

NAIST-IS-MT0351023

## 修士論文

画像特徴点による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

大江 統子

2005年3月10日

奈良先端科学技術大学院大学  
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に  
修士(工学) 授与の要件として提出した修士論文である。

大江 統子

審査委員： 横矢 直和 教授  
小笠原 司 教授  
山澤 一誠 助教授

# 画像特徴点による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\*

大江 統子

## 内容梗概

動画撮影時のカメラ位置・姿勢推定は拡張現実感やロボットナビゲーションなどの分野に応用されている。これらの分野では、カメラの絶対位置・姿勢情報が必要とされ、GPS や磁気センサを用いる手法や、カメラからの入力画像を用いる手法などが研究されている。センサ類を用いる手法では、一般的に GPS やジャイロセンサが用いられるが、カメラとセンサの同期をとることが難しいという問題や、センサによって利用できる環境が大きく異なるという問題がある。カメラからの入力画像を用いる手法としては、人工的なマーカあるいはシーン内の自然特徴点を用いる手法や物体の三次元モデルを用いる手法が挙げられる。しかし、前者では人工的なマーカを多数配置しなければ推定誤差が累積してしまうという問題があり、後者では対象となる物体の三次元モデルを事前に人手で作成しておく必要があるため、一般的な環境を対象とした場合や広範囲な環境での利用は難しいという問題がある。本研究では事前に撮影された動画画像から検出された自然特徴点をランドマークとして用いることで、センサ類や人工的なマーカを必要としない、広範囲な環境におけるカメラ位置・姿勢推定手法を提案する。本手法では事前に環境内を動画画像として撮影し、動画画像から検出された自然特徴点の画像テンプレートと、三次元復元によって得られる自然特徴点の三次元位置をランドマークデータベースに登録する。次に、これらのランドマークを入力画像内で逐次探索することでカメラ位置・姿勢の推定を行う。本研究では、まず全方位

---

\*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT0351023, 2005年3月10日.

カメラを用いて撮影した全方位動画像を用いてランドマークデータベースを構築し、次に屋内・屋外において撮影された動画像からのカメラ位置・姿勢の推定実験を通じて提案手法の有効性を検証する。

キーワード

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

# Camera Position and Posture Estimation Using Feature Landmark Database\*

Motoko Oe

## Abstract

Camera position and posture estimation from image sequences can be applied in the fields of augmented reality and robot navigation. In these fields, to obtain absolute position and posture of the camera, sensor-based and vision-based methods have been investigated. Sensor-based methods generally use GPS and gyroscope, and they have two problems. First, it is difficult to synchronize the camera and sensors. Second, environments where the system can be used are limited according to selection of sensors. On the other hand, vision-based methods using input images from the camera can be classified into two approaches: One is based on using markers and natural features and the other is based on using 3-D model of the target object in the image. The former methods have to allocate many artificial markers otherwise the estimation error will accumulate. The latter methods require a 3-D model of the target object that is modeled in advance. Therefore, it is difficult to use these methods in large and natural environments. In this research, we propose a new camera position and posture estimation method which does not require sensors and artificial markers by detecting natural feature points from image sequences taken beforehand and using them as landmarks. In this method, first, we semi-automatically reconstruct a partial 3-D model of natural feature points detected from omnidirectional image sequences, and use each feature point template and its 3-D position as a

---

\*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0351023, March 10, 2005.

landmark to create a landmark database. Secondly, camera position and posture are estimated in a sequential process by searching these landmarks from input images. In experiments, landmark databases have been created from omnidirectional image sequences captured by an omnidirectional multi-camera system, and camera position and posture have been successfully estimated in indoor and outdoor environments.

**Keywords:**

camera position and posture estimation, natural features, landmark database, augmented reality

# 目次

1. はじめに	1
2. カメラ位置・姿勢推定に関する従来研究と本研究の位置付け	3
2.1 センサを用いたカメラ位置・姿勢推定	3
2.2 画像を用いたカメラ位置・姿勢推定	4
2.2.1 人工的なマーカに基づくカメラ位置・姿勢推定	5
2.2.2 特徴点の追跡によるカメラ位置・姿勢推定手法	6
2.2.3 環境内の三次元モデルやデータベースなどの事前知識を用いた手法	8
2.3 本研究の位置付けと方針	11
2.4 提案手法の概要	12
3. ランドマークデータベースの構築	14
3.1 ランドマークデータベースの構成要素	14
3.2 全方位カメラからの三次元復元によるランドマーク情報の獲得	17
3.2.1 全方位動画画像からの環境の三次元復元	18
3.2.2 ランドマーク情報の獲得	18
4. 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	22
4.1 データベースに基づくカメラ位置・姿勢推定手法の概要	22
4.2 ランドマークの選択	22
4.3 入力画像とランドマークの対応付け	23
4.3.1 入力画像中の自然特徴点の検出	24
4.3.2 画像テンプレートとのマッチングによる対応付け	25
4.4 ランドマークに基づくカメラ位置・姿勢推定	26
5. 実験	28
5.1 屋外環境における実験	28
5.1.1 ランドマークデータベースの構築(屋外実験)	28

5.1.2	カメラパラメータ推定 (屋外実験) . . . . .	30
5.1.3	定量的な評価実験 (屋外実験) . . . . .	36
5.2	屋内環境における実験 . . . . .	41
5.2.1	ランドマークデータベースの構築 (屋内実験) . . . . .	41
5.2.2	カメラパラメータ推定 (屋内実験) . . . . .	41
5.2.3	定量的な評価実験 (屋内実験) . . . . .	41
5.3	考察 . . . . .	44
6.	まとめ . . . . .	46
	謝辞 . . . . .	47
	参考文献 . . . . .	48

## 目 次

1	Hallaway らの手法で用いる赤外線受光体 . . . . .	4
2	人工的なマーカに基づく手法 . . . . .	5
3	羽原らの手法で用いる人工的なマーカ . . . . .	6
4	Lepetit らの手法で用いる三次元モデルとキーフレーム . . . . .	9
5	Gordon らの手法による環境内の特徴点の三次元復元結果 . . . . .	10
6	提案手法のフローチャート . . . . .	13
7	ランドマークデータベースの構成要素 . . . . .	15
8	ランドマークとその画像テンプレート . . . . .	17
9	スケールの異なる画像テンプレート . . . . .	18
10	基本スケールの決定 . . . . .	21
11	ランドマークと対応付ける自然特徴点の検出 . . . . .	24
12	全方位型マルチカメラシステム Ladybug と撮影された全方位画像	29
13	推定されたカメラパスと自然特徴点の三次元位置 (屋外実験) . . .	29
14	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配置 (屋外実験) . . . . .	31
15	カメラ位置・姿勢推定に用いられたランドマーク数 (屋外実験) . .	33
16	645 フレーム目でカメラ位置・姿勢推定に用いられたランドマーク (屋外実験) . . . . .	33
17	カメラ位置・姿勢推定に用いられたランドマーク (屋外実験) . . .	34
18	カメラ位置・姿勢推定結果を用いた match move(屋外実験) . . . .	35
19	屋外実験において推定されたカメラパスと正解データ . . . . .	36
20	屋外実験において推定されたカメラパスと正解データ (一部拡大) .	36
21	それぞれのシーンのカメラパスと正解データ . . . . .	38
22	シーン 1 のカメラパスと正解データ . . . . .	39
23	シーン 2 のカメラパスと正解データ . . . . .	39
24	シーン 3 のカメラパスと正解データ . . . . .	39
25	シーン 4 のカメラパスと正解データ . . . . .	40
26	カメラ位置・姿勢推定に用いられたランドマーク (屋内実験) . . .	42

27	カメラ位置・姿勢推定結果を用いた match move(屋内実験) . . . . .	43
28	屋内実験において推定されたカメラパスと正解データ . . . . .	44

## 表 目 次

1	カメラ位置・姿勢推定手法のメリット・デメリット . . . . .	11
2	ランドマークデータベース作成の条件 . . . . .	30
3	ランドマークデータベースを用いたカメラ位置・姿勢推定の条件 . . . . .	32

## 1. はじめに

動画撮影時のカメラ位置・姿勢推定は拡張現実感における現実環境と仮想環境の位置合わせや、ロボットナビゲーションにおけるロボットの自己位置同定などの分野に応用されている。これらの分野では、カメラの絶対位置・姿勢が必要とされ、GPS や磁気センサなどのセンサ類を用いる手法 [1][6][7][8] や、カメラからの入力画像を用いる手法 [9][13][15][19][25][26][28] などが研究されている。

センサ類を用いる手法では、一般的に GPS やジャイロセンサが用いられており、広範囲な屋外環境で利用可能であるが、カメラとセンサの同期をとることが難しいという問題がある。また、センサの組み合わせによって利用できる環境が大きく異なるために、様々な環境でカメラ位置・姿勢を推定するためには、多数のセンサを組み合わせなければならず、システムが複雑になるという問題がある。一方、カメラからの入力画像を用いる手法としては、ARToolkit [9] などの人工的なマーカを用いる手法 [10][11]、入力画像上の自然特徴点を用いる手法 [13][15]、それらの手法のハイブリッド [20] などが従来研究されてきた。これらの手法では、マーカパターンの形状に基づいてカメラ位置・姿勢の推定を行うため、人工的なマーカを広範囲に多数配置する必要があり、マーカの設置に多くの人的コストがかかるという問題や、マーカによって景観が損なわれるという問題がある。このような問題を解決するために、環境内の物体の三次元モデルや自然特徴点などを事前知識として獲得し、人工的なマーカの代わりに用いる手法が提案されている [25][26]。しかし、三次元モデルを用いる手法 [25] では、対象となる物体の三次元モデルを事前に人手で作成しておく必要があるため、自然物を対象とした場合や広範囲な環境での利用は難しいという問題がある。また、自然特徴点を用いる手法 [26] では広範囲な環境への適用を考えた場合、三次元復元の推定誤差が累積するという問題があり、いずれも小物体、小領域への適用にとどまっている。

そこで、本研究では事前に撮影された全方位動画画像から検出された自然特徴点をランドマークとして用いることで、センサ類や人工的なマーカを必要としない、広範囲な環境におけるカメラ位置・姿勢推定手法を提案する。本手法では事前に環境内を動画画像として撮影し、動画画像から検出された自然特徴点の画像テンプレートと、shape-from-motion に基づく三次元復元によって得られる自然特徴

点の三次元位置をランドマークデータベースに登録する。次に、これらのランドマークを入力画像内で逐次探索することでカメラ位置・姿勢の推定を行う。本研究では、まず全方位カメラを用いて撮影した全方位動画像を用いてランドマークデータベースを構築し、次に屋内・屋外において撮影された動画像からのカメラ位置・姿勢の推定実験を通じて提案手法の有効性を検証する。

以降、2章ではカメラ位置・姿勢推定に関する従来研究と本研究の位置づけ・方針について述べる。3章では、オフライン処理で作成するランドマークデータベースの構成要素とランドマーク情報の獲得方法について述べる。4章では、3章で構築したランドマークデータベースを用いた逐次的なカメラ位置・姿勢推定手法について述べる。5章では、実際に全方位カメラによって撮影された屋外全方位動画像を用いたランドマークデータベース構築とカメラパラメータ推定実験について報告する。最後に、6章でまとめと今後の課題について述べる。

## 2. カメラ位置・姿勢推定に関する従来研究と本研究の位置付け

本章では，本研究に関連する従来研究と本研究の位置付けについて述べる．カメラ位置・姿勢推定に関する従来研究は，GPS や磁気センサなどの計測機器をビデオカメラと併用して用いることで，ビデオカメラの位置・姿勢情報を取得する手法と，ビデオカメラの入力画像から得られる情報を用いてカメラの動きを推定することで，カメラ位置・姿勢を推定する手法に大別される．

以下，2.1 節ではセンサを用いたカメラ位置・姿勢推定手法，2.2 節では画像を用いたカメラ位置・姿勢推定手法について詳述する．2.3 節では本研究の位置付けと方針，2.4 節では提案手法の概要について述べる．

### 2.1 センサを用いたカメラ位置・姿勢推定

センサを用いたカメラ位置・姿勢推定では，用いるセンサによって，屋外で利用可能な手法 [1][2][3][4][5] と屋内で利用可能な手法 [6][7][8] に分けられる．Feinerらの提案するウェアラブル型拡張現実感システム MARS[1] や Gleueらの提案するウェアラブル型拡張現実感システム ARCHEOGUIDE[2] のようなウェアラブル型拡張現実感システムでは，リアルタイムキネマティック GPS またはディファレンシャルGPSを用いてユーザの位置を計測している．また，神原ら [3] はRTK-GPSと小型慣性航法装置を用いてカメラ位置・姿勢情報を取得することで，屋外の広範囲で高精度な位置合わせが可能な拡張現実感システムを提案している．このように，広範囲な屋外でのカメラ位置情報取得にはGPSを用いることが多いが，GPSは屋内での利用が困難である．屋内で用いる手法としては，磁気センサや赤外線ビーコンを利用した手法 [6][7][8] が存在する．天目らの提案するシステム WARS[6] では，環境中に設置した赤外線マーカに位置IDを対応付け，その信号を受信することにより絶対位置を取得する．さらに，コンパスと加速度センサを組み合わせることにより，マーカ位置からの相対的な移動位置を推定し，ジャイロセンサで姿勢を計測している．また，Hallawayら [7] は，複数の赤外線受光体を図1のように同一平面上に配置し，それを装着したユーザの二次元位置と方位



図 1 Hallaway らの手法で用いる赤外線受光体 [7]

角度の推定を行っている。Welch らの HiBall[8] では、6 面のフォトダイオードからなるセンサを用いて、環境に設置した多数の赤外線 LED を、その発光と同期して撮影することによって、ユーザの位置・姿勢の推定を可能としている。しかし、一般的にこれらのデバイスは位置検出可能な範囲が狭い、事前に環境内にセンサ類を設置しなければならないという問題が存在する。

このように、センサを用いた手法では、利用するセンサの組み合わせによって利用できる環境が大きく異なる。従って、様々な環境でカメラ位置・姿勢を推定するためには、多数のセンサを組み合わせなければならず、カメラとセンサの同期をとることが難しいという問題や、システムが複雑になるという問題がある。

