

# 動画像とGPSによるハイブリッドな カメラ位置・姿勢推定法

横地 裕次 池田 聖 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

あらまし: 本稿では, 屋外環境下で取得した動画像とGPSによる位置情報からのハイブリッドなカメラ外部パラメータ推定手法を提案する. 提案手法では, 画像上で定義される誤差とGPSの位置情報から定義される誤差を同時に最小化することで, カメラ外部パラメータの推定結果に誤差が蓄積することを防ぐ. これにより, 動画像上で安定した特徴点の追跡及び, カメラ外部パラメータ推定の精度向上を図る. 実験では, シミュレーションによるカメラ外部パラメータの推定精度の評価及び, 屋外実環境下で取得した動画像とRTK-GPSによる位置情報から推定したカメラ外部パラメータを用いた仮想物体の重畳表示によって, 提案手法の有効性を確認した.

## A Hybrid Method for Estimating Extrinsic Camera Parameters Using a Video Camera and a GPS

Yuji YOKOCHI, Sei IKEDA, Tomokazu SATO, and Naokazu YOKOYA

Graduate School of Information Science, Nara Institute of Science and Technology

Abstract: This paper describes a novel method for estimating extrinsic camera parameters using an image sequence and position data acquired by a GPS (Global Positioning System). This method is based on tracking a number of image features and optimizing extrinsic camera parameters. The position data acquired by a GPS are used to prevent incorrect tracking of image features and to optimize extrinsic camera parameters over the whole image sequence. In this method, extrinsic camera parameters are estimated without accumulating an error even from a long image sequence. The validity of the proposed method has been demonstrated through experiments of estimating extrinsic camera parameters using both real data of outdoor scene and simulation data.

### 1 はじめに

現実環境を撮影した動画像からのカメラ外部パラメータの推定は, 複合現実感, 三次元形状復元, 新視点画像生成など様々な分野において基礎技術として利用される. これらの応用分野では, 近年, 広範囲な環境への手法の適用が試みられ [1-3], 広範囲に移動するカメラの外部パラメータを精度良く推定することが求められている. しかし, 動画像のみからカメラ外部パラメータを推定する手法 [4, 5] では, カメラ外部パラメータの推定結果に蓄積誤差が発生するため, 広範囲に移動する場合に著しく推定精度が低下する問題がある. また, 推定されるカメラ外部パラメータは, フレーム間での相対的な位置・姿勢関係を表すものであり, スケールについても不定である. そのため, 複数回に分けて撮影した動画像の外部パラメータの統合や,

閉曲線上の経路で撮影した動画像の外部パラメータ推定のように, 同一シーンが連続しないシーケンス上に現れる場合, 推定されるパラメータ間の不整合が問題となる.

これらの問題は, 従来から, 動画像と現実環境に関する事前知識の使用 [6-9] や動画像を用いずに複数のセンサから得られる値を直接カメラ外部パラメータとすること [10] により解決されてきた. 環境の事前知識を用いる手法では, 三次元位置関係が既知のランドマーク [6, 7] や, 三次元CADモデル [8, 9] が用いられている. 文献 [6, 7] の手法では, 三次元位置を事前に計測したランドマークと三次元位置が未知の特徴点を画像上で追跡し, その投影誤差を最小化することでカメラ外部パラメータを推定する. 文献 [8, 9] の手法は, 現実環境の三次元CADモデルの特徴点情報と画像上の特徴点を対応付けることで, カメラ外部パラメータ

を推定する．これらの手法では，ランドマークや三次元 CAD モデルの現実環境での三次元位置情報を用いることで，蓄積誤差の問題を解決している．しかし，ランドマークや三次元 CAD モデルに含まれる三次元位置情報を屋外で広範囲に取得することは困難であり，また，基準とする三次元情報の獲得のためには，人手による作業が必要であるという問題がある．

RTK-GPS，磁力計，2 軸角度センサーから得られる値を直接，動画像のカメラ外部パラメータとする手法 [10] では，RTK-GPS を用いることによって，誤差の蓄積しない数  $cm$  精度の位置情報を GPS の測位座標系で得ることができる．しかし，センサのみからカメラ外部パラメータを推定する手法では，カメラとセンサの位置・姿勢の関係を正確に計測しておく必要がある．また，一般的な GPS では，1 秒おきにしか位置情報を取得できないため，ビデオレートで正確な位置を算出することが難しいという問題がある．

本稿では，屋外環境下を広範囲に移動するカメラの外部パラメータを精度良く推定することを目的とし，動画像と RTK-GPS による位置情報からカメラ外部パラメータを推定する手法を提案する．提案手法では，動画像中の特徴点の追跡と GPS による位置情報を用いた最適化に基づき，カメラ外部パラメータを推定する．特徴点追跡では，文献 [7] の手法と同様に，特徴点の仮対応づけから推定される暫定的なカメラパラメータを用いて特徴点の画像上での探索範囲を限定し，誤追跡を防ぐ．パラメータの最適化では，動画像から推定されるカメラ外部パラメータに基づく特徴点の投影誤差及び，カメラ外部パラメータから推定される GPS 受信機の位置の推定誤差を同時に最小化することで，誤差の蓄積を防ぐ．本手法では，現実環境に関する事前知識や，GPS 受信機とカメラの間の正確なキャリブレーションを必要とせず，GPS の測位座標系におけるカメラ外部パラメータを推定できる特長がある．ただし，本手法では，カメラと GPS 受信機の位置・姿勢の関係は固定されているものとし，カメラの内部パラメータは既知とする．また，動画像と GPS の位置情報の同期はとれているものとする．

