

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2023-502972

(P2023-502972A)

(43)公表日 令和5年1月26日(2023.1.26)

(51)国際特許分類		F I			テーマコード (参考)
G 0 6 T	7/70 (2017.01)	G 0 6 T	7/70	A	5 C 0 5 4
H 0 4 N	7/18 (2006.01)	H 0 4 N	7/18	U	5 L 0 9 6

審査請求 未請求 予備審査請求 未請求 (全68頁)

(21)出願番号	特願2022-528326(P2022-528326)	(71)出願人	522190074
(86)(22)出願日	令和2年11月19日(2020.11.19)		トリゴ ビジョン リミテッド
(85)翻訳文提出日	令和4年5月13日(2022.5.13)		イスラエル国, 6 4 9 2 1 0 5 テル ア
(86)国際出願番号	PCT/IL2020/051198		ビブ, 1 5 0 メナヘム ベギン ストリ
(87)国際公開番号	WO2021/100043		ート
(87)国際公開日	令和3年5月27日(2021.5.27)	(74)代理人	100114775
(31)優先権主張番号	62/938,563		弁理士 高岡 亮一
(32)優先日	令和1年11月21日(2019.11.21)	(74)代理人	100121511
(33)優先権主張国・地域又は機関	米国(US)		弁理士 小田 直
(31)優先権主張番号	62/938,681	(74)代理人	100202751
(32)優先日	令和1年11月21日(2019.11.21)		弁理士 岩堀 明代
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100208580
			弁理士 三好 玲奈
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA	(74)代理人	100191086
	最終頁に続く		弁理士 高橋 香元
			最終頁に続く

(54)【発明の名称】 物品識別および追跡システム

(57)【要約】

物体に関連するデータを取得するための方法であって、シーンを視認するように多数のカメラを配置することであって、シーン内の少なくとも1つの基準物体が、少なくとも複数の多数のカメラによって視認可能であり、複数のカメラの各々が、それによって視認可能な基準物体の少なくとも1つの画像を取得する、配置することと、複数のカメラの各々を照射する光線交差点を発見することと、基準物体が各少なくとも1つの画像内に出現する画素位置を、シーン内の基準物体の三次元位置に関係なく、複数のカメラの各々を照射し、かつ交差領域と交差する光線に相関させることと、を含む、方法。

【選択図】 図1

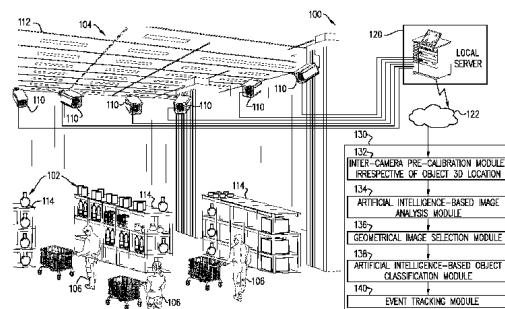


FIG. 1

【特許請求の範囲】

【請求項 1】

物体に関連するデータを取得するための方法であって、

シーンを視認するように多数のカメラを配置することであって、前記シーン内の少なくとも 1 つの基準物体が、少なくとも複数の前記多数のカメラによって視認可能であり、前記複数のカメラの各々が、それによって視認可能な前記基準物体の少なくとも 1 つの画像を取得する、配置することと、

前記複数のカメラの各々を照射する光線交差点を発見することと、

前記基準物体が各前記少なくとも 1 つの画像内に出現する画素位置を、前記シーン内の前記基準物体の三次元位置に関係なく、前記複数のカメラの各々を照射し、かつ前記交差領域と交差する前記光線に相関させることと、を含む、方法。

10

【請求項 2】

前記関連のステップが、前記複数のカメラの各々を照射し、かつ前記基準物体が各前記少なくとも 1 つの画像内に出現する前記画素位置に対応する前記光線の光線パラメータを導出することを含む、請求項 1 に記載の方法。

【請求項 3】

前記関連のステップに続いて、前記導出された光線パラメータを有する前記光線が前記交差点と所定の精度以上で交差するかどうかを確認することをさらに含む、請求項 2 に記載の方法。

【請求項 4】

前記導出された光線パラメータを有する前記光線は前記交差点と前記所定の精度以上で交差することが確認されるまで、前記発見および関連のステップが互いに対して反復的に実行される、請求項 3 に記載の方法。

20

【請求項 5】

前記交差点が、確率的な交差領域である、先行請求項のいずれか一項に記載の方法。

【請求項 6】

物体に関連するデータを取得するためのシステムであって、

シーンを視認するように配置された多数のカメラであって、前記シーン内の少なくとも 1 つの基準物体が、少なくとも複数の前記多数のカメラによって視認可能であり、前記複数のカメラの各々が、それによって視認可能な前記基準物体の少なくとも 1 つの画像を取得するように作動する、多数のカメラと、

30

前記複数のカメラの各々によって取得された前記少なくとも 1 つの画像を受信するように、かつ前記複数のカメラの各々を照射する光線交差点を発見するように作動する画像処理サブシステムと、

前記基準物体が各前記少なくとも 1 つの画像内に出現する画素位置を、前記シーン内の前記基準物体の三次元位置に関係なく、前記複数のカメラの各々を照射し、かつ前記交差点と交差する前記光線に相関させるように作動する、画素対光線較正サブシステムと、を備える、システム。

【請求項 7】

前記画素対光線較正サブシステムが、前記複数のカメラの各々を照射し、かつ前記基準物体が各前記少なくとも 1 つの画像内に出現する前記画素位置に対応する前記光線の光線パラメータを導出するように作動する、請求項 6 に記載のシステム。

40

【請求項 8】

前記システムは、前記導出された光線パラメータを有する前記光線が、前記交差点と所定の精度以上で交差するかどうかを確認するようにさらに作動する、請求項 7 に記載のシステム。

【請求項 9】

前記画像処理サブシステムおよび前記画素対光線較正サブシステムは、前記導出された光線パラメータを有する前記光線が前記交差点と前記所定の精度以上で交差することが確認されるまで、互いに対して反復的に作動する、請求項 8 に記載のシステム。

50

【請求項 10】

前記交差点が、確率的な交差領域である、請求項 6～9 のいずれか一項に記載のシステム。

【請求項 11】

物体に関連するデータを取得するための方法であって、

多数のカメラによって、複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得することと、

前記複数の画像のうちの少なくともいくつかの中の位置に出現する前記複数の物体のうちの少なくとも 1 つの物体の同一性を発見することと、

前記複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ前記複数の画像のうちの他の画像を前記共通の物体を示さないものとして拒否することであって、前記選択および前記拒否することが、前記多数のカメラの事前の相互校正に基づく、選択および拒否することと、

前記位置にある前記共通の物体を示す選択された画像のセットを出力することと、

前記少なくとも 1 つの物体の前記同一性および位置のうちの少なくとも 1 つを考慮することに基づいて、前記選択された画像のセットに示される前記共通の物体の同一性を発見することと、を含む、方法。

【請求項 12】

前記選択および前記拒否することが、前記複数の画像内の前記少なくとも 1 つの物体の前記同一性に関係なく実行される、請求項 11 に記載の方法。

【請求項 13】

前記複数の画像のうちの少なくともいくつかの中の位置に出現する前記複数の物体のうちの少なくとも 1 つの物体の同一性を発見する前記ステップに続いて、かつ前記選択するステップの前に、画像フィルタリングを実行することも含み、前記画像フィルタリングが、

前記複数の画像のうちの前記画像内の前記少なくとも 1 つの物体の前記同一性が、物体同一性のグループに属さず、前記少なくとも 1 つの物体の前記同一性が前記グループに属さない前記複数の画像のうちの前記画像が、前記選択ステップに参加する前記複数の画像に含まれていないこと、および

前記少なくとも 1 つの物体の前記同一性が事前定義された信頼水準を下回ると識別される信頼度であって、前記少なくとも 1 つの物体の前記同一性が前記事前定義された信頼水準を下回る前記信頼度で識別される前記複数の画像のうちの前記画像が、前記選択ステップに参加する前記複数の画像に含まれていない、信頼度、のうちの少なくとも 1 つに基づいて、前記複数の画像のうちの画像をフィルタリングすることを含む、請求項 11 に記載の方法。

【請求項 14】

前記グループが、類似の物体同一性の所定のグループを含む、請求項 13 に記載の方法。

【請求項 15】

前記グループが、類似の物体同一性の過去に学習されたカテゴリに基づく、請求項 13 に記載の方法。

【請求項 16】

前記複数の画像のうちの少なくともいくつかの中の位置に出現する前記複数の物体のうちの少なくとも 1 つの物体の同一性を発見する前記ステップが、

前記少なくとも 1 つの物体の前記同一性を発見するために人工知能（AI）を採用することを含み、前記 AI の採用が、

AI ネットワークを、多数の訓練画像を前記 AI ネットワークに提供することによって画像内の物体を識別するように訓練するための初期訓練段階であって、前記多数の訓練画像の各々中出现する少なくとも 1 つの物体が、前記 AI ネットワークに対して識別される、初期訓練段階と、

10

20

30

40

50

前記 A I ネットワークがその前記事前の訓練に基づいて少なくとも 1 つの物体の同一性を発見する前記ステップを実行するように作動する、後続の作動段階と、を含む、請求項 11 ~ 15 のいずれか一項に記載の方法。

【請求項 17】

前記多数の訓練画像の各々に出現し、かつ前記 A I ネットワークに対して識別された前記少なくとも 1 つの物体が、コンピュータビジョンを採用することに基づいて識別される、請求項 16 に記載の方法。

【請求項 18】

前記多数のカメラの前記事前の相互較正が、多数のカメラによって、複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得する前記ステップの前に、

前記シーンを視認するように前記多数のカメラを配置することであって、前記シーン内の少なくとも 1 つの基準物体が、少なくとも複数の前記多数のカメラによって視認可能であり、前記複数のカメラの各々が、それによって視認可能な前記基準物体の少なくとも 1 つの画像を取得する、配置することと、

前記複数のカメラの各々を照射する光線交差点を発見することと、

前記基準物体が各前記少なくとも 1 つの画像内に出現する画素位置を、前記シーン内の前記基準物体の三次元位置に関係なく、前記複数のカメラの各々を照射し、かつ前記交差領域と交差する前記光線に相関させ、それによって前記多数のカメラの前記複数のカメラに対する画素対光線較正を確立することと、

前記画素対光線較正が前記多数のカメラのすべてに対して確立されるまで、前記シーン内の前記少なくとも 1 つの基準物体を繰り返し再位置決めし、かつ前記基準物体がその各位置で視認可能である前記多数のカメラの前記複数のカメラに対して前記画素対光線較正を確立することと、を含む、請求項 11 ~ 17 のいずれか一項に記載の方法。

【請求項 19】

前記多数のカメラによって取得された複数の物体を含む前記シーンの前記複数の画像が、第 1 の解像度を有し、前記方法が、

前記複数の物体のうちの少なくとも 1 つの物体の同一性を発見する前記ステップの前に、前記複数の画像の前記第 1 の解像度を前記第 1 の解像度よりも低い第 2 の解像度に変換することであって、前記複数の画像のうちの画像を選択する前記ステップが、前記複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ前記複数の画像のうちの他の画像を前記共通の物体を示さないものとして拒否し、前記位置にある前記共通の物体を示す選択された画像のセットを出力する前記ステップが、前記第 2 の解像度を有する前記複数の画像に対して実行される、変換することと、

前記第 1 の解像度を有し、かつ前記第 2 の解像度を有する前記選択された画像のセットの画像に対応する、前記複数の画像のうちの画像を検索することと、

前記第 2 の解像度を有する前記選択された画像のセットに発見されるような前記共通の物体の前記位置に対応する領域内の前記第 1 の解像度を有する前記検索された画像を切り取ることと、

その前記切り取りに続く前記第 1 の解像度を有する前記画像の前記領域に出現する前記共通の物体の同一性を発見することと、をさらに含む、請求項 11 ~ 18 に記載の方法。

【請求項 20】

前記第 2 の解像度を有する前記画像に発見されるような前記共通の物体の前記同一性に関係なく、前記第 1 の解像度を有する前記画像の前記領域に出現する前記共通の物体の前記同一性が発見される、請求項 19 に記載の方法。

【請求項 21】

物体に関連するデータを取得するためのシステムであって、

複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得するように作動する、多数のカメラと、

前記複数の画像のうちの少なくともいくつかの中の位置に出現する前記複数の物体のうちの少なくとも 1 つの物体の同一性を発見するように作動する、画像分析モジュールと、

10

20

30

40

50

前記多数のカメラの事前の相互較正に基づいて、前記複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ前記複数の画像のうちの他の画像を前記共通の物体を示さないものとして拒否するように、かつ前記位置にある前記共通の物体を示す選択された画像のセットを出力するように作動する、画像選択モジュールと、

前記少なくとも1つの物体の前記同一性および位置のうちの少なくとも1つを考慮することに基づいて、前記選択された画像のセットに示される前記共通の物体の同一性を発見するように作動する、画像分類モジュールと、を備える、システム。

【請求項 2 2】

前記画像選択モジュールが、前記複数の画像内の前記少なくとも1つの物体の前記同一性に関係なく、前記複数の画像のうちの画像を選択し、かつ他の画像を拒否するように作動する、請求項 2 1 に記載のシステム。

10

【請求項 2 3】

前記画像分析モジュールの下流および前記画像選択モジュールの上流に画像フィルタリングモジュールも備え、前記画像フィルタリングモジュールが、

前記複数の画像のうちの前記画像内の前記少なくとも1つの物体の前記同一性が、物体同一性のグループに属さず、前記少なくとも1つの物体の前記同一性が前記グループに属さない前記複数の画像のうちの前記画像が、前記画像選択モジュールに渡されていないこと、および

前記少なくとも1つの物体の前記同一性が事前定義された信頼水準を下回ると識別される信頼度であって、前記少なくとも1つの物体の前記同一性が前記事前定義された信頼水準を下回る前記信頼度で識別される前記複数の画像のうちの前記画像が、前記画像選択モジュールに渡されていない、信頼度、のうちの少なくとも1つに基づいて、前記複数の画像のうちの画像をフィルタリングするように作動する、請求項 2 1 に記載のシステム。

20

【請求項 2 4】

前記グループが、類似の物体同一性の所定のグループを含む、請求項 2 3 に記載のシステム。

【請求項 2 5】

前記グループが、類似の物体同一性の過去に学習されたカテゴリに基づく、請求項 2 3 に記載のシステム。

【請求項 2 6】

前記画像分析モジュールが、前記少なくとも1つの物体の前記同一性を発見するために、人工知能（AI）を採用するように作動し、前記 AI の前記採用が、

AI ネットワークを、多数の訓練画像を前記 AI ネットワークに提供することによって画像内の物体を識別するように訓練するための初期訓練段階であって、前記多数の訓練画像の各々に出現する少なくとも1つの物体が、前記 AI ネットワークに対して識別される、初期訓練段階と、

前記 AI ネットワークがその前記事前の訓練に基づいて少なくとも1つの物体の同一性を発見する前記ステップを実行するように作動する、後続の作動段階と、を含む、請求項 2 1 ～ 2 5 のいずれか一項に記載のシステム。

【請求項 2 7】

前記多数の訓練画像の各々に出現し、かつ前記 AI ネットワークに対して識別された前記少なくとも1つの物体が、コンピュータビジョンを採用することに基づいて識別される、請求項 2 6 に記載のシステム。

40

【請求項 2 8】

前記多数のカメラの前記事前の相互較正が、

前記シーンを視認するように配置されている前記多数のカメラであって、前記シーン内の少なくとも1つの基準物体が、少なくとも複数の前記多数のカメラによって視認可能であり、前記複数のカメラの各々が、それによって視認可能な前記基準物体の少なくとも1つの画像を取得するように作動する、多数のカメラと、

前記複数のカメラの各々によって取得された前記少なくとも1つの画像を受信するよう

50

に、かつ前記複数のカメラの各々を照射する光線交差点を発見するように作動する、画像処理サブシステムと、

前記基準物体が各前記少なくとも1つの画像内に出現する画素位置を、前記シーン内の前記基準物体の三次元位置に関係なく、前記複数のカメラの各々を照射し、かつ前記交差点と交差する前記光線に相関させ、それによって前記多数のカメラの前記複数のカメラの画素対光線較正を確立するように作動する、画素対光線較正サブシステムと、を備え、

前記画素対光線較正が前記多数のカメラのすべてに対して確立されるまで、前記少なくとも1つの基準物体が、前記シーン内で繰り返し再位置決めされ、かつ前記画素対光線較正サブシステムは、前記基準物体がその各位置で視認可能である前記多数のカメラの前記複数のカメラに対して画素対光線較正を確立するように作動する、請求項21～27のいずれか一項に記載のシステム。

10

【請求項29】

前記多数のカメラによって取得された複数の物体を含む前記シーンの前記複数の画像が、第1の解像度を有し、前記システムが、

前記画像分析モジュールの上流にあり、かつ前記複数の画像の前記第1の解像度を前記第1の解像度よりも低い第2の解像度に変換するように作動する画像変換器であって、前記画像分析モジュールおよび前記画像選択モジュールが、前記第2の解像度を有する前記複数の画像に対して作動する、画像変換器をさらに備え、

前記画像分類モジュールが、

前記第1の解像度を有し、かつ前記第2の解像度を有する前記選択された画像のセットの画像に対応する、前記複数の画像のうちの画像を検索することと、

前記第2の解像度を有する前記選択された画像のセットに発見されるような前記共通の物体の前記位置に対応する領域において、前記第1の解像度を有する前記検索された画像を切り取ることと、

その切り取りに続く前記第1の解像度を有する前記画像の前記領域に出現する前記共通の物体の同一性を発見することと、を行うようにさらに作動する、請求項21～28のいずれか一項に記載のシステム。

20

【請求項30】

前記画像分類モジュールが、前記第2の解像度を有する前記画像に発見されるような前記共通物体の前記同一性に関係なく、前記第1の解像度を有する前記画像の前記領域に出現する前記共通の物体の前記同一性を発見するように作動する、請求項29に記載のシステム。

30

【請求項31】

物体に関連するデータを処理するための方法であって、

複数の画像の少なくとも1つの特性を、第1の特性から第2の特性に変換することと、

前記第2の特性を有する前記複数の画像のうちの少なくともいくつかの中に出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見することと、

前記少なくとも1つの特徴に基づいて、前記第1の特性を有する前記画像を修正することと、を含む、方法。

【請求項32】

前記少なくとも1つの特性が、画像解像度を含み、前記第1の特性が、第1の画像解像度を含み、前記第2の特性が、前記第1の画像解像度よりも低い第2の画像解像度を含み、請求項31に記載の方法。

40

【請求項33】

前記少なくとも1つの特徴が、前記少なくとも1つの物体の位置を含む、請求項32に記載の方法。

【請求項34】

前記修正することが、前記第2の解像度を有する前記画像に発見されるような前記少なくとも1つの物体の前記位置に基づいて、前記少なくとも1つの物体の前記位置に対応する領域内の前記第1の解像度を有する前記画像を切り取ることを含む、請求項33に記載

50

の方法。

【請求項 35】

物体に関連するデータを処理するためのシステムであって、

複数の画像の少なくとも1つの特性を、第1の特性から第2の特性に変換するように作動する、画像変換器と、

前記第2の特性を有する前記複数の画像のうちの少なくともいくつかの中に出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見するように作動する、画像分析器と、

前記少なくとも1つの特徴に基づいて前記第1の特性を有する前記画像を修正するように作動する、画像修正器と、を備える、システム。

10

【請求項 36】

前記少なくとも1つの特性が、画像解像度を含み、前記第1の特性が、第1の画像解像度を含み、前記第2の特性が、前記第1の画像解像度よりも低い第2の画像解像度を含む、請求項 35 に記載のシステム。

【請求項 37】

前記少なくとも1つの特徴が、前記少なくとも1つの物体の位置を含む、請求項 36 に記載のシステム。

【請求項 38】

前記画像修正器が、前記第2の解像度を有する前記画像に発見されるような前記少なくとも1つの物体の前記位置に基づいて、前記少なくとも1つの物体の前記位置に対応する領域内の前記第1の解像度を有する前記画像を切り取るように作動する、請求項 37 に記載のシステム。

20

【請求項 39】

物体に関連するデータを処理するための方法であって、

シーン内の少なくとも1つの物体の多次元モデルを経時的に作成することと、

前記少なくとも1つの物体の前記多次元モデルの変化を経時的に追跡することと、

少なくとも1つの所定の基準を満たす前記変化のうちの1つ以上を識別することと、

前記物体に関連する情報を導出するために、前記少なくとも1つの所定の基準を満たす前記変化に関連するデータを分析することと、を含む、方法。

【請求項 40】

30

物体に関連するデータを処理するためのシステムであって、

シーン内の少なくとも1つの物体の多次元モデルを経時的に作成するように作動するモデル作成器と、

前記少なくとも1つの物体の前記多次元モデルの変化を経時的に追跡し、かつ少なくとも1つの所定の基準を満たす前記変化のうちの1つ以上を識別するように作動する、モデル追跡器と、

前記物体に関連する情報を導出するために、前記少なくとも1つの所定の基準を満たす前記変化に関連するデータを分析するデータ分析器と、を備える、システム。

【請求項 41】

物体に関連する視覚データを処理するための方法であって、前記方法が、

40

少なくとも1台のカメラにシーンを視認させることと、

人工知能を採用することと、を含み、前記採用することは、

前記少なくとも1台のカメラからの少なくとも1つの出力が、前記シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびに

前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を含まない少なくとも1つの出力を、前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を前記出力に追加することによって強化すること、のために行われる、方法。

【請求項 42】

50

前記少なくとも1つの物体の前記存在および位置を示す前記追加情報が、前記シーンを視認するように配置された少なくとも2台の他のカメラからの出力に基づいて導出される、請求項41に記載の方法。

【請求項43】

前記強化することが、

前記少なくとも2台の他のカメラからの前記出力内の前記少なくとも1つの物体の前記存在および位置を示すバウンディングボックスと関連付けられた少なくとも1つの画素位置を識別することと、

各前記画素位置に対して、前記少なくとも2台の他のカメラについて、前記位置に対応するカメラ光線を発見することと、

前記カメラ光線のうちのカメラ光線の間でマッチングを実行して、前記カメラ光線の間の交差点を発見することであって、前記交差点が、前記物品位置に対応する、発見することと、

前記少なくとも1つの物体の前記存在および位置を示す情報を含まない出力を有する前記少なくとも1台のカメラについて、前記物品位置に対応するカメラ光線を発見することと、

前記少なくとも1つの物体の前記存在および位置を示す情報を含まない出力を有する前記少なくとも1台のカメラについて、前記カメラ光線に対応する画素位置を発見することと、

前記カメラ光線に対応する前記画素位置における前記少なくとも1台のカメラからの前記出力に新しいバウンディングボックスを挿入することであって、前記新しいバウンディングボックスが、前記少なくとも1つの物体の前記存在および位置を示す、挿入することと、を含む、請求項42に記載の方法。

【請求項44】

前記カメラ光線の前記発見、および前記カメラ光線に対応する画素位置の前記発見が、前記少なくとも1台のカメラおよび前記少なくとも2台の他のカメラの事前の相互校正に基づく、請求項43に記載の方法。

【請求項45】

前記少なくとも1つの物体が、人間以外の物品または人間の被写体のうちの少なくとも1つを含む、請求項41～44のいずれか一項に記載の方法。

【請求項46】

物体に関連するデータを処理するためのシステムであって、

シーンを視認するように配置された少なくとも1台のカメラと、

画像分析器と、を備え、前記画像分析器が、人工知能を採用し、かつ

前記少なくとも1台のカメラからの少なくとも1つの出力が、前記シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびに

前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を含まない少なくとも1つの出力を、前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を前記出力に追加することによって強化すること、を行うように作動する、システム。

【請求項47】

前記シーンを視認するように配置された少なくとも2台の他のカメラも備え、前記追加情報が、前記少なくとも2台の他のカメラからの前記出力に基づいて導出される前記少なくとも1つの物体の前記存在および位置を示す、請求項46に記載のシステム。

【請求項48】

前記画像分析器が、

前記少なくとも2台の他のカメラからの前記出力内の前記少なくとも1つの物体の前記存在および位置を示すバウンディングボックスと関連付けられた少なくとも1つの画素位置を識別することと、

10

20

30

40

50

各前記画素位置に対して、前記少なくとも 2 台の他のカメラについて、前記位置に対応するカメラ光線を発見することと、

前記カメラ光線のうちのカメラ光線の間でマッチングを実行して、前記カメラ光線の間の交差点を発見することであって、前記交差点が、前記物品位置に対応する、発見することと、

前記少なくとも 1 つの物体の前記存在および位置を示す情報を含まない出力を有する前記少なくとも 1 台のカメラについて、前記物品位置に対応するカメラ光線を発見することと、

前記少なくとも 1 つの物体の前記存在および位置を示す情報を含まない出力を有する前記少なくとも 1 台のカメラについて、前記カメラ光線に対応する画素位置を発見することと、

前記カメラ光線に対応する前記画素位置における前記少なくとも 1 台のカメラからの前記出力内に新しいバウンディングボックスを挿入することであって、前記新しいバウンディングボックスが、前記少なくとも 1 つの物体の前記存在および位置を示す、挿入することと、を行うように作動することによって、前記少なくとも 1 つの出力を強化するように作動する、請求項 47 に記載のシステム。

【請求項 49】

前記画像分析器が、前記少なくとも 1 台のカメラおよび前記少なくとも 2 台の他のカメラの事前の相互較正に基づいて、前記カメラ光線を発見し、かつ前記カメラ光線に対応する前記画素位置を発見するように作動する、請求項 48 に記載のシステム。

【請求項 50】

前記少なくとも 1 つの物体が、人間以外の物品または人間の被写体のうちの少なくとも 1 つを含む、請求項 46 ～ 49 のいずれか一項に記載のシステム。

【請求項 51】

物体に関連するデータを取得するための方法であって、

シーンを視認するように多数のカメラを配置することであって、前記シーン内の少なくとも 1 つの基準物体が、少なくとも複数の前記多数のカメラによって視認可能であり、前記複数のカメラの各々が、それによって視認可能な前記基準物体の少なくとも 1 つの画像を取得する、配置することと、

前記複数のカメラの各々を照射する光線交差点を発見することと、

前記基準物体が各前記少なくとも 1 つの画像内に出現する画素位置を、前記シーン内の前記基準物体の三次元位置に関係なく、前記複数のカメラの各々を照射し、かつ前記交差領域と交差する前記光線に相関させることと、を含み、

前記方法が、前記相関させるステップに続いて、

前記多数のカメラによって、複数の物体を含む前記シーンの複数の画像を少なくとも部分的に同時に取得すること、前記複数の画像のうちの少なくともいくつかの中の位置に出現する前記複数の物体のうちの少なくとも 1 つの物体の同一性を発見すること、前記複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ前記複数の画像のうちの他の画像を前記共通の物体を示さないものとして拒否することであって、前記選択および前記拒否することが、前記多数のカメラの事前の相互較正に基づく、選択および拒否すること、前記位置における前記共通物体を示す選択された画像のセットを出力すること、ならびに前記少なくとも 1 つの物体の前記同一性および位置のうちの少なくとも 1 つを考慮することに基づいて、前記選択された画像のセット内で示される前記共通の物体の同一性を発見すること、

ならびに

前記複数の画像のうちの少なくとも 1 つの特性を、第 1 の特性から第 2 の特性に変換すること、前記第 2 の特性を有する前記複数の画像のうちの少なくともいくつかの中に出現する少なくとも 1 つの物体の少なくとも 1 つの特徴を自動的に発見すること、および前記少なくとも 1 つの特徴に基づいて、前記第 1 の特性を有する前記画像を修正すること、

ならびに

10

20

30

40

50

前記シーン内の前記少なくとも1つの物体の多次元モデルを経時的に作成すること、前記少なくとも1つの物体の前記多次元モデルの変化を経時的に追跡すること、少なくとも1つの所定の基準を満たす前記変化のうちの1つ以上を識別すること、および前記物体に関連する情報を導出するために、前記少なくとも1つの所定の基準を満たす前記変化に関連するデータを分析すること、
ならびに

人工知能を採用することであって、前記多数のカメラの少なくとも1台のカメラからの少なくとも1つの出力が、前記シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびに前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を含まない少なくとも1つの出力を、前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を前記出力に追加することによって強化すること、のために人工知能を採用すること、のうちの少なくとも1つをさらに含む、方法。

10

【請求項52】

物体に関連するデータを取得するための方法であって、

多数のカメラによって、複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得することと、

前記複数の画像のうちの少なくともいくつかの中の位置に出現する前記複数の物体のうちの少なくとも1つの物体の同一性を発見することと、

前記複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ前記複数の画像のうちの他の画像を前記共通の物体を示さないものとして拒否することであって、前記選択および前記拒否することが、前記多数のカメラの事前の相互校正に基づく、選択および拒否することと、

20

前記位置にある前記共通の物体を示す選択された画像のセットを出力することと、

前記少なくとも1つの物体の前記同一性および位置のうちの少なくとも1つを考慮に入れることに基づいて、前記選択された画像のセットに示される前記共通の物体の同一性を発見することと、を含み、

前記方法が、

前記複数の画像のうちの少なくとも1つの特性を、第1の特性から第2の特性に変換すること、前記第2の特性を有する前記複数の画像のうちの少なくともいくつかの中に出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見すること、および前記少なくとも1つの特徴に基づいて、前記第1の特性を有する前記画像を修正すること、
ならびに

30

前記シーン内の前記少なくとも1つの物体の多次元モデルを経時的に作成すること、前記少なくとも1つの物体の前記多次元モデルの変化を経時的に追跡すること、少なくとも1つの所定の基準を満たす前記変化のうちの1つ以上を識別すること、および前記物体に関連する情報を導出するために、前記少なくとも1つの所定の基準を満たす前記変化に関連するデータを分析すること、
ならびに

