

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2016-146547
(P2016-146547A)

(43) 公開日 平成28年8月12日(2016.8.12)

(51) Int.Cl.	F 1	HO 4 R	1/40	3 2 O Z	5 D 2 2 0	テーマコード (参考)
HO 4 R 1/40 (2006.01)	HO 4 R	1/40				
GO 1 S 7/52 (2006.01)	GO 1 S	7/52		F		5 J 0 8 3
GO 1 S 15/87 (2006.01)	GO 1 S	15/87				
HO 4 R 3/00 (2006.01)	HO 4 R	3/00		3 2 O		

審査請求 未請求 請求項の数 5 O L (全 15 頁)

(21) 出願番号	特願2015-22423 (P2015-22423)	(71) 出願人	314012076 パナソニックIPマネジメント株式会社 大阪府大阪市中央区城見2丁目1番61号
(22) 出願日	平成27年2月6日 (2015.2.6)	(74) 代理人	110002000 特許業務法人栄光特許事務所
		(72) 発明者	松本 宏之 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	渡辺 周一 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	辻 寿嗣 大阪府門真市大字門真1006番地 パナソニック株式会社内
			F ターム (参考) 5D220 BA01 BC05

最終頁に続く

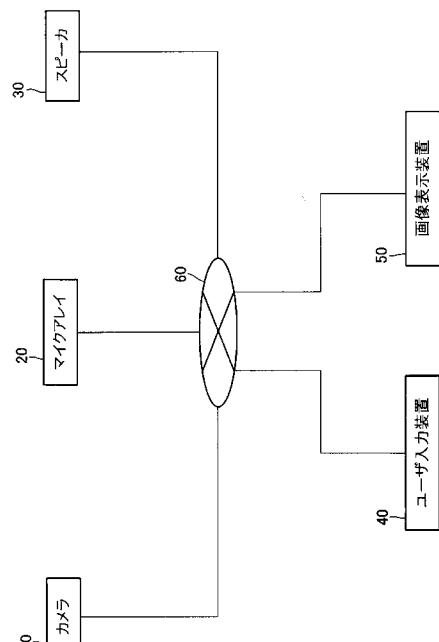
(54) 【発明の名称】 収音システム及び収音方法

(57) 【要約】

【課題】 収音システムにおいて、対象領域内の所望の音源からの音を収音する際の利便性を向上させる。

【解決手段】 対象空間の映像を撮像するカメラと、複数のマイクロホンを用いて音声を収音するマイクアレイと、収音された音声データに基づき音源位置を算出する音源解析部55と、撮像された対象空間の映像と音源位置を示す音源位置表示とを含む画像データをディスプレイ53に表示させる画像表示処理部52と、ディスプレイ53に表示された対象空間画像における指定箇所の指示を受け付けるユーザ入力装置と、収音された音声データを基に、マイクアレイから指定箇所に向かう方向に収音の指向性を形成し、指向性を有する方向の音声を強調した音声データを生成する指向性演算部57と、指向性処理後の音声データをスピーカに出力して再生させる音声出力部58と、を備える。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

対象空間の映像を撮像する撮像部と、
前記撮像部により撮像された画像データを表示する表示部と、
複数のマイクロホンを含み、前記マイクロホンを用いて音声を収音する収音部と、
前記収音部により収音された音声データを再生出力する音声再生部と、
前記音声データに基づき音源位置を算出する音源解析部と、
前記撮像された対象空間の映像と前記音源位置を示す音源位置表示とを含む画像データを前記表示部に表示させる画像表示処理部と、
前記表示部に表示された前記画像データにおける指定箇所の指示を受け付ける指示入力部と、
前記音声データを基に、前記収音部から、指定された前記画像データの指定箇所に対応する位置に向かう方向に、収音の指向性を形成し、前記指向性を有する方向の音声を強調した音声データを生成又は合成する指向性演算部と、
前記指向性に関する処理後の音声データを前記音声再生部に出力して再生させる音声出力部と、
を備える、収音システム。

【請求項 2】

請求項 1 に記載の収音システムであって、
前記音源解析部により算出された音源位置が複数ある場合に、前記画像表示処理部は、
複数の音源位置のそれぞれに対応する前記音源位置表示を生成して前記表示部に表示させる、収音システム。

【請求項 3】

請求項 2 に記載の収音システムであって、
前記指示入力部により複数の音源位置のうちの 1 つの音源位置に対応する指定箇所の指示を受け付けた場合、前記指向性演算部は、前記指定箇所の音源位置に向かう方向の音声を強調した音声データを生成又は合成する、収音システム。

【請求項 4】

請求項 1 に記載の収音システムであって、
前記画像表示処理部は、前記音源位置表示として、前記音源位置に対して所定値以下の音量レベルが小さい領域を示す静音領域表示を生成して前記表示部に表示させる、収音システム。

【請求項 5】

撮像部において対象空間の映像を撮像するステップと、
複数のマイクロホンを含む収音部において音声を収音するステップと、
前記収音された音声データに基づき音源位置を算出するステップと、
前記撮像された対象空間の映像と前記音源位置を示す音源位置表示とを含む画像データを表示部に表示させるステップと、
前記表示された前記画像データにおける指定箇所の指示を受け付けるステップと、
前記音声データを基に、前記収音部から、指定された前記画像データの指定箇所に対応する位置に向かう方向に、収音の指向性を形成し、前記指向性を有する方向の音声を強調した音声データを生成又は合成するステップと、
前記指向性に関する処理後の音声データを音声再生部に出力して再生させるステップと、
、
を有する、収音方法。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、対象領域内の所望の音源からの音を収音する収音システム及び収音方法に関

