

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6860488号  
(P6860488)

(45) 発行日 令和3年4月14日(2021.4.14)

(24) 登録日 令和3年3月30日(2021.3.30)

(51) Int.Cl.

F 1

G06T 19/00	(2011.01)	G06T	19/00	600
G06F 3/01	(2006.01)	G06F	3/01	510
G06F 3/0481	(2013.01)	G06F	3/0481	150
H04N 5/222	(2006.01)	HO4N	5/222	400
H04N 5/232	(2006.01)	HO4N	5/232	290

請求項の数 19 (全 35 頁) 最終頁に続く

(21) 出願番号	特願2017-538330 (P2017-538330)
(86) (22) 出願日	平成28年1月13日 (2016.1.13)
(65) 公表番号	特表2018-511098 (P2018-511098A)
(43) 公表日	平成30年4月19日 (2018.4.19)
(86) 國際出願番号	PCT/US2016/013127
(87) 國際公開番号	W02016/118371
(87) 國際公開日	平成28年7月28日 (2016.7.28)
審査請求日	平成30年12月13日 (2018.12.13)
(31) 優先権主張番号	62/105,701
(32) 優先日	平成27年1月20日 (2015.1.20)
(33) 優先権主張国・地域又は機関	米国(US)
(31) 優先権主張番号	14/728,650
(32) 優先日	平成27年6月2日 (2015.6.2)
(33) 優先権主張国・地域又は機関	米国(US)

(73) 特許権者	314015767
	マイクロソフト テクノロジー ライセンシング、エルエルシー
	アメリカ合衆国 ワシントン州 98052 レッドモンド ワン マイクロソフト ウェイ
(74) 代理人	100079108
	弁理士 稲葉 良幸
(74) 代理人	100109346
	弁理士 大貫 敏史
(74) 代理人	100117189
	弁理士 江口 昭彦
(74) 代理人	100134120
	弁理士 内藤 和彦

最終頁に続く

(54) 【発明の名称】複合現実システム

## (57) 【特許請求の範囲】

## 【請求項 1】

実オブジェクトと仮想オブジェクトとを含む複合現実環境を提示するためのシステムであって、

仮想環境において3次元仮想オブジェクトを表示するためのディスプレイユニットおよび1組の1つまたは複数のカメラを含むヘッドマウントディスプレイデバイスと、

前記ヘッドマウントディスプレイデバイスとは異なる位置にある1つまたは複数のカメラアセンブリであって、前記複合現実環境において実オブジェクトのビデオを取り込むための1つまたは複数のカメラアセンブリと、

前記ヘッドマウントディスプレイデバイスと前記1つまたは複数のカメラアセンブリとに動作可能に結合された処理ユニットと

を備えており、前記処理ユニットのシーンマッピングモジュールは、前記ヘッドマウントディスプレイデバイスと、前記3次元仮想オブジェクトと、前記1つまたは複数のカメラアセンブリとがレジストレーションされる3次元座標空間を含むシーンマップを生成し、前記処理ユニットと、前記1つまたは複数のカメラアセンブリのうちのカメラアセンブリとは、前記カメラアセンブリの視点からの表示のために実オブジェクトのビデオおよび前記仮想オブジェクトを含む前記複合現実環境の画像を生成する、システム。

## 【請求項 2】

前記カメラアセンブリは、前記複合現実環境において実オブジェクトのカラー画像または白黒画像を取り込むための画像取り込みデバイスを含んでいる、請求項1に記載のシス

10

20

テム。

**【請求項 3】**

前記カメラアセンブリは、前記複合現実環境の深さデータを取り込むためのイメージセンサを含んでおり、前記深さデータを使用して、前記イメージセンサの視点から前記複合現実環境の深さマップを生成する、請求項 1 に記載のシステム。

**【請求項 4】**

前記カメラアセンブリと、前記ヘッドマウントディスプレイデバイスとは、前記カメラアセンブリによって共通座標系に対してキャリブレーションされ、前記ヘッドマウントディスプレイデバイスはキャリブレーションキューの 1 つまたは複数の画像を取り込む、請求項 1 に記載のシステム。

10

**【請求項 5】**

前記処理ユニットは、前記カメラアセンブリおよび前記ヘッドマウントディスプレイデバイスの視点からキャリブレーションターゲットの前記 1 つまたは複数の画像からの深さイメージングセンサデータに基づいて、前記カメラアセンブリおよび前記ヘッドマウントディスプレイデバイスのポーズを決定する、請求項 4 に記載のシステム。

**【請求項 6】**

前記カメラアセンブリによって少なくとも部分的に生成された深さマップおよび R G B 画像のうちの少なくとも 1 つを使用して、前記カメラアセンブリは、前記シーンマップの前記 3 次元座標空間内で較正される、請求項 1 に記載のシステム。

**【請求項 7】**

前記カメラアセンブリは、前記シーンマップの前記 3 次元座標空間内に前記カメラアセンブリをレジストレーションする際に支援するための慣性測定ユニットを含む位置センサを含んでいる、請求項 1 に記載のシステム。

20

**【請求項 8】**

前記カメラアセンブリは、第 1 のカメラアセンブリであり、前記 1 つまたは複数のカメラアセンブリは、第 2 のカメラアセンブリをさらに備えており、前記システムは、前記第 1 および第 2 のカメラアセンブリのうちのいずれかの視点から前記複合現実環境の表示を選択するための選択デバイスをさらに備えている、請求項 1 に記載のシステム。

**【請求項 9】**

少なくとも第 1 の 2 次元モニタと第 2 の 2 次元モニタとをさらに備えており、前記第 1 の 2 次元モニタは、前記第 1 のカメラアセンブリの視点から前記複合現実環境の前記画像を表示し、前記第 2 の 2 次元モニタは、前記第 2 のカメラアセンブリの視点から前記複合現実環境の前記画像を表示する、請求項 8 に記載のシステム。

30

**【請求項 10】**

実オブジェクトと仮想オブジェクトとを含む複合現実環境を提示するためのシステムであって、

前記複合現実環境において、現実世界オブジェクトの画像を異なる視点から取り込むための複数の画像取り込みデバイスと、

前記複数の画像取り込みデバイスに動作可能に結合された処理ユニットとを備えており、前記処理ユニットは、互いに対して前記複数の画像取り込みデバイスのポーズを識別し、前記処理ユニットはさらに、前記複合現実環境において仮想オブジェクトを生成し、また前記複数の画像取り込みデバイスに対する前記仮想オブジェクトの位置を、3 次元座標空間を含むシーンマップ上で識別し、前記処理ユニット、および前記複数の画像取り込みデバイスのうちの画像取り込みデバイスは、前記画像取り込みデバイスの視点から 2 次元モニタで表示するための前記複合現実環境の画像を生成する、システム。

40

**【請求項 11】**

前記画像取り込みデバイスは、ヘッドマウントディスプレイデバイス上にカメラを備えており、前記ヘッドマウントディスプレイデバイスは、前記ヘッドマウントディスプレイデバイスの光学系に前記仮想オブジェクトを表示するためのディスプレイユニットをさらに備えている、請求項 10 に記載のシステム。

50

**【請求項 12】**

前記画像取り込みデバイスは、カメラアセンブリを備えている、請求項10に記載のシステム。

**【請求項 13】**

前記カメラアセンブリは、前記複合現実環境における実オブジェクトのカラー画像を取り込むためのカラー画像取込デバイスと、前記複合現実環境の深さデータを取り込むための画像センサであって、前記複合現実環境の深さデータおよび深さマップを前記画像センサの視点から生成する画像センサとを備える、請求項12に記載のシステム。

**【請求項 14】**

前記処理ユニットは、RGB画像および深さマップのうちの少なくとも一方を使用して、他の画像取り込みデバイスに対する前記カメラアセンブリの前記ポーズを識別する、請求項13に記載のシステム。 10

**【請求項 15】**

前記カメラアセンブリは、他の画像取り込みデバイスに対する前記カメラアセンブリの前記ポーズを識別する際に支援するための慣性測定ユニットを含む位置センサを含んでいる、請求項12に記載のシステム。

**【請求項 16】**

実オブジェクトと仮想オブジェクトとを含む複合現実環境を提示する方法であって、

(a) ヘッドマウントディスプレイデバイスに組み込まれた第1のカメラを経由して、第1の視点から、前記複合現実環境の第1の画像を取り込むステップと、 20

(b) カメラアセンブリに組み込まれた第2のカメラを経由して、第2の視点から、前記複合現実環境の第2の画像を取り込むステップと、

(c) 前記複合現実環境の中の位置において仮想オブジェクトを生成するステップと、

(d) 前記ヘッドマウントディスプレイデバイスの光学系に前記仮想オブジェクトを表示するステップであって、前記仮想オブジェクトは、前記ヘッドマウントディスプレイデバイスの視点から表示されている、表示するステップと、

(e) 前記カメラアセンブリから2次元ディスプレイデバイス上に前記第2の画像を前記仮想オブジェクトと一緒に表示することであって、前記仮想オブジェクトは、前記カメラアセンブリの視点から表示されている、表示するステップと、

(f) 前記第1の画像および前記第2の画像中の共通のオブジェクトを少なくとも使用して、前記ヘッドマウントディスプレイデバイス、前記カメラアセンブリ、および仮想オブジェクトのお互いに対するポーズを、3次元座標空間を含むシーンマップ上にレジストレーションするステップと  
を含む方法。 30

**【請求項 17】**

前記2次元ディスプレイデバイスの表示を、前記仮想オブジェクトと一緒に前記第1の画像を前記ヘッドマウントディスプレイデバイスから表示することへと変更するステップをさらに含み、前記仮想オブジェクトは、前記ヘッドマウントディスプレイデバイスの視点から表示されている、請求項16に記載の方法。

**【請求項 18】**

前記仮想オブジェクトの前記位置の右側、左側、下側、および上側を含む方向のうちの少なくとも1つの方向から、前記カメラアセンブリによって、前記第2の画像を取り込むステップをさらに含む、請求項16に記載の方法。 40

**【請求項 19】**

仮想オブジェクトツールを使用して、前記ヘッドマウントディスプレイデバイスの着用者を前記仮想オブジェクトとして構築するステップをさらに含み、

前記2次元モニタ上に表示される前記カメラアセンブリからの前記第2の画像は、前記構築する際に、前記仮想オブジェクトツールおよび前記仮想オブジェクトを示すビデオフィードを含む、請求項16に記載の方法。

**【発明の詳細な説明】****【背景技術】****【0001】**

[0001] 複合現実は、仮想イメージを現実世界物理環境と混合することを可能にする技術である。実オブジェクトと、ユーザの視野の中に表示される仮想オブジェクトとの複合イメージを見るために、シースルーモードのヘッドマウント複合現実ディスプレイデバイスがユーザによって着用され得る。しかしながら、ヘッドマウントディスプレイデバイスを有していない他の人々は、複合現実環境を見ることができない。彼らは、実オブジェクトを見ることができるが、仮想オブジェクトを見ることはできない。

**【発明の概要】**

10

**【課題を解決するための手段】****【0002】**

[0002] 本技術の実施形態は、一般にモニタに表示され得る様々な視点からの複合現実環境のビューを取り込むためのシステムおよび方法に関する。本システムは、複合現実環境の内部のユーザ定義された位置における1つまたは複数の物理的カメラを含んでいる。本システムは、1つまたは複数の物理的カメラと時間的に同期し、また1つまたは複数のカメラの視点から複合現実環境において仮想オブジェクトをレンダリングする。複合現実環境からの実オブジェクトと仮想オブジェクトとは、次いで、他の人々によって見られるための1つまたは複数の外部の2Dモニタに1つまたは複数のカメラの視点から表示されてもよい。

20

**【0003】**

[0003] この概要は、以下の発明を実施するための形態においてさらに説明される概念の選択を、簡略化された形態で導入するために提供される。この概要は、特許請求の範囲の主題の重要な特徴または本質的な特徴を識別することを意図してはおらず、また特許請求の範囲の主題の範囲を決定する際の助けとして使用されることも意図してはいない。

**【図面の簡単な説明】****【0004】**

【図1】[0004]実オブジェクトと、仮想オブジェクトとを含む仮想現実環境の説明図である。

30

【図2】[0005]図1のカメラアセンブリによって取り込まれる複合現実環境のモニタ上の表示を示す図である。

【図3】[0006]本技術の実施形態によるカメラアセンブリのブロック図である。

【図4】[0007]ヘッドマウントディスプレイユニットの一実施形態の斜視図である。

【図5】[0008]ヘッドマウントディスプレイユニットの一実施形態の一部分の側面図である。

【図6】[0009]ヘッドマウントディスプレイユニットのコンポーネントについての一実施形態のブロック図である。

【図7】[0010]ヘッドマウントディスプレイユニットに関連する処理ユニットのコンポーネントについての一実施形態のブロック図である。

40

【図8】[0011]ヘッドマウントディスプレイユニットに関連する処理ユニットのソフトウェアコンポーネントについての一実施形態のブロック図である。

【図9】[0012]本システムのヘッドマウントディスプレイユニットに関連する1つまたは複数の処理ユニットのオペレーションを示すフローチャートである。

【図10】[0013]図9のフローチャートの中に示される様々なステップの例についてのより詳細なフローチャートである。

【図11】図9のフローチャートの中に示される様々なステップの例についてのより詳細なフローチャートである。

【図12】図9のフローチャートの中に示される様々なステップの例についてのより詳細なフローチャートである。

【図13】[0014]本技術のカメラアセンブリによって取り込まれる複合現実環境のモニタ

50

上の表示を示す図である。

【図14】[0015]本技術の実施形態によるシステムのブロック図である。

【発明を実施するための形態】

【0005】

[0016] 本技術の実施形態について次に図面を参照して説明される。これらの実施形態は、一般に、無制限の数の人々に対してモニタに表示可能である様々な視点からの複合現実環境のビューを取り込むためのシステムおよび方法に関する。本システムは、複合現実環境の内部のユーザ定義された位置に、1つまたは複数の物理的カメラを含んでいる。本システムは、1つまたは複数の物理的カメラと時間的に同期し、また1つまたは複数のカメラの視点から、複合現実環境をレンダリングする。シーンは、次いで、他の人たちが見るために1つまたは複数の外部の2Dモニタに表示されてもよい。10

【0006】

[0017] 図1は、現実世界オブジェクトと混合される仮想オブジェクトを含む複合現実環境10を示す。実施形態においては、本技術は、仮想オブジェクトを生成するためのコンテンツ生成ソフトウェアアプリケーションと一緒に使用されることもある。示される例においては、ユーザ18は、仮想コンテンツ40を、この例においては、仮想ロボットを生成している。仮想ロボットは、テーブル42の上に構築されており、このテーブル42は、実際のものであっても、仮想的なものであってもよい。示される例においては、ユーザは、仮想ツール44を適用しており、この例においては、ロボットの一部分を着色するためのペイントイングツールを適用している。1つまたは複数のカメラの視点からの仮想コンテンツおよび実コンテンツを表示するための本技術は、コンテンツ生成アプリケーション以外の、多種多様な複合現実アプリケーションと共に使用されることもある。20

【0007】

[0018] 仮想オブジェクトと、実オブジェクトとを含む複合現実環境は、処理ユニット4と一緒に機能するヘッドマウントディスプレイデバイス2を経由してユーザ18によって見ることができる。ヘッドマウントディスプレイデバイス2は、柔軟性のあるケーブル6を経由して処理ユニット4と通信することができる。ヘッドマウントディスプレイデバイスは、代わりに、処理ユニット4とワイヤレスに通信することもできる。さらなる実施形態においては、処理ユニット4は、ヘッドマウントディスプレイデバイス2の中に一体化されていることもある。一実施形態において、めがねの形状をしている、ヘッドマウントディスプレイデバイス2は、ユーザの頭の上に着用され、その結果、ユーザは、ディスプレイを通して見ることができ、またそれによって、ユーザの前の空間についての実際の直接のビューを有することができる。ヘッドマウントディスプレイデバイス2と、処理ユニット4についてのさらなる詳細が、以下で提供される。30

