

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6814279号  
(P6814279)

(45) 発行日 令和3年1月13日(2021.1.13)

(24) 登録日 令和2年12月22日(2020.12.22)

(51) Int.Cl.

F I

G O 6 F 16/63 (2019.01)

G O 6 F 16/63

G O 6 F 16/638 (2019.01)

G O 6 F 16/638

G O 6 F 16/73 (2019.01)

G O 6 F 16/73

G O 6 F 16/738 (2019.01)

G O 6 F 16/738

G O 6 F 3/01 (2006.01)

G O 6 F 3/01 5 1 0

請求項の数 20 (全 63 頁) 最終頁に続く

(21) 出願番号 特願2019-506525 (P2019-506525)  
 (86) (22) 出願日 平成29年8月10日 (2017.8.10)  
 (65) 公表番号 特表2019-531538 (P2019-531538A)  
 (43) 公表日 令和1年10月31日 (2019.10.31)  
 (86) 国際出願番号 PCT/US2017/046251  
 (87) 国際公開番号 W02018/031745  
 (87) 国際公開日 平成30年2月15日 (2018.2.15)  
 審査請求日 令和2年8月7日 (2020.8.7)  
 (31) 優先権主張番号 62/374,183  
 (32) 優先日 平成28年8月12日 (2016.8.12)  
 (33) 優先権主張国・地域又は機関  
 米国 (US)

早期審査対象出願

(73) 特許権者 514108838  
 マジック リープ, インコーポレイテッ  
 ド  
 Magic Leap, Inc.  
 アメリカ合衆国 フロリダ 33322,  
 プランテーション, ウェスト サンライズ  
 ブールバード 7500  
 7500 W SUNRISE BLVD  
 , PLANTATION, FL 3332  
 2 USA  
 (74) 代理人 100078282  
 弁理士 山本 秀策  
 (74) 代理人 100113413  
 弁理士 森下 夏樹

最終頁に続く

(54) 【発明の名称】 ワードフロー注釈

(57) 【特許請求の範囲】

【請求項 1】

拡張現実 (AR) システムであって、前記 AR システムは、  
 仮想コンテンツを AR システムのユーザに提示するように構成された AR ディスプレイ  
 と、  
 周囲音を捕捉するように構成されたオーディオセンサと、  
 前記 AR ディスプレイおよび前記オーディオセンサと通信するハードウェアプロセッサ  
 と  
 を備え、  
 前記ハードウェアプロセッサは、  
 前記オーディオセンサによって捕捉された前記周囲音を受信することと、  
 前記周囲音内の発話の存在を検出することと、  
 前記検出された発話をテキストに変換することと、  
 前記テキストを解析し、1つ以上の単語を認識することと、  
単語頻度リスト上の前記1つ以上の単語の位置を決定することであって、前記単語頻度  
リストは、前記オーディオセンサによるそれらの検出の頻度に少なくとも部分的に基づい  
てランク付けされた単語のリストを含む、ことと、  
ユーザ入力に基づいて、前記単語頻度リスト上の閾値位置を決定することと、  
前記1つ以上の単語の位置が前記単語頻度リスト上の前記閾値位置を超えるかどうか  
に基づいて、前記1つ以上の単語を稀な単語として分類することと、

前記 1 つ以上の単語に関連付けられた減少係数を決定することであって、前記減少係数は、前記稀な単語に関連付けられた補助情報の表示を抑制する期間を含む、ことと、

前記減少係数に基づいて、前記稀な単語に関連付けられた前記補助情報を読み出すことと、

前記 A R ディスプレイに前記読み出された補助情報を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ている、A R システム。

【請求項 2】

前記稀な単語は、一般的単語辞書から除外された単語である、請求項 1 に記載の A R システム。

【請求項 3】

前記ハードウェアプロセッサは、  
前記補助情報の表示を閉じるための条件を検出することと、  
前記検出された条件に応答して、前記 A R ディスプレイに前記補助情報の表示を閉じさせることと  
を行うようにさらにプログラムされている、請求項 1 に記載の A R システム。

【請求項 4】

前記表示を閉じるための前記条件を検出するために、前記ハードウェアプロセッサは、閾値期間が経過したことを決定すること、別の稀な単語を検出すること、ユーザの姿勢を検出すること、または、ユーザ入力デバイスからの入力を受信することのうちの少なくとも 1 つを実施するようにプログラムされている、請求項 3 に記載の A R システム。

【請求項 5】

前記ハードウェアプロセッサは、  
前記検出された発話の源を決定することと、  
前記稀な単語を検出し、前記検出された発話の源が前記 A R ディスプレイのユーザ以外の話者に関連付けられていることを決定すると、前記補助情報を読み出し、表示することと  
を行うようにさらにプログラムされている、請求項 1 に記載の A R システム。

【請求項 6】

前記検出された発話の源は、1 つ以上の指向性マイクロホンを備えている前記オーディオセンサから収集されたオーディオ情報、または前記 A R システムの外向きに面した結像システムによって入手された画像のうちの少なくとも 1 つに基づいて決定される、請求項 5 に記載の A R システム。

【請求項 7】

前記検出された発話の源を決定するために、前記ハードウェアプロセッサは、音声認識を前記検出された発話に対して実施し、前記話者の識別を決定するようにプログラムされている、請求項 5 に記載の A R システム。

【請求項 8】

前記検出された発話の源は、前記ユーザの環境内の別のコンピュータ、前記話者に関連付けられた別の A R デバイス、または前記ユーザの環境内の人物のうちの少なくとも 1 つを含む、請求項 5 に記載の A R システム。

【請求項 9】

前記ユーザが前記話者とテレプレゼンスセッション中であり、前記検出された発話の源が前記話者に関連付けられた他のユーザデバイスである場合、前記ハードウェアプロセッサは、前記話者の仮想アバタをレンダリングすることと、前記補助情報を前記他のユーザの前記仮想アバタの近傍にレンダリングすることとを前記 A R ディスプレイに行わせるようにさらにプログラムされている、請求項 8 に記載の A R システム。

【請求項 10】

前記 A R ディスプレイに前記読み出された補助情報をレンダリングさせるために、前記ハードウェアプロセッサは、前記源の場所を決定することと、前記 A R ディスプレイを通して、前記源のビューを曖昧にすることなく、前記補助情報を前記源に近接した 3 D 空間

10

20

30

40

50

内の位置にレンダリングすることとを行うようにプログラムされている、請求項 5 に記載の A R システム。

【請求項 1 1】

補助情報は、前記稀な単語の説明を含む、請求項 1 に記載の A R システム。

【請求項 1 2】

前記ハードウェアプロセッサは、  
ユーザから、前記補助情報に注釈を付けるための指示を受信することと、  
前記補助情報の注釈の記憶を開始することと、  
前記稀な単語の別の存在を検出することに応答して、前記 A R ディスプレイに前記注釈および前記補助情報をレンダリングさせることと  
を行うようにさらにプログラムされている、請求項 1 に記載の A R システム。

10

【請求項 1 3】

方法であって、前記方法は、  
仮想コンテンツをユーザの環境内に提示するように構成されたハードウェアプロセッサおよび A R ディスプレイを備えている拡張現実 ( A R ) デバイスの制御下で、  
前記 A R デバイスの前記ユーザの環境を監視することと、  
前記ユーザまたは前記環境のうちの少なくとも 1 つに関連付けられたコンテキスト情報に基づいて、前記環境内の着目オブジェクトの存在を検出することと、  
単語頻度リスト上の前記着目オブジェクトの位置を決定することであって、前記単語頻度リストは、前記 A R デバイスによるそれらの検出の頻度に少なくとも部分的に基づいて  
ランク付けされた単語のリストを含む、ことと、

20

ユーザ入力に基づいて、前記単語頻度リスト上の閾値位置を決定することと、  
1 つ以上の単語の位置が前記単語頻度リスト上の前記閾値位置を超えるかどうかに基づいて、前記着目オブジェクトを稀な単語として分類することと、

前記着目オブジェクトに関連付けられた減少係数を決定することであって、前記減少係数は、前記着目オブジェクトに関連付けられた補助情報の表示を抑制する期間を含む、ことと、

前記減少係数に基づいて、前記着目オブジェクトのための補助情報を読み出すことと、  
前記ユーザの視野内の前記補助情報の表示位置を決定することであって、前記視野は、ユーザが所与の時間に知覚可能な前記環境の一部を含む、ことと、  
前記 A R ディスプレイに前記補助情報を前記表示位置にレンダリングさせることとを含む、方法。

30

【請求項 1 4】

前記着目オブジェクトの存在を検出することは、  
前記環境内の周囲音から発話を検出することと、  
前記発話を解析し、稀な単語を識別することと  
を含む、請求項 1 3 に記載の方法。

【請求項 1 5】

前記環境を監視することは、  
前記環境の周囲音を捕捉し、分析すること、または  
前記環境の画像を入手すること  
のうちの少なくとも 1 つを含む、請求項 1 3 に記載の方法。

40

【請求項 1 6】

前記コンテキスト情報は、前記ユーザの過去の体験に関連付けられ、前記単語頻度リストは、前記ユーザの過去の体験において前記ユーザが遭遇した頻度によって順序付けられた単語のリストを含む、請求項 1 3 に記載の方法。

【請求項 1 7】

前記補助情報は、前記着目オブジェクトの説明テキストを含む、請求項 1 3 に記載の方法。

【請求項 1 8】

50

前記ユーザから、前記補助情報に注釈を付けるための指示を受信することと、  
前記補助情報の注釈の記憶を開始することと、  
前記着目オブジェクトの別の存在を検出することに応答して、前記ARディスプレイに  
前記注釈および前記補助情報をレンダリングさせることと  
をさらに含む、請求項13に記載の方法。

【請求項19】

前記補助情報の表示位置を決定することは、  
前記着目オブジェクトの場所を決定することと、  
前記ARディスプレイを通して、前記着目オブジェクトのビューを曖昧にすることなく  
、前記着目オブジェクトに近接した前記環境内の位置として前記表示位置を識別すること  
と  
を含む、請求項13に記載の方法。

【請求項20】

前記単語頻度リストは、前記ユーザの過去の体験において前記ユーザが遭遇した頻度によ  
って順序付けられた単語のリストを含む、請求項1に記載のARシステム。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願の引用)

本願は、米国仮出願第62/374,183号(2016年8月12日出願、名称「WORD FLOW ANNOTATION」)の米国特許法§119(e)に基づく優先権の利益を主張し、上記出願の開示は、その全体が参照により本明細書に引用される。

【0002】

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、発話の視覚的注釈を生成し、システムのユーザによる発話の理解を向上させることに  
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、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おり、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様式、またはそのように知覚  
され得る様式でユーザに提示される。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明性を伴わずに、デジタルまたは仮想画像情報  
の提示を伴う。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間における仮  
想画像要素の快適で、自然のような感覚で、かつ豊かな提示を促進するVR、AR、また  
はMR技術の生成は、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】

【0004】

聴取者に対して最小限のレベルの注意を逸らすものおよび聴取者による取るに足りない  
レベルの努力を伴って、着目オブジェクト(例えば、聴取者に新しいまたは未知の単語)  
に関する説明または情報を(近)リアルタイムで提供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。

【0005】

10

20

30

40

50

いくつかの実施形態では、拡張現実（ＡＲ）システムは、仮想コンテンツをＡＲシステムのユーザに提示するように構成されたＡＲディスプレイと、周囲音を捕捉するように構成されたオーディオセンサと、ＡＲディスプレイおよびオーディオセンサと通信するハードウェアプロセッサとを備えていることができる。ハードウェアプロセッサは、オーディオセンサによって捕捉された周囲音を受信し、周囲音内の発話の存在を検出し、検出された発話をテキストに変換し、発話中の稀な単語を検出し、稀な単語に関連付けられた補助情報を読み出し、ＡＲディスプレイに読み出された補助情報をレンダリングさせるようにプログラムされることができる。

【０００６】

いくつかの実施形態では、方法が、仮想コンテンツをユーザの環境内に提示するように構成されたハードウェアプロセッサおよびＡＲディスプレイを備えている拡張現実（ＡＲ）デバイスの制御下で実施されることができる。方法は、ＡＲデバイスのユーザの環境を監視することと、ユーザまたは環境のうちの少なくとも１つに関連付けられたコンテキスト情報に基づいて、環境内の着目オブジェクトの存在を検出することと、着目オブジェクトのための補助情報を読み出すことと、ユーザの視野内の補助情報の表示位置を決定することとであって、視野は、ユーザが所与の時間に知覚可能な環境の一部を含む、ことと、ＡＲディスプレイに補助情報を表示位置にレンダリングさせることとを含む。

【０００７】

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、および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）

拡張現実（ＡＲ）システムであって、前記ＡＲシステムは、  
仮想コンテンツをＡＲシステムのユーザに提示するように構成されたＡＲディスプレイと、  
周囲音を捕捉するように構成されたオーディオセンサと、  
前記ＡＲディスプレイおよび前記オーディオセンサと通信するハードウェアプロセッサと  
を備え、  
前記ハードウェアプロセッサは、  
前記オーディオセンサによって捕捉された前記周囲音を受信することと、  
前記周囲音内の発話の存在を検出することと、  
前記検出された発話をテキストに変換することと、  
前記発話中の稀な単語を検出することと、  
前記稀な単語に関連付けられた補助情報を読み出すことと、  
前記ＡＲディスプレイに前記読み出された補助情報を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ている、ＡＲシステム。

（項目２）

前記稀な単語は、一般的単語辞書から除外された単語である、項目１に記載のＡＲシステム。

（項目３）

前記ハードウェアプロセッサは、  
前記補助情報の表示を閉じるための条件を検出することと、  
前記検出された条件に応答して、前記ＡＲディスプレイに前記補助情報の表示を閉じさせることと  
を行うようにさらにプログラムされている、項目１に記載のＡＲシステム。

（項目４）

前記表示を閉じるための前記条件を検出するために、前記ハードウェアプロセッサは、

閾値期間が経過したことを決定すること、別の稀な単語を検出すること、ユーザの姿勢を検出すること、または、ユーザ入力デバイスからの入力を受信することのうちの少なくとも1つを実施するようにプログラムされている、項目3に記載のARシステム。

(項目5)

前記ハードウェアプロセッサは、

前記検出された発話の源を決定することと、

前記稀な単語を検出し、前記検出された発話の源が前記ARディスプレイのユーザ以外の話者に関連付けられていることを決定すると、前記補助情報を読み出し、表示すること

を行うようにさらにプログラムされている、項目1に記載のARシステム。

10

(項目6)

前記検出された発話の源は、1つ以上の指向性マイクロホンを備えているオーディオセンサから収集されたオーディオ情報、または前記ARシステムの外向きに面した結像システムによって入手された画像のうちの少なくとも1つに基づいて決定される、項目5に記載のARシステム。

(項目7)

前記検出された発話の源を決定するために、前記ハードウェアプロセッサは、音声認識を前記検出された発話に対して実施し、前記話者の識別を決定するようにプログラムされている、項目5に記載のARシステム。

(項目8)

前記検出された発話の源は、前記ユーザの環境内の別のコンピュータ、前記話者に関連付けられた別のARデバイス、または前記ユーザの環境内の人物のうちの少なくとも1つを含む、項目5に記載のARシステム。

20

(項目9)

前記ユーザが前記話者とテレプレゼンスセッション中であり、前記検出された発話の源が前記話者に関連付けられた他のユーザデバイスである場合、前記ハードウェアプロセッサは、前記話者の仮想アバタをレンダリングすることと、前記補助情報を前記他のユーザの前記仮想アバタの近傍にレンダリングすることとを前記ARディスプレイに行わせるようにさらにプログラムされている、項目8に記載のARシステム。

(項目10)

前記ARディスプレイに前記読み出された補助情報をレンダリングさせるために、前記ハードウェアプロセッサは、前記源の場所を決定することと、前記ARディスプレイを通して、前記源のビューを曖昧にすることなく、前記補助情報を前記源に近接した3D空間内の位置にレンダリングすることとを行うようにプログラムされている、項目5に記載のARシステム。

30

(項目11)

補助情報は、前記稀な単語の説明を含む、項目1に記載のARシステム。

(項目12)

前記ハードウェアプロセッサは、

ユーザから、前記補助情報に注釈を付けるための指示を受信することと、

前記補助情報の注釈の記憶を開始することと、

前記稀な単語の別の存在を検出することに応答して、前記ARディスプレイに前記注釈および前記補助情報をレンダリングさせることと

を行うようにさらにプログラムされている、項目1に記載のARシステム。

40

(項目13)

方法であって、前記方法は、

仮想コンテンツをユーザの環境内に提示するように構成されたハードウェアプロセッサおよびARディスプレイを備えている拡張現実(AR)デバイスの制御下で、

前記ARデバイスの前記ユーザの環境を監視することと、

前記ユーザまたは前記環境のうちの少なくとも1つに関連付けられたコンテキスト情報

50

に基づいて、前記環境内の着目オブジェクトの存在を検出することと、  
前記着目オブジェクトのための補助情報を読み出すことと、  
前記ユーザの視野内の前記補助情報の表示位置を決定することであって、前記視野は、  
ユーザが所与の時間に知覚可能な前記環境の一部を含む、ことと、  
前記ARディスプレイに前記補助情報を前記表示位置にレンダリングさせることと  
を含む、方法。

(項目14)

前記着目オブジェクトは、キーワード、物理的オブジェクト、または人物を含む、項目  
13に記載の方法。

(項目15)

前記着目オブジェクトは、キーワードを含み、前記着目オブジェクトの存在を検出する  
ことは、

前記環境内の周囲音から発話を検出することと、  
前記発話を解析し、稀な単語を識別することと  
を含む、項目14に記載の方法。

(項目16)

前記環境を監視することは、  
前記環境の周囲音を捕捉し、分析すること、または  
前記環境の画像を入手すること  
のうちの少なくとも1つを含む、項目13に記載の方法。

(項目17)

前記コンテキスト情報は、前記ユーザの過去の体験に関連付けられ、前記着目オブジェ  
クトは、前記ユーザが稀にしか遭遇しないオブジェクトを含む、項目13に記載の方法。

(項目18)

前記補助情報は、前記着目オブジェクトに関連付けられた製品情報、または前記着目オ  
ブジェクトの説明テキストのうちの少なくとも1つを含む、項目13に記載の方法。

(項目19)

前記ユーザから、前記補助情報に注釈を付けるための指示を受信することと、  
前記補助情報の注釈の記憶を開始することと、  
前記着目オブジェクトの別の存在を検出することに応答して、前記ARディスプレイに  
前記注釈および前記補助情報をレンダリングさせることと  
をさらに含む、項目13に記載の方法。

(項目20)

前記補助情報の表示位置を決定することは、  
前記着目オブジェクトの場所を決定することと、  
前記ARディスプレイを通して、前記着目オブジェクトのビューを曖昧にすることなく  
、前記着目オブジェクトに近接した前記環境内の位置として前記表示位置を識別すること  
と  
を含む、項目13に記載の方法。

【図面の簡単な説明】

【0008】

【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2】図2は、ワードフロー注釈システムの実施形態を実装し得るウェアラブルシステ  
ムの例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例を図式的に図示  
する。

【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。

10

20

30

40

50

【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、または明視野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例のプロセスフロー図である。

【図 9】図 9 は、ワードフロー注射システムの実施形態含むウェアラブルシステムの別の例のブロック図である。

【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方法の例のプロセスフロー図である。

10

【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方法の例のプロセスフロー図である。

【図 12】図 12 は、互い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。

【図 13 A】図 13 A - 13 G は、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例を図示する。

【図 13 B】図 13 A - 13 G は、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例を図示する。

【図 13 C】図 13 A - 13 G は、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例を図示する。

20

【図 13 D】図 13 A - 13 G は、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例を図示する。

【図 13 E】図 13 A - 13 G は、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例を図示する。

【図 13 F】図 13 A - 13 G は、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例を図示する。

【図 13 G】図 13 A - 13 G は、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例を図示する。

【図 14 A】図 14 A および 14 B は、発話に注射を付ける例示的过程を示す。

【図 14 B】図 14 A および 14 B は、発話に注射を付ける例示的过程を示す。

30

【図 15】図 15 は、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例示的过程を図示する。

【図 16 A】図 16 A および 16 B は、キーワードに基づくスレッド識別の例を図示する。

【図 16 B】図 16 A および 16 B は、キーワードに基づくスレッド識別の例を図示する。

【図 17】図 17 は、スレッド識別の例示的过程を図示する。

【発明を実施するための形態】

【0009】

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的实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。

40

【0010】

(概要)

現代の世界における日々の発話では、それらが一般的ではない考えまたは概念、特定の問題分野からの考えまたは概念に言及するので、または聴取者に未知の人物または物事に対する適切な言及であるので、聴取者に新しい名前または言及に遭遇することが一般的である。聴取者は、単に、熟知していないこと、または知識がないことに起因して、ある単語を把握していない場合がある。必然的に、それは、会話を妨害し、当惑、混乱、または誤解につながり得る。

50



## 【 0 0 1 1 】

聴取者への最小限のレベルの注意を逸らすものおよび聴取者による取るに足りないレベルの努力を伴って、新しいまたは未知の単語のための説明もしくは情報を（近）リアルタイムで提供することができる拡張現実システムを有することが望ましくあり得る。

## 【 0 0 1 2 】

本開示は、拡張現実デバイス（ARD）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例を開示する。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を可能にすることができる。そのようなシステムは、文字、画像効果、テキスト、グラフィック、または任意の種類のビデオ等、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。ARDは、ARディスプレイを含むことができ、それは、片眼もしくは両眼またはVRディスプレイ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。ARDは、オーディオセンサも含むことができ、それは、ユーザまたはユーザの環境内の他の人物による発話を含むオーディオ情報を捕捉することができる。