10

20

30

40

50

する。

【背景技術】

【0002】

例えば、講演会場又は大講義室など、多数の聴衆者に対して講演又は講義を行う大きな会場においては、講演者又は講義者の声が会場内に広く届くように拡声するが、聴衆者の中から質問などの発言を受け付ける場合、同様に聴衆者が発する声を拡声する必要がある。このような場合、聴衆者が発する声をマイクロфонによって収音して再生し、講演者又は講義者が聴衆者の発言を容易に聞き取れるようにしたり、会場内の他の聴衆者が両者の発言内容をモニタできるようにしている。

【0003】

大きな会場内の任意の位置にいる発言者からの声を収音するためには、多数のマイクロфонを会場全体に渡って配置し、発言者の近くのマイクロфонに切り替えて収音する、或いは、マイクロфонを発言者まで移動させて収音する、などの方法が用いられる。この種の用途に利用可能な収音システムとしては、例えば特許文献1、2に開示されているものがある。特許文献1には、複数の赤外線式マイクロфонを備え、これらのマイクロфонからの赤外線信号を切り替えて受信し、複数の位置における収音が可能な赤外線受信装置が開示されている。特許文献2には、共通のネットワークにリンクされたアンテナのアレイに関連付けられたワイヤレスマイクロфонを有し、1又は複数のオーディオソースからの音を収音可能なオーディオシステムが開示されている。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2004-080121号公報

【特許文献2】特開2014-110640号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

大きな会場において、任意の位置にいる発言者からの声を収音する場合、多数のマイクロфонを配置して切り替える構成では、システムが大型化、複雑化し、多大なコストを要するという課題がある。また、マイクロфонを発言者まで移動させて収音する構成では、会場係が手持ちで運ぶなど、マイクロфонの移動に手間及び時間を要し、運用時に多くの労力がかかる課題がある。

【0006】

本発明は、上述した課題を解決するために、対象領域内の所望の音源からの音を収音する際の利便性を向上できる収音システム及び収音方法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】

【0007】

本発明は、対象空間の映像を撮像する撮像部と、前記撮像部により撮像された画像データを表示する表示部と、複数のマイクロфонを含み、前記マイクロфонを用いて音声を収音する収音部と、前記収音部により収音された音声データを再生出力する音声再生部と、前記音声データに基づき音源位置を算出する音源解析部と、前記撮像された対象空間の映像と前記音源位置を示す音源位置表示とを含む画像データを前記表示部に表示させる画像表示処理部と、前記表示部に表示された前記画像データにおける指定箇所の指示を受け付ける指示入力部と、前記音声データを基に、前記収音部から、指定された前記画像データの指定箇所に対応する位置に向かう方向に、収音の指向性を形成し、前記指向性を有する方向の音声を強調した音声データを生成又は合成する指向性演算部と、前記指向性に関する処理後の音声データを前記音声再生部に出力して再生させる音声出力部と、を備える収音システムを提供する。

【0008】

本発明は、撮像部において対象空間の映像を撮像するステップと、複数のマイクロфон

10

20

30

40

50

を含む収音部において音声を収音するステップと、前記収音された音声データに基づき音源位置を算出するステップと、前記撮像された対象空間の映像と前記音源位置を示す音源位置表示とを含む画像データを表示部に表示させるステップと、前記表示された前記画像データにおける指定箇所の指示を受け付けるステップと、前記音声データを基に、前記収音部から、指定された前記画像データの指定箇所に対応する位置に向かう方向に、収音の指向性を形成し、前記指向性を有する方向の音声を強調した音声データを生成又は合成するステップと、前記指向性に関する処理後の音声データを音声再生部に出力して再生させるステップと、を有する、収音方法を提供する。

【発明の効果】

【0009】

10

本発明によれば、対象領域内の所望の音源からの音を収音する際の利便性を向上させることができる。

【図面の簡単な説明】

【0010】

【図1】本実施形態に係る収音システムの構成を示すブロック図

【図2】カメラの外観構成の一例を示す斜視図

【図3】(A)マイクアレイの外観構成の一例を示す斜視図、(B)マイクアレイの構成の一例を示す斜視図

【図4】マイクアレイの内部構成を示すブロック図

【図5】画像表示装置の構成を示すブロック図

20

【図6】収音システムの使用形態の一例を示す模式図

【図7】収音システムの使用時の状態の一例を示す模式図

【図8】収音システムの動作手順の第1例を示すフローチャート

【図9】収音システムの動作手順の第2例を示すフローチャート

【発明を実施するための形態】

【0011】

以下、本発明に係る収音システム及び収音方法の実施形態(以下、「本実施形態」という)について、図面を参照して説明する。本実施形態の収音システムは、例えば、講演会場、大講義室、株主総会会場、議会場等の大きな室内空間を有する施設、観光バス等の乗り物の車内、或いは、その他の多数の発言者が存在し得る場所に設置される音響システムに適用される。

30

【0012】

図1は、本実施形態に係る収音システムの構成を示すブロック図である。収音システムは、カメラ10、マイクアレイ20、スピーカ30、ユーザ入力装置40、画像表示装置50を含む構成である。カメラ10と、マイクアレイ20と、スピーカ30と、ユーザ入力装置40と、画像表示装置50とは、ネットワーク60を介して相互に接続されている。

【0013】

撮像部の一例としてのカメラ10は、例えば講演会場に用いられるホール等の室内の天井に設置され、対象空間の映像を撮像する。図2は、カメラ10の外観構成の一例を示す斜視図である。図2では、室内の空間を360°の全方位に渡って撮像可能な全方位カメラの例を示している。カメラ10は、撮像レンズと、CCD(Charge Coupled Device)型イメージセンサ又はCMOS(Complementary Metal Oxide Semiconductor)型イメージセンサ等による固体撮像素子とを内蔵した撮像デバイス11を有する。カメラ10は、撮像した対象の画像データを出力し、ネットワーク60を介して画像表示装置50に送信する。

