

(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2017-173530  
(P2017-173530A)

(43) 公開日 平成29年9月28日 (2017.9.28)

(51) Int.Cl.		F I		テーマコード (参考)
<b>G 1 0 L</b> 15/22 (2006.01)		G 1 0 L	15/22 4 6 0 Z	
<b>G 0 6 F</b> 3/16 (2006.01)		G 0 6 F	3/16 6 2 0	
		G 0 6 F	3/16 6 3 0	
		G 0 6 F	3/16 6 7 0	

審査請求 未請求 請求項の数 9 O L (全 18 頁)

(21) 出願番号 特願2016-58958 (P2016-58958)  
(22) 出願日 平成28年3月23日 (2016.3.23)

(71) 出願人 00005223  
富士通株式会社  
神奈川県川崎市中原区上小田中4丁目1番1号  
(74) 代理人 110002147  
特許業務法人酒井国際特許事務所  
(72) 発明者 古賀 奨  
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

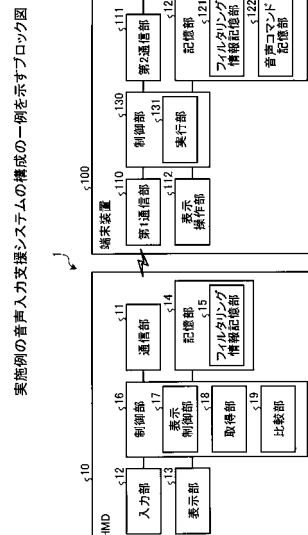
(54) 【発明の名称】 音声入力支援プログラム、ヘッドマウントディスプレイ、音声入力支援方法および音声入力支援装置

(57) 【要約】

【課題】 音声の認識精度を高めることができる音声入力支援プログラム、ヘッドマウントディスプレイ、音声入力支援方法および音声入力支援装置を提供する。

【解決手段】 音声入力支援プログラムは、マイクにより收音された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得する処理をコンピュータに実行させる。また、音声入力支援プログラムは、取得した音情報と、取得した音声パターンとを比較する処理をコンピュータに実行させる。また、音声入力支援プログラムは、比較結果を出力する処理をコンピュータに実行させる。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

マイクにより収録された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得し、

取得した前記音情報と、取得した前記音声パターンとを比較し、

比較結果を出力する、

処理をコンピュータに実行させることを特徴とする音声入力支援プログラム。

**【請求項 2】**

マイクにより収録された音情報を取得すると、複数の音声パターンをアプリの種別に対応付けて記憶する記憶部を参照して、端末の画面に表示されたアプリの種別に対応付けられた音声パターンを取得し、

取得した前記音情報と、取得した前記音声パターンとを比較し、

比較結果を出力する、

処理をコンピュータに実行させることを特徴とする音声入力支援プログラム。

**【請求項 3】**

さらに、複数の音声コマンドを音声パターンに対応付けて記憶する記憶部を参照して、前記複数の音声コマンドのうち、取得した前記音声パターンに対応付けられた音声コマンドを前記端末の画面に表示する、

ことを特徴とする請求項 1 または 2 に記載の音声入力支援プログラム。

**【請求項 4】**

前記複数の音声パターンおよび画像情報、または、前記複数の音声パターンおよびアプリの種別は、前記端末から取得して前記記憶部に記憶する、

ことを特徴とする請求項 1 または 2 に記載の音声入力支援プログラム。

**【請求項 5】**

マイクと、

ディスプレイと、

前記ディスプレイに表示される画像情報それぞれに対応付けて音声パターンを記憶する記憶部と、

前記マイクにより収録された音情報を取得すると、前記記憶部を参照して、前記ディスプレイに表示された画像情報に対応付けられた音声パターンを取得し、取得した前記音情報と、取得した前記音声パターンとの比較結果を出力する制御部と、

を備えるヘッドマウントディスプレイ。

**【請求項 6】**

マイクにより収録された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得し、

取得した前記音情報と、取得した前記音声パターンとを比較し、

比較結果を出力する、

処理をコンピュータが実行することを特徴とする音声入力支援方法。

**【請求項 7】**

マイクにより収録された音情報を取得すると、複数の音声パターンをアプリの種別に対応付けて記憶する記憶部を参照して、端末の画面に表示されたアプリの種別に対応付けられた音声パターンを取得し、

取得した前記音情報と、取得した前記音声パターンとを比較し、

比較結果を出力する、

処理をコンピュータが実行することを特徴とする音声入力支援方法。

**【請求項 8】**

マイクにより収録された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声

10

20

30

40

50

パターンを取得する取得部と、

取得された前記音情報と、取得された前記音声パターンとを比較する比較部と、  
比較結果を出力する出力制御部と、  
を有することを特徴とする音声入力支援装置。

【請求項 9】

マイクにより收音された音情報を取得すると、複数の音声パターンをアプリの種別に対応付けて記憶する記憶部を参照して、端末の画面に表示されたアプリの種別に対応付けられた音声パターンを取得する取得部と、

取得された前記音情報と、取得された前記音声パターンとを比較する比較部と、  
比較結果を出力する出力制御部と、  
を有することを特徴とする音声入力支援装置。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声入力支援プログラム、ヘッドマウントディスプレイ、音声入力支援方法および音声入力支援装置に関する。

【背景技術】

【0002】

近年、ヘッドマウントディスプレイ等の表示装置を用いて、撮像された画像上にオブジェクトを重畳表示させる A R (Augmented Reality) 技術が提案されている。ヘッドマウントディスプレイを用いる場合には、入力手段として音声認識によるコマンド入力提案されている。また、動画データを管理するために、動画の代表画像データと、音声認識されたキーワードと、動画データとを対応付けて記憶してインデックスを管理することが提案されている。

20

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開平 8 - 2 1 2 3 2 8 号公報

【特許文献 2】特開 2 0 1 0 - 0 3 4 8 9 3 号公報

【特許文献 3】特開 2 0 0 6 - 3 0 1 7 5 7 号公報

30

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、ヘッドマウントディスプレイで音声認識を行うためには、音声認識エンジンを常に稼働させているため、意図しない音声を検出してしまふ場合がある。このため、使用中のアプリケーションにおいて、意図しないコマンド入力が行われたり、対応するコマンドがないのに音声認識が行われたりといった、誤認識が多発する場合がある。

【0005】

一つの側面では、本発明は、音声の認識精度を高めることができる音声入力支援プログラム、ヘッドマウントディスプレイ、音声入力支援方法および音声入力支援装置を提供することにある。

40

【課題を解決するための手段】

【0006】

一つの態様では、音声入力支援プログラムは、マイクにより收音された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得する処理をコンピュータに実行させる。また、音声入力支援プログラムは、取得した前記音情報と、取得した前記音声パターンとを比較する処理をコンピュータに実行させる。また、音声入力支援プログラムは、比較結果を出力する処理をコンピュータに実行させる。

