

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-199285

(P2017-199285A)

(43) 公開日 平成29年11月2日(2017.11.2)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 1/00 (2006.01)	G06T 1/00 315	5B057
G10L 25/51 (2013.01)	G10L 25/51	
G10L 21/0272 (2013.01)	G10L 21/0272 100Z	

審査請求 未請求 請求項の数 14 O L (全 21 頁)

(21) 出願番号 特願2016-91608 (P2016-91608)
 (22) 出願日 平成28年4月28日 (2016. 4. 28)

(71) 出願人 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74) 代理人 100126240
 弁理士 阿部 琢磨
 (74) 代理人 100124442
 弁理士 黒岩 創吾
 (72) 発明者 降旗 久義
 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
 (72) 発明者 鈴木 雅博
 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

最終頁に続く

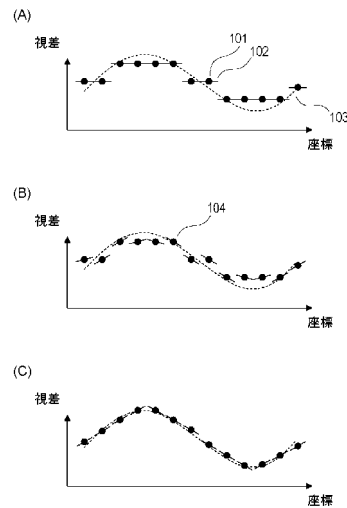
(54) 【発明の名称】 情報処理装置、情報処理方法、プログラム

(57) 【要約】

【課題】 入力データ間の対応を表わす対応データを高速に精度よく推定する。

【解決手段】 第1のデータと第2のデータとを取得する取得手段と、前記第1のデータの部分データと前記第2のデータの部分データとの対応を表す対応データに基づいて、前記対応データの勾配を表す勾配データを導出する導出手段と、前記勾配データに基づいて、前記対応データを伝播させる伝播手段とを備える。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

第 1 のデータと第 2 のデータとを取得する取得手段と、
前記第 1 のデータの部分データと前記第 2 のデータの部分データとの対応を表す対応データに基づいて、前記対応データの勾配を表す勾配データを導出する導出手段と、
前記勾配データに基づいて、前記対応データを伝播させる伝播手段とを備えることを特徴とする情報処理装置。

【請求項 2】

前記対応データが所定の条件を満たすまで前記導出手段による処理と前記伝播手段による処理を繰り返すように、前記導出手段と、前記伝播手段とを制御する制御手段とを更に備えることを特徴とする請求項 1 に記載の情報処理装置。

10

【請求項 3】

前記導出手段は、前記伝播手段による対応データの伝播により更新された対応データに基づいて新たに勾配データを導出し、
前記伝播手段は、前記新たに導出された勾配データに基づいて対応データを伝播させることを特徴とする請求項 2 に記載の情報処理装置。

【請求項 4】

前記第 1 のデータと前記第 2 のデータは、異なる撮像位置から物体を撮像して得られる第 1 の画像と第 2 の画像であり、
前記対応データは、前記第 1 の画像の部分領域と前記第 2 の画像の部分領域の間の視差を表すデータであることを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載の情報処理装置。

20

【請求項 5】

前記対応データは、前記第 1 の画像および前記第 2 の画像と同じ配列形式を有するマップとして保持されることを特徴とする請求項 4 に記載の情報処理装置。

【請求項 6】

前記視差を表すデータは、前記第 1 の画像の部分領域の座標位置または前記第 2 の画像の部分領域の座標位置のいずれかを基準としてデータであることを特徴とする請求項 4 または 5 に記載の情報処理装置。

【請求項 7】

更に、前記対応データを評価する評価手段を備え、
前記伝播手段は、前記勾配データと前記評価手段の結果とに基づいて、前記対応データを伝播させることを特徴とする請求項 5 または 6 に記載の情報処理装置。

30

【請求項 8】

前記評価手段は、前記対応データによって決定される前記第 1 のデータの部分データと前記第 2 のデータの部分データとが整合しているほど高くなるような値を評価値として求めることを特徴とする請求項 7 に記載の情報処理装置。

【請求項 9】

前記伝播手段は、前記マップにおいて、所定の対応データの評価値が近傍の対応データの評価値よりも高い場合に、前記所定の対応データで前記近傍の対応データを置き換えるように対応データを伝播させることを特徴とする請求項 8 に記載の情報処理装置。

40

【請求項 10】

前記評価手段は、更に、前記勾配データに基づいて前記対応データを評価することを特徴とする請求項 8 または 9 に記載の情報処理装置。

【請求項 11】

入力データから第 1 のデータと第 2 のデータとを取得する取得手段と、
前記第 1 のデータと前記第 2 のデータとの対応を表す対応データに基づいて、前記対応データの勾配を表す勾配データを導出する導出手段と、
前記勾配データに基づいて、前記対応データを伝播させる伝播手段とを備えることを特徴とする情報処理装置。

50

【請求項 1 2】

第 1 のデータと第 2 のデータとを取得する取得工程と、
前記第 1 のデータの部分データと前記第 2 のデータの部分データとの対応を表す対応データに基づいて、前記対応データの勾配を表す勾配データを導出する導出工程と、
前記勾配データに基づいて、前記対応データを伝播させる伝播工程とを備えることを特徴とする情報処理方法。

【請求項 1 3】

入力データから第 1 のデータと第 2 のデータとを取得する取得工程と、
前記第 1 のデータと前記第 2 のデータとの対応を表す対応データに基づいて、前記対応データの勾配を表す勾配データを導出する導出工程と、
前記勾配データに基づいて、前記対応データを伝播させる伝播工程とを備えることを特徴とする情報処理方法。

10

【請求項 1 4】

コンピュータを、請求項 1 乃至 1 1 のいずれか 1 項に記載の情報処理装置の各手段として機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データ間の対応を求めるための技術に関する。

20

【背景技術】

【0002】

従来から、複数のデータ間の対応を求めることにより、さまざまな処理に用いられてきた。例えば、複数の画像の局所領域間の対応を、画素値の整合度に基づいて推定する方法（例えば、画像のブロックマッチング）がある。これにより、ステレオ画像に基づく視差や距離の推定、画像上の物体の認識や追尾、画像の合成や編集をすることが可能になる。

【0003】

また、対象のデータは画像に限られず、画像以外のデータを対象とすることもできる。例えば、音声や文字などの 1 次元のデータ、動画像やボクセルなどの 3 次元のデータなど、様々な種類のデータについて対応の推定方法が検討されている。

