

航空写真を事前知識として用いた 地上撮影画像からのカメラの位置・姿勢推定

関井 大 気^{†1} 桑 秀 行^{†1}
佐藤 智 和^{†1} 横 矢 直 和^{†1}

本稿では、既に比較的広域のデータが存在する航空写真を事前知識として、地上で撮影された1枚の静止画像からカメラの位置・姿勢を6自由度で推定する手法を提案する。航空写真を用いる従来手法の大半はカメラの姿勢や使用する環境に制約を与え、カメラ位置・姿勢を3自由度で推定しているため、画像による三次元計測や拡張現実感など広域の屋外環境において6自由度でのカメラ位置・姿勢推定が必要なアプリケーションへの応用は難しい。本研究では、航空写真における高度が未知の二次元の参照点と地上撮影画像上の対応点から、画像中の高度方向の直線と対応点の間の再投影誤差を最小化することで、地上カメラの相対高度、航空写真上の二次元の絶対位置および絶対姿勢を6自由度で導出するPnP問題の解法を提案する。実験では、シミュレーションおよび実環境の地上撮影画像と航空写真を用いた実験を行い、従来手法と提案手法の推定精度を比較することで、提案手法の有効性を示す。

Camera Position and Posture Estimation for a Single Ground-View Image Using Aerial Images

TAIKI SEKII,^{†1} HIDEYUKI KUME,^{†1} TOMOKAZU SATO^{†1}
and NAOKAZU YOKOYA^{†1}

This report describes a method to estimate camera position and posture in 6-DOF for a still image taken on the ground using aerial images that are already existing for many places of the world. For this purpose, most of existing methods estimate camera position and posture in 3-DOF with constraints on camera posture. However, these constraints are not suitable for such applications as augmented reality and three-dimensional measurement which need 6-DOF camera parameters in a wide outdoor scene. In this paper, we present a solution for PnP (perspective-n-point) problem from 2D-to-2D point correspondences between a ground-view image and an aerial image for estimating camera position and posture in 6 DOF (3-DOF absolute posture, 2-DOF horizontal absolute

position, and 1-DOF relative height) by minimizing reprojection-error defined between points and projected vertical lines. We demonstrate the performance of our method by comparing it with a conventional method in simulation and real environment.

1. はじめに

デジタルカメラや携帯電話に内蔵されたカメラで撮影された画像の位置・姿勢情報は、画像による三次元計測や拡張現実感によるヒューマンナビゲーション・景観シュミレーションなど、カメラ位置・姿勢に依存したアプリケーションにおいて必要不可欠である。これらの分野では、高精度な6自由度のカメラ位置・姿勢推定を実現する必要がある。

広域屋外環境下で撮影された静止画像に対するカメラの絶対位置・姿勢推定手法は、携帯端末に搭載されたGPSやジャイロなどのセンサ類を用いる手法^{1),2)}と画像を用いる手法³⁾⁻⁷⁾に大別できる。センサ類を用いる手法^{1),2)}は複数のセンサにより得られる計測値を組み合わせてカメラ位置・姿勢を推定するため、環境変化に対し比較的ロバストな推定が可能であり、また、使用する環境に関する事前知識や特別な準備を必要としないという利点がある。しかし、画像による三次元計測や拡張現実感などでの利用を前提とした場合には、一般にセンサから得られるカメラ位置・姿勢の推定精度が不十分であるという問題がある。これに対し、環境に関する事前知識を格納したデータベースと入力画像を照合することで、カメラの絶対位置・姿勢を高精度に推定する手法が提案されている。このような手法における事前知識として、これまでに画像データベース³⁾や人工的なマーカ⁴⁾、ランドマークデータベース⁵⁾、航空写真^{6),7)}など様々な情報を用いる手法が提案されている。この中で、環境をあらかじめ三次元復元し自然特徴点ランドマークデータベースとして用いる手法⁵⁾は、高精度にカメラ位置・姿勢を推定できるという特長を持つが、広範囲な屋外環境におけるデータベース構築の人的コストが大きという問題がある。一方、航空写真を用いる手法^{6),7)}は、比較的広域のデータが存在する航空写真を事前知識として用いることで、特別な準備なしに地上撮影画像からカメラの位置・姿勢を推定することができるという利点がある。しかし、航空写真を用いる従来手法の大半は、カメラの姿勢や使用する環境に制約を与えカメラ位

^{†1} 奈良先端科学技術大学院大学

Nara Institute of Science and Technology

置・姿勢を3自由度で推定しているため、6自由度でのカメラ位置・姿勢推定が必要なアプリケーションへの応用は難しい。

そこで本稿では、航空写真における高度が未知の二次元の参照点と地上撮影画像上の対応点から、地上カメラの相対高度、航空写真上の二次元の絶対位置および絶対姿勢の6自由度のカメラパラメータを導出するPnP問題の解法を提案する。具体的には、まず画像中の対応点を空間に投影し、航空写真上で与えられた対応点と空間に投影された光線間の空間距離の重み付き二乗和を誤差関数として定義し、これを準線形に解くことにより相対高度を除く5自由度のカメラ位置・姿勢を求める。次に、これを初期値として画像上の再投影誤差に対する非線形最小化を行い、相対高度を除く5自由度のパラメータを決定する。最後に、各特徴点に対する相対高度を計算することで6自由度のカメラ位置・姿勢を導出する。本稿では、実環境における推定の高精度化を目的として、既知の重力方向を用いたカメラ位置・姿勢の推定手法についても述べる。

