

(19)日本国特許庁(JP)

**(12)特許公報(B2)**

(11)特許番号  
**特許第7365212号**  
**(P7365212)**

(45)発行日 令和5年10月19日(2023.10.19)

(24)登録日 令和5年10月11日(2023.10.11)

(51)国際特許分類

H 0 4 N	21/44 (2011.01)	F I	H 0 4 N	21/44
H 0 4 N	21/4385(2011.01)		H 0 4 N	21/4385

請求項の数 11 (全21頁)

(21)出願番号	特願2019-218746(P2019-218746)
(22)出願日	令和1年12月3日(2019.12.3)
(65)公開番号	特開2021-90118(P2021-90118A)
(43)公開日	令和3年6月10日(2021.6.10)
審査請求日	令和4年3月10日(2022.3.10)

(73)特許権者	310021766 株式会社ソニー・インタラクティブエンタテインメント 東京都港区港南1丁目7番1号
(74)代理人	100105924 弁理士 森下 賢樹
(74)代理人	100109047 弁理士 村田 雄祐
(74)代理人	100109081 弁理士 三木 友由
(74)代理人	100134256 弁理士 青木 武司
(72)発明者	岡田 充弘 東京都港区港南1丁目7番1号 株式会社ソニー・インタラクティブエンタテイ

最終頁に続く

(54)【発明の名称】 動画再生装置、動画再生システム、および動画再生方法

**(57)【特許請求の範囲】****【請求項1】**

サーバが再生中の、同じ空間を表す複数の動画像のデータを、時分割してなるセグメントデータの単位で時系列順に取得するデータ取得部と、

前記複数の動画像のデータのうち1つから音声データを取得し、別の動画像のデータから画像データを取得するデータ分離部と、

前記音声データの出力を継続させながら、前記画像データを同期させて出力する出力制御部と、

出力対象の画像データを、別の動画像のデータから取得したものに切り替える操作をユーザから受け付ける入力情報取得部と、

を備え、

前記入力情報取得部が前記切り替える操作を受け付けたとき、前記データ取得部は、前記サーバで再生中のセグメントから、切り替え後の動画像のデータの取得を開始し、

前記出力制御部は、前記音声データの出力を、出力可能となった状態で所定時間停止させてから開始するとともに、当該音声データのPTS (Presentation Time Stamp) が切り替わるタイミングで、前記画像データの出力を開始することを特徴とする動画再生装置。

**【請求項2】**

前記データ取得部は、前記セグメントデータの単位でPTSが付与された前記動画像のデータを取得し、

前記出力制御部は、前記サーバに繰り返して要求した前記動画像のデータのPTSが切り替わったときに取得したセグメントデータに含まれる音声データを、所定時間停止させてから出力開始することを特徴とする請求項1に記載の動画再生装置。

#### 【請求項3】

前記サーバに動画像のデータを要求し、取得したセグメントデータに付与されたPTSを検出する機能を複数備え、当該PTSが切り替わるタイミングの検出分解能を細分化することを特徴とする請求項2に記載の動画再生装置。

#### 【請求項4】

前記出力制御部は、1つ分のセグメントデータの再生時間に、前記サーバに前記動画像のデータを要求してからその画像が出力可能になるまでの時間を加算した時間だけ、前記音声データの出力を停止させることを特徴とする請求項1から3のいずれかに記載の動画再生装置。

10

#### 【請求項5】

前記出力制御部は、1つ分のセグメントデータの再生時間だけ前記音声データの出力を停止させることを特徴とする請求項1に記載の動画再生装置。

#### 【請求項6】

前記データ取得部は、出力対象外の動画像のデータも前記サーバから取得し、前記出力制御部は、前記出力対象外の動画像から取得した画像データを、前記セグメントデータの単位で出力可能な状態で停止させたうえ、前記切り替える操作がなされた際、停止させた画像データのうち切り替え後の画像データを出力することを特徴とする請求項1に記載の動画再生装置。

20

#### 【請求項7】

前記出力制御部は、前記切り替える操作に応じて、切り替え前の画像をフェードアウトさせるとともに、切り替え後の画像をフェードインさせることを特徴とする請求項1から6のいずれかに記載の動画再生装置。

#### 【請求項8】

前記データ取得部は、前記複数の動画像の時間軸のずれに係る情報を前記サーバから取得し、

前記出力制御部は、当該ずれに係る情報に基づき、前記音声データに対する前記画像データの出力開始タイミングを調整することを特徴とする請求項1から7のいずれかに記載の動画再生装置。

30

#### 【請求項9】

同じ空間を表す複数の動画像のデータをストリーム転送するサーバと、前記複数の動画像のデータを用いて動画像をディスプレイに出力する動画再生装置と、を備え、

前記動画再生装置は、

前記サーバが再生中の前記複数の動画像のデータを、時分割してなるセグメントデータの単位で時系列順に取得するデータ取得部と、

前記複数の動画像のデータのうち1つから音声データを取得し、別の動画像のデータから画像データを取得するデータ分離部と、

前記音声データの出力を継続させながら、前記画像データを同期させて出力する出力制御部と、

出力対象の画像データを、別の動画像のデータから取得したものに切り替える操作をユーザから受け付ける入力情報取得部と、

を備え、

前記入力情報取得部が前記切り替える操作を受け付けたとき、前記データ取得部は、前記サーバで再生中のセグメントから、切り替え後の動画像のデータの取得を開始し、

前記出力制御部は、前記音声データの出力を、出力可能となった状態で所定時間停止させてから開始するとともに、当該音声データのPTS(Presentation Time Stamp)が切り替わるタイミングで、前記画像データの出力を開始することを特徴とする動画再生システム。

40

50

**【請求項 10】**

サーバが再生中の、同じ空間を表す複数の動画像のデータを、時分割してなるセグメントデータの単位で時系列順に取得するステップと、

前記複数の動画像のデータのうち1つから音声データを取得し、別の動画像のデータから画像データを取得するステップと、

前記音声データの出力を継続させながら、前記画像データを同期させてディスプレイに出力するステップと、

出力対象の画像データを、別の動画像のデータから取得したものに切り替える操作をユーザから受け付けるステップと、

前記切り替える操作を受け付けたとき、前記サーバで再生中のセグメントから、切り替え後の動画像のデータの取得を開始するステップと、

を含み、

前記出力するステップは、前記音声データの出力を、出力可能となった状態で所定時間停止させてから開始するとともに、当該音声データのPTS (Presentation Time Stamp) が切り替わるタイミングで、前記画像データの出力を開始することを特徴とする、動画再生装置による動画再生方法。

**【請求項 11】**

サーバが再生中の、同じ空間を表す複数の動画像のデータを、時分割してなるセグメントデータの単位で時系列順に取得する機能と、

前記複数の動画像のデータのうち1つから音声データを取得し、別の動画像のデータから画像データを取得する機能と、

前記音声データの出力を継続させながら、前記画像データを同期させてディスプレイに出力する機能と、

出力対象の画像データを、別の動画像のデータから取得したものに切り替える操作をユーザから受け付ける機能と、

をコンピュータに実現させ、

前記受け付ける機能が前記切り替える操作を受け付けたとき、前記取得する機能は、前記サーバで再生中のセグメントから、切り替え後の動画像のデータの取得を開始し、

前記出力する機能は、前記音声データの出力を、出力可能となった状態で所定時間停止させてから開始するとともに、当該音声データのPTS (Presentation Time Stamp) が切り替わるタイミングで、前記画像データの出力を開始することを特徴とするコンピュータプログラム。

**【発明の詳細な説明】**

**【技術分野】**

**【0001】**

本発明は、ネットワークを利用して動画像を再生する動画再生装置、画像再生システム、および動画再生方法に関する。

**【背景技術】**

**【0002】**

画像処理技術の進歩やネットワーク環境の拡充により、用いる端末や環境によらず様々な動画像を気軽に楽しめるようになっている。動画像をストリーミング配信するサービスも普及し、様々な規格が策定されている。例えばHLS (HTTP Live Streaming) によれば、配信サーバは、動画像を数秒から数十秒の所定の長さで区切ってなるセグメントのデータと、そのデータの格納場所や再生順序などを規定するプレイリスト（またはインデックスファイル）を提供する。クライアントはまずプレイリストを取得し、必要なセグメントデータをサーバに要求することにより動画を再生する（例えば、特許文献1参照）。同様の規格としてMP4-DASH (Dynamic Adaptive Streaming over HTTP) やCMASF (Common Media Application Format) などがある（例えば、特許文献2、3参照）。

