

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5957760号
(P5957760)

(45) 発行日 平成28年7月27日 (2016. 7. 27)

(24) 登録日 平成28年7月1日 (2016. 7. 1)

(51) Int. Cl.	F 1
HO 4 N 21/485 (2011. 01)	HO 4 N 21/485
HO 4 N 21/436 (2011. 01)	HO 4 N 21/436

請求項の数 10 (全 24 頁)

(21) 出願番号	特願2012-51943 (P2012-51943)	(73) 特許権者	314012076
(22) 出願日	平成24年3月8日 (2012. 3. 8)		パナソニック IP マネジメント株式会社
(65) 公開番号	特開2013-187765 (P2013-187765A)		大阪府大阪市中央区城見2丁目1番61号
(43) 公開日	平成25年9月19日 (2013. 9. 19)	(74) 代理人	100109210
審査請求日	平成26年9月2日 (2014. 9. 2)		弁理士 新居 広守
		(74) 代理人	100137235
			弁理士 寺谷 英作
		(74) 代理人	100131417
			弁理士 道坂 伸一
		(72) 発明者	田中 俊介
			大阪府門真市大字門真1006番地 パナソニック株式会社内
		(72) 発明者	山田 豪
			大阪府門真市大字門真1006番地 パナソニック株式会社内

最終頁に続く

(54) 【発明の名称】 映像音声処理装置

(57) 【特許請求の範囲】

【請求項 1】

映像音声処理装置であって、
映像信号を出力する映像出力部と、
前記映像信号に対応する音声信号を出力する音声出力部と、
前記映像信号に対応する前記音声信号を、前記映像音声処理装置の外部の音声再生装置に送信する音声送信部と、

前記映像音声処理装置の動作モードを、（a）前記音声出力部から前記音声信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第一モード、および、（b）前記映像出力部から前記映像信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第二モードの一方から他方へ切り換える制御部と、

前記動作モードが前記第一モードである期間に、前記音声出力部から出力される音声信号を遅延させる量である音声遅延量を特定する遅延情報の入力を受け付ける受付部と、

前記受付部が受け付けた前記遅延情報によって特定される前記音声遅延量に応じて前記音声出力部から出力される音声信号を遅延させる音声遅延部と、

前記動作モードが前記第二モードである期間に、前記映像出力部から出力される映像信号を前記音声遅延量に応じた映像遅延量だけ遅延させる映像遅延部とを備え、

前記音声出力部は、前記第一モードである期間に、前記音声出力部から出力される音声信号の遅延の調整のための、時系列上で変化する調整用の音声信号を出力し、

前記音声送信部は、前記第一モードである期間に、前記調整用の音声信号を送信し、

10

20

前記映像出力部は、前記動作モードが前記第一モードである期間に、ユーザによる所定の操作のためのユーザインターフェース画面であって、前記調整用の音声信号の変化に同期して動く調整用の動画を含むユーザインターフェース画面を示す映像信号を出力し、
前記受付部は、ユーザの前記所定の操作により入力される前記遅延情報の入力を受け付ける

映像音声処理装置。

【請求項 2】

前記映像出力部は、音声遅延部により、前記音声出力部から出力される前記音声信号が遅延された場合、前記音声信号に追従して遅延させた前記調整用の動画を含むユーザインターフェース画面を示す映像信号を出力する

10

請求項 1 記載の映像音声処理装置。

【請求項 3】

前記映像遅延部は、前記音声遅延量以下の値である前記映像遅延量だけ前記映像出力部から出力される前記映像信号を遅延させる

請求項 1 記載の映像音声処理装置。

【請求項 4】

前記音声遅延部は、前記映像信号のフレームレートから算出される 1 フレーム分の時間の整数倍に対応する前記音声遅延量に応じて前記音声出力部から出力される前記音声信号を遅延させる

請求項 1 記載の映像音声処理装置。

20

【請求項 5】

前記映像遅延部は、前記音声遅延量より大きな前記映像遅延量だけ前記映像出力部から出力される前記映像信号を遅延させ、

前記音声送信部は、前記音声遅延量と前記映像遅延量との差分に応じた値だけ、前記音声送信部から送信される前記音声信号を遅延させる

請求項 1 記載の映像音声処理装置。

【請求項 6】

前記映像遅延部は、前記音声遅延量以下である前記映像遅延量であって、前記映像信号のフレームレートから算出される 1 フレーム分の時間の整数倍に対応する前記映像遅延量だけ前記映像出力部から出力される前記映像信号を遅延させる

30

請求項 1 記載の映像音声処理装置。

【請求項 7】

前記受付部は、前記遅延情報として、前記音声信号を受信して再生する前記外部の音声再生装置から出力される音声の信号である再生音声信号の入力を受け付け、

前記映像遅延部は、前記再生音声信号と、前記音声遅延部が遅延させる前の前記音声信号との間の遅延量である前記音声遅延量に応じた前記映像遅延量だけ、前記映像出力部から出力される前記映像信号を遅延させる

請求項 1 記載の映像音声処理装置。

【請求項 8】

さらに、前記映像遅延量を示す情報である映像遅延情報を記憶する記憶部を備え、

40

前記映像遅延部は、前記動作モードが前記第二モードである期間に、前記記憶部から読み出した前記映像遅延情報に示される前記映像遅延量だけ、前記映像出力部から出力される前記映像信号を遅延させる

請求項 1 記載の映像音声処理装置。

【請求項 9】

前記記憶部は、前記音声再生装置を含む複数の音声再生装置のそれぞれに対応する複数の映像遅延量を示す前記映像遅延情報を記憶し、

前記映像遅延部は、前記動作モードが前記第二モードであって、かつ、前記音声送信部が、前記複数の音声再生装置のそれぞれに同時に前記音声信号を送信する場合、(c) 前記記憶部に記憶されている前記映像遅延情報に示される前記複数の映像遅延量のうち、最

50

も大きな映像遅延量を選択し、(d) 前記映像出力部から出力される前記映像信号を、選択した映像遅延量だけ遅延させる

請求項 8 記載の映像音声処理装置。

【請求項 10】

映像音声処理装置によって実行される映像音声処理方法であって、

前記映像音声処理装置は、映像信号を出力する映像出力部と、前記映像信号に対応する音声信号を出力する音声出力部と、前記映像信号に対応する前記音声信号を、前記映像音声処理装置の外部の音声再生装置に送信する音声送信部とを備え、

前記映像音声処理方法は、

前記映像音声処理装置の動作モードが、前記音声出力部から前記音声信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第一モードである期間に、前記音声出力部から出力される音声信号を遅延させる量である音声遅延量を特定する遅延情報の入力を受け付ける受付ステップと、

前記受付ステップにおいて受け付けられた前記遅延情報によって特定される前記音声遅延量に応じて前記音声出力部から出力される音声信号を遅延させる音声遅延ステップと、

前記映像音声処理装置の動作モードが、前記映像出力部から前記映像信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第二モードである期間に、前記映像出力部から出力される映像信号を前記音声遅延量に応じた映像遅延量だけ遅延させる映像遅延ステップとを含み、

前記音声出力部は、前記第一モードである期間に、前記音声出力部から出力される音声信号の遅延の調整のための、時系列上で変化する調整用の音声信号を出力し、

前記音声送信部は、前記第一モードである期間に、前記調整用の音声信号を送信し、

さらに、前記動作モードが前記第一モードである期間に、前記映像出力部が、ユーザによる所定の操作のためのユーザインターフェース画面であって、前記調整用の音声信号の変化に同期して動く調整用の動画を含むユーザインターフェース画面を示す映像信号を出力する、ユーザインターフェース画面出力ステップを含み、

前記受付ステップでは、ユーザの前記所定の操作により入力される前記遅延情報の入力を受け付ける

映像音声処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本願発明は、映像音声処理装置に関し、特に再生時における映像信号と音声信号との同期のための処理を行う映像音声処理装置に関する。

【背景技術】

【0002】

従来、映像信号および音声信号を処理して出力する映像音声処理装置が存在する。このような映像音声処理装置では、例えば、映像信号と音声信号とを互いに異なる機器に出力して、それぞれに映像または音声を再生させる場合がある。この場合、再生時における映像信号と音声信号との同期（例えば、「リップシンク」と呼ばれる）が問題となる。

【0003】

そこで、再生時における映像信号と音声信号とを同期させるための技術も開示されている。例えば、特許文献 1 には、音声信号を遅延させることで再生映像と再生音とのずれを低減させる音声映像伝送装置について記載されている。

