

ジャイロセンサとランドマークデータベースを用いた モーションブラーにロバストなカメラ位置・姿勢推定

松田 幸大 池田 聖 佐藤 智和 横矢 直和
奈良先端科学技術大学院大学 情報科学研究科
{ kodai-m, sei-i, tomoka-s, yokoya }@is.naist.jp

1. はじめに

動画像からカメラの位置・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への応用が可能である。これらの応用分野では、位置・姿勢の推定に誤差が蓄積せず、マーカやセンサなど物理的なインフラの整備を必要としない手法が望まれている。これらを満たすカメラの位置・姿勢推定手法として、事前に環境内の撮影及び計測を行い、自然特徴点の三次元位置やテクスチャなどをランドマークデータベース（以下、LDB）としてデータベース化し、入力画像上の特徴点をLDB内から探索することでカメラの位置・姿勢を推定する手法が提案されている[1,2]。これらの手法では、一旦LDBが構築され、初期フレームのカメラの位置・姿勢が与えられると、以降は、人手を介すことなく、カメラの位置・姿勢を推定できる。しかし、カメラの動きによりモーションブラーが生じると、入力画像上の特徴点とLDB内のランドマークを正しく対応付けられず、カメラの位置・姿勢の推定誤差が増大する問題がある。

そこで本稿では、LDBとジャイロセンサにより得られるカメラの角速度を併用することで、モーションブラーが生じた画像に対してもカメラの位置・姿勢を正しく推定する手法を提案する。提案手法では、LDB内の自然特徴点周辺のテクスチャに対し、センサにより得られる角速度を用いて入力画像と同様のモーションブラーを再現し、マッチング時における対応付けの精度向上を図る。ただし、本稿では、カメラの内部パラメータ、カメラとジャイロセンサの同期、LDBは与えられているものとする。

2. ジャイロセンサと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

提案手法では、オフライン処理で構築したLDB、動画像の第*i*フレームに対してジャイロセンサから得られる角速度 ω_i 、第(*i*-1)フレームにおけるカメラの位置・姿勢 M_{i-1} を用いて、第*i*フレームにおけるカメラの位置・姿勢 M_i を推定する。ただし、初期フレームのカメラの位置・姿勢とシャッタースピード s は既知とする。

LDBは、各ランドマークに対する世界座標系での三次元位置 \mathbf{l} に加えて、LDB構築時に各々のランドマークを撮影したカメラ位置とそれに対応するランドマークを中心とした多重スケールの画像テンプレート、テンプレートの法線ベクトルからなる[1]。本手法におけるLDBを用いたカメラの位置・姿勢の推定手順は、Step1:推定に用いるランドマークの選択、Step2:モーションブラーを考慮したランドマークの探索、Step3:投影誤差最小化によるカメラ位置・姿勢推定、からなる。以下では、ジャイロセンサによる計測結果を新たに利用するStep1 Step2について詳述する。

Step1: 推定に用いるランドマークの選択

ジャイロセンサによる第*i*フレームの角速度 ω_i と直前フレームまでの情報を用いて、第*i*フレームにおけるカメラの暫定的な位置・姿勢 M_i を算出し、最終的なカメラの位置・姿勢 M_i の推定に利用するランドマークをLDBから選択する。ここでは第(*i*-1)フレームにおけるカメラの位置・姿勢 M_{i-1} を回転行列 R_{i-1} 及び並進ベクトル \mathbf{p}_{i-1} で表現し、ランドマークの三次元位置 \mathbf{l} とカメラの暫定的な位置・姿勢 M_i を用いて、第*i*フレームの画像上へのランドマークの予測投影位置 $\tilde{\mathbf{c}}_i = [u, v]^T$ を以下の式により算出する。

$$\begin{bmatrix} au \\ av \\ a \end{bmatrix} = R_{i-1} \{ \mathbf{l} + (\omega \times \mathbf{l}) \Delta t \} + (\mathbf{p}_{i-1} + \mathbf{v}_i \Delta t)$$

ただし、 \mathbf{v}_i は第*i*フレームでのカメラの速度であり、第(*i*-1)フレームでの速度 $\mathbf{v}_{i-1} = \mathbf{p}_{i-1} - \mathbf{p}_{i-2}$ により近似する。また、 Δt はフレーム間の時間、 a は媒介変数である。ここでは、予測投影位置 $\tilde{\mathbf{c}}_i$ がカメラの画角内に存在するランドマークのみを選択し、推定に利用するランドマーク群とする。

Step2: モーションブラーを考慮したランドマークの探索

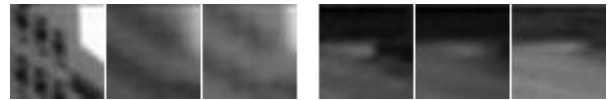
Step1で選択されたランドマークに対して、モーションブラーを再現したLDBテンプレートを作成し、 $\tilde{\mathbf{c}}_i$ を中心とする一定の範囲のウィンドウ内の各画素を注目画素とするテンプレートに対して相関値が最大のものをランドマークと対応付ける。本研究ではフレーム間における画像上での特徴点の動きをLDBテンプレートごとに直線で近似し、LDBテンプレートに対する画像上でのモーションブラーの大きさ、方向を以下のベクトル \mathbf{b} で表す。

$$\mathbf{b} = (\mathbf{c}_{i-1} - \tilde{\mathbf{c}}_i) \frac{s}{f}$$

ただし、 \mathbf{c}_{i-1} は M_{i-1} と \mathbf{l} から算出したランドマークの前フレームでの投影座標、 f は入力画像のフレームレート($1/\Delta t$)を表す。

3. 予備実験

モーションブラーが生じた入力画像に対し、カメラ位置の推定精度を検証した。本実験では、ビデオカメラ (SONY DSR-PD150) を用いて撮影した動画像 ($f=15\text{fps}$, $s=1/30$ 秒) とジャイロセンサ (TOKIMEC TISS-5-40, 方位角誤差: ± 2 度 (ドリフト誤差: ± 3 度/時), ロール角・ピッチ角誤差: ± 0.5 度) の計測値を入力とし、連続した8フレーム (フレーム間の角速度は平均22度/秒, 最大39度/秒) に対してカメラの位置・姿勢を推定し、また、手動でランドマークの対応付けを行って得た推定結果を真値とする定量評価を行った。ただし、入力画像は姿勢成分のみを変化させて撮影した。図1に2つの異なるランドマークに対して、従来手法[1]で推定に用いたテンプレート、提案手法で用いたモーションブラーを再現したテンプレート、入力画像から作成したテンプレートを示す。提案手法のテンプレートは比較的に入力画像と類似しており、モーションブラーが正しく再現されている。本実験において、モーションブラーを考慮しない従来手法ではカメラ位置推定における誤差は平均573mm 最大1679mmであったが提案手法では、平均237mm、最大359mmとなり、特に最大誤差が大幅に減少した。このことから、モーションブラーを考慮することで、カメラ位置の推定精度が向上することを確認した。



従来手法 提案手法 入力画像 従来手法 提案手法 入力画像
ランドマーク A ランドマーク B

図1: LDBテンプレートと
入力画像から作成したテンプレートの比較

参考文献

- [1] 大江 統子, 佐藤 智和, 横矢 直和: “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, 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.