

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6410769号
(P6410769)

(45) 発行日 平成30年10月24日 (2018.10.24)

(24) 登録日 平成30年10月5日 (2018.10.5)

(51) Int.Cl.		F I			
HO4N	5/60	(2006.01)	HO4N	5/60	O7O
HO4S	7/00	(2006.01)	HO4S	7/00	32O

請求項の数 18 (全 23 頁)

(21) 出願番号	特願2016-148995 (P2016-148995)	(73) 特許権者	000001007
(22) 出願日	平成28年7月28日 (2016.7.28)		キヤノン株式会社
(65) 公開番号	特開2018-19294 (P2018-19294A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成30年2月1日 (2018.2.1)	(74) 代理人	100076428
審査請求日	平成30年4月11日 (2018.4.11)		弁理士 大塚 康德
		(74) 代理人	100115071
			弁理士 大塚 康弘
		(74) 代理人	100112508
			弁理士 高柳 司郎
		(74) 代理人	100116894
			弁理士 木村 秀二
		(74) 代理人	100130409
			弁理士 下山 治
		(74) 代理人	100134175
			弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 情報処理システム及びその制御方法、コンピュータプログラム

(57) 【特許請求の範囲】

【請求項 1】

音響信号を生成する情報処理システムであって、

複数の撮影装置により撮影領域を複数の方向から撮影することにより得られる画像に基づいて生成される仮想視点画像に関する仮想視点の位置に応じた情報を取得する取得手段と、

前記取得手段により取得される情報によって特定される仮想視点に応じた視界に含まれる前記撮影領域内の地面に応じた位置を、前記音響信号の生成に係る仮想聴取点の位置として決定する決定手段と、

少なくとも前記撮影領域の一部の音を收音する複数のマイクによる收音に基づく複数の收音信号と、前記決定手段により決定される仮想聴取点の位置とに基づいて、当該仮想聴取点に応じた前記音響信号を生成する音響生成手段と
を有することを特徴とする情報処理システム。

【請求項 2】

少なくとも前記撮影領域の一部の音を收音する複数のマイクによる收音に基づく複数の收音信号の中から、前記音響信号を生成するために用いる 1 以上の收音信号を、前記取得手段により取得される前記仮想視点の位置に応じた情報に基づいて選択する選択手段を有し、

前記音響生成手段は、前記選択手段により選択される前記 1 以上の收音信号を用いて、前記決定手段により決定された仮想聴取点に応じた前記音響信号を生成することを特徴と

10

20

する請求項 1 に記載の情報処理システム。

【請求項 3】

前記決定手段は、前記取得手段により取得される情報に基づいて、前記仮想聴取点における聴取方向をさらに決定することを特徴とする請求項 1 又は 2 に記載の情報処理システム。

【請求項 4】

前記決定手段は、前記取得手段により取得される情報によって特定される前記仮想視点に応じた視界に含まれる位置であって前記撮影領域内の地面に応じた高さの位置を、前記仮想聴取点の位置として決定することを特徴とする請求項 1 乃至 3 の何れか 1 項に記載の情報処理システム。

10

【請求項 5】

音響信号を生成する情報処理システムであって、
複数の撮影装置により撮影領域を複数の方向から撮影することにより得られる画像に基づいて生成される仮想視点画像に関する仮想視点の位置に応じた情報を取得する取得手段と、

前記取得手段により取得される情報と前記撮影領域内の被写体の位置とに基づいて、前記音響信号の生成に係る仮想聴取点を決定する決定手段と、

少なくとも前記撮影領域の一部の音を收音する複数のマイクによる收音に基づく複数の收音信号と、前記決定手段により決定される仮想聴取点とに基づいて、当該仮想聴取点に応じた音響信号を生成する音響生成手段と

20

を有することを特徴とする情報処理システム。

【請求項 6】

前記決定手段は、前記取得手段により取得される情報によって特定される仮想視点に応じた視界に含まれる被写体と所定の位置関係にある位置を、前記仮想聴取点の位置として決定することを特徴とする請求項 5 に記載の情報処理システム。

【請求項 7】

前記決定手段は、前記取得手段により取得される情報によって特定される仮想視点に応じた視界に含まれる複数の被写体の位置によって規定される範囲内の位置を、前記仮想聴取点の位置として決定することを特徴とする請求項 5 に記載の情報処理システム。

【請求項 8】

30

前記複数のマイクは少なくとも前記撮影領域内の複数の收音位置の音を收音するように設置され、

前記選択手段は、前記複数の收音位置の中から 1 以上の收音位置を選択し、選択された当該 1 以上の收音位置に対応する 1 以上の收音信号を、前記音響信号を生成するために用いる 1 以上の收音信号として選択することを特徴とする請求項 2 に記載の情報処理システム。

【請求項 9】

前記選択手段は、前記音響信号を生成するために用いる前記 1 以上の收音信号に対応する前記 1 以上の收音位置として、前記決定手段により決定される前記仮想聴取点の位置を中心とする聴取範囲に含まれる收音位置を選択することを特徴とする請求項 8 に記載の情報処理システム。

40

【請求項 10】

前記選択手段は、前記音響信号を生成するために用いる前記 1 以上の收音信号に対応する前記 1 以上の收音位置として、前記取得手段により取得される情報によって特定される前記仮想視点に応じた視界に対応する聴取範囲に含まれる收音位置を選択することを特徴とする請求項 8 に記載の情報処理システム。

【請求項 11】

前記選択手段は、前記音響信号を生成するための前記 1 以上の收音信号に対応する前記 1 以上の收音位置として、前記取得手段により取得される情報と前記複数の撮影装置により撮影される被写体の位置とに基づいて決定される聴取範囲に含まれる收音位置を選択す

50

ることを特徴とする請求項 8 に記載の情報処理システム。

【請求項 1 2】

音響信号を生成する情報処理システムの制御方法であって、

複数の撮影装置により撮影領域を複数の方向から撮影することにより得られる画像に基づいて生成される仮想視点画像に関する仮想視点の位置に応じた情報を取得する取得工程と、

前記取得工程において取得される情報によって特定される仮想視点に応じた視界に含まれる前記撮影領域内の地面に応じた位置を、前記音響信号の生成に係る仮想聴取点の位置として決定する決定工程と、

少なくとも前記撮影領域の一部の音を收音する複数のマイクによる收音信号に基づく複数の收音信号と、前記決定工程において決定される仮想聴取点の位置とに基づいて、当該仮想聴取点に応じた前記音響信号を生成する音響生成工程と
を有することを特徴とする制御方法。

10

【請求項 1 3】

少なくとも前記撮影領域の一部の音を收音する複数のマイクによる收音に基づく複数の收音信号の中から、前記音響信号を生成するために用いる 1 以上の收音信号を、前記取得工程において取得される前記仮想視点の位置に応じた情報に基づいて選択する選択工程を有し、

前記音響生成工程においては、前記選択工程において選択される前記 1 以上の收音信号を用いて、前記決定工程において決定された仮想聴取点に応じた前記音響信号が生成されることを特徴とする請求項 1 2 に記載の情報処理システムの制御方法。

20

【請求項 1 4】

前記決定工程においては、前記取得工程において取得される情報によって特定される前記仮想視点に応じた視界に含まれる位置であって前記撮影領域内の地面に応じた高さの位置が、前記仮想聴取点の位置として決定されることを特徴とする請求項 1 2 又は 1 3 に記載の情報処理システムの制御方法。

【請求項 1 5】

音響信号を生成する情報処理システムの制御方法であって、

複数の撮影装置により撮影領域を複数の方向から撮影することにより得られる画像に基づいて生成される仮想視点画像に関する仮想視点の位置に応じた情報を取得する取得工程と、

前記取得工程において取得される情報と前記撮影領域内の被写体の位置とに基づいて、前記音響信号の生成に係る仮想聴取点を決定する決定工程と、

少なくとも前記撮影領域の一部の音を收音する複数のマイクによる收音に基づく複数の收音信号と、前記決定工程において決定される仮想聴取点とに基づいて、当該仮想聴取点に応じた音響信号を生成する音響生成工程とを有することを特徴とする情報処理システムの制御方法。

30

【請求項 1 6】

前記決定工程は、前記取得工程において取得される情報によって特定される仮想視点に応じた視界に含まれる被写体と所定の位置関係にある位置を、前記仮想聴取点の位置として決定することを特徴とする請求項 1 5 に記載の情報処理システムの制御方法。

40

【請求項 1 7】

前記決定工程は、前記取得工程において取得される情報によって特定される仮想視点に応じた視界に含まれる複数の被写体の位置によって規定される範囲内の位置を、前記仮想聴取点の位置として決定することを特徴とする請求項 1 5 に記載の情報処理システムの制御方法。

【請求項 1 8】

コンピュータを請求項 1 から 1 1 のいずれか 1 項に記載の情報処理システムが備える各手段として機能させるためのコンピュータプログラム。

【発明の詳細な説明】

50

【技術分野】

【0001】

本発明は情報処理システム及びその制御方法、コンピュータプログラムに関し、特に、任意視点映像にふさわしい音響再生信号を生成するための技術に関する。

【背景技術】

【0002】

近年、コンピュータグラフィックス技術等を応用することにより、スポーツ競技場などの広い範囲を取り囲むように設置された複数のカメラによる撮影映像を適宜処理することによって、任意視点の映像を生成するシステムが開発されている。

【0003】

このようなシステムによって生成される任意視点映像に対してより臨場感を持たせるために、それに見合う音響信号を生成、再生することが求められる。

【0004】

そこで、競技場などにおける任意位置を指定してその場にふさわしい音響信号を生成することが知られている（特許文献1）。この構成では、競技場に設置した複数のマイクの指向性と場所に加えて任意の視聴点及び視聴角度を入力し、各チャンネルに対する信号分配の割合と、各マイクと視聴点間の距離による影響を自動的に算出し、サラウンド音声を自動的にミキシングする。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2005-223771号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

しかしながら、上記の構成においては、視点の変化に応じて、聴取範囲および聴取点と聴取方向を自動的に変化させることができない。そのため、視点の移動にふさわしい音場の変化表現が困難である。

【0007】

例えば、特許文献1の構成では、ユーザによって聴取点や聴取方向が直接指定されるため、必ずしも任意視点映像にふさわしい聴取点や聴取方向がユーザによって指定されるとは限らない。

