

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7321736号
(P7321736)

(45)発行日 令和5年8月7日(2023.8.7)

(24)登録日 令和5年7月28日(2023.7.28)

(51)国際特許分類

H 0 4 S	7/00 (2006.01)	H 0 4 S	7/00	3 0 0
H 0 4 R	3/00 (2006.01)	H 0 4 R	3/00	3 2 0

請求項の数 18 (全19頁)

(21)出願番号 特願2019-64366(P2019-64366)
 (22)出願日 平成31年3月28日(2019.3.28)
 (65)公開番号 特開2020-167471(P2020-167471)
 A)
 (43)公開日 令和2年10月8日(2020.10.8)
 審査請求日 令和4年3月28日(2022.3.28)

(73)特許権者 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74)代理人 100126240
 弁理士 阿部 琢磨
 (74)代理人 100124442
 弁理士 黒岩 創吾
 (72)発明者 多和田 典朗
 東京都大田区下丸子3丁目30番2号キ
 ャノン株式会社内
 審査官 菊地 陽一

最終頁に続く

(54)【発明の名称】 情報処理装置、情報処理方法、及びプログラム

(57)【特許請求の範囲】**【請求項1】**

それぞれ異なる位置で収音する複数の収音部で生成された収音信号に基づく複数の音響データを取得するデータ取得手段と、

前記データ取得手段により取得された音響データに対応する仮想的な音源位置を、前記音響データと他の音響データとの相関に関する相関情報と、前記音響データに関連付けられる音源の基準位置とに基づいて設定する設定手段と、

前記データ取得手段により取得された前記複数の音響データのうち1以上の音響データを、仮想的な聴取位置と前記設定手段により前記1以上の音響データそれぞれについて設定された仮想的な音源位置とに基づいて、前記聴取位置に応じた音を再生するための音響再生用データを生成する生成手段とを有することを特徴とする情報処理装置。

【請求項2】

前記データ取得手段により取得された音響データに関する前記相関情報と前記音響データに関連付けられる音源の基準位置とに基づいて、前記音響データに対応する仮想的な音源位置を設定可能な設定範囲を決定する決定手段をさらに有し、

前記設定手段は、前記データ取得手段により取得された音響データに対応する仮想的な音源位置を、前記決定手段により決定された設定範囲内で設定することを特徴とする請求項1に記載の情報処理装置。

【請求項3】

前記データ取得手段により取得される音響データについて前記決定手段により決定され

る設定範囲は、前記音響データに関連付けられる音源の基準位置を含むことを特徴とする請求項2に記載の情報処理装置。

【請求項4】

前記基準位置は、前記収音部が設置される位置、又は、前記収音部が設置される位置から前記収音部の指向方向へ離れた位置であることを特徴とする請求項3に記載の情報処理装置。

【請求項5】

前記決定手段により決定された設定範囲内の位置を指定するユーザ操作に応じた入力を受け付ける受付手段をさらに有し、

前記設定手段は、前記データ取得手段により取得される音響データに対応する仮想的な音源位置を、前記受付手段により受け付けられた入力に応じて設定することを特徴とする請求項2乃至4の何れか1項に記載の情報処理装置。

10

【請求項6】

前記決定手段は、前記データ取得手段により取得される第1音響データに関する前記相関情報が表す音の局所性が、前記データ取得手段により取得される第2音響データに関する前記相関情報が表す音の局所性よりも高い場合に、前記第1音響データに対応する設定範囲が前記第2音響データに対応する設定範囲よりも小さくなるように決定することを特徴とする請求項2乃至5の何れか1項に記載の情報処理装置。

【請求項7】

前記決定手段は、前記複数の音響データが分類される複数のグループそれぞれに対応する位置範囲を決定し、前記データ取得手段により取得される音響データに対応する設定範囲を前記音響データが属するグループの位置範囲に決定することを特徴とする請求項2乃至6の何れか1項に記載の情報処理装置。

20

【請求項8】

前記データ取得手段により取得される音響データに関する前記相関情報は、他の音響データとの間に応じた相関係数、コサイン類似度、及びメル周波数ケプストラム係数の少なくとも何れかを示すことを特徴とする請求項1乃至7の何れか1項に記載の情報処理装置。

【請求項9】

前記設定手段は、前記データ取得手段により取得される音響データに対応する音の広がりを、前記音響データに関する前記相関情報と前記音響データの取得に係る収音部の指向性との少なくとも何れかに基づいて設定し、

前記生成手段は、前記1以上の音響データを、前記設定手段により前記1以上のデータそれぞれについて設定された音源位置と音の広がりとに基づいて、前記音響再生用データを生成することを特徴とする請求項1乃至8の何れか1項に記載の情報処理装置。

30

【請求項10】

前記設定手段は、前記データ取得手段により取得される第3音響データに関する前記相関情報が表す音の局所性が、前記データ取得手段により取得される第4音響データに関する前記相関情報が表す音の局所性よりも高い場合に、前記第3音響データに対応する音の広がりが前記第4音響データに対応する音の広がりよりも小さくなるように設定することを特徴とする請求項9に記載の情報処理装置。

40

【請求項11】

前記設定手段は、前記複数の音響データに対応する複数の仮想的な音源のうち隣り合う2つの仮想的な音源の間隔が所定範囲内となるように、前記複数の音源の位置と音の広がりとの少なくとも何れかを設定することを特徴とする請求項1乃至10の何れか1項に記載の情報処理装置。

【請求項12】

前記設定手段は、前記複数の仮想的な音源の位置の移動量が最小化されるように、前記複数の仮想的な音源の位置と音の広がりとの少なくとも何れかを設定することを特徴とする請求項11に記載の情報処理装置。

50

【請求項 1 3】

前記設定手段は、前記複数の音響データに対応する複数の仮想的な音源位置の偏りが前記複数の収音部が設置される位置の偏りよりも小さくなるように設定することを特徴とする請求項 1 乃至 1_2 の何れか 1 項に記載の情報処理装置。

【請求項 1 4】

それぞれ異なる位置で収音する複数の収音部で生成された収音信号に基づく複数の音響データを取得するデータ取得手段と、

前記データ取得手段により取得された音響データに対応する仮想的な音源位置を設定する設定手段と、

前記データ取得手段により取得された音響データの取得に係る収音部の指向性と前記収音部の位置とに基づいて、前記音響データに対応する仮想的な音源位置を設定可能な設定範囲を決定する決定手段とを有し、

前記設定手段は、前記データ取得手段により取得された音響データに対応する仮想的な音源位置を、前記決定手段により決定された設定範囲内で設定することを特徴とする情報処理装置。

【請求項 1 5】

前記決定手段は、前記データ取得手段により取得される第 1 音響データの取得に係る収音部の指向性が、前記データ取得手段により取得される第 2 音響データの取得に係る収音部の指向性よりも鋭い場合に、前記第 1 音響データに対応する設定範囲が前記第 2 音響データに対応する設定範囲よりも小さくなるように決定することを特徴とする請求項 1_4 に記載の情報処理装置。

【請求項 1 6】

それぞれ異なる位置で収音する複数の収音部で生成された収音信号に基づく複数の音響データを取得するデータ取得工程と、

前記データ取得工程において取得された音響データに対応する仮想的な音源位置を、前記音響データと他の音響データとの相関に関する相関情報と、前記音響データに関連付けられる音源の基準位置とに基づいて設定する設定工程と、

前記データ取得工程において取得された前記複数の音響データのうち 1 以上の音響データを、仮想的な聴取位置と前記設定工程において前記 1 以上の音響データそれぞれについて設定された仮想的な音源位置とに基づいて、前記聴取位置に応じた音を再生するための音響再生用データを生成する生成工程とを有することを特徴とする情報処理方法。

