単眼移動視による物体形状推定に関する検討*

野村 由司彦^{*1}, 長屋 匡 紀^{*2}
藤井 省 三^{*1}, 松村 司 郎^{*2}

Object Shape Estimation by Ego-Motion Stereo

Yoshihiko NOMURA, Masamoto NAGAYA,
Seizo FUJII and Shiro MATSUMURA

Object shapes and camera motions are simultaneously estimated using the Marquardt method from sequential images obtained through a moving camera. This method has many advantages: (1) Convergence of viewing lines leads to high estimation accuracy. (2) Viewing hardware is simple and applicable to large-scale objects. (3) The optimization algorithm is also simple and accurate. Furthermore, estimation errors are examined, and the optimum camera motion is presented for 2 views and 5 object points in such a way that the average error is minimized.

Key Words: Seeing and Hearing Mechanism, Shape Measurement, Image Processing, Recognition, Camera Motion Estimation, Ego-Motion Stereo, Marquardt Method, Convergence Characteristics, Optimization of Viewing Point

1. まえがき

はじめに、能動的に動くカメラからの時系列画像を用いた視覚技術をさまざまな観点から分類し、その特徴を概説する。(1) 視軸の輻輳¹⁻⁷⁾は、3次元計測精度を飛躍的に高める。輻輳させなくても、視点間距離の高精度な連続的な変化は、計測精度を向上させ⁸⁾、画像の対応付けを容易にする⁹⁾。(2) 単眼移動視³⁻¹⁰⁾は、複眼移動視^{1,2)}に比べて、異なる画像間での物体の対応づけをしなければならない反面、複数台のカメラ間の位置関係を較正する必要がない。(3) カメラモーションと物体の3次元形状をともに未知とする方法²⁻⁸⁾は、カメラモーションを既知とし物体の3次元形状のみを未知とする方法^{1,9,10)}に比べて、大規模な測定対象に適している。(4) カメラモーションが未知の場合、観測方程式は非線形となるが、その解法は、(a) 非線形最適化、(b) カルマンフィルタ、(c) 解析的解法がある。解法(a)は、単眼視^{6,7)}と両眼視¹⁾に適用され、初期値の与え方に注意を要するが、最適解に収束すれば計測精度は高い。解法(b)は、物体形状推定とカメラモーション推定が分離され、逐次的に行われるので手続きが複雑となり、誤差を生ずる。また線形化のために微小と仮定して求めたモーションを累積させてカメラ位置を求めるので誤差も大きくなる⁵⁾。解法(c)は、モーション前後における各特徴点への視線ベクトルとカメラ並進ベクトルの共平面拘束などを利用する⁴⁾。多段の最適

化、多段の固有値計算の過程で誤差を生じ、前後1組のモーションにしか適用できない。

本研究では、視線が輻輳するように移動する単眼の画像から、対象の3次元位置とカメラモーションを非線形最適化法を用いて同時に推定する方法を取り上げる。したがって、(1) 3次元計測精度が高い、(2) 較正の手間が少なく、大規模な測定に向く、(3) 収束性の問題があるが、線形化誤差を含まず、アルゴリズムが簡明である。非線形最適化法としては、通常は Gauss-Newton 法⁶⁾が用いられているが、収束性の問題から共役勾配法を用いる方法も提案されている⁷⁾。これに対して、本研究ではアルゴリズムが簡明で、収束性も高い Marquardt 法¹¹⁾を適用する。そして、Marquardt 法が初期値の誤差に対して頑健に収束し、実用的であることを示す。さらに、最も基本的な2視点5特徴点の条件において計測誤差を定式化し、誤差最小規範により最適なカメラモーション、つまり視点の設定方法を明らかにする。

2. カメラ画像による3次元計測方法

2.1 3次元計測のモデル化

(1) 記号の説明とパラメータの分類 Fig.1は、3次元空間中にある物体上の特徴点がレンズを通して撮像面に結像する様子を示している。ここでは像の歪曲は考えない。物体上の特徴点 i が時刻 j に撮影された時の撮像面上の2次元座標を $u_i(j) \equiv [u_i(j) \ v_i(j)]^T$ と表す。本論文で扱う2次元画像座標は、テレビカメラ

