

### 3次元形状・運動復元のための高速非線形最適化計算法

右田 剛史<sup>†,☆</sup> 天野 晃<sup>††</sup> 浅田 尚紀<sup>†</sup>

画像列から物体形状とカメラ運動を同時に復元する運動からの形状復元は、非線形最適化問題として定式化できるが、安定化（局所解回避）と高速化（計算量低減）が大きな課題である。本論文では、非線形最適化アルゴリズムの1つである共役勾配法において、ヘッセ行列をブロック対角行列で近似した前処理行列を用いて計算量を低減する手法を提案する。シミュレーションおよび実画像の12種類のデータセットを用いた実験の結果、未知数が1,000個規模の問題に対して、従来のLevenberg-Marquardt法および共役勾配法の数倍から十数倍の高速化の効果を確認した。

### A Fast Nonlinear Optimization Algorithm for 3D Shape and Motion Recovery

TSUYOSHI MIGITA,<sup>†,☆</sup> AKIRA AMANO<sup>††</sup> and NAOKI ASADA<sup>†</sup>

Simultaneous recovery of shape and motion from image sequences is formulated as a nonlinear optimization problem. This paper proposes a fast algorithm named "block diagonal matrix preconditioned conjugate gradient method" characterized by block diagonalized approximation of Hessian. Experimental results using real and synthetic image data have shown that our algorithm reduces the calculation time by 80% to 95% compared with the Levenberg-Marquardt and conjugate gradient methods.

#### 1. はじめに

物体を移動カメラで撮影した画像列から物体形状とカメラ運動を同時に復元する「運動からの形状復元」はコンピュータビジョンの基本問題であると同時に、実物体の3次元モデル化や複合現実への利用など様々な応用が期待される技術である。

形状復元のもっとも基本的な解法は、2枚の画像間の8個以上の対応点を用いた8点アルゴリズム<sup>1),2)</sup>であるが、この手法はノイズに敏感であり<sup>3)</sup>、安定な形状復元が困難であることが知られている<sup>4)</sup>。この問題を解決するには、3枚以上の画像を用いて誤差を最小化することが有効であり、その代表的な手法である因子分解法<sup>5)</sup>は、最小2乗法による形状復元を線形計算を用いて高速かつ安定に唯一の解を得ることができる。しかし、定式化の性質上、適用できる画像列には、透視投影が線形投影として十分に近似できること、同じ特徴点が全画像を通じて観測されていること、という制

約がある。これらの条件を緩和し因子分解法の適用範囲を拡張する研究が行われている<sup>5)~9)</sup>が、それらは線形解法の長所（高速性、安定性、解の唯一性）を損なうという問題がある。たとえば文献5)では、一部の特徴点に隠蔽が生じる場合への適用について検討しているが、解の唯一性が保証されないことが知られている<sup>10)</sup>。

一方、このような線形化を行わなくても、推定した物体特徴点を投影したときの座標と観測特徴点座標との誤差の2乗和を最小化することによって運動と形状を復元することができる<sup>3),11)~14)</sup>。この方法は、2乗誤差最小という意味で最適な解が得られるうえ、線形解法と比べて以下のような特徴がある<sup>11)</sup>。

- 透視投影はもちろん種々の投影モデルに適用できる。
- 画像列の途中で、新たな特徴点の出現や観測中の特徴点の隠蔽などによる特徴点の不完全対応が生じてても定式化に影響しない。
- 形状や運動に関する制約の導入、点特徴と線特徴の同時使用など、柔軟性、拡張性が高い。

しかし、非線形変換である透視投影を対象に特徴点の隠蔽も扱う場合の定式化は非線形最適化問題となり、計算上次の2つの課題がある。

- 安定性（適切な初期値の設定と局所解の回避）

<sup>†</sup> 広島市立大学情報科学部知能情報システム工学科  
Department of Intelligent Systems, Hiroshima City University

<sup>††</sup> 京都大学大学院情報学研究科システム科学専攻  
Graduate School of Informatics, Kyoto University

<sup>☆</sup> 現在、岡山大学工学部情報工学科  
Presently with Okayama University

一般に非線形最適化問題では、大局的な最適解を得ることが容易ではなく、適切な初期値の設定と局所解の回避法が重要となる。文献15)では、オプティカルフローを用いた形状復元において、鏡映対称と bas-relief ambiguity に起因する局所解が存在することを述べているが、局所解の性質やその回避法について詳細に議論した研究は少ない。従来の研究では、前述の線形解法やその拡張手法を用いて得られた近似解を初期値とし、最適解に補正する方法 (bundle adjustment)<sup>14)</sup> が用いられている。

#### ● 高速性 (反復計算に要する計算量の削減)

多くの研究では、Levenberg-Marquardt法<sup>18)</sup> (以下 LM 法) を用いて非線形最適化計算を行っているが、問題の規模が大きくなると計算負荷が急速に増大するため、高速化の工夫が必要となる。ヘッセ行列が特殊な疎構造となる性質を利用した直接法<sup>13)</sup> やスカイライン記憶法による LU 分解<sup>11)</sup> などの高速化手法が報告されているが、LM 法を用いる限り 1 反復あたりの計算量は特徴点数か画像数の 3 乗のオーダーとなる。

安定性については、対象とする問題の性質に応じた工夫が必要であり一般的な対策を議論することは難しい。それに対して、高速性については、反復計算による解探索の計算量の削減が主な課題であり、計算アルゴリズムの改良についてある程度一般的に議論することが可能である。