30

【0004】

非特許文献 1 は、ステレオ画像上の各座標の対応を示す対応データ（視差データ）のマップを推定する方法を開示している。対応データのマップは、次の 2 つの処理を繰り返すことで推定される。まず、マップ上の各座標位置が示す対応データを周囲に伝播させ、評価値の高くなる対応データに置き換える。そして、マップ上の各座標の対応データをランダムに変動させることで、評価値の高くなる対応データに置き換える。なお、評価値は、対応データが示すステレオ画像上の 2 つの局所領域の整合度に応じて計算される

【先行技術文献】

【非特許文献】

【0005】

40

【非特許文献 1】“ PatchMatch stereo - stereo matching with slanted support windows ”, BMVC 2011, Michael Bleyer 1, Christoph Rheman n 1, Carsten Rother

【発明の概要】

【発明が解決しようとする課題】

【0006】

しかしながら、非特許文献 1 の方法では、対応データをランダムに変動させているため、高い評価値を持つ対応データを引き当てる確率が低い。そのため精度の良い対応データに収束させるのに時間がかかるという課題があった。

50

【 0 0 0 7 】

本発明は、上記の課題に鑑みてなされたものであり、入力データ間の対応を表わす対応データを高速に精度よく推定することを目的とする。

【課題を解決するための手段】

【 0 0 0 8 】

上記課題を解決するために、情報処理装置は、例えば、第1のデータと第2のデータとを取得する取得手段と、前記第1のデータの部分データと前記第2のデータの部分データとの対応を表す対応データに基づいて、前記対応データの勾配を表す勾配データを導出する導出手段と、前記勾配データに基づいて、前記対応データを伝播させる伝播手段とを備える。

10

【発明の効果】

【 0 0 0 9 】

本発明により、入力データ間の対応を表わす対応データを高速に精度よく推定することを目的とする。

【図面の簡単な説明】

【 0 0 1 0 】

【図1】第1の実施形態における処理の概要を説明する図である。

【図2】第1の実施形態における情報処理装置の機能ブロックを示す図である。

【図3】第1の実施形態における情報処理装置の処理フローチャートを示す図である。

【図4】第1の実施形態におけるマップの例を表す。

20

【図5】第2の実施形態における情報処理装置の機能ブロックを示す図である。

【図6】第2の実施形態における情報処理装置の処理のフローチャートを示す図である。

【図7】本願発明の情報処理装置のハードウェア構成の例を示す図である。

【発明を実施するための形態】

【 0 0 1 1 】

まず、本発明の各実施形態の詳細について説明する前に、各実施形態の情報装置のハードウェア構成について、図7を参照して説明する。同図において、CPU710は、バス700を介して接続する各デバイスを統括的に制御する。CPU710は、読み出し専用メモリ(ROM)720に記憶された処理ステップやプログラムを読み出して実行する。オペレーティングシステム(OS)をはじめ、本実施形態に係る各処理プログラム、デバイスドライバ等はROM720に記憶されており、ランダムアクセスメモリ(RAM)730に一時記憶され、CPU710によって適宜実行される。また、入力I/F740は、外部の装置(撮像装置など)から画像を情報処理装置1で処理可能な形式で入力信号として入力する。また、出力I/F750は、外部装置に、該外部装置が処理可能な形式で、対応付けされた特徴などを出力信号として出力する。

30

【 0 0 1 2 】

(第1の実施形態)

本実施形態の概要を説明する。まず、2つの異なる視点位置から対象を撮像して得られたステレオ画像を取得する。取得したステレオ画像の一方の画像(第1の画像)の局所領域(部分データ)ごとに、他方の画像(第2の画像)の対応する局所領域を求める。そして、局所領域同士の対応を表す対応データ(本実施形態では視差データ)を推定する。そして、対応データの勾配に基づいてより整合のとれる対応データを伝播させる(置き換える)ことにより、対応データの精度を上げる処理を行う。

40

【 0 0 1 3 】

なお、本実施形態では、視差データを推定する例について説明するが、後述する変形例において説明するように、視差データ以外のものを対応データとして推定することももちろん可能である。そのため、本願発明における対応データは視差データに限られない。

【 0 0 1 4 】

また、本実施形態では、ステレオ画像上の各画素についての視差データを、画像と同じ解像度を有する配列の画像形式のマップとして保持する。さらに、マップは、視差データ

50

の勾配を表わす勾配データを保持する。視差データの勾配とは、マップ上の座標の変化に対する視差データの変化を表わす。

【0015】

次に、図2を参照して本実施形態の情報処理装置200の機能構成について説明する。本実施形態の情報処理装置200は、取得部210、設定部220、更新部230、伝播部240からなる。

【0016】

取得部210は、2つの異なる視点(撮像位置)で物体を撮影した2枚の画像を取得する。画像は撮像装置から有線、無線を介して直接取得しても良いし、あらかじめ撮影された画像が記録されたサーバ等からネットワークを介して取得しても良い。また、画像は、撮像装置により撮像されたものだけでなく、コンピュータグラフィックスのようにレンダリングされた画像でも良い。本実施形態では、2枚の画像は、水平に並べた2台の撮像装置それぞれによって撮像されたものであるとし、また、撮像装置の内部パラメータと外部パラメータはキャリブレーション済であり、既知の情報とする。また、画像間の視差データを精度良く推定するためには、撮影する物体にテクスチャが多く含まれていることが望ましいため、プロジェクタを用いて、撮影する物体にテクスチャを投影した状態で撮影した画像を用いてもよい。物体にテクスチャが投影されることで、例えば無地の物体についても正確に対応データを推定できるようになる。

10

【0017】

設定部220は、取得部210で取得した2枚の画像の局所領域の対応を示す視差の初期値を設定する。視差データは、ステレオ画像のうち片側の画像を基準とし計算される。すでに述べたように、本実施形態では、画像と同じ解像度を有する配列の画像形式のマップの各座標位置に対応させて、視差データと勾配データを保持する。本実施形態では、画像I1を基準としたマップを保持する。以降、マップは、取得した画像と同じ要素数(画素数)を持つ2次元の配列形式を有するとする。この場合、例えば、マップ上の座標(u, v)に対応付けられる視差データC(u, v)は、画像I1上の座標(u, v)と、対応する画像I2上の座標との視差を表わす。ここで、視差は、画像のX軸方向の1次元の視差を表わす。すなわち、視差データC(u, v)は、(u, v)における2枚の画像間の視差の値を1次元の数値で示している。

20

【0018】

勾配データGx(u, v)とGy(u, v)は、それぞれマップのX軸とY軸の方向におけるC(u, v)の勾配を表わす1次元の数値である。また、G(u, v)と表記した場合、Gx(u, v)とGy(u, v)の両方を含めたベクトルを表わす。