## 2.2 画像を用いたカメラ位置・姿勢推定

カメラからの入力画像を用いる手法として、ARToolkit[9] などの人工的なマーカを用いる手法や、現実環境にもともと存在する特徴的な点や形状を利用する手法などが挙げられる。従来は人工的なマーカを用いる手法が主流であったが、近年では環境内の自然特徴を用いる手法や人工的なマーカとの組み合わせによるハ

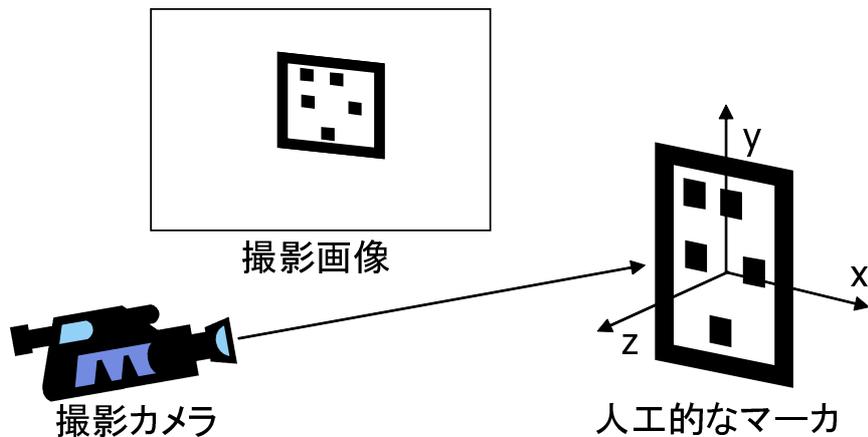


図 2 人工的なマーカに基づく手法

イブリッドな手法も多く提案されはじめている．また，環境内の物体の三次元モデルをあらかじめ取得しておくことで人工的なマーカの代わりに用いる手法なども提案されている．以下ではそれぞれの手法と原理についてそれぞれ解説する．

### 2.2.1 人工的なマーカに基づくカメラ位置・姿勢推定

人工的なマーカに基づくカメラ位置・姿勢推定手法としては，ARToolkit[9]などの形状やパターンが既知のマーカを利用する手法[10][11][12]が挙げられる．これらの手法では，図2のように，三次元位置が既知であるマーカを撮影し，カメラの入力画像を用いてマーカ上の点の三次元座標と画像上の二次元座標との対応関係を求めることでカメラの位置・姿勢を決定する．

羽原ら[10]は，図3のように環境に多数の正方形の画像マーカを配置し，ユーザが装着したヘルメットに搭載されたカメラでマーカをとらえることでユーザ位置・姿勢の推定を行う手法を提案している．この手法では，一部のマーカのみ事前に位置情報を測定し，その他のマーカの位置情報は，周囲のマーカの位置情報を利用して相対的に求めることで，マーカの三次元位置計測の手間を軽減している．また，Naimarkら[11]は，円形の画像マーカと加速度センサを組み合わせることで，安定して位置同定可能なシステムを開発している．

これらの手法では，人工的な画像マーカに基づいて位置・姿勢の推定を行うた



図 3 羽原らの手法で用いる人工的なマーカ [10]

め、マーカを広範囲に多数配置する必要があり、マーカの設置に多くの人的コストがかかるという問題や、マーカによって景観が損なわれるという問題がある。

### 2.2.2 特徴点の追跡によるカメラ位置・姿勢推定手法

自然特徴点の追跡による手法では、移動しながら撮影される動画像上の特徴点を複数のフレーム間で対応付けることで、特徴点の三次元位置とカメラの位置・姿勢を復元する。このような手法には因子分解法 [13][14] や射影的な復元を利用する手法 [15][16][17][18] が挙げられる。また、自然特徴点と人工的なマーカを併用した手法 [19][20][21] も存在する。

- 自然特徴点を用いた手法

自然特徴点を用いた手法として、因子分解法や射影的な復元を利用する手法が挙げられる。因子分解法は、カメラモデルを線形近似し、線形解法によって撮影対象の形状とカメラの運動パラメータを同時に求める手法 [13][14] で

ある．この手法は安定かつ高速に対象の復元を行うことができるが，カメラモデルを線形近似するためにカメラパラメータの復元精度に問題がある．さらに，動画像の全てのフレームに撮影されている特徴点を用いることを前提としているため，特徴点が全てのフレームに渡って連続して観測できない一般的な動画像からの三次元復元は困難である．近年，不完全なデータに同手法を適用する手法 [22] や透視投影モデルへの適用を試みる手法 [23] が提案されているが，これらの問題を同時に解決することは難しく，複雑な屋外環境への手法の適用は困難である．

また，射影的な復元を利用する手法 [15][16][17][18] では，各フレームにおいて，逐次的に射影的なカメラパラメータを復元し，射影座標系からユークリッド座標系への変換を行うことでカメラパラメータを復元する．しかし，射影的な復元では，多くの自由度を残したまま逐次的な三次元復元を行うため，特徴点の誤対応による累積的な推定誤差が大きくなるという問題がある．また，これらの手法で得られるのは，カメラの絶対位置・姿勢ではなく相対位置・姿勢であり，スケール情報も失われるため，ユーザの絶対位置が必要となるナビゲーションなどには用いることはできない．

- 自然特徴と人工的なマーカを併用した手法

Neumann ら [19] や Davison ら [20] は，人工的なマーカや環境内の自然特徴点を追跡することによってカメラ位置・姿勢推定を行う手法を提案している．しかし，このような手法では，前フレームからの相対的な動きを推定していくため，人工的なマーカが入力画像上に写っていない状態が長時間続くと，推定誤差が累積してしまい，長時間の運用や広範囲な環境への適用が困難であるという問題がある．

佐藤ら [21] は三次元位置が既知の少数の基準点と多数の自然特徴点を全方位動画像中で自動追跡し，動画像全体での最適化処理を行うことでカメラ位置・姿勢を推定する手法を提案している．この shape-from-motion 手法では，三次元位置が既知の基準点を用い，動画像全体での投影誤差の最小化を行うことで，複雑で広範囲な環境を撮影した場合でも累積的な推定誤差を含まないカメラ位置・姿勢推定が可能である．しかし，動画像全体を用

いてカメラパスを最適化するため，逐次処理が必要な AR やロボットナビゲーションに直接用いることはできない．

### 2.2.3 環境内の三次元モデルやデータベースなどの事前知識を用いた手法

人工的なマーカや入力画像から得られた自然特徴を用いる代わりに，環境内に存在する物体の三次元モデル [24][25] や特徴点データベース [26]，画像データベース [27][28][29] などを事前知識として用いる手法が提案されている．これらの手法では，環境内の情報をあらかじめ取得しておき，それらと入力画像を照合することによってカメラ位置・姿勢推定を行う．

- 環境内の三次元モデルを用いた手法

Comport ら [24] は，現実環境に存在する物体の三次元形状を用いてカメラ位置・姿勢を推定する手法を提案している．この手法では，現実環境にある線，円，円柱，球等の単純な形状を多数利用しているため，部分的なオクルージョンが発生しても位置合わせが可能であるといった特徴がある．また，Lepetit ら [25] は，図 4 で示すような物体や環境の三次元モデルと複数のキーフレーム画像から得られる特徴点情報を用いて三次元モデルと入力画像の対応関係を求めることで，カメラ位置・姿勢を推定する手法を提案している．

これらの手法では，三次元モデルと前フレームのカメラ位置・姿勢情報を用いることで，物体の見え方の変化やオクルージョンにロバストなカメラ位置・姿勢推定が可能である．しかし，対象となる物体の三次元モデルを事前に人手で作成しておく必要があるため，自然物を対象とした場合や広範囲な環境での利用が難しいという問題がある．

- 環境内の特徴点データベースを用いた手法

Gordon ら [26] は小物体を撮影した複数枚の画像から，セルフキャリブレーションによって図 5 に示すような特徴点の三次元復元を行い，スケールや回転に不変な SIFT 特徴量を用いて入力画像との対応付けを行うことで，カメラ位置・姿勢を推定する手法を提案している．この手法では特別な事前

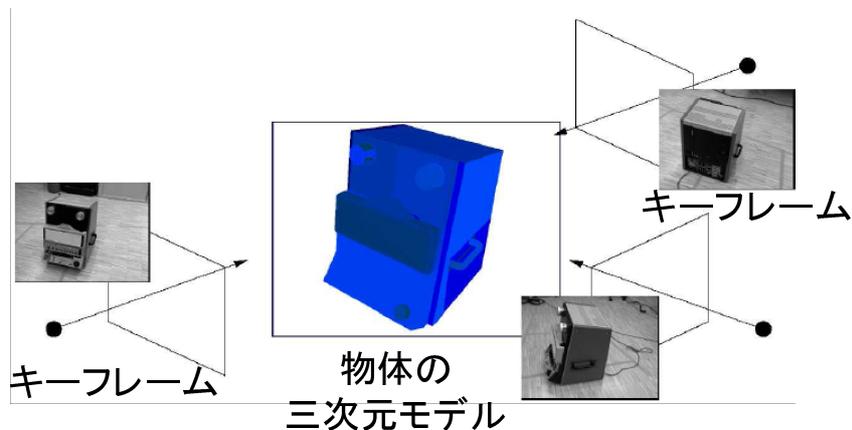


図 4 Lepetit らの手法で用いる三次元モデルとキーフレーム [25]

知識を必要としないが，広範囲な環境への適用を考えた場合，三次元復元の推定誤差が累積するという問題があり，小物体，小領域への適用にとどまっている．

- 画像データベースを用いた手法

興相らの提案するウェアラブル型拡張現実感システム Weavy[27] では，あらかじめ撮影しておいたパノラマ画像群とユーザが装着したカメラからの画像との位置合わせ結果からユーザの絶対的な位置・姿勢を推定している．しかし，パノラマ画像群のみを用いてカメラ位置・姿勢推定を行うには，密な間隔(室内オフィス環境で概ね1枚/ $1m^2$ )でパノラマ画像を撮影しておく必要がある．そのため，このシステムでは加速度計，およびジャイロセンサからの情報を組み合わせることにより，ユーザの位置・姿勢を推定している．

Cipolla ら [28] は，環境内の画像データベースを事前に作成し，それらと入力画像を対応付けることで，一枚の画像からのカメラ位置・姿勢推定を行う手法を提案している．この手法では，画像中の縦・横方向の直線とその消失点を用いて入力画像に写っている建物が画面に対して垂直に見えるように画像の垂直化(レクティフィケーション)を行い，データベース画像と入力画像の対応関係を求めることでカメラ位置・姿勢を推定する．しかし，

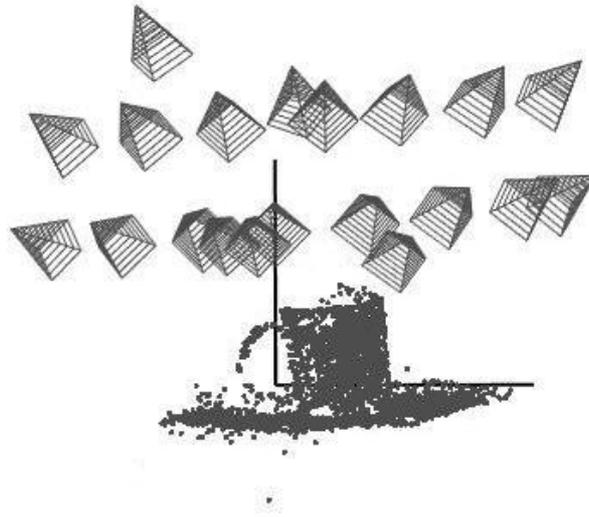


図 5 Gordon らの手法による環境内の特徴点の三次元復元結果 [26]

この手法は平行直線が画像内に複数存在していることを前提としているため、利用可能な範囲が限定されるという問題がある。

亀田ら [29] は、屋外に設置した複数台の定点カメラの映像から検出されたランドマークを数点、入力画像と対応付けることで、物理センサで得られたユーザの位置・姿勢を補正する手法を提案している。この手法では、定点カメラの映像を基に、ランドマークテクスチャをオンラインで随時更新することで安定したランドマーク検出を実現する。また、環境内の三次元モデルを用いることで、物体の陰などによって死角となる現実世界の状況を提示できる。しかし、常時環境内に複数台の定点カメラを必要とすることから設置コストがかかるという問題や、処理に時間がかかる、物理センサとの併用でしか用いることができないという問題がある。

表 1 カメラ位置・姿勢推定手法のメリット・デメリット

手法	メリット	デメリット
センサ類を用いる手法	広範囲な屋外で使用可能	センサによって利用環境が制限される システムが複雑になる
人工的なマーカを用いる手法	誤差が累積しない	マーカの設置コストがかかる 景観を損ねる
自然特徴点の追跡による手法	カメラ以外を必要としない	推定誤差が累積する
環境の事前知識を用いる手法	カメラ以外を必要としない 誤差が累積しない	事前準備に手間がかかる

### 2.3 本研究の位置付けと方針

前節までに概観したのように、従来手法は手法ごとに適用可能な条件や範囲が異なる。表 1 にそれぞれの基本的な手法のメリット・デメリットを示す。センサを用いた手法では、複数のセンサを組み合わせる必要があり、システムが複雑になってしまう。人工的なマーカを用いる手法では、環境内にマーカを多数配置する必要があるために人的コストがかかる、景観を損ねるという問題がある。逆に、環境内の自然特徴のみを用いた手法では、カメラ位置・姿勢の推定誤差が累積してしまうという問題がある。あらかじめ環境内の物体の三次元モデルや特徴点データベースなどの事前知識を獲得しておくことで、人工的なマーカを用いずに累積誤差のないカメラ位置・姿勢を推定する手法は存在するが、現時点においては、小規模な環境を対象としたものが多く、広範囲な屋内外で適用は難しい。

そこで、本研究では事前に撮影された全方位動画画像から環境の粗な三次元復元を行い、動画画像から検出された自然特徴点をランドマークとして用いることで、広範囲な環境におけるカメラ位置・姿勢推定を実現する手法を提案する [30]。提