【発明の効果】

50

## 【 0 0 0 7 】

音声の認識精度を高めることができる。

## 【 図面の簡単な説明 】

## 【 0 0 0 8 】

【 図 1 】 図 1 は、実施例の音声入力支援システムの構成の一例を示すブロック図である。

【 図 2 】 図 2 は、フィルタリング情報の通知の一例を示す図である。

【 図 3 】 図 3 は、フィルタリング情報記憶部の一例を示す図である。

【 図 4 】 図 4 は、対応する音声コマンドがある場合の一例を示す図である。

【 図 5 】 図 5 は、対応する音声コマンドがない場合の一例を示す図である。

【 図 6 】 図 6 は、音声コマンド記憶部の一例を示す図である。

【 図 7 】 図 7 は、実施例の音声入力処理の一例を示すシーケンス図である。

【 図 8 】 図 8 は、実施例の音声入力処理の一例を示すシーケンス図である。

【 図 9 】 図 9 は、音声入力支援プログラムを実行するコンピュータの一例を示す図である

10

## 【 発明を実施するための形態 】

## 【 0 0 0 9 】

以下、図面に基づいて、本願の開示する音声入力支援プログラム、ヘッドマウントディスプレイ、音声入力支援方法および音声入力支援装置の実施例を詳細に説明する。なお、本実施例により、開示技術が限定されるものではない。また、以下の実施例は、矛盾しない範囲で適宜組みあわせてもよい。

20

## 【 実施例 】

## 【 0 0 1 0 】

図 1 は、実施例の音声入力支援システムの構成の一例を示すブロック図である。図 1 に示す音声入力支援システム 1 は、ヘッドマウントディスプレイ（以下、HMD（Head Mounted Display）ともいう。）10と、端末装置100と、サーバ200とを有する。HMD10と端末装置100とは、1対1で無線または有線により接続される。すなわち、HMD10は、端末装置100の表示部の一例として機能する。なお、HMD10と端末装置100とは、1対1での接続に限定されず、1対多、多対1または多対多での接続であってもよい。また、図1では、HMD10と端末装置100との組について、1組を一例として示したが、HMD10と端末装置100との組の数は限定されず、任意の数のHMD10と端末装置100との組を有してもよい。また、HMD10および端末装置100は、音声入力支援装置の一例である。

30

## 【 0 0 1 1 】

HMD10と端末装置100との間は、例えば、Wi-Fi Direct（登録商標）等の無線LAN（Local Area Network）により相互に通信可能に接続される。また、端末装置100とサーバ200との間は、ネットワークNを介して相互に通信可能に接続される。かかるネットワークNには、有線または無線を問わず、インターネットを始め、LANやVPN（Virtual Private Network）などの任意の種類通信網を採用できる

## 【 0 0 1 2 】

HMD10は、ユーザが端末装置100とともに装着し、端末装置100から送信される表示画面を表示する。HMD10は、例えば、単眼の透過型のHMDを用いることができる。なお、HMD10は、例えば、両眼用や没入型等の各種のHMDを用いるようにしてもよい。また、HMD10は、ユーザによる音声入力を受け付けるために、入力部の一例としてマイクを有する。

40

## 【 0 0 1 3 】

HMD10は、マイクにより收音された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得する。HMD10は、取得した音情報と、取得した音声パターンとを比較し、比較結果を出力する。HMD10は、出力された比較結果が、音情報

50

と音声パターンとが一致する場合に、音声コマンドID (Identifier) を端末装置100に送信する。これにより、HMD10は、音声の認識精度を高めることができる。

【0014】

端末装置100は、ユーザが装着して操作する情報処理装置であり、例えば、タブレット端末やスマートフォン等の移動体通信端末等を用いることができる。端末装置100は、例えば、HMD10と連携して動作するARミドルウェア(以下、ARミドルともいう。)と、Webアプリケーション(以下、Webアプリともいう。)とを実行する。ARミドルは、Webアプリに対してARコンテンツの表示、表示画面における画面遷移、操作メニュー等の基本的な機能を提供する。Webアプリは、例えば、ユーザに対して設備の点検等に関連する操作画面等を提供する。なお、以下の説明では、ARミドルとWeb 10

【0015】

サーバ200は、例えば、ある工場における設備点検用のARコンテンツを管理するデータベースや、Webアプリの各画面におけるフィルタリング情報を記憶するデータベースを有する。なお、フィルタリング情報は、音声コマンドIDを画面に対応付けた情報、つまり、複数の音声パターンを画像情報に対応付けた情報である。サーバ200は、端末装置100の求めに応じて、ネットワークNを介して、ARコンテンツを端末装置100 20

【0016】

ここで、本願の音声認識による音声コマンドの入力と、従来の音声認識による音声コマンドの入力とを比較する。従来の音声認識による音声コマンドの入力では、音声認識の結果に処理が対応付けられていない場合であっても音声認識を行い、例えば認識音を鳴らすことでユーザに音声認識を行ったことを通知している。ところが、実際には、認識結果に対応する音声コマンドがないため、何も処理が行われず、ユーザは、音声認識結果や音声認識後の処理結果を把握できないこととなる。これに対し、本願の音声認識による音声コマンドの入力では、フィルタリング情報を用いて、音声認識の結果に処理が対応付けられていない場合には、フィルタリングによって、例えば認識音を鳴らさない。このため、本願の音声認識による音声コマンドの入力では、ユーザは音声入力した音声コマンドが、現 30

【0017】

本願におけるフィルタリング情報の通知について説明する。図2は、フィルタリング情報の通知の一例を示す図である。なお、図2では、HMD10に表示される表示画面のイメージをWebアプリ100bに模式的に表示しているが、実際にはHMD10の表示素子に表示される。図2の例では、Webアプリ100bからARミドル100aに対して、Webアプリ100bで使用する音声コマンドのリスト、つまりフィルタリング情報を通知する(ステップS1)。次に、ARミドル100aは、HMD10に対してARミドル100aおよびWebアプリ100bで使用するフィルタリング情報を送信する(ステップS2)。また、ARミドル100aは、HMD10に対して表示中の画面の画面ID 40

【0018】

HMD10は、ユーザから入力された音情報に対して音声コマンドの認識を行って、フィルタリング情報に含まれる音声パターンと比較する。HMD10は、比較の結果、フィルタリング情報に含まれる音声パターンと一致する場合に、一致した音声コマンドIDをARミドル100aに送信する(ステップS3)。

【0019】

ARミドル100aは、受信した音声コマンドIDに対応する音声コマンドの処理を実行する(ステップS4)。また、ARミドル100aは、受信した音声コマンドIDが、 50

Webアプリ100bで処理を実行する音声コマンドの音声コマンドIDである場合には、当該音声コマンドIDまたは対応する音声コマンドをWebアプリ100bに出力する(ステップS5)。また、ARミドル100aは、Webアプリ100bにおいて画面遷移が発生すると、遷移した画面の画面IDをHMD10に送信する(ステップS6)。HMD10は、画面IDを受信すると、画面IDに対応するフィルタリング情報に基づいて、音声認識におけるフィルタリングを開始する。

