

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6961007号
(P6961007)

(45) 発行日 令和3年11月5日(2021.11.5)

(24) 登録日 令和3年10月14日(2021.10.14)

(51) Int.Cl.

F 1

G06F	3/01	(2006.01)	G06F	3/01	510
G06T	19/00	(2011.01)	G06T	19/00	A
G06F	3/16	(2006.01)	G06F	3/16	630
H04S	7/00	(2006.01)	G06F	3/16	620
			G06F	3/16	500

請求項の数 30 (全 43 頁) 最終頁に続く

(21) 出願番号 特願2019-546368 (P2019-546368)
 (86) (22) 出願日 平成30年2月27日 (2018.2.27)
 (65) 公表番号 特表2020-509492 (P2020-509492A)
 (43) 公表日 令和2年3月26日 (2020.3.26)
 (86) 國際出願番号 PCT/US2018/020020
 (87) 國際公開番号 WO2018/160593
 (87) 國際公開日 平成30年9月7日 (2018.9.7)
 審査請求日 令和3年2月26日 (2021.2.26)
 (31) 優先権主張番号 62/464,757
 (32) 優先日 平成29年2月28日 (2017.2.28)
 (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, ウエスト サンライズ
 ブールバード 7500
 7500 W SUNRISE BLVD
 , PLANTATION, FL 33322
 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】複合現実デバイスにおける仮想および実オブジェクトの記録

(57) 【特許請求の範囲】

【請求項 1】

ユーザによる使用のための仮想画像生成システムであって、
 メモリと、
 ディスプレイサブシステムと、
 オブジェクト選択デバイスであって、前記オブジェクト選択デバイスは、入力を前記ユーザから受信し、前記ユーザ入力に応答して、3次元場面内の複数のオブジェクトから少なくとも1つのオブジェクトを選択するために構成され、前記少なくとも1つの選択されたオブジェクトは、実オブジェクトを含む、オブジェクト選択デバイスと、
 制御サブシステムであって、前記制御サブシステムは、少なくとも、

オーディオ捕捉デバイスを用いて、前記複数のオブジェクトのうちの少なくとも1つの選択されていないオブジェクトよりも前記少なくとも1つの選択されたオブジェクトから出るオーディオデータを優先的に入手することと、

少なくとも、第2のメタデータに少なくとも部分的に基づいて第1のメタデータを前記ユーザに対して局所化することによって、前記オーディオデータから空間化オーディオをレンダリングすることであって、前記第1のメタデータは、前記ユーザに対して局所化されて第1の局所化メタデータになり、少なくとも1つの選択されたオブジェクトと関連付けられ、前記仮想画像生成システムによってまたは前記オーディオ捕捉デバイスを有する異なる仮想画像生成システムによって前記オーディオデータとともに生成され、前記第2のメタデータは、前記ユーザのための、前記仮想画像生成システムによって検出される

10

20

位置特性または配向特性に関連する、ことと、

前記第1のメタデータと前記第2のメタデータとに少なくとも部分的に基づいて、レンダリングされた前記空間化オーディオと、前記ユーザに同期する前記少なくとも1つの選択されたオブジェクトを表す少なくとも1つの仮想オブジェクトとを提示することと
のために構成される、制御サブシステムと
を備える、仮想画像生成システム。

【請求項2】

少なくとも画像捕捉デバイスを用いて、前記少なくとも1つの選択されたオブジェクトのための画像データを入手することと、

少なくとも、前記第2のメタデータに部分的にまたは全体的に基づいて前記ユーザに対して前記第1のメタデータを局所化することによって、前記少なくとも1つの選択されたオブジェクトを3次元場面内の前記仮想オブジェクトとして表す、前記少なくとも1つの選択されたオブジェクトの前記画像データを複数の画像フレームとしてレンダリングすることと

10

をさらに含み、

前記第1のメタデータは、前記第2のメタデータに部分的にまたは全体的に基づいて局所化されて前記第1の局所化メタデータになり、前記仮想画像生成システムによってまたは前記画像捕捉デバイスを有する異なる仮想画像生成システムによって検出される異なる位置特性または配向特性に関連する、請求項1に記載の仮想画像生成システム。

【請求項3】

20

前記空間化オーディオおよび前記画像データは、前記空間化オーディオおよび前記仮想オブジェクトが生成される空間環境と異なる空間環境において、少なくとも、前記異なる空間環境において前記少なくとも1つの選択されたオブジェクトを再位置付けされたオブジェクトとして再位置付けすることによって、再生される、請求項2に記載の仮想画像生成システム。

【請求項4】

前記空間化オーディオおよび前記画像データは、さらに少なくとも、前記再位置付けされたオブジェクトに対応する絶対メタデータを入手することによって、異なる空間環境において再生される、請求項3に記載の仮想画像生成システム。

【請求項5】

30

前記空間化オーディオおよび前記画像データは、さらに少なくとも、前記絶対メタデータを局所化して局所化絶対メタデータになることによって、異なる空間環境において再生され、前記絶対メタデータは、前記少なくとも1つの選択されたオブジェクトに、再生ユーザに対応する、請求項4に記載の仮想画像生成システム。

【請求項6】

前記空間化オーディオおよび前記画像データは、さらに少なくとも、前記局所化絶対メタデータに少なくとも部分的に基づいて前記空間化オーディオおよび前記画像データをレンダリングすることによって、異なる空間環境において再生される、請求項5に記載の仮想画像生成システム。

【請求項7】

40

前記ユーザの頭部姿勢を追跡する1つ以上のセンサをさらに備え、前記ユーザの前記頭部姿勢は、前記第1のメタデータおよび前記オーディオデータに沿って、前記第2のメタデータとしてデータベース内に記憶され、前記少なくとも1つの選択されたオブジェクトに対応する前記第1のメタデータは、前記ユーザの前記頭部姿勢を備える前記第2のメタデータに少なくとも部分的に基づいて前記ユーザに前記空間化オーディオをレンダリングするためおよび前記画像データをレンダリングするために前記ユーザに対して局所化される、請求項2に記載の仮想画像生成システム。

【請求項8】

ビデオプロセッサをさらに備え、前記ビデオプロセッサは、データベース内に記憶される前記少なくとも1つの選択されたオブジェクトの前記画像データを入手し、前記画像デ

50

ータを処理することなく前記画像データをレコーダに送信し、前記レコーダは、処理された画像を前記メモリ内に記憶し、前記ディスプレイサブシステムは、投影サブシステムおよび部分的に透明なディスプレイ表面を含み、前記投影サブシステムは、前記複数の画像フレームを前記部分的に透明なディスプレイ表面上へ投影し、前記部分的に透明なディスプレイ表面は、エンドユーザの眼と周囲環境との間の視野内に位置付けられるために構成される、請求項2に記載の仮想画像生成システム。

【請求項9】

ビデオプロセッサおよびフレーム構造をさらに備え、前記フレーム構造は、前記ユーザによって装着されるために、および、前記ディスプレイサブシステムの少なくとも一部を担持するために構成され、前記ビデオプロセッサは、前記画像捕捉デバイスによって捕捉されたビデオデータから前記少なくとも1つの選択されたオブジェクトの前記画像データを抽出し、前記ビデオデータから抽出された前記画像データを前記少なくとも1つの仮想オブジェクトとして前記メモリ内に記憶し、前記画像データをレコーダに送信し、前記レコーダは、前記画像データを前記少なくとも1つの仮想オブジェクトとしてデータベース内に記録する、請求項2に記載の仮想画像生成システム。

10

【請求項10】

複数のスピーカをさらに備え、前記制御サブシステムは、前記空間化オーディオを前記複数のスピーカにさらに伝達する、請求項1に記載の仮想画像生成システム。

【請求項11】

前記ディスプレイサブシステムは、視野を有し、前記オブジェクト選択デバイスは、前記ユーザ入力に応答して、前記視野内の前記少なくとも1つのオブジェクトを選択するための第1の命令を受信する、請求項1に記載の仮想画像生成システム。

20

【請求項12】

前記オブジェクト選択デバイスは、前記ユーザ入力に応答して、3次元カーソルを前記ディスプレイサブシステムの前記視野内で移動させるための第2の命令を受信する、請求項1に記載の仮想画像生成システム。

【請求項13】

前記ユーザ入力は、1つ以上の音声コマンドまたは1つ以上の手のジェスチャを含み、前記オブジェクト選択デバイスは、前記1つ以上の音声コマンドまたは前記1つ以上の手のジェスチャを捕捉する1つ以上のデバイスを備える、請求項1に記載の仮想画像生成システム。

30

【請求項14】

前記仮想画像生成システムは、前記オーディオデータのタイプに少なくとも部分的に基づいて、前記少なくとも1つの選択されたオブジェクトをプロファイル化し、前記オーディオデータのタイプは、少なくとも、1つ以上の実オブジェクトからそれぞれ出る前記オーディオデータの1つ以上の特性に少なくとも部分的に基づいて前記オーディオデータの特性を分析することによって、決定される、請求項1に記載の仮想画像生成システム。

【請求項15】

前記少なくとも1つのオブジェクトは、複数のオブジェクトを含み、前記オブジェクト選択デバイスは、前記複数のオブジェクトのうちの1つのオブジェクトを個々に選択する第1の機能を前記ユーザに提供し、前記ユーザ入力に応答する範囲内で前記複数のオブジェクトのうちの複数のオブジェクトを大域的に選択する第2の機能を前記ユーザに提供し、前記ディスプレイサブシステムによって前記ユーザに提供された視野は、前記範囲を備える、請求項1に記載の仮想画像生成システム。

40

【請求項16】

前記オブジェクト選択デバイスは、前記ユーザから別の入力を受信し、前記別の入力に応答して、前記少なくとも1つの選択されたオブジェクトを選択解除する、請求項1に記載の仮想画像生成システム。

【請求項17】

前記ディスプレイサブシステムによって前記ユーザに提供された視野に対する前記少な

50

くとも 1 つの選択されたオブジェクトの場所を追跡する少なくとも 1 つのセンサをさらに備える、請求項 1 に記載の仮想画像生成システム。

【請求項 18】

前記制御サブシステムは、前記少なくとも 1 つのセンサによって追跡される前記少なくとも 1 つの選択されたオブジェクトの場所が、前記ディスプレイサブシステムによって前記ユーザに提供される前記視野外に移動すると、前記オーディオデータを前記メモリ内に記憶することを引き起こす、請求項 17 に記載の仮想画像生成システム。

【請求項 19】

前記制御サブシステムは、前記少なくとも 1 つのセンサによって追跡される前記少なくとも 1 つの選択されたオブジェクトの場所が、前記ディスプレイサブシステムによって前記ユーザに提供される前記視野外に移動すると、前記オーディオデータを前記メモリ内に記憶することを継続する、請求項 17 に記載の仮想画像生成システム。

10

【請求項 20】

オーディオ捕捉デバイスは、指向性オーディオ出力を生成し、前記制御サブシステムは、前記少なくとも 1 つの選択されていないオブジェクトよりも前記少なくとも 1 つの選択されたオブジェクトの方向に、前記オーディオ捕捉デバイスによって受信された前記オーディオデータを優先的に表すように、前記指向性オーディオ出力をさらに修正し、前記空間化オーディオは、前記制御サブシステムによって修正された前記指向性オーディオ出力から導出される、請求項 1 に記載の仮想画像生成システム。

【請求項 21】

20

前記少なくとも 1 つの選択されたオブジェクトに関連する画像データを捕捉する 1 つ以上の画像捕捉デバイスをさらに備え、前記制御サブシステムは、前記オーディオデータと同期する前記画像データを前記メモリ内に記憶し、前記少なくとも 1 つの選択されたオブジェクトは、カーソルを使用することなく前記ユーザによって選択される、請求項 1 に記載の仮想画像生成システム。

【請求項 22】

前記制御サブシステムは、前記メモリ内に記憶された前記画像データを前記少なくとも 1 つの選択されたオブジェクトのための仮想コンテンツデータに変換し、前記仮想コンテンツデータを前記メモリ内にさらに記憶する、請求項 21 に記載の仮想画像生成システム。

30

【請求項 23】

前記少なくとも 1 つの選択されたオブジェクトは、前記仮想画像生成システムによってレンダリングされた仮想オブジェクトを含む、請求項 1 に記載の仮想画像生成システム。

【請求項 24】

複数の仮想オブジェクトによって生成された前記オーディオデータを記憶するデータベースをさらに備え、前記制御サブシステムは、選択されていない仮想オブジェクトよりも選択された仮想オブジェクトに対応する前記オーディオデータを前記データベースから優先的に入手するための命令を受信し、前記データベースは、前記仮想オブジェクトの説明を記憶し、前記データベースに記憶される前記仮想オブジェクトの前記説明は、前記ユーザによって表される前記仮想オブジェクトの口頭説明に合致される、請求項 1 に記載の仮想画像生成システム。

40

【請求項 25】

前記制御サブシステムは、前記少なくとも 1 つの選択されたオブジェクトに対応する前記第 1 のメタデータをさらに生成し、前記メモリ内に記憶される前記オーディオデータは、前記オーディオデータ内のコンテンツおよび前記第 1 のメタデータを含む、請求項 1 に記載の仮想画像生成システム。

【請求項 26】

前記第 1 のメタデータは、前記少なくとも 1 つの選択されたオブジェクトのための絶対位置および絶対配向に関するデータを含む、請求項 1 に記載の仮想画像生成システム。

【請求項 27】

50

少なくとも 1 つのスピーカをさらに備え、前記仮想画像生成システムは、前記空間化オーディオをレンダリングすることによって、前記少なくとも 1 つの選択されたオブジェクトのみから入手された前記オーディオデータを前記空間化オーディオにさらに変換し、前記オーディオデータを前記空間化オーディオに変換することは、前記メモリ内に記憶された前記オーディオデータを読み出すことと、前記メモリから読み出された前記オーディオデータから空間化オーディオを導出することと、前記空間化オーディオを前記少なくとも 1 つのスピーカに伝達することと、前記実オブジェクトから生じる音と前記空間化オーディオを混合することとを含む、請求項 1 に記載の仮想画像生成システム。

【請求項 2 8】

レコーダをさらに備え、前記レコーダは、第 1 のオーディオデータを入手し、前記第 1 のオーディオデータは、前記空間化オーディオ、および、前記少なくとも 1 つの選択されたオブジェクトに対応する前記第 1 のメタデータまたは前記第 2 のメタデータを含み、前記レコーダは、第 1 のビデオデータをさらに入手し、前記第 1 のビデオデータは、前記第 1 のビデオデータの画像フレーム、および、少なくとも前記少なくとも 1 つの選択されたオブジェクトに関連する対応するメタデータを含み、前記レコーダは、異なるコンピューティングデバイス上のプレーヤにおける後続再生のために前記第 1 のオーディオデータおよび前記第 1 のビデオデータをさらに記憶する、請求項 2 7 に記載の仮想画像生成システム。

10

【請求項 2 9】

前記レコーダは、記録し、第 1 の仮想オブジェクトからの第 1 のオーディオおよび第 1 の実オブジェクトからの第 2 のオーディオを記録しないように構成され、前記空間化オーディオは、少なくとも、前記第 1 の局所化メタデータおよび大域的メタデータを前記ユーザのために前記オーディオデータに適用することによって、前記オーディオデータからレンダリングされる、請求項 2 8 に記載の仮想画像生成システム。

20

【請求項 3 0】

少なくとも、前記第 2 のメタデータに少なくとも部分的に基づいて前記ユーザに対して前記第 1 のメタデータではなく絶対メタデータを局所化することによって、前記仮想オブジェクトとして前記少なくとも 1 つの選択されたオブジェクトを表す複数の画像フレームとして、前記少なくとも 1 つの選択されたオブジェクトのための少なくとも画像捕捉デバイスによって入手された画像データをレンダリングすることをさらに含み、前記少なくとも 1 つの選択されたオブジェクトは、前記仮想画像生成システムの前記ユーザまたは前記異なる仮想画像生成システムの異なるユーザが注視する方向にかかわらず、持続的に選択されたままである、請求項 1 に記載の仮想画像生成システム。

30

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、概して、仮想現実および拡張現実システムに関する。

【背景技術】

【0 0 0 2】

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」または「拡張現実」体験のための複合現実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、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過性を伴わずに、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（すなわち、他の実際の実世界視覚的入力に対して透過性である）。故に、AR シナリオは、他の実際の実世界視覚的入力に対する透過性を伴って、デジタルまたは仮想画像情報の提示を伴う。

40

【0 0 0 3】

例えば、図 1 を参照すると、拡張現実場面 4 が、描写されており、AR 技術のユーザに

50

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8を特徴とする、実世界公園状設定6が見える。これらのアイテムに加え、AR技術のエンドユーザはまた、実世界プラットフォーム8上に立っているロボット像10、および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12を「見ている」と知覚するが、これらの要素10、12は、実世界には存在しない。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またはAR技術の生産は、困難である。

【0004】

VRおよびARシステムは、典型的には、少なくともユーザの頭部に緩く結合され、したがって、エンドユーザの頭部が移動すると移動する、頭部装着型ディスプレイ（またはヘルメット搭載型ディスプレイ、またはスマートグラス）を採用する。エンドユーザの頭部運動が、ディスプレイシステムによって検出される場合、表示されているデータは、頭部姿勢（すなわち、ユーザの頭部の配向および／または場所）の変化を考慮するように更新されることができる。AR（すなわち、仮想および実オブジェクトの同時視認）を有効にする、頭部装着型ディスプレイは、いくつかの異なるタイプの構成を有することができる。多くの場合、「ビデオシースルーディスプレイ」と称される、1つのそのような構成では、カメラが、実場面の要素を捕捉し、コンピューティングシステムが、仮想要素を捕捉された実場面上に重畠し、不透明ディスプレイが、合成画像を眼に提示する。別の構成は、多くの場合、「光学シースルーディスプレイ」と称され、エンドユーザには、ディスプレイシステム内の透明（または半透明）要素を通して、環境内の実オブジェクトからの光が透けて見え、それを直接視認することができる。多くの場合、「コンバイナ」と称される、透明要素は、ディスプレイからの光を実世界のエンドユーザのビューにわたって重畠する。

【0005】

多くの場合、VR/ARシステムのユーザは、後にオンラインで公開するために、体験をVR/ARシステム上に記録および保存することによって、その体験を他者と共有することを所望し得る（例えば、ゲームをプレーする、テレビ会議を行う、または映画を鑑賞するとき）。しかしながら、典型的には、雑音環境に起因して、雑音および他の望ましくないまたは予期しない音が記録内に存在し得る、またはそれ以外の体験に対して注意が逸らされる、非常に多くの音源が存在し得る。そのような望ましくない／予期しない音は、実オブジェクトから、例えば、VR/ARシステムの近傍で遊んでいる子供から、または仮想オブジェクトから、例えば、VR/ARシステムのコンテキストにおいて再生中の仮想テレビから生じ得る。

