

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7138361号
(P7138361)

(45)発行日 令和4年9月16日(2022.9.16)

(24)登録日 令和4年9月8日(2022.9.8)

(51)国際特許分類 F I
G 0 6 T 7/70 (2017.01) G 0 6 T 7/70 Z

請求項の数 25 (全26頁)

(21)出願番号	特願2020-548924(P2020-548924)	(73)特許権者	314000442
(86)(22)出願日	令和2年4月7日(2020.4.7)		高麗大学校産学協力団
(65)公表番号	特表2021-531524(P2021-531524 A)		KOREA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION
(43)公表日	令和3年11月18日(2021.11.18)		大韓民国 02841 ソウル ソンブクク
(86)国際出願番号	PCT/KR2020/004685		アナムロ 145
(87)国際公開番号	WO2020/251151		145, Anam-ro Seongb
(87)国際公開日	令和2年12月17日(2020.12.17)		uk-gu Seoul 02841,
審査請求日	令和2年9月14日(2020.9.14)		Republic of Korea
(31)優先権主張番号	10-2019-0070829	(74)代理人	110000383
(32)優先日	令和1年6月14日(2019.6.14)		特許業務法人エビス国際特許事務所
(33)優先権主張国・地域又は機関	韓国(KR)	(72)発明者	ト ナジュ
(31)優先権主張番号	10-2020-0039564		大韓民国、ソウル 02802、ソンブ
(32)優先日	令和2年4月1日(2020.4.1)		ク-グ、チョンナム-ロ 21-ギル、
	最終頁に続く		最終頁に続く

(54)【発明の名称】 3次元仮想空間モデルを利用したユーザポーズ推定方法および装置

(57)【特許請求の範囲】

【請求項1】

深さ測定装置および映像取得装置を利用して3次元空間に対する深さ情報および映像情報を含む空間情報を取得する段階、

前記空間情報に基づいて深さ-映像連携情報を構成し、前記深さ-映像連携情報に基づいて前記3次元空間に対応する3次元仮想空間モデルを構築する段階、

前記3次元空間でユーザデバイスによって取得された映像を含むユーザ情報を受信する段階、

前記3次元仮想空間モデル内で前記ユーザ情報に対応する対応情報を生成する段階、

前記対応情報と前記ユーザ情報との類似度を算出する段階、および

前記類似度に基づいて、前記3次元空間における映像を取得する前記ユーザデバイスの位置情報と方向情報を含むユーザポーズを推定する段階

を含む、3次元空間に対するユーザポーズ推定方法。

【請求項2】

前記3次元仮想空間モデルを構築する段階は、

前記3次元空間に対する映像情報で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分し、前記背景領域を利用して前記3次元仮想空間モデルを構築することを特徴とする、

請求項1に記載の3次元空間に対するユーザポーズ推定方法。

【請求項3】

前記対応情報を生成する段階は、

前記ユーザ情報に含まれた映像で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分する段階、

前記ユーザ情報に含まれた映像の背景領域を利用して前記ユーザ情報を加工する段階、および

前記3次元仮想空間モデル内で加工されたユーザ情報に対応する対応情報を生成する段階を含む、請求項1に記載の3次元空間に対するユーザポーズ推定方法。

【請求項4】

前記類似度を算出する段階は、

前記類似度を高める方向に前記対応情報を再生成する段階、および

再生成された対応情報に基づいて類似度を再算出する段階

を含む、請求項1に記載の3次元空間に対するユーザポーズ推定方法。

【請求項5】

前記類似度を算出する段階は、

前記ユーザ情報および前記対応情報を比較するための比較対象領域を抽出する段階、

前記ユーザ情報から抽出された比較対象領域と前記対応情報から抽出された比較対象領域で共通領域を決定する段階、および

前記共通領域に基づいて前記ユーザ情報および前記対応情報をそれぞれ再生成する段階を含む、請求項1に記載の3次元空間に対するユーザポーズ推定方法。

【請求項6】

前記類似度を算出する段階は、

前記ユーザデバイス周辺に対する追加のユーザ情報を取得する段階、および

前記ユーザ情報および追加のユーザ情報に基づいて類似度を算出する段階

を含む、請求項1に記載の3次元空間に対するユーザポーズ推定方法。

【請求項7】

前記ユーザポーズを推定する段階は、

前記ユーザデバイスによって前記ユーザポーズの推定に利用される付加的な情報であるユーザ付加情報が取得される場合、前記ユーザ情報または前記追加のユーザ情報とともに、前記ユーザ付加情報を利用して前記ユーザポーズを推定することを含む、

請求項6に記載の3次元空間に対するユーザポーズ推定方法。

【請求項8】

前記追加のユーザ情報を取得する段階は、

前記3次元仮想空間モデルに基づいて追加のユーザ情報取得のための案内情報を前記ユーザデバイスに送信することを含む、

請求項6に記載の3次元空間に対するユーザポーズ推定方法。

【請求項9】

前記案内情報は、前記3次元仮想空間モデル内の予め設定された特徴点に対するユーザ情報取得ポーズを含み、

前記追加のユーザ情報を取得する段階は、前記類似度を高める方向に繰り返し実行されることを特徴とする、

請求項8に記載の3次元空間に対するユーザポーズ推定方法。

【請求項10】

3次元空間に対するユーザデバイスの位置 (position) および方向 (orientation) 情報を含むユーザポーズを推定する方法であって、

前記3次元空間で取得された映像を含むユーザ情報を受信する段階、

前記3次元空間に対する深さ情報および映像情報を含む空間情報に基づいて構築された3次元仮想空間モデルを確認する段階、

前記3次元仮想空間モデル内で前記ユーザ情報に対応する対応情報を生成する段階、

前記対応情報と前記ユーザ情報との類似度を算出する段階、および

前記類似度に基づいてユーザポーズを推定する段階

10

20

30

40

50

を含む、３次元空間に対するユーザポーズ推定方法。

【請求項１１】

前記３次元仮想空間モデルは、前記３次元空間に対する映像情報で前記３次元空間の構造と関連する背景領域と前記３次元空間に置かれた物体に該当する非背景領域とを区分し、前記背景領域を利用して構築されたものであることを特徴とする、

請求項１０に記載の３次元空間に対するユーザポーズ推定方法。

【請求項１２】

前記対応情報を生成する段階は、

前記ユーザ情報に含まれた映像で前記３次元空間の構造と関連する背景領域と前記３次元空間に置かれた物体に該当する非背景領域とを区分する段階、

前記ユーザ情報に含まれた映像の背景領域を利用して前記ユーザ情報を加工する段階、および

前記３次元仮想空間モデル内で加工されたユーザ情報に対応する対応情報を生成する段階

を含む、請求項１０に記載の３次元空間に対するユーザポーズ推定方法。

【請求項１３】

前記類似度を算出する段階は、

前記類似度を高める方向に前記対応情報を再生成する段階、および

再生成された対応情報に基づいて類似度を再算出する段階

を含む、請求項１０に記載の３次元空間に対するユーザポーズ推定方法。

【請求項１４】

前記類似度を算出する段階は、

前記ユーザ情報および前記対応情報を比較するための比較対象領域を抽出する段階、

前記ユーザ情報から抽出された比較対象領域と前記対応情報から抽出された比較対象領域で共通領域を決定する段階、および

前記共通領域に基づいて前記ユーザ情報および前記対応情報をそれぞれ再生成する段階

を含む、請求項１０に記載の３次元空間に対するユーザポーズ推定方法。

【請求項１５】

前記類似度を算出する段階は、

ユーザデバイス周辺に対する追加のユーザ情報を取得する段階、および

前記ユーザ情報および追加のユーザ情報に基づいて類似度を算出する段階

を含む、請求項１０に記載の３次元空間に対するユーザポーズ推定方法。

【請求項１６】

前記ユーザポーズを推定する段階は、

前記ユーザデバイスによって前記ユーザポーズの推定に利用される付加的な情報であるユーザ付加情報が取得される場合、前記ユーザ情報または前記追加のユーザ情報とともに、前記ユーザ付加情報を利用して前記ユーザポーズを推定することを含む、

請求項１５に記載の３次元空間に対するユーザポーズ推定方法。

【請求項１７】

前記追加のユーザ情報を取得する段階は、

前記３次元仮想空間モデルに基づいて追加のユーザ情報取得のための案内情報を前記ユーザデバイスに送信することを含む、

請求項１５に記載の３次元空間に対するユーザポーズ推定方法。

【請求項１８】

前記案内情報は、前記３次元仮想空間モデル内の予め設定された特徴点に対するユーザ情報取得ポーズを含み、

前記追加のユーザ情報を取得する段階は、前記類似度を高める方向に繰り返し実行されることを特徴とする、

請求項１７に記載の３次元空間に対するユーザポーズ推定方法。

【請求項１９】

10

20

30

40

50

3次元空間に対する深さ情報および映像情報を含む空間情報を取得する空間情報取得部、前記空間情報に基づいて深さー映像連携情報を構成し、前記深さー映像連携情報に基づいて前記3次元空間に対応する3次元仮想空間モデルを生成する仮想空間モデル生成部、前記3次元空間でユーザデバイスによって取得された映像を含むユーザ情報を受信するユーザ情報受信部、および

前記3次元仮想空間モデル内で前記ユーザ情報に対応する対応情報を生成し、前記対応情報と前記ユーザ情報との類似度を算出し、前記類似度に基づいてユーザポーズを推定するように構成された少なくとも1つのプロセッサを含む制御部

を備える、3次元空間に対するユーザポーズ推定装置。

【請求項20】

前記仮想空間モデル生成部は、

前記3次元空間に対する映像情報で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分し、前記背景領域を利用して前記3次元仮想空間モデルを構築することの特徴とする、