* 原稿受付 平成6年9月9日。

^{*1} 正員、名古屋大学工学部 (〒464-01 名古屋市千種区不老町)。^{*2} 中部電力(株)電力技術研究所 (〒459 名古屋市緑区大高町北関山 20-1)。

の撮像素子に対応するメモリ上の座標で表し、単位は画素とする。レンズ中心 C を原点とし、撮像面の水平、垂直方向に x_c 軸、 y_c 軸、レンズ中心から視軸の方向に z_c 軸となるように定めた右手座標系をカメラ座標系とする。観測対象の3次元物体に対して適当に定めた右手座標系をワールド座標系とし、原点を W 、各座標軸を x_w, y_w, z_w と表す。特徴点 i の座標を、ワールド座標系上で ${}^w\mathbf{r}_i \equiv [{}^w x_i \ {}^w y_i \ {}^w z_i]^T$ 、カメラ座標系上で ${}^c\mathbf{r}_i \equiv [{}^c x_i \ {}^c y_i \ {}^c z_i]^T$ と表す。また、カメラ座標系の各座標軸、 x_c, y_c, z_c に関するワールド座標系の回転角を ϕ, θ, ψ (右回りを正) とする。カメラに固有の定数は、画面距離 f 、撮像素子の幅と高さ d_u と d_v 、主点 c_u, c_v である。

Fig.1 Camera model

(2) 定式化 ${}^w\mathbf{r}_i$ から ${}^c\mathbf{r}_i$ への変換は、

$$[{}^c\mathbf{r}_i] = [{}^c\mathbf{R}_w \ {}^c\mathbf{t}_w] \begin{bmatrix} {}^w\mathbf{r}_i \\ 1 \end{bmatrix} \quad (1)$$

による。ここで、 ${}^c\mathbf{R}_w$ は、3次元ワールド座標系を3次元カメラ座標系のまわりに回転させる時の回転行列であり、本論文では y_c, x_c, z_c 軸の順に回転させる。 ${}^c\mathbf{t}_w \equiv [{}^c t_x \ {}^c t_y \ {}^c t_z]^T$ はカメラ座標系から見た時のワールド座標系の原点の位置ベクトルを表す。

次に ${}^c\mathbf{r}_i$ から $u_i(j)$ への変換は、

$$u_i(j) = \frac{1}{c_{zi}} C \ {}^c\mathbf{r}_i \quad (2)$$

による。ここで、 C は透視投影を表す定数行列である。

$$C \equiv \begin{bmatrix} \frac{f}{d_u} & 0 & c_u \\ 0 & \frac{f}{d_v} & c_v \end{bmatrix} \quad (3)$$

上記の2つをまとめ、 ${}^w\mathbf{r}_i$ から $u_i(j)$ への変換は、

$$u_i(j) = g({}^w\mathbf{r}_i, \alpha(j)) \quad (4)$$

となる。ただし、

$$g({}^w\mathbf{r}_i, \alpha(j)) \equiv \frac{1}{e_3^T {}^w\mathbf{r}_i + c_{t_z}} C [{}^c\mathbf{R}_w \ {}^c\mathbf{t}_w] \begin{bmatrix} {}^w\mathbf{r}_i \\ 1 \end{bmatrix} \quad (5)$$

$$\alpha(j) \equiv [\phi(j) \ \theta(j) \ \psi(j) \ {}^c t_x(j) \ {}^c t_y(j) \ {}^c t_z(j)]^T \quad (6)$$

$$[e_1, e_2, e_3]^T \equiv {}^c\mathbf{R}_w \quad (7)$$

2.2 未知パラメータの推定法

(1) 非線形観測方程式の導出 ${}^w\mathbf{r}_i$ から $u_i(j)$ への変換式 (4) をすべての時刻におけるカメラモーションおよび特徴点についてまとめるとともに、 $u_i(j)$ が測定誤差を含んでいることを考えると、非線形観測方程式として、

