

(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

F I

H 0 4 N 21/44 (2011.01)

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)	(73)特許権者	310021766
(22)出願日	令和1年12月3日(2019.12.3)		株式会社ソニー・インタラクティブエン
(65)公開番号	特開2021-90118(P2021-90118A)		タテインメント
(43)公開日	令和3年6月10日(2021.6.10)		東京都港区港南1丁目7番1号
審査請求日	令和4年3月10日(2022.3.10)	(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 P T S 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 P T S を検出する機能を複数備え、当該 P T S 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、

40

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 P T S (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。

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MPEG-DASH (Dynamic Adaptive Streaming over HTTP) やCMAF (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もまた、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画像の時間軸がずれ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】

30

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。

【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。ただしスピーカは、ディスプレイ１６とは別に設けてもよい。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、録画済みの動画画像を配信してもよいし、撮影中あるいは作成中の動画画像をラ

50

イブ配信してもよい。この際、動画配信サーバ 1 2 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 1 0 に送信してもよい。

【 0 0 1 6 】

図 2 は、本実施の形態で動画再生装置 1 0 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 1 8 に複数のカメラ 2 0 a、2 0 b 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 2 0 a、2 0 b 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。

【 0 0 1 7 】

動画配信サーバ 1 2 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 1 0 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 1 0 から随時受け付ける。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 1 0 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。

【 0 0 1 8 】

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。

【 0 0 1 9 】

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まりを「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 1 0 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する。この際、表示上の画像を切り替えても、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することで、シームレスな切り替えを実現する。

【 0 0 2 0 】

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。

【 0 0 2 1 】

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 1 2 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 1 2 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。

【 0 0 2 2 】

図において、音声再生用の動画像のプレイリスト 1 3 2 は、音声再生用のセグメントデータ 1 3 6 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 1 3 4 a、1 3 4 b はそれぞれ、表示用のセグメントデータ 1 3 8 a、1 3 8 b 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

10

20

30

40

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。

【 0 0 2 3 】

動画配信サーバ 1 2 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クスファイル 1 3 0 を生成する。インデクスファイル 1 3 0 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 1 3 2、1 3 4 a、1 3 4 b 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、H L S、M P E G - D A S H、C M A F 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

【 0 0 2 4 】

いずれにしる本実施の形態で動画配信サーバ 1 2 は、動画再生装置 1 0 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、インデクスファイル 1 3 0、プレイリスト 1 3 2、1 3 4 a、1 3 4 b 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 1 2 はまず、必要な動画像のプレイリストを動画再生装置 1 0 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 1 0 に送信する。

【 0 0 2 5 】

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 1 2 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）1 3 8 が継続的に送信される一方、表示用の動画像のデータ列 1 3 9 a、1 3 9 b 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

【 0 0 2 6 】

一方、動画再生装置 1 0 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 1 2 側では、自らが有する S T C (S y s t e m T i m e C l o c k) 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 P T S (P r e s e n t a t i o n T i m e S t a m p) を生成し、それぞれのセグメントデータに付与する。

30

【 0 0 2 7 】

動画配信サーバ 1 2 はまた、S T C をベースとした所定周期でのカウンタ値を表す P C R (P r o g r a m C l o c k R e f e r e n c e) を生成する。動画配信サーバ 1 2 のシステム符号化部 1 4 0 は、画像や音声のセグメントデータとそれぞれの P T S、および所定周期での P C R 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 1 0 に送信する。動画再生装置 1 0 のシステム復号部 1 4 2 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 S T C 再生部 1 4 4 は、各パケットの受信時刻と P C R 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 S T C を再生する。

40

【 0 0 2 8 】

これにより動画再生装置 1 0 は、動画配信サーバ 1 2 と共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 1 0 は、当該時間軸上での P T S において、対応する画像および音声出力されるように、バッファ 1 4 6 a、1 4 6 b 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 P T S が付与されていれば、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 1 6 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 1 0 の内部回路構成を示している。動画再生装置 1 0 は、C P U (Central Processing Unit) 2 3、G P U (Graphics Processing Unit) 2 4、メインメモリ 2 6 を含む。これらの各部は、バス 3 0 を介して相互に接続されている。バス 3 0 にはさらに入出力インターフェース 2 8 が接続されている。入出力インターフェース 2 8 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 1 2 と通信を確立する通信部 3 2、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 3 4、ディスプレイ 1 6 へデータを出力する出力部 3 6、入力装置 1 4 からデータを入力する入力部 3 8、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 4 0 が接続される。