40

【0014】

収音部の一例としてのマイクアレイ20は、例えば講演会場に用いられるホール等の室内の天井に設置され、対象空間の音を収音する。図3は、マイクアレイ20の構成例を示す図であり、(A)はマイクアレイ20の外観構成の一例を示す斜視図、(B)はマイク

50

アレイ 20 の構成の一例を示す斜視図である。

【0015】

マイクアレイ 20 は、筐体 21 に複数のマイクロホン 22 が一様に配置されたマイクロホンの集合体である。図 3 (B) では、複数のマイクロホン 22 を、円盤状の筐体 21 の一面に設け、筐体 21 と同一の中心を有する小さい円状及び大きい円状の 2 個の同心円状に沿って配置した例を示している。ここで、小さな円状に沿って配置された複数のマイクロホン 22A は、互いの間隔が狭く、高い音域に適した特性を有する。また、大きな円状に沿って配置された複数のマイクロホン 22B は、直径が大きく、低い音域に適した特性を有する。マイクアレイ 20 は、各々のマイクロホン 22 を用いて、対象地点 (場所) の周囲の音を収音し、収音した音声データを出力し、ネットワーク 60 を介して画像表示装置 50 に送信する。

10

【0016】

なお、カメラ 10 とマイクアレイ 20 とは、上記例のように別体に設ける構成に限定されず、1 つの筐体にカメラ 10 及びマイクアレイ 20 を設けた一体的な構成としてもよい。

【0017】

図 4 は、マイクアレイ 20 の信号処理部の機能構成の一例を示すブロック図である。マイクアレイ 20 は、複数のマイクロホン 22、複数の増幅器 (AMP) 23、複数の A/D 変換器 24、符号化部 25、送信部 26 を含む構成であり、複数のマイクロホン 22 のそれぞれに対して、増幅器 23、A/D 変換器 24 が設けられる。増幅器 23 は、マイクロホン 22 にて収音した音声信号を増幅する。A/D 変換器 24 は、増幅器 23 の出力の音声信号をデジタルの音声データに変換する。符号化部 25 は、複数の A/D 変換器 24 の出力の音声データを符号化し、複数のマイクロホン 22 にて収音した音声データを含む符号化音声データとして出力する。送信部 26 は、符号化部 25 にて符号化した音声データをネットワーク 60 を介して画像表示装置 50 に送信する。

20

【0018】

画像表示装置 50 は、カメラ 10 にて撮像した画像データを処理する機能と、画像データを表示する機能と、マイクアレイ 20 にて収音した音声データを処理する機能とを有する。

30

【0019】

図 5 は、画像表示装置 50 の機能構成の一例を示すブロック図である。画像表示装置 50 は、画像受信部 51、画像表示処理部 52、ディスプレイ 53、音声受信部 54、音源解析部 55、ユーザ入力受信部 56、指向性演算部 57、音声出力部 58 を含む構成である。画像データ処理機能を実現する構成要素として、画像受信部 51、画像表示処理部 52 を有する。画像データ表示機能を実現する構成要素として、ディスプレイ 53 を有する。音声データ処理機能を実現する構成要素として、音声受信部 54、音源解析部 55、指向性演算部 57、音声出力部 58 を有する。

【0020】

画像受信部 51 は、カメラ 10 にて撮像した画像データを受信し、画像表示処理部 52 に出力する。音声受信部 54 は、マイクアレイ 20 にて収音した音声データを受信し、音源解析部 55 及び指向性演算部 57 に出力する。ユーザ入力受信部 56 は、ユーザ入力装置 40 から送られるユーザ指示入力を受信し、指向性演算部 57 に出力する。音源解析部 55 は、マイクアレイ 20 にて収音した音声データに基づいて音が到来する方向を解析して音源の位置を算出し、音源位置データを画像表示処理部 52 に出力する。この際、音源解析部 55 は、マイクアレイ 20 の各マイクロホン 22 にて収音した複数の音声データの位相差から音源の方向及び音源までの距離を推定し、音源位置を求める。

40

【0021】

画像表示処理部 52 は、画像データと音源位置データとに基づき、対象空間の映像を示す対象空間画像を生成する。この対象空間画像は、対象空間における音源位置を示す音源位置表示を含む。表示部の一例としてのディスプレイ 53 は、例えば液晶表示パネル等の

50

表示装置により構成され、画像表示処理部52から出力される対象空間画像データを入力し、対象空間画像を表示する。本実施形態では、音源位置表示を対象空間画像中に表示することによって、収音した音声データにおける音源の位置を視覚化する。

【0022】

指向性演算部57は、マイクアレイ20にて収音した音声データに対して、特定方向への強調処理を行って収音の指向性を形成する。この際、指向性演算部57は、ユーザ指示入力に基づき、マイクアレイ20の各マイクロホン22の位置からユーザ指示がなされた指定箇所に向かう方向の音声（音量レベル）を強調（増幅）するために、ユーザの指定方向への収音の指向性を形成した音声データを生成又は合成する。音声出力部58は、指向性を形成した音声データをアナログの音声再生信号に変換して出力する。

10

【0023】

音声再生部の一例としてのスピーカ30は、画像表示装置50の近傍に配置され、画像表示装置50から出力される指向性に関する処理後の音声再生信号を入力し、画像表示装置50を使用するユーザとしての講演者又は講義者等に対して、音声再生信号を音声出力する。これにより、画像表示装置50のユーザは、指定位置の音を強調した音声を聴取できる。なお、スピーカ30は、カメラ10及びマイクアレイ20により撮像及び収音する対象空間内など、ホール等の会場の室内の任意の位置に配置し、聴衆者等の画像表示装置50のユーザ以外の人が聴取できるようにしてもよい。また、音声再生部として、スピーカだけでなく、音声信号に関する各種処理を行う音声処理装置、音声信号を増幅するアンプなど、他の音声再生装置を備えていてもよい。

