

頑健な恒等写像学習を用いた 計量アフィン射影画像列からの運動と形状の復元

藤木 淳[†] 高橋 隆 史^{††} 栗田 多喜夫[†]

点特徴を用いた複数の画像からのカメラ運動と立体形状の復元は基本的かつ重要な問題であり、手法の頑健化も含め様々な復元アルゴリズムが提案されてきた。通常、アフィン射影に基づく復元は計測行列へのアフィン空間のあてはめとその結果を用いたユークリッド復元の2段階からなるが、追跡ミスなどにより計測行列の一部が大きな誤差を持つ場合にロバスト推定を用いてアフィン空間のあてはめを行うと過適合を起こし、ユークリッド復元結果が悪くなる場合がある。本稿では3層線型パーセプトロンに対し、結合係数に制約を入れ、かつ学習則にロバストな推定関数を組み込んだ恒等写像を学習させることによって2段階の推定を同時に行いカメラ運動と立体形状を復元する頑健な手法を提案する。提案手法では誤差が大きなデータが3割程度あっても良い精度の復元解を得ることができる。

Recovering the 3-D Shape from a Sequence of Metric Affine Projection Images via Robust Autoassociative Learning

JUN FUJIKI,[†] TAKASHI TAKAHASHI^{††} and TAKIO KURITA[†]

Recovering the camera motion and the object shape from multiple images with point correspondences is the fundamental and important problem, then many algorithms, including their robust extension, had presented. Usually, recovering under affine projection consists of two steps: the one step is fitting an affine space onto measurement matrix, and the other step is estimate Euclidean reconstruction from the affine space. When some data are contaminated by large errors owing to failure of tracking, estimation of affine space via robust estimator overfit to the measurement matrix and it derives bad Euclidean reconstruction. In this paper, we present a new robust algorithm which work well even if thirty percent of data have large error. To estimate both steps of the reconstruction simultaneously, the algorithm is based on autoassociative learning of a multilayered neural network under robust criterion with their connecting coefficients constrained.

1. はじめに

点特徴を用いた複数の2次元画像からカメラ運動と対象物体の立体形状を同時に復元する問題はコンピュータビジョンにおいて基本的かつ重要な問題である。

理論的なカメラはピンホールカメラと同等な透視射影カメラであるから、透視射影を用いたカメラ運動と立体形状の復元が理想である。しかしながら透視射影画像からの復元問題は非線型逆問題となるため、ノイ

ズに敏感で初期値依存性が高く数値計算上不安定であるという問題があり、安定して3次元形状を復元するのは難しい。

そこで、正射影モデルなどの透視射影のアフィン近似射影に基づいた復元手法が提案された。近似射影である以上モデルの近似誤差が不可避であるから、復元精度には自ずと限界がある。しかし画像の撮像条件によっては十分実用に耐えることができ、何よりも線型逆問題であるため、非線型写像の逆問題に比べて高速かつ数値計算上安定して解が得られるという利点がある。また、透視射影画像からアフィン近似射影画像を反復的に推定する手法³⁾のように、非線型最適化問題である透視射影からの復元問題の初期値として用いることができることから、アフィン近似射影からの復元問題は重要であり、多くのアフィン近似射影からの

[†] 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology

^{††} 龍谷大学理工学部数理情報学科

Faculty of Science and Technology, Ryukoku University

復元アルゴリズムが提案された。

それらの中で、因子分解法^{7),15),21)} は特に優れた手法である。アフィン近似射影のもとでは画像座標はカメラの基底ベクトルと立体の3次元座標の内積として得られる。よってこれら内積、つまり画像座標を並べてできる計測行列 (measurement matrix) と呼ばれる行列は、カメラの基底ベクトルを並べた行列と立体の3次元座標を並べた行列の積として表現することができる。よって複数の画像からのカメラ運動と立体形状の復元問題は画像座標を並べた行列の分解と同等であるというのが因子分解法の基本的な考え方である。