案手法では2.2.3節で述べた特徴点データベースを用いた手法と同様、環境内の特徴点情報をあらかじめ獲得し、それらを入力画像と対応付けるというアプローチをとる。ただし、2.2.2節で述べた誤差の累積しない三次元復元手法をデータベースの作成に用いることで、広範囲におけるカメラの位置・姿勢推定を可能にする。本研究ではまず、自然特徴点とマーカを用いたオフライン三次元復元手法 [21] を用いることで全方位動画画像から累積誤差のないカメラパスを推定し、環境内の多数の自然特徴点の三次元位置を自動で推定する。次に、動画画像から検出された自然特徴点の画像テンプレートと、三次元復元によって得られる自然特徴点の三次元位置をランドマーク情報として登録する。このようにして、事前にランドマークデータベースを構築し、それらのランドマークを三次元位置が既知のマーカの代わりに用いることで、環境内に人工的なマーカを設置することなく、累積誤差を含まないカメラ位置・姿勢を推定する。

## 2.4 提案手法の概要

提案手法は図6に示すように、オフライン処理によるランドマークデータベースの構築(ステージ1)、ランドマークデータベースを用いたカメラ位置・姿勢の逐次推定(ステージ2)、の二つのステージによって構成される。ステージ1では、事前に全方位カメラを用いて環境内を全方位動画画像として撮影し、環境の粗な三次元復元を行う。そして、動画画像から検出された自然特徴点の画像テンプレートと、三次元復元によって得られる自然特徴点の三次元位置を、ランドマーク情報としてランドマークデータベースに登録する。ステージ2では、これら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を行うことで、累積誤差を含まないカメラ位置・姿勢を逐次的に推定する。ステージ2はオンライン逐次処理の枠組みになっており、ビデオレートに近いリアルタイム処理を実現できれば、ARにおける幾何的位置合わせやロボットナビゲーションに利用できる。なお、本研究では全方位カメラおよびビデオカメラの内部パラメータはあらかじめ較正済みであるとする。

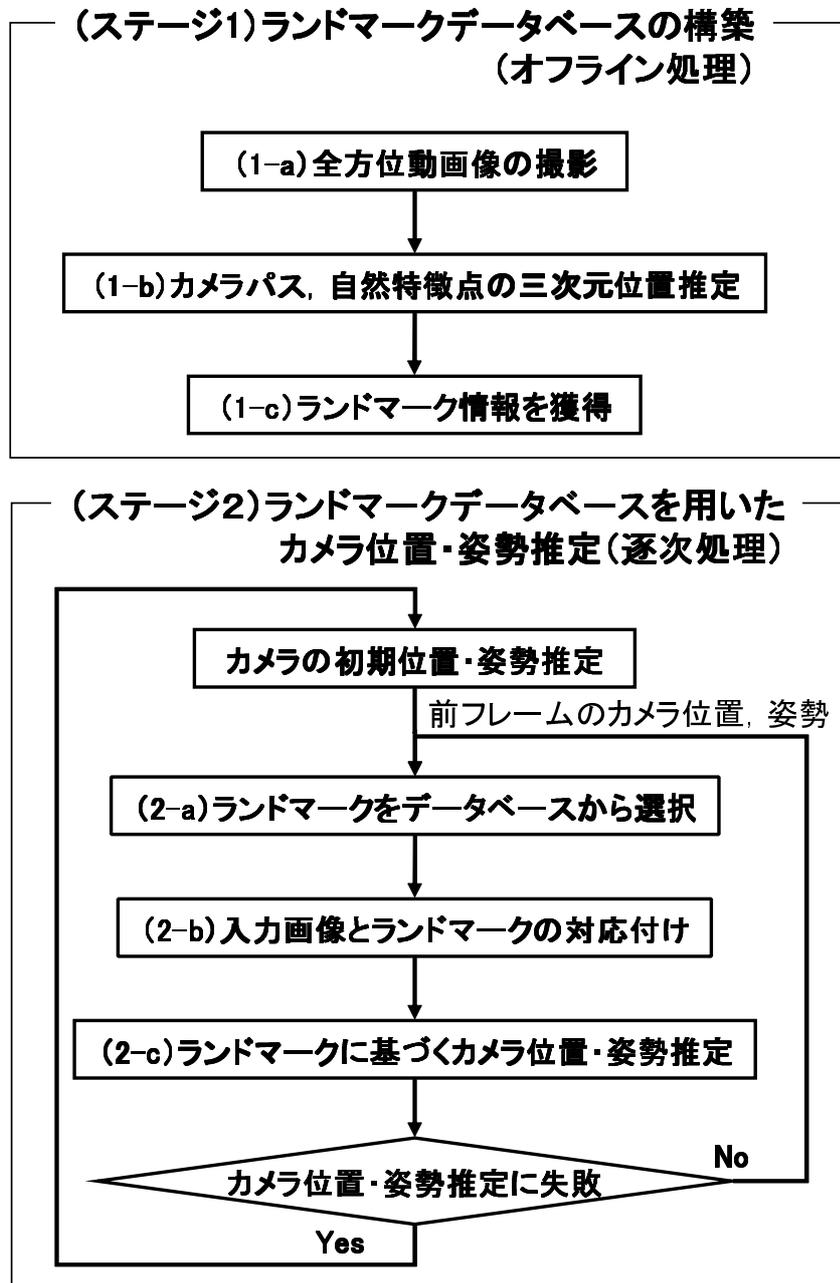


図 6 提案手法のフローチャート

### 3. ランドマークデータベースの構築

本章では、次章で述べるカメラ位置・姿勢の推定に必要なランドマークデータベースの構築方法(ステージ1)について詳述する。本手法では、環境内を撮影した動画像から自然特徴点を検出し、ランドマークとして用いる。そこで、本ステージでは、図6に示すように、まず多数のランドマーク情報を効率的に取得するために、広範囲を一度に撮影可能な全方位カメラを用いて環境内を移動しながら撮影する(1-a)。次に、全方位動画像上の自然特徴点追跡による三次元復元を行い、全方位動画像から検出された自然特徴点の三次元位置と動画像のカメラパスを推定する。これにより、自然特徴点の三次元位置、自然特徴点撮影時のカメラ位置・姿勢情報を取得する(1-b)。最後に、撮影された全方位動画像とフレームごとのカメラ位置・姿勢、自然特徴点の三次元位置情報に基づき、それぞれの自然特徴点ごとにランドマーク情報を自動的に取得し、ランドマークデータベースを構築する(1-c)。以下では、まずランドマークデータベースの構成要素について述べ、次に各ランドマークの情報を取得するための手順について述べる。

#### 3.1 ランドマークデータベースの構成要素

ランドマークデータベースは図7のように、複数のランドマークによって構成される。これらのランドマークは次章で述べるカメラ位置・姿勢推定処理において、入力画像中の特徴点と対応付けられ、カメラ位置・姿勢の推定(ステージ2)で用いられる。それぞれのランドマークは、1個の(1)ランドマークの三次元位置、複数の(2)撮影地点ごとの情報、から構成される。また、撮影地点ごとの情報は、(A)撮影時の全方位カメラの位置、(B)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、(C)画像テンプレートの法線ベクトル、(D)画像テンプレートの基本スケールから成る。以下にそれぞれの要素について詳述する。

##### (1) ランドマークの三次元位置

本研究では次章で述べるステージ2において、ランドマークの三次元位置と画像上の二次元座標の対応からカメラパラメータを推定するため、ラン

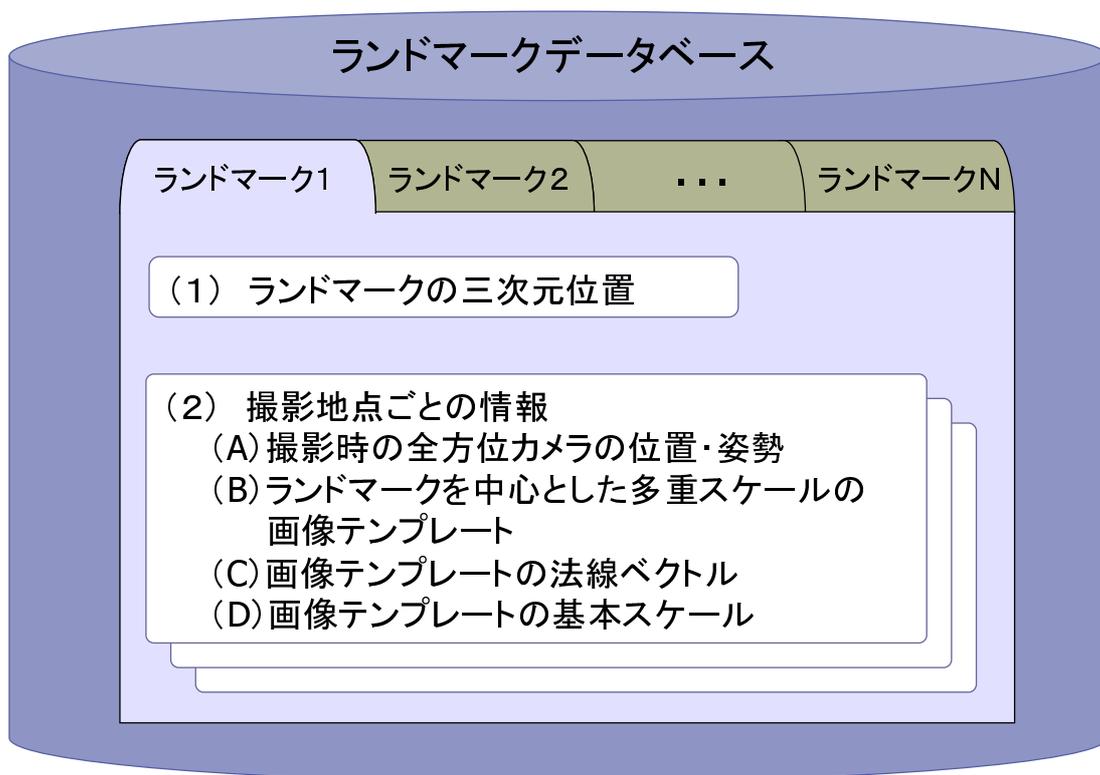


図 7 ランドマークデータベースの構成要素

ドマークの三次元位置が必要となる。ランドマークの三次元位置は，次節で述べる環境の三次元復元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ている。世界座標系は X 軸，Y 軸が実環境における地面に対して水平，Z 軸が地面に対して垂直な座標系であるとする。

## (2) 撮影地点ごとの情報

ランドマークと入力画像を対応付けるために必要な情報として，撮影地点ごとの情報を保持する。ランドマークの見え方は撮影地点によって異なるため，単一方向から撮影された画像のみでは高精度な対応付けは難しい。そこで，見え方の変化に対応するため，各ランドマークに対して複数の撮影地点ごとの情報を保持する。

### (A) 撮影時の全方位カメラの位置

入力画像と対応付けるランドマークを選択する際に用いる。全方位カメラの位置は世界座標系で保持されている。

### (B) 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

多重スケール画像テンプレートはステージ 2 において入力画像の中からのランドマークの探索に用いられる。これは，図 8 のように，世界座標系においてカメラの投影中心とランドマークを結ぶ直線に対して垂直な面に撮影画像を投影することで作成する。これにより，撮影カメラ姿勢やレンズ歪みの影響を受けない画像テンプレートを作成する。本研究では，図 9 のように，各撮影地点ごとに基本スケール(入力画像と画像テンプレートの解像度がほぼ等しくなるような画像テンプレートの実空間中での大きさ)とその 2 倍，4 倍のスケールで作成された多重スケールの画像テンプレートを保持する。また，作成する画像テンプレートの画素数は常に一定とし，スケールによって解像度が異なる画像テンプレートを作成する。

### (C) 画像テンプレートの法線ベクトル

図 8 に示すような，画像テンプレートの面の法線方向を表す。後述する画像テンプレートの選択や，入力画像の画像パターン作成に用いる。

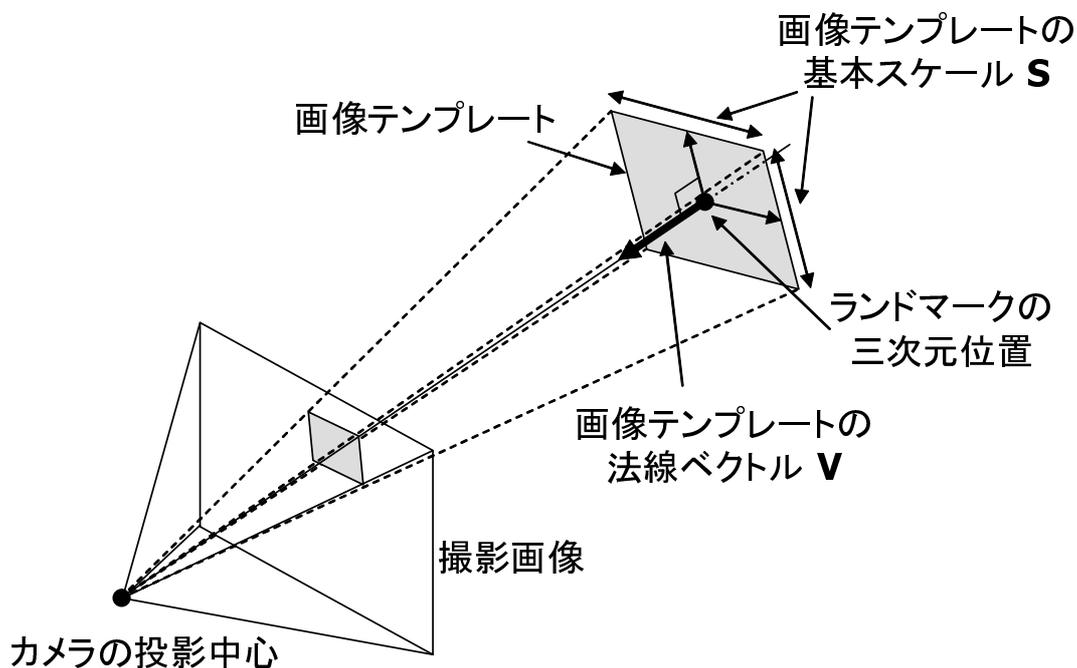


図 8 ランドマークとその画像テンプレート

(D) 画像テンプレートの基本スケール

画像テンプレートのスケールとは、図8に示すような、画像テンプレートの実空間中での大きさであると定義する。画像テンプレートの基本スケールは、(B)で述べた多重スケールの画像テンプレートの中で、最小スケールの画像テンプレートの実空間中での大きさを表す。本研究では、画像テンプレートの解像度がランドマーク作成時に撮影された全方位画像の解像度とほぼ等しくなるように基本スケールを決定する。