【0008】

[0019] 処理ユニット4は、ゲームアプリケーションや非ゲームアプリケーションなどのアプリケーションを実行するハードウェアコンポーネントおよび/またはソフトウェアコンポーネントを含むことができる。一実施形態においては、処理ユニット4は、本明細書において説明されるプロセスを実行するためのプロセッサ読取り可能ストレージデバイスの上に記憶される命令を実行することができる、標準プロセッサ、専用プロセッサ、マイクロプロセッサなどのプロセッサを含むことができる。実施形態においては、処理ユニット4は、1つまたは複数のリモートコンピューティングシステムにワイヤレスに（例えば、Wi-Fi（登録商標）通信手段、ブルートゥース（登録商標）通信手段、赤外線通信手段、または他のワイヤレス通信手段で）通信することができる。これらのリモートコンピューティングシステムは、コンピュータ、ゲームシステムまたはゲームコンソール、あるいはリモートサービスプロバイダを含むことができる。40

【0009】

[0020] ヘッドマウントディスプレイデバイス2と、処理ユニット4とは、互いに協調して、複合現実環境10においてユーザに仮想オブジェクトを、例えば、オブジェクト40、44を提示することができる。本技術の態様に従って、複合現実環境10は、他の人た50

ちが見る外部モニタに表示され得る複合現実環境 10 のビューを生成するために使用され得る 1 つまたは複数のカメラアセンブリ 50 をさらに含むことができる。具体的には、カメラアセンブリ 50 を使用して、ヘッドマウントディスプレイデバイス 2 を着用していない他の人たちにより、モニタ上で見られ得る、仮想オブジェクトと、実オブジェクトとを含む複合現実環境 10 のビューを生成することができる。

#### 【0010】

[0021] 示される例においては、カメラアセンブリ 50 は、カメラアセンブリ 50 a および 50 b を含んでいる。しかしながら、本技術は、さらなる実施形態においては、単一のカメラアセンブリ 50 、または 2 つよりも多いカメラアセンブリ 50 を用いて動作することもできることが理解される。特定のカメラビューが選択されるときに、複合現実環境の中の仮想オブジェクトと、実オブジェクトとは、選択されたカメラアセンブリの視点から、外部モニタに表示されることもある。すなわち、現実世界オブジェクトは、選択されたカメラアセンブリの物理的ロケーションのビュー視点からモニタに表示されることもある。また仮想オブジェクトは、選択されたカメラアセンブリの物理的ロケーションのビュー視点からモニタに表示されることもある。

10

#### 【0011】

[0022] 例えば、カメラアセンブリ 50 a が選択された場合、複合現実環境のビューは、図 2 に示されるように、カメラアセンブリ 50 a の物理的ロケーションの視点からのものになるであろう。例えば、ユーザ 18 、カメラアセンブリ 50 b 、場合によってはテーブル 42 などの現実世界オブジェクトは、カメラアセンブリ 50 a の錐台視野の内部にある。それゆえに、カメラアセンブリ 50 a によるそのような現実世界オブジェクトについての取り込まれる画像は、カメラアセンブリ 50 a の視点からモニタ 60 の上に表示されることもある。複数のカメラアセンブリが、同時に選択されることもあり、その結果、複数のカメラアセンブリのビューが、同時に 1 つまたは複数のモニタ 60 の上に示されることが、理解される。

20

#### 【0012】

[0023] 同様に、複合現実環境の内部の仮想オブジェクトは、3 次元空間の中に規定された位置を有しており、また仮想オブジェクトが見られるビュー視点 (viewing perspective) に応じて、異なる角度および側から見られることもある。図 2 において、仮想コンテンツ 40 や仮想ツール 44 などの仮想オブジェクトは、カメラアセンブリ 50 a の錐台視野の内部にある。したがって、そのような仮想オブジェクトは、カメラアセンブリ 50 a の位置と一致する 3 次元空間の中の位置からモニタ 60 の上に表示されることもある。結果は、カメラアセンブリ 50 a の視点からモニタ 60 の上の複合現実環境 10 のビューである。図 1 において、仮想オブジェクトと、実オブジェクトとは、与えられた視点から示される。この視点は、例えば、別のユーザによって着用される第 2 のヘッドマウントディスプレイデバイス 2 (図示されず) を通して見られる可能性がある。この視点は、代わりに、例えば、第 3 のカメラアセンブリ 50 (図示されず) の位置から取り込まれるモニタ 60 の上で見られる可能性もある。

30

#### 【0013】

[0024] 実施形態においては、処理ユニット 4 (単独で、または他の処理ユニットと組み合わされて機能している) は、部屋または他の環境の中で、ユーザの x 、 y 、 z デカルト位置を含む環境と、現実世界オブジェクトと、仮想 3 次元オブジェクトとのモデルを構築することができる。各ヘッドマウントディスプレイデバイス 2 の位置は、以下で説明されるように、複合現実環境のモデルに対してキャリブレーションされることもある。各カメラアセンブリ 50 の位置はまた、以下で説明されるように、複合現実環境のモデルに対してキャリブレーションされる可能性もある。さらなる実施形態において、1 つまたは複数のヘッドマウントディスプレイデバイス 2 と、カメラアセンブリ 50 との位置は、複合現実環境に対する代わりに、互いにキャリブレーションされることもある。本明細書においてシーンマップと称される環境の 3 次元モデル、ならびに各ユーザの F O V のすべてのトラッキング、各カメラアセンブリの錐台、および環境の中のオブジェクトは、モバイル処

40

50

理ユニット4それ自体によって生成され、または以下で説明されるように他の処理デバイスと協力して機能している可能性がある。

#### 【0014】

[0025] これらのキャリブレーションオペレーションは、システムが、複合現実環境の各ユーザの視線およびFOVと、複合現実環境の各カメラセンブリの視線および錐台視野とを決定することを可能にする。したがって、仮想画像は、各ユーザに表示されることもあるが、システムは、各ユーザの視点、またはカメラセンブリの視点からの仮想画像のディスプレイを決定する。さらに、以下で説明されるように、ヘッドマウントディスプレイデバイス2と、カメラセンブリ50とは、深センサおよび/または他のイメージセンサを含むことができる。これらのセンサは、与えられたデバイス2の視点、またはカメラセンブリ50の視点から、第1の仮想オブジェクトが、第2の仮想オブジェクトまたは実オブジェクトを閉塞しているか、あるいは第2の仮想オブジェクトまたは実オブジェクトによって閉塞されているかを決定することができる。10

#### 【0015】

[0026] 図3は、外部モニタ60およびコンピューティングシステム70と通信しているカメラセンブリ50の概略ブロック図である。コンピューティングシステム70は、処理ユニット4であるか、または処理ユニット4を含むことができ、あるいは代わりに、コンピューティングシステム70は、処理ユニット4と一緒に機能することもできる。図3に示されるカメラセンブリ50は、本明細書において説明されるカメラセンブリ50のうちのどれかとすることもできる。1つまたは複数のカメラセンブリ50は、複合現実環境10の内部の、または周囲のどこかに位置して、任意の角度から複合現実環境10の内部の実オブジェクトおよび/または仮想オブジェクトを取り込むことができる。カメラセンブリ50は、例えば、三脚の上に取り付けられ、あるいはそれ以外の方法で複合現実環境の中の、または周囲の望ましい位置で支えられて、複合現実環境の鳥瞰図を提供することができる。鳥瞰図は、複合現実環境の横から、下から、上から、または内部からの複合現実環境のビューである。カメラセンブリは、固定された位置にあることもあり、またはそれらは、複合現実環境の内部をあちこち動き回ることができる。20

#### 【0016】

[0027] カメラセンブリ50は、カメラ52と、位置センサ54と、イメージセンサ56とを含む。カメラセンブリ50は、さらなる実施形態においては、追加のコンポーネントを含むことができる。カメラ52は、カリフォルニア州、アーバイン市のRed.com社からの、例えば、Scarllet(登録商標)カメラなどのビデオカメラとすることができます。他のビデオカメラも企図される。実施形態においては、カメラ52は、動的な実オブジェクトのビデオを取り込み、またそれらが、移動し、または変化するときにそれらを表示することができる。さらなる実施形態においては、カメラセンブリ50のカメラ52は、代わりに、または追加して、実オブジェクトの静止画像を取り込むことができる。30

#### 【0017】

[0028] 位置センサ54を使用して、仮想オブジェクト、他のカメラセンブリ50、および/またはヘッドマウントディスプレイデバイス2に対して、複合現実環境10の内部のカメラセンブリ50の位置をキャリブレーションすることができる。例においては、位置センサ54は、以下で説明される慣性測定ユニット132に類似した慣性測定ユニットとすることができます。位置センサ54の慣性測定ユニットは、3軸磁力計、3軸ジャイロ、および/または3軸加速度計などの慣性センサを含むことができる。慣性測定ユニットは、カメラセンブリ50の位置、方向付け、および突然の加速度(ピッチ、ロールおよびヨー)を検知する。カメラセンブリ50(またはヘッドマウントディスプレイデバイス2)の位置および方向付けは、本明細書においては、カメラセンブリ50(またはヘッドマウントディスプレイデバイス2)のポーズと称される。慣性測定ユニットは、磁力計、ジャイロおよび加速度計に加えて、または磁力計、ジャイロおよび加速度計の代わりに他の慣性センサを含むことができる。4050

## 【0018】

[0029] 例えば、以前の時間からのカメラアセンブリ50の知られている位置および方向付けを使用して、位置センサ54の慣性測定ユニットを使用して、カメラアセンブリ50の位置および方向付けを決定することができる。次いで、慣性測定ユニットからの慣性情報を使用して、カメラアセンブリ50の位置および方向付けをアップデートすることができる。慣性測定ユニットからの情報は、カメラアセンブリについての正確な運動学的データを提供することができるが、慣性測定ユニットは、一般的に、カメラアセンブリの位置に関する絶対的な位置情報を提供することはない。「グラウンドトゥルース」とも称される、この絶対的な位置情報は、カメラアセンブリ50のカメラ52および／またはイメージセンサ56から取得される画像データから提供される可能性がある。

10

## 【0019】

[0030] 以下で説明されるように、ヘッドマウントデバイス2は、慣性測定ユニット132を含んでいる。したがって、実施形態においては、位置センサ54は、カメラ52に取り付けられたヘッドマウントディスプレイデバイス2とすることができます、またはヘッドマウントディスプレイデバイス2を含むことができる。そのような実施形態においては、慣性測定ユニット132（および場合によっては、以下で説明される他のアセンブリのいくつか）は、位置センサ54の一部分としてアクティブとすることができます、また以下で説明される他のアセンブリ（例えば、マイクロディスプレイ120）は、ディスエイブルにされ、またはそれ以外の方法で、非アクティブである可能性がある。

## 【0020】

20

[0031] イメージセンサ56を使用して、複合現実環境10の内部のカメラアセンブリ50から現実世界オブジェクトへの深さデータを取得することができる。実施形態においては、イメージセンサ56は、例えば、ワシントン州、レッドモンド市のマイクロソフト社が発売したXbox One（登録商標）プラットフォームと共に使用されるKinect（登録商標）カメラとすることができます。画像センサ56からのデータを使用して、ヘッドマウントディスプレイデバイス2に関して以下で説明される不透明フィルタ114と同一の不透明フィルタを実装することができる。そのような不透明フィルタは、カメラアセンブリ50の視点からのビューから部分的に、または完全にブロックされる仮想オブジェクトが、適切なオクルージョンと共に表示されることを保証する。

## 【0021】

30

[0032] カメラアセンブリ50は、ケーブル62を経由してモニタ60と通信することができます。モニタ60は、例えば、ユーザにビデオおよび／またはオーディオを提供することができるテレビジョン、モニタ、高精細度テレビジョン（HDTV: high-definition television）などとすることができます。例えば、カメラアセンブリ50は、グラフィックスカードなどのビデオアダプタ、および／または複合現実環境10に関連するオーディオ／ビジュアル信号を提供することができるサウンドカードなどのオーディオアダプタを含むことができます。モニタ60は、カメラアセンブリ50からオーディオ／ビジュアル信号を受信することができ、また次いで、モニタ60を見ている任意数の人々にオーディオ／ビジュアルに関連するビデオおよび／またはオーディオを出力することができる。モニタは、2次元モニタ、または3次元モニタとすることができます。さらなる実施形態においては、カメラアセンブリ50からモニタへとフィードを送信する代わりに、フィードはヘッドマウントディスプレイデバイス2へと送信されることもある。このデバイス2は、1つまたは複数のカメラアセンブリ50がある環境の中に存在するユーザ18によって着用されることもあり、あるいはこのデバイス2は、遠くにあり、ユーザ18以外のユーザによって着用されることもある。

40

## 【0022】

[0033] カメラアセンブリ50のすべてが、单一モニタ60に接続されることもある。さらに、单一モニタ60の代わりに、1つまたは複数のカメラアセンブリ50が、多数のモニタ60にオーディオ／ビデオ信号を出力することができる事が、理解される。実施形態においては、2つのモニタ60が存在しており、1つは、生カメラフィード（現実世界

50

オブジェクトだけを含んでいる)を表示しており、また第2のものは、同じ生カメラフィードを示しているが、さらに、生カメラフィードと融合された仮想オブジェクトを表示していることが、企図される。一実施形態によれば、モニタ60は、ケーブル62を経由してカメラアセンブリ50に接続されることもあり、このケーブルは、例えば、SDIビデオケーブル、S-ビデオケーブル、同軸ケーブル、HDMIケーブル、DVIケーブル、VGAケーブル、コンポーネントビデオケーブルなどとすることができる。カメラアセンブリ50からのフィードは、リアルタイム表示のためにモニタに送信されることもあり、あるいはそれは、その後の再生のために(任意の仮想オブジェクトが、同様に、その後の再生のために記憶されて)記録されることもある。アタッチされたモニタ60にフィードを送信する代わりに、フィードは、ウェブサーバに記憶されることもあり、またウェブストリームとしてアクセス可能にされ、またはそれ以外の方法でリモートロケーションにブロードキャストされることもある。

#### 【0023】

[0034] 以下で説明されるように、ヘッドマウントディスプレイデバイス2と、処理ユニット4とを使用して、複合現実環境10の内部の仮想オブジェクトと、実オブジェクトと、カメラアセンブリとの3次元座標を含むシーンマップを規定することができる。しかしながら、実施形態においては、処理ユニット4および/またはコンピューティングシステム70と一緒に機能する1つまたは複数のカメラアセンブリ50は、シーンマップを決定する際に支援することができる。コンピューティングシステム70が、処理ユニット4と分離している場合には、コンピューティングシステム70は、例えば、上記で説明されるようなケーブル62を用いて、カメラアセンブリ50に結合されることもある。さらなる実施形態(コンピューティングシステム70が、処理ユニット4と同じ、または処理ユニット4とは異なる場合)においては、コンピューティングシステム70は、カメラアセンブリ50から遠くにあり、また何らかの他の有線ネットワークまたはワイヤレスネットワークによってそれに接続されることもある。コンピューティングシステム70は、すべてのカメラアセンブリ50および/またはヘッドマウントディスプレイデバイス2と通信していることができる。

#### 【0024】

[0035] 図4および5は、ヘッドマウントディスプレイデバイス2の斜視図および側面図を示す。図5は、テンブル102と、鼻ブリッジ104とを有するデバイスの一部分を含むヘッドマウントディスプレイデバイス2の右側面だけを示す。サウンドを記録し、またそのオーディオデータを処理ユニット4に送信するためのマイクロフォン110が、以下で説明されるように鼻ブリッジ104に組み込まれている。ヘッドマウントディスプレイデバイス2の前に、ビデオ画像と静止画像とを取り込むことができる部屋に面しているビデオカメラ112が、存在している。これらの画像は、以下で説明されるように処理ユニット4に送信される。