20

【0024】

指示入力部の一例としてのユーザ入力装置40は、例えばディスプレイ53の画面に対応して配置され、ユーザの指又はスタイルスペン等によって入力操作が可能なタッチパネル又はタッチパッドを用いて構成される。ユーザ入力装置40は、ユーザの操作に応じて、音声データの音量レベルの強調を所望する1つ以上の指定箇所の座標データをユーザ指示入力として画像表示装置50に出力する。なお、ユーザ入力装置40は、マウス又はキーボード等のポインティングデバイスを用いて構成されてもよい。

20

【0025】

ここで、本実施形態における指向性に関する処理（以下、「指向性処理」と記載）について説明する。本実施形態では、例えば遅延和方式を用いた指向性処理によって、ユーザ指示がなされた指定箇所の方向への音声の強調処理を行う。指向性演算部57は、遅延器及び加算器を有し、各マイクロホン22にて収音した音声データについて、それぞれ所定の遅延時間を与えて加算する。具体的には、指向性演算部57は、それぞれの音声データに対して遅延器によって各マイクロホン22の位置における音波の到達時間差に対応する遅延時間を与えて位相を揃えた後、加算器によって遅延処理後の音声データを加算する。このとき、指向性演算部57は、それぞれの遅延器の遅延時間を変更することによって、任意の方向の音量レベルを強調した音声データを生成することができる。

30

【0026】

図6は、本実施形態の収音システムの使用形態の一例を示す模式図である。図6では、本実施形態の収音システムを講演会場において使用する場合の例を示している。この場合、収音システムのユーザとして、例えば講演者が使用する形態が想定される。講演会場に用いられるホール等の室内の天井75には、1台のカメラ10と1台のマイクアレイ20とが離れた位置に設置されている。室内の床に並べて配置された複数の席にはそれぞれ聴衆者が座っており、講演者85による講演を聴取している。そして、複数の聴衆者の中から、2人の聴衆者が発言者81、82としてほぼ同じタイミングで質問を発言した状態を想定する。

40

【0027】

この場合、講演者85が、カメラ10にて撮像され画像表示装置50に表示される対象空間画像を見ながら、画像中の所望の位置をユーザ入力装置40により指定することによって、スピーカ30から指定位置の方向への音が強調された音声が再生出力される。例え

50

ば、講演者 85 が発言者 81 の位置を指定すると、発言者 81 の方向にマイクアレイ 20 の指向性が形成され、発言者 81 からの音を強調した音声データが生成され、スピーカ 30 より再生される。このとき、画像表示装置 50 の指向性演算部 57 は、マイクアレイ 20 の各マイクロホン 22 が収音した音声データを用いて、各マイクロホン 22 の位置から指定位置に対応する音声位置に向かう指向方向の音（音量レベル）を強調（増幅）した音声データを生成する。これにより、ユーザとしての講演者 85 は、対象空間画像中の所望位置を指示するだけの簡単な操作によって、指定位置にいる発言者 81 の発言を良好に聴取できる。

【0028】

図 7 は、本実施形態の収音システムの使用時の状態の一例を示す模式図である。画像表示装置 50 のディスプレイ 53 には、カメラ 10 により複数の聴衆者が存在する領域を撮像した対象空間画像 90 が表示される。このとき、カメラ 10 にて撮像した画像データに基づいて、画像表示装置 50 の画像表示処理部 52 によって対象空間画像データが生成され、ディスプレイ 53 に出力される。

10

【0029】

また、マイクアレイ 20 により収音した音声データを用いて、画像表示装置 50 の音源解析部 55 によって音源位置が算出され、画像表示処理部 52 では、算出した音源位置データに基づいて音源位置表示の画像データが生成される。これにより、対象空間画像 90 には、音源位置を示す音源位置表示 91、92 が重畠されて表示される。音源位置表示 91、92 は、それぞれ発言者 81、82 の位置（すなわち音源位置）に対応している。ユーザとしての講演者 85 は、音源位置表示 91、92 によって音源となる発言者の位置を容易に直感的に認識可能である。

20

【0030】

ディスプレイ 53 上には、ユーザ入力装置 40 としてのタッチパネルが配置され、ユーザである講演者 85 が指又はスタイルスペン等によって対象空間画像 90 中の所望位置を指定することによって指示入力操作が可能になっている。講演者 85 は、対象空間画像 90 の音源位置表示 91、92 を見て、所望の発言者の位置を指定する。例えば、講演者 85 が図中左側の発言者 81 に対応する音源位置表示 91 の位置を指定したとする。講演者 85 による指示入力操作に対応して、対象空間画像 90 中の指定箇所の座標データがユーザ指示入力として画像表示装置 50 に入力される。そして、画像表示装置 50 の指向性演算部 57 により、マイクアレイ 20 の各マイクロホン 22 にて収音した音声データに対して、指定位置（音源位置表示 91 の位置）に対する方向への指向性処理が実行され、指定位置の方向から到来する音を強調した音声データが生成される。

30

【0031】

ユーザである講演者 85 には、スピーカ 30 から再生出力される指向性処理後の音声によって、指定位置からの音、すなわち所望の音源（発言者 81）からの音を明瞭に聴き分けることが可能である。また、スピーカ 30 を聴衆者がいる客席にも配置することにより、講演者 85 によって指定された位置からの音（発言者 81 の音声）を強調した状態で他の聴衆者が聴取できる。

