

動画画像と GPS による位置情報を用いたカメラ外部パラメータの推定

横地 祐次^{*†} 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 〒 630-0192 奈良県生駒市高山町 8916-5

E-mail: †{sei-i,tomoka-s,yokoya}@is.naist.jp

あらまし 本論文では、屋外環境下で広範囲にカメラを移動し取得した動画画像と GPS による位置情報からカメラの外部パラメータを推定する手法を提案する。提案手法は、動画画像中の特徴点の追跡により外部パラメータを推定する手法を基に、GPS による位置情報を用いた外部パラメータの最適化により、累積的な推定誤差を軽減するものである。さらに、GPS の位置情報を用いて推定された暫定的な外部パラメータを特徴点の追跡処理にフィードバックすることにより、特徴点の追跡自体についても精度向上を図る。シミュレーションによる精度評価実験および屋外実環境下で取得した動画画像と GPS による位置情報を用いた実験により、環境中の三次元位置が既知な点を用いて累積誤差を軽減する手法と同程度の精度で外部パラメータが推定可能であることを確認した。

キーワード GPS, カメラ外部パラメータ, 特徴点追跡

Extrinsic Camera Parameter Estimation Using an Image Sequence and Position Data Acquired by GPS

Yuji YOKOCHI^{*†}, Sei IKEDA[†], Tomokazu SATO[†], and Naokazu YOKOYA[†]

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

Takayama-cho 8916-5, Ikoma-shi, Nara, 630-0192 Japan

E-mail: †{sei-i,tomoka-s,yokoya}@is.naist.jp

Abstract This paper describes a novel method for estimating extrinsic camera parameters using an image sequence and sparse position data acquired by GPS. This method is based on tracking a number of image features and optimizing extrinsic parameters using position data from GPS. The position data are actually used to prevent the algorithm from incorrectly tracking image features. Moreover, the position data are also used for optimizing extrinsic parameters over the whole image sequence. The proposed method allows us to estimate extrinsic parameters accurately with little accumulated estimation errors even from a long image sequence. The scale of estimated extrinsic parameters is automatically determined by position data acquired by GPS. The validity of the proposed method is demonstrated through experiments of estimating extrinsic parameters for both synthetic and real outdoor scenes.

Key words GPS, extrinsic camera parameters, feature tracking

1. はじめに

現実環境を撮影した動画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ける基礎技術として利用されている。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], [2], 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、動画画像のみから外部パラメータ

を推定する手法 [3], [4] では、外部パラメータの推定結果に累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、スケールについては不定である。そのため、複数回に分けて撮影した動画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、動画画像と現実環境に関する事前

* 現在, 株式会社 本田技術研究所

知識の使用 [5] ~ [8] 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外部パラメータとすること [9] 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、三次元的な位置関係が既知のランドマーク群 [5], [6] や、三次元 CAD モデル [7], [8] が用いられている。文献 [5], [6] の手法では、三次元位置が既知のランドマークと三次元位置が未知の画像特徴点 (以後、特徴点) をフレーム毎に追跡し、その再投影誤差を最小化することで外部パラメータを推定する。文献 [7], [8] の手法では、現実環境の三次元 CAD モデルの輪郭と入力画像上のエッジを対応付けることで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。

一方、RTK-GPS、磁力計、2 軸角度センサーの値から算出される値を、直接動画像のカメラ外部パラメータとする手法 [9] 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 (以後、GPS 測位値) を獲得できる。しかし、センサの組み合わせによってカメラの外部パラメータを推定する手法では、カメラとセンサ類の位置関係が必要であるが、これを正確に計測することは難しい。また、一般的な GPS では、位置情報の取得レートが 1Hz であり動画像よりも低く、カメラの運動に含まれる高周波成分を復元することは難しい。

そこで本論文では、これらの問題を解決するために、屋外環境下を広範囲に移動するカメラで取得した動画像と GPS 測位値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 測位値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。特徴点の追跡では、文献 [6] 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外部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。GPS 測位値を用いた最適化では、動画像から推定される外部パラメータに基づく特徴点の再投影誤差及び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外部パラメータを推定できる特長がある。そのため、動画像に CG を重畳表示する景観シミュレーションや映画の特撮などの分野へ提案手法を適用することで、環境の事前知識を獲得するのに要する人的コストの削減に貢献できると考えられる。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、動画像と GPS 測位値の同期はとれているものとする。また、本論文で扱う全てのカメラの内部パラメータは既知であり、予めカメラのレンズ歪みや焦点距離などを考慮したキャリブレーションが行われているものとする。

以下、2 章では、提案手法の基礎である文献 [6] の手法について概説し、3 章では、提案手法の流れ及び文献 [6] の手法との相

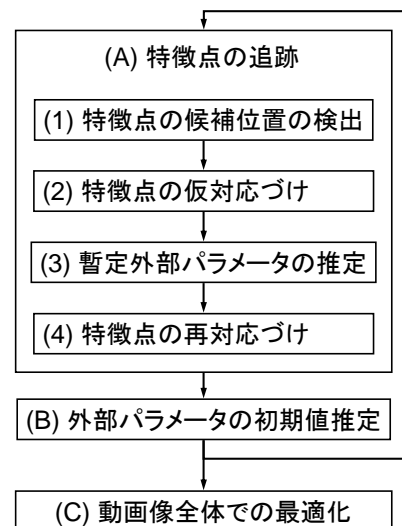


図 1 特徴点追跡による外部パラメータ推定の処理の流れ

違点について詳述する。4 章では提案手法の有効性を確認するためにシミュレーションによる実験及び実環境のデータを用いた実験を示し、最後に 5 章で本論文をまとめる。