30

【0019】

更新部230は、マップの各画素の勾配データG(u, v)を、その画素の周辺の視差データの分布に基づいて更新(導出)する。更新の方法の詳細については後述する。本実施形態においては、勾配データを視差データの分布によって推定するが、外部センサを利用して勾配データを推定しても良い。例えば、距離センサを用意して、距離センサが計測する距離画像に基づいて、マップ上の勾配データを設定しても良い。

【0020】

伝播部240は、マップ上の各座標位置(u, v)における視差データC(u, v)を、勾配データG(u, v)の示す傾きの方向に伝播させ、評価値が高くなる視差データに置き換える。

40

【0021】

以上のように、設定部220、更新部230、伝播部240では、(i)設定部220は、2つの画像の初期の視差データを設定する。(ii)そして、更新部230は、その初期の視差データに基づいて勾配データを生成する。(iii)そして、伝播部240は、生成された勾配データに基づいて視差データを伝播させる。(iv)そして、更新部230は、伝播させた視差データに基づいて勾配データを更新する。(v)そして、伝播部240は、更新された勾配データに基づいて、視差データを再度伝播させる。繰り返し計

50

算の1回目のみ(i)、(i i)の処理が必要となるが、それ以外の計算では、(i i i) ~ (v)の処理を繰り返すことにより、視差データの精度が真値に近づくように計算をする

次に、本実施形態の処理手順について説明する。図3は、本実施形態の処理手順を示すフローチャートである。

【0022】

(ステップS310)

取得部210は、異なる撮影位置で物体を撮影した2つの画像を取得する。本実施形態では、2つの画像のうち、一方の画像をI1、右眼の画像をI2と表わす。また、以降の説明では、画像上の画素値は、I1(u, v)、I2(u, v)と表わす。uとvは、画像上のX軸とY軸についての座標を表わす。

10

【0023】

(ステップS320)

設定部220は、マップ上の各座標の視差データの初期値を設定する。本実施形態では、視差データを座標ごとにランダムに設定する。具体的には、式(1)を使って設定する。

【0024】

【数1】

$$C(u, v) = \text{Rand} \cdot (C_{\max} - C_{\min}) + C_{\min}$$

20

式(1)

式(1)は、CmaxからCminまでの数値をランダムに設定する。ここで、Randは、0~1までの範囲で発生する一様乱数を表わす。CmaxとCminは、ステレオ画像の視差の上限と下限を表わす。CmaxとCminの値は、予め設定する固定値とする。ランダムに設定することで、各座標で異なる視差が設定される。

【0025】

マップ上の一部でも視差データの精度が高ければ、正しい視差データを伝播させることで、マップ全体の視差データの精度を向上できる。ランダムに設定した場合でも、マップ上の一部では精度の高い視差を設定することができるため本実施形態では上記のようにランダムに設定している。

30

【0026】

しかしながら、本願発明の初期値の設定はランダム以外の設定方法でも良い。具体的には、画像上の各座標について、対応データを探索的に計算するブロックマッチングによって設定しても良い。さらに、Graph CutやBelief Propagationなどの最適化アルゴリズムを利用して対応データを推定し、設定しても良い。他にも、ステレオ画像から視差を推定する場合には、別途距離センサを用意して、距離センサが計測する距離画像に基づいて、対応データのマップの初期値を設定しても良い。具体的には、距離センサが計測する距離画像を3次元点群に変換し、この点群をステレオ画像の座標系に投影することで、ステレオ画像上の各座標における距離値を求める。次に各座標の距離値から視差に変化し、この視差の値を対応データの初期値として設定する。(距離値と視差値とは1対1の関係があり、相互に変換可能である。)

40

(ステップS330)

更新部230は、マップ上の各座標の勾配データを、その周辺の視差データの分布に基づいて更新(生成)する。本実施形態では、X軸方向とY軸方向に対して、視差データの勾配を計量するフィルタを適用することで、マップ上の各座標の勾配データを計算する。

【0027】

具体的にマップ上の各座標の勾配データは、式(2)、(3)によって計算する。

50

【 0 0 2 8 】

【 数 2 】

$$G_x(u, v) = \frac{C(u+k, v) - C(u-k, v)}{2k}$$

式 (2)

【 0 0 2 9 】

【 数 3 】

$$G_y(u, v) = \frac{C(u, v+k) - C(u, v-k)}{2k}$$

式 (3)

10

式 (2)、(3) に示すように、勾配データは、注目する座標 (u, v) を中心に、k 離れた 2 か所の対応データの傾き (変化率) として計算する。ここで、k の値は、予め設定する固定値とする。

【 0 0 3 0 】

20

勾配データの計算は、視差データの分布に基づいて計算する方法であれば、上記以外のような方法でも良い。例えば、ソベルフィルタによって計算しても良いし、マップ上の視差データを微分することで計算しても良い。注目する座標の周辺の視差データの値に対して、面フィッティングを行うことで、勾配を計算しても良い。

【 0 0 3 1 】

以上の処理により計算されるマップ上のそれぞれの座標について勾配データを示したのが、図 1 (A) の線 102 である。本ステップでは、繰り返し計算の 1 回目では、S320 で設定された視差データの初期値に基づいて、勾配データを生成する。繰り返し計算の 2 回目以降では、S340 で伝播された視差データに基づき式 (2)、(3) で計算される勾配データによりそれまでの勾配データを更新する処理を行う。なお、繰り返し計算の 1 回目では、視差データの精度が低いため、その視差データに基づいて勾配データを計算しても計算精度が低い。そのため、繰り返し計算の 1 回目においては勾配を 0 ($G_x = 0$ 、 $G_y = 0$) として、設定しても良い。

30

【 0 0 3 2 】

また、マップ上の各座標の勾配データを、その周辺の勾配データに基づいて更に更新しても良い。例えば、注目する座標 (u, v) の周辺の視差データから勾配データを計算し、さらに座標 (u, v) の周辺の勾配データとの平均をとることで、座標 (u, v) の勾配データを計算しても良い。

(ステップ S340)

伝播部 240 は、マップ上の各座標の視差データを勾配データの示す傾きの方向に伝播させ、評価値が高くなる視差データに置き換える。更に、視差データの置き換えとともに勾配データも置き換える。まず、マップ上の各座標を順番に選択する。本実施形態では、ラスタスキャンの要領で、マップ上の左上から右下に向けて順番に選択する。図 4 の 410 は、マップを表わす。420 は、選択中の座標を表わす。以降、選択した座標は (u, v) と表わす。