$$\mathbf{y} \approx \mathbf{f}(\mathbf{x}) \quad (8)$$

を得る^{6,7)}。ここで、 $\mathbf{y}$ は観測値ベクトル、 $\mathbf{x}$ は未知パラメータベクトルである。

$$\mathbf{y} \equiv \begin{bmatrix} u_1(1) \\ u_1(2) \\ \vdots \\ u_1(n_m) \\ u_2(1) \\ u_2(2) \\ \vdots \\ u_2(n_m) \\ \vdots \\ u_{n_p}(1) \\ u_{n_p}(2) \\ \vdots \\ u_{n_p}(n_m) \end{bmatrix}, \mathbf{f}(\mathbf{x}) \equiv \begin{bmatrix} g({}^w\mathbf{r}_1, \alpha(1)) \\ g({}^w\mathbf{r}_1, \alpha(2)) \\ \vdots \\ g({}^w\mathbf{r}_1, \alpha(n_m)) \\ g({}^w\mathbf{r}_2, \alpha(1)) \\ g({}^w\mathbf{r}_2, \alpha(2)) \\ \vdots \\ g({}^w\mathbf{r}_2, \alpha(n_m)) \\ \vdots \\ g({}^w\mathbf{r}_{n_p}, \alpha(1)) \\ g({}^w\mathbf{r}_{n_p}, \alpha(2)) \\ \vdots \\ g({}^w\mathbf{r}_{n_p}, \alpha(n_m)) \end{bmatrix}$$

$$\mathbf{x} \equiv [{}^w\mathbf{r}_1^T, {}^w\mathbf{r}_2^T, \dots, {}^w\mathbf{r}_{n_p-1}^T, {}^w\mathbf{r}_{n_p-2}, {}^w\mathbf{y}_{n_p-2}, \alpha(1)^T, \dots, \alpha(n_m)^T]^T \quad (9)$$

ただし、特徴点 i が時刻 j において観測されていなければ、 $u_i(j)$ に関する式は削除される。基準となるワールド座標系の大きさと向きを決めるため、 ${}^w\mathbf{r}_{n_p}$ 、 ${}^w\mathbf{r}_{n_p-1}$ 、 ${}^w\mathbf{z}_{n_p-2}$ は既知として、 $\mathbf{x}$ を構成した。なお、 n_p は特徴点の全数、 n_m は最大時刻を表す。

(2) 非線形最小2乗法によるパラメータの推定法 アルゴリズムが簡明で収束の良い非線形最小2乗法、Marquardt 法¹¹⁾ を適用して、式 (8) の非線形観測方程式から未知パラメータを推定する。まず、 $\mathbf{f}(\mathbf{x})$ を k 次の推定値 $\hat{\mathbf{x}}^{(k)}$ を用いて1次近似して、式 (8) に代入し、線形近似観測方程式

$$\Delta \mathbf{y}^{(k)} \approx \mathbf{A}^{(k)} \Delta \mathbf{x}^{(k)} \quad (10)$$

を得る。ここで、

$$\begin{cases} \Delta \mathbf{y}^{(k)} \equiv \mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}^{(k)}) \\ \mathbf{A}^{(k)} \equiv \frac{\partial \mathbf{f}(\hat{\mathbf{x}}^{(k)})}{\partial \mathbf{x}} \\ \Delta \mathbf{x} \equiv \mathbf{x}^{(k+1)} - \hat{\mathbf{x}}^{(k)} \end{cases} \quad (11)$$

$A^{(k)}$ はヤコビアン行列と呼ばれ、具体的には、

$$\frac{\partial g(wr_i, \alpha(j))}{\partial wr_i} = \frac{\delta_{ij'}}{e_3^T wr_i + c_{i_3}} [C^c R_w - g(wr_i, \alpha(j)) e_3^T] \quad (12)$$

$$\frac{\partial g(wr_i, \alpha(j))}{\partial \phi(j')} = \frac{\delta_{ij'}}{e_3^T wr_i + c_{i_3}} \left[C \frac{\partial^c R_w}{\partial \phi(j)} wr_i - g(wr_i, \alpha(j)) wr_i^T \frac{\partial e_3}{\partial \phi(j)} \right] \quad (13)$$

$$\frac{\partial g(wr_i, \alpha(j))}{\partial^c t_w(j')} = \frac{\delta_{ij'}}{e_3^T wr_i + c_{i_3}} [C - g(wr_i, \alpha(j)) [0, 0, 1]] \quad (14)$$