## 【 0 0 1 3 】

ARDは、ワードフロー注釈システムを実装することができる。ワードフロー注釈システムは、ARDの発話認識、稀な単語の検出、および表示能力を組み合わせ、ユーザに、単語またはユーザに適切な補助情報を提供することができる。例として、第1のユーザのためのARDは、第1のユーザが第2のユーザと通信するテレプレゼンスセッション中、ネットワークを介して、オーディオ情報を受信することができる。ワードフロー注釈システムは、第2のユーザからのオーディオ情報内の発話を検出することができ、検出された発話をテキストに変換し、変換されたテキスト内の稀な単語を検索し、検出された稀な単語に関連付けられた補助情報を読み出すことができる。ARDは、補助情報をARDのARディスプレイ上に表示することができる。システムは、閾値期間が経過した後、またはシステムが別の稀な単語を検出すると、補助情報の表示を閉じることができ、それは、有利には、依然として、稀な単語についての有用な補助情報を提供しながら、第1のユーザへの注意を逸らすもののレベルを低減させ得る。稀な辞書単語に対して、単語の意味（例えば、定義）が、表示されることができる。稀な適切な名詞（例えば、地理的領域、人物の名前等）に対して、適切な源（例えば、百科事典）からの関連情報が、表示されることができる。システムが提供し得る注釈の種類は、例えば、インターネット上の膨大な利用可能な情報リソースと同様に無限であり得る。いくつかの実装では、システムは、変換されたテキスト（稀な単語についての情報に加えて、またはその代替として）を、ディスプレイの装着者に、例えば、キャプションまたは発話の吹き出し（時として、コールアウトボックスとも称される）の形態で、テキストの話者の近傍に表示する。いくつかの実装では、補助情報の表示は、視覚的（例えば、前述のキャプションまたは発話の吹き出し）である。しかしながら、他の実装では、視覚的ディスプレイに加えて、またはその代替として、補助情報は、ユーザに聴覚的に提示されることができる（視覚障害ユーザのために有利であり得る）。

## 【 0 0 1 4 】

稀な単語に加えて、またはその代替として、類似技法が、他のキーワードにも、これらのキーワードのいくつかが稀ではない場合でも、適用されることができる。ARDは、ユーザの環境内の着目オブジェクト（例えば、ユーザが典型的には遭遇しない、オブジェクト）を検出し、着目オブジェクトに関連する補助情報を表示することができる。本明細書に説明されるキーワードおよび稀な単語は、単一単語、語句、または文を含むことができる。

## 【 0 0 1 5 】

さらに、会話は、多くの場合、複数の話題を伴う。ARDは、会話中のキーワードを識別し、キーワードに基づいて、会話中の話題を識別することができる。ARDは、会話に

10

20

30

40

50

関連付けられたテキストを異なるコンテキストスレッドに分離することができ、各スレッドは、1つ以上の話題に対応する。有利には、それは、A R Dが、異なるコンテキストスレッドを異なるように提示するユーザインターフェース（U I）を提供することができる（例えば、A Rディスプレイの両側で）ので、会話に関わる考えおよび要点のユーザの理解を促進することができる。

#### 【0016】

（ウェアラブルシステムの3Dディスプレイの例）

ウェアラブルシステム（本明細書では、拡張現実（A R）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、組み合わせにおける静止画像、ビデオのフレーム、またはビデオ等であり得る。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、または組み合わせて、V R、A R、またはM R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（H M D）であることができ、それは、A Rデバイス（A R D）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「A R」は、用語「M R」と同義的に使用される。

#### 【0017】

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、M R場面100が、描写され、M R技術のユーザは、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を見る。これらのアイテムに加え、M R技術のユーザは、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140も「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。

#### 【0018】

3Dディスプレイが、真の深度感覚を生成するために、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各点のために、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、それは、不安定な結像、有害な眼精疲労、頭痛をもたらし、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。

#### 【0019】

V R、A R、およびM R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、各深度平面に対して異なり得（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）、画像は、視認者の眼によって別個に集中させられ、それによって、異なる深度平面上に位置する場面のための異なる画像特徴を焦点に集めるために要求される眼の遠近調節に基づいて、または、焦点はずれであ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。

#### 【0020】

図2は、ウェアラブルシステム200の例を図示し、それは、A R / V R / M R場면을提供するように構成されることができ、本明細書に説明されるワードフロー注釈システムの実施形態を含むことができる。ウェアラブルシステム200は、A Rシステム200とも称され得る。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、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に結合され得る。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、A R / V R / M R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ていることができる。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のオーディオセンサ（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。

【0021】

10

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 200 は、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）も含むことができる。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し得る。内向きに面した結像システム 462 は、フレーム 230 に取り付けられ得、結像システム 462 は、処理モジュール 260 または 270 と電気通信し得、処理モジュールは、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 の眼の瞳孔の直径もしくは向き、眼の移動、または眼姿勢を決定し得る。

【0022】

20

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、またはビデオであり得る。

【0023】

ディスプレイ 220 は、有線導線または無線接続等によって、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）、ローカルデータ処理モジュール 260 は、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こと、ヘルメットもしくは帽子に固定して取り付けられること、ヘッドホンに内蔵されること、または別様にユーザ 210 に除去可能に取り付けられること（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。

30

【0024】

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および不揮発性メモリ（例えば、フラッシュメモリ）等のデジタルメモリを備え得、その両方は、データの処理、キャッシング、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스코ープ等の（例えば、フレーム 230 に動作可能に結合されるか、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み出し後にディスプレイ 220 に渡すために、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）を含み得る。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得る。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、互いに動作可能に結合され得る。

40

【0025】

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得る。いくつかの実施形態

50

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、それは、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り得る。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。

#### 【0026】

ヒト視覚系は、複雑であり、深度の現実的知覚を提供することは、困難である。理論によって限定されるわけではないが、オブジェクトの視認者は、両眼離反運動と遠近調節との組み合わせにより、オブジェクトを 3 次元として知覚し得ると考えられる。互いに対する 2 つの眼の両眼離反運動移動（すなわち、眼の視線を集中させ、オブジェクトに固定する 10 ための互いに向かい、またはそこから離れる瞳孔の回転）は、眼の水晶体の焦点合わせ（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 - 両眼離反運動反射」として知られる関係下、同一距離にたいする両眼離反運動の調和した変化を自動的に生じさせるであろう。同様に、両眼離反運動の変化は、通常条件下、遠近調節の調和した変化を誘起するであろう。遠近調節と両眼離反運動との間のより良好な調和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。

#### 【0027】

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示する。図 3 を参照すると、眼 302 および 304 からの z - 軸上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焦点が合っているように、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を焦点に集める。その結果、特定の遠近調節された状態は、深度平面 306 のうちの特定の 1 つに関連付けられていると言え、それは、関連付けられた焦点距離を有し、それによって、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は、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、焦点が合っている。いくつかの実施形態では、3 次元画像は、眼 302 および 304 の各々に画像の異なる提示を提供することによって、および、深度平面の各々に対応する画像の異なる提示を提供することによっても、シミュレートされ得る。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で焦点が合ってい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各々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。

#### 【0028】

（導波管スタックアセンブリ）

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、440b 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 2 のウェアラブルシステム 200 に対応し得、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 2 のディスプレイ 220 の中に統合され得る。

#### 【0029】

図4を継続して参照すると、導波管アセンブリ480は、複数の特徴458、456、454、452を導波管間に含み得る。いくつかの実施形態では、特徴458、456、454、452は、レンズであり得る。他の実施形態では、特徴458、456、454、452は、レンズではないこともある。むしろ、それらは、単に、スペーサであり得る(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)。

#### 【0030】

導波管432b、434b、436b、438b、440b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得る。各導波管レベルは、特定の深度平面に関連付けられ得、それは、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得る。画像投入デバイス420、422、424、426、428は、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得、それらの各々は、眼410に向かって出力するために、各それぞれの導波管に入射光を分配するように構成され得る。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、各導波管は、特定の導波管に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に向かわせら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得る。

#### 【0031】

いくつかの実施形態では、画像投入デバイス420、422、424、426、428の各々は、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個別的な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の各々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。

#### 【0032】

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読み取り可能な媒体内の命令)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り得る。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)の一部であり得る。

#### 【0033】

導波管440b、438b、436b、434b、432bは、全内部反射(TIR)によって各それぞれの導波管内で光を伝搬するように構成され得る。導波管440b、438b、436b、434b、432bの各々は、主要な上部および底部表面と、それらの主要上部表面と底部表面との間に延びている縁とを伴う平面である、または別の形状(例えば、湾曲)を有し得る。図示される構成では、導波管440b、438b、436b、434b、432bの各々は、光を向け直し、各それぞれの導波管内で伝搬させ、導波管から外に画像情報を眼410に出力することによって、光を導波管から外に抽出するように構成される光抽出光学要素440a、438a、436a、434a、432aを含み得る。抽出された光は、外部結合光とも称され得、光抽出光学要素は、外部結合光学要素とも称され得る。抽出される光のビームは、導波管内を伝搬する光が光向け直し要素に衝突する場所において導波管によって出力される。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り得る。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、434

b、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、上部もしくは底部主要表面に配置され得るか、または導波管440b、438b、436b、434b、432bの体積内に直接配置され得る。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得る。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り得、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得る。

#### 【0034】

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、眼に最も近い導波管432bは、そのような導波管432bの中に投入されると、コリメートされた光を眼410に送達するように構成され得る。コリメートされた光は、光学無限遠焦点面を表し得る。次の導波管434bは、眼410に到達し得る前、第1のレンズ452（例えば、負のレンズ）を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得る。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得る。同様に、第3の上方の導波管436bは、眼410に到達する前、その出力光を第1のレンズ452および第2のレンズ454の両方を通過させる。第1および第2のレンズ452、454の組み合わせられた屈折力は、眼/脳が、第3の導波管436bから生じる光が次の上方の導波管434b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近い第2の焦点面から生じるように解釈するように、波面曲率の別の漸増量を生成するように構成され得る。

#### 【0035】

他の導波管層（例えば、導波管438b、440b）およびレンズ（例えば、レンズ456、458）も同様に構成され、スタック内の最も高い導波管440b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化する。スタックされた導波管アセンブリ480の他側の世界470から生じる光を視認/解釈するとき、レンズ458、456、454、452のスタックを補償するために、補償レンズ層430が、スタックの上部に配置され、下方のレンズスタック458、456、454、452の集約力を補償し得る。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焦点合わせ側面の両方は、静的であり得る（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、一方または両方は、電気活性特徴を使用して動的であり得る。

#### 【0036】

図4を継続して参照すると、光抽出光学要素440a、438a、436a、434a、432aは、光をそれぞれの導波管から外に向け直すことと、導波管に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伴うこの光を出力することとの両方のために構成され得る。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得る。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440a、438a、436a、434a、432aは、特定の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り得る。例えば、光抽出光学要素440a、438a、436a、434a、432aは、体積ホログラム、表面ホログラム、および/または回折格子であり得る。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 3 7 】

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、比較的到低回折効率を有し、それによって、ビームの光の一部のみが、DOE の各交差によって眼 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。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、その結果は、導波管内で跳ね返るこの特定のコリメートされたビームに対して、眼 3 0 4 に向う非常に均一なパターンの出射放出である。

## 【 0 0 3 8 】

いくつかの実施形態では、1 つ以上の DOE は、それらが能動的に回折する「オン」状態と、有意に回折しない「オフ」状態との間で切り替え可能であり得る。例えば、切り替え可能な DOE は、ポリマー分散液晶の層を備え得、その中で、微小液滴は、ホスト媒体中の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合致するように切り替えられることができ（その場合、パターンは、入射光を感知できるほどに回折しない）、微小液滴は、ホスト媒体のそれに合致しない屈折率に切り替えられることもできる（その場合、パターンは、入射光を能動的に回折する）。

## 【 0 0 3 9 】

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向きに基づいて、動的に変動させられ得る。被写界深度は、視認者の瞳孔サイズと反比例して変化し得る。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、増加し、それによって、その平面の場所が眼の焦点深度を越えるので判別不能である 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、より焦点が合って見え得る。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間隔を置かれた深度平面の数は、減少させられた瞳孔サイズに伴って減少させられ得る。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こともある。しかしながら、これらの 2 つの深度平面は、別の瞳孔サイズにおけるユーザに対して、遠近調節を変化させることなく同時に十分に焦点が合っていることもある。

## 【 0 0 4 0 】

いくつかの実施形態では、ディスプレイシステムは、瞳孔サイズまたは向きの決定に基づいて、または、特定の瞳孔サイズまたは向き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得る。例えば、ユーザの眼が、2 つの導波管に関連付けられた 2 つの深度平面間を区別不能である場合、コントローラ 4 6 0（ローカル処理およびデータモジュール 2 6 0 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、それは、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り替え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、オフ状態に切り替えられ得る。

## 【 0 0 4 1 】

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、この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、この条件は、視認者の瞳孔のサイズの決定に応答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、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せられ得る。

## 【 0 0 4 2 】

ウェアラブルシステム 4 0 0 は、世界 4 7 0 の一部を結像する外向きに面した結像シス

10

20

30

40

50

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のこの部分は、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と同一であることも、そうでないこともあり、それは、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ので、ウェアラブルシステム４００を包囲する４ ステラジアン（steradian）の立体角を含み得る。他のコンテキストでは、装着者の移動は、より抑制され得、それに応じて、装着者のＦＯＲは、より小さい立体角に対し得る。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。

10

#### 【００４３】

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。上で説明されるように、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、別の例として、指向性マイクロホンを備えていることができ、それは、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も提供することができる。ウェアラブルシステム４００は、発話源を位置特定することにおいて、または特定の瞬間におけるアクティブ話者を決定するため等に、外向きに面した結像システム４６４およびオーディオセンサ２３０の両方からの情報を使用することができる。例えば、ウェアラブルシステム４００は、単独で、または話者の反射された画像（例えば、鏡に見られるように）と組み合わせて、音声認識を使用し、話者の識別を決定することができる。別の例として、ウェアラブルシステム４００は、指向性マイクロホンから入手された音に基づいて、環境内の話者の位置を決定することができる。ウェアラブルシステム４００は、発話認識アルゴリズムを用いて、話者の位置から生じる音を解析し、発話のコンテンツを決定し、音声認識技法を使用して、話者の識別（例えば、名前または他の人口統計情報）を決定することができる。

20

30

#### 【００４４】

ウェアラブルシステム４００は、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）も含むことができる。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の瞳孔のサイズおよび／または向きを決定するために使用され得る。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）を決定することにおける使用のために、またはユーザのバイオメトリック識別のために（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、各眼のために、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得る。いくつかの他の実施形態では、単一の眼４１０のみの瞳孔直径または向きが、決定され（例えば、対の眼あたり単一カメラのみを使用して）、ユーザの両眼に対して同様であると仮定される。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得る。ウェアラブルシステム４００は、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向き）も決定し得る。

40

#### 【００４５】

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

50



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、トーテム (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な平行移動 (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感じることができる。平行移動をサポートするマルチ DOF コントローラは、3 DOF 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、6 DOF と称され得る。ある場合、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す、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 に提供し得る (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 の使用中、ユーザの手によって保持され得る。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。

#### 【0046】

図 5 は、導波管によって出力された出射ビームの例を示す。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、同様に機能し得、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 432b の入力縁 432c において導波管 432b の中に投入され、TIR によって導波管 432b 内を伝搬する。光 520 が DOE 432a 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、実質的に平行として図示されるが、それらは、導波管 432b に関連付けられた深度平面に応じて、ある角度で眼 410 に伝搬するようにも向け直され得る (例えば、発散出射ビーム形成)。実質的に平行出射ビームは、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見え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組は、より発散する出射ビームパターンを出力し得、それは、眼 410 がより近い距離に遠近調節し、それを網膜上に焦点に集めることを要求し、光学無限遠より眼 410 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。

#### 【0047】

図 6 は、導波管装置を含む光学システムと、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、制御サブシステムとを示す概略図であり、それらは、多焦点立体ディスプレイ、画像、または明視野の生成において使用される。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、または明視野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 632a (1つのみのが図 6 に示される) と、一次導波管 632a の少なくともいくつかの各々に関連付けられた 1つ以上の DOE 632b とを含むことができる。平面導波管 632b は、図 4 を参照して議論される導波管 432b、434b、436b、438b、440b に類似することができる。光学システムは、分配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得る。分配導波管装置は、例えば、分配平面導波管 622b と、分配平面導波管 622b に関連付けられた少なくとも 1つの DOE 622a (二重破線によって図示される) とを含み得る。分配平面導波管 622b は、少なくともいくつかの点において、それと異なる向きを有する一次平面導波管 632b と類似または同じであり得る。同様に、少なくとも 1つの DOE 622a は、少なくともいくつかの点において、DOE 632a と類似または同じであり得る。例えば、分配平面導波管 622b または DOE 622a は、それぞれ、一次平面導波管 632b または DOE 632a と同一材料から成り得る。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 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 4 8 】

中継され、射出瞳が拡張された光は、分配導波管装置から 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、分配平面導波管 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得る一次平面導波管 6 3 2 b にその光を通すことができる。

10

【 0 0 4 9 】

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 を含み得る。光ファイバ 6 4 0 の遠位端は、圧電材料の中空管 6 4 2 に通され、または受け取られ得る。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）に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、または直径にメッキされ得る。コア電極（図示せず）も、管 6 4 2 のコア、中心、内側周縁、または内径に位置し得る。

【 0 0 5 0 】

20

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 つの軸において曲げる。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、大きな撓みを通して掃引することができる。

【 0 0 5 1 】

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 6 4 4 の先端は、2 次元（2 - D）走査を占めるエリア内において 2 軸方向に走査される。光源 6 1 0 の強度をファイバカンチレバー 6 4 4 の走査と同期して変調することによって、ファイバカンチレバー 6 4 4 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 2 0 1 4 / 0 0 0 3 7 6 2 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。

30

【 0 0 5 2 】

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 6 4 4 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 6 4 8 によって、少なくとも 1 つの回折光学要素（DOE）6 2 2 a を含む狭い分配平面導波管 6 2 2 b 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 によって分配平面導波管 6 2 2 b 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、DOE 6 2 2 a と繰り返し交差することができる。DOE 6 2 2 a 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%）が DOE 6 2 2 a との交差の各点においてより大きい一次平面導波管 6 3 2 b の縁に向かって回折されるようにし、光の一部が TIR を介して分配平面導波管 6 2 2 b の長さに沿ってそのオリジナル軌道上で継続するようにする。

40

【 0 0 5 3 】

DOE 6 2 2 a との交差の各点において、追加の光が、一次導波管 6 3 2 b 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組に分割することによって、光の射出瞳は、分配平面導波管 6 2 2 b 内の DOE 6 2 2 a によって垂直に拡張されることができる。分配平面導波管 6 2 2 b から外部結合されたこの垂直に拡張された光は、一次平面導波管 6 3 2 b の縁に進入することができる。

50

## 【0054】

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、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につれて、複数の点において DOE 632a と交差する。DOE 632a は、有利には、線形回折パターンおよび放射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、光の偏向および集中の両方を生成するように設計または構成され得る。DOE 632a は、有利には、ビームの光の一部のみが、DOE 632a の各交差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。

10

## 【0055】

伝搬する光と DOE 632a との間の交差の各点において、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 から逃れ、一次導波管 632b の面から出て来ることを可能にする。いくつかの実施形態では、DOE 632a の放射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与え、個々のビームの光波面を成形（例えば、曲率を与える）することと、ビームを設計された焦点レベルに合致する角度に操向することとの両方を行う。

## 【0056】

故に、これらの異なる経路は、異なる角度における DOE 632a 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、光が一次平面導波管 632b の外で結合されることができるようになる。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明視野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組（例えば、3 層）が、それぞれの色（例えば、赤色、青色、緑色）を生成するために採用され得る。したがって、例えば、第 1 の 3 つの隣接する層の組が、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得る。第 2 の 3 つの隣接する層の組が、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得る。複数の組が、種々の焦点深度を伴うフル 3D または 4D カラー画像明視野を生成するために採用され得る。

20

## 【0057】

（ウェアラブルシステムの他のコンポーネント）

多くの実装では、ウェアラブルシステムは、上で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み得る。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み得る。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り得る。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得る。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得るか、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）の感覚を再現し得る。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得る（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得る。

30

40

## 【0058】

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み得る。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り得る。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、ロックスイッチ）を有していないこともある。代わりに、トーテムは、単に、物理的表面を提供し得、ウェア

50

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得る。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得る。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得る。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り得る。ユーザは、単独で、または姿勢と組み合わせで、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得る。

【0059】

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。

【0060】

(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)

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明視野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得る。仮想世界をマッピングすることにおいて、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ら捕捉されたFov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(2D点または3D点等)の組を収集し、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のユーザを包囲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。

【0061】

図7は、MR環境700の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、室内カメラ等の静止入力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)または静止室内システム(例えば、室内カメラ等)から受信するように構成され得る。ウェアラブルシステムは、種々のセンサ(例えば、加速度計、ジャイロスコープ、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。この情報は、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静止カメラからの情報でさらに補完され得る。カメラ(室内カメラ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、マッピング点の組にまとめられ得る。

【0062】

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の助けを借りて意味情報をオブジェクトに添付することができる。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得る。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、

10

20

30

40

50

WAN等)を通して互いに接続され、クラウドにアクセスすることができる。

【0063】

この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得る。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専用に使われ得る。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得る一方、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得る。

