

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7366349号
(P7366349)

(45)発行日 令和5年10月23日(2023.10.23)

(24)登録日 令和5年10月13日(2023.10.13)

(51)国際特許分類		F I	
H 0 4 N	5/222(2006.01)	H 0 4 N	5/222 1 0 0
G 0 5 D	1/00 (2006.01)	G 0 5 D	1/00 B
G 0 5 D	1/10 (2006.01)	G 0 5 D	1/10
G 0 6 T	7/55 (2017.01)	G 0 6 T	7/55
G 0 6 T	7/70 (2017.01)	G 0 6 T	7/70
請求項の数 17 (全16頁) 最終頁に続く			
(21)出願番号	特願2022-539072(P2022-539072)	(73)特許権者	000002185
(86)(22)出願日	令和3年6月25日(2021.6.25)		ソニーグループ株式会社
(65)公表番号	特表2023-508414(P2023-508414		東京都港区港南1丁目7番1号
	A)	(74)代理人	100092093
(43)公表日	令和5年3月2日(2023.3.2)		弁理士 辻居 幸一
(86)国際出願番号	PCT/US2021/039151	(74)代理人	100109070
(87)国際公開番号	WO2022/005901		弁理士 須田 洋之
(87)国際公開日	令和4年1月6日(2022.1.6)	(74)代理人	100067013
審査請求日	令和4年6月23日(2022.6.23)		弁理士 大塚 文昭
(31)優先権主張番号	16/917,013	(74)代理人	100109335
(32)優先日	令和2年6月30日(2020.6.30)		弁理士 上杉 浩
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100120525
			弁理士 近藤 直樹
		(72)発明者	リウ チェンイー
			アメリカ合衆国 カリフォルニア州 9 5
			最終頁に続く

(54)【発明の名称】 マルチドローンビジュアルコンテンツ取り込みシステム

(57)【特許請求の範囲】
【請求項1】

シーンを画像化するシステムであって、
各ドローンが前記シーン上の対応する飛行経路に沿って移動するとともに、対応する第1の姿勢及び対応する第1の時点において前記シーンの対応する第1の画像を取り込むドローンカメラを有する、複数のドローンと、
カメラコントローラによって提供された各ドローンカメラの第1の姿勢の推定値を部分的に使用することによって、前記シーン上で所望のカメラ姿勢を有するドローンの所望のパターンを作成して維持するように各ドローンの前記飛行経路を制御するフライコントローラと、
前記複数のドローンから前記シーンの対応する複数の取り込み画像を受け取り、前記受け取った複数の取り込み画像を処理して前記シーンの3D表現をシステム出力として生成し、各ドローンカメラの前記第1の姿勢の前記推定値を前記フライコントローラに提供する前記カメラコントローラと、
を備え、前記システムは、わずか1人の人間オペレータと共に完全に動作可能であり、
前記カメラコントローラは、
各ドローンエージェントが対応する1つのドローンのみに通信可能に結合されて対応する取り込まれた第1の画像を受け取る、複数のドローンエージェントと、
前記ドローンエージェントの各々及び前記フライコントローラに通信可能に結合されたグローバルオブティマイザと、

を含み、

前記カメラコントローラ内の前記ドローンエージェント及び前記グローバルオブティマイザは、各ドローンについて、第1の姿勢の推定値と、前記対応するドローンカメラによって画像化された前記シーンを特性化する深度マップとを反復的に改善して、全ての前記ドローンからの前記推定値及び深度マップを使用して前記シーンの前記3D表現を作成するように協働し、

前記フライコントローラは、前記カメラコントローラから前記ドローンカメラの各々の前記第1の姿勢の推定値を受け取り、これに応じて必要な場合に前記対応する飛行経路及びドローンカメラ姿勢を調整する、

ことを特徴とするシステム。

10

【請求項2】

各ドローンに対応する前記深度マップは、前記第1の画像と、対応する第2の姿勢及び対応する第2の時点において対応するドローンカメラによって取り込まれて対応するドローンエージェントによって受け取られた前記シーンの第2の画像とを処理することに基づいて、前記対応するドローンエージェントによって生成される、

請求項1に記載のシステム。

【請求項3】

各ドローンに対応する前記深度マップは、前記第1の画像と、前記対応するドローン内の深度センサによって生成された深度データとを処理することに基づいて、対応するドローンエージェントによって生成される、

請求項1に記載のシステム。

20

【請求項4】

各ドローンエージェントは、

前記対応するドローンによって取り込まれた前記第1の画像が、前記対応するドローンを特性化するデータ及び画像取り込みパラメータを使用して処理されて前記対応するドローンの前記第1の姿勢の推定値を生成するように、他の1つのドローンエージェントと協働し、

前記ドローンエージェントが結合されている前記ドローンの前記ドローンカメラの前記第1の姿勢推定値を反復的に改善して前記対応する深度マップを反復的に改善するように前記グローバルオブティマイザと協働する、

請求項1に記載のシステム。

30

【請求項5】

各ドローンカメラの前記第1の姿勢の推定値を生成することは、各ドローンに固有のローカル座標系で表される姿勢関連データを前記複数のドローンによって共有されるグローバル座標系に変換することを含み、前記変換は、同時位置及びマッピング (Simultaneous Location and Mapping: SLAM) と多視点三角測量 (Multiview Triangulation: MT) との組み合わせを含む、

請求項4に記載のシステム。

【請求項6】

前記グローバルオブティマイザは、

前記複数のドローンエージェントの各々からの、前記対応するドローンを特性化するデータと、前記対応する処理済みの第1の画像、第1の姿勢推定値及び深度マップとを含む入力に基づいて、前記シーンの前記3D表現を生成して反復的に改善し、

前記複数のドローンの前記ドローンカメラの前記姿勢推定値を前記フライコントローラに提供する、

請求項1に記載のシステム。

40

【請求項7】

前記グローバルオブティマイザによって実行される前記反復的改善は、前記シーンの前記3D表現が所定の品質閾値を満たすまでドローンカメラ姿勢推定値及び深度マップを継続的かつ反復的に改善するループプロセスを含む、

50

請求項 6 に記載のシステム。

【請求項 8】

シーンを画像化する方法であって、

各ドローンが前記シーン上の対応する飛行経路に沿って移動するとともに、対応する第 1 の姿勢及び対応する第 1 の時点において前記シーンの対応する第 1 の画像を取り込むカメラを有する、複数のドローンを展開するステップと、

フライコントローラを使用して、カメラコントローラによって提供された各カメラの第 1 の姿勢の推定値を部分的に使用することによって、前記シーン上で所望のカメラ姿勢を有するドローンの所望のパターンを作成して維持するように各ドローンの前記飛行経路を制御するステップと、

カメラコントローラを使用して、前記複数のドローンから前記シーンの対応する複数の取り込み画像を受け取り、前記受け取った複数の取り込み画像を処理して前記シーンの 3 D 表現をシステム出力として生成し、各カメラの前記第 1 の姿勢の前記推定値を前記フライコントローラに提供するステップと、

