

ランドマークデータベースに基づく カメラトラッキング法の高速化と安定化

樋下 航[†] 一刈 良介[†] 天目 隆平[‡] 柴田 史久[†] 田村 秀行[†]

[†]立命館大学大学院理工学研究科 [‡]立命館大学総合理工学研究機構

〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: toishita@rm.is.ritsumei.ac.jp

あらまし 本稿では、屋内外で利用可能な特徴点追跡に基づくカメラトラッキング手法を提案する。従来手法では、絶対的なカメラ位置姿勢を取得するために何点かの特徴点の3次元位置を手動で計測する必要があり、準備に多大な時間と労力を要する点が問題となっていた。本研究では、人為的マーカと背景中の特徴点を併用することで、自動的に特徴点の3次元位置を推定し、準備に要する時間を大幅に削減する。さらに、従来手法のトラッキング処理においては、特徴点追跡に多くの計算コストを要し、また一度破綻すると復帰できないという問題があった。そこで、画像中から得られる局所不変特徴量を特徴点追跡に用いることで速度と安定性を向上させる。

キーワード カメラトラッキング, ランドマークデータベース, 局所不変特徴量

Performance Improvement of Camera Tracking Method Using the Feature Landmark Database

Wataru TOISHITA[†], Ryosuke ICHIKARI[†], Ryuhei TENMOKU[‡]
Fumihisa SHIBATA[†], and Hideyuki TAMURA[†]

[†] Graduate School of Science and Engineering, Ritsumeikan University

[‡] Graduate Organization of Science and Engineering, Ritsumeikan University

1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577 Japan

E-mail: toishita@rm.is.ritsumei.ac.jp

Abstract This paper describes a camera tracking method which can be used in outdoor and indoor environments using natural feature tracking. In conventional methods, much time and human cost are needed to construct landmark database because they require measuring some absolute positions of feature beforehand. Furthermore, tracking methods have considerable problems that they require much computational cost and cannot recover from tracking failures. We propose an automatic method for constructing landmark database to reduce the time for preparation. The key idea of our method is combining use of natural features in the background and a fiducial marker. We adopt the local invariant feature for matching to achieve the improvement of speed and robustness of camera tracking process.

Keyword Camera Tracking, Feature Landmark Database, Local Invariant Feature

1. はじめに

我々の研究グループは、複合現実感技術を用いて映画撮影を支援するMR-PreViz[1]プロジェクトを推進している。MR-PreVizシステムを利用すると、例えば実際の撮影現場を背景に俳優のCGキャラクタを重畳描画するなど、本番撮影前にカメラワークを効果的に検討することが可能である。

MR-PreViz映像の撮影時には、実背景と仮想物体の幾何学的整合性を保つために実時間で高精度にカメラ位置姿勢を推定する必要がある。この位置合わせ問題は複合現実感における最重要課題の1つであり、これまで数多くの手法が提案されてきた。

近年の研究の主流は、物理的なセンサを用いる手法よりも、画像情報のみからカメラトラッキングを行うビジョンベースの手法である。中でも、背景中の特徴点を利用する手法は屋内外を問わず高精度な推定が可能であることが知られている。特徴点を利用する手法は事前知識を必要とする手法[2][3]と必要としない手法[4]に分類される。前者はトラッキング時に撮影対象となる背景の特徴点情報が既知であるため、安定したトラッキングが可能である。しかし、その事前知識の獲得に手間や時間がかかるという問題がある。一方、後者では、カメラトラッキングと特徴点の3次元位置推定を同時並行で行うアプローチがとられる。事前準備

