

三次元点群への局所平面当てはめに基づく法線情報付き 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

川戸 一希*† 武富 貴史† 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

〒 630-0192 奈良県生駒市高山町 8916-5

E-mail: †{kazuki-k,takafumi-t,tomoka-s,yokoya}@is.naist.jp

あらまし 現実環境を撮影した画像に対して CG などの仮想物体を重畳することで情報を付加するビデオスルー型拡張現実感においては、現実環境と仮想物体との幾何学的な位置合わせのためにカメラの位置・姿勢を推定する必要があり、従来から様々なカメラ位置・姿勢推定法が提案されている。中でもランドマークデータベースを用いた手法は Structure-from-Motion(SfM) 法を用いて半自動でデータベースを構築できるため、低い人的コストで広域環境に適用できるという特徴を持つ。ただし、従来手法ではランドマークと自然特徴点の対応付け処理において、カメラの姿勢の変化のみを考慮した画像の補正を行っているため、ランドマーク取得時のカメラ位置から離れた地点において、見え方が大きく変化するランドマークを正しく対応付けることができず、推定処理が失敗するという問題があった。この問題に対して、本研究では SfM 法により得られた三次元自然特徴点群に局所平面を当てはめることで、新たにその法線方向を算出する。この法線情報を用いて、平面上に存在すると判定されたランドマークについては視点位置に依存しない画像テンプレートを生成し、対応付けの成功率を向上させることで、視点位置の変化に対するカメラ位置・姿勢推定の頑健性を向上させる。

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, 自然特徴点, 拡張現実感

1. はじめに

近年、現実環境を撮影した画像中に CG などによって表現される仮想環境を位置合わせして描画することで情報を付加するビデオスルー型の拡張現実感技術に関する研究が盛んに行われている [1] ~ [4]。拡張現実感において、現実環境と仮想環境との位置合わせを実現するためには、一般にカメラの位置・姿勢を推定することが必要となり、現在までに様々な手法が提案されている。中でも、従来、我々が提案しているランドマークデータベースを用いた手法 [5] は SfM 法を用いることで広域な環境においても比較的容易にデータベースを構築できるという特長を持つ。しかし、従来手法はオンラインでのランドマークと自然特徴点の対応付け処理において、カメラの姿勢の変化のみを考慮した画像の補正を行っているため、ランドマーク取得時のカメラ位置から離れた地点において、見え方が大きく変化するランドマークを正しく対応付けることができず、推定精度が低下するという問題が残されていた。

本研究では、ランドマークデータベースを用いたカメラ位置・姿勢推定手法 [5] を基礎とし、新たに SfM 法により得られた三次元点群に対して主成分分析に基づく局所平面を当てはめを行うことで視点位置に依存しない画

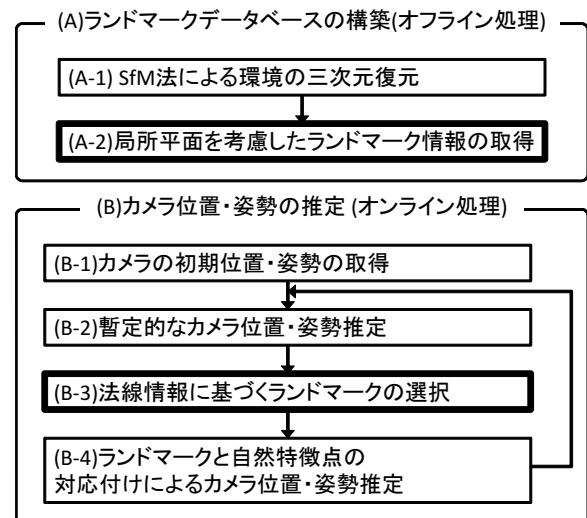


図 1 提案手法の処理の流れ

像テンプレートの生成を行う。視点位置に依存しない画像テンプレートをを用いることで、データベース構築時のカメラ位置から離れた地点においてもランドマークを正しく対応付け、推定精度の低下を防ぐ。図 1 に提案手法の処理の流れを示す。図中の太枠は、本研究で改良した処理を示す。ランドマークデータベースを用いたカメラ位置・姿勢推定手法は、オフラインでのデータベース構築処理 (A) とオンラインでのカメラ位置・姿勢推定処理 (B) から成る。提案手法では、処理 (A-2) において、平

* 現在, Ubitus Inc. Advanced Media Technology Group 所属
E-mail: kazuki.kawato@ubitusinc.com

面上に存在すると判定されたランドマークに対して、主成分分析により算出された局所平面上へ入力画像を投影することで撮影位置・姿勢に依存しない画像パターンを生成する。これにより、対応付けの成功率を向上させ、視点位置の変化に対して頑健なカメラ位置・姿勢推定を実現する。

