

位置に依存した固有空間上でのランドマーク照合による カメラの位置・姿勢推定の高速化

Reduction of computational cost for landmark based camera parameter estimation by using position-dependent eigen space

武富 貴史[†] 佐藤 智和[†] 横矢 直和[†]
Takafumi Taketomi Tomokazu Sato Naokazu Yokoya

1. はじめに

カメラの位置及び姿勢を推定する技術は、拡張現実感やロボットナビゲーションなどの分野に応用が可能である。中でも従来提案されている自然特徴点ランドマークを用いた手法 [1][2] は、物理的なインフラを用いずにカメラの位置・姿勢を推定できるという特徴を持つ。特に大江らの手法 [1] は、屋外などの広範囲な環境においても累積誤差を含まないカメラ位置・姿勢を推定でき、広い分野への応用が期待できる。しかし、大江らの手法 [1] では主に自然特徴点の対応付け処理に多くの計算コストを要するため、リアルタイムでの処理が困難であるという問題が残されている。

この問題に対して、本研究では、あらかじめカメラの位置・姿勢推定を行う空間を格子状に分割した上で、格子ごとに位置に依存した固有空間を作成し、固有空間上での距離を用いて対応点候補を絞りこむことで、照合すべき自然特徴点の数を削減し、対応付け処理の計算コストを低減する。また、自然特徴点抽出の際に、比較的容易に算出できる輝度勾配の方向を用いて照合すべき自然特徴点の削減を行う。

2. 固有空間上でのランドマーク照合による カメラ位置・姿勢推定の高速化

以下ではまず、従来のランドマークデータベース（以下 LDB）を用いたカメラ位置・姿勢推定手法 [1] について概説し、本研究で用いる高速化手法のアイデアを述べる。次に、輝度勾配の方向と位置に依存した固有空間上での距離を用いた対応点候補の絞りこみによるカメラの位置・姿勢推定の高速化手法についてより詳しく述べる。

2.1 LDB を用いたカメラ位置・姿勢推定の概要

本研究の基礎となる手法 [1] の処理の流れについて概説する。図 1 に示すように、LDB を用いたカメラ位置・姿勢の推定は、オフラインでの LDB 構築の処理 (A) とオンラインでのカメラ位置・姿勢の推定処理 (B) から成る。オフライン処理 (A) では、まず利用環境を全方位カメラを用いて撮影する (A-1)。次に、撮影した動画画像中の自然特徴点を自動追跡することで、自然特徴点の 3 次元位置と全方位カメラのカメラパラメータを推定する (A-2)。最後に、ランドマークの情報として以下の情報を登録する (A-3)。

- 自然特徴点の 3 次元位置
- 多重スケールの画像テンプレート
- 画像テンプレートの法線ベクトル

オンライン処理 (B) では、まず何らかの方法によりカメラの初期位置・姿勢を取得し (B-1)、続いてカメラ位置・姿勢推定処理を逐次繰り返す。逐次推定では、まず、前のフレームで推定されたカメラ位置・姿勢を用いて推定に用いるランドマークをデータベース中から選択する (B-2)。次に、入力画像中の自然特徴点を検出し (B-3)、ランドマークと対応付ける (B-6)。最後に、得られた対応関係を用いてカメラの位置・姿勢を推定する (B-7)。