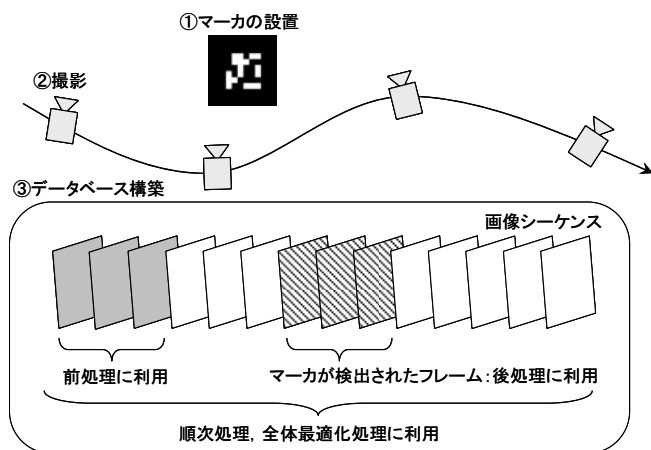


図 1: ランドマークデータベース構築手順

備がいないため手軽にトラッキングを行えるが、現実世界に対する絶対的なカメラ位置姿勢やスケール情報を取得できないといった問題がある。そのため、撮影現場において特定の場所に CG キャラクタを合成しなければならない MR-PreViz システムには適さない。

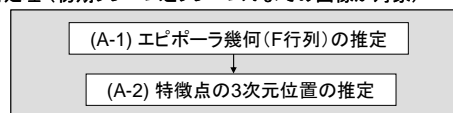
我々が選択した特徴点ランドマークデータベースに基づく手法は、事前に特徴点の 3 次元位置と特徴点周辺の画像情報をランドマークとしてデータベースに登録しておき、ランドマークと入力画像中の特徴点を毎フレーム対応付けることでトラッキングを実現する。屋内外で絶対的なカメラ位置姿勢が取得できるため、MR-PreViz 目的には最適な手法だと考えられる。

武富らの従来手法[2]を MR-PreViz に適用する上で問題となるのは、MR-PreViz では本番撮影前の短い時間で事前知識の学習を終わらせることが要求される点である。武富らの手法では、人手により環境中の複数の特徴点の 3 次元位置を計測し、あらかじめ全方位カメラで撮影した動画像中の同一点と対応づける必要がある。この準備には多大な時間と労力を要するため、MR-PreViz にそのまま適用することは出来ない。

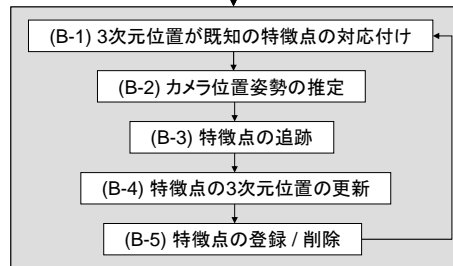
そこで本研究では、映画撮影においては本番撮影時のおおよそのカメラパスが既知であるという特性を利用し、データベース構築時とトラッキング時で同じカメラパスをとることで準備に要する時間を短縮する。広域な環境を対象とする武富らの手法とは異なり、我々の手法では本番撮影時に映る背景中のランドマークのみを登録しておけばよい。そこで、提案手法では単眼カメラで撮影した動画像中の特徴点に対し、追跡処理に基づくステレオ視を用いて 3 次元位置推定を行う。これにより、従来必要であった手動作業が自動化され、撮影現場でランドマークデータベースを高速に構築することが可能となる。

また、従来のトラッキング手法[2]で問題となるのは、トラッキングの処理速度と安定性が撮影現場で利用できる水準まで達していない点である。従来は初期フレームにおけるカメラ位置姿勢推定を広範囲で推定可能な手法で実現しているが、推定に数十秒を要するため、何度もカメラワークを試す際に問題となる。また、

前処理 (初期フレームとフレーム t までの画像が対象)



順次処理 (全フレームが対象)



後処理 (全フレームが対象)

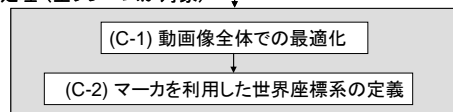


図 2: 特徴点の 3 次元位置推定アルゴリズム

トラッキング処理自体の計算コストも十分低いとは言えず、さらなる高速化が望まれる。安定性に関しては、従来はトラッキングが一度破綻すると復帰できないという問題があった。手ぶれやオクルージョン等によりトラッキングの破綻は頻繁に起きるものであり、安定性の向上には復帰処理が必要不可欠である。