【0006】

したがって、ユーザが関心がある仮想または実オブジェクトのみからの音を記録するための単純および効率的手段を提供する必要が残ったままである。

【発明の概要】

【課題を解決するための手段】

【0007】

本発明の第1の側面によると、エンドユーザによる使用のための仮想画像生成システムは、メモリと、ディスプレイサブシステムと、入力をエンドユーザから受信し、エンドユーザ入力に応答して、少なくとも1つのオブジェクト（例えば、実オブジェクトおよび／または仮想オブジェクト）を持続的に選択するために構成される、オブジェクト選択デバイスとを備える。一実施形態では、ディスプレイサブシステムは、視野を有し、オブジェクト選択デバイスは、視野内のオブジェクトを持続的に選択するために構成される。この場合、オブジェクト選択デバイスは、エンドユーザ入力の受信に応答して、3次元カーソルをディスプレイサブシステムの視野内で移動させ、オブジェクトを選択するために構成されてもよい。別の実施形態では、エンドユーザ入力は、1つ以上の音声コマンドを備え、オブジェクト選択デバイスは、音声コマンドを感知するために構成される、1つ以上の

10

20

30

40

50

マイクロホンを備える。さらに別の実施形態では、エンドユーザ入力は、1つ以上の手のジェスチャを備え、その場合、オブジェクト選択デバイスは、手のジェスチャを感知するために構成される、1つ以上のカメラを備えてもよい。

【0008】

複数のオブジェクトが選択される場合、オブジェクト選択デバイスは、エンドユーザ入力に応答して、オブジェクトを個々に選択および／または大域的に選択するために構成されてもよい。大域的に選択される場合、オブジェクト選択デバイスは、エンドユーザ入力に応答して、視野の角度範囲（視野の角度範囲全体未満であってもよい、または視野の角度範囲全体であってもよい）内の全てのオブジェクトを大域的に選択するために構成されてもよい。一実施形態では、オブジェクト選択デバイスはさらに、他のエンドユーザ入力に応答して、別の入力をエンドユーザから受信し、以前に選択されたオブジェクトを持続的に選択解除するために構成される。

10

【0009】

仮想画像生成システムはさらに、少なくとも1つの選択されたオブジェクトから生じるビデオデータを生成し、複数の画像フレームをビデオデータからの3次元場面内にレンダリングし、画像フレームをディスプレイサブシステムに伝達するために構成される、制御サブシステムを備える。一実施形態では、ディスプレイサブシステムは、エンドユーザの眼の正面に位置付けられるために構成される。別の実施形態では、ディスプレイサブシステムは、投影サブシステムおよび部分的に透明なディスプレイ表面を含む。この場合、投影サブシステムは、画像フレームを部分的に透明なディスプレイ表面上に投影させるために構成されてもよく、部分的に透明なディスプレイ表面は、エンドユーザの眼と周囲環境との間の視野内に位置付けられるために構成されてもよい。仮想画像生成システムはさらに、エンドユーザによって装着するために構成され、ディスプレイサブシステムの少なくとも一部を担持する、フレーム構造を備えてもよい。

20

【0010】

制御サブシステムはさらに、選択されたオブジェクトから生じるオーディオデータを生成することと、オーディオデータをメモリ内に記憶することとのために構成される。仮想画像生成システムはさらに、複数のスピーカを備えてもよく、その場合、制御サブシステムはさらに、生成されたオーディオデータをスピーカに伝達するために構成されてもよい。随意の実施形態では、制御サブシステムはさらに、オーディオデータと同期するビデオデータをメモリ内に記憶するために構成される。さらに別の実施形態では、仮想画像生成システムはさらに、ディスプレイサブシステムの視野に対する選択されたオブジェクトの場所を追跡するために構成される、少なくとも1つのセンサを備える。この場合、制御サブシステムは、選択されたオブジェクトの追跡された場所がディスプレイサブシステムの視野外に移動すると、オーディオデータをメモリ内に記憶することを中止するために構成される、または代替として、選択されたオブジェクトの追跡された場所がディスプレイサブシステムの視野外に移動すると、オーディオデータをメモリ内に記憶することを継続するために構成されてもよい。

30

【0011】

選択されたオブジェクトが、実オブジェクトを含む場合、仮想画像生成システムはさらに、オーディオ出力を生成するために構成される、マイクロホンアセンブリを備えてもよく、その場合、制御サブシステムはさらに、オーディオ出力の方向を修正し、選択された実オブジェクトから生じる音を優先的に感知するために構成されてもよい。オーディオデータは、修正されたオーディオ出力から導出されてもよい。仮想画像生成システムはさらに、選択された実オブジェクトから生じるビデオデータを捕捉するために構成される、1つ以上のカメラを備えてもよく、その場合、制御サブシステムはさらに、オーディオデータと同期するビデオデータをメモリ内に記憶するために構成されてもよい。制御サブシステムは、捕捉されたビデオデータを選択された実オブジェクトのための仮想コンテンツデータに変換し、仮想コンテンツをメモリ内に記憶するために構成されてもよい。

40

【0012】

50

選択されたオブジェクトが、仮想オブジェクトを含む場合、仮想画像生成システムはさらに、複数の仮想オブジェクトのための音に対応するコンテンツデータを記憶するために構成される、データベースを備えてもよく、その場合、制御サブシステムはさらに、選択された仮想オブジェクトに対応するコンテンツデータをデータベースから入手するために構成されてもよく、メモリ内に記憶されるオーディオデータは、入手されたコンテンツデータを含む。制御サブシステムはさらに、選択された仮想オブジェクトに対応するメタデータ（例えば、選択された仮想オブジェクトのための位置、配向、および音量データ）を生成するために構成されてもよく、その場合、メモリ内に記憶されるオーディオデータは、入手されたコンテンツデータおよび生成されたメタデータを含んでもよい。一実施形態では、仮想画像生成システムはさらに、エンドユーザの頭部姿勢を追跡するために構成される、1つ以上のセンサを備え、その場合、データベースは、複数の仮想オブジェクトのための絶対メタデータを記憶するために構成されてもよく、制御サブシステムはさらに、選択された仮想オブジェクトに対応する絶対メタデータを入手し、エンドユーザの追跡された頭部姿勢に基づいて、絶対メタデータをエンドユーザに対して局所化することによって、メタデータを生成するために構成されてもよい。10

【0013】

仮想画像生成システムはさらに、少なくとも1つのスピーカを備えてもよく、その場合、制御サブシステムはさらに、記憶されたオーディオデータをメモリから読み出し、オーディオを読み出されたオーディオデータから導出し、オーディオをスピーカに伝達するために構成されてもよい。メモリ内に記憶されるオーディオデータは、コンテンツデータおよびメタデータを含んでもよく、その場合、制御サブシステムはさらに、記憶されたコンテンツデータおよびメタデータをメモリから読み出し、読み出されたコンテンツデータおよびメタデータに基づいて、空間化オーディオをレンダリングし、レンダリングされた空間化オーディオをスピーカに伝達するために構成されてもよい。20

【0014】

本発明の第2の側面によると、エンドユーザによって仮想画像生成システムを動作させる方法が、提供される。本方法は、少なくとも1つのオブジェクト（例えば、実オブジェクトおよび/または仮想オブジェクト）を持続的に選択するステップを含む。1つの方法では、オブジェクトを選択するステップは、3次元カーソルをエンドユーザの視野内で移動させ、3次元カーソルを用いてオブジェクトを選択するステップを含む。別 の方法では、オブジェクトを選択するステップは、1つ以上の音声コマンドを発行するステップを含む。さらに別の方法では、少なくとも1つのオブジェクトを選択するステップは、1つ以上の手のジェスチャを行うステップを含む。複数のオブジェクトが、選択される場合、複数のオブジェクトを選択するステップは、オブジェクトを個々に選択するステップおよび/またはオブジェクトを大域的に選択するステップを含んでもよい。大域的に選択される場合、オブジェクトは、エンドユーザの視野の角度範囲を定義し（視野の角度範囲全体未満であってもよい、または視野の角度範囲全体であってもよい）、エンドユーザの視野の定義された角度範囲内のオブジェクトの全てを選択することによって、選択されてもよい。隨意の方法はさらに、以前に選択されたオブジェクトを持続的に選択解除するステップを含んでもよい。30

【0015】

本方法はさらに、選択されたオブジェクトから生じるビデオデータを生成するステップと、複数の画像フレームを生成されたビデオデータからの3次元場面内にレンダリングするステップと、画像フレームをエンドユーザに表示するステップと、少なくとも1つの選択されたオブジェクトから生じるオーディオデータを生成するステップと、少なくとも1つの選択されたオブジェクトから生じるオーディオデータをメモリ内に記憶するステップとを含む。1つの方法はさらに、選択されたオブジェクトから生じるオーディオデータをエンドユーザによる知覚のための音に変換するステップを含んでもよい。本方法は、隨意に、オーディオデータと同期するビデオデータをメモリ内に記憶するステップを含んでもよい。さらに別の方法はさらに、エンドユーザの視野に対する選択されたオブジェクトの40

場所を追跡するステップを含んでもよい。この場合、本方法はさらに、選択されたオブジェクトの追跡された場所がエンドユーザの視野外に移動すると、オーディオデータをメモリ内に記憶することを中止するステップ、または代替として、選択されたオブジェクトの追跡された場所がエンドユーザの視野外に移動すると、オーディオデータをメモリ内に記憶することを継続するステップを含んでもよい。

【0016】

選択されたオブジェクトが、実オブジェクトを含む場合、本方法はさらに、他の実オブジェクトから生じる音と比較して、選択された実オブジェクトから生じる音を優先的に感知するステップを含んでもよく、その場合、オーディオデータは、優先的に感知される音から導出されてもよい。本方法はさらに、選択された実オブジェクトから生じるビデオデータを捕捉し、オーディオデータと同期するビデオデータをメモリ内に記憶するステップを含んでもよい。捕捉されたビデオデータは、メモリ内への記憶のために、仮想コンテンツデータに変換されてもよい。

10

【0017】

選択されたオブジェクトが、仮想オブジェクトを含む場合、本方法はさらに、複数の仮想オブジェクトのための音に対応するコンテンツデータを記憶し、選択された仮想オブジェクトに対応するコンテンツデータを入手するステップを含んでもよく、その場合、メモリ内に記憶されるオーディオデータは、入手されたコンテンツデータを含んでもよい。本方法はさらに、選択された仮想オブジェクトに対応するメタデータ（例えば、選択された仮想オブジェクトのための位置、配向、および音量データ）を生成するステップを含んでもよく、その場合、メモリ内に記憶されるオーディオデータは、入手されたコンテンツデータおよび生成されたメタデータを含んでもよい。本方法はさらに、エンドユーザの頭部姿勢を追跡するステップと、複数の仮想オブジェクトのための絶対メタデータを記憶するステップとを含んでもよい。この場合、メタデータを生成するステップは、選択された仮想オブジェクトに対応する絶対メタデータを読み出し、エンドユーザの追跡された頭部姿勢に基づいて、絶対メタデータをエンドユーザに対して局所化するステップを含んでもよい。

20

【0018】

本方法はさらに、記憶されたオーディオデータを読み出すステップと、オーディオを読み出されたオーディオデータから導出するステップと、オーディオをエンドユーザによる知覚のための音に変換するステップとを含んでもよい。記憶されたオーディオデータは、コンテンツデータおよびメタデータを含んでもよく、その場合、本方法はさらに、記憶されたコンテンツデータおよびメタデータをメモリから読み出すステップと、読み出されたコンテンツデータおよびメタデータに基づいて、空間化オーディオをレンダリングするステップと、空間化オーディオをエンドユーザによる知覚のための音に変換するステップとを含んでもよい。

30

【0019】

本発明の第3の側面によると、再生ユーザによる使用のための仮想画像生成システムが、提供される。仮想画像生成システムは、オリジナル空間環境内の少なくとも1つのオブジェクト（例えば、実オブジェクトおよび/または仮想オブジェクト）から生じるオーディオコンテンツデータおよびビデオコンテンツデータを記憶するために構成される、メモリと、複数のスピーカと、ディスプレイサブシステムとを備える。一実施形態では、ディスプレイサブシステムは、エンドユーザの眼の正面に位置付けられるために構成される。別の実施形態では、ディスプレイサブシステムは、投影サブシステムおよび部分的に透明なディスプレイ表面を含む。この場合、投影サブシステムは、画像フレームを部分的に透明なディスプレイ表面上に投影するために構成されてもよく、部分的に透明なディスプレイ表面は、エンドユーザの眼と周囲環境との間の視野内に位置付けられるために構成されてもよい。仮想画像生成システムはさらに、エンドユーザによって装着するために構成され、ディスプレイサブシステムの少なくとも一部を担持する、フレーム構造を備えてよい。

40

50

【 0 0 2 0 】

仮想画像生成システムはさらに、オーディオコンテンツデータおよびビデオコンテンツデータをメモリから読み出し、読み出されたオーディオコンテンツデータおよびビデオコンテンツデータからそれぞれオーディオおよびビデオをオリジナル空間環境と異なる新しい空間環境内にレンダリングし、レンダリングされたオーディオをスピーカに、生成されたビデオデータをディスプレイサブシステムに同期して伝達するために構成される、制御サブシステムを備える。

【 0 0 2 1 】

一実施形態では、制御サブシステムは、オーディオコンテンツデータおよびビデオコンテンツデータをメモリ内に記憶するために構成される。仮想画像生成システムはさらに、10
入力をエンドユーザから受信し、メモリ内へのオーディオコンテンツデータおよびビデオコンテンツデータの記憶に先立って、エンドユーザ入力に応答して、オリジナル空間環境内のオブジェクトを持続的に選択するために構成される、オブジェクト選択デバイスを備えてよい。

【 0 0 2 2 】

オブジェクトが、実オブジェクトを含む場合、仮想画像生成システムはさらに、オーディオコンテンツデータをオリジナル空間環境内の実オブジェクトから捕捉するために構成される、マイクロホンアセンブリを備えてよい。マイクロホンアセンブリは、オーディオ出力を生成するために構成されてもよく、その場合、制御サブシステムはさらに、オーディオ出力の方向を修正し、選択された実オブジェクトから生じる音を優先的に感知するために構成されてもよい。オーディオコンテンツデータは、修正されたオーディオ出力から導出されてもよい。仮想画像生成システムはさらに、ビデオデータをオリジナル空間環境内の選択された実オブジェクトから捕捉するために構成される、1つ以上のカメラを備えてよい。随意の実施形態では、制御サブシステムは、捕捉されたビデオデータを選択された実オブジェクトのための仮想コンテンツデータに変換し、仮想コンテンツデータをビデオコンテンツデータとしてメモリ内に記憶するために構成されてもよい。

【 0 0 2 3 】

オブジェクトが、仮想オブジェクトを含む場合、仮想画像生成システムはさらに、複数の仮想オブジェクトのための音に対応するコンテンツデータを記憶するために構成される、データベースを備えててもよく、その場合、制御サブシステムはさらに、仮想オブジェクトに対応するコンテンツデータをデータベースから入手するために構成されてもよく、メモリ内に記憶されるオーディオデータは、入手されたコンテンツデータを含んでもよい。

【 0 0 2 4 】

一実施形態では、制御サブシステムは、新しい空間環境内の少なくとも1つのオブジェクトに対応する絶対メタデータを入手し、読み出されたオーディオコンテンツデータおよび絶対メタデータからのオーディオを新しい空間環境内にレンダリングするために構成される。新しい空間環境内のオブジェクトに対応する絶対メタデータを入手することは、オブジェクトを新しい空間環境内に位置付けることを含んでもよい。この場合、仮想画像生成システムはさらに、入力を再生ユーザから受信するために構成される、ユーザ入力デバイスを備えててもよく、その場合、制御サブシステムは、再生ユーザからの入力に応答して、オブジェクトを新しい空間環境内に位置付けるために構成されてもよい。仮想画像生成システムはさらに、再生ユーザの頭部姿勢を追跡するために構成される、1つ以上のセンサを備えててもよく、その場合、制御サブシステムはさらに、レンダリングされたオーディオが空間化されるように、再生ユーザの追跡された頭部姿勢に基づいて、絶対メタデータを再生ユーザに対して局所化するために構成されてもよい。

【 0 0 2 5 】

本発明の第4の側面によると、オリジナル空間環境内でオーディオコンテンツデータおよびビデオコンテンツデータとして以前に記録された少なくとも1つのオブジェクト（例えば、実オブジェクトおよび/または仮想オブジェクト）のオーディオおよびビデオを再生するために、再生ユーザによって仮想画像生成システムを動作させる方法が、提供され

10

20

30

40

50

る。本方法は、オーディオコンテンツデータおよびビデオコンテンツデータをメモリから読み出すステップを含む。1つの方法はさらに、オーディオコンテンツデータおよびビデオコンテンツデータをメモリ内に記憶するステップを含む。この場合、本方法はさらに、メモリ内へのオーディオコンテンツデータおよびビデオコンテンツデータの記憶に先立って、オリジナル空間環境内のオブジェクトを持続的に選択するステップを含んでもよい。

【0026】

オブジェクトが、実オブジェクトを含む場合、本方法はさらに、オーディオコンテンツデータを実オブジェクトから捕捉するステップを含んでもよい。この場合、本方法はさらに、他の実オブジェクトから生じる音と比較して、選択された実オブジェクトから生じる音を優先的に感知するステップを含んでもよい。オーディオコンテンツデータは、優先的に感知される音から導出される。本方法はさらに、ビデオデータを選択された実オブジェクトから捕捉するステップと、捕捉されたビデオデータを仮想コンテンツデータに変換するステップとを含んでもよい。オブジェクトが、仮想オブジェクトを含む場合、本方法はさらに、複数の仮想オブジェクトのための音に対応するコンテンツデータを記憶するステップと、仮想オブジェクトに対応するコンテンツデータをデータベースから入手するステップとを含んでもよい。メモリ内に記憶されるオーディオコンテンツデータは、入手されたコンテンツデータを含んでもよい。

【0027】

本方法はさらに、読み出されたオーディオコンテンツデータおよびビデオコンテンツデータからそれぞれオーディオおよびビデオをオリジナル空間環境と異なる新しい空間環境内にレンダリングし、オーディオおよびビデオをそれぞれ音および画像フレームに変換するステップと、音および画像フレームを再生ユーザに同期して伝達するステップとを含む。1つの方法はさらに、新しい空間環境内のオブジェクトに対応する絶対メタデータ入手するステップを含み、その場合、オーディオは、読み出されたオーディオコンテンツデータおよび絶対メタデータから新しい空間環境内にレンダリングされる。本方法はさらに、再生ユーザの頭部姿勢を追跡するステップと、再生ユーザの追跡された頭部姿勢に基づいて、絶対メタデータを再生ユーザに対して局所化するステップとを含んでもよく、その場合、オーディオは、レンダリングされたオーディオが空間化されるように、読み出されたオーディオコンテンツデータおよび局所的メタデータから新しい空間環境内にレンダリングされてもよい。新しい空間環境内のオブジェクトに対応する絶対メタデータ入手するステップは、例えば、再生ユーザからの入力に応答して、オブジェクトを新しい空間環境内に位置付けるステップを含んでもよい。

【0028】

本発明の付加的および他の目的、特徴、および利点は、詳細な説明、図、および請求項に説明される。

本発明は、例えば、以下を提供する。

(項目1)