人工知能を採用することであって、前記多数のカメラの少なくとも1台のカメラからの少なくとも1つの出力が、前記シーン内の前記少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびに前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を含まない少なくとも1つの出力を、前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を前記出力に追加することによって強化すること、のために人工知能を採用すること、のうちの少なくとも1つをさらに含む、方法。

40

【請求項53】

物体に関連するデータを取得するための方法であって、

少なくとも1つの物体を含むシーンの複数の画像のうちの少なくとも1つの特性を、第1の特性から第2の特性に変換することと、

50

前記第2の特性を有する前記複数の画像のうちの少なくともいくつかの中に出現する前記少なくとも1つの物体の少なくとも1つの特徴を自動的に発見することと、

前記少なくとも1つの特徴に基づいて、前記第1の特性を有する前記画像を修正することと、を含み、

前記方法が、

前記シーン内の前記少なくとも1つの物体の多次元モデルを経時的に作成すること、前記少なくとも1つの物体の前記多次元モデルの変化を経時的に追跡すること、少なくとも1つの所定の基準を満たす前記変化のうちの1つ以上を識別すること、および前記物体に関連する情報を導出するために、前記少なくとも1つの所定の基準を満たす前記変化に関連するデータを分析すること、

10

ならびに

人工知能を採用することであって、前記シーンを視認する多数のカメラのうちの少なくとも1台のカメラからの少なくとも1つの出力が、前記シーン内の前記少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびに前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を含まない少なくとも1つの出力を、前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を前記出力に追加することによって強化すること、のために人工知能を採用すること、のうちの少なくとも1つをさらに含む、方法。

【請求項54】

物体に関連するデータを取得するための方法であって、

20

シーン内の前記少なくとも1つの物体の多次元モデルを経時的に作成することと、

前記少なくとも1つの物体の前記多次元モデルの変化を経時的に追跡することと、

少なくとも1つの所定の基準を満たす前記変化のうちの1つ以上を識別することと、

前記物体に関連する情報を導出するために、前記少なくとも1つの所定の基準を満たす前記変化に関連するデータを分析することと、を含み、

前記方法が、

人工知能を採用することであって、前記シーンを視認する多数のカメラのうちの少なくとも1台のカメラからの少なくとも1つの出力が、前記シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびに前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を含まない少なくとも1つの出力を、前記シーン内の前記少なくとも1つの物体の前記存在および位置を示す情報を前記出力に追加することによって強化すること、のために人工知能を採用すること、のうちの少なくとも1つをさらに含む、方法。

30

【請求項55】

コンピュータによって実行されたときに、前記コンピュータに、請求項50～54のいずれか一項に記載の方法のステップのうちの少なくともいくつかのステップを実行させる命令が記憶されている、物体に関連するデータを処理するための非一時的なコンピュータ可読媒体。

【発明の詳細な説明】

【技術分野】

40

【0001】

関連出願

2019年11月21日に出願された「ITEM IDENTIFICATION AND TRACKING SYSTEM」と題する米国仮特許出願第62/938,681号、および2019年11月21日に出願された「IMAGE CONVERSION FOR IMAGE PROCESSING」と題する米国仮特許出願第62/938,563号が本明細書によって参照され、その開示は参照により本明細書に組み込まれ、その優先権は、米国特許法施行規則第1.78条(a)(4)および(5)(i)に従って本明細書によって主張される。

【0002】

50

本発明は、概して、データ分析、より具体的には、物品識別および追跡の目的でデータを分析するためのシステムおよび方法に関する。

【背景技術】

【0003】

物品を識別および追跡するための様々なタイプのシステムおよび方法が当技術分野で既知である。

【発明の概要】

【0004】

本発明は、自動化された、費用効果が高く、かつ時間効率の良い物品の識別および追跡の目的で、データの分析に関連する改善されたシステムおよび方法を提供しようとする。

【0005】

したがって、本発明の好ましい実施形態によれば、物体に関連するデータを取得するための方法であって、シーンを視認するように多数のカメラを配置することであって、シーン内の少なくとも1つの基準物体が、少なくとも複数の多数のカメラによって視認可能であり、複数のカメラの各々が、それによって視認可能な基準物体の少なくとも1つの画像を取得する、配置することと、複数のカメラの各々を照射する光線交差点を発見することと、基準物体が各少なくとも1つの画像内に出現する画素位置を、シーン内の基準物体の三次元位置に関係なく、複数のカメラの各々を照射し、かつ交差領域と交差する光線に相関させることと、を含む、方法が提供される。

【0006】

好ましくは、相関のステップは、複数のカメラの各々を照射し、かつ基準物体が各少なくとも1つの画像内に出現する画素位置に対応する光線の光線パラメータを導出することを含む。

【0007】

好ましくは、方法は、相関のステップに続いて、導出された光線パラメータを有する光線が交差点と所定の精度以上で交差するかどうかを確認することをさらに含む。

【0008】

好ましくは、導出された光線パラメータを有する光線が交差点と所定の精度以上で交差することが確認されるまで、発見および相関のステップが互いに対して反復的に実行される。

【0009】

好ましくは、交差点は、確率的な交差領域である。

【0010】

本発明の別の好ましい実施形態によれば、物体に関連するデータを取得するためのシステムであって、シーンを視認するように配置された多数のカメラであって、シーン内の少なくとも1つの基準物体が、少なくとも複数の多数のカメラによって視認可能であり、複数のカメラの各々が、それによって視認可能な基準物体の少なくとも1つの画像を取得するように作動する、多数のカメラと、複数のカメラの各々によって取得された少なくとも1つの画像を受信するように、かつ複数のカメラの各々を照射する光線交差点を発見するように作動する画像処理サブシステムと、基準物体が各少なくとも1つの画像内に出現する画素位置を、シーン内の基準物体の三次元位置に関係なく、複数のカメラの各々を照射し、かつ交差点と交差する光線に相関させるように作動する、画素対光線較正サブシステムと、を含む、システムがさらに提供される。

【0011】

好ましくは、画素対光線較正サブシステムは、複数のカメラの各々を照射し、かつ基準物体が各少なくとも1つの画像内に出現する画素位置に対応する光線の光線パラメータを導出するように作動する。

【0012】

好ましくは、システムは、導出された光線パラメータを有する光線が、交差点と所定の精度以上で交差するかどうかを確認するようにさらに作動する。

10

20

30

40

50

【0013】

好ましくは、画像処理サブシステムおよび画素対光線較正サブシステムは、導出された光線パラメータを有する光線が交差点と所定の精度以上で交差することが確認されるまで、互いに対して反復的に作動する。

【0014】

好ましくは、交差点は、確率的な交差領域である。

【0015】

本発明の別の好ましい実施形態によれば、多数のカメラによって、複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得することと、複数の画像のうちの少なくともいくつかの中の位置に出現する複数の物体のうちの少なくとも1つの物体の同一性を発見することと、複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ複数の画像のうちの他の画像を共通の物体を示さないものとして拒否することによって、選択および拒否することが、多数のカメラの事前の相互較正に基づく、選択および拒否することと、位置にある共通の物体を示す選択された画像のセットを出力することと、少なくとも1つの物体の同一性および位置のうちの少なくとも1つを考慮することに基づいて、選択された画像のセットに示される共通の物体の同一性を発見することと、を含む、方法もまた提供される。

【0016】

好ましくは、選択および拒否することは、複数の画像内の少なくとも1つの物体の同一性に関係なく実行される。

【0017】

好ましくは、この方法はまた、複数の画像のうちの少なくともいくつかの中の位置に出現する複数の物体のうちの少なくとも1つの物体の同一性を発見するステップに続いて、かつ選択するステップの前に、画像フィルタリングを実行することも含み、画像フィルタリングは、

複数の画像のうちの画像内の少なくとも1つの物体の同一性が、物体同一性のグループに属さず、少なくとも1つの物体の同一性がグループに属さない複数の画像のうちの画像が、選択ステップに参加する複数の画像に含まれていないこと、および

少なくとも1つの物体の同一性が事前定義された信頼水準を下回ると識別される信頼度であって、少なくとも1つの物体の同一性が事前定義された信頼水準を下回る信頼度で識別される複数の画像のうちの画像が、選択ステップに参加する複数の画像に含まれていない、信頼度、のうちの少なくとも1つに基づいて、複数の画像のうちの画像をフィルタリングすることを含む。

【0018】

好ましくは、グループは、類似の物体同一性の所定のグループを含む。

【0019】

代替的に、グループは、類似の物体同一性の過去に学習されたカテゴリに基づく。

【0020】

好ましくは、複数の画像のうちの少なくともいくつかの中の位置に出現する複数の物体のうちの少なくとも1つの物体の同一性を発見するステップは、少なくとも1つの物体の同一性を発見するために人工知能（AI）を採用することを含み、AIの採用は、AIネットワークを、多数の訓練画像をAIネットワークに提供することによって画像内の物体を識別するように訓練するための初期訓練段階であって、多数の訓練画像の各々中出现する少なくとも1つの物体が、AIネットワークに対して識別される、初期訓練段階と、AIネットワークがその事前の訓練に基づいて少なくとも1つの物体の同一性を発見するステップを実行するように作動する、後続の作動段階と、を含む。

【0021】

好ましくは、多数の訓練画像の各々中出现し、かつAIネットワークに対して識別された少なくとも1つの物体は、コンピュータビジョンを採用することに基づいて識別される。

10

20

30

40

50

【0022】

好ましくは、多数のカメラの事前の相互較正は、多数のカメラによって、複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得するステップの前に、シーンを視認するように多数のカメラを配置することであって、シーン内の少なくとも1つの基準物体が、少なくとも複数の多数のカメラによって視認可能であり、複数のカメラの各々が、それによって視認可能な基準物体の少なくとも1つの画像を取得する、配置することと、複数のカメラの各々を照射する光線交差点を発見することと、基準物体が各少なくとも1つの画像内に出現する画素位置を、シーン内の基準物体の三次元位置に関係なく、複数のカメラの各々を照射し、かつ交差領域と交差する光線に相関させ、それによって多数のカメラの複数のカメラに対する画素対光線較正を確立することと、画素対光線較正が多数のカメラのすべてに対して確立されるまで、シーン内の少なくとも1つの基準物体を繰り返し再位置決めし、かつ基準物体がその各位置で視認可能である多数のカメラの複数のカメラに対して画素対光線較正を確立することと、を含む。

10

【0023】

好ましくは、多数のカメラによって取得された複数の物体を含むシーンの複数の画像は、第1の解像度を有し、方法は、複数の物体のうちの少なくとも1つの物体の同一性を発見するステップの前に、複数の画像の第1の解像度を第1の解像度よりも低い第2の解像度に変換することであって、複数の画像のうちの画像を選択するステップが、複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ複数の画像のうちの他の画像を共通の物体を示さないものとして拒否し、位置にある共通の物体を示す選択された画像のセットを出力するステップが、第2の解像度を有する複数の画像に対して実行される、変換することと、第1の解像度を有し、かつ第2の解像度を有する選択された画像のセットの画像に対応する、複数の画像のうちの画像を検索することと、第2の解像度を有する選択された画像のセットに発見されるような共通の物体の位置に対応する領域内の第1の解像度を有する検索された画像を切り取ることと、その切り取りに続く第1の解像度を有する画像の領域に出現する共通の物体の同一性を発見することと、をさらに含む。

20

【0024】

好ましくは、第2の解像度を有する画像に発見されるような共通の物体の同一性に関係なく、第1の解像度を有する画像の領域に出現する共通の物体の同一性が発見される。

【0025】

本発明のさらに別の好ましい実施形態によれば、物体に関連するデータを取得するためのシステムであって、複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得するように作動する、多数のカメラと、複数の画像のうちの少なくともいくつかの中の位置に出現する複数の物体のうちの少なくとも1つの物体の同一性を発見するように作動する、画像分析モジュールと、多数のカメラの事前の相互較正に基づいて、複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ複数の画像のうちの他の画像を共通の物体を示さないものとして拒否するように、かつ位置にある共通の物体を示す選択された画像のセットを出力するように作動する、画像選択モジュールと、少なくとも1つの物体の同一性および位置のうちの少なくとも1つを考慮することに基づいて、選択された画像のセットに示される共通の物体の同一性を発見するように作動する、画像分類モジュールと、を備える、システムがまたさらに提供される。

30

40

【0026】

好ましくは、画像選択モジュールは、複数の画像内の少なくとも1つの物体の同一性に関係なく、複数の画像のうちの画像を選択し、かつ他の画像を拒否するように作動する。

【0027】

好ましくは、システムはまた、画像分析モジュールの下流および画像選択モジュールの上流に画像フィルタリングモジュールも含み、画像フィルタリングモジュールが、

複数の画像のうちの画像内の少なくとも1つの物体の同一性が、物体同一性のグループに属さず、少なくとも1つの物体の同一性がグループに属さない複数の画像のうちの画像が、画像選択モジュールに渡されていないこと、および

50

少なくとも1つの物体の同一性が事前定義された信頼水準を下回ると識別される信頼度であって、少なくとも1つの物体の同一性が事前定義された信頼水準を下回る信頼度で識別される複数の画像のうちの画像が、画像選択モジュールに渡されていない、信頼度、のうちの少なくとも1つに基づいて、複数の画像のうちの画像を除外するように作動する。

【0028】

好ましくは、グループは、類似の物体同一性の所定のグループを含む。

【0029】

代替的に、グループは、類似の物体同一性の過去に学習されたカテゴリに基づく。

【0030】

好ましくは、画像分析モジュールは、少なくとも1つの物体の同一性を発見するために、人工知能（AI）を採用するように作動し、AIの採用は、AIネットワークを、多数の訓練画像をAIネットワークに提供することによって画像内の物体を識別するように訓練するための初期訓練段階であって、多数の訓練画像の各々に出現する少なくとも1つの物体が、AIネットワークに対して識別される、初期訓練段階と、AIネットワークがその事前訓練に基づいて少なくとも1つの物体の同一性を発見するステップを実行するように作動する、後続の作動段階と、を含む。

【0031】

好ましくは、多数の訓練画像の各々に出現し、かつAIネットワークに対して識別された少なくとも1つの物体は、コンピュータビジョンを採用することに基づいて識別される。

【0032】

好ましくは、多数のカメラの事前の相互較正は、シーンを視認するように配置されている多数のカメラであって、シーン内の少なくとも1つの基準物体が、少なくとも複数の多数のカメラによって視認可能であり、複数のカメラの各々が、それによって視認可能な基準物体の少なくとも1つの画像を取得するように作動する、多数のカメラと、複数のカメラの各々によって取得された少なくとも1つの画像を受信するように、かつ複数のカメラの各々を照射する光線交差点を発見するように作動する、画像処理サブシステムと、基準物体が各少なくとも1つの画像内に出現する画素位置を、シーン内の基準物体の三次元位置に関係なく、複数のカメラの各々を照射し、かつ交差点と交差する光線に相関させ、それによって多数のカメラの複数のカメラの画素対光線較正を確立するように作動する、画素対光線較正サブシステムと、を含み、画素対光線較正が多数のカメラのすべてに対して確立されるまで、少なくとも1つの基準物体が、シーン内で繰り返し再位置決めされ、かつ画素対光線較正サブシステムが、基準物体がその各位置で視認可能である多数のカメラの複数のカメラに対して画素対光線較正を確立するように作動する。

【0033】

好ましくは、多数のカメラによって取得された複数の物体を含むシーンの複数の画像は、第1の解像度を有し、システムは、画像分析モジュールの上流にあり、かつ複数の画像の第1の解像度を第1の解像度よりも低い第2の解像度に変換するように作動する画像変換器をさらに含み、画像分析モジュールおよび画像選択モジュールが、第2の解像度を有する複数の画像に対して作動する、画像変換器と、画像分類モジュールが、第1の解像度を有し、かつ第2の解像度を有する選択された画像のセットの画像に対応する、複数の画像のうちの画像を検索することと、第2の解像度を有する選択された画像のセットに発見されるような共通物体の位置に対応する領域において、第1の解像度を有する検索された画像を切り取ることと、その切り取りに続く第1の解像度を有する画像の領域に出現する共通の物体の同一性を発見することと、を行うように作動する。

【0034】

好ましくは、画像分類モジュールは、第2の解像度を有する画像内で発見されるような共通物体の同一性に関係なく、第1の解像度を有する画像の領域に出現する共通の物体の同一性を発見するように作動する。

【0035】

10

20

30

40

50

本発明のさらに好ましい実施形態によれば、物体に関連するデータを処理するための方法であって、複数の画像の少なくとも1つの特性を、第1の特性から第2の特性に変換することと、第2の特性を有する複数の画像のうちの少なくともいくつかの中に出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見することと、少なくとも1つの特徴に基づいて、第1の特性を有する画像を修正することと、を含む、方法がさらに提供される。

【0036】

好ましくは、少なくとも1つの特性は、画像解像度を含み、第1の特性は、第1の画像解像度を含み、第2の特性は、第1の画像解像度よりも低い第2の画像解像度を含む。

【0037】

好ましくは、少なくとも1つの特徴は、少なくとも1つの物体の位置を含む。

【0038】

好ましくは、修正することは、第2の解像度を有する画像に発見されるような少なくとも1つの物体の位置に基づいて、少なくとも1つの物体の位置に対応する領域内の第1の解像度を有する画像を切り取ることを含む。

【0039】

本発明のさらに別の好ましい実施形態によれば、物体に関連するデータを処理するためのシステムであって、複数の画像の少なくとも1つの特性を、第1の特性から第2の特性に変換するように作動する、画像変換器と、第2の特性を有する複数の画像のうちの少なくともいくつかの中に出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見するように作動する、画像分析器と、少なくとも1つの特徴に基づいて第1の特性を有する画像を修正するように作動する、画像修正器と、を含む、システムもまた提供される。

【0040】

好ましくは、少なくとも1つの特性は、画像解像度を含み、第1の特性は、第1の画像解像度を含み、第2の特性は、第1の画像解像度よりも低い第2の画像解像度を含む。

【0041】

好ましくは、少なくとも1つの特徴は、少なくとも1つの物体の位置を含む。

【0042】

好ましくは、画像修正器は、第2の解像度を有する画像に発見されるような少なくとも1つの物体の位置に基づいて、少なくとも1つの物体の位置に対応する領域内の第1の解像度を有する画像を切り取るように作動する。

【0043】

本発明のまたさらに好ましい実施形態によれば、物体に関連するデータを処理するための方法であって、シーン内の少なくとも1つの物体の多次元モデルを経時的に作成することと、少なくとも1つの物体の多次元モデルの変化を経時的に追跡することと、少なくとも1つの所定の基準を満たす変化のうちの1つ以上を識別することと、物体に関連する情報を導出するために、少なくとも1つの所定の基準を満たす変化に関連するデータを分析することと、を含む、方法がまたさらに提供される。

【0044】

本発明のさらに好ましい実施形態によれば、物体に関連するデータを処理するためのシステムであって、シーン内の少なくとも1つの物体の多次元モデルを経時的に作成するモデル作成器と、少なくとも1つの物体の多次元モデルの変化を経時的に追跡し、かつ少なくとも1つの所定の基準を満たす変化のうちの1つ以上を識別するように作動する、モデル追跡器と、物体に関連する情報を導出するために、少なくとも1つの所定の基準を満たす変化に関連するデータを分析するデータ分析器と、を含む、システムがまたさらに提供される。

【0045】

本発明の別の好ましい実施形態によれば、物体に関連する視覚データを処理するための方法であって、方法が、少なくとも1つのカメラにシーンを視認させることと、人工知能

10

20

30

40

50

を採用することと、を含み、採用することが、少なくとも1台のカメラからの少なくとも1つの出力が、シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびにシーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびにシーン内の少なくとも1つの物体の存在および位置を示す情報を出力に追加することによって強化すること、のために行われる、方法がさらに提供される。

【0046】

好ましくは、少なくとも1つの物体の存在および位置を示す追加情報は、シーンを視認するように配置された少なくとも2台の他のカメラからの出力に基づいて導出される。

【0047】

好ましくは、強化することは、少なくとも2台の他のカメラからの出力内の少なくとも1つの物体の存在および位置を示すバウンディングボックスと関連付けられた少なくとも1つの画素位置を識別することと、各画素位置に対して、少なくとも2台の他のカメラについて、位置に対応するカメラ光線を発見することと、カメラ光線のうちのカメラ光線間でマッチングを実行して、カメラ光線の間の交差点を発見することであって、交差点が、物品位置に対応する、発見することと、少なくとも1つの物体の存在および位置を示す情報を含む出力を有する少なくとも1台のカメラについて、物品位置に対応するカメラ光線を発見することと、少なくとも1つの物体の存在および位置を示す情報を含む出力を有する少なくとも1台のカメラについて、カメラ光線に対応する画素位置を発見することと、カメラ光線に対応する画素位置における少なくとも1台のカメラからの出力に新しいバウンディングボックスを挿入することであって、新しいバウンディングボックスが、少なくとも1つの物体の存在および位置を示す、挿入することと、を含む。

【0048】

好ましくは、カメラ光線の発見、およびカメラ光線に対応する画素位置の発見は、少なくとも1台のカメラおよび少なくとも2台の他のカメラの事前の相互校正に基づく。

【0049】

好ましくは、少なくとも1つの物体は、人間以外の物品または人間の被写体のうちの少なくとも1つを含む。

【0050】

本発明のさらに別の好ましい実施形態によれば、物体に関連するデータを処理するためのシステムであって、シーンを視認するように配置された少なくとも1台のカメラと、画像分析器と、を備え、画像分析器が、人工知能を採用し、かつ少なくとも1つのカメラからの少なくとも1つの出力が、シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびにシーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびにシーン内の少なくとも1つの物体の存在および位置を示す情報を出力に追加することによって強化すること、を行うように作動する、システムがまたさらに提供される。

【0051】

好ましくは、システムはまた、シーンを視認するように配置された少なくとも2台の他のカメラも含み、追加情報が、少なくとも2台の他のカメラからの出力に基づいて導出される少なくとも1つの物体の存在および位置を示す。

【0052】

好ましくは、画像分析器は、少なくとも2台の他のカメラからの出力内の少なくとも1つの物体の存在および位置を示すバウンディングボックスと関連付けられた少なくとも1つの画素位置を識別することと、各画素位置に対して、少なくとも2台の他のカメラについて、位置に対応するカメラ光線を発見することと、カメラ光線のうちのカメラ光線の間でマッチングを実行して、カメラ光線の間の交差点を発見することであって、交差点が、物品位置に対応する、発見することと、少なくとも1つの物体の存在および位置を示す情報を含む出力を有する少なくとも1台のカメラについて、物品位置に対応するカメラ光線を発見することと、少なくとも1つの物体の存在および位置を示す情報を含む出力

力を有する少なくとも1台のカメラについて、カメラ光線に対応する画素位置を発見することと、カメラ光線に対応する画素位置における少なくとも1台のカメラからの出力内に新しいバウンディングボックスを挿入することであって、新しいバウンディングボックスが、少なくとも1つの物体の存在および位置を示す、挿入することと、を行うように作動することによって、少なくとも1つの出力を強化するように作動する。

【0053】

好ましくは、画像分析器は、少なくとも1台のカメラおよび少なくとも2台の他のカメラの事前の相互較正に基づいて、カメラ光線を発見し、かつカメラ光線に対応する画素位置を発見するように作動する。

【0054】

好ましくは、少なくとも1つの物体は、人間以外の物品または人間の被写体のうちの少なくとも1つを含む。

【0055】

本発明の別の好ましい実施形態によれば、物体に関連するデータを取得するための方法であって、シーンを視認するように多数のカメラを配置することであって、シーン内の少なくとも1つの基準物体が、少なくとも複数の多数のカメラによって視認可能であり、複数のカメラの各々が、それによって視認可能な基準物体の少なくとも1つの画像を取得する、配置することと、複数のカメラの各々を照射する光線交差点を発見することと、基準物体が各少なくとも1つの画像内に出現する画素位置を、シーン内の基準物体の三次元位置に関係なく、複数のカメラの各々を照射し、かつ交差領域と交差する光線に相関させることと、を含み、方法が、相関させるステップに続いて、

多数のカメラによって、複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得すること、複数の画像のうちの少なくともいくつかの中の位置に出現する複数の物体のうちの少なくとも1つの物体の同一性を発見すること、複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ複数の画像のうちの他の画像を共通の物体を示さないものとして拒否することであって、選択および拒否することが、多数のカメラの事前の相互較正に基づく、選択および拒否すること、位置における共通物体を示す選択された画像のセットを出力すること、ならびに少なくとも1つの物体の同一性および位置のうちの少なくとも1つを考慮することに基づいて、選択された画像のセット内で示される共通の物体の同一性を発見すること、

複数の画像のうちの少なくとも1つの特性を、第1の特性から第2の特性に変換すること、第2の特性を有する複数の画像のうちの少なくともいくつかの中に出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見すること、および少なくとも1つの特徴に基づいて、第1の特性を有する画像を修正すること、

シーン内の少なくとも1つの物体の多次元モデルを経時的に作成すること、少なくとも1つの物体の多次元モデルの変化を経時的に追跡すること、少なくとも1つの所定の基準を満たす変化のうちの1つ以上を識別すること、および物体に関連する情報を導出するために、少なくとも1つの所定の基準を満たす変化に関連するデータを分析すること、ならびに

人工知能を採用することであって、多数のカメラのうちの少なくとも1台のカメラからの少なくとも1つの出力が、シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびにシーン内の少なくとも1つの物体の存在および位置を示す情報を含まない少なくとも1つの出力を、シーン内の少なくとも1つの物体の存在および位置を示す情報を出力に追加することによって強化すること、のために人工知能を採用すること、のうちの少なくとも1つをさらに含む、方法がさらに提供される。

【0056】

本発明の別の好ましい実施形態によれば、物体に関連するデータを取得するための方法であって、多数のカメラによって、複数の物体を含むシーンの複数の画像を少なくとも部分的に同時に取得することと、複数の画像のうちの少なくともいくつかの中の位置に出現

10

20

30

40

50

する複数の物体のうちの少なくとも1つの物体の同一性を発見することと、複数の画像のうちの画像を共通の物体を示すものとして選択し、かつ複数の画像のうちの他の画像を共通の物体を示さないものとして拒否することと、選択および拒否することが、多数のカメラの事前の相互較正に基づく、選択および拒否することと、位置にある共通の物体を示す選択された画像のセットを出力することと、少なくとも1つの物体の同一性および位置のうちの少なくとも1つを考慮に入れることに基づいて、選択された画像のセットに示される共通の物体の同一性を発見することと、を含み、方法が、

複数の画像のうちの少なくとも1つの特性を、第1の特性から第2の特性に変換すること、第2の特性を有する複数の画像のうちの少なくともいくつかの中に出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見すること、および少なくとも1つの特徴に基づいて、第1の特性を有する画像を修正すること、

シーン内の少なくとも1つの物体の多次元モデルを経時的に作成すること、少なくとも1つの物体の多次元モデルの変化を経時的に追跡すること、少なくとも1つの所定の基準を満たす変化のうちの1つ以上を識別すること、および物体に関連する情報を導出するために、少なくとも1つの所定の基準を満たす変化に関連するデータを分析すること、ならびに

人工知能を採用することとあって、多数のカメラのうちの少なくとも1台のカメラからの少なくとも1つの出力が、シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびにシーン内の少なくとも1つの物体の存在および位置を示す情報を含まない少なくとも1つの出力を、シーン内の少なくとも1つの物体の存在および位置を示す情報を出力に追加することによって強化すること、のために人工知能を採用すること、のうちの少なくとも1つをさらに含む、方法がさらに提供される。

【0057】

本発明のさらに別の好ましい実施形態によれば、物体に関連するデータを取得するための方法とあって、少なくとも1つの物体を含むシーンの複数の画像のうちの少なくとも1つの特性を、第1の特性から第2の特性に変換することと、第2の特性を有する複数の画像のうちの少なくともいくつかの中に出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見することと、少なくとも1つの特徴に基づいて、第1の特性を有する画像を修正することと、を含み、方法が、

シーン内の少なくとも1つの物体の多次元モデルを経時的に作成すること、少なくとも1つの物体の多次元モデルの変化を経時的に追跡すること、少なくとも1つの所定の基準を満たす変化のうちの1つ以上を識別すること、および物体に関連する情報を導出するために、少なくとも1つの所定の基準を満たす変化に関連するデータを分析すること、ならびに

人工知能を採用することとあって、シーンを視認する多数のカメラのうちの少なくとも1台のカメラからの少なくとも1つの出力が、シーン内の少なくとも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびにシーン内の少なくとも1つの物体の存在および位置を示す情報を含まない少なくとも1つの出力を、シーン内の少なくとも1つの物体の存在および位置を示す情報を出力に追加することによって強化すること、のために人工知能を採用すること、のうちの少なくとも1つをさらに含む、方法がさらに提供される。

【0058】

本発明のさらに好ましい実施形態によれば、物体に関連するデータを取得するための方法とあって、シーン内の少なくとも1つの物体の多次元モデルを経時的に作成することと、少なくとも1つの物体の多次元モデルの変化を経時的に追跡することと、少なくとも1つの所定の基準を満たす変化のうちの1つ以上を識別することと、物体に関連する情報を導出するために、少なくとも1つの所定の基準を満たす変化に関連するデータを分析することと、を含み、方法が、人工知能を採用することとあって、シーンを視認する多数のカメラのうちの少なくとも1台のカメラからの少なくとも1つの出力が、シーン内の少なく

とも1つの物体の存在および位置を示す情報を含むかまたは含まないかを確認すること、ならびにシーン内の少なくとも1つの物体の存在および位置を示す情報を含まない少なくとも1つの出力を、シーン内の少なくとも1つの物体の存在および位置を示す情報を出力に追加することによって強化すること、のために人工知能を採用すること、のうちの少なくとも1つをさらに含む、方法がまたさらに提供される。