【0008】

そこで本発明は、視点の変化に応じた音響の変化をより適切に表現することが可能な技術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】

【0009】

上記目的を達成するため、本発明による情報処理システムは以下の構成を備える。即ち、音響信号を生成する情報処理システムであって、複数の撮影装置により撮影領域を複数の方向から撮影することにより得られる画像に基づいて生成される仮想視点画像に関する仮想視点の位置に応じた情報を取得する取得手段と、前記取得手段により取得される情報によって特定される仮想視点に応じた視界に含まれる前記撮影領域内の地面に応じた位置を、前記音響信号の生成に係る仮想聴取点の位置として決定する決定手段と、少なくとも前記撮影領域の一部の音を收音する複数のマイクによる收音に基づく複数の收音信号と、前記決定手段により決定される仮想聴取点の位置とに基づいて、当該仮想聴取点に応じた前記音響信号を生成する音響生成手段と、を有する。

【発明の効果】

【0010】

本発明によれば、視点の変化に応じた音響の変化をより適切に表現することが可能な技術を提供することができる。

10

20

30

40

50

【図面の簡単な説明】

【0011】

【図1】任意視点映像生成システムの一構成例を示すブロック図。

【図2】競技場における收音点の配置状況を示す模式図。

【図3】メイン処理の処理手順を示すフローチャート。

【図4】任意視点映像生成システムで用いられる情報のデータ構造を示す図。

【図5】聴取範囲決定処理の処理手順を示すフローチャート。

【図6】視点と聴取範囲、聴取点、聴取方向の関係を示す模式図。

【図7】被写体位置検知処理の処理手順を示すフローチャート。

【図8】收音点選択処理の処理手順を示すフローチャート。

10

【図9】聴取範囲内收音点選択処理の処理手順を示すフローチャート。

【図10】再生信号生成処理の処理手順を示すフローチャート。

【図11】ステレオ再生信号生成処理の処理手順を示すフローチャート。

【図12】サラウンド再生信号生成処理の処理手順を示すフローチャート。

【図13】ヘッドフォン再生信号生成処理の処理手順を示すフローチャート。

【図14】聴取範囲内收音点選択処理の処理手順を示すフローチャート。

【発明を実施するための形態】

【0012】

以下、本発明の実施形態について、図面を参照して説明する。なお、以下の実施形態は本発明を限定するものではなく、また、本実施形態で説明されている特徴の組み合わせの全てが本発明の実施に必須のものとは限らない。なお、同一の構成については、同じ符号を付して説明する。

20

【0013】

<<実施形態1>>

(任意視点映像生成システム)

本発明の一実施形態に係る任意視点映像生成システムについて、図1を参照して説明する。図1は本実施形態に係る任意視点映像生成システムの一構成例を示すブロック図である。本実施形態に係る任意視点映像生成システムは、複数の撮影装置により撮影された複数の画像信号と、複数の收音点において收音された複数の收音信号とに基づいて、任意の視点に対応する画像及び音響を出力する情報処理システムとして動作する。

30

【0014】

図1中、1は收音信号入力部であり、本システムの撮影対象である競技場にまんべんなく散りばめて設定した收音点に設置された複数のマイクロホンによる收音信号を入力し、收音信号の増幅やノイズ除去などを行う。さらに、收音点の特徴を示す付加的な情報を各收音信号に付加し、收音点情報として收音点選択部3へ出力する。

【0015】

図2は、撮影対象の競技場における收音点の設定例を模式的に示す図である。図2中、101は收音点の一つ、102は観客席、103はトラック、104はグラウンドを示す。本実施形態では、図2に示したように、收音点を競技場のあらゆる場所にまんべんなく設定し、收音点の音を常に收音している場合の例を説明する。

40

【0016】

2は聴取範囲決定部であり、視点情報指定部6によって指定された視点情報や、視点映像生成部8によって出力される視点映像、及び、被写体位置検知部9によって出力される被写体位置に基づき、聴取範囲、聴取点、及び、聴取方向を決定する。3は收音点選択部であり、聴取範囲決定部2から出力される聴取範囲、聴取点、聴取方向に従って、音響再生信号生成に用いる收音点を、收音信号入力部が出力する收音点情報の中から適宜選択する。4は音響信号生成部であり、收音点選択部3で選択された收音点の收音信号に基づいて、任意の再生フォーマットの再生音響信号を生成する音響生成を行う。そして、音響再生部11及びMUX15へ出力する。

【0017】

50

5は操作部であり、本システムに対するユーザの各操作指示を受け付ける。6は視点情報指定部であり、操作部5を介して送信されるユーザの操作指示に基づいて、視点情報を生成し、聴取範囲決定部2や視点映像生成部8へ出力する。操作部5は、キーボード、マウス等のポインティング装置、タッチパネル等により実現される。このように、本実施形態では、任意視点映像生成システムは、視点情報指定部6を介して視点情報を取得する。

【0018】

7は映像信号入力部であり、本システムの撮影対象である競技場に設置された複数のカメラによって撮影された映像信号を入力し、映像信号の増幅やノイズ除去などを行う。さらに、撮影時のカメラパラメータを各收音信号に付加し、カメラ撮影情報として視点映像生成部8へ出力する。なお、本実施形態では、任意視点画像として動画像である映像を生成する場合の例を説明するが、静止画像を対象としてもよい。

10

【0019】

8は視点映像生成部であり、視点情報指定部6によって指定された視点情報に従って、複数カメラによる映像を適宜処理することにより、任意視点映像を生成して、聴取範囲決定部2、被写体位置検知部9、映像再生部10、MUX15へ出力する。9は被写体位置検知部であり、視点映像生成部8が生成した視点映像や視点情報に基づき、視点映像に写っている被写体の位置を検知する。後述するように、被写体には、人物や人物以外の特定の物体等が含まれる。10は映像再生部であり、視点映像生成部8が生成した視点映像を再生し、表示部19へ出力する。

【0020】

20

11は音響再生部であり、音響信号生成部4によって生成された音響信号を再生環境に応じて再生する。12はステレオスピーカセットであり、音響信号生成部4が生成したステレオ音響信号を適宜増幅し、音に変えて出力する。13はサラウンドスピーカセットであり、音響信号生成部4が生成したサラウンド音響信号を適宜増幅し、音に変えて出力する。14はヘッドフォンであり、音響信号生成部4が生成したヘッドフォン信号を音に変えて出力する。本実施形態では、ステレオスピーカセット12、サラウンドスピーカセット13、及び、ヘッドフォン14のいずれかにより音響信号の音を出力する例を説明するが、音の再生環境はここに例示したものに限られない。

【0021】

15はMUX(マルチプレクサ)であり、視点映像生成部8が生成した任意視点映像信号と、音響信号生成部4が生成した音響信号を重畳して、一つの映像ストリームデータを作成し、通信部16や出力部18へ出力する。16は通信部であり、MUX15から出力される映像ストリームデータを通信網17に適宜送信する。17は通信網であり、インターネットや公衆回線網などの公共の通信網を示す。18は出力部であり、出力端子を備え、出力端子に接続された外部機器へ、MUX15から出力される映像ストリームデータを出力する。19は表示部であり、映像再生部10が再生する任意視点映像を表示する。表示部19は、液晶パネル、有機ELディスプレイ等により実現される。

30

【0022】

これらの構成要素は不図示のCPU(中央演算処理ユニット)に制御バスを介して接続されており、CPUからの命令指示に従ってその動作が統合制御される。CPUは、不図示の記憶装置からコンピュータプログラムを読み出し、コンピュータプログラムに従って装置全体を制御する。

40

【0023】

(メイン処理)

図1に示す本実施形態の構成において、視点情報に応じて聴取範囲、聴取点、聴取方向を決定し、それに基づいて音場生成を行う処理をフローチャートを参照して以下に説明する。図3は、本実施形態のメイン処理の処理手順を示すフローチャートである。以下の各ステップは、CPUの制御に基づき実行される。

【0024】

S101は、視点情報指定部6において、内部の変更命令バッファに一時蓄積されてい

50

る命令に従って視点情報を変更し、聴取範囲決定部 2 と視点映像生成部 8 へ出力する処理である。

【 0 0 2 5 】

図 4 (a) に、本実施形態における視点情報のデータ構造を示す。図 4 (a) に示すように、本実施形態の視点情報は、視点位置、視点俯角、視点方向、画角を含む。このうち、視点位置は撮影対象である競技場における視点の位置を示す三次元座標である。一例として、本実施形態では、東西に東向きに X 軸、南北に北向きに Y 軸、上下に上向きに Z 軸を取る三次元座標系を設定し、原点を競技場全体の南西角に取る場合を説明する。視点俯角は、視点が向いている俯角であり、水平方向を 0 ° として ± 9 0 ° までの範囲で指定される。視点方向は同じく視点が向いている水平面方向であり、本実施形態では、真北を 0 ° (すなわち、Y 軸正方向) とする絶対方向を基準の正面として、右回り (時計回り) を正、左回り (反時計回り) を負として示す。画角は視点から見た視点映像の上下、左右の幅を角度で示す値である。また、視点から観察される視線の 3 次元的な方向を、以下、視線の方向という。視線の方向は、視点俯角 (視点仰角) と視点の方向を合わせたものに相当する。本実施形態のシステムでは、この情報で指定される任意視点から見た映像を生成することができる。

10

【 0 0 2 6 】

S 1 0 1 の処理が終わると、S 1 0 2 から S 1 0 5 までの音響信号を生成する処理の流れと、S 1 0 6 から S 1 0 8 までの映像を生成する処理の流れと、S 1 0 9 の処理が並行して行われる。以下、これらの処理の流れをそれぞれ説明する。

20

【 0 0 2 7 】

S 1 0 2 では、收音点において取得された音響を取得し、收音信号入力部 1 において、複数のマイクロホンの收音信号の増幅やノイズ除去などを行う。さらに、收音点の特徴を示すヘッダ情報を各收音信号に付加し、收音点情報として收音点選択部 3 へ出力する。図 4 (b) に本実施形態における收音点情報のデータ構造を示す。図 4 (b) のように、本実施形態における收音点情報は、收音点 ID、收音点座標、收音信号を含む。このうち、收音点 ID は、收音点を識別するための番号 (識別情報) である。また、收音点座標は、この收音点の位置を示す座標であり、本実施形態では競技場における水平面座標を示す。收音信号はこの收音点に設置したマイクロホンで收音した音響信号そのものである。