**【先行技術文献】**

10

20

30

40

50

**【特許文献】****【0003】**

【文献】特開2018-93349号公報

特開2017-108217号公報

国際公開第2017/204109号

**【発明の概要】****【発明が解決しようとする課題】****【0004】**

動画のストリーミング配信は、複数のクライアントで共通の動画を鑑賞できるようにすることを基本としている。一方、上記のような技術により、ランダムアクセスでの再生や、ネットワーク環境に応じた画質レベルの切り替えなど、クライアント個々の事情に合わせた柔軟性も実現されつつある。今後、視聴者それぞれの好みや意図に応じて、より多様な操作を可能にすることが望まれる。

10

**【0005】**

本発明はこうした課題に鑑みてなされたものであり、その目的は、ストリーミング配信される動画像に対するユーザ操作を好適に実現する技術を提供することにある。

**【課題を解決するための手段】****【0006】**

本発明のある態様は動画再生装置に関する。この動画再生装置は、サーバからストリーム転送される、同じ空間を表す複数の動画像のデータを取得するデータ取得部と、複数の動画像のうち1つから音声データを取得し、別の動画像から画像データを取得するデータ分離部と、音声データと前記画像データを同期させて出力する出力制御部と、を備えたことを特徴とする。

20

**【0007】**

本発明の別の態様は動画再生システムに関する。この動画再生システムは、同じ空間を表す複数の動画像のデータをストリーム転送するサーバと、複数の動画像のデータを用いて動画像をディスプレイに出力する動画再生装置と、を備え、動画再生装置は、サーバから複数の動画像のデータを取得するデータ取得部と、複数の動画像のうち1つから音声データを取得し、別の動画像から画像データを取得するデータ分離部と、音声データと前記画像データを同期させて出力する出力制御部と、を備えたことを特徴とする。

30

**【0008】**

本発明のさらに別の態様は動画再生方法に関する。この動画再生方法は動画再生装置が、サーバからストリーム転送される、同じ空間を表す複数の動画像のデータを取得するステップと、複数の動画像のうち1つから音声データを取得し、別の動画像から画像データを取得するステップと、音声データと画像データを同期させてディスプレイに出力するステップと、を含むことを特徴とする。

**【0009】**

なお、以上の構成要素の任意の組合せ、本発明の表現を方法、装置、システム、コンピュータプログラム、コンピュータプログラムを記録した記録媒体などの間で変換したものまた、本発明の態様として有効である。

40

**【発明の効果】****【0010】**

本発明によると、ストリーミング配信される動画像に対するユーザ操作を好適に実現できる。

**【図面の簡単な説明】****【0011】**

【図1】本実施の形態を適用できる動画再生システムを例示する図である。

【図2】本実施の形態で動画再生装置にストリーミング配信される動画像の例を説明するための図である。

【図3】本実施の形態において配信対象とする動画像のデータ構造を説明するための図で

50

ある。

【図4】画像と音声の同期を実現するためのシステムの構成を示す図である。

【図5】本実施の形態における動画再生装置の内部回路構成を示す図である。

【図6】本実施の形態における動画再生装置および動画配信サーバの機能ブロックの構成を示す図である。

【図7】本実施の形態における画像と音声の出力の流れを例示する図である。

【図8】本実施の形態において表示画像を切り替える際の動画再生装置の処理手順を示すフロー チャートである。

【図9】図7の制御方式で切り替えに時間を要するケースの、画像と音声の出力の流れを例示する図である。

10

【図10】本実施の形態において、表示中の画像以外の画像を常に出力可能な状態にした場合の、画像と音声の出力の流れを例示する図である。

【図11】本実施の形態において音声出力に与える遅延時間を適応的に決定する場合の、画像と音声の出力の流れを例示する図である。

【図12】本実施の形態において音声出力に与える遅延時間を適応的に決定する場合の、画像と音声の出力の流れを例示する図である。

【図13】本実施の形態において、1つのコンテンツを構成する複数の動画像の時間軸が直列を有する例を模式的に示す図である。

【図14】本実施の形態における動画再生装置が、動画配信サーバが提供する動画像の時間ずれに対応するように出力タイミングを調整する際の、画像と音声の出力の流れを例示する図である。

20

#### 【発明を実施するための形態】

##### 【0012】

図1は本実施の形態を適用できる動画再生システムを例示している。図示する動画再生システムは複数の動画再生装置10a、10b、10c、…がネットワーク8を介して動画配信サーバ12に接続した構成を有する。動画再生装置10a、10b、10c、…はそれぞれ、ユーザが操作するクライアント端末であり、入力装置14a、14b、14c…と、ディスプレイ16a、16b、16c、…が有線または無線で接続される。

##### 【0013】

以後、動画再生装置10a、10b、10c、…を動画再生装置10、入力装置14a、14b、14c…を入力装置14、ディスプレイ16a、16b、16c、…をディスプレイ16と総称する場合がある。動画再生装置10、入力装置14、ディスプレイ16は、図示するようにそれぞれが別の筐体を有していてもよいし、それらの2つ以上が一体的に設けられていてもよい。例えば動画再生装置10、入力装置14、ディスプレイ16を一体的に備える携帯端末などであってもよい。

30

##### 【0014】

またディスプレイ16はテレビ受像器など一般的な平板型ディスプレイでもよいし、ヘッドマウントディスプレイなどのウェアラブルディスプレイでもよい。いずれにしろディスプレイ16は、画像を表示させる表示パネルと、音声を出力するスピーカを備えるものとする。ただしスピーカーは、ディスプレイ16とは別に設けてよい。動画再生装置10は、パーソナルコンピュータ、ゲーム機、コンテンツ再生装置などのいずれでもよい。ネットワーク8はインターネットやLAN(Local Area Network)など、その規模は限定されない。

40

##### 【0015】

いずれにしろ本実施の形態では、ユーザ操作に基づき動画再生装置10が動画配信サーバ12へ動画像の配信を要求し、動画配信サーバ12が要求された動画像をストリーミング配信することを基本とする。その限りにおいて、用いられる通信プロトコル、動画再生装置10の形態、動画配信サーバ12の構成などは特に限定されない。例えば動画配信サーバ12は、録画済みの動画像を配信してもよいし、撮影中あるいは作成中の動画像をラ

50

イプ配信してもよい。この際、動画配信サーバ12は、別のコンテンツ提供サーバと接続し、動画像のデータを取得したうえで動画再生装置10に送信してもよい。

#### 【0016】

図2は、本実施の形態で動画再生装置10にストリーミング配信される動画像の例を説明するための図である。この例では、コンサート会場18に複数のカメラ20a、20bを設置し、コンサートの様子を異なる方向から撮影して動画配信することを想定している。例えばコンサートの開始から終了までを、複数のカメラ20a、20bで並行して撮影することにより、共通の時間軸で視野の異なる複数の動画像が取得される。

#### 【0017】

動画配信サーバ12は、そのような複数の動画像のうちいずれかを動画再生装置10に配信するとともに、視野の異なる動画像への配信対象の切り替えを、動画再生装置10から隨時受け付ける。すなわち本実施の形態では、動画再生中の任意のタイミングで、別の視野の画像へ表示を切り替えることを可能にする。これにより動画再生装置10側で動画像を見ているユーザは、自分が見たい演者をメインに撮影した動画や、会場全体を俯瞰した動画などを、コンサートの進行状況などに応じて自由に切り替えることができる。

10

#### 【0018】

なお同一空間を異なる視野で表した動画像であれば、表示対象や表示目的は特に限定されない。例えば図示するようなコンサートの映像のほか、スポーツ競技や各種イベントなどの映像でもよいし、実写画像に限らず、仮想空間を別の視野で表したコンピュータグラフィックスなどでもよい。また共通の時間軸上で同期がとれていれば、一部の期間のみ別の視野の動画像へ切り替えられるようにしてもよく、切り替え可能な動画像の数も限定されない。