式 (10) から未知パラメータ x の修正ベクトル Δx の最尤推定値 $\Delta \hat{x}^{(k)}$ を、

$$\Delta \hat{x}^{(k)} = (A^{(k)T} \Sigma^{-1} A^{(k)} + \lambda I)^{-1} A^{(k)T} \Sigma^{-1} \Delta y^{(k)} \quad (15)$$

により求め、未知パラメータベクトルを反復改良していく。この Marquardt 法は、解から離れ非線形性の影響が大ききときには λ を大きくして、解を最急降下方向に小さく修正することにより頑健性を高める。解に近づくにつれて λ を小さくして、Gauss-Newton 解に近づけて大きく修正することにより高速化させる。 Σ は観測値ベクトル y の誤差の共分散行列である。

3. 収束性と計測精度

3.1 求解の条件

未知パラメータの値が決定されるための必要条件は、未知パラメータの個数 $(3(n_p - 3) + 2 + 6n_m)$ が観測データの個数 $(2n_p n_m)$ 以下となることである。特にすべての特徴点がすべての画像に撮影されている場合には⁸⁾、上記の関係を満たす最小の n_p, n_m は、それぞれ 5, 2 である。このとき、未知パラメータの個数および観測値の個数はともに 20 個となる。以下のシミュレーションおよび実験はこの基本的な条件で行なった。個々の未知パラメータについては、各特徴点が少なくとも 2 つのモーションで観測されること、各モーションにおいて、少なくとも 3 つの特徴点が観測されることが必要である。

3.2 収束性

3次元物体として Fig.2 に示すような特徴点の数 $n_p = 5$ の四角錐を考える。四角錐の頂点の高さを h とする。各特徴点を 0~4 とする。基準座標系決定のため、特徴点 2 の wz 座標および、特徴点 3, 4 の wx, wy, wz 座標の 7 個の値を既知として与える。また、Fig.3 に示すように、常に辺長 32 の範囲が視野 20 deg のカメラで観測できる位置にカメラを動かすものとする。つまり、

視野角が円弧角と等しい円周上を動くものとし、この円周は $z-x$ 平面上に存在するものとする。以下に示す収束性確認のシミュレーションでは、 h を 0, θ を +45 deg と -45 deg に設定した。なお、特徴点の座標およびカメラの位置の単位は無次元量とした。未知パラメータの初期値には、真値を中心に位置は各方向に最大 30, 角度は最大 30 deg の幅でランダムにずらした値を用いた。パラメータはいずれも数回の反復で収束した。

Fig.2 3-D object used for simulation

Fig.3 Camera motion and geodesic dome

ここで、正しく収束できる未知パラメータの初期値 $\hat{x}^{(0)}$ について調べる。具体的には未知パラメータの特徴点とカメラの位置に関する部分とカメラの回転角に関する部分について、未知パラメータの初期値を真値からある一定の範囲内でランダムにずらして収束の成否を調べた。Fig.4 は上記の試行を 3 回ずつ行った結果であり、○, ●, ▲, × はそれぞれ真値に収束した回数が 3, 2, 1, 0 回であったことを示す。Fig.4 から、初期値が 3 次元物体の最長の辺長 16 と同程度にずれても収束している。角度についても数 10 度ずれていても収束している。この他の方法として、最も広く用いられている Gauss-Newton 法を適用してみたが、Fig.4 のいずれの条件でも真値に収束しないことがあった。このように、本研究の未知パラメータ推定に際して Marquardt 法が初期値の誤差に対して頑健で、実用的であることがわかる。

3.3 誤差解析

(1) パラメータ推定値の誤差 誤差を含んだ画像座標の測定値から計算したものであるから、パラメータ推定値も誤差を含む。式 (10) を未知パラメータの真値の

Fig.4 Convergence characteristics

まわりの関係式と見なすと、パラメータの推定値の誤差行列 $\Sigma_{\hat{x}}$ は次式で与えられる¹¹⁾。

$$\Sigma_{\hat{x}} = (A^T \Sigma^{-1} A)^{-1} \quad (16)$$

(2) カメラモーションに関する条件 Fig.3 に示したように円周に沿って中心線に対称にカメラが動くものとする。カメラの2つの視軸がなす輻輳角度 2θ をパラメータとし未知パラメータの推定誤差の分散が最小となるカメラの位置関係について調べる。

