

## 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

武富 貴史<sup>†</sup> 佐藤 智和<sup>†</sup> 横矢 直和<sup>†</sup>

Real-Time Extrinsic Camera Parameter Estimation Using a Feature Landmark Database with Priorities for Augmented Reality

Takafumi TAKETOMI<sup>†</sup>, Tomokazu SATO<sup>†</sup>, and Naokazu YOKOYA<sup>†</sup>

あらまし 拡張現実感において、現実環境と仮想環境の幾何学的な位置合わせを実現するためには、カメラの位置・姿勢を推定する必要があり、現在までに様々なカメラ位置・姿勢推定手法が提案されている。なかでも、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、屋内外を問わず利用でき、広域な環境を対象とした場合にもカメラ位置・姿勢推定の誤差が累積しないという特長を持つ。しかし、従来手法ではデータベース中のランドマークと入力画像上の自然特徴点との対応付け処理に多くの計算コストを必要とするため、実時間でのカメラ位置・姿勢の推定が困難であるという問題がある。そこで、本論文では連続フレーム間でのランドマークの追跡とランドマークへの優先度情報の付加により、対応付け処理において用いられる対応点候補数を削減することで、実時間でのカメラ位置・姿勢推定処理を実現する手法を提案する。実験では、従来手法との比較によって計算コストの削減効果を検証する。また、実時間処理可能な試作システムを用いた複数のアプリケーションへの応用例を紹介し、提案手法の有効性を示す。

キーワード ランドマークデータベース、カメラ位置・姿勢推定、自然特徴点、拡張現実感

## 1. はじめに

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拡張現実感<sup>1</sup>は、ヒューマンナビゲーション [1]~[3], 各種の作業支援 [4], [5], 教育支援 [6] などの広い分野への応用が期待されており、盛んに研究されている。このような拡張現実感技術を用いたアプリケーションにおいて、現実環境と仮想環境の幾何学的な位置合わせを実現するためには、カメラの位置・姿勢を推定することが必要となり、現在までに様々なカメラ位置・姿勢推定手法が提案されている [7]~[15].

このような研究では、一般に、現実環境を撮影した画像を CG の合成対象として用いるだけでなく、カメラ位置・姿勢推定にも用いることで、精度の高い位置合わせを実現している。また、これらの手法の多くは、

環境に対する事前知識として、画像データベース、環境の三次元モデル、ランドマークデータベース等を利用することで現実環境に対する仮想物体の合成位置を決定している。なかでも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・姿勢推定手法 [12] は、複雑な屋外シーンにおいてもデータベース構築の人的コストが低く、広域環境を対象とした場合にもカメラ位置・姿勢推定の誤差が累積しないという特長を持つ。しかし、自然特徴点ランドマークデータベースを利用した従来手法 [12] では、データベースに登録されている多数のランドマークと入力画像上の自然特徴点を正しく対応付けるために多くの処理時間を必要とし、実時間でのカメラ位置・姿勢の推定が困難であるという問題が残されている。具体的には、従来手法 [12] では、データベース中より選択した  $N$  個のランドマークと入力画像中より抽出した平均  $F$  個の自然特徴点とを総当たりで照合することで対応付けているため、対応付け処理における計算コスト  $C_{prev}$  が  $N, F$  の積に比例し、このコストがカメラ位置・姿勢推定処理の大半を占めていた。著者らは従来、固有空間法によるラン

<sup>†</sup> 奈良先端科学技術大学院大学, 奈良県生駒市高山町 8916-5  
Nara Institute of Science and Technology, 8916-5 Takayama,  
Ikoma, Nara, 630-0192 Japan

ドマークの特徴次元数の削減によって1回当たりのランドマークと自然特徴点の照合コスト  $A$  を低減することを試みた [16] が、マッチングの精度を保ったままランドマークと自然特徴点の照合コスト  $A$  を低減することは難しいという結論に至った。

本研究では、対応付け処理に用いられるランドマーク数  $N$  と照合される自然特徴点数  $F$  をそれぞれ削減することで、対応付け処理における計算コスト  $C_{prev} = NFA$  を削減し実時間でカメラ位置・姿勢推定を実現する手法を提案する。具体的には、従来手法 [12] に対して以下の2つの改良を加える。

- (1) 連続フレーム間でランドマークを追跡し暫定的なカメラの位置・姿勢を推定することで照合すべき自然特徴点数  $F$  を削減する
- (2) ランドマークへ優先度情報を付加することで、正しく対応づく可能性の高い少数のランドマークを選択し、対応付け処理に用いるランドマーク数  $N$  を削減する

また、従来手法 [12] では、カメラの初期位置・姿勢を手動で与える必要があり、ユーザの補助を必要としていたが、本研究ではこの点についても改善し、ランドマークデータベースを用いた静止画像からのカメラ位置・姿勢推定手法 [17] を導入することで、カメラの初期位置・姿勢推定処理を自動化する。

提案手法の処理の流れを図1に示す。図中の太枠は、従来手法 [12] に対して本研究で追加または改良した処理を表す。ランドマークデータベースを用いたカメラ位置・姿勢推定は、オフラインでのデータベースの構築処理 (A) とオンラインでのカメラ位置・姿勢の推定処理 (B) からなる。

以下、2. では、画像に基づくカメラ位置・姿勢推定の関連研究と本研究の位置づけについて述べる。3. では、本研究で用いるランドマークデータベースの構成およびオフラインでの構築処理について述べ、4. では、優先度を付加したランドマークデータベースを用いたオンラインでのカメラ位置・姿勢推定処理について述べる。5. では提案手法の理論的な処理速度の向上効果を見積もり、6. では実験により実際の処理速度向上効果を確認する。また、試作システムを用いた実時間での仮想物体の合成実験により、実アプリケーションへの応用の可能性についても検証する。最後に、7. でまとめと今後の課題について述べる。

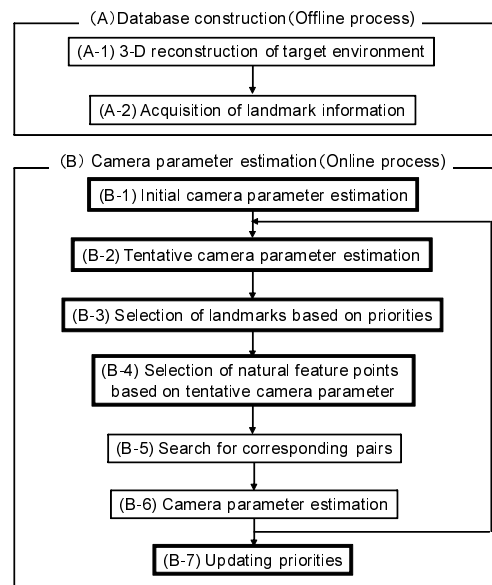


図1 提案手法の処理の流れ

Fig. 1 Flow diagram of the proposed method

## 2. 関連研究

本章では、画像に基づくカメラの位置・姿勢推定手法 [7]~[15] を環境の事前知識を用いる手法 [7]~[12] と環境の事前知識を用いない手法 [13]~[15] に大別しそれぞれの特徴を概観し、センサ融合によるそれらの手法の拡張 [18]~[20] についても述べた後、本研究の位置づけについて述べる。

