

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4191518号
(P4191518)

(45) 発行日 平成20年12月3日(2008.12.3)

(24) 登録日 平成20年9月26日(2008.9.26)

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

請求項の数 16 (全 15 頁)

(21) 出願番号	特願2003-86679 (P2003-86679)	(73) 特許権者	390019839
(22) 出願日	平成15年3月27日(2003.3.27)		三星電子株式会社
(65) 公開番号	特開2003-304589 (P2003-304589A)		SAMSUNG ELECTRONICS
(43) 公開日	平成15年10月24日(2003.10.24)		CO., LTD.
審査請求日	平成18年2月10日(2006.2.10)		大韓民国京畿道水原市靈通区梅灘洞416
(31) 優先権主張番号	2002-016692		416, Maetan-dong, Yeongtong-gu, Suwon-si,
(32) 優先日	平成14年3月27日(2002.3.27)		Gyeonggi-do 442-742
(33) 優先権主張国	韓国 (KR)		(KR)
		(74) 代理人	100070150
			弁理士 伊東 忠彦
		(72) 発明者	田 先 ▲壽▼
			大韓民国 ソウル特別市 陽川区 木1洞
			925番地 新市街地アパート 705
			棟 306号
			最終頁に続く

(54) 【発明の名称】 直交円形マイクアレイシステム及びこれを用いた音源の3次元方向検出方法

(57) 【特許請求の範囲】

【請求項1】

音源の3次元方向を検出する直交円形マイクアレイシステムであって、
前記音源より音声信号を受信する指向性マイクと、
前記音源より音声信号を受信する所定数のマイクが前記指向性マイクの周りに配列された第1マイクアレイと、

前記第1マイクアレイと直交するように、前記音源より音声信号を受信する所定数のマイクが前記指向性マイクの周りに配列された第2マイクアレイと、

前記第1及び第2マイクアレイより信号を受信し、前記信号が音声信号であるか否かを識別して音源の位置を推定する方向検出部と、

前記方向検出部において推定された前記音源の位置に応じて前記第1マイクアレイ、第2マイクアレイ及び前記指向性マイクの変える回転制御部と、

前記指向性マイクより受信された音声信号と前記第1または第2マイクアレイより受信された音声信号とに対して演算を行い、演算の行われた音声信号を出力する音声信号処理部と、を備えることを特徴とする直交円形マイクアレイシステム。

【請求項2】

前記第1及び第2マイクアレイのうち少なくとも何れか一方は円形であることを特徴とする請求項1に記載の直交円形マイクアレイシステム。

【請求項3】

前記第1及び第2マイクアレイに設けられたマイクは互いに一定の間隔を保持すること

10

20

を特徴とする請求項 1 に記載の直交円形マイクアレイシステム。

【請求項 4】

前記第 1 及び第 2 マイクアレイに設けられたマイクは指向性マイクであることを特徴とする請求項 1 に記載の直交円形マイクアレイシステム。

【請求項 5】

前記第 1 マイクアレイより入力される受信信号または前記第 2 アレイマイクより入力される受信信号を前記方向検出部の制御信号に応じて選択するスイッチをさらに備えることを特徴とする請求項 1 に記載の直交円形マイクアレイシステム。

【請求項 6】

前記方向検出部は、前記第 1 及び第 2 マイクアレイより受信された信号から音声信号を識別する音声信号識別部と、

前記音声信号識別部より受信された音声信号から、前記第 1 及び第 2 マイクアレイに設けられたマイクに受信される音声信号の受信角に応じて音源の方向を推定する音源方向推定部と、

前記音源方向推定部において推定された方向に前記第 1 及び第 2 マイクアレイを回転させるように制御信号を出力する制御信号生成部と、を備えることを特徴とする請求項 1 ないし請求項 5 のうち何れか 1 項に記載の直交円形マイクアレイシステム。

【請求項 7】

前記音源方向推定部は、前記第 1 または第 2 マイクアレイに設けられたマイクに入力される所定レベル以上の音声信号の出力値を合計して周波数領域に変換した後、周波数領域に変換された前記音声信号の出力値の合計値を音声信号の前記マイクにおける受信角を変数として変換し、最高の電力値を示す角度を音源の方向として推定することを特徴とする請求項 6 に記載の直交円形マイクアレイシステム。

【請求項 8】

前記所定レベル以上の音声信号の出力値の合計値 $y(t)$ は、アレイマイクの数 M であり、 c が媒質内における音速であり、アレイの中心から前記マイクまでの距離が r であるとき、

【数 1】

$$y(t) = \sum_{n=1}^M x_n \left(t + (n-1) \frac{2r \sin\left(\frac{\pi}{M}\right) \cos\left(\theta + \frac{2\pi(n-1)}{M}\right)}{c} \right)$$

であることを特徴とする請求項 7 に記載のマイクアレイシステム。

【請求項 9】

前記音声信号処理部は、前記第 1 または第 2 マイクアレイに設けられたマイクの各々より受信された音声信号を前記各マイク間の位置差により生じた最大遅延時間まで遅延させて合計し、前記指向性マイクより受信された音声信号を前記最大遅延時間だけ遅延させて前記合計値に加えて所望の音声信号の音質を改善することを特徴とする請求項 1 ないし請求項 5 のうち何れか 1 項に記載の直交円形マイクアレイシステム。

【請求項 10】

所定数のマイクが配列された第 1 及び第 2 マイクアレイと指向性マイクとを用いて音源の 3 次元方向を検出する方法であって、

(a) 第 1 マイクアレイより入力される信号から音声信号を識別する段階と、

(b) 前記第 1 マイクアレイに設けられたマイクに音声信号が受信される角度に応じて音源の方向を推定し、前記第 1 マイクアレイと直交する前記第 2 マイクアレイに設けられたマイクが推定された方向に向かうように前記第 2 マイクアレイを回転させる段階と、