2. 特徴点追跡による外部パラメータ推定

本章では、提案手法の基礎となる特徴点追跡によるカメラ外部パラメータ推定手法 [6] の処理手順を概説する。ただし、佐藤らによる文献 [6] 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人手により与えられていることを前提としている。文献 [6] 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) 外部パラメータの初期値推定を行う。これらの特徴点の追跡結果と推定した外部パラメータの初期値を用いて、全てのフレームに対する処理として (C) 動画像全体での最適化を行う。以下では、各ステップの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・拡大縮小等の画像の変形に対して頑健に同一の位置が検出されるエッジの角や交点を用いる。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外部パラメータを用いて、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 [10] 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。ただし、ランドマークに対しては、上述の処理 (1) の代わりに、人手による作業でキーフレーム上でのランドマークの位置が与えられ、以降の処理については特徴点と同様に扱う。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、暫定外部パラメータ

を推定する．最後に，(4) 暫定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) 外部パラメータの初期値推定

処理 (A) で求めた現フレーム (第 i フレーム) 上の特徴点の再投影誤差の二乗和を最小化することで特徴点の三次元位置，及び外部パラメータである回転 R_i および並進 t_i を推定する．特徴点 j の第 i フレーム上での再投影誤差 Φ_{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．

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

(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，全てのフレームの外部パラメータ R_i ， t_i と全ての特徴点 j の三次元位置をパラメータとし，以下の評価関数 E を最急降下法により最小化する．

$$E = \sum_i \mu_i \sum_j w_j \Phi_{ij}^2 \quad (2)$$

ただし， w_j は逐次処理で得た特徴点 j の信頼度であり，再投影誤差 Φ_{ij} の分散の逆数とする． μ_i はフレームの信頼度による重みであり，三次元位置が既知のランドマークが写っている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．

3. GPS 測位値を用いた外部パラメータ推定

提案手法では，動画と GPS 測位値から，特徴点の三次元位置，カメラと GPS 受信機の位置関係及び，カメラの外部パラメータを推定する．本手法は，特徴点追跡と全体での最適化処理の両方に GPS の位置情報を反映させることで，誤差の蓄積を防止し，精度良く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．カメラと GPS の受信機の位置関係は固定されているものとし，容易に計測できる 2 センサ間の距離については既知，角度については未知であるとする．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．

提案手法では，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 測位値を利用し，共通の最適化を行う．図 2 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の初期値推定を行い，次に一定フレーム間隔 k で (B') GPS 測位値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狭区間での最適化を行う．

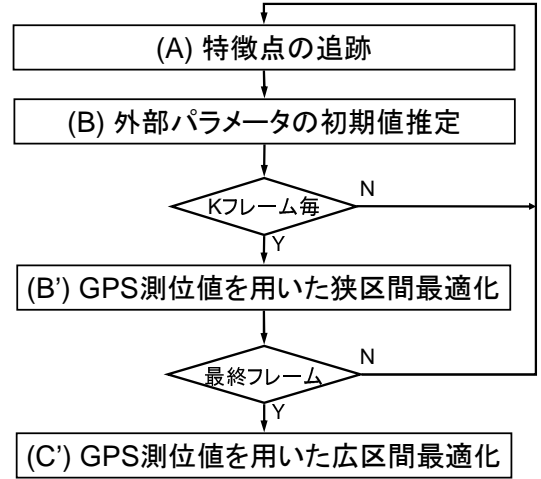


図 2 提案手法の処理手順

最後に，動画全体での最適化 (C) に対応する処理として，処理 (C') では，処理 (B') よりも広い区間で多くの GPS 測位値を用いて外部パラメータの最適化を行う．

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 測位値を用いた最適化

処理 (B') および処理 (C') では，特徴点の追跡 (A) で得られた特徴点の画像上での座標及び GPS 測位値から直前の処理 (処理 (B') については処理 (B)，処理 (C') については処理 (B')) で得られた外部パラメータを初期値として新たな外部パラメータを算出する．ここではまず，カメラと GPS 受信機の幾何学的関係をモデル化し，次に，GPS 受信機に関する誤差 Ψ_i を定義する．さらに，GPS 受信機に関する誤差 Ψ_i と画像の特徴点の再投影誤差 Φ を組み合わせて，新たな誤差関数 E' を定義し，その最小化方法について説明する．

