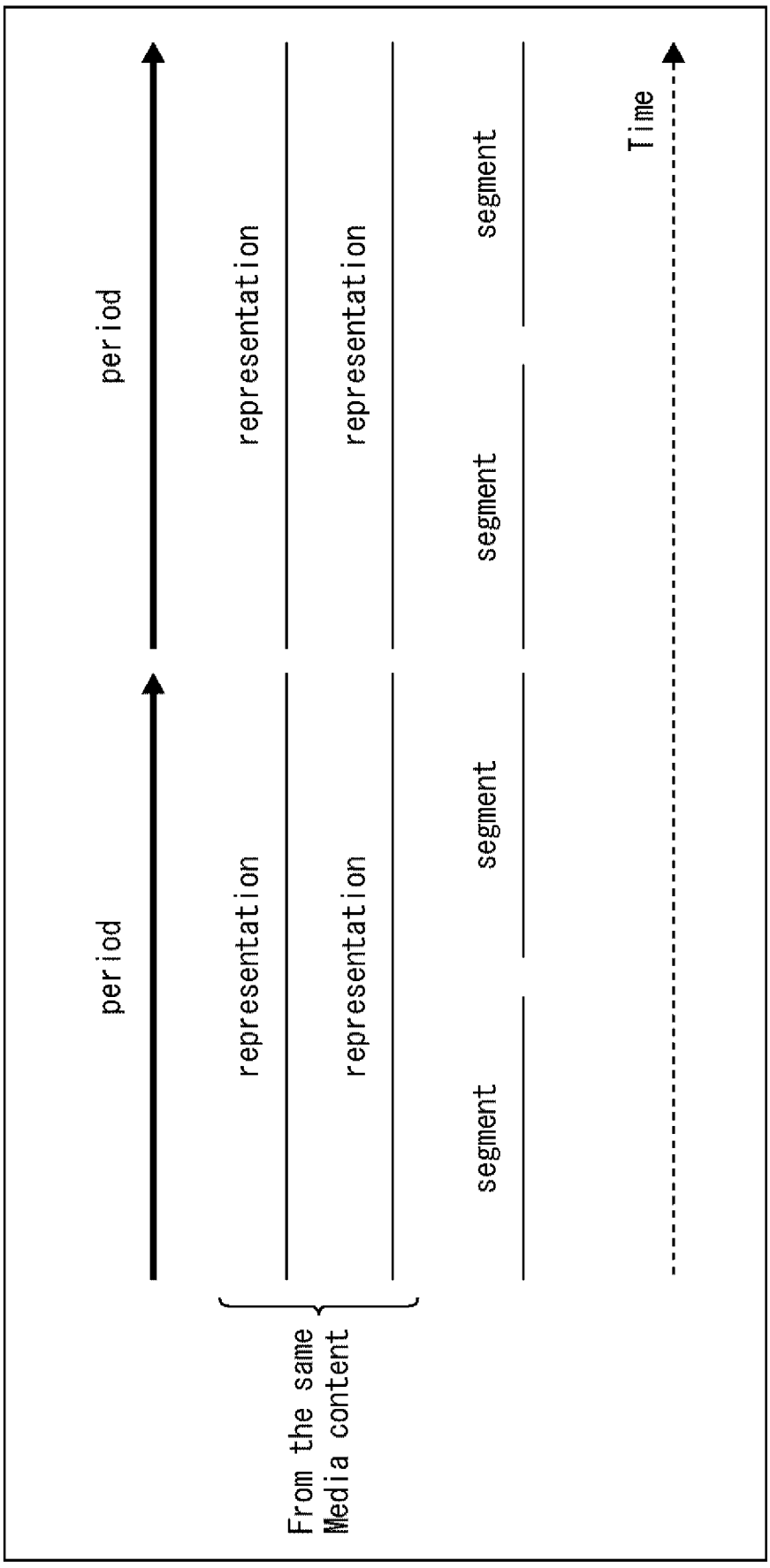
[図5]
図5



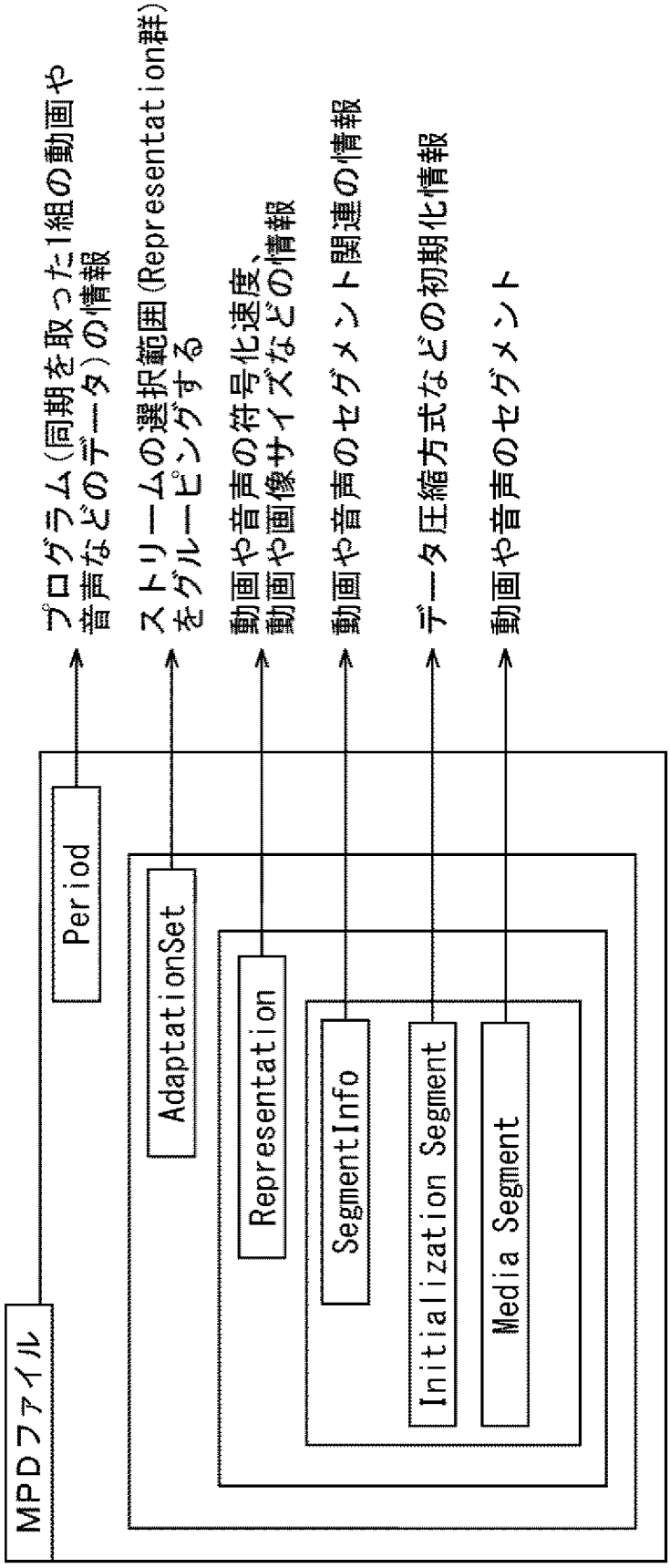
[図6]
図6


Media Presentation
Period, start=0s
...
Period, start=100s
...
Period, start=295s
...
...
Period,
·start=100
·baseURL=http://www.e.com/
...
Representation1
500kbit/s
Representation2
100kbit/s
...
Representation1
·bandwidth=500kbit/s
·width 640, height 480
...
Segment Info
duration=10s
Template:
/ahs-5-SIndex.3gs
Segment Info
Initialization Segment
http://www.e.com/ahs-5-3gp
Media Segment1
start=0s
http://www.e.com/ahs-5-1.3gm
Media Segment2
start=10s
http://www.e.com/ahs-5-2.3gm
Media Segment3
start=20s
http://www.e.com/ahs-5-3.3gm
Media Segment20
start=190s
http://www.e.com/ahs-5-20.3gm

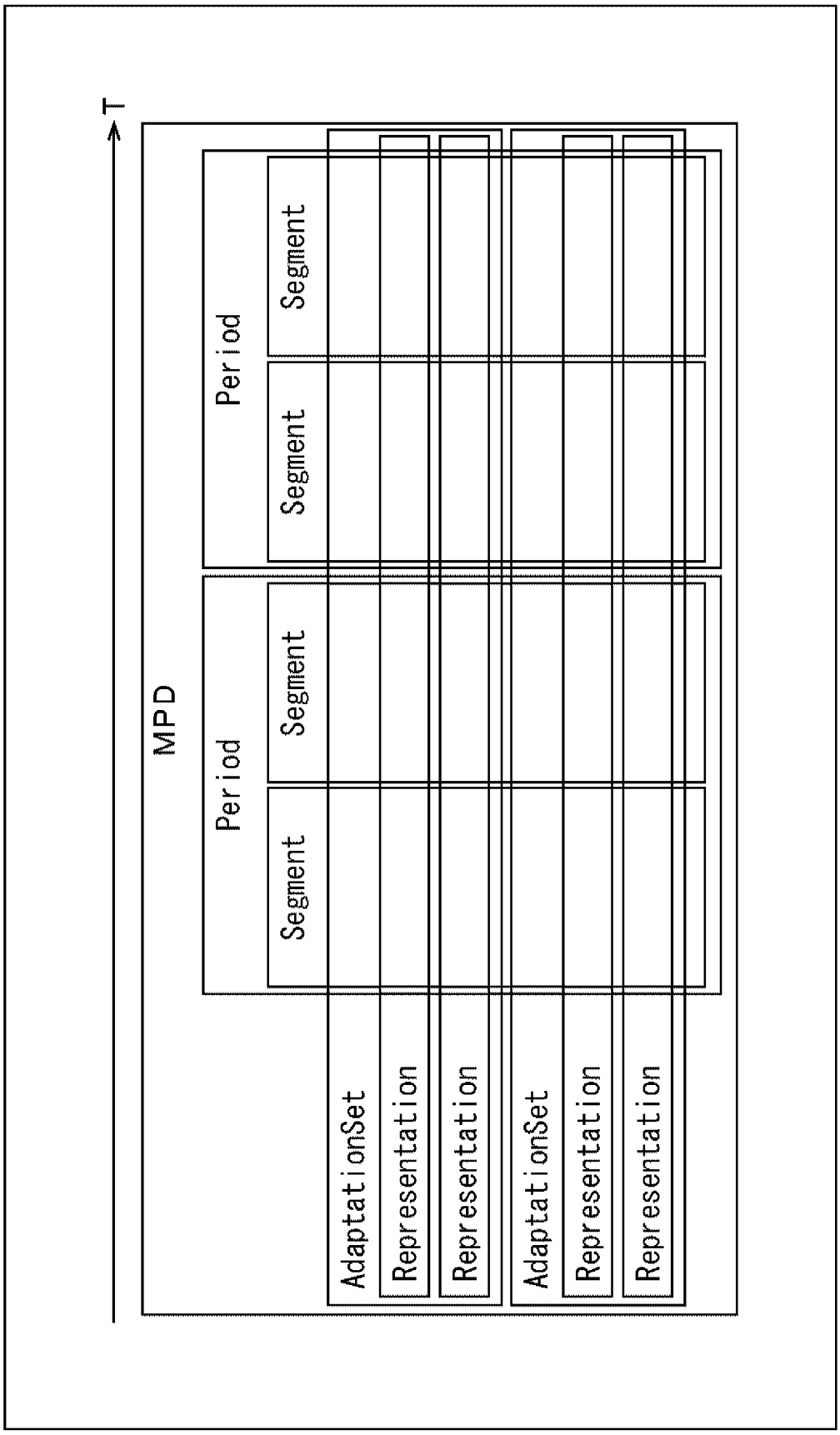
[図7]
図7


period
period
representation
representation
representation
representation
segment
segment
segment
segment
Time
From the same
Media content

[図8]
図8

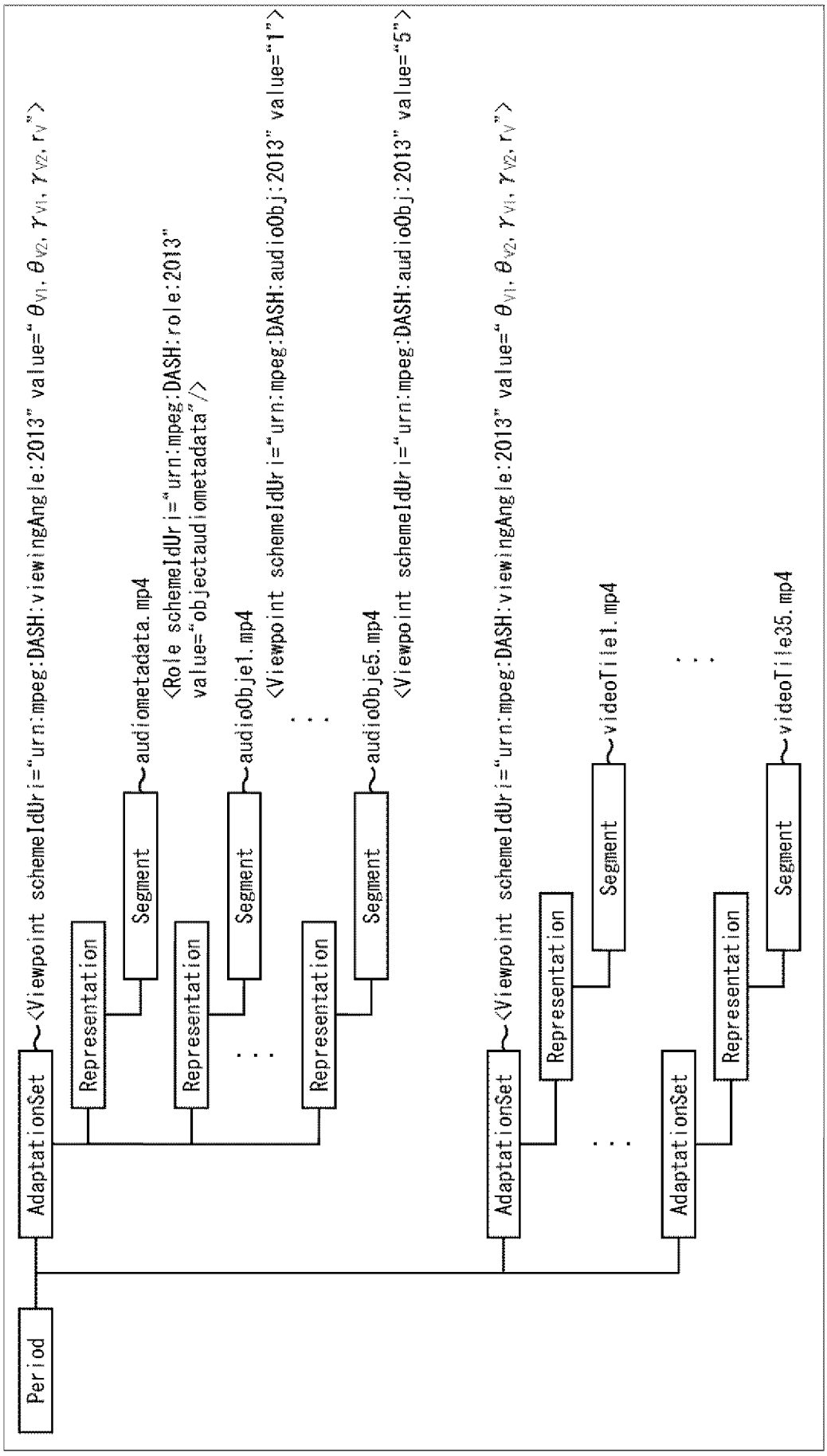

MPDファイル
Period
AdaptationSet
Representation
SegmentInfo
Initialization Segment
Media Segment
プログラム(同期を取った1組の動画や音声などのデータ)の情報
ストリームの選択範囲(Representation群)をグルーピングする
動画や音声の符号化速度、動画や画像サイズなどの情報
動画や音声のセグメント関連の情報
データ圧縮方式などの初期化情報
動画や音声のセグメント