以下、2節では本研究の関連研究について述べる。3節では提案手法に用いる誤差関数とカメラ位置・姿勢の推定方法について詳述する。4節で、既知の重力方向を用いたカメラ位置・姿勢の推定方法について述べる。5節ではシミュレーションによる実験と実環境の航空写真を用いた実験を行い、従来手法と比較することで提案手法の有効性を示す。最後に6節でまとめと今後の課題について述べる。

2. 関連研究

本節では、まず航空写真を事前知識として用いたカメラ位置・姿勢推定手法を概観し、各手法の特徴を述べる。次に、本研究に深く関連するPnP問題の解法に関する研究について考察する。

2.1 航空写真を事前知識として用いたカメラ位置・姿勢推定に関する研究

航空写真を事前知識として用いたカメラ位置・姿勢推定の従来手法は、3自由度のカメラ位置・姿勢を推定する手法と6自由度のカメラ位置・姿勢を推定する手法に分類できる。3自由度のカメラ位置・姿勢を推定する手法は、カメラのピッチ角、ロール角を既知としてカメラの位置2自由度とヨー角の3自由度のパラメータを推定する。野田ら⁶⁾は環境を平面であると仮定し、平面射影変換パラメータを算出した上で、地図上でのカメラの二次元位置と方位を算出している。この手法では、複数の地上撮影画像を用いることで、特徴点の少ない路面上での車両の位置・姿勢推定を実現している。このような3自由度のカメラ位置・姿勢推定手法を、6自由度のカメラ位置・姿勢推定が必要なアプリケーションに応用するこ

とは難しい。

これに対し、Bansalら⁷⁾は航空写真と共に航空写真よりも低空で撮影された視点の異なる鳥瞰写真を併用することで、6自由度のカメラ位置・姿勢推定を実現している。この手法は地上撮影画像を鳥瞰写真と対応付け、鳥瞰写真と航空写真の間の変換を平面射影変換により近似することで6自由度のカメラ位置・姿勢を得る。しかし、鳥瞰写真の存在しないシーンや地表面以外で撮影された画像には適用できない。

2.2 PnP問題に関する研究

本研究で扱う航空写真を事前知識として用いたカメラ位置・姿勢推定の問題は、航空写真から得た二次元の参照点と地上撮影画像上の対応点から6自由度でカメラ位置・姿勢を推定するという特殊な条件のPnP問題であると捉えられるが、一般的なPnP問題とは異なり、参照点の高度情報は利用できない。PnP問題は古くから研究されており、その基本的な解法として、P3P問題の解法^{8),9)}、P4P問題の解法^{9),10)}、6点以上の対応点によるPnP問題の解法^{8),11)–14)}がよく知られている。特にn点の対応を扱うPnP問題では、一般的に、まず空間中の誤差を最小化するカメラ位置・姿勢を線形に求め、次にこれらを初期値として参照点を画像上に投影した点と画像中の対応点間の距離(再投影誤差)の総和を誤差関数として最小化する手法が用いられている。しかし、従来のPnP問題の問題設定は、本研究が対象とする参照点の三次元座標に未知数が含まれる問題設定には直接適用できない。また、環境を平面であると仮定した上で、既知の平面上の参照点からカメラ位置・姿勢を求めるP4P問題の解法^{9),10)}を用いて、地表面上の対応点から6自由度のカメラ位置・姿勢を推定する手法が考えられるが、環境の傾斜や凹凸に対応できないという問題がある。

以上のように、環境に対する平面仮定やカメラの姿勢に対する制約なしに、航空写真のみを事前知識として用いて6自由度で直接カメラ位置・姿勢を推定する手法は、著者らの知る限り存在しない。

3. 航空写真を用いた地上撮影画像からのカメラ位置・姿勢推定

航空写真からは高度情報が得られないため、三次元の参照点を画像上に投影した点と画像上の対応点間の再投影誤差を誤差関数とする従来のPnP問題の解法を適用できない。そこで本研究では、図1のように三次元空間中において、航空写真上の参照点のX-Y座標を通る高度方向の直線を画像上に投影し、この直線と画像上の対応点間の距離の二乗和を再投影誤差として定義し、これを最小化することでカメラ位置・姿勢を算出する。ただし、この問題は非線形最小化問題となり、適切な初期値が必要となる。そこで本研究では、まず

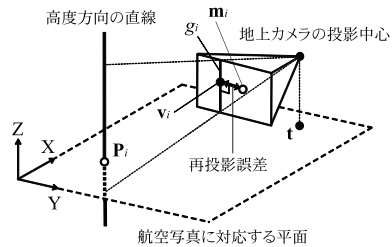


図 1 投影された直線と画像上の対応点の間で定義される再投影誤差

Fig. 1 A reprojection-error defined between a feature point and projected vertical line.

航空写真の平面上で定義された誤差関数を用いて、準線形にカメラ位置・姿勢の初期値を求め、これを用いて非線形最小化を行う。

具体的には、以下の手順により 6 自由度のカメラ位置・姿勢を推定する。

- (1) 地上撮影画像と航空写真の間で自然特徴点を対応付ける。
- (2) 航空写真上で定義された誤差関数を用いて、相対高度を除く 5 自由度のカメラ位置・姿勢の初期値を、姿勢パラメータが正規直交基底を満たすという条件の下で準線形に求める。
- (3) 画像上の直線と点の間の再投影誤差の非線形最小化を行い、5 自由度のカメラ位置・姿勢を求める。
- (4) 5 自由度のカメラ位置・姿勢から各特徴点に対する相対高度を計算する。