【先行技術文献】

【特許文献】

【0004】

【特許文献 1】特開 2004-88442 号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

ここで、例えば、ある放送番組をテレビのディスプレイに表示させながら、当該テレビから送信される当該放送番組の音声信号を、当該テレビの外部機器（外部スピーカ、または、ヘッドホンなど）に受信させて再生させる場合を想定する。この場合、当該外部機器で再生される音声信号が、当該ディスプレイに表示される映像信号に対して遅れる場合がある。

【0006】

このような場合、例えば、映像信号をどの程度遅らせればいいのか等の調整量の決定は容易ではなく、当該調整を効率よく行うことは困難である。

【0007】

本発明は、上記従来課題を考慮し、再生時における映像信号と音声信号との同期のための処理を効率よく実行することができる映像音声処理装置を提供することを目的とする。

【課題を解決するための手段】

【0008】

上記目的を達成するために、本発明の一態様に係る映像音声処理装置は、映像音声処理装置であって、映像信号を出力する映像出力部と、前記映像信号に対応する音声信号を出力する音声出力部と、前記映像信号に対応する前記音声信号を、前記映像音声処理装置の外部の音声再生装置に送信する音声送信部と、前記映像音声処理装置の動作モードを、（a）前記音声出力部から前記音声信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第一モード、および、（b）前記映像出力部から前記映像信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第二モードの一方から他方へ切り換える制御部と、前記動作モードが前記第一モードである期間に、前記音声出力部から出力される音声信号を遅延させる量である音声遅延量を特定する遅延情報の入力を受け付ける受付部と、前記受付部が受け付けた前記遅延情報によって特定される前記音声遅延量に応じて前記音声出力部から出力される音声信号を遅延させる音声遅延部と、前記動作モードが前記第二モードである期間に、前記映像出力部から出力される映像信号を前記音声遅延量に応じた映像遅延量だけ遅延させる映像遅延部とを備える。

【0009】

なお、これらの全般的または具体的な態様は、システム、方法、集積回路、コンピュータプログラムまたは記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラムおよび記録媒体の任意な組み合わせで実現されてもよい。

【発明の効果】

【0010】

本発明の映像音声処理装置によれば、再生時における映像信号と音声信号との同期のための処理を効率よく実行することができる。

【図面の簡単な説明】

【0011】

【図1】図1は、実施の形態におけるAV（Audio Visual）システムの構成概要を示す図である。

【図2】図2は、実施の形態におけるAVシステムの基本的な機能構成を示すブロック図である。

【図3】図3は、実施の形態の映像音声処理装置における基本的な処理の流れを示すフロー図である。

【図4】図4は、実施の形態の映像音声処理装置が出力するユーザインターフェース画面の一例を示す図である。

【図5】図5は、実施の形態の映像音声処理装置が行う映像信号と音声信号の同期調整を説明するための図である。

【図6】図6は、1フレーム時間と、スピーカおよびヘッドホンの間の出音タイミングのずれ量との関係を示す図である。

10

20

30

40

50

【図 7】図 7 は、実施の形態における映像音声処理装置が記憶部を備える場合の基本的な機能構成を示すブロック図である。

【図 8】図 8 は、実施の形態の A V システムが複数の音声再生装置を備える場合の構成概要を示す図である。

【図 9】図 9 は、実施の形態における映像遅延情報のデータ構成例を示す図である。

【図 10】図 10 は、実施の形態における映像音声処理装置が遅延情報として再生音信号を取得する場合の基本的な機能構成を示すブロック図である。

【発明を実施するための形態】

【0012】

(本発明の基礎となった知見)

10

本発明者は、再生時における映像信号と音声信号との同期に関し、以下の問題が生じることを見出した。

【0013】

例えばデジタルテレビ放送における放送番組は、当該放送番組に対応する映像信号および音声信号とともに、映像信号と音声信号とを同期させるための信号を含むストリームによって、各テレビに送信される。そのため、当該放送番組をテレビ単体で再生する場合には、一般にリップシンクの問題は発生しない。

【0014】

しかし、上述のように、例えば、テレビで当該映像信号を再生させながら当該テレビから送信される当該音声信号を外部機器（音声再生装置）で受信させて再生させる場合、当該テレビにおける再生映像に対して、当該音声再生装置における再生音が遅れる場合がある。

20

【0015】

この遅れは、例えば、当該テレビと当該音声再生装置との間の通信手順（通信エラー時の音声信号の再送など）、または、当該音声再生装置における音声信号の処理（音切れ防止のための音声信号のバッファリングなど）に起因して発生する。

【0016】

このように、再生映像に対して再生音が遅れた場合、上記の当該遅れの要因を考慮すると、再生音の出力を早めることは現実的ではなくかつ困難である。

【0017】

30

そのため、再生映像を遅延させることで、再生映像と再生音との再生タイミングの一致を図ることが考えられる。つまり、映像信号をディスプレイに出力し、かつ、音声信号を外部の音声再生装置に送信する映像音声処理装置において、映像信号の出力を遅延させることで、再生時における映像信号と音声信号との同期を図ることが考えられる。

【0018】

しかしながら、この場合、例えば、ユーザは、音声再生装置で再生される音を聞きながら、再生映像を遅らせるように、当該映像音声処理装置に、映像信号の遅延量を入力する必要がある。

【0019】

例えば、ユーザは、音声再生装置で再生される人物の声を聞きながら、その声と、テレビでの再生映像における当該人物の唇の動きとを合わせるように、映像信号の遅延量を調整する。

40

【0020】

つまり、聴覚と視覚とを同時に働かせながら、聴覚でとらえた音声の特徴点と、視覚でとらえた映像の特徴点とを時系列上で一致させようとする容易ではない作業が行われる。

【0021】

その結果、当該映像音声処理装置では、聴覚および視覚でとらえる音声および映像についてのユーザの違和感がなくなるまで、遅延量の増加および減少を繰り返すという、非効率的な処理が発生することとなる。

【0022】

50

さらに、音声信号を受信して再生する音声再生装置が変更になると、当該遅延量も変わるため、音声再生装置の変更のたびに、非効率的な処理が発生するという問題が生じる。

【0023】

このような問題を解決するために、本発明の一態様に係る映像音声処理装置は、映像音声処理装置であって、映像信号を出力する映像出力部と、前記映像信号に対応する音声信号を出力する音声出力部と、前記映像信号に対応する前記音声信号を、前記映像音声処理装置の外部の音声再生装置に送信する音声送信部と、前記映像音声処理装置の動作モードを、(a)前記音声出力部から前記音声信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第一モード、および、(b)前記映像出力部から前記映像信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第二モードの一方から他方へ切り換える制御部と、前記動作モードが前記第一モードである期間に、前記音声出力部から出力される音声信号を遅延させる量である音声遅延量を特定する遅延情報の入力を受け付ける受付部と、前記受付部が受け付けた前記遅延情報によって特定される前記音声遅延量に応じて前記音声出力部から出力される音声信号を遅延させる音声遅延部と、前記動作モードが前記第二モードである期間に、前記映像出力部から出力される映像信号を前記音声遅延量に応じた映像遅延量だけ遅延させる映像遅延部とを備える。

10

【0024】

この構成によれば、映像音声処理装置が第一モードで動作中に出力される、例えば音声出力部に接続されたスピーカからの音声と、外部の音声再生装置からの音声との比較の結果から得られる遅延情報を、映像音声処理装置に入力することができる。

20

【0025】

つまり、映像信号との同期の問題のない音声信号であって、音声出力部から出力される音声信号に基づく音声(第一音声)と、外部の音声再生装置からの音声(第二音声)とのずれ量(音声遅延量)を特定する遅延情報が映像音声処理装置に入力される。さらに、当該音声遅延量に応じて映像信号が遅延される。

【0026】

簡単にいうと、外部の音声再生装置からの第二音声と、映像出力部に接続されたディスプレイに表示される映像との間のずれ量が、当該第二音声と当該映像との比較ではなく、当該第二音声と、当該映像と同期が保障された第一音声との比較によって決定される。

【0027】

ここで、人間は、ある音源の発生位置等の特定に、当該音源から発生し、僅かに時間をあけて耳に到達する2つの音の時間差を利用するため、音の時間的なずれを知覚する能力が優れているという特長を有する。そのため、第一音声と第二音声とを高い精度で一致させることが可能である。すなわち、上記比較を人間が行った場合であっても、第二音声とタイミングが一致するように、第一音声を遅延させることは容易である。

30

【0028】

従って、第一音声を第二音声に同期させるための音声遅延量の決定は容易化され、その結果、第二音声と、当該映像信号に基づく再生映像との同期のための映像遅延量の決定も容易化される。