【請求項 1 7】

それぞれ異なる位置で収音する複数の収音部で生成された収音信号に基づく複数の音響データを取得するデータ取得工程と、

前記データ取得工程により取得された音響データに対応する仮想的な音源位置を設定する設定工程と、

前記データ取得工程により取得された音響データの取得に係る収音部の指向性と前記収音部の位置とに基づいて、前記音響データに対応する仮想的な音源位置を設定可能な設定範囲を決定する決定工程とを有し、

前記設定工程では、前記データ取得工程により取得された音響データに対応する仮想的な音源位置を、前記決定工程により決定された設定範囲内で設定することを特徴とする情報処理方法。

【請求項 1 8】

コンピュータを、請求項 1 乃至 1_5 の何れか 1 項に記載の情報処理装置の各手段として機能させるためのプログラム。

【発明の詳細な説明】**【技術分野】****【0 0 0 1】**

本発明は、収音により得られる音響データを処理する技術に関する。

【背景技術】

10

20

30

40

50

【0002】

収音対象となる空間の各所にマイクを配置して収音を行い、各マイクにより得られた音響データをマイクの収音位置に応じて合成することで、収音対象空間の音を再現する音響再生用データを生成できる。例えば、スタジアムに配置された複数のマイクにより得られた複数の音響データを、それぞれのマイクの収音位置に位置する音源の音のデータとして取得し、また、スタジアムの中心を仮想的な聴取位置として設定する。そして、音源の位置と聴取位置との位置関係に基づいて、サラウンドパンニング技術などを用いて各音響データを処理することで、サラウンド再生用の音響信号を生成できる。この音響信号をサラウンド再生システムで再生すれば、再生された音を聞くユーザはあたかも自分がスタジアムの中心にいるかのような音像を知覚することができる。

10

【0003】

特許文献1には、異なる位置に設置された複数のマイクにより収音された音声に、マイクの設置位置と入力された聴取位置とに応じたミキシング処理を施すことで、聴取者に対してサラウンド音声を提供することが開示されている。

【先行技術文献】**【特許文献】****【0004】****【文献】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】**

20

【0005】

複数の位置で収音された音響データに基づいて従来の方法で音響再生用データを生成する場合、収音位置の配置によっては、音響再生用データに基づいて再生される音がユーザに与える臨場感が低くなることが考えられる。例えば、収音対象のスタジアムの制約によって、設置されるマイクの配置に偏りが生じる場合がある。そのような場合に、収音された音響データをマイクの配置に合わせて合成することで生成されたデータを再生すると、聴取位置から見てマイクが存在しない方向からの音が聞こえづらくなり、スタジアムを包み込むような歓声が十分に再現できない虞がある。また、各方向から音が均等に聞こえるように、収音された音響データをマイクの配置に依らずに合成することで生成されたデータを再生すると、局所的な応援の音などが実際とは異なる方向から聞こえることにより違和感が生じる虞がある。

30

【0006】

本発明は上記課題に鑑み、複数の位置で収音された音響データから生成される音響再生用データに基づいて再生される音がユーザに与える臨場感を向上させることを目的とする。

【課題を解決するための手段】**【0007】**

上記の課題を解決するため、本発明に係る情報処理装置は、それぞれ異なる位置で収音する複数の収音部で生成された収音信号に基づく複数の音響データを取得する取得手段と、前記取得手段により取得される音響データに対応する仮想的な音源位置を、前記音響データと他の音響データとの相関に関する相関情報と、前記音響データに関連付けられる音源の基準位置に基づいて設定する設定手段と、前記取得手段により取得された前記複数の音響データのうち1以上の音響データを、仮想的な聴取位置と前記設定手段により前記1以上の音響データそれぞれについて設定された仮想的な音源位置とに基づいて、前記聴取位置に応じた音を再生するための前記音響再生用データを生成する生成手段とを有する。

40

【発明の効果】**【0008】**

本発明によれば、複数の位置で収音された音響データから生成される音響再生用データに基づいて再生される音がユーザに与える臨場感を向上させることができる。

【図面の簡単な説明】**【0009】**

50

【図1】情報処理システム10の構成例を示す図である。

【図2】情報処理装置100の動作の例を示すフローチャートである。

【図3】音源の位置及び半径の設定について説明するための図である。

【図4】音源の位置及び半径の最適化について説明するための図である。

【図5】情報処理装置100のハードウェア構成例を示す図である。

【発明を実施するための形態】

【0010】

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解決手段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。

10

【0011】

[システム構成]

図1は、本実施形態に係る情報処理システム10の構成例を示すブロック図である。情報処理システム10は、情報処理装置100と、収音部110-1から収音部110-MまでのM個の収音部と、再生部120とを有する。以降では、各収音部を特に区別しない場合には単に収音部110と表記する。なお、情報処理システム10は複数の収音部110を有していればよく、収音部110の数は限定されない。また、複数の収音部110の構成は同一であってもよいし、構成が異なる収音部110が存在してもよい。

20

【0012】

収音部110は、単一のマイク素子を有するマイクロホン又は複数のマイク素子を有するマイクアレイであり、複数の収音部110がそれぞれ異なる位置に配置される。本実施形態では、複数の収音部110がスポーツ競技を行うスタジアムに配置され、スタジアムの客席の音を収音する例を中心に説明する。ただし、収音部110による収音対象の音は客席の音に限らず、競技のフィールドで選手等が発する音であってもよい。また、収音部110の設置場所はスタジアムに限定されず、例えばコンサートホールや舞台などであってもよい。収音部110による収音で得られた収音信号は、情報処理装置100へ出力される。

【0013】

情報処理装置100は、収音部110から入力された収音信号に応じた音響データを後述する音源の情報に基づいて処理することで音響再生用データを生成し、生成されたデータに応じた再生用の信号を再生部120へ出力する。情報処理装置100のハードウェア構成例について、図5を用いて説明する。情報処理装置100は、CPU501、ROM502、RAM503、補助記憶装置504、表示部505、操作部506、通信I/F507、及びバス508を有する。

30

【0014】

CPU501は、ROM502やRAM503に格納されているコンピュータプログラムやデータを用いて情報処理装置100の全体を制御する。なお、情報処理装置100がCPU501とは異なる1又は複数の専用のハードウェアを有し、CPU501による処理の少なくとも一部を専用のハードウェアが実行してもよい。専用のハードウェアの例としては、ASIC(特定用途向け集積回路)、FPGA(フィールドプログラマブルゲートアレイ)、およびDSP(デジタルシグナルプロセッサ)などがある。ROM502は、変更を必要としないプログラムやパラメータを格納する。RAM503は、補助記憶装置504から供給されるプログラムやデータ、及び通信I/F507を介して外部から供給されるデータなどを一時記憶する。補助記憶装置504は、例えばハードディスクドライブ等で構成され、音響信号などの種々のコンテンツデータを記憶する。

40

【0015】

表示部505は、例えば液晶ディスプレイやLED等で構成され、ユーザが情報処理装置100を操作するためのGUI(Graphical User Interface)などを表示する。操作部506は、例えばキーボードやマウス、タッチパネル等で構成さ

50

れ、ユーザによる操作を受けて各種の指示を C P U 5 0 1 に入力する。通信 I / F 5 0 7 は、収音部 1 1 0 や再生部 1 2 0 などの外部の装置との通信に用いられる。例えば、情報処理装置 1 0 0 が外部の装置と有線で接続される場合には、通信用のケーブルが通信 I / F 5 0 7 に接続される。なお、情報処理装置 1 0 0 が外部の装置と無線通信する機能を有する場合、通信 I / F 5 0 7 はアンテナを備える。バス 5 0 8 は、情報処理装置 1 0 0 の各部をつないで情報を伝達する。