### 3.2 全方位カメラからの三次元復元によるランドマーク情報の獲得

本節では、全方位動画からの三次元復元を利用して、ランドマーク情報を獲得し、ランドマークデータベースを構築する方法について述べる。ランドマークデータベースの構成要素は先に図7に示した通りである。はじめに、全方位動画



図 9 スケールの異なる画像テンプレート

像から環境の三次元復元を行うことによって、ランドマークの三次元位置 (1) と撮影時の全方位カメラの位置・姿勢 (A) を取得する。次に、三次元復元で得られた情報を用いて画像テンプレートの法線ベクトル (C)、画像テンプレートの基本スケール (D) を決定し、ランドマークを中心とした多重スケール画像テンプレート (B) を作成する。以下では、これらの作成方法について順に述べる。

### 3.2.1 全方位動画像からの環境の三次元復元

ここでは、まず佐藤らの shape-from-motion 手法 [21] を用いて三次元位置が既知の少数の基準点と Harris オペレータ [31] によって検出された多数の自然特徴点を全方位動画像中で自動追跡し、カメラの動きパラメータを推定する。この手法では、まず基準点の三次元位置をトータルステーションと呼ばれる三次元測量機材を用いて計測し、少数のキーフレームの画像上で基準点を指定する。次に、全方位動画像中の自然特徴点をフレーム毎に自動追跡し、動画像全体での最適化処理を行うことで、カメラパラメータの累積的な推定誤差を最小化する。これにより、複雑で広範囲な環境を撮影した場合でも、全方位動画像のカメラパスと自然特徴点の三次元位置を基準マーカによる絶対座標系で求めることができる。

### 3.2.2 ランドマーク情報の獲得

次に、三次元復元によって得られた情報を用いてランドマークの情報を自動的に獲得し、ランドマークデータベースを構築する。佐藤らの手法により推定され

た全方位動画像のカメラパスと自然特徴点の三次元位置は，3.1 節で述べたランドマークの三次元位置 (1)，ランドマーク撮影時の全方位カメラの位置 (A) として用いる．また，撮影された全方位画像とそのカメラ位置・姿勢，自然特徴点のランドマークの三次元位置から画像テンプレートの法線ベクトル (C)，画像テンプレートの基本スケール (D) を決定し，それらを用いて多重スケール画像テンプレート (B) を作成する．以下に，それぞれの作成方法について述べる．

(1) ランドマークの三次元位置

本手法では，環境の三次元復元によって三次元位置が推定された全方位画像上の自然特徴点をランドマークとして用いる．また，ランドマークの三次元位置として，推定された自然特徴点の三次元位置を用いる．

(2) 撮影地点ごとの情報

(A) 撮影時の全方位カメラの位置

3.2.1 節の手法によって復元されたカメラパラメータに基づき，ランドマーク撮影時の全方位動画像のカメラ位置を用いる．

(B) 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

画像テンプレートは，後述する画像テンプレートの法線ベクトルと基本スケールからランドマークを中心とした面を決定し，その面上に撮影画像を投影することで作成する．まず，図 8 に示すような法線ベクトル  $V$  と，基本スケール  $S$  を求める．次に，以下の条件を満たす正方形の面を，基本スケールに対応する画像テンプレートの投影面として設定する．

- ランドマークの三次元位置を中心とする
- 法線ベクトル  $V$  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
- 世界座標系の  $XY$  平面と面の  $X$  軸が平行である

画像テンプレートの画素数を  $N \times N$  個とした場合，この投影面を等間隔な格子で  $N \times N$  個に分割する．次に，各格子点の三次元座標を撮影

画像面上に投影し，投影された画像上の画素値をその点の画素値とする．これを基本スケールの画像テンプレートとする．

同様にして，スケールが2倍，4倍の画像テンプレートを作成し，撮影地点ごとに保持する．しかし，このような手法で，スケールが2倍，4倍の画像テンプレートを作成すると，撮影画像上で疎な点を用いるため，投影された点の間に存在する画素の色情報は用いられない．従って，スケールが2倍，4倍の画像テンプレートを作成するときは，撮影画像の代わりに撮影画像を $1/2$ ， $1/4$ に縮小した画像に三次元座標を投影し，その画素値を用いることで，周辺画素の情報も考慮した色情報を取得する．

(C) 画像テンプレートの法線ベクトル

画像テンプレートの法線ベクトルはランドマークの三次元位置から撮影時の全方位カメラ位置への向きベクトルとする．

(D) 画像テンプレートの基本スケール

本研究では，画像テンプレートの1画素が，撮影画像上の1画素に対応するように画像テンプレートの基本スケールを決定する．画像テンプレートのスケールは，図10に示すようにランドマークの奥行き距離 $D$ に比例する．そこで，撮像面上における画素間の距離 $d$ を用いて，テンプレート画像平面が撮像面に対して平行であると仮定したときの，テンプレート画像平面上における1画素の大きさ $s$ を以下のように算出する．

$$s = D \times d/f \quad (1)$$

ただし， $f$ はカメラの焦点距離を表す．これにより，基本スケール $S$ を $S = s \times (N - 1)$ として決定することで，画像テンプレートの解像度と撮影画像の解像度がほぼ等しくなるような画像テンプレートを作成する．

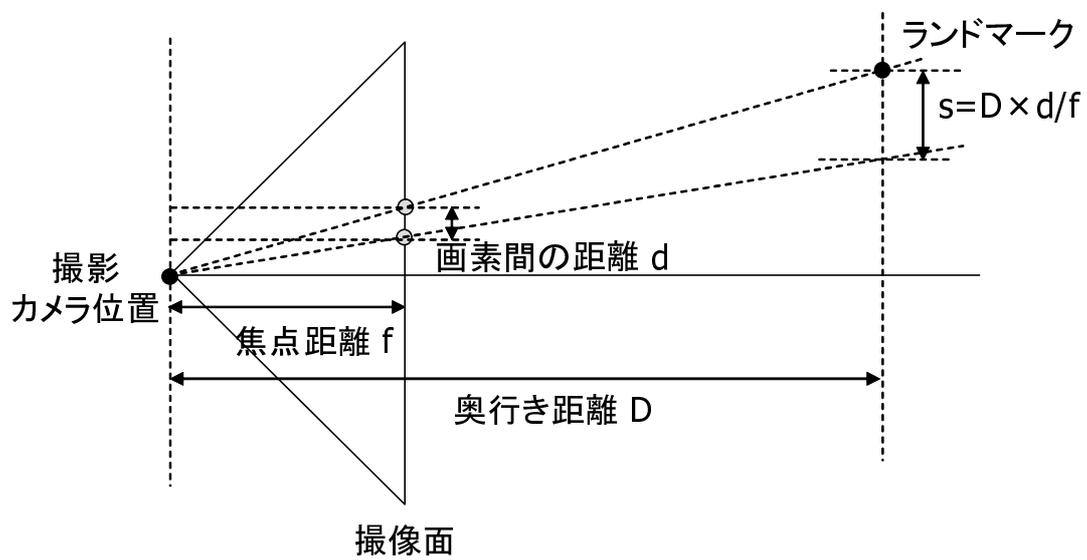


図 10 基本スケールの決定

## 4. 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

### 4.1 データベースに基づくカメラ位置・姿勢推定手法の概要

本章では，前章の手法によって事前に作成されたランドマークデータベースを用いたカメラ位置・姿勢推定手法について述べる．本手法のカメラ位置・姿勢推定処理の流れは先に示した図6の通りである．まず，入力画像の初期カメラ位置・姿勢を推定する．次に，前章で作成されたデータベースからランドマークを選択する (2-a)．入力画像から自然特徴点の検出を行い，検出された自然特徴点とランドマークの画像テンプレートとのマッチングを行うことによって入力画像とランドマークの対応付けを行う (2-b)．最後に，それらの対応関係を用いてカメラ位置・姿勢推定を行う (2-c)．カメラ位置・姿勢推定に失敗していればカメラの初期位置・姿勢推定から，成功していればランドマークの選択から処理を繰り返す．ただし，本論文では初期カメラ位置・姿勢は何らかの手法によって取得済みとする．以下，各処理について詳述する．

### 4.2 ランドマークの選択

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選択する．また，ランドマークの見え方の変化に対応するため，複数の画像テンプレートの中から入力画像と見え方が最も近いものを選択し，マッチングに使用する．本手法では，対応付けに用いるランドマークとその画像テンプレートの組を以下の4つの条件によって選択する．ただし，ここでは，現フレームの入力画像のカメラ位置・姿勢は未知であるため，前フレームのカメラ位置・姿勢を代わりに用いる．

(条件1) ランドマークが入力画像の視野範囲内に存在する：ランドマークの三次元位置を前フレームのカメラ位置・姿勢を用いて撮像面に投影する．これにより，入力画像面上に投影されたランドマークのみを選択する．

(条件 2) ランドマーク撮影時のカメラ位置と入力画像撮影時のカメラ位置の距離が閾値以下である：ランドマークの撮影位置と前フレームのカメラ位置間の距離を算出し，距離が閾値以下のランドマークを選択する．

(条件 3) ランドマークの画像テンプレートの法線ベクトルとランドマークから入力画像の撮影カメラ位置までのベクトルの成す角が閾値以下であり最小である：ランドマークの三次元位置と前フレームのカメラ位置を結ぶ直線と画像テンプレートの法線ベクトルとのなす角  $\theta$  が最小のものを選択する．直線と法線ベクトルのなす角  $\theta$  が閾値以上であれば，そのランドマークを用いない．

(条件 4) 既に選択されたランドマークと入力画像上で密接していない：まず，入力画像を格子状に分割する．ランドマークの三次元位置を前フレームのカメラ位置・姿勢を用いて入力画像上に投影したときに，同じ格子内に投影されるランドマークが既に選択されていた場合は，このランドマークを用いない．

はじめに，条件 1 を満たすランドマークの中から，条件 2 のランドマークの撮影位置と前フレームのカメラ位置間の距離が小さい順に，条件 3 を満たすランドマークを一定数選択する．選択されたランドマークの中から，条件 3 で述べる角  $\theta$  が小さい順にランドマークを選択していき，条件 4 を満たすランドマークが一定数選択されるまで処理を繰り返す．

### 4.3 入力画像とランドマークの対応付け

本節では，データベースから選択されたランドマークを入力画像内で探索し，対応付ける手法について述べる．まず，ランドマークの三次元位置を前フレームのカメラ位置・姿勢を用いて現在の入力画像上に投影し，その周辺で検出された自然特徴点を，投影したランドマークと対応付ける候補とする．次に，検出された自然特徴点を中心とした画像パターンと，ランドマークの画像テンプレートとのパターンマッチングを行うことで，前節で選択されたランドマークとの対応付けを行う．以下に，それぞれの処理手順を詳述する．

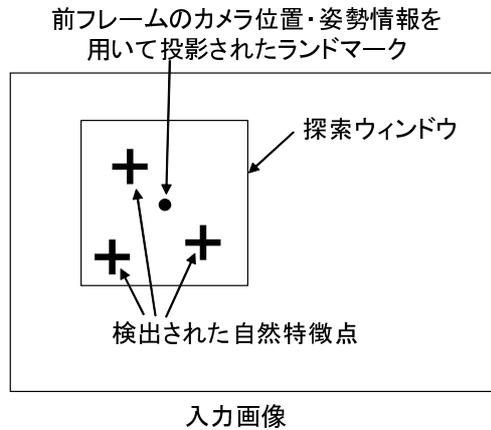


図 11 ランドマークと対応付ける自然特徴点の検出

#### 4.3.1 入力画像中の自然特徴点の検出

まず，図 11 に示すように，前フレームのカメラ位置・姿勢を用いてランドマークの三次元位置を入力画像上に投影する．投影された座標周辺にランドマークに対応する自然特徴点が存在すると仮定し，投影座標を中心とする一定サイズの探索ウィンドウ内でランドマークデータベース作成時と同じ Harris オペレータ [31] により自然特徴点を検出する．

Harris オペレータによる入力画像上の座標  $(x, y)$  の特徴量  $H_{xy}$  は，まずガウシアンオペレータによる入力画像の平滑化処理を行い，次に一定の大きさの正方形窓  $W$  において，画像上の輝度  $I$  の勾配  $I_u, I_v$  を用いて以下に示す行列  $A$  を算出し，その行列の固有値  $\lambda_1, \lambda_2$  の最小値を求めることで得られる．

$$A = \sum_{\mathbf{x} \in W} \begin{pmatrix} I_u(\mathbf{x})^2 & I_u(\mathbf{x})I_v(\mathbf{x}) \\ I_u(\mathbf{x})I_v(\mathbf{x}) & I_v(\mathbf{x})^2 \end{pmatrix} \quad (2)$$

$$H_{xy} = \min(\lambda_1, \lambda_2) \quad (3)$$

この特徴量  $H_{xy}$  が閾値以上，かつ極大である点を自然特徴点として検出する．このようにして窓  $W$  内で検出された自然特徴点をランドマークと対応付ける候補とする．

#### 4.3.2 画像テンプレートとのマッチングによる対応付け

ここではまず、各ランドマークについて、対応候補点として検出された自然特徴点ごとに、その点を中心とした画像パターンを作成する。次に、それらの画像パターンとランドマークの画像テンプレートを正規化相互相関によって比較する。最も相関値が高い画像パターンを持つ自然特徴点周辺の全画素を候補として再度、画像パターンを作成し、ランドマークの画像テンプレートとの相関値を得る。このとき、最も相関値が高い点をランドマークと対応付ける。これらの詳細を以下で述べる。

それぞれの画像パターンはランドマークの画像テンプレートと同様に、自然特徴点を中心とした面上に入力画像を投影したものであり、前フレームのカメラ位置・姿勢、対応付けるランドマークの三次元位置、画像パターンの法線ベクトルとスケール情報を用いて作成する。自然特徴点を中心とした面はランドマークの画像テンプレート作成時と同一の向きであるものとし、作成する画像パターンのスケールは対応付けるランドマークの画像テンプレートの基本スケールと同一のものと、その2倍、4倍の多重スケールとする。作成したスケールの異なる画像パターンのセットを多重スケール画像パターン  $p$  とする。このようにして作成した多重スケール画像パターン  $p$  とランドマークの多重スケール画像テンプレート  $t$  から、以下に示す正規化相互相関により相関値を求め、まず画像テンプレートと最も相関の高い画像パターンを持つ自然特徴点を選択する。

$$C(t, p) = \frac{\sum_{j=0}^y \sum_{i=0}^x (t(i, j) - \bar{t})(p(i, j) - \bar{p})}{\sqrt{\sum_{j=0}^y \sum_{i=0}^x (t(i, j) - \bar{t})^2 \sum_{j=0}^y \sum_{i=0}^x (p(i, j) - \bar{p})^2}} \quad (4)$$