以下，2 章では，提案手法の基となる動画像からの特徴点追跡に基づくカメラ外部パラメータ推定手法 [7] の処理手順を説明し，GPS の位置情報によりパラメータの推定精度の向上が図れる部分を明らかにする．3 章では，動画像からの特徴点追跡と GPS による位置情報を用いた最適化に基づきカメラ外部パラメータを推定する手法について述べる．4 章では，シミュレーションによるカメラ外部パラメータの推定精度の定量的評価実験及び，実環境データからカメラ外部パラメータを推定する実験について述べる．5 章で，本稿のまと

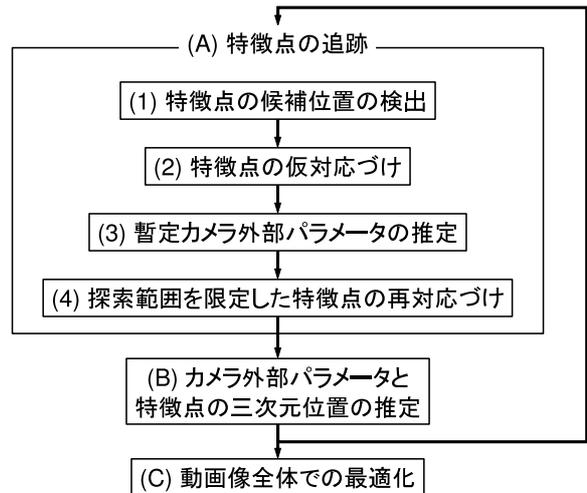


図 1: 特徴点追跡によるカメラ外部パラメータ推定の処理の流れ

めと今後の課題を述べる．

## 2 特徴点追跡によるカメラ外部パラメータ推定と GPS の利用

本章では，提案手法の基礎となっている特徴点追跡によるカメラ外部パラメータ推定手法 [7] の処理手順を概説し，GPS による位置情報を用いることにより，パラメータの推定精度の向上が図れる部分を明らかにする．ただし，文献 [7] の手法では，動画像中の一部にランドマークが写り，適当なキーフレーム上でそれらの画像上での位置が手動で与えられていることを前提としている．また，カメラのレンズ歪みや焦点距離などを含む内部パラメータは既知と仮定されている．

### 2.1 特徴点追跡によるカメラ外部パラメータの推定

文献 [7] の手法では，図 1 に示すように，各フレームに対する逐次処理として，(A) 特徴点の追跡，(B) カメラ外部パラメータと特徴点の三次元位置の推定を行う．次に，全てのフレームを用いた処理として，(C) 動画像全体での最適化を行う．以下では，各処理について説明する．

(A) 特徴点の追跡 特徴点を入力画像上で検出し，前フレームと現フレームでの対応関係を求める．特徴点には，回転・拡大縮小に強いエッジのコーナーなどの追跡が容易な点のみを検出する．さらに，特徴点の仮対応づけの結果から統計的手法によって誤追跡結果を排除し，それにより求まる暫定的なカメラ外部パラメータ

タを用いて、特徴点の対応関係を修正する．これにより、特徴点の周辺類似パターンとの誤対応を防止する．

具体的な処理手順としては、まず、(1) Harris オペレータで求められる特徴量が極大値となる位置を特徴点の候補位置とする．次に、(2) 前フレーム上の各特徴点の近傍領域を現フレーム上でパターンマッチングによって探索し、現フレームにおける仮の追跡位置を決定する．(3) 特徴点の仮の追跡位置から、LMeds 基準を用いたロバスト推定 [11] によって、暫定カメラ外部パラメータを推定する．最後に、(4) 暫定カメラ外部パラメータを用いて、前フレームにおけるすべての特徴点の三次元位置から、現フレーム上での出現位置を予測し、ステップ (2) よりも小さな探索窓を用いて対応づけを行い、追跡位置を修正する．

(B) カメラ外部パラメータと特徴点の三次元位置の推定 処理 (A) で求めた現フレーム (第  $i$  フレーム) 上の特徴点の投影誤差の和を最小化することで、カメラ外部パラメータ  $M_i$  と特徴点の三次元位置を推定する．特徴点  $j$  の第  $i$  フレーム上での投影誤差  $ER_{ij}$  は、画像上への投影座標  $\hat{q}_{ij}$  と画像上で検出される座標  $q_{ij}$  を用いて、以下のように表される．

$$ER_{ij} = |q_{ij} - \hat{q}_{ij}|^2 \quad (1)$$

(C) 動画全体での最適化 処理 (A), (B) を初期フレームから最終フレームまで繰り返した後に、動画全体で最適化を行うことで、累積的な誤差を最小化する．これにより、処理 (A), (B) の繰り返しで生じる蓄積誤差を最小化する．ここでは、全てのフレームのカメラ外部パラメータ  $M_i$  と全ての特徴点の三次元位置  $p_j$  をパラメータとし、以下の評価関数  $E$  を最小化する．

$$E = \sum_i (A_i \sum_j w_j ER_{ij}) \quad (2)$$

ただし、 $w_j$  は逐次処理で得た特徴点  $j$  の信頼度であり、ここでは、投影誤差  $ER_{ij}$  の分散の逆数とする． $A_i$  はフレームの信頼度による重みであり、キーフレームに大きな重みを与えることで、キーフレームのカメラ外部パラメータを真値とみなして動画全体を最適化する．