提案手法では、データベース構築時とトラッキング時でカメラパスが同じであることを利用し、初期フレームにおけるカメラ位置の推定範囲をデータベース構築時のカメラパス上に限定することで、処理時間を短縮する。さらに、ランドマークと入力画像中の特徴点の対応付けに局所不変特徴量の一つである SIFT 特徴量[7]を利用する。SIFT 特徴量は数ある特徴量の中で非常に識別性能が高いことが知られている。そこで、ランドマークの 3 次元情報をもとに SIFT 特徴量を高速に記述することで、トラッキング処理時間の短縮とトラッキング破綻からの実時間復帰の両立を図る。

以下、2 章ではデータベース構築手法について述べ、3 章ではトラッキング手法について述べる。4 章では、データベース構築に関しては屋内外で実験を行い、トラッキングに関しては従来手法との計算コストの比較や性能評価を行う。最後に、5 章で本稿をまとめる。

2. ランドマークデータベース構築

ランドマークデータベース構築の手順は次の通りである。まず、1 つの人為的マーカが配置された背景を単眼カメラにて撮影する (図 1 の①と②)。マーカは世界座標系を定義するために利用され、ユーザが座標系の原点としたい場所に配置する。次に、得られた画像シーケンスを用いて特徴点の 3 次元位置を推定する (図 1 の③)。そして、得られた 3 次元位置と特徴点周辺の画像情報をランドマークとしてデータベースに登録する。最後に、環境中からマーカを取り払うことでトラッキング前に必要な準備は完了となる。

2.1. 特徴点の3次元位置推定

提案手法では、従来手法[2]が手動で計測していた特徴点の3次元位置を前処理として自動的に推定する。これにより人手による作業時間が省かれ、データベース構築の大幅な時間短縮が期待される。以下で、前処理、順次処理、後処理の3段階からなる特徴点の3次元位置推定処理について詳述する(図2)。

2.1.1. 前処理

前処理では、順次処理で必要となる何点かの特徴点の3次元位置を取得する。本手法では、一般的なステレオ視に基づく手法により3次元位置推定を行う。初期フレームとある程度ベースラインが確保できるフレームとの間で、カメラの相対的な姿勢 \mathbf{R} , \mathbf{t} と特徴点の3次元位置 $\mathbf{P}_r = [x_r, y_r, z_r]^T$ を求めるためにエピポーラ幾何 (\mathbf{F} 行列) を推定する。

本手法では、8点の対応関係から \mathbf{F} 行列を線形的に解く Eight-Point Algorithm[8]と、誤対応を排除するための PROSAC 法[9]を組み合わせて \mathbf{F} 行列を求める。対応関係は、後述する順次処理と同様の特徴点追跡処理によって与えられる。 \mathbf{F} 行列はカメラの内部パラメータ \mathbf{A} と \mathbf{R} , \mathbf{t} によって式(1)のように記述される。

$$\mathbf{F} = \mathbf{A}^{-T} [\mathbf{t}]_{\times} \mathbf{R} \mathbf{A}^{-1} \quad (1)$$

\mathbf{A} は事前のカメラキャリブレーションによって既知であるため、 \mathbf{F} 行列から \mathbf{R} , \mathbf{t} が求まる。したがって、2フレーム間の姿勢から三角測量の原理を用いて特徴点の3次元位置 \mathbf{P}_r を決定することができる。なお、スケールが未知なため、カメラ間距離 $\|\mathbf{t}\|$ は1とする。

2.1.2. 順次処理

順次処理では、前処理で求めておいた特徴点の3次元位置を用いてカメラ位置姿勢を求め、フレーム毎に Structure-from-Motion 処理を行うことで新たな特徴点の3次元位置を推定していく。具体的には以下の手順で達成される(図2)。

(B-1) 3次元位置が既知の特徴点の対応付け

Harris オペレータ[10]で検出された特徴点に対し、画像テンプレートを用いて3次元位置 \mathbf{P}_r が既知の特徴点をフレーム間で対応付ける。テンプレート間の相関は各画素間の差の2乗和(以下SSD)で計算する。探索窓は、前フレームのカメラ位置姿勢を用いて、特徴点ごとの投影座標を中心に設定される。

