

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6798010号  
(P6798010)

(45) 発行日 令和2年12月9日(2020.12.9)

(24) 登録日 令和2年11月20日(2020.11.20)

(51) Int. Cl.	F I	
<b>G06F 3/01 (2006.01)</b>	G06F 3/01	510
<b>G06T 19/00 (2011.01)</b>	G06F 3/01	570
<b>G06T 7/00 (2017.01)</b>	G06T 19/00	600
<b>G06T 7/20 (2017.01)</b>	G06T 7/00	350C
<b>G09B 21/00 (2006.01)</b>	G06T 7/20	300A

請求項の数 44 (全 80 頁) 最終頁に続く

(21) 出願番号	特願2019-512813 (P2019-512813)	(73) 特許権者	514108838
(86) (22) 出願日	平成29年9月12日 (2017.9.12)		マジック リープ, インコーポレイテッド
(65) 公表番号	特表2019-535059 (P2019-535059A)		Magic Leap, Inc.
(43) 公表日	令和1年12月5日 (2019.12.5)		アメリカ合衆国 フロリダ 33322,
(86) 国際出願番号	PCT/US2017/051173		プランテーション, ウェスト サンライズ
(87) 国際公開番号	W02018/052901		ブルバード 7500
(87) 国際公開日	平成30年3月22日 (2018.3.22)		7500 W SUNRISE BLVD
審査請求日	令和2年9月11日 (2020.9.11)		, PLANTATION, FL 3332
(31) 優先権主張番号	62/440,320	(74) 代理人	100078282
(32) 優先日	平成28年12月29日 (2016.12.29)		弁理士 山本 秀策
(33) 優先権主張国・地域又は機関	米国 (US)	(74) 代理人	100113413
(31) 優先権主張番号	62/394,013		弁理士 森下 夏樹
(32) 優先日	平成28年9月13日 (2016.9.13)		
(33) 優先権主張国・地域又は機関	米国 (US)		最終頁に続く

(54) 【発明の名称】 感覚式アイウェア

(57) 【特許請求の範囲】

【請求項1】

手話認識のためのウェアラブルシステムであって、前記ウェアラブルシステムは、  
 仮想コンテンツをユーザに提示するように構成される頭部搭載型ディスプレイと、  
 前記ユーザの環境からオーディオを検出するように構成されるオーディオセンサと、  
 前記ユーザの前記環境を結像するように構成される外向きに面した結像システムと、  
 前記頭部搭載型ディスプレイおよび前記結像システムと通信するハードウェアプロセッサであって、前記ハードウェアプロセッサは、

前記外向きに面した結像システムによって捕捉された複数の画像を受信することと、  
 オブジェクト認識装置を用いて、前記複数の画像内の少なくとも1つの手のセットを検出することと、

前記少なくとも1つの手のセットの相対的なサイズを決定することと、  
 前記少なくとも1つの手のセットの前記相対的なサイズに基づいて、複数の人物の中から、前記少なくとも1つの手のセットのソースを識別することと、

前記オブジェクト認識装置を用いて、前記複数の画像内の前記少なくとも1つの手のセットによる少なくとも1つのジェスチャを検出することと、

前記少なくとも1つのジェスチャの前記ソースが、前記ユーザ以外の人物に属することを決定することと、

前記決定にตอบสนองして、手話における前記少なくとも1つのジェスチャの意味を認識することと、

前記ユーザと関連付けられたコンテキスト情報に基づいて、標的言語を識別すること  
であって、前記コンテキスト情報は、前記オーディオセンサによって取得されるオーディ  
オデータの分析から取得される前記ユーザの発話を含む、ことと、

前記認識された意味に基づいて、前記少なくとも1つのジェスチャの前記ソースが前  
記ユーザ以外の前記人物に属することの決定にさらに基づいて、前記少なくとも1つのジ  
ェスチャを前記標的言語に翻訳することと、

前記標的言語への前記ジェスチャの翻訳に少なくとも部分的に基づいて、仮想コンテ  
ンツを生成することと、

前記頭部搭載型ディスプレイに、前記仮想コンテンツを前記ユーザにレンダリングさ  
せることと

を行うようにプログラムされる、ハードウェアプロセッサと  
を備える、ウェアラブルシステム。

【請求項2】

前記結像システムは、前記ユーザの周囲を結像するように構成される広角カメラのうち  
の1つ以上のものを備える、請求項1に記載のウェアラブルシステム。

【請求項3】

前記ハードウェアプロセッサは、前記ジェスチャと関連付けられた補助情報にアクセス  
するようにさらにプログラムされ、前記頭部搭載型ディスプレイによってレンダリングさ  
れた仮想コンテンツは、前記補助情報を備える、請求項1に記載のウェアラブルシステム  
。

【請求項4】

前記ユーザと関連付けられたコンテキスト情報に基づいて、標的言語を識別するため  
に、前記ハードウェアプロセッサは、

前記ウェアラブルシステムによって捕捉されるような前記ユーザの発話、前記ユーザの  
場所、または前記言語を前記標的言語として選択する前記ユーザからの入力のうち少な  
くとも1つに基づいて、前記標的言語をユーザによって理解される言語として設定す  
る  
ようにプログラムされる、請求項1に記載のウェアラブルシステム。

【請求項5】

前記ハードウェアプロセッサは、前記標的言語が発話された言語であるかどうかを決定  
し、前記標的言語が発話された言語であることの決定に回答して、前記標的言語におい  
て、前記翻訳されたジェスチャと関連付けられた発話のオーディオストリームを再生す  
る  
ようにプログラムされる、請求項1に記載のウェアラブルシステム。

【請求項6】

前記ハードウェアプロセッサは、前記標的言語が別の手話であるかどうかを決定し、前  
記標的言語が別の手話であることの決定に回答して、前記他の手話における別のジェス  
チャのグラフィックを前記ジェスチャの翻訳として提示するようにプログラムされる、請  
求項1に記載のウェアラブルシステム。

【請求項7】

前記手話における前記ジェスチャの意味を認識するために、前記ハードウェアプロセッ  
サは、深層ニューラルネットワーク技法を前記結像システムによって捕捉された前記画像  
の少なくとも一部に適用するようにプログラムされる、請求項1に記載のウェアラブルシ  
ステム。

【請求項8】

前記ハードウェアプロセッサは、前記ユーザの場所に少なくとも部分的に基づいて、前  
記手話を候補手話のリストから識別するようにさらにプログラムされる、請求項1に記  
載のウェアラブルシステム。

【請求項9】

前記認識された意味に基づいて、前記ジェスチャを前記標的言語に翻訳するために、前  
記ハードウェアプロセッサは、前記ジェスチャを前記標的言語におけるテキスト表現に  
変換するようにプログラムされる、請求項1に記載のウェアラブルシステム。

10

20

30

40

50

## 【請求項 10】

前記ハードウェアプロセッサは、前記検出されたジェスチャのソースを決定し、前記検出されたジェスチャのソースが前記ウェアラブルシステムのユーザであることの決定に応じて、前記標的言語における前記ジェスチャの翻訳を別のユーザのウェアラブルシステムに通信するようにプログラムされる、請求項 1 に記載のウェアラブルシステム。

## 【請求項 11】

前記ハードウェアプロセッサは、前記仮想コンテンツを前記頭部搭載型ディスプレイによる表示から放逐するための条件を検出し、前記条件の検出に応答して、前記仮想コンテンツを前記頭部搭載型ディスプレイによる表示から除去するようにプログラムされる、請求項 1 に記載のウェアラブルシステム。

10

## 【請求項 12】

前記条件は、持続時間、ユーザの手のジェスチャ、またはユーザ入力デバイスからの入力のうちの少なくとも 1 つを含む、請求項 11 に記載のウェアラブルシステム。

## 【請求項 13】

前記画像は、ビデオの 1 つ以上のフレームを含む、請求項 1 に記載のウェアラブルシステム。

## 【請求項 14】

手話認識のための方法であって、前記方法は、  
ウェアラブルシステムの制御下において、前記ウェアラブルシステムは、前記ウェアラブルシステムのユーザに仮想コンテンツを提示するように構成される頭部搭載型ディスプレイを備え、

20

前記ウェアラブルシステムの結像システムによって捕捉された画像を受信することと、前記画像を分析し、人物のジェスチャを検出することと、前記人物のジェスチャと関連付けられた手のセットの相対的なサイズを決定することと

、前記相対的なサイズに基づいて、前記人物の前記ジェスチャのソースを識別することと

、前記人物のジェスチャの前記ソースが前記ユーザ以外の人物に属することを決定することと、

前記検出されたジェスチャに少なくとも部分的に基づいて、手話における通信の存在を検出することと、

30

前記人物のジェスチャが前記ユーザのジェスチャではない場合、前記手話における前記ジェスチャの意味を認識することと、

前記ジェスチャが翻訳されるであろう標的言語を識別することであって、前記標的言語を識別することは、前記ウェアラブルシステムによって取得されたオーディオデータを分析し、前記ユーザの発話を検出することを含む、ことと、

前記認識された意味に基づいて、前記人物のジェスチャの前記ソースが前記ユーザ以外の人物に属することの決定にさらに基づいて、前記ジェスチャを前記標的言語に翻訳することと、

前記標的言語への前記ジェスチャの翻訳に少なくとも部分的に基づいて、仮想コンテンツを生成することと、

40

前記頭部搭載型ディスプレイに、前記仮想コンテンツを前記ユーザにレンダリングさせることと

を含む、方法。

## 【請求項 15】

前記ウェアラブルシステムは、第 2 のウェアラブルシステムによってレンダリングするために第 2 の仮想コンテンツを前記第 2 のウェアラブルシステムに通信するようにさらに構成される、請求項 14 に記載の方法。

## 【請求項 16】

前記認識された意味に基づいて、前記ジェスチャを前記標的言語に翻訳することは、前

50

記ジェスチャを前記標的言語におけるテキスト表現に変換することを含む、請求項 1 4 に記載の方法。

【請求項 1 7】

前記仮想コンテンツは、前記標的言語におけるテキスト表現または前記標的言語における別のものを表すグラフィックを備える、請求項 1 4 に記載の方法。

【請求項 1 8】

前記手話における前記ジェスチャの意味を認識することは、深層ニューラルネットワーク技法を前記結像システムによって捕捉された画像の少なくとも一部に適用することを含む、請求項 1 4 に記載の方法。

【請求項 1 9】

前記手話における通信の存在を検出することは、  
前記手話を候補手話のリストから識別することと、  
前記検出されたジェスチャが前記手話における表現に対応することを決定することとを含む、請求項 1 4 に記載の方法。

【請求項 2 0】

前記検出されたジェスチャが前記手話における表現に対応することを決定することは、前記ジェスチャを行っている前記人物の唇の移動および前記人物が前記ジェスチャを行っている間に捕捉されたオーディオデータと関連させて前記ジェスチャを分析することを含む、請求項 1 9 に記載の方法。

【請求項 2 1】

コンピューティングシステムであって、  
ハードウェアコンピュータプロセッサと、  
その上に記憶されているソフトウェア命令を有する非一過性コンピュータ可読媒体であって、前記ソフトウェア命令は、

1 つ以上のネットワークを介して複数のウェアラブル A R デバイスと通信することと、  
前記複数のウェアラブル A R デバイスのうちの第 1 のウェアラブル A R デバイスの外向きに面した結像システムによって捕捉された画像情報を受信することであって、前記画像情報は、前記第 1 のウェアラブル A R デバイスの第 1 の装着者の 1 つ以上の手の画像を含む、ことと、

前記第 1 の装着者の前記 1 つ以上の手の画像の分析に少なくとも基づいて、前記受信された画像情報内の手話を自動的に検出することと、

前記検出された手話をテキストに変換することと、

前記 1 つ以上のネットワークを通して、前記複数のウェアラブル A R デバイスの中の第 2 のウェアラブル A R デバイスに前記変換されたテキストを伝送することであって、前記第 2 のウェアラブル A R デバイスは、前記変換されたテキストを前記第 2 のウェアラブル A R デバイスの第 2 の装着者に表示するように構成される、ことと

を含む動作を前記コンピューティングシステムに実施させるように前記ハードウェアコンピュータプロセッサによって実行可能である、システム。

【請求項 2 2】

前記ソフトウェア命令は、前記 1 つ以上のネットワークを通して、前記第 1 のウェアラブル A R デバイスの前記第 1 の装着者の世界マップを伝送することを前記コンピューティングシステムに行わせるようにさらに構成される、請求項 2 1 に記載のシステム。

【請求項 2 3】

前記第 1 の装着者の前記世界マップは、前記第 1 のウェアラブル A R デバイスの前記第 1 の装着者のアバタを含む、請求項 2 2 に記載のシステム。

【請求項 2 4】

前記ソフトウェア命令は、標的言語を識別することを前記コンピューティングシステムに行わせるようにさらに構成される、請求項 2 1 に記載のシステム。

【請求項 2 5】

前記ソフトウェア命令は、前記第 2 の A R デバイスによって捕捉されるような発話、前

10

20

30

40

50

記第2のARデバイスの場所、または、前記第2のARデバイスの装着者からの入力のうち少なくとも1つに基づいて、前記第2のARデバイスの前記第2の装着者によって理解される標的言語を識別することを前記コンピューティングシステムに行わせるようにさらに構成される、請求項21に記載のシステム。

【請求項26】

前記変換されたテキストは、前記標的言語である、請求項25に記載のシステム。

【請求項27】

手話を自動的に検出することは、前記受信された画像情報内の一連のジェスチャを検出することを含む、請求項21に記載のシステム。

【請求項28】

前記検出された手話をテキストに変換することは、手話辞書にアクセスし、検出されたジェスチャをテキストに翻訳することを含む、請求項27に記載のシステム。

【請求項29】

前記検出された手話をテキストに変換することは、機械学習アルゴリズムの使用を含む、請求項21に記載のシステム。

【請求項30】

前記ソフトウェア命令は、前記検出された手話のソースを決定することであって、前記ソースは、人物を含む、ことと、前記決定されたソースに基づいて、前記検出された手話を変換することとを前記コンピューティングシステムに行わせるようにさらに構成される、請求項21に記載のシステム。

【請求項31】

前記検出された手話のソースを決定することは、少なくとも前記画像情報内の前記1つ以上の手のサイズに基づく、請求項30に記載のシステム。

【請求項32】

前記ソフトウェア命令は、前記変換されたテキストに基づいて、オーディオストリームを前記第2のウェアラブルARデバイスに伝送することを前記コンピューティングシステムに行わせるようにさらに構成される、請求項21に記載のシステム。

【請求項33】

ウェアラブルARデバイス間の通信を促進するための方法であって、前記方法は、  
1つ以上のネットワークを介して複数のウェアラブルARデバイスと通信することと、  
前記複数のウェアラブルARデバイスのうちの第1のウェアラブルARデバイスの外向きに面した結像システムによって捕捉された画像情報を受信することであって、前記画像情報は、前記第1のウェアラブルARデバイスの第1の装着者の1つ以上の手の画像を含む、ことと、

前記第1の装着者の前記1つ以上の手の画像の分析に少なくとも基づいて、前記受信された画像情報内の手話を自動的に検出することと、

前記検出された手話をテキストに変換することと、

前記1つ以上のネットワークを通して、前記複数のウェアラブルARデバイスの中の第2のウェアラブルARデバイスに前記変換されたテキストを伝送することであって、前記第2のウェアラブルARデバイスは、前記変換されたテキストを前記第2のウェアラブルARデバイスの第2の装着者に表示するように構成される、ことと

を含む、方法。

【請求項34】

前記1つ以上のネットワークを通して、前記第1のウェアラブルARデバイスの前記第1の装着者の世界マップを伝送することを含む、請求項33に記載の方法。

【請求項35】

前記第1の装着者の前記世界マップは、前記第1のウェアラブルARデバイスの前記第1の装着者のアバタを含む、請求項34に記載の方法。

【請求項36】

標的言語を識別することを含む、請求項33に記載の方法。

10

20

30

40

50

## 【請求項 37】

前記第2のARデバイスによって捕捉されるような発話、前記第2のARデバイスの場所、または、前記第2のARデバイスの装着者からの入力のうち少なくとも1つに基づいて、前記第2のARデバイスの前記第2の装着者によって理解される標的言語を識別することを含み、請求項33に記載の方法。

## 【請求項 38】

前記変換されたテキストは、前記標的言語である、請求項37に記載の方法。

## 【請求項 39】

手話を自動的に検出することは、前記受信された画像情報内の一連のジェスチャを検出することを含み、請求項33に記載の方法。

10

## 【請求項 40】

前記検出された手話をテキストに変換することは、手話辞書にアクセスし、検出されたジェスチャをテキストに翻訳することを含み、請求項33に記載の方法。

## 【請求項 41】

前記検出された手話をテキストに変換することは、機械学習アルゴリズムの使用を含む、請求項33に記載の方法。

## 【請求項 42】

前記検出された手話のソースを決定することであって、前記ソースは、人物を含む、ことと、

前記決定されたソースに基づいて、前記検出された手話を変換することを含む、請求項33に記載の方法。

20

## 【請求項 43】

前記画像情報内の前記1つ以上の手のサイズに基づいて、前記検出された手話のソースが前記第1のウェアラブルARデバイスの装着者であることを決定することを含み、請求項33に記載の方法。

## 【請求項 44】

前記変換されたテキストに基づいて、オーディオストリームを前記第2のウェアラブルARデバイスに伝送することを含み、請求項33に記載の方法。

## 【発明の詳細な説明】

30

## 【技術分野】

## 【0001】

(関連出願の相互参照)

本願は、米国仮出願第62/394,013号、出願日2016年9月13日、発明の名称“SENSORY EYEWEAR”および米国仮出願第62/440,320号、出願日2016年12月29日、発明の名称“SYSTEMS AND METHODS FOR AUGMENTED REALITY”に対する35 U.S.C. § 119(e)のもとでの優先権の利益を主張するものであり、これらの開示の全体は、参照により本明細書中に援用される。

## 【0002】

40

本開示は、仮想現実および拡張現実結像ならびに可視化システムに関し、より具体的には、環境内の手話またはテキストを認識し、認識された手話またはテキストに基づいて、仮想コンテンツをレンダリングすることに関する。

## 【背景技術】

## 【0003】

現代のコンピューティングおよびディスプレイ技術は、デジタル的に再現された画像またはその一部が、現実であるように見える、もしくはそのように知覚され得る様式でユーザに提示される、いわゆる「仮想現実」、「拡張現実」、または「複合現実」体験のためのシステムの開発を促進している。仮想現実または「VR」シナリオは、典型的には、他の実際の実世界の視覚的入力に対する透過性を伴わずに、デジタルまたは仮想画像情報の

50

提示を伴う。拡張現実または「AR」シナリオは、典型的には、ユーザの周囲の実際の世界の可視化に対する拡張としてのデジタルまたは仮想画像情報の提示を伴う。複合現実または「MR」は、物理的および仮想オブジェクトが、共存し、リアルタイムで相互作用する、新しい環境を生成するための実世界と仮想世界の融合に関連する。結論から述べると、ヒトの視知覚系は、非常に複雑であって、他の仮想または実世界画像要素の中で仮想画像要素の快適かつ自然のような感覚で豊かな提示を促進する、VR、AR、またはMR技術の生産は、困難である。本明細書に開示されるシステムおよび方法は、VR、AR、ならびにMR技術に関連する種々の課題に対処する。

【発明の概要】

【課題を解決するための手段】

10

【0004】

環境内の手話およびテキストを認識するための複合現実システムの種々の実施形態が、開示される。これらの実施形態は、有利には、身体障害者間のより優れた相互作用を可能にし得る。

【0005】

複合現実デバイスのための感覚式アイウェアシステムは、他の人々または環境とのユーザの相互作用を促進することができる。一実施例として、感覚式アイウェアシステムは、手話を認識および解釈し、翻訳された情報を複合現実デバイスのユーザに提示することができる。ウェアラブルシステムはまた、ユーザの環境内のテキストを認識し、テキストを修正し（例えば、テキストのコンテンツまたは表示特性を変化させることによって）、修正されたテキストをレンダリングし、オリジナルテキストをオクルードすることができる。

20

【0006】

本明細書に説明される主題の1つ以上の実装の詳細が、付随の図面および以下の説明に記載される。他の特徴、側面、および利点は、説明、図面、ならびに請求項から明白となるであろう。本概要または以下の発明を実施するための形態のいずれも、本発明の主題の範囲を定義または限定することを主張するものではない。

本発明は、例えば、以下を提供する。

(項目1)

手話認識のためのウェアラブルシステムであって、前記ウェアラブルシステムは、  
仮想コンテンツをユーザに提示するように構成される頭部搭載型ディスプレイと、  
前記ユーザの環境を結像するように構成される結像システムと、  
前記頭部搭載型ディスプレイおよび前記結像システムと通信するハードウェアプロセッサであって、前記ハードウェアプロセッサは、

30

前記結像システムによって捕捉された画像を受信することと、

オブジェクト認識装置を用いて、前記画像内のジェスチャを検出することと、

手話における前記ジェスチャの意味を認識することと、

前記ユーザと関連付けられたコンテキスト情報に基づいて、標的言語を識別することと、

前記認識された意味に基づいて、前記ジェスチャを前記標的言語に翻訳することと、

40

少なくとも部分的に、前記標的言語への前記ジェスチャの翻訳に基づいて、仮想コンテンツを生成することと、

前記頭部搭載型ディスプレイに、前記仮想コンテンツをユーザにレンダリングさせることと

を行うようにプログラムされる、ハードウェアプロセッサと

を備える、ウェアラブルシステム。

(項目2)

前記結像システムは、前記ユーザの周囲を結像するように構成される広角カメラのうちの1つ以上のものを備える、項目1に記載のウェアラブルシステム。

(項目3)

50

前記ハードウェアプロセッサはさらに、前記ジェスチャと関連付けられた補助情報にアクセスするようにプログラムされ、前記頭部搭載型ディスプレイによってレンダリングされた仮想コンテンツは、前記補助情報を備える、項目 1 に記載のウェアラブルシステム。

(項目 4)

前記ユーザと関連付けられたコンテキスト情報に基づいて、標的言語を識別するために、前記ハードウェアプロセッサは、

前記ウェアラブルシステムによって捕捉されるような前記ユーザの発話、前記ユーザの場所、または前記言語を前記標的言語として選択する前記ユーザからの入力のうち少なくとも 1 つに基づいて、前記標的言語をユーザによって理解される言語として設定する

ようにプログラムされる、項目 1 に記載のウェアラブルシステム。

(項目 5)

前記ハードウェアプロセッサは、前記標的言語が発話された言語であるかどうかを決定し、前記標的言語が発話された言語であることの決定に回答して、前記標的言語において、前記翻訳されたジェスチャと関連付けられた発話のオーディオストリームを再生するようにプログラムされる、項目 1 に記載のウェアラブルシステム。

(項目 6)

前記ハードウェアプロセッサは、前記標的言語が別の手話であるかどうかを決定し、前記標的言語が別の手話であることの決定に回答して、前記他の手話における別のジェスチャのグラフィックを前記ジェスチャの翻訳として提示するようにプログラムされる、項目 1 に記載のウェアラブルシステム。

(項目 7)

前記手話における前記ジェスチャの意味を認識するために、前記ハードウェアプロセッサは、深層ニューラルネットワーク技法を前記結像システムによって捕捉された前記画像の一部に適用するようにプログラムされる、項目 1 に記載のウェアラブルシステム。

(項目 8)

前記ハードウェアプロセッサはさらに、少なくとも部分的に、前記ユーザの場所に基づいて、前記手話を候補手話のリストから識別するようにプログラムされる、項目 1 に記載のウェアラブルシステム。

(項目 9)

前記認識された意味に基づいて、前記ジェスチャを前記標的言語に翻訳するために、前記ハードウェアプロセッサは、前記ジェスチャを前記標的言語におけるテキスト表現に変換するようにプログラムされる、項目 1 に記載のウェアラブルシステム。

(項目 10)

前記ハードウェアプロセッサは、前記検出されたジェスチャ源を決定し、前記検出されたジェスチャ源が前記ウェアラブルシステムのユーザであることの決定に応じて、前記標的言語における前記ジェスチャの翻訳を別のユーザのウェアラブルシステムに通信するようにプログラムされる、項目 1 に記載のウェアラブルシステム。

(項目 11)

前記ハードウェアプロセッサは、前記仮想コンテンツを前記頭部搭載型ディスプレイによる表示から放逐することからの条件を検出し、前記条件の検出に回答して、前記仮想コンテンツを前記頭部搭載型ディスプレイによる表示から除去するようにプログラムされる、項目 1 に記載のウェアラブルシステム。

(項目 12)

前記条件は、持続時間、ユーザの手のジェスチャ、またはユーザ入力デバイスからの入力のうち少なくとも 1 つを含む、項目 11 に記載のウェアラブルシステム。

(項目 13)

前記画像は、ビデオの 1 つ以上のフレームを含む、項目 1 に記載のウェアラブルシステム。

(項目 14)

手話認識のための方法であって、前記方法は、

10

20

30

40

50

結像システムによって捕捉された画像を受信することと、  
前記画像を分析し、ユーザのジェスチャを検出することと、  
少なくとも部分的に、前記検出されたジェスチャに基づいて、手話における通信の存在を検出することと、  
前記手話における前記ジェスチャの意味を認識することと、  
前記ジェスチャが翻訳されるであろう標的言語を識別することと、  
前記認識された意味に基づいて、前記ジェスチャを前記標的言語に翻訳することと、  
少なくとも部分的に、前記標的言語への前記ジェスチャの翻訳に基づいて、仮想コンテンツを生成することと、  
頭部搭載型ディスプレイに、前記仮想コンテンツをユーザにレンダリングさせることと  
を含む、方法。

10

(項目15)

前記画像は、複合現実コンテンツを提示するように構成される第1のウェアラブルデバイスから受信される一方、前記仮想コンテンツは、レンダリングのために、第2のウェアラブルデバイスに通信され、前記第1のウェアラブルデバイスおよび前記第2のウェアラブルデバイスは、複合現実コンテンツをユーザに提示するように構成される、項目14に記載の方法。

(項目16)

前記認識された意味に基づいて、前記ジェスチャを前記標的言語に翻訳することは、前記ジェスチャを前記標的言語におけるテキスト表現に変換することを含む、項目14に記載の方法。

20

(項目17)

前記仮想コンテンツは、前記標的言語におけるテキスト表現または前記標的言語における別のものを表すグラフィックを備える、項目14に記載の方法。

(項目18)

前記手話における前記ジェスチャの意味を認識することは、深層ニューラルネットワーク技法を前記結像システムによって捕捉された画像の一部に適用することを含む、項目14に記載の方法。

(項目19)

前記手話における通信の存在を検出することは、  
前記手話を候補手話のリストから識別することと、  
前記検出されたジェスチャが前記手話における表現に対応することを決定することと  
を含む、項目14に記載の方法。

30

(項目20)

前記検出されたジェスチャが前記手話における表現に対応することを決定することは、前記ジェスチャを行っている人物の唇の移動および前記ユーザが前記ジェスチャを行っている間に捕捉されたオーディオデータと関連させて前記ジェスチャを分析することを含む、項目19に記載の方法。

【図面の簡単な説明】

【0007】

【図1】図1は、人物によって視認されるある仮想現実オブジェクトおよびある物理的オブジェクトを伴う、複合現実シナリオの例証を描写する。

【図2A】図2Aは、感覚式アイウェアシステムを実装し得る、ウェアラブルシステムの実施例を図式的に図示する。

【図2B】図2Bは、ウェアラブルシステムの種々の例示的コンポーネントを図式的に図示する。

【図3】図3は、複数の深度平面を使用して3次元画像をシミュレートするためのアプローチの側面を図式的に図示する。

【図4】図4は、画像情報をユーザに出力するための導波管スタックの実施例を図式的に図示する。

40

50

【図5】図5は、導波管によって出力され得る、例示的出射ビームを示す。

【図6】図6は、導波管装置と、光を導波管装置へまたはそこから光学的に結合するための光学結合器サブシステムと、多焦点立体ディスプレイ、画像、またはライトフィールドの生成において使用される、制御サブシステムとを含む、光学システムを示す、概略図である。

【図7】図7は、ウェアラブルシステムの実施例のブロック図である。

【図8】図8は、認識されるオブジェクトに関連して仮想コンテンツをレンダリングする方法の実施例のプロセスフロー図である。

【図9】図9は、感覚式アイウェアシステムを含む、ウェアラブルシステムの別の実施例のブロック図である。

10

【図10】図10は、ウェアラブルシステムへのユーザ入力を決定するための方法の実施例のプロセスフロー図である。

【図11】図11は、仮想ユーザインターフェースと相互作用するための方法の実施例のプロセスフロー図である。

【図12】図12は、相互に相互作用する複数のウェアラブルシステムを描写する、全体的システム図を図式的に図示する。

【図13A】図13Aは、感覚式アイウェアシステムの例示的ユーザ体験を示し、感覚式アイウェアシステムは、手話（例えば、手話者によってジェスチャされる）を解釈することができる。

【図13B】図13Bは、感覚式アイウェアシステムの別の例示的ユーザ体験を示し、標的発話および補助情報の両方が、提示される。

20

【図13C】図13Cは、テレプレゼンスセッションにおける感覚式アイウェアシステムの例示的ユーザ体験を示す。

【図13D】図13Dは、手話を解釈するための例示的仮想ユーザインターフェースを図示する。

【図14A】図14Aおよび14Bは、感覚式アイウェアシステムを用いた対人通信を促進するための例示的プロセスを図示する。

【図14B】図14Aおよび14Bは、感覚式アイウェアシステムを用いた対人通信を促進するための例示的プロセスを図示する。

【図14C】図14Cは、変換されたテキストと関連付けられた補助情報を決定し、補助情報を提示するための例示的方法のプロセスフロー図である。

30

【図15】図15は、感覚式アイウェアシステムを用いた対人通信を促進するための別の例示的プロセスを図示する。

【図16A】図16A - 16Eは、環境内のテキストを認識し、テキストの表示特性を修正し、修正されたテキストをレンダリングするように構成される、感覚式アイウェアシステムのための例示的ユーザ体験を図示する。

【図16B】図16A - 16Eは、環境内のテキストを認識し、テキストの表示特性を修正し、修正されたテキストをレンダリングするように構成される、感覚式アイウェアシステムのための例示的ユーザ体験を図示する。

【図16C】図16A - 16Eは、環境内のテキストを認識し、テキストの表示特性を修正し、修正されたテキストをレンダリングするように構成される、感覚式アイウェアシステムのための例示的ユーザ体験を図示する。

40

【図16D】図16A - 16Eは、環境内のテキストを認識し、テキストの表示特性を修正し、修正されたテキストをレンダリングするように構成される、感覚式アイウェアシステムのための例示的ユーザ体験を図示する。

【図16E】図16A - 16Eは、環境内のテキストを認識し、テキストの表示特性を修正し、修正されたテキストをレンダリングするように構成される、感覚式アイウェアシステムのための例示的ユーザ体験を図示する。

【図17】図17は、環境とのユーザの相互作用を促進するための感覚式アイウェアの例示的プロセスを図示する。

50

【図18】図18は、表記のコンテンツを修正することによって、ユーザが物理的環境内の表記を理解することを補助する実施例を図示する。

【図19】図19は、ユーザが物理的環境内の表記を理解することを補助する例示のプロセスを図示する。

【発明を実施するための形態】

【0008】

図面全体を通して、参照番号は、参照される要素間の対応を示すために再使用され得る。図面は、本明細書に説明される例示の実施形態を図示するために提供され、本開示の範囲を限定することを意図されない。

(概要)

【0009】

AR/VR/MRコンテンツを提示するように構成される、ウェアラブルシステムは、感覚式アイウェアシステムを実装し、他の人々または環境とのユーザの相互作用を向上させることができる。例示的ウェアラブルシステムは、頭部搭載型ディスプレイと、種々の結像センサと、1つ以上のハードウェアプロセッサとを備えることができる。ディスプレイは、片眼または両眼の正面に装着される、シースルーディスプレイであることができる。

【0010】

他の人々とのユーザの相互作用体験を向上させるために、ウェアラブルシステムは、手話を捕捉および解釈するように構成されることができる。手話は、主に、視覚的ジェスチャ（例えば、手の形状、手の配向、手、腕、または身体の移動、もしくは表情）を使用して、通信する。数百もの手話が、世界中で使用されている。いくつかの手話は、その他より頻繁に使用され得る。例えば、アメリカ手話（ASL）は、米国およびカナダにおいて広く使用されている。

【0011】

多くの人々は、任意の手話を把握していない。言語障害者または聴覚障害者およびその会話相手が、同一手話を熟知しているとは限らない。これは、聴覚障害者または言語障害者との会話を妨害し得る。故に、会話相手によって行われている身ぶり（例えば、ジェスチャ）を結像し、身ぶりをテキストまたはグラフィック（例えば、システムユーザが理解する手話における手話ジェスチャのグラフィック）に変換し、次いで、身ぶりに関連付けられた情報（例えば、身ぶりのユーザによって理解される言語への翻訳）を表示することができる。ウェアラブルシステムは、ユーザと会話相手との間の通信を改良することに著しく役立つことができる。さらに、ウェアラブルシステムのユーザに対して最小限のレベルの注意散漫およびそれによる取るに足りないレベルの労力を伴って、（近）リアルタイムで手話のテキストまたはグラフィカル変換を提供することができる。ウェアラブルシステムを有することが望ましくあり得る。

【0012】

本開示は、ウェアラブルシステムのコンテキストにおいて、そのような望ましいシステムの実施例を開示する。ウェアラブルシステムは、頭部搭載型コンポーネント（例えば、頭部搭載型ディスプレイ等）を含んでもよい。そのようなデバイスは、ユーザが、情報が通常の視認可能実世界とともに（またはその上部に）同時に視認可能であるような様式において、コンピューティングデバイスによって提供される情報を視覚的に受信することができる。そのようなシステムは、表音文字、画像効果、テキスト、グラフィック、または任意の種類ビデオ等、従来のコンピュータ画面上に表示され得る、任意の形態の情報を表示するために使用されることができる。

【0013】

本明細書に説明されるウェアラブルシステムは、ウェアラブルシステムの手話認識（SLR）と表示能力を組み合わせ、ユーザに検出された手話に基づく情報を提供することができる。例えば、ウェアラブルシステム上の外向きに面したカメラは、行われているジェスチャを結像し、ジェスチャ間の身ぶりを識別し、身ぶりをユーザが理解する言語に翻訳

10

20

30

40

50

し、翻訳をユーザに表示することができる。検出された手話のトランスクリプト（例えば、キャプションまたはテキスト吹き出し）が、ウェアラブルシステムによって、ユーザに表示されることができる。機械学習アルゴリズム（例えば、深層ニューラルネットワーク）は、画像を受信し、身ぶりの識別および翻訳を実施することができる。ユーザによってプロンプトされると、トランスクリプト内の単語の意味または適切なソースからの関連情報が、表示されることができる。ウェアラブルシステムが提供し得る、補助情報の種類は、例えば、インターネット上の膨大な数の利用可能な情報リソースと同様に無限であり得る。

【 0 0 1 4 】

他の人々とのユーザの相互作用体験を向上させることに加え、またはその代替として、感覚式アイウェアシステムはまた、環境とのユーザの体験を改良することができる。環境とのユーザ相互作用を改良する実施例として、感覚式アイウェアシステムを実装するウェアラブルシステムは、環境内のテキスト（例えば、商業用または公共看板等の表記上の、例えば、テキスト）を認識し、テキストの表示特性を修正（例えば、テキストのサイズを増加させることによって）またはテキストのコンテンツを修正（例えば、テキストを別の言語に翻訳することによって）し、修正されたテキストを物理的環境内のテキストにわたってレンダリングすることができる。

