

1 はじめに

異なる視点から得られた画像間でいくつかの点が対応づけられている場合、その三次元情報とカメラ運動を同時に復元する問題は、コンピュータビジョンの最も基本的な課題の一つであり、ステレオにおけるセルフキャリブレーションなど多くの応用が考えられる。カメラを透視投影に従うものとしてモデル化すると、この問題は一般に非線形最適化に帰着するため、その最適解の探索においては真値に十分近い初期値を与えることが重要である。従来のほとんどの手法は、8点アルゴリズム [1] を用いてカメラ運動パラメータの初期値を計算し、これを幾何学的制約を満たすように改善するという戦略をとっている。しかし、8点アルゴリズムは観測された特徴点の位置誤差に対し極めて敏感で、しばしば真値とかけ離れた値を与えるため、最適化の初期値としてすら使えないことがある。原点移動とスケールリングによって入力データを正規化して数値的不安定性を軽減する試み [2] もあるが、十分でない。

これに対し、本稿ではまずカメラをアフィン投影に従うものとして運動パラメータの初期値を求め、次にこれを透視投影モデルに基づいて改善する手法を提案する。これによって、8点アルゴリズムを用いて初期値を求める従来の手法に比べて、画像中の特徴点の位置誤差に対するロバスト性が向上することを示す。

2 エピポーラ拘束と基本行列

カメラの内部パラメータを既知とする。すると、シーン中の同一の点を2つの異なる視点から見たときの視線方向を表すベクトル $\mathbf{x} \equiv [x, y, 1]^T$ および $\mathbf{x}' \equiv [x', y', 1]^T$ が計算できて、それらの間には

$$\mathbf{x}'^T \mathbf{E} \mathbf{x} = 0 \quad (1)$$

というエピポーラ拘束が成り立つ。ここで、 $\mathbf{E} \in \mathcal{R}^{3 \times 3}$ は基本行列と呼ばれ、2つの視点間の運動パラメータ(並進と回転)から成る。並進の絶対的な長さがわからないので、 \mathbf{E} には定数倍の不定性が残る。さらに、 \mathbf{E} の特異値は0および非零の2重根となることが知られており、これは次の条件と等価である [3]。

$$\det \mathbf{E} = 0, \quad \|\mathbf{E}\|^2 = 2 \|\tilde{\mathbf{E}}\| \quad (2)$$

ただし、 $\tilde{\mathbf{E}}$ は \mathbf{E} の余因子行列である。

十分多くの対応 $\{\mathbf{x}_i, \mathbf{x}'_i\} (i = 1, 2, \dots, N)$ が与えられれば、制約条件 (2) のもとで

$$J \equiv \sum_{i=1}^N \frac{(\mathbf{x}'_i{}^T \mathbf{E} \mathbf{x}_i)^2}{(\mathbf{E} \mathbf{x}_i)_x^2 + (\mathbf{E} \mathbf{x}_i)_y^2 + (\mathbf{E}^T \mathbf{x}'_i)_x^2 + (\mathbf{E}^T \mathbf{x}'_i)_y^2} \quad (3)$$

3D Reconstruction from Multiple Images

- from Affine to Perspective Projection -

Toshio UESHIBA and Fumiaki TOMITA

Electrotechnical Laboratory

1-1-4 Umezono, Tsukuba, Ibaraki 305, Japan

なる評価関数を最小にすることによって、等方的 Gaussian ノイズの仮定の下での \mathbf{E} の最尤推定解が求められる。

8点アルゴリズム [1] とは、(3) 式の分母をすべて1とし、さらに制約条件 (2) を無視して \mathbf{E} を線形計算のみで求める手法である。このアルゴリズムは、実際には \mathbf{E} の自由度は5であるにもかかわらずあたかも8自由度であるかのように扱っており、この冗長性がノイズに対する不安定性の大きな原因となっている。

