

(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2017-5356

(P2017-5356A)

(43) 公開日 平成29年1月5日(2017.1.5)

(51) Int.Cl.	F I	テーマコード (参考)
HO 4 R 25/00 (2006.01)	HO 4 R 25/00 D	5 C 1 2 2
HO 4 R 3/00 (2006.01)	HO 4 R 25/00 Q	5 D 2 2 0
HO 4 N 5/225 (2006.01)	HO 4 R 3/00 3 2 0	
HO 4 N 5/232 (2006.01)	HO 4 N 5/225 F	
	HO 4 N 5/232 Z	
審査請求 有 請求項の数 15 O L (全 15 頁)		

(21) 出願番号 特願2015-114756 (P2015-114756)
 (22) 出願日 平成27年6月5日 (2015.6.5)

(71) 出願人 514135144
 リウ チン フォン
 Ching-Feng LIU
 台湾高雄市鼓山區美術南五街99巷10號
 No. 10, Ln. 99, Meis
 hu S. 5th St., Gush
 an Dist., Kaohsiung
 City, Taiwan Posta
 l code 80460

最終頁に続く

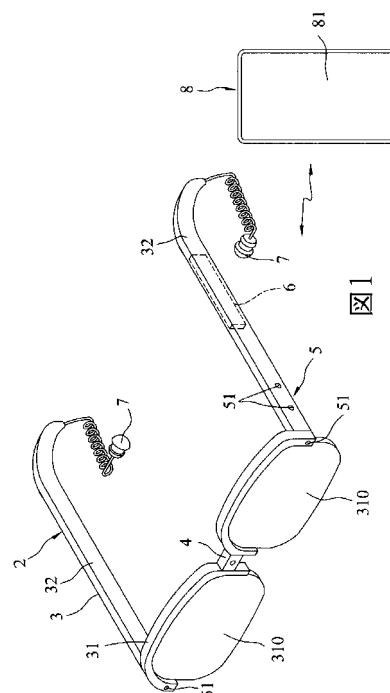
(54) 【発明の名称】 オーディオ信号を処理する方法及び補聴器システム

(57) 【要約】

【課題】複数の話者が居る中でターゲット話者に集中するように補助することができるオーディオ信号を処理する方法及び補聴器システムを提供する。

【解決手段】補聴器システム2は画像取り込みモジュール4、收音モジュール5、プロセッサ6及び音声出力装置7を備えている。装着具3は一对のレンズ310を保持しているフロントサポーター31とユーザーの耳に掛けて支持されるサイドサポーター32とから構成されたメガネ枠機構を有し、画像取り込みモジュール4がフロントサポーター31に、收音モジュール5、プロセッサ6及び音声出力装置7がサイドサポーター32にそれぞれ取り付けられている。プロセッサ6は、画像取り込みモジュール4によって取得された補聴器システム2の周りの画像40をタッチディスプレイ81に映すように外付け電子装置8と通信できる。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

画像取り込みモジュールと収音モジュールとプロセッサとを備えている補聴器システムを用いて実行される方法であって、

前記収音モジュールを用いて、前記補聴器システムの周囲の音声情報を集める、ステップ (a) と、

前記画像取り込みモジュールを用いて、前記補聴器システムの周囲の画像を取得する、ステップ (b) と、

前記プロセッサを用いて、前記収音モジュールによって集められた前記音声情報に指向性信号処理操作を実行し、前記画像取り込みモジュールによって取得された前記画像におけるターゲット対象物の位置に関連するエリアからの抽出されたボイス信号を含む出力オーディオ信号を生成する、ステップ (c) と、

前記出力オーディオ信号を出力する、ステップ (d) と、
を有することを特徴とするオーディオ信号を処理する方法。

【請求項 2】

前記ステップ (c) では、

前記プロセッサを用いて、前記画像取り込みモジュールによって取得された画像における人顔対象物の存在を識別する、サブステップ (c 1) と、

前記プロセッサを用いて、識別された前記人顔対象物それぞれに関連する対象物情報を定める、サブステップ (c 2) と、

前記プロセッサを用いて、識別された前記人顔対象物に関連する対象物情報に基づき、識別された前記人顔対象物それぞれの尤度分類を定めるステップであって、前記尤度分類とは、識別された前記人顔対象物に関連する人がターゲット話者であるとする尤度の分類である、サブステップ (c 3) と、

前記プロセッサを用いて、前記画像における識別された前記人顔対象物の一つをターゲット話者として選択するステップであって、選択された前記人顔対象物の一つは他の前記人顔対象物よりも高い尤度を有するものである、サブステップ (c 4) と、
を有することを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記サブステップ (c 2) では、識別された前記人顔対象物それぞれの前記対象物情報が、

前記画像における識別された前記人顔対象物の前記補聴器システムに対する相対深度、

前記画像取り込みモジュールの一つの軸を表す前記画像における参照軸と前記画像における識別された前記人顔対象物の位置との間に形成されたインクルードドアングル、及び、

識別された前記人顔対象物の前記画像において曲がった角度
の少なくとも一つを有することを特徴とする請求項 2 に記載の方法。

【請求項 4】

前記サブステップ (c 2) では、前記対象物情報は、関連する識別された前記人顔対象物の前記補聴器システムに対する相対深度情報と相対方向情報を含み、

前記相対深度情報において異なる相対深度には異なる重みが予め定められ、

前記相対方向情報において異なる相対方向には異なる重みが予め定められ、

識別された前記人顔対象物それぞれの前記尤度分類は、それぞれの前記相対深度情報及び前記相対方向情報に関連する重みに応じて予め定められ、

前記相対方向情報は、

前記画像取り込みモジュールの一つの軸を表す前記画像における参照軸と、前記画像における識別された前記人顔対象物の位置との間に形成されたインクルードドアングルと、

関連する識別された前記人顔対象物が前記画像において曲がった角度と、
を示す、ことを特徴とする請求項 2 又は 3 に記載の方法。

10

20

30

40

50

【請求項 5】

前記ステップ（b）で取得された複数の前記画像及び前記サブステップ（c 2）における前記対象物情報は、前記画像における関連する識別された前記人顔対象物の唇に動きがあるかどうかを示す、ことを特徴とする請求項 2 又は 3 に記載の方法。

