

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7558467号
(P7558467)

(45)発行日 令和6年9月30日(2024.9.30)

(24)登録日 令和6年9月19日(2024.9.19)

(51)国際特許分類

F I

H 0 4 S 7/00 (2006.01)

H 0 4 S 7/00 3 0 0

請求項の数 8 (全16頁)

(21)出願番号	特願2024-542112(P2024-542112)	(73)特許権者	000006013
(86)(22)出願日	令和4年9月28日(2022.9.28)		三菱電機株式会社
(86)国際出願番号	PCT/JP2022/036165		東京都千代田区丸の内二丁目7番3号
(87)国際公開番号	WO2024/069796	(74)代理人	100116964
(87)国際公開日	令和6年4月4日(2024.4.4)		弁理士 山形 洋一
審査請求日	令和6年7月12日(2024.7.12)	(74)代理人	100120477
早期審査対象出願			弁理士 佐藤 賢改
		(74)代理人	100135921
			弁理士 篠原 昌彦
		(74)代理人	100203677
			弁理士 山口 力
		(72)発明者	栗野 智治
			東京都千代田区丸の内二丁目7番3号
			三菱電機株式会社内
		(72)発明者	細谷 耕佑
			最終頁に続く

(54)【発明の名称】 音空間構築装置、音空間構築システム、プログラム及び音空間構築方法

(57)【特許請求の範囲】

【請求項1】

複数の音源からの音声を含む音声データを取得する音声取得部と、
前記音声データから、前記複数の音源の位置である複数の音源位置を判定する音源判定部と、
前記音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成する音声抽出部と、
前記複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、前記複数の音源に対応する複数の立体音を生成するフォーマット変換部と、
音声を聴く位置である聴覚位置を取得する位置取得部と、
前記聴覚位置と、前記複数の音源位置の各々との間の角度及び距離を算出する移動処理部と、
前記複数の立体音のそれぞれを、前記複数の音源位置のそれぞれに対応する角度及び距離で調整することで、前記聴覚位置における複数の立体音である複数の調整済立体音を生成する角度距離調整部と、
前記複数の調整済立体音を重畳する重畳部と、を備えること
を特徴とする音空間構築装置。

【請求項2】

前記音声抽出部は、前記音声データから、前記複数の音源に含まれる一つの音源からの音声を分離した残りのデータを、前記音声データから減算することで、前記複数の抽出音

声データの内の、前記一つの音源に対応する抽出音声データを生成すること
を特徴とする請求項 1 に記載の音空間構築装置。

【請求項 3】

前記音源判定部は、前記複数の音源が含まれる空間を撮像した画像を用いて、前記複数の音源位置を判定すること

を特徴とする請求項 1 又は 2 に記載の音空間構築装置。

【請求項 4】

前記音声データは、前記音空間構築装置に対してネットワークで接続された集音装置で捕捉された音声を示すデータであること

を特徴とする請求項 1 又は 2 に記載の音空間構築装置。

10

【請求項 5】

前記音声データに含まれている音声とは、捕捉された時及び場所の少なくとも何れか一方において異なる音声の音声データを、前記立体音響のフォーマットに変換することで生成された立体音である重畳用立体音を示す重畳用音声データを取得する別音声取得部と、
前記重畳用立体音から、前記聴覚位置における立体音である重畳用調整済立体音を生成する重畳用角度距離調整部と、をさらに備え、

前記重畳部は、前記複数の調整済立体音及び前記重畳用調整済立体音を重畳すること
を特徴とする請求項 1 又は 2 に記載の音空間構築装置。

【請求項 6】

音空間構築装置と、前記音空間構築装置に対してネットワークで接続され、複数の音源からの音声を含む音声データを生成する集音装置とを備える音空間構築システムであって、
前記音空間構築装置は、

前記集音装置と通信を行う通信部と、

前記通信部を介して、前記音声データを取得する音声取得部と、

前記音声データから、前記複数の音源の位置である複数の音源位置を判定する音源判定部と、

20

前記音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成する音声抽出部と、

前記複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、前記複数の音源に対応する複数の立体音を生成するフォーマット変換部と、

30

音声を聴く位置である聴覚位置を取得する位置取得部と、

前記聴覚位置と、前記複数の音源位置の各々との間の角度及び距離を算出する移動処理部と、

前記複数の立体音のそれぞれを、前記複数の音源位置のそれぞれに対応する角度及び距離で調整することで、前記聴覚位置における複数の立体音である複数の調整済立体音を生成する角度距離調整部と、

前記複数の調整済立体音を重畳する重畳部と、を備えること

を特徴とする音空間構築システム。

【請求項 7】

コンピュータを、

40

複数の音源からの音声を含む音声データを取得する音声取得部、

前記音声データから、前記複数の音源の位置である複数の音源位置を判定する音源判定部、

前記音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成する音声抽出部、

前記複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、前記複数の音源に対応する複数の立体音を生成するフォーマット変換部、

音声を聴く位置である聴覚位置を取得する位置取得部、

前記聴覚位置と、前記複数の音源位置の各々との間の角度及び距離を算出する移動処理部、

50

前記複数の立体音のそれぞれを、前記複数の音源位置のそれぞれに対応する角度及び距離で調整することで、前記聴覚位置における複数の立体音である複数の調整済立体音を生成する角度距離調整部、及び、

前記複数の調整済立体音を重畳する重畳部、として機能させることを特徴とするプログラム。

【請求項 8】

複数の音源からの音声を含む音声データを取得し、
前記音声データから、前記複数の音源の位置である複数の音源位置を判定し、
前記音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成し、
前記複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、前記複数の音源に対応する複数の立体音を生成し、
音声を聴く位置である聴覚位置を取得し、
前記聴覚位置と、前記複数の音源位置の各々との間の角度及び距離を算出し、
前記複数の立体音のそれぞれを、前記複数の音源位置のそれぞれに対応する角度及び距離で調整することで、前記聴覚位置における複数の立体音である複数の調整済立体音を生成し、
前記複数の調整済立体音を重畳すること
を特徴とする音空間構築方法。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、音空間構築装置、音空間構築システム、プログラム及び音空間構築方法に関する。