この分解において注意すべきことは、実際に観測される画像は透視射影画像であるということである。そのため何らかの意味で透視射影画像からアフィン近似射影画像を推定しなければならない。アフィン近似射影の因子分解法では、透視射影をアフィン射影で近似する際に生ずるモデルとしての誤差を、画像座標の観測において生ずる特徴点の追跡誤差や、デジタル画像において画像座標が量子化される際に生ずる量子化誤差と同列に扱って、観測誤差として取り除いている。

アフィン近似射影の因子分解法においてこれらの誤差を除去するには、観測された画像座標からアフィン近似射影によって構成される3次元(以下の)アフィン空間を推定すればよい。このことは、因子分解法の言葉で述べれば、すでに述べたような誤差を含む“観測された”計測行列から、アフィン射影によって得られるであろうランクが3以下の“真の”計測行列を推定することとなる。そして通常、アフィン近似射影の因子分解法ではこの推定を特異値分解 (SVD; Singular Value Decomposition) を用いて実現する。具体的には、計測行列を特異値分解し、主要3特異値のみを残し、残りの特異値を0とおくことによってランクが3以下の“真の”計測行列を推定している。この推定は、観測された計測行列に対して最小二乗 (LS; Least Squares) 基準で最も近いランク3の行列を推定していることに相当している。つまり、因子分解法における特異値分解とは、最小二乗推定を行うための簡便な方法であり、特異値分解は因子分解法にとって不可欠な操作ではなく^{6),11)}、最小二乗推定以外の推定基準で“真の”計測行列を推定するには特異値分解以外の方法を用いなければならない。実際、最小二乗推定量は外れ値に弱いので、外れ値を含む状況において“真の”計測行列を頑健に推定するためには最小二乗推定よりもロバスト推定を行う方が良い結果が得られる。

この“真の”計測行列を推定、一般的にはアフィン

部分空間のあてはめをロバストに行う方法は数多く研究されている^{2),4),9),12),13),20),22),23),25),26)}。

一般にロバスト推定はランダムサンプリングを行い RANSAC⁵⁾ などの基準により最適な空間を探索する手法^{9),13),20),22),26)} と、適当な誤差関数を設定した後に誤差関数を最小化する手法^{2),4),12),23),25)} とに分けて考えることができる。

RANSAC などをを行い最適な空間を探索する手法として、最小二乗中央値 (LMedS; Least Median of Squares) 基準¹⁷⁾ を用いて外れ値を判定して、残った点全体を用いて3次元アフィン空間を表現し、外れ値をその3次元アフィン空間に射影することによって“正しい”計測行列を推定する手法^{9),13),20)} が提案されている。これらの手法はいずれも、joint image²⁴⁾ のすべてを参照して外れ値か正常値を判定するので、正常値であるようなサンプル、つまり全画像にわたって精度良く追跡された特徴点が多量に存在すれば強力な手法だが、全画像にわたって精度良く追跡された特徴点が多量に得られない場合は良い結果が得られないという問題がある。そして外れ値の割合が多くなると joint image 全体が正常値であるような特徴点が多くなるので、推定が破綻する可能性がある。また、途中まで精度良く追跡できていたが、途中から追跡に失敗したような特徴点は外れ値と判定され、追跡に成功していた部分の追跡情報まで取り除いてしまうこととなるので、計測行列の持つ情報を活かしきれていないという欠点もある。この問題を克服するためには joint image の一部だけを用いてランダムサンプリングをするように手法を改良する必要があるが、どの一部だけを選択するかは組合せは指数関数的に増大するので、相当の工夫が必要となる。これは MLESAC²²⁾ についても同様である。また、計測行列からなす角度の大きい3つの行を抜き出してその線型結合として3次元アフィン空間を表現する手法²⁶⁾ も、抜き出された行に大きな誤差が含まれている場合は復元精度が悪くなるという問題がある。

次に適当な誤差関数を設定した後に誤差関数を最小

MLESAC²²⁾ におけるランダムサンプリングは尤度を最大化するための初期値を求めるためにあり、最終的には尤度を最大化するように繰返し計算を行う手法だから、両方に属すると考えることができる。