環境の事前知識を用いる手法 [7]~[12] には、事前知識として、画像データベースを用いる手法 [7], [8]、環境の三次元モデルを用いる手法 [9], [10]、ランドマークデータベースを用いる手法 [11], [12] がある。画像データベースを用いる手法 [7], [8] では、事前に環境内を撮影した画像とその撮影位置・姿勢をデータベースへ登録しておき、入力画像とデータベース中の画像を対応付けることでカメラの位置・姿勢を推定する。岩佐ら [7] は、入力画像と最も類似した画像をデータベース中より選び出し、その画像の撮影位置を入力画像の撮影位置として推定している。この手法では、大まかなカメラ位置の推定は可能であるが、正確なカメラ位置を推定することは難しい。Cipolla ら [8] は、文献 [7] の手法と同様にデータベース中から類似した画像を選択した上で、その画像が撮影されたカメラ位置からの相対的な位置・姿勢を推定することで推定精度を高め

る手法を提案している。しかし、この手法では撮影地点の高さを推定することができず、6自由度で正確なカメラ位置・姿勢を推定することは難しい。また、これらの手法は入力画像とデータベース中の画像の照合に多くの計算コストを必要とし、実時間での処理が難しいという問題がある。環境の三次元モデルを用いる手法[9],[10]では、入力画像中のエッジや自然特徴点などをあらかじめ作成した三次元モデルと対応付けることによって、実時間でのカメラの位置・姿勢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、データベースの構築に多くの人的コストを必要とするという問題がある。ランドマークデータベースを用いる手法[11],[12]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 structure-from-motion 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、データベースに登録されている多数のランドマークと画像上の自然特徴点を正しく対応付けるために多くの処理時間を必要とし、実時間でのカメラ位置・姿勢の推定が困難であるという問題が残されている。

これに対して、あらかじめ特徴空間を KD-Tree によって分割しておくことで、高速に対応付け処理を行う手法[21]が提案されているが、実時間での処理には到っておらず、また、広域な屋外環境のように高次元の特徴空間上で大量のデータを扱う場合には、このようなアプローチによる探索処理の高速化を行うことは難しい。また、著者らは、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、各地点から観測できるランドマークを用いて位置に依存した固有空間を作成しておき、ランドマークの照合を固有空間上で行うことにより対応付けの計算コストを低減する手法[16]を試みたが、マッチングの精度を保ったままランドマークと自然特徴点の照合コストを低減することは難しく実時間処理の実現には到らなかった。

一方、環境の事前知識を用いない手法[13]~[15]では、一般に入力画像中の自然特徴点を追跡することにより、カメラ位置・姿勢の推定と環境のモデル化を同時に行う SLAM(Simultaneous Localization and Map-

ping) と呼ばれるアプローチが用いられる。これらの手法では、事前にデータベースを構築する必要がなく、手軽に拡張現実環境を構築できるという特長があるが、広域な環境を対象としたカメラの位置・姿勢推定では、カメラ位置・姿勢の推定誤差が蓄積するという問題があり、利用できる範囲が限定される。また、SLAM によるカメラ位置・姿勢の推定手法は、それ単体では現実環境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ため、拡張現実感技術を用いたナビゲーションのような実環境の位置に依存した情報の提示を必要とするアプリケーションにそのまま用いることはできない。

また、画像に基づくカメラ位置・姿勢推定の処理速度やロバスト性を向上させるため、GPS やジャイロなどの計測機器とカメラ画像を併用したハイブリッド手法[18]~[20]が提案されている。これらの手法では、計測機器から得られる情報よりカメラの大まかな動きを推定することで、ロバスト性の高い実時間カメラ位置・姿勢推定を実現しているが、計測機器によって利用可能な環境に制約があり、様々な環境で用いるためには多数の計測機器を組み合わせる必要がある。また、現在のところ画像からの絶対位置・姿勢の取得には、データベースの構築コストやカメラ位置・姿勢推定の処理コストの問題が残されている。

本研究では、拡張現実感において現実環境と仮想環境の位置合わせの精度を高めやすいという特徴を持つ画像からの位置・姿勢推定手法の中でも、広域で複雑な環境においてもデータベースを比較的 low コストに構築でき、高精度なカメラ位置・姿勢推定を実現可能なランドマークデータベースを用いる手法[12]に着目し、従来手法[12]で問題となっていたカメラ位置・姿勢推定処理の計算コストを低減することで実時間でのカメラ位置・姿勢推定を実現する。

### 3. ランドマークデータベースの構築

本章では、オフラインでのランドマークデータベース構築処理(図1のプロセス(A))について述べる。本研究では、4.で述べるオンラインでのカメラ位置・姿勢の推定において、推定に有効なランドマークを効率的に選択するために、従来手法[12]で用いられているデータベースに優先度情報を付加した新たなランドマークデータベースを用いる。また、初期フレームのカメラ位置・姿勢を推定するために、自然特徴点の固有スケールと SIFT 特徴ベクトルの情報を登録してお

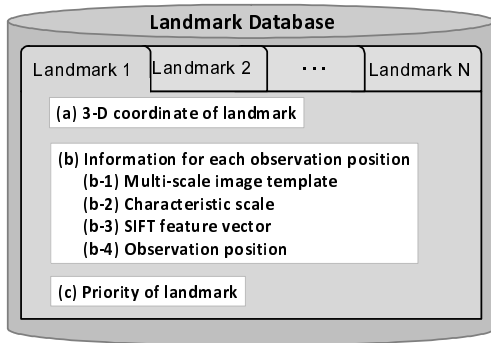


図 2 ランドマークデータベースの構成要素  
Fig. 2 Elements of landmark database

く。ここでは、ランドマークデータベースを構築するために、まず、全方位画像を用いた環境の三次元復元を行う (図 1 の (A-1))。次に、環境の三次元復元結果を用いてランドマーク情報を取得する (図 1 の (A-2))。以下では、各処理について詳しく述べる。