10

【0064】

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得る。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しるしを識別する)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得る。コンピュータビジョンアルゴリズムの非限定的例は、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)、方向付きFASTおよび回転BRISK(ORB)、パ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-Kanadeアルゴリズム、Horn-Schunckアルゴリズム、Mean-shiftアルゴリズム、視覚的同時位置推定およびマッピング(vSLAM)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(ICP)、セミグローバルマッチング(SGM)、セミグローバルブロックマッチング(SGBM)、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。

20

30

【0065】

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。訓練されると、機械学習アルゴリズムは、HMD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)、インスタンスベ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、Perceptron等)、深層学習アルゴリズム(例えば、Deep Boltzmann Machine、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、Stacked Generalization等)、および/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、個々のデータ組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)、データ組(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追加の画像の組)、条件付き状況、または他の変形例に特定の追加のモデルを生成し得る。いくつかの実施形態では

40

50

、ウェアラブルHMD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み得る。

#### 【0066】

マップデータベース内のこの情報および点の集合に基づいて、オブジェクト認識装置708a-708nは、オブジェクトを認識し、オブジェクトを意味情報で補完し、生命をオブジェクトに与え得る。例えば、オブジェクト認識装置が、点の組がドアであることを認識する場合、システムは、いくつかの意味情報を添付し得る（例えば、ドアは、ヒンジを有し、ヒンジを中心として90度移動を有する）。オブジェクト認識装置が、点の組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情報を添付し得る。意味情報は、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、意味情報は、オブジェクトの法線を含み得る。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か、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。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され得る。例えば、MR環境700は、Californiaで起こっている場面についての情報を含み得る。環境700は、New Yorkにおける1人以上のユーザに伝送され得る。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「引き渡され」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、場所特定目的のために、トポロジマップも使用し得る。

#### 【0067】

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得る。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欲し得るか、またはCaliforniaに存在する友人と散歩に行くことを欲し得る。

#### 【0068】

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得る。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得る。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、この情報に基づいて、まばらな点を決定し得る。まばらな点は、ユーザの周囲における種々のオブジェクトの向きおよび位置を表示およびに理解することにおいて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）を決定することにおいて使用され得る。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ジングし、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得る。この情報は、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得、所望の仮想場面が、ブロック850において、適宜、ユーザに表示され得る。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、適切な向き、位置等において表示され得る。

#### 【0069】

図9は、ウェアラブルシステムの別の例のブロック図である。この例では、ウェアラブルシステム900は、マップ920を備え、それは、世界のためのマップデータを含むマップデータベース710を含み得る。マップは、ウェアラブルシステム上にローカルに部

分的に常駐し得、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）に部分的に常駐し得る。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向きを決定するために、マップ 920 からのデータを利用し得る。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ように、オンザフライで収集されたデータから算出され得る。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てい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得る。

【0070】

10

まばらな点の表現は、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り得る。システムは、世界内の種々のコンポーネントの場所のみではなく、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、マップにデータ投入すること、およびマップからのデータを使用することを含む多くの目標を達成する構築ブロックであり得る。

【0071】

一実施形態では、まばらな点の位置は、それ自体では完全に適正でないこともあり、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高密度表現が、少なくとも部分的にこの間隙を充填するために利用され得る。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得、深度情報が、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ト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、一緒に融合され得、このうちのいくつかは、表面表現を用いて要約され得る。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、大点群に対して）かつ要約しやすい入力であり得る。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得る。姿勢 910 は、同様に、この融合プロセス 930 への入力であり得、融合 930 の出力は、マッププロセス 920 にデータ投入するための入力となる。サブ表面が、トポグラフィマッピング等において、互いに接続し、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。

20

30

【0072】

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、利用され得る。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こと、または種々の条件下で逃げていること（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等を決定するための入力であり得る。世界マップは、オブジェクトの場所に関する情報またはオブジェクトの意味情報を含み得、世界マップは、複合現実に対する別の有用な入力であることができる。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。

40

【0073】

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。動き回るために、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、ユーザが行うことを欲すること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はなく、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

50

よって追跡され得る。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、場所および向きだけではなく、ユーザが、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得、それは、そのようなアクティビティがどのカメラの視野内にないときでも、生じている状況を決定することを補助し得る。）

手のジェスチャ追跡または認識も、入力情報を提供し得る。ウェアラブルシステム900は、ボタン押し下げに対して、左または右、停止、握持、保持等をジェスチャすることに対して、手のジェスチャを追跡および解釈するように構成され得る。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またはカレンダーをめくることが、または別の人物もしくはプレーヤと「フィストバンプ」を行うことを欲し得る。ウェアラブルシステム900は、動的であることも、そうでないことも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得る。例えば、ジェスチャは、停止を示すために手を広げること、OKを示すために親指を上げること、OKではないことを示すために親指を下げることに、または方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こと等、単純な静的ジェスチャであり得る。

#### 【0074】

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、特定の深度または範囲にレンダリングする）。一実施形態では、眼の両眼離反運動が、三角測量を使用して決定され得、次いで、その特定の人物のために開発された両眼離反運動/遠近調節モデルを使用して、遠近調節が、決定され得る。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向き）を決定することができる。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、他の技法も、眼追跡のために使用されることができる。

#### 【0075】

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、音声認識を含み得る。システム900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900は、発話している人物（例えば、発話がARDの装着者または別の人物もしくは音声（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と、言われていることを決定するための発話認識技術と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260または遠隔処理モジュール270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（DTW）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。ある場合、発話は、複数の源から、例えば、ユーザの近傍の別の人物から、人物の近傍で放送されているテレビ上のアナウンサーから、およびスピーカ240を介してARDのユーザに再生されている発話コンテンツから生じるであろう。以下にさらに説明されるように、これらの異なる発話源（例えば、この例では、人物、テレビのアナウンサー、およびオーディオストリーム）は、コンテンツ分析され得、異なる話題が、ARDのユーザインターフェースによって、ユーザに異なるように提示され得る（例えば、異なるスレッドの中に編成される異なる話題、異なるスレッドの中に編成される異なる話者による発話、またはこれらの組み合わせ）。

#### 【0076】



ローカルデータおよび処理モジュール260または遠隔処理モジュール270も、音声認識アルゴリズムを適用することができ、それは、話者がウェアラブルシステム900のユーザ210であるか、ユーザが会話している別の人物である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(DTW)技法を含むことができる。音声認識技法は、コホートモデルおよび世界モデル等のアンチ話者技法も含むことができる。スペクトル特徴は、話者特性を表すことにおいて使用され得る。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270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。

10

#### 【0077】

ワードフロー注釈システム970の実装は、ユーザインターフェース(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、使用され、例えば、補助情報の表示を閉じること、または単語を一般的単語辞書に追加することができる。そのような実装およびこれらの使用の例は、以下にさらに説明される。

#### 【0078】

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つの対のカメラを含むことができる: 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; ユーザの正面に向けられ、立体視結像プロセス940をハンドリングし、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を捕捉することも行うための異なる対のカメラ。ステレオプロセス940に対する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)の一部であり得る。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向けられる眼追跡カメラ(図4に示される内向きに面した結像システム462の一部であり得る)を含むことができる。ウェアラブルシステム900は、1つ以上のテクスチャ光プロジェクト(赤外線(IR)プロジェクト等)を備え、テクスチャを場面の中に投入し得る。

20

#### 【0079】

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方法1000の例のプロセスフロー図である。この例では、ユーザは、トーテムと相互作用し得る。ユーザは、複数のトーテムを有し得る。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得る。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得る。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得るか、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)を通して検出され得る。

30

#### 【0080】

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、向き、または移動を検出する。基準フレームは、マップ点の組であり得、それに基づいて、ウェアラブルシステムは、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。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。

40

#### 【0081】

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、仮想ページをめくり、次のページに移動すること、または1つのユーザインターフェース(UI)デ

50

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得る。別の例として、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得る。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得る。いくつかの実装では、ユーザの眼の両眼離反運動が、追跡されることができ、遠近調節/両眼離反運動モデルが、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、ユーザの眼は、ユーザが集中している深度平面上に情報を提供する。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こと、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。

10

#### 【0082】

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2におけるディスプレイ220等）によって投影され得る。それは、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得る。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得る。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）追跡され得る。

20

#### 【0083】

図11は、仮想ユーザインターフェースと相互作用する方法1100の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得る。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る。

#### 【0084】

ブロック1110では、ウェアラブルシステムは、特定のUIを識別し得る。UIのタイプは、ユーザによって事前に決定され得る。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）に基づいてデータ投入される必要があることを識別し得る。UIは、セキュリティシナリオに特定であることができ、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得る。例えば、UIの境界、一般的構造、形状等に関連付けられたデータが、生成され得る。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得る。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か、または平面UIが壁上もしく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、頭部姿勢、または眼姿勢の座標を決定し得る。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包囲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、決定され得る。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得る。

30

40

#### 【0085】

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得るか、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得る。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、明視野ディスプレイは、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。仮想UIが生成されると、ウェアラブルシステムは、ブ

50

ロック 1 1 5 0 において、単に、より多くの仮想コンテンツを仮想 UI 上に生成するユーザからのコマンドを待ち得る。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り得る。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）を待ち得、認識される場合（ブロック 1 1 6 0）、コマンドに関連付けられた仮想コンテンツが、ユーザに表示され得る（ブロック 1 1 7 0）。

#### 【 0 0 8 6 】

ウェアラブルシステム、UI、およびユーザ体験（UX）の追加の例は、米国特許公開第 2 0 1 5 / 0 0 1 6 7 7 7 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。

#### 【 0 0 8 7 】

（複数のウェアラブルシステム間の例示的通信）

図 1 2 は、互い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 1 2 0 0 は、ユーザデバイス 1 2 3 0 a、1 2 3 0 b、1 2 3 0 c を含む。ユーザデバイス 1 2 3 0 a、1 2 3 0 b、および 1 2 3 0 c は、ネットワーク 1 2 9 0 を通して、互いに通信することができる。ユーザデバイス 1 2 3 0 a - 1 2 3 0 c の各々は、ネットワークインターフェースを含み、ネットワーク 1 2 9 0 を介して、遠隔コンピューティングシステム 1 2 2 0（ネットワークインターフェース 1 2 7 1 も含み得る）と通信することができる。ネットワーク 1 2 9 0 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り得る。コンピューティング環境 1 2 0 0 は、1 つ以上の遠隔コンピューティングシステム 1 2 2 0 も含むことができる。遠隔コンピューティングシステム 1 2 2 0 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み得る。ユーザデバイス 1 2 3 0 a、1 2 3 0 b、および 1 2 3 0 c は、ネットワーク 1 2 9 0 を介して、遠隔コンピューティングシステム 1 2 2 0 と通信し得る。

#### 【 0 0 8 8 】

遠隔コンピューティングシステム 1 2 2 0 は、遠隔データリポジトリ 1 2 8 0 を含み得、それは、特定のユーザの物理的および / 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 1 2 8 0 は、一般的単語辞書、補助情報源等のワードフロー注釈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 2 に示される遠隔データリポジトリ 2 8 0 の実施形態であり得る。遠隔コンピューティングシステム 1 2 2 0 は、遠隔処理モジュール 1 2 7 0 も含み得る。遠隔処理モジュール 1 2 7 0 は、図 2 に示される遠隔処理モジュール 2 7 0 の実施形態であり得る。遠隔処理モジュール 1 2 7 0 は、1 つ以上のプロセッサを含み得、それは、ユーザデバイス（1 2 3 0 a、1 2 3 0 b、1 2 3 0 c）および遠隔データリポジトリ 1 2 8 0 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 2 6 0（図 2 に示されるような）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 1 2 2 0 は、所与のユーザが、特定のユーザ自身の物理的および / 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得る。

#### 【 0 0 8 9 】

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 D 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り得る。例えば、ユーザデバイス 1 2 3 0 b および 1 2 3 0 c は、図 2 に示されるウェアラブルシステム 2 0 0（または図 4 に示されるウェアラブルシステム 4 0 0）の実施形態であり得、それは、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。

#### 【 0 0 9 0 】

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 4 6 6 と共に使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 4

10

20

30

40

50

6 4 を使用して)。ユーザデバイスおよび／または遠隔コンピューティングシステム 1 2 2 0 は、ユーザから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 1 2 2 0 に送信し得る。ユーザデバイスは、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 1 2 2 0 に送信し得る。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 1 2 2 0 から受信し、ユーザに投影する前、最終処理を提供し得る。ユーザデバイスは、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得る。ユーザデバイスは、入手された情報を処理する一方、遠隔データリポジトリ 1 2 8 0 と通信し得る。複数のユーザデバイスおよび／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および／または処理に関与し得る。

10

#### 【0091】

物理的世界に関する情報は、経時的に展開され得、異なるユーザデバイスによって収集される情報に基づき得る。仮想世界のモデルも、経時的に展開され、異なるユーザの入力に基づき得る。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 1 2 1 0 を構築するために使用され得る。世界マップ 1 2 1 0 は、図 9 に説明されるマップ 9 2 0 の少なくとも一部を含み得る。種々のオブジェクト認識装置（例えば、7 0 8 a、7 0 8 b、7 0 8 c ... 7 0 8 n）が、オブジェクトおよびタグ画像を認識するために、および、意味情報をオブジェクトに添付するために使用され得る。これらのオブジェクト認識装置は、図 7 にも説明される。

20

#### 【0092】

遠隔データリポジトリ 1 2 8 0 は、データを記憶し、世界マップ 1 2 1 0 の構築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 1 2 1 0 についての情報を受信することができる。世界マップ 1 2 1 0 は、ユーザまたは別の人物によって作成され得る。本明細書に議論されるように、ユーザデバイス（例えば、1 2 3 0 a、1 2 3 0 b、1 2 3 0 c）および遠隔コンピューティングシステム 1 2 2 0 は、単独で、または組み合わせて、世界マップ 1 2 1 0 を構築および／または更新し得る。例えば、ユーザデバイスは、遠隔処理モジュール 1 2 7 0 および遠隔データリポジトリ 1 2 8 0 と通信し得る。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入手および／または処理し得る。遠隔処理モジュール 1 2 7 0 は、遠隔データリポジトリ 1 2 8 0 およびユーザデバイス（例えば、1 2 3 0 a、1 2 3 0 b、1 2 3 0 c）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得る。遠隔コンピューティングシステム 1 2 2 0 は、例えば、ユーザの画像を選択的にクロッピングすること、ユーザの背景を修正すること、仮想オブジェクトをユーザの環境に追加すること、ユーザの発話に補助情報で注釈を付けること等、ユーザデバイス（例えば、1 2 3 0 a、1 2 3 0 b、1 2 3 0 c）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 1 2 2 0 は、処理された情報を同一および／または異なるユーザデバイスに送信することができる。

30

#### 【0093】

ワードフロー注釈システムの実施形態の種々の機能性が、ここで説明されるであろう。

40

#### 【0094】

（例示的発話捕捉）

開示されるワードフロー注釈システムの例示的実装は、対面通信のコンテキストにおいて生じる。そのようなコンテキストでは、検出された発話の源は、ワードフロー注釈システムの物理的近傍にある。システムのオーディオセンサは、周囲オーディオ情報を検出することができる。コンピューティングモジュール（例えば、ワードフロー注釈システム 9 7 0）は、発話を周囲オーディオ情報から検出することができる。このモジュールは、本明細書に説明されるスキームを含む自動発話認識のアルゴリズム（図 9 を参照して説明されるように）を実装することができる。コンピューティングモジュールは、ユーザの発話

50

(例えば、発話をテキストに変換することによって)を解析し、ユーザの発話中のキーワードを識別することができる。キーワードは、稀な単語またはコンテキストキーワードを含むことができる。キーワードは、単独で、または組み合わせて、キーワードデータベースにアクセスすることによって、またはコンテキスト情報を決定することによって、識別され得る。コンテキスト情報は、ユーザもしくは会話の他の参加者に関連付けられた情報(例えば、人口統計情報、過去のアクティビティ/体験、発話挙動等)、または環境に関連付けられた情報(例えば、会話の場所、環境内のオブジェクトもしくは人々等)を含み得る。キーワードは、ウェアラブルシステム200のユーザへの提示のための仮想コンテンツを決定するために使用され得る。例えば、ウェアラブルシステムは、キーワードに関連付けられた補助情報(例えば、説明テキストまたは画像等)を提示することができる。

10

#### 【0095】

開示されるワードフロー注釈システムの別の例示の実装は、遠隔通信のコンテキストにおいて生じる。そのような例示的シナリオでは、話者は、ワードフロー注釈システム(およびそのユーザ)から物理的に遠隔に存在する。話者に関連付けられたオーディオ情報は、多くのチャネルのうちの1つを通して、聴取者に到着し得る。例えば、話者および聴取者は、電話を介して、またはインターネットオーディオもしくはオーディオ-ビデオチャットセッションを通して、会話していることもある。話者および聴取者は、図12に図示されるように、ネットワークを通して通信するARシステムを使用して、会話していることもある(例えば、テレプレゼンスセッション等)。別の例として、話者は、無線、テレビ、またはインターネットプログラミング等上に示される者等、オーディオまたはA/Vプログラム内の人物であり得る。話者の音声、聴取者の場所において周囲音の中で聞こえる場合、ワードフロー注釈システムは、対面通信コンテキストと同一方法で音を捕捉することができる。話者の音声、聴取者の場所において周囲音の中で聞こえない(例えば、ワードフロー注釈システム970のオーディオ増幅器232を通して、聴取者に話者の声が聞こえる)場合、システムは、異なる方法で音を捕捉するか、または音を捕捉するステップをスキップすることができる(例えば、システムは、直接、発話に関連付けられた情報が内蔵される電気信号から、発話を検出することができる)。遠隔通信使用のさらなる説明は、以下のテレプレゼンスと題された小節にある。

20

#### 【0096】

(話者音声否認および源位置特定)

30

ARシステムは、音をその源またはその源の場所に関連付ける能力を通して、追加の有用用途を実施することができる。いくつかのそのような用途は、以下に説明される。

#### 【0097】

例示の実装は、検出された発話がユーザ自身の発話であるかどうかを決定する。これは、複数のオーディオセンサ、例えば、2つ以上のマイクロホンの使用を通して、または指向性オーディオセンサ、例えば、指向性マイクロホンの使用を通して、達成されることができる。そのようなオーディオセンサを用いて、システムは、音源の場所を識別することができる。したがって、システムは、ユーザ自身の発話を他者のものと区別することができる。システムは、稀な単語検出のために、ユーザ自身の発話を処理しないように構成されることができる。

40

#### 【0098】

ARシステムの実装は、例えば、カメラ464および466、または静止カメラ(例えば、静止入力704)を通して供給される情報を用いて、音源の場所とユーザのFORに対する源の位置とを識別することができる。システムは、稀な単語検出のために、ユーザのFOV外の人物からの発話を処理しないように構成されることができる。

#### 【0099】

発話をその源に関連付けるための実装の能力は、コンテキスト依存補助情報を提供することにおいても有用であり得る。ある話者は、特定の分野またはコンテキストにおける単語を発する可能性がより高い。例えば、カメラ店の従業員が「SD」と言っているとき、それは、San Diegoの都市よりもセキュア化されたデジタルメモリカードについ

50

て話している可能性がより高い。したがって、話者をカメラ店従業員として識別可能な実装は、コンテキスト上正しい補助情報をユーザに提供することがより高くあり得る。話者を識別する能力は、いくつかの方法において実装され得る。例えば、発話（例えば、対話または独り言）中の参加者の名前は、多くの場合、発話の開始時の導入で述べられる。または、ユーザは、話者の名前をUI相互作用を通してシステムに入力し得る。外向きに面したカメラを装備するシステムは、話者の画像を捕捉し、画像認識を通して、話者を識別することができる。システムは、音声認識を実施し、図9に説明される種々の音声認識アルゴリズムを使用して、話者の識別を識別することもできる。

#### 【0100】

（発話からテキストへの例示的変換）

キーワードを識別するために、ウェアラブルシステム200は、オーディオストリームをテキストに変換することができる（例えば、ワードフロー注釈システム970によって）。発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得る）または図9を参照して説明される他の発話認識アルゴリズム等のアルゴリズムを使用して実施されることができる。ウェアラブルシステム200は、発話/テキスト機能性を含むことができるか（例えば、ローカル処理モジュール260内に）、または発話/テキスト機能性に遠隔でアクセスすることができる（例えば、遠隔処理モジュール270を介して）。ウェアラブルシステム200は、商業用発話/テキストサービスへの無線接続を利用し得る（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）。ワードフロー注釈技術は、任意のそのような発話/テキストアルゴリズムを利用し得るが、非慣習的な方法で利用し、発話/テキスト変換をユーザから生じていないオーディオに適用する。

#### 【0101】

ワードフロー注釈実装は、ローカルまたは遠隔で、例えば、ローカル処理およびデータモジュール260を使用するウェアラブルデバイス上で、または遠隔サーバ（例えば、遠隔コンピューティングシステム1220を含む）上で、発話からテキストへの変換を実施し得る。変換がローカルで実施されるか、遠隔で実施されるかにかかわらず、キーワード検出および補助情報読み出し等の他の処理ステップも、変換が行われる場所から独立して、ローカルまたは遠隔で行われ得る。例えば、発話/テキスト変換が、遠隔で行われ、キーワード検出がローカルで行われる場合、捕捉されたオーディオストリームは、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、そこで変換が行われる。変換されたテキスト文字列は、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。別の例として、発話/テキスト変換、キーワード検出、および補助情報読み出しが、遠隔で行われる場合、捕捉されたオーディオストリームは、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が、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。

#### 【0102】

本明細書に説明される例では、ARシステム200は、発話/テキスト変換を使用して、キーワードを検出することができるが、いくつかの実装では、ARシステム200は、オーディオデータを使用して、直接、キーワードを検出することができる。例えば、ARシステム200は、発話をテキストに変換する必要なく、キーワードの人物/ユーザの発音に基づいて、キーワードをルックアップすることができる。

