

ロバスト推定を用いた一般アフィン射影モデルの因子分解法

蔵田 武志 藤木 淳 坂上 勝彦

電子技術総合研究所, 〒 305-8568 つくば市梅園 1-1-4

e-mail: {kurata, fujiki, sakaue}@etl.go.jp

あらまし 時系列画像から形状と運動を復元する因子分解法は、数値計算上安定であり良い復元結果を与えるが、通常、最小自乗誤差基準を用いるため外れ値の影響を大きく受けてしまう。そこで本稿では、ロバスト推定手法の一種である LMedS 推定を用いた一般アフィン射影モデルの因子分解法を提案する。まず、因子分解法の第 1 段階である動きと形状の分離において、LMedS 推定を用いて特徴点を選択することにより、外れ値の影響を抑制する。次に、この手法を実時間で適用するために筆者らが提案した逐次型因子分解法へ拡張する。シミュレーションと実画像を用いた実験により本手法の有効性を示す。

キーワード shape from motion, 一般アフィン射影モデル, 逐次型因子分解法, ロバスト推定, LMedS 推定, 実時間処理

A Factorization Method for Generalized Affine Projection Model Using Robust Estimation

Takeshi Kurata Jun Fujiki Katsuhiko Sakaue

Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba-shi, Ibaraki 305-8568, JAPAN

e-mail: {kurata, fujiki, sakaue}@etl.go.jp

Abstract The factorization method for the structure-from-motion is stable in calculation and gives an accurate reconstruction from a tracking result which contains only a single object, however, it is affected by the outliers such as false tracking and other objects because of using the least squares criterion. Then we present a robust factorization method for GAP (generalized affine projection) model. Using the LMedS (Least Median of Squares) criterion, we can discard feature points which is regarded as outliers, then a measurement matrix is stably decomposed into the motion and shape matrices in the first stage of the factorization method. We revise this method into the iterative factorization method for real time processing. Some experiments using synthetic data and a real sequence are given to show the performance of our proposed robust factorization method.

Key words shape from motion, generalized affine projection model, iterative factorization method, robust statistics, LMedS estimation, real-time processing

1 はじめに

時系列画像からカメラと対象物の相対運動及び対象物の3次元形状を同時に復元する structure from motion 問題は、コンピュータビジョンやその周辺領域において、基本的かつ重要な問題である。

この問題を解くために、Tomasi and Kanade[15]により正射影モデルにおいて提案され、Poelman and Kanade[12]により scaled orthographic 射影モデル、paraperspective 射影モデルの各アフィン射影モデルに拡張された因子分解法は、数値計算的に安定で、全画像を均等に扱え、簡潔に定式化が行えるという利点を持っている。最近ではこの枠組を、透視射影モデルに拡張する手法も試みられている [16]。

これらの手法は単一物体を対象としており、実世界へ適用するには必然的に制約が多くなる。そこで、形状行列の性質を利用して、複数物体を分離するための手法が Costeira and Kanade により提案されている [1]。ただし、各物体の持つ運動に制約があり [6]、また、観測誤差や外れ値に弱いという報告もされている [5]。

Zhang[17]では、代表的なロバスト推定手法である M 推定 (M-estimation) [4] と LMedS 推定 (Least Median of Squares method) [14] を用いたエピポラ幾何の復元手法などについて述べられている。この中では、LMedS 推定には、明らかに間違っただけ対応付けられた点を外れ値として除去する効果が強く、M 推定には大きな観測誤差としての外れ値の影響を抑制する効果が高いという特性を用いてエピポラ幾何復元の安定化が図られている。

また、因子分解法は基本的にバッチ処理のため実時間応用には向かない。そこで、過去の復元結果を圧縮しながら復元解を逐次更新する因子分解法が提案されている [9][3]。

本稿では、特に外れ値と実時間処理に焦点をあて、それらを考慮した因子分解法について検討する。まず、因子分解法の第1段階である動きと形状の分離において、LMedS 推定を用いて特徴点を選択することにより、外れ値の影響を抑制する。次に、この手法を実時間で適用するために筆者らが提案した逐次型因子分解法へ拡張する [3]。

以下に本稿の構成を示す。2節で一般アフィン射影モデルの因子分解法について概説し、3節で LMedS 推定を用いた因子分解法について述べる。4節では、その逐次型因子分解法への拡張について述べ、5節でシミュレーションと実画像による提案手法の評価実験を行う。

2 一般アフィン射影モデルの因子分解法

まず、一般アフィン射影モデルの因子分解法について簡単に説明する [2]。なお、本稿では簡単のために、光軸と画像平面は直交し、それらの交点は既知とし、アスペクト比は 1:1 とする。

2.1 計測行列

第 f 画像上の第 p 特徴点の座標を (x_{fp}, y_{fp}) ($f = 1, \dots, F$) とする。

$$W = \begin{pmatrix} W_1 \\ \vdots \\ W_F \end{pmatrix}_{(2F \times P)}, \quad W_f = \begin{pmatrix} x_{f1} & \cdots & x_{fP} \\ y_{f1} & \cdots & y_{fP} \end{pmatrix}$$

で定義される行列 W を計測行列と呼び、それらの座標原点を特徴点の重心 (x_f, y_f) に移動した行列

$$W' = \begin{pmatrix} W'_1 \\ \vdots \\ W'_F \end{pmatrix}, \quad W'_f = \begin{pmatrix} x'_{f1} & \cdots & x'_{fP} \\ y'_{f1} & \cdots & y'_{fP} \end{pmatrix}$$

