

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6363204号  
(P6363204)

(45) 発行日 平成30年7月25日(2018.7.25)

(24) 登録日 平成30年7月6日(2018.7.6)

(51) Int.Cl.

F 1

HO4N 5/232 (2006.01)  
GO6T 7/20 (2017.01)HO4N 5/232 290  
GO6T 7/20 300  
HO4N 5/232 945  
HO4N 5/232 190

請求項の数 29 (全 27 頁)

(21) 出願番号 特願2016-541312 (P2016-541312)  
 (86) (22) 出願日 平成26年12月12日 (2014.12.12)  
 (65) 公表番号 特表2017-502589 (P2017-502589A)  
 (43) 公表日 平成29年1月19日 (2017.1.19)  
 (86) 國際出願番号 PCT/US2014/070078  
 (87) 國際公開番号 WO2015/094977  
 (87) 國際公開日 平成27年6月25日 (2015.6.25)  
 審査請求日 平成29年11月15日 (2017.11.15)  
 (31) 優先権主張番号 61/919,754  
 (32) 優先日 平成25年12月21日 (2013.12.21)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 14/567,119  
 (32) 優先日 平成26年12月11日 (2014.12.11)  
 (33) 優先権主張国 米国(US)

早期審査対象出願

(73) 特許権者 595020643  
 クアアルコム・インコーポレイテッド  
 QUALCOMM INCORPORATED  
 アメリカ合衆国、カリフォルニア州 92  
 121-1714、サン・ディエゴ、モア  
 ハウス・ドライブ 5775  
 (74) 代理人 100108855  
 弁理士 蔵田 昌俊  
 (74) 代理人 100109830  
 弁理士 福原 淑弘  
 (74) 代理人 100158805  
 弁理士 井関 守三  
 (74) 代理人 100112807  
 弁理士 岡田 貴志

最終頁に続く

(54) 【発明の名称】対象トラッキングボックスの表示を安定させるためのシステムおよび方法

## (57) 【特許請求の範囲】

## 【請求項 1】

方法であって、

画像のシーケンスの第1の画像のための第1の境界ボックスを定義する第1のデータを受信することと、前記第1の境界ボックスは、関心領域に対応し、トラックされた対象を含み、

画像の前記シーケンスの第2の画像に関する対象トラッキングデータを受信することと、前記対象トラッキングデータは、第2の境界ボックスを定義し、前記第2の境界ボックスは、前記関心領域に対応し、前記第2の画像の中の前記トラックされた対象を含み、

前記第2の画像に関する前記対象トラッキングデータを受信することおよび前記第2の境界ボックスを定義することに基づいて、前記第2の画像のための検索領域を前記第2の境界ボックスに基づいて決定することと、前記検索領域は、前記第2の画像の一部分に対応し、前記第2の境界ボックス内の複数のピクセルと、前記第2の境界ボックスに隣接する複数のピクセルとを含み、

前記第2の画像のための前記検索領域を決定することに基づいて、複数の類似のメトリックを前記検索領域に対応する複数の検索境界ボックスに基づいて決定することと、ここにおいて、前記検索領域は、前記複数の検索境界ボックスを含み、前記複数の類似のメトリックの各類似のメトリックは、前記第1の境界ボックス内の複数の第1のピクセル、および前記類似のメトリックに対応する前記複数の検索境界ボックスのうちの1つの検索境界ボックス内の複数の検索ピクセルに基づいて決定され、前記複数の検索境界ボックスの

10

20

各々の検索座標は、1つ以上の方にシフトされる前記第2の境界ボックスの第2の座標に基づき、前記複数の類似のメトリックを決定することは、前記複数の第1のピクセルの列和ベクトルを決定すること、ここにおいて、前記列和ベクトルの特定の要素は、前記第1の境界ボックスの特定の列に対応するピクセル値の和に対応する、をさらに備え、

前記複数の類似のメトリックに基づいて、修正された第2の境界ボックスを決定すること、

を備える、方法。

#### 【請求項2】

前記第1のデータは、前記第1の境界ボックスの第1の座標および第1のディメンションを含み、前記対象トラッキングデータは、前記第2の境界ボックスの前記第2の座標および第2のディメンションを含む、請求項1に記載の方法。 10

#### 【請求項3】

前記複数の検索境界ボックスの各々の検索ディメンションは、前記第1の境界ボックスの前記第1のディメンションに対応し、前記複数の検索境界ボックスは、前記第2の境界ボックスを含む、請求項2に記載の方法。

#### 【請求項4】

前記複数の類似のメトリックは、差分絶対値和(S A D)メトリックを含み、前記S A Dメトリックは、前記複数の第1のピクセルの第1の値、および前記複数の検索ピクセルの第2の値に基づき、前記第1の値および前記第2の値は、対応する第1のピクセルまたは対応する検索ピクセルについて、ピクセル輝度、ピクセルの色サブコンポーネント、またはそれらの組み合わせを識別する、請求項1に記載の方法。 20

#### 【請求項5】

前記複数の類似のメトリックのうちの特定の類似のメトリックに基づいて、前記修正された第2の境界ボックスである前記複数の検索境界ボックスの特定の検索境界ボックスを選択することをさらに備え、前記特定の類似のメトリックは、前記特定の検索境界ボックスに対応する、請求項1に記載の方法。

#### 【請求項6】

前記複数の第1のピクセルに対応する第1のピクセル特性、および前記特定の検索境界ボックス内の複数の特定の検索ピクセルに対応する第2のピクセル特性の差分絶対値和(S A D)に少なくとも部分的にに基づいて、前記特定の類似のメトリックを計算することをさらに備える、請求項5に記載の方法。 30

#### 【請求項7】

前記複数の第1のピクセルの第1の列和ベクトル、および前記特定の検索境界ボックス内の複数の特定の検索ピクセルの第2の列和ベクトルの第1の差分絶対値和(S A D)と、

前記複数の第1のピクセルの第1の列和差分ベクトル、および前記複数の特定の検索ピクセルの第2の列和差分ベクトルの第2のS A Dと、

前記複数の第1のピクセルの第1の行和ベクトル、および前記複数の特定の検索ピクセルの第2の行和ベクトルの第3のS A Dと、

前記複数の第1のピクセルの第1の行和差分ベクトル、および前記複数の特定の検索ピクセルの第2の行和差分ベクトルの第4のS A Dと、 40

を加算することによって前記特定の類似のメトリックを計算することをさらに備える、請求項5に記載の方法。

#### 【請求項8】

前記複数の類似のメトリックを決定することは、

前記複数の第1のピクセルの列和差分ベクトルを決定すること、ここにおいて、前記列和差分ベクトルの特定の要素は、前記列和ベクトルの第1の要素と、前記列和ベクトルの第2の要素間の差分に対応し、前記第1の要素および前記第2の要素は、前記第1の境界ボックスの近接した列に対応する、

をさらに備える、請求項1に記載の方法。 50

**【請求項 9】**

前記複数の類似のメトリックを決定することは、

前記複数の第1のピクセルの行和ベクトルを決定すること、ここにおいて、前記行和ベクトルの特定の要素は、前記第1の境界ボックスの特定の行に対応するピクセル値の和に対応する、

をさらに備える、請求項1に記載の方法。

**【請求項 10】**

前記複数の類似のメトリックを決定することは、

前記複数の第1のピクセルの行和差分ベクトルを決定すること、ここにおいて、前記行和差分ベクトルの特定の要素は、前記行和ベクトルの第1の要素と、前記行和ベクトルの第2の要素間の差分に対応し、前記第1の要素および前記第2の要素は、前記第1の境界ボックスの近接した行に対応する、

10

をさらに備える、請求項9に記載の方法。

**【請求項 11】**

複数の画像に対応するメディアンディメンションを決定すること、ここにおいて、前記複数の画像は、画像の前記シーケンスの中の前記第2の画像に先行し、前記修正された第2の境界ボックスの第2のディメンションは、前記メディアンディメンションに対応する、

をさらに備える、請求項1に記載の方法。

**【請求項 12】**

20

装置であって、

命令を記憶するように構成されたメモリと、

少なくとも第1の画像および第2の画像を含む画像のシーケンス中の前記第2の画像に関する対象トラッキングデータの受信に基づいて、ここで、前記対象トラッキングデータは、第2の境界ボックスを定義し、

前記第2の画像のための検索領域を前記第2の境界ボックスに基づいて決定することと、前記検索領域は、前記第2の画像の一部分に対応し、前記第2の境界ボックス内の複数のピクセルと、前記第2の境界ボックスに隣接する複数のピクセルとを含み、

複数の類似のメトリックを前記検索領域に対応する複数の検索境界ボックスに基づいて決定することと、ここにおいて、前記検索領域は、前記複数の検索境界ボックスを含み、前記複数の類似のメトリックの各類似のメトリックは、前記第1の画像の第1の境界ボックス内の複数の第1のピクセル、および前記類似のメトリックに対応する複数の検索境界ボックスのうちの1つの検索境界ボックス内の複数の検索ピクセルに基づいて決定され、前記第1の境界ボックスは、関心領域に対応し、トラックされた対象を含み、前記複数の検索境界ボックスの各々の検索座標は、1つ以上の方向にシフトされる前記第2の境界ボックスの第2の座標に基づき、前記第1の画像は、画像の前記シーケンスの中の前記第2の画像に先行し、前記第2の境界ボックスは、前記関心領域に対応し、前記第2の画像の中の前記トラックされた対象を含み、前記複数の類似のメトリックを決定することは、前記複数の第1のピクセルの列和ベクトルを決定すること、ここにおいて、前記列和ベクトルの特定の要素は、前記第1の境界ボックスの特定の列に対応するピクセル値の和に対応する、をさらに備え、

30

前記複数の類似のメトリックに基づいて、修正された第2の境界ボックスを決定することと、

を行うための前記命令を実行するように構成されたプロセッサと、

前記修正された第2の境界ボックスをディスプレイするように構成されたディスプレイデバイスと

を備える、装置。

**【請求項 13】**

前記プロセッサは、前記複数の類似のメトリックのうちの特定の類似のメトリックに基づいて、前記修正された第2の境界ボックスである前記複数の検索境界ボックスの特定の

40

50

検索境界ボックスを選択するようにさらに構成される、請求項1\_2に記載の装置。

**【請求項 1 4】**

前記プロセッサは、前記複数の第1のピクセルの第1のピクセル特性、および前記特定の検索境界ボックスの複数の特定の検索ピクセルの第2のピクセル特性の差分絶対値和(S A D)に少なくとも部分的に基づいて、前記特定の類似のメトリックを計算するようにさらに構成される、請求項1\_3に記載の装置。

**【請求項 1 5】**

前記プロセッサは、前記複数の第1のピクセルの第1のピクセル輝度、および前記特定の検索境界ボックスの複数の特定の検索ピクセルの第2のピクセル輝度に少なくとも部分的に基づいて、前記特定の類似のメトリックを計算するようにさらに構成される、請求項1\_3に記載の装置。10

**【請求項 1 6】**

画像の前記シーケンスは、ビデオストリームに対応する、請求項1\_2に記載の装置。

**【請求項 1 7】**

前記プロセッサは、前記第1の画像のための第1のデータを取得するように、および前記第2の画像に関する対象トラッキングデータを取得するようにさらに構成され、

前記第1のデータは、前記第1の境界ボックスを定義し、

前記対象トラッキングデータは、前記第2の境界ボックスを定義し、

前記第1のデータは、前記第1の境界ボックスの第1の座標および第1のディメンションを含み、20

前記対象トラッキングデータは、前記第2の境界ボックスの前記第2の座標および第2のディメンションを含み、

前記複数の検索境界ボックスは、前記第2の境界ボックスを含む、請求項1\_2に記載の装置。

**【請求項 1 8】**

前記第1の画像および前記第2の画像を生成するように構成されたカメラをさらに備える、請求項1\_2に記載の装置。

**【請求項 1 9】**

前記複数の検索境界ボックスの各々の検索ディメンションは、前記第1の境界ボックスの第1のディメンションに対応する、請求項1\_2に記載の装置。30

**【請求項 2 0】**

前記プロセッサは、複数の画像に対応するメディアンディメンションを決定すること、ここにおいて、前記複数の画像は、画像の前記シーケンスの中の前記第2の画像に先行し、前記修正された第2の境界ボックスの第2のディメンションは、前記メディアンディメンションに対応する、