【請求項 6】

前記ステップ（c）は、

前記画像取り込みモジュールによって取得された前記画像における前記人顔対象物の存在を特定し、それぞれ識別された前記人顔対象物に関連する対象物情報を定め、識別された前記人顔対象物に関連する前記対象物情報に基づき、それぞれの識別された前記人顔対象物の尤度分類を定め、前記尤度分類とは、識別された前記人顔対象物に関連する人がターゲット話者であるとする尤度の分類であり、前記画像において、他の識別された前記人顔対象物よりも高い尤度分類を有する少なくとも 1 つの識別された前記人顔対象物を前記ターゲット対象物として選ぶ、第 1 のモード、及び

10

前記画像取り込みモジュールによって取得された前記画像を外付け電子装置に送り、前記外付け電子装置から、前記画像において前記ターゲット対象物の位置を示す外部制御信号を受け取る第 2 のモードの内の選ばれた一つのモードにおいて前記プロセッサが作動するサブステップを有する、

ことを特徴とする請求項 1 ～ 5 のいずれかの一項に記載の方法。

【請求項 7】

前記ステップ（b）では、前記画像取り込みモジュールは、前記收音モジュールによって集められた前記音声情報にスピーチコンテンツが含まれていることが前記プロセッサにより検出された後で、前記画像を取得することができることを特徴とする請求項 1 ～ 6 のいずれかの一項に記載の方法。

20

【請求項 8】

聴覚障害者の聞き取りを補助する補聴器システムであって、

前記補聴器システムの周囲の音声情報を集める、收音モジュールと、

前記補聴器システムの周囲の画像を取り入れる、画像取り込みモジュールと、

前記收音モジュール及び前記画像取り込みモジュールと連結され、前記收音モジュールによって集められた前記音声情報に対して指向性信号処理操作を施し、出力オーディオ信号を生成する音声処理モジュールを有し、前記出力オーディオ信号は、前記画像取り込みモジュールによって取得された画像におけるターゲット対象物の位置に関連するエリアからの抽出されたボイス信号を含む、プロセッサと、

30

前記プロセッサと連結され、前記出力オーディオ信号を出力する、音声出力装置とを備えていることを特徴とする補聴器システム。

【請求項 9】

前記プロセッサは更に、

前記画像取り込みモジュールによって取得された画像における人顔対象物の存在を識別するように構成され、識別された前記人顔対象物に関連する対象物情報を定めると共に、前記対象物情報に基づき識別された前記人顔対象物それぞれの尤度分類を定める、画像分析モジュールを有し、

40

前記画像における識別された前記人顔対象物のそれぞれに対して定められる前記尤度分類は、識別された前記人顔対象物に関連する人がターゲット話者であるとする尤度の分類であり、

前記音声処理モジュールは、前記画像における識別された前記人顔対象物における少なくとも一つを前記ターゲット対象物として選択し、前記ターゲット対象物として選択された少なくとも一つの識別された前記人顔対象物は、前記画像における前記ターゲット対象物として選択された少なくとも一つ以外の他の識別された前記人顔対象物よりも高い前記尤度を有するものであり、

前記画像分析モジュールは、

前記画像における識別された前記人顔対象物の前記補聴器システムに対しての相対深度

50

を定める、深度分析ユニット、

前記画像取り込みモジュールの軸を表す参照軸と、前記画像における識別された前記画像における前記人顔対象物の位置との間に形成されたインクルードドアングルを定める、角度分析ユニット、及び、

識別された前記人顔対象物が前記画像において曲がった角度を定める、ターン分析ユニット、の少なくとも一つを有することを特徴とする請求項 8 に記載の補聴器システム。

【請求項 10】

前記画像分析モジュールによって定められた前記対象物情報は、関連する識別された前記人顔対象物の前記補聴器システムに対する相対深度情報及び相対方向情報を有し、

前記相対深度情報の相対深度は、異なる相対深度に異なる重みが予め定められ、

前記相対方向情報の相対方向は、異なる相対方向に異なる重みが予め定められ、

前記画像分析モジュールは、前記対象物情報それぞれの前記相対深度情報及び前記相対方向情報に関連する重みに応じて、識別された前記人顔対象物の尤度分類を定める尤度分類器を有する、

ことを特徴とする請求項 8 又は 9 に記載の補聴器システム。

【請求項 11】

前記相対方向情報は、

前記画像取り込みモジュールの一つの軸を表す前記画像における参照軸と前記画像における識別された前記人顔対象物の位置との間に形成されたインクルードドアングルと、

前記画像において前記関連する識別された前記人顔対象物の曲がった角度と、
を示すことを特徴とする請求項 10 に記載の補聴器システム。

【請求項 12】

前記画像取り込みモジュールは、複数の画像を取り入れるように構成され、

前記画像分析モジュールは、唇動き検出ユニットを有し、

前記画像分析モジュールによって定められた前記対象物情報は、前記唇動き検出ユニットによって検出される、対応する識別された前記人顔対象物の唇の動きがあるかどうかを示す情報を含むものである、ことを特徴とする請求項 8 又は 9 に記載の補聴器システム。

【請求項 13】

前記プロセッサは、

前記画像取り込みモジュールによって取得された前記画像における前記人顔対象物の存在を特定し、それぞれ識別された前記人顔対象物に関連する対象物情報を定め、識別された前記人顔対象物に関連する前記対象物情報に基づき、それぞれの識別された前記人顔対象物の尤度分類を定め、前記尤度分類は、識別された前記人顔対象物に関連する人がターゲット話者であるとする尤度によって分類を定めるものであり、前記画像において、他の識別された前記人顔対象物よりも高い尤度分類を有する少なくとも 1 つの識別された前記人顔対象物を前記ターゲット対象物として選ぶ、第 1 のモード、及び