#### 【0103】

（稀な単語の例示的検出）

ARシステムは、稀な単語等のキーワードに対して、変換されたテキストを検索することができる。例示的な稀な単語は、ARシステムのユーザに対して低頻度であるか、または一般的ではない単語を含むことができる。稀な単語は、限定ではないが、一般的に使用されない単語、別の言語からの単語（例えば、英語の文中で使用される中国語の単語「g

10

20

30

40

50

uanxi」)、技術用語、ARシステムのユーザが通常取り組むコンテキスト外の特  
殊分野からの単語等を含むことができる。例えば、ARシステムは、単語がARシステム  
によって検出された発話中で低頻度で使用される場合、単語を稀であると識別し得る。その  
ような稀な単語のための定義、翻訳、またはコンテキスト情報を提供することによって、  
ARシステムは、有利には、ユーザがユーザに聞こえる発話をより良好に理解することを  
補助することができる。

#### 【0104】

稀な単語は、一般的単語の既知のリストとの比較によって検出され得る。例えば、単語  
は、それが一般的単語辞書またはデータベース内に見出されない場合、稀であると見なされ  
得る。一般的単語辞書またはデータベースは、ローカルで記憶されるか(例えば、図2  
におけるデータモジュール71内に)、または遠隔で記憶され得る(例えば、遠隔データ  
リポジトリ74内に)。一般的単語辞書またはデータベースは、第三者源から利用可能で  
あり、例えば、インターネット上で利用可能であり得る。

#### 【0105】

稀な単語は、言語におけるその発声頻度または単語頻度リスト内の関連付けられたラン  
クを通して検出され得る。Zipfの法則は、発声頻度を単語頻度リスト内のランクと関  
連させる例であり、すなわち、単語の発声頻度は、単語頻度リスト内のそのランクに反比  
例する。稀な単語は、発話中の閾値頻度未満のその発生に基づいて、ARシステムによ  
って識別され得る。ワードフローシステムは、ユーザが、ユーザの言語知識に基づいて、単  
語を検出するようにシステムを構成することを可能にすることができる。例えば、ユーザ  
は、5,000または7,000個(または別の値)の最も一般的な単語のリストにない  
単語を検出するようにシステムを構成し得る。別の例として、ユーザは、その発生の見込  
みが言語中で0.01%、0.003%、0.001%、または別の値等の閾値未満であ  
る単語を検出するようにシステムを構成し得る。ARシステムは、デフォルト閾値(例え  
ば、0.005%)を搭載し得、それは、ユーザ構成可能である。典型的には、閾値が高  
いほど、より多くの単語が稀であるとフラグ付けされ、閾値が低いほど、より少ない単語  
がフラグ付けされるであろう。ユーザは、閾値を、稀な単語に関して提示される補助情報  
が、ユーザに有用であると考えられ、過度に注意を逸らすものではない値に設定すること  
ができる。ARシステムは、ユーザが、一時的に、閾値をゼロに設定する(例えば、稀な  
単語へのフラグ付けを停止する)ことを可能にし得、それは、ある状況では有用であり得  
る(例えば、ユーザが、慣れていない環境内で車両を運転している、ユーザが、注意を逸  
らされたくない宗教的もしくは文化的イベントに出席している、または、ユーザが、大部  
分の単語がユーザに対して稀な単語であり得るイベントに出席している(例えば、専門家  
ではないユーザが、話者が日常的に特殊医療用語および語句を話す、医療業界の講演に出  
席する)等)。

#### 【0106】

いくつかの実装では、稀な単語は、使用統計を使用閾値と比較することによって、単語  
の低頻度使用によって決定され、単語が、一般に使用される単語の使用閾値より低頻度で  
生じる場合、単語は、稀な単語としてフラグ付けされることができる。稀な単語および/  
または一般に使用される単語は、システムによって記憶されることができ、稀なおよび/  
または一般的単語のリストは、使用に基づいて、更新されることができる。

#### 【0107】

稀な単語は、ユーザの既知の言語における低頻度で使用される単語であり得る。例えば  
、単語「ossify」は、多くの英語話者に対して稀な単語であり得る。稀な単語は、  
ユーザに既知の言語と異なる言語における単語であり得る。例えば、図13Aに示される  
中国語語句「guanxi」は、多くの英語話者に対して稀な単語であり得る。稀な単語  
は、特殊分野における用語、例えば、薬物の名前、クロケット等の一般的ではないスポー  
ツにおいて使用される用語等であり得る。稀な単語は、頭字語、企業または非営利組織の  
名前、株式表示記号であり得る。稀な単語は、単一単語である必要はなく、単語のグルー  
プ(例えば、「free electron laser」)、語句(例えば、「bri

10

20

30

40

50

ng your own device」またはその頭字語形態「BYOD」)等であり得る。実際、ワードフロー注釈システムは、ユーザが、あるカテゴリの単語(上記の例のうちのいくつか等)を注釈のための候補として選定することを可能にするように構成されることができる。

#### 【0108】

稀な単語に加えて、またはその代替として、キーワードも、コンテキストキーワードであることができる。コンテキストキーワードは、会話の話題に関連する単語または会話中に頻繁に使用される単語を含み得る。例として、ARシステム200のユーザは、アフリカ(Africa)旅行についてその友人と話している場合がある。単語「Africa」は、英語言語では、稀な単語ではないこともあるが、ARシステム200は、会話がAfricaに関連するので、それをコンテキストキーワードとしてフラグ付けすることができる。別の例として、ユーザが、自分のアフリカ旅行について話している間、ユーザは、アフリカでの自分のサーフィン体験について述べる。故に、ユーザは、単語「surf」をその発話中で繰り返し使用し得る。ARシステム200は、文中で頻繁に使用されるので、単語「surf」をコンテキストキーワードとしてフラグ付けすることができる。

10

#### 【0109】

いくつかの単語は、コンテキストキーワードおよび稀な単語の両方であり得る。例えば、ある人物が、法理に関する講義をしているとき、語句「res judicata」が、複数回、現れ得る。語句「res judicata」は、ラテン語語句であり、日々の発話において一般に使用されないの、稀な単語と見なされ得るが、この語句は、法的専門家の発話中におけるその頻繁な発生に起因して、コンテキストキーワードとも見なされ得る。

20

#### 【0110】

ARシステムは、単語および語句間の関係を決定し、関係に基づいて、キーワードを検出することができる。例えば、ユーザの友人が、「I surfed in Africa」と言い得る。ARシステムは、単語「surf」および単語「Africa」の両方がキーワードである可能性が高いことを決定することができる。しかし、ユーザは、自分のアフリカ旅行について話していたので、ARシステムは、単語「surf」ではなく、単語「Africa」をキーワードとして識別し得る。しかしながら、ユーザが、サーフィンについてあまり把握していない場合、ARシステムは、単語「Africa」ではなく、単語「surf」をキーワードとして識別し得る。

30

#### 【0111】

キーワードは、事前に指定され得る。ARシステムは、ユーザまたは別の当事者(例えば、アプリケーション開発者、買物ウェブサイト等)によって指定されるコンテキストキーワードのデータベースを含むことができる。キーワードのデータベースは、ユーザによって編集可能であり、例えば、キーワード(もしくは稀な単語)を追加または除去し得る。例として、ARシステム200のユーザは、最近、Lindaと名付けられた赤ちゃんを出産していることもある。ユーザは、語句「my daughter Linda」をコンテキストキーワードとして設定することができる。したがって、ARシステムが、ユーザが「my daughter Linda」と言うことを検出すると、ARシステムは、例えば、赤ちゃんが睡眠中であるか、空腹であるか等、ユーザの赤ちゃんに関する補助情報を自動的に提示することができる。別の例として、ARシステムは、語句「utility bill」をキーワードとしてタグ付けし得る。故に、ARシステムは、発話中の語句「utility bill」の検出時、期限および額を含むユーザの公共料金の請求書(utility bill)を自動的に提示することができる。

40

#### 【0112】

キーワードは、ユーザ、環境、または会話の参加者に関連付けられたコンテキスト情報に基づいて、決定されることができる。例示的ワードフロー注釈システムは、少なくとも部分的に、キーワードの検出において、ユーザ挙動を利用することができる。例えば、ユーザは、ユーザが理解していない単語を聞くと、一時的に、ある方向を凝視し得る。その

50



ような例示的システムは、例えば、凝視に先立って生じる発話中の稀な可能性が高い単語に対する検索において、内向きに面したカメラ466を使用して、この特定の凝視を検出することができる。別の例として、ユーザは、別の人物によって話された単語を単独で繰り返し、システムに単語をキーワードとして処理するように促し得る。システムは、キーワードのその検出を補助するために、他のユーザ挙動を使用することができる。実際、実装は、ユーザが、キーワード検出における実装によって組み込まれる挙動をカスタマイズすることを可能にすることができる。例えば、ARシステムは、ユーザインターフェースを提供することができ、それによって、ユーザは、単語を一般的単語または稀な単語のリストに追加すること、もしくはそれから除去することができる。ARシステムは、稀な単語分類のために、ユーザ自身の発話を監視することができる。例えば、暗号使用者の発話は、頭字語AES (Advanced Encryption Standard) を含み得る。複数回 (例えば、3もしくは5回) または構成可能持続時間内で複数回 (例えば、5分以内に3回、1時間以内に7回等)、暗号使用者/ユーザの頭字語AESを検出すると、ARシステムは、頭字語AESをこのユーザのために一般的単語として再分類することができる。ARシステムは、それが一般的または稀な単語のリストを検出および更新する単語の頻度を監視することができる。例えば、単語「simvastatin」(コレステロール治療薬用) は、多くのユーザに対して稀な単語であり得るが、保健医療専門家であるユーザに対して、そうではない。保健医療専門家であるユーザのために、ARシステムは、この単語が、保健医療専門家のARシステムによって検出された発話中で低頻度で使用されないわけではないことを検出し、この単語をこのユーザのために一般的単語として再分類し得る。再分類は、例えば、単語をユーザのための一般的単語辞書に追加することによって達成され得る。ユーザのための一般的単語辞書は、図2に図示されるシステム200内の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に記憶され得る。

#### 【0113】

この例は、単語が稀な単語であるか、一般的単語であるかを分類することを参照して説明されたが、いくつかの実施形態では、分類/再分類は、一般に、キーワードのためである。ARシステムが、単語が複数回使用されたことを検出すると、または稀な単語が会話中に現れると、ARシステムは、単語をキーワードとしてフラグ付けし得る。例えば、ARシステムが、単語「petrify」がユーザの会話中に複数回使用されたことを検出し得るとき、ARシステムは、単語「petrify」をキーワードとして分類することができる。しかしながら、ARシステムは、後に、ユーザが化石に関連する会話に関わっていることを決定し得る。ARシステムは、ARシステムが、ユーザが単語「petrify」の意味を学習 (または別様に把握) したことを決定することができるので、単語「petrify」をもはやキーワードではなくなるように再分類し得る。

#### 【0114】

いくつかの実装では、稀な単語は、補助情報がユーザに頻繁に提示された場合、ユーザが稀な単語の意味を学習中であるという仮定の下、稀な単語に関連付けられた補助情報の提示が減少またはさらに停止する傾向であるように (少なくとも部分的に、減少係数に基づいて)、減少係数に関連付けられ得る。減少係数は、稀な単語が、その期間中に閾値回数を上回って述べられた場合、補助情報の表示が減少または停止するように、期間を含み得る。減少係数は、稀な単語が述べられた回数が数値カウントを超える場合、補助情報の表示が減少または停止するように、数値カウントを含み得る。減少係数、期間、または数値カウント (もしくは関連付けられた閾値) は、ユーザ構成可能であり得る。

#### 【0115】

キーワードに加えて、またはその代替として、ARシステムは、ユーザの環境内の着目オブジェクト (例えば、オブジェクトまたは人物等) を検出することもできる。ARシステムは、図7に説明される1つ以上のオブジェクト認識装置708を使用して、着目オブジェクトを検出することができる。例として、ARシステム200のユーザは、Floridaに住んでいる。ユーザは、休暇でCaliforniaに出掛け得る。ユーザは、

10

20

30

40

50

例えば、ARディスプレイ220を通して、Californiaで浜辺に横たわっているアシカを見ることがある。AR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、アシカを検出し、オブジェクト認識装置708を使用して、アシカを識別することができる。ARシステム200は、ユーザが、典型的には、Floridaでアシカに遭遇しないので、またはアシカを見ないので、アシカを着目オブジェクトとしてフラグ付けすることができる。

#### 【0116】

(補助情報の例示的読み出し)

ワードフロー注釈システムは、システムが検出する着目オブジェクトに関連付けられた補助情報を読み出すことができる。補助情報は、テキスト、画像、または他の聴覚的もしくは視覚的情報を含み得る。システムは、視覚的に(例えば、ディスプレイ220を介して)、オーディオ(例えば、オーディオ増幅器232またはスピーカ240を介して)で、触覚的(例えば、触知)フィードバックを介して等、補助情報を提示することができる。

10

#### 【0117】

ワードフロー注釈システムは、着目オブジェクトおよびその関連付けられた補助情報を含む情報のデータリポジトリ(例えば、データベース)を備え得る。例えば、データリポジトリは、一般的単語、稀な単語、他のコンテキストキーワード、ユーザの環境内の一般的オブジェクト(ユーザが、多くの場合、相互作用する)等を記憶し得る。補助情報は、意味情報(図7を参照して説明されるように)、意味または説明、オブジェクトに関連付けられた購入情報、ステータスまたは統計、ユーザが明白に知覚していないこともある着目オブジェクトの追加の詳細、もしくは他の情報を含むことができ、それは、着目オブジェクトの理解およびそれとの相互作用を促進する。補助情報の例は、図13A-13E、15A、および15Bにさらに説明される。

20

#### 【0118】

情報のそのようなデータベースは、ローカルで、例えば、図2におけるデータモジュール260内に記憶されるか、または遠隔で、例えば、遠隔データリポジトリ280内に記憶され得る。ワードフロー注釈システムは、公にアクセス可能な情報、例えば、インターネット上の情報を利用して、稀な単語に関連付けられた補助情報を読み出すことができる。この状況では、システムは、ネットワークにアクセスし、稀な単語のためのクエリを、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り得る。

30

#### 【0119】

データリポジトリ内の着目オブジェクトは、ARシステム200によって更新されることができる。例として、ARシステム200のワードフロー注釈システム970は、単語をシステムが稀な単語として識別する単語の組から除去するように構成されることができる。これは、例えば、ユーザが特殊知識を有する状況(例えば、薬剤師は、多くの薬物の名前を把握している可能性が高い)、ユーザが関連付けまたは使用を通して単語に慣れ親しんでいる状況(例えば、Acme Corporationの従業員は、Acme Corporationに関する補助情報を必要としない可能性が高い)において有用であり得る。例示的ワードフロー注釈システムは、単語を稀な単語の組から除去することができる。これは、単語を一般的単語辞書に追加することによって実装され得る。これは、システムが稀な単語を検出した後の追加のチェックとして実装され得る。検出された稀な単語が、除去されるべき単語の組内に見出される場合、システムは、単語に関連付けられた補助情報を読み出さないか、または表示しない。

40

#### 【0120】

ユーザは、UI相互作用を通して、特定の単語を稀な単語の組から除去し得る。ユーザ

50

の入力は、例えば、稀な単語およびその関連付けられた補助情報が表示されるとき、ユーザ入力デバイス 504 を通して入力され得る。例示的システムは、ある回数、例えば、3 または 5 回、もしくはユーザによって構成されるある回数にわたって、単語およびその関連付けられた補助情報を表示した後、特定の単語をその一般的単語辞書に追加するように構成されることができる。

#### 【0121】

UI 相互作用は、例えば、稀な単語を除去または追加するために使用され得る。例えば、ユーザは、「単語「phoropter」を一般的単語として追加する」と言い得、AR システムは、ユーザの発話を検出することができ、単語を一般的単語リストに追加し、および/または、それを稀な単語リストから除去することができる。

10

#### 【0122】

データリポジトリを更新するこれらの例は、稀な単語を参照して説明されたが、類似技法は、データリポジトリ内の他の着目オブジェクトを更新するためにも使用されることができる。

#### 【0123】

(補助情報との例示的相互作用)

AR システム 200 は、補助情報との種々のユーザ相互作用をサポートするように構成されることができる。例えば、AR システムは、補助情報を表示すること、閉じること、または呼び戻すことができる。ユーザは、補助情報の一部にマークすること、または補助情報に注釈を付けることができる。

20

#### 【0124】

AR システムは、着目オブジェクトの異なる発生に対して特定の着目オブジェクトに関連付けられた補助情報の異なる部分を表示するように構成されることができる。これは、すぐに表示され得るものより多くの補助情報が利用可能であるとき、有用であり得る。例えば、実装は、稀な単語の第 1 の発生時、補助情報をウェブサイト # 1 から読み出し、その第 2 の発生時、ウェブサイト # 2 から読み出す等を行うように構成されることができる。別の例として、実装は、その第 1 の発生時、薬物の使用を表示し、その第 2 の発生時、薬物の副作用を表示する等を行うように構成されることができる。表示する補助情報の部分を決定することにおいて有用なデータは、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 内に記憶され得る。補助情報の異なる部分をユーザに連続して提供することによって、AR システム 200 は、表示される補助情報がユーザの注意を逸らす(例えば、ディスプレイの FOV のあまりに多くを覆うことによって)、または難しすぎてユーザによって理解されない(例えば、時間が短すぎてユーザが情報の全てを読むことができない補助情報を表示することによって)であろう見込みを低減させることができる。

30

#### 【0125】

AR システムは、ユーザの構成に基づいて(例えば、ユーザの知識レベルに従って)、特定の稀な単語に関連付けられた異なる補助情報を表示するように構成されることができる。例えば、システムは、ユーザ(例えば、薬剤師)が、特殊薬物情報のためのシステムを構成している場合、薬物に関する特殊情報を読み出し、表示することができる。そうでなければ、システムは、薬物に関する一般的情報を読み出し、表示することができ、それは、特殊薬物情報のためのシステムを構成していないユーザのために適切である可能性が高い。システムは、異なる情報源を特殊または一般的情報に関連付け、適切な情報を提供することができる。例えば、辞書は、一般的情報源であり得、薬物ウェブサイトは、特殊情報源であり得る。

40

#### 【0126】

ワードフロー注釈システムは、発話のコンテキストに基づいて、稀な単語に関連付けられた補助情報を表示するように構成されることができる。稀な単語の近くのコンテキストからの追加の単語が、最も可能性が高い意味または最も有用であり得る補助情報の種類を決定するために使用されることができる。これは、ユーザに発話に関連する情報を提供す

50

るために、または稀な単語に関連付けられた曖昧性を解決することにおいて、有用であり得る。例えば、名前 Acme Corporation が、図 13D に図示されるように、会社役員のコンテキストを伴う発話中に現れる場合、実装は、ボックス 1340 に示されるように、Acme Corporation の役員の名前および肩書を返すことができる。名前 Acme Corporation が、図 13E に図示されるように、株価のコンテキストを伴う発話中に現れる場合、実装は、ボックス 1354 に示されるように、Acme Corporation の株式情報を返すことができる。別の例として、頭字語「IMDB」は、例えば、「internet movie database」、「immigrant database」、「in-memory database」等の種々の異なる意味を有し得る。ワードフロー注釈システムは、コンテキスト情報を決定し、正しい意味を決定することができる。一例では、ワードフロー注釈システムは、話者の識別を決定し、話者の識別と関係する意味を識別することができる。別の例では、ワードフロー注釈システムは、稀な単語を包囲する弁別的単語を使用して、それらの単語とその単語の特定の解釈に関連付けられた補助情報との間の相関を算出することができる。ワードフロー注釈システムは、最も一般的な N 個の単語の中にない単語を識別および解釈し、発話のコンテキストを決定することができる。上記の IMDB 例では、単語「IMDB」には、話者の発話中に「visa」、「legislation」、および「deportation」等の単語が付随する。全てのこれらの付随の単語は、「IMDB」の「internet movie database」解釈ではなく、「immigrant database」解釈に関連付けられる可能性が高い。したがって、AR システムは、「immigrant database」としての「IMDB」の補助情報を表示することを自動的に選定することができる。

#### 【0127】

ワードフロー注釈システムは、稀な単語のコンテキストと、ユーザのために読み出される、または表示されるべき候補である補助情報との間の相関を算出することができる。コンテキスト分析に関連付けられた処理が、ローカルまたは遠隔で、例えば、ローカル処理モジュール 71 または遠隔処理モジュール 72 によって、実施され得る。コンテキスト分析は、稀な単語を包囲する弁別的単語の使用に基づいて、それらの単語とその単語の特定の解釈に関連付けられた補助情報との間の相関を算出することができる。弁別的単語は、最も一般的な N 個の単語の中にはない単語であり得、ある N の値は、一般的単語辞書内のエントリの数または発声頻度ランクの閾値より小さくあり得る。高相関は、関連補助情報を示し得る。低相関は、非関連補助情報を示し得る。例えば、稀な単語が、「IMDB」であり、他の付随の単語が、「visa」、「legislation」、および「deportation」を含む場合、単語は、「Internet Movie Database」、「in-memory database」、またはある他の解釈の中ではなく、「Immigrant Database」解釈に対する補助情報の中で見出される可能性が高く、システムは、より高い相関に基づいて、「Immigrant Database」のための補助情報を表示するように構成されることができる。これは、コンテキスト依存補助情報のタイプである頭字語曖昧性解消の例である。