請求項19に記載の3次元空間に対するユーザポーズ推定装置。

【請求項21】

前記制御部は、

前記ユーザ情報に含まれた映像で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分し、前記ユーザ情報に含まれた映像の背景領域を利用して前記ユーザ情報を加工し、前記3次元仮想空間モデル内で加工されたユーザ情報に対応する対応情報を生成する、

請求項19に記載の3次元空間に対するユーザポーズ推定装置。

【請求項22】

3次元空間に対するユーザデバイスの位置 (p o s i t i o n) および方向 (o r i e n t a t i o n) 情報を含むユーザポーズを推定する装置であって、

前記3次元空間に対する深さ情報および映像情報を含む空間情報に基づいて構築された3次元仮想空間モデルを提供する仮想空間モデル提供部、

前記3次元空間で前記ユーザデバイスによって取得された映像を含むユーザ情報を受信するユーザ情報受信部、および

前記3次元仮想空間モデル内で前記ユーザ情報に対応する対応情報を生成し、前記対応情報と前記ユーザ情報との類似度を算出し、前記類似度に基づいて前記ユーザポーズを推定するように構成された少なくとも1つのプロセッサを含む制御部

を備える、3次元空間に対するユーザポーズ推定装置。

【請求項23】

前記3次元仮想空間モデルは、前記3次元空間に対する映像情報で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分し、前記背景領域を利用して構築されたものであることを特徴とする、

請求項22に記載の3次元空間に対するユーザポーズ推定装置。

【請求項24】

3次元空間に対する映像を含むユーザ情報を生成するユーザ情報生成部、

前記ユーザ情報をユーザポーズ推定サーバに送信し、3次元仮想空間モデルで推定されたユーザポーズに関する情報を前記サーバから受信する通信部、および

前記ユーザ情報生成部および通信部の動作を制御し、前記ユーザポーズに関する情報を現在実行中のアプリケーションまたは駆動システムに伝達するように構成された少なくとも1つのプロセッサを含む制御部

を備える、3次元空間に対するユーザポーズ推定装置。

【請求項25】

前記3次元仮想空間モデルは、前記3次元空間に対する深さ情報および映像情報を含む空間情報に基づいて生成され、前記3次元空間に対する映像情報で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分し、

10

20

30

40

50

前記背景領域を利用して構築されたものであることを特徴とする、
請求項 24 に記載の 3 次元空間に対するユーザポーズ推定装置。

【発明の詳細な説明】

【技術分野】

【0001】

技術分野は、空間地図の生成および活用に関し、より詳細には、3次元仮想空間モデルを利用して現実空間内のユーザポーズを推定する方法および装置に関する。

【背景技術】

【0002】

空間地図を活用してユーザポーズ (pose) を推定する方法は、地形値 (geometry) を利用する方式、映像値を利用する方式、地形値と映像値との混合を利用する方式がある。

このとき、LiDAR (ライダー) あるいはこれと類似の動作原理をもつ深さ測定装置を活用して点群情報を取得したり、カメラあるいはこれと類似の動作原理をもつ映像測定装置を活用して映像情報を取得したり、Kinect (キネクト) あるいはこれと類似の動作原理をもつ深さ-映像測定装置を活用して色一点群情報を取得したり、またはこれらの組み合わせを利用して、現実空間を表現する空間地図を構成することができる。

【0003】

現実空間に対する映像情報、深さ情報、および深さ-映像連携情報は、「空間情報」と称される。

ユーザポーズは、ユーザデバイスが現実空間で取得したユーザ情報を空間地図と比べることによって推定される。

【0004】

ここで、「ユーザ情報」とは、ユーザデバイスが現実空間で取得した映像を含む情報である。また、「ポーズ」とは、位置 (position) と方向 (orientation) の両方を含む概念である。したがって、「ユーザポーズ」とは、現実空間に対して映像情報を取得した位置情報と、映像情報を取得した方向情報を含む情報であると言える。

ただ、従来技術における空間地図を利用したユーザポーズの推定は、次のような問題を抱えている。

【0005】

第一に、空間地図は、空間情報を取得するポーズに敏感なことがある。したがって、空間地図が空間情報を取得するポーズに敏感な場合、ユーザポーズ推定の精密度は低下するようになる。例えば、理論的に可能なすべてのポーズで空間情報を取得して空間地図を構成する場合は、精密度の高いユーザポーズが推定可能となる。

【0006】

ところが、現実空間に対してすべてのポーズで空間情報を取得することは事実上不可能である。空間地図を構成するために現実空間から均等な分布で多数のポーズの空間情報を取得した場合、空間情報を取得するポーズの分布によってユーザポーズ推定の敏感度は低下する。しかし、このような場合には、空間情報の取得時間、空間情報の容量、および処理時間などのようなシステム負荷問題が発生する恐れがある。

【0007】

一方、システムの負荷問題を考慮しながら少数のポーズで空間情報を取得する場合には、空間地図が現実空間を十分に表現することができない。さらに、空間地図を取得する経路が変われば空間地図の信頼性が低下し、現実空間を強靱に表現することができなくなる。現実空間が強靱に表現されていない空間地図は、ユーザポーズ推定の精密度の低下に繋がる。

第二に、不連続的な空間地図は、ユーザポーズ推定の精密度を低下させることがある。図 1 は、点群情報で構成された不連続的な空間地図の例を示した図である。

【0008】

図 1 に示すように、点群情報を活用して空間地図を構成する場合、空間情報の取得範囲

10

20

30

40

50

や経路に応じて点群情報を稠密に取得できないことがある。点群情報を稠密に取得できない場合には連続性のない空間地図が生成されるようになり、これはユーザポーズ推定の精密度を低下させることに繋がる。

第三に、空間地図を構成するための空間情報が取得された時点とユーザ情報を取得する時点との差により、ユーザポーズ推定の精密度が下がることもある。

【0009】

図2および図3は、時間の流れによる空間の変化を示した例示図である。

図2は、光または照明が時間の変化によって変化する例を示している。

より具体的に、図2の(a)、(b)、および(c)は、同じ空間において、照明や外部から流入する光の量が時間の流れによって変化する例を示している。

また、図3の(a)および(b)は、同じ空間において、時間の流れによって物体が変化した例を示している。

【0010】

図3の(a)ではテーブル210の上に何も置かれていないが、図3の(b)ではテーブル220上に物体が置かれていることを示している。

例えば、図2に示した空間に対し、空間地図を構成するための空間情報は(a)から取得され、ユーザ情報は(c)から取得されることがある。また、図3に示した空間に対し、空間地図を構成するための空間情報は(a)から取得され、ユーザ情報は(b)から取得されることがある。

【0011】

このように、同じ空間であっても、空間情報が取得された時点とユーザ情報を取得する時点(time)との差により映像情報が一致しないことがある。したがって、同じ空間であっても、空間情報が取得された時点とユーザ情報を取得する時点との差により、ユーザポーズ推定の精密度は低下するようになる。

現実空間は、時間の流れにより、光または照明の変化、人物などの動的な動きの変化、物体またはインテリアなどの変化が発生する。このような変化が更新されていない空間地図を使用する場合にはユーザ情報との類似性が低下するようになり、これはユーザポーズ推定の精密度を低下させることに繋がる。

したがって、空間地図に基づいてユーザポーズを推定するときに発生する従来の問題を解決する方法が必要となっている。

【発明の概要】

【発明が解決しようとする課題】

【0012】

本発明は、上述した問題点を解決するために、現実空間で取得された空間情報に基づいて構成された3次元仮想空間モデルとユーザが取得したユーザ情報とを活用してユーザポーズを推定する方法及び装置を提供する。

【課題を解決するための手段】

【0013】

一実施形態に係る3次元空間に対するユーザポーズ推定方法は、深さ測定装置および映像取得装置を利用して3次元空間に対する深さ情報および映像情報を含む空間情報を取得する段階、前記空間情報に基づいて深さ-映像連携情報を構成し、前記深さ-映像連携情報に基づいて前記3次元空間に対応する3次元仮想空間モデルを構築する段階、前記3次元空間でユーザデバイスによって取得された映像を含むユーザ情報を受信する段階、前記3次元仮想空間モデル内で前記ユーザ情報に対応する対応情報を生成する段階、前記対応情報と前記ユーザ情報との類似度を算出する段階、および前記類似度に基づいてユーザポーズを推定する段階を含む。

【0014】

前記3次元仮想空間モデルを構築する段階は、前記3次元空間に対する映像情報で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分し、前記背景領域を利用して前記3次元仮想空間モデルを構築することを含

10

20

30

40

50

んでよい。

【0015】

前記対応情報を生成する段階は、前記ユーザ情報に含まれた映像で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分する段階、前記ユーザ情報に含まれた映像の背景領域を利用して前記ユーザ情報を加工する段階、および前記3次元仮想空間モデル内で加工されたユーザ情報に対応する対応情報を生成する段階を含んでよい。

【0016】

前記類似度を算出する段階は、前記類似度を高める方向に前記対応情報を再生成する段階、および再生成された対応情報に基づいて類似度を再算出する段階を含んでよい。

10

【0017】

前記類似度を算出する段階は、前記ユーザ情報と前記対応情報とを比較するための比較対象領域を抽出する段階、前記ユーザ情報から抽出された比較対象領域と前記対応情報から抽出された比較対象領域で共通領域を決定する段階、前記共通領域に基づいて前記ユーザ情報および前記対応情報をそれぞれ再生成する段階を含んでよい。

【0018】

前記類似度を算出する段階は、前記ユーザデバイス周辺に対する追加のユーザ情報を取得する段階、および前記ユーザ情報および追加のユーザ情報に基づいて類似度を算出する段階を含んでよい。

【0019】

前記ユーザポーズを推定する段階は、前記ユーザデバイスによって前記ユーザポーズの推定に利用される付加的な情報であるユーザ付加情報が取得される場合、前記ユーザ情報または前記追加のユーザ情報とともに、前記ユーザ付加情報を利用して前記ユーザポーズを推定することを含んでよい。