【0059】

本発明の好ましい実施形態によれば、コンピュータによって実行されたときに、コンピュータに、上述された方法のステップのうちの少なくともいくつかのステップを実行させる命令が記憶されている、物体に関連するデータを処理するための非一時的なコンピュータ可読媒体もまた提供される。

【図面の簡単な説明】

【0060】

本発明は、図面と併せて以下の詳細な説明に基づいて、より完全に理解および認識されるであろう。

【0061】

【図1】本発明の好ましい実施形態に従って構築され、作動する、物品識別および追跡のためのシステムの部分的に簡略化されかつ部分的に絵で示されたブロック説明図である。

【図2A】図1に示されるタイプの物品識別および追跡システムの一部を形成する較正サブシステムの構成要素の簡略化された概略説明図である。

【図2B】図2Aに示されるタイプの較正サブシステムの操作におけるステップを示す簡略化されたフローチャートである。

【図3】図1に示されるタイプの物品識別および追跡システムの一部を形成する、画像分析および分類機能を有するシステム構成要素の簡略化されたブロック説明図である。

【図4】図3に示されるタイプの画像分析および分類構成要素によって実行される機械学習を示す簡略化されたブロック図である。

【図5】図1に示されるタイプの物品識別および追跡システムで採用されるカメラの配置を最適化するのに役立つスコアリングを示す簡略化されたグラフである。

【図6】図1に示されるタイプの物品識別および追跡システムにおいて有用な物品を追跡するためのモデルの簡略化された概略説明図である。

【図7-8】図6に示されるタイプのモデルに基づいた物品の追跡におけるステップを示す簡略化されたそれぞれのフローチャートである。

【図9A-9B】本発明のシステムの実施形態の実装形態の簡略化された斜視説明図および側面説明図である。

【図10A-10B、11A-11B、12A-12B】図9Aおよび9Bの実装形態の一部の簡略図であり、本発明の好ましい実施形態による、AIALゴリズムによって注釈付けされた、人間の被写体を含むコンテキストにおける複数のカメラの視線および対応する画像を示している。

【図13A】本発明の好ましい実施形態による、画像修正によるAI失敗ケースの改善を示す、図9Aおよび9Bの実装形態の一部の簡略図である。

【図13B】図13Aの実施形態に従って実行されるAI失敗ケースの改善におけるステップを示す簡略化されたフローチャート図である。

【図14A-14B】本発明の好ましい実施形態による、3DモデリングによるAI失敗ケースの改善を示す、図9Aおよび9Bの実装形態の一部の簡略化された図である。

【図14C】図14Bの実施形態に従って実行されるAI失敗ケースの改善におけるステップを示す簡略化されたフローチャート図である。

【図14D】図14Bの実施形態に従って実行されるAI失敗ケースの改善の段階に含まれるステップを示す簡略化されたフローチャート図である。

【図15A-15B】図9Aおよび9Bの実装形態の一部の簡略図であり、本発明の別の好ましい実施形態による、AIALゴリズムによって注釈付けされた、人間の被写体を含む文脈における複数のカメラの視線および対応する画像を示している。

10

20

30

40

50

【図 1 6 A】本発明のさらに別の好ましい実施形態による、画像修正による A I 失敗ケースの改善を示す、図 9 A および 9 B の実装形態の一部の簡略化された図である。

【図 1 6 B】図 1 6 A の実施形態に従って実行される A I 失敗ケースの改善におけるステップを示す簡略化されたフローチャート図である。

【図 1 7 A - 1 7 B】本発明のさらに好ましい実施形態による、3 D モデリングによる A I 失敗ケースの改善を示す、図 9 A および 9 B の実装形態の一部の簡略図である。

【図 1 7 C】図 1 7 B の実施形態に従って実行される A I 失敗ケースの改善におけるステップを示す簡略化されたフローチャート図である。

【図 1 7 D】図 1 7 B の実施形態に従って実行される A I 失敗ケースの改善の段階に含まれるステップを示す簡略化されたフローチャート図である。

【図 1 8】バウンディングボックスの厳密化を示す図 9 A および 9 B の実装形態の一部の簡略化された図である。

【発明を実施するための形態】

【0 0 6 2】

次に図 1 を参照すると、図 1 は、本発明の好ましい実施形態に従って構築され、作動する、物品識別および追跡のためのシステムの部分的に簡略化されかつ部分的に絵で示されたブロック説明図である。

【0 0 6 3】

図 1 に見られるように、物品 1 0 2 の識別および追跡のためのシステム 1 0 0 が提供される。物品 1 0 2 は、建物 1 0 4 内に位置付けることができ、ここでは、例としてのみ、買い物施設 1 0 4 内に位置付けられ、買い物客 1 0 6 によって取り扱われる消費者物品 1 0 2 として具体化される。しかしながら、システム 1 0 0 は、倉庫、空港、製造施設、およびレストラン含むがこれらに限定されない任意の好適な施設内で手動または自動で処理され得る任意のタイプの単一または複数の物品の識別および追跡のために実装され得ることが理解される。

【0 0 6 4】

システム 1 0 0 は、好ましくは、建物 1 0 4 などのシーンを視認し、その中の少なくとも 1 つの物体、例えば、物品 1 0 2 および／または買い物客 1 0 6 のうちの少なくとも 1 つに関連する視覚データを取得するように配置された多数のカメラ 1 1 0 を含む。多数のカメラ 1 1 0 は、好ましくは、必ずしもではないが、建物 1 0 4 内の固定表面、例えば、建物 1 0 4 の天井 1 1 2 または棚 1 1 4 上に取り付けられ、好ましくは、建物 1 0 4 をそれに対して様々な角度から視認するように相互に空間的に分散される。

【0 0 6 5】

カメラ 1 1 0 は、好ましくは、内部に位置付けられた物品 1 0 2 および／または買い物客 1 0 6 を含む、建物 1 0 4 の画像を取得するように作動する。カメラ 1 1 0 は、システム 1 0 0 の画像化要件に応じて、建物 1 0 4 の画像を少なくともほぼリアルタイムで連続的に取得するように連続的に作動するか、または相互に調整されたストロボ方式で作動することができる。一例として、カメラ 1 1 0 は、2 5 画像フレーム／カメラ／秒など、時間内の単位ごとにカメラごとに複数の画像を取得するように作動することができる。本発明の特定の実施形態では、カメラ 1 1 0 は 2 D カメラであり得る。代替的に、本発明の特に好ましい実施形態では、カメラ 1 1 0 は、深度画像を出力する 3 D 深度カメラであり得る。

【0 0 6 6】

カメラ 1 1 0 によって取得された画像のうちの少なくともいくつかは、物品識別器および追跡器 1 3 0 によって処理するために、ローカルサーバ 1 2 0 および／またはクラウドベースのサーバ 1 2 2 に提供され得る。画像は、サーバ 1 2 0 において完全にローカルに処理されるか、サーバ 1 2 0 において部分的に処理されるか、クラウドベースのサーバ 1 2 2 において部分的に処理されるか、または処理のためにクラウドベースのサーバ 1 2 2 に完全に転送され得る。ローカルサーバ 1 2 0 および／またはクラウドベースのサーバ 1 2 2 に提供される画像のうちの少なくともいくつかは、所定の基準に基づいて、ローカル

10

20

30

40

50

またはリモートのいずれかで保存され得る。

【0067】

物品識別器および追跡器130の機能は、プロセッサによって、例えば、ローカルサーバ120および／またはクラウドベースのサーバ122のプロセッサによって実行され得る。本発明の実施形態によれば、コンピュータ可読媒体（例えば、レジスタメモリ、プロセッサキャッシュ、RAM、ROM、ハードドライブ、フラッシュメモリ、CD ROM、磁気媒体など）に記憶されたコンピュータプログラムアプリケーションは、実行されたときに、コントローラまたはプロセッサに、自動化された物品識別および追跡のための方法など、本明細書で論じられた機能および方法のうちの1つ以上を実行するように指示するまたは引き起こすことができるコードまたは実行可能命令を含み得る。コンピュータ可読媒体は、すべての形態およびタイプのコンピュータ可読媒体を含む非一時的コンピュータ可読媒体であり得る。

10

【0068】

物品識別器および追跡器130は、多数のカメラ110を較正するために、カメラ間較正モジュール132を含み得る。較正モジュール132は、好ましくは、互いに関して多数のカメラ110を較正するように作動可能である。本発明の特に好ましい実施形態では、較正モジュール132は、好ましくは、カメラ110が較正されるフィデューシャル物体の三次元位置とは独立してかつ関係なく、多数のカメラ110を互いに関して繰り返し較正するように作動する。較正モジュール132は、好ましくは、カメラ110の高精度の画素対光線較正を確立するために、物品識別器および追跡器130の他の画像処理機能を採用する前にカメラ110を較正するように作動し、これに基づいて、物品識別器および追跡器130の後続の画像処理機能が実行され得る。較正モジュール132の好ましい作動に関するさらなる詳細は、図2Aおよび2Bを参照して以降で提供される。

20

【0069】

物品識別器および追跡器130は、画像分析モジュール134をさらに含み得、好ましくは、較正されたカメラ110によって取得された建物104の画像の分析のために人工知能を採用する。カメラ110によって取得された画像のすべてまたはいくつかは、追跡器130に提供され得る。

【0070】

画像分析モジュール134は、好ましくは、物品102および／または買い物客106のうちの少なくとも1つの存在および位置を検出し、カメラ110によって取得された画像のうちの少なくともいくつかに出現する物品102のうちの少なくとも1つの物品の予備的識別を提供するように作動する。

30

【0071】

カメラ110によって同時に取得された画像のうちの様々な画像に出現する物品は、画像分析モジュール134が様々な物品が識別された多数の画像を出力するように、画像のうちの異なる画像において画像分析モジュール134によって異なって識別され得ることが理解される。画像分析モジュール134の作動に関するさらなる詳細は、図3を参照して以降で提供される。

【0072】

物品識別器および追跡器130は、画像分析モジュール134によって物品が識別された画像のうちの画像を選択し、他の画像を拒否するための画像選択モジュール136をさらに含み得る。画像選択モジュール136は、好ましくは、較正モジュール132によって実行されるように、好ましくはカメラ110の相互較正を考慮することに基づいて、画像の選択を実行するように作動する。より具体的には、画像選択モジュール136は、対応する位置を占め、したがって共通の物品に対応する物品を識別するために、画像分析モジュール134によって識別され、かつカメラ110を照射する物体から発せられる光線の幾何学的交差、および／またはカメラ110によって取得された深度画像内で識別される空間における点の幾何学的交差を含む幾何学的考察に基づいて画像を選択し得る。したがって、画像選択モジュール136は、フィルタとして効果的に機能し、それらの共通の

40

50

交差点を発見することに基づいて、画像分析モジュール 134 によって提供される画像のうちの選択された画像を選択する。画像選択モジュール 136 の作動に関するさらなる詳細は、図 3 を参照して以降で提供される。

【0073】

物品識別器および追跡器 130 は、建物 104 内の共通の場所に位置付けられている共通の物品の単一の同一性を導出するために、画像選択モジュール 136 によって選択された画像内の物品の様々な可能性のある多様な同一性を融合するための物体分類モジュール 138 をさらに含み得る。物体分類モジュール 138 は、画像分析モジュール 134 によって識別されるように、異なる画像に出現する物品の様々な多様な同一性を比較検討するために人工知能を採用し、単一の最終的な物品分類を導出し得る。物体分類モジュール 138 は、追加的または代替的に、物体識別を再実行し、異なる画像に出現する物品の様々な多様な同一性を融合し、それによって識別されるようにし得る。単一の導出分類は、物品識別器および追跡器 130 によって発見されたような、所与の位置にある所与の物品の同一性であると考えられる。物体分類モジュール 138 の好ましい作動に関するさらなる詳細は、図 4 を参照して以降で提供される。

【0074】

物品識別器および追跡器 130 は、買い物客 106 および／または物品 102 のうちの 1 つと関連する対象のイベントを追跡するためのイベント追跡モジュール 140 をさらに含み得る。本発明の 1 つの好ましい実施形態によれば、イベント追跡モジュール 140 は、買い物客 106 による物品 102 の取り扱いと関連するイベントを自動的に追跡するための買い物リスト作成モジュールとして具体化され得る。買い物リスト作成モジュール 140 は、目録を作成および維持し得、その目録は、物体分類モジュール 138 によって発見されるような、少なくとも 1 つの物品 102 の同一性および位置に基づいて、特定の買い物客 106 と関連付けられた買い物リストであり得る。買い物客 106 と関連付けられていることが発見された物品は、買い物客 106 の買い物リストに属すると考えることができ、したがって、例えば、建物 104 を出る前に、買い物客 106 による買い物リスト上の物品の自動化された目録管理および支払いを容易にする。追加的または代替的に、イベント追跡モジュール 140 は、分析の目的で、建物 104 内の買い物客 106 および／または物品 102 の通過と関連する対象のイベントを追跡し得る。イベント追跡モジュール 140 の好ましい作動に関するさらなる詳細は、図 6～8 を参照して以降で提供される。

【0075】

較正モジュール 132、画像分析モジュール 134、画像選択モジュール 136、物体分類モジュール 138、およびイベント追跡モジュール 140 の間に分散されるものとして説明された様々な機能は、その説明を明確にする目的で細分化され、様々な機能は、本発明の範囲から逸脱せずに、代替的に単一の作動モジュールによって組み合わせて実行されるか、物品識別器および追跡器 130 の様々なサブモジュール間で異なって分散されてもよいことが、理解される。

【0076】

本発明のいくつかの実施形態では、物品識別器および追跡器 130 に含まれるモジュール (132～140) のいくつかまたはすべての機能は、互いに独立して、独立したモジュールとして、または互いにいくつかのみと組み合わせて、または本明細書に記載されていない他の構成要素と組み合わせて提供されてもよいことがさらに理解される。例えば、較正モジュール 132 によって提供され、以降でより詳細に説明されるカメラ較正機能は、物品識別器および追跡器 130 内での使用、およびその中に含まれていると説明される他のすべてのモジュールとの組み合わせでの使用に限定されない。むしろ、較正モジュール 132 によって提供されるカメラ較正機能は、様々なカメラシステムにおけるカメラ較正に有用であり得、システム 100 の他の要素のすべてまたはいずれかを必ずしも伴わずにその中で適用され得る。

【0077】

同様に、画像分析モジュール１３４によって提供され、以降でより詳細に説明される画像分析機能は、物品識別器および追跡器１３０内での使用、およびその中に含まれていると説明される他のすべてのモジュールとの組み合わせでの使用に限定されない。むしろ、画像分析モジュール１３２によって提供される画像分析機能は、様々な画像化システムにおける画像分析に有用であり得、システム１００の他の要素のすべてまたはいずれかを必ずしも伴わずにその中で適用され得る。

【００７８】

同様に、画像選択モジュール１３６によって提供され、以降でより詳細に説明される画像選択機能は、物品識別器および追跡器１３０内での使用、およびそこに含まれていると説明される他のすべてのモジュールとの組み合わせでの使用に限定されない。むしろ、画像選択モジュール１３６によって提供される画像選択機能は、様々なタイプの画像化システムにおける画像選択に有用であり得、システム１００の他の要素のすべてまたはいずれかを必ずしも伴わずにその中で適用され得る。

【００７９】

同様に、分類モジュール１３８によって提供され、以降でより詳細に説明される分類機能は、物品識別器および追跡器１３０内での使用、およびそこに含まれていると説明される他のすべてのモジュールとの組み合わせでの使用に限定されない。むしろ、分類モジュール１３８によって提供される画像分類機能は、様々なタイプの画像化システムの文脈における様々なタイプの画像の分類に有用であり得、システム１００の他の要素のすべてまたはいずれかを必ずしも伴わずにその中で適用され得る。

【００８０】

同様に、イベント追跡モジュール１４０によって提供され、以降でより詳細に説明されるイベント追跡機能は、物品識別器および追跡器１３０内での使用、およびその中に含まれていると説明される他のすべてのモジュールとの組み合わせでの使用に限定されない。むしろ、イベント追跡モジュール１４０によって提供されるイベント追跡機能は、様々な画像化および非画像化システムにおけるイベント追跡に有用であり得、システム１００の他の要素のすべてまたはいずれかを必ずしも伴わずにその中で適用され得る。

【００８１】

次に図２Ａを参照すると、図２Ａは、図１に示されるタイプの物品識別および追跡システムの一部を形成し得る較正サブシステムの構成要素の簡略化された概略図であり、を参照すると、図２Ｂは、その作動のステップを示す簡略化されたフローチャートである。

【００８２】

ここで図２Ａに戻ると、物品識別および追跡システム１００の一部を形成する多数のカメラ１１０は、６台のカメラ１１０ａ～１１０ｆを含むことが見られる。カメラ１１０のそのような数量は、その相互較正の説明の明確さの目的で例示的であり、最小化されることが理解される。実際には、システム１００は、典型的には、建物１０４のサイズおよびその画像化仕様に応じて、１０未満または数百を超えるカメラ１１０、またはその間の任意の数を含み得る。

【００８３】

当業者によって理解されるように、カメラ１１０は、好ましくは、システム１００における物品の識別および追跡の目的で、画像取得のためにそれらを採用する前に較正される。そのような較正は、好ましくは、カメラ較正モジュール１３２によって実行され、好ましくは、カメラ１１０の各々によって取得された画像に物体が出現する画素位置と、画像化された物体から発せられ、かつ対応するカメラセンサ画素に照射する光線との相関を含み、したがって、空間内の画像化された物体の３Ｄ位置に対するセンサ画素位置の相関を可能にする。

【００８４】

本発明の好ましい実施形態によれば、多数のカメラ１１０を較正するために、多数のカメラ１１０は、好ましくは、ここでは例として建物１０４内に位置着けられる基準物体２００として具現化される、少なくとも１つの基準物体を含むシーンを視認するように配置

される。基準物体 200 は、少なくとも複数の多数のカメラ 110 の視線内にあるように位置付けられた、ボールまたはパターン化された物品などの単純な基準物体であり得る。ここで、例として、基準物体 200 は、図 2 A において、第 1 のカメラ 110 a、第 3 のカメラ 110 c、および第 5 のカメラ 110 e によって視認可能であるが、第 2 のカメラ 110 b、第 4 のカメラ 110 d、および第 6 のカメラ 110 f によっては視認可能でないことが見られる。特定のセットアップでは、多数のカメラ 110 a ~ 110 f のすべてが基準物体 200 を視認することが可能であり得るが、建物 104 内のその所与の位置で基準物体 200 が視認可能である複数のカメラは、必ずしも多数のカメラ 110 a ~ 110 f のすべてを含まないことが理解される。

【0085】

ここで第 1、第 3、および第 5 のカメラ 110 a、110 c、および 110 e として具体化される、基準物体 200 が視認可能である複数のカメラの各々は、好ましくは、基準物体 200 の少なくとも 1 つの画像を取得するように作動する。基準物体 200 から取得された画像は、好ましくは、カメラ較正モジュール 132 の一部を形成する画像処理サブモジュール 220 に供給される。画像処理サブモジュール 220 は、好ましくは、基準物体 200 を視認することが可能である複数のカメラによって取得された基準物体 200 の画像を受信するように、かつ複数のカメラの各カメラが照射される、光線の推定される共通の光線交差点を発見するように作動する。このような交差点を発見するために、少なくとも 2 台のカメラが互いに対して較正され得るように、少なくとも 2 台のカメラを照射する光線の共通の交差点を発見することを可能にするように、基準物体 200 が多数のカメラ 110 の少なくとも 2 台のカメラによって視認可能でなければならないことが理解される。光線交差点は、任意の適切な方法によって発見され得、その様々な方法が当技術分野で既知である。例えば、光線交差点は、光線交差点の有効点である、複数の光線までの距離が最小である点を計算するために、中点法などの方法を使用して発見され得る。

【0086】

好ましくは、画像処理サブモジュール 220 によって最初に発見された推定交差点は、3D 空間内の特定の点ではなく、複数のカメラを照射する光線が交差する可能性が高い確率的な交差領域に対応する。確率的な交差領域は、基準物体 200 が位置付けられる可能性が最も高い 3D 空間内の物理的領域に対応すると考えられる。カメラの位置および／または向きは、交差領域の初期推定値の発見を支援するために画像処理サブモジュール 220 に入力され得るが、必ずしも提供されとは限らない。画像処理サブモジュール 220 によって最初に発見された交差点は、少なくとも部分的に、代替の、より精度の低い較正方法に基づいてもよく、この較正方法は、基準物体 200 の位置の初期推定値を提供し得、この初期推定は、その後、本発明の好ましい方法に従って改良されることになる。交差点の初期推定値は、較正パラメータを推定するための *Perspective-n-Point* (PNP) ソリューションなどの当技術分野で既知である較正方法に基づいて発見され得る。

【0087】

好ましくはカメラ較正モジュール 132 の一部も形成する画素対光線較正サブモジュール 240 は、好ましくは、次に、複数のカメラ 110 a、110 c、および 110 e によって取得された各々の少なくとも 1 つの画像内に基準物体 200 が出現する画素位置を、複数のカメラの各々を照射し、かつ画像処理サブモジュール 220 によって発見された交差点と交差する光線に相関させるように作動する。画素対光線較正サブモジュール 240 は、それに関するカメラ 110 a、110 c、および 110 e の較正を可能にするために、入力パラメータとして、建物 104 内の基準物体 200 の 3D 位置を必要としないことが理解される。むしろ、基準物体 200 が視認可能である複数のカメラのうちの個々のカメラ 110 a、110 c、および 110 e は、それらの間の光線の交差に基づいて、互いに対して較正される。したがって、複数のカメラ 110 a、110 c、および 110 e は、シーン内の基準物体の絶対的な三次元位置に関係なく、互いに関して較正される。しかしながら、本発明の特定の実施形態では、カメラ間較正をとリンクするために、カメラ較

正モジュール 132 は、カメラ間較正を外部基準フレームとリンクさせるために、外部絶対基準フレームに対する基準物体 200 の 3D 位置を備えてもよいことに留意されたい。

【0088】

好ましくは、画像処理サブモジュール 220 および画素対光線較正サブモジュール 240 は、光線交差領域が画像処理サブモジュール 220 によって最初に推定され、それと交差して複数のカメラ 110 a、110 c、110 e を照射する光線のパラメータが最初に導出されるように、反復的に作動する。画素対光線較正サブモジュール 240 によって導出された光線パラメータは、その後、好ましくは、画像処理サブモジュール 220 によって光線交差領域をさらに精密化するために画像処理サブモジュール 220 にフィードバックされ、光線パラメータはその後新たに描写された交差領域に基づいて再導出される。そのような反復プロセスは、反復的に導出された光線パラメータを有する光線が、許容可能な精度、例えば 1 cm 以内で定義された交差領域と交差することが発見されるまで繰り返し実行され得る。

10

【0089】

複数のカメラ 110 は、建物 104 への他の修正または追加を必要とせずに、その中のシステム 100 の実装を容易にするために、既存の建物 104 に有利に後付けされ得ることが理解される。したがって、システム 100 は、多種多様な建物に簡単かつ効率的に実装することができる。複数のカメラ 110 は、システム 100 が容易に拡張可能であり、内部に 102 および／または買い物客 106 の密集した配置を有する非常に大きい建物を含む、任意のサイズの建物で実施され得るように、建物 104 のサイズおよびその中の物品 102 および買い物客 106 の分布に応じて、任意の数のカメラを含み得ることがさらに理解される。

20

【0090】

画素対光線較正サブモジュール 240 によって確認された光線パラメータは、外部および／または内部からの光線パラメータを含み得る。固有光線パラメータは、光学歪みパラメータなど、カメラ 110 a、110 c、および 110 e の各々の光学パラメータを含み得る。外部からの光線パラメータは、基準物品 200 およびカメラ 110 のうちの他のカメラに対するカメラ位置、ならびにそれらの相対的な x、y およびシータ座標に関して各カメラセンサ画素を照射する光線の特徴付けるパラメータを含み得る。内部および外部からの光線パラメータの両方は、画素対光線較正サブモジュール 240 によって確認され得る。代替的に、固有光線パラメータは、別のシステムによって測定され、画素対光線較正サブモジュール 240 に供給され得る。外部および内部からの光線パラメータ 240 の両方は、好ましくは、カメラ 110 a、110 c、および 110 e の各々について個別に導出される。

30

【0091】

次に図 2B に戻ると、カメラ較正モジュール 132 の画像処理サブモジュール 220 および画素対光線較正サブモジュール 240 によって実行され得るような、本発明のカメラ間較正に含まれるステップがプロセスフローチャート 250 に示されている。

【0092】

第 1 のステップ 260 に見られるように、物体 200 などの基準物体は、好ましくは、シーン内に位置付けられる。基準物体 200 は、シーンの画像を取得するように配置されたカメラによって視認可能な任意の単純な受動物体であり得、複雑なまたは能動物体である必要はない。

40

【0093】

第 2 のステップ 262 に見られるように、シーンの画像は、好ましくは、次に、N 台のカメラによって取得され、N は、シーンを視認するように配置されたカメラの総数を表す。次に、第 3 のステップ 264 に見られるように、基準物体が出現する画像を取得したカメラのグループが識別される。通常、カメラのグループには、シーンを視認している N 台のカメラすべてが含まれているわけではなく、基準物体 200 は N 台のカメラのサブセットのみによって視認可能である。

50

【0094】

次に、第4のステップ266に見られるように、基準物体が画素位置に出現する画像を取得したカメラのグループを照射する光線の交差領域が推定される。第5のステップ268に見られるように、光線パラメータは、次に、好ましくは、カメラのグループのカメラの各々について導出され、その光線パラメータは、第4のステップ264で発見された交差領域と交差し、かつカメラのグループの各カメラによってその取得された画像に基準物体が出現する画素位置に対応する光線を定義する。このような光線パラメータには、内部および内部からのパラメータが含まれ得る。

【0095】

次に、問い合わせ270に見られるように、プロセス250は、第5のステップ268で定義された光線が、許容可能な精度で第4のステップ266で定義された交差領域に接近して交差するかどうかを問い合わせする。

【0096】

交差領域および光線パラメータが十分なレベルの精度で確認されたことが発見された場合、この交差領域は基準物体の位置に対応していると考えられ、画素対光線較正は、その位置で基準物体を画像化したこれらのカメラについて導出される。

【0097】

第5のステップ268で導出された光線パラメータを有する光線が十分な水準の精度で交点と交差することが発見されない場合、第6のステップ272に見られるように、第5のステップ268で導出された光線パラメータに基づいて交差領域が精緻化され、第5のステップ268および第6のステップ272が反復的に繰り返され得、それによって、問い合わせ270で確認されるように、光線パラメータが十分な精度で交差点と交差することが発見されるまで、光線パラメータは、最も最近定義された交差点に基づいて再導出され、交差点は、次に、新たに導出された光線パラメータに基づいて再定義される。

【0098】

問い合わせ270で確認されたように、十分な精度で画素対光線較正の実行に続いて、プロセス250は、好ましくは、追加の問い合わせ274に見られるように、シーンを視認するすべてのN台のカメラについて画素対光線較正が導出されたかどうかを確認する。N台のカメラのすべてが較正されていない場合では、第7のステップ276に見られるように、異なるグループのカメラによって視認可能であるように、基準物体は、好ましくは、シーン内に再位置付けされる。次に、ステップ262～274を、N台のカメラすべてについて画素対光線較正が実行され、較正が完了したと考えられるまで繰り返され得る。

【0099】

基準物体がその新しい位置で視認可能であるカメラのグループは、基準物体が以前に視認可能であったカメラのグループと重複してもよく、全N台のカメラのうちのカメラが複数回較正されるが、毎回異なる基準物体位置およびカメラの異なるグループに関して較正されることが理解される。

【0100】

図2Aおよび2Bを参照して本明細書に記載されたカメラ間較正の特定の方法は、本発明の物品識別および追跡システムの文脈内での実装形態に限定されず、むしろ、共通のシーンを視認する少なくとも2台のカメラの迅速かつ正確な較正から利益を得る他の様々な用途におけるカメラ間較正に有用であり得ることが理解される。

【0101】

次に図3を参照すると、図3は、画像分析および分類機能を実行するシステム構成要素の簡略化されたブロック図であり、図1に示されるタイプの物品識別および追跡システムの一部を形成し得る。

【0102】

図3に見られるように、物品識別および追跡システム100の多数のカメラ110は、ここでは、カメラ1からカメラNまでを含むように示されている。システム100は、建物104（図1）のサイズに応じて、100を超えるカメラなどの多数のカメラ、または

10

20

30

40

50

少数のカメラを含み得る。

【0103】

図1を参照して上述されたように、カメラ110は、好ましくは、消費者物品102および個人106（図1）などの複数の物体を含むシーンを視認するように、かつそれらの画像を取得するように配置される。カメラ1～Nは、好ましくは、少なくとも部分的に相互に同時にシーンの複数の画像を取得するように作動する。ここで、例として、カメラ1～Nは、画像1～Nを取得するように作動することが示されている。ただし、カメラ1～Nの各々によって複数の画像が取得され得るが、複数のカメラ1～Nは、画像を取得するために必然的に作動することが理解される。カメラ110は、従来の2Dカメラであり得るか、またはステレオカメラ、飛行時間型カメラ、または構造化光カメラなどの深度センサを含み、深度情報を含む3D画像を提供するカメラであり得る。

10

【0104】

図3に示されるようなカメラ1～Nによる画像の取得は、所与の時間枠におけるカメラ1～Nの同時作動を表すことが理解される。システム100の作動において、カメラ1～Nは、好ましくは、そのような画像を1秒間に複数回取得するように作動し、この画像は、好ましくは、図3を参照して本明細書で以下に概説されるように処理される。

【0105】

カメラ1～Nは、好ましくは、それによって画像1～Nを取得する前に較正されることが理解される。カメラ1～Nは、図2A～2Bを参照して本明細書で上述されたように相互較正され得る。