#### 【0128】

補助情報の表示は、種々の方法で提示され得る。例えば、システムは、図 13A におけるボックス 1308 に図示されるように、稀な単語および関連付けられた補助情報のみを表示することができる。システムは、発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。この構成では、例えば、ユーザの注意が一時的に逸らされた場合、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。本開示では、用語「補助情報」は、トランスクリプトを含む。

#### 【0129】

会話相手の一連の発話は、映画内の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得る。利用可能である場合、話者の名前も同様に、表示され得る。システムが話者の名前を取得し得る方法の例は、上で説明されている。実装は、情報（例え

10

20

30

40

50

ば、発話のトランスクリプトまたは補助情報)を発話の吹き出し、例えば、図13に図示されるボックス1308等の話者の近傍に幾何学的に局所化されたテキスト内に設置することができる。

#### 【0130】

発話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、太字にされたテキスト等、稀な単語をハイライトすることができる。システムは、ユーザが、UI相互作用を通して、現在または過去の稀な単語を選択し、関連付けられた補助情報を読み出すこと、または呼び戻すことを可能にすることができる。

#### 【0131】

実装は、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化するために、補助情報を設置し得る(例えば、発話の吹き出し内に、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、単純化され、ユーザは、その注意を話者から遠くに引き離す必要がない。補助情報は、読み取りアクションを会話相手に最小限に見えるようにするように設置され、そうすることによって、補助情報へのユーザのアクセスをあらわにすることなく、あまり注意を逸らさず、より良好な通信を提供し得る。例えば、話者の場所を決定可能な実装は、補助情報を話者の隣に設置し得る。ARシステムの外向きに面したカメラからの画像は、例えば、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を助けることができる。ARシステムは、図8に図示されるプロセッサー800を使用して、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2310において認識されたオブジェクトは、その発話が、注釈のために処理されるべき話者であり得る。

#### 【0132】

補助情報がオーディオで提示される場合、ユーザまたは会話相手によって被られる注意を逸らすものを低減させる別の例として、ARシステムは、ユーザに聞こえるために十分であるが、会話相手には聞こえない大きさの音量で情報を提示すること、および/または、ユーザもしくは会話相手のいずれも話していないときに情報を提示することができる。

#### 【0133】

読み出された補助情報は、ARシステム200のディスプレイによって、ユーザに表示されることができる。表示される補助情報は、条件が満たされるまで、見えたままであり得る。例えば、表示される情報は、固定時間量にわたって、次の稀な単語(または別の着目オブジェクト)が表示されるまで、またはユーザアクションによって閉じられるまで、見えたままであり得る。ユーザアクションは、受動的、例えば、眼移動であり得る。内向きに面したカメラ(例えば、図4におけるカメラ462)が、ユーザの眼移動を検出または追跡するために使用されることができる。表示は、ユーザが、システムによって、補助情報の表示エリア全体を追跡したことが観察された後、眼追跡を介して、閉じられ得る。例えば、補助情報が、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、左右および上下)を追跡することができる。表示は、ユーザが、システムによって、補助情報の表示エリアから眼を離れた(またはそれを見ていない)ことが観察された後、眼追跡を介して、閉じられ得る。ユーザアクションは、能動的、例えば、図4におけるユーザ入力デバイス504を通じた入力アクションであり得る。ワードフロー注釈システムは、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)相互作用の組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要素の形態をとり得る。ボタンは、実際の物理的ボタンまたはAR内に表示されるものであることができる。UI相互作用は、例えば、図4に関連して上で説明されるように、頭部姿勢の形態をとり得る。UI相互作用検出の例は、図10に関連して上で説明される。

#### 【0134】

ワードフロー注釈システムは、ユーザに、補助情報表示を閉じることを遅らせるように

10

20

30

40

50

促すことができる。例えば、システムは、表示される補助情報の明度を低減させ、または配色を変化させ、ディスプレイが、間もなく、例えば、数秒以内に閉じられるであろうことをユーザに通知し得る。上で説明されるもの等のUI相互作用は、表示を閉じることを延期するために使用され得る。例えば、ARシステムが、眼追跡を通して、ユーザが補助情報を積極的に視認中であることを検出する場合、ARシステムは、表示を閉じることを延期することができる。

#### 【0135】

上で説明されるもの等のUI相互作用は、閉じられた補助情報を呼び戻すために使用され得る。例えば、ユーザ入力デバイスを通した入力アクションが、直近で表示された補助情報を呼び戻すために使用されるか、または呼び戻すために、前に表示されていた補助情報の特定のアイテムを選択するために使用されることができる。

10

#### 【0136】

補助情報を表示している間、ARシステムは、ユーザが、補助情報に注釈を付けることを可能にすることができる。例として、ARシステムは、ユーザが、後に視認するために、補助情報の一部をマークすることを可能にすることができる。例えば、ユーザは、より詳細に読み込むために、稀な単語の説明の一部をハイライトすることができる。別の例として、ユーザが、プロバスケットボールの試合を鑑賞しており、コメンテータが特定の選手について話している間、ARシステムは、その選手によって公認された製品を補助情報として提示することができる。ユーザは、ARシステムによって提示された製品を保存またはブックマークし、製品を購入するかどうかを後に決定することができる。

20

#### 【0137】

ユーザは、補助情報に関連付けられた仮想コンテンツを追加または描写することによって、着目オブジェクトに注釈を付けることもできる。例えば、ユーザは、ユーザ入力デバイス466または手ジェスチャを使用して、メモもしくは図を補助情報に追加すること、または補助情報の一部を除去することができる（例えば、その部分を消去または×を付けることによって）。ARシステムは、ユーザの注釈を保存し、着目オブジェクトが後の時間に検出されると、ユーザの注釈を補助情報とともにレンダリングすることができる。

#### 【0138】

（着目オブジェクトとの相互作用のための例示的ユーザ体験）

図13A - 13Gは、着目オブジェクトとの相互作用および着目オブジェクトの補助情報の例を図示する。

30

#### 【0139】

（キーワードの存在に基づいて仮想コンテンツを提示する例）

図13Aは、キーワードを検出し、キーワードのための補助情報を提示する例を図示する。図13Aでは、ユーザのARシステムのワードフロー注釈システム970は、キーワードを検出し、キーワードの補助情報を提供するように構成されることができる。ARシステムのユーザに表示される情報を乱雑することを回避するために、ARシステムは、稀な単語の検出時、補助情報のみを提供することができる。

#### 【0140】

図13Aでは、ARDを装着しているユーザ（図示せず）が、女性1302を自分の環境内で知覚し得る。女性は、コールアウトボックス1304に示される文を話す。文内に、単語「guanxi」が存在し、それは、稀な単語であり得る。ARDは、女性によって話された文を捕捉し（例えば、オーディオセンサ232を介して）、文をテキストに変換し、「guanxi」を稀な単語として検出し、「guanxi」に関連付けられた補助情報（例えば、定義）を読み出し、補助情報をディスプレイ（例えば、ディスプレイ220）上に表示することができる。例証では、補助情報は、ボックス1308内に表示される。ユーザは、このボックス1308がARディスプレイ220によって提示される仮想コンテンツの一部として見る。ボックス1308内の補助情報は、ユーザが、単語「Guanxi」を使用することによって、女性が意味する内容を理解することに役立ち得る。

40

50

## 【 0 1 4 1 】

( テレプレゼンスセッションにおける例示的ユーザ体験 )

複数の A R システムの複数のユーザが、システムの補助を通して、遠隔で会話を実践し得る。例えば、図 1 2 に図示されるように、( 人工デバイスの補助なしでは、互いに直接見えない、または聞こえないような ) 2 つの物理的場所における ( デバイス 1 0 0 a および 1 0 0 b の ) 2 人のユーザの両者が、A R デバイスを装備することができ、それは、ワードフロー注釈システム 9 7 0 を装備し得る。各ユーザの発話は、ユーザのそれぞれのデバイスのオーディオセンサによって捕捉され、ネットワーク 1 2 9 0 を通して伝送され得る。ユーザ C の発話は、例えば、トランスクリプトとして、または稀な単語 + 補助情報として、ユーザ B のデバイス上に、その逆も同様に、表示され得る。

10

## 【 0 1 4 2 】

ワードフロー注釈デバイスは、発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 1 2 9 0 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、またはテキストを発話に逆変換するかのいずれかを行うことができる。これは、対応する発話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ので、ネットワーク 1 2 9 0 の帯域幅が制約される場合に有利であり得る。ワードフロー注釈システムは、ユーザ、すなわち、話者の音声情報を含み、その音声情報を使用して、テキストを発話に逆変換し得る。このように、再構成された発話は、話者自身の音声のように聞こえ得る。

## 【 0 1 4 3 】

ワードフロー注釈システムは、A R ディスプレイ 2 2 0 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もできる。例えば、A R ディスプレイは、遠隔話者のアバタを補助情報とともに提示し、参加者の視覚的ならびに聴覚的感覚に参与することができる。上の図 4 に関連して説明されるように、内向きに面した結像システム 4 6 2 を装備する A R D は、H M D によって塞がれた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、それは、テレプレゼンスセッション中、第 1 の発呼者に、第 2 の発呼者の塞がれていない顔が見えるように、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 のユーザに関連付けられた世界マップ情報が、ワードフロー注釈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得る。これは、A R D 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

## 【 0 1 4 4 】

テレプレゼンス用途では、周囲音の捕捉は、対面シナリオにおけるユーザ、すなわち、聴取者に関連付けられたデバイスではなく、ユーザ、すなわち、話者に関連付けられたデバイスによって実施される。発話の存在の検出およびテキストへの発話の変換は、いずれかのユーザに関連付けられたデバイスによって実施されることができる。発話源は、音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 のデバイスが音を捕捉するとき、ユーザ A が話している。

## 【 0 1 4 5 】

図 1 3 B は、テレプレゼンスセッションにおけるワードフロー注釈システムの別の例示的用途を示す。この例では、女性 1 3 2 6 は、ユーザ 1 3 2 8 および 1 3 2 0 と会議を行っているマーケティンググループ内に存在する。ユーザ 1 3 2 8 および 1 3 2 0 は、そのそれぞれの A R D 1 3 2 2 b および 1 3 2 2 c を装着している。A R D 1 3 2 2 b および 1 3 2 2 c は、図 2 に示される A R システムの 2 0 0 の少なくとも一部を含むことができる。

40

## 【 0 1 4 6 】

女性 1 3 2 6 が、( 時として、発話の吹き出しとも称されるコールアウトボックス 1 3 1 2 に示されるように ) 「 B r u g e s において p h o r o p t e r s を販売するための広告キャンペーンを議論しましょう」と述べている。A R D 1 3 2 2 b および 1 3 2 2 c は、このオーディオを捕捉し、オーディオ内のキーワードを識別することができる ( 例え

50

ば、ワードフロー注釈システム 970 を介して)。いくつかの実施形態では、女性は、単独で、または組み合わせて、女性のオーディオデバイスまたは ARD が、オーディオを捕捉し、オーディオストリームを処理し得る（例えば、それをテキストに変換する、またはキーワードを識別する）ように、オーディオデバイス（例えば、マイクロホン）を自分の近傍に有するか、または ARD を装着し得る。いくつかの実施形態では、オーディオファイルをテキストに変換すること、またはキーワードを識別すること等のオーディオ処理の少なくとも一部は、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、それは、各ユーザの ARD による同一オーディオストリームの重複処理を回避することができる。

#### 【0147】

ARD 1322b および 1322c は、女性の発話に関連付けられた補助情報を表示することができる。この例では、単語「phoropter」（眼科専門家によって眼鏡の処方箋を決定するために使用される医療デバイス）が、ユーザ 1320 に対して稀な単語であると決定され、ARD 1322c は、単語「phoropter」の定義をボックス 1316 内に表示する。単語「phoropter」が、ARD のユーザに対して稀な単語ではない場合、ARD は、ユーザのために、ボックス 1316 を表示しないこともある。この例では、単語「phoropter」は、ユーザが眼科部署におけるエンジニアであるので、ユーザ 1328 に対して稀な単語ではない。その結果、ARD 1322b は、補助情報（ボックス 1316 に示される）を装着者 1328 に提示しないであろう。この例を継続すると、ARD 1322b は、単語「Bruges」（ベルギー内の都市）がユーザ 1328 に対して稀な単語であることを決定することができ、都市についての情報および都市の場所の小地図提供するグラフィック 1324 を表示することができる。

#### 【0148】

##### （言語翻訳）

ワードフロー注釈システムは、言語翻訳のために使用されることができる。図 13C は、ワードフロー注釈システムの別の例示的用途を示し、ここでは、ワードフロー注釈システムは、第 1 の言語（例えば、スペイン語）における語句を第 2 の言語（例えば、英語を話すユーザのための英語）に翻訳することができる。この例では、AR システム 200 の英語を話すユーザ（図示せず）は、外国を訪問中であり、話者 1390 が、破線ボックス 1332 に示されるような「El supermercado esta al otro lado de la calle」と言うことを聞いている。AR システム 200 は、必要ではないが、話者の発話をユーザに表示し得る（例えば、ボックス 1332 は、典型的には、AR システムのユーザに表示されないが、いくつかの実装では、表示されることができる）。AR システム 200 は、話者 1390 の発話がスペイン語であることを決定し、発話を英語（「雑貨屋は、通りの向こう側にあります」）に翻訳し、キャプション 1336（または発話の吹き出し）を翻訳されたテキストとともに表示する。

#### 【0149】

AR システムは、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、または遠隔データリポジトリ 280 を使用して、1 つの言語から別の言語への翻訳をサポートすることができる。例えば、AR システムは、ワードフロー注釈システム 970 を実装し得、それは、ユーザの主要言語以外の言語をユーザの主要言語に翻訳することができる。ユーザの主要言語は、ユーザの母国語または会話中のユーザの好ましい言語であり得る。ユーザの主要言語以外の言語は、外国語と見なされ得る。この構成では、ユーザは、図 13C に図示されるように、コンテキスト補助情報ではなく、入力単語の翻訳を見ることを選択することができる。ユーザは、ユーザの主要言語、話者の言語、両言語（例えば、行間注釈フォーマットで）における検出された発話のトランスクリプト、言語を混合したトランスクリプト（例えば、話者の言語における一般的単語、話者の言語における一般的ではない単語、または両言語）等を見ることを選択することができる。システムは、一般的外国語単語辞書のカスタマイズをサポートすることができ、例えば、ユーザの母国語の一般的単語辞書より小さいサイズを可能にする。



例えば、ユーザは、特定の言語で最も使用される 1,000 または 2,500 個（もしくはある他の数）ほど一般的ではない任意の単語の翻訳を見ることを選択し得る。同一ユーザは、ユーザ自身の言語の一般的単語辞書が 7,000 または 8,000 個（もしくはある他の数）の単語のサイズを有することを選択し得る。

#### 【0150】

ワードフロー注釈システムは、ユーザが外国語における発話を理解することに役立ち得るだけでなく、ユーザが外国語を話すことにも役立ち得る。例えば、システムは、ユーザ自身の発話を外国語に翻訳するように構成されることができる。この翻訳は、ユーザの発話全体（例えば、全文）またはユーザの発話中の選択された単語に適用されることができる。例えば、ユーザは、「What is the word for 'patent' ? 」と言うことによって（または別の UI 相互作用を通して）、翻訳を提供するようにシステムに促し得る。システムは、着目外国語における「patent」の単語を、例えば、ディスプレイ 220 上に提供することによって応答することができる。ユーザは、外国語における単語を見、単語を自身で話すことができる。ワードフロー注釈システムは、オーディオ増幅器（例えば、スピーカ 240）を含み、翻訳された単語をオーディオで提供することができる。一実装では、ユーザのみが、例えば、イヤホンの形態におけるオーディオ増幅器を通して、このオーディオを聞き得る。別の実装では、ユーザなおよび会話相手が、ラウドスピーカを通して、このオーディオを聞き得る。

#### 【0151】

（コンテキスト情報に基づく補助情報の提示の例）

図 13D および 13E は、コンテキスト情報に基づく補助情報の提示の例を図示する。図 13A に説明されるように、ARD を装着しているユーザ（図示せず）は、女性 1302 を知覚し得る。女性 1302 は、ユーザの物理的環境内に存在し得る。ユーザは、女性 1302 と企業関連の会話中であり得る。図 13D におけるテキストボックス 1342 に示されるように、女性 1302 は、「Acme Corp. の幹部に何が起こるだろう」と言う。ユーザの ARD は、女性 1302 によって話されたこの文を捕捉することができる。ワードフロー注釈システム 970 は、女性 1302 の発話を分析し、女性 1302 によって話される文中のキーワードを検出することができる。この例では、ユーザが企業関連の会話中であるので、ARD は、それが企業関連であるので、「Acme Corp. 's executives」をキーワードとして識別することができる。ARD が、キーワードを決定すると、ARD は、キーワードに関連付けられた補助情報にアクセスし、それを読み出すことができる。例えば、ARD は、データリポジトリ（例えば、遠隔データリポジトリ 280）内で検索を実施すること、または公共のデータベース検索を実施することができる。テキストボックス 1340 に示されるように、ARD は、Acme Corp の重役の名前を読み出している。ARD は、ボックス 1340 に描写されるように、名前を補助情報として表示することができる。

#### 【0152】

図 13E では、女性 1302 は、「Acme Corp の株価の動きが激しい！」と言う。ユーザが企業関連の会話中であるので、ARD は、女性 1302 の発話中に述べられた企業上の情報をキーワードとして識別することができる。この例では、ARD は、語句「Acme Corp. 's stock」が、単語「tremendously」ではなく、この語句が企業関連であるので、キーワードであることを識別する。故に、ARD は、キーワードに関連付けられた補助情報を検索すること、またはそれにアクセスすることができる。テキストボックス 1354 に示されるように、ARD は、現在の株価、増加のパーセンテージ、変動範囲、および前日の終値をユーザに仮想コンテンツとして提示することができる。

#### 【0153】

図 13F は、コンテキスト情報に基づいて、着目オブジェクトに関連付けられた補助情報を提示する例を図示する。ユーザ（図示せず）は、ARD を装着しながら、野球の試合を観戦することができる。ARD は、シースルーディスプレイを含むことができるので、

ユーザは、野球の試合をライブで観戦すること、またはA R Dの外部のディスプレイ（例えば、コンピュータ画面もしくはテレビ等）上で再生することができる。野球の試合も、いくつかの状況では、A R Dによってレンダリングされ得る。ユーザが、野球の試合を観戦している間、ユーザは、選手P a u l 1 3 6 2が、打席に向かっており、コメンテータが、選手P a u l 1 3 6 2について話していることを知覚し得る。故に、A R Dは、選手P a u l 1 3 6 2が着目オブジェクトであることを決定することができる。この例では、A R Dは、単独で、または組み合わせて、種々の技法を使用して、着目オブジェクトを決定することができる。例えば、A R Dは、コメンテータの発話中で選手P a u l 1 3 6 2について頻繁に述べられていることを検出し、選手P a u l 1 3 6 2が着目オブジェクトであることの一因であることを決定することができる。別の例として、A R Dは、ゲーム進行を監視することによって、打席に向かっていない選手を追跡し、打席に向かっていない選手を着目オブジェクトとして識別することができる。さらに別の例として、ユーザが、野球の試合をテレビで観戦している間、A R Dは、外向きに面した結像システム4 6 4を使用して、テレビ上に表示されるような野球の試合の画像を入手することができる。A R Dは、1つ以上のオブジェクト認識装置7 0 8を使用して、画像を解析し、入手された画像から、打席に向かっていない選手を識別することができる。

10

#### 【0 1 5 4】

A R Dが、選手P a u l 1 3 6 2が着目オブジェクトであることを決定すると、A R Dは、選手P a u l 1 3 6 2に関連する情報を読み出し、ユーザに提示することができる。図1 3 Fのボックス1 3 6 4に示されるように、A R Dは、ゲーム統計（例えば、出場試合数および平均打率）、守備位置、および選手P a u l 1 3 6 2が属するチーム等の選手P a u l 1 3 6 2の野球関連情報を提示することができる。

20

#### 【0 1 5 5】

補助情報は、例えば、統計、ステータス、説明、製品情報、着目オブジェクトに関連付けられた追加の詳細／事実等、種々の情報を含むことができる。図1 3 Gは、補助情報が製品情報を含む例を図示する。図1 3 Fに説明されるように、A R Dは、選手P a u l 1 3 6 2が着目オブジェクトであることを決定することができる。選手P a u l 1 3 6 2の統計を表示する（図1 3 Gに示されるように）ことに加え、またはその代替として、A R Dは、選手P a u l 1 3 6 2（またはそのチーム）が公認している製品（例えば、Tシャツ1 3 4 0）を決定し、製品情報（Tシャツ1 3 4 0の画像、Tシャツ1 3 4 0の価格1 3 4 2、およびTシャツ1 3 4 0が購入され得るウェブサイト1 3 4 4等）を注釈ボックス1 3 6 8内に表示することができる。レンダリングすべき製品を決定し、製品情報にアクセスするために、A R Dは、公共データベースまたはA R Dに関連付けられたデータリポジトリの検索を実施し、選手P a u l 1 3 6 2が現れる広告を見出すことができる。A R Dは、故に、広告内で識別された製品を選手P a u l 1 3 6 2が公認している製品として決定することができる。いくつかの状況では、選手P a u l 1 3 6 2は、2つ以上の製品を公認し得る。A R Dは、製品のサブセットを取り上げ、補助情報として提示し得る。例えば、A R Dは、選手P a u l 1 3 6 2が過去1年に公認した製品を取り上げ得る。A R Dは、コンテキスト情報に基づいて、製品を取り上げることもできる。例えば、ユーザは、スポーツイベントを観戦しているので、A R Dは、選手P a u l 1 3 6 2によって公認されているスポーツ商品を決定し、これらのスポーツ商品の情報をユーザに提示し得る。A R Dは、購入決定をユーザから受け取るように構成され得る。例えば、ユーザが、Tシャツ1 3 4 0の購入を望む場合、ユーザは、例えば、「Tシャツを購入する」と言うか、またはユーザ入力デバイスを作動させ、ウェブサイト1 3 4 4へのリンクを選択し得、A R Dは、ユーザの購入選択を受け取り、トランザクションを完了するステップを行うであろう（例えば、ウェブサイト1 3 4 4にアクセスし、自動的に、Tシャツ1 3 4 0の購入を注文することによって）。