20

【0020】

前記追加のユーザ情報を取得する段階は、前記3次元仮想空間モデルに基づいて追加のユーザ情報を取得するための案内情報を前記ユーザデバイスに送信することを含んでよい。

【0021】

前記案内情報は、前記3次元仮想空間モデル内の予め設定された特徴点に対するユーザ情報取得ポーズを含み、前記追加のユーザ情報を取得する段階は、前記類似度を高める方向に繰り返し実行されてよい。

30

他の実施形態に係る3次元空間に対するユーザの位置 (position) および方向 (orientation) 情報を含むユーザポーズを推定する方法は、前記3次元空間で取得された映像を含むユーザ情報を受信する段階、前記3次元空間に対する深さ情報および映像情報を含む空間情報に基づいて構築された3次元仮想空間モデルを確認する段階、前記3次元仮想空間モデル内で前記ユーザ情報に対応する対応情報を生成する段階、前記対応情報と前記ユーザ情報との類似度を算出する段階、および前記類似度に基づいてユーザポーズを推定する段階を含む。

【0022】

一実施形態に係る3次元空間に対するユーザポーズ推定装置は、3次元空間に対する深さ情報および映像情報を含む空間情報を取得する空間情報取得部、前記空間情報に基づいて深さ-映像連携情報を構成し、前記深さ-映像連携情報に基づいて前記3次元空間に対応する3次元仮想空間モデルを生成する仮想空間モデル生成部、前記3次元空間でユーザデバイスによって取得された映像を含むユーザ情報を受信するユーザ情報受信部、および前記3次元仮想空間モデル内で前記ユーザ情報に対応する対応情報を生成し、前記対応情報と前記ユーザ情報との類似度を算出し、前記類似度に基づいて前記ユーザポーズを推定するように構成された少なくとも1つのプロセッサを含む制御部を備える。

40

【0023】

前記空間モデル生成部は、前記3次元空間に対する映像情報で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分し、前

50

記背景領域を利用して前記３次元仮想空間モデルを構築してよい。

【００２４】

前記制御部は、前記ユーザ情報に含まれた映像で前記３次元空間の構造と関連する背景領域と前記３次元空間に置かれた物体に該当する非背景領域とを区分し、前記ユーザ情報に含まれた映像の背景領域を利用して前記ユーザ情報を加工し、前記３次元仮想空間モデル内で加工されたユーザ情報に対応する対応情報を生成してよい。

【００２５】

他の実施形態に係る３次元空間に対するユーザの位置（*position*）および方向（*orientation*）情報を含むユーザポーズを推定する装置は、前記３次元空間に対する深さ情報および映像情報を含む空間情報に基づいて構築された３次元仮想空間モデルを提供する仮想空間モデル提供部、前記３次元空間でユーザデバイスによって取得された映像を含むユーザ情報を受信するユーザ情報受信部、前記３次元仮想空間モデル内で前記ユーザ情報に対応する対応情報を生成し、前記対応情報と前記ユーザ情報との類似度を算出し、前記類似度に基づいて前記ユーザポーズを推定するように構成された少なくとも１つのプロセッサを含む制御部を備える。

【００２６】

また他の実施形態に係る３次元空間に対するユーザポーズ推定装置は、３次元空間に対する映像を含むユーザ情報を生成するユーザ情報生成部、前記ユーザ情報をユーザポーズ推定サーバに送信し、３次元仮想空間モデルで推定されたユーザポーズに関する情報を前記サーバから受信する通信部、前記ユーザ情報生成部および通信部の動作を制御し、前記ユーザポーズに関する情報を現在実行中のアプリケーションまたは駆動システムに伝達するように構成された少なくとも１つのプロセッサを含む制御部を備える。

【発明の効果】

【００２７】

本発明の実施形態は、空間地図として３次元仮想空間モデルを使用することにより、空間情報の取得経路に強靱な３次元仮想空間モデルを構築することができ、空間情報取得ポーズによるユーザポーズ推定精密度の感度を減らすことができる。

【００２８】

また、本発明の実施形態に係る３次元仮想空間モデルは、現実空間と類似に構成可能である上に、空間情報の取得時間、空間情報の容量、データの処理時間などを減らすことができる。

また、時間の流れるによる現実空間の変化にも強靱なユーザポーズ推定方法を提供することができる。

また、本発明の実施形態は、複合現実（*Mixed Reality*）でユーザポーズを推定するときに活用することができる。

さらに、精密なユーザポーズ推定によって現実空間と仮想空間との違和感を軽減することができ、複合現実におけるユーザの没入度を高めることができる。したがって、本発明の実施形態は、複合現実の関連技術の商用化および発展に寄与することができる。

【図面の簡単な説明】

【００２９】

【図１】点群情報を活用して構成された不連続的な空間地図の例を示した図である。

【図２】時間の流れるによる空間の変化を示した例示図である。

【図３】時間に流れるによる空間の変化を示した他の例示図である。

【図４】本発明の一実施形態における、３次元仮想空間モデルの例を示した図である。

【図５】一実施形態における、３次元仮想空間モデルを生成する例を説明するための図である。

【図６】一実施形態における、３次元仮想空間モデルを利用したユーザポーズ推定システムを説明するための図である。

【図７】一実施形態における、３次元空間に対するユーザポーズ推定装置の構成を説明するための図である。

【図 8】一実施形態における、ユーザデバイスの構成を説明するための図である。

【図 9】一実施形態における、ポーズの概念を説明するための例示図である。

【図 10】一実施形態における、3次元空間に対するユーザポーズ推定方法を説明するためのフローチャートである。

【図 11】他の実施形態における、3次元空間に対するユーザポーズ推定方法を説明するためのフローチャートである。

【図 12】一実施形態における、追加でユーザポーズを取得する方法の例を説明するための図である。

【発明を実施するための形態】

【0030】

以下、添付の図面および添付の図面に記載された内容を参照しながら本発明の実施形態について詳細に説明するが、本発明が実施形態によって制限あるいは限定されることはない。

【0031】

本明細書で使用される用語は、実施形態を説明するためのものに過ぎず、本発明を制限するためのものではない。本明細書に記載する単数形は、文面で特に明記されない限り複数形も含む。明細書で使用される「含む (comprises)」および／または「含む (comprising)」は、記載された構成要素、段階、動作、および／または素子において、1つ以上の他の構成要素、段階、動作、および／または素子の存在または追加を排除しない。

【0032】

本明細書で使用される「実施例」、「例」、「側面」、「例示」などは、記載された任意の態様 (aspect) または設計が、他の態様または設計よりも好ましいまたは有利であると解釈されてはならない。

また、「または」という用語は、排他的論理合「exclusive or」というよりは包含的論理合「inclusive or」を意味する。つまり、別途記載されない限りまたは文脈から明らかでない限り、「xがaまたはbを利用する」という表現は、自然な包含的順列 (natural inclusive permutations) のうちのいずれか1つを意味する。

【0033】

また、本明細書および特許請求の範囲で使用される第1や第2などの用語は、多様な構成要素を説明するために使用されるものであるが、前記構成要素が前記用語によって限定されてはならない。前記用語は、1つの構成要素を他の構成要素と区別する目的としてのみ使用される。

【0034】

別途の定義なければ、本明細書で使用されるすべての用語（技術および科学的用語を含む）は、本発明が属する技術分野において通常の知識を有する者が共通して理解することのできる意味として使用されるであろう。また、一般的に使用される事前に定義されている用語は、明らかに特に定義されていない限り、理想的または過度に解釈されてはならない。

【0035】

なお、本発明を説明するにあたり、関連する公知機能または構成についての具体的な説明が本発明の要旨を不必要に不明瞭にし得ると判断される場合には、その詳細な説明は省略する。さらに、本明細書で使用される用語 (terminology) は、本発明の実施形態を適切に表現するための用語であり、これは、ユーザ、運用者の意図、または本発明が属する分野の慣例などによって異なることもある。したがって、本明細書で使用される用語は、本明細書の全般にわたる内容に基づいて定義されなければならない。

【0036】

図 4 は、本発明の一実施形態における、3次元仮想空間モデルの例を示した図である。

図 5 は、一実施形態における、3次元仮想空間モデルを生成する例を説明するための図

10

20

30

40

50

である。

図 4 を参照すると、通常の 3 次元仮想空間モデルは、obj、x3d などのような深さ一映像連携情報を利用して構成された現実空間に対するモデルを指す。例えば、3 次元仮想空間モデルは、韓国登録特許第 10-1835434 号公報（発明の名称：投映イメージ生成方法およびその装置、イメージピクセルと深さ値とのマッピング方法）により、obj、x3d などが生成されたモデルまたは「TeVR モデル」を含んでもよい。

【0037】

このとき、一実施形態に係る 3 次元仮想空間モデルは、図 5 に示すように、3 次元仮想空間モデルを構成するために背景領域と非背景領域とを区分し、背景領域だけで構成されてよい。

10

図 5 において、(a) は空間情報に含まれた映像情報であり、(b) は非背景領域を除いた映像を、(c) は背景領域を確張して映像情報を生成した例を示している。

例えば、背景領域とは、3 次元空間を形成する建築物自体の構造を意味するか、またはドアや窓のように建築物に付随した構造物であってよい。したがって、映像情報において、背景領域は、3 次元空間の構造と関連する領域であると定義されてよい。

【0038】

図 5 の (a) において、室内空間に位置する多様な物体（机、本棚、椅子など）に対応する領域は、非背景領域に該当する。図 5 の (b) は、室内空間に位置する多様な物体（白色で表示された領域）が取り除かれた非背景領域を示している。

3 次元仮想空間モデルは、室内空間と室外空間の両方を含む概念であって、独立的な室内空間、独立的な室外空間、または室内と室外とが連結した空間であってもよい。3 次元仮想空間モデルに obj、x3d などのようなモデル（人物、事物など）を追加してもよく、3 次元仮想空間モデルは、モデルが追加された 3 次元仮想空間モデルを含む概念であってもよい。一方、空間地図は、3 次元の代わりとして、次元を低めて 2 次元仮想空間モデルの使用も可能である。