ただし、 $\bar{t}$ 、 $\bar{p}$  はそれぞれ  $t$ 、 $p$  の画素値の平均であり、 $x$ 、 $y$  は画像パターン及び画像テンプレートの座標である。

入力画像上の自然特徴点がランドマークを投影した座標と同じ位置に検出されるとは限らないため、次に、相関値が最大となる自然特徴点周辺の全画素を対応付け候補として再度相関値を算出し、最終的に相関値が最も高い点をランドマークと対応付ける。これにより、データベース内の自然特徴点の三次元座標と入力

画像上の二次元座標を対応付けることができる。なお、相関値が閾値以下の場合には対応付けは行わず、ランドマークを棄却する。

#### 4.4 ランドマークに基づくカメラ位置・姿勢推定

入力画像とランドマークの対応付けによって得られた二次元座標と三次元座標の対応関係を用いてカメラ位置・姿勢推定を行う。まず、RANSAC[32]により誤対応を排除する。そして、誤対応を除いた正しい対応関係のみを用いてカメラ位置・姿勢推定を行い、再投影誤差が最小となるカメラパラメータを最終的な推定結果とする。RANSACによる誤対応の排除では、以下の処理を行う。

- (1) 対応付けられたランドマークの三次元座標と二次元座標の組からランダムに6組以上を選択し、暫定的なカメラパラメータを推定する。
- (2) 全ての組について、ランドマークの二次元座標と三次元位置を暫定的なカメラパラメータによって画像上に投影した座標との距離の二乗誤差を再投影誤差として求める。

ランドマーク  $i$  の二次元座標を  $x_i$ 、ランドマーク  $i$  の三次元座標を暫定的なカメラパラメータで画像上に投影した座標を  $\hat{x}_i$  とすると、ランドマーク  $i$  の再投影誤差  $R_i$  と再投影誤差の中間値  $RM$  は以下の式によって表される。

$$R_i = |x_i - \hat{x}_i|^2 \quad (5)$$

$$RM = \text{median}(R_1, R_2, \dots, R_n) \quad (6)$$

ただし、 $n$  は対応付けられたランドマークの数とする。

- (3) 上記の(1)、(2)の処理を複数回繰り返し、再投影誤差の中間値  $RM$  が最小となる暫定カメラパラメータを選択する。
- (4) 選択されたカメラパラメータによって各ランドマークの投影誤差を評価し、閾値を超える結果を誤対応として削除する。

最後に、得られた正しいと判定された対応関係のみを用いて、再投影誤差が最小となるカメラパラメータを求める。ここでは出口らの手法 [33] を利用し、線形最小二乗法によってカメラパラメータの初期値を算出し、補正を行った後に、再投影誤差の非線形最小化を行う。このようにして得られたカメラパラメータを最終的なカメラ位置・姿勢推定結果とする。

## 5. 実験

入力画像の自然特徴点とランドマークデータベースとの対応付けによるカメラ位置・姿勢推定手法の有効性を検証するために，屋外・屋内環境における実験と精度評価を行った．まず，表 2 に示す条件で，実際に屋外・屋内環境を全方位カメラを用いて撮影し，環境内の自然特徴点とカメラパスの三次元復元を行うことで，ランドマークデータベースを構築した．次に，表 3 に示す条件で，構築したランドマークデータベースを用いた動画像のカメラ位置・姿勢推定を行った．また，カメラ位置・姿勢推定結果を用いて CG を重畳表示することで，カメラ位置・姿勢が正しく推定できているかを確認した．評価実験では，実際に三次元座標を計測した点のみを用いて正解データを作成し，ランドマークデータベースによるカメラ位置・姿勢推定結果との比較を行った．

### 5.1 屋外環境における実験

#### 5.1.1 ランドマークデータベースの構築 (屋外実験)

まず，図 12 左に示すような水平方向に 5 個，上方向に 1 個の CCD カメラを配置した全方位型マルチカメラシステム (Point Grey Research 社 Ladybug) を用いて図 12 右に示すような屋外環境 (本学キャンパス) を複数の動画像系列 (1251 フレーム × 6 台) として撮影した．次に，佐藤らの手法 [21] によってカメラパスと自然特徴点の三次元座標を推定し，表 2 に示す条件でランドマークデータベースを構築した．ただし，データ量削減のため，三次元復元で得られた全フレームのデータに対し，10 フレームおきのデータを用いた．また，それぞれのランドマークにつき，3 つの 15×15 画素の画像テンプレートからなる多重スケール画像テンプレートを撮影地点ごとに作成した．本実験で得られたランドマークの数は約 12400 個であり，一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．推定された全方位動画像のカメラパスと自然特徴点の三次元位置を図 13 に示す．図中の曲線は推定されたカメラパス，錐台は 100 フレームおきのカメラ位置・姿勢を表す．



図 12 全方位型マルチカメラシステム Ladybug と撮影された全方位画像

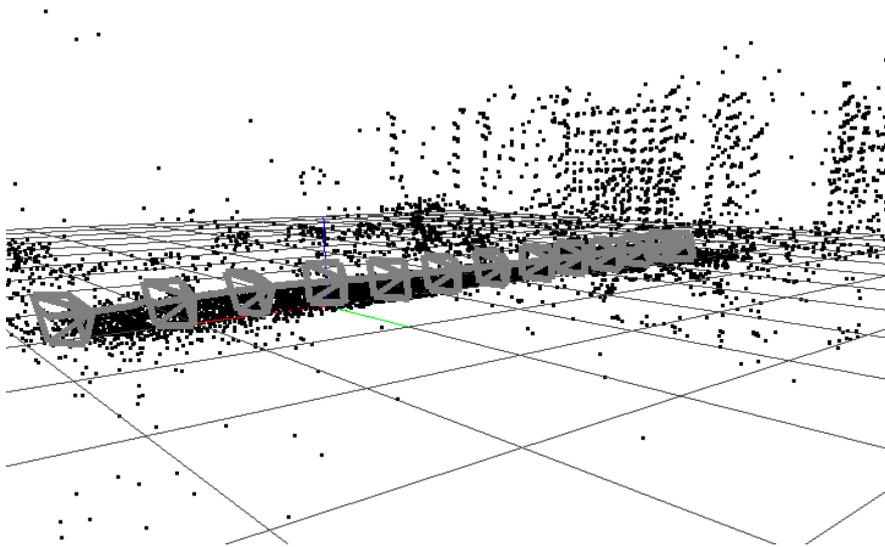


図 13 推定されたカメラパスと自然特徴点の三次元位置 (屋外実験)

表 2 ランドマークデータベース作成の条件

ランドマークデータベース作成時の条件

	屋外実験	屋内実験
撮影全方位カメラ	PointGrey Ladybug	
撮影全方位画像	768 × 1024 画素 × 6, 15fps	
撮影したフレーム数	1251 フレーム (10 フレームおきに使用)	701 フレーム (5 フレームおきに使用)

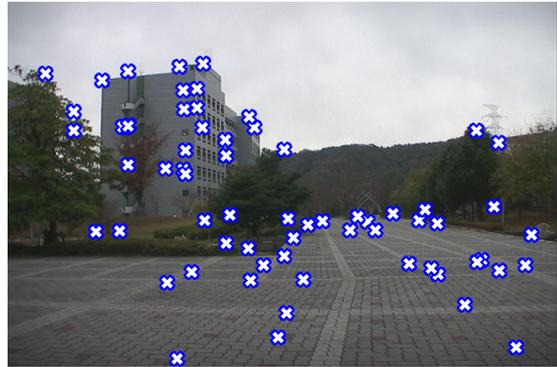
### 5.1.2 カメラパラメータ推定 (屋外実験)

次に，広角レンズ (SONY VCL-HG0758) を取り付けたビデオカメラ (SONY DSR-DP-150) を用いて 1000 フレームの単眼動画像 (解像度 720× 480 画素，プログレッシブスキャン，15fps) を撮影し，4 章で述べた手法による逐次的なカメラ位置・姿勢推定を行った．ただし，本実験では初期フレームのカメラ位置・姿勢はあらかじめ既知とし，手動で与えた．実験時の条件は表 3 に示すとおりである．本実験においてデータベースから一度に選択するランドマークの数はフレーム当たり最大 100 個，入力画像中の対応付け候補の探索範囲は 120×60 画素，RANSAC による繰り返し回数は 500 回とした．逐次推定の結果，1 フレームの処理時間は PC (CPU: Pentium4 3GHz，Memory: 1.5GB) を用いて平均 2.8 秒であった．

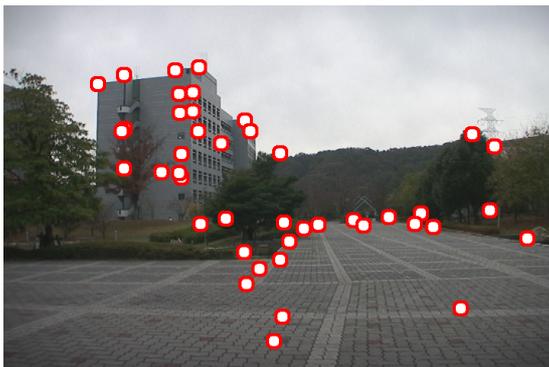
あるフレームのカメラ位置・姿勢推定に用いられたランドマークの画像上の位置を図 14 に，カメラ位置・姿勢推定に用いられたランドマーク数を図 15 に示す．図 14 は，(a) がデータベースから選択されたランドマーク，(b) が選択されたランドマークのうち，入力画像と対応付けられたランドマーク，(c) が対応付けられたランドマークの中から RANSAC によって誤対応を排除した後のランドマーク，(d) がそれらをまとめたものである．これらのランドマークは，実際に入力画像と対応付けられたランドマークは対応付けが行われた位置に，データベースから選択されたが入力画像と対応付けられなかったランドマークは最終的に推定されたカメラ位置・姿勢情報を用いて入力画像上に投影された位置に表示している．



(a) データベースから選択された  
ランドマーク



(b) 入力画像と対応付けられた  
ランドマーク



(c) RANSAC によって誤対応を排除した  
後のランドマーク



(d) (a) ~ (c) をまとめて表示した結果

- データベースから選択されたランドマーク
- ✕ 入力画像と対応付けられたランドマーク
- RANSAC によって誤対応を排除した後のランドマーク

図 14 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配置 (屋外実験)

表 3 ランドマークデータベースを用いたカメラ位置・姿勢推定の条件

ランドマークデータベースを用いたカメラ位置・姿勢推定の条件

	屋外実験	屋内実験
入力撮影カメラ	SONY DSR-DP-150	
広角レンズ	SONY VCL-HG0758	
入力画像	720 × 480 画素, 15fps	
画像上の特徴点の探索範囲	120 × 60 画素	60 × 60 画素
データベースから選択する ランドマーク数	100 個/フレーム	100 個/フレーム
マシンスペック	CPU:Pentium4 3.0GHz , Memory:1.5GB	

図 15 も同様に，上から，データベースから選択されたランドマーク，入力画像と対応付けられたランドマーク，対応付けられたランドマークの中から RANSAC によって誤対応を排除した後のランドマークである．結果，1 フレーム中で入力画像と対応付けられたランドマークは平均 78 個，最終的にカメラ位置・姿勢推定に使われたランドマークは平均 57 個であった．データベースから選択されるランドマーク数は 1 フレームにつき 100 個としたが，645 フレーム目 (図 16 参照) のように，データベースからランドマークを選択するための条件を満たすランドマークが 100 個以下の場合もあった．原因としては，撮影環境内に自然特徴点が少ない場合や，撮影カメラ位置から撮影物体までの距離が短く，ランドマーク選択時の条件 3 を満たすランドマークが少ない場合などが考えられる．

最後に，200 フレームおきのカメラ位置・姿勢推定に用いられたランドマークの画像上の配置を図 17 に，推定されたカメラ位置・姿勢を用いて仮想物体を一定位置に重畳表示させた match move の結果を図 18 に示す．図 17 で示すランドマークの表示位置は図 14 と同様にして決定している．図 18 から，CG で描かれた人物が入力画像と幾何学的に正しく合成されており，カメラ位置・姿勢の推定結果に大きな誤差が含まれていないことが分かる．

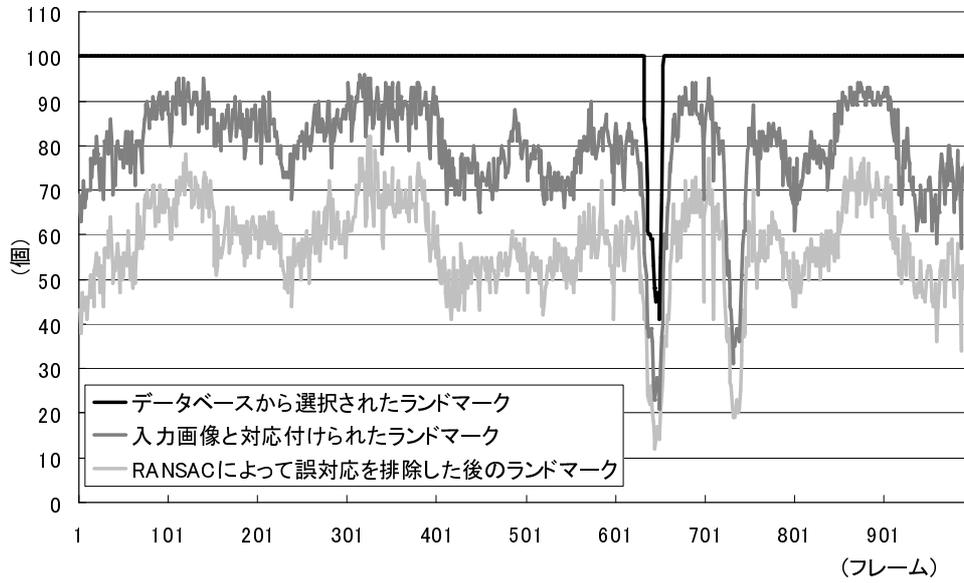


図 15 カメラ位置・姿勢推定に用いられたランドマーク数 (屋外実験)