40

【 0 0 3 3 】

次に、選択した座標 (u, v) の周辺 (近傍) の座標から視差データを伝播させる。本実施形態では、選択した座標 (u, v) の左と上の座標、すなわち座標 (u-1, v) と座標 (u, v-1) の 2 か所の視差データを座標 (u, v) に向かって伝播させる。このとき視差データは、勾配データの示す傾きの方向に変化させる。具体的には、伝播させる

50

元の座標を (s , t) としたとき、式 (4)、(5) によって計算する。

【 0 0 3 4 】

【 数 4 】

$$C'(u,v) = C(s,t) + (u-s) \cdot Gx(s,t) + (v-t) \cdot Gy(s,t)$$

式 (4)

【 0 0 3 5 】

【 数 5 】

$$G'(u,v) = G(s,t)$$

式 (5)

10

ここで、C (u , v)、G (u , v) は、座標 (s , t) の視差データと勾配データとを、選択した座標 (u , v) に向かって伝播させたそれぞれのデータを示している。式 (4) に示すように、視差データ C (u , v) は、伝播元の視差データ C (s , t) に勾配データ G (s , t) の傾き分の変化を加えて計算する。なお、勾配データは伝播元の勾配データと同じであると推定して伝播させる。以降の説明で、座標 (u , v) に元々ある (すなわち、更新前) 視差データを C 1 とする。座標 (u - 1 , v) と座標 (u , v - 1) の 2 か所の視差データを、伝播させることにより生成した視差データを C 2、C 3 とする。

20

【 0 0 3 6 】

次に、C 1、C 2、C 3 について、評価値を計算する。視差データの評価は、基準となる画像 I 1 上の座標 (u , v) と、視差を加えた (視差を考慮することで決定された) 画像 I 2 上の座標 (u + C (u , v) , v) との局所領域の整合度によって計算する。具体的に、本実施形態では、マップ上の座標 (u , v) における視差データ C の評価値 E を、式 (6)、式 (7) に示す S A D (S u m o f A b s o l u t e D i f f e r e n c e) に基づいて計算する。

30

【 0 0 3 7 】

【 数 6 】

$$E(u,v,P) = -SAD(u,v,P) = - \sum_{m=-w}^w \sum_{n=-w}^w D(u,v,m,n,C)$$

式 (6)

【 0 0 3 8 】

【 数 7 】

$$D(u,v,m,n,C) = |I1(u+m,v+n) - I2(u+m+C(u,v),v+n)|$$

式 (7)

40

式 (6) における評価値 E は、I 1 と I 2 とで比較する局所領域が整合するほど、高い値となる。w は、S A D を計算する局所領域のサイズを表わす。w の値は、予め設定する固定値とする。式 (7) の D は、座標 (u , v) を中心とした局所領域内の位置 (m , n) について、視差データの示す 2 つの座標の画素値の差を計量する関数である。本実施形態では、視差データ C の値だけずれた 2 つの座標の画素値の差の絶対値を計量する。

【 0 0 3 9 】

50

次に、座標 (u, v) における、C1、C2、C3の視差データのうち、最も評価値の高い視差データを、座標 (u, v) の新しい視差データとして、マップ上に割り当てる。この処理により、座標 (u, v) に元々ある視差データC1の評価値に比べ、伝播による視差データC2やC3の評価値が高い場合に、マップ上の視差データが置き換わる(更新される)。この場合、式(4)、(5)によって推定した視差データを、新しい視差データとして設定する。更に、座標(u,v)の視差データの置き換えとともに、勾配データも置き換える。なお、勾配データの置き換えを行わず、視差データのみ置き換えるようにしても良い。

【0040】

以上の操作を、選択する座標 (u, v) を変えて行うことで、マップ全体について視差データおよび勾配データの置き換えを行う。

10

【0041】

なお、本発明は、マップ上の左上から右下に向けて順番に選択することに限られない。例えば、マップ上の右下から左上に向けて選択しても良い。選択する方向は、何れの方向でも良い。画像を複数のブロック分割し、そのブロック毎に所定の方向に順番に選択しても良い。ランダムに座標を選択しても良い。また、伝播部を含む繰り返しの毎に、伝播部における座標の選択方法を変えても良い。例えば、奇数番の繰り返しでは、マップ上の左上から右下に向けて選択し、偶数番の繰り返しでは、マップ上の右下から左上に向けて選択するようにしても良い。

【0042】

20

また、選択した座標 (u, v) の左と上の座標、すなわち座標 (u - 1, v) と座標 (u, v - 1) の2か所の視差データを座標 (u, v) に向かって伝播させる処理を説明したが、伝播させる元の座標は、左と上以外の座標でも良い。選択した座標 (u, v) の右と下の座標、すなわち座標 (u + 1, v) と座標 (u, v + 1) の座標から伝播させても良い。2か所以上の複数の座標から伝播させても良いし、1つの座標から伝播させても良い。1つ以上離れた座標から伝播させても良い。また、周辺の座標から選択した座標に向かって伝播させても良いし、これとは逆に、選択した座標から周辺の座標に向かって伝播させても良い。

【0043】

(ステップS350)

30

S350では、伝播部240は、終了条件を満たしているかの判定を行う。本実施形態では、S330における更新部230の処理とS340における伝播部240の処理を、所定回数以上繰り返していることを終了条件とする。所定回数以上繰り返している場合には、処理を終了し、所定回数未満であれば、ステップ330の処理に戻る。所定回数の繰り返しの後、図1(C)のように、黒丸で示される求められた視差データは、破線で示される真値に収束する。

【0044】

なお、繰り返し処理の終了の判定は、マップ上の視差データの変化量に基づいて行っても良い。例えば、繰り返し計算の度に、マップ上の視差データの変化量を計算し、変化量が所定値以下であれば終了するように、判定を行っても良い。マップ上の視差データの変化量は、異なる繰り返し時点のマップの差に基づく指標であれば、どのような指標でも良い。マップ上の各座標の要素の差の平均値でも良いし、最大値や最小値でも良い。本実施形態の場合、更新部230の処理の後のマップの状態に基づいて変化量を計算しても良いし、伝播部240の処理の後のマップの状態に基づいて変化量を計算しても良い。

40

【0045】

本願発明は、各座標の周辺の視差データから勾配データを更新し、勾配データの示す傾きの方向に視差データを伝播させることにより、視差データの伝播をより精度よく行うことができる。