【 0 0 1 5 】

本明細書にさらに説明されるように、ウェアラブルシステムは、ユーザの環境の画像を受信することができる。画像は、ウェアラブルシステムの外向きに面した結像システムまたはウェアラブルシステムと関連付けられたトーテムによって入手されてもよい。ウェアラブルシステムは、画像は、1つ以上の表音文字または表意文字を含むかどうかを決定し、1つ以上の表音文字または表意文字をテキストに変換することができる。ウェアラブルシステムは、例えば、機械学習アルゴリズムまたは光学文字認識（OCR）アルゴリズム等の種々の技法を使用して、画像が表音文字または表意文字を含むかどうかを決定してもよい。ウェアラブルシステムは、オブジェクト認識装置（例えば、図7に説明される）を使用して、表音文字および表意文字を識別し、それらをテキストに変換してもよい。

【 0 0 1 6 】

ある実施形態では、テキストは、ユーザにウェアラブルシステムを伴わずに見えるであろうもの異なるように、ユーザのための表示されることができる。例えば、ウェアラブルシステムは、頭部搭載型ディスプレイに、オリジナル画像と関連付けられた表音文字または表意文字と関連付けられたフォントサイズと異なるフォントサイズにおいて、テキストを表示させることができる。ウェアラブルシステムはまた、テキストの表示品質を改良することができる。例えば、霧、霞、雨、明光、低光量、表音文字と周囲画像との間の低光量または色コントラスト等の種々の環境要因は、ウェアラブルシステムを用いなければ、環境内のテキストのユーザのクリアなビューを妨げ得る。ウェアラブルシステムは、テキストの明確性を増加させるであろう、標識を提示してもよい（例えば、増加されたコントラスト比またはより大きいフォントを用いて）。

【 0 0 1 7 】

ウェアラブルシステムはまた、テキスト（例えば、テキストを上に表示）をそのオリジナル言語から標的言語に翻訳することができる。例えば、テキストは、ユーザが理解していない言語からユーザが理解している言語に翻訳されてもよい。翻訳されたテキストは、ユーザが、ユーザが理解可能な言語におけるテキストを容易に視認し得るように、オリジナルテキストにわたってレンダリングされてもよい。  
(ウェアラブルシステムの3Dディスプレイの実施例)

【 0 0 1 8 】

ウェアラブルシステム（本明細書では、拡張現実（AR）システムとも称される）は、2Dまたは3D仮想画像をユーザに提示するために構成されることができる。画像は、組み合わせまたは同等物における、静止画像、ビデオのフレーム、またはビデオであってもよい。ウェアラブルシステムの少なくとも一部は、ユーザ相互作用のために、単独で、ま

10

20

30

40

50

たは組み合わせて、VR、AR、またはMR環境を提示し得る、ウェアラブルデバイス上に実装されることができる。ウェアラブルデバイスは、頭部搭載型デバイス(HMD)であることができ、これは、ARデバイス(ARD)と同義的に使用される。さらに、本開示の目的のために、用語「AR」は、用語「MR」と同義的に使用される。

#### 【0019】

図1は、人物によって視認される、ある仮想現実オブジェクトおよびある物理的オブジェクトを伴う、複合現実シナリオの例証を描写する。図1では、MR場面100が、描写され、MR技術のユーザには、人々、木々、背景における建物、およびコンクリートプラットフォーム120を特徴とする、実世界公園状設定110が見える。これらのアイテムに加え、MR技術のユーザはまた、実世界プラットフォーム120上に立っているロボット像130と、マルハナバチの擬人化のように見える、飛んでいる漫画のようなアバタキャラクター140とが「見える」と知覚するが、これらの要素は、実世界には存在しない。

10

#### 【0020】

3Dディスプレイが、真の深度感覚、より具体的には、表面深度のシミュレートされた感覚を生成するために、ディスプレイの視野内の点毎に、その仮想深度に対応する遠近調節応答を生成することが望ましくあり得る。ディスプレイ点に対する遠近調節応答が、収束および立体視の両眼深度キューによって決定されるようなその点の仮想深度に対応しない場合、ヒトの眼は、遠近調節衝突を体験し、不安定な結像、有害な眼精疲労、頭痛、および遠近調節情報の不在下では、表面深度のほぼ完全な欠如をもたらし得る。

#### 【0021】

20

VR、AR、およびMR体験は、複数の深度平面に対応する画像が視認者に提供されるディスプレイを有する、ディスプレイシステムによって提供されることができる。画像は、深度平面毎に異なってもよく(例えば、場面またはオブジェクトの若干異なる提示を提供する)、視認者の眼によって別個に集束され、それによって、異なる深度平面上に位置する場面に関する異なる画像特徴に合焦させるために要求される眼の遠近調節に基づいて、または合焦からずれている異なる深度平面上の異なる画像特徴を観察することに基づいて、ユーザに深度キューを提供することに役立ち得る。本明細書のいずれかに議論されるように、そのような深度キューは、信用できる深度の知覚を提供する。

#### 【0022】

図2Aは、ウェアラブルシステム200の実施例を図示し、これは、AR/VR/MR場面を提供するように構成されることができる。ウェアラブルシステム200はまた、ARシステム200と称され得る。ウェアラブルシステム200は、ディスプレイ220と、ディスプレイ220の機能をサポートするための種々の機械的ならびに電子的モジュールおよびシステムとを含む。ディスプレイ220は、ユーザ、装着者、または視認者210によって装着可能である、フレーム230に結合されてもよい。ディスプレイ220は、ユーザ210の眼の正面に位置付けられることができる。ディスプレイ220は、AR/VR/MRコンテンツをユーザに提示することができる。ディスプレイ220は、ユーザの頭部上に装着される、頭部搭載型ディスプレイを備えることができる。いくつかの実施形態では、スピーカ240が、フレーム230に結合され、ユーザの外耳道に隣接して位置付けられる(いくつかの実施形態では、示されない別のスピーカが、ユーザの他方の外耳道に隣接して位置付けられ、ステレオ/成形可能音響制御を提供する)。ディスプレイ220は、環境からオーディオストリームを検出し、周囲音を捕捉するために、オーディオセンサ(例えば、マイクロホン)232を含むことができる。いくつかの実施形態では、示されない1つ以上の他のオーディオセンサが、ステレオ音受信を提供するために位置付けられる。ステレオ音受信は、音源の場所を決定するために使用されることができる。ウェアラブルシステム200は、音声または発話認識をオーディオストリームに実施することができる。

30

40

#### 【0023】

ウェアラブルシステム200は、ユーザの周囲の環境内の世界を観察する、外向きに面した結像システム464(図4に示される)を含むことができる。ウェアラブルシステム

50

200はまた、ユーザの眼移動を追跡することができる、内向きに面した結像システム462(図4に示される)を含むことができる。内向きに面した結像システムは、一方の眼の移動または両方の眼の移動のいずれかを追跡することができる。内向きに面した結像システム462は、フレーム230に取り付けられてもよく、内向きに面した結像システムによって入手された画像情報を処理し、例えば、ユーザ210の眼、眼の移動、または眼姿勢の瞳孔直径もしくは配向を決定し得る、処理モジュール260または270と電気通信してもよい。

【0024】

実施例として、ウェアラブルシステム200は、外向きに面した結像システム464または内向きに面した結像システム462を使用して、ユーザの姿勢の画像を入手することができる。画像は、静止画像、ビデオのフレーム、またはビデオであってもよい。

10

【0025】

ディスプレイ220は、有線導線または無線接続等によって、フレーム230に固定して取り付けられる、ユーザによって装着されるヘルメットもしくは帽子に固定して取り付けられる、ヘッドホンに内蔵される、または別様にユーザ210に除去可能に取り付けられる(例えば、リュック式構成において、ベルト結合式構成において)等、種々の構成において搭載され得る、ローカルデータ処理モジュール260に動作可能に結合されることができる(250)。

【0026】

ローカル処理およびデータモジュール260は、ハードウェアプロセッサならびに不揮発性メモリ(例えば、フラッシュメモリ)等のデジタルメモリを備えてもよく、その両方とも、データの処理、キャッシュ、および記憶を補助するために利用され得る。データは、画像捕捉デバイス(例えば、内向きに面した結像システムおよび/または外向きに面した結像システム内のカメラ)、オーディオセンサ(例えば、マイクロホン)、慣性測定ユニット(IMU)、加速度計、コンパス、全地球測位システム(GPS)ユニット、無線デバイス、もしくはジャイロスコープ等の(例えば、フレーム230に動作可能に結合される、または別様にユーザ210に取り付けられ得る)センサから捕捉されるデータ(a)または、場合によっては処理もしくは読出後にディスプレイ220への通過のために、遠隔処理モジュール270もしくは遠隔データリポジトリ280を使用して入手もしくは処理されるデータ(b)を含んでもよい。ローカル処理およびデータモジュール260は、これらの遠隔モジュールがローカル処理およびデータモジュール260へのリソースとして利用可能であるように、有線または無線通信リンク等を介して、通信リンク262または264を遠隔処理モジュール270または遠隔データリポジトリ280に動作可能に結合されてもよい。加えて、遠隔処理モジュール280および遠隔データリポジトリ280は、相互に動作可能に結合されてもよい。

20

30

【0027】

いくつかの実施形態では、遠隔処理モジュール270は、データまたは画像情報を分析および処理するように構成される、1つ以上のプロセッサを備えてもよい。いくつかの実施形態では、遠隔データリポジトリ280は、デジタルデータ記憶設備を備え得、これは、インターネットまたは「クラウド」リソース構成における他のネットワーキング構成を通して利用可能であってもよい。いくつかの実施形態では、全てのデータが、記憶され、全ての算出が、ローカル処理およびデータモジュールにおいて実施され、遠隔モジュールからの完全に自律的な使用を可能にする。

40

【0028】

図2Bは、ウェアラブルシステム200を示し、これは、ディスプレイ220と、フレーム230とを含むことができる。引き伸ばし図202は、ウェアラブルシステム200の種々のコンポーネントを図式的に図示する。ある実装では、図2Bに図示されるコンポーネントのうちの1つ以上のものは、ディスプレイ220の一部であることができる。種々のコンポーネントは、単独で、または組み合わせて、ウェアラブルシステム200のユーザまたはユーザの環境と関連付けられた種々のデータ(例えば、聴覚的または視覚的デ

50

ータ等)を収集することができる。他の実施形態は、ウェアラブルシステムが使用される用途に応じて、付加的またはより少ないコンポーネントを有してもよいことを理解されたい。なお、図2Bは、種々のコンポーネントのうちのいくつかと、ウェアラブルシステムを通して収集、分析、および記憶され得る、データのタイプとを提供する基本概念を提供する。

#### 【0029】

図2Bは、例示的ウェアラブルシステム200を示し、これは、ディスプレイ220を含むことができる。ディスプレイ220は、ユーザの頭部、またはフレーム230に対応する、筐体もしくはフレーム108に搭載され得る、ディスプレイレンズ106を備えることができる。ディスプレイレンズ106は、筐体108によって、ユーザの眼302、304の正面に位置付けられる、1つ以上の透明ミラーを備えてもよく、投影された光38を眼302、304の中にバウンスさせ、ビーム成形を促進しながら、また、ローカル環境からの少なくとも一部の光の透過を可能にするように構成されてもよい。投影された光ビーム38の波面は、投影された光の所望の焦点距離と一致するように屈曲または集束されてもよい。図示されるように、2つの広視野マシビジョンカメラ16(世界カメラとも称される)が、筐体108に結合され、ユーザの周囲の環境を結像することができる。これらのカメラ16は、二重捕捉式可視光/非可視(例えば、赤外線)光カメラであることができる。カメラ16は、図4に示される外向きに面した結像システム464の一部であってもよい。世界カメラ16によって入手された画像は、姿勢プロセッサ36によって処理されることができる。例えば、姿勢プロセッサ36は、1つ以上のオブジェクト認識装置708(例えば、図7に示される)を実装し、ユーザまたはユーザの環境内の別の人物の姿勢を識別する、またはユーザの環境内の物理的オブジェクトを識別することができる。

#### 【0030】

図2Bを継続して参照すると、光38を眼302、304の中に投影するように構成される、ディスプレイミラーおよび光学を伴う、一对の走査式レーザ成形波面(例えば、深度のために)光プロジェクタモジュールが、示される。描写される図はまた、ユーザの眼302、304を追跡し、レンダリングおよびユーザ入力をサポート可能であるように構成される、赤外線光源26(発光ダイオード「LED」等)とペアリングされる、2つの小型赤外線カメラ24を示す。カメラ24は、図4に示される、内向きに面した結像システム462の一部であってもよい。ウェアラブルシステム200はさらに、センサアセンブリ39を特徴とすることができ、これは、X、Y、およびZ軸加速度計能力ならびに磁気コンパスおよびX、Y、およびZ軸ジャイロスコープ能力を備え、好ましくは、200Hz等の比較的到高周波数でデータを提供し得る。センサアセンブリ39は、図2Aを参照して説明される、IMUの一部であってもよい。描写されるシステム200はまた、ASIC(特定用途向け集積回路)、FPGA(フィールドプログラマブルゲートアレイ)、またはARMプロセッサ(高度縮小命令セット機械)等の頭部姿勢プロセッサ36を備えることができ、これは、リアルタイムまたは近リアルタイムユーザ頭部姿勢を捕捉デバイス16から出力された広視野画像情報からを計算するように構成されてもよい。頭部姿勢プロセッサ36は、ハードウェアプロセッサであることができ、図2Aに示されるローカル処理およびデータモジュール260の一部として実装されることができる。

#### 【0031】

また、示されるのは、デジタルまたはアナログ処理を実行し、姿勢をセンサアセンブリ39からのジャイロスコープ、コンパス、または加速度計データから導出するように構成される、プロセッサ32である。プロセッサ32は、図2Aに示される、ローカル処理およびデータモジュール260の一部であってもよい。ウェアラブルシステム200はまた、図2Bに示されるように、例えば、GPS37(全地球測位システム)等の測位システムを含み、姿勢および測位分析を補助することができる。加えて、GPSはさらに、ユーザの環境についての遠隔ベース(例えば、クラウドベース)の情報を提供してもよい。本情報は、ユーザの環境内のオブジェクトまたは情報を認識するために使用されてもよい。

10

20

30

40

50

## 【 0 0 3 2 】

ウェアラブルシステムは、GPS 37および遠隔コンピューティングシステム（例えば、遠隔処理モジュール270、別のユーザのARD等）によって入手されたデータを組み合わせてもよく、これは、ユーザの環境についてのより多くの情報を提供することができる。一実施例として、ウェアラブルシステムは、GPSデータに基づいて、ユーザの場所を決定し、ユーザの場所と関連付けられた仮想オブジェクトを含む、世界マップを読み出すことができる（例えば、遠隔処理モジュール270と通信することによって）。別の実施例として、ウェアラブルシステム200は、世界カメラ16（図4に示される外向きに面した結像システム464の一部であってもよい）を使用して、環境を監視することができる。世界カメラ16によって入手された画像に基づいて、ウェアラブルシステム200は、環境内の表意文字を検出することができる（例えば、図7に示される1つ以上のオブジェクト認識装置708を使用することによって）。ウェアラブルシステムはさらに、GPS 37によって入手されたデータを使用して、表意文字を解釈することができる。例えば、ウェアラブルシステム200は、表意文字が位置する、地理的領域を識別し、地理的領域と関連付けられた1つ以上の言語を識別することができる。ウェアラブルシステムは、故に、識別された言語に基づいて、例えば、識別された言語と関連付けられた統語論、文法、文構造、スペル、句読点等に基づいて、表意文字を解釈することができる。一実施例では、ドイツ滞在ユーザ210は、アウトバーンを運転中、交通標識を知覚し得る。ウェアラブルシステム200は、ユーザ210がドイツに滞在しており、結像された交通標識からのテキストが、GPS 37から入手されたデータに基づいて（単独で、または世界カメラ16によって入手された画像と組み合わせて）、ドイツ語である可能性が高いことを識別することができる。

10

20

## 【 0 0 3 3 】

いくつかの状況では、世界カメラ16によって入手された画像は、ユーザの環境内のオブジェクトの不完全な情報を含み得る。例えば、画像は、霞のある大気、テキスト中の染みまたは誤植、低光量、ぼやけた画像、オクルージョン、世界カメラ16の限定FOV等に起因して、不完全なテキスト（例えば、文、表音文字、または語句）を含み得る。ウェアラブルシステム200は、GPS 37によって入手されたデータを画像内のテキストを認識する際のコンテキスト手掛かりとして使用し得る。

## 【 0 0 3 4 】

ウェアラブルシステム200はまた、レンダリングエンジン34を備えてもよく、これは、世界のユーザのビューのために、ユーザにローカルなレンダリング情報を提供し、スキャナの動作およびユーザの眼の中への結像を促進するように構成されることができる。レンダリングエンジン34は、ハードウェアプロセッサ（例えば、中央処理ユニットまたはグラフィック処理ユニット等）によって実装されてもよい。いくつかの実施形態では、レンダリングエンジンは、ローカル処理およびデータモジュール260の一部である。レンダリングエンジン34は、ウェアラブルシステム200の他のコンポーネントに通信可能に結合されることができる（例えば、有線または無線リンクを介して）。例えば、レンダリングエンジン34は、通信リンク102を介して、眼カメラ24に結合され、通信リンク104を介して、投影サブシステム18（網膜走査ディスプレイに類似する様式において、走査レーザ配列を介して、光をユーザの眼302、304の中に投影することができる）に結合されることができる。レンダリングエンジン34はまた、それぞれ、リンク105および94を介して、例えば、センサ姿勢プロセッサ32および画像姿勢プロセッサ36等の他の処理ユニットと通信することができる。

30

40

## 【 0 0 3 5 】

カメラ24（例えば、小型赤外線カメラ）は、眼姿勢を追跡し、レンダリングおよびユーザ入力をサポートするために利用されてもよい。いくつかの例示的眼姿勢は、ユーザが見ている場所または合焦させている深度（眼の輻輳・開散運動（vergence）を用いて推定されてもよい）を含んでもよい。GPS 37、ジャイロスコープ、コンパス、および加速度計39は、大まかなまたは高速姿勢推定を提供するために利用されてもよい。

50

カメラ16のうちの1つ以上のものは、画像および姿勢を入手することができ、これは、関連付けられたクラウドコンピューティングリソースからのデータと併せて、ローカル環境をマッピングし、ユーザビューを他者と共有するために利用されてもよい。

【0036】

図2Bに描写される例示的コンポーネントは、例証目的のためだけのものである。複数のセンサおよび他の機能モジュールが、例証および説明の容易性のために、ともに示される。いくつかの実施形態は、これらのセンサまたはモジュールの1つのみまたはサブセットを含んでもよい。さらに、これらのコンポーネントの場所は、図2Bに描写される位置に限定されない。いくつかのコンポーネントは、ベルト搭載型コンポーネント、ハンドヘルドコンポーネント、またはヘルメットコンポーネント等、他のコンポーネント内に搭載または格納されてもよい。一実施例として、画像姿勢プロセッサ36、センサ姿勢プロセッサ32、およびレンダリングエンジン34は、ベルトパック内に位置付けられ、超広帯域、Wi-Fi、Bluetooth（登録商標）等の無線通信を介して、または有線通信を介して、ウェアラブルシステムの他のコンポーネントと通信するように構成されてもよい。描写される筐体108は、好ましくは、ユーザによって頭部搭載可能かつ装着可能である。しかしながら、ウェアラブルシステム200のいくつかのコンポーネントは、ユーザの身体の他の部分に装着されてもよい。例えば、スピーカ240が、ユーザの耳の中に挿入され、音をユーザに提供してもよい。

【0037】

ユーザの眼302、304の中への光38の投影に関して、いくつかの実施形態では、カメラ24は、一般に、眼302、304の焦点の位置または「焦点深度」と一致する、ユーザの眼302、304の中心が幾何学的に輻輳される場所を測定するために利用されてもよい。眼が輻輳する全ての点の3次元表面は、「単視軌跡」と称され得る。焦点距離は、有限数の深度をとり得る、または無限に変動し得る。輻輳・開散運動距離から投影された光は、対象の眼302、304に集束されるように現れる一方、輻輳・開散運動距離の正面または背後の光は、ぼかされる。本開示のウェアラブルシステムおよび他のディスプレイシステムの実施例はまた、米国特許公開第2016/0270656号（参照することによってその全体として本明細書に組み込まれる）に説明される。

【0038】

ヒト視覚系は、複雑であって、深度の現実的知覚を提供することは、困難である。オブジェクトの視認者は、輻輳・開散運動移動と遠近調節（accommodation）の組み合わせに起因して、オブジェクトを3次元として知覚し得ると考えられる。相互に対する2つの眼の輻輳・開散運動移動（例えば、瞳孔が、相互に向かって、またはそこから離れるように移動し、眼の視線を収束させ、オブジェクトを固視するような瞳孔の回転）は、眼の水晶体の合焦（または「遠近調節」）と緊密に関連付けられる。通常条件下、焦点を1つのオブジェクトから異なる距離における別のオブジェクトに変化させるための眼のレンズの焦点の変化または眼の遠近調節は、「遠近調節 - 輻輳・開散運動反射」として知られる関係下、輻輳・開散運動の整合変化を自動的に同一距離に生じさせるであろう。同様に、輻輳・開散運動の変化は、通常条件下、遠近調節の整合変化を誘起するであろう。遠近調節と輻輳・開散運動との間のより良好な整合を提供するディスプレイシステムは、3次元画像のより現実的かつ快適なシミュレーションを形成し得る。

【0039】

さらに、約0.7ミリメートル未満のビーム直径を伴う、空間的にコヒーレントな光は、眼が合焦している場所にかかわらず、ヒトの眼によって正しく解決されることができると。したがって、適切な焦点深度の錯覚を作成するために、眼の輻輳・開散運動が、カメラ24を用いて追跡されてもよく、レンダリングエンジン34および投影サブシステム18は、単視軌跡上またはそれに近接する全てのオブジェクトを合焦させてレンダリングし、全ての他のオブジェクトを可変程度に焦点をずらしてレンダリングするために利用されてもよい（例えば、意図的に作成されたぼけを使用して）。好ましくは、システム220は、ユーザに、約60フレーム/秒以上のフレームレートでレンダリングする。上記に説明

10

20

30

40

50

されるように、好ましくは、カメラ24は、眼追跡のために利用されてもよく、ソフトウェアは、輻輳・開散運動幾何学形状だけではなく、また、ユーザ入力としての役割を果たすための焦点場所キューも取り上げるように構成されてもよい。好ましくは、そのようなディスプレイシステムは、昼間または夜間の使用のために好適な明度およびコントラストを用いて構成される。

#### 【0040】

いくつかの実施形態では、ディスプレイシステムは、好ましくは、視覚的オブジェクト整合のために約20ミリ秒未満の待ち時間、約0.1度未満の角度整合、および約1弧分の分解能を有し、これは、理論によって限定されるわけではないが、ヒトの眼のほぼ限界であると考えられる。ディスプレイシステム220は、位置特定システムと統合されてもよく、これは、GPS要素、光学追跡、コンパス、加速度計、または他のデータソースを伴い、位置および姿勢決定を補助し得る。位置特定情報は、関連世界のユーザのビュー内における正確なレンダリングを促進するために利用されてもよい（例えば、そのような情報は、眼鏡が実世界に対する場所を把握することを促進するであろう）。

10

#### 【0041】

いくつかの実施形態では、ウェアラブルシステム200は、ユーザの眼の遠近調節に基づいて、1つ以上の仮想画像を表示するように構成される。ユーザに画像が投影されている場所に合焦させるように強制する、従来の3Dディスプレイアプローチと異なり、いくつかの実施形態では、ウェアラブルシステムは、投影された仮想コンテンツの焦点を自動的に変動させ、ユーザに提示される1つ以上の画像のより快適な視認を可能にするように構成される。例えば、ユーザの眼が、1mの現在の焦点を有する場合、画像は、ユーザの焦点と一致するように投影されてもよい。ユーザが、焦点を3mに偏移させる場合、画像は、新しい焦点と一致するように投影される。したがって、ユーザに所定の焦点を強制するのではなく、いくつかの実施形態のウェアラブルシステム200は、ユーザの眼がより自然な様式において機能することを可能にする。

20

#### 【0042】

そのようなウェアラブルシステム200は、仮想現実デバイスに対して典型的に観察される、眼精疲労、頭痛、および他の生理学的症状の発生率を排除または低減させ得る。これを達成するために、ウェアラブルシステム200の種々の実施形態は、1つ以上の変焦点要素(VFE)を通して、仮想画像を可変焦点距離に投影するように構成される。1つ以上の実施形態では、3D知覚は、画像をユーザから固定された焦点面に投影する、多平面焦点システムを通して達成されてもよい。他の実施形態は、可変平面焦点を採用し、焦点面は、ユーザの焦点の現在の状態と一致するように、z-方向に往復して移動される。

30

#### 【0043】

多平面焦点システムおよび可変平面焦点システムの両方において、ウェアラブルシステム200は、眼追跡を採用し、ユーザの眼の輻輳・開散運動を決定し、ユーザの現在の焦点を決定し、仮想画像を決定された焦点に投影してもよい。他の実施形態では、ウェアラブルシステム200は、ファイバスキャナまたは他の光生成源を通して、網膜を横断して、可変焦点の光ビームをラスタパターンで可変に投影する、光変調器を備える。したがって、画像を可変焦点距離に投影するウェアラブルシステム200のディスプレイの能力は、ユーザがオブジェクトを3Dにおいて視認するための遠近調節を容易にするだけではなく、また、米国特許公開第2016/0270656号（参照することによってその全体として本明細書に組み込まれる）にさらに説明されるように、ユーザの眼球異常を補償するために使用されてもよい。いくつかの他の実施形態では、空間光変調器は、種々の光学コンポーネントを通して、画像をユーザに投影してもよい。例えば、以下にさらに説明されるように、空間光変調器は、画像を1つ以上の導波管上に投影してもよく、これは、次いで、画像をユーザに伝送する。

40

#### 【0044】

図3は、複数の深度平面を使用して3次元画像をシミュレートするためのアプローチの

50

側面を図示する。図3を参照すると、z-軸上の眼302および304からの種々の距離におけるオブジェクトは、それらのオブジェクトが合焦するように、眼302および304によって遠近調節される。眼302および304は、特定の遠近調節された状態をとり、オブジェクトをz-軸に沿った異なる距離に合焦させる。その結果、特定の遠近調節された状態は、特定の深度平面におけるオブジェクトまたはオブジェクトの一部が、眼がその深度平面に対して遠近調節された状態にあるとき、合焦するように、関連付けられた焦点距離を有する、深度平面306のうちの特定の1つと関連付けられると言え得る。いくつかの実施形態では、3次元画像は、眼302および304毎に、画像の異なる提示を提供することによって、また、深度平面のそれぞれに対応する画像の異なる提示を提供することによって、シミュレートされてもよい。例証を明確にするために、別個であるように示されるが、眼302および304の視野は、例えば、z-軸に沿った距離が増加するにつれて、重複し得ることを理解されたい。加えて、例証を容易にするために、平坦であるように示されるが、深度平面の等高線は、深度平面内の全ての特徴が特定の遠近調節された状態における眼と合焦するように、物理的空間内で湾曲され得ることを理解されたい。理論によって限定されるわけではないが、ヒトの眼は、典型的には、有限数の深度平面を解釈し、深度知覚を提供することができると考えられる。その結果、知覚された深度の高度に真実味のあるシミュレーションが、眼にこれらの限定数の深度平面のそれぞれに対応する画像の異なる提示を提供することによって達成され得る。

(導波管スタックアセンブリ)

【0045】

図4は、画像情報をユーザに出力するための導波管スタックの実施例を図示する。ウェアラブルシステム400は、複数の導波管432b、434b、436b、438b、4400bを使用して、3次元知覚を眼/脳に提供するために利用され得る、導波管のスタックまたはスタックされた導波管アセンブリ480を含む。いくつかの実施形態では、ウェアラブルシステム400は、図2Aのウェアラブルシステム200に対応してもよく、図4は、そのウェアラブルシステム200のいくつかの部分をもより詳細に概略的に示す。例えば、いくつかの実施形態では、導波管アセンブリ480は、図2Aのディスプレイ220の中に統合されてもよい。

【0046】

図4を継続して参照すると、導波管アセンブリ480はまた、複数の特徴458、456、454、452を導波管間に含んでもよい。いくつかの実施形態では、特徴458、456、454、452は、レンズであってもよい。他の実施形態では、特徴458、456、454、452は、レンズではなくてもよい。むしろ、それらは、単に、スペーサであってもよい(例えば、空気間隙を形成するためのクラディング層または構造)。

【0047】

導波管432b、434b、436b、438b、4400bまたは複数のレンズ458、456、454、452は、種々のレベルの波面曲率または光線発散を用いて、画像情報を眼に送信するように構成されてもよい。各導波管レベルは、特定の深度平面と関連付けられてもよく、その深度平面に対応する画像情報を出力するように構成されてもよい。画像投入デバイス420、422、424、426、428は、それぞれ、眼410に向かって出力するために、各個別の導波管を横断して入射光を分散させるように構成され得る、導波管440b、438b、436b、434b、432bの中に画像情報を投入するために利用されてもよい。光は、画像投入デバイス420、422、424、426、428の出力表面から出射し、導波管440b、438b、436b、434b、432bの対応する入力縁の中に投入される。いくつかの実施形態では、光の単一ビーム(例えば、コリメートされたビーム)が、各導波管の中に投入され、特定の導波管と関連付けられた深度平面に対応する特定の角度(および発散量)において眼410に向かって指向される、クローン化されたコリメートビームの場合全体を出力してもよい。

【0048】

いくつかの実施形態では、画像投入デバイス420、422、424、426、428

10

20

30

40

50

は、それぞれ、それぞれの対応する導波管440b、438b、436b、434b、432bの中への投入のための画像情報を生成する、離散ディスプレイである。いくつかの他の実施形態では、画像投入デバイス420、422、424、426、428は、例えば、画像情報を1つ以上の光学導管（光ファイバケーブル等）を介して、画像投入デバイス420、422、424、426、428のそれぞれに送り得る、単一の多重化されたディスプレイの出力端である。

#### 【0049】

コントローラ460が、スタックされた導波管アセンブリ480および画像投入デバイス420、422、424、426、428の動作を制御する。コントローラ460は、導波管440b、438b、436b、434b、432bへの画像情報のタイミングおよび提供を調整する、プログラミング（例えば、非一過性コンピュータ可読媒体内の命令）を含む。いくつかの実施形態では、コントローラ460は、単一体型デバイスまたは有線もしくは無線通信チャネルによって接続される分散型システムであってもよい。コントローラ460は、いくつかの実施形態では、処理モジュール260または270（図2に図示される）の一部であってもよい。

#### 【0050】

導波管440b、438b、436b、434b、432bは、全内部反射（TIR）によって各個別の導波管内で光を伝搬するように構成されてもよい。導波管440b、438b、436b、434b、432bはそれぞれ、主要な上部および底部表面ならびにそれらの主要上部表面と底部表面との間に延在する縁を伴う、平面である、または別の形状（例えば、湾曲）を有してもよい。図示される構成では、導波管440b、438b、436b、434b、432bはそれぞれ、光を再指向させ、各個別の導波管内で伝搬させ、導波管から、画像情報を眼410に出力することによって、光を導波管から抽出するように構成される、光抽出光学要素440a、438a、436a、434a、432aを含んでもよい。抽出された光はまた、外部結合光と称され得、光抽出光学要素はまた、外部結合光学要素と称され得る。抽出された光のビームは、導波管によって、導波管内を伝搬する光が光再指向要素に衝突する場所に出力される。光抽出光学要素（440a、438a、436a、434a、432a）は、例えば、反射または回折光学特徴であってもよい。説明の容易性および図面の明確性のために、導波管440b、438b、436b、434b、432bの底部主要表面に配置されて図示されるが、いくつかの実施形態では、光抽出光学要素440a、438a、436a、434a、432aは、上部または底部主要表面に配置されてもよい、または導波管440b、438b、436b、434b、432bの容積内に直接配置されてもよい。いくつかの実施形態では、光抽出光学要素440a、438a、436a、434a、432aは、透明基板に取り付けられ、導波管440b、438b、436b、434b、432bを形成する、材料の層内に形成されてもよい。いくつかの他の実施形態では、導波管440b、438b、436b、434b、432bは、材料のモノリシック部品であってもよく、光抽出光学要素440a、438a、436a、434a、432aは、材料のその部品の表面上および/またはその内部に形成されてもよい。

#### 【0051】

図4を継続して参照すると、本明細書に議論されるように、各導波管440b、438b、436b、434b、432bは、光を出力し、特定の深度平面に対応する画像を形成するように構成される。例えば、眼の最近傍の導波管432bは、そのような導波管432bの中に投入されるにつれて、コリメートされた光を眼410に送達するように構成されてもよい。コリメートされた光は、光学無限遠焦点面を表し得る。次の上方の導波管434bは、眼410に到達し得る前に、第1のレンズ452（例えば、負のレンズ）を通して通過する、コリメートされた光を送出するように構成されてもよい。第1のレンズ452は、眼/脳が、その次の上方の導波管434bから生じる光を光学無限遠から眼410に向かって内向きにより近い第1の焦点面から生じるように解釈するように、若干の凸面波面曲率を生成するように構成されてもよい。同様に、第3の上方の導波管436b

10

20

30

40

50

は、眼 4 1 0 に到達する前に、その出力光を第 1 のレンズ 4 5 2 および第 2 のレンズ 4 5 4 の両方を通して通過させる。第 1 および第 2 のレンズ 4 5 2 ならびに 4 5 4 の組み合わせられた屈折力は、眼 / 脳が、第 3 の導波管 4 3 6 b から生じる光が次の上方の導波管 4 3 4 b からの光であった光学無限遠から人物に向かって内向きにさらにより近い第 2 の焦点面から生じるように解釈するように、別の漸増量の波面曲率を生成するように構成されてもよい。

#### 【 0 0 5 2 】

他の導波管層（例えば、導波管 4 3 8 b、4 4 0 b）およびレンズ（例えば、レンズ 4 5 6、4 5 8）も同様に構成され、スタック内の最高導波管 4 4 0 b を用いて、人物に最も近い焦点面を表す集約焦点力のために、その出力をそれと眼との間のレンズの全てを通して送出する。スタックされた導波管アセンブリ 4 8 0 の他側の世界 4 7 0 から生じる光を視認 / 解釈するとき、レンズ 4 5 8、4 5 6、4 5 4、4 5 2 のスタックを補償するために、補償レンズ層 4 3 0 が、スタックの上部に配置され、下方のレンズスタック 4 5 8、4 5 6、4 5 4、4 5 2 の集約力を補償してもよい。そのような構成は、利用可能な導波管 / レンズ対と同じ数の知覚される焦点面を提供する。導波管の光抽出光学要素およびレンズの集束側面は両方とも、静的であってもよい（例えば、動的または電気活性ではない）。いくつかの代替実施形態では、一方または両方とも、電気活性特徴を使用して動的であってもよい。

