

K-030

## 動画像とGPS測位データからのカメラ運動の推定 Camera Motion Estimation from Video Images and GPS Position Data

横地 裕次<sup>†</sup>  
Yuji YOKOCHI

池田 聖<sup>†</sup>  
Sei IKEDA

佐藤 智和<sup>†</sup>  
Tomokazu SATO

横矢 直和<sup>†</sup>  
Naokazu YOKOYA

### 1. はじめに

現実環境を撮影した動画像からのカメラ運動の推定は、複合現実感、三次元形状復元、ロボットナビゲーションなど様々な分野での応用が期待され、多くの研究が成されている。しかし、動画像のみからカメラ運動を推定する手法には、カメラの位置・姿勢の推定結果に蓄積誤差が生じるため、広範囲に移動する場合に推定精度が著しく低下する問題があった。また、推定されるカメラ運動のスケールは不定であり、複数の動画像から推定されるカメラ運動の座標系を精度良く統合することは困難であった。これらの問題に対して、環境中に複数のマーカを配置する手法 [1, 2] や、複数のセンサを組み合わせる手法 [3] などが提案されている。[1, 2] の手法は、三次元位置が既知のマーカを用いることで、蓄積誤差とスケールの不定性の問題を解決したが、現実環境中にあらかじめ複数のマーカを配置し、計測する必要があった。他方、RTK-GPS、磁力計、2軸角度センサーを組み合わせることで、動画像を用いずにカメラの位置姿勢を推定する手法 [3] 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 測地座標系で得ることができたため、スケールの決定及び、複数のカメラ運動の推定結果を統合することが容易である。しかし、RTK-GPS によって位置が計測されるのは約 1 秒間隔である。また、各センサとカメラの正確な位置・姿勢の関係が考慮されていないため、カメラと GPS 受信機の取り付け位置が離れている場合、精度の高いカメラ運動の推定は困難である。

本稿では、屋外環境下で取得した動画像と GPS の相対測位法による測位データからカメラ運動を推定する手法を提案する。提案手法では、動画像から推定されるカメラ運動に基づく特徴点の投影誤差及び、カメラの位置・姿勢から推定される GPS 受信機の位置の推定誤差を同時に最小化することで、特徴点の三次元位置、カメラと GPS の位置・姿勢の関係及び、カメラ運動を同時に推定する。本手法では、現実世界中にマーカを設置することなく、動画像のすべてのフレームにおいてスケールの不定性のないカメラの位置・姿勢を推定できる。

### 2. 動画像とGPSを用いたカメラ運動の推定

本研究では、カメラと GPS 受信機の位置・姿勢の関係を固定して撮影した動画像と GPS の測位データから、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、カメラの位置・姿勢を推定する。提案手法では、まず、動画像中の特徴点を追跡することで、逐次的にスケールの任意性を含む相対的なカメラ運動を推定する。次に、動画像全体で特徴点の投影誤差と、カメラの位置・姿勢から推定される GPS 受信機の位置推定誤差を同時に最小化する。ただし、本稿では、動画像と GPS 測位データの同期が取れているものとして、カメラ運動を推定する。

<sup>†</sup>奈良先端科学技術大学院大学  
Nara Institute of Science and Technology (NAIST)

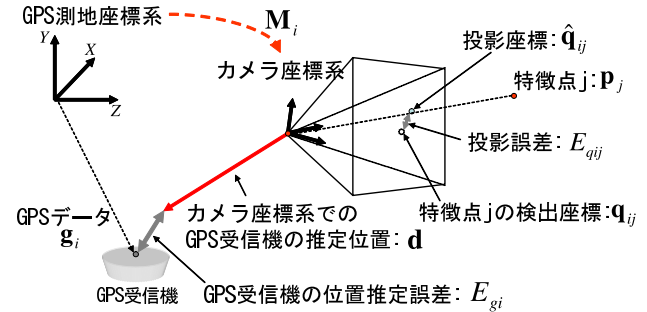


図 1: GPS 測地座標系とカメラ座標系及び誤差の定義

#### 2.1 GPS 測地座標系とカメラ座標系の定義

本研究で用いる GPS の測地座標系とカメラ座標系の関係を図 1 に示す。本稿では、GPS 測地座標系から第  $i$  フレームのカメラ座標系への変換行列  $M_i$  を、姿勢  $(\alpha, \beta, \gamma)$ 、位置  $(t_x, t_y, t_z)^T$  の 6 つのパラメータを用いて、以下のように定義する。

$$M_i = \begin{pmatrix} \mathbf{R}(\alpha, \beta, \gamma) & (t_x, t_y, t_z)^T \\ \mathbf{0} & 1 \end{pmatrix} \quad (1)$$