本稿では、航空写真上において世界座標が定義されており、高度方向が Z 軸、高度方向に垂直な平面が X-Y 軸に対応しているものとする。また、入力画像取得時のカメラ内部パラメータを既知とし、上記のステップ (1) における特徴点はあらかじめ (手動で) 対応付けられているものとする。画像間の対応付けの例を図 2 に示す。航空写真は直交投影で撮影されており、基準となる高度の地表面上に射影されていると仮定する。ただし、図 2 の航空写真中の建物のように、航空写真を撮影したカメラが鉛直下向きでない場合、基準となる高度の地表面上に存在しない点 (立体物上の点など) が航空写真上に傾いて撮像されてしまうため、このような点は地表面上の対応する位置の点を対応付ける。例えば、図 2 の地上撮影画像の建物上部の点は、航空写真上の赤で示された点ではなく地表面上の建物下部の点を対応付ける。また、カメラの高度は手法の原理上各特徴点との相対的な関係のみを、その他の 5 自由度のカメラ位置・姿勢から計算することができる。以下、本研究で用いる誤差関数の定義、上記のステップ (2), (3), (4) に対応するカメラ位置・姿勢の推定方法、について順

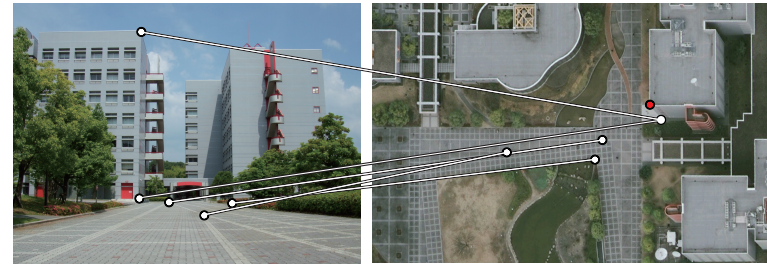


図 2 地上撮影画像と航空写真の間の点の対応付けの例

Fig. 2 An example of matching points between an aerial image and a ground-view image.

に述べる。

3.1 カメラ位置・姿勢推定に用いる誤差関数の定義

本節では、カメラ位置・姿勢推定処理において最終的に用いる再投影誤差と、初期値の推定に用いる誤差関数の定義について詳述する。

3.1.1 地上撮影画像上での再投影誤差の定義

図 1 に示したように、本研究では、地上撮影画像上の特徴点 m_i と、航空写真上の対応点 P_i を通る高度方向の直線を地上撮影画像上に投影した直線 g_i 上で m_i から最も近い点 v_i の間の距離の二乗和を、再投影誤差 E_v として以下のように定義する。

$$E_v = \sum_i^n |m_i - v_i|^2 \quad (1)$$

ただし、 n は全ての対応点の数である。 E_v の最小化は非線形最小化問題であるため、適切なカメラ位置・姿勢の初期値が必要となる。

3.1.2 初期値の推定に用いる航空写真上での誤差関数の定義

ここでは、再投影誤差 E_v の最小化に必要なカメラ位置・姿勢の初期値を得るために、 E_v を近似した誤差関数 E_a を、航空写真上に定義された X-Y 平面上で定義する。具体的には、図 3 に示すように、地上カメラの投影中心 t と撮像面上の特徴点 p_i を通る直線 (光線) s_i と、航空写真上の対応点 $P_i = (X_i, Y_i)^T$ を通る高度方向の直線との距離 d_i を用いて、誤差関数 E_a を次式のように定義する。

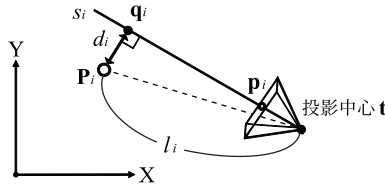


図3 航空写真上で定義される誤差関数
Fig.3 Proposed error function defined on an aerial image.

$$E_a = \sum_i^n a_i^2, \quad a_i^2 = \frac{d_i^2}{l_i^2} \quad (2)$$

ただし, l_i はカメラから点 P_i のへの航空写真上での距離を表す. また, d_i は次式のように表現できる.

$$d_i = |\mathbf{q}_i - \mathbf{P}_i| \quad (3)$$

点 \mathbf{q}_i は直線 s_i 上における P_i から最も近い点を $X - Y$ 平面に射影した二次元点である. ここで, カメラ座標系から世界座標系への回転行列を $\mathbf{R} = (\mathbf{r}_1^T, \mathbf{r}_2^T, \mathbf{r}_3^T)^T$, 航空写真上のカメラ位置を $\mathbf{t} = (t_1, t_2)^T$ とおけば, ベクトルの直交条件および平行条件より, 次式が成り立つ.

$$(\mathbf{q}_i - \mathbf{P}_i) \cdot (\mathbf{q}_i - \mathbf{t}) = 0 \quad (4)$$

$$(\mathbf{q}_i - \mathbf{t}) = \lambda_i (\mathbf{p}_i \cdot \mathbf{r}_1, \mathbf{p}_i \cdot \mathbf{r}_2)^T \quad (5)$$

ただし, \cdot はベクトルの内積を表す. ベクトル $(\mathbf{p}_i \cdot \mathbf{r}_1, \mathbf{p}_i \cdot \mathbf{r}_2)^T$ は航空写真上における投影中心から撮像面上の特徴点 \mathbf{p}_i へのベクトルを $X - Y$ 平面に射影したものであり, λ_i は媒介係数である. 以上により, 式 (2)(3)(4)(5) から a_i を求めると, 次式のように表される.

$$a_i = l_i^{-1} w_i \{ \mathbf{p}_i \cdot (t_2 \mathbf{r}_1 - t_1 \mathbf{r}_2) - Y_i (\mathbf{p}_i \cdot \mathbf{r}_1) + X_i (\mathbf{p}_i \cdot \mathbf{r}_2) \} \quad (6)$$

$$w_i = \{ (\mathbf{p}_i \cdot \mathbf{r}_1)^2 + (\mathbf{p}_i \cdot \mathbf{r}_2)^2 \}^{-\frac{1}{2}} \quad (7)$$

3.2 6 自由度のカメラ位置・姿勢の推定

本研究では, まず姿勢パラメータが正規直交基底を満たすという条件の下で, 式 (2) で定義した誤差関数 E_a を準線形に解き, 5 自由度のカメラ位置・姿勢の初期値を求める. 次に, それらを初期値として誤差関数 E_a, E_v の非線形最小化を行う. 以下, 具体的な手順について詳述する.