#### 【 0 0 5 3 】

図 4 を継続して参照すると、光抽出光学要素 4 4 0 a、4 3 8 a、4 3 6 a、4 3 4 a、4 3 2 a は、光をその個別の導波管から再指向し、かつ導波管と関連付けられた特定の深度平面のための適切な発散またはコリメーション量を用いて本光を出力することの両方のために構成されてもよい。その結果、異なる関連付けられた深度平面を有する導波管は、関連付けられた深度平面に応じて、異なる量の発散を伴う光を出力する、異なる構成の光抽出光学要素を有してもよい。いくつかの実施形態では、本明細書に議論されるように、光抽出光学要素 4 4 0 a、4 3 8 a、4 3 6 a、4 3 4 a、4 3 2 a は、具体的角度において光を出力するように構成され得る、立体または表面特徴であってもよい。例えば、光抽出光学要素 4 4 0 a、4 3 8 a、4 3 6 a、4 3 4 a、4 3 2 a は、体積ホログラム、表面ホログラム、または回折格子であってもよい。回折格子等の光抽出光学要素は、2015年6月25日に公開された米国特許公開第 2015 / 0178939 号（参照することによってその全体として本明細書に組み込まれる）に説明される。

#### 【 0 0 5 4 】

いくつかの実施形態では、光抽出光学要素 4 4 0 a、4 3 8 a、4 3 6 a、4 3 4 a、4 3 2 a は、回折パターンを形成する回折特徴、すなわち、「回折光学要素」（本明細書では、「DOE」とも称される）である。好ましくは、DOE は、ビームの光の一部のみが、DOE の各交差点を用いて、眼 4 1 0 に向かって偏向される一方、残りが、全内部反射を介して、導波管を通して移動し続けるように、比較的到低回折効率を有する。画像情報を搬送する光は、したがって、複数の場所において導波管から出射する、いくつかの関連出射ビームに分割され、その結果、導波管内でバウンスする本特定のコレートされたビームに関して、眼 3 0 4 に向かって非常に均一パターンの出射放出となることができる。

#### 【 0 0 5 5 】

いくつかの実施形態では、1つ以上のDOEは、能動的に回折する「オン」状態と有意に回折しない「オフ」状態との間で切替可能であってもよい。例えば、切替可能なDOEは、ポリマー分散液晶の層を備えてもよく、その中で微小液滴は、ホスト媒体中に回折パターンを備え、微小液滴の屈折率は、ホスト材料の屈折率に実質的に整合するように切り替えられることができる（その場合、パターンは、入射光を著しく回折させない）、または微小液滴は、ホスト媒体のものに整合しない屈折率に切り替えられることができる（その場合、パターンは、入射光を能動的に回折させる）。

#### 【 0 0 5 6 】

10

20

30

40

50

いくつかの実施形態では、深度平面または被写界深度の数および分布は、視認者の眼の瞳孔サイズまたは配向に基づいて、動的に変動されてもよい。被写界深度は、視認者の瞳孔サイズと反比例して変化してもよい。その結果、視認者の眼の瞳孔のサイズが減少するにつれて、被写界深度は、その平面の場所が眼の焦点深度を越えるため判別不能である1つの平面が、判別可能となり、瞳孔サイズの低減および被写界深度の相当する増加に伴って、より合焦して現れ得るように増加する。同様に、異なる画像を視認者に提示するために使用される、離間される深度平面の数は、減少された瞳孔サイズに伴って減少されてもよい。例えば、視認者は、一方の深度平面から他方の深度平面への眼の遠近調節を調節せずに、第1の深度平面および第2の深度平面の両方の詳細を1つの瞳孔サイズにおいて明確に知覚することが可能ではない場合がある。しかしながら、これらの2つの深度平面は、同時に、遠近調節を変化させずに、別の瞳孔サイズにおいてユーザに合焦するには十分であり得る。

10

**【0057】**

いくつかの実施形態では、ディスプレイシステムは、瞳孔サイズまたは配向の決定に基づいて、もしくは特定の瞳孔サイズまたは配向を示す電気信号の受信に応じて、画像情報を受信する導波管の数を変動させてもよい。例えば、ユーザの眼が、2つの導波管と関連付けられた2つの深度平面間を区別不能である場合、コントローラ460（ローカル処理およびデータモジュール260の実施形態であり得る）は、これらの導波管のうちの1つへの画像情報の提供を停止するように構成またはプログラムされることができる。有利には、これは、システムへの処理負担を低減させ、それによって、システムの応答性を増加させ得る。導波管のためのDOEがオンおよびオフ状態間で切替可能である実施形態では、DOEは、導波管が画像情報を受信するとき、オフ状態に切り替えられてもよい。

20

**【0058】**

いくつかの実施形態では、出射ビームに視認者の眼の直径未満の直径を有するという条件を満たさせることが望ましくあり得る。しかしながら、本条件を満たすことは、視認者の瞳孔のサイズの変動性に照らして、困難であり得る。いくつかの実施形態では、本条件は、視認者の瞳孔のサイズの決定にตอบสนองして出射ビームのサイズを変動させることによって、広範囲の瞳孔サイズにわたって満たされる。例えば、瞳孔サイズが減少するにつれて、出射ビームのサイズもまた、減少し得る。いくつかの実施形態では、出射ビームサイズは、可変開口を使用して変動されてもよい。

30

**【0059】**

ウェアラブルシステム400は、世界470の一部を結像する、外向きに面した結像システム464（例えば、デジタルカメラ）を含むことができる。世界470の本部分は、世界カメラの視野（FOV）と称され得、結像システム464は、時として、FOVカメラとも称される。世界カメラのFOVは、視認者210のFOVと同一である場合とそうではない場合があり、これは、視認者210が所与の時間に知覚する、世界470の一部を包含する。例えば、いくつかの状況では、世界カメラのFOVは、ウェアラブルシステム400の視認者210の視野より大きくあり得る。視認者による視認または結像のために利用可能な領域全体は、動眼視野（FOR）と称され得る。FORは、装着者が、その身体、頭部、または眼を移動させ、空間内の実質的に任意の方向を知覚することができるため、ウェアラブルシステム400を圍繞する4ステラジアン（steradian）の立体角を含んでもよい。他のコンテキストでは、装着者の移動は、より抑制されてもよく、それに応じて、装着者のFORは、より小さい立体角に接し得る。外向きに面した結像システム464から得られた画像は、ユーザによって行われるジェスチャ（例えば、手または指のジェスチャ）を追跡し、ユーザの正面における世界470内のオブジェクトを検出する等のために、使用されることができる。

40

**【0060】**

ウェアラブルシステム400は、オーディオセンサ232、例えば、マイクロホンを含み、周囲音を捕捉することができる。上記に説明されるように、いくつかの実施形態では、1つ以上の他のオーディオセンサが、発話源の場所の決定に有用なステレオ音受信を提

50

供するために位置付けられることができる。オーディオセンサ 232 は、別の実施例として、指向性マイクロホンを備えることができ、これはまた、オーディオ源が位置する場所に関するそのような有用な指向性情報を提供することができる。

#### 【0061】

ウェアラブルシステム 400 はまた、眼移動および顔移動等のユーザの移動を観察する、内向きに面した結像システム 466 (例えば、デジタルカメラ)を含むことができる。内向きに面した結像システム 466 は、眼 410 の画像を捕捉し、眼 304 の瞳孔のサイズまたは配向を決定するために使用されてもよい。内向きに面した結像システム 466 は、ユーザが見ている方向(例えば、眼姿勢)を決定するために使用するため、またはユーザのバイオメトリック識別のため(例えば、虹彩識別を介して)、画像を得るために使用されることができる。いくつかの実施形態では、少なくとも 1つのカメラが、眼毎に、独立して、各眼の瞳孔サイズまたは眼姿勢を別個に決定し、それによって、各眼への画像情報の提示がその眼に対して動的に調整されることを可能にするために利用されてもよい。いくつかの他の実施形態では、単一眼 410 のみの瞳孔直径または配向(例えば、対の眼あたり単一カメラのみを使用して)が、決定され、ユーザの両眼に関して類似すると仮定される。内向きに面した結像システム 466 によって得られる画像は、ユーザに提示されるべきオーディオまたは視覚的コンテンツを決定するためにウェアラブルシステム 400 によって使用され得る、ユーザの眼姿勢または気分を決定するために分析されてもよい。ウェアラブルシステム 400 はまた、IMU、加速度計、ジャイロスコープ等のセンサを使用して、頭部姿勢(例えば、頭部位置または頭部配向)を決定してもよい。

#### 【0062】

ウェアラブルシステム 400 は、ユーザが、コマンドをコントローラ 460 に入力し、ウェアラブルシステム 400 と相互作用し得る、ユーザ入力デバイス 466 を含むことができる。例えば、ユーザ入力デバイス 466 は、トラックパッド、タッチスクリーン、ジョイスティック、多自由度(DOF)コントローラ、容量感知デバイス、ゲームコントローラ、キーボード、マウス、指向性パッド(Dパッド)、ワンド、触知デバイス、トータム(例えば、仮想ユーザ入力デバイスとして機能する)等を含むことができる。マルチDOFコントローラは、コントローラの一部または全部の可能性として考えられる平行移動(例えば、左/右、前方/後方、もしくは上/下)または回転(例えば、ヨー、ピッチ、もしくはロール)におけるユーザ入力を検知することができる。平行移動をサポートする、マルチDOFコントローラは、3DOFと称され得る一方、平行移動および回転をサポートする、マルチDOFコントローラは、6DOFと称され得る。ある場合には、ユーザは、指(例えば、親指)を使用して、タッチセンサ式入力デバイスを押下またはその上でスワイプし、入力をウェアラブルシステム 400 に提供してもよい(例えば、ユーザ入力をウェアラブルシステム 400 によって提供されるユーザインターフェースに提供するために)。ユーザ入力デバイス 466 は、ウェアラブルシステム 400 の使用の間、ユーザの手によって保持されてもよい。ユーザ入力デバイス 466 は、ウェアラブルシステム 400 と有線または無線通信することができる。

#### 【0063】

図5は、導波管によって出力された出射ビームの実施例を示す。1つの導波管が、図示されるが、導波管アセンブリ 480 内の他の導波管も、同様に機能してもよく、導波管アセンブリ 480 は、複数の導波管を含むことを理解されたい。光 520 が、導波管 432 b の入力縁 432 c において導波管 432 b の中に投入され、TIRによって導波管 432 b 内を伝搬する。光 520 がDOE 432 a に衝突する点において、光の一部が、出射ビーム 510 として導波管から出射する。出射ビーム 510 は、略平行として図示されるが、それらはまた、導波管 432 b と関連付けられた深度平面に応じて、ある角度で眼 410 に伝搬するように再指向されてもよい(例えば、発散出射ビーム形成)。略平行出射ビームは、光を外部結合し、眼 410 から遠距離(例えば、光学無限遠)における深度平面に設定されるように現れる画像を形成する光抽出光学要素を伴う、導波管を示し得ることを理解されたい。他の導波管または他の光抽出光学要素のセットは、より発散する、出

10

20

30

40

50

射ビームパターンを出力してもよく、これは、眼 4 1 0 がより近い距離に遠近調節し、網膜に合焦させることを要求し、光学無限遠より眼 4 1 0 に近い距離からの光として脳によって解釈されるであろう。

#### 【 0 0 6 4 】

図 6 は、導波管装置と、光を導波管装置へまたはそこから光学的に結合するための光学結合器サブシステムと、多焦点立体ディスプレイ、画像、またはライトフィールドの生成において使用される制御サブシステムとを含む、光学システムを示す、概略図である。光学システムは、導波管装置と、光を導波管装置にまたはそこから光学的に結合するための光学結合器サブシステムと、制御サブシステムとを含むことができる。光学システムは、多焦点立体、画像、またはライトフィールドを生成するために使用されることができる。光学システムは、1つ以上の一次平面導波管 6 3 2 a (1つのみが図 6 に示される)と、一次導波管 6 3 2 a の少なくともいくつかのそれぞれと関連付けられた1つ以上の D O E 6 3 2 b とを含むことができる。平面導波管 6 3 2 b は、図 4 を参照して議論される導波管 4 3 2 b、4 3 4 b、4 3 6 b、4 3 8 b、4 4 0 b に類似することができる。光学システムは、分散導波管装置を採用し、光を第 1 の軸 (図 6 の図では、垂直または Y - 軸) に沿って中継し、第 1 の軸 (例えば、Y - 軸) に沿って光の有効射出瞳を拡張させてもよい。分散導波管装置は、例えば、分散平面導波管 6 2 2 b と、分散平面導波管 6 2 2 b と関連付けられた少なくとも1つの D O E 6 2 2 a (二重破線によって図示される)とを含んでもよい。分散平面導波管 6 2 2 b は、少なくともいくつかの点において、それと異なる配向を有する一次平面導波管 6 3 2 b と類似または同じであってもよい。同様に、少なくとも1つの D O E 6 2 2 a は、少なくともいくつかの点において、D O E 6 3 2 a と類似または同じであってもよい。例えば、分散平面導波管 6 2 2 b または D O E 6 2 2 a は、それぞれ、一次平面導波管 6 3 2 b または D O E 6 3 2 a と同一材料から成ってもよい。図 6 に示される光学ディスプレイシステム 6 0 0 の実施形態は、図 2 A に示されるウェアラブルシステム 2 0 0 の中に統合されることができる。

#### 【 0 0 6 5 】

中継され、射出瞳が拡張された光は、分散導波管装置から1つ以上の一次平面導波管 6 3 2 b の中に光学的に結合され得る。一次平面導波管 6 3 2 b は、好ましくは、第 1 の軸に直交する、第 2 の軸 (例えば、図 6 の図では、水平または X - 軸) に沿って、光を中継することができる。着目すべきこととして、第 2 の軸は、第 1 の軸に対して非直交軸であることができる。一次平面導波管 6 3 2 b は、その第 2 の軸 (例えば、X - 軸) に沿って、光の有効射出瞳を拡張させる。例えば、分散平面導波管 6 2 2 b は、光を垂直または Y - 軸に沿って中継および拡張させ、光を水平または X - 軸に沿って中継および拡張させる、一次平面導波管 6 3 2 b にその光を通過させることができる。

#### 【 0 0 6 6 】

光学システムは、単一モード光ファイバ 6 4 0 の近位端の中に光学的に結合され得る、1つ以上の着色光源 (例えば、赤色、緑色、および青色レーザ光) 6 1 0 を含んでもよい。光ファイバ 6 4 0 の遠位端は、圧電材料の中空管 6 4 2 を通して螺合または受容されてもよい。遠位端は、固定されない可撓性カンチレバー 6 4 4 として、管 6 4 2 から突出する。圧電管 6 4 2 は、4つの象限電極 (図示せず) と関連付けられることができる。電極は、例えば、管 6 4 2 の外側、外側表面もしくは外側周縁、または直径に鍍着されてもよい。コア電極 (図示せず) もまた、管 6 4 2 のコア、中心、内側周縁、または内径に位置してもよい。

#### 【 0 0 6 7 】

例えば、ワイヤ 6 6 0 を介して電氣的に結合される、駆動電子機器 6 5 0 は、対向する対の電極を駆動し、圧電管 6 4 2 を独立して2つの軸において屈曲させる。光ファイバ 6 4 4 の突出する遠位先端は、機械的共鳴モードを有する。共鳴の周波数は、光ファイバ 6 4 4 の直径、長さ、および材料性質に依存し得る。圧電管 6 4 2 をファイバカンチレバー 6 4 4 の第 1 の機械的共鳴モードの近傍で振動させることによって、ファイバカンチレバー 6 4 4 は、振動させられ、大偏向を通して掃引し得る。

## 【 0 0 6 8 】

2つの軸において共振振動を刺激することによって、ファイバカンチレバー644の先端は、2次元(2-D)走査を充填する面積内において2軸方向に走査される。光源610の強度をファイバカンチレバー644の走査と同期して変調させることによって、ファイバカンチレバー644から発せられる光は、画像を形成することができる。そのような設定の説明は、米国特許公開第2014/0003762号(参照することによってその全体として本明細書に組み込まれる)に提供されている。

## 【 0 0 6 9 】

光学結合器サブシステムのコンポーネントは、走査ファイバカンチレバー644から発せられる光をコリメートすることができる。コリメートされた光は、鏡面表面648によって、少なくとも1つの回折光学要素(DOE)622aを含有する、狭分散平面導波管622bの中に反射されることができる。コリメートされた光は、TIRによって分散平面導波管622bに沿って(図6の図に対して)垂直に伝搬し、そうすることによって、DOE622aと繰り返し交差することができる。DOE622aは、好ましくは、低回折効率を有する。これは、光の一部(例えば、10%)をDOE622aとの交差点の各点においてより大きい一次平面導波管632bの縁に向かって回折させ、光の一部をTIRを介して分散平面導波管622bの長さを辿ってそのオリジナル軌道上で継続させることができる。

## 【 0 0 7 0 】

DOE622aとの交差点の各点において、付加的光が、一次導波管632bの入口に向かって回折されることができる。入射光を複数の外部結合セットに分割することによって、光の射出瞳は、分散平面導波管622b内のDOE622aによって垂直に拡張されることができる。分散平面導波管622bから外部結合された本垂直に拡張された光は、一次平面導波管632bの縁に進入することができる。

## 【 0 0 7 1 】

一次導波管632bに進入する光は、TIRを介して、一次導波管632bに沿って(図6の図に対して)水平に伝搬することができる。光は、複数の点においてDOE632aと交差するにつれて、TIRを介して、一次導波管632bの長さの少なくとも一部に沿って水平に伝搬する。DOE632aは、有利には、線形回折パターンおよび半径方向対称回折パターンの総和である、位相プロファイルを有し、光の偏向および集束の両方を生成するように設計または構成され得る。DOE632aは、有利には、ビームの光の一部のみが、DOE632aの各交差点において視認者の眼に向かって偏向される一方、光の残りが、TIRを介して、一次導波管632bを通して伝搬し続けるように、低回折効率(例えば、10%)を有し得る。

## 【 0 0 7 2 】

伝搬する光とDOE632aとの間の交差点の各点において、光の一部は、一次導波管632bの隣接面に向かって回折され、光がTIRから逃散し、一次導波管632bの面から発せられることを可能にする。いくつかの実施形態では、DOE632aの半径方向対称回折パターンは、加えて、ある焦点レベルを回折された光に付与し、個々のビームの光波面を成形(例えば、曲率を付与する)し、かつビームを設計される焦点レベルに合致する角度に操向することの両方を行う。

## 【 0 0 7 3 】

故に、これらの異なる経路は、異なる角度におけるDOE632aの多重度、焦点レベル、または射出瞳において異なる充填パターンをもたらすことによって、光を一次平面導波管632bの外部で結合させることができる。射出瞳における異なる充填パターンは、有利には、複数の深度平面を伴うライトフィールドディスプレイを生成するために使用されることができる。導波管アセンブリ内の各層またはスタック内の層のセット(例えば、3層)が、個別の色(例えば、赤色、青色、緑色)を生成するために採用されてもよい。したがって、例えば、第1の3つの隣接する層のセットが、それぞれ、赤色、青色、および緑色光を第1の焦点深度において生成するために採用されてもよい。第2の3つの隣接

10

20

30

40

50

する層のセットが、それぞれ、赤色、青色、および緑色光を第2の焦点深度において生成するために採用されてもよい。複数のセットが、種々の焦点深度を伴うフル3Dまたは4Dカラー画像ライトフィールドを生成するために採用されてもよい。

(ウェアラブルシステムの他のコンポーネント)

【0074】

多くの実装では、ウェアラブルシステムは、上記に説明されるウェアラブルシステムのコンポーネントに加えて、またはその代替として、他のコンポーネントを含んでもよい。ウェアラブルシステムは、例えば、1つ以上の触知デバイスまたはコンポーネントを含んでもよい。触知デバイスまたはコンポーネントは、触覚をユーザに提供するように動作可能であってもよい。例えば、触知デバイスまたはコンポーネントは、仮想コンテンツ(例えば、仮想オブジェクト、仮想ツール、他の仮想構造)に触れると、圧力またはテクスチャの感覚を提供してもよい。触覚は、仮想オブジェクトが表す物理的オブジェクトの感覚を再現してもよい、または仮想コンテンツが表す想像上のオブジェクトもしくはキャラクタ(例えば、ドラゴン)の感覚を再現してもよい。いくつかの実装では、触知デバイスまたはコンポーネントは、ユーザによって装着されてもよい(例えば、ユーザウェアラブルグローブ)。いくつかの実装では、触知デバイスまたはコンポーネントは、ユーザによって保持されてもよい。

【0075】

ウェアラブルシステムは、例えば、ユーザによって操作可能であって、ウェアラブルシステムへの入力またはそれとの相互作用を可能にする、1つ以上の物理的オブジェクトを含んでもよい。これらの物理的オブジェクトは、本明細書では、トーテムと称され得る。いくつかのトーテムは、例えば、金属またはプラスチック片、壁、テーブルの表面等、無生物オブジェクトの形態をとってもよい。ある実装では、トーテムは、実際には、任意の物理的入力構造(例えば、キー、トリガ、ジョイスティック、トラックボール、ロッカスイッチ)を有していなくてもよい。代わりに、トーテムは、単に、物理的表面を提供してもよく、ウェアラブルシステムは、ユーザにトーテムの1つ以上の表面上にあるように見えるように、ユーザインターフェースをレンダリングしてもよい。例えば、ウェアラブルシステムは、トーテムの1つ以上の表面上に常駐するように見えるように、コンピュータキーボードおよびトラックパッドの画像をレンダリングしてもよい。例えば、ウェアラブルシステムは、トーテムとしての役割を果たす、アルミニウムの薄い長方形プレートの表面上に見えるように、仮想コンピュータキーボードおよび仮想トラックパッドをレンダリングしてもよい。長方形プレート自体は、任意の物理的キーまたはトラックパッドもしくはセンサを有していない。しかしながら、ウェアラブルシステムは、仮想キーボードまたは仮想トラックパッドを介して行われた選択または入力として、長方形プレートを用いたユーザ操作または相互作用もしくはタッチを検出し得る。ユーザ入力デバイス466(図4に示される)は、トラックパッド、タッチパッド、トリガ、ジョイスティック、トラックボール、ロッカもしくは仮想スイッチ、マウス、キーボード、多自由度コントローラ、または別の物理的入力デバイスを含み得る、トーテムの実施形態であってもよい。ユーザは、単独で、または姿勢と組み合わせて、トーテムを使用し、ウェアラブルシステムまたは他のユーザと相互作用してもよい。

【0076】

本開示のウェアラブルデバイス、HMD、およびディスプレイシステムと使用可能な触知デバイスおよびトーテムの実施例は、米国特許公開第2015/0016777号(参照することによってその全体として本明細書に組み込まれる)に説明されている。

(例示的ウェアラブルシステム、環境、およびインターフェース)

【0077】

ウェアラブルシステムは、高被写界深度をレンダリングされたライトフィールド内で達成するために、種々のマッピング関連技法を採用してもよい。仮想世界をマッピングする際、実世界内の全ての特徴および点を把握し、仮想オブジェクトを実世界に関連して正確に描くことが有利である。この目的を達成するために、ウェアラブルシステムのユーザか

10

20

30

40

50

ら捕捉されたF O V画像が、実世界の種々の点および特徴についての情報を伝達する新しい写真を含むことによって、世界モデルに追加されることができる。例えば、ウェアラブルシステムは、マップ点(2 D点または3 D点等)のセットを収集し、新しいマップ点を見出し、世界モデルのより正確なバージョンをレンダリングすることができる。第1のユーザの世界モデルは、第2のユーザが第1のユーザを囲繞する世界を体験し得るように、(例えば、クラウドネットワーク等のネットワークを経由して)第2のユーザに通信されることができる。

【0078】

図7は、MR環境700の実施例のブロック図である。MR環境700は、入力(例えば、ユーザのウェアラブルシステムからの視覚的入力702、室内カメラ等の定常入力704、種々のセンサからの感覚入力706、ユーザ入力デバイス466からのジェスチャ、トーテム、眼追跡、ユーザ入力等)を1つ以上のユーザウェアラブルシステム(例えば、ウェアラブルシステム200もしくはディスプレイシステム220)または定常室内システム(例えば、室内カメラ等)から受信するように構成されてもよい。ウェアラブルシステムは、種々のセンサ(例えば、加速度計、ジャイロ스코ープ、温度センサ、移動センサ、深度センサ、GPSセンサ、内向きに面した結像システム、外向きに面した結像システム等)を使用して、ユーザの環境の場所および種々の他の属性を決定することができる。本情報はさらに、異なる視点からの画像または種々のキューを提供し得る、部屋内の定常カメラからの情報で補完されてもよい。カメラ(室内カメラおよび/または外向きに面した結像システムのカメラ等)によって入手された画像データは、マッピング点のセットに低減されてもよい。

【0079】

1つ以上のオブジェクト認識装置708が、受信されたデータ(例えば、点の集合)を通してクローリングし、点を認識またはマッピングし、画像をタグ付けし、マップデータベース710を用いて、意味論情報をオブジェクトに結び付けることができる。マップデータベース710は、経時的に収集された種々の点およびその対応するオブジェクトを備えてもよい。種々のデバイスおよびマップデータベースは、ネットワーク(例えば、LAN、WAN等)を通して相互に接続され、クラウドにアクセスすることができる。

【0080】

本情報およびマップデータベース内の点集合に基づいて、オブジェクト認識装置708a-708nは、環境内のオブジェクトを認識してもよい。例えば、オブジェクト認識装置は、顔、人物、窓、壁、ユーザ入力デバイス、テレビ、ドキュメント(例えば、本明細書におけるセキュリティ実施例において説明されるような旅券、運転免許証、パスポート)、ユーザの環境内の他のオブジェクト等を認識することができる。1つ以上のオブジェクト認識装置が、ある特性を伴うオブジェクトのために特殊化されてもよい。例えば、オブジェクト認識装置708aは、顔を認識するために使用されてもよい一方、別のオブジェクト認識装置は、ドキュメントを認識するために使用されてもよい。

【0081】

オブジェクト認識は、種々のコンピュータビジョン技法を使用して実施されてもよい。例えば、ウェアラブルシステムは、外向きに面した結像システム464(図4に示される)によって入手された画像を分析し、場面再構成、イベント検出、ビデオ追跡、オブジェクト認識(例えば、人物またはドキュメント)、オブジェクト姿勢推定、顔認識(例えば、環境内の人物またはドキュメント上の画像から)、学習、インデックス化、運動推定、または画像分析(例えば、写真、署名、識別情報、旅行情報等のドキュメント内の印を識別する)等を実施することができる。1つ以上のコンピュータビジョンアルゴリズムが、これらのタスクを実施するために使用されてもよい。コンピュータビジョンアルゴリズムの非限定的実施例は、スケール不変特徴変換(SIFT)、スピードアップロバスト特徴(SURF)、配向FASTおよび回転BRIBEF(ORB)、バイナリロバスト不変スケラブルキーポイント(BRISK)、高速網膜キーポイント(FREAK)、Viola-Jonesアルゴリズム、Eigenfacesアプローチ、Lucas-Kan

10

20

30

40

50

a d e アルゴリズム、H o r n - S c h u n k アルゴリズム、M e a n - s h i f t アルゴリズム、視覚的同時位置推定およびマッピング ( v S L A M ) 技法、シーケンシャルベイズ推定器 ( 例えば、カルマンフィルタ、拡張カルマンフィルタ等 )、バンドル調節、適応閾値化 ( および他の閾値化技法 )、反復最近傍点 ( I C P )、セミグローバルマッチング ( S G M )、セミグローバルブロックマッチング ( S G B M )、特徴点ヒストグラム、種々の機械学習アルゴリズム ( 例えば、サポートベクトルマシン、k 最近傍アルゴリズム、単純ベイズ、ニューラルネットワーク ( 畳み込みまたは深層ニューラルネットワークを含む )、または他の教師あり / 教師なしモデル等 ) 等を含む。

#### 【 0 0 8 2 】

1 つ以上のオブジェクト認識装置 7 0 8 はまた、種々のテキスト認識アルゴリズムを実装し、テキストを画像から識別および抽出することができる。いくつかの例示的テキスト認識アルゴリズムは、光学文字認識 ( O C R ) アルゴリズム、深層学習アルゴリズム ( 深層ニューラルネットワーク等 )、パターンマッチングアルゴリズム、事前処理のためのアルゴリズム等を含む。

#### 【 0 0 8 3 】

オブジェクト認識は、加えて、または代替として、種々の機械学習アルゴリズムによって実施されることができる。いったん訓練されると、機械学習アルゴリズムは、H M D によって記憶されることができる。機械学習アルゴリズムのいくつかの実施例は、教師ありまたは教師なし機械学習アルゴリズムを含むことができ、回帰アルゴリズム ( 例えば、通常の最小 2 乗回帰等 )、インスタンススペースのアルゴリズム ( 例えば、学習ベクトル量子化等 )、決定ツリーアルゴリズム ( 例えば、分類および回帰ツリー等 )、ベイズアルゴリズム ( 例えば、単純ベイズ等 )、クラスタリングアルゴリズム ( 例えば、k - 平均クラスタリング等 )、関連付けルール学習アルゴリズム ( 例えば、アプリアルゴリズム等 )、人工ニューラルネットワークアルゴリズム ( 例えば、P e r c e p t r o n 等 )、深層学習アルゴリズム ( 例えば、D e e p B o l t z m a n n M a c h i n e、すなわち、深層ニューラルネットワーク等 )、次元削減アルゴリズム ( 例えば、主成分分析等 )、アンサンブルアルゴリズム ( 例えば、S t a c k e d G n e r a l i z a t i o n 等 )、または他の機械学習アルゴリズムを含む。いくつかの実施形態では、個々のモデルは、個々のデータセットのためにカスタマイズされることができる。例えば、ウェアラブルデバイスは、ベースモデルを生成または記憶することができる。ベースモデルは、開始点として使用され、データタイプ ( 例えば、テレプレゼンスセッション内の特定のユーザ )、データセット ( 例えば、テレプレゼンスセッション内のユーザの取得される付加的画像のセット )、条件付き状況、または他の変形例に特有の付加的モデルを生成してもよい。いくつかの実施形態では、ウェアラブル H M D は、複数の技法を利用して、集約されたデータの分析のためのモデルを生成するように構成されることができる。他の技法は、事前に定義された閾値またはデータ値を使用することを含んでもよい。

#### 【 0 0 8 4 】

マップデータベース内の本情報および点の集合に基づいて、オブジェクト認識装置 7 0 8 a - 7 0 8 n は、オブジェクトを認識し、オブジェクトを意味論情報で補完し、生命をオブジェクトに与えてもよい。例えば、オブジェクト認識装置が、点のセットがドアであることを認識する場合、システムは、いくつかの意味論情報を結び付けてもよい ( 例えば、ドアは、ヒンジを有し、ヒンジを中心として 9 0 度移動を有する )。オブジェクト認識装置が、点のセットが鏡であることを認識する場合、システムは、鏡が、部屋内のオブジェクトの画像を反射させ得る、反射表面を有するという意味論情報を結び付けてもよい。意味論情報は、本明細書に説明されるように、オブジェクトのアフォーダンスを含むことができる。例えば、意味論情報は、オブジェクトの法線を含んでもよい。システムは、ベクトルを割り当てることができ、その方向は、オブジェクトの法線を示す。経時的に、マップデータベースは、システム ( ローカルに常駐し得る、または無線ネットワークを通してアクセス可能であり得る ) がより多くのデータを世界から蓄積するにつれて成長する。いったんオブジェクトが認識されると、情報は、1 つ以上のウェアラブルシステムに伝送

10

20

30

40

50

されてもよい。例えば、MR環境700は、Californiaで生成している場面についての情報を含んでもよい。環境700は、New Yorkにおける1人以上のユーザに伝送されてもよい。FOVカメラおよび他の入力から受信されたデータに基づいて、オブジェクト認識装置および他のソフトウェアコンポーネントは、場面が世界の異なる部分に存在し得る第2のユーザに正確に「パス」され得るように、種々の画像から収集された点をマッピングし、オブジェクトを認識すること等ができる。環境700はまた、位置特定目的のために、トポロジマップを使用してもよい。

#### 【0085】

図8は、認識されたオブジェクトに関連して仮想コンテンツをレンダリングする方法800の実施例のプロセスフロー図である。方法800は、仮想場面がウェアラブルシステムのユーザに提示され得る方法を説明する。ユーザは、その場面から地理的に遠隔に存在してもよい。例えば、ユーザは、New Yorkに存在し得るが、Californiaで現在起こっている場面を視認することを所望し得る、またはCaliforniaに存在する友人と散歩に行くことを所望し得る。

#### 【0086】

ブロック810では、ウェアラブルシステムは、ユーザの環境に関する入力をユーザおよび他のユーザから受信してもよい。これは、種々の入力デバイスおよびマップデータベース内にすでに保有されている知識を通して達成されてもよい。ユーザのFOVカメラ、センサ、GPS、眼追跡等が、ブロック810において、情報をシステムに伝達する。システムは、ブロック820において、本情報に基づいて、疎点を決定してもよい。疎点は、ユーザの周囲における種々のオブジェクトの配向および位置を表示ならびに理解する際に使用され得る、姿勢データ（例えば、頭部姿勢、眼姿勢、身体姿勢、または手のジェスチャ）を決定する際に使用されてもよい。オブジェクト認識装置708a、708nは、ブロック830において、これらの収集された点を通してクローリングし、マップデータベースを使用して、1つ以上のオブジェクトを認識してもよい。本情報は、次いで、ブロック840において、ユーザの個々のウェアラブルシステムに伝達されてもよく、所望の仮想場面が、ブロック850において、適宜、ユーザに表示されてもよい。例えば、所望の仮想場面（例えば、CAにおけるユーザ）が、New Yorkにおけるユーザの種々のオブジェクトおよび他の周囲に関連して、適切な配向、位置等において表示されてもよい。

#### 【0087】

図9は、ウェアラブルシステムの別の実施例のブロック図である。本実施例では、ウェアラブルシステム900は、世界に関するマップデータを含み得る、マップ920を備える、マップデータベース710を含んでもよい。マップは、部分的に、ウェアラブルシステム上にローカルに常駐してもよく、部分的に、有線または無線ネットワークによってアクセス可能なネットワーク化された記憶場所（例えば、クラウドシステム内）に常駐してもよい。姿勢プロセス910が、ウェアラブルコンピューティングアーキテクチャ（例えば、処理モジュール260またはコントローラ460）上で実行され、ウェアラブルコンピューティングハードウェアまたはユーザの位置および配向を決定するために、マップ920からのデータを利用してもよい。姿勢データは、ユーザが、システムを体験し、その世界内で動作するにつれて、オンザフライで収集されたデータから算出されてもよい。データは、実または仮想環境内のオブジェクトに関する画像、センサ（概して、加速度計およびジャイロスコープコンポーネントを備える、慣性測定ユニット等）からのデータ、および表面情報を備えてもよい。

#### 【0088】

疎点表現は、同時位置特定およびマッピング（例えば、入力が画像/視覚のみである構成を指す、SLAMまたはvSLAM）プロセスの出力であってもよい。システムは、世界内の種々のコンポーネントの場所だけではなく、世界が構成される内容も見出すように構成されることができる。姿勢は、マップへの取込およびマップからのデータを使用することを含め、多くの目標を達成する、構築ブロックであり得る。

