

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4867579号
(P4867579)

(45) 発行日 平成24年2月1日(2012.2.1)

(24) 登録日 平成23年11月25日(2011.11.25)

(51) Int.Cl.		F I			
HO4R	3/00	(2006.01)	HO4R	3/00	320
HO4R	1/40	(2006.01)	HO4R	1/40	320A
HO4M	3/56	(2006.01)	HO4R	1/40	310
			HO4M	3/56	Z

請求項の数 2 (全 15 頁)

(21) 出願番号	特願2006-294683 (P2006-294683)	(73) 特許権者	000004075 ヤマハ株式会社 静岡県浜松市中区中沢町10番1号
(22) 出願日	平成18年10月30日(2006.10.30)	(74) 代理人	110000970 特許業務法人 楓国際特許事務所
(65) 公開番号	特開2007-151103 (P2007-151103A)	(74) 代理人	100084548 弁理士 小森 久夫
(43) 公開日	平成19年6月14日(2007.6.14)	(72) 発明者	田中 良 静岡県浜松市中沢町10番1号 ヤマハ株式会社内
審査請求日	平成21年8月20日(2009.8.20)	(72) 発明者	田丸 卓也 静岡県浜松市中沢町10番1号 ヤマハ株式会社内
(31) 優先権主張番号	特願2005-320042 (P2005-320042)		
(32) 優先日	平成17年11月2日(2005.11.2)		
(33) 優先権主張国	日本国(JP)		

最終頁に続く

(54) 【発明の名称】 遠隔会議装置

(57) 【特許請求の範囲】

【請求項1】

音声を受音し、受音信号として出力する複数のマイクをアレイ状に配置したマイクアレイと、

前記複数のマイクから出力した受音信号を遅延加算することにより複数の受音エリアへ指向させた受音ビームをそれぞれ形成する受音ビーム形成部と、

前記複数の受音ビームのうち最大音量を示す受音ビームに対応する受音エリアを位置情報として検出する位置情報検出手段と、

前記マイクの受音信号の出力と前記位置情報とを送信する送信手段と、

複数のスピーカをアレイ状に配置したアレイスピーカと、

外部より音声信号および位置情報を受信する受信部と、

前記受信した位置情報に基づき決定される位置を仮想音源位置とする放音ビームが形成されるよう前記受信した音声信号を処理して前記複数のスピーカに供給する信号処理部と

を備え、

前記位置情報検出手段は、前記最大音量となる受音ビームの受音エリアを検出した後、該受音エリアをさらに細分した複数の細分化受音エリアへ指向させた取得用受音ビームを各々形成し、前記取得用受音ビームの音量の大きい順に複数選択した取得用受音ビームに対応する細分化受音エリアに基づいて前記位置情報を検出する遠隔会議装置。

【請求項2】

前記位置情報検出手段は、前記取得用收音ビームの音量の大きい順に複数選択した取得用收音ビームに対応する細分化收音エリアの間を、前記選択した取得用收音ビームの強弱に応じた比例配分により、前記位置情報を検出すると共に、

前記選択した複数の取得用收音ビームの出力を前記比例配分により合成し、前記マイクの收音信号の出力として前記送信手段に送信させる信号加算手段を備えた請求項1に記載の遠隔会議装置。

【発明の詳細な説明】

【技術分野】

【0001】

10 音声を位置情報と共に送信して、受信した音声とその音場を再現する遠隔会議装置に関する。

【背景技術】

【0002】

従来、送信側の音声を受信して、送信側の音声の音場を再現する遠隔会議装置が提案されている（特許文献1～2参照。）。

【0003】

特許文献1には、以下の装置が記載されている。マイクによって話者の音声をとらえ、マイクから得られる話者情報によって話者位置情報を形成し、この話者位置情報を音声情報と共に多重化して伝送させる。また、受信側では送られて来る話者位置情報により鳴動させるスピーカの位置を切り替え、話者の声と位置を受信側で再現する。

20

【0004】

また、特許文献2では、複数のマイクアレイで受け取った音声情報を送信して、これと同数のスピーカアレイに出力することにより、送信元の音場を再現する立体音声情報の創出方法等が開示されている。

【特許文献1】特開平9-261351号公報

【特許文献2】特開平2-114799号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

30 しかしながら、特許文献1では、再生する音声を音像定位させる方法が複数のスピーカの音量バランスのみで調整することによるため、正確な音像定位を行うことが困難であった。

【0006】

また、特許文献2では、複数のマイクアレイで受け取った音場情報をそのまま出力するので、自然な音声を出力することができる反面、アレイスピーカの数と同数のチャンネル数の信号を遠隔地に伝送しなければならないため、多くの伝送回線リソースを消費してしまう欠点があった。

【0007】

40 そこで、本発明は、送信先の音場を再現する装置において、受信した音源位置情報に基づいて音声信号をスピーカアレイに出力することにより、伝送回線リソースの消費を抑えつつ、送信先の音場を正確に再現することを目的とする。

【課題を解決するための手段】

【0008】

本発明は、上述の課題を解決するための手段を以下のように構成している。

【0009】

(1) 本発明は、

音声を收音し、收音信号として出力する複数のマイクをアレイ状に配置したマイクアレイと、

前記複数のマイクから出力した收音信号を遅延加算することにより複数の收音エリアへ指向させた收音ビームをそれぞれ形成する收音ビーム形成部と、

50

前記複数の收音ビームのうち最大音量を示す收音ビームに対応する收音エリアを位置情報として検出する位置情報検出手段と、

前記マイクの收音信号の出力と前記位置情報とを送信する送信手段と、

複数のスピーカをアレイ状に配置したアレイスピーカと、

外部より音声信号および位置情報を受信する受信部と、

前記受信した位置情報に基づき決定される位置を仮想音源位置とする放音ビームが形成されるよう前記受信した音声信号を処理して前記複数のスピーカに供給する信号処理部と、

を備え、

前記位置情報検出手段は、前記最大音量となる收音ビームの收音エリアを検出した後、該收音エリアをさらに細分した複数の細分化收音エリアへ指向させた取得用收音ビームを各々形成し、前記取得用收音ビームの音量の大きい順に複数選択した取得用收音ビームに対応する細分化收音エリアに基づいて前記位置情報を検出することを特徴とする。

10

【0010】

本発明の装置は、2台以上の遠隔会議装置を通信回線やネットワークを介して相互に接続するための、1台分の機能を備えている。即ち、受信側ユニットと送信側ユニットとを備えている。