【 0 0 1 6 】

情報処理装置 1 0 0 は、図 1 に示すように、その機能的な構成要素として記憶部 1 0 1 、信号処理部 1 0 2 、入力部 1 0 3 、表示制御部 1 0 4 、出力部 1 0 6 、及び操作検出部 1 0 5 を有する。これらの各機能部は、図 5 に示したハードウェア構成要素により実現される。入力部 1 0 3 は、収音部 1 1 0 からの入力を受け付け、収音に基づく音響データを記憶部 1 0 1 に記憶する。記憶部 1 0 1 に記憶される音響データは、収音部 1 1 0 から入力された収音信号のデータであってもよいし、収音信号に対してノイズ除去等の処理を行うことで得られるデータであってもよい。また記憶部 1 0 1 は、複数の収音部 1 1 0 の位置や種別等、収音に関する各種の情報も記憶する。

10

【 0 0 1 7 】

信号処理部 1 0 2 は、記憶部 1 0 1 に記憶された音響データに対して後述する各種の処理を行い、再生部 1 2 0 により音を再生するための音響再生用データを生成する。表示制御部 1 0 4 は、表示部 5 0 5 に各種の情報を表示させる。操作検出部 1 0 5 は、操作部 5 0 6 を介して入力されたユーザ操作を検出する。出力部 1 0 6 は、再生用の音響信号を再生部 1 2 0 へ出力する。出力部が出力する音響信号は、信号処理部 1 0 2 により生成された音響再生用データの信号であってもよいし、音響再生用データに対してフォーマット変換などの処理を行うことで得られる信号であってもよい。

20

【 0 0 1 8 】

再生部 1 2 0 は、複数のスピーカを有するサラウンドスピーカシステムやヘッドホンで構成され、情報処理装置 1 0 0 から入力された音響信号に基づいて音を再生する。特に本実施形態における再生部 1 2 0 は、ユーザに対して複数の方向から音を聴かせることで立体音響を実現する。また、再生部 1 2 0 は、情報処理装置 1 0 0 から入力された音響信号に対して D A 変換や増幅処理を行う機能を有していてもよい。

30

【 0 0 1 9 】

なお、図 1 では収音部 1 1 0 と情報処理装置 1 0 0 が直接接続されており、情報処理装置 1 0 0 と再生部 1 2 0 とが直接接続されている例を示すが、情報処理システム 1 0 の構成はこれに限定されない。例えば、収音部 1 1 0 による収音に基づく音響データが情報処理装置 1 0 0 と接続可能な記憶装置（不図示）に記憶され、情報処理装置 1 0 0 はその記憶装置から音響データを取得してもよい。また例えば、情報処理装置 1 0 0 は音響再生用データを情報処理装置 1 0 0 と接続可能な音響機器（不図示）に出力し、その音響機器が音響再生用データに基づく音響信号を再生部 1 2 0 へ出力してもよい。また、情報処理装置 1 0 0 が再生部 1 2 0 を有していてもよい。

【 0 0 2 0 】

40

[情報処理装置の動作]

以下では、本実施形態に係る情報処理装置 1 0 0 の動作について、図 2 のフローチャートを用いて説明する。図 2 に示す処理は、情報処理装置 1 0 0 に収音信号が入力され、音響再生用データの生成に係る処理を開始するための指示を情報処理装置 1 0 0 が受け付けたタイミングで開始される。処理を開始するための指示は、情報処理装置 1 0 0 の操作部 5 0 6 を介したユーザ操作により行われてもよいし、他の装置から指示が入力されてもよい。ただし図 2 に示す処理の実行タイミングは上記タイミングに限定されない。

【 0 0 2 1 】

図 2 に示す処理は収音部 1 1 0 による収音と並行して実行されてもよいし、収音部 1 1 0 による収音が終了した後に実行されてもよい。図 2 に示す処理は、C P U 5 0 1 が R O M 5 0 2 に格納されたプログラムを R A M 5 0 3 に展開して実行することで実現される。

50

なお、図2に示す処理の少なくとも一部を、CPU501とは異なる1又は複数の専用のハードウェアにより実現してもよい。

【0022】

S200では、入力部103が、複数の収音部110からの入力を受け付けて、それぞれ異なる位置で収音する複数の収音部110による収音に基づく複数の音響データを取得する。S201では、信号処理部102が、複数の収音部110による収音対象となる収音エリアの情報を取得する。本実施形態では、収音部110による収音対象がスタジアムの観客席の音であるものとし、収音エリアの情報として図3(a)に示すスタジアム300の観客席エリア302の位置や形状等の幾何情報が取得されるものとする。また、収音エリアの情報には、観客席エリア302を構成する、ホーム側エリア303およびアウェイ側エリア304の幾何情報が含まれていてもよい。

10

【0023】

収音エリアの情報は、記憶部101にあらかじめ記憶されているものとする。ただし、操作部506を介したユーザ操作により、情報処理装置100に収音エリアの情報が入力されてもよい。また、情報処理装置100は、スタジアム全体を撮影する不図示の撮影部(カメラ)から取得した映像信号を解析することで、収音エリアの情報を取得してもよい。このとき、観客のユニフォームの色などから、ホーム側エリア303およびアウェイ側エリア304を識別するようにしてよい。

【0024】

S202では、信号処理部102が、複数の収音部110の情報を取得する。収音部110の情報には、少なくとも各収音部110の位置を示す情報が含まれる。さらに、収音部110の情報には、各収音部110が指向性を有するか否かの種別、指向性を有する場合の指向方向、及び指向性の鋭さに関する情報が含まれていてもよい。

20

【0025】

ここで、収音部110の位置や指向方向は、S201で取得される収音エリアの情報に対応する座標系で記述するものとする。例えば、複数の収音部110が位置する空間であるスタジアムの中心301にグローバル座標系の原点を取り、観客席エリア302の直線部と平行および垂直となるようにx軸およびy軸を取って、それらの軸と垂直な鉛直上方向にz軸を取る。このように定義された座標系における座標値で、各収音部110の位置や指向方向が表される。ただし、収音エリアの情報及び収音部110の情報の形式はこれに限定されない。

30

【0026】

収音部110の情報は、記憶部101に予め記憶されているものとする。ただし、操作部506を介したユーザ操作により信号処理部102に収音部110の情報が入力されてもよい。また、信号処理部102は、不図示の撮影部から取得した映像信号を解析することで、収音部110を検出し、収音部110の情報を取得してもよい。このとき、あらかじめ種々の収音部110の映像を用いて行った学習の結果に基づいて映像解析を行ってよい。また、信号処理部102は、収音部110の情報をその収音部110から直接取得してもよい。このとき、各収音部110が、GPSや姿勢センサを用いて取得した位置や方向の情報を出力してもよい。

40

【0027】

S203からS205の処理は、S200で取得された、複数の収音部110に対応する複数の音響データそれぞれに対して行われる。ここでは処理対象とする音響データを変更しながら繰り返し処理が行われるものとするが、複数の音響データに対する処理が並行して行われてもよい。

【0028】

以降では、各音響データに対応する仮想的な音源を設定して処理が行われる。すなわち、1つの収音部110により収音された音が、収音エリア内の1つの音源から発される音として処理される。例えば、図3(a)における無指向性のマイク110-1は、その周囲で発されてマイクに届いた音を収音するが、情報処理装置100は、そのマイクで収音

50

された音を、位置 321 に存在する仮想的な音源から発された音として処理する。音源の位置は、立体音響を実現するための音響再生用データを生成するために用いられる。すなわち情報処理装置 100 は、後述の処理において、収音部 110 による収音に基づく音響データを、その音響データに対応する音源の位置と仮想的な聴取位置とに基づいて処理することで、聴取位置に応じた音を再生するための音響再生用データを生成する。仮想的な聴取位置は例えばスタジアムの中央などに設定することができ、また、聴取位置を変更することも可能である。