エンドユーザによる使用のための仮想画像生成システムであって、

メモリと、

ディスプレイサブシステムと、

オブジェクト選択デバイスであって、前記オブジェクト選択デバイスは、入力を前記エンドユーザから受信し、エンドユーザ入力に応答して、少なくとも1つのオブジェクトを持続的に選択するために構成される、オブジェクト選択デバイスと、

制御サブシステムであって、前記制御サブシステムは、少なくとも1つの選択されたオブジェクトから生じるビデオデータを生成し、複数の画像フレームを前記ビデオデータからの3次元場面内にレンダリングし、前記画像フレームを前記ディスプレイサブシステムに伝達し、前記少なくとも1つの選択されたオブジェクトから生じるオーディオデータを生成することと、前記オーディオデータを前記メモリ内に記憶することとのために構成される、制御サブシステムと

を備える、仮想画像生成システム。

10

20

30

40

50

(項目 2)

前記制御サブシステムはさらに、前記オーディオデータと同期する前記ビデオデータを前記メモリ内に記憶するために構成される、項目1に記載の仮想画像生成システム。

(項目 3)

複数のスピーカをさらに備え、前記制御サブシステムはさらに、前記生成されたオーディオデータを前記スピーカに伝達するために構成される、項目1に記載の仮想画像生成システム。

(項目 4)

前記ディスプレイサブシステムは、視野を有し、前記オブジェクト選択デバイスは、前記視野内の前記少なくとも1つのオブジェクトを持続的に選択するために構成される、項目1に記載の仮想画像生成システム。

10

(項目 5)

前記オブジェクト選択デバイスは、3次元カーソルを前記ディスプレイサブシステムの視野内で移動させ、前記エンドユーザ入力の受信に応答して、前記少なくとも1つのオブジェクトを選択するために構成される、項目4に記載の仮想画像生成システム。

(項目 6)

前記エンドユーザ入力は、1つ以上の音声コマンドを含み、前記オブジェクト選択デバイスは、前記1つ以上の音声コマンドを感知するために構成される1つ以上のマイクロホンを備える、項目1に記載の仮想画像生成システム。

20

(項目 7)

前記エンドユーザ入力は、1つ以上の手のジェスチャを含み、前記オブジェクト選択デバイスは、前記1つ以上の手のジェスチャを感知するために構成される1つ以上のカメラを備える、項目1に記載の仮想画像生成システム。

(項目 8)

前記少なくとも1つのオブジェクトは、複数のオブジェクトを含み、前記オブジェクト選択デバイスは、前記エンドユーザ入力に応答して、前記オブジェクトを個々に選択するために構成される、項目1に記載の仮想画像生成システム。

(項目 9)

前記少なくとも1つのオブジェクトは、複数のオブジェクトを含み、前記オブジェクト選択デバイスは、前記エンドユーザ入力に応答して、前記オブジェクトを大域的に選択するために構成される、項目1に記載の仮想画像生成システム。

30

(項目 10)

前記オブジェクト選択デバイスは、前記エンドユーザ入力に応答して、前記視野の角度範囲内の全てのオブジェクトを大域的に選択するために構成される、項目9に記載の仮想画像生成システム。

(項目 11)

前記角度範囲は、前記視野の角度範囲全体未満である、項目10に記載の仮想画像生成システム。

(項目 12)

前記角度範囲は、前記視野の角度範囲全体である、項目10に記載の仮想画像生成システム。

40

(項目 13)

前記オブジェクト選択デバイスはさらに、別の入力を前記エンドユーザから受信し、他のエンドユーザ入力に応答して、少なくとも1つの以前に選択されたオブジェクトを持続的に選択解除するために構成される、項目1に記載の仮想画像生成システム。

(項目 14)

前記ディスプレイサブシステムの視野に対する前記少なくとも1つの選択されたオブジェクトの場所を追跡するために構成される少なくとも1つのセンサをさらに備える、項目1に記載の仮想画像生成システム。

(項目 15)

50

前記制御サブシステムは、前記少なくとも1つの選択されたオブジェクトの追跡された場所が前記ディスプレイサブシステムの視野外に移動すると、前記オーディオデータを前記メモリ内に記憶することを中止するために構成される、項目14に記載の仮想画像生成システム。

(項目16)

前記制御サブシステムは、前記少なくとも1つの選択されたオブジェクトの追跡された場所が前記ディスプレイサブシステムの視野外に移動すると、前記オーディオデータを前記メモリ内に記憶することを継続するために構成される、項目14に記載の仮想画像生成システム。

(項目17)

前記少なくとも1つの選択されたオブジェクトは、実オブジェクトを含む、項目1に記載の仮想画像生成システム。

(項目18)

オーディオ出力を生成するために構成されるマイクロホンアセンブリをさらに備え、前記制御サブシステムはさらに、オーディオ出力の方向を修正し、前記選択された実オブジェクトから生じる音を優先的に感知するために構成され、前記オーディオデータは、前記修正されたオーディオ出力から導出される、項目17に記載の仮想画像生成システム。

(項目19)

前記選択された実オブジェクトから生じるビデオデータを捕捉するために構成される1つ以上のカメラをさらに備え、前記制御サブシステムはさらに、前記オーディオデータと同期する前記ビデオデータを前記メモリ内に記憶するために構成される、項目17に記載の仮想画像生成システム。

(項目20)

前記制御サブシステムは、前記捕捉されたビデオデータを前記選択された実オブジェクトのための仮想コンテンツデータに変換し、前記仮想コンテンツを前記メモリ内に記憶するために構成される、項目19に記載の仮想画像生成システム。

(項目21)

前記少なくとも1つの選択されたオブジェクトは、仮想オブジェクトを含む、項目1に記載の仮想画像生成システム。

(項目22)

複数の仮想オブジェクトのための音に対応するコンテンツデータを記憶するために構成されるデータベースをさらに備え、前記制御サブシステムはさらに、前記選択された仮想オブジェクトに対応するコンテンツデータを前記データベースから入手するために構成され、前記メモリ内に記憶されるオーディオデータは、前記入手されたコンテンツデータを含む、項目21に記載の仮想画像生成システム。

(項目23)

前記制御サブシステムはさらに、前記選択された仮想オブジェクトに対応するメタデータを生成するために構成され、前記メモリ内に記憶されるオーディオデータは、前記入手されたコンテンツデータおよび生成されたメタデータを含む、項目22に記載の仮想画像生成システム。

(項目24)

前記メタデータは、前記選択された仮想オブジェクトのための位置、配向、および音量データを含む、項目23に記載の仮想画像生成システム。

(項目25)

前記エンドユーザの頭部姿勢を追跡するために構成される1つ以上のセンサをさらに備え、前記データベースは、前記複数の仮想オブジェクトのための絶対メタデータを記憶するために構成され、前記制御サブシステムはさらに、前記選択された仮想オブジェクトに対応する絶対メタデータを入手し、前記エンドユーザの追跡された頭部姿勢に基づいて、前記絶対メタデータを前記エンドユーザに対して局所化することによって、前記メタデータを生成するために構成される、項目23に記載の仮想画像生成システム。

10

20

30

40

50

(項目 2 6)

少なくとも 1 つのスピーカをさらに備え、前記制御サブシステムはさらに、前記記憶されたオーディオデータを前記メモリから読み出し、オーディオを前記読み出されたオーディオデータから導出し、前記オーディオを前記少なくとも 1 つのスピーカに伝達するため構成される、項目 1 に記載の仮想画像生成システム。

(項目 2 7)

前記メモリ内に記憶されるオーディオデータは、コンテンツデータおよびメタデータを含み、前記制御サブシステムはさらに、前記記憶されたコンテンツデータおよびメタデータを前記メモリから読み出し、前記読み出されたコンテンツデータおよびメタデータに基づいて、空間化オーディオをレンダリングし、前記レンダリングされた空間化オーディオを前記少なくとも 1 つのスピーカに伝達するために構成される、項目 2 6 に記載の仮想画像生成システム。

10

(項目 2 8)

前記ディスプレイサブシステムは、前記エンドユーザの眼の正面に位置付けられるために構成される、項目 1 に記載の仮想画像生成システム。

(項目 2 9)

前記ディスプレイサブシステムは、投影サブシステムおよび部分的に透明なディスプレイ表面を含み、前記投影サブシステムは、前記画像フレームを前記部分的に透明なディスプレイ表面上に投影するために構成され、前記部分的に透明なディスプレイ表面は、前記エンドユーザの眼と周囲環境との間の視野内に位置付けられるために構成される、項目 2 8 に記載の仮想画像生成システム。

20

(項目 3 0)

前記エンドユーザによって装着するために構成されるフレーム構造をさらに備え、前記フレーム構造は、前記ディスプレイサブシステムの少なくとも一部を担持する、項目 2 8 に記載の仮想画像生成システム。

(項目 3 1)

エンドユーザによって仮想画像生成システムを動作させる方法であって、
少なくとも 1 つのオブジェクトを持続的に選択することと、
少なくとも 1 つの選択されたオブジェクトから生じるビデオデータを生成することと、
複数の画像フレームを前記生成されたビデオデータからの 3 次元場面内にレンダリングすることと、

30

前記画像フレームを前記エンドユーザに表示することと、
前記少なくとも 1 つの選択されたオブジェクトから生じるオーディオデータを生成することと、

前記少なくとも 1 つの選択されたオブジェクトから生じるオーディオデータをメモリ内に記憶することと
を含む、方法。

(項目 3 2)

前記オーディオデータと同期する前記ビデオデータを前記メモリ内に記憶することをさらに含む、項目 3 1 に記載の方法。

40

(項目 3 3)

前記少なくとも 1 つの選択されたオブジェクトから生じるオーディオデータを前記エンドユーザによる知覚のための音に変換することをさらに含む、項目 3 1 に記載の方法。

(項目 3 4)

前記少なくとも 1 つのオブジェクトは、前記エンドユーザの視野内で選択される、項目 3 1 に記載の方法。

(項目 3 5)

前記少なくとも 1 つのオブジェクトを選択することは、3 次元カーソルを前記エンドユーザの視野内で移動させ、前記 3 次元カーソルを用いて前記少なくとも 1 つのオブジェクトを選択することを含む、項目 3 4 に記載の方法。

50

(項目 3 6)

前記少なくとも 1 つのオブジェクトを選択することは、1 つ以上の音声コマンドを発行することを含む、項目 3 1 に記載の方法。

(項目 3 7)

前記少なくとも 1 つのオブジェクトを選択することは、1 つ以上の手のジェスチャを行うことを含む、項目 3 1 に記載の方法。

(項目 3 8)

前記少なくとも 1 つのオブジェクトは、複数のオブジェクトを含み、前記複数のオブジェクトを選択することは、前記オブジェクトを個々に選択することを含む、項目 3 1 に記載の方法。

10

(項目 3 9)

前記少なくとも 1 つのオブジェクトは、複数のオブジェクトを含み、前記複数のオブジェクトを選択することは、前記オブジェクトを大域的に選択することを含む、項目 3 1 に記載の方法。

(項目 4 0)

前記オブジェクトを大域的に選択することは、前記エンドユーザの視野の角度範囲を定義し、前記エンドユーザの視野の定義された角度範囲内の前記オブジェクトの全てを選択することを含む、項目 3 9 に記載の方法。

20

(項目 4 1)

前記定義された角度範囲は、前記エンドユーザの視野の角度範囲全体未満である、項目 4 0 に記載の方法。

(項目 4 2)

前記定義された角度範囲は、前記エンドユーザの視野の角度範囲全体である、項目 4 0 に記載の方法。

(項目 4 3)

少なくとも 1 つの以前に選択されたオブジェクトを持続的に選択解除することをさらに含む、項目 3 1 に記載の方法。

(項目 4 4)

前記エンドユーザの視野に対する前記少なくとも 1 つの選択されたオブジェクトの場所を追跡することをさらに含む、項目 3 1 に記載の方法。

30

(項目 4 5)

前記少なくとも 1 つの選択されたオブジェクトの追跡された場所が前記エンドユーザの視野外に移動すると、前記少なくとも 1 つの選択されたオブジェクトから生じるオーディオデータを前記メモリ内に記憶することを中止することをさらに含む、項目 4 4 に記載の方法。

(項目 4 6)

前記少なくとも 1 つの選択されたオブジェクトの追跡された場所が前記エンドユーザの視野外に移動すると、前記少なくとも 1 つの選択されたオブジェクトから生じるオーディオデータを前記メモリ内に記憶することを継続することをさらに含む、項目 4 4 に記載の方法。

40

(項目 4 7)

前記少なくとも 1 つの選択されたオブジェクトは、実オブジェクトを含む、項目 3 1 に記載の方法。

(項目 4 8)

他の実オブジェクトから生じる音と比較して、前記選択された実オブジェクトから生じる音を優先的に感知することをさらに含み、前記オーディオデータは、前記優先的に感知される音から導出される、項目 4 7 に記載の方法。

(項目 4 9)

前記選択された実オブジェクトから生じるビデオデータを捕捉することと、前記オーディオデータと同期する前記ビデオデータを前記メモリ内に記憶することと

50

をさらに含む、項目 4 7 に記載の方法。

(項目 5 0)

前記捕捉されたビデオデータを仮想コンテンツデータに変換し、前記仮想コンテンツデータを前記メモリ内に記憶することをさらに含む、項目 4 9 に記載の方法。

(項目 5 1)

前記少なくとも 1 つの選択されたオブジェクトは、仮想オブジェクトを含む、項目 3 1 に記載の方法。

(項目 5 2)

複数の仮想オブジェクトのための音に対応するコンテンツデータを記憶することと、前記選択された仮想オブジェクトに対応するコンテンツデータを入手することであって、前記メモリ内に記憶されるオーディオデータは、前記入手されたコンテンツデータを含む、こと

をさらに含む、項目 5 1 に記載の方法。

(項目 5 3)

前記選択された仮想オブジェクトに対応するメタデータを生成することであって、前記メモリ内に記憶されるオーディオデータは、前記入手されたコンテンツデータおよび前記生成されたメタデータを含む、こと

をさらに含む、項目 5 2 に記載の方法。

(項目 5 4)

前記メタデータは、前記選択された仮想オブジェクトのための位置、配向、および音量データを含む、項目 5 3 に記載の方法。

(項目 5 5)

前記エンドユーザの頭部姿勢を追跡することと、

前記複数の仮想オブジェクトのための絶対メタデータを記憶することと

をさらに含み、

前記メタデータを生成することは、前記選択された仮想オブジェクトに対応する絶対メタデータを読み出し、前記エンドユーザの追跡された頭部姿勢に基づいて、前記絶対メタデータを前記エンドユーザに対して局所化することを含む、項目 5 3 に記載の方法。

(項目 5 6)

前記記憶されたオーディオデータを読み出し、オーディオを前記読み出されたオーディオデータから導出し、前記オーディオを前記エンドユーザによる知覚のための音に変換することをさらに含む、項目 3 1 に記載の方法。

(項目 5 7)

前記記憶されたオーディオデータは、コンテンツデータおよびメタデータを含み、前記方法はさらに、

前記記憶されたコンテンツデータおよびメタデータを前記メモリから読み出すことと、

前記読み出されたコンテンツデータおよびメタデータに基づいて、空間化オーディオをレンダリングすることと、

前記空間化オーディオを前記エンドユーザによる知覚のための音に変換することと

を含む、項目 3 1 に記載の方法。

(項目 5 8)

再生ユーザによる使用のための仮想画像生成システムであって、

メモリであって、前記メモリは、オリジナル空間環境内の少なくとも 1 つのオブジェクトから生じるオーディオコンテンツデータおよびビデオコンテンツデータを記憶するため構成される、メモリと、

複数のスピーカと、

ディスプレイサブシステムと、

制御サブシステムであって、前記制御サブシステムは、前記オーディオコンテンツデータおよびビデオコンテンツデータを前記メモリから読み出し、前記読み出されたオーディオコンテンツデータおよびビデオコンテンツデータからそれぞれオーディオおよびビデオ

10

20

30

40

50

を前記オリジナル空間環境と異なる新しい空間環境内にレンダリングし、前記レンダリングされたオーディオを前記スピーカに、前記生成されたビデオデータを前記ディスプレイス
サブシステムに同期して伝達するために構成される、制御サブシステムと
を備える、仮想画像生成システム。

(項目 5 9)

前記制御サブシステムは、前記オーディオコンテンツデータおよびビデオコンテンツデータを前記メモリ内に記憶するために構成される、項目 5 8 に記載の仮想画像生成システム。

(項目 6 0)

前記メモリ内への前記オーディオコンテンツデータおよびビデオコンテンツデータの記憶に先立って、入力をエンドユーザから受信し、エンドユーザ入力に応答して、前記オリジナル空間環境内の前記少なくとも 1 つのオブジェクトを持続的に選択するために構成されるオブジェクト選択デバイスをさらに備える、項目 5 9 に記載の仮想画像生成システム。

(項目 6 1)

前記少なくとも 1 つのオブジェクトは、実オブジェクトを含む、項目 5 9 に記載の仮想画像生成システム。

(項目 6 2)

前記オーディオコンテンツデータを前記オリジナル空間環境内の実オブジェクトから捕捉するために構成されるマイクロホンアセンブリをさらに備える、項目 6 1 に記載の仮想画像生成システム。

(項目 6 3)

前記マイクロホンアセンブリは、オーディオ出力を生成するために構成され、前記制御サブシステムはさらに、前記オーディオ出力の方向を修正し、選択された実オブジェクトから生じる音を優先的に感知するために構成され、前記オーディオコンテンツデータは、前記修正されたオーディオ出力から導出される、項目 6 2 に記載の仮想画像生成システム。

(項目 6 4)

前記ビデオデータを前記オリジナル空間環境内の選択された実オブジェクトから捕捉するために構成される 1 つ以上のカメラをさらに備える、項目 6 1 に記載の仮想画像生成システム。

(項目 6 5)

前記制御サブシステムは、前記捕捉されたビデオデータを前記選択された実オブジェクトのための仮想コンテンツデータに変換し、前記仮想コンテンツデータを前記ビデオコンテンツデータとして前記メモリ内に記憶するために構成される、項目 6 4 に記載の仮想画像生成システム。

(項目 6 6)

前記少なくとも 1 つのオブジェクトは、仮想オブジェクトを含む、項目 5 8 に記載の仮想画像生成システム。

(項目 6 7)

複数の仮想オブジェクトのための音に対応するコンテンツデータを記憶するために構成されるデータベースをさらに備え、前記制御サブシステムはさらに、前記仮想オブジェクトに対応するコンテンツデータを前記データベースから入手するために構成され、前記メモリ内に記憶されるオーディオデータは、前記入手されたコンテンツデータを含む、項目 6 6 に記載の仮想画像生成システム。

(項目 6 8)

前記制御サブシステムは、前記新しい空間環境内の前記少なくとも 1 つのオブジェクトに対応する絶対メタデータ入手し、前記読み出されたオーディオコンテンツデータおよび前記絶対メタデータからのオーディオを前記新しい空間環境内にレンダリングするために構成される、項目 5 8 に記載の仮想画像生成システム。

10

20

30

40

50

(項目 6 9)