[図9]
図9

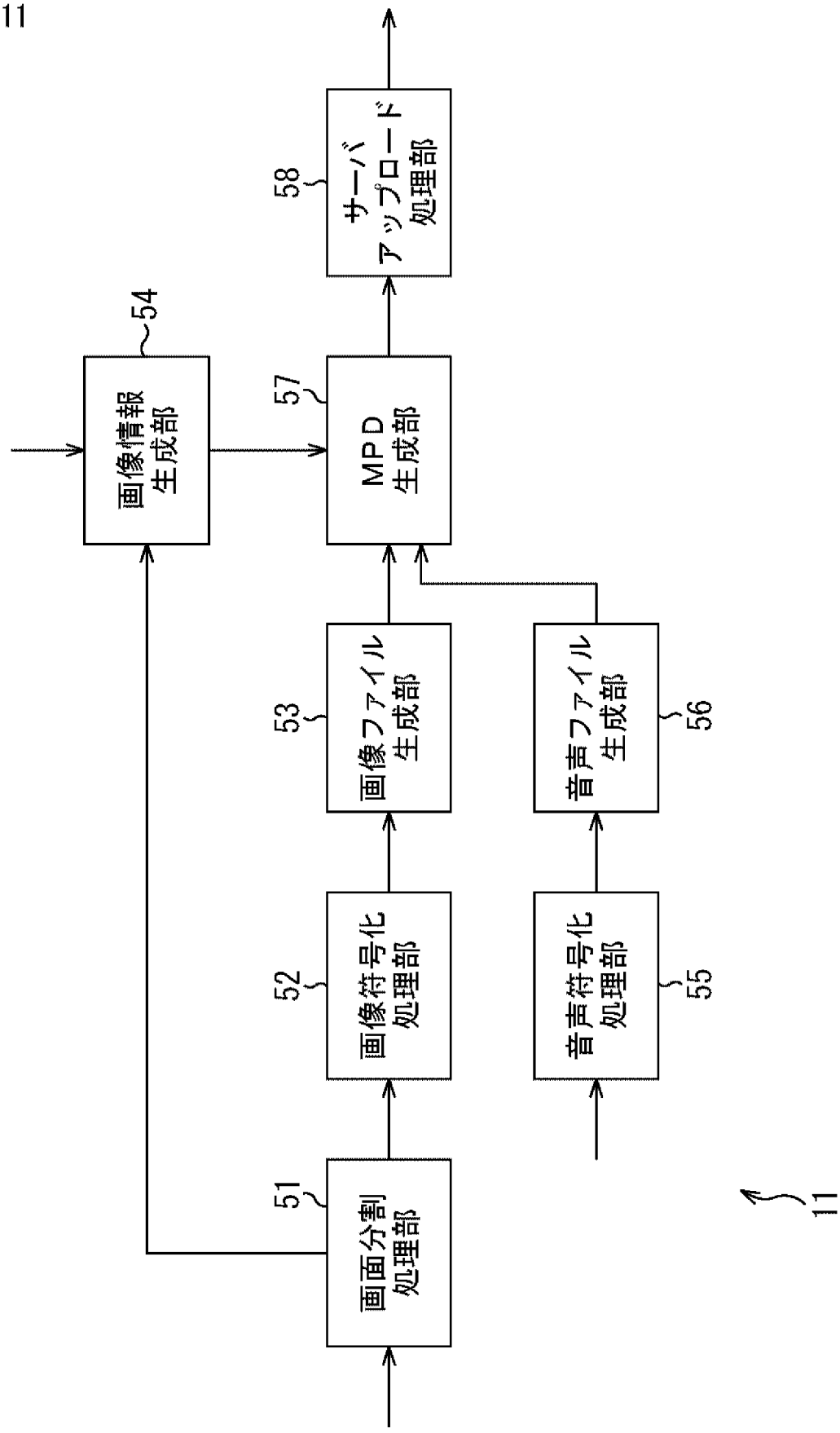

T
MPD
Period
Segment
Segment
Segment
Period
Segment
Segment
AdaptationSet
Representation
Representation
AdaptationSet
Representation
Representation

[図10]

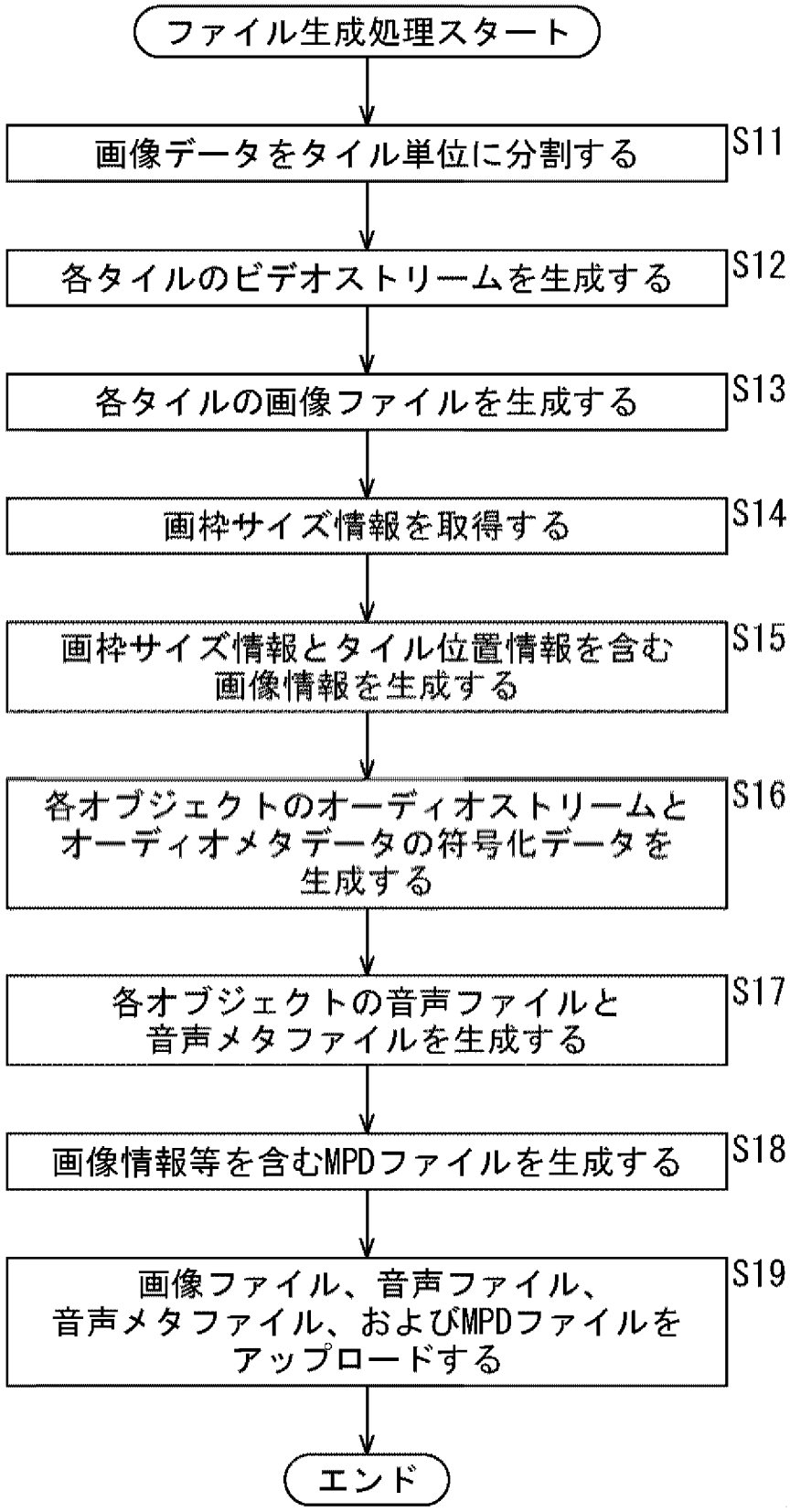
図10


Period
AdaptationSet
<Viewpoint schemeIdUri="urn:mpeg:DASH:viewingAngle:2013" value="θV1, θV2, γV1, γV2, rV">
Representation
Segment
audiometadata.mp4
<Role schemeIdUri="urn:mpeg:DASH:role:2013"
value="objectaudiometadata"/>
audioObje1.mp4
<Viewpoint schemeIdUri="urn:mpeg:DASH:audioObj:2013" value="1">
audioObje5.mp4
<Viewpoint schemeIdUri="urn:mpeg:DASH:audioObj:2013" value="5">
videoTile1.mp4
videoTile35.mp4

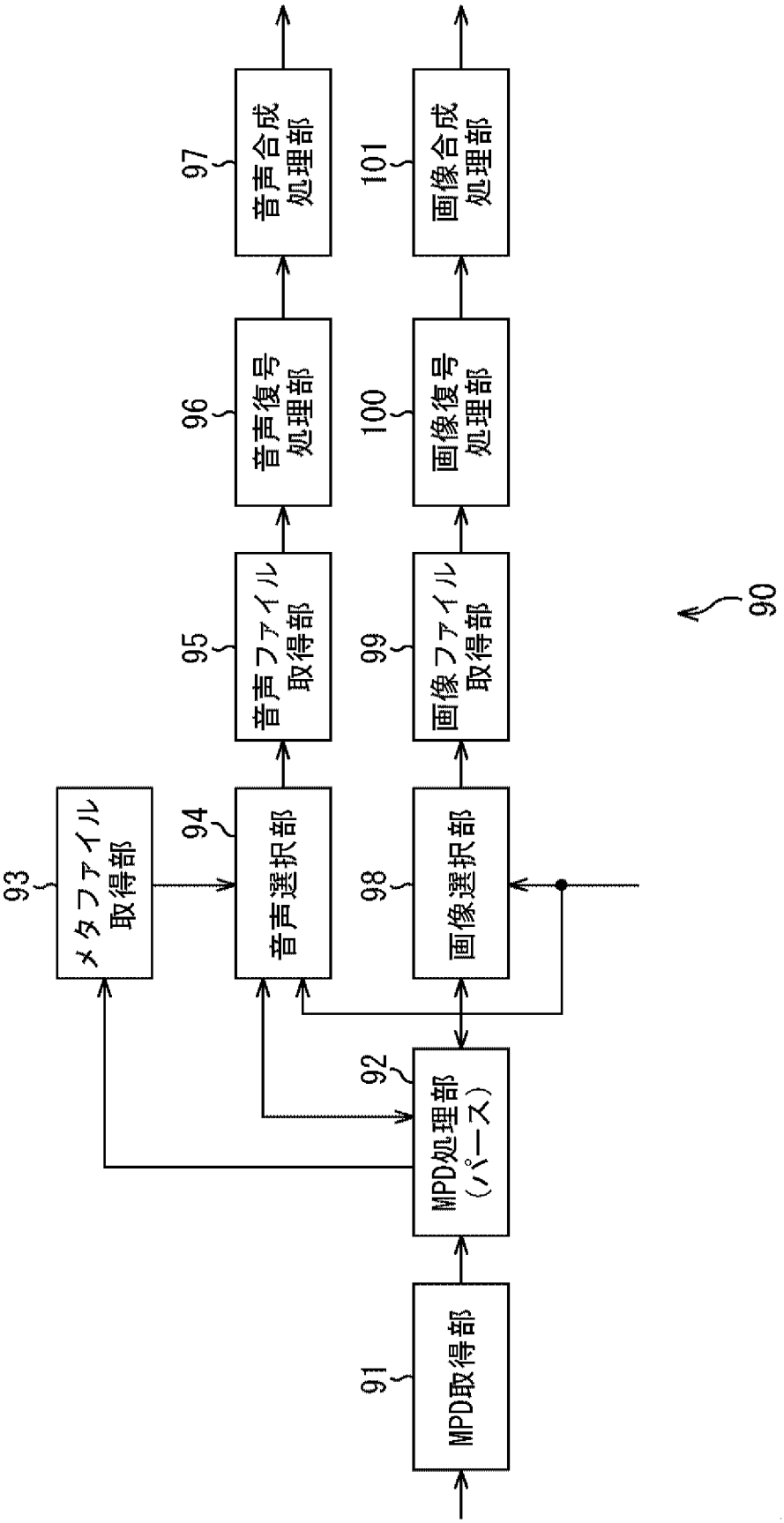
[図11]
図11


画面分割
処理部
51
画像符号化
処理部
52
音声符号化
処理部
55
画像ファイル
生成部
53
音声ファイル
生成部
56
画像情報
生成部
54
MPD
生成部
57
サーバ
アップロード
処理部
58
11

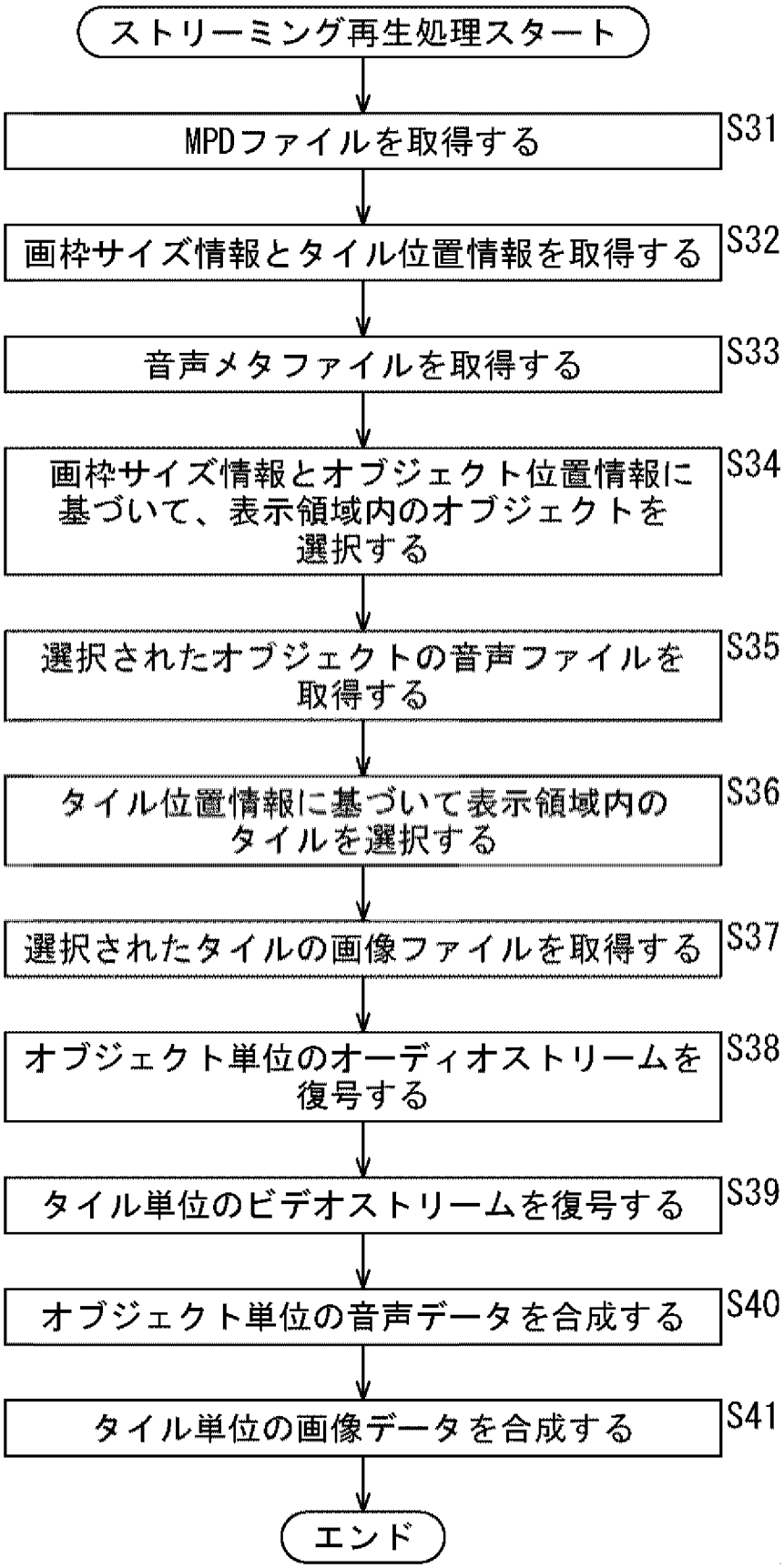
[図12]
図12


ファイル生成処理スタート
画像データをタイル単位に分割する
S11
各タイルのビデオストリームを生成する
S12
各タイルの画像ファイルを生成する
S13
画枠サイズ情報を取得する
S14
画枠サイズ情報とタイル位置情報を含む
画像情報を生成する
S15
各オブジェクトのオーディオストリームと
オーディオメタデータの符号化データを
生成する
S16
各オブジェクトの音声ファイルと
音声メタファイルを生成する
S17
画像情報等を含むMPDファイルを生成する
S18
画像ファイル、音声ファイル、
音声メタファイル、およびMPDファイルを
アップロードする
S19
エンド

