

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3670180号  
(P3670180)

(45) 発行日 平成17年7月13日(2005.7.13)

(24) 登録日 平成17年4月22日(2005.4.22)

(51) Int.Cl.<sup>7</sup>

F I

G 1 O L 15/00

G 1 O L 3/00 5 5 1 C

G 1 O L 15/22

H O 4 R 25/00 M

H O 4 R 25/00

G 1 O L 3/00 5 5 1 E

G 1 O L 3/00 5 6 1 D

請求項の数 10 (全 27 頁)

(21) 出願番号	特願平11-338458	(73) 特許権者	399011357
(22) 出願日	平成11年11月29日(1999.11.29)		有限会社ジーエムアンドエム
(65) 公開番号	特開2000-308198(P2000-308198A)		東京都品川区東五反田五丁目10番5-1
(43) 公開日	平成12年11月2日(2000.11.2)		04号
審査請求日	平成11年11月29日(1999.11.29)	(74) 代理人	100089705
(31) 優先権主張番号	特願平11-37559		弁理士 社本 一夫
(32) 優先日	平成11年2月16日(1999.2.16)	(74) 代理人	100076691
(33) 優先権主張国	日本国(JP)		弁理士 増井 忠式
		(74) 代理人	100075270
			弁理士 小林 泰
		(74) 代理人	100080137
			弁理士 千葉 昭男
		(74) 代理人	100096013
			弁理士 富田 博行

最終頁に続く

(54) 【発明の名称】 補聴器

(57) 【特許請求の範囲】

【請求項1】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び/又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び/又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記出力制御手段は、上記認識手段からの認識結果に基づいて、上記認識手段による認識結果及び/又は認識結果を上記変換手段により加工変換した認識結果を、再度上記出力手段から出力するように制御することを特徴とする補聴器。

【請求項2】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び/又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に

10

20

基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記出力手段は人工内耳機構からなり、上記出力制御手段は、上記認識結果及び／又は上記変換された認識結果を電気信号として出力するように上記制御信号を生成することを特徴とする補聴器。

【請求項 3】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記出力手段は圧挺板からなり、上記出力制御手段は、上記圧挺板に上記認識結果及び／又は上記変換された認識結果を振動として出力するように上記制御信号を生成することを特徴とする補聴器。

10

【請求項 4】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記出力手段は人工中耳機構からなり、上記出力制御手段は、上記認識結果及び／又は上記変換された認識結果を電気信号として出力するように上記制御信号を生成することを特徴とする補聴器。

20

【請求項 5】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記出力手段は、超音波出力機構からなり、上記出力制御手段は、上記超音波出力機構に上記認識結果及び／又は上記変換された認識結果を電気信号として出力するように上記制御信号を生成することを特徴とする補聴器。

30

40

【請求項 6】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記出力手

50

段は、タクトイルエイド用振動子アレイからなり、上記出力制御手段は、上記振動子アレイに上記認識結果及び／又は上記変換された認識結果を電気信号として出力するように上記制御信号を生成することを特徴とする補聴器。

【請求項 7】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記出力手段は、聴性脳幹インプラント機構からなり、上記出力制御手段は、上記聴性脳幹インプラント機構に上記認識結果及び／又は上記変換された認識結果を電気信号として出力するように上記制御信号を生成することを特徴とする補聴器。

10

【請求項 8】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記音響電気変換手段、認識手段、変換手段、出力制御手段及び出力手段を複数の装置にて実現し、該複数の装置間を無線通信網にて接続し、少なくとも音声情報の送受信を行うことを特徴とする補聴器。

20

【請求項 9】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段とを備え、上記音響電気変換手段は外部から検出された音声から生成した電気信号を用いて上記音声信号を生成するものである補聴器。

30

【請求項 10】

検出した音声を聴力障害者用に要約して提示する補聴器であって、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、上記認識手段からの認識結果の意味内容が使用者にとって分かり易くなるよう、上記認識結果である音声情報を要約する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を出力して使用者に音声を提示するための出力手段と、表示部と、上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果である音声情報を上記表示部へ出力する手段とを備える補聴器。

40

【発明の詳細な説明】

50

## 【 0 0 0 1 】

## 【 発明の属する技術分野 】

本発明は、マイクロホン等により検出した音声を聴力障害者が理解しやすい形式に加工変換して提示する補聴器、音声言語障害を持つ者より発せられた音声や音声言語障害を是正するために用いる補助的装置や手段（例として喉頭摘出後の代用発声法（speech production substitutes））により発せられた音声を加工変換して出力する補聴器に関する。

## 【 0 0 0 2 】

## 【 従来技術 】

補聴器には、気導方式と、骨導方式とが従来から使用されている。補聴器の種類としては、箱形補聴器、耳かけ補聴器、CROS（Contra-lateral Routing of Signal）補聴器、耳穴形補聴器がある。また、従来の処理方式として分けると、アナログ補聴器とデジタル補聴器とがある。また、補聴器には、小寺の報告によると集団で使用する大型のもの（卓上訓練用補聴器、集団訓練用補聴器）、個人的に使用する小型のものがある（小寺一興、補聴器の選択と評価 図説耳鼻咽喉科new approach メジカルビュー - ,39,1996参照）。

10

## 【 0 0 0 3 】

このデジタル補聴器は、マイクロホンで検出した音声を先ずA/D（analog/digital）変換処理することでデジタルデータを生成する。そして、このデジタル補聴器は、例えばフーリエ変換処理を施すことにより入力されたデジタルデータを周波数スペクトルに分解することで解析を行い、各周波数帯域毎に音声の感覚的な大きさに基づいた増幅度の算出を行う。そして、このデジタル補聴器は、各周波数帯域毎に増幅されたデジタルデータをデジタルフィルターに通過させてD/A変換処理を行って再び音声を使用者の耳に出力するように構成されている。これにより、デジタル補聴器は、話し手の音声を雑音の少ない状態で使用者に聞かせていた。

20

## 【 0 0 0 4 】

また、従来において、例えば喉頭摘出により音声障害をもつ人は、通常の声帯振動による発声機構を失い、音声生成が困難になる。

## 【 0 0 0 5 】

現在まで、喉頭摘出後の代用発声法として、音源としての振動体の性質から大別するとゴム膜（笛式人工喉頭）やブザー（電気人工喉頭（経皮型、埋込み型））等の人工材料を用いる方法と、下咽頭や食道粘膜を使用する方法（食道発声、気管食道瘻発声、ボイスプロステシス（voice prostheses）使用の気管食道瘻発声）がある。また、その他の代用発声法としては、口唇を動かしたときに生じる筋電図を利用したものや聴力障害による発声障害者のために種々の音声処理技術を利用した発声発話訓練装置、パラトグラフ（palatograph）によるものや口腔内の振動子によるものが報告されている。

30

## 【 0 0 0 6 】

## 【 発明が解決しようとする課題 】

しかし、上述したデジタル補聴器では、各周波数帯域毎にデジタルデータを増幅させる処理を行っているだけなので、マイクロホンにより周囲の音が無作為に收音し、雑音をそのまま再生してしまい使用者の不快感が残り、アナログ補聴器と比べても、種々の聴力検査において大幅な改善はなかった。また、従来のデジタル補聴器では、難聴者の身体状態、利用状態及び使用目的に応じて検出した音声に対する処理を適応させることはなされていなかった。

40

## 【 0 0 0 7 】

そこで、本発明の目的は、使用者の身体状態、利用状態及び使用目的に応じて音声認識の結果を提示するとともに、ノイズが少ない状態で認識結果を提示することができる補聴器を提供することにある。

## 【 0 0 0 8 】

また、上記代用発声法に共通してみられるのは、喉頭摘出前の本人自身の本来の正常な状態での声帯振動によるものではないので、生成する音声の音質が悪く、本来正常であった本人が発していた声とはかけ離れているという問題点が挙げられる。

50

## 【 0 0 0 9 】

そこで、本発明は、上述したような実情に鑑みて提案されたものであり、喉頭摘出や舌口腔底切除や構音障害等による音声言語障害を有する人達が本来自身がもつ、或いは自在に変換させて自然な音声で発声することを可能とするとともに、外部からの音声を使用者に出力して自然な会話を行わせることができる補聴器を提供することを目的とする。

## 【 0 0 1 0 】

## 【課題を解決するための手段】

上述の課題を解決する本発明に係る補聴器は、音声言語障害を有した使用者から発せられた音声及び／又は外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号に基づいて音声認識をする処理を行う音声認識手段と、音声データを記憶する記憶手段と、上記音声認識手段からの認識結果に基づいて上記記憶手段に記憶された音声データを組み合わせ、出力する音声を示す音声情報を生成する音声情報生成手段と、上記音声情報生成手段で生成された音声情報を音声に変換して外部に出力する使用者音声出力手段と、上記音声認識手段で認識された認識結果を上記外部からの音声として使用者に出力する外部音声出力手段とを備えることを特徴とするものである。

10

## 【 0 0 1 1 】

このような補聴器は、外部からの音声を使用者に出力するとともに、障害を有して発せられた音声を発声した使用者に出力する。

## 【 0 0 1 2 】

本発明に係る補聴器は、外部からの音声を検出して音声信号を生成する音響電気変換手段と、上記音響電気変換手段からの音声信号を用いて音声認識処理を行う認識手段と、使用者の身体状態、利用状態及び使用目的に応じて、上記認識手段からの認識結果の内容を変更するように変換する変換手段と、上記認識手段による認識結果及び／又は認識結果を上記変換手段により変換した認識結果を出力させる制御信号を生成する出力制御手段と、上記出力制御手段で生成された制御信号に基づいて上記認識手段による認識結果及び／又は上記変換手段により変換された認識結果、である音声情報を出力して使用者に音声を提示するための認識結果を出力して認識結果である音声情報を使用者に提示する出力手段とを備えることを特徴とするものである。

20

## 【 0 0 1 3 】

このような補聴器は、変換手段で認識結果の内容を変更することで出力結果を変更して使用者に変換手段で変更された音声等を提示する。このような補聴器によれば、使用者の身体状態、利用状態及び使用目的に応じて自在に変換方式を変更して認識結果を提示する。

30

## 【 0 0 1 4 】

## 【発明の実施の形態】

以下、本発明の実施の形態について図面を参照しながら詳細に説明する。

## 【 0 0 1 5 】

本発明は、例えば図 1 及び図 2 に示すように構成された補聴器 1 に適用される。この補聴器 1 は、図 1 に示すように、ヘッドマウントディスプレイ（head-mounted display: HMD）2 と、音声認識、音声情報の生成等を行うコンピュータ部 3 との間を光ファイバーケーブル 4 で接続してなる携帯型のものである。また、コンピュータ部 3 は、例えば使用者の腰部に装着されるような支持部 5 に付属して配設されており、当該支持部 5 に付属したバッテリー 6 から電力が供給されることで駆動するとともに、HMD 2 を駆動させる。

40

## 【 0 0 1 6 】

HMD 2 は、使用者の目前に配置されるディスプレイ部 7 と、使用者からの音声を検出する使用者用マイクロホン 8 と、使用者に音声を出力する音声出力部 9 と、使用者の頭部に上述の各部を配置させるように支持する支持部 5 と、外部からの音声等を検出する外部用マイクロホン 11 とを備える。