【0029】

もちろん、人間ではなく、機械的に音声遅延量を決定する場合であっても、例えば、第一音声および第二音声の音圧レベルのピークのタイミングの比較等によって、容易に特定することができる。つまり、音声解析の結果と、映像解析の結果とを比較するような複雑な処理なしに、音声遅延量は決定され、その結果、第二音声と、当該映像信号に基づく再生映像との同期のための映像遅延量の決定も容易化される。

40

【0030】

以上のように、本態様の映像音声処理装置は、再生時における映像信号と音声信号との同期のための映像遅延量を効率よく特定することができ、その結果、当該同期のための処理を効率よく実行することができる。

【0031】

50

また、例えば、前記映像出力部は、前記動作モードが前記第一モードである期間に、ユーザによる所定の操作のためのユーザインターフェース画面を示す映像信号を出力し、前記受付部は、ユーザの前記所定の操作により入力される前記遅延情報の入力を受け付けるとしてもよい。

【0032】

この構成によれば、映像音声処理装置は、例えば、リップシンクのための調整作業を、ユーザに効率よく行わせることができる。

【0033】

また、例えば、前記映像遅延部は、前記音声遅延量以下の値である前記映像遅延量だけ前記映像出力部から出力される前記映像信号を遅延させるとしてもよい。

10

【0034】

この構成によれば、映像遅延量は、リップシンクのための厳密な遅延量よりも小さくなる可能性があるが、少なくとも、音声映像に先行する事態が防止される。例えば、人物が喋っている映像において、人物が口を動かす前に発話音が外部の音声再生装置で再生されるような、極めて不自然な状況の発生が防止される。

【0035】

また、例えば、前記音声遅延部は、前記映像信号のフレームレートから算出される1フレーム分の時間の整数倍に対応する前記音声遅延量に応じて前記音声出力部から出力される前記音声信号を遅延させるとしてもよい。

【0036】

20

この構成によれば、例えば、映像の遅延がフレーム単位で行われる場合に、音声遅延量をそのまま映像遅延量として用いることができる。つまり、映像信号と音声信号との同期に係る処理負荷が軽減される。

【0037】

また、例えば、前記映像遅延部は、前記音声遅延量より大きな前記映像遅延量だけ前記映像出力部から出力される前記映像信号を遅延させ、前記音声送信部は、前記音声遅延量と前記映像遅延量との差分に応じた値だけ、前記送信部から送信される前記音声信号を遅延させるとしてもよい。

【0038】

この構成によれば、例えば以下のような効果を生ずる。例えば映像遅延量が定数の整数倍として決定される場合、音声遅延量が厳密な遅延量と同一視できる場合であっても、映像遅延量を、音声遅延量と一致させられない場合がある。

30

【0039】

このような場合であっても、映像遅延量を音声遅延量より大きな値として決定し、かつ、音声送信部から送信される音声信号を遅延させることで、映像遅延量を厳密な遅延量に近づける場合と同じ効果が生ずる。つまり、リップシンクの精度が向上される。

【0040】

また、例えば、前記映像遅延部は、前記音声遅延量以下である前記映像遅延量であって、前記映像信号のフレームレートから算出される1フレーム分の時間の整数倍に対応する前記映像遅延量だけ前記映像出力部から出力される前記映像信号を遅延させるとしてもよい。

40

【0041】

この構成によれば、映像遅延量が、当該映像信号におけるフレームレートに応じて決定されるため、当該映像信号の遅延処理がフレーム単位で行われる。つまり、当該遅延処理の煩雑化が抑制される。

【0042】

また、例えば、前記受付部は、前記遅延情報として、前記音声信号を受信して再生する前記外部の音声再生装置から出力される音声の信号である再生音声信号の入力を受け付け、前記映像遅延部は、前記再生音声信号と、前記音声遅延部が遅延させる前の前記音声信号との間の遅延量である前記音声遅延量に応じた前記映像遅延量だけ、前記映像出力部から出

50

力される前記映像信号を遅延させるとしてもよい。

【0043】

この構成によれば、外部の音声再生装置から得られる再生音信号が遅延情報として用いられる。そのため、例えば、映像音声処理装置による、リップシンクの自動化が可能となる。

【0044】

また、例えば、本発明の一態様に係る映像音声処理装置はさらに、前記映像遅延量を示す情報である映像遅延情報を記憶する記憶部を備え、前記映像遅延部は、前記動作モードが前記第二モードである期間に、前記記憶部から読み出した前記映像遅延情報に示される前記映像遅延量だけ、前記映像出力部から出力される前記映像信号を遅延させるとしてもよい。

10

【0045】

この構成によれば、映像音声処理装置において決定された映像遅延量が記憶される。そのため、例えば、映像音声処理装置からの音声信号の送信先として複数の音声再生装置が存在する場合、当該複数の音声再生装置のそれぞれに対応する映像遅延量を記憶部に記憶しておくことができる。

【0046】

その結果、映像音声処理装置は、音声信号の送信先の音声再生装置が変更になった場合であっても、適切な映像遅延量を用いた映像信号の遅延処理を行うことができる。

【0047】

20

また、例えば、前記記憶部は、前記音声再生装置を含む複数の音声再生装置のそれぞれに対応する複数の映像遅延量を示す前記映像遅延情報を記憶し、前記映像遅延部は、前記動作モードが前記第二モードであって、かつ、前記音声送信部が、前記複数の音声再生装置のそれぞれに同時に前記音声信号を送信する場合、(c)前記記憶部に記憶されている前記映像遅延情報に示される前記複数の映像遅延量のうち、最も大きな映像遅延量を選択し、(d)前記映像出力部から出力される前記映像信号を、選択した映像遅延量だけ遅延させるとしてもよい。

【0048】

この構成によれば、例えば以下のような効果を生ずる。例えば複数のユーザが、映像出力部に接続された一つのディスプレイに表示された映像を見ながら、それぞれが装着する、映像音声処理装置と無線通信するヘッドホンで音声を聞く場合を想定する。

30

【0049】

この場合、ヘッドホンごとに音声の遅延量が異なるため、これらヘッドホンそれぞれに対応する映像遅延量も異なるが、これら映像遅延量のうちの最大値に応じて映像信号が遅延される。つまり、少なくとも、これらヘッドホンそれぞれからの再生音が、当該ディスプレイに表示される映像に先行するような極めて不自然な事態の発生は抑制される。

【0050】

なお、これらの全般的または具体的な態様は、システム、方法、集積回路、コンピュータプログラムまたは記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラムおよび記録媒体の任意の組み合わせで実現されてもよい。

40

【0051】

(実施の形態)

以下、実施の形態の映像音声処理装置を、図面を参照しつつ説明する。なお、各図は、模式図であり、必ずしも厳密に図示したものではない。

【0052】

また、以下で説明する実施の形態は、いずれも本発明の一具体例を示すものである。以下の実施の形態で示される数値、形状、材料、構成要素、構成要素の配置位置および接続形態、ステップ、ステップの順序などは、一例であり、本発明を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、最上位概念を示す独立請求項に記載されていない構成要素については、任意の構成要素として説明される。

50

【0053】

図1は、実施の形態におけるAV（Audio Visual）システム10の構成概要を示す図である。

【0054】

図2は、実施の形態におけるAVシステム10の基本的な機能構成を示すブロック図である。

【0055】

図1に示すように、実施の形態におけるAVシステム10は、テレビ100とヘッドホン200とを備える。

【0056】

テレビ100は、放送番組等のAVコンテンツを受信して再生する装置であり、映像音声処理装置110と、ディスプレイ150と、スピーカ160とを備える。

【0057】

ヘッドホン200は、映像音声処理装置110の外部の音声再生装置の一例である。ヘッドホン200は、映像音声処理装置110から送信される音声信号を受信する受信部210と、受信部210が受信した音声信号の再生音を出力するスピーカ220とを有する。

【0058】

なお、ヘッドホン200は右耳用と左耳用の2つのスピーカ220を有しているが、図2では、いずれか一方のスピーカ220の図示は省略している。

【0059】

ユーザは、ヘッドホン200でAVコンテンツの音声を聞きながら、当該AVコンテンツの映像をテレビ100のディスプレイ150で見ることができる。

【0060】

なお、映像音声処理装置110とヘッドホン200との間の通信規格としては、例えば、Bluetooth（登録商標）が採用される。

【0061】

映像音声処理装置110は、図2に示すように、映像出力部111と、音声出力部112と、音声送信部113と、制御部114と、受付部115と、音声遅延部116と、映像遅延部117と、映像信号処理部118と、音声信号処理部119とを備える。