一般に，GPS 測位値と外部パラメータの推定結果に誤差が含まれない場合，第 i フレームの外部パラメータ R_i ， t_i とそれに対応する GPS 測位値 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．

$$R_i g_i + t_i = d \quad (i \in \mathcal{F}) \quad (3)$$

ただし， \mathcal{F} 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，GPS 測位値 g_i 及び外部パラメータ R_i ， t_i に推定誤差があると，式 (3) は成り立たない．本研究では，外部パラメータ R_i ， t_i から推定される GPS 受信機の位置と GPS 測位値との差を GPS 受信機に関する誤差関数 Ψ_i とする．

$$\Psi_i = |R_i g_i + t_i - d| \quad (4)$$

次に，新たな誤差関数 E' を定義し，GPS 受信機に関する誤差 Ψ_i と画像特徴点の再投影誤差 Φ の組み合わせ方について説明する．ここでは，外部パラメータ推定に用いる最適化の誤差関数 E' を，前章の式 (1) で示した画像上の再投影誤差 Φ_{ij} 及び，式 (4) の GPS 位置の誤差 Ψ_i を用いて以下のように定義

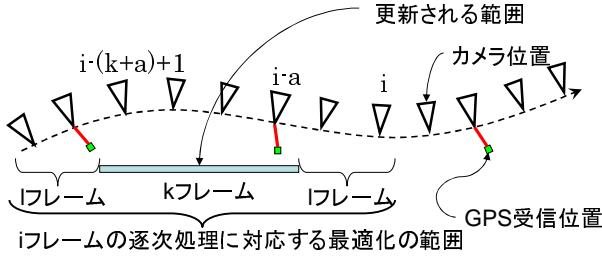


図 3 最適化の範囲

する．

$$E' = \frac{\omega}{|\mathcal{F}|} \sum_{i \in \mathcal{F}} \Psi_i^2 + \frac{1}{\sum_i |S_i|} \sum_i \mu_i \sum_{j \in S_i} w_j \Phi_{ij}^2 \quad (5)$$

ただし， ω は Ψ_i 及び Φ_{ij} に対する重み係数， μ_i は各フレームに対する重み係数， S_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．式 (5) の誤差関数 E' を用いた最適化では，外部パラメータ R_i ， t_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，式 (5) の誤差関数 E' には，局所解が多数存在するため，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，最適化の各フレームの重み μ_i を変化させることで，局所解を回避する．具体的には，GPS 測位値があるフレームの重みを大きく設定することおよび，すべてのフレームの重みを等価に設定することを，最適化処理が収束するステップごとに交互に繰り返す．誤差関数 E' の重み μ_i を変化させても大域最小解を与えるパラメータに大きな差はなく，局所解の位置のみが大きく変化するため，局所解を回避できると考えられる．

3.2 GPS 測位値を用いた狭区間最適化 (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，前節で定義した E' を最小化することで，GPS 測位値を反映させた外部パラメータ及び特徴点の三次元位置を算出し，処理 (A) にフィードバックさせることで，特徴点追跡の精度向上を図る．

具体的には，図 3 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，処理 (B') では，第 $i - (k + 2l) + 1$ フレームから第 i フレームの特徴点追跡結果および GPS 測位値を用いて最適化を行い，第 $i - (k + l) + 1$ フレームから第 $i - l$ フレームの k フレーム分の外部パラメータのみを更新する．これにより，更新する区間に対して，前後のフレームにおける特徴点の追跡結果及び，GPS 測位値を最適化に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．ただし，多数の追跡済みの特徴点がある場合，推定結果を更新する区間のフレームに写るように， l は数十フレームに設定する．また，特徴点追跡が外部パラメータの蓄積誤差により破綻する前に GPS 測位値を反映させる必要があるため， k は数フレームに設定する．

3.3 GPS 測位値を用いた広範囲最適化 (C')

処理 (C') では，前章で述べた手法 [6] の動画像全体での最適

化 (C) と同様に，処理 (B') よりも広い区間で特徴点の追跡結果及び GPS 測位値を用いて，外部パラメータの最適化を行うことにより，両情報に含まれる局所的な誤差の平均化を図る．ただし，本研究では，長時間かつ広範囲にカメラを移動させて撮影した動画像を入力として想定している．このような場合，時間的に十分離れた 2 フレームに関する外部パラメータや特徴点の三次元位置などの情報は，互いに影響しないと考えられる．そこで処理 (B') での最適化範囲 $k + 2l$ よりも長い複数の $k' + 2l'$ フレームのシーケンスに動画像を分割し，各シーケンスを独立に最適化する．ただし，処理 (B') と同様に最適化区間に，外部パラメータを更新する区間 k' フレームとその前後に更新しない区間 l' フレームを設け，区間 l' フレームを十分大きくとることで，各シーケンスの最適化における独立性を確保する．ただし，十分な数の GPS 測位値を用いて測位誤差を平均化するために， l' は数百フレームに設定する．また，分割されたシーケンス間の独立性を確保するために， k' についても数百フレームに設定する．

4. 実験

本章では，まず，式 (5) で定義した誤差関数の最小化によって，最適化処理 (C') で推定される外部パラメータの精度を定量的に評価するために，実環境から作成したデータを用いたシミュレーション実験を行う．次に，提案手法における特徴点検出位置及び GPS 測位値の誤差モデルや特徴点追跡の精度を含めて，提案手法の有効性を確認するために，実環境で取得した動画像と GPS 測位値から外部パラメータを推定する．また，推定された外部パラメータを用いて動画像上に仮想物体を重畳表示するマッチムーブを行うことで，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．最後に，動画像全体の一括最適化によって推定される外部パラメータと，動画像を複数のシーケンスへ分割して最適化し，最後に統合して得られる外部パラメータの比較を行うことで，分割による最適化の有効性を確認する．

