

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6271741号
(P6271741)

(45) 発行日 平成30年1月31日 (2018. 1. 31)

(24) 登録日 平成30年1月12日 (2018. 1. 12)

(51) Int. Cl.

F I

G 0 6 F 17/30 (2006.01)

G 0 6 F 17/30 2 2 O B

G 0 6 F 17/30 1 7 O D

請求項の数 30 (全 90 頁)

(21) 出願番号	特願2016-541344 (P2016-541344)	(73) 特許権者	595020643
(86) (22) 出願日	平成26年12月19日 (2014. 12. 19)		クァアルコム・インコーポレイテッド
(65) 公表番号	特表2017-502412 (P2017-502412A)		Q U A L C O M M I N C O R P O R A T E D
(43) 公表日	平成29年1月19日 (2017. 1. 19)		アメリカ合衆国、カリフォルニア州 9 2
(86) 国際出願番号	PCT/US2014/071642		1 2 1 - 1 7 1 4、サン・ディエゴ、モア
(87) 国際公開番号	W02015/095781		ハウス・ドライブ 5 7 7 5
(87) 国際公開日	平成27年6月25日 (2015. 6. 25)	(74) 代理人	100108855
審査請求日	平成29年5月23日 (2017. 5. 23)		弁理士 蔵田 昌俊
(31) 優先権主張番号	61/919, 627	(74) 代理人	100109830
(32) 優先日	平成25年12月20日 (2013. 12. 20)		弁理士 福原 淑弘
(33) 優先権主張国	米国 (US)	(74) 代理人	100158805
(31) 優先権主張番号	62/036, 502		弁理士 井関 守三
(32) 優先日	平成26年8月12日 (2014. 8. 12)	(74) 代理人	100112807
(33) 優先権主張国	米国 (US)		弁理士 岡田 貴志

最終頁に続く

(54) 【発明の名称】 オブジェクト編成を符号化するためのシステム、方法、および装置

(57) 【特許請求の範囲】

【請求項 1】

メタデータを生成する方法であって、

ユーザインターフェースを介して、少なくとも1つのプロセッサによって、ビデオストリームの特定のフレームに示される複数のオブジェクトの選択を受信することと、ここにおいて、前記特定のフレームは物理空間に対応する位置空間中の前記複数のオブジェクトを示す、

前記少なくとも1つのプロセッサによって、前記位置空間中の前記複数のオブジェクトのそれぞれの位置を示す位置情報に基づいて、前記複数のオブジェクトが、前記位置空間中の特定の幾何学的配置上に配置されていることを決定することと、ここにおいて、前記特定の幾何学的配置は、前記複数のオブジェクトのうちのオブジェクトの第1のペア間の第1の距離と、前記複数のオブジェクトのうちのオブジェクトの第2のペア間の第2の距離との間の比較に基づいて、複数の幾何学的配置の候補から選択され、

前記少なくとも1つのプロセッサによって、前記特定のフレーム中の前記複数のオブジェクトの前記特定の幾何学的配置を特定する前記メタデータを生成することと、
を備える方法。

【請求項 2】

前記位置空間中の前記複数のオブジェクトについて基準位置を決定することと、ここにおいて、前記位置空間は、複数の領域を備える、

前記基準位置に最も近いセントロイドを有する、前記複数の領域のうちの特定の領域を

特定することと、ここにおいて、前記メタデータは前記特定の領域を示す、
をさらに備える、請求項 1 に記載の方法。

【請求項 3】

前記複数の領域の各領域は、コードブックの、異なるコードワードに対応し、第 1 のコードワードは第 1 の数のビットを含み、第 2 のコードワードは、前記第 1 の数のビットとは異なる第 2 の数のビットを含む、請求項 2 に記載の方法。

【請求項 4】

前記位置空間は複数の領域を備え、前記複数の領域の各領域は、コードブックの、異なるコードワードに対応し、前記複数のオブジェクトの各オブジェクトについて、前記メタデータは、前記複数の領域のうち、前記オブジェクトを含む特定の領域を特定する特定のコードワードを含む、請求項 1 に記載のメタデータを生成する方法。

10

【請求項 5】

前記位置空間中の前記複数のオブジェクトについて基準位置を決定することをさらに備え、

前記基準位置は、前記ビデオストリームのピクセル座標空間または前記物理空間中の前記複数のオブジェクトの位置に基づく、請求項 1 に記載の方法。

【請求項 6】

前記メタデータを含むファイルを生成することをさらに備え、ここにおいて、前記ファイルは、前記ビデオストリームとは分離される、請求項 1 に記載の方法。

【請求項 7】

20

前記複数のオブジェクトの総数、前記位置空間の 1 つまたは複数の領域に含まれるオブジェクトの数、またはそれらの組合せに基づいて、1 つまたは複数の尺度を決定することと、

前記 1 つまたは複数の尺度に基づいて、前記特定の幾何学的配置を選択するために決定木を使用することと、

をさらに備える、請求項 1 に記載の方法。

【請求項 8】

前記位置空間は、前記ビデオストリームのピクセル座標空間を備える、請求項 1 に記載の方法。

【請求項 9】

30

前記選択を受信することに応答して、前記ビデオストリーム内で、前記複数のオブジェクトをトラッキングすることをさらに備える、請求項 1 に記載の方法。

【請求項 10】

プロセッサによって実行されると、前記プロセッサに、

ビデオストリームの特定のフレームに示される複数のオブジェクトの選択を受信することと、ここにおいて、前記特定のフレームは物理空間に対応する位置空間中の前記複数のオブジェクトを示す、

前記複数のオブジェクトのそれぞれの位置を示す位置情報に基づいて、前記複数のオブジェクトが、前記位置空間中の特定の幾何学的配置上に配置されていることを決定することと、ここにおいて、前記特定の幾何学的配置は、前記複数のオブジェクトのうちのオブジェクトの第 1 のペア間の第 1 の距離と、前記複数のオブジェクトのうちのオブジェクトの第 2 のペア間の第 2 の距離との間の比較に基づいて、複数の幾何学的配置の候補から選択され、

40

前記特定のフレーム中の前記複数のオブジェクトの前記特定の幾何学的配置を特定するメタデータを生成することと、

を行わせる命令を記憶する非一時的コンピュータ可読媒体。

【請求項 11】

前記複数のオブジェクトは、前記ビデオストリームに示されるトラッキングされるオブジェクトのセットから選択される、請求項 10 に記載の非一時的コンピュータ可読媒体。

【請求項 12】

50

オブジェクトの前記第 1 のペアは、前記複数のオブジェクトのうちの第 1 のオブジェクトと前記複数のオブジェクトのうちの第 2 のオブジェクトとを含み、オブジェクトの前記第 2 のペアは、前記第 2 のオブジェクトと前記複数のオブジェクトのうちの第 3 のオブジェクトとを含む、請求項 10 に記載の非一時的コンピュータ可読媒体。

【請求項 13】

前記メタデータは、前記特定の幾何学的配置に適用されるべきスケール係数、前記特定の幾何学的配置に関する方向の向きのインジケータ、および前記特定の幾何学的配置が軸の回りで反転されるべきであるかどうかを示すミラーフラグインジケータを示す、請求項 10 に記載の非一時的コンピュータ可読媒体。

【請求項 14】

前記命令は、前記プロセッサによって実行されると、さらに前記プロセッサに、前記複数のオブジェクトの基準位置を、地面における前記複数のオブジェクトの平均位置として計算させる、請求項 10 に記載の非一時的コンピュータ可読媒体。

【請求項 15】

前記メタデータは、前記ビデオストリームのキャプチャの間に生成され、ここにおいて、前記命令は、前記プロセッサによって実行されると、さらに前記プロセッサに、前記複数のオブジェクトを特定するために、前記メタデータを探索させる、請求項 10 に記載の非一時的コンピュータ可読媒体。

【請求項 16】

前記命令は、前記プロセッサによって実行されると、さらに前記プロセッサに、前記ビデオストリーム内で前記ビデオストリームの非連続的なフレームにおける前記複数のオブジェクトの幾何学的配置をトラッキングすることと、各非連続的なフレームについて、

ディスプレイを介して、前記幾何学的配置の指示を提示することと、

前記非連続的なフレームについて前記複数の幾何学的配置の候補のうちの 1 つの幾何学的配置を特定する、対応するメタデータを生成することと、

を行わせる、請求項 10 に記載の非一時的コンピュータ可読媒体。

【請求項 17】

装置であって、

ビデオストリームの特定のフレームに示される複数のオブジェクトの選択を受信するための手段と、ここにおいて、前記特定のフレームは物理空間に対応する位置空間中の前記複数のオブジェクトを示す、

前記複数のオブジェクトのそれぞれの位置を示す位置情報に基づいて、前記複数のオブジェクトが、前記位置空間中の特定の幾何学的配置上に配置されていることを決定するための手段と、ここにおいて、前記特定の幾何学的配置は、前記複数のオブジェクトのうちのオブジェクトの第 1 のペア間の第 1 の距離と、前記複数のオブジェクトのうちのオブジェクトの第 2 のペア間の第 2 の距離との間の比較に基づいて、複数の幾何学的配置の候補から選択され、

前記特定のフレーム中の前記複数のオブジェクトの前記特定の幾何学的配置を特定するメタデータを生成するための手段と、

を備える装置。

【請求項 18】

前記位置空間中の前記複数のオブジェクトについて基準位置を決定するための手段をさらに備え、

ここにおいて、前記基準位置は、前記複数のオブジェクトの重心に基づいて決定される、請求項 17 に記載の装置。

【請求項 19】

前記メタデータは、前記複数のオブジェクトの総数または前記複数のオブジェクトの前記総数の近似値を示す、請求項 17 に記載の装置。

【請求項 20】

前記メタデータはマルチビット数を備え、前記マルチビット数の各ビット位置は、前記位置空間の複数の領域のうちの異なる領域に対応し、ここにおいて、前記マルチビット数の各ビットについて、前記ビットの値は、対応する領域が、前記複数のオブジェクトのうちの少なくとも１つのオブジェクトによって占有されているかどうかを示す、請求項１７に記載の装置。

【請求項２１】

前記位置空間が複数の重複しない領域を備える、請求項１７に記載の装置。

【請求項２２】

前記特定の幾何学的配置は、前記複数のオブジェクトの位置によって定義される形状を備え、ここにおいて、前記複数のオブジェクトの各オブジェクトについて、前記オブジェクトの特定の位置は、前記形状の異なる頂点に対応する、請求項１７に記載の装置。

10

【請求項２３】

前記ビデオストリーム内で、前記複数のオブジェクトをトラッキングするための手段をさらに備える、請求項１７に記載の装置。

【請求項２４】

装置であって、

ビデオストリームの特定のフレーム内の複数のオブジェクトの選択を受信することと、ここにおいて、前記特定のフレームは物理空間に対応する位置空間中の前記オブジェクトを示す、

前記位置空間中の前記複数のオブジェクトのそれぞれの位置を示す位置情報を取得することと

20

を行うように構成される判別器と、

前記位置情報に基づいて、前記複数のオブジェクトが、前記位置空間中の特定の幾何学的配置上に配置されていることを決定することと、ここにおいて、前記特定の幾何学的配置は、前記複数のオブジェクトのうちのオブジェクトの第１のペア間の第１の距離と、前記複数のオブジェクトのうちのオブジェクトの第２のペア間の第２の距離との間の比較に基づいて、複数の幾何学的配置の候補から選択され、

前記特定のフレーム中の前記複数のオブジェクトの前記特定の幾何学的配置を特定するメタデータを生成することと

を行うように構成される選択器と、

30

を備える、装置。

【請求項２５】

前記位置空間中の前記複数のオブジェクトについて基準位置を決定するように構成されるオブジェクトトラッカーをさらに備え、

前記位置情報は、前記基準位置に対する前記複数のオブジェクトの位置を示す、請求項２４に記載の装置。

【請求項２６】

前記選択器は、前記複数の幾何学的配置の候補のうちのどれが、前記複数のオブジェクトの前記位置によって表される形状と最も類似しているかを決定するように構成される、請求項２４に記載の装置。

40

【請求項２７】

プロセッサに結合され、データを送信および受信するように構成される通信回路をさらに備える、請求項２４に記載の装置。

【請求項２８】

前記選択器は、前記複数の幾何学的配置の候補のうちのどれが、前記位置空間中の前記複数のオブジェクトの分布と最も類似しているかを決定するように構成され、

前記ビデオストリーム内で、前記複数のオブジェクトをトラッキングするように構成されるオブジェクトトラッカーをさらに備え、前記位置空間は地面を備える、請求項２４に記載の装置。

【請求項２９】

50

前記メタデータを前記ビデオストリームの前記特定のフレームと関連付けるように構成されるビデオストリームエディタをさらに備え、

前記選択器は、前記位置空間中の前記複数のオブジェクトを示す第2のフレーム中の前記複数のオブジェクトの各々の第2の位置を示す情報に基づいて、前記複数のオブジェクトの前記複数の幾何学的配置の候補のうちの異なる1つを特定する第2のメタデータを生成するように構成され、

前記ビデオストリームエディタは、前記第2のメタデータを前記ビデオストリームの前記第2のフレームと関連付けるように構成される、請求項24に記載の装置。

【請求項30】

プロセッサに結合され、前記ビデオストリームを生成するように構成されるカメラをさらに備える、請求項24に記載の装置。

【発明の詳細な説明】

【優先権の主張】

【0001】

関連出願の相互参照

本出願は、その内容全体が参照により本明細書に明確に組み込まれる、同一出願人により所有される2013年12月20日出願された米国仮特許出願第61/919,627号、2014年8月12日出願された米国仮特許出願第62/036,502号、2014年10月14日出願された米国仮特許出願第62/063,498号、および2014年12月18日出願された米国非仮特許出願第14/575,986号の優先権を主張する。

【技術分野】

【0002】

本開示は全般に、画像およびビデオメタデータに関する。

【背景技術】

【0003】

[0003]技術の進歩は、より小さく、より強力なコンピューティングデバイスをもたらした。たとえば、現在、小型で、軽量で、ユーザによって容易に持ち運ばれるポータブルワイヤレス電話、携帯情報端末(PDA)、およびページングデバイスなどのワイヤレスコンピューティングデバイスを含む、様々なポータブルパーソナルコンピューティングデバイスが存在する。より具体的には、携帯電話およびインターネットプロトコル(IP)電話などのポータブルワイヤレス電話機は、ワイヤレスネットワークを通じて音声パケットとデータパケットとを通信することができる。さらに、多くのそのようなワイヤレス電話は、内蔵されている他のタイプのデバイスを含む。たとえば、ワイヤレス電話はまた、デジタルスチルカメラと、デジタルビデオカメラと、デジタルレコーダと、オーディオファイルプレーヤとを含み得る。また、そのようなワイヤレス電話は、インターネットにアクセスするために使用され得るウェブブラウザアプリケーションなどのソフトウェアアプリケーションを含む、実行可能命令を処理することができる。したがって、これらのワイヤレス電話は、かなりのコンピューティング能力を含み得る。

【発明の概要】

【0004】

[0004]特定の実施形態では、方法は、モバイルデバイスにおいて、特定のフレームにおける第1のオブジェクトの第1の位置と、特定のフレームにおける第2のオブジェクトの第2の位置とをトラッキングすることを含む。方法はまた、第1の位置と関連付けられる第1の座標データと、第2の位置と関連付けられる第2の座標データとを生成することを含む。方法はさらに、第1の符号化された位置データと第2の符号化された位置データとを生成するために、コードブックを使用して第1の座標データと第2の座標データとを符号化することを含む。方法はまた、第1の位置および第2の位置に対応する入力パラメータに基づく特定のフレームの検索を可能にするために、第1の符号化された位置データと第2の符号化された位置データとを記憶することを含む。

【 0 0 0 5 】

[0005]別の特定の実施形態では、モバイルデバイスは、メモリと、メモリに結合されたプロセッサとを含む。プロセッサは、モバイルデバイスにおいて、特定のフレームにおける第1のオブジェクトの第1の位置と、特定のフレームにおける第2のオブジェクトの第2の位置とをトラッキングするように構成される。プロセッサはまた、第1の位置と関連付けられる第1の座標データと、第2の位置と関連付けられる第2の座標データとを生成するように構成される。プロセッサはさらに、第1の符号化された位置データと第2の符号化された位置データとを生成するために、コードブックを使用して第1の座標データと第2の座標データとを符号化するように構成される。プロセッサはまた、第1の位置および第2の位置に対応する入力パラメータに基づく特定のフレームの検索を可能にするために、第1の符号化された位置データと第2の符号化された位置データとを記憶するように構成される。

10

【 0 0 0 6 】

[0006]別の特定の実施形態では、モバイルデバイスは、特定のフレームにおける第1のオブジェクトの第1の位置と、特定のフレームにおける第2のオブジェクトの第2の位置とをトラッキングするための手段を含む。モバイルデバイスはまた、第1の位置と関連付けられる第1の座標データと、第2の位置と関連付けられる第2の座標データとを生成するための手段を含む。モバイルデバイスはさらに、第1の符号化された位置データと第2の符号化された位置データとを生成するために、コードブックを使用して第1の座標データと第2の座標データとを符号化するための手段を含む。モバイルデバイスはまた、第1の位置および第2の位置に対応する入力パラメータに基づく特定のフレームの検索を可能にするために、第1の符号化された位置データと第2の符号化された位置データとを記憶するための手段を含む。

20

【 0 0 0 7 】

[0007]別の特定の実施形態では、非一時的コンピュータ可読媒体は、プロセッサによって実行されると、プロセッサに、モバイルデバイスにおいて、特定のフレームにおける第1のオブジェクトの第1の位置と、特定のフレームにおける第2のオブジェクトの第2の位置とをトラッキングさせる命令を含む。命令はまた、プロセッサに、第1の位置と関連付けられる第1の座標データと、第2の位置と関連付けられる第2の座標データとを生成させるように実行可能である。命令はさらに、プロセッサに、第1の符号化された位置データと第2の符号化された位置データとを生成するために、コードブックを使用して第1の座標データと第2の座標データとを符号化させるように実行可能である。命令はまた、プロセッサに、第1の位置および第2の位置に対応する入力パラメータに基づく特定のフレームの検索を可能にするために、第1の符号化された位置データと第2の符号化された位置データとを記憶させるように実行可能である。

30

【 0 0 0 8 】

[0008]一般的な構成に従ってメタデータを生成する方法は、複数のオブジェクトを選択するために少なくとも1つのデジタル画像を使用することを含み、少なくとも1つのデジタル画像は物理空間との関連で当該複数のオブジェクトを示す。この方法はまた、少なくとも1つのプロセッサによって、位置空間中の選択されたオブジェクトの位置を示す情報に基づいて、選択されたオブジェクトの複数の幾何学的構成 (geometrical arrangements) の候補のうちの1つを特定するメタデータを生成することを含む。命令を実行する機械にそのような方法を実行させる命令を有する、コンピュータ可読記憶媒体 (たとえば、非一時的媒体) も開示される。

40

【 0 0 0 9 】

[0009]一般的な構成に従ってメタデータを生成するための装置は、複数のオブジェクトを選択するために少なくとも1つのデジタル画像を使用するための手段を含み、少なくとも1つのデジタル画像は物理空間との関連で当該複数のオブジェクトを示す。この装置はまた、位置空間中の選択されたオブジェクトの位置を示す情報に基づいて、選択されたオブジェクトの複数の幾何学的構成の候補のうちの1つを特定するメタデータを生成するた

50

めの手段を含む。

【 0 0 1 0 】

[0010]別の一般的な構成に従ってメタデータを生成するための装置は、複数のオブジェクトを選択するために少なくとも1つのデジタル画像を使用するように構成される判別器を含み、少なくとも1つのデジタル画像は物理空間との関連で当該複数のオブジェクトを示す。方法はまた、位置空間中の選択されたオブジェクトの位置を示す情報に基づいて、選択されたオブジェクトの複数の幾何学的構成の候補のうちの1つを特定するメタデータを生成するように構成される選択器を含む。

【 0 0 1 1 】

[0011]開示される実施形態の少なくとも1つによってもたらされる1つの具体的な利点は、モバイルデバイス上での画像トラッキングとフレーム検索とを使用したユーザ体験の改善である。本開示の他の態様、利点、および特徴は、以下のセクション、すなわち、「図面の簡単な説明」と、「発明を実施するための形態」と、「特許請求の範囲」とを含めて本出願全体を検討した後に明白になる。

【図面の簡単な説明】

【 0 0 1 2 】

【図1】電子デバイスを示すブロック図。

【図2A】オブジェクトおよび検出モジュールの特定の例示的な実施形態を示すブロック図。

【図2B】図2のオブジェクトおよび検出モジュール内のプロセッサ実装コンポーネントの特定の例示的な実施形態を示すブロック図。

【図3】動きに基づくトラッキングとオブジェクト検出とを実行するための方法の特定の例示的な実施形態のフローチャート。

【図4】動きに基づくトラッキングを実行するための方法の特定の例示的な実施形態のフローチャート。

【図5】前方後方誤差に基づいて動きに基づくトラッキングにおけるトラッキング誤差を推定するための方法の特定の例示的な実施形態のフローチャート。

【図6】オブジェクト検出を実行するための方法の特定の例示的な実施形態のフローチャート。

【図7】異なるウィンドウサイズを有する画像ウィンドウの特定の例示的な実施形態の図。

【図8】オブジェクトトラッキングおよび検出モジュールの特定の例示的な実施形態を示すブロック図。

【図9】平滑化モジュールの特定の例示的な実施形態を示すブロック図。

【図10】動きトラッキングにおける平滑化ジッタのための方法の特定の例示的な実施形態のフローチャート。

【図11A】一般的な構成によるメタデータを生成する方法M100のフローチャート。

【図11B】方法M100の実装形態M110のフローチャート。

【図11C】方法M100の実装形態M120のフローチャート。

【図11D】方法M110およびM120の実装形態M130のフローチャート。

【図12】3つのオブジェクトを選択するための一連のユーザの動作の例を示す図。

【図13】3つのオブジェクトを選択するための一連のユーザの動作の別の例を示す図。

【図14】フレーム内のオブジェクトとオブジェクトの強調された表示の例を示す図。

【図15A】方向感知デバイスを使用して選択点を操作する例を示す図。

【図15B】方向感知デバイスを使用して選択点を操作する例を示す図。

【図15C】方向感知デバイスを使用して選択点を操作する例を示す図。

【図16】実際の距離と比較されたときのオブジェクトの画像間の距離の不一致の例を示す図。

【図17A】オブジェクトの下部境界線へとオブジェクトの重心を投影して対応する決定された位置を得る例を示す図。

10

20

30

40

50

【図 1 7 B】地面の決定された位置へのリンクされた点の投影の例を示す図。

【図 1 8 A】ピクセル座標空間における距離と地面における距離との対応付けを示す図。

【図 1 8 B】ピクセル座標空間における距離と地面における距離との対応付けを示す図。

【図 1 9 A】バスケットボールコートの上図。

【図 1 9 B】コートの位置空間の不均一な分割のある例を示す図。

【図 1 9 C】コートの位置空間の不均一な分割のある例を示す図。

【図 1 9 D】コートの位置空間の不均一な分割のある例を示す図。

【図 2 0】オブジェクトトラッキングを使用したビデオ処理を示す別の特定の実施形態の図。

【図 2 1】クラスタ化に基づいてフレームを検索するために使用されるモバイルデバイスの画面の特定の実施形態を示す図。 10

【図 2 2】オブジェクトトラッキングを使用したビデオ処理方法の特定の例示的な実施形態の別のフローチャート。

【図 2 3】オブジェクトトラッキングを使用したビデオ処理方法の特定の例示的な実施形態のフローチャート。

【図 2 4】矩形座標を使用してオブジェクトの位置のためのメタデータを符号化する特定の実施形態を示す図。

【図 2 5】極座標を使用してオブジェクトの位置のためのメタデータを符号化する特定の実施形態を示す図。

【図 2 6】メタデータを符号化するための方法の特定の例示的な実施形態のフローチャート。 20

【図 2 7】オブジェクトトラッキングを使用したビデオ処理方法の特定の例示的な実施形態の別のフローチャート。

【図 2 8】オブジェクトトラッキングを使用したビデオ処理を示す別の特定の実施形態の図。

【図 2 9】オブジェクトトラッキングを使用したビデオ処理方法の特定の例示的な実施形態の別のフローチャート。

【図 3 0 A】一般的な構成による装置 A 1 0 0 のブロック図。

【図 3 0 B】装置 A 1 0 0 の実装形態 A 1 1 0 のブロック図。

【図 3 0 C】装置 A 1 0 0 の実装形態 A 1 2 0 のブロック図。 30

【図 3 0 D】装置 A 1 1 0 および A 1 2 0 の実装形態 A 1 3 0 のブロック図。

【図 3 1】ビデオ処理技法を実行するように動作可能なコンポーネントを含むワイヤレスデバイスのブロック図。

【図 3 2 A】一般的な構成による装置 M F 1 0 0 のブロック図。

【図 3 2 B】装置 M F 1 0 0 の実装形態 M F 1 1 0 のブロック図。

【図 3 2 C】装置 M F 1 0 0 の実装形態 M F 1 2 0 のブロック図。

【図 3 2 D】装置 M F 1 1 0 および M F 1 2 0 の実装形態 M F 1 3 0 のブロック図。

【図 3 3 A】テンプレートのセットの例を示す図。

【図 3 3 B】テンプレートへの修正の例を示す図。

【図 3 4 A】選択されたオブジェクトの異なる編成 (formations) の例を示す図。 40

【図 3 4 B】一意ではないマッピングの例を示す図。

【図 3 4 C】選択されたオブジェクトの一意な編成を示す図。

【図 3 5】決定の尺度 (メトリック) のセットの例を示す図。

【図 3 6】決定の尺度のセットの別の例を示す図。

【図 3 7 A】方法 M 2 0 0 の実装形態 M 2 0 0 のフローチャート。

【図 3 7 B】方法 M 1 1 0 の実装形態 M 2 1 0 のフローチャート。

【図 3 7 C】方法 M 1 2 0 の実装形態 M 2 2 0 のフローチャート。

【図 3 7 D】方法 M 1 3 0 の実装形態 M 2 3 0 のフローチャート。

【図 3 8 A】一般的な構成によるメタデータを生成するための装置 A 1 0 0 のブロック図

【図 3 8 B】装置 A 1 0 0 の実装形態 A 1 1 0 のブロック図。
 【図 3 8 C】装置 A 1 0 0 の実装形態 A 1 2 0 のブロック図。
 【図 3 8 D】装置 A 1 1 0 および A 1 2 0 の実装形態 A 1 3 0 のブロック図。
 【図 3 8 E】装置 A 1 0 0 の実装形態 A 2 0 0 のブロック図。
 【図 3 8 F】装置 A 1 1 0 の実装形態 A 2 1 0 のブロック図。
 【図 3 8 G】装置 A 1 2 0 の実装形態 A 2 2 0 のブロック図。
 【図 3 9 A】装置 A 1 3 0 の実装形態 A 2 3 0 のブロック図。
 【図 3 9 B】装置 A 1 0 0 の実装形態 A 3 0 0 のブロック図。
 【図 3 9 C】装置 A 2 0 0 の実装形態 A 3 1 0 のブロック図。
 【図 3 9 D】装置 A 2 3 0 の実装形態 A 3 3 0 のブロック図。
 【図 4 0 A】一般的な構成によるメタデータを生成するための装置 M F 1 0 0 のブロック図。
 【図 4 0 B】装置 M F 1 0 0 の実装形態 M F 1 1 0 のブロック図。
 【図 4 0 C】装置 M F 1 0 0 の実装形態 M F 1 2 0 のブロック図。
 【図 4 0 D】装置 M F 1 0 0 の実装形態 M F 1 3 0 のブロック図。
 【図 4 0 E】装置 M F 1 0 0 の実装形態 M F 2 0 0 のブロック図。
 【図 4 0 F】装置 M F 1 1 0 の実装形態 M F 2 1 0 のブロック図。
 【図 4 0 G】装置 M F 1 2 0 の実装形態 M F 2 2 0 のブロック図。
 【図 4 1 A】装置 M F 1 3 0 の実装形態 M F 2 3 0 のブロック図。
 【図 4 1 B】装置 M F 1 0 0 の実装形態 M F 4 0 0 のブロック図。
 【図 4 1 C】装置 M F 1 0 0 の実装形態 M F 3 0 0 のブロック図。
 【図 4 1 D】装置 M F 2 0 0 の実装形態 M F 3 1 0 のブロック図。
 【図 4 1 E】装置 M F 2 3 0 の実装形態 M F 3 3 0 のブロック図。
 【図 4 2】オブジェクトトラッキングを使用したビデオ処理を示す特定の実施形態の図。
 【図 4 3】オブジェクトの位置のためのメタデータを符号化する特定の実施形態を示す図

10

20

【図 4 4】ビデオ処理方法の特定の実施形態を示すフローチャート。

【発明を実施するための形態】

【0013】

[0082]ワイヤレス電話または他のモバイルデバイスは、カメラによってビデオストリームをキャプチャし、ならびに / または、別のデバイスから、および / もしくはネットワークを介して、ビデオストリームを受信することができる。ビデオストリーム内のオブジェクトをトラッキングするための、新しい特徴および / または改善された特徴が望まれ得る。

30

【0014】

[0083]文脈によって明確に限定されない限り、「信号」という用語は、本明細書では、ワイヤ、バス、または他の伝送媒体上で表されるメモリ位置（またはメモリ位置のセット）の状態を含む、その通常の意味のいずれをも示すために使用される。文脈によって明確に限定されない限り、「発生（generating）」という用語は、本明細書では、算出（computing）または別様の生成（producing）のような、その通常の意味のいずれをも示すために使用される。文脈によって明確に限定されない限り、「計算（calculating）」という用語は、本明細書では、複数の値からの算出（computing）、評価、推定、および / または選択のような、その通常の意味のいずれをも示すために使用される。文脈によって明確に限定されない限り、「取得（obtaining）」という用語は、計算（calculating）、導出、（たとえば、外部デバイスからの）受信、および / または（たとえば、記憶要素のレイからの）検索のような、その通常の意味のいずれをも示すために使用される。文脈によって明確に限定されない限り、「選択（selecting）」という用語は、2 つ以上のセットの少なくとも 1 つ、およびすべてよりも少数の識別、指示、適用、および / または使用のような、その通常の意味のいずれをも示すために使用される。「備える（comprising）」という用語は、本明細書と特許請求の範囲とにおいて使用される場合、他の要素または動

40

50

作を除外するものではない。「に基づく」(「AはBに基づく」など)という用語は、(i)「から導出される」(たとえば、「BはAの前身である」)、(ii)「少なくとも~に基づく」(たとえば、「Aは少なくともBに基づく」)、および特定の文脈で適当な場合に、(iii)「に等しい」(たとえば、「AはBに等しい」)という場合を含む、その通常の意味のいずれかを示すのに使用される。同様に、「に応答して」という用語は、「少なくとも~に応答して」を含む、その通常の意味のいずれをも示すために使用される。

【0015】

[0084]別段規定されていない限り、特定の特徴を有する装置の動作のいずれの開示も、類似の特徴を有する方法を開示する(その逆も同様)ことが明確に意図され、特定の構成による装置の動作のいずれの開示も、類似の構成による方法を開示する(その逆も同様)ことが明確に意図される。「構成」という用語は、その具体的な文脈によって示されるように、方法、装置、および/またはシステムに関して使用され得る。「方法」、「処理」、「手順」、および「技法」という用語は、具体的な文脈によって別段に規定されていない限り、一般的、および互換的に使用される。「装置」および「デバイス」という用語も、具体的な文脈によって別段に規定されていない限り、一般的、および互換的に使用される。「要素」および「モジュール」という用語は、一般に、より大きな構成の一部を示すために使用される。文脈によって明確に限定されない限り、「システム」という用語は、本明細書では、「共通の目的を果たすために相互作用する要素のグループ」を含む、その通常の意味のいずれをも示すために使用される。

【0016】

[0085]別段に規定されていない限り、「一連」という用語は、2つ以上の項目の順列を示すために使用される。定冠詞によって最初に導入されない限り、請求要素を修飾するのに使用される順序を示す用語(たとえば、「第1の」、「第2の」、「第3の」など)は、それ自体、別のものに対する請求要素の優先順位または順序を示しておらず、請求要素を、(順序を示す用語の使用を別にすれば)同じ名前を有する別の請求要素から区別しているだけである。文脈によって明確に限定されない限り、「複数」および「セット」という用語の各々は、本明細書では、1よりも大きい整数量を示すために使用される。

【0017】

[0086]図1を参照すると、電子デバイス102を示すブロック図が示される。電子デバイス102はまた、ワイヤレス通信デバイス、モバイルデバイス、移動局、加入者局、クライアント、クライアント局、ユーザ機器(UE)、遠隔局、アクセス端末、モバイル端末、端末、ユーザ端末、加入者ユニットなどとも呼ばれ得る。電子デバイスの例としては、ラップトップまたはデスクトップコンピュータ、携帯電話、スマートフォン、ワイヤレスモデム、電子リーダー、タブレットデバイス、ゲームシステムなどがある。これらのデバイスのいくつかは、1つまたは複数の業界規格に従って動作し得る。

【0018】

[0087]スマートフォンまたはタブレットコンピュータのような電子デバイス102は、カメラを含み得る。カメラは、画像センサ114と、光学系118の視野内に位置するオブジェクトの画像を画像センサ114上に合焦させる光学系118(たとえば、レンズ)とを含み得る。電子デバイス102はまた、カメラソフトウェアアプリケーションと表示画面とを含み得る。カメラアプリケーションが動作しているとき、光学系118の視野内に位置するオブジェクトの画像は画像センサ114によって記録され得る。画像センサ114によって記録されている画像は、表示画面上に表示され得る。これらの画像は、任意の所与の瞬間において、光学系118の視野内に位置するオブジェクトが表示画面上に表示されるように、比較的高いフレームレートで高速に連続して表示され得る。実施形態はキャプチャされたフレーム(たとえば、ビデオフレーム)に関して説明されるが、本明細書で論じられる技法は、任意のデジタル画像に対して使用され得る。したがって、「フレーム」および「デジタル画像」という用語は、本明細書において互換的に使用され得る。

【0019】

[0088]カメラアプリケーションのユーザインターフェース120は、表示画面上に表示されている1つまたは複数のオブジェクトがトラッキングされることを可能にし得る。電子デバイス102のユーザは、トラッキングされるべきである(1つまたは複数の)オブジェクトを選択することを許可され得る。さらに、(1つまたは複数の)選択されたオブジェクトは、オブジェクトを後で検出するための基準として使用され得る。

【0020】

[0089]一構成では、ディスプレイは、たとえば、指、スタイラスまたは他のツールによる、物理的タッチからの入力を受信するタッチスクリーン116である。タッチスクリーン116は、トラッキングされるべきターゲットオブジェクトを定義するタッチ入力を受け取り得る。たとえば、電子デバイス102が、対象の動物を含む自然のシーンをキャプチャしている場合、ユーザは、必要な場合、その動物がトラッキングまたは検出されるべきであるという希望を示すバウンディングボックスをその動物の周りに描き得る。ターゲットオブジェクトは、任意の好適な方法で選択され得る。たとえば、トラッキング、検出、または両方が行われるべきであるターゲットオブジェクトを選択するために、顔認識、歩行者認識などが使用され得る。一構成では、複数のオブジェクトがトラッキングされ得る。ユーザインターフェース120は、たとえば、1つまたは複数のターゲットオブジェクトを選択する(すなわち、定義する)ために、ユーザがオブジェクトトラッキングおよび検出モジュール104と対話することを可能にし得る。タッチスクリーン116は、ビューファインダ131を含み得る。ビューファインダ131は、ビデオストリームまたはライブフィードを表示するタッチスクリーン116の部分を指し得る。たとえば、ビューファインダ131は、電子デバイス102上のカメラによって取得されたビューを表示することができる。

【0021】

[0090]電子デバイス102は、ビデオフレーム中の選択されたオブジェクトをトラッキングし、および/またはそのオブジェクトを検出するための、オブジェクトトラッキングおよび検出モジュール104を含み得る。オブジェクトトラッキングおよび検出モジュール104は、1つまたは複数のオブジェクトをトラッキングするための動きトラッカー106を含み得る。動きトラッカー106は、前のビデオフレームと現在のビデオフレームとの間でのターゲットオブジェクトの位置および/または位置の変化を推定するために、フレームごとに画像(たとえば、ビデオフレーム)上の点の動きをトラッキングするために、動きに基づき得る。

【0022】

[0091]オブジェクトトラッキングおよび検出モジュールはまた、ビデオフレーム中のオブジェクトを検出するためのオブジェクト検出器108を含み得る。オブジェクト検出器108は、現在のビデオフレームのすべてまたは一部分を、(たとえば、ビデオフレームのシーケンス中の)キャプチャされた前のビデオフレーム112の選択されたオブジェクトまたは部分と比較することによって、オブジェクトを検出するために、動きに基づくモデルではなく、オブジェクトモデルを使用し得る。オブジェクト検出器108は、ビデオフレーム内の複数のオブジェクトを検出するために使用され得る。

【0023】

[0092]オブジェクトトラッキングおよび検出モジュール104は、メモリバッファ110も含み得る。メモリバッファ110は、1つまたは複数のキャプチャされたフレームと、キャプチャされたビデオフレームと関連付けられるデータとを記憶し得る。一例では、メモリバッファ110は、前のキャプチャされたビデオフレーム112を記憶し得る。オブジェクトトラッキングおよび検出モジュール104は、動きに基づくトラッキングおよび/またはオブジェクト検出を実行する際に、キャプチャされた前のビデオフレーム112についての、メモリバッファ110から与えられたデータを使用し得る。ターゲットオブジェクトをより正確にトラッキングおよび/または検出するように動きに基づくトラッキングとオブジェクト検出とを調整するために、メモリバッファ110からのフィードバックを介して、データが動きトラッカー106またはオブジェクト検出器108に与えら

れ得る。たとえば、メモリバッファ 110 は、オブジェクトをトラッキングまたは検出するときにオブジェクトの位置とサイズとをより正確に特定するために使用され得る。1 つまたは複数のパラメータを動きトラッカー 106 およびオブジェクト検出器 108 に与えるために、位置とウィンドウサイズのデータを動きトラッカー 106 およびオブジェクト検出器 108 に与え得る。

【0024】

[0093] 上で述べられたように、電子デバイス 102 は動きに基づくトラッキングを実行し得る。動きに基づくトラッキングは、様々な方法を使用して実行され得る。一例では、トラッキングはメジアンフロー方法によって実行され、その方法では、動きトラッカー 106 は、画像のペア I_t 、 I_{t+1} （たとえば、ビデオフレーム）とバウンディングボックス b_t とを受け入れ、バウンディングボックス b_{t+1} を出力する。点のセットがバウンディングボックス b_t 内の矩形格子上で初期設定されてよく、 I_t と I_{t+1} との間の希薄な動きフローを生成するためにそれらの点がトラッキングされてよい。点予測の品質が推定されてよく、各点が誤差を割り当てられてよい。最悪の予測の一部分（たとえば、50%）はフィルタで除去されてよく、残りの予測は、バウンディングボックス全体の変位を推定するために使用される。動きトラッカー 106 は、電子デバイス 102 によってキャプチャされた各ビデオフレームに対する動きに基づくトラッキングを実行し得る。同様の方法において、動きに基づくトラッキングは、1 つまたは複数の勾配（たとえば、 x 勾配および y 勾配）を計算すること、および時間勾配を計算するためにフレームのペア間の差を使用すること、および現在のビデオフレーム内のターゲットオブジェクトを正確にトラッキングするために複数の勾配値を使用することによって実行され得る。動きに基づくトラッキングに関するさらなる詳細が、以下で与えられる。

【0025】

[0094] 動きに基づくトラッキングを実行するとき、動きトラッカー 106 は、動きトラッキング方法の計算または推定された精度に基づいてトラッキング信頼性値を決定し得る。いくつかの構成では、トラッキング信頼性値は、ターゲットオブジェクトが現在のビデオフレームまたはそのビデオフレームの定義されたウィンドウ内に入る可能性または確率に対応する、0 から 1 の間の実数であり得る。トラッキング信頼性値は、トラッキング閾値と比較され得る。トラッキング信頼性値がトラッキング閾値よりも大きい場合、ターゲットオブジェクトが現在のビデオフレーム内で見出される可能性は高いことがある。代替的に、トラッキング信頼性値がトラッキング閾値以下である場合、ターゲットオブジェクトが現在のビデオフレーム内で見出される可能性は低いまたは不確実であることがある。トラッキング信頼性値を決定するための様々な方法が使用され得る。一構成では、トラッキング信頼性値は、現在のビデオフレーム中のトラッキングされたウィンドウ（たとえば、トラッキングパッチウィンドウ）と、前にキャプチャされたビデオフレームからの前に記憶された画像パッチとの間の正規化相互相関（NCC: normalized cross correlation）を計算することによって決定される。トラッキング信頼性値を決定することに関するさらなる詳細が、以下で与えられる。