そこで本論文では、非線形最適化問題の高速化に焦点を絞り、運動からの形状復元における高速計算アルゴリズムとして、LM 法と共役勾配法の長所を組み合わせた「ブロック対角行列による前処理つき共役勾配法<sup>19)</sup>」を提案する。このアルゴリズムの計算量は特徴点数と画像数の積に比例するため、問題規模が大きくなるほど LM 法よりも有利になり、未知数が 1000 程度のデータでは LM 法の約 1/12 の計算時間で解を求めることができる。以下、2 章では透視投影における形状復元問題の定式化について述べ、3 章で非線形最適化手法の改良と前処理つき共役勾配法<sup>\*</sup>の特徴について述べる。4 章でシミュレーションおよび実データによる実験と評価を行い、5 章でまとめを述べる。

<sup>\*</sup> 文献 14) では、LM 法において現れる連立一次方程式を解くために前処理つき共役勾配法を用いているが、全体の計算量は特徴点数か画像数の 3 乗のオーダーのままである。本論文では、非線形最適化の解法としての共役勾配法に対して前処理を適用している点異なる。

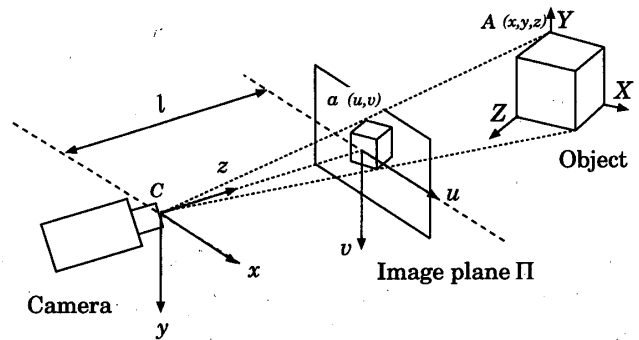


図 1 投影モデル

Fig. 1 Projection model.

## 2. 非線形定式化による形状復元

### 2.1 剛体運動と透視投影

物体を剛体とし、 $P$  個の特徴点で形状を表す。各特徴点は物体座標系に対して静止しており、その座標を  $s_p = (x_p, y_p, z_p)^T$ , ( $1 \leq p \leq P$ ) とする。物体を撮影した  $F$  枚の画像列を考え、第  $f$  画像 ( $1 \leq f \leq F$ ) の物体座標系とカメラ座標系の関係が、回転行列  $R_f$  と並進ベクトル  $t_f = (t_{fx}, t_{fy}, t_{fz})^T$  で表されるとする。これらのパラメータを用いると、第  $f$  画像における第  $p$  特徴点のカメラ座標  $s_{fp} = (x_{fp}, y_{fp}, z_{fp})^T$  は、次式の剛体変換によって表される。

$$s_{fp} = R_f s_p + t_f \quad (1)$$

次に、この点がピンホールカメラで投影される場合を考える (図 1)。レンズ中心を  $C$ 、画像面を  $\Pi$ 、焦点距離を  $l$  とし、 $C$  から  $\Pi$  の方向を  $z$  軸正の方向、カメラ座標系  $x, y$  軸に対応する画像面の座標系を  $u, v$  軸とする。

第  $f$  画像における第  $p$  特徴点の 3 次元座標  $s_{fp}$  と、その特徴点の画像上の 2 次元座標  $u_{fp} = (u_{fp}, v_{fp})^T$  の関係は、投影変換の作用素を  $\mathcal{P}$  とすると次式で表される。

$$u_{fp} = \mathcal{P} s_{fp} = \mathcal{P}(R_f s_p + t_f) \quad (2)$$

透視投影の場合、 $\mathcal{P}$  は次式となる。

$$\mathcal{P} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{l}{z} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3)$$

非線形最適化計算において回転行列を表現する場合、4 元数<sup>☆☆</sup>を用いると都合がよい<sup>11), 12), 17)</sup>。4 元数  $q =$

<sup>☆☆</sup> 4 元数は、他の表現法 (Euler 角, Roll-pitch-yaw, 回転軸と回転角, 歪対称行列の指数関数など<sup>22)</sup>) と比べて、三角関数を必要としない、微分計算が容易である、パラメータの対称性が高く特異点や不均一性がない<sup>14)</sup> などの利点がある。ただし、本来自由度が 3 しかない回転行列を 4 つのパラメータで表すため、冗長な表現となっている。

$$Q(\mathbf{q}) = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix}$$

図2 4元数とスケール回転行列

Fig. 2 A quaternion and its associated scaled-rotation matrix.

$(q_0, q_1, q_2, q_3)$  は、図2に示す回転行列の定数倍（スケール回転行列）に対応し、ノルム  $(q_0^2 + q_1^2 + q_2^2 + q_3^2)$  を1に正規化した4元数は回転行列を表す。

ここで、透視変換（式(3)）では任意の実数  $k \neq 0$  に対して  $\mathcal{P}\mathbf{s} = \mathcal{P}(k\mathbf{s})$  が成立するスケール不変性に着目すると、回転並進に続いて透視変換が作用する式(2)は次のように変形することができる。

$$\begin{aligned} \mathbf{u}_{fp} &= \mathcal{P}[(k_f R_f) \mathbf{s}_p + (k_f \mathbf{t}_f)] \\ &= \mathcal{P} \left( Q(\mathbf{q}_f) \mathbf{s}_p + \begin{bmatrix} X_f \\ Y_f \\ Z_f \end{bmatrix} \right) \end{aligned} \quad (4)$$

この式では、スケール回転行列を表す4元数と  $k$  倍された並進ベクトル  $(X_f, Y_f, Z_f)^T$  の7つのパラメータによってカメラの位置・姿勢を表現している。本来6自由度の位置・姿勢を7つのパラメータ  $(q_0, q_1, q_2, q_3, X, Y, Z)$  で表現しているため、任意の実数  $k$  に対して曲線  $(kq_0, kq_1, kq_2, kq_3, k^2X, k^2Y, k^2Z)$  上の点はすべて同一の位置・姿勢を表していることになるが、非線形最適化計算ではこの曲線上のどの点に収束してもよい。経験的にも、1自由度増えることによる収束性への悪影響がないことを確認している<sup>\*</sup>。4元数を正規化しないことにより、計算量が低減でき、負の数の平方根や0による除算などの問題を避けることもできるため、本研究では投影モデルとして式(4)を用いることとした。