20

【0039】

3 次元空間モデルは、事前に構築された obj、x3d のようなモデルを利用してもよいし、空間情報を新たに取得して構築してもよいし、事前に構築されたモデルを更新して利用してもよい。3 次元空間モデルは、現実空間と類似するものと判断されてよい。

【0040】

図 6 は、一実施形態における、3 次元仮想空間モデルを利用したユーザポーズ推定システムを説明するための図である。

30

図 6 を参照すると、3 次元仮想空間モデルを利用したユーザポーズ推定システムは、ユーザデバイス 610 およびユーザポーズ推定装置 620 を備える。一実施形態に係るユーザポーズ推定装置 620 は、サーバ（図示せず）に備えられてもよいし、ユーザデバイス 610 に備えられてもよい。

【0041】

ユーザデバイス 610 は、現実空間 601 でユーザ情報 611 を取得し、ユーザデバイス 610 にユーザ情報 611 を送信してよい。

ユーザポーズ推定装置 620 は、装置内部または外部のストレージシステム 602 に記録された 3 次元仮想空間モデル 630 およびユーザ情報 611 を利用してユーザポーズを推定してよい。

40

【0042】

ユーザポーズ推定装置 620 は、3 次元仮想空間モデル 630 でユーザポーズに該当する確率が高い対応情報 621 をユーザ情報 611 と比較することで、正確なユーザポーズを推定することが可能となる。

【0043】

図 7 は、一実施形態における、3 次元空間に対するユーザポーズ推定装置の構成を説明するための図である。

図 7 を参照すると、一実施形態に係る 3 次元空間に対するユーザポーズ推定装置 620

50

は、仮想空間モデル提供部 730、制御部 740、およびユーザ情報受信部 750を備える。また、ユーザポーズ推定装置 620は、空間情報取得部 710および仮想空間モデル生成部 720をさらに備えてよい。さらに、ユーザポーズ推定装置 620は、ユーザ情報要請部 760をさらに備えてよい。

空間情報取得部 710は、3次元空間に関する深さ情報および映像情報を含む空間情報を取得する。例えば、空間情報は、深さ測定装置および映像測定装置を利用して取得されてよい。

【0044】

3次元仮想空間モデルを構成するための空間情報を取得する深さ測定装置または映像測定装置などで構成された測定装置の視野角（ $Fov: Field\ of\ View$ ）が現実空間を確保することのできる経路で空間情報を取得すれば、3次元仮想空間モデルは現実空間と類似するように構成されるようになる上に、空間情報の取得時間、空間情報の容量、データの処理時間などを減らすことができ効率的である。

【0045】

映像情報は、3次元空間に対する2次元イメージとして、2自由度の基底ベクトル（ $basis\ vector$ ）で表現可能な形態であってよく、カメラのように3次元を2次元で表現する形態、またはカメラに赤外線フィルタを装着して3次元列情報を2次元で表現した形態であってよい。

【0046】

深さ情報は、3自由度の基底ベクトルで表現可能な点形態であり、深さ測定装置を利用して取得されてもよく、互いに異なる場所で撮影された2つ以上のイメージを活用して推定されてもよい。前者の例としては、 $LIDAR$ （ライダー）、 $SONAR$ （ソナー）、赤外線（ $Infrared$ ）、 TOF （ $Time\ Of\ Flight$ ）距離探知機を利用して取得された深さ情報があり、後者の例としては、ステレオカメラ、マルチカメラ、全方向ステレオカメラなどを利用して取得された深さ情報がある。一方、 $Kinect$ 、 $JUMP$ 、 $PrimeSense$ 、 $Project\ Beyond$ などのデバイスを利用すれば、深さ情報および映像情報を同時に取得することも可能である。

【0047】

例えば、本発明の一実施形態では、深さ測定装置を利用して取得した深さ情報だけでなく、内挿法（ $interpolation$ ）によって深さ情報を新たに推定して使用してもよい。より具体的には、取得した複数の深さ情報のうちから3つ以上の深さ情報を選別して多角形（三角形を含む）メッシュ（ $Mesh$ ）を構成した後、多角形メッシュの内部に内挿法によって新たな深さ情報を推定して追加するのである。

【0048】

一方、本発明の一実施形態に係る取得深さ情報および映像情報は、統合センサシステムを利用して同時に取得されてもよい。複数の測定装置を使用する場合、センサ間の座標関係を求めるキャリブレーション過程が必要となることがある。

【0049】

空間情報を取得する過程で慣性測定装置（ IMU ）などを追加で使用してもよく、タイヤ型移動ロボットにセンサを装着して測定する場合には距離情報（ $odometry$ ）を活用してもよい。現実空間が測定装置の視野角よりも広い場合、センサの回転、移動、またはこれを組み合わせて空間情報を取得してよい。このとき、個別空間情報が取得された3次元ポーズ（ $pose$ ）がそれぞれ異なる場合があり、個別空間情報が取得されたポーズを予測するために、スラム（ $SLAM: Simultaneous\ Localization\ And\ Mapping$ ）、視覚慣性オドメトリ（ $VIO: Visual\ Inertial\ Odometry$ ）、視覚オドメトリ（ $VO: Visual\ Odometry$ ）などの手法が活用されてよい。

【0050】

一方、空間情報の構成は、測定装置の種類に応じて異なってよい。一例として、測定装置が単一カメラだけで構成された場合、事前測定情報はカメラ映像情報で構成され、該当

10

20

30

40

50

の映像情報を利用して単一カメラである場合にはピクセル間の相対的距離を、複数のカメラの場合にはピクセル間の絶対距離の予測が可能である。特に、特徴点を抽出せずに単一カメラの場合は、累積した映像情報を利用してピクセルの深さを予測することができ、複数のカメラの場合は、複数のカメラ映像やこの累積した映像情報を利用してピクセルの深さを予測することもできる。

【0051】

さらに、追加の深さ情報、慣性情報などの情報がともに活用された場合、各測定装置の固有の特性に合うように空間情報処理が可能となる。一例として、慣性測定装置によって慣性情報が取得可能な場合は、これを活用してスラムの性能を向上させたり、映像情報処理時の映像取得ポーズに対する推定（prediction）情報として使用して映像取得ポーズに対する補正（correction）がより容易に行われるようにできる。また、慣性情報の加速度値あるいは角速度値を活用して実際の移動距離を予想することもでき、これを単一カメラあるいは複数のカメラから抽出された深さ情報の大きさ（scale）を補正するのに活用することもできる。

【0052】

仮想空間モデル生成部720は、空間情報に基づいて深さ－映像連携情報を構成し、前記深さ－映像連携情報に基づいて前記3次元空間に対応する3次元仮想空間モデルを生成する。

【0053】

例えば、室内空間の3次元仮想空間モデルを生成する場合、空間情報取得部710は、室内空間に対する空間情報である室内空間イメージを取得してよい。このとき、室内空間イメージは、室内空間内部の多様な位置で撮影されたイメージであってよい。

【0054】

このとき、仮想空間モデル生成部720は、室内空間の構造に対応する領域である背景領域と、室内空間に位置する物体または移動中の人物に対応する領域である非背景領域とを区分してよい。

仮想空間モデル生成部720は、室内空間イメージを構成するイメージのピクセル値に基づいて背景領域と非背景領域とを区分してよい。

【0055】

背景領域とは、他の要素によって一部が隠れているため完全なデータではないが、隠れていない部分と類似性があるものと推論され、隠れていない部分をホールフィリング（hole filling）やインペインティング（inpainting）手法によって再構成することのできる部分に該当してよい。さらに、背景領域とは、建物内部の大きな看板や案内デスクなどのような他の物体を隠すことがあるが、該当の物体のエッジ部分で映像と地形との整合度がすべてのデータ内で一致するか、別の整合過程によって一致させることのできる部分であってよい。

【0056】

仮想空間モデル生成部720は、室内空間イメージで背景領域を非背景領域に確張して少なくとも1つの拡張室内空間イメージを生成してよい。例えば、図5の（b）において、非背景領域が取り除かれた白色で表現された部分を背景領域として確張してよい。

【0057】

仮想空間モデル生成部720は、背景領域に含まれたエッジが非背景領域との境界線で断絶した場合にエッジの延長線が背景領域と非背景領域との境界線を越えて非背景領域に繋がるという推論に基づき、拡張されたイメージを生成してよい。

【0058】

このとき、特定の室内空間イメージの他に1つ以上の室内空間イメージを背景の補完イメージとして指定し、特定の室内空間イメージの非背景領域に該当する領域を、背景の補完イメージの情報を利用して減らしてよい。

【0059】

仮想空間モデル生成部720は、少なくとも1つの拡張された室内空間イメージと室内

10

20

30

40

50

空間に関する深さ値の情報を含む地形情報に基づいて深さ－映像連携情報を生成してよい。深さ－映像連携情報は、少なくとも1つの拡張された室内空間イメージのピクセルに対応する室内空間の深さ値がマッチングされた情報であってよい。

【0060】

仮想空間モデル生成部720は、少なくとも1つの拡張された室内空間イメージと地形情報の他にも、少なくとも1つの拡張された室内空間イメージと地形情報それぞれの取得位置、および取得角度に関する情報を含むイメージ取得ポーズと深さ取得ポーズをさらに利用して深さ－イメージ連携情報を生成してよい。

【0061】

仮想空間モデル生成部720は、少なくとも1つの拡張された室内空間イメージ、地形情報、および深さ－映像連携情報を利用して現実の3次元空間に対する3次元仮想空間モデルを生成する。

仮想空間モデル提供部730は、ユーザポーズの推定が必要な場合、3次元空間に関する深さ情報および映像情報を含む空間情報に基づいて構築された3次元仮想空間モデルを提供する。