を含み、前記方法の完全な動作のためにわずか 1 人の人間オペレータしか必要でなく、

前記カメラコントローラは、

各ドローンエージェントが対応する 1 つのドローンのみに通信可能に結合されて対応する取り込まれた第 1 の画像を受け取る、複数のドローンエージェントと、

前記ドローンエージェントの各々及び前記フライコントローラに通信可能に結合されたグローバルオブティマイザと、

を含み、

前記カメラコントローラ内の前記ドローンエージェント及び前記グローバルオブティマイザは、各ドローンについて、前記第 1 の姿勢の推定値と、前記対応するドローンカメラによって画像化された前記シーンを特性化する深度マップとを反復的に改善して、全ての前記ドローンからの前記推定値及び深度マップを使用して前記シーンの前記 3 D 表現を作成するように協働し、

前記フライコントローラは、前記カメラコントローラから前記ドローンカメラの各々の前記改善された第 1 の姿勢の推定値を受け取り、これに応じて必要な場合に前記対応する飛行経路及びドローンカメラ姿勢を調整する、

ことを特徴とする方法。

【請求項 9】

各ドローンに対応する前記深度マップは、前記第 1 の画像と、対応する第 2 の姿勢及び対応する第 2 の時点において対応するドローンカメラによって取り込まれて対応するドローンエージェントによって受け取られた前記シーンの第 2 の画像とを処理することに基づいて、前記対応するドローンエージェントによって生成される、

請求項 8 に記載の方法。

【請求項 10】

各ドローンに対応する前記深度マップは、前記第 1 の画像と、対応するドローン内の深度センサによって生成された深度データとを処理することに基づいて、対応するドローンエージェントによって生成される、

請求項 8 に記載の方法。

【請求項 11】

前記協働は、

各ドローンエージェントが、前記対応するドローンによって取り込まれた前記第 1 の画像を前記ドローンを特性化するデータ及び前記対応する取り込み画像の画像取り込みパラメータを使用して処理して前記対応するドローンの前記第 1 の姿勢の推定値を生成するように他の 1 つのドローンエージェントと協働し、

各ドローンエージェントが、前記ドローンエージェントが結合されている前記ドローンの前記ドローンカメラの前記第 1 の姿勢推定値を反復的に改善して前記対応する深度マップを反復的に改善するように前記グローバルオブティマイザと協働する、

10

20

30

40

50

ことを含む、請求項 8 に記載の方法。

【請求項 1 2】

各ドローンカメラの前記第 1 の姿勢の推定値を生成するステップは、各ドローンに固有のローカル座標系で表される姿勢関連データを前記複数のドローンによって共有されるグローバル座標系に変換するステップを含み、前記変換は、同時位置及びマッピング (Simultaneous Location and Mapping: SLAM) と多視点三角測量 (Multiview Triangulation: MT) との組み合わせを含む、

請求項 1 1 に記載の方法。

【請求項 1 3】

前記グローバルオブティマイザは、

前記複数のドローンエージェントの各々からの、前記対応するドローンを特性化するデータと、前記対応する処理済みの第 1 の画像、第 1 の姿勢推定値及び深度マップとを含む入力に基づいて、前記シーンの前記 3 D 表現を生成して反復的に改善し、

前記複数のドローンカメラの前記第 1 の姿勢推定値を前記フライコントローラに提供する、

請求項 9 に記載の方法。

【請求項 1 4】

前記グローバルオブティマイザによって実行される前記反復的改善は、前記シーンの前記 3 D 表現が所定の品質閾値を満たすまでドローンカメラ姿勢推定値及び深度マップを継続的かつ反復的に改善するループプロセスを含む、

請求項 1 3 に記載の方法。

【請求項 1 5】

前記協働の前に、部分的に、

前記複数のドローンのカメラの各々からの電気信号又は視覚信号を比較して時間的同期を可能にし、

各ドローンについて SLAM プロセスを実行して各ドローンのローカル座標系を確立し、多視点三角測量プロセスを実行して前記複数のドローンによって共有されるグローバル座標フレームワークを定める、

ことによって前記複数のドローン間の時間的及び空間的關係を確立するステップをさらに含む、

請求項 8 に記載の方法。

【請求項 1 6】

1 又は 2 以上のプロセッサと、

前記 1 又は 2 以上のプロセッサによる実行のための、1 又は 2 以上の非一時的媒体に符号化されたロジックと、

を備えた装置であって、前記ロジックは、実行されると、

各ドローンがシーン上の対応する飛行経路に沿って移動するとともに、対応する第 1 の姿勢及び対応する第 1 の時点において前記シーンの対応する第 1 の画像を取り込むカメラを有する複数のドローンを展開し、

フライコントローラを使用して、カメラコントローラによって提供された各カメラの第 1 の姿勢の推定値を部分的に使用することによって、前記シーン上で所望のカメラ姿勢を有するドローンの所望のパターンを作成して維持するように各ドローンの前記飛行経路を制御し、

カメラコントローラを使用して、前記複数のドローンから前記シーンの対応する複数の取り込み画像を受け取り、前記受け取った複数の取り込み画像を処理して前記シーンの 3 D 表現をシステム出力として生成し、各カメラの前記第 1 の姿勢の前記推定値を前記フライコントローラに提供する、

ことによって前記シーンを画像化するように動作可能であり、

前記装置の完全な動作のためにわずか 1 人の人間オペレータしか必要でなく、

10

20

30

40

50

前記カメラコントローラは、

各ドローンエージェントが対応する1つのドローンのみに通信可能に結合されて前記対応する取り込まれた第1の画像を受け取る、複数のドローンエージェントと、

前記ドローンエージェントの各々及び前記フライコントローラに通信可能に結合されたグローバルオブティマイザと、

を含み、

前記カメラコントローラ内の前記ドローンエージェント及び前記グローバルオブティマイザは、各ドローンについて、前記第1の姿勢の推定値と、前記対応するドローンカメラによって画像化された前記シーンを特性化する深度マップとを反復的に改善して、全ての前記ドローンからの前記推定値及び深度マップを使用して前記シーンの前記3D表現を作成するように協働し、

前記フライコントローラは、前記カメラコントローラから前記ドローンカメラの各々の前記改善された第1の姿勢の推定値を受け取り、これに応じて必要な場合に前記対応する飛行経路及びドローンカメラ姿勢を調整する、
ことを特徴とする装置。

【請求項17】

各ドローンに対応する前記深度マップは、

前記第1の画像と、対応する第2の姿勢及び対応する第2の時点において対応するドローンカメラによって取り込まれて対応するドローンエージェントによって受け取られた前記シーンの第2の画像とを処理すること、又は、