【背景技術】

【0002】

現在、立体音響技術の開発が進んでいる。例えば、アンビソニックス方式を用いることで、マイク位置での360度方向の音場を再現することが可能である。アンビソニックス方式の実現には、通常、アンビソニックスマイクが用いられる。アンビソニックスマイクが固定されていると、仮想空間内を体験者が自由に動いた際に、移動した場所での音場を再現することができない。

【0003】

これに対して、特許文献1には、方向性オーディオを捕捉するマイクロフォンシステムの空間データに応答して、捕捉された方向性オーディオの方向特性を修正するように適応された装置が開示されている。これにより、視聴位置の移動に応じた方向制オーディオの方向特性を修正することができる。

【先行技術文献】

【特許文献】

【0004】

【文献】特表2022-509761号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、従来の技術では、音源が二つ以上ある場合、視聴位置の移動に対するアンビソニックスBフォーマットの空間追従を行うことはできない。

【0006】

そこで、本開示の位置又は複数の態様は、集音装置が固定された状態において、自由位置での音場を再現できるようにすることを目的とする。

【課題を解決するための手段】

【0007】

本開示の一態様に係る音空間構築装置は、複数の音源からの音声を含む音声データを取得する音声取得部と、前記音声データから、前記複数の音源の位置である複数の音源位置を判定する音源判定部と、前記音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成する音声抽出部と、前記複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、前記複数の音源に対応する複数の立体音を生成するフォーマット変換部と、音声を聴く位置である聴覚位置を取得する位置取得部と、前記聴覚位置と、前記複数の音源位置の各々との間の角度及び距離を算出する移動処理部と、前記複数の立体音のそれぞれを、前記複数の音源位置のそれぞれに対応する角度及び距離で調整することで、前記聴覚位置における複数の立体音である複数の調整済立体音を生成する角度距離調整部と、前記複数の調整済立体音を重畳する重畳部と、を備えることを特徴とする。

10

【0008】

本開示の一態様に係る音空間構築システムは、音空間構築装置と、前記音空間構築装置に対してネットワークで接続され、複数の音源からの音声を含む音声データを生成する集音装置とを備える音空間構築システムであって、前記音空間構築装置は、前記集音装置と通信を行う通信部と、前記通信部を介して、前記音声データを取得する音声取得部と、前記音声データから、前記複数の音源の位置である複数の音源位置を判定する音源判定部と、前記音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成する音声抽出部と、前記複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、前記複数の音源に対応する複数の立体音を生成するフォーマット変換部と、音声を聴く位置である聴覚位置を取得する位置取得部と、前記聴覚位置と、前記複数の音源位置の各々との間の角度及び距離を算出する移動処理部と、前記複数の立体音のそれぞれを、前記複数の音源位置のそれぞれに対応する角度及び距離で調整することで、前記聴覚位置における複数の立体音である複数の調整済立体音を生成する角度距離調整部と、前記複数の調整済立体音を重畳する重畳部と、を備えることを特徴とする。

20

【0009】

本開示の一態様に係るプログラムは、コンピュータを、複数の音源からの音声を含む音声データを取得する音声取得部、前記音声データから、前記複数の音源の位置である複数の音源位置を判定する音源判定部、前記音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成する音声抽出部、前記複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、前記複数の音源に対応する複数の立体音を生成するフォーマット変換部、音声を聴く位置である聴覚位置を取得する位置取得部、前記聴覚位置と、前記複数の音源位置の各々との間の角度及び距離を算出する移動処理部、前記複数の立体音のそれぞれを、前記複数の音源位置のそれぞれに対応する角度及び距離で調整することで、前記聴覚位置における複数の立体音である複数の調整済立体音を生成する角度距離調整部、及び、前記複数の調整済立体音を重畳する重畳部、として機能させることを特徴とする。

30

【0010】

本開示の一態様に係る音空間構築方法は、複数の音源からの音声を含む音声データを取得し、前記音声データから、前記複数の音源の位置である複数の音源位置を判定し、前記音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成し、前記複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、前記複数の音源に対応する複数の立体音を生成し、音声を聴く位置である聴覚位置を取得し、前記聴覚位置と、前記複数の音源位置の各々との間の角度及び距離を算出し、前記複数の立体音のそれぞれを、前記複数の音源位置のそれぞれに対応する角度及び距離で調整することで、前記聴覚位置における複数の立体音である複数の調整済立体音を生成し、前記複数の調整済立体音を重畳することを特徴とする。

40

【発明の効果】

50

【 0 0 1 1 】

本開示の一又は複数の態様によれば、集音装置が固定された状態において、自由位置での音場を再現することができる。

【図面の簡単な説明】

【 0 0 1 2 】

【図 1】実施の形態 1 に係る音空間構築装置の構成を概略的に示すブロック図である。

【図 2】音声抽出部の構成を概略的に示すブロック図である。

【図 3】コンピュータの構成を概略的に示すブロック図である。

【図 4】聴覚位置の移動に伴う処理例を説明するための第 1 の例である。

【図 5】聴覚位置の移動に伴う処理例を説明するための第 2 の例である。

10

【図 6】聴覚位置の移動に伴う処理例を説明するための第 3 の例である。

【図 7】実施の形態 2 に係る音空間構築システムの構成を概略的に示すブロック図である。

【図 8】実施の形態 2 における集音装置の構成を概略的に示すブロック図である。

【図 9】実施の形態 2 における音空間構築装置の構成を概略的に示すブロック図である。

【図 10】実施の形態 3 に係る音空間構築装置の構成を概略的に示すブロック図である。

【発明を実施するための形態】

【 0 0 1 3 】

実施の形態 1 .