20

【0106】

画像1～Nは、好ましくは、画像分析モジュールに提供され、好ましくは、物品識別および追跡器130の画像分析モジュール134として具体化される（図1）。

【0107】

本発明の1つの好ましい実施形態によれば、カメラ110によって取得された画像1～Nは、好ましくは、画像分析モジュール134へのその提供の前に、前処理される。好ましくは、前処理は、複数の画像1～Nの少なくとも1つの特性を第1の特性から第2の特性に変換することを含む。本発明の1つの好ましい実施形態によれば、画像1～Nは、画像分析モジュール134によるその分析の前に、解像度が低減される。ここで、例として、第1の解像度を有する画像1～Nは、第2の解像度を有する対応する低解像度画像1～Nに変換され、第2の解像度は第1の解像度よりも低いと見られる。例えば、画像解像度の低下を含むそのような前処理は、画像分析モジュール134自体によって、またはシステム100の一部を形成する画像変換モジュールなどの別の構成要素によって実行され得ることが理解される。

30

【0108】

画像1～Nの前処理およびその少なくとも1つの特性の変換は、好ましくは、画像分析モジュール134によって画像1～Nを処理するために必要とされる計算電力を低下するために実行される。ここで、例として、画像の解像度の低下1～Nは、画像を処理するために必要な計算電力を低下させる。しかしながら、画像を処理するために必要とされる計算電力を低下させるために、解像度以外の画像1～Nの特性が代替的に変換され得ることが理解される。ほんの一例として、画像を圧縮するか、または画像内の対象の区域（例えば、動きが検出された区域）に焦点を当てるために追加情報を使用することができる。しかしながら、元の、例えば、カメラ1～Nによって取得されたより高い解像度の画像は、好ましくは、システム100に保持および記憶され、任意選択で、以下でさらに詳述されるように、画像分類モジュール138によって処理される。

40

【0109】

好ましくは、システム100は、本明細書でさらに後述されるように、解像度の低下などの第2の特性を有する複数の画像1～Nのうちの少なくともいくつかに出現する少なくとも1つの物体の少なくとも1つの特徴を自動的に発見し、発見された少なくとも1つの特徴に基づいて、元のより高い解像度などの第1の特性を有する画像を修正するように作

50

動する。

【0110】

画像分析モジュール134は、好ましくは、複数の画像1～N、特に好ましくは複数の低解像度画像1～Nを受信し、カメラ1～Nによって画像化され、複数の画像1～Nのうちの少なくともいくつかに出現する複数の物体のうちの少なくとも1つの物体の存在を検出し、その同一性を発見するように作動する。画像化された対象の物体は、物品102（図1）などの1つ以上の消費者物品、買い物客106（図1）などの1人以上の個人、または両方でもあり得ることは理解される。ここで、例として、画像分析モジュール134は、物体が物品102などの物品である場合、好ましくは、物体同一性1に対応する画像1～N内の物体同一性Nを識別する。対象の物体が人である場合の画像分析モジュール134の作動に関するさらなる詳細は、以降で提供される。

10

【0111】

物体同一性1～Nは、多様な同一性である場合もあれば、少なくともいくつかの共通の物体同一性を含む場合があることも理解される。物体同一性1～Nは、異なる物体を示す画像1～Nに起因して多様な同一性である可能性があり、その異なる物体は、画像分析モジュール134によって対応して異なるように識別される。追加的または代替的に、物体同一性1～Nは、少なくとも部分的に同じ物体を示す画像1～Nのうちの少なくともいくつかにもかかわらず、画像分析モジュール134によって画像化された物体が異なって識別されるため、多様な同一性であり得る。これは、同じ物体が、カメラ1～Nによって異なる画像化角度で、異なる視点から、および異なる画像化パラメータで、画像1～Nのうちの少なくともいくつかにおいて画像化されることに起因し、画像1～Nのうちの2つ以上における同じ物体が画像解析モジュール134によって異なる同一性を割り当てられることになる場合がある。画像1～Nは複数の物体を示し得ること、および画像分析モジュール134は、画像1～Nの各々の中の複数の物体同一性を識別するために作動し得ることが理解される。

20

【0112】

本発明の特に好ましい実施形態によれば、画像分析モジュール134は、好ましくは、物体の存在を検出し、その同一性、例えば、低解像度画像1～Nにおける物体同一性1～Nなどの物体の同一性を発見するように作動する。好ましくは、必ずしもそうではないが、画像分析モジュール134は、人工知能を採用する。画像の人工知能ベースの分析は、2Dまたは3D画像に対して実行され得ることが理解される。好ましくは、画像分析モジュール134は、内部の物体の位置を示すバウンディングボックスをその中に挿入することによって、画像1～Nに注釈付けするように作動する。さらに好ましくは、画像分析モジュール134は、それによって識別された各物体同一性に信頼水準を割り当てるように作動し、その信頼水準は、好ましくは、物体の存在が検出された信頼度と、その検出された物体の同一性が分析モジュール134によって識別された信頼度との両方を表現している。

30

【0113】

このような検出および識別の人工知能ベースのプロセスは、非常に計算量が多く、通常、ローカルサーバ120およびクラウド122（図1）の一方または両方に位置付けられたコンピューティングデバイスによって実行するのにかなりの時間がかかることが理解される。したがって、画像分析モジュール134によるそのような画像分析の実行前に、カメラ110によって取得された高解像度画像をより低い解像度の画像に変換することは、物体の検出および識別に必要な計算電力および処理時間がそれによって大幅に削減されるため、非常に有利である。しかしながら、場合によっては、そのような解像度の低下は必要ない場合があり、高解像度画像1～Nは、その解像度のいかなる事前の低下も伴わずに、画像分析モジュール134に直接提供され、画像分析モジュール134によって処理され得る。

40

【0114】

典型的には、本発明の画像における物体識別の目的で人工知能を採用する場合、画像分

50

析モジュール 134 は、画像内の物体を自動的に識別することができるように最初に訓練される。そのような訓練は、画像分析モジュール 134 に含まれ得る AI ネットワークへの多数の訓練画像の提供を含み得、多数の訓練画像の各々に出現する少なくとも 1 つの物体は、画像分析モジュール 134 の AI ネットワークについて識別される。訓練画像に出現し、画像分析モジュール 134 に提供される物体の識別は、人間の識別を含み得る。より好ましくは、訓練画像に出現し、画像分析モジュール 134 に提供される物体の識別は、マシンビジョンに基づく機械識別を含み得る。そのようなアルゴリズムを訓練するためのデータの生成に関するさらなる詳細は、図 9A ～図 18 を参照して以降で提供される。

【0115】

訓練に続いて、画像分析モジュール 134 は、好ましくは、人工知能に基づいて、そこに提供される画像内の物体を自動的に識別するように作動する。

10

【0116】

任意選択で、物体識別画像 1 ～ N は、画像分析モジュール 134 の下流および画像選択モジュール 136 の上流の画像フィルタリングモジュール 300 に供給され得る。画像フィルタリングモジュール 300 は、好ましくは、内部で少なくとも 1 つの物体が識別された複数の画像を画像分析モジュール 134 から受信するように作動し、ここでは、画像分析モジュール 134 は、物体識別画像 1 ～ N として具体化される。画像フィルタリングモジュール 300 は、好ましくは、複数の画像のうちの 1 つ以上の画像を除外するように作動する。

【0117】

画像フィルタリングモジュール 300 によって実行されるフィルタリングは、物体同一性の定義されたグループに属していないものとして画像内で識別された少なくとも 1 つの物体の同一性に基づき得る。内部で識別された物体の同一性が物体同一性の定義されたグループに属さない少なくとも 1 つの画像は、好ましくは、画像フィルタリングモジュール 300 によって画像選択モジュール 136 に提供されない。

20

【0118】

ここで、例として、画像フィルタリングモジュール 300 は、物体識別画像 1、2、3、4、および N 受信することが示されている。例えば、物体同一性 1、3、4、および N は、物体同一性の定義されたグループのメンバーであり得、一方、物体同一性 2 は、物体同一性 1、3、4、および N と非常に異質であることに起因して物体同一性のその定義されたグループ外であり得る。この場合、システム 100 の下流構成要素に物体識別画像 1、3、4、および N のみ提供され、物体識別画像 2 は提供されない。このようにして、シーンから同時に取得された他の画像で識別された物体と明確かつ明白に異なる物体が内部で識別された画像は拒否される。

30

【0119】

画像フィルタリングモジュール 300 によって適用される物体同一性の定義されたグループは、クラスタ化された類似の物体同一性の所定のグループであり得るか、またはシステム 100 の作動の過程の間に定義された動的グループであり得る。追加的または代替的に、定義されたグループは、システム 100 またはそれに類似の他のシステムの過去の作動に基づいて、類似の物体同一性の過去に学習されたカテゴリに基づいて定義され得る。

40

【0120】

追加的または代替的に、画像フィルタリングモジュール 300 は、物品が画像分析モジュール 134 によって検出および識別される信頼度に閾値信頼水準を適用することに基づいて、画像分析モジュール 134 によってそこに提供される複数の画像のうちの 1 つ以上を除外するように作動し得る。画像フィルタリングモジュール 300 のこの作動モードでは、事前定義された信頼水準未満の信頼度で検出および識別された物品は、フィルタリングモジュール 300 によって除外され、システム 100 の下流構成要素に提供されない。

【0121】

任意選択で画像フィルタリングモジュール 300 によってフィルタリングされた物体識別画像 1 ～ N は、好ましくは、画像相関モジュール 302 に供給される。画像相関モジュ

50

ール 302 は、好ましくは、そこに供給された画像の各々において識別された少なくとも 1 つの物体が出現する複数の画素位置を、複数のカメラを照射する対応する複数の光線に相関させ、複数の光線のうちの少なくともいくつかの交差点を発見するように作動する。光線交差点は、任意の適切な方法によって発見され得、その様々な方法が当技術分野で知られている。例えば、光線交差点は、光線交差点の有効点である、複数の光線までの距離が最小である点を計算するために、中点法などの方法を使用して発見され得る。

【0122】

ここで、例として、画像相関モジュール 302 は、好ましくは、対応する画素位置で内部で識別された物体を有する画像 1、3、4、および N を受信し、物体がカメラ 1、3、4、および N を照射する光線に対して出現する画素位置を関連付けるように作動し、画像相関モジュール 302 は、好ましくは、光線の少なくともいくつかの交差点を発見する。

10

【0123】

物体が画像内に出現する画素位置と、画像が取得されるカメラを照射する光線のパラメータとのそのような相関は、好ましくは、カメラ校正モジュール 132 によって最初に行われた画素対光線校正に基づき、この画素対光線校正は事前に多数のカメラ 110 の各カメラについての画素対光線の相関を確立したことが理解される。

【0124】

好ましくは、画像選択モジュール 136 (図 1) として具現化される画像選択モジュールは、好ましくは次いで、それらの画像のみを選択するように次に作動し、それらの画像において、画像相関モジュール 302 によって発見された交差点と交差する光線に対応する画素位置に少なくとも 1 つの識別された物体が出現する。複数の光線のうちのいくつかの交差点は、空間内の共通の個々の物体の 3D 位置に対応すると考えられ、その結果、交差点と交差する光線に対応する画素位置に物体を示す画像は、その光線交差点に位置付けられているその共通の物体を示す画像であると考えられ得ることが理解される。これは、カメラを照射する様々な光線が実際に空間内の同じ画像化された物理的点から発せられない限り、カメラセンサ画素を照射し、3D 空間で交差する様々な光線の確率が無視できるという理解に基づいていることが理解される。

20

【0125】

したがって、画像選択モジュール 136 は、好ましくは、共通の物体を示す複数の画像のうちの画像を選択し、共通の物体を示さない複数の画像のうちの他の画像を拒否するように作動する。画像選択モジュール 136 によって実行される選択および拒否は、好ましくは、複数のカメラの事前の相互校正に基づき、この事前の相互校正は、カメラ 110 によって取得された画像の画素位置とカメラ 110 を照射する光線のパラメータとの間の相関を確立するのに役立つ。したがって、画像選択モジュール 136 は、好ましくは、幾何学的選択モジュールとして機能し、3D 空間におけるカメラ光線の幾何学的交差に基づいて画像を選択するように作動することが理解される。

30

【0126】

画像選択モジュール 136 は、好ましくは、光線交差点に対応する位置にある共通の物体を示すものとして識別された選択された画像のセットを出力するように作動する。発見された交差点に対応する 3D 空間内の位置は、物体の 3D 位置に対応すると考えられる。本発明の 1 つの好ましい実施形態では、物体は、図 6 ~ 図 8 を参照して以降でさらに詳述されるように、光線交差点に位置付けられている 3D 空間内の点としてモデル化され得る。

40

【0127】

ここで、例として、画像選択モジュール 136 は、物体識別画像 1、3、および 4 を含む選択された画像のセット 304 を出力するように示されている。選択された画像のセット 304 は、物体識別画像 N を含まず、この物体識別画像 N は、画像相関モジュール 302 によって発見された光線交差点と交差しない照射光線に対応する画素位置に物品を示すものとして、画像選択モジュール 136 によって拒否された。

【0128】

50

画像選択モジュール136によって実行される選択および拒否は、好ましくは、事前のカメラ間較正に基づき、より具体的には、好ましくは、事前のカメラ間較正によって発見される光線交差に基づく幾何学的フィルタリングに基づき、したがって、複数の画像内の少なくとも1つの物体の同一性に関係なく実行され得ることが理解される。画像選択モジュール136は、画像分析モジュール134によって発見されるような物体の同一性ではなく、その中の物体の位置に基づいて画像を選択および拒否するように作動し得る。

【0129】

代替的に、画像選択モジュール136は、それによって実行される幾何学的画像選択を増強するために、画像分析モジュール134によって発見されるような物体の同一性を考慮するように作動し得る。特定の実施形態では、画像は、画像選択モジュール136でク
10
リークにグループ化され得、それによって、各クリークは、内部で識別された同様の物体を有する画像から構成される。次に、画像選択モジュール136は、物品の位置を比較し、すべての画像ではなく、各クリーク内のそれらの画像についてのみ光線交差点に位置付けられている物体を示す画像を識別するように作動し得る。そのようなアプローチは、画像選択モジュール136によって実行される幾何学的画像選択を促進し得る。

【0130】

光線交差に基づく画像関連モジュール302および画像選択モジュール134の上述された作動は、カメラ110によって取得された2D画像の処理に特によく適していることが理解される。しかしながら、カメラ110が、物体識別画像1~Nが深度情報を含むよ
20
うに深度カメラとして具体化される場合、画像関連モジュール302および画像選択モジュール134は、代替の方法で作動し得る。

【0131】

本発明のこの実施形態では、画像関連モジュール302は、そこに入力された物体識別画像1~Nのうちの任意の所与の1つで識別される物品の3D位置を、物体識別画像1~Nに含まれる深度情報に基づいて発見するように作動し得る。次に、画像選択モジュール134は、所与の画像内の3D物品の位置が、物体識別画像1~Nの他のいずれかの物品の3D物品の位置と一致するかどうかを確認し得る。物品の位置が合致すると発見された場合、物品は空間における共通の位置にあり、したがって同じ物理物品に対応すると考え
30
られる。

【0132】

画像選択モジュール136によって選択された画像は、2D画像の場合は光線交差、または3D深度画像の場合は合致する3D位置に基づいて、したがって、同じ物理的位置にある同じ物品を示すと考えられる画像のセットを形成すると考えることができる。したがって、画像選択モジュール136は、幾何学的フィルタリングモジュールとして効果的に機能し、同じ位置に同じ物品を表示することが発見されない画像を拒否することが理解される。本発明のシステムの作動のこの段階では、物品の特定の最終的な同一性はまだ確認
40
されていないことが理解される。

【0133】

次に、画像選択モジュール136によって選択された画像は、好ましくは、物品分類モジュールに提供され、好ましくは、物体分類モジュール138として具体化される（図1）。分類モジュール138は、好ましくは、画像選択モジュール136によって出力された選択された画像のセットを受信し、選択された画像のセットを示す少なくとも1つの物体の同一性を発見するように作動し、この物体は、好ましくは、3D空間の所与の領域を占める共通の物体である。それ自体によって受信された画像内の物体同一性を考慮しなくてもよい画像選択モジュール136とは対照的に、分類モジュール138は、好ましくは、選択された画像のセット内の物体同一性を考慮することによって物体同一性を導出するように作動する。複数のカメラによって取得された画像に基づいて物体同一性を確立することは、単一のカメラによって取得された画像に基づいて物体同一性を確立することとは対照的に、誤差を減少させ、物品識別の精度を改善するのに有利に役立つことが理解される。
50

【0134】

本明細書で上述されたように、選択された画像が、空間内の共通点を占める共通物体をすべて示すことに基づいて選択されたにもかかわらず、画像分析モジュール134は、その画像のうちの異なる画像における共通物体の外観の変動に起因して、共通物体に多様な物体同一性を割り当てている可能性がある。分類モジュール138は、好ましくは、選択された画像のセット内の様々な多様な物体同一性を融合し、単一の物体同一性306を導出するように作動し、この単一の物体同一性は、共通の物体の最終的な確認された同一性であると考えられる。分類モジュール138は、好ましくは、異なる物体同一性の組み合わせおよび重み付けに基づいて、様々な多様な物体同一性を融合するように作動可能である。

10

【0135】

本発明の1つの好ましい実施形態では、分類モジュール138は、画像分析モジュール134によって発見されるような、選択された画像のセット内の様々な物体同一性を受信するように、かつ単一の最終的な物体同一性を導出するために、画像分析モジュール134によって各物体同一性に割り当てられた信頼水準に少なくとも部分的に基づいて、様々な同一性を融合するように作動し得る。このような同一性の融合は、機械学習に基づき得る。画像分類モジュール138によって導出された単一の最終的な物体同一性はまた、それに関連する信頼水準を有し得る。画像分析モジュール134によって分析される画像の解像度が低下することに起因して、その中で識別される様々な物体同一性と関連付けられた信頼度、したがって、それに基づく最終的な融合された物体同一性と関連付けられた信頼度が許容できないほど低くなる可能性があることが理解される。

20

【0136】

そのような場合、本発明の追加の好ましい実施形態によれば、分類モジュール138は、物体の識別を再実行するように作動し得る。画像分析モジュール134によって分析され、選択された画像のセット304を形成する低解像度画像に対応する初期の高解像度画像が検索され、分類モジュール138に提供され得る。分類モジュール138は、例えば、画像分析モジュール134によって識別される物品位置の領域で高解像度画像を切り取ることによって、高解像度画像を修正するように作動し得る。次に、分類モジュール138は、切り取られた高解像度画像内の物品の同一性を再発見するように作動し得る。

【0137】

分類モジュール138によって処理される画像のより高い解像度に起因して、それによって発見される物品同一性の精度は、典型的には、画像分析モジュール134によって対応する低解像度画像に発見される物品同一性の精度よりも高いことが理解される。さらに、切り取られる高解像度画像に起因して、分類モジュール138は、より限定された画像領域内で物品の同一性を分類するように作動し、したがって、必要な計算電力および処理時間を低減させる。結果として、高解像度画像の画像分析に基づく分類モジュール138によって出力される最終物体分類と関連付けられた信頼性は、より低い解像度の画像に基づく画像分析モジュール134によって提供される物品同一性に基づく信頼性よりも大きくなり得る。したがって、画像分析モジュール134によって分析される画像は、その中に出現する物品の同一性をより正確に発見することを可能にするために、より高い解像度の画像の修正、例えば切り取りを指示するのに役立つ。

30

40

【0138】

図3に見られるように、例として、その中で識別された対応する物体1、3、および4を有する画像1、3、および4を含む低解像度画像304の選択されたセットは、画像分類モジュール138に直接提供され得る。次に、分類モジュール138は、対応する高解像度画像1、3、および4を検索し、高解像度画像1、3、および4を切り取り、物品同一性306を導出するために、切り取られた高解像度画像に対して物品識別および分類を再実行し得る。この実施形態では、分類モジュール138は、より低い解像度の画像に基づいて、元のより高い解像度の画像を修正するための画像修正器として機能し得ることが理解される。

50

【0139】

分類モジュール138が元の高解像度画像上で物品識別を再実行する場合、最終的な物品同一性306の導出は、画像分析モジュール134によって処理されるような低解像度画像に基づいて導出された物品同一性を考慮する場合と考慮しない場合があることが理解される。

【0140】

解像度以外の元の画像1～Nの特性が変換された場合、分類モジュール138は、より高い精度で内部に出現する物品の同一性を発見することを可能にするために、オリジナルの非修正画像を検索し、元の画像の分析を指示するために、より少ない処理電力を必要とする変換された画像を使用するように、本明細書で説明されるものと同様の方法で作動し得ることが理解される。

10

【0141】

分類モジュール138の作動、特に単一の最終的な物体同一性を導出するために分類モジュール138が多様な物体同一性のバランスをとるために作動する方法に関するさらなる詳細が、図4を参照して以降で提供される。

【0142】

図3に示されるシステム100の構成要素に関して上述された画像処理機能は、建物104（図1）内の物品102などの物品の位置および同一性の識別に関して説明されることが理解される。しかしながら、場合によっては、カメラ110による買い物客106の画像化に基づいて、建物104（図1）内の買い物客106などの個人の位置を追加的または代替的に発見することが望ましい場合があることが理解される。

20

【0143】

買い物客106（図1）などの画像化された個人の存在および位置を発見する場合、買い物客106の同一性の発見は必ずしも関連性がないことが理解される。したがって、この場合、人工知能画像分析モジュール134は、好ましくは、買い物客106の存在および位置を発見するためにのみ作動する。好ましくは、これは、好ましくは、AIアルゴリズムの採用に基づいて、買い物客106の存在および位置を発見する画像分析モジュール134によって達成される。好ましくは、画像分析モジュール134は、買い物客106の身体を表す姿勢グラフを作成する。そのような姿勢グラフは、買い物客106の身体の主要位置に対応するいくつかの主要点を含み得る。これらの主要点は、買い物客106の身体上のそれらの位置を示すために好ましくはラベル付けされている。内部の個人の存在および位置を識別するための画像のAI処理に関するさらなる詳細は、図15A～17Dを参照して以降で提供される。

30

【0144】

内部で姿勢グラフによって識別された買い物客106などの個人を有する画像は、画像相関モジュール302に供給され得ることがさらに理解される。画像相関モジュール302は、好ましくは、そこに供給された画像の各々における姿勢グラフ内の主要点が出現する複数の画素位置を、複数のカメラを照射する対応する複数の光線に相関させ、複数の光線のうちの少なくともいくつかの交差点を発見するように作動する。したがって、共通の光線交差点に対応する主要点は、買い物客106上の同じ身体位置に対応すると考えられる。

40

【0145】

画像における解像度の低下およびその後の元のより高い解像度の画像の検索は、人間以外の物体の画像と比較して、そのような画像を迅速かつ正確に処理するAIアルゴリズムの能力が向上していることに起因して、人間の被写体の画像の処理に関して必ずしも必要ではないことが理解される。

【0146】

次に図4を参照すると、図4は、図3に示されるタイプの画像分析および分類構成要素によって実行される機械学習を示す簡略化されたブロック図である。

【0147】

50

図4に見られるように、内に識別された少なくとも1つの物体を有する入力画像のセット400は、好ましくは、画像フィルタリングモジュール300、画像相関モジュール302、および画像選択モジュール136（図3）を介して、例えば、画像分析モジュール134によって分類モジュール138に提供される。例えば、入力画像のセット400は、図3の選択された画像304に対応し得る。入力画像のセット400において識別される少なくとも1つの物体は、好ましくは、画像分析モジュール134によって採用されるように、人工知能に基づいて識別される。内部に入力された画像内の物体を識別するために画像分析モジュール134によって採用される人工知能は、深層学習、ニューラルネットワーク、または物体同一性への任意の他のタイプの人工知能アプローチを含み得る。

【0148】

画像分析モジュール134は、好ましくは、シーンを視認するカメラ110（図1および3）などの複数のカメラによって取得された複数の画像を受信し、図3を参照して本明細書で上述されたように、複数の画像に出現する対応する複数の物体の複数の同一性を自動的に発見するために人工知能を採用する。複数のカメラによって取得された複数の画像の解像度は、必要な計算電力を低減し、画像処理を加速するために、画像分析モジュール134によるさらなる処理の前に低下され得る。

【0149】

画像選択モジュール136は、好ましくは、さらに図3を参照して本明細書で上述されたように、カメラによって画像化され、選択された画像の各画像に出現する複数の物体の一部を形成する物体の位置に関して、複数のカメラの較正に基づいて複数の画像のうちの選択された画像を選択する。

【0150】

画像分類モジュール138は、好ましくは、複数の画像のうちの選択された画像を受信し、選択された画像に出現する物体の同一性を導出するために機械学習を採用する。

【0151】

分類モジュール138は、好ましくは、機械学習を採用して発見された異なる物体同一性を比較検討および組み合わせることに基づいて単一の最終的な物体同一性402を導出するように作動し、その異なる物体識別は、好ましくは、人工知能を採用することによって発見される。したがって、本発明のシステムは、好ましくは、「二重」機械学習を採用し、単一の究極の物体同一性を導出するために、機械学習を採用することに基づいて、人工知能を採用することに基づいて最初に識別された多様な複数の物体同一性が融合される。

【0152】

本発明の1つの可能な実施形態では、分類モジュール138は、所与の同一性を有する共通の物体を示すと識別された画像の相対数を考慮することに基づいて、多様な物体同一性を重み付けする。例えば、第1の同一性が入力画像400の80%において共通の物品に割り当てられ、第2の異なる同一性が入力画像400の20%において同じ物品に割り当てられた場合、分類モジュール138は、好ましくは、第1の同一性に大きい重みを与えるように作動する。

【0153】

本発明の1つの好ましい実施形態では、図3を参照して本明細書で上述されたように、分類モジュール138は、画像分析モジュール134によって発見された、複数の画像のうちの選択された画像における対応する複数の物体の複数の同一性を融合することに基づいて、少なくとも複数の画像のうちの選択された画像に出現する物体の同一性402を導出するように作動する。

【0154】

図3を参照して本明細書で上述されたように、本発明の特定の実施形態では、画像分析モジュール134は、解像度の低下した入力画像に基づいて物体の検出および識別を実行するように作動し、必要な計算電力および処理時間を最小化するために画像の解像度を低下させることが望ましい。ただし、入力画像の解像度が低いことに起因して、入力画像に

10

20

30

40

50

基づいて発見された物体同一性と関連付けられた信頼度は比較的低くなる可能性がある。したがって、低解像度の入力画像を使用して画像分析モジュール134によって発見された物体同一性に基づいて導出された物体同一性402に関連付けられた信頼度は、許容できないほど低くなる可能性がある。

【0155】

物体同一性402と関連付けられた信頼度を高めるために、分類モジュール138は、図3を参照して本明細書でも上述されたように、任意選択で、高解像度画像を切り取り、続いて、切り取られた高解像度画像上で物体検出および識別を再実行するように作動するサブモジュール403を含み得る。

【0156】

サブモジュール403は、カメラ110によって捕捉された、第1のより高い解像度を有する複数の初期画像のうちの少なくともいくつかを検索するように作動し得る。次に、サブモジュール403は、第2のより低い解像度の画像の対応するそれぞれの画像に出現する識別された物体の位置に対応する領域内のより高い解像度の画像の各々を切り取り得る。したがって、例えば、画像分析モジュール134（図3）によって実行されるようなより低い解像度の画像における物体識別は、例えばサブモジュール403によって実行されるような元の対応するより高い解像度の画像の切り取りを指示するのに役立つ。

【0157】

次に、サブモジュール403は、人工知能を採用して、切り取られた第1のより高い解像度の画像内の物体の同一性を自動的に発見し、少なくとも、選択された切り取られた複数のより高い解像度の画像内の対応する複数の物体の複数の同一性を融合することに基づいて、より高い解像度の画像内に出現する物体の同一性を導出するために人工知能を採用し得る。より高い解像度の画像が切り取られたことに起因して、画像の検出および識別はより小さな画像領域に限定され、したがって必要な計算時間および電力が低減される。それによって最終的な物体同一性402は、切り取られたより高い解像度の画像に基づいて、より高い信頼水準で導出され得る。対応するより低い解像度の画像に基づいて、より低い関連する信頼水準で導出されたであろう物体同一性402は、より高い解像度の画像に基づいて物体同一性402を導出する際に考慮される場合と考慮されない場合があることが理解される。より低い解像度の画像に基づく物体同一性402が十分な水準の信頼度で発見された場合、サブモジュール403による追加の処理は不必要であり得ることもまた

【0158】

入力された多様な物体同一性を融合するために分類モジュール138によって採用される機械学習は、物体識別が、本明細書で上述されたようにより高い解像度および／またはより低い解像度の画像に基づき得、様々な入力によって増強され得る。例として、単一の最終的な物体同一性を導出するために多様な物体同一性を融合する際に、分類モジュール138は、過去に発見された物体同一性を1つの入力係数404として考慮し得る。この場合、集約された買い物客の統計は、以前に所与の物品を選択した買い物客によって通常選択される物品の範囲を示し得る。買い物リスト作成モジュール140（図1）によって好ましくは生成されるような買い物客106によってすでに選択された物品同一性に基づいて、分類モジュール138は、買い物客106によって以前に選択された他の物品に基づいて、所与の買い物客106によって選択されると典型的に予想される物品に対応する物品同一性により大きい重みを与え得る。そのような集約された買い物客の統計は、好ましくは、集約された匿名の統計であり、所与の買い物客106に固有ではないことが理解される。

【0159】

追加的または代替的に、単一の最終的な物体同一性を導出するために多様な物体同一性を融合する際に、分類モジュール138は、追加の入力係数406として、所与の物体が画像に出現する頻度に関連する集約された履歴データを考慮し得る。例えば、入力画像400に出現数共通の物品が、画像分析モジュール134によって相互に等しい重みを有す