【0026】

[0095] 電子デバイス 102 は、オブジェクト検出も実行し得る。オブジェクト検出は、様々な方法を使用して実行され得る。一構成では、オブジェクト検出はスライディングウィンドウ方法を使用して実行され、その方法では、ターゲットオブジェクトが現在のビデオフレーム中で、または現在のビデオフレームの特定のウィンドウもしくはウィンドウのサブセット内で見出されるかどうかを決定するために、ビデオフレーム内のウィンドウの複数のサブセットのコンテンツが調査される。すべての可能なウィンドウ位置およびサイズのすべてまたはサブセットが、ビデオフレーム中で探索され得る。たとえば、各ウィンドウはデータのピクセルに対応してよく、オブジェクト検出器 108 は、ターゲットオブジェクトが特定のウィンドウまたはサブウィンドウ内にある信頼性のレベル（たとえば、バイナリインジケータ）を決定するために、データのピクセルを使用して 1 つまたは複数の計算を実行することができる。1 つまたは複数のウィンドウと関連付けられる信頼性の

レベルに基づいて、検出器の信頼性値が現在のビデオフレームのために取得され得る。さらに、オブジェクト検出の精度または効率を増加させるために、追加の技法が使用され得る。これらの技法のいくつかが、以下で説明される。

【 0 0 2 7 】

[0096]いくつかの構成では、動きトラッカー 1 0 6 およびオブジェクト検出器 1 0 8 は、並列にではなく順次的に動作し得る。たとえば、電子デバイス 1 0 2 は、選択されたオブジェクト（たとえば、ターゲットオブジェクト）の動きに基づくトラッキングを実行し、続いて、トラッキングされたパラメータに基づいて、選択されたオブジェクトのオブジェクト検出を実行し得る。一構成では、電子デバイス 1 0 2 は、現在のビデオフレームに対する動きに基づくトラッキングを実行し得る。電子デバイス 1 0 2 は次いで、トラッキングされたパラメータに基づいて、現在のフレームに対するオブジェクト検出を実行し得る。一構成では、トラッキングされたパラメータは、信頼性値と閾値との比較に基づき得る。たとえば、トラッキング信頼性値がトラッキング閾値を下回る場合、電子デバイス 1 0 2 はオブジェクト検出を実行し得る。代替的に、トラッキング信頼性値がトラッキング閾値を上回る場合、電子デバイス 1 0 2 は、現在のビデオフレームのためのオブジェクト検出をスキップし、現在のビデオフレームの動きトラッキング結果に基づいて次のビデオフレームに対する動きに基づくトラッキングを実行し続け得る。言い換えれば、オブジェクト検出は、動きに基づくトラッキングがあまり良好でなく、たとえば、トラッキング信頼性値がトラッキング閾値を下回るときのみ、実行され得る。オブジェクト検出が実行されるかどうかおよび / またはどのように実行されるかを考慮するとき、他のトラッキングされたパラメータが使用され得る。トラッキングされるパラメータの例は、ターゲットオブジェクトの領域、ウィンドウ位置、ウィンドウサイズ、スケールレベル、ターゲットサイズ、トラッキングおよび / もしくは検出信頼性値、または、ターゲットオブジェクトの効率的なトラッキングおよび / または検出を支援するために使用され得る他のパラメータを含み得る。

【 0 0 2 8 】

[0097]トラッキングされたパラメータに基づいて動きに基づくトラッキングとオブジェクト検出とを順次的に実行することは、電子デバイス 1 0 2 が、大量の計算を実行することなくビデオフレーム内のターゲットオブジェクトをトラッキングおよび / または検出することを可能にし得る。具体的には、動きに基づくトラッキングはオブジェクト検出よりも計算集約的でないことがあるので、電子デバイス 1 0 2 はオブジェクト検出の実行をスキップすることができ、ここで、現在のビデオフレーム内のターゲットオブジェクトを正確にトラッキングするために動きに基づくトラッキングが使用され得る。たとえば、トラッキング信頼性値が特定の目標閾値を超えると電子デバイス 1 0 2 が決定する場合、電子デバイス 1 0 2 は、現在のビデオフレーム内のターゲットオブジェクトの位置または存在を正確に決定するために、現在のビデオフレームに対してオブジェクト検出が必要とされないことを決定し得る。さらに、オブジェクト検出は多くの場合有益であり得るので、電子デバイス 1 0 2 は、ターゲットオブジェクトをより正確に検出するために、またはトラッキング閾値との比較に基づく、動きに基づくトラッキングが不十分である場合、オブジェクト検出を実行するために、オブジェクト検出が使用され得る場合を決定し得る。

【 0 0 2 9 】

[0098]いくつかの構成では、現在のビデオフレームに対するオブジェクト検出をスキップするのではなく、メモリバッファ 1 1 0 によって与えられる動きに基づくトラッキングの結果および / または追加の情報が、オブジェクト検出を実行するプロセスを限定または調整するために使用され得る。たとえば、ターゲットオブジェクトが、動きに基づくトラッキング方法を使用して正確にトラッキングされ得ない場合、電子デバイス 1 0 2 は、依然として、動きに基づくトラッキングを介して与えられたパラメータがない場合よりも少ない計算能力を使用してオブジェクトをより正確に検出するためにオブジェクト検出中に使用され得る、ターゲットオブジェクトと関連付けられる位置、ウィンドウスケールまたは他のトラッキングされたパラメータについての情報を推定または取得し得る。したがっ

10

20

30

40

50

て、動きに基づくトラッキングが、トラッキング閾値を超えるトラッキング信頼性値を与えない場合でも、動きに基づくトラッキングの結果は、その後にオブジェクト検出を実行するときに使用され得る。

【0030】

[0099]電子デバイス102上のビューファインダ131は、第1のトラッキングエリア133と第2のトラッキングエリア135とを含み得る。第1のトラッキングエリア133と第2のトラッキングエリア135の両方が、タッチスクリーン116を使用してユーザによって指定され得る。たとえば、ユーザは、第1のトラッキングエリア133および第2のトラッキングエリア135の所望の位置へと、タッチスクリーン116上の焦点リングをドラッグすることができる。必要とはされないが、トラッキングエリアの1つは静止していてもよい。たとえば、第1のトラッキングエリア133はオブジェクト（たとえば、歩いている人）をトラッキングすることができ、第2のトラッキングエリア135は静止している木をカバーすることができる。一構成では、第2のトラッキングエリア135は、電子デバイス102上のタッチスクリーン116全体をカバーすることができる。

【0031】

[00100]電子デバイス102は、ビデオ処理モジュール137を含み得る。ビデオ処理モジュール137は、重複（オーバーラップ）143を含み得る。重複143は、第1のトラッキングエリア133と第2のトラッキングエリア135の重複（オーバーラップ）の量を反映し得る。たとえば、第1のトラッキングエリア133および第2のトラッキングエリア135が互いにまったく重複しない場合、重複143は0%であり得る。同様に、第1のトラッキングエリア133が第2のトラッキングエリア135と完全に重複する場合（または、どちらのトラッキングエリアがより大きいかによっては、第2のトラッキングエリア135が第1のトラッキングエリア133と完全に重複する場合）、重複143は100%であり得る。ビデオ処理モジュール137は、閾値145を含み得る。図13に関して説明されるように、重複143は、ビデオ処理が実行されるべきかどうかを決定するために、閾値145と比較され得る。

【0032】

[00101]ビデオ処理モジュール137はまた、画面区分147の機能を含み得る。たとえば、画面区分147は、図11、図13、図15、および図21に関して説明されるように、第1のトラッキングエリア133および第2のトラッキングエリア135と関連付けられる個々のビデオストリームを表示するために、ビューファインダ132を複数のウィンドウへと区分することができる。ビデオ処理モジュール137はまた、幾何学的配置（geometry）トラッキング149の機能を含み得る。たとえば、幾何学的配置トラッキング149の機能は、図17および図18に関して説明されるように、第1のトラッキングエリア133と第2のトラッキングエリア135との間の幾何学的配置をトラッキングすることができる。幾何学的配置は、ビューファインダ131上に表示され得る。ビデオ処理モジュール137はまた、クラスタ情報151の機能を含み得る。クラスタ情報151の機能は、幾何学的配置と関連付けられる少なくとも1つのパラメータ（たとえば、重心）に基づいてクラスタを生成することができる。たとえば、各クラスタは、図17および図18に関して説明されるように、第1のトラッキングエリア133と第2のトラッキングエリア135との間で実質的に同様の幾何学的配置を有するビデオフレームを含み得る。

【0033】

[00102]図2Aを参照すると、オブジェクトトラッキングおよび検出モジュール204を示すブロック図が示されている。オブジェクトトラッキングおよび検出モジュール204は、電子デバイスまたはワイヤレスデバイス内に実装され得る。オブジェクトトラッキングおよび検出モジュール204は、オブティカルフローモジュール226とトラッキング信頼性値228とを有する動きトラッカー206を含み得る。オブジェクトトラッキングおよび検出モジュール204は、スキャナロケータ230と、スキャナスケラ236と、分類器238と、検出信頼性値240とを有するオブジェクト検出器208も含み得

る。メモリバッファ 210 は、動きトラッカー 206 とオブジェクト検出器 208 とに与えられ得るキャプチャされた前のビデオフレーム 212 と関連付けられるデータを記憶し得る。オブジェクトトラッキングおよび検出モジュール 204、動きトラッカー 206、オブジェクト検出器 208、ならびにメモリバッファ 210 は、図 1 に関して上記で説明されたオブジェクトトラッキングおよび検出モジュール 104、動きトラッカー 106、オブジェクト検出器 108、ならびにメモリバッファ 110 の構成であり得る。

【0034】

[00103] 動きトラッカー 206 は、現在のビデオフレーム (N) 224 に対する動きに基づくトラッキングを実行するために使用され得る。たとえば、前のビデオフレーム (N-1) 222 および現在のビデオフレーム (N) 224 が (たとえば、電子デバイス 102 によって) 受信され得る。前のビデオフレーム (N-1) 222 は、ビデオフレームのシーケンス中の現在のビデオフレーム (N) 224 の直前にあり得る。追加のビデオフレームが、オブジェクトトラッキングおよび検出モジュール 204 によって取得され処理され得る。前のビデオフレーム (N-1) 222 は、動きトラッカー 206 に与えられ得る。さらに、メモリバッファ 210 は、本明細書ではキャプチャされた前のビデオフレーム 212 と呼ばれる、前のビデオフレーム (N-1) 222 と関連付けられるデータを記憶し得る。いくつかの構成では、メモリバッファ 210 は、電子デバイス 102 から (たとえば、カメラから) 前のビデオフレーム (N-1) 222 についての情報を直接取得し得る。メモリバッファ 210 はまた、融合モジュール 260 から前のビデオフレーム (N-1) 222 についてのトラッキング結果を取得することができ、トラッキング結果は、オブジェクトが前のビデオフレーム (N-1) 222 の中でどこでトラッキングおよび / または検出されたかを指定し得る。前のビデオフレーム (N-1) 222 または他の前にキャプチャされたビデオフレームについてのこの情報は、メモリバッファ 210 に記憶され得る。

【0035】

[00104] 動きトラッカー 206 はその後、ビデオフレームのシーケンス中の現在のビデオフレーム (N) 224 を受信し得る。動きトラッカー 206 は、(たとえば、メモリバッファ 210 から与えられた情報を使用して) 現在のビデオフレーム (N) 224 を前のビデオフレーム (N-1) 222 と比較し得る。動きトラッカー 206 は、オブティカルフローモジュール 226 を使用して、現在のビデオフレーム (N) 224 上のオブジェクトの動きをトラッキングし得る。オブティカルフローモジュール 226 は、現在のビデオフレーム (N) 224 上のオブジェクトの動きに基づくトラッキングを実行するためのハードウェアおよび / またはソフトウェアを含み得る。前のビデオフレーム (N-1) 222 と現在のビデオフレーム (N) 224 とを比較することによって、動きトラッカー 206 は、ターゲットオブジェクトが現在のビデオフレーム (N) 224 中にある可能性と関連付けられるトラッキング信頼性値 228 を決定し得る。一例では、トラッキング信頼性値 228 は、ターゲットオブジェクトが現在のビデオフレーム (N) 224 内に、または現在のビデオフレーム (N) 224 内のウィンドウ内にある確実性の百分率に基づく (たとえば、0 から 1 の間の) 実数である。

【0036】

[00105] オブジェクト検出器 208 は、現在のビデオフレーム (N) 224 上のオブジェクトを検出するために使用され得る。たとえば、オブジェクト検出器 208 は、ビデオフレームのシーケンス中の現在のビデオフレーム (N) 224 を受信し得る。オブジェクト検出器 208 は、トラッキングされたパラメータに基づいて現在のビデオフレーム (N) 224 に対するオブジェクト検出を実行し得る。トラッキングされたパラメータは、ターゲットオブジェクトが正確にトラッキングされている可能性に対応するトラッキング信頼性値 228 を含み得る。より具体的には、トラッキングされたパラメータは、トラッキング信頼性値 228 とトラッキング閾値 250 との比較を含み得る。トラッキングされたパラメータは、メモリバッファ 210 から与えられた情報も含み得る。オブジェクトを検出するときに使用され得るトラッキングされたパラメータのいくつかの例は、領域、ウィ

ンドウ位置、ウィンドウサイズ、またはオブジェクト検出を実行するときにパラメータとしてオブジェクト検出器 208 によって使用され得る他の情報を含む。

【0037】

[00106]オブジェクト検出器 208 は、スキャナロケータ 230 を含み得る。スキャナロケータ 230 は、ウィンドウ位置選択器 232 とランダムマイザ 234 とを含み得る。ウィンドウ位置選択器 232 は、ビデオフレーム内の複数のウィンドウを選択し得る。たとえば、ビデオフレームは、関連する位置とサイズとを各々が有する、複数のウィンドウを含み得る。一構成では、各ビデオフレームは、ビデオフレーム中の総ピクセルの一部を各々が含む、複数の（たとえば、約 10,000 個の）重複するウィンドウに分割される。代替的に、任意の好適な数のウィンドウがあつてよく、それらは重複しないことがある。スキャナロケータ 230 内のウィンドウ位置選択器 232 は、ターゲットオブジェクトを識別することをその中で試みる、ウィンドウの位置を選択し得る。ランダムマイザ 234 は、オブジェクトを検出するために異なるサイズおよび位置のウィンドウをランダムに選択し得る。いくつかの構成では、ランダムマイザ 234 は、ビデオフレーム内のウィンドウをランダムに選択する。代替的に、ランダムマイザ 234 は、1 つまたは複数の要因に基づいてウィンドウをより正確に選択し得る。たとえば、ランダムマイザ 234 は、オブジェクトが大抵位置する領域、サイズまたは概略的な位置に基づいて、ウィンドウの選択を限定し得る。この情報は、メモリバッファ 210 を介して取得されてよく、または、完全に依拠するのに十分正確ではないが、オブジェクト検出を実行するときに役立つ情報を与え得る、動きに基づくトラッキングを介して取得され得る。したがって、ランダムマイザ 234 は、探索すべき複数のウィンドウをランダムに選択し得るが、ウィンドウの選択は、オブジェクト検出器 208 に与えられた情報に基づいて限定され、したがって完全にランダムではないことがある。

【0038】

[00107]オブジェクト検出器 208 は、あるサイズのウィンドウを引き出すかまたは選択するために使用され得る、スキャナスケラ 236 も含み得る。ウィンドウサイズは、オブジェクトを検出するとき、または画像が特定のウィンドウ内にあるかどうかを検出するためにウィンドウの選択を元の画像と比較するとき、ウィンドウのサイズを狭くするためにスキャナロケータ 230 によって使用され得る。スキャナスケラ 236 は、オブジェクトを定義するとき、初めにあるサイズまたはスケールレベルの 1 つもしくは複数のウィンドウを選択し、または代替的に、メモリバッファ 210 から与えられた情報に基づいて、あるサイズまたはスケールレベルの 1 つもしくは複数のウィンドウを引き出し得る。

【0039】

[00108]分類器 238 は、ターゲットオブジェクトの一部またはすべてが特定のウィンドウ中で見出されるどうかを決定するために使用され得る。いくつかの構成では、分類器 238 は、ターゲットオブジェクトが特定のウィンドウまたはサブウィンドウ内で検出されるかどうかを示すために、各ウィンドウについてバイナリ値を生成し得る。この分類（たとえば、バイナリ分類）は、オブジェクト検出器 208 によって探索された各ウィンドウに対して実行され得る。具体的には、分類器 238 は、オブジェクトが検出される各ウィンドウに対してバイナリ 1 を生成し、オブジェクトが検出されない各ウィンドウに対してバイナリ 0 を生成し得る。1 および 0 の数または組合せに基づいて、オブジェクト検出器 208 は、ターゲットオブジェクトが現在のビデオフレーム（N）224 内に存在する可能性を示す検出信頼性値 240 を決定し得る。いくつかの構成では、検出信頼性値 240 は、オブジェクトが正確に検出された百分率または確率を示す 0 から 1 の間の実数である。

【0040】

[00109]オブジェクト検出器 208 は、領域、ターゲットサイズ、ウィンドウサイズ、スケールレベル、ウィンドウ位置、および 1 つまたは複数の信頼性値を含む、様々なトラッキングされたパラメータに従ってオブジェクト検出を実行し得る。ビデオフレームのウィンドウまたはウィンドウのサブセットが探索され、オブジェクト検出器 208 が、各々

の探索されたウィンドウに対するバイナリ値を取得すると、オブジェクト検出器 208 は、最高の信頼性を有する現在のビデオフレーム上のウィンドウサイズならびに位置または領域を決定し得る。この位置およびウィンドウサイズは、ターゲットオブジェクトをより正確にトラッキングおよび / または検出するために後のトラッキングおよび検出において使用され得る。

【0041】

[00110] 様々な技法が、ターゲットオブジェクトを検出する際にオブジェクト検出器 208 によって使用され得る。一構成では、ターゲットオブジェクトを検出することは、あらゆる可能なウィンドウ位置およびあらゆる可能なウィンドウサイズにおいてウィンドウのためにバイナリ分類を実行することを含み得る。しかしながら、あらゆる可能なウィンドウを探索することはリソース集約的である。したがって、別の構成では、オブジェクト検出器は、ビデオフレーム中のすべての可能なウィンドウではなく、ウィンドウ位置およびサイズのサブセットを探索し得る。たとえば、オブジェクト検出器 208 は、すべての可能なウィンドウの 1 % を探索し得る。次いで、検出が不成功である（たとえば、検出信頼性値 240 が検出閾値 252 よりも小さい）場合、後のキャプチャされたフレームにおいて、より高い百分率のウィンドウ位置、たとえば 2 % が探索され得る。探索されるウィンドウ位置の百分率の刻みは、均一であってよく、不均一であってよく、低速または高速であってよく、すなわち、連続するフレームが、1 %、2 %、3 %、4 % または 1 %、2 %、4 %、8 % を有し得る。一構成では、探索されるフレームの百分率は、高い検出信頼性値に回答して極めて高く（たとえば、80 %、90 %、100 %）、すなわち、ターゲットオブジェクトが次のビデオフレームであることを保証するように、設定され得る。たとえば、探索されるフレームの百分率は、検出およびトラッキング閾値 256 を超える検出およびトラッキング信頼性値に回答して、少なくとも 80 % に飛び得る。代替的に、百分率は、60 %、70 %、90 % などに飛び得る。加えて、検出およびトラッキング閾値の任意の好適な値、たとえば、0.6、0.65、0.7、0.75、0.8、0.85 などが使用され得る。さらに、探索されるウィンドウの百分率は、ランダムイザ 234（乱数発生器）に基づいてランダムに決定されてよく、たとえば、1 % から 15 % の間のランダムな百分率のウィンドウが、キャプチャされたフレーム中で探索され得る。すべてのウィンドウ位置のサブセットを探索することによって、オブジェクト検出は、電子デバイス 102 において少数のリソースを使用し得る。

【0042】

[00111] 本明細書で説明される技法は、各位置に対してウィンドウサイズのサブセットを探索し得る。各ウィンドウサイズは、本明細書ではスケールレベルと呼ばれることがあり、各スケールレベルは特定のウィンドウサイズに対応する。たとえば、20 個の可能なスケールレベルがあり得る。すべての 20 個のスケールレベルを探索するのではなく、スケールレベルまたはウィンドウサイズのサブセットが各ウィンドウ位置において探索され得る。

【0043】

[00112] 本明細書で説明される技法は、探索されるウィンドウ位置およびサイズを調整するためにメモリバッファ 210 からのフィードバックも使用し得る。言い換えれば、その中でターゲットオブジェクトが正常に検出および / またはトラッキングされた、最後のキャプチャされたビデオフレームの位置およびサイズが、現在のビデオフレーム (N) 224 を探索するための開始点として使用され得る。たとえば、ターゲットオブジェクトが最近のビデオフレーム中で検出され、トラッキングされた（すなわち、最近のキャプチャされたビデオフレームのための検出およびトラッキング信頼性値 256 が検出およびトラッキング閾値を上回る）場合、スキャナロケータは、最近のフレームと関連付けられる位置およびサイズにおいて、現在のキャプチャされたフレームを探索し始め得る。たとえば、ターゲットオブジェクトが、光学系の視野外に移動する場合、またはある距離において消える場合、ターゲットオブジェクトは、ターゲットオブジェクトが光学系の視野から出たとき、またはある距離において消えたときと同じサイズで再び現れる可能性が高いこと

がある。したがって、あるサイズまたはある範囲のサイズが、オブジェクト検出を実行するときに後のビデオフレーム中のターゲットオブジェクトを検出するために予測され得る。

【 0 0 4 4 】

[00113]キャプチャされたビデオフレーム (N) 2 2 4 中で探索されるウィンドウ位置およびウィンドウサイズの探索範囲が、最近のビデオフレーム (たとえば、前のビデオフレーム (N - 1) 2 2 2) 中のターゲットオブジェクトと関連付けられるウィンドウ位置およびウィンドウサイズと同様のものに限定され得る。本明細書で使用される「探索範囲」という用語は、ビデオフレーム中のターゲットオブジェクトを検出および/またはトラッキングするときに利用され得る候補ウィンドウ位置または候補ウィンドウサイズ (または両方) のセットを指す。たとえば、探索されるウィンドウ位置のサブセットは、ターゲットオブジェクトが最近のビデオフレーム中のどこで見出されたかに基づいて現在のビデオフレーム (N) 2 2 4 の一部分、たとえば、現在のビデオフレーム (N) 2 2 4 の4分の1または半分のうちの1つの中から選択され得る。言い換えれば、探索空間は、ターゲットオブジェクトが最後にトラッキングまたは検出されたところの近くに限定され得る。同様に、各ウィンドウ位置のために探索されるフレームのサイズは、ターゲットにされたオブジェクトが見出された、最近のビデオフレーム中のウィンドウのサイズに基づいて限定され得る。たとえば、オブジェクトが、8というスケールレベルを有するウィンドウを使用して最近のフレーム中で検出された場合、スキャナスケラ 2 3 6 は、8プラスまたはマイナス3という、現在のビデオフレーム (N) 2 2 4 のためのウィンドウスケールレベル、すなわち、スケールレベル5 ~ 11のみを選択し得る。これはさらに、低い確率の探索をなくし、オブジェクト検出の効率を増加させ得る。代替的に、最近の (非現在の) ビデオフレームがターゲットオブジェクトを検出しなかった (すなわち、最近のビデオフレームのための検出およびトラッキング信頼性値 2 5 6 が検出およびトラッキング閾値を下回る) 場合、オブジェクト検出器 2 0 8 は、探索される探索空間 (ウィンドウ位置) を拡大することができ、たとえば、画像のより広い範囲または画像全体が探索の対象になり得る。

【 0 0 4 5 】

[00114]オブジェクトトラッキングおよび検出モジュール 2 0 4 は、複数のウィンドウを統合して単一のウィンドウを形成するための融合モジュール 2 6 0 を含み得る。初めに2つの信頼性値、すなわち、オブジェクト検出器 2 0 8 からの検出信頼性値 2 4 0 および動きトラッカー 2 0 6 からのトラッキング信頼性値 2 2 5 がある。融合モジュール 2 6 0 は、2つの信頼性値を組み合わせて (たとえば、より大きい信頼性値を選んで)、検出およびトラッキング信頼性値 2 5 6 にすることができる。検出およびトラッキング信頼性値 2 5 6 は、ターゲットオブジェクトがビデオフレーム上で識別されたかどうかを示し得る。一構成では、検出およびトラッキング信頼性値 2 5 6 は0から1の間の実数であってよく、ただし、0は、ターゲットオブジェクトが特定のビデオフレーム中で識別されたことの、最低のあり得る信頼性を示し、1は、ターゲットオブジェクトが特定のビデオフレーム中で識別されたことの、最高のあり得る信頼性を示す。言い換えれば、検出およびトラッキング信頼性値 2 5 6 は、ターゲットオブジェクトが見つけられた可能性の全体的な指示として働き得る。さらに、検出およびトラッキング信頼性値 2 5 6 は、次のビデオフレーム中で探索すべきウィンドウ位置、ウィンドウサイズ、またはウィンドウの百分率を決定するために使用されるパラメータであり得る。融合モジュール 2 6 0 は、現在のビデオフレーム (N) 2 2 4 についての情報をメモリバッファ 2 1 0 に与えるために使用され得る。一例では、融合モジュール 2 6 0 は、トラッキングされたウィンドウ 2 4 2 についての情報 (たとえば、ウィンドウ位置 2 4 4、ウィンドウサイズ 2 4 6 など) と検出およびトラッキング信頼性値 2 5 6 とをメモリバッファ 2 1 0 に与え得る。融合モジュール 2 6 0 は、組み合わせられたトラッキング結果 (たとえば、バウンディングボックス) を形成して検出およびトラッキング信頼性値 2 5 6 を計算するために、動きトラッカー 2 0 6 およびオブジェクト検出器 2 0 8 からのトラッキング結果 (たとえば、バウンディングボック

10

20

30

40

50

ス)を使用し得る。

【0046】

[00115]メモリバッファ210は、前のビデオフレーム(N-1)222、現在のビデオフレーム(N)224、または他のキャプチャされたビデオフレームと関連付けられる1つまたは複数の値を記憶し得る。一構成では、メモリバッファ210は、前のビデオフレーム(N-1)222に対応する情報を含み得る、キャプチャされた前のビデオフレーム212を記憶する。キャプチャされた前のビデオフレーム212は、各ウィンドウ242について位置244と、ウィンドウサイズ246と、(たとえば、分類器238からの)バイナリ決定248とを含む、1つまたは複数のウィンドウ242についての情報を含み得る。キャプチャされた前のビデオフレーム212はまた、トラッキング閾値250と、検出閾値252と、検出およびトラッキング閾値254とを含み得る。トラッキング閾値250は、トラッキング信頼性レベルがトラッキング閾値250よりも大きいかどうかを決定するために(258)、動きトラッカー206またはオブジェクトトラッキングおよび検出モジュール204上の回路(たとえば、信頼性レベル比較器)に与えられ得る。検出閾値252は、検出信頼性値240が検出閾値252よりも大きいかどうかを決定するために、オブジェクト検出器208またはオブジェクトトラッキングおよび検出モジュール204上の他の回路に与えられ得る。検出およびトラッキング閾値254は、トラッキング閾値250および検出閾値252に基づく組み合わせられた値であり得る。検出およびトラッキング閾値254は、動きに基づくトラッキングおよびオブジェクト検出のための組み合わせられた信頼性値を決定するために、検出およびトラッキング信頼性値256と比較され得る。閾値の各々は、ターゲットオブジェクトがビデオフレーム内に位置する可能性に基づき得る。オブジェクトトラッキングおよび検出モジュール204は、特定の検出およびトラッキング信頼性値256が取得されるまで、現在のビデオフレーム(N)224に対する動きに基づくトラッキングおよび/または検出を実行し得る。さらに、動きに基づくトラッキングおよびオブジェクト検出は、複数のビデオフレームのシーケンス中の各ビデオフレームに対して実行され得る。

【0047】

[00116]動きに基づくトラッキングおよびオブジェクト検出を実行することは、動きに基づくトラッキングの後に、トラッキングされたパラメータに基づくオブジェクト検出を順次的に実行することを含み得る。特に、本システムおよび方法は、2ステップのトラッキングおよび検出手法を実施し得る。動きに基づくトラッキングが、使用されるオブジェクト検出としての実際のオブジェクト識別ではなく、シーンの相対的な動きに基づくので、動きに基づくトラッキングは、電子デバイスにおいて、オブジェクト検出を実行することほどリソース集約的でないことがある。したがって、オブジェクト検出器208の代わりに動きトラッカー206を使用することがより効率的であることがあり、ここで、ターゲットオブジェクトが、オブジェクト検出を実行することもなく正確にトラッキングされ得る。

【0048】

[00117]したがって、オブジェクト検出器208と並列に動きトラッカー206を使用するのではなく、オブジェクトトラッキングおよび検出モジュール204は、動きトラッカー206が不十分である場合にオブジェクト検出器208を使用するにすぎず、すなわち、動きトラッキングおよびオブジェクト検出は(実行されるとしても)並列にではなく順次的に実行される。トラッキングが実行される各ビデオフレームに対して、動きトラッカー206は、ターゲットオブジェクトが現在のビデオフレーム(N)224中にある可能性を示す0から1の間の実数であり得る、トラッキング信頼性値228を生成し得る。

【0049】

[00118]2ステップのトラッキングおよび検出手法の一構成では、動きトラッカー206は、最初に現在のビデオフレーム(N)224に対する動きに基づくトラッキングを実行し得る。動きトラッカー206は、動きに基づくトラッキングのプロセスに基づいてトラッキング信頼性値228を決定し得る。トラッキング信頼性値228とメモリバッファ

210によって与えられたトラッキング閾値250とを使用して、オブジェクトトラッキングおよび検出モジュール204内の回路（たとえば、信頼性レベル比較器）は、トラッキング信頼性値228がトラッキング閾値250を超えるかどうかを決定し得る（258）。トラッキング信頼性値228がトラッキング閾値250よりも大きい場合、オブジェクトトラッキングおよび検出モジュール204は、オブジェクト検出の実行をスキップし、出力262を生成するために融合モジュール260にトラッキング結果を与え得る。出力262は、ターゲットオブジェクトが現在のビデオフレーム（N）224内にあるという指示を含み得る。さらに、出力262は、ターゲットオブジェクトについての追加の情報を含み得る。

【0050】

10

[00119]トラッキング信頼性値228がトラッキング閾値250を超えない場合、オブジェクト検出器208はその後、現在のビデオフレーム（N）224に対するオブジェクト検出を実行し得る。オブジェクト検出は、現在のビデオフレーム（N）224内のウィンドウのすべてまたはサブセットに対して実行され得る。オブジェクト検出器208はまた、動きに基づくトラッキングの結果および/またはメモリバッファ210から与えられた情報に基づいて、ウィンドウのサブセット、ウィンドウサイズ、または他の検出基準を選択し得る。オブジェクト検出は、オブジェクト検出器208に与えられた1つまたは複数のトラッキングされたパラメータに基づいて、幾分ロバストなプロセスを使用して実行され得る。オブジェクト検出器208は、検出信頼性値240を決定し、検出信頼性値240を検出閾値252と比較し得る。検出信頼性値240が検出閾値252を上回る場合、オブジェクト検出器208は、出力262を生成するために融合モジュール260に検出結果を与え得る。出力262は、ターゲットオブジェクトが現在のビデオフレーム（N）224内にあるという指示を含み、および/または検出されたオブジェクトについての追加の情報を含み得る。

20

【0051】

[00120]代替的に、検出信頼性値240が検出閾値252以下である場合、オブジェクト検出器208は、現在のビデオフレーム（N）224内のより多数のウィンドウを探索することのような、よりロバストな方法を使用してオブジェクト検出を再び実行し得る。オブジェクト検出器208は、満足な検出信頼性値240が取得されるまで、オブジェクト検出のプロセスを繰り返し得る。満足な検出信頼性値240が取得され、その結果、現在のビデオフレーム内のターゲットオブジェクトが識別されると、オブジェクトトラッキングおよび検出モジュール204が、次のビデオフレームに対するトラッキングおよび検出を実行するために使用され得る。

30

【0052】

[00121]図2Bを参照すると、オブジェクトトラッキングおよび検出モジュール204内のコンポーネントを実装するプロセッサ264の特定の例示的な実施形態が示されている。図2Aに示されるように、オブジェクトトラッキングおよび検出モジュール204は、プロセッサ264によって実装され得る。異なるプロセッサが、異なるコンポーネントを実装するために使用され得る（たとえば、あるプロセッサは動きトラッカー206を実装することができ、別のプロセッサはオブジェクト検出器208を実装するために使用されてよく、また別のプロセッサはメモリバッファ210を実装するために使用されてよい）。

40

【0053】

[00122]図3を参照すると、動きに基づくトラッキングとオブジェクト検出とを実行するための方法300の特定の例示的な実施形態のフローチャートが示されている。方法300は、電子デバイス102、たとえば、オブジェクトトラッキングおよび検出モジュール104によって実装され得る。電子デバイス102は、前のビデオフレーム（N-1）222と現在のビデオフレーム（N）224とを比較することによって、現在のビデオフレーム（N）224のための動きに基づくトラッキングを実行することができる（302）。オブジェクトをトラッキングすることは、画像のペア間で点をトラッキングすること

50

によってメジアンフロー方法を使用して実行され得る。動きに基づくトラッキングの他の方法も使用され得る。加えて、動きに基づくトラッキングは、メモリバッファ 110 を介して与えられたキャプチャされた前のビデオフレーム 112 についての情報を使用して現在のビデオフレーム (N) 224 のために実行され得る。

【0054】

[00123] 電子デバイス 102 は、トラッキング信頼性値 228 を決定することができる (304)。トラッキング信頼性値 228 は、ターゲットオブジェクトが正確にトラッキングされた可能性または確実性を示し得る。電子デバイス 102 は、トラッキング信頼性値 228 がトラッキング閾値 250 よりも大きいかどうかを決定することができる (306)。トラッキング信頼性値 228 がトラッキング閾値 250 よりも大きい場合、電子デバイス 102 は、次のビデオフレームのための動きに基づくトラッキングを実行することができる (308)。さらに、電子デバイス 102 は、動きに基づくトラッキングの結果に基づいて、現在のビデオフレーム (N) 224 に対するオブジェクト検出を実行することをスキップし得る。言い換えれば、オブジェクト検出は、動きトラッキングがあまり良好でないときのみ、すなわち、トラッキング信頼性値 228 がトラッキング閾値 250 よりも大きくない場合に、現在のビデオフレーム (N) 224 のために実行され得る。しかしながら、トラッキング信頼性値 228 がトラッキング閾値 250 よりも大きくない場合、電子デバイス 102 は、現在のビデオフレーム (N) 224 のためのオブジェクト検出を実行することができる (310)。電子デバイス 102 は、動きに基づくトラッキングに続いてオブジェクト検出を実行し得る。いくつかの構成では、オブジェクト検出は、より高い検出信頼性値 240 を取得するために、異なるロバスト性とともに複数回実行され得る。

【0055】

[00124] 図 4 を参照すると、動きに基づくトラッキングを実行するための方法 400 の特定の例示的な実施形態のフローチャートが示されている。方法 400 は、電子デバイス 102、たとえば、オブジェクトトラッキングおよび検出モジュール 104 によって実装され得る。電子デバイス 102 は、バウンディング (bounding) ボックスを使用してターゲットオブジェクトを識別することができる (402)。オブジェクトを識別すること (402) は、タッチスクリーン 116、または対象のオブジェクトがそれにより選択される他の入力方法を使用して、手動で実行され得る。複数のオブジェクトが、同様の方法で識別され得る。さらに、トラッキングされるべきオブジェクトを識別するために、他の入力方法が使用され得る。一例では、オブジェクトは、ターゲットオブジェクトの周りにバウンディングボックスを手動で描くことによって識別される。

【0056】

[00125] 電子デバイス 102 は、バウンディングボックス内の格子上に点を初期設定することができる (404)。格子上の点は、バウンディングボックス全体にわたって一様に離隔され得る。さらに、2つの画像 (たとえば、前のビデオフレーム (N-1) 222 と現在のビデオフレーム (N) 224) 間の格子上の点がトラッキングされ得る (406)。一例では、それらの点は、画像間の希薄な動きフローを生成する Lucas-Kanade トラッカーによってトラッキングされる。電子デバイス 102 は、2つの画像 (たとえば、前のビデオフレーム (N-1) 222 と現在のビデオフレーム (N) 224) 間のトラッキング誤差を推定することができる (408)。トラッキング誤差を推定すること (408) は、トラッキングされた点の各点に誤差値を割り当てることを含み得る。さらに、トラッキング誤差を推定すること (408) は、たとえば、前方後方誤差と、正規化相互相関 (NCC) と、2乗差分和とを含む、様々な方法を使用して実行され得る。推定されたトラッキング誤差は、トラッキング信頼性値 228 を取得し、最終的に、ターゲットオブジェクトが現在のビデオフレーム (N) 224 中にある可能性を決定するために、使用され得る。一構成では、トラッキング信頼性値 228 は、現在のビデオフレーム (N) 224 中のトラッキングされたウィンドウと前のビデオフレーム (N-1) 222 中のトラッキングされたウィンドウとの間の正規化相互相関 (NCC) を計算することによ

って取得され得る。トラッキング誤差はまた、図 5 に関して以下でより詳細に説明される前方後方誤差推定を含む、追加の技法を使用して推定され得る。さらに、電子デバイス 102 は、範囲外の点 (outlying point) の予測をフィルタで除去することができる (410)。たとえば、電子デバイスは、最悪の予測の 50% をフィルタで除去し得る。残りの予測は、バウンディングボックスの変位を推定するために使用され得る。

【0057】

[00126] 電子デバイス 102 は、バウンディングボックスを更新することができる (412)。バウンディングボックスを更新すること (412) は、更新されたバウンディングボックスが次のビデオフレームのための新しいバウンディングボックスになるように、実行され得る。動きに基づくトラッキングのプロセスは次いで、次のビデオフレームのために繰り返されてよく、または、トラッキング信頼性値 228 がトラッキング閾値 250 以下である場合、動きに基づくトラッキングのプロセスは、ターゲットオブジェクトが正確にトラッキングされ得るまで、次のビデオフレームに対しては中止され得る。いくつかの構成では、現在のビデオフレーム (N) 224 のための動きに基づくトラッキングが満足な結果を与えない場合、電子デバイス 102 は、ターゲットオブジェクトを位置特定する際に信頼性のより高いレベルを取得するために現在のビデオフレーム (N) 224 に対するオブジェクト検出を実行し得る。いくつかの構成では、動きに基づくトラッキングが満足な結果を生成することができない場合 (たとえば、ターゲットオブジェクトがビデオフレームの範囲外に移動したとき)、オブジェクト検出は、ターゲットオブジェクトが検出されるまで、任意の後のビデオフレームに対して実行され得る。

【0058】

[00127] 図 5 を参照すると、前方後方誤差に基づいて動きに基づくトラッキングにおけるトラッキング誤差を推定するための方法 500 の特定の例示的な実施形態のフローチャートが示されている。方法 500 は、電子デバイス 102 (たとえば、オブジェクトトラッキングおよび検出モジュール 104) によって実施され得る。いくつかの構成では、電子デバイス 102 は、トラッキングされたウィンドウ間の正規化相互相関 (NCC) を計算し得る。正規化相互相関 (NCC) は、トラッキング信頼性値 228 を決定するために使用され得る。電子デバイス 102 はまた、正規化相互相関 (NCC) を補足する様々なトラッキング誤差推定技法 (たとえば、前方後方誤差、2 乗差分和) を使用し得る。前方後方誤差推定を使用する例では、電子デバイス 102 は、前方軌道 (forward trajectory) を決定するために前のビデオフレーム (N-1) 222 と現在のビデオフレーム (N) 224 との間で前方トラッキングを実行することができる (502)。前方トラッキングは、k 個のステップ前方の画像をトラッキングすることを含み得る。得られた前方軌道は $(x_t, x_{t+1}, \dots, x_{t+k})$ に等しくてよく、ただし、 x_t は時間的な点位置であり、k は画像のシーケンスの長さを示す。電子デバイス 102 は、後方軌道 (backward trajectory) を決定するために現在のビデオフレーム (N) 224 と前のビデオフレーム (N-1) 222 との間で後方トラッキングを実行することができる (504)。得られる後方軌道は、

【0059】

【数 1】

$$(\hat{x}_t, \hat{x}_{t+1}, \dots, \hat{x}_{t+k}), \text{ ここで } \hat{x}_{t+k} = x_{t+k},$$

【0060】

に等しくてよい。

【0061】

[00128] 電子デバイス 102 は、前方軌道と後方軌道との間の前方後方誤差を決定することができる (506)。前方後方誤差は、前方軌道と後方軌道との間の距離として定義され得る。さらに、様々な距離が軌道比較のために定義され得る。一構成では、検証軌道

の始点と終点との間のユークリッド距離が、前方後方誤差を決定するときに使用され得る。一構成では、前方後方誤差は、トラッキング信頼性値 2 2 8 を決定するために使用され得る、トラッキング誤差として使用され得る。

【 0 0 6 2 】

[00129] 図 6 を参照すると、オブジェクト検出を実行するための方法 6 0 0 の特定の例示的な実施形態のフローチャートが示されている。方法 6 0 0 は、電子デバイス 1 0 2 (たとえば、オブジェクトトラッキングおよび検出モジュール 1 0 4) によって実装され得る。電子デバイス 1 0 2 は、現在のビデオフレーム (N) 2 2 4 中のウィンドウ位置およびサイズのサブセットを探索することによって、現在のビデオフレーム (N) 2 2 4 に対するオブジェクト検出と動きに基づくトラッキングとを実行することができる (6 0 2)