10

20

30

40

50

## 【0089】

一実施形態では、疎点位置は、それ自体では完全に適正であり得ず、さらなる情報が、多焦点AR、VR、またはMR体験を生成するために必要とされ得る。概して、深度マップ情報を指す、稠密表現が、少なくとも部分的に、本間隙を充填するために利用されてもよい。そのような情報は、立体視940と称されるプロセスから算出されてもよく、深度情報は、三角測量または飛行時間感知等の技法を使用して決定される。画像情報およびアクティブパターン（アクティブプロジェクタを使用して生成される赤外線パターン等）、画像カメラから入手された画像、または手ジェスチャ/トーテム950が、立体視プロセス940への入力としての役割を果たし得る。有意な量の深度マップ情報が、ともに融合されてもよく、このうちのいくつかは、表面表現を用いて要約されてもよい。例えば、数学的に定義可能な表面は、ゲームエンジンのような他の処理デバイスへの効率的（例えば、大規模点群に対して）かつ要約しやすい入力であってもよい。したがって、立体視プロセス（例えば、深度マップ）940の出力は、融合プロセス930において組み合わせられてもよい。姿勢950は、同様に、本融合プロセス930への入力であってもよく、融合930の出力は、マップ取込プロセス920への入力となる。サブ表面が、トポグラフィマッピング等において相互に接続し、より大きい表面を形成し得、マップは、点および表面の大規模ハイブリッドとなる。

10

## 【0090】

複合現実プロセス960における種々の側面を解決するために、種々の入力が、利用されてもよい。例えば、図9に描写される実施形態では、ゲームパラメータは、システムのユーザが1匹以上のモンスターと種々の場所においてモンスターバトルゲームをプレーしていること、モンスターが死んでいる、または種々の条件下で逃げている（ユーザがモンスターを撃つ場合等）、種々の場所における壁または他のオブジェクト、および同等物を決定するための入力であってもよい。世界マップは、複合現実に対する別の有用な入力となる、そのようなオブジェクトが相互に対して存在する場所に関する情報を含んでもよい。世界に対する姿勢は、同様に、入力となり、ほぼあらゆる双方向システムに対して重要な役割を果たす。

20

## 【0091】

ユーザからの制御または入力は、ウェアラブルシステム900への別の入力である。本明細書に説明されるように、ユーザ入力は、視覚的入力、ジェスチャ、トーテム、オーディオ入力、感覚入力等を含むことができる。動き回るまたはゲームをプレーするために、例えば、ユーザは、ウェアラブルシステム900に、所望する対象に関して命令する必要がある。空間内で自ら移動するだけでなく、利用され得る種々の形態のユーザ制御が、存在する。一実施形態では、トーテム（例えば、ユーザ入力デバイス）、または玩具銃等のオブジェクトが、ユーザによって保持され、システムによって追跡されてもよい。システムは、好ましくは、ユーザがアイテムを保持していることを把握し、ユーザがアイテムと行っている相互作用の種類を理解するように構成されるであろう（例えば、トーテムまたはオブジェクトが、銃である場合、システムは、場所および配向だけではなく、ユーザが、そのようなアクティビティがカメラのいずれかの視野内にないときでも、生じている状況の決定を補助し得る、IMU等のセンサを装備し得る、トリガまたは他の感知ボタンもしくは要素をクリックしているかどうかを理解するように構成されてもよい。）

30

40

## 【0092】

手のジェスチャ追跡または認識もまた、入力情報を提供してもよい。ウェアラブルシステム900は、ボタン押下のため、左または右、停止、握持、保持等をジェスチャするために、手のジェスチャを追跡および解釈するように構成されてもよい。例えば、1つの構成では、ユーザは、非ゲーム環境において電子メールもしくはカレンダーを通して捲る、または別の人物もしくはプレーヤと「フィストバンプ」を行うことを所望し得る。ウェアラブルシステム900は、動的である場合とそうではない場合がある、最小量の手のジェスチャを活用するように構成されてもよい。例えば、ジェスチャは、停止を示すために手を広げる、OKを示すために親指を上げる、OKではないことを示すために親指を下げる、

50

または指向性コマンドを示すために左右もしくは上下に手をフリップする等、単純な静的ジェスチャであってもよい。手のジェスチャ追跡は、ジェスチャを行い、手話を用いて通信する他者等のユーザの環境内の他者によって行われるジェスチャの追跡を含むことができる（例えば、図 1 3 A 参照）。

#### 【 0 0 9 3 】

眼追跡は、別の入力である（例えば、ユーザが見ている場所を追跡し、ディスプレイ技術を制御し、具体的深度または範囲においてレンダリングする）。一実施形態では、眼の輻輳・開散運動が、三角測量を使用して決定されてもよく、次いで、その特定の人物のために開発された輻輳・開散運動/遠近調節モデルを使用して、遠近調節が、決定されてもよい。眼追跡は、眼カメラによって実施され、眼視線（例えば、片眼または両眼の方向もしくはは配向）を決定することができる。他の技法も、例えば、眼の近傍に設置された電極による電位の測定（例えば、電気眼球図記録）等、眼追跡のために使用されることができる。

10

#### 【 0 0 9 4 】

発話追跡は、単独で、または他の入力（例えば、トーテム追跡、眼追跡、ジェスチャ追跡等）と組み合わせて使用され得る、別の入力であり得る。発話追跡は、単独で、または組み合わせて、発話認識、音声認識を含んでもよい。システム 9 0 0 は、オーディオストリームを環境から受信する、オーディオセンサ（例えば、マイクロホン）を含むことができる。システム 9 0 0 は、発話している人物（例えば、発話が A R D の装着者または別の人物もしくはは音声（例えば、環境内のラウドスピーカによって伝送される記録された音声）からのものであるかどうか）を決定するための音声認識技術ならびに言われていることを決定するための発話認識技術を組み込むことができる。ローカルデータおよび処理モジュール 2 6 0 または遠隔処理モジュール 2 7 0 は、マイクロホンからのオーディオデータ（または、例えば、ユーザによって鑑賞されているビデオストリーム等の別のストリーム内のオーディオデータ）を処理し、例えば、隠れマルコフモデル、動的時間伸縮法（D T W）ベースの発話認識、ニューラルネットワーク、ディープフィードフォワードおよび再帰ニューラルネットワーク等の深層学習アルゴリズム、エンドツーエンド自動発話認識、機械学習アルゴリズム（図 7 を参照して説明される）、もしくは音響モデル化または言語モデル化等を使用する、他のアルゴリズム等の種々の発話認識アルゴリズムを適用することによって、発話のコンテンツを認識することができる。

20

30

#### 【 0 0 9 5 】

複合現実プロセス 9 6 0 への別の入力は、環境内の表記を追跡するステップを含むことができる。表記は、商業用または公共看板を含むことができる。図 1 6 A - 1 9 を参照して説明されるように、システムは、表記を認識し、表記内のテキストを識別し、テキストの特性を調節し（例えば、テキストのフォントサイズを増加させ、可読性を改良する）、テキストのコンテンツを修正する（例えば、テキストを外国言語からユーザによって理解される言語に翻訳する）等を行うことができる。

#### 【 0 0 9 6 】

ローカルデータおよび処理モジュール 2 6 0 または遠隔処理モジュール 2 7 0 はまた、音声認識アルゴリズムを適用することができる、これは、話者がウェアラブルシステム 9 0 0 のユーザ 2 1 0 またはユーザが会話している別の人物であるかどうか等の話者の識別を識別することができる。いくつかの例示的音声認識アルゴリズムは、頻度推定、隠れマルコフモデル、ガウス混合モデル、パターンマッチングアルゴリズム、ニューラルネットワーク、マトリクス表現、ベクトル量子化、話者ダイアライゼーション、決定ツリー、および動的時間伸縮（D T W）技法を含むことができる。音声認識技法はまた、コホートモデルおよび世界モデル等のアンチ話者技法を含むことができる。スペクトル特徴は、話者特性を表す際に使用されてもよい。ローカルデータおよび処理モジュールまたは遠隔データ処理モジュール 2 7 0 は、図 7 を参照して説明される種々の機械学習アルゴリズムを使用して、音声認識を実施することができる。

40

#### 【 0 0 9 7 】

50

システム900はまた、他の人々または環境とのユーザの相互作用を促進するための感覚式アイウェアシステム970を含むことができる。感覚式アイウェアシステム970の実装は、UIを介して、これらのユーザ制御または入力を使用することができる。UI要素（例えば、制御、ポップアップウィンドウ、吹き出し、データエントリフィールド等）が、例えば、情報、例えば、変換されたテキスト、グラフィック、または補助情報の表示を放逐するかまたは補助情報の表示を要求するために使用される。UIは、感覚式アイウェアシステム970が、手話において会話相手によって行われる身ぶりを翻訳する際に使用するための言語を把握するように、ユーザが、ユーザが理解する1つ以上の言語のリストを入力することを可能にすることができる。そのような実装およびこれらの使用の実施例が、以下にさらに説明される。

10

**【0098】**

感覚式アイウェアシステム970はまた、テキスト認識、修正、およびレンダリング特徴を備えることができる。そのような特徴は、ウェアラブルシステムの種々の他のコンポーネントと組み合わせて、環境とのユーザの相互作用を向上させ得る。例えば、HMDは、ユーザの物理的環境の画像から識別されたテキストに基づいて、画像をディスプレイ上に投影するように（例えば、投影された画像が物理的環境からのオリジナルテキストをオクルードするように）構成される、1つ以上の光源11を含むことができる。光学的に透過性の接眼レンズ106は、1つ以上の光源11からの光をユーザ210に画像として伝送するように構成されることができる。画像は、HMDシステム200が画像を表示し得る、多くの可能性として考えられる深度のうちの1つにすぎない、特定の深度にあるかのように現れ得る。HMDシステム100は、異なる深度平面306上にあるかのように現れ得る、いくつかの異なる深度に現れるように画像を投影可能であり得る（図3参照）。接眼レンズ106が光学的に透過性である、いくつかの実施形態では、接眼レンズ106は、環境からの光がユーザの眼に入射することを可能にすることができる。したがって、そのような実施形態では、ユーザ210には、環境からの画像の一部が、1つ以上の光源11からの投影された画像とともに見え得る。

20

**【0099】**

カメラシステムに関して、図9に示される例示的ウェアラブルシステム900は、3つの対のカメラ、すなわち、ユーザの顔の両側に配列される相対的広FOVまたは受動SLAM対のカメラと、ユーザの正面に配向され、立体視結像プロセス940をハンドリングし、また、ユーザの顔の正面の手のジェスチャおよびトーテム/オブジェクトの軌道捕捉するための異なる対のカメラとを含むことができる。立体視プロセス940のためのFOVカメラまたは対のカメラはまた、カメラ16と称され得る。立体視プロセス940のためのFOVカメラおよび対のカメラは、外向きに面した結像システム464（図4に示される）の一部であってもよい。ウェアラブルシステム900は、眼ベクトルおよび他の情報を三角測量するために、ユーザの眼に向かって配向される、眼追跡カメラ（また、眼カメラ24としても示され、図4に示される内向きに面した結像システム462の一部であってもよい）を含むことができる。ウェアラブルシステム900はまた、1つ以上のテクスチャ化光プロジェクタ（赤外線（IR）プロジェクタ等）を備え、テクスチャを場面の中に投入してもよい。

30

40

**【0100】**

図10は、ウェアラブルシステムへのユーザ入力を決定するための方法1000の実施例のプロセスフロー図である。本実施例では、ユーザは、トーテムと相互作用してもよい。ユーザは、複数のトーテムを有してもよい。例えば、ユーザは、ソーシャルメディアアプリケーションのための指定される1つのトーテム、ゲームをプレーするための別のトーテム等を有してもよい。ブロック1010では、ウェアラブルシステムは、トーテムの運動を検出してもよい。トーテムの移動は、外向きに面した結像システムを通して認識されてもよい、またはセンサ（例えば、触知グローブ、画像センサ、手追跡デバイス、眼追跡カメラ、頭部姿勢センサ等）を通して検出されてもよい。

**【0101】**

50

少なくとも部分的に、検出されたジェスチャ、眼姿勢、頭部姿勢、またはトーテムを通じた入力に基づいて、ウェアラブルシステムは、ブロック 1020 において、基準フレームに対するトーテム（またはユーザの眼もしくは頭部もしくはジェスチャ）の位置、配向、または移動を検出する。基準フレームは、それに基づいてウェアラブルシステムがトーテム（またはユーザ）の移動をアクションまたはコマンドに変換する、マップ点のセットであってもよい。ブロック 1030 では、トーテムとのユーザの相互作用が、マッピングされる。基準フレーム 1020 に対するユーザ相互作用のマッピングに基づいて、システムは、ブロック 1040 において、ユーザ入力を決定する。

#### 【0102】

例えば、ユーザは、トーテムまたは物理的オブジェクトを前後に移動させ、仮想ページを捲り、次のページに移動する、または 1 つのユーザインターフェース（UI）ディスプレイ画面から別の UI 画面に移動することを示してもよい。別の実施例として、ユーザは、ユーザの FOR 内の異なる実または仮想オブジェクトを見るために、その頭部または眼を移動させてもよい。特定の実または仮想オブジェクトにおけるユーザの注視が、閾値時間より長い場合、その実または仮想オブジェクトは、ユーザ入力として選択されてもよい。いくつかの実装では、ユーザの眼の輻輳・開散運動が、追跡されることができ、遠近調節／輻輳・開散運動モデルが、ユーザが合焦している深度平面に関する情報を提供する、ユーザの眼の遠近調節状態を決定するために使用されることができ、いくつかの実装では、ウェアラブルシステムは、レイキャスティング技法を使用して、ユーザの頭部姿勢または眼姿勢の方向に沿っている実または仮想オブジェクトを決定することができる。種々の実装では、レイキャスティング技法は、実質的に殆ど横幅を伴わない細い光線束を投じる、または実質的横幅を伴う光線（例えば、円錐または円錐台）を投じることを含むことができる。

#### 【0103】

ユーザインターフェースは、本明細書に説明されるようなディスプレイシステム（図 2A におけるディスプレイ 220 等）によって投影されてもよい。また、1 つ以上のプロジェクタ等の種々の他の技法を使用して表示されてもよい。プロジェクタは、画像をキャンバスまたは球体等の物理的オブジェクト上に投影してもよい。ユーザインターフェースとの相互作用は、システムの外部またはシステムの一部の 1 つ以上のカメラを使用して（例えば、内向きに面した結像システム 462 または外向きに面した結像システム 464 を使用して）追跡されてもよい。

#### 【0104】

図 11 は、仮想ユーザインターフェースと相互作用するための方法 1100 の実施例のプロセスフロー図である。方法 1100 は、本明細書に説明されるウェアラブルシステムによって行われてもよい。方法 1100 の実施形態は、ウェアラブルシステムの FOV 内の人物またはドキュメントを検出するためにウェアラブルシステムによって使用されることができ、

#### 【0105】

ブロック 1110 では、ウェアラブルシステムは、特定の UI を識別してもよい。UI のタイプは、ユーザによって与えられてもよい。ウェアラブルシステムは、特定の UI がユーザ入力（例えば、ジェスチャ、視覚的データ、オーディオデータ、感覚データ、直接コマンド等）に基づいて取り込まれる必要があることを識別してもよい。UI は、セキュリティシナリオに特有であることができ、システムの装着者は、ドキュメントを装着者に提示するユーザを観察する（例えば、旅客検問所において）。ブロック 1120 では、ウェアラブルシステムは、仮想 UI のためのデータを生成してもよい。例えば、UI の境界、一般的構造、形状等と関連付けられたデータが、生成されてもよい。加えて、ウェアラブルシステムは、ウェアラブルシステムがユーザの物理的場所に関連して UI を表示し得るように、ユーザの物理的場所のマップ座標を決定してもよい。例えば、UI が、身体中心である場合、ウェアラブルシステムは、リング UI がユーザの周囲に表示され得る、または平面 UI が壁上またはユーザの正面に表示され得るように、ユーザの物理的立ち位置

10

20

30

40

50

、頭部姿勢、または眼姿勢の座標を決定してもよい。本明細書に説明されるセキュリティコンテキストでは、UIは、装着者が、旅行者および旅行者のドキュメントを見ている間、UIを容易に視認し得るように、UIがドキュメントをシステムの装着者に提示する旅行者を囲繞しているかのように表示され得る。UIが、手中心の場合、ユーザの手のマップ座標が、決定されてもよい。これらのマップ点は、FOVカメラ、感覚入力を通して受信されたデータ、または任意の他のタイプの収集されたデータを通して導出されてもよい。

#### 【0106】

ブロック1130では、ウェアラブルシステムは、データをクラウドからディスプレイに送信してもよい、またはデータは、ローカルデータベースからディスプレイコンポーネントに送信されてもよい。ブロック1140では、UIは、送信されたデータに基づいて、ユーザに表示される。例えば、ライトフィールドディスプレイは、仮想UIをユーザの眼の一方または両方の中に投影することができる。いったん仮想UIが生成されると、ウェアラブルシステムは、ブロック1150において、単に、ユーザからのコマンドを待機し、より多くの仮想コンテンツを仮想UI上に生成してもよい。例えば、UIは、ユーザの身体またはユーザの環境内の人物（例えば、旅行者）の身体の周囲の身体中心リングであってもよい。ウェアラブルシステムは、次いで、コマンド（ジェスチャ、頭部または眼移動、音声コマンド、ユーザ入力デバイスからの入力等）を待機してもよく、認識される場合（ブロック1160）、コマンドと関連付けられた仮想コンテンツが、ユーザに表示されてもよい（ブロック1170）。

#### 【0107】

ウェアラブルシステム、UI、およびユーザ体験（UX）の付加的実施例は、米国特許公開第2015/0016777号（参照することによってその全体として本明細書に組み込まれる）に説明されている。

（複数のウェアラブルシステム間の例示的通信）

#### 【0108】

図12は、相互に相互作用する複数のユーザデバイスを描写する、全体的システム図を図式的に図示する。コンピューティング環境1200は、ユーザデバイス1230a、1230b、1230cを含む。ユーザデバイス1230a、1230b、および1230cは、ネットワーク1290を通して、相互に通信することができる。ユーザデバイス1230a - 1230cはそれぞれ、ネットワークインターフェースを含み、ネットワーク1290を介して、遠隔コンピューティングシステム1220（また、ネットワークインターフェース1271を含んでもよい）と通信することができる。ネットワーク1290は、LAN、WAN、ピアツーピアネットワーク、無線、Bluetooth（登録商標）、または任意の他のネットワークであってもよい。コンピューティング環境1200はまた、1つ以上の遠隔コンピューティングシステム1220を含むことができる。遠隔コンピューティングシステム1220は、クラスタ化され、異なる地理的場所に位置する、サーバコンピュータシステムを含んでもよい。ユーザデバイス1230a、1230b、および1230cは、ネットワーク1290を介して、遠隔コンピューティングシステム1220と通信してもよい。

#### 【0109】

遠隔コンピューティングシステム1220は、遠隔データリポジトリ1280を含んでもよく、これは、具体的ユーザの物理または仮想世界についての情報を維持することができる。データ記憶装置1280は、手話辞書、補助情報源等、感覚式アイウェアに有用な情報を含むことができる。遠隔データリポジトリは、図2Aに示される遠隔データリポジトリ280の実施形態であってもよい。遠隔コンピューティングシステム1220はまた、遠隔処理モジュール1270を含んでもよい。遠隔処理モジュール1270は、図2Aに示される遠隔処理モジュール270の実施形態であってもよい。いくつかの実装では、遠隔コンピューティングシステム1220は、ウェアラブルシステム200と無関係の第三者システムであってもよい。

10

20

30

40

50

## 【0110】

遠隔処理モジュール1270は、1つ以上のプロセッサを含んでもよく、これは、ユーザデバイス(1230a、1230b、1230c)および遠隔データリポジトリ1280と通信することができる。プロセッサは、ユーザデバイスおよび他の源から取得される情報を処理することができる。いくつかの実装では、処理または記憶の少なくとも一部は、ローカル処理およびデータモジュール260(図2Aに示されるように)によって提供されることができる。遠隔コンピューティングシステム1220は、所与のユーザが、具体的ユーザ自身の物理的または仮想世界についての情報を別のユーザと共有することを可能にしてもよい。

## 【0111】

ユーザデバイスは、単独で、または組み合わせて、ウェアラブルデバイス(HMDもしくはAR等)、コンピュータ、モバイルデバイス、または任意の他のデバイスであってもよい。例えば、ユーザデバイス1230bおよび1230cは、図2Aに示されるウェアラブルシステム200(または図4に示されるウェアラブルシステム400)の実施形態であってもよく、これは、AR/VR/MRコンテンツを提示するように構成されることができる。

## 【0112】

ユーザデバイスのうちの1つ以上のものは、図4に示されるユーザ入力デバイス466と併用されることができる。ユーザデバイスは、ユーザおよびユーザの環境についての情報を取得することができる(例えば、図4に示される外向きに面した結像システム464を使用して)。ユーザデバイスまたは遠隔コンピューティングシステム1220は、ユーザデバイスから取得される情報を使用して、画像、点、および他の情報の集合を構築、更新、および建造することができる。例えば、ユーザデバイスは、入手された未加工情報を処理し、さらなる処理のために、処理された情報を遠隔コンピューティングシステム1220に送信してもよい。ユーザデバイスはまた、処理のために、未加工情報を遠隔コンピューティングシステム1220に送信してもよい。ユーザデバイスは、処理された情報を遠隔コンピューティングシステム1220から受信し、ユーザに投影させる前に、最終処理を提供してもよい。ユーザデバイスはまた、取得された情報を処理し、処理された情報を他のユーザデバイスに渡してもよい。ユーザデバイスは、入手された情報を処理しながら、遠隔データリポジトリ1280と通信してもよい。複数のユーザデバイスまたは複数のサーバコンピュータシステムが、入手された画像の構築もしくは処理に関与してもよい。

## 【0113】

物理的世界に関する情報は、経時的に展開されてもよく、異なるユーザデバイスによって収集される情報に基づいてもよい。仮想世界のモデルはまた、経時的に展開され、異なるユーザの入力に基づいてもよい。そのような情報およびモデルは、時として、本明細書では、世界マップまたは世界モデルと称され得る。図7および9を参照して説明されるように、ユーザデバイスによって入手された情報は、世界マップ1210を構築するために使用されてもよい。世界マップ1210は、図9に説明されるマップ920の少なくとも一部を含んでもよい。種々のオブジェクト認識装置(例えば、708a、708b、708c...708n)が、オブジェクトおよびタグ画像を認識するため、ならびに意味論情報をオブジェクトに付加するために使用されてもよい。これらのオブジェクト認識装置はまた、図7に説明される。

## 【0114】

遠隔データリポジトリ1280は、データを記憶し、世界マップ1210の構造を促進するために使用されることができる。ユーザデバイスは、ユーザの環境についての情報を常に更新し、世界マップ1210についての情報を受信することができる。世界マップ1210は、ユーザまたは別の人物によって作成されてもよい。本明細書に議論されるように、ユーザデバイス(例えば、1230a、1230b、1230c)および遠隔コンピューティングシステム1220は、単独で、または組み合わせて、世界マップ1210を

10

20

30

40

50

構築および/または更新してもよい。例えば、ユーザデバイスは、遠隔処理モジュール 1270 および遠隔データリポジトリ 1280 と通信してもよい。ユーザデバイスは、ユーザおよびユーザの環境についての情報を入手または処理してもよい。遠隔処理モジュール 1270 は、遠隔データリポジトリ 1280 およびユーザデバイス（例えば、1230a、1230b、1230c）と通信し、ユーザおよびユーザの環境についての情報を処理してもよい。遠隔コンピューティングシステム 1220 は、例えば、ユーザの画像を選択的にクロッピングする、ユーザの背景を修正する、仮想オブジェクトをユーザの環境に追加する、ユーザの発話に補助情報で注釈を付ける等、ユーザデバイス（例えば、1230a、1230b、1230c）によって入手された情報を修正することができる。遠隔コンピューティングシステム 1220 は、処理された情報を同一または異なるユーザデバイスに送信することができる。

10

**【0115】**

感覚式アイウェアシステムの実施形態の種々の機能性が、以下にさらに説明される。  
(ユーザ相互作用を促進するための例示的感覚式アイウェア)

**【0116】**

ウェアラブルシステム 200 は、他の人々または環境とのユーザの相互作用を促進するための感覚式アイウェアシステム 970 を実装することができる。他の人々と相互作用する一実施例として、ウェアラブルシステム 200 は、例えば、手話を構成し得るジェスチャを検出し、手話を別の言語（例えば、別の手話または発話された言語）に翻訳し、翻訳された情報をウェアラブルシステムのユーザに提示することによって、手話を解釈することができる。別の実施例として、感覚式アイウェアシステム 970 は、発話を手話に翻訳し、手話をユーザに提示することができる。

20

**【0117】**

ウェアラブルシステム 970 はまた、環境内のオブジェクトを認識し、（仮想環境内の）オブジェクトの特性を修正し、修正されたオブジェクトを仮想オブジェクトとしてユーザに提示することによって、環境とのユーザの相互作用を促進することができる。例えば、ウェアラブルシステム 200 は、外向きに面した結像システム 464 によって入手された画像に基づいて、ユーザの環境内の標識（例えば、交通標識、店舗正面の標識等）を認識し、ユーザの環境内の標識の特性を修正し、修正された標識をユーザに提示することができる。修正された標識は、オリジナル標識がオクルードされ得るように、ユーザの 3D 環境上にオーバーレイされてもよい。

30

(対人通信のためのツールとしての例示的感覚式アイウェアシステム)

**【0118】**

いくつかの状況では、会話内の 1 人以上の人々は、手または身体ジェスチャ（例えば、手話等）を使用して、自身を表し得る。会話は、テレプレゼンスセッションの間または人々が相互に物理的近傍に居るときに生じ得る。ウェアラブルシステム 200 は、ユーザが手話者と通信するとき、ウェアラブルシステム 200 のユーザ（観察者とも称される）のために、手話者の手話を解釈することができる。ウェアラブルシステム 200 はまた、手話者が観察者の発話を理解し得るように、口頭または手話ベースの発話をグラフィック（例えば、手のジェスチャの画像等）に翻訳し、グラフィックを手話者に提示することができる。例えば、頭部搭載型ディスプレイを装着している観察者は、低減された視野を有し得、したがって、観察者は、手話を使用して手話者によって行われる完全ジェスチャを観察不可能である場合がある。ウェアラブルシステム 200 は、外向きに面した結像システム 464 を使用して、手話者によるジェスチャを捕捉することができる（ユーザが頭部搭載型ディスプレイを通して知覚することができるものより広い視野を有する、カメラを有し得るため）。ウェアラブルシステム 200 は、捕捉されたジェスチャを仮想グラフィックとして観察者に示し、または捕捉されたジェスチャから変換されたテキストを示し、観察者の手話者の発話の理解を促進することができる。さらに、ウェアラブルシステム 200 は、1 つの手話を別の手話に翻訳するように構成されることができる。例えば、会話内の 1 人の人物は、アメリカ手話を使用し得、他の人物は、ドゴン手話を使用し得る。ウエ

40

50

アラブルシステム 200 は、ドゴン手話を使用する人物のために、アメリカ手話をドゴン手話に翻訳し、アメリカ手話を使用する人物のために、ドゴン手話をアメリカ手話に翻訳することができる。

(例示的手話捕捉)

【0119】

ウェアラブルシステムは、種々の技法を使用して、オリジナル発話を捕捉し、オリジナル発話を標的発話に翻訳することができる。発話は、手または身体ジェスチャもしくは可聴音の形態であってもよい。本明細書に説明されるように、オリジナル発話は、手話であってもよく、標的発話は、別の手話または発話された言語であってもよい。代替として、オリジナル発話は、発話された言語であってもよい一方、標的発話は、手話である。ウェアラブルシステム 200 は、発話のコンテキスト（例えば、発話が対人または電気通信を介したものであるかどうか）に応じて、外向きに面した結像システム 464、オーディオセンサ 232 を使用して、またはネットワーク 1290 を介して、別のコンピューティングデバイスと通信することによって、オリジナル発話を捕捉することができる。

10

【0120】

検出された手話の手話者が感覚式アイウェアシステムの物理的近傍に居る、対人通信の間のオリジナル発話を捕捉する実施例として、外向きに面した結像システム 464 は、ユーザの環境の画像を捕捉することができる。ウェアラブルシステム 200 は、画像情報から、手話を構成し得る、ジェスチャ（例えば、手/身体ジェスチャまたは唇移動）を検出することができる。ウェアラブルシステム 200 は、例えば、深層ニューラルネットワーク、隠れマルコフモデル、動的プログラミングマッチング等のアルゴリズムを使用して、ジェスチャを認識し、話者によって行われるジェスチャによって表される身ぶりを認識することができる。図 7 を参照して説明されるように、ジェスチャ認識は、1 つ以上のオブジェクト認識装置 708 によって実施されてもよい。

20

【0121】

遠隔通信のコンテキストにおいてオリジナル発話を捕捉する実施例として、ウェアラブルシステム 200 は、遠隔コンピューティングデバイス（例えば、別のウェアラブルシステム）から受信されたデータを分析することによって、または外向きに面した結像システム 464 によって捕捉されたデータ（例えば、オリジナル発話がテレビ上に存在する場合）を分析することによって、オリジナル発話の存在を捕捉および認識することができる。一実施例では、手話者および観察者は、インターネットビデオチャットセッションを通して、会話してもよい。手話者および観察者はそれぞれ、その個別の HMD を装着することができる。HMD は、ネットワーク 1290 を介して、相互に通信することができる（図 12 に示される）。手話者が、反射表面（例えば、鏡）の正面に居る場合、手話者の HMD は、外向きに面した結像システム 464 を介して、手話者の反射された画像を入手することによって、手話者のジェスチャを捕捉することができる。手話者の反射された画像は、手話の認識および解釈のために、観察者の HMD または遠隔コンピューティングシステム 1220 に送信されてもよい。別の実施例として、手話者は、テレビまたはインターネットプログラミング等上に提示されるもの等、ビデオプログラム内の人物であってもよい。手話者のジェスチャが、観察者の場所において視覚的に観察され得る場合、ウェアラブルシステム 464 は、対人通信コンテキストと同一方法において（例えば、オーディオセンサ 232 または外向きに面した結像システム 464 を介して）、手話ジェスチャを捕捉することができる。

30

40

【0122】

別の人物によってジェスチャされる手話のテキストまたはグラフィカル翻訳をウェアラブルシステム 200 のユーザに表示することに加え、またはその代替として、ウェアラブルシステムのユーザ 200 はまた、手話をを用いて通信してもよい。この場合、ウェアラブルシステムは、外向きに面した結像システム 464 によって、ユーザ自身の手話ジェスチャ（一人称視点から）を捕捉することができる。ウェアラブルシステムは、手話を、テキスト、オーディオ、画像等のフォーマットで表され得る、標的発話に変換することができ

50

る。ウェアラブルシステム200は、別のユーザへの提示のために、結果を別のウェアラブルシステムに伝送することができる。本明細書に説明されるように、オリジナル発話から標的発話への変換は、単独で、または組み合わせて、ユーザのウェアラブルシステム、別のユーザのウェアラブルシステム、または遠隔コンピューティングシステム1220によって実施されることができる。例えば、ユーザのウェアラブルシステムは、ユーザの手のジェスチャを捕捉し、捕捉されたビデオまたは画像（手話ジェスチャを含有する）を別のユーザのウェアラブルシステムまたは遠隔コンピューティングシステム120に伝送することができる。これは、手話をビデオまたは画像から抽出し、手話を発話言語または別の手話のための視聴覚コンテンツに変換することができる。視聴覚コンテンツは、テキスト、グラフィック、ビデオ、動画、音等を含むことができる。

10

（手話者ジェスチャ除外および源位置特定）

#### 【0123】

ウェアラブルシステムは、例えば、オーディオセンサ232、外向きに面した結像システム464、定常入力704、またはユーザの環境内の他のセンサ等の種々のセンサを使用して、ジェスチャまたは手話源を識別することができる。一実施例として、ウェアラブルシステムは、一連の手のジェスチャならびに唇移動を外向きに面した結像システム464によって入手されたデータから検出してもよい。ウェアラブルシステムは、手話者もまた対応する唇移動を有するため、手のジェスチャが手話者と関連付けられることを見出し得る。別の実施例として、ウェアラブルシステムは、ユーザとジェスチャとの間の距離を測定し、ジェスチャ源を決定することができる。例えば、ウェアラブルシステムは、手が外向きに面した結像システム464によって入手された画像内で比較的大きく現れるため、一連のジェスチャがユーザから生じることを決定することができる。しかし、手が、比較的小さく現れる場合、ウェアラブルシステムは、ジェスチャがユーザ以外の人物から生じることを見出し得る。さらに別の実施例として、ウェアラブルシステムは、視聴覚コンテンツを再生しているオブジェクトを認識することによって（例えば、オブジェクト認識装置708を使用して、テレビを認識することによって）、ジェスチャが視聴覚コンテンツ（例えば、テレビ内）から生じていることを見出し得る。

20

#### 【0124】

ジェスチャ源に基づいて、ウェアラブルシステム200は、ある人々からのジェスチャを処理しないように構成されることができる。例えば、ウェアラブルシステムは、ジェスチャをユーザの環境内の複数の人々から捕捉してもよいが、ウェアラブルシステムは、手話認識のために、ユーザのFOVの中心外の人物からの手話を処理しないように構成されることができる。別の実施例として、ウェアラブルシステムは、ユーザ自身の手話を処理しないように構成されてもよい。

30

#### 【0125】

いくつかの実施形態では、ウェアラブルシステムは、例えば、外向きに面した結像システム464がユーザの手のジェスチャを捕捉するために、ユーザがその手を挙げる必要がないような角度において、カメラを外向きに面した結像システム464内に位置付けること等によって、ユーザ自身の手話を検出するようにセンサを構成することができる。センサはまた、ユーザ自身の手話を検出しないように構成されることができる。例えば、非検出は、ユーザ自身の手の方向（典型的には、ユーザのFOVの下方）における画像を捕捉しない、またはそのような方向における画像をフィルタリング除去する（例えば、クロッピングによって）ことを通して達成されることができる。したがって、システムは、ユーザ自身の手話と他者のものを区別することができる。

40

（手話からテキストへの例示的変換）

#### 【0126】

ウェアラブルシステム200は、捕捉された手話を、ユーザに提示または別の言語に翻訳され得る、テキストに変換することができる。手話からテキストへの変換は、深層学習（深層ニューラルネットワークを利用してよい）、隠れマルコフモデル、動的プログラミングマッチング等のアルゴリズムを使用して実施されることができる。例えば、深層学

50

習方法（ある場合には、畳み込みニューラルネットワーク）は、身ぶりを表す特徴を決定し、学習された特徴に基づいて、分類モデルを構築するように、既知の身ぶりを含有する画像またはビデオ上で訓練されることができる（教師あり学習）。そのような訓練された深層学習方法は、次いで、ローカル処理およびデータモジュール260または遠隔処理モジュールおよびウェアラブルシステム200のデータリポジトリ270、280によって、外向きに面した結像サブシステムによって検出された手話者の画像に適用されることができる。