3.2.1 5 自由度のカメラ位置・姿勢の準線形解法

ここでは, 式 (6) について a_i を n 組並べた列ベクトルを考える. a_i を n 組並べた列ベクトルを次式に示す.

$$(a_1, a_2, \dots, a_n)^T = \mathbf{A} \mathbf{x} \quad (8)$$

$$\mathbf{A} = \begin{pmatrix} l_1^{-1} w_1 & \dots & 0 \\ \vdots & l_2^{-1} w_2 & \vdots \\ & & \ddots \\ 0 & \dots & l_n^{-1} w_n \end{pmatrix} \begin{pmatrix} \mathbf{p}_1^T & -Y_1 \mathbf{p}_1^T & X_1 \mathbf{p}_1^T \\ \mathbf{p}_2^T & -Y_2 \mathbf{p}_2^T & X_2 \mathbf{p}_2^T \\ \vdots & \vdots & \vdots \\ \mathbf{p}_n^T & -Y_n \mathbf{p}_n^T & X_n \mathbf{p}_n^T \end{pmatrix} \quad (9)$$

$$\mathbf{x} = (t_2 \mathbf{r}_1 - t_1 \mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_2)^T \quad (10)$$

式 (8) において, \mathbf{A} は重み行列と観測値を要素とする行列の積, \mathbf{x} は求めるパラメータを要素とする行列である. ここで, \mathbf{A} が全て既知のパラメータであれば, $E_a = \sum_i^n a_i^2$ を最小化する \mathbf{x} は, 擬似逆行列を用いて $\mathbf{A} \mathbf{x} = \mathbf{0}$ を解くことで算出できる. このとき, 原理的に 8 組以上の対応点が必要となる.

しかし, $\mathbf{A} \mathbf{x} = \mathbf{0}$ を解くにあたり複数の問題が存在する. まず, \mathbf{A} が未知数を含んでいる. また, \mathbf{x} には未知数の積が含まれ, カメラ位置である t_1, t_2 が一意には求まらない. そこで本手法では, 前者の問題に対し, \mathbf{A} に含まれる重み行列の初期値を単位行列とし, 擬似逆行列を用いて \mathbf{x} を求め, 重み行列を更新する. また, 後者の問題に対し, 先に $\mathbf{r}_1, \mathbf{r}_2$ を求め, それらを正規直交化し, $\mathbf{r}_1, \mathbf{r}_2$ を定数として再度 t_1, t_2 を求める. このような重み行列の更新と t_1, t_2 の算出を繰り返すことで, 準線形に解を得る. 最後に, $\mathbf{r}_1, \mathbf{r}_2$ から \mathbf{r}_3 を求めることで, 5 自由度のカメラ位置・姿勢を得る.

3.2.2 5 自由度のカメラ位置・姿勢の非線形最小化

準線形解法により求めた 5 自由度のカメラ位置・姿勢を初期値として, 誤差関数 E_a, E_v の順に非線形最小化を行う. 非線形最小化の手法には, 多自由度の非線形関数においてパラメータを分散させずに効率よく探索する方法として知られる Levenberg-Marquardt 法¹⁵⁾を用いる.

3.2.3 特徴点に対する相対高度の導出

カメラと各特徴点との相対高度は, 求めた 5 自由度のカメラ位置・姿勢から点毎に独立に計算することが可能である. 対応点 i の相対高度 \hat{t}_{3i} は, 世界座標系における投影中心から撮像面上の対応点 i へのベクトルの高さであることから, 次式で求める.

$$\hat{t}_{3i} = \lambda_i (\mathbf{p}_i \cdot \mathbf{r}_3) \quad (11)$$

ここで、 λ_i は求めた回転行列 \mathbf{R} 、カメラ位置 \mathbf{t} から、式 (2)、(3)、(4)、(5) により次式のように表される。

$$\lambda_i = w_i^2 |(\mathbf{p}_i \cdot \mathbf{r}_1)(X_i - t_1) + (\mathbf{p}_i \cdot \mathbf{r}_2)(Y_i - t_2)| \quad (12)$$

4. 既知の重力方向を用いた 5 自由度のカメラ位置・姿勢の推定

3 節の手法によりカメラ位置・姿勢を推定する際、画像上で対応点が同一直線上に存在する場合には、式 (9) で定義した行列 \mathbf{A} がランク落ち解が得られない。実用上問題となるケースとしては、地面上の点のみを対応付けた場合、画像上で一部の領域に対応点が集中することで、この条件に近づき推定精度が低下する。このとき、方位を除くカメラ姿勢の変化に対する再投影誤差の変化が小さくなり、カメラから得られる重力方向が正しく求まらない。このような場合において、重力方向を何らかの手法により得ることができれば、精度低下を回避できる。以下では、このような場合を想定した重力方向の推定手法、既知の重力方向を用いた 5 自由度のカメラ位置・姿勢の推定について詳述する。