### 3.1 全方位画像を用いた環境の三次元復元 (A-1)

本研究では、三次元位置が既知の少数の基準点や、GPS などの外部指標を用いた structure-from-motion 法 [22], [23] を用いることで、広域で複雑な環境においても蓄積誤差の生じない環境の三次元復元を行う。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 [24] 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、structure-from-motion 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 [22] または GPS から得られる絶対位置 [23] を利用して、動画像全体での最適化処理を行うことで、カメラ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。このような手法を用いることで、広域で複雑な環境においても蓄積誤差なく比較的容易に環境の三次元復元を行うことができる。

### 3.2 ランドマーク情報の取得 (A-2)

本研究で用いるランドマークデータベースには図 2 に示すように、(a) 自然特徴点の三次元位置、(b) 撮影地点ごとの情報、(c) ランドマークの優先度の情報が保持される。撮影地点ごとの情報は、(b-1) 多重スケールの画像テンプレート、(b-2) 自然特徴点の固有スケール、(b-3) SIFT 特徴ベクトル、(b-4) ランドマーク撮影時の全方位カメラの位置から成る。以下ではそ

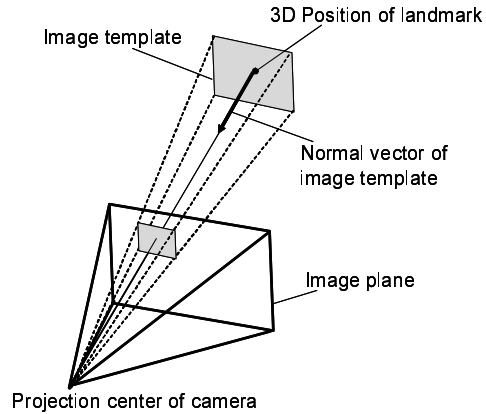


図 3 画像テンプレートの作成  
Fig. 3 Generation of image template.

れぞれの情報の取得方法について述べる。

(a) **ランドマークの三次元位置**: 4. で述べるオンラインでのカメラ位置・姿勢推定では、ランドマークの三次元位置と画像上の二次元座標を対応付けることによりカメラの位置・姿勢を推定するため、ランドマークの三次元位置情報が必要となる。ランドマークの三次元位置は 3.1 で述べた処理によって得られるものであり、環境に固定された世界座標系で保持される。

(b) **撮影地点ごとの情報**: 撮影地点によるランドマークの見え方の違いに対応するため、撮影地点ごとに多重スケールの画像テンプレートを作成し、データベースへ登録する。ここでは、図 3 に示すように、世界座標系において、カメラの投影中心とランドマークの三次元位置を結ぶ直線に対して垂直な面に画像上での自然特徴点周辺のパターンを投影することで、カメラの撮影姿勢に依存しない画像テンプレートを作成し登録する (b-1)。また、カメラの初期位置・姿勢を推定するために、Harris-Laplacian [25] によって算出される自然特徴点の固有スケール (b-2)、SIFT-descriptor [26] による SIFT 特徴ベクトル (b-3) を算出し、データベースへ登録する [17]。ここで、自然特徴点の固有スケールは、自然特徴点周辺のテクスチャの局所構造から決定されるものであり、これを用いることで、一定の空間領域に対応する画像の局所領域を切り出し、自然特徴点を正しく対応づけることができる。さらに、オンライン処理でのランドマーク選択処理 (B-3) において、撮影可能なランドマークを選択するために、ランドマーク撮影時の全方位カメラの位置をデータベースへ登録する (b-4)。

(c) **ランドマークの優先度**：4. で述べるオンラインでのカメラ位置・姿勢の推定において、正しく対応づく可能性の高いランドマークをデータベース中から効率的に選択するために、ランドマークに優先度情報を付加する。優先度は、過去に同じ場所でユーザによって撮影された入力に対する推定結果を利用して、各ランドマークがカメラ位置・姿勢推定に利用される確率を算出することで決定する。ただし、ランドマークデータベースの構築時点においては優先度を算出することはできないため、全てのランドマークに対して同一の優先度を設定しておく。4. で述べるオンライン処理によって、ユーザのカメラ位置・姿勢が推定された場合には、データベースからランドマーク  $i$  が選択された回数  $D_i$  とカメラ位置・姿勢推定の際に誤対応として排除されずに推定に用いられた回数  $E_i$  を用いて、ランドマーク  $i$  の優先度  $P_i$  を以下の式で更新する。

$$P_i = \frac{E_i}{D_i} \quad (1)$$

本研究では、データベースの管理者が、あらかじめデータベース構築時にいくつかの学習データを与えることで、一般ユーザが利用する前に優先度を更新しておくことを想定する。この優先度の更新処理については4.6 で詳しく述べる。

#### 4. 優先度情報を用いた実時間でのカメラ位置・姿勢推定

本章ではオンラインでのカメラ位置・姿勢推定処理(図1のプロセス(B))について述べる。本研究では、従来手法[12]と同等の推定精度を保ったまま、計算コストを削減するために、推定に有効なランドマークおよび自然特徴点を残しながら対応点候補数を削減する。具体的には、ランドマークの選択処理(B-3)において、優先度情報を用いることで、入力画像中の自然特徴点と正しく対応づく可能性の高いランドマークを効率的にデータベース中から取り出す。これにより、データベース中の全ランドマークを対等に扱っていた従来手法[12]に対して照合すべきランドマークの数を削減することができる。また、従来手法[12]で考慮されていなかった連続フレーム間でのカメラの動きをランドマークの追跡処理により推定することで、自然特徴点の選択処理(B-4)における自然特徴点の探索範囲を小領域に限定し、推定に有効な自然特徴点を残したまま照合すべき自然特徴点の数を削減する。以下では、オンラインでのカメラ位置・姿勢推定の各処理について

詳しく述べる。

##### 4.1 カメラの初期位置・姿勢の取得 (B-1)

本研究では、初期フレームのカメラ位置・姿勢は、ランドマークデータベースを用いた静止画像からのカメラ位置・姿勢推定手法[17]によって自動で取得する(B-1)。ここでは、まず、入力画像中の自然特徴点の固有スケール、SIFT 特徴ベクトルを算出した後、これらをデータベース中の全てのランドマークの SIFT 特徴ベクトルと照合することで自然特徴点とランドマークを対応付ける。次に、対応付けられた自然特徴点の固有スケールとランドマークの固有スケールの比を用いて、ランドマークからカメラまでのおおよその距離を算出する[17]。最後に、ランドマークの観測地点の整合性を利用して、対応付けられたランドマークから誤対応を排除した後に、PnP 問題[27]を解くことで6自由度のカメラ位置・姿勢を推定する。