[図13]
図13

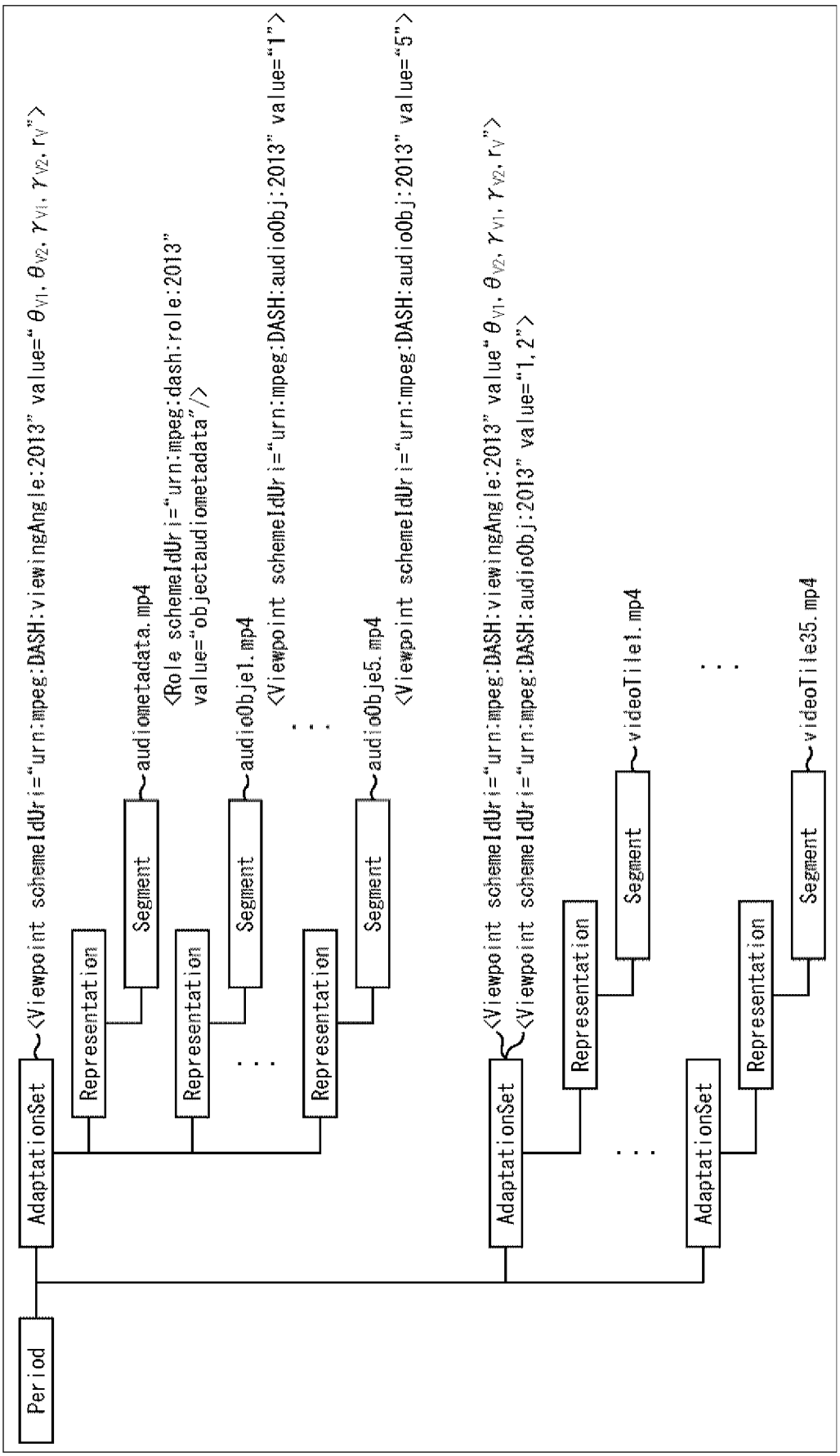

MPD取得部
91
MPD処理部（パース）
92
メタファイル取得部
93
音声選択部
94
音声ファイル取得部
95
音声復号処理部
96
音声合成処理部
97
画像選択部
98
画像ファイル取得部
99
画像復号処理部
100
画像合成処理部
101
90

[図14]
図14

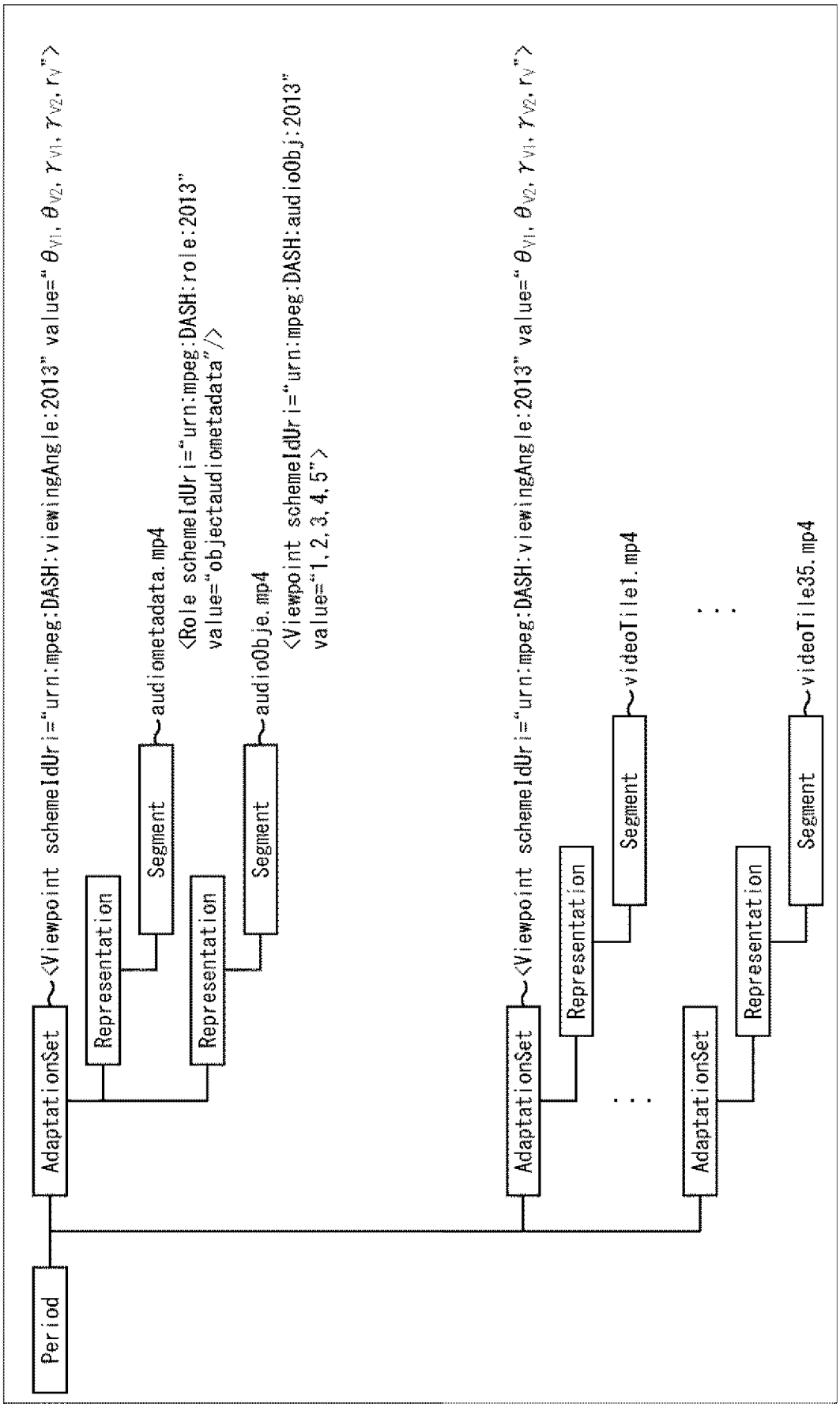

ストリーミング再生処理スタート
MPDファイルを取得する
S31
画枠サイズ情報とタイル位置情報を取得する
S32
音声メタファイルを取得する
S33
画枠サイズ情報とオブジェクト位置情報に基づいて、表示領域内のオブジェクトを選択する
S34
選択されたオブジェクトの音声ファイルを取得する
S35
タイル位置情報に基づいて表示領域内のタイルを選択する
S36
選択されたタイルの画像ファイルを取得する
S37
オブジェクト単位のオーディオストリームを復号する
S38
タイル単位のビデオストリームを復号する
S39
オブジェクト単位の音声データを合成する
S40
タイル単位の画像データを合成する
S41
エンド

[図15]

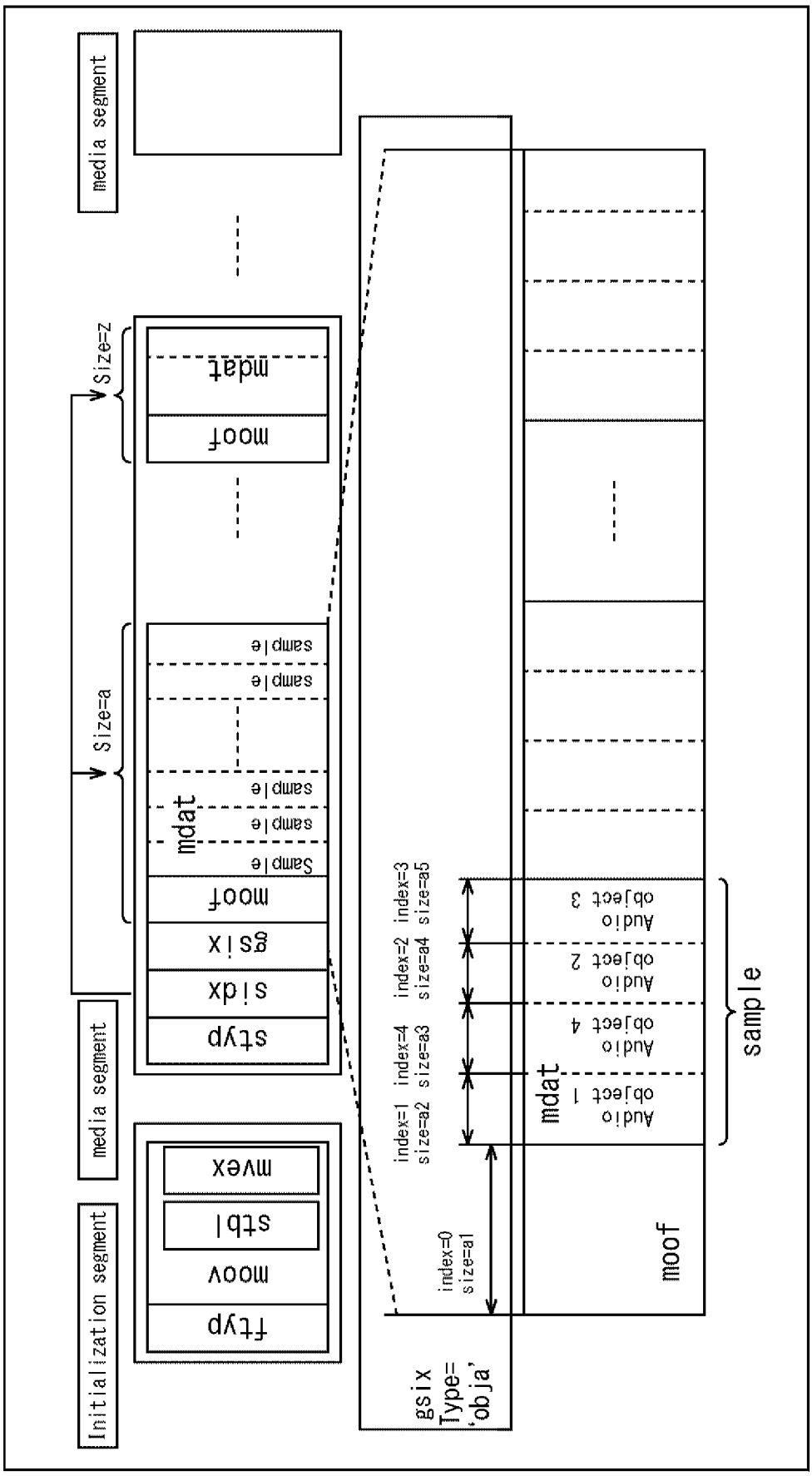
図15


Period
AdaptationSet
<Viewpoint schemeIdUri="urn:mpeg:DASH:viewingAngle:2013" value="θV1, θV2, γV1, γV2, rV">
Representation
Segment
audiometadata.mp4
<Role schemeIdUri="urn:mpeg:dash:role:2013" value="objectaudiometadata"/>
Representation
Segment
audioObje1.mp4
<Viewpoint schemeIdUri="urn:mpeg:DASH:audioObj:2013" value="1">
Representation
Segment
audioObje5.mp4
<Viewpoint schemeIdUri="urn:mpeg:DASH:audioObj:2013" value="5">
AdaptationSet
<Viewpoint schemeIdUri="urn:mpeg:DASH:viewingAngle:2013" value="θV1, θV2, γV1, γV2, rV">
<Viewpoint schemeIdUri="urn:mpeg:DASH:audioObj:2013" value="1,2">
Representation
Segment
videoTile1.mp4
AdaptationSet
Representation
Segment
videoTile35.mp4

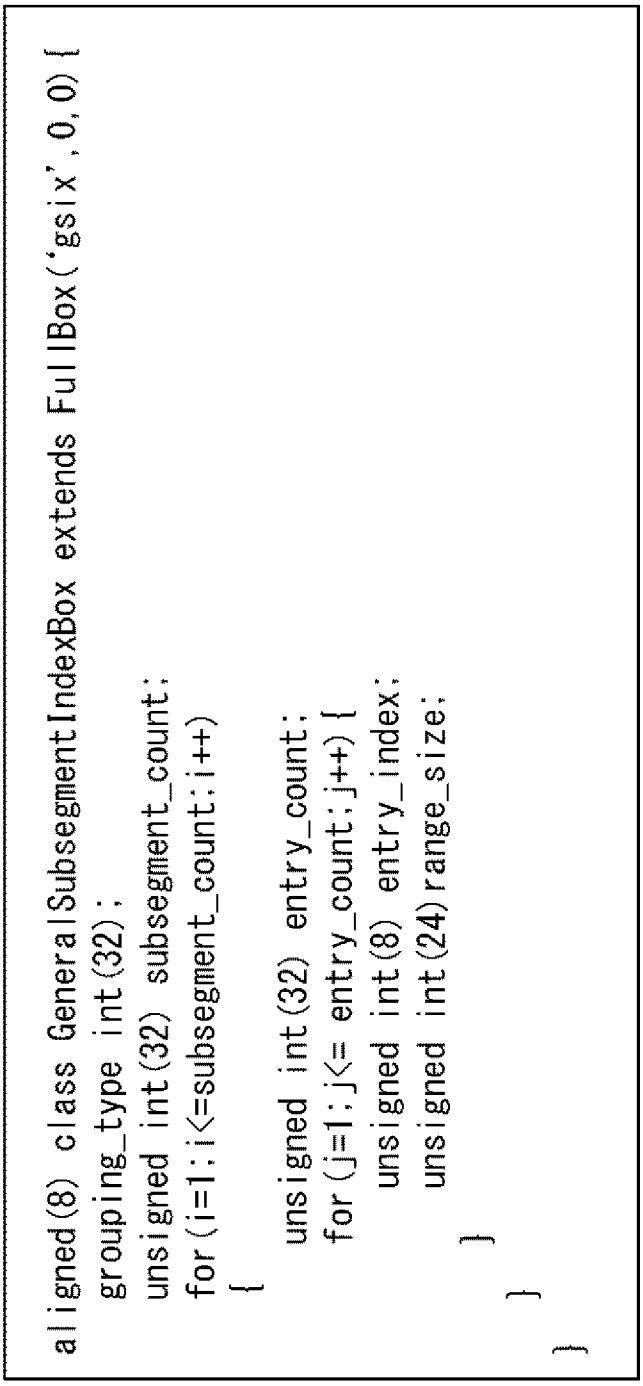
[図16]
図16


Period
AdaptationSet
<Viewpoint schemeIdUri="urn:mpeg:DASH:viewingAngle:2013" value="θv1, θv2, γv1, γv2, rv">
Representation
Segment
audiometadata.mp4
<Role schemeIdUri="urn:mpeg:DASH:role:2013"
value="objectaudiometadata"/>
Representation
Segment
audioObje.mp4
<Viewpoint schemeIdUri="urn:mpeg:DASH:audioObj:2013"
value="1,2,3,4,5">
AdaptationSet
<Viewpoint schemeIdUri="urn:mpeg:DASH:viewingAngle:2013" value="θv1, θv2, γv1, γv2, rv">
Representation
Segment
videoTile1.mp4
AdaptationSet
Representation
Segment
videoTile35.mp4