## 【 0 0 1 7 】

ディスプレイ部 7 は、使用者の目前に配されることで例えば使用者用マイクロホン 8 及び

50

／又は後述の外部用マイクロホン１１で検出した音声の意味内容等を表示する。なお、このディスプレイ部７は、コンピュータ部３からの命令に応じて、上述の音声の意味内容のみならず、他の情報を表示しても良い。

【００１８】

使用者用マイクロホン８は、使用者の口元付近に配設され、使用者が発した音声を検出する。そして、この使用者用マイクロホン８は、使用者からの音声を電気信号に変換してコンピュータ部３に出力する。

【００１９】

外部用マイクロホン１１は、丸板状に形成された音声出力部９の側面に設けられる。この外部用マイクロホン１１は、外部からの音声を検出して電気信号に変換してコンピュータ部３に出力する。

10

【００２０】

この使用者用マイクロホン８及び外部用マイクロホン１１は、配設する位置を問わず、使用者の操作に応じて、種々のマイク（骨導マイク、気導音と骨導音を拾い上げるマイクをもつ超小型送受話一体ユニットのマイク（日本電信電話株式会社製）、無指向性マイク、単一指向性（超指向性等）マイク、双指向性マイク、ダイナミックマイク、コンデンサマイク（エレクトレットマイク）、ズームマイク、ステレオマイク、ＭＳステレオマイク、ワイヤレスマイク）、セラミックマイク、マグネティックマイク、マイクロフォンアレイを用いても良い。また、イヤホンとしては、マグネティックイヤホンが使用可能である。これらのマイクの收音技術として、また、伝送技術としてエコーキャンセラ等を用いても良い。また、これらのマイクロホン８，１１は、従来より採用されている利得調整器と音声調整器と出力制御装置（maximam output power control式、automatic recruitment control コンプレッション式等）を適用したものが使用可能である。

20

【００２１】

更に、使用者用マイクロホン８及び外部用マイクロホン１１は、図１に示すように、別個に設ける一例のみならず、一体に構成されたものであっても良い。

【００２２】

支持部５は、例えば形状記憶合金等の弾性材料等からなり、使用者の頭部に固定可能とすることで、上述のディスプレイ部７，使用者用マイクロホン８，音声出力部９を所定の位置に配設可能とする。なお、この図１に示した支持部５は、使用者の額から後頭部に亘って支持部材を配設することでディスプレイ部７等を所定位置に配設するものの一例について説明したが、所謂ヘッドホン型の支持部であっても良いことは勿論であり、音声出力部９を両耳について設けても良い。

30

【００２３】

コンピュータ部３は、例えば使用者の腰部に装着される支持部５に付属されてなる。このコンピュータ部３は、図２に示すように、例えばマイクロホン８，１１で検出して生成した電気信号が入力される。このコンピュータ部３は、電気信号を処理するためのプログラムを格納した記録媒体、この記録媒体に格納されたプログラムに従って音声認識、音声情報の生成処理を行うＣＰＵ（Central Processing Unit）等を備えてなる。なお、このコンピュータ部３は、腰部のみならず、頭部のＨＭＤ２と一体化しても良い。

40

【００２４】

コンピュータ部３は、使用者用マイクロホン８及び／又は外部用マイクロホン１１で検出した音声から生成した電気信号に基づいて、記録媒体に格納されたプログラムを起動することで、ＣＰＵにより音声認識処理を行うことで、認識結果を得る。これにより、コンピュータ部３は、ＣＰＵにより、使用者用マイクロホン８及び／又は外部用マイクロホン１１で検出した音声の内容を得る。

【００２５】

つぎに、本発明を適用した補聴器１の電氣的な構成について図２を用いて説明する。この補聴器１は、音声を検出して音声信号を生成する上述のマイクロホン８，１１に相当するマイクロホン２１と、マイクロホン２１で生成された音声信号が入力され音声認識処理を

50

行う上述のコンピュータ部 3 に含まれる信号処理部 2 2、信号処理部 2 2 からの認識結果に基づいて音声情報を生成する上述のコンピュータ部 3 に含まれる音声情報生成部 2 3 と、音声データが記憶され信号処理部 2 2 及び音声情報生成部 2 3 にその内容が読み込まれる上述のコンピュータ部 3 に含まれる記憶部 2 4 と、音声情報生成部 2 3 からの音声情報を用いて音声出力する上述の音声出力部 9 に相当するスピーカ部 2 5 と、音声情報生成部 2 3 からの音声情報を用いて当該音声情報が示す内容を表示する上述のディスプレイ部 7 に相当する表示部 2 6 とを備える。

【0026】

上記マイクロホン 2 1 は、例えば喉頭摘出後の代用発声法を用いて発せられた使用者からの音声又は外部からの音声を検出して、当該音声に基づく音声信号を生成する。そして、このマイクロホン 2 1 は、生成した音声信号を信号処理部 2 2 に出力する。

10

【0027】

また、このマイクロホン 2 1 は、使用者の口元付近に配設され、使用者が発した音声を検出する。また、このマイクロホン 2 1 は、外部からの音声を検出して音声信号を生成する。なお、以下の説明においては、使用者の音声を検出するマイクロホンを上述と同様に使用者用マイクロホン 8 と呼び、外部からの音声を検出するマイクロホンを上述と同様に外部用マイクロホン 1 1 と呼び、双方を総称するときには単にマイクロホン 2 1 と呼ぶ。

【0028】

上記代用発声法としては、例えば人工喉頭（電気式、笛式）、食道発声及び種々の音声再建術を実現するための機構である。

20

【0029】

上記信号処理部 2 2 は、マイクロホン 2 1 からの音声信号を用いて音声認識処理を行う。この信号処理部 2 2 は、例えば内部に備えられたメモリに格納した音声認識処理を行うためのプログラムに従った処理を行うことにより音声認識処理を実行する。具体的には、この信号処理部 2 2 は、使用者の音声をサンプリングして生成し記憶部 2 4 に格納された音声データを参照し、マイクロホン 2 1 からの音声信号を言語として認識する処理を行う。この結果、この信号処理部 2 2 は、マイクロホン 2 1 からの音声信号に応じて認識結果を生成する。

【0030】

この信号処理部 2 2 は、例えば認識対象音声による分類と対象話者による分類の音声認識処理があり、認識対象音声による分類の音声認識処理では単語音声認識（isolated word recognition）と連続音声認識（continuous speech recognition）がある。また、音声情報生成部 2 3 は、連続単語音声認識には連続単語音声認識（continuous word recognition）と文音声認識（sentence speech recognition）、会話音声認識（conversational speech recognition）、音声理解（speech understanding）がある。また対象話者による分類では不特定話者型（speaker independent）、特定話者型（speaker dependent）、話者適応型（speaker adaptive）等がある。この信号処理部 2 2 が行う音声認識手法としては、ダイナミックプログラミング（Dynamic Programming）マッチングによるもの、音声の特徴によるもの、隠れマルコフモデル（HMM）によるものがある。

30

【0031】

また、信号処理部 2 2 は、入力した音声を用いて話者認識を行う。このとき、信号処理部 2 2 は、使用者の話者からの音声の特徴を抽出する処理や音声の周波数特性を用いて話者認識結果を生成して音声情報生成部 2 3 に出力する。また、信号処理部 2 2 は、話者による変動が小さな特徴量を用いる方法、マルチテンプレート法、統計的手法を用いて不特定話者認識を行う。また、話者適応には、個人差の正規化法、話者間の音声データの対応関係によるもの、モデルパラメータの更新によるもの、話者選択によるものがある。この信号処理部 2 2 では、以上の音声認識を使用者の身体状態、利用状態及び使用目的に応じて行う。

40

【0032】

ここで、使用者の身体状態とは使用者の難聴や言語障害の程度等を意味し、利用状態とは

50

使用者が補聴器 1 を使用する環境（室内、野外、騒音下）等を意味し、使用目的とは使用者が補聴器 1 を利用するときの目的、すなわち認識の向上させることや、使用者が理解しやすいようにすること等であって、例えば普段話す人と対話することや、不特定多数の人と対話することや、音楽（オペラ、演歌）を観覧することや、講演を聴くことや、言語障害者と対話することである。

【 0 0 3 3 】

また、この信号処理部 2 2 は、マイクロホン 2 1 に入力した音声を記憶し、学習する機能を有する。具体的には、信号処理部 2 2 は、マイクロホン 2 1 で検出した音声の波形データを保持しておき、後の音声認識処理に用いる。これにより、信号処理部 2 2 は、更に音声認識を向上させる。更に、この信号処理部 2 2 は、学習機能を備えることで出力する結果を正確にすることができる。

10

【 0 0 3 4 】

上記記憶部 2 4 には、上記信号処理部 2 2 が入力された音声を認識するときに、入力された音声を検出することで生成した音声波形と比較される音声モデルを示すデータが格納されている。また、記憶部 2 4 には、例えば喉頭摘出前に発声した声帯振動による発声機構を持つ使用者の音声や、出力することを希望する音声を予めサンプリングして得たデータが音声データとして格納されている。

【 0 0 3 5 】

音声情報生成部 2 3 は、信号処理部 2 2 からの認識結果及び記憶部 2 4 に格納された使用者の音声を示す音声データを用いて、音声情報を生成する。このとき音声情報生成部 2 3 は、認識結果に応じて、記憶部 2 4 に格納された音声データを組み合わせるとともに、認識結果を加工変換して音声情報を生成する。このとき、音声情報生成部 2 3 は、内蔵した CPU、音声情報生成プログラムを用いて音声情報を生成する。

20

【 0 0 3 6 】

また、この音声情報生成部 2 3 は、認識結果を用いて音声から音声分析し、当該音声分析した音声の内容に応じて、音声データを再構成するという処理を行うことで、音声を示す音声情報を生成する。そして、音声情報生成部 2 3 は、生成した音声情報をスピーカ部 2 5 及び表示部 2 6 に出力する。

【 0 0 3 7 】

更に、音声情報生成部 2 3 は、信号処理部 2 2 からの認識結果を、使用者の身体状態、利用状態及び使用目的に応じて加工、変換、合成等をして音声情報を生成する処理を行う。更に、この音声情報生成部 2 3 は、マイクロホン 2 1 で検出された音声を使用者に提示するための処理を認識結果及び / 又は加工等をして得た認識結果について行う。

30

【 0 0 3 8 】

更にまた、音声情報生成部 2 3 は、認識結果から生成した音声情報を修飾して新たな音声情報を生成しても良い。このとき、音声情報生成部 2 3 は、使用者の身体状態、利用状態及び使用目的に基づいて、更に使用者が理解し易い言葉を付け加えることで、使用者の音声の認識を更に向上させる。

【 0 0 3 9 】

更にまた、この音声情報生成部 2 3 は、音声情報を表示部 2 6 に出力するときに音声の意味内容を画像として表示部 2 6 に出力する。音声情報生成部 2 3 は、例えば使用者又は使用者の話者及び外部からの音声が入力されて信号処理部 2 2 からの認識結果として物体を示す認識結果が入力されたときには、当該物体を示す画像データを表示部 2 6 に出力して表示させる処理を行う。