40

【0032】

図 8 は、本実施形態の収音システムの動作手順の第 1 例を示すフローチャートである。ここでは、画像表示装置 50 の動作を中心に説明する。画像表示装置 50 は、カメラ 10 にて撮像した画像データ（カメラ画像）を画像受信部 51 にて受信する（ステップ S11）。そして、画像表示装置 50 は、画像表示処理部 52 にて画像データを処理し、ディスプレイ 53 に対象空間画像 90 を表示する（ステップ S12）。

【0033】

また、画像表示装置 50 は、マイクアレイ 20 にて収音した音声データ（マイク音声）を音声受信部 54 にて受信する（ステップ S13）。そして、画像表示装置 50 は、音源解析部 55 によって音声データから音源位置を解析し、音源位置データを算出する（ステップ S14）。続いて、画像表示装置 50 は、画像表示処理部 52 によって音源位置データ

50

タに基づいて音源位置を示す画像データを生成し、音源位置表示 91、92 を対象空間画像 90 に表示する（ステップ S15）。なお、音源位置を表示する場合、対象空間画像 90 として、音源位置表示 91、92 の近傍を拡大した（ズームした）画像を表示してもよい。

【0034】

次に、画像表示装置 50 は、ユーザ入力受信部 56 におけるユーザ指示入力の受信を監視し、ユーザ入力を検出する（ステップ S16）。このユーザ入力検出処理において、ユーザ指示入力の受信によるユーザ入力の有無を判定する（ステップ S17）。ステップ S17 においてユーザ入力有りと判定された場合、画像表示装置 50 は、指向性演算部 57 によって指向性処理を行い、ユーザ指示入力に対応する指定位置の方向への指向性を形成し、この指定方向に強調処理した音声データを生成する（ステップ S18）。そして、画像表示装置 50 は、音声出力部 58 により指向性処理後の音声データを出力し、スピーカ 30 より指定箇所からの音を強調した音声再生信号を音声出力させる（ステップ S19）。

。

【0035】

このように、本実施形態では、カメラ 10 で撮像した対象空間画像において、マイクアレイ 20 で収音した音声の音源位置を可視化できる。また、可視化した音源位置表示を目安にして、ユーザが聴取したい所望の音源位置を指定し、指定方向に指向性を形成した音声データを再生できる。

【0036】

図 9 は、本実施形態の収音システムの動作手順の第 2 例を示すフローチャートである。第 2 例は、ユーザによる指定箇所が実際の音源位置とずれている場合の動作を示すものである。

【0037】

ステップ S11～S17までの動作は上述した第 1 例と同様であり、カメラ 10 にて撮像した画像データに基づく対象空間画像 90 の表示、マイクアレイ 20 にて収音した音声データに基づく音源解析及び音源位置表示 91、92 の表示が実行される。

【0038】

ステップ S17 においてユーザ入力有りと判定された場合、画像表示装置 50 は、指向性演算部 57 によって指向性処理を行い、ユーザ指示入力に対応する指定位置の方向への指向性を形成し、この指定方向に強調処理した音声データを生成する（ステップ S18）。次に、画像表示装置 50 は、指向性演算部 57 による指向性処理後の音声データの音量レベルを参照し、指定方向の音量が所定のしきい値以上かどうかを判定する（ステップ S21）。

【0039】

ステップ S21 において指定方向の音量がしきい値未満である場合、画像表示装置 50 は、音源位置指定の再入力を促す通知を行う（ステップ S22）。通知としては、例えば、画像表示処理部 52 によって表示用のメッセージの画像データを生成し、対象空間画像 90 上に再入力を促すメッセージを表示する。なお、メッセージの表示に限らず、音声による通知など、他の各種通知手段を用いることが可能である。そして、ステップ S16、S17 に戻り、ユーザ入力検出処理及びユーザ入力判定処理を実行する。

【0040】

ステップ S21 において指定方向の音量がしきい値以上である場合、画像表示装置 50 は、音声出力部 58 により指向性処理後の音声データを出力し、スピーカ 30 より指定位置からの音を強調した音声再生信号を音声出力させる（ステップ S19）。

【0041】

第 2 例では、指定方向に指向性を形成した音声データの音量レベルが小さい場合に、音源位置指定を再度実行させ、指定方向に指向性を形成した音声データを良好に再生できる。なお、ユーザによる指定位置の入力を再度実行させる代わりに、指定位置を少しづつ移動させて音声データの音量レベルが大きくなる方向に指向性を調整し、指向性処理後の音

10

20

30

40

50

声データとして所定値以上の音量レベルが得られるようにする機能を持たせることも可能である。

【0042】

従来では、講演会等の会場において、任意の位置にいる発言者からの声を収音する場合、マイクロホンを発言者まで移動させて収音するなど、マイクロホンの移動に手間及び時間を要し、運用時に多くの労力がかかる課題があった。また、マイクロホンを移動させる時にマイクロホンが破損したり、マイクロホンが壊されてしまう課題も生じていた。これに対し、本実施形態では、撮像した対象空間画像中に音源位置を視覚化して表示し、講演者等のユーザが音源位置表示を確認して所望の発言者の位置を指定することにより、収音した音声データにおいて、所望の指定位置からの音を強調した音声を再生できる。これによって、対象領域内の所望の発言者からの音を簡便に収音して再生でき、ユーザの利便性を向上できる。

10

【0043】

次に、本実施形態の他の適用例を示す。本実施形態の収音システムは、観光バスの車内など、乗り物の室内に設けることも可能である。この場合、収音システムのユーザとして、例えばバスガイドが使用する形態が想定される。本例では、天井等に配置したカメラによって車内全体を撮像し、ユーザであるバスガイドの近傍に配置したディスプレイに対象空間画像を表示する。また、天井等に配置したマイクアレイによって車内の音を収音し、収音した音声データを基に音源位置を算出して、ディスプレイの対象空間画像中に音源位置表示を表示する。