送信側ユニットでは、音源位置情報検出手段は音源の位置を検出すると共に、マイクはその音源の音声信号を收音し、この音声を電気信号に変換する。

【0011】

20

音源位置情報検出手段は、複数の位置へ指向させた收音ビームを前記複数のマイクの收音信号から形成し、その複数ある收音ビームの最大音量から收音ビームの收音エリアの位置を話者の位置情報として検出する。送信部は、前記アレイマイクで得られた音声信号と前記位置情報とを送信する。

【0012】

受信側ユニットは、受信部が相手装置から音源の音声信号と前記音源の位置情報とを受信する。信号処理部は、前記スピーカユニットの後方に、この受信した音源の位置情報によって決定される位置を仮想音源位置とする放音ビームが形成されるよう、前記受信した音声信号を処理して前記スピーカユニットに供給する。例えば、この仮想音源の座標は、装置中央を原点とし、背面方向をY軸とする座標軸平面上とすることができる。

30

以上、2台の遠隔会議装置のうちの一について説明したが、逆方向への音声信号・位置関係の收音・放音も同様である。

【0013】

本発明のこれらの構成により、会議参加者に対して、遠隔会議装置をはさんだ反対側に設けた相手装置側の会議室の仮想的な音源と対話しているかのような、リアルな位置関係で対話することができる。また、前記音源の位置情報を、前記音源の音声信号と共に受信するから、特許文献2のように、多くの伝送回線リソースを消費してしまう欠点を解消できる。また、特許文献2のように、スピーカが送信側と受信側で同じ数に限定する必要がない。

【0014】

40

なお、送信手段が送信する前記マイクの收音信号の出力は、マイクの收音信号であればよく、音源が複数のマイク音声信号のうちいずれか1つ、または複数を合成するもの（例えば単純加算したもの）でも良いし、また複数のマイクからの收音信号を遅延加算した收音ビームを出力するものでも良いし、收音ビームを複数合成するものでも良い。

【0016】

この構成では、位置情報検出手段は、細分化收音エリアへ指向させた取得用收音ビームを形成し、その收音ビームの音量が大きいものを選択するから、位置情報としてより細かいエリアを探索できる。また、この構成では、初めから細かいエリアそれぞれに対する收音ビームを形成して選択するのではなくて、2段階に絞り込んで選択している。したがって、形成すべき收音ビームの数を少なくすることができるから、位置情報検出手段に用い

50

るハードウェアを簡素化できる。

【0017】

(3)本発明は、

前記位置情報検出手段は、前記取得用収音ビームの音量の大きい順に複数選択した取得用収音ビームに対応する細分化収音エリアの間を、前記選択した取得用収音ビームの強弱に応じた比例配分により、前記位置情報を検出すると共に、

前記選択した複数の取得用収音ビームの出力を前記比例配分により合成し、前記マイクの収音信号の出力として前記送信手段に送信させる信号加算手段を備えたことを特徴とする。

【0018】

この構成では、位置情報検出手段は、選択した細分化収音エリアの間を、前記選択した取得用収音ビームの強弱に対応した比例配分により位置情報を検出するので、細分化収音エリアの中間部分を補間でき、より精度の高い位置情報を得ることができる。また、信号処理部は前記選択した複数の取得用収音ビームの出力を前記比例配分に基づいて合成するので、各収音ビームの収音エリアの中間部分を補間できる。また、このように常にバランスをとって補間することで、音源が移動した場合にも、滑らかに収音ビームの切り替えをすることができる。

【発明の効果】

【0019】

本発明によれば、会議参加者に対して、遠隔会議装置をはさんだ反対側に設けた相手装置側の会議室の仮想的な音源と対話しているかのような、リアルな位置関係で対話をすることができる。

【発明を実施するための最良の形態】

【0020】

図1を用いて、本実施形態の音声通信に用いる遠隔会議装置について説明する。図1は、本実施形態の音声通信に用いる遠隔会議装置の外観図(A)と、概略機能図(B)である。

【0021】

まず、図1(A)を用いて、遠隔会議装置1の構成を説明する。図1(A)に示すように、遠隔会議装置1は、外観上、複数のスピーカ SP_i ($i = 1 \sim N$ 、 N は整数)からなるスピーカアレイと、複数のマイク M_i ($i = 1 \sim N$)からなるマイクアレイとを備えており、例えば、机101の上に設置されるものである。遠隔地に設置された2台の遠隔会議装置と通信回線で接続して、相互に音声通信を行う。

【0022】

なお、スピーカアレイを構成する複数のスピーカ SP_i ($i = 1 \sim N$)は、それぞれ独立して音声信号を扱うことができる音声処理系統に接続され、独立に音声出力が可能となっており、マイクアレイを構成する複数のマイク M_i ($i = 1 \sim N$)は、それぞれ独立に音声出力をデジタルデータに変換可能である。なお、座標系 X 、 Y 、 Z についての説明は後述する。

【0023】

次に、図1(B)を用いて、遠隔会議装置1の機能の概略を説明する。図1(B)の左側の図に示すように、音声を送信する場合には、送信側の話者102Aは、遠隔会議装置1に向かって語りかける。そうすると、送信側ユニット1Aでは、マイクアレイを構成する複数のマイク M_i ($i = 1 \sim N$)がそれぞれ音声を収音するとともに、これら収音された音声に基づいて送信側の話者102Aの位置を検出し、収音した音声信号とその位置を示す位置情報とを送信する。この位置情報は、収音した電気信号それぞれにディレイを設定して合成した(即ち遅延加算した)収音ビームを複数パターン形成し、そのパターンの強弱から計算する。この位置情報の座標系は、例えば、図1(A)、図1(B)に示すように、 XYZ 軸座標系で与えることができる。具体的には、 XYZ 軸座標系は、遠隔会議装置1の受音面に対して、上下方向が Z 軸、左右方向が X 軸、この受音面に垂直な奥行き