## 2.2 形状復元の定式化と高速化の準備

形状復元は、投影モデルに基づいて物体特徴点を投影した画像座標（式(4)）と、観測された画像上の特徴点座標  $\tilde{\mathbf{u}}_{fp} = (\tilde{u}_{fp}, \tilde{v}_{fp})^T$  の差の2乗和を最小化する形状パラメータ  $\mathbf{s}_p$  と運動のパラメータ（4元数  $\mathbf{q}_f$  とスケール並進ベクトル  $(X_f, Y_f, Z_f)^T$ ）を求める問題として定式化される。

<sup>\*</sup> 文献14)では、原点から非常に遠い特徴点は3パラメータ  $(x, y, z)$  で表現するよりも、射影的に4パラメータ  $(x, y, z, w) \rightarrow (x/w, y/w, z/w)$  で表現する方がよいとしている。これは、3パラメータ表現ではきわめて大きな数に収束するには多数の反復を要するが、4パラメータ表現では  $w$  を0に近づけるだけで無限遠近くの点を表現でき、この方が最適化の反復回数にはるかに少なく済むためである。本論文の場合も、これと同様の冗長性であり不安定性を生じるものではない。

$$\arg \min_{\mathbf{s}_p, \mathbf{q}_f, X_f, Y_f, Z_f} \sum_{f,p} |\tilde{\mathbf{u}}_{fp} - \mathbf{u}_{fp}|^2 \quad (5)$$

ここで、運動と形状のパラメータを合わせた  $N = 3P + 7F$  個の未知数を含むベクトル  $\mathbf{w}$

$$\mathbf{w} = (x_1, y_1, z_1, \dots, x_P, y_P, z_P, X_1, q_{10}, q_{11}, q_{12}, q_{13}, Z_1, Y_1, \dots, X_F, q_{F0}, q_{F1}, q_{F2}, q_{F3}, Z_F, Y_F)^T \quad (6)$$

と、推定座標  $\mathbf{u}_{fp}$  と観測座標  $\tilde{\mathbf{u}}_{fp}$  の残差を  $FP$  個並べたベクトル  $\mathbf{r}$

$$\mathbf{r}(\mathbf{w}) = (r_{111}, r_{112}, \dots, r_{FP1}, r_{FP2})^T \quad (7)$$

$$\begin{bmatrix} r_{fp1} \\ r_{fp2} \end{bmatrix} = \begin{bmatrix} u_{fp} - \tilde{u}_{fp} \\ v_{fp} - \tilde{v}_{fp} \end{bmatrix} \quad (8)$$

を定義すると<sup>\*\*</sup>、式(5)は次の関数  $E$  の最小化問題と等価となる<sup>\*\*\*</sup>。

$$E(\mathbf{w}) = \frac{1}{2} \mathbf{r}(\mathbf{w})^T \mathbf{r}(\mathbf{w}) \quad (9)$$

ここで、点  $\mathbf{w}$  の近傍  $\mathbf{w} - \mathbf{d}$  における  $E$  の2次近似は

$$E(\mathbf{w} - \mathbf{d}) \approx E(\mathbf{w}) - \mathbf{d}^T \mathbf{J}^T \mathbf{r}(\mathbf{w}) + \frac{1}{2} \mathbf{d}^T (\mathbf{J}^T \mathbf{J}) \mathbf{d} \quad (10)$$

となり、ヤコビ行列  $\mathbf{J}$  は次式で与えられる。

$$\mathbf{J} = \begin{bmatrix} \frac{\partial r_{111}}{\partial w_1} & \frac{\partial r_{111}}{\partial w_2} & \dots & \frac{\partial r_{111}}{\partial w_N} \\ \frac{\partial r_{112}}{\partial w_1} & \frac{\partial r_{112}}{\partial w_2} & \dots & \frac{\partial r_{112}}{\partial w_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_{FP2}}{\partial w_1} & \frac{\partial r_{FP2}}{\partial w_2} & \dots & \frac{\partial r_{FP2}}{\partial w_N} \end{bmatrix} \quad (11)$$

## 3. 非線形最適化計算の高速化

式(9)の最小化、すなわち非線形最適化問題の解探索には、勾配に基づく反復法を用いる。 $\mathbf{w}_k$  をある時点での推定値、 $\mathbf{d}_k$  を勾配  $\mathbf{g}_k = \nabla E(\mathbf{w}_k)$  から計算される探索方向とすると、反復計算は

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \alpha_k \mathbf{d}_k \quad (12)$$

ただし、 $\alpha_k = \arg \min_{\alpha} E(\mathbf{w}_k - \alpha \mathbf{d}_k)$

<sup>\*\*</sup> 隠蔽などにより  $\tilde{\mathbf{u}}_{fp}$  が観測されない時は  $r_{fp1} = r_{fp2} = 0$  とする。

<sup>\*\*\*</sup> 計算の便宜上係数を  $(1/2)$  としている。

表1 非線形最小二乗法における探索方向

Table 1 Search direction of non-linear least squares method.