#### 【0020】

次に、HMD10の構成について説明する。図1に示すように、HMD10は、通信部11と、入力部12と、表示部13と、記憶部14と、制御部16とを有する。なお、HMD10は、図1に示す機能部以外にも、例えば各種の入力デバイスや音声出力デバイス等の機能部を有することとしてもかまわない。

10

#### 【0021】

通信部11は、例えば、無線LAN等の通信モジュール等によって実現される。通信部11は、例えば、Wi-Fi Direct(登録商標)により端末装置100と無線で接続され、端末装置100との間で情報の通信を司る通信インタフェースである。通信部11は、端末装置100から、フィルタリング情報、終了情報、表示画面および画面IDを受信する。通信部11は、受信したフィルタリング情報、終了情報、表示画面および画面IDを制御部16に出力する。また、通信部11は、制御部16から入力された音声コマンドIDを端末装置100に送信する。

#### 【0022】

入力部12は、例えば、マイクロフォン(以下、マイクともいう。)であり、ユーザが発声する音声を收音する。入力部12は、マイクとして、例えば、エレクトレットコンデンサマイク等の各種のマイクを用いることができる。入力部12は、收音した音声である音情報を制御部16に出力する。

20

#### 【0023】

表示部13は、各種情報を表示するための表示デバイスである。表示部13は、例えば、ハーフミラーに映像が投影され、ユーザが映像とともに外部の風景を透過して見ることができる透過型のHMDの表示素子に相当する。なお、表示部13は、没入型、ビデオ透過型、網膜投影型等のHMDに対応する表示素子でもよい。

#### 【0024】

記憶部14は、例えば、RAM(Random Access Memory)、フラッシュメモリ(Flash Memory)等の半導体メモリ素子等の記憶装置によって実現される。記憶部14は、フィルタリング情報記憶部15を有する。また、記憶部14は、制御部16での処理に用いる情報を記憶する。

30

#### 【0025】

フィルタリング情報記憶部15は、端末装置100から受信したフィルタリング情報を記憶する。なお、フィルタリング情報記憶部15は、音声コマンドの辞書の一例である。図3は、フィルタリング情報記憶部の一例を示す図である。図3に示すように、フィルタリング情報記憶部15は、画面ID管理テーブル15aと、音声コマンドID管理テーブル15bとを有する。画面ID管理テーブル15aは、画面IDとフィルタリングIDとを対応付けて記憶する。つまり、画面ID管理テーブル15aは、「画面ID」、「フィルタリングID」といった項目を有する。

40

#### 【0026】

「画面ID」は、HMD10に表示する画面を識別する識別子である。「フィルタリングID」は、表示される画面における音声コマンドのセットを識別する識別子である。なお、画面ID管理テーブル15aは、例えば、「画面ID」に代えてWebアプリ100bの種別を識別する「アプリID」を用いてもよい。この場合、「フィルタリングID」は、Webアプリ100bにおける音声コマンドのセットを識別する識別子である。

#### 【0027】

音声コマンドID管理テーブル15bは、フィルタリングIDと音声コマンドIDとを

50

対応付けて記憶する。つまり、音声コマンドID管理テーブル15bは、「フィルタリングID」、「音声コマンドID」といった項目を有する。

【0028】

「フィルタリングID」は、表示される画面における音声コマンドのセットを識別する識別子である。「音声コマンドID」は、音声コマンドを識別する識別子である。また、「音声コマンドID」には、図示しない音声パターンが対応付けて記憶されている。

【0029】

図1の説明に戻って、制御部16は、例えば、CPU (Central Processing Unit) やMPU (Micro Processing Unit) 等によって、内部の記憶装置に記憶されているプログラムがRAMを作業領域として実行されることにより実現される。また、制御部16は、例えば、ASIC (Application Specific Integrated Circuit) やFPGA (Field Programmable Gate Array) 等の集積回路により実現されるようにしてもよい。制御部16は、表示制御部17と、取得部18と、比較部19とを有し、以下に説明する情報処理の機能や作用を実現または実行する。なお、制御部16の内部構成は、図1に示した構成に限られず、後述する情報処理を行う構成であれば他の構成であってもよい。

10

【0030】

表示制御部17は、例えば、ユーザによって電源が投入され、表示画面の受信を開始すると、音声認識エンジンを起動する起動指示を取得部18に出力する。また、表示制御部17は、通信部11を介して端末装置100からフィルタリング情報、表示画面および画面IDを受信する。表示制御部17は、受信したフィルタリング情報をフィルタリング情報記憶部15に記憶する。また、表示制御部17は、通信部11を介して端末装置100から画面IDが対応付けられた表示画面を受信すると、画面IDを取得部18に出力するとともに、表示画面を表示部13に表示させる。

20

【0031】

さらに、表示制御部17は、画面IDが対応付けられた表示画面について、画面遷移が発生すると、遷移後の表示画面および画面IDも同様に、画面IDを取得部18に出力するとともに、表示画面を表示部13に表示させる。なお、表示制御部17は、例えば、Webアプリ100bが起動されていない状態の表示画面等の画面IDが対応付けられていない表示画面を受信すると、受信した表示画面を表示部13に表示させる。

【0032】

表示制御部17は、画面IDが対応付けられた表示画面の表示中に、表示画面中に認識した音声コマンドが含まれる表示画面に更新されると、更新された表示画面を表示部13に表示させる。すなわち、表示制御部17は、複数の音声コマンドのうち、取得した音声パターンに対応付けられた音声コマンドを表示画面に表示する。また、表示制御部17は、通信部11を介して端末装置100から終了情報を受信したか否かを判定する。表示制御部17は、終了情報を受信していない場合には、音情報の取得を待機する。表示制御部17は、終了情報を受信した場合には、取得部18に対して終了指示を出力する。

30

【0033】

取得部18は、表示制御部17から起動指示が入力されると、音声認識エンジンを起動して、入力部12により收音された音情報の取得を開始する。取得部18は、音声認識エンジンを用いて、取得した音情報をフィルタリング情報記憶部15に記憶された音声パターンと比較可能な音情報に変換する。すなわち、取得部18は、音声コマンドを認識する。取得部18は、表示制御部17から画面IDが入力されると、フィルタリング情報記憶部15を参照して、画面IDに対応付けられた、1つ以上の音声コマンドIDおよび音声パターンを取得する。取得部18は、変換後の音情報、音声コマンドIDおよび音声パターンを比較部19に出力する。すなわち、取得部18は、フィルタリング情報を用いて取得された音情報のフィルタリングを開始する。また、取得部18は、表示制御部17から終了指示が入力されると、音声認識エンジンを停止する。