10

20

30

40

50

方向をY軸方向とする。また、X軸、Z軸方向の代わりに、この受信面の中心からの距離Rと角度の極座標で与えることも可能である。

【0024】

また、図1(B)の右側の図に示すように、遠隔会議装置1の受信側では、この送信された音声信号と位置情報とを基に、受信側ユニット1Bは、仮想音源VSの位置を計算し、これに基づいて、各スピーカSPiに出力する信号へのディレイ量をそれぞれ算出し、送信側ユニット1A側で話された音声にディレイを付与してスピーカSPiに出力する。これにより、送信側ユニット1A側で話された音声の音場を正確に再現して、受信側の聴取者102Bに、送信側の音場を伝える。ここで、仮想音源VSは、受信側ユニット1Bに話しかける受信側の聴取者102Bと対向する位置に設定するのが自然であるから、受信側の聴取者102Bから見て左右方向のX軸座標及び奥行き方向のY軸座標は、実際に送信側ユニット1Aで得られた値と、受信側ユニット1Bで仮想的な位置を設定する座標の値が正負逆になるようにする。前述の極座標を用いる場合も同様である。

10

【0025】

なお、以上の図1の説明では、送信側ユニット1Aと受信側ユニット1Bを概念的に区別するため別の符号を用いて記述したが、実際には、本実施形態の遠隔会議装置1は、双方向通信が可能であり、受信側の聴取者102Bも、送信側の話者102Aに音声を受信側ユニット1Bのマイクアレイを通じて送信することができる。すなわち、各遠隔会議装置は、図1(B)の送信側ユニット1Aと受信側ユニット1Bを一体に備えたものである。

20

また、本実施形態では、遠隔会議装置1が伝達する音声として、人声を例に挙げているが、本実施形態の装置1の対象は、人声に限らない。例えば、音楽等であっても良い。以下、同じである。

【0026】

ここで、図5(A)を参照して送信側ユニット1Aで行なう話者の位置の探索方法について簡単に説明する(図5の詳細は後述する。)。遠隔会議装置1内部の検出用ビーム形成部(後述図2の検出用ビーム形成部21に相当)は、図5(A)に示すように、このディレイパターンそれぞれに従い、送信側の話者102Aの位置として想定する、複数の收音エリア111~114へ指向させた收音ビームを各々形成する。

【0027】

この收音ビームは、マイクMi(i=1~N)それぞれからの距離差によるディレイ量差をその特定の位置に合うように調整してデジタル音声信号を重ね合わせることにより形成する。これにより、その特定の位置の音声は互いに強められて入力されると共に、この特定の位置以外から来る音声は互いに打ち消されることになり、この特定の位置の音声への指向性を有することになる。

30

【0028】

この收音エリア111~114と送信側の話者102Aの位置とが合致していれば、これらの收音ビームの中から、音量が最も大きな收音ビームが得られることになる。そこで、遠隔会議装置1内部の検出用ビーム形成部は、收音エリア111~114に指向させた4個の收音ビームを同時に形成し、音量の大きい方向を探索する。また、遠隔会議装置1内部の取得用ビーム形成部(後述図2の取得用ビーム形成部22に相当し、検出用ビーム形成部と同様の回路構成を持つ)は、このようにして得られた音量の大きい方向(図5ではエリア114)を更に分割する細分化エリア131~134を設定する。また、この取得用ビーム形成部は、これらの細分化エリア131~134に指向させた收音ビームを各々形成する。遠隔会議装置1は、この細分化エリア131~134に向けた收音ビームから音量の大きい收音ビームを選択する。このように2段階で探索することにより、話者102Aの位置を正確且つ迅速に探索できる。

40

【0029】

このビーム形成部は、常時稼動するか、または、遠隔会議装置1が所定の間隔ごと、例えば0.5秒おきにビーム位置算出部24が位置検出を行い、そのときにビーム位置算出

50

部 2 4 から指示を受けて、稼動するようにする。

【 0 0 3 0 】

次に、図 2 を用いて、遠隔会議装置 1 の内部構成について説明する。図 2 は、この内部構成を表すブロック図である。

まず、図 2 (A) を用いて、送信側ユニット 1 A の構成について説明する。送信側ユニット 1 A は、收音した音声をデジタルで加工するために、マイク M i と、A / D 変換器 2 9 とを備える。また、話者が話している位置を推定し、位置情報を制御信号 2 4 1 として出力することができるよう、話者がいると想定されるマイク前方の複数の位置へ指向させた收音ビームを形成する検出用ビーム形成部 2 1 と、B P F 2 3 (バンドパスフィルタ) とビーム位置算出部 2 4 を備える。また、ビーム位置算出部 2 4 で検出した位置の近傍をさらに詳しく探索し、その位置へ指向させた收音ビームを形成することができるよう、取得用ビーム形成部 2 2 と信号選択部 2 5 と信号加算部 2 6 を備える。また、位置情報 2 7 A と收音ビームを多重化するために多重化部 2 8 を備える。以下それぞれの構成を説明する。

10

【 0 0 3 1 】

図 2 (A) のマイク M i (i = 1 ~ N) は、図 1 に示すように N 個のマイクであり、これらのマイク M i により送信側の話者 1 0 2 A の音声を音声信号に変換する。

A / D 変換器 2 9 は、A / D 変換用の I C で構成でき、マイク M i により取得した音声信号をデジタル音声信号に変換して、検出用ビーム形成部 2 1 、取得用ビーム形成部 2 2 に送る。

20

【 0 0 3 2 】

検出用ビーム形成部 2 1 は、話者がいると想定されるエリアへ指向させた收音ビームを複数同時に形成する。そのため、例えば、ディレイ処理を行うためのリングバッファ用 R A M またはそれに相当するものを備えて、マイク M i 各々で收音した音声信号のディレイ量を調整する。また、このディレイ処理を行うため、リングバッファ用 R A M の書き込み、読み出しを制御するプログラムおよびディレイ制御用データを格納した R O M を備える。またビーム形成部 2 1 はこのプログラムを実行させるための計算部を備えている。遠隔会議装置 1 は、遠隔会議装置 1 の信号処理を動作させる D S P を設けており、検出用ビーム形成部 2 1 の計算部はその一機能として構成する。

【 0 0 3 3 】

そこで、図 2 (A) の検出用ビーム形成部 2 1 は、前述したように送信側ユニット 1 A の前にある話者の位置を複数想定して收音ビームの收音エリアを設定し、その位置へ指向させた收音ビームを形成する。

30

【 0 0 3 4 】

なお、以下の説明では、検出用ビーム形成部 2 1 で形成する收音ビームと取得用ビーム形成部 2 2 で形成する收音ビームを区別するため、後者の收音ビームを取得用ビームと称する。

【 0 0 3 5 】

図 2 (A) の検出用ビーム形成部 2 1 は、粗い範囲の探索をするために收音エリア (図 5 (A) の 1 1 1 ~ 1 1 4 に相当) へ指向させた收音ビームを形成する。