を修正計測行列と呼ぶ。

2.2 一般アフィン射影モデル

3次元空間中の第 p 特徴点 s_p ($p = 1, \dots, P$) と、 (x_{fp}, y_{fp}) の関係を一般アフィン射影モデルで表すと、

$$\begin{pmatrix} x_{fp} \\ y_{fp} \end{pmatrix} = \underset{(2 \times 3)}{A_f} \underset{(3 \times 3)}{C_f} s_p + \mathbf{f}_f$$

となる。ここで、 $C_f = (\mathbf{i}_f, \mathbf{j}_f, \mathbf{k}_f)^T$ は画像面 Π_f 上の正規直交基底 $\{\mathbf{i}_f^T, \mathbf{j}_f^T\}$ とカメラの光軸方向の単位ベクトル \mathbf{k}_f^T を行ベクトルとするカメラ座標基底行列、 A_f はアフィンモデル行列で具体的には以下のように記述される。なお、 l は焦点距離、 z_f は物体重心とカメラの相対奥行きである。

正射影モデル

$$A_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Scaled Orthographic 射影モデル

$$A_f = \frac{l}{z_f} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Paraperspective 射影モデル

$$A_f = \frac{1}{z_f} \begin{pmatrix} l & 0 & -x_f \\ 0 & l & -y_f \end{pmatrix}$$

P 個の s_p の重心 \mathbf{g} は、 (x_f, y_f) に射影されるので、第 p 特徴点の \mathbf{g} からの相対座標を s'_p とすると、

$$\begin{pmatrix} x'_{fp} \\ y'_{fp} \end{pmatrix} = A_f C_f s'_p$$

となる。ここで $S' = (s'_1, \dots, s'_P)$ とすると、

$$W'_f = A_f C_f S' = M_f S', \quad M_f = A_f C_f = \begin{pmatrix} \mathbf{m}_f^T \\ \mathbf{n}_f^T \end{pmatrix}$$

$$W' = \begin{pmatrix} A_1 C_1 \\ \vdots \\ A_F C_F \end{pmatrix} S' = \begin{pmatrix} M_1 \\ \vdots \\ M_F \end{pmatrix} S'$$

が得られるので、以下のように W' をカメラ運動を表す行列 M と特徴点の3次元情報を表す行列 S' の積に因子分解できることがわかる。

$$W' = \underset{(2F \times 3)}{M} \underset{(3 \times P)}{S'}, \quad M = \begin{pmatrix} M_1 \\ \vdots \\ M_F \end{pmatrix}$$

2.3 計量拘束

W' からユークリッド計量による動きと形状を復元するのに用いることのできる拘束条件は、カメラ座標系の基底が正規直交基底をなす、つまり、 $C_f C_f^T = I_3$ ということなので、

$$M_f M_f^T = A_f A_f^T \quad (1)$$

が成立する。

ところが、一度の操作でこの条件を満たすように W' を分解するのは難しいので、通常、2段階にわけて分解する。まず第1段階では、 W' が rank3 であることを利用し、特異値分解などを用いて、

$$W' = \begin{matrix} \hat{M} & \hat{S}' \\ (2F \times P) & (2F \times 3) (3 \times P) \end{matrix}$$

のように分解する。ここで、任意の3次の正則行列を A とすると、 $W' = \hat{M} A A^{-1} \hat{S}'$ のような分解も可能である。いいかえれば、この分解の時点ですでにアフィン変換の自由度を持つ運動と形状が分離できていることになる。

第2段階では、拘束条件(1)を用いて A を求め、ユークリッド計量による復元解を一意*に決定する。ここで、

$$\hat{M} = \begin{pmatrix} \hat{M}_1 \\ \vdots \\ \hat{M}_f \end{pmatrix}, \quad \hat{M}_f = \begin{pmatrix} \hat{m}_f^T \\ \hat{n}_f^T \end{pmatrix}$$

とし、 $Q = A A^T$ とすると(1)から、

$$\hat{M}_f Q \hat{M}_f^T = A_f A_f^T$$

という拘束条件を得る。この計量行列 Q は、本質的に異なる3枚以上の画像があれば決定できる。

3 LMedS 推定を用いた因子分解法

3.1 特徴点(計測行列の列)の選択

3次元復元(キャリブレーション)を自動化するためには、通常、特徴点を追跡処理により対応付けする必要がある[†]。しかし、追跡結果に様々な要因による誤差や外れ値が混入することは避けられず、その結果を修正せずに通常の最小自乗法による因子分解法を適用するのは困難である。

本稿では、因子分解法の第1段階である動きと形状の分離において、LMedS 推定を用いて特徴点を選択することにより、外れ値の影響を抑制する。LMedS 推定は後述のとおり、入力データのランダムサンプリングによる試行を繰り返し行う。因子分解法の場合は W の行と列のどちらかまたは両方の選択による試行が繰り返されることになる。行と列を同時に選択する場合はその組み合わせ数が増大し計算コストが高くなるため、特に実時間応用を考えた場合には現実的ではない。そこで本稿では列(特徴点)の選択による LMedS 推定を行う。