LMedS 基準を用いた逐次型因子分解法¹³⁾ では、ある時点での復元結果を用いて、新しく入力された1枚の画像の中でのみ外れ値の判定を行うので、全画像にわたって精度良く追跡された特徴点が多量に得られないことがあるという問題は回避できる。また、追跡に成功していた部分の追跡情報まで取り除いてしまうという欠点も克服できる。

化する手法^{2),4),12),23),25)}であるが、これらの違いは、どのように誤差関数を設定するか、そして誤差関数をどのように最小化するかである。

ロバスト推定における誤差関数としては、 L_1 ノルム¹²⁾、 L_1 ノルムの近似曲線である Huber の推定関数、対数双曲線余弦関数、Geman-McClure の推定関数^{2),23)} など多くの M 推定 (M-estimation) に基づく誤差関数が提案されており、ほかに観測データに外れ値が正常値のラベルを与え、正常値に対しては二乗誤差を、外れ値に対しては定数誤差としての罰金項を計算したものを誤差関数とする手法²⁵⁾ や、これを拡張した、データの正常値らしさに相当する重みに二乗誤差を乗じたものに正常値らしさから計算される罰金項を加えた誤差関数を最小化するようにあてはめべきアフィン空間と正常値らしさを定める手法^{2),23)} も提案されている。この誤差関数を最小化するようにアフィン空間と正常値らしさを定める手法は、罰金項から定まる M 推定の推定関数を最小化するようにアフィン空間を定める問題に帰着することができる^{2),23)}。ここで、どの誤差関数を選択するのが適切であるかは、誤差の持つ確率分布に依存するため、単純に判断することはできない。

なお、提案手法では、推定関数として対数双曲線余弦関数を用いてアフィン空間をあてはめ、あてはめたアフィン空間を用いて外れ値が正常値のラベル付けをし、正常値のみを用いてアフィン空間の再推定を行う。

誤差関数を最小化のための手法としては、最急降下法^{2),23)}、ニュートン法、凸計画法¹²⁾などが考えられるが、提案手法では単峰である凸関数に対して大域的収束性が保証されている最急降下法を用い、その最急降下法を多層パーセプトロン (MLP; Multi-Layer Perceptron) の恒等写像学習 (Autoassociative learning) を用いて実現する。

つまり本稿では、画像ごとに特徴点が例外値か否かを判定できるようにするため、最小二乗推定としての因子分解法を M 推定としての因子分解法に拡張する。具体的には、"観測された" 計測行列に最も近い "真の" 計測行列の推定量を、最小二乗推定量から M 推定量 (M-estimator) へと変更し、M 推定の誤差関数を最小化するための最急降下法を 3 層線型パーセプトロンの恒等写像学習を用いて実現する。

ニューロコンピューティングでは、3 層線型パーセ

プトロンに対して最小二乗基準で恒等写像学習を行うと、その中間層には主成分分析 (PCA; Principal Component Analysis) と等価な特徴が自己組織化される¹⁾ ことが知られている。より詳しくいえば、入力層と出力層の素子の個数が N 、中間層の素子の個数が $H (< N)$ である 3 層線型パーセプトロンに対して最小二乗基準で恒等写像学習を行った後にデータを入力すると、出力層はデータ列を第 1 ~ 第 H 主成分ベクトルで張られる空間への正射影ベクトルとなり、中間層はその正射影を第 1 ~ 第 H 主成分ベクトルの線型結合で表現したときの係数ベクトルをアフィン変換したものとなる。ここで、主成分分析の表現が特異値分解によって得られることに着目すると 3 層パーセプトロンを利用してアフィン近似射影の因子分解法が実装できることが分かる⁸⁾。ここで、アフィン近似射影の因子分解法では 3 次元アフィン空間を推定するので、中間層の素子数が 3 である 3 層パーセプトロンに対して joint image を最小二乗基準により恒等写像学習させることになり、このとき中間層は立体の 3 次元座標をアフィン変換したものとなり、出力層は推定された "真の" データとなる。