【0062】

映像出力部111は、映像信号を出力する。本実施の形態では、映像出力部111は、映像信号処理部118から映像遅延部117を介して取得した映像信号を、ディスプレイ150に出力する。その結果、当該映像信号に基づく映像がディスプレイ150に表示される。

【0063】

音声出力部112は、当該映像信号に対応する音声信号を出力する。本実施の形態では、音声出力部112は、音声信号処理部119から音声遅延部116を介して取得した音声信号を、スピーカ160に出力する。その結果、当該音声信号に基づく音声、つまり、ディスプレイ150に表示される映像に対応する音声がスピーカ160から出力される。

【0064】

音声送信部113は、当該映像信号に対応する音声信号を、映像音声処理装置110の外部の音声再生装置であるヘッドホン200に送信する。

【0065】

本実施の形態では、音声送信部113は、音声信号処理部119から取得した音声信号を、ヘッドホン200に送信する。その結果、当該音声信号に基づく音声がヘッドホン200から出力される。

【0066】

具体的には、ヘッドホン200では、音声送信部113から送信された音声信号を受信部210が受信し、音声の再生のための所定の処理を行う。これにより、例えばディスプ

10

20

30

40

50

レイ 150 に表示される映像に対応する音声、ヘッドホン 200 が備えるスピーカ 220 から出力される。

【0067】

制御部 114 は、映像音声処理装置 110 の動作モードを調整モードおよび視聴モードの一方から他方へ切り換える。また、制御部 114 は、映像出力部 111 等の映像音声処理装置 110 が備える各構成要素の制御も行う。

【0068】

なお、調整モードは、第一モードの一例であり、音声出力部 112 から音声信号が出力され、かつ、音声送信部 113 から音声信号が送信される動作モードである。具体的には、後述する、リップシンクのための調整（以下、「同期調整」という。）を実行する場合の動作モードである。

10

【0069】

また、視聴モードは、第二モードの一例であり、映像出力部 111 から映像信号が出力され、かつ、音声送信部 113 から音声信号が送信される動作モードである。つまり、ユーザが、ディスプレイ 150 に表示された映像を見ながら、ヘッドホン 200 で当該映像に対応する音声を聞く場合の動作モードである。

【0070】

本実施の形態では、視聴モードでは、スピーカ 160 からの音声出力は停止される。

【0071】

なお、映像音声処理装置 110 は、ディスプレイ 150 とスピーカ 160 とを用いて AV コンテンツをユーザに視聴させる通常の動作モード（通常モード）でも動作する。しかし、当該通常モードは、テレビ 100 としての一般的な動作モードであるためその説明は省略する。

20

【0072】

受付部 115 は、動作モードが調整モードである期間に、音声出力部 112 から出力される音声信号を遅延させる量である音声遅延量を特定する遅延情報の入力を受け付ける。

【0073】

本実施の形態では、調整モードにおいて、当該遅延情報の入力のためのユーザインターフェース画面がディスプレイ 150 に表示される。ユーザインターフェース画面については図 4 を用いて後述する。

30

【0074】

音声遅延部 116 は、受付部 115 が受け付けた遅延情報によって特定される音声遅延量に応じて、音声出力部 112 から出力される音声信号を遅延させる。つまり、スピーカ 160 から出力される音声、当該音声遅延量に応じて遅延される。

【0075】

映像遅延部 117 は、動作モードが視聴モードである期間に、映像出力部 111 から出力される映像信号を当該音声遅延量に応じた映像遅延量だけ遅延させる。つまり、調整モードにおいて決定された音声遅延量に応じて、ディスプレイ 150 に表示される映像が遅延される。

【0076】

40

本実施の形態の映像音声処理装置 110 は、上記構成を有することで、リップシンクのための処理、つまり、再生時における映像信号と音声信号との同期のための処理を効率よく実行することができる。

【0077】

具体的には、本実施の形態では、上述のように音声信号の送信の通信規格として、Bluetooth（登録商標）が採用されている。

【0078】

また、ヘッドホン 200 では音声信号のバッファリングが行われており（バッファは図 2 に図示せず）、これにより、ヘッドホン 200 では、ヘッドホン 200 で再生されるべき音声信号が途切れることなく再生される。

50

【0079】

しかしながら、音声信号のバッファリング等の処理に起因して、ヘッドホン200での再生音が本来的な再生タイミングより遅れて再生され得る。その結果、ディスプレイ150での再生映像と、ヘッドホン200での再生音との間にずれが生じ得る。

【0080】

そこで、本実施の形態の映像音声処理装置110では、音声遅延部116と映像遅延部117との処理により、ディスプレイ150での再生映像と、ヘッドホン200での再生音との間の同期のための処理を効率よく実行することができる。

【0081】

なお、本実施の形態では、映像信号処理部118は、例えば、テレビ100が有するチューナ（図示せず）から受け取ったストリームから映像信号を取得し、映像遅延部117に出力する。また、音声信号処理部119は、当該ストリームから音声信号を取得し、音声遅延部116に出力する。

10

【0082】

つまり、映像信号処理部118および音声信号処理部119は、テレビ100で再生される映像および音声のソースである信号を映像音声処理装置110に与える装置であり、映像音声処理装置110の外部に備えられていてもよい。つまり、映像信号処理部118および音声信号処理部119は、映像音声処理装置110に必須の要素ではない。

【0083】

以下、図3～図6を用いて、本実施の形態の映像音声処理装置110の処理の流れを説明する。

20

【0084】

図3は、実施の形態の映像音声処理装置110における基本的な処理の流れを示すフロー図である。

【0085】

制御部114は、例えばユーザからの指示により、映像音声処理装置110の動作モードを視聴モードから調整モードに切り換える（S1）。

【0086】

動作モードが調整モードである期間に、受付部115は、ユーザの所定の操作により、遅延情報の入力を受け付ける（S2）。例えば、“200millisecond（ms）”という音声遅延量そのもの、または、“+12”等の音声遅延量の大きさを表す数値等が、遅延情報として入力される。

30

【0087】

音声遅延部116は、当該遅延情報によって特定される音声遅延量に応じて音声出力部112から出力される音声信号を遅延させる（S3）。

【0088】

その後、制御部114は、例えばユーザからの指示により、音声遅延量を決定した後で映像音声処理装置110の動作モードを調整モードから視聴モードに切り換える（S4）。

【0089】

映像遅延部117は、動作モードが視聴モードである期間に、映像出力部111から出力される映像信号を当該音声遅延量に応じた映像遅延量だけ遅延させる（S5）。

40

【0090】

以上の処理を実行する映像音声処理装置110の具体的な動作を、図4および図5を参照しながら説明する。

【0091】

図4は、実施の形態の映像音声処理装置110が出力するユーザインターフェース画面151の一例を示す図である。

【0092】

図5は、実施の形態の映像音声処理装置110が行う同期調整を説明するための図であ

50

る。

【0093】

映像音声処理装置110は、調整モードで動作する場合、図4に示すようなユーザインターフェース画面151を、ディスプレイ150に出力する。

【0094】

また、この調整モードでは、スピーカ160およびヘッドホン200の双方から、同期調整のための音声として、例えば所定の間隔ごと（例えば、1sec～2secごと）のパルス音出力される。

【0095】

また、スピーカ160から出力される所定の間隔ごとのパルス音に同期して、例えば、図5に示すような、ボールが床面で跳ね返されることで単振動する動画がユーザインターフェース画面151に表示される。具体的には、ボールが床面に当たるタイミングでスピーカ160からパルス音出力される。

10

【0096】

また、同期調整が完了されていない時点では、図5の(a)に示すように、ヘッドホン200からの出音のタイミングは、スピーカ160からの出音のタイミングに対して遅れている。

【0097】

従って、ユーザは、例えば、左右の一方の耳で、ヘッドホン200からの音声を聞きながら、他方の耳で、スピーカ160からの音声を聞いた場合、左右の耳で知覚される音声の間に時間的なずれがあることが認識される。

20

【0098】

このような状況において、例えばユーザがテレビ100のリモコン170の十字キーを操作することで、遅延情報が映像音声処理装置110に入力される。

【0099】

図4に示す例では、ユーザインターフェース画面151には、音声遅延量を特定する遅延情報として、設定値“+12”が、設定値表示フィールド152に表示されている。この設定値は、例えばユーザがリモコン170の十字キーを操作することで変更される。さらに、当該設定値が遅延情報として受付部115に受け付けられる。