## 2.2 GPS による位置情報の利用による手法の拡張

前節の各処理 (A), (B), (C) において GPS による位置情報を用いることで、精度向上が図れる部分を明らかにする．まず、GPS による位置情報を用いるための条件について考察する．GPS による位置情報は、1

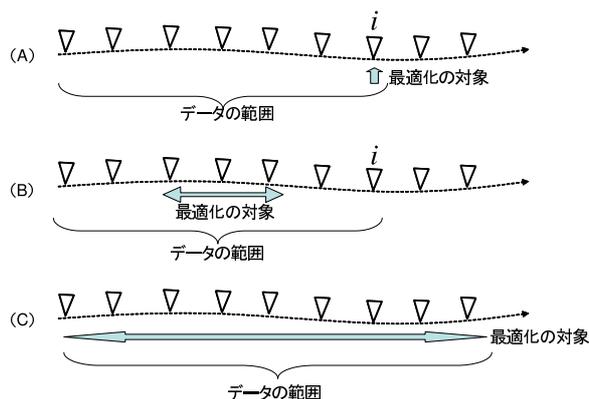


図 2: 拡張手法で用いるデータの範囲

秒間隔で得られ計測誤差が含まれる．このことから、1 地点のみの位置情報を用いると GPS の計測誤差の影響によりカメラ位置の推定精度が低下する恐れがある．そこで、文献 [7] の手法における各処理を GPS による位置情報を利用できる形に拡張する場合は、GPS の計測誤差を平均化するために、複数地点での位置情報を同時に用いる必要がある．そこで、GPS の位置情報を用いる拡張手法では、投影誤差の最小化処理において、ランドマークの代わりに GPS の位置情報を用いることで、画像上の誤差に加え、GPS の位置情報から定義される誤差を同時に考慮する．

次に、各処理での GPS による位置情報を用いた最適化による精度向上の見込み及び、計算量について考察する．図 2 に各処理において用いることのできるデータの範囲を示す．第  $i$  フレームにおける特徴点の追跡 (A) では、特徴点の仮対応づけ (2) と暫定カメラ外部パラメータの推定 (3) を 1 フレームごとに行うことにより、ロバストな特徴点追跡を実現している．GPS を用いた拡張手法において、この暫定カメラ外部パラメータの推定に GPS の位置情報を用いる場合、既に述べたように GPS の計測誤差を平均化するためには、一度に複数フレームを用いる最適化処理を、1 フレームごとに行うことになり、元的手法と比べて計算量が大幅に増大する．また、逐次処理においては、図 2(A) に示すように第  $(i+1)$  フレーム以降のフレームにおける特徴点の追跡情報及び、カメラ外部パラメータの初期値が利用できないため、GPS の位置情報に含まれる誤差の影響が大きく、処理 (A) で特徴点の追跡精度を向上させることは難しい．

カメラ外部パラメータと特徴点の三次元位置の推定 (B) では、第  $i$  フレームまでのすべてのカメラ外部パラメータを用いて、特徴点の三次元位置を推定するが、カメラ外部パラメータは第  $i$  フレームのみ推定する．

拡張手法においては、第  $i$  フレーム以前の複数フレームのカメラ外部パラメータを GPS の位置情報を用いて最適化することで、誤差の蓄積を防止できる。この最適化では、図 2(B) に示すように、第  $i$  フレーム以前の複数フレームのカメラ外部パラメータを同時に最適化することで、最適化の対象フレームの前後のシーケンスにおける特徴点の追跡情報を用いることができるため、パラメータの推定精度向上を図ることができる。さらに、数フレーム間隔で最適化することで、1 フレームごとの最適化に比べて計算量を削減できる。

全体最適化 (C) についても処理 (B) と同様に、GPS の位置情報を用いてカメラ外部パラメータを最適化することで累積誤差を最小化できると考えられる。この最適化は、逐次処理の終了後に 1 度だけ行えばよい。

以上のことから、第  $i$  フレームの逐次処理における特徴点追跡 (A) では、第  $i$  フレーム以降の特徴点追跡情報を用いることができず、精度の向上は期待できない。また、計算量の面で問題がある。処理 (B)、(C) では、最適化の対象フレームを第  $i$  フレーム以前とすることで前後の特徴点追跡情報を用いることができるため、精度の向上が期待できる。したがって、提案手法では、従来手法の処理 (B) 及び (C) を、GPS の位置情報を用いて拡張する。