[図17]
図17

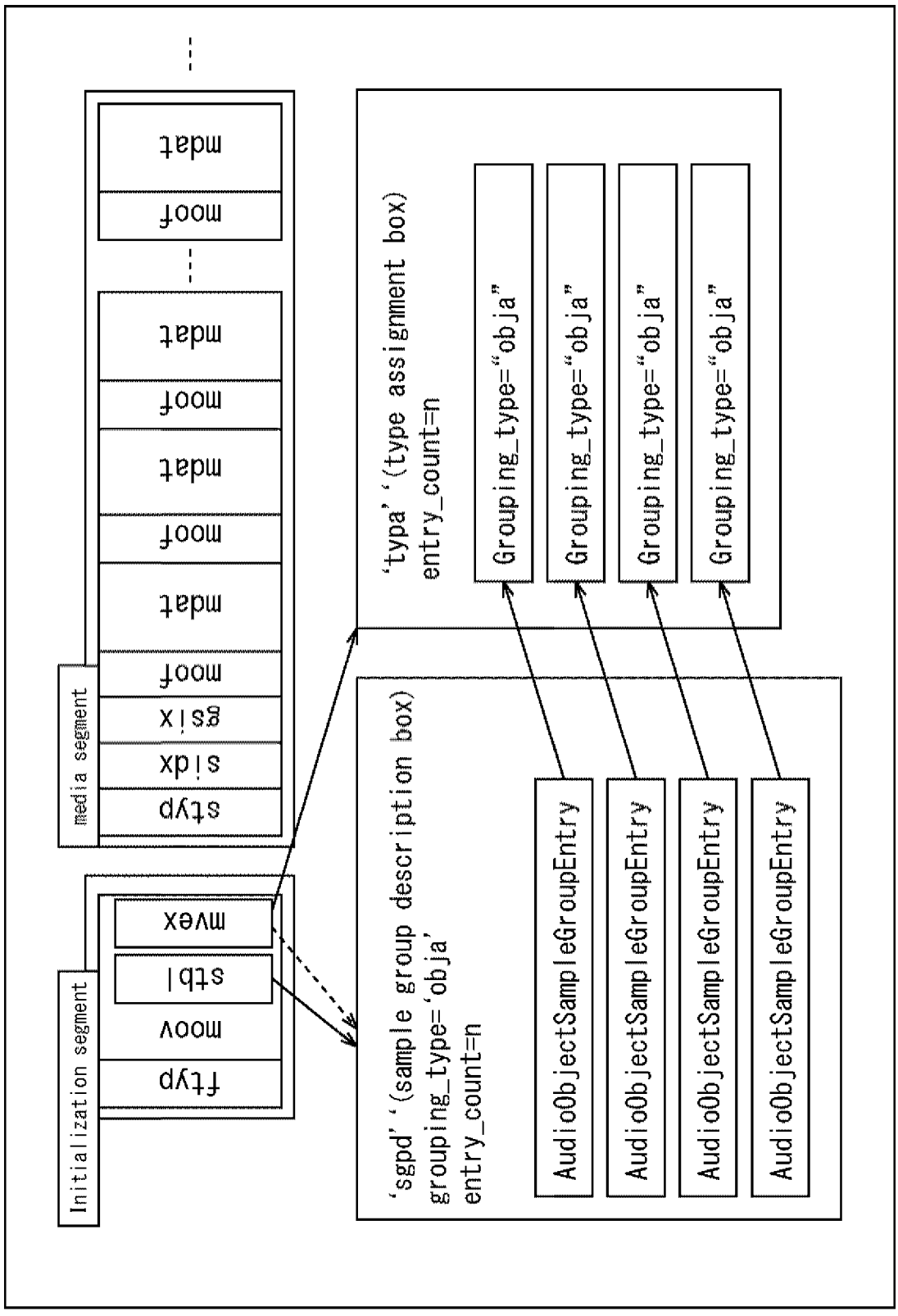

Initialization segment
media segment
media segment
ftyp
moov
stbl
mvex
styp
sidx
gsix
moof
mdat
Sample
sample
sample
sample
sample
sample
Size=a
moof
mdat
Size=z
media segment
gsix
Type=
'obja'
index=0
size=a1
index=1 index=4 index=2 index=3
size=a2 size=a3 size=a4 size=a5
moof
mdat
Audio
object 1
Audio
object 4
Audio
object 2
Audio
object 3
sample

[図18]
図18

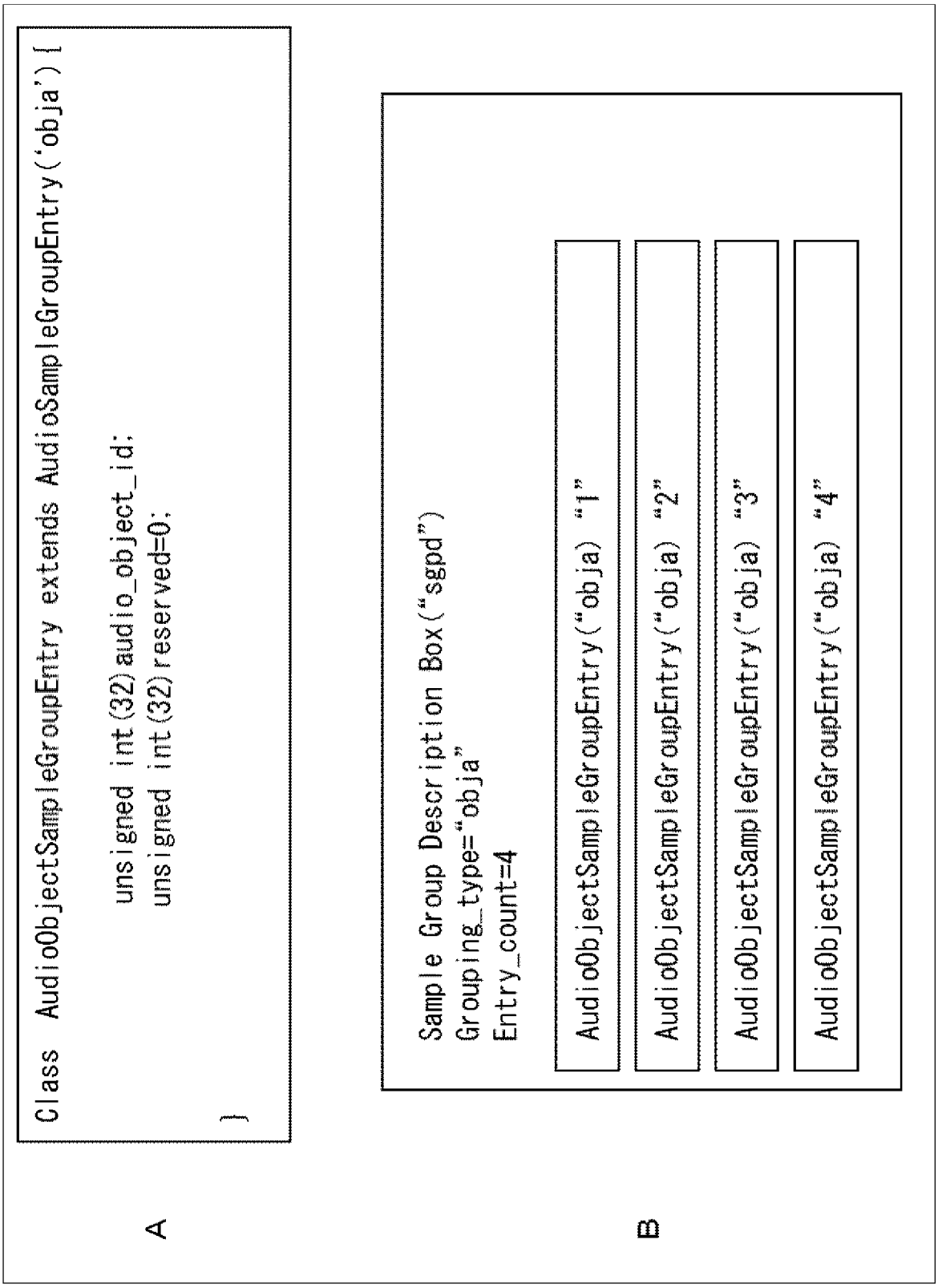


```
aligned(8) class GeneralSubsegmentIndexBox extends FullBox('gsix', 0, 0) {
    grouping_type int(32);
    unsigned int(32) subsegment_count;
    for(i=1; i<=subsegment_count; i++)
        {
            unsigned int(32) entry_count;
            for(j=1; j<= entry_count; j++) {
                unsigned int(8) entry_index;
                unsigned int(24) range_size;
            }
        }
    }
```

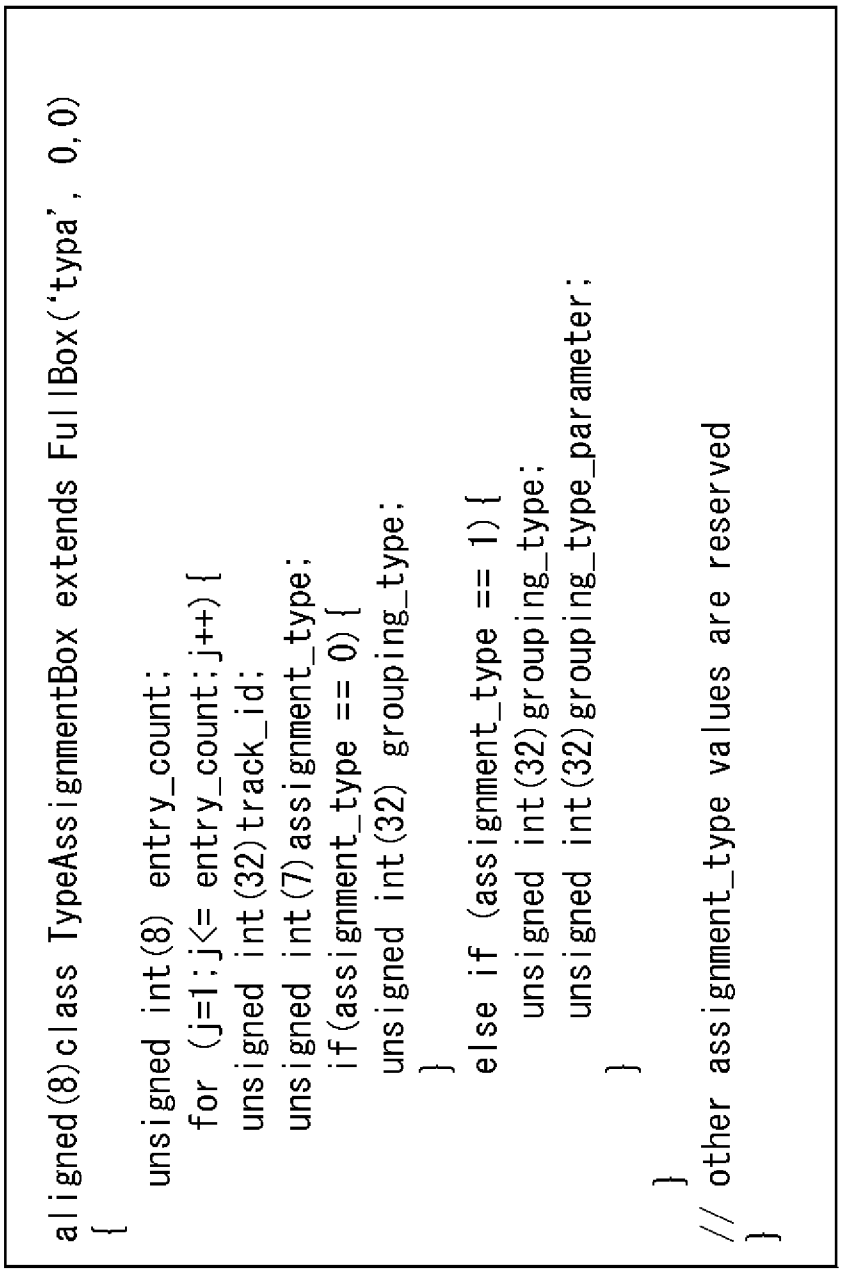
[図19]
図19


Initialization segment
ftyp
moov
stbl
mvex
media segment
styp
sidx
ssix
moof
mdat
moof
mdat
moof
mdat
moof
mdat
'sgpd' (sample group description box)
grouping_type='obja'
entry_count=n
AudioObjectSampleGroupEntry
AudioObjectSampleGroupEntry
AudioObjectSampleGroupEntry
AudioObjectSampleGroupEntry
'typa' (type assignment box)
entry_count=n
Grouping_type="obja"
Grouping_type="obja"
Grouping_type="obja"
Grouping_type="obja"