【0100】

30

具体的には、正の整数である設定値に単位遅延量dを乗じた値が、音声遅延量として扱われる。単位遅延量dは、例えば、映像出力部111から出力される映像信号のフレームレートから算出される1フレーム分の時間（以下、「1フレーム時間」という。）である。

【0101】

例えば、当該フレームレートが60Frames/secである場合、単位遅延量dは、1フレーム時間である(50/3(=16.66666...))msecである。そのため、設定値が“12”である場合、(50/3)msecに12を乗じた結果である200msecが音声遅延量として算出される。なお、この算出は、例えば、受付部115、制御部114、または音声遅延部116によって行われる。

40

【0102】

音声遅延部116は、このようにして得られた音声遅延量に応じて、音声信号処理部119から受信する音声信号を遅延させる。その結果、音声出力部112から出力される音声信号が遅延される。

【0103】

例えば、音声遅延部116による遅延の単位が0.1msecであれば、音声出力部112から出力される音声信号が、200.0msecだけ遅延される。なお、音声遅延量と実際の音声の遅延量とが厳密に一致しなくてもよい。例えば、音声遅延部116による遅延の単位が3msecである場合、実際の音声の遅延量が201msecであってもよい。

50

【0104】

このように、ユーザにより入力される遅延情報（設定値）に応じて、音声出力部112から出力される音声信号が遅延され、その結果、スピーカ160から出力される音声が遅延される。

【0105】

そのため、ユーザは、スピーカ160からの音声とヘッドホン200からの音声の間に時間的なずれが最も小さくなるように、設定値を変更することができる。

【0106】

その結果、図5の（b）に示すように、当該時間的なずれが最も小さくなると知覚される音声遅延量Dが決定される。例えば、設定値が“+12”である場合に、ユーザによりリモコン170の所定のボタンが押下されることで、設定値“+12”が、同期調整のための遅延情報として決定される。つまり、設定値“+12”に対応する“200msec”が、音声遅延量Dとして特定される。

10

【0107】

また、このように特定された音声遅延量Dだけ音声出力部112から出力される音声信号が遅延され、これにより、図5の（b）に示すように、ヘッドホン200からの出音のタイミングとスピーカ160からの出音のタイミングとが一致（略一致も含む、以下同じ）する。つまり、ヘッドホン200とスピーカ160との間における音声信号の同期がなされる。

【0108】

制御部114は、このように特定された音声遅延量Dを取得し、映像遅延部117に送信する。

20

【0109】

映像遅延部117は、受信した音声遅延量Dに応じて映像遅延量Vを決定し、映像信号処理部118から受信する映像信号を映像遅延量Vだけ遅延させる。その結果、映像出力部111から出力される映像信号が映像遅延量Vだけ遅延される。

【0110】

ここで、上記のように、単位遅延量dが映像信号における1フレーム時間である場合、つまり、音声遅延量Dが、当該1フレーム時間の整数倍の値である場合、例えば、音声遅延量Dがそのまま映像遅延量Vとして扱われる。

30

【0111】

例えば、音声遅延量Dが“200msec”である場合、映像遅延量Vも“200msec”と決定される。

【0112】

この場合、映像遅延部117は、映像信号を映像信号処理部118から受信して12フレーム遅延させて映像出力部111に出力する。これにより、映像出力部111からディスプレイ150に出力される映像信号は、映像遅延量Vである“200msec”だけ遅延される。

【0113】

なお、単位遅延量dが映像信号における1フレーム時間である場合、映像遅延部117は、音声遅延量Dそのものではなく、単位遅延量dに乗算される設定値を受け取ってもよい。例えば、設定値が“+12”である場合、当該設定値を受け取った映像遅延部117は、映像遅延量Vとして“+12”を決定し、上記のように映像信号を12フレーム遅延させる。

40

【0114】

その結果、映像出力部111からディスプレイ150に出力される映像信号は、音声遅延量Dと同じ値である“200msec”だけ遅延される。

【0115】

このように、映像遅延部117が、音声遅延量Dに応じた映像遅延量Vだけ映像信号を遅延させることで、ヘッドホン200からの出音のタイミングとディスプレイ150での

50

映像の表示のタイミングとが一致する。

【0116】

ここで、本実施の形態においては、以上の映像遅延部117による映像信号の遅延処理は、調整モードでの動作期間中にも実行される。つまり、スピーカ160からのパルス音の出力のタイミングの変化に追従して、図4に示すユーザインターフェース画面151に表示されたボールが床面に当たるタイミングが変化する。

【0117】

なお、映像遅延部117による映像信号の遅延処理は、少なくとも、映像音声処理装置110の動作モードが視聴モードである期間に行われればよい。つまり、調整モードにおいてスピーカ160およびヘッドホン200から出力される、同期調整用の音声に対応する映像（例えば、図4におけるボールの映像）は、ユーザインターフェース画面151に表示されていなくてもよい。

10

【0118】

また、ユーザインターフェース画面151は、ディスプレイ150の表示領域の一部のみに表示されてもよい。例えば、通常の放送番組の映像に重畳させて、設定値表示フィールド152等の、設定値の入力および確認に必要なユーザインターフェース用の映像を表示させてもよい。この場合、同期調整用の音声として、当該放送番組の音声が用いられてもよい。

【0119】

また、ユーザインターフェース画面151は必須ではなく、ディスプレイ150に表示される画像、テレビ100に設けられたランプ、またはスピーカ160からの音声等を介して、調整モードで動作中であることをユーザに知覚させてもよい。

20

【0120】

この場合、ユーザは、調整モードで動作中であることを認識できるため、例えばリモコン170の十字キーを操作することで、スピーカ160からの音声を、ヘッドホン200からの音声と同期させるように、スピーカ160からの音声を遅延させることができる。

【0121】

また、調整モードから視聴モードへの切り替えは、例えば、上述の、設定値の決定のための、リモコン170の所定のボタンの押下をトリガとして実行される。また、例えば、受付部115が受け付ける遅延情報（設定値）が変更されない期間が閾値を越えたことをトリガとして、調整モードから視聴モードに切り替えられてもよい。

30

【0122】

映像音声処理装置110の動作モードが視聴モードとなった場合、ユーザは、上記のように遅延された映像をディスプレイ150で見ることができる。

【0123】

具体的には、ヘッドホン200から出力される音声の、本来的な出力タイミングからのずれ量だけ遅延された映像がディスプレイ150に表示される。その結果、ヘッドホン200での再生音とディスプレイ150に表示される再生映像とは同期される。

【0124】

なお、映像音声処理装置110の動作モードが視聴モードである場合に、スピーカ160からの音声出力を停止させずに継続させてもよい。この場合、音声遅延部116は、例えば上記の音声遅延量D（または映像遅延量V）だけ、音声出力部112から出力される音声信号を遅延させればよい。

40

【0125】

これにより、ヘッドホン200を装着しているユーザが視聴するAVコンテンツを、当該ヘッドホン200を装着していないユーザにも視聴させることができる。つまり、当該ヘッドホン200を装着していないユーザに、ディスプレイ150に表示された再生映像と同期した音声を、スピーカ160によって提供することができる。

【0126】

また、映像音声処理装置110とヘッドホン200との通信が終了した場合、例えば、

50

当該通信の終了をトリガとして、映像音声処理装置 110 の視聴モードから通常モードに切り替わる。また、映像遅延部 117 は映像信号の遅延処理を終了する。

【0127】

このように、本実施の形態の映像音声処理装置 110 は、同期調整において、スピーカ 160 からの音声とヘッドホン 200 からの音声との比較により、ヘッドホン 200 から出力される音声の本来的な出力タイミングからのずれ量が決定される。

【0128】

つまり、ヘッドホン 200 での再生音との比較対象として、ヘッドホン 200 での再生音と同期されるべき映像そのものではなく、当該映像と同期が保障されたスピーカ 160 からの出力音が用いられ、これにより、当該映像の遅延量が決定される。

10

【0129】

従って、本実施の形態の映像音声処理装置 110 によれば、再生時における映像信号と音声信号との同期のための処理を効率よく行うことができる。

【0130】

なお、上記説明では、単位遅延量 d として、映像信号のフレームレートから算出される 1 フレーム時間が用いられるとした。しかしながら、単位遅延量 d に特に限定はなく、例えば、1 msec などの、映像音声処理装置 110 が扱う映像信号の 1 フレーム時間よりも小さな数値であってもよい。

【0131】

これにより、例えば、遅延情報として映像音声処理装置 110 に入力される値の厳密性を向上させることができる。つまり、音声遅延量 D として、同期調整のためのより正確な値を決定することが可能となる。