10

【 0 0 6 3 】

[00130] 電子デバイス 1 0 2 は、検出およびトラッキング信頼性値 2 5 6 を決定することができる (6 0 4)。検出およびトラッキング信頼性値 2 5 6 は、ターゲットオブジェクトが現在のビデオフレーム (N) 2 2 4 の中で、または特定のウィンドウ内で見出されるかどうかの、信頼性のレベルを与え得る。電子デバイス 1 0 2 は、検出および信頼性値 2 5 6 が検出およびトラッキング閾値 2 5 4 よりも大きいかどうかも決定することができる (6 0 6)。検出および信頼性値 2 5 6 が検出およびトラッキング閾値 2 5 4 よりも大きい場合、電子デバイス 1 0 2 は、次のビデオフレームにおけるウィンドウとサイズのサブセット (たとえば、同じサブセット) とを使用して、次のビデオフレームに対するオブジェクト検出を実行することができる (6 0 8)。代替的に、検出および信頼性値 2 5 6 が検出およびトラッキング閾値 2 5 4 よりも小さい場合、電子デバイス 1 0 2 は、次のビデオフレームにおけるウィンドウ位置およびサイズのより大きいサブセットを使用して、次のビデオフレームに対するオブジェクト検出を実行することができる (6 1 0)。いくつかの構成では、信頼性値 2 5 6 が検出およびトラッキング閾値 2 5 4 よりも小さい場合、電子デバイス 1 0 2 は、次のビデオフレームの探索空間全体および / またはすべてのウィンドウを使用して、次のビデオフレームに対するオブジェクト検出を実行することができる (6 1 0)。

20

【 0 0 6 4 】

[00131] 図 7 を参照すると、異なるウィンドウサイズ 7 6 6 を有する画像ウィンドウ 7 0 0 の特定の実施形態が示されている。具体的には、図 7 は、1 0 個の可能なウィンドウサイズ 7 6 6 a ~ j のセットを示している。各ウィンドウサイズ 7 6 6 は、スケールレベル (たとえば、1 ~ 1 0) に対応し得る。本明細書では矩形として示されているが、探索されるウィンドウは、任意の形状、たとえば、正方形、矩形、円形、楕円、ユーザ定義されたものなどであり得る。さらに、任意の数、たとえば、5 個、1 5 個、2 0 個、3 0 個などのウィンドウサイズ 7 6 6 またはスケールレベルが利用可能であり得る。

30

【 0 0 6 5 】

[00132] 探索範囲は、特定の位置に対して使用されるウィンドウサイズのサブセットによって示されてよく、たとえば、現在のビデオフレーム (N) 2 2 4 中で探索されるウィンドウサイズは、最近のフレーム中のターゲットオブジェクトと関連付けられるウィンドウ位置およびウィンドウサイズと同様のものに限定され得る。たとえば、フィードバックを伴わずに、オブジェクト検出器 2 0 8 は、各々の選択されたウィンドウ位置についてすべての 1 0 個のウィンドウサイズ 7 6 6 a ~ j を探索し得る。しかしながら、オブジェクトが、第 5 のウィンドウサイズ 7 6 6 e を有するウィンドウを使用して最近の (非現在の) ビデオフレーム中で検出された場合、スキャナスケラ 2 3 6 は、5 プラスまたはマイナス 3 という現在のキャプチャされたフレームのためのウィンドウサイズ、すなわち、ウィンドウサイズ 2 ~ 8 だけを選択し得る。言い換えれば、第 1 のウィンドウサイズ 7 6 6 a、第 9 のウィンドウサイズ 7 6 6 i、および第 1 0 のウィンドウサイズ 7 6 6 j を有するウィンドウは、最近または前のビデオフレーム (N - 1) 2 2 2 からのフィードバックに基づいて探索されないことがある。これはさらに、低い確率の探索をなくし、オブジェ

40

50

クト検出の効率を増加させ得る。言い換えれば、最近のビデオフレームからのフィードバックを使用することは、実行される計算を減らすのに役立ち得る。代替的に、最近のビデオフレームがターゲットオブジェクトを検出しなかった（すなわち、最近のキャプチャされたフレームのための検出およびトラッキング信頼性値 256 が検出およびトラッキング閾値 254 よりも小さい）場合、オブジェクト検出器 208 は、サイズレベルのサブセットを使用することによって探索範囲を限定しないことがある。

【0066】

[00133]図 8 を参照すると、オブジェクトトラッキングおよび検出モジュール 804 の特定の例示的な実施形態が示されている。図 8 に示されたオブジェクトトラッキングおよび検出モジュール 804 は、図 2 に示されたオブジェクトトラッキングおよび検出モジュール 204 と同様のモジュールを含み、同様の機能を実行し得る。具体的には、図 8 に示されたオブジェクト検出器 808、動きトラッカー 806、スキャナロケータ 830、ウィンドウ位置選択器 832、ランダムマイザ 834、スキャナスケラ 836、分類器 838、融合モジュール 860、メモリバッファ 810、キャプチャされた前のビデオフレーム 812、ウィンドウ 842、位置 844、サイズ 846、バイナリ決定 848、トラッキング閾値 850、検出閾値 852、検出およびトラッキング閾値 854、検出信頼性値 840、トラッキング信頼性値 828、ならびに検出およびトラッキング信頼性値 856 は、図 2 に示されたオブジェクト検出器 208、動きトラッカー 206、スキャナロケータ 230、ウィンドウ位置選択器 232、ランダムマイザ 234、スキャナスケラ 236、分類器 238、融合モジュール 260、メモリバッファ 210、キャプチャされた前のビデオフレーム 212、ウィンドウ 242、位置 244、サイズ 246、バイナリ決定 248、トラッキング閾値 250、検出閾値 252、検出およびトラッキング閾値 254、検出信頼性値 240、トラッキング信頼性値 228、ならびに検出およびトラッキング信頼性値 256 に対応し、それらと同様の機能を有し得る。

【0067】

[00134]オブジェクトトラッキングおよび検出モジュール 804 は、ターゲットの動きおよびトラッキングエラーによるジッタの影響を低減するために使用される平滑化モジュール 861 を含み得る。言い換えれば、平滑化モジュール 861 は、トラッキング結果を平滑化し、探索ウィンドウが、位置 (x, y) 844 とサイズ (幅、高さ) 846 の両方においてより滑らかな軌道を有するようにする。平滑化モジュール 861 は、単純な移動平均 (MA) フィルタまたは自動回帰 (AR) フィルタであり得る。位置 844 に対する平滑化の程度およびサイズ 846 に対する平滑化の程度は異なり得る。カルマンフィルタのような予測フィルタも、位置 844 の平滑化に好適であり得る。したがって、平滑化モジュール 861 は、平滑化されていない位置 863 と平滑化されていないサイズ 865 とを入力として受信し、平滑化された位置 867 と平滑化されたサイズ 869 とを出力し得る。

【0068】

[00135]図 9 を参照すると、平滑化モジュール 961 の特定の例示的な実施形態が示されている。平滑化モジュール 961 は、ターゲットの動きとトラッキングエラーとによるジッタの影響を低減するために使用されてよく、すなわち、したがって、トラッキング結果 (バウンディングボックス) は、位置 (x, y) とサイズ (幅、高さ) の両方においてより滑らかな軌道を有する。一構成では、位置平滑化フィルタ 971 およびサイズ平滑化フィルタ 973 は、平滑化されていない位置 963 と平滑化されていないサイズ 965 とを入力として受け取り、平滑化された位置 967 と平滑化されたサイズ 969 とを出力するために、自動回帰 (AR) モデルを使用して実装される。

【0069】

[00136]自動回帰 (AR) モデルでは、X は、平滑化されるべき変数であり、位置またはサイズのいずれかであると仮定する。さらに、X' をオブジェクトトラッカーによる X の出力とする。この構成では、時間 t における X の平滑化されたフィルタリング X_t が、式 (1) に従って表され得る。

$$X_t = W * X'_{t-1} + (1 - W) * X_{t-1} \quad (1)$$

ここで、 X'_{t-1} は、時間 t における X のトラッカー出力であり、 X_{t-1} は、時間 $t - 1$ における X の平滑化された結果であり、 W ($0 < W < 1$) は、平滑化の効果を制御する平滑化重みである。たとえば、 X'_{t-1} は、現在のビデオフレーム (N) 224 のために選択されたウィンドウ位置またはウィンドウサイズであってよく、 X_{t-1} は、前のビデオフレーム ($N - 1$) 222 のために使用されたウィンドウ位置またはウィンドウサイズであってよい。

【0070】

[00137]異なる平滑化重み W が、位置平滑化フィルタ 971 およびサイズ平滑化フィルタ 973 のために使用され得る。たとえば、一実装形態では、ウィンドウ位置に対しては平滑化効果がより少ないが、ウィンドウサイズに対して平滑化効果がより強いように、 $W_{\text{location}} = 0.8$ および $W_{\text{size}} = 0.4$ である。平滑化重みのこの選択は、より少ないトラッキング遅延とより少ないジッタの両方をもたらす。

【0071】

[00138]平滑化重みのその選択はまた、検出およびトラッキング信頼性値 856 がある閾値（たとえば、検出およびトラッキング閾値 854）を下回るときに低減され得る。これは、潜在的トラッキングまたは検出誤差が高いとき、より強いフィルタリングを引き起こし得る。たとえば、低いトラッキングの信頼性に応答して（たとえば、検出およびトラッキング信頼性値 856 が検出およびトラッキング閾値 854 を下回る）、位置のための平滑化重みおよびサイズのための平滑化重みは、それぞれ $W_{\text{location}} = 0.65$ および $W_{\text{size}} = 0.2$ に設定され得る。言い換えれば、重みの一方または両方が下げられてよく、これにより、ウィンドウ位置およびサイズ選択は、現在のビデオフレームのウィンドウ位置およびサイズよりも、前のビデオフレームのウィンドウ位置およびサイズにより大きく依存するようになり得る。

【0072】

[00139]重み付けは、検出およびトラッキング信頼性値 856 ではなく、トラッキング信頼性値 828 または検出信頼性値 840 に基づき得る。たとえば、平滑化重み W_{location} および W_{size} は、トラッキング信頼性値 828 がトラッキング閾値 850 を下回ったことに応答して下げられてよく、すなわち、より強いフィルタリングが、不十分な動きトラッキングに応答して使用されてよい。代替的に、平滑化重みは、検出信頼性値 840 が検出閾値 852 を下回ったことに応答して下げられてよく、すなわち、より強いフィルタリングが、不十分なオブジェクト検出に応答して使用されてよい。

【0073】

[00140]別の構成では、カルマンフィルタリングが、ウィンドウ位置を平滑化するために使用され得る。そのような構成では、フィルタリングは式 (2) ~ (7) に従って定義され得る。

$$x_k = F_k x_{k-1} + w_k \quad (2)$$

$$z_k = H x_{k-1} + v_k \quad (3)$$

ここで、 x_{k-1} は時間 $k - 1$ における前の状態であり、 x_k は

【0074】

【数2】

$$X_k = [x, y, \dot{x}, \dot{y}]$$

【0075】

によって定義される現在の状態であり、

(x, y) はバウンディングボックスの中心の位置であり、

【0076】

10

20

30

40

【数 3】

$$(\dot{x}, \dot{y})$$

【0077】

は、各方向における速度である。さらに、状態遷移モデル F_k 、および観測モデル H は、それぞれ式 (4) ~ (5) によって定義され得る。

【0078】

【数 4】

10

$$F_k = \begin{bmatrix} 1, & 0, & \Delta t, & 0 \\ 0, & 1, & 0, & \Delta t \\ 0, & 0, & 1, & 0 \\ 0, & 0, & 0, & 1 \end{bmatrix} \quad (4)$$

$$H = \begin{bmatrix} 1, & 0, & 0, & 0 \\ 0, & 1, & 0, & 0 \end{bmatrix} \quad (5)$$

20

【0079】

上式で、 Δt は調節可能なパラメータである。加えて、 w_k は、式 (6) に従った共分散 Q を有するゼロ平均多変量正規分布 (zero mean multivariate normal distribution) から導出されると考えられるプロセス雑音である (すなわち、 $w_k \sim N(0, Q)$)。

【0080】

【数 5】

30

$$Q = \begin{bmatrix} 1, & 0, & 0, & 0 \\ 0, & 1, & 0, & 0 \\ 0, & 0, & 1, & 0 \\ 0, & 0, & 0, & 1 \end{bmatrix} * \sigma_1^2 \quad (6)$$

【0081】

ここで、 σ_1 は調節可能なパラメータである。同様に、 w_k は、式 (7) に従った共分散 R を有するゼロ平均ガウスホワイトノイズ (zero mean Gaussian white noise) であると

40

【0082】

【数 6】

$$R = \begin{bmatrix} 1, & 0, & 0, & 0 \\ 0, & 1, & 0, & 0 \\ 0, & 0, & 1, & 0 \\ 0, & 0, & 0, & 1 \end{bmatrix} * \sigma_2^2 \quad (7)$$

【0083】

50

ここで、 α_2 は調節可能なパラメータである。

【0084】

[00141]図10を参照すると、動きトラッキングの結果のジッタを平滑化するための方法1000の特定の例示的な実施形態のフローチャートが示されている。方法1000は、電子デバイス102、たとえば、電子デバイス102中のオブジェクトトラッキングおよび検出モジュール804によって実行され得る。電子デバイス102は、現在のビデオフレーム224と関連付けられる1つまたは複数のウィンドウ位置と1つまたは複数のウィンドウサイズとを決定し、たとえば、平滑化されていない位置863と平滑化されていないサイズ865とを決定することができる(1002)。電子デバイス102はまた、1つまたは複数の平滑化されたウィンドウ位置867と1つまたは複数の平滑化されたウィンドウサイズ869とを生成するために、1つまたは複数のウィンドウ位置と1つまたは複数のウィンドウサイズとをフィルタリングすることができる(1004)。たとえば、これは、移動平均フィルタ、自動回帰フィルタ、またはカルマンフィルタを使用することを含み得る。一構成では、低いトラッキング信頼性に応答して(たとえば、検出およびトラッキング信頼性値856が検出およびトラッキング閾値854を下回る)、位置のための平滑化重みおよびサイズのための平滑化重みは低減され得る。代替的に、平滑化重みは、検出信頼性値840またはトラッキング信頼性値828に基づいて低減され得る。電子デバイスはまた、1つまたは複数の平滑化されたウィンドウ位置867および1つまたは複数の平滑化されたサイズ869によって定義された1つまたは複数のウィンドウを使用して、現在のビデオフレーム224内のターゲットオブジェクトを検出することができる(1006)。

【0085】

[00142]撮影されたシーンの文脈では、「オブジェクト」という用語は、シーン内の物理的な物体を指す。ビデオストリームの文脈では、「オブジェクト」という用語は、ビデオストリーム内での物体の表現(たとえば、ビデオストリームのフレーム中の物体の画像)を指す。本明細書で使用される「モバイルデバイス」という用語は、次のフォームファクタ、すなわち、携帯可能(たとえば、スマートフォン)、運転可能(たとえば、車両またはロボット)、装着可能(たとえば、衣服またはアクセサリ)、および飛行可能(たとえば、ドローン)のいずれかのデバイスを含む。モバイルデバイスは、1つまたは複数の画面(たとえば、タッチスクリーン)および/または1つまたは複数の画像キャプチャデバイス(たとえば、カメラ)を含み得る。

【0086】

[00143]デジタル画像(たとえば、ビデオストリームのフレーム)は、1つまたは複数のオブジェクトを含み得るシーンを示す。通常、オブジェクトは物理的な有形の物体(たとえば、人々)である。図12では、たとえば、シーン中のオブジェクトは、3人の人物と、星と、木を含む。

【0087】

[00144]1つまたは複数のデジタル画像(たとえば、ビデオストリームのフレーム)から、1つまたは複数の画像内の選択されたオブジェクトの位置を経時的に表す情報を有する、関連するストリームまたはファイルを生成するのが望ましいことがある。一例では、そのような情報は、経時的に変化し得る、対象の点(たとえば、選択されたオブジェクトの重心)の絶対的な位置を含む。別の例では、そのような情報は、経時的に変化し得る、対象の点(たとえば、選択されたオブジェクトの重心)の位置に対する選択されたオブジェクトの位置を含む。対象の点の位置(たとえば、選択されたオブジェクトの位置、または複数のオブジェクトの重心)は、メタデータ(たとえば、1つまたは複数のそれぞれの位置コードブックへの、コードワードのような1つまたは複数のインデックス)として符号化され得る。

【0088】

[00145]図11Aは、タスクT100とT300とを含む第1の構成によるメタデータを生成する方法M100のフローチャートを示す。タスクT100は、複数のオブジェク

トを選択するために少なくとも1つのデジタル画像を使用し、ビデオストリームは物理空間との関連でオブジェクトを示す。位置空間中の選択されたオブジェクトの位置を示す情報に基づいて、タスクT300は、選択されたオブジェクトの複数の幾何学的構成の候補のうちの1つを特定するメタデータを生成する。タスクT100およびT300は、たとえば、図1の電子デバイス102、図2Bのプロセッサ264によって実装されるオブジェクトトラッキングおよび検出モジュール204、プロセッサによって実装される図8のオブジェクトトラッキングおよび検出モジュール804、プロセッサによって実装される図9の平滑化モジュール961、またはこれらの任意の組合せによって実行され得る。

【0089】

[00146]タスクT100は、複数のオブジェクトを選択するために少なくとも1つのデジタル画像を使用し、少なくとも1つのデジタル画像は物理空間との関連でオブジェクトを示す。この選択は、ユーザによる直接の動作にตอบสนองして実行され得る。そのような場合、少なくとも1つのデジタル画像をユーザに表示するために画面が使用されてよく、ユーザは、表示に現れているオブジェクトの中から特定の所望のオブジェクトを選択するために指示を入力することができる。図11Bは、少なくとも1つのデジタル画像を表示するタスクT50を含む方法M100の実装形態M110のフローチャートを示す。表示がタッチスクリーン上で実行される場合、ユーザは、画面をタッチして所望のオブジェクト内のある点を選択することによって、選択を示すことができる。

【0090】

[00147]少なくとも1つのデジタル画像は、ビデオストリームの1つまたは複数のフレームであり得る。ビデオストリームは一連のフレームを表し、ここで各フレームはピクセル座標空間中の画像を表す。ビデオストリームは通常、ストリームからフレームを復元するために使用され得る他の情報（たとえば、各フレームに対する、対応するフレーム開始コードまたはパケットおよびフレーム終了コードまたはパケット）を含む。ビデオストリームはまた、特定のフレームと関連付けられ得る埋め込まれたデータ（たとえば、メタデータ）を含み得る。ビデオストリームは、（可視の波長および/または他の波長を感知し得る）カメラまたは他の撮像デバイスによって生成されてよく、別のデバイスからストリーミングされてよく、またはデコードによって（たとえば、磁気媒体または光学媒体に記憶されている情報から）生成されてよく、圧縮された形式または圧縮されていない形式であってよい。ビデオストリームは、構築された光イメージャまたは他のデプスカメラ（たとえば、Microsoft Kinect）によってキャプチャされた画像に基づくビデオストリームのような、深度情報を含み得る。そのようなビデオストリームは、たとえば、各ピクセルの深度値を対応する色と対応付けることによって、タッチスクリーン上に表示され得る。ビデオストリームは、生であることがあり、遅延していることがあり、または記憶装置から検索される（たとえば、事前に記録されている）ことがある。

【0091】

[00148]ビデオストリームは、デジタル画像（「フレーム」）の時間シーケンスである。一例では、画像のシーケンスは、特定のフレームレートで（たとえば、同じフレームレートでの提示のために）キャプチャされる。ビデオストリームは、各フレームの開始と終了とを示す情報を含み得る。一例では、ビデオストリームは、各フレームの最初のピクセルの前のフレーム開始コードと、各フレームの最後のピクセルの後のフレーム終了コードとを含む。別の例では、ビデオストリームは、ビデオストリームのフレームのピクセル情報を搬送する1つまたは複数のデータ信号と、ビデオストリームの各フレームの開始と終了とを示す1つまたは複数の同期信号とを含む。プライマリビデオストリームの特定のフォーマットに応じて、プライマリビデオストリームは追加の情報（たとえば、ヘッダ、メタデータ）を含み得る。プライマリビデオストリームが、フレーム開始コードとフレーム終了コードとの間にフレームデータ（すなわち、ピクセル値）以外のデータを含むことも可能である。そのような他のデータは、たとえば、センサ構成の詳細および/または画像の統計値を含み得る。

【0092】

10

20

30

40

50

[00149]ビデオストリームの解析は、（たとえば、フレーム開始情報とフレーム終了情報とを使用して）ビデオストリームのフレームを抽出することと、フレームを処理のために利用可能にすることとを含む。たとえば、解析することは、バッファのような対応する指定された位置に各フレームを記憶することを含み得る。図 1 1 C は、少なくとも 1 つのデジタル画像を生成するためにビデオストリームを解析するタスク T 3 0 を含む、方法 M 1 0 0 の実装形態 M 1 2 0 のフローチャートを示す。代替的に、ビデオストリームの 1 つまたは複数のフレームが任意の時間において方法 M 1 0 0 に対して（たとえば、メモリ中のバッファにおいて）利用可能であるように、別のプロセスが（たとえば、表示のために）ビデオストリームを解析するために実行され得る。図 1 1 D は、方法 M 1 1 0 および M 1 2 0 の実装形態 M 1 3 0 のフローチャートを示す。

10

【 0 0 9 3 】

[00150]そのようなバッファ（たとえば、メモリバッファ 1 1 0 ）は通常、ある時間において少なくとも 1 つの完全なフレームを記憶することが可能である。プライマリビデオストリームの解析は、複数のフレームバッファを交互にするように構成され得る。たとえば、1 つのそのようなフレームバッファは処理のために完全なフレームを記憶することができるが、シーケンス中の後続のフレームが抽出されて別のフレームバッファに記憶される。代替的に、タスク T 3 0 または解析プロセスは、新たなフレームが抽出されるにつれてバッファが更新される（および古いフレームが上書きされる）ように、解析されたフレームを循環バッファに記憶するように構成され得る。

【 0 0 9 4 】

20

[00151]プライマリビデオストリームを解析することは、1 つまたは複数の関連するオーディオストリームおよび / またはメタデータストリームからプライマリビデオストリームを分離するための、逆多重化動作を含み得る。代替的に、そのような動作は、タスク T 3 0 または解析プロセスの上流で実行され得る（たとえば、プライマリビデオストリームはデマルチプレクサによって提供され得る）。

【 0 0 9 5 】

[00152]図 1 2 は、3 つのオブジェクトを選択するための一連のユーザの動作の例を示す。パネル A は、デバイスのタッチスクリーン上に表示されるようなシーンを示す。デバイスは、選択モードを開始するようにユーザによって操作され得る。たとえば、タッチスクリーンは、ユーザが適切なアイコンをタッチするときに、選択モードに入るように構成され得る。パネル A は、フレームの右上の角にある選択ツールアイコンの例を示す。パネル B において、ユーザは、選択モードを開始するために選択ツールアイコンをタッチする。パネル C において、デバイスが選択モードにあることを示すために、タッチ動作にตอบสนองしてツールアイコンがハイライトされており、ユーザは第 1 のオブジェクトをタッチして第 1 のオブジェクトを選択する。パネル D において、オブジェクトが選択されていることを示すために、タッチ動作にตอบสนองして第 1 のオブジェクトがハイライトされており（たとえば、示されるようにシルエットが付けられている、または輪郭が付けられている）、ユーザは第 2 のオブジェクトをタッチして第 2 のオブジェクトを選択する。パネル E において、オブジェクトが選択されていることを示すために、タッチ動作にตอบสนองして第 2 のオブジェクトがハイライトされており、ユーザは第 3 のオブジェクトをタッチして第 3 のオブジェクトを選択する。パネル F において、オブジェクトが選択されていることを示すために、タッチ動作にตอบสนองして第 3 のオブジェクトがハイライトされており、ユーザは選択ツールアイコンを再びタッチして選択モードを終了する。デバイスは、ユーザが選択モードにおいて再び選択されたオブジェクトをタッチすると、選択されたオブジェクトの選択を解除するように構成され得る。

30

40

【 0 0 9 6 】

[00153]図 1 3 は、3 つのオブジェクトを選択するための一連のユーザの動作の別の例を示す。この場合、デバイスは、図 1 2 に示されるような選択モードとは異なるように動作する。パネル A において、ユーザは、選択モードを開始するために選択ツールアイコンをタッチする。パネル B において、デバイスが選択モードにあることを示すために、タッ

50

チ動作にตอบสนองしてツールアイコンがハイライトされており、ユーザは第1のオブジェクトをタッチして第1のオブジェクトを選択する。パネルCにおいて、オブジェクトが選択されていることを示すために、タッチ動作にตอบสนองして第1のオブジェクトがハイライトされており、選択にตอบสนองして、選択モードが終了され、ハイライトがツールアイコンから除去される。ユーザは選択ツールアイコンを再びタッチして選択モードを開始し、パネルDにおいて、デバイスが選択モードにあることを示すために、タッチ動作にตอบสนองしてツールアイコンがハイライトされており、ユーザは第2のオブジェクトをタッチして第2のオブジェクトを選択する。パネルEにおいて、オブジェクトが選択されていることを示すために、タッチ動作にตอบสนองして第2のオブジェクトがハイライトされており、選択にตอบสนองして、選択モードが終了され、ハイライトがツールアイコンから除去される。ユーザは選択ツールアイコンを再びタッチして選択モードを開始し、パネルFにおいて、デバイスが選択モードにあることを示すために、タッチ動作にตอบสนองしてツールアイコンがハイライトされており、ユーザは第3のオブジェクトをタッチして第3のオブジェクトを選択する。

10

【0097】

[00154]別の例では、タッチスクリーンは、2動作の選択にตอบสนองして選択モードに入るように構成されてよく、第1の動作はメニューモードを選択し（たとえば、ユーザは画面の指定された領域をタッチしてメニューを表示させる）、第2の動作はメニューから選択モードを選択する。さらなる例では、選択ツールアイコンをタッチアンドホールドし、アイコンを画面上の所望の点へとドラッグし、次いでアイコンを離してその点を選択することによって、ユーザがオブジェクトを選択するように、タッチスクリーンが構成され得る。代替的に、ユーザは、所望のオブジェクトの少なくとも一部分を含む画像のエリア（たとえば、バウンディングボックスまたは楕円またはなげなわ）を選択することによって選択を示すために、タッチスクリーンを操作することができる。そのような場合、ユーザは、所望のオブジェクトを個々に、および/またはグループとして選択することができる。

20

【0098】

[00155]ユーザが選択モードの間に表示されるビューのフィールドを操作できるようにすること（たとえば、ビューのフィールドを広げるためのピンチ動作、およびビューのフィールドを狭めるためのズーム動作を行うこと）が、望ましいことがある。タッチスクリーンの操作以外に、直接のユーザによる選択のための他の可能な方式は、1つまたは複数のボタンもしくは他のスイッチを作動させること、および/またはジェスチャ認識を含む。ユーザによるそのような直接の動作に対する代替形態として、タスクT100は、1つまたは複数の所定の基準を少なくとも1つのデジタル画像（たとえば、ビデオストリームの1つまたは複数のフレーム）からの情報に適用することによって、自動的に選択を実行することができる。そのような基準は、たとえば、選手のジャージの番号の認識、顔認識、および/または（たとえば、スポーツイベントにおけるボールのような高速に移動するオブジェクトを識別するための）速度の検出を含み得る。

30

【0099】

[00156]方法M100は、タスクT100における選択にตอบสนองして、オブジェクトをトラッキングするように実施され得る。たとえば、下で説明されるような位置決定タスクT200は、（たとえば、オブジェクトトラッキングおよび検出モジュール204ならびに/またはオブジェクトトラッキングおよび検出モジュール804を参照して上で説明されたように）ビデオストリーム内の1つまたは複数の選択されたオブジェクトをトラッキングするためにも実施され得る。代替的に、オブジェクトトラッキングはタスクT100よりも前に開始してよい。そのような場合、タスクT100は、トラッキングされるオブジェクトのセットの中からの選択のために実施され得る。たとえば、ユーザは、トラッキングされたオブジェクトの表示の中から選択することができる。そのようなオブジェクトトラッキング（たとえば、タスクT200による）のために使用され得るアルゴリズムの例は、Lucas-Kanade法（およびKanade-Lucas-Tomasiトラッカーのような変形）と、ミーンシフトトラッキングとを含む。オブジェクトのトラッキングは、プライマリビデオストリームのフレームの他の部分からオブジェクトを区画する

40

50

ことを含み得るが、オブジェクトトラッキングは、そのような区画化がなくても実行され得る。

【 0 1 0 0 】

[00157] デジタル画像（たとえば、ビデオストリームのフレーム）の他の部分からのオブジェクトの区画化は、フレームの他の部分に対するオブジェクトの強調された表示（たとえば、シルエット付与または輪郭付与）をサポートするために使用され得る。区画化は、タスク T 1 0 0 におけるオブジェクトの選択に応答して（たとえば、表示タスク T 5 0 によって）実行され得る。そのような区画化は、オブジェクトの選択を構成することによって、ユーザに視覚的なフィードバックを提供するために使用され得る。区画化は、選択が開始する前にも実行され得る。そのような場合、タスク T 1 0 0 における選択は、すでに区画されているオブジェクトのセット（たとえば、区画されたオブジェクトが強調されている表示）の中から行われ得る。

10

【 0 1 0 1 】

[00158] 図 1 4 の左側のパネルは、オブジェクトを含むシーンを示すビデオストリームからのフレームの 2 つの例を示す。オブジェクトの強調された表示は、タスク T 1 0 0 における選択の前に、またはそのような選択に応答して（たとえば、表示タスク T 5 0 によって）実行され得る。強調された表示の一例はオブジェクトのシルエットであり、これは、背景とは対照的な色（たとえば、図 1 4 の中央のパネルにおいて示されるように、黒または黄）で表示され得る。この対照的な色は、すべてのオブジェクトに対して同じであってよく、または、（たとえば、各オブジェクトの局所的な背景と対照的となるように）オブジェクトによって異なっていてよい。強調された表示の他の例（対照的な色と同様に行われ得る）は、オブジェクトに輪郭を付与すること、およびオブジェクトにおいて、オブジェクト内に、またはオブジェクトの周囲にグラフィカル要素を表示すること（たとえば、オブジェクトの頂点または重心における三角形、図 1 4 の右側のパネルに示されるようなオブジェクトの周囲の楕円など）を含む。

20

【 0 1 0 2 】

[00159] ビデオストリームの表示をユーザによるオブジェクトの直接の選択のための環境として使用することは、低速で動いているオブジェクトに対しては適切であり得る。しかしながら、所望のオブジェクトのいずれかが表示の中で高速に動いている場合、そのような構成はユーザにとって不愉快であり得る。そのような高速な移動の場合、代わりに、ユーザによるオブジェクトの直接の選択のための環境として、ビデオストリームの単一のフレームを表示することが望ましいことがある。上で説明されたように、選択は、フレーム中のオブジェクトが区画され強調されている表示から行われ得る。ストリーム表示の選択とフレーム表示の選択の両方がサポートされる場合、これらの表示モードからの選択は、ユーザによって、および/または自動的に実行され得る。自動的な選択の場合、ストリームにおける動きの程度が、たとえばオプティカルフロー、動きベクトル、フレーム間のピクセル差分などのような尺度（メトリック）を使用して、定量化され得る。そのような場合、フレーム表示モードは、尺度の値が閾値を上回るときに選択されてよく、ストリーム表示モードは、尺度の値が閾値を下回るときに選択されてよい。

30

【 0 1 0 3 】

[00160] 位置空間中の選択されたオブジェクトの位置を示す情報に基づいて、タスク T 3 0 0 は、選択されたオブジェクトの複数の幾何学的構成の候補の中から 1 つを特定するメタデータを生成する。位置空間は、デジタル画像（たとえば、ビデオストリームのフレーム）の（ x , y ）直交座標空間のような、（たとえば、ビデオストリームの）少なくとも 1 つのデジタル画像のピクセル座標空間であり得る。デジタル画像（たとえば、ビデオフレーム）のピクセル座標空間の原点は通常、画像の左上または左下の角として指定される。選択されるオブジェクトの位置は、たとえば、ユーザがオブジェクトを選択するために画面をタッチした位置であり得る。代替的に、選択されるオブジェクトの位置は、本明細書で説明されるようなオブジェクトトラッキングおよび検出モジュール 2 0 4 または 8 0 4 のような、オブジェクトトラッカーによって提供され得る。

40

50

【 0 1 0 4 】

[00161]幾何学的構成は、選択されるオブジェクトの空間における構成である。たとえば、幾何学的構成は、（たとえば、多角形の頂点のような）選択されるオブジェクトの位置によって表される形状であり得る。タスク T 3 0 0 は、オブジェクトの位置に 1 つまたは複数の尺度を適用することによって、幾何学的構成の候補のうちの適切な 1 つを特定するように実施され得る。そのような場合、タスク T 3 0 0 は、尺度の値に従って候補を選択するために、決定木を使用することができる。一例では、（たとえば、すでに分類されている）幾何学的構成の訓練セットが、決定木を生成するために使用される。生成されるメタデータはまた、候補に適用されるべきスケール係数および / またはアスペクト係数、候補の方向の向き、および、軸の回りで候補が反転されるべきであるかどうかを示すミラーフラグのうちのいずれかのような、追加の情報を含み得る。

10

【 0 1 0 5 】

[00162]タスク T 3 0 0 は、複数の幾何学的構成の候補のうちのいずれが選択されたオブジェクトの位置によって表される形状と最も類似しているかを決定するように実施され得る。選択されたオブジェクトのすべてが一意である場合、タスク T 3 0 0 によって使用され得る類似性の 1 つの尺度は、オブジェクトの観測される位置と、（たとえば、方向およびスケールリングの補償後の）幾何学的構成の候補における対応する位置との間の距離を、各々の選択されたオブジェクトについて合計したものである。図 3 5 は、「等距離(equidistance)」、「近 - 遠 - 遠(near-far-far)」、および「近 - 遠 - 遠(near-far-far)」と標識される、3 つの幾何学的構成の候補のセットの 1 つと 3 つの位置 A、B、C によって表される形状を照合するためにタスク T 3 0 0 が実施され得る、別の例を示す。

20

【 0 1 0 6 】

[00163]この例では、点の間の 3 つの距離（すなわち、A B、B C、および A C）が、それらの中で最小のものと、中間のものと、最大のものとを決定するために計算され、大きさにより分類される。最小の距離は、最大の距離の 3 分の 2 である値と比較される。最小の距離の方が大きい場合、タスク T 3 0 0 は、図 3 5 の左上にあるボックスによって示されるように、候補「等距離」を選択する。最小の距離の方が小さい場合、中間の距離が、最小の距離および最大の距離の合計の半分である値と比較される。中間の距離の方が大きい場合、タスク T 3 0 0 は、図 3 5 の下にあるボックスによって示されるように、候補「近 - 遠 - 遠」を選択する。中間の距離の方が小さい場合、タスク T 3 0 0 は、図 3 5 の

30

【 0 1 0 7 】

[00164]図 3 5 はまた、決定の境界に当たる 3 つの場合を示しており、これらは、望み通りに、隣接する候補のいずれかへとタスク T 3 0 0 によって割り当てられ得る。この場合、適切な候補を特定するメタデータは、各々が 3 つの幾何学的構成の候補の対応する 1 つを示す、3 つのあり得る値の 1 つを有するパラメータであり得る。メタデータはまた、（たとえば、回転として）候補に適用されるべき方向を示し得る。一例では、方向は、最小の距離の反対側の点を通る、幾何学的構成の中心からの線の方法として示される。別の例では、方向は、選択されたオブジェクトの特定の 1 つに対応する点を通る、幾何学的構成の中心からの線の方法として示される。

40

【 0 1 0 8 】

[00165]図 3 7 A は、タスク T 2 0 0 を含む方法 M 2 0 0 の実装形態 M 2 0 0 のフローチャートを示し、タスク T 2 0 0 は、位置空間中の選択されたオブジェクトの位置を示す情報を取得する。タスク T 2 0 0 は、タスク T 1 0 0 における選択に応答して実行するように実施され得る。位置空間は、デジタル画像（たとえば、ビデオストリームのフレーム）の（x, y）直交座標空間のような、（たとえば、ビデオストリームの）少なくとも 1 つのデジタル画像のピクセル座標空間であり得る。デジタル画像（たとえば、ビデオフレーム）のピクセル座標空間の原点は通常、画像の左上または左下の角として指定される。図 3 7 B ~ 図 3 7 D は、タスク T 2 0 0 を含む、方法 M 1 1 0 の実装形態 M 2 1 0、方法 M 1 2 0 の実装形態 M 2 2 0、および方法 M 1 3 0 の実装形態 M 2 3 0 のフローチャート

50

をそれぞれ示す。

【0109】

[00166]表示のピクセル座標空間は位置空間と同じであってよく、または位置空間は表示のピクセル座標空間を含んでよい。たとえば、表示のフレームは、（たとえば、ビューのフィールドを狭めるためのズーム動作に回答して）ビデオストリームの対応するフレームの一部分だけを含み得る。代替的に、表示のピクセル座標空間は位置空間を含み得る。たとえば、ビデオストリームは、表示のフレームよりも小さい表示のウィンドウに表示されてよく、他のコンテンツは表示の別のウィンドウに表示される。

【0110】

[00167]タスクT200は、ピクセル座標空間中の選択されたオブジェクトの画像の位置として、選択されたオブジェクトの位置を決定するように実施され得る。たとえば、決定された位置はオブジェクトの重心であり得る。一定の密度が仮定される場合、重心はセントロイドと同じであるが、密度の差は、たとえば色および/または透明度の差によっても示され得る。決定された位置の他の例は、オブジェクトの上部または下部のピクセル、オブジェクトの上部または下部へのオブジェクトの重心の垂直方向の投影、またはオブジェクトの特定の特徴の位置を含む。図17Aは、オブジェクトの下部境界線へとオブジェクトの重心C10を投影して対応する決定された位置DP10を得る例を示す。

【0111】

[00168]代替的に、または加えて、タスクT200は、複数のオブジェクトの編成（formation）の基準の位置を決定するように実施され得る。たとえば、タスクT200は、選択されたオブジェクトの編成の位置として基準の位置を計算するように実施され得る。そのような場合、決定された基準の位置は、編成の重心（異なるオブジェクトが同じまたは異なる密度を有する）であってよく、オブジェクトに属する上部もしくは下部のピクセル、編成の上部もしくは下部への編成の重心の垂直方向の投影、またはオブジェクトの特定の1つの位置であり得る。そのような場合、タスクT200は、（たとえば、基準の位置を原点として使用するために）基準の位置に対する位置として選択されたオブジェクトの位置を決定することができる。

【0112】

[00169]選択されたオブジェクトは、共通のオブジェクトの一部ではないように、互いに物理的に離れていてよい。たとえば、オブジェクトの少なくとも1つは、他のオブジェクトとは独立に動くことが可能であり得る。特定の例では、オブジェクトの1つまたは複数は、異なる人、動物、および/または車両である。

【0113】

[00170]ビデオストリームまたは表示のピクセル座標空間は、2次元または3次元であり得る。3次元表示の例は、立体視的（たとえば、左/右）なボリュメトリック（volumetric）表示を含む。そのような位置空間にあるオブジェクトのユーザ選択は、たとえば、手袋、ジェスチャ（たとえば、方向感知デバイスを傾けること）、または圧力感知タッチスクリーンを使用して、3次元で選択点を操作することを含み得る。

【0114】

[00171]代替的に、位置空間は物理空間（たとえば、シーン空間）であり得る。1つのそのような例では、決定された基準の位置は、構築された光イメージまたは他のデプス（depth）カメラ（たとえば、Microsoft Kinect）からのビデオストリームのような、深度（depth）情報を含むビデオストリームを使用して選択される、少なくとも1つの点に基づく。そのようなビデオストリームは、たとえば、各ピクセルの深度値を対応する色と対応付けることによって、タッチスクリーン上に表示され得る。決定された基準の位置は、物理空間中の各々の選択されたオブジェクトの画像の位置に基づいてよく、または、上の例のように、選択されたオブジェクトの編成（formation）の位置（たとえば、重心）であってよい。そのような位置空間にあるオブジェクトのユーザ選択は、たとえば、手袋、ジェスチャ（たとえば、方向感知デバイスを傾けること）、または圧力感知タッチスクリーンを使用して、3次元で選択点を操作することを含み得る。

10

20

30

40

50

【 0 1 1 5 】

[00172]図 1 5 A ~ 図 1 5 C は、方向感知デバイスを使用して表示されるシーンの深度次元において選択点を操作する例を示す。そのような方向感知デバイスは、たとえば、重力軸（たとえば、デバイスおよび地球の中心を通る軸）に対するデバイスの方向を示す 1 つまたは複数の方向センサを有するモバイルデバイスであり得る。そのような 1 つまたは複数の方向センサは、たとえば、1 つまたは複数の慣性センサ（たとえば、ジャイロ스코プおよび / または加速度計）および / または 1 つまたは複数の磁場センサ（たとえば、磁力計）を含み得る。