(B-2) カメラ位置姿勢の推定

得られた3D-2Dの対応関係からPnP問題[11]を解くことでカメラ位置姿勢 \mathbf{M}_r を求める。誤対応はPROSAC法を用いて排除する。

(B-3) 特徴点の追跡

3次元位置が未知の点に対して、前フレームと現フレームでのテンプレートマッチングを行う。追跡精度を高めるため、画像全体で主要なオプティカルフローを求める[5]。フローにより予測された出現位置を中心に探索窓を設定することで、周囲に類似したパターンが存在する場合でも誤対応の少ない追跡が行える。

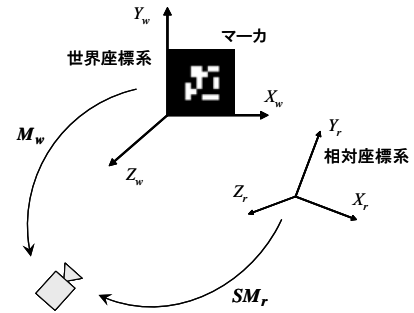


図3: 相対座標系と世界座標系の関係

(B-4) 特徴点の3次元位置の更新

十分なフレーム数追跡された特徴点に対し、追跡された全フレームのカメラ位置姿勢を用いて、マルチベースラインステレオ法により3次元位置を推定する。

(B-5) 特徴点の登録と削除

3次元位置と各フレームのカメラ位置を結ぶ直線の成す最大角について、閾値以上の特徴点をランドマークとして登録する。また、再投影誤差の標準偏差が閾値以上のランドマークは3次元位置の推定精度が低いとみなしてデータベースから削除する。

2.1.3. 後処理

順次処理の段階で推定された特徴点の3次元位置は相対的な値であるため、実世界とのスケールを合わせ、絶対的な座標値へ変換する必要がある。そこで後処理ではシーン中に配置されたマーカーを利用し、相対座標系から世界座標系への変換、およびスケール合わせを行う。本手法では、マーカーには ARToolKitPlus[6]を用い、マーカー座標系を世界座標系とする(図3)。

相対座標系における座標 \mathbf{P}_r から世界座標系における座標 \mathbf{P}_w への変換は、マーカーが検出されたフレームにおける相対座標系でのカメラ位置姿勢行列 \mathbf{M}_r とマーカー座標系でのカメラ位置姿勢行列 \mathbf{M}_w , そしてスケール変換行列 \mathbf{S} を用いて以下の式(2)で表される。

$$\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \mathbf{M}_w^{-1} \mathbf{S} \mathbf{M}_r \begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix}, \quad \mathbf{S} = \text{diag}(s, s, s, 1) \quad (2)$$

ここで、スケール s は式(3)より決まる。

$$s = \frac{l_w}{l_r} \quad (3)$$

ただし、 l_w は世界座標系におけるマーカー1辺の長さであり、既知とする。相対座標系におけるマーカー1辺の長さ l_r は、マルチベースラインステレオによって推定されたマーカー2頂点の3次元距離として計算される。

2.2. ランドマークデータベースの構成要素

データベースは、複数のランドマークによって構成され、各ランドマークは以下の情報を持つ。

- (1) ランドマークの3次元位置 \mathbf{P}_w
- (2) 撮影地点ごとの情報
 - (a) SIFT 特徴量 (128次元ベクトル \mathbf{V})
 - (b) SIFT スケール決定係数 a
 - (c) 登録時のカメラ位置姿勢 \mathbf{M}_w