2. 関連研究

本節では、画像に基づくカメラ位置・姿勢推定手法を事前知識を用いない手法と事前知識を用いる手法に大別し概観した後、本研究の位置づけについて述べる。

事前知識を用いない手法は visual-SLAM (Simultaneous Localization and Mapping) と呼ばれ、入力画像中の自然特徴点を追跡することによって、自然特徴点の三次元位置およびカメラの位置・姿勢を同時に推定する [6] ~ [8]。この手法は、事前準備無しに、その場で拡張現実感環境を構築できるという特長を持つが、visual-SLAM 単体では、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ため、位置に依存した情報を利用するナビゲーションなどのアプリケーションへの応用には向かない。このような問題に対して、Bleser ら [7] は、環境の一部のワイヤースケルトモデルと visual-SLAM を併用することで基準となる座標系を設定している。また、Klein ら [8] は、初期化の際にインタラクティブな操作によって、おおまかなスケール情報を取得している。しかし、これらの手法においても、広域で複雑な環境においてはモデル作成のための多くの人的コストが必要であるという問題や事前に基準となる座標系を定義できないために位置に依存した情報を利用できないという問題が残っている。

一方、環境の事前知識を用いる手法は、環境の三次元モデル [9], [10] やランドマークデータベース [5] を用いることで、絶対的なカメラ位置・姿勢を推定している。Drummond ら [9] は、ワイヤースケルトモデルを用い、入力画像中のエッジとワイヤースケルトモデルを対応付けることでカメラ位置・姿勢の推定を行っている。また、Reitmayr ら [10] は、テクスチャ付のポリゴンモデルを作成し、モデルの形状情報とテクスチャ中のエッジを併用することでロバストなカメラ位置・姿勢の推定を実現している。しかし、一般に広域で複雑な屋外環境の三次元モデルを作成することは難しく、事前知識の取得に多くの人的コストを必要とするという問題がある。これに対して、我々が提案しているランドマークデータベースを用いる手法 [5] では、自然特徴点の三次元位置と自然特徴点周辺の局所的な画像情報をランドマークとしてデータベースへ登録しておき、ランドマークと入力画像中の自然特徴点を対応付けることでカメラ位置・姿勢の推定を行う。自然特徴点の三次元位置は SfM 法によって自動で推定されるため、データベース構築の人的コストが低く、広域で複雑な屋外環境を対象とした場合にも比較的

容易にデータベースを作成することができる。しかし、従来手法では、撮影位置に依存した画像テンプレートの生成を行っているため、ユーザの近くに存在するランドマークの見え方を補正することが困難であり、高精度な位置合わせを実現することが難しい。このような問題に対して、我々は、高精度な位置合わせが必要な場所においてセンサによって取得された密な奥行き情報を用いる手法 [11] を提案しているが、密な奥行き情報の取得には多くの人的コストが必要となるため、広域環境への適用は困難である。

そこで、本研究では、ランドマークデータベースを用いた手法 [5] を基礎とし、従来手法で用いられている視点位置に依存した画像テンプレートに加えて、新たに視点位置に依存しない画像テンプレートを生成しデータベースへ登録する。オンライン処理では、生成した 2 種類の画像テンプレートを適応的に利用することで、ランドマーク取得時のカメラ位置から離れた地点におけるカメラ位置・姿勢推定精度の低下を防ぐ。

3. 局所平面当てはめに基づくランドマークデータベースの構築

本節では、オフラインでのデータベース構築処理 (A) について述べる。本研究では、ランドマークデータベース構築時のカメラパスから離れた場合においても頑健なカメラ位置・姿勢推定を実現するために、新たにランドマークの法線情報を付加したデータベースを構築する。ここでは、(A-1) SfM 法による環境の三次元復元、(A-2) 局所平面当てはめに基づくランドマーク情報の取得について述べる。

3.1 SfM 法による環境の三次元復元

本研究では、佐藤らの手法 [12] または Ikeda らの手法 [13] を用いることにより、広域で複雑な環境においても蓄積誤差の生じない三次元復元を行う。これらの手法では、まず、対象となる環境を移動しながら全方位カメラを用いて撮影する。次に、Harris オペレータ [14] によって検出された動画像中の自然特徴点を自動追跡することで、SfM 法によって自然特徴点の三次元位置と画像上の座標、および全方位カメラの外部カメラパラメータを推定する。最後に、三次元位置が既知の少数の基準点 [12]、または、GPS から得られる位置情報 [13] を利用して、動画像全体での最適化処理を行い、カメラの外部パラメータと自然特徴点の三次元位置の累積的な推定誤差を最小化することで、蓄積誤差の生じない三次元復元を行う。