前記再生ユーザの頭部姿勢を追跡するために構成される 1 つ以上のセンサをさらに備え、前記制御サブシステムはさらに、前記レンダリングされたオーディオが空間化されるように、前記再生ユーザの追跡された頭部姿勢に基づいて、前記絶対メタデータを前記再生ユーザに対して局所化するために構成される、項目 6 8 に記載の仮想画像生成システム。

(項目 7 0)

前記新しい空間環境内の少なくとも 1 つのオブジェクトに対応する絶対メタデータ入手することは、前記少なくとも 1 つのオブジェクトを前記新しい空間環境内に位置付けることを含む、項目 6 8 に記載の仮想画像生成システム。

(項目 7 1)

入力を前記再生ユーザから受信するために構成されるユーザ入力デバイスをさらに備え、前記制御サブシステムは、前記再生ユーザからの入力に応答して、前記少なくとも 1 つのオブジェクトを前記新しい空間環境内に位置付けるために構成される、項目 7 0 に記載の仮想画像生成システム。

(項目 7 2)

前記ディスプレイサブシステムは、前記再生ユーザの眼の正面に位置付けられるために構成される、項目 5 8 に記載の仮想画像生成システム。

(項目 7 3)

前記ディスプレイサブシステムは、投影サブシステムおよび部分的に透明なディスプレイ表面を含み、前記投影サブシステムは、前記ビデオデータに応答して、画像フレームを前記部分的に透明なディスプレイ表面上に投影させるために構成され、前記部分的に透明なディスプレイ表面は、前記再生ユーザの眼と前記新しい空間環境との間の視野内に位置付けられるために構成される、項目 7 2 に記載の仮想画像生成システム。

(項目 7 4)

前記再生ユーザによって装着されるために構成されるフレーム構造をさらに備え、前記フレーム構造は、前記ディスプレイサブシステムの少なくとも一部を担持する、項目 7 2 に記載の仮想画像生成システム。

(項目 7 5)

オリジナル空間環境内にオーディオコンテンツデータおよびビデオコンテンツデータとして以前に記録された少なくとも 1 つのオブジェクトのオーディオおよびビデオを再生するために、再生ユーザによって仮想画像生成システムを動作させる方法であって、前記方法は、

前記オーディオコンテンツデータおよびビデオコンテンツデータをメモリから読み出すこと、

前記読み出されたオーディオコンテンツデータおよびビデオコンテンツデータからそれぞれオーディオおよびビデオを前記オリジナル空間環境と異なる新しい空間環境内にレンダリングすること、

前記オーディオおよびビデオをそれぞれ音および画像フレームに変換すること、

前記音および画像フレームを前記再生ユーザに同期して伝達すること

を含む、方法。

(項目 7 6)

前記オーディオコンテンツデータおよびビデオコンテンツデータを前記メモリ内に記憶することをさらに含む、項目 7 5 に記載の方法。

(項目 7 7)

前記メモリ内への前記オーディオコンテンツデータおよびビデオコンテンツデータの記憶に先立って、前記オリジナル空間環境内の前記少なくとも 1 つのオブジェクトを持続的に選択することをさらに含む、項目 7 6 に記載の方法。

(項目 7 8)

前記少なくとも 1 つのオブジェクトは、実オブジェクトを含む、項目 7 6 に記載の方法

(項目 7 9)

前記オーディオコンテンツデータを前記実オブジェクトから捕捉することをさらに含む、項目 7 8 に記載の方法。

(項目 8 0)

他の実オブジェクトから生じる音と比較して、選択された実オブジェクトから生じる音を優先的に感知することであって、前記オーディオコンテンツデータは、前記優先的に感知される音から導出される、ことをさらに含む、項目 7 9 に記載の方法。

(項目 8 1)

ビデオデータを前記選択された実オブジェクトから捕捉することと、前記捕捉されたビデオデータを前記仮想コンテンツデータに変換することとをさらに含む、項目 7 8 に記載の方法。

10

(項目 8 2)

前記少なくとも 1 つのオブジェクトは、仮想オブジェクトを含む、項目 7 6 に記載の方法。

(項目 8 3)

複数の仮想オブジェクトのための音に対応するコンテンツデータを記憶し、前記仮想オブジェクトに対応するコンテンツデータを前記データベースから入手することであって、前記メモリ内に記憶されるオーディオコンテンツデータは、前記入手されたコンテンツデータを含む、こと

をさらに含む、項目 8 2 に記載の方法。

20

(項目 8 4)

前記新しい空間環境内の前記少なくとも 1 つのオブジェクトに対応する絶対メタデータを入手することであって、前記オーディオは、前記読み出されたオーディオコンテンツデータおよび前記絶対メタデータから前記新しい空間環境内にレンダリングされる、ことをさらに含む、項目 7 5 に記載の方法。

(項目 8 5)

前記再生ユーザの頭部姿勢を追跡することと、

前記再生ユーザの追跡された頭部姿勢に基づいて、前記絶対メタデータを前記再生ユーザに対して局所化することであって、前記オーディオは、前記レンダリングされたオーディオが空間化されるように、前記読み出されたオーディオコンテンツデータおよび前記局所的メタデータから前記新しい空間環境内にレンダリングされる、ことをさらに含む、項目 8 4 に記載の方法。

30

(項目 8 6)

前記新しい空間環境内の少なくとも 1 つのオブジェクトに対応する絶対メタデータを入手することは、前記少なくとも 1 つのオブジェクトを前記新しい空間環境内に位置付けることを含む、項目 8 4 に記載の方法。

(項目 8 7)

入力を前記再生ユーザから受信することであって、前記少なくとも 1 つのオブジェクトは、前記再生ユーザからの入力に応答して、前記新しい空間環境内に位置付けられる、ことをさらに含む、項目 8 6 に記載の方法。

40

【図面の簡単な説明】**【0029】**

図面は、本発明の好ましい実施形態の設計および有用性を図示し、類似要素は、共通参考番号によって参照される。本発明の上記および他の利点および目的が取得される方法をより深く理解するために、上記に簡単に説明された本発明のより具体的な説明が、付随の図面に図示される、その具体的な実施形態を参考することによって与えられるであろう。これらの図面は、本発明の典型的な実施形態のみを描写し、したがって、その範囲の限定と見なされるべきではないことを理解した上で、本発明は、付随の図面の使用を通して、附加的特異性および詳細とともに、説明および解説されるであろう。

50

【0030】

【図1】図1は、先行技術の拡張現実生成デバイスによってエンドユーザに表示され得る、3次元拡張現実場面の図である。

【0031】

【図2】図2は、本発明の一実施形態に従って構築された、拡張現実システムの斜視図である。

【0032】

【図3】図3は、図2の拡張現実システムのブロック図である。

【0033】

【図4】図4は、図2の拡張現実システム内で使用される空間化スピーカシステムの一実施形態の平面図である。 10

【0034】

【図5】図5は、エンドユーザがオブジェクトを個々に選択することを可能にするために図2の拡張現実システムによって使用される、1つの技法を図示する、平面図である。

【0035】

【図6】図6は、エンドユーザがオブジェクトを個々に選択することを可能にするために図2の拡張現実システムによって使用される、別の技法を図示する、平面図である。

【0036】

【図7】図7は、エンドユーザがオブジェクトを個々に選択することを可能にするために図2の拡張現実システムによって使用される、さらに別の技法を図示する、平面図である。 20

【0037】

【図8】図8は、エンドユーザが複数のオブジェクトを大域的に選択することを可能にするために図2の拡張現実システムによって使用される、技法を図示する、平面図である。

【0038】

【図9】図9は、エンドユーザが複数のオブジェクトを大域的に選択することを可能にするために図2の拡張現実システムによって使用される、別の技法を図示する、平面図である。

【0039】

【図10a】図10aは、図2の拡張現実システムを装着するために使用され得る、1つの技法の平面図である。 30

【0040】

【図10b】図10bは、図2の拡張現実システムを装着するために使用され得る、別の技法の平面図である。

【0041】

【図10c】図10cは、図2の拡張現実システムを装着するために使用され得る、さらに別の技法の平面図である。

【0042】

【図10d】図10dは、図2の拡張現実システムを装着するために使用され得る、さらに別の技法の平面図である。 40

【0043】

【図11】図11は、種々の例示的な種々の仮想音および実音と相互作用する、図2の拡張現実システムを図示する、ブロック図である。

【0044】

【図12】図12は、図2の拡張現実システム内で使用されるオーディオプロセッサの一実施形態を図示する、ブロック図である。

【0045】

【図13】図13は、図2の拡張現実システムによって選択された仮想および実オブジェクトに対応するコンテンツデータおよびメタデータを記録する、メモリの略図である。

【0046】

【図14】図14は、音を実オブジェクトから優先的に受信するために図2の拡張現実システム内で使用される、マイクロホンアセンブリおよび対応するオーディオ処理モジュールの概略である。

【0047】

【図15a】図15aは、音をエンドユーザに対して第1の配向を有する2つのオブジェクトから優先的に受信するために図2の拡張現実システムのオーディオプロセッサによって生成された、指向性パターンの平面図である。

【0048】

【図15b】図15bは、音をエンドユーザに対して第2の配向を有する2つのオブジェクトから優先的に受信するために図2の拡張現実システムのオーディオプロセッサによって生成された、指向性パターンの平面図である。

【0049】

【図16a】図16aは、エンドユーザに対してオリジナル空間環境内で分散されたオブジェクトのブロック図である。

【0050】

【図16b】図16bは、エンドユーザに対して新しい空間環境内で分散された図17aのオブジェクトのブロック図である。

【0051】

【図17】図17は、図2の拡張現実システムを動作させ、仮想および実オブジェクトのオーディオおよびビデオを選択および記録する1つの方法を図示する、フロー図である。

20

【0052】

【図18】図18は、図2の拡張現実システムを動作させ、図17において記録されたオーディオおよびビデオを新しい空間環境内で再生する1つの方法を図示する、フロー図である。

【発明を実施するための形態】

【0053】

続く説明は、拡張現実システム内で使用されるべきディスプレイシステムおよび方法に関する。しかしながら、本発明は、拡張現実システムにおける用途に非常に適しているが、本発明は、その最も広い側面において、そのように限定され得ないことを理解されたい。例えば、本発明は、仮想現実システムに適用されることができる。したがって、多くの場合、拡張現実システムの観点から本明細書に説明されるが、本教示は、そのような使用のそのようなシステムに限定されるべきではない。拡張現実システムは、例えば、ビデオゲーム、仮想および実人物の組み合わせとのビデオ会議、または映画の鑑賞のコンテキストにおいて動作されてもよい。

30

【0054】

本明細書に説明される拡張現実システムは、エンドユーザが、エンドユーザによって持続的に選択される少なくとも1つのオブジェクト（仮想または実際のいずれか）から生じるオーディオデータを記録することを可能にする。そのような記録されたオーディオデータは、続いて、同一または異なるエンドユーザによって再生されることができる。記録されたオーディオデータから生じる音は、オーディオデータが元々記録された実環境内の同一または異なるエンドユーザに再生されてもよい。オーディオデータのコンテンツを記録することに加え、オーディオコンテンツが元々記録された環境を特徴付けるメタデータおよびエンドユーザの頭部姿勢が、再生の間、オーディオは、再レンダリングされ、エンドユーザがオリジナル記録の間に空間化音を聴覚的に体験したものと同一様式において聴覚的に体験される、空間化音に変換され得るように、そのようなオーディオデータと関連付けて記録されてもよい。随意に、オーディオは、同一または異なるエンドユーザが新しい環境に適切な聴覚的体験を有し得るように、再レンダリングされ、新しい仮想または実環境内の同一または異なるエンドユーザによる知覚のための空間化音に変換されてもよい。オーディオデータは、周囲環境内の仮想オブジェクトおよび実オブジェクトから生じるビデオデータと同期して記録されてもよい。

40

50

【0055】

本明細書に説明される拡張現実システムは、エンドユーザの視野内の実（または物理的）オブジェクトと混合された仮想オブジェクトの画像を提供し、かつ（視野内または外のいずれかの）実（または物理的）源から生じる実音と混合された（視野内または外のいずれかの）仮想源から生じる仮想音を提供するように動作されてもよい。この目的を達成するため、本発明に従って構築された拡張現実システム100の一実施形態が、ここで、図2および3を参照して説明されるであろう。拡張現実システム100は、ディスプレイサブシステム102を備え、これは、ディスプレイ画面104と、画像をディスプレイ画面104上に投影する、投影サブシステム（図示せず）とを含む。

【0056】

10

図示される実施形態では、ディスプレイ画面104は、部分的に透明なディスプレイ画面であって、それを通して、周囲環境内の実オブジェクトが、エンドユーザ50によって見られることができ、その上に、仮想オブジェクトの画像が、表示されてもよい。拡張現実システム100はさらに、ディスプレイ画面104が、エンドユーザ50の眼52の正面、特に、エンドユーザ50の眼52と周囲環境との間のエンドユーザ50の視野内に位置付けられるように、部分的に透明なディスプレイ画面104を担持する、エンドユーザ50によって装着される、フレーム構造106を備える。

【0057】

ディスプレイサブシステム102は、エンドユーザ50の眼52に、高レベルの画質および3次元知覚を伴って、かつ2次元コンテンツを提示することも可能である、物理的現実に対する拡張として快適に知覚され得る、光ベースの放射パターンを提示するように設計される。ディスプレイサブシステム102は、單一コヒーレント場面の知覚を提供する、フレームのシーケンスを高周波数で提示する。

20

【0058】

代替実施形態では、拡張現実システム100は、1つ以上の結像機（例えば、カメラ）を採用し、周囲環境の画像を捕捉し、それをビデオデータに変換してもよく、これは、次いで、仮想オブジェクトを表すビデオデータと混合することができ、その場合、拡張現実システム100は、混合されたビデオデータを表す画像をエンドユーザ50に不透明ディスプレイ表面上で表示してもよい。

【0059】

30

ディスプレイサブシステムを説明するさらなる詳細は、「Display Subsystem and Method」と題された米国仮特許出願第14/212,961号および「Planar Waveguide Apparatus With Diffraction Element(s) and Subsystem Employing Same」と題された米国仮特許出願第14/331,216号（参照することによって明示的に本明細書に組み込まれる）に提供される。

【0060】

拡張現実システム100はさらに、仮想オブジェクトから音のみをエンドユーザ50に提示しながら、エンドユーザ50が、直接、実オブジェクトからの音を聞くことを可能にするための1つ以上のスピーカ108を備える。代替実施形態では、拡張現実システム100は、1つ以上のマイクロホン（図示せず）を備え、周囲環境から生じる実音を捕捉し、それをオーディオデータに変換してもよく、これは、仮想音からオーディオデータと混合することができ、その場合、スピーカ108は、混合されたオーディオデータを表す音をエンドユーザ50に伝達してもよい。

40

【0061】

いずれの場合も、スピーカ108は、スピーカ108が、例えば、イヤーバッドまたはヘッドホンとして、エンドユーザ50の外耳道に隣接して（その中または周囲に）位置付けられるように、フレーム構造106によって担持される。スピーカ108は、ステレオ/調節可能音制御を提供してもよい。スピーカ108は、外耳道に隣接して位置付けられるように説明されるが、外耳道に隣接して位置しない、他のタイプのスピーカも、音を工

50

ンドユーザ50に伝達するために使用されることができる。例えば、スピーカは、例えば、骨伝導技術を使用して、外耳道からある距離に設置されてもよい。図4に図示される随意の実施形態では、複数の空間化スピーカ108（例えば、4つのスピーカ108-1、108-2、108-3、および108-4）が、エンドユーザ50の頭部54を中心として位置し、音を頭部54の左、右、正面、および背面から受信するために構成され、エンドユーザ50の左および右耳56に向かって向けられてもよい。拡張現実システムのために使用され得る、空間化スピーカに関するさらなる詳細は、「Mixed Reality System with Spatialized Audio」と題された米国仮特許出願第62/369,561号（参照することによって明示的に本明細書に組み込まれる）に説明される。

10

【0062】

重要なこととして、拡張現実システム100は、エンドユーザ50が、これらの選択されたオブジェクトからのみの音の記録のために、（仮想または実際のいずれかの）1つ、いくつか、または全てのオブジェクトを選択することを可能にするために構成される。この目的を達成するために、拡張現実システム100はさらに、エンドユーザ50からの入力に応答して、そこからの音を記録するために、1つ以上の実オブジェクト（すなわち、そこから実音が生じる、実オブジェクト）および仮想オブジェクト（すなわち、そこから仮想音が生じる、仮想オブジェクト）を選択するために構成される、オブジェクト選択デバイス110を備える。オブジェクト選択デバイス110は、エンドユーザ50の視野内の実オブジェクトまたは仮想オブジェクトを個々に選択する、および／または、エンドユーザ50の視野内の実オブジェクトまたは仮想オブジェクトのサブセットまたは全てを大域的に選択するように設計されてもよい。オブジェクト選択デバイス110はまた、エンドユーザ50からの付加的入力に応答して、1つ以上の以前に選択された実オブジェクトまたは仮想オブジェクトを選択解除するために構成されてもよい。この場合、オブジェクト選択デバイス110は、それらが以前に選択されたものと同一様式において、実オブジェクトまたは仮想オブジェクトを選択解除するように設計されてもよい。いずれの場合も、具体的オブジェクトは、持続的に選択され、意図的に選択解除されるまで、具体的オブジェクトが選択された状態のままであることを意味する。

20

【0063】

一実施形態では、ディスプレイサブシステム102は、3次元カーソルをエンドユーザ50の視野内に表示してもよく、これは、オブジェクト選択デバイス110への入力に応答して、拡張現実場面内の具体的実オブジェクトまたは仮想オブジェクトを選択する目的のために、エンドユーザ50の視野内で変位され得る。

30

【0064】

例えば、図5に示されるように、4つの仮想オブジェクト（V1-V4）および2つの実オブジェクト（R1-R2）が、ディスプレイ画面104の視野60内に位置する。ディスプレイサブシステム102は、3Dカーソル62を視野60内に表示してもよく、これは、円形の形態をとるよう図示される。3Dカーソル62は、オブジェクト選択デバイス110へのエンドユーザ50による入力に応答して、オブジェクトのうちの1つにわたって、この場合は、仮想オブジェクトV3にわたって移動され、それによって、3Dカーソル62とそのオブジェクトを関連付けてもよい。関連付けられたオブジェクトは、次いで、オブジェクト選択デバイス110へのエンドユーザ50による付加的入力に応答して、選択されることができる。具体的オブジェクト（この場合、仮想オブジェクトV3）が、3Dカーソル62と関連付けられ、選択の準備ができたことの視覚的フィードバックを提供するために、関連付けられたオブジェクトまたはさらに3Dカーソル62自体が、ハイライトされてもよい（例えば、色または陰影の変化）。いったん選択されると、オブジェクトは、選択解除されるまで、ハイライトされたままであってもよい。当然ながら、仮想オブジェクトV3の代わりに、または仮想オブジェクトV3に加え、実オブジェクトを含む、拡張現実場面4内の他のオブジェクトも、3Dカーソル62をこれらの他のオブジェクトのいずれかにわたって設置し、3Dカーソル62内でオブジェクトを選択するこ