20

#### 【0019】

図3は、本実施の形態において配信対象とする動画像のデータ構造を説明するための図である。なお以後の説明において、「動画像」は画像と音声を含むものとする。また、切り替え可能な複数の動画像のまとめを「コンテンツ」とする。上述のとおり本実施の形態では、1つの空間を表す、視野の異なる動画像を複数準備することにより、動画再生装置10において表示される視野を任意のタイミングで切り替えられるようになる。この際、表示上の画像を切り替えて、音声については1つの動画像のものを出力しつづけるようにして、シームレスな切り替えを実現する。

30

#### 【0020】

例えば音声再生用の動画像と表示用の複数の動画像を別に準備し、前者の音声を出力する一方、後者のうちいずれかの画像を表示する。視野の切り替え操作に対しては、後者のうちいずれかの画像に表示対象を切り替える。あるいは表示用の複数の動画像のうちいずれか1つを、音声再生用として兼用してもよい。以後の説明では、音声再生用の動画像を別途準備する態様に主眼を置く。この場合、当該動画像の画像自体は表示されないため、低ビットレートで表すことによりデータサイズを抑えることができる。

#### 【0021】

一方、表示用、音声再生用に関わらず同様のデータ構造の動画像として準備することにより、動画配信サーバ12は、従来と同様のプロトコルでデータ送信が可能となる。ここで各動画像のデータは、数秒から数十秒程度の所定時間ごとに分割した状態で保持、送信される。以後、分割後のデータのそれぞれを「セグメントデータ」と呼ぶ。動画配信サーバ12は、動画像ごとに、それを分割してなる複数のセグメントデータと、各セグメントデータの定義情報であるプレイリストを生成する。

40

#### 【0022】

図において、音声再生用の動画像のプレイリスト132は、音声再生用のセグメントデータ136のそれぞれの格納場所、再生時間、再生順序などを定義する。同様に複数の表示用動画像のプレイリスト134a、134bはそれぞれ、表示用のセグメントデータ138a、138bのそれぞれの格納場所、再生時間、再生順序などを定義する。録画済みの動画の場合、各プレイリストは静的なデータである。撮影中の動画の場合、時間経過に

50

対し新たなセグメントデータが発生し、それに応じて各プレイリストが更新される。

#### 【0023】

動画配信サーバ12はさらに、それらの動画像を対応づけ、1つのコンテンツとして定義するインデックスファイル130を生成する。インデックスファイル130は、コンテンツの内容、当該コンテンツとして準備されている複数の動画像のプレイリスト132、134a、134bの格納場所や視野情報など、各動画像に係る情報を記載する。動画像を時分割して配信する技術として、HLS、MPEG-DASH、CMAFなどいくつかの規格が実用化されており、本実施の形態ではそのいずれを採用してもよい。規格によって、準備するファイルの呼称や記載形式が様々であることは当業者には理解されるところである。

10

#### 【0024】

いずれにしろ本実施の形態で動画配信サーバ12は、動画再生装置10が指定するコンテンツのうち、音声再生用の動画像のデータと、要求された視野の表示用の動画像のデータを、インデックスファイル130、プレイリスト132、134a、134bを辿ることにより特定し、送信する。詳細には動画配信サーバ12はまず、必要な動画像のプレイリストを動画再生装置10に送信し、そのうち必要な時間におけるセグメントデータを指定した送信要求を受け付けることにより、当該セグメントデータを動画再生装置10に送信する。

#### 【0025】

各セグメントデータは、所定時間ごとの画像データと音声データを含む。動画配信サーバ12は、送信対象のセグメントデータを時系列順にパケット化して送信する。詳細には図の右側に示すように、画像データのパケットと音声データのパケットが、多重ストリームの形式で順番に送出される。本実施の形態では特に、音声再生用の動画像のデータ列(ストリーム)138が継続的に送信される一方、表示用の動画像のデータ列139a、139bは、ユーザによる切り替え操作に応じて切り替えて送信される。

20

#### 【0026】

一方、動画再生装置10は、要求した動画像のパケットを順次取得すると、画像と音声にずれが生じないように、両者を同期させて出力する。本実施の形態では、異なる動画像の画像と音声を組み合わせて出力するが、その同期のさせ方には、1つの動画像の画像と音声を同期させる従来のしくみを利用できる。図4は、画像と音声の同期を実現するためのシステムの構成を示している。まず動画配信サーバ12側では、自らが有するSTC(System Time Clock)をベースとして、画像および音声の出力タイミングを規定するPTS(Presentation Time Stamp)を生成し、それぞれのセグメントデータに付与する。

30

#### 【0027】

動画配信サーバ12はまた、STCをベースとした所定周期でのカウンタ値を表すPCR(Program Clock Reference)を生成する。動画配信サーバ12のシステム符号化部140は、画像や音声のセグメントデータとそれぞれのPTS、および所定周期でのPCRを含めたパケットの列からなる多重化ストリームを生成し、動画再生装置10に送信する。動画再生装置10のシステム復号部142は、送信された多重化ストリームからそれらのデータを分離する。そしてSTC再生部144は、各パケットの受信時刻とPCRが示すカウンタ値を対応づけるように発振器の周波数を調整することによりSTCを再生する。

40

#### 【0028】

これにより動画再生装置10は、動画配信サーバ12と共に通化された時間軸上での動作を実現する。具体的には動画再生装置10は、当該時間軸上でのPTSにおいて、対応する画像および音声が出力されるように、バッファ146a、146bによって各データの出力タイミングを調整する。1つのコンテンツを構成する全ての動画像に対し同じ時間軸でのPTSが付与されれば、それに基づき出力タイミングを調整することで、異なる動画像であっても画像と音声をずれなくディスプレイ16に出力できる。なお本図では、データの符号化処理と復号処理については図示を省略している。

50

**【 0 0 2 9 】**

図5は、動画再生装置10の内部回路構成を示している。動画再生装置10は、C P U (Central Processing Unit) 23、G P U (Graphics Processing Unit) 24、メインメモリ26を含む。これらの各部は、バス30を介して相互に接続されている。バス30にはさらに入出力インターフェース28が接続されている。入出力インターフェース28には、U S BやI E E E 1 3 9 4などの周辺機器インターフェースや、有線又は無線L A Nのネットワークインターフェースからなり、動画配信サーバ12と通信を確立する通信部32、ハードディスクドライブや不揮発性メモリなどの記憶部34、ディスプレイ16へデータを出力する出力部36、入力装置14からデータを入力する入力部38、磁気ディスク、光ディスクまたは半導体メモリなどのリムーバブル記録媒体を駆動する記録媒体駆動部40が接続される。

10

**【 0 0 3 0 】**

C P U 2 3は、記憶部34に記憶されているオペレーティングシステムを実行することにより動画再生装置10の全体を制御する。C P U 2 3はまた、リムーバブル記録媒体から読み出されてメインメモリ26にロードされた、あるいは通信部32を介してダウンロードされた各種プログラムを実行する。G P U 2 4は、ジオメトリエンジンの機能とレンダリングプロセッサの機能とを有し、C P U 2 3からの描画命令に従って描画処理を行い、出力部36に出力する。メインメモリ26はR A M (Random Access Memory)により構成され、処理に必要なプログラムやデータを記憶する。なお動画配信サーバ12も同様の回路構成としてよい。

20

**【 0 0 3 1 】**

図6は、動画再生装置10および動画配信サーバ12の機能ブロックの構成を示している。同図に示す各機能ブロックは、ハードウェア的には図5で示したC P U 2 3、G P U 2 4、メインメモリ26などで実現でき、ソフトウェア的には、記録媒体からメモリにロードした、情報処理機能、画像描画機能、データ入出力機能、通信機能などの諸機能を発揮するプログラムで実現される。したがって、これらの機能ブロックがハードウェアのみ、ソフトウェアのみ、またはそれらの組合せによっていろいろな形で実現できることは当業者には理解されるところであり、いずれかに限定されるものではない。

**【 0 0 3 2 】**

動画配信サーバ12は、動画再生装置10からの要求を取得する要求取得部50、要求に従ったデータを準備するデータ準備部52、コンテンツのデータを記憶するデータ記憶部54、および、準備されたデータを動画再生装置10に送信するデータ送信部56を含む。要求取得部50は、表示対象の切り替えを含む、動画像の送信要求を動画再生装置10から取得する。そのため要求取得部50は、選択可能なコンテンツやそれを表す画像の視野に係る情報など、動画像の選択に必要な情報をあらかじめ動画再生装置10に送信してもよい。

