

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6973652号  
(P6973652)

(45) 発行日 令和3年12月1日(2021.12.1)

(24) 登録日 令和3年11月8日(2021.11.8)

(51) Int. Cl.		F I			
<b>G 1 0 L</b>	<b>15/20</b>	<b>(2006.01)</b>	<b>G 1 0 L</b>	15/20	2 0 0 Z
<b>G 1 0 L</b>	<b>15/28</b>	<b>(2013.01)</b>	<b>G 1 0 L</b>	15/28	4 0 0
<b>G 1 0 L</b>	<b>15/32</b>	<b>(2013.01)</b>	<b>G 1 0 L</b>	15/32	2 2 0 Z

請求項の数 8 (全 12 頁)

(21) 出願番号	特願2020-538869 (P2020-538869)	(73) 特許権者	000004237
(86) (22) 出願日	平成30年2月14日 (2018.2.14)		日本電気株式会社
(65) 公表番号	特表2021-510427 (P2021-510427A)		東京都港区芝五丁目7番1号
(43) 公表日	令和3年4月22日 (2021.4.22)	(74) 代理人	100103090
(86) 国際出願番号	PCT/JP2018/005057		弁理士 岩壁 冬樹
(87) 国際公開番号	W02019/159253	(74) 代理人	100124501
(87) 国際公開日	令和1年8月22日 (2019.8.22)		弁理士 塩川 誠人
審査請求日	令和2年7月13日 (2020.7.13)	(72) 発明者	ワン チョンチョン
			東京都港区芝五丁目7番1号 日本電気株式会社社内
		(72) 発明者	越仲 孝文
			東京都港区芝五丁目7番1号 日本電気株式会社社内
		審査官	富澤 直樹

最終頁に続く

(54) 【発明の名称】 音声処理装置、方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

空気を介して音声を取得する気導マイクロフォンからの音声を認識する気導マイクロフォン音声認識手段と、

ウェアラブルマイクロフォンからの音声を認識するウェアラブルマイクロフォン音声認識手段と、

少なくとも前記気導マイクロフォンと話者との間の距離を含む環境状況を測定する検知手段と、

前記環境状況に基づいて、前記気導マイクロフォン音声認識手段の認識結果の重みおよび前記ウェアラブルマイクロフォン音声認識手段の認識結果の重みを計算する重み決定手段と、

前記重みを用いて、前記気導マイクロフォン音声認識手段から出力された認識結果と前記ウェアラブルマイクロフォン音声認識手段から出力された認識結果とを結合する結合手段とを備えた音声処理装置。

【請求項2】

前記重み決定手段は、前記気導マイクロフォンと話者との間の距離に従って、前記ウェアラブルマイクロフォン音声認識手段の認識結果の重みを決定する

請求項1に記載の音声処理装置。

【請求項3】

前記気導マイクロフォン音声認識手段の認識結果の重みは、前記気導マイクロフォンと

話者との間の距離に関して、単調非減少関数である

請求項 2 に記載の音声処理装置。

【請求項 4】

前記気導マイクロフォン音声認識手段および前記ウェアラブルマイクロフォン音声認識手段は、音声認識、話者認識、言語認識、方言認識、感情認識、性別認識、年齢認識のいずれかを実行する。

請求項 1 から請求項 3 のうちのいずれか 1 項に記載の音声処理装置。

【請求項 5】

空気を介して音声を取得する気導マイクロフォンからの音声を認識し、  
ウェアラブルマイクロフォンからの音声を認識し、  
少なくとも前記気導マイクロフォンと話者との間の距離を含む環境状況を測定し、  
前記環境状況に基づいて、前記気導マイクロフォンからの音声を認識した結果の重みと前記ウェアラブルマイクロフォンからの音声を認識した結果の重みとを計算し、  
前記重みを用いて、前記気導マイクロフォンからの音声を認識した結果と前記ウェアラブルマイクロフォンからの音声を認識した結果とを結合する  
音声処理方法。