3.2 主成分分析による三次元点群への局所平面の当てはめ

本研究では、視点位置に依存しない画像テンプレートを生成するために、SfM 法によって取得された三次元点

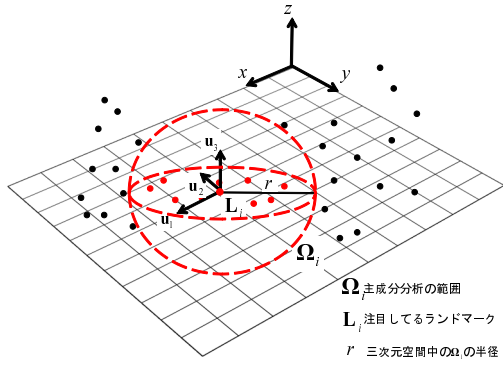


図 2 主成分分析の範囲

群に対して、主成分分析を用いて局所平面を当てはめ、平面上に存在するランドマークの判定とその法線方向の算出を行う。具体的には、まず、図 2 に示すように、ランドマーク $L_i(x_i, y_i, z_i)$ を中心とする半径 r 内に存在するランドマーク群 Ω_i の三次元座標を主成分分析することで、第一、第二、第三主成分に対応する固有値 $\lambda_1, \lambda_2, \lambda_3$ と固有ベクトル u_1, u_2, u_3 を算出する。ランドマーク群 Ω_i が同一平面上に分布している場合、第三主成分の方向の単位ベクトル u_3 は当てはめべき平面の法線方向を表し、全ての点が平面上に存在する場合、固有値 λ_3 は λ_2 に比べ、非常に小さな値を取る。そこで、本研究では、以下の条件を用いて平面上に存在するランドマークを検出する。

$$\left| \frac{\lambda_2}{\lambda_3} \right| > m \quad (1)$$

ここで、 m は正の値を取る定数である。

3.3 局所平面当てはめに基づくランドマーク情報の取得

本研究で用いるランドマークデータベースの構成要素を図 3 に示す。本研究では、従来の視点位置に依存した画像テンプレート (ii) に加え、平面上に存在すると判定されたランドマークに対して視点位置に依存しない画像テンプレート (iii) を生成しデータベースへ登録する。以下では、具体的な画像テンプレートの生成方法について述べる。

3.3.1 視点位置に依存した画像テンプレートの生成

視線方向の変化による見え方の変化に対応するために、従来手法と同様の方法で視点位置に依存した画像テンプレート $T_{dependent}$ を作成し、データベースへ登録する。ここでは、図 4 に示すように、世界座標系において、以下の条件を満たす正方形の平面を仮定し、この面上に入力画像を投影することで画像テンプレートを作成する。

- ランドマークの三次元位置を中心とする
- ランドマークの三次元位置とランドマーク撮影時のカメラ位置とを結ぶ直線に対して垂直である。
- 世界座標系での大きさが $S \times S$ である。

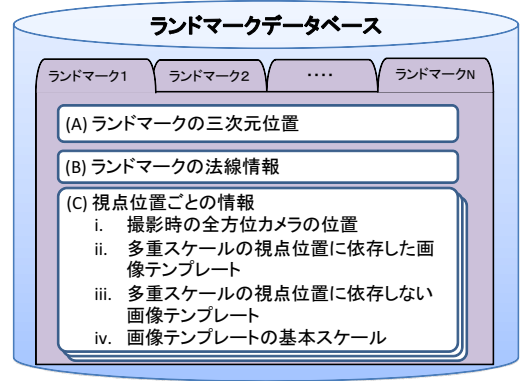


図 3 ランドマークデータベースの構成要素

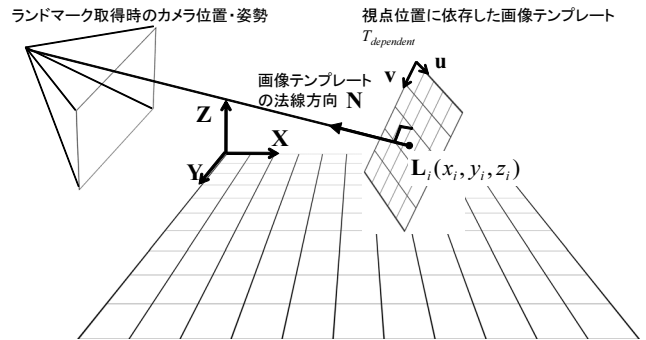


図 4 視点位置に依存した画像テンプレートの生成方法

ここで、画像テンプレートの空間中でのサイズ S は、入力画像の 1 画素と画像テンプレートの 1 画素が対応するように以下の式により決定する。

$$S = \frac{D}{f} 2^{(p-1)d} \quad (2)$$

ただし、 d はカメラの CCD における画素間の距離、 f はカメラの焦点距離、 D はランドマークの奥行き距離、 p は多重スケールテンプレートのスケール値 (1, 2, 3) を表す。また、画像テンプレートの基底 u, v は、以下のよう決定する

$$v = \frac{\hat{u} \times N}{\|\hat{u} \times N\|} \quad (3)$$

$$u = \frac{N \times v}{\|N \times v\|} \quad (4)$$

$$\hat{u} = Z \times N \quad (5)$$

ただし、 N は画像テンプレートの法線方向を表す。ここで、画像テンプレートの法線ベクトル N が Z 軸に対して平行である場合、画像テンプレートの座標 u, v を一意に決定することができない。そのような場合には、以下の式で u, v を決定する。