40

50

とによって、選択されることがある。また、図5における3Dカーソル62は、円形の形態をとるが、3Dカーソル62は、具体的オブジェクトを指示するためにエンドユーザ50によって使用され得る、矢印を含む、任意の形状であることができるこを理解されたい。視野60内の以前に選択されたオブジェクトのいずれかは、3Dカーソル62をその以前に選択されたオブジェクトにわたって移動させ、それを選択解除することによって、選択解除されることがある。

【0065】

オブジェクト選択デバイス110は、エンドユーザ50が、3Dカーソル62を具体的オブジェクトにわたって移動させ、続いて、その具体的オブジェクトを選択することを可能にする、任意のデバイスの形態をとることができる。一実施形態では、オブジェクト選択デバイス110は、物理的に操作され、3Dカーソル62を具体的オブジェクトにわたって移動させ、「クリック」し、具体的オブジェクトを選択し得る、マウス、タッチパッド、ジョイスティック、指向性ボタン等の従来の物理的コントローラの形態をとる。

【0066】

別の実施形態では、オブジェクト選択デバイス110は、音声コマンドに応答して、3Dカーソル62を具体的オブジェクトにわたって移動させ、次いで、具体的オブジェクトを選択し得る、マイクロホンおよび対応する音声解釈モジュールを備えてもよい。例えば、エンドユーザ50は、指向性コマンド、例えば、「左に移動」または「右に移動」を発し、3Dカーソル62を具体的オブジェクトにわたって漸次的に移動させ、次いで、「選択」等のコマンドを発し、具体的オブジェクトを選択してもよい。

【0067】

さらに別の実施形態では、オブジェクト選択デバイス110は、フレーム構造106に搭載される、1つ以上のカメラ（例えば、前向きに面したカメラ112）と、具体的オブジェクトの選択のために、対応して、3Dカーソル62を具体的オブジェクトにわたって移動させる、エンドユーザ50による物理的ジェスチャ（例えば、指移動）を追跡することが可能な対応するプロセッサ（図示せず）とを備えてもよい。例えば、エンドユーザ50は、指を使用して、視野60内の3Dカーソル62を具体的オブジェクトにわたって「ドラッグ」し、次いで、3Dカーソル62を「タップ」し、具体的オブジェクトを選択してもよい。または、前向きに面したカメラ112が、例えば、少なくとも部分的に、具体的オブジェクトの選択のために、対応して、3Dカーソル62を具体的オブジェクトにわたって移動させる、エンドユーザ50の頭部54の配向に基づいて、例えば、エンドユーザ50の注意の中心を検出または推測するために採用されてもよい。例えば、エンドユーザ50は、その頭部50を移動させ、視野60内の3Dカーソル62を具体的オブジェクトにわたって「ドラッグ」し、次いで、その頭部50を素早く点頭させ、具体的オブジェクトを選択してもよい。

【0068】

さらに別の実施形態では、オブジェクト選択デバイス110は、1つ以上のカメラ（例えば、後ろ向きに面したカメラ114（図2に示される））と、エンドユーザ50の眼52、特に、その具体的オブジェクトの選択のために、対応して、3Dカーソル62を具体的オブジェクトにわたって移動させる、エンドユーザ50が合焦している方向および/または距離を追跡する、対応するプロセッサとを備えてもよい。後ろ向きに面したカメラ114は、エンドユーザ50の眼52の角位置（片眼または両眼が向いている方向）、瞬目、および焦点深度を追跡してもよい（眼収束を検出することによって）。例えば、エンドユーザ50は、その眼54を視野内で移動させ、3Dカーソルを具体的オブジェクトにわたって「ドラッグ」し、次いで、瞬目し、具体的オブジェクトを選択してもよい。そのような眼追跡情報は、例えば、光をエンドユーザの眼に投影させ、その投影された光の少なくとも一部の戻りまたは反射を検出することによって、判別されてもよい。眼追跡デバイスについて議論するさらなる詳細は、「Display Subsystem and Method」と題された米国仮特許出願第14/212,961号、「Methods and Subsystem for Creating Focal Planes

10

20

30

40

50

in Virtual and Augumented Reality」と題された米国特許出願第14/726,429号、および「Subsystem and Method for Augumented and Virtual Reality」と題された米国特許出願第14/205,126号（参照することによって明示的に本明細書に組み込まれる）に提供される。

【0069】

代替実施形態では、オブジェクト選択デバイス110は、従来の物理的コントローラ、マイクロホン／音声解釈モジュール、および／またはカメラを組み合わせて、3Dカーソル62を移動させ、使用し、オブジェクトを選択してもよい。例えば、物理的コントローラ、指ジェスチャ、または眼移動が、3Dカーソル62を具体的オブジェクトにわたって移動させるために使用することができ、音声コマンドが、その具体的オブジェクトを選択するために使用することができる。

【0070】

3Dカーソル62を使用して、エンドユーザ50の視野内のオブジェクトを選択するのではなく、具体的オブジェクトが、その具体的オブジェクトを意味論的に識別すること、またはエンドユーザ50に表示されるメニューを介してオブジェクトを選択することによって、選択されてもよく、その場合、オブジェクトは、エンドユーザ50の視野内にある必要はない。この場合、オブジェクト選択デバイス110は、具体的オブジェクトが、エンドユーザ50によって提供される口頭コマンドを変換するように意味論的に識別される場合、マイクロホンおよび音声解釈モジュールの形態をとる。例えば、仮想オブジェクトV3が、ドラムに対応する場合、エンドユーザ50は、「ドラムを選択」と発してもよく、それに応答して、ドラムV3が、選択されるであろう。口頭コマンドに対応するオブジェクトの選択を促進するために、視野内の全ての関連オブジェクトを識別する意味論情報が、好ましくは、エンドユーザ50によって口頭で表されるオブジェクトの説明が、データベース内に記憶されるオブジェクトの説明に合致され得るように、データベース内に記憶される。意味論情報を含む、メタデータが、データベース内の仮想オブジェクトと事前に関連付けができる一方、視野内の実オブジェクトは、「Method and System for Inserting Recognized Object Data into a Virtual World」と題された米国特許出願第14/704,800号（参照することによって明示的に本明細書に組み込まれる）に説明される様式において、意味論情報と事前にマッピングされ、関連付けられてもよい。

【0071】

代替として、具体的オブジェクトは、3Dカーソル62を使用せずに、単に、指ジェスチャを使用してそれを指差すまたは「クリック」することによって、選択されてもよい。この場合、オブジェクト選択デバイス110は、1つ以上のカメラ（例えば、前向きに面したカメラ114）と、具体的オブジェクトの選択のための指ジェスチャを追跡する、対応するプロセッサとを備えてもよい。例えば、エンドユーザ50は、単に、図6に示されるように、それを指差すことによって、具体的オブジェクト（この場合、仮想オブジェクトV3）を選択してもよい。別の実施形態では、具体的オブジェクトは、3Dカーソル62を使用せずに、図7に示されるように、少なくとも2本の指（例えば、人差し指および親指）を使用して円形または部分的円形を形成することによって、選択されてもよい。

【0072】

3Dカーソル62は、一度に1つのみのオブジェクトを選択するために使用されるように説明されたが、その代替または随意の実施形態では、3Dカーソル62は、一度に複数のオブジェクトを選択するために使用されてもよい。例えば、図8に図示されるように、線64が、3Dカーソル62を使用して、あるグループのオブジェクトの周囲に、例えば、実オブジェクトR1および仮想オブジェクトV3およびV4の周囲に描かれ、それによって、これらのグループのオブジェクトを選択することができる。3Dカーソル62は、例えば、オブジェクトを個々に選択するために、上記に説明されるものと同一手段を使用して制御されることができる。代替として、線が、3Dカーソル62を使用せずに、例え

10

20

30

40

50

ば、指ジェスチャを使用することによって、あるグループのオブジェクトの周囲に描かれることができる。

【0073】

代替実施形態では、エンドユーザ50の事前に定義された視野の角度範囲内のオブジェクトのグループが、選択されてもよく、その場合、オブジェクト選択デバイス110は、例えば、エンドユーザ50によって作動され、これらのオブジェクトを選択し得る、単一物理的または仮想選択ボタンの形態をとることができる。視野の角度範囲は、エンドユーザ50によって事前に定義されてもよい、または拡張現実システム100の中に事前にプログラムされてもよい。例えば、図9に示されるように、60度（視野の中心から±30度）の角度範囲66が、120度視野60のコンテキストに示される。視野60の角度範囲64内の全てのオブジェクト（この場合、仮想オブジェクトV1、V2、およびV3）が、選択ボタンの作動に応じて、大域的に選択されることができる一方、視野60の角度範囲64外の全てのオブジェクト（この場合、実オブジェクトR1およびR2および仮想オブジェクトV4）は、選択ボタンの作動に応じて、選択されないであろう。一実施形態では、エンドユーザ50は、例えば、視野60の中心線に向かってまたはそこから離れる角度範囲を定義する縁の一方または両方をドラッグすることによって、角度範囲を修正してもよい（矢印によって示される）。エンドユーザ50は、例えば、角度範囲を最小の0度から視野全体（例えば、120度）まで調節してもよい。代替として、視野60の角度範囲64は、エンドユーザ50がそれを調節する能力を伴わずに、事前にプログラムされてもよい。例えば、視野60の全体内の全てのオブジェクトが、選択ボタンの作動に応答して、選択されてもよい。

【0074】

拡張現実システム100はさらに、周囲環境内の実オブジェクトからの音をオーディオ信号に変換するために構成される、1つ以上のマイクロホンを備える。特に、拡張現実システム100は、オブジェクト選択デバイス110を介してエンドユーザ50によって選択された1つ以上の実オブジェクトの方向および距離に対応する、特定の方向および/または特定の距離における音を優先的に受信するために構成される、マイクロホンアセンブリ116を備える。マイクロホンアセンブリ116は、図2に図示されるように（2つのみが示される）、フレーム構造106に搭載される、マイクロホン要素118のアレイ（例えば、4つのマイクロホン）を備える。マイクロホンアセンブリ116に関する詳細は、下記にさらに詳細に説明されるであろう。拡張現実システム100はさらに、例えば、コマンドまたは叙述をエンドユーザ50から受信するために、エンドユーザ50の発話をオーディオ信号に変換するために構成される、専用マイクロホン122を備える。

【0075】

拡張現実システム100は、拡張現実システムに対する選択された実オブジェクトの位置または配向が変化しても、これらの実オブジェクトから生じる音が、マイクロホンアセンブリ116によって選択されていない実オブジェクトと比較して、優先的かつ継続的に感知され得るように、既知の座標系内の選択された実オブジェクトの位置および配向を追跡する。既知の座標系内の全ての仮想オブジェクトの位置および場所は、典型的には、拡張現実システム100に対して「既知」であって（すなわち、その中に記録される）、したがって、概して、能動的に追跡される必要がない。

【0076】

図示される実施形態では、拡張現実システム100は、仮想オブジェクトに対応する空間化オーディオを実および物理的3次元（3D）空間内の既知の仮想場所および配向でレンダリングおよび提示し、音の明確性または現実性に影響を及ぼすように、音が実オブジェクトの仮想場所から生じるようにエンドユーザ50に現れさせる、空間化オーディオシステムを採用する。拡張現実システム100は、種々の仮想オブジェクトと関連付けられたオーディオがその仮想位置から生じるように現れるように、エンドユーザ50の位置を追跡し、空間化オーディオをより正確にレンダリングする。さらに、拡張現実システム100は、種々の仮想オブジェクトと関連付けられた指向性オーディオが個別の仮想オブジ

10

20

30

40

50

エクトに適切な仮想方向に（例えば、仮想キャラクタの頭部の背後からではなく、仮想キャラクタの口から）伝搬するように現れるように、エンドユーザ50の頭部姿勢を追跡し、空間化オーディオをより正確にレンダリングする。さらに、拡張現実システム100は、種々の仮想オブジェクトと関連付けられたオーディオが、実際の物理的および仮想オブジェクトから適切に反射する、またはそれによってオクルードまたは妨害されるように現れるように、空間化オーディオをレンダリングする際、他の実際の物理的および仮想オブジェクトを考慮する。

【0077】

この目的を達成するために、拡張現実システム100はさらに、仮想3次元場面に対するエンドユーザ50の頭部54の位置および配向を追跡し、かつエンドユーザ50の頭部54に対する実オブジェクトの位置および配向を追跡するための頭部／オブジェクト追跡サブシステム120を備える。例えば、頭部／オブジェクト追跡サブシステム120は、エンドユーザ50の頭部姿勢データ（位置および配向）を収集するために構成される、1つ以上のセンサと、センサ120によって収集された頭部姿勢データに基づいて、既知の座標系内のエンドユーザ50の頭部姿勢を決定するために構成される、プロセッサ（図示せず）とを備えてもよい。センサは、画像捕捉デバイス（可視および赤外線光カメラ等）、慣性測定ユニット（加速度計およびジャイロスコープを含む）、コンパス、マイクロホン、GPSユニット、または無線デバイスのうちの1つ以上のものを含んでもよい。図示される実施形態では、センサは、前向きに面したカメラ112（図2に示される）を備える。このように頭部に装着されると、前向きに面したカメラ120は、特に、エンドユーザ50が位置する環境に対するエンドユーザ50の頭部54の距離および角位置（すなわち、頭部が向けられている方向）を示す情報を捕捉するために好適である。頭部配向は、任意の方向（例えば、エンドユーザ50の基準フレームに対して上／下、左、右）において検出されてもよい。下記にさらに詳細に説明されるであろうように、前向きに面したカメラ114はまた、周囲環境内の実オブジェクトのビデオデータを入手し、拡張現実システム100のビデオ記録機能を促進するために構成される。カメラはまた、周囲環境内の実オブジェクトを追跡するために提供されてもよい。フレーム構造106は、カメラがフレーム構造106の正面および背面上に搭載され得るように、設計されてもよい。このように、カメラのアレイは、エンドユーザ50の頭部54を包囲し、関連オブジェクトの全ての方向を網羅してもよい。

【0078】

拡張現実システム100はさらに、仮想3次元場面を記憶するために構成される、3次元データベース124を備え、これは、仮想オブジェクト（仮想オブジェクトのコンテンツデータと、これらの仮想オブジェクトと関連付けられた絶対メタデータ、例えば、3D場面内のこれらの仮想オブジェクトの絶対位置および配向との両方）と、仮想オブジェクト（仮想オブジェクトのコンテンツデータと、これらの仮想オブジェクトと関連付けられた絶対メタデータ、例えば、3D場面内のこれらの仮想オブジェクトの音量および絶対位置および配向、および仮想源の近傍の任意の仮想または実オブジェクトを含む、各仮想オブジェクトを囲繞する空間音響、部屋寸法、壁／床材料等との両方）とを備える。

【0079】

拡張現実システム100はさらに、視野内に現れる仮想オブジェクトおよび実オブジェクトから生じるビデオデータを記録することに加え、エンドユーザ50がオブジェクト選択デバイス110を介して選択したこれらの仮想オブジェクトおよび実オブジェクトからのみ生じるオーディオデータを記録する、制御サブシステムを備える。拡張現実システム100はまた、同期されたビデオおよびオーディオが、再生の間、正確に再レンダリングされ得るように、ビデオデータおよびオーディオデータと関連付けられたメタデータを記録してもよい。

【0080】

この目的を達成するために、制御サブシステムは、仮想オブジェクトと関連付けられたビデオコンテンツおよび絶対メタデータを3次元データベース124から入手し、エンド

10

20

30

40

50

ユーザ50の頭部姿勢データ（下記にさらに詳細に説明されるように、ビデオのための絶対メタデータをエンドユーザ50の頭部54に対して局所化するために使用されるであろう）を頭部／オブジェクト追跡サブシステム120から入手し、次いで、エンドユーザ50の視野内の周囲環境内の実オブジェクトから生じる画像と混合された画像への変換のために、ディスプレイサブシステム102に伝達される、そこからのビデオをレンダリングするために構成される、ビデオプロセッサ126を備える。ビデオプロセッサ126はまた、下記にさらに説明されるであろうように、仮想オブジェクトから生じるビデオデータとともに、続いて記録されるであろう、周囲環境の実オブジェクトから生じるビデオデータを前向きに面したカメラ112から入手するために構成される。

【0081】

10

同様に、オーディオプロセッサ128は、仮想オブジェクトと関連付けられたオーディオコンテンツおよびメタデータを3次元データベース124から入手し、エンドユーザ50の頭部姿勢データ（下記にさらに詳細に説明されるように、オーディオのための絶対メタデータをエンドユーザ50の頭部54に対して局所化するために使用されるであろう）を頭部／オブジェクト追跡サブシステム120から入手し、次いで、周囲環境内の実オブジェクトから生じる音と混合された空間化音への変換のために、スピーカ108に伝達される、そこからの空間化オーディオをレンダリングするために構成される。

【0082】

オーディオプロセッサ128はまた、下記にさらに説明されるであろうように、選択された仮想オブジェクトからの空間化オーディオデータとともに、仮想オブジェクト毎のエンドユーザ50の頭部54に対して局所化された任意の結果として生じるメタデータ（例えば、位置、配向、および音量データ）とともに、および、大域的メタデータ（例えば、拡張現実システム100またはエンドユーザ50によって大域的に設定された音量データ）とともに、続いて記録されるであろう、周囲環境内の選択された実オブジェクトからのみ生じるオーディオデータをマイクロホンアセンブリ116から入手するために構成される。

20

【0083】

拡張現実システム100はさらに、メモリ130と、ビデオおよびオーディオをメモリ130内に記憶するために構成される、レコーダ132と、エンドユーザ50または他のエンドユーザへの後続再生のために、ビデオおよびオーディオをメモリ130から読み出すために構成される、プレーヤ134とを備える。レコーダ132は、選択された仮想および実オブジェクトに対応する、空間化オーディオデータ（オーディオコンテンツオーディオデータとメタデータの両方）をオーディオプロセッサ128から入手し、本オーディオデータをメモリ130内に記憶し、選択された仮想および実オブジェクトと一致する仮想および実オブジェクトに対応する、ビデオデータ（ビデオコンテンツデータとメタデータの両方）をさらに入手する。プレーヤ134は、レコーダ132およびメモリ130が位置する同一ARシステム100に位置するように図示されるが、プレーヤは、サードパーティARシステム内に、またはさらにARシステム100によって以前に記録されたビデオおよびオーディオを再生する、スマートフォンまたはコンピュータ上に位置してもよいことを理解したい。

30

【0084】

40

ビデオプロセッサ126、オーディオプロセッサ128、レコーダ132、およびプレーヤ134の機能を実施する、制御サブシステムは、多種多様な形態のいずれかをとってもよく、いくつかのコントローラ、例えば1つ以上のマイクロコントローラ、マイクロプロセッサまたは中央処理ユニット（CPU）、デジタル信号プロセッサ、グラフィック処理ユニット（GPU）、他の集積回路コントローラ、例えば、特定用途向け集積回路（ASIC）、プログラマブルゲートアレイ（PGA）、例えば、フィールドPGA（FPGA）、および／またはプログラマブル論理コントローラ（PLU）を含んでもよい。