【0029】

S203 では、信号処理部 102 が、対象の音響データのデータ取得に用いられた収音部 110 の位置に基づいて、その音響データに関連付けられる音源の基準位置を決定する。例えば、対象の音響データが無指向性のマイク 110-1 により取得された場合、その音響データに関連づけられる基準位置は、マイク 110-1 の位置と一致する位置 321 となる。一方、対象の音響データがマイクアレイ 110-2 により取得された場合、その音響データに関連付けられる基準位置は、マイクアレイ 110-2 の位置からその指向方向に離れた位置であって観客席エリア 302 内である位置 322 となる。マイクアレイ 110-2 は、信号処理により指向性を形成することができ、収音した音から信号処理により特定の指向方向の音を抽出可能なマイクである。

10

【0030】

また、対象の音響データが指向性マイク 110-3 により取得された場合、その音響データに関連づけられる基準位置は、マイク 110-3 の位置からその指向方向に離れた位置であって観客席エリア 302 内である位置 323 となる。また、対象の音響データが観客席エリア 302 の外に位置する無指向性のマイク 110-4 により取得された場合、無指向性のマイク 110-4 の位置は所定の収音対象エリアである観客席エリア 302 の外となる。そこで、そのような音響データに関連づけられる基準位置は、マイク 110-4 の位置と観客席エリア 302 のエリア内でマイク 110-4 に最も近い位置を結ぶ線上の位置であって観客席エリア 302 内である位置 323 とする。なお、上記の基準位置の決め方は一例であり、これに限定されない。

20

【0031】

S204 では、信号処理部 102 が、S203 で決定した音源の基準位置に基づいて、音源の位置を基準位置から変更可能な範囲（音源の位置を設定可能な範囲）を表す設定範囲を決定する。設定範囲には基準位置が含まれ、例えば基準位置が設定範囲の中心となる。本実施形態において設定範囲は例えば以下のようない目的で設定される。情報処理装置 100 は、音源の位置を基準位置から変更することを可能とする。これにより、音源の位置に従って後述の処理により生成される音響再生用データにおいて、収音位置の偏りに起因する歎声等の偏りを低減することができる。しかしながら、全周の音のバランスを改善するために音源を基準位置から全く自由に動かせるようにしてしまふと、実際に音が発された位置と設定された音源位置との差が大きくなることで、再生される音に違和感が生じる可能性がある。

30

【0032】

特に、スタジアムの観客席エリアの音を収音する例においては、楽器の演奏を含む応援の音などの局所的な（本実施形態では局所性が高いと表現する）音も収音され、このような音が聞こえる位置が大きく変更されると違和感が生じやすい。なお、このような局所性の高い音に対応する音源の位置変更による音場の空間再現性への影響は、ある視点位置から見た映像の再生と共に、その視点位置を仮想的な聴取位置とする音が再生される場合に、視聴者により知覚されやすくなる。このような再生される音の臨場感の低下を抑制することを目的として、設定範囲が設定される。

40

【0033】

音の偏りを低減しつつ音源の位置変更による違和感を抑制するために、信号処理部 102 は、局所性が高い音に対応する音源ほど基準位置からの移動可能量が小さくなるように、設定範囲を決定する。局所性の判定基準としては、例えば収音部 110 の特徴である指

50

向性の鋭さを用いることができる。すなわち、信号処理部 102 は、対象の音響データの取得に用いられた収音部 110 の指向性が鋭いほど、その音響データに対応する音源の設定範囲を小さくする。このようにする理由は、局所性の高い音の収音には指向性の鋭い収音部 110 が用いられることが多いためである。また、指向性の鋭い収音部 110 で収音した音には、指向方向以外の方向からの音の混入が少ないため、他の収音部 110 で収音される音とは異なる特徴的な音が収音されることが多いのである。

【 0 0 3 4 】

図 3 (a) に示す例において、無指向性のマイク 110 - 1 により取得される音響データに対応する音源の設定範囲としては、基準位置 321 から音源位置を大きく変更可能な広い設定範囲 331 が設定される。無指向性のマイク 110 - 4 により取得される音響データに対応する音源の設定範囲としては、同様に、広い設定範囲 334 が設定される。一方、比較的緩い指向性を有する指向性マイク 110 - 3 により取得される音響データに対応する音源の設定範囲として、設定範囲 331 及び設定範囲 334 よりは小さい設定範囲 333 が設定される。また、鋭い指向性を有するマイクアレイ 110 - 2 により取得される音響データに対応する音源の設定範囲としては、設定範囲 333 よりさらに小さい設定範囲 332 が設定される。

10

【 0 0 3 5 】

なお、上記では収音部 110 の特徴として指向性を用いるものとしたが、これに限らず、収音部 110 による収音可能範囲の広さなどの特徴を用いてもよい。また、設定範囲を決定するための局所性の判定基準は上記の例に限定されず、音響データを解析することで得られる特徴量を判定基準として用いてもよい。音響データの特徴量は、例えば、対象の音響データと別の収音部 110 により取得された別の音響データとの間の相関係数やコサイン類似度などを示す相関情報である。相関係数を用いる場合、各音響データに表れる音の伝搬遅延の影響を補正するため、相互相関関数が最大となる遅延補正量のもとで、相関係数を算出するのが好適である。そして、他の音響データとの相関係数の平均値が小さい、すなわち他の音源との相関が低いほど、その音響データに対応する音源の設定範囲を小さくしてもよい。

20

【 0 0 3 6 】

また、相関情報として、各音響データからメル周波数ケプストラム係数 (M F C C) といった特徴量を算出し、公知のクラスタリング技術を用いて、複数の音響データを複数のグループ (特徴量クラスタ) に分類してもよい。そして信号処理部 102 は、各グループに対応する位置範囲を決定し、各音響データに対応する音源の設定範囲を、その音響データが属するグループ (音源が属するクラスタ) に対応する範囲に決定してもよい。図 4 (a) は、8つの音源が4つのクラスタにクラスタリングされた場合の例を示す。図 4 (a)において、同一クラスタに属する音源の基準位置は、同じ图形で表されている。すなわち、黒丸で表される基準位置 411 ~ 413 の音源は第1クラスタに属し、黒星で表される基準位置 414 の音源は第2クラスタに属する。黒三角で表される基準位置 415 ~ 416 の音源は第3クラスタに属し、黒ひし形で表される基準位置 417 ~ 418 の音源は第4クラスタに属する。

30

【 0 0 3 7 】

例えば、第1クラスタに属する音源の設定範囲と、第2クラスタに属する音源の設定範囲との境界は、隣接するクラスタに属する基準位置 413 と基準位置 414 の音源の間にスタジアムの中心 301 から引いた境界 421 のように定められる。このとき、音源の数が少ないクラスタに属する音源ほど局所性が高いものとして、そのクラスタの設定範囲が小さくなるよう、双方向矢印 422 で模式的に表されるように、境界 421 をより局所性が高い音源の基準位置 414 の方に寄せてよい。また、基準位置 411 ~ 413 は、観客席エリア 302 の中のホーム側エリア 303 内であるため、第1クラスタに属する音源の設定範囲をホーム側エリア 303 内に限定してもよい。これらの条件により、第1クラスタに属する音源の設定範囲が、基準位置 411 ~ 413 を含む灰色で塗られた設定範囲 423 に決定される。

