

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6007487号
(P6007487)

(45) 発行日 平成28年10月12日 (2016. 10. 12)

(24) 登録日 平成28年9月23日 (2016. 9. 23)

(51) Int. Cl.		F I			
HO4R	3/00	(2006.01)	HO4R	3/00	320
HO4R	1/40	(2006.01)	HO4R	1/40	320A
HO4R	1/06	(2006.01)	HO4R	1/06	320

請求項の数 5 (全 14 頁)

(21) 出願番号	特願2011-268288 (P2011-268288)	(73) 特許権者	000005496
(22) 出願日	平成23年12月7日 (2011. 12. 7)		富士ゼロックス株式会社
(65) 公開番号	特開2013-121078 (P2013-121078A)		東京都港区赤坂九丁目7番3号
(43) 公開日	平成25年6月17日 (2013. 6. 17)	(74) 代理人	100104880
審査請求日	平成26年11月21日 (2014. 11. 21)		弁理士 古部 次郎
		(74) 代理人	100125346
			弁理士 尾形 文雄
		(74) 代理人	100166981
			弁理士 砂田 岳彦
		(72) 発明者	西野 洋平
			神奈川県足柄上郡中井町境430 グリー
			ンテクなかい 富士ゼロックス株式会社内
		(72) 発明者	原田 陽雄
			神奈川県足柄上郡中井町境430 グリー
			ンテクなかい 富士ゼロックス株式会社内
			最終頁に続く

(54) 【発明の名称】 情報処理装置、情報処理システムおよびプログラム

(57) 【特許請求の範囲】

【請求項1】

話者の音声を取得する少なくとも2以上の音声取得手段が取得した話者の音声の情報に基づいて、話者が端末装置を装着する装着者が当該装着者とは異なる人である他者かを識別した自他識別情報と、当該装着者と当該他者との向き合う角度である対面角度の情報を取得する取得手段と、

前記取得手段が取得した前記自他識別情報と前記対面角度の情報とに基づいて、前記装着者と前記他者との対話状況を特定する特定手段と、

を備えたことを特徴とする情報処理装置。

【請求項2】

前記特定手段は、前記装着者と前記他者とが会話を行なっているか否かを特定することを特徴とする請求項1記載の情報処理装置。

【請求項3】

話者の音声を取得する少なくとも2以上の音声取得手段と、

前記少なくとも2以上の音声取得手段が取得した話者の音声の情報に基づいて、話者が本対面角度出力手段を装着する装着者が当該装着者とは異なる人である他者かを識別した自他識別情報と、当該装着者と当該話者との向き合う角度である対面角度の情報を取得する取得手段と、

前記取得手段により取得された前記自他識別情報と前記対面角度の情報とを送信する送信手段と、

を備える対面角度出力手段と、
 前記対面角度出力手段の前記送信手段により送信された前記自他識別情報と前記対面角度の情報とを受信する受信手段と、
前記受信手段が受信した前記自他識別情報と前記対面角度の情報とに基づいて、前記装着者と前記他者との対話状況を特定する特定手段と、
 を備えることを特徴とする情報処理システム。

【請求項 4】

話者の音声を取得する少なくとも 2 以上の音声取得手段と、
 前記少なくとも 2 以上の音声取得手段により取得された前記音声の情報を送信する送信手段と、
 前記送信手段により送信された前記音声の情報を受信する受信手段と、
 前記受信手段により受信された複数の前記音声の情報に基づいて、話者が端末装置を装着する装着者か当該装着者とは異なる人である他者かを識別した自他識別情報と、当該装着者と当該他者との向き合う角度である対面角度の情報とを取得する取得手段と、
前記取得手段が取得した前記自他識別情報と前記対面角度の情報とに基づいて、前記装着者と前記他者との対話状況を特定する特定手段と、
 を備えることを特徴とする情報処理システム。

【請求項 5】

コンピュータに、
 話者の音声を取得する少なくとも 2 以上の音声取得手段が取得した話者の音声の情報に基づいて、話者が端末装置を装着する装着者か当該装着者とは異なる人である他者かを識別した自他識別情報と、当該装着者と当該他者との向き合う角度である対面角度の情報とを取得する機能と、
取得した前記自他識別情報と前記対面角度の情報とに基づいて、前記装着者と前記他者との対話状況を特定する機能と、
 を実現させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、対面角度出力装置、対面角度出力システムおよびプログラムに関する。

【背景技術】

【0002】

特許文献 1 には、ミラー部とミラー部を可動に支えて自動車内に設けられるミラー支持部とを含むルームミラーと、ミラー部の背後に隠れることなくミラー支持部に設けられた第 1 のマイクロフォンと、ミラー部に設けられた第 2 のマイクロフォンとを含み、第 1 のマイクロフォンは、マイクロフォン支持部を介してミラー支持部に設けられる自動車用マイクロフォンアレイ装置が開示されている。

特許文献 2 には、複数のマイクロホンで受音した信号を処理し、音源の位置を推定する方法において、まず、受音信号の相互相関関数を全てのマイクロホンの組について計算し、次に、上記相互相関関数について、基準マイクロホンと他のマイクロホンとの間の相互相関関数の最大値を与える時間差を求め、これを予備推定時間差とし、そして、全てのマイクロホンについての遅延和のパワーを最大にする時間差を上記予備推定時間差の近傍で探索して、これを推定時間差とし、最後に、上記推定時間差に基づいて音源位置を計算する音源位置推定方法が開示されている。