20

【0044】

ここで、ユーザであるバスガイドが所望の位置を指定すると、指定箇所を示すユーザ指示入力に応じて、指定箇所の位置（音源位置など）に向かう方向に指向性を形成し、指向方向の音声を強調した音声データを生成する。そして、指向性処理後の音声データを車内に配置したスピーカに出力し、指定箇所の方向に指向性を持たせた収音音声を再生する。これにより、ユーザであるバスガイドが対象空間画像中の所望位置を指示するだけの簡単な操作によって、指定した位置にいる乗客が発した音声を明瞭に再生して聴取できる。また、スピーカを車内の任意の位置に複数箇所配置することにより、車内にいるバスガイド及び乗客が指定位置からの音を強調した状態で聴取できる。

30

【0045】

従来では、観光バスの車内において収音して再生する場合、マイクロホンを発言者まで手渡しで移動させて収音するなど、マイクロホンの移動に手間及び時間を要し、不便であるという課題があった。また、マイクロホンを移動させる時に落下などによりマイクロホンが破損する課題もあった。これに対し、本実施形態では、撮像した対象空間画像中に音源位置を視覚化して表示し、バスガイド等のユーザが音源位置表示を確認して所望の発言者の位置を指定することにより、収音した車内の音声データにおいて、所望の指定位置からの音を強調した音声を再生できる。これによって、対象領域内の所望の発言者からの音を簡便に収音して再生でき、ユーザの利便性を向上できる。

【0046】

なお、ディスプレイに表示した対象空間画像中に音源位置表示を表示する場合、音源自身の位置を示す音源位置表示に限定されない。音源位置表示の他の例として、音源位置に対して音量レベルが所定値以下の小さい領域を示す静音領域表示を生成し、対象空間画像中に表示するものであってもよい。対象空間内の注目すべき場所としては、音源位置はもちろん、音源位置とは反対に誰も発言していない場所、静かな場所などを表示して、ユーザに認識させることも可能である。この場合、例えば音源位置を含む音量レベルが所定値以上の領域を除外し、消去法によって静音領域を算出可能である。また、時間軸上の音量レベルの統計値に基づいて静音領域を算出することも可能である。本例によれば、音量レベルが小さい領域を視覚化でき、ユーザに容易に認識させることができる。

40

【0047】

以上のように、本実施形態の収音システムは、対象空間の映像を撮像するカメラと

50

、カメラ10により撮像された画像データを表示するディスプレイ53と、複数のマイクロホン22を含み、マイクロホン22を用いて音声を収音するマイクアレイ20と、マイクアレイ20により収音された音声データを再生出力するスピーカ30と、音声データに基づき音源位置を算出する音源解析部55と、撮像された対象空間画像90と音源位置を示す音源位置表示91、92とを含む画像データをディスプレイ53に表示させる画像表示処理部52と、ディスプレイ53に表示された画像データにおける指定箇所の指示を受け付けるユーザ入力装置40と、音声データを基に、マイクアレイ20から指定箇所に対応する位置に向かう方向に、収音の指向性を形成し、指向性を有する方向の音声を強調した音声データを生成又は合成する指向性演算部57と、指向性処理後の音声データをスピーカ30に出力して再生させる音声出力部58と、を備える。

10

【0048】

このように、撮像した対象空間の映像とともに音源位置を示す音源位置表示を表示することによって、音源位置を可視化でき、ユーザが容易に音源の位置を認識できる。また、指定箇所の指示入力に応じて収音の指向性を形成し、指向性を有する方向の音声を強調した音声データを生成又は合成することによって、ユーザが所望の音源位置を指定し、指定方向に指向性を形成した音声データを再生できる。これによって、対象領域内の所望の位置の音源からの音を簡便に収音して再生でき、ユーザの利便性を向上できる。

【0049】

また、本実施形態の収音システムでは、音源解析部55により算出された音源位置が複数ある場合に、画像表示処理部52は、複数の音源位置のそれぞれに対応する音源位置表示91、92を生成してディスプレイ53に表示させる。

20

【0050】

これにより、ユーザが音源位置表示によって複数の音源のそれぞれの音源位置を容易に認識できる。

【0051】

また、本実施形態の収音システムでは、ユーザ入力装置40により複数の音源位置のうちの1つの音源位置に対応する指定箇所の指示を受け付けた場合、指向性演算部57は、指定箇所の音源位置に向かう方向の音声を強調した音声データを生成又は合成する。

30

【0052】

これにより、ユーザが指定した指定箇所に対応する方向への指向性を形成し、指定された音源位置からの音を明瞭に再生できる。

【0053】

また、本実施形態の収音システムでは、画像表示処理部52は、音源位置表示として、音源位置に対して所定値以下の音量レベルが小さい領域を示す静音領域表示を生成してディスプレイ53に表示させる。

【0054】

これにより、対象空間において音量レベルが小さい領域を視覚化でき、ユーザに容易に認識させることができる。

【0055】

また、本実施形態の収音方法は、カメラ10において対象空間の映像を撮像するステップと、複数のマイクロホン22を含むマイクアレイ20において音声を収音するステップと、収音された音声データに基づき音源位置を算出するステップと、撮像された対象空間画像90と音源位置を示す音源位置表示91、92とを含む画像データをディスプレイ53に表示させるステップと、表示された画像データにおける指定箇所の指示を受け付けるステップと、音声データを基に、マイクアレイ20から指定箇所に対応する位置に向かう方向に、収音の指向性を形成し、指向性を有する方向の音声を強調した音声データを生成又は合成するステップと、指向性処理後の音声データをスピーカ30に出力して再生させるステップと、を有する。