(c) 前記第 2 マイクアレイに設けられたマイクに音声信号が入力される角度に応じて前記音源の方向を推定する段階と、

10

20

30

40

50

(d) 前記指向性マイクを (b) 及び (c) 段階において推定された音源の方向に移動させて音声信号を受信し、受信された音声信号を出力する段階と、

(e) 音源の位置変化及び音源の音声発話が終了したか否かを検出する段階と、を含むことを特徴とする音源の 3 次元方向検出方法。

【請求項 1 1】

前記第 1 及び第 2 マイクアレイのうち少なくとも何れか一方は円形であることを特徴とする請求項 1 0 に記載の音源の 3 次元方向検出方法。

【請求項 1 2】

前記第 1 及び第 2 マイクアレイに設けられたマイクは互いに一定の間隔を保持することを特徴とする請求項 1 0 に記載の音源の 3 次元方向検出方法。

10

【請求項 1 3】

前記第 1 及び第 2 マイクアレイに設けられたマイクは指向性マイクであることを特徴とする請求項 1 0 に記載の音源の 3 次元方向検出方法。

【請求項 1 4】

前記 (b) 及び (c) 段階は、前記第 1 または第 2 マイクアレイに設けられたマイクに入力される所定レベル以上の音声信号の出力値を合計して周波数領域に変換した後、周波数領域に変換された前記音声信号の出力値の合計値を音声信号の前記マイクにおける受信角を変数として変換し、最高の電力値を示す角度を音源の方向として推定することを特徴とする請求項 1 0 ないし請求項 1 3 のうち何れか 1 項に記載の音源の 3 次元方向検出方法。

20

【請求項 1 5】

前記所定レベル以上の音声信号の出力値の合計値 $y(t)$ は、アレイマイクの数 M であり、 c が媒質内における音速であり、アレイの中心から前記マイクまでの距離が r であるとき、

【数 2】

$$y(t) = \sum_{n=1}^M x_n \left(t + (n-1) \frac{2r \sin\left(\frac{\pi}{M}\right) \cos\left(\theta + \frac{2\pi(n-1)}{M}\right)}{c} \right)$$

30

であることを特徴とする請求項 1 4 に記載の音源の 3 次元方向検出方法。

【請求項 1 6】

前記 (d) 段階は、前記第 1 または第 2 マイクアレイに設けられたマイクの各々より受信された音声信号を前記各マイク間の位置差により生じた最大遅延時間だけ遅延させて合計し、前記指向性マイクより受信された音声信号を前記最大遅延時間だけ遅延させて前記合計値に加えて所望の音声信号の音質を改善する段階を含むことを特徴とする請求項 1 0 ないし請求項 1 3 のうち何れか 1 項に記載の音源の 3 次元方向検出方法。

【発明の詳細な説明】

【0001】

40

【発明の属する技術分野】

本発明は音源の 3 次元方向を検出するシステム及び方法に関する。

【0002】

【従来の技術】

本発明への理解のために、以下では、本発明の方向推定の対象である音源を話者と指称して例示的に説明する。

【0003】

一般にマイクは音響信号を全ての方向で受信する特性を有する。このような特性を有するマイクは、通常、無指向性（全方向性）マイクと呼ばれ、使用されているが、受信しようとする音声信号以外に周りの雑音及び反響まで受信してしまい、所望の音声信号が歪めら

50

れるという問題がある。かかる通常のマイクの問題点を解決するために使われるのが指向性マイクである。

【0004】

指向性マイクはマイクの軸方向から一定の角度（指向角）内に受信される音響だけを受信する特性を有するので、話者が指向性マイクの指向角内でマイクに向かって話をすれば、話者の音声は周りの雑音に比べて大きくマイクを介して受信され、指向角内に存在しない雑音は受信されない。

【0005】

最近、このような指向性マイクは、遠隔会議においてよく使われる。しかし、指向性マイクを遠隔会議に使う場合、指向性マイクの特性のゆえに、話者はマイクの指向角内でマイクに向かって話さなければならないという制約がある。すなわち、話者は自分の席に座って話す場合にも設けられたマイクの指向角から外れる方向に話せず、指向角から外れて会議室内を移動しつつ話せないため、話者が自由に話せないという不具合が招かれる。

10

【0006】

上述した問題点を解決するために、マイクを一定の間隔にて配列して一定の空間内で移動する話者の音声信号を受信するマイクアレイシステムが考案されている。

【0007】

図1Aに示すように、平面形マイクアレイシステムは、受音しようとする空間の一側面に設けられて前方で移動する話者の音声を受信する。すなわち、平面形マイクアレイシステムは前方180°ほどの範囲内で移動する話者の音声を受信することができる。しかし、話者がマイクアレイシステムの後方で移動する場合には話者の音声を受信できなくなるという制限が依然として残る。

20

【0008】

平面形マイクアレイシステムの上述した制限を克服した円形マイクアレイを図1Bに示す。円形マイクアレイシステムは、マイクが設けられた平面内においてその中心から360°の範囲内で移動する話者の音声を受信することができる。しかし、円形マイクアレイは、マイクが設けられた平面をXY平面とした時、XY平面上の話者の位置は考慮するとはいえ、Z軸上の話者の位置は考慮しておらず、全ての方向から発せられる信号を受信することになり、話者の位置とは関係ないZ軸上の方向からの雑音及び反響まで受信してしまうため、音声の歪みを引き起こすという問題点が依然として存在する。