#### 【0127】

テキスト変換機能性は、単独で、または組み合わせて、ローカル処理およびデータモジュール260、遠隔処理モジュール270、遠隔データリポジトリ280、または遠隔コンピューティングシステム1220によって実装されることができる。例えば、ウェアラブルシステム200は、HMD上に実装される手話/テキスト機能性を含むことができる。一実施例として、ウェアラブルシステムは、手話辞書をローカルデータモジュール260または遠隔データリポジトリ280内に記憶することができる。ウェアラブルシステムは、故に、手話辞書にアクセスし、検出されたジェスチャをテキストに翻訳することができる。別の実施例として、ウェアラブルシステム200は、遠隔コンピューティングシステム1220によって実装される手話/テキスト機能性にアクセスすることができる。ウェアラブルシステム200は、商業用手話/テキストサービスまたはデータリポジトリとの無線接続を利用してよい（例えば、アプリケーションプログラミングインターフェース（API）を介して）。例えば、ウェアラブルシステム200は、捕捉されたジェスチャを遠隔コンピューティングシステム1220に提供し、対応するテキストを遠隔コンピューティングシステム1220から受信することができる。

#### 【0128】

変換がローカルまたは遠隔で実施されるかどうかにかかわらず、変換されたテキストを表示する、補助情報（以下にさらに説明される）を読み出す等の他の処理ステップは、テキスト変換が実施される場所から独立して、ローカルまたは遠隔で行われてもよい。例えば、手話/テキスト変換が、遠隔で行われ、変換されたテキストが、ローカルで表示されるべき（例えば、システムのユーザが、観察者である）場合、捕捉されたビデオストリームは、ネットワークを介して変換を実施する、遠隔処理モジュール270または遠隔サーバに送信されることができる。変換されたテキスト文字列は、表示のために、システムのローカルコンポーネント（例えば、ローカル処理およびデータモジュール260）に返される。別の実施例として、手話/テキスト変換および補助情報読出が、遠隔で行われる場合、捕捉されたビデオストリームは、ネットワークを介して、遠隔処理モジュール270または遠隔サーバに送信されることができ、読み出された補助情報は、システムのローカルコンポーネントに返されることができる。ローカル/遠隔処理の他の組み合わせも、実行可能である。

#### 【0129】

これらの実施例は、身ぶりをテキストに変換することを参照して説明されるが、身ぶりは、例えば、グラフィック、動画、オーディオ、または他のタイプの視聴覚コンテンツ等の種々の他のフォーマットに変換されてもよい。さらに、身ぶりの翻訳は、身ぶりが最初にテキストに翻訳されることを要求しない。

（1つの手話を別の手話に変換する実施例）

#### 【0130】

本明細書に記載されるように、数百もの手話が、世界中に存在する。故に、本明細書に説明されるウェアラブルシステムはまた、会話相手の両方が、身ぶりを行なっているが、異なる手話システムであるときに使用されることができる。有利には、各そのような手話者は、その独自のウェアラブルシステムを使用して、他の手話者の身ぶりをユーザ自身の手話システムに翻訳することができる。ウェアラブルシステムは、身ぶりをユーザによって理解されるテキストまたはユーザ自身の手話のグラフィック表現に翻訳してもよい。

#### 【0131】

ウェアラブルシステム200は、特定の手話、例えば、アメリカ手話(ASL)を認識するように構成されてもよい。ウェアラブルシステム200はまた、複数の手話、例えば、ASL、イギリス手話、中国手話、ドゴン手話等を認識するように構成されてもよい。いくつかの実装では、ウェアラブルシステム200は、例えば、感覚式アイウェアシステムの場所情報に基づいて、手話認識の再構成をサポートする。ウェアラブルシステムは、単独で、または手話辞書と組み合わせて、例えば、オブジェクト認識装置708を利用して、外向きに面した結像システム464によって知覚されたジェスチャを認識する、システムがユーザ自身のまたは好ましい手話を認識する方法に類似する手段を通して、外国語手話を認識してもよい。ウェアラブルシステムは、ユーザが知覚する手話をユーザの優勢手話に変換することができる。ユーザの優勢手話は、会話におけるユーザの第1の手話またはユーザの好ましい手話であることができる。ユーザの優勢手話以外の手話は、外国語手話と見なされ得る。ウェアラブルシステムは、ユーザが、外国語手話の変換されたテキストを選択することを可能にすることができる。例えば、ユーザは、外国語手話を選択することができ、ウェアラブルシステムは、外国語手話におけるジェスチャの意味をテキストとしてウェアラブルシステムのユーザに提示することができる。

10

**【0132】**

ウェアラブルシステムは、環境内で発話された言語または場所情報の補助を通して、外国語手話を認識してもよい。例えば、ウェアラブルシステムは、イタリア語がユーザの環境内で発話されたことを検出する、またはGPSによって入手されたデータに基づいて、ユーザがイタリアに滞在していることを決定する。本情報に基づいて、ウェアラブルシステムは、イタリア手話を認識するための機能を自動的にアクティブ化することができる。別の実施例として、ウェアラブルシステムは、ウェアラブルシステムがサポート可能な手話の選好の順序を有してもよい。本実施例では、ASLは、ユーザが米国出身であるため、イタリア手話より優勢であり得る。しかしながら、いったんウェアラブルシステムが、ユーザが、イタリア語話者によって圍繞されている、または物理的にイタリアに居ることを検出すると、ウェアラブルシステムは、イタリア手話がここではASLより先に来るように、選好の順序を変化させることができる。したがって、ウェアラブルシステムは、イタリア手話を英語テキストまたはASLと関連付けられたグラフィックに翻訳することができる。

20

**【0133】**

感覚式アイウェアシステムは、ユーザが外国語手話を理解することに役立ち得るだけでなく、また、ユーザが外国語手話の身ぶりを行うことに役立ち得る。例えば、ウェアラブルシステムは、ユーザ自身の言語を外国語手話に翻訳するように構成されることができる。システムは、外国語手話ジェスチャ(例えば、翻訳された手話)をディスプレイ上に表示することができる。ユーザには、外国語手話におけるジェスチャが見え、ジェスチャを模倣することができる。例えば、ユーザは、聴覚障害手話者と会話し得る。ウェアラブルシステムは、ユーザの発話を捕捉し、ユーザに、手話者が理解する手話における対応するジェスチャを表示することができる。ユーザは、故に、ディスプレイによって提示されるようにジェスチャを行い、手話者と通信することができる。いくつかの実施形態では、ジェスチャをユーザに示すのではなく、ウェアラブルシステムは、手話者がユーザの口頭発話を理解可能であるように、代わりに、ユーザの発話に対応する身ぶりを手話者に通信することができる。

30

40

**【0134】**

ウェアラブルシステムは、オーディオ増幅器(例えば、スピーカ240)を含み、認識された手話をオーディオで提供することができる。例えば、ウェアラブルシステムは、ウェアラブルシステム200のユーザへの再生のために、手話者による手話をオーディオストリームに変換することができる。

(手話と関連付けられた補助情報を決定する実施例)

**【0135】**

人々が、手話を伴う会話を含む、会話中に、単語または語句を把握または理解しないこ

50

とは、稀ではない。ウェアラブルシステムは、表示されるテキストの一部と関連付けられた補助情報を表示し、ユーザの理解を向上させることができる。補助情報は、定義のコンテキストを拡張し、それに追加される、定義、翻訳、説明等の情報を含むことができる。補助情報は、例えば、テキスト、画像、グラフィック、動画、または他の聴覚的もしくは視覚的情報等の種々の形態で存在してもよい。システムは、例えば、図 2 A におけるディスプレイ 220 を介して、補助情報を視覚的に提示することができる。システムは、例えば、図 2 A におけるオーディオ増幅器 240 を介して、聴覚障害者ではないユーザに、補助情報をオーディオとして提示することができる。そのような単語または語句に関する定義、翻訳、説明、または他の情報を提供することによって、ウェアラブルシステムは、有利には、ユーザが、ユーザが観察する手話をより良好に理解することを補助することができる。

10

## 【 0 1 3 6 】

補助情報は、ユーザの環境のコンテキスト情報、発話のコンテキスト等に基づいて決定されてもよい。実施例として、ウェアラブルシステムは、会話相手の身ぶりに関連付けられた補助情報を表示するかどうかを決定する際、少なくとも部分的に、ユーザ挙動を利用することができる。例えば、ユーザは、一時的に、ある方向を（例えば、手話者または手話者の手に向かって）注視し得る。ウェアラブルシステムは、ユーザの視線方向（例えば、内向きに面した結像システム 462 を使用して）を検出することができ、それに応答して、会話相手の身ぶりに関連付けられた補助情報を読み出し、表示することができる。

## 【 0 1 3 7 】

ウェアラブルシステムは、補助情報のデータリポジトリ（例えば、データベース）を備えてもよい。ウェアラブルシステムは、データリポジトリにアクセスすることによって、表示されるテキストと関連付けられた補助情報を読み出すことができる。情報のそのようなデータベースは、ウェアラブルシステムのローカルで、例えば、図 2 A におけるデータモジュール 260 内に記憶される、または遠隔で、例えば、遠隔データリポジトリ 270 内に記憶されてもよい。ウェアラブルシステムは、公共でアクセス可能な情報、例えば、インターネットに関する情報を利用して、補助情報を決定することができる。例えば、ウェアラブルシステムは、ネットワークにアクセスし、会話内の単語 / 語句に関するクエリを辞書、百科事典、または他の類似リソース等のインターネット上のリソースに送信することができる。そのようなリソースは、一般的（例えば、Wikipedia 等の汎用百科事典）、または、例えば、rxlist.com もしくは鉱物学データベース（例えば、webmineral.com）上のもの等の薬物のインデックス等の特殊なものであってもよい。

20

（変換された身ぶりまたは補助情報の例示的表示、放逐、および呼び戻し）

## 【 0 1 3 8 】

ウェアラブルシステムは、単独で、または補助情報と組み合わせて、変換された身ぶり（例えば、テキストまたはグラフィカルフォーマットにおける）をウェアラブルシステム 200 のユーザに提示することができる。例えば、ウェアラブルシステム 200 は、補助情報を手話の変換された身ぶりとともに表示する、変換された身ぶりまたは補助情報を個々に表示する（例えば、補助情報が表示される持続時間の間、補助情報のみを表示する）、または 2 つの表示モード間で切り替えるように構成されることができる。単独で、または補助情報と組み合わせて、変換された身ぶりは、時として、表示アイテムと称され得る。

30

40

## 【 0 1 3 9 】

変換されたテキストまたは補助情報は、種々の方法で提示されてもよい。一実施例では、ウェアラブルシステム 200 は、変換されたテキストまたは補助情報をテキスト吹き出し内に設置することができ、例えば、テキストは、図 13 におけるグラフィック 1355 に図示されるように、手話者の近傍に幾何学的に局所化される。別の実施例として、ウェアラブルシステム 200 は、検出された手話のローリングトランスクリプトを表示するように構成されることができる。本構成では、例えば、ユーザの注意が一時的に逸らされた

50

場合、見逃された単語またはさらに文は、迅速に再読されることができる。手話者の変換されたテキストトランスクリプトは、映画におけるエンドクレジットの提示に類似する、ローリングテキストとして表示されてもよい。

#### 【0140】

変換されたテキストのトランスクリプトを表示するシステムは、いくつかの方法において、例えば、下線が引かれた、着色された、太字テキスト等、補助情報が要求される、単語または語句をハイライトすることができる。そのようなハイライトは、補助情報が読み出される、または表示される前に表示されることができる。本表示モードにおいて構成される、いくつかの実施形態は、ユーザが、ハイライトされたテキストの要求を確認またはキャンセルすることを可能にすることができる。代替として、または加えて、そのようなハイライトは、補助情報とともに表示されることができる。本表示モードは、ユーザに、補助情報が関連付けられるテキストを明確にすることができる。システムは、以下にさらに説明されるように、ユーザが、UI相互作用を通して、現在または過去の変換されたテキストを選択し、関連付けられた補助情報を読み出す、もしくは呼び戻すことを可能にすることができる。

10

#### 【0141】

ウェアラブルシステム200は、UI相互作用を介して情報にアクセスするために、ユーザの眼移動を最小限にするように、変換されたテキストまたは補助情報を設置することができる（例えば、テキスト吹き出し内に、またはローリングトランスクリプトとして）。このように、UIは、簡略化され、ユーザは、その注意を手話者から遠くに引き離す必要がない。変換されたテキストまたは補助情報は、読取アクションを会話相手に最小限に可視にするように設置され、そうすることによって、変換されたテキストまたは補助情報へのユーザのアクセスを露見させることなく、あまり注意散漫とならず、より良好な通信を提供し得る。例えば、手話者の場所を決定することが可能な実装は、変換されたテキストまたは補助情報を手話者の隣に設置することができる。外向きに面した結像システム464からの画像は、例えば、手話者の顔、ジェスチャ等を、例えば、曖昧にしない、適切な場所の決定に役立ち得る。ウェアラブルシステム200は、図8に図示されるプロセスフローを使用して、変換されたテキストまたは補助情報表示の場所を決定することができる。例えば、ブロック850において認識されるオブジェクトは、その手話が手話認識のために処理されるべき手話者であることができる。

20

30

#### 【0142】

変換されたテキストまたは補助情報がオーディオとして提示される場合（例えば、ユーザが、聴覚障害者ではなく、会話相手が、手話を使用する場合）、ユーザまたは会話相手によって被られる注意散漫を低減させる別の実施例として、ウェアラブルシステムは、ユーザに聞こえるために十分であるが、（言語障害者であるが、聴覚障害者ではない）会話相手に聞こえるためには十分ではない、大きさの音量で情報を提示する、またはユーザもしくは会話相手のいずれも発話していないとき、情報を提示することができる。

#### 【0143】

表示されるアイテムは、条件が満たされるまで、可視のままであってもよい。例えば、表示されるアイテムは、固定時間量にわたって、次の表示されるアイテムが表示されるまで、またはユーザアクションによって放逐されるまで、可視のままであってもよい。ユーザアクションは、受動的（例えば、内向きに面した結像システム462によって捕捉されるような眼移動）であってもよい。ウェアラブルシステムは、ユーザが表示されるアイテムを精査したことを決定すると、表示を放逐することができる。例えば、表示されるアイテムが、テキストである場合、システムは、テキストを通じたユーザの眼移動（例えば、左右または上下）を追跡することができる。いったんウェアラブルシステムが、ユーザが表示されるアイテム全体（または表示されるアイテムの大部分）に目を通じたことを決定すると、ウェアラブルシステムは、故に、表示されるアイテムを放逐することができる。別の実施例として、表示されるアイテムは、ユーザが表示されるアイテムによって占有されたエリアから眼を逸らしている（またはそれを見ていない）ことがシステムによって観

40

50

察された後に放逐されてもよい。ユーザアクションはまた、能動的（例えば、外向きに面した結像システム464によって捕捉されるような手のジェスチャ、オーディオセンサ232によって受信されるような音声入力、またはユーザ入力デバイス466からの入力によって）であってもよい。例えば、いったんウェアラブルシステムが、ユーザによるスワイプジェスチャを検出すると、ウェアラブルシステムは、表示されるアイテムを自動的に放逐することができる。

#### 【0144】

ウェアラブルシステムは、特定のユーザのために、カスタマイズされたユーザインターフェース（UI）相互作用のセットをサポートするように構成されることができる。UI相互作用は、指、ポインタ、またはある種類のスタイラスを用いて、眼によるボタン上への注視および後続の固視によって、もしくはその他のいずれかで作動される、ボタンに類似するUI要素の形態をとってもよい。ボタンは、（例えば、キーボード上の）実際の物理的ボタンまたはディスプレイ220によって表示される仮想ボタンであることができる。UI相互作用は、例えば、図4に関連して上記に説明されるように、頭部姿勢の形態をとってもよい。UI相互作用検出の実施例は、図10に関連して上記に説明される。

#### 【0145】

ウェアラブルシステムは、ユーザに、表示されるアイテムの放逐を遅延させるようにプロンプトすることができる。例えば、ウェアラブルシステム200は、表示されるアイテムの明度を低減させ、または配色を変化させ、表示されるアイテムが、間もなく、例えば、数秒以内に放逐されるであろうことをユーザに通知してもよい。上記に説明されるもの等のUI相互作用は、放逐することを延期するために使用されてもよい。例えば、ウェアラブルシステムは、ユーザが表示されるアイテムから眼を逸らしたことを検出してもよい。したがって、ウェアラブルシステムは、表示されるアイテムの透明度を増加させ、表示されるアイテムが間もなく放逐されるであろうことをユーザに通知することができる。しかしながら、ウェアラブルシステムが、眼追跡を通して、ユーザが表示されるアイテムに眼を戻したことを検出する場合、ARシステムは、表示を放逐することを延期することができる。

#### 【0146】

上記に説明されるもの等のUI相互作用はまた、放逐された表示されるアイテムを呼び戻すために使用されてもよい。例えば、ユーザ入力デバイスを通じた入力アクション（例えば、キーボード上のバックスペースの作動）は、直近に表示されたアイテムを呼び戻すために使用される、または呼び戻すのために特定の表示されるアイテムを選択するために使用されることができる。

（感覚式アイウェアシステムの例示的ユーザ体験）

#### 【0147】

図13Aは、感覚式アイウェアシステムの例示的ユーザ体験を示し、感覚式アイウェアシステムは、ユーザのウェアラブルシステムのために、手話（例えば、手話者によってジェスチャされる）を解釈することができる。本実施例は、感覚式アイウェアシステムのユーザが観察する、手話者1301を示す。ユーザは、手話者1301が、場面1305、1310、および1315に示されるように、手のジェスチャのシーケンス1300を行なっていることを知覚することができる。場面1305における手のジェスチャは、単語「how」を表し、場面1310における手のジェスチャは、単語「are」を表し、場面1315における手のジェスチャは、単語「you」を表す。したがって、シーケンス1300は、「How are you」として解釈されることができる。シーケンス1320および1340は、シーケンス1300と同一ジェスチャを示す。ジェスチャ1305は、ジェスチャ1325および1345に対応し、ジェスチャ1310は、ジェスチャ1330および1350に対応し、ジェスチャ1315は、ジェスチャ1335および1355に対応する。しかしながら、シーケンス1300、1320、および1340は、以下にさらに説明されるように、異なるユーザディスプレイ体験を図示する。

#### 【0148】

シーケンス1300における手のジェスチャを英語語句「How are you」に翻訳するために、ウェアラブルシステム200の外向きに面した結像システム464は、一連の画像またはビデオのいずれかとして、ジェスチャのシーケンスを捕捉することができる。ウェアラブルシステムは、ジェスチャを一連の画像またはビデオから抽出することができる。ウェアラブルシステムは、手話認識を抽出されたジェスチャに、例えば、オブジェクト認識装置708または深層学習アルゴリズムを適用することを通して、実施することができる。手話を認識するプロセスでは、ウェアラブルシステムは、ローカルまたは遠隔記憶装置内に記憶される手話辞書にアクセスすることができる。ウェアラブルシステムは、認識された手話から変換されたテキスト（または身ぶりのグラフィック表現）を、ディスプレイ220を介して、ユーザ（図示せず）に表示することができる。感覚式アイウェアシステムはまた、変換された身ぶりに関連付けられた補助情報の要求を受信し、本明細書に説明される技法を使用して、補助情報を読み出し、表示することができる。

10

## 【0149】

図13Aに図示されるグラフィカルシーケンスでは、単語「how」を表すことは、例えば、グラフィック1305および1310に示されるように、2つの明確に異なるジェスチャを行う。ウェアラブルシステムは、第2のジェスチャ（場面1310における）後、単語「how」を表示する前に待機してもよい（場面1305においてジェスチャされるように）。加えて、または代替として、ウェアラブルシステムは、グラフィカルシーケンス1320に示されるように、文または語句が完了されるまで、テキスト変換または表示を延期してもよく、語句「How are you」は、場面1335の最後に示される。ウェアラブルシステムは、例えば、グラフィック1355に示されるように、変換されたテキストまたは補助情報をキャプションまたはテキスト吹き出しとして表示することができる。キャプションまたはテキスト吹き出しは、ユーザのFOV内に位置付けられ、例えば、手話者の顔のユーザのビューを曖昧にせず、手話者に近接近して、ユーザへの注意散漫を最小限にすることができる。

20

## 【0150】

図13Bは、感覚式アイウェアシステムの別の例示的ユーザ体験を示し、標的発話および補助情報の両方が、提示される。本実施例では、ユーザ（図示せず）は、HMDを装着し、手話者1362を知覚することができる。手話者は、手話を使用して、質問「Where is the PTO?」を尋ねている（手話者は、質問の最後に表音文字「O」をジェスチャするように描写されている）。ウェアラブルシステムは、手話者によって行われるジェスチャを認識し、それらをテキストに変換し、変換されたテキストをテキスト吹き出し1360内においてウェアラブルシステムのユーザに表示することができる。ウェアラブルシステムは、「PTO」が頭字語であって、ユーザが日常の発話において頻繁に使用しない単語であることを決定することができる。例えば、ウェアラブルシステムは、一般に使用される単語および語句の辞書を維持し、「PTO」が辞書内にないことを決定することができる。単語「PTO」が辞書内にないことの検出に応じて、ウェアラブルシステムは、語句「PTO」と関連付けられた補助情報のアクセスを開始することができる。

30

## 【0151】

ウェアラブルシステムは、コンテキスト情報に基づいて、頭字語に関する補助情報を読み出すことができる。本実施例では、システムは、その場所情報に依拠してもよく、例えば、システム（およびそのユーザ）は、現在、Virginia州のAlexandriaにある（居る）。システムは、頭字語「PTO」に関する補助情報として、「特許商標庁」を読み出す。システムは、ディスプレイ220を介して、補助情報を仮想バナー1365としてユーザに表示する。図13Bに示される変換されたテキストおよび補助情報の表示モードは、例証にすぎない。いくつかの実施形態は、それらを異なるように表示してもよい、例えば、両方が、キャプションとして連続して表示される。

40

## 【0152】

複数のウェアラブルシステムの複数のユーザは、その個別のウェアラブルシステムの補

50

助を通して、遠隔で通信することができる。図13Cは、テレプレゼンスセッションにおける感覚式アイウェアシステムの例示的ユーザ体験を示す。例えば、図13Cに図示されるように、(人工デバイスの補助なしでは、相互に直接見えない、または聞こえないような)2つの物理的場所1370a、1370bにおける2人のユーザ1372a、1372bは両方とも、それぞれ、ウェアラブルシステム1374a、1374bを装着することができる。ユーザ1372a、1372bの一方または両方は、手話を使用して会話し得る。手のジェスチャは、ユーザの個別のウェアラブルシステムの結像システムによって捕捉され、ネットワーク1290を通して伝送されてもよい。ユーザA1372aの手話は、変換されたテキストとしてユーザB1372bのデバイス上に表示されてもよく、その逆も同様である。

10

**【0153】**

感覚式アイウェアシステムは、検出された手話をテキストにローカルで変換し、変換されたテキストのみをネットワーク1290を通して伝送することができる。他方のユーザのデバイスは、テキストを表示するか、または他方のユーザが聴覚障害者ではない場合、テキストを可聴発話に変換するかのいずれかを行うことができる。これは、対応する画像、ビデオ、またはオーディオを伝送するためよりも少量のデータが、テキストを伝送するために要求されるため、ネットワーク1290の帯域幅が制約される場合に有利であり得る。

**【0154】**

ウェアラブルシステムはまた、ディスプレイ220上に提示される画像を通して、テレプレゼンス会話を向上させることができる。例えば、ディスプレイ220は、遠隔手話者のアバタを変換されたテキストまたは補助情報とともに提示し、参加者の視覚的感覚を引き付けることができる。例えば、内向きに面した結像システム464を装備するウェアラブルシステムは、HMDによってオクルードされる装着者の顔の領域を置換するための画像を捕捉することができ、これは、第1のユーザに、テレプレゼンスセッションの間、第2のユーザのオクルードされていない顔が見え得るように、およびその逆も同様であるように、使用されることができる。第1のユーザと関連付けられた世界マップ情報が、感覚式アイウェアシステムを伴うテレプレゼンスセッションの第2のユーザに通信されてもよい。これは、HMD装着者によって見られる遠隔ユーザの画像の作成を通して、ユーザ体験を向上させることができる。

20

30

**【0155】**

テレプレゼンス用途では、画像情報を捕捉することは、観察者であるユーザと関連付けられたデバイスではなく、手話者であるユーザと関連付けられたデバイスによって実施され(例えば、一人称視点から)、これは、対人シナリオにおいて典型的であり得る。手話の存在の検出および手話からテキストへの変換は、いずれかのユーザと関連付けられたデバイスによって、または遠隔システム、例えば、サーバコンピュータシステム1220によって実施されることができる。手話源は、画像を捕捉するデバイスに基づいて決定されることができ、例えば、ユーザAのデバイスが画像を捕捉するとき、ユーザAが、身ぶりを行なっている。

**【0156】**

図13Dは、手話を解釈するための例示的仮想ユーザインターフェースを図示する。本実施例では、ユーザ1392は、ウェアラブルシステム1380(ウェアラブルシステム200の少なくとも一部を含んでもよい)を装着している。本実施例では、ユーザ1392は、カウンタの背後に居て、カウンタに接近する人物1394を知覚する。例えば、ユーザ1392は、医療施設における看護師または窓口担当、宿泊客を補助するホテル従業員(例えば、コンシェルジュ)等であってもよい。人物1394は、体調が悪く、薬局への行き方等、医療機関を探している場合がある。ウェアラブルシステム1380は、図13Dに示されるように、ユーザ1394による手のジェスチャを観察することができる(例えば、外向きに面した結像システム464を介して)。ウェアラブルシステム1380は、(例えば、オブジェクト認識装置708を使用して)手のジェスチャが、示されるよ

40

50

うに、手話における表現であることを自動的に検出し、手のジェスチャと関連付けられた意味を認識し、手のジェスチャの翻訳をユーザ1392が理解する標的言語（例えば、英語）で提供することができる。ウェアラブルシステム1380は、仮想ユーザインターフェース1382を提示し、ウェアラブルシステムによって捕捉された入力1384a、入力1384aに対応する翻訳1384b（例えば、「近くに薬局はありますか？体調が悪くて。」）を示すことができる。ウェアラブルシステムはまた、ユーザ入力要素1384cおよび1384dを仮想ユーザインターフェース1382上に提供することができる。例えば、ユーザ1392は、手のジェスチャ（例えば、押下ジェスチャ）を使用して、ユーザ入力要素1384cを選択してもよい。ユーザ入力要素1384cの作動はウェアラブルシステムに、例えば、近くの薬局の場所または「分かりません」等の応答のリストを提供させてもよい。いくつかの実施形態では、ウェアラブルシステム1380は、応答のための手話における対応するグラフィックを示すことができる。ユーザ1392は、故に、グラフィックに示されるように、手のジェスチャを使用して、人物1394に応答することができる。別の実施例として、ユーザ入力要素1384dが、作動される場合、ウェアラブルシステムは、例えば、ユーザインターフェース要素1382を放逐するかまたは援助を要請する等、オプションのリストを提供することができる。いくつかの実施形態では、インターフェース1382のエリア1384aは、ユーザ1392が人物1394と通信するために実施し得る、手話ジェスチャ（例えば、「薬局は、通りの向い側にあります」という身ぶり）を示す、出力グラフィックを含んでもよい。

（対人通信のためのツールとしての感覚式アイウェアシステムのための例示的プロセス）

【0157】

10

20

図14Aおよび14Bは、感覚式アイウェアシステムを用いた対人通信を促進するための例示的プロセスを図示する。図14Aおよび14Bにおける例示的プロセス1400および1440は、図2Aに示されるウェアラブルシステムによって実施されることができる。

【0158】

ブロック1404では、ウェアラブルシステムは、環境内の画像情報を捕捉することができる。本明細書に説明されるように、ウェアラブルシステムは、外向きに面した結像システム464を使用して、ユーザの周囲内の画像情報を捕捉することができる。ウェアラブルシステムはまた、環境内のオーディオ情報を捕捉することができる。オーディオ情報は、外向きに面した結像システム464によって入手されたデータと併用され、発話またはジェスチャ源を決定する、または手話の存在を検出することができる。

30

【0159】

ブロック1408では、ウェアラブルシステムは、捕捉された画像情報内に手話の存在を検出する。本検出処理は、ローカル（例えば、ローカル処理モジュール71によって）または遠隔（例えば、遠隔処理モジュール72によって）で行われてもよい。ウェアラブルシステムは、種々のオブジェクト認識装置を使用して、手のジェスチャの存在を検出することができる。例えば、ウェアラブルシステムは、手のジェスチャのシーケンスが手話における語句または文を構成し得ることを見出し得る。別の実施例として、ウェアラブルシステムは、一連の手のジェスチャならびに唇移動を検出してよい。ウェアラブルシステムは、そのようなジェスチャおよび唇移動がオーディオ情報を付随しないため、手のジェスチャおよび唇移動が手話と関連付けられることを見出し得る。

40

【0160】

いくつかの実施形態では、ウェアラブルシステムは、コンテキスト情報に基づいて、手話を検出および解釈することができる。例えば、ウェアラブルシステムは、（例えば、夕食の会話の）オーディオ信号を受信し、それらの信号を言語に変換し、または意味をその言語から抽出し、それによって、議論の話題のジャンル（または他の属性）を推測することができる。これは、手話を解釈するために使用されることができる（例えば、議論の話題と整合するように手のジェスチャを解釈するため等）。

【0161】

50

ウェアラブルシステムは、ユーザ自身の手話を検出または無視するように構成されることができ、ブロック1408の機能は、ユーザ自身の手話が比較的近距离において一人称視点から捕捉され得るため、本構成に基づいて異なり得る。例えば、システムが、ユーザ自身の手話を捕捉するように構成される場合、ユーザの手に下向きに指向される付加的な外向きに面したカメラが、オンにされてもよい、または外向きに面した結像システムが、ユーザの手の画像を捕捉するように広角モードに構成されてもよい。

【0162】

ブロック1412では、システムは、手話が検出されるかどうかを決定する。手話が、検出される場合、プロセスフロー1400は、ブロック1416に継続する。手話が、検出されない場合、フローは、ブロック1408（図示されるように）またはブロック1404（図示せず）に戻る。

10

【0163】

ブロック1404 - 1412における動作は、ウェアラブルシステム（その結像システムを含む）がオンになると、または手話認識機能が有効にされると、持続的または周期的に（例えば、あるサンプリング頻度で）、実施されてもよい。これらの動作は、フローチャート1400および1440内の他のブロックと並行して（例えば、時限割込によって駆動されるバックグラウンドタスクとして）実施されることができ、それらは、例証目的のために、処理フローシーケンス内で離散ブロックとして示される。しかし、それらは、図示されるシーケンスによって限定されない。上記に説明される実施例以外の多くの処理フローも、システム設計者の裁量に応じて可能性として考えられる。

20

【0164】

ブロック1416では、ウェアラブルシステムは、検出された手話源（例えば、手話者）を決定することができる。源は、ユーザの物理的近傍に居る人物、ユーザ、またはユーザが知覚する視覚的コンテンツ内の人物であってもよい。手話源は、例えば、システムが、ウェアラブルシステムのFOVの中心またはその近傍の人物のみからの手話を処理するように構成される場合、関連し得る（FOVの中心の外側の人々からの手話は、例えば、複数の人物が並行して手話で会話しているとき、破棄され、さらに処理され得ない）。別の実施例として、ウェアラブルシステムは、ユーザが見ている人物のみに関する手話認識のために、ジェスチャを処理することができ、これは、FOVの中心における人物である場合とそうではない場合がある。ウェアラブルシステムは、内向きに面した結像システム462および外向きに面した結像システム464によって入手されたデータに基づいて、ユーザが見ている人物を識別することができる。例えば、外向きに面したカメラは、ユーザに対する手話者の位置を含む、情報を提供することができる。内向きに面したカメラは、ユーザが見ている方向を含む、情報を提供することができる。両方のカメラからの情報を使用することによって、ウェアラブルシステムは、ユーザが見ている人物と、その人物が手話源であるかどうかとを決定することができる。

30

【0165】

ブロック1420では、システムは、手話源が変化したかどうかを決定する。変化した場合、フロー1400は、図14Bに示されるように、ブロック1424 - ブロック1444に継続する。手話源が変化しなかった場合、フローは、ブロック1428に移行し、手話認識処理を継続し、これは、画像情報を捕捉するステップ（ブロック1404）と、手話の存在を検出するステップ（ブロック1408）と、図14Bに示される処理ステップとを含むことができる。例えば、システムが、ジェスチャが同一手話者から生じ続けていると決定する場合、システムは、画像情報を捕捉し、手話を検出し続けることに加え、ブロック1448から開始する機能を実施し続けることができる。

40

【0166】

ブロック1448では、ウェアラブルシステムは、手話をユーザによって理解される言語に翻訳することができる。例えば、システムは、認識された手話をテキストに変換することができ、これは、システムによって表示される（例えば、テキスト吹き出しまたはキャプションとして）と、ユーザによって読み取られることができる。ある場合には、ユー

50

ザが、異なる手話を理解する場合、例えば、グラフィックがユーザ自身の手話における身ぶりに変換された身ぶりを示すように、他の手話者の身ぶりのグラフィック表現が、ユーザに表示されることができる。

【0167】

ブロック1452では、例示的システムは、システムが、ユーザ自身ならびに会話相手の手話を検出するように構成されるとき、検出された手話がユーザ自身のものであるかどうかを決定することができる。該当する場合、プロセスは、ブロック1484に進み、システムは、変換されたテキストを観察者/会話相手のディスプレイデバイスに伝送することができる。

【0168】

ブロック1484から、システムは、ブロック1488に進み、処理を継続することができる。システムが、ユーザ自身の手話を無視するように構成されるとき、ブロック1452および1484は両方とも、フローから省略されることができる。検出された手話が、ユーザ自身のものではない場合、フローは、ブロック1456に継続する。

【0169】

ブロック1456では、ウェアラブルシステムは、上記に説明されるように、ディスプレイによって、変換されたテキストを表示することができる。システムのユーザが、聴覚障害者ではない場合、テキストは、視覚的ディスプレイに加え、またはその代替として、例えば、オーディオ増幅器240を通して、オーディオとして提示されることができる。

【0170】

ブロック1460では、ウェアラブルシステムは、変換されたテキストに関する補助情報の要求を監視することができる。補助情報の要求は、トリガ条件の検出に応じて、ユーザのウェアラブルシステムによって送信されてもよい。いくつかの例示的トリガ条件は、ユーザのインジェクション、例えば、ユーザのジェスチャまたはユーザ入力デバイス466の作動を含む、またはユーザが理解し得ない単語（または語句）の検出に応じてもよい。

【0171】

ブロック1464では、システムは、要求が受信されたかどうかを決定する。要求が受信されない場合、フローは、ブロック1476に移行し、これが、以下にさらに説明される。

【0172】

要求が受信された場合、ブロック1468では、システムは、変換されたテキスト（またはその要求される一部）と関連付けられた補助情報を読み出すことができる。本明細書に説明されるように、補助情報は、例えば、本明細書に説明されるように、ユーザの場所、発話のコンテキスト、または他のタイプの情報等のコンテキスト情報に基づいて、決定され、読み出されてもよい。

【0173】

ブロック1472では、ウェアラブルシステムは、ウェアラブルシステムのディスプレイ220を介して、読み出された補助情報を表示することができる。いくつかの実装では、ウェアラブルシステムは、補助情報を表示する前に、変換されたテキストの表示を放逐してもよい。

【0174】

フローは、ブロック1464または1472から、ブロック1476に進んでもよい。ブロック1476では、システムは、変換されたテキストまたは補助情報表示を放逐するための条件を検出することができる。そのような条件が検出されると、ブロック1480では、システムは、変換されたテキストまたは補助情報の表示を放逐し、ブロック1488に継続することができる。ブロック1488では、手話認識の処理は、上記のブロック1428の説明に類似する様式において継続する。