10

【請求項 6】

前記ウェアラブルマイクロフォンからの音声を認識した結果の重みを、前記気導マイクロフォンと話者との間の距離に従って決定する

請求項 5 に記載の音声処理方法。

20

【請求項 7】

コンピュータに

空気を介して音声を取得する気導マイクロフォンからの音声を認識する処理と、  
ウェアラブルマイクロフォンからの音声を認識する処理と、  
少なくとも前記気導マイクロフォンと話者との間の距離を含む環境状況を測定する処理と、

前記環境状況に基づいて、前記気導マイクロフォンからの音声を認識した結果の重みと前記ウェアラブルマイクロフォンからの音声を認識した結果の重みとを計算する処理と、

前記重みを用いて、前記気導マイクロフォンからの音声を認識した結果と前記ウェアラブルマイクロフォンからの音声を認識した結果とを結合する処理と

を実行させるための音声処理プログラム。

30

【請求項 8】

前記ウェアラブルマイクロフォンからの音声を認識した結果の重みは、前記気導マイクロフォンと話者との間の距離に従って決定される

請求項 7 に記載の音声処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声処理装置、音声処理方法および音声処理プログラムに関する。

【背景技術】

【0002】

音声処理において音声を取得するために使用される最も一般的な方法は、マイクロフォンの使用である。気導マイクロフォン (air conduction microphone)、咽喉マイクロフォン、骨伝導マイクロフォン、皮膚伝導マイクロフォンなど、さまざまな伝導方式に応じた多くの種類のマイクロフォンがある。それらのうち、空気を介して伝播する音声を取得する気導マイクロフォンは、音声処理分野で最も一般的に使用されている。気導マイクロフォンは、一般的な音声処理システムが良好な性能を得ることができる全帯域、例えば 20 MHz の音声を捉える。しかし、気導マイクロフォンは、環境ノイズおよび話者と見なされる音源からの距離の影響を受けやすい (非特許文献 1 参照)。そのような厳しい状況での性能を改善する方法は、音声処理研究者にとって最も興味をかき立てるテーマの 1 つ

40

50

である。特に、現在、グーグルアシスタント (Google Assistant)、アマゾンアレクサ (Amazon Alexa)、アップルシリ (Apple Siri) などの A I スピーカの急速な普及によって、最新かつ緊急の解決課題になっている (Google、Amazon、Alexa、Apple、Siri はそれぞれ登録商標)。

【0003】

気導マイクロフォンの代表例は、会議室やホールなどの空間で使用される集音マイクロフォンである。ただし、気導マイクロフォンの利用は、会議やコンベンションなどに限定されない。

【0004】

一方、「ウェアラブル」デバイスの人気が高まっている。ウェアラブルデバイスは、通信目的のワイヤレス送信、音楽の聴取、身体活動量の計測、生体認証など、多目的に使用できる電子デバイスである。ウェアラブルデバイスのハードウェア構成には、一般に、マイクロフォンが含まれる。本明細書では、ウェアラブルデバイスのマイクロフォンを「ウェアラブルマイクロフォン」という。一般に、「ウェアラブルマイクロフォン」は、周囲のノイズに対して感度が低く、常に話者に近接すなわち近距離にあるので、上述したようなノイズが多い環境で使用できる。しかし、皮膚や骨を伝導する高周波成分が伝送損失により大幅に減衰するので、ウェアラブルマイクロフォンで得られる音声信号の品質は比較的低い。非特許文献 2 には、ウェアラブルマイクロフォンの性能が、ノイズがない状態での話者認識アプリケーションで用いられる気導マイクロフォンよりも劣化することが示されている。

【先行技術文献】

【特許文献】

【0005】

【特許文献 1】特開 2003 - 264883 号公報

【非特許文献】

【0006】

【非特許文献 1】J. Ming, T. J. Hazen, J. R. Glass, and D. A. Reynolds, "Robust speaker recognition in noisy conditions", Audio, Speech, and Language Processing, IEEE Transactions, 2007

【非特許文献 2】S. Tsuge, S. Kuroiwa, "Speaker recognition using multiple types of speech", The Seventh Symposium on Biometrics, Recognition and Authentication, 2017

【非特許文献 3】E. Georgati, T. May, S. Par., et al. "Speaker Distance Detection Using a Single Microphone", IEEE TASLP 2011

【非特許文献 4】J.G. Fiscus, "A post-processing system to yield reduced word error rates: Recognizer Output Voting Error Reduction (ROVER)", Automatic Speech Recognition and Understanding, 1997