【 0 1 1 6 】

[00173]図 1 5 A ~ 図 1 5 C の各々において、左パネルは、方向感知デバイスの画面上の同じデジタル画像の表示を示し、右パネルは、重力軸 g に対するデバイスの対応する方向（太線で指示されている、表示の垂直方向の線を通る断面図において示されている）を示す。これらの図面の各々に表示されるようなシーンは、前景の（すなわち、キャプチャデバイスに最も近い）オブジェクトと、中景のオブジェクトと、背景の（すなわち、キャプチャデバイスから最も遠い）オブジェクトとを含む。これらの例の各々において、傾きの度合いは、表示の左側にあるスライダーによって示される。デバイスが図 1 5 A において示されるように地面に対して垂直であるとき、傾きの度合いは低く、（左パネルにおいてハイライトによって示されるように）前景のオブジェクトが選択される。デバイスが図 1 5 C において示されるように地面に対して水平であるとき、傾きの度合いは高く、（左パネルにおいてハイライトによって示されるように）背景のオブジェクトが選択される。デバイスが図 1 5 B において示されるようにこれらの方向の間にあり地面に対して 4 5 度の角度にあるとき、傾きの度合いは中間点であり、（左パネルにおいてハイライトによって示されるように）中景のオブジェクトが選択される。

【 0 1 1 7 】

[00174]別の例では、物理空間中のオブジェクトの位置は、オブジェクトの中の、またはオブジェクト上の 1 つまたは複数の位置センサ（たとえば、GPS センサ）からの情報によって示され得る。たとえば、オブジェクトは、1 つまたは複数のそのようなセンサを含んでよく（たとえば、ボール）、または、1 つまたは複数のそのようなセンサを装着してよい（たとえば、人または動物）。表示のピクセル座標空間中のユーザにより選択される位置を、物理空間中のオブジェクトの感知される位置とリンクすることが、タッチスクリーンデバイス中の位置センサおよび / もしくは方向センサ、ならびに / またはオブジェクトの特徴（たとえば、色、ジャージの番号）の認識を使用して実行され得る。

【 0 1 1 8 】

[00175]加えて、または代替的に、物理空間中のオブジェクトの位置は、ピクセル座標空間中のオブジェクトの画像の位置から投影され得る。表示面におけるオブジェクト画像の位置の間の空間的な関係は、物理空間中の対応するオブジェクトの位置の間の空間的な関係とは良好に相関しないことがある。そのような不一致は、表示面（たとえば、ビデオストリームのキャプチャの間の焦点面）が地面に対して垂直であるときに最も大きくなり得る。

【 0 1 1 9 】

[00176]図 1 6 は、シーン空間中の実際のオブジェクトの間の距離と比較されたときの、2 D ピクセル座標空間中のオブジェクトの画像間の距離の不一致の例を示す。上パネル A はデジタル画像（たとえば、ビデオストリームのフレーム）を示し、下パネル A は示されるようなシーンの左側の 3 人の人のセントロイドの間の距離を示す。これらの距離は、一番左の人と他の 2 人の各々との間の距離が、他の 2 人の間の距離よりも遠いことを示す。上パネル B は、上から直接見られた場合の同じシーンを示し、下パネル B は、真ん中の人と他の 2 人との間の距離が他の 2 人の間の距離よりも実際にははるかに遠いことを示すので、下パネル A において示される距離の関係は、シーン空間における実際の距離の関係に対して不正確である。

【 0 1 2 0 】

[00177]物理空間は、シーン空間中の地面のような2次元空間であり得る。1つのそのような例では、位置空間はシーン空間中の地面であり、ピクセル座標空間中の点は地面における対応する点に投影される。たとえば、地面は、図16のパネルCに示されるようなバスケットボールコートなどの、運動競技場であり得る。ビューの異なるフィールドを有する、および/または異なる視点からキャプチャされる、同じまたは同様の空間中の同様のイベントのビデオストリームの間での、オブジェクトの位置の一貫性をサポートするために、そのような位置空間を使用するのが望ましいことがある。

【0121】

[00178]デジタル画像のピクセル座標空間中の点を地面における対応する点と関連付けることは、シーン空間中のリンクされた点（たとえば、1つまたは複数のセンサおよび/またはデプスカメラからのデータによって示されるような）を地面へと下方向に投影する（または、適切であれば、リンクされた点を地面へと上方向に投影する）ことによって、実行され得る。図17Bは、地面における所定の位置DP20へのリンクされる点LP10（上で論じられたようなオブジェクト上のGPSセンサおよび/または他の位置センサによって示され得る）のそのような投影の例を示す。

【0122】

[00179]代替的に、デジタル画像のピクセル座標空間中の点を地面における対応する点と関連付けることは、図18Aおよび図18Bに示されるような地面にピクセル座標空間からの点を投影することによって実行されてよく、図18Aおよび図18Bは、地面に対して垂直な面におけるキャプチャデバイスの焦点軸を通る断面を示す。そのような投影は、地面に対するキャプチャデバイスの方向角 α が（たとえば、上で説明されたような1つまたは複数の方向センサから）知られている場合には、直接実行され得る。代替的に、方向角 α は、地面の基準（たとえば、競技場の境界線）、水平線（たとえば、明るさの違いによって示される）、オブジェクトの境界となる水平方向の特徴（たとえば、舞台の前および後を示す線）、または少なくとも後方の境界となる水平方向の特徴のような、ビデオフレーム中の基準情報から推測され得る。図18Bに示されるように、地面における距離 r と s の比は、方向角 α によって、焦点面における対応する距離 p と q の間の比と関連する。視野の角度 β が知られており、焦点面における距離 p と q の比が表示空間において映像化されるような対応する距離と等しい（すなわち、これらの距離が、互いに対して、表示のピクセル座標空間中の垂直な線に沿って現れるので）ことが仮定され得る。

【0123】

[00180]1つのそのような例では、タスクT200は、（たとえば、上で説明されたように）キャプチャデバイスの方向および/または画像基準情報を使用して、表示面を地面にマッピングするように実施される。この実装形態では、タスクT200は、（たとえば、図17Aに示されるように）オブジェクトの重心を表示面において下方向にオブジェクトの下部へと投影することによって、表示面における各々の選択されたオブジェクトの位置を決定し、この表示位置を地面の対応する位置にマッピングすることによって、位置空間におけるオブジェクトの位置を決定する。この場合、上で説明された例のいずれかによれば、タスクT200は、地面におけるオブジェクトの位置の平均として、または、地面におけるオブジェクトの位置の編成の位置として、決定される基準の位置を計算するように実施され得る。

【0124】

[00181]タスクT300は、メタデータを少なくとも1つのデジタル画像と（たとえば、ビデオストリームと）関連付けるように実施され得る。そのようなメタデータはまた、選択されたオブジェクトのラベル（たとえば、人の名前）、選択されたオブジェクトの間の幾何学的構成の識別などのような、トラッキングされるオブジェクトに関する他の情報を含み得る。加えて、または代替的に、メタデータは、デバイスの他の要素から、および/または受信されるような少なくとも1つの画像またはビデオストリーム中のメタデータから取得され得る情報（たとえば、カメラの設定、カメラの向き、センサの特性、キャプチャの時間および/または位置）を含み得る。

【 0 1 2 5 】

[00182]タスクT300のそのような実施は、XML（拡張マークアップ言語）、KL V（Key - Length - Value）、*.srt（SubRipファイルフォーマット）、および*.vtt（ウェブビデオテキストトラックフォーマット）の中のいずれかのような、所望のフォーマットへとメタデータを符号化することができる。いくつかのフォーマットでは、メタデータの少なくともいくつかは、デジタル画像またはビデオストリーム内に（たとえば、埋め込まれたデータのために確保されているフレームデータの一部分の中に）含まれ得る。代替的に、タスクT300は、別のファイルとして、ビデオストリームと同期されるメタデータストリームとしてメタデータをパッケージ化するように実施され得る。

10

【 0 1 2 6 】

[00183]ビデオストリームの各フレームについて、たとえば、または各々のn番目のフレーム（ここでnは1よりも大きな整数である）について、または指定される事象（たとえば、選択されたオブジェクトが動いたことのトラッキングタスクによる検出）の発生後に、ビデオストリームに対するタスクT300の繰り返される実体（たとえば、タスクT200およびT300の繰り返される実体）を含むように方法M100を実施するのが望ましいことがある。そのような場合、メタデータファイルまたはストリーム内のメタデータの項目は、そのような項目とビデオストリームの対応するフレームとの所望の関連付けを示すために、タイムスタンプを付与され得る。そのような場合、タスクT300はまた、ビデオとメタデータストリーム（場合によっては1つまたは複数のオーディオストリームも伴う）とを多重化することを含み得る。

20

【 0 1 2 7 】

[00184]メタデータを含むストリームが送信および/または記憶されるべきである場合、他の情報および/またはフォーマット化が、特定のストリームフォーマットに適合するように実行され得る（たとえば、メタデータおよび/またはストリームは、ヘッダプラスペイロードフォーマットへとパケット化され得る）。タスクT300は、メタデータを含むストリームを、ストリームを符号化するように構成され得る圧縮器または他のエンコーダに提供するように実施され得る。そのようなストリームを記憶するために使用され得る記憶ファイルフォーマットの例は、次のビデオコンテナフォーマット、すなわち、AVI、WMV、MOV、MKV、MPG、およびMP4のいずれかを含む。

30

【 0 1 2 8 】

[00185]いくつかの適用例では、選択されたオブジェクトの幾何学的構成は、位置空間における多数の選択されたオブジェクト（たとえば、公共空間における群衆、動物の大群、空中または水上の乗物または他のオブジェクトの大群）の分布であり得る。タスクT300が各々の個々のオブジェクトの位置を示すメタデータを生成することは不要であることがあり、選択されたオブジェクトの厳密な総数も不要であることがある。

【 0 1 2 9 】

[00186]そのような場合、タスクT300は、複数の幾何学的構成の候補のうちのいずれが位置空間中の選択されたオブジェクトの分布と最も類似しているかを決定するように実施され得る。たとえば、基準の位置（たとえば、群衆の中の話者の位置）に対する選択されたオブジェクトの所与の（たとえば、観測される）分布に最も近い候補を決定するように、タスクT300を実施するのが望ましいことがある。

40

【 0 1 3 0 】

[00187]図36は、図面の右側の10個の候補の各々の中心にある点によって示される基準の位置（たとえば、対象のオブジェクトまたは「プライマリ」オブジェクトの位置）に対して配置されている多数のオブジェクトの分布を照合するようにタスクT300が実施され得る、一例を示している。この場合、図の左側の図に示されるように、円形の位置空間が16個の領域に分割され、各領域中のオブジェクトの数が決定される。位置空間は円形である必要はない（たとえば、正方形、長方形、楕円などであってよい）。

【 0 1 3 1 】

50

[00188]図36の左の4つの大きな図は、この選択のための決定基準として使用され得る4つの占有度の基準のセット、すなわち、(A)内側の領域のオブジェクトの数と外側の領域のオブジェクトの数との比、(B)最も混雑している半分の中のオブジェクトの数と他の半分の中のオブジェクトの数との比(代替的には、最も混雑している半分におけるオブジェクトの数とオブジェクトの総数との比)、(C)最も混雑している4分の1におけるオブジェクトの数と反対側の4分の1におけるオブジェクトの数との比、および、(D)最も混雑している対向する2つの4分の1におけるオブジェクトの数と他の2つの4分の1におけるオブジェクトの数との比(代替的には、最も混雑している対向する2つの4分の1におけるオブジェクトの数とオブジェクトの総数との比)の一例を示す。

【0132】

10

[00189]選択されたオブジェクトの数が大きいつきでも、位置空間におけるオブジェクトの特定の分布に対するこれらの4つの尺度の値は、非常に低い計算的な複雑さで計算され得る。次いで、たとえば、所与の分布に対するこれらの尺度の値に従って、図36の右側にある10個の幾何学的構成の候補の中から選択するために、決定木が使用され得る。そのような場合、タスクT300は、選択された候補を特定し選択されたオブジェクトの総数を示すメタデータを生成するように実施され得る。メタデータのサイズをさらに低減するために、タスクT300は、選択されたオブジェクトの総数の近似を(たとえば、総数を指定された量子化係数で割り、その結果を最も近い整数へと丸めることによって)示すために、メタデータを生成するように実施され得る。

【0133】

20

[00190]上の4つの占有度の尺度に対する代替形態として、タスクT300は、(A)オブジェクトの総数の指示(場合によっては上で説明されたような近似)、および(B)対応する領域が占有されるかどうかを各ビットの状態が示す16ビットの数として、分布を符号化するメタデータを生成するように実施され得る。そのような場合、近隣の領域が他の占有されている領域よりもはるかに多く占有されている場合、エンコーダが空いている領域の占有度を示すのが望ましいことがある。他の代替形態は、分布の方向の軸(たとえば、2次モーメント)を計算することを含み得る。

【0134】

[00191]メタデータにタスクT300によって含められ得る他の特性は、領域の間の境界の可変性の指示と、各々の個別のオブジェクトの経時的な移動の程度(ブラウン運動)とを含む。メタデータは、分布の追加の態様を示すために(たとえば、オブジェクトのグループ内での異なる位置における異なる程度の活動の指示を符号化するために)、重複する形状(たとえば、候補の異なるセットから場合によっては選択される、各分布に対する2つ以上の候補)を含み得る。

30

【0135】

[00192]図36の右に示される10個の候補の例示的なセットにおいて、プライマリオブジェクトに向かって集中している3つの形状は、周辺に配置された異常値オブジェクトの存在を示唆する(そうでなければ、そのような集中のない形状が選択されているであろう)。対応するデコーダは、周辺の中にオブジェクトの総数のうちの小さな部分を分布させることによって、そのような異常値を表すように構成され得る。代替的には、エンコーダは、スケーリングのときにそのような異常値を無視するように構成され得る。

40

【0136】

[00193]タスクT300はまた、図33A、図33B、図34A~図34Cと、編成、編成データ、および編成コードブックの関連する議論とを参照して下で説明されるように、幾何学的構成の中から選択する追加の例のいずれかを実行するように実施され得る。本明細書で説明されるような方法M100の実装形態は、選択されたオブジェクトの間の特定の幾何学的構成に対する一致を示すものとしてインデクシングされるフレームが、メタデータを探索することにより検索のために容易に識別され得るようにし、それによってビデオストリーム全体を確認する必要をなくすために、(たとえば、ビデオストリームのキャプチャの間に)メタデータを提供するために使用され得る。

50

【 0 1 3 7 】

[00194]図 2 0 を参照すると、オブジェクトトラッキングを使用したビデオ処理を示す特定の実施形態が示される。図 2 0 に示される実施形態に関して説明されるビデオ処理技法は、図 1 の電子デバイス 1 0 2、図 2 B のプロセッサ 2 6 4 によって実装されるオブジェクトトラッキングおよび検出モジュール 2 0 4、プロセッサによって実装される図 8 のオブジェクトトラッキングおよび検出モジュール 8 0 4、プロセッサによって実装される図 9 の平滑化モジュール 9 6 1、またはこれらの任意の組合せによって実行され得る。

【 0 1 3 8 】

[00195]図 2 0 は、第 1 の時刻においてキャプチャされるようなシーンを示す、デジタル画像 1 7 0 0 を示す。たとえば、画像 1 7 0 0 は、モバイルデバイスの画面に表示されるべきビデオストリームに対応し得る。モバイルデバイスは、ストリームを（たとえば、モバイルデバイスのカメラによって）キャプチャし、または別のデバイスからストリームを受信するように構成され得る。フレーム 1 7 0 0 に示されるシーンは、第 1 のオブジェクト 1 7 0 2 と、第 2 のオブジェクト 1 7 0 4 と、第 3 のオブジェクト 1 7 0 6 と、星とを含む。示される実施形態では、第 1 のオブジェクト 1 7 0 2 は第 1 の人物に対応してよく、第 2 のオブジェクト 1 7 0 4 は木に対応してよく、第 3 のオブジェクト 1 7 0 6 は第 2 の人物に対応してよい。別の実施形態では、第 1 のオブジェクト 1 7 0 2 および第 3 のオブジェクト 1 7 0 6 の少なくとも 1 つは、モバイルデバイスを介して制御されるロボットに対応し得る。

【 0 1 3 9 】

[00196]図 2 0 はまた、モバイルデバイスの画面上の画像 1 7 0 0 の表示 1 7 1 0 を示す。モバイルデバイスの画面（たとえば、タッチスクリーン）は、ビューファインダに対応し得る。表示 1 7 1 0 は、ユーザが（たとえば、タスク T 2 0 0 による）トラッキングのためにオブジェクトを選択することを可能にする、ユーザ選択機構 1 7 1 1 を含み得る。たとえば、ユーザは、画面上のユーザ選択機構 1 7 1 1 をタッチし、第 1 のオブジェクト 1 7 0 2 のトラッキングを選択して可能にするために第 1 のオブジェクト 1 7 0 2 の周囲で第 1 の焦点リング 1 7 1 2 をドラッグすることによって、タスク T 1 0 0 を実行することができる。第 1 の焦点リング 1 7 1 2 のそのような配置に応答して、モバイルデバイスは、トラッキングのためにフレーム 1 7 0 0 中の第 1 のオブジェクト 1 7 0 2 を選択することができる。同様の方式で、ユーザはまた、画面上のユーザ選択機構 1 7 1 1 をタッチし、第 2 のオブジェクト 1 7 0 4 および第 3 のオブジェクト 1 7 0 6 の周囲で第 2 の焦点リング 1 7 1 4 および第 3 の焦点リング 1 7 1 6 をドラッグして、それぞれ、第 2 のオブジェクト 1 7 0 4 および第 3 のオブジェクト 1 7 0 6 のトラッキングを選択して可能にすることができる。第 2 の焦点リング 1 7 1 4 および第 3 の焦点リング 1 7 1 6 のそのような配置に応答して、モバイルデバイスは、トラッキングのために、それぞれ第 2 のオブジェクト 1 7 0 4 と第 3 のオブジェクト 1 7 0 6 とを選択することができる。

【 0 1 4 0 】

[00197]図 2 0 はまた、（たとえば、第 1 の時刻の後の）第 2 の時刻にキャプチャされるようなシーンを示す、画像 1 7 2 0 を示す。画像 1 7 2 0 において、各オブジェクト 1 7 0 2 ~ 1 7 0 6 の位置は、第 1 の時刻におけるシーンを示す画像 1 7 0 0 中の対応する位置に対して変化している。たとえば、画像 1 7 2 0 において、第 1 のオブジェクト 1 7 0 2 および第 3 のオブジェクト 1 7 0 6 が移動している。第 2 のオブジェクト（たとえば、木）は静止しているオブジェクトであるが、第 2 のオブジェクト 1 7 0 4 は動いたように見えることがある。たとえば、モバイルデバイス（または他のキャプチャデバイス）の位置が動くことがあり、それによって、第 2 のオブジェクト 1 7 0 4 が動いたという錯覚が生まれることがある。

【 0 1 4 1 】

[00198]モバイルデバイスは、選択されたオブジェクト 1 7 0 2 ~ 1 7 0 6 の間の 1 つまたは複数の空間的な関係（「幾何学的配置（geometry）」とも呼ばれる）を監視することができる。たとえば、モバイルデバイスは、第 1 のオブジェクト 1 7 0 2 と第 2 のオブ

ジェクト1704との間の第1の距離と、第2のオブジェクト1704と第3のオブジェクト1706との間の第2の距離と、第3のオブジェクト1706と第1のオブジェクト1702との間の第3の距離とをトラッキングして測定することができる。特定の実施形態において、（たとえば、タスクT200によって）各々の選択されたオブジェクト1702～1706の間の幾何学的配置をトラッキングすることは、トリガ事象があると開始し得る。非限定的な例として、各々の選択されるオブジェクト1702～1706の間の幾何学的配置をトラッキングすることは、第1のオブジェクト1702が第2のオブジェクト1704と交差すると開始し得る。他の実施形態において、各々の選択されたオブジェクト1702～1706の間の幾何学的配置をトラッキングすることは、ユーザ入力があると開始し得る。非限定的な例として、各々の選択されるオブジェクト1702～1706の間の幾何学的配置をトラッキングすることは、オブジェクト1702～1706の2つ以上をユーザが選択すると開始し得る。別の例として、各々の選択されたオブジェクト1702～1706の間の幾何学的配置をトラッキングすることは、トラッキングモードを開始することをユーザが選択すると開始し得る。

【0142】

[00199]図20はまた、（たとえば、表示タスクT50による）モバイルデバイスの画面上の画像1720の表示1730を示す。モバイルデバイスは、表示1730上での幾何学的配置の指示を含み得る。たとえば、第1のオブジェクト1702と第2のオブジェクト1704との間の第1の距離の第1の指示(L_{12})は、表示1730に含まれ得る。第1の距離は、画像1720の中での第1のオブジェクト1702の位置、および画像1720の中での第2のオブジェクト1704の位置に基づいて決定され得る。たとえば、モバイルデバイスは、第1のオブジェクト1702の位置と、第2のオブジェクト1704の位置とをトラッキングして、第1の指示を生成することができる(L_{12})。第2のオブジェクト1704と第3のオブジェクト1706との間の第2の距離の第2の指示(L_{23})も、表示1730に含まれ得る。第2の距離は、画像1720の中での第2のオブジェクト1704の位置と、画像1720の中での第3のオブジェクト1706の位置とに基づいて決定され得る。たとえば、モバイルデバイスは、第2のオブジェクト1704の位置と、第3のオブジェクト1706の位置とをトラッキングして、第2の指示を生成することができる(L_{23})。第1のオブジェクト1702と第3のオブジェクト1706との間の第3の距離の第3の指示(L_{13})も、表示1730に含まれ得る。第3の距離は、フレーム1720の中での第1のオブジェクト1702の位置と、フレーム1720の中での第3のオブジェクト1706の位置とに基づいて決定され得る。たとえば、モバイルデバイスは、第1のオブジェクト1702の位置と、第2のオブジェクト1704の位置とをトラッキングして、第3の指示を生成することができる(L_{13})。

【0143】

[00200]モバイルデバイスはまた、オブジェクト1702～1706の複合的な幾何学的配置（たとえば、選択されたオブジェクトの幾何学的構成）をトラッキングすることができる。たとえば、示される実施形態では、オブジェクト1702～1706の幾何学的構成は、各指示(L_{12} 、 L_{23} 、 L_{13})によって形成される三角形として特徴付けられ得る。各オブジェクト間の幾何学的配置の少なくとも1つのパラメータ、または、複合的な幾何学的配置の少なくとも1つのパラメータが、画像（たとえば、1つまたは複数のビデオストリームのフレーム）をクラスタ化するために使用され得る。非限定的な例として、2つの特定のオブジェクト1702～1706の間の距離（たとえば、第1の距離、第2の距離、および/または第3の距離）に基づいて、特定の画像がクラスタ化され得る。別の非限定的な例として、複数のオブジェクト1702～1706の重心(C_{M2})に基づいて、特定の画像がクラスタ化され得る。たとえば、モバイルデバイスは、各画像に対する各指示(L_{12} 、 L_{23} 、 L_{13})によって形成される三角形の重心(C_{M2})を決定することができる。特定の実施形態では、重心(C_{M2})は、計算され、インデクシングされ、画面1730に表示され得る。別の実施形態では、重心(C_{M2})は、計算されインデクシングされるだけであり得る。実質的に同様の重心を有するフレームは、メモリの中へと一緒にクラ

スタ化されインデクシングされ得る。そのようなトラッキングの間、表示の中心が上記の重心と一致するように、表示変換動作を制御するのが望ましいことがある。加えて、または代替的に、選択されたオブジェクトのすべてが表示内で可視のままであるように、カメラの光学および／またはデジタルズーム動作を制御するのが望ましいことがある。

【0144】

[00201] 特定の実施形態では、モバイルデバイスは、各オブジェクト1702～1706の間の幾何学的配置および／または各フレームに対する複合的な幾何学的配置をトラッキングする（たとえば、フレームごとに幾何学的配置をトラッキングする）ことができる。他の実施形態では、モバイルデバイスは、非連続的な間隔において各オブジェクト1702～1706の間の幾何学的配置および／または複合的な幾何学的配置をトラッキングする（たとえば、非連続的なフレームにおいて幾何学的配置をトラッキングする）ことができる。非連続的なフレームにおいて幾何学的配置をトラッキングすることは、モバイルデバイスにおいて消費される電力の量を減らし（バッテリー持続時間を延長し）、幾何学的配置情報および／またはクラスタ化情報と関連付けられる情報をインデクシングするために使用されるメモリの量を減らすことができる。

【0145】

[00202] たとえば、特定の実施形態では、モバイルデバイスは、各フレームにおいて各オブジェクト1702～1706（たとえば、3つのオブジェクト）をトラッキングすることができる。トラッキングの間、各オブジェクト1702～1706は、オブジェクト1702～1706の位置およびオブジェクト1702～1706のサイズと関連付けられる少なくとも3つの値によって表され得る。たとえば、ある値はオブジェクト1702～1706のx座標に対応してよく、ある値はオブジェクト1702～1706のy座標に対応してよく、別の値はオブジェクト1702～1706のサイズ（たとえば、バウンディングボックス）に対応してよい。各値は、32ビット（たとえば、4バイト）の数に対応してよい。したがって、少なくとも288ビットのデータ（3つのオブジェクト×3つの値／オブジェクト×32ビット／値）が、トラッキングの間に収集され、各フレームのためのメモリへとロギングされ得る。モバイルデバイスが毎秒30フレームをキャプチャすると仮定すると、1時間（たとえば、3600秒）フレームごとに3つのオブジェクト1702～1706をトラッキングしてインデクシングするには、少なくとも3,888,000バイトのデータが収集され、メモリへとロギングされる。

【0146】

[00203] しかしながら、非連続的なフレームにおける幾何学的配置をトラッキングすることは、メモリの要件を緩和し得る。非限定的な例として、1時間、15フレームごとの間隔で各オブジェクト1702～1706をトラッキングしてインデクシングすることは、メモリ空間に93.33%の節約をもたらし得る。非連続的なフレームにおいて幾何学的配置をトラッキングすることは、オブジェクト1702～1706と関連付けられる動きが比較的少量である状況において、特に有益であり得る。モバイルデバイスは、ユーザ入力（たとえば、ユーザ選好）に基づいてフレームをトラッキングしてインデクシングする頻度を決定することができ、および／または、たとえば、選択されたオブジェクトの1つまたは複数の各々の位置におけるユーザ指示および／または経時的な変化の程度に基づいて、その頻度を変化させることができる。

【0147】

[00204] 各オブジェクト1702～1706はまた、画面1730上のオブジェクト1702～1706の位置に対応する、個々の重心を有し得る。たとえば、第1のオブジェクト1702は、第1のオブジェクト1702の位置に重心(C_{01})を有してよく、第2のオブジェクト1704は、第2のオブジェクト1704の位置に重心(C_{02})を有してよく、以下同様である。各オブジェクト1702～1706の重心はまた、モバイルデバイスがトラッキングする幾何学的パラメータであり得る。

【0148】

[00205] 図20はまた、（たとえば、第2の時刻の後の）第3の時刻にキャプチャされ

10

20

30

40

50

るようなシーンを示す、画像 1740 を示す。画像 1740 において、各オブジェクト 1702 ~ 1706 の位置は、第 2 の時刻における画像 1720 の中で示される対応する位置に対して変化している。

【0149】

[00206]図 20 はまた、(たとえば、表示タスク T50 による) モバイルデバイスの画面上の画像 1740 の表示 1750 を示す。モバイルデバイスは、表示 1750 に幾何学的配置の指示を表示し得る。たとえば、第 1 のオブジェクト 1702 と第 2 のオブジェクト 1704 との間の第 1 の距離の第 1 の指示 (L_{12}) は表示 1750 に含まれてよく、第 2 のオブジェクト 1704 と第 3 のオブジェクト 1706 との間の第 2 の距離の第 2 の指示 (L_{23}) は表示 1750 に含まれてよく、第 1 のオブジェクト 1702 と第 3 のオブジェクト 1706 との間の第 3 の距離の第 3 の指示 (L_{13}) は表示 1750 に含まれてよい。加えて、モバイルデバイスはまた、オブジェクト 1702 ~ 1706 の複合的な幾何学的配置をトラッキングすることができる。たとえば、示される実施形態では、オブジェクト 1702 ~ 1706 の複合的な幾何学的配置は、各指示 (L_{12} 、 L_{23} 、 L_{13}) によって形成される三角形に対応し得る。

【0150】

[00207]別の実施形態では、モバイルデバイスは、各オブジェクト 1702 ~ 1706 の動きをトラッキングし、各オブジェクト 1702 ~ 1706 の経時的な動きを表示するプロットのシーケンスを生成することができる。モバイルデバイスは、各フレームにおいて各オブジェクト 1702 ~ 1706 に対するプロットを生成することができ、上で説明されたように、メモリの要件を緩和して電力の節約を改善するために、非連続的なフレームにおいて各オブジェクト 1702 ~ 1706 に対するプロットを生成することができる。

【0151】

[00208]表示上で現在見えていない(たとえば、現在、ビデオストリームによってキャプチャされるシーン内にない)オブジェクトに対する 1 つまたは複数の選択されたオブジェクトの空間的な関係をトラッキングするようにデバイスを構成すること(たとえば、タスク T200 を実施すること)も可能である。たとえば、現在画面上にない固定された(基準の)オブジェクトに対する移動するオブジェクトの空間的な関係がトラッキングされ得る。特定の例では、デバイスは、選択された選手がゴールまたはバスケットから 20 フィート以内にいるとき、および/または、ゴール/バスケットに向かって動いているときには、ゴール/バスケットが現在画面上にないとしても、常に指し示すように構成され得る。そのような場合、デバイスは、現在表示されているシーンよりも大きな地図を作成するように構成され得る。カメラの位置が(たとえば、GPS センサのような位置センサによって示されるように)知られている場合、カメラの 1 つまたは複数の方向センサが、現在のビューがより大きなシーンの地図とどのように関連するかを決定するために使用され得る。そのような 1 つまたは複数の方向センサは、重力軸(たとえば、デバイスおよび地球の中心を通る軸)に対するカメラの方向を示すための、1 つまたは複数の慣性センサ(たとえば、ジャイロスコープおよび/または加速度計)、および/または 1 つまたは複数の磁場センサ(たとえば、磁力計)を含み得る。加えて、または代替的に、スポーツのフィールドまたはコート上の印のような、または舞台のシーンの固定物のような、固定された基準のオブジェクトである可視の地物が、基準として使用され得る。基準のオブジェクトの外観は、そのオブジェクトが見られる角度が変化するにつれて変化し得るが、そのようなマッピングは、動いているカメラに対しても行われ得る。

【0152】

[00209]図 20 に示される実施形態は、モバイルデバイスが、図 21 に関して下で説明されるように、画像 1700、1720、1740 の中のオブジェクト 1702 ~ 1706 の幾何学的配置をトラッキングし、実質的に同様の幾何学的配置を有するフレームを検索のためにクラスタ化することを可能にし得る。モバイルデバイスは、単一のカメラを使用して、または複数のカメラを使用して、図 20 に関して説明される技法を実行すること

ができる。

【0153】

[00210]本明細書で説明されるようなメタデータを生成するための使用法および適用例は、スポーツイベント（たとえば、サッカー、バスケットボール、フットボール、ホッケー）、社会的なイベント（たとえば、結婚式、パーティ、ダンス）、演芸（たとえば、演劇、学芸会、コンサート、デュエット）、およびセキュリティまたは調査のための監視のいずれかを含み得る。モバイルデバイスはまた、頭部に装着されるディスプレイであってよく、ユーザの命令を検出するために、ユーザの選択が、たとえば、ジョイスティック、ジェスチャ認識カメラシステム、またはタッチセンサおよび/もしくは方向センサ（上で説明されたような1つまたは複数の方向センサ）を装備した手袋を使用することにより実行される。ビデオストリームは、生であってよく（たとえば、デバイスによってキャプチャされ、または、Bluetooth（登録商標）のような短距離接続を介して、またはインターネットへのWi-Fi（登録商標）接続のようなネットワークを介して、別のキャプチャデバイスからストリーミングされてよく）、または（やはりそのデバイスまたは別のデバイスによって）記録されてよい。

10

【0154】

[00211]ストリーム内での選択されたオブジェクトの間の相対的な位置および/または空間的な関係に従って、ビデオストリームのフレームを分類するのが望ましいことがある。図21を参照すると、クラスタ化に基づいてフレームを検索するために使用されるモバイルデバイスの表示1800の特定の実施形態が示されている。表示1800は、水平軸と垂直軸とを有するグラフを含み得る。水平軸は、モバイルデバイスによってキャプチャされる異なるビデオファイルに対応し得る。たとえば、代替的な実施形態では、モバイルデバイスは、第1のビデオファイルと第2のビデオファイルとをキャプチャした。垂直軸は、ビデオファイルの時間インデックスに対応し得る。

20

【0155】

[00212]各ビデオファイルは、ビデオファイル中の選択されたオブジェクトの幾何学的配置と関連付けられる少なくとも1つのパラメータに基づく、クラスタを含み得る。たとえば、ビデオファイル中の各クラスタは、選択されたオブジェクト間で実質的に同様の幾何学的配置を有する、フレームのグループを含み得る。図21に示されるクラスタは、オブジェクト間の重心に基づき得る。しかしながら、他の実施形態では、クラスタは1つまたは複数の異なるパラメータ（たとえば、形状、距離、空間的な特性、カラースキームなど）に基づき得る。パラメータはユーザによって選択され得る。

30

【0156】

[00213]第1のビデオの10:00前後のマークまでの、第1のビデオの初めにあるフレームは、第1のクラスタ（たとえば、クラスタ1）へとグループ化され得る。第1のクラスタは、選択されたオブジェクト間の画面1800上での重心が150前後のx座標と250前後のy座標とを有するような、フレームに対応し得る。図20の第2のフレームを参照すると、選択されたオブジェクト1702~1706の重心（ C_{M2} ）（たとえば、三角形の重心（ C_{M2} ））は、画面1730上の150前後のx座標と、画面1730上の250前後のy座標とを有し得る。したがって、第2のフレーム（およびオブジェクト1702~1706の間の実質的に同様の重心を有する他のフレーム）が、図21の第1のクラスタに配置され得る。したがって、ユーザが第1のクラスタを選択する場合、モバイルデバイスは、選択されたオブジェクト1702~1706が図20の第2のフレームと実質的に同様の構成を有する、対応するビデオストリーム（たとえば、ビデオの部分）を表示することができる。

40

【0157】

[00214]第1のビデオの10:00前後のマークから第1のビデオの20:00前後のマークまでのフレームは、第2のクラスタ（たとえば、クラスタ2）へとグループ化され得る。第2のクラスタは、選択されたオブジェクト間の画面1800上での重心が100前後のx座標と250前後のy座標とを有するような、フレームに対応し得る。図20の

50

第3のフレームを参照すると、選択されたオブジェクト1702～1706の重心(C_{M3})は、画面1750上の100前後のx座標と、画面1750上の250前後のy座標とを有し得る。したがって、第3のフレーム(およびオブジェクト1702～1706の間の実質的に同様の重心を有する他のフレーム)が、図21の第2のクラスタに配置される。したがって、ユーザが第2のクラスタを選択する場合、モバイルデバイスは、選択されたオブジェクト1702～1706が図21の第3のフレームと実質的に同様の構成を有する、対応するビデオストリーム(たとえば、ビデオの部分)を表示することができる。

【0158】

[00215]追加のクラスタ(たとえば、第3のクラスタ)は、異なる重心(たとえば、175のx座標および325のy座標)を有するフレームに対応し得る。ユーザは、ビデオの特定の部分と関連付けられる構成(たとえば、重心)を有するクラスタを選択することによって、ビデオの特定の部分を検索する(たとえば、位置特定する)ことができる。したがって、選択されたオブジェクト1702～1706の幾何学的パラメータ(たとえば、距離、各々のおよび/または選択されたオブジェクトの重心)に基づくクラスタへとフレームをインデクシングすることによって、モバイルデバイスのユーザは、ビデオを見進める(たとえば、再生する、早送りする、巻き戻すなど)必要なく、ビデオの所望の部分容易に位置特定することができる。クラスタへとフレームをインデクシングすることで、ユーザは、どの構成(たとえば、幾何学的配置)が最も頻繁に発生したか、最も稀に発生したかなどを、見る事が可能になり得る。そのような幾何学的パラメータのそのようなインデクシングおよび/または値は、上で説明されたようなタスクT300によって生成されるメタデータの中に含まれ得る。

【0159】

[00216]図21はまた、ビデオストリームのフレームのクラスタと関連付けられる領域の表示1802を含む。たとえば、クラスタの表示1802は、図20のシーンのビデオストリームに対応し得る。表示1802は10個の領域($C1 \sim C10$)を含む。示される実施形態では、領域($C1 \sim C10$)は、特定のフレームにおけるオブジェクト1702～1706の重心に基づく。他の実施形態では、領域($C1 \sim C10$)は、他の幾何学的パラメータ(たとえば、各オブジェクト1702～1706の重心、形状、距離、空間的特性、カラスキームなど)に基づき得る。特定の実施形態では、各領域($C1 \sim C10$)は等しいサイズを有し得る(たとえば、各領域($C1 \sim C10$)は「ボロノイセル(Voronoi Cell)」であり得る)。

【0160】

[00217]各領域($C1 \sim C10$)は、特定のクラスタと関連付けられる重心に対応するx座標とy座標とを含み得る。非限定的な例として、第1の領域($C1$)に対応する重心は、75というx座標と580というy座標とを有し得る。加えて、第2の領域($C2$)に対応する重心は、215というx座標と580というy座標とを有し得る。図20のシーンのビデオストリーム中のフレームは、オブジェクト1702～1706の重心によってインデクシングされ得る。各フレームは、重心に基づいて領域($C1 \sim C10$)の1つに対応するクラスタに配置され得る。たとえば、第1の領域($C1$)の重心に最も近い重心を有するフレームは、対応する第1のクラスタ内に配置され得る。そのようなクラスタ化情報は、タスクT300によって生成され、(たとえば、本明細書で説明されるようなビデオストリームに同期されるメタデータストリーム中の)メタデータとして対応するフレームと関連付けられ得る。そのような場合、特定のクラスタ内にインデクシングされるフレームは、メタデータを探索することによって、検索のために容易に識別されることが可能であり、ビデオストリーム全体を確認する必要をなくす。

【0161】

[00218]モバイルデバイスは、特定のクラスタをユーザが選択したことに応答して、特定の幾何学的配置(たとえば、重心)を有するビデオフレームを表示することができる。たとえば、ユーザが第1の領域($C1$)を選択する(たとえば、タッチする)場合、モバ

イルデバイスは、オブジェクト1702～1706の重心が第1の(C1)領域内にある、ビデオフレームのクラスタを表示することができる。特定の実施形態では、ユーザは、第1の領域(C1)において画面にタッチし、閾値を超える時間の期間(たとえば、3秒)、第1の領域(C1)をホールドすることができる。3秒間第1の領域をホールドした後、第1の領域(C1)は表示1802上で拡大され得る。たとえば、表示1802は、1804に示されるように第1の領域(C1)を示し得る。このモードでは、第1の領域1804は、特定のフレームの重心($C_{M1023-1026}$)に基づいて、特定のフレーム(たとえば、フレーム1023からフレーム1026)を示し得る。特定の実施形態では、ユーザは特定のフレーム(たとえば、フレーム1024)を選択することができ、モバイルデバイスはフレーム1024においてビデオストリームの再生を開始することができる。

10

【0162】

[00219]図22を参照すると、オブジェクトトラッキングを使用してビデオ処理方法1900の特定の実施形態を示すフローチャートが示される。代替的な実施形態では、方法1900は、図1の電子デバイス102、図2Bのプロセッサ264によって実装されるオブジェクトトラッキングおよび検出モジュール204、プロセッサによって実装される図8のオブジェクトトラッキングおよび検出モジュール804、プロセッサによって実装される図9の平滑化モジュール961、またはこれらの任意の組合せを使用して実行され得る。

【0163】

[00220]方法1900は、1902において、モバイルデバイスにおいて、シーン中の複数のオブジェクトの選択を受け取ることを含む。たとえば、図20を参照すると、ユーザは、画面1710上のユーザ選択機構1711をタッチし、第1のオブジェクト1702の周囲で第1の焦点リング1712をドラッグし、第2のオブジェクト1704の周囲で第2の焦点リング1714をドラッグし、第3のオブジェクト1706の周囲で第3の焦点リング1716をドラッグして、それぞれ、第1、第2、および第3のオブジェクト1702～1706のトラッキングを可能にすることができる。表示されるシーン1700は、キャプチャされモバイルデバイスの画面に表示されるビデオストリームに対応し得る。モバイルデバイスは、ストリームを(たとえば、モバイルデバイスのカメラによって)キャプチャし、または別のデバイスからストリームを受信するように構成され得る。