30

40

#### 【0 1 5 6】

（発話に注釈を付けるための例示的プロセス）

図1 4 Aおよび1 4 Bは、発話に注釈を付ける例示的プロセスを示す。処理フロー1 4

50

00は、図2に示されるようなARシステム200によって実施されることができる。

【0157】

ブロック1404では、システムのオーディオセンサは、周囲音を捕捉することができる。周囲音は、ユーザの環境内の1人以上の人物の発話を含むことができる。発話は、ユーザの発話、人物（ユーザ以外）の発話、または視聴覚コンテンツ（例えば、映画、音楽等）からの発話であり得る。周囲音は、雑音等の背景音も含み得る。

【0158】

ブロック1408では、システムは、周囲音を監視し、発話の存在を検出する。この検出処理は、ローカル（例えば、ローカル処理モジュール260によって）または遠隔で（例えば、遠隔処理モジュール270によって）、行われ得る。システムは、捕捉された周囲音内の単語のストリームを識別することによって、発話の存在を検出することができる。システムは、ユーザの姿勢または発話の存在を示すユーザ入力デバイス466からの入力等のユーザ入力にตอบสนองして、検出を開始または実施することもできる。

【0159】

ブロック1412では、システムは、発話を検出されたかどうかを決定する。発話を検出される場合、フローは、ブロック1416に継続する。発話を検出されない場合、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に戻る。

【0160】

ボックス1404から1412内の動作は、システムおよびオーディオセンサがオンにされると、例えば、ワードフロー注釈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）、実施され得る。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、定時割り込み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る。それらは、例証目的のために、処理フローシーケンス内で個別的なボックスとして示される。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。上で説明される例以外の多くの処理フローも、システム設計者の裁量で可能である。

【0161】

ブロック1416では、ARシステムは、検出された発話の源を決定することができる。発話源は、例えば、システムが、ユーザ自身の発話を処理しないように構成されるかどうか、またはシステムが、ARシステムのFOV内の人物からの発話のみを処理するように構成されるかどうか（FOV外の人物から生じる発話が、破棄され、さらに処理され得ない）に関連し得る。例えば、ARシステムは、ユーザの正面の人々からの発話を処理し、ユーザの背後の人々からの発話を無視し得る。発話源は、発話の補助情報またはテキストの表示場所を決定するために使用されることもできる。例えば、ARシステムは、発話に関連付けられた補助情報または発話のトランスクリプトを話している人物の近傍に表示し得る。

【0162】

ブロック1420では、システムは、発話源の変化が生じたかどうかを決定する。例えば、2人のユーザが、会話中であるとき、システムは、文が前の話者から生じているか、別の話者から生じているかを決定することができる。変化が生じた場合、フローは、ブロック1424からブロック1444に継続する。発話源の変化が生じていない場合、フローは、ブロック1428に進み、ワードフロー注釈処理を継続し、それは、周囲音を捕捉するステップ（ブロック1404）、発話の存在を検出するステップ（ブロック1408）、および図14Bに示される処理ステップを含むことができる。例えば、システムが、同一話者が話し続けていることを決定する場合、システムは、音の捕捉および発話の検出の継続に加え、ブロック1456から開始する機能を継続して実施することができる。

【0163】

ブロック1448では、例示的システムは、検出された発話がユーザ自身の発話であるかどうかを決定することができる。該当する場合、かつ、システムがユーザ自身の発話を処理しないように構成される場合、処理は、ブロック1484を通過してブロック1404

10

20

30

40

50

に戻り得る。検出された発話が、ユーザ自身の発話ではない場合、フローは、ブロック 1452 に継続する。ブロック 1452 では、システムは、発話がユーザの環境内の源から生じているかどうかを決定する。ユーザの環境は、例えば、内向きに面した結像システム 462、外向きに面した結像システム 464、または静止カメラによって捕捉された情報を通して決定されるようなユーザが見ているまたは会話している人物を含み得る。外向きに面した結像システム 464 は、ユーザに対する話者の位置を含む情報を提供することができる。内向きに面した結像システムは、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、ワードフロー注釈システムは、ユーザが見ている人物を決定することができる。

#### 【0164】

ユーザの環境は、別の例として、音圧閾値を上回る発話を含み得る。ワードフロー注釈システムは、構成可能音圧閾値、例えば、通常の会話に対して 60 ~ 70 dBA、囁きに対して 30 ~ 40 dBA をサポートすることができ、dBA は、異なる音周波数に対するヒトの耳の変動する感度を考慮するようにスケールされた A 加重デシベル測定値である。音圧閾値またはそれを上回る発話が、ユーザの環境内に存在し得、以下に説明されるように、ワードフロー注釈システムによって処理されることができる。音圧閾値を下回る発話は、ユーザの直近環境内に存在しないこともあり（例えば、背景会話または遠距離源からの発話）、処理されないこともある。

#### 【0165】

ユーザの環境は、第 3 の例として、特定の源からの発話を含み得る。例えば、ユーザは、オフィス内のスピーカーホンを通して、遠隔人物と会話を実行し得る。ユーザは、ユーザが、スピーカーホンを通して会話を実行している間、ドキュメントまたはコンピュータ画面を見ているとき等、スピーカーホンを見ていないこともある。AR システムは、ユーザが、ユーザの FOV がスピーカーホンを含まない場合でも、オーディオが処理されるであろうように、スピーカーホンをユーザの環境として指定することを可能にすることができる。別の例として、ワークフロー注釈システムの旅行者ユーザは、例えば、ツアーガイドを見ずに（または時折しか見ずに）、旅行者ユーザが観光地の景観を楽しんでいる間、ツアーガイドの発話に注釈を付けることを欲し得る。ワードフロー注釈システムは、旅行者ユーザが、ツアーガイドをユーザの環境として指定することを可能にすることができる。ツアーガイドの発話は、旅行者ユーザにとって外国語であり得る。そのような指定は、

#### 【0166】

ワードフロー注釈システムの実装は、ユーザの環境変化を促進することができる。スピーカーホン例では、システムは、会話の持続時間にわたって、スピーカーホンをユーザの環境として指定するように構成され、会話が終了した後、音圧閾値に基づいて、ユーザの環境を決定するように再構成されることができる。

#### 【0167】

別の例として、ユーザの環境が、音圧閾値に基づいて決定され、ユーザの FOV 内の発話源に限定される場合、実装は、ユーザの FOV 外からの閾値を上回る捕捉された発話の一部をバッファリングすることができる。このバッファリングされた情報は、例えば、ユーザが方向転換し、ユーザの FOV 外にあった捕捉された発話源を見た場合、読み出され、処理されることができる。

#### 【0168】

ユーザの環境内の発話に対して、ブロック 1456 では、システムは、上で説明されるように、発話をテキストに変換することができる。

#### 【0169】

ブロック 1460 では、システムは、稀な単語検出を変換されたテキストに実施することができる。ブロック 1464 では、システムは、稀な単語が検出されたかどうかを決定する。これらのボックス 1460 および 1464 内で実施される機能は、前節に詳細に説明されている。

10

20

30

40

50

## 【 0 1 7 0 】

稀な単語が検出されない場合、フローは、ブロック 1 4 8 8 に進み、上記のブロック 1 4 2 8 の説明に類似する様式において、ワードフロー注釈処理を継続する。

## 【 0 1 7 1 】

稀な単語が、ブロック 1 4 6 8 において検出される場合、システムは、稀な単語に関連付けられた補助情報を読み出すことができる。ブロック 1 4 7 2 では、システムは、読み出された補助情報をシステムの A R ディスプレイ上に表示することができる。ブロック 1 4 7 6 では、システムは、補助情報を閉じ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 1 4 8 0 では、システムは、補助情報の表示を閉じ、ブロック 1 4 8 8 に継続することができる。これらのボックス内で実施される機能は、上記および図 1 3 A - 1 3 G を参照して説明される。

10

## 【 0 1 7 2 】

ボックス 1 4 0 4 から 1 4 1 2 に関して上で記載されるものと同様に、フローチャート 1 4 4 0 内の動作は、他のボックスフローチャート 1 4 0 0 および 1 4 4 0 と並行して実施され得る。それらは、例証目的のために、処理フローシーケンス内で個別的な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、稀な単語のための補助情報を表示し得る（ブロック 1 4 7 2 において）一方、システムは、追加の発話をテキストに変換する（ブロック 1 4 5 6 において）か、稀な単語検出を追加の発話に対して実施する（ブロック 1 4 6 0 において）か、または別の稀な単語のための補助情報を読み出す（ブロック 1 4 6 8 において）。別の例として、システムは、発話をテキストに変換することができる（ブロック 1 4 5 6 において）一方、以前に検出された稀な単語のための補助情報を読み出す（ブロック 1 4 6 8 において）。多くの他の処理フローも、可システム設計者の裁量に応じて可能である。

20

## 【 0 1 7 3 】

（着目オブジェクトに注釈を付ける例示的プロセス）

図 1 5 は、着目オブジェクトおよび着目オブジェクトの補助情報との相互作用の例示的プロセスを図示する。例示的プロセス 1 5 0 0 は、本明細書に説明される A R システムによって実施されることができる。

## 【 0 1 7 4 】

ブロック 1 5 1 0 では、A R システムは、ユーザの環境内の着目オブジェクトを識別することができる。着目オブジェクトは、人物、オブジェクト（物理的もしくは仮想）、または周囲音等を含むことができる。A R システムは、外向きに面した結像システム 4 6 4、内向きに面した結像システム 4 6 2、または静止カメラから入手された画像を使用して、着目オブジェクトを識別することができる。A R システムは、マイクロホン 2 3 2 から入手されたオーディオ情報、または別のデバイス（例えば、別のユーザの A R D、環境内のマイクロホン）等から受信されたオーディオ情報を分析することによって、着目オブジェクトを識別することができる。図 1 3 F を参照して説明されるように、オーディオ情報は、画像と組み合わせて使用され、着目オブジェクトを決定することができる。いくつかの実装では、A R システムは、コンテキスト情報に基づいて、着目オブジェクトを識別することができる。本明細書に説明されるように、コンテキスト情報は、ユーザ、着目オブジェクト、ユーザの環境、ユーザの環境内のオブジェクトまたは人物等に関連付けられた情報を含み得る。別の例として、ユーザの環境内のオブジェクトは、ユーザの経験により、ユーザに対して稀であり得る（したがって、着目オブジェクトであり得る）が、別のユーザに対して稀ではないこともある。組織用鉗子は、あるユーザに対して、多くの場合、それらが日常の生活において見られないので、ユーザに対して稀なものであり得るが、それらは、外科医に対して稀なではないこともある。

30

40

## 【 0 1 7 5 】

ブロック 1 5 2 0 では、A R システムは、コンテキスト情報に基づいて、着目オブジェクトに関連付けられた補助情報を決定することができる。A R システムは、例えば、遠隔データリポジトリ 2 8 0、A R デバイスにローカルのデータリポジトリ、または別のコン

50

ピューティングデバイスによって管理されるデータ記憶装置等の1つ以上のデータリポジトリからの補助情報にアクセスすることができる。ARシステムは、検索エンジンを用いた検索を実施することによって、補助情報を決定することもできる。ARシステムは、コンテキスト情報に基づいて、補助情報の一部のみを提示することができる。例えば、ユーザがスポーツイベントを観戦しているとき、選手の補助情報は、他のタイプの情報（例えば、選手の家族に関する情報）ではなく、選手のスポーツ関連情報（例えば、試合統計）を含み得る。ARシステムは、コンテキスト情報に基づいて、補助情報にアクセスすること、またはそれを検索することもできる。例えば、ユーザが、株式取引に関連する会話中である場合、ユーザに語句「Shop Corp.」が聞こえるとき、ARシステムは、「Shop Corp.」の株式情報にアクセスすることができる。しかし、ユーザが、買物について会話中である場合、ARが語句「Shop Corp.」を検出するとき、ARシステムは、Shop Corp.によって販売されている製品にアクセスすることができる。

10

#### 【0176】

ブロック1530では、ARシステムは、補助情報をユーザに提示することができる。補助情報は、ユーザの物理的環境上にオーバーレイされた仮想コンテンツとして提示されることができる。例えば、ユーザが、ユーザの環境内に物理的に存在する別のユーザと会話中である間、ARシステムが他のユーザによって話される稀な単語を検出すると、ARシステムは、発話の吹き出しを他のユーザのすぐ隣に示すことができる。会話がテレプレゼンスセッションにおいて生じる場合、ARシステムは、補助情報を他のユーザの仮想アバターの近傍に提示することができる。

20

#### 【0177】

ブロック1540では、ARシステムは、随意に、補助情報とのユーザ相互作用を検出することができる。例えば、ユーザは、メモ/図を補助情報に追加することができる。ユーザは、補助情報を閉じること、追加の補助情報を要求すること（例えば、手ジェスチャを使用して、補助情報を示すユーザインターフェース要素を拡張させる）、または補助情報をマーク/編集することができる。

#### 【0178】

ブロック1550では、ARシステムは、随意に、補助情報とのユーザの相互作用に基づいて、ユーザインターフェース動作を実施することができる。例えば、ARシステムは、補助情報を閉じるためのユーザ入力に応答して、補助情報を隠すことができる。ARシステムは、ブロック1540において、ユーザ相互作用に従って、ユーザのメモを表示すること、または補助情報の一部をハイライトすることもできる。

30

#### 【0179】

（スレッド識別の例）

補助情報を決定することに加え、またはその代替として、ARシステムは、キーワードを使用して、会話中のスレッドを識別することもできる。図16Aおよび16Bは、キーワードに基づくスレッド識別の例を図示する。これらの図では、2人のユーザ1328および1320は、彼らのそれぞれのARD1322bおよび1322cを装着することができる。2人のユーザ1328および1320は、対面会話中であり得るか、またはテレプレゼンスセッション中であり得る。2人のユーザ1328および1320はまた、テキストメッセージングを使用して（例えば、テキストメッセージングアプリケーションを介して等）、互いに通信することもできる。

40

#### 【0180】

発話の吹き出し1612、1614、1616、1618は、ユーザ1328および1320によって話された一連の文を図示する。これらの発話の吹き出しは、ユーザ1328と1320との間の会話のコンテンツを図示する。ARシステムは、これらの発話の吹き出しを視覚的コンテンツとしてユーザ1328および1320に提示しないこともある。発話吹き出し1612および1616内のコンテンツは、ユーザ1328によって話され、発話吹き出し1614および1618内のコンテンツは、ユーザ1320によって話

50

される。発話の吹き出し 1 6 1 2、1 6 1 4、1 6 1 6、および 1 6 1 8 に示されるように、2 人のユーザ間の会話は、2 つの話題を伴い：1 つは、ニューラルネットワークプログラミングに関連し、もう 1 つは、ユーザ 1 3 2 8 の今後のアフリカ旅行に関連する。しかしながら、2 つの話題は、両者の発話内にひとまとめにされている。

#### 【0 1 8 1】

前節に説明されるように、A R システム 2 0 0 は、キーワードに基づいて、2 つの別個の話題が会話中に存在することを識別することができる。この例では、A R システムは、1 つのキーワードを「ニューラルネットワーク」として、別のキーワードを「アフリカ」として識別することができる。キーワード「ニューラルネットワーク」は、アルゴリズムおよびコンピュータプログラミングに関連する一方、キーワード「アフリカ」は、地理的領域を説明するので、A R システム 2 0 0 は、これらの 2 つのキーワードが互いに関連せず、したがって、各キーワードが別個の話題に関連付けられ得ることを決定し得る。いくつかの状況では、A R システムが、複数のキーワードを識別している場合でも、キーワードのうちのいくつかは、同じ話題に関連付けられ得る。例えば、A R システムは、1 つのキーワード「vacation」および別のキーワード「beach」を検出し得る。A R システムは、それにもかかわらず、ユーザが浜辺での休暇について話していることもあるので、1 つの話題を識別し得る。

#### 【0 1 8 2】

複数の話題が、会話中で検出される場合、A R システムは、話題を分離し、異なる会話スレッドを作成する。図 1 6 B に示されるように、図 1 6 A におけるユーザ 1 3 2 8 と 1 3 2 0 との間の会話は、2 つのスレッド 1 6 3 0 および 1 6 4 0 に分離され、スレッド 1 6 3 0 は、ニューラルネットワーク会話に対応し、スレッド 1 6 4 0 は、今後のアフリカ旅行に対応する。発話の吹き出し 1 6 3 2 および 1 6 4 2 は、発話の吹き出し 1 6 1 2 に対応し、発話の吹き出し 1 6 3 4 および 1 6 4 4 は、発話の吹き出し 1 6 1 4 に対応し、発話の吹き出し 1 6 3 6 および 1 6 4 6 は、発話の吹き出し 1 6 1 6 に対応し、発話の吹き出し 1 6 3 8 および 1 6 4 8 は、発話の吹き出し 1 6 1 8 に対応する。

#### 【0 1 8 3】

A R システムは、会話を 2 つの別個のスレッドとしてユーザ（例えば、ユーザ 1 3 2 0）に提示することができる。例えば、スレッド 1 6 3 0 は、ユーザの F O V の左側に表示されることができる一方、スレッド 1 6 4 0 は、ユーザの F O V の右側に表示されることができる。テキスト、発話の吹き出し等のための異なる色またはグラフィックが、異なるスレッドを区別するために使用され得る。有利には、いくつかの実施形態では、会話を異なるスレッドに分離することによって、ユーザは、ユーザが A R D を装着している間、各話題のための会話のフローを知覚し、他のユーザとのユーザの相互作用中の混乱を低減させることが可能である。2 つのスレッドが、図 1 6 B に示されるが、それは、例証のためであって、スレッドの数は、3 つ以上であることができる。いくつかの実装では、A R システムが、例えば、ユーザの眼の視線方向を分析することを介して、ユーザ 1 3 2 0 が一方のスレッド（例えば、スレッド 1 6 3 0）を見ていることを決定する場合、A R システムは、例えば、スレッドの明度またはコントラストを低減させるか、もしくは他のスレッドを全く表示しないことによって、他方のスレッド（例えば、スレッド 1 6 4 0）の外観を強調解除し得る。

#### 【0 1 8 4】

いくつかの状況では、1 つ（または複数）のスレッドは、複数のサブスレッドを有し得、各サブスレッドは、スレッドから派生する話題を含み得る。例えば、ユーザ 1 3 2 8 が、自分の今後のアフリカ旅行について話している間、ユーザ 1 3 2 8 は、アフリカでサーフィンをすることを計画していると言い得る。ユーザ 1 3 2 0 は、サーフィンの話題を継続し、サーフィンの道具および場所について議論し得る。A R システムは、ユーザがサーフィンについて議論していることを検出すると、サーフィンのための完全に新しいスレッドを有するのではなく、アフリカにおけるサーフィンのためのサブスレッドを提供することができる。

10

20

30

40

50

## 【 0 1 8 5 】

図 1 3 A - 1 3 G および図 1 6 A - 1 6 B を参照して説明される例は、発話の吹き出しおよびテキストボックスを使用するが、これらの発話の吹き出しおよびテキストボックスは、1 つの可能なユーザインターフェース例である。A R システムは、種々の視覚的グラフィックを使用して、補助情報およびテキストスレッド、例えば、字幕、バナー等の提示を提供することができる。

## 【 0 1 8 6 】

(スレッド識別の例示的プロセス)

図 1 7 は、スレッド識別の例示的プロセスを図示する。プロセス 1 7 0 0 は、本明細書に説明される A R システム 2 0 0 の実施形態によって実施され得る。

10

## 【 0 1 8 7 】

ブロック 1 7 1 0 では、A R システムは、第 1 のオーディオストリーム（例えば、第 1 の人物に関連付けられた）および第 2 のオーディオストリーム（例えば、第 2 の人物に関連付けられた）を受信することができる。この例では、第 1 の人物および第 2 の人物は、互いに会話することができ、A R システムは、各人物によって話されるオーディオストリームを捕捉することができる。テレプレゼンスの別の例では、第 1 の人物は、遠隔場所に存在する第 2 の人物に話し掛けている。第 1 の人物の A R システム上のオーディオセンサは、第 1 の人物の発話を検出し、第 1 のオーディオストリームを生成することができ、第 2 の人物の A R システム（または電話）上のオーディオセンサは、第 2 の人物の発話を検出し、第 2 のオーディオストリームを（例えば、ネットワーク 1 2 9 0 を介して）第 1 の人物の A R システムに通信することができる。いくつかの実施形態では、第 1 の人物または第 2 の人物の両者は、A R システムのユーザであることができる。他の実施形態では、ユーザは、第 1 の人物と第 2 の人物との間の会話を観察し得る（例えば、第 1 の人物と第 2 の人物との間の討論を見ている）。A R システムは、オーディオセンサ 2 3 2 を使用して、またはネットワーク 1 2 9 0 を介してオーディオストリーム（もしくは書き起こされたオーディオストリーム）を受信することによって、第 1 のオーディオストリームまたは第 2 のオーディオストリームを入手することができる。