前記画像取り込みモジュールによって取得された前記画像を外付け電子装置に送り、前記外付け電子装置から、前記画像において前記ターゲット対象物の位置を示す外部制御信号を受け取る第 2 のモードの内の選ばれた一つのモードにおいて操作可能に構成されている

ことを特徴とする請求項 8 ～ 12 のいずれかの一項に記載の補聴器システム。

【請求項 14】

前記音声処理モジュールは、前記收音モジュールによって集められた音声情報に含まれるスピーチコンテンツを検出するスピーチ検出ユニットを有し、

前記画像取り込みモジュールは、前記スピーチ検出ユニットにより前記音声情報にスピーチコンテンツが検出された後で、前記画像を取得することができることを特徴とする請求項 8 ～ 13 のいずれかの一項に記載の補聴器システム。

【請求項 15】

前記画像取り込みモジュールと前記收音モジュールと前記プロセッサが取り付けられた装着具を更に有し、

前記装着具は、前記画像取り込みモジュールが取り付けられたフロントサポーターと、前記フロントサポーターの両側にそれぞれ連結された2つのサイドサポーターとからなり、

前記収音モジュールは、複数のマイクロホンを含むマイクロホンアレイであり、

前記マイクロホンのそれぞれが前記フロントサポーター及び2つの前記サイドサポーターのいずれかに設けられ、

前記プロセッサは、2つの前記サイドサポーターの一方に設けられ、

前記音声出力装置は、2つの前記サイドサポーターのそれぞれに連結されて設けられた一対のイヤホンを有し、

前記装着具は、前記フロントサポーターとするレンズ枠を有する眼鏡フレームと、前記サイドサポーターとする一対のテンプルとを備えていることを特徴とする請求項8～14のいずれかの一項に記載の補聴器システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、オーディオ信号を処理する方法及び補聴器システムに関し、特に、画像情報及び音声情報を用いて集音するオーディオ信号を処理する方法及び補聴器システムに関する。

【背景技術】

【0002】

カクテルパーティー効果 (cocktail party effect) とは、カクテルパーティーのようにたくさんの人がそれぞれに雑談している中でも、自分が興味のある人の発話を聞き取ることができるように、音声を選択的に聞き取ることができることを指す。これは、一般に人の聴覚系が音源の位置を限定し、該音源の位置から音声情報を抽出することができるようになっていていることによる。このように、騒々しい環境でも、特定の相手に集中することによって簡単にターゲット話者と会話することが可能である。

【0003】

しかし、聴覚障害を有する人によっては、一般の人よりも可聴値 (threshold of hearing) が高く、音源の位置を特定することができないため、カクテルパーティー効果が得られない。また、片方の耳だけが不自由な場合でもうまくカクテルパーティー効果を得ることができない。

【先行技術文献】

【非特許文献】

【0004】

【非特許文献1】書名：Digital Hearing Aids、出版社：Thieme Medical Publishers, INC.、著者：Arthur Schaub、出版日：2008年6月16日、ISBN：978-1-60406-006-5

【発明の概要】

【発明が解決しようとする課題】

【0005】

従来の補聴器システムには、音源の位置を特定することにより、聴覚障害者の聞き取りを補助するものもある。このような補聴器システムは、ユーザーの前方からの音声情報を抽出し、抽出された音声情報を処理し、信号対ノイズ比 (SNR) を上げるように構成されている。聴覚障害のあるユーザーは、補聴器システムによって、騒々しい環境でも、前方に居る話者の声がよく聞き取れるようになるが、ユーザーの前方に複数の人が居る場合、従来の補聴器システムは、ユーザーのターゲット話者がだれであるかを特定することができない。

【0006】

本発明は、複数の話者が居る場合でもユーザーがターゲット話者の声を集中して聞くこ

10

20

30

40

50

とができるように補助することができるオーディオ信号を処理する方法及び補聴器システム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】

上記目的を達成するために、一の観点によれば、

本発明に係るオーディオ信号を処理する方法は、画像取り込みモジュールと收音モジュールとプロセッサとを備えている補聴器システムを用いて実行される方法であって、前記收音モジュールを用いて、前記補聴器システムの周囲の音声情報を集める、ステップ（a）と、前記画像取り込みモジュールを用いて、前記補聴器システムの周囲の画像を取得する、ステップ（b）と、前記プロセッサを用いて、前記收音モジュールによって集められた前記音声情報に指向性信号処理操作を実行し、前記画像取り込みモジュールによって取得された前記画像におけるターゲット対象物の位置に関連するエリアからの抽出されたボイス信号を含む出力オーディオ信号を生成する、ステップ（c）と、前記出力オーディオ信号を出力する、ステップ（d）と、を有することを特徴とする。

10

【0008】

また、他の観点によれば、本発明に係る補聴器システムは、聴覚障害者の聞き取りを補助する補聴器システムであって、前記補聴器システムの周囲の音声情報を集める、收音モジュールと、前記補聴器システムの周囲の画像を取り入れる、画像取り込みモジュールと、前記收音モジュール及び前記画像取り込みモジュールと連結され、前記收音モジュールによって集められた前記音声情報に対して指向性信号処理操作を施し、出力オーディオ信号を生成する音声処理モジュールを有し、前記出力オーディオ信号は、前記画像取り込みモジュールによって取得された画像におけるターゲット対象物の位置に関連するエリアからの抽出されたボイス信号を含む、プロセッサと、前記プロセッサと連結され、前記出力オーディオ信号を出力する、音声出力装置と、を備えていることを特徴とする。

20

【発明の効果】

【0009】

本発明に係るオーディオ信号を処理する方法及び補聴器システムによって、ユーザーは、騒々しい環境、又は、ユーザーの前方に複数の人が居る場合でも、ターゲット話者を特定してそのターゲット話者の声がよく聞き取れるようになる。