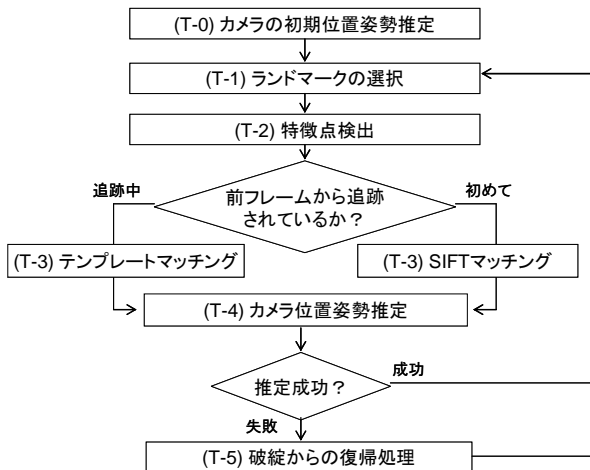


図 4: カメラトラッキング手法の流れ

(a) はトラッキング時に高精度なランドマークの対応付けを実現するためのものである。(b) はデータベース構築時とトラッキング時で同じ SIFT 特徴量を抽出するための値である。実際には、カメラとランドマーク間の 3次元距離と SIFT スケールを乗じた値である。なお、SIFT スケールとは画像上での特徴量記述領域の大きさとする。(c) はランドマークの選択時に用いる。

3. カメラトラッキング

本章では、ランドマークデータベースを用いたトラッキング手法について述べる。従来手法の計算コストが大きかったのは、対応付けに多重スケールな画像テンプレートを用いるためである。また、テンプレート作成には前フレームでのカメラ位置姿勢が要するため、トラッキングが破綻すると復帰することができなかった。本研究では、ランドマークの対応付けに SIFT 特徴量[7]を利用し、さらに破綻からの復帰処理を導入することで、トラッキングの高速化と安定化を図る。

3.1. トラッキングアルゴリズム

本手法では、最初にデータベース構築時とトラッキング時のカメラパスはほぼ同じという制約を利用し、カメラの初期位置姿勢を高速に推定する。次に、ランドマークの対応付け結果からカメラ位置姿勢を毎フレーム推定することでカメラトラッキングを実現する(図 4)。以下に各処理について詳述する。

(T-0) カメラの初期位置姿勢推定

初期フレームでは、前フレームのカメラ位置姿勢が与えられないため、トラッキングとは異なる方法で推定を行う必要がある。本稿では、データベース構築に用いる画像シーケンスから任意に選択された複数のキーフレームを利用することで、データベース構築時のカメラパス付近であれば高速にカメラの初期位置姿勢を推定可能な手法を提案する。

処理は 2つのステップから構成される。第 1ステップでは、入力画像と最も類似したキーフレームを検索する。画像間の類似度 c は

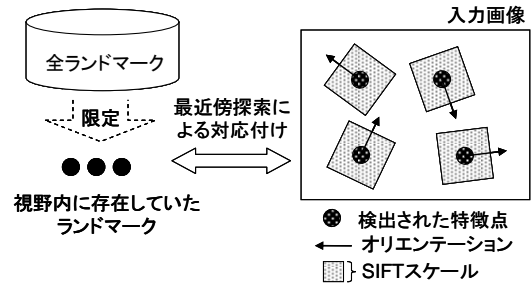


図 5: 破綻からの復帰における対応付け処理

$$c = \sum_{i=1}^N \frac{1}{SSD(V_i, V')} \quad (4)$$

から計算する。ただし、 N はキーフレーム内に存在する全ランドマーク数、 V_i は各ランドマークの SIFT 特徴量であり、 V' は最近傍である特徴点の特徴量を表す。

第 2 ステップでは、類似度 c が最大のキーフレーム内に存在するランドマークと入力画像中の特徴点を、kd 木を用いた最近傍探索によって対応付けることでカメラの初期位置姿勢を推定する。

(T-1) ランドマークの選択

利用するランドマークを限定するため、多数のランドマークから以下の条件を全て満たすものを選択する。

- (条件 1) 視野内に存在
- (条件 2) ランドマーク撮影時のカメラ位置と前フレームのカメラ位置の距離が閾値以下
- (条件 3) 入力画像におけるランドマークどうしの距離が閾値以上

フレーム毎の計算量を一定にするため、選択されるランドマーク数に上限を設ける。上限を越える場合には、前フレームで対応付けられたものを優先させ、フレーム間で利用するランドマークに連続性を持たせる。これは、表示した仮想物体の振動を抑えるためである。

(T-2~3) ランドマークの対応付け

まず、各ランドマークの対応付け候補を限定するため、入力画像中の探索窓内で検出された特徴点を各ランドマークの対応付け候補とする。探索窓は各ランドマークの投影座標を中心とした円として設定され、画像上への投影には前フレームのカメラ位置姿勢を使う。