20

【0164】

[00221]1904において、各オブジェクト間の幾何学的配置がトラッキングされ得る。たとえば、図20を参照すると、モバイルデバイスは、各々の選択されたオブジェクト1702～1706間の幾何学的配置をトラッキングすることができる。たとえば、モバイルデバイスは、第1のオブジェクト1702と第2のオブジェクト1704との間の第1の距離と、第2のオブジェクト1704と第3のオブジェクト1706との間の第2の距離と、第3のオブジェクト1706と第1のオブジェクト1702との間の第3の距離とをトラッキングして測定することができる。

30

【0165】

[00222]1906において、幾何学的配置の指示が画面上に表示され得る。たとえば、図20を参照すると、第1のオブジェクト1702と第2のオブジェクト1704との間の第1の距離の第1の指示(L_{12})は、表示1730に含まれ得る。第1の距離は、画像1720の中での第1のオブジェクト1702の位置と、画像1720中での第2のオブジェクト1704の位置とに基づいて決定され得る。たとえば、モバイルデバイスは、第1のオブジェクト1702の位置と、第2のオブジェクト1704の位置とをトラッキングして、第1の指示を生成することができる(L_{12})。第2のオブジェクト1704と第3のオブジェクト1706との間の第2の距離の第2の指示(L_{23})も、表示1730に含まれ得る。第2の距離は、画像1720の中での第2のオブジェクト1704の位置と、画像1720の中での第3のオブジェクト1706の位置とに基づいて決定され得る。たとえば、モバイルデバイスは、第2のオブジェクト1704の位置と、第3のオブジェクト1706の位置とをトラッキングして、第2の指示を生成することができる(L_{23})

40

50

。第1のオブジェクト1702と第3のオブジェクト1706との間の第3の距離の第3の指示(L_{13})も、表示1730に含まれ得る。第3の距離は、画像1720の中での第1のオブジェクト1702の位置と、画像1720の中での第3のオブジェクト1706の位置とに基づいて決定され得る。たとえば、モバイルデバイスは、第1のオブジェクト1702の位置と、第2のオブジェクト1704の位置とをトラッキングして、第3の指示を生成することができる(L_{13})。

【0166】

[00223]図22の方法1900は、モバイルデバイスが、画像1700、1720、1740の中のオブジェクト1702~1706の幾何学的配置をトラッキングし、実質的に同様の幾何学的配置を有するフレームを検索のためにクラスタ化することを可能にし得る。モバイルデバイスは、単一のカメラを使用して、または複数のカメラを使用して、方法1900を実行することができる。さらなる例では、方法は、1つまたは複数のオブジェクトの間の指定された空間的な関係が(たとえば、タスクT200によって)検出されるとき、たとえば、第1の選択されたオブジェクトと第2の選択されたオブジェクトとの間の距離が指定された閾値の距離よりも小さくなる(代替的には大きくなる)とき、警告条件を示すことを含み得る。

【0167】

[00224]図23を参照すると、オブジェクトトラッキングを使用したビデオ処理方法2000、2010の特定の実施形態を示すフローチャートが示される。代替的な実施形態では、方法2000、2010は、図1の電子デバイス102、図2Bのプロセッサ264によって実装されるオブジェクトトラッキングおよび検出モジュール204、プロセッサによって実装される図8のオブジェクトトラッキングおよび検出モジュール804、プロセッサによって実装される図9の平滑化モジュール961、またはこれらの任意の組合せを使用して実行され得る。

【0168】

[00225]方法2000は、2002において、モバイルデバイスにおいて、シーン中の複数のオブジェクトの選択を受け取ることを含む。たとえば、図20を参照すると、ユーザは、画面1710上のユーザ選択機構1711をタッチし、第1のオブジェクト1702の周囲で第1の焦点リング1712をドラッグし、第2のオブジェクト1704の周囲で第2の焦点リング1714をドラッグし、第3のオブジェクト1706の周囲で第3の焦点リング1716をドラッグして、それぞれ、第1、第2、および第3のオブジェクト1702~1706のトラッキングを可能にすることができる。表示されるシーン1700は、キャプチャされモバイルデバイスの画面に表示されるビデオストリームに対応し得る。モバイルデバイスは、ストリームを(たとえば、モバイルデバイスのカメラによって)キャプチャし、または別のデバイスからストリームを受信するように構成され得る。

【0169】

[00226]2004において、各オブジェクト間の幾何学的配置がトラッキングされ得る。たとえば、図20を参照すると、モバイルデバイスは、各々の選択されたオブジェクト1702~1706間の幾何学的配置をトラッキングすることができる。たとえば、モバイルデバイスは、第1のオブジェクト1702と第2のオブジェクト1704との間の第1の距離と、第2のオブジェクト1704と第3のオブジェクト1706との間の第2の距離と、第3のオブジェクト1706と第1のオブジェクト1702との間の第3の距離とをトラッキングして測定することができる。加えて、モバイルデバイスはまた、オブジェクト1702~1706の複合的な幾何学的配置をトラッキングすることができる。たとえば、示される実施形態では、オブジェクト1702~1706の複合的な幾何学的配置は、3つの指示(L_{12} 、 L_{23} 、 L_{13})によって形成され得る三角形に対応し得る。

【0170】

[00227]2006において、ビデオストリームのフレームは、幾何学的配置と関連付けられる少なくとも1つのパラメータに基づいてクラスタ化され得る。たとえば、図21を参照すると、ビデオファイル中の各クラスタは、選択されたオブジェクト1706~17

08間で実質的に同様の幾何学的配置を有する、フレームのグループを含み得る。図21に示されるクラスタは、オブジェクト1706~1708の間の重心(C_{M2} 、 C_{M3}) (たとえば、三角形の重心)に基づき得る。

【0171】

[00228] 特定の実施形態では、方法2000は、オブジェクトが特定の幾何学的配置にあるときに、フレームを検索することを含み得る。たとえば、図21を参照すると、モバイルデバイスは、フレーム中のオブジェクト1702~1706の重心(C_{M2} 、 C_{M3})に基づいてフレームをインデクシングすることができる。特定の重心(たとえば、特定の幾何学的配置)を有するフレームは、特定の重心と関連付けられるクラスタを選択することによって容易に実現され得る。たとえば、モバイルデバイスは、図21の画面1800上の第1のクラスタ(たとえば、クラスタ1)を選択することによって、オブジェクト1702~1706が図20の第2のフレーム中のオブジェクト1702~1706と実質的に同様の幾何学的配置を有するフレームを検索することができる。たとえば、第1のクラスタをユーザが選択したことに応答して、モバイルデバイスは、オブジェクト1702~1706の重心が150前後のx座標と250前後のy座標とを有するフレームを検索することができる。

10

【0172】

[00229] 方法2010は、2012において、モバイルデバイスにおいて、ビデオストリーム中の選択されたオブジェクトの特定の幾何学的配置の指示を受信することを含み得る。非限定的な例として、図21を参照すると、モバイルデバイスは、オブジェクト1702~1704の重心が第1の領域($C1$)の重心に最も近いビデオストリームを表示するために、指示(たとえば、表示1802に示される第1の領域($C1$))においてユーザが画面をタッチすること)を受け取り得る。

20

【0173】

[00230] 2014において、ビデオストリームのフレームは、特定の幾何学的配置に基づいて検索され得る。たとえば、図21を参照すると、モバイルデバイスは、オブジェクト1702~1704の重心が第1の領域($C1$)に対応するエリアの中にある、図20のシーン中のビデオストリームのフレームを検索することができる。

【0174】

[00231] 2016において、検索されたフレームはモバイルデバイスの画面上に表示され得る。たとえば、図21を参照すると、モバイルデバイスは、オブジェクト1702~1704の重心が第1の領域($C1$)に対応するエリアの中にある、ビデオストリームのフレームを表示する(たとえば、再生する)ことができる。特定の実施形態では、モバイルデバイスは、順次的な順序でビデオストリームを再生することができる。

30

【0175】

[00232] 図23の方法2000、2010は、ユーザが、ビデオの特定の部分と関連付けられる構成(たとえば、重心)を有するクラスタを選択することによって、ビデオの特定の部分を検索する(たとえば、位置特定する)ことを可能にし得る。したがって、選択されたオブジェクト1702~1706の幾何学的パラメータに基づくクラスタへとフレームをインデクシングする(たとえば、本明細書で説明されるようなタスクT300によってメタデータを生成する)ことによって、モバイルデバイスのユーザは、ビデオを見進める(たとえば、再生する、早送りする、巻き戻すなど)必要なく、ビデオの所望の部分を容易に位置特定することができる。

40

【0176】

[00233] タスクT300は、決定された基準の位置に基づいて、位置空間を分割する複数の別々の領域のうちの1つを識別するメタデータを生成するように実施されてよく、複数の別々の領域は等しくないサイズの領域を含む。上で説明されたように、位置空間は、2次元または3次元における、少なくとも1つのデジタル画像(たとえば、ビデオストリームの少なくとも1つのフレームのピクセル座標空間)または物理空間(たとえば、シーン空間の地面)のピクセル座標空間であってよく、個々のフレームにおいて表される空間

50

を超えて延び得る。そのようなメタデータは、たとえば、基準の位置を含む位置空間の領域を示す（たとえば、1つまたは複数の量子化コードブックへの）1つまたは複数のインデックスとして実装され得る。

【0177】

[00234]分割方式は、位置空間を不均一に分割する（たとえば、不均一なサイズおよび/または形状の領域を含む）ように構成され得る。たとえば、各々の別々の領域がm個の観測されたデータ点を含むように（たとえば、発生の頻度に従った分割）、位置空間を分割する分割方式を生成するために、観測されたデータ点（たとえば、オブジェクトの位置）のセットが使用され得る。代替的に、または加えて、分割方式は、関心の高い空間のエリアが関心の低いエリアよりも多数の別々の領域へと分割されるように、位置空間を分割

10

【0178】

[00235]分割方式によって示される様々な領域は、位置空間のそれぞれのエリアと関連付けられる異なる度合いの関心に従って集中し得る。図19B~図19Dは、バスケットボールコート（たとえば、図19Aに示されるような）に適用されるような位置空間の不均一な分割のいくつかの例を示す。そのような分割方式は、（たとえば、図19Bおよび図19Dのような）特定のバスケットに近いより小さな領域への、または、（たとえば、図19Cのような）いずれかのバスケットに近いより小さな領域への、コート空間の分割を符号化することができる。事前の（たとえば、ストリームがキャプチャされる前の）位置空間の関心に基づく不均一な分割を示すことが可能であり得るので、方法M100はデジタル画像またはビデオストリームのキャプチャの間に実行され得る。

20

【0179】

[00236]同様の方式で、分割方式によって示される様々な領域は、位置空間のそれぞれのエリア内のマッピングされたデータ点（たとえば、オブジェクトの位置または複数のオブジェクトの重心）の発生の頻度の差に従って集中し得る。そのような場合、位置空間の不均一な分割は事後に（たとえば、キャプチャされたストリームへの訓練の結果として）取得され得る。等しい数の観測結果が各々の別々の領域において見出されるように、分割が実施され得る。

【0180】

[00237]一例では、分割方式は、位置空間におけるオブジェクトの位置（たとえば、観測されるオブジェクトの位置）の、k個のクラスタへのk平均クラスタ化を実行することによって得られる。たとえば、分割方式は、ビデオストリームの複数のフレームにおいて観測されるようなオブジェクトの位置（たとえば、選択されたオブジェクトの1つまたは複数の位置）の、k個のクラスタへのk平均クラスタ化を実行することによって得られ得る。そのようなクラスタ化は、位置空間をボロノイセルへと区分し、ここで、得られるk平均はセルの生成要素（たとえば、セントロイド）であり、そのような方式で（たとえば、クラスタ化を使用して）得られる分割方式は量子化コードブックとも呼ばれ得る。クラスタは、別のクラスタとは異なる数の観測されたオブジェクトの位置を有し得る。初期条件（すなわち、初期のk平均）を得るために、そのような分割方式生成タスクが、観測された位置に対する階層的なクラスタ化方法（たとえば、分岐型クラスタリングまたは凝集型クラスタリング）を実行するように実施され得る。不均一な分割を得るために使用され得る類似性の尺度の一例は、位置空間における観測されたオブジェクトの位置の間のユークリッド距離である。

30

40

【0181】

[00238]（たとえば、異常値の存在下での）分割方式生成タスクによって使用され得るk平均クラスタ化に対する1つの代替形態は、k-medoidクラスタ化である。別の代替形態は平均シフトクラスタ化であり、これは、分布のモードを異なる領域へと分離するために使用され得る。タスクT300は、複数の分割方式（たとえば、複数のコードブック）の中から選択することを含んでよく、これは、異なる長さであってよく、および/または異なるサイズの領域に対応してよい。たとえば、タスクT300は、第1のレベル

50

の分割方式が複数の別の領域の1つを示し、また、複数の第2のレベルの分割方式のうちの示される領域に対応する第2のレベルの分割方式を示すように、分割方式の階層を使用することができる。そのような階層は、複数の別の領域の1つを互いに異なるように再分割するために使用され得る。1つのそのような例では、第1のコードブックは、観測された位置の分布の異なるモードに対応する異なる量子化領域をインデクシングし、第2のコードブックは、これらのモード領域の1つまたは複数の各々を副領域へと分割するために使用されるので、タスクT300は、決定された位置に対する2つ以上のコードワードを含むメタデータを生成することができる。

【0182】

[00239]タスクT300は、たとえば、そのセントロイドが決定された基準の位置に最も近い領域を選択することによって、位置空間の対応する領域に、決定された基準の位置をマッピングするように実施され得る。インデックス選択のために使用される類似性の尺度（たとえば、位置空間におけるユークリッド距離）は、区分のために使用される類似性の尺度と同じであってよく、または異なっていてよい。分割方式が量子化コードブックによって表される場合、そのようなマッピングは、コードブック（たとえば、選択された領域を示すコードワード）へのインデックスを選択することを含み得る。

【0183】

[00240]図24を参照すると、矩形座標を使用してオブジェクトの位置に対するメタデータを生成するための方法M100の特定の適用形態が示されている。図24の実施形態は、フレームごとに複数のオブジェクトの決定された位置を符号化するために使用され得るコードブック2102を示す。決定された位置は、位置空間2106の中の各オブジェクトの位置を示し得る。コードブック2102は、位置空間2106の中の位置（たとえば、2132、2134、2136）を、等しくないサイズの領域を含む別々の領域への位置空間の分割2108の中の対応する領域（たとえば、2142、2144、2146）にマッピングする。

【0184】

[00241]コードブック2102は、複数のフレームまたは他の画像2104に対するメタデータを生成するために使用され得る。そのようなメタデータは、ビデオストリームに組み込まれてよく、および/または別々にストリーミングされてよく、および/または別のファイルに記憶されてよい。決定された位置は、フレーム中のオブジェクトのための矩形座標データ（たとえば、x座標およびy座標）を含み得る。特定の実施形態では、決定された位置はまた、フレーム中のオブジェクトに対する3次元座標データ（たとえば、3次元直交座標系のz座標）を含み得る。示される実施形態では、コードブック2102は、10000個のフレーム（または他の画像）にわたって3つのオブジェクトに対する決定された位置をクラスタ化することによって生成され得る。これらの決定された位置はまた、メタデータとして対応するフレームと関連付けられていることがある。特定の実施形態では、第1のフレームは、図20の第1の画像（たとえば、表示1710）に対応してよく、第2のフレームは、図20の第2の画像（たとえば、表示1730）に対応してよく、第3のフレームは、図20の第3の画像（たとえば、表示1750）に対応してよい。この例では、10000個のフレームに対する決定された位置はコードブック2102を得るために使用されるが、他の実施形態では、コードブック2102は、より少数の（または追加の）フレームに対する決定された位置に基づき得る。代替的に、または加えて、コードブックは、上で説明されたような（たとえば、図19B～図19Dに示されるように）関心に基づく分割方式、または、観測される決定された位置のセットをクラスタ化することによって得られない別の分割方式に基づき得る。加えて、決定された位置は、より少数の（または追加の）オブジェクトに対する位置の情報を含み得る。図24の以下の説明は、第3の画像に対応する決定された位置に基づく。しかしながら、同様の技法が他の決定された位置に適用可能であり得る。

【0185】

[00242]第3の画像における決定された位置は、第1の矩形座標データ2132と、第

10

20

30

40

50

2の矩形座標データ2134と、第3の矩形座標データ2136とを含み得る。第1の矩形座標データ2132は、図20の第3の画像における第1のオブジェクト1702のx座標の位置と、第3の画像における第1のオブジェクト1702のy座標の位置とを含み得る。第2の矩形座標データ2134は、第3の画像における第2のオブジェクト1704のx座標の位置と、第3の画像における第2のオブジェクト1704のy座標の位置とを含み得る。第3の矩形座標データ2136は、第3の画像における第3のオブジェクト1706のx座標の位置と、第3の画像における第3のオブジェクト1706のy座標の位置とを含み得る。

【0186】

[00243]別の特定の実施形態では、矩形座標データ2132~2136における各オブジェクト1702~1706のx座標およびy座標は、第3の画像におけるオブジェクト1702~1706の重心(C_{M3})に対するものであり得る。たとえば、重心(C_{M3})は原点として指定されてよく、各オブジェクト1702~1706の位置(たとえば、x座標およびy座標)は原点に対するものであってよい。そのような場合、図19Aに示されるようなバスケットボールコートまたは他の競技場の例のように、図19B~図19Dに示されるような不均一な分割方式に従って、選択されたオブジェクト(たとえば、選手)のそのような重心の位置を量子化するのが望ましいことがある。

【0187】

[00244]別の特定の実施形態では、特定のオブジェクトは原点として指定されてよく、他のオブジェクトの位置(たとえば、x座標およびy座標)は原点に対するものであってよい。特定のオブジェクト(「主要対象のオブジェクト」と呼ばれる)は、次の例、すなわち、ボール(たとえば、スポーツイベントのビデオにおける)、選択された相手(たとえば、スポーツイベントにおける)、犯罪被害者(たとえば、監視ビデオにおける)、花嫁(たとえば、結婚式のビデオにおける)のいずれをも含み得る。主要対象のオブジェクトは、時間とともに移動することがあり(たとえば、人)、または、空間中の固定された位置を有する物体であることがある(たとえば、スポーツのビデオにおけるネット、バスケット、または他のゴール)。非限定的な例として、第1のオブジェクト1702は原点として指定されてよく、第2のオブジェクト1704および第3のオブジェクト1706の位置は原点に対するものであってよい。

【0188】

[00245]プロセッサは、等しくないサイズの領域を含む別々の領域への位置空間の分割2108に従ってメタデータを生成するために、コードブック2102を使用して矩形座標データ2132~2136を符号化することができる。例示すると、プロセッサは、第1の符号化された位置データ2142へと第1の矩形座標データ2132を符号化する(たとえば、量子化する)ことができる。たとえば、プロセッサは、第1の矩形座標データ2132を対応する値(たとえば、コードブック2102のコードワード)にマッピングして、第1の矩形座標データ2132を値(たとえば、第1の符号化された位置データ2142)として符号化することができる。加えて、プロセッサは、第2の符号化された位置データ2144へと第2の矩形座標データ2134を符号化することができる。たとえば、プロセッサは、コードブック2102を使用して第2の矩形座標データ2134を対応する値にマッピングして、第2の矩形座標データ2134を値(たとえば、第2の符号化された位置データ2144)として符号化することができる。同様の方式で、プロセッサは、第3の符号化された位置データ2146へと第3の矩形座標データ2136を符号化することができる。たとえば、プロセッサは、コードブック2102を使用して第3の矩形座標データ2136を対応する値にマッピングして、第3の矩形座標データ2136を値(たとえば、第3の符号化された位置データ2146)として符号化することができる。

【0189】

[00246]第1の符号化された位置データ2142に基づいて、図20の第3の画像における第1のオブジェクト1702(たとえば、オブジェクト1)の位置は、格子2108

における対応する位置（たとえば、示される領域のセントロイド）として、生成されたメタデータにおいて表され得る。第2の符号化された位置データ2144に基づいて、第3の画像における第2のオブジェクト1704（たとえば、オブジェクト2）の位置は、格子2108における対応する位置として、生成されたメタデータにおいて表され得る。第3の符号化された位置データ2146に基づいて、第3の画像における第3のオブジェクト1706（たとえば、オブジェクト3）の位置は、格子2108における対応する位置として、生成されたメタデータにおいて表され得る。

【0190】

[00247]上で述べられたように、決定された位置は、複数の選択されたオブジェクトの幾何学的構成または「編成」の基準の位置であり得る。1つまたは複数の位置（または「基準」）コードブックを使用することに加えて、またはその代わりに、タスクT300は、1つまたは複数の編成（または「相対的」）コードブックを使用するように実施されてよく、この編成コードブックは、基準の位置（たとえば、オブジェクトの重心の位置）に対する空間中の選択されたオブジェクトの編成（たとえば、幾何学的構成、またはこの構成を表す形状）を符号化する。3つの選択されたオブジェクトの非限定的な例では、基準の位置として、2D空間における選択されたオブジェクトの重心の位置を記憶し、2D空間における各々の選択されたオブジェクトの、基準の位置に対する位置を、全体で8の自由度で示すメタデータを生成するのが望ましいことがある。これらの値は、（たとえば、8D空間における）単一のコードワードとして量子化され得る。代替的に、これらの値は、上で説明されたような1つまたは複数の位置コードブックを使用して（たとえば、2D空間における）基準の位置として、および1つまたは複数の編成コードブック（すなわち、幾何学的構成の候補のインデクシングされたセット）を使用して（たとえば、6D空間における）オブジェクト編成として、別々に量子化され得る。

【0191】

[00248]代替的に、編成コードブックまたは相対的コードブックへのインデックスは、テンプレートのセット（すなわち、幾何学的構成の候補）の1つを特定することができる。そのような場合、示されたテンプレートへの修正（たとえば、スケーリング、回転または他の方向付け、アスペクト比など）は別々に量子化され得る。図33Aは3つのテンプレートの例を示し、図33Bは図33Aの一番左のテンプレートへの修正の3つの例を示す。

【0192】

[00249]編成コードブックは、たとえば、複数のあり得る編成の間の発生の頻度の差、複数のあり得る編成の間の関心のレベル（たとえば、所定の編成に対する類似性）の差、特定の所定の事象（たとえば、得点事象）またはパターン（たとえば、フットボールチームの特定のスターティングフォーメーション）に対するあり得る編成の異なる関連度、空間的な基準から最も遠い選択されたオブジェクトの距離、選択されたオブジェクトの選択されたサブセットの互いからの距離、および/または、選択されたオブジェクトの任意の2つの間の最大の距離に従って、不均一に編成空間（すなわち、すべてのあり得る編成の空間のある部分）を分割するように構成され得る。

【0193】

[00250]コードワード（たとえば、符号化された位置データ2142～2146）は、固定幅コードワードまたは可変幅コードワードであり得る。特定の実施形態では、プロセッサは、オブジェクトの位置に基づいて（たとえば、オブジェクトが「高密度エリア」にあるかまたは「低密度エリア」にあるかに基づいて）、可変幅コードワードを使用することができる。高密度エリアは、オブジェクト1702～1706の少なくとも1つによって頻繁に専有される図20中の画像のエリアに対応し得る。メタデータ中のビットの数を減らすために、位置空間2106の中の高密度エリアに対応するコードワードは、低密度エリアに対応するコードワードよりも少数のビット（すなわち、より短いコードワード）を有し得る（すなわち、占有される頻度の低い格子位置がより長いコードワードを有する）。代替的に、または加えて、不均一な分割方式2108は、高密度領域においてより多

くの副領域（すなわち、高解像度をもたらすためのより多数の別個のコードワード）と、低密度領域においてより少数の副領域（すなわち、低解像度をもたらすためのより少数の別個のコードワード）とを有する、階層的な方式であり得る。

【0194】

[00251]オブジェクトが高密度エリアにある場合、オブジェクトに対するコードワードは、オブジェクトのより「正確な」位置を生成するための相対的に小さな領域を表し得る。たとえば、高密度エリアのオブジェクト1702～1706の位置をトラッキングして、フレームごとにより正確なトラッキングを得るために、より多くの数のコードワードが使用され得る。オブジェクトが低密度エリアにある場合、オブジェクトに対するコードワードは、比較的大きな領域を表し得る。特定の実施形態では、高密度エリアおよび低密度エリアは、履歴データの分析に基づいて（たとえば、画面のどのエリアがオブジェクトにより頻繁に占有された履歴を有するかを分析して）決定され得る。別の特定の実施形態では、高密度エリアおよび低密度エリアは事前に決定されていてよい。

10

【0195】

[00252]編成コードブックへのコードワードは、 n 個の選択されたオブジェクトの n 個の位置への一意なマッピングを区別することができる（たとえば、各々のそのようなマッピングに対する一意なコードワードを含み得る）。たとえば、図34Aは、一意なコードワードとして各々符号化され得る（すなわち、異なるそれぞれの幾何学的構成の候補によって表され得る）3つの選択されたオブジェクトの4つの異なる編成の例を示す。これらの例では、選択されたオブジェクトの各位置は、そのオブジェクトを一意に識別するアイコン（たとえば、円、四角形、または三角形）によって示される。代替的に、編成コードブックへのコードワードは、いくつかの異なる一意なマッピングを符号化することができる。たとえば、図34Bは、図34Aの編成に対応する一意ではないマッピングの例を示し、ここで任意の選択されたオブジェクトの位置が x によって示される。この場合、図34Cは、図34Bの一番左に示される単一の一意ではないマッピング（すなわち、幾何学的構成の候補）によって示され得る、選択されたオブジェクトの6個の一意な編成を示す。

20

【0196】

[00253]コードブック2102は、固定されたコードブックまたは適応的なコードブックであり得る。適応的なコードブックは、オブジェクトの位置を表す履歴データに基づいてコードワードの長さを調整することができる（たとえば、可変幅のコードワードを生成することができる）。たとえば、適応コードブックは、画面の頻繁に占有されるエリアの中のオブジェクトに対してより短いコードワードを生成することができ、履歴データに基づいて画面のより稀に占有されるエリアの中のオブジェクトに対してより長いコードワードを生成することができる。対照的に、固定されたコードブックは、履歴データに基づいて変化しなくてよく、固定幅コードワードまたは可変幅コードワードを使用することができる。

30

【0197】

[00254]それぞれの基準の位置に対する複数のあり得る編成の間の発生の頻度または関心のレベルの差を利用する複数の編成コードブックを使用して、編成データを符号化するのが望ましいことがある。たとえば、可変サイズの編成コードブックのセットの中の1つが、基準の位置のコードワードに応じて選択され得る（たとえば、基準の位置がコートのある中心にあるときにはより小さいコードブック v_s 基準の位置がゴールの近くであるときはより大きいコードブック）。

40

【0198】

[00255]図24の実施形態は、コードブック2102を使用してタスクT300によって生成されるメタデータに基づいて、ビデオの特定の部分のような、特定の画像をユーザが検索する（たとえば、位置特定する）ことを可能にし得る。たとえば、位置空間2106は、ユーザが異なるオブジェクトに対する異なる位置（たとえば、パラメータ）を選択することを可能にする、ユーザ対話型の格子として画面に表示され得る。この選択に基づ

50

いて、モバイルデバイスのユーザは、ビデオを見進める（たとえば、再生する、早送りする、巻き戻すなど）必要なく、同様のオブジェクトの位置構成を有するビデオの所望の部分（たとえば、フレーム）を容易に位置特定することができる。たとえば、方法M100は、生成されたメタデータを、分割方式において領域の1つを指定する探索要求と比較した結果に基づいて、1つまたは複数の画像を（たとえば、記憶装置から）検索するタスクを含むように実施され得る。特定の実施形態では、位置空間2106のそのような表示は、図21中のクラスタの画面1802と実質的に同様の方式で使用され得る。加えて、生成されたメタデータ2142～2146は、圧縮されてエンコーダからデコーダに送信され得る。メタデータ2142～2146を圧縮してデコーダに送信することで、デコーダは、比較的少量のデータ（たとえば、量子化された値）を使用して、トラッキングされたオブジェクトの位置を決定することが可能になり得る。たとえば、デコーダは、コードブック2102の中の量子化された値に基づいて、オブジェクト1702～1706と関連付けられるビデオを復号するように、レンダリング動作を実行することができる。

10

【0199】

[00256]一例では、検索されるフレームは、フットボールゲームまたは他のスポーツイベントにおいて、特定のプレーもしくはプレーのセット、または特定のスターティングフォーメーションの間にだけ発生する、選択されたプレーヤーの間の選択された編成を含む。別の例では、検索されるフレームは、ラップトラック上の近接している特定のレーシングカーの間の選択された編成を含む。そのような選択的な検索によって促されるビデオ分析はまた、特定の編成の発生とのある事象（事前に特定されていないことがある）の相関を検出するために使用され得る。たとえば、潜在的な適用例は、集会（たとえば、立法会議、街頭デモ、社会的な集まり）の中の人々をトラッキングし、（たとえば、意思決定者、トラブルメーカー、触媒の働きをする人物を識別するために）選択された人々の周囲を指定されたイベントと関連付けることを含む。本明細書で説明されるような方法M100の実装形態はまた、（たとえば、研究所または動物園の）捕獲されている動物または野生の動物の間の社会的な交流の分析に適用され得る。

20

【0200】

[00257]図25を参照すると、極座標を使用してオブジェクトの位置に対するメタデータを生成するための方法M100の別の特定の適用形態が示されている。図25の実施形態は、フレームごとに複数のオブジェクトの決定された位置を符号化するために使用され得るコードブック2202を示す。決定された位置は、極位置空間2206の中の各オブジェクトの位置を示し得る。コードブック2202は、位置空間2206の中の位置（たとえば、2232、2234、2236）を、等しくないサイズの領域を含む別々の領域への位置空間2206の分割2208の中の対応する領域（たとえば、2242、2244、2246）にマッピングすることができる。

30

【0201】

[00258]コードブック2202は、複数のフレームまたは他の画像2204に対するメタデータを生成するために使用され得る。そのようなメタデータは、ビデオストリームに組み込まれてよく、および/または別々にストリーミングされてよく、および/または別のファイルに記憶されてよい。決定された位置は、フレーム中のオブジェクトのための極座標データを含み得る。特定の実施形態では、決定された位置はまた、フレーム中のオブジェクトに対する球面または円柱座標データ（たとえば、3D座標データ）を含み得る。示される実施形態では、コードブック2102は、10000個のフレーム（または他の画像）にわたって3つのオブジェクトに対する決定された位置をクラスタ化することによって生成され得る。これらの決定された位置はまた、メタデータとして対応するフレームと関連付けられていることがある。特定の実施形態では、第1のフレームは、図20の第1の画像（たとえば、表示1710）に対応してよく、第2のフレームは、図20の第2の画像（たとえば、表示1730）に対応してよく、第3のフレームは、図20の第3の画像（たとえば、表示1750）に対応してよい。この例では、10000個のフレームに対する決定された位置はコードブック2202を得るために使用されるが、他の実施形

40

50

態では、コードブック 2 2 0 2 は、より少数の（または追加の）フレームに対する決定された位置に基づき得る。代替的に、または加えて、コードブックは、上で説明されたような（たとえば、図 1 9 B ~ 図 1 9 D に示されるように）関心に基づく分割方式、または、観測される決定された位置のセットをクラスタ化することによって得られない別の分割方式に基づき得る。加えて、決定された位置は、より少数の（または追加の）オブジェクトに対する位置の情報を含み得る。図 2 5 の以下の説明は、第 3 の画像に対応する決定された位置に基づく。しかしながら、同様の技法が他の決定された位置に適用可能であり得る。

【 0 2 0 2 】

[00259] 第 3 の画像における決定された位置は、極方向データ 2 2 3 0 と、第 1 の極座標データ 2 2 3 2 と、第 2 の極座標データ 2 2 3 4 と、第 3 の極座標データ 2 2 3 6 とを含み得る。極方向データ 2 2 3 0 は、重心 (C_{M3})（たとえば、図 2 0 の第 3 の画像の重心 (C_{M3}) の x 座標 (C_x)、および第 3 の画像の重心 (C_{M3}) の y 座標 (C_y) の位置を示し得る。極方向データ 2 2 3 0 はまた、中心として重心 (C_{M3}) を有する円の半径 (R) を示し得る。スケール係数として実装され得る半径 (R) は、各々のトラッキングされたオブジェクトを包含するのに十分大きいことがある。説明のための例が、極位置空間 2 2 0 6 に関して図示されている。一例では、タスク T 3 0 0 は、等しくないサイズの領域を含む複数の領域の 1 つを示すメタデータとして方向の位置（たとえば、重心の位置）を符号化するように実施されてよく、それらの領域は、位置空間 2 2 0 6（たとえば、図 2 4 に示されるような矩形座標空間）とは異なり得る、その中で方向の位置が示される対応する位置空間を分割する。

【 0 2 0 3 】

[00260] 第 1 の極座標データ 2 2 3 2 は、第 1 のオブジェクト 1 7 0 2（オブジェクト 1）の第 1 の極座標を示し得る。第 1 の極座標は、重心 (C_{M3}) からの第 1 のオブジェクト 1 7 0 2 の第 1 の距離 (a) と、0 度の基準から測定される第 1 のオブジェクト 1 7 0 2 の第 1 の角度 (α)（たとえば、水平線からの角度方向の変位）とを含み得る。第 2 の極座標データ 2 2 3 4 は、第 2 のオブジェクト 1 7 0 4（オブジェクト 2）の第 2 の極座標を示し得る。第 2 の極座標は、重心 (C_{M3}) からの第 2 のオブジェクト 1 7 0 4 の第 2 の距離 (b) と、水平線から測定される第 2 のオブジェクト 1 7 0 4 の第 2 の角度 (β) とを含み得る。第 3 の極座標データ 2 2 3 6 は、第 3 のオブジェクト 1 7 0 6（オブジェクト 3）の第 3 の極座標を示し得る。第 3 の極座標は、重心 (C_{M3}) からの第 3 のオブジェクト 1 7 0 6 の第 3 の距離 (c) と、水平線から測定される第 1 のオブジェクト 1 7 0 2 の第 3 の角度 (γ) とを含み得る。

【 0 2 0 4 】

[00261] 特定の実施形態では、極方向データ 2 2 3 0 の半径 (R) は、重心 (C_{M3}) から最も離れているオブジェクトの距離に基づいて正規化され得る。たとえば、半径 (R) は、第 1 の距離 (a) が第 2 の距離 (b) および第 3 の距離 (c) より大きい場合、第 1 の距離 (a) に等しくてよい。

【 0 2 0 5 】

[00262] プロセッサは、等しくないサイズの領域を含む領域への位置空間の分割 2 1 0 8 に従ってメタデータを生成するために、コードブック 2 2 0 2 を使用して極座標データ 2 2 3 2 ~ 2 2 3 6 を符号化することができる。例示すると、プロセッサは、第 1 の符号化された位置データ 2 2 4 2 へと第 1 の極座標データ 2 2 3 2 を符号化する（たとえば、量子化する）ことができる。たとえば、プロセッサは、第 1 の極座標データ 2 2 3 2（極方向データ 2 2 3 0 に基づく）を対応する値（たとえば、コードブック 2 2 0 2 のコードワード）にマッピングして、第 1 の極座標データ 2 2 3 2 を値（たとえば、第 1 の符号化された位置データ 2 2 4 2）として符号化することができる。加えて、プロセッサは、第 2 の符号化された位置データ 2 2 4 4 へと第 2 の極座標データ 2 2 3 4 を符号化することができる。たとえば、プロセッサは、コードブック 2 2 0 2 を使用して第 2 の極座標データ 2 2 3 4（極方向データ 2 2 3 0 に基づく）を対応する値にマッピングして、第 2 の極

座標データ 2 2 3 4 を値（たとえば、第 2 の符号化された位置データ 2 2 4 4 ）として符号化することができる。同様の方式で、プロセッサは、第 3 の符号化された位置データ 2 2 4 6 へと第 3 の極座標データ 2 2 3 6 を符号化することができる。たとえば、プロセッサは、コードブック 2 2 0 2 を使用して第 3 の極座標データ 2 2 3 6 を対応する値にマッピングして、第 3 の極座標データ 2 2 3 6 を値（たとえば、第 3 の符号化された位置データ 2 2 4 6 ）として符号化することができる。

【 0 2 0 6 】

[00263] 第 1 の符号化された位置データ 2 2 4 2 に基づいて、図 2 0 の第 3 の画像における第 1 のオブジェクト 1 7 0 2（たとえば、オブジェクト 1）の位置は、極格子 2 2 0 8 における対応する位置（たとえば、示される領域のセントロイド）として、生成されたメタデータにおいて表され得る。第 2 の符号化された位置データ 2 2 4 4 に基づいて、第 3 の画像における第 2 のオブジェクト 1 7 0 4（たとえば、オブジェクト 2）の位置は、極格子 2 2 0 8 における対応する位置として、生成されたメタデータにおいて表され得る。第 3 の符号化された位置データ 2 2 4 6 に基づいて、第 3 の画像における第 3 のオブジェクト 1 7 0 6（たとえば、オブジェクト 3）の位置は、極格子 2 2 0 8 における対応する位置として、生成されたメタデータにおいて表され得る。たとえば、極格子 2 2 0 8 は、格子の原点からの半径方向の距離の範囲に基づいて（たとえば、重心（ C_{M3} ）からの半径方向の距離の範囲に基づいて）、および 0 度の基準からの角度方向の変位の範囲に基づいて、等しくないサイズの領域を含む複数の領域へと分割され得る。

【 0 2 0 7 】

[00264] 図 2 5 の実施形態は、コードブック 2 2 0 2 を使用してタスク T 3 0 0 によって生成されるメタデータに基づいて、特定の画像（たとえば、ビデオのフレーム）をユーザが検索する（たとえば、位置特定する）ことを可能にし得る。たとえば、極位置空間 2 2 0 6 は、ユーザが異なるオブジェクトに対する異なる位置（たとえば、パラメータ）を選択することを可能にする、ユーザ対話型の格子として画面に表示され得る。この選択に基づいて、モバイルデバイスのユーザは、ビデオを見進める（たとえば、再生数、早送りする、巻き戻すなど）必要なく、同様の異議の位置構成を有するビデオの所望の部分（たとえば、フレーム）を容易に位置特定することができる。たとえば、方法 M 1 0 0 は、生成されたメタデータを、分割方式において領域の 1 つを指定する探索要求と比較した結果に基づいて、1 つまたは複数の画像を（たとえば、記憶装置から）検索するタスクを含むように実施され得る。特定の実施形態では、極位置空間 2 2 0 6 のそのような表示は、図 2 1 中のクラスタの画面 1 8 0 2 と実質的に同様の方式で使用され得る。加えて、生成されたメタデータ 2 2 4 2 ~ 2 2 4 6 は、圧縮されてエンコーダからデコーダに送信され得る。生成されたメタデータ 2 2 4 2 ~ 2 2 4 6 を圧縮してデコーダに送信することで、デコーダは、比較的少量のデータ（たとえば、量子化された値）を使用して、トラッキングされたオブジェクトの位置を決定することが可能になり得る。たとえば、デコーダは、コードブック 2 2 0 2 の中の量子化された値に基づいて、オブジェクト 1 7 0 2 ~ 1 7 0 6 と関連付けられるビデオを復号するように、レンダリング動作を実行することができる。

【 0 2 0 8 】

[00265] コードワード（たとえば、符号化された位置データ 2 2 4 2 ~ 2 2 4 6）は、固定幅コードワードまたは可変幅コードワードであり得る。特定の実施形態では、プロセッサは、オブジェクトの位置に基づいて（たとえば、オブジェクトが「高密度エリア」にあるかまたは「低密度エリア」にあるかに基づいて）、可変幅コードワードを使用することができる。高密度エリアは、オブジェクト 1 7 0 2 ~ 1 7 0 6 の少なくとも 1 つによって頻繁に専有される図 2 0 中の画像のエリアに対応し得る。メタデータ中のビットの数を減らすために、極位置空間 2 2 0 6 の中の高密度エリアに対応するコードワードは、低密度エリアに対応するコードワードよりも少数のビット（すなわち、より短いコードワード）を有し得る（すなわち、占有される頻度の低い格子位置がより長いコードワードを有する）。代替的に、または加えて、不均一な分割方式 2 2 0 8 は、高密度領域においてより

多くの副領域（すなわち、高解像度をもたらすためのより多数の別個のコードワード）と、低密度領域においてより少数の副領域（すなわち、低解像度をもたらすためのより少数の別個のコードワード）とを有する、階層的な方式であり得る。

【0209】

[00266]オブジェクトが高密度エリアにある場合、オブジェクトに対するコードワードは、オブジェクトのより「正確な」位置を生成するための相対的に小さな領域を表し得る。たとえば、高密度エリアのオブジェクト1702～1706の位置をトラッキングして、フレームごとにより正確なトラッキングを得るために、より多くの数のコードワードが使用され得る。オブジェクトが低密度エリアにある場合、オブジェクトに対するコードワードは、比較的大きな領域を表し得る。特定の実施形態では、高密度エリアおよび低密度エリアは、履歴データの分析に基づいて（たとえば、画面のどのエリアがオブジェクトにより頻繁に占有された履歴を有するかを分析して）決定され得る。別の特定の実施形態では、高密度エリアおよび低密度エリアは事前に決定されていてよい。