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2009 - 184480 号公報

【特許文献 2】特開平 11 304906 号公報

【発明の概要】

10

20

30

40

50

【発明が解決しようとする課題】

【0004】

本発明は、装着者の音声か他者の音声かを識別し、他者の音声の場合に、装着者と音声を発生する話者との向き合う角度である対面角度を、把握することを目的とする。

【課題を解決するための手段】

【0005】

請求項1に記載の発明は、話者の音声を取得する少なくとも2以上の音声取得手段が取得した話者の音声の情報に基づいて、話者が端末装置を装着する装着者か当該装着者とは異なる人である他者かを識別した自他識別情報と、当該装着者と当該他者との向き合う角度である対面角度の情報とを取得する取得手段と、前記取得手段が取得した前記自他識別情報と前記対面角度の情報とに基づいて、前記装着者と前記他者との対話状況を特定する特定手段と、を備えたことを特徴とする情報処理装置である。

10

【0006】

請求項2に記載の発明は、前記特定手段は、前記装着者と前記他者とが会話を行なっているか否かを特定することを特徴とする請求項1記載の情報処理装置である。

請求項3に記載の発明は、話者の音声を取得する少なくとも2以上の音声取得手段と、前記少なくとも2以上の音声取得手段が取得した話者の音声の情報に基づいて、話者が本対面角度出力手段を装着する装着者か当該装着者とは異なる人である他者かを識別した自他識別情報と、当該装着者と当該話者との向き合う角度である対面角度の情報とを取得する取得手段と、前記取得手段により取得された前記自他識別情報と前記対面角度の情報とを送信する送信手段と、を備える対面角度出力手段と、前記対面角度出力手段の前記送信手段により送信された前記自他識別情報と前記対面角度の情報とを受信する受信手段と、前記受信手段が受信した前記自他識別情報と前記対面角度の情報とに基づいて、前記装着者と前記他者との対話状況を特定する特定手段と、を備えることを特徴とする情報処理システムである。

20

【0007】

請求項4に記載の発明は、話者の音声を取得する少なくとも2以上の音声取得手段と、前記少なくとも2以上の音声取得手段により取得された前記音声の情報を送信する送信手段と、前記送信手段により送信された前記音声の情報を受信する受信手段と、前記受信手段により受信された複数の前記音声の情報に基づいて、話者が端末装置を装着する装着者か当該装着者とは異なる人である他者かを識別した自他識別情報と、当該装着者と当該他者との向き合う角度である対面角度の情報とを取得する取得手段と、前記取得手段が取得した前記自他識別情報と前記対面角度の情報とに基づいて、前記装着者と前記他者との対話状況を特定する特定手段と、を備えることを特徴とする情報処理システムである。

30

【0008】

請求項5に記載の発明は、コンピュータに、話者の音声を取得する少なくとも2以上の音声取得手段が取得した話者の音声の情報に基づいて、話者が端末装置を装着する装着者か当該装着者とは異なる人である他者かを識別した自他識別情報と、当該装着者と当該他者との向き合う角度である対面角度の情報とを取得する機能と、取得した前記自他識別情報と前記対面角度の情報とに基づいて、前記装着者と前記他者との対話状況を特定する機能と、を実現させるプログラムである。

40

【発明の効果】

【0009】

請求項1の発明によれば、装着者と他者の対話状況を特定することができる情報処理を提供できる。

請求項3の発明によれば、装着者と他者の対話状況を特定することができるシステムを構築できる。

請求項4の発明によれば、装着者と他者の対話状況を特定することができるシステムを構築できる。

請求項5の発明によれば、装着者と他者の対話状況を特定することができる機能をコン

50

コンピュータにより実現できる。

【図面の簡単な説明】

【0010】

【図1】本実施形態による対面角度出力システムの構成例を示す図である。

【図2】端末装置の構成例を示す図である。

【図3】本実施の形態における対面角度について説明した図である。

【図4】第1マイクロフォンおよび第2マイクロフォンを使用して対面角度を求める方法について説明した図である。

【図5】(a)～(c)は、本実施の形態において時間差を求める方法を説明した図である。

【図6】対面角度と第1マイクロフォンと第2マイクロフォンに話者との音声が届く時間差との関係を示した図である。

【図7】端末装置の動作について説明したフローチャートである。

【図8】(a)～(b)は、実際に使用される端末装置の具体例を示した図である。

【発明を実施するための形態】

【0011】

以下、添付図面を参照して、本発明の実施形態について詳細に説明する。

<システム構成例>

図1は、本実施形態による対面角度出力システムの構成例を示す図である。

図1に示すように、本実施形態の対面角度出力システム1は、対面角度出力装置の一例である端末装置10とホスト装置20とを備えて構成される。端末装置10とホスト装置20とは、無線通信回線を介して接続されている。無線通信回線の種類としては、Wi-Fi (Wireless Fidelity)、Bluetooth (登録商標)、ZigBee、UWB (Ultra Wideband) 等の既存の方式による回線を用いて良い。また、図示の例では、端末装置10が1台のみ記載されているが、詳しくは後述するように、端末装置10は、使用者各人が装着して使用するものであり、実際には使用者数分の端末装置10が用意される。

【0012】

端末装置10は、話者の音声を取得する音声取得手段として、話者の音声を取得する複数のマイクロフォン(第1マイクロフォン11a、第2マイクロフォン11b、第3マイクロフォン11c)と、増幅器(第1増幅器13a、第2増幅器13b、および第3増幅器13c)とを備える。また、端末装置10は、収録音声を解析する音声解析部15と、解析結果をホスト装置20に送信するためのデータ送信部16とを備え、さらに電源部17とを備える。

【0013】

本実施の形態において、第1マイクロフォン11aと第2マイクロフォン11bとは、水平方向に予め定められた距離だけ離間して配される。ここでは、第1マイクロフォン11aと第2マイクロフォン11bとは装着者の口に近い位置に水平方向に並んで配置され、この間の距離は例えば、10cm～20cmである。また第3マイクロフォン11cは、第1マイクロフォン11a、第2マイクロフォン11bに対し装着者の口(発声部位)からの距離が遠い位置に配置される。ここでは、第3マイクロフォン11cは、第1マイクロフォン11a、第2マイクロフォン11bから、例えば35cm程度の距離をおいて下方に配置される。つまり本実施の形態では、端末装置10に配されるマイクロフォンは、装着者の口からの距離が互いに異なる2つ、および水平方向の距離が離れた2つの双方を選択可能となっている。ここでは前者として、第1マイクロフォン11aと第3マイクロフォン11cの組、および第2マイクロフォン11bと第3マイクロフォン11cの組が選択できる。さらに後者として、第1マイクロフォン11aと第2マイクロフォン11bの組が選択できる。