【発明の概要】

【発明が解決しようとする課題】

【0007】

気導マイクロフォンのみが使用されている非特許文献 1 には、ノイズの影響や残響ノイズを発生させる話者と気導マイクロフォンとの距離が遠いと、音声処理の性能が低下することが示されている。音声処理に使用される気導マイクロフォンは、遠距離やノイズに弱いことが証明されている。非特許文献 2 では、インイヤ型骨伝導マイクロフォン (ウェアラブルマイクロフォンの一種) と気導マイクロフォンとが比較され、インイヤ型骨伝導マイクロフォンは、ノイズがなく近接した状態での話者認識において、気導マイクロフォンと競争にならないことが示されている。

【0008】

本発明の一態様は、上記課題を解決することを目的とした、マイクロフォンの種類を活かして、異なるマイクロフォンの組み合わせを利用する音声処理装置である。

## 【0009】

なお、特許文献1には、気導音響マイクロフォンと骨伝導音響マイクロフォンという異なるマイクロフォンを使用することが開示されている。しかし、特許文献1に開示された装置は、騒音測定の結果に応じて、気導音響マイクロフォンから出力される音声信号を気導音響マイクロフォンから出力される音声信号に加算する。この装置は、気導音響マイクロフォンが話者から離れているときに通常発生する残響の影響を考慮していない。当該装置で実行されているような、異なるマイクロフォンからの音声信号の追加は、必ずしも認識性能を向上させるわけではない。なお、認識システムは、一般に、音声信号を捉えるマイクロフォンに合わせて調整または最適化される。

## 【課題を解決するための手段】

10

## 【0010】

音声処理装置の例示的な態様は、空気を介して音声を取得する気導マイクロフォンからの音声を認識する気導マイクロフォン音声認識手段と、ウェアラブルマイクロフォンからの音声を認識するウェアラブルマイクロフォン音声認識手段と、少なくとも気導マイクロフォンと話者との間の距離を含む環境状況を測定する検知手段と、環境状況に基づいて、気導マイクロフォン音声認識手段の認識結果の重みおよびウェアラブルマイクロフォン音声認識手段の認識結果の重みを計算する重み決定手段と、重みを用いて、気導マイクロフォン音声認識手段から出力された認識結果とウェアラブルマイクロフォン音声認識手段から出力された認識結果とを結合する結合手段とを備える。

## 【0011】

20

音声処理方法の例示的な態様は、空気を介して音声を取得する気導マイクロフォンからの音声を認識し、ウェアラブルマイクロフォンからの音声を認識し、少なくとも気導マイクロフォンと話者との間の距離を含む環境状況を測定し、環境状況に基づいて、気導マイクロフォンからの音声を認識した結果の重みとウェアラブルマイクロフォンからの音声を認識した結果の重みとを計算し、重みを用いて、気導マイクロフォンからの音声を認識した結果とウェアラブルマイクロフォンからの音声を認識した結果とを結合する。

## 【0012】

音声処理プログラムの例示的な態様は、コンピュータに、空気を介して音声を取得する気導マイクロフォンからの音声を認識する処理と、ウェアラブルマイクロフォンからの音声を認識する処理と、少なくとも気導マイクロフォンと話者との間の距離を含む環境状況を測定する処理と、環境状況に基づいて、気導マイクロフォンからの音声を認識した結果の重みとウェアラブルマイクロフォンからの音声を認識した結果の重みとを計算する処理と、重みを用いて、気導マイクロフォンからの音声を認識した結果とウェアラブルマイクロフォンからの音声を認識した結果とを結合する処理とを実行させる。

30

## 【発明の効果】

## 【0013】

本発明によれば、音声処理装置、音声処理方法およびプログラムが、各種のマイクロフォンを利用することによって、厳しい条件において音声処理についてより優れた性能を発揮することができる。