図 1 は、実施の形態 1 に係る音空間構築装置 1 0 0 の構成を概略的に示すブロック図である。

20

音空間構築装置 1 0 0 は、音声取得部 1 0 1 と、音源判定部 1 0 2 と、音声抽出部 1 0 3 と、フォーマット変換部 1 0 4 と、位置取得部 1 0 5 と、移動処理部 1 0 6 と、角度距離調整部 1 0 7 と、重畳部 1 0 8 と、出力処理部 1 0 9 とを備える。

【 0 0 1 4 】

音声取得部 1 0 1 は、複数の音源からの音声を含む音声データを取得する。

例えば、音声取得部 1 0 1 は、マイク等の集音装置（図示しない）で生成された音声データを取得する。音声データの音声は、アンビソニックス方式に対応したマイクであるアンビソニックスマイクで捕捉されることが望ましいが、複数の無指向マイクで捕捉されてもよい。また、音声取得部 1 0 1 は、図示しない接続 I / F（ I n t e r F a c e ）を介して、集音装置から音声データを取得してもよく、図示しない通信 I / F を介して、インターネット等のネットワークから音声データを取得してもよい。取得された音声データは、音源判定部 1 0 2 に与えられる。

30

【 0 0 1 5 】

音源判定部 1 0 2 は、音声データから、複数の音源の位置である複数の音源位置を判定する。

例えば、音源判定部 1 0 2 は、音声データに含まれている音源の数を判定する音源数判定と、音声データに含まれている音源の位置である音源位置を推定する音源位置推定とを行う。

【 0 0 1 6 】

音源数判定には、公知の技術が用いられればよい。例えば、下記の文献 1 には、音源数判定として、独立成分分析による音源数推定法が記載されている。

40

【 0 0 1 7 】

また、音源判定部 1 0 2 は、図示しないカメラ等の撮像装置から得られた画像データで示される画像を解析することで、音源を特定して、その音源の数を判定してもよい。言い換えると、音源判定部 1 0 2 は、複数の音源が含まれる空間を撮像した画像を用いて、複数の音源位置を判定してもよい。例えば、音源となる物体の方向及び大きさにより、その物体の位置を判定することができる。

【 0 0 1 8 】

音源位置推定にも、公知の技術が用いられればよい。例えば、下記の文献 2 には、ビームフォーミング法及び M U S I C 法による音源位置の推定方法が記載されている。

50

【 0 0 1 9 】

音声データ及びその音声データに対する音源数判定による音源数を示す音源数データは、音声抽出部 1 0 3 に与えられる。

音源位置推定による音源位置を示す音源位置データは、移動処理部 1 0 6 に与えられる。

【 0 0 2 0 】

音声抽出部 1 0 3 は、音声データで示される音声を音源毎に抽出して、抽出された音声を示す抽出音声データを生成することで、複数の抽出音声データを生成する。複数の抽出音声データのそれぞれは、複数の音源のそれぞれに対応する。

例えば、音声抽出部 1 0 3 は、音声データから音源毎の音声データである抽出音声データを抽出する。具体的には、音声抽出部 1 0 3 は、音声データから、複数の音源に含まれる一つの音源からの音声を分離した残りのデータを、その音声データから減算することで、複数の抽出音声データの内、その一つの音源に対応する抽出音声データを生成する。抽出音声データは、フォーマット変換部 1 0 4 に与えられる。

10

【 0 0 2 1 】

図 2 は、音声抽出部 1 0 3 の構成を概略的に示すブロック図である。

音声抽出部 1 0 3 は、騒音低減部 1 1 0 と、抽出処理部 1 1 1 とを備える。

騒音低減部 1 1 0 は、音声データから騒音を低減する。騒音の低減方法は、公知の技術が使用されればよい。例えば、騒音低減部 1 1 0 は、下記の文献 5 に記載されている G S C (G l o b a l S i d e l o b e C a n c e l l e r) を用いて、騒音を低減すればよい。音声データから騒音が低減された処理済音声データは、抽出処理部 1 1 1 に与えられる。

20

【 0 0 2 2 】

抽出処理部 1 1 1 は、処理済音声データから、音源毎の音声データである抽出音声データを抽出する。

抽出処理部 1 1 1 は、音源分離部 1 1 2 と、位相調整部 1 1 3 と、減算部 1 1 4 とを備える。

【 0 0 2 3 】

音源分離部 1 1 2 は、処理済音声データから、音源毎の音声データを分離することで、分離音声データを生成する。音源毎の音声データを分離する方法については、公知の方法が使用されればよい。例えば、音源分離部 1 1 2 は、下記の文献 4 に記載されている I L R M A (I n d e p e n d e n t L o w - R a n k M a t r i x A n a l y s i s) という技術を用いて分離を行う。

30

【 0 0 2 4 】

位相調整部 1 1 3 は、音源分離部 1 1 2 における音源分離に用いた信号処理で、音源毎に、与えられた位相回転を抽出し、その位相回転をキャンセルする逆側の位相回転を、処理済音声データに与えることで、位相調整済音声データを生成する。位相調整済音声データは、減算部 1 1 4 に与えられる。

【 0 0 2 5 】

減算部 1 1 4 は、音源毎に、処理済音声データから位相調整済音声データを減算することで、音源毎の音声データである抽出音声データを抽出する。

40

【 0 0 2 6 】

図 1 に戻り、フォーマット変換部 1 0 4 は、複数の抽出音声データのフォーマットを、立体音響のフォーマットに変換することで、複数の音源に対応する複数の立体音を生成する。

例えば、フォーマット変換部 1 0 4 は、抽出音声データを立体音響フォーマットに変換する。ここでは、フォーマット変換部 1 0 4 は、抽出音声データのフォーマットを、立体音響フォーマットであるアンビソニックス B フォーマットに変換することで、立体音を示す立体音データを生成する。

【 0 0 2 7 】

なお、音声アンビソニックスマイクで捕捉されている場合には、フォーマット変換部

50

104は、抽出音声データのアンビソニックスAフォーマットを、アンビソニックスBフォーマットに変換すればよい。アンビソニックスAフォーマットからアンビソニックスBフォーマットへの変換方法は、公知の技術が使用されればよい。例えば、下記の文献5には、アンビソニックスAフォーマットからアンビソニックスBフォーマットへの変換方法が記載されている。