【0175】

ブロック1404 - 1412に関して上記に記載されるものと同様に、フローチャート

10

20

30

40

50

1440内の動作は、フローチャート1400および1440の他のブロックと並行して実施されてもよい。それらは、例証目的のために、処理フローシーケンス内で離散ボックスとして示されるが、それらは、図示されるシーケンスによって限定されない。例えば、システムは、変換されたテキストに関する補助情報を表示してもよい(ブロック1472において)一方、システムは、付加的手話をテキストに変換する(ブロック1456において)、補助情報要求監視を実施する(ブロック1460において)、または別の変換されたテキストに関する補助情報を読み出す(ブロック1468において)。別の実施例として、システムは、手話をテキストに変換することができる(ブロック1448において)一方、以前に要求された(補助情報に関する)変換されたテキストに関する補助情報を読み出す(ブロック1468において)。多くの他の処理フローも、システム設計者の裁量に応じて可能性として考えられる。

10

**【0176】**

図14Cは、変換されたテキストと関連付けられた補助情報を決定し、補助情報を提示するための例示的方法のプロセスフロー図である。本プロセス1490は、本明細書に説明されるウェアラブルシステム200または別のコンピューティングデバイス上で実行されることができ、それ自体が、手話認識機能性を有する場合とそうではない場合がある。本プロセス1490は、1つの感覚式アイウェアシステムを使用して、手話を検出し、手話をテキストに変換し、変換されたテキストを別のデバイスまたはシステム上に表示することがより有利な状況に適用可能であり得る。例示的状況は、手話者が第2の人物と遠隔で通信することを所望する場合であり得る。手話者のウェアラブルシステムは、手話者自身の手話をテキストに変換することができる。ウェアラブルシステムは、変換されたテキストを第2の人物によって視認可能な遠隔システムに伝送することができる。変換されたテキストは、対応する画像またはビデオよりはるかに少ない情報ビットで伝送されることができ、そのようなプロセスは、有利には、伝送媒体からはるかに低い帯域幅を要求する、またははるかに信頼性のある通信をもたらすことができる。

20

**【0177】**

プロセス1490は、ブロック1492から開始し、デバイスまたはシステムは、ある種類の処理を実施し、これは、手話処理に関連する場合とそうではない場合がある。ブロック1494では、デバイスまたはシステムは、テキストがウェアラブルシステムから受信されたかどうかを決定することができる。該当しない場合、プロセスは、ブロック1492に戻ることができる。該当する場合、プロセスは、ブロック1496に進むことができる。ブロック1496では、デバイスまたはシステムは、テキストをウェアラブルシステムから受信し、テキストをレンダリングすることができる。プロセスは、次いで、ブロック1456に進むことができる。レンダリングデバイスが、HMDを備える場合、レンダリングデバイスは、テキストをユーザの物理的環境上にオーバーレイされる仮想コンテンツとして提示することができる。ブロック1456 - 1480における処理は、図14Bに関連して上記に説明されるように、同様に進むことができる。

30

**【0178】**

図15は、感覚式アイウェアシステムを用いた対人通信を促進するための別の例示的プロセスを図示する。例示的プロセス1500は、本明細書に説明されるウェアラブルシステム200の1つ以上のコンポーネント(例えば、単独で、または組み合わせて、ローカル処理およびデータモジュール260、遠隔処理モジュール270)によって実施されてもよい。図12を参照して説明されるように、本図15に説明されるステップのうちの一つ以上のものは、例えば、別のユーザのウェアラブルシステムまたは第三者のサーバシステム等、ユーザのウェアラブルシステムの一部ではない、1つ以上のコンピューティングデバイスによって実施されることができる。

40

**【0179】**

ブロック1510では、ウェアラブルシステムは、環境内の発話を識別および認識することができる。発話は、手話の形態であってもよい。例えば、ウェアラブルシステムは、外向きに面した結像システム464によって入手されたデータを分析し、手話の一部であ

50

る、手のジェスチャを識別することができる。ウェアラブルシステムはまた、オーディオセンサ 232 によって入手されたオーディオデータを分析することができ、これは、ユーザの環境内の人物による発話を含んでもよい。ウェアラブルシステムは、オブジェクト認識装置 708 を使用して、発話を認識することができる。例えば、ウェアラブルシステムは、オブジェクト認識装置を使用して、手話の画像を分析することによって、語句または単語の存在を認識することができる。ウェアラブルシステムはまた、図 7 に説明される種々の発話認識アルゴリズムを使用して、オーディオデータを認識することができる。

【0180】

ブロック 1520 では、ウェアラブルシステムは、標的言語を識別することができる。標的言語は、ウェアラブルシステムのユーザが通信するために使用する、言語であってもよい。例えば、ユーザは、英語を使用して他の人々と通信し得る一方、認識されたオリジナル発話（別のユーザによって使用される）は、手話である。標的言語はまた、ユーザまたはウェアラブルシステムによって選択された言語であってもよい。例えば、ユーザは、ユーザが別の言語を話す場合でも、ユーザが、手話を使用して、別の人物と通信することを所望し得るため、ASL を標的言語として選択してもよい。別の実施例として、ウェアラブルシステムは、ユーザの場所に基づいて、言語を自動的に選択してもよい。例えば、ウェアラブルシステムは、ユーザが滞在している国を決定し、その国の公用語を標的言語として選択することができる。

10

【0181】

ブロック 1530 では、ウェアラブルシステムは、検出された発話を標的言語に変換することができる。ウェアラブルシステムは、例えば、辞書翻訳等の本明細書に説明される種々の技法を使用し、そのような変換を実施することができる。

20

【0182】

ブロック 1540 では、ウェアラブルシステムは、ウェアラブルシステムのユーザへの提示のために、変換された発話と関連付けられた視聴覚コンテンツを決定することができる。一実施例として、視聴覚コンテンツは、標的言語におけるテキストを含んでもよい。別の実施例として、視聴覚コンテンツは、標的言語におけるオーディオストリームであってもよく、変換された発話は、発話された言語である。さらに別の実施例として、視聴覚コンテンツは、標的言語が手話である場合、グラフィックまたは動画であってもよい。

【0183】

随意のブロック 1550 では、ウェアラブルシステムは、提示のために、視聴覚コンテンツを頭部搭載型ディスプレイに通信することができる。例えば、視聴覚コンテンツは、1人のユーザのウェアラブルシステムから別のユーザのウェアラブルシステムに通信されてもよい。本実施例では、第1のユーザのウェアラブルシステムは、第1のユーザを捕捉し、発話を標的言語に変換し、変換された発話を第2のユーザのウェアラブルシステムに通信することができる。

30

（環境と相互作用するためのツールとしての例示的感覚式アイウェアシステム）

【0184】

別の人物によるジェスチャを認識することに加えて、またはその代替として、本明細書に説明されるウェアラブルシステムはまた、例えば、図 7 を参照して説明される種々のテキスト認識アルゴリズムを用いて、環境内の標識を認識することができる。ウェアラブルシステムはまた、テキストを修正し（例えば、表示特性またはテキストのコンテンツを修正する）、修正されたテキストをユーザの物理的環境上にレンダリングすることができる。例えば、修正されたテキストは、ユーザが、オリジナルテキストではなく、修正されたテキストを知覚するであろうように、オリジナルテキストにオーバーレイし、それをオクルードするようにレンダリングされてもよい。

40

（テキストの表示特性を修正する実施例）

【0185】

図 16A - 16E は、環境内のテキストを認識し、テキストと関連付けられた表示特性を修正し、修正されたテキストをレンダリングするように構成される、感覚式アイウェア

50

システムのための例示的ユーザ体験を図示する。図16Aを参照すると、ユーザ210は、ウェアラブルシステム(図16Aには図示せず)を装着することができ、ディスプレイ220を介して、環境内の物理的オブジェクト1606が見え得る。ウェアラブルシステムは、外向きに面した結像システム464を含むことができ、これは、画像1602内にオブジェクト1606を備える、画像1602を捕捉することができる。外向きに面した結像システム464に加え、またはその代替として、ウェアラブルシステム200は、他のセンサまたはデバイスを使用して、物理的オブジェクトの画像を捕捉することができる。例えば、ユーザ入力デバイス466(例えば、トーテム)は、結像能力を有してもよく、画像1602を捕捉することができる、これは、オブジェクト1606の画像を含む。オブジェクト1606は、手書き文字、表音文字、記号、または表意文字1610をその上または中に含有し得る、標識または他のオブジェクトを含んでもよい。例えば、表音文字が、オブジェクト上に書かれている、またはオブジェクトから、それとともに成形される、もしくはそれに内蔵されてもよい。テキストはまた、静的または点滅光のシーケンス、もしくは1つ以上の物理的オブジェクトの配列であってもよい。図16A-16Eに示される実施例では、オブジェクト1606は、交通一時停止標識である。他の実施例では、限定ではないが、オブジェクト1606は、任意のタイプの表記(例えば、商業用または公共看板)、書籍、雑誌、紙片、コンピュータディスプレイ画面、テレビ画面等であり得る。

10

**【0186】**

ウェアラブルシステム200は、例えば、図7を参照して説明されるように、1つ以上のオブジェクト認識装置708を使用して、画像1602を分析し、オブジェクト1606を認識することができる。一実施例として、ウェアラブルシステムは、オブジェクト1606が交通標識であることを認識することができる(例えば、図16Aでは八角形である、オブジェクト1606の形状に基づいて)。別の実施例として、ウェアラブルシステムは、オブジェクト1606内のテキストの存在を認識することができる。ウェアラブルシステムは、テキストのフォーマットにかかわらず、テキストを認識することができる(例えば、テキストがオブジェクト上にある、またはテキストを投影する光のシーケンス(例えば、ネオンライト、LEDライト等)によって表されるかどうかにかかわらず)。

20

**【0187】**

図18を参照してさらに説明されるであろうように、ある実施形態では、ウェアラブルシステム200は、テキストの意味を認識し、テキストをオリジナル言語から標的言語に変換することができる。例えば、ウェアラブルシステム200は、例えば、英語、中国語、スペイン語、ドイツ語、アラブ語、ヒンディー語等、種々の言語からの表音文字、記号、または表意文字を識別し、オリジナルの表示される言語からのテキストを別の言語に翻訳することができる。いくつかの実施形態では、そのような翻訳は、以前に規定された設定(例えば、ユーザの選好またはユーザの人口統計もしくは地理的情報等)に従って自動的に生じることができる。いくつかの実施形態では、翻訳は、ユーザからのコマンド(例えば、口頭またはジェスチャ)に回答して行われることができる。

30

**【0188】**

ウェアラブルシステム200は、オブジェクト認識装置708を使用して、テキスト1610の特性を分析することができる。例えば、ウェアラブルシステム200は、テキスト1610と関連付けられたフォントサイズまたは書体を認識することができる。ウェアラブルシステムは、テキスト1610の特性を調節し、修正されたテキストを生成することができる。例えば、ウェアラブルシステム200は、テキスト1610のサイズを調節し、テキスト1610を拡大または縮小してもよい。修正されたテキストのサイズは、部分的に、眼210からオリジナルテキスト1610までの距離またはユーザの特性に依存してもよい。例えば、テキスト1610が、ユーザから遠く離れている場合、ウェアラブルシステムは、テキスト1610を拡大することができる。別の実施例として、ユーザの眼の能力に応じて、システムは、テキストのサイズを調節する方法に関する決定を行うことができる。ウェアラブルシステムは、ユーザから以前に入手された情報に基づいて、人

40

50

物の眼の能力を決定することができる。例えば、ユーザは、眼の任意の視覚問題が存在するかどうかを入力することができる。ウェアラブルシステムはまた、視覚試験をユーザに実施し（例えば、仮想オブジェクトを異なる深度平面およびサイズにおいて表示し、ユーザが仮想オブジェクトを明確に知覚することができるかどうかを決定することによって）、ユーザの眼の能力を決定することができる。ユーザの眼の能力に基づいて、ウェアラブルシステムは、テキストの特性（例えば、距離/場所、色、サイズ、フォント等）に基づいて、ユーザがテキスト1610を知覚する可能性が高いであろうかどうかを決定することができる。例えば、ウェアラブルシステムは、ウェアラブルシステムが、ユーザがテキストを明確に知覚することができないことを決定する場合（例えば、テキストが焦点からずれているとき）、テキストを拡大または太字にすることができる。ユーザが近視であるが、テキストがユーザから遠く離れている場合、ウェアラブルシステムは、ユーザがテキストをより容易に知覚し得るように、テキストのサイズを拡大することができる。サイズ調節は、近視の程度に対応してもよい。サイズは、ユーザの近視の程度が大きい場合、より大きい拡大率と関連付けられてもよい一方、サイズは、ユーザの近視の程度が小さい場合、より小さい拡大率と関連付けられてもよい。本明細書にさらに説明されるように、ウェアラブルシステムはまた、ユーザの眼の能力に基づいて、修正されたテキストの表示場所を変化させることができる。図3を参照すると、ディスプレイシステム220は、複数の深度平面を含むことができる。ユーザが遠視であるが、テキストがユーザに近接する場合、ウェアラブルシステムは、修正されたテキストがユーザから遠く離れて現れるように、修正されたテキストをオリジナル深度平面よりユーザから遠く離れた深度平面306にレンダリングすることができる。サイズ調節は、テキストのフォントサイズを変化させることによって生じることができる（例えば、テキストが文字列として認識される場合）。サイズ調節はまた、テキスト1610を含有する画像1602の一部をズームインまたはアウト（例えば、デジタルズーム）することによって生じることができる（例えば、テキストが、テキスト文字列ではなく、画像として分析される場合）。

10

20

**【0189】**

ウェアラブルシステム200は、修正されたテキストをユーザにレンダリングすることができる。図16Aを継続して参照すると、HMDを装着しているユーザには、オブジェクト1606のレンダリングされたバージョン1618を含有する、仮想画像1614（HMDによってレンダリングされるように）が見え得る。いくつかの実装では、オブジェクト1606のレンダリングされたバージョン1618は、オリジナルテキストをオクルードすることができる。図16Aに示されるように、レンダリングされたテキスト1622は、「STOP（一時停止）」であって、オリジナルテキスト1610と比較して拡大されている。HMDは、拡大されたテキスト1622をオリジナルテキスト1610上にオーバーレイされるようにレンダリングすることができ、したがって、ユーザは、オリジナルテキスト1610を知覚し得ない。本実施例では、テキストサイズを増加させることによって、ユーザは、有利には、実際には、はるかに小さく、知覚が困難であり得る、下層テキスト1610をより容易に知覚し、理解し、それに応答することができる。

30

**【0190】**

図16Bは、ユーザの環境内のテキストの特性を修正する別の実施例を図示する。レンダリングされたテキスト1634によって示されるように、ウェアラブルシステム200は、オリジナルテキスト1610のフォントを太字にすることができる。フォントを太字にすることに加え、またはその代替として、例えば、テキスト色、陰影、輪郭、フォーマット（例えば、イタリック体、下線、アライメント、両端揃え等）等を変化させること等、オリジナルテキスト1610への他の改変も、同様に行われることができる。ウェアラブルシステム200は、レンダリングされたテキスト1634を点滅、回転させる等、テキスト1610と関連付けられたグラフィック要素を追加（または修正）してもよい。

40

**【0191】**

図16Cは、修正されたテキストを焦点インジケータ1640とともにレンダリングする実施例を図示する。焦点インジケータ1640は、黒丸、十字線、後光、色、知覚され

50

る深度の変化（例えば、レンダリングされたテキストをより近くに現れさせる）、テキストの背景の追加または変化、動画、または他の視覚的効果等の視覚的効果を備えることができ、これは、ユーザの注意を引き付ける。図16Cに示される実施例では、ウェアラブルシステム200は、焦点インジケータ1640をテキスト1638がレンダリングされる背景1650として表示するように構成されてもよい。背景1650は、境界領域1642と、内側領域1646とを備えることができる。境界領域1642は、内側領域1646を境界することができる。示される実施形態では、仮想表音文字1638が、内側領域1646内に表示される。テキスト背景1650は、テキスト背景1650が、ユーザにHMDを伴わずに見えないであろうものと異なる背景であるように、表示される画像1614内にレンダリングされることができる。いくつかの実施形態では、内側領域1646および境界領域1642のうちの一つ以上のものは、モノクロ（例えば、白色、黒色、または灰色）である。システムは、オリジナルテキスト1610ではなく、レンダリングされたテキスト1638が、ユーザによって見えるように、背景1650を改変することができる。例えば、背景は、オリジナルテキスト1610をオクルードし得るように、不透明であってもよい。処理電子機器はまた、画像1614の残りの中に混成するように、背景1650を表示するように構成されることができる。例えば、背景1650は、画像1614の残りと同じ色およびテクスチャ効果を有してもよい。ウェアラブルシステムはまた、例えば、後光をテキスト1638または背景1650の周囲に表示する等、テキスト1638または背景1650をハイライトするように、背景1650およびテキスト1638を表示することができる。そのような場合、背景1650は、画像1614の残りの中にシームレスに統合し得ない。例えば、内側領域1646は、背景1650またはテキスト1638を強調するために、境界領域1642によって輪郭が付けられ得る。

#### 【0192】

ある状況下では、オリジナルテキストの視覚的外観は、例えば、ユーザとオブジェクト1606との間の環境効果（例えば、雨、霧）のため、明確ではない場合がある。図16Dは、より読みやすいように、テキストの特性を修正し、修正されたテキストをレンダリングする実施例を図示する。本図では、テキスト1626は、ユーザ210にぼけて現れる。テキストは、種々の理由から、ぼけて知覚され得る。例えば、低視力を伴うユーザは、特定の距離においてテキストがはっきり見えない場合がある。近視を伴うユーザは、近傍に現れるテキストの画像が比較的明確である一方、遠くに現れるテキストがぼけることを見出し得る。同様に、遠視を伴うユーザには、遠くに現れるテキストが明確に見え得る一方、近傍に現れるテキストに合焦させることが困難である。しかし、眼条件は、画像がぼけて現れ得る唯一の理由ではあり得ない。眼210が遠近調節することができるものより近くまたはより遠くに現れるテキストもまた、ぼけて現れ得る。テキストが、ユーザに対して急速に移動するように現れる場合、テキスト1626は、ぼけて現れ得る。気候または天候要因ならびに画像を入手したカメラの分解能等の上記に説明されるもの以外の要因もまた、一因となり得る。

#### 【0193】

本実施例では、ウェアラブルシステム200は、ぼけたテキスト1626またはそうでなければ読取が困難であるテキストをより明確にまたはより読みやすくすることができる。テキストが、ユーザにぼけて現れるが、ウェアラブルシステムによって受信された画像ではそうではない場合、ウェアラブルシステムは、図13Aを参照して説明される類似技法を使用して、外向きに面した結像システム464または別のデバイス（例えば、ユーザ入力デバイス466または車載カメラ等のウェアラブルシステムの外部のカメラ等）によって入手された画像を分析し、テキスト1626を識別することができる。ウェアラブルシステムは、事実上、テキスト1630によって示されるように、テキストをレンダリングすることができる。ある実装では、ウェアラブルシステムは、ユーザまたは環境の条件に基づいて、仮想テキスト1630の特性を調節することができる。例えば、ユーザが、近視である場合、ウェアラブルシステムは、テキスト1626のフォントを拡大する、またはユーザのより近く（例えば、より近い深度平面上）に現れるようにテキストをレンダ

10

20

30

40

50

リングすることができる。別の実施例として、環境が暗いとき、ウェアラブルシステムは、テキスト1630と仮想画像1614の他の領域との間のコントラスト比を増加させることができる。

#### 【0194】

いくつかの状況では、テキスト1626は、ウェアラブルシステムによって取得される画像1602がぼけているため（例えば、高速運転速度に起因して、またはカメラの分解能が低いとき）、ぼけて現れる。本明細書に説明されるように、ウェアラブルシステムは、オブジェクト認識装置708を使用して、ぼけたテキスト1626の存在を識別することができる。例えば、ウェアラブルシステムは、オブジェクト1606内または上のテキストの存在に関する尤度を決定することができる。いくつかの状況では、尤度が閾値を超える場合、ウェアラブルシステムは、例えば、OCRアルゴリズム等の図7を参照して説明される、1つ以上のテキスト認識アルゴリズムを使用して、ぼけたテキスト1626に対応する最も可能性が高い表音文字1630を識別することができる。

10

#### 【0195】

図16Eは、オリジナルテキスト1610が、障害物1654に起因して、部分的に判読不能であるときのシナリオを図示する。示されるように、障害物1654は、オリジナル画像1602内のオリジナルテキスト1610の一部を被覆する。しかしながら、障害物1654は、1つ以上の種々の形態をとることができる。例えば、障害物1654は、例えば、支柱、建物等、眼210またはディスプレイと画像1602との間のある物理的障害物であり得る。障害物1654はまた、上記に説明されるもの等の環境または天候障害物であり得る。障害物1654はまた、オブジェクト1606上にあり得る（例えば、テキスト1610の一部が、標識1606上の別のオブジェクトによってオクルードされる、またはテキスト1610の一部が、消去される、欠損している、またはステッカーによって被覆される）。これは、例えば、蓄積された埃または泥を有する表面、書字1610が見出されるオブジェクト1606の表面への損傷、インクのしみ（例えば、プリンタから）、オリジナルテキスト1610内の歪曲、または任意の他の類似障害物1654を含み得る。

20

#### 【0196】

システムは、オリジナルテキスト1610が伝える内容を決定する際、コンテキスト情報（時として、本明細書では、コンテキスト手掛かりとも称される）を使用してもよい。本明細書に説明される種々のコンテキスト手掛かりが、個々に、または組み合わせて、ウェアラブルシステムによって使用され、テキスト1610に関する完全テキストを決定してもよい。例示的コンテキスト手掛かりは、ユーザの場所である。例えば、上記に説明されるように、GPSシステム37（図2B参照）は、ユーザの場所データを入手ことができ、場所データに基づいて、ウェアラブルシステムは、テキストの言語に関する初期推量を提供することができる。適用可能である場合、いくつかの実施形態では、ウェアラブルシステムは、付加的情報を可視スペクトル外の波長（例えば、赤外線、紫外線）における1つ以上の光源から受信された信号から得てもよい。例えば、ウェアラブルシステムは、紫外線光を標識1606に向かって放出し、紫外線光下のみで可視の表記情報を露見させてもよい（または別の源（例えば、太陽）によって表記から反射された紫外線光を検出してよい）。いくつかの実施形態では、システムは、システムがオリジナルテキスト1610の可視部分をチェックし得る、単語のデータベースへのアクセスを有する。そのような実施例では、ウェアラブルシステム200は、表音文字の候補または最も可能性が高い単語を決定可能であり得る。例えば、図16Eに示されるように、システムは、部分的に、オブジェクト1606の八面体形状またはオブジェクト1606の赤色（図示せず）に起因して、表音文字が「STOP（一時停止）」であると推測する。

30

40

#### 【0197】

ウェアラブルシステムは、コンテキスト手掛かりとして、周囲の単語、記号、句読点、または表意文字に依拠し、オリジナルテキスト1610が伝える内容を決定可能であり得る。ある実施形態では、システムは、例えば、機械学習技法を使用して、場所特有のコン

50

テキスト手掛かりを識別可能である。例えば、システムは、ユーザが路上を運転中であって、路上標識上で頻繁に使用される単語に対するテキストの識別に先入観があり得ることを検出可能であり得る。ウェアラブルシステムは、ローカル処理およびデータモジュール 270 または遠隔処理モジュール 280 (例えば、図 2 A 参照) によってアクセスされる、データベースを備えてもよい。データベースは、ユーザが従事している特定のアクティビティ (例えば、スキー)、ユーザの地理的場所、ユーザの進行速度、ユーザの高度、システムによって受信される周囲雑音の音量またはタイプ、システムによって受信されたエリア内の可視または他の光のレベルまたはタイプ、システムの周囲の温度または気候、ユーザからのテキストの知覚された距離、もしくはシステムが取り上げる別の当事者によって発話された単語のカテゴリまたは複数のカテゴリと関連付けられた単語のカテゴリを記憶してもよい。いくつかの実施形態では、ウェアラブルシステムは、本情報をコンテキスト手掛かりとして使用して、上記に説明される 1 つ以上の関連付けに従って、ユーザによって視認されるテキストの単語または言語に関して可能性がより高い候補をより正確に絞り込むことができる。いくつかの実施形態では、ウェアラブルシステムは、機械学習アルゴリズム (例えば、深層ニューラルネットワーク) を使用して、種々の状況における以前の単語から「学習」し、本状況に基づいて、可能性が高い単語を識別することができる。故に、本学習を実施することによって、ウェアラブルシステム 200 は、ユーザの挙動に具体化されることができ、テキストをより迅速または効率的に決定することができる。

【0198】

図 16 A - 16 E に説明される実施例では、システムは、ユーザからのオリジナル表音文字が現れる知覚された距離に基づいて、テキストを表示すべき深度を決定することができる。オリジナル表音文字とユーザとの間の知覚された距離は、立体視アルゴリズムを適用する (例えば、外向きに面した結像システムによって入手されたデータに)、または深度センサ (例えば、ライダ) によって入手されたデータを分析する等によって、種々の技法を使用して測定されてもよい。立体視アルゴリズムは、ブロックマッチングアルゴリズム、セミグローバルマッチングアルゴリズム、セミグローバルブロックマッチングアルゴリズムアルゴリズム、視差マップ、三角測量、深度マップ、ニューラルネットワークアルゴリズム、同時位置特定およびマッピングアルゴリズム (例えば、SLAM または v-SLAM) 等を含むことができる。ユーザに近接して知覚される表音文字は、ディスプレイシステム 220 上の近傍深度に表示されてもよい。いくつかの実施形態では、ユーザから第 1 の距離閾値 (例えば、約 800 cm) より近くに現れる表音文字は、第 1 の深度において、システム上に表示される。いくつかの実施形態では、第 1 の距離閾値は、約 200 cm より近くに現れる表音文字が第 1 の深度に表示されるように、200 cm である。いくつかの実施形態では、第 1 の距離閾値は、約 80 cm である。表音文字が第 1 の深度にあるかのように表示されるかどうかまたは使用される第 1 の距離閾値は、いくつかの要因に依存し得る。1 つの要因は、システムが表示可能な異なる深度の数であり得る。例えば、実施形態がオブジェクトを 2 つの異なる深度にのみ表示する場合、より短い第 1 の距離閾値が、使用されてもよい一方、実施形態がテキストをより多数の異なる深度に表示し得るとき、より小さい範囲が、使用されてもよい。例えば、ユーザが、新聞を読んでいる場合、システムは、テキストをユーザの近くに知覚し、したがって、新聞上の表音文字は、近い深度にあるかのように、システム上に表示されるであろう。図 3 に示されるように、ディスプレイシステム 220 は、複数の深度平面 306 を備えてもよく、これは、仮想オブジェクトをユーザから異なる距離に現れさせることができる。ある実装では、ウェアラブルシステムは、ユーザの眼の能力に基づいて、修正されたテキストのレンダリング場所を調節することができる。例えば、ユーザが、近視である場合、ウェアラブルシステムは、修正されたテキストを、テキストが元々対応する深度平面よりユーザに近い深度平面にレンダリングすることができる。別の実施例として、ユーザが、遠視である場合、ウェアラブルシステムは、修正されたテキストを、オリジナルテキストが現れる場所よりユーザから遠く離れた深度平面にレンダリングすることができる。

【0199】

同様に、ユーザから遠くに知覚される表音文字は、ディスプレイシステム上の遠い深度に表示されてもよい。いくつかの実施形態では、ユーザからほぼ第2の距離閾値より遠くに現れる表音文字は、第1の深度より遠く離れて現れる第2の深度においてシステム上に表示される。いくつかの実施形態では、第2の距離閾値は、約300cmである。いくつかの実施形態では、第2の距離閾値は、約600cmである。いくつかの実施形態では、第2の距離閾値は、約10mである。例えば、運転中に広告板上で視認されるテキストは、第2の深度にレンダリングされてもよい。

#### 【0200】

第1および第2の距離閾値間の差異は、種々の実施形態において異なり得る。差異の大きさは、例えば、システムがテキストを表示することができる深度の数、実世界オブジェクトまたはテキストから距離を知覚するシステムの能力の精度または正確度、もしくは手動または工場設定等のいくつかの要因に基づき得る。いくつかの実施形態では、差異は、100m未満である。いくつかの実施形態では、差異は、700cm未満である。いくつかの実施形態では、差異は、30cm未満である。ある実施形態では、差異は、ゼロである（例えば、第1の距離閾値および第2の距離閾値は、同一である）。

#### 【0201】

いくつかの実施形態では、システムは、負の差異を取り扱うことが可能である。すなわち、オブジェクトまたはテキストが第1の深度および第2の深度の両方に表示されるための基準に適合する、ある重複が存在する。そのような実施形態では、ウェアラブルシステムは、コンテキスト手掛かりを使用して、最もシームレスな視認体験をユーザに提供するのである深度を決定することができる。例えば、最初に、ユーザに近接して現れるが、ユーザから離れるように急速に移動する、オブジェクトは、最初に、第1の深度に表示されるための基準に適合し得る。しかしながら、システムは、オブジェクトの場所の機動のため、オブジェクトを第2の深度に表示するのであることを決定し得る。

#### 【0202】

ウェアラブルシステムのいくつかの実施形態は、テキストを3つ以上の深度に表示可能である。そのような場合、第1および第2の深度間の第3、第4等の深度に対応する、中間距離閾値または距離の範囲が、含まれることができる。例えば、いくつかの実施形態では、テキストは、表音文字が、例えば、ディスプレイ220から約100cm~300cm離れるように現れるとき、第3の深度にレンダリングされてもよい。

#### 【0203】

ウェアラブルシステム200は、自動的に、またはユーザ入力に応答して、テキストを画像から識別または認識するように構成されることができる。テキストが自動的に識別される、実施形態では、ユーザは、画像をテキストとともに視認することができ、システムは、ユーザによるコマンドを伴わずに、本明細書に説明されるように、テキストを識別および表示することができる。テキストがユーザ入力に応答して識別される、実施形態では、ユーザは、種々のコマンドを使用して、テキストの識別または表示を開始することができる。例えば、コマンドは、口頭キュー、手のジェスチャ、頭部運動（例えば、點頭）、眼移動（例えば、瞬目）等であってもよい。

（テキストの表示特性を修正する例示的プロセス）

#### 【0204】

図17は、環境とのユーザの相互作用を促進するための感覚式アイウェアの例示的プロセスを図示する。プロセス1700は、ウェアラブルシステム200の1つ以上のコンポーネント（例えば、単独で、または組み合わせて、ローカル処理およびデータモジュール260、遠隔処理モジュール270）によって実施されることができる。

#### 【0205】

ブロック1704では、ウェアラブルシステムは、1つ以上のカメラを通して、光学信号を受信することができる。1つ以上のカメラは外向きに面した結像システム464の一部または車載カメラもしくはユーザ入力デバイス466等の別のコンピューティングデバイスの一部であってもよい。

10

20

30

40

50

## 【0206】

ブロック1708では、ウェアラブルシステムは、画像を信号から識別するステップを含むことができる。例えば、ウェアラブルシステムは、光学信号をヒト可読画像に変換することができる。いくつかの実施形態では、画像を信号から識別するステップはまた、例えば、1つ以上のオブジェクト認識装置708を使用して、光学文字認識(OCR)を画像に実施する等、画像のコンテンツを認識するステップを含むことができる。ある実施形態では、光学文字認識プロセスは、1つ以上の表音文字または表意文字のテキストまたは言語の可能性が高い候補を識別するステップを含む。光学文字認識プロセスは、種々のコンテキスト情報(例えば、コンテキスト手掛かり)を使用して、認識を実施してもよい。いくつかの例示的コンテキスト情報は、ユーザまたはユーザの近傍の他者が従事している

10

## 【0207】

図17を継続して参照すると、プロセス1700はさらに、ブロック1712に示されるように、画像が表音文字または表意文字を含むかどうかを決定するステップを含むことができる。いくつかの実施形態では、プロセス1700が、画像が表音文字または表意文字を備えていないことを決定する場合、プロセスは、ブロック1704に戻ることができる。プロセス1700が、画像が表音文字または表意文字を含むことを決定する場合、方法は、ブロック1716に継続する。

20

## 【0208】

ブロック1716では、ウェアラブルシステムは、表音文字または表意文字をテキストに変換することができる。これは、例えば、テキストを第1の言語と異なる第2の言語において表示するステップを含むことができる(図18および19を参照してさらに説明されるように)。いくつかの実施形態では、1つ以上の表音文字または表意文字(画像から)をテキストに変換するステップは、ユーザまたは別の人物からの入力またはコマンドの受信にตอบสนองして行われることができる。そのような入力またはコマンドは、例えば、口頭

30

## 【0209】

ブロック1720では、ウェアラブルシステムは、ディスプレイに、ユーザから複数の深度のうちの第1の深度に現れるようにテキストをレンダリングするように命令することができる。いくつかの実施形態では、テキストを表示するステップは、光学的に透過性の接眼レンズを通して、光をユーザに画像として伝送するステップを含む。接眼レンズは、本明細書に説明されるもののいずれかであることができる。例えば、光は、ユーザの眼の中に指向され、画像を眼内に形成し得る。ウェアラブルシステムは、本明細書に説明されるように、ファイバ走査プロジェクタまたは他のプロジェクタを使用してよい。いくつかの実施形態では、方法は、場所データをGPSシステム37(図2Bを参照して説明される)から受信してもよい。本場所データは、図16A-16Eを参照して本明細書にさらに説明されるように、システムが画像から抽出されたテキストを推測することに役立つために使用されることができる。

40

## 【0210】

ウェアラブルシステムはまた、テキストを修正し、修正されたテキストをレンダリングすることができる(例えば、ディスプレイ220からの光をユーザの眼に向かって投影する)。例えば、方法は、オリジナル表音文字または表意文字に対して、異なるフォント、フォントサイズ、色、背景または背景色、フォーマット、明確性のレベル、言語、または

50

明度において、テキストを表示することができる。いくつかの実施形態では、方法は、テキストを動画化するステップまたはテキストと相互作用する仮想オブジェクトを組み込むステップを含むことができる。

(表記のコンテンツを修正する実施例)

【0211】

テキストの表示特性を修正することに加え、またはその代替として、ウェアラブルシステムはまた、例えば、テキストをから1つの言語から別の言語に翻訳し、修正されたテキストを表示すること等によって、テキストのコンテンツを修正することができる。図18は、表記のコンテンツを修正することによって、ユーザが物理的環境内の表記を理解することを補助する実施例を図示し、表記は、ローカル言語からウェアラブルシステムのユーザが理解可能な標的言語に翻訳される。

10

【0212】

図18は、2つの場面1800aおよび1800bを図示する。場面1800aは、本明細書に説明されるHMDを装着せずに、ユーザによって知覚されることができる。場面1800bは、HMDを装着しているユーザによって知覚されることができる(例えば、ディスプレイ220を通して、説明される翻訳プロセスを伴わずに)。図示されるように、両場面1800aおよび1800bは、路上1802と、歩行者1804とを含む。場面1800aはまた、路上標識1810aおよび1820aを示し、これは、簡体字中国語表意文字を含む。標識1820aはまた、英語表意文字を含む。しかしながら、HMDのユーザ(図18には図示せず)は、英語話者であり得、中国語表意文字を理解し得ない。有利には、いくつかの実施形態では、ウェアラブルシステムは、路上標識1810aおよび1820b上のテキストを自動的に認識し、路上標識の外国語言語テキスト部分をユーザが理解する言語に変換することができる。ウェアラブルシステムはまた、場面1800bに示されるように、翻訳された表記を仮想画像として物理的標識にわたって提示することができる。故に、ユーザは、HMDが、下層中国語テキストが知覚されないような十分な明度を伴って、仮想画像を表示するため(英語テキストを用いて)、標識1810a、1820a内の中国語テキストを知覚せず、代わりに、標識1810b、1820bに示される英語テキストを知覚するであろう。