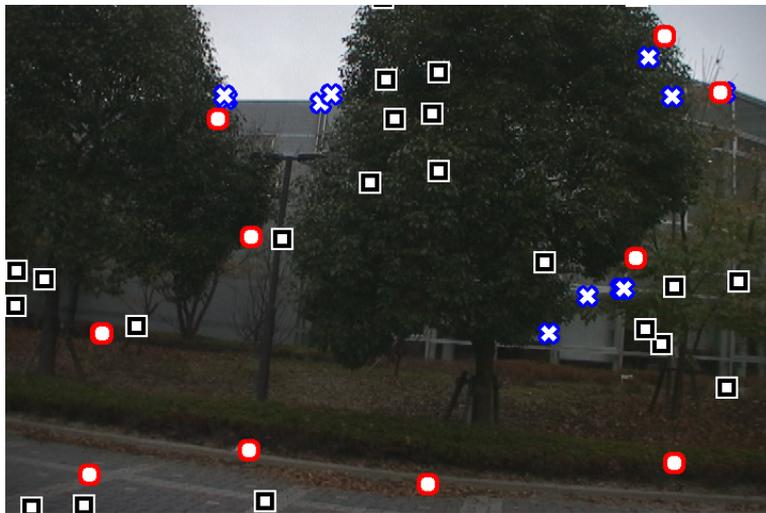
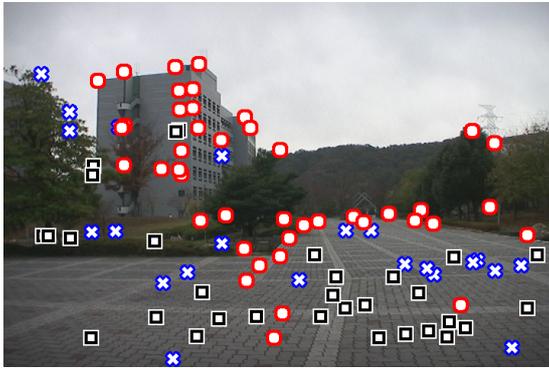
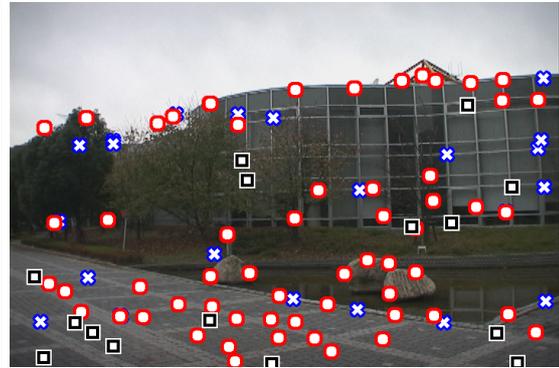


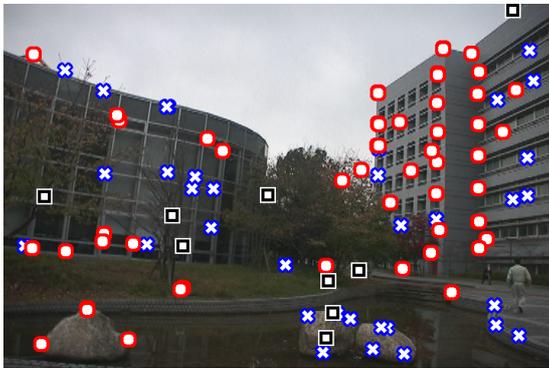
図 16 645 フレーム目でカメラ位置・姿勢推定に用いられたランドマーク (屋外実験)



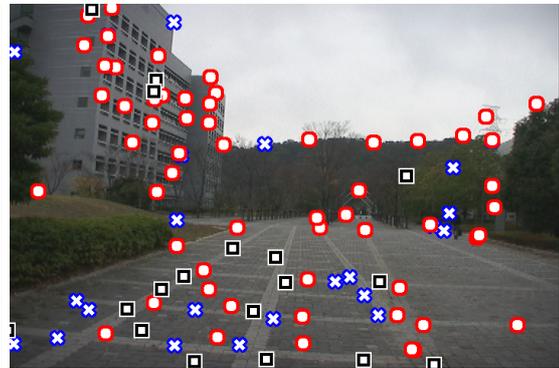
(a)1 フレーム目



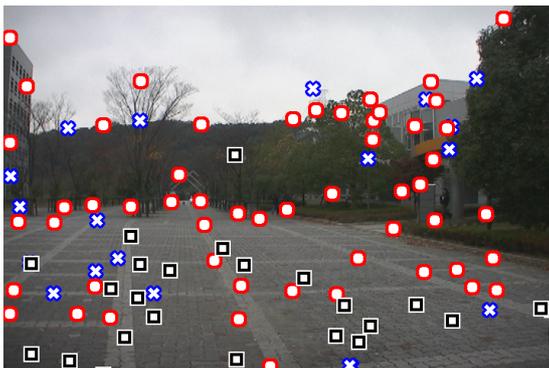
(b)200 フレーム目



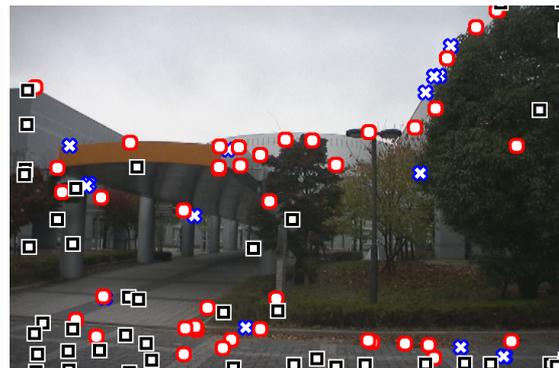
(c)400 フレーム目



(d)600 フレーム目



(e)800 フレーム目



(f)1000 フレーム目

図 17 カメラ位置・姿勢推定に用いられたランドマーク (屋外実験)



(a)1 フレーム目



(b)200 フレーム目



(c)400 フレーム目



(d)600 フレーム目



(e)800 フレーム目



(f)1000 フレーム目

図 18 カメラ位置・姿勢推定結果を用いた match move(屋外実験)

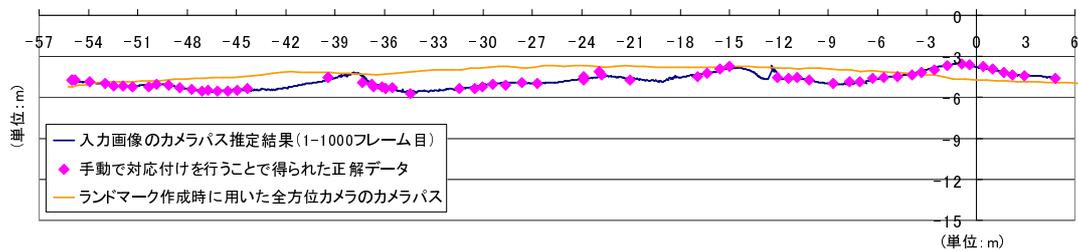


図 19 屋外実験において推定されたカメラパスと正解データ

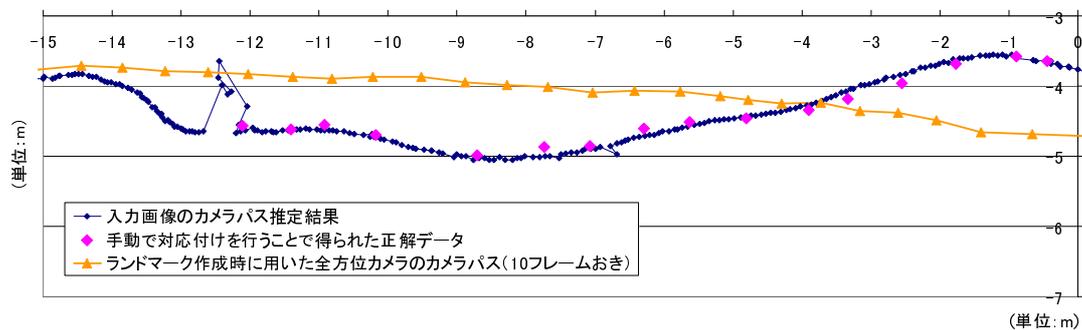


図 20 屋外実験において推定されたカメラパスと正解データ (一部拡大)

### 5.1.3 定量的な評価実験 (屋外実験)

次に、推定されたカメラ位置・姿勢を正解データと比較することで推定精度を評価した。本実験では、環境内の特徴点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の位置を手動で対応付け、4.4節に述べた手法を用いることでカメラパラメータの正解データを10フレームおきに作成した。ただし、自然物が入力画像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少ないフレームや、手動で対応付けを行った結果による再投影誤差が1.5画素以上であるフレームに関しては評価対象とはしなかった。

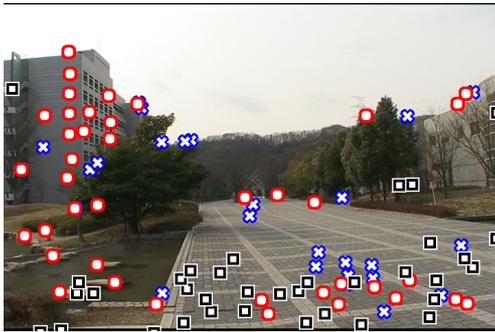
屋外実験の結果、1000フレーム中で正解データを作成した66地点におけるカメラの位置推定誤差は平均220mm、カメラの光軸角度の推定誤差は平均0.37度であった。図19に初期フレームから1000フレーム目までのカメラパラメータと正解データを、図20に図19の一部を拡大したものを示す。図19、20のX軸、Y軸はそれぞれランドマークデータベース構築時の世界座標系におけるX軸、Y軸

に対応している。図 19, 20 から, 推定されたカメラ位置・姿勢に累積的な推定誤差が生じていないことを確認できる。

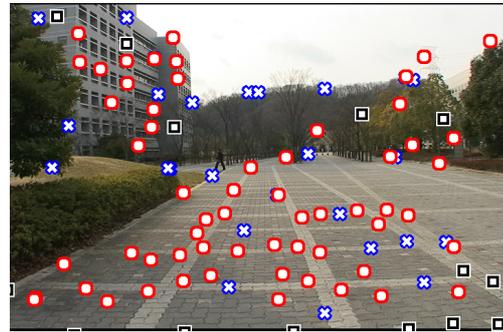
また, ランドマーク作成時の全方位カメラパスから離れるに従い, カメラ位置・姿勢推定を行う入力画像におけるランドマークの見え方が変化すると考えられる。そこで, ランドマーク作成時の全方位カメラパスから入力画像の撮影位置までの距離の変化によるカメラ位置・姿勢推定精度の異なりを検証するため, 撮影位置が異なる 4 つの入力動画像 (シーン 1 からシーン 4) のカメラ位置・姿勢推定を行い, 正解データとの比較を行った。図 21 にそれぞれのシーンの初期フレームと最終フレームを示す。図中のマークは図 14 と同様に, カメラ位置・姿勢推定に用いられたランドマークを表している。

シーン 1 では, ランドマーク作成時の全方位カメラパスに近い位置において, カメラを一定方向に向けたまま移動した。シーン 2 では, ランドマーク作成時の全方位カメラパスから 3m 程度離れた位置をシーン 1 と同じ向きにカメラを固定して移動した。また, シーン 3 では, シーン 2 と同じような位置をカメラをシーン 2 とは逆方向にカメラを固定して移動した。シーン 4 では, ランドマーク作成時の全方位カメラパスに対して垂直に, カメラを一定方向に向けて移動した。

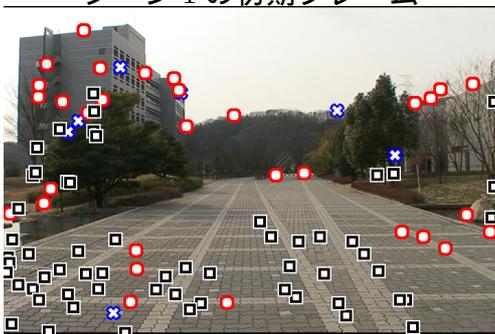
実験の結果, 図 22 から図 25 に示すようなカメラパスが推定された。シーン 1 のカメラパス推定結果 (図 22) とシーン 2 のカメラパス推定結果 (図 23) を比較すると, シーン 2 の推定精度が悪いことが分かる。これは, ランドマーク作成時の全方位カメラパスから離れることによって, 選択されるランドマークに偏りが出たことが主な原因と考えられる。また, シーン 2 のカメラパス推定結果 (図 23) とシーン 3 のカメラパス推定結果 (図 24) を比較すると, ランドマーク作成時の全方位カメラパスからの距離は同程度であるが, シーン 3 の方が推定精度が良い。これは, シーン 2 ではランドマークが類似パターンの多い地面や遠方にしか存在していないのに比べて, シーン 3 では常に正面に建物が存在し, 見え方の変化が少なく, 対応付けが容易なランドマークが多く存在したためと考えられる。シーン 4 では, ランドマーク作成時の全方位カメラパスに対して垂直に移動している。シーン 4 のカメラパス推定結果 (図 25) においても, ランドマーク作成時の全方位カメラパスに近い位置で撮影された入力画像のほうが推定精度が良いと言える。