3.2 LMedS 基準によるアフィン運動行列の復元

まず、LMedS 推定のアルゴリズムについて簡単に説明する。未知パラメータを推定するために必要な数の

*ここでは鏡映の議論はしない

[†]線特徴を用いた復元手法については [13][10] 参照

観測データをランダムサンプリングにより選び、そのデータ集合で推定されるパラメータにおける全データの自乗誤差のメディアンを求める。この試行を十分な回数繰り返し、各試行中で最も小さい自乗誤差のメディアンを取るパラメータを選択する。これにより得られる解の breakdown point (外れ値に対するロバストさ)は0.5である。これは、全データの50%までの外れ値に対応できることを表している。

一般アフィン射影モデルの場合、動きと形状の分離には最低4点あればよい。そこで、試行 $j (= 1, \dots, J)$ では特徴点の観測値の時系列である $2F$ 次元列ベクトル $(x_{1p}, y_{1p}, x_{2p}, y_{2p}, \dots, x_{Fp}, y_{Fp})^T$ を P 個の中から4個、ランダムサンプリングにより選び、列部分計測行列 $W^{[j]}$ を生成する。 $W^{[j]}$ を SVD により $U^{[j]} \Sigma^{[j]} V^{[j]T}$ のように分解すると、 $U^{[j]}$ をそのサンプル集合における試行推定運動行列とみなすことができる。

なお、試行推定運動行列 $U^{[j]}$ に平行移動成分が含まれる場合はランク4、含まれない場合はランク3である必要があるため、まったくランダムに4点をサンプルするのは有効ではない。そこで、 $W^{[j]}$ の修正計測行列 $W'^{[j]}$ のランク3らしさにより、その $W^{[j]}$ による試行を行うかどうか判定することにする。

次に、この試行推定値 $U^{[j]}$ を用いた場合の全データの誤差を

$$dW^{[j]} = \begin{pmatrix} I & -U^{[j]} & U^{[j]T} \\ (2F \times 2F) & (2F \times 4) & (4 \times 2F) \end{pmatrix} W \quad (2F \times P)$$

により求め、各特徴点の時系列自乗誤差 $r_p^{[j]2}$ 、そのメディアン $\mu^{[j]}$ を計算する。

$$r_p^{[j]2} = \sum_f dW_{fp}^{[j]2}$$

$$\mu^{[j]} = \text{med}_p(r_p^{[j]2})$$

この試行を充分繰り返し、最小メディアン値を持つ試行での推定運動行列 $U^{[L]}$ を LMedS 基準により得られる準最適解とする。

必要な試行の回数は次のように求めることができる。 ϵ を任意に1点選ぶ時に外れ値を選ぶ確率、 n を未知パラメータの推定のために必要なサンプルデータの数、 J を試行の回数とすると、 n 個のサンプルデータ全てが外れ値でない試行が少なくとも1回起こる確率 R は、

$$R = 1 - \{1 - (1 - \epsilon)^n\}^J$$

のように計算できる。先に述べたように一般アフィン射影モデルの場合、 $n = 4$ で形状と動きの分離が可能であり、例えば、 $R \geq 0.999$ とするためには、 $\epsilon = 0.4$ の場合で $J \geq 50$ 、 $\epsilon = 0.5$ の場合で $J \geq 108$ であればよいことになる。

3.3 外れ値として振舞う特徴点の除去

1節でも触れたように LMedS 推定により得られるパラメータは、明らかに間違っ対応付けられた点や対象物体以外の物体に属している点としての外れ値には有効であるが、比較的大きな対応のずれや、ガウシアンノイズに対してはあまり有効ではないという特性を持つ。そこで、LMedS 推定により得られる結果から外れ値とみなされる特徴点を除去し他のすべての点を用いて最小自乗解を求める。

LMedS 基準により第 L 番目の試行が選ばれたとすると、

$$\hat{\sigma} = 1.4826\{1.0 + 5.0/(P - n)\}\sqrt{\mu_p^{[L]}}$$

により、外れ値以外のデータの誤差が正規分布に従う場合の推定標準偏差値を求めることができる。この値をもとに本稿では以下の基準を満たす特徴点（計測行列の第 p 列）を全て選択する。

$$r_p^{[L]^2} \leq (2.5\hat{\sigma})^2$$

そして、選択された特徴点のみから成る計測行列に通常の因子分解法を適用し、その運動行列と形状行列を求める。なお、LMedS 推定で外れ値を除去した後、M 推定などにより復元解を求めてもよい。

4 特徴点選択による逐次型ロバスト因子分解法

因子分解法は通常、すべての画像が入力されてから計算をはじめめるバッチ処理の形を取るため実時間応用には向かない。そこで推定値の逐次更新型因子分解法が提案されている [9][3]。本稿では、前節で提案した特徴点選択による因子分解法を、筆者らの提案した逐次型因子分解法 [3] に拡張する。