30

【0009】

【発明が解決しようとする課題】

本発明は、平面的に移動する話者の位置だけではなく、話者の3次元空間上の動きまで考慮して、話者がいかなる方向で発音しても話者の音声を受信することのできるマイクアレイシステム及び方法を提供することを課題とする。

【0010】

本発明が解決しようとする他の課題は、受信される話者の音声を最大化し、且つ、話者の音声を除いた周りの雑音及び反響を最小化して話者の音声を明確に認識させることにより、音声認識の性能を向上させることのできるマイクアレイシステム及び方法を提供することである。

40

【0011】

【課題を解決するための手段】

前記課題を達成するために、本発明は、音源の3次元方向を検出する直交円形マイクアレイシステムであって、前記音源より音声信号を受信する指向性マイクと、前記音源より音声信号を受信する所定数のマイクが前記指向性マイクの周りに配列された第1マイクアレイと、前記第1マイクアレイと直交するように、前記音源より音声信号を受信する所定数のマイクが前記指向性マイクの周りに配列された第2マイクアレイと、前記第1及び第2マイクアレイより信号を受信し、前記信号が音声信号であるか否かを識別して音源位置を推定する方向検出部と、前記方向検出部において推定された前記音源位置に応じて前記第1マイクアレイ、第2マイクアレイ及び前記指向性マイクの方法を変える回転制御部と

50

、前記指向性マイクより受信された音声信号と前記第1または第2マイクアレイより受信された音声信号とに対して演算を行い、演算の行われた音声信号を出力する音声信号処理部と、を備えることを特徴とする直交円形マイクアレイシステムを提供する。

【0012】

前記直交円形マイクアレイシステムは、好ましくは、前記方向検出部より入力される音声信号であって、前記第1マイクアレイより入力される受信信号または前記第2アレイマイクより入力される受信信号を前記方向検出部の制御信号に応じて選択するスイッチをさらに備える。また、前記方向検出部は、好ましくは、前記第1及び第2マイクアレイより受信された信号から音声信号を識別する音声信号識別部と、前記音声信号識別部より受信された音声信号から、前記第1及び第2マイクアレイに設けられたマイクに受信される音声信号の受信角に応じて音源の方向を推定する音源方向推定部と、前記音源方向推定部において推定された方向に前記第1及び第2マイクアレイを回転させるように制御信号を出力する制御信号生成部と、を備える。

10

【0013】

さらに、前記目的を達成するために、本発明は、(a)第1マイクアレイより入力される信号から音声信号を識別する段階と、(b)前記第1マイクアレイに設けられたマイクに音声信号が受信される角度に応じて音源の方向を推定し、前記第1マイクアレイと直交する前記第2マイクアレイに設けられたマイクが推定された方向に向かうように前記第2マイクアレイを回転させる段階と、(c)前記第2マイクアレイに設けられたマイクに音声信号が入力される角度に応じて前記音源の方向を推定する段階と、(d)前記超指向性マイクを(b)及び(c)段階において推定された音源の方向に移動させて音声信号を受信し、受信された音声信号を出力する段階と、(e)音源の位置変化及び音源の音声発話が終了したか否かを検出する段階と、を含む音源の3次元方向検出方法を提供する。

20

【0014】

前記直交円形マイクアレイシステム及び音源の3次元方向検出方法において、前記第1及び第2マイクアレイのうち少なくとも何れか一方は円形であり、前記第1及び第2マイクアレイに設けられたマイクは互いに一定の間隔を保持し、前記第1及び第2マイクアレイに設けられたマイクは指向性マイクであることが望ましい。

【0015】

前記直交円形マイクアレイシステムにおける音源方向推定部において行われる前記音源の3次元方向検出方法における前記(b)及び(c)段階は、好ましくは、前記第1または第2マイクアレイに設けられたマイクに入力される所定レベル以上の音声信号の出力値を合計して周波数領域に変換した後、周波数領域に変換された前記音声信号の出力値の合計値を音声信号の前記マイクにおける受信角を変数として変換し、最高の電力値を示す角度を音源の方向として推定するものである。

30

【0016】

前記所定レベル以上の音声信号の出力値の合計値 $y(t)$ は、好ましくは、アレイマイクの数 M であり、 c が媒質内における音速であり、アレイの中心から前記マイクまでの距離が r であるとき、

$$y(t) = \sum_{n=1}^M x_n(t + (n-1) \frac{2r \sin(\frac{\pi}{M}) \cos(\theta + \frac{2\pi(n-1)}{M})}{c})$$

40

である。

【0017】

前記直交円形マイクアレイシステムにおける音声信号処理部において行われる前記音源の3次元方向検出方法における(d)段階は、好ましくは、前記第1または第2マイクアレイに設けられたマイクの各々より受信された音声信号を前記各マイク間の位置差により生じた最大遅延時間だけ遅延させて合計し、前記指向性マイクより受信された音声信号を前記最大遅延時間だけ遅延させて前記合計値に加えて所望の音声信号の音質を改善する段

50

階を含む。

【0018】

【発明の実施の形態】

以下、添付した図面に基づき、本発明の好適な実施形態について説明する。

【0019】

図2Aは、本発明による直交円形マイクアレイの構造を示す図面であり、図2Bは、図2Aの直交円形マイクアレイの構造をロボットに対して適用し具現した例である。

【0020】

本発明において、緯度型マイクアレイ201と経度型マイクアレイ202とは物理的に互いと直交するように配され、図2Aに示すように、3次元の形構造を有する。また、マイクアレイシステムは、図2Bに示すように、ロボットや人形などの様々な形態のものに対して応用可能である。