## 【図面の簡単な説明】

40

## 【0014】

【図1】第1の実施形態の音声処理装置のブロック図である。

【図2】第1の実施形態の音声処理装置の動作を示すフローチャートである。

【図3】音声処理装置の機能を実現可能な情報処理システムの構成例を示すブロック図である。

【図4】音声処理装置の主要部を示すブロック図である。

## 【発明を実施するための形態】

## 【0015】

以下、本発明の一実施形態を、図面を参照して説明する。以下の詳細な説明は単なる例示であり、本発明または本発明の用途および使用を限定することを意図していない。さら

50

に、上記の発明の背景または以下の詳細な説明に示されている考え方に拘束される意図はない。さらに、図面は、詳細な説明とともに、本発明の方法の考え方を説明するためのものである。図面は、例示のためのものであり、技術の適用を制限するものではない。

#### 【0016】

##### 第1の実施形態

第1の実施形態の音声処理装置は、ノイズや話者と気導マイクロフォンとの間の遠距離といった厳しい状況において、高性能の音声処理を提供することができる。

#### 【0017】

##### <音声処理装置の構成>

図1は、第1の実施形態の音声処理装置100のブロック図である。音声処理装置100は、気導マイクロフォン音声認識部(一般的な気導マイクロフォンを用いた認識部)101、ウェアラブルマイクロフォン音声認識部102、検知部103、重み決定部104、および結合部105を備える。

#### 【0018】

気導マイクロフォン音声認識部101は、音声処理を、気導マイクロフォン200からの音声データを用いる認識に適用する。気導マイクロフォン音声認識部101は、発言者を予測する話者認識の場合、気導マイクロフォン200からの音声データを用いて、話者候補のスコアを計算する。スコアとして対数尤度がよく用いられる。音声内容を予測する音声認識の場合、気導マイクロフォン音声認識部101は、気導マイクロフォン200からの音声を用いて、言語単位、例えば音素、音節、単語の事後確率を算出する。どのような種類の音声処理も可能である。例えば、音声認識、感情認識、言語認識、方言認識、年齢認識、性別認識などである。すべての認識の詳細は、当業者にはよく知られている。

#### 【0019】

ウェアラブルマイクロフォン音声認識部102は、音声処理を、ウェアラブルマイクロフォン300からの音声データを用いる認識に適用する。ウェアラブルマイクロフォン300の一例は、インイヤ型骨伝導マイクロフォンである。話者認識の場合、ウェアラブルマイクロフォン音声認識部102は、ウェアラブルマイクロフォン300からの音声データを用いて、話者候補のスコアを計算する。音声認識の場合、ウェアラブルマイクロフォン音声認識部102は、ウェアラブルマイクロフォン300からの音声データを用いて、言語単位の事後確率をスコアとして算出する。

#### 【0020】

検知部103は、信号雑音比(SNR)および気導マイクロフォン200と話者との間の距離を含む環境状況を測定する。SNR測定のために、検知部103は、まず、気導マイクロフォン音声における音声区間および非音声区間を検出し、次に、音声区間の平均振幅と非音声区間の平均振幅との比を計算する。距離測定のために、種々の方法を使用可能である。例えば、図示されていない、ウェアラブルマイクロフォン300とともに組込まれる宅内位置センサが、話者の位置を取得し、それと気導マイクロフォン200の既知の位置との間の距離を計算する。非特許文献3に開示されているような、単一のマイクロフォンを使用した話者距離検出を使用してもよい。

#### 【0021】

重み決定部104は、検知部103による検知結果に応じて重みを決定する。ここでの重みは、気導マイクロフォン音声認識部101およびウェアラブルマイクロフォン音声認識部102から出力された結果を用いた重みである。例えば、重み $w_a$ および $w_s$ は、次の(1)式および(2)式を使用して計算される。

#### 【0022】

$$w_a = a + c \cdot \text{SNR} \quad (1)$$

$$w_s = b + d \cdot \text{distance} \quad (2)$$

#### 【0023】

ここで、「距離」は、検知部103が測定した気導マイクロフォン200と話者との距離である。 $w_a$ は、気導マイクロフォン音声認識部101からのスコア $s_a$ ごとの重みで

10

20

30

40

50

ある。 $w_s$  は、ウェアラブルマイクロフォン音声認識部 102 からのスコア  $s_s$  ごとの重みである。 $a$ 、 $b$ 、 $c$  および  $d$  は、実験で決定可能な定数である。 $c$  および  $d$  は、正数である。