を行うようにさらに構成される、請求項1\_2に記載の装置。

**【請求項 2 1】**

カメラディスプレイを備えるカメラをさらに備え、前記プロセッサは、

画像の前記シーケンスの中の前記第2の画像に先行する前記複数の画像の数が、しきい値を満たすと決定することに応答して、前記決定されたメディアンディメンションに基づいて前記修正された第2の境界ボックスの前記第2のディメンションを定義する修正された第2の境界ボックスデータを決定することと、40

前記修正された第2の境界ボックスデータを前記カメラディスプレイに送ることと、ここにおいて、前記カメラディスプレイは、前記修正された第2の境界ボックスを用いて前記第2の画像をディスプレイする、

を行うようにさらに構成される、請求項2\_0に記載の装置。

**【請求項 2 2】**

命令を記憶するコンピュータ可読記憶デバイスであって、前記命令がプロセッサによって実行されるとき、前記プロセッサに、

少なくとも第1の画像および第2の画像を含む画像のシーケンス中の前記第2の画像に50

関する対象トラッキングデータの受信に基づいて、ここで、前記対象トラッキングデータは、第2の境界ボックスを定義し、

前記第2の画像のための検索領域を前記第2の境界ボックスに基づいて決定することと、前記検索領域は、前記第2の画像の一部分に対応し、前記第2の境界ボックス内の複数のピクセルと、前記第2の境界ボックスに隣接する複数のピクセルとを含み、

複数の類似のメトリックを前記検索領域に対応する複数の検索境界ボックスに基づいて決定することと、ここにおいて、前記検索領域は、前記複数の検索境界ボックスを含み、前記複数の類似のメトリックの各類似のメトリックは、前記第1の画像の第1の境界ボックス内の複数の第1のピクセル、および前記類似のメトリックに対応する複数の検索境界ボックスのうちの1つの検索境界ボックス内の複数の検索ピクセルに基づいて決定され、前記第1の境界ボックスは、関心領域に対応し、トラックされた対象を含み、前記複数の検索境界ボックスの各々の検索座標は、1つ以上の方に向にシフトされる前記第2の境界ボックスの第2の座標に基づき、前記第1の画像は、画像の前記シーケンスの中の前記第2の画像に先行し、前記第2の境界ボックスは、前記関心領域に対応し、前記第2の画像の中の前記トラックされた対象を含み、前記複数の類似のメトリックを決定することは、前記複数の第1のピクセルの列和ベクトルを決定すること、ここにおいて、前記列和ベクトルの特定の要素は、前記第1の境界ボックスの特定の列に対応するピクセル値の和に対応する、をさらに備え、

前記複数の類似のメトリックに基づいて、修正された第2の境界ボックスを決定することと、

を備えるオペレーションを行わせる、命令を記憶したコンピュータ可読記憶デバイス。  
【請求項23】

前記オペレーションは、複数の画像に対応するメディアンディメンションを決定することと、ここにおいて、前記複数の画像は、画像の前記シーケンスの中の前記第2の画像に先行し、前記修正された第2の境界ボックスの第2のディメンションは、前記メディアンディメンションに対応する、

をさらに備える、請求項22に記載のコンピュータ可読記憶デバイス。

【請求項24】

前記オペレーションは、  
前記第1の画像のための第1のデータを受信することと、ここにおいて、前記第1のデータは、前記第1の境界ボックスを定義し、

前記第2の画像に関する対象トラッキングデータを受信することと、ここにおいて、前記対象トラッキングデータは、前記第2の境界ボックスを定義する、

をさらに備える、請求項22に記載のコンピュータ可読記憶デバイス。

【請求項25】

前記第1のデータは、前記第1の境界ボックスの第1の座標および第1のディメンションを含み、前記対象トラッキングデータは、前記第2の境界ボックスの前記第2の座標および第2のディメンションを含む、請求項24に記載のコンピュータ可読記憶デバイス。

【請求項26】

前記複数の検索境界ボックスの各々の検索ディメンションは、前記第1の境界ボックスの第1のディメンションに対応し、前記複数の検索境界ボックスは、前記第2の境界ボックスを含む、請求項22に記載のコンピュータ可読記憶デバイス。

【請求項27】

前記オペレーションは、  
特定の検索境界ボックスに対応する前記複数の類似のメトリックのうちの特定の類似のメトリックに基づいて、前記修正された第2の境界ボックスである前記複数の検索境界ボックスの前記特定の検索境界ボックスを選択することをさらに備える、請求項22に記載のコンピュータ可読記憶デバイス。

【請求項28】

装置であって、

画像のシーケンスの第1の画像のための第1の境界ボックスを定義する第1のデータを受信するための手段と、ここにおいて、前記第1の境界ボックスは、関心領域に対応し、トラックされた対象を含み、

画像の前記シーケンスの第2の画像に関する対象トラッキングデータを生成するための手段と、前記対象トラッキングデータは、第2の境界ボックスを定義し、ここにおいて、前記第2の境界ボックスは、前記関心領域に対応し、前記第2の画像の中の前記トラックされた対象を含み、

前記第2の画像に関する前記対象トラッキングデータを受信することに基づいて、

前記第2の境界ボックスに基づく前記第2の画像のための検索領域と、前記検索領域は、前記第2の画像の一部分に対応し、前記第2の境界ボックス内の複数のピクセルと、  
前記第2の境界ボックスに隣接する複数のピクセルとを含み、

前記検索領域に対応する複数の検索境界ボックスに基づく複数の類似のメトリックと、ここにおいて、前記検索領域は、前記複数の検索境界ボックスを含み、前記複数の類似のメトリックの各類似のメトリックは、前記第1の境界ボックス内の複数の第1のピクセル、および前記類似のメトリックに対応する前記複数の検索境界ボックスのうちの1つの検索境界ボックス内の複数の検索ピクセルに基づいて決定され、前記複数の検索境界ボックスの各々の検索座標は、1つ以上の方向にシフトされる前記第2の境界ボックスの第2の座標に基づき、

前記複数の類似のメトリックに基づく修正された第2の境界ボックスと、

を決定するための手段と、前記決定するための手段は、前記複数の第1のピクセルの列和ベクトルを決定するための手段、ここにおいて、前記列和ベクトルの特定の要素は、前記第1の境界ボックスの特定の列に対応するピクセル値の和に対応する、をさらに備え、  
を含む、装置。

#### 【請求項29】

受信するための前記手段、生成するための前記手段、および決定するための前記手段は、モバイル電話、セットトップボックス、ミュージックプレーヤ、ビデオプレーヤ、エンターテイメントユニット、ナビゲーションデバイス、通信デバイス、携帯情報端末（PDA）、固定ロケーションデータユニット、またはコンピュータのうちの少なくとも1つに組み込まれる、請求項28に記載の装置。

#### 【発明の詳細な説明】

#### 【関連出願の相互参照】

#### 【0001】

[0001] 本願は、2013年12月21日に出願された所有者共通の（commonly owned）米国仮特許出願第61/919,754号からの、および2014年12月11日に出願された米国非仮特許出願第14/567,119号からの優先権を主張し、その内容は、全文を参照により明確に本明細書に組み込まれる。

#### 【技術分野】

#### 【0002】

[0002] 本開示は、対象トラッキングボックス（object tracking box）の表示を安定化させることに一般的に關する。

#### 【背景技術】

#### 【0003】

[0003] 技術における進歩は、より小型で、より強力なコンピューティングデバイスをもたらした。例えば、小型で、軽量、かつユーザによって容易に持ち運ばれる、例えば携帯無線電話（portable wireless telephones）、携帯情報端末（PDAs）、およびペーディングデバイス等の、無線コンピューティングデバイスを含む、様々な携帯用パーソナルコンピューティングデバイスが現在存在する。より具体的には、例えばセルラ電話およびインターネットプロトコル（IP）電話等の携帯無線電話は、無線ネットワークにわたって音声およびデータパケットを通信することができる。さらに、多くのそのような無線電話は、そこに組み込まれる他のタイプのデバイスを含む。例えば、無線電話はまた、デジ

10

20

30

40

50

タル静止カメラ、デジタルビデオカメラ、デジタルレコーダ、およびオーディオファイルプレイヤを含むことができる。また、そのような無線電話は、インターネットにアクセスするために使用されることができる、例えばウェブブラウザアプリケーション等のソフトウェアアプリケーションを含む実行可能な命令を処理することができる。そのようなものとして、これらの無線電話は、重要なコンピューティング機能 (computing capabilities) を含むことができる。

#### 【0004】

[0004] 例えば無線電話等の電子デバイスは、カメラを含み得る。カメラは、ユーザがカメラディスプレイの中で見ることができる画像のシーケンスをキャプチャし得る。ユーザは、カメラディスプレイの領域を選択することによって画像の中の任意の対象 (arbitrary object) を選択し得る。トラッキングアルゴリズムは、後の画像にわたって対象の動きをトラックすることができ、カメラディスプレイ上でトラックされた対象上に (over) ボックスをディスプレイし得る。ディスプレイされたボックスは、画像間のボックスのロケーション、および / またはサイズにおける急速な変化のため不安定であるように見え得る。例えば、ユーザは、カメラを揺るがすように無線電話を持っている可能性がある。別の例として、対象は、画像間で大量の転移を持って (with a high amount of displacement) 動き得る。10

#### 【発明の概要】

#### 【0005】

[0005] 対象トラッキングボックスの表示を安定化させるシステムおよび方法は、開示される。ユーザは、画像がディスプレイされるカメラディスプレイの領域 (例えば、正方形または長方形) を選択することによって画像の中の対象を選択し得る。カメラディスプレイは、選択された対象を取り囲んでいる境界ボックス (bounding box) を示し得る。対象、カメラ、またはそれら両方は、画像のシーケンスがキャプチャされる間、動いている可能性がある。トラッカーは、境界ボックスが後の画像にわたって対象をおおよそトラックするような境界ボックスの座標および / またはディメンションをアップデートし得る。座標および / またはディメンションをアップデートすることは、1つの画像から別のものへ「ジャンプする」ように見える境界ボックスをもたらし得る。スタビライザー (stabilizer) は、第1の画像から後の画像への境界ボックスの表示を「スムーズにする」 (例えば、ジッターを低減する) ことができる。例えば、スタビライザーは、第1の画像に対応する第1の境界ボックスの第1の座標を受信することができ、トラッカーからの後の画像に対応する第2の境界ボックスのアップデートされた座標 (例えば、第2の座標) を受信し得る。スタビライザーは、第2の境界ボックスの周りの検索領域 (search region) を決定することができ、検索領域に対応する複数の検索境界ボックスを決定し得る。検索境界ボックスの各々は、ジッターを低減するために第2の境界ボックスに置き換えるための候補境界ボックス (candidate bounding box) に対応し得る。スタビライザーは、類似のメトリックに基づいて第1の境界ボックスと最も類似している特定の検索境界ボックスを選択するために検索境界ボックスの各々の検索ピクセルを、第1の境界ボックスの第1のピクセルと比較し得る。スタビライザーは、第2の境界ボックスを、選択された検索境界ボックスに置き換えることができ、それは、画像のシーケンスの中の境界ボックス (boundary boxes) の表示と関連する目に見えるジッター (visual jitter) を低減する。203040

#### 【0006】

[0006] 特定の態様では、方法は、画像のシーケンスの第1の画像のための第1の境界ボックスを定義する第1のデータを受信することを含む。第1の境界ボックスは、トラックされた対象を含む関心領域 (region of interest) に対応する。方法はまた、画像のシーケンスの第2の画像に関する対象トラッキングデータを受信すること、対象トラッキングデータは、第2の境界ボックスを定義し、を含む。第2の境界ボックスは、第2の画像の中のトラックされた対象を含む関心領域に対応する。方法はさらに、第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのための類似のメトリックを決定することを含む。検索境界ボックスの各々の検索座標は、1つ以50