$$v = \frac{\hat{u} \times N}{\|\hat{u} \times N\|} \quad (6)$$

$$u = \frac{N \times v}{\|N \times v\|} \quad (7)$$

$$\hat{u} = X \times N \quad (8)$$

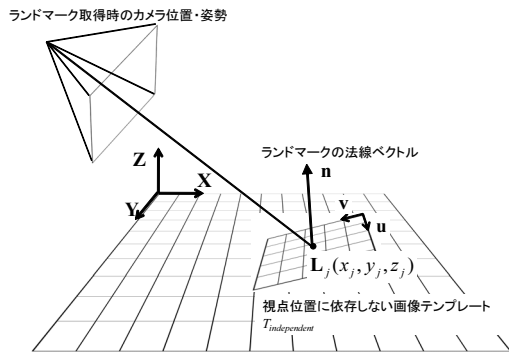


図 5 視点位置に依存しない画像テンプレートの生成方法

最後に、仮定した平面上に撮影画像を投影することで、撮影時のカメラ姿勢やレンズ歪みの影響を排除した画像テンプレートを作成する。画像テンプレートの画素数を $N \times N$ とした場合、この投影面を等間隔な格子で $N \times N$ に分割し、各格子点の三次元座標を撮影画像面上に投影した上で、その投影座標の画素値をその点の画素値とする。このとき、画像テンプレートの 1 画素が入力画像上の 1 画素と対応するように画像テンプレートの大きさを決定し、これを基本スケールの画像テンプレートとする。このような手法により、多重スケール ($p = 1, 2, 3$) の画像テンプレートを作成し、これらを撮影地点ごとに保持する。

3.3.2 視点位置に依存しない画像テンプレートの生成

主成分分析を用いた平面当てはめによって平面上に存在すると判定されたランドマークに対して、視点位置に依存しない画像テンプレート $T_{independent}$ を生成する。ここでは、図 5 に示すように、主成分分析により算出したランドマークの法線ベクトル n に対して垂直な正方形の平面を仮定し、3.3.1 節と同様の方法で u, v を算出する。ここで、視点位置に依存しない画像テンプレートは、カメラの投影中心とランドマークを結ぶ直線に対して大きく傾くことがある。そのため、従来と同様の手法によって、スケールの大きい画像テンプレートを生じた場合、画像テンプレートの一部の解像度が低下する。そこで、本研究では、高解像度な画像テンプレートを作成し、その画像テンプレートから必要な範囲を切り出し、縮小することで視点位置に依存しない画像テンプレートを生成する。これにより、画像テンプレート生成時の解像度の低下を抑える。

4. 局所形状を考慮したランドマーク選択に基づくカメラ位置・姿勢推定

オンラインでのカメラ位置姿勢推定処理は、従来手法と同様に、まず、初期フレームでのカメラ位置・姿勢を何らかの方法により取得する (B-1)。次に、暫定的なカメラ位置・姿勢推定 (B-2)、ランドマークの選択 (B-3)、ランドマークと自然特徴点との対応付けによるカメラ位置・姿勢推定 (B-4) を繰り返す。本手法において、平面

上に存在するランドマークは 2 種類の画像テンプレートを保持するため、ランドマークの選択処理において、画像テンプレートの法線情報、および、ランドマークとカメラの距離に基づき、一方の画像テンプレートを適応的に選択する。

4.1 暫定的なカメラ位置・姿勢の推定

連続フレーム間でランドマークを追跡することで、現フレームの暫定的なカメラ位置・姿勢を推定する。暫定的なカメラ位置・姿勢推定では、まず、前のフレームで自然特徴点と対応付けられ、誤対応として排除されずにカメラ位置・姿勢の推定に用いられたランドマークを選択し、現フレームにおける対応点を探索する。ランドマークの対応点探索において、連続フレーム間ではパターンの変形や明るさの変化が微小であることから、ランドマークの前フレームの画像上の位置を中心とする一定のウィンドウ内に存在する自然特徴点を対象として SSD (Sum of Squared Differences) を評価尺度としたテンプレートマッチングによりランドマークの追跡を行う。

次に、追跡した自然特徴点を用いて現フレームの暫定的なカメラ位置・姿勢を推定する。カメラ位置・姿勢の推定では、まず、LMedS 基準 [15] を用いて誤対応を排除する。次に、正しい対応関係を用いて、カメラ位置・姿勢の推定を行い、再投影誤差を最小化することによって現フレームの暫定的なカメラ位置・姿勢を推定する [16]。

4.2 法線情報に基づくランドマークの選択

ここでは、対応付けの候補となる N_{max} 個のランドマークとその画像テンプレートをデータベースより選択する。提案手法では、平面上に存在すると判定されたランドマークは、視点位置に依存した画像テンプレート $T_{dependent}$ に加え、視点位置に依存しない画像テンプレート $T_{independent}$ を保持している。そのため、それらの画像テンプレートの法線情報に基づき、対応付けに用いる画像テンプレートの種類を適応的に決定する。

ランドマークの選択では、以下の条件に基づき、ランドマークとその画像テンプレートを選択する。

条件 1: ランドマークの三次元位置を暫定的なカメラ位置・姿勢を用いて撮像面に投影する。これにより、入力画像の画角内に存在するランドマークのみを用いる。

条件 2: 図 6 に示すように、ランドマークの三次元位置とランドマークが撮影されたカメラ位置とを結ぶ直線とランドマークの三次元位置と暫定カメラ位置を結ぶ直線とのなす角 θ_c が閾値以上であれば、そのランドマークを用いない。しかし、 θ_c が閾値以上であっても、そのランドマークが平面上に存在し、ランドマークの三次元位置と暫定カメラ位置との距離 d が閾値 D_{max} 以下の場合、このランドマークを用いる。