20

## 【 0 1 8 8 】

ブロック 1 7 2 0 では、A R システムは、第 1 のオーディオストリームおよび第 2 のオーディオストリームを解析することができる。第 1 の人物と第 2 の人物との間の会話は、複数の話題を伴い得るので、A R システムは、第 1 の話題に関連付けられた第 1 のキーワードと、第 2 の話題に関連付けられた第 2 のキーワードとを識別することができる。キーワードは、前節に説明される種々の技法を使用して識別されることができる。いくつかの実施形態では、話題は、2 つ以上のキーワードを含み得る。

30

## 【 0 1 8 9 】

ブロック 1 7 3 0 では、A R システムは、第 1 の話題に関連付けられた第 1 のスレッドと、第 2 の話題に関連付けられた第 2 のスレッドとを生成することができる。A R システムは、第 1 の話題に関連する会話の部分を第 1 のスレッドにまとめ、第 2 の話題に関連する会話の部分を第 2 のスレッドにまとめることができる。

## 【 0 1 9 0 】

ブロック 1 7 4 0 では、A R システムは、第 1 のスレッドまたは第 2 のスレッドのうちの少なくとも 1 つを A R D によってレンダリングさせることができる。図 1 6 B に図示されるように、A R D は、第 1 のスレッドおよび第 2 のスレッドをユーザの F O V 内の異なる表示場所にレンダリングすることができる。いくつかの実施形態では、A R システムは、2 つのスレッドのうちの 1 つのみを提示し得る。例えば、A R システムは、その関連付けられた話題が第 1 の人物および第 2 の人物によって現在議論されているスレッドを提示し得る。

40

## 【 0 1 9 1 】

(追加の側面)

第 1 の側面では、拡張現実システムを通して、稀な単語のための情報を提供する方法で

50



あって、方法は、オーディオセンサを備えている拡張現実（ＡＲ）システムの制御下、オーディオセンサを介して、周囲音を捕捉することと、周囲音内の発話の存在を検出することと、検出された発話をテキストに変換することと、変換されたテキスト内の稀な単語を検出することと、稀な単語に関連付けられた補助情報を読み出すことと、ＡＲシステムを使用して、補助情報を表示することと、補助情報の表示を閉じるための条件を検出することと、補助情報の表示を閉じることとを含む、方法。

【０１９２】

第２の側面では、検出された発話がＡＲシステムのユーザのユーザ環境内である場合、変換するステップからのステップが、実施される、側面１に記載の方法。

【０１９３】

第３の側面では、稀な単語は、一般的単語辞書から除外された単語である、側面１－３のいずれか１項に記載の方法。

【０１９４】

第４の側面では、一般的単語辞書は、稀な単語を含むように更新される、側面３に記載の方法。

【０１９５】

第５の側面では、稀な単語は、ＡＲシステムのユーザの着目言語における構成可能閾値より低頻度で生じる単語である、側面１－４のいずれか１項に記載の方法。

【０１９６】

第６の側面では、稀な単語は、ＡＲシステムのユーザに対して外国の言語における単語であり、補助情報は、ユーザの主要言語における稀な単語の翻訳を含む、側面１－５のいずれか１項に記載の方法。

【０１９７】

第７の側面では、検出された発話中の稀な単語のコンテキストを決定し、コンテキストに基づいて、稀な単語に関連付けられた補助情報を読み出すことをさらに含む、側面１－６のいずれか１項に記載の方法。

【０１９８】

第８の側面では、コンテキストを決定することは、少なくとも部分的に、検出された発話中の稀な単語に時間的に近接して現れる単語に基づく、側面７に記載の方法。

【０１９９】

第９の側面では、補助情報の表示を閉じるための条件は、ユーザインターフェース相互作用に基づく、側面１－８のいずれか１項に記載の方法。

【０２００】

第１０の側面では、ユーザインターフェース相互作用は、ＡＲシステムのユーザの眼移動に基づく、側面９に記載の方法。

【０２０１】

第１１の側面では、補助情報の表示を閉じるための条件は、持続時間に基づく、側面１－１０のいずれか１項に記載の方法。

【０２０２】

第１２の側面では、補助情報の表示を閉じるための条件は、少なくとも部分的に、別の稀な単語の検出に基づく、側面１－１１のいずれか１項に記載の方法。

【０２０３】

第１３の側面では、補助情報の閉じられた表示を再表示するための条件を検出することと、補助情報の閉じられた表示を再表示することとをさらに含む、側面１－１２のいずれか１項に記載の方法。

【０２０４】

第１４の側面では、ＡＲディスプレイと、オーディオセンサと、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た非一過性データ記憶装置と、データ記憶装置と通信するハードウェアプロセッサとを備えている稀な単語のための情報を提供するための拡張現実（ＡＲ）装置であって、コンピュータ実行可能命令は、実行されると、オー

10

20

30

40

50

ディオセンサによって検出されたオーディオ情報を受信することと、受信されたオーディオ情報内の発話を検出することと、検出された発話中の稀な単語を検出することと、稀な単語に基づいて、補助情報を読み出すことと、ARディスプレイを使用して、読み出された補助情報を表示することとをプロセッサに行わせる、装置。

【0205】

第15の側面では、通信ネットワークを経由して通信するための通信コンポーネントをさらに備え、補助情報は、通信コンポーネントを介して、通信ネットワークを通して読み出される、側面14に記載の装置。

【0206】

第16の側面では、プロセッサは、遠隔プロセッサが稀な単語を検出するために、検出された発話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、稀な単語を検出する、側面15に記載の装置。

10

【0207】

第17の側面では、オーディオセンサは、複数のマイクロホンまたは指向性マイクロホンを備えている、側面14 - 16のいずれか1項に記載の装置。

【0208】

第18の側面では、プロセッサは、検出された発話の源を決定することと、稀な単語を検出することと、検出された発話の源がAR装置のユーザ以外であることを決定すると、補助情報を読み出し、表示することとを行うようにさらに構成されている、側面14 - 17のいずれか1項に記載の装置。

20

【0209】

第19の側面では、ユーザインターフェース相互作用を受信するためのユーザ入力デバイスをさらに備えている、側面14 - 18のいずれか1項に記載の装置。

【0210】

第20の側面では、稀な単語に関する情報を提供するための拡張現実(AR)システムであって、システムは、第1のウェアラブルARデバイスを備え、第1のウェアラブルARデバイスは、ARディスプレイと、オーディオセンサと、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たデータ記憶装置と、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、オーディオセンサによって検出されたオーディオ情報を受信することと、受信されたオーディオ情報内の発話を検出することと、検出された発話中の稀な単語を検出することと、稀な単語に基づいて、補助情報を読み出すことと、ARディスプレイ上に読み出された補助情報を表示することとをプロセッサに行わせる、システム。

30

【0211】

第21の側面では、内向きに面した結像システムをさらに備えている、側面20に記載のシステム。

【0212】

第22の側面では、外向きに面した結像システムをさらに備えている、側面20 - 21のいずれか1項に記載のシステム。

【0213】

40

第23の側面では、プロセッサは、外向きに面した結像システムからの画像情報を使用して、第1のウェアラブルARデバイスに関連付けられたユーザの環境に対する発話源の場所を決定する、側面22に記載のシステム。

【0214】

第24の側面では、プロセッサは、外向きに面した結像システムからの画像情報を使用して、発話源の場所と、ARディスプレイを通して、源のビューを曖昧にせず、ARディスプレイ上の発話源に空間内で近接した補助情報の表示位置とを決定するようにさらに構成されている、側面22に記載のシステム。

【0215】

第25の側面では、ARディスプレイと、オーディオセンサと、通信ネットワークを経

50

由して通信するように構成された通信コンポーネントとを備えている第2のウェアラブルARデバイスをさらに備え、第1のウェアラブルARデバイスは、通信ネットワークを経由して通信するように構成された通信コンポーネントをさらに備え、第1のウェアラブルARデバイスのオーディオセンサは、オーディオ情報を検出し、読み出された補助情報は、第2のウェアラブルARデバイスのARディスプレイ上に表示される、側面20-24のいずれか1項に記載のシステム。

【0216】

第26の側面では、オーディオストリーム内のスレッドを識別する方法であって、方法は、第1のオーディオストリームおよび第2のオーディオストリームを受信することと、第1のオーディオストリームおよび第2のオーディオストリームを解析し、第1の話題に関連付けられた第1のキーワードおよび第2の話題に関連付けられた第2のオーディオストリームを識別することと、第1の話題に関連付けられた第1のスレッドおよび第2の話題に関連付けられた第2のスレッドを生成することと、第1のスレッドまたは第2のスレッドのうちの少なくとも1つが拡張現実デバイスによってレンダリングされるようにすることを含む、方法。

10

【0217】

第27の側面では、第1のオーディオストリームまたは第2のオーディオストリームは、人物または視聴覚コンテンツのうちの少なくとも1つから生じる、側面26に記載の方法。

【0218】

20

第28の側面では、第1のオーディオストリームは、第1の人物から生じ、第2のオーディオストリームは、第2の人物から生じる、側面26-27のいずれか1項に記載の方法。

【0219】

第29の側面では、第1の話題は、複数のサブ話題をさらに含む、側面26-28のいずれか1項に記載の方法。

【0220】

第30の側面では、ハードウェアプロセッサと、拡張現実ディスプレイとを備えている拡張現実システムであって、ハードウェアプロセッサは、側面26-29に記載のいずれか1項を実施するようにプログラムされている、拡張現実システム。

30

【0221】

第31の側面では、拡張現実(AR)システムであって、システムは、仮想コンテンツをARシステムのユーザに提示するように構成されたARディスプレイと、周囲音を捕捉するように構成されたオーディオセンサと、ARディスプレイおよびオーディオセンサと通信するハードウェアプロセッサとを備え、ハードウェアプロセッサは、オーディオセンサによって捕捉された周囲音を受信することと、周囲音内の発話の存在を検出することと、検出された発話をテキストに変換することと、発話中の稀な単語を検出することと、稀な単語に関連付けられた補助情報を読み出すことと、ARディスプレイに読み出された補助情報を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ている、ARシステム。

【0222】

40

第32の側面では、稀な単語は、一般的単語辞書から除外された単語である、側面31に記載のARシステム。

【0223】

第33の側面では、ハードウェアプロセッサは、補助情報の表示を閉じるための条件を検出することと、検出された条件に応答して、ARディスプレイに補助情報の表示を閉じさせることとを行うようにさらにプログラムされている、側面31-32のいずれか1項に記載のARシステム。

【0224】

第34の側面では、表示を閉じるための条件を検出するために、ハードウェアプロセッサは、閾値期間が経過したことを決定すること、別の稀な単語を検出すること、ユーザの

50

姿勢を検出すること、またはユーザ入力デバイスからの入力を受信することのうちの少なくとも1つを実施するようにプログラムされている、側面33に記載のARシステム。

【0225】

第35の側面では、ハードウェアプロセッサは、検出された発話の源を決定することと、稀な単語を検出することと、検出された発話の源がARディスプレイのユーザ以外の話者に関連付けられていることを決定すると、補助情報を読み出し、表示することとを行うようにさらにプログラムされている、側面31-34のいずれか1項に記載のARシステム。

【0226】

第36の側面では、検出された発話の源は、1つ以上の指向性マイクロホンを備えているオーディオセンサから収集されたオーディオ情報またはARシステムの外向きに面した結像システムによって入手された画像のうちの少なくとも1つに基づいて決定される、側面35に記載のARシステム。

10

【0227】

第37の側面では、検出された発話の源を決定するために、ハードウェアプロセッサは、音声認識を検出された発話に対して実施し、話者の識別を決定するようにプログラムされている、側面35-36のいずれか1項に記載のARシステム。

【0228】

第38の側面では、検出された発話の源は、ユーザの環境内の別のコンピュータ、話者に関連付けられた別のARデバイス、またはユーザの環境内の人物のうちの少なくとも1つを含む、側面35-37のいずれか1項に記載のARシステム。

20

【0229】

第39の側面では、ユーザが話者とテレプレゼンスセッション中であり、検出された発話の源が話者に関連付けられた他のユーザデバイスである場合、ハードウェアプロセッサは、ARディスプレイに、話者の仮想アバタをレンダリングさせ、補助情報を他のユーザの仮想アバタの近傍にレンダリングさせるようにさらにプログラムされている、側面38に記載のARシステム。

【0230】

第40の側面では、ARディスプレイに読み出された補助情報をレンダリングするために、ハードウェアプロセッサは、源の場所を決定することと、ARディスプレイを通して、源のビューを曖昧にすることなく、補助情報を源に近接した3D空間内の位置にレンダリングすることとを行うようにプログラムされている、側面35-39のいずれか1項に記載のARシステム。

30

【0231】

第41の側面では、補助情報は、稀な単語の説明を含む、側面31-40のいずれか1項に記載のARシステム。

【0232】

第42の側面では、ハードウェアプロセッサは、ユーザから、補助情報に注釈を付けるための指示を受信することと、補助情報の注釈の記憶を開始することと、別の稀な単語の存在を検出することとにตอบสนองして、ARディスプレイに注釈および補助情報をレンダリングさせることとを行うようにさらにプログラムされている、側面31-41のいずれか1項に記載のARシステム。

40

【0233】

第43の側面では、方法であって、仮想コンテンツをユーザの環境内に提示するように構成されたハードウェアプロセッサおよびARディスプレイを備えている拡張現実(AR)デバイスの制御下で、ARデバイスのユーザの環境を監視することと、ユーザまたは環境のうちの少なくとも1つに関連付けられたコンテキスト情報に基づいて、環境内の着目オブジェクトの存在を検出することと、着目オブジェクトのための補助情報を読み出すことと、ユーザの視野内の補助情報の表示位置を決定することとであって、視野は、ユーザが所与の時間に知覚可能な環境の一部を含む、ことと、ARディスプレイに補助情報を表示

50

位置にレンダリングさせることとを含む、方法。

【0234】

第44の側面では、着目オブジェクトは、キーワード、物理的オブジェクト、または人物を含む、側面43に記載の方法。

【0235】

第45の側面では、着目オブジェクトは、キーワードを含み、着目オブジェクトの存在を検出することは、環境内の周囲音から発話を検出することと、発話を解析し、稀な単語を識別することとを含む、側面44に記載の方法。

【0236】

第46の側面では、環境を監視することは、環境の周囲音を捕捉し、分析すること、または環境の画像を入手することのうちの少なくとも1つを含む、側面43 - 45のいずれか1項に記載の方法。

10

【0237】

第47の側面では、コンテキスト情報は、ユーザの過去の体験に関連付けられ、着目オブジェクトは、ユーザが稀にしか遭遇しないオブジェクトを含む、側面43 - 46のいずれか1項に記載の方法。

【0238】

第48の側面では、補助情報は、着目オブジェクトに関連付けられた製品情報、または着目オブジェクトの説明テキストのうちの少なくとも1つを含む、側面43 - 47のいずれか1項に記載の方法。

20

【0239】

第49の側面では、ユーザから、補助情報に注釈を付けるための指示を受信することと、補助情報の注釈の記憶を開始することと、着目オブジェクトの別の存在を検出することに応答して、ARディスプレイに注釈および補助情報をレンダリングさせることとさらに含む、側面43 - 48のいずれか1項に記載の方法。

【0240】

第50の側面では、補助情報の表示位置を決定することは、着目オブジェクトの場所を決定することと、ARディスプレイを通して、着目オブジェクトのビューを曖昧にすることなく、着目オブジェクトに近接する環境内の位置として表示位置を識別することとを含む、側面43 - 49のいずれか1項に記載の方法。

30

【0241】

(他の考慮点)

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プロセス、方法、およびアルゴリズムの各々は、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、および/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か、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、所与の機能に特定の回路によって実施され得る。

40

【0242】

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ので、(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

50

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 2 4 3 】

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)、読み取り専用メモリ (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等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読み取り可能な媒体上に記憶され得る。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読み取り可能な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、または複数の個別的な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か、またはコンピュータ読み取り可能な伝送媒体を介して通信され得る。

10

【 0 2 4 4 】

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、追加の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、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もできる。タスクまたはイベントが、開示される例示的实施形態に追加されるか、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、全ての实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。

20

30

【 0 2 4 5 】

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。

40

【 0 2 4 6 】

本開示のシステムおよび方法の各々は、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。上で説明される種々の特徴およびプロセスは、互いに独立して使用され得るか、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

50

、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。

【0247】

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、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もできる。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、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。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で説明され、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もしくは必須ではない。

【0248】

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.)」等等、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、および/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、および/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、および/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「～を備えている」、「～を含む」、「～を有する」等は、同義語であり、非限定的方式で包括的に使用され、追加の要素、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1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 사용되는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。

【0249】

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を対象と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。

【0250】

同様に、動作は、特定の順序で図面に描写され得るが、それは、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つ以上の追加の動作が、図示される動作のいずれかの前、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。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、上で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。