4.1 シミュレーションによる精度評価

シミュレーションによる評価では，提案手法の処理 (C') によって推定される最終的な外部パラメータの推定精度を定量的に評価する．また，ランドマークを用いる手法と精度を比較することにより，提案手法の有効性を確認する．

本シミュレーションで用いる仮想環境を図 4 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し，点群は仮想環境中に配置した特徴点の三次元座標を表す．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．このときのカメラの移動経路とは異なる経路で仮想カメラを同図に示すように a 地点から b 地点へ移動させ，仮想的に 990 フレーム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一のものを用いた．本シ

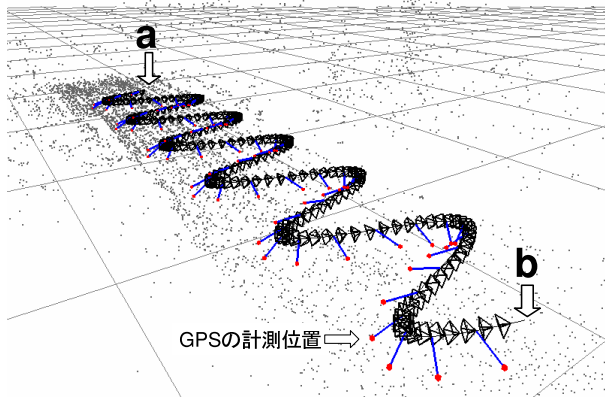


図 4 シミュレーション環境

シミュレーションにおける入力では、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た。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 mm) と設定し、GPS 測位値には、上記 GPS 受信機の位置に誤差を付加し、これを 1Hz に相当するフレーム間隔として 15 フレーム毎に与えた。特徴点の画像上での検出座標 q_{ij} および、GPS 測位値 g_i に含まれる誤差を表 1 に示す。また、外部パラメータ R_i, t_i の初期値は、表 2 に示す誤差を真値に付加して作成した。3.1 節で定義した式 (5) 誤差関数 E' 中の重み係数 ω は、両誤差関数 Φ, Ψ の単位の違いを考慮して 10^{-9} とした。また、式 (5) 中の各フレームの重み係数 μ_i については、GPS 測位値が得られていないフレームの場合 1.0 とし、GPS 測位値が得られているフレームの場合は、前章 3.1 節で述べた最適化処理において 500 ステップ毎に 2.0 と 1.0 を交互に用いた。逐次処理中に行う最適化処理 (B') においては、更新対象フレーム $k = 5$ 、重複フレーム $l = 22$ フレームと設定した。処理 (C') では、 $k' = 3600$ 、 $l' = 0$ とし、動画全体を一括で最適化した。初期フレームにおける外部パラメータは、手で大まかな姿勢を与え、位置に関しては GPS 測位値を用いた。カメラと GPS 受信機間の距離は既知とし、角度については未知とした。また、提案手法の精度評価のために、ランドマークを用いる手法 [6] によって上述と同様の条件で外部パラメータを推定した。ただし、手法 [6] においては、キーフレームの数や設定位置により推定精度が変化する可能性があるため、全てのフレームをキーフレームとし、

表 1 観測データに付加する誤差

特徴点の検出座標	標準偏差 0.6 画素のガウスノイズ+量子化誤差
GPS の計測誤差	標準偏差 30mm のガウスノイズ

表 2 外部パラメータの初期値に付加する誤差

カメラ位置	標準偏差 500mm のガウスノイズ
カメラ光軸の向き	標準偏差 0.020rad のガウスノイズ

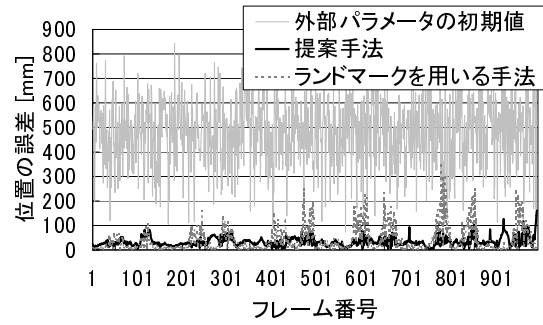


図 5 カメラ位置の推定誤差

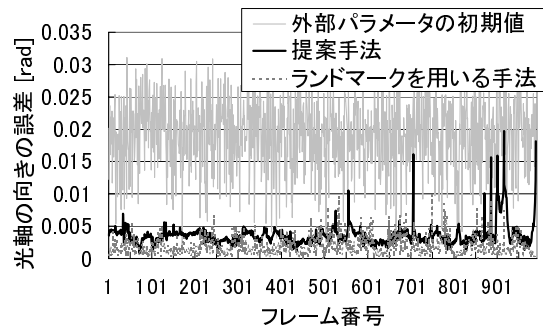


図 6 カメラ光軸方向の推定誤差

全てのフレームで同数のランドマークが写るように設定した。この実験では、各フレームで 15 個以上のランドマークが撮影されるように仮想環境中の特徴点からランダムに複数個を選択し、全ての画像上での位置を与えた。