手法	探索方向 $\mathbf{d}_k$
Levenberg-Marquardt 法	$\dot{H}_k^{-1} \mathbf{g}_k$
共役勾配法	$\mathbf{g}_k + \beta_k \mathbf{d}_{k-1}$
前処理つき共役勾配法	$\ddot{H}^{-1} \mathbf{g}_k + \beta_k \mathbf{d}_{k-1}$

と表され、 $\mathbf{w}_k$  が収束するまでこの計算を繰り返す。探索方向  $\mathbf{d}_k$  の選び方によって計算効率が大きく左右されるが、非線形最適化の代表的手法である LM 法と共役勾配法 (Conjugate Gradient 法, 以下 CG 法) は、表 1 に示す値を用いる。表中、 $\dot{H}_k$ ,  $\ddot{H}$  はヘッセ行列 ( $\partial^2 E / \partial w_i \partial w_j$ ) の近似、 $\beta_k$  はパラメータである。

式 (12) 中の  $\arg \min_{\alpha} E(\mathbf{w}_k - \alpha \mathbf{d}_k)$  は、 $\alpha$  の関数  $E(\alpha)$  を最適化する直線探索を行うが、厳密に  $E(\alpha)$  を最小化する必要はなく、たとえば Newton 法や LM 法では探索せずに  $\alpha = 1$  とすることもできる。しかし、CG 法では 2 次関数に対する最適性に基づいて適切な  $\alpha$  を求める必要があるため、本研究では 2 次近似による方法で直線探索を行うこととした。詳細は付録に示すが、この方法は黄金分割法<sup>18)</sup>よりはるかに少ない回数で探索を終了する。

### 3.1 Levenberg-Marquardt 法 (LM 法)

LM 法は、ヘッセ行列を  $\dot{H}_k = (J_k^T J_k + \mu_k I)$ , ( $\mu_k$  はパラメータ) で近似するもので、解付近では 2 次収束に近い性能を示す優れた手法である。しかし、反復ごとに大規模な連立方程式を解く必要があるため、計算量が未知パラメータの数  $N$  の 3 乗のオーダーとなり、大規模問題には適さない。

形状復元問題では  $\dot{H} = J^T J + \mu I$  は次のような構造の  $N$  次対称行列となる。空白部分は 0 要素を表す。

$$\dot{H} = \begin{array}{|c|c|} \hline \begin{array}{c} S \\ \hline V \end{array} & \begin{array}{c} U \\ \hline T \end{array} \\ \hline \end{array} \quad (13)$$

$S$  と  $T$  の部分はブロック対角行列であり、ブロックの大きさはそれぞれ  $3 \times 3$  と  $7 \times 7$  である。この行列の特殊構造を利用して、密行列に対するよりも高速に逆行列を求める方法が知られている<sup>11),13),14)</sup>。しかし、この行列には 0 要素が多いものの、Cholesky 分解の計算量を  $O(\min(F^3, P^3))$  以下に削減することはできない。文献 13) では、時系列画像を仮定した近似によって、特徴点数に比例する計算量に抑えているが、この近似が一般に有効であるとは限らない。

### 3.2 共役勾配法 (CG 法)

CG 法は、反復回数は多いが連立方程式を解く必要がないという特徴を持つ。LM 法との優劣は、対象とする問題の規模や性質に左右されるため一概にはいえないが、本研究で対象とする規模 (未知数が数千程度) の形状復元問題に対しては、ほぼ同等の性能であることを後述の実験により確認している。

なお、文献 12) では LM 法は CG 法よりも格段に効率が良いと述べているが、これは問題規模が小さく Cholesky 分解の手間が大きな問題にならないためと考えられる。

### 3.3 前処理つき共役勾配法

線形連立方程式に対する CG 法は、係数行列の固有値を密集させる前処理によって格段に高速化されることが知られている<sup>20)</sup>が、非線形最適化問題に対しても前処理行列を導入することにより同様の効果が期待できる。この前処理つき共役勾配法 (Preconditioned CG 法, 以下 PCG 法) の計算は以下ようになる。

初期探索点  $\mathbf{w}_0$  では、その点での関数  $E$  の勾配  $\mathbf{g}_0$  と前処理行列  $\ddot{H}$  を用いて探索方向  $\mathbf{d}_0 = \ddot{H}^{-1} \mathbf{g}_0$  を求め、2 回目以降は次式を用いて探索方向を決定する。

$$\mathbf{d}_k = \ddot{H}^{-1} \mathbf{g}_k + \beta_k \mathbf{d}_{k-1} \quad (14)$$

$$\text{ただし、} \beta_k = \frac{(\ddot{H}^{-1} \mathbf{g}_k)^T (\mathbf{g}_k - \mathbf{g}_{k-1})}{\mathbf{g}_{k-1}^T \ddot{H}^{-1} \mathbf{g}_{k-1}}$$

$\beta$  の決定は Polak-Rebière-Polyak の式<sup>18)</sup>を拡張したものであり、 $\ddot{H}$  を単位行列とすると PCG 法は CG 法に一致する (表 1)。式 (14) で得られた探索方向  $\mathbf{d}_k$  を式 (12) に代入することによって探索点  $\mathbf{w}$  が更新される。

PCG 法では、効果的な前処理行列を構成することが求められるが、我々は、形状復元問題におけるヘッセ行列が特殊な疎構造であることに着目し、ブロック対角行列による前処理つき共役勾配法 (Block Diagonal matrix preconditioned CG method, 以下 BDCG 法) を考案した。

### 3.4 BDCG 法

形状復元問題のヘッセ行列の近似は式 (13) の構造を持ち、 $U$ ,  $V$  部分に比べて  $S$ ,  $T$  のブロック対角部分の要素の絶対値が非常に大きいことが知られている<sup>13)</sup>。そこで、 $\dot{H}$  の  $U$  と  $V$  の部分を 0 で置き換えた行列を、前処理行列  $\ddot{H}$  として用いることとした。

このブロック対角行列は、 $O(N)$  の計算量で Cholesky 分解が可能であり、必要な記憶容量も ( $N^2 \rightarrow 7N$ ) と大幅に削減することができる。なお、 $U$  や  $V$  の各要素は 2 次元ベクトルの内積であるのに

対して、対角部分  $S, T$  はそれぞれ  $6P$  次元,  $14F$  次元ベクトルの内積で表されるため,  $F$  と  $P$  がともに大きいほど対角成分の値も大きくなり,  $U, V$  を無視することによる相対誤差は小さくなる.  $U, V$  を省略することにより LM 法の  $\dot{H}^{-1}\mathbf{g}$  と比べて探索方向に誤差が含まれることになるが, この影響を共役方向生成 (表 1 の  $\beta$  の項) によって補うのが本手法の特徴である.