40

50

【 0 0 3 8 】

同様に、境界 4 2 4 およびホーム側エリア 3 0 3 に応じて、第 3 クラスタに属する音源の設定範囲が、基準位置 4 1 5 ~ 4 1 6 を含む灰色で塗られた設定範囲 4 2 6 に決定される。また、ホーム側エリア 3 0 3 内で設定範囲 4 2 3 と設定範囲 4 2 6 の間の領域である設定範囲 4 2 7 が、第 2 クラスタに属する音源の設定範囲として決定される。また、第 4 クラスタに属する音源の設定範囲 4 2 8 は、基準位置 4 1 7 ~ 4 1 8 を含むアウェイ側エリア 3 0 4 と一致する。

【 0 0 3 9 】

上記のホーム側エリア 3 0 3 やアウェイ側エリア 3 0 4 の例のように、収音対象の空間を予め分割した分割エリアの情報に基づいて音源の設定範囲を限定することができる。これにより、音響再生用データに基づいて音を再生した場合に例えばホーム側の応援歌がアウェイ側から聞こえてくるといった違和感を抑制できる。同様の限定は、収音部 1 1 0 の指向性に基づいて音源の設定範囲を決定する場合にも適用できる。図 3 (a) の例において、基準位置 3 2 1 の音源の設定範囲 3 3 1 は、ホーム側エリア 3 0 3 内に限定されることで、灰色で示す設定範囲 3 4 1 となる。同様に、基準位置 3 2 4 の音源の設定範囲 3 3 4 は、アウェイ側エリア 3 0 4 内に限定されることで、灰色で示す設定範囲 3 4 4 となる。

10

【 0 0 4 0 】

なお、上記の例では、対象の音響データに対応する音源の設定範囲が、その音響データの特徴及びその音響データの取得に係る収音部 1 1 0 の特徴の何れかに基づいて決定されるものとした。ただしこれに限らず、音響データの特徴と収音部 1 1 0 の特徴との両方に基づいて設定範囲が決定されてもよい。

20

【 0 0 4 1 】

図 2 の S 2 0 5 では、信号処理部 1 0 2 が、対象の音響データに対応する音の広がりを制御するための、音源半径の制御範囲を、音響データの特徴及び収音部 1 1 0 の特徴の少なくとも何れかに基づいて決定する。音の広がりを制御するとは、サラウンド再生用の信号の生成において、MDAP (M u l t i p l e - D i r e c t i o n A m p l i t u d e P a n n i n g) などの公知のパンニング技術を用いて、広い角度範囲のスピーカに音源信号を配分することである。MDAP を用いる場合、VBAP (V e c t o r B a s e A m p l i t u d e P a n n i n g) を用いる場合よりも、再生された音を聞くユーザに単一の音源の音の広がりを感じさせることができる。そのため、特にスタジアムの歓声のように、実際には複数の位置の観客から発される音（非点音源の音）を単一の音源として扱う場合には、音の広がりを大きくすることで再生音の違和感を低減できる。

30

【 0 0 4 2 】

本実施形態において、再生音を聞くユーザが知覚する音の広がりは、信号処理部 1 0 2 により設定される音源半径に応じて変化する。例えば、スタジアムの中心 3 0 1 を仮想的な聴取位置として音響再生用データを生成する場合、対象の音響データに対応する音源半径に応じた大きさの音源を中心 3 0 1 から見た場合の見込み角を算出する。そして、その見込み角に応じた角度範囲のスピーカに対象の音響データに応じた音源信号を配分することで、音の広がりを制御できる。

【 0 0 4 3 】

このように、音源位置の制御に加えて音源半径の制御をすることで、仮想的な聴取位置から見て音源が存在しない方向の範囲を狭めることができ、全周の音のバランスをより改善できる。ただし、音源位置の制御の場合と同じく、収音対象の空間において実際に音が発される位置の範囲と設定された音源半径との差が大きくなると、再生される音に違和感が生じる虞がある。このような違和感の発生を抑制するために、信号処理部 1 0 2 は、音源半径を変更可能な制御範囲を決定する。

40

【 0 0 4 4 】

制御範囲の決定方法の一例として、信号処理部 1 0 2 は、局所性が高い音に対応する音源ほど音源半径が小さくなるように制御範囲を決定する。例えば、音響データの取得に用いられる収音部 1 1 0 の指向性が鋭いほど、その音響データに対応する音源半径が小さく

50

なるように制御範囲を決定する。これは、鋭い指向性で収音した音は、狭い範囲内の位置から発された音である可能性が高いためである。逆に、緩い指向性で収音した音は、広い範囲内の位置から発された音である可能性が高い。そのため、音響データの取得に用いられる収音部 110 の指向性が緩い場合、もしくは収音部 110 が無指向性である場合、音の広がりを表現できるよう、その音響データに対応する音源半径が大きくなるように制御範囲が設定される。

【0045】

図 3 (b) は、音源半径の制御の例を示す。無指向性のマイク 110 - 1 により取得される音響データに対応する音源には、音源の位置 321 を中心として、音源半径が大きくなるような灰色の制御範囲 351 が設定される。そして、制御範囲 351 の中間となる音源半径が、基準半径 361 として設定される。基準半径 361 は音源半径の初期値であり、音源半径は制御範囲 351 内で基準半径 361 から変更可能である。なお、S204 の説明において、音源位置の設定範囲をホーム側エリア 303 内やアウェイ側エリア 304 内に限定する例を示したが、音源半径の制御範囲についても同様の限定を課してもよい。例えば、制御範囲 351 に含まれる最大の音源半径 371 の円のうち、ホーム側エリア 303 外となる部分の割合が所定値（例えば 20 %）以下となるように、制御範囲 351 が限定されてもよい。

10

【0046】

また、無指向性のマイク 110 - 4 により取得される音響データに対応する音源には、音源の位置 324 を中心として灰色の制御範囲 354 が設定され、制御範囲 354 の中間となる音源半径が基準半径 364 として設定される。図 3 (b) の例では、制御範囲 354 に含まれる最大の音源半径 374 の円の全体がアウェイ側エリア 304 内となるような限定を課している。このように制御範囲 354 を決定することで、音響再生用データに基づいて音を再生した場合に例えばアウェイ側の応援歌がホーム側エリアまで広がって聞こえるといった違和感を抑制できる。

20

【0047】

また、緩い指向性を有する指向性マイク 110 - 3 により取得される音響データに対応する音源には、音源の位置 323 を中心として、音源半径が中程度の大きさになるような灰色の制御範囲 353 が設定される。そして、制御範囲 353 の中間となる音源半径が、基準半径 363 として設定される。一方、鋭い指向性を有するマイクアレイ 110 - 2 により取得される音響データに対応する音源には、音源の位置 322 を中心として、音源半径が小さくなるような灰色の制御範囲 352 が設定される。そして、制御範囲 352 の中間となる音源半径が、基準半径 362 として設定される。制御範囲 352 に内で設定される音源半径は、半径の大きさが 0 であってもよい。なお、音源半径の制御範囲は音源位置を中心として設定されるため、音源の位置が基準位置から変更された場合は、変更された音源位置を中心として音源半径の制御範囲が設定される。

30

【0048】

なお、音源位置の設定範囲を決定する場合と同様に、音源半径の制御範囲を決定する場合においても、音響データの局所性が用いられてもよい。例えば、対象となる音響データと他の音響データとの相関度合に応じた局所性が高いほど、その音響データに対応する音源半径が小さくなるように、制御範囲が決定されてもよい。

40

【0049】

S206 では、表示制御部 104 が、音源の基準位置及び設定範囲を示す図 3 (a) や図 4 (a) のような画像、又は、音源の基準半径及び制御範囲を示す図 3 (b) のような画像など、音源の情報に応じた画像を生成して、表示部 505 に表示させる。なお、図 4 (a) に示すような表示において、音源の位置や設定範囲が、音源が属するクラスタごとに色分けして表示されてもよい。また、ホーム側エリア 303 とアウェイ側エリア 304 で色分けした表示がされてもよい。表示制御部 104 による表示の内容は、図 3 や図 4 の例に限定されず、音源の設定範囲及び制御範囲の少なくとも何れかを識別可能にする情報が表示されればよい。

