

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第5335574号  
(P5335574)

(45) 発行日 平成25年11月6日(2013.11.6)

(24) 登録日 平成25年8月9日(2013.8.9)

(51) Int.Cl.

F 1

|             |              |                  |        |       |         |
|-------------|--------------|------------------|--------|-------|---------|
| <b>G06T</b> | <b>7/20</b>  | <b>(2006.01)</b> | GO 6 T | 7/20  | B       |
| <b>HO4N</b> | <b>5/232</b> | <b>(2006.01)</b> | HO 4 N | 5/232 | Z       |
| <b>G06T</b> | <b>1/00</b>  | <b>(2006.01)</b> | GO 6 T | 1/00  | 3 1 5   |
| <b>G06T</b> | <b>7/60</b>  | <b>(2006.01)</b> | GO 6 T | 7/60  | 1 5 O S |

請求項の数 9 (全 21 頁)

|           |                              |
|-----------|------------------------------|
| (21) 出願番号 | 特願2009-145822 (P2009-145822) |
| (22) 出願日  | 平成21年6月18日 (2009.6.18)       |
| (65) 公開番号 | 特開2011-3029 (P2011-3029A)    |
| (43) 公開日  | 平成23年1月6日 (2011.1.6)         |
| 審査請求日     | 平成24年6月18日 (2012.6.18)       |

|           |  |
|-----------|--|
| (73) 特許権者 | 000001007<br>キヤノン株式会社<br>東京都大田区下丸子3丁目30番2号 |
| (74) 代理人  | 100076428<br>弁理士 大塚 康徳                     |
| (74) 代理人  | 100112508<br>弁理士 高柳 司郎                     |
| (74) 代理人  | 100115071<br>弁理士 大塚 康弘                     |
| (74) 代理人  | 100116894<br>弁理士 木村 秀二                     |
| (74) 代理人  | 100130409<br>弁理士 下山 治                      |
| (74) 代理人  | 100134175<br>弁理士 永川 行光                     |

最終頁に続く

(54) 【発明の名称】 画像処理装置及びその制御方法

## (57) 【特許請求の範囲】

## 【請求項 1】

撮像手段によって撮像された複数の撮像画像を入力する画像取得手段と、  
前記入力された複数の撮像画像の各々に対して、各画素の属性に従って複数の領域を抽出する領域抽出手段と、

前記領域抽出手段により抽出された複数の領域の各々の属性に従って、撮像画像間で対応する領域を決定する領域対応付け手段と、

前記対応する領域の画像内の特徴点の座標に基づいて特徴点の3次元位置を推定することにより、前記領域の形状を推定する領域形状推定手段と、

前記対応する領域における各特徴点の3次元位置に基づいて前記撮像手段の3次元運動パラメータ、または、前記対応する領域の3次元運動パラメータを計算することにより、前記対応する領域の剛体運動を推定する領域剛体運動推定手段と、

前記対応する領域の剛体運動の精度として、第1の評価値を算出する第1の算出手段と、

前記対応する領域を他の領域と統合した場合における、前記統合された領域について推定された剛体運動の精度として、第2の評価値を算出する第2の算出手段と、

前記第2の評価値と、前記第1の評価値とに基づき、統合した場合の方が剛体運動の精度が高いと判定された場合に、前記対応する領域と前記他の領域とを統合する領域変更手段と、

を備えることを特徴とする画像処理装置。

10

20

**【請求項 2】**

前記領域変更手段は、前記複数の領域のうち 1 つの領域に関して推定された剛体運動の精度が所定のレベルを下回った場合に、前記 1 つの領域を複数の領域に分離することを特徴とする請求項 1 に記載の画像処理装置。

**【請求項 3】**

前記領域変更手段により変更された領域について、前記領域形状推定手段と前記領域剛体運動推定手段と前記領域変更手段を繰り返し機能させる繰り返し制御手段を更に備え、

前記繰り返し制御手段は、繰り返しの前後における前記領域形状推定手段による推定結果の差が所定の推定結果に関する閾値より小さくなった場合に前記繰り返しを停止することを特徴とする請求項 2 に記載の画像処理装置。 10

**【請求項 4】**

前記繰り返し手段は、繰り返しの前後における前記領域変更手段による変更後の領域の個数の差が所定の領域の個数に関する閾値よりも小さくなった場合に前記繰り返しを停止することを特徴とする請求項 3 に記載の画像処理装置。

**【請求項 5】**

入力手段と、抽出手段と、領域対応付け手段と、形状推定手段と、運動推定手段と、第 1 の算出手段と、第 2 の算出手段と、領域変更手段とを備える画像処理装置の制御方法であって、

前記入力手段が、撮像手段によって撮像された複数の撮像画像を入力する入力工程と、

前記抽出手段が、前記入力された複数の撮像画像の各々に対して、各画素の属性に従つて複数の領域を抽出する抽出工程と、 20

前記領域対応付け手段が、前記抽出された複数の領域の各々の属性に従つて、撮像画像間で対応する領域を決定する領域対応付け工程と、

前記形状推定手段が、前記対応する領域の画像内の特徴点の座標に基づいて特徴点の 3 次元位置を推定することにより、前記領域の形状を推定する形状推定工程と、

前記運動推定手段が、前記対応する領域における各特徴点の 3 次元位置に基づいて前記撮像手段の 3 次元運動パラメータ、または、前記対応する領域の 3 次元運動パラメータを計算することにより、前記対応する領域の剛体運動を推定する運動推定工程と、前記第 1 の算出手段が、前記対応する領域の剛体運動の精度として、第 1 の評価値を算出する第 1 の算出工程と、 30

前記第 2 の算出手段が、前記対応する領域を他の領域と統合した場合における、前記統合された領域について推定された剛体運動の精度として、第 2 の評価値を算出する第 2 の算出工程と、

前記領域変更手段が、前記第 2 の評価値と前記第 1 の評価値とに基づき、統合した場合の方が剛体運動の精度が高いと判定された場合に、前記対応する領域と前記他の領域とを統合する領域変更工程と、

を備えることを特徴とする画像処理装置の制御方法。

**【請求項 6】**

請求項 5 に記載の画像処理装置の制御方法の各工程をコンピュータに実行させるためのプログラム。 40

**【請求項 7】**

請求項 6 のプログラムを記録したコンピュータ読み取り可能な記録媒体。

**【請求項 8】**

複数の撮像画像を入力する入力手段と、

前記入力された複数の撮像画像の各々に対して、各画素の属性に従つて複数の領域を抽出する抽出手段と、

前記抽出された複数の領域の各々の属性に従つて、撮像画像間で対応する領域を決定する領域対応付け手段と、

前記対応する領域の画像内の特徴点の座標に基づいて特徴点の 3 次元位置を推定することにより、前記領域の形状を推定する形状推定手段と、

10

20

30

40

50

前記対応する領域における各特徴点の3次元位置に基づいて運動を計算することにより、前記対応する領域の剛体運動を推定する運動推定手段と、  
前記複数の領域のうちの1つの領域と他の領域とを統合した場合に推定される剛体運動の精度が、前記1つの領域と前記他の領域とのそれぞれについて前記推定された剛体運動の精度の平均よりも高いと判定された場合に、前記1つの領域と前記他の領域とを統合する領域変更手段と、  
を備えることを特徴とする画像処理装置。

## 【請求項9】

前記領域変更手段は、1つの領域についての前記剛体運動の精度の時間変化が閾値を超えた場合に、前記1つの領域を複数の領域へ分離することを特徴とする請求項8に記載の画像処理装置。10

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、撮像画像列に含まれる移動する撮影対象を含む領域を検出する画像処理装置および画像処理方法に関する。20

## 【背景技術】

## 【0002】

撮影対象のサイズが撮像装置から撮影対象までの距離と比較して十分小さい場合、または撮像装置（以下、カメラと称す）の移動量が撮影対象までの距離より十分小さい場合は、観察する対象をほぼ平面として見ることができる。つまり、撮影対象の空間的な広がりに対して撮影装置の移動による変化が小さい場合には、観察している対象の変化が小さいという前提で複数の射影近似が利用できる。射影近似には、中心射影を線形近似した弱中心射影や擬似中心射影、さらには平行投影などがある。20