【0028】

一方、音声が複数の無指向マイクで捕捉されている場合には、フォーマット変換部104には、公知の技術を利用して、抽出音声データのフォーマットを、アンビソニックスBフォーマットに変換すればよい。例えば、下記の文献6には、無指向マイクで集音された結果をビームフォーミングにより双指向性を生成することで、アンビソニックスBフォーマットを生成する方法が記載されている。

10

【0029】

位置取得部105は、音声を聴く位置である聴覚位置を取得する。例えば、位置取得部105は、図示しないマウス又はキーボード等の入力I/Fを介して、ユーザから、ユーザが仮想空間において音を聴く聴覚位置の特定を受けることで、その聴覚位置を取得する。ここでは、ユーザは、仮想空間を移動できることが前提となっているため、位置取得部105は、定期的に、又は、ユーザの移動が検知される毎に、その聴覚位置を取得する。

そして、位置取得部105は、取得された聴覚位置を示す位置データを移動処理部106に与える。

【0030】

20

移動処理部106は、聴覚位置と、複数の音源位置の各々との間の角度及び距離を算出する。

例えば、移動処理部106は、位置データで示される聴覚位置と、音源位置データで示される音源位置とから、音源位置毎に、聴覚位置との間の角度及び距離を算出する。

そして、移動処理部106は、音源毎に、算出された角度及び距離を示す角度距離データを角度距離調整部107に与える。

【0031】

角度距離調整部107は、複数の立体音のそれぞれを、複数の音源位置のそれぞれに対応する角度及び距離で調整することで、聴覚位置における複数の立体音である複数の調整済立体音を生成する。

30

例えば、角度距離調整部107は、音源毎に、立体音データを、角度距離データで示される角度及び距離となるように調整する。

例えば、角度距離調整部107は、アンビソニックスの規格に従って、アンビソニックスBフォーマットにおける音源からの音の到来方向に対応する角度を、容易に変更することができる。

【0032】

また、角度距離調整部107は、立体音データにおける振幅を、角度距離データで示される距離に応じて調整する。例えば、聴覚位置と音源との距離が、音声データが取得された際における捕捉位置と音源との距離の半分になるのであれば、角度距離調整部107は、その振幅を6dB大きくする。言い換えると、角度距離調整部107は、距離と振幅との関係を、例えば二乗則に従って調整すればよい。

40

【0033】

角度距離調整部107は、音源毎に、角度及び距離を調整した立体音である調整済立体音を示す調整済立体音データを重畳部108に与える。

【0034】

重畳部108は、複数の調整済立体音を重畳する。

例えば、重畳部108は、音源毎の調整済立体音データを重畳する。具体的には、重畳部108は、音源毎の調整済立体音データで示される音信号を足し合わせる。これにより、重畳部108は、足し合わされた音信号を示す合成音データを生成する。合成音データは、出力処理部109に与えられる。

50

【 0 0 3 5 】

出力処理部 1 0 9 は、合成音データで示されるチャンネルベースの音を、両耳で授聴するための音であるバイノーラル音に変換することで、出力音を示す出力音データを生成する。チャンネルベースの音を、バイノーラル音に変換する方法については、公知の方法が使用されればよい。例えば、下記の文献 7 には、チャンネルベースの音を、バイノーラル音に変換する方法が記載されている。

【 0 0 3 6 】

そして、出力処理部 1 0 9 は、例えば、図示しない接続 I / F を介して、スピーカ等の音声出力装置に出力音データを出力する。または、出力処理部 1 0 9 は、図示しない通信 I / F を介して、スピーカ等の音声出力装置に出力音データを出力する。

10

【 0 0 3 7 】

以上に記載された音空間構築装置 1 0 0 は、図 3 に示されているようなコンピュータ 1 0 により実現することができる。

コンピュータ 1 0 は、例えば、HDD (Hard Disk Drive) 及び SSD (Solid State Drive) 等の補助記憶装置 1 1 と、メモリ 1 2 と、CPU (Central Processing Unit) 等のプロセッサ 1 3 と、キーボード及びマウス等の入力 I / F 1 4 と、USB (Universal Serial Bus) 等の接続 I / F 1 5 と、NIC (Network Interface Card) 等の通信 I / F 1 6 とを備える。

【 0 0 3 8 】

20

具体的には、音声取得部 1 0 1、音源判定部 1 0 2、音声抽出部 1 0 3、フォーマット変換部 1 0 4、位置取得部 1 0 5、移動処理部 1 0 6、角度距離調整部 1 0 7、重畳部 1 0 8 及び出力処理部 1 0 9 は、プロセッサ 1 3 が補助記憶装置 1 1 に記憶されているプログラムをメモリ 1 2 にロードして、そのプログラムを実行することで実現可能である。

【 0 0 3 9 】

そのプログラムは、図示しないリーダー/ライターを介して記録媒体から、あるいは、通信 I / F 1 6 を介してネットワークから、補助記憶装置 1 1 にダウンロードされ、それから、メモリ 1 2 上にロードされてプロセッサ 1 3 により実行されるようにしてもよい。また、そのプログラムは、リーダー/ライターを介して記録媒体から、あるいは、通信 I / F 1 6 を介してネットワークから、メモリ 1 2 上に直接ロードされ、プロセッサ 1 3 により実行されるようにしてもよい。

30

【 0 0 4 0 】

アンビソニックス方式では、ユーザが向いている方向に応じて、音源からの音の到来方向を変更できるようになっている。

しかしながら、図 4 に示されているように、第 1 の音源 2 0 及び第 2 の音源 2 1 のように複数の音源がある場合において、ユーザ 2 2 が第 1 の聴覚位置 2 3 から第 2 の聴覚位置 2 4 に移動すると、ユーザ 2 2 と第 1 の音源 2 0 との間の角度は、角度 θ_1 から角度 θ_2 に変わり、ユーザ 2 2 と第 2 の音源 2 1 との間の角度は、角度 θ_3 から角度 θ_4 に変わる。

【 0 0 4 1 】

従来からのアンビソニックス方式では、ユーザの向きの変更等のように、一様な角度の変更は収容可能であるが、図 4 に示されているように、音源毎の角度の変更を行うことはできない。