【図面の簡単な説明】

30

【0010】

【図1】本発明に係る補聴器システムの一つの実施形態を実施する一例を概略的に示す斜視図である。

【図2】本発明に係る補聴器システムの構成を概略的に示すブロック図である。

【図3】本発明に係るオーディオ信号を処理する方法のフローチャートを示す図である。

【図4】画像取り込みモジュールによって取り入れられた補聴器システムの周囲の画像を示す図である。

【図5】三つの識別された人顔対象物及び人顔対象物それぞれの参照軸に対するインクルードッドアングル $\theta 1$ 、 $\theta 2$ 、 $\theta 3$ を示す図である。

【発明を実施するための形態】

40

【0011】

以下添付図面に従って本発明に係る補聴器システムの一つの好ましい実施の形態について詳説する。

（実施形態）

【0012】

図1は、本発明に係る補聴器システムの一つの実施形態を概略的に示す斜視図である。図示の如く、この実施形態に係る補聴器システム2は、眼鏡等の装着具3によってユーザーに装着可能に構成することができる。

【0013】

該実施形態に係る補聴器システム2は、図2に示されているように、画像取り込みモジ

50

ジュール 4、收音モジュール 5、プロセッサ 6 及び音声出力装置 7 等の電子部品を備えている。補聴器システム 2 に備えている電子部品が装着具 3 に取り付けられている。

【0014】

装着具 3 は、画像取り込みモジュール 4 が取り付けられる横長のフロントサポーター 3 1 と、ユーザーの耳に掛けて支持されるようにフロントサポーター 3 1 の両側にそれぞれ連結されているサイドサポーター 3 2、3 2 とから構成された枠機構を備えている。装着具 3 は、この実施形態では、フロントサポーター 3 1 とするレンズ枠を有する眼鏡フレームと、サイドサポーター 3 2、3 2 とする一对のテンプルとを備えてメガネのように構成され、フロントサポーター 3 1 は一对のレンズ 3 1 0、3 1 0 を保持可能に構成されている。

10

【0015】

画像取り込みモジュール 4 は、図 4 に示されているように、補聴器システム 2 の周りの画像 4 0 を取り込むように構成され、複数の画像 4 0 を取り込み可能に一对のレンズ 3 1 0、3 1 0 の間に取り付けられたカメラ部を有する。

【0016】

收音モジュール 5 は、1 例としては、複数のマイクロホン 5 1 を含むマイクロホンアレイを有する。複数のマイクロホン 5 1 は、補聴器システム 2 の周囲の情報を集めるようにフロントサポーター 3 1 とサイドサポーター 3 2、3 2 に取り付けられている。なお、マイクロホン 5 1 は、さまざまな指向性、例えば全方向性、単一指向性又はそれらの混合型の指向性を備えたものが用いられることが可能であることに留意されたい。

20

【0017】

プロセッサ 6 は、画像取り込みモジュール 4 及び收音モジュール 5 と連結されるようにサイドサポーター 3 2、3 2 のいずれかに取り付けられ、図 2 に示されているように、画像分析モジュール 6 1 と音声処理モジュール 6 2 とを有する。

【0018】

画像分析モジュール 6 1 は、画像取り込みモジュール 4 によって取り込まれた画像 4 0 における人の顔の存在（図 4 参照）を人顔対象物 4 0 1 として識別するように構成されている。画像分析モジュール 6 1 はまた、人顔対象物 4 0 1 のそれぞれに関連する対象物情報を定めると共に、人顔対象物 4 0 1 の対象物情報に基づいて人顔対象物 4 0 1 のそれぞれを識別するように尤度分類を定める。なお、ここでいう尤度とは、人顔対象物 4 0 1 のそれぞれに応じて、識別された人顔対象物 4 0 1 に関連する人がターゲット話者である可能性を示す尤度であり、以下に詳説するように、尤度の高低によって分類する。

30

【0019】

画像分析モジュール 6 1 によって定められた対象物情報は、この実施形態では、識別された人顔対象物 4 0 1 に関連する、補聴器システム 2 に対する相対深度情報と相対方向情報とを含む。この実施形態では、相対方向情報とは、画像 4 0 における参照軸 4 0 2 と画像 4 0 における識別された人顔対象物 4 0 1 との位置の間の角度（インクルードドアングル）を示す。参照軸 4 0 2 とは、図 1 に示されているように一对のレンズ 3 1 0 間のフロントサポーター 3 1 に配置された画像取り込みモジュール 4 の軸を意味する。相対方向情報はまた、画像 4 0 において識別された人顔対象物 4 0 1 のターン（顔の向き）による角度をも示す。

40

【0020】

画像分析モジュール 6 1 は、上述の対象物情報を得るために、図 2 に示されているように、深度分析ユニット 6 1 1 と、角度分析ユニット 6 1 2 と、ターン分析ユニット 6 1 3 と、尤度分類器 6 1 4 と、唇動き検出ユニット 6 1 5 とを有する。

【0021】

深度分析ユニット 6 1 1 は、画像 4 0 における識別された人顔対象物 4 0 1 それぞれの補聴器システム 2 に対する相対深度を定めるように構成されている。より詳しく説明すると、深度分析ユニット 6 1 1 は、識別された人顔対象物 4 0 1 のサイズと識別された人顔対象物 4 0 1 の相対深度の関係に関連する情報を含むデータベースを有する。1 例として

50

は、深度分析ユニット611は、例えば、画像40におけるそれぞれの識別された人顔対象物401の領域を算出することによって、それぞれの識別された人顔対象物401のサイズを得るように構成される。その後、深度分析ユニット611は、該データベースを参照することで、画像40におけるそれぞれの識別された人顔対象物401の相対深度を定めるように構成される。なお、この実施形態では、深度分析ユニット611を用いて相対深度を定めるが、他の実施形態としては深度分析ユニット611以外の構成を用いてもよいことに留意されたい。