## 【0003】

非特許文献1によれば、射影近似において画像中の特徴点の3次元位置は、透視投影の計算を線形化して表現することができる。ここでは、静止したカメラ座標系を世界座標系と同一視し、XY面を画像面、Z軸をカメラの光軸とみなす。物体に任意に物体座標系を固定し、番目の特徴点 $p_\alpha$ の物体座標系に関する座標を $(a_\alpha, b_\alpha, c_\alpha)$ とする。時刻 $t$ での物体座標系の原点の位置ベクトルを $t$ 、各座標基底ベクトルを $\{i_\alpha, j_\alpha, k_\alpha\}$ とすると、特徴点 $p_\alpha$ の時刻 $t$ における位置 $r$ は、次式(1)のように示される。30

## 【0004】

## 【数1】

$$\mathbf{r}_{\alpha} = \mathbf{t}_\alpha + a_\alpha \mathbf{i}_\alpha + b_\alpha \mathbf{j}_\alpha + c_\alpha \mathbf{k}_\alpha \quad \dots \quad (1)$$

## 【0005】

中心射影の近似である平行投影を仮定すると、点 $r = (X, Y, Z)^T$ の画像座標は $(X, Y)$ となる。ベクトル $t$ ,  $\{i_\alpha, j_\alpha, k_\alpha\}$ の投影、すなわちZ軸座標をのぞいた2次元ベクトルを $t'$ ,  $\{i'_\alpha, j'_\alpha, k'_\alpha\}$ とする。それらを $=1, \dots, M$ に渡って縦に並べた2M次元ベクトルを $m_0, m_1, m_2, m_3$ とすると、式(1)で定義した2M次元ベクトル $p$ が式(2)のように表せる。40

## 【0006】

## 【数2】

$$\mathbf{p}_\alpha = \mathbf{m}_0 + a_\alpha \mathbf{m}_1 + b_\alpha \mathbf{m}_2 + c_\alpha \mathbf{m}_3 \quad \dots \quad (2)$$

## 【0007】

各特徴点の移動軌跡は2M次元空間の1点として表せる。そしてN個の点 $p$ が $\{m_0, m_1, m_2, m_3\}$ の張る4次元部分空間に含まれることになる。

## 【0008】

複数物体の分離は、2M次元空間の点集合をお互いに異なる4次元部分空間に分割することになる。

#### 【0009】

非特許文献2では、上述の2M次元空間の分離について因子分解を用いて分離する方法が開示されている。

#### 【0010】

一方、複数物体の分離は画面座標における特徴点の2次元的な分布を用いても分離することができる。非特許文献3では、道路を走行する車両追跡法として特徴点の軌跡群をグラフ分割アルゴリズムによりクラスタリングすることで車両の追跡を行う方法が開示されている。画面上にある特徴点の群をグラフにより表現し、過去フレームにおける追跡情報を拘束条件としてグラフ分割問題として定式化し、複数物体の分離する方法を開示している。10

#### 【0011】

さらに、幾何的な特徴によらず、画素の色情報を用いて領域を検出するセグメンテーションの技術も画面上の物体領域の分離をすることができる。非特許文献4では、グラフカットと呼ばれる物体と背景のセグメンテーションに関する技術が開示されている。画素の集合と近傍のピクセルの集合との関係をグラフで表現し、エッジ上の画素に対してどちらのグラフノードに属するかのコストを計算してどちらの領域に属するかを計算する。

#### 【先行技術文献】

#### 【非特許文献】

20

#### 【0012】

【非特許文献1】「因子分解を用いない因子分解法：複数物体分離」金谷健一、電子情報通信学会技術研究報告 PRMU98-26 pp. 1-8 1998.

【非特許文献2】「A multiple factorization method for motion analysis」 J. Costeira and T. Kanade, Proc. 5th Int. Conf. Computer Vision (ICCV95), pp.1071-1076, 1995

【非特許文献3】「拘束付きグラフ分割を用いたオクルージョンに強い車両追跡」安部、小沢、電子情報通信学会論文誌A、Vol.J90-A No.12 pp.948-959 2007

【非特許文献4】「Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images」 Y. Boykov, M. Jolly, Proc. of International Conference of Computer Vision, vol.1, pp.105-112, 200130

#### 【非特許文献5】「ロボットビジョンの基礎」出口光一郎、コロナ社，2000年6月

#### 【発明の概要】

#### 【発明が解決しようとする課題】

#### 【0013】

非特許文献1、2に述べられている技術は、カメラと撮影対象の関係が撮影画像列の前後では大きく変化しないという前提をもとに、平行射影により特徴点の射影を近似できる範囲においては上手く働く。しかし、実際には画像から観測される撮影対象に属する特徴点の様子には、平行射影で近似できない「見え具合」が生じる。すなわち、撮影対象のサイズとカメラとの位置、もしくは複数の撮影対象とカメラの相対的な運動により、「見え具合」の違いや遮蔽が発生する。特に、カメラが大きく回り込んで撮影する場合や、撮影対象が回転している場合には、特徴点の追跡に失敗することも多くなる。さらに、射影近似による3次元の推定誤差の影響も大きいため、カメラの焦点距離と撮影対象との距離が近接している場合には、物体の形状推定の精度が低くなる。40

#### 【0014】

また射影変換によらず、2次元的な領域の変化を求める非特許文献3の方法では、領域が重なる場合に、再度分離して動く対象を適切に分離することが難しい。それは、領域の3次元的な運動を想定していないため、それぞれの領域の奥行きが違っていても画面座標では同じになるため、オクルージョンが発生した際にそれらの領域が同じ領域として判断されてしまう。50

**【0015】**

非特許文献4では、事前の知識として背景領域と前景領域の色の属性を指示する必要があり、非特許文献3と同様に対象の3次元的な運動を想定していないので、領域が混在する場合にはその領域を分離することが難しくなる。

**【0016】**

このように従来技術では、撮影対象の奥行き方向の空間的な広がりが、カメラの移動量に対して大きい場合、特にカメラと対象空間の距離が近い場合には、領域の対応付けが困難となる。それは、一般的なビデオカメラによる撮影において、撮影者がカメラを手に持つて、前の人や物を撮影した画像を撮影する場合に発生する。その撮影した画像列から領域の対応を求める際に、平行射影では近似できない撮影対象の動きに対して、非特許文献1や2による方法では領域の対応付けがうまく行かない。  
10

**【0017】**

上述の問題点は、撮影対象の空間の3次元的な位置を考慮して、領域の3次元運動を推定しながら対応付けを行うことで、基本的には解決することができる。そこで、本発明は上記問題点に鑑み、カメラと撮影対象が移動する状態での撮影画像列から3次元的な撮影対象の形状と運動を基に撮影対象に対応付けられた領域検出を行うことを目的とする。

**【課題を解決するための手段】****【0018】**

上記の目的を達成する本発明に係る画像処理装置は、

撮像手段によって撮像された複数の撮像画像を入力する画像取得手段と、

20

前記入力された複数の撮像画像の各々に対して、各画素の属性に従って複数の領域を抽出する領域抽出手段と、

前記領域抽出手段により抽出された複数の領域の各々の属性に従って、撮像画像間で対応する領域を決定する領域対応付け手段と、

前記対応する領域の画像内の特徴点の座標に基づいて特徴点の3次元位置を推定することにより、前記領域の形状を推定する領域形状推定手段と、

前記対応する領域における各特徴点の3次元位置に基づいて前記撮像手段の3次元運動パラメータ、または、前記対応する領域の3次元運動パラメータを計算することにより、前記対応する領域の剛体運動を推定する領域剛体運動推定手段と、

前記対応する領域の剛体運動の精度として、第1の評価値を算出する第1の算出手段と

30

前記対応する領域を他の領域と統合した場合における、前記統合された領域について推定された剛体運動の精度として、第2の評価値を算出する第2の算出手段と、

前記第2の評価値と、前記第1の評価値とに基づき、統合した場合の方が剛体運動の精度が高いと判定された場合に、前記対応する領域と前記他の領域とを統合する領域変更手段と、