【0046】

以上説明した本発明にかかる処理によって視差データおよび勾配データが真値に近づい

50

ていく様子を図1を参照して説明する。

図1は、簡単のため1次元で表した各座標を横軸にし、各座標に対して視差の値を縦軸にして表したグラフである。黒丸101はそれぞれの座標に対応する視差データを表し、線102は視差データの勾配データを示す。図1(A)は、繰り返し計算の1回目の処理が終わった後の視差データ、および勾配データである。上記で説明したように、視差データの初期値はランダムに与えられるため、1回の計算ではまだ破線103によって示される視差データの真値かかれている。図1(B)は、2回目の繰り返し計算における勾配データの更新後のデータ104を示す。すなわち、図1(A)の視差データの分布に基づいて勾配データを更新した後のデータである。図1(C)は図1(B)の後に再び伝播の処理を行った後のデータであり真値に収束していることを示している。すなわち、上記説明したように、勾配データの更新処理、視差データおよび勾配データの伝播処理を繰り返すことにより、視差データを真値に収束させることができる。ここでは簡単のため、2回の繰り返し計算で真値に収束した例を示したが、物体が複雑な形状をしている場合には更に多くの繰り返し計算が必要となる。

10

【0047】

以上述べたように、本発明では、入力データの対応を表わすマップを推定する上で、対応データの分布から計算する勾配データに応じて対応データを伝播させる。これにより、精度の高いマップを推定できる。

【0048】

(第2の実施形態)

第1の実施形態では、(i)設定部220は、2つの画像の初期の視差データを設定する。(ii)そして、更新部230は、その初期の視差データに基づいて勾配データを生成する。(iii)そして、伝播部240は、生成された勾配データに基づいて視差データを伝播させる。(iv)そして、更新部230は、伝播させた視差データに基づいて勾配データを計算し、それまで保持していた勾配データを更新する。(v)そして、伝播部240は、更新された勾配データに基づいて、視差データを再度伝播させる。繰り返し計算の1回目のみ(i)、(ii)の処理が必要となるが、それ以外の計算では、(iii)~(v)の処理を繰り返すことにより、視差データの精度が真値に近づくように計算をした。

20

【0049】

それに対して、第2の実施形態では、(iv)で求められた勾配データを保持することをせず、(v)において伝播によって変化する視差データの分布に応じて、その都度勾配データを計算しながら、視差データの伝播を行う。すなわち、一部の視差データが伝播によって変更された場合に、勾配データを計算しなおす。そして、計算された新たな勾配データでその他の視差データを伝播させる処理を行う。これにより、伝播される視差データの分布が即座に反映された勾配データを計算するので、より少ない繰り返し計算で真値に収束する効果が期待できる。

30

【0050】

本実施形態では、第1の実施形態と同様に、ステレオ画像を入力データとして、画像上の各座標の対応データを高精度に推定する方法を説明する。上記で説明したように、本実施形態では、第1の実施形態とは異なり、マップは視差データのみを保持し、勾配データは保持しない。勾配データは、伝播によって変化する視差データの分布に応じてその都度周辺の視差データより計算する。

40

【0051】

次に本実施形態にかかる情報処理装置500の機能構成を、図5を参照しながら説明する。報処理装置500は、取得部510、設定部520、伝播部530からなる。

【0052】

取得部510と設定部520の処理は、第1の実施形態における取得部210、設定部220と同じである。

【0053】

50

伝播部 530 は、マップ上の各座標の視差データを、勾配データの示す傾きの方向に伝播させ、評価値が高くなる視差データに置き換える。そして、視差データが置き換えられた場合には、新しい視差データの分布から勾配データを計算しなおす

本実施形態の以降の説明では、画像上の画素の表現や、マップ上の各座標の対応データの表現は、第 1 の実施形態と同じである。

【0054】

次に、本実施形態の処理手順について説明する。図 6 は、本実施形態の処理手順を示すフローチャートである。

【0055】

(ステップ S610)

取得部 510 は、異なる撮影位置で物体を撮影した 2 枚の画像 I1 と I2 を取得する。本ステップは第 1 の実施形態における S310 と同様である。

【0056】

(ステップ S620)

設定部 220 は、マップ上の各座標の視差データの初期値を設定する。本ステップは、第 1 の実施形態の S320 と同じである。

【0057】

(ステップ S630)

伝播部 250 は、まず、マップ上の各座標の勾配データを、周辺の視差データの分布から計算(生成)する。そして生成された勾配データの示す傾きの方向に視差データを伝播させ、評価値が高くなる視差データで設定部 220 で設定された視差データの初期値を置き換える。

【0058】

具体的には、マップ上の各座標位置について、その周辺の視差データから勾配フィルタを使って勾配データを計算し、勾配データの傾きの方向に視差データを伝播させる。

【0059】

これを行うために、まず、マップ上の各座標を順番に選択する。ここでは、第 1 の実施形態と同じように、ラスタスキャンの要領で、マップ上の左上から右下に向けて順番に選択する。図 4 の 410 は、マップを表わす。320 は、選択中の座標を表わす。以降、選択した座標は (u, v) と表わす。

【0060】

次に、選択した座標 (u, v) の周辺の座標から視差データを伝播させる。本実施形態では、選択した座標 (u, v) の左と上の座標、すなわち座標 $(u - 1, v)$ と座標 $(u, v - 1)$ の 2 か所の視差データを座標 (u, v) に向かって伝播させる。このとき視差データは、勾配データの示す傾きの方向に変化させる。具体的には、伝播させる元の座標を (s, t) としたとき、式(8)によって計算する。ただし、第 1 の実施形態とは異なり、マップは勾配データを保持していないため、勾配データは現在のマップ上の視差データの分布に応じて計算する。具体的に、式(8)の勾配データ $G_x(s, t)$ 、 $G_y(s, t)$ は、式(9)、(10)に示す勾配フィルタを使って計算する。

【0061】

【数 8】

$$C'(u, v) = C(s, t) + (u - s) \cdot G_x(s, t) + (v - t) \cdot G_y(s, t)$$

式(8)

【0062】

10

20

30

40

【数 9】

$$G_x(s, t) = \frac{C(s+k, t) - C(s-k, t)}{2k}$$

式 (9)

【0063】

【数 10】

$$G_y(s, t) = \frac{C(s, t+k) - C(s, t-k)}{2k}$$

式 (10)

10

ここで、 $C(u, v)$ は、座標 (s, t) の視差データを、選択した座標 (u, v) に向かって伝播させた対応データを示す。式 (8) に示すように、対応データ $C(u, v)$ は、伝播元の対応データ $C(s, t)$ に勾配データ $G(s, t)$ の傾き分の変化を加えて計算する。

20

以降の説明で、座標 (u, v) に元々ある視差データを C_1 とする。座標 $(u-1, v)$ と座標 $(u, v-1)$ の2か所の対応データを、伝播させることにより生成した対応データを C_2 、 C_3 とする。

【0064】

次に、 C_1 、 C_2 、 C_3 について、評価値を計算する。具体的な評価値の計算は、第1の実施形態と同じであり、式 (6)、(7) に基づいて計算する。

【0065】

次に、座標 (u, v) における、 C_1 、 C_2 、 C_3 の視差データのうち、最も評価値の高い視差データを、座標 (u, v) の新しい視差データとして、マップ上に割り当てる。