20

【0132】

ここで、このように、単位遅延量 d として 1 フレーム時間よりも小さな数値が採用された場合、音声遅延量 D は、1 フレーム時間の整数倍にならない場合がある。

【0133】

つまり、音声遅延量 D として、厳密な同期のための正確な値が決定された場合であっても、上述のように映像遅延部 117 が映像信号をフレーム単位で遅延させる場合、ヘッドホン 200 での再生音とディスプレイ 150 に表示される再生映像とが、厳密には同期しないことになる。

30

【0134】

そのため、映像遅延部 117 は、映像信号をフレーム単位で遅延させるのではなく、1 フレーム時間よりも小さな単位で映像信号を遅延させてもよい。これにより、ヘッドホン 200 での再生音とディスプレイ 150 での再生映像とのより厳密な同期が可能となる。

【0135】

また、例えば映像遅延部 117 による映像遅延の処理負荷を増加させないために、映像信号をフレーム単位で遅延させることを維持する場合、映像遅延量 V を、音声遅延量 D よりも小さな値に決定してもよい。これにより、少なくとも、ヘッドホン 200 からの再生音がディスプレイ 150 に表示される再生映像に先行するような極めて不自然な事象の発生は防止される。

40

【0136】

図 6 は、1 フレーム時間 S と、スピーカ 160 およびヘッドホン 200 の間の出音タイミングのずれ量との関係を示す図である。

【0137】

例えば、スピーカ 160 からの音声とヘッドホン 200 からの音声との間の厳密な時間的なずれ量が $D1$ である場合を想定する。この場合、スピーカ 160 からの音声とディスプレイ 150 の再生映像とは同期されているため、ディスプレイ 150 の再生映像とヘッドホン 200 からの音声との間の厳密な時間的なずれ量も $D1$ であるとみなされる。

【0138】

ここで、この $D1$ を音声遅延量として示す遅延情報が、映像音声処理装置 110 に入力

50

された場合を想定する。

【0139】

この想定において、映像遅延部117が映像信号をフレーム単位で遅延させた場合、映像遅延量Vは1フレーム時間Sの整数倍である。つまり、図6において、 $t(0)$ を起点（映像遅延量=0）とした場合、 $t(0)$ と、 $t(1)$ ・・・、 $t(n+1)$ 、・・・のいずれかの値が、映像遅延量Vとして決定される。なお、 $t(n) = S \cdot n$ （ n は正の整数）である。

【0140】

この場合、例えば制御部114または映像遅延部117は、遅延情報によって特定される音声遅延量D1以下の値を、映像遅延量Vとして決定する。

10

【0141】

図6に示す場合、音声遅延量D1以下であって、かつ、音声遅延量D1に最も近い、1フレーム時間Sの n 倍である、 $t(n)$ が、映像遅延量Vとして決定される。

【0142】

例えば、音声遅延量D1が210msecであり、1フレーム時間Sが $(50/3)$ msecである場合、210msec以下であり、かつ、210msecに最も近い、 $(50/3)$ msecの12倍である、“200msec”が、映像遅延量Vとして決定される。なお、この場合、映像遅延量Vとしては、上述のように、“200msec”に対応するフレーム数である“12”が決定されてもよい。

【0143】

20

このように、映像音声処理装置110に入力される遅延情報から特定される音声遅延量Dが定数の整数倍である場合、当該定数が小さいほど、音声遅延量Dは、リップシンクのための本来的な遅延量により近い値をとりうる。つまり、音声遅延量Dの厳密性を向上させることができる。

【0144】

また、映像遅延量Vが1フレーム時間の整数倍である場合など、映像遅延量Vが、厳密性の高い音声遅延量Dの値と一致する値をとり得ない場合、上記のように、音声遅延量D以下であり、かつ、音声遅延量Dに近い値を映像遅延量Vとして決定する。これにより、ディスプレイ150とヘッドホン200との間のリップシンクの問題の発生は実質的に防止され、かつ、音声映像が先行するような極めて不自然な状況の発生は防止される。

30

【0145】

また、図6に示す場合において、音声遅延量D1より大きな値が、映像遅延量Vとして決定されてもよい。例えば、音声遅延量D1に最も近い、1フレーム時間Sの整数倍である、 $t(n+1)$ が、映像遅延量Vとして決定されてもよい。

【0146】

この場合、例えば、音声送信部113からヘッドホン200に送信する音声信号を遅延させることで、ヘッドホン200での再生音とディスプレイ150での再生映像とを同期させることができる。

【0147】

例えば、音声遅延量D1が186msecであり、1フレーム時間Sが $(50/3)$ msecである場合において、186msecより大きく、かつ、 $(50/3)$ msecの整数倍（12倍）である、“200msec”が、映像遅延量Vとして決定された場合を想定する。

40

【0148】

この場合、映像遅延量Vが音声遅延量D1よりも14msec大きいため、なんら手当てをしない場合、ヘッドホン200での再生音が、14msecだけ、ディスプレイ150での再生映像に先行することになる。

【0149】

そこで、音声送信部113からヘッドホン200に送信する音声信号を、14msecだけ遅延させる。

50

【0150】

これにより、音声遅延量D1が、リップシンクのための本来的な遅延量と同一である場合、理論上、ヘッドホン200での再生音とディスプレイ150での再生映像とは完全に同期する。また、音声遅延量D1の誤差を考慮した場合であっても、ヘッドホン200とディスプレイ150との間のリップシンクの厳密性は向上する。

【0151】

つまり、簡単にいうと、映像音声処理装置110は、映像の遅延量を本来的な必要量よりも大きくし、かつ、ヘッドホン200への音声を遅延させることで、ヘッドホン200での再生音とディスプレイ150での再生映像とを厳密に同期させることも可能である。

【0152】

また、以上説明した、映像音声処理装置110の同期調整において用いられた映像遅延量Vを記憶しておいてもよい。

【0153】

図7は、実施の形態における映像音声処理装置110が記憶部130を備える場合の基本的な機能構成を示すブロック図である。

【0154】

例えば、映像音声処理装置110の制御部114が、上記の同期調整において決定された映像遅延量Vを、映像遅延情報131として記憶部130に記憶させる。

【0155】

これにより、その後、ヘッドホン200と映像音声処理装置110との間の通信が一旦終了した後に、ヘッドホン200と映像音声処理装置110とが通信を再開した場合、記憶された映像遅延量Vが用いられた自動的な同期調整が実行される。つまり、制御部114は、記憶部130から映像遅延量Vを読み出して、映像遅延部117に送信し、映像遅延部117に、映像遅延量Vに応じた映像信号の遅延を行わせることができる。

【0156】

なお、記憶部130に記憶される映像遅延情報131は、映像遅延量Vそのものを示さなくてもよい。例えば、映像遅延量Vに対応する音声遅延量Dを示す映像遅延情報131が、記憶部130に記憶されてもよい。

【0157】

また、映像音声処理装置110が、複数の音声再生装置と通信する場合、記憶部130は、これら複数の音声再生装置のそれぞれに対応する複数の映像遅延量を示す映像遅延情報131を記憶してもよい。

【0158】

図8は、実施の形態のAVシステム10が複数の音声再生装置を備える場合の構成概要を示す図である。

【0159】

図9は、実施の形態における映像遅延情報131のデータ構成例を示す図である。

【0160】

図8に示すように、映像音声処理装置110を備えるテレビ100が、上記のヘッドホン200の他に、2つのヘッドホン(201、202)と通信する場合を想定する。

【0161】

なお、ヘッドホン201および202のそれぞれについても、例えば映像音声処理装置110とのペアリングが終了した後に、図3～図5を用いて説明した同期調整が実行されている。そのため、ヘッドホン201および202のそれぞれに対応する映像遅延量Vが求められている。

【0162】

また、これら3つのヘッドホン(200、201、202)は、互いに機種が異なるため、または個体差により、それぞれの再生音の本来的な再生タイミングからのずれ量(遅延量)が互いに異なる。

【0163】

10

20

30

40

50

そのため、図9に示すように、これら3つのヘッドホン(200、201、202)のそれぞれに対応する映像遅延量Vを示す映像遅延情報131を、それぞれの識別子である外部機器IDと対応付けて記憶部130に記憶させておく。なお、各ヘッドホン(200、201、202)の外部機器IDは、映像音声処理装置110と通信を開始する場合に、各ヘッドホン(200、201、202)から映像音声処理装置110に通知される。

【0164】

また、図9に示す例では、ヘッドホン200の外部機器IDは、“H-A”であり、ヘッドホン201の外部機器IDは、“H-B”であり、ヘッドホン202の外部機器IDは、“H-C”である。