ここで、図 6 に示すように、多くの場合、ランドマークの三次元位置と暫定カメラ位置を結ぶ直線とランド

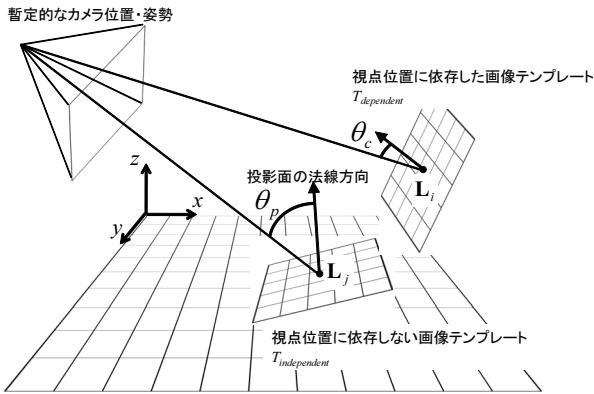


図 6 カメラに対する画像テンプレートの傾き

マークの法線ベクトルとの成す角 θ_p はランドマークの三次元位置とランドマークが撮影されたカメラ位置とを結ぶ直線とランドマークの三次元位置と暫定カメラ位置とを結ぶ直線とのなす角 θ_c に対して大きくなる．そのため、従来手法 [5] と同様に θ_c, θ_p の大きさを基準に画像テンプレートの選択を行うと、視点位置に依存した画像テンプレート $T_{dependent}$ が優先的に用いられるという問題が生じる．そこで、本研究では以下の式で定義する優先度 C を新たに定義し、この優先度に基づいてテンプレートの選択を行う．具体的には、まず (条件 1)、(条件 2) を満たすランドマークを、ランドマーク取得時のカメラ位置と暫定カメラ位置の距離 d_{view} でソートし、 d_{view} が小さい順に N'_{max} 個のランドマークを選択する．次に、選択されたランドマークを以下に定義する優先度 C を用いてソートし、優先度の高い N_{max} 個のランドマークを選択する．

$$C = \begin{cases} (w\theta_c + 1)^{-1} & (T_{dependent} \text{ の場合}) \\ (\theta_p + 1)^{-1} & (T_{independent} \text{ の場合}) \end{cases} \quad (9)$$

ここで、 w は θ_c と θ_p のバランスを取るための重みである．

4.3 ランドマークと自然特徴点の対応付けによるカメラ位置・姿勢の推定

ここでは、まず、暫定的なカメラパラメータを用いて、処理 (B-3) で選択されたランドマークを入力画像上へ投影する．次に、投影したランドマークの画像上の位置を中心とする一定の範囲内に存在する自然特徴点に対してデータベース構築時と同様の方法で画像パターンを生成し、ランドマークの画像テンプレートと正規化相互相関法によって照合する．最後に、ランドマークと対応付けられた自然特徴点を用いて、現フレームのカメラ位置・姿勢を推定する．カメラ位置・姿勢の推定では、暫定的なカメラ位置・姿勢の推定処理と同様に、LMedS 基準による誤対応の排除後、正しい対応関係を用いて、再投影誤差を最小化することによって現フレームのカメラ位置・姿勢を推定する．

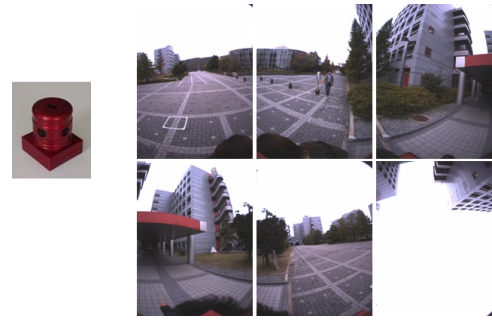


図 7 全方位型マルチカメラシステムと全方位画像の 1 フレーム

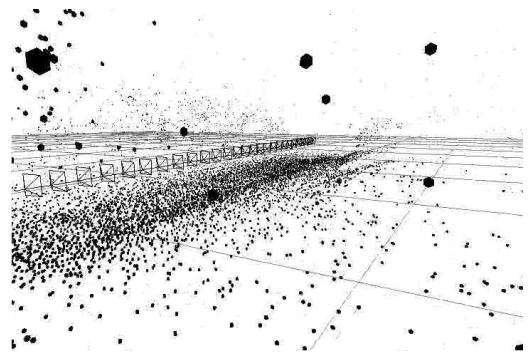


図 8 推定された自然特徴点の三次元位置と全方位カメラの位置・姿勢

5. 実験

提案手法の有効性を示すために、従来手法 [5] との比較により、カメラパスから離れた場合におけるカメラ位置・姿勢推定の頑健性の向上効果について検証した．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、次に単眼カメラで撮影した動画像を用いた実験について述べる．

