

# 自然特徴点ランドマークデータベースと姿勢センサに基づく 高速な回転にロバストなカメラ位置・姿勢推定

松田 幸大<sup>†</sup> 池田 聖<sup>†</sup> 佐藤 智和<sup>†</sup> 横矢 直和<sup>†</sup>

<sup>†</sup> 奈良先端科学技術大学院大学 情報科学研究科 〒 630-0192 奈良県生駒市高山町 8916-5

E-mail: †{kodai-m,sei-i,tomoka-s,yokoya}@is.naist.jp

あらまし カメラの位置及び姿勢を推定する技術は、拡張現実感、ロボットナビゲーションなど様々な分野で応用が可能である。従来提案されている、環境の特徴点の三次元位置や画像テンプレートが登録されたランドマークデータベースを用いたカメラ位置・姿勢推定法は、物理的なインフラを用いず、カメラの絶対的な位置・姿勢が得られる特長がある。しかし、カメラが高速な回転を伴う場合、モーションブラーや対応点の探索範囲が拡大することによって誤対応が生じ、推定誤差が増大する問題がある。本稿では、この問題を解決するために、ランドマークデータベースと姿勢センサから得られる角速度情報を併用したカメラ位置・姿勢推定法を提案する。提案手法では、まず角速度情報を用いてデータベース中の多数のランドマークから推定に用いるランドマークを効果的に選択する。次に、データベース中の画像テンプレートに対してモーションブラーを再現することで、モーションブラーによる類似度の低下を防ぎ、誤対応を低減する。

キーワード カメラ位置・姿勢推定, ランドマークデータベース, 姿勢センサ, 拡張現実感

## Robust Estimation of Camera Position and Posture with High-speed Rotation Based on Feature Landmark Database and Inertial Sensor

Kodai MATSUDA<sup>†</sup>, Sei IKEDA<sup>†</sup>, Tomokazu SATO<sup>†</sup>, and Naokazu YOKOYA<sup>†</sup>

<sup>†</sup> Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara, 630-0192 Japan

E-mail: †{kodai-m,sei-i,tomoka-s,yokoya}@is.naist.jp

**Abstract** This paper describes a novel camera position and posture estimation method using an inertial sensor and feature landmark database which basically consists of three-dimensional positions of natural feature points (landmarks) and their image templates. This method is widely applicable to a number of fields such as robot navigation and augmented reality system for wearable computer users. In such fields, it is frequently required to estimate absolute position and posture of cameras for environment without any physical infrastructures. A conventional method using feature landmark database has already been satisfying this requirement. However, the estimation error is increased due to miss tracking of image features caused by high-speed rotation of camera. To solve this problem, in the proposed method, the landmarks are first selected from the database by using angular velocity acquired from the inertial sensor. The motion blur is then reproduced on image templates of landmarks, and blurred templates are used for matching with those in the input images.

**Key words** camera position and posture estimation, landmark database, inertial sensor, augmented reality

### 1. はじめに

カメラの位置及び姿勢を実時間で推定する技術は、拡張現実

感 [1], [2] やロボットナビゲーション [3] など、様々な分野への応用が可能である。特に、屋外などの広範囲な環境を対象としたシステムでは、カメラの絶対的な位置及び姿勢の取得が必

要である。従来から、赤外線マーカなどのセンサ [4], [5] や画像マーカ [6], [7] などの物理的なインフラを用いた手法、予め作成された環境のモデルを用いた動画像処理による手法 [8] ~ [11] など、多くの手法が提案されてきた。物理的なインフラを用いた手法 [4] ~ [7] では、センサからの信号やマーカのパターンから個々のインフラを識別することが容易であるため、時系列的な情報なしでカメラの位置及び姿勢の推定が可能であり、高速な動きに対してロバストな推定が実現できる。しかし、これらの手法では、環境中に物理的なインフラを整備する必要があり、自然環境などインフラの設置が困難な環境においては利用できない。一方、環境のモデルを用いた手法では、センサやマーカなどのインフラを用いずに、カメラの位置及び姿勢の推定が可能である。環境のモデルとして、ワイヤースケルトモデル [8] や自然特徴点群 [9], [10], テクスチャ付きのポリゴンモデルを用いた手法 [11] が提案されており、これらの手法では、エッジや自然特徴点などの入力画像中の特徴とモデルとを対応付けることによりカメラの絶対的な位置及び姿勢が推定される。一般的に、こういった環境モデルの生成には多大な労力を要するが、自然特徴点を用いた手法 [9], [10] では、動画像から structure-from-motion の技術を用いて環境の三次元モデルが自動生成されるため、広範囲な景観を容易にモデル化できるという特長を持つ。しかし、これらの環境の三次元モデルを用いた手法 [8] ~ [11] では、画像中の個々の特徴を一意に識別することが難しく、カメラの回転及び並進が大きい場合、モーションブラーや特徴の探索範囲の拡大によりモデルと入力画像において誤対応が生じ、カメラ位置及び姿勢の推定誤差が増大する問題がある。