#### 【0025】

[0036] ヘッドマウントディスプレイデバイス2のフレームの一部分は、ディスプレイ(1つまたは複数のレンズを含む)を取り囲むであろう。ヘッドマウントディスプレイデバイス2のコンポーネントを示すために、ディスプレイを取り囲むフレームの一部分は、示されていない。ディスプレイは、光誘導光学的要素115と、不透明フィルタ114と、シースルーレンズ116と、シースルーレンズ118とを含む。一実施形態においては、不透明フィルタ114は、シースルーレンズ116の背後にあり、またシースルーレンズ116と位置合わせされており、光誘導光学的要素115は、不透明フィルタ114の背後にあり、また不透明フィルタ114と位置合わせされており、またシースルーレンズ118は、光誘導光学的要素115の背後にあり、また光誘導光学的要素115と位置合わせされている。シースルーレンズ116および118は、めがねで使用される標準レンズであり、またどのような処方箋(処方箋の無い場合を含む)に対しても用いられる可能性がある。一実施形態においては、シースルーレンズ116および118は、可変の処方レンズによって置換される可能性がある。不透明フィルタ114は、(ピクセルごとに基づ

10

20

30

40

50

いて、または一様に、のいずれかで)自然光にフィルタを掛けて取り除いて、仮想イメージのコントラストを向上させる。光誘導光学的要素115は、目に人工的な光を導く。不透明フィルタ114と光誘導光学的要素115について、さらなる詳細が以下で提供される。

#### 【0026】

[0037] テンプル102に対して、またはテンプル102の内側に取り付けられて、画像ソースがあり、この画像ソースは、(一実施形態においては)仮想画像を投影するためのマイクロディスプレイ120と、マイクロディスプレイ120から光誘導光学的要素115へと画像を方向付けるためのレンズ122とを含んでいる。一実施形態においては、レンズ122は、コリメータレンズである。10

#### 【0027】

[0038] 制御回路136は、ヘッドマウントディスプレイデバイス2の他のコンポーネントをサポートする様々なエレクトロニクスを提供する。制御回路136のさらなる詳細は、図6に関して、以下で提供される。イヤホン130と、慣性測定ユニット132と、温度センサ138とが、テンプル102の内側にあり、またはテンプル102に取り付けられている。図6に示される一実施形態においては、慣性測定ユニット132(すなわちIMU132)は、3軸磁力計132A、3軸ジャイロ132B、3軸加速度計132Cなどの慣性センサを含んでいる。慣性測定ユニット132は、ヘッドマウントディスプレイデバイス2の位置と、方向付けと、突然の加速度(ピッチ、ロールおよびヨー)とを検知する。IMU132は、磁力計132A、ジャイロ132B、および加速度計132Cに加えて、または磁力計132A、ジャイロ132B、および加速度計132Cの代わりに他の慣性センサを含むことができる。20

#### 【0028】

[0039] マイクロディスプレイ120は、レンズ122を通して画像を投影する。マイクロディスプレイ120を実装するために使用され得る異なる画像生成技術が、存在している。例えば、マイクロディスプレイ120は、透過型投影技術を使用する際に実装される可能性があり、ここで光源は、白色光のバックライト付きの、光学的にアクティブな材料によって変調される。これらの技術は、通常、強力なバックライトと、高い光エネルギー密度とを有するLCDタイプディスプレイを使用して、実装される。マイクロディスプレイ120は、外部光が、光学的にアクティブな材料によって反射され、また変調される反射技術を使用して実装される可能性もある。照明は、技術に応じて、白色光源またはRGB光源のいずれかによって点灯されて転送される。デジタル光処理(DLP: digital light processing)技術と、液晶オンシリコン(LCOS: liquid crystal on silicon)技術と、クアルコム社のMirasol(登録商標)ディスプレイ技術は、ほとんどのエネルギーが、変調された構造から反射され、また本システムにおいて使用され得るので、効率のよい反射技術の例である。さらに、マイクロディスプレイ120は、光が、ディスプレイによって生成される場合の放射技術を使用して実装される可能性がある。例えば、マイクロビジョン社のPicop(登録商標)ディスプレイエンジンは、透過型要素としての役割を果たす微少スクリーンの上へ、または目の中に直接ビーム送信(例えば、レーザ)、のいずれかのマイクロミラーステアリングを用いてレーザ信号を放射する。30

#### 【0029】

[0040] 光誘導光学的要素115は、マイクロディスプレイ120からヘッドマウントディスプレイデバイス2を着用するユーザの目140へと光を送信する。光誘導光学的要素115はまた、ヘッドマウントディスプレイデバイス2の前からの光が、矢印142によって示されるように、光誘導光学的要素115を通して目140に送信されることを可能にしており、それによって、ユーザが、マイクロディスプレイ120から仮想画像を受信することに加えて、ヘッドマウントディスプレイデバイス2の前の空間に実際の直接のビューポートを有することを可能にしている。それゆえに、光誘導光学的要素115の壁は、シースルーである。光誘導光学的要素115は、第1の反射表面124(例えば、ミラー表面または他の表面)を含んでいる。マイクロディスプレイ120からの光は、レンズ1224050

を通過し、反射表面 124 の上に入射するようになる。反射表面 124 は、光が、内部反射によって光誘導光学的要素 115 を備えている平面基板の内側にトラップされるようになり、マイクロディスプレイ 120 からの入射光を反射する。基板の表面からのいくつかの反射の後に、トラップされた光の波は、選択的に反射する表面 126 のアレイに到達する。5つの表面のうちの1つだけが、ラベル 126 付けされて、図面が過密になることを防止していることに注意すべきである。反射表面 126 は、これらの反射表面に入射する光の波を基板からユーザの目 140 へと結合する。

### 【0030】

[0041] 異なる光線が、異なる角度で伝わり、また基板の内側から反射することになるので、異なる光線は、異なる角度で、様々な反射表面 126 に当たることになる。それゆえに、異なる光線は、反射表面のうちの異なる反射表面によって基板の外側に反射されることになる。どの光線が、どの表面 126 によって基板の外側に反射されることになるかについての選択は、表面 126 の適切な角度を選択することによって設計される。光誘導光学的要素についてのさらなる詳細は、2008年11月20日に公開された「Substrate-Guided Optical Devices」という名称の米国特許公開第2008/0285140号の中で見出される可能性がある。一実施形態においては、それぞれの目は、それ自体の光誘導光学的要素 115 を有するであろう。ヘッドマウントディスプレイデバイス 2 が、2つの光誘導光学的要素を有するときに、それぞれの目は、両方の目に同じ画像を、または2つの目に異なる画像を表示することができるそれ自体のマイクロディスプレイ 120 を有することができる。別の実施形態においては、両方の目の中に光を反射する1つの光誘導光学的要素が存在する可能性もある。10

### 【0031】

[0042] 光誘導光学的要素 115 と位置合わせされている不透明フィルタ 114 は、一様に、またはピクセル当たりに基づいて、のいずれかで、光誘導光学的要素 115 を通過しないように自然光を選択的にブロックする。不透明フィルタ 114 の一例についての詳細は、2010年9月21日に出願された「Opacity Filter For See-Through Mounted Display」という名称の Bar-Zeev 他による米国特許公開第2012/0068913号の中で提供されている。しかしながら、一般に、不透明フィルタ 114 の一実施形態は、シースルーライドパネル、エレクトロクロミックフィルム、または不透明フィルタとしての機能を果たすことができる類似したデバイスとすることができます。不透明フィルタ 114 は、ピクセルの高密度グリッドを含むことができ、ここで、各ピクセルの光透過率は、最小透過率と、最大透過率との間で、個別に制御可能である。0 ~ 100 % の透過率の範囲が理想的であるが、例えば、ピクセル当たりに約 50 % から 90 % までなどのより制限された範囲もまた、許容可能である。20

### 【0032】

[0043] アルファ値のマスクが、現実世界オブジェクトについてのプロキシを用いた z - バッファリングの後に、レンダリングパイプラインから使用される可能性がある。システムが、複合現実ディスプレイについてのシーンをレンダリングするときは、以下で説明されるように、どの現実世界オブジェクトが、どの仮想オブジェクトの前にあるかに注意する。仮想オブジェクトが、現実世界オブジェクトの前にある場合、そのときには不透明性が、仮想オブジェクトのカバレッジエリアについて、オンである可能性がある。仮想オブジェクトが、現実世界オブジェクトの背後に（仮想的に）ある場合、そのときには不透明性は、そのピクセルについてのどのようなカラーとも同様に、オフである可能性があり、そのようにして、ユーザは、実際光のその対応するエリア（サイズが1ピクセル以上）について、現実世界オブジェクトのみを見ることになる。カバレッジは、ピクセルごとにに基づいたものになり、そのようにして、システムは、仮想オブジェクトの一部分が、現実世界オブジェクトの前にあり、仮想オブジェクトの一部分が、現実世界オブジェクトの背後にあり、また仮想オブジェクトの一部分が、現実世界オブジェクトと一致している場合を取り扱うことができる。低いコスト、電力、および重さで 0 % から 100 % までの不透明4050

性に進むことができるディスプレイは、この使用のために最も望ましい。さらに、不透明フィルタは、カラーLCDを用いて、または有機LEDなどの他のディスプレイを用いて、カラーでレンダリングされる可能性もある。

### 【0033】

[0044] ヘッドマウントディスプレイデバイス2はまた、ユーザの目の位置を追跡するためのシステムを含んでいる。以下で説明されることになるように、本システムは、ユーザの位置と方向付けとを追跡することになり、その結果、本システムは、ユーザのFOVを決定することができる。代替的な実施形態においては、本システムは、ユーザのFOVの測定を改良するためにユーザの目の位置を追跡するための技術を含むことができる。例えば、ヘッドマウントディスプレイデバイス2は、アイトラッキングセンブリ（図示されず）を含むことができ、このアイトラッキングセンブリは、アイトラッキング照明デバイスと、アイトラッキングカメラとを有する。一実施形態においては、アイトラッキング照明デバイスは、1つまたは複数の赤外線（IR: infrared）エミッタを含んでおり、この赤外線（IR）エミッタは、目に向かってIR光を放射する。アイトラッキングカメラは、反射されたIR光を検知する1つまたは複数のカメラを含んでいる。ひとみの位置は、角膜の反射を検出すると知られているイメージング技法によって識別される可能性がある。例えば、2008年7月22日に発行された「Head Mounted Eye Tracking and Display System」という名称の米国特許第7,401,920号を参照のこと。そのような技法は、カメラを追跡することに関連した目の中心の位置を位置づけることができる。一般に、アイトラッキングは、目の画像を取得すること、およびコンピュータビジョン技法を使用して、眼窩の内部のひとみのロケーションを決定することを含む。一実施形態においては、目は、通常、調和して動くので、1つの目のロケーションを追跡することで、十分である。しかしながら、それぞれの目を別々に追跡することも可能である。10

### 【0034】

[0045] 図5は、ヘッドマウントディスプレイデバイス2の2分の1を示しているにすぎない。完全なヘッドマウントディスプレイデバイスは、もう1組のシースルーレンズと、もう1つの不透明フィルタと、もう1つの光誘導光学的要素と、もう1つのマイクロディスプレイ120と、もう1つのレンズ122と、部屋に面したカメラと、イヤホンと、温度センサとを含むことができる。20

### 【0035】

[0046] 図6は、ヘッドマウントディスプレイデバイス2の様々なコンポーネントを示すブロック図である。図7は、処理ユニット4の様々なコンポーネントを説明するブロック図である。そのコンポーネントが図6に示されるヘッドマウントディスプレイデバイス2を使用して、現実世界のユーザのビューと、1つまたは複数の仮想画像をシームレスに融合することにより、ユーザに仮想体験を提供している。さらに、図6のヘッドマウントディスプレイデバイスコンポーネントは、様々な状態を追跡する多数のセンサを含んでいる。ヘッドマウントディスプレイデバイス2は、処理ユニット4からの仮想画像についての命令を受信することになり、またセンサ情報を処理ユニット4に提供して戻すことになる。処理ユニット4は、図6のヘッドマウントディスプレイデバイスに従って、どこで、またいつ仮想画像をユーザに、提供し、また命令を送信すべきかを決定することができる。30

### 【0036】

[0047] 図6のコンポーネントのうちのいくつか（例えば、部屋に面したカメラ112と、マイクロディスプレイ120と、不透明フィルタ114と、イヤホン130と、温度センサ138と）は、シャドウの形で示されて、ヘッドマウントディスプレイデバイス2の左側に1つと、右側に1つ、これらのデバイスが、それぞれ2つ存在していることを示している。図6は、電力管理回路202と通信している制御回路200を示すものである。制御回路200は、プロセッサ210と、メモリ214（例えば、D-RAM）と通信しているメモリ制御装置212と、カーラインインターフェース216と、カメラバッファ218と、ディスプレイドライバ220と、ディスプレイフォーマッタ222と、タイミングジェ40  
50

ネレータ 226 と、ディスプレイアウトインターフェース 228 と、ディスプレイインインターフェース 230 とを含む。

#### 【0037】

[0048] 一実施形態においては、制御回路 200 のコンポーネントは、専用の回線、または 1つまたは複数のバスを経由して、互いに通信している。別の実施形態においては、制御回路 200 のコンポーネントは、プロセッサ 210 と通信している。カメラインターフェース 216 は、2つの部屋に面したカメラ 112 にインターフェースを提供し、また部屋に面したカメラから受信される画像をカメラバッファ 218 に記憶する。ディスプレイドライバ 220 は、マイクロディスプレイ 120 を駆動するであろう。ディスプレイフォーマッタ 222 は、マイクロディスプレイ 120 の上に表示されている仮想画像についての情報を不透明制御回路 224 に提供し、この不透明制御回路 224 は、不透明フィルタ 114 を制御する。タイミングジェネレータ 226 を使用して、システムについてのタイミングデータを提供する。ディスプレイアウトインターフェース 228 は、部屋に面したカメラ 112 から処理ユニット 4 へと画像を提供するためのバッファである。ディスプレイインインターフェース 230 は、マイクロディスプレイ 120 の上に表示されるべき仮想画像などの画像を受信するためのバッファである。ディスプレイアウトインターフェース 228 と、ディスプレイインインターフェース 230 とは、処理ユニット 4 に対するインターフェースであるバンドインターフェース 232 と通信する。10

#### 【0038】

[0049] 電力管理回路 202 は、電圧レギュレータ 234 と、アイトラッキング照明ドライバ 236 と、オーディオの DAC および増幅器 238 と、マイクロフォン前置増幅器およびオーディオ ADC 240 と、温度センサインターフェース 242 と、クロックジェネレータ 244 とを含む。電圧レギュレータ 234 は、処理ユニット 4 からバンドインターフェース 232 を経由して電力を受け取り、またその電力をヘッドマウントディスプレイデバイス 2 の他のコンポーネントに供給する。オーディオの DAC および増幅器 238 は、オーディオ情報をイヤホン 130 に出力する。マイクロフォン前置増幅器およびオーディオ ADC 240 は、マイクロフォン 110 についてのインターフェースを提供する。温度センサインターフェース 242 は、温度センサ 138 についてのインターフェースである。電力管理回路 202 はまた、電力を供給し、3軸磁力計 132A と、3軸ジャイロ 132B と、3軸加速度計 132C とからデータを受信して戻す。20

#### 【0039】

[0050] 図 7 は、処理ユニット 4 の様々なコンポーネントを説明するブロック図である。図 7 は、電力管理回路 306 と通信している制御回路 304 を示す。制御回路 304 は、中央演算処理装置 (CPU: central processing unit) 320 と、グラフィックス処理ユニット (GPU: graphics processing unit) 322 と、キャッシュ 324 と、RAM 326 と、メモリ 330 (例えば、D-RAM) と通信しているメモリ制御装置 328 と、フラッシュメモリ 334 (または他のタイプの不揮発性ストレージ) と通信しているフラッシュメモリ制御装置 332 と、バンドインターフェース 302 およびバンドインターフェース 232 を経由してヘッドマウントディスプレイデバイス 2 と通信しているディスプレイアウトバッファ 336 と、バンドインターフェース 302 およびバンドインターフェース 232 を経由してヘッドマウントディスプレイデバイス 2 と通信しているディスプレイインバッファ 338 と、マイクロフォンに接続するための外部マイクロフォンコネクタ 342 と通信しているマイクロフォンインターフェース 340 と、ワイヤレス通信デバイス 346 に接続するための PCI エクスプレスインターフェースと、USB ポート (単数または複数) 348 とを含む。一実施形態においては、ワイヤレス通信デバイス 346 は、Wi-Fi (登録商標) イネーブルド通信デバイス、ブルートゥース (登録商標) 通信デバイス、赤外線通信デバイスなどを含むことができる。処理ユニット 4 の上にデータまたはソフトウェアをロードし、同様に処理ユニット 4 に充電するために、USB ポートを使用して、処理ユニット 4 を処理ユニットコンピューティングシステム 22 にドッキングさせることができる。一実施形態においては、CPU 320 および GPU 322 は、仮想 3 次元オブ4050