前記第1の画像と、前記対応するドローン内の深度センサによって生成された深度データとを処理すること、

のいずれかに基づいて、前記対応するドローンエージェントによって生成される、

請求項16に記載の装置。

【発明の詳細な説明】

【背景技術】

【0001】

カメラを備えたドローンの可用性が高まるにつれ、これまでアクセスが困難であったシーンの画像を取り込むことに基づく新たなスタイルの映画撮影術が生まれた。従来、専門家らは、外部パラメータを十分に制御した正確なカメラ軌道を使用することによって高品質画像を撮影していたが、ドローン上のカメラは、ドローンがホバリングしている時であっても常に動いている。この理由は、ドローンの空気力学的特性によって連続的な動き変動が避けられないからである。関与するドローンが1機のみであれば、ロボット工学分野で周知の自己位置推定及びマッピングの同時実行 (Simultaneous Localization and Mapping: SLAM) によってカメラ姿勢 (位置及び向き) の6D的組み合わせを推定することが依然として可能である。しかしながら、多くの場合は、異なるビュースポットで同時に複数のカメラを使用して複雑な編集及び完全な3Dシーン再構成を可能にすることが望ましい。従来のSLAM手法は、単一ドローン、単一カメラの状況ではうまく機能するが、複数ドローン、複数カメラの状況にかかわる全ての姿勢の推定には適していない。

【0002】

マルチドローン撮影における他の課題としては、複数のドローンによって取り込まれた画像のビデオストリームを統合する複雑さ、並びに所望のフォーメーション (又は群れパターン) 及びこのフォーメーションのいずれかの経時的な望ましい変化を達成できるように全てのドローンの飛行経路を制御する必要性が挙げられる。ドローンを伴う専門的な撮影のための現行実務では、飛行パラメータを制御するコントローラとカメラ姿勢を制御するコントローラという各ドローン上の2つの独立したコントローラを人間のオペレータが操作する必要がある。ドローンのサイズ、重量及びコスト、全体的なシステムの信頼性、並びに出力されるシーン再構成の品質に関する多くのマイナス要素が存在する。

【発明の概要】

【発明が解決しようとする課題】

【0003】

従って、関心シーンを再構成するのに必要なビジュアルコンテンツを効率的に取り込んで処理できるように、複数の移動中のドローン上のカメラによって取り込まれた画像を統合してこれらのドローンを（また場合によってはドローンとは無関係にカメラを）正確に制御する改善されたシステム及び方法が必要とされている。ビジュアルコンテンツの統合は、ドローンから離れた場所で自動的に行われ、制御は、やはりドローンから離れた場所であるが必ずしも同じである必要はない場所で行われ、風などの要因による空気力学的ノイズに適応した高精度なドローンの位置決めを達成するように自動フィードバック制御機構を伴うことが理想的である。また、時にはシステム運用に必要な人間オペレータの数を最小化することも有益となり得る。

10

【課題を解決するための手段】

【0004】

実施形態は、一般に複数のドローンによって取り込まれた画像に基づいてシーンを3Dで画像化する方法及びシステムに関する。

【0005】

1つの実施形態では、システムが、複数のドローンと、フライコントローラと、カメラコントローラとを含み、システムは、わずか1人の人間オペレータと共に完全に動作可能である。各ドローンは、シーン上の対応する飛行経路に沿って移動するとともに、対応する第1の姿勢及び対応する第1時点においてシーンの対応する第1の画像を取り込むドローンカメラを有する。フライコントローラは、カメラコントローラによって提供された各ドローンカメラの第1の姿勢の推定値を部分的に使用することによって、シーン上で所望のカメラ姿勢を有するドローンの所望のパターンを作成して維持するように各ドローンの飛行経路を制御する。カメラコントローラは、複数のドローンからシーンの対応する複数の取り込み画像を受け取り、受け取った画像を処理してシーンの3D表現をシステム出力として生成し、各ドローンカメラの第1の姿勢の推定値をフライコントローラに提供する。

20

【0006】

別の実施形態では、シーンを画像化する方法が、各ドローンがシーン上の対応する飛行経路に沿って移動するとともに、対応する第1の姿勢及び対応する第1の時点においてシーンの対応する第1の画像を取り込むカメラを有する複数のドローンを展開するステップと、フライコントローラを使用して、カメラコントローラによって提供された各カメラの姿勢の推定値を部分的に使用することによって、シーン上で所望のカメラ姿勢を有するドローンの所望のパターンを作成して維持するように各ドローンの飛行経路を制御するステップと、カメラコントローラを使用して、複数のドローンからシーンの対応する複数の取り込み画像を受け取り、受け取った複数の取り込み画像を処理してシーンの3D表現をシステム出力として生成し、各カメラの第1の姿勢の推定値をフライコントローラに提供するステップとを含む。方法の完全な動作のためには、わずか1人の人間オペレータしか必要でない。

30

【0007】

別の実施形態では、装置が、1又は2以上のプロセッサと、1又は2以上のプロセッサによる実行のための、1又は2以上の非一時的媒体に符号化されたロジックとを含む。ロジックは、実行されると、各ドローンがシーン上の対応する飛行経路に沿って移動するとともに、対応する第1の姿勢及び対応する第1の時点においてシーンの対応する第1の画像を取り込むカメラを有する複数のドローンを展開し、フライコントローラを使用して、カメラコントローラによって提供される各カメラの姿勢の推定値を部分的に使用することによって、シーン上で所望のカメラ姿勢を有するドローンの所望のパターンを作成して維持するように各ドローンの飛行経路を制御し、カメラコントローラを使用して、複数のドローンからシーンの対応する複数の取り込み画像を受け取り、受け取った画像を処理してシーンの3D表現をシステム出力として生成し、各カメラの姿勢の推定値をフライコントローラに提供することによってシーンを画像化するように動作可能である。シーンを画像