【図 1】

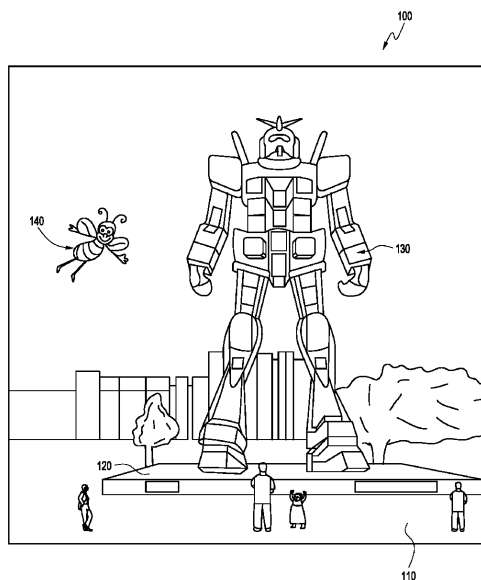


FIG. 1

【図 2】

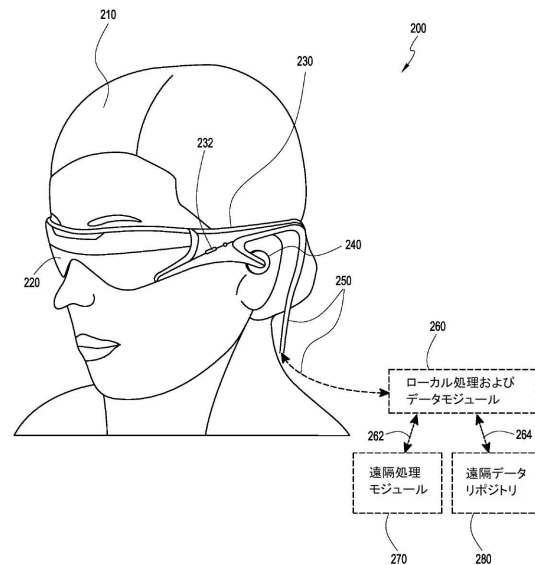


FIG. 2



【図 3】

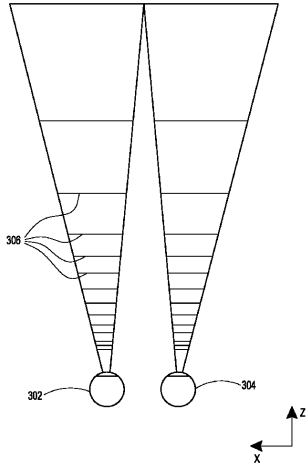


FIG. 3

【図 4】

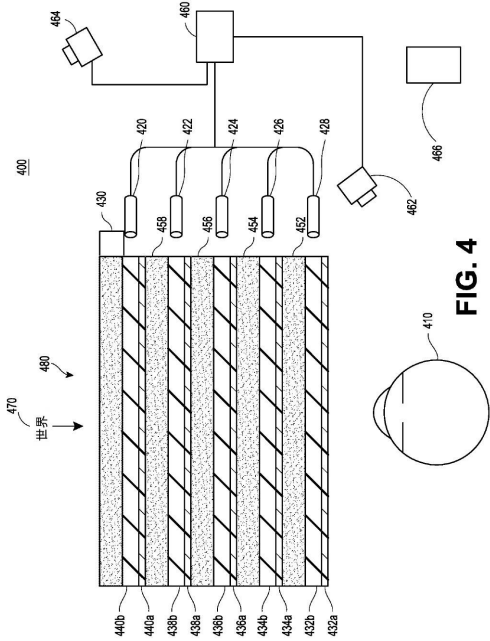


FIG. 4

【図 5】

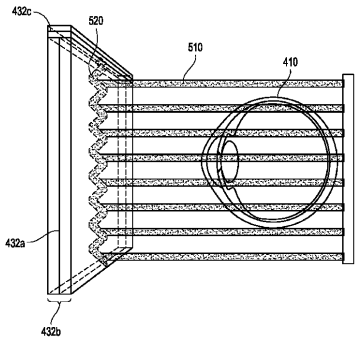


FIG. 5

【図 6】

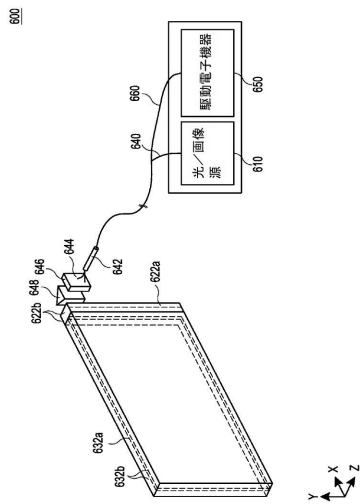


FIG. 6

【図 7】

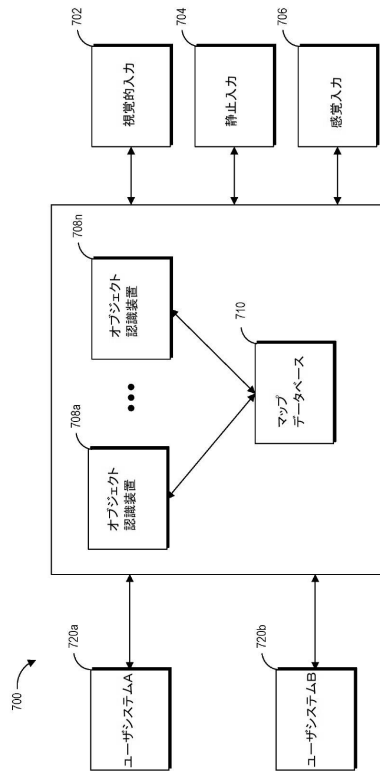


FIG. 7

【図 8】

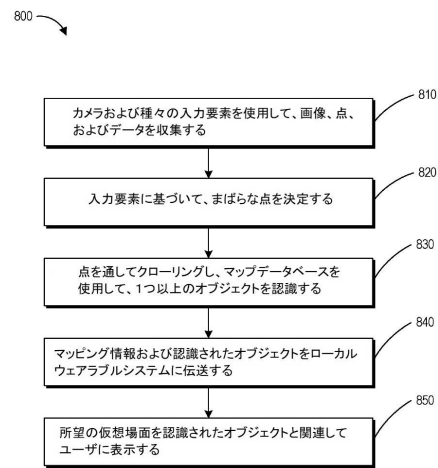


FIG. 8

【図 9】

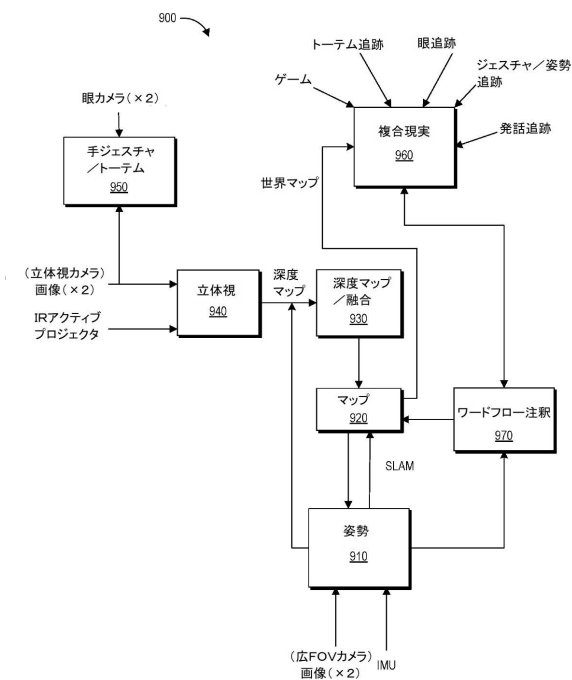


FIG. 9

【図 10】

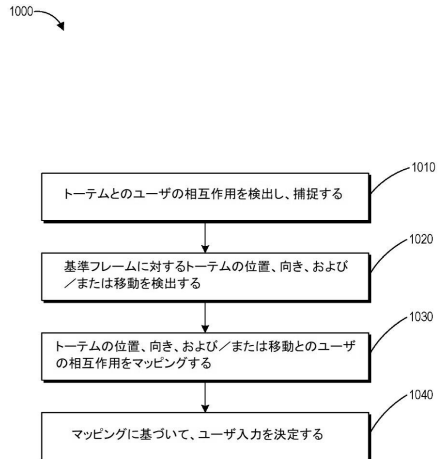


FIG. 10

【図 1 1】

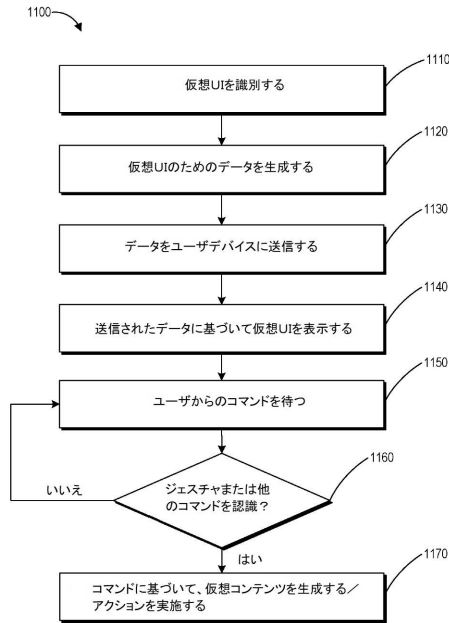


FIG. 11

【図 1 3 A】

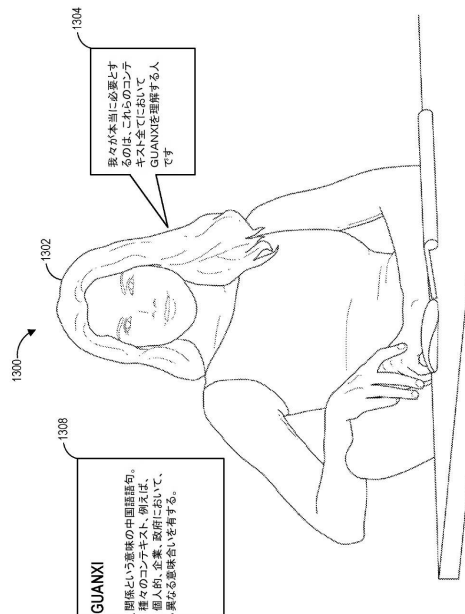


FIG. 13A

【図 1 2】

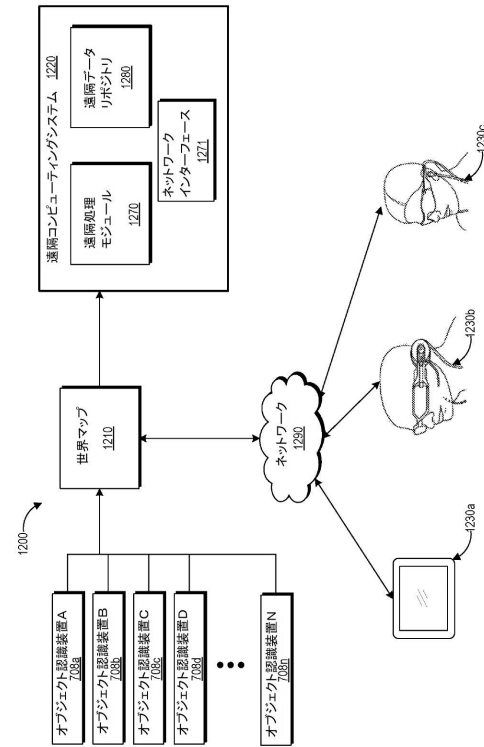


FIG. 12

【図 1 3 B】

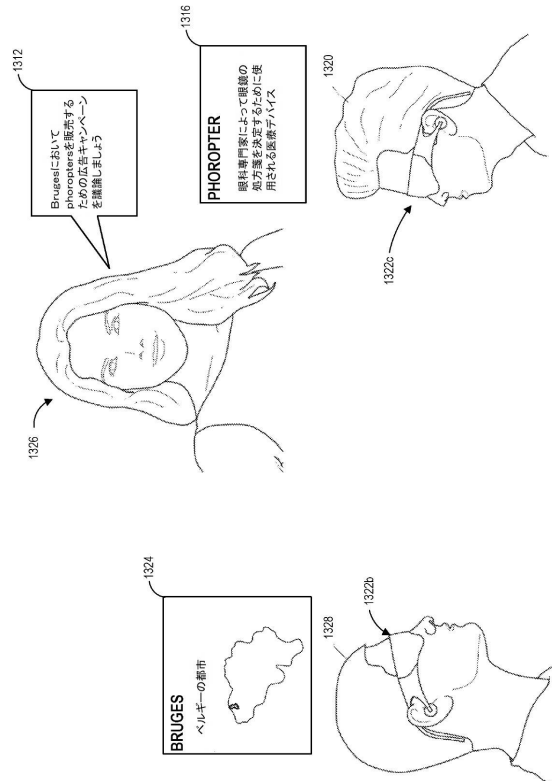


FIG. 13B

【図 13C】

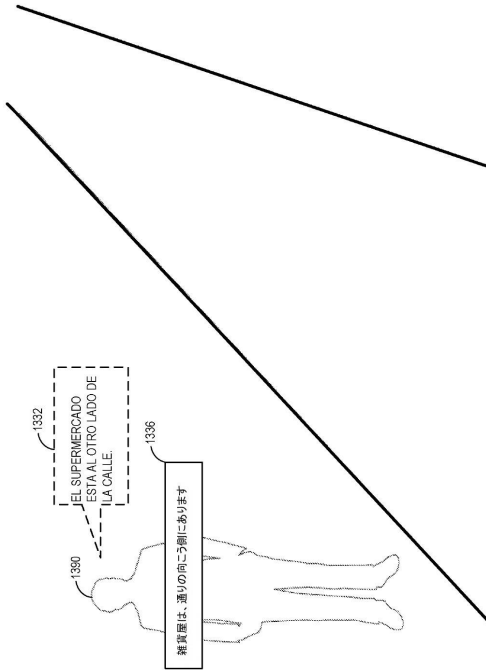


FIG. 13C

【図 13D】

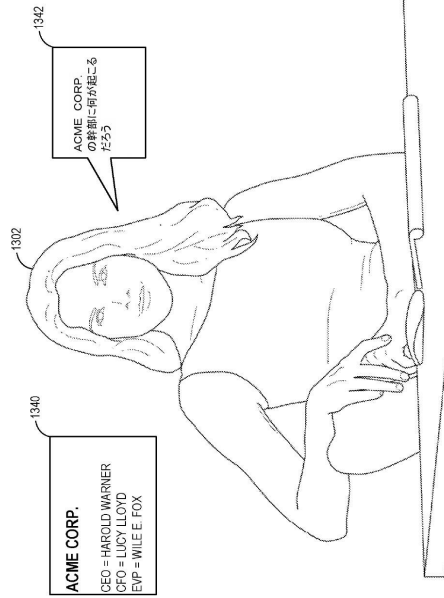


FIG. 13D

【図 13E】

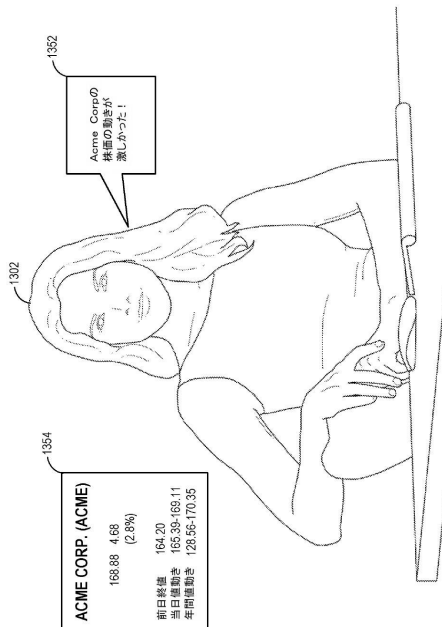


FIG. 13E

【図 13F】

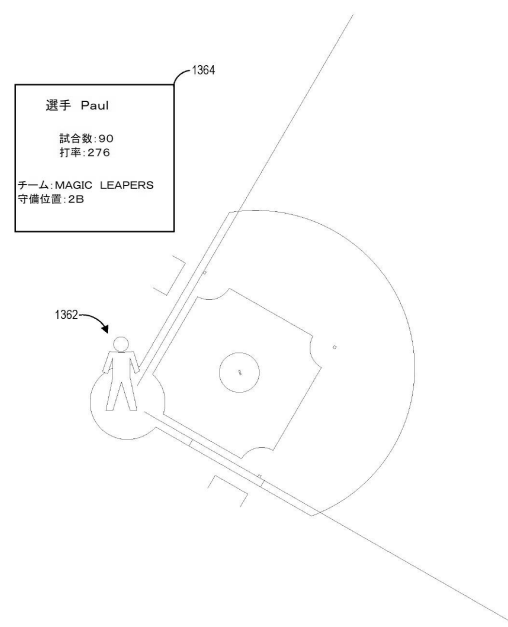


FIG. 13F

【図 13 G】

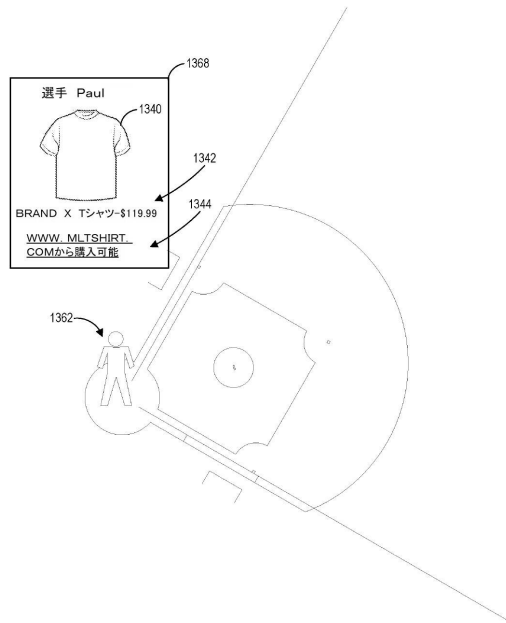


FIG. 13G

【図 14 A】

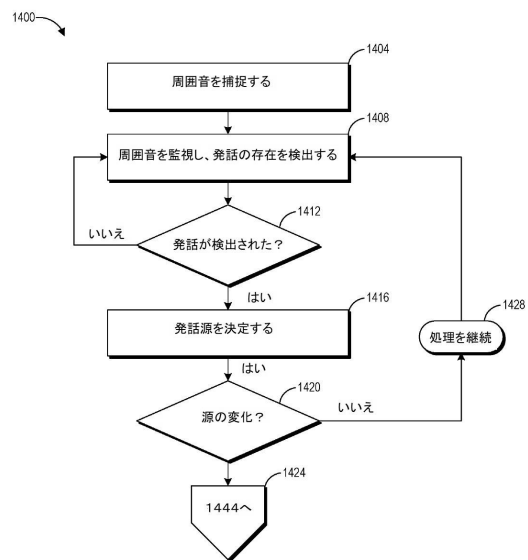


FIG. 14A

【図 14 B】

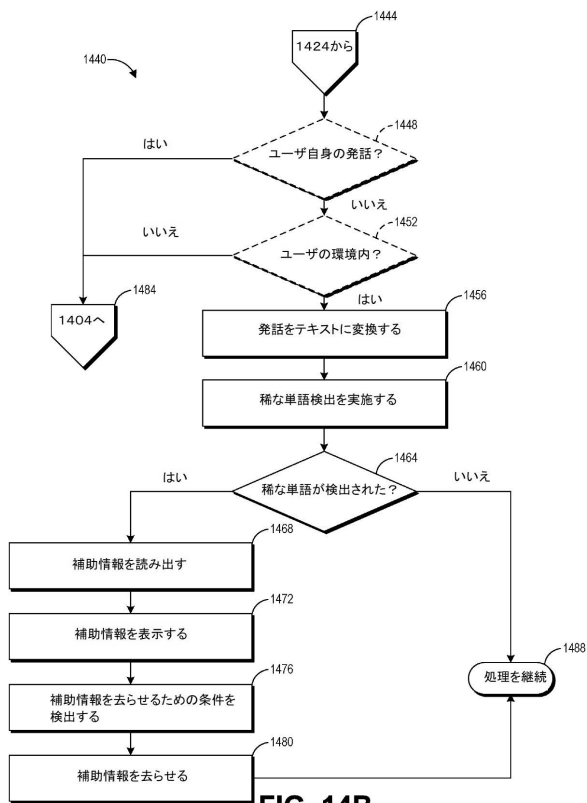


FIG. 14B

【図 15】

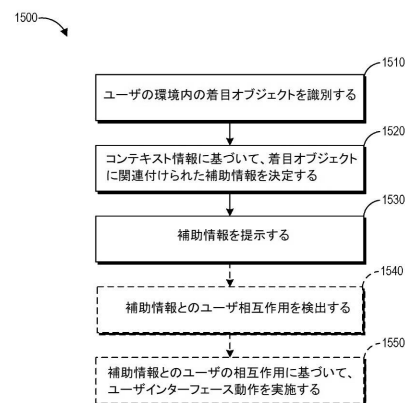


FIG. 15

【図 16A】

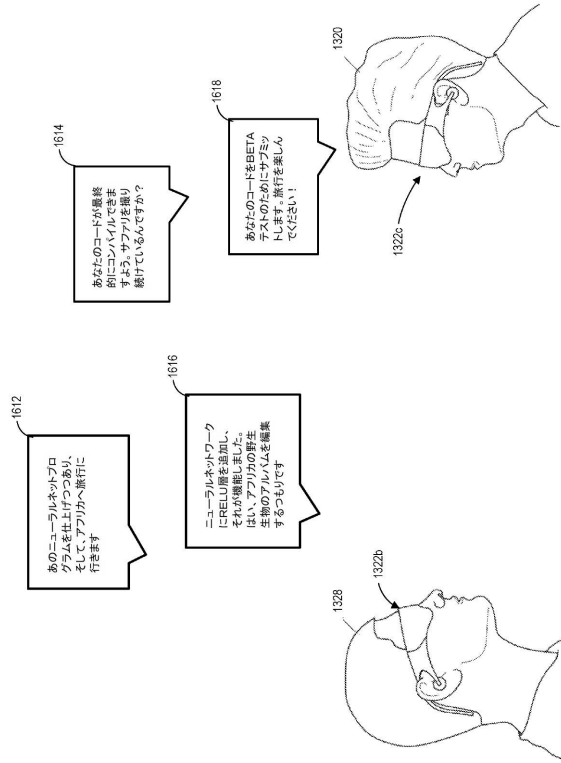


FIG. 16A

【図 16B】

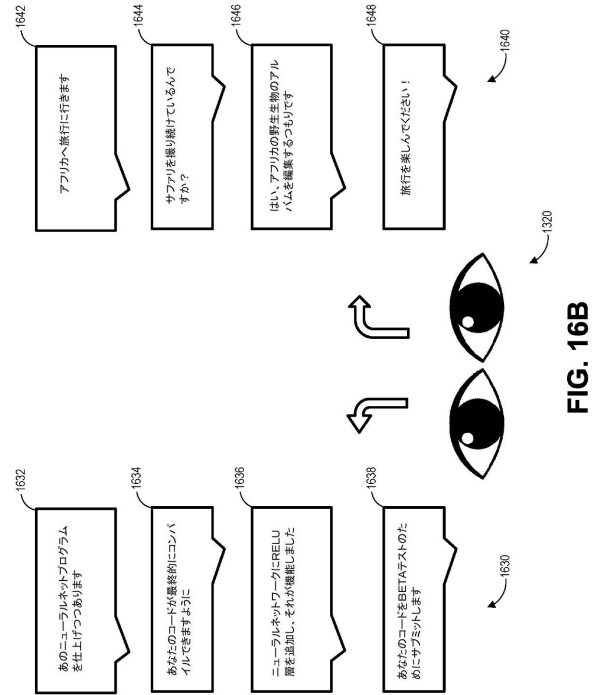


FIG. 16B

【図 17】

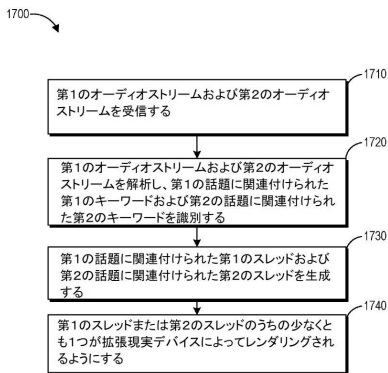


FIG. 17

## フロントページの続き

(51)Int.Cl.		F I			
<b>G 1 0 L</b>	<b>15/22</b>	<b>(2006.01)</b>	<b>G 1 0 L</b>	<b>15/22</b>	<b>4 5 3</b>
<b>G 1 0 L</b>	<b>15/10</b>	<b>(2006.01)</b>	<b>G 1 0 L</b>	<b>15/10</b>	<b>2 0 0 W</b>
<b>G 0 6 F</b>	<b>3/16</b>	<b>(2006.01)</b>	<b>G 0 6 F</b>	<b>3/16</b>	<b>6 5 0</b>
			<b>G 0 6 F</b>	<b>3/16</b>	<b>6 4 0</b>
			<b>G 0 6 F</b>	<b>3/16</b>	<b>6 2 0</b>
			<b>G 0 6 F</b>	<b>3/16</b>	<b>6 1 0</b>

(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 ソマーズ, ジェフリー

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  
ールバード 7 5 0 0

(72)発明者 ディバイン, ジェニファー エム.アール.

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  
ールバード 7 5 0 0

(72)発明者 スーク, ジョセフ ウェイン

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  
ールバード 7 5 0 0

(72)発明者 ケーラー, エイドリアン

アメリカ合衆国 カリフォルニア 9 0 0 2 7, ロサンゼルス, エヌ. ウェスタン アベニ  
ュー 1 9 4 0

審査官 原 秀人

(56)参考文献 特開2011-118531(JP,A)

特開2010-108099(JP,A)

特開2005-115289(JP,A)

特開2015-118710(JP,A)

特開2011-248860(JP,A)

米国特許出願公開第2014/0223279(US,A1)

米国特許出願公開第2010/0023476(US,A1)

(58)調査した分野(Int.Cl., DB名)

G 0 6 F 1 6 / 0 0 - 1 6 / 9 5 8

G 0 6 F 3 / 0 1

G 0 6 F 3 / 0 4 8 - 3 / 0 4 8 9

G 0 6 F 3 / 1 6

G 1 0 L 1 5 / 1 0

G 1 0 L 1 5 / 2 2