4.1 重力方向の推定

重力方向の推定手法として、画像中の直線群の利用¹⁶⁾ や、携帯端末上の加速度センサの利用が考えられる。画像中の直線群を用いる手法¹⁶⁾ は、構造物の直線を検出し重力方向の直線群から得られる消失点からその方向を求めるものであり、多くの構造物が撮影された画像に対して有効である。また、携帯端末上の加速度センサを用いる方法では、一般に普及している加速度センサでも比較的高精度に推定でき、その推定精度は安価なものでは約 0.5 (度)、高価なものでは約 0.02 (度) である。後述する実験では、これらの計測誤差がカメラ位置・姿勢推定精度に与える影響について検証する。

4.2 既知の重力方向を用いた誤差関数の最小化による 5 自由度のカメラ位置・姿勢の推定

何らかの手法により既知の重力方向が得られる場合における、カメラ位置・姿勢の推定手法を述べる。本項では、まずカメラ姿勢の重力方向に対応する成分を定数とする場合において、誤差関数 E_a を算出するためのパラメータを再定義する。次に、これらを用いて誤差関数 E_a を最小化する 5 自由度のカメラ位置・姿勢を推定する手法について述べる。

カメラ座標系における重力方向の単位ベクトルは、世界座標系からカメラ座標系への回転行列 \mathbf{R}^T と世界座標系での重力方向のベクトル $\mathbf{g} = (0, 0, -1)^T$ の積 $\mathbf{R}^T \mathbf{g}$ で現され、これは回転行列 \mathbf{R} の行ベクトル $\mathbf{r}_3 = (r_{31}, r_{32}, r_{33})$ に対応することから、 $\mathbf{r}_3 = \mathbf{R}^T \mathbf{g}$ より \mathbf{r}_3 は定数となる。また、 \mathbf{r}_2 は次式を満たす。

$$\mathbf{r}_2 = \mathbf{r}_3 \times \mathbf{r}_1 \quad (13)$$

式 (13) を式 (6) の右辺に代入することで、誤差関数 E_a を算出するために必要なパラメータを以下のように再定義する。

$$a_i = l_i^{-1} w_i \{ \mathbf{p}_i \cdot (t_2 \mathbf{r}_1 - t_1 \mathbf{r}_2) + \mathbf{b}_i \cdot \mathbf{r}_1 \} \quad (14)$$

$$\mathbf{b}_i = (X_i(r_{32} + q_i r_{33}) - Y_i p_i, X_i(p_i r_{33} - r_{31}) - Y_i q_i, X_i(q_i r_{31} - p_i r_{32}) - Y_i) \quad (15)$$

$$\mathbf{p}_i = (p_i, q_i, 1)^T \quad (16)$$

$$\mathbf{A} = \begin{pmatrix} l_1^{-1} w_1 & \cdots & \mathbf{0} \\ \vdots & l_2^{-1} w_2 & \vdots \\ & & \ddots \\ \mathbf{0} & \cdots & l_n^{-1} w_n \end{pmatrix} \begin{pmatrix} \mathbf{p}_1^T & \mathbf{b}_1 \\ \mathbf{p}_2^T & \mathbf{b}_2 \\ \vdots & \vdots \\ \mathbf{p}_n^T & \mathbf{b}_n \end{pmatrix} \quad (17)$$

$$\mathbf{x} = (t_2 \mathbf{r}_1 - t_1 \mathbf{r}_2, \mathbf{r}_1)^T \quad (18)$$

これらを用いて、3.2.1 項と同様の方法で、 \mathbf{r}_3 を定数として更新した誤差関数 E_a を最小化するカメラ位置・姿勢の初期値を準線形に求める。このとき、原理的に 5 組以上の対応点が必要となる。次に、3.2.2 項の手順において、カメラ姿勢の重力方向に対応する成分を定数として、準線形に求めたカメラ位置・姿勢の初期値から誤差関数 E_a 、 E_v の非線形最小化を行い、5 自由度のカメラ位置・姿勢を求める。

5. 実験

本節では、シミュレーションによる実験、および実環境の地上撮影画像と航空写真を用いた実験により、提案手法のカメラ位置・姿勢の推定精度を評価する。本実験では評価尺度として、相対高度を除くカメラの位置 2 自由度については航空写真上での真値との距離、姿勢 3 自由度についてはカメラの光軸の真値との角度を用い、カメラの相対高度は評価しない。また、従来手法として、2.2 節で述べた平面上の参照点を用いた P4P 問題の解法^{9),10)} を用い、推定精度を比較する。以下では、P4P 問題の解法を用いる従来手法、重力方向が未知である提案手法、重力方向を既知とした提案手法をそれぞれ、P4P、P8P2D、P5P2D と呼ぶ。この中で、P8P2D、P5P2D はロバスト推定手法の代表的なアルゴリズムである RANSAC⁹⁾ により、カメラ位置・姿勢を推定する。

5.1 シミュレーションによる実験

航空写真を事前知識として用いる従来手法^{9),10)} は、環境の平面上の参照点を用い、参照点の平面は航空写真と平行であると仮定しており、環境の傾斜や凹凸は考慮されない。本実験

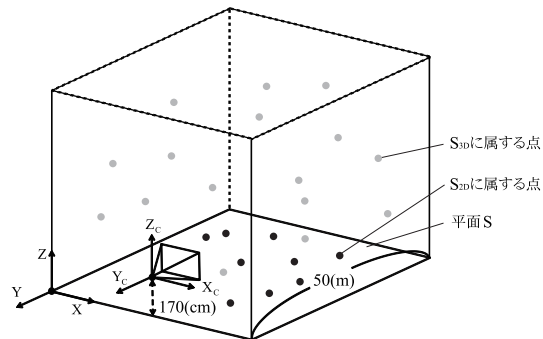


図4 シミュレーションの仮想環境
Fig. 4 Virtual environment for simulation.