【0066】

この処理により、座標 (u, v) に元々ある視差データ C_1 の評価値に比べ、伝播による視差データ C_2 や C_3 の評価値が高い場合に、マップ上の視差データが置き換わる。

30

【0067】

以上の操作を、選択する座標 (u, v) を変えて行うことで、マップ全体について視差データの置き換えを行う。

【0068】

(ステップ S640)

S640では、本フローの終了判定を行う。具体的な内容は、第1の実施形態と同じである。

【0069】

本実施形態によれば、伝播される視差データの分布が即座に反映された勾配データを計算するので、より少ない繰り返し計算で真値に収束する効果が期待できる。

40

【0070】

(第3の実施形態)

本実施形態では、上記実施形態において求められた視差データを利用するアプリケーションについて説明する。

【0071】

上記実施形態では、マップとしてステレオ画像における部分領域の視差を推定した。推定した視差を以下の式を用いて計算することにより、距離値を求めることが可能となる。すなわち、ステレオ画像を撮像したカメラの相対的な位置姿勢が既知の場合、画像間の視

50

差と距離値は1対1の関係を持つため、視差と距離値は相互に変換できる。

【0072】

【数11】

$$D = \frac{f \cdot B}{C}$$

式(11)

10

ここで、 f はカメラの焦点距離、 B はステレオカメラのベースライン(ステレオカメラの間隔)、 D は距離値、 C は視差を表わす。また、距離値は、3次元座標に変換して出力しても良い。距離値や3次元座標を出力することで、物体の3次元形状を計測するために利用できる。上記のように、対応データを変換する場合、上記実施形態の情報処理装置は、さらに対応データ変換部を備えても良い。また、対応データとして、視差ではなく距離値をマップ上に保持するように変更しても良い。

【0073】

また、出力した距離値や3次元座標に基づいて、画像上で観測する物体の位置姿勢を計測しても良い。例えば、画像上で観測する物体の3次元モデルが既知の場合、3次元モデルと推定した距離値とのICP(Iterative Closest Point)を行うことで、物体の位置姿勢を推定できる。

20

【0074】

さらに、別の利用アプリケーションを説明する。2つの画像間の視差データに基づいて、画像間の視差データの示すずれを最小化するように2つの画像を重ねることで、2枚の画像からパノラマ画像を合成する方法として利用できる。また、異なる視差データの傾向を閾値によって分けることで、画像上に映る背景と前景の物体を分けて、物体を検出する方法として利用できる。例えば、ステレオ画像から計算した距離値の情報に基づいて、距離センサを利用するSLAM(Simultaneous Localization and Mapping)の入力データに利用できる。例えば、3次元のボクセルデータの対応データを表わすマップを推定する場合、複数の3次元のボクセルデータを対応データに基づいて合成し、精度を向上させることができる。

30

【0075】

以降、第1と第2の実施形態に対する変形例を説明する。なお、以下の変形例による変形が施された実施形態の処理結果を用いて、第3の実施形態の処理を行うことはもちろん可能である。

【0076】

(変形例1)

本変形例では、対応データとして視差データ以外の情報を推定する例を説明する。まず、2次元のベクトルを用いる例を説明する。

40

【0077】

上記実施形態では、画像のX軸方向の1次元の視差を対応データとした。本変形例では、画像のX軸とY軸方向における2次元の視差を対応データとする。なお、2次元の視差を推定する場合、画像上の局所領域の対応が必ずしも画像のX軸方向にない条件においても、2枚の画像の対応データを精度良く推定できる。

【0078】

本変形例では、2枚の画像の対応を表わすマップを推定する。対応を計算する2枚の画像のうち、片方をI1、もう片方をI2と表わす。画像上の画素値は、I1(u, v)、I2(u, v)と表わす。 u と v は、画像上のX軸とY軸上の座標を表わす。

【0079】

50

また、マップは、2枚の画像のうち片側の画像を基準としたパラメータを保持する。本変形例では、画像I1を基準としたマップを保持する。従って、マップは、画像と同じ要素数（画素数）を持つ2次元の配列とする。この場合、例えば、マップ上の座標（ u, v ）にある2次元の視差データは、画像I1上の座標（ u, v ）と、対応する画像I2上の座標との差を表わす。ここで、2次元の視差データは、X軸とY軸についての座標の差を表わす2次元のベクトルである。

【0080】

また、マップ上の各座標の視差データを $C(u, v)$ 、勾配データを $G_x(u, v)$ 、 $G_y(u, v)$ と表わす。対応データ $C(u, v)$ は、2枚の画像上の視差を表わす2次元ベクトルである。ここで、 $C(u, v)$ は、X軸方向の差 $C_x(u, v)$ と、Y軸方向の差 $C_y(u, v)$ の2つの成分を含む。勾配データ $G_x(u, v)$ と $G_y(u, v)$ は、マップのX軸とY軸の方向における $C(u, v)$ の勾配を表わす2次元のベクトルである。また、 $G(u, v)$ と表記した場合、 $G_x(u, v)$ と $G_y(u, v)$ の両方を含めたベクトルを表わす。なお、第2の実施形態に示すように、マップ上に勾配データを保持しないようにしても良い。

10

【0081】

本変形例の構成は、第1或いは第2の実施形態と同じである。本変形例の処理手順は、第1或いは第2の実施形態で説明した計算式を、本変形例の対応データの次元に変更して計算する。例えば、設定部では、2次元の対応データの2つの成分について、初期値を設定する。更新部では、2次元の対応データの2つの成分について、勾配を計算する。伝播部では、マップ上の各座標の視差データを、勾配データの2つの成分の示す傾きの方向に伝播させ、評価値が高くなる視差データに置き換える。また伝播部の、式(6)、式(7)で求める評価値については、2次元の対応データの2つの成分を用いて、式(7)を変形した式(12)により計算する。

20

【0082】

【数12】

$$D(u, v, m, n, C) = |I1(u + m, v + n) - I2(u + m + C_x(u, v), v + n + C_y(u, v))|$$

式(12)

30

ここで、 $C_x(u, v)$ と $C_y(u, v)$ は、視差データ $C(u, v)$ に含まれるX軸とY軸に関する視差を表わす。式(11)は、X軸とY軸それぞれに視差のある条件について、画素値の差を計量する。