40

取得用ビーム形成部 2 2 は、検出用ビーム形成部 2 1 と同様の構成を備えている。ただし、取得用ビーム形成部 2 2 は、検出用ビーム形成部 2 1 の收音ビームを解析するビーム位置算出部 2 4 の計算結果に基づき、さらに細かい範囲に收音エリア (図 5 (A) の收音エリア 1 1 4 の中の細分化收音エリア 1 3 1 ~ 1 3 4) を設定する。取得用ビーム形成部 2 2 は、細分化收音エリア 1 3 1 ~ 1 3 4 に指向させた取得用ビームを各々形成し、信号選択部 2 5 に出力する。このように、検出用ビーム形成部 2 1 、取得用ビーム形成部 2 2 で段階的に細かく收音ビーム、取得用ビームを形成し、送信側の話者 1 0 2 A の位置を探索しているので、初めから細かく探索するよりも、ハードウェアの構成を簡略化できると共に、計算速度が速くなり、送信側の話者 1 0 2 A の位置の変動に対する応答を速くすることができる。

50

図2(A)のBPF23は、検出用ビーム形成部21の收音ビームMB1~MB4に対し、送信側の話者102Aの検出に必要な音声帯域以外の帯域をカットするフィルタを積み込み演算し、收音ビームMB'1~MB'4を出力する。これにより計算量を削減することができ、ビーム位置算出部24が行なう探索を高速化できる。なお、BPF23では、このような帯域をカットしても、前述の検出用ビーム形成部21では、粗い收音エリア111~114を設定しているので問題ない。

【0036】

ビーム位置算出部24は、BPF23で出力した收音ビームMB'1~MB'4のうちから最も音量が大きいパターンを選択する。これにより、送信側の話者102Aの位置を特定できる。ビーム位置算出部24は、当該特定された位置情報を制御信号241として取得用ビーム形成部22に出力する。具体的には、送信側の話者102Aがいると想定される複数の位置のパターンについて予めIDコードを設けておき、ビーム位置算出部24は、收音ビームの音量が大きい位置のIDコードを取得用ビーム形成部22に出力する。

10

【0037】

なお、このビーム位置算出部24における音量の計算は、デジタル時系列音声データをFFT変換して、複数の特定周波数ごとのゲインの2乗を足し合わせることで代用することができる。また、ビーム位置算出部24は、所定間隔ごと、例えば0.5秒ごとに動作する。これにより、ビーム位置算出部24は、話者102Aの移動を制御信号241として検出できる。

【0038】

20

図2(A)の取得用ビーム形成部22は、ビーム位置算出部24からの制御信号241を受けて、検出用ビーム形成部21で合成した收音ビームのうちから最も音量が大きい收音ビームに対応する收音エリアをさらに細分する複数(本実施例では4つ)の收音エリアへ指向させた取得用ビームを形成する。

また、信号選択部25は、取得用ビーム形成部22が出力した取得用ビームのうち音量が大きい2つの取得用ビーム(図5の細分化收音エリア133、134に相当)を選択し、これらの取得用ビームの音量に応じて比例配分した位置を音源の位置102Aとして決定し、これを位置情報27Aとして多重化部28に出力する。

【0039】

図2(A)の信号加算部26は、信号選択部25が出力した2つの取得用ビームを音量に応じて比例配分して合成し、この合成したデジタルの音声信号27Bを多重化部28に送る。このように常に2つ以上の取得用ビームを合成することにより、話者が移動したり、入れ替わったりしたときも音声スムーズに移動するようにすることができる。

30

【0040】

図2(A)の多重化部28は、信号選択部25で生成した位置情報27Aと、信号加算部26で生成した音声信号27Bを多重化して、受信側ユニット1Bに送信20を行う。

【0041】

次に図2(B)を用いて、受信側ユニット1Bの内部構成について説明する。受信側ユニット1Bは、受信部31と、パラメータ算出部32と、仮想音源生成信号処理部33と、DAC34i(i=1~N)と、AMP35i(i=1~N)とを備え、外部のスピーカSPi(i=1~N)を接続している。以下、それぞれの構成を説明する。

40

【0042】

図2(B)の受信部31は、送信側ユニット1Aから、位置情報27Aと、音声信号27Bについて受信信号30の受信を行う。この音声信号27Bは、仮想音源生成信号処理部33に送られると共に、位置情報27Aは、パラメータ算出部32に送られる。

パラメータ算出部32は、位置情報27Aを基に仮想音源VS(図1(B)参照。)を設定して、仮想音源VSの位置から各スピーカSPiの位置までの距離36i(i=1~N)を計算する。そして、この距離36iを音速で除算することにより、ディレイ量のパラメータを設定する。

【0043】

50

図2(B)の仮想音源生成信号処理部33、DAC34*i* (*i* = 1 ~ *N*)、AMP35*i* (*i* = 1 ~ *N*)について説明する。

仮想音源生成信号処理部33は、パラメータ算出部32で設定するパラメータに基づいて、受信部31で受信した音声、即ち、送信側ユニット1Aの多重化部28で生成した音声信号27Bを、スピーカSP*i* (*i* = 1 ~ *N*)の出力系統ごとに信号処理し、DAC34*i* (*i* = 1 ~ *N*)にそれぞれ出力する。受信部31に inputs する音声信号27Bは1系統であるが、音声信号27Bに対しパラメータ算出部32のパラメータ(ディレイ量設定)に基づいて、スピーカSP*i*の出力系統ごとに、遅延処理を行う。これにより、仮想音源VSを焦点とする放音ビームをスピーカSP*i*から出力することができ、あたかも、受信側ユニット1Bの前方に仮想音源があるような音像定位を実現できる。

10

【0044】

DAC34*i*は、仮想音源生成信号処理部33で得られた音声デジタルデータをアナログ信号に変換して出力する。

AMP35*i* (*i* = 1 ~ *N*)は、DAC34*i* (*i* = 1 ~ *N*)の出力をそれぞれ増幅して、スピーカSP*i* (*i* = 1 ~ *N*)に出力する。

【0045】

ここで、送信側ユニット1Aの信号選択部25による絞り込み探索について、補足説明をする。信号選択部25は、例えば取得用ビーム形成部22が、特定時間、話者が静止していると判断した場合にのみ行なっても良い。処理能力の問題上、この絞り込み探索をする時間的余裕がない場合があるからである。この場合には、音声信号27Bの生成を以下(A)~(B)のように処理する。