40

【0056】

以上、図面を参照しながら各種の実施形態について説明したが、本発明はかかる例に限

50

定されないことは言うまでもない。当業者であれば、特許請求の範囲に記載された範疇内において、各種の変更例又は修正例に想到し得ることは明らかであり、それらについても当然に本発明の技術的範囲に属するものと了解される。また、本発明の趣旨を逸脱しない範囲において、上記実施形態における各構成要素を任意に組み合わせてもよい。

【産業上の利用可能性】

【0057】

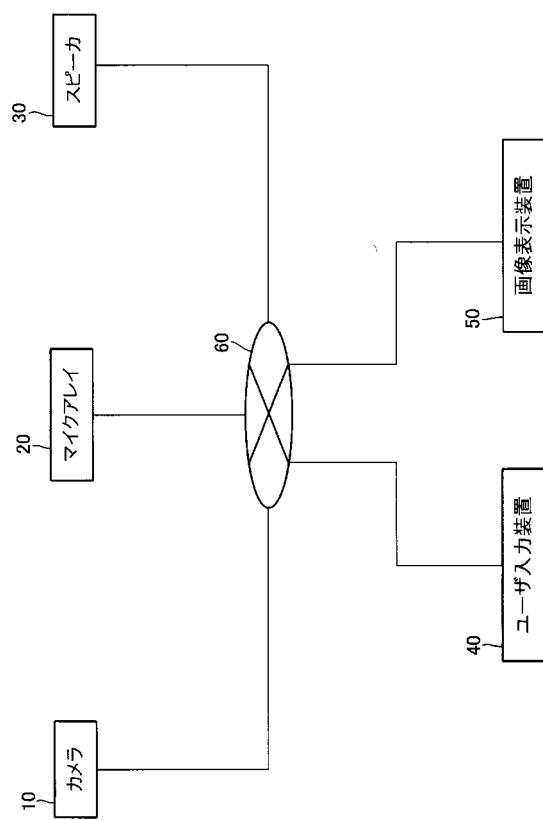
本発明は、対象領域内の所望の音源からの音を収音する際の利便性を向上できる収音システム及び収音方法等として有用である。

【符号の説明】

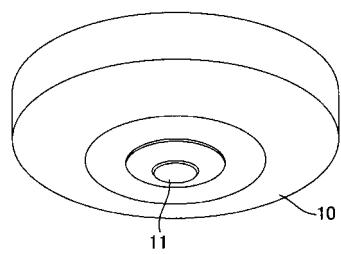
【0058】

- | | | |
|-----------------|-----------|----|
| 1 0 | カメラ | 10 |
| 1 1 | 撮像デバイス | |
| 2 0 | マイクアレイ | |
| 2 2、2 2 A、2 2 B | マイクロホン | |
| 2 3 | 増幅器 | |
| 2 4 | A / D 変換器 | |
| 2 5 | 符号化部 | |
| 2 6 | 送信部 | |
| 3 0 | スピーカ | |
| 4 0 | ユーザ入力装置 | 20 |
| 5 0 | 画像表示装置 | |
| 5 1 | 画像受信部 | |
| 5 2 | 画像表示処理部 | |
| 5 3 | ディスプレイ | |
| 5 4 | 音声受信部 | |
| 5 5 | 音源解析部 | |
| 5 6 | ユーザ入力受信部 | |
| 5 7 | 指向性演算部 | |
| 5 8 | 音声出力部 | |
| 6 0 | ネットワーク | 30 |
| 8 1、8 2 | 発言者 | |
| 8 5 | 講演者 | |
| 9 0 | 対象空間画像 | |
| 9 1、9 2 | 音源位置表示 | |

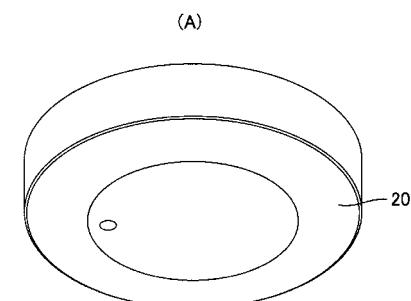
【図1】



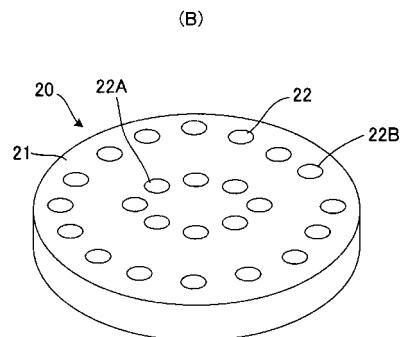
【図2】



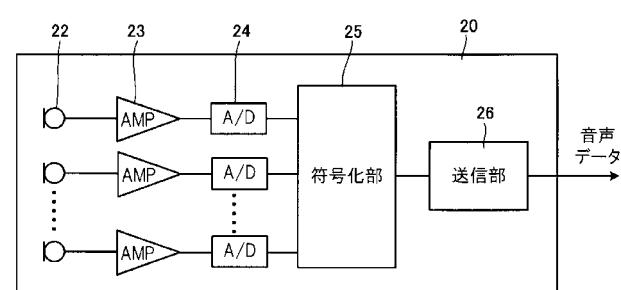
【図3】



(A)