上の方向にシフトされる第2の境界ボックスの第2の座標に対応する。方法はまた、類似のメトリックに基づいて、修正された第2の境界ボックスを決定することを含む。

#### 【0007】

[0007] 別の特定の態様では、装置は、メモリおよびプロセッサを含む。メモリは、命令を記憶するように構成される。プロセッサは、第1の画像の第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのための類似のメトリックを決定するための命令を実行するように構成される。第1の境界ボックスは、トラックされた対象を含む関心領域に対応する。検索境界ボックスの各々の検索座標は、1つ以上の方向にシフトされる第2の境界ボックスの第2の座標に対応する。第1の画像は、画像のシーケンスの中の第2の画像に先行する。第2の境界ボックスは、第2の画像の中のトラックされた対象を含む関心領域に対応する。プロセッサはまた、類似のメトリックに基づいて、修正された第2の境界ボックスを決定するための命令を実行するように構成される。  
10

#### 【0008】

[0008] 別の特定の態様では、コンピュータ可読記憶デバイスは、プロセッサによって実行されるとき、プロセッサに、第1の画像の第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのための類似のメトリックを決定することを含むオペレーションを行わせる命令を記憶する。第1の境界ボックスは、トラックされた対象を含む関心領域に対応する。検索境界ボックスの各々の検索座標は、1つ以上の方向にシフトされる第2の境界ボックスの第2の座標に対応する。第1の画像は、画像のシーケンスの中の第2の画像に先行する。第2の境界ボックスは、第2の画像中のトラックされた対象を含む関心領域に対応する。オペレーションはまた、類似のメトリックに基づいて、修正された第2の境界ボックスを決定することを含む。  
20

#### 【0009】

[0009] 提供される1つの特定の利点は、対象トラッキングボックス（例えば、境界ボックス）が画像のシーケンスの中の1つの画像から別のものに安定化されることである。例えば、境界ボックスの座標は、修正された境界ボックス内のピクセルが、先行する画像の先行する境界ボックス内のピクセルにより類似しているような修正された境界ボックスを生成するために修正され得る。別の例として、境界ボックスの修正されたサイズ（例えば、ディメンション）は、複数の先行する画像のメディアンディメンション（median dimensions）に対応し得る。本開示の他の態様、利点、および特徴は、以下のセクション：図面の簡単な説明、発明を実施するための形態、および特許請求の範囲を含む、本願全体を再考察した後に明らかとなるだろう。  
30

#### 【図面の簡単な説明】

#### 【0010】

【図1】[0010] 図1は、対象トラッキングボックスの表示を安定化させるように動作可能であるシステムの特定の例示的な実施形態のブロック図である。

【図2】[0011] 図2は、トラックされた対象を含む特定の例示的な画像の図表である。

【図3】[0012] 図3は、図2のトラックされた対象を含む別の例示的な画像の図表である。  
40

【図4】[0013] 図4は、図2のトラックされた対象を含む別の例示的な画像の図表である。

【図5】[0014] 図5は、対象トラッキングボックスの表示を安定化させる方法の特定の例示的な実施形態の流れ図である。

【図6】[0015] 図6は、対象トラッキングボックスのディスプレイを安定化させる方法の別の例示的な実施形態の流れ図である。

【図7】[0016] 図7は、図1-6のシステムおよび方法にしたがって対象トラッキングボックスの表示を安定化せしめるように動作可能であるデバイスのブロック図である。

#### 【発明を実施するための形態】

#### 【0011】

[0017] 図1を参照すると、対象トラッキングボックスの表示を安定化させるように動作可能であるシステムの特定の例示的な実施形態が開示され、概して100と示される。システム100は、カメラ112と、およびトラッカー160と結合されるスタビライザー102を含む。スタビライザー102は、メモリ120を含み得る。特定の実施形態では、システム100の1つ以上のコンポーネントは、モバイル電話、セットトップボックス、ミュージックプレーヤ、ビデオプレーヤ、エンターテイメントユニット、ナビゲーションデバイス、通信デバイス、携帯情報端末(PDA)、固定ロケーションデータユニット、コンピュータ、またはそれらの組み合わせに組み込まれ得る。

#### 【0012】

[0018] 以下の説明では、図1のシステム100によって行われる様々な機能が、あるコンポーネントまたはモジュールによって行われるように説明されることに留意されるべきである。しかしながら、コンポーネントおよびモジュールのこの区分(division)は、例示のためだけである。代替の実施形態では、特定のコンポーネントまたはモジュールによって行われる機能は、複数のコンポーネントまたはモジュールの間で分けられ得る。さらに、代替の実施形態では、図1の2つ以上のコンポーネントまたはモジュールは、単一のコンポーネントまたはモジュールに組み込まれ得る。図1に例示された各コンポーネントまたはモジュールは、ハードウェア(例えば、フィールドプログラマブルゲートアレイ(FPGA)デバイス、特定用途向け集積回路(ASIC)、デジタルシグナルプロセッサ(DSP)、コントローラ等)、ソフトウェア(例えば、プロセッサによって実行可能な命令)、またはそれらの任意の組み合わせを使用してインプリメントされ得る。

10

#### 【0013】

[0019] オペレーションの間に、カメラ112は、画像のシーケンス104をキャプチャし得る。特定の実施形態では、画像のシーケンス104は、ユーザが記録している(例えば、メモリに記憶している)ビデオストリームに対応し得る。別の実施形態では、画像のシーケンス104は、時間期間にわたって(例えば、ビューファインダディスプレイ(viewfinder display)に対応している)カメラディスプレイによってディスプレイされる画像データに対応し得る。例えば、ユーザは、画像データを見ることができ、続いて写真を撮り得る(例えば、メモリに特定の画像を記憶する)。

#### 【0014】

[0020] 画像のシーケンス104は、第1の画像106を含み得る。第1の画像106は、カメラディスプレイ(示されていない)を介してユーザ150にディスプレイされ得る。ユーザ150は、第1の画像106の中の関心領域162を選択することによってカメラディスプレイでディスプレイされる対象110(例えば、図1の車)を選択し得る。関心領域162は、対象110を含み得る。スタビライザー102は、関心領域162に対応する第1の境界ボックス116を定義するユーザ150からの第1のデータ122を受信し得る。例えば、第1のデータ122は、第1の画像106の中の第1の境界ボックス116の第1の座標(例えば、水平軸(x軸)座標および垂直軸(y軸)座標)を含み得る。特定の実施形態では、第1の座標は、第1の画像106の中の第1の境界ボックス116の左上隅に対応し得る。第1のデータ122はまた、第1の境界ボックス116の第1のディメンションを含み得る。例えば、第1のデータは、第1の境界ボックス116の(例えば、第1の画像106のx軸に沿った)幅および(例えば、第1の画像106のy軸に沿った)高さを含み得る。この例では、第1の境界ボックス116は、正方形または長方形に対応し得る。

30

#### 【0015】

[0021] 画像のシーケンス104は、第2の画像108を含み得る。第2の画像108はまた、対象110を含む関心領域162を含み(例えば、描写し)得る。トラッカー160は、対象トラッキングボックス(例えば、第2の境界ボックス118)を定義する対象トラッキングデータ124を生成し得る。第2の境界ボックス118は、第2の画像108の中の関心領域162に対応し得る。例えば、対象トラッキングデータ124は、第2の画像108の中の第2の境界ボックス118の第2の座標(例えば、x軸座標および

40

50

*y* 軸座標)を含み得る。特定の実施形態では、第2の座標は、第2の画像108の中の第2の境界ボックス118の左上隅に対応する。対象トラッキングデータ124はまた、第2の境界ボックス118の第2のディメンションを含み得る。例えば、対象トラッキングデータ124は、第2の境界ボックス118の(例えば、第2の画像108のx軸に沿った)幅および(例えば、第2の画像108の*y*軸に沿った)高さを含み得る。

#### 【0016】

[0022] スタビライザー102は、第2の境界ボックス118に基づいて第2の画像108のための検索領域を決定し得る。例えば、検索領域は、図4に関して説明されるように、第2の境界ボックス118内のピクセル、および第2の境界ボックス118に実質的に近いピクセルを含み得る。スタビライザー102は、図4に関して説明されるように、検索領域内の複数の検索境界ボックスを決定し得る。例えば、検索境界ボックスの各々の座標は、1つ以上の方に向シフトされる第2の境界ボックス118の第2の座標に対応し得る。検索境界ボックスの各々は、ジッターを低減するために、第2の画像108をディスプレイするとき、第2の境界ボックス118を置き換えるための候補境界ボックスに対応し得る。スタビライザー102は、メモリ120に検索境界ボックスの検索座標130を記憶し得る。検索境界ボックスの各々のディメンションは、第1の境界ボックス116の第1のディメンションに対応し得る。10

#### 【0017】

[0023] スタビライザー102は、第1の境界ボックス116内の第1のピクセル、および検索境界ボックスの各々内の検索ピクセルのための類似のメトリックを決定し得る。例えば、類似のメトリックは、差分絶対値和(sum of absolute differences)(SAD)メトリックを含み得る。例示するために、スタビライザー102は、第1のピクセルに対応する第1のピクセル特性(first pixel characteristics)(例えば、ピクセル輝度(a pixel intensity)、ピクセルの色(例えば、赤、緑、青、シアン、マゼンタ、黄、または黒)サブコンポーネント、またはそれらの組み合わせ)および候補検索ピクセルに対応する第2のピクセル特性のSADに少なくとも部分的に基づいて、候補検索境界ボックス内の候補検索ピクセルおよび第1のピクセルのための特定の類似のメトリックを計算し得る。20

#### 【0018】

[0024] 特定の実施形態では、スタビライザー102は、図2に関して説明されるように、第1の境界ボックス116の第1のピクセルの第1の列和ベクトル(first column sum vector)、第1の列和差分ベクトル(first column sum difference vector)、第1の行和ベクトル(first row sum vector)、および/または第1の行和差分ベクトルを計算し得る。同様に、スタビライザー102はまた、候補検索境界ボックスの候補検索ピクセルの第2の列和ベクトル、第2の列和差分ベクトル、第2の行和ベクトル、および/または第2の行和差分ベクトルを計算し得る。スタビライザー102は、メモリ120に、列和ベクトル132(例えば、第1の列和ベクトルおよび第2の列和ベクトル)、列和差分ベクトル134(例えば、第1の列和差分ベクトルおよび第2の列和差分ベクトル)、行和ベクトル136(例えば、第1の行和ベクトルおよび第2の行和ベクトル)、行和差分ベクトル138(例えば、第1の行和差分ベクトルおよび第2の行和差分ベクトル)、またはそれらの組み合わせを記憶し得る。30

#### 【0019】

[0025] スタビライザー102は、第1の列和ベクトルおよび第2の列和ベクトルの第1のSAD、第1の列和差分ベクトルおよび第2の列和差分ベクトルの第2のSAD、第1の行和ベクトルおよび第2の行和ベクトルの第3のSAD、および/または第1の行和差分ベクトルおよび第2の行和差分ベクトルの第4のSADを合わせることによって第1の境界ボックス116の第1のピクセルおよび候補検索境界ボックスの候補検索ピクセルのための特定の類似のメトリックを決定し得る。スタビライザー102は、メモリ120に検索境界ボックスの各々に対応する類似のメトリック(similarity metrics)128を記憶し得る。40

**【 0 0 2 0 】**

[0026] スタビライザー 102 は、第 1 の境界ボックス 116 の第 1 のピクセルに最も類似している検索ピクセルを含む特定の検索境界ボックスを選択し得る。例えば、スタビライザー 102 は、対応する類似のメトリックが類似のメトリック 128 の（例えば、最も低い値を有する）第 1 のピクセルと最も高い類似 (highest similarity) を示すことを決定することに応答して特定の検索境界ボックスを選択し得る。

**【 0 0 2 1 】**

[0027] スタビライザー 102 は、選択された検索境界ボックスに基づいて、修正された第 2 の境界ボックスを決定し得る。例えば、スタビライザー 102 は、修正された第 2 の境界ボックスデータ 126 を生成し得る。修正された第 2 の境界ボックスデータ 126 は、修正された第 2 の境界ボックスの修正された座標を示し得る。修正された座標は、選択された検索境界ボックスの座標に対応し得る。特定の実施形態では、修正された座標は、第 2 の境界ボックス 118 の第 2 の座標と同一であり得る。ゆえに、特定の検索境界ボックスが第 1 の境界ボックス 116 に最も類似しているように決定されるため、特定の検索境界ボックスは、複数の検索（例えば、候補）境界ボックスの中から選択することができ、その結果、画像のシーケンス 104 の中の画像間の境界ボックスの配置の中の目に見えるジッターを低減する。