40

【0034】

比較部19は、取得部18から変換後の音情報、音声コマンドIDおよび音声パターン

50

が入力されると、変換後の音情報と、音声パターンとを比較する。比較部 19 は、変換後の音情報が、1 つ以上の音声パターンのうちいずれかと一致する場合には、一致した音声パターンに対応する音声コマンド ID を含む、一致した旨の比較結果を生成する。比較部 19 は、変換後の音情報が、1 つ以上の音声パターンのうちいずれとも一致しない場合には、一致しない旨の比較結果を生成する。比較部 19 は、生成した比較結果を出力する。すなわち、比較部 19 は、出力制御部でもあり、通信部 11 を介して、端末装置 100 に生成した比較結果を送信する。

#### 【0035】

言い換えると、比較部 19 は、変換後の音情報がフィルタリング情報と一致するか否かを判定する。比較部 19 は、変換後の音情報がフィルタリング情報と一致する場合には、一致した音声パターンに対応する音声コマンド ID を含む、一致した旨の比較結果を生成し、生成した比較結果を端末装置 100 に送信する。比較部 19 は、変換後の音情報がフィルタリング情報と一致しない場合には、一致しない旨の比較結果を生成し、生成した比較結果を端末装置 100 に送信する。

10

#### 【0036】

また、比較部 19 は、生成した比較結果が一致した旨の比較結果である場合には、例えば、図示しないイヤホン等に対して認識音を出力する。さらに、比較部 19 は、生成した比較結果が一致しない旨の比較結果である場合に、例えば、図示しないイヤホン等に対して「認識できません」といった音声を出力する。なお、比較部 19 は、生成した比較結果が一致しない旨の比較結果である場合に、認識音や音声を出力しないようにしてもよい。

20

#### 【0037】

ここで、図 4 および図 5 を用いてフィルタリング情報の音声パターンに対応する音声コマンドの有無における表示画面例について説明する。図 4 は、対応する音声コマンドがある場合の一例を示す図である。なお、図 4 および図 5 では、説明のために、HMD 10 の表示素子に表示される表示画面を模式的に端末装置 100 内に記載している。

#### 【0038】

図 4 に示すように、ユーザ 5 が「メニュー」と発声すると、HMD 10 は、「メニュー」の音情報がフィルタリング情報と一致するか否かを判定する。図 4 の例では、「メニュー」の音情報がフィルタリング情報と一致するので、HMD 10 は、認識結果、つまり一致した旨の比較結果を端末装置 100 に送信する。端末装置 100 は、比較結果に含まれる音声コマンド ID に基づいて、メニュー画面 21 を HMD 10 に送信して表示させる。

30

#### 【0039】

図 5 は、対応する音声コマンドがない場合の一例を示す図である。図 5 に示すように、ユーザ 5 が「番号 1」と発声すると、HMD 10 は、「番号 1」の音情報がフィルタリング情報と一致するか否かを判定する。図 5 の例では、「番号 1」の音情報がフィルタリング情報と一致しないので、HMD 10 は、認識結果、つまり一致しない旨の比較結果を端末装置 100 に送信する。端末装置 100 は、一致しない旨の比較結果に基づいて、エラー画面 22 を HMD 10 に送信して表示させる。

#### 【0040】

続いて、端末装置 100 の構成について説明する。図 1 に示すように、端末装置 100 は、第 1 通信部 110 と、第 2 通信部 111 と、表示操作部 112 と、記憶部 120 と、制御部 130 とを有する。なお、端末装置 100 は、図 1 に示す機能部以外にも既知のコンピュータが有する各種の機能部、例えば各種の入力デバイスや音声出力デバイス等の機能部を有することとしてもかまわない。

40

#### 【0041】

第 1 通信部 110 は、例えば、無線 LAN 等の通信モジュール等によって実現される。第 1 通信部 110 は、例えば、Wi-Fi Direct (登録商標) により HMD 10 と無線で接続され、HMD 10 との間で情報の通信を司る通信インタフェースである。第 1 通信部 110 は、HMD 10 から比較結果を受信する。第 1 通信部 110 は、受信した比較結果を制御部 130 に出力する。また、第 1 通信部 110 は、制御部 130 から入力

50

されたフィルタリング情報、終了情報、表示画面および画面IDをHMD10に送信する。

【0042】

第2通信部111は、例えば、第3世代移動通信システム、LTE (Long Term Evolution) 等の携帯電話回線および無線LAN等の通信モジュール等によって実現される。

第2通信部111は、ネットワークNを介してサーバ200と無線で接続され、サーバ200との間で情報の通信を司る通信インタフェースである。第2通信部111は、制御部130から入力されたデータ取得指示およびフィルタリング情報取得指示を、ネットワークNを介して、サーバ200に送信する。また、第2通信部111は、データ取得指示に応じたARコンテンツ、および、フィルタリング情報取得指示に応じたフィルタリング情報

10

【0043】

表示操作部112は、各種情報を表示するための表示デバイス、および、ユーザから各種操作を受け付ける入力デバイスである。例えば、表示操作部112は、表示デバイスとして液晶ディスプレイ等によって実現される。また、例えば、表示操作部112は、入力デバイスとして、タッチパネル等によって実現される。つまり、表示操作部112は、表示デバイスと入力デバイスとが一体化される。表示操作部112は、ユーザによって入力された操作を操作情報として制御部130に出力する。なお、表示操作部112は、HMD10と同様の画面を表示してもよいし、HMD10と異なる画面を表示してもよい。

20

【0044】

記憶部120は、例えば、RAM、フラッシュメモリ等の半導体メモリ素子、ハードディスクや光ディスク等の記憶装置によって実現される。記憶部120は、フィルタリング情報記憶部121と、音声コマンド記憶部122とを有する。また、記憶部120は、制御部130での処理に用いる情報を記憶する。

【0045】

フィルタリング情報記憶部121は、サーバ200から取得したフィルタリング情報を記憶する。なお、フィルタリング情報記憶部121の構成については、HMD10のフィルタリング情報記憶部15と同様であるので、その説明を省略する。

【0046】

音声コマンド記憶部122は、音声コマンドIDと音声コマンドとを対応付けて記憶する。図6は、音声コマンド記憶部の一例を示す図である。図6に示すように、音声コマンド記憶部122は、「音声コマンドID」、「音声コマンド」といった項目を有する。音声コマンド記憶部122は、例えば音声コマンドIDごとに1レコードとして記憶する。

30

【0047】

「音声コマンドID」は、音声コマンドを識別する識別子である。「音声コマンド」は、例えば「メニュー表示」、「番号1選択」といったコマンドを示す情報である。

【0048】

図1の説明に戻って、制御部130は、例えば、CPUやMPU等によって、内部の記憶装置に記憶されているプログラムがRAMを作業領域として実行されることにより実現される。また、制御部130は、例えば、ASICやFPGA等の集積回路により実現されるようにしてもよい。制御部130は、実行部131を有し、以下に説明する情報処理の機能や作用を実現または実行する。なお、制御部130の内部構成は、図1に示した構成に限られず、後述する情報処理を行う構成であれば他の構成であってもよい。