【0022】

角度分析ユニット612は、図5に示されているように、画像取り込みモジュール4の一つの軸を表す画像40における参照軸402と画像40における識別された人顔対象物401の位置との間に形成されたインクルードドアングルを定めるように構成されている。

【0023】

画像取り込みモジュール4は、この実施形態において一对のレンズ310の間に取り付けられているので、参照軸402は、装着具3を掛けているユーザーの視野内の線を示す。

【0024】

図5は、三つの識別された人顔対象物401及び参照軸402に対する各人顔対象物401それぞれのインクルードドアングル θ_1 、 θ_2 、 θ_3 を示している。画像分析モジュール61は、識別された人顔対象物401それぞれの相対深度及びインクルードドアングルを使って、識別された人顔対象物401の補聴器システム2に対する相対位置を得ることができる。なお、この実施形態では、画像分析モジュール61を用いて人顔対象物401の相対位置を定めるが、他の実施形態としては画像分析モジュール61以外の構成を用いてもよいことに留意されたい。

【0025】

ターン分析ユニット613は、画像40において識別された人顔対象物401の曲がった角度、つまり、装着具3を掛けたユーザーから見た場合に人顔対象物401が正面を向いている場合を基準として人顔対象物401の顔の向きがその基準から曲がった（ターンした）角度を定めるように構成されている。

【0026】

尤度分類器614は、識別された人顔対象物401に関連する対象物情報に基づいて、識別された人顔対象物401それぞれの尤度分類を定めるように構成されている。なお、画像分析モジュール61によって定められた識別された人顔対象物401に関連する対象物情報は、相対深度情報と相対方向情報とを含む。

【0027】

この実施形態では、異なる相対深度には異なる重みが予め定められ、異なる相対方向には異なる重みが予め定められ、識別された人顔401それぞれの尤度分類は、それぞれの相対深度情報及び相対方向情報に関連する重みに応じて、尤度分類器614によって予め定められる。

【0028】

一般的には、人との距離に関して、パーソナルスペースという近接学研究の用語があり、人と人の間の距離は、夫婦、恋人、親子、親友の間等0.5m以内の密接距離、相手の表情が読み取れ且つ相手に手が届く0.5m～1.5m以内の個体距離、例えば会議等の相手に手は届きづらいが容易に会話ができる1.5～3m以内の社会距離、例えば講演等の複数の相手が見渡せる3m以上の公共距離等の四つのゾーンに大別される。これによれば、親密な相手又はより重要な対話者ほどユーザーに近くなり、且つユーザーに近いほど、ユーザーに向かってほぼ対面すると共に、ユーザーの視野中心に近づいてくる。

【0029】

これによって、この実施形態では、尤度分類器614は、相対深度が小さく、インクルードドアングルも小さく、画像40において識別された人顔対象物401の顔の向きの

10

20

30

40

50

傾いた角度も小さいことに高い重みを対応付ける。次いで、尤度分類器 6 1 4 は、識別された人顔対象物 4 0 1 それぞれに対応付けられた重みを基にして尤度分類を定める。この実施形態では、識別された人顔対象物 4 0 1 それぞれが、高尤度、中尤度及び低尤度のいずれかに分類される。

【0030】

唇動き検出ユニット 6 1 5 は、識別された人顔対象物 4 0 1 に唇の動きがあるかどうかを検出するものであり、画像分析モジュール 6 1 によって定められた対象物情報は、唇動き検出ユニット 6 1 5 によって検出された、対応する識別された人顔対象物 4 0 1 に唇の動きがあるかどうかを示す情報を含むものである。例えば本実施形態では、唇動き検出ユニット 6 1 5 は、高尤度に分類された識別された人顔対象物 4 0 1 において唇の動きがあるかどうかを検出するように構成されている。唇動き検出ユニット 6 1 5 によって、一つの識別された人顔対象物 4 0 1 の唇の動きが検出されると、該人顔対象物 4 0 1 を発話者であるとみなし、ターゲット話者とするのが考えられる。

10

【0031】

唇動き検出ユニット 6 1 5 はまた、高尤度に分類された識別された人顔対象物 4 0 1 がなかった場合、中尤度に分類された識別された人顔対象物 4 0 1 に唇の動きが存在するかどうかを検出するように構成されている。高尤度に分類された識別された人顔対象物 4 0 1 も中尤度に分類された識別された人顔対象物 4 0 1 もなかった場合、唇動き検出ユニット 6 1 5 は判断を中止する。この場合、画像 4 0 に取得されて識別された人顔対象物 4 0 1 においてターゲット話者が存在する可能性が低いと考えられる。

20

【0032】

この実施形態では、唇動き検出ユニット 6 1 5 を用いて識別された人顔対象物 4 0 1 を話者であるか否かと検出するが、場合によって唇動き検出ユニット 6 1 5 を省略してもよい。この場合、プロセッサ 6 は、一つの識別された人顔対象物 4 0 1 をターゲット対象として選ぶ。この際に選ばれる識別された人顔対象物 4 0 1 は、画像 4 0 において他の一つの識別された人顔対象物 4 0 1 と比較してより高い尤度を有するものである。画像分析モジュール 6 1 は、ターゲット対象と思われる識別された人顔対象物 4 0 1 を示す信号を音声処理モジュール 6 2 に送る。

【0033】

音声処理モジュール 6 2 は、この実施形態では、收音モジュール 5 によって集められた音声情報を受信・処理するように構成されている。音声処理モジュール 6 2 は、受信された音声情報を処理する際、例えば、アナログ・デジタル変換処理や、ノイズ低減処理等を行う。この実施形態では、音声処理モジュール 6 2 は、スピーチ検出ユニット 6 2 1 とサウンドプロセッサ 6 2 2 とを有する。