10

【0021】

各円形マイクアレイ201、202は、本発明に用いられる指向性マイクの指向角及びマイクアレイが具現される物体の大きさを考慮して所定数のマイクが円形に配列されてなる。図2Cに示すように、XYまたはYZ平面の話者の位置を推定するマイクアレイ構造に取り付けられる指向性マイクの指向角 θ_1 が 90° であり、円形マイクアレイ構造の半径がRであるとした時、4個の指向性マイクが設けられた場合、表示された話者はマイクの指向角から外れて位置するため、話者の音声はマイクアレイに取り付けられたマイクにより受信されなくなる。

20

【0022】

一方、マイクの指向角が 90° より大きい(指向角が θ_2 である場合)か、あるいはマイクアレイの半径がより小さい時(半径がrである場合)には、同じ位置の話者により発音される音声はマイクアレイに取り付けられたマイクにより受信される。図2Cに示すところから明らかなように、アレイに取り付けられるマイクの指向角、話者との距離及びアレイが実装される物体の大きさを考慮してマイクアレイを構成しなければならない。マイクアレイは、指向性マイクの指向角に応じて少なくとも $(2/\sin\theta + 1)$ 個のマイクを備えれば全体 360° の範囲内に存在する話者の位置が検出可能になるが、マイクアレイが実装された物体と話者との間には一定の距離を保持する必要がある。

30

【0023】

図2に示すマイクアレイシステムにおいて、緯度型マイクアレイ201は、XY平面上において話者より音声を受信するので、話者のXY平面上の2次元位置が推定可能である。XY平面上において話者の2次元位置が推定されると、経度型マイクアレイ202は、推定された2次元位置に向かって回転して話者より音声を受信し、話者の3次元位置が推定可能となる。

【0024】

図3を参照し、直交円形マイクアレイを用いて話者の位置を推定し、話者の音声を受信する本発明によるシステム構成について説明する。

【0025】

本発明によるシステムは、話者の音声をXY平面上において2次元的に受信する緯度型マイクアレイ201と、推定された話者の2次元的位置に向かって話者の音声をYZ平面上において3次元的に受信する経度型マイクアレイ202と、マイクアレイ201、202より受信された信号から話者の位置を推定し、それによる制御信号を出力する方向検出部304と、緯度型マイクアレイ201より入力される音声信号と経度型マイクアレイ202より入力される音声信号とを選択的に方向検出部304に伝えるスイッチ303と、推定された話者の位置から音声を受信する超指向性マイク308と、超指向性マイク308及び経度型アレイマイク202より受信された音質を改善する音声処理部305と、経度型マイクアレイ202の回転方向及び角度を制御する第1回転制御部306及び超指向性マイク308の回転方向及び角度を制御する第2回転制御部307を備える。

40

【0026】

50

また、方向検出部 304 は、緯度型マイクアレイ 201 及び経度型マイクアレイ 202 に受信された信号から音声信号を識別する音声信号識別部 3041 と、音声信号識別部 3041 より受信された音声信号から、緯度型マイクアレイ 201 及び経度型マイクアレイ 202 より入力される音声信号の受信角に応じて音源の方向を推定する音源方向推定部 3042 と、音源方向推定部 3042 において推定された方向に経度型マイクアレイ 202 を回転させるように制御信号を出力し、スイッチ 303 に入力されるマイクアレイ信号をいつ切り換えるかを決定する制御信号を出力し、音声処理部 305 に音質改善された音声信号をいつ出力するかを決定する制御信号を出力する制御信号生成部 3043 を備える。

【0027】

以下、図 3 及び図 4 を参照し、本発明による話者の位置推定方法について説明する。

10

【0028】

本発明によるマイクアレイシステムが通電されると、まず、緯度型マイクアレイ 201 が作動して周りから信号を受信する（第 400 段階）。緯度型マイクアレイ 201 に設けられた指向性マイクは指向角度内で入力される信号を受信し、受信されたアナログ信号は A/D 変換器 310 によりデジタル信号に変換されてスイッチ 303 に出力される。最初の動作時に、スイッチ 303 は、緯度型マイクアレイ 201 より入力される信号を方向検出部 304 に伝える。

【0029】

方向検出部 304 に組み込まれた音声信号識別部 3041 は、スイッチ 303 を介して入力されたデジタル信号中に音声信号が存在するか否かを識別する（第 410 段階）。マイクアレイを介して話者の音声をできる限りきれいに受信して音声認識の性能を向上させることが本発明の目的であるという点に鑑みると、音声信号識別部 3041 が現在マイク 301 より入力された信号のうち音声区間だけを正確に検出して音声処理部 305 を介して音声認識部 320 に入力することは、音声認識の性能と関連して極めて重要である。

20

【0030】

音声認識は、音声のない区間が続いていて音声が入り始めたときにその瞬間をいかに正確に把握して音声信号が開始する瞬間を知らせるかという機能と、音声が続いていて音声のない区間が始まったときにこれをいかに正確に把握して音声信号が終了する瞬間を知らせるかという機能とに大きく分けられる。これら 2 つの機能を実現するものとして下記のような技術が公知である。

30

【0031】

まず、音声信号終了の瞬間を知らせる機能を実現するための方法として、マイクを介して入力される信号を一定のフレーム区間（例えば、30ms）に合わせて分割し、これら信号のエネルギーを計算した後にこのエネルギー値が以前のエネルギー値に比べて格段に下がったときに、音声信号がそれ以降発せられないと判断して音声信号終了の瞬間として処理する方法が知られている。この時、エネルギーが格段に小さくなることを判別する臨界値として一つの固定値を用いれば、大きく発音された音声と小さく発音された音声との差を無視できるので、過去の音声区間を観察してその臨界値を適応的に変え、この臨界値を用いて現在入力された信号が音声であるか否かを検出する方法が提案されている（R. ハリハラン（Hariharan, R.）, J. ハッキネン（Hakkinen, J.）, K. ローリラ（Laurila, K.）共著, 「ロウバスト・エンド・オブ・アタランス・デテクション・フォー・リアルタイム・スピーチ・レコグニション・アプリケーションズ（Robust end-of-utterance detection for real-time speech recognition applications）」, 2001 年度アイトリプリー・インターナショナル・コンファレンス・オン・アコースティクス・スピーチ・アンド・シグナルプロセッシング、論集（2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings）, 2001 年, 第 1 巻, 249 - 252 頁参照）。