10

【 0 0 3 0 】

C P U 2 3 は、記憶部 3 4 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 1 0 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 2 6 にロードされた、あるいは通信部 3 2 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 3 6 に出力する。メインメモリ 2 6 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 1 2 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 1 0 および動画配信サーバ 1 2 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 2 6 などを実現でき、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 1 2 は、動画再生装置 1 0 からの要求を取得する要求取得部 5 0、要求に従ったデータを準備するデータ準備部 5 2、コンテンツのデータを記憶するデータ記憶部 5 4、および、準備されたデータを動画再生装置 1 0 に送信するデータ送信部 5 6 を含む。要求取得部 5 0 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 1 0 から取得する。そのため要求取得部 5 0 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 1 0 に送信してもよい。

30

【 0 0 3 3 】

データ準備部 5 2 は要求取得部 5 0 と連携し、動画再生装置 1 0 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 5 2 は例えば、選択されたコンテンツに対応するインデ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 5 0 に与えることにより、動画再生装置 1 0 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 5 2 は、対応するプレイリストを取得し、要求取得部 5 0 を介して動画再生装置 1 0 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 5 2 は、音声再生用の動画像と、動画再生装置 1 0 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 5 2 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。

【 0 0 3 5 】

データ記憶部 5 4 は図 3 に示したように、インデクスファイル、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。

【 0 0 3 6 】

データ送信部 5 6 は、データ準備部 5 2 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 1 0 に送信する。この際、データ送信部 5 6 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 P T S を付与するとともに、周期的に P C R も付与する。データ送信部 5 6 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。

【 0 0 3 7 】

動画再生装置 1 0 は、ユーザ操作の内容を取得する入力情報取得部 6 0、動画像のストリームを取得するデータ取得部 6 2、ストリームからデータを分離するデータ分離部 6 4、画像データを復号する画像復号部 6 6、音声データを復号する音声復号部 6 8、動画像の出力を制御する出力制御部 7 0 を含む。入力情報取得部 6 0 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 1 4 から取得し、必要なデータを動画配信サーバ 1 2 に要求する。

【 0 0 3 8 】

上述のとおり入力情報取得部 6 0 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 1 2 から取得し、出力制御部 7 0 を介してディスプレイ 1 6 に選択肢として表示させてもよい。入力情報取得部 6 0 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 1 2 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 1 2 に要求する。

【 0 0 3 9 】

データ取得部 6 2 は、ユーザ操作に応じて動画配信サーバ 1 2 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 P T S 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 6 4 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 1 2 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。

【 0 0 4 0 】

画像復号部 6 6 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 6 8 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 7 0 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 1 6 に出力する。詳細には出力制御部 7 0 は、P T S 検出部 7 2、時間調整部 7 4、画像加工部 7 6、および出力部 7 8 を含む。P T S 検出部 7 2 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 P T S を検出する。

【 0 0 4 1 】

時間調整部 7 4 は、画像と音声にずれが生じないように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

10

20

30

40

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ずれ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、ディスプレイ１６のスピーカーに出力する。

【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を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。

【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 t_1 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 t_0 より遅延する。

【００４６】

従来技術においては、出力可能となった時刻 t_1 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 t_2 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($t_2 - t_1$)は、１つ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 t_2 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 t_3 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。

【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、遅延させたことにより、時刻 t_3 は、PTS「２」の音声データの出力開始時刻 t_4 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 t_4 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

10

20

30

40

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。また図示するように後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。

【 0 0 4 9 】

なお画像の出力開始に当たっては、画像加工部 7 6 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 t_5 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 1 0 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 1 2 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 7 6 は第 1 の画像をフェードアウトさせる。