上記の初期位置・姿勢推定処理では、データベース中の全てのランドマークと入力画像上の自然特徴点を照合するため、推定処理の完了に数十秒を要する。従って、本研究では初期位置・姿勢の推定が完了するまでの間カメラを動かさないものとする。

##### 4.2 暫定的なカメラ位置・姿勢の推定 (B-2)

ここでは、入力画像中のランドマークを追跡することにより、現フレームの暫定的なカメラ位置・姿勢を推定する(B-2)。これにより、後述する4.4で述べるランドマークと照合すべき自然特徴点の選択処理(B-4)において、ランドマークの探索範囲を小領域に限定することで照合すべき自然特徴点数を削減することができる。暫定的なカメラ位置・姿勢の推定では、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の探索において、連続フレーム間ではパターンの変形や明るさの変化は微小であることから、前フレームにおけるランドマークの位置を中心とする  $T \times T$  画素の画像をテンプレートとして用い、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  $W_1$  内で 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。ただし、テンプレートマッチングによるランドマークの追跡処理において、SSD の算出はウィンドウ  $W_1$  内の全ての画素に対して行うのではなく、Harris オペレータによって検出された自然特徴点のみを対象として行う。

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パラメータ  $\hat{\mathbf{M}}$  を推定する。暫定的なカメラパラメータ  $\hat{\mathbf{M}}$  を推定することによって、フレーム間でのカメラのおおよその動きを推定する。ただし、本研究では線形解法を用いてカメラパラメータの推定処理を高速化しているため、6点以上のランドマークが追跡されなかった場合には、暫定的なカメラ位置・姿勢を推定することができない[28]。また、6点以上のランドマークが追跡された場合でも、追跡されたランドマークの中に誤対応が半数以上含まれる場合には誤対応を排除することができず、暫定的なカメラ位置・姿勢を正しく推定できない。従って、本手法では追跡されたランドマーク数が6点未満または再投影誤差が一定の閾値以上となった場合に、暫定的なカメラ位置・姿勢の推定に失敗したものと判定する。この場合には、手法[14]や手法[20]と同様に、追跡処理に失敗する直前のカメラ位置・姿勢推定結果を用いて追跡処理の復帰を行う。ただし、後述する実験において用いる試作システムでは、この復帰処理についての実装および検証は行っていない。

### 4.3 ランドマークの選択 (B-3)

本研究では、まず、幾何学的な制約を用いて暫定的なカメラ位置・姿勢から観測可能なランドマークをデータベースより選択する。次に、優先度情報を用いて観測可能なランドマークの中からランドマークと正しく対応づく可能性の高いランドマークを選択する。

**4.3.1 幾何学的な制約によるランドマークの選択**  
ランドマークの選択処理 (B-3) では、まず処理 (B-2) で推定された暫定的なカメラ位置・姿勢とランドマークが撮影された全方位カメラの位置を用いて、以下の幾何学的な条件を満たす画像テンプレートをデータベース中より選択する。

- ・ ランドマークが入力画像の画角内に存在する
- ・ ランドマーク撮影時のカメラ位置とランドマークを結ぶ直線と、暫定的なカメラ位置とランドマークを結ぶ直線の成す角が閾値  $\theta$  以下 (図 4)

このような幾何学的な制約によって暫定的なカメラ位置・姿勢から観測可能なランドマークを選択することができる。

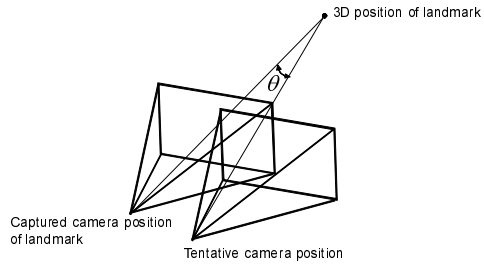


図 4 幾何学的な制約によるランドマークの選択  
Fig.4 Landmark selection based on geometric constraint.

### 4.3.2 優先度に基づくランドマークの選択

幾何学的な制約を用いて選択されたランドマークの候補中より推定に有効なランドマークを優先度に基づいて選択する。優先度に基づくランドマークの選択処理では、まず、入力画像を  $G$  個の格子に分割し、幾何学的な制約を用いて選択されたランドマークの候補中から各格子に対して最も優先度の高いランドマークを選択する。次に、画像全体で優先度  $P_i$  の高い順にランドマークを  $N_{proj}$  個選択する。ただし、 $N_{proj}$  は最大  $N_{prior}$  個とし、 $N_{proj} = N_{prior} - N'_{track}$  とする。ここで、 $N'_{track}$  は処理 (B-2) で誤対応として排除されず、暫定的なカメラ位置・姿勢の推定に用いられたランドマークの数である。優先度を用いてランドマークを選択することにより、自然特徴点と正しく対応づく可能性の高いランドマークを優先的に選択することができる。これにより、照合すべきランドマークの数を削減することができる。

### 4.4 自然特徴点の選択 (B-4)

ここでは、暫定的なカメラ位置・姿勢を用いて、処理 (B-3) で選択した優先度の高いランドマークと照合すべき入力画像中の自然特徴点を選択する。ランドマークと自然特徴点の選択処理では、まず処理 (B-2) で推定した暫定的なカメラパラメータ  $\hat{\mathbf{M}}$  を用いて、処理 (B-3) で選択されたランドマーク  $j(j = 1, 2, \dots, N)$  を以下の式により入力画像上へ投影する。

$$\begin{bmatrix} a_j u_j \\ a_j v_j \\ a_j \end{bmatrix} = \hat{\mathbf{M}} \begin{bmatrix} x_j \\ y_j \\ z_j \\ 1 \end{bmatrix} \quad (2)$$

ただし、 $(x_j, y_j, z_j)$  はランドマーク  $j$  の世界座標系での三次元位置、 $(u_j, v_j)$  はランドマーク  $j$  の画像上への投影位置、 $a_j$  はカメラ座標系でのランドマーク  $j$

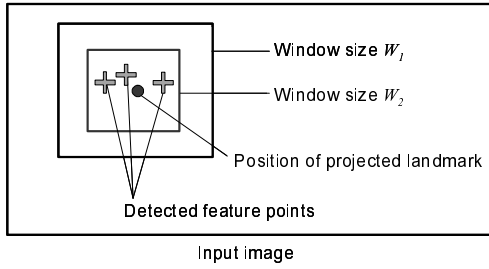


図5 ランドマークの対応点の探索範囲  
Fig.5 Search range of corresponding point.