ジェクトをどこで、いつ、またどのようにしてユーザの視野に挿入すべきかを決定するための主要な役に立つ機械である。さらなる詳細が、以下で提供される。

#### 【0040】

[0051] 電力管理回路306は、クロックジェネレータ360と、アナログデジタル変換器362と、バッテリ充電器364と、電圧レギュレータ366と、ヘッドマウントディスプレイ電源376と、温度センサ374（場合によっては処理ユニット4のリストバンドの上に位置している）と通信している温度センサインターフェース372とを含む。アナログデジタル変換器362を使用して、バッテリ電圧、温度センサを監視し、またバッテリ充電機能を制御する。電圧レギュレータ366は、電力をシステムに供給するためのバッテリ368と通信している。バッテリ充電器364を使用して、充電ジャック370から電力を受け取るとすぐに、（電圧レギュレータ366を経由して）バッテリ368を充電する。HMD電源376は、ヘッドマウントディスプレイデバイス2に電力を供給する。10

#### 【0041】

[0052] 図8は、ディスプレイデバイス2の部屋に面したカメラ112と、処理ユニット4の上のソフトウェアモジュールのうちのいくつかと、を含むモバイル複合現実アセンブリ30の高レベルブロック図を示す。これらのソフトウェアモジュールのうちのいくつかまたはすべては、代わりに、ヘッドマウントディスプレイデバイス2のプロセッサ210の上に実装されることもある。20

#### 【0042】

[0053] 示されるように、部屋に面したカメラ112は、ヘッドマウントディスプレイデバイス2の中のプロセッサ210に画像データを提供する。一実施形態においては、部屋に面したカメラ112は、深さカメラと、RGBカラー画像カメラと、IR光コンポーネントとを含んでいて、シーンの画像データを取り込むことができる。以下で説明されるように、部屋に面したカメラ112は、これらのコンポーネントのすべてよりも少ないものを含むことができる。20

#### 【0043】

[0054] 例えば、飛行時間分析を使用して、IR光コンポーネントは、シーンの上に赤外光を放射することができ、また次いで、センサ（図示されず）を使用して、例えば、深さカメラ、および／またはRGBカメラを使用して、シーンの中の1つまたは複数のオブジェクトの表面からの後方散乱された光を検出することができる。いくつかの実施形態においては、パルス赤外光が、発信光パルスと、対応する着信光パルスとの間の時間が測定され、またその時間を使用して、例えば、ユーザの手を含めて、部屋に面したカメラ112からシーンの中のオブジェクトの上の特定のロケーションへの物理的距離を決定することができる。さらに、他の例示の実施形態においては、発信する光波の位相は、着信する光波の位相と比較されて、位相シフトを決定することができる。次いで、位相シフトを使用して、取り込みデバイスからターゲットまたはオブジェクトの上の特定のロケーションに対する物理的距離を決定することができる。30

#### 【0044】

[0055] 別の例示の実施形態によれば、飛行時間分析を使用して、例えば、シャッターされた光パルスイメージングを含む様々な技法を経由して、時間と共に光の反射ビームの強度を分析することにより、部屋に面したカメラ112からオブジェクトの上の特定のロケーションまでの物理的距離を間接的に決定することができる。40

#### 【0045】

[0056] 別の例示の実施形態においては、部屋に面したカメラ112は、構造化された光を使用して、深さ情報を取り込むことができる。そのような分析においては、パターン化された光（すなわち、グリッドパターン、ストライプパターン、異なるパターンなどの知られているパターンとして表示される光）が、例えば、IR光コンポーネントを経由して、シーンの上に投影されることもある。シーンの中の1つまたは複数のターゲットまたはオブジェクトの表面に当たったすぐ後に、パターンは、それに応じて変形されるようにな50

る可能性がある。パターンのそのような変形は、例えば、3Dカメラおよび／またはRGBカメラ（および／または他のセンサ）によって取り込まれることもあり、また次いで分析されて、部屋に面したカメラ112からオブジェクトの上の特定のロケーションまでの物理的距離を決定することができる。いくつかの実装形態においては、IR光コンポーネントは、深さカメラおよび／またはRGBカメラから移動させられ、そのようにして三角測量を使用して、深さカメラおよび／またはRGBカメラからの距離を決定することができる。いくつかの実装形態においては、部屋に面したカメラ112は、IR光を検知する専用のIRセンサ、またはIRフィルタを有するセンサを含むことができる。

#### 【0046】

[0057] 本技術は、深さカメラと、RGBカメラと、IR光コンポーネント各々を用いることなく、オブジェクトと、それらオブジェクトの3次元位置とを検知することができることが、理解される。実施形態においては、部屋に面したカメラ112は、例えば、標準の画像カメラ（RGBまたは白黒）のみを用いて機能することができる。そのような実施形態は、個別に、または組み合わされて使用される様々な画像トラッキング技法により、動作することができる。例えば、単一の、標準の画像の、部屋に面したカメラ112は、特徴の識別とトラッキングとを使用することができる。すなわち、標準カメラからの画像データを使用して、シーンについての興味のある領域、または特徴を抽出することが可能である。ある期間にわたってこれらの同じ特徴を調べることにより、オブジェクトについての情報は、3次元空間において、決定されることもある。

#### 【0047】

[0058] 実施形態においては、ヘッドマウントディスプレイデバイス2は、2つの間隔の開けられた標準画像の部屋に面したカメラ112を含むことができる。この例においては、シーンの中のオブジェクトに対する深さは、2つのカメラのステレオ効果によって決定されることもある。各カメラは、何らかのオーバーラップした1組の特徴を撮像することができ、また深さは、それらの視野の中の視差の差から計算される可能性がある。

#### 【0048】

[0059] 知られていない環境の内部の位置情報を用いてシーンマップを決定するためのさらなる方法は、同時の位置特定およびマッピング（SLAM: simultaneous localization and mapping）として知られている。SLAMの1つの例は、「Systems and Methods for Landmark Generation for Visual Simultaneous Localization and Mapping」という名称の米国特許第7,774,158号の中で開示されている。さらに、IMU132からのデータを使用して、視覚トラッキングデータをもっと正確に解釈することができる。

#### 【0049】

[0060] 処理ユニット4は、シーンマッピングモジュール452を含むことができる。上記で説明されるような前に面したカメラ（単数または複数）112からのデータを使用して、シーンマッピングモジュールは、シーン（ユーザの手のうちの一方または両方を含む）の中のオブジェクトを3次元座標系にマッピングすることができる。シーンマッピングモジュールについてのさらなる詳細は、以下で説明される。

#### 【0050】

[0061] シーンの内部のユーザの位置を追跡するために、ユーザは、画像データから認識されることもある。処理ユニット4は、骨格の認識およびトラッキングモジュール448を実装することができる。骨格のトラッキングモジュール448の一例は、2012年3月2日に出願された「Skeletal Joint Recognition And Tracking System」という名称の米国特許公開第2012/0162065号の中に開示されている。そのようなシステムは、ユーザの手を追跡することもできる。しかしながら、実施形態においては、処理ユニット4は、さらに、手の認識およびトラッキングモジュール450を実行することもできる。モジュール450は、部屋に面したカメラ112から画像データを受信し、またFOVの中のユーザの手と、ユーザの手の

10

20

30

40

50

位置とを識別することができる。手の認識およびトラッキングモジュール450の一例は、2011年11月18日に出願された「System for Recognizing an Open or Closed Hand」という名称の米国特許公開第2012/0308140号の中で開示されている。一般に、モジュール450は、画像データを検査して、指、指と谷間との間の空間とすることができるオブジェクトの幅および長さを識別することができ、ここで、指は、それらの様々な位置においてユーザの手を識別し、また追跡するために一緒になる。

#### 【0051】

[0062] 処理ユニット4は、シーンの中の1人または複数人のユーザについての骨格モデルおよび／またはハンドデータを受信し、ユーザが処理ユニット4の上で実行されるアプリケーションに影響を及ぼす所定のジェスチャを実行しているか、またはアプリケーション制御の動作を実行しているかを決定するためのジェスチャ認識エンジン454をさらに含むことができる。ジェスチャ認識エンジン454についてのさらなる情報は、2009年4月13日に出願された「Gesture Recognizer System Architecture」という名称の米国特許出願第12/422,661号の中に見出される可能性がある。10

#### 【0052】

[0063] 上記で述べられるように、ユーザは、オブジェクトを選択し、場合によってはこれらのオブジェクトを修正する口語コマンドの形式の、様々な口頭によるジェスチャを実行することができる。それに応じて、本システムは、音声認識エンジン456をさらに含んでいる。音声認識エンジン456は、様々な知られている技術のどれかに従って動作することができる。20

#### 【0053】

[0064] 処理ユニット4は、同期ソフトウェアエンジン480と、シーンマップに合わせて1つまたは複数のヘッドマウントディスプレイデバイス2とカメラアセンブリ50とをキャリブレーションするためのキャリブレーションソフトウェアエンジン482とをさらに実装することができる。同期エンジン480と、キャリブレーションエンジン482とのオペレーションは、以下で説明される。

#### 【0054】

[0065] 上記で指摘されるように、カメラアセンブリ50は、イメージセンサ56を含んでいる。一実施形態においては、イメージセンサ56は、深さカメラと、IR光コンポーネントとを含んでいて、シーンの画像データを取り込むことができる。これらのコンポーネントを使用して、イメージセンサ56は、部屋に面したカメラ112が画像データを取り込む上記で説明された方法のうちのどれかに従って、カメラアセンブリ50の視点から、シーンの深さデータを取り込むことができる。イメージセンサ56が、例えば、深センサだけ、RGBカメラだけ、白黒カメラだけ、または一緒に機能する2つ以上のこれらのセンサ／カメラを使用して、様々な他の技法のうちのどれかによってシーンマップを構築することができるが、さらに理解される。例えば、深さデータは、2つのカラーカメラとステレオ分析とを使用して生成されることもある。そのようなシーンマップは、単一のカメラアセンブリ、または一緒に機能する複数のカメラアセンブリ（および場合によってはヘッドマウントディスプレイデバイス2）からのデータを使用して、構築されることがある。3040

#### 【0055】

[0066] 上記で指摘されるように、カメラアセンブリ50は、コンピューティングシステム70と通信することができる。コンピューティングシステム70と処理ユニット4とが、互いに別々になっている場合に、コンピューティングシステム70は、処理ユニット4の代わりに、または処理ユニット4に追加して、図8に関して上記で説明される処理ユニット4のソフトウェアモジュール448、450、452、454、456、480および／または482のうちのいくつかまたはすべてを実行することができる。

#### 【0056】

50

[0067] 例示の一実施形態においては、ヘッドマウントディスプレイデバイス2と、処理ユニット4とは、一緒に機能して、ユーザが中にあり、またその環境の中の様々な移動するオブジェクト、または静止したオブジェクトを追跡する環境についてのシーンマップまたはモデルを生成する。実施形態においては、1つまたは複数のカメラアセンブリ50のセンサからのデータは、シーンマップを生成する際に助けとなることができる。これらの特徴は、図9のフローチャートに関して、以下でより詳細に説明される。

#### 【0057】

[0068] さらに、処理ユニット4は、ヘッドマウントディスプレイデバイス2の位置および方向付けを追跡することにより、ユーザ18によって着用されるヘッドマウントディスプレイデバイス2のFOVを追跡する。ヘッドマウントディスプレイデバイス2によって取得される、例えば、部屋に面したカメラ112とIMU132からのセンサ情報は、処理ユニット4に送信される。処理ユニット4は、データを処理し、シーンモデルをアップデートする。処理ユニット4は、さらに、どこに、いつ、などのようにして、任意の仮想3次元オブジェクトを挿入すべきかについての命令をヘッドマウントディスプレイデバイス2に提供する。同様にして、コンピューティングシステム70は、1つまたは複数のカメラアセンブリ50の位置と、方向付けと、錐台視野とを追跡し、それに応じてシーンモデルをアップデートすることができる。これらの特徴はまた、図9のフローチャートに関して、より詳細に説明されることもある。

#### 【0058】

[0069] 図9は、画像データの單一フレームを生成し、レンダリングし、表示するためにかかる時間などの離散的期間中の、処理ユニット4と、ヘッドマウントディスプレイデバイス2と、1つまたは複数のカメラアセンブリ50と、コンピューティングシステム70とのオペレーションと対話性とについての高レベルフローチャートである。実施形態においては、データは、60Hzのレートで、リフレッシュされることもあるが、データは、さらなる実施形態においては、より頻繁に、またはそれほど頻繁ではないように、リフレッシュされることもある。

#### 【0059】

[0070] 仮想環境を表現するためのシステムは、ステップ600において構成されていることもある。例えば、システムのユーザ18またはオペレータは、提示されるべき仮想コンテンツと、どこに仮想コンテンツが提示されるべきかとを指定することができる。ユーザが、コンテンツ生成ソフトウェアエンジン458を用いて仮想コンテンツを生成している場合、複数の仮想ツールを含む仮想ツールのケースが、開かれ、また表示されることもある。

#### 【0060】

[0071] ステップ604において、処理ユニット4は、シーンからデータを収集する。これは、ヘッドマウントディスプレイデバイス2により、またとりわけ、部屋に面したカメラ112とIMU132とにより、検知されるデータとすることができる。これはまた、カメラアセンブリ50により、またとりわけ、カメラ50と、位置センサ54と、イメージセンサ56とにより、検知される画像と慣性データとを含むこともできる。

#### 【0061】

[0072] 以下で説明されるように、1つまたは複数のヘッドマウントディスプレイデバイス2と、1つまたは複数のカメラアセンブリ50の位置は、シーンマップの内部で互いにキャリブレーションされることもある。しかしながら、1つまたは複数のカメラアセンブリ50のカメラ52とイメージセンサ56と、1つまたは複数のヘッドマウントディスプレイデバイス2の部屋に面したカメラ112との内で受信される画像データが、異なるレートで受信されることが、起こる可能性がある。したがって、カメラアセンブリと、ヘッドマウントディスプレイデバイスとが、シーンマップに対してキャリブレーションされる可能性がある前に、本技術の実施形態は、ステップ606において、変数としての時間を取り除くために、異なるカメラアセンブリの画像取り込みデバイスと、ヘッドマウントディスプレイデバイスとから受信される画像データを時間的に同期させる。このようにして

10

20

30

40

50

、キャリブレーションエンジン 482 は、同時に取り込まれる、異なる画像取り込みデバイスからの画像データについての分析を実行することができる。以下の説明においては、カメラ 52、イメージセンサ 56、および／または部屋に面したカメラ 112 は、一括して、画像取り込みデバイスと称される。

#### 【0062】

[0073] 同期化ステップ 606 は、同期化エンジン 480 によって実行される。ステップ 606 に関するさらなる詳細は、次に、図 10 のフローチャートを参照して説明される。以下の説明は、異なる画像取り込みデバイスの深さカメラと、RGB カメラとから受信されるデータの同期化に対して適用されることもある。したがって、1つまたは複数のヘッドマウントディスプレイデバイス 2 と、1つまたは複数のカメラアセンブリ 50 とのすべての画像取り込みデバイスは、一緒に同期化されることもある。10

#### 【0063】

[0074] 各画像取り込みデバイスの内部で、シーン情報のフレームは、ステップ 720 において、深さカメラと、RGB カメラとによって取り込まれる。ステップ 724 において、各取り込みデバイスについての深さデータと RGB フレームデータとは、処理ユニット 4 および／またはコンピューティングシステム 70 に送信される。ステップ 726 において、着信する深さ／RGB データは、処理ユニット 4 またはコンピューティングシステム 70 のいずれかにおいて、単一マスタクロックでタイムスタンプを押される。