30

**【 0 0 3 3 】**

データ準備部52は要求取得部50と連携し、動画再生装置10から取得した要求の内容に応じたデータを準備する。データ準備部52は例えば、選択されたコンテンツに対応するインデックスファイルを取得し、それを表す画像（視野）の選択肢を特定する。この情報を要求取得部50に与えることにより、動画再生装置10からさらに、視野を指定した動画像の転送要求を得る。これに応じてデータ準備部52は、対応するプレイリストを取得し、要求取得部50を介して動画再生装置10に送信することで、必要なセグメントデータの要求を受け付ける。

40

**【 0 0 3 4 】**

例えば動画像を最初から再生する場合、データ準備部52は、音声再生用の動画像と、動画再生装置10が指定した視野の、表示用の動画像のセグメントデータを、それぞれのプレイリストの先頭に記載されているものから順に取得していく。途中で表示対象が切り替えられた場合、データ準備部52は、音声再生用の動画像については引き続き後続のセグメントデータを取得するとともに、表示用の動画像については、切り替え後の動画像の

50

プレイリストに記載された、対応する時刻以降のセグメントデータを取得する。

#### 【0035】

データ記憶部54は図3に示したように、インデックスファイル、複数の動画像のプレイリストおよびそのセグメントデータを、コンテンツごとに格納する。ここで各データは、適宜圧縮符号化されているとする。また撮影中の動画像の場合、それらのデータは随時更新されるが、その更新手段については図示を省略している。データの元となる動画像は、図示しない別のサーバなどから取得してもよく、その取得タイミングは特に限定されない。

#### 【0036】

データ送信部56は、データ準備部52が準備したセグメントデータを順次パケット化して動画再生装置10に送信する。この際、データ送信部56は、上述のとおりセグメントデータ単位など所定の単位で、画像および音声のデータにPTSを付与するとともに、周期的にPCRも付与する。データ送信部56はそのほか、一般的なストリーミング転送において付与される情報を適宜付与してよい。

10

#### 【0037】

動画再生装置10は、ユーザ操作の内容を取得する入力情報取得部60、動画像のストリームを取得するデータ取得部62、ストリームからデータを分離するデータ分離部64、画像データを復号する画像復号部66、音声データを復号する音声復号部68、動画像の出力を制御する出力制御部70を含む。入力情報取得部60は、コンテンツの選択、表示画像の選択、および途中での表示対象の切り替えなどのユーザ操作の内容を入力装置14から取得し、必要なデータを動画配信サーバ12に要求する。

20

#### 【0038】

上述のとおり入力情報取得部60はあらかじめ、選択可能なコンテンツやそれを表す画像に係る情報を動画配信サーバ12から取得し、出力制御部70を介してディスプレイ16に選択肢として表示させててもよい。入力情報取得部60はまた、音声再生用の動画像や、表示用に選択された動画像のプレイリストを動画配信サーバ12から取得し、再生開始操作や表示切り替え操作がなされた時刻に対応するセグメントデータを動画配信サーバ12に要求する。

#### 【0039】

データ取得部62は、ユーザ操作に応じて動画配信サーバ12からストリーミング転送された動画像のデータを継続的に取得していく。当該データには、音声再生用の動画像と表示用の動画像のデータが含まれる。各動画像のストリームには、所定の単位でPTSが付与された、画像データと音声データが含まれる。データ分離部64は、そのような多重化されたデータを動画像ごとに分離し、さらに画像データと音声データに分離する。動画配信サーバ12から送信されるストリームには、それらのデータを識別するための情報が含まれる。当該情報を用いたデータ分離には、従来実用化されている手法を適用できる。

30

#### 【0040】

画像復号部66は、分離されたデータのうち、ユーザが選択した表示対象の動画像に含まれる画像データを復号する。音声復号部68は、分離されたデータのうち、音声再生用の動画像に含まれる音声データを復号する。出力制御部70は、復号された画像と音声を、適切なタイミングで順次、ディスプレイ16に出力する。詳細には出力制御部70は、PTS検出部72、時間調整部74、画像加工部76、および出力部78を含む。PTS検出部72は、出力対象の画像データおよび音声データに付与されているPTSを検出する。

40

#### 【0041】

時間調整部74は、画像と音声にずれが生じないよう出力タイミングを調整する。一般的なストリーミング転送では、図4に示すように、画像と音声のデータがほぼ同時期に送信され、パケット間の僅かな時間差を調整する以外は即時の出力を基本とする。一方、本実施の形態では、異なる動画像に含まれる画像と音声を同期させて出力するとともに、途中で別の動画像に含まれる画像へ表示を切り替えることを許容する。この切り替えにおいても画像と音声の同期を維持しながら、切り替え前後の画像ができるだけ途切れないよう

50

にすることにより、1つのコンテンツとしての連続性を表すことができる。

#### 【0042】

そこで時間調整部74は、音声再生用動画像の音声の出力を、動画配信サーバ12からのデータ取得タイミングに対し所定時間遅らせたうえで、そのPTSに合うように画像を出力する。これにより、画像の切り替え操作がなされたときに表示していた画像のPTSと、動画配信サーバ12から送信される、切り替え後の画像の先頭のPTSの差を小さくできるとともに、出力されつづける音声と当該先頭の画像とのズレが生じないようにできる。時間調整の具体例は後に述べる。

#### 【0043】

画像加工部76は、画像の切り替え操作に応じて、表示中の画像をフェードアウトさせるとともに、切り替え後の画像をフェードインさせる加工を行う。これにより、切り替え操作に対しシームレスな表示変化を実現する。出力部78は、表示用動画像の画像をディスプレイ16の表示パネルに出力するとともに、音声再生用の動画像の音声を、ディスプレイ16のスピーカーに出力する。

10

#### 【0044】

図7は、本実施の形態における画像と音声の出力の流れを例示している。図の横方向は時間経過を表し、各矩形はセグメントデータの時間長を示している。矩形内の数字はPTSを表象しており、簡易的に同じPTSと同じ値の自然数で表しているが、実際のPTSのフォーマットを限定するものではない。最上段は動画配信サーバ12において再生されている動画データの流れであり、システムクロックに対する共通のPTSが、1つのコンテンツを表す全ての動画像のセグメントデータに同様に与えられる。

20

#### 【0045】

そのように再生された動画像は、並行して複数の動画再生装置10に配信される。ここである動画再生装置10が矢印150のタイミングで動画配信を要求すると、動画配信サーバ12は、当該要求を受信した時点で再生中の、PTS「1」のセグメントから、動画像データの送信を開始する。図示する例では、まず音声再生用の動画像のデータのみが送信されている。送信要求のタイミング、データ転送、復号処理などに依存して、動画再生装置10においてPTS「1」のセグメントデータが出力可能となる時刻t1は、動画配信サーバ12における対応するデータの再生開始時刻t0より遅延する。

30

#### 【0046】

従来技術においては、出力可能となった時刻t1において即時に音声や画像を出力するが、本実施の形態における時間調整部74は、出力可能となった状態で音声出力を一時停止させたうえ、所定時間後の時刻t2において一時停止を解除し出力を開始させる。この例で一時停止させる時間(t2-t1)は、1分のセグメントデータの時間とし、例えば3秒間などである。このように音声の出力に猶予を持たせることにより、後の表示画像の切り替えにおいて、切り替え後の画像が音声に間に合うようにする。

#### 【0047】

図示する例では、動画配信サーバ12がまず音声出力用の動画像データのみを送信することにより、時刻t2の時点では、動画再生装置10からPTS「1」の音声のみが出力される。ここで動画再生装置10が第1の表示用の動画像を要求すると、矢印152のように、その時点で再生されているPTS「2」のセグメントから当該表示用動画像のデータが送信される。表示用の動画像についても、送信要求のタイミング、データ転送、復号処理などにより、PTS「2」の先頭の画像が出力可能となる時刻t3は、動画配信サーバ12における対応するデータの再生開始時刻より遅延する。