40

【 0 0 4 0 】

更にまた、この音声情報生成部 2 3 は、信号処理部 2 2 からの認識結果に応じて、以前にスピーカ部 2 5 又は表示部 2 6 に出力した音声情報を再度出力する。音声情報生成部 2 3 は、音声情報を出力した後に、使用者又は使用者に対する話者がもう一度聞き直したいことに応じて発した音声を示す認識結果が入力されたとき判定したときには、スピーカ部 2 5 又は表示部 2 6 に出力した音声情報を再度出力する処理を行う。また、音声情報生成部 2

50



3は、例えば使用者の話者からの音声の特徴を抽出する処理や音声の周波数特性を用いた話者認識結果に基づいて、以前にスピーカ部25又は表示部26に出力した音声情報を再度出力しても良い。更に、音声情報生成部23は、人工知能の機能を用いて音声対話を行うことで、スピーカ部25又は表示部26に出力した音声情報を再度出力しても良い。

【0041】

更にまた、音声情報生成部23は、再度出力する処理を行うか否かを操作入力部28からの操作入力命令に応じて切り換えても良い。すなわち、使用者が再度出力する処理を行うか否かの切換を操作入力部28を操作することで決定し、操作入力部28をスイッチとして用いる。

【0042】

また、この音声情報生成部23は、再度音声情報を出力するとき、以前に出力した音声情報を再度出力するか、以前に出力した音声情報とは異なる音声情報を出力するかを、信号処理部22を介して入力される操作入力部28からの操作入力信号に応じて選択する。

【0043】

表示部26は、上記音声情報生成部23で生成した音声情報が示す音声、カメラ機構29で撮像した画像等を表示する。

【0044】

操作入力部28は、スイッチ、キーボード、マウス等でも良く、使用者に操作されることで、操作入力信号を生成する。

【0045】

このような補聴器1は、マイクロホン21で検出した音声について信号処理部22で音声認識処理をして、認識結果に基づいて音声情報生成部23でプログラムを起動することで使用者に応じた処理を行うことができる。これにより、補聴器1は、スピーカ部25にマイクロホン21からの音声出力するとともに、表示部26に表示するので、音声に対する使用者の認識を向上させることができる。視覚聴覚同時に矛盾する音韻情報を提示した場合にいずれの情報とも異なった音韻に異聴が生ずるというMuGurk効果(MuGurk H and MacDonald J: Hearing lips and seeing voice, Nature 264, 746-8, 1976参照)や、乳児がすでに聴覚からの音声情報と視覚からの口形の情報との対応関係を獲得しているとの報告(Kuhl PK et al. Human processing of auditory-visual information in speech perception. ICSLP'94 S11.4, Yokohama, 1994)や視覚が音源方向の知覚に影響を与える(腹話術効果)、及び人間は無意識のうちに音源かどうかを学習し、区別するなどの報告は人間のコミュニケーションが本来マルチモーダルなものであるとする仮説を支持するものである(Saitou H and Mori T: 視覚認知と聴覚認知 Ohmsha, 119-20, 1999参照)。以上のことは、視覚が聴覚に影響を及ぼしていることを意味し、表示部26に認識結果等を表示することで音声情報を補足し、音声に対する使用者の認識を向上させる。この補聴器1では、音声のみならず、表示部26に表示する画像を通じて話者に音声の意味内容を伝達し、対話することができる。

【0046】

更に、この補聴器1によれば、使用者用マイクロホン8及び/又は外部用マイクロホン11で検出した音声を認識した結果に応じて表示部26に表示する音声の意味内容及びスピーカ部25から出力する音声の内容を変更させることができるので、更に音声に対する使用者の認識を向上させることができる。従って、この補聴器1によれば、音声情報生成部23により音声認識処理を変更するプログラムを実行することにより、身体状態(難聴の程度等)、利用状態及び使用目的に応じて認識処理を変更することで、使用者が理解しやすい音声の意味的な情報を表示することで更に認識を向上させることができる。

【0047】

スピーカ部25は、上記音声情報生成部23で生成した音声出力する。このスピーカ部25としては、例えば使用者から話し手に対して音声出力するものであっても良く、更には、使用者が発した音声を使用者の耳に対して発声するように音声出力するものであっても良い。また、使用者の耳に対して発声するように音声出力するスピーカ部25は

10

20

30

40

50

、スピーカユニットの変換方式としてダイナミック型や静電型（コンデンサ型、エレクトロスタティック型）によるものでも良く、形状としてはヘッドフォン（オープンエア型、クローズド型、カナルタイプ等のイン・ザ・イヤertype等）によるものでも良い。また、スピーカ部 25 は、従来の補聴器、拡声器、集音器のスピーカによるものでも良く、使用者から話者に対して音声を出力するスピーカ部 25 は従来から用いられているスピーカ装置でよい。

【0048】

また、スピーカ部 25 は、音声情報に基づいて出力する音声と逆位相の音を出力するようにしても良い。これにより、スピーカ部 25 から出力する音声に含まれる雑音成分を除去し、使用者及び / 又は使用者に対する話者に雑音の少ない音声を出力する。

10

【0049】

また、この補聴器 1 は、外部の通信ネットワークと接続された通信回路 27 を備えている。この通信回路 27 は、電話、携帯電話、インターネットや無線、衛星通信等の通信ネットワークを介して例えば音声言語障害を有する者から発せられた音声が入力される。この通信回路 27 は、外部からの音声や音声を示すデータを信号処理部 22 に入力する。また、この通信回路 27 は、音声情報生成部 23 で生成した音声情報を外部のネットワークに出力する。

【0050】

また、この通信回路 27 は、信号処理部 22、音声情報生成部 23 を介して文字放送、文字ラジオを表示部 26 で表示させても良い。このとき、通信回路 27 は、文字放送等を受信するためのチューナ機能を備え、使用者の所望のデータを受信する。

20

【0051】

このように構成された補聴器 1 は、例えば喉頭摘出後に電気式人工喉頭を使って発声された音声マイクロホン 21 に入力された場合であっても、信号処理部 22 で音声認識し、記憶部 24 に格納された喉頭摘出前にサンプリングした音声を示す音声データを用いて音声情報生成部 23 で出力する音声を示す音声情報を生成するので、スピーカ部 25 から喉頭摘出前の使用者の音声に近似した音声を出力することができる。

【0052】

なお、上述した本発明を適用した補聴器 1 の説明においては、マイクロホン 21 で検出される喉頭摘出した人の音声である一例について説明したが、聴力障害による言語障害の一つである構音障害（articulation disorders）を持つ者からの音声を検出したときであっても良い。このとき、補聴器 1 は、言語障害の音声を音声データとして記憶部 24 に記憶しておき、当該発声者が発声したことに応じて記憶部 24 に格納された発声者の音声を示す音声データを参照して信号処理部 22 で音声認識処理を行い、音声情報生成部 23 で認識結果に応じて音声データを組み合わせることで音声情報を生成する処理を行うことにより、スピーカ部 25 から音声言語障害のない音声を出力するとともに、表示部 26 により音声情報に基づいた音声内容を表示することができる。

30

【0053】

したがってこの補聴器 1 によれば、例えば喉頭摘出者が代用発声法により発生した音声を表示部 26 に表示することで不自然な音声を訂正させることができる。

40

【0054】

更に、補聴器 1 は、例えば聴力障害による構音障害を持つ者は発声のためのフィードバックが得られず、「きょうは（今日は）」という音声「きょんわあ」となってしまうのを上述した処理を行うことにより正常な「きょうは（今日は）」という音声に訂正してスピーカ部 25 から出力することができる。

【0055】

更に、この補聴器 1 は、表示部 26 を備えているので、発声者の音声をスピーカ部 25 から正常な音声にして出力するとともに、発声者の音声内容を表示することにより音声障害者や難聴者の言語訓練学習にとって好適なシステムを提供することができる。

【0056】

50

つぎに、上述の音声情報生成部 2 3 が信号処理部 2 2 からの認識結果を加工、変換して音声情報を生成する処理、音声データを組み合わせる処理で適用することができる種々の例について述べる。なお、変換処理等の種々の例は、以下に述べる例に限定するものではない。

#### 【0057】

音声情報生成部 2 3 は、信号処理部 2 2 からの認識結果を変換するとき、人工知能技術を用いて認識結果を加工変換して音声情報を生成しても良い。音声情報生成部 2 3 は、例えば音声対話システムを用いる。ここで、特に聴力の低下した老人は相手話者の言ったことを再度聞き直すことがあるが、このシステムを用いて認識結果を加工変換することにより、補聴器 1 と使用者とが対話して以前に記憶した相手話者の言ったことの情報を得て、使用者の音声認識を向上させることができ、聞き直す手間を省略することができる。

10

#### 【0058】

このようなシステムは、マルチモーダル対話システムである表情つき音声対話システムを用いることで実現可能である。このマルチモーダル対話システムでは、ポインティングデバイスとタブレットを利用する入力技術である直接操作・ペンジェスチャ技術、テキスト入力技術、音声認識等の音声入出力技術、人間の視覚、聴覚、触覚、力覚を利用した仮想現実感技術、ノンバーバルモダリティ技術の技術要素をモダリティとし組み合わせて用いる。このとき、音声情報生成部 2 3 は、言語情報を補足する手段、対話の文脈情報（或いはその補足手段）、使用者の認知的負担或いは心理的抵抗感を軽減する手段として各モダリティを用いる。なお、ノンバーバルインターフェースとして身振り（gesture）インターフェースを用いてもよい。その場合ジェスチャーインターフェースの計測として装着型センサによる身振り計測には身振りトラッキングが必要であり手袋型デバイス、磁気や光学的位置計測を用い、身振りの非接触計測にはマーカを立体解析したりする映像や 3 D 再構成によるものを用いてもよい。

20

#### 【0059】

なお、このマルチモーダル対話システムの詳細は文献「Nagao K and Takeuchi A, Speech dialogue with facial displays: Multimodal human-computer conversation. Proc. 32nd Ann Meeting of the Association for Computational Linguistics, 102-9, Morgan Kaufmann Publishers, 1994 及び Takeuchi A and Nagao K, Communicative facial displays as a new conversational modality. Proc ACM/IFIP Conf on Human Factors in Computing Systems (INTERCHI'93), 187-93, ACM Press, 1993」に記載されている。