この問題を解決するために、環境のモデルとセンサから得られる情報を併用してカメラの位置及び姿勢を推定するハイブリッドな手法 [12] ~ [14] が研究されている。探索範囲外にモデルとの対応点が存在することによる誤対応を防ぐ手法として、Reitmayr ら [12] や Jiang [13] らは、姿勢センサによる計測結果と動画像処理による推定を拡張カルマンフィルタにより統合することで、カメラの回転に対してロバストに位置及び姿勢の推定を行う手法を提案している。Klein ら [14] は姿勢センサにより得られる角速度情報からブラーを考慮したエッジ探索ウィンドウを作成することで、ワイヤースケルトモデルと入力画像中の物体のエッジがモーションブラーによって誤対応を起こす問題を解決し、モーションブラーが生じた入力画像に対するカメラの位置及び姿勢の推定を可能にしている。しかし、従来のポリゴンモデル [12] やワイヤースケルトモデル [13], [14] の利用を前提とした手法では、対象とする環境の規模が大きく複雑な場合、モデルの作成自体が難しいという問題がある。

本稿では、自然特徴点群の三次元位置と画像テンプレートをデータベース化したランドマークデータベース [9] 及び姿勢センサにより得られる角速度情報を併用することで、カメラの高速な回転が生じた場合においても位置及び姿勢をロバストに推定する手法を提案する。提案手法では、まず、現フレームにお

けるカメラの暫定的な位置及び姿勢を求め、フレーム毎にデータベース中の多数のランドマークから推定に用いるランドマークを適切に限定する。次に、限定されたランドマークの画像テンプレートに対して入力画像と同様のモーションブラーを再現し、モーションブラーによる入力画像とランドマーク間の類似度の低下を防ぐ。さらに、カメラの並進を考慮した大きさの検出ウィンドウによって検出された入力画像内の特徴点とランドマークをテンプレートマッチングによって対応付け、カメラの位置及び姿勢を推定する。実験では、並進移動するカメラに高速な回転を与えて取得した動画像からカメラの位置及び姿勢を推定し、特徴点とランドマークの対応や推定されたカメラの位置及び姿勢の推定精度を検証することにより提案手法の有効性を示す。ただし、本稿では、カメラの内部パラメータは既知であり、カメラと姿勢センサの同期及び位置と姿勢の関係は与えられているものとする。

以下、2 節では、ランドマークデータベースと姿勢センサを用いたカメラ位置及び姿勢推定について述べ、3 節では、提案手法の精度評価実験について報告する。最後に、4 節でまとめと今後の課題について述べる。

## 2. ランドマークデータベースと姿勢センサを用いたカメラ位置及び姿勢推定

提案手法では、予め構築したランドマークデータベースを用いてカメラ位置及び姿勢を推定する。以下では、まずランドマークデータベースの構築について述べる。次に、オンライン逐次処理の枠組みである、ランドマークデータベースと姿勢センサによるカメラの位置及び姿勢の推定手法について述べる。

### 2.1 ランドマークデータベースの構築

ランドマークデータベースは、多数のランドマークから成り、各ランドマークの情報として、ランドマークの 3 次元位置と外観情報を保持している。様々な位置からのランドマークの見え方の変化に対応するために、外観情報は各ランドマークの撮影地点毎に、以下の情報を含む [9]。

- ・ ランドマークを撮影したカメラ位置
- ・ 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
- ・ 画像テンプレートのスケール

以上の情報は、事前に環境の撮影及び基準となる自然特徴点の三次元位置の計測を行い、環境内の自然特徴点を三次元復元することで取得する [15]。

### 2.2 カメラ位置及び姿勢の推定手法

上記のランドマークデータベース、姿勢センサからの角速度情報及び入力画像から、カメラの絶対的な位置及び姿勢を推定する手法について述べる。本手法の処理の流れを図 1 に示す。各フレームにおけるステップは以下の通りである。まず、姿勢センサで得られた角速度情報から暫定的なカメラの位置及び姿勢を算出し (A)、データベースから推定に利用するランドマークを限定する (B)。限定されたランドマークの画像テンプレ