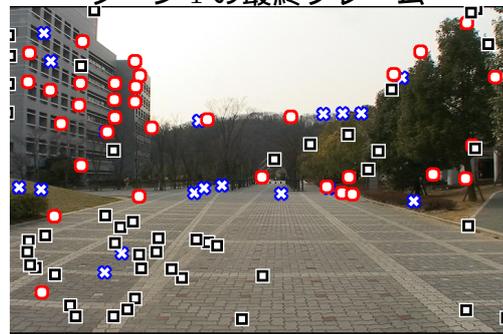
シーン1の初期フレーム



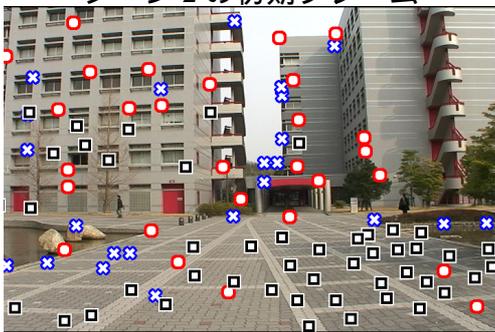
シーン1の最終フレーム



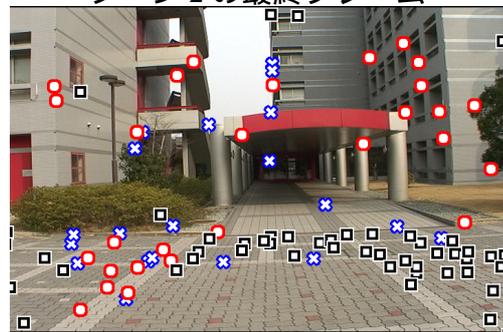
シーン2の初期フレーム



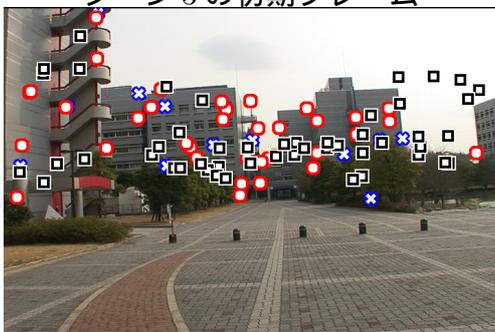
シーン2の最終フレーム



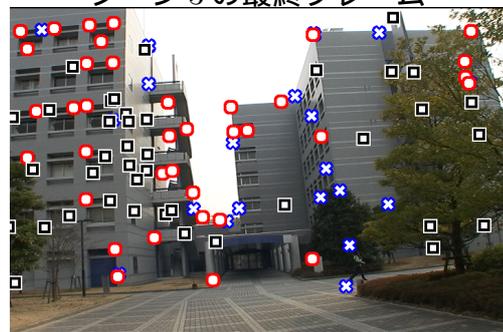
シーン3の初期フレーム



シーン3の最終フレーム



シーン4の初期フレーム



シーン4の最終フレーム

図 21 それぞれのシーンのカメラパスと正解データ

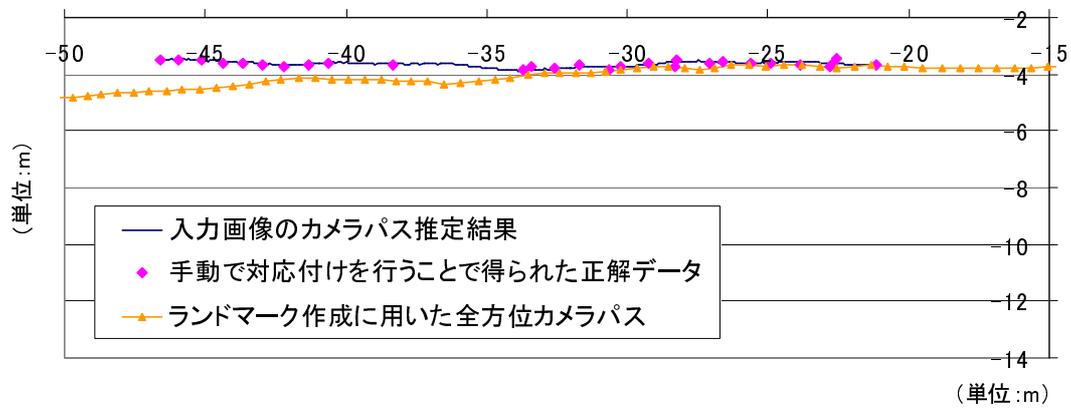


図 22 シーン 1 のカメラパスと正解データ

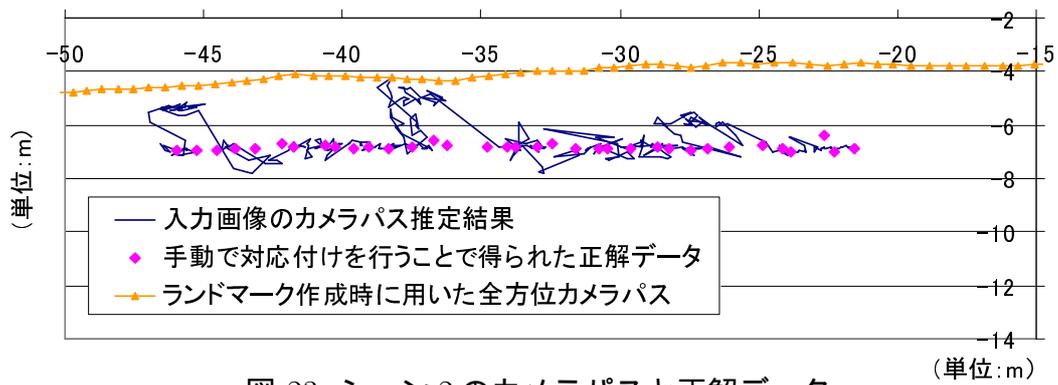


図 23 シーン 2 のカメラパスと正解データ

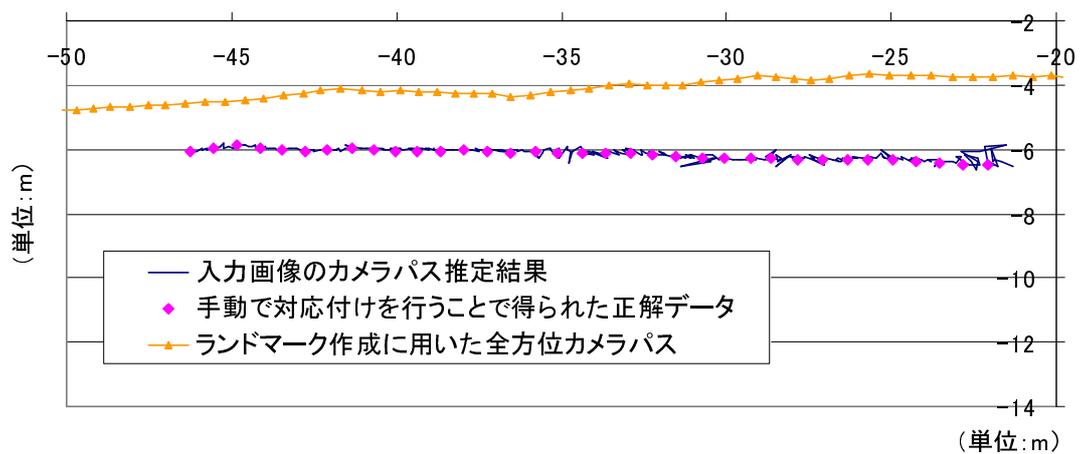


図 24 シーン 3 のカメラパスと正解データ

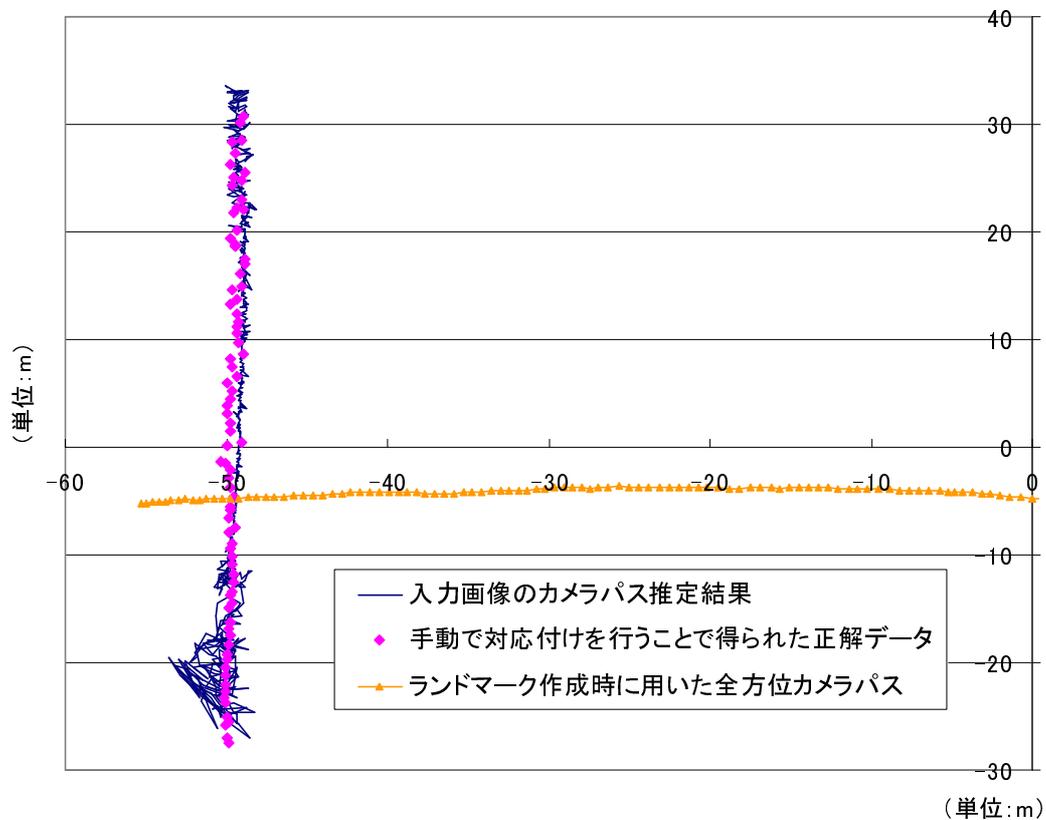


図 25 シーン 4 のカメラパスと正解データ

しかし、このシーンにおいては、正面に存在する建物から検出されたランドマークが多数存在している。これらのランドマークはシーン 4 での移動方向による見え方の変化が少ないため、30m 離れた位置においてもある程度良好なカメラ位置・姿勢推定が得られている。また、シーン後半では、カメラパスから離れるに従って撮影位置からランドマークへの距離が短くなっている。そのため、シーン前半とシーン後半を比べると、ランドマーク作成時の全方位カメラパスから同程度離れている場合でも、シーン後半の方がカメラ位置・姿勢推定精度が良いのではないかと考えられる。

## 5.2 屋内環境における実験

次に，屋外環境と同様の手順で屋内環境における実験を行った．

### 5.2.1 ランドマークデータベースの構築 (屋内実験)

表 2 で示したように，屋内環境における実験で用いた全方位動画像のフレーム数は 701 フレームであり，ランドマークデータベース作成には三次元復元で得られた全フレームのデータに対して 5 フレームおきのデータを用いた．また，屋外実験と同様に，それぞれのランドマークにつき 3 つの  $15 \times 15$  画素の画像テンプレートからなる多重スケール画像テンプレートを撮影地点ごとに作成した．実験で得られたランドマークの数は約 2100 個であり，一つのランドマークにつき平均 9.4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．

### 5.2.2 カメラパラメータ推定 (屋内実験)

ランドマークデータベースを用いたカメラ位置・姿勢では，900 フレームの単眼動画像をフレーム毎に逐次推定した．また，データベースから一度に選択するランドマークの数は最大 100 個，入力画像中の対応付け候補の探索範囲は  $60 \times 60$  画素，RANSAC による繰り返し回数は 500 回とした．逐次推定の結果，1 フレームの処理時間は屋外実験と同じマシンを用いて平均 2.5 秒であった．

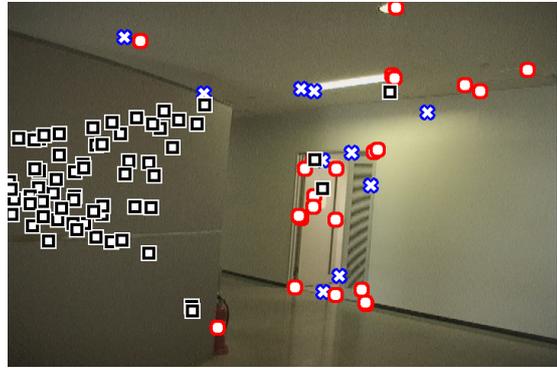
カメラ位置・姿勢推定に用いられたランドマークの空間配置を図 26 に，推定されたカメラ位置・姿勢を用いて仮想物体を一定位置に重畳表示させた結果を図 27 に示す．

### 5.2.3 定量的な評価実験 (屋内実験)

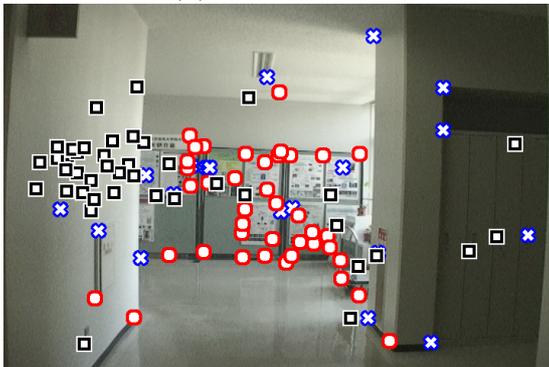
屋外実験と同様に，環境内の特徴点を測定し，各フレームの画像上でそれらの特徴点の位置を手動で対応付けることでカメラパラメータの正解データを 50 フレームおきに作成した．結果，900 フレーム中で正解データを作成した 14 地点におけるカメラの位置推定誤差は平均 52mm，カメラの光軸角度の推定誤差は平均



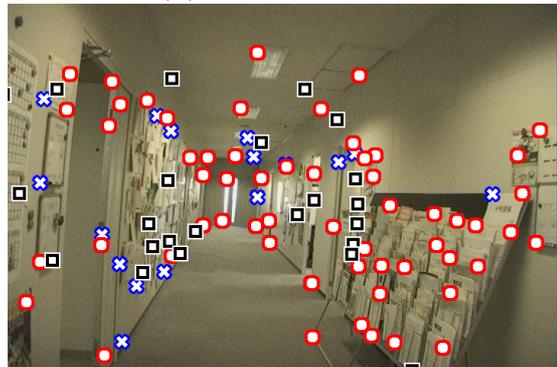
(a)1 フレーム目



(b)200 フレーム目



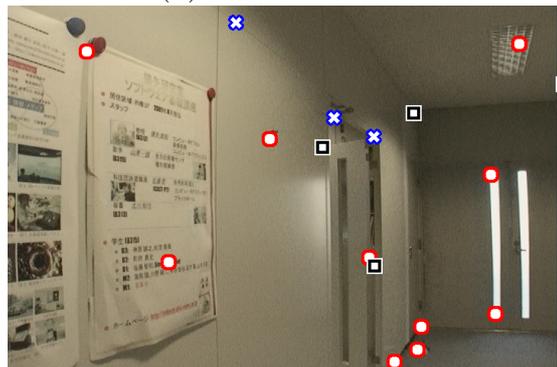
(c)400 フレーム目



(d)600 フレーム目



(e)800 フレーム目



(f)900 フレーム目

図 26 カメラ位置・姿勢推定に用いられたランドマーク (屋内実験)