なお, 式 (14) において  $\ddot{H}^{-1}$  は毎回更新するわけではなく共役方向生成によって探索を進めるため, 探索点が  $\ddot{H}$  を求めた点から離れるに従って探索方向  $\mathbf{d}_k$  が最適探索方向からずれ, 解の収束性が低下する. そこで, 通常の CG 法における定期的な再出発<sup>18)</sup>を拡張し, 再出発と同時に  $\ddot{H}$  を更新することとした. 解の収束性を高めるためには, 反復計算における探索方向の精度をある程度維持する必要があるため, 通常の CG 法の再出発に比べて周期を短く設定し, 後述の実験では 16 反復ごとに  $\ddot{H}$  の再計算を行った. この再出発による解収束の効果と計算コストを LM 法と比較すると次のように考えることができる.  $\ddot{H}$  の再計算の計算量が  $O(FP)$  であるのに対して LM 法の 1 反復の計算量が  $O(\min(F^3, P^3))$  であることを考えると,  $\ddot{H}$  の計算から次の再計算までの間の解探索によって LM 法の 1 反復と同等の解収束の効果が得られれば, 計算コストの点で十分有利である.

次に, BDCG 法は CG 法と同様に LM 法に比べて多くの反復回数を要する点について考察する. 線形連立方程式に対する前処理付き共役勾配法の反復回数は, 未知数の数よりも係数行列の固有値の密集具合によることが知られている<sup>20)</sup>. 同様に, 非線形最適化問題においても反復回数を決定する要素はヘッセ行列の固有値の分布となる. 文献 4) の解析を考慮すると, 反復回数は特徴点の配置と運動の性質によって決まるものと考えられ, 実験的にも反復回数が特徴点や画像数にほとんど影響されないことを確認している. したがって, 反復回数は少ないが計算量が  $O(\min(F^3, P^3))$  の LM 法より, 反復回数が問題の規模にほぼ無関係で計算量が  $O(FP)$  の BDCG 法の方が, 未知パラメータの数  $N$  が大きな問題になるほど計算速度の点で有利になる.

なお, BDCG 法は解の探索方向を求める際にヘッセ行列の近似を行っているが, 式 (9) そのものの最小化については何ら近似を行っていないため, 理論的にも実験的にも LM 法や CG 法と同じ精度の解が得られることを確認している.

## 4. 実 験

提案手法の有効性を確認するためシミュレーションおよび実画像を用いた実験を行った. ただし, 従来の LM 法, CG 法と本論文で提案する BDCG 法の計算速度を比較することが目的であるため, 入力画像間で対応関係が既知の特徴点集合とし, 次節で述べる初期値を用いて大局的な最適解が得られる問題に適用した. 実験に用いた計算機は, Pentium 4, 2.0 GHz, メモリ 1 GB, OS は Vine Linux 2.5 である.

なお, 性能比較のために実装した LM 法は行列の疎構造を利用した Cholesky 分解を用いたものであり, 従来の高速化 LM 法<sup>11),13)</sup>と同等の性能である.

### 4.1 初期値の設定

非線形最適化問題で正しい解を得るには, 適切な初期値を与えて局所解を回避する必要がある. 因子分解法や 8 点アルゴリズムに基づく手法<sup>3),23),24)</sup>が適用可能な場合, それらの解を初期値とすることができ, 実問題では隠蔽が生じやすく必ずしも線形解法が適用できるわけではない. そのため, 実用上は一般的な初期値からの収束性能が重要であることを考慮して実験では次の初期値を用いた.

$$\begin{aligned} \mathbf{s}_p &= (0, 0, 0)^T, & (X_f, Y_f, Z_f) &= (0, 0, 1), \\ \mathbf{q}_f &= (\cos(\omega f), 0, 0, \sin(\omega f))^T \end{aligned}$$

これらの値は, 静止物体の周囲を移動しながらカメラで撮影する状況を想定したもので, カメラが物体を注視して一定距離 (並進ベクトル一定) を保ったまま周囲を角度  $2\omega F$  ( $\omega$  は角速度,  $F$  は画像数) だけ移動することを表している. ただし, 物体形状は未知であるため, 全特徴点が原点に縮退した状態 ( $\mathbf{s}_p = (0, 0, 0)^T$ ) を初期値としている. このような単純な初期値を用いても, 多くの場合で最適解に収束することを実験的に確認している. 以下, この初期値をゼロ初期値と呼ぶ.

### 4.2 シミュレーション実験

物体形状は球とし, 表面上に  $P = 60, 120, 240$  個の特徴点をほぼ均等に配置した 3 種類の物体に対して, それぞれの周囲を 360 度回転する間に画像数  $F = 25, 50, 100$  枚を撮影した 3 画像列を作成し, 全部で 9 例のデータセットを用いて形状復元を行った. 図 3 が 240 点の球の例であり三角形パッチの頂点が特徴点を表す. 9 種類のデータの未知数の個数 ( $N = 3P + 7F$ ) は表 2 に示すように最小 355, 最大 1420 である. このデータでは, 半径 100 の球が距離 1000 程度離れて投影されているので射影歪みは大きくないが, どの画像も全特徴点の半数が球の裏側で見えないという隠蔽を生じている. なお, 画像上での物体の像の大きさは

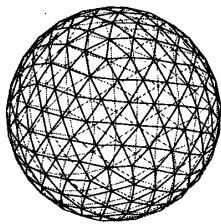


図3 シミュレーション実験に用いた球  
Fig. 3 Spherical object for simulation study.

表2 未知数の個数  
Table 2 Number of unknown variables.