では、図4に示すような平面上に分布する点群 S_{2D} 、三次元空間中に分布する点群 S_{3D} 、平面上と三次元空間中の点群 $S_{2D} \cup S_{3D}$ の、異なる三つのデータセットに対し、P4P、P8P2D、P5P2Dを用いたカメラ位置・姿勢推定を行い、それぞれの推定精度を評価する。また、 S_{2D} のデータセットを用い、平面の傾きを変化させた場合における、P4P、P5P2Dの推定精度を比較する。加えて、重力方向の推定に誤差が含まれることを想定したP5P2Dの精度評価も行う。

図4の仮想環境では、世界座標系 $X-Y-Z$ 空間において、 $Z=0$ の50(m)×50(m)の平面 S 上、 S を底面とする50(m)×50(m)×50(m)の立方体内にランダムに推定に十分な数の点を分布させ、それぞれ S_{2D} 、 S_{3D} とした。カメラは平面 S から170(cm)の位置(25.0, 0.0, 1.7)に設置し、カメラ座標系の $X_C - Y_C - Z_C$ 軸方向が世界座標系の $X - Y - Z$ 軸方向と一致するようにカメラの姿勢を設定した。対応点の撮像面上の検出座標、高度 Z 座標を除く航空写真に対応する二次元の $X - Y$ 座標にそれぞれ、1(pixel)、10(cm)の標準偏差の正規分布に従うノイズを加え、それらを入力として実験に用いた。P4Pに与える点群の高度は常に0とした。

5.1.1 異なるデータセットに対する実験

S_{2D} 、 S_{3D} 、 $S_{2D} \cup S_{3D}$ の各データセットに対して、カメラ位置・姿勢を推定した結果を表1に示す。ここでは、P5P2Dに使用する重力方向を真値とした。

実験結果より、P4Pは S_{2D} のように平面仮定が成立する場合にはカメラ位置・姿勢を高精度に推定できるが、高度が未知の点群を含む S_{3D} 、 $S_{2D} \cup S_{3D}$ では解が得られなかった。

表1 異なるデータセットに対する推定誤差の比較
Table 1 Comparison of estimation error for different datasets.

手法	カメラ位置の推定誤差 (mm)			カメラ光軸方向の推定誤差 (度)		
	S_{2D}	S_{3D}	$S_{2D} \cup S_{3D}$	S_{2D}	S_{3D}	$S_{2D} \cup S_{3D}$
P4P	11.2	-	-	0.031	-	-
P8P2D	101.1	65.5	52.2	2.41	0.073	0.057
P5P2D	11.6	11.3	10.9	0.041	0.039	0.037

提案手法は参照点の高度情報を用いないため、全データセットでカメラ位置・姿勢を推定可能であり、環境の凹凸に対応可能であることが分かる。ただし、P8P2Dは点が画像上で偏る S_{2D} を用いた場合において、推定精度が低い。これに対して、P5P2Dは既知の重力方向を用いることで、どのような場合も高精度なカメラ位置・姿勢推定を実現できることが分かる。

5.1.2 平面の傾きを考慮した実験

S_{2D} のデータセットを用いて、平面 S の X_C, Y_C 軸のカメラの投影中心周りの回転角をそれぞれ $\theta_{X_C}, \theta_{Y_C}$ (度)として、 $\theta_{X_C}, \theta_{Y_C}$ を変化させた入力データに対し、P4P、P5P2Dによりカメラ位置・姿勢を推定した結果を図5に示す。ここでは、P5P2Dに使用する重力方向を真値とした。実験結果より、P4Pでは $\theta_{X_C}, \theta_{Y_C}$ の値の変化に応じて推定精度が低下している。これは参照点の存在する平面が、航空写真上の $X - Y$ 平面と平行であるとした仮定からの逸脱が原因であり、P4Pは環境の傾斜に対応できないことが分かる。一方で、P5P2Dでは $\theta_{X_C}, \theta_{Y_C}$ の値の変化が推定精度に影響しないことが確認できる。

5.1.3 誤差を含む重力方向を用いた実験

S_{2D} のデータセットを用いて、P5P2Dに使用する重力方向に標準偏差 δ (度)の正規分布に従うノイズを与え、 δ を変化させながらそれぞれ100回カメラ位置・姿勢を推定した結果を図6に示す。同図より、加速度センサを用いる重力方向の推定誤差が0.02(度)~0.5(度)程度の範囲の中であれば、位置推定誤差が25(mm)程度となっており、これは屋外ARナビゲーションなどの屋外拡張現実感アプリケーションを実現するために十分な精度であると考えられる。

5.2 実環境の地上撮影画像と航空写真を用いた実験

本実験では、拡張現実感のためのベンチマークである TrakMark¹⁷⁾ [trakmark.net] の NAIST Campus Package 02 に含まれる動画画像から30枚の画像を選び、各画像に対しP4P、P8P2D、P5P2Dを適用し5自由度のカメラ位置・姿勢推定の実験を行う。NAIST