【0024】

ここで示される例は、重み  $w_a$  および  $w_s$  の計算における 1 次多項式関数に基づいている。ただし、SNR と距離に関して、当然、単調非減少である他の関数を使用することができる。

【0025】

そのような関数の別の例は、0 または 1 の 2 つの値のみを取ることができるステップ関数である。その場合、気導マイクロフォン音声認識部 101 またはウェアラブルマイクロフォン音声認識部 104 が、SNR および距離に従って排他的に選択される。

10

【0026】

結合部 105 は、重み決定部 104 から出力された重みを用いて、気導マイクロフォン音声認識部 101 からのスコアとウェアラブルマイクロフォン音声認識部 102 からのスコアとを結合し、最終的な認識結果を出力する。話者認識の場合、結合部 105 は、例えば、次の式 (3) を用いて、話者候補ごとに、結合スコア  $s$  を算出する。

【0027】

$$s = (w_a s_a + w_s s_s) / (w_a + w_s) \quad (3)$$

【0028】

結合部 105 は、話者候補全員のスコアを比較することによって、スコアが最も高い話者を示す情報 (データ) を認識結果として出力する。音声認識の場合、非特許文献 4 における Recognizer Output Voting Error Reduction (ROVER) を結合部 105 で使用することができる。

20

【0029】

< 音声処理装置の操作 >

次に、図 2 を参照して、音声処理装置 100 の動作を説明する。

【0030】

ステップ S11 において、気導マイクロフォン音声認識部 101 およびウェアラブルマイクロフォン音声認識部 102 は、音声データを受信する。

【0031】

ステップ S12 において、気導マイクロフォン音声認識部 101 は、気導マイクロフォン 200 からの音声データを用いて認識結果を推定する。すなわち、気導マイクロフォン音声認識部 101 は、例えば、スコア  $s_a$  を算出する。話者認識の場合、気導マイクロフォン音声認識部 101 は、話者候補のスコアを計算する。音声認識、感情認識、言語認識、方言認識、年齢認識、または性別認識の場合、気導マイクロフォン音声認識部 101 は、所定の音声候補、感情候補、単語候補、方言候補、各年齢についての単語候補、または、各性別についての単語候補のスコアを計算する。

30

【0032】

ステップ S13 において、ウェアラブルマイクロフォン音声認識部 102 は、気導マイクロフォン音声認識部 101 と同様に、ウェアラブルマイクロフォン 300 からの音声データを用いて、例えば認識結果を推定する。すなわち、ウェアラブルマイクロフォン音声認識部 102 は、例えば、スコア  $s_s$  を算出する。

40

【0033】

ステップ S14 において、検知部 103 は、SNR および気導マイクロフォンと話者との間の距離を含む環境状況を測定する。

【0034】

ステップ S15 において、重み決定部 104 は、気導マイクロフォン音声認識部 101 の認識結果およびウェアラブルマイクロフォン音声認識部 102 の認識結果の重みを決定する。

【0035】

50

なお、図2に示された形態に限定されず、ステップA02～A03とステップA04～A05の順番を入れ替えることも可能であり、ステップS12とステップS13とを入れ替えることも可能である。

【0036】

ステップS16において、結合部105は、重み決定部104から出力された重みを反映しながら、気導マイクロフォン認識音声部101から出力された認識結果とウェアラブルマイクロフォン音声認識部102から出力された認識結果とを結合する。

【0037】

以上に説明したように、第1の実施形態によれば、厳しい状況での認識のための音声処理の性能を向上させることができる。全帯域の音声を拾う気導マイクロフォン200と、周囲のノイズに強いウェアラブルマイクロフォン300との両方が利用される。つまり、上記の実施形態の音声処理装置、音声処理方法およびプログラムは、全帯域の音声を拾う気導マイクロフォン200と、周囲のノイズに強いウェアラブルマイクロフォン300という異なる種類のマイクロフォンを利用して、厳しい状況下での音声処理の性能を向上させることができる。