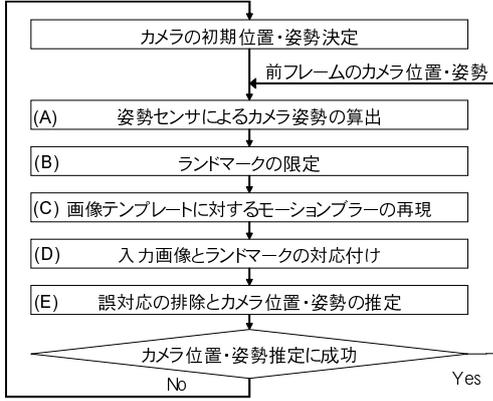


図 1 全体の処理の流れ

トに対して入力画像と同様のモーションブラーを再現する (C) . 生成したテンプレートと入力画像のテンプレートマッチングにより, 入力画像中の自然特徴点とランドマークとを対応付け (D), 最後に, それらの対応関係を用いて PnP 問題 [16] を解くことによりカメラの位置及び姿勢を推定する (E) . ただし, 本稿では, 初期フレームのカメラの位置及び姿勢は何らかの手法によって取得出来ることを前提としており, カメラ位置及び姿勢の推定が失敗した場合には, 再度, カメラの初期位置・姿勢の推定から各ステップの処理を繰り返す. このカメラ位置・姿勢推定の成功・失敗判定については, ステップ (E) の説明で述べる. 以下, ステップ (A) ~ (E) について詳述する.

#### (A) 姿勢センサによるカメラ姿勢の算出

姿勢センサから得られる姿勢情報のみを用いると, ドリフトによる累積的な誤差の影響を受けるため, 本ステップでは, ドリフト誤差を含まないカメラの姿勢を計測するために, 第  $(i-1)$  フレームでランドマークを用いて推定された絶対的なカメラの姿勢  $\mathbf{R}_{i-1}^c$  と微小変化量である角速度  $\omega_i$ , および予めキャリブレーションにより得たカメラ座標系からセンサ座標系への変換  $\mathbf{R}^{cg}$  のみからカメラの姿勢  $\tilde{\mathbf{R}}_i^c$  を算出する. 以下では, カメラの姿勢  $\tilde{\mathbf{R}}_i^c$  の算出方法について述べる.

図 2 に示すように, 第  $i$  フレームにおけるカメラの姿勢  $\mathbf{R}_i^c$  は, 姿勢センサの姿勢  $\mathbf{R}_i^g$  とカメラ座標系から姿勢センサ座標系への変換  $\mathbf{R}^{cg}$  から  $\mathbf{R}_i^c = (\mathbf{R}^{cg})^{-1}\mathbf{R}_i^g$  で表される. また, 第  $(i-1)$  フレームから第  $i$  フレームへの姿勢センサの回転を  $\mathbf{R}_{i-1,i}^g$  としてあらわすと,  $\mathbf{R}_i^g = \mathbf{R}_{i-1,i}^g\mathbf{R}_{i-1}^g$  が成立する. 以上より, 第  $i$  フレームにおけるカメラの暫定的な姿勢  $\tilde{\mathbf{R}}_i^c$  は,  $\Omega_i = (\mathbf{R}_{i-1}^g)^{-1}\mathbf{R}_{i-1,i}^g$  を用いて以下の式で表される.

$$\tilde{\mathbf{R}}_i^c = \mathbf{R}_{i-1}^c\Omega_i\mathbf{R}^{cg}\mathbf{R}_{i-1}^c \quad (1)$$

$\Omega_i$  は第  $(i-1)$  フレームから第  $i$  フレームにおける微小な回転であり, 第  $i$  フレームにおける角速度  $\omega_i = (\omega_{ix}, \omega_{iy}, \omega_{iz})$  の要素を用いて, 以下の様に  $3 \times 3$  の歪対象行列で表現する.

$$\Omega_i = \begin{pmatrix} 0 & -\omega_{iz} & \omega_{iy} \\ \omega_{iz} & 0 & -\omega_{ix} \\ -\omega_{iy} & \omega_{ix} & 0 \end{pmatrix} \quad (2)$$