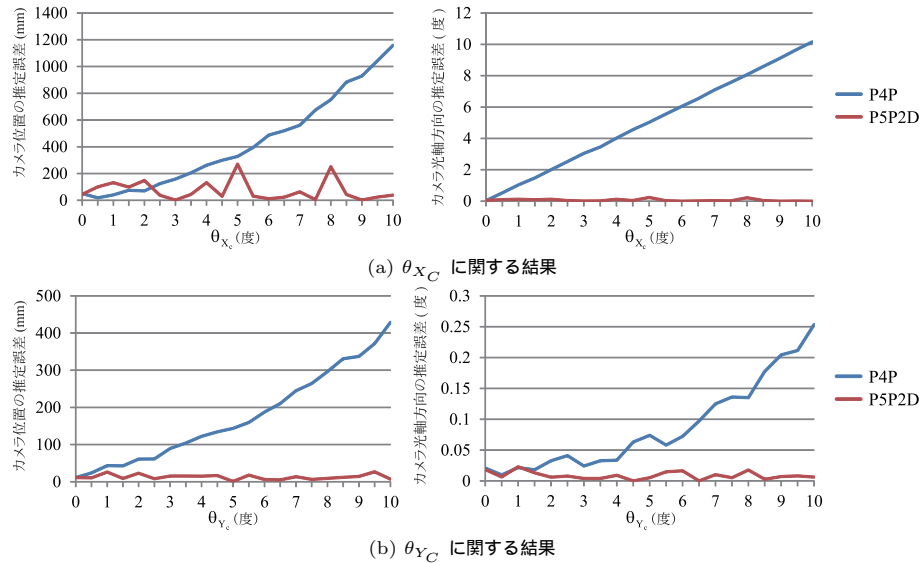


図 5 平面を傾けた場合の推定精度の比較

Fig. 5 Experimental results under the influence of plane gradient.

Campus Package 02 では、対応点の各画像上の検出座標と対応する三次元位置が与えられており、それらから従来の PnP 問題の解法により計算された各画像のカメラ位置・姿勢の参照値が与えられている。本実験では、このうちカメラ位置・姿勢の参照値を真値として用いる。航空写真は Google Map [maps.google.com] から取得したものを使用し、これを 30 枚の地上撮影画像と手動で対応付けた。本実験では、対応付けた点群を、実際に地表面上に存在する点群 R_{2D} 、それ以外の点群 R_{3D} 、に分類し評価に用いる。P5P2D では、カメラの姿勢の真値に正規分布に従う標準偏差 δ (度) の誤差を与え、 δ を変化させながらそれぞれ 100 回カメラ位置・姿勢を推定した。 R_{2D} 、 R_{3D} 、 $R_{2D} \cup R_{3D}$ の各データセットに対し、P4P、P8P2D、P5P2D を適用した実験結果を表 2 に示す。また、地上撮影画像と航空写真の間の点の対応付けの例と P5P2D によって得られた実験結果を図 7 に示す。同図中において、地上撮影画像および航空写真中の白点は与えた対応点、縦の白線は近傍の点に対応する投影された高度方向の直線である。また、航空写真中の赤で表示された三角形はカメラ位置・姿勢の真値であり、青で表示された三角形は R_{2D} のデータセットで P5P2D により推

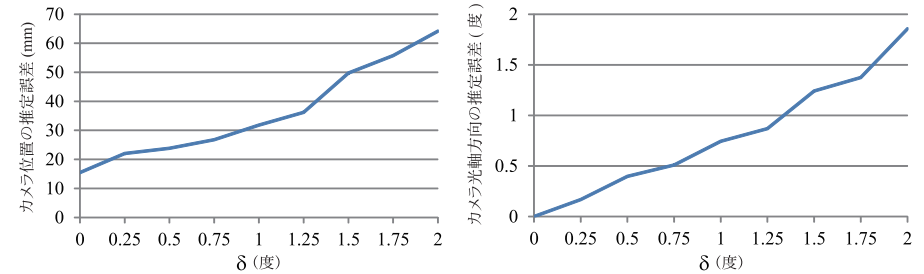


図 6 誤差を含む重力方向を用いた実験結果

Fig. 6 Experimental results under the influence of error of vertical direction.

表 2 実環境の航空写真を用いた実験結果
Table 2 Experimental results in real environment.

手法	カメラ位置の推定誤差 (mm)			カメラ光軸方向の推定誤差 (度)		
	R_{2D}	R_{3D}	$R_{2D} \cup R_{3D}$	R_{2D}	R_{3D}	$R_{2D} \cup R_{3D}$
P4P	252.7	-	-	1.981	-	-
P8P2D	1201.1	165.5	149.8	2.414	0.679	0.592
P5P2D($\delta = 0.0$)	240.2	142.9	138.4	0.374	0.325	0.321
P5P2D($\delta = 0.5$)	270.3	156.8	149.2	0.563	0.574	0.553
P5P2D($\delta = 1.0$)	289.3	181.1	173.6	0.694	0.715	0.827

定されたカメラ位置・姿勢である。

本実験では、P4P と高度が未知の点群を含む R_{3D} 、 $R_{2D} \cup R_{3D}$ を用いた場合では、解が得られなかった。また、 R_{2D} を用いた場合においても、地表面が僅かな傾斜・凹凸を持つことにより、環境を平面とした仮定が成立しないため、P4P は推定精度が悪いことが確認できる。これに対し、P8P2D は平面上の点のみを用いなければ(少なくとも立体物などの点を用いれば)、 R_{2D} を用いた P4P に比べ高精度に推定できることが分かる。また、P5P2D を用いた場合、重力方向に計測誤差が含まれる状況下でも、全データセットに対して高精度にカメラ位置・姿勢を推定できている。以上のことから、提案手法は対応点の高度が未知の環境下において、環境の傾斜・凹凸の変化に頑健なカメラ位置・姿勢推定を実現できることが分かる。