を備えることを特徴とする。

**【発明の効果】****【0019】**

本発明による画像処理装置により撮像画像から撮影対象に対応付けられた領域検出を行うことができる。

40

**【図面の簡単な説明】****【0020】**

【図1】実施形態1に係る画像処理装置の主要部分の構成を示す図である。

【図2】実施形態1に係る画像処理装置を利用する構成を示す図である。

【図3】実施形態2に係る画像処理装置の主要部分の構成を示す図である。

【図4】実施形態2に係る統合分離制御部の処理手続きのフローを示す図である。

【図5】本発明の他の実施形態に係る画像処理方法を実装したプログラムを利用して撮影対象のデータを共有する構成を示す図である。

**【発明を実施するための形態】**

50

**【0021】**

以下の説明では、本発明の例示的実施形態を開示する。

**【0022】**

## &lt;実施形態1&gt;

まず、本発明の実施形態1について説明する。

## (構成)

図1は、実施形態1に係る、画像処理装置のうちの一つである画像処理装置の主要部の構成例を示すものである。図1に示すように、画像処理装置は、画像取得部10、領域抽出部20、領域対応付け部30、領域形状推定部40、領域剛体運動推定部50および領域統合分離部60から構成される。

10

**【0023】**

画像取得部10は、例えば撮像装置により取得された2以上の画像をメモリに書き込んで画像データとする。また領域抽出部20は、その取得された画像から属性に基づいて領域を抽出する。領域対応付け部30は、その抽出されたそれぞれの領域の対応付けを行い、領域形状推定部40はその結果を用いて領域の形状を推定する。領域剛体運動推定部50は、領域形状推定部40で推定された複数の形状推定結果を用い領域の剛体運動を推定する。また領域統合分離部60は、領域剛体運動推定部50により推定された剛体運動の結果を用いて、複数の領域の統合または分離を行う。これにより、撮像画像から撮影対象に対応付けられた領域検出を行い、撮影対象の位置と姿勢を推定することができる

図2は、実施形態1における、画像処理装置と接続された他の主要機能の構成を示す。本発明の画像処理装置は、図2に示すような画像処理装置100のような形態で利用することができる。画像処理装置100には、撮像装置200からの撮影対象を映した映像が入力される。この画像処理装置100においてこの映像情報から領域情報が生成される。さらに、画像処理装置100には、この領域情報をを利用して画像を合成する画像合成装置300が接続されている。画像合成装置300で生成された合成画像は、画像提示装置400により確認することができる。なお、撮像装置200、画像合成装置300、画像提示装置400は、本発明の実施例を利用する例の一つとしてあげたものであり、画像処理装置の入力と出力に関する機構や信号のフォーマットを規定するものではない。

20

**【0024】**

例えば、撮像装置200としては光・電子変換素子であるCCDやCMOSの半導体素子を利用した装置を使うことができる。なお、撮像装置を構成するレンズには、一般的には光学歪みが存在するが、これらは校正パターンを用いてカメラキャリブレーションを行い、事前に補正值を取得することが可能である。また、一般的にはビデオカメラ等で撮影し、任意の媒体に記録した動画ファイルから撮像画像列を抽出して利用することもできる。さらに、撮影されている画像列を、ネットワークを経由して受信して利用することもできる。

30

**【0025】**

画像合成装置300の例としては、画像信号を利用してコンピュータグラフィックスと合成することができれば良いので、画像入力インターフェースポートを有するパーソナルコンピュータを利用することができます。また、画像合成装置のほか、移動する撮影物体の領域を記録装置に保存して、撮影した画像の情報として利用しても構わない。

40

**【0026】**

図2の構成では、それぞれの装置が別々の機器として分離して構成されているが、入出力用のケーブルで装置間が接続されても構わない。また、プリント基板上に構成されるバスを利用して情報をやりとりしても構わない。例えば、デジタルカメラのように撮像機能と画像を表示する画像表示機能を有する場合に、本発明の画像処理装置の機能を内蔵していくても構わない。

## (画像取得部)

まず、画像取得部10について説明する。ここでは、撮影対象を含む撮像画像とは、2次元の画像として例えばカラー画像を取得するものである。カラー画像とは、画素と呼ば

50

れる単位で構成されており、画素には例えば R G B の色の情報が格納されている。この画素を縦横の 2 次元方向の配列に並べることで実現されている。例えば V G A ( V i d e o G r a p h i c s A r r a y ) サイズのカラー画像は、x 軸（横）方向 640 画素、y 軸（縦）方向 480 画素の 2 次元配列で表現されており、各画素内に、その位置における色の情報が例えば R G B 形式で格納されている。なお、カラー画像の代わりに、白黒画像を用いる場合は、当該白黒画像の画素値は、各撮像素子への光量を示す濃度値となる。

#### 【 0027 】

画像取得部 10 では、撮影対象を含む画像を撮像装置より取得できるようになっていれば良く、実際には画素配列の大きさやカラーの配置や階調数、さらには撮像装置のカメラパラメータは既知とする。

10

#### （領域検出部）

領域検出部 20 は、画像取得部 10 により得られた撮像画像から領域を抽出する処理を行う。ここで言う領域の抽出とは、2 次元画像上の属性が共通している小領域を検出することである。この時点において、領域は、移動している撮影対象の一部なのか、それとも背景の一部なのかは不明である。画素の属性は、色や濃度勾配を利用することができる。その属性に関しては撮影対象の色や柄などに依存する場合もあるので、複数の属性を用いて領域検出をすることが可能である。

#### 【 0028 】

例えば、色が類似する領域を検出するには、画素の色である R G B の色情報を H S V 表色系に変換し、その際の色相情報を利用することで、同じ色を有する隣接領域を検出することができる。一般に、カラーラベリングと呼ばれる画像処理を用いて実現することができる。また、テクスチャ情報を利用することもできる。濃度分布の局所領域の周期性や方向性を抽出する画像特徴量もしくは濃度勾配の値を用いて、同じ柄の領域を検出することができる。

20

#### 【 0029 】

さらに、非特許文献 4 で開示されているグラフカットによる領域検出を利用することで、複数の色の属性を有する領域を一つの領域として検出することが可能である。ここでは、領域検出のために類似した属性を有する隣接画素の範囲が検出できれば良い。従って既存の領域検出が利用できる。

#### （領域対応付け部）

30

領域対応付け部 30 は、前記領域検出部 20 により検出された撮像画像間の画素の領域に対して、画像特徴量を利用して領域の対応付けを決定する。前記の領域検出部 20 により検出された領域の境界付近では照明不足や他の物体による遮蔽による影響で恒常に安定して検出できるものでは無い。そこで、濃度勾配の特徴的な領域を、領域検出部 20 で検出された領域ごとに抽出する。

#### 【 0030 】

例えば、画像特徴量としては局所的な領域の画素の濃度勾配がコーナーの形状を有しているかどうかを利用できる。例えば、H a r r i s オペレータを用いることで、注目画素の近傍領域の濃度勾配を計算し、そのヘッシャン行列の値を用いて濃度勾配の曲率を算出し、角や隅の特徴を有する領域の画像特徴量を算出することができる。また、画像中の濃度勾配として輪郭や線分を検出する S o b e l フィルター、C a n n y フィルター、また G a b o r 特徴量などを用いることで、エッジ成分の画像特徴量を算出することができる。ここでは、画像処理の分野で利用されている画像特徴量の算出方法を利用する。

40

#### 【 0031 】

非特許文献 1, 2 で述べられているように、撮影対象と撮像装置との距離が十分にあって撮影対象の動きが平行移動に近い場合は、撮像画像列の前後では、画像特徴として検出した特徴量の変化は少ないことが前提となっている。

#### 【 0032 】

しかし、一般的には手動による撮影では撮影対象と撮像装置との関係や撮影対象の動きは制約されていないので、このように検出した画像特徴量の比較だけでは画像特徴量を精

50

度良く算出することは難しい。そこで、複数の画像特徴量を用いて領域の初期の対応を求めることがある。

### 【0033】

時刻における番目の領域 $L$ に属する番目の特徴点 $P$ の画面座標を $X = \{X_r, Y_g, Y_b\}$ とする。撮像画像の座標ベクトル $X$ における濃度を $I(X)$ で表すと、特徴点における局所画像の各画素の濃度値は $I(X + s)$ で表せる。ここで、 $s$ は局所画像の範囲を示すもので、 $s = \{sx, sy\}$ とする。 $s_x$ は特徴点の $X$ 軸方向の相対位置で、 $s_y$ は $Y$ 軸方向の相対位置である。撮影画像がカラー画像であるとすると、局所画像の濃度 $I$ には、赤、緑、青の要素が含まれ、 $I(X + s) = \{I(X + s)_r, I(X + s)_g, I(X + s)_b\}$ となる。領域 $L$ に対して、 $N$ 個の特徴点が検出されたとすると、 $= 1, \dots, N$ となる。局所画像の範囲 $s$ を $-S$ から $S$ とし、番目の領域 $L$ に含まれる色情報の平均を計算することとする。赤の成分の平均 $L_r$ は式(3)、緑の成分 $L_g$ は式(4)、青の成分 $L_b$ は式(5)のようになる。