10

20

30

40

50

る2つの可能な同一性を割り当てられた場合、集約された履歴データに基づいて、買い物客によってより頻繁に典型的に選択される物品に対応する物体同一性は、分類モジュール138によってより大きい重みが与えられ得る。

【0160】

さらに追加的または代替的に、単一の最終的な物体同一性を導出するために多様な物体同一性を融合する際に、分類モジュール138は、追加の入力係数408として、シーン内の少なくとも1つの物体の所定のモデルに基づいてシーン内の物体の位置を考慮し得る。例えば、建物104内の物品の初期位置は、分類モジュール138に提供され得、および、建物プラングラムに基づいて、様々な物体同一性の中からそれに応じて物品の同一性が導出され得る。

10

【0161】

さらに、体積または重量センサ410は、物体識別を支援するデータを提供するために、建物104内に位置付けられてもよいが、特に、買い物客106の手の注文サイズ以下の小さな物品の場合では、これらの物品は画像内での同一性に対して困難であり得る。体積または重量センサ410は、物品102が載っている表面の初期重量、および物品がそこから取り出した後の表面の後続の重量を記録し得る。重量の違いは、取り出された物品の重量を示し得、したがって、その識別に役立つ。そのような体積または重量データは、多様な物体同一性の融合を支援するために分類モジュール138に提供され得る。そのような体積または重量データは、追加的または代替的に画像分析モジュール134に提供され、それによって実行される予備的な物体識別を支援するようにし得る。重量センサが表面から取り出された物品の重量を正確に記録するために、物品の重心が重量センサの上になければならないことが理解される。小さな物品の場合、これは、小さな物品を容器内に位置決めすることによって達成することができ、その容器は、重量センサに対して中央に位置付けることができる。

20

【0162】

追加的または代替的に、しるし読み取りモジュール412は、本発明のシステムに組み込まれ得る。しるし読み取りモジュール412は、別個のモジュールであり得るか、またはカメラ110の機能に含まれ得る。しるし読み取りモジュール412は、好ましくは、物体の最終的な同一性を導出するために、画像化された物体の一部を形成するしるしの少なくとも一部分を少なくとも部分的に読み取るように作動する。

30

【0163】

例として、画像化された物体に含まれるしるしは、物体に印刷されたテキストまたは凡例であり得る。テキストの少なくとも一部を読み取ることは、分類モジュール138による物体同一性の分類を支援し得る。例えば、物体上の特定の文字の出現は、分類モジュール138によって使用されて、画像分析モジュール134によって提供される可能な物品同一性の範囲から物品同一性を区別し得る。

【0164】

さらに例として、画像化された物体に含まれるしるしは、所与の物品と関連付けられたロゴなどの画像の形態であり得る。物品上の特定の記号、エンブレム、または色の外観は、部分的にしか識別されない場合でも、分類モジュール138によって使用されて、画像分析モジュール134によって提供される可能な物品同一性の範囲から物品同一性を区別し得る。

40

【0165】

さらに追加的にまたは代替的に、画像背景減算モジュール414が本発明のシステムに組み込まえ得る。そのような画像背景減算は、分類モジュール138に提供された複数の物体同一性の中から単一の物体同一性を導出するのに支援するのに有用であり得る。画像背景減算モジュール414は、時間的背景減算を実行し得、シーンの少なくとも1つの初期画像が第1の時点で取得され、シーンの少なくとも1つの後続画像が第2の時点で取得され、後続の画像が第1の時点と第2の時点との間のシーンからの物体の取り出しを検出するために、初期画像から減算される。例として、棚114（図1）は、そこから物品

50

を除去する前および後に画像化され得、棚 1 1 4 から取り出される物品の性質および量の識別を支援するために使用される 2 つの画像の差。

【0 1 6 6】

画像背景減算モジュール 4 1 4 は、追加的または代替的に、空間的背景減算を実行し得、画像分析を単純化するために、不動要素をその所定のモデルに基づいてシーンの画像から減算し得る。

【0 1 6 7】

次に図 5 を参照すると、図 5 は、図 1 に示されるタイプの物品識別および追跡システムで採用されるカメラの配置を最適化するのに役立つスコアリングを示す簡略化されたグラフである。

【0 1 6 8】

図 1 を参照して本明細書で上述されたように、システム 1 0 0 は、好ましくは、建物 1 0 4 (図 1) などのシーンを視認し、その中の少なくとも 1 つの物体に関連する視覚データを取得するように配置された多数のカメラ 1 1 0 を含む。本発明の好ましい実施形態によれば、カメラ 1 1 0 の複数の可能な配置は、好ましくは、死角が最小化され、建物 1 0 4 内の物品の視野角の範囲が最大化されるカメラ 1 1 0 の最適な配置を発見するために、その中にカメラ 1 1 0 を設置する前に、建物 1 0 4 に関してコンピュータシミュレーションされる。

【0 1 6 9】

そのようなコンピュータシミュレーションは、建物 1 0 4 のコンピュータモデルを個々のユニットのグリッドに細分化することを含み得る。次に、第 1 のカメラ配置は、建物 1 0 4 のユニットに関して定義され得る。次に、すべてのユニットに、そのユニットが視認可能であるカメラの数に応じてサブスコアが割り当てられ得る。次に、建物を構成するすべてのユニットのサブスコアの合計は、そのカメラ構成の合計スコアを表す。追加的代替的な可能なカメラ配置をモデル化し、各々にスコアを割り当て、そのスコアは、所与のモデル内のユニットが見られる全体的なカメラ密度を表現する。

【0 1 7 0】

このようなスコアリングを示す非常に概略的な簡略化された例示的なグラフが図 5 に示されている。図 5 に示されるように、各構成に対して合計スコアが割り当てられる。合計スコアは、建物モデルの各ユニットの個々のサブスコアの合計であり、個々のサブスコアは、各それぞれのユニットが見られるカメラの数を表現する。図 5 に示されるデータの場合、第 1 1 のカメラ構成が最も高い構成スコアを持っているため、建物内の視認に最適であることが理解される。

【0 1 7 1】

建物 1 0 4 は、ユニットに分割することができ、いかなる個人 1 0 6 も存在しない場合、それに関してモデル化された様々なカメラ構成となる。追加的または代替的に、建物 1 0 4 は、建物内の個人 1 0 6 の存在を考慮することに基づいて、それに関してモデル化されたユニットおよびカメラ構成に分割され得る。構成スコア、したがって最適に識別されたカメラ配置は、建物内の個人 1 0 6 の存在または不在に応じて異なり得ることが理解される。

【0 1 7 2】

次に図 6 を参照すると、図 6 は、図 1 に示されるタイプの物品識別および追跡システムで有用であり、そのイベント追跡モジュール 1 4 0 内で特に有用である物品追跡のためのモデルの簡略化された概略図である。物品のそのような追跡は、本明細書で後述されるように、建物内の特定の個人と関連する物品の目録を提供するために有用であり得る。

【0 1 7 3】

図 6 に見られるように、多数のカメラ 1 1 0 は、好ましくは、建物 1 0 4 を画像化するように作動する。カメラ 1 1 0 によって取得された画像に対する画像処理の実行の結果として、特に図 3 を参照して本明細書で上述された画像処理の実行の結果として、建物 1 0 4 内の物品 1 0 2 および個人 1 0 6 (図 1) は、3 D 空間モデル 6 0 2 内の点 6 0 0 とし

10

20

30

40

50

てモデル化され得る。点 6 0 0 は、ここでは黒丸で示されている「物品点」6 1 0 の場合であることが示されているように、物品の存在と関連付けられた点であり得る。代替的に、点 6 0 0 は、ここでは塗りつぶされていない円として示される「人点」6 1 2 の場合であるように、個人の存在、より具体的には姿勢グラフ内の主要点によって表される個人の特定の身体部分と関連付けられた点であり得る。建物 1 0 4 が人と物品との両方を含む場合、物品点 6 1 0 と人点 6 1 2 との両方が点 6 0 0 の合計を含み得る。

【0 1 7 4】

次に図 7 に戻ると、モデル 6 0 2 に関連して実行されるプロセス 7 0 0 は、カメラ 1 1 0 が好ましくは 1 秒当たり 2 0 または 2 5 の画像などの高速フレームレートで建物 1 0 4 の画像を取得するように作動する第 1 のステップ 7 0 2 で開始することが見られる。カメラ 1 1 0 が作動する特定のフレームレートは、建物 1 0 4 の画像化要件に従って変化し得ることが理解される。

10

【0 1 7 5】

第 2 のステップ 7 0 4 に見られるように、各時点でカメラ 1 1 0 によって同時に取得された画像の各セットについて、図 3 ~ 4 に関して本明細書で上述されたような画像処理が、好ましくは、建物 1 0 4 内の、物体が点 6 0 0 によって表される無生物物品または人であり得る物体の分布の多次元、好ましくは 3 D の、モデルを生成するように実行される。そのような 3 D モデルは、建物 1 0 4 内の対象の領域に応じて、建物 1 0 4 全体または建物 1 0 4 の特定のサブの部分に対して生成され得ることが理解される。そのような 3 D モデルは、イベント追跡モジュール 1 4 0 に含まれ得るモデル生成器によって生成され得る。

20

【0 1 7 6】

カメラ 1 1 0 のうちの様々なカメラは、建物 1 0 4 内の物体のうちの異なる物体の画像を提供し得、同じ物体は、建物 1 0 4 内の物体の位置に応じて、異なる時間枠でカメラ 1 1 0 のうちの異なるカメラによって画像化され得ることが理解される。したがって、3 D 空間における所与の物体の位置に対応する点 6 0 0 のうちの特定の点は、第 1 の時点でのカメラ 1 1 0 のうちの特定のカメラからの画像に基づいて、および第 2 の時点でのカメラ 1 1 0 のうちの特定の他のカメラからの画像に基づいて、第 1 の時点と第 2 の時点との間の所与の物体の動きに起因して、第 2 のステップ 7 0 4 で生成され得る。

【0 1 7 7】

本発明のこの特徴は、建物 1 0 4 が密集した環境である場合に特に有利であることが理解される。建物 1 0 4 内の物体のおそらく密集した分布にもかかわらず、対応する点 6 0 0 によって表される物体が任意の所与の画像化時間フレームで少なくとも 2 つのカメラ 1 1 0 の視線に留まるという条件で、点 6 0 0 の 3 D モデルが連続的に生成され得る。

30

【0 1 7 8】

好ましくは、連続する時点に対して生成された多次元モデルは、少なくとも 1 つの物体の多次元モデルの変化を経時的に追跡するために比較される。より具体的には、点 6 0 0 のうちの各点の動きを追跡するために、連続する時点に対応するそれぞれの 3 D モデルにおける点 6 0 0 のうちの各点の位置が比較される。そのような追跡は、画像フレーム間の時間ギャップが、特定の物体が隣接するフレーム内のその物体の位置に対して大幅に移動するのにかかる時間よりも短いという仮定に基づいていることが理解される。このような追跡は、イベント追跡モジュール 1 4 0 に含まれ得るモデル追跡器によって実行され得る。

40

【0 1 7 9】

そのような追跡は、1 つの時点で生成された 3 D モデルからの情報が、2 つの時点が直接隣接していないにもかかわらず、別の時点で生成された 3 D モデルからの情報を増強するために使用されることを含み得ることが理解される。例えば、人点 6 1 2 のうちの特定の点がいくつかの連続するフレームから消え、その後再び出現する場合、その点 6 1 2 は、間にあるフレームからの点がないにもかかわらず、その点のより前の時間の時間インスタンスに追跡され得る。これは、例えば、1 つ以上の点 6 1 2 によって表される個人が一

50

時的に出て、その後、建物 104 に戻る場合に関連し得る。

【0180】

したがって、各時点に対して生成された 3D 空間内の点 600 の各 3D モデルは静的であるが、3D 空間内の点 600 の動的通過の連続 3D モデルは、カメラ 110 によって取得された建物 104 の画像のストリームに基づいて構築され得る、点 600 の通過は、第 3 のステップ 706 で見られるように、隣接するフレーム間の点位置の変化の比較によって追跡されることが理解される。

【0181】

次に、第 4 のステップ 708 に見られるように、目録は、第 3 のステップ 706 で構築された連続 3D モデルに基づいて生成され得る。3D モデルは、好ましくは、建物 104 内のすべての点 600 に対して生成されるが、目録は、好ましくは、対象のイベントなどの変化を定義する所定の「トリガ」基準を満たす点 600 のうちの点の位置の選択された変化のみに基づくことが理解される。そのような目録は、イベント追跡モジュール 140 に含まれ得るデータ分析器によって生成され得る。

10

【0182】

例として、図 6 に戻ると、建物 104 が買い物施設である場合、物品点 610 は、人点 612 によってモデル化された買い物客が購入できる物品を表し得る。そのようなシナリオでは、対象のイベントを表すと考えられ、それに基づいて目録を作成することができる点の動きのトリガ基準は、例えば図 6 の棚 114 によって表されるような購入できる物品の位置から所定の距離内に人点 612 のうちの 1 つが接近するようなイベントであり得る。そのようなトリガ基準を満たすイベントは、領域 720 で概略的に示され、そこでは、人点 612 が棚 114 に接近するのが見られる。この場合、3D モデルは、人点 612 によって表されるように、買い物客によって取り扱われる物品点 610 によって表されるように、物体の同一性を決定するために分析されることが理解される。そのようなイベント追跡は、買い物客と関連付けられた買い物リストを生成するために使用され得るか、または建物 104 内の買い物客の通過を理解するために分析の目的で使用され得る。

20

【0183】

さらに例として、建物 104 が倉庫である場合、物品点 610 は、人点 612 によって表される労働者によって取り扱われる物品を表し得る。そのようなシナリオでは、目録を作成することができる点の動きのトリガ基準は、人点 612 のうちの 1 つと関連付けられた置場から所定の距離内にある物品点 610 の 1 つの接近であり得る。この場合、3D モデルは、1 つ以上の人点 612 と関連付けられた置場に入出入りした物品点 610 によって表される物体の同一性を決定するために分析され得る。

30

【0184】

トリガ基準は、3D モデルにおける点 600 のやり取りに関連するイベントに限定されないことが理解される。例として、プロセス 700 は、物体に関連するデータが外部ソースから取得される第 5 のステップ 709 を含み得る。そのような外部ソースは、例として、棚 114 に位置付けられたセンサであり得る。棚 114 の 1 つ上の物品の重量の所与の変化は、物体が棚から取り出されたことを示すトリガ基準として定義され得る。そのようなトリガ基準に基づいて、3D モデルは、重量の変化を受けた棚 114 から、人点 612 によって表されるような買い物客によって取り出された、物品点 610 によって表されるような物体の同一性を決定するために分析されてもよい。

40

【0185】

これらのトリガ基準は、例としてのみ提供されており、対象のやり取りに対応する 3D モデル内の変化を区別するために、任意の適切なトリガ基準を使用することができることが理解される。

【0186】

次に図 8 に戻ると、物品の目録が生成される図 7 の第 4 のステップ 708 内の好ましいサブステップが示されている。第 1 のサブステップ 800 に見られるように、連続 3D モデル内の物体の位置が追跡されることが好ましく、その物体は物品または個人であり得る

50

。第2のサブステップ802に見られるように、所定のトリガ基準を満たす物体位置の変化が好ましくは識別される。

【0187】

次に、第3のサブステップ804に見られるように、所定の「トリガ」基準を満たす位置の変化を受けた物体の同一性および数量が発見される。第3のサブステップ804は、少なくとも1つのモデル化された物体に関連する情報を導出するために、第2のサブステップ802において発見されるように、少なくとも1つの所定の基準を満たす変化に関連するデータを分析することを含むことが理解される。物体の同一性および数量の発見は、例として、図4を参照して本明細書で上で概説されたように、自動的に実行され得ることが理解される。場合によっては、所与の物体の同一性または数量を確認するために、3Dモデルの分析において人間の入力があり得ることがさらに理解される。そのような人間の入力は、物体の数量および／または同一性に関して自動的に生成された情報を増強するために供給され得る。

10

【0188】

次に、第4のサブステップ806に見られるように、物体の数量および同一性は、買い物リストなどの目録に追加され得る。追加的または代替的に、所定のトリガ基準を満たすイベントに関連する情報は、建物内の物品および／または個人の通過の理解を得るために分析の目的で記憶され得る。第4のステップ806で買い物リストが生成される場合では、買い物リストは、好ましくは、建物104内の1つ以上の人点612によって表されるように、買い物客の通過中に連続的に更新される。例として、買い物リストは、買い物客が建物104から出る前に、支払いのために買い物客に自動的に提示され得る。

20

【0189】

したがって、所与の買い物客のための買い物リストの生成は、好ましくは、買い物客と関連付けられた物品の追跡に基づいており、3Dモデルで事前定義された「トリガ」基準を満たす買い物客のやり取りの分析に基づいており、必ずしも買い物客の識別を伴わないことが好ましいことが理解される。したがって、システム100（図1）は、入力として買い物客の同一性を必要とせずに、匿名化された方法で作動し得る。

【0190】

本明細書で上述されたように、システム100（図1）の構成要素によって実行される画像処理は、好ましくは、特に画像内の物体の存在および位置を検出するために、AIアルゴリズムに基づく画像の分析を含む。特に、多数の物品102および買い物客106を含む大規模な建物である建物104の場合では、必要な精度の水準で作動するようにそのようなAIアルゴリズムを訓練するために、非常に大量のデータが必要とされることが理解される。本発明の特に好ましい実施形態によれば、AIアルゴリズムのそのような訓練は、取得されたデータに基づく追加データの自動化された生成に基づいて、効率的、迅速かつ自動的に可能になる。そのような追加のデータは、取得されたデータの供給に加えて、その訓練中にAIアルゴリズムに供給され得、したがって、アルゴリズムの訓練を促進し、本発明を、大規模な物理的建物における、非常に大規模な実装に特によく適しているものにする。

30

【0191】

次に、図9Aおよび図10Bを参照すると、図9Aおよび10Bは、本発明のシステムの実施形態の実装形態の簡略化された斜視図および側面図である。

40

【0192】

図9Aおよび9Bに見られるように、物品102は、好ましくは、建物104の棚114上に配置され、天井112に取り付けられた多数のカメラ110によって視認可能である。物品102およびカメラ110の配置は、本発明の実施形態の以下の説明を明確にする目的で、非常に簡略化された方法で図9Aおよび9Bに示されており、物品102およびカメラ110は、本明細書に示されているものとは別の構成で配置され得ることが理解される。

【0193】

50

次に図 10 A および 10 B、11 A および 11 B、ならびに 12 A および 12 B を参照すると、図 10 A および 10 B、11 A および 11 B、ならびに 12 A および 12 B は、図 9 A および 9 B の実装形態の一部の簡略図であり、本発明の好ましい実施形態による、A I アルゴリズムによって注釈付けされた、人間の被写体を含む文脈における複数のカメラの視線および対応する画像を示している。

【0194】

次に図 10 A および 10 B に戻ると、第 1 のシナリオでは、買い物客 106 などの人間の被写体が物品 102 のうちの 1 つを取り扱っていることが示されている。例としてのみ、物品 102 を取り扱う買い物客 106 は、建物 104 において、カメラ 1 ~ 4 として示される 4 台のカメラによって同時に視認可能であるように示されている。カメラ 1 ~ 4 による買い物客 106 の画像化は、少なくとも本明細書で上述された A I アルゴリズムの訓練のための画像を提供する目的で、この文脈で好ましくは実行されることが理解される。そのような訓練の目的は、A I アルゴリズムが、カメラ 110 によって取得されたその画像に基づいて、建物 104 内の物品 102 の存在および位置を自動的に正確に識別することを可能にすることである。A I アルゴリズムが、十分な精度でカメラ出力内の物品 102 の存在および位置を識別することができるように訓練されると、A I アルゴリズムは、図 1 ~ 8 を参照して本明細書で上述されたように、システム 100 内、より具体的にはその物品識別器および追跡器 130 (図 1) 内に実装され得る。

【0195】

図 10 A および 10 B に見られるように、カメラ 1 はシーンの画像 1001 を出力し、カメラ 2 はシーンの画像 1002 を出力し、カメラ 3 はシーンの画像 1003 を出力し、カメラ 4 はシーンの画像 1004 を出力する。画像 1001、1002、1003、および 1004 は、好ましくは、その中の物品 102 の存在および位置を識別するために、A I アルゴリズムによって処理される。このような画像処理に使用される A I アルゴリズムは、シード A I または抽象 A I ネットワークと呼ばれることがあり、シード A I は、限られた範囲で、大きい程度の誤差で物品の存在および位置を識別することができることが理解される。本明細書で後述される本発明の実施形態の目的は、シード A I の性能を改善し、したがってシステム 100 内でのその実装を可能にするためにシード A I を訓練するための追加のデータを提供することである。

【0196】

A I アルゴリズムによる画像 1001、1002、1003、および 1004 の処理の結果として、画像 1001、1002、1003、および 1004 は、好ましくは、物品 102 の位置にその中に描かれた注釈付きのバウンディングボックスである。図 10 B に見られるように、バウンディングボックス 1010 は、好ましくは、画像 1002、1003、および 1004 の各々に追加される。しかしながら、物品 102 が買い物客 106 によって画像 1001 で妨害され、したがってカメラ 1 の視線 1012 に沿っていないため、バウンディングボックスは画像 1001 に追加されない。

【0197】

同様に、図 10 A および 10 B に示されるものとは買い物客 106 およびカメラ 1 ~ 4 に関して物品 102 の異なる配置を含む第 2 のシナリオにおける図 11 A および 11 B に見られるように、カメラ 1 は画像 1101 を出力し、カメラ 2 は画像 1102 を出力し、カメラ 3 は画像 1103 を出力し、カメラ 4 はシーンの画像 1104 を出力する。画像 1101、1102、1103、および 1104 は、好ましくは、その中の物品 102 の存在および位置を識別するために、A I アルゴリズムによって処理される。

【0198】

A I アルゴリズムによる画像 1101、1102、1103、および 1104 の処理の結果として、画像 1101、1102、1103、および 1104 は、好ましくは、物品 102 の位置にその中に描かれた注釈付きのバウンディングボックスである。図 11 B に見られるように、バウンディングボックス 1110 は、好ましくは、画像 1101、1102、および 1103 の各々に追加される。しかしながら、物品 102 が買い物客 106

によって画像 1104 で妨害され、したがってカメラ 4 の視線 1112 に沿っていないため、バウンディングボックスは画像 1104 に追加されない。

【0199】

ここで図 12 A および 12 B に戻ると、図 11 A に示されるものと同じ買い物客 106 およびカメラ 1 ~ 4 に関する物品 102 の配置を含む第 3 のシナリオが示されている。しかしながら、図 12 B に見られるように、ここでの A I アルゴリズムによる画像 1101、1102、1103、および 1104 の処理は、画像 1102 および 1103 にのみバウンディングボックス 1110 の挿入をもたらす。物品 102 が画像 1101 に出現しているにもかかわらず、画像 1101 にはバウンディングボックスが挿入されていない。このような場合は、A I 失敗ケースと呼ばれることがあり、A I アルゴリズムによって分析されたカメラ出力、ここでは画像 1101 は、物品 102 の存在および位置を示す情報を含むことが、そのような情報がカメラ出力に存在するにもかかわらず、発見されないことを意味する。このような A I 失敗は、A I アルゴリズムが、十分な水準の精度で画像内の物品の存在および位置を識別するように訓練されていない結果、A I アルゴリズムの限界に起因して発生する可能性がある。

10

【0200】

画像 1101 に関して図示された A I 失敗の場合は、画像 1104 の場合とは異なることが理解される。どちらの場合も、A I アルゴリズムは画像内の物品 102 の存在および位置を識別せず、それぞれの画像にバウンディングボックスは追加されない。しかしながら、画像 1104 の場合では、物品 102 はまったく画像内に出現せず、したがって A I アルゴリズムは画像内の物品 102 を正しく識別しないのに対し、画像 1101 の場合では、物品 102 は画像内に出現し、画像内の物品 102 の存在および位置の識別の欠如は、画像の生来の特性ではなく、A I アルゴリズムの失敗に起因するものである。

20

【0201】

図 10 A ~ 12 B に示されるシナリオにおける画像 1001、1002、1003、1004、および 1101、1102、1103、1104 の処理の前に、A I アルゴリズムは、好ましくは、物品 102 の同一性を備えていることが理解される。物品 102 の同一性は、好ましくは、人間の専門家によって発見され、A I アルゴリズムを訓練するために A I アルゴリズムに入力される。しかしながら、本発明のいくつかの実施形態では、物品 102 の同一性は、必ずしも人間の専門家によって A I アルゴリズムに入力され得ず、むしろ機械学習に基づいて自動的に生成され得ることが想定される。

30

【0202】

図 12 A および 12 B の例の画像 1101 などの A I 失敗ケースを改善するために、画像修正が採用され得る。図 13 A は、本発明の好ましい実施形態による、画像修正による A I 失敗ケースの改善を示す、図 9 A および 9 B の実装形態の一部の簡略化された図である。

【0203】

図 13 A に見られるように、画像 1101 などの A I 失敗ケース画像は、シーンを視認するカメラ 110 のうちの少なくとも 1 台のカメラによって出力され得る。図 13 A および 13 B を参照して本明細書で説明されるような画像修正方法は、1 台以上のカメラの出力に対して実行され得ることが理解される。図 12 B を参照して説明された人工知能を採用することに基づいて、カメラ出力がシーン内の少なくとも 1 つの物体の存在および位置を示す情報を含まないことを確認した後、カメラ出力が修正され得る。

40

【0204】

好ましくは、カメラ出力の修正は、A I 失敗画像を相互に異なる多数の角度で回転させ、A I を採用して、所与の角度で回転されるように、修正された各出力を分析して、画像内の物品の存在および位置を発見することを含む。

【0205】

ここで、例として、元の画像 1101 は、4 つの異なる角度で回転することによって修正されて、回転画像 1302、1304、1306、および 1308 を生成することが示

50

されている。画像 1 1 0 1 を分析するために元々うまく採用されなかった A I アルゴリズムは、好ましくは、回転画像 1 3 0 2、1 3 0 4、1 3 0 6、および 1 3 0 8 の各々を分析するために再採用される。修正された A I 失敗画像に A I アルゴリズムを採用すると、場合によっては、画像内の物品の存在および場所の識別に成功し、したがって、バウンディングボックスによる画像の注釈付けに成功することが発見されている。ここで、例として、バウンディングボックス 1 3 1 0 が、その中の物品 1 0 2 の位置を識別する画像 1 3 0 2、1 3 0 4、1 3 0 6、および 1 3 0 8 の各々に、それによる A I アルゴリズムによる処理の結果として追加されることが示されている。

【0206】

次に、注釈付き画像は、好ましくは、フレーム 1 3 2 0 に示されるように、バウンディングボックス 1 3 1 0 の重複領域における物品 1 0 2 の位置に対応する、おそらく不規則な輪郭を定義するために、逆回転によって再位置合わせされ、その中に現在存在するバウンディングボックス 1 3 1 0 は重ね合わされる。次に、フレーム 1 3 2 4 に示されるように、新しいデカルトバウンディングボックス 1 3 2 2 が物品の位置に挿入され得る。最終的なバウンディングボックス 1 3 2 2 は、物品位置のより精緻化されたレンダリングに基づいており、この精緻化された物品位置は、順番に好ましくは個々のバウンディングボックス 1 3 1 0 の重ね合わせに基づいているので、フレーム 1 3 2 4 に示される最終的なバウンディングボックス 1 3 2 2 は、様々な個々のバウンディングボックス 1 3 1 0 が重なり合っていない場合よりも正確であることが理解される。物品 1 0 2 およびバウンディングボックス 1 3 2 2 は、その提示を明確にする目的で、人間の被写体に関してフレーム 1 3 2 4 内に誇張されたサイズで示されていることが理解される。

【0207】

本明細書に記載される画像修正は、元の画像の回転に関して詳述されているが、画像カラスケールの変化、画像ワーピング、および画像倍率の選択的变化を例として含む、追加または代替のタイプの画像修正も可能であることが理解される。

【0208】

本明細書で説明される画像修正は、画像 1 1 0 1 などの A I 失敗画像への適用に限定されることがさらに理解される。むしろ、そのような画像修正は、バウンディングボックスの位置をより良く定義するために、A I アルゴリズムが最初に物品の存在および位置をうまく識別した画像にも適用され得る。このような場合では、A I アルゴリズムによって処理された元の修正されていない画像には、物品の存在と場所を示すバウンディングボックスで注釈付けされ得るが、バウンディングボックスのサイズおよび／または位置は不正確になる可能性がある。本明細書で説明される画像修正方法の使用は、バウンディングボックスを厳密化するために有利である可能性がある。

【0209】

次に図 1 3 B を参照すると、図 1 3 B は、図 1 3 A の実施形態に従って実行される A I 失敗ケースの改善および／またはバウンディングボックスの厳密化のステップを示す簡略化されたフローチャート図である。

【0210】

図 1 3 B に見られるように、画像修正に基づく A I 失敗ケースおよび／またはバウンディングボックス厳密化の改善のための方法 1 3 4 0 は、修正のために A I 処理された画像が選択される第 1 のステップ 1 3 4 2 で開始し得る。次に、方法 1 3 4 0 は、第 2 の問い合わせステップ 1 3 4 4 において、第 1 のステップ 1 3 4 2 において選択された画像がバウンディングボックスを含むかどうかのチェックに進む。問い合わせステップ 1 3 4 4 は、画像の A I 処理後にバウンディングボックスが描画されなかった A I 失敗画像と、すでにバウンディングボックスを含む A I 成功処理画像とを区別するのに役立つことが理解される。

【0211】

第 2 の問い合わせステップ 1 3 4 4 において、画像がバウンディングボックスを含むことが発見された場合、画像は第 3 のステップ 1 3 4 6 で回転され、第 4 のステップ 1 3 4