【0062】

このとき、ユーザポーズ推定は、ユーザデバイス610またはユーザポーズ推定装置620にインストールされたアプリケーションの実行後に行われてよい。仮想空間モデル提供部730は、ユーザデバイス610またはユーザポーズ推定装置620で実行中のアプリケーションまたは該当の装置の駆動システムに3次元仮想空間モデルを提供してよい。

【0063】

制御部740は、少なくとも1つのプロセッサを含んでよい。このとき、制御部740は、命令語(instructions)またはプログラムが記録された少なくとも1つのコンピュータ読み取り可能なストレージ(one or more computer-readable storage media)と連結してよい。

【0064】

したがって、制御部740は、3次元仮想空間モデル内でユーザ情報に対応する対応情報を生成し、対応情報とユーザ情報との類似度を算出し、類似度に基づいてユーザポーズを推定するように構成された少なくとも1つのプロセッサを含む。

【0065】

一実施形態に係るユーザポーズ推定は、ディープラーニング(Deep Learning)またはニューラルネットワークを活用して3次元仮想空間モデルを学習することによって実行されてよい。

【0066】

学習問題の形態に応じて強化学習(reinforcement learning)、教師あり学習(supervised learning)、教師なし学習(unsupervised learning)に分けられてよい。学習段階では膨大な量の訓練データ(training test)が必要となることがあり、訓練データは、映像情報が含まれたデータとそのデータが取得されたポーズが含まれたデータなどで構成されてよく、学習データの量を増加させるために前記2つの種類のデータにノイズ(noise)を追加して変形させて使用してもよい。畳み込みニューラルネットワーク(CNN)または多様なニューラルネットワークの全体または一部を使用してもよい。ディープラーニングの性能または速度向上のために、1つ以上のGPUを使用してもよく、並列演算を実行してもよい。ディープラーニングの結果はスカラー、ベクトル、確率などで表現されてよく、この結果を利用することにより、ユーザ情報が取得されたポーズとして予想されるユーザポーズを推定してよい。入力によってユーザ情報の映像情報を使用してよく、ユーザ付加情報をともに使用してよい。ユーザ付加情報をともに使用する場合、ニューラルネットワークにレイヤ(layer)を追加したり、関数を変化させたり、パラメータ数を調節したり、その値を変更したりしてよい。ニューラルネットワークを構成するために、パイソン、C言語、MATLAB(マトラボ)などや、これらの組み合わせからなるコン

10

20

30

40

50

コンピュータ言語を使用してもよい。

【0067】

ユーザ情報が順に取得される場合、3次元仮想空間モデルを基盤として粒子フィルタ (Particle Filter)、EKF、EIF、UKFなどの手法を活用してユーザポーズを推定してよい。ユーザ付加情報として慣性情報または距離情報が取得される場合、推定されたユーザポーズを補正してよい。順に取得されたユーザ情報によって粒子フィルタの値が特定のポーズとして収斂されてよく、このとき収斂された地点をユーザポーズとして推定してよい。ユーザポーズの推定時に加重値 (Weight) を付加してよく、多数の収斂ポーズのうちからユーザポーズを決定してもよい。

【0068】

ディープラーニングによって推定したポーズと粒子フィルタなどで推定されたポーズとを融合してユーザポーズを推定してよい。例えば、ディープラーニングで推定したポーズの周りで粒子フィルタを実行してユーザポーズを推定してよく、反対の方法として、粒子フィルタによって収斂されたポーズの周りでディープラーニングを利用してユーザポーズを推定してよい。ユーザポーズの推定時に加重値 (Weight) を付加してよく、多数の収斂ポーズのうちからユーザポーズを決定してもよい。

【0069】

類似度とは、3次元仮想空間モデルで生成した対応情報とユーザ情報との類似の程度を意味し、類似度が高いほど前記対応情報と前記ユーザ情報が類似すると見なしてよく、類似度が高い対応情報が生成された3次元仮想空間モデルのポーズをユーザ情報が取得されたユーザポーズとして推定してよい。類似度は、スカラー (scalar)、ベクトル (vector)、共分散行列 (covariance matrix) などで表現されてよく、ユークリッド距離 (euclidean distance)、マンハッタン距離 (manhattan distance)、マハラノビス距離 (mahalanobis distance)、構造的類似性 (SSIM: structural similarity)、正規化情報距離 (NID: normalized information distance)、最小平均2乗誤差推定 (MMSE: Minimum Mean Square Error)、エントロピ (Entropy) などによって算出されてもよい。

【0070】

類似度算出およびユーザポーズ推定について、図10を参照しながらより詳しく説明する。

【0071】

このとき、3次元仮想空間モデルは、3次元空間に対する映像情報で3次元空間の構造と関連する背景領域と3次元空間に置かれた物体に該当する非背景領域とを区分し、背景領域で構築されたものであってよい。

【0072】

ユーザ情報受信部750は、3次元空間でユーザデバイスによって取得された映像を含むユーザ情報を受信する。

ユーザ情報は、映像情報が含まれた情報であって、1つ以上の映像測定装置を含み、深さ測定装置または付加装置などを利用して取得してよい。測定装置の視野角が狭くて十分なユーザ情報が取得されない場合、測定装置の回転、移動、またはこれを組み合わせてユーザ情報を取得してよい。ユーザ情報は、単一または複数の映像センサ (カメラ) によって取得してよく、ピンホール (pin-hole) モデルや魚眼レンズ、またはパノラマ形式で取得されてもよい。単一の映像情報、複数の映像情報、または順列の映像情報が取得されてよい。取得したユーザ情報を利用して映像情報、深さ情報、または深さ-映像連携情報などを構成してよい。

【0073】

例えば、単一映像測定装置を使用すれば映像情報を取得することができ、順に取得された映像情報を使用すれば深さ情報を計算することができ、これにより深さ-映像連携情報を構成することができる。

10

20

30

40

50

例えば、複数の映像測定装置を使用すれば、各映像測定装置で取得された映像情報と映像測定装置との関係を活用して深さ情報を計算することができ、これにより深さ－映像連携情報を構成することができる。映像測定装置との関係は、映像測定装置間のキャリブレーション（Calibration）情報、または各映像測定装置で取得された映像情報間の変換情報（Homography matrix）であってよい。

【0074】

例えば、少なくとも1つ以上の映像測定装置と少なくとも1つ以上の深さ測定装置を使用する場合、2つの装置間のキャリブレーション情報を利用して深さ－映像連携情報を構成してよい。ディープラーニング（Deep Learning）を活用して映像情報から深さ情報を抽出してもよい。ニューラルネットワークが構成されてよく、畳み込みニューラルネットワークが使用されてよい。学習および試験のために多数のデータが必要となることがあり、ニューラルネットワークは、線形関数、非線形関数、多数のレイヤなどで構成されてよく、ディープラーニングの結果は、確率、スカラー、ベクトルなどで表現されてよい。繰り返し学習が実行されてよく、パラメータチューニングが必要となることもある。ディープラーニングによって抽出された深さ情報を活用して深さ－映像連携情報を構成してよい。映像情報を加工した、加工された映像情報を使用してもよく、例えば、イメージの明暗や彩度などを変化させたり、パノラマイメージを校正イメージ（Rectified Image）に変換させたりする作業が実行されてよい。

【0075】

ユーザ情報要請部760は、追加のユーザ情報の取得が必要な場合、ユーザデバイス610に案内情報を送信してよい。案内情報については、図12を参照しながら詳しく説明する。

【0076】

図8は、一実施形態における、ユーザデバイスの構成を説明するための図である。

図8を参照すると、ユーザデバイス610は、ユーザ情報生成部810、通信部820、および制御部830を備える。ユーザデバイス610は、ユーザとインタフェースするためのディスプレイ、入力手段、および出力手段を含むユーザインタフェース部840をさらに備えてよい。

【0077】

ユーザ情報生成部810は、3次元空間に対する映像を含むユーザ情報を生成する。したがって、ユーザ情報生成部810は、映像測定装置および深さ測定装置のうちの少なくとも1つを含んでよい。

【0078】

通信部820は、ユーザ情報をユーザポーズ推定サーバに送信し、3次元仮想空間モデルで推定されたユーザポーズに関する情報を前記サーバから受信する。

このとき、ユーザポーズ推定サーバは、図7に示したユーザポーズ推定装置620であってよく、ユーザポーズ推定サービスを提供する別のサーバであってもよい。

【0079】

制御部830は、ユーザ情報生成部810および通信部820の動作を制御し、ユーザポーズに関する情報を現在実行中のアプリケーションまたは駆動システムに伝達するように構成された少なくとも1つのプロセッサを含む。

図9は、一実施形態における、ポーズの概念を説明するための例示図である。

3次元仮想空間モデルを構築するために使用される空間情報は、現実空間の一部ポーズで取得された不連続情報として考慮されてよい。ここで、ポーズ（pose）とは、位置（position）と方向（orientation）の両方を含む概念である。一例として、2次元において、ポーズは、測定装置の位置である x 、 y と、測定装置の角度 a とで表現されてよい。

【0080】

図9に示した例は、横と縦が1mである正方形の平面で、測定装置が x 軸、 y 軸に対して0～1mの範囲で10cm間隔に移動し、回転角度は0～360度の範囲で10度ずつ

10

20

30

40

50

回転する例を示している。

このとき、可能な全体ポーズの場合の数は、 $11 \times 11 \times 37$ 、すなわち4,477種類となる。同じように、3次元において、ポーズは、センサの位置であるx、y、zと、測定装置の角度ロール（roll）、ピッチ（pitch）、ヨー（yaw）で表現されてよい。

【0081】

横、縦、高さが各1mである正六面体の空間で、センサがx軸、y軸、z軸に対して0～1mの範囲で10cm間隔に動き、回転角度は0～360度の範囲でそれぞれ10度ずつ回転すると仮定するとき、可能な全体ポーズの場合の数は $11 \times 11 \times 11 \times 37 \times 37 \times 19$ 、すなわち、約3400万種類の場合の数が存在することになる。