### 3 動画と GPS の位置情報からのカメラ外部パラメータ推定

提案手法では、動画と GPS による位置情報から、特徴点の三次元位置、カメラと GPS 受信機の位置・姿勢の関係及び、カメラ外部パラメータを推定する。本手法は、逐次処理と全体での最適化処理の両方で GPS の位置情報を用いた最適化を行うことで、誤差の蓄積を防止し、精度良くカメラ外部パラメータを推定する。本手法は、ランドマークの三次元位置のような事前知識を必要としない。また、従来手法では計算機による一括処理が難しいような、広範囲にカメラが移動する環境を想定する。ただし、本手法では、カメラの内部パラメータは既知とする。また、カメラと GPS の受信機の位置・姿勢の関係は固定されているものとし、その距離については既知、角度については未知とする。さらに、動画と GPS による位置情報の同期が取れているものとする。

前章 2.2 に述べたように、処理 (A) に GPS による位置情報を用いると計算コストの面で問題があるため、提案手法では、カメラ外部パラメータと特徴点の三次元位置の推定 (B) 及び、全体での最適化 (C) に対応する処理において、GPS の位置情報を利用した共通の最

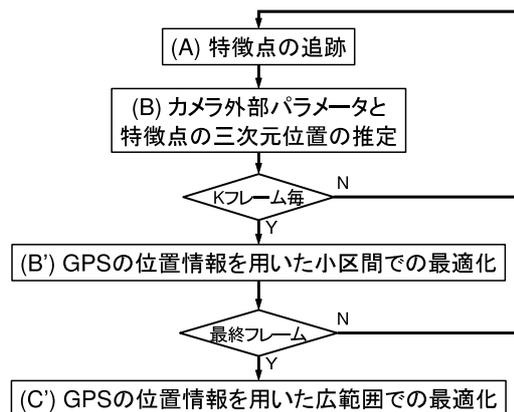


図 3: 提案手法の処理手順

適化を行う。図 3 に示すように、本手法ではまず、フレーム毎の処理として、(A) 特徴点の追跡、(B) カメラ外部パラメータと特徴点の三次元位置の推定を行い、次に一定フレーム間隔で (B') GPS による位置情報を用いた最適化を行う。処理 (B') では、蓄積誤差の影響によって後続のフレームで特徴点追跡の精度が低下することを防ぐために、小範囲での最適化を行う。これらを初期フレームから最終フレームまで繰り返す。最後に処理 (C) として動画全体で最適化する。処理 (C) では、カメラ外部パラメータの推定精度の向上のために、処理 (B') よりも大きな区間で多くの GPS の位置情報を用いて最適化を行う。ただし、本研究では、長時間の動画を投入として想定しているため、動画全体を一括で最適化することが困難な場合には、全フレームを処理 (B') よりも十分に大きな区間に分割し、それぞれを独立に最適化する。

以下では、まず処理 (B')、(C) で共通に用いる GPS による位置情報を用いた最適化について述べる。次に、処理 (B')、(C) についてそれぞれ詳述する。

#### 3.1 GPS による位置情報を用いた最適化

本研究では、画像上で定義される誤差と GPS の位置情報から定義される誤差を同時に最小化することで、カメラ外部パラメータの推定結果に生じる蓄積誤差の低減を図る。まず、図 4 に示す GPS 受信機の位置推定誤差  $EG_i$  について定義する。第  $i$  フレームのカメラ外部パラメータ  $M_i$  とそれに対応する GPS による位置  $g_i$  及び、カメラ座標系での GPS 受信機の位置  $d$  には、以下の関係が成り立つ。

$$M_i g_i = d \quad (i \in F_g) \quad (3)$$

ただし、 $F_g$  は、GPS の計測時刻に対応する動画のフレームの集合を表す。ここで、カメラ外部パラメー

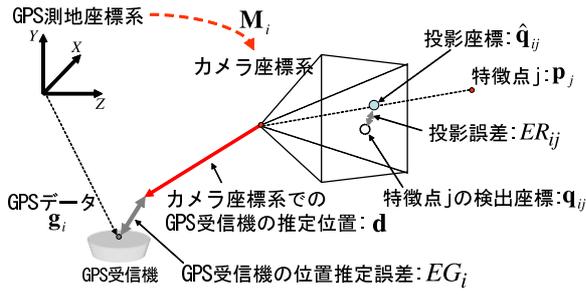


図 4: 座標系と誤差の定義

タ  $M_i$  に推定誤差があると,  $M_i$  によって推定される GPS 受信機の位置と実際の GPS による計測位置は一致しない. このカメラ外部パラメータ  $M_i$  から推定される GPS 受信機の位置の推定誤差を誤差関数  $EG_i$  とし, 以下のように定義する.

$$EG_i = |d - M_i g_i|^2 \quad (4)$$

本研究では, カメラ外部パラメータ推定に用いる最適化の誤差関数  $E$  を, 前章 2.1 節の式 (1) で示した画像上の投影誤差  $ER_{ij}$  及び, 式 (4) の GPS 位置の誤差  $EG_i$  を用いて, 以下のように定義する.

$$E = \frac{\omega}{|\mathbf{F}_g|} \sum_{i \in \mathbf{F}_g} EG_i + \frac{(1-\omega)}{\sum_i |S_{p_i}|} \sum_i (A_i \sum_{j \in S_{p_i}} ER_{ij}) \quad (5)$$

ただし,  $\omega$  は  $EG_i$  及び  $ER_{ij}$  に対する重み係数,  $A_i$  は各フレームに対する重み係数,  $S_{p_i}$  は第  $i$  フレームの画像上で検出される特徴点の集合を表す. 重み係数  $\omega$  は, GPS による位置情報の信頼度に基づき決定する. 式 (5) の誤差関数  $E$  を用いた最適化では, カメラ外部パラメータ  $M_i$ , 特徴点の三次元位置  $p_j$  及び, カメラと GPS 受信機の位置関係  $d$  をパラメータとして, 誤差関数  $E$  を勾配法によって最小化する.