【 0 0 2 8 】

30

次に、S 1 0 3 では、聴取範囲決定部 2 において、S 1 0 1 で視点情報指定部 6 から送信された視点情報に基づいて、聴取範囲、聴取位置、聴取方向を決定する。この処理 (聴取範囲決定処理) の詳細は図 5、図 6 を用いて後述する。次に、S 1 0 4 では、收音点選択部 3 において、視点映像に見合う音響信号を生成するために用いる收音信号を收音した收音点を選択する。この処理 (收音点選択処理) の詳細は図 8 と図 9 を用いて後述する。次に、S 1 0 5 では、音響信号生成部 4 において、各音響再生環境に応じた音響信号を生成する。この処理 (再生信号生成処理) の詳細は図 1 0、図 1 1、図 1 2、図 1 3 を用いて後述する。生成した音響信号は、音響再生部 1 1 や M U X 1 5 へ出力される。処理を終えると、S 1 1 0 へ進む。

【 0 0 2 9 】

40

一方、S 1 0 6 では、複数のカメラにより撮影された映像信号を取得し、映像信号入力部 7 において、映像信号のノイズ除去や輝度調整等を行う。さらに、各カメラの撮影パラメータをヘッダ情報として映像信号に付加し、カメラ撮影情報として收音点選択部 3 へ出力する。図 4 (c) に、本実施形態におけるカメラ撮影情報のデータ構造を示す。図 4 (c) に示すように、本実施形態のカメラ撮影情報は、カメラ位置、カメラ俯角、カメラ方向、画角、焦点距離、映像信号を含む。このうち、カメラ位置は撮影対象である競技場におけるカメラの位置を示す三次元座標である。カメラ俯角は、視点が向いている俯角であり、水平方向を 0 ° として ± 9 0 ° までの範囲で指定される。カメラ方向は同じくカメラが向いている水平面方向であり、本実施形態では、真北 (すなわち、Y 軸正方向) を 0 ° とする絶対方向を基準の正面として、右回り (時計回り) を正、左回り (反時計回り)

50

を負として示す。画角は撮影映像の幅を角度で示す値である。焦点距離は、カメラレンズの光学中心から撮像面までの距離を示す値である。映像信号は、このカメラで撮影した映像そのものの信号である。

【 0 0 3 0 】

次に、S 1 0 7 では、視点映像生成部 8 において、S 1 0 1 で視点情報指定部 6 から送信された視点情報に基づいて、S 1 0 6 で受信した複数カメラ映像を適宜処理、合成し、任意視点映像を生成する。すなわち、複数の画像信号に基づいて、視点位置及び視線の方向に応じた画像を生成する画像生成を行う。このような、複数のカメラ映像から任意視点の映像を合成する手法は公知であり、当分野において一般的に行われているため、詳細な説明は行わない。

10

【 0 0 3 1 】

次に、S 1 0 8 では、被写体位置検知部 9 において、S 1 0 6 で受信した複数カメラの映像と S 1 0 7 で生成した任意視点映像を解析することにより、任意視点映像に写っている被写体が実際に存在する競技場における位置を検知する。この処理（被写体位置検知処理）の詳細は図 7 を用いて後述する。処理を終えると、S 1 1 0 へ進む。

【 0 0 3 2 】

さらに、S 1 0 9 では、視点情報指定部 6 において、ユーザが操作部 5 を介して入力した視点変更指示を受け付けて、視点情報変更命令に変換して内部の変更命令バッファに一次蓄積する。処理を終えると、S 1 1 0 へ進む。

【 0 0 3 3 】

20

S 1 1 0 は、M U X 1 5 において、S 1 0 5 で生成した音響再生信号と、S 1 0 7 で生成した任意視点映像信号を重畳・合成し、一つの映像ストリームデータとしてまとめて、通信部 1 6 や出力部 1 8 へ出力する処理である。処理を終えると、S 1 1 1 へ進む。

【 0 0 3 4 】

S 1 1 1 は、不図示の C P U において、本実施形態のシステムの出力先を判定する処理である。出力先が再生機器の場合は、S 1 1 2 へ進む。出力先が通信網の場合は、S 1 1 3 へ進む。出力先が外部機器の場合は、S 1 1 4 へ進む。

【 0 0 3 5 】

S 1 1 2 は、音響再生部 1 1 と、映像再生部 1 0 において、S 1 0 5 で生成した音響再生信号と、S 1 0 7 で生成した任意視点映像信号を、各々音響再生環境や表示部 1 9 に同期して出力する処理である。このような処理は一般的な映像出力装置において一般的に行われており、公知であるため、詳細な説明はしない。この処理により、任意視点映像とそれに見合う音響信号が同期して再生されるため、映像再生時の臨場感を高めることができる。処理を終えると、S 1 1 5 へ進む。

30

【 0 0 3 6 】

S 1 1 3 は、通信部 1 6 において、S 1 1 0 で作成した映像ストリームデータを、通信網 1 7 を経由して外部に送信する処理である。処理を終えると、S 1 1 5 へ進む。

【 0 0 3 7 】

S 1 1 4 は、出力部 1 8 において、S 1 1 0 で作成した映像ストリームを外部出力端子に繋がれた外部機器へ出力する処理である。処理を終えると、S 1 1 5 へ進む。

40

【 0 0 3 8 】

S 1 1 5 は、不図示の C P U において、本フロー全体で行われているメイン処理を終るかどうかを判定する処理である。この判定の結果、処理を終了する場合（S 1 1 5 で Y E S ）は、本フローの処理を終了する。終了しない場合（S 1 1 5 で N O ）は、S 1 0 1 へ戻る。

【 0 0 3 9 】

（聴取範囲決定処理）

図 5 は、本実施形態における S 1 0 3 の聴取範囲決定処理の詳細な処理手順を示すフローチャートである。なお、本フローチャートにおける処理は全て聴取範囲決定部 2 において行われる。

50

【 0 0 4 0 】

まず、S 2 0 1では、聴取範囲決定部2の内部RAM（不図示）に格納されている聴取点情報を初期化する。ここで、聴取点情報のデータ構造を図4（d）に示す。本実施形態における聴取点情報は、聴取範囲、聴取点、聴取方向を含む。このうち、聴取範囲は、撮影対象である競技場において、視点映像に没入した場合にすぐ近くに聴こえるであろう音が発生している範囲を示すデータである。本実施形態では、水平面上の4点の座標が格納されており、これらを結ぶことによってできる四角形で囲まれた範囲を聴取範囲とする。後述するように、聴取範囲は、音響信号を生成するために用いる収音信号の収音点を選択するための基準となる場所的範囲として機能する。また、聴取点は後の処理で再生信号を生成する際に、収音信号を配置する際の基準とする点であり、同じく水平面上の座標が格納される。また、聴取方向は同じく後の処理で再生信号を生成する際に、収音信号を配置する際の基準とする方向である。本実施形態では、聴取点からみた正面方向を視点方向と同じく絶対角度で示す。S 2 0 1ではこの聴取点情報の各データを全て初期化する。

10

【 0 0 4 1 】

具体的には、S 2 0 2ではS 1 0 1で視点情報指定部6によって指定された視点情報の仰角が -10° より下かどうかを判定する。この処理の趣旨は、指定された視点情報が俯瞰視点であるか、水平視点であるかどうかを判定することである。俯瞰視点とは、視点が、スタジアム等の撮影対象を情報から俯瞰的に観察する位置にある場合をいい、水平視点とは、視点が、撮影対象をほぼ真横から水平に観察する位置にある場合をいう。よって、本実施形態では視点切り替えの判定基準を仰角 -10° としているが、これは一例であり、実施状況に応じて別の判定基準を用いてもよい。例えば、撮影対象が体育館におけるバスケットコートであれば、判定基準となる仰角を -20° のようにもっと深くしてもよい。そうすることで観察対象の性質上観察の範囲が上下に大きい場合に、俯瞰視点として判定する場合を適切に設定して、聴取範囲を適切に決定し、臨場感ある音響信号を表現することが可能となる。また、後の処理であるS 2 0 3によって計算される画角の水平面投影範囲を先に計算しておき、投影範囲が予め定めた範囲に収まれば俯瞰視点とし、収まらなければ水平視点としてもよい。このように、本実施形態では、視線の仰角が予め定められた負の値を下回るか否かを判定し、その判定結果に応じて、聴取範囲を決定するための処理を分岐させることで、音響信号生成の基礎となる聴取範囲を適切に決定することができる。

20

30

【 0 0 4 2 】

このような判定の結果、 -10° より下、つまり、俯瞰視点の場合（S 2 0 2でYES）はS 2 0 3へ進む。そうでない場合、つまり、水平視点の場合（S 2 0 2でNO）は、S 2 0 5へ進む。

【 0 0 4 3 】

S 2 0 3では、この視点情報の画角を撮影対象の競技場等の競技面、つまり、水平面に投影した場合の範囲を計算し、この範囲を聴取範囲とする。一例として、視点位置（15、0、10）、視点俯角 -45° 、視点方向 0° （すなわち、Y軸正方向）、水平画角 20° 、垂直画角 14° の視点の視点映像を、高さ0の競技面に投影する場合を説明する。このとき、投影面（ $Z=0$ ）の中心点の水平面座標は（15、10）であり、投影面は、上端のY座標が $10 \times \tan 52^\circ = 12.8$ 、下端のY座標が $10 \times \tan 38^\circ = 7.8$ の台形になる。ここで、垂直画角 14° の半分は $14^\circ / 2 = 7^\circ$ であり、 $52^\circ = 45^\circ + 7^\circ$ 、 $38^\circ = 45^\circ - 7^\circ$ である。

40

【 0 0 4 4 】

台形の上端、下端の中心点と視点との距離は、各々 $10 / \cos 52^\circ = 16.2$ 、 $10 / \cos 38^\circ = 12.7$ となる。よって、それぞれから左右 10° （水平画角 20° の半分）ずつ開くと、台形の上端の長さは $16.2 \times \tan 10^\circ \times 2 = 5.7$ 、下端の長さは $12.7 \times \tan 10^\circ \times 2 = 4.5$ となる。