の奥行きを表す. 次に, 図5に示すように, 投影したランドマーク  $j$  の画像上の位置  $(u_j, v_j)$  を中心とする一定ウィンドウ  $W_2$  内の自然特徴点を選択する (B-4). ただし, ここではウィンドウ  $W_2$  のサイズをウィンドウ  $W_1$  のサイズよりも小さく設定することで対応点候補となる自然特徴点の探索範囲を小領域に限定する.

#### 4.5 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・姿勢推定 (B-5), (B-6)

処理 (B-4) においてウィンドウ内で検出された全ての自然特徴点について, データベース構築時と同様に, カメラの撮影姿勢に依存しない画像テンプレートを作成し, 作成した画像テンプレートとランドマークの画像テンプレートを照合する. ここでは, データベース構築時とカメラ位置・姿勢推定時の明るさの変化に対応するために, 正規化相互相関法による照合を行い, 相関値が最大となる自然特徴点をランドマークと対応付ける (B-5).

次に, ランドマークと対応付けられた自然特徴点と, 処理 (B-2) で暫定的なカメラ位置・姿勢の推定に用い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. 現フレームのカメラ位置・姿勢は処理 (B-2) と同様に, LMedS 基準を用いて誤対応を排除した後に, 再投影誤差を最小化することによって推定する (B-6).

#### 4.6 推定結果のフィードバックによる優先度の更新 (B-7)

カメラの位置・姿勢推定の完了後, ランドマークが利用された頻度の情報をデータベースへフィードバックすることで, データベース中の各ランドマークの優先度を更新する (B-7). ここでは頻度の情報として, カメラ位置・姿勢推定の逐次処理において, ランドマーク  $i$  が処理 (B-3) でデータベース中から選択された回数  $D_{inew}$  と処理 (B-6) において LMedS 基準により

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回数  $E_{inew}$  を利用することでデータベース中のランドマークの優先度を更新する. ランドマーク  $i$  の優先度  $P_i$  は, ユーザより新たに取得した頻度情報  $(D_{inew}, E_{inew})$  と, 過去に同じ環境内でカメラ位置・姿勢の推定を行ったユーザより得られている頻度情報  $(D_{iold}, E_{iold})$  を用いて, 次式で更新する.

$$P_i = \frac{E_{iold} + E_{inew}}{D_{iold} + D_{inew}} \quad (3)$$

## 5. 計算コストの比較

本章では, 提案手法によるランドマークと自然特徴点の対応付けの理論的なコストの低減量について述べる. 従来手法における対応付け処理の計算コストは  $C_{prev} = NFA$  であるが, 本研究では照合される自然特徴点数  $F$  と対応付け処理に用いられるランドマーク数  $N$  を削減することで, 対応付けの計算コスト  $C_{prev}$  を低減する. 提案手法では, ランドマークの探索範囲を限定するために, 新たに暫定的なカメラパラメータを推定することが必要となるため, ランドマーク対応付けの総コスト  $C_{new}$  は,

$$C_{new} = C_{track} + C_{proj} \quad (4)$$

となる. ただし,  $C_{track}$ ,  $C_{proj}$  はそれぞれ, 暫定的なカメラパラメータの推定 (B-2), 暫定的なカメラパラメータを用いたランドマークの対応付け (B-5), に必要な計算コストである. ここで, 処理 (B-2) の計算コスト  $C_{track}$  は, 追跡に用いるランドマーク数を  $N_{track}$  個とすると,

$$C_{track} = N_{track}FB + E_{LMedS} \quad (5)$$

となる. ただし,  $E_{LMedS}$  は誤対応の排除とカメラ位置・姿勢の推定にかかる計算コストであり,  $B$  は追跡処理においてランドマークと自然特徴点の1組のペアに対して SSD を算出する計算コストである. このコストは, 幾何学的な変形や輝度値の正規化を必要とするランドマークの照合コスト  $A$  に対してはるかに小さく, また, 追跡に用いるランドマーク数  $N_{track}$  は  $N$  に対して小さいため, 提案手法で新たに導入した処理 (B-2) のオーバーヘッドは小さい.

処理 (B-4) では, 処理 (B-2) で推定された暫定的なカメラ位置・姿勢を用いて画像上の探索範囲を限定することで, 対応点候補として用いられる自然特徴点の

平均数を  $S_2/S_1$  ( $S_1, S_2$  はそれぞれウィンドウ  $W_1, W_2$  の面積を表す) に、また、優先度に基づくランドマークの選択処理と、処理 (B-2) で追跡されたランドマークを対応付けるランドマークから除外することにより、ランドマークの数  $N$  を  $(N_{prior} - N'_{track})/N$  に削減できる ( $N_{prior} \leq N$ )。これにより、処理 (B-5) の計算コスト  $C_{proj}$  は、

$$C_{proj} = (N_{prior} - N'_{track}) \frac{S_2}{S_1} FA \quad (6)$$

となる。

ここで、従来の対応付け処理に対する高速化の効果 (処理コストの比) は以下の式で算出できる。

$$\frac{C_{new}}{C_{prev}} = \frac{C_{track} + C_{proj}}{C_{prev}} \quad (7)$$

$$= \frac{C_{track}}{C_{prev}} + \frac{(N_{prior} - N'_{track}) S_2}{N S_1} \quad (8)$$

式 (8) において、第 1 項は処理 (B-2) によるオーバーヘッドを、第 2 項は処理 (B-5) における計算コストの削減効果を表している。実際には各処理の繰り返しにおいて一定のオーバーヘッドが生じるため、第 2 項の計算コストの削減効果は 6. で述べる実験結果と完全には一致しない。

## 6. 実験

提案手法の有効性を示すために、屋外環境において実際にランドマークデータベースを構築し、実環境での実験を行った。ここでは、まず、従来手法 [12] と提案手法のカメラ位置・姿勢推定の計算コストおよび推定精度の比較を行った。次に、提案手法を実アプリケーションに適用し、実アプリケーションへの応用の可能性を検証した。ただし、現時点において構築したシステム (CPU: Core 2 Extreme 2.93GHz, メモリ: 2GB) では、初期位置・姿勢の推定に 60 秒程度を必要とするため、後述する実験では、初期位置・姿勢の推定中は環境内に固定された三脚にカメラを設置しておき、初期位置・姿勢推定の完了後にカメラを移動させている。

本実験で用いるランドマークデータベースは、全方位型マルチカメラシステム (Point Grey Research 社 Ladybug) で屋外環境 (約 75m) を歩きながら撮影した全方位動画画像 (1260 フレーム) を入力とし、3. で述べた手法を用いることで構築した。図 6 に、データベースの構築に用いた全方位動画画像の一部を示す。ま



図 6 データベースの構築に用いた全方位動画画像の一部  
Fig.6 Sampled images taken by omnidirectional multi-camera system.