4.1 逐次型因子分解法の概説

4.1.1 初期段階：最初の $k (\geq 3)$ 画像の扱い

カメラ運動と形状を決定する計量を求めるためには、本質的に異なる 3 枚以上の画像が必要である。そこでまず初めの k 画像から運動行列 $M_{[k]}$ と形状行列 $S'_{[k]}$ を求める。この初期段階の世界座標系を基準世界座標系とする。 $M_{[k]}$ の SVD を $F_{[k]}\Lambda_{[k]}E_{[k]}$ としたとき、第 k 主成分運動行列 $\mathcal{M}_{[k]}$ 、第 k 主成分修正計測行列 $\mathcal{W}'_{[k]}$ を以下のようにして求める。

$$\begin{aligned}\mathcal{M}_{[k]} &= \Lambda_{[k]}E_{[k]} = F_{[k]}^T M_{[k]} \\ \mathcal{W}'_{[k]} &= \mathcal{M}_{[k]}S'_{[k]} = F_{[k]}^T \mathcal{W}'_{[k]}\end{aligned}$$

4.1.2 更新段階

第 f 逐次修正計測行列 $W'_{[f]}$ ($f > k$) を

$$W'_{[f]} = (\mathcal{W}'_{[f-1]}^T, W_f^T)^T$$

(5×P)

と定め、 $W'_{[f]} = \hat{M}_{[f]}\hat{S}'_{[f]}$ のように一時的に分解する。ここで、

$$\hat{M}_{[f]} = (\hat{\mathcal{M}}_{[f]}^T, \hat{M}_f^T)^T$$

である。この $\hat{M}_{[f]}$ についての以下の拘束条件からユークリッド計量を計算し、基準世界座標に固定されていない運動行列 $\bar{M}_{[f]}$ と形状行列 $\bar{S}'_{[f]}$ を求める。

$$\begin{aligned}\hat{M}_{[f]}Q_{[f]}\hat{M}_{[f]}^T &= \Lambda_{[f-1]}^2 \\ \hat{M}_fQ_{[f]}\hat{M}_f^T &= A_fA_f^T\end{aligned}\quad (2)$$

次に、 $\bar{S}'_{[f]}$ を $S'_{[f-1]}$ に変換する直交行列 \mathcal{E}_f を求め、 $\bar{M}_f, \bar{S}'_{[f]}$ を基準世界座標系に固定する。

$$\begin{aligned}M_f &= \bar{M}_f\mathcal{E}_f^T \\ S'_{[f]} &= \mathcal{E}_f\bar{S}'_{[f]}\end{aligned}\quad (3)$$

$M_{[f]}$ の SVD を $F_{[f]}\Lambda_{[f]}E_{[f]}$ とすると、第 f 主成分運動行列 $\mathcal{M}_{[f]}$ および第 f 主成分修正計測行列 $\mathcal{W}'_{[f]}$ は以下のようにして求められる。

$$\begin{aligned}\mathcal{M}_{[f]} &= F_{[f]}^T M_{[f]} \\ \mathcal{W}'_{[f]} &= \mathcal{M}_{[f]}S'_{[f]} = F_{[f]}^T \mathcal{W}'_{[f]}\end{aligned}\quad (4)$$

4.2 特徴点選択処理の導入

3 節の手法と同様、LMedS 推定により外れ値とみなされる特徴点を除去した後、最小自乗解を得るという処理を初期段階と更新段階のそれぞれに導入する。

4.2.1 初期段階

初期段階では、 k フレームまでの計測行列 W を用いて 3 節の手法による特徴点を選択し、初期値となる運動行列 $M_{[k]}$ と形状行列 $S'_{[k]}$ を求め、その主成分運動行列 $\mathcal{M}_{[k]}$ 、主成分修正計測行列 $\mathcal{W}'_{[k]}$ を得る。 k の決定方法については 5 節で述べる。

4.2.2 更新段階

更新段階では、逐次計測行列 $W_{[f]}$ から同様の手法で特徴点選択を行い、選択された特徴点から成る計測行列 $W_{\text{sub}[f]}$ の一時的分解によって、 $\hat{M}_{[f]}$ および $\hat{S}'_{\text{sub}[f]}$ を得る。

計量の拘束条件は特徴点数に依存しないのでやはり (2) によりユークリッド計量を計算し、 $\bar{M}_{[f]}$ と $\bar{S}'_{\text{sub}[f]}$ を求めることができる。

$\bar{M}_{[f]}$ と $\bar{S}'_{\text{sub}[f]}$ の基準世界座標系への固定は、 $\bar{S}'_{\text{sub}[f]}$ を $S'_{\text{sub}[f-1]}$ に変換する直交行列 \mathcal{E}_f を求めることで、(3) と同様に計算することができる。

$$\begin{aligned}M_f &\leftarrow \bar{M}_f\mathcal{E}_f^T \\ S'_{\text{sub}[f]} &\leftarrow \mathcal{E}_f\bar{S}'_{\text{sub}[f]}\end{aligned}$$

ここで、 $S'_{[f-1]}$ に含まれて $S'_{\text{sub}[f]}$ に含まれない特徴点、つまり第 f フレームで外れ値と判定されたり追跡結果が得られずデータが欠落した特徴点の復元結果については第 $f-1$ フレームの結果を $S'_{[f]}$ で採用することとする。この処理により、LMedS 推定による特徴点選択に誤って外れ値と判定された場合でも次のフレームで回復することが可能となり、また、追跡処理へのフィードバック情報としての応用も考えられる。第 f 主成分運動行列 $\mathcal{M}_{[f]}$ および第 f 主成分修正計測行列 $\mathcal{W}'_{[f]}$ は (4) により得ることができる。