Frames\Points	60	120	240
25	355	535	895
50	530	710	1070
100	880	1060	1420

表3 計算時間の比較 (単位は  $10^{-3}$  秒)  
Table 3 Comparison of calculation time.

Frames\Points	60	120	240	
25	LM	67.7	122.5	278.4
	CG	27.3	49.3	146.6
	BDCG	9.0	17.2	34.8
50	LM	132.1	491.2	877.1
	CG	55.3	101.6	242.3
	BDCG	18.6	35.1	69.1
100	LM	270.9	860.8	7533.9
	CG	138.2	208.8	499.4
	BDCG	38.8	72.0	141.5

150画素程度であり、観測された特徴点位置には標準偏差  $\sqrt{2}$  画素の正規乱数の誤差を与えた。

ゼロ初期値を用いて形状復元した結果は、いずれの手法も式(9)の残差は小数点以下3桁まで一致し、形状の3次元誤差はすべての例で球の半径の1/110以下であった。このことから、3手法とも同じ精度の解が得られることが分かる。

形状復元に要した3手法の計算時間を表3に示す。いずれのデータセットにおいてもLM, CG, BDCGの順に計算時間が減少し、3手法の中ではBDCG法が最も高速である。 $F, P$ の増加に対する計算時間の変化をグラフにしたものを図4に示す。 $F, P$ がそれぞれ2倍, 4倍に増えたとき、BDCG法, CG法の計算時間がほぼ線形に増加するのに対して、LM法は急速に増大する傾向がある。これは、CG法とBDCG法の計算量が $FP$ に比例するのに対して、LM法では $F^3, P^3$ に比例するためである。なお、CG法とBDCG法の計算時間にはほぼ定数倍の差があり、この実験では3倍程度であった。

一方、3手法の反復回数を比べたものを表4に示す。LM法はいずれのデータセットに対しても反復計算は

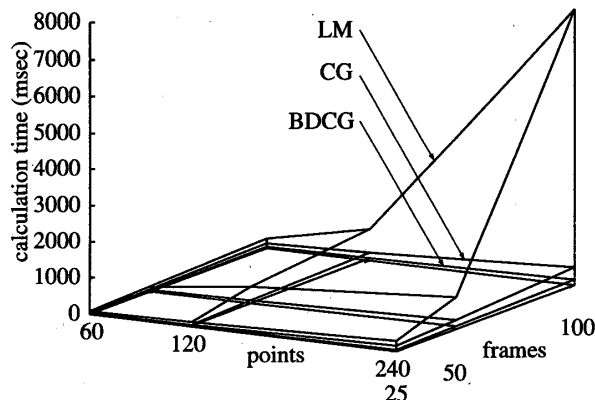


図4 LM法, CG法, BDCG法の計算時間の比較  
Fig. 4 Comparison of calculation time between LM, CG and BDCG methods.

表4 反復回数の比較  
Table 4 Comparison of iteration.

Frames\Points	60	120	240	
25	LM	5	5	5
	CG	105	88	124
	BDCG	28	28	29
50	LM	5	5	5
	CG	108	101	113
	BDCG	29	29	29
100	LM	5	5	5
	CG	139	107	127
	BDCG	32	31	31

5回と収束効率が最も良い。これに対して、CG法は120回程度、BDCG法は30回程度の反復計算を行っている。しかし、上述のように反復計算1回あたりの計算量はLM法が多く、CG法とBDCG法は少ないため、全体の計算時間ではCG法やBDCG法の方が有利である。しかも、LM法とBDCG法は $F, P$ の増加に対して反復回数がほとんど変化しないため、問題の規模が大きくなるほどBDCG法が有利となる。

#### 4.3 実画像を用いた実験

実画像からの形状復元に3手法を適用した結果を示す。対象物体として車の模型(65×70×155mm)を用い、カメラの焦点距離と物体との距離を変化させて射影歪みの異なる3種類の画像列を撮影した。図5に各画像列の第1フレームを示す。画像サイズは640×480であり、各画像列からKLT法<sup>25)</sup>を用いて物体上の特徴点を検出・追跡した。各画像列の性質を表す各種パラメータを表5に示す。

ゼロ初期値を用いて形状復元したときの3手法の計算時間、反復回数、残差(RMS)の比較を表6に示す。3手法とも残差が一致していることから、復元形状に精度の差がないことが分かる。3手法の計算時間と反復回数の違いは、シミュレーション実験と同じ傾

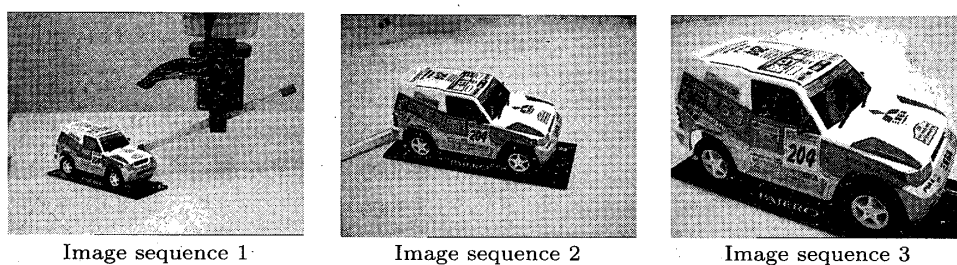


図5 画像列1, 2, 3の第1フレーム

Fig. 5 The first frames of sequence 1, 2 and 3.

表5 画像列1, 2, 3の性質

Table 5 Characteristic parameters of image sequences.

	画像列1	画像列2	画像列3
特徴点数 × 画像数	90×81	187×19	269×16
未知数の個数	837	694	919
対象物までの距離 (mm)	1,100	300	150
カメラの焦点距離 (pixel)	1,800	870	870
物体の回転量 (度)	33	24	20
画像中の対象物の画角 (度)	8	25	40

表6 ゼロ初期値による3手法の性能比較 (計算時間の単位は  $10^{-3}$ 秒)

Table 6 Comparison of calculation time, number of iteration and RMS errors between LM, CG and BDCG methods initialized by general value.