ただし, 誤差関数  $E$  には, 局所解が数多く存在するため, 単純な勾配法では, 大域最小解を得ることが難しい. そこで, 特徴点の三次元位置を最適化する際に, 各フレームの重み  $A_i$  を動的に変化させることで, 局所解を回避する. 具体的には, GPS の位置情報があるフレームの重みを大きく設定すること及び, すべてのフレームの重みを等価に設定することを, 最適化の一定ステップごとに交互に繰り返す. 重み  $A_i$  の変動による最適化の効果を図 5 に示す. 図 5 は, 真値に誤差を付加して作成したカメラ外部パラメータを勾配法で最適化した際の, 推定パラメータの誤差の変化を表す. グラフの傾きが不連続な部分が重みの変動を行った瞬間である. 同図から, 重み変動を行わない場合は, カメラ位置の推定誤差が  $200\text{mm}$  程度のときに局所解に

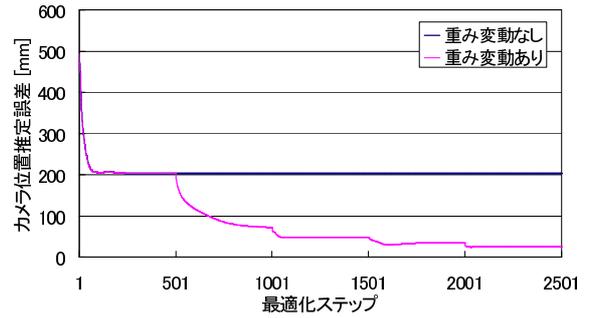


図 5: 重み変動の効果

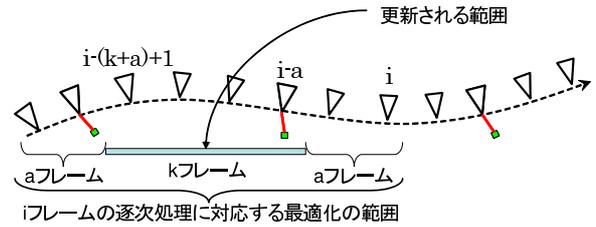


図 6: 最適化の範囲

陥っているが, 重み変動を行うことにより, 局所解が回避されていることが分かる.

### 3.2 GPS の位置情報を用いた小区間での最適化 (B')

処理 (B') では, 処理 (A), 処理 (B) で得られるカメラ外部パラメータを初期値として, 前節で述べた最適化により, GPS の位置情報を反映させたカメラ外部パラメータを算出する. さらに, この処理で算出されたパラメータを処理 (A) に反映させることで, 追跡精度の向上を図る.

具体的には, 図 6 に示すように, 処理 (A) 及び処理 (B) で処理の対象となる第  $i$  フレームに対して, 第  $(i-(k+2a)+1)$  フレームから第  $i$  フレームの情報を用いて小区間での同時最適化を行い, 第  $(i-(k+a)+1)$  フレームから第  $(i-a)$  フレームの  $k$  フレーム分の情報のみを更新する. これにより, 更新するフレームに対して, 前後のフレームにおける特徴点の追跡結果及び, GPS の位置情報を反映させることができる. さらに, 一度に複数のフレーム ( $k$  フレーム) を更新し,  $k$  フレーム間隔で本処理 (B') を行うことで, 計算量の増大を軽減する.

### 3.3 GPS の位置情報を用いた広範囲での最適化 (C')

前節で述べた小区間でのカメラ外部パラメータの最適化では, 最適化に利用できる範囲は逐次処理の第  $i$

フレームまでである．ここでは，逐次処理終了後に全てのフレームにおける特徴点の追跡情報と GPS の位置情報を用いることで，GPS の局所的な誤差を平均化し，カメラ外部パラメータの推定精度の向上を図る．ただし，本研究では，従来手法において計算機による一括処理が難しいような長時間の動画を入力として想定している．そこで，全てのフレームを一括で処理することが困難な場合には，動画像全体を処理 (B') の最適化範囲よりも長い複数のシーケンスに分割し，各シーケンスで独立に 3.1 節で述べた最適化を実行する．

本手法では，GPS の位置情報によりカメラ外部パラメータの絶対座標が推定されているため，複数シーケンスの統合に特別な処理は不要である．

## 4 カメラ外部パラメータの推定実験

本章ではまず，計算機シミュレーションにより，提案手法の全体最適化 (C') で推定されるカメラ外部パラメータの推定精度を定量的に評価する．次に，実環境で取得した動画像と GPS による位置情報から，カメラ外部パラメータを推定する．また，推定されたカメラ外部パラメータを用いて動画像に仮想物体を合成するマッシュアップを行うことで，カメラ外部パラメータが正しく推定されていることを確認する．