10

【0210】

[00267]コードブック2202は、固定されたコードブックまたは適応的なコードブックであり得る。適応的なコードブックは、オブジェクトの位置を表す履歴データに基づいてコードワードの長さを調整することができる（たとえば、可変幅のコードワードを生成することができる）。たとえば、適応コードブックは、画面の頻繁に占有されるエリアの中のオブジェクトに対してより短いコードワードを生成することができ、履歴データに基づいて画面のより稀に占有されるエリアの中のオブジェクトに対してより長いコードワードを生成することができる。対照的に、固定されたコードブックは、履歴データに基づいて変化しなくてよく、固定幅コードワードまたは可変幅コードワードを使用することができる。

20

【0211】

[00268]図26を参照すると、オブジェクトトラッキングを使用したビデオ処理方法2300の特定の実施形態を示すフローチャートが示される。代替的な実施形態では、方法2300は、図1の電子デバイス102、図2Bのプロセッサ264によって実装されるオブジェクトトラッキングおよび検出モジュール204、プロセッサによって実装される図8のオブジェクトトラッキングおよび検出モジュール804、プロセッサによって実装される図9の平滑化モジュール961、またはこれらの任意の組合せを使用して実行され得る。

30

【0212】

[00269]方法2300は、2302において、モバイルデバイスにおいて、特定のフレームにおける第1のオブジェクトの第1の位置と、特定のフレームにおける第2のオブジェクトの第2の位置とをトラッキングすることを含み得る。たとえば、図20を参照すると、第1のオブジェクト1702の第1の位置は第3のフレームにおいてトラッキングされてよく、第2のオブジェクト1704の第2の位置は第3のフレームにおいてトラッキングされてよい。加えて、第3のオブジェクト1706の第3の位置は第3のフレームにおいてトラッキングされ得る。

【0213】

40

[00270]2304において、第1の位置と関連付けられる第1の座標データが生成されてよく、第2の位置と関連付けられる座標データが生成されてよい。たとえば、図24を参照すると、プロセッサは、第3のフレーム中の第1のオブジェクト1702の位置と関連付けられる第1の矩形座標データ2132と、第3のフレーム中の第2のオブジェクト1704の位置と関連付けられる第2の矩形座標データ2134とを生成することができる。加えて、プロセッサは、第3のフレーム中の第3のオブジェクト1706の位置と関連付けられる第3の矩形座標データ2136を生成することができる。

【0214】

[00271]別の例として、図25を参照すると、プロセッサは、第3のフレーム中の第1のオブジェクト1702の位置と関連付けられる第1の極座標データ2232と、第3の

50

フレーム中の第2のオブジェクト1704の位置と関連付けられる第2の極座標データ2234とを生成することができる。加えて、プロセッサは、第3のフレーム中の第3のオブジェクト1706の位置と関連付けられる第3の極座標データ2236を生成することができる。

【0215】

[00272] 2306において、第1の座標データおよび第2の座標データは、第1の符号化された位置データおよび第2の符号化された位置データとしてメタデータを生成するために、コードブックを使用して符号化され得る。たとえば、図24を参照すると、プロセッサは、第1の符号化された位置データ2142（たとえば、コードブック2102のコードワード）を生成するために、コードブック2102を使用して第1の矩形座標データ2132を符号化することができ、プロセッサは、第2の符号化された位置データ2144を生成するために、コードブック2102を使用して第2の矩形座標データ2134を符号化することができる。加えて、プロセッサは、第3の符号化された位置データ2146を生成するために、コードブック2102を使用して第3の矩形座標データ2132を符号化することができる。

10

【0216】

[00273] 別の例として、図25を参照すると、プロセッサは、第1の符号化された位置データ2242（たとえば、コードブック2202のコードワード）としてメタデータを生成するために、コードブック2202を使用して第1の極座標データ2232を符号化することができ、プロセッサは、第2の符号化された位置データ2244としてメタデータを生成するために、コードブック2202を使用して第2の極座標データ2234を符号化することができる。加えて、プロセッサは、第3の符号化された位置データ2246としてメタデータを生成するために、コードブック2202を使用して第3の極座標データ2232を符号化することができる。

20

【0217】

[00274] 2308において、メタデータ（たとえば、第1の符号化された位置データおよび第2の符号化された位置データ）は、入力パラメータに基づく特定のフレームの検索を可能にするために記憶され得る。たとえば、図24を参照すると、プロセッサは、第1の符号化された位置データ2142と第2の符号化された位置データ2144とをメモリに記憶することができる。モバイルデバイスのユーザは、パラメータ（たとえば、第1のオブジェクト1702および第2のオブジェクト1704の所望の位置）を入力することができる。入力パラメータに基づいて、モバイルデバイスは、第1の位置および第2の位置に対応する入力パラメータに回答して、再生のために第3のフレームを検索することができる。

30

【0218】

[00275] 別の例として、図25を参照すると、プロセッサは、第1の符号化された位置データ2242と第2の符号化された位置データ2244とをメモリに記憶することができる。モバイルデバイスのユーザは、パラメータ（たとえば、第1のオブジェクト1702および第2のオブジェクト1704の所望の位置）を入力することができる。入力パラメータに基づいて、モバイルデバイスは、第1の位置および第2の位置に対応する入力パラメータに回答して、再生のために第3のフレームを検索することができる。

40

【0219】

[00276] 特定の実施形態では、方法2300はまた、複数のフレームにわたってシーンの第1のエリアの第1の密度をトラッキングすることを含み得る。第1の密度は、第1のオブジェクト1702または第2のオブジェクト1704が第1のエリアの中にある回数に基づき得る。方法2300はまた、複数のフレームにわたってシーンの第2のエリアの第2の密度をトラッキングすることを含み得る。第2の密度は、第1のオブジェクト1702または第2のオブジェクト1704が第2のエリアの中にある回数に基づき得る。方法2300はさらに、第1の密度を第2の密度と比較することと、第1の密度が第2の密度よりも大きい場合、コードブック2102中の第1の量のメタデータを第1のエリアに

50

割り振り、コードブック 2 1 0 2 中の第 2 の量のメタデータを第 2 のエリアに割り振ることを含み得る。第 1 の量のメタデータは、第 2 の量のメタデータよりも大きいことがある。たとえば、コードブック 2 1 0 2 中の多数のコードワードは第 1 のエリア（たとえば、高解像度エリア）に対応してよく、少数のコードワードはより密度の低いエリアに対するコードワードの数を減らすための第 2 のエリア（たとえば、より低解像度のエリア）に対応してよい。

【 0 2 2 0 】

[00277] 図 2 6 の方法 2 3 0 0 は、コードブック 2 1 0 2 を使用して符号化されたビデオの特定の部分（フレーム）をユーザが検索する（たとえば、位置特定する）ことを可能にし得る。たとえば、格子 2 1 0 6 は、ユーザが異なるオブジェクトに対する異なる位置（たとえば、パラメータ）を選択することを可能にする、画面上に表示されるユーザ対話型の格子であり得る。この選択に基づいて、モバイルデバイスのユーザは、ビデオを見進める（たとえば、再生する、早送りする、巻き戻すなど）必要なく、同様の異議の位置構成を有するビデオの所望の部分の位置特定することができる。

10

【 0 2 2 1 】

[00278] 図 2 7 を参照すると、オブジェクトトラッキングを使用したビデオ処理方法 2 6 0 0 の特定の実施形態を示すフローチャートが示される。例示的な実施形態では、方法 2 6 0 0 は、ロボットを使用して実行され得る。

【 0 2 2 2 】

[00279] 2 6 0 2 において、方法 2 6 0 0 は、ロボットにおいて、複数のオブジェクトの間の幾何学的配置をトラッキングすることを含み得る。たとえば、ユーザは、トラッキングされるべき特定のオブジェクトをロボットに伝えることができる。特定の実施形態では、ロボットはディスプレイインターフェースを含んでよく、ユーザはディスプレイインターフェースとの対話を介してロボットと交流することができる。別の例として、ロボットは、環境因子に基づいてどのオブジェクトがトラッキングされるべきかを、独立に決定することができる。たとえば、ロボットは、動きを検出するためのセンサを含み得る。検出された動きに基づいて、ロボットは、動きと関連付けられるオブジェクトをトラッキングすることを選択し得る。

20

【 0 2 2 3 】

[00280] 2 6 0 4 において、処理機能が幾何学的配置に基づいて実行され得る。たとえば、ロボットは、幾何学的配置に基づいてオブジェクト間の関係を決定することができる。別の例として、ロボットは、オブジェクトの幾何学的配置に基づいて設定（たとえば、シーン）を決定することができる。

30

【 0 2 2 4 】

[00281] 図 2 8 を参照すると、オブジェクトトラッキングを使用したビデオ処理を示す特定の実施形態が示される。図 2 8 に示される実施形態に関して説明されるビデオ処理技法は、図 1 の電子デバイス 1 0 2、図 2 B のプロセッサ 2 6 4 によって実装されるオブジェクトトラッキングおよび検出モジュール 2 0 4、プロセッサによって実装される図 8 のオブジェクトトラッキングおよび検出モジュール 8 0 4、プロセッサによって実装される図 9 の平滑化モジュール 9 6 1、またはこれらの任意の組合せによって実行され得る。

40

【 0 2 2 5 】

[00282] 図 2 8 は、第 1 の時刻においてキャプチャされるようなシーンを示す、画像 2 7 0 0 を示す。たとえば、画像 2 7 0 0 は、モバイルデバイスの画面に表示されるべきビデオストリームのフレームに対応し得る。モバイルデバイスは、ストリームを（たとえば、モバイルデバイスのカメラによって）キャプチャし、または別のデバイスからストリームを受信するように構成され得る。フレーム 2 4 0 0 に示されるシーンは、第 1 のオブジェクト 2 7 0 2 と、第 2 のオブジェクト 2 7 0 4 と、第 3 のオブジェクト 2 7 0 6 と、星と、木とを含む。示される実施形態では、第 1 のオブジェクト 2 7 0 2 は第 1 の人物に対応してよく、第 2 のオブジェクト 2 7 0 4 は第 2 の人物に対応してよく、第 3 のオブジェクト 2 7 0 6 は第 3 の人物に対応してよい。別の実施形態では、オブジェクト 2 7 0 2 ~

50

2706の少なくとも1つは、モバイルデバイスを介して制御されるロボットに対応し得る。

【0226】

[00283]図28はまた、モバイルデバイスの画面上のフレーム2700の表示2710を示す。モバイルデバイスの画面（たとえば、タッチスクリーン）は、ビューファインダに対応し得る。表示2710は、ユーザが（たとえば、タスクT200による）トラッキングのためにオブジェクトを選択することを可能にする、ユーザ選択機構2711を含み得る。たとえば、ユーザは、画面2710上のユーザ選択機構2711をタッチし、第1のオブジェクト2702のトラッキングを可能にするために第1のオブジェクト2702の周囲で第1の焦点リング2712をドラッグすることができる。第1の焦点リング2712のそのような配置に応答して、モバイルデバイスは、トラッキングのために第1のオブジェクト2712を選択することができる。同様の方式で、ユーザはまた、画面上のユーザ選択機構2711をタッチし、第2のオブジェクト2704および第3のオブジェクト2706の周囲で第2の焦点リング2714および第3の焦点リング2716をドラッグして、それぞれ、第2のオブジェクト2704および第3のオブジェクト2706のトラッキングを可能にすることができる。第2の焦点リング2714および第3の焦点リング2716のそのような配置に応答して、モバイルデバイスは、トラッキングのために、それぞれ第2のオブジェクト2704と第3のオブジェクト2706とを選択することができる。

10

【0227】

20

[00284]図28はまた、第2の時刻においてキャプチャされるようなシーンを示す、画像（たとえば、フレーム）2720を示す。画像2720において、各オブジェクト2702～2706の位置は、第1の時刻におけるシーンを示す画像2700に示される対応する位置に対して変化している。たとえば、第1のオブジェクト2702は星に近づくように移動しており、第2のオブジェクト2704は木に近づくように移動しており、第3のオブジェクト2706は画像の下部に近づくように移動している。

【0228】

[00285]図24はまた、モバイルデバイスの画面上の画像2720の表示2730を示す。表示2730は、第1のウィンドウ2732、第2のウィンドウ2734、第3のウィンドウ2736、および第4のウィンドウ2738へと（たとえば、表示タスクT50によって）区分され得る。ユーザ選択機構2711による3つのオブジェクト2702～2706の選択に応答して、たとえば、モバイルデバイスは、3つの対応するウィンドウ2732～2736および第4のウィンドウ2738へと画面2730を区分することができる。

30

【0229】

[00286]第1のウィンドウ2732は、第1のオブジェクト2702を含む画像2720の第1の部分を表示することができる。画像2720の第1の部分は、画像2720の中の第1のオブジェクト2702の位置に基づいて決定され得る。たとえば、モバイルデバイスは第1のオブジェクト2702の位置をトラッキングすることができ、第1のオブジェクト2702の位置に注目しているビデオストリームが第1のウィンドウ2732に表示され得る。第2のウィンドウ2734は、第2のオブジェクト2704を含む画像2720の第2の部分を表示することができる。画像2720の第2の部分は、画像2720の中の第2のオブジェクト2704の位置に基づいて決定され得る。たとえば、モバイルデバイスは第2のオブジェクト2704の位置をトラッキングすることができ、第2のオブジェクト2704の位置に注目しているビデオストリームが第2のウィンドウ2734に表示され得る。第3のウィンドウ2736は、第3のオブジェクト2706を含む画像2720の第3の部分を表示することができる。画像2720の第3の部分は、画像2720の中の第3のオブジェクト2706の位置に基づいて決定され得る。たとえば、モバイルデバイスは第3のオブジェクト2706の位置をトラッキングすることができ、第3のオブジェクト2706の位置に注目しているビデオストリームが第3のウィンドウ2

40

50

7 3 6 に表示され得る。

【 0 2 3 0 】

[00287]第4のウィンドウ2 7 3 8は、各オブジェクト2 7 0 2 ~ 2 7 0 6の間の幾何学的配置（たとえば、重心（ C_{M2} ））の指示を表示することができる。たとえば、第4のウィンドウ2 7 3 8は、画像2 7 2 0中のエリアに対応する複数のセクション（グレースケールの線として表されている）を含む格子を含み得る。たとえば、格子の左下のセクションは、画像2 7 2 0の左下のセクションに対応し得る。図2 1および図2 2に関して説明されるような同様の方式で、モバイルデバイスは、画像2 7 2 0中のオブジェクトの重心（ C_{M2} ）を決定することができる。重心（ C_{M2} ）は、格子の対応するセクションに表示され得る。

10

【 0 2 3 1 】

[00288]図2 8はまた、（たとえば、第2の時刻の後の）第3の時刻にキャプチャされるようなシーンを示す、画像2 7 4 0を示す。画像2 7 4 0において、各オブジェクト2 7 0 2 ~ 2 7 0 6の位置は、第2の時刻におけるシーンを示す画像2 7 2 0中の対応する位置に対して変化している。たとえば、第1のオブジェクト2 7 0 2は木に近づくように移動しており、第2のオブジェクト2 7 0 4はシーン2 7 4 0の下部に近づくように移動しており、第3のオブジェクト2 7 0 6は星に近づくように移動している。

【 0 2 3 2 】

[00289]図2 8はまた、モバイルデバイスによって生成されるウィンドウ2 7 3 2 ~ 2 7 3 6を使用した、モバイルデバイスの画面上の画像2 7 4 0の表示2 7 5 0を示す。たとえば、第1のウィンドウ2 7 3 2は、第1のオブジェクト2 7 0 2を含む画像2 7 4 0の第1の部分を表示し、第2のウィンドウ2 7 3 4は、第2のオブジェクト2 7 0 4を含む画像2 7 4 0の第2の部分を表示し、第3のウィンドウ2 7 3 6は、第3のオブジェクト2 7 0 6を含む画像2 7 4 0の第3の部分を表示する。

20

【 0 2 3 3 】

[00290]第4のウィンドウ2 7 3 8は、第3の時刻における画像2 7 4 0を反映するように、指示（たとえば、重心（ C_{M2} ））を更新することができる。モバイルデバイスは、フレームごとに、または非連続的な時間間隔で、幾何学的配置をトラッキングし、更新された幾何学的配置の指示を第4のウィンドウ2 7 3 8に表示することができる。たとえば、第3のフレームにおいて、第4のウィンドウ2 7 3 8は、第3の時刻における画像2 7 4 0中のオブジェクト2 7 0 2 ~ 2 7 0 6の重心（ C_{M3} ）を表示することができる。第4のウィンドウ2 7 3 8はまた、第2のフレームと第3のフレームとの間で（たとえば、タスクT 2 0 0によって示されるように）オブジェクトの重心をトラッキングする、経路（たとえば、点線）を表示することができる。

30

【 0 2 3 4 】

[00291]図2 8に示される実施形態は、モバイルデバイスが、フレーム2 7 0 0、2 7 2 0、2 7 4 0中のオブジェクト2 7 0 2 ~ 2 7 0 6をトラッキングすることを可能にでき、オブジェクト2 7 0 2 ~ 2 7 0 6に対応する個別化された（たとえば、専心的な）ビデオストリームを別のウィンドウ2 7 3 2 ~ 2 7 3 6に表示することができる。実施形態はまた、ユーザが第4のウィンドウ2 7 3 8中のオブジェクト2 7 0 2 ~ 2 7 0 6の幾何学的配置を見ることを可能にする。モバイルデバイスは、単一のカメラを使用して、または複数のカメラを使用して、図2 8に関して説明される技法を実行することができる。

40

【 0 2 3 5 】

[00292]図2 9を参照すると、オブジェクトトラッキングを使用したビデオ処理方法2 8 0 0の特定の実施形態を示すフローチャートが示される。代替的な実施形態では、方法2 8 0 0は、図1の電子デバイス1 0 2、図2 Bのプロセッサ2 6 4によって実装されるオブジェクトトラッキングおよび検出モジュール2 0 4、プロセッサによって実装される図8のオブジェクトトラッキングおよび検出モジュール8 0 4、プロセッサによって実装される図9の平滑化モジュール9 6 1、またはこれらの任意の組合せを使用して実行され得る。

50

【 0 2 3 6 】

[00293]方法 2 8 0 0 は、2 8 0 2 において、モバイルデバイスにおいて、シーン中の第 1 のオブジェクト、シーン中の第 2 のオブジェクト、およびシーン中の第 3 のオブジェクトの選択を受け取ることを含む。たとえば、図 2 8 を参照すると、ユーザは、画面 2 7 1 0 上のユーザ選択機構 2 7 1 1 をタッチし、第 1 のオブジェクト 2 7 0 2 の周囲で第 1 の焦点リング 2 7 1 2 をドラッグし、第 2 のオブジェクト 2 7 0 4 の周囲で第 2 の焦点リング 2 7 1 4 をドラッグし、第 3 のオブジェクト 2 7 0 6 の周囲で第 3 の焦点リング 2 7 1 6 をドラッグして、それぞれ、第 1、第 2、および第 3 のオブジェクト 2 7 0 2 ~ 2 7 0 6 のトラッキングを可能にすることができる。表示されるシーンは、キャプチャされモバイルデバイスの画面に表示されるビデオストリームに対応し得る。

10

【 0 2 3 7 】

[00294]表示はまた、2 8 0 4 において、アット第 1 のウィンドウ、第 2 のウィンドウ、第 3 のウィンドウ、および第 4 のウィンドウへと区分され得る。たとえば、図 2 8 を参照すると、表示 2 7 3 0 は、第 1 のウィンドウ 2 7 3 2、第 2 のウィンドウ 2 7 3 4、第 3 のウィンドウ 2 7 3 6、および第 4 のウィンドウに区分され得る。

【 0 2 3 8 】

[00295]2 8 0 6 において、第 1 のオブジェクトを含む表示されるシーンの第 1 の部分は、第 1 のウィンドウに表示され得る。たとえば、図 2 8 を参照すると、第 1 のウィンドウ 2 7 3 2 は、第 1 のオブジェクト 2 7 0 2 を含む画像 2 7 2 0 の第 1 の部分を表示することができる。画像 2 7 2 0 の第 1 の部分は、画像 2 7 2 0 の中の第 1 のオブジェクト 2 7 0 2 の位置に基づいて決定され得る。たとえば、モバイルデバイスは第 1 のオブジェクト 2 7 0 2 の位置をトラッキングすることができ、第 1 のオブジェクト 2 7 0 2 の位置に注目しているビデオストリームが第 1 のウィンドウ 2 7 3 2 に表示され得る。

20

【 0 2 3 9 】

[00296]2 8 0 8 において、第 2 のオブジェクトを含む表示されるシーンの第 2 の部分は、第 2 のウィンドウに表示され得る。たとえば、図 2 8 を参照すると、第 2 のウィンドウ 2 7 3 4 は、第 2 のオブジェクト 2 7 0 4 を含む画像 2 7 2 0 の第 2 の部分を表示することができる。画像 2 7 2 0 の第 2 の部分は、画像 2 7 2 0 の中の第 2 のオブジェクト 2 7 0 4 の位置に基づいて決定され得る。たとえば、モバイルデバイスは第 2 のオブジェクト 2 7 0 4 の位置をトラッキングすることができ、第 2 のオブジェクト 2 7 0 4 の位置に注目しているビデオストリームが第 2 のウィンドウ 2 7 3 4 に表示され得る。

30

【 0 2 4 0 】

[00297]2 8 1 0 において、第 3 のオブジェクトを含む表示されるシーンの第 3 の部分は、第 3 のウィンドウに表示され得る。たとえば、図 2 8 を参照すると、第 3 のウィンドウ 2 7 3 6 は、第 3 のオブジェクト 2 7 0 6 を含む画像 2 7 2 0 の第 3 の部分を表示することができる。画像 2 7 2 0 の第 3 の部分は、画像 2 7 2 0 の中の第 3 のオブジェクト 2 7 0 6 の位置に基づいて決定され得る。たとえば、モバイルデバイスは第 3 のオブジェクト 2 7 0 6 の位置をトラッキングすることができ、第 3 のオブジェクト 2 7 0 6 の位置に注目しているビデオストリームが第 3 のウィンドウ 2 7 3 6 に表示され得る。

40

【 0 2 4 1 】

[00298]2 8 1 2 において、各オブジェクト間の幾何学的配置の指示が第 4 のウィンドウに表示され得る。たとえば、図 2 8 を参照すると、第 4 のウィンドウ 2 7 3 8 は、各オブジェクト 2 7 0 2 ~ 2 7 0 6 間の重心 (C_{M2}) を表示することができる。第 4 のウィンドウ 2 7 3 8 は、第 3 の時刻におけるシーンを反映するように重心 (C_{M2}) を更新することができる。たとえば、第 3 のフレームにおいて、第 4 のウィンドウ 2 7 3 8 は、第 3 の時刻における画像 2 7 4 0 中のオブジェクト 2 7 0 2 ~ 2 7 0 6 の重心 (C_{M3}) を表示することができる。第 4 のウィンドウ 2 7 3 8 はまた、第 2 のフレームと第 3 のフレームとの間でオブジェクトの重心をトラッキングする、経路 (たとえば、点線) を表示することができる。

【 0 2 4 2 】

50

[00299]図29の方法2800は、モバイルデバイスが、シーン2700、2720、2740中のオブジェクト2702~2706をトラッキングすることと、オブジェクト2702~2706に対応する個別化された(たとえば、専心的な)ビデオストリームを別のウィンドウ2732~2736に表示することとを可能にし得る。方法2800はまた、ユーザが第4のウィンドウ2738中のオブジェクト2702~2706の幾何学的配置を見ることを可能にする。モバイルデバイスは、単一のカメラを使用して、または複数のカメラを使用して、方法2800を実行することができる。

【0243】

[00300]図42を参照すると、オブジェクトトラッキングを使用したビデオ処理を示す特定の実施形態が示される。図42に示される実施形態に関して説明されるビデオ処理技法は、図1の電子デバイス102、図2Bのプロセッサ264によって実装されるオブジェクトトラッキングおよび検出モジュール204、プロセッサによって実装される図8のオブジェクトトラッキングおよび検出モジュール804、プロセッサによって実装される図9の平滑化モジュール961、またはこれらの任意の組合せによって実行され得る。

【0244】

[00301]図42は、第1の時刻におけるシーンを示す、デジタル画像1100を示す。たとえば、画像1100は、キャプチャされモバイルデバイスの画面に表示されるビデオストリームに対応し得る。モバイルデバイスは、ストリームを(たとえば、モバイルデバイスのカメラによって)キャプチャし、または別のデバイスからストリームを受信するように構成され得る。画像1100に示されるシーンは、第1の関心領域(ROI: region of interest)1102と、第2のROI1104と、第3のROI1106とを含み得る。各ROI1102~1106は、本明細書で説明される技法(たとえば、図1~図10を参照すると、オブジェクトトラッキングおよび検出モジュール204および804、ならびに/またはタスクT200)に従ってトラッキングされる1つまたは複数のオブジェクトを含むシーン内の領域に対応し得る。

【0245】

[00302]例示すると、第1のROI1102は、トラッキングされたオブジェクトの第1のグループ(たとえば、第1のトラッキングされたオブジェクト1110および第3のトラッキングされたオブジェクト1114)を含み得る。加えて、第2のROI1104は、トラッキングされたオブジェクトの第2のグループ(たとえば、第4のトラッキングされたオブジェクト1120、第5のトラッキングされたオブジェクト1122、および第6のトラッキングされたオブジェクト1124)を含んでよく、第3のROI1106は、トラッキングされたオブジェクトの第3のグループ(たとえば、第7のトラッキングされたオブジェクト1130、第8のトラッキングされたオブジェクト1132、および第9のトラッキングされたオブジェクト1134)を含み得る。第2のトラッキングされたオブジェクト1112は、ROI1102~1104の外側に位置し得る。

【0246】

[00303]メタデータは、シーン1100中の各々のトラッキングされたオブジェクト1110~1114、1120~1124、1130~1134の位置を特定するために(たとえば、タスクT300によって)生成され得る。特定の実施形態では、メタデータは矩形座標データに対応する。たとえば、第1のメタデータは、第1のトラッキングされたオブジェクト1110の位置(たとえば、x座標位置およびy座標位置)を特定するために生成されてよく、第2のメタデータは、第2のトラッキングされたオブジェクト1112の位置を特定するために生成されてよく、以下同様である。他の実施形態では、メタデータは、極座標データ、球面座標データ、円柱座標データ、または3次元直交座標データに対応し得る。

【0247】

[00304]各ROI1102~1106は、ROI1102~1106中のトラッキングされたオブジェクトの重心によって定義され得る。たとえば、第1のROI1102は、第1のトラッキングされたオブジェクト1110と第3のトラッキングされたオブジェク

10

20

30

40

50

ト 1 1 1 4 との間の第 1 の重心 (C_{M1}) によって定義され得る。第 2 の R O I 1 1 0 4 は、第 4 のトラッキングされたオブジェクト 1 1 2 0、第 5 のトラッキングされたオブジェクト 1 1 2 2、および第 6 のトラッキングされたオブジェクト 1 1 2 4 の間の第 2 の重心 (C_{M2}) によって定義され得る。第 3 の R O I 1 1 0 6 は、第 7 のトラッキングされたオブジェクト 1 1 3 0、第 8 のトラッキングされたオブジェクト 1 1 3 2、および第 9 のトラッキングされたオブジェクト 1 1 3 4 の間の第 3 の重心 (C_{M3}) によって定義され得る。

【 0 2 4 8 】

[00305] 図 4 3 に関してより詳細に説明されるように、シーン 1 1 0 0 中の各々のトラッキングされたオブジェクト 1 1 1 0 ~ 1 1 1 4、1 1 2 0 ~ 1 1 2 4、1 1 3 0 ~ 1 1 3 4 の位置のためのメタデータを符号化するために、コードブックが使用され得る。R O I 1 1 0 2 ~ 1 1 0 6 の 1 つの内側に位置するトラッキングされたオブジェクトのためのメタデータは、R O I 1 1 0 2 ~ 1 1 0 6 の外側に位置するトラッキングされたオブジェクト (たとえば、第 2 のトラッキングされたオブジェクト 1 1 1 2) のためのメタデータよりも高いビットレートで符号化され得る。より高いビットレートで R O I 1 1 0 2 ~ 1 1 0 6 の 1 つの内部に位置するトラッキングされたオブジェクトのためのメタデータを符号化することは、R O I 1 1 0 2 ~ 1 1 0 6 中のトラッキングされたオブジェクトの位置が、他のトラッキングされたオブジェクトの位置よりも高いレベルの精度で符号化されることを可能にし得る。

【 0 2 4 9 】

[00306] 図 4 3 を参照すると、オブジェクトの位置のためのメタデータを符号化する特定の実施形態が示される。図 4 3 の実施形態は、フレームごとに複数のオブジェクトの位置のためのメタデータを符号化するために使用され得るコードブック 1 2 0 2 を示す。メタデータは、格子中の各オブジェクトの位置を示すために使用され得る。

【 0 2 5 0 】

[00307] コードブック 1 2 0 2 は、複数のフレーム 1 2 0 4 に対するメタデータを符号化するために使用され得る。メタデータは、フレーム中のオブジェクトのための矩形座標データ (たとえば、x 座標および y 座標)、フレーム中のオブジェクトのための極座標、フレーム中のオブジェクトのための球面座標、フレーム中のオブジェクトのための円柱座標、またはフレーム中のオブジェクトのための 3 次元直交座標を含み得る。示される実施形態では、コードブック 1 2 0 2 は、1 0 0 0 0 個のフレームにわたって 3 つのオブジェクトのためのメタデータを符号化することができる。特定の実施形態では、第 3 のフレームは、図 4 3 の第 1 のフレーム (たとえば、シーン 1 1 0 0) に対応し得る。1 0 0 0 0 個のフレームに対するメタデータがコードブック 1 2 0 2 において示されるが、他の実施形態では、コードブック 1 2 0 2 は、より少数の (または追加の) フレームに対するメタデータを符号化することができる。加えて、メタデータは、より少数の (または追加の) オブジェクトに対する位置情報を含み得る。図 4 3 の以下の説明は、第 3 のフレームに対応するメタデータに基づく。しかしながら、同様の技法が他のメタデータに適用可能であり得る。

【 0 2 5 1 】

[00308] 第 3 のフレーム中のメタデータは、第 1 の座標データ 1 2 3 2 と、第 2 の座標データ 1 2 3 4 と、第 3 の座標データ 1 2 3 6 と、第 4 の座標データ 1 2 3 7 と、...、第 9 の座標データ 1 2 3 8 とを含み得る。第 1 の座標データ 1 2 3 2 は、図 4 2 の画像 1 1 0 0 中の第 1 のトラッキングされたオブジェクト 1 1 1 0 の座標の位置を含んでよく、第 2 の座標データ 1 2 3 4 は、画像 1 1 0 0 中の第 2 のトラッキングされたオブジェクト 1 1 1 2 の座標の位置を含んでよく、第 3 の座標データ 1 2 3 6 は、画像 1 1 0 0 中の第 3 のトラッキングされたオブジェクト 1 1 1 4 の座標の位置を含んでよく、第 4 の座標データ 1 2 3 7 は、画像 1 1 0 0 中の第 4 のトラッキングされたオブジェクト 1 1 2 0 の座標の位置を含んでよく、以下同様である。

【 0 2 5 2 】

[00309] プロセッサは、座標データ 1 2 3 2 ~ 1 2 3 6 を符号化するためにコードブック 1 2 0 2 を使用することができる。例示すると、プロセッサは、第 1 のビットレート（たとえば、比較的高いビットレート）を使用して、第 1 の座標データ 1 2 3 2 を第 1 の符号化された位置データ 1 2 4 2 へと符号化（たとえば、量子化）することができる。たとえば、プロセッサは、第 1 の座標データ 1 2 3 2 を対応する値（たとえば、コードブック 1 2 0 2 のコードワード）にマッピングすることができ、第 1 の座標データ 1 2 3 2 を値（たとえば、第 1 の符号化された位置データ 1 2 4 2）として符号化することができる。第 1 のトラッキングされたオブジェクト 1 1 1 0 は R O I（たとえば、図 4 2 の第 1 の R O I 1 1 0 2）の内部にあるので、プロセッサは第 1 のビットレートを使用することができる。

10

【 0 2 5 3 】

[00310] 加えて、プロセッサは、第 1 のビットレートよりも低い第 2 のビットレートを使用して、第 2 の座標データ 1 2 3 4 を第 2 の符号化された位置データ 1 2 4 4 へと符号化することができる。たとえば、プロセッサは、コードブック 1 2 0 2 を使用して第 2 の座標データ 1 2 3 4 を対応する値にマッピングすることができ、第 2 の座標データ 1 2 3 4 を値（たとえば、第 2 の符号化された位置データ 1 2 4 4）として符号化することができる。第 2 のトラッキングされたオブジェクト 1 1 1 2 は図 4 2 の R O I 1 1 0 2 ~ 1 1 0 6 の各々の外部にあるので、プロセッサは第 2 のビットレートを使用することができる。

【 0 2 5 4 】

20

[00311] プロセッサは、第 2 のビットレートよりも高い第 3 のビットレートを使用して、第 3 の座標データ 1 2 3 6 を第 3 の符号化された位置データ 1 2 4 6 へと符号化することができる。たとえば、プロセッサは、コードブック 1 2 0 2 を使用して第 3 の座標データ 1 2 3 6 を対応する値にマッピングして、第 3 の矩形座標データ 1 2 3 6 を値（たとえば、第 3 の符号化された位置データ 1 2 4 6）として符号化することができる。第 3 のトラッキングされたオブジェクト 1 1 1 4 は R O I（たとえば、図 4 2 の第 1 の R O I 1 1 0 2）の内部にあるので、プロセッサは第 3 のビットレートを使用することができる。特定の実施形態では、第 3 のビットレートは、第 1 のビットレートに等しくてよく、または実質的に等しくてよい。

【 0 2 5 5 】

30

[00312] プロセッサは、第 2 のビットレートよりも高い第 4 のビットレートを使用して、第 4 の座標データ 1 2 3 7 を第 4 の符号化された位置データ 1 2 4 7 へと符号化することができる。たとえば、プロセッサは、コードブック 1 2 0 2 を使用して第 4 の座標データ 1 2 3 7 を対応する値にマッピングすることができ、第 4 の座標データ 1 2 3 7 を値（たとえば、第 4 の符号化された位置データ 1 2 4 7）として符号化することができる。第 4 のトラッキングされたオブジェクト 1 1 2 0 は R O I（たとえば、図 4 2 の第 2 の R O I 1 1 0 4）の内部にあるので、プロセッサは第 4 のビットレートを使用することができる。特定の実施形態では、第 4 のビットレートは、第 1 のビットレートに実質的に等しくてよい。

【 0 2 5 6 】

40

[00313] プロセッサは、第 2 のビットレートよりも高い第 9 のビットレートを使用して、第 9 の座標データ 1 2 3 8 を第 9 の符号化された位置データ 1 2 4 8 へと符号化することができる。たとえば、プロセッサは、コードブック 1 2 0 2 を使用して第 9 の座標データ 1 2 3 8 を対応する値にマッピングすることができ、第 9 の座標データ 1 2 3 8 を値（たとえば、第 9 の符号化された位置データ 1 2 4 8）として符号化することができる。第 9 のトラッキングされたオブジェクト 1 1 3 4 は R O I（たとえば、図 4 2 の第 3 の R O I 1 1 0 6）の内部にあるので、プロセッサは第 9 のビットレートを使用することができる。特定の実施形態では、第 9 のビットレートは、第 1 のビットレートに実質的に等しくてよい。

【 0 2 5 7 】

50

[00314]同様の符号化技法が、図42における他のトラッキングされたオブジェクトの座標データのために実行され得る。ROI1102~1106の1つの内側に位置するトラッキングされたオブジェクトのためのメタデータ(たとえば、座標データ)は、ROI1102~1106の外側に位置するトラッキングされたオブジェクト(たとえば、第2のトラッキングされたオブジェクト1112)のためのメタデータよりも高いビットレートで符号化され得る。より高いビットレートでROI1102~1106の1つの内部に位置するトラッキングされたオブジェクトのためのメタデータを符号化することは、トラッキングされたオブジェクトの位置が比較的高いレベルの精度で符号化されることを可能にし得る。

【0258】

10

[00315]図43の実施形態は、ユーザが、コードブック1202を使用して符号化されたデータに基づいてビデオの特定の部分を検索する(たとえば、位置特定する)ことを可能にし得る。たとえば、位置空間は、ユーザが異なるオブジェクトに対する異なる位置(たとえば、パラメータ)を選択することを可能にする、ユーザ対話型の格子として画面に表示され得る。この選択に基づいて、モバイルデバイスのユーザは、ビデオを見進める(たとえば、再生する、早送りする、巻き戻すなど)必要なく、同様のオブジェクトの位置構成を有するビデオの所望の部分(たとえば、フレーム)を容易に位置特定することができる。加えて、符号化された位置データ1242~1248は、圧縮されてエンコードからデコードに送信され得る。符号化された位置データ1242~1248を圧縮してデコードに送信することで、デコードは、比較的少量のデータ(たとえば、量子化された値)を使用して、トラッキングされたオブジェクトの位置を決定することが可能になり得る。

20

【0259】

[00316]図44を参照すると、オブジェクトトラッキングを使用したビデオ処理方法1300の特定の実施形態を示すフローチャートが示される。代替的な実施形態では、方法1300は、図1の電子デバイス102、図2Bのプロセッサ264によって実装されるオブジェクトトラッキングおよび検出モジュール204、プロセッサによって実装される図8のオブジェクトトラッキングおよび検出モジュール804、プロセッサによって実装される図9の平滑化モジュール961、図43のコードブック1202、またはこれらの任意の組合せを使用して実行され得る。

【0260】

30

[00317]方法1300は、1302において、モバイルデバイスにおいて、特定のフレームの関心領域に位置する第1のオブジェクトの第1の位置を(たとえば、タスクT200によって)トラッキングすることを含み得る。たとえば、図42を参照すると、第1のトラッキングされたオブジェクト1110の第1の位置がトラッキングされ得る。1304において、特定のフレームの関心領域に位置しない第2のオブジェクトの第2の位置が、トラッキングされ得る。たとえば、図42を参照すると、第2のトラッキングされたオブジェクト1112の第2の位置がトラッキングされ得る。

【0261】

[00318]1306において、第1の位置と関連付けられる第1の座標データおよび第2の位置と関連付けられる座標データが生成され得る。たとえば、図43を参照すると、プロセッサは、シーン1100中の第1のトラッキングされたオブジェクト1110の位置と関連付けられる第1の座標データ1232と、シーン1100中の第2のトラッキングされたオブジェクト1112の位置と関連付けられる第2の座標データ1234とを生成することができる。

40

【0262】

[00319]1308において、第1の座標データおよび第2の座標データは、第1の符号化された位置データと第2の符号化された位置データとを生成するために、コードブックを使用して(たとえば、タスクT300によって)符号化され得る。たとえば、図43を参照すると、プロセッサは、第1の符号化された位置データ1242(たとえば、コードブック1202のコードワード)を生成するために、コードブック1202を使用して第

50

1の座標データ1232を符号化することができ、プロセッサは、第2の符号化された位置データ1244を生成するために、コードブック1202を使用して第2の座標データ1234を符号化することができる。第1の座標データ1232は第1のビットレート（たとえば、高ビットレート）で符号化されてよく、第2の座標データ1234は第2のビットレート（たとえば、低ビットレート）で符号化されてよい。第1のトラッキングされたオブジェクト1110はROI（たとえば、図42の第1のROI1102）の内部にあるので、プロセッサは第1のビットレートを使用して第1の座標データ1232を符号化することができる。第2のトラッキングされたオブジェクト1112は図42のROI1102～1106の各々の外部にあるので、プロセッサは第2のビットレートを使用して第2の座標データ1234を符号化することができる。

10

【0263】

[00320]1310において、第1の符号化された位置データおよび第2の符号化された位置データは、入力パラメータに基づく特定のフレームの検索を可能にするために記憶され得る。たとえば、図43を参照すると、プロセッサは、第1の符号化された位置データ1242と第2の符号化された位置データ1244とをメモリに記憶することができる。モバイルデバイスのユーザは、パラメータ（たとえば、第1のトラッキングされたオブジェクト1110および第2のトラッキングされたオブジェクト1112の所望の位置）を入力することができる。入力パラメータに基づいて、モバイルデバイスは、第1の位置および第2の位置に対応する入力パラメータに応答して、再生のために第3のフレームを検索することができる。

20

【0264】

[00321]特定の実施形態では、方法1300は、特定のフレームの関心領域に位置する第3のオブジェクトの第3の位置をトラッキングすることを含み得る。たとえば、第3のトラッキングされたオブジェクト1114の第3の位置がトラッキングされ得る。第3の位置と関連付けられる第3の座標データが生成され得る。第3の座標データは、第3の符号化された位置データを生成するためにコードブックを使用して符号化され得る。第2のビットレートよりも高い第3のビットレートを使用して、第3の座標データが符号化され得る。第3の符号化された位置データは、入力パラメータに基づく特定のフレームの検索を可能にするために記憶され得る。