30

#### 【0060】

このような人工知能機能を用いた音声対話システムとしては、マイクホン 2 1 で検出した音声を、信号処理部 2 2 で A / D 変換、音響分析、ベクトル量子化の後、音声認識モジュールによって、上位スコアをもつ単語レベルの最良仮説を生成するシステムが使用可能である。ここで、音声情報生成部 2 3 は、隠れマルコフモデル（HMM）に基づく音韻モデルを用いて、ベクトル量子コードから音素を推定し、単語列を生成する。音声情報生成部 2 3 は、生成した単語列を、構文・意味解析モジュールにより意味表現に変換する。このとき、音声情報生成部 2 3 は、単一化文法を用いて構文解析を行い、次にフレーム型知識ベースと事例ベース（例文を解析して得られた文パターン）を用いて曖昧さの解消を行う。発話の意味内容の決定後、プラン認識モジュールにより使用者の意図を認識する。これは対話の進行に従い動的に修正・拡張されていく使用者の信念モデルと対話のゴールに関するプランに基づいている。意図を認識する課程で、主題の管理や、代名詞の照応解消、省略の補完などを行う。そして使用者の意図に基づいて協調的な応答を生成するモジュールが起動する。このモジュールはあらかじめ用意されたテンプレートの発話パターンに領域知識により得られた応答に関する情報を埋め込むことにより発話を生成する。この応答は音声合成モジュールにより音声となる。なお、この信号処理部 2 2 及び音声情報生成部 2 3 が行う処理としては、例えば文献（Nagao N, A preferential constraint satisfaction technique for natural language analysis. Proc 10th European Conf on Artificial Intelligence, 523-7, John Wiley & Sons, 1992）、（Tanaka H, Natural language proces

40

50

sing and its applications,330-5,1999,電子情報通信学会編 コロナ社)、(Nagao K, Abduction and dynamic preference in plan-based dialogue understanding.Proc 13th Int joint Conf on Artificial Intelligence,1186-92,Morgan Kaufmann Publishers,1993)に記載された処理を行うことでも実現可能である。

#### 【0061】

また、音声情報生成部23は、人工知能機能を用いて行う処理として、システムの擬人化を行い、音声認識、構文・意味解析、プラン認識より表情パラメータ調節、表情アニメーションを表示部26を用いて行うことにより、視覚的手段を用いて音声対話に対して使用者の認知的負担、心理的抵抗感を軽減する。なお、この音声情報生成部23が行う処理としては、FACS(Facial Action Coding System; Ekman P and Friesen WV, Facial Action Coding System.Consulting Psychologists Press Palo Alto,Calif,1978)に記載された処理を行うことができる。

10

#### 【0062】

更にまた、音声情報生成部23は、音声対話コンピュータシステム(参照Nakano M et al,柔軟な話者交代を行う音声対話システムDUG-1,言語処理学会第5回年次大会論文集,161-4,1999)としては話し言葉を理解する逐次理解方式(Incremental Utterance Understanding: Nakano M, Understanding unsegmented user utterances in real-time spoken dialogue systems.Proc of the 37th Ann meeting of the association for computational linguistics,200-7)と内容の逐次変更が可能な逐次生成方式(Incremental Utterance Production: Dohsaka K and Shimazu A,A computational model of incremental utterance production in task-oriented dialogues. Proc of the 16th Int Conf on Computational Linguistics, 304-9, 1996. 及びDohsaka K and Shimazu A,System architecture for spoken utterance production in collaborative dialogue. Working Notes of IJCAI 1997 Workshop on Collaboration, Cooperation and Conflict in Dialogue Systems, 1997及び Dohsaka K et al,複数の対話ドメインにおける協調的対話原則の分析、電子情報通信学会技術研究報告NLC-97-58,25-32,1998)による音声と画像を用いる人工知能システムである。ここで、音声情報生成部23は、理解と応答のプロセスが平行動作する。また、音声情報生成部23は、ISTARプロトコール(参照Hirasawa J,Implementation of coordinative nodding behavior on spoken dialogue systems, ICSLP-98,2347-50,1998)を用いて音声認識と同時に単語候補を言語処理部に逐次的に送る。

20

30

#### 【0063】

すなわち、音声対話システムDUG-1で用いている技術を用いることにより、補聴器1では、例えば所定のデータ量(文節)ごとに使用者及び/又外部からの音声を音声認識するとともに、音声情報を生成する処理を行う。音声情報生成部23では、使用者及び/又は外部からの音声に応じて、音声認識処理、音声情報認識処理を随時中止、開始することができ、効率的な処理を行うことができる。更に、この補聴器1では、使用者の音声に応じて、音声認識処理、音声情報生成処理を制御することができるので、柔軟に話者の交替を実現することができる。すなわち、音声情報を生成している最中に使用者及び/又は外部からの音声を検出することで処理を変更し、使用者に提示する音声情報の内容を変更等の処理を行うことができる。

40

#### 【0064】

更にまた、音声情報生成部23は、キーワードスポティングを用いて使用者の自由な発話を理解する処理を行っても良い(Takabayashi Y,音声自由対話システム TOSBURG II - 使用者中心のマルチモーダルインターフェースの実現に向けて-.信学論 vol J77-D-II No.8 1417-28,1994)。

#### 【0065】

この音声情報生成部23は、例えばアクセント等の処理を行うように変換処理を行って音声情報を出力しても良い。このとき、音声情報生成部23は、必要に応じて、特定の発音についてはアクセントの強弱を変化させるように音声情報を変換して出力するようにする。

50

## 【 0 0 6 6 】

音声情報生成部 2 3 は、音声データを合成するとき、どのような内容の音声でも合成するときには規則による音声合成、滑らかな音声を合成するために可変長単位を用いた音声合成、自然な音声を合成するための韻律制御、また音声の個人性付与のために音質変換を行って音声情報を生成しても良い。これは、例えば書籍「"自動翻訳電話" ATR国際電気通信基礎技術研究所編 pp.177-209, 1994オーム社」に記載されている技術を適用することにより実現可能である。

## 【 0 0 6 7 】

また、ボコーダ (vocoder) 処理を用いても高品質の音声を合成することが可能である。例えば音声分析変換合成法 STRAIGHT (speech transformation and representation based on adaptive interpolation of weighted spectrogram) 等を施すことで実現可能である (文献「Maeda N et al, Voice Conversion with STRAIGHT. TECHNICAL REPORT OF IEICE, EA98-9, 31-6, 1998」参照)。

## 【 0 0 6 8 】

更に、この音声情報生成部 2 3 は、文字情報から音声を作り出す音声合成 (text to speech synthesis) 技術を用いることにより話の内容に関する情報 (音韻性情報) や音の高さや大きさに関する情報 (韻律情報) を聴力障害者の難聴の特性に合わせてその人の最も聞き易い音の高さに調整することも可能であり、他に話速変換技術 (voice speed converting)、周波数圧縮 (frequency compress) 処理などの音声特徴量の変換処理を行う。また出力する音声の帯域を調整する帯域拡張 (frequency band expansion) 処理や、音声強調 (speech enhancement) 処理等を音声情報に施す。帯域拡張処理、音声強調処理としては、例えば「Abe M, "Speech Modification Methods for Fundamental Frequency, Duration and Speaker Individuality," TECHNICAL REPORT OF IEICE, SP93-137, 69-75, 1994」にて示されている技術を用いることで実現可能である。なお、上述したように、信号処理部 2 2 及び音声情報生成部 2 3 で音声認識処理をして認識結果を加工変換する場合のみならず、上記処理のみを行ってスピーカ部 2 5 に出力しても良い。また、この補聴器 1 では、認識結果及び / 又は上記処理のみを行った結果を同時に又は時間差を付けて出力しても良い。また、この補聴器 1 では、認識結果及び / 又は上記処理のみを行った結果をスピーカ部 2 5 又は表示部 2 6 の右チャンネルと左チャンネルとで異なる内容を出力しても良い。

## 【 0 0 6 9 】

更にまた、上記音声情報生成部 2 3 は、認識結果を用いて音声から言語を理解し、当該理解した言語を用いて音声データから音声情報を構成するという処理を行うのみならず、他の処理を認識結果に基づいて理解した言語を必要に応じて加工変換する処理を行っても良い。すなわち、この音声情報生成部 2 3 は、音声情報を構成するとともに、音声情報としてスピーカ部 2 5 に出力するときの速度を変化させる話速変換処理を行っても良い。すなわち、この話速変換処理は、使用者の状態に応じて適当な話速を選択することによりなされる。

## 【 0 0 7 0 】

更にまた、この音声情報生成部 2 3 は、認識結果に応じて、例えば日本語の音声情報を英語の音声情報に変換して出力するような翻訳処理を行って出力しても良く、通信機能と合わせて自動翻訳電話にも応用可能である。更には音声情報生成部 2 3 は自動要約 (automatic abstracting) を行い、「United States of America」を「USA」と要約するように変換して音声情報を出力しても良い。

## 【 0 0 7 1 】

音声情報生成部 2 3 が行う他の自動要約処理としては、例えば文章内から要約に役立ちそうな手がかり表現を拾い出し、それらをもとに読解可能な文表現を生成する生成派の処理 (文献「McKeown K and Radev DR, Generating Summaries of Multiple News Articles. In Proc of 18th Ann Int ACM SIGIR Conf on Res and Development in Information Retrieval, 74-82, 1995 及び Hovy E, Automated Discourse Generation using Discourse Stru

10

20

30

40

50

cture Relations, Artificial Intelligence, 63, 341-85, 1993」参照)、要約を「切り抜き」と考えて処理し客観的評価が可能となるように問題を設定しようという立場の抽出派の処理(文献「Kupiec J et al, A Trainable Document Summarizer, In Proc of 14th AnnInt ACM SIGIR Conf on Res and Development in Information Retrieval, 68-73, 1995」、及び「Miike S, et al, A Full-text Retrieval System with a Dynamic Abstract Generation Function. Proc of 17th Ann Int ACM SIGIR Conference on Res and Development in Information Retrieval, 152-9, 1994」及び「Edmundson HP, New Method in Automatic Extracting. J of the ACM, 16, 264-85, 1969」参照)がある。更に、この音声情報生成部23は、例えば文献「Nakazawa M, et al. Text summary generation system from spontaneous speech, 日本音響学会講演論文集 1-6-1, 1-2, 1998」に記載されている手法(Partial Matching MethodとIncremental Reference Interval-Free連続DPを用いて重要キーワードの抽出を行い、Incremental Path Methodを用いて単語認識を行う)を用いることが可能である。

10

#### 【0072】

更にまた、この音声情報生成部23は、認識結果に応じて、特定の音素、母音、子音、アクセント等において、消去したり、音声を出力することに代えてブザー音、あくび音、せき音、単調な音等を音声情報とともに出力するように制御しても良い。このとき、音声情報生成部23は、例えば文献「Warren RM and Perceptual Restoration of Missing Speech Sounds, Science vol.167, 392-393, 1970」や文献「Warren RM, Obusek CJ, "Speech perception and phonemic restoration," Perception and psychophysics vol.9, 358-362, 1971」に記載されている手法を実現した処理を音声情報について行う。

20

#### 【0073】

更にまた、音声情報生成部23は、認識結果を用いてホーン調となるように音質を変換させて音声情報を出力しても良い。上記ホーン調とは、集音管を使ったもので、約2000Hz以下の帯域の音声を増幅させて、利得を約15dB程度とすることである。すなわち、このホーン調とは、管共鳴を用いた重低音を再生する技術により出力される音質である。この音声情報生成部23は、例えばUS PATENT 4628528により公知となされているアコースティックウェーブ・ガイド(acoustic wave guide)技術を用いて出力される音質に近似した音に変換して音声情報を出力する。ここで、音声情報生成部23は、例えば低音のみを通過させるフィルター処理を行って音声情報を出力する処理を行っても良く、例