**【 0 0 2 2 】**

[0028] 修正された第 2 の境界ボックスデータ 126 は、修正された第 2 の境界ボックスのディメンションを示し得る。特定の実施形態では、修正された第 2 の境界ボックスのディメンションは、第 1 の境界ボックス 116 の第 1 のディメンション、または第 2 の境界ボックス 118 の第 2 のディメンションに対応し得る。別の実施形態では、修正された第 2 の境界ボックスのディメンションは、第 2 の画像 108 に先行する複数の画像に対応するメディアンディメンション 140 に対応し得る。スタビライザー 102 は、先行する画像の数がしきい値を満たすことを決定することに応答して、修正された第 2 の境界ボックスのディメンションとしてメディアンディメンション 140 を使用し得る。先行する画像のしきい値の数は、デフォルト値であり得る。スタビライザー 102 は、修正された第 2 の境界ボックスデータ 126 をカメラディスプレイに送り得る。例えば、カメラディスプレイは、修正された第 2 の境界ボックスで第 2 の画像 108 をディスプレイし得る。

**【 0 0 2 3 】**

[0029] 特定の実施形態では、スタビライザー 102 は、第 2 の画像 108 の後で画像（例えば、画像のシーケンス 104 の第 3 の画像）を受信することを予期して、修正された第 2 の境界ボックスに対応するピクセル特性を記憶し得る。修正された第 2 の境界ボックスの修正されたディメンションが選択された検索境界ボックスのディメンション（すなわち、第 1 の境界ボックス 116 の第 1 のディメンション）に対応するとき、修正された第 2 の境界ボックスは、選択された検索境界ボックスに対応し得る。スタビライザー 102 は、修正された第 2 の境界ボックスの修正されたディメンションが選択された検索境界ボックスのディメンション（または、第 1 の境界ボックス 116 の第 1 のディメンション）に対応することを決定することに応答して、選択された検索境界ボックスのピクセル特性を記憶し得る。

**【 0 0 2 4 】**

[0030] 特定の実施形態では、修正された第 2 の境界ボックスの修正されたディメンションは、選択された検索境界ボックスのディメンション（または第 1 の境界ボックス 116 の第 1 のディメンション）と区別され得る。例えば、修正された第 2 の境界ボックスの修正されたディメンションは、第 2 の境界ボックス 118 の第 2 のディメンションに、またはメディアンディメンション 140 に対応し得る。修正された第 2 の境界ボックスのディメンションが、選択された検索境界ボックスのディメンション（または、第 1 のディメンション）に対応しないとき、スタビライザー 102 は、修正された第 2 の境界ボックスのピクセル特性を生成、および記憶し得る。例えば、スタビライザー 102 は、修正された第 2 の境界ボックスに対応して、図 2 に関して説明されるように、行和ベクトル、行和

10

20

30

40

50

差分ベクトル、列和ベクトル、および／または列和差分ベクトルを生成し得る。

#### 【0025】

[0031] 画像のシーケンス 104 の追加の画像が受信されるとき、追加の境界ボックスは、(1つまたは複数の)前回の(previous)境界ボックスとの類似に基づいて選択され得る。例えば、スタビライザー 102 は、第3の画像を受信することができ、第3の画像に対応する第3の境界ボックスを定義する対象トラッキングデータ 124 を受信し得る。スタビライザー 102 は、修正された第2の境界ボックス、および第3の境界ボックスに基づいて、修正された第3の境界ボックスを決定し得る。例えば、スタビライザー 102 は、修正された第3の境界ボックスを決定するために使用されることができる追加の類似のメトリックを生成するために、修正された第2の境界ボックスの記憶されたピクセル特性を使用し得る。10

#### 【0026】

[0032] ゆえに、特定の境界ボックスが先行する画像(例えば、第1の画像 106)の境界ボックスに最も類似しているように決定されるため、スタビライザー 102 は、複数の検索(例えば、候補)境界ボックスの中から特定の境界ボックスを選択することができ、その結果、画像のシーケンス 104 の中の画像間の境界ボックスの配置の中の目に見えるジッターを低減する。

#### 【0027】

[0033] 図2を参照すると、トラックされた対象を含む例示の画像の図表が開示され、概して 200 と示される。特定の実施形態では、画像 200 は、図1の第1の画像 106、または第2の画像 108 に対応し得る。画像 200 は、ピクセルの複数の列(例えば、列 0 - 列 5)、および複数の行(例えば、行 0 - 行 4)を含む。画像 200 は、ピクセルの列のセット(例えば、列 0 - 列 2)、および行のセット(例えば、行 0 - 行 2)を含む第1の境界ボックス 202 を含むことができ、それらは、枠で囲まれたピクセル(boxed pixels)として例示される。特定の実施形態では、枠で囲まれたピクセルは、第1の境界ボックス 116 のピクセル、1つ以上の検索境界ボックスのピクセル、またはそれら両方に対応し得る。特定のピクセルは、その座標(例えば、x, y 座標)で識別され得る。画像 200 は、各ピクセル(x, y)の特定のピクセル特性 I を示す。特定の実施形態では、ピクセル特性は、ピクセル輝度に対応し得る。例えば、I(0, 0) は、列 0、および行 0 でピクセルのピクセル輝度を示すことができ、I(0, 1) は、列 0、および行 1 でピクセルのピクセル輝度を示し得る。特定の実施形態では、画像 200 は、グレースケール画像であり得る。2030

#### 【0028】

[0034] 第1の境界ボックス 202 は、画像 200 のために定義され得る。(0, 0) でピクセルは、第1の境界ボックス 202 の左上隅を定義する。第1の境界ボックス 202 のディメンションは、3ピクセルの高さ、および3ピクセルの幅を含み得る。スタビライザー 102 は、左上隅の座標、およびディメンションに基づいて第1の境界ボックス 202 の右下隅(例えば、(2, 2))を決定し得る。特定の実施形態では、第1の境界ボックス 202 は、図1の第1の境界ボックス 116 に対応し得る。別の実施形態では、第1の境界ボックス 202 は、図1に関して説明された検索境界ボックスのうちの1つ以上に対応し得る。40

#### 【0029】

[0035] スタビライザー 102 は、境界ボックスの列和ベクトル c(x, y) を決定し得る。境界ボックスの左上隅の座標は、(x, y) に対応し得る。例えば、スタビライザ-102 は、第1の境界ボックス 202 の枠で囲まれたピクセルの第1の列和ベクトル(c(0, 0)) 204 を決定し得る。図2に示された例では、I(0, 0) = 1、I(1, 0) = 2、I(2, 0) = 3、I(3, 0) = 10、I(0, 1) = 4、I(1, 1) = 5、I(2, 1) = 6、I(3, 1) = 11、I(0, 2) = 7、I(1, 2) = 8、I(2, 2) = 9、I(3, 2) = 12、I(0, 3) = 13、I(1, 3) = 14、I(2, 3) = 15、および I(3, 3) = 16。c(0, 0) 204 の各要素(Each ele50

ment)  $c_j(0, 0)$  は、第 1 の境界ボックス 202 の列  $j$  のピクセル値の和と等しくあり得る。ゆえに、図 2 に示された例に関して、 $c_j(0, 0) 204$  の値は：

$$\begin{aligned}c_0(0, 0) &= I(0, 0) + I(0, 1) + I(0, 2) = 12, \\c_1(0, 0) &= I(1, 0) + I(1, 1) + I(1, 2) = 15, \text{ および} \\c_2(0, 0) &= I(2, 0) + I(2, 1) + I(2, 2) = 18.\end{aligned}$$

### 【0030】

[0036] スタビライザー 102 は、第 1 の境界ボックス 202 の枠で囲まれたピクセルの列和差分ベクトル  $c_{delta}(0, 0) 206$  を決定し得る。 $c_{delta}(0, 0) 206$  の各要素  $c_{delta_j}(0, 0)$  は、 $j > 0$  の場合  $c_j(0, 0) - c_{j-1}(0, 0)$  に等しくあることができ、 $j = 0$  の場合 0 に等しくあり得る。ゆえに、図 2 に示された例に関して、 $c_{delta}(0, 0) 206$  の値は：

$$\begin{aligned}c_{delta_0}(0, 0) &= 0, \\c_{delta_1}(0, 0) &= c_1(0, 0) - c_0(0, 0) = 15 - 12 = 3, \text{ および} \\c_{delta_2}(0, 0) &= c_2(0, 0) - c_1(0, 0) = 18 - 15 = 3.\end{aligned}$$

### 【0031】

[0037] スタビライザー 102 は、第 1 の境界ボックス 202 の枠で囲まれたピクセルの行和ベクトル  $r(0, 0) 208$  を決定し得る。 $r(0, 0) 208$  の各要素  $r_i(0, 0)$  は、第 1 の境界ボックス 202 の行  $i$  のピクセル値の和と等しくあり得る。ゆえに、図 2 に示される例に関して、 $r(0, 0) 208$  の値は：

$$\begin{aligned}r_0(0, 0) &= I(0, 0) + I(1, 0) + I(2, 0) = 6, \\r_1(0, 0) &= I(0, 1) + I(1, 1) + I(2, 1) = 15, \text{ および} \\r_2(0, 0) &= I(0, 2) + I(1, 2) + I(2, 2) = 24.\end{aligned}$$

### 【0032】

[0038] スタビライザー 102 は、第 1 の境界ボックス 202 の枠で囲まれたピクセルの第 1 の行差分ベクトル  $r_{delta}(0, 0) 210$  を決定し得る。 $r_{delta}(0, 0) 210$  の各要素  $r_{delta_i}(0, 0)$  は、 $i > 0$  の場合  $r_i(0, 0) - r_{i-1}(0, 0)$  に等しくあることができ、 $i = 0$  の場合 0 に等しくあり得る。ゆえに、図 2 に示された例に関して、 $r_{delta}(0, 0) 210$  の値は：

$$\begin{aligned}r_{delta_0}(0, 0) &= 0, \\r_{delta_1}(0, 0) &= r_1(0, 0) - r_0(0, 0) = 15 - 6 = 9, \text{ および} \\r_{delta_2}(0, 0) &= r_2(0, 0) - r_1(0, 0) = 24 - 15 = 9.\end{aligned}$$

### 【0033】

[0039] 特定の実施形態では、スタビライザー 102 は、画像 200 に対応する積分画像 (integral image) を生成し得る。積分画像の各特定のピクセルの値は、特定のピクセル、ならびに特定のピクセルの上のピクセル、および特定のピクセルの左のピクセル、のピクセル値の和に等しい。積分画像  $Int$  の中のピクセル  $(x, y)$  の値は、 $Int(x, y)$  で表示され得る。 $Int(x, y)$  は、式  $Int(x, y) = I(x, y) + Int(x-1, y) + Int(x, y-1) - Int(x-1, y-1)$  によって決定され得る。

### 【0034】

[0040] 図 2 に示された例では、積分画像のピクセルのサブセットの値は：

$$\begin{aligned}Int(0, 0) &= I(0, 0) = 1 \\Int(0, 1) &= I(0, 1) + I(0, 0) = 5 \\Int(0, 2) &= I(0, 2) + I(0, 1) + I(0, 0) = 12 \\Int(0, 3) &= I(0, 3) + I(0, 2) + I(0, 1) + I(0, 0) = 25 \\Int(1, 0) &= I(1, 0) + I(0, 0) = 3 \\Int(1, 1) &= I(1, 1) + I(0, 1) + I(0, 0) + I(1, 0) = 12 \\Int(1, 2) &= I(1, 2) + I(1, 1) + I(0, 1) + I(0, 0) + I(1, 0) + I(0, 2) = 27 \\Int(1, 3) &= I(1, 3) + I(1, 2) + I(1, 1) + I(0, 1) + I(0, 0) = 50\end{aligned}$$