8において回転画像上にバウンディングボックスを再描画するために、A Iアルゴリズムが適用される。次に、第4のステップ1348において回転画像上に描画された新しいバウンディングボックスのパラメータが、第5のステップ1350に見られるように、画像に存在する元のバウンディングボックスのパラメータと比較される。このようなパラメータは、バウンディングボックスの寸法、バウンディングボックスの位置、または関連するその他のパラメータが含み得る。

【0212】

次に、第6の問い合わせステップ1352に見られるように、方法1340は、2つのバウンディングボックス、すなわち、回転画像内の元のバウンディングボックスおよび新しく描画されたバウンディングボックスのパラメータが、所定の閾値を超えて異なるかどうかを確認する。パラメータが所定の閾値を超えて異なる場合では、第7のステップ1354に見られるように、元のバウンディングボックスは許容可能であると考えられ、元の画像のさらなる処理は必要でないと考えられる。この場合、図14A~14Dを参照して本明細書で以下に概説されるように、画像内のバウンディングボックスは、バウンディングボックスをさらに厳密化する必要がなく、画像をさらに処理する準備ができているように、十分に正確に描写されていると考えられる。

【0213】

新しいバウンディングボックスと比較した元のバウンディングボックスのパラメータが、第6の問い合わせステップ1352で発見されたように、所定の閾値を超えて異なる場合、または画像がバウンディングボックスを含まない場合では、第2の問い合わせステップ1344において、方法1340は、第8のステップ1360に進み、そこで、画像は、N個の異なる角度だけ回転され、Nは、任意の整数であり得る。第9のステップ1362に見られるように、各回転画像について、画像に注釈付けし、その中に物品の存在および位置を示すバウンディングボックスを描画することを試みるために、回転画像は好ましくはA Iアルゴリズムによって処理される。場合によっては、第9のステップ1362が成功し、回転画像に追加のバウンディングボックスが描画されることがあるが、他の場合には、第9のステップ1362は、A Iアルゴリズムが画像内に出現する物品を識別することに失敗すること、または物品が実際に画像内に出現しないことのいずれかに起因して失敗し得ることが理解される。

【0214】

さらに、第9のステップ1362が、以前に処理に成功しなかった画像に新しいバウンディングボックスを描くことに成功した場合、方法1340は、画像修正によってA I失敗ケースを改善することに成功したことが理解される。このようなアプローチは、A Iアルゴリズムによってうまく処理された画像を修正して、そこからさらに情報を抽出できる従来の画像修正アプローチとは異なる。対照的に、本発明の好ましい実施形態では、A Iアルゴリズムによってうまく処理されなかった画像は、画像を改善し、以前にうまく処理されなかった画像をうまく処理された画像に変換するために修正される。しかしながら、本発明のいくつかの実施形態では、方法1340は、これらの画像に基づいてさらに多くのデータを生成するために、うまく分析された画像に対してさらに実行され得ることが理解される。

【0215】

第10のステップ1364および第11のステップ1366に見られるように、そのA I処理に続くバウンディングボックスを含むすべての画像が選択され、空間的に位置合わせされる。そのような空間的位置合わせは、第12のステップ1368に見られるように、すべての画像を元の方向に戻し、それによってその後続の重ね合わせを可能にするように、画像の逆回転を含み得る。第10のステップ1364は、シーン内の少なくとも1つの物体の存在および位置を示す情報を含む少なくとも1台のカメラの出力と、画像のA I処理によって発見されるような、当該シーン内の少なくとも1つの物体の存在および位置を示す情報を含まない少なくとも1台のカメラの出力とを区別するのに役立つことが理解される。

【0216】

好ましくは画像内のバウンディングボックスによって定義されるように、物体の存在および位置を示す情報を含むことが発見されたこれらの画像のみが、第11のステップ1366において相互に空間的に位置合わせされ、第12のステップ1368において重ね合わされる画像のセットにまとめられる。

【0217】

次に、第13のステップ1370に見られるように、多角形は、好ましくは、バウンディングボックスの重複領域に基づいて定義される。多角形のサイズおよび位置は、好ましくは、物品のサイズおよび位置に対応することが理解される。第14のステップ1372に見られるように、多角形の位置に対応する新しいデカルトバウンディングボックスが追加され、この新しいバウンディングボックスは、好ましくは、個々の画像のいずれかに存在する元のバウンディングボックスよりも正確に物品の位置に対応する。

【0218】

次に図14Aおよび14Bを参照すると、図14Aおよび14Bは、本発明の好ましい実施形態による、3DモデリングによるAI失敗ケースの改善を示す、図9Aおよび9Bの実装形態の一部の簡略図である。

【0219】

図14Aに見られるように、図12Aに示されるものと同じ買い物客106およびカメラ1～4に関する物品102の配置を含む第3のシナリオが示されている。カメラ出力1101、1102、1103、および1104は、好ましくは、その中の物品102の存在および位置を発見するためにAIアルゴリズムによって処理され、この物品は、画像に注釈付けするバウンディングボックスによって示され得る。そのような処理は、図10A～12Bを参照して本明細書で上述された元の画像のAI処理、ならびに図13Aおよび13Bを参照して本明細書で上述された画像修正に基づくAI処理を含み得ることが理解される。

【0220】

この場合、例として、バウンディングボックス1400は、画像1101および1103に挿入されているように見え、バウンディングボックスは、画像1102および1104に挿入されているようには見えない。画像1101および1103の場合に示されるバウンディングボックス1400は、その初期AI分析に続いて画像に最初に追加されたバウンディングボックスであり得、画像修正の方法でその後改善された初期バウンディングボックスに基づいて厳密化されたバウンディングボックスであり得、または図13Aおよび13Bを参照して本明細書で上述されたように、画像修正に基づいて改善された初期AI失敗画像に基づく新しいバウンディングボックスであり得ることが理解される。

【0221】

次に、画像1101および1103に存在するバウンディングボックス1400の特定の性質に関係なく、3D投影モデリングステップ1402で概略的に示されるように、好ましくはカメラ1～4によって出力されるすべての画像に対して3Dモデリング投影が実行されることが好ましい。3D投影モデリングステップ1402は、好ましくは、AIアルゴリズムによって分析されるように、物品102の存在および位置を示す情報を含むことが発見されるカメラの出力と、AIアルゴリズムによって分析されるように発見されないカメラの出力とを区別し、物品102の存在および位置を示す情報を含まない出力を、その出力に物品102の存在および位置を示す情報を追加することによって強化する。より具体的には、3D投影モデリングは、好ましくは、バウンディングボックスと関連付けられた画素に対応する光線の交差によって発見されるように、対応する画像内のバウンディングボックスの位置に基づいて、バウンディングボックスを含まない画像にバウンディングボックスを追加することを含む。3D投影モデリングは、人工知能を採用して画像分析器で実行され得る。しかしながら、3D投影モデリングは、バウンディングボックスによる画像の注釈に限定されず、任意の輪郭を描く形状または領域による画像内の物品の存在の指示を含み得ることが理解される。3D投影モデリングに関するさらなる詳細は、図

1 4 C を参照して本明細書で以下に提供される。ここで、例として、3 D 投影モデリングの実行に続いて、新しいバウンディングボックス 1 4 0 4 が画像 1 1 0 2 および 1 1 0 4 に挿入されているのが見られる。

【0 2 2 2】

少なくともバウンディングボックスによって明らかに描写された物品が実際には画像に出現しない可能性があるため、3 D 投影モデリングによって実行されるようなバウンディングボックス 1 4 0 4 などのバウンディングボックスの画像への追加は必ずしも正確ではないことが理解される。ここで、例として、物品 1 0 2 は、買い物客 1 0 6 によるその妨害のために画像 1 1 0 4 に出現せず、したがって、バウンディングボックス 1 4 0 4 による画像 1 1 0 4 の注釈は不正確である。

10

【0 2 2 3】

3 D モデリング投影によってうまく改善された画像と 3 D モデリング投影によってうまく改善されなかった画像とを区別するために、3 D モデリング投影によって出力された画像は、好ましくは、フィルタリングステップ 1 4 1 0 によって概略的に示されるようにフィルタリングされる。そのようなフィルタリングは、画像の視覚的または幾何学的特性に基づき得、図 1 4 D を参照して本明細書で以下でさらに詳述される。

【0 2 2 4】

ここで、例として、画像フィルタリングの結果として、画像 1 1 0 4 は、バウンディングボックス 1 4 0 4 を含むものとして誤って拒否され、画像 1 1 0 1、1 1 0 2、および 1 1 0 3 は、物品 1 0 2 の存在および位置に関する正確な情報を含む、うまく処理された画像として受け入れられる。うまく処理された画像 1 1 0 1、1 1 0 2、および 1 1 0 3 は、好ましくは、許容可能な精度で物品存在および位置を識別するためにアルゴリズムをさらに訓練するために、A I 画像処理アルゴリズムにフィードバックされる。このような A I 処理および訓練は、好ましくは、シード A I アルゴリズムが訓練され、成熟したアルゴリズムに発展するまで反復的に実行され、さらなる訓練が不要になるように十分なレベルの精度で実行される。

20

【0 2 2 5】

ここで、図 1 4 C を参照すると、図 1 4 C は、図 1 4 B の実施形態に従って実行された A I 失敗ケースの改善の 3 D 投影モデリング段階に含まれるステップを示す簡略化されたフローチャート図である。

30

【0 2 2 6】

図 1 4 C に見られるように、3 D 投影モデリング 1 4 0 2 のための方法 1 4 2 0 は、好ましくは、第 1 のステップ 1 4 2 2 で開始し、バウンディングボックスを含む各画像内のバウンディングボックスと関連付けられた少なくとも 1 つの主要画素位置が発見される。主要画素位置は、バウンディングボックス内、例えば、バウンディングボックスの中心またはバウンディングボックスの四隅の各々にあり得るか、またはそれに近接してバウンディングボックスの外側にあり得ることが理解される。

【0 2 2 7】

次に、第 2 のステップ 1 4 2 4 に見られるように、各主要画素位置に対応するカメラ光線が発見されることが好ましい。各画素位置に対応するカメラ光線の発見は、図 2 A および 2 B を参照して本明細書で上述されるように、好ましくは、最初のカメラ間較正に基づくことが理解される。このようなカメラ間較正は、画素対光線較正を確立するのに役立つ、シーン内の物品を視認する特定のカメラを照射する光線のパラメータを、物品がカメラセンサに出現する画素位置に相関させてもよい。第 1 のステップ 1 4 2 2 で所与のバウンディングボックスに対して複数の主要な画素位置が発見される場合、対応する複数の光線が好ましくは、第 2 のステップ 1 4 2 4 で発見され、画像化された物品の位置および形状の両方を確認することを可能にする。

40

【0 2 2 8】

第 3 のステップ 1 4 2 6 に見られるように、次に、好ましくは、第 2 のステップ 1 4 2 4 において識別された光線の間でマッチングが実行され、光線交差点または光線交差領域

50

が好ましくは確認される。光線交差点または光線交差領域は、画像化された物品の位置に対応すると考えられ得ることが理解される。少なくとも2つの光線の交差は、光線交差領域において画像化された物品102の存在を確立するのに十分であると考えられ得る。

【0229】

第4のステップ1428に見られるように、第3のステップ1426で発見された交差点と交差しない光線に相関する主要な画素位置に対応するバウンディングボックスは、誤って注釈付けされたバウンディングボックスとみなされ、したがって拒否される。第4のステップ1428は、AIアルゴリズムが実際に物品を含まない画像にバウンディングボックスを誤って配置するか、または画像内の誤った位置にバウンディングボックスを配置する可能性がある誤検知を除外するのに役立つことが理解される。

10

【0230】

第5のステップ1430に見られるように、光線交差に基づいて物品位置が確立されると、バウンディングボックスを含まない出力を有するカメラの場合、物品位置に対応するカメラ光線が発見され得る。カメラ光線に対応する画素位置は、第6のステップ1432に見られるように、既知の画素対光線較正に基づいて確認され得る。続いて、第7のステップ1434に見られるように、新しいバウンディングボックスは、第6のステップ1432で識別されたその画素位置に以前はバウンディングボックスを含まなかった画像に描画され得る。

【0231】

本発明の特定の実施形態では、方法1420のステップが省略され得ることが理解される。例えば、カメラ110が深度カメラとして具体化される場合、そのようなカメラの画素対光線相関は自動的に知られ、したがって、ステップ1424および1432などの相関ステップは省略され得る。

20

【0232】

次に、図14Dを参照すると、図14Dは、図14Bの実施形態に従って実行されるAI失敗ケースの改善のフィルタリング段階に含まれるステップを示す簡略化されたフローチャート図である。

【0233】

図14Dに見られるように、3D投影モデリング1402（図14B）の出力をフィルタリングする1410（図14B）のための方法1480は、好ましくは、第1のステップ1482で開始し、そこで3D投影モデリングによって出力された画像が提供される。これらの画像は、様々な可能なフィルタリングアプローチのうちの少なくとも1つによってフィルタリングされ得る。

30

【0234】

第2のステップ1484に見られるように、画像は、背景減算に基づいてフィルタリングされ得、静的背景を有する時系列画像は、買い物客106によって取り扱われている物品102を示さないとして拒否され、なぜならそのような画像は、物品の位置の変化に起因して経時的に変化する動的背景を有すると予想されるからである。

【0235】

追加的または代替的に、第3のステップ1486に見られるように、画像は、その共通の視覚的特性に基づいてフィルタリングされ得る。第3のステップ1486に従って、その視覚的特性の類似性を評価するために、第1のステップ1482において入力された画像内のバウンディングボックスの視覚的特性が比較される。このような視覚的特性は、色、テクスチャ、またはその他の特性を含み得る。画像内のバウンディングボックスが、第1のステップ1482において提供された画像のうちの他の画像内のバウンディングボックスの視覚的特性と比較して著しく異なる視覚的特性を有することが発見された場合、バウンディングボックスは、他のバウンディングボックスによって識別されるものと同一の物品に対応しないものと考えられ、拒否される。

40

【0236】

追加的または代替的に、第4のステップ1488に見られるように、画像は、その幾何

50

学的特性に基づいてフィルタリングされ得る。特に、最初のカメラの相互較正に基づいて位置が既知である所与のカメラと、対応するカメラ画像内のバウンディングボックスの位置に基づく見かけ上の物品の位置との間の距離が発見され得る。カメラと見かけ上の物品位置との間の距離は、図14Cを参照して説明されたように、カメラと物品位置との間の予想距離との整合性のために比較され得、この距離は、3Dモデリング投影によって発見される物品位置に基づいて既知である。カメラと物品距離が許容できないほど小さい場合、これは、画像1104（図14B）の場合に見られるように、別の物体が対象の物品を妨害していることを示していると考えられ、画像は拒否される。深度カメラの使用は、それによって提供される深度情報に起因して、幾何学的フィルタリングを容易にするのに特に有利であることが理解される。

10

【0237】

第5のステップ1490に見られるように、フィルタリングステップ1484、1486、および1488のすべてを通過した画像のみが、好ましくは、存在および位置物品102を示す正確に配置されたバウンディングボックスを含む画像であると考えられる。第2、第3、および第4のステップ1484、1486、および1488は、並行して実行され得るか、または連続して実行され得、第1のタイプのフィルタリングに基づいて許容可能であると考えられる画像のみが後続のフィルタリングステップに渡されることが理解される。第5のステップ1490において出力されたこれらの画像は、好ましくは、その後、AIALゴリズムのさらなる訓練のために元のAIALゴリズムにフィードバックされる。

20

【0238】

図9A～14Dを参照して本明細書で上述された本発明の実施形態では、AIALゴリズムによって物品102バウンディングボックスが好ましくは挿入される対象の物品102は、買い物客106によって取り扱われる典型的な無生物物品にあることが理解される。しかしながら、特定の場合において、例えば、建物104内の買い物客106の追跡を容易にするために、それによって取り扱われる物品102ではなく／それに加えて、買い物客106の存在および位置を識別するようにAIALゴリズムを訓練することが望ましい場合がある（図9Aおよび9B）。次に、AIALゴリズムが、AI失敗ケースの改善に基づいて個人106の存在および位置を特定するように訓練される本発明の好ましい実施形態について、図15A～17Dを参照しながら説明する。

30

【0239】

ここで図15Aおよび15Bに戻ると、買い物客106およびカメラ1～4に関する物品102の配置を含む追加のシナリオが示されている。カメラ1～4は、好ましくは、物品102を取り扱う買い物客106を同時に画像化し、それぞれの出力画像1501、1502、1503、および1504を提供する。

【0240】

図15Bに見られるように、画像1501、1502、1503、および1504は、好ましくは、その中の買い物客106の存在および位置を識別するために、AIALゴリズムによって処理される。買い物客106の存在および位置は、買い物客106の姿勢グラフを形成する主要点1510によって示される。典型的には、画像は、買い物客106の体の上側部分を表す姿勢グラフを形成するために、AIALゴリズムによって8つの主要点によって注釈付けされ得る。しかしながら、特定の画像処理要件に応じて、より多くのまたはより少ない数の主要点が姿勢グラフに含まれ得ることが理解される。

40

【0241】

ここで、例として、8つの主要点1510が画像1501および1502の各々に追加され、買い物客106の頭、腕、および胴体の上側部分を描写する姿勢グラフを形成することが見られる。画像1503と画像1504のどちらにも主要点は追加されていない。画像1503の場合では、買い物客106が画像1503に出現するにもかかわらず、AIALゴリズムによる画像の処理に続いて主要点は追加されない。このような場合は、AI失敗ケースと呼ばれることがあり、AIALゴリズムによって分析されたカメラ出力、

50

ここでは画像1503は、個人106の存在および位置を示す情報を含むことが、そのような情報がカメラ出力に存在するにもかかわらず、発見されないことを意味する。このようなAI失敗は、AIアルゴリズムが、十分な水準の精度で画像内の物品、ここでは人間物品の存在および位置を識別するように訓練されていない結果、AIアルゴリズムの限界に起因して発生する可能性がある。

【0242】

画像1503に関して図示されたAI失敗の場合は、画像1504の場合とは異なることが理解される。どちらの場合も、AIアルゴリズムは、画像内の物品（つまり、買い物客106）の存在および位置を識別せず、それぞれの画像に主要点は追加されない。しかしながら、画像1504の場合では、個人106はまったく画像内に出現せず、したがってAIアルゴリズムは画像内の個人106を正しく識別しないのに対し、画像1503の場合、個人106は画像内に出現し、画像内の個人106の存在および位置の識別の欠如は、画像の生来の特性ではなく、AIアルゴリズムの失敗に起因するものである。

【0243】

図15Aおよび15Bの例の画像1503などのAI失敗の場合を改善するために、画像修正が採用され得る。図16Aは、本発明の好ましい実施形態による、画像修正によるAI失敗ケースの改善を示す、図9Aおよび9Bの実装形態の一部の簡略化された図である。

【0244】

図16Aに見られるように、画像1503などのAI失敗ケース画像は、シーンを視認する少なくとも1台のカメラによって出力され得る。図16Aおよび16Bを参照して本明細書で説明されるような画像修正方法は、1台以上のカメラの出力に対して実行され得ることが理解される。図15Bを参照して説明された人工知能を採用することに基づいて、出力がシーン内の少なくとも1つの人間物体の存在および位置を示す情報を含まないことを確認した後、カメラ出力が修正され得る。

【0245】

好ましくは、修正することは、AI失敗画像を相互に異なる多数の角度で回転させ、AIを採用して、所与の角度で回転されるように、修正された各出力を分析して、画像内の人間物品の存在および位置を発見することを含む。

【0246】

ここで、例として、元の画像1503は、4つの異なる角度で回転されて、回転画像1602、1604、1606、および1608を生成するように示されている。画像1503を分析するために元のうまく採用されなかったAIアルゴリズムは、好ましくは、回転画像1602、1604、1606および1608の各々を分析するために再採用される。修正されたAI失敗画像にAIアルゴリズムを採用すると、場合によっては、画像の注釈付けに成功し、画像内の物品の存在および位置の識別をもたらすことが発見されている。ここで、例として、複数の主要点1610が、画像1602、1604、1606、および1608の各々に追加され、それによるAIアルゴリズムによる処理の結果として、その中の個人106の位置を少なくとも部分的に識別することが示されている。

【0247】

画像1602、1604、および1608の場合に示されるように、修正されたAI失敗画像におけるAIアルゴリズムの採用が部分的に成功し、画像内にすべてではないがいくつかの主要点が挿入される可能性があることが理解される。修正されたAI失敗画像に対するAIアルゴリズムの採用は、最大数の8つの主要点によって注釈付けされると見られる画像1606の場合に示されるように、画像内に最大選択数の主要点の挿入をもたらし、完全に成功し得ることが、さらに理解される。

【0248】

次に、注釈付き画像は、好ましくは、フレーム1620に示されるように、逆回転によって再位置合わせされ、その中に存在する主要点が重ね合わされて、対応する主要点の重複領域1618における各主要点の位置に対応するおそらく不規則な輪郭を定義する。次

に、フレーム 1 6 2 4 に示されるように、1 つ以上の新しい主要点 1 6 2 2 を、各主要点の重複位置に挿入することができる。フレーム 1 6 2 4 の最終的な姿勢グラフは、個々の対応する主要点の重ね合わせに基づく主要点位置のより精緻化されたレンダリングに順番に基づいているので、フレーム 1 6 2 4 に示される主要点の新しいセットの最終的な姿勢グラフは、様々な個々の主要点が重ね合わされなかった場合よりも正確であることが理解される。

【0 2 4 9】

本明細書に記載される画像修正は、元の画像の回転に関して詳述されているが、画像ラースケールの変化、画像ワーピング、および画像倍率の選択的变化を例として含む、追加または代替のタイプの画像修正も可能であることが理解される。

10

【0 2 5 0】

本明細書で説明される画像修正は、画像 1 5 0 3 などの A I 失敗画像への適用に限定されないことがさらに理解される。むしろ、そのような画像修正は、主要点の位置をより良く定義するために、A I アルゴリズムが最初に物品の存在および位置をうまく識別した画像にも適用され得る。このような場合、A I アルゴリズムによって処理された元の修正されていない画像には、人間物品の存在および位置を示す主要点のセットで注釈付けされ得るが、主要点のうちの 1 つ以上の位置が不正確になる可能性がある。本明細書で説明される画像修正方法の使用は、主要点を含む姿勢グラフを改善するために有利である可能性がある。本明細書で説明される画像修正方法の使用は、最大数ではないがいくつかの主要点を含む注釈付き画像の場合にも有利である可能性がある。そのような場合、画像修正を使用して、追加の主要点を補足することによって画像を改善することができ、したがって、買い物客 1 0 6 を表すより完全な姿勢グラフの生成につながる。

20

【0 2 5 1】

次に図 1 6 B を参照すると、図 1 6 B は、図 1 6 A の実施形態に従って実行される A I 失敗ケースの改善および／または姿勢グラフの改善におけるステップを示す簡略化されたフローチャート図である。

【0 2 5 2】

図 1 6 B に見られるように、A I 失敗ケースの改善および／または画像修正に基づく姿勢グラフ改善のための方法 1 6 4 0 は、修正のために A I 処理された画像が選択される第 1 のステップ 1 6 4 2 で開始し得る。次に、方法 1 6 4 0 は、第 2 の問い合わせステップ 1 6 4 4 において、第 1 のステップ 1 6 4 2 において選択された画像が、8 つの主要点などの可能な主要点の総数を含むかどうかのチェックに進む。問い合わせステップ 1 6 4 4 は、主要点の最大数を含む完全な姿勢グラフを含むうまく処理された画像と、画像の A I 処理後に主要点が最大よりも少ないかまたはまったく追加されなかった A I 失敗または部分的失敗画像とを区別するのに役立つことが理解される。

30

【0 2 5 3】

第 2 の問い合わせステップ 1 6 4 4 において、画像が主要点の総数を含むことが発見された場合、画像は第 3 のステップ 1 6 4 6 で回転され、第 4 のステップ 1 6 4 8 において回転画像上に主要点を再描画するために、A I アルゴリズムが適用される。次に、第 4 のステップ 1 6 4 8 において回転画像上に描かれた新しい主要点のパラメータが、第 5 のステップ 1 6 5 0 に見られるように、画像に存在する対応する元の主要点のパラメータと比較される。このようなパラメータには、主要点位置または他のパラメータが含まれる場合がある。

40

【0 2 5 4】

次に、第 6 の問い合わせステップ 1 6 5 2 に見られるように、方法 1 6 4 0 は、主要点の対応するペアのパラメータ、すなわち、回転画像における元の主要点および対応する新しく描かれた主要点が、所定の閾値を超えて異なるかどうかを確認する。パラメータが所定の閾値を超えて異なる場合では、第 7 のステップ 1 6 5 4 に見られるように、元の主要点のセットは許容可能であると考えられ、元の画像のさらなる処理は必要でないと考えられる。この場合、画像内の姿勢グラフは、図 1 7 A ～ 1 7 D を参照して本明細書で以

50

下で概説されるように、さらなる修正が不要であり、画像がさらなる処理の準備ができていように、十分に正確に描写されていると考えられる。

【0255】

新しい主要点と比較した元の主要点のパラメータが、第6の問い合わせステップ1652で発見されたように、所定の閾値を超えて異なる場合、または画像が主要点の総数を含まない場合では、第2の問い合わせステップ1644において、方法1640は、第8のステップ1660に進み、そこで、画像は、N個の異なる角度だけ回転され、Nは、任意の整数であり得る。第9のステップ1662に見られるように、各回転画像について、画像に注釈付けし、その中に人間物品の存在および位置を示す主要点のセットを描画することを試みるために、回転画像は好ましくはAIALゴリズムによって処理される。場合によっては、第9のステップ1662が成功し、回転画像に追加の主要点が描画されることがあるが、他の場合には、第9のステップ1662は、AIALゴリズムが画像内に出現する個人を識別することに失敗することに起因して、または個人が実際に画像内に出現しないことに起因して失敗し得ることが理解される。

10

【0256】

さらに、第9のステップ1662が、以前は失敗したAI処理された画像において新しい主要点を描画することに成功した場合、方法1640は、画像修正によってAI失敗ケースを改善することに成功したことが理解される。このようなアプローチは、AIALゴリズムによってうまく処理された画像を修正して、そこからさらに情報を抽出できる従来の画像修正アプローチとは異なる。対照的に、本発明の好ましい実施形態では、AIALゴリズムによってうまく処理されなかった画像は、画像を改善し、以前にうまく処理されなかった画像をうまく処理された画像に変換するために修正される。

20

【0257】

第10のステップ1664および第11のステップ1666で見られるように、そのAI処理に続く主要点を含むすべての画像が選択され、空間的に位置合わせされる。そのような空間的位置合わせは、第12のステップ1668に見られるように、すべての画像を元の方向に戻し、それによってその後続の重ね合わせを可能にするように、画像の逆回転を含み得る。第10のステップ1664は、シーン内の少なくとも1つの人間物体の存在および位置を示す情報を含む少なくとも1台のカメラの出力と、画像のAI処理によって発見されるような、当該シーン内の少なくとも1つの人間物体の存在および位置を示す情報を含まない少なくとも1台のカメラの出力とを区別するのに役立つことが理解される。

30

【0258】

好ましくは画像内の主要点によって定義されるように、物体の存在および位置を示す情報を含むこれらの画像のみが、第11のステップ1666において相互に空間的に位置合わせされ、第12のステップ1668において重ね合わせられる画像のセットにまとめられる。

【0259】

次に、第13のステップ1670に見られるように、多角形は、好ましくは、重ねられた主要点の各セットの重複領域に基づいて定義される。次に、第14のステップ1672に見られるように、各多角形の位置に対応する新しい主要点が追加され、この新しい主要点は、好ましくは、個々の画像のいずれかに存在する元の主要点よりも買い物客106の身体に関連部分により正確に対応する。

40

【0260】

次に図17Aおよび17Bを参照すると、図17Aおよび17Bは、本発明の好ましい実施形態による、3DモデリングによるAI失敗ケースの改善を示す、図9Aおよび9Bの実装形態の一部の簡略図である。

【0261】

図17Aに見られるように、図15Aに示されるものと同じ買い物客106およびカメラ1〜4に関する物品102の配置を含む第3のシナリオが示されている。カメラ出力1501、1502、1503、および1504は、好ましくは、その中の個人106の存

50

在および位置を発見するために A I アルゴリズムによって処理され、その個人は、画像に注釈付けする主要点のセットを含む姿勢グラフによって示され得る。そのような処理は、図 15 B を参照して本明細書で上述された元の画像の A I 処理、ならびに図 16 A および 16 B を参照して本明細書で上述された画像修正に基づく A I 処理を含み得ることが理解される。

【0262】

この場合、例として、複数の主要点 1700 が画像 1501 および 1503 に挿入されているように見え、主要点は画像 1502 および 1504 に挿入されていないように見える。画像 1501 および 1503 の場合に示される複数の主要点 1700 は、その初期 A I 分析に続いて画像に最初に追加された主要点であり得、画像修正の方法でその後改善された初期姿勢グラフに基づいて厳密化された主要点であり得、または図 16 A および 16 B を参照して本明細書で上述されたように、画像修正に基づいて改善された初期 A I 失敗画像に基づく新しい主要点のセットであり得ることが理解される。