40

【0032】

音声認識に関する他の方法としては、OOV（Out-Of-Vocabulary）に該当する語彙に対してガーベージ（非音声）モデルを予め構成しておいて、実際に使う時にマイクを介して

50

入った信号がこのモデルにいかほど適しているかを判断してガーベジであるか、それとも音声信号であるかを判別する方法が公知である。この方法は、音声ではない音を前もって学習することによってガーベジモデルを構成し、現在入力された信号がこのガーベジモデルにいかほど適しているかを判断して音声ノガーベジ区間を決定する方法である。また、雑音入り音声と雑音のっていない音声との関係を神経網や線形回帰分析を用いて推定した後、変換により雑音を除去する方法を用いることもある（J. カミネロ (Camirero, J.)、D. デラトッレ (De La Torre, D.)、L. ヴィラルビア (Villarrubia, L.)、C. マーチン (Martin, C.)、L. エルナンデス (Hernandez, L.) 共著、「オンライン・ガーベジ・モデリング・ウィズ・ディスクリミネント・アナリシス・フォー・アタランス・ベリフィケーション (On-line Garbage Modeling with Discriminant Analysis for Utterance Verification)」, 第4回インターナショナル・コンファレンス・オン・スポークン・ランゲージ、ICSLP 96 論集 (Fourth International Conference on Spoken Language, ICSLP 96 Proceedings)」, 1996年, 第4巻, 2111~2114頁参照)。

【0033】

前述した方法などを用い、音声信号識別部3041は、緯度型マイクアレイ201を介して一定レベル以上の音声信号値が入力されなければ現在音声が入力されていないと見なす。緯度型マイクアレイ201に設けられたマイク301のうち一部、例えば、 n 個のマイクから一定レベル以上の音声信号値が検出され、他のマイクからは信号値が入力されない場合には音声信号が検出されると判断し、話者が $(n+1) \times$ (ここで、 θ は指向角) 内に存在すると判断して入力された信号を音源方向推定部3042に出力する。

【0034】

図5A及び図5Bを参照し、話者の方向を推定する方法について説明する。話者から本発明によるマイクアレイへと入力される音声信号は、マイクアレイ201、202に設けられた各マイク301、302に達する時、音声信号が最初に達したマイクに比べて一定の遅延時間をもって受信される。この遅延時間はマイクの指向角 θ 及び話者の位置、すなわち、音声信号が入力される角度 α に応じて決定される。

【0035】

本実施形態では指向性マイクの特性を考慮し、音が一定の信号レベル未満に受信されるマイクの場合には、話者が各該当マイクの指向角内にないと判断し、これらマイクが担当する角度は話者の位置推定角度から除外される。

【0036】

音源方向推定部3042は、話者の位置を推定するために、図5Aに示されたように、まず、ある指向性マイクを基準マイクとして選択し、アレイの中心から基準となる指向性マイクをつなぐ仮想の線(基準線)から話者の音声を受信される角度 θ を測定する。基準となるマイク以外のマイクに対しては基準線と平行な仮想の線からマイクに受信される角度を測定する。もし、アレイの実装された物体が音源よりかなり大きくなければ、音声を受信する各マイクにおける音声受信角 θ_i は実質的に同じでありうる。

【0037】

一定レベル以上の音を受信されるマイクにおいて受信された全ての音を合計し、これを高速フーリエ変換(FFT)を通じて周波数領域に変換した後にこれを再び f の領域に変換した時、最大の電力値を有する f_m が話者の位置する方向を表わす。

【0038】

ある時間領域において一定の遅延時間をもって n 番目のマイクに入力される受信音声信号を $x_n(t)$ とし、各々のマイクの音声信号値を合計した出力信号を $y(t)$ とした時、 $y(t)$ は下記式により得られる。

【0039】

【数1】

10

20

30

40

$$y(t) = \sum_{n=1}^M x_n(t + \text{delay}_n)$$

$$= \sum_{n=1}^M x_n(t + (n-1) \frac{2r \sin(\frac{\pi}{M}) \cos(\theta + \frac{2\pi(n-1)}{M})}{c})$$

【 0 0 4 0 】

$y(t)$ を周波数領域に変換して得られた $Y(f)$ は、下記の通りである。

【 0 0 4 1 】

10

【 数 2 】

$$Y(f) = \sum_{n=1}^M X_n(f) \exp(j 2\pi f_{\text{speech}}(n-1) \frac{2r \sin(\frac{\pi}{M}) \cos(\theta + \frac{2\pi(n-1)}{M})}{c})$$

【 0 0 4 2 】

この時、 c は音源から音声伝えらるる媒質における音速を表わし、 r はマイクアレイに設けられたマイク間隔を表わし、 M はアレイに設けられたマイクの数を表わし、 θ はマイクに受信される音声信号の入射角を各々表わし、 $\theta = 2\pi / M$ が成立つ。

20

【 0 0 4 3 】

周波数領域に変換された $Y(f)$ を変数 θ により表わし、すなわち、 $Y(f)$ を再び θ の領域に変換し、 θ の領域において受信された音声信号のエネルギーを求めれば、下記の通りである。