さらに、M 推定の誤差関数を最小化するための最急降下法を 3 層線型パーセプトロンの恒等写像学習を用いて実現する際に、3 層線型パーセプトロンの中間層の持つアフィン変換の自由度を制御し、中間層に対象物体のユークリッド座標 (正規直交座標) が自己組織化されるように、パーセプトロンの結合係数に因子分解法の計量拘束 (metric constraint) に対応する制約を導入する。

一般に、因子分解法の枠組みで記述されるほとんどの手法は観測された計測行列から "真の" 計測行列を推定する第 1 段階と "真の" 計測行列を用いて運動と形状を推定する第 2 段階に分けられている。しかし、それぞれの推定において最良の推定を行ったとしても全体として最良の推定を行っているとは限らず、第 1 段階における推定結果が過適合することによって第 2 段階における推定結果が悪くなる ことがある。しかし提案手法では、パーセプトロンの結合係数に制約を導入することによりこの 2 つ推定を同時に行うため、第 1 段階における推定の過適合によって第 2 段階における推定結果が悪くなることを避けることができる。

以上のことから、本稿では、3 層線型パーセプトロ

文献 2) では、例として誤差 x 、正常値らしさ z に対して $E = zx^2 + (\sqrt{z} - 1)^2$ を最小化するように z を定めると、 $E = \frac{x^2}{1+x^2}$ と Geman-McClure の誤差関数を利用した M 推定と一致することを示している。

計測行列が負または 0 の固有値を持つこともある¹⁶⁾。もちろん、別々に行うこともできる。

ンに対して M 推定関数を用いた基準で結合係数に制約を入れた恒等写像学習を行い、中間層に形状のユークリッド復元解が自己組織化されるような M 推定としての因子分解法を実現する。

この提案手法により、追跡ミスなどにより誤差が大きくなったデータが 25%~30% 程度あっても良い推定結果を与えることが人工データを用いた実験より得られた。また、 M 推定を用いた推定結果を用いた例外値判定を利用して判定された正常値のみを用いて復元値を再推定することにより、誤差が大きくなったデータが 30%~35% 程度あっても良い推定結果を与えることが人工データを用いた実験より得られた。また、実画像を用いた実験においても良好な結果を得ることが分かった。また、提案手法における M 推定の推定関数を下に凸な関数として選び、結合係数の学習則を最急降下法で与えると、大域最適解への収束が保証されるため、提案手法は数値計算上安定して動作する。

本稿の構成は以下のとおりである。2 章では計量アフィン射影 (MAP; Metric Affine Projection)⁷⁾ と、その特殊な場合である scaled orthographic 射影について説明し、3 章ではバッチ処理の因子分解法^{7),21)} について概観する。そして 4 章では 3 層線型パーセプトロンを用いた、最小二乗基準による恒等写像学習を用いた復元手法⁸⁾、および提案手法である M 推定量を用いた基準による恒等写像学習を用いた復元手法について説明する。5 章では人工データによる実験を、そして 6 章では実画像による実験を行う。

2. 計量アフィン射影 (MAP)

本章では MAP と scaled orthographic 射影について説明する。後に述べるように、MAP は scaled orthographic 射影に帰着できるので、先に scaled orthographic 射影について説明する。本稿で用いるカメラは校正済みであり、光軸と画像平面は直交し、光軸と画像平面の交点は既知で画像面上の原点であり、アスペクト比は 1 であるとする。

ここで、カメラの焦点距離を l 、第 f 画像面上の正規直交基底 (カメラ座標枠の X 軸、 Y 軸方向の単位ベクトルの世界座標) を $\{i_f, j_f\}$ 、第 f カメラの光軸方向の単位ベクトル (カメラ座標枠の Z 軸方向の単位ベクトルの世界座標) を k_f 、第 f カメラのカメラ基底行列を $C_f = (i_f, j_f, k_f)^\top$ とし、第 f カメラのカメラ中心の世界座標を t_f とする。また、第 p

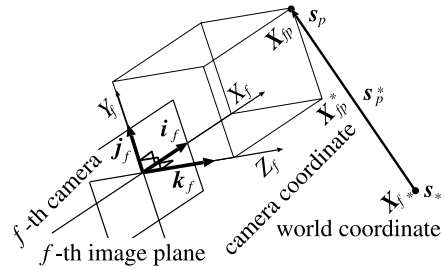


図 1 カメラ座標系と世界座標系

Fig. 1 Camera coordinate and world coordinate.