【0085】

ビデオプロセッサ126、オーディオプロセッサ128、レコーダ132、およびプレ

50

ーヤ 134 の機能は、それぞれ、単一の統合されたデバイスによって実施されてもよく、ビデオプロセッサ 126、オーディオプロセッサ 128、レコーダ 132、および／またはプレーヤ 134 の機能の少なくともいくつかは、単一の統合されたデバイスの中に組み合わせられてもよい、またはビデオプロセッサ 126、オーディオプロセッサ 128、レコーダ 132、またはプレーヤ 134 のそれぞれの機能は、いくつかのデバイス間に分散されてもよい。例えば、ビデオプロセッサ 126 は、仮想オブジェクトのビデオデータを 3 次元データベース 124 から入手し、そこからの合成ビデオフレームをレンダリングする、グラフィック処理ユニット (GPU) と、実オブジェクトのビデオフレームを前向きに面したカメラ 112 から入手する、中央処理ユニット (CPU) とを備えてもよい。同様に、オーディオプロセッサ 128 は、マイクロホンアセンブリ 116 およびユーザマイクロホン 122 から入手されたオーディオデータを処理する、デジタル信号プロセッサ (DSP) と、3 次元データベース 124 から入手されたオーディオデータを処理する、CPU とを備えてもよい。レコーダ 132 の記録機能およびプレーヤ 134 の再生機能は、CPU によって実施されてもよい。

【0086】

さらに、拡張現実システム 100 の種々の処理コンポーネントは、分散型サブシステム内に物理的に含有されてもよい。例えば、図 10a ~ 10d に図示されるように、拡張現実システム 100 は、有線導線または無線コネクティビティ 152 等によって、エンドユーザ 50 の頭部 54 に搭載されるコンポーネント（例えば、ディスプレイサブシステム 102 の投影サブシステム、マイクロホンアセンブリ 116、スピーカ 104、およびカメラ 114、118）に動作可能に結合される、ローカル処理およびデータモジュール 150 を備える。ローカル処理およびデータモジュール 150 は、フレーム構造 106（図 10a）に固定して取り付けられる、ヘルメットまたは帽子 106a（図 10b）に固定して取り付けられる、ヘッドホンに内蔵される、エンドユーザ 50 の胴体 58 に除去可能に取り付けられる（図 10c）、またはベルト結合式構成においてエンドユーザ 50 の腰部 59 に除去可能に取り付けられる（図 10d）等、種々の構成において搭載されてもよい。拡張現実システム 100 はさらに、遠隔モジュール 154、156 が、相互に動作可能に結合され、ローカル処理およびデータモジュール 150 へのリソースとして利用可能であるように、有線導線または無線コネクティビティ 158、160 等によって、ローカル処理およびデータモジュール 150 に動作可能に結合される、遠隔処理モジュール 154 および遠隔データリポジトリ 156 を備える。

【0087】

ローカル処理およびデータモジュール 150 は、電力効率的プロセッサまたはコントローラと、フラッシュメモリ等のデジタルメモリとを備えてもよく、両方とも、可能性として、処理または読み出後、ディスプレイサブシステム 102 への通過のために、センサから捕捉された、および／または遠隔処理モジュール 154 および／または遠隔データリポジトリ 156 を使用して入手および／または処理されたデータの処理、キャッシュ、および記憶を補助するために利用されてもよい。遠隔処理モジュール 154 は、データおよび／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 つ以上の比較的に強力なプロセッサまたはコントローラを備えてもよい。遠隔データリポジトリ 156 は、比較的大規模デジタルデータ記憶設備を備えてもよく、これは、インターネットまたは「クラウド」リソース構成における他のネットワーキング構成を通して利用可能であってもよい。一実施形態では、全てのデータは、記憶され、全ての算出は、ローカル処理およびデータモジュール 150 内で実施され、任意の遠隔モジュールから完全に自律的使用を可能にする。

【0088】

上記に説明される種々のコンポーネント間の結合 152、158、160 は、有線または光学通信を提供するための 1 つ以上の有線インターフェースまたはポート、または無線通信を提供するための RF、マイクロ波、および IR 等を介した 1 つ以上の無線インターフェースまたはポートを含んでもよい。いくつかの実装では、全ての通信は、有線であつ

10

20

30

40

50

てもよい一方、他の実装では、全ての通信は、ディスプレイサブシステム 102 内で使用される光ファイバを除き、無線であってもよい。なおもさらなる実装では、有線および無線通信の選択肢は、図 10a ~ 10d に図示されるものと異なり得る。したがって、有線または無線通信の特定の選択肢は、限定と見なされるべきではない。

【0089】

図示される実施形態では、ディスプレイサブシステム 102 の光源および駆動電子機器(図示せず)、オブジェクト追跡サブシステム 120 およびオブジェクト選択デバイス 110 の処理コンポーネント、およびオーディオプロセッサ 128 の DSP は、ローカル処理およびデータモジュール 150 内に含有されてもよい。ビデオプロセッサ 126 の GPU およびビデオプロセッサ 126 およびオーディオプロセッサ 128 の CPU は、遠隔処理モジュール 154 内に含有されてもよいが、代替実施形態では、これらのコンポーネントまたはその一部は、ローカル処理およびデータモジュール 150 内に含有されてもよい。3 次元データベース 124 およびメモリ 130 は、遠隔データリポジトリ 156 と関連付けられることができる。

【0090】

図 3 に図示されるオーディオプロセッサ 128 は、エンドユーザ 50 によって選択された仮想および実オブジェクトからのオーディオデータを処理および記録することにおいてさらに詳細に説明されるであろう。図 11 に示される例示的シナリオでは、エンドユーザ 50 (例えば、両親) が、仮想ドラマ - V2 オブジェクトと、実際のボーカル R2 、例えば、子供と、仮想ギタリスト V3 と、仮想ベーシスト V4 とを含む、4 人組バンドからの音を記録することを所望し、仮想テレビからの音を記録せずに、仮想テレビ V1 上のニュースまたはスポーツを監視することを所望し、さらに、例えば、誰かが料理をしている実際の台所 R1 からの音を記録することを所望していないとする。

【0091】

図 12 に図示される実施形態では、オーディオプロセッサ 128 の機能は、仮想オブジェクトから生じるオーディオを処理する、CPU 180 と、実オブジェクトから生じるオーディオを処理する、DSP 182 との間に分散される。CPU 180 は、個々の仮想オブジェクト V1 - Vn に対応する空間化オーディオデータ EX - V1 ~ EX - Vn を生成するために構成される、1 つ以上の特殊効果モジュール 184 (この場合、特殊効果モジュール 1 - n) を備える。この目的を達成するために、特殊効果モジュール 184 は、仮想オブジェクト V1 - Vn に対応するオーディオコンテンツデータ AUD - V1 ~ AUD - Vn および絶対メタデータ MDa - V1 ~ MDa - Vn を 3D データベース 124 から、および頭部姿勢データを頭部 / オブジェクト追跡サブシステム 120 から入手し、頭部姿勢データに基づいて、絶対メタデータ MDa - V1 ~ MDa - Vn をエンドユーザ 50 の頭部 54 に対して局所化し、局所的メタデータ (例えば、位置、配向、および音量データ) をオーディオコンテンツデータに適用し、仮想オブジェクト V1 - Vn のための空間化オーディオデータを生成する。

【0092】

CPU 180 はさらに、個別の特殊効果モジュール 184 から受信された空間化オーディオデータ EX - V1 ~ EX - Vn を混合し、混合されたオーディオデータ EX を取得するために構成される、ミキサ 186 と、大域的メタデータ MD - OUT (例えば、大域的音量) を混合された空間化オーディオデータに適用し、複数の音チャネルを通してスピーカ 108 に出力される、最終空間化オーディオ AUD - OUT EX を取得するために構成される、大域的特殊効果モジュール 188 とを備える。

【0093】

重要なこととして、特殊効果モジュール 184 は、メモリ 130 (図 2 に示される) 内への記憶のために、オブジェクト選択デバイス 110 を介してエンドユーザ 50 によって選択された仮想オブジェクトから生じるオーディオコンテンツデータと、これらの選択された仮想オブジェクトに対応するメタデータ (局所的および / または絶対) とをレコーダ 132 に送信するために構成され、大域的特殊効果モジュール 188 は、メモリ 130 内

への記憶のために、大域的メタデータMD-OUTをレコーダ132に送信するために構成される。例示的実施形態では、仮想オーディオコンテンツデータAUD-V2（すなわち、仮想ドラマ）、AUD-V3（すなわち、仮想ギタリスト）、AUD-V4（すなわち、仮想ベーシスト）は、記録のために選択される一方、オーディオコンテンツデータAUD-V1（すなわち、仮想テレビ）は、記録のために選択されない。したがって、オーディオコンテンツデータAUD-V2、AUD-V3、およびAUD-V4および対応する局所的メタデータMD-V2、MD-V3、およびMD-V4は、図13に示されるように、メモリ130内に記憶される。

【0094】

代替実施形態では、選択された仮想オブジェクトからのオーディオコンテンツデータおよび対応する局所的/絶対メタデータおよび大域的メタデータをメモリ130内に個々に記憶する代わりに、またはそれに加え、CPU180は、加えて、選択された仮想オブジェクトAUD-V2、AUD-V3、およびAUD-V4のみに対応する空間化オーディオデータEFX-V2、EFX-V3、EFX-V4を混合し、大域的メタデータMD-OUTを本混合された空間化オーディオデータに適用し、選択された仮想オブジェクトAUD-V2、AUD-V3、およびAUD-V4からのみのオーディオを含む、空間化オーディオを取得することによって生成される、空間化オーディオを出力する。しかしながら、この場合、付加的オーディオ混合機能が、CPU180の中に組み込まれる必要があるであろう。

【0095】

DSP182は、特定の方向から、この場合、オブジェクト選択デバイス110を介して、エンドユーザ50によって選択された各実オブジェクトの方向から、マイクロホンアセンブリ116によって受信された音を優先的に表す、マイクロホンアセンブリ116から入手されたオーディオ信号を処理し、オーディオ信号を出力するために構成される。実オブジェクトの位置および/または配向は、エンドユーザ50の頭部54に対して移動し得るため、実オブジェクト追跡データは、エンドユーザ50の頭部54に対する実オブジェクトの位置および/または配向の任意の変化が考慮され得、DSP182が、オーディオ出力を動的に修正し、比較的に移動する実オブジェクトの方向からマイクロホンアセンブリ116によって受信された音を優先的に表し得るように、頭部/オブジェクト追跡サブシステム120から受信されてもよい。例えば、実オブジェクトが選択されたとき、エンドユーザ50が、その頭部54を頭部54の配向に対して反時計回りに90度移動させる場合、DSP182から出力されたオーディオの優先的方向は、時計回りに90度動的に偏移されることができる。

【0096】

図14を参照すると、マイクロホンアセンブリ116のマイクロホン要素118は、それぞれ、周囲音信号を検出し、オーディオ信号に変換するために構成される、マイクロホン要素の位相アレイ（この場合、マイクロホン要素M1～Mn）の形態をとる。図示される実施形態では、マイクロホン要素118は、性質上、デジタルであって、したがって、周囲音信号をデジタルオーディオ信号、この場合、パルス密度変調（PDM）信号に変換する。好ましくは、マイクロホン要素118は、相互から離間され、オーディオ出力の指向性を最大限にする。例えば、図2に示されるように、マイクロホン要素118のうちの2つが、フレーム構造106の各アームに搭載されてもよいが、4つ等の2つを上回るマイクロホン要素118が、フレーム構造106の各アームに搭載されてもよい。代替として、フレーム構造106は、マイクロホン要素118がフレーム構造106の正面および背面上に搭載され得るように、設計されてもよい。このように、マイクロホン要素118のアレイは、エンドユーザ50の頭部54を包囲し、音の潜在的源の全ての方向を網羅し得る。

【0097】

マイクロホンアセンブリ116はさらに、それぞれ、個別のデジタルオーディオ信号を対応するマイクロホン要素118から受信し、「デシメーション」と称されるデジタルフ

10

20

30

40

50

イルタ動作を実施し、デジタルオーディオ信号を P D M フォーマットからより容易に操作可能なパルスコード変調 (P C M) に変換するために構成される、複数のデジタルマイクロホンインターフェース (D M I C) 1 9 0 (この場合、マイクロホン要素 M 毎に 1 つずつの D M I C 1 ~ D M I C n) を備える。 D M I C 1 9 0 はそれぞれまた、固定利得制御をデジタルオーディオ信号上で実施する。

【 0 0 9 8 】

D S P 1 8 2 は、それぞれ、マイクロホンアセンブリ 1 1 6 によって出力されるデジタルオーディオ信号を処理し、選択された実オブジェクト (R 1 ~ R m のうちの 1 つ) の方向にマイクロホンアセンブリ 1 1 6 によって受信された音を優先的に表す、指向性オーディオ信号 A U D - R (指向性オーディオ信号 A U D - R 1 ~ A U D - R m のうちの 1 つ) を出力するために構成される、複数のオーディオ処理モジュール 2 0 0 を備える。個別のオーディオ処理モジュール 2 0 0 によって出力される指向性オーディオ信号 A U D - R 1 ~ A U D - R m は、指向性オーディオ出力 A U D - O U T M I C の中に組み合わせられ、これは、全ての選択された実オブジェクトから生じる音を優先的に表す。図示される実施形態では、D S P 1 8 2 は、オブジェクト選択デバイス 1 1 0 を介してエンドユーザ 5 0 によって選択された実オブジェクト毎に、オーディオ処理モジュール 2 0 0 の 1 つのインスタンスを作成する。

【 0 0 9 9 】

この目的を達成するために、オーディオ処理モジュール 2 0 0 はそれぞれ、複数の遅延要素 1 9 4 (この場合、マイクロホン要素 M 毎に 1 つずつの遅延要素 D 1 ~ D n) 、複数の利得要素 1 9 6 (この場合、マイクロホン要素 M 毎に 1 つずつの利得要素 G 1 ~ G n) 、および総和器 1 9 8 の形態において、処理パラメータを備える。遅延要素 1 9 4 は、それぞれ、遅延係数をマイクロホンアセンブリ 1 1 6 の対応する利得増幅器 1 9 2 から受信された増幅されたデジタル信号に適用し、利得要素 1 9 6 は、それぞれ、利得係数を遅延されたデジタル信号に適用する。総和器 1 9 8 (S) は、利得調節および遅延された信号を加算し、それぞれ、個別の指向性オーディオ信号 A U D - R を生成する。

【 0 1 0 0 】

マイクロホン要素 1 1 8 は、空間的に配列され、各オーディオ処理モジュール 2 0 0 の遅延要素 1 9 4 および利得要素 1 9 6 は、指向性極性パターンに従って周囲音の受信をもたらす様式において、マイクロホンアセンブリ 1 1 6 から受信されたデジタルオーディオ信号に適用される (すなわち、特定の角度方向または複数の方向から到着した音は、他の角度方向から到着した音より強調されるであろう) 。 D S P 1 8 2 は、遅延要素 1 9 4 の遅延係数および利得要素 1 9 6 の利得係数を変化させることによって、指向性オーディオ信号 A U D - R 1 ~ A U D - R m 、したがって、組み合わせられた指向性オーディオ出力 A U D - O U T M I C の指向性を修正するために構成される。

【 0 1 0 1 】

したがって、オーディオ出力 A U D - O U T M I C の指向性は、選択された実オブジェクトに基づいて修正される、例えば、音が優先的に受信される方向または複数の方向は、選択された実オブジェクトまたは源の方向に沿って設定されてもよいことが理解され得る。

【 0 1 0 2 】

例えば、図 1 5 a を参照すると、それぞれ、2 つの特定の方向 D a および D b に沿った2 つの実オブジェクト R a および R b が、選択される場合、D S P 1 8 2 は、オーディオ処理モジュール 2 0 0 の 2 つのインスタンスを生成し、これらのオーディオ処理モジュール 2 0 0 のそれぞれ内において、実オブジェクト R a および R b の方向 D a および D b と整合される 2 つのロープを有する、受信利得パターンが、生成されるように、各オーディオ処理モジュール 2 0 0 内の遅延要素 1 9 4 および利得要素 1 9 6 の全てのための個別の遅延係数および利得係数を選択するであろう。エンドユーザ 5 0 の頭部 5 4 に対する実オブジェクト R a および R b の配向が、変化する場合、実オブジェクト R a および R b の特定の方向は、変化し得、その場合、D S P 1 8 2 は、受信利得パターンが、図 1 5 b に図

10

20

30

40

50

示されるように、方向 D_c および D_d と整合される 2 つのロープを有するように、各オーディオ処理モジュール 200 内の遅延要素 194 および利得要素 196 の全てのための異なる遅延係数および利得係数を選択してもよい。

【0103】

オーディオ出力 AUD - OUT MIC の指向性のそのような動的修正を促進するためには、遅延 / 利得値の異なるセットおよび対応する優先的方向が、DSP182 によるアクセスのために、メモリ 130 内に記憶されてもよい。すなわち、DSP182 は、各選択された実オブジェクト R の方向とメモリ 130 内に記憶される最も近い指向性値を合致させ、その選択された方向のための遅延 / 利得係数の対応するセットを選択する。

【0104】

マイクロホン要素 118 は、デジタルとして説明されるが、マイクロホン要素 118 は、代替として、アナログであってもよいことに留意されたい。さらに、遅延要素 194、利得要素 196、および総和器 198 は、DSP182 内に常駐するソフトウェアコンポーネントとして開示および図示されるが、遅延要素 194、利得要素 196、および総和器 198 のうちの任意の 1 つ以上のものは、DSP182 外に常駐するが、その制御下にある、アナログハードウェアコンポーネントを備えてよい。しかしながら、ソフトウェアベースのオーディオ処理モジュール 200 の使用は、いくつかの明確に異なる実オブジェクトからの音が、同時に、優先的に受信および処理されることを可能にする。

【0105】

図 12 に従って参照すると、DSP182 はまた、音声データをユーザマイクロホン 122 から受信し、それと指向性オーディオ出力 AUD - OUT MIC を組み合わせる。随意の実施形態では、DSP182 は、音響エコーキャンセル (AEC) および雑音抑制 (NS) 機能を仮想オブジェクトから生じるスピーカ 108 からの音に対して実施するために構成される。すなわち、マイクロホンアセンブリ 116 は、音が優先的に受信される方向がスピーカ 108 と一致し得ない場合でも、スピーカ 108 によって放出される音を感知し得る。この目的を達成するために、大域的特殊効果モジュール 188 によってスピーカ 108 に出力される空間化オーディオデータはまた、DSP182 の中に入力され、これは、空間化オーディオデータを使用して、スピーカ 108 によってマイクロホンアセンブリ 116 に出力される結果として生じる音（雑音と見なされる）を抑制し、スピーカ 108 からマイクロホンアセンブリ 116 へのフィードバックから生じる任意のエコーをキャンセルする。