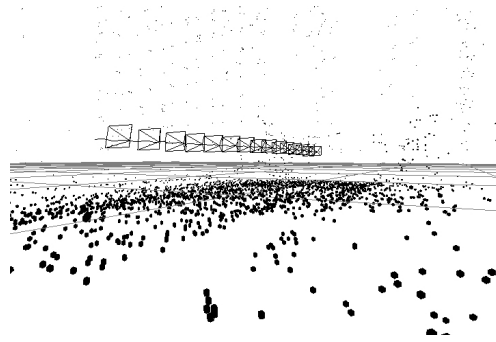


図 7 三次元復元の結果  
Fig.7 Result of 3-D reconstruction.

た、データベース構築時のカメラパスと復元されたランドマークの三次元位置を図 7 に示す。

データベースに登録されたランドマーク数は約 12400 個であり、各ランドマークに対して、平均 8ヶ所の異なる地点で撮影された画像テンプレートが登録された。本実験ではビデオカメラ (SONY DSR-PD-150) を用いて撮影した 4 本の動画画像 (解像度  $720 \times 480$  画素, プログレッシブスキャン, 15fps) に対してカメラの位置・姿勢推定を行った。そのうち、3 本の動画画像の推定結果を優先度設定のために用い、残りの 1 本の動画画像 (1000 フレーム) を評価用として用いた。カメラ位置・姿勢推定に用いた各パラメータを表 1 に示す。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あらかじめ Tsai の手法 [29] によって校正した。カメラ位置・姿勢推定精度の定量的な評価に用いるカメラ位置・姿勢の真値は、環境内の特徴点の



表 1 実験で用いたパラメータ  
Table 1 Parameters in experiment.

	Oe's method [12]	Proposed method
Image template size $T$ (pixel)	-	15
Window size $W_1$ (pixel)	-	$120 \times 60$
Window size $W_2$ (pixel)	$120 \times 60$	$20 \times 20$
Angle threshold $\theta$ (degree)	15	
Number of grids $G$	$72 \times 48$	
Training video	-	Three video sequences
Initial value of priorities	-	1/2

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。ただし、真値作成の際に、三次元位置を与えた特徴点の画像上の再投影誤差が平均 1.5 画素以上となったフレームや自然物が入力画像中の大半を占めるなどの理由で、トータルステーションによって計測できる特徴点が 6 点未満となったフレームについては真値が信頼できないため評価対象にしなかった。

### 6.1 処理速度向上効果の検証

提案手法の有効性を示すために、以下の 4 つの手法を比較し、計算コストの削減効果について検証を行った。

- (手法 A) 高速化処理無し（従来手法 [12]）
- (手法 B) 提案手法において、暫定的なカメラ位置・姿勢の推定による探索範囲の限定のみを行った場合（対応点候補となる自然特徴点数のみを削減）
- (手法 C) 提案手法において、優先度を用いたランドマークの選択のみを用いた場合（対応点候補となるランドマーク数のみを削減）
- (手法 D) 暫定的なカメラ位置・姿勢の推定による探索範囲の限定と優先度を用いたランドマークの選択を両方行った場合（提案手法：自然特徴点・ランドマークの双方の削減）

本実験では、まず、データベースから選択するランドマークの最大数  $N_{prior}$  を変化させ、各手法に対する

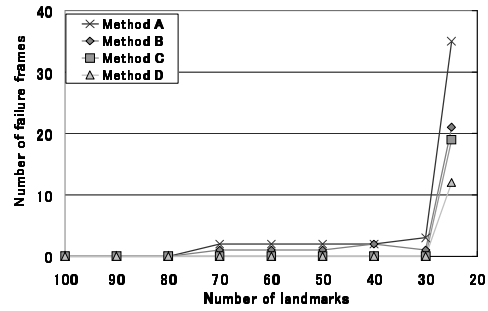


図 8 ランドマーク数と推定に失敗したフレーム数の関係  
Fig.8 Relation between number of landmarks and failure frames.

推定失敗率を算出することで、処理 (B-3) において選択するランドマーク数  $N_{prior}$  を決定した。ただし、優先度情報を用いない手法 A, B では、従来手法 [12] と同様に幾何学的な制約のみを用いて撮影可能なランドマークを  $N$  個選択した。次に、決定したランドマーク数  $N_{prior}$ ,  $N$  を用いて各手法の処理時間を比較し、計算コストの削減効果について検証を行った。

#### 6.1.1 推定に用いるランドマーク数の決定

ここでは、処理 (B-3) においてデータベースから選択するランドマークの最大数  $N_{prior}$ ,  $N$  を決定するために、 $N_{prior}$ ,  $N$  を変化させ、各手法に対する推定失敗率を比較した。本実験では、カメラ位置・姿勢の逐次推定処理において、データベース中のランドマークと入力画像中の自然特徴点の対応関係が 6 組以上得られなかった場合にカメラ位置・姿勢の推定に失敗したとみなした。また、推定に失敗した場合には、手動で失敗したフレームのカメラ位置・姿勢を与えることにより、カメラ位置・姿勢の逐次推定処理を再開し、最終フレーム（1000 フレーム目）まで推定を完了させた上で推定失敗の回数を算出した。

図 8 にデータベースから選択するランドマーク数と推定に失敗したフレーム数の関係を示す。優先度情報を用いない手法 A, B は、データベース中より選択するランドマーク数  $N$  が 70 個以下の場合にカメラ位置・姿勢の推定に失敗するフレームが発生するが、優先度を用いてデータベース中よりランドマークを選択する手法 C, D では、ランドマーク数を 30 個まで減らした場合においても最終フレームまで失敗せずにカメラ位置・姿勢を推定できている。このことから、優先度を用いることで、正しく対応づく可能性の高いランドマークを優先的に選択できていることが確認でき

表 2 1 フレーム当たりの処理時間の比較 (ミリ秒)  
Table 2 Comparison of processing time for one frame (ms).