【 0 0 4 5 】

よって、聴取範囲は、 $Z=0$ の投影面において、（12.15、12.8）、（17.、

50

85、12.8)、(12.75、7.8)、(17.25、7.8)の4点で囲まれた領域になる。ここで、投影面に係る台形の上端辺の両頂点のx座標は、 $15 - 5.7 / 2 = 12.15$ 、 $15 + 5.7 / 2 = 17.85$ となる。投影面に係る台形の下端辺の両頂点のx座標は、 $15 - 4.5 / 2 = 12.75$ 、 $15 + 4.5 / 2 = 17.25$ となる。このように計算された聴取範囲は、聴取範囲決定部2の内部RAMに格納されている聴取情報に格納される。

【0046】

次に、S204では、S203で決定した聴取範囲において、画角において上方向に対応する投影面(水平面)上の方向を聴取方向とする。先ほどの例の場合は、 0° (すなわち、Y軸正方向)となる。処理を終えるとS208へ進む。

10

【0047】

一方、S205では、S108で被写体位置検知部9によって検知された被写体位置座標を取得する。次に、S206では、取得した被写体位置を囲む範囲を計算し、聴取範囲とする。例えば、S205で取得した被写体位置の数が3つで、その水平面座標が(2, 2)、(6, 3)、(5, 6)とすると、これを囲む聴取範囲を(1, 1)、(1, 7)、(7, 1)、(7, 7)の4点で囲まれた領域として設定する。つまり、ここでの例では、全ての被写体位置におけるX座標の最小値をXmin、X座標の最大値をXmax、Y座標の最小値をYmin、Y座標の最大値をYmaxとすると、聴取範囲は次のようになる。すなわち、(Xmin - 1, Ymin - 1)、(Xmin - 1, Ymax + 1)、(Xmax + 1, Ymin - 1)、(Xmax + 1, Ymax + 1)を各頂点座標とする四角形を聴取範囲とする。もっとも、聴取範囲は、取得した被写体位置を全て包含する必要最小限の大きさを有する領域であれば、四角形でなくてもよい。

20

【0048】

次に、S207では、S101で送信された視点情報の視点方向を、そのまま聴取方向として聴取点情報に格納する。このようにすることで、水平視点の場合は、視点映像における視点方向と再生信号における音の方向が合致するようになる。処理を終えるとS208へ進む。

【0049】

S208は、S203やS206で決定した聴取範囲の中心点を聴取点として聴取点情報に格納する処理である。本実施形態の例では、聴取範囲が四角形であるため、対角線が交差する点を算出し、これを聴取点として内部RAMの聴取点情報に格納する。なお、聴取点の座標は、聴取範囲を定める四角形の各頂点の平均座標としてもよい。次に、S209では、内部RAMに格納されている聴取点情報を、收音点選択部3へ出力し、聴取範囲決定処理を終えてリターンする。

30

【0050】

図6は、本実施形態における視点と聴取範囲、聴取点、聴取方向の関係を示す模式図である。図6中、聴取範囲を点線の矩形で示す。また、聴取点を八角の星型で示し、聴取方向を黒い矢印で示す。図6(a)は俯瞰視点の場合を示しており、画角を水平面に投影した範囲を聴取範囲とし、その対角線の交点を聴取点、画角の上方向に対応するY軸正方向を聴取方向とする。一方、図6(b)は水平視点の場合を示しており、視点映像における被写体の位置を検知して、その位置を囲むように聴取範囲を設定する。聴取点は聴取範囲の対角線の交点とし、視点方向(Y軸正方向)を聴取方向とする。

40

【0051】

以上説明したように、本実施形態の聴取範囲決定処理では、視点情報及び被写体の位置から、任意視点映像に応じた聴取範囲、聴取点、聴取位置を自動的に決定する。すなわち、視点位置及び視線の方向に応じて、視点映像生成部8が生成した画像に対応する音響信号を生成する基準となる聴取点を決定し、この聴取点に応じた音響信号を複数の收音信号に基づいて生成する。このように、視点の位置、視線の方向、画角、被写体の位置等に応じて聴取点、聴取範囲、聴取の方向を自動的に決定して、任意視点の音響を自動的に生成することで、視点の変化に応じた音響の変化を忠実に表現可能である。なお、本実施形態

50

では、画角の投影範囲や被写体の位置検出により聴取範囲及び聴取位置を決定しているが、単純に任意視点映像の中央を常に聴取点として扱うようにしてもよい。

【 0 0 5 2 】

また、本実施形態では、視線の仰角が予め定められた負の値を下回り、俯瞰視点であると判定されたときは、視点から見た画角に対応する撮影対象における範囲を聴取範囲として決定する。また、視線の仰角が予め定められた負の値を下回らず、水平視点であると判定されたときは、撮影対象における被写体の位置を包含する範囲を聴取範囲として決定する。このように、視線の仰角に応じて聴取範囲の決定方式を分岐させることで、視線の仰角に応じた臨場感ある音響を再現することが可能である。

【 0 0 5 3 】

また、本実施形態では、視線の方向に基づいて、聴取点における聴取の方向を示す聴取方向を決定し、音響信号生成の際には、この聴取方向に応じた音響信号を生成する。具体的には、俯瞰視点の場合は画角の上方向を聴取方向とし、水平視点の場合は視点方向を聴取方向として決定する。このため、視線の方向に対応する音響信号を生成することができる。その他、本発明の主旨を逸脱しない範囲で実施することが可能である。

【 0 0 5 4 】

(被写体位置検知処理)

図 7 は、本実施形態における S 1 0 8 の被写体検知処理の詳細な処理手順を示すフローチャートである。なお、本フローチャートにおける処理は全て被写体位置検知部 9 によって実行される。

【 0 0 5 5 】

まず、S 3 0 1 では、被写体位置検知部の内部 R A M に一次保存されているデータを全て初期化する。次に、S 3 0 2 では、S 1 0 7 において視点映像生成部 8 によって生成された視点映像を解析し、視点映像において焦点が合っている被写体を検知して、全て抽出する。例えば、視点映像をコントラスト画像、エッジ画像などに変換することにより、輪郭が明確な被写体、つまり、焦点が合っている被写体を検出する。抽出する被写体は選手などの人物に限らず、例えば車やバイク等の物体であってもよい。また、数は一つでも複数でもよい。S 3 0 2 では、このような焦点が合っている被写体を全て抽出し、抽出した各画像の特徴を被写体情報として内部 R A M に一次記憶する。

【 0 0 5 6 】

次に、S 3 0 3 から S 3 0 6 までは、S 3 0 2 で抽出した個々の被写体情報に対するループ処理を行う。まず、S 3 0 3 でループ処理を開始する。S 3 0 4 では、任意視点映像を生成するために用いた複数のカメラ画像のうち、処理対象となっている被写体が写っている画像を撮影したカメラ画像を複数特定して、そのカメラ位置座標と被写体方向を求める。次に、S 3 0 5 では、S 3 0 4 で求めた複数のカメラ位置座標と被写体方向から、処理対象となっている被写体の位置座標を三角測量法で算出する。算出した座標は被写体位置検出部の内部 R A M に被写体位置座標として保存される。

【 0 0 5 7 】

S 3 0 6 で全ての被写体情報に対して処理が終了したかどうかを判定し、終了した場合はループを抜け、被写体位置検知処理を終了してリターンする。なお、内部 R A M に格納されている被写体位置座標は、聴取範囲決定部 2 からの要求に応じて適宜、聴取範囲決定部 2 へ出力される。

【 0 0 5 8 】

このように、本実施形態では、視点映像生成部により生成された任意視点画像を解析して、その画像に写り込んでいる被写体の位置を検出する。そのため、被写体の位置を検知するための専用のセンサ等を設けることなく、被写体の位置を検出して聴取範囲を適切に決定することができる。なお、位置センサ等を用いて被写体の位置を検出しても構わない。

【 0 0 5 9 】

(收音点選択処理)

図 8 は、本実施形態における S 1 0 4 の收音点選択処理の詳細な処理手順を示すフローチャートである。なお、本フローチャートにおける処理は全て收音点選択部 3 によって実行される。

【 0 0 6 0 】

まず、S 4 0 1 では、收音点選択部 3 の内部 R A M に保存されている選択收音点情報リストを初期化する。選択收音点情報リストとは、選択された收音点の情報を記憶する領域である。図 4 (e) に、本実施形態における選択收音点情報のデータ構造の一例を示す。図 4 (e) に示すように、選択收音点情報は、收音点 I D 、対応音源配置方向 I D 、聴取点から見た方向を含む。このうち、收音点 I D は、收音点を識別するための I D (識別情報) であり、図 4 (b) を用いて説明した收音点情報に格納されている同名のデータと共通のデータが用いられる。対応音源配置方向 I D は、この收音点がカバーする音源配置方向を示す番号 (識別情報) である。なお音源配置方向については後述する。聴取点から見た方向は、聴取点から見たこの收音点の方向を、聴取方向を基準として計算したものである。

10

【 0 0 6 1 】

次に、S 4 0 2 では、S 1 0 3 で決定された聴取点情報に基づいて、再生時に周囲に配置する音源の大まかな配置方向を決定する。本実施形態では、音源の配置方向を、聴取方向を 0 ° として起点とし、水平面を 4 5 ° おきに一周した 8 方向を音源配置方向として設定する。

【 0 0 6 2 】

20

次に、S 4 0 3 から S 4 1 0 までは、S 4 0 2 で設定した各音源配置方向に対するループ処理を行う。S 4 0 3 でループ処理を開始する。

【 0 0 6 3 】

S 4 0 4 では、聴取点から見て、対象音源配置方向の $\pm 22.5^\circ$ の角度の範囲内の領域を收音点探索範囲とする。この音源配置方向に対応する收音点は、この探索範囲から探索される。

【 0 0 6 4 】

次に、S 4 0 5 では、S 4 0 4 で設定した探索範囲内に收音点があるかどうかを判定する。探索範囲内に收音点がある場合 (S 4 0 5 で Y E S) は、処理は S 4 0 6 へ進む。そうでない場合 (S 4 0 5 で N O) は、対象となっている音源配置方向に收音点を割り当てないことに決定し、S 4 1 0 へ進む。