50

【 0 0 5 0 】

S 2 0 7 では、音源位置及び音源半径の調整に係る調整モードの設定を判定し、調整モードが手動モードであれば S 2 0 8 へ、調整モードが自動モードであれば S 2 1 1 へ進む。調整モードの設定は、情報処理装置 1 0 0 に対するユーザ操作に基づいて行われる。ただし、情報処理装置 1 0 0 が收音部 1 1 0 の数や種別などの情報に応じて調整モードを設定してもよい。

【 0 0 5 1 】

S 2 0 8 から S 2 1 0 の処理は、操作検出部 1 0 5 が検出したユーザ操作に応じた入力それぞれに対して行われる。すなわち、ユーザにより複数の操作が行われた場合、操作の数に応じて S 2 0 8 から S 2 1 0 の処理が繰り返される。S 2 0 8 では、操作検出部 1 0 5 が、ユーザ操作により指定された音源と、その音源についての変更指示の内容を判断する。音源についての変更指示の内容とは、音源位置の変更や音源半径の変更である。そして、信号処理部 1 0 2 が、ユーザ操作に応じた音源位置又は音源半径の変更指示が、S 2 0 4 で設定された設定範囲及び S 2 0 5 で設定された制御範囲の条件を満たすかを判定する。信号処理部 1 0 2 は、設定範囲内の位置を新たな音源位置として指定するユーザ操作に応じた入力を受け付けた場合など、条件を満たす場合には、その入力に応じて音源の位置又は半径を設定し、S 2 1 0 へ進む。一方、条件を満たさない場合は S 2 0 9 へ進む。

10

【 0 0 5 2 】

S 2 0 9 では、信号処理部 1 0 2 が、上記の条件を満たすように音源の位置または半径を制御する。例えば、ユーザ操作に応じた変更指示が、音源の位置を設定範囲外に移動させるようなものであった場合、信号処理部 1 0 2 は音源位置を設定範囲の境界に移動させる。同様に、ユーザ操作に応じた変更指示が、音源の半径を制御範囲外の大きさに変更させるようなものであった場合、信号処理部 1 0 2 は音源半径を制御範囲の境界（最大半径または最小半径）に変更する。ただしこれに限らず、ユーザ操作に応じた変更指示が設定範囲や制御範囲の条件を満たさない場合に、信号処理部 1 0 2 は、その変更指示を受付は行わず、音源の位置及び半径を変更しなくてもよい。S 2 0 8 及び S 2 0 9 処理の結果、対象の音響データに対応する音源位置が、S 2 0 4 で決定された設定範囲内で設定される。また、対象の音響データに対応する音源半径が、S 2 0 5 で決定された制御範囲内で設定される。

20

【 0 0 5 3 】

S 2 1 0 では、表示制御部 1 0 4 が、音源の位置や半径に係る表示を更新する。例えば図 3 (a) に示すように、音源の位置が基準位置 3 2 3 から位置 3 8 3 に変更された場合、変更後の音源位置 3 8 3 および十字矢印 3 9 3 による移動可能方向が表示される。ここで、変更後の音源位置 3 8 3 が設定範囲 3 3 3 の境界に近い場合は、音源位置 3 8 3 、設定範囲 3 3 3 、及び十字矢印 3 9 3 の少なくとも何れかの表示色や線幅を変えたり、表示を点滅させたりして目立たせるようにしてもよい。また、十字矢印 3 9 3 について、設定範囲 3 3 3 の境界から離れる方向の矢印のみ表示したり、その方向の矢印を相対的に大きく表示したりしてもよい。

30

【 0 0 5 4 】

また、例えば図 3 (b) に示すように、音源の半径が基準半径 3 6 1 から変更される場合、双方向矢印 3 9 1 により調整可能方向が表示される。ここで、変更後の音源半径が制御範囲 3 5 1 の境界に近い場合は、音源半径、制御範囲 3 5 1 、及び双方向矢印 3 9 1 の少なくとも何れかの表示色や線幅を変えたり、表示を点滅させたりして目立たせるようにしてもよい。また、双方向矢印 3 9 1 について、制御範囲 3 5 1 の境界から離れる方向の矢印のみ表示したり、その方向の矢印を相対的に大きく表示したりしてもよい。なお、ユーザ操作に応じた入力が設定範囲又は制御範囲の条件を満たさない場合に、表示制御部 1 0 4 は、その旨を示す表示やエラー表示を行ってもよい。

40

【 0 0 5 5 】

以上のように手動モードで調整を行うことで、設定範囲及び制御範囲の条件のもと、ユーザ操作に応じて音源の位置や半径を変更することで、再生音が違和感を生じさせること

50

を抑制しつつユーザの意図に応じた音の調整を実現できる。一方、以下で説明する自動モードによる調整によれば、全周の音のバランスについて定量化を行い、制約条件および評価関数を定めることで、ユーザ操作の手間を削減しつつ音源の位置及び半径を最適化することができる。

【0056】

S 2 1 1 では、信号処理部 1 0 2 が、全周の音のバランスを表す指標として、複数の音響データに対応する各音源の間隔を特定する。具体的には、図 4 (b) に示すように、音源半径 4 3 1 ~ 4 3 8 を有する位置 4 1 1 ~ 4 1 8 の音源に対し、仮想的な聴取位置 (図 4 (b) の例ではスタジアムの中心 3 0 1) から接線を引いて、隣接する接線の成す角を音源間隔 とする。例えば、音源半径 4 3 1 を有する位置 4 1 1 の音源と、音源半径 4 3 2 を有する基準位置 4 1 2 の音源との音源間隔は 12 である。なお、音源半径 4 3 2 を有する基準位置 4 1 2 の音源と、音源半径 4 3 3 を有する基準位置 4 1 3 の音源のように、各音源半径に応じた円に重複がある場合は、音源間隔 23 を負の値で表すものとする。

【0057】

ここで、12 や 56 のように音源間隔が大きいと、その方向の音が再生されなくなるため、全周の音のバランスが偏り、音場における包まれ感が十分再現できない。また、音源間隔が絶対値が大きい負の値を取ると、複数の異なる音源の音が略同じ方向から再生されることになるため、この場合も全周の音のバランスが偏り、音場の方向感が十分再現できない。そこで、信号処理部 1 0 2 は、複数の音響データに対応する複数の音源の間隔が所定範囲内となるように、音源位置及び音源半径を設定する。具体的には、全周の音のバランスを表す指標として音源間隔 L を用い、音源間隔 H の好適な範囲を L < H のように定める。これにより、L < H の場合に音場の方向感が十分でなく、H < L の場合に音場における包まれ感が十分でないと判定できる。なお、各音源の音源半径に応じた円について多少の重複を許し、L を負の値としてもよい。

【0058】

S 2 1 2 では、信号処理部 1 0 2 が、音源位置および音源半径を最適化変数とする最適化問題を定式化して解くことで、音源位置・半径の最適化を行う。まず、音源の位置及び半径に係る制約条件として、S 2 0 4 及び S 2 0 5 で設定された設定範囲の条件 (1) と制御範囲の条件 (2) に加えて、全周の音のバランスを改善するための制約条件 (3) 「 $L_i < H (i = 1 \sim N)$ 」を課す。ここで、 i は隣接する音源の音源間隔であり、音源位置および音源半径の関数で表される。また、N は音源の数である。