本実施形態の第1マイクロフォン11a、第2マイクロフォン11b、第3マイクロフォン11cとして用いられるマイクロフォンの種類としては、ダイナミック型、コンデンサ型等、既存の種々のものを用いて良い。とくに無指向性のMEMS (Micro Electro Me

10

20

30

40

50

chanical Systems) 型マイクロフォンが好ましい。

【0014】

第1増幅器13a、第2増幅器13b、および第3増幅器13cは、それぞれ第1マイクロフォン11a、第2マイクロフォン11b、第3マイクロフォン11cが取得音声に応じて出力する電気信号を増幅する。本実施形態の第1増幅器13a、第2増幅器13b、第3増幅器13cとして用いられる増幅器としては、既存のオペアンプ等を用いて良い。

【0015】

音声解析部15は、第1増幅器13a、第2増幅器13b、第3増幅器13cから出力された電気信号を解析する。そして詳しくは後述するが、話者が装着者か他者かを識別し、話者が他者と識別されたときに装着者と話者との向き合う角度である対面角度を出力する。即ち、音声解析部15は、話者が装着者か他者かを識別する識別手段として機能するとともに、装着者と話者との向き合う角度である対面角度を出力する出力手段として機能する。

10

【0016】

データ送信部16は、音声解析部15による解析結果を含む取得データと端末IDを、上記の無線通信回線を介してホスト装置20へ送信する。ホスト装置20へ送信する情報としては、ホスト装置20において行われる処理の内容に応じて、上記の解析結果の他、例えば、第1マイクロフォン11a、第2マイクロフォン11b、第3マイクロフォン11cによる音声の取得時刻、取得音声の音圧等の情報を含めて良い。また端末装置10に音声解析部15による解析結果を蓄積するデータ蓄積部を設け、一定期間の保存データを一括送信しても良い。なお有線回線で送信しても良い。本実施の形態では、データ送信部16は、対面角度の情報を送信する送信手段として機能する。

20

【0017】

電源部17は、上記の第1マイクロフォン11a、第2マイクロフォン11b、第3マイクロフォン11c、第1増幅器13a、第2増幅器13b、第3増幅器13c、音声解析部15およびデータ送信部16に電力を供給する。電源としては、例えば乾電池や充電電池等の既存の電源が用いられる。また、電源部17は、必要に応じて、電圧変換回路および充電制御回路等の周知の回路を含む。

【0018】

ホスト装置20は、端末装置10から送信されたデータを受信するデータ受信部21と、受信したデータを蓄積するデータ蓄積部22と、蓄積したデータを解析するデータ解析部23と、解析結果を出力する出力部24とを備える。このホスト装置20は、例えばパーソナルコンピュータ等の情報処理装置により実現される。また、上記のように本実施形態では複数台の端末装置10が使用され、ホスト装置20は、その複数台の端末装置10の各々からデータを受信する。

30

【0019】

データ受信部21は、上記の無線回線に対応しており、各端末装置10からデータを受信してデータ蓄積部22へ送る。本実施の形態では、データ受信部21は、データ送信部16により出力された対面角度の情報を受信する受信手段として機能する。データ蓄積部22は、データ受信部21から取得した受信データを発話者別に蓄積する。ここで、発話者の識別は、端末装置10から送信される端末IDと、あらかじめホスト装置20に登録されている発話者名と端末IDの照合により行う。また、端末装置10から端末IDのかわりに装着者状態を送信するようにしてもよい。

40

【0020】

データ解析部23は、データ蓄積部22に蓄積されたデータを解析する。具体的な解析内容および解析手法は、本実施形態のシステムの利用目的や利用態様に依りて種々の内容および手法を取り得る。例えば、端末装置10の装着者どうしの対話頻度や各装着者の対話相手の傾向を分析したり、対話における個々の発話の長さや音圧の情報から対話者の関係を類推したりすることが行われる。

50

【 0 0 2 1 】

出力部 2 4 は、データ解析部 2 3 による解析結果を出力したり、解析結果に基づく出力を行ったりする。この解析結果等を出力する手段は、システムの利用目的や利用態様、解析結果の内容や形式等に応じて、ディスプレイ表示、プリンタによる印刷出力、音声出力等、種々の手段を取り得る。

【 0 0 2 2 】

< 端末装置の構成例 >

図 2 は、端末装置 1 0 の構成例を示す図である。

上記のように、端末装置 1 0 は、各使用者に装着されて使用される。使用者が装着可能とするため、本実施形態の端末装置 1 0 は、図 2 に示すように、装置本体 3 0 と、装置本体 3 0 に両端が接続することにより環形状をなす提げ紐 4 0 とを備えた構成とする。図示の構成において、使用者は、提げ紐 4 0 に首を通し、装置本体 3 0 を首から提げて装着する。そのため本実施の形態では、装置本体 3 0 および提げ紐 4 0 を「人間の首に掛けられる部材」と表現できる。また本実施の形態では、端末装置 1 0 を装着する使用者を装着者として表現することがある。

10

【 0 0 2 3 】

装置本体 3 0 は、金属や樹脂等で形成された薄い直方体のケース 3 1 に、少なくとも第 1 増幅器 1 3 a、第 2 増幅器 1 3 b、第 3 増幅器 1 3 c、音声解析部 1 5、データ送信部 1 6 および電源部 1 7 を実現する回路と電源部 1 7 の電源（電池）とを収納して構成される。また本実施の形態では、ケース 3 1 に第 3 マイクロフォン 1 1 c が設けられる。さらにケース 3 1 には、装着者の氏名や所属等の ID 情報を表示した ID カード等を挿入するポケットを設けても良い。また、ケース 3 1 自体の表面にそのような ID 情報等を記載しても良い。