5.1 ランドマークデータベースの構築

本実験では全方位マルチカメラシステム (Point Grey Research 社 Ladybug, 解像度 $768 \times 1024 \times 6$ 台) を用いて屋外環境を移動撮影することにより、全方位動画像 (621 フレーム, カメラパス約 35m) を撮影し、SfM 法 [12] により自然特徴点の三次元位置と全方位カメラのカメラパラメータを推定した．図 7, 図 8 に取得された全方位画像と三次元復元の結果を示す．提案手法では、これらの結果に対して、局所平面を当てはめ、法線情報の算出を行うことでランドマークデータベースを構築した．ここで、本実験では主成分分析を行う範囲を半径 $r = 3$ メートルとした．本実験において、データベースに登録されたランドマークの数は約 6500 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 6 個所であった．主成分分析による平面当てはめの結果を図 9 に示す．図中の青色の点は平面以外と判定されたランドマーク、赤色の点は平面と判定されたランドマーク、赤色の線はその

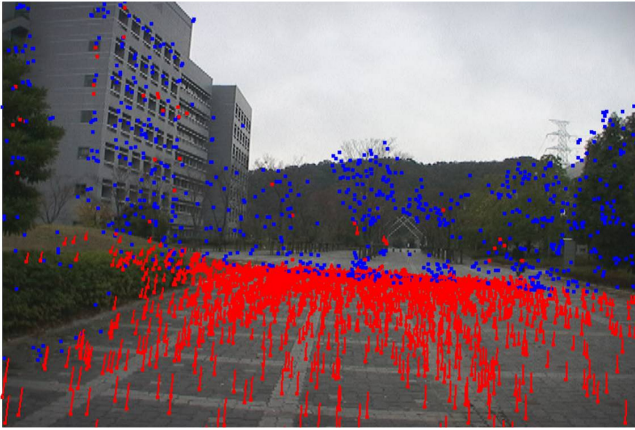


図 9 再投影されたランドマークとその法線情報

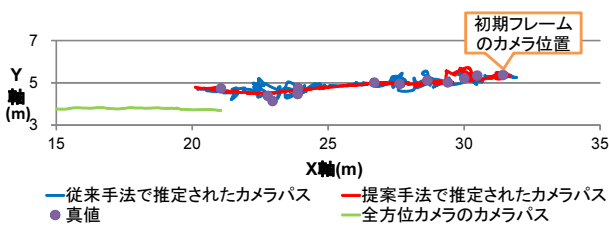


図 10 カメラの初期位置と推定されたカメラパス

ランドマークの法線方向を表す。同図は、4.2 節以降で用いる入力画像の 1 フレームに対して、後述するカメラパラメータの真値を用いてランドマークを重畳合成している。この結果から、法線が推定されているランドマークは主に地面上に存在するランドマークであり、それらの法線方向がおおむね正しく推定されていることが分かる。

5.2 カメラ位置・姿勢の推定精度の比較

本実験では、表 1 に示すような機器を用い、図 10 に示すように、ランドマークデータベース構築時のカメラパスから約 10m 離れた地点からランドマークデータベース構築時のカメラパスに近づくような経路で、動画像 (解像度 720×480 画素, プログレッシブスキャン, 15fps, 170 フレーム) を撮影し、推定されたカメラ位置・姿勢を従来手法 [5] と比較した。ただし、本実験では、提案手法、従来手法ともにカメラの内部パラメータは校正済みかつ既知とし、初期フレームのカメラ位置・姿勢は環境内の自然特徴点の三次元位置をトータルステーションにより計測し、画像上での自然特徴点と手動で対応付け、PnP 問題を解くことで算出した。本実験で用いたパラメータを表 2 に示す。世界座標系は、X 軸、Y が現実環境における地面に対して水平、Z が地面に対して垂直で、上向きの座標系とした。

まず、自然特徴点とランドマークの対応付けの成功率

表 1 カメラ位置・姿勢推定に用いた機器

入力撮影カメラ	SONY DSR-PD-150
計算機	CPU:Xeon(TM) 2.67GHz, Memory:12.0GB

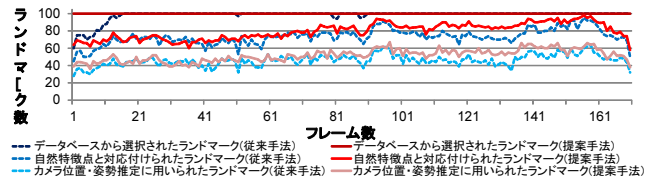


図 11 各フレームにおける対応点数

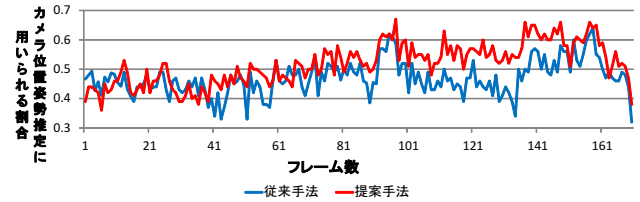


図 12 各フレームにおけるカメラ位置姿勢推定に用いられる割合

を検証するために、データベースから選択されたランドマーク、自然特徴点と対応付けられたランドマーク、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従来手法と提案手法で比較した。得られたランドマークの数の推移を図 11、各フレームにおける、データベースから選択されたランドマークの内、カメラ位置・姿勢推定に用いられる割合を図 12、1 フレームあたりの平均数を表 3 に示す。これらの結果から、提案手法は従来手法と比較して、多くのランドマークが最終的なカメラ位置・姿勢推定に用いられているのに対して、データベースから選択されているランドマーク数はほぼ同数である。このことから、提案手法は従来手法と比較して多くのランドマークが頑健に対応付いていることがわかる。各フレームにおける、データベースから選択されたランドマークの内、カメラ位置・姿勢推定に用いられる割合に対して有意水準 1% とした t 検定を行い、提案手法と従来手法との比較を行った結果、有意差が認められたことから、平均的に提案手法がより多くのランドマークを頑健に対応付けることができると考えられる。