【0059】

次に、最小化する評価関数については、各音源の基準位置からの移動量の和とする。これは、制約条件を満たしていれば、音源の基準位置からの移動量は少ない方が違和感を感じにくいためである。このとき、局所性が高い音源ほど基準位置からの移動量を小さくするために、局所性が高い音源ほど大きい値となるような重み係数を移動量に対して設定し、重み付き移動量の和を用いるのが好適である。同様の目的で、移動量を設定範囲の大きさなどで正規化した、正規化移動量の和を用いるようにしてもよい。

【0060】

なお、信号処理部 1 0 2 は、操作検出部 1 0 5 により検出したユーザ操作に応じた入力に基づいて、制約条件 (3) における L, H を調整してもよい。これにより例えば、ユーザが H を L に近付けることで音場の方向感を重視したり、L を H に近付けることで音場における包まれ感を重視したりすることができる。

【0061】

信号処理部 1 0 2 は、上記のような最適化問題を解くことで、全音源の位置及び半径の最適化を行う。なお、上記の最適化処理は一例であり、信号処理部 1 0 2 はその他の方法により自動で音源の位置及び半径を制御してもよい。また、信号処理部 1 0 2 はすべての音源に対して制御を行わなくてもよく、複数の音源の中から選択された音源について位置や半径の制御を行ってもよい。S 2 1 1 及び S 2 1 2 処理の結果、対象の音響データに対応する音源位置が、S 2 0 4 で決定された設定範囲内で設定される。また、対象の音響データ

10

20

30

40

50

ータに対応する音源半径が、S 2 0 5 で決定された制御範囲内で設定される。

【0 0 6 2】

S 2 1 3 では、表示制御部 1 0 4 が、音源の位置や半径に係る表示を更新する。例えば、最適化前は図 4 (b) に示すように音源半径 4 3 1 ~ 4 3 8 を有する基準位置 4 1 1 ~ 4 1 8 にあった音源が、最適化後は図 4 (c) に示すように音源半径 4 5 1 ~ 4 5 8 を有する音源位置 4 4 1 ~ 4 4 8 の音源となる。図 4 (b) に示す最適化前の音源の設定では、音源間隔 1 2 、 5 6 、及び 7 8 が H より大きく、また音源間隔 2 3 (< 0) が L より小さかったため、S 2 1 2 で説明した制約条件 (3) を満たしていない。一方、図 4 (c) に示す最適化後の音源の設定では、以下のように音源の位置及び半径の制御を行ったことにより、音源間隔 ' 1 2 、 ' 2 3 、 ' 5 6 、及び ' 7 8 を含め、全ての音源間隔が L ~ H の範囲に収まっている。10

【0 0 6 3】

まず、図 4 (b) において基準半径 4 3 1 ~ 4 3 3 を有する基準位置 4 1 1 ~ 4 1 3 にあった音源は、図 4 (c) において音源位置 4 4 1 ~ 4 4 3 に移動して音源半径 4 5 1 ~ 4 5 3 を有する音源となっている。これらの音源は局所性が低い音に対応し、基準半径 4 3 1 ~ 4 3 3 が大きい。そのため、基準半径に応じた円とホーム側エリア 3 0 3 の境界との間は狭く、音源半径を大きくすることで音源間隔 1 2 を埋めることは難しい。そこで、これらの音源の位置を変更することで、音源間隔 ' 1 2 、 ' 2 3 が制約条件を満たすようにしている。

【0 0 6 4】

また、図 4 (b) において基準半径 4 3 4 を有する基準位置 4 1 4 にあった音源は、局所性が高い音に対応し、音源半径や音源位置の大きな変更に適さない。そのため、図 4 (c) において音源の位置及び半径とともに図 4 (b) の状態を維持している。すなわち、図 4 (c) の音源位置 4 4 4 および音源半径 4 5 4 は、図 4 (b) の基準位置 4 1 4 および基準半径 4 3 4 とそれぞれ同じである。20

【0 0 6 5】

また、図 4 (b) において基準半径 4 3 5 ~ 4 3 6 を有する基準位置 4 1 5 ~ 4 1 6 にあった音源は、図 4 (c) において音源位置 4 4 5 ~ 4 4 6 に移動して音源半径 4 5 5 ~ 4 5 6 を有する音源となっている。これらの音源の基準半径 4 3 5 ~ 4 3 6 に応じた円とホーム側エリア 3 0 3 の境界との間には間隔があるため、図 4 (c) における音源半径 4 5 5 ~ 4 5 6 の大きさは、図 4 (b) における基準半径 4 3 5 ~ 4 3 6 よりも大きくなっている。すなわち、これらの音源については、音源位置および音源半径を制御することで、音源間隔 ' 5 6 が制約条件を満たすようにしている。30

【0 0 6 6】

また、図 4 (b) において基準半径 4 3 7 ~ 4 3 8 を有する基準位置 4 1 7 ~ 4 1 8 にあった音源は、図 4 (c) において音源半径 4 5 7 ~ 4 5 8 を有する音源位置 4 4 7 ~ 4 4 8 の音源となっている。このとき、基準半径 4 3 7 ~ 4 3 8 に応じた円とアウェイ側エリア 3 0 4 の境界との間には間隔があるため、図 4 (c) における音源半径 4 5 7 ~ 4 5 8 の大きさは、図 4 (b) における基準半径 4 3 7 ~ 4 3 8 よりも大きくなっている。一方、図 4 (c) における音源位置 4 4 7 ~ 4 4 8 は、図 4 (b) における基準位置 4 1 7 ~ 4 1 8 をそれぞれ維持している。すなわち、これらの音源については、音源半径を制御することで、音源間隔 ' 7 8 が制約条件を満たすようにしている。40

【0 0 6 7】

上述した手動モードによる調整と自動モードによる調整は、併用されてもよい。例えば、ユーザ操作に応じて音源半径や音源位置が変更される音源以外の音源については、自動で音源位置及び半径が逐次最適化されるようにもよい。また、自動モードによる調整処理 (S 2 1 1 から S 2 1 3 の処理) が行われた後に手動モードによる調整処理 (S 2 0 8 から S 2 1 0 の処理) が行われてもよい。

【0 0 6 8】

S 2 1 4 では、信号処理部 1 0 2 が、複数の收音部 1 1 0 が位置する空間における仮想

10

20

30

40

50

的な聴取位置及び聴取方向を示す情報を取得する。この聴取位置及び聴取方向は、操作検出部 105 が検出したユーザ操作に基づいて決定される。すなわち、ユーザは収音対象の空間内の仮想的な聴取位置及び聴取方向を任意に指定することができる。なお、聴取位置及び聴取方向の少なくとも何れかが、ユーザ操作に依らずに所定の位置及び方向に決定されてもよい。例えば、聴取位置がスタジアムの中心に決定され、聴取方向がスタジアムの中心からメインスタンドに向かう方向に決定されてもよい。また、情報処理装置 100 は、外部の装置から聴取位置及び聴取方向を決定するための情報取得を行ってもよい。例えば、情報処理装置 100 は、音響再生用データと共に再生される映像データに係る視点の位置及び方向を示す視点情報を取得し、その視点情報に基づいて聴取点の位置及び方向を決定してもよい。

10

【 0 0 6 9 】

そして信号処理部 102 は、収音部 110 により取得された音響データを、上述のように設定された音源位置及び音源半径の情報と、仮想的な聴取位置及び聴取方向とに基づいて処理することで、音響再生用データを生成する。これにより、仮想的な聴取位置及び聴取方向に応じた音を再生するための音響再生用データが得られる。具体的には、聴取位置及び聴取方向に対する各音源の方向と見込み角に基づいて、MDAP などの公知のパンニング技術を用いて収音に基づく音響データを処理することで、サラウンド再生用のデータが生成される。なお、信号処理部 102 は、このようなサラウンド再生用のデータに頭部伝達関数 (HRTF) を適用することで、バイノーラル化されたヘッドホン再生用のデータを生成してもよい。