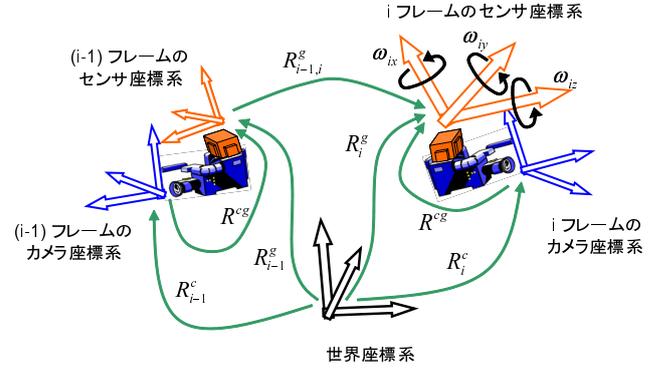


図 2 座標系の定義

本研究では, 姿勢センサとカメラ間の変換行列  $\mathbf{R}^{cg}$  を既知とし, カメラの姿勢  $\mathbf{R}_{i-1}^c$  及び行列  $\Omega$  は累積誤差のない回転を表すので, 式 (2) を用いることで前フレームのカメラ姿勢と現フレームにおける角速度情報から絶対的なカメラの姿勢を表す  $\tilde{\mathbf{R}}_i^c$  を算出できる.

#### (B) ランドマークの限定

ステップ (A) で算出した暫定的なカメラの姿勢  $\tilde{\mathbf{R}}_i^c$  を用いて, ランドマークデータベースに登録されている多数のランドマークから, 入力画像との対応付けに用いるランドマークを限定する処理について述べる. ここでは, 第  $i$  フレームに対して予測されるカメラの位置  $\tilde{\mathbf{P}}_i^c$  及び姿勢  $\tilde{\mathbf{R}}_i^c$  から, 画角内に入るランドマークを選択し, 入力画像と対応付ける候補とする. 具体的には, 第  $(i-1)$  フレームでのカメラの位置を  $\mathbf{P}_{i-1}^c$ , ランドマーク  $j$  の三次元位置を  $\mathbf{l}_{(j)}$  とすると, 第  $i$  フレームにおける画像上へのランドマーク  $j$  の予測投影位置  $\tilde{\mathbf{c}}_{(j)i} = [x_{(j)i}, y_{(j)i}]^t$  は, ランドマークの奥行き  $a_{(j)i}$  を用いて, 以下の式で表される.

$$\begin{pmatrix} a_{(j)i}x_{(j)i} \\ a_{(j)i}y_{(j)i} \\ a_{(j)i} \end{pmatrix} = \tilde{\mathbf{R}}_i^c\mathbf{l}_{(j)} + \tilde{\mathbf{P}}_i^c \quad (3)$$

ただし, 第  $i$  フレームにおけるカメラの予測位置  $\tilde{\mathbf{P}}_i^c$  は次式により算出する.

$$\tilde{\mathbf{P}}_i^c = 2\mathbf{P}_{i-1}^c - \mathbf{P}_{i-2}^c \quad (4)$$

また,  $a_{(j)i}$  は第  $i$  フレームにおけるカメラ座標系でのランドマーク  $j$  の奥行きを表す. 本ステップでは, 上式で得られる予測投影位置  $\tilde{\mathbf{c}}_{(j)i}$  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とする.

#### (C) 画像テンプレートに対するモーションブラーの再現

ステップ (B) により選択されたランドマークの画像テンプレートから, 入力画像のモーションブラーを再現した画像テンプレートを生成する処理について述べる. 本研究では, 等速直線運動を仮定した PSF モデルを用いて, 各ランドマークに対してデータベースの画像テンプレートにモーションブラーを再現する. この処理により, 図 3 に示すように, 入力画像とランドマーク間の画像テンプレートの類似度を高めることが出来る.

対応付けに用いた 画像テンプレート	ブラーの 再現なし	ブラーの 再現あり	入力画像
正規化相互相関値	0.76	0.95	

図3 テンプレート間の正規化相互相関値に対するモーションブラー再現の効果の確認

モーションブラーの大きさ, 方向は, 以下のベクトル  $\mathbf{b}$  で表される.

$$\mathbf{b} = (\mathbf{c}_{(j)i-1} - \tilde{\mathbf{c}}_{(j)i})s\Delta t \quad (5)$$

ただし,  $\mathbf{c}_{(j)i-1}$  は第  $(i-1)$  フレームにおけるカメラ位置  $\mathbf{P}_{i-1}^c$  及び姿勢  $\mathbf{R}_{i-1}^c$  と三次元位置  $\mathbf{l}_{(j)}$  から算出したランドマークの第  $(i-1)$  フレームにおける画像への投影座標,  $\Delta t$  はフレーム間の時間を表し, シャッタースピード  $s$  は既知とする.