#### 【0064】

[0075] スタンプ 730 において、同期化エンジン 480 は、与えられたフレーム番号についてのデバイス 2 とアセンブリ 50 とのすべてから受信されるフレームデータを調べる。与えられたフレームについてのそのようなデータについてのタイムスタンプ、ならびにカメラのそれについての知られている解像度を使用して、同期化エンジンは、データを生成する各デバイス／アセンブリからの深さデータおよび RGB データについての必要とされる時間オフセットを決定する。これから、同期化エンジンは、各画像取り込みデバイスのタイミングにおいて必要とされる調整を決定することができる。同期化エンジンは、単一デバイス（例えば、そのフレームについての第 1 のデータを生成したデバイス）のフレームデータからのタイムスタンプを基準として選択し、またそのタイムスタンプを基準タイムスタンプとして設定することができる。次いで、すべての他のデバイスについてのフレームデータは、その基準タイムスタンプに合わせて調整されることもある。代わりに、同期化エンジンは、内部クロック基準を生成し、その内部クロック基準に対するすべてのフレームデータについての調整を設定することができる。2030

#### 【0065】

[0076] ひとたび各デバイスについての調整が、ステップ 730 において決定された後に、同期化エンジンは、ステップ 734 において、そのデバイスがそのフレームデータを生成するレートを調整するコマンドを各デバイス 2／アセンブリ 50 に送信することができる。フレームデータの生成を調整するために使用され得る様々なメカニズムが存在している。一実施形態においては、同期化エンジンは、フレームデータを生成する取り込みデバイスについてのプランキング間隔に追加し、またはそのプランキング間隔から差し引くことができ、その結果、すべての取り込みデバイスからのすべてのフレームデータが、同時に生成され、送信される。同期化エンジン 480 が、1つのデバイスまたはアセンブリからのフレームがより早くやってくるようにさせ、例えば、バッファアンダーフローを回避する必要がある場合に、同期化エンジン 480 は、そのデバイスについてのプランキング間隔を低減させるコマンドをデバイス／アセンブリに送信することができる。同期化エンジン 480 が、デバイスまたはアセンブリについてのフレームデータが、後で生成されるようにさせ、例えば、バッファオーバーフローを回避する必要がある場合に、同期化エンジン 480 は、そのデバイス／アセンブリについての垂直プランキング間隔を増大させることができる。40

#### 【0066】

[0077] ステップ 606 における同期化エンジン 480 のオペレーションに関連したさら50

なる詳細は、2010年5月3日に出願された「Heterogeneous Image Sensor Synchronization」という名称の出願者の同時係属の米国特許出願第12/772,802号の中に開示されている。本明細書における開示を仮定すると、当業者なら、ヘッドマウントディスプレイデバイス2の異なる取り込みデバイスと、カメラアセンブリ50とをキャリブレーションプロセスのために一緒に同期させるために使用され得る様々な他のオペレーションを正しく理解するであろう。

#### 【0067】

[0078] さらなる実施形態においては、同期化ステップが、初期キャリブレーションプロセスから省略させられ得ることが、考えられる。そのような実施形態は、運動ベクトルを使用して、取り込みポイント以外の時に、取り込まれた画像データポイントのロケーションを予測することができる。計算的に困難であり、また処理集約的であるが、そのような実施形態は、可能である。10

#### 【0068】

[0079] 次に、図9に戻ると、ひとたび画像データフレームが、ステップ606においてデータの移動するシーンについて同期させられた後に、取り込みデバイスは、ステップ608において、互いに、またシーンに対してキャリブレーションされることもある。以下の説明においては、z-軸は、カメラレンズからそのまま伸びるように規定され、またx-軸と、y-軸とは、それぞれ、z-軸からの水平オフセットと垂直オフセットとである。x軸と、y軸と、z軸とのこの定義は、例としてにすぎず、またそれらの軸は、さらなる実施形態においては、異なるように方向付けされることもある。各取り込みデバイスは、他の取り込みデバイスの空間とは異なるx、y、zデカルト空間を有することができる。以下で説明されるように、ひとたびすべての取り込みデバイスが、キャリブレーションされた後に、画像取り込みデバイスのすべてを含む単一の3D現実世界デカルト座標系が、規定されることもある。20

#### 【0069】

[0080] 実施形態においては、キャリブレーションオペレーションは、異なる取り込みデバイスからの深さ情報と、RGB情報との両方を使用して、キャリブレーションエンジン482によって実行される。深さカメラは、画像の中のポイントのx位置と、y位置と、z位置についての情報を提供するが、低い解像度を有する可能性がある。RGBカメラは、z方向における深さ情報を提供することはないが、一般的に、1つの取り込みデバイスのビューを別のものに相互に関連づけるために使用される不連続性またはキーを識別するための高い解像度と、テクスチャとを有している。それに応じて、本技術の実施形態は、キャリブレーションプロセスのために取り込みデバイスからの深さカメラと、RGBカメラとの両方を使用することができる。30

#### 【0070】

[0081] しかしながら、より計算集約的であり、またシーンの中のオブジェクトに依存しているが、取り込みデバイスを互いに、また2つ以上の取り込みデバイスからの深さ情報だけを使用したシーンに対してキャリブレーションすることが、可能なこともある。同様に、依然としてより計算集約的であり、またシーンの中のオブジェクトに依存しているが、2つ以上の取り込みデバイスからのRGB情報だけを使用して、互いにデバイスをキャリブレーションすることが、可能であることもある。RGB情報だけを使用してカメラビューをキャリブレーションするための技術は、例えば、2007年5月17日に公開された、「Navigating Images Using Image Based Geometric Alignment and Object Based Controls」という名称の米国特許公開第2007/0110338号の中に説明されており、この公開は、マイクロソフト社のPhotosynth（登録商標）画像認識ソフトウェアの技術の中で使用される。40

#### 【0071】

[0082] 深さデータと、RGBデータとを使用したキャリブレーションエンジン482による、画像取り込みデバイスのキャリブレーションの一実施形態に関連したさらなる詳細50

は、次に、図11のフローチャートを参照して説明される。ステップ740において、異なる画像取り込みデバイスは、シーンからの深さデータと、RGB画像データとを取り込む。ステップ742において、異なる取り込みデバイスについて、深さカメラとRGBカメラそれぞれのカメラビューは、カメラ中心のビューから生成するカメラに特有の直交3D世界ビューへと変換される。いくつかの深さカメラは、基準の垂直フレームを使用して、取り込まれたデータを測定する。すなわち、カメラは、カメラからさらに離れているオブジェクトについて増大するz-方向の中で平面を定義する。 $x$ 次元と、 $y$ 次元とは、 $z$ -軸からの水平分散と、垂直分散とを表す。他の深さカメラは、基準の放射状フレームを使用して、取り込まれたデータを測定する。すなわち、カメラは、カメラにおける第1のポイント( $x_1, y_1, z_1$ )と、オブジェクトにおける第2のポイント( $x_2, y_2, z_2$ )とを有するベクトルを測定する。

#### 【0072】

[0083] どのようにしてカメラが、その画像データポイントを測定するかにかかわらず、理想的には、測定における任意の2つの隣接する座標の間の距離は、同じである。しかしながら、カメラレンズひずみに起因して、これは、当てはまらない可能性もあるであろう。それに応じて、ステップ742において、キャリブレーションエンジン482は、次に、デバイス2と、センブリ50との中の画像取り込みデバイスの中の深さカメラと、RGBカメラとの両方についてのカメラひずみを補正することができる。与えられたカメラ(深さまたはRGB)のひずみは、カメラ製造業者によって提供される、知られている特性とすることができます。そうでない場合には、例えば、カメラの視野の内部の異なるロケーションにおけるチェックカードパターンなど、知られている次元のオブジェクトを撮像することを含めて、カメラのひずみを算出するためのアルゴリズムが、知られている。その画像の中のポイントのカメラビュー座標における偏差は、カメラレンズひずみの結果になるであろう。ひとたびレンズひずみの程度が知られた後に、ひずみは、与えられた取り込みデバイスについてのポイントクラウドの中のポイントの一様なカメラビューマップをもたらす、知られている逆行列変換によって補正されることもある。

#### 【0073】

[0084] キャリブレーションエンジン482は、次に、ステップ746において、カメラビューから、直交3D世界ビューへと取り込みデバイスによって取り込まれるひずみ補正された画像データポイントを変換することができる。この直交3D世界ビューは、取り込みデバイスに関して直交 $x$ 、 $y$ 、 $z$ デカルト座標系における画像取り込みデバイスによって取り込まれるすべてのデータのポイントクラウドマップである(例えば、取り込みデバイスは、3D世界ビューの中の(0、0、0)にあるようにすることができる)。直交3D世界ビューに対してカメラビューを変換するための行列変換式は、知られている。例えば、David H. Eberly、「3d Game Engine Design: A Practical Approach To Real-Time Computer Graphics」、Morgan Kaufman Publishers(2000)を見て欲しい。

#### 【0074】

[0085] 取り込みデバイス深さカメラからのデータは、ステップ746において直交3D世界ビューを構築するために十分である可能性がある。しかしながら、取り込みデバイスRGBカメラからのデータが、ステップ746において3D世界ビューを構築する際に使用される可能性もある。RGBカメラは、データポイントの2Dマップに戻り、そのようにしてRGBデータだけから3D世界マップを導き出すことは、計算的に、より難しい。しかしながら、シーンの中のオブジェクトに応じて、RGBデータから3D世界マップを構築することは、可能であることもある。例えば、以前に述べられた米国特許公開第2007/0110338号を見て欲しい。

#### 【0075】

[0086] ヘッドマウントデバイス2と、カメラセンブリ50との中の各画像取り込みデバイスは、ステップ746において、直交3D世界ビューを構築することができる。与え

10

20

30

40

50

られた取り込みデバイスからのデータポイントのx、y、z世界座標は、依然として、ステップ746の結論におけるその取り込みデバイスの視点からのものであり、またデバイス2とカメラアセンブリ50との中の他の取り込みデバイスからのデータポイントのx、y、z世界座標にまだ相関づけられていない。さらなる実施形態においては、カメラレンズひずみについて補正するステップは、カメラビューを直交3D世界ビューに変換するステップ746の後に実行されることもある。

#### 【0076】

[0087] 次のステップは、異なる画像取り込みデバイスの様々な直交3D世界ビューをヘッドマウントディスプレイデバイス2とカメラアセンブリ50において、すべての画像取り込みデバイスによって共用される単一の全体の3D世界ビューへと変換することである。これを遂行するために、キャリブレーションエンジン482の実施形態は、次にステップ748で、それぞれの取り込みデバイスの世界ビューのポイントクラウドにおけるキーポイントの不連続性またはキーを調べ、次いで、ステップ750において、異なる取り込みデバイスの異なるポイントクラウドの間で同じであるキーを識別する。10

#### 【0077】

[0088] ひとたびキャリブレーションエンジン482が、2つの異なる取り込みデバイスの2つの世界ビューが同じキャリブレーションキーを含むことを決定することができた後に、エンジン482は、ステップ752において、互いに関する、またキーに関する2つの取り込みデバイスについての位置と、方向付けと、焦点距離とを決定することができる。実施形態においては、1つまたは複数のデバイス2および/またはカメラアセンブリ50の中の必ずしもすべての画像取り込みデバイスが、同じ共通キーを共用することになるとは限らない。しかしながら、第1の取り込みデバイスと、第2の取り込みデバイスとが、キーを共用しており、また第2の取り込みデバイスと、第3の取り込みデバイスとが、キーを共用している限り、キャリブレーションエンジンは、互いに、また単一の全体の3D世界ビューに、第1の取り込みデバイスと、第2の取り込みデバイスと、第3の取り込みデバイスとについての位置と、方向付けと、焦点距離とを決定することができる。同じことが、さらなる取り込みデバイスについても、当てはまる。20

#### 【0078】

[0089] このプロセスにおける第1のステップ748は、各取り込みデバイスについての3D現実世界ビューポイントクラウドの中でキーを識別することである。キーは、隣接するポイントと不連続であるデータポイントとることができ、またそれゆえに、異なる取り込みデバイスからのポイントクラウドの間で簡単に識別される。理想的には、いくつかのそのようなキーは、各ポイントクラウドの内部で識別され、また場合によっては、ポイントクラウドの内部のオブジェクトのエッジおよび/またはコーナーを規定することができる。画像ポイントクラウドの内部の他のオブジェクトの様子は、同様にキーとすることができます。様々な知られているアルゴリズムは、画像ポイントクラウドからキーを識別するために存在する。そのようなアルゴリズムは、例えば、K. Mikola jczyk、C. Schmid、「A Performance Evaluation Of Local Descriptors」、IEEE Transaction on Pattern Analysis & Machine Intelligence、27、10、1615～1630頁(2005年)の中で説明される。画像データを有するキーを検出するさらなる一方法は、スケール不变特徴変換(SIFT: Scale-Invariant Feature Transform)アルゴリズムである。SIFTアルゴリズムは、例えば、David G. Loweによる2004年3月23日に発行された「Method and Apparatus for Identifying Scale Invariant Features in an Image and Use of Same for Locating an Object in an Image」という名称の米国特許第6,711,293号の中で説明される。別のキー検出器方法は、最大限に安定した極値領域(MSER: Maximally Stable Extremal Regions)アルゴリズムである。MSERアルゴリズムは、例えば、J. Matas、O. Chum、M. U30  
40

rba、およびT. Pajdlaによる論文、「Robust Wide Baseline Stereo From Maximally Stable External Regions」、Proc. of British Machine Vision Conference、384~396頁(2002年)の中で説明される。

#### 【0079】

[0090] ステップ750において、2つ以上の取り込みデバイスからのポイントクラウドの間で共用されるキューラが、識別される。概念的には、第1の組のベクトルが、第1の取り込みデバイスのデカルト座標系の中の、第1の取り込みデバイスと、1組のキューラとの間に存在しており、また第2の組のベクトルが、第2の取り込みデバイスのデカルト座標系の中の、第2の取り込みデバイスと、その同じ組のキューラとの間に存在している場合に、2つの座標系は、互いにに対して、両方の取り込みデバイスを含む単一のデカルト座標系へと解決され得る(resolve)。いくつかの知られている技法が、2つ以上の取り込みデバイスからのポイントクラウド間で共用されたキューラを見出すために存在している。そのような技法は、例えば、S. Arya、D.M. Mount、N.S. Netanyahu、R. Silverman、およびA.Y. Wu、「An Optimal Algorithm For Approximate Nearest Neighbor Searching Fixed Dimensions」、Journal of the ACM 45、6、891~923頁(1998年)の中で示される。他の技法が、それだけには限定されないが、ハッシングまたはコンテキスト依存のハッシングを含めて、上記で述べられるArya他の近似的な最も近い隣接ソリューションの代わりに、またはそれに追加して、使用される可能性がある。

10

#### 【0080】

[0091] 2つの異なる取り込みデバイスからのポイントクラウドが、十分に多数のマッチしたキューラを共用する場合に、2つのポイントクラウドを一緒に相互に関連づける行列は、例えば、ランダムサンプリングコンセンサス(RANSAC: Random Sampling Consensus)、または様々な他の推定技法によって推定されることもある。次いで、回復された基本的な行列に対する異常値であるマッチは、取り除かれることがある。1対のポイントクラウドの間の1組の仮定された幾何学的に一貫したマッチを見出した後に、マッチは、それぞれのポイントクラウドについての1組のトラックへと組織化されることもあり、ここで、トラックは、ポイントクラウドの間の1組の相互にマッチしたキューラである。その組の中の第1のトラックは、第1のポイントクラウドの中の各共通キューラの投影を含むことができる。組の中の第2のトラックは、第2のポイントクラウドの中の各共通キューラの投影を含むことができる。