【0265】

30

[00322]特定の実施形態では、方法1300は、特定のフレームの第2の関心領域に位置する第4のオブジェクトの第4の位置をトラッキングすることを含み得る。たとえば、第4のトラッキングされたオブジェクト1120の第4の位置がトラッキングされ得る。第4の位置と関連付けられる第4の座標データが生成され得る。第4の座標データは、第4の符号化された位置データを生成するためにコードブックを使用して符号化され得る。第2のビットレートよりも高い第4のビットレートを使用して、第4の座標データが符号化され得る。第4の符号化された位置データは、入力パラメータに基づく特定のフレームの検索を可能にするために記憶され得る。

【0266】

[00323]図44の方法1300は、コードブック1202を使用して符号化されたビデオの特定の部分（フレーム）をユーザが検索する（たとえば、位置特定する）ことを可能にし得る。加えて、方法1300は、ROIの内側のトラッキングされたオブジェクトのためのメタデータが、ROIの外側のオブジェクトよりも高いビットレートで符号化されることを可能にし得る。より高いビットレートでROIの内部に位置するトラッキングされたオブジェクトのためのメタデータを符号化することは、ROIの中のトラッキングされたオブジェクトの位置が、他のトラッキングされたオブジェクトの位置よりも高いレベルの精度で符号化されることを可能にし得る。

40

【0267】

[00324]図38Aは、一般的な構成によるメタデータを生成するための装置A100のブロック図を示す。装置A100は、複数のオブジェクトを選択するために少なくとも1

50

つのデジタル画像を使用するように構成される判別器 1 0 0 を含み、（たとえば、タスク T 1 0 0 を参照して本明細書で説明されるように）少なくとも 1 つのデジタル画像は物理空間との関連でオブジェクトを示す。装置 A 1 0 0 はまた、位置空間中の選択されたオブジェクトの位置を示す情報に基づいて、（たとえば、タスク T 3 0 0 を参照して本明細書で説明される）選択されたオブジェクトの複数の幾何学的構成の候補の中から 1 つを特定するメタデータを生成するように構成される選択器 3 0 0 を含む。

【 0 2 6 8 】

[00325] 図 3 8 B は、装置 A 1 0 0 の実装形態 A 1 1 0 のブロック図を示す。装置 A 1 1 0 は、（たとえば、表示タスク T 5 0 を参照して本明細書で説明されるように）少なくとも 1 つのデジタル画像を表示するように構成される表示モジュール 5 0 を含む。

10

【 0 2 6 9 】

[00326] 図 3 8 C は、装置 A 1 0 0 の実装形態 A 1 2 0 のブロック図を示す。装置 A 1 2 0 は、（たとえば、解析タスク T 3 0 を参照して本明細書で説明されるように）少なくとも 1 つのデジタル画像を生成するためにビデオストリームを解析するように構成される、インターフェース 3 0 を含む。図 3 8 D は、装置 A 1 1 0 および A 1 2 0 の実装形態 A 1 3 0 のブロック図を示す。

【 0 2 7 0 】

[00327] 図 3 8 E は、装置 A 1 0 0 の実装形態 A 2 0 0 のブロック図を示す。装置 A 2 0 0 は、（たとえば、タスク T 2 0 0 を参照して本明細書で説明されるように）位置空間中の複数のオブジェクトのための基準の位置を決定するように構成されるオブジェクトトラッカー 2 0 0 を含む。オブジェクトトラッカー 2 0 0 はまた、ビデオストリーム内で、（たとえば、タスク T 2 0 0 を参照して本明細書で説明されたように）前記複数のオブジェクトをトラッキングするように構成され得る。図 3 8 F、図 3 8 G、および図 3 9 A は、それぞれ、オブジェクトトラッカー 2 0 0 の実体を含む、装置 A 1 1 0、A 1 2 0、および A 1 3 0 の実装形態 A 2 1 0、A 2 2 0、および A 2 3 0 のブロック図を示す。

20

【 0 2 7 1 】

[00328] 図 3 9 B は、装置 A 1 0 0 の実装形態 A 3 0 0 のブロック図を示す。装置 A 2 0 0 は、（たとえば、タスク T 3 0 0 を参照して本明細書で説明されるように）選択器 3 0 0 によって生成されるメタデータをビデオストリームの対応するフレームと関連付けるように構成されるビデオストリームエディタ 4 0 0 を含む。図 3 9 C および図 3 9 D は、ビデオストリームエディタ 4 0 0 の実体をそれぞれ含む、装置 A 2 0 0 および A 2 3 0 の実装形態 A 3 1 0 および A 3 3 0 のブロック図を示す。

30

【 0 2 7 2 】

[00329] 本明細書で開示されるような装置 A 1 0 0 の実装形態のいずれかは、次のフォームファクタ、すなわち、携帯可能（たとえば、スマートフォン）、運転可能（たとえば、車両またはロボット）、装着可能（たとえば、衣服またはアクセサリ）、および飛行可能（たとえば、ドローン）のいずれかのデバイスのような、モバイルデバイス内に実装され得る。モバイルデバイスは、少なくとも 1 つのデジタル画像を表示するように構成される 1 つもしくは複数の画面（たとえば、タッチスクリーン）および / または少なくとも 1 つのデジタル画像（たとえば、ビデオストリーム）をキャプチャするように構成される 1 つもしくは複数の撮像（imaging）ユニット（たとえば、カメラまたは他の画像キャプチャデバイス）を含み得る。

40

【 0 2 7 3 】

[00330] 図 3 1 を参照すると、上で説明されたビデオ処理技法を実行するように動作可能なコンポーネントを含むワイヤレスデバイス 2 9 0 0 のブロック図が示されている。デバイス 2 9 0 0 は、メモリ 2 9 3 2 に結合されたデジタル信号プロセッサ（DSP）のようなプロセッサ 2 9 1 0 を含む。デバイス 2 6 0 0 は、本明細書で説明されるような装置 A 1 0 0 の実装形態のいずれかの要素を含むように実装され得るモバイルデバイスの一例である。

【 0 2 7 4 】

50

[00331]図31はまた、プロセッサ2910とディスプレイ2928とに結合されたディスプレイコントローラ2926を示している。特定の実施形態では、表示は、上で説明されたビューファインダまたは画面に対応し得る。メモリ2932は、実行可能命令2956を含む有形の非一時的プロセッサ可読記憶媒体であり得る。命令2956はプロセッサ2910のようなプロセッサによって実行され得る。プロセッサ2910またはディスプレイコントローラ2926のようなプロセッサは、図3の方法300、図4の方法400、図5の方法500、図6の方法600、図10の方法1000、図12の方法1200、図14の方法1400、図16の方法1600、図19の方法1900、図20の方法2000、2010、図23の方法2300、図25の方法2500、図26の方法2600、図28の方法2800、方法M100、方法M110、方法M120、方法M130、またはこれらの任意の組合せの、少なくとも一部分を実行するように命令2956を実行することができる。

【0275】

[00332]コーダ/デコーダ(コーデック)2934も、プロセッサ2910に結合され得る。スピーカー2936およびマイクロフォン2938が、コーデック2934に結合され得る。図31は、ワイヤレスコントローラ2940がプロセッサ2910とアンテナ2942とに結合され得ることも示す。特定の実施形態では、高周波(RF)インターフェース2980が、ワイヤレスコントローラ2940とアンテナ2942との間に配設される。

【0276】

[00333]特定の実施形態では、プロセッサ2910、ディスプレイコントローラ2926、メモリ2932、コーデック2934、およびワイヤレスコントローラ2940は、システムインパッケージまたはシステムオンチップデバイス2922に含まれる。特定の実施形態では、入力デバイス2930(たとえば、上で説明されたようなシーンをキャプチャするために使用されるカメラ)および電源2944が、システムオンチップデバイス2922に結合される。その上、特定の実施形態では、図31に示されるように、ディスプレイ2928、入力デバイス2930、スピーカー2936、マイクロフォン2938、アンテナ2942、電源2944、およびRFインターフェース2980は、システムオンチップデバイス2922の外部にある。たとえば、RFインターフェース2980は別のチップであってよい。しかしながら、ディスプレイ2928、入力デバイス2930、スピーカー2936、マイクロフォン2938、アンテナ2942、電源2944、およびRFインターフェース2980の各々は、インターフェースまたはコントローラのような、システムオンチップデバイス2922のコンポーネントに結合され得る。

【0277】

[00334]図40Aは、一般的な構成によるメタデータを生成するための装置MF100のブロック図を示す。装置MF100は、複数のオブジェクトを選択するために少なくとも1つのデジタル画像を使用するための手段F100を含み、(たとえば、タスクT100を参照して本明細書で説明されるように)少なくとも1つのデジタル画像は物理空間との関連でオブジェクトを示す。装置MF100はまた、位置空間中の選択されたオブジェクトの位置を示す情報に基づいて、(たとえば、タスクT300を参照して本明細書で説明される)選択されたオブジェクトの複数の幾何学的構成の候補の中から1つを特定するメタデータを生成するための手段F300を含む。

【0278】

[00335]図40Bは、装置MF100の実装形態MF110のブロック図を示す。装置MF110は、(たとえば、表示タスクT50を参照して本明細書で説明されるように)少なくとも1つのデジタル画像を表示するための手段F50を含む。

【0279】

[00336]図40Cは、装置MF100の実装形態MF120のブロック図を示す。装置MF120は、(たとえば、解析タスクT30を参照して本明細書で説明されるように)少なくとも1つのデジタル画像を生成するためにビデオストリームを解析するための手段

F 3 0を含む。図 4 0 Dは、装置 M F 1 1 0および M F 1 2 0の実装形態 M F 1 3 0のブロック図を示す。

【 0 2 8 0 】

[00337]図 4 0 Eは、装置 M F 1 0 0の実装形態 M F 2 0 0のブロック図を示す。装置 M F 2 0 0は、（たとえば、タスク T 2 0 0を参照して本明細書で説明されるように）位置空間中の複数のオブジェクトのための基準の位置を決定するための手段 F 2 0 0を含む。図 4 1 Bは、（たとえば、タスク T 2 0 0を参照して本明細書で説明されたように）ビデオストリーム内で前記複数のオブジェクトをトラッキングするための手段 F 2 1 0を含む、装置 M F 1 0 0の実装形態 M F 4 0 0のブロック図を示す。図 4 0 F、図 4 0 G、および図 4 1 Aは、手段 F 2 0 0の実体をそれぞれ含む、装置 M F 1 1 0、M F 1 2 0、および M F 1 3 0の実装形態 M F 2 1 0、M F 2 2 0、および M F 2 3 0のブロック図を示す。そのような装置はまた、手段 F 2 0 0に加えて、またはその代わりに、手段 F 2 1 0を含むように実施され得る。

【 0 2 8 1 】

[00338]図 4 1 Cは、装置 M F 1 0 0の実装形態 M F 3 0 0のブロック図を示す。装置 M F 2 0 0は、（たとえば、タスク T 3 0 0を参照して本明細書で説明されるように）手段 F 3 0 0によって生成されるメタデータをビデオストリームの対応するフレームと関連付けるための手段 F 4 0 0を含む。図 4 1 Dおよび図 4 1 Eは、手段 F 4 0 0の実体をそれぞれ含む、装置 M F 2 0 0および M F 2 3 0の実装形態 M F 3 1 0および M F 3 3 0のブロック図を示す。

【 0 2 8 2 】

[00339]さらに、本明細書で開示された実施形態に関して説明された様々な例示的な論理ブロック、構成、モジュール、回路、およびアルゴリズムステップは、電子ハードウェア、プロセッサによって実行されるコンピュータソフトウェア、または両方の組合せとして実装され得ることを、当業者は理解されよう。本明細書で開示された装置（たとえば、装置 M F 1 0 0、M F 1 1 0、M F 1 2 0、M F 1 3 0、M F 2 0 0、M F 2 1 0、M F 2 2 0、M F 2 3 0、M F 3 0 0、M F 3 1 0、M F 3 3 0、M F 4 0 0、A 1 0 0、A 1 1 0、A 1 2 0、A 1 3 0、A 2 0 0、A 2 1 0、A 2 2 0、A 2 3 0、A 3 0 0、A 3 1 0、および A 3 3 0の中のいずれか）の様々な実装形態の1つまたは複数の要素は、全体または一部が、マイクロプロセッサ、組込みプロセッサ、IPコア、デジタル信号プロセッサ、FPGA（フィールドプログラマブルゲートアレイ）、ASSP（特定用途向け標準製品）、およびASIC（特定用途向け集積回路）のような、論理要素（たとえば、プロセッサ 2 6 1 0）の1つまたは複数の固定アレイまたはプログラマブルアレイ上で実行されるように構成された命令の1つまたは複数のセットとして実装され得る。様々な例示的なコンポーネント、ブロック、構成、モジュール、回路、およびステップが、上では概して、それらの機能に関して説明された。そのような機能が、ハードウェアとして実装されるかまたはプロセッサ実行可能命令として実装されるかは、具体的な適用例、およびシステム全体に課せられる設計制約に依存する。当業者は、説明されている機能を具体的な適用例ごとに様々な方法で実装することができるが、そのような実装の決定は、本開示の範囲からの逸脱を引き起こすと解釈されるべきではない。

【 0 2 8 3 】

[00340]本明細書で開示された実施形態に関して説明された方法またはアルゴリズムのステップは、直接ハードウェアで、プロセッサによって実行されるソフトウェアモジュールで、またはそれら2つの組合せで具現化され得る。ソフトウェアモジュールは、ランダムアクセスメモリ（RAM）、フラッシュメモリ、読取り専用メモリ（ROM）、プログラマブル読取り専用メモリ（PROM）、消去可能プログラマブル読取り専用メモリ（EEPROM（登録商標））、レジスタ、ハードディスク、リムーバブルディスク、コンパクトディスク読取り専用メモリ（CD-ROM）、または当技術分野で知られている任意の他の形態の非一時的記憶媒体に存在し得る。例示的な記憶媒体は、プロセッサが記憶媒体から情報を読み取り、

記憶媒体に情報を書き込むことができるようにプロセッサに結合される。代替として、記憶媒体は、プロセッサと一体であってよい。プロセッサおよび記憶媒体は、特定用途向け集積回路（ASIC）の中に存在し得る。ASICはコンピューティングデバイスまたはユーザ端末の中に存在し得る。代替として、プロセッサおよび記憶媒体は、コンピューティングデバイスまたはユーザ端末内に個別のコンポーネントとして存在し得る。

【0284】

[00341]開示された実施形態の上記の説明は、当業者が開示された実施形態を製作または使用することを可能にするために提供されている。これらの実施形態に対する様々な修正は、当業者には容易に明らかであり、本明細書で定義されている原理は、本開示の範囲から逸脱することなく、他の実施形態に適用され得る。したがって、本開示は、本明細書

10

に示された実施形態に限定されることを意図されておらず、以下の特許請求の範囲によって定義される原理および新規な特徴と一致する可能な最も広い範囲を与えられるべきである。

以下に本願の出願当初の特許請求の範囲に記載された発明を付記する。

[C1] メタデータを生成する方法であって、

複数のオブジェクトを選択するために少なくとも1つのデジタル画像を使用することと、
ここにおいて、前記少なくとも1つのデジタル画像は物理空間との関連で前記複数のオブジェクトを示す、

少なくとも1つのプロセッサによって、位置空間中の前記選択されたオブジェクトの位置を示す情報に基づいて、前記選択されたオブジェクトの複数の幾何学的構成の候補のうちの1つを特定するメタデータを生成することと、

20

を備える方法。

[C2] 前記方法は、前記位置空間中の前記複数のオブジェクトについて基準位置を決定することを備え、

前記情報は、前記決定された基準位置に対する前記選択されたオブジェクトの位置を示す、

C1に記載のメタデータを生成する方法。

[C3] メタデータを前記生成することは、前記複数の幾何学的構成の候補のうちのどれが、前記選択されたオブジェクトの前記位置によって表される形状と最も類似しているかを決定することを備える、C1に記載のメタデータを生成する方法。

30

[C4] メタデータを生成することは、前記複数の幾何学的構成の候補のうちのどれが、前記位置空間中の前記選択されたオブジェクトの分布と最も類似しているかを決定することを備える、C1に記載のメタデータを生成する方法。

[C5] 前記位置空間中の前記複数のオブジェクトについて基準位置を決定することを備え、

前記決定された基準位置は、(A)前記少なくとも1つのデジタル画像のピクセル座標空間および(B)前記物理空間のうちの1つの中の前記選択されたオブジェクトの位置に基づく、C1に記載のメタデータを生成する方法。

[C6] 前記位置空間は地面である、C1に記載のメタデータを生成する方法。

[C7] 前記生成されたメタデータをビデオストリームのフレームと関連付けることと、
前記位置空間中の前記選択されたオブジェクトの各々の第2の位置を示す情報に基づいて、前記選択されたオブジェクトの前記複数の幾何学的構成の候補のうちの異なる1つを特定する第2のメタデータを生成することと、

40

前記生成された第2のメタデータを前記ビデオストリームの第2のフレームと関連付けることと、

を備える、C1に記載のメタデータを生成する方法。

[C8] 前記位置空間は、(A)前記少なくとも1つのデジタル画像のピクセル座標空間および(B)前記物理空間のうちの1つである、C1に記載のメタデータを生成する方法。

[C9] ビデオストリーム内で、前記複数のオブジェクトをトラッキングすることを備え

50

る、C 1 に記載のメタデータを生成する方法。

[C 1 0] プロセッサによって実行されると、前記プロセッサに、

複数のオブジェクトを選択するために少なくとも1つのデジタル画像を使用することと、
ここにおいて、前記少なくとも1つのデジタル画像は物理空間との関連で前記複数のオブジェクトを示し、

位置空間中の前記選択されたオブジェクトの位置を示す情報に基づいて、前記選択されたオブジェクトの複数の幾何学的構成の候補のうちの1つを特定するメタデータを生成することと、

を行わせる命令を記憶する非一時的コンピュータ可読媒体。

[C 1 1] 前記プロセッサによって実行されると、前記プロセッサに、前記位置空間中の前記複数のオブジェクトについて基準位置を決定させる命令を記憶し、

前記情報は、前記決定された基準位置に対する前記選択されたオブジェクトの位置を示す、C 1 0 に記載の非一時的コンピュータ可読媒体。

[C 1 2] メタデータを前記生成することは、前記複数の幾何学的構成の候補のうちのどれが前記選択されたオブジェクトの前記位置によって表される形状と最も類似しているかを決定することを備える、C 1 0 に記載の非一時的コンピュータ可読媒体。

[C 1 3] メタデータを生成することは、前記複数の幾何学的構成の候補のうちのどれが前記位置空間中の前記選択されたオブジェクトの分布と最も類似しているかを決定することを備える、C 1 0 に記載の非一時的コンピュータ可読媒体。

[C 1 4] 前記位置空間が地面である、C 1 0 に記載の非一時的コンピュータ可読媒体。

[C 1 5] 前記プロセッサによって実行されると、前記プロセッサに、

前記生成されたメタデータをビデオストリームのフレームと関連付けることと、

前記位置空間中の前記選択されたオブジェクトの各々の第2の位置を示す情報に基づいて、前記選択されたオブジェクトの前記複数の幾何学的構成の候補のうちの異なる1つを特定する第2のメタデータを生成することと、

前記生成された第2のメタデータを前記ビデオストリームの第2のフレームと関連付けることと、

を行わせる命令を記憶する、C 1 0 に記載の非一時的コンピュータ可読媒体。

[C 1 6] 前記プロセッサによって実行されると、前記プロセッサに、ビデオストリーム内で前記複数のオブジェクトをトラッキングさせる命令を記憶する、C 1 0 に記載の非一時的コンピュータ可読媒体。

[C 1 7] メタデータを生成するための装置であって、

複数のオブジェクトを選択するために少なくとも1つのデジタル画像を使用するための手段と、ここにおいて、前記少なくとも1つのデジタル画像は物理空間との関連で前記複数のオブジェクトを示し、

位置空間中の前記選択されたオブジェクトの位置を示す情報に基づいて、前記選択されたオブジェクトの複数の幾何学的構成の候補のうちの1つを特定するメタデータを生成するための手段と、

を備える装置。

[C 1 8] 前記位置空間中の前記複数のオブジェクトについて基準位置を決定するための手段を備え、

前記情報は、前記決定された基準位置に対する前記選択されたオブジェクトの位置を示す、C 1 7 に記載の装置。

[C 1 9] メタデータを前記生成することは、前記複数の幾何学的構成の候補のうちのどれが、前記選択されたオブジェクトの前記位置によって表される形状と最も類似しているかを決定することを備える、C 1 7 に記載の装置。

[C 2 0] メタデータを生成することが、前記複数の幾何学的構成の候補のうちのどれが、前記位置空間中の前記選択されたオブジェクトの分布と最も類似しているかを決定することを備える、C 1 7 に記載の装置。

[C 2 1] 前記位置空間は地面である、C 1 7 に記載の装置。

10

20

30

40

50

[C 2 2] 前記生成されたメタデータをビデオストリームのフレームと関連付けるための手段と、

前記位置空間中の前記選択されたオブジェクトの各々の第 2 の位置を示す情報に基づいて、前記選択されたオブジェクトの前記複数の幾何学的構成の候補のうちの異なる 1 つを特定する第 2 のメタデータを生成するための手段と、

前記生成された第 2 のメタデータを前記ビデオストリームの第 2 のフレームと関連付けるための手段とを備える、C 1 7 に記載の装置。

[C 2 3] 前記ビデオストリーム内で、前記複数のオブジェクトをトラッキングするための手段を備える、C 1 7 に記載の装置。

[C 2 4] メタデータを生成するための装置であって、

複数のオブジェクトを選択するために少なくとも 1 つのデジタル画像を使用するように構成される判別器と、ここにおいて、前記少なくとも 1 つのデジタル画像は物理空間との関連で前記複数のオブジェクトを示し、

位置空間中の前記選択されたオブジェクトの位置を示す情報に基づいて、前記選択されたオブジェクトの複数の幾何学的構成の候補のうちの 1 つを特定するメタデータを生成するように構成される選択器と、を備える、装置。

[C 2 5] 前記位置空間中の前記複数のオブジェクトについて基準位置を決定するように構成されるオブジェクトトラッカーを備え、

前記情報は、前記決定された基準位置に対する前記選択されたオブジェクトの位置を示す、C 2 4 に記載の装置。

[C 2 6] 前記選択器は、前記複数の幾何学的構成の候補のうちのどれが、前記選択されたオブジェクトの前記位置によって表される形状と最も類似しているかを決定するように構成される、C 2 4 に記載の装置。

[C 2 7] 前記選択器は、前記複数の幾何学的構成の候補のうちのどれが、前記位置空間中の前記選択されたオブジェクトの分布と最も類似しているかを決定するように構成される、C 2 4 に記載の装置。

[C 2 8] 前記位置空間は地面である、C 2 4 に記載の装置。

[C 2 9] 前記生成されたメタデータをビデオストリームのフレームと関連付けるように構成されるビデオストリームエディタを備え、

前記選択器は、前記位置空間中の前記選択されたオブジェクトの各々の第 2 の位置を示す情報に基づいて、前記選択されたオブジェクトの前記複数の幾何学的構成の候補のうちの異なる 1 つを特定する第 2 のメタデータを生成するように構成され、

前記ビデオストリームエディタは、前記生成された第 2 のメタデータを前記ビデオストリームの第 2 のフレームと関連付けるように構成される、C 2 4 に記載の装置。

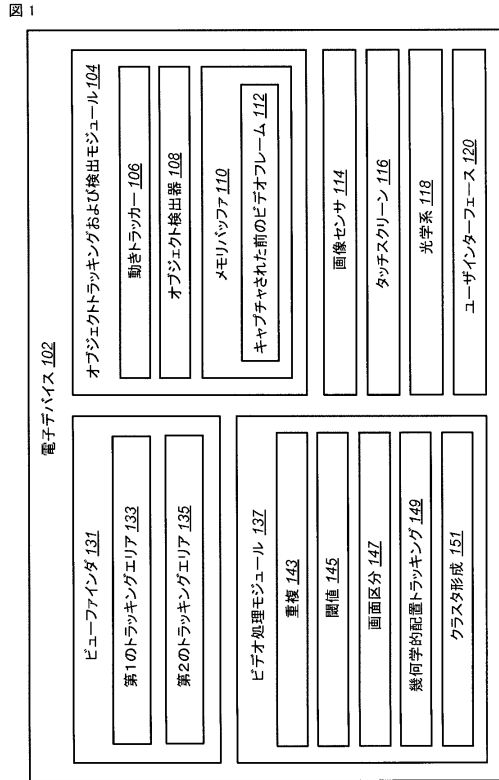
[C 3 0] ビデオストリーム内で、前記複数のオブジェクトをトラッキングするように構成されるオブジェクトトラッカーを備える、C 2 4 に記載の装置。

10

20

30

【図 1】



【図 2 A】

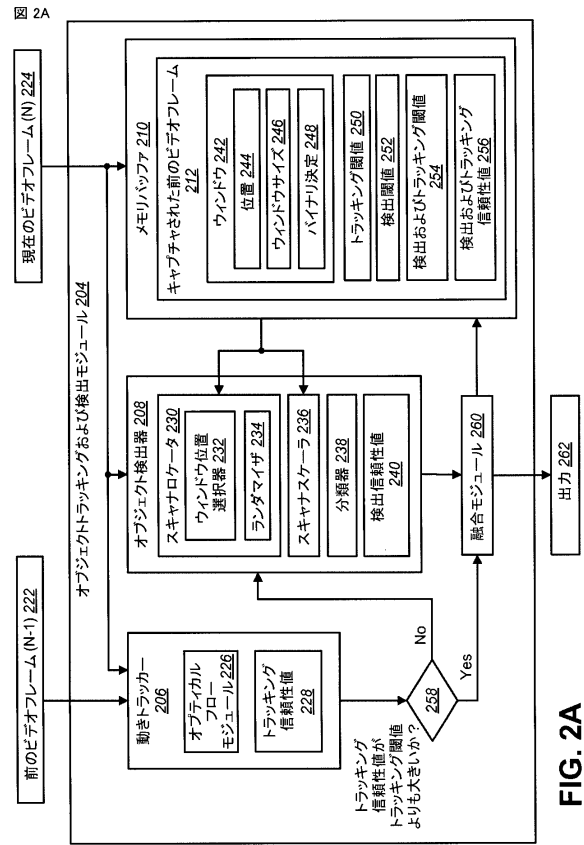


FIG. 2A

【図 2 B】

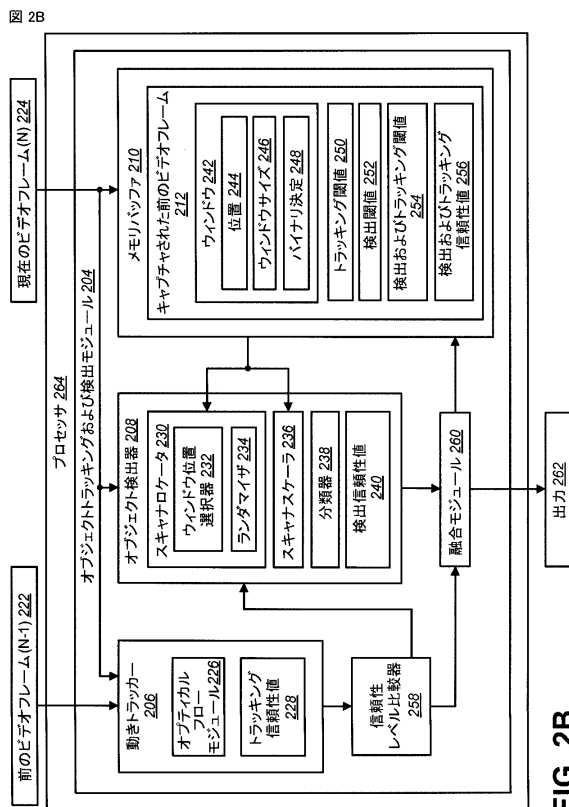


FIG. 2B

【図 3】

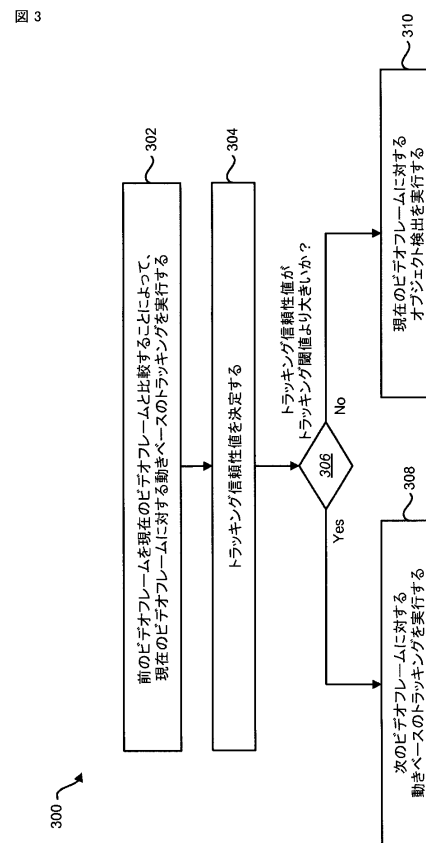


FIG. 3

【図 4】

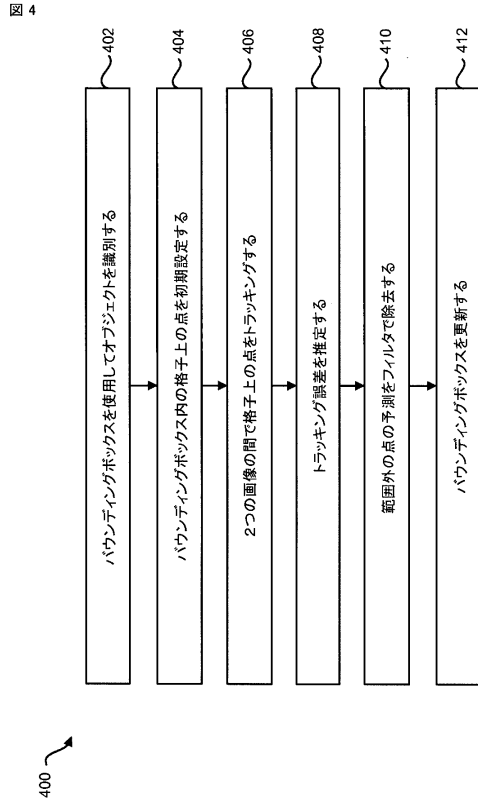


FIG. 4

【図 5】

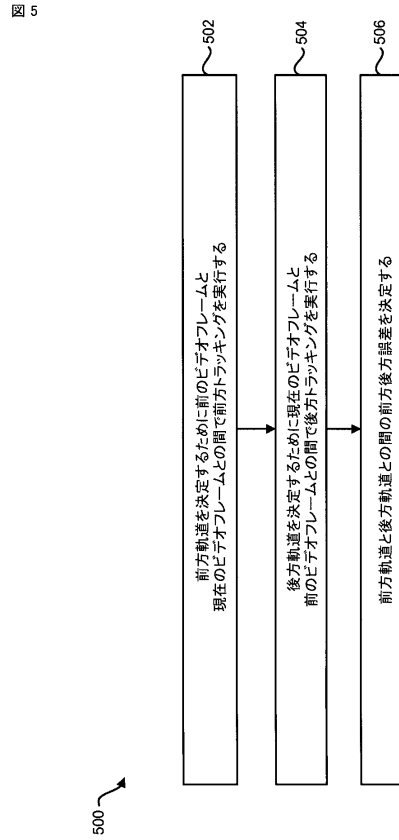


FIG. 5

【図 6】

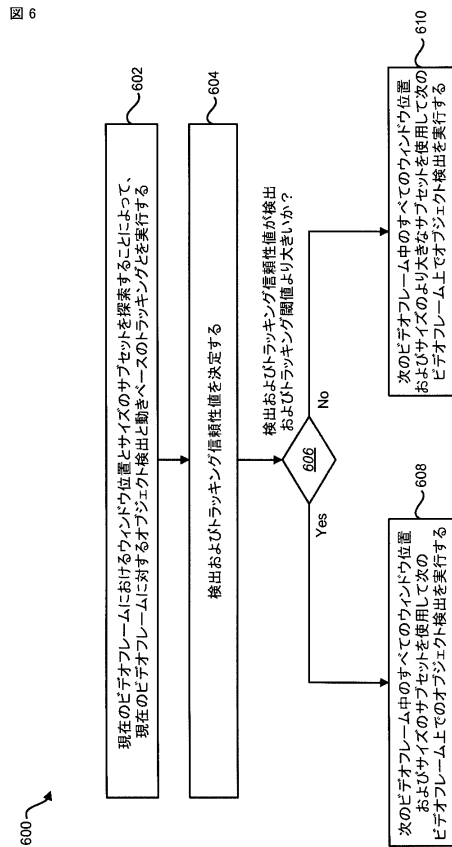


FIG. 6

【図 7】

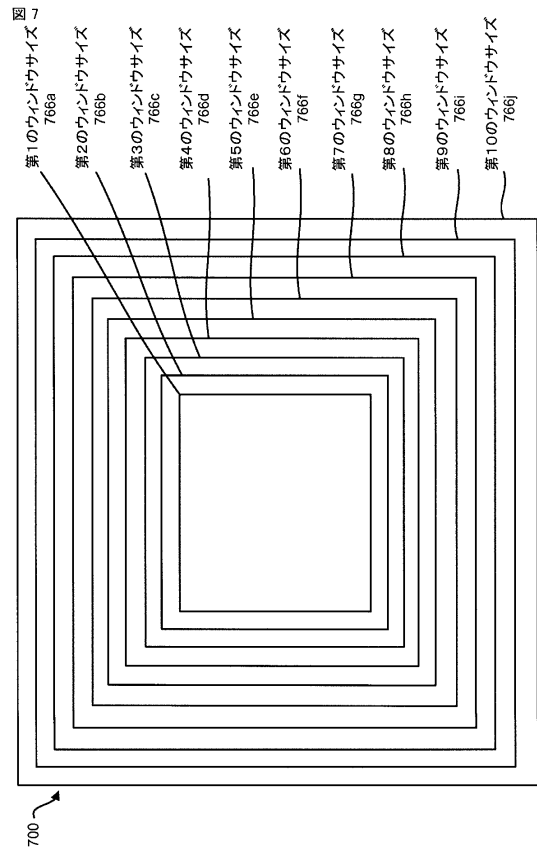


FIG. 7

【図 8】

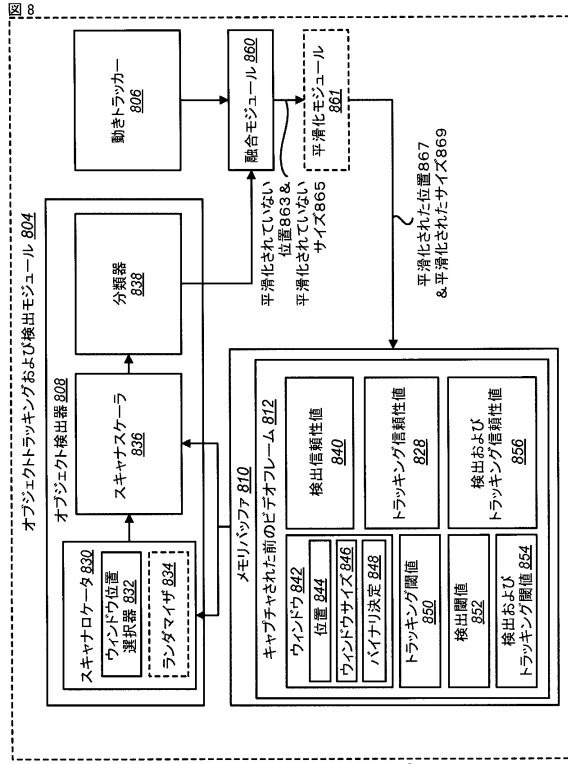


FIG. 8

【図 9】

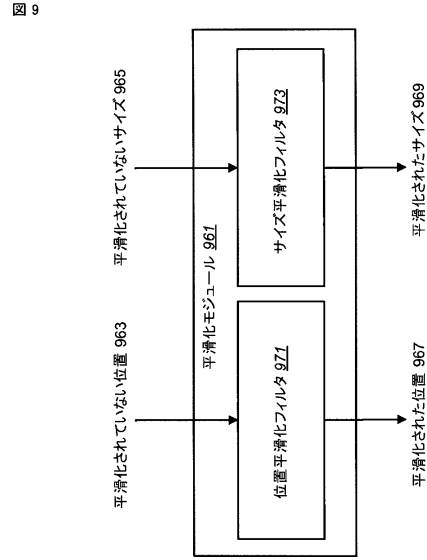


FIG. 9

【図 10】

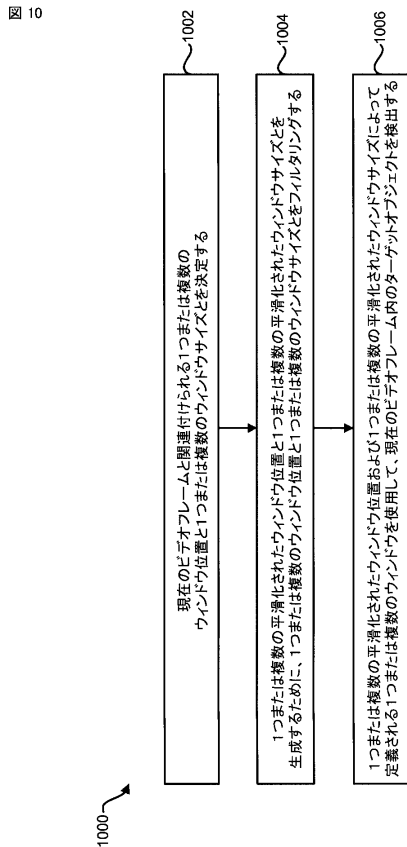


FIG. 10

【図 11 A】

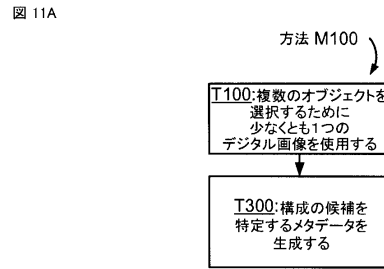


FIG. 11A

【図 11 B】

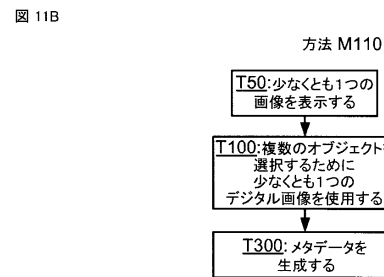


FIG. 11B

【図 11C】

図 11C

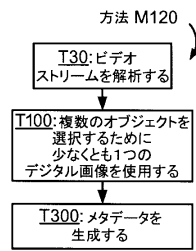


FIG. 11C

【図 11D】

図 11D

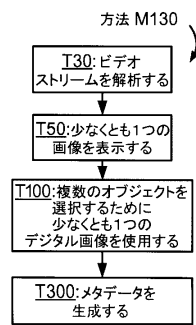


FIG. 11D

【図 12】

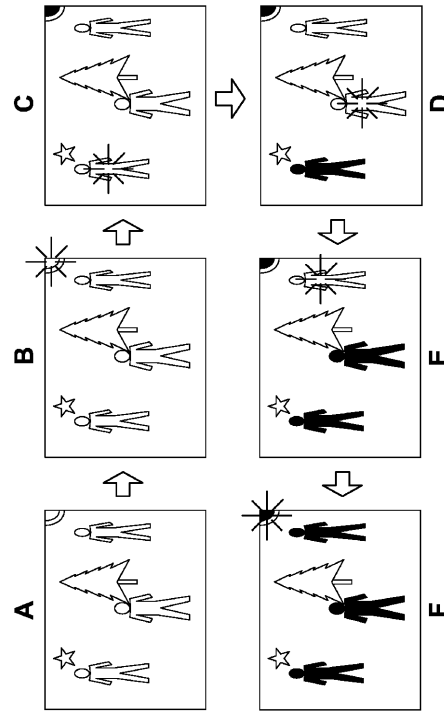


FIG. 12

【図 13】

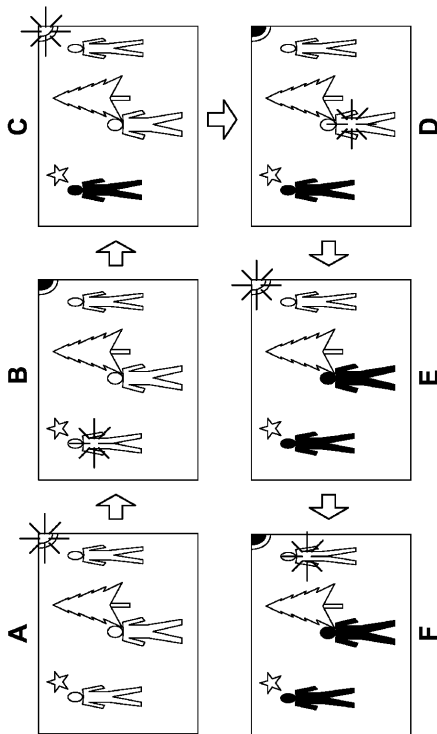


FIG. 13

【図 14】

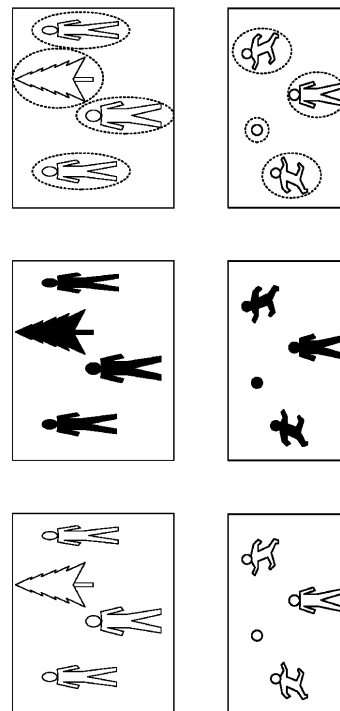
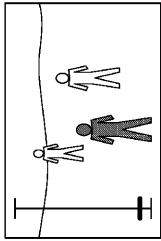
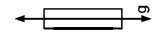