40

50

化する装置の完全な動作のためには、わずか1人の人間オペレータしか必要でない。

【0008】

本明細書の残り部分及び添付図面を参照することにより、本明細書に開示する特定の実施形態の性質及び利点のさらなる理解を実現することができる。

【図面の簡単な説明】

【0009】

【図1】いくつかの実施形態によるシーンの画像化を示す図である。

【図2】図1の実施形態によるシーンの画像化を示す図である。

【図3】いくつかの実施形態による、ドローンエージェントがどのように機能できるかについての例を示す図である。

【図4】いくつかの実施形態による、一対のドローンカメラ間の変換計算の概要を示す図である。

【図5】いくつかの実施形態による、2つのカメラ位置間の複数のベクトルの交差を推定するために適用される最小二乗法の数学的詳細を示す図である。

【図6】いくつかの実施形態による、2つのカメラについてスケーリングの初期解をどのように達成できるかを示す図である。

【図7】いくつかの実施形態による、2つのカメラの座標間の初期回転をどのように計算できるかを示す図である。

【図8】いくつかの実施形態による、2つのカメラの座標（位置、回転及びスケーリング）を完全に整合させる計算の最終ステップを要約した図である。

【図9】いくつかの実施形態による、ドローンエージェントがどのように深度マップを生成するかを示す図である。

【図10】いくつかの実施形態における、フライコントローラとカメラコントローラとの間の相互作用を示す図である。

【図11】いくつかの実施形態による、ドローン群の飛行及び姿勢制御がどのように達成されるかを示す図である。

【図12】いくつかの実施形態における、システムのコンポーネント間の高水準データフローを示す図である。

【発明を実施するための形態】

【0010】

図1に、本発明のいくつかの実施形態による、シーン120を画像化するシステム100を示す。図2には、システム100のコンポーネントを異なる詳細レベルで示す。複数のドローンを示しており、各ドローン105は対応する経路110に沿って移動する。図1には、人間160によって操作される、各ドローンと無線通信しているフライコントローラ130を示す。ドローンは、カメラコントローラ140とも無線通信して取り込み画像をカメラコントローラ140に送信する。カメラコントローラ140からは、飛行制御を容易にするデータがフライコントローラ130に送信される。任意に、フライコントローラ130からは、カメラコントローラ140内の画像処理を容易にする他のデータがカメラコントローラ140に送信される。システム出力は、シーン120の3D再構成150の形態で提供される。

【0011】

図2には、複数のドローンエージェント142及びグローバルオプティマイザ144を含むカメラコントローラ140の内部機構の一部と、システムのコンポーネント間のフィードバックループを含むデータの流れとを示す。シーン120及びシーン再構成150については、単純化のために図1よりも抽象的に表している。

【0012】

各ドローンエージェント142は、1つのドローンのみと「マッチアップ」し、そのドローン105内の又はそのドローン105に取り付けられたドローンカメラ115から画像を受け取る。単純化のために、図2には様々なドローン上に同じ相対的位置及び向きのドローンカメラを示しているが、実際には必ずしもそうであるとは限らない。各ドローン

10

20

30

40

50

エージェントは、対応するドローンカメラから（場合によっては、フライコントローラ 130 から受け取られたフライコマンド情報と組み合わせて）受け取られた各画像（又はビデオストリームからのフレーム）を、ドローン、ドローンカメラ及び取り込み画像を特性化するデータと共に処理して、そのドローンに固有の座標フレームにおける、本開示では 3D 位置と 3D 定位との組み合わせとして定められるドローンカメラ姿勢の推定値を（例えば、SLAM 技術を使用して）生成する。上述した特性データは、典型的にはドローン ID、内部カメラパラメータ、並びに画像タイムスタンプ、サイズ、符号化及び取り込みレート（fps）などの画像取り込みパラメータを含む。

【0013】

次に、各ドローンエージェントは、推定されるカメラ姿勢をグローバル座標系で表して各ドローンが共有できるように、少なくとも 1 つの他のドローンエージェントと協働して自機のドローンカメラに固有の座標変換を計算する。この計算は、以下で図 3 及び図 4 を参照しながらさらに詳細に説明する新規のロバスト座標整列アルゴリズム（robust coordinate aligning algorithm）を使用して実行することができる。

【0014】

各ドローンエージェントは、対応するドローンカメラが観察するようなシーン 120 の、対応する画像が取り込まれた各姿勢についての高密度深度¹マップも生成する。深度マップは、グローバル座標系で計算され表現される。いくつかの事例では、このマップが、同じドローンカメラからわずかに異なる時点及び姿勢で受け取られた、立体ペアとして機能できるほど十分に重なり合う視野を有する画像のペアをドローンエージェントが処理することによって生成される。後述する図 9 に示すように、ドローンエージェントは、このようなペアを処理して対応する深度マップを生成するために、周知の技術を使用することができる。他の事例では、深度測定値が RGB 画像画素とともに送信されて、深度マップを生成するためにドローンエージェントが処理する（単純な RGB 画像ではなく）RGBD 画像が形成されるように、ドローンが何らかのタイプの深度センサを含む。さらに他の事例では、両オプションを提示して、深度センサからの情報を、立体ペア処理によって既に生成されている深度マップを改善するための補助として使用することもできる。内蔵型深度センサの例としては、LiDAR システム、飛行時間型、及び立体カメラによって提供される深度センサが挙げられる。

（1. 本明細書で使用する「高密度」という単語は、深度マップの解像度がその導出元である RGB 画像の解像度に等しく又は非常に近いことを意味する。一般に、LiDAR 又は RGB-D のようなモダリティは、RGB よりもはるかに低い（VGA よりも低い）解像度で深度マップを生成する。ビジュアルキーポイントベースの方法は、深さを有するさらにまばらな点を生成する。）

【0015】

各ドローンエージェントは、対応するドローンを本質的に特性化するデータと共に、ドローンカメラ姿勢の推定値及び対応する深度マップをいずれもグローバル座標でグローバルオプティマイザ 144 に送信する。グローバルオプティマイザ 144 は、各ドローンエージェントからこれらの全てのデータ及び RGB 画像を受け取るとこれらのデータをまとめて処理し、より多くの画像及びデータが受け取られるにつれて時間と共に拡張、修正及び改善できる 3D 点群表現を生成する。3D 点群内に既に画像のキーポイントが存在して一致が確認された場合、このキーポイントは「登録」されたとみなされる。この処理の主な目的は、複数の画像にわたる 3D 点群画像データを検証し、これに対応して各ドローンカメラの推定姿勢及び深度マップを調整することである。このようにして、画像化されたシーン再構成の「構造」と、ドローンカメラの空間的及び時間的な「動き」又は位置との同時最適化を達成することができる。

【0016】

このグローバル最適化は、例えば異なる姿勢で取り込まれた複数の画像から 3D 点群再構成を生成するグラフベースのオプティマイザ Bundle Fusion などの、現在利

10

20

30

40

50

用可能な様々な最先端のSLAM又は動きからの構造 (Structure from Motion: SfM) オプティマイザのいずれか一方の使用に部分的に依存する。

【0017】

本発明では、このようなオプティマイザが、プロセスレベルの反復オプティマイザ (iterative optimizer) に埋め込まれて、各サイクル後に更新 (改善) されたカメラ姿勢推定値及び深度マップをフライコントローラに送信し、必要時にこのフライコントローラを使用して飛行経路及び姿勢の調整を行うことができる。その後、ドローンエージェントは、ドローンからドローンエージェントに送信された後続の画像を、各ドローンエージェントが少なくとも1つの他のドローンエージェントと協力することを伴って上述したように処理して、さらに改善された深度マップ及びドローンカメラ姿勢推定値をもたらす、これらがさらにグローバルオプティマイザに送信されて次の反復サイクルにおいて使用され、以下同様である。このように、カメラ姿勢推定値及び深度マップの精度はサイクル毎に改善され、さらにドローンの飛行経路の制御及び3D点群の再構築の品質も改善される。この再構成が所定の品質閾値を満たしたとみなされると反復サイクルは停止し、この時点での再構成を最終的なシステム出力として提供することができる。この出力の用途については、例えば映画撮影のための3Dシーン再構成又はビュー変更体験を含む多くのものを容易に想定することができる。