10

【0082】

測定装置の移動間隔と回転角度の間隔を減少させれば不連続情報を連続情報のように見せることはできるが、可能なポーズの数は幾何級数的に増加するはずであり、現実空間の体積は 1m^3 よりも極めて大きくなるため、可能なすべてのポーズで空間情報を取得することは現実的に不可能である。

【0083】

このような理由により、空間情報を取得する段階では、現実空間を十分に含むことのできる一部のポーズでデータを取得し、前記取得された空間情報に基づいて深さ-映像連携情報を構成して3次元仮想空間モデルを構築することにより、一部のポーズで取得した空間情報を拡張させることができる。

20

【0084】

3次元仮想空間モデルは、一部のポーズで取得された空間情報に基づいて構成されてよいが、空間情報から構成された深さ-映像連携情報を活用するため、3次元仮想空間モデル内のいかなるポーズでも、同一の現実空間内のポーズで取得された空間情報または構成された深さ-映像連携情報と類似する対応情報を生成することができる。

【0085】

すなわち、空間情報が取得されないポーズであっても、3次元仮想空間モデルを使用すれば、該当のポーズでの深さ-映像連携情報または空間情報と類似の対応情報を生成することができ、前記生成された対応情報は、現実空間の同じポーズで取得した空間情報から構成された深さ-映像連携情報と類似すると考慮されてよい。3次元仮想空間モデルは、不連続情報である空間情報を連続情報である対応情報に変換させることができる。

30

【0086】

図10は、一実施形態における、3次元空間に対するユーザポーズ推定方法を説明するためのフローチャートである。

図10に示した方法は、図7に示したユーザポーズ推定装置620によって実行されてよい。

段階S1010で、装置は、深さ測定装置および映像取得装置を利用して3次元空間に対する深さ情報および映像情報を含む空間情報を取得する。

段階S1020で、装置は、空間情報に基づいて深さ-映像連携情報を構成し、深さ-映像連携情報に基づいて3次元空間に対応する3次元仮想空間モデルを構築する。

40

段階S1030で、装置は、3次元空間でユーザデバイスによって取得された映像を含むユーザ情報を受信する。このとき、ユーザ情報は、取得された映像に対応する空間の深さ情報をさらに含んでよい。

段階S1040で、装置は、3次元仮想空間モデル内でユーザ情報に対応する対応情報を生成する。

【0087】

3次元仮想空間モデルを利用すれば、空間情報が取得されないポーズであっても、該当のポーズで取得された空間情報または深さ-映像連携情報と類似する対応情報を生成することができる。

【0088】

50

対応情報は、深さ情報、映像情報、または深さ－映像連携情報で表現されてよい。3次元仮想空間モデル内で3自由度の基底ベクトルとして表現されるポーズで対応情報を生成してよい。

【0089】

例えば、ユーザ情報取得ポーズの高さが変わらないのであれば、3次元仮想空間モデル内で2自由度の基底ベクトルによって表現されるポーズで対応情報を生成してよい。対応情報は、視野角、映像情報変換、深さ情報変換などの過程を経て生成されてもよい。

【0090】

このとき、対応情報を生成する段階S1040は、ユーザ情報に含まれた映像で前記3次元空間の構造と関連する背景領域と前記3次元空間に置かれた物体に該当する非背景領域とを区分する段階、ユーザ情報に含まれた映像の背景領域を利用して前記ユーザ情報を加工する段階、および3次元仮想空間モデル内で加工されたユーザ情報に対応する対応情報を生成する段階を含んでよい。

10

【0091】

ユーザが3次元仮想空間モデルの背景となった現実空間でユーザ情報を取得するときの現実空間は、3次元仮想空間モデルを構成するために空間情報を取得した時点と等しくないことがあり、人物、事物、インテリアなどのような空間の様子が変わることがある。

【0092】

したがって、ユーザ情報で背景部分と非背景部分とを区分してユーザ情報から非背景部分を取り除いてもよく、背景部分を使用してユーザ情報を変換させてよい。照明や光などによる効果を取り除くためにユーザ情報を加工して使用してよい。ユーザ情報を3次元空間モデルで生成した対応情報と比較する過程において、ユーザ情報または対応情報の形態を変換して比較してよい。

20

【0093】

段階S1050で、装置は、対応情報とユーザ情報との類似度を算出する。

このとき、類似度を算出する段階は、類似度を高める方向に対応情報を再生成し、再生成された対応情報に基づいて類似度を再算出する段階を含んでよい。このとき、類似度を高める方向は、ユーザ情報を再取得したり、ユーザ情報に対応する対応情報を再生成したり、ユーザ情報の他に付加的な情報を使用したりすることを含む。

【0094】

類似度を高めるために、類似度を算出する段階S1050は、ユーザ情報および対応情報を比較するための比較対象領域を抽出する段階、ユーザ情報から抽出された比較対象領域と対応情報から抽出された比較対象領域で共通領域を決定する段階、および共通領域に基づいてユーザ情報および対応情報をそれぞれ再生成する段階を含んでよい。

30

【0095】

例えば、対応情報から構造の単純化によって歪曲した領域などのような所定の基準による一定の領域を取り除き、ユーザ情報の非背景部分に対応する領域を取り除くことにより、比較過程に使用される対応情報を再生成してよい。また、ユーザ情報で、非背景部分とともに、対応情報の歪曲した領域などに該当する領域を取り除くことにより、比較過程に使用されるユーザ情報を再生成してよい。

40

3次元仮想空間モデルで生成された対応情報とユーザが取得したユーザ情報との類似性を計算する方法は、対応情報の映像情報とユーザ情報の映像情報とを比較する方法、対応情報の深さ情報とユーザ情報の深さ情報とを比較する方法、または深さ－映像連携情報を比較する方法などであってよい。

【0096】

このとき、対応情報とユーザ情報の大きさ（scale）が異なることがあるため、正規化（Normalization）が要求されたり、相対的な比較が必要となることもある。

一方、映像情報の比較は、各映像情報の形式に類似するようにするための映像情報の変換が必要となることがある。例えば、パノラマイメージ（Panorama Image

50

）と補正イメージ（*Rectified image*）との変換があつてよく、映像情報の大きさを正規化してよく、視野角の変換があつてよい。

【0097】

これとは反対に、停留するイメージをパノラマ形式に変形させて使用することも可能である。2つの映像情報から、RANSAC、SIFT、FAST、SURFなどの手法、あるいはこれらの組み合わせを利用して映像情報の特徴点（*Feature*）を見つけ出し、類似する特徴点のペアを連結してよい。特徴点は、エッジ（*Edge*）、直線、線分、コーナー（*Corner*）、円、楕円など、またはこれらの組み合わせであつてよく、大きさ（*scale*）、回転（*rotation*）などが異なつてよい。特徴点マッチング（*Feature Matching*）、SSIM（*Structural Similarity*）、NID（*Normalized Information Distance*）、ホモグラフィ行列（*Homography Matrix*）などの手法によって映像情報の類似度を計算してよい。

10

【0098】

特徴点マッチングによって結ばれた多数のピクセル座標を使用してホモグラフィ行列を計算してよく、これを利用して2つのイメージ情報の差（誤差、*error*）を計算してよい。SSIMは2つのイメージの類似度を計算する方法であり、NIDは確率的な計算法である。

【0099】

ユーザ情報から深さ情報を抽出することができる場合、対応情報の深さ情報との類似度を比較してよい。深さ情報は、3次元点群情報（PCD：*Point Cloud Data*）、深さ地図（*Depth Map*）、メッシュ（*Mesh*）などで表現されてよく、2つの深さ情報の形式を統一する過程を必要としてもよい。深さ情報は、ピクセル間（ポイントあたり）が比較されてよく、周辺領域を考慮して比較してもよい。内挿法（*interpolation*）によって深さ情報を新たに推定し比較してもよく、加重値を付加して算出してもよい。

20

【0100】

ユーザ情報で深さー映像連携情報が構成可能な場合、対応情報の深さー映像連携情報と比較してよい。深さ情報と映像情報をそれぞれ比較して各類似度を算出し、全体類似度を算出してよく、各類似度に加重値を与えて算出してよい。また、深さー映像連携情報を複合的に比較してよく、深さ情報と映像情報の類似度を算出する方法を組み合わせで実行されてよい。

30

【0101】

3次元仮想空間モデルを構築するための空間情報が取得された時点とユーザ情報が取得された時点とが異なることがあるため、同一のポーズであっても、対応情報とユーザ情報が異なることがある。したがって、対応情報とユーザ情報で強靱な特徴点を比較してよい。例えば、対応情報とユーザ情報で背景部分と非背景部分とを区分し、背景部分を使用して類似度を算出してよく、背景部分を利用して構築された3次元仮想空間モデルで対応情報を生成してユーザ情報の背景部分との類似度を算出してよい。対応情報とユーザ情報から照明または光に対する光源効果を取り除いて類似度を算出してよく、光源効果に強靱な特徴を比較して類似度を算出してよい。

40

【0102】

このとき、類似度を算出する段階S1050は、ユーザデバイス周辺に対する追加のユーザ情報を取得する段階、およびユーザ情報および追加のユーザ情報に基づいて類似度を算出する段階を含んでよい。追加のユーザ情報を取得するために、図12に示すように案内情報を利用してよい。

【0103】

段階S1060で、装置は、類似度が予め設定された値以上である候補対応情報を確認し、前記候補対応情報にマッチングされるポーズをユーザポーズとして推定する。

【0104】

50

類似度が高いほど、対応情報が生成された3次元仮想空間モデルのポーズとユーザ情報が取得されたポーズが同じであると考慮されてよい。または、類似度が基準値(threshold)よりも高ければ、2つのデータが取得されて再構成されたポーズはほぼ同じであると思えてよく、基準値は現実空間の環境によって異なってもよい。または、多数の候補ポーズで生成された対応情報のうち、ユーザポーズと類似度が最も高いかいずれかの判断方式によって選択されたポーズが、ユーザポーズとして考慮されてもよい。