【 0 0 5 0 】

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 1 2 からは矢印 1 5 4 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 1 0 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 t_6 は、PTS「6」の音声データの出力開始時刻 t_7 よりは前になる。

【 0 0 5 1 】

そこで時間調整部 7 4 は、PTS「6」の音声データの出力が開始される時刻 t_7 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 t_7 において一時停止を解除し、画像加工部 7 6、出力部 7 8 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

【 0 0 5 2 】

たうえで、同じ PTS の音声出力されるのを待ってフェードインさせる。

図 8 は、表示画像を切り替える際の動画再生装置 1 0 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 1 6 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 6 0 は、表示画像を切り替えるユーザ操作を待機する (S 1 0 の N)。当該ユーザ操作がなされたら (S 1 0 の Y)、出力制御部 7 0 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 (S 1 2)。

【 0 0 5 3 】

一方、入力情報取得部 6 0 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 1 2 に要求することにより、データ取得部 6 2 は、切り替え後の動画像データのうち、動画配信サーバ 1 2 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 (S 1 4)。そして画像復号部 6 6 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 (S 1 6)。出力制御部 7 0 は、復号された画像の出力を一時停止させるとともに (S 1 8)、その PTS を検出する (S 2 0)。

【 0 0 5 4 】

そして出力制御部 7 0 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 (S 2 2 の N)。結果として、ディスプレイ 1 6 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 (S 2 2 の Y)、出力制御部 7 0 は画像出力の一時停止を解除するとともに (S 2 4)、画像をフェードインさせながら表示させる (S 2 6)。なお S 2 4 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。

【 0 0 5 5 】

図 7 で示した態様は、音声出力を遅延させる時間を、1 つ分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 1 2 での再生時間と動画再生装置 1 0 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 1 2 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

10

20

30

40

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 1 2 での再生時間と動画再生装置 1 0 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 1 2 における動画の再生時間に対し、動画再生装置 1 0 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 1 2 において P T S 「 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 1 0 において出力可能となる時刻 t_9 は、動画配信サーバ 1 2 における再生開始時刻 t_8 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を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 1 2 と動画再生装置 1 0 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 t_{10} 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 1 0 において第 1 の画像の P T S 「 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 1 2 では、それより 2 つ後の、P T S 「 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 1 0 は、切り替え後の第 2 の動画データのうち P T S 「 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 t_{12} から、P T S 「 6 」 の音声データの出力開始時刻 t_{13} 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ふ。

20

【 0 0 5 9 】

切り替え前の第 1 の画像の表示は、時刻 t_{12} より前の時刻 t_{11} ですすでに停止していることから、時刻 t_{11} から時刻 t_{13} 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 1 0 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 1 2 から音声再生用の動画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 の表示用動画データ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 の表示用動画データも並行して取得する。

【 0 0 6 1 】

すなわち動画再生装置 1 0 の入力情報取得部 6 0 は、矢印 1 5 6 a、1 5 6 b、1 5 6 c に示すように、出力中の動画データ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 の表示用動画データのセグメント単位での送信を、動画配信サーバ 1 2 に要求する。そして動画再生装置 1 0 は、画像の切り替え操作がなされないうちは第 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 P T S 「 3 」 の先頭の画像が時刻 t_{14} で出力可能となり、P T S 「 4 」 の先頭の画像が時刻 t_{15} で出力可能となっている。それぞれの画像の出力は、同じ P T S 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 $t_{16} \sim t_{18}$ 中の時刻 t_{17} に、画像を切り替える操作がなされたことから、時刻 t_{18} 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とうえで、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で、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。

【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。

【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。

【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。

【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

10

20

30

40

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つ分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画像も全て準備せずとも、常に短時間で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の画像を準備し終わることができ、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、最小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が保証できる。

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つ分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つ分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。

【 0 0 7 6 】

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

【 0 0 7 7 】

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ずれを測定する。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。

【 0 0 7 8 】

20

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。

【 0 0 7 9 】

図ではPTS「1」の音声、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

【 0 0 8 0 】

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

【 0 0 8 1 】

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ても音声途切れることなく、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。

【 0 0 8 2 】

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。

【 0 0 8 3 】

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させることにより、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。