20

【 0 0 2 4 】

提げ紐 4 0 には、第 1 マイクロフォン 1 1 a、第 2 マイクロフォン 1 1 b が設けられる（以下、第 1 マイクロフォン 1 1 a、第 2 マイクロフォン 1 1 b、および第 3 マイクロフォン 1 1 c をそれぞれ区別しない場合には、マイクロフォン 1 1 a、1 1 b、1 1 c と記載）。提げ紐 4 0 の材質としては、革、合成皮革、木綿その他の天然繊維や樹脂等による合成繊維、金属等、既存の種々の材質を用いて良い。また、シリコン樹脂やフッ素樹脂等を用いたコーティング処理が施されていても良い。

30

【 0 0 2 5 】

この提げ紐 4 0 は、筒状の構造を有し、提げ紐 4 0 の内部にマイクロフォン 1 1 a、1 1 b を収納している。マイクロフォン 1 1 a、1 1 b を提げ紐 4 0 の内部に設けることにより、マイクロフォン 1 1 a、1 1 b の損傷や汚れを防ぎ、対話者がマイクロフォン 1 1 a、1 1 b の存在を意識することが抑制される。

【 0 0 2 6 】

< 話者が装着者か他者であるかを識別する方法の説明 >

以上の構成において、話者が装着者か装着者以外の者である他者であるかを識別（自他識別）する方法について説明する。

本実施形態のシステムは、端末装置 1 0 に設けられたマイクロフォンのうち、例えば、第 1 マイクロフォン 1 1 a と第 3 マイクロフォン 1 1 c により収録された音声の情報を用いて、端末装置 1 0 の装着者自身の発話音声と他者の発話音声とを識別する。言い換えれば、本実施形態は、収録音声の発話者に関して自他の別を識別する。また、本実施形態では、収録音声の情報のうち、形態素解析や辞書情報を用いて得られる言語情報ではなく、音圧（第 1 マイクロフォン 1 1 a と第 3 マイクロフォン 1 1 c への入力音量）等の非言語情報に基づいて発話者を識別する。言い換えれば、言語情報により特定される発話内容ではなく、非言語情報により特定される発話状況から音声の発話者を識別する。

40

【 0 0 2 7 】

図 1 および図 2 を参照して説明したように、本実施形態において、端末装置 1 0 の第 3 マイクロフォン 1 1 c は装着者の口（発声部位）から遠い位置に配置され、第 1 マイクロ

50

フォン11aは装着者の口（発声部位）に近い位置に配置される。すなわち、装着者の口（発声部位）を音源とすると、第1マイクロフォン11aと音源との間の距離と、第3マイクロフォン11cと音源との間の距離が大きく異なる。例えば、第1マイクロフォン11aと音源との間の距離は、第3マイクロフォン11cと音源との間の距離の1.5～4倍程度に設定することができる。ここで、マイクロフォン11a、11cにおける収録音声の音圧は、マイクロフォン11a、11cと音源との間の距離が大きくなるにしたがって減衰（距離減衰）する。したがって、装着者の発話音声に関して、第1マイクロフォン11aにおける収録音声の音圧と第3マイクロフォン11cにおける収録音声の音圧とは大きく異なる。

【0028】

一方、装着者以外の者（他者）の口（発声部位）を音源とした場合を考えると、その他者が装着者から離れているため、第1マイクロフォン11aと音源との間の距離と、第3マイクロフォン11cと音源との間の距離は、大きく変わらない。装着者に対する他者の位置によっては、両距離の差は生じ得るが、装着者の口（発声部位）を音源とした場合のように、第1マイクロフォン11aと音源との間の距離が第3マイクロフォン11cと音源との間の距離の数倍となることはない。したがって、他者の発話音声に関して、第1マイクロフォン11aにおける収録音声の音圧と第3マイクロフォン11cにおける収録音声の音圧とは、装着者の発話音声の場合のように大きく異なることはない。

【0029】

そこで、本実施形態では、第1マイクロフォン11aにおける収録音声の音圧と第3マイクロフォン11cにおける収録音声の音圧の比である音圧比を求める。そしてこの音圧比の差を用いて、収録音声における装着者自身の発話音声と他者の発話音声とを識別する。より具体的には、本実施形態では、第3マイクロフォン11cの音圧と第1マイクロフォン11aの音圧との比に閾値を設定する。そして、音圧比が閾値よりも大きい音声は装着者自身の発話音声と判断し、音圧比が閾値よりも小さい音声は他者の発話音声と判断する。

【0030】

なお上述した例では、第1マイクロフォン11aと第3マイクロフォン11cを使用して自他識別の判断を行なったが、これに限られるものではなく、第2マイクロフォン11bと第3マイクロフォン11cを使用しても同様であることは勿論である。

また上述した例では、第1マイクロフォン11aおよび第3マイクロフォン11cによる収録音声の音圧を基に自他識別の判断を行なったが、これに収録音声の位相差の情報を加味することも考えられる。つまり装着者の口（発声部位）を音源とすると、上述した通り、第1マイクロフォン11aと音源との間の距離と、第3マイクロフォン11cと音源との間の距離は大きく異なる。そのため第1マイクロフォン11aによる収録音声と、第3マイクロフォン11cによる収録音声の位相差は大きくなる。一方、装着者以外の者（他者）の口（発声部位）を音源とした場合は、上述の通り、他者が装着者から離れているため、第1マイクロフォン11aと音源との間の距離と、第3マイクロフォン11cと音源との間の距離は、大きく変わらない。そのため第1マイクロフォン11aによる収録音声と、第3マイクロフォン11cによる収録音声の位相差は小さくなる。よって第1マイクロフォン11aおよび第3マイクロフォン11cによる収録音声の位相差を考慮することで自他識別の判断の精度が向上する。