40

#### 【0048】

一方、動画再生装置10においてPTS「1」のセグメントの音声出力を所定時間、遅延させることにより、時刻t3は、PTS「2」の音声データの出力開始時刻t4よりは前になる。そこで時間調整部74は、PTS「2」の第1の画像が出力可能となった状態で一時停止させたうえ、PTS「2」の音声データの出力が開始される時刻t4で、同じPTS「2」の第1の画像の出力を開始させる。これにより、元の動画像データが異なっ

50

ていても、画像と音声にずれのない状態で動画像を表すことができる。また図示するよう後にから画像のデータを取得しても、先に出力していた音声に間に合わせることができる。

#### 【0049】

なお画像の出力開始に当たっては、画像加工部76がフェードインさせることにより唐突感を軽減させる。このようにして画像と音声を出力している状態で、時刻t5において、ユーザが別の画像へ表示を切り替える操作を行ったら、動画再生装置10は、新たに選択された第2の表示用の動画像を動画配信サーバ12に要求するとともに、それまで表示していた第1の画像の出力を適当なタイミングで停止させる。この際、画像加工部76は第1の画像をフェードアウトさせる。

#### 【0050】

図示する例では、PTS「4」の画像がフェードアウトする。一方、動画配信サーバ12からは矢印154のように、要求された時点で再生中のPTS「6」のセグメントから、切り替え後の表示用動画像のデータが送信される。ここでも動画再生装置10においてPTS「1」のセグメントの音声出力を所定時間、遅延させたことにより、第2の動画像のうちPTS「6」の先頭の画像が出力可能となる時刻t6は、PTS「6」の音声データの出力開始時刻t7よりは前になる。

10

#### 【0051】

そこで時間調整部74は、PTS「6」の音声データの出力が開始される時刻t7まで、PTS「6」の第2の画像の出力を一時停止させる。そして時刻t7において一時停止を解除し、画像加工部76、出力部78により、PTS「6」の第2の画像をフェードインさせながら出力を開始する。以後同様に、表示対象を切り替える操作がなされる都度、表示中の画像をフェードアウトさせ、切り替え後の画像を出力可能な状態で一時停止させたうえで、同じPTSの音声が出力されるのを待ってフェードインさせる。

20

#### 【0052】

図8は、表示画像を切り替える際の動画再生装置10の処理手順を示すフローチャートである。このフローチャートは、音声再生用の動画像から得られる音声と、表示用の動画像から得られる画像をディスプレイ16に出力している状態において開始される。この状態で入力情報取得部60は、表示画像を切り替えるユーザ操作を待機する(S10のN)。当該ユーザ操作がなされたら(S10のY)、出力制御部70は出力中の画像をフェードアウトさせて出力を停止する(S12)。

30

#### 【0053】

一方、入力情報取得部60が、送信対象の表示用動画像の切り替えを動画配信サーバ12に要求することにより、データ取得部62は、切り替え後の動画像データのうち、動画配信サーバ12において要求時に再生されていたセグメントからデータ取得を開始する(S14)。そして画像復号部66は、当該動画像データから取り出された画像のセグメントデータの復号を開始する(S16)。出力制御部70は、復号された画像の出力を一時停止させるとともに(S18)、そのPTSを検出する(S20)。

#### 【0054】

そして出力制御部70は、検出したPTSと、出力中の音声のPTSを比較をしていく。両者が異なっているうちは、画像の出力を一時停止させたままとする(S22のN)。結果として、ディスプレイ16に何も表示されないブラックアウトの状態が短時間発生し得る。PTSが一致したら(S22のY)、出力制御部70は画像出力の一時停止を解除するとともに(S24)、画像をフェードインさせながら表示させる(S26)。なおS24の一時停止の解除は厳密には、画像と音声のPTSが一致するタイミングを予測して、その直前に実施してよい。

40

#### 【0055】

図7で示した態様は、音声出力を遅延させる時間を、1つのセグメント時間に固定していた。この制御方式によれば、動画配信サーバ12での再生時間と動画再生装置10での表示時間との間に、少なくとも1セグメント分のずれが生じるため、動画配信サーバ12から送信される、切り替え後の画像データの先頭は、切り替え操作時に表示されていた

50

セグメントより後のセグメントとなることが保証される。結果として、切り替え後の画像の表示開始が音声に間に合わなくなるのを回避できる。

#### 【 0 0 5 6 】

一方、この制御方式では、表示開始操作のタイミングによっては、動画配信サーバ12での再生時間と動画再生装置10での表示時間との間に必要以上のずれが生じ、切り替え後の画像の表示開始までに余計な待機時間を要してしまう可能性がある。図9は、図7の制御方式で切り替えに時間を要するケースの、画像と音声の出力の流れを例示している。図の表し方は図7と同様であるが、この例では、動画配信サーバ12における動画の再生時間に対し、動画再生装置10における出力時間の遅延が、図7のケースより大きくなっている。

10

#### 【 0 0 5 7 】

このようなずれの大きさは、表示開始操作がなされたタイミングによって変化する。すなわち図9の場合、動画配信サーバ12においてPTS「1」のセグメントの再生終了間際に開始操作がなされたため、当該セグメントのデータが動画再生装置10において出力可能となる時刻t9は、動画配信サーバ12における再生開始時刻t8より、およそ1セグメント分、すでに遅延している。このような場合にも、さらに1セグメント分の固定の遅延時間と与えてから音声出力を開始すると、結果的には動画配信サーバ12と動画再生装置10において、動画再生に2セグメント分のずれが生じることになる。

#### 【 0 0 5 8 】

例えば時刻t10において画像の切り替え操作がなされたとき、動画再生装置10において第1の画像のPTS「4」のセグメントが表示されているのに対し、動画配信サーバ12では、それより2つ後の、PTS「6」のセグメントが再生されている。そのため動画再生装置10は、切り替え後の第2の動画像データのうちPTS「6」のセグメントからデータを取得せざるを得ない。結果として、当該画像が出力可能になった時刻t12から、PTS「6」の音声データの出力開始時刻t13までの時間が、1セグメント分の時間より長くなってしまう。

20

#### 【 0 0 5 9 】

切り替え前の第1の画像の表示は、時刻t12より前の時刻t11ですでに停止していることから、時刻t11から時刻t13までの比較的長い期間、何も表示されないブラックアウトの状態となる。このような状況を回避するための方策の一つとして、表示中の画像以外の画像を、セグメント単位で常に出力可能な状態にしておくことが考えられる。

30

#### 【 0 0 6 0 】

図10は、表示中の画像以外の画像を常に出力可能な状態にした場合の、画像と音声の出力の流れを例示している。図の表し方は図7と同様である。また動画配信サーバ12から音声再生用の動画像データを取得して所定の遅延時間を設けてから音声を出力し、第1の表示用動画像のデータを取得して音声に対応するタイミングで第1の画像の出力を開始する流れも図7と同様である。一方、この制御方式では、画像の切り替え操作にかかわらず、第2の表示用動画像も並行して取得する。

#### 【 0 0 6 1 】

すなわち動画再生装置10の入力情報取得部60は、矢印156a、156b、156cに示すように、出力中の動画像におけるセグメントの切り替わりのタイミングで、表示されていない第2の表示用動画像のセグメント単位での送信を、動画配信サーバ12に要求する。そして動画再生装置10は、画像の切り替え操作がなされない場合は第1の画像の出力を継続するとともに、第2の画像データを投機的に復号し、出力可能な状態にしておく。

40

#### 【 0 0 6 2 】

図示する例では、第2の画像のうちPTS「3」の先頭の画像が時刻t14で出力可能となり、PTS「4」の先頭の画像が時刻t15で出力可能となっている。それぞれの画像の出力は、同じPTSの音声の出力が開始されるまで一時停止とし、その直前のセグメント出力期間に画像を切り替えるユーザ操作がなされなければデータを破棄する。一時停

50

止している第2の画像のPTSの直前のセグメントの出力期間に切り替え操作がなされたら、一時停止を解除して当該第2の画像の出力を開始する。

#### 【0063】

図示する例では、第2の画像のうちPTS「5」のデータを一時停止させる一方、その直前のPTS「4」の第1の画像の出力期間t16～t18中の時刻t17に、画像を切り替える操作がなされたことから、時刻t18においてPTS「5」の第2の画像データを破棄せず出力している。この場合も出力制御部70は、切り替え前の第1の画像のうちPTS「4」の画像をフェードアウトさせて停止させたうえで、切り替え後の第2の画像をフェードインさせる。