[図20]
図20

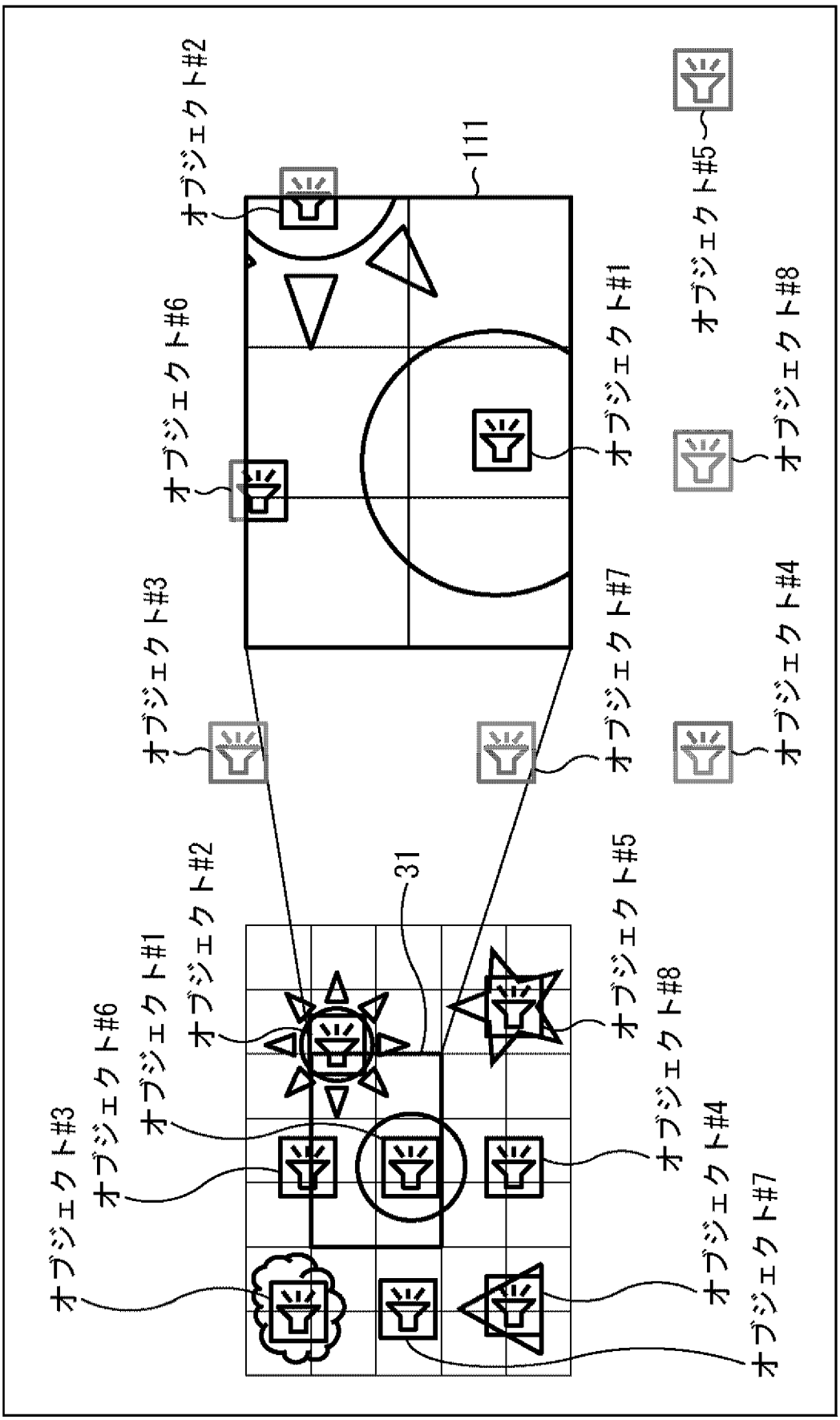

A
Class AudioObjectSampleGroupEntry extends AudioSampleGroupEntry('objа') {
unsigned int(32) audio_object_id;
unsigned int(32) reserved=0;
}
B
Sample Group Description Box("sgpd")
Grouping_type="objа"
Entry_count=4
AudioObjectSampleGroupEntry("objа") "1"
AudioObjectSampleGroupEntry("objа") "2"
AudioObjectSampleGroupEntry("objа") "3"
AudioObjectSampleGroupEntry("objа") "4"

[図21]
図21

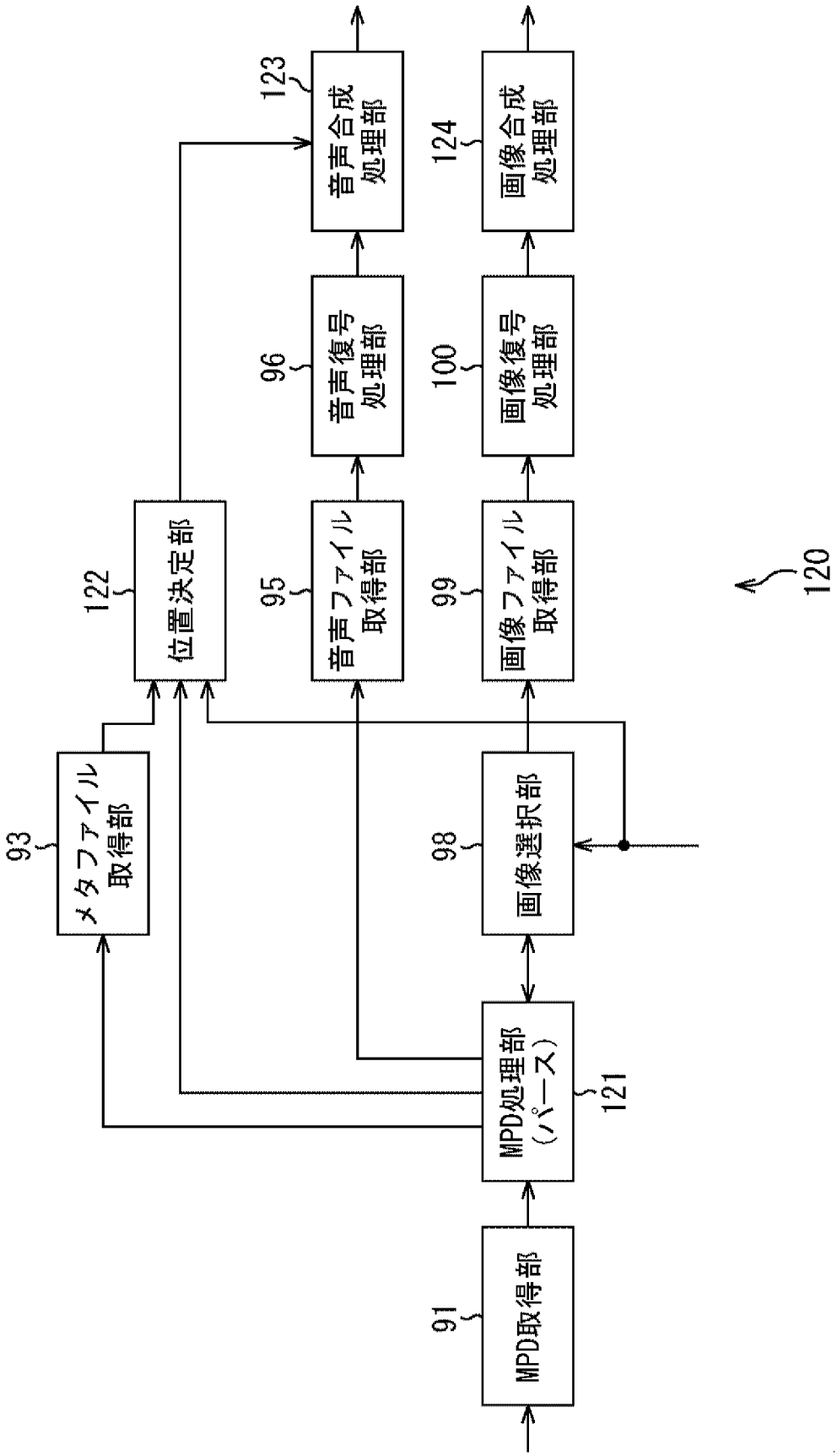


```
aligned(8)class TypeAssignmentBox extends FullBox('typa', 0,0)
{
    unsigned int(8) entry_count;
        for (j=1;j<= entry_count;j++){
    unsigned int(32)track_id;
    unsigned int(7)assignment_type;
        if(assignment_type == 0){
    unsigned int(32) grouping_type;
        }
        else if (assignment_type == 1){
            unsigned int(32)grouping_type;
            unsigned int(32)grouping_type_parameter;
        }
    }
// other assignment_type values are reserved
}
```

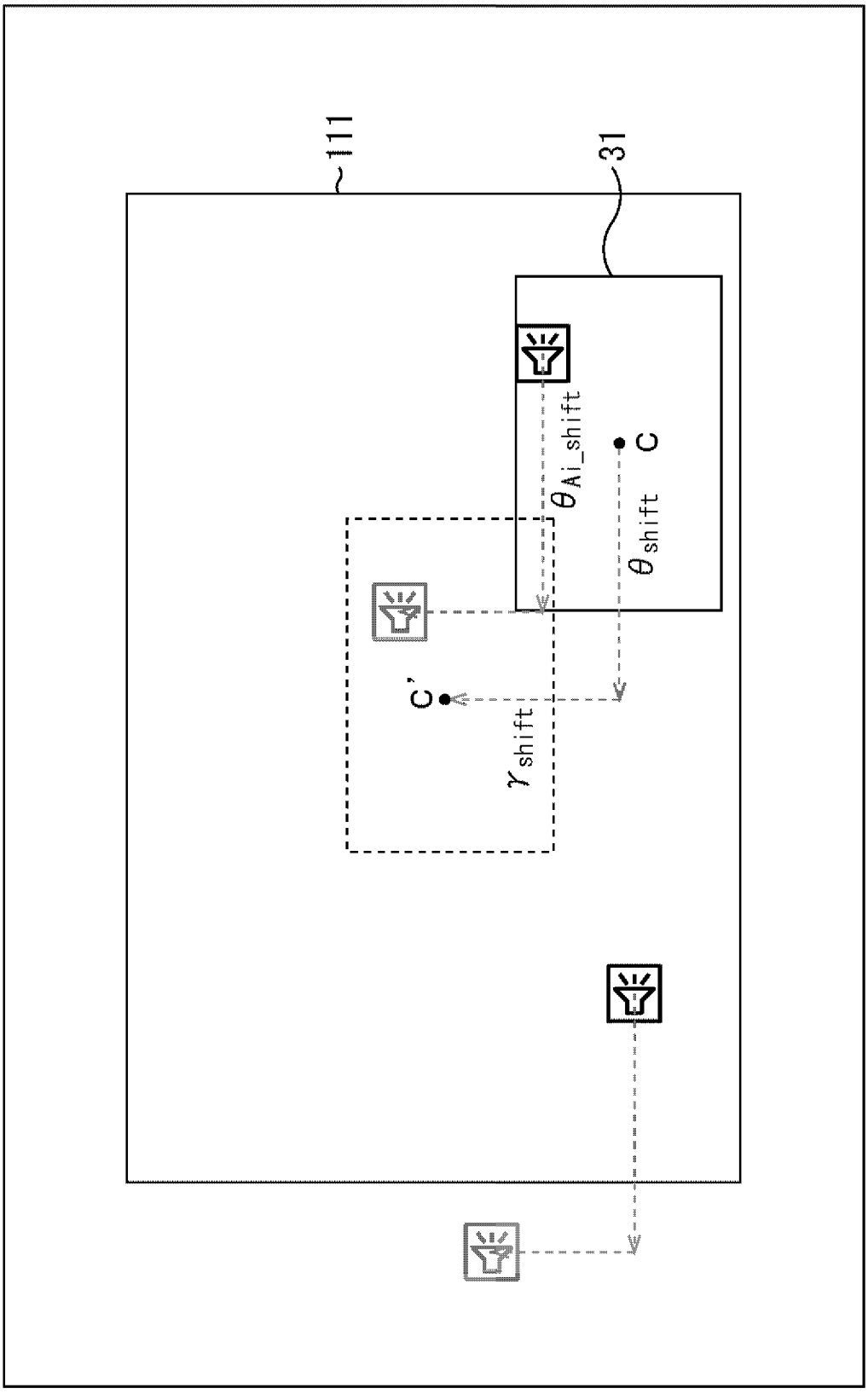
[図22]
図22


111
31
オブジェクト#1
オブジェクト#2
オブジェクト#3
オブジェクト#4
オブジェクト#5
オブジェクト#6
オブジェクト#7
オブジェクト#8

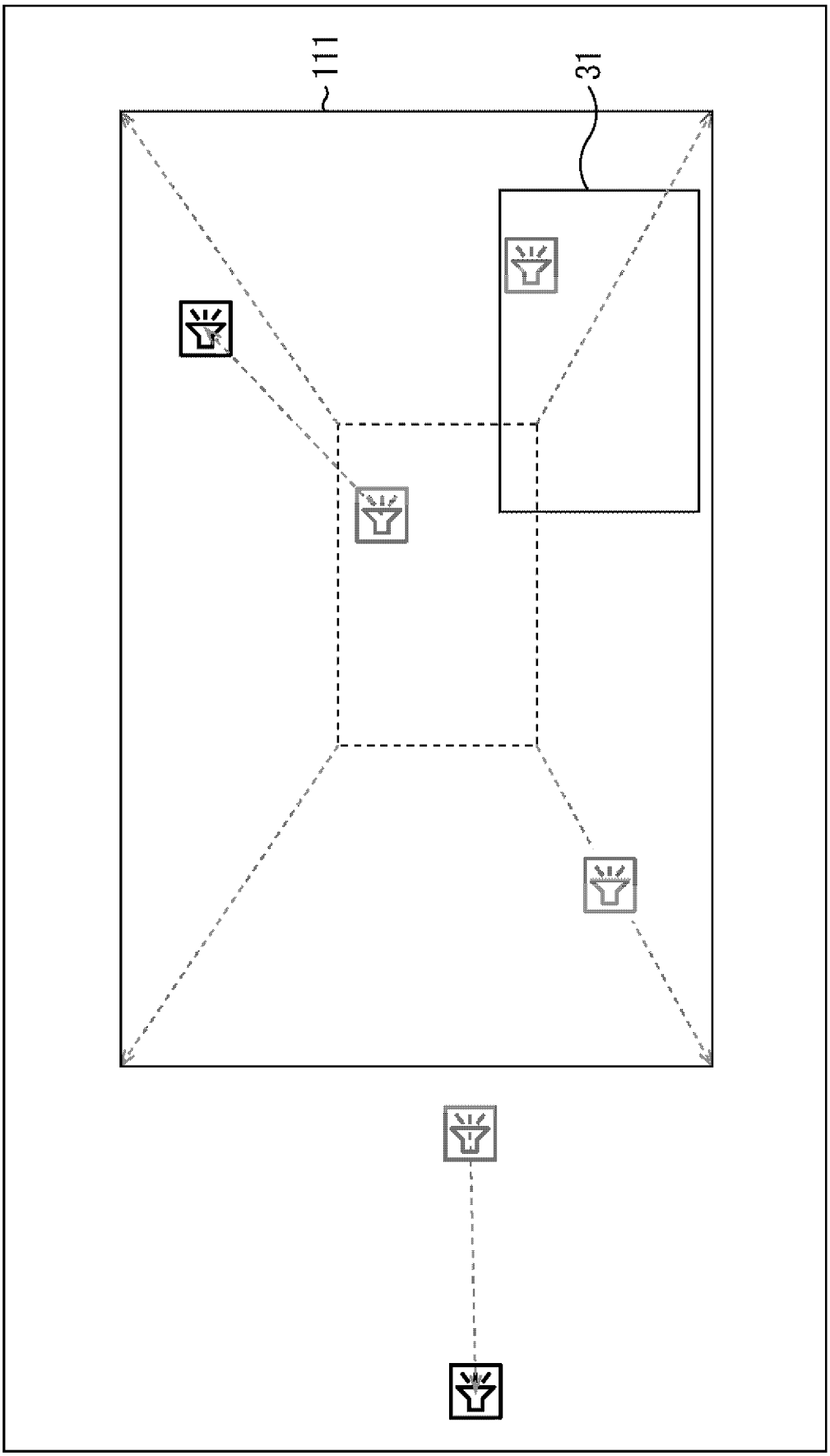
[図23]
図23


MPD取得部
91
MPD処理部（パース）
121
メタファイル取得部
93
位置決定部
122
音声ファイル取得部
95
音声復号処理部
96
音声合成処理部
123
画像選択部
98
画像ファイル取得部
99
画像復号処理部
100
画像合成処理部
124
120