【0031】

< 対面角度の説明 >

図3は、本実施の形態における対面角度について説明した図である。

本実施の形態において対面角度とは、端末装置10の装着者と話者との向き合う角度である。そして本実施の形態で定義される対面角度の一例として図3では、水平方向の対面角度を示している。つまり図3は、装着者と話者を頭上から見た図である。そして本実施の形態は、本実施の形態では、対面角度として、2個の音声取得手段である第1マイクロフォン11aと第2マイクロフォン11bとを結ぶ線分と、この線分の中点および話者

10

20

30

40

50

を結ぶ線分との角度を採用する。これにより対面角度の数学的取り扱いがより簡単になる。そしてこの定義を採用した場合、例えば、装着者と話者とがそれぞれ正面を向いて対向しているときは、この二人の対面角度は、 90° となる。

【0032】

<対面角度を求める方法の説明>

図4は、第1マイクロフォン11aおよび第2マイクロフォン11bを使用して対面角度を求める方法について説明した図である。

ここで、点Sが話者の位置、より正確には、話者の音声の音源である発声点の位置であるとする。そして発声点から発した音声は、点Sから同心円状に広がる。ただし音声は有限の速度である音速で広がるため音声は第1マイクロフォン11aに到達した時間と第2マイクロフォン11bに到達した時間とは異なり、音声の行路差に対応した時間差 t が生じる。そして第1マイクロフォン11aと第2マイクロフォン11bとの間の距離をD、中点Cと点Sの間の距離をLとするとこれらの間には次の(1)式が成立する。

【0033】

$$= (L^2 + LD \cos \theta + D^2 / 4)^{0.5} - (L^2 - LD \cos \theta + D^2 / 4)^{0.5} \quad \dots (1)$$

【0034】

この(1)式は、 $L > D$ の場合、Lによる影響が小さいため、下記(2)式に近似できる。

【0035】

$$D \cos \theta \quad \dots (2)$$

【0036】

また音速 c と時間差 t を使用すると、下記(3)式が成立する。

【0037】

$$= c \cdot t \quad \dots (3)$$

【0038】

つまりこの(2)式および(3)式を使用することで、対面角度 θ が求まる。つまり2つの音声取得手段である第1マイクロフォン11aと第2マイクロフォン11bに話者の音声は到達する時間差 t および第1マイクロフォン11aと第2マイクロフォン11bが離間する距離Dに基づいて、装着者と話者との向き合う角度である対面角度 θ を計算することができる。

【0039】

また第1マイクロフォン11aと第2マイクロフォン11bに話者の音声は到達する時間差 t は、以下のようにして求めることができる。

図5(a)~(c)は、本実施の形態において時間差 t を求める方法を説明した図である。このうち図5(a)は、第1マイクロフォン11aと第2マイクロフォン11bに到達する話者の音声をサンプリング周波数1MHzでサンプリングし、そのデータの中から連続する5000ポイントを抜き出した図である。

ここで横軸は5000ポイントのデータにそれぞれに付けられたデータ番号を表わし、縦軸は、話者の音声の振幅を表わす。そして実線は、第1マイクロフォン11aに到達した話者の音声の波形信号であり、点線は、第2マイクロフォン11bに到達した話者の音声の波形信号である。

【0040】

本実施の形態では、この2つの波形信号の相互相関関数を求める。つまり一方の波形信号を固定し、他方の波形信号をシフトしつつ積和をとる計算を行なう。図5(b)~(c)は、この2つの波形信号に対する相互相関関数を示した図である。このうち図5(b)は、サンプリングした5000ポイントのデータ全体の相互相関関数であり、図5(c)は、図5(b)に示した相互相関関数のピーク付近を拡大した図である。なお図5(b)~(c)では、第1マイクロフォン11aに到達した話者の音声の波形信号を固定し、第2マイクロフォン11bに到達した話者の音声の波形信号をシフトして相互相関関数を求

10

20

30

40

50

めた場合を示している。

図5(c)に示すようにデータ番号0を基準にしてピーク位置は、-227ポイントずれている。これは第1マイクロフォン11aを基準にして第2マイクロフォン11bに到達する話者の音声、この分遅延して到達することを意味する。本実施の形態においてサンプリング周波数は、上述の通り1MHzなので、サンプリングした各データ間の時間は、 1×10^{-6} (s)である。よってこの遅延時間としては $227 \times 1 \times 10^{-6}$ (s) = 227 (μ s)となる。つまりこの場合、時間差 t は、227 (μ s)である。

【0041】

また本実施の形態では、振幅を予め定められた周波数帯毎に分け、最も振幅が大きい周波数帯について大きい重み付けをして相互相関関数を求めている。これにより求められる時間差 t がより正確になる。またこの時間差 t をより正確に求めるために第1マイクロフォン11aと第2マイクロフォン11bの距離は1cm~100cmの範囲内になることが好ましい。第1マイクロフォン11aと第2マイクロフォン11bの距離が1cm未満であると時間差 t が小さくなりすぎ、この後導出される対面角度の誤差が大きくなりやすい。また100cmより大きいと時間差 t を導出する際に、反射音の影響を受けやすくなる。また相互相関関数を求める際に、より長い時間幅について計算を行なう必要があるため計算に要する負荷が大きくなる。

【0042】

図6は、対面角度 θ と第1マイクロフォン11aと第2マイクロフォン11bに話者の音声到達する時間差 t との関係を示した図である。図6において、横軸は対面角度 θ (°)を表わし、縦軸は、時間差 t を表わす。

図6により例えば、装着者と話者とが正面同士で向き合っているとき、つまり対面角度が90°の場合は、時間差 t は、0であり、対面角度 θ が90°の場合からずれるに従い時間差 t は大きくなることわかる。