40

【 0 0 4 2 】

このため、実施の形態 1 は、例えば、図 5 及び図 6 に示されているように、音声データから、第 1 の音源 2 0 からの抽出音声データと、第 2 の音源 2 1 からの抽出音声データとを抽出して、処理を行う。

【 0 0 4 3 】

具体的には、図 5 に示されているように、ユーザ 2 2 が第 1 の聴覚位置 2 3 から第 2 の聴覚位置 2 4 に移動した場合、実施の形態 1 は、ユーザ 2 2 と第 1 の音源 2 0 との間の角度を第 1 の角度 θ_1 から第 2 の角度 θ_2 に変更する。さらに、実施の形態 1 は、第 1 の音

50

源 2 0 からの音の強さも、第 1 の聴覚位置 2 3 と、第 1 の音源 2 0 との間の第 1 の距離 d_1 から、第 2 の聴覚位置 2 4 と、第 1 の音源 2 0 との間の第 2 の距離 d_2 への変化に応じて変更している。

【 0 0 4 4 】

また、図 6 に示されているように、ユーザ 2 2 が第 1 の聴覚位置 2 3 から第 2 の聴覚位置 2 4 に移動した場合、実施の形態 1 は、ユーザ 2 2 と第 2 の音源 2 1 との間の角度を第 3 の角度 θ_3 から第 4 の角度 θ_4 に変更する。さらに、実施の形態 1 は、第 2 の音源 2 1 からの音の強さも、第 1 の聴覚位置 2 3 と、第 2 の音源 2 1 との間の第 3 の距離 d_3 から、第 2 の聴覚位置 2 4 と、第 2 の音源 2 1 との間の第 4 の距離 d_4 への変化に応じて変更している。

10

【 0 0 4 5 】

そして、実施の形態 1 は、以上のようにして、音源毎に処理されたデータを重畳することで、ユーザの移動に伴って音を変化させている。

このため、実施の形態 1 によれば、複数の音源が存在していても、仮想空間における自由位置での音場を再現することができる。

【 0 0 4 6 】

実施の形態 2 .

図 7 は、実施の形態 2 に係る音空間構築システム 2 3 0 の構成を概略的に示すブロック図である。

音空間構築システム 2 3 0 は、音空間構築装置 2 0 0 と、集音装置 2 4 0 とを備える。

20

音空間構築装置 2 0 0 と、集音装置 2 4 0 とは、インターネット等のネットワーク 2 3 1 で接続されている。

【 0 0 4 7 】

集音装置 2 4 0 は、音空間構築装置 2 0 0 とは離れた空間における音声を捕捉して、その音声を示す音声データを、ネットワーク 2 3 1 を介して、音空間構築装置 2 0 0 に送信する。

【 0 0 4 8 】

図 8 は、集音装置 2 4 0 の構成を概略的に示すブロック図である。

集音装置 2 4 0 は、集音部 2 4 1 と、制御部 2 4 2 と、通信部 2 4 3 とを備える。

【 0 0 4 9 】

30

集音部 2 4 1 は、集音装置 2 4 0 が設置された空間における音声を捕捉する。集音部 2 4 1 は、例えば、アンビソニックスマイク又は複数の無指向マイクで構成することができる。

【 0 0 5 0 】

制御部 2 4 2 は、集音装置 2 4 0 での処理を制御する。

例えば、制御部 2 4 2 は、集音部 2 4 1 で捕捉された音声を示す音声データを生成して、通信部 2 4 3 を介して、その音声データを音空間構築装置 2 0 0 に送る。

【 0 0 5 1 】

また、制御部 2 4 2 は、通信部 2 4 3 を介して、音空間構築装置 2 0 0 から、音声を捕捉する方向が示された場合には、集音部 2 4 1 を制御することで、その方向からの音声を示す音声データを生成して、音空間構築装置 2 0 0 に送る。これは、音空間構築装置 2 0 0 でビームフォーミングが行われる際の処理である。

40

【 0 0 5 2 】

以上に記載された制御部 2 4 2 の一部又は全部は、図示されていないが、メモリと、メモリに格納されているプログラムを実行する CPU (Central Processing Unit) 等のプロセッサとにより構成することができる。このようなプログラムは、ネットワークを通じて提供されてもよく、また、記録媒体に記録されて提供されてもよい。即ち、このようなプログラムは、例えば、プログラムプロダクトとして提供されてもよい。

【 0 0 5 3 】

50

また、制御部 242 の一部又は全部は、図示されていないが、単回路、複合回路、プログラムで動作するプロセッサ、プログラムで動作する並列プロセッサ、ASIC (Application Specific Integrated Circuit) 又は FPGA (Field Programmable Gate Array) 等の処理回路で構成することもできる。

以上のように、制御部 242 は、処理回路網により実現することができる。

【0054】

通信部 243 は、ネットワーク 231 を介して、音空間構築装置 200 と通信を行う。

例えば、通信部 243 は、ネットワーク 231 を介して、音空間構築装置 200 に音声データを送信する。

また、通信部 243 は、ネットワーク 231 を介して、音空間構築装置 200 からの指示を受信し、その指示を制御部 242 に与える。

【0055】

ここで、通信部 243 は、図示されていないが、NIC 等の通信 I/F により実現することができる。

【0056】

図 9 は、実施の形態 2 における音空間構築装置 200 の構成を概略的に示すブロック図である。

音空間構築装置 200 は、音声取得部 201 と、音源判定部 202 と、音声抽出部 103 と、フォーマット変換部 104 と、位置取得部 105 と、移動処理部 106 と、角度距離調整部 107 と、重畳部 108 と、出力処理部 109 と、通信部 220 とを備える。

【0057】

実施の形態 2 における音空間構築装置 200 の音声抽出部 103、フォーマット変換部 104、位置取得部 105、移動処理部 106、角度距離調整部 107、重畳部 108 及び出力処理部 109 は、実施の形態 1 における音空間構築装置 100 の音声抽出部 103、フォーマット変換部 104、位置取得部 105、移動処理部 106、角度距離調整部 107、重畳部 108 及び出力処理部 109 と同様である。