30

#### 【0074】

更にまた、この音声情報生成部23は、例えばマイクロホン21に音楽が入力されたと判断したときには、音声情報を変換して表示部26に音符や色を表示するように処理を行っても良い。また、この音声情報生成部23は、音声のリズムなどが分かるために変換した音声のリズムを信号が点滅するように音声情報を変換して表示部26に表示しても良い。

#### 【0075】

更にまた、この音声情報生成部23は、例えば警報等の発信音がマイクロホン21に入力されたと判断したときには、音声情報を変換することで表示部26に警報等がマイクロホン21で検出された旨の表示を行ったり、スピーカ部25に警報の内容を知らせるような内容を出力しても良く、例えば救急車や非常ベルのサイレンを聞いたら表示するだけでなく大音量で「救急車ですよ」や「火事ですよ」とスピーカ部25から出力するとともに、表示部26に救急車や火事を示す画像を表示することで難聴者に非常事態を伝えることができ、最悪の事態を避けることができる。

40

#### 【0076】

更にまた、音声情報生成部23は、過去に行った変換合成処理について記憶する機能を備えていても良い。これにより、音声情報生成部23は、過去に行った変換合成処理の改良を自動的に行う学習処理を行うことができ、変換合成処理の処理効率を向上させることが

50

できる。

【0077】

更にまた、この信号処理部22及び音声情報生成部23は、話し手の音声のみについての認識結果を生成して音声情報を生成し、スピーカ部25及び/又はディスプレイ部7に提示することで使用者に知らせる一例のみならず、例えば特定の雑音に対してのみ音声認識を行っても良い。要するに、信号処理部22及び音声情報生成部23は、入力した音について音声認識処理を行って、認識結果を使用者の身体状態、利用状態及び使用目的に応じて変換することで使用者が理解し易い表現で音声情報を生成して出力する処理を行う。

【0078】

更にまた、上述した本発明を適用した補聴器1の説明においては、記憶部24に予めサンプリングして格納した音声データを音声情報生成部23により組み合わせることにより音声情報を生成して出力するものの一例について説明したが、上記音声情報生成部23は、記憶部24に記憶された音声データを組み合わせる音声情報を生成するときに格納された音声データに変換処理を施す音声データ変換部を備えていても良い。このような音声データ変換部を備えた補聴器1は、例えばスピーカ部25から出力する音声の音質を変化させることができる。

【0079】

更にまた、上述した本発明を適用した補聴器1の説明においては、例えば喉頭摘出前の使用者の音声を予めサンプリングすることにより得た音声データを記憶部24に格納するものの一例について説明したが、記憶部24には、一つの音声データのみならず複数の音声データを予めサンプリングして格納しても良い。すなわち記憶部24には、例えば喉頭摘出前に発せられた音声を予めサンプリングした音声データ、及び前記喉頭摘出前に発せられた音声に近似した音声データを格納しても良く、更には全く異なる音質の音声データを格納しても良く、更にまた、喉頭摘出前の音声データを生成し易い音声データを格納しても良い。このように複数の音声データが記憶部24に格納されているとき、音声情報生成部23は、各音声データの関係を例えば関係式等を用いて関連づけを行って選択的に音声データを用いて音声情報を生成しても良い。

【0080】

また、上述の補聴器1は、サンプリングして記憶部24に格納した音声データを合成することで音声情報を生成して出力する一例について説明したが、記憶部24に記憶されている音声データを合成することで生成した音声情報に、音声情報生成部23によりボコーダ処理を施すことにより、サンプリングして記憶されている音声データが示す音声とは異なる音質の音声に変換して出力しても良い。このとき、音声情報生成部23は、ボコーダ処理を用いた例としてSTRAIGHTを施す。

【0081】

更にまた、信号処理部22は、話者認識(speaker recognition)処理を入力される音声について行って各話者に対応した認識結果を生成しても良い。そして、この信号処理部22では、各話者に関する情報を認識結果とともにスピーカ部25や表示部26に出力することで使用者に提示しても良い。

【0082】

補聴器1で話者認識を行うときには、ベクトル量子化(文献Soong FK and Rosenberg AE, On the use of instantaneous and transitional spectral information in speaker recognition. Proc of ICASSP '86, 877-80, 1986)によるものでも良い。このベクトル量子化を利用した話者認識では、準備段階の処理として登録話者用の学習用音声データからスペクトルの特徴を表すパラメータを抽出して、これらをクラスタリングすることによりコードブックを作成する。ベクトル量子化による方法は話者の特徴が作成された符号帳に反映されていると考える手法である。認識時には入力された音声と全ての登録話者のコードブックを用いてベクトル量子化を行い、入力音声全体に対して量子化ひずみ(スペクトルの誤差)を計算する。この結果を用いて話者の識別や照合の判定を行う。

【0083】

また、補聴器 1 で話者認識を行うときには、HMM（文献Zheng YC and Yuan BZ,Text-dependent speaker identification using circular hidden Markov models.Proc of ICASSP '88,580-2,1988）による方法であっても良い。この方法では、準備段階の処理として登録話者の学習用音声データからHMMを作成する。HMMを用いる方法では話者の特徴は状態間の遷移確率とシンボルの出力確率に反映されると考える。話者認識の段階では入力音声を用いて全ての登録話者のHMMによる尤度を計算して判定を行う。HMMの構造としてleft to rightモデルに対してエルゴディックなHMMを用いてもよい。

#### 【0084】

更にまた、補聴器 1 では、ATR-MATRIX system（ATR 音声翻訳通信研究所製：参照 Takezawa T et al, ATR-MATRIX: A spontaneous speech translation system between English and Japanese. ATR J2,29-33,June1999）で用いられている音声認識（ATRSPREC）、音声合成（CHATR）、言語翻訳（TDMT）を行うことで、マイクロホン 21 で入力した音声を翻訳して出力することができる。

#### 【0085】

上記音声認識（ATRSPREC）では、大語彙連続音声認識を行い、音声認識ツールを用いて音声認識に必要な音響モデルと言語モデルの構築、及び信号処理から探索までの工程を処理する。この音声認識では、行った処理をツール群として完結し、ツール同士の組み合わせることができる。また、この音声認識を行うとき、不特定話者の音声認識を行っても良い。

#### 【0086】

上記音声合成（CHATR）では、あらかじめデータベース化された多量の音声単位から、出力したい文に最も適した単位を選択してつなぎあわせ、音声を合成する。このため、滑らかな音声が出力することができる。この音声合成では、話し手の声に最も近い音声データを用いて話し手の声に似た声で合成することができる。また、この音声合成を行うときには、音声情報生成部 23 は、入力された音声から話し手が男性か女性かを判断し、それに応じた声で音声合成を行っても良い。

#### 【0087】

上記言語翻訳（TDMT）では、文の構造を判断する処理、対話用例を用いた対話特有のくだけた表現などの多様な表現を扱って言語翻訳を行う。また、この言語翻訳では、マイクロホン 21 が一部聞き取れなかった部分があっても、翻訳できる部分はなるべく翻訳する部分翻訳処理を行い、一文全体を正確に翻訳できない場合でも、話し手が伝えたい内容をかなりの程度相手に伝える。

#### 【0088】

また、上記音声認識、音声合成、言語翻訳を行うときには、通信回路 27 を介して携帯電話等の通信機器と接続して双方向の対話可能である。

#### 【0089】

上記音声認識、音声合成、言語翻訳を行う補聴器 1 では、例えば日英双方向の音声翻訳システムの利用、ほぼリアルタイムの認識、翻訳、合成、話し始めの指示をシステムに与える必要がなく、全二重の対話が可能自然な発話に対する、質の高い認識、翻訳、合成「あの一」、「えーと」といった言葉や、多少くだけた表現があっても認識が可能となる。

#### 【0090】

更にまた、音声情報生成部 23 は、上記音声認識（ATRSPREC）において、信号処理部 22 からの認識結果に基づいて文の構造を判断するだけでなく、対話用例を用いることにより、対話特有のくだけた表現などの多様な表現に対応した音声情報を生成する。また、音声情報生成部 23 は、マイクロホン 21 で会話中の一部が聞き取れなかった部分があっても、音声情報を生成することができる部分はなるべく音声情報を生成する。これにより、音声情報生成部 23 は、一文全体の音声情報を正確に生成できない場合でも、話し手が伝えたい内容をかなりの程度相手に伝える。このとき、音声情報生成部 23 は、翻訳処理（部分翻訳機能）を行って音声情報を生成しても良い。

#### 【0091】



また、音声情報生成部 23 は、上記音声合成 (CHATR) において、予めデータベース化して記憶された多量の音声単位の音声データから、出力したい文に最も適した単位を選択してつなぎあわせ、音声を合成して音声情報を生成する。これにより、音声情報生成部 23 は、滑らかな音声を出力するための音声情報を生成する。また、音声情報生成部 23 は、話し手の声に最も近い音声データを用いて話し手の声に似た声で合成処理を行っても良く、入力された音声から話し手が男性か女性かを判断し、それに応じた声で音声合成を行って音声情報を生成しても良い。

【0092】

更にまた、音声情報生成部 23 は、マイクロホン 21 からの音声から、特定の音源の音のみを抽出してスピーカ部 25 及び / 又は表示部 26 に出力しても良い。これにより、補聴器 1 は、複数の音源から到来する音の混合の中から、特定の音源の音のみを抽出して聞くことができるカクテルパーティ現象を人工的に作ることができる。

10

【0093】

更にまた、音声情報生成部 23 は、音韻的に近い例を用いて誤りを含んだ認識結果を訂正する手法を用いて聞き間違いを修正して音声情報を生成しても良い (文献 Ishikawa K, Sumida E: A computer recovering its own misheard-Guessing the original sentence from a recognition result based on familiar expressions- ATR J 37,10-11,1999)。このとき、音声情報生成部 23 は、使用者の身体状態、利用状態及び使用目的に応じて処理を行って、使用者にとってわかりやすい形態に加工変換する。

【0094】

20

なお、上述した補聴器 1 の説明においては、マイクロホン 21 で検出した音声について音声認識処理、音声生成処理を行う一例について説明したが、使用者等により操作される操作入力部 28 を備え当該操作入力部 28 に入力されたデータを音声及び / 又は画像とるように信号処理部 22 により変換しても良い。また、この操作入力部 28 は、例えば使用者の指に装着され、指の動きを検出することでデータを生成して信号処理部 22 に出力するものであっても良い。

【0095】

また、この補聴器 1 は、例えば使用者が液晶画面等をペンにより接触させることで文字及び / 又は画像を描き、その軌跡を取り込むことによる画像に基づいて文字及び / 又は画像データを生成する文字及び / 又は画像データ生成機構を備えていても良い。補聴器 1 は、生成した文字及び / 又は画像データを信号処理部 22 及び音声情報生成部 23 により認識・変換等の処理を行って出力する。