$, 0) + I(1, 0) + I(0, 2) + I(0, 3) = 54$   
 $I_{nt}(2, 0) = I(2, 0) + I(1, 0) + I(0, 0) = 6$   
 $I_{nt}(2, 1) = I(2, 1) + I(2, 0) + I(1, 0) + I(0, 0) + I(1, 1) + I(0, 1) = 21$   
 $I_{nt}(2, 2) = I(2, 2) + I(2, 1) + I(2, 0) + I(1, 0) + I(0, 0) + I(1, 1) + I(0, 1) + I(1, 2) + I(0, 2) = 45$   
 $I_{nt}(2, 3) = I(2, 3) + I(2, 2) + I(2, 1) + I(2, 0) + I(1, 0) + I(0, 0) + I(1, 1) + I(0, 1) + I(1, 2) + I(0, 2) + I(0, 3) + I(1, 3) = 87$   
 $I_{nt}(3, 0) = I(3, 0) + I(2, 0) + I(1, 0) + I(0, 0) = 16$  10  
 $I_{nt}(3, 1) = I(3, 1) + I(3, 0) + I(2, 1) + I(2, 0) + I(1, 0) + I(0, 0) + I(1, 1) + I(0, 1) = 42$   
 $I_{nt}(3, 2) = I(3, 2) + I(3, 1) + I(3, 0) + I(2, 2) + I(2, 1) + I(2, 0) + I(1, 0) + I(0, 0) + I(1, 1) + I(0, 1) + I(1, 2) + I(0, 2) = 78$   
 $I_{nt}(3, 3) = I(3, 3) + I(3, 2) + I(3, 1) + I(3, 0) + I(2, 3) + I(2, 2) + I(2, 1) + I(2, 0) + I(1, 0) + I(0, 0) + I(1, 1) + I(0, 1) + I(1, 2) + I(0, 2) + I(0, 3) + I(1, 3) = 136$

## 【0035】

20

[0041] スタビライザー 102 は、積分画像から  $c(0, 0) 204$  および  $r(0, 0) 208$  の値を決定し得る。例えば、 $c(0, 0) 204$  の特定の要素は、積分画像の第 1 の要素と積分画像の第 2 の要素の差分に対応することができ、ここで、第 1 の要素および第 2 の要素は、積分画像の近接した列に対応する。例示するために、 $c_0(0, 0)$  は、 $I_{nt}(0, 2)$  に対応し、 $c_1(0, 0)$  は、 $I_{nt}(1, 2) - I_{nt}(0, 2)$  に対応し、および  $c_2(0, 0)$  は、 $I_{nt}(2, 2) - I_{nt}(1, 2)$  に対応する。

## 【0036】

[0042] 特定の実施形態では、左上の座標 ( $x, 0$ ) を持つ境界ボックス（例えば、第 1 の境界ボックス 202）の  $c_i(x, 0)$  の値は：

$x + i = 0$  の場合、 $c_i(x, 0) = I_{nt}(0, \text{高さ} - 1)$ 、および

30

$x + i > 0$  の場合、 $c_i(x, 0) = I_{nt}(x + i, \text{高さ} - 1) - I_{nt}(x + i - 1, \text{高さ} - 1)$ 、

ここで、高さ（例えば、3）は、境界ボックスの行の数に対応する。

## 【0037】

[0043] 別の例として、 $r(0, 0) 208$  の特定の要素は、積分画像の第 1 の要素と積分画像の第 2 の要素の差分に対応することができ、ここで、第 1 の要素および第 2 の要素は、積分画像の近接した行に対応する。例示するために、 $r_0(0, 0)$  は、 $I_{nt}(2, 0)$  に対応し、 $r_1(0, 0)$  は、 $I_{nt}(2, 1) - I_{nt}(2, 0)$  に対応し、および  $r_2(0, 0)$  は、 $I_{nt}(2, 2) - I_{nt}(2, 1)$  に対応する。

## 【0038】

40

[0044] 特定の実施形態では、左上の座標 ( $0, y$ ) を持つ境界ボックス（例えば、第 1 の境界ボックス 202）の  $r_i(0, y)$  の値は：

$y + j = 0$  の場合、 $r_j(0, y) = I_{nt}(\text{幅} - 1, 0)$ 、および

$y + j > 0$  の場合、 $r_j(0, y) = I_{nt}(\text{幅} - 1, y + j) - I_{nt}(\text{幅} - 1, y + j - 1)$ 、

ここで、幅（例えば、3）は、境界ボックスの列の数に対応する。

## 【0039】

[0045] 別の例として、スタビライザー 102 は、積分画像を使用して第 2 の境界ボックス 204 のために行和ベクトル  $r(1, 1)$ 、および列和ベクトル  $c(1, 1)$  を決定し得る。例えば、 $c(1, 1)$  の特定の要素は、積分画像の第 1 の要素と積分画像の第 2

50

の要素の差分に対応することができ、ここで、第1の要素および第2の要素は、積分画像の近接した列に対応する。例示するために、 $c_0(1, 1)$ は、 $\text{Int}(1, 3) - \text{Int}(0, 3) - \text{Int}(1, 0) + \text{Int}(0, 0)$ に対応し、 $c_1(1, 1)$ は、 $\text{Int}(2, 3) - \text{Int}(1, 3) - \text{Int}(2, 0) + \text{Int}(1, 0)$ に対応し、および $c_2(1, 1)$ は、 $\text{Int}(3, 3) - \text{Int}(2, 3) - \text{Int}(3, 0) + \text{Int}(2, 0)$ に対応する。

#### 【0040】

[0046] 特定の実施形態では、左上の座標( $x, y$ )を持つ境界ボックス(例えば、第2の境界ボックス204)の $c_i(x, y)$ の値、ここで $y > 0$ は：

$x + i = 0$ の場合、 $c_i(x, y) = \text{Int}(0, y + \text{高さ} - 1) - \text{Int}(0, y - 1)$  10  
）、および

$x + i > 0$ の場合、 $c_i(x, y) = \text{Int}(x + i, y + \text{高さ} - 1) - \text{Int}(x + i - 1, y + \text{高さ} - 1) - \text{Int}(x + i, y - 1) + \text{Int}(x + i - 1, y - 1)$ 、  
ここで、高さ(例えば、3)は、境界ボックスの行の数に対応する。

#### 【0041】

[0047] 別の例として、 $r(1, 1)$ の特定の要素は、積分画像の第1の要素と積分画像の第2の要素の差分に対応することができ、ここで、第1の要素および第2の要素は、積分画像の近接した行に対応する。例示するために、 $r_0(1, 1)$ は、 $\text{Int}(3, 1) - \text{Int}(0, 1) - \text{Int}(3, 0) + \text{Int}(0, 0)$ に対応し、 $r_1(1, 1)$ は、 $\text{Int}(3, 2) - \text{Int}(0, 2) - \text{Int}(3, 1) + \text{Int}(0, 1)$ に対応し、および $r_2(1, 1)$ は、 $\text{Int}(3, 3) - \text{Int}(0, 3) - \text{Int}(3, 2) + \text{Int}(0, 2)$ に対応する。 20

#### 【0042】

[0048] 特定の実施形態では、左上の座標( $x, y$ )を持つ境界ボックス(例えば、第2の境界ボックス204)の $r_j(x, y)$ の値、ここで $x > 0$ は：

$y + j = 0$ の場合、 $r_j(x, y) = \text{Int}(x + \text{幅} - 1, 0) - \text{Int}(x - 1, 0)$ 、および

$y + j > 0$ の場合、 $r_j(x, y) = \text{Int}(x + \text{幅} - 1, y + j) - \text{Int}(x + \text{幅} - 1, y + j - 1) - \text{Int}(x - 1, y + j) + \text{Int}(x - 1, y + j - 1)$ 、  
ここで、幅(例えば、3)は、境界ボックスの列の数に対応する。 30

#### 【0043】

[0049] 特定の実施形態では、スタビライザー102は、第2の画像108に対応する積分画像を生成し得る。積分画像からの検索境界ボックスの各々に対応する行和ベクトル、および列和ベクトルを計算することは、第2の画像108のピクセル値から直接ベクトルを計算することより速くある(例えば、一定の時間の中で算出される)ことができ、およびより少ない処理リソースを使用し得る。

#### 【0044】

[0050] 図3を参照すると、図2のトラックされた対象を含む例示の画像の図が開示され、概して300と示される。画像300は、図1の第2の画像108に対応し得る。左上の座標(0, 1)およびディメンション $4 \times 3$ (ピクセルの幅×ピクセルの高さ)は、画像300のための第2の境界ボックス118を定義する。各ピクセル値 $I_2(x, y)$ は、画像300の対応するピクセル( $x, y$ )のピクセル特性に対応し得る。添字2は、ピクセル特性が第2の画像108に対応することを示す。 40

#### 【0045】

[0051] スタビライザー102は、第2の境界ボックス118に基づいて検索領域304を決定し得る。例えば、検索領域304は、第2の境界ボックス118、および第2の境界ボックス118に関連して1つ以上の方向に追加のピクセルを含み得る。例えば、スタビライザー102は、検索領域304を生成するために、第2の境界ボックス118の右に、および第2の境界ボックス118の左にピクセルの第1の数(例えば、1)を合計し、ならびに第2の境界ボックス118の上に、および第2の境界ボックス118の下に 50

ピクセルの第 2 の数（例えば、1）を合計し得る。

**【0046】**

[0052] 図 4 を参照すると、図 2 のトラックされた対象を含む例示的な画像の図表が開示され、概して 400 と示される。画像 400 は、図 1 の第 2 の画像 108 に対応し得る。検索領域 304 は、画像 400 のために定義され得る。

**【0047】**

[0053] スタビライザー 102 は、検索領域 304 に基づいて画像 400 内で複数の検索境界ボックス（例えば、第 1 の検索境界ボックス 402、第 2 の検索境界ボックス 404、および第 3 の検索境界ボックス 406）を生成し得る。検索境界ボックスの各々のディメンションは、第 1 の境界ボックス 116 のディメンションと等しくあり得る。スタビライザー 102 は、第 1 の範囲（例えば、0 - 3）から選択された左上のピクセルの x 座標、および第 2 の範囲（例えば、0 - 2）から選択された左上のピクセルの y 座標を持つ検索境界ボックスを生成し得る。検索領域 304 は、第 1 の範囲の外の x 座標、または第 2 の範囲の外の y 座標を持つピクセルを含み得る。しかしながら、スタビライザー 102 は、そのような検索境界ボックスが画像 400 内で完全に一致しないため、第 1 の範囲の外の x 座標、第 2 の範囲の外の y 座標、またはそれら両方を有する左上のピクセルを持つ検索境界ボックスを生成することを控え得る。10

**【0048】**

[0054] スタビライザー 102 は、図 1 - 2 に関してさらに説明されたように、検索境界ボックス 402 - 406 の各々に対応する検索行和ベクトル、検索行和差分ベクトル、検索列和ベクトル、および / または検索列和差分ベクトルを生成し得る。スタビライザー 102 は、図 1 に関して説明されたように、類似のメトリック 128 を生成することができ、第 1 の境界ボックス 116 に「最も類似している」として特定の検索境界ボックスを選択し得る。例えば、スタビライザー 102 は、第 2 の検索境界ボックス 404 の類似のメトリックに基づいて第 2 の検索境界ボックス 404 を選択し得る。ゆえに、スタビライザー 102 は、ユーザ 150 のためにディスプレイされた、修正された境界ボックスがロケーション (1, 1) における左上のピクセルを有するべきことを示すために、修正された第 2 の境界ボックスデータ 126 を生成し得る。20

**【0049】**

[0055] 特定の実施形態では、修正された第 2 の境界ボックスのディメンションは、図 1 の第 1 の境界ボックス 116 の第 1 のディメンション、図 1 の第 2 の境界ボックス 118 の第 2 のディメンション、または図 1 のメディアンディメンション 140 に対応し得る。修正された第 2 の境界ボックスデータ 126 は、左上のピクセルの座標 (1, 1)、および修正された第 2 の境界ボックスのディメンションを示し得る。スタビライザー 102 は、修正された第 2 の境界ボックスデータ 126 をカメラディスプレイに提供し得る。カメラディスプレイは、修正された第 2 の境界ボックスで第 2 の画像 108 をディスプレイし得る。修正された第 2 の境界ボックスは、第 2 の境界ボックス 118 より安定した境界ボックスに対応し得る。例えば、修正された第 2 の境界ボックスに対応するピクセルは、第 1 の境界ボックス 116 の第 1 のピクセルにより類似している可能性がある。別の例として、修正された第 2 の境界ボックスのディメンションは、複数の先行する画像のディメンションにより密接に対応し得る。30

**【0050】**

[0056] 図 5 を参照すると、対象トラッキングボックスの表示を安定化させる方法の特定の例示的な実施形態の流れ図が示され、概して 500 と示される。方法 500 は、図 1 のシステム 100 の 1 つ以上のコンポーネントによって行われ得る。