30

【 0 0 6 5 】

S 4 0 6 は、探索範囲の角度内で、かつ、聴取範囲外に收音点があるかどうかを判定する処理である。この判定の結果、收音点がある場合 (S 4 0 6 で Y E S) は S 4 0 7 へ処理は進む。そうでない場合 (S 4 0 6 で N O) は、S 4 0 8 へ処理は進む。

【 0 0 6 6 】

S 4 0 7 は、探索範囲内かつ聴取範囲外で、聴取点により近い收音点をこの音源配置方向の音源として選択する処理である。收音点を選択されたら、收音点選択部 3 の内部 R A M に保存されている選択收音点情報リストに新規要素を追加して、選択された收音点の收音点 I D と、対応音源配置方向 I D を格納する。処理を終えると、S 4 0 9 へ進む。

40

【 0 0 6 7 】

一方、S 4 0 8 では、探索範囲内かつ、聴取範囲内で、聴取点に最も遠い收音点をこの音源配置方向の音源として選択する。この場合も選択した收音点の情報を格納する要素を作成し、選択收音点情報リストに追加する。処理を終えると、S 4 0 9 へ進む。

【 0 0 6 8 】

S 4 0 9 は、S 4 0 7 又は S 4 0 8 で選択した收音点の、聴取方向から見た方向を計算して、前処理で新規追加した選択收音点情報に格納する処理である。例えば、聴取点の座標が (1 , 1) かつ收音点座標が (2 , 1 + 3) で、聴取方向が 6 0 ° の場合、聴取点から收音点を見た方向は - 3 0 ° となるので、この角度を選択收音点情報に格納する。処理を終えると、S 4 1 0 へ進む。

50

【 0 0 6 9 】

S 4 1 0 では、全ての音源配置方向に対する処理が済んだかどうかを判定し、全ての処理が終了した場合は、ループを終了する。そして、S 4 1 1 へ進む。

【 0 0 7 0 】

S 4 1 1 は、聴取範囲内の收音点のうち、再生に用いる收音点を選択するとともに、配置方向を決定する処理である。この処理（聴取範囲内收音点選択処理）の詳細は図 9 を用いて後述する。

【 0 0 7 1 】

S 4 1 2 では、ここまでの処理で作成した選択收音点情報リストを音響信号生成部 4 へ出力する。そして、收音点選択処理を終了してリターンする。

10

【 0 0 7 2 】

以上のように、本実施形態では、視点位置及び視線の方向に応じて決定された聴取範囲に基づき、複数の收音点の中から音響信号を生成するために用いる收音点を選択し、その收音点において收音された收音信号を用いて、音響信号を生成する。このように、再生信号を生成するために必要な收音点を S 1 0 5 において再生信号を生成する処理前に選択することによって、再生信号生成に必要な処理を削減できる。さらに、視点映像において注目している範囲に見合う收音点を自動的に選択することにより、視点映像に見合うより没入感の高い再生音響信号を生成することができる。したがって、任意視点映像に見合う臨場感のある音響を小さい処理量で生成することが可能である。

【 0 0 7 3 】

20

また、本実施形態では、視点位置及び視線の方向に応じて決定された聴取点から見た撮影対象を複数の領域に区分し、その複数の領域の各々から、聴取範囲に基づき收音点を選択する。このため、聴取点を基準に、聴取点の周囲の收音点をまんべんなく選択することができる。

【 0 0 7 4 】

また、聴取点から見た領域における聴取範囲の内部に收音点が存在する場合は、その聴取範囲の内部に存在する收音点のうち、聴取点から最も離れた收音点を選択する。一方、聴取点から見た領域における聴取範囲の内部に收音点が存在しない場合は、その領域に存在する收音点のうち、聴取点に最も近い收音点を選択する。このため、聴取範囲の拡がりに応じた收音点を適切に選択して、再生信号を適切に生成することが可能となる。

30

【 0 0 7 5 】

なお、本実施形態では再生音響信号において周囲に配置する音源を 8 方向として選択した例を説明したが、音源の方向は 8 方向に限定されず、これよりも多くても少なくともよい。また、本実施形態では音源配置方向として聴取点の周囲を等分した場合の例を説明したが、等分ではなく、例えば、音響再生環境におけるチャンネルの方向に応じた方向に分割して選択してもよい。

【 0 0 7 6 】

（聴取範囲内收音点選択処理）

図 9 は、本実施形態における S 4 1 1 の聴取範囲内收音点選択処理の詳細な処理手順を示すフローチャートである。なお、本フローチャートにおける処理も全て收音点選択部 3 によって実行される。

40

【 0 0 7 7 】

まず、S 5 0 1 では、聴取範囲に含まれている收音点をリストアップして、收音点選択部 3 の内部 R A M に一次保存する。次に、S 5 0 2 から S 5 0 5 までは、S 5 0 1 でリストアップした個々の收音点に対するループ処理を行う。S 5 0 2 でループ処理を開始する。

【 0 0 7 8 】

S 5 0 3 では、処理対象の收音点が内部 R A M に保存されている選択收音点情報リストに含まれているかどうかを判定する。対象の收音点が、選択收音点リストに含まれていない場合（S 5 0 3 で N O ）は、S 5 0 4 へ進む。含まれている場合（S 5 0 3 で Y E S ）

50

は、既に再生に用いる收音点として選択されているため、S 5 0 5 へ進んでループ処理を終了する。

【 0 0 7 9 】

S 5 0 4 では、内部 R A M に保存されている選択收音点リストに新規要素を追加して、この收音点の收音点 I D と、聴取点から見た方向として 0 ° を格納する。これにより、本実施形態では、後の再生信号生成処理において、聴取範囲内の收音点で收音した信号が、聴取者の正面に定位するように再生されることになる。処理を終えると、S 5 0 5 へ進む。

【 0 0 8 0 】

S 5 0 5 では、S 5 0 1 でリストアップされた收音点全てに対して処理が済んだかどうかを確認し、全て終了した場合はループ処理を抜け、聴取範囲内收音点選択処理を終了してリターンする。

10

【 0 0 8 1 】

このように、本実施形態では、聴取範囲の内部に存在する全ての收音点を選択して、再生信号を生成するため、聴取範囲に応じた臨場感ある音響を生成することができる。また、聴取点から見た複数の領域の各々において、聴取点に最も近い收音点を選択するため、聴取点と收音点との位置関係に基づく臨場感ある音響信号を生成することが可能となる。

【 0 0 8 2 】

(再生信号生成処理)

図 1 0 は、本実施形態における S 1 0 5 の再生信号生成処理の詳細な処理手順を示すフローチャートである。再生信号生成処理では、複数の收音信号に基づいて、聴取点及び聴取範囲に応じた音響信号を生成する。なお、本フローチャートにおける処理は全て音響信号生成部 4 によって実行される。

20

【 0 0 8 3 】

S 6 0 1 は、音響信号生成部 4 の内部にある出力バッファを初期化してクリアする処理である。出力バッファは再生音響信号の出力チャンネル毎のバッファになっており、生成した音響信号を出力チャンネル毎に蓄積する。処理を終えると、S 6 0 2 へ進む。

【 0 0 8 4 】

S 6 0 2 は、これから生成する音響信号を再生する環境を判定する処理である。前述のように、本実施形態の例では音響信号を再生する環境として、ステレオ再生環境、サラウンド再生環境、及び、ヘッドフォン再生環境が設けられており、音響再生フォーマットもこれらの環境のいずれかに合わせて設定されている。

30

【 0 0 8 5 】

ステレオ再生環境である場合は、S 6 0 3 へ進む。サラウンド再生環境である場合は、S 6 0 4 へ進む。ヘッドフォン再生環境である場合は、S 6 0 5 へ進む。

【 0 0 8 6 】

S 6 0 3 は、S 1 0 4 で選択した收音点の收音信号を用いてステレオ再生信号を生成する処理である。この処理の詳細は図 1 1 を用いて後述する。処理を終えると、S 6 0 6 へ進む。

【 0 0 8 7 】

40

S 6 0 4 は、S 1 0 4 で選択した收音点の收音信号を用いてサラウンド再生信号を生成する処理である。この処理の詳細は図 1 2 を用いて後述する。処理を終えると、S 6 0 6 へ進む。

【 0 0 8 8 】

S 6 0 5 は、S 1 0 4 で選択した收音点の收音信号を用いてヘッドフォン再生信号を生成する処理である。この処理の詳細は図 1 3 を用いて後述する。処理を終えると、S 6 0 6 へ進む。

【 0 0 8 9 】

S 6 0 6 は、直前の処理で生成した再生音響信号を、音響再生部 1 1 や M U X 1 5 へ出力する処理である。処理を終えると、再生信号生成処理を終了し、リターンする。なお、

50

図10のフローチャートでは、どれか一つの再生フォーマットを選択して生成する例を示しているが、これらのフォーマットを逐次的に全て生成するようにしてもよい。

【0090】

上記のように、本実施形態では、視線の方向に基づいて、聴取点における聴取の方向を示す聴取方向をさらに決定し、S104において選択された収音点において収音された収音信号を用いて、聴取方向の正面から聞こえる音響信号を生成する。このため、任意視点画像に対応する音響を、その方向を考慮して再現よく表現することができる。

【0091】

(ステレオ再生信号生成処理)

図11は、本実施形態におけるS603のステレオ再生信号生成処理の詳細な処理手順を示すフローチャートである。なお、本フローチャートにおける処理も全て音響信号生成部4によって実行される。

【0092】

S701からS709までは、S104において収音点選択部3から出力された選択収音点情報リストに格納されている個々の選択収音点情報に対してループ処理を行う。

【0093】

S701でループ処理が開始される。S702では、処理対象の選択収音点情報の聴取点から見た方向が、 -90° から 90° の範囲内かどうか、つまり、対象の収音点が聴取点から見て真横から前方に位置するかどうかを判定する。この処理の結果、範囲内にない、つまり、収音点が後方にある場合(S702でNO)は、S703へ処理は進む。そうでない場合、つまり、前方にある場合(S702でYES)は、S707へ処理は進む。