【0058】

通信部 220 は、ネットワーク 231 を介して、集音装置 240 と通信を行う。

例えば、通信部 220 は、ネットワーク 231 を介して、集音装置 240 からの音声データを受信する。

また、通信部 220 は、ネットワーク 231 を介して、集音装置 240 に指示を送信する。

なお、通信部 220 は、図 3 に示されている通信 I/F 16 により実現することができる。

【0059】

音声取得部 201 は、通信部 220 を介して、集音装置 240 から音声データを取得する。取得された音声データは、音源判定部 202 に与えられる。実施の形態 2 では、音声データは、音空間構築装置 200 に対してネットワーク 231 で接続された集音装置 240 で捕捉された音声を示すデータである。

【0060】

音源判定部 202 は、音声データに含まれている音源の数を判定する音源数判定と、音声データに含まれている音源の位置である音源位置を推定する音源位置推定とを行う。音源数判定及び音源位置推定は、実施の形態 1 と同様の処理で行われればよい。

なお、音源判定部 202 は、例えば、ビームフォーミング法及び MUSIC 法により音源位置の推定を行う場合には、音声を捕捉する方向を示す指示を、通信部 220 を介して、集音装置 240 に送る。

【0061】

以上のように、実施の形態 2 によれば、集音装置 240 を遠隔地に設置することで、遠隔地から送信されてきた音声を用いて、仮想空間を構築することができる。

10

20

30

40

50

【 0 0 6 2 】

実施の形態 3 .

図 1 0 は、実施の形態 3 に係る音空間構築装置 3 0 0 の構成を概略的に示すブロック図である。

音空間構築装置 3 0 0 は、音声取得部 1 0 1 と、音源判定部 1 0 2 と、音声抽出部 1 0 3 と、フォーマット変換部 1 0 4 と、位置取得部 1 0 5 と、移動処理部 1 0 6 と、角度距離調整部 1 0 7 と、重畳部 3 0 8 と、出力処理部 1 0 9 と、別音声取得部 3 2 1 と、角度距離調整部 3 2 2 とを備える。

【 0 0 6 3 】

実施の形態 3 に係る音空間構築装置 3 0 0 の音声取得部 1 0 1、音源判定部 1 0 2、音声抽出部 1 0 3、フォーマット変換部 1 0 4、位置取得部 1 0 5、移動処理部 1 0 6、角度距離調整部 1 0 7 及び出力処理部 1 0 9 は、実施の形態 1 に係る音空間構築装置 1 0 0 の音声取得部 1 0 1、音源判定部 1 0 2、音声抽出部 1 0 3、フォーマット変換部 1 0 4、位置取得部 1 0 5、移動処理部 1 0 6、角度距離調整部 1 0 7 及び出力処理部 1 0 9 と同様である。

但し、移動処理部 1 0 6 は、角度距離データを角度距離調整部 3 2 2 にも与える。

【 0 0 6 4 】

別音声取得部 3 2 1 は、マイク等の集音装置（図示しない）で生成された音声データを取得する。別音声取得部 3 2 1 で取得される音声データは、音声取得部 1 0 1 で取得される音声データとは、少なくとも捕捉された時間及び位置の何れか一方が異なる音声データであるものとする。別音声取得部 3 2 1 で取得される音声データを、重畳用音声データともいう。

【 0 0 6 5 】

ここで、重畳用音声データは、実施の形態 1 における音源判定部 1 0 2、音声抽出部 1 0 3 及びフォーマット変換部 1 0 4 での処理と同様の処理により、音源毎に分離されて、アンビソニックス B フォーマットに変換されたデータであるものとする。

【 0 0 6 6 】

言い換えると、別音声取得部 3 2 1 は、音声取得部 1 0 1 で取得される音声データに含まれている音声とは、捕捉された時及び場所の少なくとも何れか一方において異なる音声の音声データを、立体音響のフォーマットに変換することで生成された立体音である重畳用立体音を示す重畳用音声データを取得する。

【 0 0 6 7 】

重畳用音声データの音声は、アンビソニックス方式に対応したマイクであるアンビソニックスマイクで捕捉されることが望ましいが、複数の無指向マイクで捕捉されてもよい。また、別音声取得部 3 2 1 は、図示しない接続 I / F を介して、集音装置から音声データを取得してもよく、図示しない通信 I / F を介して、インターネット等のネットワークから音声データを取得してもよい。さらに、別音声取得部 3 2 1 は、図示しない記憶部から重畳用音声データを取得してもよい。取得された重畳用音声データは、角度距離調整部 3 2 2 に与えられる。

【 0 0 6 8 】

角度距離調整部 3 2 2 は、重畳用立体音から、聴覚位置における立体音である重畳用調整済立体音を生成する重畳用角度距離調整部として機能する。

角度距離調整部 3 2 2 は、音源毎に、重畳用音声データを、角度距離データで示される角度及び距離となるように調整する。例えば、重畳用音声データが、音声取得部 1 0 1 で取得される音声データの音声と同じ場所における過去の音声を示す場合には、角度距離調整部 3 2 2 は、角度距離データに従って、角度及び振幅を調整すればよい。角度及び振幅の調整方法については、実施の形態 1 における角度距離調整部 1 0 7 での調整方法と同様である。

【 0 0 6 9 】

一方、重畳用音声データが、音声取得部 1 0 1 で取得される音声データの音声とは、異

10

20

30

40

50

なる場所における音声を示す場合には、角度距離データで示される角度及び距離に応じて、音源毎に角度及び振幅を調整する基準が予め定められており、その基準に従って、角度距離調整部 322 は、重畳用音声データの角度及び振幅を調整すればよい。

【0070】

角度距離調整部 322 は、音源毎に、角度及び距離を調整した重畳用立体音である重畳用調整済立体音を示す重畳用調整済音声データを重畳部 308 に与える。

【0071】

重畳部 308 は、複数の調整済立体音及び重畳用調整済立体音を重畳する。

例えば、重畳部 308 は、音源毎の調整済立体音データ及び重畳用調整済音声データを重畳する。具体的には、重畳部 308 は、音源毎の調整済立体音データで示される音信号及び調整済重畳用音声データで示される音信号を足し合わせる。これにより、重畳部 308 は、足し合わされた音信号を示す合成音データを生成する。合成音データは、出力処理部 109 に与えられる。