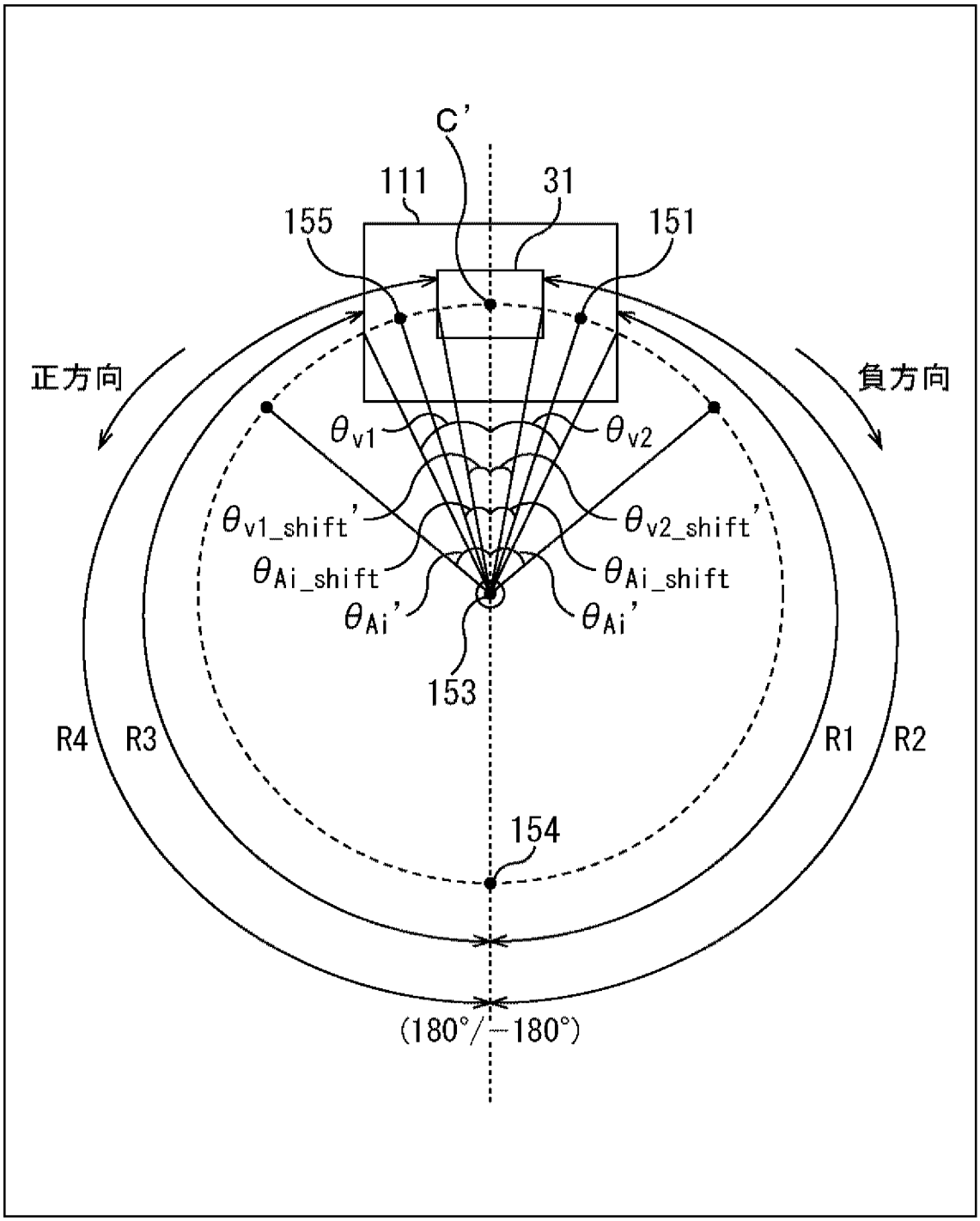
[図24]
図24



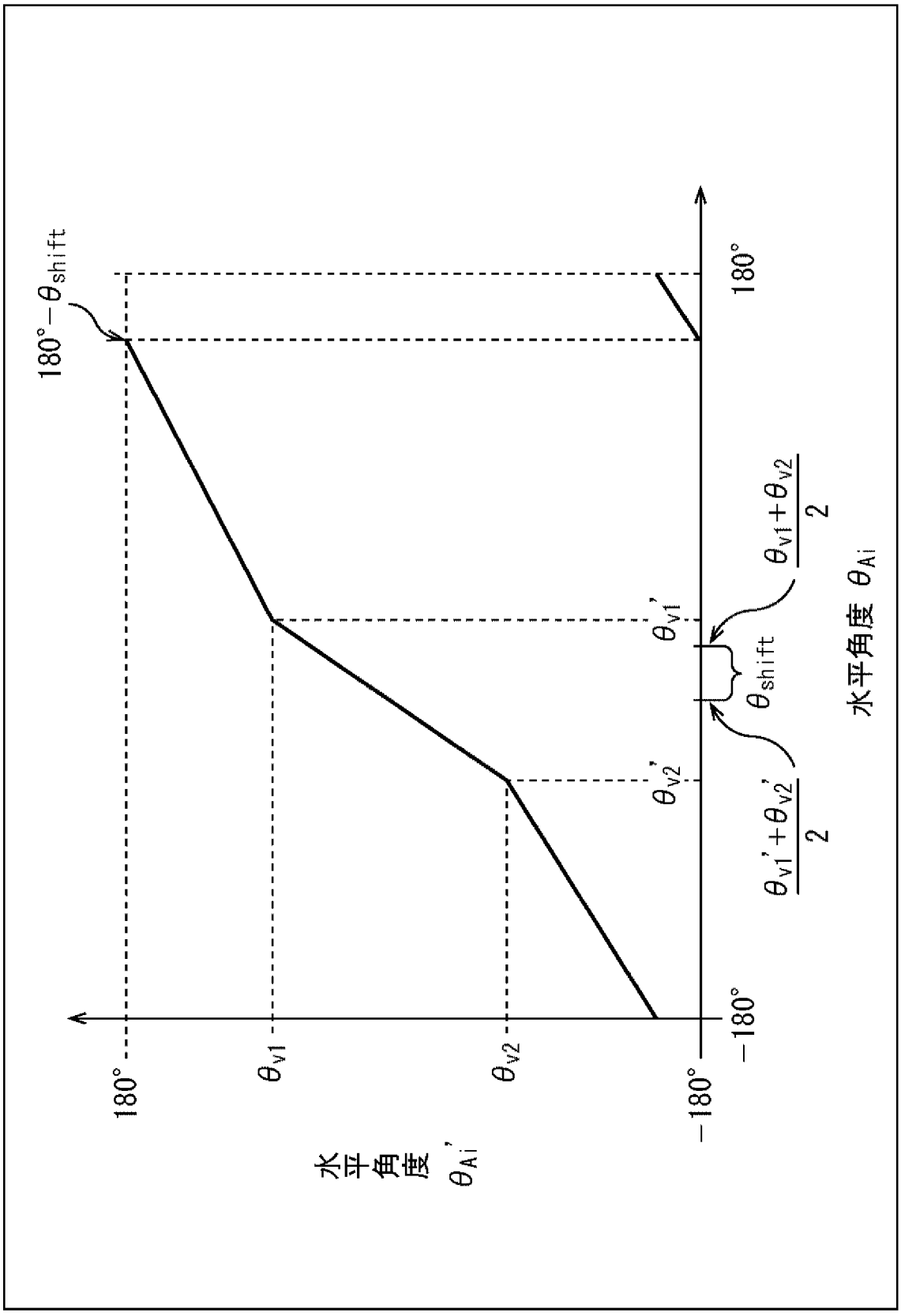
[図25]
図25



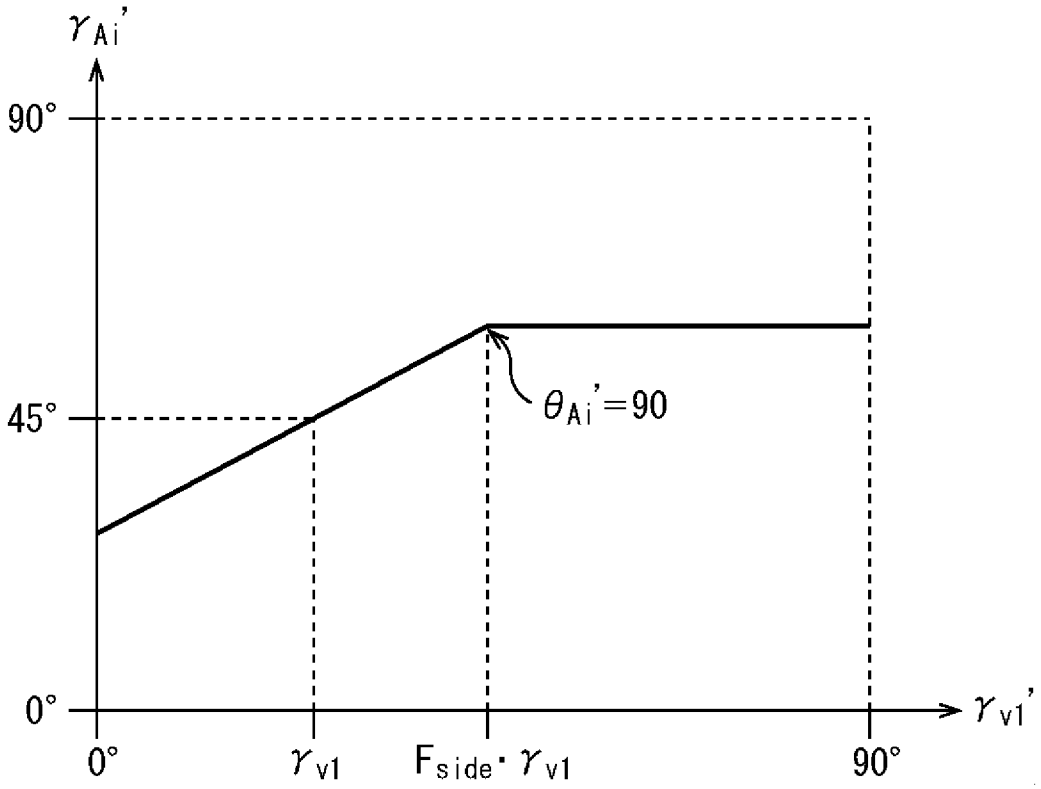
[図26]
図26


C'
111
31
155
151
正方向
負方向
θv1
θv2
θv1_shift'
θv2_shift'
θAi_shift
θAi_shift
θAi'
θAi'
153
R4
R3
R1
R2
154
(180°/−180°)

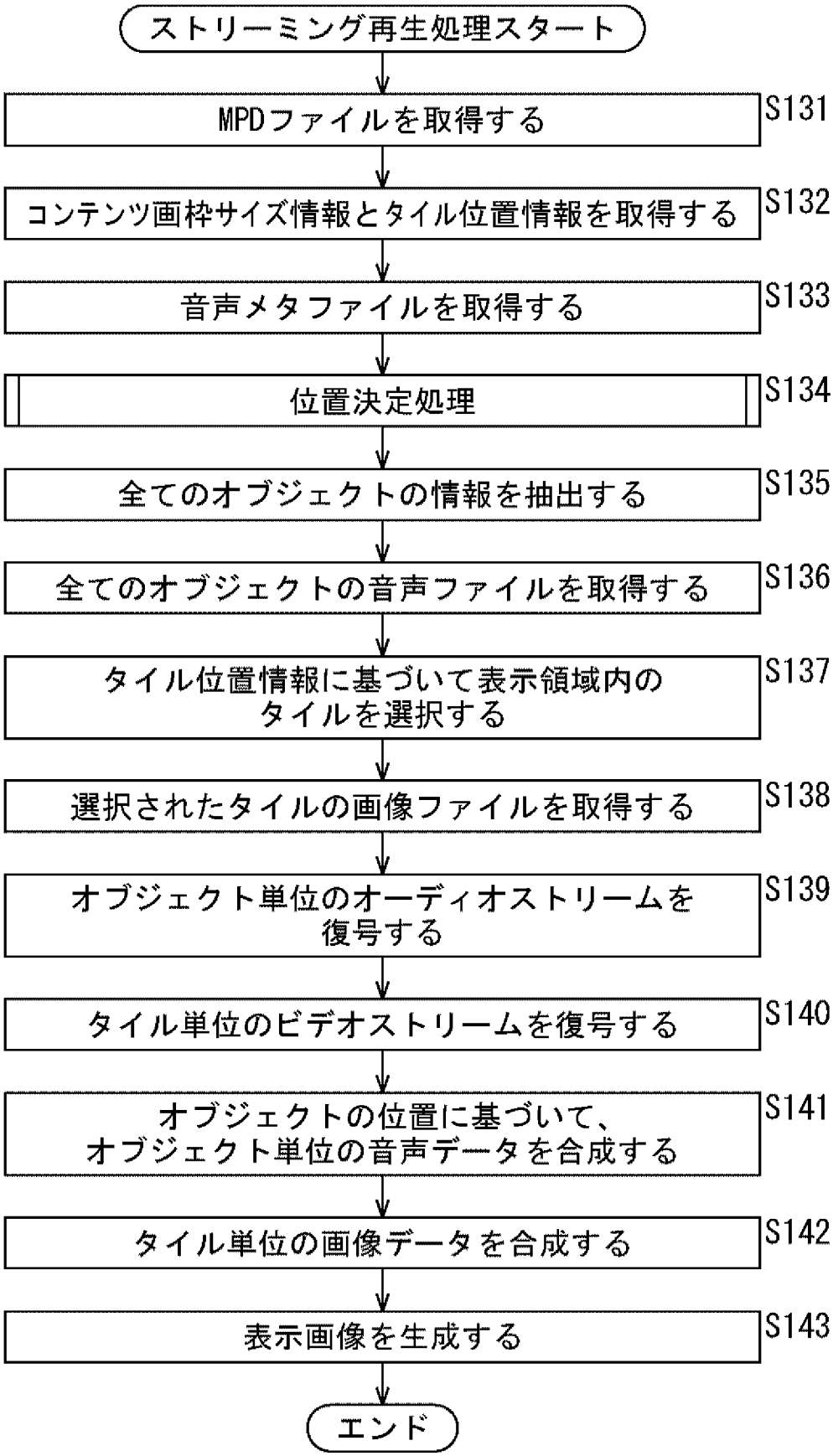
[図27]
図27


180°−θshift
180°
θv1
θv2
−180°
水平角度 θAi'
θv1'
θv2'
θshift
(θv1+θv2)/2
(θv1'+θv2')/2
水平角度 θAi

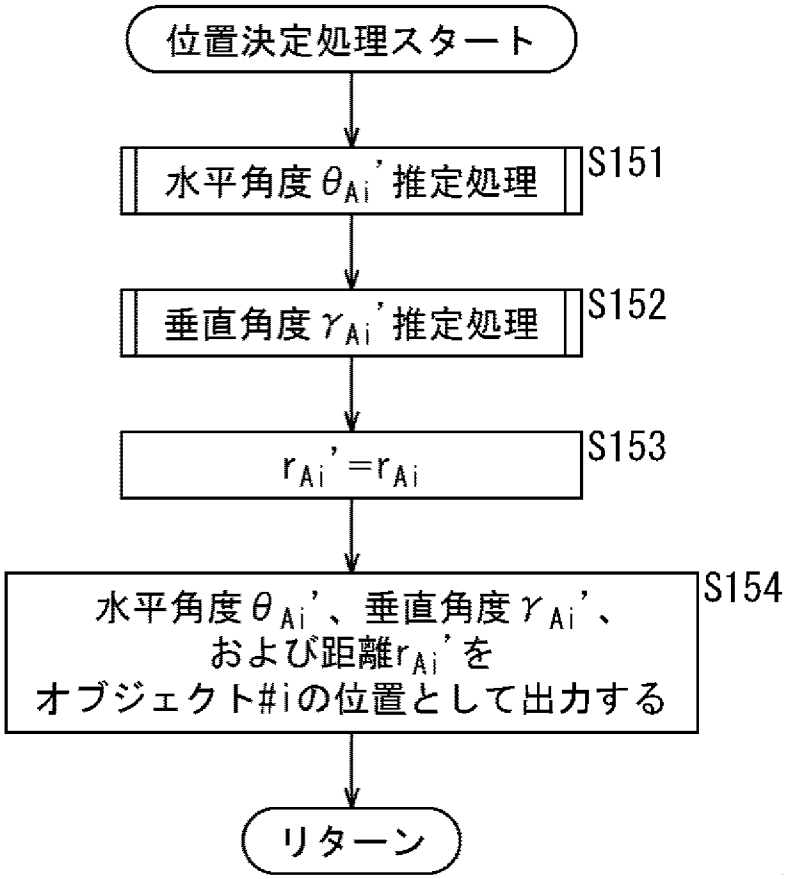
[図28]
図28


γ_{Ai}'
90°
45°
0°
$\theta_{Ai}'=90$
γ_{v1}'
0°
γ_{v1}
$F_{side}\cdot\gamma_{v1}$
90°