【0263】

次に、画像 1501 および 1503 に存在する主要点 1700 の特定の性質に関係なく、3D 投影モデリングステップ 1702 で概略的に示されるように、好ましくはカメラ 1 ~ 4 によって出力されるすべての画像に対して 3D モデリング投影が実行されることが好ましい。3D 投影モデリングは、好ましくは、A I アルゴリズムによって分析されるように、個人 106 の存在および位置を示す情報を含むことが発見されるカメラの出力と、A I アルゴリズムによって分析されるように発見されないカメラの出力とを区別し、個人 106 の存在および位置を示す情報を含まない出力を、その出力に個人 106 の存在および位置を示す情報を追加することによって強化する。より具体的には、3D 投影モデリングは、好ましくは、主要点と関連付けられた画素に対応する光線の交差によって発見される、対応する画像内の主要点の位置に基づいて、主要点を含まない画像に主要点を追加することを含む。3D 投影モデリングは、人工知能を採用して画像分析器で実行され得る。3D 投影モデリングに関するさらなる詳細は、好ましくは、図 17 C を参照して以下に提供される。ここで、例として、3D 投影モデリングの実行に続いて、主要点 1704 の新しいセットが画像 1502 および 1504 に挿入されているのが見られる。

【0264】

少なくとも主要点によって明らかに描写された人間物品が実際には画像に出現しない可能性があるため、3D 投影モデリングによって実行されるような主要点 1704 などの新しい主要点のセットの画像への追加は必ずしも正確ではないことが理解される。ここで、例として、買い物客 106 は、物品 102 によるその妨害のために画像 1504 に出現せず、したがって、主要点 1704 による画像 1504 の注釈は不正確である。

【0265】

3D モデリング投影によってうまく改善された画像と 3D モデリング投影によってうまく改善されなかった画像とを区別するために、3D モデリング投影によって出力された画像は、好ましくは、フィルタリングステップ 1710 によって概略的に示されるようにフィルタリングされる。そのようなフィルタリングは、画像の視覚的または幾何学的特性に基づき得、図 17 D を参照して本明細書で以下でさらに詳述される。

【0266】

ここで、例として、画像フィルタリングの結果として、画像 1504 は、主要点 1704 を含むものとして誤って拒否され、画像 1501、1502、および 1503 は、個人 106 の存在および位置に関する正確な情報を含む、うまく処理された画像として受け入れられる。うまく処理された画像 1501、1502、および 1503 は、好ましくは、許容可能な精度で個人の存在および位置を識別するためにアルゴリズムをさらに訓練するために、A I 画像処理アルゴリズムにフィードバックされる。このような A I 処理および訓練は、好ましくは、シード A I アルゴリズムが訓練され、成熟したアルゴリズムに発展するまで反復的に実行され、さらなる訓練が不要になるように十分なレベルの精度で実行される。

【0267】

次に図17Cを参照すると、図17Cは、図17Bの実施形態に従って実行されるAI失敗ケースの改善の3D投影モデリング段階に含まれるステップを示す簡略化されたフローチャート図である。

【0268】

図17Cに見られるように、3D投影モデリング1702のための方法1720は、好ましくは、第1のステップ1721で開始し、少なくとも1つの主要点を含む各画像内の各主要点は、主要点に対応する身体上の位置に関してラベル付けされる。好ましくは、主要点のラベル付けに続いて、各主要点と関連付けられた少なくとも1つの主要画素位置が、第2のステップ1722で見出される。主要画素位置は、主要点内、例えば、主要点の中心または主要点の外側にあり得ることが理解される。

10

【0269】

次に、第3のステップ1724に見られるように、各主要画素位置に対応するカメラ光線が発見されることが好ましい。各画素位置に対応するカメラ光線の発見は、図2Aおよび2Bを参照して本明細書で上述されるように、最初のカメラ間較正に基づき得ることが理解される。このようなカメラ間較正は、画素対光線較正を確立するのに役立ち、シーン内の物品を視認する特定のカメラを照射する光線のパラメータを、物品がカメラ画像に出現する画素位置に相関させてもよい。

【0270】

第4のステップ1726に見られるように、次に、好ましくは、第3のステップ1724において識別された光線の間でマッチングが実行され、光線交差点または光線交差領域が好ましくは確認される。光線交差点または光線交差領域は、各それぞれのラベル付けされた主要点によって示される人間の身体上の位置に対応すると考えられ得ることが理解される。少なくとも2つの光線の交差は、光線交差領域で身体部分の存在を確立するのに十分であると考えられ得る。

20

【0271】

第5のステップ1728に見られるように、交差点と交差しない光線に相関する主要画素位置に対応する主要点は、誤って配置された主要点と考えられ、したがって拒否される。第5のステップ1728は、AIアルゴリズムは、実際には個人を含まない画像に主要点を誤って配置するか、または主要点に誤ってラベル付ける可能性がある誤検知を除外するのに役立つことが理解される。

30

【0272】

第6のステップ1730に見られるように、光線交差に基づいて身体部分の位置が確立されると、主要点を含まない出力を有するカメラの場合、身体部分の位置に対応するカメラ光線が発見され得る。カメラ光線に対応する画素位置は、第7のステップ1732に見られるように、既知の画素対光線較正に基づいて確認され得る。続いて、第8のステップ1734に見られるように、主要点の新しいセットは、第7のステップ1732で識別されたその画素位置に以前は主要点を含まなかった画像に描画され得る。

【0273】

次に、図17Dを参照すると、図17Dは、図17Bの実施形態に従って実行されるAI失敗ケースの改善のフィルタリング段階に含まれるステップを示す簡略化されたフローチャート図である。

40

【0274】

図17Dに見られるように、3D投影モデリング1702の出力をフィルタリングするための方法1780は、好ましくは、第1のステップ1782で開始し、そこで3D投影モデリングによって出力された画像が提供される。これらの画像は、様々な可能なフィルタリングアプローチのうちの少なくとも1つによってフィルタリングされ得る。

【0275】

第2のステップ1784に見られるように、画像は、背景減算に基づいてフィルタリングされ得、静的背景を有する時系列画像は、そのような画像が動的背景を有すると予想さ

50

れるので拒否される。

【0276】

追加的または代替的に、第3のステップ1786に見られるように、画像は、その共通の視覚的特性に基づいてフィルタリングされ得る。第3のステップ1786に従って、その視覚特性の類似性を評価するために、第1のステップ1782において入力された画像内の対応する主要点の視覚特性が比較される。このような視覚的特性は、色、テクスチャ、またはその他の特性を含み得る。画像内の主要点が、第1のステップ1782において提供された画像のうちの他の画像内の対応する主要点の視覚的特性と比較して著しく異なる視覚的特性を有することが発見された場合、主要点は、他の主要点によって識別されるものと同じ身体部分に対応しないものとみなされ、拒否される。

10

【0277】

追加的または代替的に、第4のステップ1788に見られるように、画像は、その幾何学的特性に基づいてフィルタリングされ得る。特に、最初のカメラの相互較正に基づいて位置が既知である所与のカメラと、主要点の位置に基づいて見かけ上の身体部分の位置との間の距離が発見され得る。カメラと見かけ上の身体部分との間の距離は、図17Cを参照して説明されたように、カメラと身体部分位置との間の予想距離との整合性のために比較され得、この距離は、3Dモデリング投影によって発見される身体部分の位置に基づいて既知である。カメラから身体部分までの距離が許容できないほど小さい場合、これは、画像1504（図17B）の場合に見られるように、別の物体が対象の身体部分を妨害していることを示していると考えられ、画像は拒否される。カメラ110としての深度カメラの使用は、それによって提供される深度情報に起因して、幾何学的フィルタリングステップ1888を容易にするのに特に有利であることが理解される。

20

【0278】

第5のステップ1790に見られるように、フィルタリングステップ1784、1786、および1788のすべてを生き延びた画像のみが、好ましくは、個人106の存在および位置を示す正確に配置された主要点を含む画像であると考えられる。第2、第3、および第4のステップ1784、1786、および1788は、並行して実行され得るか、または連続して実行され得、第1のタイプのフィルタリングに基づいて許容可能であると考えられる画像のみが後続のフィルタリングステップに渡されることが理解される。第5のステップ1790において出力されたこれらの画像は、好ましくは、その後、Aリアルゴリズムのさらなる訓練のために元のAリアルゴリズムにフィードバックされる。

30

【0279】

いくつかのシナリオでは、カメラ110によって画像化されるシーンは、複数の物品102または個人106を含み得ることが理解される。このような場合、Aリアルゴリズムによって識別される対象の物体が物品102である場合、対象の物品102の数に対応する複数のバウンディングボックスが描画され得る。代替的に、拡大された単一のバウンディングボックスが描画され得、アルゴリズムは、バウンディングボックスが複数の対象の物品を含むことを示す出力を含むことができる。

【0280】

Aリアルゴリズムによって識別される対象の物体が買い物客106などの個人である場合、複数の主要点が生成され得、主要点は、それらのラベル付けに基づいて、個々の姿勢グラフにグループ化され得る。

40

【0281】

物品が物体102であるかまたは人間の被写体106であるかにかかわらず、物品の存在および位置を識別するためのAリアルゴリズムが十分に訓練されると、Aリアルゴリズムは、図1～8を参照して本明細書で上述されたようにシステム100内で作動し得ることが理解される。

【0282】

次に図18を参照すると、図18は、バウンディングボックスの厳密化を示す図9Aおよび9Bの実装形態の一部の簡略化された図である。

50

【0283】

A I アルゴリズムによる画像の成功した処理（この成功した処理は、本明細書で上述されるような画像改良を含み得る）に続いて、成功した処理画像は、アルゴリズムを訓練するために、アルゴリズムにフィードバックされることが好ましい。さらなる訓練の目的で画像をアルゴリズムにフィードバックする前に、精度の高いバウンディングボックスがアルゴリズムのさらなる訓練に使用されることを保証するために、バウンディングボックスを含む画像に対してバウンディングボックスの厳密化を実行することが有益であり得る。

【0284】

バウンディングボックス1802を含む、うまく処理された画像1800の一部が図18に示されている。図18に見られるように、バウンディングボックス1802は、その中に、対象の物品102と、物品102を取り扱う買い物客106の手1804の一部とを含む。バウンディングボックス1802の精度を改善するために、画素セグメンテーションを使用して、手1804に対応する画像の部分と物品102に対応する画像の部分とを区別してもよい。さらに、画像1800の静止部分を除去するために、背景減算技術が適用され得る。そのような技術の結果として、バウンディングボックス1802は、その厳密化後の元のバウンディングボックス1802に基づく新しいバウンディングボックス1820の場合に見られるように、サイズが縮小され、より正確に位置付けられ得る。

10

【0285】

バウンディングボックスの厳密化は、当技術分野で知られているように、画素セグメンテーションおよびバックグラウンド減算以外の追加または代替の技術によって達成され得ることが理解される。

20

【0286】

本発明は、以下に特に特許請求されているものによって限定されないことが当業者には理解されよう。むしろ、本発明の範囲には、図面を参照して前述の説明を読むと、当業者が思いつくであろう、本明細書で上述された特徴の様々な組み合わせおよびサブ組み合わせ、ならびにそれらの修正および変形が含まれ、これらは先行技術にはないものである。

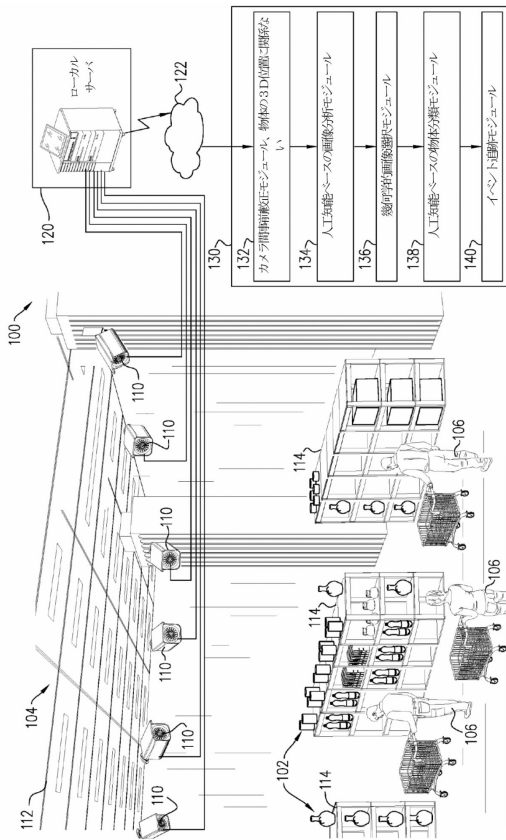
30

40

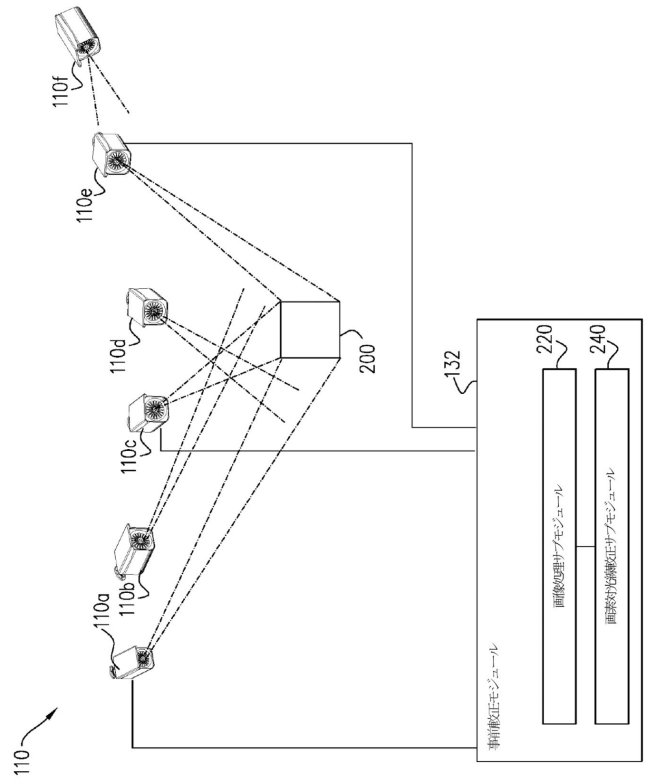
50

【図面】

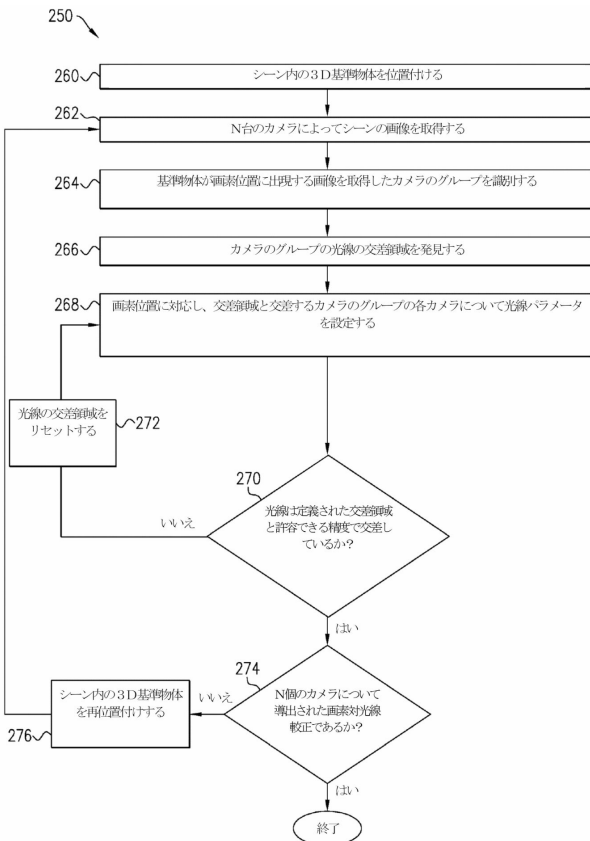
【図 1】



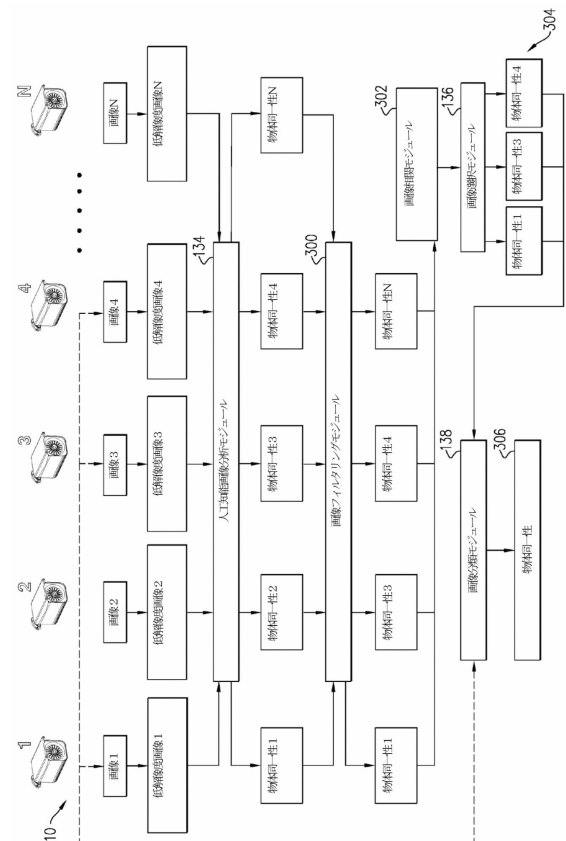
【図 2 A】



【図 2 B】



【図 3】



10

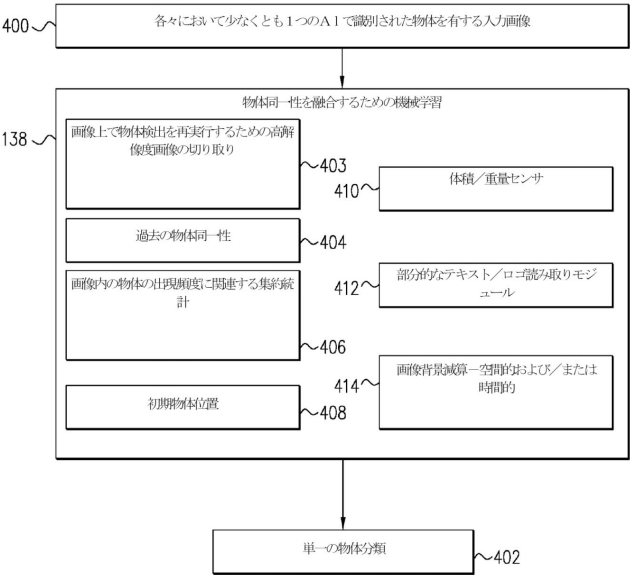
20

30

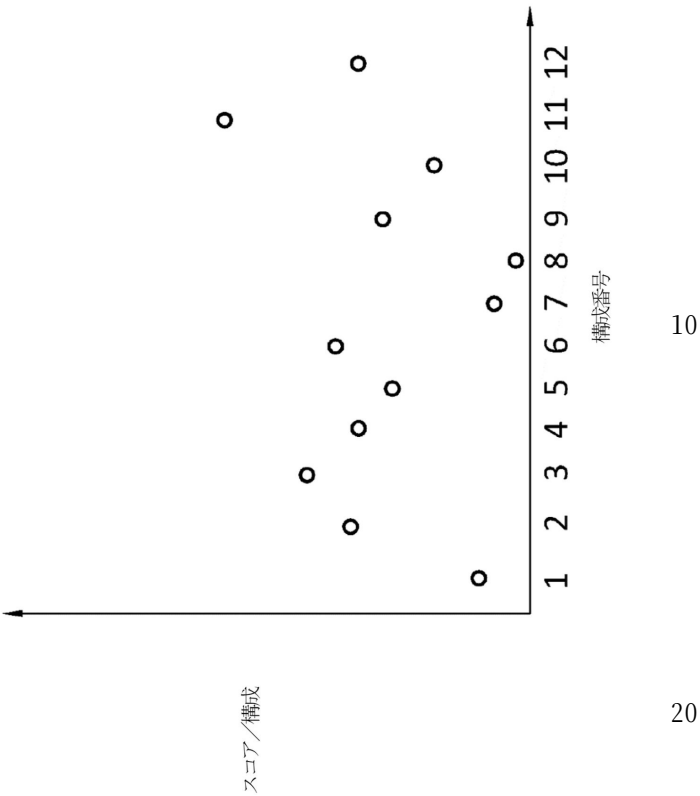
40

50

【図 4】



【図 5】



【図 6】

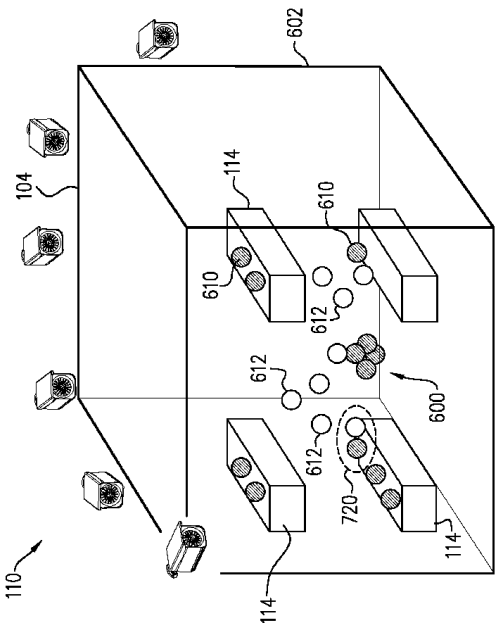
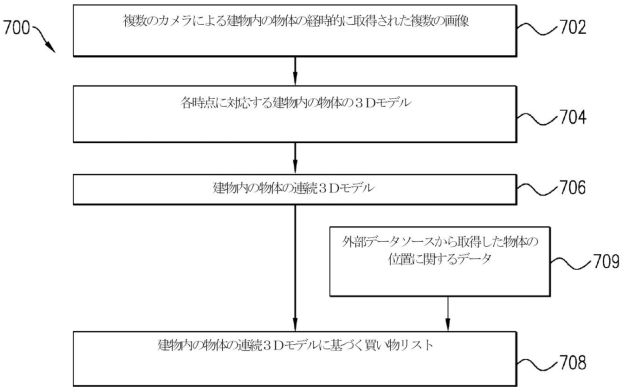


FIG. 6

【図 7】



10

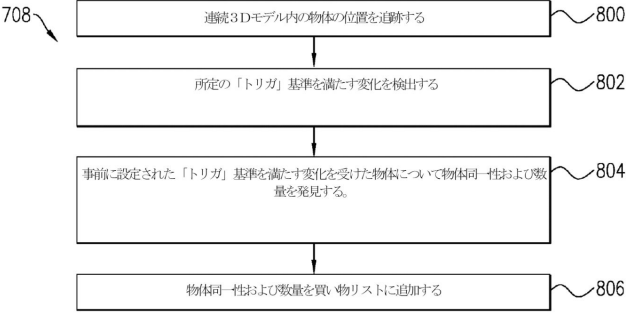
20

30

40

50

【図 8】



【図 9 A】

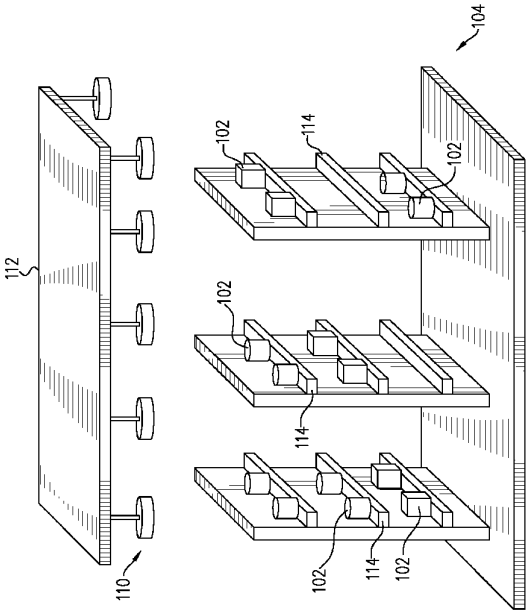


FIG. 9A

10

20

【図 9 B】

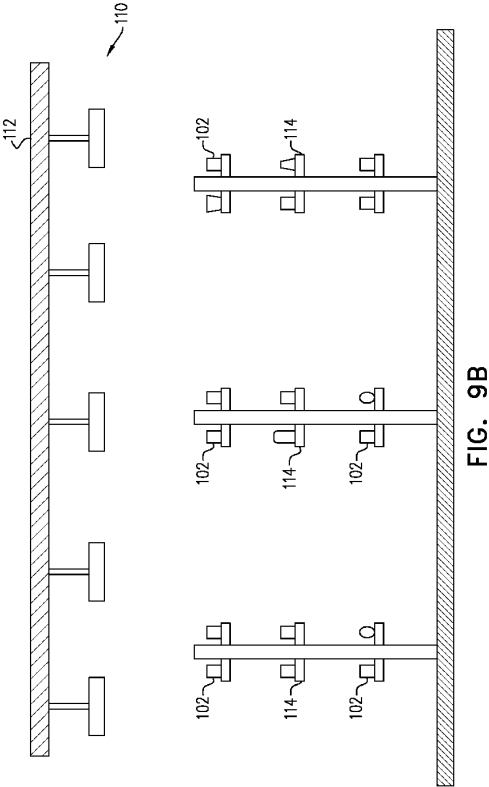
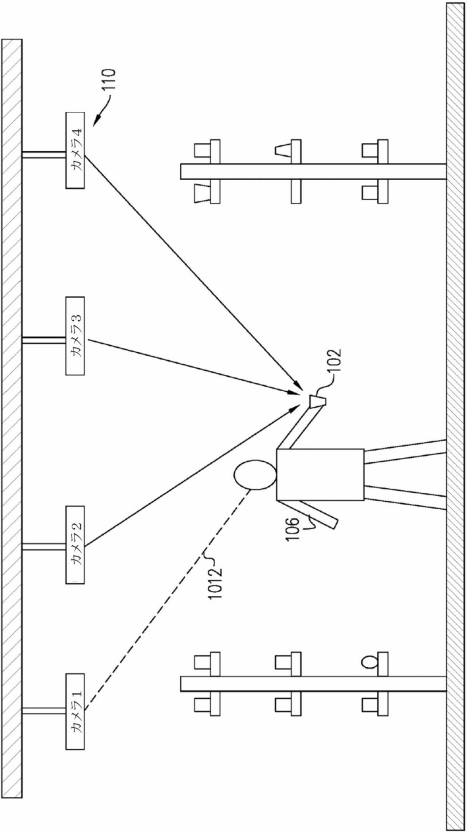