5 実験結果

提案手法である特徴点選択による逐次型ロバスト因子分解法のシミュレーションと実画像による評価実験を行った。

5.1 初期段階 k の決定

初期値の計算に用いるフレーム数 k を決定するために、本稿では以下のような実装を行った。

1. 第 k フレームまでの計測行列を用いて特徴点を選択する。

2. 外れ値を除いた修正計測行列の特異値が以下の基準を満す場合は3へ進み、満さなければ $k \leftarrow k+5$ とし、1へ戻る。

$$\frac{\text{第4特異値}}{\text{第3特異値}} < \alpha$$

3. 計量 Q の第3固有値がある閾値 β より大きければ、初期値を計算して更新段階に移る。そうでなければ、 $k \leftarrow k+5$ とし1へ戻る。

この判定処理は、本質的に異なる3画像が入ってくるまで更新段階に移らないようにするものである。なお本実装では、経験的に $\alpha = 0.2$, $\beta = 0.2$ とした。また、計算コストの削減のために計測行列の特異値ならびに計量の固有値の評価は5フレームおきとし、その評価のための計測行列は等間隔でフレームを間引き、 $10 \times P$ 行列とした。そして、更新段階に移る時に、全てのフレームを用いて初期復元解を求めた。

5.2 シミュレーション

シミュレーションの設定は以下の通りである。

- $\{s_p\}_{p=1}^{20}$ は一辺が 200mm の立方体から生成
- カメラはピンホールカメラを仮定
- 焦点距離は 35mm カメラ相当で 84mm
- 画像サイズは 640×480 で各ピクセルは一辺 8 μ m の正方形
- 画像のフレーム数 $f = 120$
- 第 f カメラ座標基底行列 C_f
 $f = 1, \dots, 60$

$$C_f^T = R_x \left(\frac{4}{3}f \right) R_y(0) R_z \left(\frac{1}{3}f \right)$$

$$f = 61, \dots, 120$$

$$C_f^T = R_x \left(160 - \frac{4}{3}f \right) R_y(0) R_z \left(\frac{1}{3}f \right)$$

- 奥行きは $z_1 = 2000\text{mm}$, $z_{120} = 1600\text{mm}$ となるように等速移動
- $u_1 = 115\text{pixel}$, $u_{60} = 0\text{pixel}$, $u_{120} = 101\text{pixel}$ となるように滑らかに移動

観測誤差と外れ値の設定については、特徴点を3組に分けそれぞれ以下のように設定した。

データセット A [12点]: 第1から第120フレームまでの正しい追跡結果に平均0、標準偏差 1pixel のガウシアンノイズを加えた点

データセット B [4点]: 第1から第60フレームまでの正しい追跡結果に平均0、標準偏差 3pixel のガウシアンノイズを加え、第61フレームからは最大 2pixel/frame でランダムに動く点

データセット C [4点]: 第1から第120フレームまで最大 2pixel/frame でランダムに動く点

以上の設定条件のもとで、形状誤差の推移、平均奥行き復元誤差、運動誤差、計算し時間を評価した。復元には paraperspective 射影モデルを用いた。また、誤差評価の比較対象として、もっとも条件のよいデータセット A の12点のみを入力としたバッチ処理型の因子分解法の復元結果を用いた。なお、3節の LMedS 推定による因子分解法に120フレーム全てを一度に入力した場合、自動的にこの12点を選択することができた。

5.2.1 特徴点選択結果

試行回数 J は100回とした。初期値計算のためのフレーム数決定処理により本実験では $k = 23$ と設定され、第1フレームからランダムに動き回るデータセット C の4点が初期段階での外れ値として除去された。

その後、第60フレームまでは残りの16点のうち1、2点が頻繁に外れ値として判断された。特に標準偏差 3pixel の誤差を持つデータセット B の4点が外れ値として選ばれる傾向が強かったが、条件のよいデータセット A についても外れ値と判断されることがあった。しかし、前節で述べたように直前のフレームでの復元結果を用いることで外れ値と判断された点についてもその後のフレームで復帰し復元結果を更新することができた。

第61フレームからデータセット B の4点は徐々に外れ値として判定され、第69フレーム以降はすべて外れ値として判定された。そのため、第70フレーム以降それらの点に関しては復元結果の更新は行われず、データセット A のみにより復元結果が更新された。

5.2.2 形状復元

形状復元誤差 $e_{\text{shape}}(f)$ は以下の式で評価した。

$$e_{\text{shape}}(f) = \frac{\sum_{p=1}^P \left\| \mathbf{s}_p^{\text{true}} - \frac{z_f^{\text{true}}}{z_f^{\text{estimated}}} \mathbf{s}_{[f],p} \right\|}{\sum_{p=1}^P \left\| \mathbf{s}_p^{\text{true}} \right\|} \times 100(\%).$$

ここで、 $\mathbf{s}_{[f],p}$ は $S_{[f]}$ の第 p 列であり、 S と $S_{[f]}$ の重心は世界座標の原点で一致している。バッチ処理による形状復元誤差は3%であった。提案手法では前述の通り、初期段階でデータセット A, B の16点が選択され復元された。第24フレームで最大8.5%の復元誤差を含んでいるが、最終的に3.3%まで収束した。これにより本手法が安定に収束していくことが確認された(図1)。

5.2.3 奥行き復元 (相対誤差)