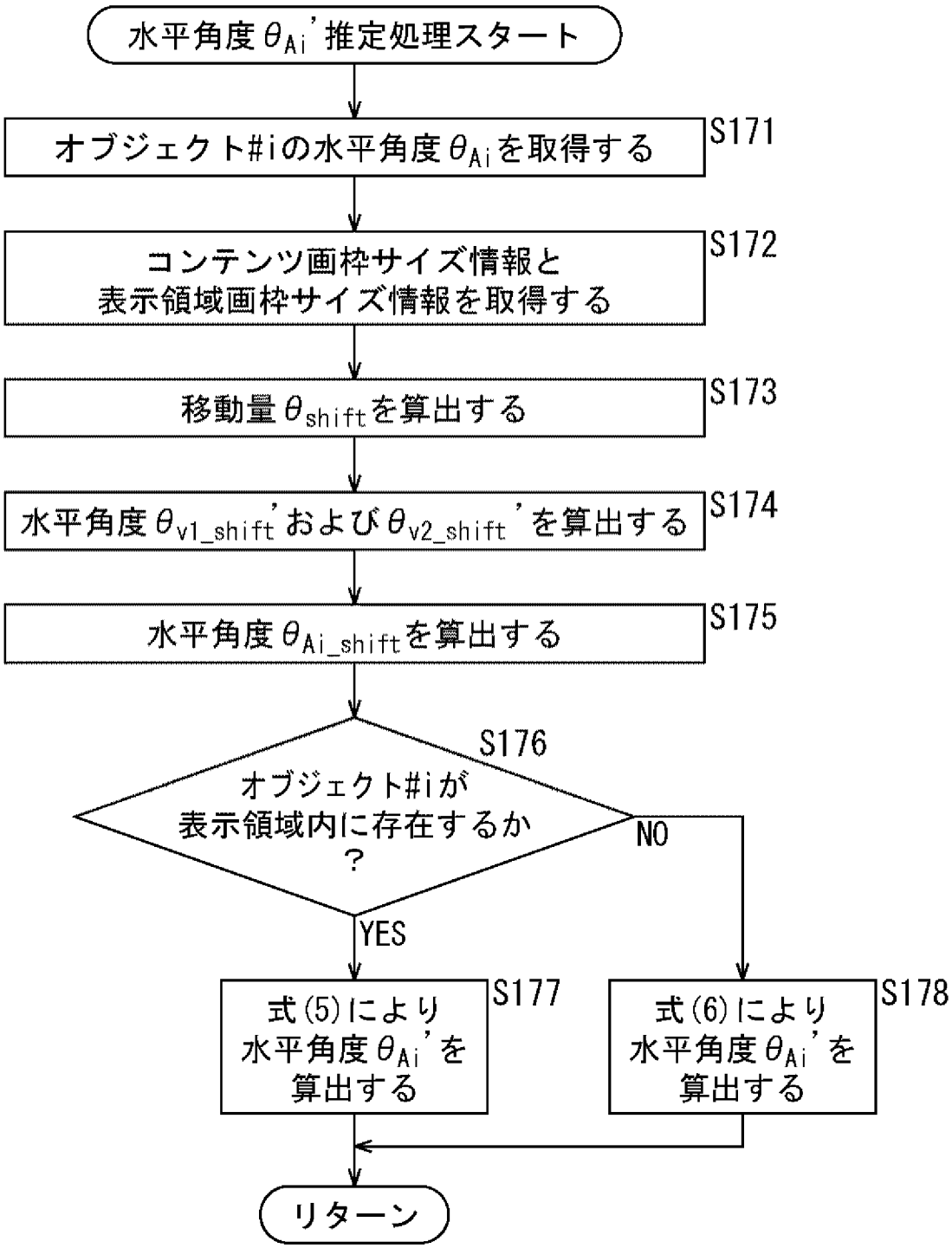
[図29]
図29


ストリーミング再生処理スタート
MPDファイルを取得する
S131
コンテンツ画枠サイズ情報とタイル位置情報を取得する
S132
音声メタファイルを取得する
S133
位置決定処理
S134
全てのオブジェクトの情報を抽出する
S135
全てのオブジェクトの音声ファイルを取得する
S136
タイル位置情報に基づいて表示領域内の
タイルを選択する
S137
選択されたタイルの画像ファイルを取得する
S138
オブジェクト単位のオーディオストリームを
復号する
S139
タイル単位のビデオストリームを復号する
S140
オブジェクトの位置に基づいて、
オブジェクト単位の音声データを合成する
S141
タイル単位の画像データを合成する
S142
表示画像を生成する
S143
エンド

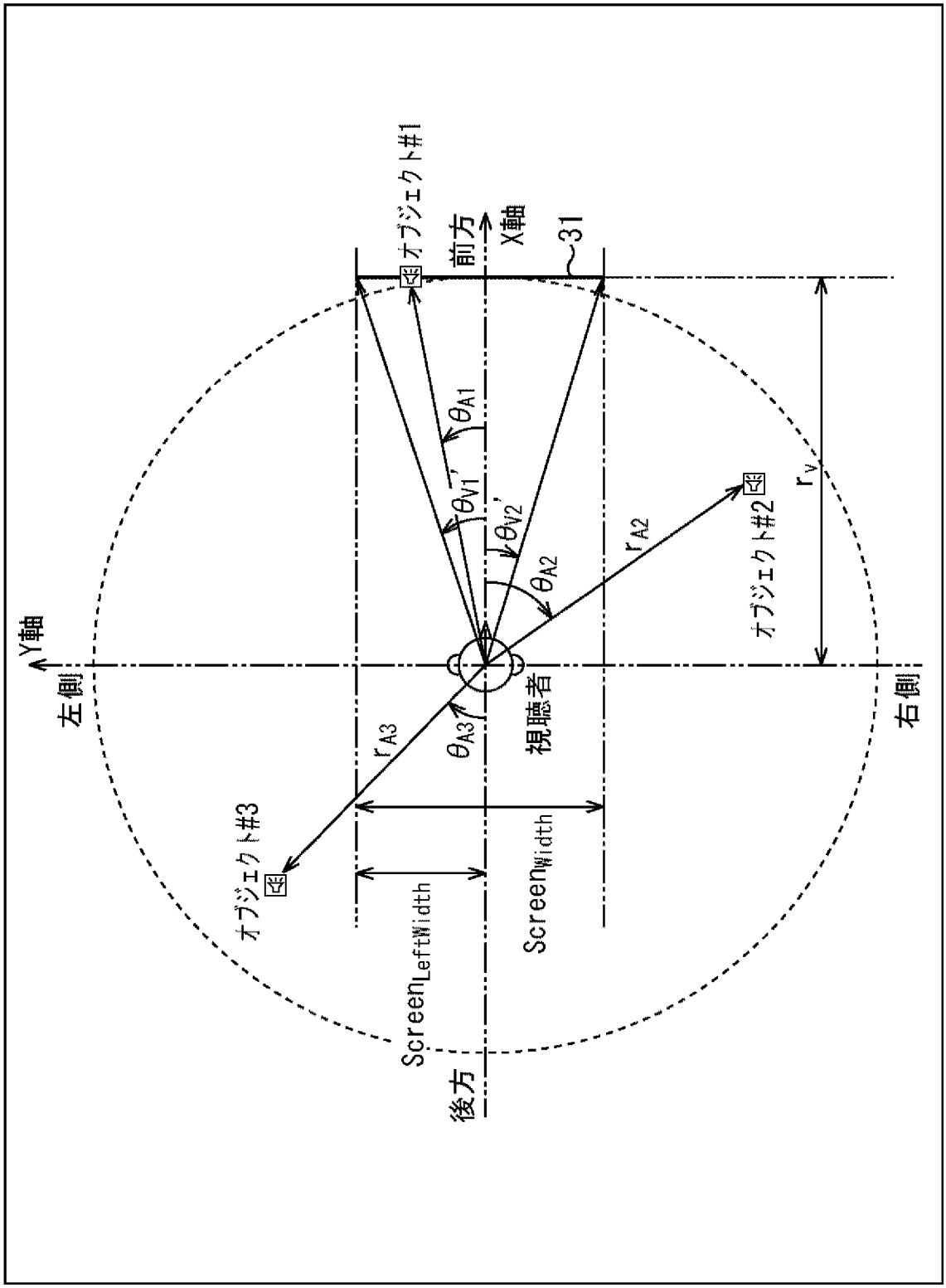
[図30]
図30


位置決定処理スタート
水平角度θAi'推定処理
S151
垂直角度γAi'推定処理
S152
rAi'=rAi
S153
水平角度θAi'、垂直角度γAi'、
および距離rAi'を
オブジェクト#iの位置として出力する
S154
リターン

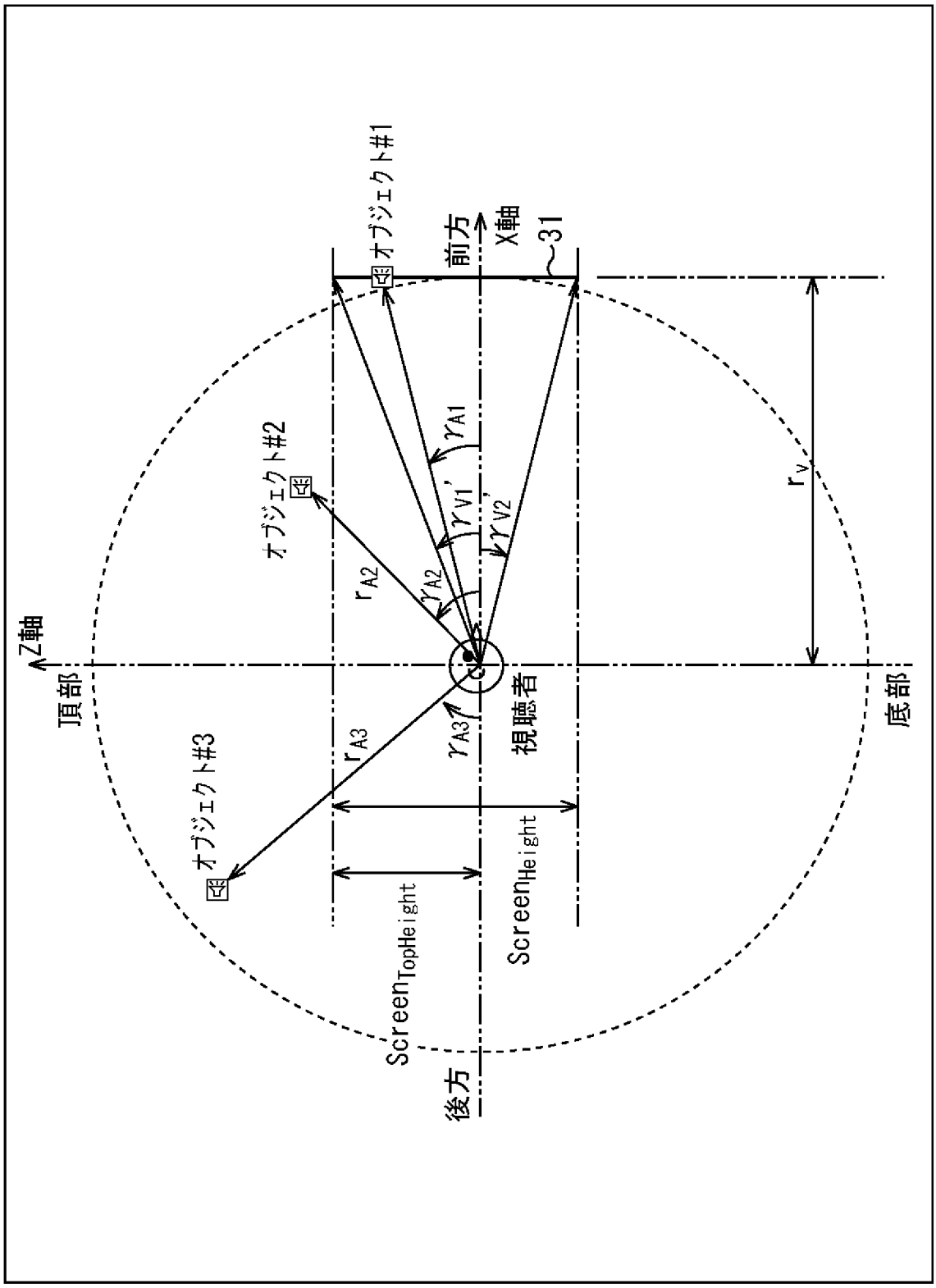
[図31]
図31


水平角度θAi'推定処理スタート
オブジェクト#iの水平角度θAiを取得する
S171
コンテンツ画枠サイズ情報と
表示領域画枠サイズ情報を取得する
S172
移動量θshiftを算出する
S173
水平角度θv1_shift'およびθv2_shift'を算出する
S174
水平角度θAi_shiftを算出する
S175
S176
オブジェクト#iが
表示領域内に存在するか
？
NO
YES
式(5)により
水平角度θAi'を
算出する
S177
式(6)により
水平角度θAi'を
算出する
S178
リターン

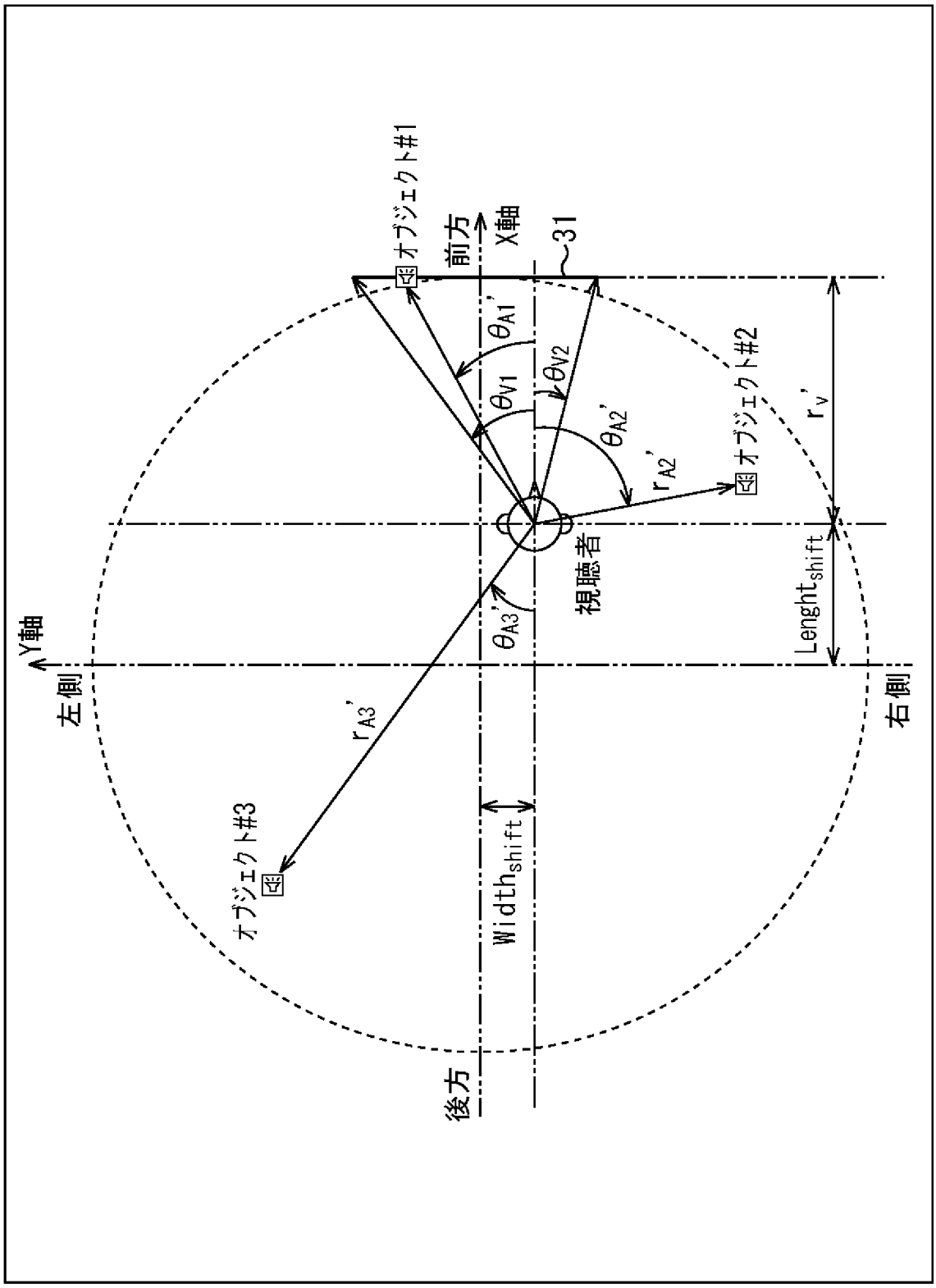
[図32]
図32


オブジェクト#1
前方
X軸
31
θA1
θV1'
θV2'
θA2
rA2
rV
オブジェクト#2
Y軸
左側
右側
視聴者
θA3
rA3
オブジェクト#3
ScreenLeftWidth
ScreenWidth
後方