Method	A	B	C	D
Process (B-2)	-	26	-	21
Process (B-3)	12	3	2	1
Process (B-4)	0	0	0	0
Process (B-5)	316	51	131	15
Process (B-6)	61	16	16	17
Overhead	4	4	4	5
Total cost	329	100	153	59

る。この結果に基づき、後述する処理速度および推定精度の比較実験では、データベースより選択するランドマーク数を手法 A, B では 80 個、手法 C, D では 30 個とする。

### 6.1.2 処理速度の比較

ここでは、PC(CPU: Core 2 Extreme 2.93GHz, メモリ: 2GB) を用いて、手法 A~D の処理時間と推定精度の比較を行った。カメラ位置・姿勢の逐次推定処理 (B-2)~(B-6) の各処理時間を表 2 に示す。

提案手法である手法 D では、全体の処理時間が平均 59ms となり、1 秒間に約 17 フレームの実時間でカメラ位置・姿勢推定を実現することができている。提案手法では、暫定的なカメラ位置・姿勢の推定 (B-2) に 21ms を要するが、これによってランドマークと入力画像中の自然特徴点の対応付け (B-4) の計算コストは従来手法 A の約 1/21 となった。なお、式 (8) を用いれば、提案手法によるランドマークと入力画像中の自然特徴点の対応付け処理 (B-5) の計算コストは、 $N = 80, N_{prior} - N_{track} \leq 30, S_2/S_1 = 1/18$  より、理論的には従来手法の 1/48 以下 (7ms 以下) となり実験結果 (15ms) と一致しない。これは、入力画像中に自然特徴点が一様に分布していないことや各処理の繰り返しにおいて一定のオーバーヘッドが生じるためであると考えられる。

### 6.1.3 推定精度の比較

各手法で推定されたカメラ位置・姿勢の精度および処理 (B-6) においてアウトライヤとして排除されたランドマークの発生率を表 3 に示す。有意水準を 5% と設定した t 検定によって手法 A と手法 B, C, D の平均位置誤差の間には有意な差が見られ、従来手法 A に対して、提案手法 B, C, D ではカメラ位置推定精度の向上が確認できる。これは、暫定的なカメラパラメータの推定による探索範囲の限定と優先度情報に基づくランドマークの選択を行うことによって、ランド

表 3 推定精度・誤対応の発生率の比較  
Table 3 Comparison of accuracy.

Method	A	B	C	D
Average position error (mm)	360	257	231	256
Std.dev. position error (mm)	528	137	204	181
Average posture error (deg.)	0.84	0.95	1.13	0.91
Std.dev. posture error (deg.)	0.71	1.20	1.16	0.91
Average re-projection error (pixel)	2.5	2.3	2.1	1.8
Rate of mismatched landmark (%)	32	23	27	20

マークと自然特徴点の対応付けにおける誤対応の発生率が低下したためであると考えられる。このことから、暫定的なカメラパラメータの推定による探索範囲の限定や優先度情報に基づくランドマークの選択は、ランドマークと入力画像中の自然特徴点の対応付け処理 (B-5) の計算コストを削減できるだけでなく、カメラ位置推定精度の向上にも貢献していることが確認できる。

### 6.2 実アプリケーションへの応用

提案手法を用いた拡張現実感技術の応用例として、提案手法をカーナビゲーションおよび Pre-visualization に適用し、実アプリケーションへの応用の可能性を検証した。図 9 は、提案手法によって推定されたカメラ位置・姿勢に基づいて、CG を入力画像上に重畳表示した結果である。図 9 中の○印はランドマークと対応付けられた自然特徴点の位置を示している。

カーナビゲーションへの応用では、事前に三次元位置が与えられた注釈情報を推定されたカメラパラメータを用いて撮影された画像中に重畳表示することでノート型 PC のディスプレイ上への AR 合成画像の提示を行った。提案手法 (位置誤差: 約 20cm, 更新周期: 1/17 秒) を用いることによって、車体の位置・姿勢を一般に用いられている GPS (位置誤差: 約 10m, 更新周期: 1 秒) よりも高精度かつ高頻度に推定することができた。しかし、提案手法では、車両が高速で移動した場合には、連続フレーム間でのパターンの変形が大きくなり、ランドマークの追跡を行うことが困難となった。一般的に、車には急激な速度変化や姿勢の変化が発生しないため、次フレームのカメラ位置・姿勢を予測することによるパターンの変形を考慮した追跡が可能である。実利用においては、この点での改良とともに、車速パルスや車載センサから得られる情報の併用などにより推定処理のロバスト性を向上させる必要がある。

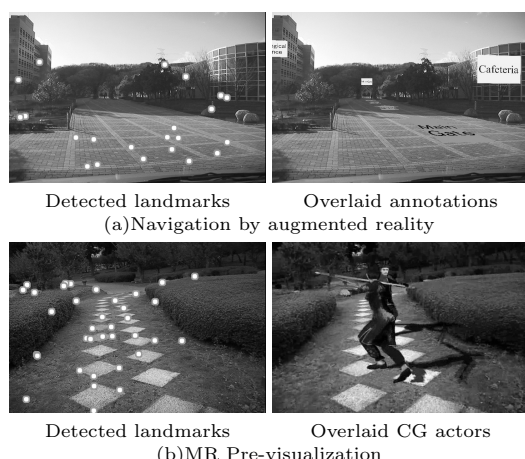


図9 実アプリケーションへの応用例  
Fig.9 Examples of applications.

図9(b)は、映画制作の初期段階において、シーンの見え方をシミュレートすることにより映画作成の支援を行うMR-PreVizへの応用例を示している。ここでは、事前に三次元位置や動きが与えられた役者のCGモデルをカメラで撮影した画像上にレンダリングすることでMR-PreViz映像の生成を行った。なお、利用するCGモデルの三次元位置・姿勢・スケール等は、ランドマークと同一座標系で、あらかじめ定義してある。試作システムでは、図9に示すような、自然物が多く存在する複雑な環境においても多くのランドマークが自然特徴点と対応付けられ、実時間でカメラの位置・姿勢を実現することができた。本手法では撮影からMR-PreViz映像の提示までの間に、理論上、最低でも1/17秒以上の遅延が発生するが、主観的にはMR-PreVizでの利用において問題になるレベルの遅延は感じられず、提案手法が有効に働くことを確認した。

## 7. まとめ

本論文では、自然特徴点の追跡による暫定的なカメラの位置・姿勢推定による探索範囲の限定を行うとともに優先度情報をランドマークへ付加することによって、ランドマークデータベースに基づくカメラ位置・姿勢推定の高速化を図る手法を提案した。実験により、実時間でのカメラ位置・姿勢推定が実現できることを確認した。また、ランドマークの対応点の探索範囲を小領域に限定したことおよび、優先度に基づいて正しく対応づく可能性の高いランドマークを優先的にデー

タベース中より選択したことにより、誤対応の発生率が低減し、カメラ位置・姿勢推定の精度が向上することも確認した。さらに、実験では提案手法を実アプリケーションに適用し、広域で自然物が多く存在するような複雑な環境においてもカメラ位置・姿勢の推定が可能であることを確認した。