(3) 計算結果 式(16)により与えられる $\Sigma_{\hat{x}}$ の対角成分において特徴点の位置に関する部分の平均を縦軸に、輻輳角度 2θ を横軸にとって Fig.5 に示す。また、 Σ は単位行列、画像座標の測定誤差の分散は 1pixel^2 とする。図から、 $h=0$ のとき $2\theta = 136.0\text{deg}$ で極小となることがわかる。 $h=10$ のとき $2\theta = 134.0\text{deg}$ で分散が大きくなるのは、カメラと特徴点の位置が Fig.6 のような関係となり、特徴点0の x 座標が決定できないためである。いずれにおいても、成瀬らが測定対象を中心とする円周上を両眼カメラが運動する場合について求めた角度 90deg ¹⁾よりやや大きくなっている。これはカメラが対象に近接して精度が向上する効果が重畳されるためである。なお Fig.5 には、 $h=0, 2\theta=100\text{deg}, 20\text{deg}$ の条件で、各々10回行った実験結果の結果が●で示されている。実線で示される理論値はこの実験結果に良く一致しており、妥当であることがわかる。

Fig.5 Optimizing camera motion

Fig.6 Camera positions in ill-condition

4. むすび

視線が輻輳するように移動する単眼の画像から対象の3次元位置とカメラモーションを、Marquardt法を用いて同時に推定する方法について、

(1) 解の収束性 特徴点の位置およびカメラの並進移動の初期値が真値から3次元物体を構成する特徴点間隔と同程度ずれ、回転角の初期値が真値から 20deg ずれていても正しく収束した。これにより、未知パラメータの初期値の誤差に対する頑健性が確認できた。

(2) 誤差解析 (2視点5特徴点の最簡条件に対するカメラモーションの最適化) 対象とする3次元物体の全体像を常に視野に収めるように、カメラモーションを円周上に限定した場合について、特徴点の位置に関する未知パラメータ1個あたりの分散を調べた。その結果、2つの視軸のなす輻輳角が 90deg よりやや大きいときに最適となることが明らかになった。

今後の課題としては、3次元的なカメラモーションの最適化、大規模3次元物体を用いた実験などがあげられる。本論文をまとめるにあたり、御助言、御協力いただいた橋本明治氏に感謝の意を表します。

参考文献

- (1) 成瀬・立田・野引・藪田：高精度多視点ステレオ画像計測方法，日本機械学会論文誌 C, Vol.59, No.557, pp.112-118 (1992)
- (2) Z.Zhang and O.D.Faugeras: 3-D Dynamic Scene Analysis, Springer Verlag (1992)
- (3) 高橋・富田：ロボットビジョンのためのカメラキャリブレーション，日本ロボット学会誌, Vol.10, No.2, pp.177-184 (1992)
- (4) 金谷：画像理解，森北出版, pp.79-114 (1990)
- (5) 木下・出口：能動視覚による3次元形状認識，計測自動制御学会論文誌, Vol.28, No.1, pp.144-153 (1992)
- (6) 服部：非線形写真測量のキャリブレーション法，写真測量とリモートセンシング, Vol.25, No.4, pp.19-25 (1986) 計測自動制御学会論文誌, Vol.28, No.1, pp.144-153 (1992)
- (7) T.J.Broida and R.Challappa: Performance Bounds for Estimating Three-dimensional Motion Parameters from a Sequences of Noisy Images, Journal of Opt.Soc.Am.A, Vol.6, No.6, pp.879-889 (1989)
- (8) A.Z.Meiri: On monocular perception of 3-D moving object, IEEE Trans. PAMI, Vol.2, No.6, pp.582-583 (1980)
- (9) L.Matthies and T.Kanade: Kalman Filter-based Algorithms for Estimating Depth from Image Sequences, International Journal of Computer Vision, 3, pp.209-236 (1989)
- (10) 奥富・金出：複数の基線長を利用したステレオマッチング，電子情報通信学会論文誌 D-II, Vol.J75-D-II, No.8, pp.1317-1327 (1992)
- (11) 中川・小柳：最小二乗法による実験データ解析，東京大学出版会, pp.29-53, 95-99 (1989)