ただし、 $\mathbf{R}$  は  $3 \times 3$  の回転行列である。

#### 2.2 特徴点の投影誤差と GPS 受信機の位置推定誤差

図 1 に示す特徴点  $j$  の投影誤差  $E_{qij}$  及び、GPS 受信機の位置推定誤差  $E_{gi}$  について、以下で述べる。まず、特徴点の投影誤差  $E_{qij}$  を定義する。特徴点  $j$  の三次元位置  $\mathbf{p}_j = (X_j, Y_j, Z_j, 1)$  と第  $i$  フレームの画像上への投影座標  $\hat{\mathbf{q}}_{ij} = (\hat{x}_j, \hat{y}_j, 1)$  には、

$$\lambda_{ij} \hat{\mathbf{q}}_{ij} = \begin{pmatrix} \lambda_{ij} \hat{x}_j \\ \lambda_{ij} \hat{y}_j \\ \lambda_{ij} \\ 1 \end{pmatrix} = M_i \mathbf{p}_j \quad (2)$$

の関係が成り立つ。ただし、 $\lambda_{ij}$  は媒介変数である。一般に式 (2) によって求まる  $\hat{\mathbf{q}}_{ij}$  と画像上で検出される特徴点  $j$  の座標  $\mathbf{q}_{ij}$  は一致しない。この投影誤差を評価尺度とし、誤差関数  $E_{qij} = |\mathbf{q}_{ij} - \hat{\mathbf{q}}_{ij}|^2$  と定義する。

次に、カメラ座標系における GPS 受信機の位置推定誤差  $E_{gi}$  を定義する。カメラと GPS 受信機の位置及び姿勢の関係が固定されているものとする。第  $i$  フレームのカメラパラメータ  $M_i$  とそれに対応する GPS 測位データ  $\mathbf{g}_i = (X_{gi}, Y_{gi}, Z_{gi}, 1)$  には、

$$M_i \mathbf{g}_i = \mathbf{d} \quad (i \in F_g) \quad (3)$$

の関係が成り立つ。式 (3) は、カメラ座標系における GPS 受信機の位置が常に一定の位置  $\mathbf{d} = (x, y, z, 1)$  であることを表す。しかし、カメラパラメータ  $M_i$  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 データは一致しない。このカメラパラメータ  $M_i$  によって推定される GPS 受信機の位置推定誤差を誤差関数  $E_{gi} = |\mathbf{d} - M_i \mathbf{g}_i|^2$  と定義する。

### 2.3 特徴点の追跡によるカメラ運動の逐次推定

動画像からのカメラ運動の推定手法 [2] では、画像上の特徴点を追跡することで、その特徴点の三次元位置とカメラの位置姿勢を逐次的に推定する。第  $i$  フレームにおけるカメラの位置姿勢は、第  $i-1$  フレームまでの三次元位置が推定された 6 個以上の特徴点を用いて、誤差関数  $E_{qij}$  を最小化することによって推定される。また、新たに検出された特徴点の三次元位置は、第  $i$  フレームまでのカメラ運動と画像上での検出座標から、誤差関数  $E_{qij}$  を最小化することによって求められる。本研究においても同様に、逐次処理によってカメラ運動の初期値を算出する。ただし、従来手法では、初期フレームにおけるカメラの位置姿勢を、三次元位置が既知の基準マーカによって推定するが、提案手法では、GPS 測位データから初期値を推定する。

### 2.4 動画像全体での最適化

逐次的なカメラ運動の推定では、誤差が蓄積する問題がある。そこで、動画像全体での最適化を行うことで、累積的な推定誤差を最小化する。提案手法では、誤差関数  $E_{gi}$  と  $E_{qij}$  を用いて評価関数  $E$  を、以下のように定義する。

$$E = \frac{\omega}{|\mathbf{F}_g|} \sum_{i \in \mathbf{F}_g} E_{gi} + \frac{(1-\omega)}{\sum_i |\mathbf{S}_{p_i}|} \sum_i \sum_{j \in \mathbf{S}_{p_i}} E_{p_{ij}} \quad (4)$$

ただし、 $\omega$  は重み係数、 $\mathbf{S}_{p_i}$  は第  $i$  フレームに観測される特徴点の集合、 $\mathbf{F}_g$  は GPS 測位データが計測されたフレームの集合を表す。カメラ運動  $M_i$ 、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、評価関数  $E$  を勾配法によって最小化する。

## 3. カメラ運動推定の予備実験

提案手法によるカメラ運動の推定精度を評価するため、シミュレーション実験を行った。本実験では、カメラ運動の初期値が逐次推定処理によって得られているものとし、動画像全体での最適化を行う。ここでは、評価関数  $E$  が収束したときの、カメラ運動の推定精度を評価する。ただし、今回の実験では、カメラと GPS 受信機の位置関係  $d$  は真値が得られているものとした。