### 4.1 シミュレーションによるカメラ外部パラメータの推定精度評価

本シミュレーションでは，提案手法の処理 (C') によって推定されるカメラ外部パラメータの推定精度を定量的に評価する．本実験で用いる仮想環境を図 7 に示す．ここで，図中の曲線はカメラの移動経路を，錘台は 5 フレーム毎のカメラの姿勢を表す．また，グリッドの一边は実環境での 20m に相当する．仮想環境中の特徴点には，事前に屋外環境を撮影した全方位動画像を入力として，佐藤らの手法 [12] を用いて三次元復元を行い，これによって推定された特徴点の三次元位置を本シミュレーションの真値として用いた．仮想カメラは，同図に示すように a 地点から b 地点へ移動させ，仮想的に 990 枚の特徴点検出画像を取得した．ただし，仮想カメラの内部パラメータには，次節で述べる実環境で用いたカメラと同じものを用いた．本環境では 1 フレームあたり平均 200 個の特徴点が画像上で観測される．また，カメラ座標系における GPS 受信機の位置を (600,600,600)(単位は mm) と設定し，GPS の位置情報の真値を 1 秒間隔 (15 フレーム毎) で作成した．さらに，特徴点の画像上での検出座標  $q_{ij}$ ，GPS の位置情報  $g_i$  及び，カメラ外部パラメータ  $M_i$  の初期値は，表 1 に示すガウスノイズをそれぞれの真値に付加して

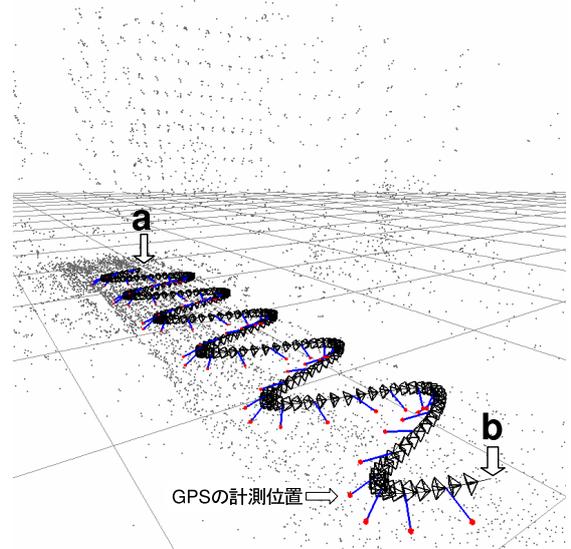


図 7: シミュレーション環境

表 1: 初期値に付加するガウスノイズの標準偏差

特徴点の検出座標	1 画素
GPS の計測誤差	30mm
カメラ位置	500mm
カメラ光軸の向き	0.02rad

作成した．3.1 節で定義した誤差関数  $E$  中の重み係数  $\omega$  は，予備実験により経験的に決定した値 ( $\omega=10^{-9}$ ) を用い，カメラと GPS 受信機間の距離は計測されているものと仮定し，角度については未知とした．

以上の条件で，特徴点の三次元位置，カメラ外部パラメータ，カメラと GPS 受信機の位置関係をパラメータとし，誤差関数  $E$  を最小化した．図 8 にカメラ位置の推定誤差，図 9 にカメラ光軸の向きの推定誤差を示す．実験の結果，カメラ外部パラメータは，カメラ位置に関して平均誤差 32.9mm 程度，カメラ光軸の向きに関して平均誤差 0.00355rad 程度で推定された．推定されたカメラ外部パラメータの誤差は，フレームが進んでも増加しないことから，GPS による位置情報を用いた最適化により，誤差の蓄積が防止されていると考えられる．

### 4.2 実環境データを用いたカメラ外部パラメータの推定

本節では，提案手法によって安定に特徴点の追跡及びカメラ外部パラメータの推定が行われることを確認するため，屋外実環境下で取得した動画像と GPS の位置情報を用いて実験を行う．実験では，ワイドレンズ (Sony VCL-HG0758) を取り付けたビデオカメラ (Sony