	画像列1			画像列2			画像列3		
	計算時間	反復回数	残差	計算時間	反復回数	残差	計算時間	反復回数	残差
LM	915	9	1.132	209	7	1.190	394	11	1.336
CG	378	157	1.132	193	159	1.190	318	226	1.336
BDCG	104	45	1.132	53	45	1.190	81	58	1.336

表7 因子分解法を初期値とした3手法の性能比較 (計算時間の単位は  $10^{-3}$ 秒)

Table 7 Comparison of calculation time, number of iteration and RMS errors between LM, CG and BDCG methods initialized by factorization.

	画像列1			画像列2			画像列3		
	計算時間	反復回数	残差	計算時間	反復回数	残差	計算時間	反復回数	残差
因子分解法	1.9	—	2.36	0.95	—	6.90	1.2	—	22.7
LM	1028	6	1.132	531	10	1.190	326	5	1.336
CG	129	48	1.132	195	140	1.191	444	263	1.337
BDCG	56	18	1.132	37	26	1.190	48	27	1.336

向を示し、計算時間はLM法 > CG法 > BDCG法の順に短く、反復回数はCG法 > BDCG法 > LM法の順に少ない。いずれの画像列もBDCG法の計算が最も速く、最大でCG法の3.9倍、LM法の8.8倍の高速化が実現できている。

なお、この実験では特徴点の隠蔽が生じない画像列を用いたため、因子分解法を適用することができる。そこで、因子分解法で求めた形状(近似解)を初期値として非線形最適化の3手法に適用した結果を表7に示す。因子分解法の計算時間は格段に短いが、残差を比べると射影歪みが強くなるほど線形近似による誤差が大きくなるのに対して、非線形解法ではほとんど精度が低下しないことが分かる。3手法の計算時間と反

復回数を比べると表6と同様の傾向を示しているが、LM法やCG法では近似解を初期値とすると、計算時間や反復回数が増え、性能が低下する現象が観測された(画像列2のLM法、画像列3のCG法)。これに対してBDCG法では、いずれのデータに対しても近似解を初期値とすることによる性能向上が見られ、前処理行列による高速化と探索方向の補正の効果が安定に機能することを確認した。この実験においても、BDCG法の計算時間が最も短く、最大でCG法の9.5倍、LM法の18.4倍の高速化が実現できている。

## 5. まとめ

本研究では、画像列からの3次元形状復元問題を対

象に、非線形最適化計算の高速アルゴリズムとして「ブロック対角行列による前処理つき共役勾配法 (BDCG法)」を提案しその評価を行った。

この手法の特徴は、ヘッセ行列をブロック対角行列で近似することによって逆行列の計算コストを大幅に低減し、近似による解の探索方向の誤差を共役方向の生成によって補正することで、全体として計算速度の向上を実現した点である。シミュレーションおよび実画像の12種類のデータセットを用いて実験を行った結果、いずれのデータに対しても従来のLevenberg-Marquardt法および共役勾配法を上回る性能が得られた。また提案手法では、近似解を初期値とすることによって反復回数も計算時間も改善され、従来手法より安定した性質が見られた。

コンピュータビジョンの多くの問題は、不完全で不正確な部分情報から全体情報を求めるものであり、モデル(全体)と観測データ(部分)の2乗誤差の最小化という定式化が有効である。モデルや定式化の制約が少ない非線形最適化手法が強力な解法であることは知られていたが、安定化(適切な初期値と局所解回避)と高速化(計算量の削減)が実用上大きな問題となっていた。本論文で提案したBDCG法は、画像数×特徴点数<10000程度の形状復元問題では0.1秒以内で計算が終了し、実用上十分な高速化の性能が得られた。

最近では、より高精度で高密度な3次元復元を目指して、動画像や高解像度画像が用いられるようになり、画像数と特徴点数がそれぞれ数百から数千の大規模データが扱われるようになってきている。本手法は、問題の規模が大きくなるほど従来手法より有利になるので、大規模問題への非線形最適化手法の適用に貢献することが期待される。なお、初期値設定と局所解回避については本論文で触れることができなかった。この問題については稿を改めて述べたい。

謝辞 本研究の一部は広島市立大学特定研究費(No.3104)の援助を受けて行った。

### 参考文献

- 1) Longuet-Higgins, H.C.: A computer algorithm for reconstructing a scene from two projections, *Nature*, Vol.293, pp.133-135 (1981).
- 2) Hartley, R.I.: In Defence of the 8-point Algorithm, *Proc. ICCV*, pp.1064-1070 (1995).
- 3) Weng, J., Ahuja, N. and Huang, T.S.: Optimal motion and structure estimation, *IEEE Trans. PAMI*, Vol.15, No.9, pp.864-884 (1993).
- 4) Szeliski, R. and Kang, S.B.: Shape Ambigu-