20

#### 【0081】

[0092] それぞれのポイントクラウドの中のキューラが、グラフの頂点の組を形成し、またマッチしたキューラの各対の間のグラフの中に、エッジが存在している場合、そのときにはこのグラフのあらゆる接続されたコンポーネントは、トラックを含んでいる。しかしながら、いくつかの接続されたコンポーネントに関するトラックは、一貫していない可能性があり、とりわけ、トラックは、それが同じ画像についての複数のキューラを含む場合に、一貫していない。実施形態においては、少なくとも2つのキューラを含んでいる一貫したトラックだけが、以下で説明される相対的な取り込みデバイスの位置を決定するステップ752のために、保持される。

30

#### 【0082】

[0093] ステップ748および750から的情報を使用して、異なる取り込みデバイスからのポイントクラウドは、単一の直交3D現実世界ビュー内の単一のポイントクラウドへと解決され得る。すべての取り込みデバイスの位置と方向付けとは、この単一のポイントクラウドと、単一の直交3D現実世界ビューとに関してキャリブレーションされる。様々なポイントクラウドを一緒に解決するために、2つのポイントクラウドについてのトラックの組の中のキューラの投影は、分析される。これらの投影から、キャリブレーションエンジン482は、キューラに関する第1の取り込みデバイスの視点を決定することができ、ま

40

50

たキューに関する第2の取り込みデバイスの視点を決定することもできる。それから、キャリレーションエンジン482は、単一のポイントクラウドの最良の推定値、および両方のポイントクラウドからのキューと他のデータポイントとを含む単一の直交3D現実世界ビューへとポイントクラウドを解決することができる。

#### 【0083】

[0094] このプロセスは、単一の直交3D現実世界ビューが、すべての画像取り込みデバイスを含むまで、任意の他の取り込みデバイスについて反復される。ひとたびこれが行われた後に、キャリレーションエンジンは、単一の直交3D現実世界ビューに対する、また互いに対する、取り込みデバイスの相対的な位置および方向付けを決定することができる。キャリレーションエンジンは、さらに、単一の直交3D現実世界ビューに関して各取り込みデバイスの焦点距離を決定することができる。10

#### 【0084】

[0095] 取り込みデバイスについてのポーズと、レンズ視野と、ひずみパラメータとを決定するキャリレーションプロセスの一例として、2つ以上の取り込みデバイス（複数のカメラアセンブリあるいは1つまたは複数のカメラアセンブリ、およびヘッドマウントディスプレイデバイス）は、キャリレーションキューの1つまたは複数の画像を取り込むことができる。このキャリレーションキューは、複数の異なる方向付けから取り込まれることもある。取り込まれた画像は、各画像取り込みデバイスの位置／方向付けおよびレンズの特性を決定するために、上記で説明されるように、処理される。

#### 【0085】

[0096] キャリレーションステップの完了のすぐ後に、様々なヘッドマウントディスプレイデバイスおよび／またはカメラアセンブリの位置と方向付けとは、知られていることもある。ヘッドマウントディスプレイデバイスのIMU132からの、またカメラアセンブリ50の位置センサ54からのIMUデータを使用して、1つまたは複数のヘッドマウントディスプレイデバイス2とカメラアセンブリ50との位置および方向付けを洗練し、かつ／または確認することができる。20

#### 【0086】

[0097] キャリレーションエンジン482によって実行され、また図9ないし11に関して、上記で説明されるキャリレーションステップは、1つまたは複数のヘッドマウントディスプレイデバイス2とカメラアセンブリ50との初期使用のすぐ後に一度実行されることもある。シーンの内部のユーザと、オブジェクトとは、変化することができ、またこのようにしてそれぞれの取り込みデバイスによって生成されるポイントクラウドは、変化することができるが、互いに関するカメラのキャリレーションは、同じままである可能性がある。カメラアセンブリが移動される場合に、キャリレーションステップは、反復されることもある。実施形態においては、キャリレーションエンジンは、様々な取り込みデバイスのポイントクラウドからのキューに関するデータを収集し続けることができ、またキャリレーションエンジンが、2つ以上の取り込みデバイスからのよりよいキューを識別する場合には、キャリレーションエンジンは、取り込みデバイスのキャリレーションを改善する試みの中で、上記のステップを反復することができる。30

#### 【0087】

[0098] シーンマップは、シーンの形状、ならびに形状、デバイス2とカメラアセンブリ50との位置、およびシーンの内部のオブジェクトの位置を識別するステップ610の中で作成されることもある。実施形態においては、与えられたフレームの中で生成されるシーンマップは、ユーザの手（単数または複数）のx位置、y位置、およびz位置と、シーンの中の他の現実世界オブジェクトおよび仮想オブジェクトとを含むことができる。深さデータと位置データとを収集するための方法は、上記で説明されてきている。40

#### 【0088】

[0099] ステップ612において、システムは、上記で説明されるように、ユーザの骨格、および／または手を検出し、追跡し、また移動する身体の部分と、他の移動するオブジェクトとの位置に基づいて、シーンマップをアップデートすることができる。ステップ650

14においては、処理ユニット4は、シーンの内部のヘッドマウントディスプレイデバイス2のx位置、y位置、およびz位置と、方向付けと、FOVとを決定する。ステップ614のさらなる詳細は、次に、図12のフローチャートに関して説明される。

#### 【0089】

[00100] ステップ760において、シーンについての画像データは、ユーザの頭部位置と、ユーザの顔からまっすぐに外を見ている顔単位ベクトルとの両方を決定するために、処理ユニット4によって分析される。頭部位置は、ヘッドマウントディスプレイデバイス2からのフィードバックから識別されることもあり、またこれから、顔単位ベクトルは、構築されることもある。顔単位ベクトルを使用して、ユーザの頭部方向付けを規定することができ、また例においては、ユーザについてのFOVの中心と考えられることもある。顔単位ベクトルはさらに、または代わりに、ヘッドマウントディスプレイデバイス2の上で部屋に面したカメラ112から戻されるカメラ画像データから識別されることもある。とりわけ、ヘッドマウントディスプレイデバイス2の上のカメラ112が何を見るかに基づいて、処理ユニット4は、ユーザの頭部方向付けを表す顔単位ベクトルを決定することができる。10

#### 【0090】

[00101] ステップ764において、ユーザの頭部の位置および方向付けは、さらに、または代わりに、より早い時間から（フレームの中で早期に、または先行するフレームからのいすれかで）ユーザの頭部の位置および方向付けについての分析から決定されることもあり、また次いでIMU132からの慣性情報を使用して、ユーザの頭部の位置および方向付けをアップデートしている。IMU132からの情報は、ユーザの頭部についての正確な運動学的データを提供することができるが、IMUは、一般的に、ユーザの頭部に関する絶対的な位置情報を提供することはない。この絶対的な位置情報、またはグラウンドトゥルースは、ヘッドマウントディスプレイデバイス2の上のカメラから取得される画像データから提供されることもある。20

#### 【0091】

[00102] 実施形態においては、ユーザの頭部の位置および方向付けは、相前後して動作するステップ760と764とによって決定されることもある。さらなる実施形態においては、ステップ769と764とのうちの一方またはもう一方を使用して、ユーザの頭部についての頭部の位置および方向付けを決定することができる。30

#### 【0092】

[00103] ユーザがまっすぐに前を見ていないことが、起こる可能性がある。したがって、ユーザの頭部の位置および方向付けを識別することに加えて、処理ユニットは、さらに、ユーザの頭部の中のユーザの目の位置を考慮することができる。この情報は、上記で説明されるアイトラッキングアセンブリによって提供されることもある。アイトラッキングアセンブリは、ユーザの目の位置を識別することができ、このユーザの目の位置は、ユーザの目が集中され、またまっすぐ前を見ている（すなわち、顔単位ベクトル）位置からの左方偏差、右方偏差、上方偏差、および/または下方偏差を示している目単位ベクトルとして表される可能性がある。顔単位ベクトルは、目単位ベクトルに合わせて調整されて、ユーザがどこを見ているかを規定することができる。40

#### 【0093】

[00104] ステップ770において、ユーザのFOVは、次に、決定される可能性がある。ヘッドマウントディスプレイデバイス2のユーザのビューの範囲は、仮説的ユーザの上方周辺ビジョンと、下方周辺ビジョンと、左方周辺ビジョンと、右方周辺ビジョンとに基づいて、あらかじめ規定されることもある。与えられたユーザについて算出されたFOVが、特定のユーザが、FOVの範囲において見ることができる可能性があるオブジェクトを含むことを保証するために、この仮説的ユーザは、最大の可能な周辺ビジョンを有するものと解釈されることもある。何らかの所定の余分なFOVが、これに追加されて、十分なデータが、実施形態における与えられたユーザのために取り込まれることを保証することができる。50

**【 0 0 9 4 】**

[00105] 与えられた瞬間ににおけるユーザについての F O V は、次いで、目単位ベクトルの任意の偏差によって調整される、顔単位ベクトルの周囲に、視界を取ること、およびそれを集中させることにより、算出されることもある。与えられた瞬間に、何をユーザが見ているかを規定することに加えて、ユーザの F O V についてのこの決定はまた、何がユーザにとって目に見えない可能性があるかを決定するために、有用であることもある。以下で説明されるように、特定のユーザの F O V の内部にあるこれらのエリアに仮想オブジェクトの処理を限定することは、処理速度を改善させることができ、またレイテンシを低減させる。

**【 0 0 9 5 】**

[00106] 図 9 をもう一度参照すると、ステップ 6 2 0 において、1つまたは複数のカメラセンブリ 5 0 の錐台視野が、決定される可能性がある。この決定は、ステップ 7 6 4 および 7 7 0 に関して、上記で説明されるものと同じやり方、または類似したやり方で行われることもある。

**【 0 0 9 6 】**

[00107] 本技術の態様に従って、ヘッドマウントディスプレイデバイス 2 またはカメラアセンブリ 5 0 のうちのどれかの視点からの、複合現実シーンの仮想オブジェクトおよび実オブジェクトのビューが、1つまたは複数のモニタ 6 0 の上に表示される可能性がある。ステップ 6 2 4 において、処理ユニット 4 および / またはコンピューティングシステム 7 0 は、ヘッドマウントディスプレイデバイス 2 またはカメラアセンブリ 5 0 のユーザまたはオペレータからの選択を受信することができる。その後に、仮想オブジェクトをレンダリングし、また表示するためのステップ 6 3 0 ~ 6 5 8 は、選択されたヘッドマウントディスプレイデバイス 2 またはカメラアセンブリ 5 0 の視点を使用する。以下のステップ 6 3 0 ~ 6 5 8 は、1つまたは複数のモニタ 6 0 の上に画像を表示することに関する。しかしながら、ステップ 6 3 0 ~ 6 5 8 の類似した組はまた、ヘッドマウントディスプレイデバイス 2 の視点から、ヘッドマウントディスプレイデバイス 2 に画像を表示するために並列に実行される可能性もある。

**【 0 0 9 7 】**

[00108] ステップ 6 3 0 において、処理ユニット 4 は、レンダリングオペレーションを選択することができ、その結果、選択されたカメラアセンブリまたはヘッドマウントディスプレイデバイスの最終的な錐台視野の内部に現れる可能性があるこれらの仮想オブジェクトだけが、レンダリングされる。他の仮想オブジェクトの位置が、依然として追跡される可能性もあるが、それらは、レンダリングされない。さらなる実施形態においては、ステップ 6 3 0 が、一緒にスキップされる可能性があり、全体の画像がレンダリングされることもまた考えられる。

**【 0 0 9 8 】**

[00109] 処理ユニット 4 および / またはコンピューティングシステム 7 0 は、次に、レンダリングセットアップステップ 6 3 8 を実行することができ、ここでセットアップレンダリングオペレーションは、ステップ 6 1 0 、 6 1 4 および 6 2 0 において受信されるシーンマップと F O V とを使用して実行される。ひとたび仮想オブジェクトデータが受信された後に、処理ユニットは、モニタ 6 0 の上でレンダリングされるべきである仮想オブジェクトについてのステップ 6 3 8 において、レンダリングセットアップオペレーションを実行することができる。ステップ 6 3 8 におけるセットアップレンダリングオペレーションは、表示される仮想オブジェクト（単数または複数）に関連する共通のレンダリングタスクを含むことができる。これらのレンダリングタスクは、例えば、シャドウマップ生成と、照明と、アニメーションとを含むことができる。実施形態においては、レンダリングセットアップステップ 6 3 8 は、仮想オブジェクトが、予測された最終的な F O V の中に表示される頂点バッファ、テクスチャ、状態など、ありそうな描画情報についてのコンパイルをさらに含むことができる。

**【 0 0 9 9 】**

10

20

30

40

50

[00110] 3 D シーンマップの中のオブジェクトのロケーションに関する情報を使用して、処理ユニット 4 / コンピューティングシステム 7 0 は、次に、ステップ 6 4 4において選択されたデバイス 2 / アセンブリ 5 0 のビューからのオクルージョンと、シェーディングとを決定することができる。とりわけ、シーンマップは、どのようにも移動する仮想オブジェクトまたは実オブジェクトと、移動していない仮想オブジェクトまたは実オブジェクトとを含めて、シーンの中にオブジェクトの x 位置と、y 位置と、z 位置とを有する。FOV の中の選択されたビューのロケーションと、オブジェクトに対するその視線とを知つて、処理ユニット 4 および / またはコンピューティングシステム 7 0 は、次いで、仮想オブジェクトが、現実世界オブジェクトのビューを部分的に閉塞するか、または完全に閉塞するかを決定することができる。さらに、処理ユニット 4 および / またはコンピューティングシステム 7 0 は、現実世界オブジェクトが、仮想オブジェクトのビューを部分的に閉塞するか、または完全に閉塞するかを決定することができる。10

#### 【0100】

[00111] ステップ 6 4 6 において、処理ユニット 4 の GPU 3 2 2 ( またはコンピューティングシステム 7 0 の中の GPU ) は、次に、モニタ 6 0 の上に表示される画像をレンダリングすることができる。レンダリングオペレーションの部分は、すでに、レンダリングセットアップステップ 6 3 8 において実行されてきており、また定期的にアップデートされてきている可能性がある。どのようなオクルージョンが引き起こされた仮想オブジェクトもレンダリングされないこともあります、またはそれらは、レンダリングされることもある。レンダリングされる場合に、オクルージョンが引き起こされたオブジェクトは、上記で説明されるように、不透明フィルタ 1 1 4 により、表示から除外されることになる。20

#### 【0101】

[00112] ステップ 6 5 0 において、処理ユニット 4 および / またはコンピューティングシステム 7 0 は、モニタ 6 0 にレンダリングされた画像を送信する時間であるかどうか、あるいはヘッドマウントディスプレイデバイス 2 および / またはカメラアセンブリ 5 0 からのもっと最近の位置フィードバックデータを使用して、画像のさらなる洗練のための時間が依然として存在するかどうかをチェックする。6 0 ヘルツのフレームリフレッシュレートを使用したシステムにおいては、単一フレームは、約 1 6 ms である。

#### 【0102】

[00113] アップデートされた画像を表示する時間である場合、選択されたカメラアセンブリ 5 0 またはヘッドマウントデバイス 2 からの画像は、モニタ 6 0 に送信される。例えば、カメラアセンブリ 5 0 が選択される場合には、カメラ 5 2 によって取り込まれる現実世界画像は、表示のためにモニタ 6 0 に送信される。ヘッドマウントディスプレイが、選択される場合に、部屋に面したカメラ 1 1 2 の RGB カメラによって取り込まれる現実世界画像は、表示のためにモニタに送信される。さらに、決定された錐台視野の中のどのような仮想オブジェクトもまた、モニタ 6 0 に送信されて、適切なピクセルにおいて表示され、視点とオクルージョンとを明らかにする。このときに、不透明フィルタについての制御データがまた、処理ユニット 4 および / またはコンピューティングシステム 7 0 からモニタ 6 0 へと送信されて、不透明フィルタ設定を適用する。モニタ 6 0 は、次いで、ステップ 6 5 8 において、実オブジェクトと、仮想オブジェクトとの融合された画像を表示することができる。3040

#### 【0103】

[00114] 上記で説明される実施形態においては、カメラアセンブリ 5 0 からの深さデータおよび他の画像データを使用して、カメラアセンブリのビューにおけるオクルージョンを決定する。処理ユニット 4 と、ヘッドマウントディスプレイデバイス 2 とによって確立されるシーンマップからのデータは、シーンマップの中のオブジェクトの表面メッシュを決定する際に、使用される可能性もある。このデータと表面メッシュとは、カメラアセンブリのビューにおいて、オクルージョンを決定するときに、使用される可能性もある。