**【0051】**

[0057] 方法 500 は、502 において、画像のシーケンスの第 1 の画像のための第 1 の境界ボックスを定義する第 1 のデータを受信することを含む。第 1 の境界ボックスは、トラックされた対象を含む関心領域に対応し得る。例えば、図 1 のスタビライザー 102 は、図 1 に関して説明されたように第 1 のデータ 122 を受信し得る。第 1 のデータ 1240

2は、画像のシーケンス104の第1の画像106のための第1の境界ボックス116を定義し得る。第1の境界ボックス116は、対象110を含む関心領域162に対応し得る。

#### 【0052】

[0058] 方法500はまた、504において、画像のシーケンスの第2の画像に関する対象トラッキングデータを受信することを含む。対象トラッキングデータは、第2の境界ボックスを定義し得る。第2の境界ボックスは、第2の画像の中のトラックされた対象を含む関心領域に対応し得る。例えば、図1のスタビライザー102は、図1に関して説明されたように、第2の画像108のための対象トラッキングデータ124を受信し得る。対象トラッキングデータ124は、第2の境界ボックス118を定義し得る。第2の境界ボックス118は、第2の画像108の中の対象110を含む関心領域162に対応し得る。10

#### 【0053】

[0059] 方法500はさらに、506において、第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのための類似のメトリックを決定することを含む。検索境界ボックスの各々の検索座標は、1つ以上の方にシフトされる第2の境界ボックスの第2の座標に対応し得る。例えば、図1のスタビライザー102は、図1-2および4に関して説明されたように、第1の境界ボックス116内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのための類似のメトリックを決定し得る。20

#### 【0054】

[0060] 方法500はまた、508において、類似のメトリックに基づいて、修正された第2の境界ボックスを決定することを含む。例えば、図1のスタビライザー102は、図1および4に関して説明されたように、類似のメトリックに基づいて、修正された第2の境界ボックスを決定し得る。

#### 【0055】

[0061] 方法500はさらに、510において、複数の画像に対応するメディアンディメンションを決定することを含む。複数の画像は、画像のシーケンスの中の第2の画像に先行し得る。修正された第2の境界ボックスの第2のディメンションは、メディアンディメンションに対応し得る。例えば、図1のスタビライザー102は、図1に関して説明されたように、第2の画像108に先行する複数の画像に対応するメディアンディメンション140を決定し得る。修正された第2の境界ボックスの第2のディメンションは、図1および4に関して説明されたように、メディアンディメンション140に対応し得る。30

#### 【0056】

[0062] ゆえに、特定の境界ボックスが先行する画像（例えば、第1の画像106）の境界ボックスに最も類似しているように決定されるため、方法500は、複数の検索（例えば、候補）境界ボックスの中からの特定の境界ボックスの選択を含み、その結果、画像のシーケンス104の中の画像間の境界ボックスの配置における目に見えるジッターを低減する。

#### 【0057】

[0063] 図5の方法500は、フィールドプログラマブルゲートアレイ（FPGA）デバイス、特定用途向け集積回路（ASIC）、例えば中央処理装置（CPU）等の処理装置、デジタルシグナルプロセッサ（DSP）、コントローラ、別のハードウェアデバイス、ファームウェアデバイス、またはそれらの任意の組み合わせによってインプリメントされ得る。一例として、図5の方法500は、図7に関して説明されるように、命令を実行するプロセッサによって行われることができる。40

#### 【0058】

[0064] 図6を参照すると、対象トラッキングボックスの表示を安定化させる方法の特定の例示的な実施形態の流れ図が示され、概して600と示される。方法600は、図1のシステム100の1つ以上のコンポーネントによって行われ得る。特定の実施形態では50

、方法 600 は、図 5 のステップ 506 において例示されたオペレーションに対応し得る。

#### 【0059】

[0065] 方法 600 は、602 において、第 1 のピクセルの第 1 の列和ベクトルを決定することを含む。例えば、図 1 のスタビライザー 102 は、図 1 - 2 に関して説明されたように、第 1 の境界ボックス 116 の第 1 のピクセルの第 1 の列和ベクトルを決定し得る。第 1 の列和ベクトルの特定の要素は、第 1 の境界ボックス 116 の特定の列に対応するピクセル値の和に対応し得る。

#### 【0060】

[0066] 方法 600 はまた、604 において、第 1 のピクセルの第 1 の列和差分ベクトルを決定することを含む。例えば、図 1 のスタビライザー 102 は、図 1 - 2 に関して説明されたように、第 1 の境界ボックス 116 の第 1 のピクセルの第 1 の列差分ベクトルを決定し得る。第 1 の列和差分ベクトルの特定の要素は、第 1 の列和ベクトルの第 1 の要素と、第 1 の列和ベクトルの第 2 の要素間の差分に対応し得る。第 1 の要素および第 2 の要素は、第 1 の境界ボックス 116 の近接した列に対応し得る。10

#### 【0061】

[0067] 方法 600 はさらに、606 において、第 1 のピクセルの第 1 の行和ベクトルを決定することを含む。例えば、図 1 のスタビライザー 102 は、図 1 に関して説明されたように、第 1 の境界ボックス 116 の第 1 のピクセルの第 1 の行和ベクトルを決定し得る。第 1 の行和ベクトルの特定の要素は、第 1 の境界ボックス 116 の特定の行に対応するピクセル値の和に対応し得る。20

#### 【0062】

[0068] 方法 600 はまた、608 において、第 1 のピクセルの第 1 の行和差分ベクトルを決定することを含む。例えば、図 1 のスタビライザー 102 は、図 1 - 2 に関して説明されたように、第 1 の境界ボックス 116 の第 1 のピクセルの第 1 の行和差分ベクトルを決定し得る。第 1 の行和差分ベクトルの特定の要素は、第 1 の行和ベクトルの第 1 の要素と、行和ベクトルの第 2 の要素間の差分に対応し得る。第 1 の要素および第 2 の要素は、第 1 の境界ボックス 116 の近接した行に対応し得る。

#### 【0063】

[0069] 方法 600 はさらに、610 において、第 1 のピクセルの第 1 の列和ベクトル、および特定の検索ピクセルの第 2 の列和ベクトルの第 1 の差分絶対値和 ( S A D )、第 1 のピクセルの第 1 の列和差分ベクトル、および特定の検索ピクセルの第 2 の列和差分ベクトルの第 2 の S A D、第 1 のピクセルの第 1 の行和ベクトル、および特定の検索ピクセルの第 2 の行和ベクトルの第 3 の S A D、ならびに第 1 のピクセルの第 1 の行和差分ベクトル、および特定の検索ピクセルの第 2 の行和差分ベクトルの第 4 の S A D を加算することによって、特定の検索境界ボックス内の第 1 のピクセル、および特定の検索ピクセルのための特定の類似のメトリックを計算することを含む。例えば、図 1 のスタビライザー 102 は、図 1 - 2 および 4 に関して説明されたように、第 1 の境界ボックス 116 の第 1 のピクセル、および特定の検索境界ボックス内の特定の検索ピクセルのための特定の類似のメトリックを計算し得る。30

#### 【0064】

[0070] 様々な実施形態が列和ベクトル、列和差分ベクトル、行和ベクトル、および行和差分ベクトルを利用するように説明されたが、これは、例のみのためであり、限定すると見なされないことに留意されるべきである。代替の実施形態では、類似性は、より少ない、より多い、または異なる計算およびデータ構造に基づいて決定され得る。

#### 【0065】

[0071] 図 6 の方法 600 は、フィールドプログラマブルゲートアレイ ( F P G A ) デバイス、特定用途向け集積回路 ( A S I C )、例えば中央処理装置 ( C P U ) 等の処理装置、デジタルシグナルプロセッサ ( D S P )、コントローラ、別のハードウェアデバイス、ファームウェアデバイス、またはそれらの任意の組み合わせによってインプリメントさ40

れ得る。一例として、図6の方法600は、図7に関して説明されるように、命令を実行するプロセッサによって行われることができる。

#### 【0066】

[0072] 図7を参照すると、デバイス（例えば、無線通信デバイス）の特定の例示的な実施形態のブロック図が、描写され、概して700と示される。デバイス700は、メモリ732に結合される、例えばデジタルシグナルプロセッサ（DSP）、または中央処理装置（CPU）等のプロセッサ710を含む。プロセッサ710は、図1のスタビライザー102および図1のトラッカー160、またはそれら両方を含み得る、および／または実行し得る。

#### 【0067】

[0073] デバイス700の1つ以上のコンポーネントは、1つ以上のタスクを行うための命令を実行するプロセッサによって、専用ハードウェア（例えば、電気回路）を介して、あるいはそれらの組み合わせでインプリメントされ得る。一例として、メモリ732、またはスタビライザー102および／またはトラッカー160の1つ以上のコンポーネントは、例えばランダムアクセスメモリ（RAM）、磁気抵抗ランダムアクセスメモリ（MRAM）、スピントルクトランスファMRAM（spin-torque transfer MRAM）（STT-MRAM）、フラッシュメモリ、読み取り専用メモリ（ROM）、プログラマブル読み取り専用メモリ（PROM）、消去可能プログラマブル読み取り専用メモリ（EPROM（登録商標））、電気的消去可能プログラマブル読み取り専用メモリ（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、またはコンパクトディスク読み取り専用メモリ（CD-ROM）等のメモリデバイスであり得る。メモリデバイスは、コンピュータ（例えば、プロセッサ710）によって実行されるとき、コンピュータに、図5の方法500、図6の方法600、またはそれらの組み合わせの少なくとも一部を行わせ得る命令を含み得る。一例として、メモリ732、またはスタビライザー102の1つ以上のコンポーネントは、コンピュータ（例えば、プロセッサ710）によって実行されるとき、コンピュータに、図5の方法500、図6の方法600、またはそれらの組み合わせの少なくとも一部を行わせる命令を含む非一時的コンピュータ可読媒体であり得る。

#### 【0068】

[0074] 図7はまた、プロセッサ710に、およびディスプレイ728に結合されるディスプレイコントローラ726を示す。例えば、図1のカメラ112等のカメラは、プロセッサ710に結合され得る。コーダ／デコーダ（CODEC）734もまた、プロセッサ710に結合されることができる。スピーカ736およびマイクロフォン738は、CODEC734に結合されることができる。

#### 【0069】

[0075] 図7はまた、無線コントローラ740が、プロセッサ710に、および無線アンテナ742に結合されることを示す。特定の実施形態では、プロセッサ710、ディスプレイコントローラ726、メモリ732、CODEC734、および無線コントローラ740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722に含まれる。特定の実施形態では、カメラ112、入力デバイス730および電源744は、システムオンチップデバイス722に結合される。さらに、特定の実施形態では、図7に例示されるように、ディスプレイ728、カメラ112、スタビライザー102、トラッカー160、入力デバイス730、スピーカ736、マイクロフォン738、無線アンテナ742、および電源744は、システムオンチップデバイス722の外部にある。しかしながら、ディスプレイ728、カメラ112、スタビライザー102、トラッカー160、入力デバイス730、スピーカ736、マイクロフォン738、無線アンテナ742、および電源744の各々は、例えばインタフェースまたはコントローラ等の、システムオンチップデバイス722のコンポーネントに結合されることがある。

#### 【0070】

[0076] 説明された実施形態と併せて、第1のデータを受信するための手段を含むシステムは、開示される。第1のデータは、画像のシーケンスの第1の画像のための第1の境

10

20

30

40

50

界ボックスを定義し得る。第1の境界ボックスは、トラックされた対象を含む関心領域に対応し得る。受信するための手段は、図7の入力デバイス730、境界ボックスを定義するデータを受信するように構成される1つ以上の他のデバイスまたは回路（例えば、モバイル電話のタッチスクリーン）、あるいはそれらの任意の組み合わせを含み得る。

#### 【0071】

[0077] システムはまた、対象トラッキングデータを生成するための手段を含み得る。対象トラッキングデータは、画像のシーケンスの第2の画像に対応し得る。対象トラッキングデータは、第2の境界ボックスを定義し得る。第2の境界ボックスは、第2の画像の中のトラックされた対象を含む関心領域に対応し得る。生成するための手段は、図1および図7のトラッカー160、対象トラッキングデータを生成するように構成される1つ以上の他のデバイスまたは回路、あるいはそれらの任意の組み合わせを含み得る。10