なお対面角度 θ は、上述のようにして算出してもよいが、例えば、図6を基にして時間差 t に対する対面角度 θ をLUT (Look up Table)として、端末装置10内に保持しておき、 t を求めた後、このLUTを参照することにより対面角度 θ を求めてもよい。

【0043】

図7は、端末装置10の動作について説明したフローチャートである。以下、図2および図7を使用して端末装置10の動作について説明を行なう。

まずマイクロフォン11a、11b、11cが、話者の音声を取得する(ステップ101)。そして第1増幅器13a、第2増幅器13b、第3増幅器13cがそれぞれマイクロフォン11a、11b、11cからの音声信号を増幅する(ステップ102)。

次に増幅された音声信号は、音声解析部15に送られ、上述した手法により話者が他者であるか否かを判断する(ステップ103)。話者が装着者であると識別(即ち他者ではないと識別)された場合(ステップ103でNo)は、ステップ101に戻る。一方、話者が他者であると識別された場合(ステップ103でYes)は、図5で説明した手法により、まず第1マイクロフォン11aと第2マイクロフォン11bに話者の音声到達する時間差 t が求められる(ステップ104)。さらに図4で説明した手法により、この時間差 t および第1マイクロフォン11aと第2マイクロフォン11bが離間する距離 D に基づいて、装着者と話者との向き合う角度である対面角度 θ を求める(ステップ105)。この対面角度 θ の情報は、データ送信部16によりホスト装置20に出力する(ステップ106)。

【0044】

< 端末装置の具体例 >

図8(a)~(b)は、実際に使用される端末装置10の具体例を示した図である。

このうち図8(a)は、図2に示した端末装置10と同様のものであり、第1マイクロフォン11a、第2マイクロフォン11b、第3マイクロフォン11cの3つのマイクロフォンを配したものである。そして第1マイクロフォン11aと第2マイクロフォン11bとの距離 D_{12} を10cmとし、第2マイクロフォン11bと第3マイクロフォン11

10

20

30

40

50

cの距離 D_{23} および第3マイクロフォン11cと第1マイクロフォン11aの距離 D_{31} を共に35cmとしている。

【0045】

図8(a)のようにマイクロフォンが3箇所以上に配された端末装置10を使用することで、別々のマイクロフォンの組を使用して、上述した自他識別の判断、および対面角度の出力を行なうことができる。これにより音声解析部15で自他識別の判断や対面角度の出力を行なうためのアルゴリズムをより簡略化することができる。

さらにこの形態によれば、複数の対面角度を出力することができる。つまり3つのマイクロフォンにより2つのマイクロフォンからなる組を3つ構成することができるため、これにより3通りの対面角度が出力できることになる。よって本実施の形態では、水平方向の角度のみならず、第2マイクロフォン11bと第3マイクロフォン11cの組または第3マイクロフォン11cと第1マイクロフォン11aの組を使用することで、上下方向の対面角度が可能となる。つまり本構成によれば、音声解析部15(図1参照)は、2つのマイクロフォンからなる組を2組以上使用することにより複数の対面角度を出力することができる。

【0046】

なお、上述した自他識別の判断、および対面角度の出力を行なうために、必要なマイクロフォンの個数は、3つであるとは限らず、3つ以上でもよい。さらに2つでもよい。

図8(b)は、マイクロフォンが2箇所に配された端末装置10について説明した図である。

図8(b)に示した端末装置10では、図8(a)に示した端末装置10に対し、第1マイクロフォン11aは同様の位置に配置されている。一方、第2マイクロフォン11bは、図8(a)における第3マイクロフォン11cがあった位置に配置されている。つまり第2マイクロフォン11bは、ケース31に設けられ、第1マイクロフォン11aとの距離 D_{12} を35cmとしている。さらに図8(a)における第2マイクロフォン11bがあった位置には、マイクロフォンは配置されていない。この端末装置10では、第1マイクロフォン11aと第2マイクロフォン11bは、装着者の口からの距離が互いに異なる2つのマイクロフォン、および水平方向の距離が離れた2つのマイクロフォンの双方を兼ねている。

【0047】

この構成において自他識別の判断を行なうには、第1マイクロフォン11aと第2マイクロフォン11bを使用し、上述した場合と同様に行なうことができる。

一方、対面角度についても同様に出力することができる。ただしこの対面角度は、第1マイクロフォン11aと第2マイクロフォン11bとを結ぶ線を通る面を基準としたものとなる。よって水平面を基準とした対面角度を求めるためには、第1マイクロフォン11aと第2マイクロフォン11bとを結ぶ線と水平面との角度を予め把握しておき、それを基に角度の補正をすればよい。ただし、この補正は必ずしも行なう必要はない。

【0048】

以上詳述したような対面角度出力システム1により、対面角度の情報を装着者と話者との関係を判定する情報として利用できる。

ここで装着者と話者との関係として判定を行なう事項としては、例えば、装着者と他者とのコミュニケーション関係である。例えば、装着者と話者が近い距離に位置し、さらに対面角度の情報により向き合っていることがわかれば、装着者と話者が会話を行なっている可能性が高い。また装着者と話者が反対方向を向いていれば、会話を行っていない可能性が高い。実際には、他に話者や装着者の音声を取得されるタイミングや、間合いなど他の情報も使用して判定が行なわれる。また装着者と話者との関係として、上下方向の対面角度を使用して、一方が他方を上から見下ろしている関係にあるなどの判定も可能である。また複数の端末装置10からの情報をもとに、会話を行っている複数人の間での位置の相互関係などの処理をしても良い。