【 0 0 8 4 】

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なこと、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 】

【 0 0 8 5 】

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。

10

20

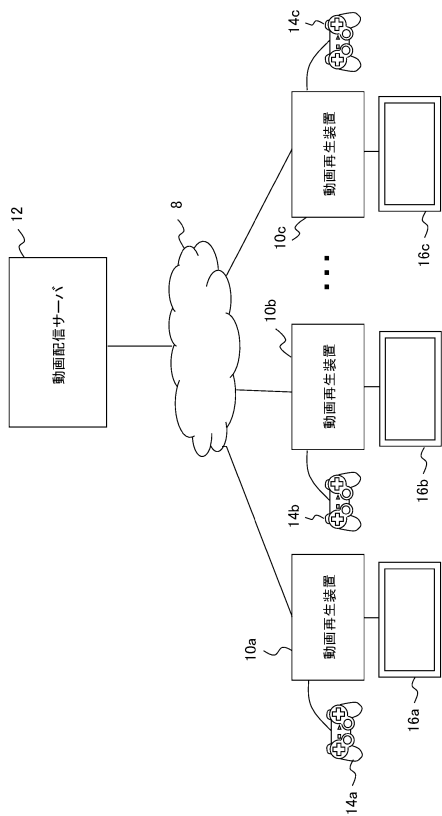
30

40

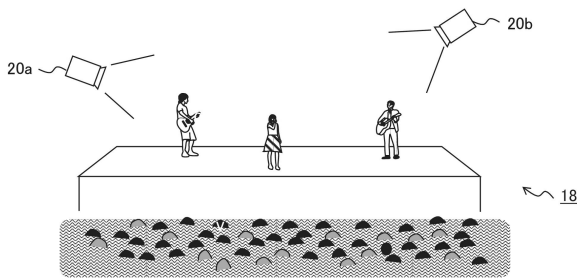
50

【図面】

【図 1】



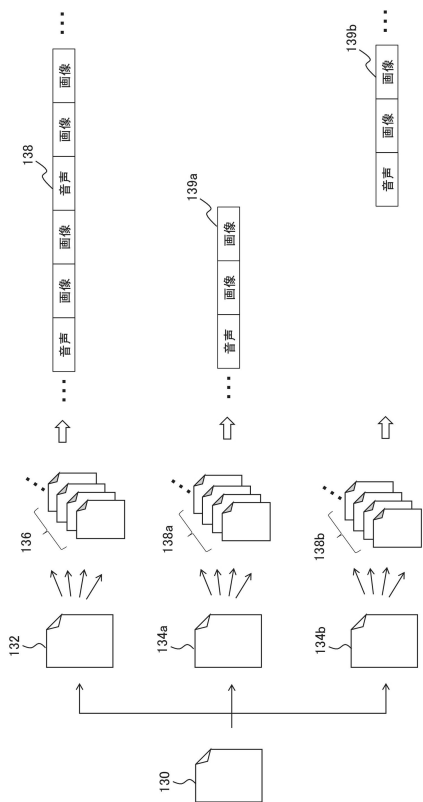
【図 2】



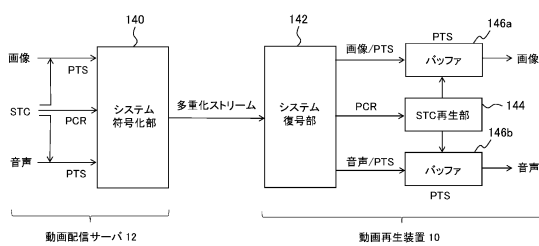
10

20

【図 3】



【図 4】