【0106】

重要なこととして、DSP182 はさらに、オーディオコンテンツデータとしてのメモリ 130（図 2 に示される）内への記憶のために、指向性オーディオ出力 AUD - OUT MIC および局所的メタデータ（例えば、指向性オーディオ出力 AUD - OUT MIC が生じた実オブジェクトの場所および配向）をレコーダ 132 に送信するために構成される。図 11 に図示される例示的実施形態では、局所的メタデータは、実オブジェクト AUD - R2（すなわち、実際のボーカル）に対応する。したがって、指向性オーディオ出力 AUD - OUT MIC（実オブジェクト AUD - R2 に優先的に対応する）および対応する局所的メタデータ MD - R2 は、図 13 に示されるように、メモリ 130 内に記憶される。

【0107】

随意の実施形態では、指向性オーディオ出力 AUD - OUT MIC（空間化されてもよい）は、エンドユーザ 50 への再生のために、スピーカ 108 または他のスピーカの中に入力されてもよい。指向性オーディオ出力 AUD - OUT MIC は、音の明確性または現実性に影響を及ぼすように、音が実オブジェクトの場所から生じるように、仮想源から生じる空間化オーディオデータがエンドユーザ 50 に現れるものと同一様式で空間化されてもよい。すなわち、局所的メタデータ（例えば、指向性オーディオ出力 AUD - OUT MIC が優先的に生じた実オブジェクトの場所および配向）は、指向性オーディオ出力 AUD - OUT MIC に適用され、空間化オーディオデータを取得してもよい。

10

20

30

40

50

【0108】

別の随意の実施形態では、実オブジェクトまたはさらにエンドユーザ50によって選択された仮想オブジェクトから生じる音は、プロファイル化されてもよい。特に、DSP182は、標的音のタイプを決定するために、選択されたオブジェクトからの音の特性と他の実オブジェクトから生じる音の特性を分析および比較してもよい。DSP182は、次いで、所望に応じて、レコーダ132によってメモリ130（図2に示される）の中に記録するために、これらの実オブジェクトから生じる全てのオーディオデータを指向性オーディオ出力AUD-OUT MIC内に含めることができる。例えば、エンドユーザ50が、音楽オブジェクト（AUD-V2、AUD-V3、AUD-V4、AUD-R2）のいずれかを選択した場合、DSP182は、マイクロホンアセンブリ116を制御し、全ての音楽実オブジェクトを優先的に感知することができる。 10

【0109】

図示される実施形態では、DSP182は、エンドユーザ50によって選択された実オブジェクト198が、ディスプレイサブシステム102の視野外に移動した場合でも（頭部／オブジェクト追跡サブシステム120から受信された実オブジェクト追跡データによって示されるように）、メモリ130内への記録のために、指向性オーディオ出力AUD-OUT MICをレコーダ130に出力することを継続する。代替実施形態では、DSP182は、エンドユーザ50によって選択された実オブジェクト198が、ディスプレイサブシステム102の視野外に移動するとすぐに、メモリ130内への記録のために指向性オーディオ出力AUD-OUT MICをレコーダ130に出力することを中止し、エンドユーザ50によって選択された実オブジェクト198が、ディスプレイサブシステム102の視野の中に戻るとすぐに、メモリ130内への記録のために指向性オーディオ出力AUD-OUT MICをレコーダ130に出力することを再開する。 20

【0110】

オーディオプロセッサ128（図示される実施形態では、CPU180およびDSP182）が、メモリ130内への記憶のために、選択された仮想オブジェクトおよび実オブジェクトから生じるオーディオコンテンツデータ（例示的の場合では、オーディオコンテンツデータAUD-V2、AUD-V3、およびAUD-V4、およびAUD-MIC）および局所的メタデータ（例示的の場合では、MD-V2、MD-V3、MD-V4、およびMD-R2）および大域的メタデータ（MD-OUT）をレコーダ132に送信するものと同様に、ビデオプロセッサ126は、図13に図示されるように、仮想オブジェクトおよび実オブジェクトから生じるビデオコンテンツデータ（例示的の場合では、ビデオコンテンツデータVID-V2、VID-V3、VID-V4、およびVID-R2）を送信してもよい。仮想オブジェクトの場合、ビデオプロセッサ126は、単に、さらなる処理を伴わずに、仮想オブジェクトを3Dデータベース124から入手し、メモリ130内への記憶のために、これらの仮想オブジェクトをレコーダ132に送信する。実オブジェクトの場合、ビデオプロセッサ126は、選択された実オブジェクトのいずれかをカメラ112から入手されたビデオから抽出または「切り取り」、これらの実オブジェクトを仮想オブジェクトとしてメモリ130内に記憶してもよい。図11に図示される例示的場合では、実際のボーカルR2のためのビデオは、仮想オブジェクトVID-R2として記録されてもよい。随意の実施形態では、ビデオプロセッサ126は、メモリ130内への記憶のために、カメラ112から入手されたビデオ全体（選択されていない仮想および実オブジェクトに対応するビデオを含む）をレコーダ132に送信する。 30 40

【0111】

プレーヤ134は、メモリ130内に記録されたビデオおよび／またはオーディオを、ビデオ／オーディオを記録したオリジナルエンドユーザ50またはサードパーティユーザであり得る、再生ユーザ50'（図16aに示される）に再生するために構成される。オーディオ／ビデオは、再生ユーザ50'によって与えられるコマンド、例えば、ユーザマイクロホン122を介した音声コマンドに応答して、プレーヤ134によって選択的に再生されてもよい。例えば、再生ユーザ50'は、「仮想オーディオオン／オフ」コマンド 50

を使用して、仮想オーディオ再生をオンまたはオフにする、または「ディスプレイオン／オフ」コマンドを使用して、仮想ビデオ再生をオンまたはオフにする、または「実オーディオオン／オフ」コマンドを使用して、実オーディオ再生をオンまたはオフにしてもよい。

【0112】

図示される実施形態では、オーディオプロセッサ128は、オーディオコンテンツデータおよびメタデータ（選択された仮想および実オブジェクトに対応する）をメモリ130から読み出し、オーディオコンテンツデータおよびメタデータからの空間化オーディオをレンダリングし、再生スピーカ108を介したユーザ50'への再生のために、空間化オーディオをプレーヤ134に伝達する。混合空間化オーディオデータ（コンテンツおよびメタデータの代わりに）が記憶される、代替実施形態では、プレーヤ134は、単に、オーディオデータを再レンダリングせずに、または別様にさらに処理せずに、再生ユーザ50'に再生するために、オーディオデータをメモリ130から入手してもよい。

10

【0113】

さらに、図示される実施形態では、ビデオプロセッサ126は、ビデオコンテンツデータおよびメタデータ（選択された仮想および実オブジェクトに対応する）を読み出し、ビデオコンテンツデータおよびメタデータからのビデオをレンダリングし、スピーカ108を介したオーディオの再生と同期したディスプレイサブシステム102を介した再生ユーザ50'への再生のために、ビデオをプレーヤ134に伝達する。隨意に、カメラ112によって捕捉されたビデオデータの全てが記憶される場合、プレーヤ134は、単に、ビデオデータをレンダリングせずに、または別様にさらに処理せずに、再生ユーザ50'に再生するために、ビデオデータをメモリ130から入手してもよい。拡張現実システム10は、再生ユーザ50'に、選択された仮想および実オブジェクトに対応するビデオのみの再生またはカメラ112によって捕捉されたビデオ全体の再生のいずれかのオプションを提供してもよい。

20

【0114】

一実施形態では、再生ユーザ50'の現在の頭部姿勢は、ビデオ／オーディオの再生の間、考慮されない。代わりに、ビデオ／オーディオは、ビデオ／オーディオデータの記録の間に元々検出された頭部姿勢を使用して、再生ユーザ50'に再生され、これは、オーディオ／ビデオコンテンツデータとともにメモリ130内に記憶される局所的メタデータ内に反映されるであろう、または混合空間化オーディオが、メタデータを伴わずに記録される場合、頭部姿勢は、メモリ130内に記憶される混合空間化オーディオ内に反映されるであろう。この場合、再生ユーザ50'は、オリジナルエンドユーザ50がビデオ／オーディオを体験したものと同一様式において、ビデオ／オーディオを体験するであろうが、但し、オリジナルエンドユーザ50によって選択された仮想および実オブジェクトから生じるオーディオのみおよび隨意にビデオのみが、再生されるであろう。この場合、再生ユーザ50'は、再生ユーザ50'の頭部姿勢が考慮されるであろうため、拡張現実に没入し得ない。むしろ、再生ユーザ50'は、ヘッドセットを使用して、オーディオ再生を体験してもよい（したがって、オーディオは、環境によって影響されないであろう）、または再生ユーザ50'は、静かな部屋内でオーディオ再生を体験してもよい。

30

【0115】

代替実施形態では、再生ユーザ50'の現在の頭部姿勢は、ビデオ／オーディオの再生の間、考慮されてもよい。この場合、ビデオ／オーディオの記録の間の再生ユーザ50'の頭部姿勢は、再生の間に検出された再生ユーザ50'の現在の頭部姿勢が、ビデオ／オーディオデータを再レンダリングするために使用されるであろうため、ビデオ／オーディオコンテンツデータとともにメモリ130内に記憶されるメタデータの中に組み込まれる必要はない。代わりに、メモリ130内に記憶される絶対メタデータ（例えば、3D場面内のこれらの仮想オブジェクトの音量および絶対位置および配向、および仮想源の近傍の任意の仮想または実オブジェクトを含む、各仮想オブジェクトを囲繞する空間音響、部屋寸法、壁／床材料等）が、再生ユーザ50'の現在の頭部姿勢を使用して、再生ユーザ5

40

50

0'の頭部姿勢に対して局所化され、次いで、オーディオ／ビデオをレンダリングするために使用されるであろう。したがって、再生ユーザ50'は、ビデオ／オーディオの再生の間、拡張現実に没入するであろう。

【0116】

再生ユーザ50'は、ビデオ／オーディオが記録された（例えば、「同一物理的部屋」）オリジナル空間環境内で拡張現実を体験してもよい、または新しい物理的または仮想空間環境（例えば、「異なる物理的または仮想部屋」）内で拡張現実を体験してもよい。

【0117】

拡張現実が、ビデオ／オーディオが記録されたオリジナル空間環境内で再生ユーザ50'によって体験される場合、選択されたオブジェクトと関連付けられた絶対メタデータは、空間化オーディオの正確な再生のために修正される必要はない。対照的に、拡張現実が、新しい空間環境内で再生ユーザ50'によって体験される場合、オブジェクトと関連付けられた絶対メタデータは、新しい空間環境でのオーディオ／ビデオの正確なレンダリングのために修正される必要があり得る。

【0118】

例えば、例示的実施形態では、仮想オブジェクトAUD-V2（すなわち、仮想ドramaー）、AUD-V3（すなわち、仮想ギタリスト）、AUD-V4（すなわち、仮想ベーシスト）、および実オブジェクト（すなわち、実際のボーカル）からのオーディオ／ビデオコンテンツは、図16aに図示されるように、小部屋250内で記録されてもよい。仮想オブジェクトAUD-V2（すなわち、仮想ドramaー）、AUD-V3（すなわち、仮想ギタリスト）、AUD-V4（すなわち、仮想ベーシスト）、および実オブジェクト（すなわち、実際のボーカル）から以前に記録されたオーディオは、図16bに図示されるように、コンサートホール252で再生されてもよい。拡張現実システム10は、オブジェクトをコンサートホール252内の任意の場所に再位置付けしてもよく、コンサートホール252内の各オブジェクトの新しい位置およびコンサートホール252内の各オブジェクトを囲繞する空間音響を含む、絶対メタデータが、生成または別様に入手されてもよい。本絶対メタデータは、次いで、再生ユーザ50'の現在の頭部姿勢を使用して、局所化され、次いで、再生ユーザ50'への再生のために、オーディオおよびビデオをコンサートホール252内にレンダリングするために使用することができる。

【0119】

拡張現実システム100の配列および機能を説明したので、拡張現実システム100を使用して、少なくとも1つのオブジェクトを選択し、これらの選択されたオブジェクトからのオーディオおよびビデオを記録する1つの方法300が、ここで、図17に関して説明されるであろう。

【0120】

最初に、エンドユーザ50が、オブジェクト選択デバイス110を介して、空間環境内の少なくとも1つのオブジェクト（例えば、実際および／または仮想）を持続的に選択する（ステップ302）。オブジェクトは、例えば、3次元カーソル62（図5に示される）をエンドユーザ50の視野60内で移動させ、3次元カーソル62を用いて、オブジェクトを選択することによって、エンドユーザ50の視野60内で選択されることができる。または、オブジェクトは、手のジェスチャ（図6または7に示される）を使用して、または音声コマンドを使用して、選択されることができる。複数のオブジェクトは、個々に選択されてもよい、または、例えば、線64をオブジェクトの周囲に引くことによって（図8に示される）、またはエンドユーザ50の視野60の角度範囲66（エンドユーザ50の視野60の角度範囲全体未満であってもよい）（図9に示される）を定義し、エンドユーザ50の視野60の定義された角度範囲66内のオブジェクトの全てを選択することによって、大域的に選択されてもよい。

【0121】

次に、空間環境内の全ての仮想オブジェクトのためのオーディオおよびビデオコンテンツ、および仮想オブジェクトと関連付けられた絶対メタデータが、入手される（ステップ

10

20

30

40

50

304)。次に、エンドユーザ50の現在の頭部姿勢が、追跡され(ステップ306)、絶対メタデータが、現在の頭部姿勢データを使用して、エンドユーザ50の頭部54に対して局所化され(ステップ308)、仮想オブジェクトのオーディオおよびビデオコンテンツに適用され、個別の仮想オブジェクトの全てのためのビデオデータおよび空間化オーディオデータを取得する(ステップ310)。3D場面内の個別の仮想オブジェクトの全てのための空間化オーディオデータは、混合され(ステップ312)、大域的メタデータは、混合空間化オーディオデータに適用され、3D場面内の全ての仮想オブジェクトのための最終空間化オーディオを取得し(ステップ314)、これは、次いで、エンドユーザ50による知覚のための音に変換される(ステップ316)。次に、ステップ310において取得されたビデオデータは、エンドユーザ50による知覚のための画像フレームに変換される(ステップ318)。次に、ステップ302においてエンドユーザ50によって選択された全ての仮想オブジェクトのためのオーディオ/ビデオコンテンツおよび全ての関連付けられたメタデータ(絶対および局所的メタデータの両方)が、記録される(ステップ320)。

【0122】

ステップ304～320と並行して、エンドユーザ50の頭部54に対する選択された実オブジェクトの位置および/または配向が、追跡され(ステップ322)、選択された実オブジェクトから生じる音は、実オブジェクトの追跡された位置および配向に基づいて、優先的に感知される(ステップ324)。次に、選択された実オブジェクトの画像が、捕捉され(ステップ326)、随意に、仮想ビデオコンテンツに変換される。次に、選択された実オブジェクトからの優先的に感知された音と関連付けられたオーディオコンテンツおよび選択された実オブジェクトの捕捉された画像と関連付けられたビデオコンテンツ、および選択された実オブジェクト毎の全ての関連付けられたメタデータ(実オブジェクトの場所および配向)が、記録される(ステップ328)。

【0123】

再生ユーザ50'のために、拡張現実システム100を使用して、少なくとも1つのオブジェクトの以前に記録されたオーディオおよびビデオを再生する1つの方法400が、ここで、図18に関して説明されるであろう。そのようなオーディオおよびビデオは、上記の図17の方法300に説明される様式において、オーディオコンテンツデータおよびビデオコンテンツデータとして事前に記録されていてもよい。オブジェクトは、実際および/または仮想のものであってもよく、エンドユーザ50によって持続的に選択されてもよい。例示的方法400では、オーディオおよびビデオは、図16aおよび16bに関して説明されるように、小部屋250等のオリジナル空間環境内で事前に記録され、コンサートホール252等のオリジナル空間環境と異なる新しい空間環境内で再生されてもよい。

【0124】

最初に、以前に記録されたオーディオコンテンツデータおよびビデオコンテンツデータが、入手される(ステップ402)。新しい空間環境が、少なくとも部分的に仮想である場合、新しい空間環境と関連付けられた付加的仮想コンテンツ(オーディオまたはビデオのいずれか)もまた、入手されてもよい。次いで、オブジェクトは、新しい空間環境内に再位置付けられ、これは、再生ユーザ50'からの入力に応答してもよい(ステップ404)。次いで、新しい空間環境内に位置付けられるオブジェクトに対応する絶対メタデータが、入手され(ステップ406)、再生ユーザ50'の頭部姿勢が、新しい空間環境内で追跡され(ステップ408)、絶対メタデータが、再生ユーザ50'の追跡された頭部姿勢に基づいて、再生ユーザ50'に対して局所化される(ステップ410)。次に、オーディオおよびビデオが、新しい空間環境内の局所的メタデータに基づいて、読み出されたオーディオコンテンツデータおよびビデオコンテンツデータからレンダリングされる(ステップ412)。レンダリングされたオーディオおよびビデオは、次いで、再生ユーザ50'による同期知覚のために、それぞれ、音および画像フレームに変換される(ステップ414)。

10

20

30

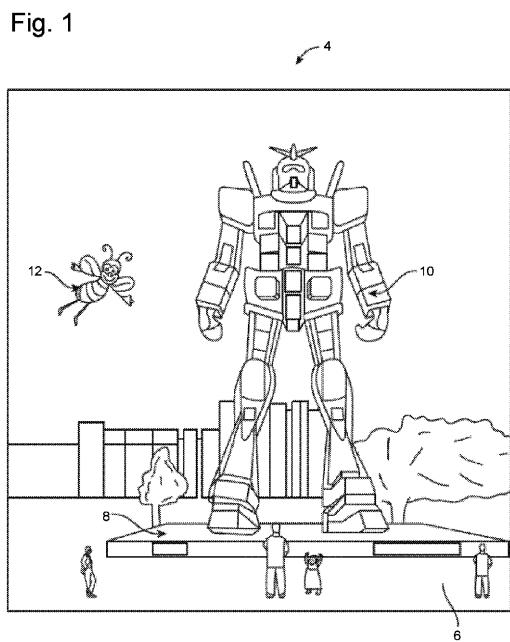
40

50

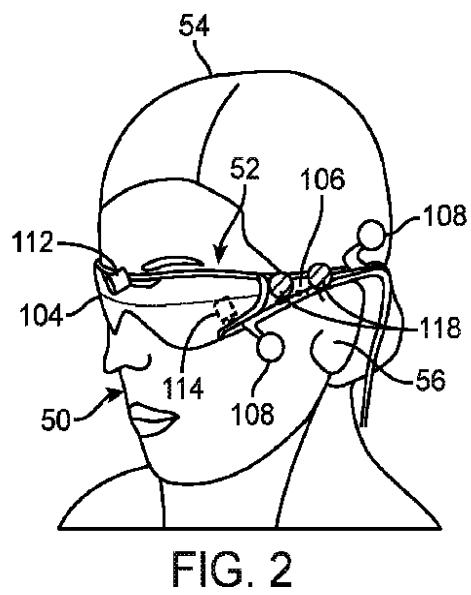
【0125】

前述の明細書では、本発明は、その具体的実施形態を参照して説明された。しかしながら、種々の修正および変更が、本発明のより広義の精神および範囲から逸脱することなく、そこに成されてもよいことは、明白であろう。例えば、上記に説明されるプロセスフローは、プロセスアクションの特定の順序を参照して説明されている。しかしながら、説明されるプロセスアクションの多くの順序は、本発明の範囲または動作に影響を及ぼすことなく、変更されてもよい。明細書および図面は、故に、限定的意味ではなく、例証的意味と見なされるべきである。