【0094】

S703では、対象となっている選択収音点情報に格納されている収音信号を逆位相化する。これにより、聴取者が後方に音像を感じることはないが、後方からの音が通常の音と違って聞こえる演出を行うことができる。次に、S704では、聴取点から見た収音点の方向が正か否かを判定する。正でない場合、すなわち、聴取点から見て収音点が左方にある場合(S704でNO)は、S705へ進む。正の場合、すなわち、聴取点から見て収音点が右方にある場合(S704でYES)は、S706へ進む。

【0095】

S705では、聴取点から見た方向に 180° を加えて符号を反転する。一方、S706では、聴取点から見た方向から 180° を減じて符号を反転する。S705とS706の処理により、聴取点を中心とする円の後ろ半分を前に折り返すようにして、後方の方向を前方の方向に変換することができる。処理を終えると、S707へ進む。

【0096】

S707では、 -90° から 90° の範囲にある聴取点から見た方向に対するステレオパンニング計算を行い、得られたL, Rチャンネルの振幅分配率に従って収音信号をL, Rチャンネルに分配する。一般に、標準のステレオ再生環境では $\pm 30^{\circ}$ に左右スピーカーを配置するため、 $\pm 90^{\circ}$ の範囲にある聴取点から見た方向を、 $\pm 30^{\circ}$ の範囲に線形に投射することにより、ステレオパンニング計算を行う。本実施形態では、このパンニング計算を、聴取点から見た方向を θ とすると、サイン則を用いて以下のように行う。

$$wL = (\sin 30^{\circ} - \sin(\theta * 30 / 90)) / 2\sin 30^{\circ} = 1/2 - \sin(\theta / 3)$$

$$wR = (\sin 30^{\circ} + \sin(\theta * 30 / 90)) / 2\sin 30^{\circ} = 1/2 + \sin(\theta / 3) \quad (1)$$

ただし、wLは左チャンネルに対する振幅分配率、wRは右チャンネルに対する振幅分配率である。

【0097】

次に、S708では、S707で分配したチャンネル信号を、各チャンネルの出力バッファに各々加算する。S709では、選択収音点情報リストに含まれている全ての選択収音点情報の処理が終了したかを確認する。全て終了した場合は、ループ処理を抜けてステレオ再生信号生成処理を終了し、リターンする。

10

20

30

40

50

【 0 0 9 8 】

(サラウンド再生信号生成処理)

図 1 2 は、本実施形態における S 6 0 4 のサラウンド再生信号生成処理の詳細な処理手順を示すフローチャートである。なお、本フローチャートにおける処理も全て音響信号生成部 4 によって実行される。

【 0 0 9 9 】

S 8 0 1 から S 8 0 7 までは、選択收音点情報リストに含まれている各選択收音点情報に対するループ処理を行う。

【 0 1 0 0 】

S 8 0 1 でループ処理を開始する。S 8 0 2 では、聴取点から見た收音点の方向が既定のチャンネル配置方向かどうかを判定する。例えば、再生環境が 5 . 1 チャンネルサラウンド再生環境だとすると、既定のチャンネル配置角度は 0 °、± 3 0 °、± 1 1 0 ° ~ 1 3 0 ° になる。対象の選択收音点情報により示される收音点の聴取点から見た方向がこの範囲の角度に該当する場合は、チャンネル配置方向であると判定する。この場合 (S 8 0 2 で Y E S) は S 8 0 6 へ処理は進む。そうでない場合 (S 8 0 2 で N O) は、S 8 0 3 へ処理は進む。

【 0 1 0 1 】

S 8 0 3 は、聴取点から見た收音点の方向の角度を挟み込む方向にある二つのチャンネルを選択する処理である。例えば、聴取点から見た收音点の方向が 5 0 ° であるとして、それを挟み込むチャンネルとして、3 0 ° の R チャンネルと、1 2 0 ° の S R チャンネルが選択される。

【 0 1 0 2 】

次に、S 8 0 4 では、S 8 0 3 で選択したチャンネル間で振幅パンニング計算を行い、二つのチャンネルにこの收音点の收音信号を分配する。本実施形態ではサイン則によって振幅パンニング計算を行う。先ほどの例で説明すると、R と S R の方向の中心方向は 7 5 ° であり、中心方向と各チャンネルの方向との開き角は 4 5 ° となる。また、チャンネル間の中心方向から見ると、5 0 ° は $5 0 ^{\circ} - 7 5 ^{\circ} = - 2 5 ^{\circ}$ となる。R、S R チャンネルへの各分配率 w_R 、 w_{SR} はサイン則よりそれぞれ次式で求められる。

$$w_R = (\sin 4 5 ^{\circ} - \sin (- 2 5 ^{\circ})) / 2 \sin 4 5 ^{\circ} \quad 0 . 6 4 9$$

$$w_{SR} = (\sin 4 5 ^{\circ} + \sin (- 2 5 ^{\circ})) / 2 \sin 4 5 ^{\circ} \quad 0 . 3 5 1 \quad (2)$$

次に、S 8 0 5 では、S 8 0 4 で分配した各チャンネルの信号を、チャンネル毎に出力バッファに加算する。一方、S 8 0 6 では、S 8 0 2 で判定された方向が同じチャンネルの出力バッファに、收音信号をそのまま加算する。

【 0 1 0 3 】

S 8 0 7 では、選択收音点情報リストに含まれる全ての選択收音点情報に対する処理が終了したかどうかを確認する。全ての処理が終了した場合は、ループ処理を終了し、S 8 0 8 へ進む。

【 0 1 0 4 】

S 8 0 8 では、各チャンネルの出力バッファに蓄積されている音響信号に対して、ローパスフィルタ (L P F) を掛けて加算することにより、L F E (L o w F r e q u e n c y E l e m e n t) 信号を生成する。L F E 信号は低域信号であり、通常は 8 0 H z 以下の信号をローパスフィルタで取り出すようにする。この信号は、サラウンドスピーカセットに含まれるサブウーファによって再生される。生成された L F E 信号は、L F E チャンネル用の出力バッファに蓄積される。処理を終えると、サラウンド再生信号生成処理を終了し、リターンする。

【 0 1 0 5 】

(ヘッドフォン再生信号生成処理)

図 1 3 は、本実施形態における S 6 0 5 のヘッドフォン再生信号生成処理の詳細な処理手順を示すフローチャートである。なお、本フローチャートにおける処理も全て 4 の音響信号生成部によって実行される。

【0106】

S901からS904までは、選択收音点情報リストに含まれている各選択收音点情報に対するループ処理を行う。

【0107】

S901でループ処理を開始する。S902では、聴取点から見た方向のH R I Rを收音信号に畳み込むことにより、両耳信号を計算する。H R I Rとは頭部伝達関数(Head Related Impulse Response)である。H R I Rは、音源方向によって変化する人間の頭部や耳介による音の回り込みを測定して両耳分のインパルス応答としたものである。收音信号に対してこれを畳み込むことにより、ヘッドフォンで聴いた場合に聴取点から見た方向に收音信号が定位する立体音響信号を作成することができる。なお、本実施形態では、方向毎のH R I Rを格納したデータベースが音響信号生成部4の内部ROMに格納されており、任意の方向を入力して検索することにより、両耳分のH R I Rを読み出して用いることができる。

10

【0108】

次に、S903では、S902で生成した両耳信号を、L、Rの出力チャンネル毎に出力バッファに加算する。

【0109】

S904では、選択收音点情報リストに含まれる全ての選択收音点情報に対する処理が終了したかどうかを確認する。全ての処理が終了した場合は、ループを抜けて、ヘッドフォン再生信号生成処理を終了し、リターンする。

20

【0110】

なお、本実施形態では、ステレオ再生処理においても全方向の收音点の收音信号を用いて再生信号を生成しているが、例えば、ステレオならば前方にある收音点の收音信号のみを用いて再生信号を生成するようにしてもよい。

【0111】

また、本実施形態では、收音点の位置に收音用のマイクロホンを設置しているが、收音の実現手法はこのような態様に限られない。例えば、遠方の微小な音も收音できるマイクロホンを複数用いて、收音点に対応する位置を狙って收音、処理することにより、ピンポイントで遠距離から狙った收音点の音を收音してもよい。

【0112】

30

また、本実施形態では收音信号や撮影映像信号をすぐに処理して任意視点映像及びそれに見合う音響信号を生成、再生しているが、收音信号や撮影映像信号を一旦記憶装置に記憶しておき、後で処理するようにしてもよい。

【0113】

以上説明したように、本実施形態にかかる構成により、視点情報から任意視点映像に応じた聴取範囲、聴取点、聴取位置を自動的に決定することで、任意視点の動きに応じて変化する臨場感のある音場再生を実現することができる。

【0114】

また、視点情報から任意視点映像に応じた聴取範囲を決定し、聴取範囲に応じた最小限の收音点を選択して再生音場に適宜配置することで、処理量を抑制しつつ映像に見合う臨場感のある音場再生を実現できる。すなわち、聴取範囲に基づき音響信号を生成するために用いる收音点を選択して、音響信号を生成することで、必要最小限の收音信号を選択して任意視点映像に対応する臨場感ある音響信号を自動的に生成することができる。

40

【0115】

<<その他の実施形態>>

実施形態1では、聴取範囲内の收音点による收音信号を全て用いて再生信号を生成しているが、聴取範囲内の收音信号のうち重要な收音信号を選択して用いることもできる。ここでは、重要な收音信号の一例として、人の声(以下、「音声」という)を含む收音信号を選択する例を説明する。以下、この場合の実施形態について説明する。

【0116】

50

本実施形態と実施形態 1 の差分は図 8 における S 4 1 1 の聴取範囲内收音点選択処理であり、他は同一であるため説明を省略し、実施形態 1 と異なる点を中心に簡潔に説明する。

【 0 1 1 7 】

図 1 4 は本実施形態における聴取範囲内收音点選択処理の詳細な処理手順を示すフローチャートである。S 1 0 0 1 から S 1 0 0 3 までの処理は図 9 の S 5 0 1 から S 5 0 3 までの処理と同一であるため説明を省略する。

【 0 1 1 8 】

S 1 0 0 4 は、対象となる收音点情報の收音信号を解析し、音声が含まれているかどうかを判定する処理である。收音信号に対してピッチ検出処理やフォルマント検出処理等を行うことにより、收音信号に音声が含まれているかどうかを判定する。判定の結果、音声が含まれている場合は S 1 0 0 6 に進む。そうでない場合は、S 1 0 0 5 へ進む。