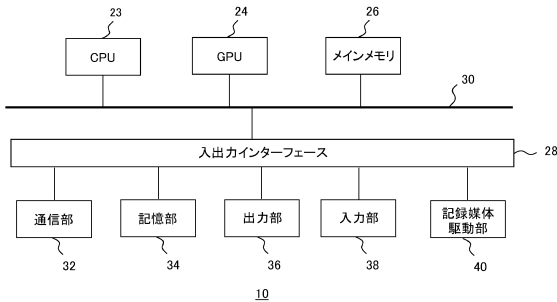


30

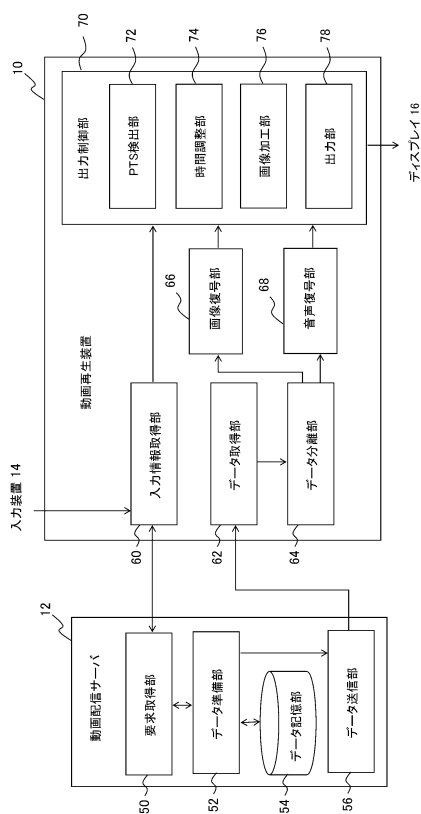
40

50

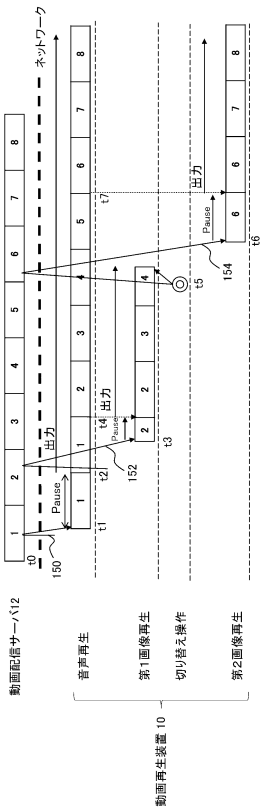
【図 5】



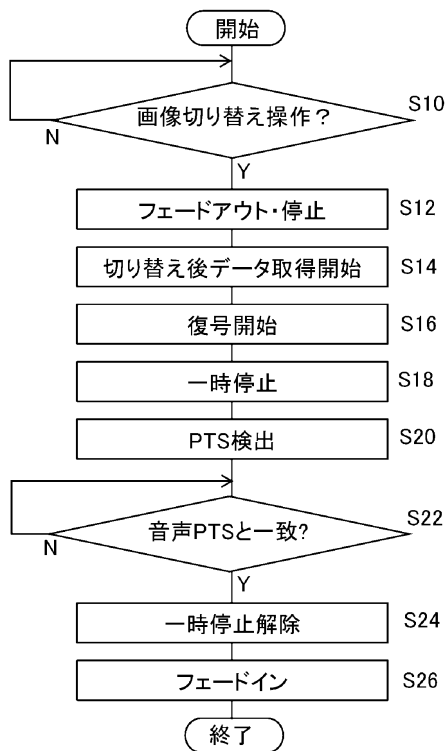
【図 6】



【図 7】



【図 8】



10

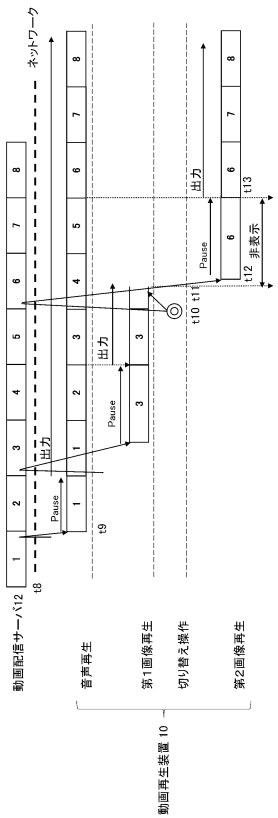
20

30

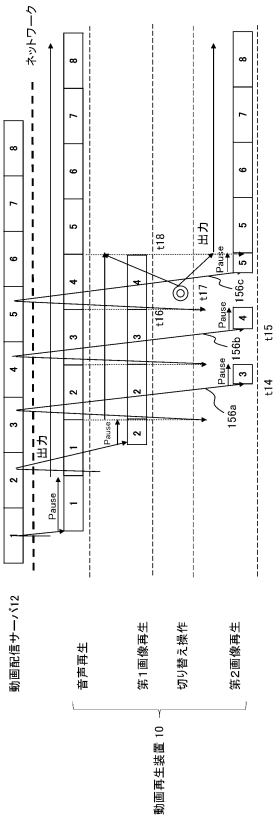
40

50

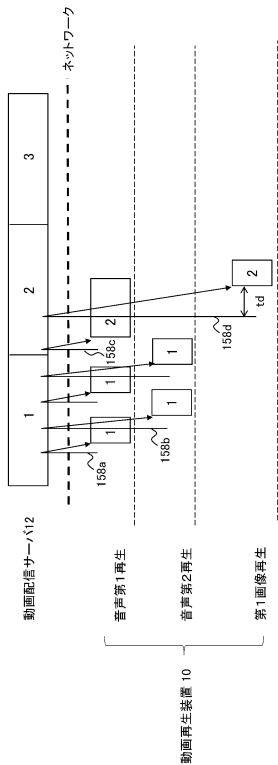
【図 9】



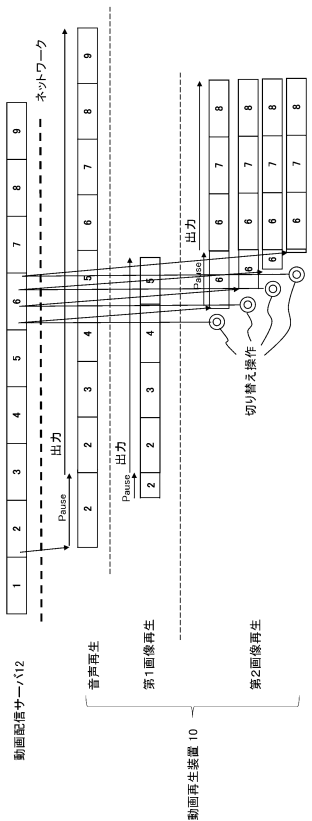
【図 10】



【図 11】



【図 12】



10