3 アルゴリズム

3.1 アフィン投影モデルによる初期値の推定

アフィン投影とは、三次元シーンから二次元画像面への線形写像として表される投影であり、orthographic, weak perspective, paraperspective 等の総称である。異なる2視点からのアフィン投影によって得た画像の対応点の間にも (1) 式のエピポーラ拘束が成立するが、次式で表されるように基本行列の左上 2×2 部分が0であるという著しい特徴を有する [4]。

$$\mathbf{E}_A \equiv \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & d \\ a & b & e \end{bmatrix} \quad (4)$$

\mathbf{E} と同様、 \mathbf{E}_A にも定数倍の不定性があるので、その自由度は4である。この最適解は、次式を最小にする a, b, c, d, e として線形計算で求められる [4]。

$$J_A \equiv \sum_{i=1}^N \frac{(ax_i + by_i + cx'_i + dy'_i + e)^2}{a^2 + b^2 + c^2 + d^2} \quad (5)$$

この計算は 4×4 対称行列の固有値問題に帰着し、冗長な自由度がないので極めて安定である。

3.2 透視投影モデルに基づく運動パラメータの改善

§3.1で求められたアフィン基本行列を初期値として、(2) 式の制約を満たすように透視投影モデルに基づいてこれを逐次的に改善することを考える。

\mathbf{E} の3つの行ベクトルを並べて9次元ベクトルとしたものを \mathbf{e} とすれば、(3) 式は

$$J = (\mathbf{e}^T \mathbf{A} \mathbf{e}) / (\mathbf{e}^T \mathbf{B} \mathbf{e}) \quad (6)$$

という形式で表される。ここで、 \mathbf{A}, \mathbf{B} は $\mathbf{x}_i, \mathbf{x}'_i$ から成る 9×9 行列である。一方、 $\tilde{\mathbf{E}}$ の3つの列ベクトルを並べて9次元ベクトルとしたものを $\tilde{\mathbf{e}}$ とすると、(2) 式は

$$\mathbf{e}^T \tilde{\mathbf{e}} = 0, \quad \|\mathbf{e}\|^2 = 2 \|\tilde{\mathbf{e}}\| \quad (7)$$

となる。さらに、定数倍の不定性を除去するために $\mathbf{e}^T \mathbf{e} = 1$ という制約を課せば、Lagrange 乗数 $\lambda_1, \lambda_2, \lambda_3$ を導入することによって、基本行列の推定問題は

$$J' \equiv J + \lambda_1 (1 - \mathbf{e}^T \mathbf{e}) + \lambda_2 (\mathbf{e}^T \tilde{\mathbf{e}}) + \lambda_3 ((\mathbf{e}^T \mathbf{e})^2 - 4\tilde{\mathbf{e}}^T \tilde{\mathbf{e}}) \quad (8)$$

を停留させることに帰着される。 $\partial J' / \partial \mathbf{e} = 0$ となるような \mathbf{e} を Newton 法で逐次的に求めれば基本行列が決定できる。

- [4] L. S. Shapiro, A. Zissermann, and M. Brady. 3D Motion Recovery via Affine Epipolar Geometry. *International Journal of Computer Vision*, Vol. 16, No. 2, pp. 147-182, 1995.

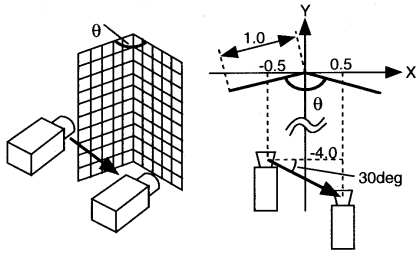


図 1: シミュレーションの状況設定

4 実験結果

シミュレーションによって本手法のロバスト性を評価した。 θ (deg)で交差する2平面が成す屏風状物体を対象とし、その上に描かれた121個の格子点を斜め後方に並進運動するカメラによって2箇所から観測した(図1)。カメラの焦点距離は1000(pixel)とし、観測された特徴点の座標に平均値0、分散 σ^2 (pixel²)のGaussianノイズを加えた。同一の σ について100回の試行を行い、カメラの並進ベクトルの推定値と真値との角度差が15deg以上になるものをoutlierとして、その割合を評価した。正規化処理を併用した8点アルゴリズム[2]を用いて初期値を求める手法と、アフィン投影モデルから初期値を求める本手法の比較を図2に示す。本手法の方が、 σ の増加に対してより安定であることがわかる。また、 θ が180degに近づくとき対象が平面的になって解の一意性がくずれるため推定が困難になるが、このような場合でも本手法の方がノイズに対して安定であることもわかる。さらに、本手法による並進ベクトルの推定誤差を図3に示す。