30

【0034】

スピーチ検出ユニット 6 2 1 は、收音モジュール 5 によって集められた音声情報に含まれ得るスピーチコンテンツを検出するように構成されている。音声情報におけるスピーチコンテンツの検出とは、補聴器システム 2 の周囲環境における人声の存在を検出することを示す。この実施形態では、画像取り込みモジュール 4 は、スピーチ検出ユニット 6 2 1 によるスピーチコンテンツの検出を示す信号が送られ後で、画像 4 0 を取得することができるようになっている。

40

【0035】

サウンドプロセッサ 6 2 2 は、音声情報に対して指向性信号処理操作を施し、出力オーディオ信号を生成するように構成されている。当該指向性信号処理操作によって、画像 4 0 におけるターゲット対象の位置に関連するエリアから来た音声を残すと共に、ターゲット対象の位置に関連するエリアから来た音声以外の方向から来た音声をろ過して除去する。この実施形態では、当該指向性信号処理操作とは、ビーム形成操作をいう。

【0036】

出力オーディオ信号は、当該指向性信号処理操作の後、画像取り込みモジュール 4 によって取得された画像 4 0 におけるターゲット対象の位置に関連するエリアから来るもので

50

ある抽出されたボイス信号を含む。サウンドプロセッサ622はまた、該抽出されたボイス信号に対してノイズの低減及び／又は増幅を行った上で音声出力装置7に送るように構成されている。この実施形態では、音声出力装置7は、一対のイヤホンの形をして、それぞれがサイドサポーター32それぞれに連結され、プロセッサ6に結合され、プロセッサ6から出力されたボイス信号を受信して外部に音声として出力するように構成されている。こうして、ユーザーは本発明に係る補聴器システム2を掛けると、ターゲット話者からの音声を選択的に聞くことができる。

【0037】

この実施形態では、プロセッサ6は更に、外付け電子装置8と通信できる通信ユニット9を有する（図2参照）。通信ユニット9は、Wi-Fi、Zigbee（登録商標）、Bluetooth（登録商標）、近距離無線通信（near field communication／NFC）のいずれかを介して、外付け電子装置8と通信するように構成される。

【0038】

外付け電子装置8は、モバイル端末、タブレットコンピュータ、ノートブックコンピュータ又は他の携帯機器を用いて実行され、コントロールユニット（図示せず）、タッチディスプレイ81及び通信ユニット82とを有する。通信ユニット82は、外付け電子装置8にインストールされたアプリケーションプログラムによって実行される。

【0039】

対象とする補聴器システム2が外付け電子装置8と通信する場合、ユーザーはタッチディスプレイ81を操作し、通信ユニット82を制御するコマンドを入力する。これによって、補聴器システム2及び外付け電子装置8の間に通信チャネルを設置させようとする。通信チャネルが設置されると、プロセッサ6は、通信ユニット9を介して、画像取り込みモジュール4によって取得された画像40を外付け電子装置8に送る。

【0040】

外付け電子装置8のコントロールユニットは、タッチディスプレイ81を制御して、人顔対象物401を含む画像40を映すと共に、ユーザーに人顔対象物401の一つをターゲット対象として選ばせる。その後、外付け電子装置8は、画像40におけるターゲット対象の位置を示す外部制御信号を生成し、通信ユニット82を介して補聴器システム2に送る。プロセッサ6が外部制御信号を受け取ると、画像分析モジュール61は、識別された人顔対象物401からのターゲット対象選択に関連する操作をバイパスし、音声処理モジュール62は外部制御信号に基づいて操作し始める。

【0041】

なお、装着具3は、更に装着具3に設けられた電子部品に電力を供給するための手段例えば電池を有することは、当業者ならすぐ理解できるだろう。また、装着具3は、ユーザーに補聴器システム2をオンオフすることを可能にする物理スイッチを設けてもよい。

【0042】

図1と図2に示されたように構成された本発明に係る補聴器システム2は、図3に示されているように、本発明に係るオーディオ信号処理方法の一例を実行する。

【0043】

ステップ100では、例えばユーザーがメガネの形等をした装着具3を掛けた後で、補聴器システム2が作動される。收音モジュール5は、起動後、その周りの音声情報を集め始める。

【0044】

ステップ102では、スピーチ検出ユニット621は、收音モジュール5によって集められた音声情報にスピーチコンテンツが含まれているか否かを検出する。スピーチコンテンツが検出された場合、ステップ104に進む。スピーチコンテンツが検出されなかった場合、再度收音モジュール5によって集められた音声情報に対する検出を行なう。

【0045】

ステップ104では、画像取り込みモジュール4が画像40を取得するようにする。

【0046】

ステップ106では、プロセッサ6は、補聴器システム2及び外付け電子装置8の間に通信チャンネルが設置されたかどうかを定める。通信チャンネルが設置されていない場合、補聴器システム2は第1のモードで作動し、ステップ110に進む。通信チャンネルが設置されている場合、補聴器システム2は第2のモードで作動し、ステップ130に進む。

【0047】

第1のモードにて作動すると、ステップ110では、画像分析モジュール61は、画像取り込みモジュール4によって取得された画像40における人顔対象物401を識別する。

【0048】

ステップ112では、画像分析モジュール61は、識別された人顔対象物401にそれぞれ対応する対象物情報を定める。深度分析ユニット611は、画像40において識別された人顔対象物401それぞれの補聴器システム2に対しての相対深度を定める。角度分析ユニット612は、画像40における参照軸402と画像40における識別された人顔対象物401の位置との間におけるインクルードドアングルを定める（図5参照）。ターン分析ユニット613は、画像40において識別された人顔対象物401それぞれの曲がった（ターンした）角度を定める。

【0049】

ステップ114では、尤度分類器614は、対象物情報に基づき、識別された対象物情報それぞれの尤度分類を定める。

【0050】

ステップ116では、唇動き検出ユニット615は、高尤度に分類された（又は高尤度に分類された識別された人顔対象物401がない場合は中尤度に分類された）識別された人顔対象物401のそれぞれの唇の動きの有無を検出する。唇の動きがあると検出された識別された人顔対象物401がターゲット対象物として選択される。