#### (D) 入力画像とランドマークの対応付け

ステップ(B)で限定された各ランドマークに対して入力画像中の自然特徴点を対応付ける処理について述べる. ここでは, ステップ(B)で算出されたランドマーク  $j$  の予測投影位置  $\tilde{\mathbf{c}}_{(j)i}$  を中心とする大きさ  $\delta$  のウィンドウ内において, ステップ(C)で生成されたテンプレートを用いて, 正規化相互相関を指標としたテンプレートマッチングを行う. ウィンドウの大きさ  $\delta$  は, カメラの並進移動と姿勢センサの計測精度を考慮して設定する. 具体的には, ランドマーク  $j$  のカメラからの奥行き  $a_{(j)i}$ , アプリケーションに合わせて設定されたカメラの最大並進速度  $V_{max}$ , 姿勢センサの精度から設定される  $\delta'$  を用い, 検出ウィンドウの大きさ  $\delta_{(j)i}$  を以下の式で算出する.

$$\delta_{(j)i} = \frac{V_{max}\Delta t}{a_{(j)i}} + \delta' \quad (6)$$

ただし, 多重スケール画像テンプレート間の正規化相互相関は, 各テンプレート内の全ての画素の輝度値を並べたベクトルを  $\mathbf{p}$ ,  $\mathbf{q}$  とし, その平均値を並べたベクトルを  $\bar{\mathbf{p}}$ ,  $\bar{\mathbf{q}}$  とした時, 以下の式により算出される.

$$C(\mathbf{p}, \mathbf{q}) = \frac{(\mathbf{p} - \bar{\mathbf{p}}) \cdot (\mathbf{q} - \bar{\mathbf{q}})}{\|\mathbf{p} - \bar{\mathbf{p}}\| \|\mathbf{q} - \bar{\mathbf{q}}\|} \quad (7)$$

この相関値が最大かつ閾値以上である対応候補点をランドマークと対応付ける.

#### (E) 誤対応の排除とカメラ位置及び姿勢推定

本ステップでは, RANSAC[17]を用いたロバスト推定によってこれまでのステップで求めたランドマークと自然特徴点の対応関係から誤対応を除去し, 正しい対応関係のみから再投影誤差を最小化することで, カメラの位置及び姿勢を推定する. 再投影誤差は, ランドマーク  $j$  の画像上への投影座標  $\tilde{\mathbf{c}}_{(j)}$  と, ランドマークに対応づけられた自然特徴点の画像上での検出座標  $\mathbf{c}_{(j)}$  の自乗距離の和  $E = \sum_j \|\tilde{\mathbf{c}}_{(j)} - \mathbf{c}_{(j)}\|^2$  とし, これを最小化することで最適なカメラの位置・姿勢を求める. 再投影誤

差の最小化は非線形最小化問題であり, 局所解の問題や計算量の問題を伴うため, まず線形解法を用いてまずカメラ位置・姿勢の初期値を求め[16], 次に, 勾配法を用いた非線形最小化によって再投影誤差  $E$  を最小化する. ここで, 最終的に利用可能な対応点数が6点を下回るか再投影誤差が十分に最小化されない場合, 推定が失敗したとみなす.

### 3. カメラ位置・姿勢推定の精度評価実験

提案手法の有効性を検証するために, 実動画像を用いた精度評価実験を行った. 以下, まず実験環境について述べ, 次に, 提案手法の定量的な評価実験について述べる.

#### 3.1 実験環境

本実験で入力として用いた動画像, 角速度情報, ランドマークデータベースのデータの取得方法について述べる. 動画像は手持ちのビデオカメラ (SONY DSR-PD150, 解像度  $720 \times 480$  pixel, 15fps), 姿勢情報はカメラに固定された姿勢センサ (TOKIMEC TISS-5-40, 方位角誤差:  $\pm 2$  度, ドリフト誤差:  $\pm 3$  度/時, ロール角・ピッチ角誤差:  $\pm 0.5$  度) で得た. ランドマークデータベースの生成には, 放射状に固定された複数のカメラにより全天球の動画像を取得できる全方位型マルチカメラシステム (Point Grey Research 社 Ladybug, 解像度  $768 \times 1024 \times 6$  台) を用いた. このカメラシステムにより屋外環境を移動撮影し, 佐藤らの手法[15]によりカメラパスと自然特徴点の三次元位置を推定することでランドマークデータベースを構築した. データベース中のランドマーク数は約7500個であり, 1つのランドマークあたりの撮影地点数は平均8ヶ所であった. また, 定量的な評価に用いるカメラ位置及び姿勢の真値は, 環境内の特徴点の三次元位置をトータルステーションを用いて測定し, 各フレームの画像上でそれらの特徴点を手動で対応付け, PnP問題を解くことで算出した.