奥行きは相対誤差により評価した。

$$e_{\text{depth}}(f) = \frac{z_f^{\text{estimated}} - z_f^{\text{true}}}{z_f^{\text{true}}} \times 100(\%)$$

ここで、 $z_k^{\text{estimated}} = z_k^{\text{true}}$ によって、スケールの自由度を固定した。提案手法による奥行き比の復元誤差についても、形状復元誤差同様、フレーム数が増えるにしたがってバッチ処理の結果に近づいていることがわかる(図2)。

5.2.4 運動復元 (2つのベクトルの成す角)

運動復元誤差は、

$$e_i(f) = \cos^{-1} \left\| (\mathbf{i}_f^{\text{estimated}})^T \mathbf{i}_f^{\text{true}} \right\|. \quad (\text{度})$$

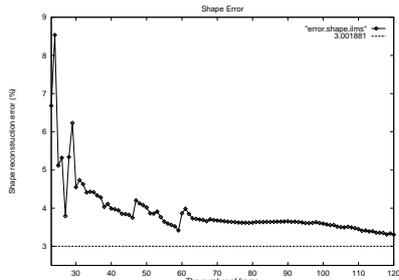


図 1: 形状復元誤差

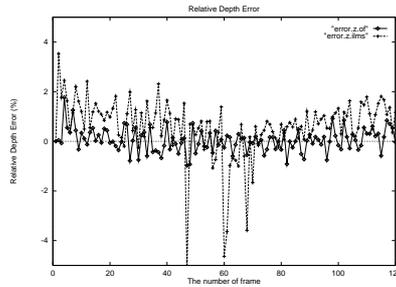


図 2: 平均奥行き復元誤差

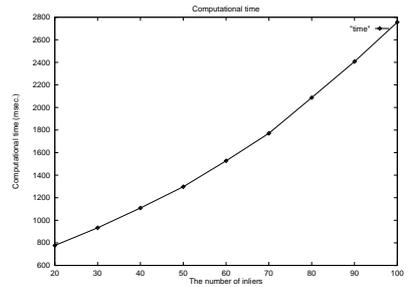


図 3: 特徴点数と計算時間

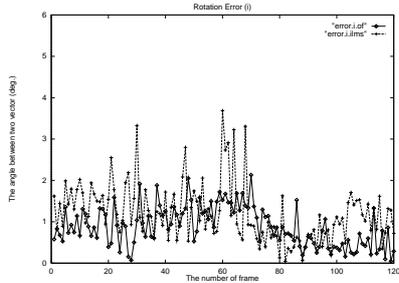


図 4: 運動復元誤差 (i_f)

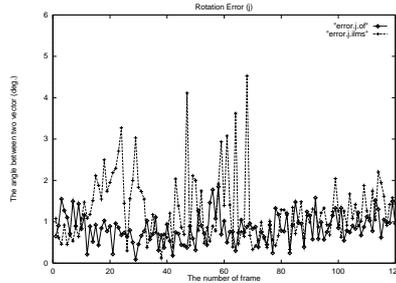


図 5: 運動復元誤差 (j_f)

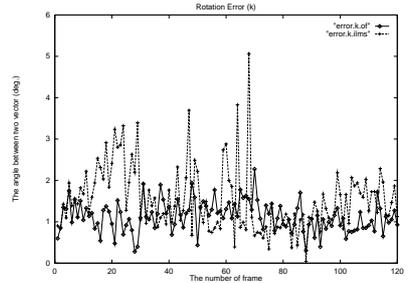
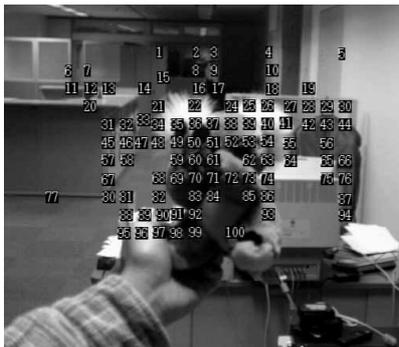
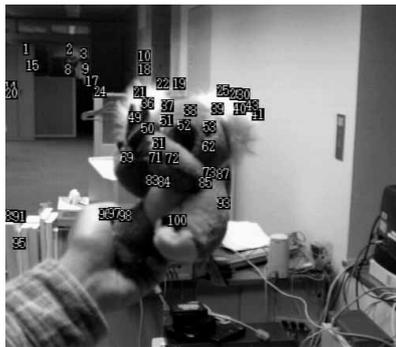


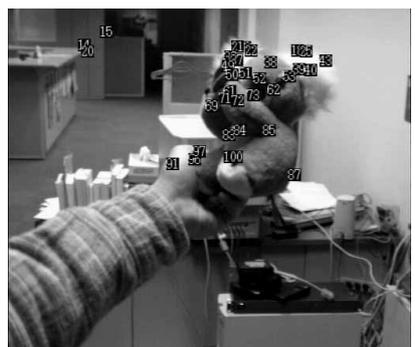
図 6: 運動復元誤差 (k_f)



A: [frame 1: 100 points]



B: [frame 33: 48 points]



C: [frame 99: 35 points]

図 7: 入力時系列画像 (99 フレーム)

により評価した。 j_f , k_f についても同様に評価した。 これらも形状復元誤差や奥行き比の復元誤差同様、フレーム数が増えるにしたがってパッチ処理の結果に近づいている。ただし、 j_f , k_f については第 60 フレーム前後で大きな誤差が見られた (図 4, 5, 6)。