【0051】

ステップ118では、サウンドプロセッサ622は音声情報に対してビーム形成操作を施すと、画像40におけるターゲット対象物に関連するエリアから抽出されたボイス信号を含む出力オーディオ信号を生成する。

【0052】

ステップ120では、当該出力オーディオ信号を音声出力装置7に送信し、出力する。

【0053】

一方、第2のモードにて（例えば、補聴器システム2が外付け電子装置8と通信したとき）操作すると、ステップ130では、プロセッサ6は、画像40を通信ユニット9を介して外付け電子装置8に送る。外付け電子装置8が画像40を受け取ると、画像40をタッチディスプレイ81に映し、補聴器システム2がユーザーによる手動で一つ又は一つ以上の識別された人顔対象物401がターゲット対象物として選択される。

【0054】

ステップ132では、プロセッサ6は、外付け電子装置8からユーザーによって選ばれた画像40におけるターゲット対象物の位置を示す外部制御信号を受け取る。その後、ステップ118に進む。

【0055】

なお、唇動き検出ユニット615を用いると、ターゲット対象物の選択の正確さを大いに高めることができるが、例えばいくつかの所定の対象物からの音声が入って自動的にろ過されて除去されないように、唇動き検出ユニット615を起動しなくてもよいことに留意されたい。

【0056】

唇動き検出ユニット615を起動しない場合、高尤度を有する識別された人顔対象物401がターゲット対象物と見なされる。高尤度を有する識別された人顔対象物401が現れない場合、中尤度を有する識別された人顔対象物401がターゲット対象物と見なされ

10

20

30

40

50

る。高尤度を有する識別された人顔対象物 4 0 1 も中尤度を有する識別された人顔対象物 4 0 1 も現れない場合、サウンドプロセッサ 6 2 2 はただ正面方向の音声情報を抽出するようにすることができる。

【0057】

また、場合によっては、補聴器システム 2 にターン分析ユニット 6 1 3 を設けなくてもよく、同様に、本発明に係る方法でもターン分析ユニット 6 1 3 に関連する操作を省略してもよい。

【0058】

なお、本実施形態では、画像取り込みモジュール 4 が装着具 3 の一対のレンズ 3 1 0、3 1 0 の間に取り付けられているが、場合によって装着具 3 の別のところに設けられてもよい。

10

【0059】

以上により、本発明に係る補聴器システム 2 は、取得された画像 4 0 からターゲットとする話者を選ぶ選択肢を 2 つ提供することができる。第 1 のモードでは、画像分析モジュール 6 1 は対象物情報に基づき自動的に選択操作を実行する。第 2 のモードでは、ユーザーによってターゲットとする話者が選別される。これによって、本発明に係る補聴器システム 2 を用いると、ユーザーは、ノイズのある環境において複数の話者からターゲットとする一人か数人の話者の居る位置からの音声を選択的に聞くことができる。

【符号の説明】

【0060】

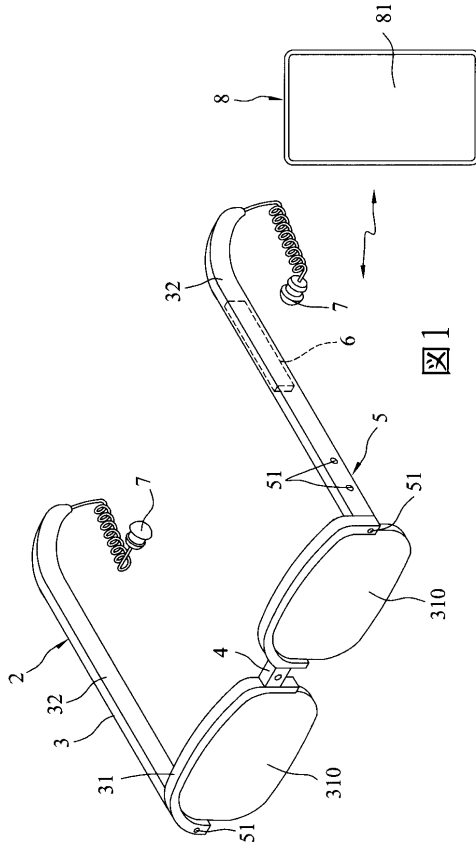
20

- 2 補聴器システム
- 3 装着具
- 3 1 フロントサポーター
- 3 1 0 レンズ
- 3 2 サイドサポーター
- 4 画像取り込みモジュール
- 4 0 画像
- 4 0 1 人顔対象物
- 4 0 2 参照軸
- 5 収音モジュール
- 5 1 マイクロホン
- 6 プロセッサ
- 6 1 画像分析モジュール
- 6 1 1 深度分析ユニット
- 6 1 2 角度分析ユニット
- 6 1 3 ターン分析ユニット
- 6 1 4 尤度分類器
- 6 1 5 唇動き検出ユニット
- 6 2 音声処理モジュール
- 6 2 1 スピーチ検出ユニット
- 6 2 2 サウンドプロセッサ
- 7 音声出力装置
- 8 外付け電子装置
- 8 1 タッチディスプレイ
- 8 2 通信ユニット
- 9 通信ユニット
- θ 1、θ 2、θ 3 インクルードッドアングル