#### 3.2 定量的な評価実験

上記の実験環境において, 以下に述べる2つの比較手法と提案手法で推定されたカメラの位置及び姿勢を真値と比較することで推定精度を調べた. ただし, モーションブラーが生じた画像から直接カメラの位置・姿勢の真値を求め, 推定誤差を計算することは困難なので, 推定に用いた入力動画像は, 以下の2種類の方法で撮影したものをを用いた.

入力1 回転のみを伴うカメラによる撮影

入力2 回転と並進を伴うカメラによる撮影

回転のみを伴うカメラによって得られた動画像 (入力1) は, モーションブラーが生じていない初期フレームから算出したカメラ位置を真値とすることで, 動画像中の全フレームに対して推定位置誤差を算出した. 一方, 回転と並進を伴うカメラによって得られた動画像 (入力2) は, 推定誤差を算出する各フレームにおいてカメラの位置・姿勢の真値を算出し, 推定位置誤差・推定姿勢誤差を算出した. ただし, 真値計算時の再投影誤差が1.5pixel以上となったフレームは, 信頼性が低いと判断し, 評価対象から除外した. 入力1, 2に対応する動画像の一部の

フレームをそれぞれ図 4, 図 5 に示す。これらは、入力画像撮影時において自動設定されたシャッタースピード 1/15 秒で撮影した。また、本実験において、カメラの最大並進速度  $V_{max}$  は、一般的な人間のジョギングの速度である  $8km/h$  とし、姿勢センサの精度から算出した検出ウィンドウの大きさ  $\delta'$  を 16 画素に設定した。

本実験では、上記の入力画像を用い、提案手法と 2 つの比較手法とを比較する。2 つの比較手法は、本研究で新たに導入したステップ (B), (C) の有効性を確認するために、提案手法中のステップ (B), (C) のうち、1 つのステップを以下に示す処理に置き換えたものとする。

比較手法 (1) ステップ (B) において、前フレームで推定したカメラの位置及び姿勢からランドマークを限定する

比較手法 (2) ステップ (C) において、画像テンプレートに対してモーションブラーの再現を行わない

### 3.2.1 回転のみを伴うカメラで撮影された動画像 (入力 1)

各フレームにおける角速度と最終的なカメラ位置・姿勢推定に用いられたランドマークの数を図 6 に、各手法の平均推定誤差と標準偏差を表 1 に示す。比較手法 (1) では、277 フレームにおいてカメラ位置姿勢推定が途中で失敗した。図 6 から、比較手法 (1) の推定処理の失敗は、大きな角速度が発生した直後に発生していることが分かる。比較手法 (2) では、大きな角速度が発生した 70 フレームや 270 フレーム付近における対応付けられたランドマーク数の減少が見られる。一方、提案手法にも対応付けられたランドマーク数の変化が見られるが比較的安定しており、全てのフレームで推定が成功した。このことから、姿勢センサを用いたステップ (B), (C) の導入により、高速な回転に対するロバスト性が向上していることが分かる。また、各比較手法に比べ、提案手法は各フレームにおける誤差のばらつきが少なく、大きな誤差を生じることなくカメラの位置を推定できていることが分かる。統計的に推定誤差の検証を行うため、有意水準を 5% として提案手法に対して t 検定を行った結果、比較手法 (1) との t 値が -7.64、比較手法 (2) との t 値が -7.93 であり、提案手法と比較手法の推定誤差に有意な差があることを確認した。

### 3.2.2 回転と並進を伴うカメラで撮影された動画像 (入力 2)

各フレームにおける角速度と最終的なカメラ位置・姿勢推定に用いられたランドマーク数を図 7 に、各手法の平均推定誤差を表 2 に示す。比較手法 (1) 及び (2) では、それぞれ 44 フレーム、84 フレームにおいて、カメラの位置及び姿勢の推定が途中で失敗した。これに対して、提案手法では、最終フレームまでカメラ位置姿勢を推定することが出来た。図 7 から、比較手法 (1) では、前述で述べた実験と同様に、大きな角速度が発生した直後に推定処理の失敗が発生し、表 2 から位置及び姿勢の推定誤差も増大していることが分かる。比較手法 (2) では、表 2 から提案手法と同程度の精度が得られているが、比較手法 (1) と同様に、大きな角速度が発生した直後に推定処理の失敗が見