以上の条件で、特徴点の三次元位置、外部パラメータ、カメラと GPS 受信機の位置関係を最適化した。実験の結果、外部パラメータは、カメラ位置に関して平均誤差 32.9mm、標準偏差 18.0mm、カメラ光軸の向きに関して平均誤差 0.0036rad、標準偏差 0.0016rad で推定された。

図 5 および図 6 に初期値として与えた真値からの誤差、提案手法における誤差、ランドマークを用いる手法における誤差を示す。ただし、図 5 は、外部パラメータの位置に関する誤差、図 6 については、光軸の向きに関する誤差を示す。実験の結果、ランドマークを用いる手法では、カメラ位置に関しては平均誤差 39.8mm、カメラ光軸の向きに関しては平均誤差 0.0019rad の精度で推定された。これに対して提案手法では、カメラ位置に関しては平均誤差 32.9mm、カメラ光軸の向きに関しては平均誤差 0.0036rad の精度で推定された。提案手法では、人手による計測が必要な事前知識を用いることなく、ランドマークを用いる手法と同程度の精度が得られていることが分かる。

4.2 実環境データを用いた外部パラメータの推定

本節では、提案手法における特徴点検出位置及び GPS 測位値の誤差モデルや特徴点追跡の精度を含めて、提案手法の有効性を確認するために、実環境で取得した動画と GPS 測位値から外部パラメータを推定する。また、推定された外部パラメータを用いて動画上に仮想物体を重畳表示するマッシュアップを行うことで、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。

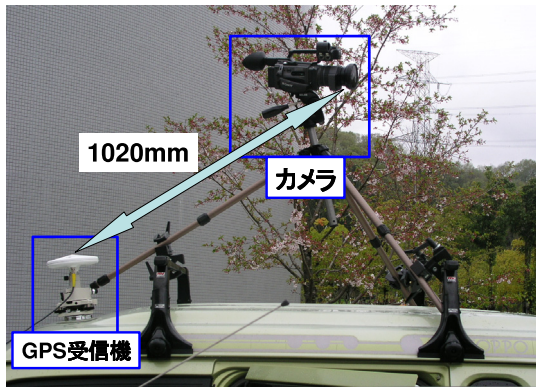


図7 ビデオカメラとGPS受信機の位置関係

実験では、ワイドレンズ (Sony VCL-HG0758) を取り付け、ビデオカメラ (Sony DSR-PD-150) と GPS 受信機 (Nikon LogPakII, 測位精度 $\pm 3.0\text{cm}$) を図7に示すように車に固定して取り付けた。このような車両を用いて、約1.1kmの距離を平均時速16.5kmの速度で走行し、得られた3600枚の画像 (解像度 720×480 画素, 14.985fps, プログレッシブ撮影) 及び、RTK法により1秒間隔で計測されたGPS測位値を入力として用いた。ただし、動画画像とGPS測位値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予めTsaiの手法[13]によりキャリブレーションし、その値を用いた。また、カメラとGPS受信機間の設置距離は、手動で計測した1020mm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測位値を用いた。逐次処理中に行う最適化処理 (B') においては、更新対象フレーム $k = 5$ 、重複フレーム $l = 22$ フレームと設定した。処理 (C') では、 $k' = 3600$ 、 $l' = 0$ とし、動画画像全体を一括で最適化した。

まず、逐次処理においてGPS測位値を用いた最適化を行うことにより、安定した特徴点追跡と、誤差の蓄積しない外部パラメータ推定が実現されることを確認するため、逐次処理においてGPS測位値を用いて最適化を行う場合と、動画画像のみから推定する場合を比較した。ただし、蓄積誤差の影響を確認するために、動画画像のみから推定する場合は、ランドマーク情報は与えず、GPS測位値を用いて最適化を行う場合と、動画画像のみから推定する場合の両方において、第1フレームと第15フレームの位置と姿勢を与えた。両フレームのカメラ位置については、GPS測位値を用い、姿勢については予め計測しておいた値を用いた。動画画像のみから外部パラメータを推定した結果と、GPS測位値を用いた最適化を行う逐次処理で推定された外部パラメータを図8に示す。実験の結果、動画画像のみから外部パラメータを推定した場合は、図8に示すように第1409フレームで追跡される特徴点が減少し、以降の外部パラメータは推定されなかった。これに対し、GPS測位値を用いる提案手法の逐次処理では、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。このことから、GPS測位値の利用により特徴点追跡自体の精度が向上していることが分かる。