20

【 0 0 7 0 】

また、音響再生用データの生成には、情報処理システム 10 が有する複数の収音部 110 で取得された複数の音響データのうち、仮想的な聴取位置や聴取方向に基づいて選択された 1 以上の音響データが用いられてもよい。このような方法によれば、情報処理装置 100 が処理するデータ量を削減できる。ただし、取得された複数の音響データのすべてを用いて音響再生用データが生成されてもよい。S215 では、出力部 106 が、S214 で生成された音響再生用データに基づく再生用の音響信号を再生部 120 へ出力する。

30

【 0 0 7 1 】

以上で図 2 の説明を終わる。なお、以上の説明では、音源の位置と音源の半径との両方を制御するものとしたが、情報処理装置 100 はこれらの一方のみを制御してもよい。また、以上の説明では分かり易さのために音源の位置や半径を 2 次元的に制御する例を示したが、本実施形態は音源の位置や半径を 3 次元的に制御する場合にも適用できる。例えば情報処理装置 100 は、各音源の位置を高さ方向に制御してもよい。このとき、音源位置を変更可能な設定範囲も 3 次元的に設定されてもよい。また、各音源は音源半径を有する球として表される。

【 0 0 7 2 】

以上説明したように、本実施形態に係る情報処理装置 100 は、それぞれ異なる位置で収音する複数の収音部 110 による収音に基づく複数の音響データを取得する。また情報処理装置 100 は、音響データに対応する仮想的な音源位置を、当該音響データの特徴及び当該音響データの取得に係る収音部 110 の特徴の少なくとも何れかと当該収音部 110 の位置に基づいてきまる設定範囲内で設定する。そして情報処理装置 100 は、取得した複数の音響データのうち 1 以上の音響データを、当該 1 以上の音響データそれぞれについて設定された音源位置と仮想的な聴取位置とに基づいて処理する。この処理により、当該聴取位置に応じた音を再生するための音響再生用データが生成される。

40

【 0 0 7 3 】

以上のような構成によれば、複数の位置で収音された音響データから生成される音響再生用データに基づいて再生される音がユーザに与える臨場感を向上させることができる。例えば、以上のように生成された音響再生用データに基づいて音の再生を行うことで、局所的な音が実際の収音環境における方向とは異なる方向から聞こえることによる違和感を抑制しつつ、音の偏りを低減することでユーザを包み込むような音を再現できる。

50

【 0 0 7 4 】

なお、本実施形態において、S 2 0 0 で取得される音響データは、収音部1 1 0 による収音に基づくデータであるものとした。しかし、実際に収音された音の音響データに代えて、コンピュータにより生成された音響データを用いてもよい。この場合、音響データに対応する音源の位置は、例えばユーザ操作に基づいて設定される。

【 0 0 7 5 】

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ー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、A S I C 等）によっても実現可能である。また、そのプログラムをコンピュータにより読み取り可能な記録媒体に記録して提供してもよい。

10

【 符号の説明 】**【 0 0 7 6 】**

1 0 情報処理システム

1 0 0 情報処理装置

1 1 0 収音部

20

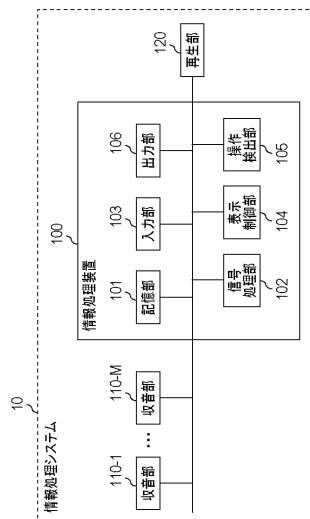
30

40

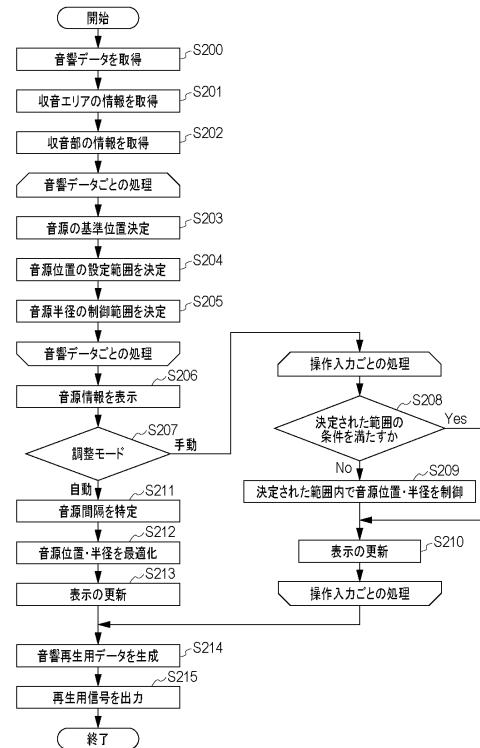
50

【四面】

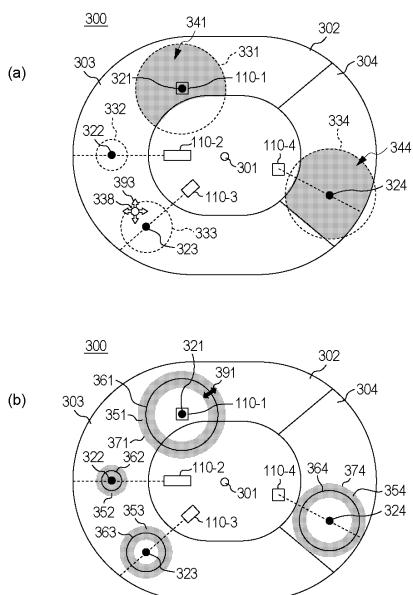
【 四 1 】



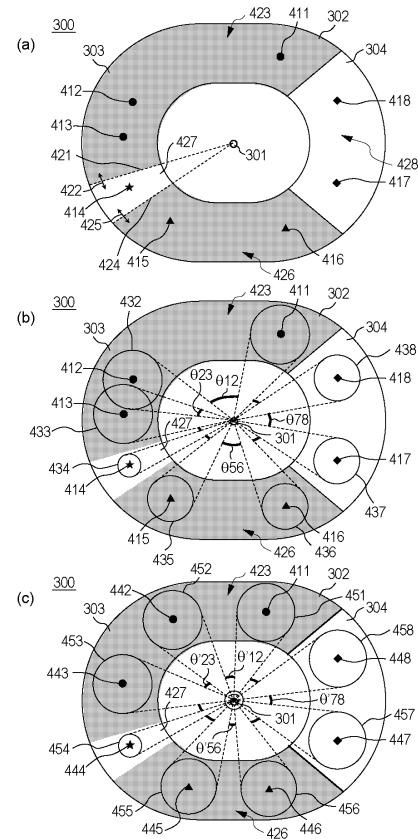
【 図 2 】



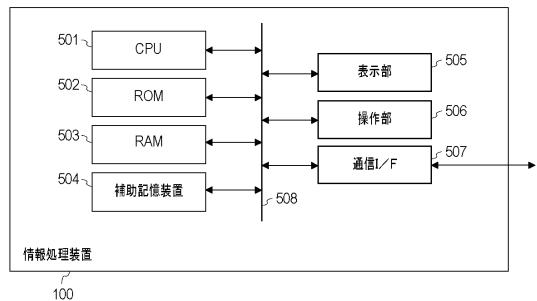
【図3】



【図4】



【図 5】



10

20

30

40

50

フロントページの続き

(56)参考文献 特開2018-191127(JP,A)

特開2006-180251(JP,A)

特開2015-082844(JP,A)

(58)調査した分野 (Int.Cl., DB名)

H04S 7/00

H04R 3/00