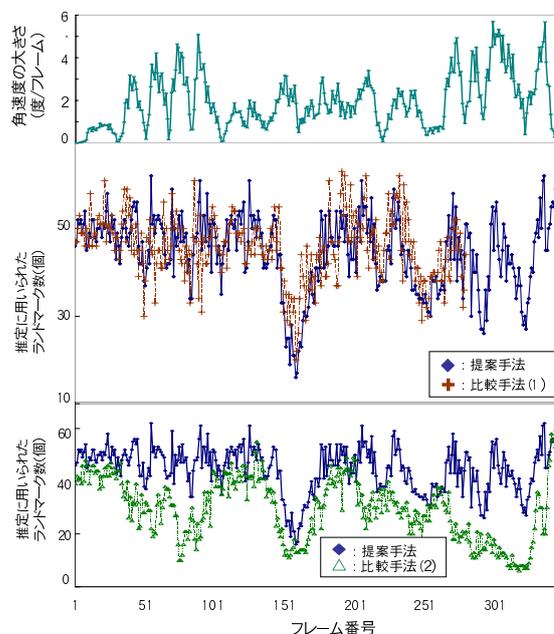


図 6 角速度の大きさと対応付けられたランドマーク数 (入力 1)

表 1 平均推定誤差と標準偏差 (入力 1)

	平均位置誤差 (mm)	標準偏差 (mm)
提案手法	394	309
比較手法 (1)	706	607
比較手法 (2)	1239	1960

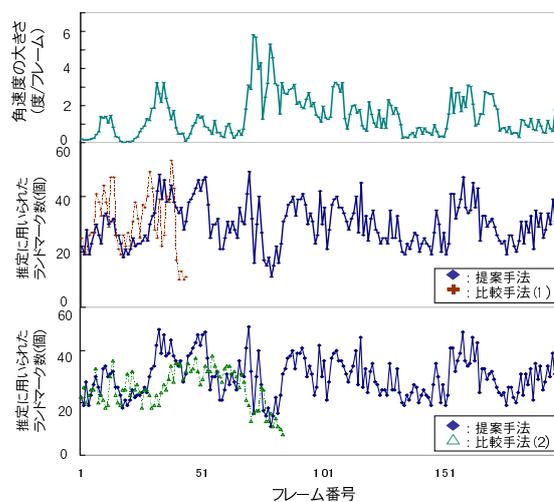


図 7 角速度の大きさと対応付けられたランドマーク数 (入力 2)

られる。今回の実験では、モーションブラーが生じたフレームでは推定誤差が検証できないため、表 2 はブラーの発生していないフレームにおける平均推定誤差を表しているが、比較手法 (1), (2) における推定の失敗は、いずれもモーションブラーが生じたフレームにおいて発生しており、推定に用いられるランドマーク数が減少しているため、これらのフレーム付近において比較手法の推定誤差が増大している可能性がある。

## 4. まとめと今後の課題

自然特徴点ランドマークデータベースと姿勢センサによる角

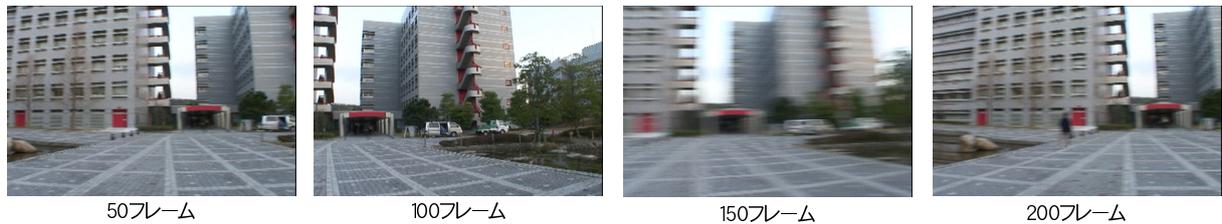


図 4 入力画像 (入力 1)

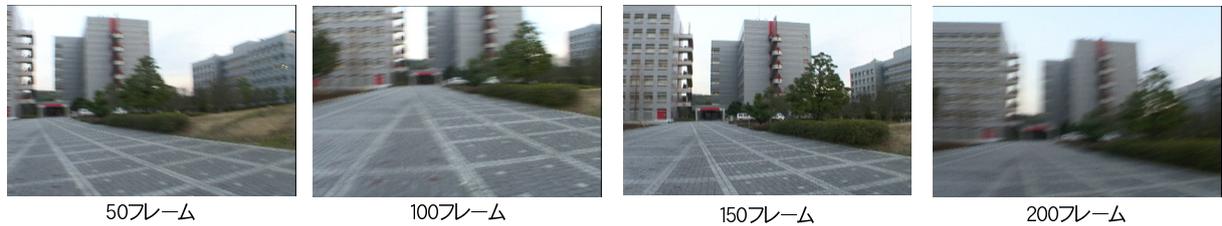


図 5 入力画像 (入力 2)