【0083】

以上のように操作を変更することで、対応データは2つの成分を持つ2次元のベクトルに拡張可能である。

【0084】

以上示すように、本発明における対応データは2次元のベクトルでも良い。同様に、本発明における対応データは2次元以上のベクトルでも良い。例えば、画像上の局所領域のスケールを表わす情報や、局所領域の回転を表わす情報などを、対応データに加えても良い。勾配データは、対応データの次元数に応じて保持する。すなわち、対応データの次元の成分毎に、勾配データを保持する。

40

【0085】

例えば、視差データとしてスケールや回転を表わす情報を保持する場合、視差データの評価値は、視差データに含まれるスケールや回転に応じて計算する。具体的には、例えば、スケール C_s を視差データの成分として加える場合、式(12)は式(13)のように変形する。

【0086】

【数 1 3】

$$D(u, v, m, n, C)$$

$$= |I1(u + m, v + n) - I2(u + Cs(u, v) \cdot m + Cx(u, v), v + Cs(u, v) \cdot n + Cy(u, v))|$$

式 (13)

式 (12) では、比較を行う画像 I 1 上の注目座標 (u, v) を基準に、局所領域の内の各座標にスケール G s 分の変化を加えて、画像 I 2 上の座標との画素の比較を行う。また、視差データに局所領域の回転を表わす情報を加える場合も同様で、回転による座標の変化を加えて、I 1 と I 2 上の画素の比較を行う。

10

【0087】

また、対応データとして画像間の視差データの代わりに、距離値を利用しても良い。ステレオカメラの相対的な位置姿勢が既知の場合、画像間の視差と距離値は 1 対 1 の関係を持つため、視差と距離値は相互に変換できる。

【0088】

(変形例 2)

上記実施形態では、2次元データであるステレオ画像について、各画像間における座標の対応を表わす 2次元のマッピングを推定する例を説明した。ただし、本発明の方法は、2次元以外を入力データにも拡張できる。2次元以外のデータのバリエーションとしては、例えば、音声や文字列などの 1次元の入力データや、動画像や 3次元のボクセルなどの 3次元の入力データがある。

20

【0089】

1次元のデータを入力とする場合、第 1 と第 2 の実施形態の処理は、次のように変更する必要がある。まず、入力データの各座標の対応を表わすマッピングは、入力データの次元数と同じ 1次元のマッピングで表現する。つまり、1次元の入力データの各座標についての対応データを 1次元のマッピングとして保持する。また、対応データの勾配を表わす勾配データは、1次元の座標に対する対応データの勾配を計算する (第 1 と第 2 の実施形態では、入力データは 2次元であるため、X軸と Y軸のそれぞれに対する対応データの勾配を計算する)。

30

【0090】

具体的に、1次元の入力データとして、2つの音声データから、各座標 (音声データの場合、時刻とも表現できる) の対応データを推定する例を説明する。音声の場合、入力データの各座標の値は画素値ではなく、周波数で表わすものとする。この場合、マッピングは、音声の要素数と同じ要素数を持つ 1次元の配列とする。対応データは、2つの音声データの座標の差を 1次元のベクトルとして表わす。勾配データは、マッピング上の 1次元の座標空間における対応データの勾配を表わす。上記データ構造に基づいて、第 1 或いは第 2 の実施形態で説明した処理の順序で、マッピング上の対応データの分布から勾配データを計算し、勾配データに応じて視差データを伝播させることで、マッピング上の対応データの精度を向上させる。

40

【0091】

また、多次元のデータを入力とする場合、第 1 と第 2 の実施形態の処理は、次のように変更する必要がある。まず、入力データの各座標の対応を表わすマッピングは、入力データの次元数と同じ次元数で表現する。対応データの勾配を表わす勾配データは、入力データの各次元の方向についての勾配を保持する。

【0092】

具体的に、3次元の入力データを扱う例として、2つの動画像から、各座標の対応データを推定する変形例を説明する。この場合、マッピングは、動画像と同じ要素数 (画素数 × フレーム数) を持つ 3次元の配列とする。対応データは、注目する座標と対応する座標との差を 3次元のベクトルとして表わす。勾配データは、対応データの 3つの成分のそれぞれ

50

について、X軸とY軸と時間軸のそれぞれの方向に対する勾配を表わす。上記データ構造に基づいて、第1或いは第2の実施形態で説明した処理の順序で、マップ上の対応データの分布から勾配データを計算し、勾配データに応じて対応データを伝播させることで、マップ上の対応データの精度を向上させる。

【0093】

(変形例3)

第1と第2の実施形態では、ステレオ画像が入力された場合に、それぞれの画像の部分領域同士の対応を見つける処理を示した。しかしながら、本願発明は、ステレオ画像のように複数のデータの入力に限られるものではなく、入力データは1つでも良い。例えば、1つの画像を入力として、各座標について最も類似する領域を、同一の画像内で対応付けるように計算しても良い。この場合、ステレオ画像の代わりに、同一の画像を入力するように、第1と第2の実施形態を変更する(具体的には、 $I_1 = I_2$ として処理を行えば良い)。

10

【0094】

ただし、画像上の局所領域を同一の画像内で対応付ける場合、第1と第2の実施形態をそのまま適応すると、マップ上の対応データは視差=0として推定されることが予想される。これは、視差=0の状態が最も対応データの評価値が高いためである。これを避けるためには、対応データの視差の値がほぼ0の時に、視差データの評価値が悪くなるように、評価値の計算式にペナルティを加えるようにすればよい。

【0095】

20

また、入力データは3つ以上でも良い。例えば、3台のカメラで撮影した画像に基づいてマップを推定する場合を考える。この場合、所定のカメラの画像の座標系を基準にマップを作成し、マップ上で保持する視差データが示す3台のカメラの画像の各領域の整合度を利用して評価値を計算する。3台分のカメラの画像情報を利用することで、マップ上の視差データの評価値をより正確に計算できる。

【0096】

(変形例4)

上記実施形態では、視差データの評価値を行う際に、SADを利用して画像領域の比較を行う処理をもって説明した。ただし、視差データの評価は、比較する画像上の局所領域が整合するほど、評価値が高くなるような計算方法であれば、どのような方法でも良い。SSD(Sum of Squared Difference)を利用しても良いし、NCC(Normalized Cross-Correlation)を利用しても良い。

30

【0097】

また、局所領域の比較計算は、勾配データに応じて、比較する局所領域の部位ごとに視差データを変化させても良い。この場合、式(6)、式(7)で求める評価値は、式(7)を変形した式(14)、(15)により計算する。

【0098】

【数14】

$$D(u, v, m, n, C, G) = |I1(u + m, v + n) - I2(u + m + F(u, v, m, n, C, G), v + n)|$$

40

式(14)

【0099】

【数15】

$$F(u, v, m, n, C, G) = C(u, v) + G_x(u, v) \cdot m + G_y(u, v) \cdot n$$

式(15)