(a)1 フレーム目



(b)200 フレーム目



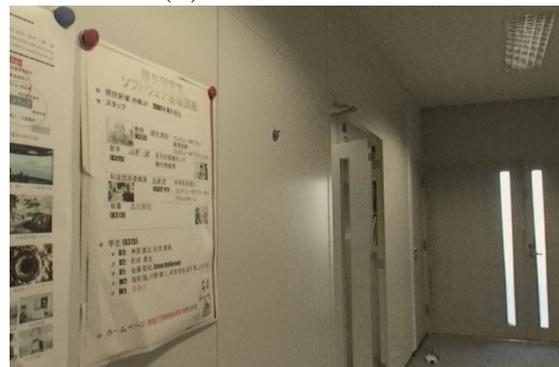
(c)400 フレーム目



(d)600 フレーム目



(e)800 フレーム目



(f)900 フレーム目

図 27 カメラ位置・姿勢推定結果を用いた match move(屋内実験)

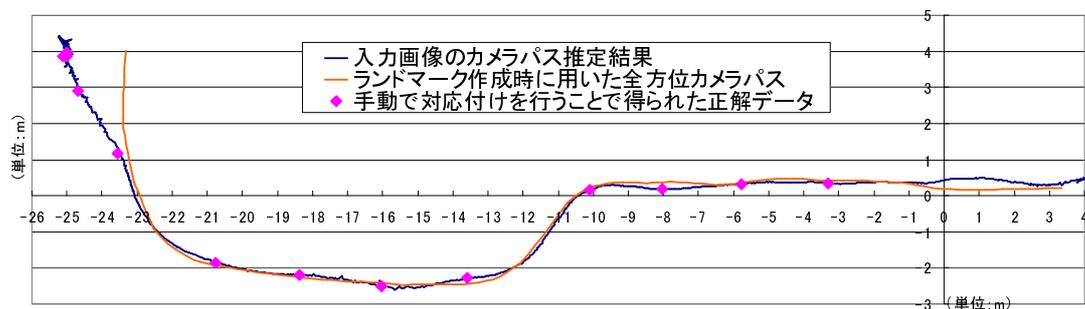


図 28 屋内実験において推定されたカメラパスと正解データ

0.40 度であった。図 28 に初期フレームから 900 フレーム目までのカメラパラメータと正解データを示す。図 28 の X 軸，Y 軸はそれぞれランドマークデータベース構築時の世界座標系における X 軸，Y 軸に対応している。図 28 から，屋内環境においても推定されたカメラ位置・姿勢に累積的な推定誤差が生じていないことを確認できる。

### 5.3 考察

屋外・屋内環境における実験により，自然特徴点によるランドマークデータベースを構築し，構築したランドマークデータベースを用いることで広範囲な環境において動画からのカメラ位置・姿勢推定が可能であることを確認した。また，図 19 から，撮影位置がランドマークデータベース作成時の全方位カメラパスに近い動画においては，長時間かつ長距離の移動を伴う場合においてもカメラ位置の軌跡は概ね滑らかで，かつ，誤差が累積することなく推定されていることが確認できる。また，その間一度もカメラ位置・姿勢推定に失敗することはなかった。しかし，図 20 で示すように，一部では 1m 程度の推定誤差が生じているフレームが存在した。推定誤差が大きいフレームでは，カメラ位置・姿勢推定に撮影カメラ位置から離れたランドマークのみを用いている，ランドマークが入力画像上で偏った配置にある，などの傾向が見られた。そこで，今後はカメラ位置・姿勢推定に用いるランドマークの選択手法を検証し，精度の向上を図る必要があると考えられる。

屋外の評価実験では，ランドマーク作成時の全方位カメラパス付近から撮影された 60m 程度の移動を伴う屋外環境の撮影において，カメラの位置に関して 220mm 程度，カメラの光軸方向に関しては 0.37 度程度の誤差でカメラパラメータを得ることが可能であることを確認した．このことから，屋外環境で撮影された入力画像において，データベース作成時の全方位カメラパス付近でカメラ位置の誤差のみを考慮した場合，投影誤差が 1 画素以内となるように CG を重畳表示するためには，カメラから 14m 以上離れた地点であれば良いことが分かる．また，図 22 から図 25 より，場合によってはランドマークデータベース作成時の全方位カメラパスから数 m ~ 数十 m 離れた位置においても，カメラ位置・姿勢の推定が可能であることを確認した．屋内の評価実験では，ランドマーク作成時の全方位カメラパス付近から撮影された動画像から，カメラの位置に関して 52mm 程度，カメラの光軸方向に関しては 0.40 度程度の誤差でカメラパラメータを得ることが可能であることを確認した．従って，本手法は，それほど厳密な位置合わせを必要としないナビゲーションや AR システムにおける建物等への注釈付加などのアプリケーションに用いるには十分な精度であると考えられる．

## 6. まとめ

本論文では自然特徴点から成るランドマークデータベースを事前に構築し、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位置・姿勢をフレームごとに逐次推定する手法を提案した。ランドマーク作成時には全方位動画画像を用い、画像上の基準点を手動で与えることによって精度の良い三次元復元を行い、半自動的に多数のランドマークを作成する。また、ランドマークの画像テンプレートとして、撮影時のカメラ姿勢やレンズ歪みによる画像の見え方の違いを考慮した多重スケール画像テンプレートを用いた。提案手法では、広範囲な環境を対象とした場合においても、センサ類や人工的なマーカを用いることなく、推定誤差の累積しないカメラの絶対位置・姿勢推定が可能である。

実験では、実際の屋外環境と屋内環境における実験と精度評価を行った。まず、環境内の全方位動画画像を撮影し、特徴点ランドマークデータベースの構築を行った。次に、構築したランドマークを用いて単眼カメラからの動画画像のカメラ位置・姿勢推定を行った。推定されたカメラ位置・姿勢を用いてCGを重畳表示することで、カメラ位置・姿勢の推定に大きな推定誤差が含まれていないことを確認した。評価実験では、手動で作成した正解データと比較することによって、累積誤差が蓄積していないことを確認した。

今後の課題としては、カメラの初期位置・姿勢決定の自動化や逐次推定(ステージ2)のリアルタイム化などが挙げられる。カメラの初期位置・姿勢推定に関しては、画像データベースに基づくカメラ位置・姿勢推定手法[28][34]を応用する、GPSなどのセンサ類を併用することで大まかな位置・姿勢情報を取得する、などの方法が考えられる。また、リアルタイム化には、より高速、高精度なデータベースからのランドマーク選択手法やランドマークと入力画像の自然特徴点の対応付け手法の検討が必要である。このような課題を解決し、リアルタイムでのカメラ位置・姿勢推定を行うことができれば、広範囲な環境内でのARにおける幾何的位置合わせやロボットナビゲーションに利用できると考えられる。

## 謝辞

本研究の全過程を通して、懇切なる御指導，御鞭撻を賜った視覚情報メディア講座横矢直和教授に心より感謝いたします。本研究を進めるにあたり，有益な御助言，御指導を頂いたロボティクス講座小笠原司教授に厚く御礼申し上げます。本研究を通じて，温かい御指導をして頂いた視覚情報メディア講座山澤一誠助教授に深く感謝いたします。そして，本研究の遂行に適切な御助言を頂きました視覚情報メディア講座神原誠之助手に深く御礼申し上げます。また，本研究を行うにあたり，多大なる御助言，御鞭撻を賜った視覚情報メディア講座佐藤智和助手に心より感謝いたします。特に，本研究のテーマ設定から本論文の執筆，その他の発表論文の添削に至るまで細やかな御指導をいただきました。

研究室において物心両面において常に温かい御支援を頂いた，視覚情報メディア講座守屋知代女史に厚く御礼申し上げます。最後に，研究活動だけでなく日々の生活においても大変お世話になった視覚情報メディア講座の諸氏に深く感謝いたします。

## 参考文献

- [1] T. H. S. Feiner, B. MacIntyre and A. Webster: “A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment,” Proc. 1st Int. Symp. on Wearable Computers, pp. 74–81, 1997.
- [2] T. Gleue and P. Daehne: “Design and implementation of a mobile device for outdoor augmented reality in the archeoguide project,” Proc. VAST’01, pp. 28–30, 2001.
- [3] 神原誠之, 横矢直和: “Rtk-gps と慣性航法装置を併用したハイブリッドセンサによる屋外型拡張現実感”, 電子情報通信学会 技術研究報告, PRMU2004-160, pp. 37–38, 2005.
- [4] P. Daehne and J. Karigiannis: “Archeoguide: System architecture of a mobile outdoor augmented reality system,” Proc. 1st IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2002), pp. 263–264, 2002.
- [5] 小田島太郎, 神原誠之, 横矢直和: “拡張現実感技術を用いた屋外型ウェアラブル注釈提示システム”, 画像電子学会誌, Vol. 32, No. 6, pp. 832–840, 2003.
- [6] R. Tenmoku, M. Kanbara and N. Yokoya: “A wearable augmented reality system using positioning infrastructures and a pedometer,” Proc. 7th IEEE Int. Symp. on Wearable Computers (ISWC’03), pp. 110–117, 2003.
- [7] D. Hallaway, T. Höllerer and S. Feiner: “Coarse, inexpensive, infrared tracking for wearable computing,” Proc. 7th IEEE Int. Symp. on Wearable Computers (ISWC’03), pp. 69–78, 2003.
- [8] G. Welch, G. Bishop, L. Vicci, S. Brumback, K. Keller and D. Colucci: “High-performance wide-area optical tracking -the hiball tracking system,” Presence: Teleoperators and Virtual Environments, Vol. 10, No. 1, pp. 1–21, 2001.

- [9] H. Kato and H. Billinghurst: “Marker tracking and hmd calibration for a video-based augmented reality conferencing system,” Proc. 2nd IEEE/ACM Int. Workshop on Augmented Reality (IWAR’99), pp. 85–94, 1999.
- [10] 羽原寿和, 町田貴史, 小川剛史, 竹村治雄: “画像マーカを用いた屋内位置検出機構とその評価”, 電子情報通信学会技術研究報告, Vol. 102, No. 216, pp. 65–70, 2002.
- [11] L. Naimark and E. Foxlin: “Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker,” Proc. 1st IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2002), pp. 27–36, 2002.
- [12] G. Baratoff, A. Neubeck and H. Regenbrecht: “Interactive multi-marker calibration for augmented reality applications,” Proc. 1st IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2002), pp. 107–116, 2002.
- [13] C. Tomasi and T. Kanade: “Shape and motion from image streams under orthography: A factorization method,” Int. J. of Computer Vision, Vol. 9, No. 2, pp. 137–154, 1992.
- [14] J. Poleman and T. Kanade: “A paraperspective factorization method for shape and motion recovery,” Tech. Rep. CMU-CS-93-219, Carnegie-Mellon Univ., 1993.
- [15] P. Beardsley, A. Zisserman and D. Murray: “Sequential updating of projective and affine structure from motion,” Int. J. of Computer Vision, Vol. 23, No. 3, pp. 235–259, 1997.
- [16] H. S. Sawhney, Y. Guo, J. Asmuth and R. Kumar: “Multi-view 3d estimation and application to match move,” Proc. IEEE Workshop on Multi-view Modeling and Analysis of Visual Scenes, pp. 21–28, 1999.

- [17] G. Roth and A. Whitehead: “Using projective vision to find camera positions in an image sequence,” Proc. 13th Int. Conf. on Vision Interface, pp. 87–94, 2000.
- [18] M. Pollefeys, R. Koch, M. Vergauwen, A. A. Deknuydt and L. J. V. Gool: “Three-dimensional scene reconstruction from images,” Proc. SPIE, Vol. 3958, pp. 215–226, 2000.
- [19] U. Neumann and S. You: “Natural feature tracking for augmented-reality,” IEEE Transactions on Multimedia, Vol. 1, No. 1, pp. 53–64, 1999.
- [20] A. J. Davison, Y. G. Cid and N. Kita: “Real-time 3d slam with wide-angle vision,” Proc. IFAC Symp. on Intelligent Autonomous Vehicles, 2004.
- [21] 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.
- [22] D. W. Jacobs: “Linear fitting with missing data: Applications to structure-from-motion and to characterizing intensity images,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 206–212, 1997.
- [23] 出口光一郎: “透視投影条件下での因子分解法による多視点画像からの形状復元”, 情処研報 CVIM106-6, 1997.
- [24] A. I. Comport, E. Marchand and F. Chaumette: “A real-time tracker for markerless augmented reality,” Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR2003), pp. 36–45, 2003.
- [25] 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 (ISMAR2003), pp. 93–102, 2003.

- [26] 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.
- [27] M. Kourogi and T. Kurata: “Personal positioning based on walking locomotion analysis with self-contained sensors and wearable camera,” Proc. 2nd IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2003), pp. 103–112, 2003.
- [28] R. Cipolla, D. Robertson and B. Tordoff: “Image-based localization,” Proc. 10th Int. Conf. Virtual Systems and Multimedia (VSMM 2004), pp. 22–29, 2004.
- [29] Y. Kameda, T. Takemasa and Y. Ohta: “Outdoor see-through vision utilizing surveillance cameras,” Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR2004), pp. 151–160, 2004.
- [30] 大江統子, 佐藤智和, 横矢直和: “画像特徴点によるランドマークデータベースを用いた動画像からのカメラ位置・姿勢推定”, 電子情報通信学会 技術研究報告, PRMU2004-157, pp. 19–24, 2005.
- [31] C. Harris and M. Stephens: “A combined corner and edge detector,” Proc. Alvey Vision Conf., pp. 147–151, 1988.
- [32] M. A. Fischler and R. C. Bolles: “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” Comm. of the ACM, Vol. 24, pp. 381–395, 1981.
- [33] 出口光一郎: “射影幾何学による P n P 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.
- [34] 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.