20

(A)音源の位置が、粗く設定した收音エリア111~114の1つに特定時間留まる場合には、取得用ビーム形成部22は、制御信号241が所定期間一定であることを検知し、ビーム位置算出部24で選択した收音エリアのさらに内側の細分化收音エリア(図5(A)の131~134に相当)へ指向させた複数の取得用ビームを形成する。

(B)音源の位置が、粗く設定した收音エリア111~114間で移動した場合には、取得用ビーム形成部22は、制御信号241の変動を検知して、取得用ビーム形成部22に検出用ビーム形成部21と同様の複数の粗い範囲を設定した收音エリア111~114へ指向させた取得用ビームを形成する。この場合の取得用ビームは、上記收音ビームと同様に粗い範囲の收音エリアへ指向する。

30

【0046】

これら(A)、(B)いずれの場合においても、信号選択部25は、取得用ビーム形成部22から出力された複数の取得用ビームの中で、音量が大きい2つを選択して信号加算部26に出力し、信号加算部26はこれらの取得用ビームを音量に応じて比例配分して合成する。

また、図2(A)の信号加算部26は、音源が大きく移動する場合に音量の大きい2つを選択して単純に合成するのではなく、切り替え前の音声をフェードアウトすると共に切り替え後の音声をフェードインして、音声をクロスフェードするようにしてもよい。このように音源が大きく移動する場合、音声を自然に接続できる。このクロスフェードを行う時間中は、送信側ユニット1A、受信側ユニット1B間で、收音エリア111~114の補間位置と、クロスフェード開始時、終了時を送受信する。パラメータ算出部32は、仮想音源の位置VSについても、このクロスフェード時間内で次第に移動するようにし、この移動に基づいてパラメータを設定する。これにより音像の移動が自然になる。

40

【0047】

なお、音源の位置が粗く設定した收音エリア111~114間で移動した場合は、取得用ビーム形成部22に検出用ビーム形成部21と同様の複数の粗い範囲を設定するが、この場合、信号選択部25の出力する位置情報27Aは、ビーム位置算出部24の出力する位置情報である制御情報241と同等になる。したがって、信号選択部25が出力する位置情報27Aの代わりに、ビーム位置算出部24にて得られる制御情報241を、多重化部28に27C(図2(A)参照)として出力しても良い。

50

【0048】

図3を用いて、図2のパラメータ算出部32におけるパラメータの設定方法について、具体的に説明する。図3は、この設定方法についての概念図である。

まず、図3(A)で示すように、仮想音源VSと受信側ユニット1Bの各スピーカSP_iとの距離36_iを計算する。次に、図3(B)で示すように、パラメータ算出部32のパラメータを設定する。例えば、SP₁について、距離36₁を音速Vで除算した値とする。

【0049】

以上、図2、図3の説明で説明したように、伝送元のマイクM_iの個数N個分の音声系統を伝送しなくとも、受信側ユニット1Bは、送信側ユニット1Aの音声信号27Bと、位置情報27Aのみに基づいて、スピーカSP_i($i = 1 \sim N$)上に、音場を正確に再現できる。

10

【0050】

<送信側装置の音声の收音の実施例>

以下、図4、図5を参照して、図2(A)で説明した検出用ビーム形成部21についての收音の実施例についてさらに具体的に説明を行う。

【0051】

図4を用いて、検出用ビーム形成部21が、收音エリア111~114へ指向させた收音ビームを形成する方法について説明する。図4では、X₁~X_Nの入力にディレイを付与して加算する方法を説明するため、仮想の計算テーブル2111~2114を示している。形成すべき收音ビームMB₁~MB₄は、X₁~X_Nの入力に所定のディレイを付与して合成したものである。検出用ビーム形成部21の図示しないROM内には、この收音エリアに收音ビームを指向させるよう、ディレイパターンのデータ2121~2124が、收音エリア111~114に対応して記録されている。検出用ビーム形成部21は、ディレイパターンのデータ2121~2124に基づいてディレイD_j_i($j = 1 \sim 4, i = 1 \sim N$)にディレイ量を設定する。マイク入力合成部2141~2144は、ディレイX₁~X_Nの入力をD_j₁~D_j_Nに通した信号を、計算テーブル2111~2114ごとに加算して、收音ビームMB₁~MB₄として出力する。

20

【0052】

なお、以上では説明上、收音ビームMB₁~MB₄毎に計算テーブル2111~2114を示したが、実装上は、D₁₁~D₄₁をまとめて1つのリングバッファ用のメモリ等で構成するのが簡便である。このリングバッファには、X₁を入力し、D₁₁~D₄₁の出力タップを設ける。D₁_i~D₄_i($i = 2 \sim N$)も同様に、それぞれ1つのX_iを入力するリングバッファを構成し、D₁_i~D₄_iの出力タップを設ける。また、実装上、ディレイパターンのデータ2121~2124を、收音ビームMB₁~MB₄毎でなく、入力X_i($i = 1 \sim N$)毎に用意するのが好ましい。

30

【0053】

次に、図5(A)を参照して音源位置の検出方法について説明する。検出用ビーム形成部21が、複数の收音エリア111~114へ指向させた收音ビームを同時に形成する。ビーム位置算出部24は、各收音ビームの音量の大小を比較して、この收音エリアのいずれかの領域に話者がいるかを監視する。

40

【0054】

図5(A)において、各マイクM_i($i = 1 \sim N$)が收音した音声信号を適当なディレイ時間だけを遅延させて合成することにより、收音エリア111~114へ指向させた收音ビームを形成するが、以下、各マイクに設定するディレイ時間(すなわち、各マイクが收音した音声を遅延させるディレイ時間)の設定方法について説明する。ここでは、收音エリア111に収束する收音ビームを例に挙げて説明する。図5(A)に示すように、收音エリア111の中心からマイクM_i($i = 1 \sim N$)までの距離121_i($i = 1 \sim N$)を計算する。この距離を音速(340m/秒)で除算して收音エリア111から各マイクまでの音波の伝播時間を計算する。計算した伝搬時間のうち、最長のものを基準時間とし

50

て、この基準時間と各マイクまでの伝播時間との差を、そのマイクに設定するディレイ時間とする。このディレイ時間が検出用ビーム形成部 2 1 に設定される。

【 0 0 5 5 】

このように設定されたディレイ時間で各マイクが收音した音声信号を遅延させて合成することにより、收音エリア 1 1 1 から各マイクへ伝搬した音声を、位相をそろえて合成することができ、收音エリア 1 1 1 の音声を高いゲインで出力することができる。一方、他のエリアから伝搬した音声は、位相がずれた状態で合成されるため、振幅が相殺されて低ゲインとなる。