30

【0096】

更に、上述の補聴器 1 は、マイクロホン 21 等からの音声を用いて信号処理部 22 により音声認識処理を行う一例に限らず、例えば使用者及び / 又は使用者以外の人が装着する鼻音センサ、呼気流センサ、頸部振動センサからの検出信号及びマイクロホン 21 等からの信号を用いて音声認識処理を行っても良い。このように、補聴器 1 は、マイクロホン 21 のみならず上記各センサを用いることにより、信号処理部 22 による認識率を更に向上させることができる。

【0097】

40

更に、この補聴器 1 は、例えば自動焦点機能やズーム機能を搭載したデジタルカメラにより動画像や静止画像等を撮像するカメラ機構 29 を図 2 に示すように備え、表示部 26 に表示するものであっても良い。このカメラ機構 29 は例えば図 1 のディスプレイ部 7 と一体に搭載されても良い。また、上記カメラ機構 29 としては、デジタルカメラを用いても良い。

【0098】

また、この補聴器 1 に備えられたカメラ機構 29 は、撮像した画像を使用者の視力や乱視等の状態に合わせて歪ませたり拡大させたりする画像変換処理を施して表示部 26 に表示する眼鏡機能を備えていても良い。

【0099】

50

このような補聴器 1 は、例えばカメラ機構 29 から CPU 等からなる信号処理回路を経由して表示部 26 に撮像した画像を表示する。この補聴器 1 は、このようなカメラ機構 29 により例えば話者を撮像した画像を使用者に提示することで、使用者の認識を向上させる。また、この補聴器 1 は、撮像した画像を通信回路 27 を介して外部のネットワークに出力しても良く、更には外部のネットワークからカメラ機構 29 で撮像した画像を入力して通信回路 27 及び信号処理回路等を介して表示部 26 に表示しても良い。

#### 【0100】

更に、この補聴器 1 では、話者を撮像した画像を用いて信号処理部 22 で顔面認識処理、物体認識処理を行って音声情報生成部 23 を介して表示部 26 に表示しても良い。これにより、補聴器 1 では、撮像対象者の口唇、顔の表情、全体の雰囲気等を使用者に提示して、使用者の音声認識を向上させる。

10

#### 【0101】

撮像機能を用いた顔の認識において顔の個人性特徴を抽出して個人認識をおこなうものとして、以下の方法があるがこれらに限られるものではない。

#### 【0102】

濃淡画像のマッチングにより識別するための特徴表現の一つとしてパターンをモザイク化し、各ブロック内の画素の平均濃度をブロックの代表値とすることで濃淡画像を低次元ベクトルに情報圧縮して表現する方法で M 特徴といわれている方法である。また、K I 特徴という濃淡顔画像の特徴表現で、Karhunen-Loeve (K L) 展開を顔画像の標本集合に適用して求められる直交基底画像を固有顔とよび、任意の顔画像をこの固有顔を用いて展開した係数から構成される低次元の特徴ベクトルで記述する方法である。更に、顔画像集合の K L 展開による次元圧縮に基づく K I 特徴によるもの照合パターンをまずフーリエスペクトルに変換し K I 特徴の場合と同様に標本集合を K L 展開することで次元圧縮を行って得られる低次元の特徴スペクトルである K F 特徴による識別を行う方法がある。以上の方法によるものが顔画像認識に用いることが可能であり、それらを用いて顔の認識を行うことは対話者が誰であるかという個人識別情報をコンピュータに与えることになり、使用者にとって対話者に対する情報が得られ、音声情報に対する認識が増す。なお、このような処理は、文献「小杉 信：“ニューラルネットを用いた顔画像の識別と特徴抽出”，情処学 C V 研報，73-2 (1991-07)」、文献「Turk MA and Pentland AP, Face recognition using eigenface. Proc CVPR, 586-91 (1991-06)」、文献「Akamatsu S et al, Robust face identification by pattern matching Based on KL expansion of the Fourier Spectrum 信学論 vol J76-D-II No.7, 1363-73, 1993」、文献「Edwards GJ et al, Learning to identify and track faces in image sequences, Proc of FG '98, 260-5, 1998」に記載されている。

20

30

#### 【0103】

この補聴器 1 では、物体認識を行うときには、物体を示すパターンをモザイク化しておき、実際に撮像した画像とマッチングを取ることで物体の識別を行う。そして、この補聴器 1 では、マッチングがとれた物体の動きベクトルを検出することで、物体の追尾を行う。これにより、物体から発せられる音声から生成される音声情報に対する認識が増す。この物体認識処理は、Sony CSL から提案されている Ubiquitous Talker (文献 Nagao K and Rekimoto J, Ubiquitous Talker: Spoken language interaction with real world objects. Proc 14th IJCAI-95, 1284-90, Morgan Kaufmann Publishers, 1995) で用いられている技術を採用することができる。

40

#### 【0104】

更に、この補聴器 1 は、静止画撮像用デジタルカメラのようにシャッターを押すことで静止画を撮像しても良い。更に、カメラ機構 29 は、動画像を生成して信号処理部 22 に出力しても良い。このカメラ機構 29 により動画像を撮像するときの信号方式としては、例えば M P E G (Moving Picture Experts Group) 方式などを用いる。更にまた、この補聴器 1 に備えられるカメラ機構 29 は、3 次元画像を撮像することで、話者や話者の口唇を撮像して表示部 26 に表示させることで更に使用者の認識を向上させることができる。

50

## 【 0 1 0 5 】

このような補聴器 1 は、使用者自身の発した音声や相手の発した音声等及び / 又はその場の情景を撮像した画像を記録し再生することで、言語学習における復習をすることができ言語学習に役立てることができる。

## 【 0 1 0 6 】

また、この補聴器 1 によれば、画像を拡大処理等して表示部 2 6 に表示することで相手を確認し全体の雰囲気をつかめ音声聴取の正確さが向上し、更に読唇 ( lip reading ) を行うことが可能となり認識を上昇させる。

## 【 0 1 0 7 】

更にまた、この補聴器 1 は、例えばスイッチ機構が設けられており、マイクロホン 2 1 で検出した音声をスピーカ部 2 5 により出力するか、カメラ機構 2 9 により撮像した画等像を表示部 2 6 により出力するか、又は音声及び画像の双方を出力するかを使用者により制御可能としても良い。このときスイッチ機構は、使用者に操作されることで、音声情報生成部 2 3 から出力を制御する。

## 【 0 1 0 8 】

また例として、スイッチ機構は、使用者及び / 又は使用者以外の音声を検出して、例えば「音声」という音声を検出したときにはマイクロホン 2 1 で検出した音声をスピーカ部 2 5 により出力するように切り換え、例えば「画像」という音声を検出したときにはカメラ機構 2 9 により撮像した画等像を表示部 2 6 により出力するように切り換え、「音声、画像」という音声を検出したときには音声及び画像の双方を出力するように切り換えても良く、以上のような音声認識を用いたスイッチ制御機構を備えていても良い。また、ジェスチャーインターフェースを用いることで、ジェスチャー認識によるスイッチ制御システムとしても良い。

## 【 0 1 0 9 】

更にまた、このスイッチ機構は、カメラ機構 2 9 のズーム状態等のパラメータを切り換えることでカメラ機構 2 9 で画像を撮像するときの状態を切り換える機能を備えていても良い。

## 【 0 1 1 0 】

つぎに、この補聴器 1 において、音声情報生成部 2 3 により作成した音声情報を出力する機構の種々の例について説明する。なお、本発明は、以下に説明する出力する機構に限られることはないことは勿論である。

## 【 0 1 1 1 】

すなわち、この補聴器 1 において、音声情報を出力する機構としてはスピーカ部 2 5 や表示部 2 6 に限らず、例えば骨導や皮膚刺激を利用したものであっても良い。この音声情報を出力する機構は、例えば小さな磁石を鼓膜等に装着し、磁石を振動させるものや、骨を通して信号を蝸牛に伝達するものであっても良い。

## 【 0 1 1 2 】

このような補聴器 1 は、例えば圧挺板を備え、音声情報生成部 2 3 により変換することにより得た信号を前記圧挺板に出力するようにしたものや、皮膚刺激を用いたタクタイルエイド ( Tactile Aid ) 等の触覚による補償技術を利用したものであっても良く、これらの骨振動や皮膚刺激等を用いた技術を利用することで、音声情報生成部 2 3 からの信号を使用者に伝達することができる。皮膚刺激を利用した補聴器 1 においては、音声情報生成部 2 3 からの音声情報が入力されるタクタイルエイド用振動子アレイが備えられており、タクタイルエイドと当該振動子アレイを介してスピーカ部 2 5 から出力する音声を出力しても良い。

## 【 0 1 1 3 】

また、上述した補聴器 1 の説明においては、音声情報を音声として出力するときの処理の一例について説明したが、これに限らず、例えば人工中耳により使用者に認識結果を提示するものであっても良い。すなわち、この補聴器 1 は、音声情報を電気信号としてコイル、振動子を介して使用者に提示しても良い。

10

20

30

40

50

## 【 0 1 1 4 】

更には、この補聴器 1 は、人工内耳機構を備え、人工内耳により使用者に認識結果を提示するものであっても良い。すなわち、この補聴器 1 は、例えば埋め込み電極、スピーチプロセッサ等からなる人工内耳システムに音声情報を電気信号として供給して使用者に提示しても良い。

## 【 0 1 1 5 】

更には、この補聴器 1 は、脳幹インプラント (Auditory Brainstem Implant) 機構を備え、聴性脳幹インプラントにより使用者に音声情報を提示するものであっても良い。すなわち、この補聴器 1 は、例えば埋め込み電極、スピーチプロセッサ等からなる脳幹インプラントシステムに音声情報を電気信号として供給して使用者に提示しても良い。

10

## 【 0 1 1 6 】

更にまた、この補聴器 1 は、使用者の身体状態、利用状態及び使用目的に応じて、例えば超音波帯域の音声認識可能な難聴者に対しては認識結果及び加工変換した認識結果を音声情報として超音波帯域の音声に変調・加工変換して出力しても良い。更にまた、この補聴器 1 は、超音波出力機構 (bone conduction ultrasound) を用いて超音波周波数帯域の信号を生成し、超音波振動子等を介して使用者に出力しても良い。

## 【 0 1 1 7 】

更にまた、この補聴器 1 は、ヘッドホンの接触子を耳珠に当て、骨伝導をおこしさらに耳珠、外耳道内壁の振動が気導音となるシステムである骨伝導ユニットを備え、当該骨伝導ユニットを使用して音声情報を使用者に提示しても良い。この骨導ユニットとしては、聴覚障害者用ヘッドホンシステムであるライブホン (日本電信電話株式会社製) が使用可能である。

20

## 【 0 1 1 8 】

更にまた、この補聴器 1 は、スピーカ部 2 5、表示部 2 6 等の複数の出力手段を備える一例について説明したが、これらの出力手段を組み合わせ用いても良く、更には各出力手段を単独で出力しても良い。また、この補聴器 1 では、マイクロホン 2 1 に入力した音声の音圧レベルを変化させる従来の補聴器の機能を用いて音声出力するとともに、上述した他の出力手段で認識結果を提示しても良い。