【0105】

対応情報生成、類似度計算段階を1度だけ実行してユーザポーズを推定してもよく、繰り返し実行してもよい。繰り返し実行は、選ばれたポーズの周りで精密に再推定してもよく、全体領域に対して無作為(random)に再推定してもよく、加重値を付加して再推定する新たなポーズを選択してもよい。このような段階は、定められた回数だけが繰り返されてもよく、類似度が基準値以上であるか、繰り返されて推定されたポーズが収束されるまで繰り返してもよい。類似度が高まるように最適化技法(Optimization)が利用されてもよい。

10

【0106】

類似度が高まるように対応情報を再生成してもよく、再生成された対応情報は、既存の対応情報が生成されたポーズと類似度との関係によってユーザポーズとして予想されるポーズで再生成されてよい。対応情報の再生成後に類似度を算出し、必要な場合には対応情報再生成および類似度算出過程を繰り返してよい。

【0107】

20

慣性情報および距離情報などのユーザ付加情報を使用することで類似度が高まるようになるため、期待されるポーズで対応情報を生成および再生成してよい。以後、対応情報とユーザ情報との類似度を算出し、必要な場合にはユーザ付加情報を活用して対応情報を再生成して類似度算出過程を繰り返してよい。

【0108】

このとき、ユーザ付加情報は、ユーザが取得した映像情報の他に、ユーザポーズの推定をサポートする情報であって、慣性情報(IMU)、距離情報(odometry)などで構成されてよい。一例として、慣性測定装置によって慣性情報の取得が可能な場合に、映像情報の処理時に映像取得ポーズに対する推定(prediction)情報として使用することにより、映像取得ポーズに対する補正(correction)をより容易に行うことができる。

30

【0109】

したがって、類似度を算出する段階S1050またはユーザポーズを推定する段階S1060は、ユーザデバイスによって前記ユーザポーズの推定に利用される付加的な情報であるユーザ付加情報が取得される場合、前記ユーザ情報または前記追加のユーザ情報とともに、前記ユーザ付加情報を利用して前記ユーザポーズを推定することを含んでよい。

【0110】

このとき、慣性情報の加速度値あるいは角速度値を活用して実際の移動距離を予想してよく、これを単一あるいは複数の映像測定装置から抽出された深さ情報の大きさ(scale)を補正するのに活用してもよい。

40

【0111】

距離情報は、ユーザが取得した映像情報に基づいて構成されたVO(Visual Odometry)、VIO(Visual Inertial Odometry)を活用して予測した距離情報であってよく、タイヤ型移動ロボットに測定装置を装着してユーザ情報を取得する場合、距離情報は、移動ロボットの距離情報となってよい。これにより、慣性情報を活用すれば、前記方法によって抽出された距離情報を補正するのに活用することが可能となる。

【0112】

ユーザの代わりにタイヤ型移動ロボットにセンサを装着してユーザ情報を取得する場合、ユーザが移動ロボットを操縦してもよく、移動ロボットが自律走行してもよく、この2

50

つの組み合わせによってユーザ情報を取得してよい。移動ロボットポーズをユーザポーズとして考慮してよく、移動ロボットとユーザ視野との座標変換関係が認知されていたり、座標変換が可能であれば、移動ロボットポーズをユーザポーズに変換させることが可能となる。

【0113】

移動ロボットは、映像が含まれたユーザ情報だけでなく、モバイルロボットの距離情報 (odometry) をユーザ付加情報として取得してよい。距離情報を活用してユーザポーズを補正してよい。順に取得した距離情報を活用して移動ロボットの相対予想ポーズを予測してよく、EKF、EIF、UKFなどの手法または類似の方法を活用して共分散行列 (Covariance matrix) などの情報を計算してよく、この情報を更新してユーザポーズを補正してよい。

10

移動ロボットを利用する場合、移動ロボットの動作、運転、操縦、移動、データ取得、記録、および処理など関連アルゴリズムは、ロボットオペレーティングシステム (ROS: robot operating system) 上で実行されてよい。

【0114】

空間情報、深さー映像連携情報、3次元仮想空間モデル、ユーザ情報、ユーザ付加情報などは、外部のサーバ (Server) に記録されて処理されてよい。

空間情報の取得と同時に深さー映像連携情報が構成されて3次元仮想空間モデルが構築されてよく、ユーザ情報の取得と同時にユーザポーズをリアルタイムで推定してもよく、遅延 (Latency) を発生させてもよく、ユーザポーズの取得が完了した後に処理されてもよい。

20

【0115】

3次元仮想空間モデルが構築されていれば、追加空間情報を取得しなくてもよく、一部の空間に対して追加空間情報が取得されてよい。追加空間情報が取得されなかった場合には構築された3次元仮想空間モデルを使用してよく、追加空間情報が取得された場合には、構築された3次元仮想空間モデルの部分または全体を更新して再構築して使用してよい。

【0116】

ユーザ情報を先に取得した後に空間情報を取得して3次元仮想空間モデルを構築してユーザポーズを推定してもよく、3次元仮想空間モデルを構築するための空間情報を先に取得した後にユーザ情報を取得してユーザポーズを推定してよい。

30

本発明は、センサシステムとコンピュータとが融合されたシステムで実行されてもよいし、独立的なセンサシステムとコンピュータで実行されてもよい。

【0117】

ユーザ情報の取得時に、各測定装置のポーズとユーザセンサシステム全体のポーズとが異なることがあるが、各測定装置およびセンサシステムの座標変換関係を利用して変換が可能である。例えば、ユーザセンサシステムの中心または適切な位置をユーザポーズとして仮定したり、ユーザセンサシステムを基準にユーザポーズとして仮定したりしてよい。この場合、必要なキャリブレーション情報またはユーザセンサシステムからユーザポーズまでの相対ポーズを認知しているか、いずれかの値で仮定してよい。

【0118】

図11は、他の実施形態における、3次元空間に対するユーザポーズ推定方法を説明するためのフローチャートである。

40

図11に示した方法は、図7に示したユーザポーズ推定装置620によって実行されてよい。

段階1110で、装置は、3次元空間で取得された映像を含むユーザ情報を受信する。

段階1120で、装置は、3次元空間に対する深さ情報および映像情報を含む空間情報に基づいて構築された3次元仮想空間モデルを確認する。このとき、3次元仮想空間モデルは、図7の仮想空間モデル提供部730によって提供されてよい。

【0119】

段階1130で、装置は、3次元仮想空間モデル内で前記ユーザ情報に対応する対応情

50

報を生成する。

段階 S 1 1 4 0 で、装置は、対応情報と前記ユーザ情報との類似度を算出する。

段階 S 1 1 5 0 で、装置は、類似度に基づいてユーザポーズを推定する。このとき、ユーザポーズは、例えば、ユーザ情報との類似度が最も高い対応情報のポーズであってよい。

【0 1 2 0】

図 1 2 は、一実施形態における、追加のユーザポーズ取得方法の例を説明するための図である。

類似度を向上させるために追加のユーザ情報を取得してよく、3次元仮想空間モデルを活用して追加のユーザ情報取得ポーズをユーザに案内してよく、ユーザは、案内されたポーズで追加のユーザ情報を取得してよい。

【0 1 2 1】

したがって、図 1 0 の説明において、追加のユーザ情報を取得する段階は、3次元仮想空間モデルに基づいて追加のユーザ情報取得のための案内情報をユーザデバイス 6 1 0 に送信することを含んでよい。

このとき、案内情報は、3次元仮想空間モデル内の予め設定された特徴点に対するユーザ情報取得ポーズを含み、追加のユーザ情報を取得する段階は、類似度を高める方向に繰り返し実行されてよい。

【0 1 2 2】

例えば、図 1 2 に示すように、類似の環境が多い長い廊下の場合、3次元仮想空間モデルで特徴点を考慮して追加のユーザ情報取得ポーズを案内してよい。

図 1 2 で、追加のユーザ情報取得ポーズは、特徴点 1、2、3 に対して順に映像を取得するポーズであるか、特徴点 1、2、3 のうちのいずれか 1 つに対するポーズであってよい。

【0 1 2 3】

上述した装置は、ハードウェア構成要素、ソフトウェア構成要素、および／またはハードウェア構成要素とソフトウェア構成要素との組み合わせによって実現されてよい。例えば、実施形態で説明された装置および構成要素は、例えば、プロセッサ、コントローラ、ALU (arithmetic logic unit)、デジタル信号プロセッサ、マイクロコンピュータ、FPGA (field programmable gate array)、PLU (programmable logic unit)、マイクロプロセッサ、または命令を実行して応答することができる様々な装置のように、1つ以上の汎用コンピュータまたは特殊目的コンピュータを利用して実現されてよい。処理装置は、オペレーティングシステム (OS) および前記 OS 上で実行される1つ以上のソフトウェアアプリケーションを実行してよい。また、処理装置は、ソフトウェアの実行に応答し、データにアクセスし、データを格納、操作、処理、および生成してもよい。理解の便宜のために、1つの処理装置が使用されるとして説明される場合もあるが、当業者は、処理装置が複数の処理要素および／または複数種類の処理要素を含んでもよいことが理解できるであろう。例えば、処理装置は、複数のプロセッサまたは1つのプロセッサおよび1つのコントローラを含んでもよい。また、並列プロセッサのような、他の処理構成も可能である。