FIG. 14

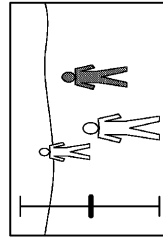
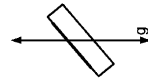
【図 15 A】

FIG. 15A



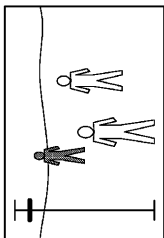
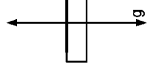
【図 15 B】

FIG. 15B

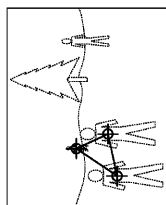
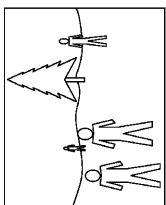


【図 15 C】

FIG. 15C

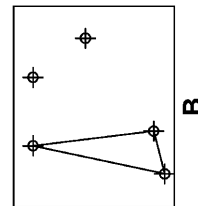
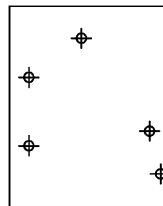


【図 16 A】



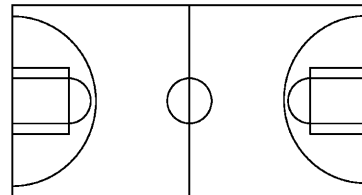
A

【図 16 B】



B

【図 16 C】

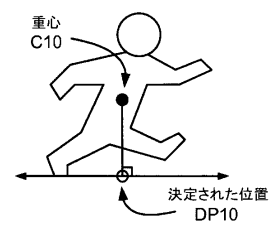


C

【図 17 A】

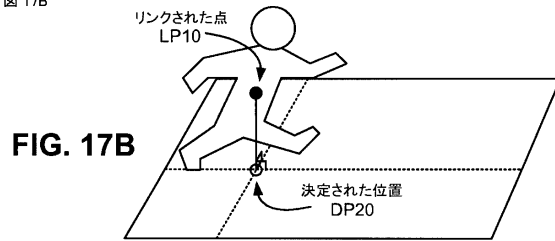
図 17A

FIG. 17A



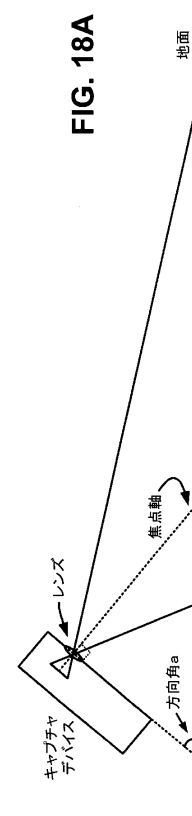
【図 17 B】

図 17B



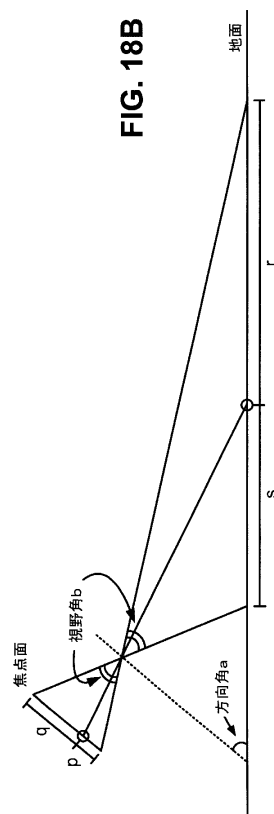
【図 18 A】

図 18A



【図 18 B】

図 18B



【図 19 A】

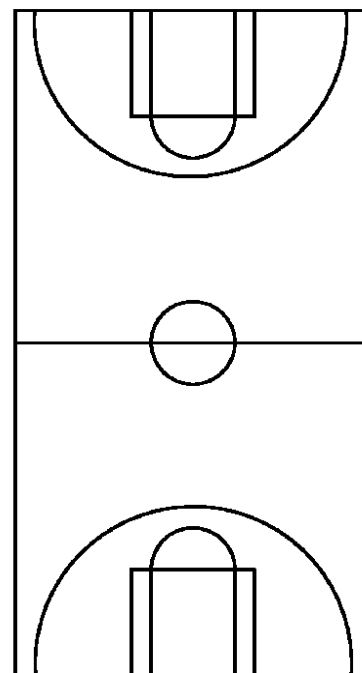


FIG. 19A

【 図 1 9 B 】

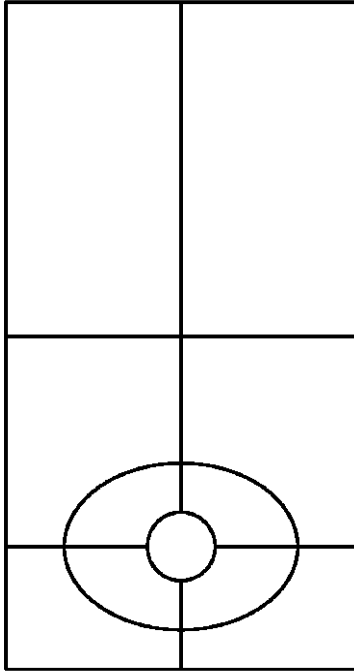


FIG. 19B

【 図 1 9 C 】

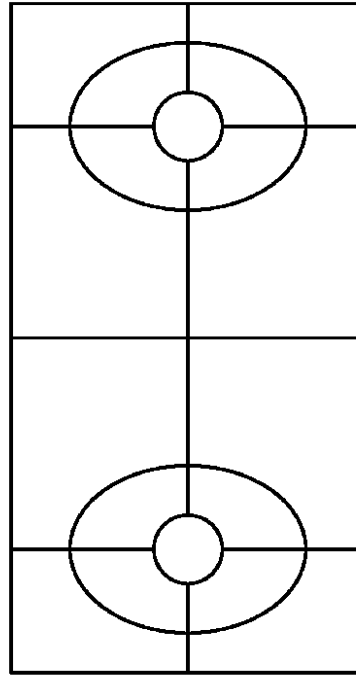


FIG. 19C

【 図 1 9 D 】

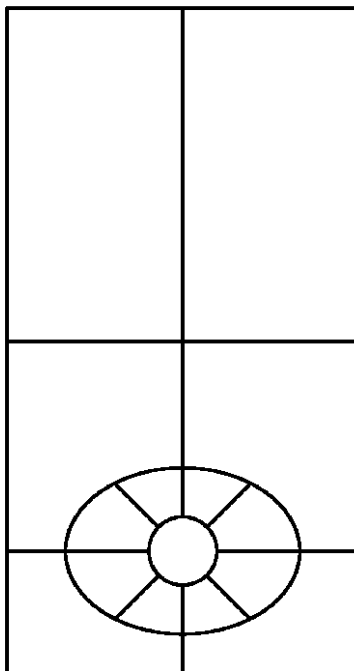


FIG. 19D

【 図 2 0 】

图 20

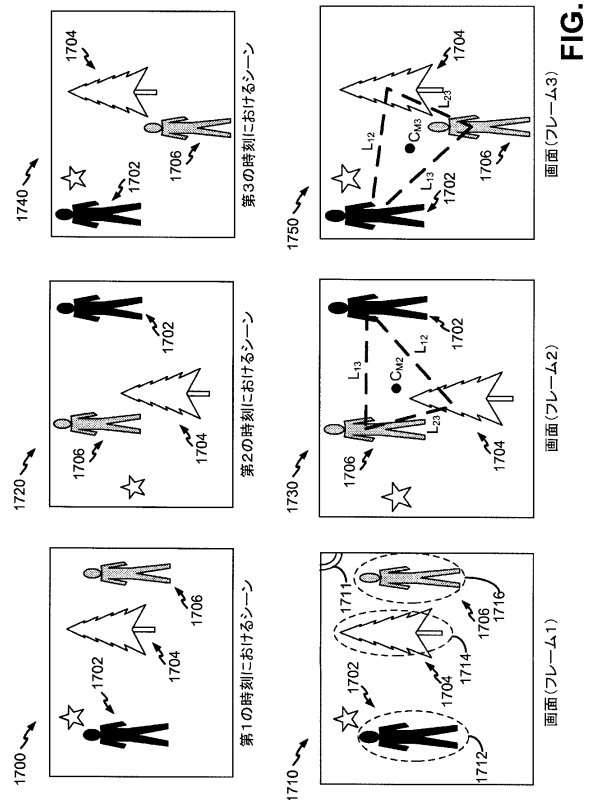


FIG. 20

【図 2 1】

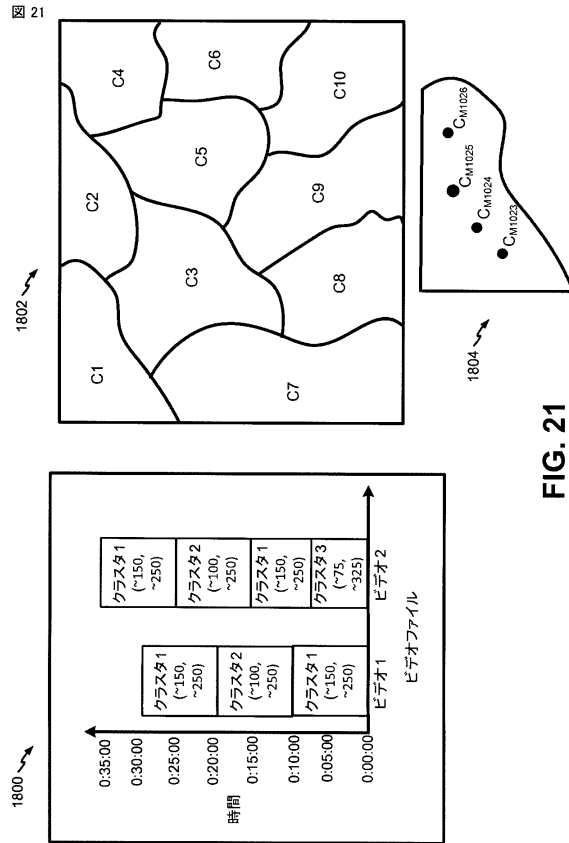


FIG. 21

【図 2 2】

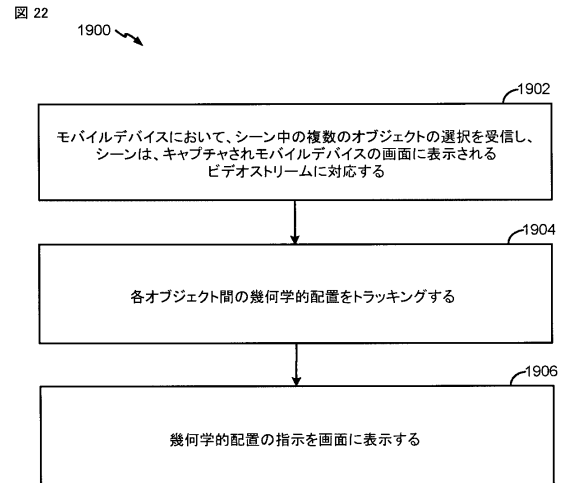


FIG. 22

【図 2 3】

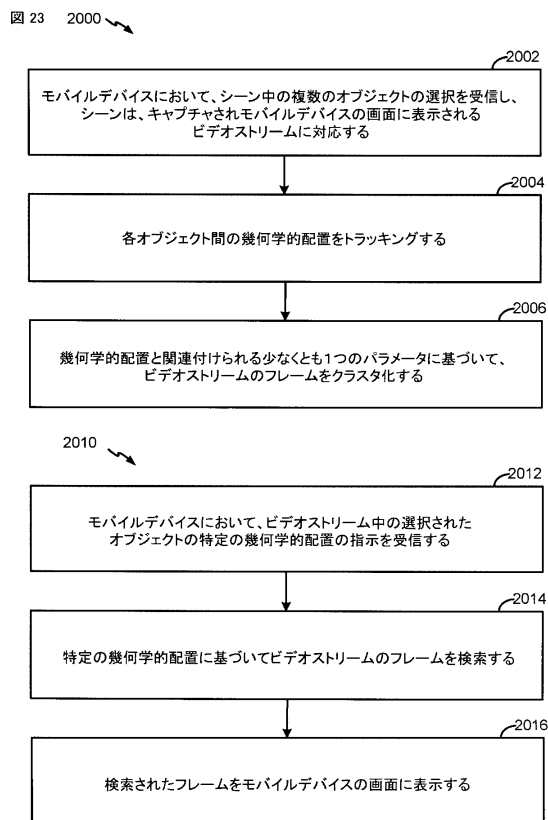


FIG. 23

【図 2 4】

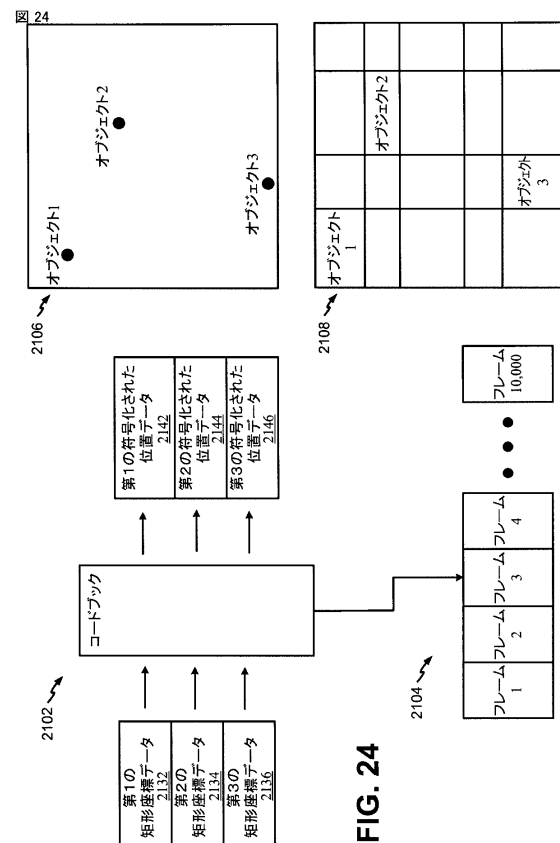


FIG. 24

【図 25】

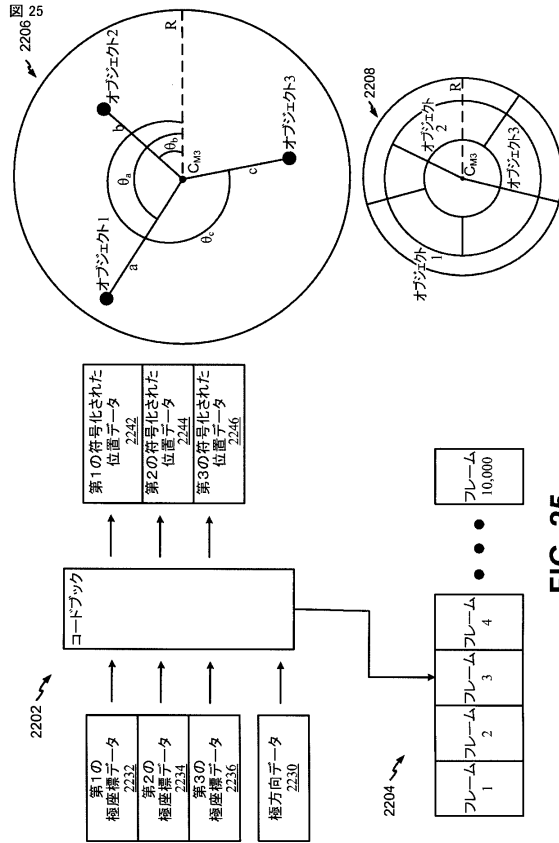


FIG. 25

【図 26】

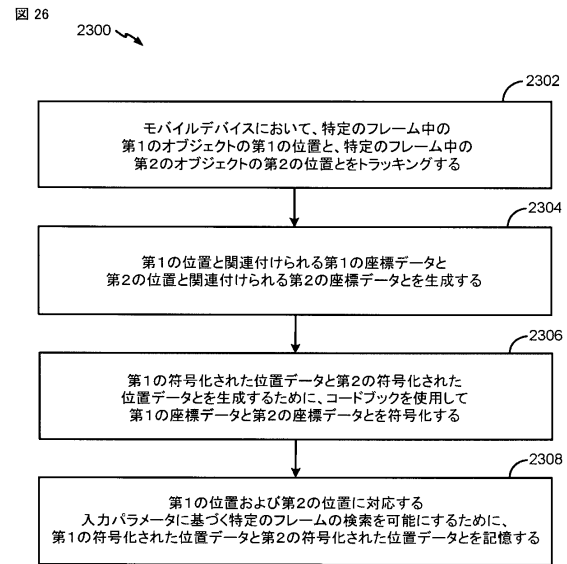


FIG. 26

【図 27】

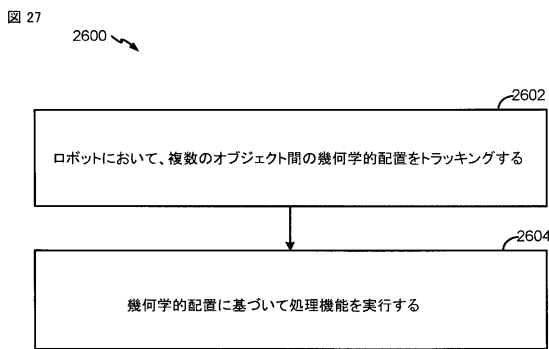


FIG. 27

【図 28】

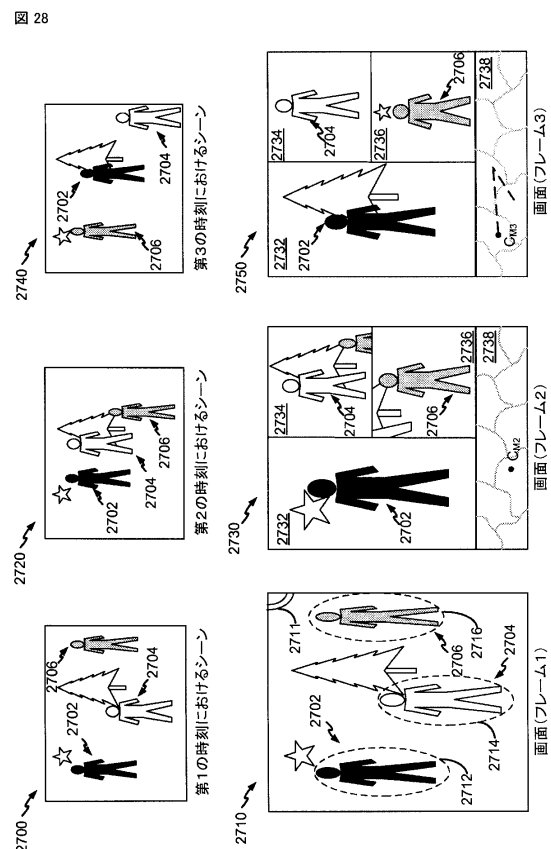


FIG. 28

【図 29】

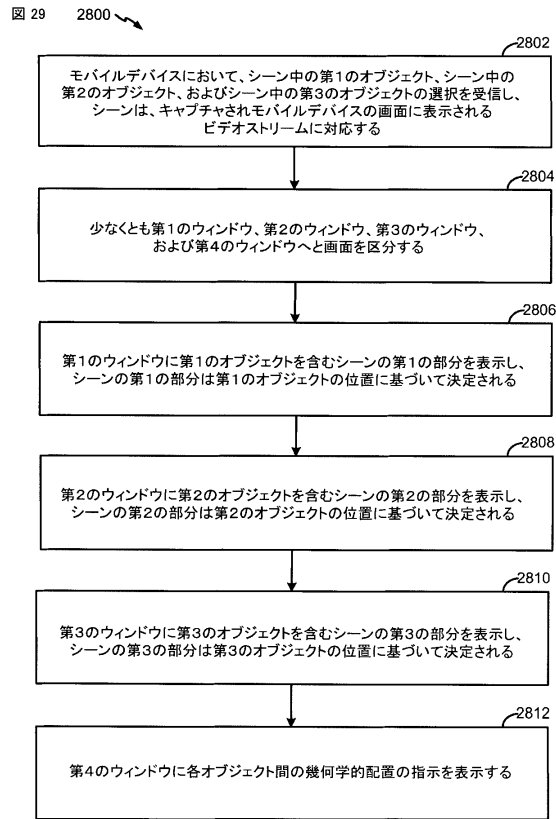


FIG. 29

【図 30 A】

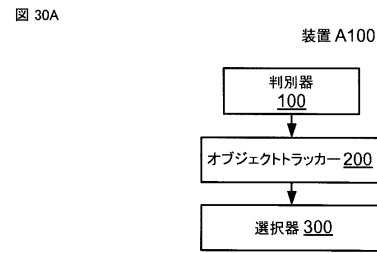


FIG. 30A

【図 30 B】

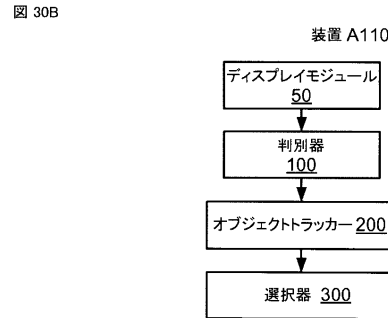


FIG. 30B

【図 30 C】

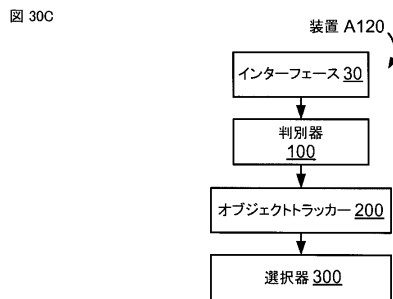


FIG. 30C

【図 30 D】

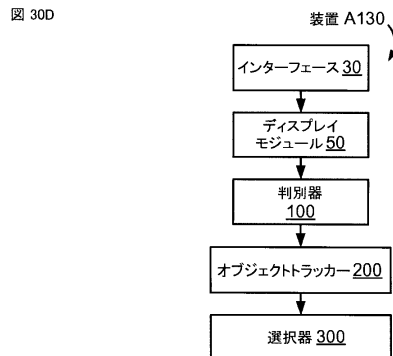


FIG. 30D

【図 31】

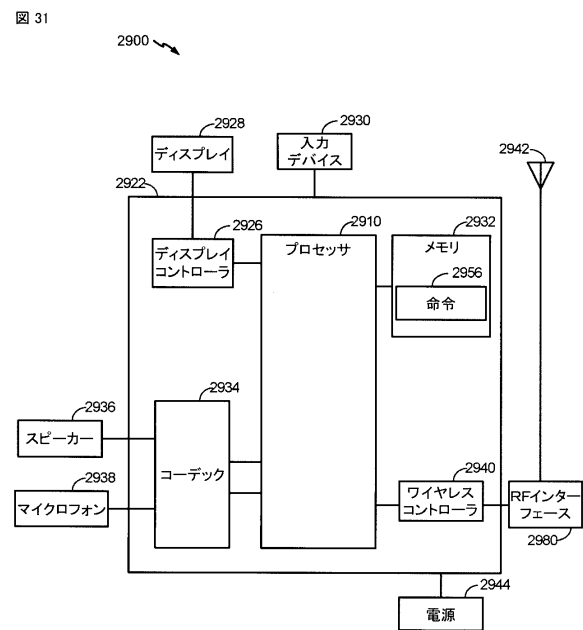


FIG. 31

【図 3 2 A】

図 32A

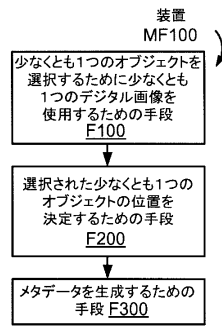


FIG. 32A

【図 3 2 B】

図 32B

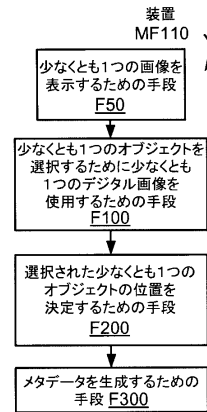


FIG. 32B

【図 3 2 C】

図 32C

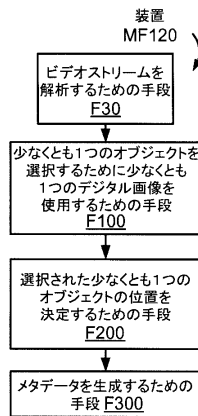


FIG. 32C

【図 3 2 D】

図 32D

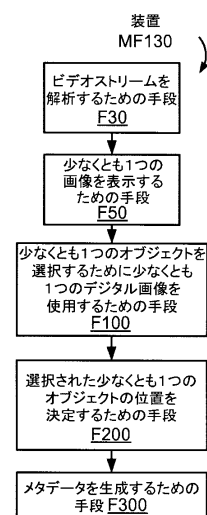


FIG. 32D

【図 3 3 A】

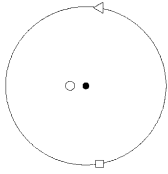
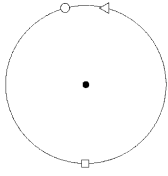
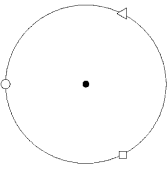


FIG. 33A



【図 3 3 B】

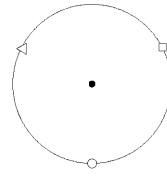
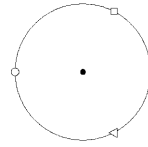


FIG. 33B



【図 3 4 A】

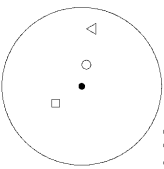
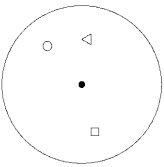
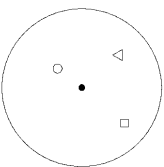
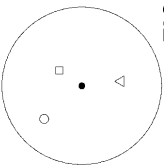


FIG. 34A



【図 3 4 B】

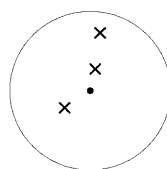
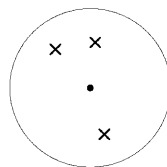
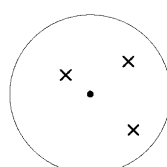
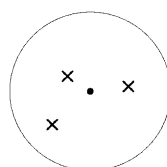
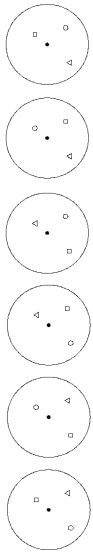


FIG. 34B



【図 34 C】

FIG. 34C



【図 36】

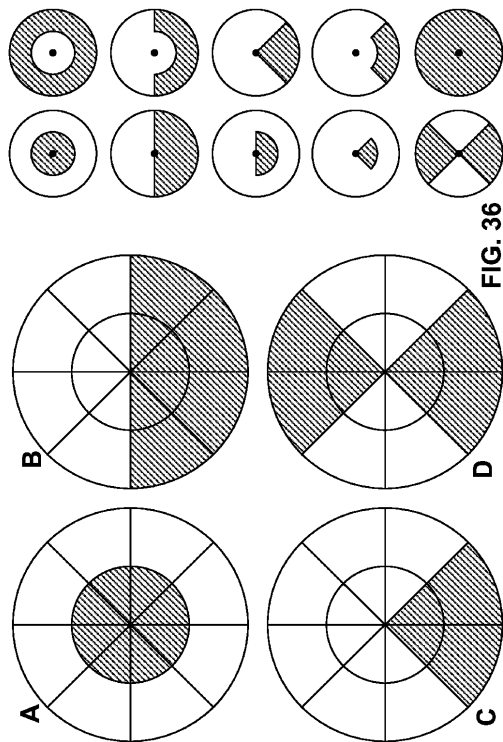


FIG. 36

【図 35】

図 35

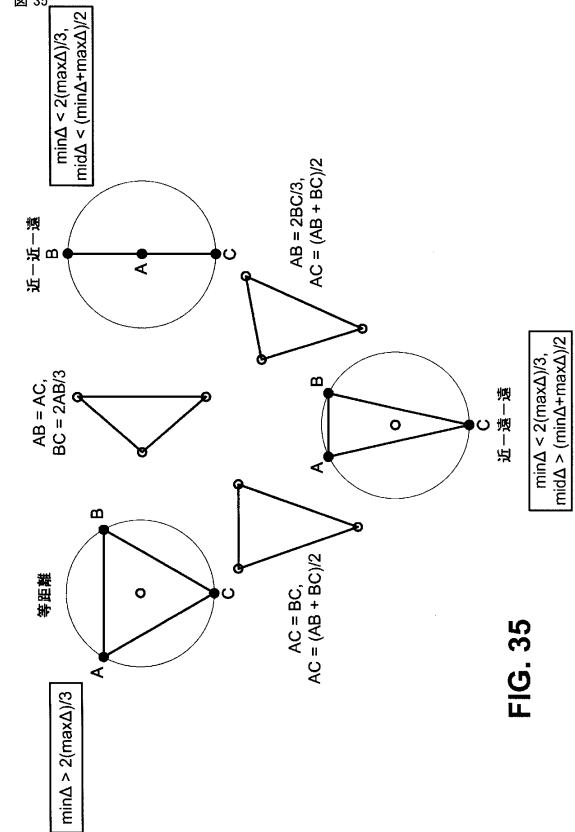


FIG. 35

【図 37 A】

図 37A

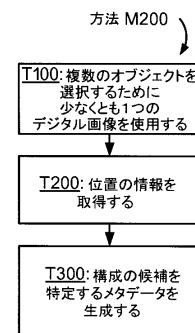


FIG. 37A

【図 37B】

図 37B

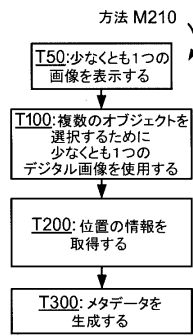


FIG. 37B

【図 37C】

図 37C

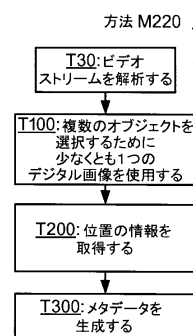


FIG. 37C

【図 37D】

図 37D

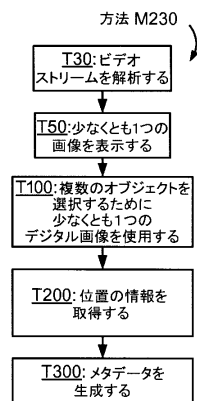


FIG. 37D

【図 38B】

図 38B

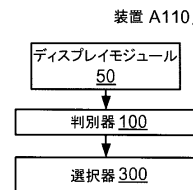


FIG. 38B

【図 38C】

図 38C

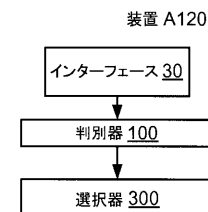


FIG. 38C

【図 38A】

図 38A

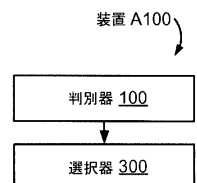


FIG. 38A

【図 38 D】

図 38D

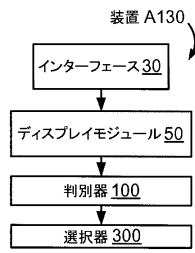


FIG. 38D

【図 38 F】

図 38F

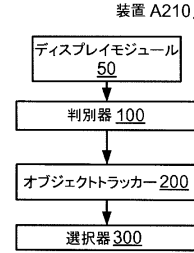


FIG. 38F

【図 38 E】

図 38E

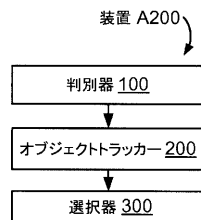


FIG. 38E

【図 38 G】

図 38G

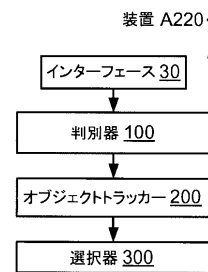


FIG. 38G

【図 39 A】

図 39A

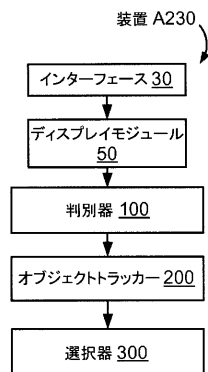


FIG. 39A

【図 39 C】

図 39C

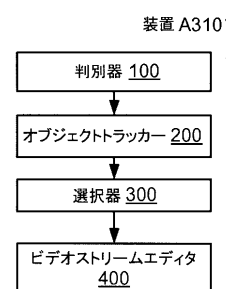


FIG. 39C

【図 39 B】

図 39B

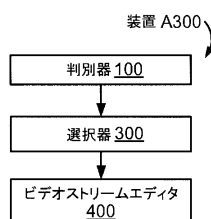


FIG. 39B

【図 39 D】

図 39D

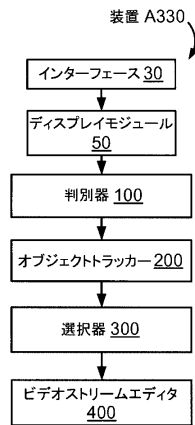


FIG. 39D

【図 40 A】

図 40A

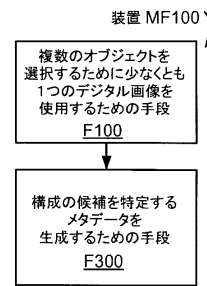


FIG. 40A

【図 40 B】

図 40B

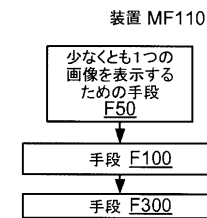


FIG. 40B

【図 40 C】

図 40C

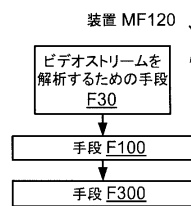


FIG. 40C

【図 40 E】

図 40E

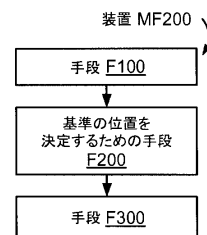


FIG. 40E

【図 40 D】

図 40D

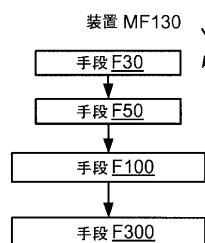


FIG. 40D

【図 40 F】

図 40F

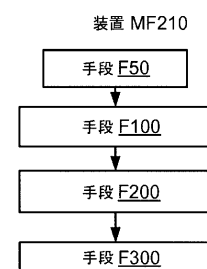


FIG. 40F

【図 40 G】

図 40G

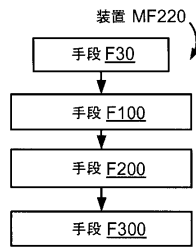


FIG. 40G

【図 41 A】

図 41A

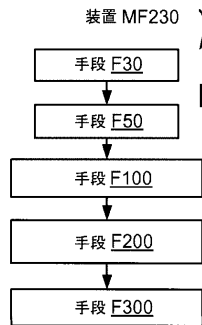


FIG. 41A

【図 41 D】

図 41D

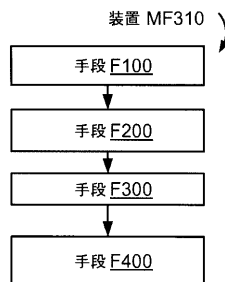


FIG. 41D

【図 41 B】

図 41B

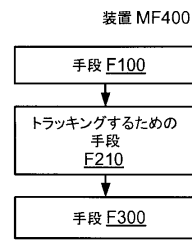


FIG. 41B

【図 41 C】

図 41C

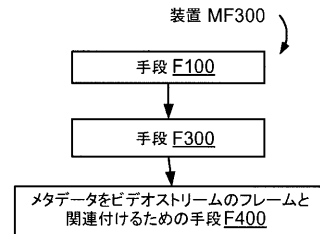


FIG. 41C

【図 41 E】

図 41E

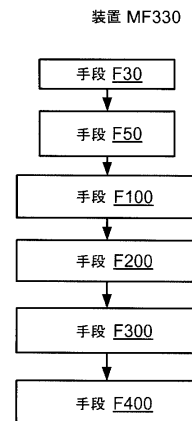


FIG. 41E

【図 4 2】

図 42

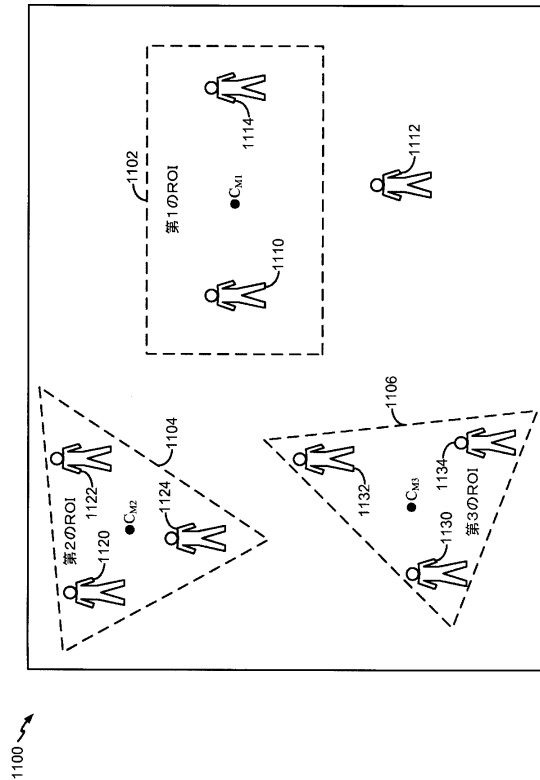


FIG. 42

【図 4 3】

図 43

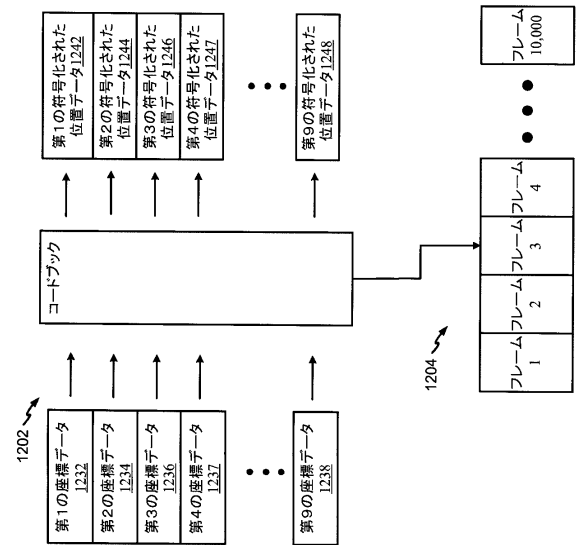


FIG. 43

【図 4 4】

図 44

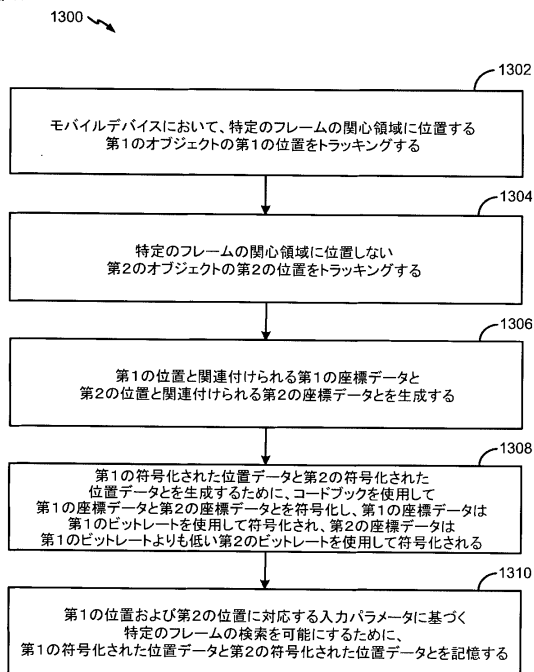


FIG. 44

フロントページの続き

- (31)優先権主張番号 62/063,498
(32)優先日 平成26年10月14日(2014.10.14)
(33)優先権主張国 米国(US)
(31)優先権主張番号 14/575,986
(32)優先日 平成26年12月18日(2014.12.18)
(33)優先権主張国 米国(US)

早期審査対象出願

- (72)発明者 ガオ、ダシャン
アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5
(72)発明者 ジョン、シン
アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5
(72)発明者 カンドハダイ、アナンサパドマナブハン・アラサニパライ
アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5
(72)発明者 レイバー、スティーブン・ダグラス
アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5

審査官 田中 秀樹

- (56)参考文献 米国特許出願公開第2011/0305394(US,A1)
米国特許出願公開第2009/0144772(US,A1)

- (58)調査した分野(Int.Cl., DB名)
G 0 6 F 1 7 / 3 0
G 0 6 T 7 / 0 0 - 7 / 6 0