10

【0018】

次に、システム100に示すドローンエージェント142が様々な実施形態においてどのように動作するかについてのさらなる詳細を説明する。

20

【0019】

本発明では、複数のドローンカメラの位置付け及び動きを制御する方法の問題に、SLAMと多視点三角測量 (Multi View Triangulation: MVT) との組み合わせによって対処する。図3に、単独でとらえた2つの技術の長所及び短所を示しており、提案する組み合わせの1つの実施形態の詳細は、画像シーケンス (又はビデオ) が既に時間的に同期していると仮定した上で、最初に各ドローン上でSLAMプロセス (例えば: ORBSLAM2) を実行して各画像におけるローカルなドローンカメラ姿勢 (以下、ローカルSLAM姿勢) を生成し、次にドローン毎に少数 (例えば、5つ) のRGB画像フレーム及び対応するローカルSLAM姿勢をロードすることを伴う。これにより、そのドローンカメラの一貫した「ローカル」座標及び「ローカル」スケールが決定される。次に、複数のドローンについてロバストMVTアルゴリズムを実行し、図4に、第1のドローンのSLAMによって定められる座標に第2のドローンのローカルSLAM姿勢を整合させるために必要な変換 (回転、スケール、平行移動) をどのように計算できるかを概略的に示す。次に、これを複数のドローン内の他の各ドローンに拡張する。その後、各ローカルSLAM姿勢に適した変換を適用する。この結果、複数のドローンのカメラ全体から取り込まれた画像の空間的及び時間的一貫性が達成される。

30

【0020】

2つのカメラ間の変換を決定するために必要な様々な計算に伴うステップの数学的詳細については図5～図8に示す。

【0021】

図5には、カメラ位置間の複数のベクトルの交点を推定するためにどのように最小二乗法を適用できるかを示す。図6には、2つのカメラについてスケージングの初期解をどのように達成できるかを示す。図7には、2つのカメラの座標間の初期回転をどのように計算できるかを示す。計算された回転行列が不偏であることを保証するために、当業で周知の技術を使用して3つの回転自由度全てにわたって平均化を行う。図8には、2つのカメラの座標 (位置、回転及びスケージング) を完全に整合させる計算の最終ステップを要約する。

40

【0022】

単純化のために、ドローンエージェントのうちの1つを、「マスター」ドローンカメラを表す「マスター」ドローンエージェントとみなすことができ、その座標は、上述した技

50

術を使用して他の全てのドローンカメラ画像が整合するグローバル座標とみなすことができる。

【0023】

図9には、上述したような技術を使用して対応するカメラの画像をマスタードローンカメラに整合させ、その過程において対応するカメラ姿勢をおおまかに推定した後にドローンエージェントが実行できる内部機能ステップを概略的な形で示す。いくつかの実施形態によれば、図の下部の4つのブロックに表す姿勢推定後のステップは、例えば第1の画像及び第2の画像などの連続して取り込まれた画像の疑似立体ペアに基づいて深度マップを生成する。その後に実行される一連の動作は、(わずかに異なる時点でドローンカメラによって取り込まれた画像を比較する)画像修正(image rectification)、PSMnet、SGMなどの様々な周知のツールのいずれかを使用した深度推定、そして最後にペアの第1の画像の画素に計算された深度を割り当てる逆修正(Un-Rectification)である。

10

【0024】

図10には、システム100のいくつかの実施形態における、フライコントローラとカメラコントローラとの間の相互作用の高水準の態様を要約する。これらの相互作用は、フライコントローラとカメラコントローラとの間のフィードバックループの形態をとり、フライコントローラは、カメラコントローラによって測定された最新の視覚的姿勢を使用してその制御モデルを更新し、カメラコントローラは、フライコントローラによって送信されたコマンドをカメラ姿勢のSLAM計算において考慮する。

20

【0025】

図11には、2つのコントローラ間の継続的フィードバックに依存するという理由でフィードバック群制御(feedback swarm control)と呼ばれる、複数のドローンの飛行経路又は姿勢の制御を達成するための典型的なプロセスの詳細を示す。結果として得られる本発明システムの重要な側面は、以下のように列挙することができる。

【0026】

(1) 制御が、カメラ位置設定のための最新かつ最も正確な視覚的基準として機能するグローバルオブティマイザの3Dマップに基づく。

(2) フライコントローラが3Dマップ情報を使用して、マップ内で明らかになった位置誤差を補正するコマンドを各ドローンに対して生成する。

30

(3) ドローンから画像が到着すると、ドローンエージェントが予想位置の「周囲」の「測定」された位置を計算し始めることによって、実現しそうでない解決策を避けることができる。

(4) ドローン群フォーメーションについては、フィードバック機構が各ドローンの姿勢を絶えず目測によって調整し、ドリフトによるフォーメーション歪み(format ion distortion)が制限される。

【0027】

図12には、フライコントローラ130とカメラコントローラ140との間の、システム100のこれらの2つの主要コンポーネントを統合する「外部」制御フィードバックループ、及びグローバルオブティマイザ144と各ドローンエージェント142との間の「内部」フィードバックループを情報の流れにラベル付けして示す。カメラコントローラ140内のグローバルオブティマイザは、完全に最適化された姿勢データ(回転+位置)を観察チャネルとしてフライコントローラに提供し、フライコントローラは、その制御パラメータ推定においてこれらの観察結果を考慮し、従ってフライコントローラによって送信されるドローンコマンドは最新の姿勢不確実性に対応するようになる。フライコントローラは、外部フィードバックループを継続しながらカメラコントローラ内のドローンエージェント142との間で動作コマンドを共有する。これらのコマンドは、カメラコントローラ内部の次のカメラ姿勢計算の最適化の抑制及び加速を行うための事前情報である。この図では、グローバルオブティマイザ144と各ドローンエージェント142との間の内部フィードバックループを、これらのコンポーネント間の両頭矢印で示す。

40

50

【0028】

本明細書で説明する実施形態は、複数のカメラ付きドローンを使用してビジュアルコンテンツを取り込んで統合するシステム及び方法において様々な利益をもたらす。具体的には、実施形態は、提案する反復的グローバルオブティマイザによって容易になるような、カメラが取り込んだ画像のビジュアルコンテンツに純粹に基づくドローン軌道及びこれらのカメラ姿勢の自動空間整合又は調整、並びに全てのドローン間における一貫した3D点群、深度マップ及びカメラ姿勢の計算を可能にする。カメラコントローラ内の提案するSLAM-MT機構は、複数の（さらには2機よりもはるかに多くの）ドローンからの連続取り込み画像のビジュアルコンテンツを単純に使用して一貫したスケールのRGB-D画像データを生成できるため、正常な動作は深度センサの存在に依拠しない（ただしこれらは有用な補助であることができる）。このようなデータは、現代の高品質な3Dシーン再構築では非常に貴重である。