次に、現フレームで新たに選択されたランドマークと、前フレームから追跡中のランドマークとで場合分けをし、それぞれ異なる特徴量により対応付けを行う。

(T-4) カメラ位置姿勢の推定

得られた 3D-2D の対応関係から PROSAC 法によって誤対応を排除し、カメラ位置姿勢を求める。ここでも表示した仮想物体の振動を抑えるため、算出されたカメラ位置姿勢に対しスムージングフィルタの一種である Double Exponential Smoothing Predictor[12]を適用し、最終的なカメラ位置姿勢として出力する。

3.2. 対応付けの高速化

本手法では、現フレームで新たに選択されたランドマークに対し、高速化した SIFT アルゴリズムによる対応付けを一度だけ行う。そして、以降のフレームで

の追跡をフレーム間でのテンプレートマッチングに切り替えることで対応づけ処理の計算コストを削減する。

一般的に SIFT 特徴量は識別性能が非常に高い一方で、計算コストが高く実時間処理には向かない。その原因は、特徴量記述領域を定めるためのスケール算出に計算コストが高い Difference of Gaussian (以下 DoG) 計算を行うからである。そこで本手法では、DoG 計算の代わりにランドマークとカメラの 3 次元位置から SIFT スケール L_{SIFT} を求める。 L_{SIFT} は以下の単純な式 (5) から算出できるため、高速な特徴量の記述が可能である。

$$L_{SIFT} = \frac{\alpha}{d} \quad (5)$$

ただし、 α はデータベースに登録された SIFT スケール決定係数であり、 d はカメラ位置とランドマークの 3 次元距離である。

各ランドマークに対し、データベースに登録された SIFT 特徴量と、入力画像から上記の方法で得た SIFT 特徴量を比較し、SIFT ベクトル間のなす角が最小となる候補を対応点とする。

3.3. トラッキング破綻からの復帰

本手法では、破綻時のカメラ位置と光軸方向は破綻前と比べてその変化量は小さいと仮定することで、手ぶれやオクルージョン等の原因により瞬間的に生じる破綻からの実時間復帰を実現する (図 5)。

まず、ランドマークと入力画像から検出された特徴点との照合に要する計算コストを抑えるため、対応付けに用いるランドマークは破綻する直前まで視野内に存在していたものに限定する。

次に、各ランドマークに対して破綻前後で同じ SIFT 特徴量を抽出するため、破綻前に求めた各ランドマークの SIFT スケールの平均を用いて特徴量を抽出する。カメラ位置の変化が破綻前後で小さければ、SIFT 特徴量が有する回転不変性により、カメラの姿勢情報が無くとも同じ特徴量を記述可能である。

最後に、効率的に対応付けるため得られた特徴量群から張られる特徴量空間内で近似最近傍探索による照合を行い、実時間での復帰を実現する。

4. 提案手法の性能評価

ランドマークデータベース構築に要する時間と、カメラトラッキングの速度と安定性を実験にて確認する。

実験環境として、PC (CPU: Core 2 Extreme 2.80 GHz, メモリ: 4GB), ビデオカメラ (SONY DSR-PD170, 720×480, プログレッシブ, 15fps) を使用し、カメラの内部パラメータは事前に求めた。

4.1. ランドマークデータベース構築の実験

屋内外の両環境においてマーカを配置し、データベース構築に要した時間をそれぞれ計測した。

屋内の画像シーケンスは、奥行き 5.00m, 幅 5.00m, 高さ 2.80m の映画撮影用セットに対し、カメラを約 3m の直線レール上で移動させて撮影した。得られた総フレーム数は 120 フレームであり、構築結果を図 6 に示す。構築に要した時間は約 71 秒、得られたランドマーク数は 126 個であった。また、3 次元位置の推定精度を確認した。2 つのランドマーク間の 3 次元距離を手動で計測した値を真値とし、推定値と比較する作業を広範囲から無作為に選んだ 10 組に対して行った。その結果、真値との平均誤差は 17.73mm であった。