【0072】

以上に記載された別音声取得部 321 及び角度距離調整部 322 も、図 3 に示されているプロセッサ 13 が補助記憶装置 11 に記憶されているプログラムをメモリ 12 にロードして、そのプログラムを実行することで実現可能である。

【0073】

以上のように、実施の形態 3 によれば、現実には発生していない別の音声も仮想空間に付加することができるため、例えば、遠隔旅行等の価値を向上することができる。具体的には、ユーザは、仮想空間における聴覚位置での過去の音声、又は、仮想空間とは別の空間での音声を聞くことができる。例えば、ユーザは、現在はない首里城の中で収録された音声を、仮想空間において聞くことができる。

【0074】

文献 1：澤田他、「独立成分分析を用いた音源数推定法」、日本音響学会、秋季研究発表会、2004 年

文献 2：浅野 太、「音のアレイ信号処理 - 音源の定位・追跡と分離」、4.5 章、コロナ社、2011 年

文献 3：浅野 太、「音のアレイ信号処理 - 音源の定位・追跡と分離」、4.5 章、コロナ社、2011 年

文献 4：北村他、「独立低ランク行列分析に基づくブラインド音源分離」、IEICE Technical Report、EA2017-56、vol. 117、No. 255、pp. 73-80、Toyama, October 2017

文献 5：西村 竜一、「アンビソニックス」、映像情報メディア学会誌、Vol. 68、No. 8、pp. 616-620、2014 年

文献 6：特許第 6742535 号公報

文献 7：特許第 4969978 号公報

【符号の説明】

【0075】

100, 200, 300 音空間構築装置、101, 201 音声取得部、102, 202 音源判定部、103 音声抽出部、104 フォーマット変換部、105 位置取得部、106 移動処理部、107 角度距離調整部、108, 308 重畳部、109 出力処理部、110 騒音低減部、111 抽出処理部、112 音源分離部、113 位相調整部、114 減算部、220 通信部、321 別音声取得部、322 角度距離調整部、230 音空間構築システム、231 ネットワーク、240 集音装置、241 集音部、242 制御部、243 通信部。