#### 【0072】

[0078] システムはさらに、類似のメトリック、および類似のメトリックに基づいて修正された第2の境界ボックスを決定するための手段を含み得る。類似のメトリックは、第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのために決定され得る。検索境界ボックスの各々の検索座標は、1つ以上の方に向シフトされる第2の境界ボックスの第2の座標に対応し得る。決定するための手段は、図7のプロセッサ710、図1および7のスタビライザー102、類似のメトリックおよび修正された境界ボックスを決定するように構成される1つ以上の他のデバイスまたは回路、あるいはそれらの任意の組み合わせを含み得る。20

#### 【0073】

[0079] 当業者は、本明細書に開示された実施形態に関連して説明された様々な例示的な論理ロック、構成、モジュール、回路、およびアルゴリズムステップが、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、またはそれら両方の組み合わせとしてインプリメントされ得ることをさらに理解することとなる。様々な例示的なコンポーネント、ロック、構成、モジュール、回路、およびステップは、一般的にそれらの機能性の観点から上記に説明された。そのような機能性が、ハードウェアとして、またはプロセッサが実行可能な命令としてインプリメントされるか否かは、特定のアプリケーションおよび全体のシステムに課せられる設計の制約に依存する。当業者は、各特定のアプリケーションに関して様々な方法で説明された機能性をインプリメントし得るが、そのようなインプリメンテーションの決定は、本開示の範囲からの逸脱をもたらすと解釈されるべきではない。30

#### 【0074】

[0080] 本明細書に開示された実施形態に関連して説明された方法またはアルゴリズムのステップは、直接ハードウェアにおいて、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において、またはこれら2つの組み合わせにおいて、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、読み取り専用メモリ(ROM)、プログラマブル読み取り専用メモリ( PROM)、消去可能なプログラマブル読み取り専用メモリ(E PROM)、電気的に消去可能なプログラマブル読み取り専用メモリ(EEPROM)、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み取り専用メモリ(CD-ROM)、または当技術分野において既知の非一時的記憶媒体の任意の他の形式内に存在し得る。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、および記憶媒体に情報を書き込むことができるようなプロセッサに結合される。代替では、記憶媒体は、プロセッサと一体化され得る。プロセッサおよび記憶媒体は、特定用途向け集積回路(ASSIC)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に存在し得る。代替では、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の中の個別のコンポーネントとして存在し得る。40

#### 【0075】

[0081] 開示された実施形態の先の説明は、開示された実施形態を製造する、または使50

用することを当業者に可能にするために提供される。これらの実施形態への様々な修正は、当業者に容易に明らかであり、本明細書に定義された原理は、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。ゆえに、本開示は、本明細書に示された実施形態に限定されるようには意図されず、下記の特許請求の範囲によって定義されるような原理および新規な特徴と一致する可能な限り最も広い範囲を与えられることとなる。

以下に、本願出願の当初の特許請求の範囲に記載された発明を付記する。

[ C 1 ]

方法であって、

画像のシーケンスの第1の画像のための第1の境界ボックスを定義する第1のデータを受信することと、ここにおいて、前記第1の境界ボックスは、トラックされた対象を含む関心領域に対応し、10

画像の前記シーケンスの第2の画像に関する対象トラッキングデータを受信することと、前記対象トラッキングデータは、第2の境界ボックスを定義し、ここにおいて、前記第2の境界ボックスは、前記第2の画像の中の前記トラックされた対象を含む前記関心領域に対応し、

前記第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのための類似のメトリックを決定することと、ここにおいて、前記検索境界ボックスの各々の検索座標は、1つ以上の方にシフトされる前記第2の境界ボックスの第2の座標に対応し、

前記類似のメトリックに基づいて、修正された第2の境界ボックスを決定することと、20  
を備える、方法。

[ C 2 ]

前記第1のデータは、前記第1の境界ボックスの第1の座標および第1のディメンションを含み、前記対象トラッキングデータは、前記第2の境界ボックスの前記第2の座標および第2のディメンションを含む、C 1に記載の方法。

[ C 3 ]

前記検索境界ボックスの各々の検索ディメンションは、前記第1の境界ボックスの前記第1のディメンションに対応する、C 2に記載の方法。

[ C 4 ]

前記類似のメトリックは、差分絶対値和(S A D)メトリックを含む、C 1に記載の方法。30

[ C 5 ]

特定の検索境界ボックス内の前記第1のピクセル、および特定の検索ピクセルのための特定の類似のメトリックに基づいて前記検索境界ボックスの前記特定の検索境界ボックスを選択すること、

をさらに備え、

前記修正された第2の境界ボックスの修正された第2の座標は、前記特定の検索境界ボックスの特定の検索座標に対応する、C 1に記載の方法。

[ C 6 ]

前記第1のピクセルに対する第1のピクセル特性、および前記特定の検索ピクセルに対する第2のピクセル特性の差分絶対値和(S A D)に少なくとも部分的に基づいて前記特定の類似のメトリックを計算すること、40

をさらに備える、C 5に記載の方法。

[ C 7 ]

前記第1のピクセルの第1の列和ベクトル、および前記特定の検索ピクセルの第2の列和ベクトルの第1の差分絶対値和(S A D)と、

前記第1のピクセルの第1の列和差分ベクトル、および前記特定の検索ピクセルの第2の列和差分ベクトルの第2のS A Dと、

前記第1のピクセルの第1の行和ベクトル、および前記特定の検索ピクセルの第2の行和ベクトルの第3のS A Dと、50

前記第1のピクセルの第1の行和差分ベクトル、および前記特定の検索ピクセルの第2の行和差分ベクトルの第4のS A Dと、

を加算することによって前記特定の類似のメトリックを計算すること、  
をさらに備える、C 5に記載の方法。

## [ C 8 ]

前記第1のピクセルの列和ベクトルを決定すること、ここにおいて、前記列和ベクトルの特定の要素は、前記第1の境界ボックスの特定の列に対応するピクセル値の和に対応し  
、  
をさらに備える、C 1に記載の方法。

## [ C 9 ]

10

前記ピクセル値の各々は、対応するピクセルのピクセル輝度を識別する、C 8に記載の  
方法。

## [ C 10 ]

前記第1のピクセルの列和差分ベクトルを決定すること、ここにおいて、前記列和差分  
ベクトルの特定の要素は、前記列和ベクトルの第1の要素と、前記列和ベクトルの第2の  
要素間の差分に対応し、前記第1の要素および前記第2の要素は、前記第1の境界ボックス  
の近接した列に対応し、  
をさらに備える、C 1に記載の方法。

## [ C 11 ]

20

前記第1のピクセルの行和ベクトルを決定すること、ここにおいて、前記行和ベクトル  
の特定の要素は、前記第1の境界ボックスの特定の行に対応するピクセル値の和に対応し  
、  
をさらに備える、C 1に記載の方法。

## [ C 12 ]

前記第1のピクセルの行和差分ベクトルを決定すること、ここにおいて、前記行和差分  
ベクトルの特定の要素は、前記行和ベクトルの第1の要素と、前記行和ベクトルの第2の  
要素間の差分に対応し、前記第1の要素および前記第2の要素は、前記第1の境界ボックス  
の近接した行に対応し、  
をさらに備える、C 1に記載の方法。

## [ C 13 ]

30

複数の画像に対応するメディアンディメンションを決定すること、ここにおいて、前記  
複数の画像は、画像の前記シーケンスの中の前記第2の画像に先行し、前記修正された第  
2の境界ボックスの第2のディメンションは、前記メディアンディメンションに対応し、  
をさらに備える、C 1に記載の方法。

## [ C 14 ]

装置であって、

命令を記憶するように構成されたメモリと、

第1の画像の第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックス  
の各々内の検索ピクセルのための類似のメトリックを決定することと、ここにおいて、  
前記第1の境界ボックスは、トラックされた対象を含む関心領域に対応し、前記検索境界  
ボックスの各々の検索座標は、1つ以上の方にシフトされる第2の境界ボックスの第2  
の座標に対応し、前記第1の画像は、画像のシーケンスの中の第2の画像に先行し、前記  
第2の境界ボックスは、前記第2の画像の中の前記トラックされた対象を含む前記関心領域  
に対応し、  
前記類似のメトリックに基づいて、修正された第2の境界ボックスを決定することと  
、

40

を行うための前記命令を実行するように構成されたプロセッサと、  
を備える、装置。

## [ C 15 ]

前記プロセッサは、

特定の検索境界ボックス内の前記第1のピクセル、および特定の検索ピクセルのための  
特定の類似のメトリックに基づいて前記検索境界ボックスの前記特定の検索境界ボックス

50

を選択するようにさらに構成され、

前記修正された第2の境界ボックスの修正された第2の座標は、前記特定の検索境界ボックスの特定の検索座標に対応する、C14に記載の装置。

[ C 1 6 ]

前記プロセッサは、前記第1のピクセルの第1のピクセル特性、および前記特定の検索ピクセルの第2のピクセル特性の差分絶対値和(S A D)に少なくとも部分的に基づいて前記特定の類似のメトリックを計算するようにさらに構成される、C15に記載の装置。

[ C 1 7 ]

前記プロセッサは、前記第1のピクセルの第1のピクセル輝度、および前記特定の検索ピクセルの第2のピクセル輝度に少なくとも部分的に基づいて前記特定の類似のメトリックを計算するようにさらに構成される、C15に記載の装置。 10

[ C 1 8 ]

画像の前記シーケンスは、ビデオストリームに対応する、C14に記載の装置。

[ C 1 9 ]

前記プロセッサは、前記第1の画像のための第1のデータを取得するように、および前記第2の画像に関する対象トラッキングデータを取得するようにさらに構成され、

ここにおいて、前記第1のデータは、前記第1の境界ボックスを定義し、

前記対象トラッキングデータは、前記第2の境界ボックスを定義する、C14に記載の装置。

[ C 2 0 ]

20

前記第1のデータは、前記第1の境界ボックスの第1の座標および第1のディメンションを含み、前記対象トラッキングデータは、前記第2の境界ボックスの前記第2の座標および第2のディメンションを含む、C19に記載の装置。

[ C 2 1 ]

前記検索境界ボックスの各々の検索ディメンションは、前記第1の境界ボックスの第1のディメンションに対応する、C14に記載の装置。

[ C 2 2 ]

前記プロセッサは、複数の画像に対応するメディアンディメンションを決定すること、ここにおいて、前記複数の画像は、画像の前記シーケンスの中の前記第2の画像に先行し、前記修正された第2の境界ボックスの第2のディメンションは、前記メディアンディメンションに対応し、を行うようにさらに構成される、C14に記載の装置。 30

[ C 2 3 ]

命令を記憶するコンピュータ可読記憶デバイスであって、前記命令がプロセッサによって実行されるとき、前記プロセッサに、

第1の画像の第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのための類似のメトリックを決定することと、ここにおいて、前記第1の境界ボックスは、トラックされた対象を含む関心領域に対応し、前記検索境界ボックスの各々の検索座標は、1つ以上の方向にシフトされる第2の境界ボックスの第2の座標に対応し、前記第1の画像は、画像のシーケンスの中の第2の画像に先行し、前記第2の境界ボックスは、前記第2の画像の中の前記トラックされた対象を含む前記関心領域に対応し、 40

前記類似のメトリックに基づいて、修正された第2の境界ボックスを決定することと、を備えるオペレーションを行わせる、コンピュータ可読記憶デバイス。

[ C 2 4 ]

前記オペレーションは、複数の画像に対応するメディアンディメンションを決定すること、ここにおいて、前記複数の画像は、画像の前記シーケンスの中の前記第2の画像に先行し、前記修正された第2の境界ボックスの第2のディメンションは、前記メディアンディメンションに対応し、をさらに備える、C23に記載のコンピュータ可読記憶デバイス。

[ C 2 5 ]

50

前記オペレーションは、

前記第1の画像のための第1のデータを受信することと、ここにおいて、前記第1のデータは、前記第1の境界ボックスを定義し、

前記第2の画像に関する対象トラッキングデータを受信することと、ここにおいて、前記対象トラッキングデータは、前記第2の境界ボックスを定義し、

をさらに備える、C 2 3 に記載のコンピュータ可読記憶デバイス。

[ C 2 6 ]