【0049】

また上述した例では、自他識別の判断や対面角度の出力を端末装置 10 で行なっていたが、これに限られるものではなく、ホスト装置 20 の方で行なってもよい。この形態における対面角度出力システム 1 としては、図 1 のものに対し、音声解析部 15 で行なっていた自他識別の判定および対面角度の出力を、例えば、ホスト装置 20 のデータ解析部 23 で行なう。この対面角度出力システム 1 においては、音声解析部 15 は不要となる。そしてこの場合データ解析部 23 は、話者が装着者か他者かを識別する識別手段として機能するとともに、装着者と話者との向き合う角度である対面角度を出力する出力手段として機能することになる。さらにこの場合、端末装置 10 のデータ送信部 16 は、マイクロフォン 11 a、11 b、11 c により取得された音声の情報を、音声解析部 15 を介さずを受け、データ送信部 21 に対し送信する。またホスト装置 20 のデータ受信部 21 は、データ送信部 16 により送信された音声の情報を受信する。即ちデータ送信部 16 は、マイクロフォン 11 a、11 b、11 c により取得された音声の情報を送信する送信手段として機能し、データ受信部 21 は、データ送信部 16 により送信された音声の情報を受信する受信手段として機能する。

10

【0050】

<プログラムの説明>

なお図 7 で説明を行った本実施の形態における端末装置 10 が行なう処理は、ソフトウェアとハードウェア資源とが協働することにより実現される。即ち、端末装置 10 に設けられた制御用コンピュータ内部の図示しない CPU が、端末装置 10 の各機能を実現するプログラムを実行し、これらの各機能を実現させる。

20

【0051】

よって図 7 で説明を行った端末装置 10 が行なう処理は、コンピュータに、装着者の口からの距離が互いに異なる 2 つおよび水平方向の距離が離れた 2 つを少なくとも含む、話者の音声を取得するための複数のマイクロフォンから音声の情報を取得する機能と、複数のマイクロフォンのうち、装着者の口からの距離が互いに異なる 2 つのマイクロフォンで取得される音声の音圧差に基づいて、話者が装着者か他者かを識別する機能と、話者が他者と識別された場合に、複数のマイクロフォンのうち水平方向の距離が離れた 2 つのマイクロフォンに話者の音声が届く時間差を導出する機能と、導出された時間差および水平方向の距離が離れた 2 つのマイクロフォン間の距離に基づいて、装着者と話者との向き合う角度である対面角度を出力する機能と、を実現させるプログラムとして捉えることもできる。

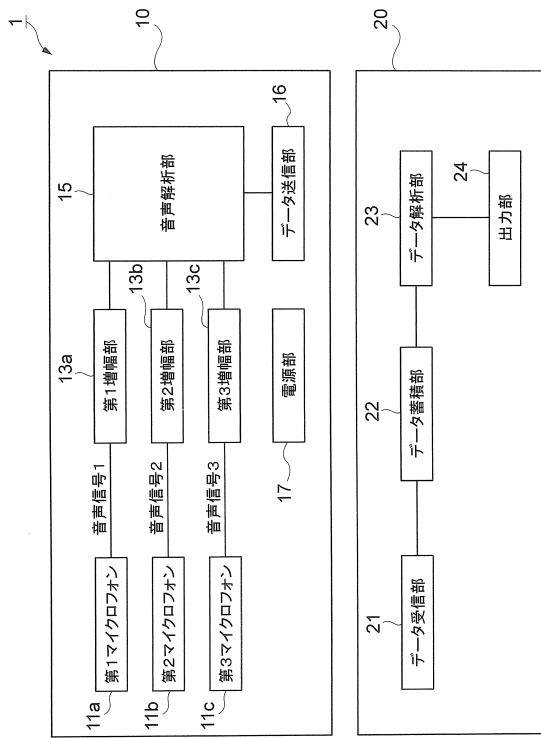
30

【符号の説明】

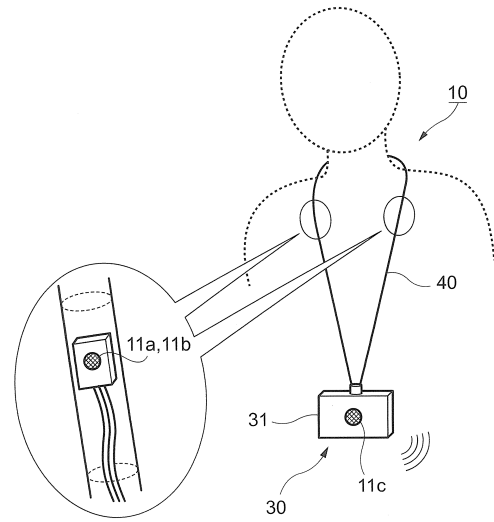
【0052】

1 ... 対面角度出力システム、10 ... 端末装置、11 a ... 第 1 マイクロフォン、11 b ... 第 2 マイクロフォン、11 c ... 第 3 マイクロフォン、15 ... 音声解析部、16 ... データ送信部、20 ... ホスト装置、30 ... 装置本体、40 ... 提げ紐