6. ま と め

本報告では、航空写真を事前知識として、航空写真における高度が未知の二次元の参照点

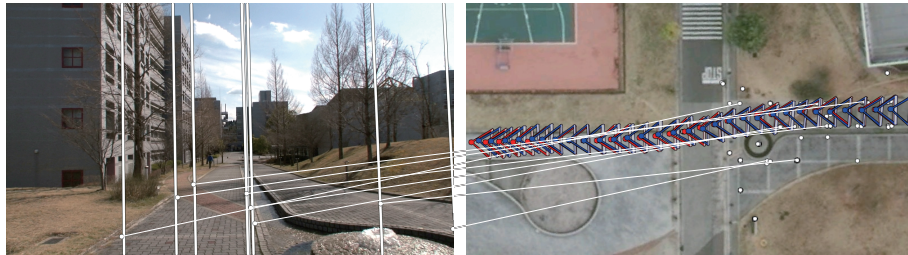


図 7 地上撮影画像と航空写真の間の点の対応付けの例と P5P2D によって得られた実験結果

Fig. 7 An example of matching points between a ground-view image and an aerial image and experimental results of P5P2D.

と画像上での対応点から、画像中の高度方向の直線と対応点の間の再投影誤差を誤差関数として定義し最小化することで、カメラの位置・姿勢を 6 自由度で推定する手法を提案した。また、既知の重力方向を用いて、求めるカメラの位置・姿勢を 3 自由度として誤差関数を最小化することにより、カメラの位置・姿勢を推定する手法を提案した。実験では、シミュレーションにより参照点の空間中での分布の異なるデータセットを用いて推定精度を評価し、平面上の点群を用いる従来手法に対する提案手法の優位性を示した。また、実環境の地上撮影画像と航空写真を用いた実験を行い、平面上の点群を用いる従来手法に比べ、提案手法が環境変化に頑健な推定が可能であることを示した。今後の課題として、航空写真と地上撮影画像間の点の対応付けの自動化、想定するアプリケーションを用いた実験による有用性の実証などが挙げられる。

謝辞 本研究の一部は科学研究費補助金(基盤(A), No.23240024, 若手(B), No.23700208)による。

参 考 文 献

- 1) Harter, A., Hopper, A., Steggle, P., Ward, A. and Webster, P.: The anatomy of a context-aware application, *Proc. ACM/IEEE Int. Conf. on Mobile Computing and Networking*, pp.59–68 (1999).
- 2) Addelee, M., Curwen, R., Hodges, S., Newman, J., Steggle, P., A.Ward and Hopper, A.: Implementing a sentient computing system, *IEEE Computer Magazine*, Vol.34, No.8, pp.50–56 (2001).
- 3) Cipolla, R., Robertson, D. and Tordoff: Image-based localization, *Proc. Int. Conf. on Virtual Systems and Multimedia*, pp.2–29 (2004).

- 4) Kato, H. and Billinghurst, H.: Marker tracking and hmd calibration for a video-based augmented reality conferencing system, *Proc. IEEE/ACM Int. Workshop on Augmented Reality*, pp.85–94 (1999).
- 5) 薄 充孝, 中川知香, 佐藤智和, 横矢直和: ランドマークデータベースに基づく静止画像からのカメラ位置・姿勢推定, *日本バーチャルリアリティ学会論文誌*, Vol.13, No.2, pp.161–170 (2008).
- 6) Noda, M., Takahashi, T., Ide, D. D.I., Murase, H., Kojima, Y. and Naito, T.: Vehicle ego-localization by matching in-vehicle camera images to an aerial image, *Proc. ACCV2010 Workshop on Computer Vision in Vehicle Technology: From Earth to Mars*, pp.1–10 (2010).
- 7) Bansal, M., Sawhney, H., Cheng, H. and Daniilidis, K.: Geo-localization of street views with aerial image databases, *Proc. ACM Multimedia 2011*, pp.1125–1128 (2011).
- 8) Slama, C.C.: *Manual of Photogrammetry*, Am. Soc. Photogrammetry and Remote Sensing, fourth edition (1980).
- 9) Fischler, M. and Bolles, R.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of ACM*, Vol.24, No.6, pp.381–395 (1981).
- 10) Abidi, M. and Chandra, T.: A new efficient and direct solution for pose estimation using quadrangular targets: Algorithm and evaluation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.17, No.5, pp.534–538 (1995).
- 11) Sutherland, I.: Three-dimensional data input by tablet, *Proc. IEEE*, Vol.62, pp.453–461 (1974).
- 12) Ganapathy, S.: Decomposition of Transformation Matrices for Robot Vision, *Proc. IEEE Conf. on Robotics and Automation*, pp.130–139 (1984).
- 13) Lepetit, V., Moreno-Noguer, F. and Fua, P.: EPnP: An accurate $O(n)$ solution to the PnP Problem, *Int. J. of Computer Vision*, Vol.81, No.2, pp.155–166 (2009).
- 14) Hesch, J.A. and Roumeliotis, S.I.: A direct least-squares (DLS) solution for PnP, *Proc. Int. Conf. on Computer Vision* (2011).
- 15) Bertsekas, D.P.: *Nonlinear programming*, Athena Scientific (1995).
- 16) Criminisi, A., Reid, I. and Zisserman, A.: Single view metrology, *Int. J. of Computer Vision*, Vol.40, No.2, pp.123–148 (2000).
- 17) Tamura, H. and Kato, H.: Proposal of international voluntary activities on establishing benchmark test schemes for AR/MR geometric registration and tracking methods, *Proc. Int. Symp. on Mixed and Augmented Reality*, pp.233–236 (2009).