【0165】

このような情報を含む映像遅延情報131を、記憶部130に記憶させておくことで、映像音声処理装置110は、音声信号の送信先のヘッドホンが変更になった場合であっても、適切な映像遅延量を用いた映像信号の遅延処理を行うことができる。

【0166】

ここで、これら3つのヘッドホン(200、201、202)のうちの少なくとも2つが同時に映像音声処理装置110と通信する場合も考えられる。

【0167】

例えば、3人のユーザのそれぞれがヘッドホン(200、201、または、202)を装着して、テレビ100のディスプレイ150に表示される映像を見る場合が考えられる。

【0168】

この場合、映像音声処理装置110の映像遅延部117は、以下の処理を実行する。すなわち、映像遅延部117は、動作モードが視聴モードであって、かつ、音声送信部113が、3つのヘッドホン(200、201、202)のそれぞれに同時に音声信号を送信する場合、(a)記憶部130に記憶されている映像遅延情報131に示される複数の映像遅延量Vのうち、最も大きな映像遅延量Vを選択し、(b)映像出力部111から出力される映像信号を、選択した映像遅延量Vだけ遅延させる。

【0169】

例えば、3つのヘッドホン(200、201、202)それぞれに対応する映像遅延量が、図9に示す値である場合、映像遅延部117が用いる映像遅延量Vとして、ヘッドホン202に対応する“201 msec”が採用される。

【0170】

つまり、映像音声処理装置110は、音声信号の送信先の装置が複数ある場合、その複数の装置のうちの、最も音声の遅延量の大きな装置に合わせて、映像音声処理装置110から出力する映像信号を遅延させる。

【0171】

これにより、少なくとも、これらヘッドホン(200、201、202)それぞれからの再生音が、ディスプレイ150に表示される再生映像に先行するような極めて不自然な事態の発生は抑制される。

【0172】

また、この場合、例えば音声送信部113が、ヘッドホン200および201のそれぞれに送信される音声信号を遅延させてもよい。これにより、ディスプレイ150に表示される当該映像と、これらヘッドホン200および201のそれぞれでの再生音との同期をより厳密にすることができる。

【0173】

例えば、上記のように、映像遅延部117が用いる映像遅延量Vとして“201 msec”が採用された場合を想定する。この場合、音声送信部113は、映像遅延量V“197 msec”に対応するヘッドホン200に対しては、4 msecだけ音声信号を遅延させる。

【0174】

10

20

30

40

50

また、音声送信部 113 は、映像遅延量 V “189 msec” に対応するヘッドホン 200 に対しては、12 msec だけ音声信号を遅延させる。

【0175】

つまり、ヘッドホン 200 および 201 との関係で、相対的に遅く設定された映像遅延量 V に合わせるように、ヘッドホン 200 および 201 のそれぞれに送信される音声信号を遅延させる。これにより、これら 3 つのヘッドホン (200、201、202) の全てについて、リップシンクの問題がより確実に解消される。

【0176】

また、遅延情報は、ユーザインターフェース画面 151 を介して映像音声処理装置 110 に入力されなくてもよい。例えば、ヘッドホン 200 の再生音を示す再生音信号が、遅延情報として映像音声処理装置 110 に入力されてもよい。

10

【0177】

図 10 は、実施の形態における映像音声処理装置 110 が遅延情報として再生音信号を取得する場合の基本的な機能構成を示すブロック図である。

【0178】

図 10 に示すように、ヘッドホン 200 の再生音を示す再生音信号が、遅延情報として、受付部 115 に受け付けられる。

【0179】

例えば、受付部 115 に接続されたマイクロフォン (図示せず) を介して、再生音信号が受付部 115 に入力される。または、受付部 115 に接続された音声入力端子 (図示せず) を介して、再生音信号が受付部 115 に入力される。

20

【0180】

この場合、例えば制御部 114 は、音声遅延部 116 から出力される音声信号に示される音圧レベルのピークのタイミングと、当該再生音信号に入力される音声信号に示される音圧レベルのピークのタイミングとの間の時間的な差分から、音声遅延量 D を特定する。

【0181】

なお、これら 2 種類の信号を用いた音声遅延量 D の特定の手法は、上記手法に限定されない。また、制御部 114 ではなく、例えば、音声遅延部 116 または受付部 115 によって、音声遅延量 D が特定されてもよい。

【0182】

30

また、音声遅延量 D の特定は、一回の上記比較処理によって行われてもよい。また、当該音声遅延量 D の特定は、音声遅延部 116 による音声信号の遅延量を変化させながら、音声遅延部 116 から出力される音声信号と、再生音信号とのずれ量をフィードバックすることによって行われてもよい。

【0183】

また、本実施の形態では、映像音声処理装置 110 とヘッドホン 200 との間の通信規格として、Bluetooth (登録商標) が採用されたとしたが、当該通信規格として Bluetooth (登録商標) 以外の通信規格が採用されてもよい。また、映像音声処理装置 110 とヘッドホン 200 との間が無線通信ではなく有線通信であってもよい。

【0184】

40

つまり、通信規格が採用する手順等の都合により、ヘッドホン 200 での再生音と、ディスプレイ 150 での再生映像との間に、人間が知覚できる程度のずれが生じる場合、当該通信規格の種類に関係なく、映像音声処理装置 110 による同期調整は有効である。

【0185】

また、映像音声処理装置 110 は、テレビ 100 以外の種類の装置に備えられてもよい。例えば、Blu-ray Disc (登録商標) 等の光ディスクまたはハードディスクに記憶された AV コンテンツを再生するレコーダまたはプレーヤに、映像音声処理装置 110 が備えられてもよい。

【0186】

また、映像音声処理装置 110 からの音声信号の送信先の装置は、ヘッドホン 200 以

50

外の種類の音声再生装置であってもよい。

【0187】

例えば、複数のスピーカを備え、映像音声処理装置110と無線または有線で通信するサラウンドシステムに、映像音声処理装置110からの音声信号が送信されてもよい。つまり、映像音声処理装置110による同期調整における音声の発生元である音声再生装置の種類は、ヘッドホンに限定されない。

【0188】

また、上記の実施の形態において、各構成要素は、専用のハードウェアで構成されるか、各構成要素に適したソフトウェアプログラムを実行することによって実現されてもよい。各構成要素は、CPU (Central Processing Unit) またはプロセッサなどのプログラム実行部が、ハードディスクまたは半導体メモリなどの記録媒体に記録されたソフトウェアプログラムを読み出して実行することによって実現されてもよい。ここで、上記実施の形態の映像音声処理装置を実現するソフトウェアは、次のようなプログラムである。

【0189】

すなわち、このプログラムは、コンピュータに、以下の映像音声処理方法を実行させる。

【0190】

当該映像音声処理方法は、映像音声処理装置によって実行される映像音声処理方法であって、前記映像音声処理装置は、映像信号を出力する映像出力部と、前記映像信号に対応する音声信号を出力する音声出力部と、前記映像信号に対応する前記音声信号を、前記映像音声処理装置の外部の音声再生装置に送信する音声送信部とを備え、前記映像音声処理方法は、前記映像音声処理装置の動作モードが、前記音声出力部から前記音声信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第一モードである期間に、前記音声出力部から出力される音声信号を遅延させる量である音声遅延量を特定する遅延情報の入力を受け付ける受付ステップと、前記受付ステップにおいて受け付けられた前記遅延情報によって特定される前記音声遅延量に応じて前記音声出力部から出力される音声信号を遅延させる音声遅延ステップと、前記映像音声処理装置の動作モードが、前記映像出力部から前記映像信号が出力され、かつ、前記音声送信部から前記音声信号が送信される第二モードである期間に、前記映像出力部から出力される映像信号を前記音声遅延量に応じた映像遅延量だけ遅延させる映像遅延ステップとを含む映像音声処理方法である。

【0191】

以上、本発明の一態様に係る映像音声処理装置について、実施の形態に基づいて説明したが、本発明は、この実施の形態に限定されるものではない。本発明の趣旨を逸脱しない限り、当業者が思いつく各種変形を本実施の形態に施したものや、異なる実施の形態における構成要素を組み合わせて構築される形態も、本発明の一態様の範囲内に含まれてもよい。

【産業上の利用可能性】

【0192】

本発明は、放送波またはネットワーク経由で送信されるAVコンテンツを再生するテレビ、および、Blu-ray Disc (登録商標) 等の光ディスク、フラッシュメモリ等の半導体メモリ、またはハードディスク等の記録媒体に記憶されたAVコンテンツを再生するレコーダまたはプレーヤ等のAV機器が備える映像音声処理装置として有用である。