### 【0034】

#### 【数3】

$$L_{\phi k r} = \frac{1}{4S^2 N} \sum_{a=1}^N \sum_{sx=-S}^S \sum_{sy=-S}^S I(\mathbf{X}_{\phi k a} + \Delta s)_r \quad \dots \quad (3)$$

10

### 【0035】

#### 【数4】

$$L_{\phi k g} = \frac{1}{4S^2 N} \sum_{a=1}^N \sum_{sx=-S}^S \sum_{sy=-S}^S I(\mathbf{X}_{\phi k a} + \Delta s)_g \quad \dots \quad (4)$$

20

### 【0036】

#### 【数5】

$$L_{\phi k b} = \frac{1}{4S^2 N} \sum_{a=1}^N \sum_{sx=-S}^S \sum_{sy=-S}^S I(\mathbf{X}_{\phi k a} + \Delta s)_b \quad \dots \quad (5)$$

30

### 【0037】

時刻における番目の領域 $L$ の色情報ベクトル $L = \{L_r, L_g, L_b\}$ は3つの要素からなる。色の恒常性が撮影画像列においては保たれると仮定すると、フレーム間の領域の対応は、色情報ベクトルの距離が近いものから対応候補として選択することができる。時刻における番目の色情報ベクトルと時刻'における'番目の色情報ベクトルとの差 $D_L$ は、ベクトルのノルムを表す記号、

$$\|\|$$

40

を用いると式(6)のようになる。

### 【0038】

#### 【数6】

$$DL_{k\phi}(k', \phi') = \|L_{\phi k} - L_{\phi' k'}\| \quad \dots \quad (6)$$

### 【0039】

色情報ベクトルの差 $D_L$ を計算し、値の小さいものから対応候補とすることができます。候補は複数あるほうが、照明による色の変化がある場合に対応する範囲を広げて検査できる。ここで、色情報ベクトルを用いて時刻における番目の領域の番目の特徴点 $P$

50

と、時刻 'におけると '番目の領域の '番目の点 P の対応を求める。単純な画素比較とすると、カメラもしくは撮影対象の動きが画像面に対してほぼ平行に移動する場合、2つの局所画像の濃度差 G( , ' )は式(7)のように表せる。なお、|x|はxの絶対値である。

【0040】

【数7】

$$G(\phi\alpha\kappa, \phi'\alpha'\kappa') = \frac{1}{4S^2} \sum_{sx=-S}^S \sum_{sy=-S}^S |I(\mathbf{X}_{\phi\alpha\kappa} + \Delta s) - I'(\mathbf{X}_{\phi'\alpha'\kappa'} + \Delta s)| \quad \dots \quad (7)$$

【0041】

10

初期の領域対応候補が求まるとき、2枚の画像のうち、濃度差 G を最小とする ' が特徴点の対応となる。式(7)は、非特許文献1、2と同様にカメラの観察対象の空間的広がりがあまり変化しない場合には有効である。しかし、課題でも述べたように、撮影対象が回転している場合には、条件を満たす対応は十分に得られない場合がある。

【0042】

そこで、本発明の実施例として領域対応付け部30では、領域の剛体運動の推定結果を利用して局所画像領域の対応をとる。ここで、時間的な連続性を考慮した過去の画像から推定した剛体運動を利用する。具体的には、時刻 の画像には時刻 1, -2, ..., 1 の剛体運動の推定結果を用いる。それらの剛体運動は、後述の領域剛体運動推定部50で推定する。なお、時刻 = 0 での初期値としては静止しているものとしても構わないし、乱数や定数で剛体運動として想定内の値を与えて良い。

20

【0043】

次に、時刻 における領域の対応付けを考える。なお、時刻 nにおいて領域剛体運動推定部50による剛体運動の推定値は得られているものとしているが、時系列的に画像を処理することにより、前の時刻の推定結果を利用することができます。

【0044】

非特許文献5によると、カメラの運動は画像中の特徴点の画像列における動きから再現することができる。

【0045】

時刻 n のとき、カメラ座標系での3次元位置が既知の N 個の点 p<sub>1</sub>, ..., p<sub>N</sub>, p<sub>n</sub> のそれぞれの位置 x<sub>1</sub>, ..., x<sub>N</sub>, x<sub>n</sub> を用いる。カメラが平行移動 V<sub>n</sub>、回転移動 R<sub>n</sub> という運動をしたとき、カメラ座標系における各点 p<sub>i</sub>, p<sub>n</sub> のカメラ座標系での速度 v<sub>i</sub>, v<sub>n</sub> は、式(8)のようになる。ただし、添字 i は 1 から N とする。

30

【0046】

【数8】

$$\Delta \mathbf{p}_{ik-n} = -\mathbf{V}_{k-n} - \boldsymbol{\Omega}_{k-n} \times \mathbf{x}_{ik-n} \quad \dots \quad (8)$$

【0047】

カメラ移動前の点 p<sub>i</sub> の位置 x<sub>i</sub>, 1 と移動後の位置 x<sub>i</sub>, n の間の関係は、 x<sub>i</sub>, n = x<sub>i</sub>, 1 + p<sub>i</sub>, n から、式(9)のように表せる。

40

【0048】

【数9】

$$\mathbf{x}_{ik} = F_{k-n} \mathbf{x}_{ik-n} - \mathbf{V}_{k-n} \quad \dots \quad (9)$$

【0049】

ただし、F<sub>k-n</sub> は、式(10)のような行列となる。

【0050】

## 【数10】

$$F_{\kappa-n} = \begin{bmatrix} 1 & \Omega_{z\kappa-n} & -\Omega_{y\kappa-n} \\ -\Omega_{z\kappa-n} & 1 & \Omega_{x\kappa-n} \\ \Omega_{y\kappa-n} & -\Omega_{x\kappa-n} & 1 \end{bmatrix} \cdots (10)$$

## 【0051】

また、点  $p_i$  の画像への投影点  $P_i = X_i$  は、焦点距離を  $f$  として、中心射影の式から式(11)のように表せる。

## 【0052】

## 【数11】

$$\mathbf{X}_{ik} = \begin{pmatrix} f \frac{x_{ik}}{z_{ik}} \\ f \frac{y_{ik}}{z_{ik}} \end{pmatrix} \cdots (11)$$

20

## 【0053】

時刻  $n$  における運動パラメータとしての平行移動成分  $V_n$  と回転移動成分  $\Omega_n$  とカメラ座標系での点  $p_i$  から時刻  $n$  における投影点  $P_{ik}$  の位置  $X_{ik}$  を計算する。具体的には、式(11)に式(9)を代入して式(12)が得られる。

## 【0054】

## 【数12】

$$\mathbf{X}_{ik} = \begin{pmatrix} f \frac{x_{ik-n} + \Omega_{z\kappa-n} v_{ik-n} - \Omega_{y\kappa-n} z_{ik-n} - V_{x\kappa-n}}{\Omega_{y\kappa-n} x_{ik-n} - \Omega_{x\kappa-n} v_{ik-n} + z_{ik-n} - V_{z\kappa-n}} \\ f \frac{-\Omega_{z\kappa-n} x_{ik-n} + v_{ik-n} + \Omega_{x\kappa-n} z_{ik-n} - V_{x\kappa-n}}{\Omega_{y\kappa-n} x_{ik-n} - \Omega_{x\kappa-n} v_{ik-n} + z_{ik-n} - V_{z\kappa-n}} \end{pmatrix} \cdots (12)$$

30

## 【0055】

式(12)は、既知の3次元位置が運動パラメータを用いて、撮像画像の位置を推定する式である。実際には、特徴点の3次元位置も運動パラメータも初期状態においては全て未知のため、式(12)をそのまま適用することはできない。しかし、後述の形状推定および運動推定を行った結果であるそれぞれの推定値を利用することが可能である。

## 【0056】

時刻  $n$  における 番目の領域の 番目の特徴  $p$  の画像座標である特徴点  $P$  は、そのカメラ座標系での推定3次元位置  $X = \{x, y, z\}$  と運動パラメータ  $V = \{V_x, V_y, V_z\}$  と  $Q = \{Q_x, Q_y, Q_z\}$  を用いて、次式(13)のように推定位置  $J$  を求めることができる。