【 0 0 5 6 】

以上では收音エリア 1 1 1 に収束する收音ビームについて説明したが、收音エリア 1 1 2、1 1 3、1 1 4 に収束する收音ビームについても同様である。

【 0 0 5 7 】

次に、図 5 (B) を用いて、図 2 (A) で示したビーム位置算出部 2 4 の算出方法について、より具体的に説明する。以下の説明では、送信側の話者が図 5 (A) に示す 1 0 2 A (收音エリア 1 1 4 の中の細分化收音エリア 1 3 3 内) の位置に存在すると仮定する。また、音源発生信号の波形が符号 2 4 0 で示した形状と仮定する。

まず、検出用ビーム形成部 2 1 は、收音エリア 1 1 1 ~ 1 1 4 を設定して、送信側の話者 1 0 2 A の位置を探索している。図 5 (A) の 1 0 2 A に話者がいるとすれば、この收音エリア 1 1 4 へ指向させた收音ビーム M B ' 4 の音量は最も大きくなる。他方、この收音エリア 1 1 4 から遠ざかる收音エリア 1 1 3、1 1 2、1 1 1 へ指向させた收音ビームは、波形 2 4 3 波形 2 4 2 波形 2 4 1 に示すように、音量は小さくなる。そこで、ビーム位置算出部 2 4 は、図 2 に示す收音ビーム M B ' 1 ~ M B ' 4 のうちで音量が最大のものに対応する收音エリア (図 5 では 1 1 4) を選択する。

また、信号選択部 2 5 は、ビーム位置算出部 2 4 と同様にして、音量の最も大きな取得用ビーム出力を選択して出力する。ただし、信号選択部 2 5 は、取得用ビーム形成部 2 2 が設定した細分化收音エリア 1 3 1 ~ 1 3 4 の内で音量が最大のものから順に 2 つの取得用ビームを選択する。この取得用ビームは、図 5 (A) では、細分化收音エリア 1 3 3、1 3 4 に対応するビームとなる。

【 0 0 5 8 】

以上で説明した構成により、それぞれ、図 5 (A) に示すような、收音エリア 1 1 j の方向へ指向させた收音ビーム M B j を形成することができる。図 2 (A) の取得用ビーム形成部 2 2 も、これと同様の構成をとることができる。

【 0 0 5 9 】

< 具体的使用形態と音場再現例 >

図 6 を用いて、本実施形態の装置の具体的使用形態と音場再現例について説明する。図 6 は、この具体的使用形態を表す図である。前述の図 1 の説明のとおり遠隔会議装置 1 は、通信回線やネットワークを介して相互に接続するための、1 台分の機能を備えており、これと同等の 2 台またはそれ以上の遠隔会議装置を、ネットワークを介して相互に接続するようにする。

【 0 0 6 0 】

図 6 (A)、(B) に示す 2 台の送信側ユニット 1 A、受信側ユニット 1 B は、音声信号と共に位置情報を送受信するので、互いに同じスピーカの数 $S P i$ 、あるいはマイク $M i$ を用意しなくとも送信側の話者 1 0 2 A の音場を伝送できるが、以下では、同一の構成の遠隔会議装置 1 を用いるものとして説明する。また、簡単のため、以下の説明では X Y 平面の 2 次元座標を用いることとし、送信側ユニット 1 A、受信側ユニット 1 B の中心の位置を図 6 に示すようにそれぞれ原点の座標とする。

【 0 0 6 1 】

図 6 (A) に示すように、送信側の話者 1 0 2 A の X、Y 座標が $(X 1, - Y 1)$ であったとする。なお、この実施形態では、話者から見て、送信側ユニット 1 A、受信側ユニット 1 B の右側が X プラス方向 (すなわち、図 6 (A) では紙面の左方向、図 6 (B) で

10

20

30

40

50

は紙面の右方向がXプラス方向)としている。遠隔会議装置1の送信側ユニット1Aは、送信側の話者102Aから発せられ、マイクMiで得られた音声信号を解析して、送信側の話者102Aの位置情報を得る。この位置情報を得る方法は、前述した図4、図5の説明のようにマイクMi(i=1~N)の音声信号の解析により行うことができる。図6の以下の説明では、この位置情報を得た結果、送信側の話者102Aは、図6(A)に示すように、送信側の話者102AのX軸のプラスの位置にいと仮定して説明する。

【0062】

図6(B)に示すように、受信側の聴取者102Bから見れば、送信側の話者102Aは向かい合って会話しているようにするのが自然であるから、送信側の話者102Aの代わりとなる仮想音源VSを設定すべき位置のX座標の値は、原点より左の-X1であり、また、Y座標の値は、奥行き側後方であるからY1となる。結果的に仮想音源VSを設定すべき位置の座標は、(-X1,Y1)となる。そして、前述の図3(B)で説明したようにして、図2(B)のパラメータ算出部32は、この仮想音源VSの位置に基づいて、スピーカSpi(i=1~N)ごとにそのディレイ量、音量の各パラメータを計算する。そして、図2に示すように、仮想音源生成信号処理部33においてこれを設定して、DAC34i(i=1~N)、AMP35i(i=1~N)を経てスピーカSpi(i=1~N)から音声として出力する。

【0063】

以上の実施形態の説明について、補足する。

以上の説明では、説明の容易のため、送信側ユニット1Aは送信を行うものとして説明したが、送信側ユニット1A、1Bは送信のみ、または受信のみ行うものではない。遠隔会議装置1は、双方向通信を行う機能を備えており、受信側の聴取者102Bが話しかけた音声と受信側の聴取者102Bの位置情報は、受信側ユニット1Bを用いて取得され、ネットワークを通じて送信側ユニット1Aに送信される。その機能は、以上の説明と同様である。

【0064】

また、以上の説明では、送信側の話者102Aは1人であり、1つの音源として説明したが、複数人いても良い。この場合には、位置情報を複数用意して、それぞれの位置での音量を検出して伝送する。受信側ユニットは、上述の図3で示したようなパラメータ設定を話者それぞれについて行い、これに基づいて算出したスピーカSpi(i=1~N)の出力をユニットごとに加算する。