今後の課題として、ユーザの移動を考慮したカメラ位置・姿勢の初期化手法の開発や、ジャイロなどのセンサを併用することによるロバスト性の向上などが挙げられる。

**謝辞** 本研究の一部は、科学技術振興機構(JST)の戦略的創造研究推進事業(CREST)「デジタルメディア作品の制作を支援する基盤技術」プログラム、総務省・戦略的情報通信研究開発推進制度(SCOPE)の支援による。

## 文 献

- [1] 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.
- [2] D. Wagner and D. Schmalstieg: "First steps towards handheld augmented reality," Proc. Int. Symp. on Wearable Computers, pp. 21–23, 2003.
- [3] 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.
- [4] M. Shin, B. S. Kim and J. Park: "AR storyboard: An augmented reality based storyboard authoring tool," Proc. Int. Symp. on Mixed and Augmented Reality, pp. 98–99, 2005.
- [5] G. Thomas: "Mixed reality techniques for TV and their application for on-set and pre-visualization in film production," DVD-ROM Proc. Int. Workshop on Mixed Reality Technology for Filmmaking, pp. 31–36, 2005.
- [6] 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.
- [7] 岩佐英彦, 栗飯原述宏, 横矢直和, 竹村治雄: "全方位画像を用いた記憶に基づく位置推定", 電子情報通信学会論文誌(D-II), Vol. J84-D-II, No. 2, pp. 310–320, 2001.
- [8] R. Cipolla, D. Robertson and B. Tordoff: "Image-based localization," Proc. Int. Conf. Virtual Systems and Multimedia, pp. 22–29, 2004.
- [9] V. Lepetit, L. Vacchetti, D. Thalmann and P. Fua: "Fully automated and stable registration for augmented reality applications," Proc. Int. Symp. on Mixed and Augmented Reality, pp. 93–102, 2003.

- [10] L. Vacchetti, V. Lepetit and P. Fua: "Combining edge and texture information for real-time accurate 3D camera tracking," Proc. Int. Symp. on Mixed and Augmented Reality, pp. 48–57, 2004.
- [11] I. Skrypnik and D. G. Lowe: "Scene modelling, recognition and tracking with invariant image features," Proc. Int. Symp. on Mixed and Augmented Reality, pp. 110–119, 2004.
- [12] 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.
- [13] A. Davison, Y. G. Cid and N. Kita: "Real-time 3D slam with wide-angle vision," Proc. IFAC Symp. on Intelligent Autonomous Vehicles, 2004.
- [14] B. Williams, G. Klein and I. Reid: "Real-time SLAM relocalisation," Proc. Int. Conf. on Computer Vision, 2007.
- [15] 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.
- [16] 武富貴史, 佐藤智和, 横矢直和: "位置に依存した固有空間上でのランドマークの照合によるカメラの位置・姿勢推定の高速化", 情報科学技術フォーラム (FIT) 一般講演論文集, pp. 95–96, 2007.
- [17] 薄充孝, 中川知香, 佐藤智和, 横矢直和: "ランドマークデータベースに基づく静止画像からのカメラ位置・姿勢推定", 日本バーチャルリアリティ学会論文誌, Vol. 13, No. 2, pp. 161–170, 2008.
- [18] Y. Kameda, T. Takemasa and Y. Ohta: "Outdoor see-through vision utilizing surveillance cameras," Proc. Int. Symp. on Mixed and Augmented Reality, pp. 151–160, 2004.
- [19] B. Jiang, U. Neumann and S. You: "A robust hybrid tracking system for outdoor augmented reality," Proc. IEEE Virtual Reality, pp. 3–10, 2004.
- [20] 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.
- [21] J. S.Beis and D.G.Lowe: "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," Proc. Conf. on Computer Vision and Pattern Recognition, pp. 1000–1006, 1997.
- [22] 佐藤智和, 池田聖, 横矢直和: "複数動画像からの全方位型マルチカメラシステムの位置・姿勢パラメータの推定", 電子情報通信学会論文誌 (D-II), Vol. J88-D-II, No. 2, pp. 347–357, 2005.
- [23] 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.
- [24] C. Harris and M. Stephens: "A combined corner and edge detector," Proc. Alvey Vision Conf., pp. 147–151, 1988.
- [25] K. Mikolajczyk and C. Schmid: "Scale & affine invariant interest point detectors," Int. Journal of Computer Vision, Vo. 60, No. 1, pp. 63–86, 2004.
- [26] D. G. Lowe: "Distinctive image features from scale-invariant keypoints," Int. Journal of Computer Vision, Vo. 60, No. 2, pp. 91–100, 2004.
- [27] R. Klette, K. Schluns and A. Koschan Eds: Computer Vision: Three-dimensional Data from Image, 1998.
- [28] 出口光一郎: "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.
- [29] R. Y. Tsai: "An efficient and accurate camera calibration technique for 3D machine vision," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 364–374, 1986.

(平成 xx 年 xx 月 xx 日受付)

**武富 貴史** (学生員)

2006 年佐世保高専・専攻科・電気電子工学専攻卒。2008 年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。現在、同大博士後期課程に在学中。情報処理学会, 日本バーチャルリアリティ学会, IEEE 各会員。

**佐藤 智和** (正員)

1999 年阪府大・工・情報工卒。2003 年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。現在、同大情報科学研究科助教。コンピュータビジョンの研究に従事。2001 年電子情報通信学会学術奨励賞受賞。博士 (工学)。情報処理学会, IEEE 各会員。

**横矢 直和** (正員: フェロー)

1974 年阪大・基礎工・情報工卒。1979 年同大大学院博士後期課程修了。同年電子技術総合技術研究所入所。以来、画像処理ソフトウェア, 画像データベース, コンピュータビジョンの研究に従事。1986~87 年マッギル大・知能機械研究センター客員教授。1992 年奈良先端科学技術大学院大学・情報科学センター教授。現在、同大情報科学研究科教授。1990 年・2007 年情報処理学会論文賞受賞。2005 年情報処理学会フェロー。工学博士。情報処理学会, 日本バーチャルリアリティ学会, 人工知能学会, 日本認知科学学会, 映像情報メディア学会, IEEE 各会員。

**Abstract** In the field of augmented reality, many kinds of image-based camera parameter estimation methods have been proposed to achieve accurate geometric registration. As one of these methods, a landmark based camera parameter estimation method has been studied. In this method, database construction cost is lower than other approaches and the method can be used in various environments. However, computational cost is expensive and thus it cannot work in real-time. To achieve real-time camera parameter estimation, the present study reduces the number of matching candidates by using priorities of landmarks that are determined from previously captured video sequences. To demonstrate the effectiveness of the proposed method, the computational cost is compared between the proposed and previous methods.

**Key words** landmark database, camera parameter estimation, natural features, augmented reality