5.2.5 計算時間

提案手法の計算時間を計測した。計算には、Pentium II 450MHz PC を用いた。フレーム数は 120 枚に固定し、特徴点数を 20 点から 100 点まで変化させた。提案手法の計算コストは外れ値を除く特徴点数 (inlier) に依存するため、全ての点が選択されるよう誤差のないデータを用いた。なお、計算時間には特徴点の選択や追跡にかかった時間は含まれていない。100 特徴点の場合で約 2.8 秒の計算時間を要したが、30frame/sec 換算で 4 秒分のデータであるため、すでに実時間処理が可能であることがわかる (図 3)。

5.3 実画像実験

次に 99 フレームの実時系列画像 (図 7) に提案手法を適用した。

特徴点の自動抽出と追跡処理には日立の IP5000 を用いた。現在のところ、15frame/sec で動作している。本稿では特徴点の追加についてまだ扱っていないため、特徴点の自動抽出は第 1 フレームのみで行い、100 点の特徴点が抽出された。

5.1 節の初期値計算のためのフレーム数決定処理によって、第 33 フレームが初期段階と設定され、その際、29 点が選択され (図 8)、19 点が外れ値として除去された (図 10)。

最終的に第 99 フレームでは 23 点が対象物体上の特徴点として選択されていた (図 9)。初期段階で選択された後、外れ値として除去された点および追跡処理に失敗した点の第 33 フレームでの位置を図 11 に示す。その中の 2 点は第 99 フレームまで存在していたが (図 12)、対応付けがかなりずれているため外れ値と判定された。

提案手法の応用の 1 つとして、復元された対象物体

のユークリッド計量, 運動行列を用いた密な3次元形状復元を行った. 具体的には, 蔵田・藤木 [8] の因子分解法の記述を用いたアフィンエピポラ拘束を利用し, 第64, 65, 66 フレームの3画像による仮想3眼ステレオマッチングを行った. つまり, ユークリッド計量は全99画像により求められ, 密な対応付けはこの3画像で得られることになる.

カメラパラメータが未知であるので, 形状復元には scaled orthographic 射影モデルを用いた. いくつかの視点からの復元結果を図13に示す. ステレオマッチングやCGモデル生成に単純な手法を用いていることもあり, よい結果とは言えないが, 提案手法によってユークリッド計量ならびに運動, 形状が実用可能な精度で復元できていることがわかる.

6 むすび

本稿では, LMedS 推定を用いて特徴点を選択し, 外れ値を含むデータに対してロバストに運動と形状を推定することができる因子分解法を提案した. また, 実時間で適用するためにこの手法を逐次型因子分解法へ拡張した.

シミュレーションと実画像を用いた評価実験の結果, 提案手法による特徴点選択が有効であることが確認できた. また, 処理速度についてはPCでも十分に実時間処理できることを示した.

本稿では更新段階の時点で新たに追跡される特徴点については扱っておらず, また, 提案手法の breakdown point についての評価も行っていない. 実世界応用のためには必ず検討すべきである.

また, 動きによる領域分離 [7] や特徴点追跡処理などと提案手法との協調処理による, 複数物体への対応や追跡精度の向上, エピポラ拘束による特徴点の再探索による復元精度向上なども今後の課題である.

謝辞 本研究の機会を与えて下さった知能情報部長大津展之博士, 情報科学部長中島秀之博士, 情報ダイナミクスラボリーダー梅山伸二博士, 早稲田大学理工学部村岡洋一教授に深く感謝致します. また, 様々な御討論, ご協力を頂きました, 適応ビジョンラボ, 情報ダイナミクスラボの皆様, ならびに早稲田大学理工学部興相正克氏に深く感謝致します.

参考文献

- [1] J. Costeira and T. Kanade: A Multi-body Factorization Method for Independently Motion Objects, *IJCV*, Vol.29, No.3, pp.159-179 (1998).
- [2] 藤木, 蔵田: 一般アフィン射影モデルの因子分解法の数理, 信学技法, PRMU97-142, pp.101-108 (1997).
- [3] 藤木, 蔵田, 田中: 一般アフィン射影モデルの逐次型因子分解法, 信学技法, PRMU98-118, pp.45-52 (1998)
- [4] P. J. Huber: *Robust Statistics*, John Wiley & Sons, New York (1981).
- [5] T. Kanade and D. D. Morris: Factorization Methods for Structure from Motion, *Phil. Trans. New Geometric Techniques in Computer Vision*, J. Lasenby, A. Zisserman, R. Cipolla, and H. C. Longuet-Higgins (Eds.), R. Soc. Lond. A, Vol. 356, pp.1153-1173, (1998).
- [6] 金谷: 因子分解法を用いない因子分解法: 複数物体分離, 信学技法, PRMU98-117, pp.37-44 (1998).
- [7] 興相, 村岡: 動画像の大域的/局所的アフィン動きパラメータの実時間推定, 信学技報, PRMU98-91, pp.53-60 (1998)
- [8] T. Kurata, J. Fujiki, and K. Sakaue: Affine Epipolar Geometry via Factorization Method. *Proc. ICPR98*, Vol.1, pp.862-866 (1998).
- [9] T. Morita, and T. Kanade: A Sequential Factorization Method for Recovering Shape and Motion From Image Streams, *IEEE Trans. PAMI*, Vol.19, No.8, pp.858-867 (1997).
- [10] D. D. Morris and T. Kanade: A Unified Factorization Algorithm for Points, Line Segments and Planes with Uncertainty Models. *Proc. ICCV98*, pp.696-702 (1998).
- [11] J. L. Mundy and A. Zisserman: *Geometric Invariance from Motion*. MIT Press, Cambridge, MA. (1992).
- [12] C. J. Poelman and T. Kanade: A Paraperspective Factorization Method for Shape and Motion Recovery, *IEEE Trans. PAMI*, Vol.19, No.3, pp.206-218 (1997).
- [13] L. Quan and T. Kanade: A Factorization Method for Affine Structure from Line Correspondences. *Proc. CVPR96*, pp.803-808 (1996).
- [14] P. J. Rousseeuw and A. M. Leroy: *Robust Regression and Outlier Detection*, John Wiley & Sons, New York (1987).
- [15] C. Tomasi and T. Kanade: Shape and Motion from Image Streams under Orthography: A Factorization Method, *IJCV*, Vol.9, No.2, pp.137-154 (1992).
- [16] T. Ueshiba and F. Tomita: A Factorization Method for Projective and Euclidean Reconstruction from Multiple Perspective Views via Iterative Depth Estimation, *Proc. 5th ECCV*, Vol.1, pp.296-310 (1998).
- [17] Z. Zhang, Determining the Epipolar Geometry and its Uncertainty: A Review, *IJCV*, Vol.27, No.2, pp.161-195 (1998).