## 【 0 1 1 9 】

更にまた、この補聴器 1 は、スピーカ部 2 5 及び / 又は表示部 2 6 から出力する出力結果を同時に或いは時間差を持たせて出力してするように音声情報生成部 2 3 で制御するスイッチ機構を備えていても良く、複数回に亘って出力結果を出力するか一回に限って出力結果を出力するかを制御するスイッチ機構を備えていても良い。

30

## 【 0 1 2 0 】

また、この補聴器 1 の説明においては、図 2 に示したような一例について説明したが、入力された音声について上述した種々の加工変換処理を行って表示部 2 6 に表示させる第 1 の処理を行う CPU と、入力された音声について上述した種々の加工変換処理を行ってスピーカ部 2 5 に出力結果を出力するための第 2 の処理を行う CPU と、カメラ機構 2 9 で撮像した画像を表示するための第 3 の処理を行う CPU とを備えたものであっても良い。

## 【 0 1 2 1 】

このような補聴器 1 は、各処理を行う CPU を独立に動作させて第 1 の処理又は第 2 の処理を行わせて出力させても良く、更には各処理を行う CPU を同時に動作させて第 1 の処理、第 2 の処理、及び第 3 の処理を行わせて出力させても良く、更には、第 1 及び第 2 の処理、第 1 及び第 3 の処理又は第 2 及び第 3 の処理を行う CPU を同時に動作させて出力させても良い。

40

## 【 0 1 2 2 】

更にまた、補聴器 1 は、使用者の身体状態、利用状態及び使用目的に応じて上述した種々の出力機構からの出力結果を同時に或いは時間差を持たせて出力してするように音声情報生成部 2 3 で制御しても良い。

## 【 0 1 2 3 】

50

更に、この補聴器 1 は、複数の CPU を有し、上述した複数の CPU で行う第 1 ~ 第 3 処理のうち、少なくとも 1 の処理をひとつの CPU で行うとともに、残りの処理を他の CPU で行っても良い。

【 0 1 2 4 】

例えば、この補聴器 1 において、ひとつの CPU が入力された音声を文字データとして加工変換を行って表示部 2 6 に出力する処理 (text to speech synthesis) を行うとともに、又はひとつの CPU が入力された音声に対して文字データとして加工変換を行って他の CPU が入力された同じ音声に対して STRAIGHT 処理を行ったりしてスピーカ部 2 5 に出力する処理を行い、他の CPU が入力された音声に対してボコーダ処理のうち、例えば音声分析合成法 STRAIGHT を用いた処理を行ってスピーカ部 2 5 に出力する処理を行っても良い。すなわちこの補聴器 1 は、スピーカ部 2 5 に出力する信号と、表示部 2 6 に出力信号とで異なる処理を異なる CPU により行うものであっても良い。

10

【 0 1 2 5 】

更に、この補聴器 1 においては、上述した種々の加工変換処理を行って上述の種々の出力機構に出力する処理を行う CPU を有するとともに、加工変換処理を施さないでマイクロホン 2 1 に入力された音声を出力しても良い。

【 0 1 2 6 】

更に、この補聴器 1 においては、上述した種々の加工変換処理の一を行うための CPU と、他の加工変換処理を行う CPU とを別個に備えていても良い。

【 0 1 2 7 】

20

更に、この補聴器 1 においては、上述のように認識結果や加工変換した認識結果や撮像した画像等について音声情報生成部 2 3 で変換する処理を行うとともに、従来と同様に音声を検出して得た電気信号を増幅させて音質調整、利得調整や圧縮調整等を行いスピーカ部 2 5 に出力するものであっても良い。

【 0 1 2 8 】

なお、この補聴器 1 において、信号処理部 2 2 及び音声情報生成部 2 3 で行う処理を、例えばフーリエ変換、ボコーダ処理 (STRAIGHT 等) の処理を組み合わせることで、上述した処理を行っても良い。

【 0 1 2 9 】

また、本発明を適用した補聴器 1 では、個人的に使用する小型のタイプの補聴器について説明したが、集団で用いる大型のもの (卓上訓練用補聴器や集団訓練用補聴器) にも用いてもよい。

30

【 0 1 3 0 】

視覚ディスプレイとして HMD、Head-coupled display (頭部結合型表示装置) があげられる。以下に例を示す。双眼式 HMD (左右眼毎に視差画像を提示し立体視を可能とするものや左右眼双方に同じ画像を提示し見かけ上の大画面を与えるもの)、単眼式、シースルー型 HMD、視覚補助や視覚強調機能のついたディスプレイ、眼鏡型の双眼望遠鏡に自動焦点機能付で Visual filter を用いたもの、接眼部にコンタクトレンズを使用するシステム、網膜投影型 (Virtual Retinal Display、Retinal projection display、網膜投影型の中間型)、視線入力機能付き HMD (製品名 HAQ-200 (島津製作所)) や頭部以外 (首、肩、顔面、眼、腕、手など) にマウントするディスプレイ、立体ディスプレイ (投影式オブジェクト指向型ディスプレイ (例 head-mounted projector: linami M et al., Head-mounted projector (II) - implementation Proc 4th Ann Conf Of Virtual Reality Society of Japan 59-62, 1999)、リンク式の立体ディスプレイ、大画面のディスプレイ (spatial immersive display) (例 omnimax、CAVE (Cruz-Neira C et al. Surrounded-screen projection-based virtual reality: The design and implementation of the CAVE, Proc of SIGGRAPH'93, 135-42, 1993 参照)、CAVE 型立体映像表示装置である CABIN (Hirose M et al. 電子情報通信学会論文誌 Vol J81-D-II No.5. 888-96, 1998)、CAVE 等の投影ディスプレイと HMD の両方の特徴をもつ小型超広視野ディスプレイ (Endo T et al. Ultra wide field of view compact display. Proc 4th Ann Conf of Vir

40

50

tual Reality Society of Japan,55-58,1999)、アーチスクリーン)が使用可能である。

【0131】

特に大画面のディスプレイのものは大型補聴器として用いるときに使用してもよい。また、上述した補聴器1では、音の再現方法としてバイノーラル方式(3次元音響システムはHead-Related Transfer Functionを用いた空間音源定位システムを用いる:例 Convolution & Acoustetron II(Crystal River Engineering),ダイナミック型ドライバユニットとエレクトレットマイクロフォンを使用した補聴器TE-H50(Sony))を使用してもよく、実際と近い音場をつくったり、トランスオーラル方式(トラッキング機能付きのトランスオーラル方式が3次元映像再現におけるCAVEに対応する)を用いたりするものは主に大型の補聴器システムの場合に用いるのが好ましい。

10

【0132】

更にまた、上述のHMD2は、頭頂部に3次元位置検出センサーを備えていても良い。このようなHMD2を備えた補聴器1では、使用者の頭の動きに合わせてディスプレイ表示を変化させることが可能となる。

【0133】

強調現実感(Augmented reality(AR))を利用した補聴器1では、使用者の動作に関するセンサを備え、センサで検出した情報、マイクロホン21で検出し音声情報生成部23で生成した音声情報とを用いることで、AR空間を生成する。音声情報生成部23は、種々のセンサシステムとVR形成システムを統合するシステムとディスプレイシステムによりなるバーチャルリアリティ(Virtual reality(VR))システムとを協調的に用いることにより、実空間にVR空間を適切に重畳することで、現実感を強調するARの空間をつくることが可能となる。これにより補聴器1では視覚ディスプレイを用いるときに、顔面部にある画像からの情報を、情報が来るたびに大幅に視線をはずすことなく、ただ画像が目の前にあるだけでなく、画像情報が、いかにもそこにあるように自然に受け入れるようになり自然な状態で視覚からの情報を受け取ることが可能となる。以上を実行するには以下のシステムがある。

20

【0134】

このような補聴器1は、図3に示すように、AR空間を形成するためには、仮想環境映像生成のための3Dグラフィックアクセラレータを音声情報生成部23の内部に搭載することでコンピュータグラフィックスの立体視が可能な構成とし、更に無線通信システムを搭載する。この補聴器1に使用者の位置と姿勢の情報を取得するため、センサ31として頭部に小型ジャイロセンサ(例 データテックGU-3011)を、使用者の腰に加速度センサ(例 データテックGU-3012)を接続する。以上のセンサ31からの情報を音声情報生成部23で処理を行った後、使用者の右目、左目に対応するスキャンコンバータ32a、32bで処理をして表示部26に映像が行くというシステム(Ban Y et al, Manual-less operation with wearable augmented reality system.Proc 3th Ann Conf of Virtual Reality society of Japan,313-4,1998参照)を用いることで可能となる。

30

【0135】

また、この補聴器1では、センサ31に加えて状況認識システム(例 Ubiquitous Talker(Sony CSL))とVRシステムを形成する他のシステムである以下の種々のセンサシステムとVR形成システムを統合するシステムとディスプレイシステム、及び、この補聴器1とを協調的に用いることにより、AR空間を強化することも可能であり、マルチモダリティを用いて音声情報を補足可能となる。

40

【0136】

このようなVR等の空間を形成するには、先ず、使用者がセンサ31に本人から情報を送り、その情報がVR形成システムを統合するシステムに送られ、ディスプレイシステムから使用者に情報が送られることで実現する。

【0137】

上記センサ31(情報入力システム)として以下のデバイスがある。

【0138】

50

特に人体の動きの取り込みや、空間に作用するデバイスとして光学式 3 次元位置センサ (ExpertVision HiRES & Face Tracker (MotionAnalysis))、磁気式 3 次元位置センサ (InsideTrack (Polhemus)、3SPACE system (POLHEMUS)、Bird (Ascension Tech.))、機械式 3 D デジタイザ (MicroScribe 3D Extra (Immersion))、磁気式 3 D デジタイザ (Model 350 (Polhemus))、音波式 3 D デジタイザ (Sonic Digitizer (Science Accessories))、光学式 3 D スキャナー (3D Laser Scanner (アステックス))、生体センサ (体内の電気で測る) サイバーフィンガー (N T T ヒューマンインタフェース研究所)、手袋型デバイス (DetaGlove (VPL Res)、Super Glove (日商エレクトロニクス) Cyber Glove (Virtual Tech))、フォースフィードバック (Haptic Master (日商エレクトロニクス))、PHANTOM (SensAble Devices))、3 D マウス (Space Controller (Logitech))、視線センサ (眼球運動分析装置 (ATR 視聴覚機構研究所製))、体全体の動きの計測に関するシステム (DateSuit (VPL Res))、モーションキャプチャーシステム (HiRES (Motion Analysis))、加速度センサ (三次元半導体加速度センサ (NEC 製))、視線入力機能付き HMD がある。

#### 【 0 1 3 9 】

また、AR を実現するためには、表示部 2 6 のみならず、触覚を利用した触覚ディスプレイ、触圧ディスプレイ、力覚ディスプレイがある。触覚ディスプレイにより音声を触覚により伝え、聴覚だけでなく触覚をも加えることで音声の認識をあげることが可能となる。この触覚ディスプレイとしては、例えば振動子アレイ (オプタコンや触覚マウス、タクチュアルボコーダ等)、触知ピンアレイ (ペーパーレスブレイル等) などが使用可能である。他に water jet、air jet、PHANTOM (SensAble Devices)、Haptic Master (日商エレクトロニクス) などがある。具体的には、補聴器 1 は、VR な空間で VR キーボードを表示し、信号処理部 2 2 及び音声情報生成部 2 3 での処理を VR キーボードまたは VR スイッチにより制御する。これにより、わざわざキーボードを用意したり、スイッチまで手を伸ばしたりすることが無くなり、使用者の操作を楽にし、耳に装着するのみの補聴器と近い装用感を得ることができる。