前記第1のデータは、前記第1の境界ボックスの第1の座標および第1のディメンションを含み、前記対象トラッキングデータは、前記第2の境界ボックスの前記第2の座標および第2のディメンションを含む、C 2 5 に記載のコンピュータ可読記憶デバイス。 10

[ C 2 7 ]

前記検索境界ボックスの各々の検索ディメンションは、前記第1の境界ボックスの第1のディメンションに対応する、C 2 3 に記載のコンピュータ可読記憶デバイス。

[ C 2 8 ]

前記オペレーションは、

特定の検索境界ボックス内の前記第1のピクセル、および特定の検索ピクセルのための特定の類似のメトリックに基づいて前記検索境界ボックスの前記特定の検索境界ボックスを選択すること、

ここにおいて、前記類似のメトリックは、前記第1のピクセルの第1のピクセル輝度、および前記特定の検索ピクセルの第2のピクセル輝度に少なくとも部分的に基づいて決定され、 20

をさらに備え、

前記修正された第2の境界ボックスの修正された第2の座標は、前記特定の検索境界ボックスの特定の検索座標に対応する、C 2 3 に記載のコンピュータ可読記憶デバイス。

[ C 2 9 ]

装置であって、

画像のシーケンスの第1の画像のための第1の境界ボックスを定義する第1のデータを受信するための手段と、ここにおいて、前記第1の境界ボックスは、トラックされた対象を含む関心領域に対応し、

画像の前記シーケンスの第2の画像に関する対象トラッキングデータを生成するための手段と、前記対象トラッキングデータは、第2の境界ボックスを定義し、ここにおいて、前記第2の境界ボックスは、前記第2の画像の中の前記トラックされた対象を含む前記関心領域に対応し、 30

前記第1の境界ボックス内の第1のピクセル、および複数の検索境界ボックスの各々内の検索ピクセルのための類似のメトリックと、ここにおいて、前記検索境界ボックスの各々の検索座標は、1つ以上の方向にシフトされる前記第2の境界ボックスの第2の座標に対応し、

前記類似のメトリックに基づいて、修正された第2の境界ボックスと、

を決定するための手段と、

を含む、装置。 40

[ C 3 0 ]

受信するための前記手段、生成するための前記手段、および決定するための前記手段は、モバイル電話、セットトップボックス、ミュージックプレーヤ、ビデオプレーヤ、エンターテイメントユニット、ナビゲーションデバイス、通信デバイス、携帯情報端末(PDA)、固定ロケーションデータユニット、またはコンピュータのうちの少なくとも1つに組み込まれる、C 2 9 に記載の装置。

【図1】

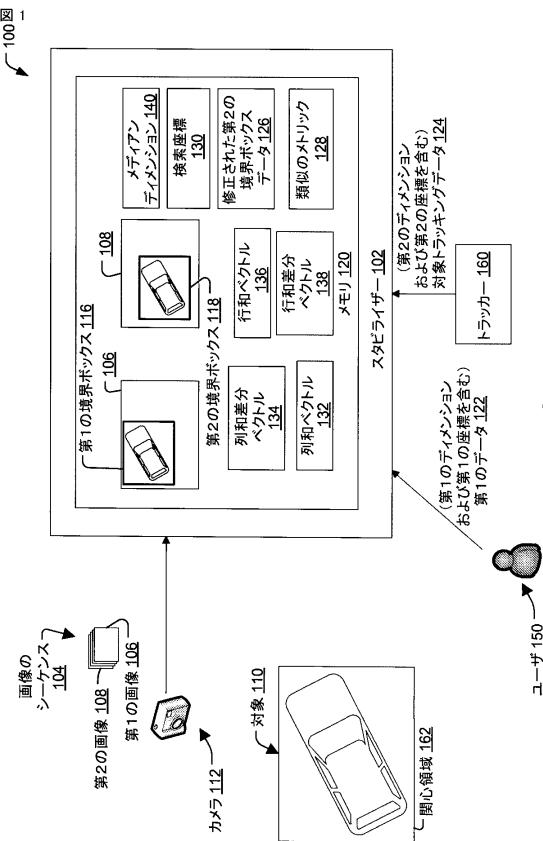


FIG. 1

【図2】

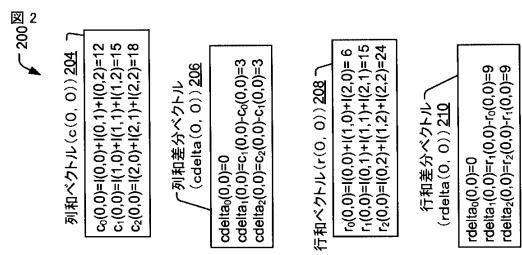


FIG. 2

【図3】

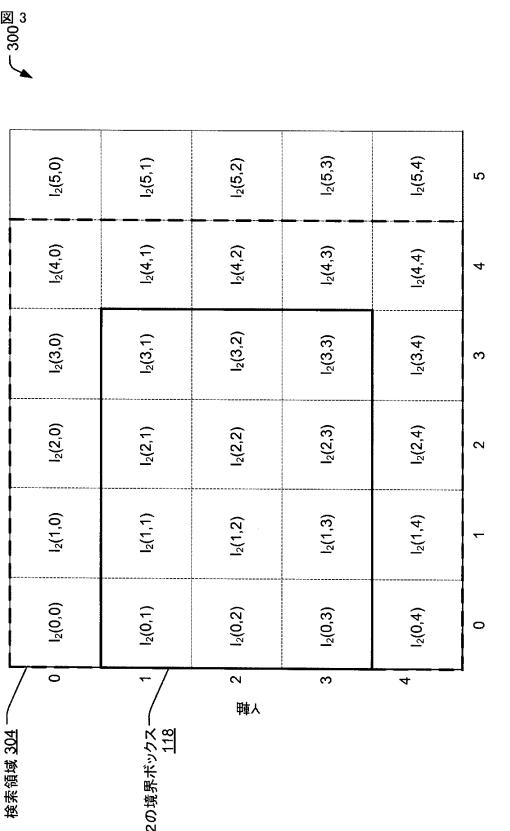


FIG. 3

【図4】

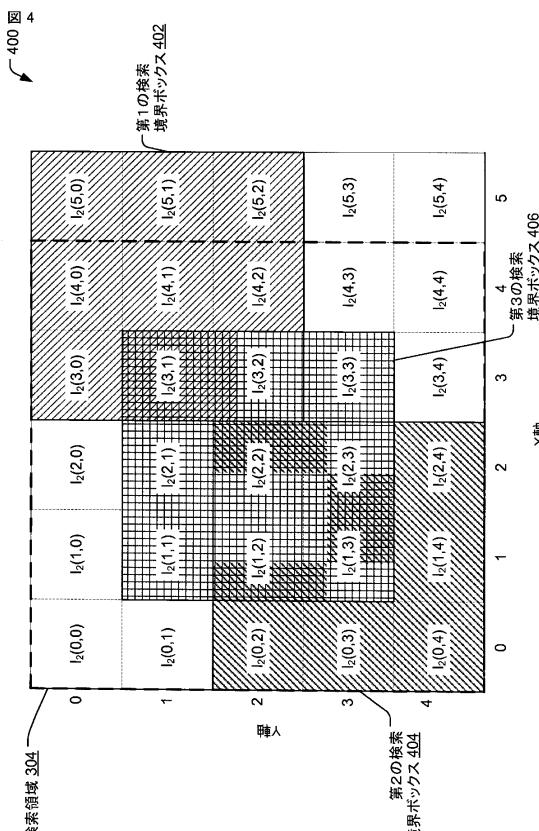


FIG. 4

【図5】

図5

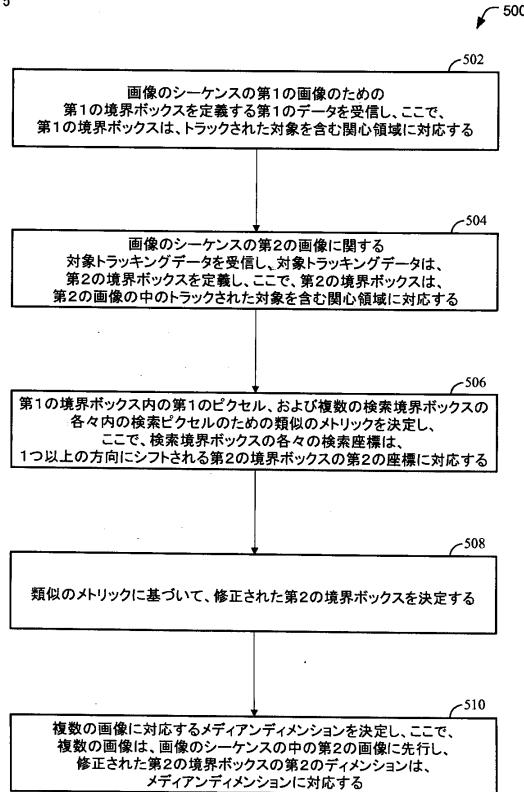


FIG. 5

【図6】

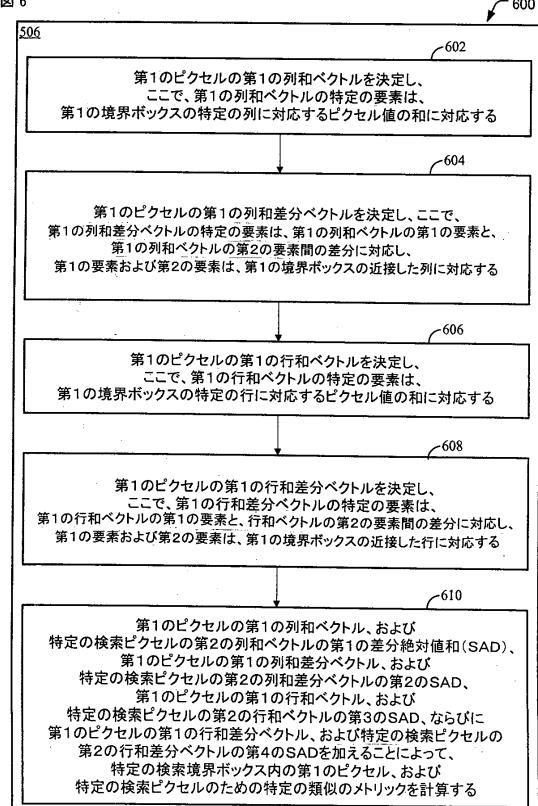


FIG. 6

【図7】

図7

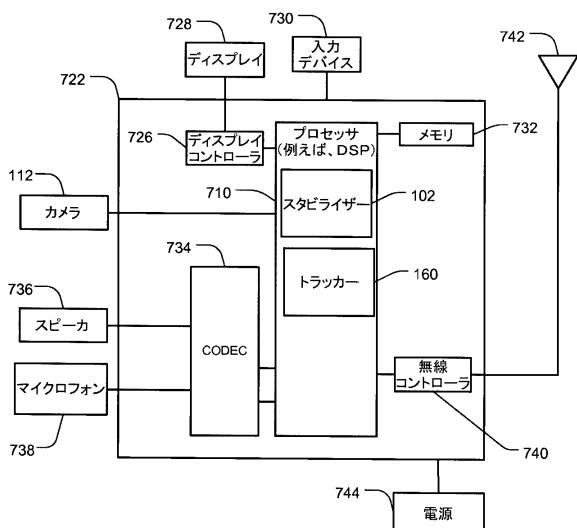


FIG. 7

---

フロントページの続き

(72)発明者 リ、クリストファー

アメリカ合衆国、カリフォルニア州 94041、マウンテン・ビュー、チーカ・アベニュー  
210

(72)発明者 ジョン、シン

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドラ  
イブ 5775

(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 92121-1714、サン・ディエゴ、モアハウス・ドラ  
イブ 5775

(72)発明者 チ、インヨン

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドラ  
イブ 5775

(72)発明者 グオ、カイ

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドラ  
イブ 5775

審査官 大西 宏

(56)参考文献 特開2011-160379(JP,A)

特表2009-522591(JP,A)

米国特許出願公開第2010/073502(US,A1)

(58)調査した分野(Int.Cl., DB名)

H04N 5/222-5/257

G06T 7/00 - 7/90