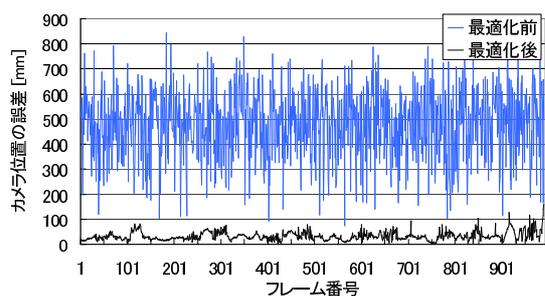


図 8: カメラ位置の推定誤差

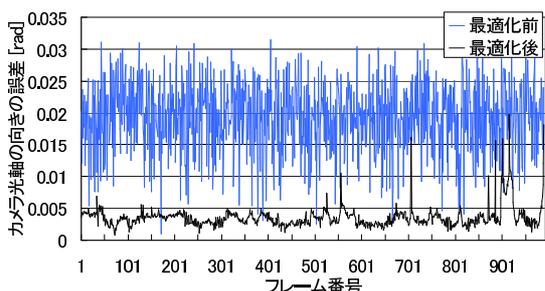


図 9: カメラ光軸の推定誤差

DSR-DP-150) と GPS 受信機 (Nikon LogPakII) を図 10 に示すように車に固定して取り付けた。そして、約  $1.1\text{km}$  の距離を平均時速  $16.5\text{km}$  の速度で走行し、得られた 3600 枚の画像 (解像度  $720 \times 480$  画素,  $14.985\text{fps}$ , プログレッシブ撮影) 及び、1 秒間隔で計測された GPS による位置情報を入力として用いた。ただし、動画像と GPS の位置情報は手動で開始時刻を同期させ、カメラの内部パラメータは、Tsai の手法 [13] によりあらかじめ得た値を用いた。また、カメラと GPS 受信機間の設置距離は、手動で計測した  $1020\text{mm}$  を設定し、角度については未知とした。初期フレームにおけるカメラ外部パラメータは、手動で大まかな姿勢を与え、位置に関しては GPS の位置情報を用いた。逐次処理で行う最適化 (B') の範囲は、 $k = 5$ 、 $a = 22$  フレームと設定した。処理 (C') では、動画像全体を対象とし、 $k = 3600$ 、 $a = 0$  フレームと設定した。

実験の結果、逐次処理では、初期フレームから最終フレームまで途切れることなく、1 フレームあたり平均約 300 個の特徴点が追跡された。提案手法によって推定された特徴点の三次元位置とカメラ外部パラメータを図 11 に示す。前節と同様に、図中の曲線は推定されたカメラの移動経路を、錘台は 5 フレーム毎のカメラの姿勢を表す。同図から分かるように、カメラパスは滑らかに推定されており、GPS による位置情報が計測されていないフレームにおいても連続的にカメラ外部パラメータが推定されている。