20

【0213】

HMD(例えば、ウェアラブルシステム200)は、図16A-17を参照して説明されるような類似技法を使用して、ユーザの環境内の標識を識別し、標識を認識することができる。いくつかの状況では、ウェアラブルシステム200は、標識の一部のみを翻訳するように構成されてもよい。例えば、ウェアラブルシステム200は、英語部分がユーザによって理解され得るため(例えば、ユーザの標的言語であるため)、中国語テキストを有する標識1820aの部分のみを翻訳し、英語テキスト(「GOLDSTAR」)を有する標識1820aの部分の翻訳しない。しかしながら、ユーザが英語および簡体字中国語の両方を読むことができるような、ユーザがバイリンガルである状況では、ウェアラブルシステム200は、標識1810aおよび1820a上のいずれのテキストも標識1810bおよび1820bに翻訳しないように構成されてもよい。

30

【0214】

図16A-16Eを参照して説明されるように、ウェアラブルシステム200は、標識の表示特性を調節するように構成されることができる。例えば、標識1820aの中国語部分の翻訳から生じたテキストは、標識1820a上のオリジナル中国語表意文字より長くなり得る。その結果、ウェアラブルシステムは、レンダリングされたテキスト(標識1820bに示されるように)がオリジナル標識の境界内に適合し得るように、翻訳されたテキスト(例えば、「ショッピングセンター」)のフォントサイズを縮小させてもよい。

40

【0215】

図18は、簡体字中国語および英語表意文字を示すが、これは、例証のためのものであって、限定ではない。ウェアラブルディスプレイシステム200の実施形態によって認識および変換される言語は、例えば、英語、中国語(簡体字または繁体字)、日本語、韓国

50

語、フランス語、スペイン語、ドイツ語、ロシア語、アラブ語、ロマンス諸語、インド・ヨーロッパ語族、シナ・チベット語族、アフロ・アジア語族、ヘブライ語、マレー・ポリネシア語族等の任意の言語を含むことができる。

(表記のコンテンツを修正する例示的プロセス)

【0216】

図19は、ユーザが物理的環境内の表記を理解することを補助する例示的プロセスを図示する。例示的プロセス1900は、ウェアラブルシステム200の1つ以上のコンポーネント(例えば、単独で、または組み合わせて、ローカル処理およびデータモジュール260、遠隔処理モジュール270)によって実施されてもよい。

【0217】

ブロック1910では、ウェアラブルシステムは、ユーザの環境の画像を受信することができる。画像は、外向きに面した結像システム464、ユーザ入力デバイス466、またはウェアラブルシステムの外部の別のデバイス上のカメラによって捕捉されることができる。画像は、静止画像、ビデオのフレーム、またはビデオであってもよい。

【0218】

ブロック1920では、ウェアラブルシステムは、画像を分析し、ユーザの環境内の表記を識別することができる。ウェアラブルシステムは、オブジェクト認識装置708を使用して、そのような識別を実施することができる。例えば、オブジェクト認識装置708は、オブジェクト上のテキストの存在を検出し、したがって、オブジェクトを標識として分類することができる、または表記の規則的境界を認識することができる(例えば、図18における長方形標識1810a、1810b)。

【0219】

ブロック1930では、ウェアラブルシステムは、表記1930上のテキストを認識することができる。例えば、ウェアラブルシステムは、標識上に存在する表意文字または表音文字を決定することができる。別の実施例として、ウェアラブルシステムは、テキストの言語を決定することができる。ウェアラブルシステムは、例えば、ユーザの場所、統語論、文法、テキストのスペル等、ユーザまたは標識と関連付けられたコンテキスト手掛かりに基づいて、そのような決定を行うことができる。ウェアラブルシステムはさらに、ブロック1930において、テキストの意味を決定することができる(例えば、辞書をルックアップすることによって)。

【0220】

ブロック1940では、ウェアラブルシステムは、テキストの少なくとも一部を標的言語に変換することができる。標的言語は、ユーザの選好またはユーザの人口統計情報に基づいて決定されてもよい。例えば、標的言語は、ユーザの母国と関連付けられた公用語、ユーザの母国語、ユーザによって最も頻繁に使用される言語、またはユーザが発話した言語(例えば、ウェアラブルシステムへの音声コマンドまたは別のユーザとの会話において)等であってもよい。標的言語はまた、ユーザの選好にしたがって設定されることができる。例えば、ユーザは、ユーザの母国語がフランス語であっても、標識が英語に翻訳されることを好み得る。

【0221】

随意的ブロック1950では、ウェアラブルシステムは、テキストと関連付けられた表示特性を修正することができる。例えば、ウェアラブルシステムは、焦点インジケータをテキスト(またはテキストと関連付けられた背景)に追加し、かつテキストのフォントサイズまたは色を変化させることができる。表示特性の例示的修正は、図16A-17を参照してさらに説明される。

【0222】

随意的ブロック1960では、ウェアラブルシステムは、複合現実デバイスによって、テキストを標的言語においてレンダリングさせることができる。MRデバイスは、本明細書に説明されるHMDであってもよい。表示特性が、修正される場合、ウェアラブルシステムはまた、修正された表示特性をレンダリングさせることができる。テキストの一部の

10

20

30

40

50

みが、標的言語に翻訳される状況では、ウェアラブルシステムは、翻訳されたテキストの一部のみを表示するか、または翻訳された部分と翻訳されていないオリジナルテキストの部分の両方を表示するかのいずれかであることができる。修正されたテキストは、オリジナルテキストがユーザのビューからオクルードされ得るように、物理的表記上のオリジナルテキストにわたってレンダリングされてもよい。

【 0 2 2 3 】

図 1 8 および 1 9 における実施例は、表記上のテキストを翻訳することを参照して説明されるが、類似技法はまた、他のタイプの媒体（書籍、テレビ、コンピュータモニタ等）に具現化されるテキストにも適用されることができる。

（手話に関連する付加的側面）

10

【 0 2 2 4 】

手話における感覚式アイウェアの用途の付加的側面が、以下にさらに提供される。

【 0 2 2 5 】

第 1 の側面では、拡張現実システムを通して、手話から変換されたテキストを提供するための方法であって、結像システムを備える、拡張現実（AR）システムの制御下、結像システムを介して、画像情報を捕捉するステップと、画像情報内のジェスチャを検出するステップであって、ジェスチャは、手話認識のための候補である、ステップと、検出されたジェスチャ内の手話を認識するステップと、認識された手話をテキストに変換するステップと、変換されたテキストを表示するステップとを含む、方法。

20

【 0 2 2 6 】

第 2 の側面では、変換されたテキストに関する補助情報の要求を受信するステップと、要求される変換されたテキストと関連付けられた補助情報を読み出すステップと、ARシステムを使用して、補助情報を表示するステップと、変換されたテキストまたは補助情報の表示を放逐するための条件を検出するステップと、変換されたテキストまたは補助情報の表示を放逐するステップとをさらに含む、側面 1 に記載の方法。

【 0 2 2 7 】

第 3 の側面では、変換されたテキストまたは補助情報の表示を放逐するための条件は、ユーザインターフェース相互作用に基づく、側面 2 に記載の方法。

【 0 2 2 8 】

第 4 の側面では、ユーザインターフェース相互作用は、少なくとも部分的に、ARシステムのユーザの眼移動に基づく、側面 3 に記載の方法。

30

【 0 2 2 9 】

第 5 の側面では、変換されたテキストまたは補助情報の表示を放逐するための条件は、少なくとも部分的に、持続時間に基づく、側面 2 - 4 のいずれか 1 項に記載の方法。

【 0 2 3 0 】

第 6 の側面では、変換されたテキストまたは補助情報の表示を放逐するための条件は、少なくとも部分的に、付加的な手話ジェスチャの変換または付加的補助情報の受信に基づく、側面 2 - 5 のいずれか 1 項に記載の方法。

【 0 2 3 1 】

第 7 の側面では、変換されたテキストまたは補助情報の放逐された表示を再表示するための条件を検出するステップと、変換されたテキストまたは補助情報の放逐された表示を再表示するステップとをさらに含む、側面 2 - 6 のいずれか 1 項に記載の方法。

40

【 0 2 3 2 】

第 8 の側面では、認識された手話をテキストに変換するステップは、深層学習技法を適用するステップを含む、側面 1 - 7 のいずれか 1 項に記載の方法。

【 0 2 3 3 】

第 9 の側面では、深層学習技法は、ニューラルネットワークを含む、側面 8 に記載の方法。

【 0 2 3 4 】

第 1 0 の側面では、ARシステムは、手話認識およびテキスト変換において、手話辞書

50

を使用する、側面 1 - 9 のいずれか 1 項に記載の方法。

【 0 2 3 5 】

第 1 1 の側面では、A R システムは、A R システムのユーザにとって外国語である、手話を認識する、側面 1 - 1 0 のいずれか 1 項に記載の方法。

【 0 2 3 6 】

第 1 2 の側面では、A R システムは、候補手話のリストを吟味することによって、手話を認識し、リストは、少なくとも部分的に、A R システムの場所に基づいて、優先順位が付けられる、側面 1 1 に記載の方法。

【 0 2 3 7 】

第 1 3 の側面では、A R システムは、候補手話のリストを吟味することによって、手話を認識し、リストは、少なくとも部分的に、A R システムの環境内で検出された発話された言語に基づいて、優先順位が付けられる、側面 1 1 - 1 2 のいずれか 1 項に記載の方法。

10

【 0 2 3 8 】

第 1 4 の側面では、手話を翻訳するための拡張現実 (A R) 装置であって、A R ディスプレイと、結像システムと、コンピュータ実行可能命令およびデータを記憶するように構成される、データ記憶装置と、データ記憶装置と通信するプロセッサとを備え、コンピュータ実行可能命令は、実行されると、プロセッサに、結像システムによって捕捉された画像情報を受信させ、受信された画像またはビデオ情報内のジェスチャを検出させ、検出されたジェスチャ内の手話を認識させ、認識された手話を A R 装置のユーザによって理解される言語に翻訳させ、A R ディスプレイを使用して、翻訳された手話と関連付けられた情報を表示させる、装置。

20

【 0 2 3 9 】

第 1 5 の側面では、コンピュータ実行可能命令は、実行されると、さらにプロセッサに、翻訳された手話に関する補助情報の要求を受信させ、要求される手話に関連する補助情報を読み出させ、A R ディスプレイを使用して、読み出された補助情報を表示させる、側面 1 4 に記載の装置。

【 0 2 4 0 】

第 1 6 の側面では、プロセッサは、ジェスチャを検出し、遠隔プロセッサが、ジェスチャを検出し、手話を認識するために、受信された画像情報を通信ネットワークを通して遠隔プロセッサに伝送することによって、手話を認識する、側面 1 4 - 1 5 のいずれか 1 項に記載の装置。

30

【 0 2 4 1 】

第 1 7 の側面では、結像システムは、複数のカメラまたは広角カメラを備える、側面 1 4 - 1 6 のいずれか 1 項に記載の装置。

【 0 2 4 2 】

第 1 8 の側面では、プロセッサはさらに、検出されたジェスチャ源を決定し、検出されたジェスチャ源が A R 装置のユーザであることの決定に応じて、表示のために、翻訳された手話を別のデバイスに伝送するように構成される、側面 1 4 - 1 7 のいずれか 1 項に記載の装置。

40

【 0 2 4 3 】

第 1 9 の側面では、オーディオ増幅器をさらに備え、プロセッサはさらに、オーディオ増幅器を通して、翻訳された手話をオーディオとして提示するようにプログラムされる、側面 1 4 - 1 8 のいずれか 1 項に記載の装置。

【 0 2 4 4 】

第 2 0 の側面では、プロセッサはさらに、オーディオ増幅器を通して、補助情報をオーディオとして提示するように構成される、側面 1 9 に記載の装置。

【 0 2 4 5 】

第 2 1 の側面では、A R 装置のユーザによって理解される言語は、認識された手話と異なる手話を含む、側面 1 4 - 2 0 のいずれか 1 項に記載の装置。

50

## 【 0 2 4 6 】

第 2 2 の側面では、1 つ以上の手話を伴う遠隔通信を促進するための拡張現実 ( A R ) システムであって、それぞれ、A R ディスプレイと、結像システムと、通信ネットワークを経由して通信するための通信システムとを備える、複数のウェアラブル A R デバイスと、コンピュータ実行可能命令およびデータを記憶するように構成される、1 つ以上のデータ記憶装置と、データ記憶装置と通信する、1 つ以上のプロセッサとを備え、コンピュータ実行可能命令は、実行されると、複数のウェアラブル A R デバイス内の第 1 のウェアラブル A R デバイスの結像システムによって捕捉された画像情報を受信し、受信された画像情報内の手話を検出し、検出された手話をテキストに変換し、通信ネットワークを通して、変換されたテキストを複数のウェアラブル A R デバイス内の第 2 のウェアラブル A R デ

10

## 【 0 2 4 7 】

第 2 3 の側面では、第 2 のウェアラブル A R デバイスはさらに、第 1 のユーザの世界マップを表示する、側面 2 2 に記載のシステム。

## 【 0 2 4 8 】

第 2 4 の側面では、第 1 のユーザの世界マップは、第 1 のユーザのアバタを備える、側面 2 3 に記載のシステム。

## 【 0 2 4 9 】

第 2 5 の側面では、複数のウェアラブル A R デバイスはそれぞれ、1 つ以上のデータ記憶装置と、1 つ以上のプロセッサとを含み、プロセッサ機能性は、ローカルプロセッサによって実施される、側面 2 2 - 2 4 のいずれか 1 項に記載のシステム。

20

## 【 0 2 5 0 】

第 2 6 の側面では、手話認識のためのウェアラブルシステムであって、仮想コンテンツをユーザに提示するように構成される、頭部搭載型ディスプレイと、ユーザの環境を結像するように構成される、結像システムと、頭部搭載型ディスプレイおよび結像システムと通信し、結像システムによって捕捉された画像を受信し、オブジェクト認識装置を用いて、画像内のジェスチャを検出し、手話におけるジェスチャの意味を認識し、ユーザと関連付けられたコンテキスト情報に基づいて、標的言語を識別し、認識された意味に基づいて、ジェスチャを標的言語に翻訳し、少なくとも部分的に、標的言語へのジェスチャの翻訳

30

## 【 0 2 5 1 】

第 2 7 の側面では、結像システムは、ユーザの周囲を結像するように構成される広角カメラのうちの 1 つ以上のものを備える、側面 2 6 に記載のウェアラブルシステム。

## 【 0 2 5 2 】

第 2 8 の側面では、ハードウェアプロセッサはさらに、ジェスチャと関連付けられた補助情報にアクセスするようにプログラムされ、頭部搭載型ディスプレイによってレンダリングされた仮想コンテンツは、補助情報を備える、側面 2 6 - 2 7 のいずれか 1 項に記載のウェアラブルシステム。

40

## 【 0 2 5 3 】

第 2 9 の側面では、ユーザと関連付けられたコンテキスト情報に基づいて、標的言語を識別するために、ハードウェアプロセッサは、ウェアラブルシステムによって捕捉されるようなユーザの発話、ユーザの場所、またはある言語を標的言語として選択するユーザからの入力の中の少なくとも 1 つに基づいて、標的言語をユーザによって理解される言語として設定するようにプログラムされる、側面 2 6 - 2 8 のいずれか 1 項に記載のウェアラブルシステム。

## 【 0 2 5 4 】

第 3 0 の側面では、ハードウェアプロセッサは、標的言語が発話された言語であるかど

50

うかを決定し、標的言語が発話された言語であることの決定に回答して、標的言語において、翻訳されたジェスチャと関連付けられた発話のオーディオストリームを再生するようにプログラムされる、側面 26 - 29 のいずれか 1 項に記載のウェアラブルシステム。

【0255】

第 31 の側面では、ハードウェアプロセッサは、標的言語が別の手話であるかどうかを決定し、標的言語が別の手話であることの決定に回答して、他の手話における別のジェスチャのグラフィックをジェスチャの翻訳として提示するようにプログラムされる、側面 26 - 29 のいずれか 1 項に記載のウェアラブルシステム。

【0256】

第 32 の側面では、手話におけるジェスチャの意味を認識するために、ハードウェアプロセッサは、深層ニューラルネットワーク技法を結像システムによって捕捉された画像の一部に適用するようにプログラムされる、側面 26 - 31 のいずれか 1 項に記載のウェアラブルシステム。

【0257】

第 33 の側面では、ハードウェアプロセッサはさらに、少なくとも部分的に、ユーザの場所に基づいて、手話を候補手話のリストから識別するようにプログラムされる、側面 26 - 32 のいずれか 1 項に記載のウェアラブルシステム。

【0258】

第 34 の側面では、認識された意味に基づいて、ジェスチャを標的言語に翻訳するために、ハードウェアプロセッサは、ジェスチャを標的言語におけるテキスト表現に変換するようにプログラムされる、側面 26 - 33 のいずれか 1 項に記載のウェアラブルシステム。

【0259】

第 35 の側面では、ハードウェアプロセッサは、検出されたジェスチャ源を決定し、検出されたジェスチャ源がウェアラブルシステムのユーザであることの決定に応じて、標的言語におけるジェスチャの翻訳を別のユーザのウェアラブルシステムに通信するようにプログラムされる、側面 26 - 34 のいずれか 1 項に記載のウェアラブルシステム。

【0260】

第 36 の側面では、ハードウェアプロセッサは、仮想コンテンツを頭部搭載型ディスプレイによる表示から放逐することからの条件を検出し、条件の検出に回答して、仮想コンテンツを頭部搭載型ディスプレイによる表示から除去するようにプログラムされる、側面 26 - 35 のいずれか 1 項に記載のウェアラブルシステム。

【0261】

第 37 の側面では、条件は、持続時間、ユーザの手のジェスチャ、またはユーザ入力デバイスからの入力のうちの少なくとも 1 つを含む、側面 36 に記載のウェアラブルシステム。

【0262】

第 38 の側面では、画像は、ビデオの 1 つ以上のフレームを含む、側面 26 - 37 のいずれか 1 項に記載のウェアラブルシステム。

【0263】

第 39 の側面では、手話認識のための方法であって、結像システムによって捕捉された画像を受信するステップと、画像を分析し、ユーザのジェスチャを検出するステップと、少なくとも部分的に、検出されたジェスチャに基づいて、手話における通信の存在を検出するステップと、手話におけるジェスチャの意味を認識するステップと、ジェスチャが翻訳されるであろう標的言語を識別するステップと、認識された意味に基づいて、ジェスチャを標的言語に翻訳するステップと、少なくとも部分的に、標的言語へのジェスチャの翻訳に基づいて、仮想コンテンツを生成し、頭部搭載型ディスプレイに、仮想コンテンツをユーザにレンダリングさせるステップとを含む、方法。

【0264】

第 40 の側面では、画像は、複合現実コンテンツを提示するように構成される、第 1 の

10

20

30

40

50

ウェアラブルシステムから受信される一方、仮想コンテンツは、レンダリングのために、第2のウェアラブルシステムに通信され、第1のウェアラブルシステムおよび第2のウェアラブルシステムは、複合現実コンテンツをユーザに提示するように構成される、側面39に記載の方法。

【0265】

第41の側面では、認識された意味に基づいて、ジェスチャを標的言語に翻訳するステップは、ジェスチャを標的言語におけるテキスト表現に変換するステップを含む、側面39に記載の方法。

【0266】

第42の側面では、仮想コンテンツは、標的言語におけるテキスト表現または標的言語における別のものを図示するグラフィックを備える、側面39-41のいずれか1項に記載の方法。

10

【0267】

第43の側面では、手話におけるジェスチャの意味を認識するステップは、深層ニューラルネットワーク技法を結像システムによって捕捉された画像の一部に適用するステップを含む、側面39-42のいずれか1項に記載の方法。

【0268】

第44の側面では、手話における通信の存在を検出するステップは、手話を候補手話のリストから識別するステップと、検出されたジェスチャが手話における表現に対応することを決定するステップとを含む、側面39-43のいずれか1項に記載の方法。

20

【0269】

第45の側面では、検出されたジェスチャが手話における表現に対応することを決定するステップは、ジェスチャをジェスチャを行っている人物の唇の移動およびユーザがジェスチャを行っている間に捕捉されたオーディオデータと関連させて分析するステップを含む、側面44に記載の方法。

(テキスト修正に関連する付加的側面)

【0270】

感覚式アイウェアによるテキストの特性を修正する、付加的側面が、以下にさらに説明される。

【0271】

30

第1の側面では、拡張現実画像コンテンツを投影するように構成される、頭部搭載型ディスプレイデバイスであって、ユーザの頭部上に装着可能であるように構成され、ディスプレイをユーザの眼の正面に支持するように構成される、フレームと、光学信号を受信するように構成される、1つ以上のカメラと、信号を1つ以上のカメラから受信し、画像を信号から識別し、画像がテキスト(例えば、1つ以上の表音文字または表意文字)を備えるかどうかを決定し、テキストを修正されるテキストに変換し、ディスプレイに、修正されたテキストをレンダリングするように命令するように構成される、処理電子機器とを備える、ディスプレイデバイス。

【0272】

第2の側面では、ディスプレイは、1つ以上の光源と、光をユーザの眼の中に指向し、画像を眼内に形成するように構成される、1つ以上の導波管スタックとを備える、側面1に記載の頭部搭載型ディスプレイデバイス。

40

【0273】

第3の側面では、1つ以上の光源は、光を導波管スタックの中に指向するように構成される、側面2に記載の頭部搭載型ディスプレイデバイス。

【0274】

第4の側面では、1つ以上の光源は、ファイバ走査プロジェクタを備える、側面2-3のいずれか1項に記載の頭部搭載型ディスプレイデバイス。

【0275】

第5の側面では、1つ以上のカメラは、1つ以上のビデオカメラを備える、側面1-4

50

のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0276】

第 6 の側面では、処理電子機器は、光学文字認識アルゴリズムを使用して、画像内の 1 つ以上の表音文字または表意文字をテキストに変換するように構成される、側面 1 - 5 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0277】

第 7 の側面では、処理電子機器は、データベースにアクセスし、1 つ以上の表音文字または表意文字のテキストまたは言語の可能性が高い候補を識別するように構成される、側面 6 に記載の頭部搭載型ディスプレイデバイス。

【0278】

第 8 の側面では、処理電子機器は、ユーザが従事しているアクティビティ、ユーザの地理的場所、ユーザの進行速度、ユーザの高度、ディスプレイによって検出される周囲雑音の音量またはタイプ、ディスプレイによって受信されたエリア内の可視または他の光のレベルまたはタイプ、ディスプレイによって検出された温度または気候、ユーザからのテキストの知覚された距離、もしくはディスプレイによって検出された単語のカテゴリのうちの 1 つ以上のものと関連付けられた入力を受信するように構成される、側面 6 - 7 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0279】

第 9 の側面では、GPS システムをさらに備える、側面 1 - 8 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0280】

第 10 の側面では、修正されたテキストは、第 1 のフォントテキストのサイズと異なる第 2 のフォントサイズであって、第 2 のフォントサイズは、第 1 のフォントサイズより大きくあり得る、側面 1 - 9 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0281】

第 11 の側面では、修正されたテキストは、テキストよりユーザに読みやすい、側面 1 - 10 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0282】

第 12 の側面では、処理電子機器は、グラフィカル要素をテキストに追加し、部分的に、修正されたテキストを形成するように構成される、側面 1 - 11 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0283】

第 13 の側面では、処理電子機器は、テキストの 1 つ以上の表音文字または表意文字を 1 つ以上の表音文字または表意文字の第 1 のフォントと異なる第 2 のフォントにおいて表示するように構成される、側面 1 - 12 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0284】

第 14 の側面では、処理電子機器は、ユーザに頭部搭載型ディスプレイを伴わずに見えるであろうものに対して、テキストの 1 つ以上の表音文字または表意文字を拡大するように構成される、側面 1 - 13 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0285】

第 15 の側面では、処理電子機器は、境界領域を表示するように構成され、境界領域は、内側領域を境界する、側面 1 - 14 のいずれか 1 項に記載の頭部搭載型ディスプレイデバイス。

【0286】

第 16 の側面では、処理電子機器は、1 つ以上の表音文字または表意文字を内側領域内に表示するように構成される、側面 15 に記載の頭部搭載型ディスプレイデバイス。

【0287】

第 17 の側面では、処理電子機器は、ユーザが頭部搭載型ディスプレイを伴わずに 1 つ以上の表音文字または表意文字を読み取るであろう、第 1 の背景と異なる第 2 の背景に対

10

20

30

40

50

して、テキストの1つ以上の表音文字または表意文字を表示するように構成される、側面1-16のいずれか1項に記載の頭部搭載型ディスプレイデバイス。

【0288】

第18の側面では、第2の背景は、モノクロ背景を含む、側面17に記載の頭部搭載型ディスプレイデバイス。

【0289】

第19の側面では、モノクロ背景は、白色を含む、側面18に記載の頭部搭載型ディスプレイデバイス。

【0290】

第20の側面では、第1の背景は、ユーザに頭部搭載型ディスプレイを伴わずに見えるであろうものを含む、側面17-19のいずれか1項に記載の頭部搭載型ディスプレイデバイス。

10

【0291】

第21の側面では、テキストは、テキストエディタによって編集可能であるように適合される、側面1-20のいずれか1項に記載の頭部搭載型ディスプレイデバイス。

【0292】

側面1-21は、頭部搭載型ディスプレイを参照して説明されるが、これらの側面に説明される類似機能はまた、図2Aを参照して説明される頭部搭載型デバイスまたはウェアラブルシステムを用いて実装されることができる。さらに、ディスプレイは、複数の深度平面を備えることができ、頭部搭載型デバイスは、少なくとも部分的に、ユーザの眼の能力に基づいて、修正されたテキストをレンダリングするための深度平面を識別するように構成される。

20

【0293】

第22の側面では、頭部搭載型ディスプレイを使用して、拡張現実画像コンテンツを投影させるための方法であって、ハードウェアプロセッサの制御下で、光学信号を1つ以上のカメラから受信するステップと、光学文字認識モジュールを使用して、画像を信号から識別するステップと、画像が1つ以上の表音文字または表意文字を含むかどうかを決定するステップと、1つ以上の表音文字または表意文字をテキストに変換するステップと、テキストを頭部搭載型ディスプレイ上に表示するステップとを含み、テキストを表示するステップは、光学的に透過性の接眼レンズを通して、光をユーザに画像として伝送するステップを含む、方法。

30

【0294】

第23の側面では、光をユーザの眼の中に指向し、画像を眼内に形成するステップをさらに含む、側面22に記載の方法。

【0295】

第24の側面では、ファイバ走査プロジェクタを使用して、光を接眼レンズの中に指向するステップをさらに含む、側面22-23のいずれか1項に記載の方法。

【0296】

第25の側面では、光学文字認識モジュールを使用するステップは、1つ以上の表音文字または表意文字のテキストまたは言語の可能性が高い候補を識別するステップを含む、側面22-24のいずれか1項に記載の方法。

40

【0297】

第26の側面では、光学文字認識モジュールを使用するステップは、ユーザが従事しているアクティビティ、ユーザの地理的場所、ユーザの進行速度、ユーザの高度、ディスプレイによって受信される周囲雑音の音量またはタイプ、ディスプレイによって検出されたエリア内の可視または他の光のレベルまたはタイプ、ディスプレイによって検出された温度または気候、ユーザからの1つ以上の表音文字または表意文字の知覚された距離、もしくはディスプレイによって検出された単語のカテゴリのうちの1つ以上のものと関連付けられた情報を含む、入力を受信するステップを含む、側面22-25のいずれか1項に記載の方法。

50

## 【0298】

第27の側面では、1つ以上の表音文字または表意文字をテキストに変換するステップは、テキストを1つ以上の表音文字または表意文字と関連付けられた第1の言語と異なる第2の言語において表示するステップを含む、側面22 - 26のいずれか1項に記載の方法。第27の側面のいくつかの実装では、本方法は、テキストを第2の言語に翻訳するステップを含む。

## 【0299】

第28の側面では、場所データをGPSシステムから受信するステップをさらに含む、側面22 - 27のいずれか1項に記載の方法。

## 【0300】

第29の側面では、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイ上に表示するステップは、1つ以上の表音文字または表意文字を1つ以上の表音文字または表意文字の第1のフォントサイズと異なる第2のフォントサイズにおいて表示するステップを含む、側面22 - 28のいずれか1項に記載の方法。

## 【0301】

第30の側面では、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイ上に表示するステップは、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイを伴わない場合よりユーザに読みやすく表示するステップを含む、側面22 - 29のいずれか1項に記載の方法。

## 【0302】

第31の側面では、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイ上に表示するステップは、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイを伴わずにユーザに現れるであろう場合より大きいフォントサイズにおいて表示するステップを含む、側面22 - 30のいずれか1項に記載の方法。

## 【0303】

第32の側面では、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイ上に表示するステップは、1つ以上の表音文字または表意文字を1つ以上の表音文字または表意文字の第1のフォントと異なる第2のフォントにおいて表示するステップを含む、側面22 - 31のいずれか1項に記載の方法。

## 【0304】

第33の側面では、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイ上に表示するステップは、ユーザに頭部搭載型ディスプレイを伴わずに見えるであろうものに対して1つ以上の表音文字または表意文字を拡大するステップを含む、側面22 - 32のいずれか1項に記載の方法。

## 【0305】

第34の側面では、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイ上に表示するステップは、境界領域を表示するステップを含み、境界領域は、内側領域を境界する、側面22 - 33のいずれか1項に記載の方法。

## 【0306】

第35の側面では、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイ上に表示するステップは、1つ以上の表音文字または表意文字を内側領域内に表示するステップを含む、側面34に記載の方法。

## 【0307】

第36の側面では、1つ以上の表音文字または表意文字を頭部搭載型ディスプレイ上に表示するステップは、ユーザが頭部搭載型ディスプレイを伴わずに1つ以上の表音文字または表意文字を読み取るであろう第1の背景と異なる第2の背景に対して、1つ以上の表音文字または表意文字を表示するステップを含む、側面22 - 35のいずれか1項に記載の方法。

## 【0308】

第37の側面では、第2の背景は、モノクロ背景を含む、側面36に記載の方法。

10

20

30

40

50

## 【0309】

第38の側面では、モノクロ背景は、白色を含む、側面37に記載の方法。

## 【0310】

第39の側面では、第1の背景は、ユーザに頭部搭載型ディスプレイを伴わずに見えるであろうものを含む、側面36 - 38のいずれか1項に記載の方法。

## 【0311】

第40の側面では、テキストは、テキストエディタによって編集可能であるように適合される、側面22 - 39のいずれか1項に記載の方法。

## 【0312】

第41の側面では、1つ以上の表音文字または表意文字をテキストに変換するステップは、入力をユーザから受信するステップを含む、側面22 - 40のいずれか1項に記載の方法。

10

## 【0313】

第42の側面では、入力をユーザから受信するステップは、口頭コマンド、手のジェスチャ、頭部の運動、またはユーザの眼の一方もしくは両方の移動のうちの1つ以上のものを受信するステップを含む、側面41に記載の方法。

## 【0314】

第43の側面では、テキストは、1つ以上の表音文字または表意文字が第1の距離閾値より近くに現れる場合、第2の深度より近くに現れる、第1の深度に表示される、側面22 - 42のいずれか1項に記載の方法。

20

## 【0315】

第44の側面では、テキストは、1つ以上の表音文字または表意文字が第2の距離閾値からより遠く離れて現れる場合、第1の深度より遠く離れて現れる、第2の深度に表示される、側面22 - 43のいずれか1項に記載の方法。

## 【0316】

第45の側面では、テキストは、1つ以上の表音文字または表意文字が、第1の距離閾値より遠く離れて、かつ第2の距離閾値より近くに現れる場合、第1の深度より遠く離れているが、第2の深度より近くに現れる、第3の深度に表示される、側面43 - 44のいずれか1項に記載の方法。

## 【0317】

第46の側面では、第1の距離閾値は、80cmである、側面43 - 45のいずれか1項に記載の方法。

30

## 【0318】

第47の側面では、第2の距離閾値は、600cmである、側面43 - 46のいずれか1項に記載の方法。

## 【0319】

第48の側面では、第2の距離閾値と第1の距離閾値との間の差異は、100m未満である、側面43 - 47のいずれか1項に記載の方法。

(表記修正に関連する付加的側面)

## 【0320】

第1の側面では、拡張現実システムであって、外向きに面した結像システムと、外向きに面した結像システムによって取得される画像を記憶するように構成される、非一過性メモリと、外向きに面した結像システムによって取得される拡張現実システムのユーザの環境の画像を受信し、画像を分析し、ユーザの環境内の表記を識別し、表記上のテキストを認識し、テキストの少なくとも一部を標的言語に変換し、ディスプレイに、変換されたテキストをユーザにレンダリングするように命令するようにプログラムされる、ハードウェアプロセッサとを備える、拡張現実システム。

40

## 【0321】

第2の側面では、ハードウェアプロセッサは、テキストと関連付けられた表示特性を修正するようにプログラムされる、側面1に記載の拡張現実システム。

50

## 【 0 3 2 2 】

第3の側面では、テキストの少なくとも一部を標的言語に変換するために、ハードウェアプロセッサは、表記上のテキストの言語を識別し、言語を標的言語に変換するようにプログラムされる、側面1または2に記載の拡張現実システム。

## 【 0 3 2 3 】

第4の側面では、ハードウェアプロセッサは、少なくとも部分的に、ユーザの場所に基づいて、標的言語を決定するようにプログラムされる、側面1 - 3のいずれか1項に記載の拡張現実システム。

## 【 0 3 2 4 】

第5の側面では、表記上のテキストを認識するために、ハードウェアプロセッサは、標的言語におけるテキストを認識するようにプログラムされる、側面1 - 4のいずれか1項に記載の拡張現実システム。

## 【 0 3 2 5 】

第6の側面では、ハードウェアプロセッサは、標的言語におけるテキストを変換しないようにプログラムされる、側面5に記載の拡張現実システム。

(他の考慮点)

## 【 0 3 2 6 】

本明細書に説明される、ならびに/または添付される図に描写されるプロセス、方法、およびアルゴリズムはそれぞれ、具体的かつ特定のコンピュータ命令を実行するように構成される、1つ以上の物理的コンピューティングシステム、ハードウェアコンピュータプロセッサ、特定用途向け回路、もしくは電子ハードウェアによって実行される、コードモジュールにおいて具現化され、それによって完全もしくは部分的に自動化され得る。例えば、コンピューティングシステムは、具体的コンピュータ命令とともにプログラムされた汎用コンピュータ(例えば、サーバ)または専用コンピュータ、専用回路等を含むことができる。コードモジュールは、実行可能プログラムにコンパイルおよびリンクされ得る、動的リンクライブラリ内にインストールされ得る、または解釈されるプログラミング言語において書き込まれ得る。いくつかの実装では、特定の動作および方法が、所与の機能に特有の回路によって実施され得る。