40

【0049】

実行部131は、ARアプリ、すなわちARミドル100aおよびWebアプリ100bを実行する。実行部131は、例えば、端末装置100の電源が投入されると、HMD10に対する表示画面の送信を開始する。ARミドル100aは、例えば、ユーザによって表示操作部112から入力された操作情報に基づいて、Webアプリ100bの起動を指示する。ARミドル100aは、Webアプリ100bからフィルタリング情報が入力

50

されると、入力されたフィルタリング情報を、第1通信部110を介してHMD10に送信する。また、ARミドル100aは、表示画面および画面IDを、第1通信部110を介してHMD10に送信する。

【0050】

ARミドル100aは、第1通信部110を介して、HMD10から比較結果を受信すると、比較結果に応じた処理を実行する。ARミドル100aは、音声コマンドIDを含む一致する旨の比較結果を受信した場合には、音声コマンド記憶部122を参照し、当該音声コマンドIDに対応する音声コマンドをARミドル100aで処理するか否かを判定する。ARミドル100aは、ARミドル100aで処理する場合には、音声コマンドに対応する処理を実行する。

10

【0051】

ARミドル100aは、ARミドル100aで処理しない場合には、音声コマンドをWebアプリ100bに出力する。なお、ARミドル100aは、一致しない旨の比較結果を受信した場合には、音声認識できない旨のメッセージを表示画面に表示させてもよいし、特になにも処理をしなくてもよい。

【0052】

ARミドル100aは、音声コマンドに対応する処理について画面遷移があるか否かを判定する。ARミドル100aは、画面遷移がある場合には、遷移後の表示画面の画面IDを、第1通信部110を介してHMD10に送信する。ARミドル100aは、画面遷移がない場合には、Webアプリ100bが終了したか否かを判定する。

20

【0053】

ARミドル100aは、Webアプリ100bが終了していない場合には、HMD10から比較結果を受信できるように待機する。ARミドル100aは、Webアプリ100bが終了した場合には、第1通信部110を介して、終了情報をHMD10に送信する。

【0054】

Webアプリ100bは、ARミドル100aからの起動指示により起動する。Webアプリ100bは、起動すると、第2通信部111およびネットワークNを介して、データ取得指示およびフィルタリング情報取得指示をサーバ200に送信する。Webアプリ100bは、第2通信部111およびネットワークNを介して、サーバ200からデータ取得指示に応じたARコンテンツ、および、フィルタリング情報取得指示に応じたフィルタリング情報を取得する。

30

【0055】

Webアプリ100bは、ARミドル100aと協働してARコンテンツを含む表示画面を生成し、生成した表示画面を、第1通信部110を介してHMD10に送信して表示させる。また、Webアプリ100bは、取得したフィルタリング情報をARミドル100aに出力する。Webアプリ100bは、ARミドル100aから音声コマンドが入力されると、音声コマンドに対応する処理を実行する。

【0056】

次に、実施例の音声入力支援システム1の動作について説明する。図7および図8は、実施例の音声入力処理の一例を示すシーケンス図である。

40

【0057】

HMD10の表示制御部17は、例えば、ユーザによって電源が投入され、表示画面の受信を開始すると、音声認識エンジンを起動する起動指示を取得部18に出力する。取得部18は、表示制御部17から起動指示が入力されると、音声認識エンジンを起動して、入力部12により収録された音情報の取得を開始する(ステップS11)。

【0058】

端末装置100の実行部131は、例えば、端末装置100の電源が投入されると、HMD10に対する表示画面の送信を開始する。実行部131で実行されるARミドル100aは、例えば、ユーザによって表示操作部112から入力された操作情報に基づいて、Webアプリ100bの起動を指示する(ステップS12)。Webアプリ100bは、

50

ARミドル100aからの起動指示により起動する(ステップS13)。Webアプリ100bは、起動すると、データ取得指示およびフィルタリング情報取得指示をサーバ200に送信する。Webアプリ100bは、サーバ200からデータ取得指示に応じたARコンテンツ、および、フィルタリング情報取得指示に応じたフィルタリング情報を取得する(ステップS14)。

【0059】

ARミドル100aは、Webアプリ100bからフィルタリング情報が入力されると、入力されたフィルタリング情報をHMD10に送信する(ステップS15)。HMD10の表示制御部17は、フィルタリング情報を受信すると、受信したフィルタリング情報をフィルタリング情報記憶部15に記憶する(ステップS16)。

10

【0060】

また、端末装置100のARミドル100aは、表示画面および画面IDをHMD10に送信する(ステップS17)。HMD10の表示制御部17は、端末装置100から表示画面および画面IDを受信する(ステップS18)。表示制御部17は、表示画面および画面IDを受信すると、画面IDを取得部18に出力するとともに、表示画面を表示部13に表示させる。取得部18は、フィルタリング情報記憶部15を参照し、フィルタリング情報を用いて取得された音情報のフィルタリングを開始する(ステップS19)。取得部18は、音情報を取得したか否かを判定する(ステップS20)。取得部18は、音情報を取得した場合には(ステップS20:肯定)、音声認識エンジンを用いて、取得した音情報をフィルタリング情報記憶部15に記憶された音声パターンと比較可能な音情報に変換する。すなわち、取得部18は、音声コマンドを認識する(ステップS21)。取得部18は、音情報を取得しない場合には(ステップS20:否定)、ステップS32に進む。

20

【0061】

取得部18は、表示制御部17から画面IDが入力されると、フィルタリング情報記憶部15を参照して、画面IDに対応付けられた、1つ以上の音声コマンドIDおよび音声パターンを取得する。取得部18は、変換後の音情報、音声コマンドIDおよび音声パターンを比較部19に出力する。比較部19は、取得部18から変換後の音情報、音声コマンドIDおよび音声パターンが入力されると、変換後の音情報が音声パターン、つまりフィルタリング情報と一致するか否かを判定する(ステップS22)。

30

【0062】

比較部19は、変換後の音情報がフィルタリング情報と一致する場合には(ステップS22:肯定)、一致した音声パターンに対応する音声コマンドIDを含む一致した旨の比較結果を端末装置100に送信する(ステップS23)。比較部19は、変換後の音情報がフィルタリング情報と一致しない場合には(ステップS22:否定)、一致しない旨の比較結果を端末装置100に送信してステップS32に進む。

【0063】

端末装置100のARミドル100aは、音声コマンドIDを含む一致する旨の比較結果をHMD10から受信する(ステップS24)。ARミドル100aは、音声コマンドIDを含む一致する旨の比較結果を受信すると、音声コマンド記憶部122を参照し、当該音声コマンドIDに対応する音声コマンドをARミドル100aで処理するか否かを判定する(ステップS25)。ARミドル100aは、ARミドル100aで処理する場合には(ステップS25:肯定)、音声コマンドに対応する処理を実行する(ステップS26)。