【0065】

また、図2、図5では、同時に出力する收音ビーム、取得ビーム、收音エリアの数をそれぞれ4つとしたが、いずれもそれ以上であっても良く、それ以下であっても良い。同時に形成する收音ビーム、取得ビームの数は4つに限定されない。また、同時でなく、時分割で收音ビームを出力して、音量を比較するのであれば、收音ビーム、取得用ビームは2つ以上でなく、1つでも足りる。

【0066】

また、話者が複数ある場合を想定する場合は、このような段階的な位置の絞込みをするときに、少なくとも最初の段階では、話者の位置を1つに絞り込むのではなく、複数の候補を用意してそれぞれの近傍をさらに探索しても良い。

【0067】

また、以上の説明では、スピーカSpi、マイクMi(i=1~N)の数はそれぞれ同じNとしたが、必ずしも同じにする必要はない。伝送経路では、特許文献2のようなものと異なり、音源の位置情報を用いて伝送しているからこれらの数が異なっても、送信元の音場を再現できる。

【0068】

また、本実施形態では、検出用ビーム形成部21、取得用ビーム形成部22により2段階に收音ビーム、取得用ビームを設定し、送信側の話者102Aを探索したが、3段階以上にエリアを絞り込む設定をして、探索しても良い。

【 0 0 6 9 】

また、取得用ビーム形成部 2 2 を設けなくても、ビーム位置算出部 2 4 だけで位置情報 2 7 C を検出し出力することも可能である。ビーム位置算出部 2 4 単独でも収音ビームを比較すれば収音エリア 1 1 1 ~ 1 1 4 のいずれに音源 1 0 2 A が含まれるかを検出できる。また、ビーム位置算出部 2 4 は、必ずしも比例配分により求めなくても、単に最大の音量となる取得用ビームを検出して、これに対応する位置情報 2 7 C を検出することができる。

【 0 0 7 0 】

さらに、以上の説明では、図 2 の送信側ユニット 1 A では、音声信号 2 7 B として、収音ビームまたは取得用ビームを送信するとしたが、単にマイク M 1 ~ M N のいずれかの出力を送信しても良い。また、マイク M 1 ~ M N のいずれかまたは全部を足し合わせたものを送信しても良い。その他、マイク M 1 ~ M N が出力する収音信号を加算等して用いたものであればどのような音声信号でも良い。このようにしても、受信側ユニット 1 B は、位置情報 2 7 A または 2 7 C に基づいて伝送回線リソースの消費を抑えつつ、送信先の音場を正確に再現することができる。ただし、話者 1 0 2 A に向けた収音ビームまたは取得用ビームを送信する方が、遠隔会議装置 1 の耐ノイズ性能が向上する。

【 図面の簡単な説明 】

【 0 0 7 1 】

【 図 1 】 本実施形態の音声通信に用いる遠隔会議装置の外観図と、概略機能図である。

【 図 2 】 本実施形態の音声通信に用いる遠隔会議装置の内部構成を表すブロック図である。

【 図 3 】 パラメータ算出部 3 2 におけるパラメータの設定方法についての概念図である。

【 図 4 】 本実施形態の音声通信に用いる遠隔会議装置の検出用ビーム形成部の内部構成図を示す。

【 図 5 】 話者の位置を特定するための説明図である。

【 図 6 】 具体的使用形態と音場再現例を説明する図である。

【 符号の説明 】

【 0 0 7 2 】

- 1 - 遠隔会議装置、 1 A - 送信側ユニット、 1 B - 受信側ユニット、
- 1 0 1 - 机、 1 0 2 A - 送信側の話者、 1 0 2 B - 受信側の聴取者、
- 1 1 1 ~ 1 1 4 収音エリア、 1 2 j i (j = 1 ~ 4 , i = 1 ~ N) - 距離、
- 1 3 1 ~ 1 3 4 - 細分化収音エリア、 S P i (i = 1 ~ N) - スピーカ、
- M i (i = 1 ~ N) - マイク、 M B j (j = 1 ~ 4) - 収音ビーム、
- V S - 仮想音源、 2 0 - 送信、 2 1 - 検出用ビーム形成部、
- 2 1 1 j (j = 1 ~ 4) - 計算テーブル、
- 2 1 2 j (j = 1 ~ 4) - ディレイパターンのデータ、
- D j i (j = 1 ~ 4 , i = 1 ~ N) - ディレイ、
- 2 1 4 j (j = 1 ~ 4) - マイク入力合成部、
- 2 2 - 取得用ビーム形成部、 2 3 - B P F、 2 4 - ビーム位置算出部、
- 2 4 1 - 制御信号、 2 5 - 信号選択部、 2 6 - 信号加算部、
- 2 7 A - 位置情報、 2 7 B - 音声信号、 2 8 - 多重化部、 2 9 - A / D 変換器、
- 3 0 - 受信信号、 3 1 - 受信部、 3 2 - パラメータ算出部、
- 3 3 - 仮想音源生成信号処理部、 3 4 i (i = 1 ~ N) - D A C、
- 3 5 i (i = 1 ~ N) - A M P、 3 6 i (i = 1 ~ N) - 距離