図 2 にシミュレーション実験で用いたカメラ運動及び、特徴点の三次元位置を示す。350 個の特徴点を原点を中心とする半径  $10m$  の範囲でランダムに配置し、カメラは、原点からの距離が長軸  $40m$ 、短軸  $20m$  の楕円上に、91 フレームで  $90^\circ$  回転移動した。特徴点  $j$  の第  $i$  フレームの画像上での検出座標  $q_{ij}$  は逐次処理によって得られているものとし、GPS 測位データは 15 フレーム毎に計測されているものとした。また、現実環境で得られるデータを想定して、特徴点の検出座標と GPS の測位データ及び、カメラ運動の初期値は、各々の真値に表 1 に示すガウシアンノイズを加えることにより生成した。以上の

表 1: 初期値に与えるガウシアンノイズ

	平均	標準偏差
特徴点の検出座標	0	0.6 [画素]
GPS 測位データ	0	20 [mm]
カメラ位置	0	1000 [mm]
カメラ光軸の向き	0	0.02 [rad]

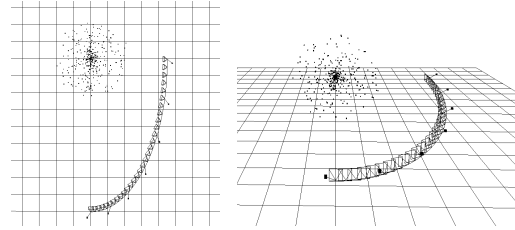


図 2: 推定するカメラ運動と特徴点

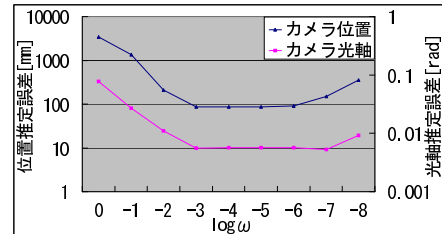


図 3:  $\omega$  とカメラ運動の推定精度の関係

条件で、カメラ運動  $M_i$  と特徴点の三次元位置  $p_j$  をパラメータとして、誤差  $E$  を最小化した。

式 (4) 中の  $\omega$  によるカメラ運動の推定精度の変化を図 3 に示す。実験の結果、 $\omega$  の値によって、カメラ運動の推定精度が大きく変化した。同図より、今回の予備実験では、 $\omega$  を  $10^{-3}$  から  $10^{-6}$  に設定したとき、他に比べて高い推定精度でカメラ運動が得られることを確認した。また、このときの標準偏差は、カメラの位置に関して  $55[mm]$ 、光軸に関して  $0.007[rad]$  程度であり、GPS 測位データが計測されていないフレームにおいても、カメラの位置・姿勢に大きな誤差を含むことは無かった。これにより、 $\omega$  を最適な値に設定すれば、真値に近いカメラの位置・姿勢が推定されることを確認した。

## 4. まとめと今後の課題

屋外環境下で取得した動画像と GPS の測位データからカメラ運動を推定する手法を提案した。提案手法は、動画像上で自然特徴点を追跡することにより、その特徴点の三次元位置とカメラ運動を逐次推定する。そして、推定された特徴点の投影誤差及び、GPS 受信機の位置推定誤差を動画像全体で最小化することにより、カメラ運動を推定する。

シミュレーション実験の結果、 $\omega$  を適切に設定すれば、GPS データの計測されていないフレームにおいてもカメラの位置・姿勢が推定されることを確認した。今後の課題は、今回の予備実験で真値を与えたカメラと GPS 受信機の位置関係を、自動的に求める手法の検討及び、カメラ運動の初期値を動画像と GPS 測位データから自動で算出することである。また、動画像と GPS 測位データの同期の問題を解決する必要がある。

## 参考文献

- [1] A.J.Davison: "Real-Time Simultaneous Localisation and Mapping with a Single Camera," Proc. 9th IEEE Int. Conf. on Computer Vision, Vol. 2, pp. 1403-1410, 2003.
- [2] T. Sato, M. Kanbara, N. Yokoya, H. Takemura: "Dense 3-D Reconstruction of an Outdoor Scene by Hundreds-baseline Stereo using a Hand-held Video Camera," Int. Jour. of Computer Vision, Vol. 47, No. 1-3, pp. 119-129, 2002.
- [3] S. Feiner, B. MacIntyre, T. Hollerer, T. Webster: "A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment," Proc. 1st. IEEE Int. Symp. on Wearable Computers, pp. 208-217, 1997.