40

【0064】

ARミドル100aは、ARミドルで処理しない場合には(ステップS25:否定)、音声コマンドをWebアプリ100bに出力する(ステップS27)。Webアプリ100bは、ARミドル100aから音声コマンドが入力されると、音声コマンドに対応する処理を実行する(ステップS28)。

【0065】

50

ARミドル100aは、音声コマンドに対応する処理について画面遷移があるか否かを判定する(ステップS29)。ARミドル100aは、画面遷移がある場合には(ステップS29:肯定)、ステップS17に戻り、遷移後の表示画面の画面IDをHMD10に送信する。ARミドル100aは、画面遷移がない場合には(ステップS29:否定)、Webアプリ100bが終了したか否かを判定する(ステップS30)。

【0066】

ARミドル100aは、Webアプリ100bが終了していない場合には(ステップS30:否定)、ステップS24に戻り、HMD10から比較結果を受信できるように待機する。ARミドル100aは、Webアプリ100bが終了した場合には(ステップS30:肯定)、終了情報をHMD10に送信する(ステップS31)。

10

【0067】

HMD10の表示制御部17は、端末装置100から終了情報を受信したか否かを判定する(ステップS32)。表示制御部17は、終了情報を受信していない場合には(ステップS32:否定)、ステップS20に戻る。表示制御部17は、終了情報を受信した場合には(ステップS32:肯定)、取得部18に対して終了指示を出力する。取得部18は、表示制御部17から終了指示が入力されると、音声認識エンジンを停止し、音声入力処理を終了する。これにより、HMD10および端末装置100は、音声の認識精度を高めることができる。

【0068】

なお、上記実施例では、フィルタリング情報記憶部15において、画面IDとフィルタリングIDとを対応付けた画面ID管理テーブル15aを用いたが、これに限定されない。例えば、「画面ID」に代えてWebアプリ100bの種別を識別する「アプリID」を用いるアプリID管理テーブルを用いてもよい。

20

【0069】

このように、HMD10は、マイクにより收音された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部14を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得する。また、HMD10は、取得した音情報と、取得した音声パターンとを比較し、比較結果を出力する。その結果、音声の認識精度を高めることができる。

【0070】

また、HMD10は、マイクにより收音された音情報を取得すると、複数の音声パターンをアプリの種別に対応付けて記憶する記憶部14を参照して、端末の画面に表示されたアプリの種別に対応付けられた音声パターンを取得する。また、HMD10は、取得した音情報と、取得した音声パターンとを比較し、比較結果を出力する。その結果、音声の認識精度を高めることができる。

30

【0071】

また、HMD10および端末装置100は、さらに、複数の音声コマンドを音声パターンに対応付けて記憶する記憶部120を参照して、複数の音声コマンドのうち、取得した音声パターンに対応付けられた音声コマンドを端末の画面に表示する。その結果、ユーザが入力された音声コマンドを確認できる。

40

【0072】

また、HMD10は、複数の音声パターンおよび画像情報、または、複数の音声パターンおよびアプリの種別は、端末装置100から取得して記憶部14に記憶する。その結果、画像情報またはアプリの種別に応じて音声認識の結果をフィルタリングできる。

【0073】

また、HMD10は、マイクと、ディスプレイと、ディスプレイに表示される画像情報それぞれに対応付けて音声パターンを記憶する記憶部14とを備える。また、HMD10は、マイクにより收音された音情報を取得すると、記憶部14を参照して、ディスプレイに表示された画像情報に対応付けられた音声パターンを取得し、取得した音情報と、取得した音声パターンとの比較結果を出力する制御部を備える。その結果、音声の認識精度を

50

高めることができる。

【0074】

なお、上記実施例では、端末装置100とHMD10とをユーザが装着する形態として説明したが、これに限定されない。例えば、HMD10を用いずに、例えばスマートフォンである端末装置100で音声認識を行うようにしてもよい。

【0075】

また、図示した各部の各構成要素は、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、各部の分散・統合の具体的形態は図示のものに限られず、その全部または一部を、各種の負荷や使用状況等に応じて、任意の単位で機能的または物理的に分散・統合して構成することができる。例えば、取得部18と比較部19とを統合してもよい。また、図示した各処理は、上記の順番に限定されるものではなく、処理内容を矛盾させない範囲において、同時に実施してもよく、順序を入れ替えて実施してもよい。

10

【0076】

さらに、各装置で行われる各種処理機能は、CPU（又はMPU、MCU（Micro Controller Unit）等のマイクロ・コンピュータ）上で、その全部または任意の一部を実行するようにしてもよい。また、各種処理機能は、CPU（またはMPU、MCU等のマイクロ・コンピュータ）で解析実行されるプログラム上、またはワイヤードロジックによるハードウェア上で、その全部又は任意の一部を実行するようにしてもよいことは言うまでもない。

【0077】

ところで、上記の実施例で説明した各種の処理は、予め用意されたプログラムをコンピュータで実行することで実現できる。そこで、以下では、上記の実施例と同様の機能を有するプログラムを実行するコンピュータの一例を説明する。図9は、音声入力支援プログラムを実行するコンピュータの一例を示す図である。

20

【0078】

図9に示すように、コンピュータ300は、各種演算処理を実行するCPU301と、データ入力を受け付ける入力装置302と、モニタ303とを有する。また、コンピュータ300は、記憶媒体からプログラム等を読み取る媒体読取装置304と、各種装置と接続するためのインタフェース装置305と、他の情報処理装置等と有線または無線により接続するための通信装置306とを有する。また、コンピュータ300は、各種情報を一時記憶するRAM307と、フラッシュメモリ308とを有する。また、各装置301～308は、バス309に接続される。

30

【0079】

フラッシュメモリ308には、図1に示した表示制御部17、取得部18および比較部19の各処理部と同様の機能を有する音声入力支援プログラムが記憶される。また、フラッシュメモリ308には、フィルタリング情報記憶部15、および、音声入力支援プログラムを実現するための各種データが記憶される。入力装置302は、例えば、コンピュータ300のユーザから音声等の音情報の入力を受け付ける。モニタ303は、例えば、コンピュータ300のユーザに対して表示画面等の各種画面を表示する。インタフェース装置305は、例えばヘッドホン等が接続される。通信装置306は、例えば、図1に示した通信部11と同様の機能を有し端末装置100と接続され、端末装置100と各種情報をやりとりする。

40

【0080】

CPU301は、フラッシュメモリ308に記憶された各プログラムを読み出して、RAM307に展開して実行することで、各種の処理を行う。また、これらのプログラムは、コンピュータ300を図1に示した表示制御部17、取得部18および比較部19として機能させることができる。