FIG. 9B

【図 10 A】

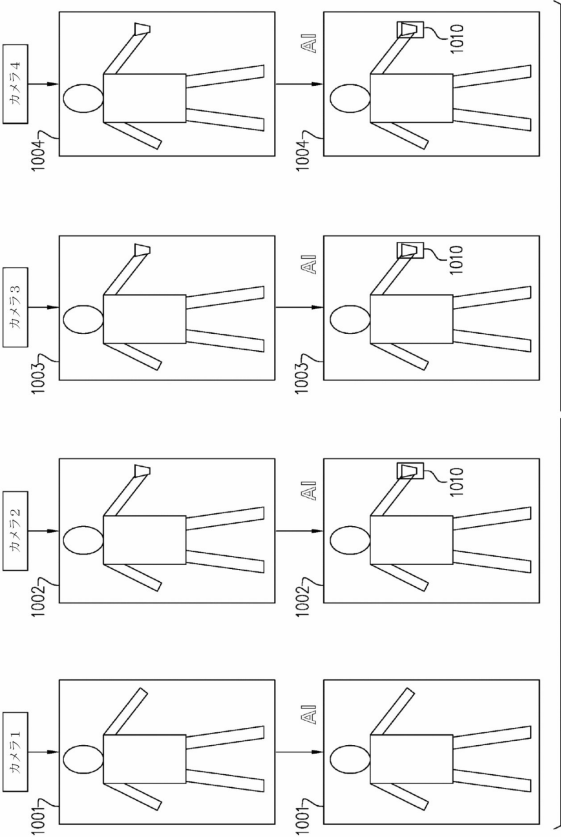


30

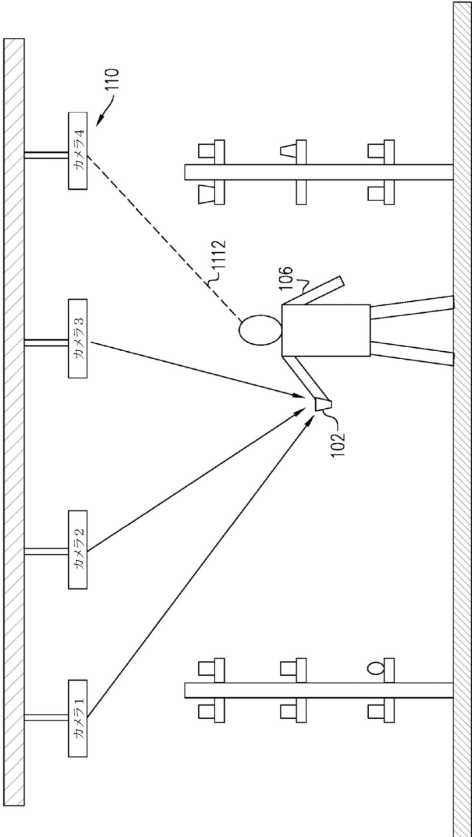
40

50

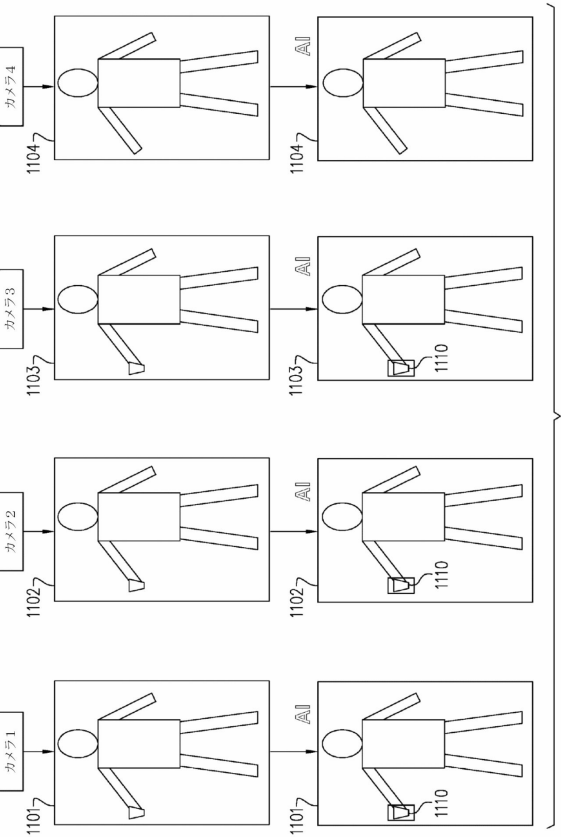
【図 1 0 B】



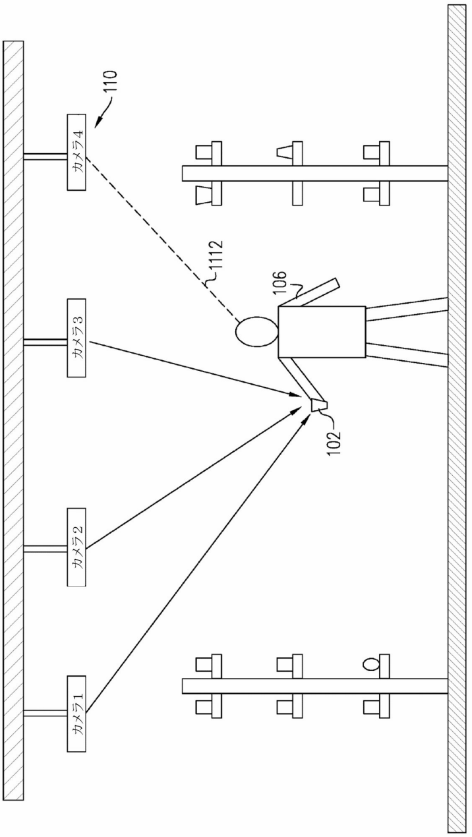
【図 1 1 A】



【図 1 1 B】



【図 1 2 A】



10

20

30

40

50

【図 1 2 B】

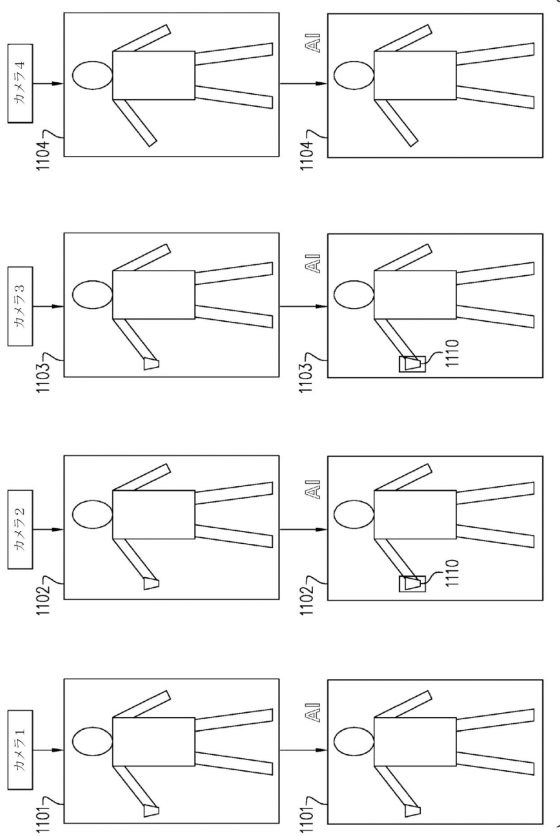
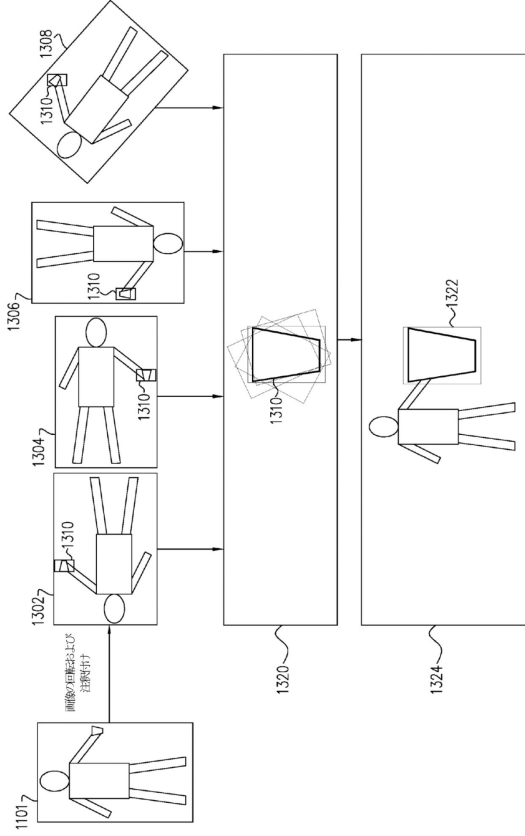
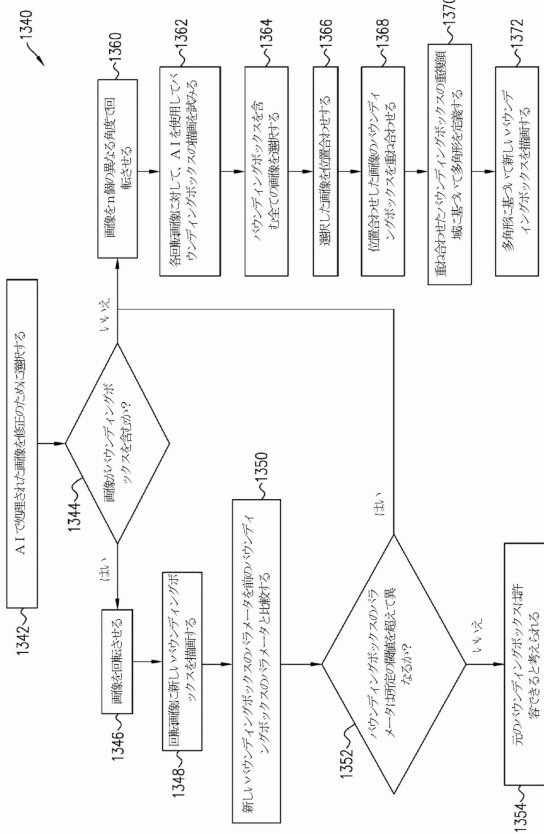


図 1 2 B

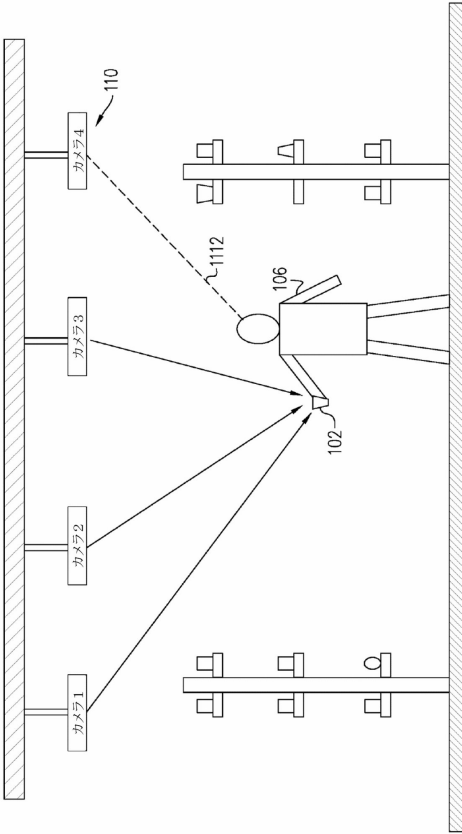
【図 1 3 A】



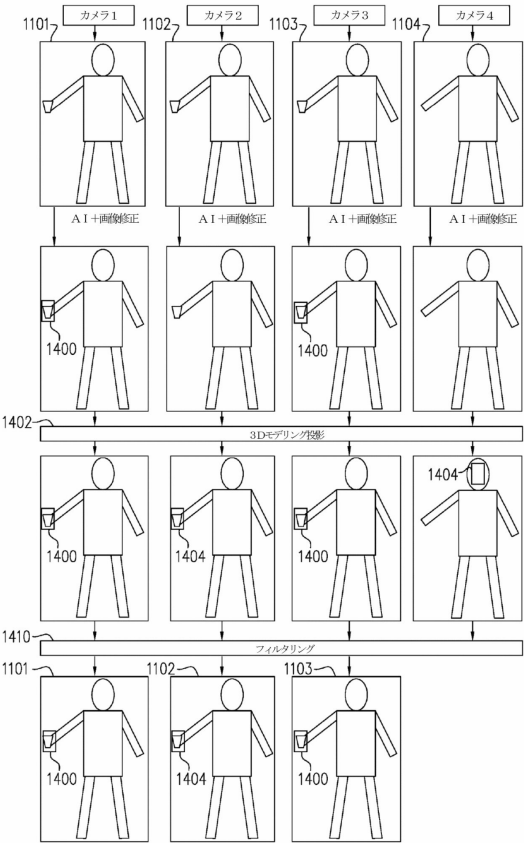
【図 1 3 B】



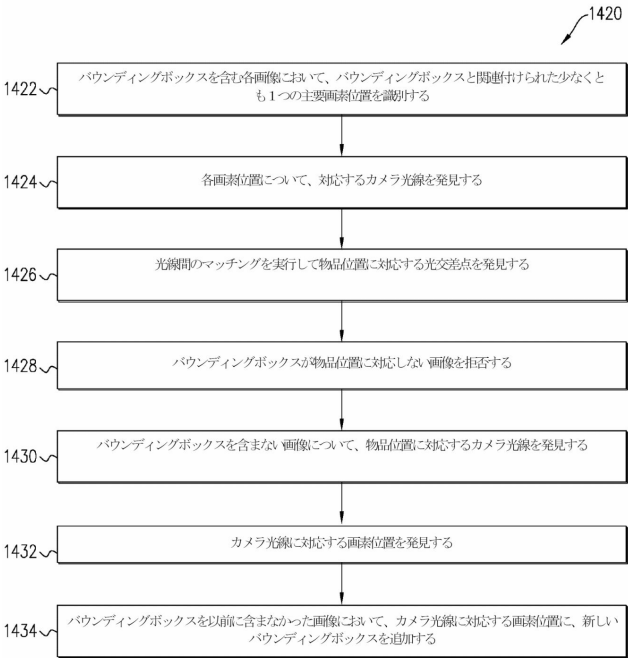
【図 1 4 A】



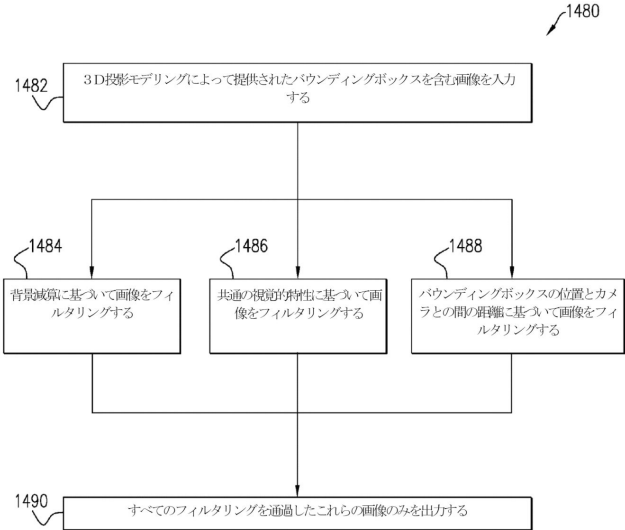
【図 1 4 B】



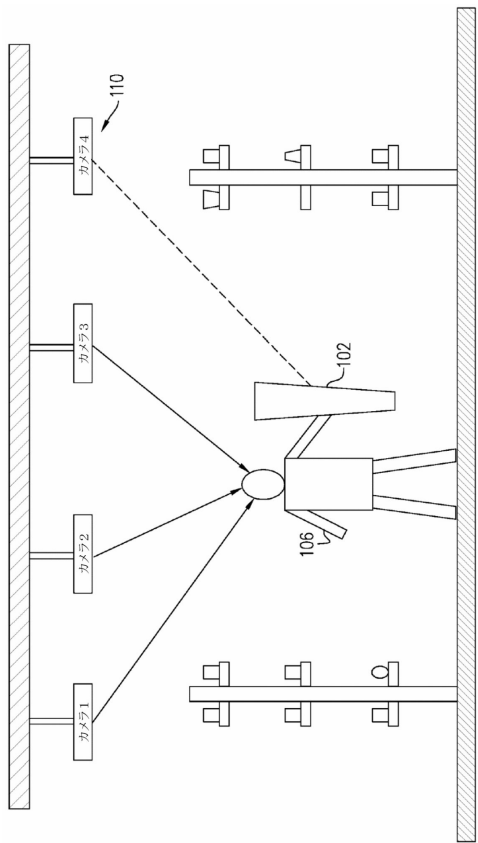
【図 1 4 C】



【図 1 4 D】



【図 1 5 A】



10

20

30

40

50

【図 15 B】

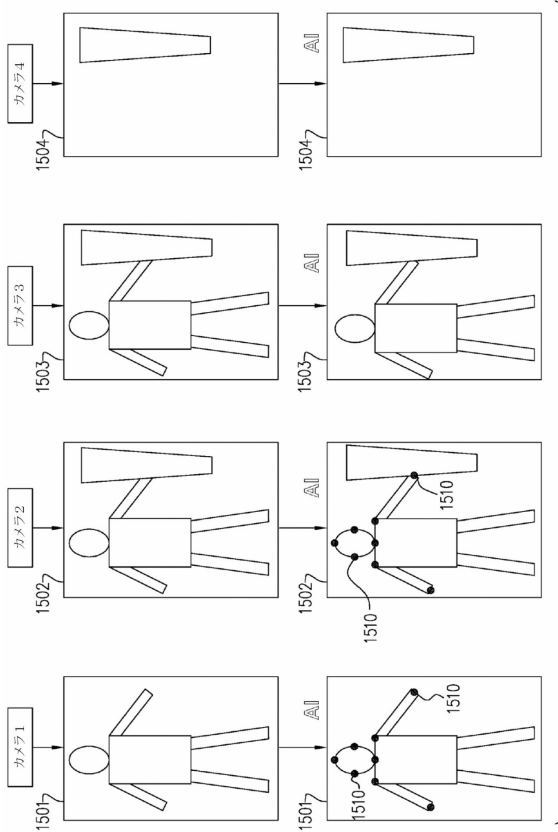
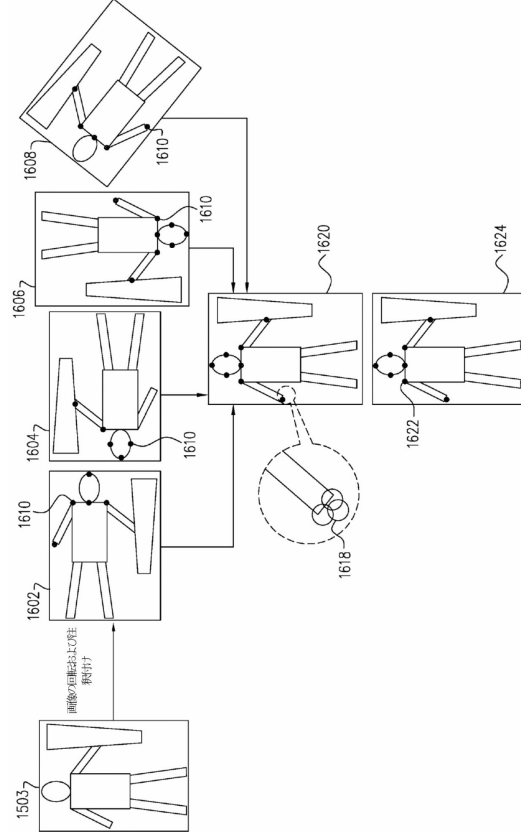


図15B

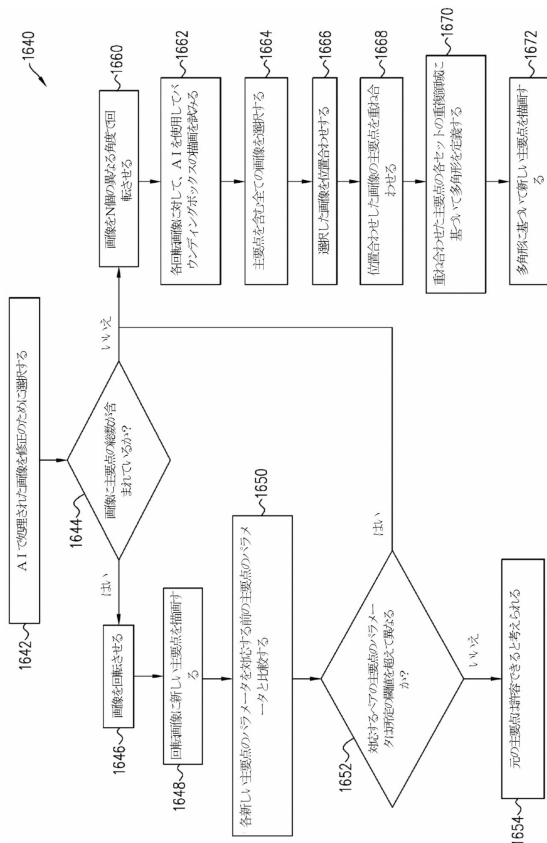
【図 16 A】



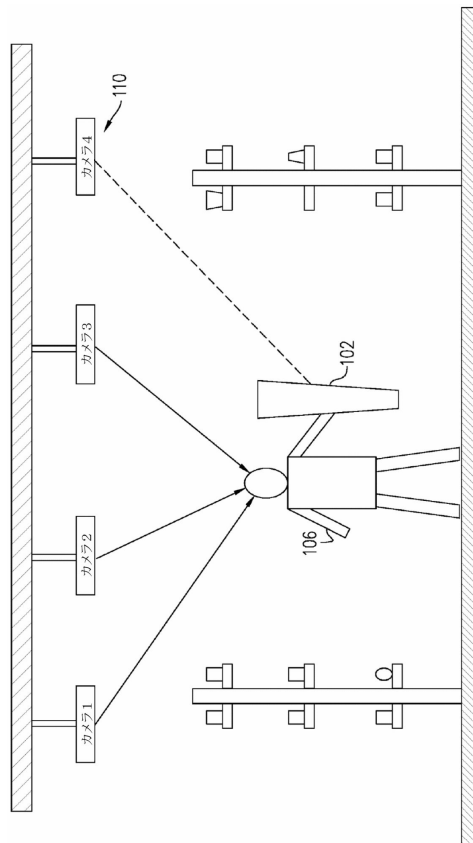
10

20

【図 16 B】



【図 17 A】

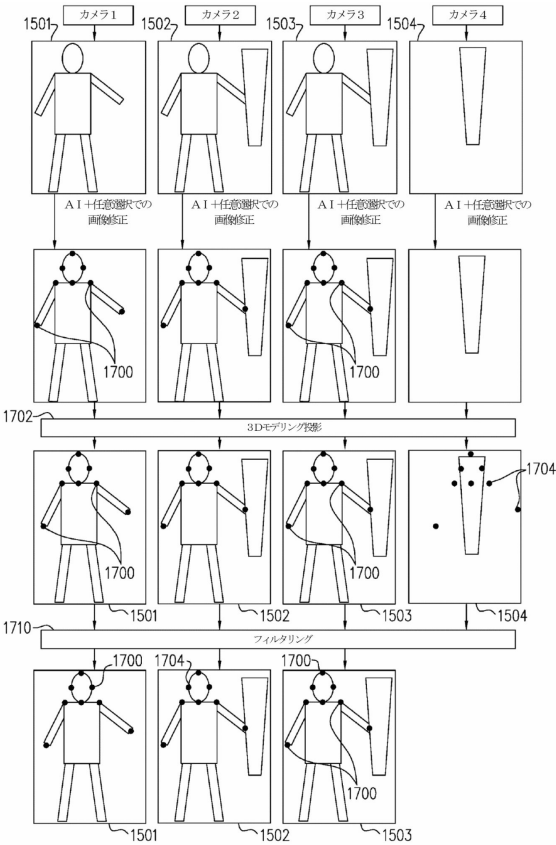


30

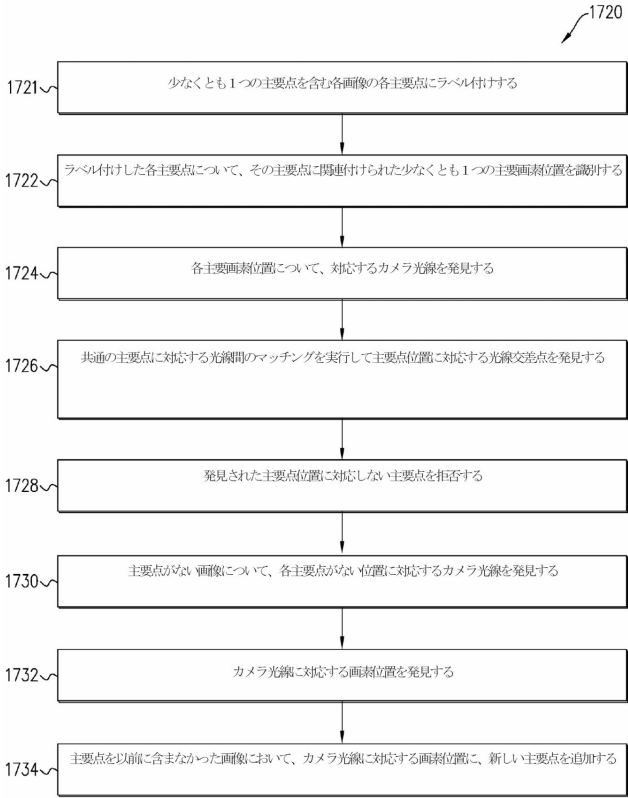
40

50

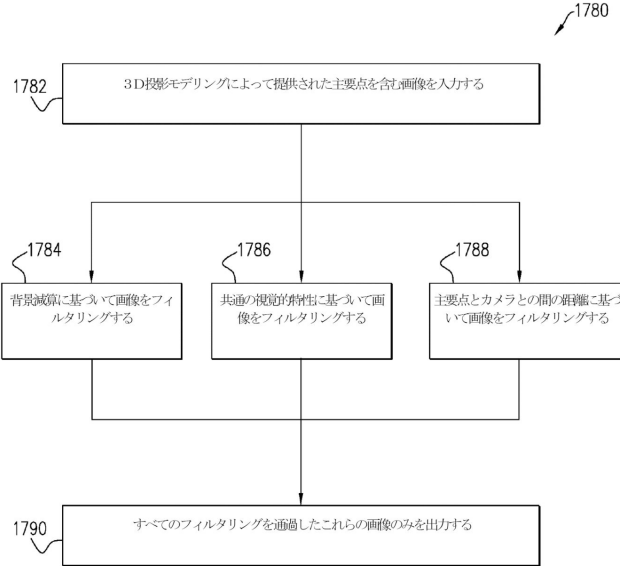
【図 17 B】



【図 17 C】



【図 17 D】



【図 18】

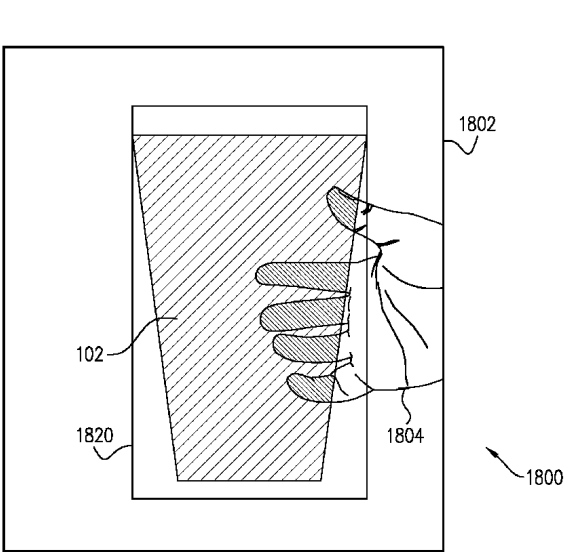


FIG. 18

10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IL2020/051198

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - G06K 9/00; H04N 7/18; G06K 9/36; H04N 17/00; H04N 7/00 (2021.01)

CPC - G06K 9/00; H04N 13/246; H04N 13/243; H04N 7/181; H04N 17/002 (2021.02)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

see Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

see Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

see Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2010/0177936 A1 (EBLING et al) 15 July 2010 (15.07.2010) entire document	1-10, 51
A	US 2014/0253738 A1 (PELICAN IMAGING CORPORATION) 11 September 2014 (11.09.2014) entire document	1-10, 51
A	US 2013/0141547 A1 (FUJITSU LIMITED et al) 06 June 2013 (06.06.2013) entire document	1-10, 51
A	US 2011/0134253 A1 (AOYAMA) 09 June 2011 (09.06.2011) entire document	1-10, 51
A	US 2009/0303077 A1 (ONOME et al) 10 December 2009 (10.12.2009) entire document	1-10, 51

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

21 April 2021

Date of mailing of the international search report

MAY 11 2021

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450
Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

Telephone No. PCT Helpdesk: 571-272-4300

Form PCT/ISA/210 (second sheet) (July 2019)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IL2020/051198

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 18-20, 28-30, 55
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet(s).

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10, 51

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IL2020/051198

Continued from Box No. III Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-10, 51, is drawn to a method for acquiring data relating to an object comprising finding a point of intersection of light rays illuminating each of said plurality of cameras.

Group II, claims 11-17, 21-27, 51-52, is drawn to a method for acquiring data relating to an object comprising: finding an identity of at least one object of said multiple objects appearing at a location within at least some of said plurality of images.

Group III, claims 31-38, 51-53, is drawn to a method for processing data relating to an object comprising: converting at least one characteristic of a plurality of images from a first characteristic to a second characteristic.

Group IV, claims 39-40, 51-54, is drawn to a method for processing data relating an object comprising: creating a multi-dimensional model of at least one object in a scene over time.

Group V, claims 41-50, 51-54, is drawn to a method for processing visual data relating to an object, the method comprising: employing artificial intelligence for: ascertaining whether at least one output from said at least one camera does or does not contain information indicating the presence and location of at least one object in said scene.

The inventions listed as Groups I, II, III, IV, or V do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical feature of the Group I invention: finding a point of intersection of light rays illuminating each of said plurality of cameras; and correlating a pixel location at which said reference object appears within each said at least one image to said light rays illuminating each of said plurality of cameras and intersecting with said region of intersection, irrespective of a three-dimensional location of said reference object within said scene as claimed therein is not present in the invention of Groups II, III, IV, or V. The special technical feature of the Group II invention: finding an identity of at least one object of said multiple objects appearing at a location within at least some of said plurality of images; selecting ones of said plurality of images as showing a common object and rejecting other ones of said plurality of images as not showing said common object, said selecting and said rejecting being based on prior inter-calibration of said multiplicity of cameras; outputting a selected set of images showing said common object at said location; and finding an identity of said common object shown in said selected set of images, based on taking into account at least one of said identity and location of said at least one object as claimed therein is not present in the invention of Groups I, III, IV, or V. The special technical feature of the Group III invention: converting at least one characteristic of a plurality of images from a first characteristic to a second characteristic; automatically finding at least one feature of at least one object appearing in at least some of said plurality of images having said second characteristic; and modifying said images having said first characteristic based on said at least one feature as claimed therein is not present in the invention of Groups I, II, IV, or V. The special technical feature of the Group IV invention: creating a multi-dimensional model of at least one object in a scene over time; tracking changes in said multi-dimensional model of said at least one object over time; identifying one or more of said changes satisfying at least one predetermined criterion; and analyzing data relating to said changes satisfying said at least one predetermined criterion in order to derive information relating to said object as claimed therein is not present in the invention of Groups I, II, III, or V. The special technical feature of the Group V invention: employing artificial intelligence for: ascertaining whether at least one output from said at least one camera does or does not contain information indicating the presence and location of at least one object in said scene, and enhancing at least one output that does not contain information indicating said presence and location of said at least one object in said scene by adding thereto information indicating said presence and location of said at least one object in said scene as claimed therein is not present in the invention of Groups I, II, III, or IV.

Groups I, II, III, IV, and V lack unity of invention because even though the inventions of these groups require the technical feature of a method for acquiring data relating to an object comprising: acquiring, by a multiplicity of cameras, at least one image of a scene, this technical feature is not a special technical feature as it does not make a contribution over the prior art.

Specifically, US 2019/0177936 to Ebling et al. teaches a method for acquiring data relating to an object comprising: acquiring, by a multiplicity of cameras, at least one image of a scene (Paras. [0006-0007]).

Since none of the special technical features of the Group I, II, III, IV or V inventions are found in more than one of the inventions, unity of invention is lacking.

フロントページの続き

,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,D
K,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),O
A(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,B
B,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD
,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,IT,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,
LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,
RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,Z
W

- (72)発明者 ガバイ, ダニエル
イスラエル国, エベン イェフダ 4050307, 37 ハーツマウト ストリート
- (72)発明者 スナッパー, ユバル
イスラエル国, テル アビブ 6404518, アパートメント 7, 14 レンブラント ストリート
- (72)発明者 ドレフ, シャケッド
イスラエル国, 6248410 テル アビブ, 10 ナフム ソコロフ ストリート
- (72)発明者 ケレット, シバン
イスラエル国, ジバタイム, 4 アフドウト ハアヴォダ ストリート
- (72)発明者 ベン ツビ, ニール
イスラエル国, 6618911 テル アビブ, 17 ハハッシマル
- (72)発明者 パベルマン カシャーニ, リタ
イスラエル国, 8502500 メイタール, 9 オルバニ
- F ターム (参考) 5C054 CA04 CC02 CF06 FC12 FD01 HA19
5L096 AA06 BA02 CA05 DA02 FA69 HA05 HA09