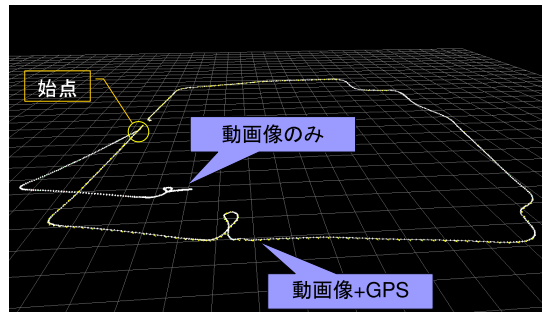
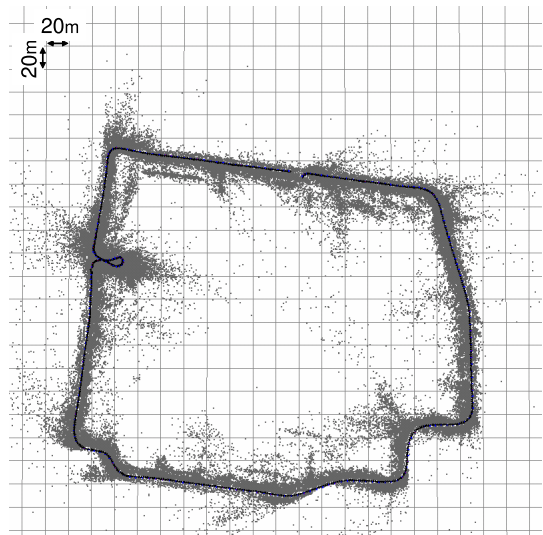
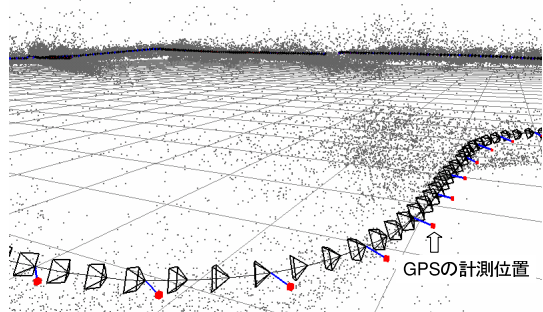


図8 逐次処理で得られた外部パラメータ



(a) 上から見た場合

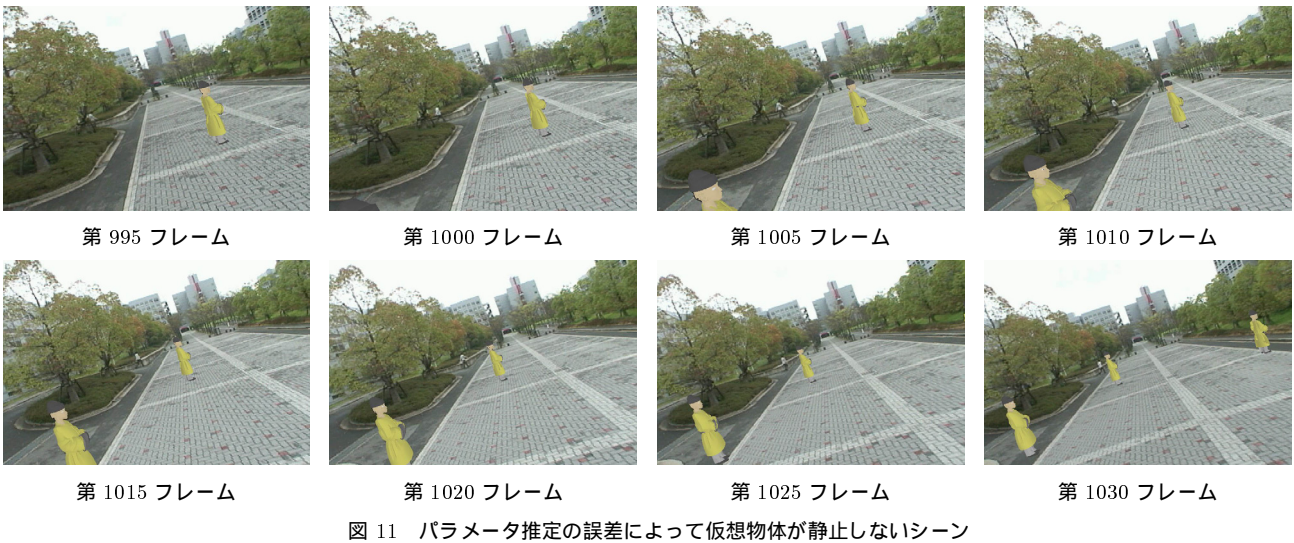
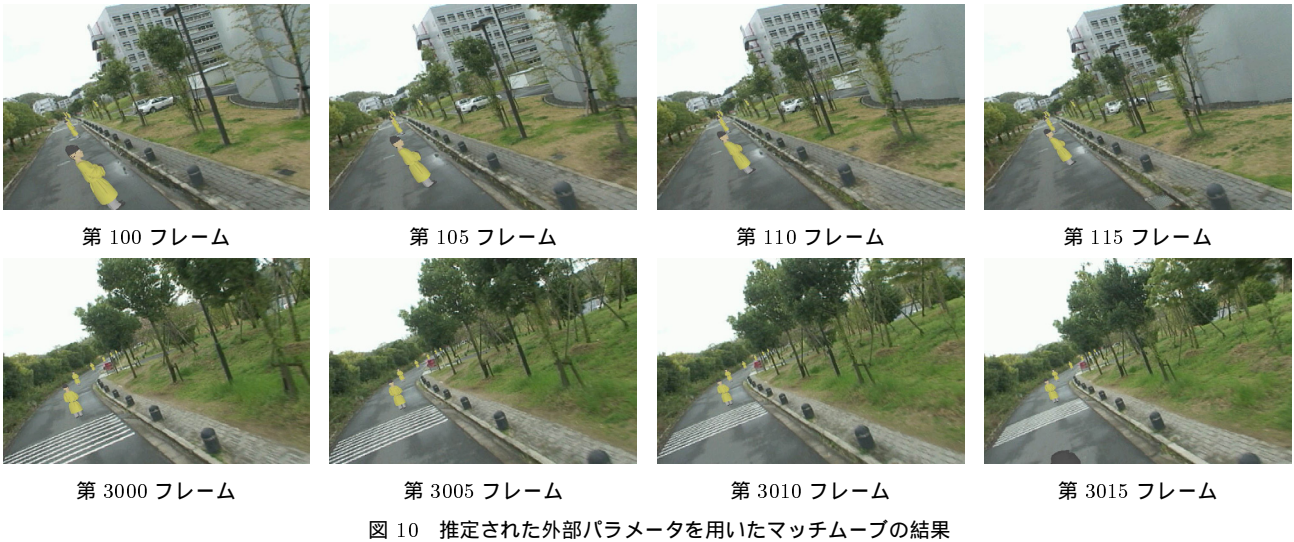