ここで、式(15)の関数Fに示すように、勾配データ G_x 、 G_y と、局所領域の中心を基準とした位置 (m, n) に応じて、視差データ $C(u, n)$ の値(視差の値)を増減

50

させる。勾配データに基づいて、局所領域の部位ごとに比較する座標を計算することで、より精度よく評価値を計算できる。

【0100】

(変形例5)

第1の実施形態では、更新部230と伝播部240とを含む処理を繰り返す例を説明した。しかし、本発明では、繰り返しの度に、マップ上の視差データ、勾配データをリセットしなければ、更新部230と伝播部240の順序はどちらでも良い。繰り返し計算の中で、少なくとも1回は更新部230で計算した勾配データを、伝播部240で利用する方法であればどのような処理順序でも良い。この条件下であれば、繰り返し毎に更新部230と伝播部240の処理の並びを変えても良いし、更新部230と伝播部240の処理の回数を減らしても増やしても良い。

10

【0101】

具体的には、例えば、繰り返しの1回目には伝播部240の処理のみを行い、繰り返しの2回目以降には更新部230と伝播部240の処理を行うようにしても良い。

【0102】

(その他の実施形態)

本発明は、上述の実施形態の1以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける1つ以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1以上の機能を実現する回路(例えば、ASIC)によっても実現可能である。

20

【0103】

<効果>

本発明では、入力データの対応を表わすマップを推定する上で、対応データの分布から計算する勾配データに応じて対応データを伝播させる。これにより、精度の高いマップを推定できる。

【0104】

<定義>

本発明において、入力データとは、対応を推定する対象である。入力データは、情報を所定の次元で保持する。例えば、周波数を1次元のデータとして保持する音声データや、輝度値を2次元のデータとして保持する画像データなどがある。対応データは、入力データの2つの領域の対応関係を表わす情報であれば、どのような情報でも良い。入力データの2つの領域の座標の差を表わす情報でも良いし、さらにスケールや回転を加えた情報でも良い。マップは、入力データの少なくとも一部の座標についての対応データを要素として保持する。勾配データは、マップ上の対応データの勾配を表わす。ここで、勾配とは、マップ上の座標の変化に対する対応データの変化を表わす。マップは、さらに勾配データを保持しても良い。

30

【0105】

設定部は、マップ上の対応データの初期値を設定する。設定部は、マップ上の少なくとも一部の対応データについて、真値に近い値を設定できる方法であれば、どのような方法でも良い。ランダムに設定しても良いし、ブロックマッチングなどの他の対応データの推定方法を利用して設定しても良い。

40

【0106】

更新部は、対応データの分布に基づいて、勾配データを計算する。

【0107】

伝播部は、マップ上の各座標の対応データを、勾配データの示す傾きの方向に伝播させる。このとき、伝播させた対応データの評価値が、マップ上に元々あった対応データの評価値よりより高くなる場合に、マップ上の対応データを置き換える。伝播部では、対応データのみを伝播させても良いし、さらに勾配データを伝播させても良い。

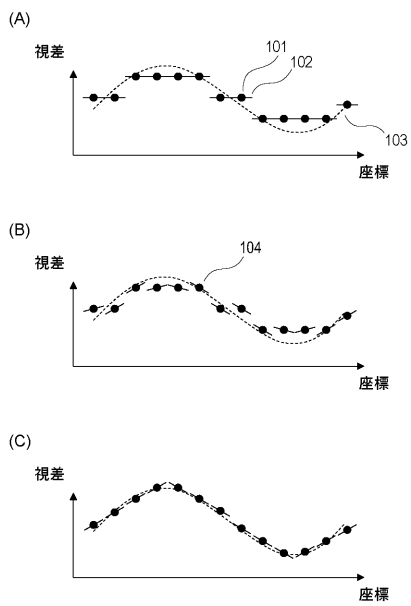
【0108】

評価値は、対応データが示す入力データの2つの領域の整合度に応じて計算する。評価

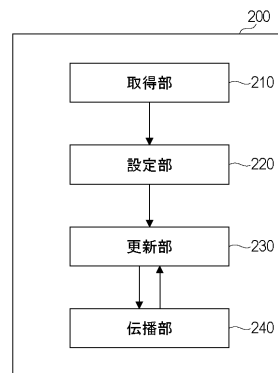
50

値は、2つの領域が整合する場合に高くなる計算方法であれば、どのような計算方法でも良い。

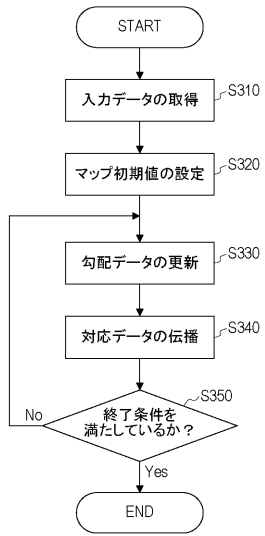
【図1】



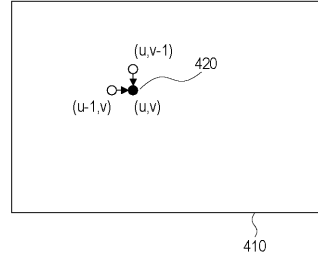
【図2】



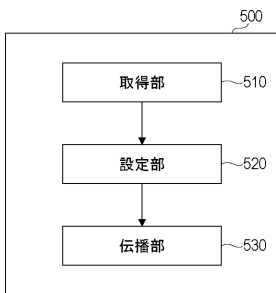
【 図 3 】



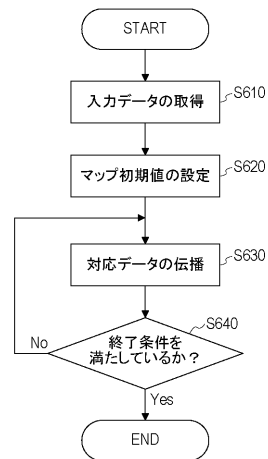
【 図 4 】



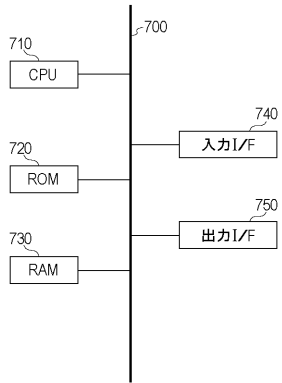
【 図 5 】



【 図 6 】



【 図 7 】



フロントページの続き

(72)発明者 片山 昭宏

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

(72)発明者 藤木 真和

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

Fターム(参考) 5B057 BA02 BA04 CA08 CA13 CB08 CB13 DB03 DB09 DC08