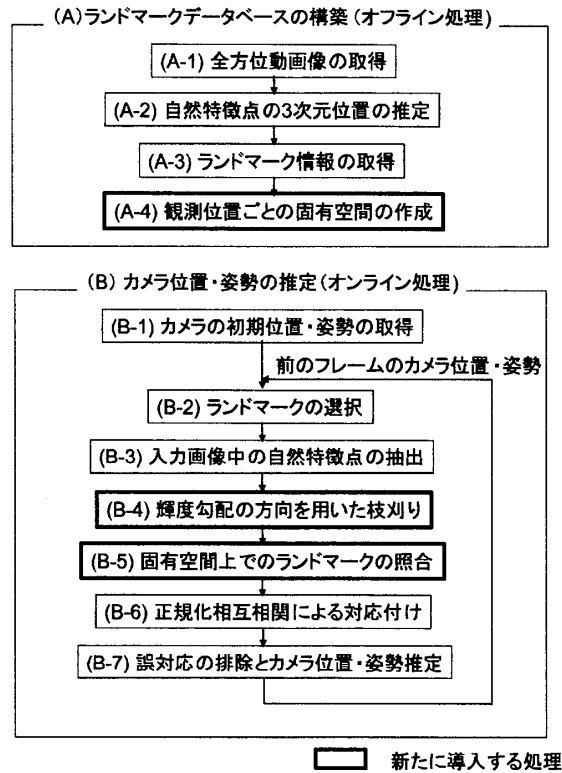


図 1: 全体の処理の流れ

従来手法では、主に自然特徴点とランドマークの対応付け処理 (B-6) に比較的多くの処理時間を必要としており、リアルタイムでの処理が困難である。この処理では、処理 (B-3) で検出された全ての自然特徴点に対して、画像の撮影姿勢に依存しない画像テンプレートの作成を行っている。このような画像テンプレートは、入力画像をカメラ位置を中心とする球面上に投影することで作成するが、各画素の投影計算に多くの計算コストがかかるという問題がある。そこで本稿では、ランドマークに対する対応付け候補となる自然特徴点を絞りこみ、高速化を図る。具体的には、まずオフライン処理 (A) において位置に依存した固有空間をあらかじめ作成しておく (A-4)。次に、オンライン処理 (B) では、自然特徴点の輝度勾配の方向を用いて対応点候補の絞りこみを行う (B-4)。また、(A-4) で作成した固有空間上での距離を用いて対応点候補の絞りこみを行うことで、照合すべき自然特徴点の数を削減し (B-5)、対応付け処理 (B-6) の計算コストを低減する。

2.2 輝度勾配の方向を用いた 対応点候補の絞りこみ

本研究では、Harris オペレータを用いて自然特徴点の抽出を行うが、Harris オペレータによる自然特徴点の抽出処理では、特徴点位置の決定のために、一定ウィンドウ内における輝度勾配ベクトルの分散が最大となる方向 (以下、主要な輝度勾配方向) を用いている。この輝度勾配の情報は、一般的には特徴点位置の抽出後に破棄され

[†] 奈良先端科学技術大学院大学
Nara Institute of Science and Technology (NAIST)

るが、本研究では特徴点抽出時に算出された主要な輝度勾配方向を保存しておき、対応点候補の枝刈りに用いる。具体的には、ランドマークの画像テンプレートから算出される主要な輝度勾配方向と、入力画像中の自然特徴点から算出される主要な輝度勾配方向の成す角が一定の閾値を超える自然特徴点を、対応付けの候補から排除することで対応点候補を削減する。

2.3 固有空間上でのランドマークの照合による対応点候補の枝刈り

本研究では、あらかじめカメラ位置・姿勢推定を行う空間を格子状に分割した上で、格子ごとに位置に依存した固有空間を作成し(A-4)、固有空間上での距離を用いて対応点候補を枝刈りする。以下では、(1)位置に依存した固有空間の生成(オフライン処理)、(2)固有空間上での非類似度を用いた対応点候補の枝刈り(オンライン処理)、について詳しく述べる。

(1) 位置に依存した固有空間の生成

ここではまず、カメラ位置・姿勢推定を行う空間を $N \times N$ の格子に分割し、各格子 $G_{i,j}$ ($i = 1, 2, \dots, N, j = 1, 2, \dots, N$) から観測できるランドマークのリスト $L_{i,j}$ を作成する。リストを作成した後、各格子ごとに以下の手順で固有空間を生成する。まず、各格子から観測できる全てのランドマークについて $n \times n$ 画素の画像テンプレートを、 n^2 次元のベクトルとして主成分分析することで、固有ベクトル \mathbf{x}_k ($k = 1, 2, \dots, n^2$) を算出する。次に、固有ベクトルを固有値の大きい順に並べ替え、累積寄与率が一定の閾値以上となるように K 個の固有ベクトルを選択する。最後に、選択された固有ベクトルから、以下のように格子 $G_{i,j}$ の固有空間への射影行列 $M_{i,j}$ を求める。

$$M_{i,j} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_K]^T \quad (1)$$

ここで、 K は固有空間の次元数に対応する。本研究では、各格子 $G_{i,j}$ に対して作成したランドマークのリスト $L_{i,j}$ と固有空間への射影行列 $M_{i,j}$ をデータベースに登録する。

(2) 固有空間上での距離を用いた対応点候補の枝刈り

オンライン処理では、(1)で作成した K 次元の固有空間上に全てのランドマークと入力画像上の自然特徴点の画像パターンを射影し、低次元の空間上で高速にランドマークを照合することで、対応付けの候補となる自然特徴点の枝刈りを行う。具体的には、まず前のフレームで推定されたカメラ位置が属する格子 $G_{i,j}$ に登録されているリスト $L_{i,j}$ を利用して、推定に用いるランドマークを高速に選択する(B-2)。次に、入力画像中の自然特徴点を抽出し(B-3)、ランドマークの画像パターン \mathbf{p} 、入力画像中の自然特徴点を中心とする画像パターン \mathbf{v} を、次式を用いて固有空間に射影する。