(b) 斜め上から見た場合

図9 カメラ外部パラメータの推定結果

次に、逐次処理で得られた外部パラメータを初期値として動画画像全体で最適化した。全体最適化によって推定された特徴点の三次元位置と外部パラメータを図9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。また、点群は外部パラメータ推定に用いられた特徴点を表す。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測位値が計測されていないフレームにおいても連続的に外部パラメータが推定されている。

推定された外部パラメータを用いて仮想物体を動画画像に合成したマッシュアップの結果を図10に示す。仮想物体は、GPSによる計測位置およびGPS受信機設置時の道路面からの高さから



計算される道路面上に配置した。図 10 に示すシーンでは、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、外部パラメータに大きな推定誤差が含まれないことを確認した (http://yokoya.naist.jp/pub/movie/yokochi/match_move.mpg)。しかし、図 11 に示すように、第 995 フレームから第 1030 フレームでは、仮想物体が一定の位置には静止せず、フレームが進むにつれ場所が移動した。これは、RTK-GPS の測位精度が環境中の遮蔽物により低下したことが考えられる。実際に、図 12 に示すように、RTK-GPS から出力される信頼度を表す標準偏差が第 995 フレームから第 1030 フレームにおいて増加している。この問題を解決するためには、GPS 測位自体もオフラインで処理し、推定された特徴点の三次元位置情報から周囲の環境の形状を推定し、有効な衛星の選択や、電波の反射状況の推定に用いる方法が考えられる。

4.3 全体最適化とシーケンス分割による最適化の比較

本節では、動画像全体を一括で最適化する結果と複数に分割して最適化する結果の比較を行い、前章 3.3 節の処理 (C') の有効性を確認する。実験には、前節の逐次処理で得られた外部パラメータを初期値とし、全体を一括で最適化した結果と、複

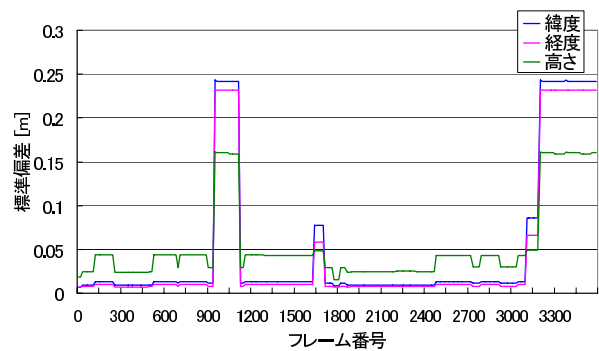


図 12 GPS 測位値に含まれる標準偏差

数のパスに分割して最適化した後に一つに統合した結果を比較する。

図 13 に示すように、重複フレーム l' を 200 フレーム、更新フレーム k' を 900 フレームとし、動画像を 4 つのシーケンスに分割し、各シーケンスを独立に最適化した後に各シーケンスの更新フレームを統合し、動画像全体の外部パラメータとした。動画像を一括で最適化して得られる外部パラメータを真値とし、分割最適化によって得られた外部パラメータの誤差を図 14 に