10

【0029】

上述した新規のローカル-グローバル座標変換法は、多対1のグローバルマッチ（global match）が行われるように複数の画像ペアをマッチさせることに基づき、これによってロバスト性をもたらす。対応するカメラ姿勢及び深度マップを計算するためにドローンエージェントによって実行される画像処理は、先行技術のシステムとは対照的にグローバル3Dマップの可用性に依存しない。各ドローンエージェントは、RGB画像及びその対応するカメラ姿勢の対を所与として単独で高密度の深度マップを生成し、これらの深度マップ及びカメラ姿勢をグローバル座標に変換した後に、結果をグローバルオブティマイザに提供することができる。従って、本発明のグローバルオブティマイザの動作は、統一座標系においてカメラ姿勢及び深度マップに対処する単純なものである。

20

【0030】

なお、データ転送には2つのループが関与する。外部ループは、フライコントローラとカメラコントローラとの間で動作してグローバルな位置決め精度を提供し、（複数のサブループで構成される）内部ループは、カメラコントローラ内のドローンエージェントとグローバルオブティマイザとの間で動作して構造及び動きの精度を提供する。

【0031】

特定の実施形態に関して説明を行ったが、これらの特定の実施形態は例示にすぎず、限定的なものではない。応用例としては、専門的な3Dシーンの取り込み、デジタルコンテンツ資産の生成、スタジオ撮影のためのリアルタイムレビューツール、ドローン群フォーメーション及び制御が挙げられる。さらに、本発明は、複雑な3D運動軌道を実行する複数のドローンを取り扱うことができるので、ロボットチームによるスキャンなどの低次元軌道のプロセス事例にも適用することができる。

30

【0032】

特定の実施形態のルーチンの実行には、C、C++、Java、アセンブリ言語などを含むあらゆる好適なプログラム言語を使用することができる。手続き型又はオブジェクト指向型などの異なるプログラミング法を使用することもできる。これらのルーチンは、単一の処理装置又は複数のプロセッサ上で実行することができる。ステップ、動作又は計算については特定の順序で示しているが、異なる特定の実施形態ではこの順序を変更することもできる。いくつかの特定の実施形態では、本明細書に逐次的なものとして示す複数のステップを同時に実行することもできる。

40

【0033】

特定の実施形態は、命令実行システム、装置、システム又はデバイスが使用する、或いはこれらに関連するコンピュータ可読記憶媒体に実装することができる。特定の実施形態は、ソフトウェア、ハードウェア又はこれらの組み合わせにおける制御ロジックの形で実装することができる。制御ロジックは、1又は2以上のプロセッサによって実行された時に、特定の実施形態において説明した内容を実行することができる。

【0034】

特定の実施形態は、特定用途向け集積回路、プログラマブルロジックデバイス、フィー

50

ルドプログラマブルゲートアレイ、光学、化学、生物学、量子又はナノエンジニアリングシステム、コンポーネント及び機構を使用してプログラムされた汎用デジタルコンピュータを使用することによって実装することができる。一般に、特定の実施形態の機能は、当業で周知のいずれかの手段によって実現することができる。分散システム、ネットワークシステム、コンポーネント及び／又は回路を使用することもできる。データの通信又は転送は、有線、無線、又は他のいずれかの手段によるものとすることができる。

【0035】

また、特定の用途に従って有用である時には、図面／図に示す要素のうちの1つ又は2つ以上をより分離又は統合された形で実装し、或いはいくつかの事例では除去又は動作不能とすることもできると理解されるであろう。上述した方法のいずれかをコンピュータが実行できるようにする、機械可読媒体に記憶できるプログラム又はコードを実装することも本発明の趣旨及び範囲に含まれる。

【0036】

「プロセッサ」は、データ、信号又はその他の情報を処理するあらゆる好適なハードウェア及び／又はソフトウェアシステム、機構又はコンポーネントを含む。プロセッサは、汎用中央処理装置、複数の処理装置、機能を実現するための専用回路又はその他のシステムを有するシステムを含むことができる。処理は、地理的位置に制限される必要も、又は時間的制限を有する必要もない。例えば、プロセッサは、その機能を「リアルタイム」、「オフライン」、「バッチモード」などで実行することができる。処理の一部は、異なる（又は同じ）処理システムが異なる時点に異なる場所で実行することもできる。処理システムの例は、サーバ、クライアント、エンドユーザ装置、ルータ、スイッチ、ネットワークストレージなどを含むことができる。コンピュータは、メモリと通信するあらゆるプロセッサとすることができる。メモリは、ランダムアクセスメモリ（RAM）、リードオンリメモリ（ROM）、磁気又は光学ディスク、或いはプロセッサが実行する命令を記憶するのに適した他の非一時的媒体などのあらゆる好適なプロセッサ可読記憶媒体とすることができる。

【0037】

本明細書及び以下の特許請求の範囲全体を通じて使用する「1つの（英文不定冠詞）」及び「その（英文定冠詞）」は、文脈において別途明確に示していない限り複数形の照応を含む。また、本明細書及び以下の特許請求の範囲全体を通じて使用する「～内（in）」の意味は、文脈において別途明確に示していない限り、「～内（in）」及び「～上（on）」の意味を含む。

【0038】

以上、本明細書では特定の実施形態について説明したが、上述した開示では修正、様々な変更及び置換の自由が意図されており、いくつかの例では、記載した範囲及び趣旨から逸脱することなく、特定の実施形態のいくつかの特徴が対応する他の特徴の使用を伴わずに使用されると理解されたい。従って、特定の状況又は材料を基本的範囲及び趣旨に適合させるように多くの修正を行うことができる。

【符号の説明】

【0039】

- 100 システム
- 105 ドローン
- 110 経路
- 120 シーン
- 130 フライコントローラ
- 140 カメラコントローラ
- 150 シーンの3D再構成
- 160 人間