【 0 0 4 4 】

【 数 3 】

$$P(\theta, k; m) = \left| \sum_{n=1}^M X_i(k; m) \exp(j 2\pi f_{\text{speech}}(n-1) \frac{2r \sin(\frac{\pi}{M}) \cos(\theta + \frac{2\pi(n-1)}{M})}{c}) \right|$$

30

【 0 0 4 5 】

但し、 θ は 0 から 2π の値を有し、 $Y(f)$ を再び θ の領域に変換する時、周波数領域における負の最大値が θ の領域における 0° に、周波数領域における 0° が θ の領域における $(n+1) \times 2\pi / M$ に、周波数領域における正の最大値が θ の領域における $(n+1) \times 2\pi / M$ にマッピングされるように周波数領域を θ の領域に変換する。

【 0 0 4 6 】

マイクアレイの出力である $P(\theta, k; m)$ により θ による出力エネルギーの大きさが分かり、最大出力時の θ を決定できるので、受信される音声から音声信号の直声強度が分かる。この式を全ての周波数 k に対してまとめれば、電力スペクトル値 $P(\theta; m)$ は下記の通りである。

40

【 0 0 4 7 】

【 数 4 】

$$P(\theta; m) = \sum_{k=0}^{K-1} \left| \sum_{n=1}^M X_i(k, m) \exp(j 2\pi f_{\text{speech}}(n-1) \frac{2r \sin(\frac{\pi}{M}) \cos(\theta + \frac{2\pi(n-1)}{M})}{c}) \right|^2$$

【 0 0 4 8 】

結論的に、全ての周波数領域において最大エネルギーを有する話者の方向を θ_s とした時

50

、 $s = \arg \max P(\theta; m)$ のように話者の方向を決定することができる（第 4 20 段階）。

【0049】

上述したように、緯度型マイクアレイ 201 より入力された音声信号から話者の緯度方向の 2 次元位置が推定されれば、音源方向推定部 3042 は制御信号生成部 3043 に検出された話者の方向 s を出力し、制御信号生成部 3043 は第 1 回転制御部 306 に話者の方向である s だけ経度型マイクアレイ 202 が回転するように制御信号を出力する。第 1 回転制御部 306 は経度型マイクアレイ 202 を s だけ回転させて経度型マイクアレイ 202 が 2 次元上の話者に正面から向かうように配置される。経度型マイクアレイ 202 を話者の方向に回転させる場合、経度型マイクアレイ 202 だけが回転することも可能であるが、緯度型マイクアレイ 201 と経度型マイクアレイ 202 とが共に回転することが好ましい。この場合には、緯度型マイクアレイ 201 と経度型マイクアレイ 202 とにより共用されるマイクが話者に向かえば適当に回転されたと判断できる（第 430 段階）。

10

【0050】

一方、経度型マイクアレイ 202 の回転が終了すれば、制御信号生成部 3043 はスイッチ 303 に制御信号を出力し、経度型マイクアレイ 202 より入力される話者の音声信号を音声信号識別部 3041 に伝える。方向検出部 304 は経度型マイクアレイ 202 より入力される音声信号を用い、第 420 段階における方法と同様にして話者の 3 次元空間上の位置を推定し、結果的に、話者の 3 次元空間における位置が、図 5B に示すように決定される。

20

【0051】

話者の 3 次元方向が決定されれば、制御信号生成部 3043 は第 2 回転制御部 307 に制御信号を出力し、超指向性マイク 308 を話者の 3 次元方向に直接的に向かうように回転させる（第 450 段階）。

【0052】

超指向性マイク 308 を介して受信された話者の音声信号は A/D 変換器 309 によりデジタル信号に変換された後、音声処理部 305 に入力される。入力された超指向性マイク信号は音声処理部 305 において経度型マイクアレイより受信された話者の音声信号と共に音質改善の過程に適用できる（第 460 段階）。

30

【0053】

本発明が適用される環境を示す図 6 及び音質改善の過程を示す図 7 を参照し、第 460 段階の音質改善過程について説明する。

【0054】

図 6 に示すように、本発明によるシステムには、マイクアレイを介して話者の音声信号だけではなく、壁体などの反射体からの反響信号及び機械などの雑音源からの雑音が受信される。本発明においては、超指向性マイク 308 により感知された信号とアレイマイクにより処理された音声信号とを共に処理して音質改善の効果を最大限に高めることができる。

【0055】

併せて、本発明においては、いったん話者の方向を決定し、話者の方向に超指向性マイク 308 を向かわせて超指向性マイク 308 により話者の音声信号を受信すれば、経度型マイクアレイ 202 または緯度型マイクアレイ 201 に受信された雑音や反響が音声処理部 305 に入力されることを防ぐために超指向性マイク 308 より受信された信号のみを処理することも可能である。但し、話者が位置を突然変えた場合、前述した段階を再び行って話者の変わった位置を決定するのに時間がかかり、その期間中には話者の音声処理されない場合がある。

40

【0056】

このような状況に対応し、本発明によるシステムは、図 7 に示すブラインド分離回路に緯度型マイクアレイ 201 または経度型マイクアレイ 202 より受信された話者の音声信号

50

と超指向性マイク308より受信された音声信号とを入力することにより、各マイクより入力された話者の音声信号と周りの雑音信号とを分離し、受信された音声信号の音質を向上させる。

【0057】

図7に示すように、マイクアレイに受信された信号だけではなく、超指向性マイク308に受信された音声信号が遅延時間をもって話者の音声信号を受信するアレイマイクの遅延時間まで遅延されて合計され、また処理される。