## 【 0 3 2 7 】

さらに、本開示の機能性のある実装は、十分に数学的、コンピュータ的、または技術的に複雑であるため、(適切な特殊化された実行可能命令を利用する)特定用途向けハードウェアまたは1つ以上の物理的コンピューティングデバイスは、例えば、関与する計算の量もしくは複雑性に起因して、または結果を実質的にリアルタイムで提供するために、機能性を実施する必要がある。例えば、動画またはビデオは、多くのフレームを含み、各フレームは、数百万のピクセルを有し得、具体的にプログラムされたコンピュータハードウェアは、商業的に妥当な時間量において所望の画像処理タスクまたは用途を提供するようにビデオデータを処理する必要がある。

## 【 0 3 2 8 】

コードモジュールまたは任意のタイプのデータは、ハードドライブ、ソリッドステートメモリ、ランダムアクセスメモリ(RAM)、読取専用メモリ(ROM)、光学ディスク、揮発性もしくは不揮発性記憶装置、同一物の組み合わせ、または同等物を含む、物理的コンピュータ記憶装置等の任意のタイプの非一過性コンピュータ可読媒体上に記憶され得る。本方法およびモジュール(またはデータ)はまた、無線ベースおよび有線/ケーブルベースの媒体を含む、種々のコンピュータ可読伝送媒体上で生成されたデータ信号として(例えば、搬送波または他のアナログもしくはデジタル伝搬信号の一部として)伝送され得、種々の形態(例えば、単一もしくは多重化アナログ信号の一部として、または複数の離散デジタルパケットもしくはフレームとして)をとり得る。開示されるプロセスまたはプロセスステップの結果は、任意のタイプの非一過性有形コンピュータ記憶装置内に持続的もしくは別様に記憶され得る、またはコンピュータ可読伝送媒体を介して通信され得る。

。

10

20

30

40

50

## 【0329】

本明細書に説明される、および/または添付される図に描写されるフロー図における任意のプロセス、ブロック、状態、ステップ、もしくは機能性は、プロセスにおいて具体的機能（例えば、論理もしくは算術）またはステップを実装するための1つ以上の実行可能命令を含む、コードモジュール、セグメント、またはコードの一部を潜在的に表すものとして理解されたい。種々のプロセス、ブロック、状態、ステップ、または機能性は、組み合わせられる、再配列される、追加される、削除される、修正される、または別様に本明細書に提供される例証的实施例から変更されることができる。いくつかの実施形態では、付加的または異なるコンピューティングシステムもしくはコードモジュールが、本明細書に説明される機能性のいくつかまたは全てを実施し得る。本明細書に説明される方法およびプロセスはまた、任意の特定のシーケンスに限定されず、それに関連するブロック、ステップ、または状態は、適切な他のシーケンスで、例えば、連続して、並行して、またはある他の様式で実施されることができる。タスクまたはイベントが、開示される例証的实施形態に追加される、またはそれから除去され得る。さらに、本明細書に説明される実装における種々のシステムコンポーネントの分離は、例証を目的とし、全ての実装においてそのような分離を要求するものとして理解されるべきではない。説明されるプログラムコンポーネント、方法、およびシステムは、概して、単一のコンピュータ製品においてともに統合される、または複数のコンピュータ製品にパッケージ化され得ることを理解されたい。多くの実装変形例が、可能である。

10

## 【0330】

本プロセス、方法、およびシステムは、ネットワーク（または分散）コンピューティング環境において実装され得る。ネットワーク環境は、企業全体コンピュータネットワーク、イントラネット、ローカルエリアネットワーク（LAN）、広域ネットワーク（WAN）、パーソナルエリアネットワーク（PAN）、クラウドコンピューティングネットワーク、クラウドソースコンピューティングネットワーク、インターネット、およびワールドワイドウェブを含む。ネットワークは、有線もしくは無線ネットワークまたは任意の他のタイプの通信ネットワークであり得る。

20

## 【0331】

本開示のシステムおよび方法は、それぞれ、いくつかの革新的側面を有し、そのうちのいかなるものも、本明細書に開示される望ましい属性に単独で関与しない、またはそのために要求されない。上記に説明される種々の特徴およびプロセスは、相互に独立して使用され得る、または種々の方法で組み合わせられ得る。全ての可能な組み合わせおよび副次的組み合わせが、本開示の範囲内に該当することが意図される。本開示に説明される実装の種々の修正が、当業者に容易に明白であり得、本明細書に定義される一般原理は、本開示の精神または範囲から逸脱することなく、他の実装に適用され得る。したがって、請求項は、本明細書に示される実装または実施形態に限定されることを意図されず、本明細書に開示される本開示、原理、および新規の特徴と一貫する最も広い範囲を与えられるべきである。

30

## 【0332】

別個の実装または実施形態の文脈において本明細書に説明されるある特徴はまた、単一の実装または実施形態における組み合わせにおいて実装されることができる。逆に、単一の実装または実施形態の文脈において説明される種々の特徴もまた、複数の実装または実施形態において別個に、もしくは任意の好適な副次的組み合わせにおいて実装されることができる。さらに、特徴がある組み合わせにおいて作用するものとして上記に説明され、さらに、そのようなものとして最初に請求され得るが、請求される組み合わせからの1つ以上の特徴は、いくつかの場合では、組み合わせから削除されることができ、請求される組み合わせは、副次的組み合わせまたは副次的組み合わせの変形例を対象とし得る。いかなる単一の特徴または特徴のグループも、あらゆる実施形態に必要なもしくは必須ではない。

40

## 【0333】

50

とりわけ、「～できる (can)」、「～し得る (could)」、「～し得る (might)」、「～し得る (may)」、「例えば (e.g.)」、および同等物等、本明細書で使用される条件文は、別様に具体的に記載されない限り、または使用されるような文脈内で別様に理解されない限り、概して、ある実施形態がある特徴、要素、またはステップを含む一方、他の実施形態がそれらを含まないことを伝えることが意図される。したがって、そのような条件文は、概して、特徴、要素、もしくはステップが、1つ以上の実施形態に対していかようにも要求されること、または1つ以上の実施形態が、著者の入力または促しの有無を問わず、これらの特徴、要素、もしくはステップが任意の特定の実施形態において含まれる、もしくは実施されるべきかどうかを決定するための論理を必然的に含むことを示唆することを意図されない。用語「～を備える」、「～を含む」、「～を有する」、および同等物は、同義語であり、非限定的方式で包括的に使用され、付加的要素、特徴、行為、動作等を除外しない。また、用語「または」は、その包括的意味において使用され（およびその排他的意味において使用されず）、したがって、例えば、要素のリストを接続するために使用されると、用語「または」は、リスト内の要素のうちの一つ、いくつか、または全てを意味する。加えて、本願および添付される請求項で使用されるような冠詞「a」、「an」、および「the」は、別様に規定されない限り、「1つ以上の」もしくは「少なくとも1つ」を意味するように解釈されるべきである。

10

**【0334】**

本明細書で使用されるように、項目のリスト「～のうちの少なくとも1つ」を指す語句は、単一の要素を含む、それらの項目の任意の組み合わせを指す。ある実施例として、「A、B、またはCのうちの少なくとも1つ」は、A、B、C、AおよびB、AおよびC、BおよびC、ならびにA、B、およびCを網羅することが意図される。語句「X、Y、およびZのうちの少なくとも1つ」等の接続文は、別様に具体的に記載されない限り、概して、項目、用語等がX、Y、またはZのうちの少なくとも1つであり得ることを伝えるために使用されるような文脈で別様に理解される。したがって、そのような接続文は、概して、ある実施形態が、Xのうちの少なくとも1つ、Yのうちの少なくとも1つ、およびZのうちの少なくとも1つがそれぞれ存在するように要求することを示唆することを意図されない。

20

**【0335】**

同様に、動作は、特定の順序で図面に描写され得るが、これは、望ましい結果を達成するために、そのような動作が示される特定の順序で、もしくは連続的順序で実施される、または全ての図示される動作が実施される必要はないと認識されるべきである。さらに、図面は、フローチャートの形態で1つ以上の例示的プロセスを図式的に描写し得る。しかしながら、描写されない他の動作も、図式的に図示される例示的方法およびプロセス内に組み込まれることができる。例えば、1つ以上の付加的動作が、図示される動作のいずれかの前に、その後に、それと同時に、またはその間に実施されることができる。加えて、動作は、他の実装において再配列される、または再順序付けられ得る。ある状況では、マルチタスクおよび並列処理が、有利であり得る。さらに、上記に説明される実装における種々のシステムコンポーネントの分離は、全ての実装におけるそのような分離を要求するものとして理解されるべきではなく、説明されるプログラムコンポーネントおよびシステムは、概して、単一のソフトウェア製品においてともに統合される、または複数のソフトウェア製品にパッケージ化され得ることを理解されたい。加えて、他の実装も、以下の請求項の範囲内である。いくつかの場合では、請求項に列挙されるアクションは、異なる順序で実施され、依然として、望ましい結果を達成することができる。

30

40

【図 1】

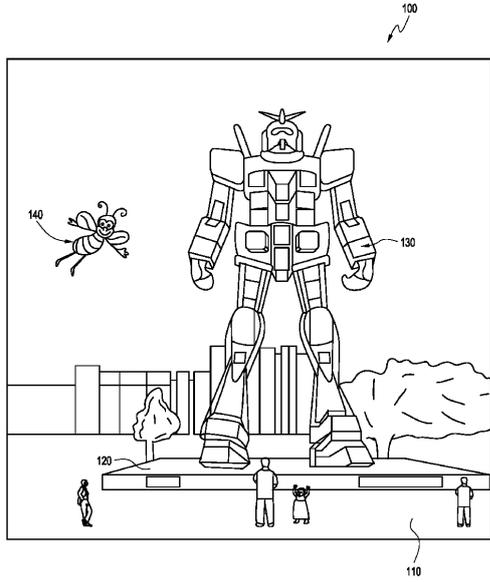


FIG. 1

【図 2 A】

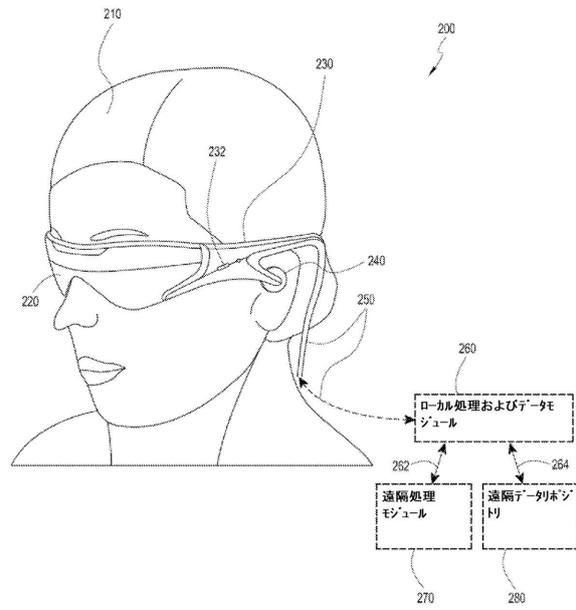


FIG. 2A

【図 2 B】

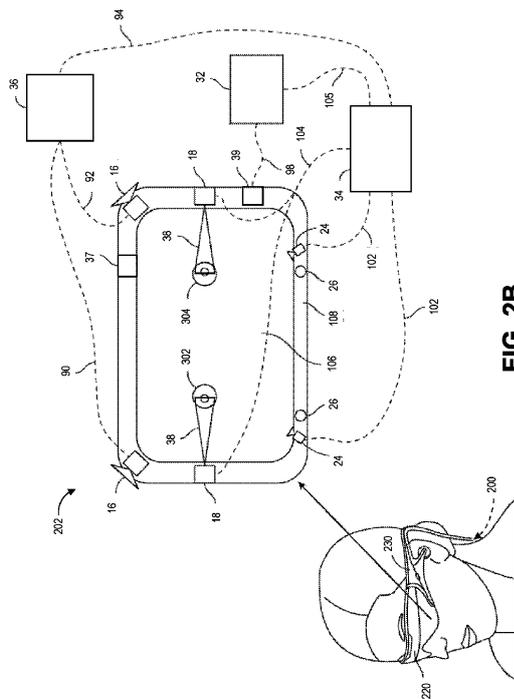


FIG. 2B

【図 3】

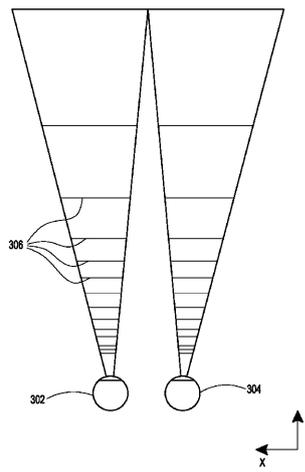


FIG. 3

【 図 4 】

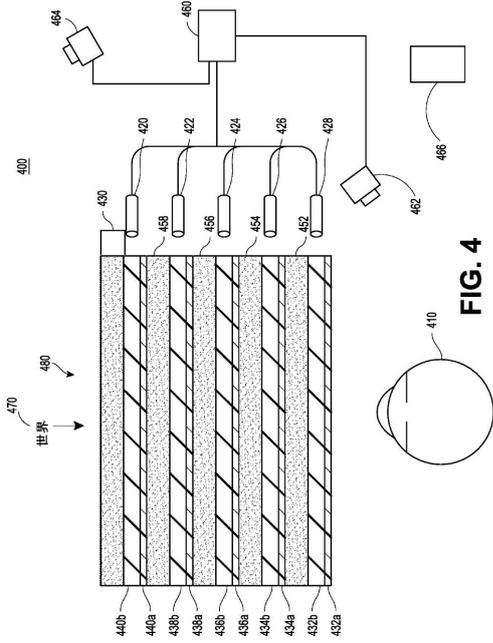


FIG. 4

【 図 5 】

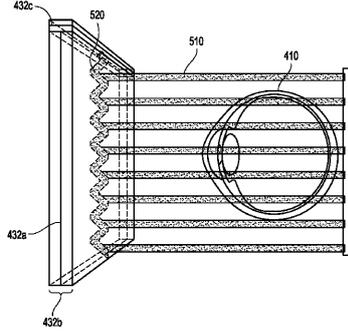


FIG. 5

【 図 6 】

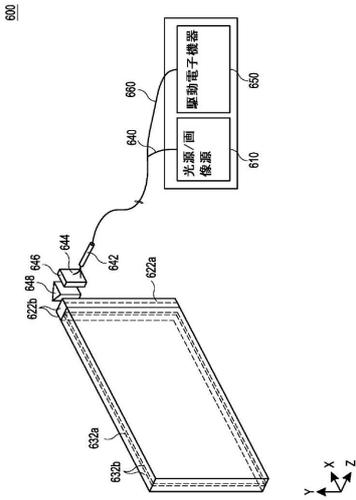


FIG. 6

【 図 7 】

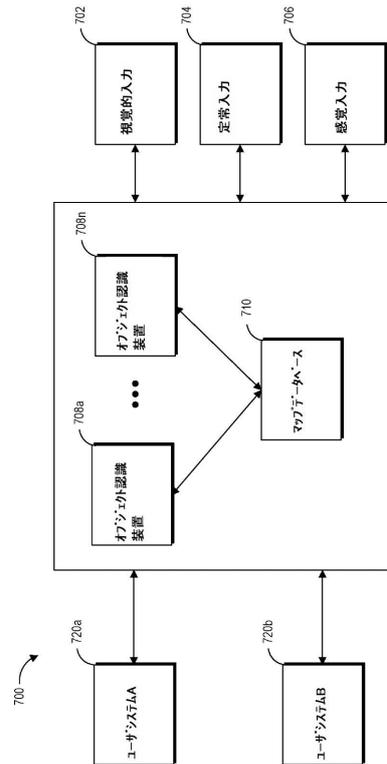


FIG. 7

【 図 8 】

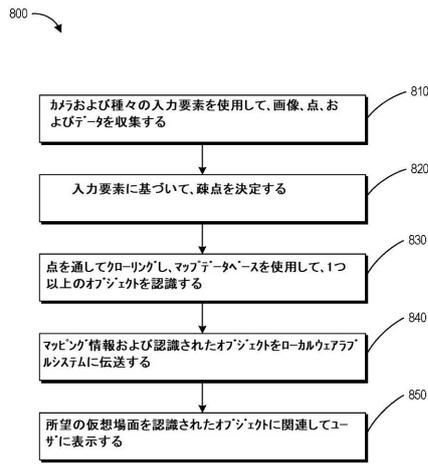


FIG. 8

【 図 9 】

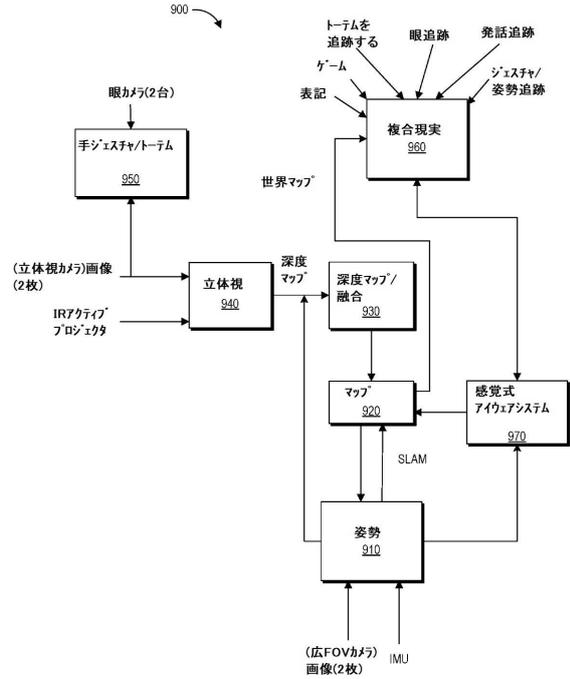


FIG. 9

【 図 10 】

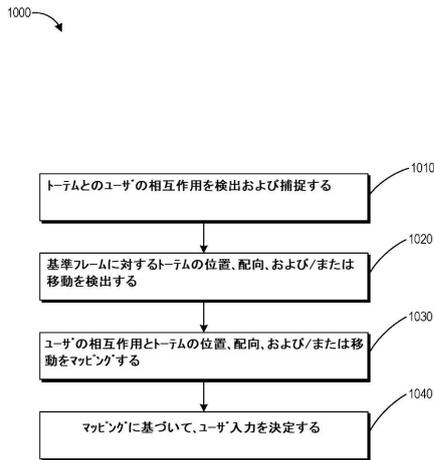


FIG. 10

【 図 11 】

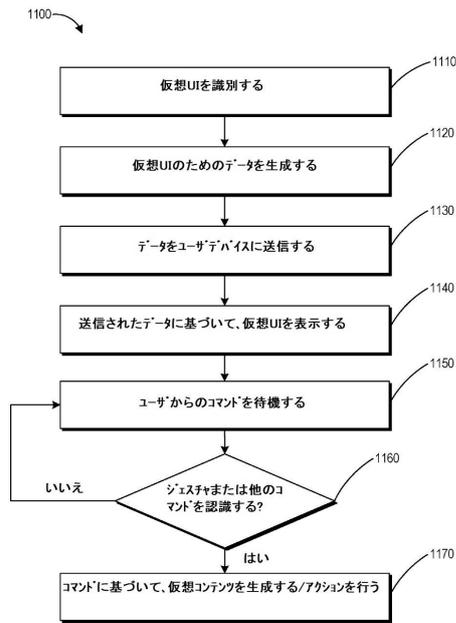
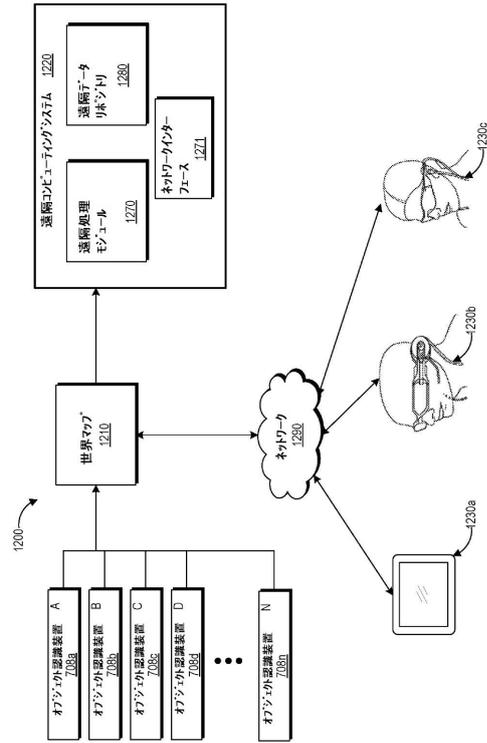


FIG. 11

【 図 1 2 】



【 図 1 3 A 】

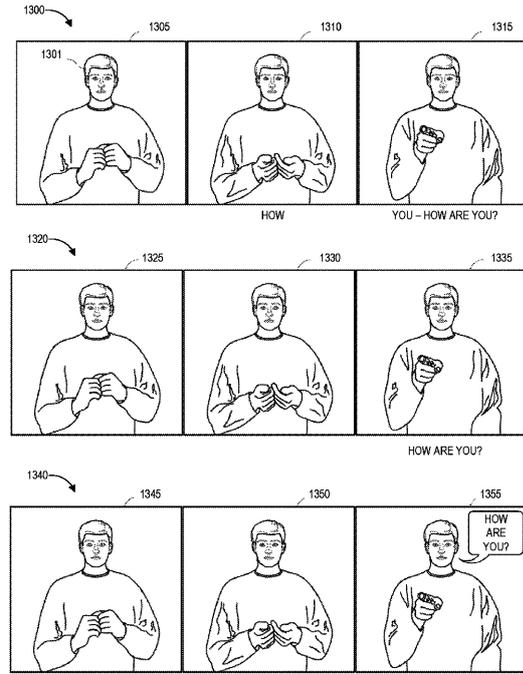


FIG. 12

FIG. 13A

【 図 1 3 B 】

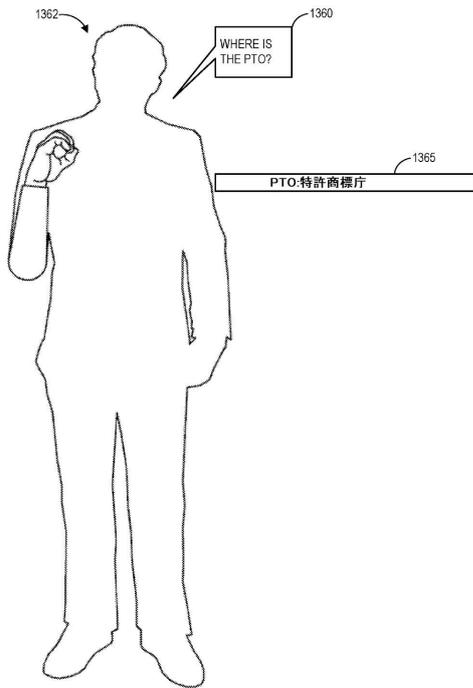


FIG. 13B

【 図 1 3 C 】

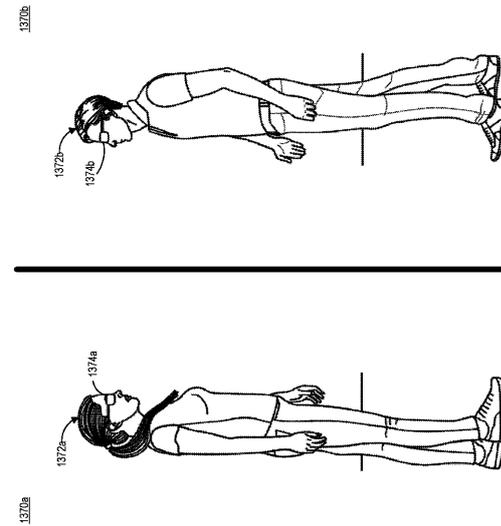


FIG. 13C

【図13D】



FIG. 13D

【図14A】

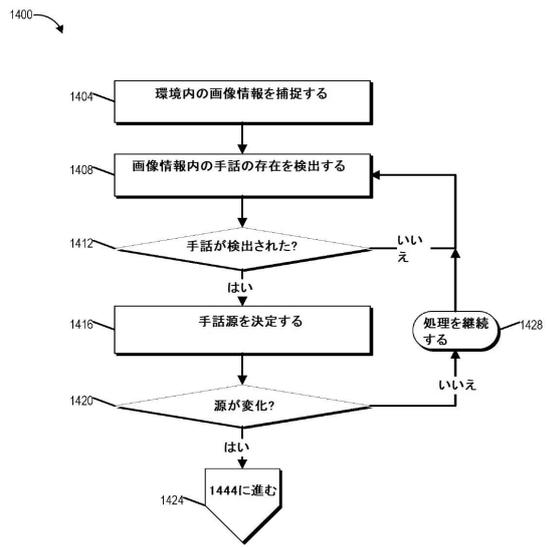


FIG. 14A

【図14B】

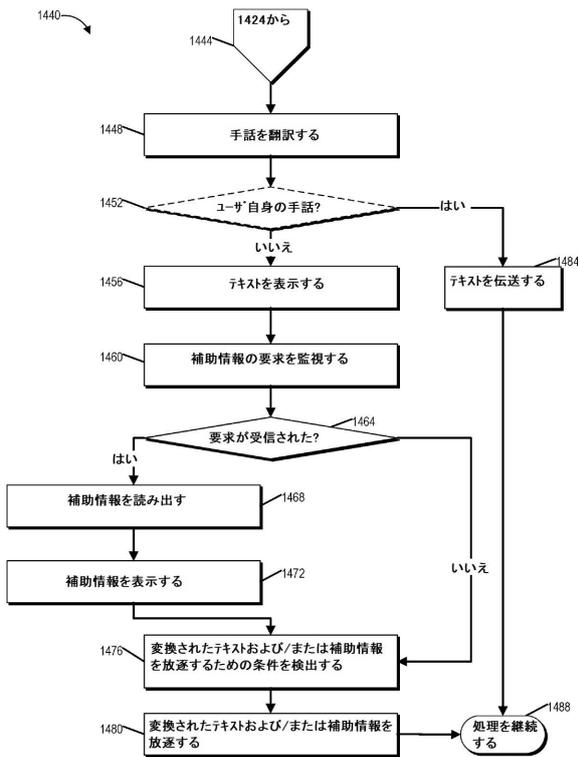


FIG. 14B

【図14C】

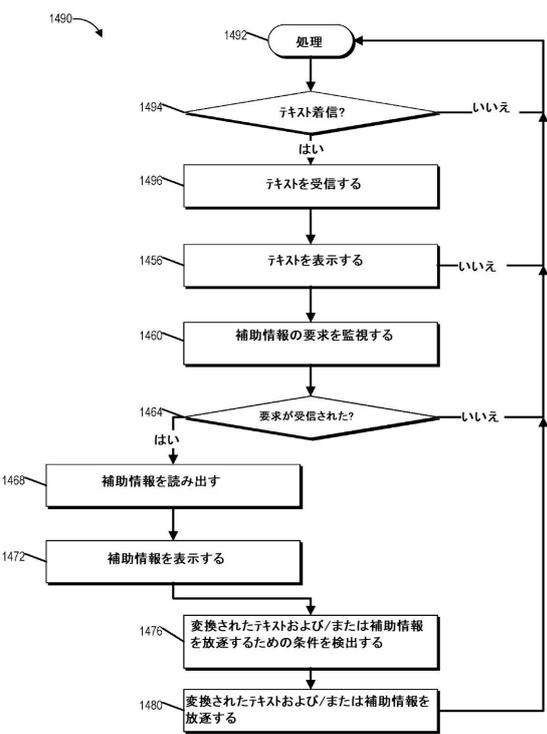


FIG. 14C

【 図 15 】

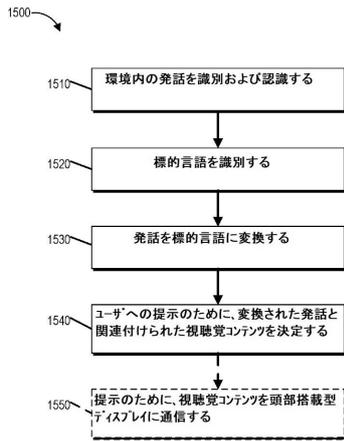


FIG. 15

【 図 16 A 】

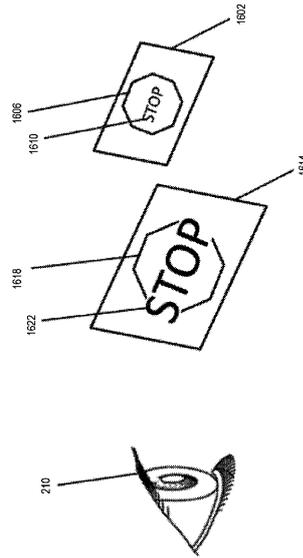


FIG. 16A

【 図 16 B 】

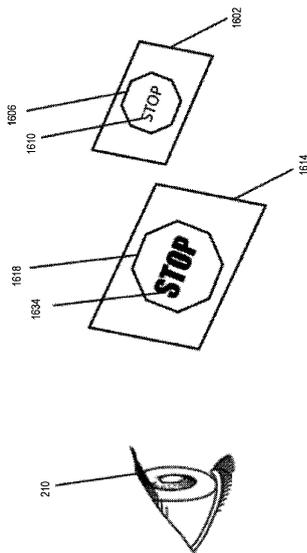


FIG. 16B

【 図 16 C 】

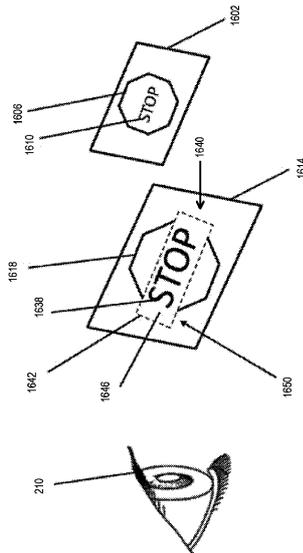


FIG. 16C

【 図 16 D 】

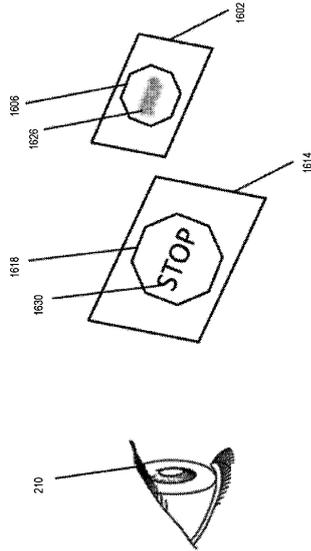


FIG. 16D

【 図 16 E 】

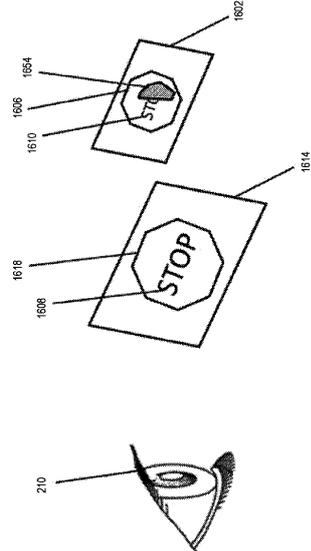


FIG. 16E

【 図 17 】

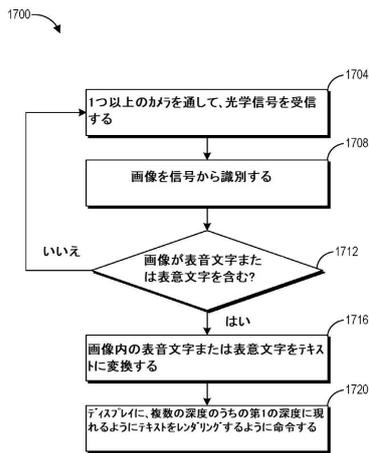


FIG. 17

【 図 18 】

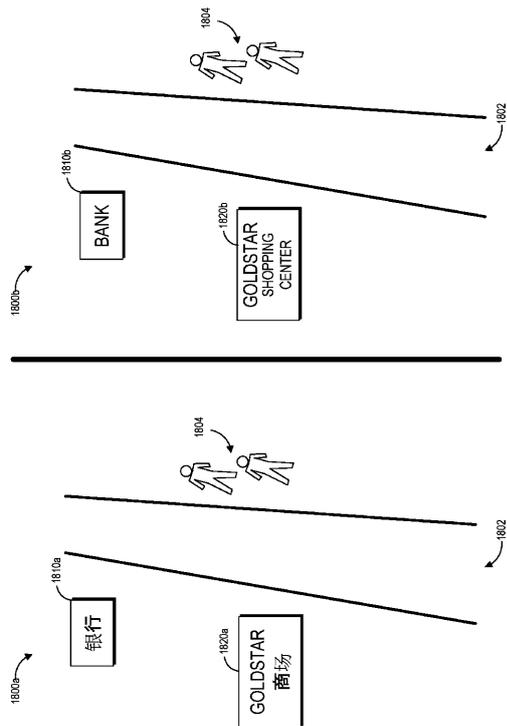


FIG. 18

【 図 19 】

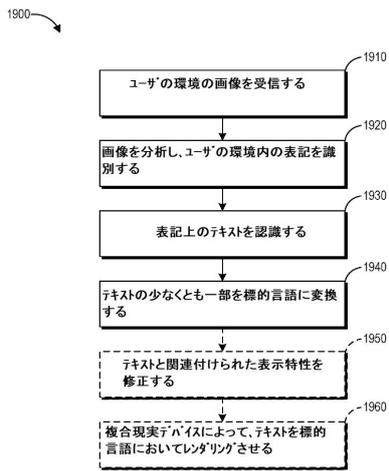


FIG. 19

## フロントページの続き

(51)Int.Cl.		F I		
G 0 6 F 40/40	(2020.01)	G 0 9 B	21/00	F
G 0 6 F 40/58	(2020.01)	G 0 6 F	40/40	
		G 0 6 F	40/58	

## 早期審査対象出願

- (74)代理人 100181674  
弁理士 飯田 貴敏
- (74)代理人 100181641  
弁理士 石川 大輔
- (74)代理人 230113332  
弁護士 山本 健策
- (72)発明者 プロウィ, エリック  
アメリカ合衆国 フロリダ 3 3 3 2 2, プランテーション, ダブリュー. サンライズ ブ  
ールバード 7 5 0 0
- (72)発明者 ウッズ, マイケル ヤヌシュ  
アメリカ合衆国 フロリダ 3 3 3 2 2, プランテーション, ダブリュー. サンライズ ブ  
ールバード 7 5 0 0
- (72)発明者 ラビノビッチ, アンドリュー  
アメリカ合衆国 フロリダ 3 3 3 2 2, プランテーション, ダブリュー. サンライズ ブ  
ールバード 7 5 0 0

審査官 円子 英紀

- (56)参考文献 特開2015-169814(JP,A)  
特開2012-252581(JP,A)  
特開2004-254218(JP,A)  
特開2005-197889(JP,A)  
特開2006-139138(JP,A)  
特開2015-076774(JP,A)

## (58)調査した分野(Int.Cl., DB名)

G 0 6 F 3 / 0 1  
G 0 6 F 4 0 / 4 0  
G 0 6 F 4 0 / 5 8  
G 0 6 T 7 / 0 0  
G 0 6 T 7 / 2 0  
G 0 6 T 1 9 / 0 0  
G 0 9 B 2 1 / 0 0