$$\mathbf{p}' = M_{i,j} \mathbf{p} \quad (2)$$

$$\mathbf{v}' = M_{i,j} \mathbf{v} \quad (3)$$

ここで、 \mathbf{p}' 、 \mathbf{v}' は固有空間上でのランドマークと自然特徴点のベクトルを表す。

最後に、固有空間上での非類似度 $d(\mathbf{p}', \mathbf{v}')$ を用いて、 $d(\mathbf{p}', \mathbf{v}')$ が閾値未満の特徴点を候補から排除することで自然特徴点の枝刈りを行う。本研究では明るさの変化に対応するために、 $d(\mathbf{p}', \mathbf{v}')$ を以下の \cos 類似度で定義する。

$$d(\mathbf{p}', \mathbf{v}') = \frac{\mathbf{p}' \cdot \mathbf{v}'}{\|\mathbf{p}'\| \|\mathbf{v}'\|} \quad (4)$$

3. 実験

提案手法の有効性を示すために、処理時間と位置推定精度について従来手法との比較を行った。本実験ではビデオカメラ (SONY DSR-DP-150) を用いて撮影した動

画像 (解像度 720×480 画素、プログレッシブスキャン、15fps、300 フレーム) に対してカメラの位置・姿勢推定を行った。推定に使用するランドマークの画像テンプレートは 15×15 画素、カメラの内部パラメータは既知とし、提案手法、従来手法ともに初期フレームのカメラ位置・姿勢は手動で与えた。また、提案手法において各格子の大きさは $5m \times 5m$ とし、固有空間の生成における累積寄与率は 95% とした。これにより生成された、固有空間の次元数は平均 12 次元となった。輝度勾配ベクトルの方向を用いた枝刈りを行う際の閾値は 20 度とし、固有空間上でランドマークの照合を行う際の非類似度 $d(\mathbf{p}', \mathbf{v}')$ の閾値は 0.99 とした。

PC (Xeon 3.0GHz 4CPU、メモリ 16GB) を用いた処理時間の比較を表 1 に示す。提案手法では、主要な輝度勾配方向を算出・保存するため、自然特徴点の抽出処理 (B-3) に 29ms を要するが、輝度勾配の方向を用いた枝刈りにより、約 40% の自然特徴点に対応点候補から排除された。また、固有空間上でのランドマークの照合 (B-5) に 51ms を要するが、固有空間を用いた自然特徴点の枝刈りによって約 40% のランドマークに対応点候補から排除された。その結果、全体で 77% の自然特徴点が枝刈りされ、対応付け処理 (B-6) に要する時間が 141ms (従来手法の約 1/3) となった。さらに、推定に用いるランドマークを、あらかじめ観測位置によってリスト化しておくことで、ランドマーク選択の処理時間も 2ms (従来手法の約 1/10) に短縮された。提案手法の全体の処理時間は 350ms (従来手法の約 3/5) であり、自然特徴点の枝刈りを行うことで、処理の高速化が実現できた。ただし、提案手法を用いても、1秒間に処理できる画像の枚数は約 3 枚であり、更なる高速化が必要である。

次に、従来手法および提案手法によって推定されたカメラの位置・姿勢を正解データと比較することによって推定精度の比較を行った。正解データは、10 フレームおきに手動でランドマークの対応付けを行い作成し、正解データのうち再投影誤差が 1.5 画素未満のフレームを評価対象とした。提案手法で推定されたカメラの平均位置誤差は 264mm であり、従来手法で推定されたカメラの平均位置誤差 220mm に対してわずかに増加した。

表 1: 処理時間の比較 (ms)

処理	従来手法	提案手法	処理時間の比
(B-2)	30	2	0.07
(B-3)	17	29	1.71
(B-5)	-	51	-
(B-6)	441	141	0.32
(B-7)	111	127	1.14
合計	599	350	0.58

4. まとめ

本稿では、位置に依存した固有空間上でランドマークの照合と自然特徴点の輝度勾配方向を用いた枝刈りによってカメラの位置・姿勢推定の高速化を行う手法を提案した。実験により従来手法と比較して、高速にカメラ位置・姿勢推定が行えることを確認した。ただし、推定処理のリアルタイム化には、更なる高速化が必要である。

参考文献

- [1] 大江, 佐藤, 横矢, "幾何学的位置合わせのための自然特徴点ランドマークデータベースを用いたカメラ位置・姿勢推定", 日本バーチャルリアリティ学会論文誌, Vol. 10, No. 3, pp. 285-294, Sep. 2005
- [2] I. Gordon and D. G. Lowe: "Scene modelling, recognition and tracking with invariant image features," Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR2004), pp. 110-119, 2004.