【0058】

図7に示す回路の動作について説明すれば、音声処理部305は、ブラインド分離回路にマイクアレイより入力される信号 $x^{array}(t)$ と超指向性マイクより入力される信号 $x^{direction}(t)$ とを入力する。これら2つの入力信号には話者の音声成分と周りの雑音成分との2種類が存在する。これが図7のブラインド分離回路に入力されると、雑音成分と話者の音声成分とを分離して $y_1(t)$ 及び $y_2(t)$ を出力する。出力される $y_1(t)$ 及び $y_2(t)$ は、下記式の通りである。

【0059】

【数5】

$$y_1(t) = x^{array}(t) + \sum_{j+i} \sum_{k=0}^L w_{array,j}(k) y_2(t-k)$$

$$y_2(t) = x^{direction}(t) + \sum_{j+i} \sum_{k=0}^L w_{direction,j}(k) y_1(t-k)$$

【0060】

前記式は $w_{array,j}(k) = -\mu \tanh(y_1(t)) y_j(t-k)$ 、 $w_{direction,j}(k) = -\mu \tanh(y_2(t)) y_1(t-k)$ により決定され、ウェイト w は最尤(ML: Maximum Likelihood)推定法に基づくものであり、相異なる信号成分が統計的に分離されるように学習された値が使われる。この時、 $\tanh()$ は非線形的なシグモイド(Sigmoid)関数を意味し、 μ は収束定数であって、ウェイトが最適の値を推定する度合いを決定する。

【0061】

話者の音声出力される間に、音源方向推定部3042は、緯度型マイクアレイ201及び経度型マイクアレイ202に受信される話者の音声から話者の位置が変わったか否かをチェックする。もし、話者の位置が変わったならば、第420段階へ進んで再び話者のXY平面上の位置とYZ平面上の位置とを推定する。但し、本発明の実施形態により話者のYZ平面上の位置のみが変わったならば、第440段階へ直ちに進んでも良い(第470段階)。

【0062】

話者の位置が変わっていない場合には、音声信号識別部3041は、第410段階における方法と同様にして話者の音声発話が終了したか否かを検出する。ここで、話者の音声発話が終了していなければ、再び話者の位置が変わったか否かを検出する(第480段階)

【0063】

【発明の効果】

上述したように、本発明によれば、指向性マイクが円形に一定の間隔をもって配置された緯度型マイクアレイと経度型マイクアレイとを互いと直交するように配置することにより、平面的に移動する話者の位置だけではなく話者の3次元空間上の動きまで考慮でき、話者がいかなる方向で発音しても話者の音声を受信できる効果がある。

【0064】

併せて、3次元的な話者の位置が決定されれば、超指向性マイクを話者の方向に向かわせて話者の音声信号を受信することにより、受信される話者の音声を最大化し、話者が発音

10

20

30

40

50

する時に生じる反響や周辺雑音の影響を最小化して話者の音声を明確に認識させることにより、音声認識の性能を向上させられる。

【0065】

さらに、超指向性マイクより受信される話者の音声信号だけではなく緯度型マイクアレイまたは経度型マイクアレイより受信されてから各マイクごとに一定の遅延間隔をもって遅延されて合計された信号を超指向性マイクより受信された信号と共に出力することにより、出力効率を高められる効果がある。