#### 【0104】

[00115] ユーザ 1 8 が、選択されたカメラアセンブリ 5 0 の視点から仮想オブジェクト50

のビューをブロックしている可能性がある。実施形態においては、カメラアセンブリ 50 からの深さデータおよび他の画像データ、および / または処理ユニット 4 とヘッドマウントディスプレイデバイス 2 からのシーンマップを使用して、ユーザ 18 が、選択されたカメラアセンブリ 50 の視点から、仮想オブジェクトの一部分またはすべてをブロックしていることに起因した、仮想オブジェクトのオクルージョンを決定することができる。

#### 【0105】

[00116] さらなる実施形態においては、カメラアセンブリ 50 からの深さデータもしくは画像データ、またはシーンマップを使用する代わりに、ユーザ 18 の知られているロケーションを使用して、選択されたカメラアセンブリ 50 の視点からのオクルージョンを決定することができる。例えば、ユーザ 18 によって着用されるヘッドマウントディスプレイデバイス 2 の位置は、知られている。ユーザ 18 の位置は、代わりに、または追加して、骨格トラッキングアルゴリズムによって、検出されることがある。その後に、ユーザ 18 のサイズと形状とを近似する幾何学的形状は、選択されたカメラアセンブリ 50 の視点からユーザによる仮想コンテンツのオクルージョンを決定する目的で構築され、また使用されることがある。10

#### 【0106】

[00117] 他方では、ステップ 650において表示される画像データのフレームを送信する時間ではまだない場合に、処理ユニットおよび / またはコンピューティングシステム 70 は、より最近のセンサデータが、最終的なビューの予測と、ビューの中のオブジェクトの最終的な位置とを改良するために、ループバックすることができる。とりわけ、ステップ 650において、依然として時間がある場合に、処理ユニット 4 および / またはコンピューティングシステム 70 は、ステップ 604 に戻って、ヘッドマウントディスプレイデバイス 2 とカメラアセンブリ 50 とからより多くの最近のセンサデータを獲得することができる。20

#### 【0107】

[00118] 処理ステップ 600 ないし 658 は、例としてのみ上記で説明される。これらのステップのうちの 1 つまたは複数は、さらなる実施形態においては省略されることもあり、ステップは、異なる順序で実行されることもあり、あるいは追加のステップが、追加されることもあることが、理解される。

#### 【0108】

[00119] 上記で説明されるシステムを使用して、仮想オブジェクトと実オブジェクトとを含む複合現実体験が、ヘッドマウントディスプレイデバイスを着用していない人々が見るためのモニタに表示される可能性もある。さらに、複合現実シーンの周囲の望ましい位置にカメラアセンブリ 50 を位置づけることにより、複合現実環境の異なるビューが、選択され、モニタに表示されることもある。30

#### 【0109】

[00120] 図 13 は、第 1 のカメラアセンブリ 50a ( 図示されず ) によって取り込まれ、第 1 のモニタ 60a に表示されるスクリーンショットを示す。スクリーンショットは、実オブジェクトと、仮想オブジェクト ( 仮想コンテンツ 40 と仮想ツール 44 を含む ) とを示している。第 2 のカメラアセンブリ 50b はまた、シーンの内部にもあり、またスクリーンショットの中に取り込まれる。第 1 のモニタ 60a はまた、やはりシーンの中にある第 2 のモニタ 60b を表示している。第 2 のモニタ 60b は、図 13 の中に示されるような同じ画像 ( すなわち、第 1 のモニタ 60a に示されている同じ画像 ) を示している。40

#### 【0110】

[00121] 図 14 は、本技術によるシステムの一実施形態についてのブロック図を示す。コンポーネント間の通信が、特有の有線接続またはワイヤレス接続を用いて示されるが、コンポーネント間の通信は、さらなる実施形態においては有線またはワイヤレスとすることができることが、理解される。

#### 【0111】

50

[00122] 本技術のいくつかの他の特徴は、複合現実環境のビデオフィードが、ネットワーク接続の上でリモートロケーションに送信され得ることを含んでいる。複合現実環境は、次いで、モニタで、またはリモートロケーションにあるヘッドマウントディスプレイデバイスを経由してのいずれかで、見られることがある。

#### 【0112】

[00123] さらなる実施形態においては、本技術は、固体世界にロックされた仮想オブジェクトを獲得するためにセンサがそれ自体を露出する瞬間に於いてポーズを収集する時間同期化システムおよび方法を使用している。仮想オブジェクトは、現実世界画像が、カメラ52から取得される時よりも後の時間にレンダリングされる。それは、カメラアセンブリの上のトラッキングデバイスからのポーズが、60Hzで受信されるからである。実施形態においては、本システムは、はっきりと異なる連続的な期間において、カメラポーズをとることができる。本システムは、次いで、これらの期間の間で補間して、サブミリ秒の精度で、与えられた時間の瞬間に於けるカメラポーズを正確に知ることができる。正確な時間における瞬間に於いて知られているカメラアセンブリの位置を用いて、仮想オブジェクトは、その正確な時間における瞬間にわたってレンダリングされて、現実世界ビデオフィードに対する仮想オブジェクトの極めて近いマッチングを提供している。そのような方法は、オブジェクトをレンダリングする際にわずかの遅延、例えば、100msから150msまでの遅延を使用するが、カメラアセンブリ50によって生成される画像の中で、仮想オブジェクトと、実オブジェクトとのシームレスな融合を提供する。

#### 【0113】

[00124] さらなる実施形態においては、時間におけるどのような与えられた瞬間に於いても、カメラアセンブリ50のカメラ52のズーム／焦点／露出／白色バランス／ISO設定は、処理ユニット4および／またはコンピューティングシステム70に電子的に通信される可能性があり、そのようにして、適切なズーム設定、ぼやけ設定、および他の設定は、その正確なフレームのために使用される可能性がある。これは、ハイエンドカメラのセンサの上でレンズによって生成される画像に、幾何学的に、また質的に、マッチする仮想画像のレンダリングを可能にする。

#### 【0114】

[00125] 上記で説明される実施形態は、カメラアセンブリ50と、ヘッドマウントディスプレイデバイス2を着用するユーザとを含んでいている。さらなる実施形態においては、ヘッドマウントディスプレイデバイス2は、省略されることもある。そのような一実施形態においては、ユーザ18は、空間の中のデバイスの位置を追跡することを可能にするセンサを含むデバイスを有することができる。そのようなデバイスは、例えば、セルラー方式電話またはコンピューティングタブレットとすることができる。シーンマップは、1つまたは複数のカメラアセンブリ50、および／またはユーザ18によって保持されるデバイスの中のセンサによって構築される可能性がある。複合現実環境は、モニタ60またはユーザによって保持されるデバイスを経由してユーザに表示されるべきである。

#### 【0115】

[00126] 要約すると、一実施形態においては、本技術は、実オブジェクトと仮想オブジェクトとを含む複合現実環境を提示するためのシステムであって、仮想環境において3次元仮想オブジェクトを表示するためのディスプレイユニットを含むヘッドマウントディスプレイデバイスと、1つまたは複数のカメラアセンブリと、ディスプレイデバイスと1つまたは複数のカメラアセンブリとに動作可能に結合された処理ユニットとを備えており、処理ユニットは、ヘッドマウントディスプレイと、3次元仮想オブジェクトと、1つまたは複数のカメラアセンブリとがレジストレーションされる(register)3次元座標空間を含むシーンマップを生成し、処理ユニットと、1つまたは複数のカメラアセンブリのうちのカメラアセンブリとは、カメラアセンブリの視点から、シーンマップを使用して、複合現実環境の表示のための画像を生成する、システムに関する。

#### 【0116】

[00127] さらなる実施形態においては、本技術は、実オブジェクトと仮想オブジェクト

10

20

30

40

50

とを含む複合現実環境を提示するためのシステムであって、複合現実環境において現実世界オブジェクトの画像を異なる視点から取り込むための複数の画像取り込みデバイスと、複数の画像取り込みデバイスに動作可能に結合された処理ユニットとを備えており、処理ユニットは、互いに対して複数の画像取り込みデバイスの位置を識別し、処理ユニットは、さらに、複合現実環境において仮想オブジェクトを生成し、また複数の画像取り込みデバイスに対して仮想オブジェクトの位置を識別し、処理ユニットと、複数の画像取り込みデバイスのうちの画像取り込みデバイスとは、画像取り込みデバイスの視点から2次元モニタで表示するための複合現実環境の画像を生成する、システムに関する。

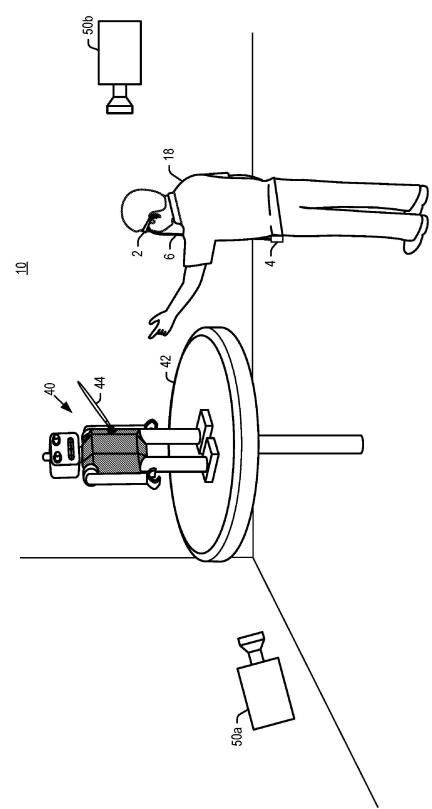
#### 【0117】

[00128] 別の実施形態においては、本技術は、実オブジェクトと仮想オブジェクトとを含む複合現実環境を提示する方法であって、(a) ヘッドマウントディスプレイデバイスに組み込まれた第1のカメラを経由して、第1の視点から、複合現実環境の第1の画像を取り込むステップと、(b) カメラアセンブリに組み込まれた第2のカメラを経由して、第2の視点から、複合現実環境の第2の画像を取り込むステップと、(c) 複合現実環境の中の位置において仮想オブジェクトを生成するステップと、(d) ヘッドマウントディスプレイデバイスの光学系に仮想オブジェクトを表示するステップであって、仮想オブジェクトは、ヘッドマウントディスプレイデバイスの視点から表示されている、表示するステップと、(e) 仮想オブジェクトと一緒にカメラアセンブリから2次元ディスプレイデバイスの上に第2の画像を表示するステップであって、仮想オブジェクトは、カメラアセンブリの視点から表示されている、表示するステップとを含む方法に関する。 10

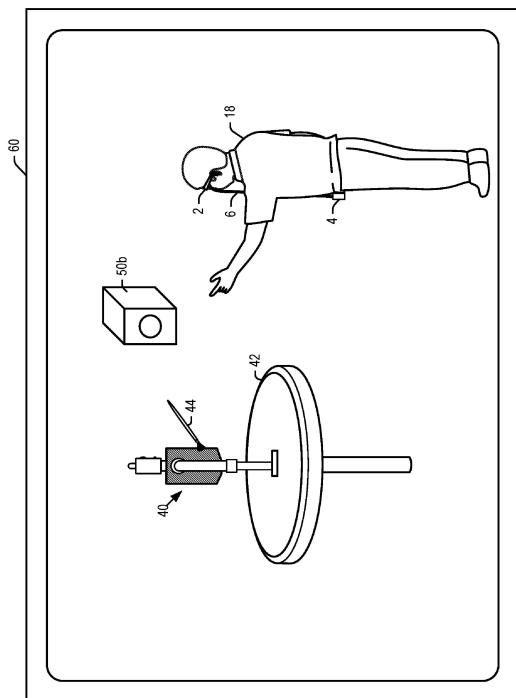
#### 【0118】

[00129] 本主題は、構造的特徴および／または方法論的動作に特有の言語において説明されてきているが、添付の特許請求の範囲の中で規定される主題は、必ずしも上記で説明される特定の特徴または動作だけに限定されるものとは限らないことを理解すべきである。もっと正確に言えば、上記で説明される特定の特徴および動作は、特許請求の範囲を実装する例示の形態として開示される。本発明の範囲は、この文書に添付された特許請求の範囲によって規定されることを意図している。 20