次に、実画像による結果を示す。横方向に移動しながらビルを2箇所から撮影し、21個の特徴点を抽出して手動で対応付けを行った後に本手法を適用した。図4に求められたエピポーラ線を示す。ビルに対してほぼ真横に移動しているため、エピポーラ線はビル正面の水平方向線に一致するはずであるが、実際そのように正しく推定されている。また、原画像を我々が開発中のステレオビジョンシステムに入力し、本手法で求めた運動パラメータを用いて3次元復元を行った結果を図5に示す。

5 おわりに

アフィン投影を初期カメラモデルとすることにより、2枚の画像の対応点からカメラ運動とシーンの3次元情報を安定に復元できることを示した。本稿ではカメラの内部パラメータを既知としたが、同様の考え方を内部パラメータ未知の2台のカメラについてfundamental matrixを推定する問題にも適用できる。

参考文献

- [1] H. C. Longuet-Higgins. A Computer Algorithm for Reconstructing a Scene from Two Projections. *Nature*, Vol. 293, No. 10, pp. 133-135, Sep 1981.
- [2] R. I. Hartley. In Defense of the 8-point Algorithm. In *Proc. 5th International Conference on Computer Vision*, pp. 1064-1070, 1995.
- [3] T. S. Huang and O. D. Faugeras. Some Properties of the E Matrix in Two-View Motion Estimation. *IEEE Trans. on*

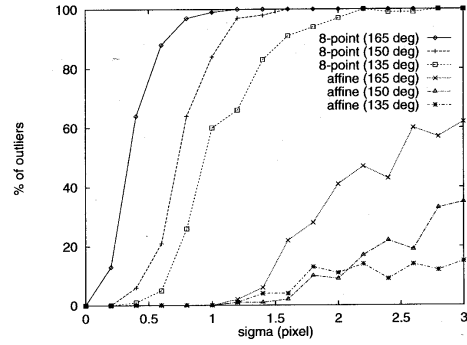


図 2: outlier の比率

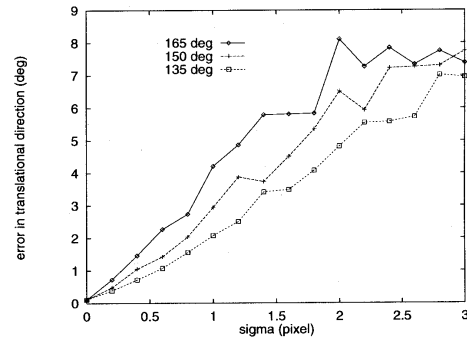


図 3: 並進ベクトルの推定誤差

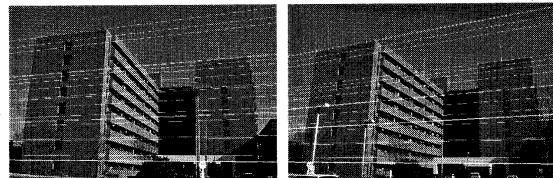


図 4: 推定されたエピポーラ線

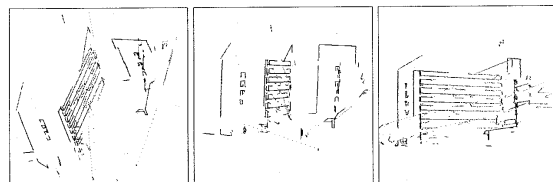


図 5: 3次元情報の復元