#### 【0064】

このように、表示対象でない動画像のデータも並行して取得したうえ、セグメントごとに先頭の画像を常に準備しておくことにより、動画配信サーバ12と動画再生装置10における出力時間のずれによらず、各セグメントの先頭で切り替えが可能になる。結果として切り替えに要する時間を最小限にできる。一方、この方式では、1つのコンテンツを表す動画像が増えるほど、また、画像の解像度が上がるほど、データ転送や復号処理に負荷がかかることになる。したがって好適には、選択肢となる動画像の数と、通信環境や動画再生装置10の処理性能などに応じて適用の是非を決定する。

10

#### 【0065】

切り替え時間を最小限にする方策の別の例として、音声出力に与える遅延時間を適応的に決定することが考えられる。図11、12は、音声出力に与える遅延時間を適応的に決定する場合の、画像と音声の出力の流れを例示している。図の表し方は図7と同様である。この方式では初期処理として、セグメントの先頭データを取得できる最速のタイミングと、表示用動画像の要求から出力可能となるまでの時間を実測する。

20

#### 【0066】

このため図11に示すように、動画再生装置10において、音声再生用の動画像を要求し、その音声データを復号してPTSの検出を行う機能ブロックを複数、準備する。すなわち入力情報取得部60、データ取得部62、データ分離部64、音声復号部68、PTS検出部72、時間調整部74からなる機能ブロックの組を複数設ける。これにより、動画再生装置10においてPTSが切り替わる最速のタイミングの検出単位を細分化する。図示する例では、当該機能ブロックを2つ準備し、それぞれ「音声第1再生」、「音声第2再生」としている。

30

#### 【0067】

まず音声第1再生を行うブロックが、矢印158aのように音声再生用動画像のデータを要求すると、動画配信サーバ12からはその時点で再生中のPTS「1」のセグメントからデータが送信される。音声第1再生を行うブロックはその先頭を復号して一時停止させるとともに、PTSを検出する。その間に、音声第2再生を行うブロックが、矢印158bのように音声再生用動画像のデータを要求し、動画配信サーバ12からその時点で再生中のセグメントデータが送信されたら、その先頭を復号して一時停止させ PTSを検出する。

40

#### 【0068】

図示する例ではこの時点で検出されるPTSも「1」である。このような処理を交互に繰り返すことにより、いずれは検出されるPTSが切り替わる。図では矢印158cの要求に応じて送信されたデータにおいて、PTSが「2」に切り替わっている。音声第1再生および音声第2再生の機能ブロックは、PTSの切り替わりが発生するまで、一時停止させた音声データを破棄しながら、次のデータ要求を繰り返す。そのような機能ブロックを2つ以上準備することにより、切り替わりのタイミングを細かい粒度で検出できる。

#### 【0069】

さらに動画再生装置10は、矢印158dのように表示用の動画像を動画配信サーバ12に要求し、要求された時点で再生中のセグメントデータを動画配信サーバ12から取得する。図ではPTS「2」のデータから送信されている。動画再生装置10の画像復号部

50

66は、送信されたデータの復号を開始し、出力制御部70は先頭画像の出力が可能な状態で一時停止させる。この際、時間調整部74は、動画配信サーバ12へのデータ要求から画像の出力が可能になるまでの時間 $t_d$ を、内部のタイマにより計測する。

#### 【0070】

図11は初期処理のみを示しているが、動画再生装置10はそれに続けて図12に示すように動画像の出力処理を実施する。すなわち、初期処理により動画配信サーバ12との時間差が最小の状態で音声再生用動画像を取得できるため、時間調整部74は、先頭のPTS「2」のデータを一時停止させた状態で所定時間、停止させる。この際の停止時間は、1つのセグメントデータの出力時間に、初期処理で取得した、データ要求から画像の出力が可能になるまでの時間 $t_d$ を加算した値とする。

10

#### 【0071】

表示用の動画像データも初期処理で取得し、時間調整部74が一時停止させたPTS「2」の画像を用い、出力部78がPTS「2」の音声出力とともに出力する。この制御方式によれば、図10のように表示対象外の動画像も全て準備せずとも、常に短時間での切り替えを実現できる。すなわち図の「第2画像生成」に示すように、表示の切り替え操作がどのタイミングでなされても、切り替え前のPTSの音声が出力されている期間に、切り替え後のPTSの画像を準備し終えることができ、当該画像を次のPTSの切り替えのタイミングで出力できる。

#### 【0072】

図示する例では最下段に示すように、動画配信サーバ12においてPTS「6」の再生終了間際に切り替え操作がなされても、動画再生装置10において、同じPTS「6」の音声出力に、切り替え後の画像出力を間に合わせることができている。したがって、1つのコンテンツを構成する動画像の数や解像度の高さによらず同程度の処理負荷で、最短時間での切り替えが可能になる。なお音声出力に設ける遅延時間には、実際にはいくらかのマージンを加算しておいてもよい。また画像表示の開始や停止において、画像加工部76がフェードイン、フェードアウトさせる加工を行う点などは上述と同様とする。

20

#### 【0073】

なお図11では第1の表示用画像についてのみ、動画配信サーバ12にデータを要求してから出力可能となるまでの時間 $t_d$ を計測したが、コンテンツを構成する全ての表示用動画像について同じ計測を行ってもよい。例えば表示用動画像の画像サイズやビットレートが異なる場合、データを要求してから出力可能となるまでの時間が異なることが考えられる。この場合、動画像ごとに測定した時間 $t_d$ のうち最も長い時間を採用して、音声出力に与える遅延時間に加算することにより、切り替え後の画像を音声に間に合わせることを保証できる。

30

#### 【0074】

さらにネットワーク8の状況によっては、同じ動画像であっても、動画配信サーバ12にデータを要求してから出力可能となるまでの時間 $t_d$ が変動することが考えられる。そのため時間 $t_d$ を複数回測定するか、定期的に測定し、最も長い時間を採用して、音声出力に与える遅延時間に加算してもよい。例えばストリーミング転送中にネットワーク8の状況が悪化したら、動画像の出力途中であっても音声出力の遅延時間を増加方向に調整し、切り替え後の画像を音声に間に合わせるようにしてもよい。このような場合も、測定対象は1つの動画であっても全ての動画であってもよい。

40

#### 【0075】

また音声出力に与える遅延時間に含まれる、1つのセグメントの出力時間は、動画配信サーバ12から送信されるプレイリストから取得できるが、動画再生装置10で実測することもできる。この場合、図11で示した初期処理において、PTSの切り替わりを2回検出し、両者の時間差を1つのセグメントの出力時間として取得する。ただし当該出力時間や、動画配信サーバ12にデータを要求してから出力可能となるまでの時間 $t_d$ の計測回数を増やすほど、初期処理に時間を要することになる。したがってそれらの値を事前に取得しておき、動画再生時はそれを読み出すのみとすることにより、再生開始までの

50

時間を短縮するようにしてもよい。

#### 【0076】

これまでの説明では、1つのコンテンツを構成する複数の動画像の同期が完全にとれていることを前提としていたが、個別のカメラで撮影した動画像であれば特に、時間軸が微小量ずれていることが考えられる。図13は、1つのコンテンツを構成する複数の動画像の時間軸がずれを有する例を模式的に示している。図の横軸は時間経過を表し、音声再生用動画像、第1の表示用動画像、第2の表示用動画像の、動画配信サーバ12での再生時間の流れを、矩形の長さを各PTSの再生時間として示している。図示する例では、音声再生用動画像の時間軸に対し、第1の表示用動画像は時間D1だけ遅れており、第2の表示用動画像は時間D2だけ進んでいる。

10

#### 【0077】

本実施の形態では、この時間ずれを動画配信サーバ12で取得しておく。例えばそれぞれの動画像の音声が共通であることを利用し、同じ音の発生時刻を動画像間で比較することによりそれを測定する。動画配信サーバ12は、例えば音声再生用動画像の音声を基準とし、他の表示用動画像の画像のずれ量とずれの方向を、動画再生装置10からアクセス可能な記憶領域に格納しておく。動画再生装置10の入力情報取得部60は、ユーザによるコンテンツの選択に応じ、動画配信サーバ12に動画像のデータとともに、全ての表示用動画像の時間ずれに係る情報の送信を要求する。そしてデータ取得部62は、動画像のストリームに加え、時間ずれに係る情報も取得する。