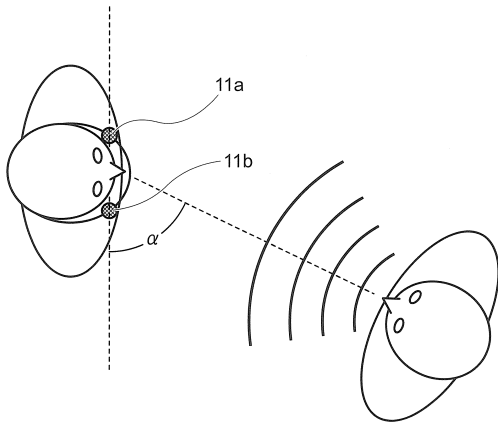
【図1】



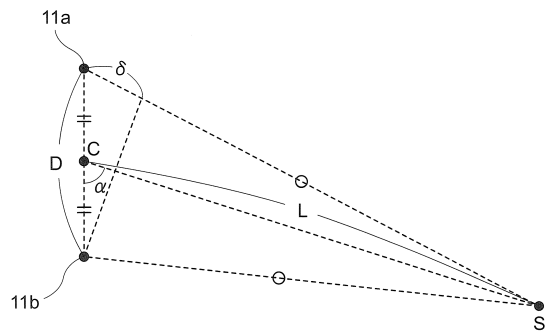
【図2】



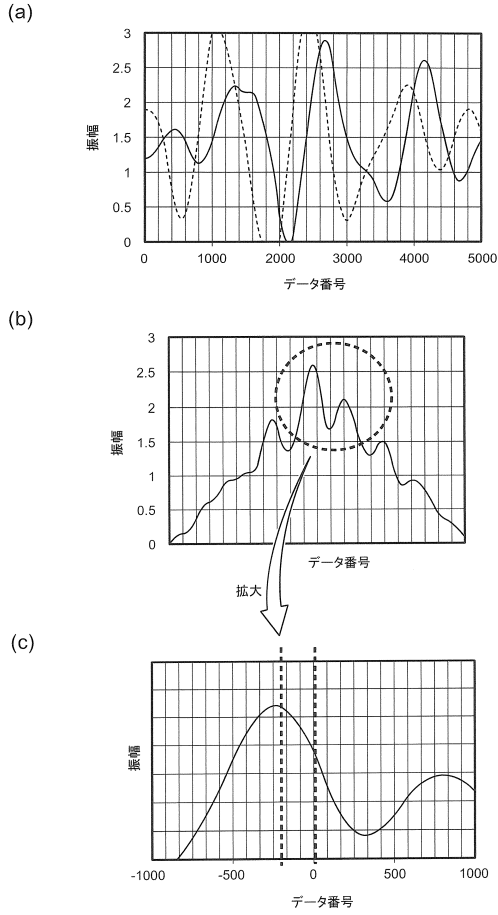
【図3】



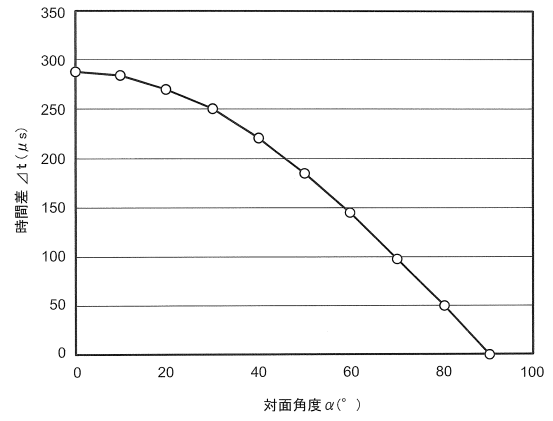
【図4】



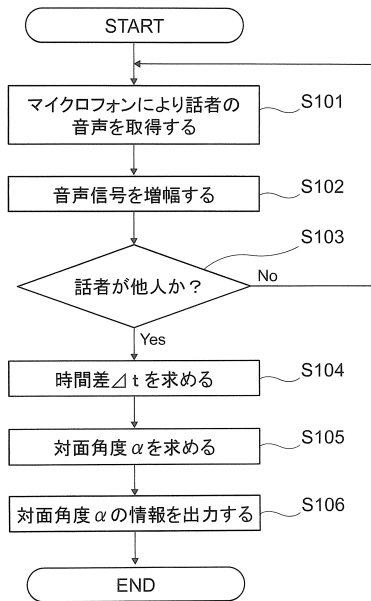
【図5】



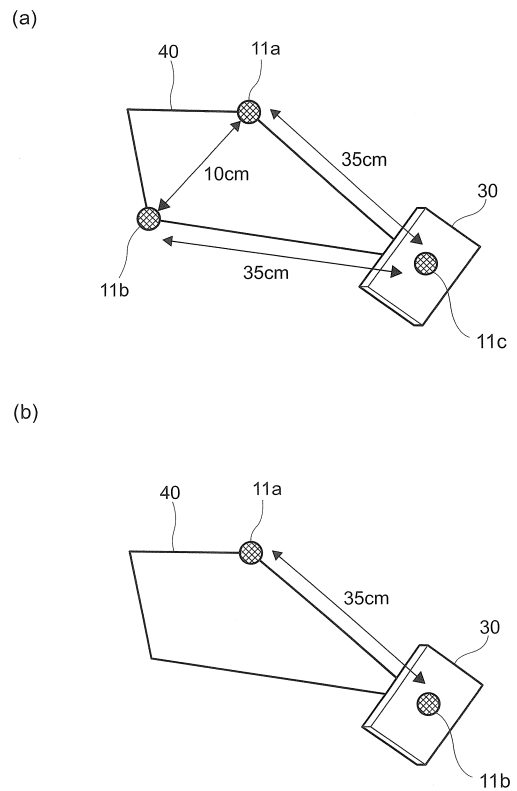
【図6】



【図7】



【図8】



フロントページの続き

- (72)発明者 米山 博人
神奈川県足柄上郡中井町境430 グリーンテクなかい 富士ゼロックス株式会社内
- (72)発明者 下谷 啓
神奈川県足柄上郡中井町境430 グリーンテクなかい 富士ゼロックス株式会社内
- (72)発明者 飯田 靖
神奈川県足柄上郡中井町境430 グリーンテクなかい 富士ゼロックス株式会社内
- (72)発明者 内藤 孝雄
神奈川県足柄上郡中井町境430 グリーンテクなかい 富士ゼロックス株式会社内

審査官 富澤 直樹

- (56)参考文献 特開2007-264473(JP,A)
特開2001-236092(JP,A)
特開平11-304906(JP,A)
特開2010-026361(JP,A)
特開2007-108518(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04R 3/00
H04R 1/06
H04R 1/40
G10L 15/00 - 15/34
G10L 21/0208 - 21/0308
G01S 3/808
G01S 5/22