【0038】

最後に、ここで説明および図示されたプロセス、技術、および方法は、特定の装置に限定または関連していないことは明らかである。コンポーネントの組み合わせを使用して実装可能である。また、ここでの教示に従って、様々なタイプの汎用的な装置を使用可能である。本発明は、特定の例を使用して説明されたが、それらは単なる例示であり、限定を意図するものではない。例えば、記載されたソフトウェアは、C/C++、Java（登録商標）、MATLABおよびPythonなどの多種多様な言語で実装されてもよい。さらに、本発明の技術の他の応用は、当業者には明らかであろう。

【0039】

上記の実施形態は、ハードウェアで構成することができるが、コンピュータプログラムによって実施することも可能である。

【0040】

図3に示す情報処理システムは、CPU（Central Processing Unit）1000、I/O（Input/Output）デバイス1001、メモリ1002などの1つ以上のプロセッサを含む。メモリ1002として、ハードディスクなどの磁気記憶媒体、または、ROM（read only memory）やフラッシュメモリなどの半導体メモリを用いることができる。I/Oデバイス1001は、入力音声データのための出力部と認識結果を出力するための入力部の機能を実現可能である。

【0041】

メモリ1002は、図1に示す音声処理装置100の各部の機能を実現するためのプログラムを記憶する。メモリ1002において、少なくともプログラムが記憶される領域は、非一時的な有形の記憶領域（非一時的な記憶領域）である。また、メモリ1002は、一時的な情報記憶部を実現しうる。CPU1000は、メモリ1002に格納されたプログラムに従って処理を実行することによって、図1に示された気導マイクロフォン音声認識部101、ウェアラブルマイクロフォン音声認識部102、検知部103、重み決定部104、および結合部105の機能を実現する。

【0042】

図4は、本発明による音声処理装置の主要部を示すブロック図である。図4に示すように、音声処理装置10は、空気を介して音声を取得する気導マイクロフォンからの音声を認識する気導マイクロフォン音声認識手段11（実施形態では、気導マイクロフォン音声認識部101で実現される。）と、ウェアラブルマイクロフォンからの音声を認識するウェアラブルマイクロフォン音声認識手段12（実施形態では、ウェアラブルマイクロフォン音声認識部102で実現される。）と、少なくとも気導マイクロフォンと話者との間の距離を含む環境状況を測定する検知手段13（実施形態では、検知部103で実現される。）と、環境状況に基づいて、気導マイクロフォン音声認識手段11の認識結果の重みお

10

20

30

40

50

よびウェアラブルマイクロフォン音声認識手段 1 2 の認識結果の重みを計算する重み決定手段 1 4 (実施形態では、重み決定部 1 0 4 で実現される。)と、重みを用いて、気導マイクロフォン音声認識手段 1 1 から出力された認識結果とウェアラブルマイクロフォン音声認識手段 1 2 から出力された認識結果とを結合する結合手段 1 5 (実施形態では、結合部 1 0 5 で実現される。)とを備えている。

【 0 0 4 3 】

上記の実施形態の一部または全部は、以下の付記のようにも記載され得るが、以下に限定されるわけではない。

【 0 0 4 4 】

(付記 1) 空気を介して音声を取得する気導マイクロフォンからの音声を認識する気導マイクロフォン音声認識手段と、

ウェアラブルマイクロフォンからの音声を認識するウェアラブルマイクロフォン音声認識手段と、

少なくとも前記気導マイクロフォンと話者との間の距離を含む環境状況を測定する検知手段と、

前記環境状況に基づいて、前記気導マイクロフォン音声認識手段の認識結果の重みおよび前記ウェアラブルマイクロフォン音声認識手段の認識結果の重みを計算する重み決定手段と、

前記重みを用いて、前記気導マイクロフォン音声認識手段から出力された認識結果と前記ウェアラブルマイクロフォン音声認識手段から出力された認識結果とを結合する結合手段とを備えた音声処理装置。

【 0 0 4 5 】

(付記 2) 前記重み決定手段は、前記気導マイクロフォンと話者との間の距離に従って、前記ウェアラブルマイクロフォン音声認識手段の認識結果の重みを決定する

付記 1 に記載の音声処理装置。

【 0 0 4 6 】

(付記 3) 前記気導マイクロフォン音声認識手段の認識結果の重みは、前記気導マイクロフォンと話者との間の距離に関して、単調非減少関数である

付記 2 に記載の音声処理装置。

【 0 0 4 7 】

(付記 4) 前記気導マイクロフォン音声認識手段および前記ウェアラブルマイクロフォン音声認識手段は、音声認識、話者認識、言語認識、方言認識、感情認識、性別認識、年齢認識のいずれかを実行する