#### 【 0 1 4 0 】

前庭感覚ディスプレイとしては、ウオッシュアウトとウオッシュバックにより狭い動作範囲の装置でも多様な加速度表現ができるシステム (モーションベット) が使用可能である。

#### 【 0 1 4 1 】

VR システムを統合するシステムとしては、以下のものがあり、それら限定されることはないが、C、C++ のライブラリとして供給され、表示とそのデータベース、デバイス入力、干渉計算、イベント管理などをサポートし、アプリケーションの部分は使用者がライブラリを使用してプログラミングするものや、ユーザプログラミングを必要とせずデータベースやイベント設定をアプリケーションツールで行い、そのまま VR シミュレーションを実行するシステムなどを使用してもよい。またこの補聴器 1 に関する個々のシステム間を通信にてつなげてよい。また、状況を高い臨場感を保って伝送するのに広帯域の通信路を使用しても良い。また、補聴器 1 では、3 D コンピュータグラフィックスの分野で用いられている以下の技術を用いてもよい。現実に関わり得ることを忠実に画像として提示し、非現実的な空間を作り、実際には不可能なことも画像として提示することがコンセプトとなる。3 D コンピュータグラフィックスでは、以下のモデリング技術、レンダリング技術、アニメーション技術により可能となる。複雑で精密なモデルを作るモデリング技術としては、ワイヤーフレームモデリング、サーフェスモデリング、ソリッドモデリング、ベジエ曲線、B - スプライン曲線、NURBS 曲線、ブール演算 (ブーリアン演算)、自由形状変形、自由形状モデリング、パーティクル、スweep、フィレット、ロフティング、メタボール等がある。また、質感や陰影をつけリアルな物体を追求するためにレンダリング技術としては、シェーディング、テクスチャマッピング、レンダリングアルゴリズム、モーションブラー、アンチエイリアシング、デプスキューイングがある。また、作成したモデルを動かし、現実の世界をシミュレーションするためのアニメーション技術として

は、キーフレーム法、インバースキネマティクス、モーフィング、シュリンクラップアニメーション、チャンネルがある。また、サウンドレンダリングとして「文献Takala T, Computer Graphics (Proc SIGGRAPH 1992) Vol26, No2, 211-20」に記載されている技術を用いても良い。

#### 【0142】

このようなVRシステムを統合するシステムとしては、例えばDivision Incのシステム(VRランタイムソフトウェア[dVS]、VR空間構築ソフトウェア[dVISE]、VR開発用ライブラリ[VC Toolkit]、SENSE 8社のWorldToolKitと、WorldUp、Superscape社のVRT、ソリッドレイ社のRealMaster、モデルなしのVRの生成として、文献「Hirose M et al. A study of image editing technology for synthetic sensation. Proc ICAT'94, 63-70, 1994」に記載されている方法等を使用しても良い。

#### 【0143】

また、本実施の形態では、HMD 2と、コンピュータ部3との間を光ファイバーケーブル4で接続してなる携帯型の補聴器1について説明したが、HMD 2とコンピュータ部3との間をワイヤレスとし、HMD 2とコンピュータ部3との間を無線や赤外線を用いた信号伝送方式等により情報の送受信を行っても良い。更に、この補聴器1においては、HMD 2とコンピュータ部3との間をワイヤレスとする場合のみならず、図2に示した各部が行う機能毎に分割して複数の装置とし、各装置間をワイヤレスとしても良く、少なくともコンピュータ部3を使用者に装着させずにHMD 2と情報の送受信を行っても良い。更にまた、この補聴器1においては、使用者の身体状態、利用状態、使用目的に応じて、図2に示した各部が行う機能毎に分割して複数の装置とし、各装置間をワイヤレスとしても良い。これにより、補聴器1は、使用者が装着する装置の重量、体積を軽減し、使用者の身体

#### 【0144】

また、補聴器1では、通信回路27を介して信号処理部22及び音声情報生成部23で行う処理の制御及びバージョンアップ、修理等をして良い。これにより、補聴器1では、通信回路27を通じて視覚ディスプレイ、聴覚ディスプレイ等を通じて修理、制御、調整等を受けることができる。

#### 【0145】

また、本発明を適用した補聴器1によれば、合成した音声を表示することで使用者に提示することができるので、例えば事務(ウェアブルコンピュータとして)、通信(自動翻訳電話への応用など)、産業医学領域(メンタルヘルスなど)、医療現場(聴力検査への利用)、外国語学習、言語訓練、娯楽(テレビゲーム)、個人用のホームシアター、コンサートや試合等の観戦、番組制作(アニメーション、実写映像、ニュース、音楽制作)、水中(ダイビングでの水中における会話など)、諜報活動や軍事、騒音下などの悪条件での作業業務(建築現場工場など)、スポーツ(自動車やヨット等のレースや、山や海等の冒険時、選手の試合時や練習時での選手同士や選手とコーチ間の意志疎通や情報変換)、や宇宙空間での作業、運輸(宇宙船や飛行機のパイロット)、カーナビゲーションシステム、VRとARとを用いた種々のシミュレーション作業(遠隔手術(マイクロサージャリー)など)等、教育、トレーニング、内科治療、傷病治療、政治、旅行、買い物、マーケティング、広告、宗教、デザインの分野、アミューズメントパーク等におけるFish-tank VR display、裸眼立体視システム、トレイグジスタンス視覚システムなどを用いたVRやARや、テレグジスタンスやアールキューブを利用したもの、電話やインターネットでの対応業務にも適用可能であり、音声言語障害者のみならず、重病患者、重度身体障害者のコミュニケーション、介護学校等の広い分野で使用可能である。

#### 【0146】

##### 【発明の効果】

以上詳細に説明したように、本発明に係る補聴器は、音声言語障害者を検出して得た認識結果に基づいて予め記憶した音声データを組み合わせて音声情報を音声に変換して外部に出力するとともに、外部からの音声を使用者に出力することができるので、喉頭摘出や舌



口腔底切除や構音障害等による音声言語障害を有する人達が本来自身がもつ、或いは自在に変換させて自然な音声で発声することを可能とするとともに、外部からの音声を使用者に出力することで使用者の聴覚を補うことができる。

【0147】

本発明に係る補聴器は、使用者の身体状態、利用状態及び使用目的に応じて、上記認識手段からの認識結果の内容を変更するように加工変換する変換手段を備えているので、使用者の身体状態、利用状態及び使用目的に応じて音声認識の結果を提示するとともに、ノイズが少ない状態で認識結果を提示することができる。

【図面の簡単な説明】

【図1】本発明を適用した補聴器の外観の一例を示すブロック図である。

10

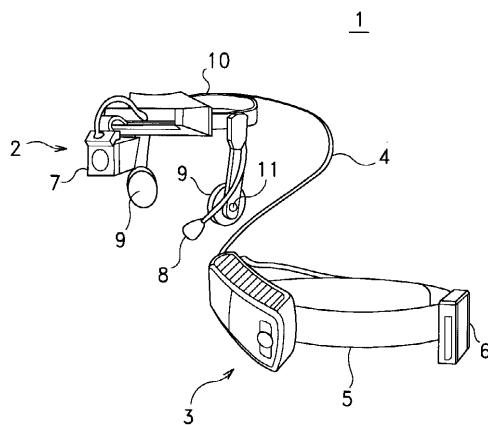
【図2】本発明を適用した補聴器の構成を示すブロック図である。

【図3】本発明を適用した補聴器でAR空間を作るための構成について示すブロック図である。

【符号の説明】

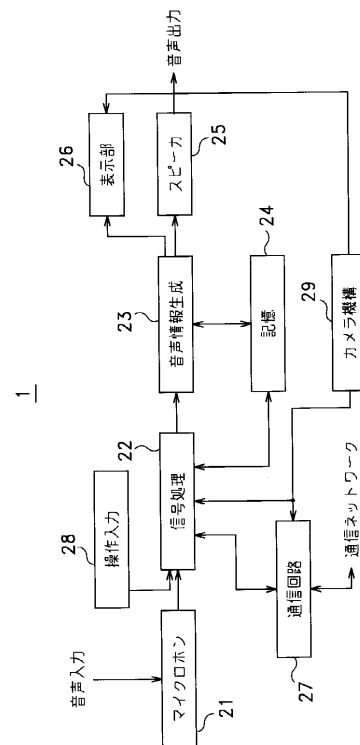
1 音声生成装置、2 ヘッドマウントディスプレイ、3 コンピュータ部、7 ディスプレイ部、8 使用者用マイクロホン、11 外部用マイクロホン、21 マイクロホン、23 音声情報生成部、24 記憶部、25 スピーカ部、26 表示部

【図1】



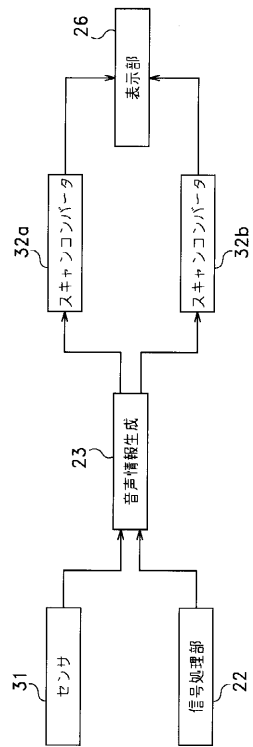
本発明を適用した補聴器

【図2】



本発明を適用した補聴器

【図 3】



AR空間を作る補聴器

## フロントページの続き

(74)代理人 100091063

弁理士 田中 英夫

(72)発明者 大場 俊彦

東京都杉並区善福寺2丁目8番3号

審査官 石丸 昌平

- (56)参考文献 実開昭61-114472(JP,U)  
特公平06-042760(JP,B2)  
特公平07-053168(JP,B2)  
特表平10-508442(JP,A)  
特開平02-097200(JP,A)  
特開平03-035296(JP,A)  
特開平04-156033(JP,A)  
特開平04-249990(JP,A)  
特開平05-056499(JP,A)  
特開平05-289608(JP,A)  
特開平07-013582(JP,A)  
特開平07-084592(JP,A)  
特開平07-181888(JP,A)  
特開平07-191599(JP,A)  
特開平07-327213(JP,A)  
特開平08-065647(JP,A)  
特開平08-079897(JP,A)  
特開平08-212228(JP,A)  
特開平09-292971(JP,A)  
特開平10-108152(JP,A)  
特開平10-228367(JP,A)  
特開昭51-055604(JP,A)  
特開昭60-143100(JP,A)  
特開昭62-224349(JP,A)  
特開昭62-231981(JP,A)  
特開昭63-249560(JP,A)

(58)調査した分野(Int.Cl.<sup>7</sup>, DB名)

G10L 15/00-15/28

G09B 21/00

H04R 25/00

H04N 7/14