【0066】

上述した本発明の好適な実施形態は例示的なものであり、これに対するあらゆる変形及び変更は特許請求の範囲に属すると理解されるべきである。

10

【図面の簡単な説明】

【図1A】従来の技術による平面形マイクアレイシステムの構造を示す図面である。

【図1B】従来の技術による円形マイクアレイシステムの構造を示す図面である。

【図2A】本発明による直交円形マイクアレイの構造を示す図面である。

【図2B】本発明によるマイクアレイを具体的に適用した例を示す図面である。

【図2C】マイクアレイ上にマイクを配置する時の考慮事項を説明するための図面である。

【図3】本発明による直交円形マイクアレイシステムの構成を示すブロック図である。

【図4】本発明による音源の3次元方向検出方法を示すフローチャートである。

【図5A】本発明により音源の方向を推定するために音源の角度を分析する例を示す図面である。

20

【図5B】最終的に決定された話者の位置を示す図面である。

【図6】本発明によるシステムが適用される環境を示す図面である。

【図7】音源より受信された音声信号を分離して音質の向上を図るブラインド分離回路を示す図面である。

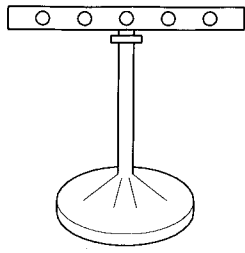
【符号の説明】

201 緯度型マイクアレイ

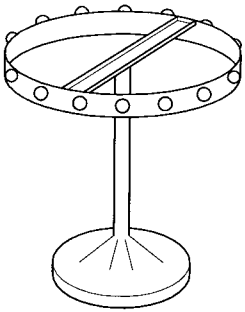
202 経度型マイクアレイ

304 方向検出部

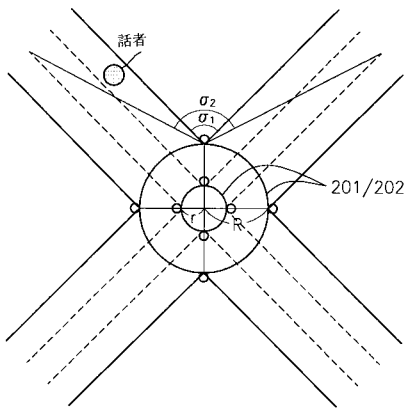
【図 1 A】



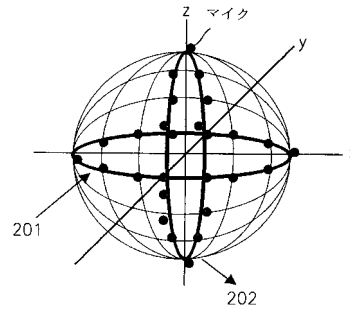
【図 1 B】



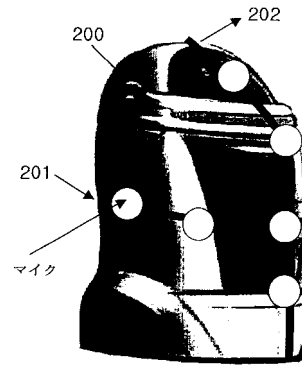
【図 2 C】



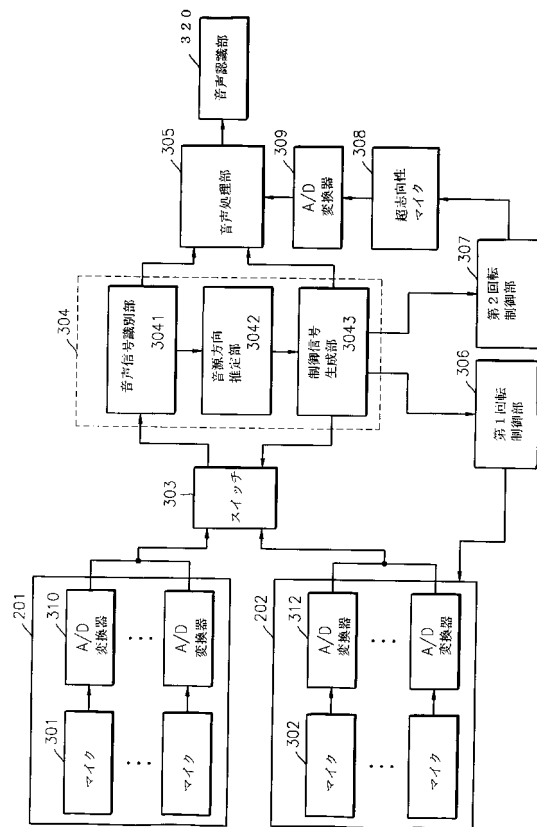
【図 2 A】



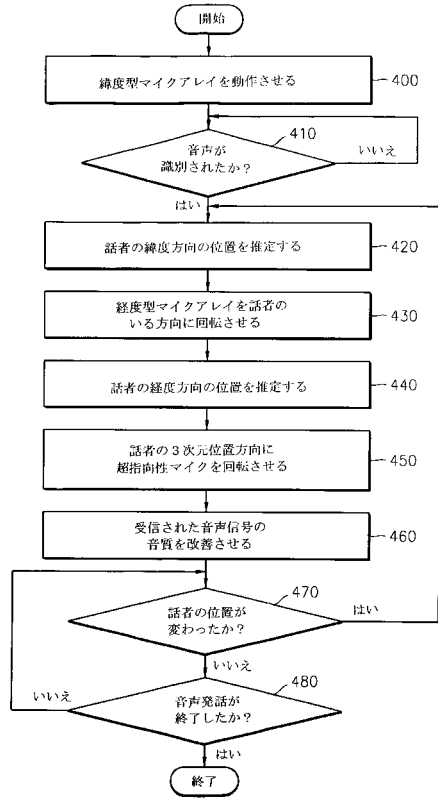
【図 2 B】



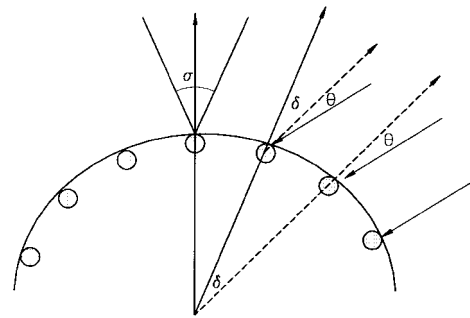
【図 3】



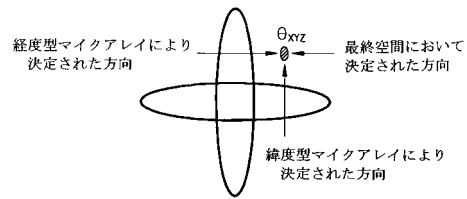
【図4】



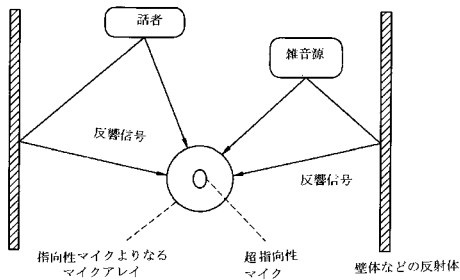
【図5A】



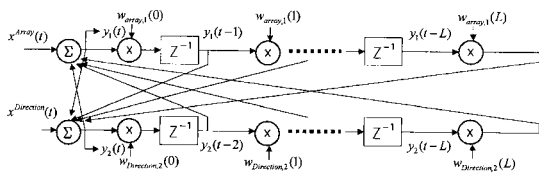
【図5B】



【図6】



【図7】



フロントページの続き

- (72)発明者 金 載 祐
大韓民国 京畿道 竜仁市 器興邑 永德里 15番地 信一アパート 109棟 2006号
- (72)発明者 金 尚 龍
大韓民国 京畿道 竜仁市 水枝邑 上 硯 里 30番地 盛源アパート 118棟 1204号

審査官 新川 圭二

- (56)参考文献 特開昭60-090499(JP,A)
特開2000-152372(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04R 1/40
H04R 3/00