#### 【0078】

図14は、動画再生装置10が、動画配信サーバ12が提供する動画像の時間ずれに対応するように出力タイミングを調整する際の、画像と音声の出力の流れを例示している。図の表し方は図7と同様であるが、最上段に示した、動画配信サーバ12における再生の流れは、音声再生用動画像のものとする。したがって動画再生装置10は、図7、10、12と同様の手続きにより音声再生用動画を取得し、先頭のデータ出力を一時停止させることにより所定時間の遅延を与える。

20

#### 【0079】

図ではPTS「1」の音声が、所定時間の遅延の後、出力されている。ただし図11で示したような初期処理を実施する場合、PTSの数値はこの限りではない。そして動画再生装置10は、さらに第1の表示用動画像を取得し、その先頭のデータを出力可能な状態で一時停止させた後、出力するが、時間調整部74はその際のタイミングを、あらかじめ取得しておいた動画像の時間ずれに合わせて調整する。図13に示すように、第1の表示用動画像が時間D1だけ遅れていれば、時間調整部74は図示するように、同じPTS「2」の音声の出力が開始されてから時間D1だけ遅らせてPTS「2」の画像の出力を開始する。

30

#### 【0080】

表示の切り替え操作がなされ、切り替え後の画像を表示させる場合も同様の時間調整をする。すなわち図13に示すように、第2の表示用動画像が時間D2だけ進んでいれば、図示するように時間調整部74は、同じPTS「6」の音声の出力が開始されるより時間D2だけ早いタイミングでPTS「6」の画像の出力を開始する。これらの調整により、異なる動画像の音声と表示を、厳密にずれなく出力しつづけることができる。なおこの手法は、動画像の撮影や生成時に生じるずれに限らず、動画像の復号処理に起因したずれに対しても同様の効果を奏する。

40

#### 【0081】

以上述べた本実施の形態によれば、1つのコンテンツを表す複数の動画像をストリーミング転送の対象とし、クライアント端末である動画再生装置において、1つの動画像の音声と別の動画像の画像を組み合わせて出力させる。これにより、ユーザが表示対象の動画像を切り替えるても音声が途切れることなく、1つコンテンツとしての連続性を保つことができる。また、動画再生装置における音声の出力を、1セグメント程度遅延させることにより、切り替え後の画像データの転送や復号に時間を要しても、出力しつづけている音

50

声に間に合わせることができる。

【0082】

さらに動画配信サーバにおける再生と動画再生装置での出力の時間ずれの変動を抑えるように、取得したセグメントのPTSの切り替わりを細かい時間間隔で検出する。そして切り替わり後のPTSが最速で得られるタイミングから起算して、音声出力に遅延時間を与えることにより、切り替え操作のタイミングによらず、取得したセグメントの画像が切り替え前のセグメントの次に、音声と同期させて出力できる。

【0083】

また、切り替え前の画像は停止時にフェードアウトさせ、切り替え後の画像をフェードインすることにより、任意の切り替えであるにも関わらず滑らかで自然な変化を演出できる。本実施の形態は、1つの空間を表す動画像を準備すれば、それをストリーム転送するサーバ側の処理については通常と変わりなく実現できる。したがって導入障壁が低く、従来の配信システムへの適用が容易である。

10

【0084】

以上、本発明を実施の形態をもとに説明した。上記実施の形態は例示であり、それらの各構成要素や各処理プロセスの組合せにいろいろな変形例が可能のこと、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。

【符号の説明】

【0085】

8 ネットワーク、 10 動画再生装置、 12 動画配信サーバ、 14 入力装置、  
16 ディスプレイ、 23 CPU、 24 GPU、 26 メインメモリ、 32 通信部、  
34 記憶部、 36 出力部、 38 入力部、 40 記録媒体駆動部、 50  
要求取得部、 52 データ準備部、 54 データ記憶部、 56 データ送信部、 60  
入力情報取得部、 62 データ取得部、 64 データ分離部、 66 画像復号部、 6  
8 音声復号部、 70 出力制御部、 72 PTS検出部、 74 時間調整部、 76  
画像加工部、 78 出力部。