表 2 平均推定誤差 (入力 2)

	平均位置誤差 (mm)	平均姿勢誤差 (度)
提案手法	459	0.36
比較手法 (1)	1406	1.42
比較手法 (2)	571	0.39

速度情報、動画像を入力とし、データベースと入力画像中の自然特徴点を対応付けることにより、カメラの位置・姿勢を推定する手法を提案した。提案手法では、姿勢センサによる角速度情報を推定に用いるランドマークの選択やモーションブラーを考慮したテンプレートマッチングに利用することにより、人工的なマーカを用いることなく、高速な回転を伴うカメラの絶対的な位置及び姿勢をロバストに推定することができる。今後の課題として、カメラ位置・姿勢推定のリアルタイム化、動画像と姿勢センサによる姿勢情報の同期の自動化が挙げられる。

#### 文 献

- [1] S. Feiner, B. MacIntyre, T. Höllerer and A. Webster: "A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment," Proc. 1st IEEE Int. Symp. on Wearable Computers, pp. 74-81, 1997.
- [2] P. Daehne and J. Karigiannis: "Archeoguide: System architecture of a mobile outdoor augmented reality system," Proc. 1st IEEE Int. Symp. on Mixed and Augmented Reality, pp. 263-264, 2002.
- [3] 松本, 稲葉, 井上: "ビューベースアプローチに基づく移動ロボットナビゲーション", 日本ロボット学会誌, Vol.20, No.5, pp. 506-514, 2002.
- [4] A. Harter, A. Hopper, P. Steggle, A. Ward and P. Webster: "The anatomy of a context-aware application," Proc. 5th IEEE Int. Conf. on Mobile Computing and Networking, pp. 59-68, 1999.
- [5] InterSense Inc.: "Is-900," <http://www.isense.com/products/prec/is900/>.
- [6] 加藤, M. Billinghamurst, 浅野, 橋: "マーカー追跡に基づく拡張現実感システムとそのキャリブレーション", 日本バーチャルリアリティ学会論文誌, Vol.4, No.4, pp. 607-616, 1999.
- [7] L. Naimark and E. Foxlin: "Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker," Proc. 1st IEEE Int. Symp. on Mixed and Augmented Reality, pp. 27-36, 2002.
- [8] E. Rosten and T. Drummond: "Fusing points and lines for high performance real-time tracking," Proc. 10th IEEE Int. Conf. on Computer Vision, pp. 1508-1515, 2005.
- [9] 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, 2005.
- [10] I. Gordon and D. G. Lowe: "Scene modeling, recognition and tracking with invariant image features," Proc. 3rd IEEE Int. Symp. on Mixed and Augmented Reality, pp. 110-119, 2004.
- [11] V. Lepetit, L. Vacchetti, D. Thalmann and P. Fua: "Fully automated and stable registration for augmented reality applications," Proc. 2nd IEEE Int. Symp. on Mixed and Augmented Reality, pp. 93-102, 2003.
- [12] G. Reitmayr and T. W. Drummond: "Going out: Robust model-based tracking for outdoor augmented reality," Proc. 5th IEEE Int. Symp. on Mixed and Augmented Reality, pp. 109-118, 2006.
- [13] B. Jiang, U. Neumann and S. You: "A robust hybrid tracking system for outdoor augmented reality," Proc. IEEE Conf. on Virtual Reality, pp. 3-10, 2004.
- [14] G. S. W. Klein and T. W. Drummond: "Tightly integrated sensor fusion for robust visual tracking," Image and Vision Computing, Vol.22, pp. 769-776, 2004.
- [15] T. Sato, S. Ikeda and N. Yokoya: "Extrinsic camera parameter recovery from multiple image sequences captured by an omni-directional multi-camera system," Proc. European Conf. on Computer Vision, Vol.2, pp. 326-340, 2004.
- [16] 出口: "射影幾何学による PnP カメラ補正問題の統一的解法", 情処シンポジウム, Vol.90, pp. 41-50, 1990.
- [17] M. A. Fischler and R. C. Bolles: "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Communications of the ACM, Vol.24, pp. 381-395, 1981.