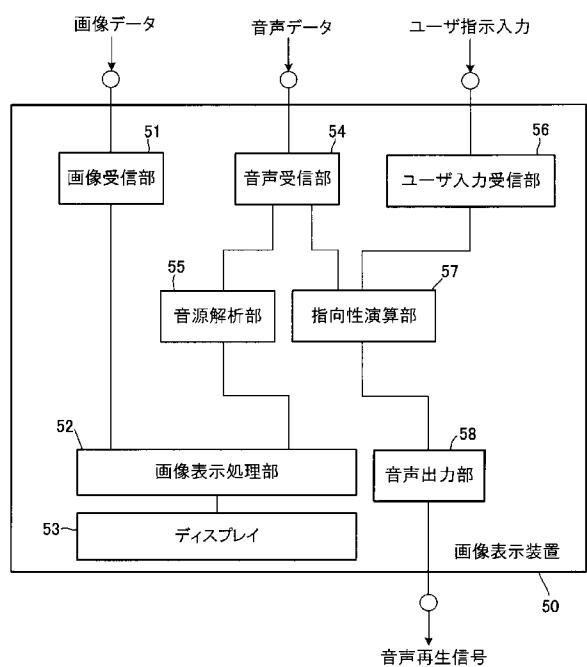


(B)

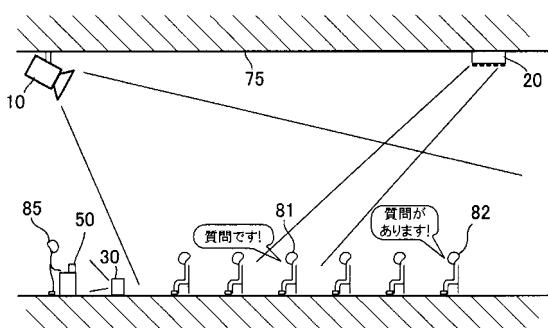


【図4】

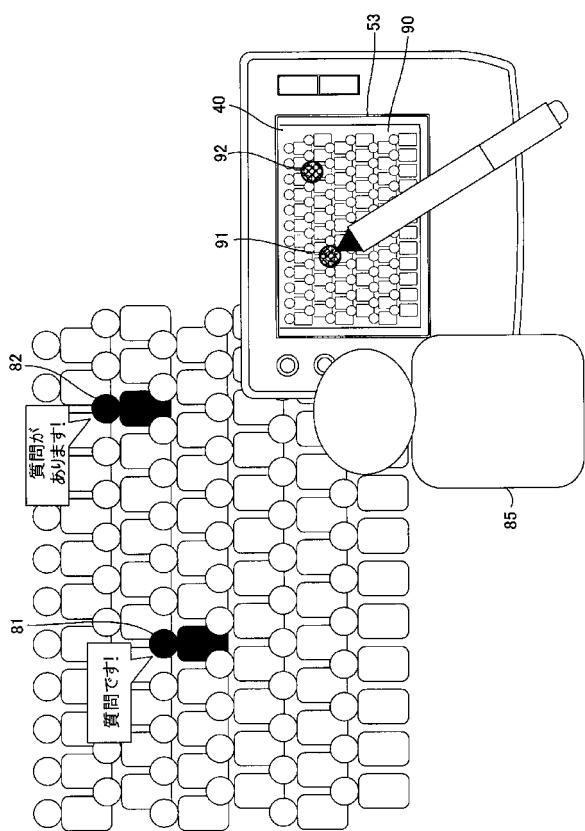
【図5】



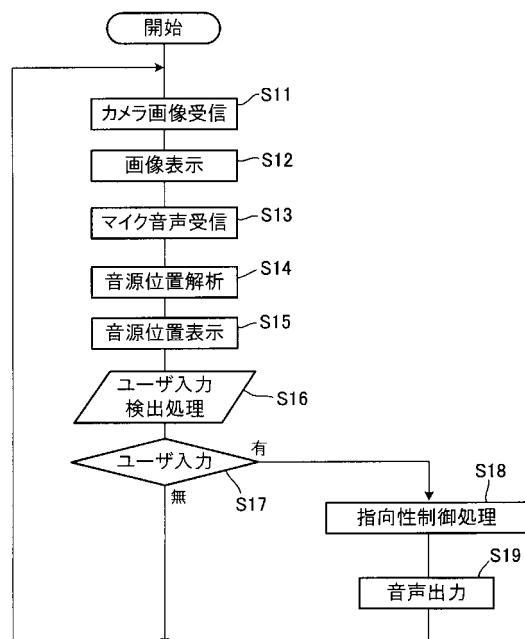
【図6】



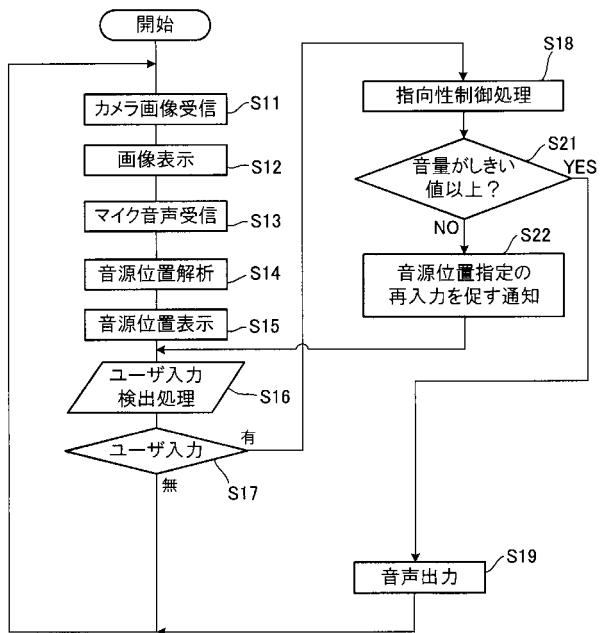
【図7】



【図8】



【図9】



フロントページの続き

F ターム(参考) 5J083 AA05 AB09 AC31 AD02 AE07 AF01 AG05 BC11 CA10 CA12
EB03