表 2 実験に用いたパラメータ

	従来手法	提案手法
$T_{dependent}$ のサイズ (画素)	15×15	15×15
$T_{independent}$ のサイズ (画素)	-	15×15
$T_{independent}$ のスケール (倍)	-	1, 5, 9
N_{max}	100	100
θ_c の閾値	15°	15°
重み w	-	9

表 3 1 フレームあたりのランドマーク数

	提案手法	従来手法
データベースから選択されたランドマーク数	100	98.5
自然特徴点と対応付けられたランドマーク数	79.6	73.3
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	51.7	46.1
選択されたランドマークの内、カメラ位置・姿勢推定に用いられる割合	0.80	0.76

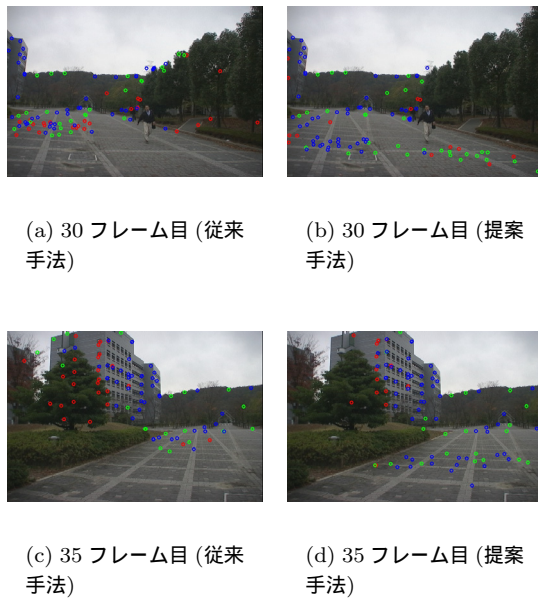


図 13 カメラ位置・姿勢推定に用いられたランドマークの比較

図 13 にカメラ位置・姿勢推定に用いられたランドマークの画像上の位置を示す．丸印は，3.2 節で述べた (条件 1)，(条件 2) を満たすランドマークを表しており，図中の赤の丸印は正規化相互相関を用いた照合により入力画像の自然特徴点と対応付けられなかったランドマーク，緑の丸印は処理 (B-4) で LMedS 基準により誤対応と判定されたランドマーク，青の丸印は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を表す．提案手法では，従来手法よりも視点位置に近い位置に存在するランドマークを数多く対応付けることができていることが確認できる．これは，視点位置に依存しない画像テンプレートを用いることで，画像パターンが適切に補正されたためだと考えられる．

従来手法と提案手法により推定されたカメラパスの比較を行った．推定されたカメラパスと真値，データベース構築時の全方位カメラのカメラパスを図 10 に示す．真値は，初期フレームにおけるカメラの位置・姿勢と同様に手動で対応点を与え，PnP 問題を解くことで算出した．この結果から，提案手法により推定されたカメラパスは真値から大きく離れることなく，おおむね正しく推定できていることが確認できる．それに対して，従来手法により推定されたカメラパスでは一部のフレームにおいて，真値から大きく外れるような誤差が生じていることが分かる．これは提案手法において，カメラの位置を決定するために重要なユーザの近くに存在するランドマークが正しく対応づいたことにより，改善された点であると考えられる．

図 14 に，従来手法，提案手法によって推定されたカメラ位置・姿勢を用いて仮想物体を重畳した結果と真値を用いて仮想物体を重畳した結果を示す．111 フレーム目，121 フレーム目では，真値と比較して従来手法の仮