【0081】

なお、上記の音声入力支援プログラムは、必ずしもフラッシュメモリ308に記憶されている必要はない。例えば、コンピュータ300が読み取り可能な記憶媒体に記憶された

50

プログラムを、コンピュータ300が読み出して実行するようにしてもよい。コンピュータ300が読み取り可能な記憶媒体は、例えば、CD-ROMやDVDディスク、USB(Universal Serial Bus)メモリ等の可搬型記録媒体、フラッシュメモリ等の半導体メモリ、ハードディスクドライブ等が対応する。また、公衆回線、インターネット、LAN等に接続された装置にこの音声入力支援プログラムを記憶させておき、コンピュータ300がこれらから音声入力支援プログラムを読み出して実行するようにしてもよい。

【0082】

以上、本実施例を含む実施の形態に関し、さらに以下の付記を開示する。

【0083】

(付記1) マイクにより収録された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得し、

取得した前記音情報と、取得した前記音声パターンとを比較し、

比較結果を出力する、

処理をコンピュータに実行させることを特徴とする音声入力支援プログラム。

10

【0084】

(付記2) マイクにより収録された音情報を取得すると、複数の音声パターンをアプリの種別に対応付けて記憶する記憶部を参照して、端末の画面に表示されたアプリの種別に対応付けられた音声パターンを取得し、

取得した前記音情報と、取得した前記音声パターンとを比較し、

比較結果を出力する、

処理をコンピュータに実行させることを特徴とする音声入力支援プログラム。

20

【0085】

(付記3) さらに、複数の音声コマンドを音声パターンに対応付けて記憶する記憶部を参照して、前記複数の音声コマンドのうち、取得した前記音声パターンに対応付けられた音声コマンドを前記端末の画面に表示する、

ことを特徴とする付記1または2に記載の音声入力支援プログラム。

【0086】

(付記4) 前記複数の音声パターンおよび画像情報、または、前記複数の音声パターンおよびアプリの種別は、前記端末から取得して前記記憶部に記憶する、

ことを特徴とする付記1または2に記載の音声入力支援プログラム。

30

【0087】

(付記5) マイクと、

ディスプレイと、

前記ディスプレイに表示される画像情報それぞれに対応付けて音声パターンを記憶する記憶部と、

前記マイクにより収録された音情報を取得すると、前記記憶部を参照して、前記ディスプレイに表示された画像情報に対応付けられた音声パターンを取得し、取得した前記音情報と、取得した前記音声パターンとの比較結果を出力する制御部と、

を備えるヘッドマウントディスプレイ。

40

【0088】

(付記6) マイクにより収録された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得し、

取得した前記音情報と、取得した前記音声パターンとを比較し、

比較結果を出力する、

処理をコンピュータが実行することを特徴とする音声入力支援方法。

【0089】

(付記7) マイクにより収録された音情報を取得すると、複数の音声パターンをアプリの種別に対応付けて記憶する記憶部を参照して、端末の画面に表示されたアプリの種別に対

50

応付けられた音声パターンを取得し、

取得した前記音情報と、取得した前記音声パターンとを比較し、  
比較結果を出力する、

処理をコンピュータが実行することを特徴とする音声入力支援方法。

【0090】

(付記8)さらに、複数の音声コマンドを音声パターンに対応付けて記憶する記憶部を参照して、前記複数の音声コマンドのうち、取得した前記音声パターンに対応付けられた音声コマンドを前記端末の画面に表示する、

ことを特徴とする付記6または7に記載の音声入力支援方法。

【0091】

(付記9)前記複数の音声パターンおよび画像情報、または、前記複数の音声パターンおよびアプリの種別は、前記端末から取得して前記記憶部に記憶する、

ことを特徴とする付記6または7に記載の音声入力支援方法。

【0092】

(付記10)マイクにより收音された音情報を取得すると、複数の音声パターンを画像情報に対応付けて記憶する記憶部を参照して、端末の画面に表示された画像情報に対応付けられた音声パターンを取得する取得部と、

取得された前記音情報と、取得された前記音声パターンとを比較する比較部と、  
比較結果を出力する出力制御部と、

を有することを特徴とする音声入力支援装置。

【0093】

(付記11)マイクにより收音された音情報を取得すると、複数の音声パターンをアプリの種別に対応付けて記憶する記憶部を参照して、端末の画面に表示されたアプリの種別に対応付けられた音声パターンを取得する取得部と、

取得された前記音情報と、取得された前記音声パターンとを比較する比較部と、  
比較結果を出力する出力制御部と、

を有することを特徴とする音声入力支援装置。

【0094】

(付記12)さらに、複数の音声コマンドを音声パターンに対応付けて記憶する記憶部を参照して、前記複数の音声コマンドのうち、取得した前記音声パターンに対応付けられた音声コマンドを前記端末の画面に表示する表示制御部、

を有することを特徴とする付記10または11に記載の音声入力支援装置。

【0095】

(付記13)前記複数の音声パターンおよび画像情報、または、前記複数の音声パターンおよびアプリの種別は、前記端末から取得して前記記憶部に記憶する、

ことを特徴とする付記10または11に記載の音声入力支援装置。

【符号の説明】

【0096】

- 1 音声入力支援システム
- 10 ヘッドマウントディスプレイ
- 11 通信部
- 12 入力部
- 13 表示部
- 14 記憶部
- 15 フィルタリング情報記憶部
- 16 制御部
- 17 表示制御部
- 18 取得部
- 19 比較部
- 100 端末装置

10

20

30

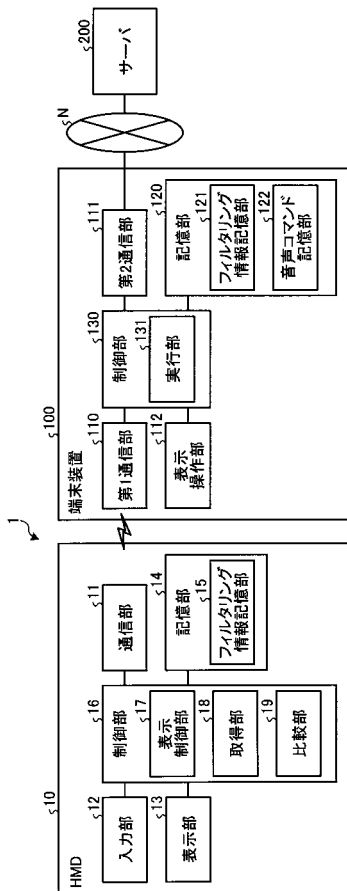
40

50

- 1 1 0 第 1 通 信 部
- 1 1 1 第 2 通 信 部
- 1 1 2 表 示 操 作 部
- 1 2 0 記 憶 部
- 1 2 1 フ ィ ル タ リ ン グ 情 報 記 憶 部
- 1 2 2 音 声 コ マ ン ド 記 憶 部
- 1 3 0 制 御 部
- 1 3 1 実 行 部
- 2 0 0 サ ー バ
- N ネ ッ ト ワ ー ク