【0 1 2 4】

ソフトウェアは、コンピュータプログラム、コード、命令、またはこれらのうちの1つ以上の組み合わせを含んでもよく、思うままに動作するように処理装置を構成したり、独立的または集散的に処理装置に命令したりしてよい。ソフトウェアおよび／またはデータは、処理装置に基づいて解釈されたり、処理装置に命令またはデータを提供したりするために、いかなる種類の機械、コンポーネント、物理装置、仮想装置、コンピュータ格納媒体または装置、または伝送される信号波に永久的または一時的に具現化されてよい。ソフトウェアは、ネットワークによって接続されたコンピュータシステム上に分散され、分散された状態で格納されても実行されてもよい。ソフトウェアおよびデータは、1つ以上のコンピュータ読み取り可能な記録媒体に格納されてもよい。

【0 1 2 5】

10

20

30

40

50

実施形態に係る方法は、多様なコンピュータ手段によって実行可能なプログラム命令の形態で実現されてコンピュータ読み取り可能な媒体に記録されてよい。前記コンピュータ読み取り可能な媒体は、プログラム命令、データファイル、データ構造などを単独でまたは組み合わせて含んでよい。前記媒体に記録されるプログラム命令は、実施形態のために特別に設計されて構成されたものであっても、コンピュータソフトウェアの当業者に公知な使用可能なものであってもよい。コンピュータ読み取り可能な記録媒体の例としては、ハードディスク、フロッピーディスク、および磁気テープのような磁気媒体、CD-ROM、DVDのような光媒体、フロプティカルディスク (floptical disk) のような光磁気媒体、およびROM、RAM、フラッシュメモリなどのようなプログラム命令を格納して実行するように特別に構成されたハードウェア装置が含まれる。プログラム命令の例は、コンパイラによって生成されるもののような機械語コードだけではなく、インタプリタなどを使用してコンピュータによって実行される高級言語コードを含む。上述したハードウェア装置は、実施形態の動作を実行するために1つ以上のソフトウェアモジュールとして動作するように構成されてもよく、その逆も同じである。

10

【0126】

以上のように、実施形態を、限定された実施形態と図面に基づいて説明したが、当業者であれば、上述した記載から多様な修正および変形が可能であろう。例えば、説明された技術が、説明された方法とは異なる順序で実行されたり、かつ／あるいは、説明されたシステム、構造、装置、回路などの構成要素が、説明された方法とは異なる形態で結合されたりまたは組み合わされたり、他の構成要素または均等物によって対置されたり置換されたりとしても、適切な結果を達成することができる。

20

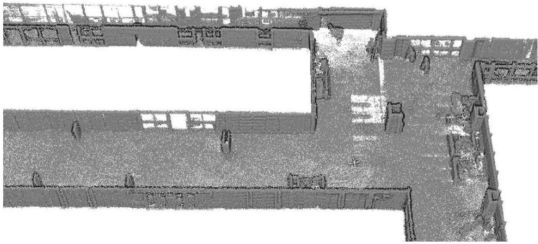
したがって、異なる実施形態であっても、特許請求の範囲と均等なものであれば、添付される特許請求の範囲に属する。

30

40

50

【図面】
【図 1】



【図 2】



(a)



(b)



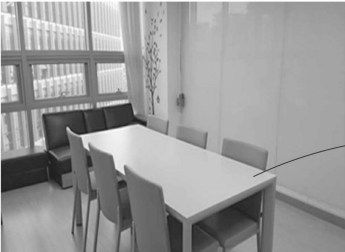
(c)

10

20

【図 3】

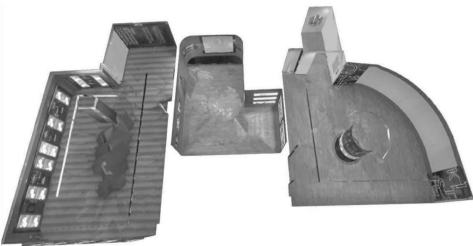
【図 4】



(a)



(b)



30

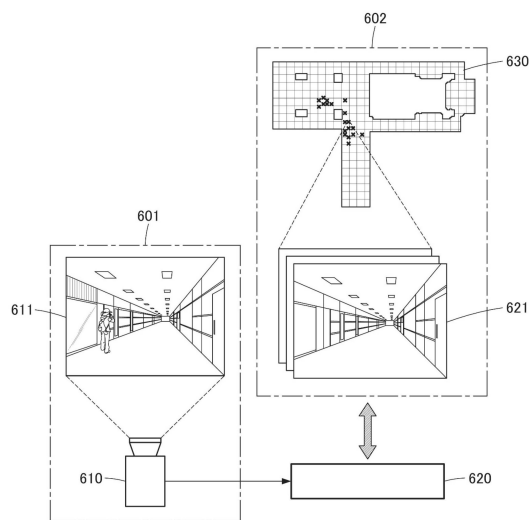
40

50

【図 5】



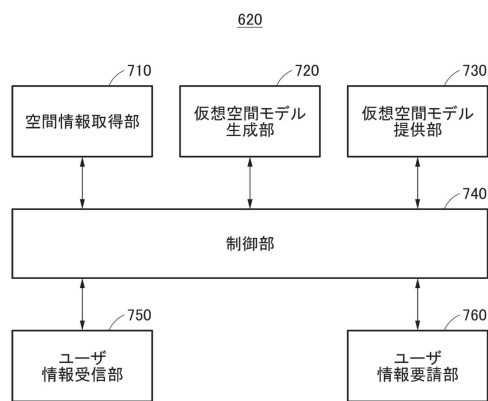
【図 6】



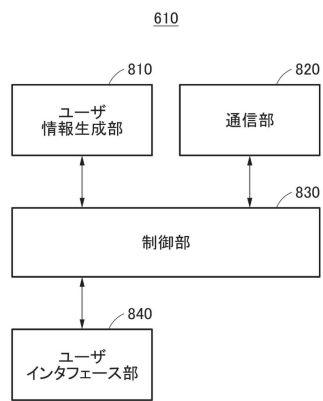
10

20

【図 7】



【図 8】

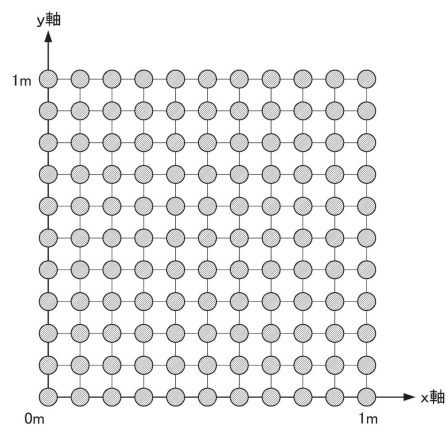


30

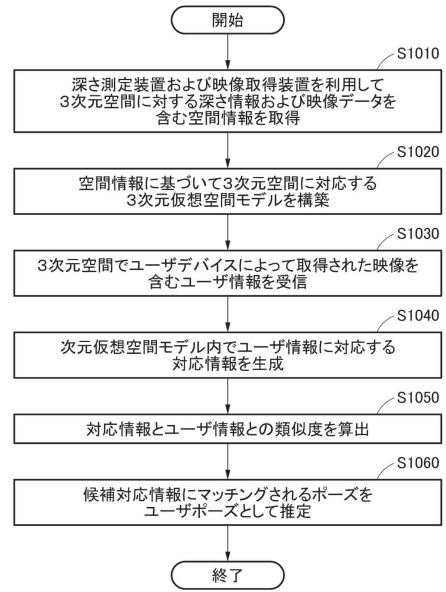
40

50

【図 9】



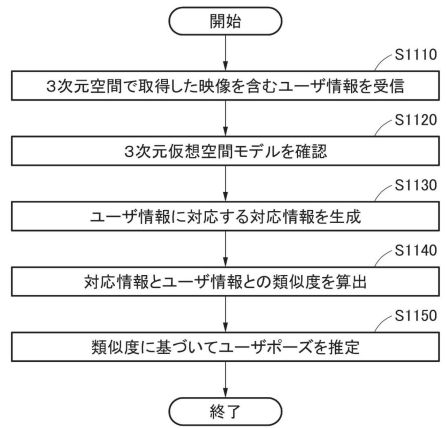
【図 10】



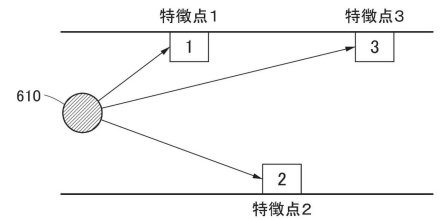
10

20

【図 11】



【図 12】



30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

韓国(KR)

ヒュンダイ アイ - パーク 1 チャ アパートメント 127

(72)発明者 イム ガヒョン

大韓民国、ソウル 02855、ソンブク - グ、インチョン - ロ 22 - ギル、5、ナンバー 407 アナム - ドン 5 - ガ

(72)発明者 ヒョン チャンフン

大韓民国、ソウル 02843、ソンブク - グ、コリョデ - ロ 17ガ - ギル、15 - 6、ナンバー 301

(72)発明者 キム ドンウ

大韓民国、ソウル 02473、トンデムン - グ、チェギ - ロ 5 - ギル、46、ナンバー 302

(72)発明者 チャン ブンチョル

大韓民国、ソウル 02578、トンデムン - グ、ムハク - ロ 47 - ギル、10、ナンバー 302

(72)発明者 チェ ヒョンエー

大韓民国、ソウル 02580、トンデムン - グ、ハンビト - ロ 20、ナンバー 1106

審査官 小池 正彦

(56)参考文献

特開2014-095553 (JP, A)

国際公開第2018/235219 (WO, A1)

特開2016-194779 (JP, A)

特表2018-528509 (JP, A)

特表2018-504663 (JP, A)

韓国公開特許第10-2017-0012316 (KR, A)

韓国登録特許第10-1835434 (KR, B1)

韓国公開特許第10-2017-0008505 (KR, A)

大江統子等、幾何学的位置合わせのための自然特徴点ランドマークデータベースを用いたカメラ位置・姿勢推定、日本バーチャルリアリティ学会論文誌、2005年10巻3号、日本、日本バーチャルリアリティ学会、285-294

(58)調査した分野 (Int.Cl., DB名)

G06T 7/70