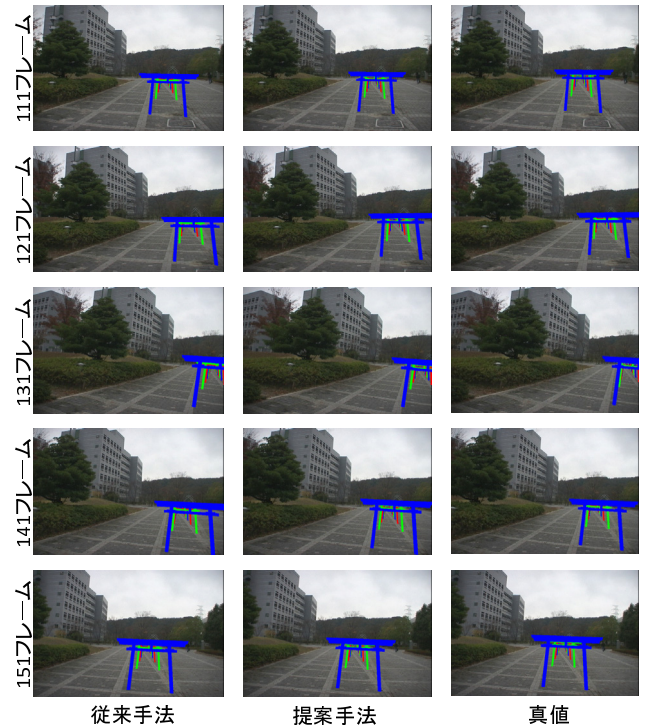


図 14 仮想物体の重畳結果

想物体が手前に重畳されているのに対して，提案手法ではおおむね正しく重畳されている．また，141 フレーム目では，従来手法，提案手法ともに，仮想物体が真値と比較して仮想物体が手前に重畳されているが，提案手法は従来手法と比較して，位置ずれが少ないことが確認できる．この結果から，提案手法は従来手法と比較して正しく仮想物体を重畳していることが確認できる．

6. ま と め

本論文では，ランドマークデータベースを用いた動画からのカメラ位置・姿勢推定において，ランドマークの三次元位置に局所平面を当てはめ，これを用いて幾何学的に補正した画像テンプレート利用することで，ランドマークデータベース構築時の全方位カメラの経路から離れた地点で撮影した入力画像に対しても頑健なカメラ位置・姿勢推定を実現する手法を提案した．入力画像とランドマークデータベースの対応付けにおいて，平面上に存在するランドマークに対して視点位置に依存しない画像テンプレートを用いることで，視点位置に近いランドマークの頑健な対応付けを実現し，これにより，カメラ位置・姿勢の推定精度の向上を実現させた．

今後の課題として，ジャイロなどのセンサを併用することによる頑健性の向上などが挙げられる．

謝辞 本研究の一部は科学研究費補助金 (基盤研究 (A), No.23240024, 若手研究 (B), No.23700208), 文部科学省特別教育研究経費「アンビエント環境知能研究創出事業」による．

文 献

- [1] D. Wagner and D. Schmalstieg: “First steps towards handheld augmented reality”, Proc. Int. Symp. on Wearable Computers, pp. 21–23 (2003).
- [2] T. Miyashita, P. Meier, T. Tachikawa, S. Orlic, T. Eble, V. Scholz, A. Gapel, O. Gerl, S. Arnaudov and S. Lieberknecht: “An augmented reality museum guide”, Proc. Int. Symp. on Mixed and Augmented Reality, pp. 103–106 (2008).
- [3] P. Daehne and J. Karigiannis: “Archeoguide: System architecture of a mobile outdoor augmented reality system”, Proc. Int. Symp. on Mixed and Augmented Reality, pp. 263–264 (2002).
- [4] H. Kaufmann and A. Dunser: “Summary of usability evaluations of an educational augmented reality application”, Proc. Int. Conf. on Human-Computer Interaction, pp. 22–27 (2007).
- [5] 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), **J92-D**, 8, pp. 1440–1451 (2009).
- [6] A. Davison, Y. G. Cid and N. Kita: “Real-time 3D slam with wide-angle vision”, Proc. IFAC Symp. on Intelligent Autonomous Vehicles (2004).
- [7] G. Bleser, H. Wuest and D. Stricker: “Online camera pose estimation in partially known and dynamic scenes”, Proc. Int. Symp. on Mixed and Augmented Reality, pp. 56–65 (2006).
- [8] G. Klein and D. Murray: “Parallel tracking and mapping for small AR workspaces”, Proc. Int. Symp. on Mixed and Augmented Reality, pp. 225–234 (2007).
- [9] T. Drummond and R. Cipolla: “Real-time visual tracking of complex structures”, IEEE Transactions on Pattern Analysis and Machine Intelligence, **Vol. 27**, pp. 932–946 (2002).
- [10] G. Reitmayr and T. Drummond: “Going Out: Robust model-based tracking for outdoor augmented reality”, Proc. Int. Symp. on Mixed and Augmented Reality, pp. 109–118 (2006).
- [11] T. Taketomi, T. Sato and N. Yokoya: “AR Cultural Heritage Reconstruction Based on Feature Landmark Database Constructed by using Omnidirectional Range Sensor”, Proc. Asian Conf. on Computer Vision Workshop on e-Heritage (2010).
- [12] 佐藤智和, 池田聖, 横矢直和: “複数動画画像からの全方位型マルチシステムの位置・姿勢パラメータの推定”, 電子情報通信学会論文誌 (D-II), **J88-D-II**, 2, pp. 347–357 (2005).
- [13] S. Ikeda, T. Sato, K. Yamaguchi and N. Yokoya: “Construction of Feature Landmark Database Using Omnidirectional Videos and GPS Positions”, Proc. Int. Conf. on 3-D Digital Imaging and Modeling, pp. 249–256 (2007).
- [14] C. Harris and M. Stephens: “A combined corner and edge detector”, Proc. Alvey Vision Conf., pp. 147–151 (1988).
- [15] P. J. Rousseeuw: “Least median of squares regression”, J. of the American Statistical Association, **Vol. 79**, pp. 871–880 (1984).
- [16] R. Klette, K. Schluns and A. Koschan Eds: “Computer Vision: Three-dimensional Data from Image”, Springer (1998).