付記 1 から付記 3 のうちのいずれかに記載の音声処理装置。

【 0 0 4 8 】

(付記 5) 空気を介して音声を取得する気導マイクロフォンからの音声を認識し、

ウェアラブルマイクロフォンからの音声を認識し、

少なくとも前記気導マイクロフォンと話者との間の距離を含む環境状況を測定し、

前記環境状況に基づいて、前記気導マイクロフォンからの音声を認識した結果の重みと前記ウェアラブルマイクロフォンからの音声を認識した結果の重みとを計算し、

前記重みを用いて、前記気導マイクロフォンからの音声を認識した結果と前記ウェアラブルマイクロフォンからの音声を認識した結果とを結合する

音声処理方法。

【 0 0 4 9 】

(付記 6) 前記ウェアラブルマイクロフォンからの音声を認識した結果の重みを、前記気導マイクロフォンと話者との間の距離に従って決定する

付記 5 に記載の音声処理方法。

【 0 0 5 0 】

(付記 7) 前記気導マイクロフォンからの音声を認識した結果の重みは、前記気導マイクロフォンと話者との間の距離に関して、単調非減少関数である

10

20

30

40

50

付記 6 に記載の音声処理方法。

【 0 0 5 1 】

( 付記 8 ) コンピュータに

空気を介して音声を取得する気導マイクロフォンからの音声を認識する処理と、

ウェアラブルマイクロフォンからの音声を認識する処理と、

少なくとも前記気導マイクロフォンと話者との間の距離を含む環境状況を測定する処理と、

前記環境状況に基づいて、前記気導マイクロフォンからの音声を認識した結果の重みと前記ウェアラブルマイクロフォンからの音声を認識した結果の重みとを計算する処理と、

前記重みを用いて、前記気導マイクロフォンからの音声を認識した結果と前記ウェアラブルマイクロフォンからの音声を認識した結果とを結合する処理と

を実行させるための音声処理プログラム。

10

【 0 0 5 2 】

( 付記 9 ) 前記ウェアラブルマイクロフォンからの音声を認識した結果の重みは、前記気導マイクロフォンと話者との間の距離に従って決定される

付記 8 に記載の音声処理プログラム。

【 0 0 5 3 】

( 付記 1 0 ) 前記気導マイクロフォンからの音声を認識した結果の重みは、前記気導マイクロフォンと話者との間の距離に関して、単調非減少関数である

付記 9 に記載の音声処理プログラム。

20

【 0 0 5 4 】

( 付記 1 1 ) プロセッサで実行されるときに、

空気を介して音声を取得する気導マイクロフォンからの音声を認識する処理と、

ウェアラブルマイクロフォンからの音声を認識する処理と、

少なくとも前記気導マイクロフォンと話者との間の距離を含む環境状況を測定する処理と、

前記環境状況に基づいて、前記気導マイクロフォンからの音声を認識した結果の重みと前記ウェアラブルマイクロフォンからの音声を認識した結果の重みとを計算する処理と、

前記重みを用いて、前記気導マイクロフォンからの音声を認識した結果と前記ウェアラブルマイクロフォンからの音声を認識した結果とを結合する処理と

を実行する音声処理プログラムが格納された非一時的なコンピュータ読み取り可能な情報記録媒体。

30

【 0 0 5 5 】

( 付記 1 2 ) 前記ウェアラブルマイクロフォンからの音声を認識した結果の重みは、前記気導マイクロフォンと話者との間の距離に従って決定される

付記 1 1 に記載のコンピュータ読み取り可能な情報記録媒体。

【 0 0 5 6 】

( 付記 1 3 ) 前記気導マイクロフォンからの音声を認識した結果の重みは、前記気導マイクロフォンと話者との間の距離に関して、単調非減少関数である

付記 1 2 に記載のコンピュータ読み取り可能な情報記録媒体。

40

【 0 0 5 7 】

以上、実施形態を参照して本願発明を説明したが、本願発明は上記の実施形態に限定されるものではない。本願発明の構成や詳細には、本願発明のスコープ内で当業者が理解し得る様々な変更をすることができる。