10

【 0 1 1 9 】

S 1 0 0 5 は、收音信号の平均振幅が予め定めた既定値（閾値）を超えているかどうかを判定する処理である。判定の結果、閾値を超えている場合は S 1 0 0 6 に進む。そうでない場合は、S 1 0 0 7 へ進む。

【 0 1 2 0 】

S 1 0 0 6 と S 1 0 0 7 は図 9 における S 5 0 4 と S 5 0 5 と同一の処理であるため説明を省略する。

【 0 1 2 1 】

20

以上説明した処理制御を行うことで、聴取範囲内の收音点で收音した信号のうち、重要な情報が含まれていると推定される声が混じっている信号や、音源の発生源に近いと推定できる平均振幅の大きな信号のみを選択して再生することができる。よって、再生信号生成処理にかかる処理量をさらに削減しつつ、重要な音のみを選択して再生することができる。

【 0 1 2 2 】

また、上記構成では、聴取範囲内の收音点を選択する場合に、人の音声等の音響的に重要な收音点を検知する例を説明した。さらに、視点映像を解析してボールを蹴ったりスクラムを組むなどの所定のイベントが生じている場所を特定し、その場所に最も近い收音点を選択するようにしてもよい。イベントの発生は、ユーザの指定やセンサの使用等により検知することができる。

30

【 0 1 2 3 】

また、実施形態 1 では、俯瞰視点の場合に画面の上方向の音を前方に配置しているが、水平面の他に、上層にもチャンネルを配置する再生フォーマットへ出力する場合には、上層チャンネルに配置するようにしてもよい。このように、S 1 0 4 において選択された收音点の視点から見た位置関係に基づいて、收音点において收音された收音信号を合成して、音響信号を生成することで、收音点の配置に応じた臨場感ある音響を再現することができる。その他、本発明の主旨を逸脱しない範囲で実施することが可能である。

【 0 1 2 4 】

< その他の実施形態 >

40

本発明は、上述の実施形態の 1 以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける 1 つ以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1 以上の機能を実現する回路（例えば、ASIC）によっても実現可能である。

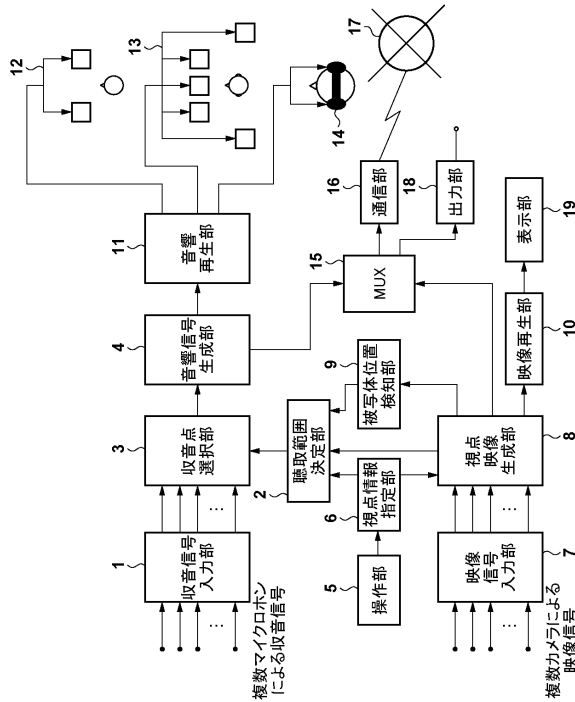
【 符号の説明 】

【 0 1 2 5 】

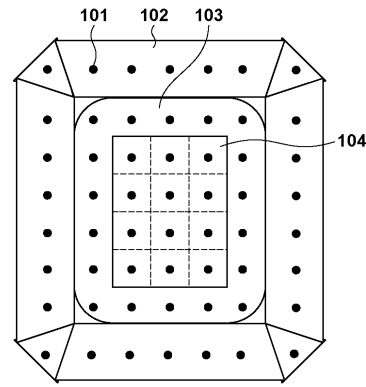
收音信号入力部：1、聴取範囲決定部：2、收音点選択部：3、音響信号生成部：4、操作部：5、視点情報指定部：6、映像信号入力部：7、視点映像生成部：8、被写体位置検知部：9、映像再生部：10、音響再生部：11、ステレオスピーカーセット：12、サラウンドスピーカーセット：13、ヘッドフォン：14、表示部：19