## 【0057】

40

## 【数13】

$$J(\phi, \alpha, \kappa) = \begin{cases} f \frac{X_{\phi\alpha\kappa}/z_{\phi\alpha\kappa} + \Omega_{z\phi\kappa} Y_{\phi\alpha\kappa}/z_{\phi\alpha\kappa} - \Omega_{y\phi\kappa} z_{\phi\alpha\kappa} - V_{x\phi\kappa}}{\Omega_{y\phi\kappa}(X_{\phi\alpha\kappa}/z_{\phi\alpha\kappa}) - \Omega_{x\phi\kappa}(Y_{\phi\alpha\kappa}/z_{\phi\alpha\kappa}) + z_{\phi\alpha\kappa} - V_{z\phi\kappa}} \\ f \frac{-\Omega_{z\phi\kappa}(X_{\phi\alpha\kappa}/z_{\phi\alpha\kappa}) + (Y_{\phi\alpha\kappa}/z_{\phi\alpha\kappa}) + \Omega_{x\phi\kappa} z_{\phi\alpha\kappa} - V_{x\phi\kappa}}{\Omega_{\phi\kappa}(X_{\phi\alpha\kappa}/z_{\phi\alpha\kappa}) - \Omega_{x\phi\kappa}(Y_{\phi\alpha\kappa}/z_{\phi\alpha\kappa}) + z_{\phi\alpha\kappa} - V_{z\phi\kappa}} \end{cases} \quad \dots (13)$$

## 【0058】

式(13)の推定位置Jを式(7)のXに代入して、運動と3次元位置を考慮した濃度差G'は、次式(14)のように表せる。

10

## 【0059】

## 【数14】

$$G'(\phi\alpha\kappa, \phi'\alpha'\kappa') = \frac{1}{4S^2} \sum_{sx=-S}^S \sum_{sy=-S}^S |I(J(\phi, \alpha, \kappa - n) + \Delta s) - I'(\mathbf{X}_{\phi'\alpha'} + \Delta s)| \quad \dots (14)$$

## 【0060】

式(14)の濃度差G'は、時刻nの運動パラメータから推定される画面座標を用いて対応する濃度値の計算を行っており、静止している運動パラメータの値が0となるので式(7)と同じになる。なお、式(14)は局所画像Iにおける隣接画素の奥行き値は注目画素と同じとして計算しても良いし、もし後述の領域形状推定で位置が求まっているればそれを利用しても構わない。

20

## (領域形状推定部)

領域形状推定部40では、領域の画像内の特徴点の3次元座標を求める。カメラ座標系で形状推定は特徴点の奥行き値zを推定することである。

## 【0061】

時刻iでの対応づけられた第i番目の特徴点のオプティカルフローは式(15)のように示される。

## 【0062】

## 【数15】

$$\dot{\mathbf{X}}_{ik} = (X_{ik} - X_{ik-1}, Y_{ik} - Y_{ik-1}) \quad \dots (15)$$

30

## 【0063】

さらに、時刻1において推定された撮像装置の運動パラメータV<sub>1,1</sub>を利用すると、式(9)および式(11)から、式(16)で表される、時刻1での奥行きz<sub>i,1</sub>が得られる。なお、ここでfはカメラの焦点距離である。

## 【0064】

## 【数16】

$$z_{ik-1} = \frac{f X_{ik} V_{xk-1} - f^2 V_{xk-1}}{X_{ik} (f - \Omega_{xk-1} Y_{ik-1} + \Omega_{yk-1} X_{ik-1}) - f X_{ik-1} - \Omega_{z_{k-1}} f Y_{ik-1} + \Omega_{y_{k-1}} f^2} \quad \dots (16)$$

40

## 【0065】

また、初期段階では運動パラメータは推定できていないため、後述の領域剛体運動推定部50での処理を先に実行しても良い。

## (領域剛体運動推定部)

中心射影を仮定すると、カメラ座標系での3次元位置p<sub>i</sub>と撮像画像の座標には式(11)の関係がある。式(11)に式(9)を代入して、x<sub>i</sub>、y<sub>i</sub>、z<sub>i</sub>を消去して、カメラの運動パラメータに関して整理する。

## 【0066】

50

時刻  $n$  における点  $p_i$  のカメラ座標系での 3 次元座標、カメラの平行運動成分  $V_n$ 、回転運動成分  $\Omega_n$ 、および奥行きの推定値  $z_{ik-n}$ 、そして画像画面上の特徴点  $P_i$  の座標  $X_{i-n}$  と  $P_i$  の座標  $X_{i-n}$  との関係を示す式として、式(17)および式(18)を得る。

【0067】

【数17】

$$-X_{ik}y_{ik-n}\Omega_{xk-n} + (X_{ik}x_{ik-n} + fz_{ik-n})\Omega_{yk-n} - fy_{ik-n}\Omega_{zk-n} + fV_{xk-n} - X_{ik}V_{zk-n} = fx_{ik-n} - X_{ik}z_{ik-n} \quad \dots (17)$$

【0068】

【数18】

$$(-Y_{ik}y_{ik-n} - fz_{ik-n})\Omega_{xk-n} + Y_{ik}x_{ik-n}\Omega_{yk-n} + fx_{ik}\Omega_{zk-n} + fV_{yk-n} - Y_{ik}V_{zk-n} = fy_{ik-n} - Y_{ik}z_{ik-n} \quad \dots (18)$$

【0069】

既知の  $M$  個の点に関して式(17)および式(18)の 2 つの式が得られるので、全体として  $2M$  の方程式となる。この  $2M$  の方程式は、未知数が運動パラメータ  $V$ 、 $\Omega$  のそれぞれ 3 個の要素からなるので、合計 6 個に対する連立方程式となる。よって、最低 3 点の対応するオプティカルフローがあれば、 $V$ 、 $\Omega$  を計算することができる。3 点より多い場合は最小 2 乗法を用いて解くことができる。

【0070】

点の位置が不定の場合は、点の位置も未知変数として推定することができる。そこで、特徴点の 3 次元位置とスケールの運動成分をのぞいた 5 つの運動パラメータを未知として計算する。つまり、 $3M + 5$  の未知数を解くための方程式は、1 つの点につき式(17)、式(18)が得られるので、5 点の対応から構成される連立方程式を解くことで求めることができる。

【0071】

領域形状推定部 40 では、領域抽出部 20 により求められた対応する領域における特徴点のうち、5 点をランダムにサンプルして式(17)、式(18)からなる連立方程式を計算する。そして、領域の特徴点の 3 次元位置および領域の運動パラメータを推定することとする。

【0072】

ただし、1 回のサンプリングでは領域の誤対応による推定の誤差が大きくなる場合も想定されるため、複数回のサンプリングを行い、その中で誤差の小さいものを採用すると誤対応による影響を小さくすることができる。

(領域統合分離部)

実際の画像では、領域検出部 20 において小さい領域が多数検出される場合がある。それぞれ領域ごとに領域形状推定部 40 および領域剛体運動推定部 50 の処理を行うことはできるが画面に占める観測面積が小さいと誤差の影響を受けやすい。そこで、複数の領域を一つの剛体運動で近似できる場合は、統合処理により推定精度を向上する。また、領域に別の剛体運動物体が含まれる場合は、推定精度が低下するので、その領域を検出し、分離することとする。

【0073】

ここで、領域 A と領域 B が時刻 1 において同じ剛体運動をしていると仮定すると、領域 A の平行移動成分  $V_A$  と領域 B の平行移動成分  $V_B$  は同じ値になり、同様に回転運動成分  $\Omega_A$  と  $\Omega_B$  も同じになる。しかし、特徴点検出時の誤差や計算誤差の影響により全く同じになることは少ない。そこで、ほぼ同じ運動をする領域のそれぞれの剛体運動を 1 つの剛体運動として統合すると、観察する領域の観測対象が増加

するため、相対的に誤差が少なくなる可能性がある。また、逆に領域の一部が異なる運動をすると、異なる運動を始める前に比べて誤差が相対的に大きくなる。つまり、誤差の相対的な変化を観察することで、領域を統合するか分離するかを決めることができる。

## 【0074】

領域統合分離部 60 では、複数の領域の剛体運動を統合する処理と分離する処理を行う。まず、同じ剛体運動をしている領域の統合処理について述べる。

## 【0075】

ここで、同時刻における画像からの検出領域のうち、A 番目と B 番目の領域との運動パラメータの差を D ( A , B ) とする。D ( A , B ) は、平行移動成分と回転運動成分のそれぞれのベクトルのノルムの 2 乗を用いて式 ( 19 ) のように表現できる。