【 符号の説明 】

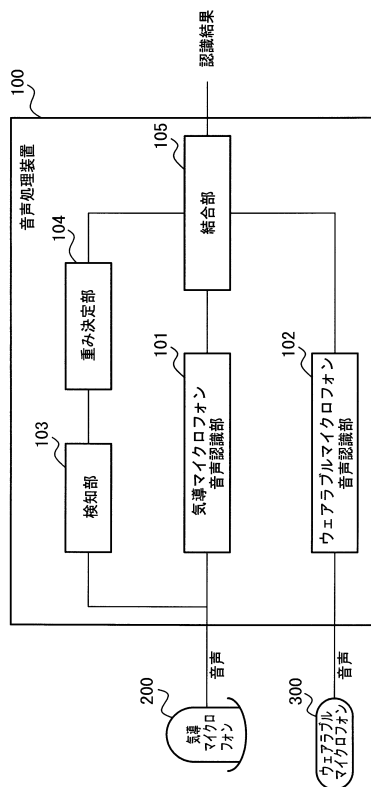
【 0 0 5 8 】

- 1 0 音声処理装置
- 1 1 気導マイクロフォン音声認識手段
- 1 2 ウェアラブルマイクロフォン音声認識手段
- 1 3 検知手段

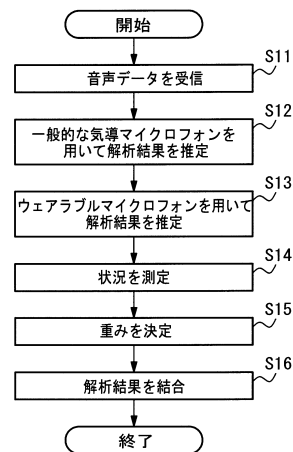
50

- 1 4 重み決定手段
- 1 5 結合手段
- 1 0 0 音声処理装置
- 1 0 1 気導マイクロフォン音声認識部
- 1 0 2 ウェアラブルマイクロフォン音声認識部
- 1 0 3 検知部
- 1 0 4 重み決定部
- 1 0 5 結合部
- 2 0 0 気導マイクロフォン
- 3 0 0 ウェアラブルマイクロフォン
- 1 0 0 0 CPU
- 1 0 0 1 I/Oデバイス
- 1 0 0 2 メモリ

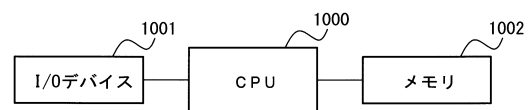
【図1】



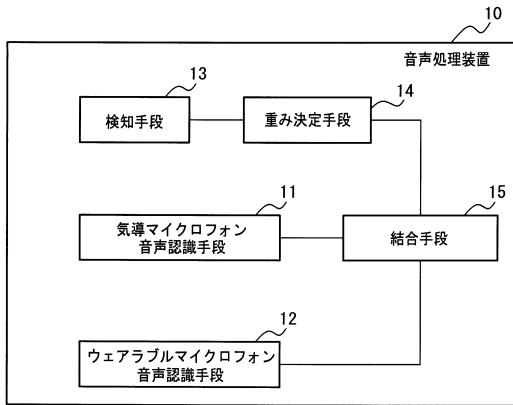
【図2】



【図3】



【図4】



---

フロントページの続き

- (56)参考文献 特開平10-232691(JP,A)  
特開2004-198831(JP,A)  
特開平11-305792(JP,A)  
特開2003-264883(JP,A)  
特開2012-145636(JP,A)  
特開2007-041371(JP,A)  
特開2005-244968(JP,A)  
特開2006-276604(JP,A)  
米国特許出願公開第2017/0025121(US,A1)  
TSUGE, Satoru et al., Speaker verification method using bone-conduction and air-conduction speech, 2009 International Symposium on Intelligent Signal Processing and Communication Systems(ISPACS2009), 米国, IEEE, 2009年12月09日, pp.449-452, ISBN:978-1-4244-5016-9

(58)調査した分野(Int.Cl., DB名)

G10L 15/00 - 15/34