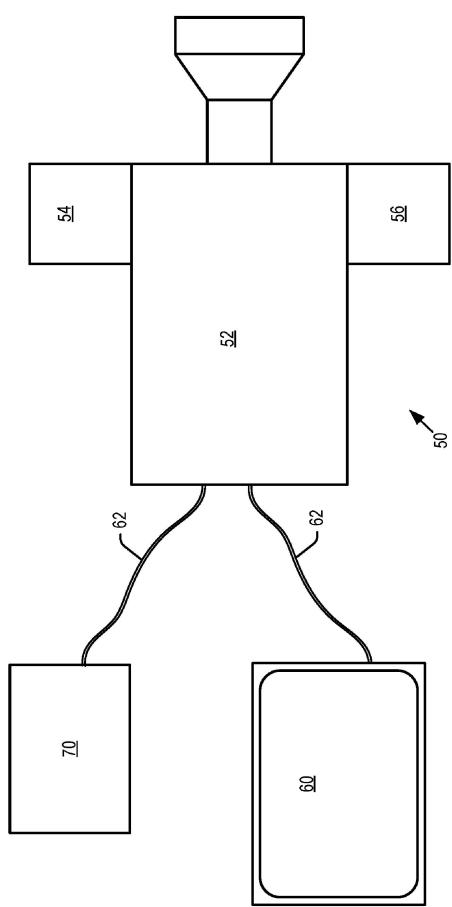
【図1】



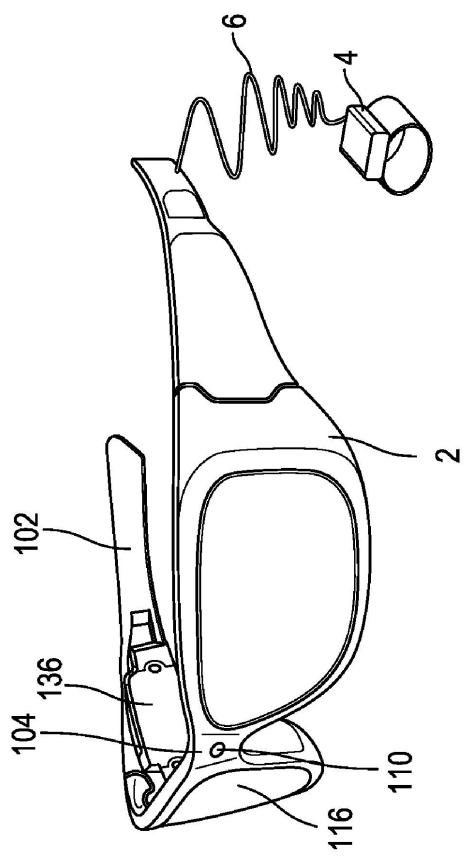
【図2】



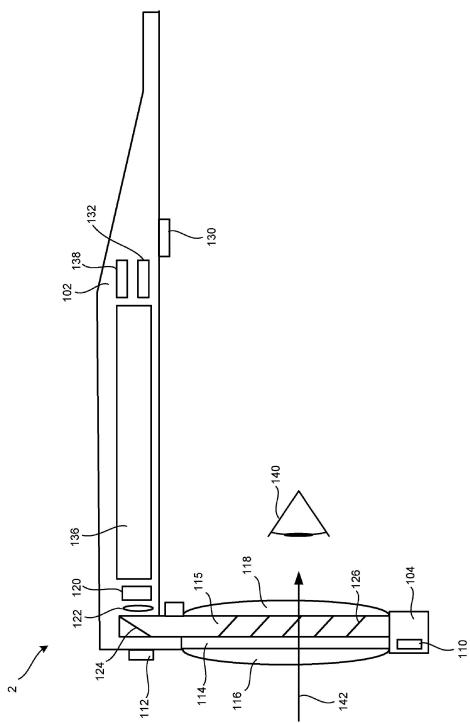
【図3】



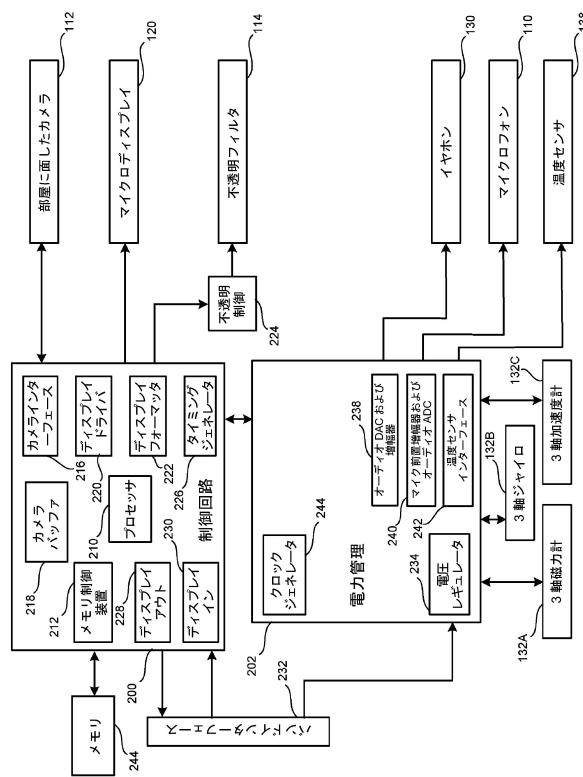
【図4】



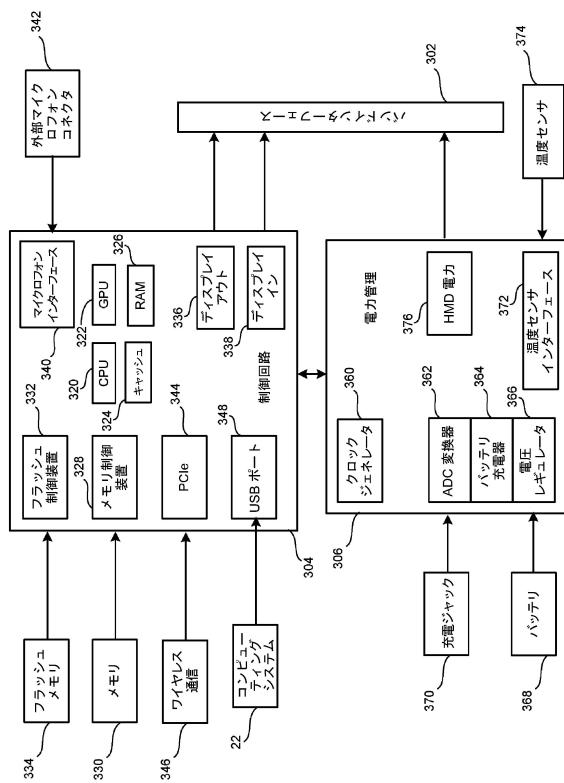
【図5】



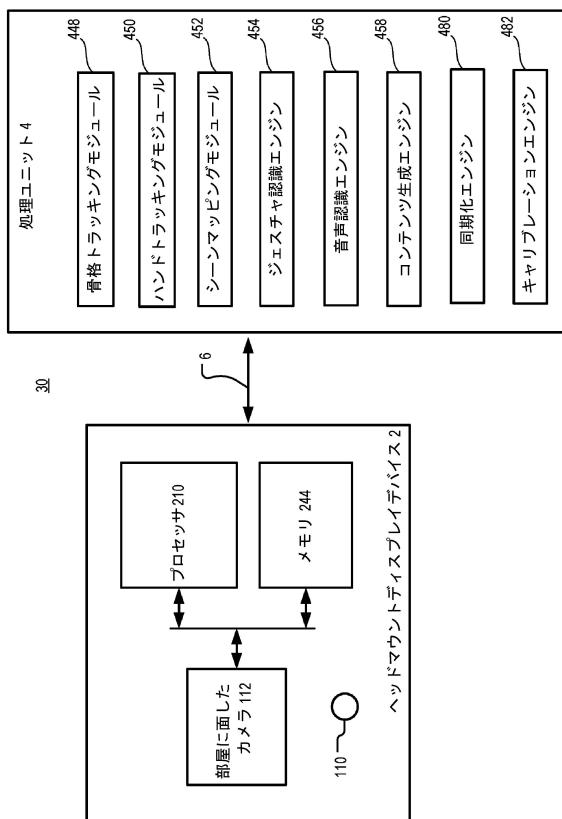
【図6】



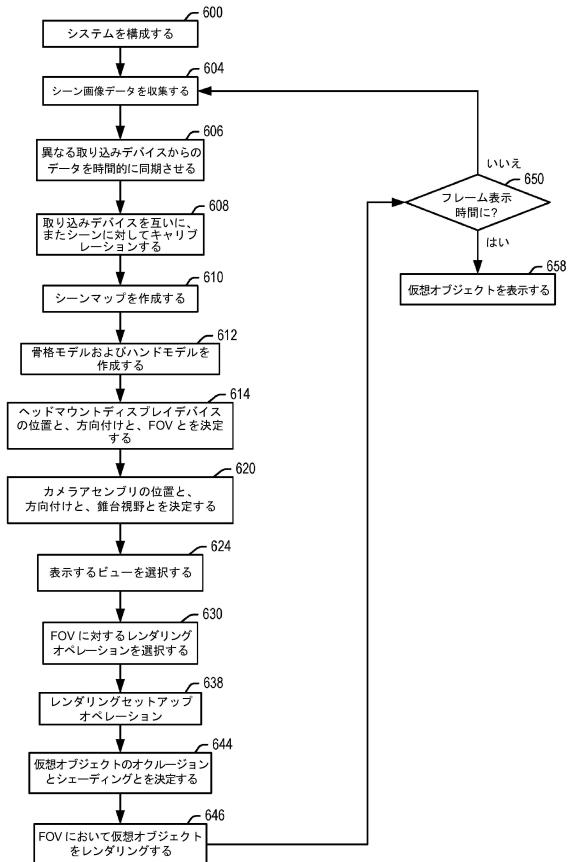
【 四 7 】



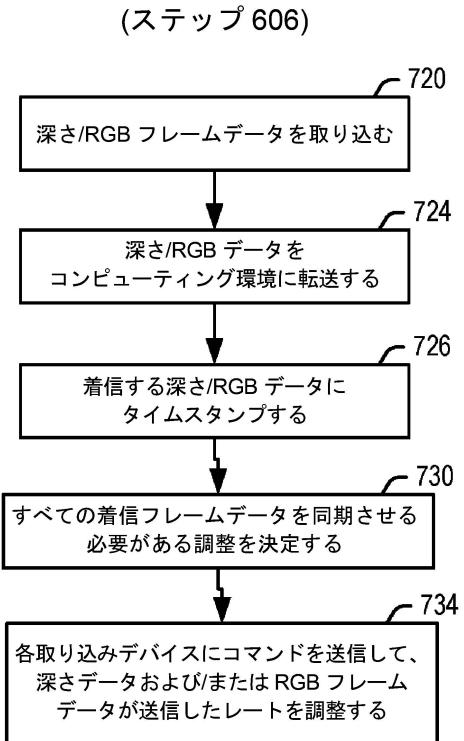
【 図 8 】



【図9】

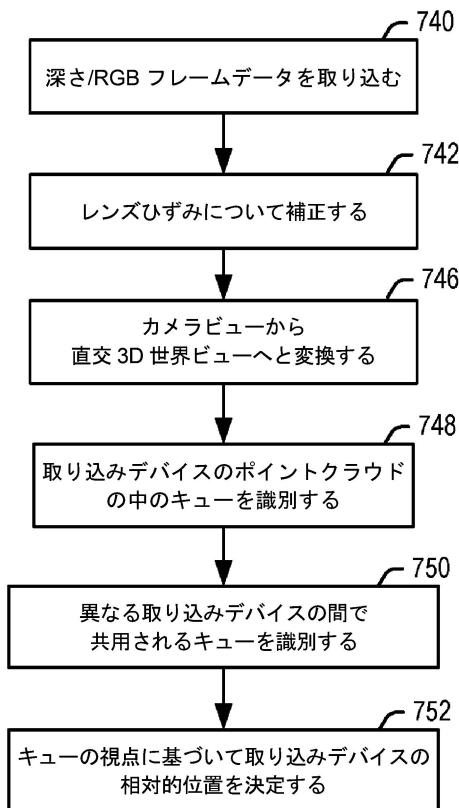


【図10】



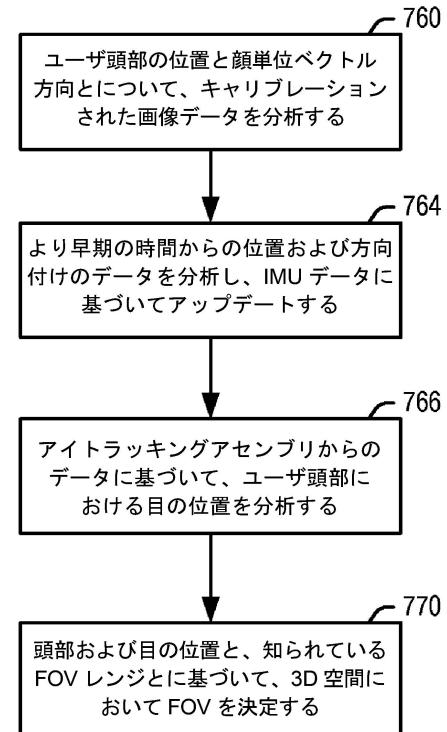
【図11】

(ステップ608)

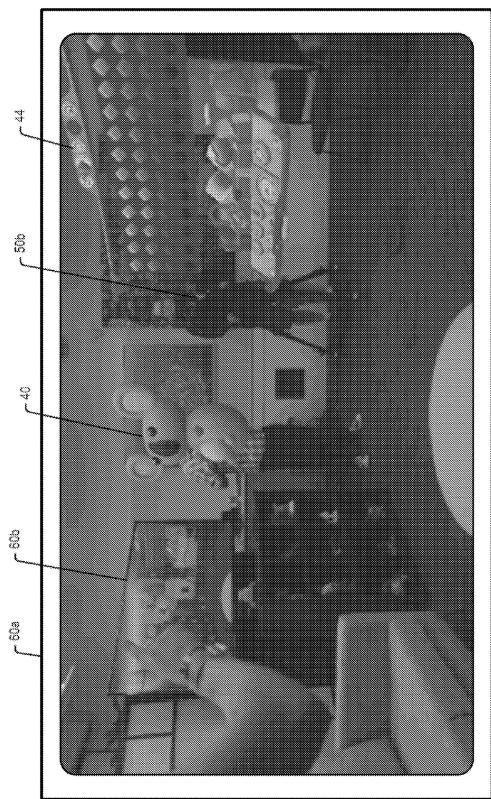


【図12】

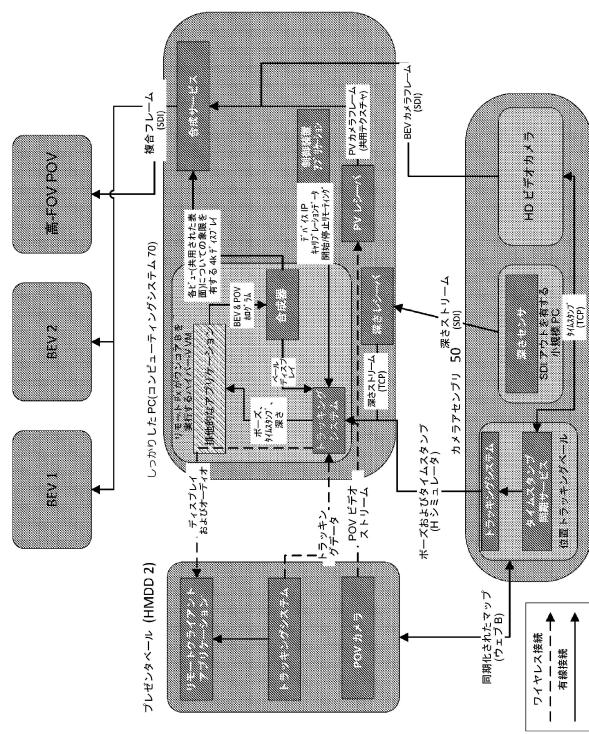
(ステップ614)



【図13】



【図14】



## フロントページの続き

(51)Int.Cl. F I  
H 04N 5/232 939

- (74)代理人 100108213  
弁理士 阿部 豊隆
- (74)代理人 100210125  
弁理士 小宮山 真世
- (72)発明者 トムリン, アーサー チャールズ  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 キーブラー, エヴァン マイケル  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 ファイト, ニコラス ガーバス  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 マウント, ブライアン ジェイ.  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 アルト, グレゴリー ローウェル  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 トサール, ジョージ  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 ライオンズ, ジョナサン マイケル  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 アンプラス, アンソニー ジェイ.  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 エグパート, キャメロン クイン  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 ゲイマン, ウィル  
アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト  
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ  
:パテント グループ ドケティング (ビルディング 8/1000)
- (72)発明者 マクグリン, ジェフ ダブリュー.

アメリカ合衆国，ワシントン州 98052-6399，レッドモンド，ワン マイクロソフト  
ウェイ，マイクロソフト テクノロジー ライセンシング，エルエルシー，エーティーティーエヌ  
：パテント グループ ドケティング (ビルディング 8/1000)

(72)発明者 ハンス，ジェレミー

アメリカ合衆国，ワシントン州 98052-6399，レッドモンド，ワン マイクロソフト  
ウェイ，マイクロソフト テクノロジー ライセンシング，エルエルシー，エーティーティーエヌ  
：パテント グループ ドケティング (ビルディング 8/1000)

(72)発明者 シルヴァン，ロジャー セバスチャン - ケヴィン

アメリカ合衆国，ワシントン州 98052-6399，レッドモンド，ワン マイクロソフト  
ウェイ，マイクロソフト テクノロジー ライセンシング，エルエルシー，エーティーティーエヌ  
：パテント グループ ドケティング (ビルディング 8/1000)

(72)発明者 プファッフェ，アレクサンダー ゲオルク

アメリカ合衆国，ワシントン州 98052-6399，レッドモンド，ワン マイクロソフト  
ウェイ，マイクロソフト テクノロジー ライセンシング，エルエルシー，エーティーティーエヌ  
：パテント グループ ドケティング (ビルディング 8/1000)

(72)発明者 クロイマン，ダン

アメリカ合衆国，ワシントン州 98052-6399，レッドモンド，ワン マイクロソフト  
ウェイ，マイクロソフト テクノロジー ライセンシング，エルエルシー，エーティーティーエヌ  
：パテント グループ ドケティング (ビルディング 8/1000)

(72)発明者 サルトウェル，エリック アンドリュー

アメリカ合衆国，ワシントン州 98052-6399，レッドモンド，ワン マイクロソフト  
ウェイ，マイクロソフト テクノロジー ライセンシング，エルエルシー，エーティーティーエヌ  
：パテント グループ ドケティング (ビルディング 8/1000)

(72)発明者 ワード，クリス

アメリカ合衆国，ワシントン州 98052-6399，レッドモンド，ワン マイクロソフト  
ウェイ，マイクロソフト テクノロジー ライセンシング，エルエルシー，エーティーティーエヌ  
：パテント グループ ドケティング (ビルディング 8/1000)

審査官 田中 幸雄

(56)参考文献 特開2001-195601 (JP, A)

米国特許出願公開第2013/0335405 (US, A1)

堀田亮介ほか, MR-Cyborg Soldiers 2: 観客・プレイヤ協調型複合現実感ゲーム, エンタインメントコンピューティングシンポジウム論文集, 日本, 情報処理学会, 2014年 9月12日, 211-213ページ

(58)調査した分野(Int.Cl., DB名)

G 06 T 19/00

G 06 F 3/01

G 06 F 3/0481

H 04 N 5/222

H 04 N 5/232