20

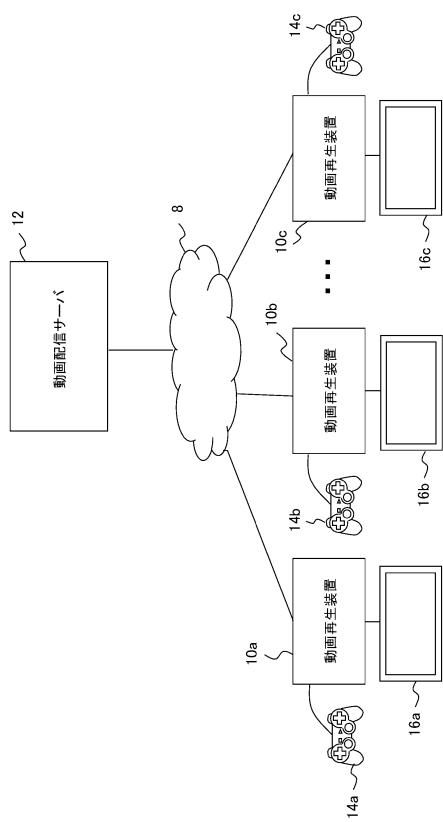
30

40

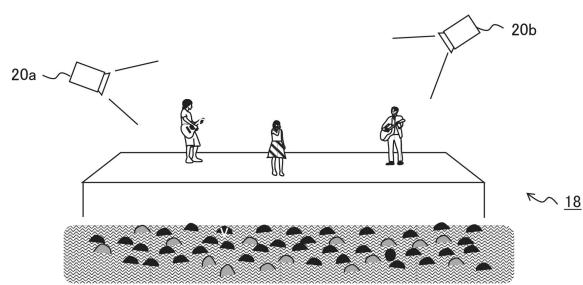
50

【図面】

【図 1】



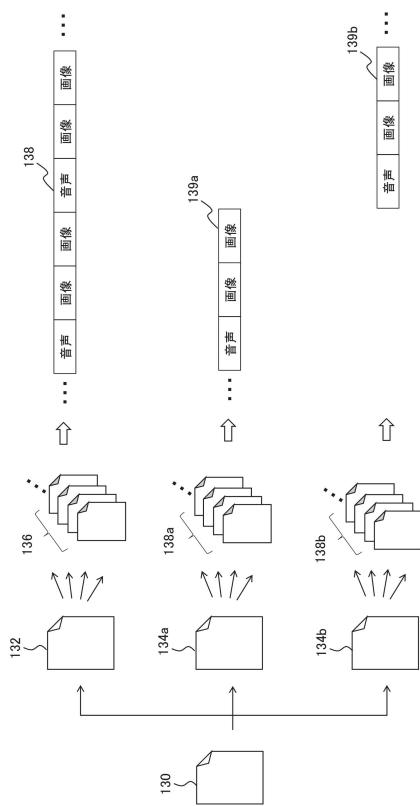
【図 2】



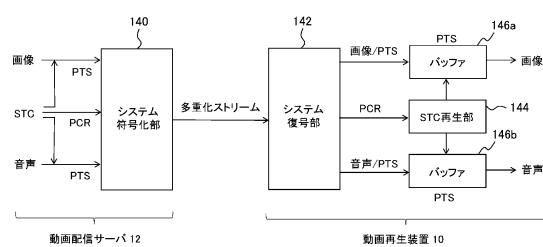
10

20

【図 3】



【図 4】

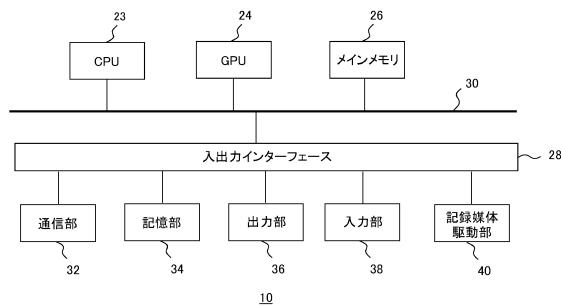


30

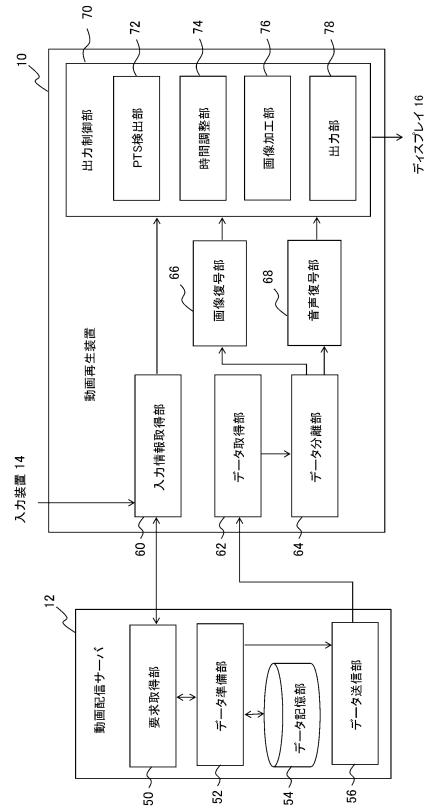
40

50

【図5】



【図6】



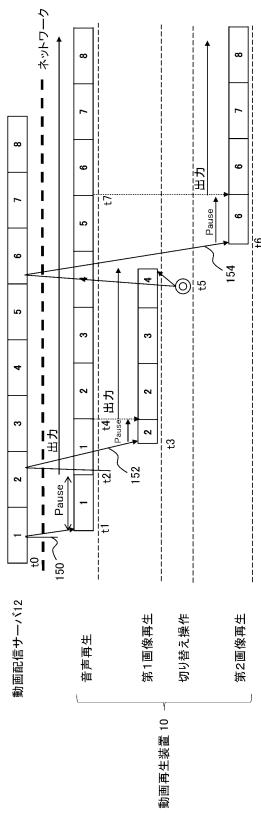
10

20

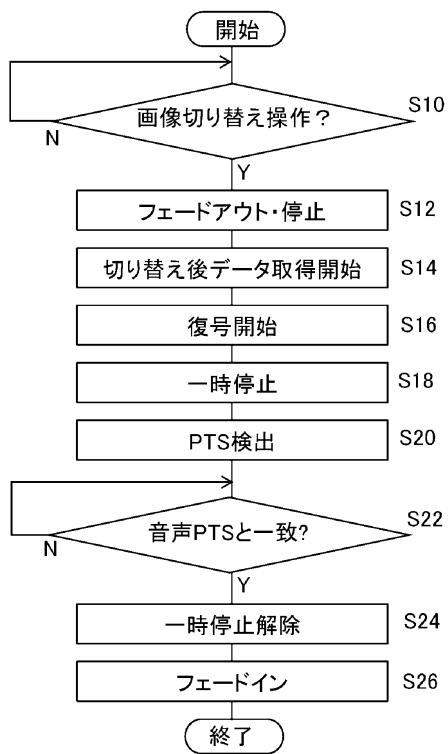
30

40

【図7】

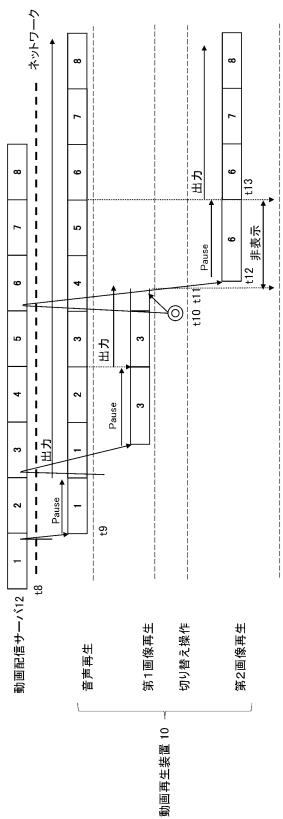


【図8】

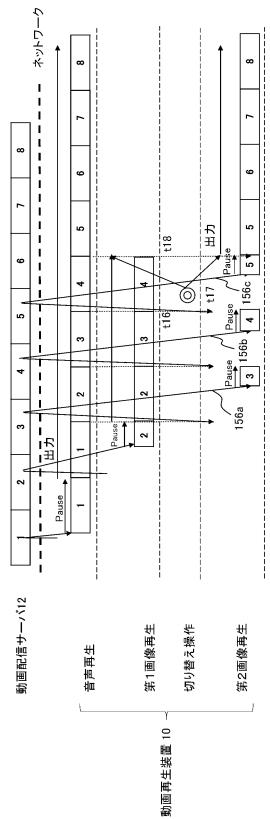


50

【図 9】



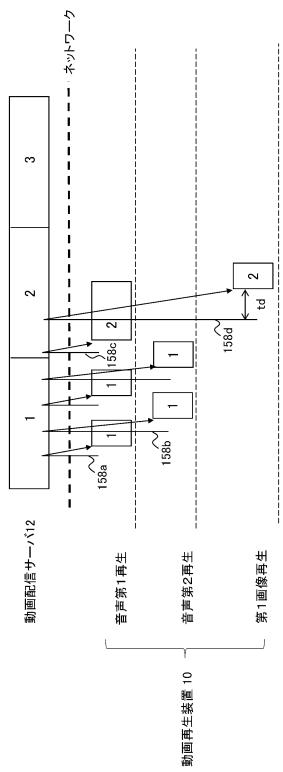
【図 10】



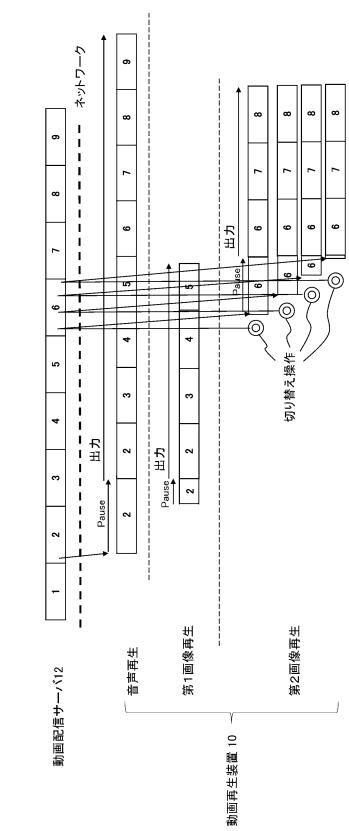
10

20

【図 11】



【図 12】

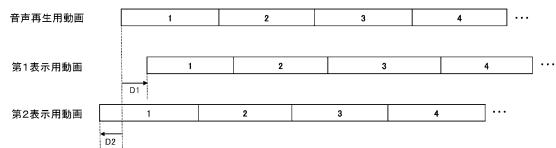


30

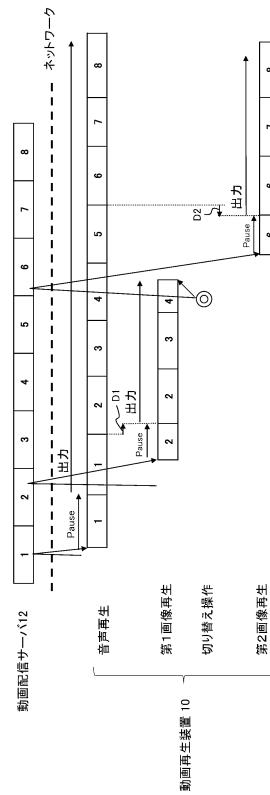
40

50

【図 1 3】



【図 1 4】



10

20

30

40

50

---

フロントページの続き

メント内

審査官 醍醐 一貴

- (56)参考文献
- 特表2018-523386 (JP, A)
  - 特開2017-225044 (JP, A)
  - 特開2012-105236 (JP, A)
  - 特開2008-005112 (JP, A)
  - 特開2018-125621 (JP, A)
  - 国際公開第2004/077825 (WO, A1)
- (58)調査した分野 (Int.Cl., DB名)
- H04N 21/00 - 21/858