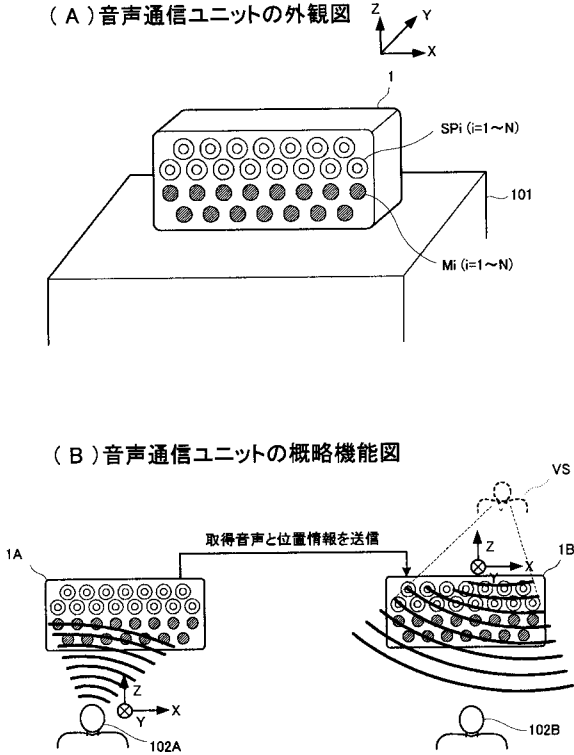
10

20

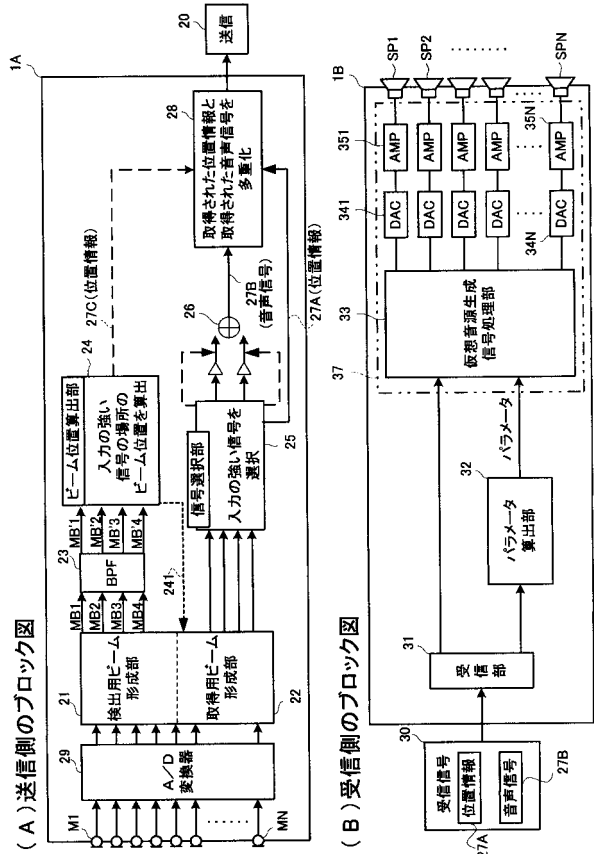
30

40

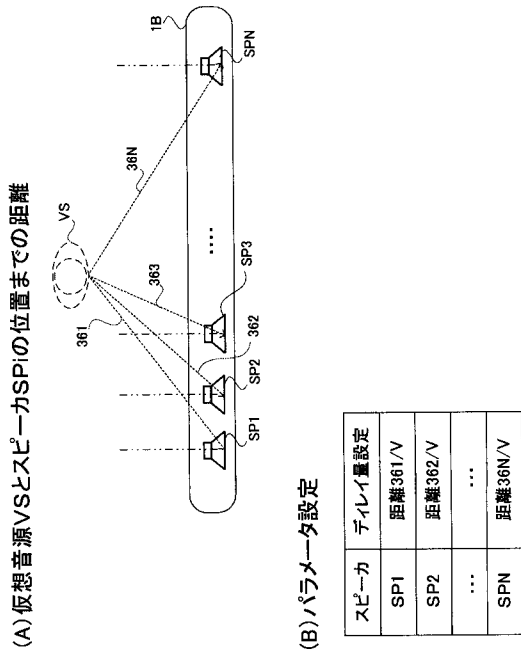
【図1】



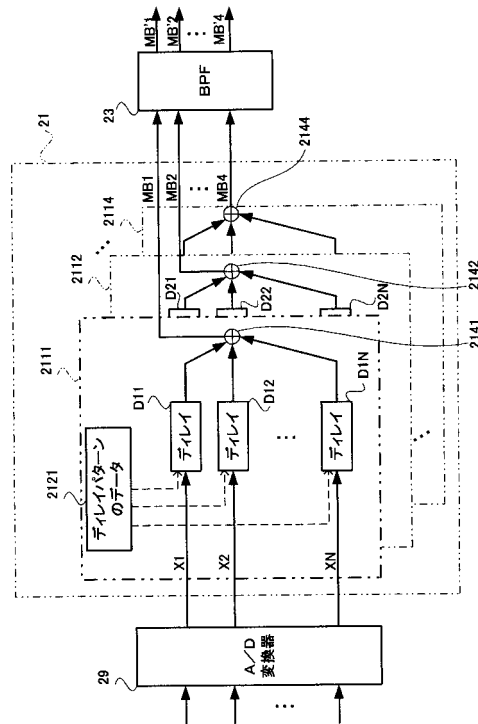
【図2】



【図3】

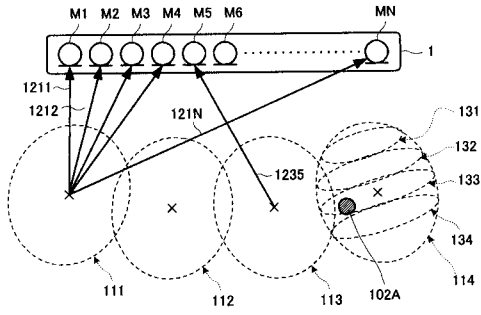


【図4】



【図5】

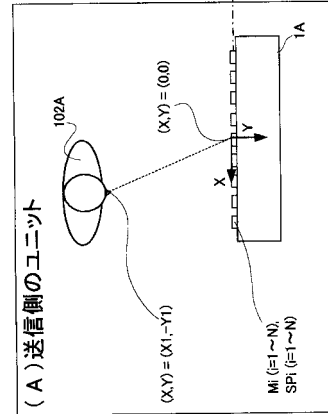
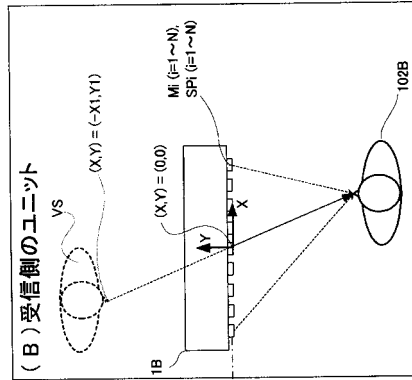
(A) アレイマイクと音源の想定位置との位置関係



(B) 音源発生音声信号とマイクビームとの関係

音源102A 発生音声信号		240
収音ビームMB'1 (収音エリア111 に対応)		241
収音ビームMB'2 (収音エリア112 に対応)		242
収音ビームMB'3 (収音エリア113 に対応)		243
収音ビームMB'4 (収音エリア114 に対応)		244

【図6】



フロントページの続き

(72)発明者 刑部 勝一

静岡県浜松市中沢町10番1号

ヤマハ株式会社内

審査官 境 周一

(56)参考文献 特開平08-505745(JP,A)
特開平09-121400(JP,A)
特開平09-261351(JP,A)
特開平05-316239(JP,A)
特開平11-018192(JP,A)
特開2003-270034(JP,A)
特開2005-197896(JP,A)
特開平09-261792(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04R 1/00-31/00
H04M 1/00-1/82
H04M 3/56