10

## 【0076】

## 【数 19】

$$D(A, B) = \| \mathbf{V}_A - \mathbf{V}_B \|^2 + \| \boldsymbol{\Omega}_A - \boldsymbol{\Omega}_B \|^2 \quad \dots (19)$$

## 【0077】

いま、A 番目の領域に対して、画面から検出される A 番目以外の領域に対して運動パラメータの差 D を計算し、その計算値の小さい順から整列することとする。これにより、運動パラメータ A と近い運動をする候補が選択することができる。なお、A 番目の領域選択の順序としては、画面上の領域の面積が大きいほうから順次求めると誤差の影響を受けにくく推定精度が向上する。

20

## 【0078】

式 ( 12 ) を用いると時刻 n の運動パラメータとカメラ座標系での3次元位置から時刻 n の画面座標を推定できる。そこで、時刻 n の領域 A の特徴点 p\_i\_n と、その時の運動パラメータ V\_n と n を用いて、時刻 n に推定される画面座標を X'\_i ( A ) とする。時刻 n において撮像画像から検出された領域 A の特徴点 P\_i の画面座標 X\_i ( A ) との差を投影面誤差 E\_i と定義すると次式 ( 20 ) のように表せる。

## 【0079】

## 【数 20】

$$E_{Ai\kappa} = | \mathbf{X}_{i\kappa}(A) - \mathbf{X}'_{i\kappa}(A) | \quad \dots (20)$$

30

## 【0080】

領域 A における特徴点から n 個無作為にサンプリングした集合を C\_{nA} と表す。そこで、集合 C\_{nA} の投影画面誤差の和を E\_A ( n ) として式 ( 21 ) のように計算する。

## 【0081】

## 【数 21】

$$\sum E_{Ai\kappa}(n) = \sum_{i \in C_{nA}} E_{Ai\kappa} = \sum_{i \in C_{nA}} | \mathbf{X}_{i\kappa}(A) - \mathbf{X}'_{i\kappa}(A) | \quad \dots (21)$$

40

## 【0082】

同様に、領域 B における投影画面誤差の和を E\_B ( n ) とすると、式 ( 22 ) のように表せる。

## 【0083】

## 【数 22】

$$\sum E_{Bi\kappa}(n) = \sum_{i \in C_{nB}} | \mathbf{X}_{i\kappa}(B) - \mathbf{X}'_{i\kappa}(B) | \quad \dots (22)$$

50

## 【0084】

次に、領域 A と領域 B が統合した状態で、投影画面誤差の和を求める。そのためには、時刻  $n$ において、領域 A と領域 B が同じ剛体運動をしている一つの領域として考える。つまり、式(17)および式(18)における運動パラメータ推定において、領域 A と領域 B を結合した領域 A $\cap$ B から無作為に  $n$  個選択した集合  $C_{nA\cap B}$  を用いてパラメータ推定を行う。投影画面誤差の和を  $E_{A\cap B}(n)$  とすると式(23)のように表せる。

## 【0085】

## 【数23】

$$\sum E_{A\cap B_k}(n) = \sum_{i \in C_{nA\cap B}} |\mathbf{X}_{ik}(A\cap B) - \mathbf{X}'_{ik}(A\cap B)| \quad \dots \quad (23)$$
10

## 【0086】

統合処理においては、領域 A と領域 B の統合した領域から剛体運動の推定結果が、それぞれ領域ごとに求めた推定結果の平均より良くなる場合に領域 A と領域 B を統合し、領域変更を行う。推定結果が良いということは、投影画面誤差が少なくなることなので、次式(24)の関係式を満たすときに統合することとする。

## 【0087】

## 【数24】

$$\sum E_{A\cap B_k}(n) < \frac{1}{2} (\sum E_{Ak}(n) + \sum E_{Bk}(n)) \quad \dots \quad (24)$$
20

## 【0088】

ここで、 $n$  の値としては 5 点以上あれば良いが、多すぎると選択された特徴点のうちに誤対応の点が含まれる可能性が高くなり、一つの誤差の影響で式(24)が常に満たされない場合が発生しやすい。そこで、 $n$  は 5 点 ~ 10 点程度にして、式(22)から式(23)までの処理を複数回実行し、それぞれの結果を整列したときの中間の値を採用することで、誤対応による影響を削減することができる。

## 【0089】

次に、領域に別の運動が含まれる際の分離について述べる。時刻 1 まで同一の剛体運動をしていた領域 A が、時刻  $n$  において別の剛体運動をする領域 A の局所領域 B が含まれるとする。これまで同一の剛体運動として推定していた運動パラメータによる投影画面誤差の和の値が局所領域 B の影響により大きくなる。分離の判定としては、投影画面誤差の和の誤差の時間変化を検査すれば良く、式(25)のように表せる。

## 【0090】

## 【数25】

$$\left| \sum E_{Ak}(n) - \sum E_{Ak-1}(n) \right|^2 > econst \quad \dots \quad (25)$$

## 【0091】

式(25)の  $econst$  は、実験的に求めた定数であり、所定のレベルでもある。式(25)の条件を満たし、複数の領域のうち 1 つの領域に関して推定された剛体運動の精度が所定のレベルを下回るとき、時刻  $n$  には別の剛体運動による領域が存在するとして分離処理を行う。

## 【0092】

分離処理では、領域に含まれる他の剛体運動の部分を抽出することを行う。領域 A に含まれる特徴点の集合はインライアとして、集合 A に登録する。具体的には、無作為に抽出した集合  $C_{nA}$  が式(25)を満たさない場合には、インライアとして領域 A の一部として  $C'_{nA}$  として登録する。

## 【0093】

10

20

30

40

50

一方、選択した特徴点の集合の中に他の剛体運動が領域に含まれる場合、その他の剛体運動をしている特徴点はアウトライアとなり、画面投影誤差を大きくする要因となる。そこで、その特徴点を抽出して、領域Aとは別の集合に登録する。具体的には、無作為に抽出した集合 $C_{n_A}$ が式(25)を満たす場合は集合の中にアウトライアの特徴点が含まれているものとする。

#### 【0094】

具体的なアウトライアを選別するために、該 $C_{n_A}$ から1つ特徴点の情報を抽出し、既に領域Aとして登録されている集合 $C'_{n_A}$ から $n-1$ 個の特徴点を選択して、式(25)を満たすかどうか検査する。該 $C_{n_A}$ から抽出した特徴点の一つがアウトライアならば、 $n-1$ 個の特徴点が $C'_{n_A}$ の値がAに属している場合でも式(25)を満たす可能性が高い。これを該 $C_{n_A}$ の全ての特徴点に対して順次繰り返し検査する。アウトライアとして検出されたものは、アウトライアの集合 $C'_{n_B}$ に登録する。10

#### 【0095】

領域Aに含まれる特徴点全てに対して処理が終了したら、アウトライアの集合 $C'_{n_B}$ を用いて、式(24)のBを $C'_{n_B}$ に置き換えて検査を実行する。もし式(24)を満たさなければ、アウトライアの集合 $C'_{n_B}$ は別の剛体運動である可能性が高いので、以降の画像列においては別の剛体運動として登録する。

#### 【0096】

<実施形態2>

実施形態1において、領域統合分離部60の領域統合および分離は、探索的に行われているので再度領域の対応付けからの処理を繰り返すことで、推定精度の向上が期待できる。そこで、本発明の実施形態2においては、それぞれの推定結果を再度繰り返して処理して、推定精度を向上するための制御を行う統合分離制御部を有する画像処理装置の実施例について述べる。20

#### 【0097】

図3は、統合分離制御部70を有する画像処理装置のうちの一つである画像処理装置の主要部の構成例を示したものである。ここで、画像取得部10、領域抽出部20、領域対応付け部30、領域形状推定部40、領域剛体運動推定部50および領域統合分離部60統合分離制御部70の機能は、基本的に図1で述べた機能と同じであるので説明は省略する。30

#### 【0098】

統合分離制御部70は、領域対応付け部30、領域形状推定部40、領域剛体運動推定部50、領域統合分離部60のそれぞれの手段を実現するための処理部に対して、領域の推定結果を渡して、その処理結果を用いて制御を行うものである。