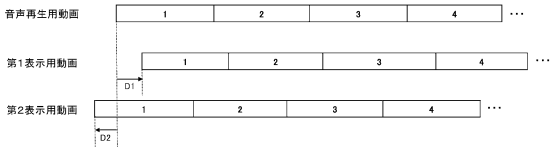
20

30

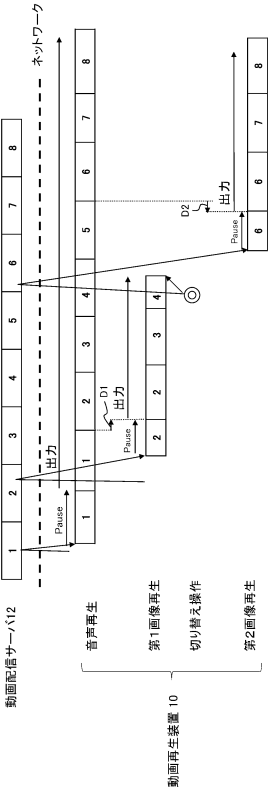
40

50

【図 13】



【図 14】



10

20

30

40

50

フロントページの続き

ンメント内

審査官 醍醐 一貴

- (56)参考文献 特表 2 0 1 8 - 5 2 3 3 8 6 (J P , A)
特開 2 0 1 7 - 2 2 5 0 4 4 (J P , A)
特開 2 0 1 2 - 1 0 5 2 3 6 (J P , A)
特開 2 0 0 8 - 0 0 5 1 1 2 (J P , A)
特開 2 0 1 8 - 1 2 5 6 2 1 (J P , A)
国際公開第 2 0 0 4 / 0 7 7 8 2 5 (W O , A 1)
- (58)調査した分野 (Int.Cl. , D B 名)
H 0 4 N 2 1 / 0 0 - 2 1 / 8 5 8