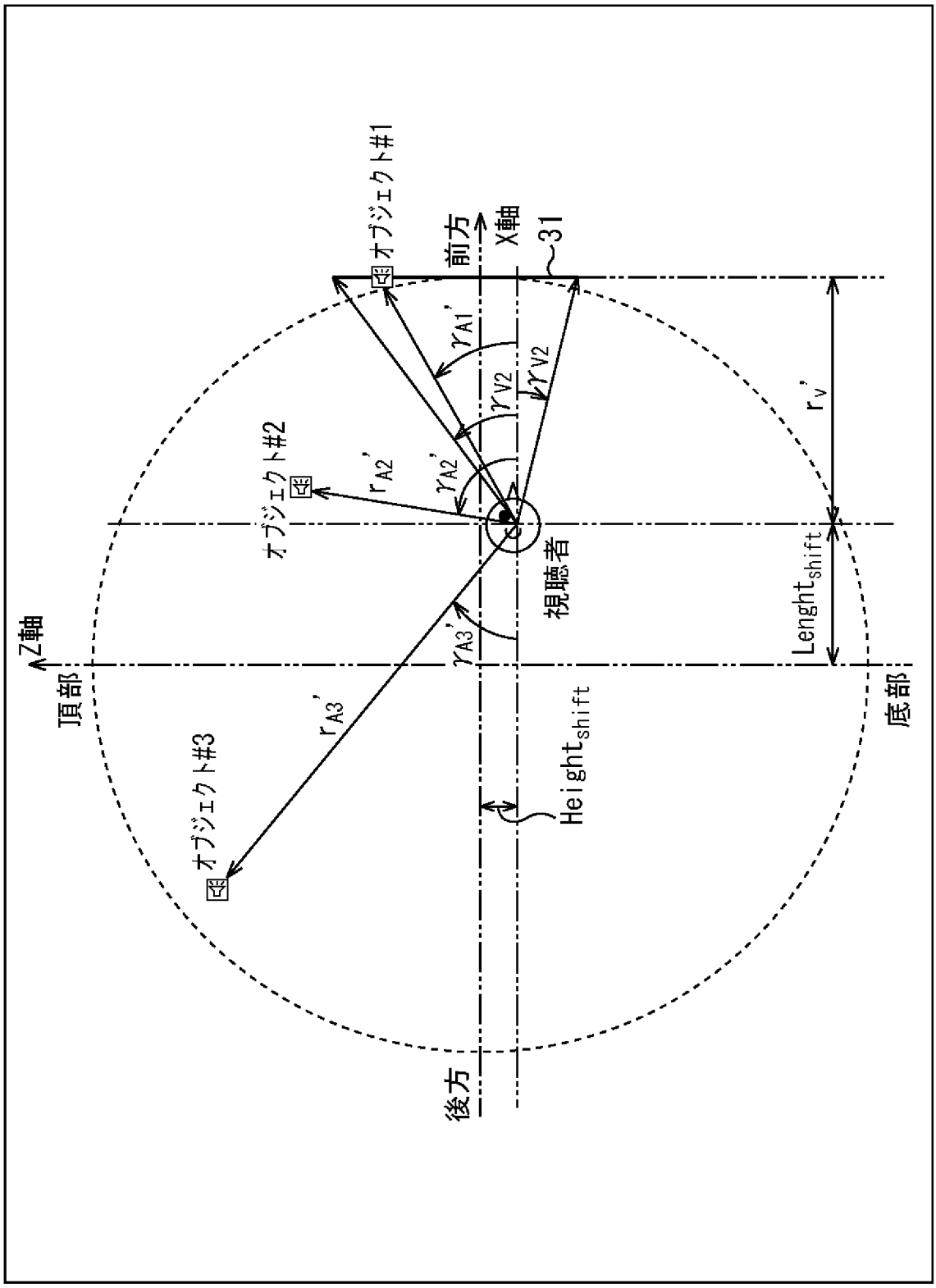
[図33]
図33


オブジェクト#1
前方
X軸
31
オブジェクト#2
オブジェクト#3
Z軸
頂部
底部
視聴者
後方
r_{A1}
r_{A2}
r_{A3}
r_V
γ_{A1}
γ_{A2}
γ_{A3}
γ_{V1}'
γ_{V2}'
$Screen_{TopHeight}$
$Screen_{Height}$

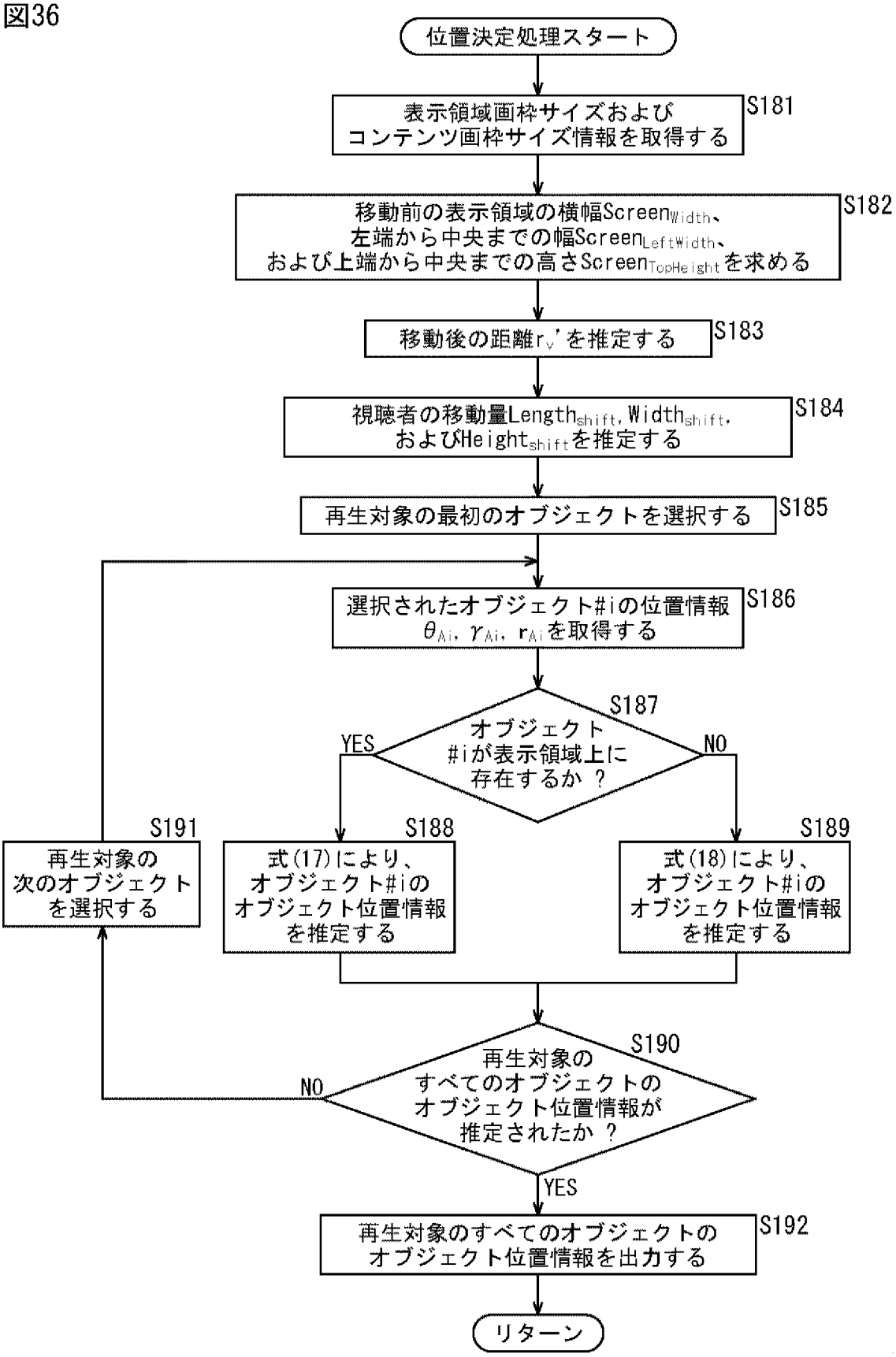
[図34]
図34


オブジェクト#1
前方
X軸
31
θA1'
θV1
θV2
θA2'
rA2'
オブジェクト#2
rv'
視聴者
Lenghtshift
Y軸
左側
θA3'
rA3'
右側
オブジェクト#3
Widthshift
後方

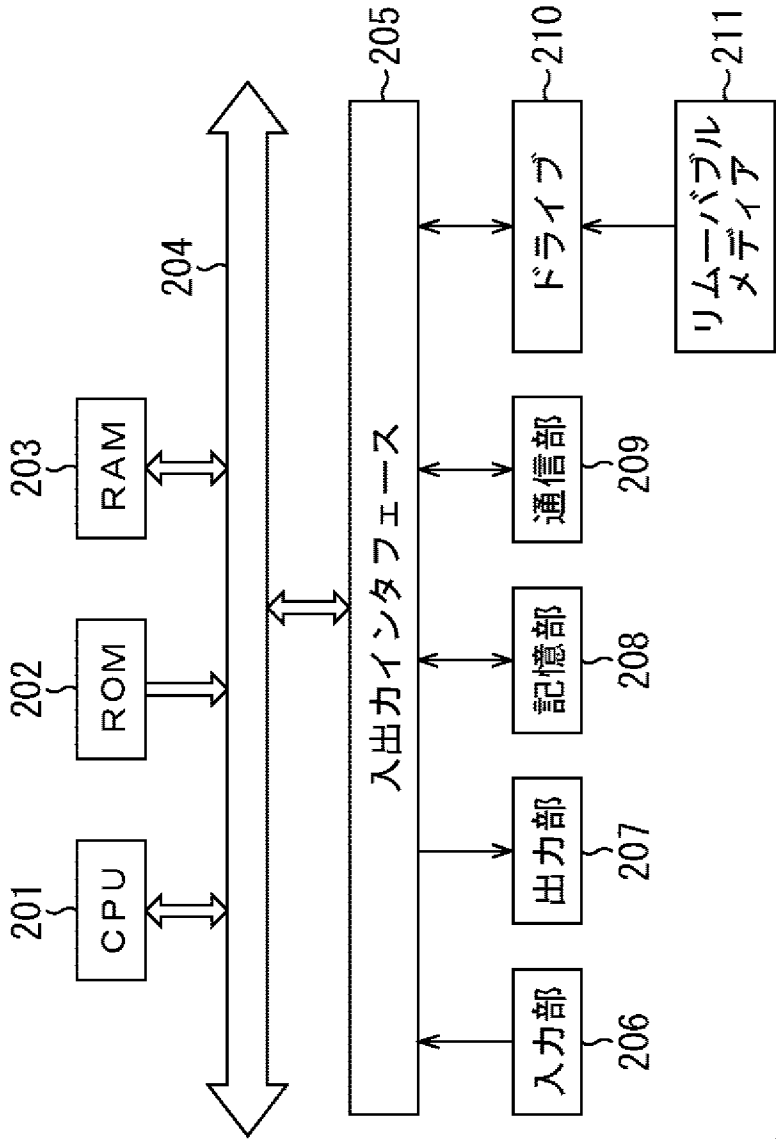
[図35]
図35


オブジェクト#1
前方
X軸
31
γ_{A1}'
γ_{V2}
γ_{V2}
r_v'
オブジェクト#2
r_{A2}'
γ_{A2}'
視聴者
$Lenght_{shift}$
Z軸
頂部
γ_{A3}'
底部
r_{A3}'
オブジェクト#3
$Height_{shift}$
後方

[図36]
図36


位置決定処理スタート
表示領域画枠サイズおよび
コンテンツ画枠サイズ情報を取得する
S181
移動前の表示領域の横幅$Screen_{Width}$、
左端から中央までの幅$Screen_{LeftWidth}$、
および上端から中央までの高さ$Screen_{TopHeight}$を求める
S182
移動後の距離r_v'を推定する
S183
視聴者の移動量$Length_{shift}$, $Width_{shift}$,
および$Height_{shift}$を推定する
S184
再生対象の最初のオブジェクトを選択する
S185
選択されたオブジェクト#iの位置情報
θ_{Ai}, γ_{Ai}, r_{Ai}を取得する
S186
オブジェクト
#iが表示領域上に
存在するか ？
S187
YES
NO
式(17)により、
オブジェクト#iの
オブジェクト位置情報
を推定する
S188
式(18)により、
オブジェクト#iの
オブジェクト位置情報
を推定する
S189
再生対象の
次のオブジェクト
を選択する
S191
再生対象の
すべてのオブジェクトの
オブジェクト位置情報が
推定されたか ？
S190
NO
YES
再生対象のすべてのオブジェクトの
オブジェクト位置情報を出力する
S192
リターン

[図37]
図37


201
CPU
202
ROM
203
RAM
204
205
入出力インタフェース
206
入力部
207
出力部
208
記憶部
209
通信部
210
ドライブ
211
リムーバブル
メディア

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/067508

A. CLASSIFICATION OF SUBJECT MATTER
H04N21/84(2011.01)i, *H04N5/765*(2006.01)i, *H04N5/91*(2006.01)i, *H04N21/439*(2011.01)i, *H04N21/4728*(2011.01)i, *H04N21/845*(2011.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N21/84, H04N5/765, H04N5/91, H04N21/439, H04N21/4728, H04N21/845

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2014
Kokai Jitsuyo Shinan Koho 1971-2014 Toroku Jitsuyo Shinan Koho 1994-2014

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 2001-197598 A (Mitsubishi Electric Corp.), 19 July 2001 (19.07.2001), paragraphs [0015] to [0026]; fig. 1 (Family: none)	1-3,9-13,17, 20 4-8,14-16, 18,19
A	JP 2004-7211 A (Victor Company of Japan, Ltd.), 08 January 2004 (08.01.2004), paragraphs [0014] to [0076]; fig. 1 to 7 (Family: none)	1-20
A	JP 2007-27846 A (Oki Electric Industry Co., Ltd.), 01 February 2007 (01.02.2007), paragraphs [0037] to [0089]; fig. 1 to 5 & US 2007/0025703 A1	1-20

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 July, 2014 (30.07.14)	12 August, 2014 (12.08.14)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/067508

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 6-301390 A (Sanyo Electric Co., Ltd.), 28 October 1994 (28.10.1994), paragraphs [0007] to [0025]; fig. 1 to 6 (Family: none)	1-20

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１４／０６７５０８

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　H04N21/84(2011.01)i, H04N5/765(2006.01)i, H04N5/91(2006.01)i, H04N21/439(2011.01)i, H04N21/4728(2011.01)i, H04N21/845(2011.01)i

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　H04N21/84, H04N5/765, H04N5/91, H04N21/439, H04N21/4728, H04N21/845

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１４年
日本国実用新案登録公報	１９９６－２０１４年
日本国登録実用新案公報	１９９４－２０１４年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	JP 2001-197598 A（三菱電機株式会社）2001.07.19, 段落【００１５】～【００２６】、第１図（ファミリーなし）	1-3,9-13,17,20
A		4-8,14-16,18,19
A	JP 2004-7211 A（日本ビクター株式会社）2004.01.08, 段落【００１４】～【００７６】、第１～７図（ファミリーなし）	1-20

☑　Ｃ欄の続きにも文献が列挙されている。

☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの
「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「Ｏ」口頭による開示、使用、展示等に言及する文献
「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「＆」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
３０．０７．２０１４	１２．０８．２０１４

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｃ　４６８７
福西　章人
電話番号　０３－３５８１－１１０１　内線　３５４１

Ｃ（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
A	JP 2007-27846 A（沖電気工業株式会社）2007.02.01，段落【００３７】～【００８９】、第１～５図 & US 2007/0025703 A1	1-20
A	JP 6-301390 A（三洋電機株式会社）1994.10.28，段落【０００７】～【００２５】、第１～６図（ファミリーなし）	1-20