#### 【0099】

図4は、統合分離制御部70の内部処理手続きの主要な部分を記載したフロー図である。具体的な処理手続きに関して、このフロー図を用いて説明する。なお、このフローに記載されている内容は、統合分離制御部の内部処理手続きの主要な部分を記載したものであり、実際にはそれぞれの処理結果のデータ保持等に関わる手続きが必要となる。

#### 【0100】

まず、統合分離制御部70は、画像取得部10および領域抽出部20による領域検出が行われた状態で、以下のステップ動作を開始する。

#### 【0101】

ステップS10では、統合分離制御部の開始が領域抽出部20の出力が得られた時に開始する。ステップS11では、統合分離制御部の繰り返し回数*i*を0に初期化する。

#### 【0102】

ステップS20では、領域の統合・分離による対応付けの変更を行う。初期値としては、過去の画像列による推定結果を保持している場合は、それを利用することができる。ここでは、前の画像の処理結果もしくは領域統合分離部60による領域の統合・分離について、領域対応付け部30の処理を再度実行する。領域の統合・分離により、画像面投影誤4050

差が小さくなるように推定された運動パラメータを利用することができるので、対応付けをより良くすることができます。

#### 【0103】

ステップS30では、前記ステップS20により対応付けが変更された結果を用いて、領域形状推定部40の処理を再度実行する。推定結果としては、カメラ座標系での各特徴点の奥行き値が得られる。

#### 【0104】

ステップS40では、前記ステップS30により推定された推定形状と、前回繰り返しにより推定した過去の領域形状推定部40の推定形状との差を計算する。具体的には、カメラ座標系における対応する各特徴点の奥行き値の差の2乗和を計算する。これは、運動パラメータの推定が十分に精度が良くなっている場合は、形状推定の値があまり変化しなくなることから、次のステップで処理手続きの終了を判断するかどうかの値を計算することとなる。

10

#### 【0105】

ステップS50では、前記ステップS40で計算した形状推定誤差の値が設定した閾値より小さいかどうかの判別を行う。閾値の設定は実験的に設定することができる。処理による誤差の変化が小さいときには、これ以上推定を繰り返すことは無駄なことなので、ステップS100の繰り返し停止の制御に移る。設定した閾値より誤差の変化が大きい場合は、まだ精度を向上する余地があるとしてステップS60の処理に移行する。

#### 【0106】

20

前記ステップS60の処理では、領域の剛体運動推定を行う。ここでは、ステップS40の処理で得られた形状推定結果を領域剛体運動推定部50で再度処理することを行う。領域の剛体運動推定には特徴点の奥行き値の推定値が必要なので、この値が精度良く求まつていると剛体運動の推定精度も向上する。

#### 【0107】

ステップS70では、前記ステップS60の剛体運動推定結果を利用して、領域の統合・分離の制御を行う。具体的には、領域統合分離部60に対してステップS60の推定結果を使って再度処理することを行う。剛体運動の推定精度が向上していると、投影画面誤差の計算の精度も向上する。それにより、領域変更である統合・分離の処理にも影響を与える。

30

#### 【0108】

ステップS80では、前記ステップS70において処理された領域の変更後の統合や分離に関する処理による変化を検査して、処理の制御を行う。前回の繰り返し時、もしくは前の画像列における領域統合分離部60の結果を利用し、統合と分離のそれぞれの処理個数の変化を調べる。そして処理個数の差を計算し、その差が設定した閾値より小さい場合は、ステップS100の繰り返し停止の制御を実行する。統合・分離の処理個数の変化が閾値より大きい場合は、インライア・アウトライアの分離等が十分に行われていない可能性が考えられる。そこで、ステップS90の制御に移る。

#### 【0109】

ステップS90では、繰り返し回数を保持している変数iの値を1つ増加させる。ステップS95では、繰り返し回数が閾値以上であるかどうかを判断する。繰り返しにより統合・分離の変化が継続している状態には、そもそもその推定において誤対応が多く含まれて推定の各ステップによる計算が破綻している場合が考えられる。その場合は、継続して繰り返し制御を実行することは困難なので、繰り返し回数が設定した閾値を越えた場合は、処理を停止するステップS100に制御が移行する。それ以外は、ステップS20の処理手続きを継続して行うこととする。なお、ステップS95における閾値も実験的に設定することができる。

40

#### 【0110】

上述の各ステップは、基本的には本発明の実施形態1の手段を実現する処理部を繰り返して実行する制御を組み合わせたものであり、本発明の効果をより良くするための一例で

50

ある。ステップS50およびステップS80における処理の判断に用いられる閾値は、実験的に定めた値を用いても構わないし、事前にシーンに対する知識が得られるならば、それに適した事前に用意した値を利用しても良い。また、ステップS50では得られた形状推定の計測範囲をパラメータとした閾値を設定しても構わないし、ステップS80では、統合・分離の個数をパラメータとした閾値を設定しても良い。

#### 【0111】

<他の実施形態>

本発明は、実施形態1で説明した例のように画像処理装置として他の撮像装置と組み合わせて利用する他に、コンピュータのプログラムとしても実施できる。

#### 【0112】

本発明の一実施例であるところの構成をコンピュータのプログラムとして実施し、移動する撮影対象領域を検出し、撮影対象の領域のみを伝送することに利用できる。ネットワークを利用する場合は、画像のような情報量が多い場合は、撮像領域のみを限定することで、情報量を削減することができる。

#### 【0113】

従来の撮影対象の領域の検出には、ブルーバックと呼ばれる背景を均一の色で塗りつぶした環境において、前景である撮影対象を撮影する方法が放送分野においては一般的に利用されている。しかし、これらは放送分野では一般的であるが、それ以外の一般的なビデオカメラの利用者にとっては撮影準備が煩雑であり、そのような準備をして撮影することは少ない。

#### 【0114】

本発明の実施例では、移動する撮影対象を移動しながら撮影した場合にでも、撮影対象の領域を動的に得ることができる。図5を用いて本発明の好適な利用方法の一例について説明する。

#### 【0115】

パーソナルコンピュータ530は、CPU、記録素子、外部記録素子、それらを結ぶバスなどのハードウェアとOSが動作する機能を具備し、入力装置としてキーボード・マウスを、そして画像出力として液晶ディスプレイを有しているものとする。

本発明の画像処理方法をプログラムのアプリケーションとして実装し、該OSにおいて利用できるようにしたとする。そのアプリケーションは、パーソナルコンピュータ530の記録領域にロードされて実行できるものとする。該アプリケーションは、本発明の画像処理方法に係る処理パラメータの変更、動作の指示、処理結果の確認が画面に表示できるようにしたものである。該アプリケーションのGUI530は、パーソナルコンピュータ530のキーボード・マウスの入力装置を使って利用者が操作できるものとする。

#### 【0116】

撮像装置200は、パーソナルコンピュータ530の外部入力インターフェースとケーブル501で接続されている。一般的にはUSBカメラや、IEEE1394カメラを利用することができます。パーソナルコンピュータ530に該撮像装置を利用するデバイスドライバが設定されており、撮像画像を取得できる状態にあるものとする。

#### 【0117】

撮影者510は、撮像装置200を手にもち、撮影対象を撮影するものとする。撮像装置200は、矢印520のように移動させながら移動する撮影対象を撮影するものとする。

#### 【0118】

移動する撮影対象600として、ここでは車とした。撮影対象600は、矢印510に示す移動をしながら、撮影者510の前を通過する動作を行う。さらに、撮影するシーンには、静止している撮影対象600も撮影画像中に撮影されているものとする。

#### 【0119】

そこで、撮影を開始するために、パーソナルコンピュータ530にて本発明の実施例である画像処理方法を実装したアプリケーションを実行し、GUI230からマウス若しく

10

20

30

40

50

はキー ボードを用いて撮影開始を指示することとする。

#### 【0120】

移動する撮影対象600を撮影した後、GUI530から撮影終了をマウス若しくはキー ボードを用いて指示する。GUI530は、本発明の実施例1で述べた処理を行い、領域情報出力部120およびそれに付随する領域形状推定部40の結果をグラフィックスライブラリを用いて3次元グラフィックスとして表示する。なお、グラフィックスライブラリはOpenGLなどの汎用3次元描画ライブラリを利用できるが、パーソナルコンピュータ530に、その機能を有していない場合でもCPUを用いて画像を生成することは可能である。

#### 【0121】

利用者は、GUI530に表示された撮影対象600の領域情報を確認したのち、ネットワークのサーバに対して撮影対象600に関する情報をアップロードすることができる。ここでは、パーソナルコンピュータ530に装備している無線LANモジュール550を利用し、無線LANルーター560を介してインターネット570の通信経路上に存在するサーバ575にデータを送信できる。