10

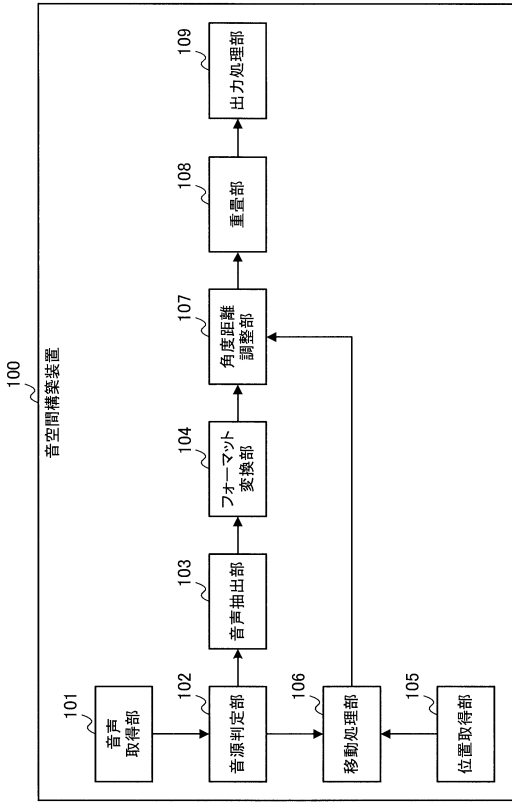
20

30

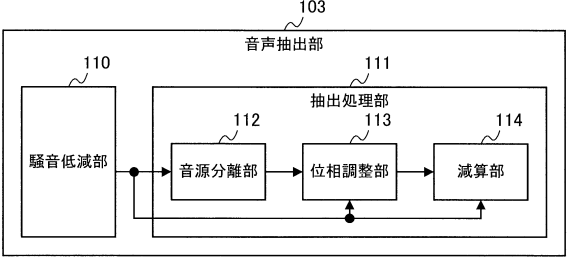
40

【 図 面 】

【 図 1 】



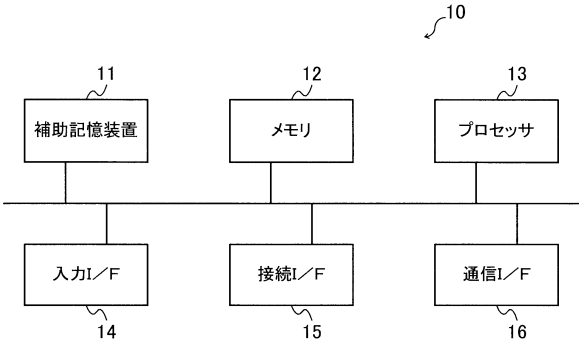
【 図 2 】



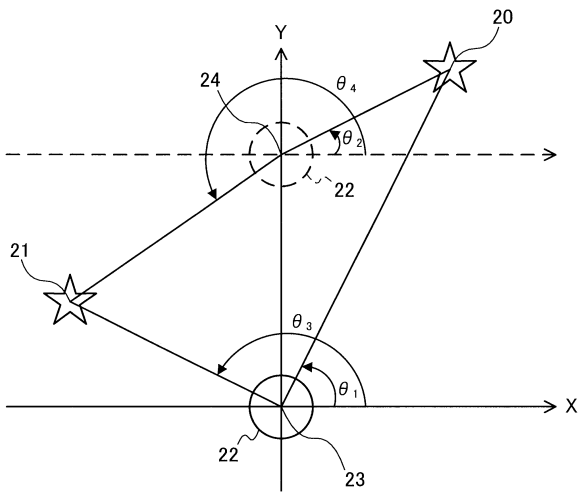
10

20

【 図 3 】



【 図 4 】

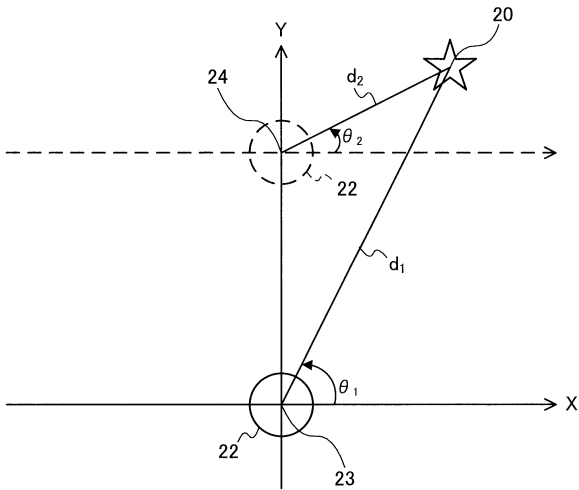


30

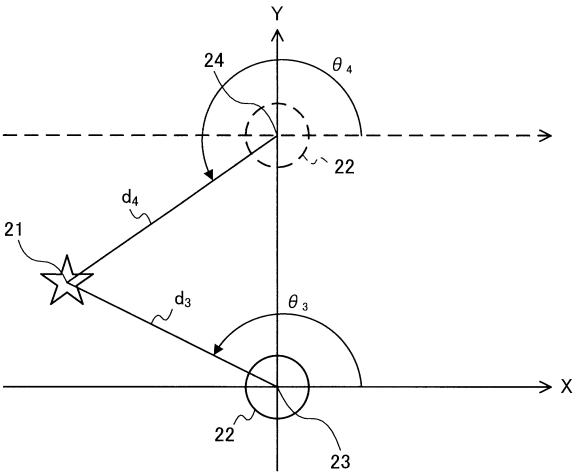
40

50

【図 5】

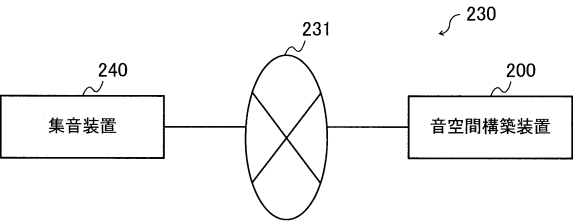


【図 6】

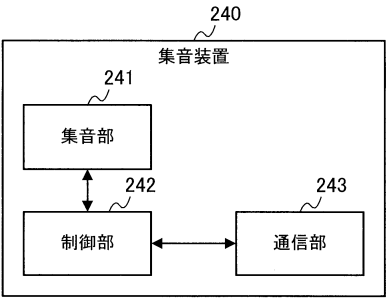


10

【図 7】



【図 8】



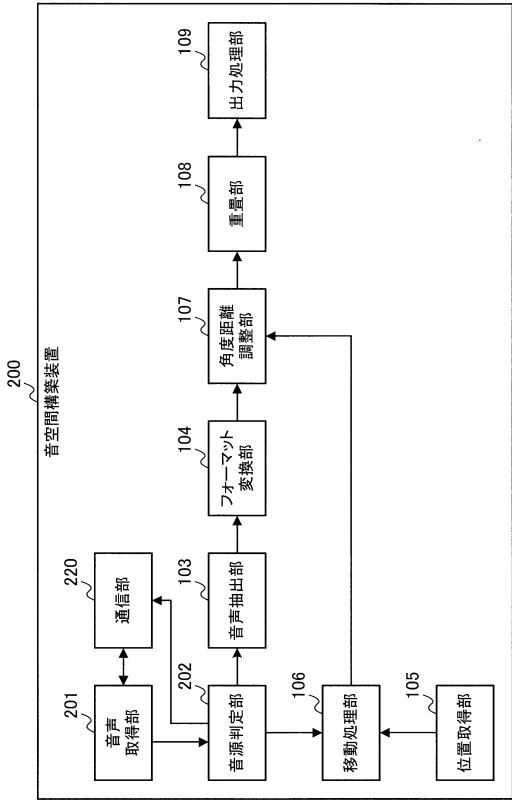
20

30

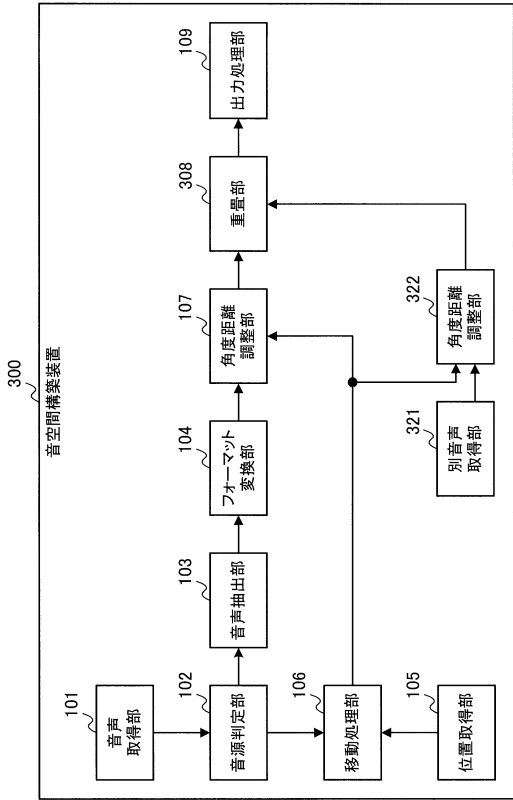
40

50

【図 9】



【図 10】



10

20

30

40

50

フロントページの続き

東京都千代田区丸の内二丁目 7 番 3 号 三菱電機株式会社内
(72)発明者 小川 勇
東京都千代田区丸の内二丁目 7 番 3 号 三菱電機株式会社内
審査官 徳 田 賢二
(56)参考文献 特開 2 0 2 0 - 1 6 7 4 7 1 (J P , A)
特表 2 0 2 0 - 5 2 7 8 8 7 (J P , A)
特表 2 0 2 0 - 5 3 6 2 8 6 (J P , A)
特表 2 0 2 1 - 5 0 8 1 9 7 (J P , A)
(58)調査した分野 (Int.Cl. , D B 名)
H 0 4 S 7 / 0 0