屋外の画像シーケンスは、カメラを手持ちで約 10m 移動させて撮影した。得られた総フレーム数は 220 フレーム、構築に要した時間は約 160 秒、得られたランドマーク数は 169 個であった。図 7 の結果から世界座標系の原点がマーカ頂点に正確に設定されていないことが分かる。これは屋内実験と比べて移動距離が増えたことにより、累積誤差も増加したからと考えられる。

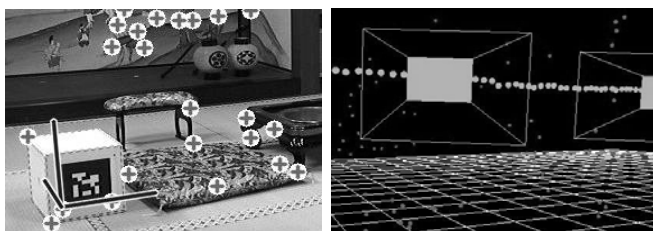
以上の結果より、提案手法は高速なデータベース構築を行えることを確認した。ただし、より高精度な推定を行うために累積誤差を抑える工夫が必要である。

4.2. カメラトラッキングの実験

トラッキング処理時間の計測と破綻からの復帰処理の実験を屋内で行った。トラッキング処理に用いる各種パラメータは次のように設定した。1 フレームで利用するランドマークの上限は 60 個とし、データベース登録時の SIFT スケールは 24 画素とした。また、追跡に用いる画像テンプレートのサイズは 11×11 画素とし、探索窓は半径 30 画素の円とした。



撮影された画像シーケンス 時間

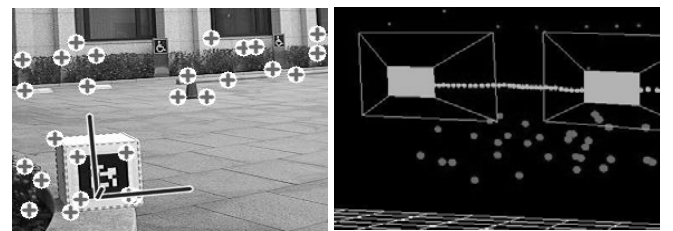


ランドマークの再投影結果
(マーカ上の線分は座標軸) 推定されたカメラパス

図 6: 屋内でのデータベース構築結果



撮影された画像シーケンス 時間



ランドマークの再投影結果
(マーカ上の線分は座標軸) 推定されたカメラパス

図 7: 屋外でのデータベース構築結果

表 1: トラッキングの 1 フレームあたりの処理時間

処理	時間 [ms]
(T-1) ランドマークの選択	0.21
(T-2) 特徴点検出	15.60
(T-3) 対応付け	1.68
(T-4) カメラ位置姿勢推定	0.56
合計	18.05

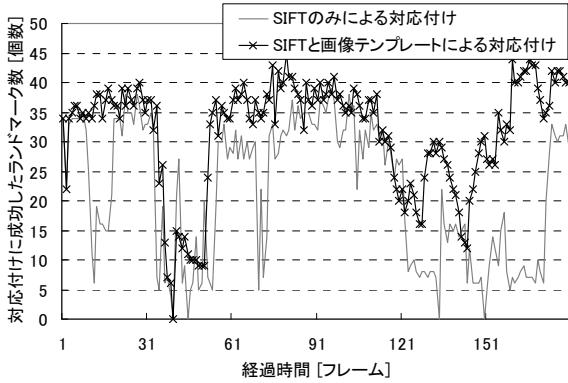


図 8: トラッキング破綻からの復帰

まず、提案手法の処理時間を表 1 に示す。従来手法は処理時間の合計は約 59ms だったのに対し、提案手法では約 18ms となり、対応付け処理の計算コストを抑えることで全体の処理時間が短縮されていることが分かる。なお、全てのランドマークの対応付けに SIFT 特徴量を用いた場合でも、対応付け処理に要した時間は 2.61ms となり、十分計算コストは削減されている。また、キーフレームを 10 枚利用した初期カメラ位置姿勢の推定には約 45ms の時間を要した。データベース構築時のカメラパスから平均 780mm 以内であれば、従来手法より短時間で推定可能であることを確認した。