10

20

30

40

50

【図面】
【図 1】

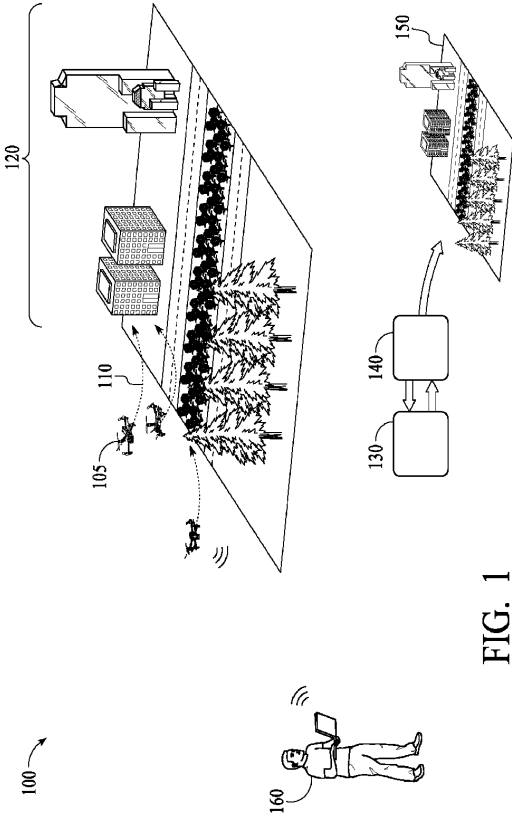


FIG. 1

【図 2】

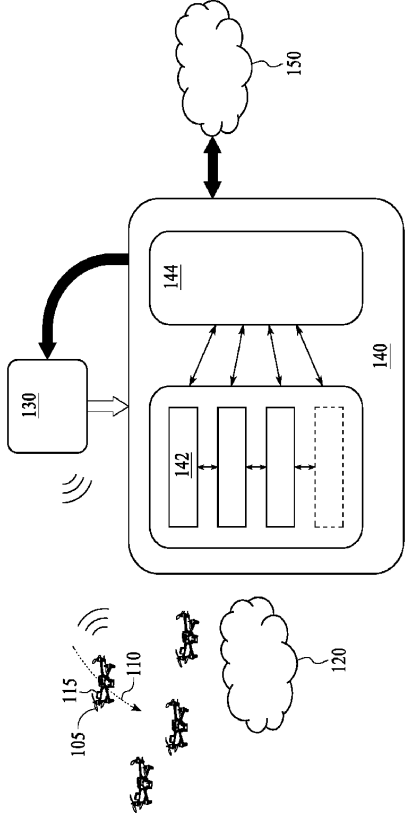


FIG. 2

【図 3】

速度	SLAM	多視点三角測量
精度	3Dマップの使用により高速 3Dマップの使用により良好であるが、 時間と共にわずかにドリフトする可能性	低速 目標カメラ距離に対する基線、及び特徴 マッチングの性能による
カメラ座標及びスケール	実行毎に困難な可能性	常に同じ座標に整合
画像重複要件	高（通常は〜90%の同様の視点）	低（60%でOK、200の異なる視点まで）
移動オブジェクト	よりロバスト	敏感

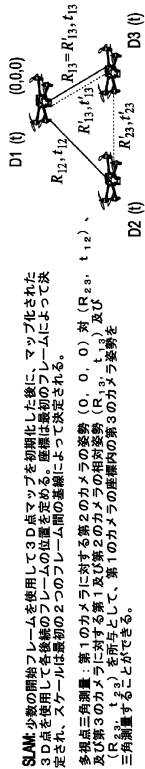


FIG. 3

【図 4】

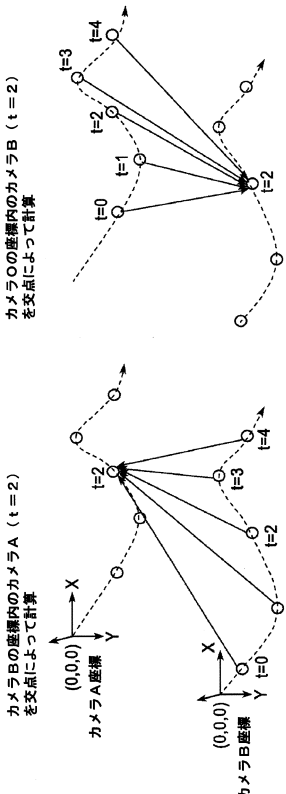


FIG. 4

BRISK特徴などのいずれかの特徴ベースの姿勢推定によって並進ベクトルからカメラ対カメラベクトルを導出することができる
各カメラ軌道からの中間サンプル（例えば、 $t=2$ ）のカメラ位置が望ましい