【符号の説明】

【0193】

- 10 AVシステム
- 100 テレビ
- 110 映像音声処理装置
- 111 映像出力部

10

20

30

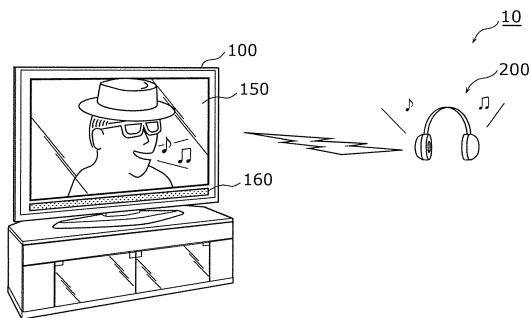
40

50

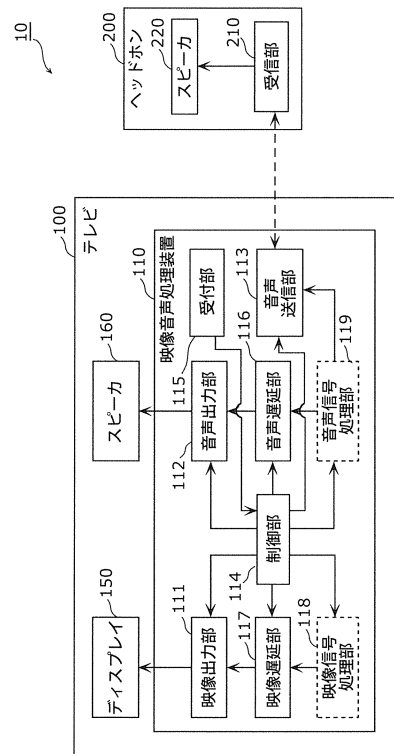
- 112 音声出力部
- 113 音声送信部
- 114 制御部
- 115 受付部
- 116 音声遅延部
- 117 映像遅延部
- 118 映像信号処理部
- 119 音声信号処理部
- 130 記憶部
- 131 映像遅延情報
- 150 ディスプレイ
- 151 ユーザーインターフェース画面
- 152 設定値表示フィールド
- 160、220 スピーカ
- 170 リモコン
- 200、201、202 ヘッドホン
- 210 受信部

10

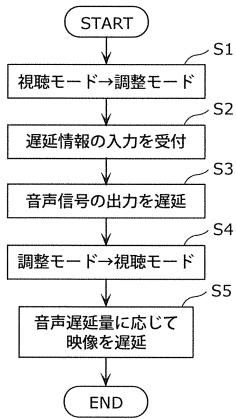
【図 1】



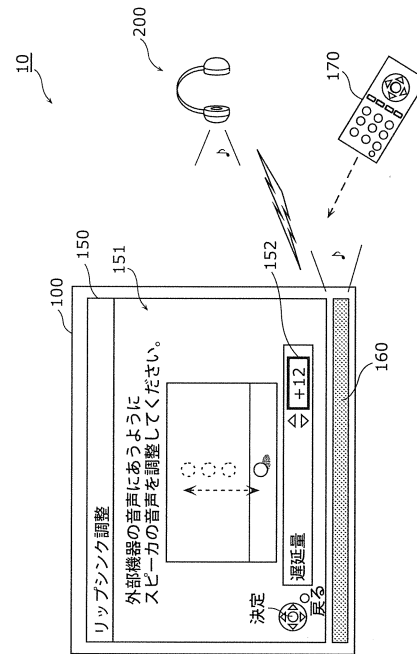
【図 2】



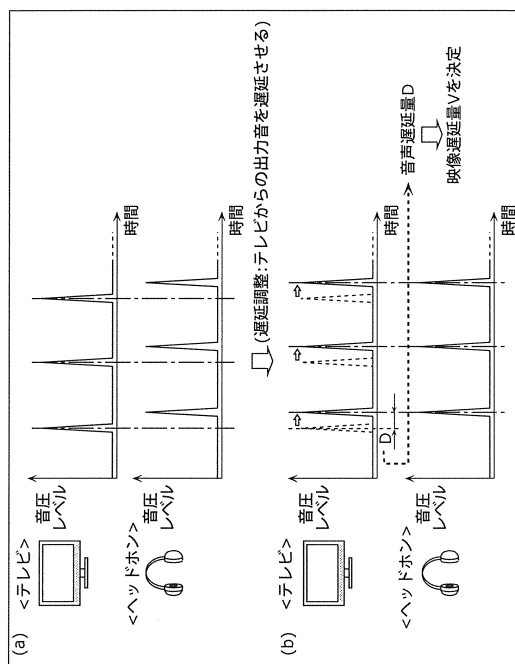
【図 3】



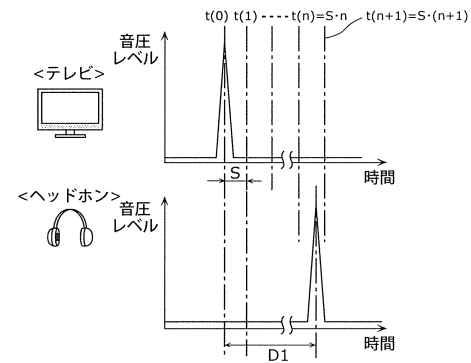
【図 4】



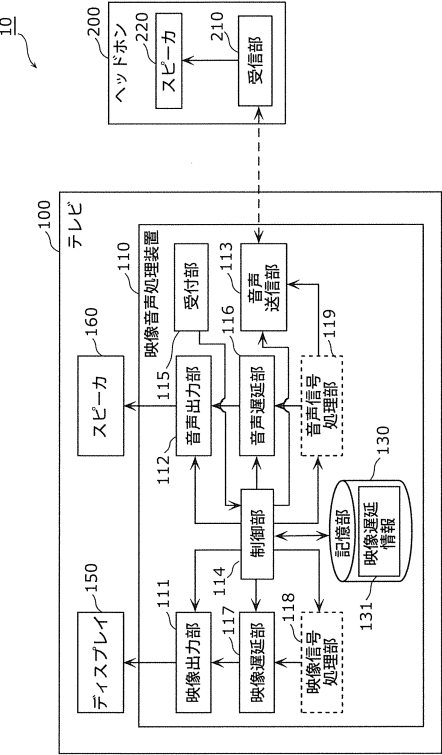
【図 5】



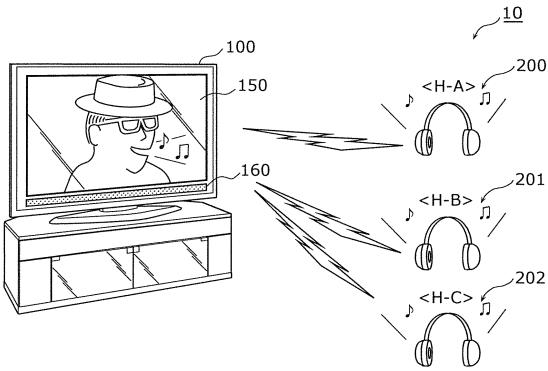
【図 6】



【図 7】



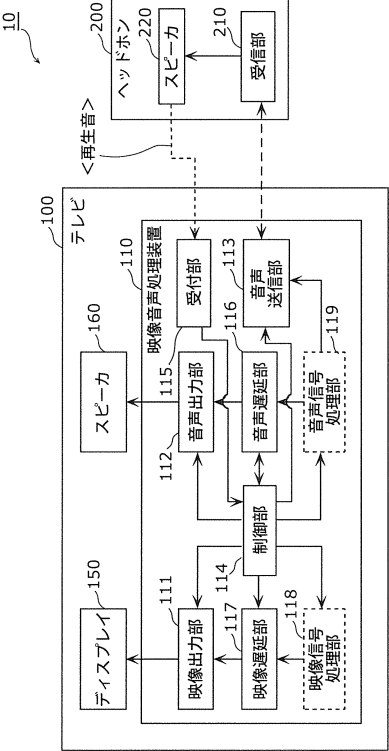
【図 8】



【図 9】

記憶部	
外部機器ID	映像遅延量
H-A	197
H-B	189
H-C	201
...	...

【図 10】



フロントページの続き

審査官 後藤 嘉宏

- (56)参考文献 特開2010-219783 (JP, A)
特表2010-541323 (JP, A)
特開2009-182912 (JP, A)
特開2010-119076 (JP, A)
特開平7-222209 (JP, A)
特開2007-43652 (JP, A)

- (58)調査した分野(Int.Cl., DB名)
H04N 21/00-21/858