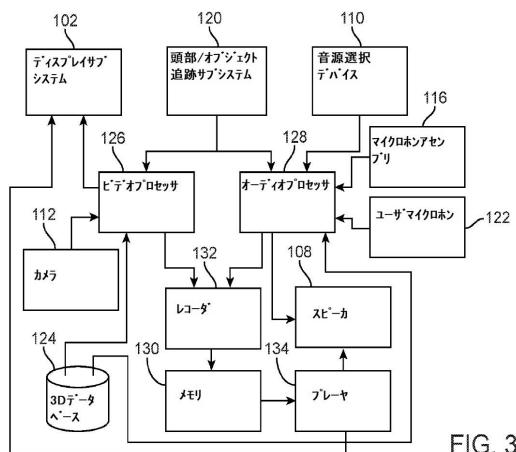
【図1】



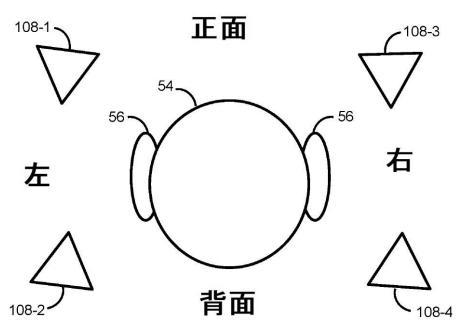
【図2】



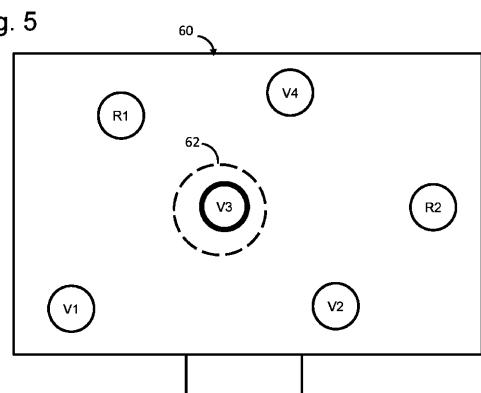
【図3】



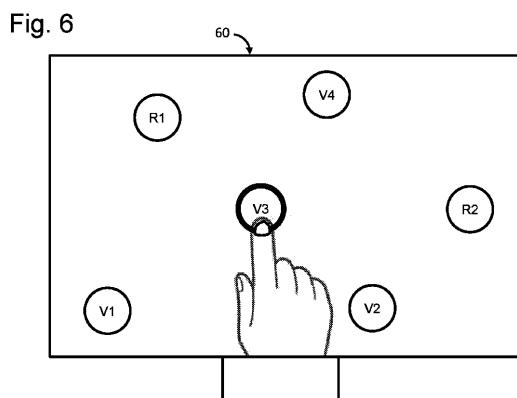
【図4】



【図5】



【図6】



【図7】

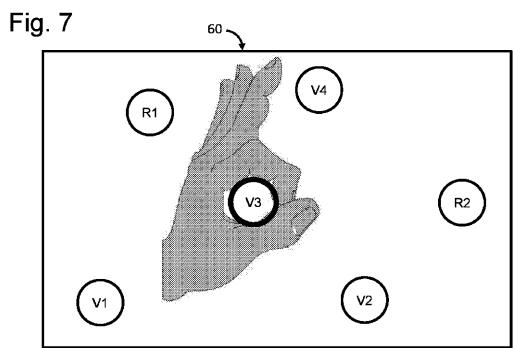
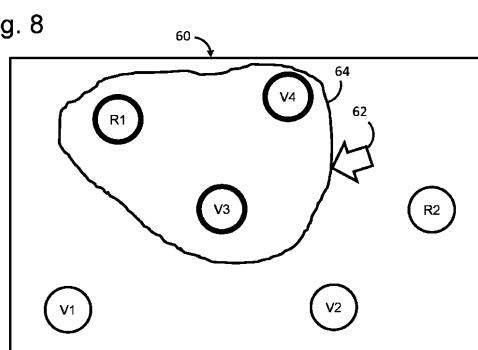
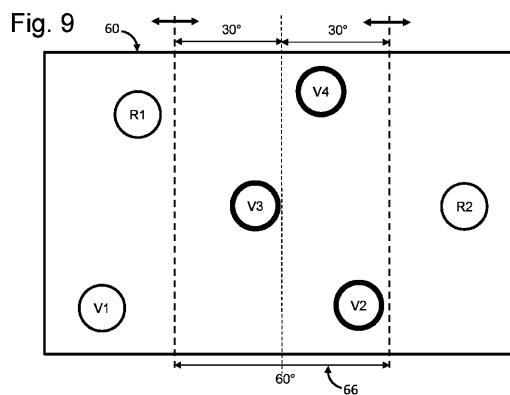


Fig. 5

【図8】



【図9】



【図 10 a】

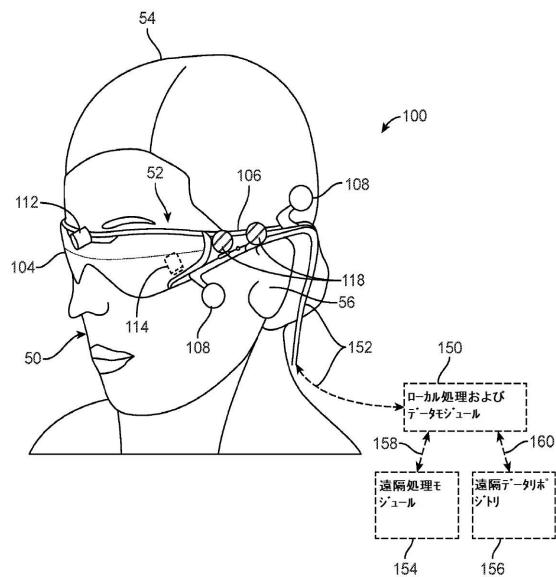


FIG. 10a

【図 10 b】

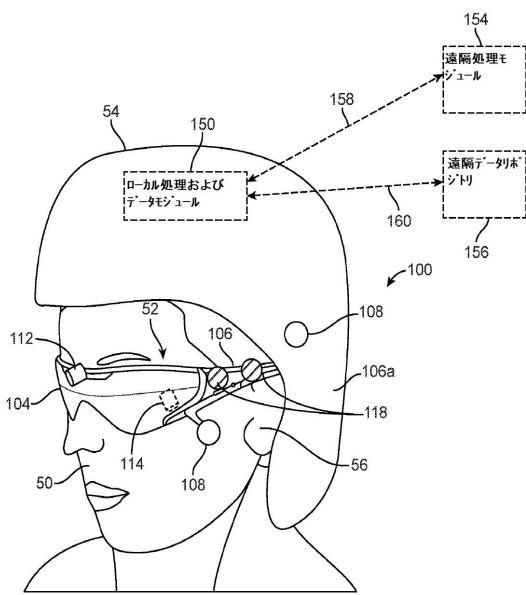


FIG. 10b

【図 10 c】

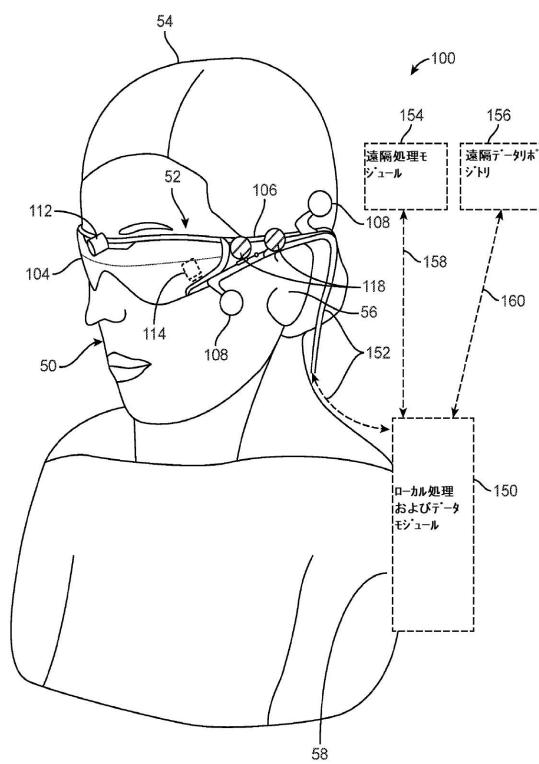


FIG. 10c

【図 10 d】

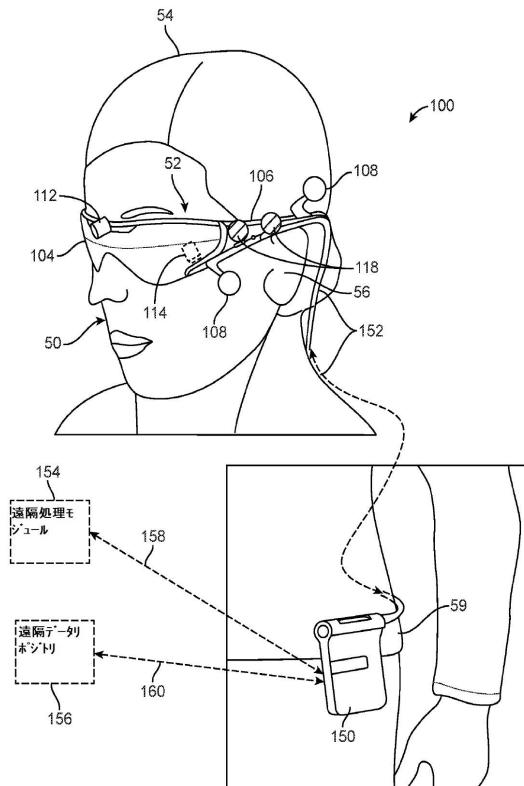


FIG. 10d

【図 1 1】

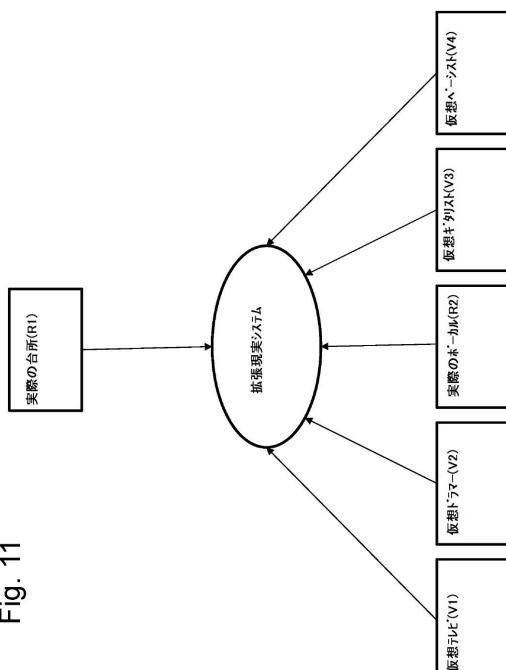
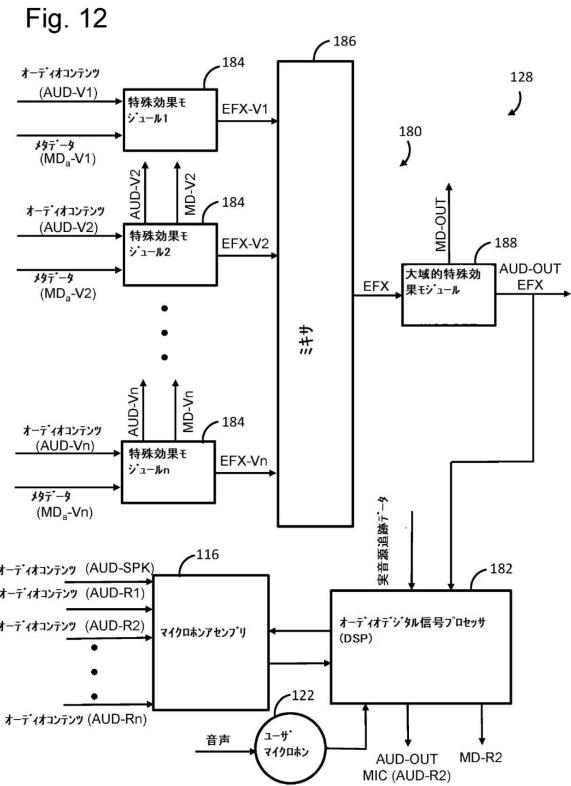
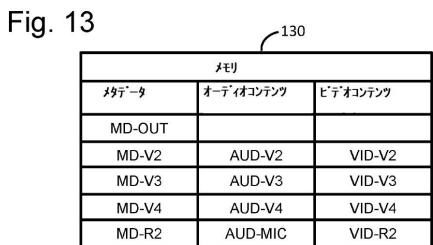


Fig. 11

【図 1 2】



【図 1 3】



【図 1 4】

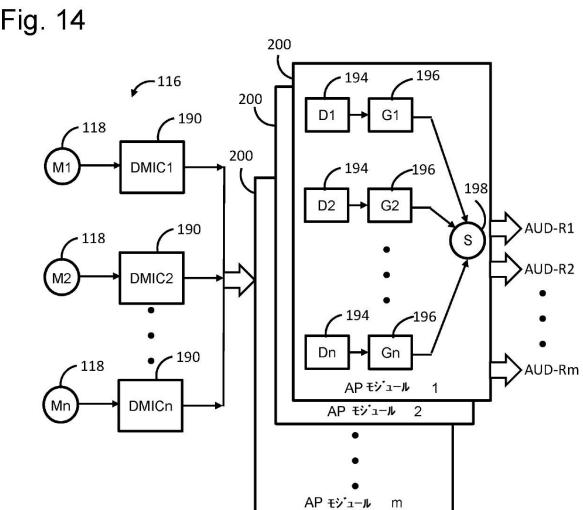
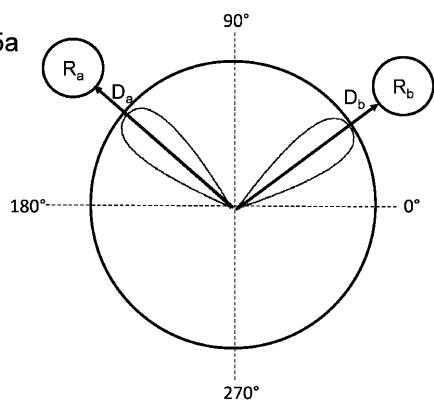


Fig. 14

【図 1 5 a】



【図 1 5 b】

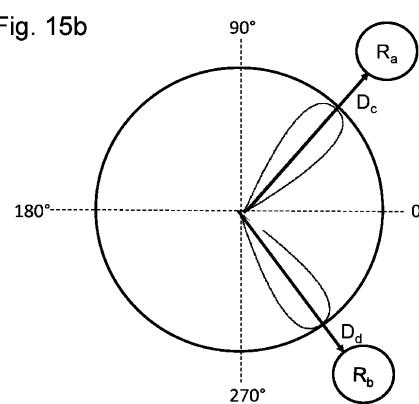
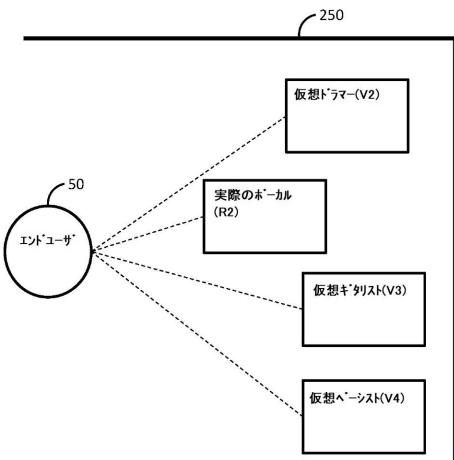


Fig. 15b

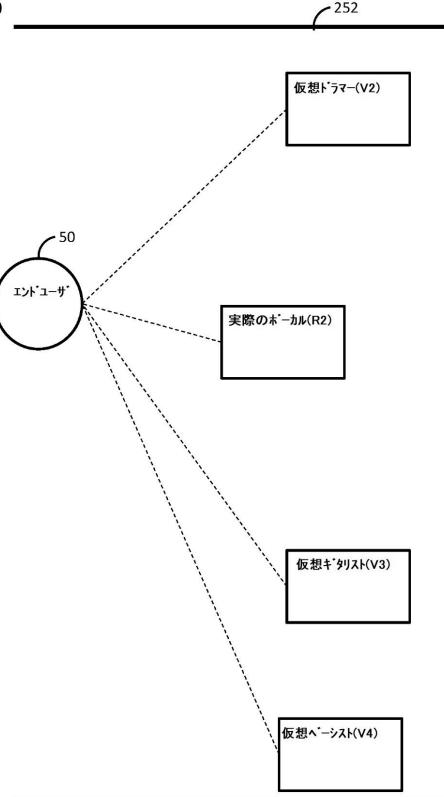
【図 16 a】

Fig. 16a



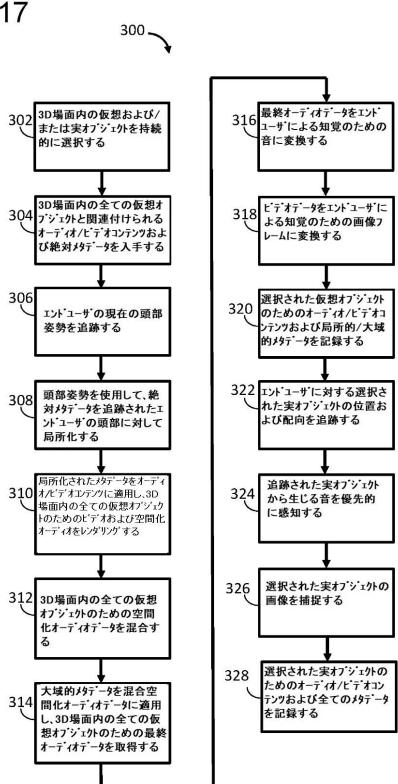
【図 16 b】

Fig. 16b



【図 17】

Fig. 17



【図 18】

Fig. 18



フロントページの続き

(51)Int.Cl.

F I

G 0 6 F	3/01	5 7 0
H 0 4 S	7/00	3 0 0

(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 フアン, ジチャン

アメリカ合衆国 フロリダ 33332, ウエストン, エヌ ハイビスカス ストリート 1
9226

審査官 堀 洋介

(56)参考文献 特開2015-019371(JP, A)

米国特許出願公開第2017/0045941(US, A1)

特開平08-298635(JP, A)

特開2013-162285(JP, A)

米国特許出願公開第2009/0174946(US, A1)

米国特許出願公開第2016/0093108(US, A1)

特表2016-522463(JP, A)

特表2015-503141(JP, A)

(58)調査した分野(Int.Cl., DB名)

G 0 6 F 3 / 0 1

G 0 6 T 1 9 / 0 0

G 0 6 F 3 / 1 6

H 0 4 S 7 / 0 0