【図 5】

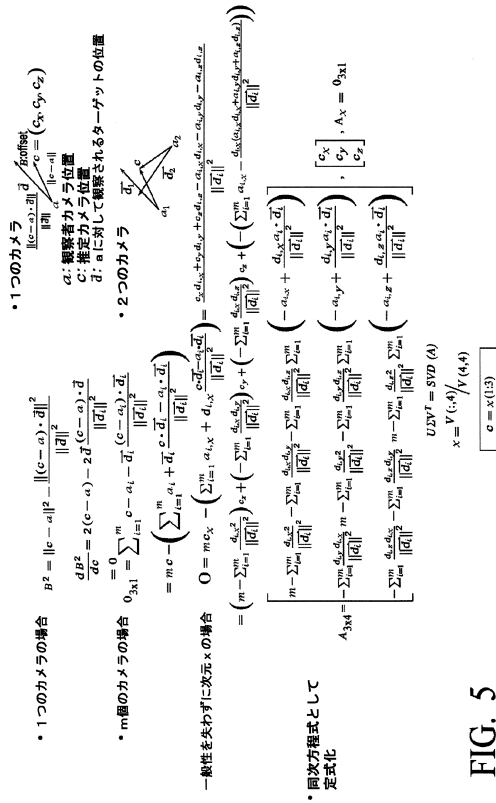


FIG. 5

【図 6】

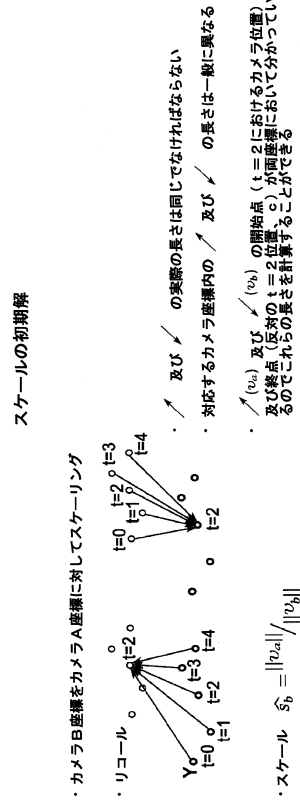


FIG. 6

【圖 7】

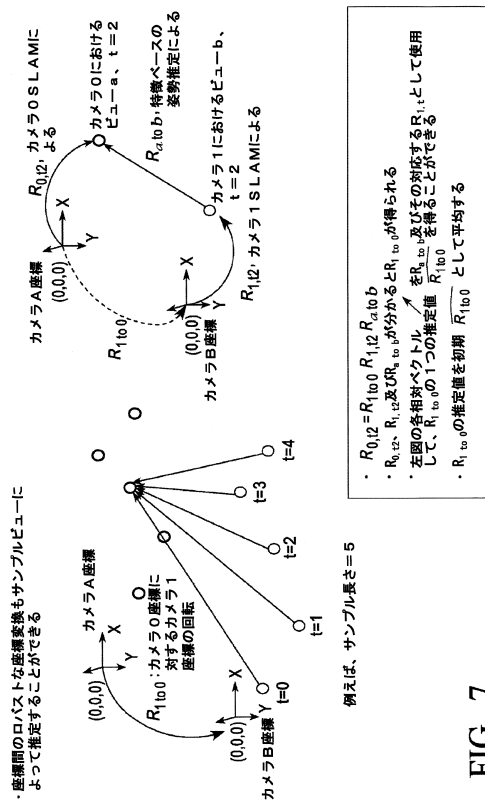


FIG. 7

【図 8】

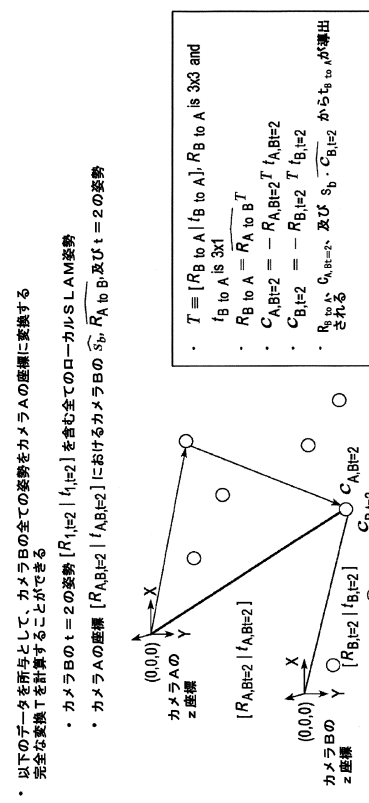


FIG. 8

【図 9】

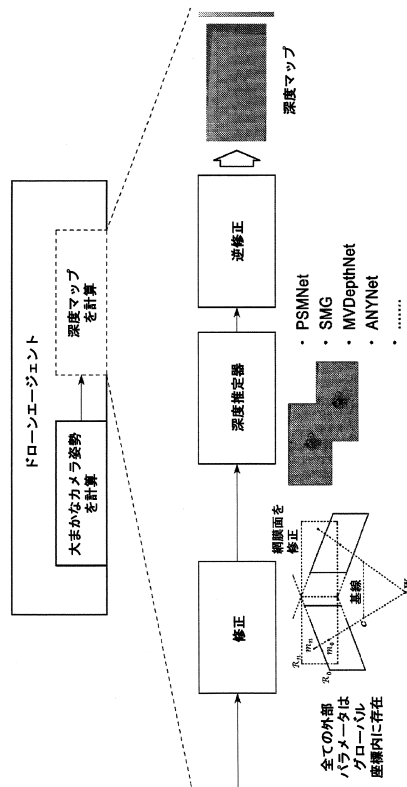


FIG. 9

【図 10】

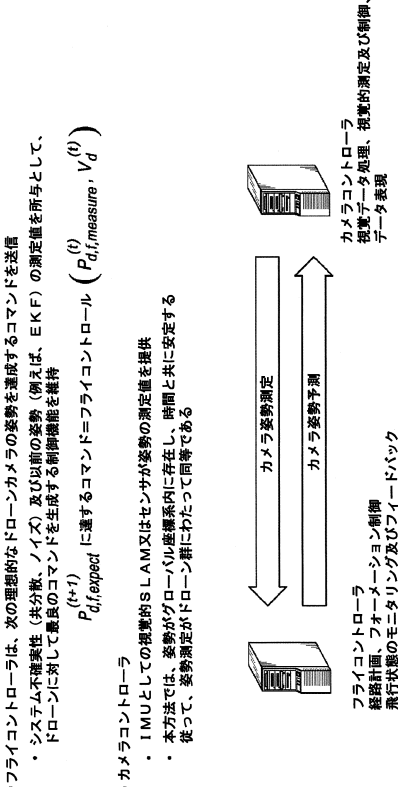


FIG. 10

【図 11】

- 初期化
 - 図8のドローン座標整合のためのフレームは、フライコントローラ座標をグローバル座標に整合させる変換Tの計算にも使用される
- フィードバック
 - フライコントローラは、時点 $t + w$ におけるコマンドがドローン d に送信されると、カメラコントローラ内のドローンエージェントに $P_{d,expect}^{(t+w)}$ を送信する
 - カメラコントローラは、 $T(P_{d,expect}^{(t+w)})$ を以前の姿勢として使用することにより、時点 $t + w$ におけるドローン d のセンサデータをを使用して $P_{d,measure}^{(t+w)}$ の計算を加速させる
 - 計算された姿勢 $P_{d,measure}^{(t+w)}$ 及びその時間タグ $t + w$ は、フライコントローラに送信される
 - フライコントローラは、 $T^{-1}(P_{d,measure}^{(t+w)})$ と $P_{d,expect}^{(t+w)}$ との間の相違を使用して $V_d^{(t)}$ を $V_d^{(t+w)}$ に更新する
 - フライコントローラは、 $V_d^{(t+w)}$ を所与として、予想される群フォーメーションを生成するのに最適なコマンドを計算する
 - カメラは全てのフレームをフィードバックする必要はなく、従って $w > 1$ であることができる

FIG. 11

【図 12】

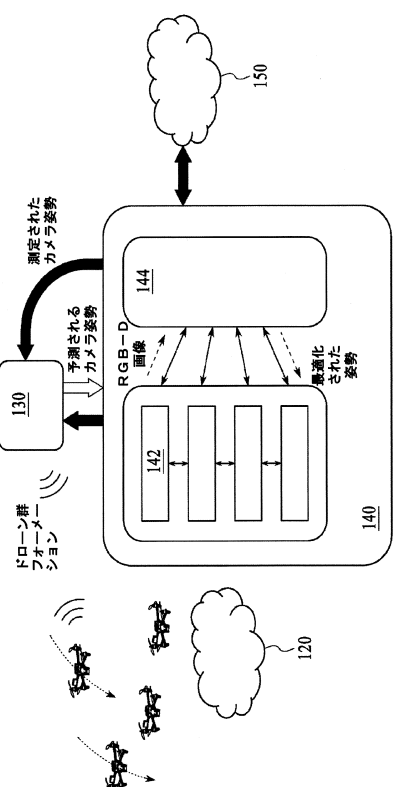


FIG. 12

フロントページの続き

(51)国際特許分類

F I

G 0 6 T 17/00 (2006.01)

G 0 6 T 17/00

H 0 4 N 23/60 (2023.01)

H 0 4 N 23/60 3 0 0

H 0 4 N 23/60 5 0 0

1 2 1 サンノゼ ブエナ ビュー コート 2 7 3 4

(72)発明者 ベレストフ アレクサンダー

アメリカ合衆国 カリフォルニア州 9 5 1 3 8 サン ノゼ マンダーストン ドライヴ 5 3 1 3

審査官 佐藤 直樹

(56)参考文献

国際公開第 2 0 2 0 / 0 4 0 6 7 9 (W O , A 1)

国際公開第 2 0 2 0 / 1 1 0 4 0 1 (W O , A 1)

特開 2 0 1 4 - 0 8 9 1 6 0 (J P , A)

(58)調査した分野 (Int.Cl., D B名)

H 0 4 N 5 / 2 2 2

G 0 5 D 1 / 0 0

G 0 5 D 1 / 1 0

G 0 6 T 7 / 5 5

G 0 6 T 7 / 7 0

G 0 6 T 1 7 / 0 0

H 0 4 N 2 3 / 6 0