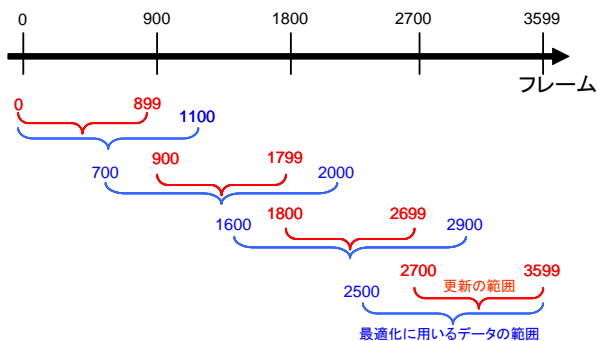


図 13 分割最適化における各シーケンスの分割範囲

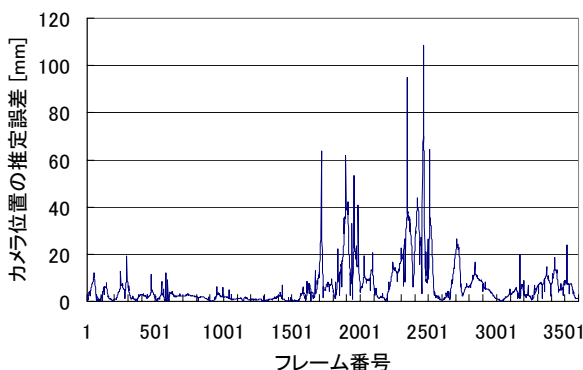


図 14 動画全体の一括最適化と分割最適化の比較

示す．分割最適化によって得られた外部パラメータの誤差は平均 6mm であった．4.1 節で示した実験において，カメラ位置の推定誤差が 32.9mm であり上述の誤差よりも大きいことから，分割最適化により一括での最適化に近い値が得られていることが分かる．このときの計算時間は，Pentium4 3GHz，メモリ 1GB の PC を用いた場合，一括最適化が約 7 時間であったのに対して，分割最適化では各シーケンスがそれぞれ約 1.5 時間であった．誤差が大きなフレームに関しては，3 番目のシーケンスにおいて推定をした外部パラメータを更新しない区間 l' フレームの大きさが不十分であった可能性がある．今回は，経験的にこの値を設定したが，今後はフレーム間の影響を定量化し，自動的に区間 l' の大きさを決定する必要がある．以上により，計算機に搭載されるメモリの量の制限など，一括で動画全体を最適化できないような場合においても，最適化の際に外部パラメータを更新しない区間 l' フレームのフレーム数を十分大きく設定することで，一括で最適化する場合に近い結果が得られることが分かった．

5. まとめ

動画と GPS 測位値からカメラの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再投影誤差と GPS 測位値によって定義される位置誤差を同時に最小化することで，外部パラメータを各フレームに対して推定することができる．

シミュレーション実験及び実環境データを用いた実験により，提案手法により推定された外部パラメータが三次元位置が既知

な点であるランドマークを用いる手法と同程度の精度であることを確認した．提案手法により，現実環境に関する事前知識の取得に要する人的コストの大幅な削減が可能である．しかし，GPS の電波をさえぎる遮蔽物に囲まれるような状況下では，外部パラメータの推定精度の低下が見られた．

今後は，上記のような環境中に遮蔽物が多数存在する場合に対応するために，推定される特徴点の三次元位置情報を利用して衛星を選択することや，異なるシーケンスにおいて同一シーンが撮影される場合に再出現する特徴点の対応づけを行うことにより，さらなる精度向上を図る．

文 献

- [1] Feiner, S., MacIntyre, B., Höllerer, T. and Webster, A.: 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] 遠藤隆明, 片山昭宏, 田村秀行, 廣瀬通孝: 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*, Vol. 7, No. 2, pp. 185-192 (2002).
- [3] Fitzgibbon, A. W. and Zisserman, A.: Automatic Camera Recovery for Closed or Open Image Sequences, *Pro. 5th European Conference on Computer Vision (ECCV)*, pp. 311 - 326 (1998).
- [4] Pollefeys, M., Koch, R., Vergauwen, M., Dekeyser, B. and Gool, L. V.: Three-dimensional Scene Reconstruction from Images, *Proc. SPIE*, Vol. 3958, pp. 215-226 (2000).
- [5] Davison, A. J.: Real-Time Simultaneous Localisation and Mapping with a Single Camera, *Proc. 9th IEEE Int. Conf. on Computer Vision*, Vol. 2, pp. 1403-1410 (2003).
- [6] 佐藤智和, 神原誠之, 横矢直和, 竹村治雄: 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)*, Vol. J86-D-II, No. 10, pp. 1431-1440 (2003).
- [7] Comport, A. I., Marchand, É. and Chaumette, F.: A Real-time Tracker for Markerless Augmented Reality, *Proc. 2nd ACM/IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR2003)*, pp. 36-45 (2003).
- [8] Vacchetti, L., Lepetit, V. and Fua, P.: 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).
- [9] Güven, S. and Feiner, S.: Authoring 3D Hypermedia for Wearable Augmented and Virtual Reality, *Proc. 7th IEEE Int. Symp. on Wearable Computers (ISWC2003)*, pp. 118-126 (2003).
- [10] Harris, C. and Stephens, M.: A Combined Corner and Edge Detector, *Proc. Alvey Vision Conf.*, pp. 147-151 (1988).
- [11] 栗田多喜夫, 坂上勝彦: 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. Vol. I, pp. 65-70 (2000).
- [12] 佐藤智和, 池田聖, 横矢直和: 複数動画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).
- [13] Tsai, R. Y.: An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 364-374 (1986).