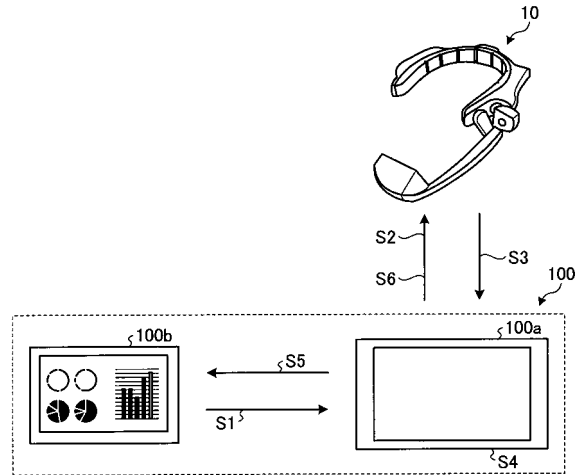
【 図 1 】

実施例の音声入力支援システムの構成の一例を示すブロック図



【 図 2 】

フィルタリング情報の通知の一例を示す図



【 図 3 】

フィルタリング情報記憶部の一例を示す図

15

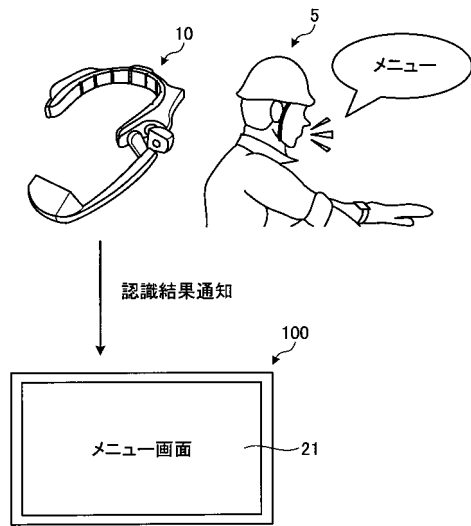
15a	
画面ID	フィルタリングID
1	1
2	2
3	3

15b	
フィルタリングID	音声コマンドID
1	1, 2, 4, 6
2	2, 5, 10, 15, 25
3	3, 8, 10, 17, 20, 21, 22, 25, 30

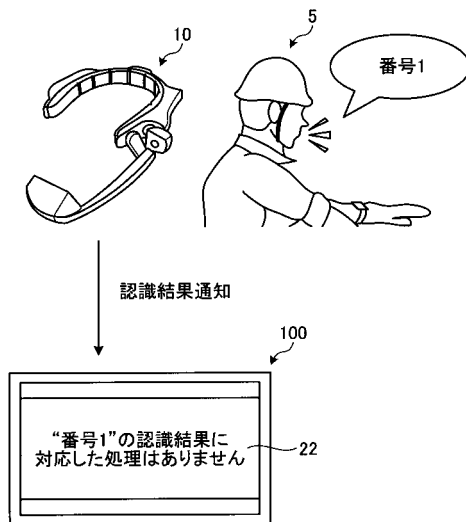
【 図 4 】

対応する音声コマンドがある場合の一例を示す図



【 図 5 】

対応する音声コマンドがない場合の一例を示す図



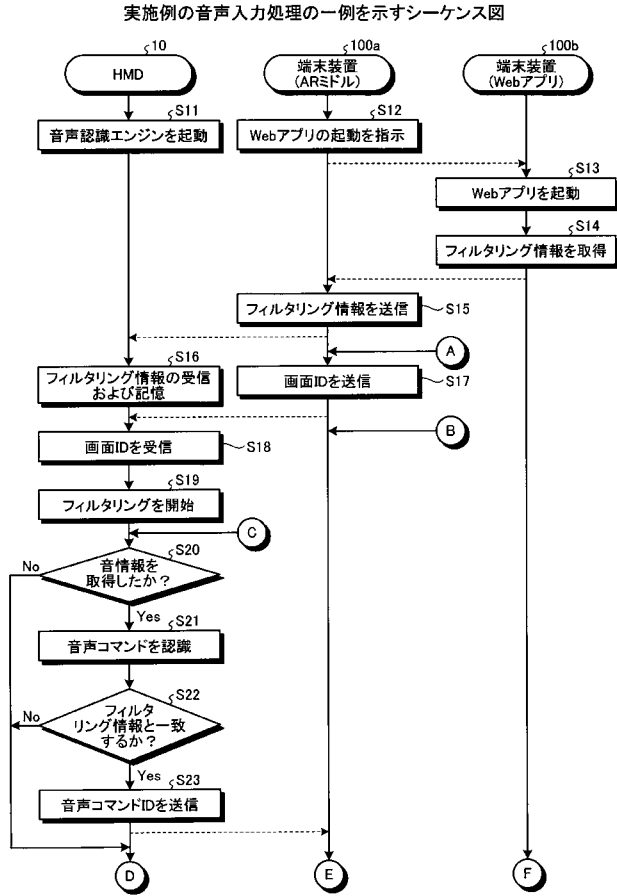
【 図 6 】

音声コマンド記憶部の一例を示す図

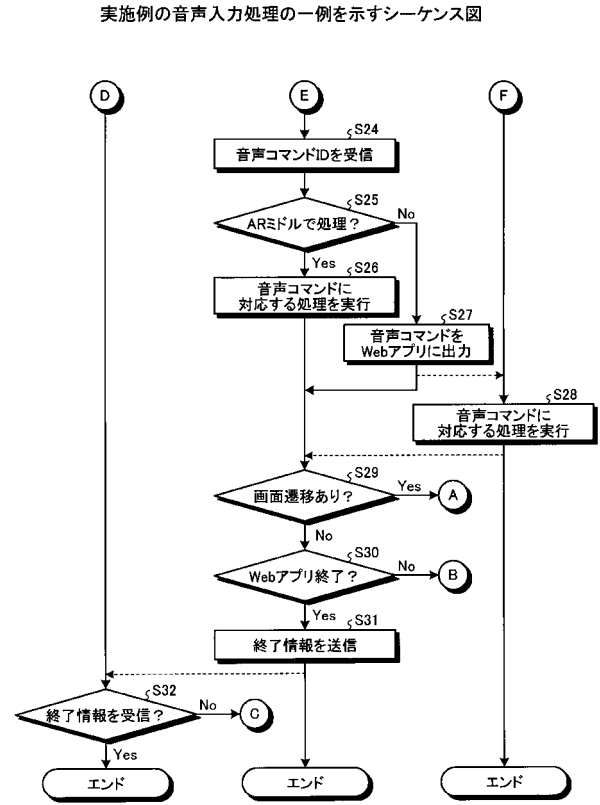
122

音声コマンドID	音声コマンド
1	メニュー表示
2	番号1選択
3	番号2選択
...	...

【 図 7 】



【 図 8 】



【 図 9 】

音声入力支援プログラムを実行するコンピュータの一例を示す図