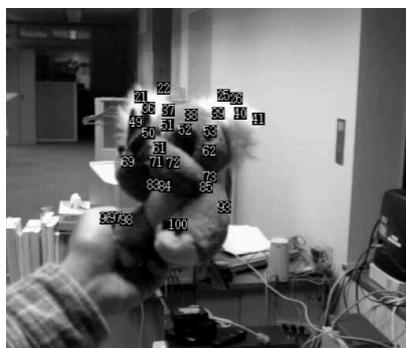


図 8: 選択された点
[frame 33: 29 points]

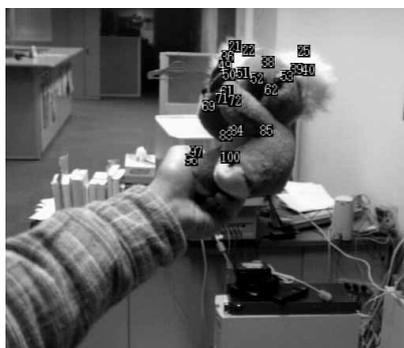


図 9: 選択された点
[frame 99: 23 points]



図 10: 外れ値
[frame 33: 19 points]



図 11: 外れ値 [frame 33: 6 points]

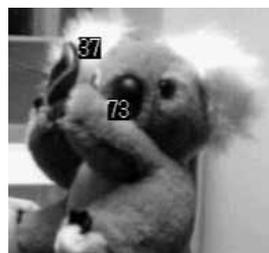


図 12: 外れ値 [frame 99: 2 points]

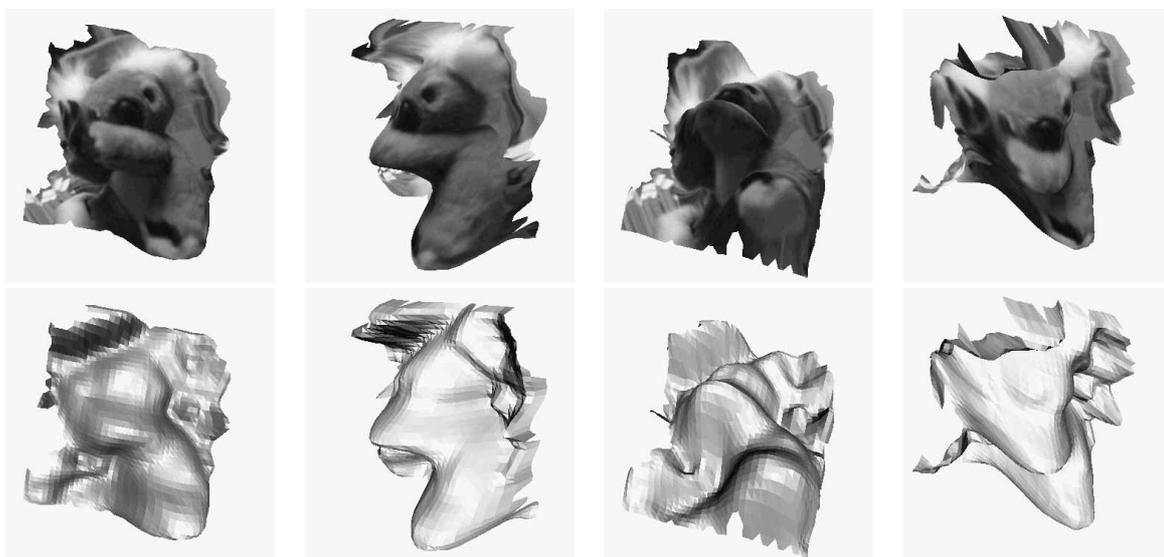


図 13: 第 64,65,66 フレームを用いた仮想 3 眼ステレオにより復元された対象物体の密な 3 次元形状