次に、トラッキング破綻からの復帰が正しく行われるか確認するため、フレーム毎の対応付けに成功したランドマーク数を調べた (図 8)。本実験では、対応関係が 6 組以上得られないフレームを破綻とみなした。40 フレーム目で破綻しているが、その後復帰に成功していることが確認できる。また、130 フレーム周辺ではデータベース構築時のカメラパスから大きく離れたことで、対応付けられた数が減少している。SIFT のみで対応付けた場合には破綻するが、SIFT とテンプレートを切り替える提案手法では破綻していない。これは、フレーム間でのランドマークの見え方の変化は微小であるため、ランドマーク撮影時と比べて現フレームでの見え方が大きく変化していても追跡を続けられるからである。最後に、得られたカメラ位置姿勢を用いて CG キャラクタを重畳描画した結果を図 9 に示す。

5. むすび

本稿では、MR-PreViz に最適な特徴点ランドマークデータベースに基づくカメラトラッキング手法を提案した。提案手法では、従来手動で行っていた特徴点の 3 次元位置計測を自動化させることで、データベース構築に要する時間を大幅に削減することが可能である。トラッキング手法においては、SIFT 特徴量を用いるこ



100th frame



160th frame

図 9: カメラトラッキングに基づく MR 合成 (左: ランドマークの対応付け, 右: 合成結果)

とで従来以上に高速かつ安定したカメラ位置姿勢推定が可能となった。本手法は MR-PreViz への適用を想定しているが、準備の容易さやトラッキング性能は他の MR システムにも十分適用できるものと考えられる。

謝辞 本研究は、JST の CREST「映画制作を支援する複合現実型可視化技術」の支援による。

文献

- [1] 一刈良介, 川野圭祐, 天目隆平, 大島登志一, 柴田史久, 田村秀行: “映画制作を支援する複合現実型プレビューアライゼーションとカメラワーク・オーサリング”, 日本 VR 学会論文誌, Vol. 12, No. 3, pp. 343 - 354, 2007.
- [2] 武富貴史, 佐藤智和, 横矢直和: “拡張現実感のための優先度情報を付加した自然特徴点ランドマークデータベースを用いた実時間カメラ位置・姿勢推定”, 電子情報通信学会論文誌 (D), Vol. J92-D, No. 8, pp. 1440 - 1451, 2009.
- [3] Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar, and N. Navab: “Marker-less tracking for AR: a learning-based approach,” Proc. ISMAR, pp. 295 - 305, 2002.
- [4] G. Klein, and D. Murray: “Parallel tracking and mapping for small AR workspaces,” Proc. ISMAR, pp. 225 - 234, 2007.
- [5] D. Wagner, T. Langlotz, and D. Schmalstieg: “Robust and unobtrusive marker tracking on mobile phones,” Proc. ISMAR, pp. 121 - 124, 2008.
- [6] D. Wagner, and D. Schmalstieg: “ARToolKitPlus for pose tracking on mobile devices,” Proc. Computer Vision Winter Workshop, pp. 139 - 146, 2007.
- [7] D. Lowe: “Distinctive image features from scale-invariant keypoints,” Int. Journal of Computer Vision 60(2), pp. 91 - 100, 2004.
- [8] R. Hartley: “In defense of the eight-point algorithm,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, No. 6, pp. 580 - 593, 1997.
- [9] O. Chum, and J. Matas: “Matching with PROSAC - progressive sample consensus,” Proc. Conf. on CVPR, Vol. 1, pp. 220 - 226, 2005.
- [10] C. Harris, and M. Stephens: “A combined corner and edge detector,” Proc. Alvey Vision Conf., pp. 147 - 151, 1988.
- [11] F. Moreno-Noguer, V. Lepetit, and P. Fua: “Accurate non-iterative O(n) solution to the PnP problem,” Proc. Int. Conf. on Computer Vision, pp. 1 - 8, 2007.
- [12] J. LaViola: “Double exponential smoothing: an alternative to kalman filter-based predictive tracking,” Proc. the Workshop on Virtual Environments, pp. 199 - 206, 2003.