#### 【0122】

データの送信形式は、無線LANやインターネットで既定されているプロトコルをそのまま利用することができる。HTTPを使うと、プロキシーを利用している場合でも簡単にデータを送信することができる。

#### 【0123】

さらに、撮影対象600に対する利用者のコメントを追加することもできる。一例として、撮影対象の属性や、撮影された場所に関する情報、利用した機材の情報を追加することで、後ほど該撮影対象を確認するときの利便性を向上することができる。

#### 【0124】

サーバ575は、撮影対象600に関する領域と付加情報を受信すると、それをサーバ575のWebサーバに閲覧できるように登録を行う。例えば、利用者のコメントと、第1フレームの撮影対象600の画像をスナップ画像としてHTML形式で記述したファイルをサーバの閲覧可能フォルダに置くことで実現できる。

#### 【0125】

インターネット570に接続しているパーソナルコンピュータ580を利用している閲覧者は、サーバ575が提供する情報をWEBブラウザを利用して閲覧することができるものとする。

#### 【産業上の利用可能性】

#### 【0126】

実施例の一例として上述の構成例を説明したが、一般的にはカメラと本発明の処理手段を実施する装置があれば良い。本発明の処理構成を携帯電話や携帯型のコンピュータで実施するプログラムという形で処理を実行させても同じ効果が得られる。

#### 【0127】

なお、前述した実施形態の機能を実現するソフトウェアのプログラムを、記録媒体から直接、あるいは有線／無線通信を用いて当該プログラムを実行可能なコンピュータを有するシステム又は装置に供給する。そして、そのシステムあるいは装置のコンピュータが該供給されたプログラムを実行することによって同等の機能が達成される場合も本発明に含む。

#### 【0128】

したがって、本発明の機能処理をコンピュータで実現するために、該コンピュータに供給、インストールされるプログラムコード自体も本発明を実現するものである。つまり、本発明の機能処理を実現するためのコンピュータプログラム自体も本発明に含まれる。

#### 【0129】

その場合、プログラムの機能を有していれば、オブジェクトコード、インタプリタにより実現されるプログラム、OSに供給するスクリプトデータ等、プログラムの形態を問わ

10

20

30

40

50

ない。

【0130】

プログラムを供給するコンピュータ読み取り可能な記憶媒体としては、例えば、フレキシブルディスク、ハードディスク、磁気テープ等の磁気記録媒体、MO、CD-ROM、CD-R、CD-RW、DVD-ROM、DVD-R、DVD-RW等の光／光磁気記録媒体、不揮発性の半導体メモリなどがある。

【0131】

有線／無線通信を用いたプログラムの供給方法として、例えばクライアントコンピュータ上で本発明を形成するコンピュータプログラムとなり得るプログラムデータファイルを記録し、接続が適合するコンピュータにそのファイルをダウンロードする方法がある。この場合、プログラムデータファイルを複数のセグメントファイルに分割し、セグメントファイルを異なるサーバに配置することも可能であり、ファイルを圧縮して置くことも可能である。10

【0132】

つまり、本発明の機能処理をコンピュータで実現するためにプログラムデータファイルを複数のユーザに対してダウンロードさせるサーバ装置も本発明に含む。

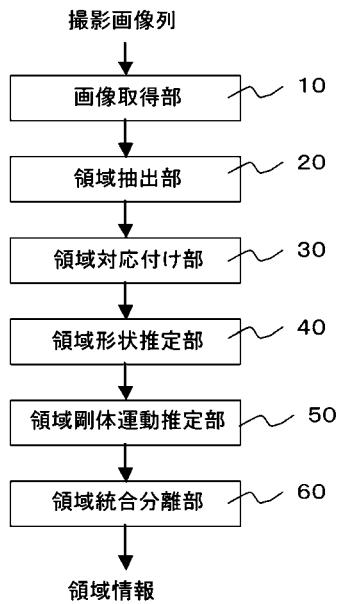
【0133】

また、本発明のプログラムを暗号化してCD-ROM等の記録媒体に格納してユーザに配布し、所定の条件を満たしたユーザに対して暗号化を解く鍵情報を、例えばインターネットを介してホームページからダウンロードされることによって供給する。そして、その鍵情報を使用することにより暗号化されたプログラムを実行してコンピュータにインストールさせて実現することも可能である。更に、コンピュータが読み出したプログラムを実現することによって、前述した実施形態の機能が実現される。また、そのプログラムの指示に基づき、コンピュータ上で稼働しているOSなどが、実際の処理の一部又は全部を行い、その処理によって前述した実施形態の機能が実現され得る。20

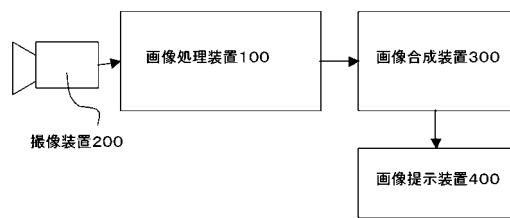
【0134】

更に、記録媒体から読み出されたプログラムを、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込む。その後、そのプログラムの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部又は全部を行い、その処理によっても前述した実施形態の機能が実現され得る。また、監視やセキュリティの用途としても撮像画像から多くの情報量を取得できる本装置の有用性を利用することができる。例えば、撮影対象として人を対象として背景領域から移動物体として検出することにも利用できる。30

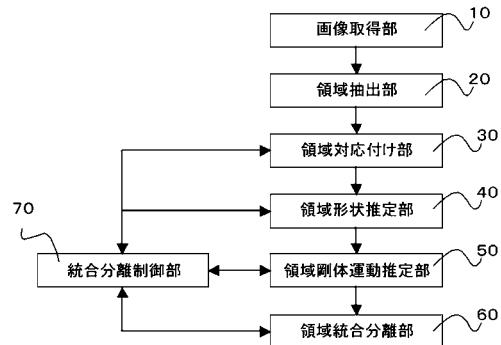
【図1】



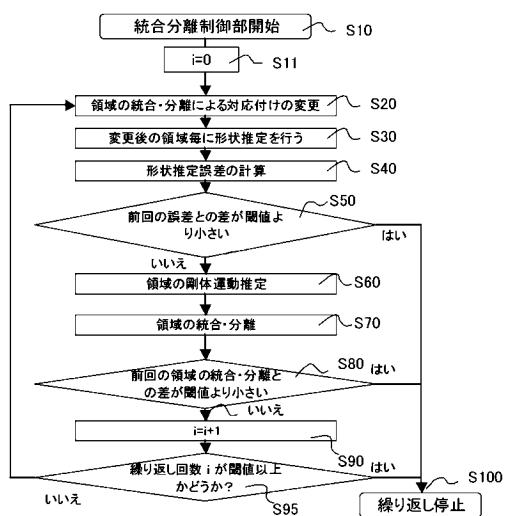
【図2】



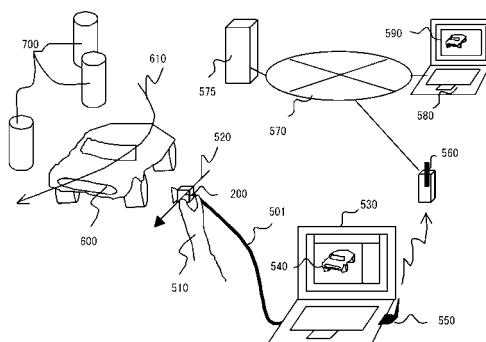
【図3】



【図4】



【図5】



---

フロントページの続き

(72)発明者 小林 一彦

東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 板垣 有紀

(56)参考文献 特開平09-212652(JP,A)

特開平08-287263(JP,A)

特表2002-525735(JP,A)

前 泰志, オプティカルフローとエッジを用いた三次元運動物体の複数平面近似による抽出と追跡, 日本ロボット学会誌, 日本, 社団法人日本ロボット学会, 1999年10月15日, 第17巻第7号, p.45-51

上田 剛弘, 平面の3次元配置に基づく動画像の構造利用符号化, 電子情報通信学会論文誌, 日本, 社団法人電子情報通信学会, 1997年 2月25日, 第J80-D-II巻 第2号, p.487-496

(58)調査した分野(Int.Cl., DB名)

G 06 T 1 / 0 0 - 7 / 6 0

H 04 N 5 / 2 3 2