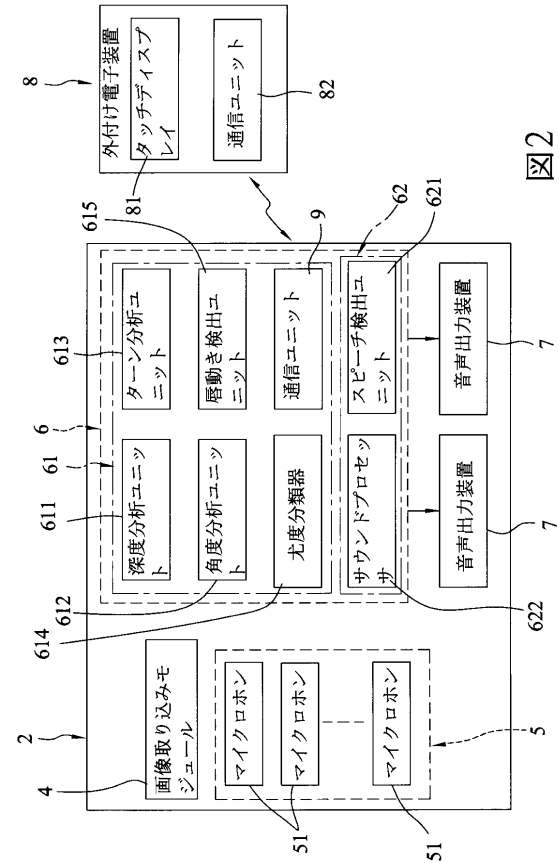
30

40

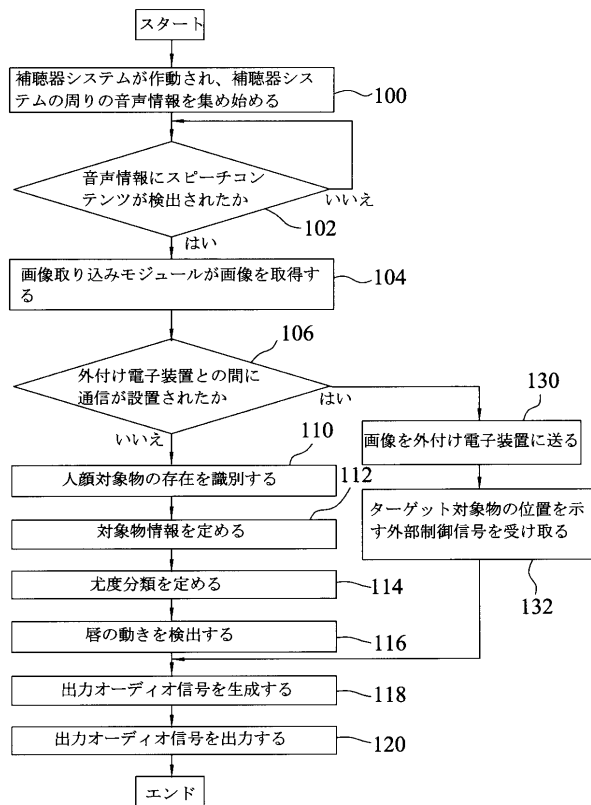
【図 1】



【図 2】



【図 3】



【図 4】

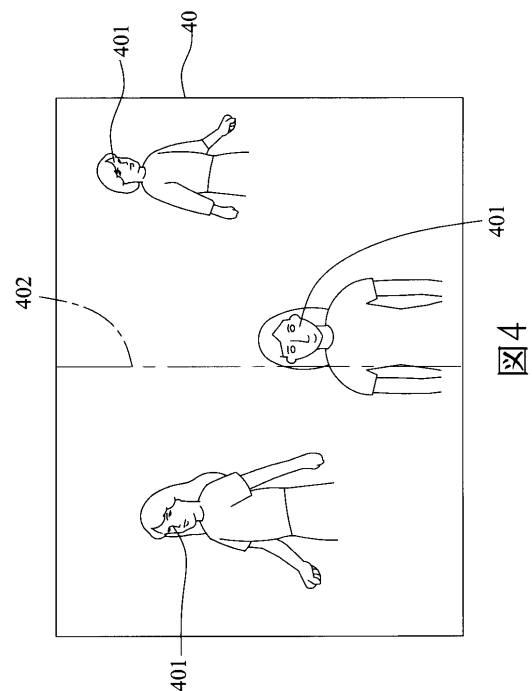
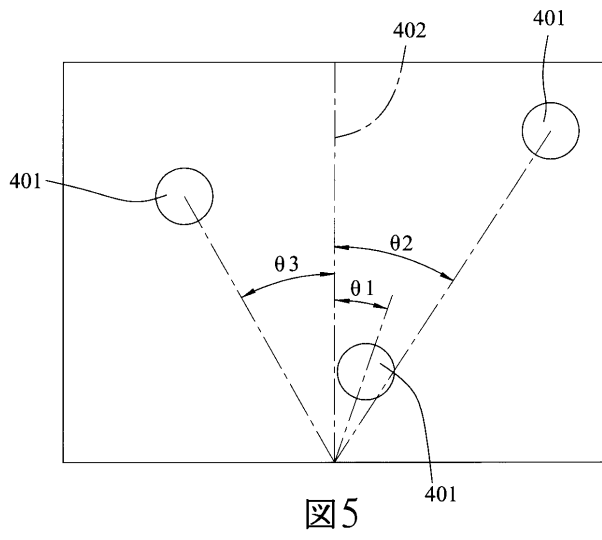


図3

【図 5】



フロントページの続き

(71)出願人 514135155

チェン シアオ ハン

Hsiao-Han CHEN

台湾台南市仁徳區忠義三街24巷21號

No. 21, Ln. 24, Jhongyi 3rd St., Rende Dist.,
Tainan City, Taiwan Postal code 71753

(74)代理人 100099759

弁理士 青木 篤

(74)代理人 100092624

弁理士 鶴田 準一

(74)代理人 100114018

弁理士 南山 知広

(74)代理人 100165191

弁理士 河合 章

(74)代理人 100151459

弁理士 中村 健一

(74)代理人 100196601

弁理士 酒井 祐市

(72)発明者 リウ チン フォン

台湾, カオションシー クーシャンチュイ メイシューナンウーチェ 99シアン 10ハオ

(72)発明者 チェン シアオ ハン

台湾, タイナンシー レントーチュイ チョンイーサンチェ 24シアン 21ハオ

Fターム(参考) 5C122 DA03 EA59 FH10 FH12 FH14 FJ01 FJ09 HB01

5D220 BA01 BC05