50

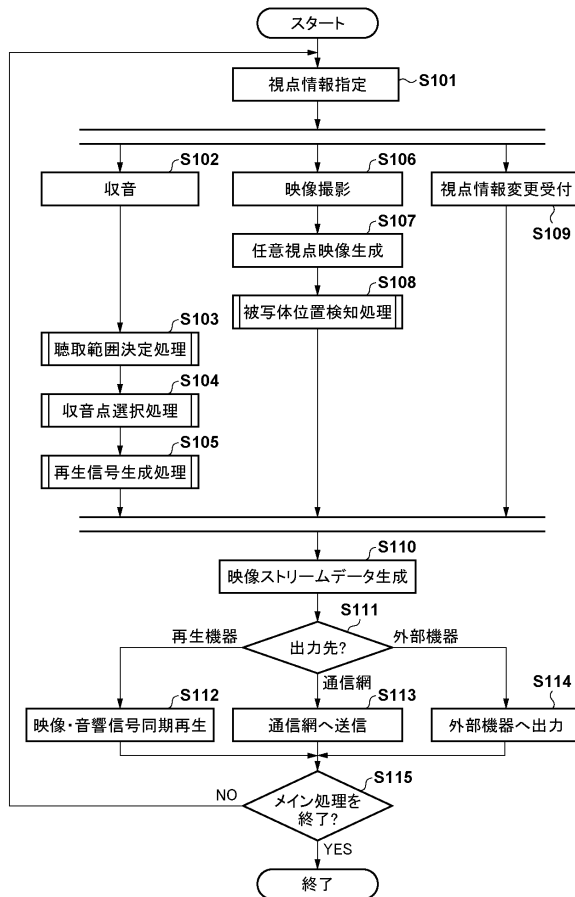
【図 1】



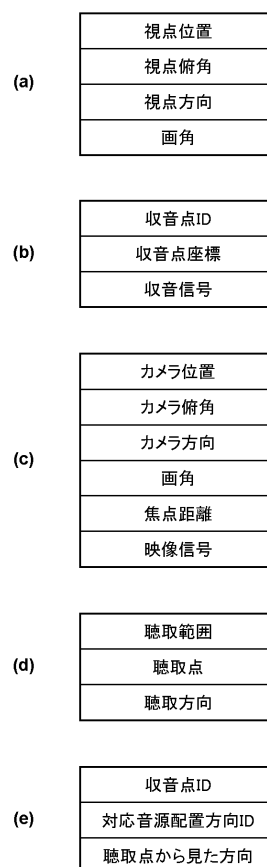
【図 2】



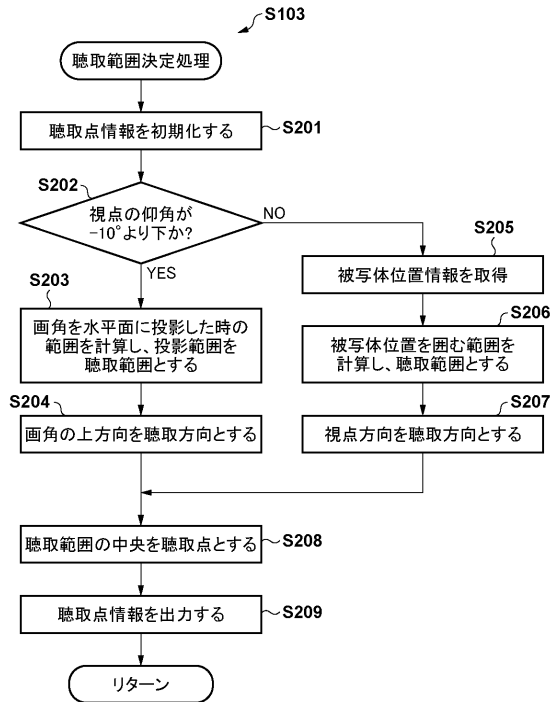
【図 3】



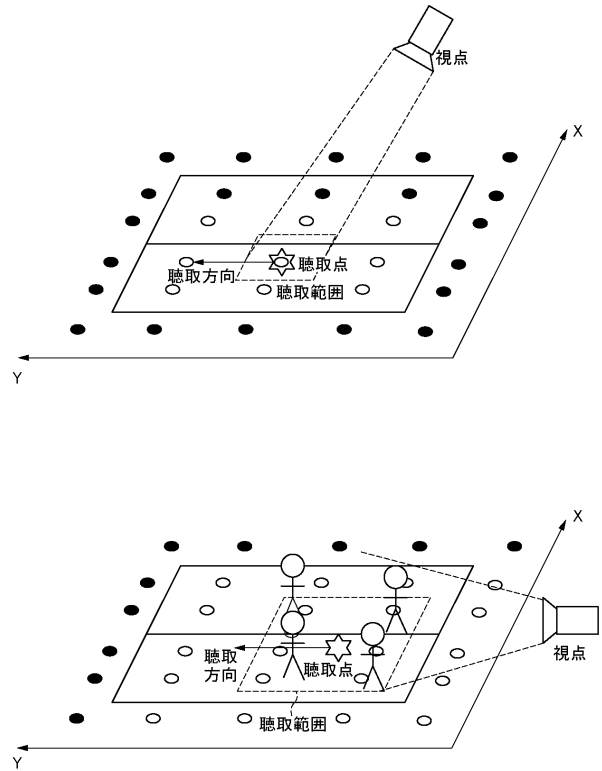
【図 4】



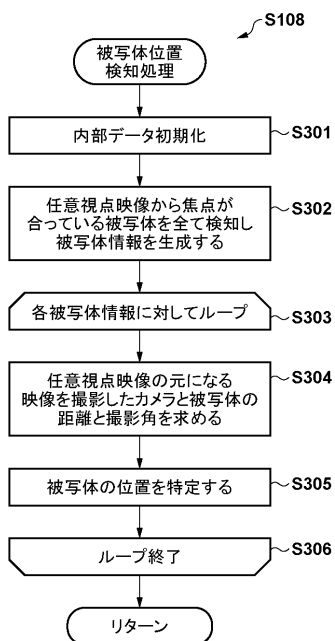
【図 5】



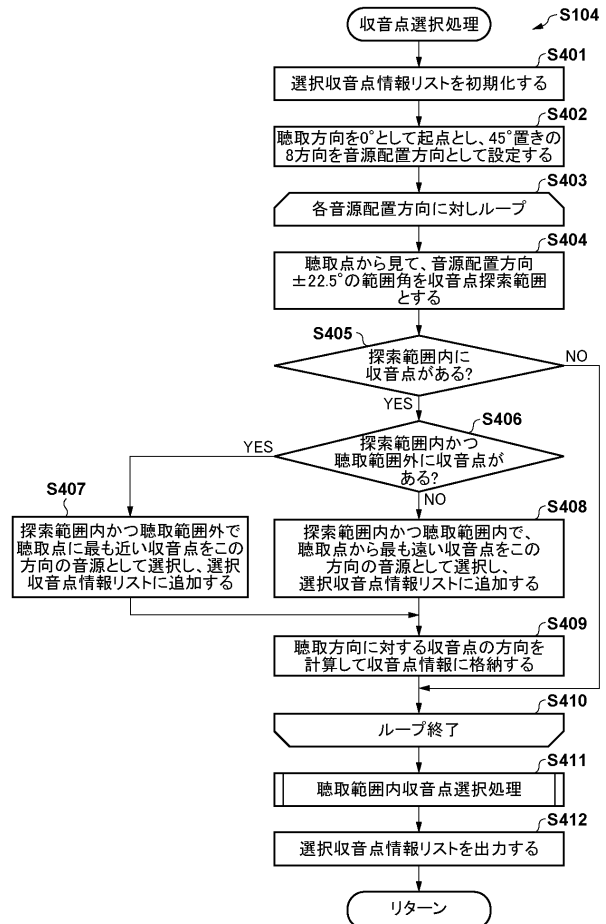
【図 6】



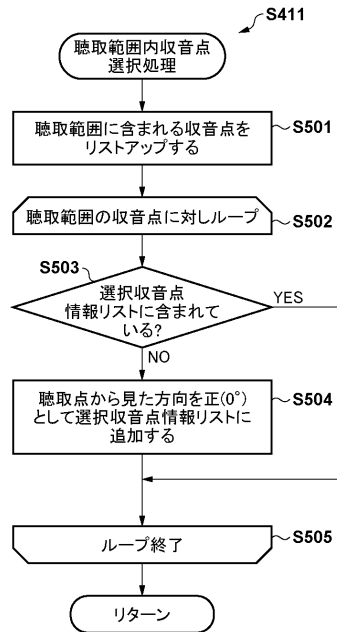
【図 7】



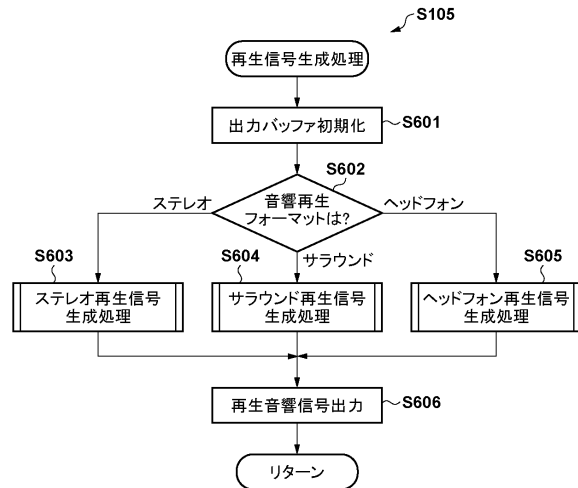
【図 8】



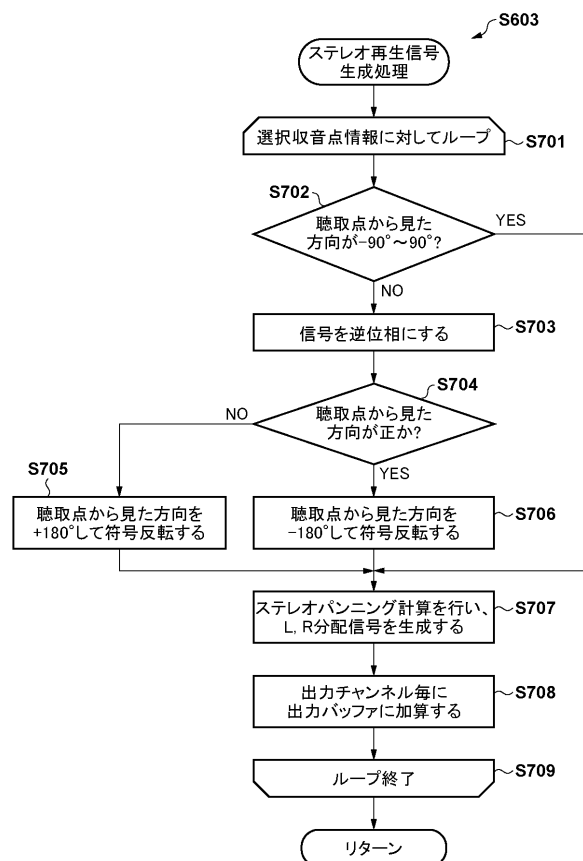
【図 9】



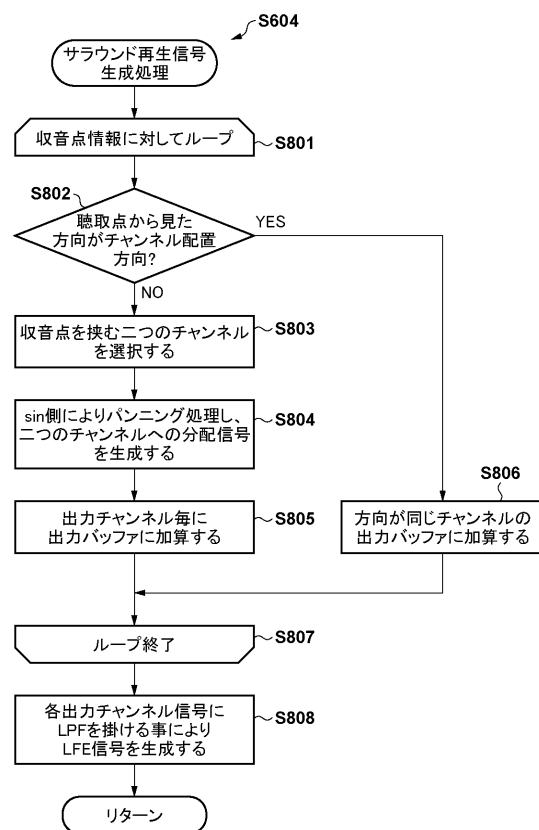
【図 10】



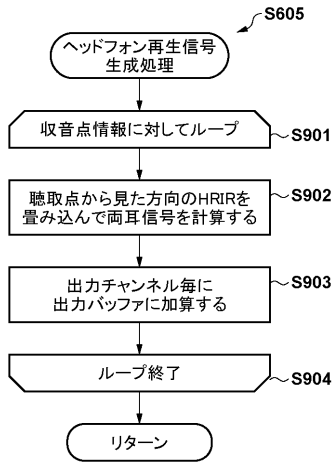
【図 11】



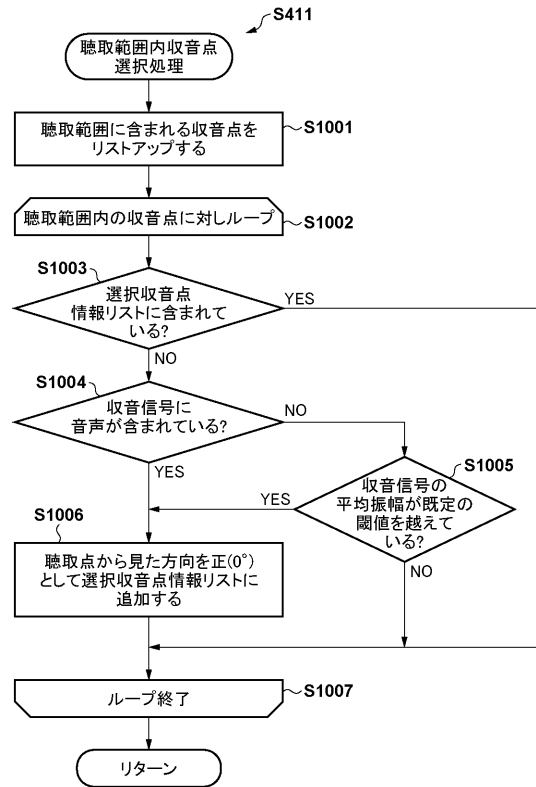
【図 12】



【図 13】



【図 14】



フロントページの続き

(72)発明者 船越 正伸

東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 益戸 宏

(56)参考文献 特開2005-223771(JP,A)

国際公開第96/021321(WO,A1)

特開2013-034107(JP,A)

堀内俊治 他3名,自由視聴点技術に基づくスマートフォン向けミュージックビデオアプリケーション,2012年映像情報メディア学会年次大会講演予稿集,2012年 8月

福島慶繁 他6名,多視点・多聴点データ取得システムを用いた自由視聴点映像生成,電子情報通信学会論文誌,2008年,Vol.J91-D No.8,pp.2039-2041

馬渡隆行 他3名,A-16-37 拡張現実感における音声注釈情報取得のための仮想マイクロフォンの提案,2004年電子情報通信学会総合大会講演論文集 基礎・境界,2004年 3月 8日,p.346

小宮山 他4名,VR技術による3次元音響再生システム Interactive 3D Sound Reproduction System by Virtual Acoustics,映像情報メディア学会技術報告,日本,(社)映像情報メディア学会 The Institute of Image Information and Television Engineers,2000年11月,Vol.24, No.71, pp.51-56

(58)調査した分野(Int.Cl.,DB名)

H04N 5/60

H04S 7/00

H04S 1/00

H04S 5/00

H04N 7/18

H04N 21/00

G06T 19/00