- ities in Structure from Motion, *IEEE Trans. PAMI*, Vol.19, No.5, pp.506-512 (1997).
- 5) Tomasi, C. and Kanade, T.: Shape and Motion from Image Streams under Orthography: a Factorization Method, *Int. J. of Computer Vision*, Vol.9, No.2, pp.137-154 (1992).
- 6) Sturm, P. and Triggs, B.: A factorization based algorithm for multi-image projective structure and motion, *Proc. ECCV*, pp.B-709-720 (1996).
- 7) 金谷健一: 因子分解を用いない因子分解法: 平行投影から透視変換へ, 電子情報通信学会PRMU 98-26, pp.1-8 (1998).
- 8) 植芝俊夫, 富田文明: 奥行きパラメータの逐次推定による多視点透視投影画像のための因子分解法, 電子情報通信学会論文誌, Vol.J81-D-II, No.8, pp.1718-1726 (1998).
- 9) Morris, D.D. and Kanade, T.: A Unified Factorization Algorithm for Points, Line Segments and Planes with Uncertainty Models, 画像の認識・理解シンポジウム論文集, Vol.I, pp.401-406 (1998).
- 10) Jacobs, D.: Linear fitting with missing data: Application to structure-from-motion and to characterizing intensity images, *Proc. CVPR*, pp.206-212 (1997).
- 11) Szeliski, R. and Kang, S.B.: Recovering 3D shape and motion from image streams using nonlinear least squares, *J. Visual Communication and Image Representation*, Vol.5, No.1, pp.10-28 (1994).
- 12) Kumar, E.V.E., Tirumalai, A. and Jain, C.: A Non-linear Optimization Algorithm for the Estimation of Structure and Motion Parameters, *Proc. CVPR*, pp.136-153 (1989).
- 13) McLauchlan, P.F.: A Batch/Recursive Algorithm for 3D Scene Reconstruction, *Proc. CVPR*, pp.738-743 (2000).
- 14) Triggs, B., McLauchlan, P.F., Hartley, R. and Fitzgibbon, A.W.: Bundle adjustment — A modern synthesis, *Lecture Notes in Computer Science*, No.1883, pp.298-375 (2000).
- 15) Chiuso, A., Brockett, R. and Soatto, S.: Optimal Structure from Motion: Local Ambiguities and Global Estimates, *Int. J. of Computer Vision*, Vol.39, No.3, pp.195-228 (2000).
- 16) Shum, H.-Y., Ikeuchi, K. and Reddy, R.: Principal Component Analysis with Missing Data and Its Application to Polyhedral Object Modeling, *IEEE Trans. PAMI*, Vol.17, No.9, pp.854-867 (1995).
- 17) Jebara, T., Azarbayejani, A. and Pentland, A.: 3D Structure from 2D Motion, *IEEE Signal Processing Magazine*, pp.66-84 (1999).



- 18) 今野 浩, 山下 浩: 非線形計画法, 日科技連 (1978).
- 19) 右田剛史, 天野 晃, 浅田尚紀: 非線形最適化計算を用いた運動からの形状復元, 画像の認識・理解シンポジウム論文集, Vol.II, pp.199-204 (2000).
- 20) 名取 亮: 線形計算, 朝倉書店 (1993).
- 21) Oliensis, J.: A Multi-Frame Structure-from-Motion Algorithm under Perspective Projection, *Int. J. of Computer Vision*, Vol.34, No.2, pp.163-192 (1999).
- 22) 徐 剛, 辻 三朗: 3次元ビジョン, 共立出版 (1998).
- 23) Xu, G. and Zhang, Z.: *Epipolar Geometry in Stereo, Motion and Object Recognition, A Unified Approach*, Kluwer Academic Publishers (1996).
- 24) Kanatani, K.: Gauge-Based Reliability Analysis of 3-D Reconstruction from Two Uncalibrated Perspective Views, *Proc. ICPR*, Vol.1, pp.76-79 (2000).
- 25) Tomasi, C. and Kanade, T.: Detection and Tracking of Point Features, CMU-CS-91-132 (1991).

## 付 録

式 (12) 中の  $\arg \min E(\mathbf{w}_k - \alpha \mathbf{d}_k)$  は, 目的関数  $E(\alpha)$  を 2 次関数  $G(\alpha)$  で近似し,  $G'(\alpha) = 0$  となる  $\alpha$  を求める. 近似の条件は,  $E(\alpha)$  と  $G(\alpha)$  が 3 点で一致すること ( $E(\alpha) = G(\alpha), \alpha \in \{0, \delta, 2\delta\}$ ) であり,  $\delta$  は条件 ( $G(\delta) < G(0)$  かつ  $G(\delta) < G(2\delta)$ ) を満たすように選ぶ. まず, 適当な初期値 ( $\delta = 1$  や前回の  $\delta$  の値など) を選び, 上記の条件を満たすまで  $\delta$  を 2 倍あるいは 1/2 倍に変更する. この手順は, 上記の条件が緩やかであるため, 黄金分割法よりも少ない反復回数で  $G(\alpha)$  が定まり  $\alpha$  が得られる.

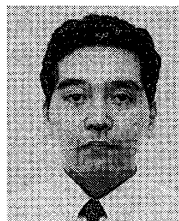
(平成 15 年 1 月 17 日受付)

(平成 15 年 9 月 5 日採録)



右田 剛史 (学生会員)

1998 年広島市立大学情報科学部知能情報システム工学科卒業. 2000 年同大学院修士課程修了. 同大学院情報科学研究科博士後期課程に進学. 2003 年 6 月岡山大学工学部情報工学科助手. コンピュータビジョン, 数値計算の高速アルゴリズムに興味を持つ. 電子情報通信学会会員.



天野 晃 (正会員)

1988 年京都大学工学部電気工学科卒業. 1990 年同大学院修士課程修了. 1993 年同大学院博士課程学位修得退学. 同年京都大学工学部助手. 1995 年広島市立大学助教授. 2002 年京都大学大学院情報学研究科助教授. コンピュータビジョン, 文書画像処理の研究に従事. IEEE CS, 電子情報通信学会, 人工知能学会, ME 学会各会員.



浅田 尚紀 (正会員)

1979 年京都大学工学部電気工学科卒業. 同大学院博士課程修了. 福井医科大学助手, 京都大学工学部助手, 岡山大学工学部助教授を経て 1995 年から広島市立大学情報科学部教授, 現在に至る. 1989 年シカゴ大学客員研究員. 工学博士. コンピュータビジョン, コンピュータグラフィクス, 人工知能, 医用画像理解の研究に従事. 1993 年情報処理学会論文賞. IEEE CS, ACM, 電子情報通信学会, 人工知能学会, 医用画像工学会各会員.