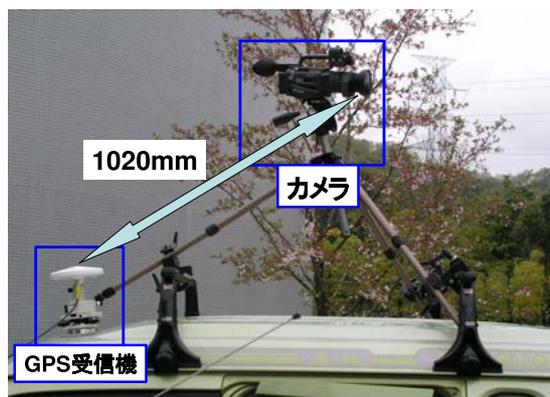


図 10: ビデオカメラと GPS 受信機の位置関係

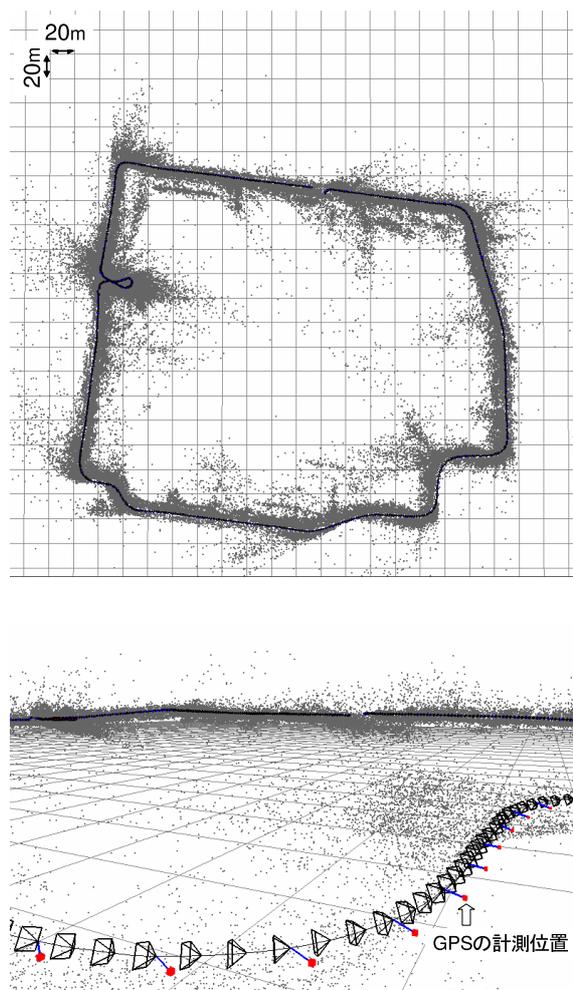


図 11: カメラ外部パラメータの推定結果

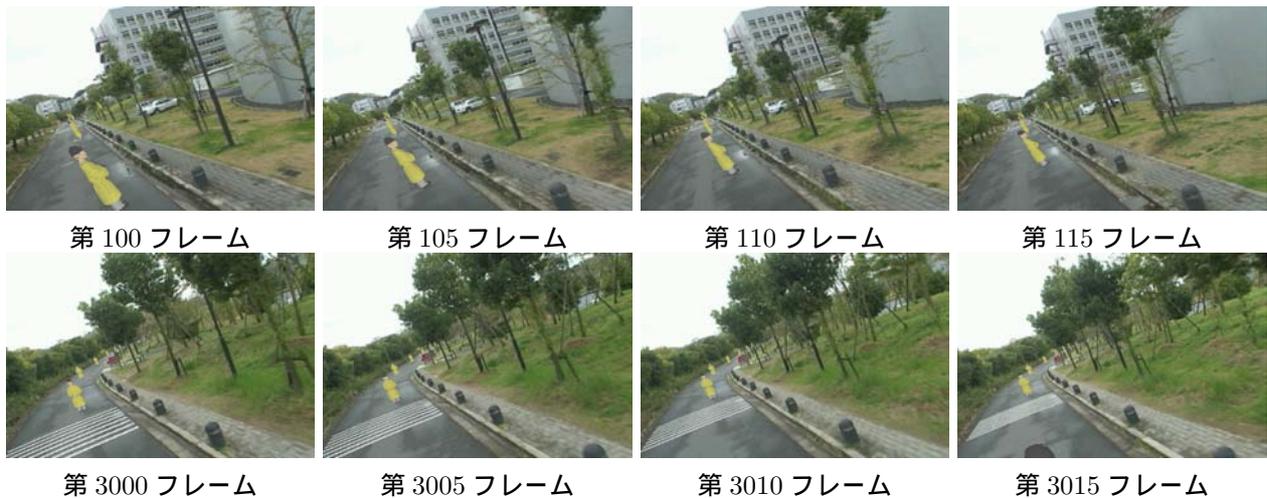


図 12: 推定されたカメラ外部パラメータを用いたマッチムーブの結果

推定されたカメラ外部パラメータを用いて仮想物体を動画に合成した結果 (マッチムーブ) を図 12 に示す。仮想物体は, GPS による計測位置から計算される道路面上に配置した。仮想物体が現実環境中でほぼ一定の位置に合成されていることから, カメラ外部パラメータが正しく推定されていることを確認できる。

## 5 まとめ

動画と GPS による位置情報からカメラ外部パラメータを推定する手法を提案した。提案手法では, 画像上で定義される特徴点の投影誤差と GPS による位置情報によって定義される位置推定誤差を同時に最小化することで, 推定誤差の累積しないカメラ外部パラメータを全てのフレームで推定することができる。

シミュレーション実験では, 提案手法で推定されるカメラ外部パラメータの推定精度を定量的に評価した。実験の結果, カメラ位置に関して平均誤差  $32.9\text{mm}$  程度, カメラ光軸の向きに関して平均誤差  $0.00355\text{rad}$  程度でカメラ外部パラメータ推定が可能であることを確認した。また, 実環境データを用いた実験では, 提案手法によって安定に特徴点が追跡されることを確認した。また, マッチムーブで仮想物体が現実環境で一定の場所に合成され続けることから, カメラ外部パラメータが正しく推定されていることを確認した。今後の課題として, 今回手動で与えた初期フレームにおけるカメラ姿勢の自動推定及び, 動画と GPS による位置情報の同期の自動化があげられる。

## 参考文献

[1] S. Feiner, B. MacIntyre, T. Höllerer and A. Webster: "A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment," Proc. 1st IEEE Int. Symp. on Wearable Computers

(ISWC1997), pp. 208–217, 1997.

[2] T. Asai, M. Kanbara and N. Yokoya: "3d modeling of wide area outdoor environments by integrating omnidirectional range and color images," Proc. 3rd IEEE and ACM Int. Symp. on Mixed Augmented Reality (ISMAR2004), pp. 264–265, 2004.

[3] 遠藤, 片山, 田村, 廣瀬: "写実的な広域仮想空間構築のための画像補間手法", 日本バーチャルリアリティ学会論文誌, 7, pp. 185–192, 2002.

[4] C. Tomasi and T. Kanade: "Shape and motion from image streams under orthography: A factorization method," Int. Journal of Computer Vision, 9, 2, pp. 137–154, 1992.

[5] M. Pollefeys, R. Koch, M. Vergauwen, B. Dekenuydt and L. V. Gool: "Three-dimensional scene reconstruction from images," Proc. SPIE, 3958, pp. 215–226, 2000.

[6] A. J. Davison: "Real-time simultaneous localisation and mapping with a single camera," Proc. 9th IEEE Int. Conf. on Computer Vision, 2, pp. 1403–1410, 2003.

[7] 佐藤, 神原, 横矢, 竹村: "マーカと自然特徴点の追跡による動画からのカメラ移動パラメータの復元", 電子情報通信学会論文誌 (D-II), J86-D-II, 10, pp. 1431–1440, 2003.

[8] A. I. Comport, É. Marchand and F. Chaumette: "A real-time tracker for markerless augmented reality," Proc. 2nd ACM/IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR2003), pp. 36–45, 2003.

[9] L. Vacchetti, V. Lepetit and P. Fua: "Combining edge and texture information for real-time accurate 3d camera tracking," Proc. 3rd IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2004), pp. 48–57, 2004.

[10] S. Güven and S. Feiner: "Authoring 3d hypermedia for wearable augmented and virtual reality," Proc. 7th IEEE Int. Symp. on Wearable Computers (ISWC2003), pp. 118–126, 2003.

[11] 栗田, 坂上: "ロバスト統計とその画像理解への応用", 画像の認識・理解シンポジウム (MIRU2000) 講演論文集, Vol. 1, pp. 65–70, 2000.

[12] 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 (ECCV2004), 2, pp. 326–340, 2004.

[13] 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.