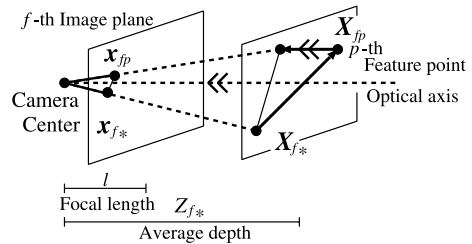


図 2 Scaled orthographic 射影モデル

Fig. 2 Scaled orthographic projection.

特徴点の世界座標を s_p 、第 f 画像面上の第 p 特徴点のカメラ座標、画像座標をそれぞれ

$$X_{fp} = \begin{pmatrix} X_{fp} \\ Y_{fp} \\ Z_{fp} \end{pmatrix}, \quad x_{fp} = \begin{pmatrix} x_{fp} \\ y_{fp} \end{pmatrix}$$

とする ($f = 1, \dots, F; p = 1, \dots, P$).

このとき、ある特定の特徴点である第 * 特徴点 (特徴点の重心を選んでよい) およびその射影からの相対座標を

$$s_p^* = s_p - s_*,$$

$$X_{fp}^* = X_{fp} - X_{f*}, \quad x_{fp}^* = x_{fp} - x_{f*}$$

のように表すと、世界座標から見たカメラ座標は基底 i_f, j_f, k_f における成分表示であるから、

$$s_p^* = X_{fp}^* i_f + Y_{fp}^* j_f + Z_{fp}^* k_f = C_f^\top X_{fp}^*$$

より

$$X_{fp}^* = C_f s_p^* \quad (1)$$

となる (図 1 参照)。

また、行列 X のムーア・ペンローズ逆行列を X^+ で表し、 2×3 行列 $Y = (Y_1, Y_2)^\top$ に対して 3×3 行列 \tilde{Y} を $\tilde{Y} = (Y_1, Y_2, Y_1 \times Y_2)^\top$ で定めるものとする。

2.1 Scaled orthographic 射影

Scaled orthographic 射影は正射影にカメラと物体 (第 * 特徴点) との距離 Z_{f*} を考慮した射影 (図 2) であり、

アフィン座標系における原点と各軸の方向の単位長さ (同伴線型空間の基底) の組を座標枠という¹⁰⁾。

$$\mathbf{x}_{fp}^* = \frac{l}{Z_{f*}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \mathbf{X}_{fp}^* \quad (2)$$

と表現される．グローバルスケールの不定性のため $\{Z_{f*}\}_{f=1}^F$ の比しか求まらないので $\lambda_{f*} = \frac{Z_{f*}}{Z_{1*}}$ ，つまり $\lambda_{1*} = 1$ となるようにグローバルスケールを調整した scaled orthographic 射影モデル

$$\mathbf{x}_{fp}^* = \frac{1}{\lambda_{f*}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \mathbf{X}_{fp}^* = \frac{1}{\lambda_{f*}} \begin{pmatrix} i_f^\top \\ j_f^\top \end{pmatrix} \mathbf{s}_p^*$$

を本稿では用いる．

2.2 計量アフィン射影

MAP⁷⁾ は、ユークリッド復元を可能とする計量が定義されている正射影，scaled orthographic 射影，paraperspective 射影¹⁵⁾ を含む射影であり，計量が定義されている点において校正されていないカメラモデルとしてのアフィン射影¹⁴⁾ とは異なる射影である．MAP の相対座標による具体的な表現は，

$$\mathbf{x}_{fp}^* = A_f \mathbf{X}_{fp}^* = \frac{1}{\lambda_{f*}} B_f \mathbf{X}_{fp}^* \quad (B_f \text{ が既知}) \quad (3)$$

である．ここで 2×3 行列 A_f は第 f カメラのアフィン射影を表す行列であり， λ_{f*} はグローバルスケールパラメータを含む平均奥行き (average depth) である．正射影モデルの場合，平均奥行きは $\lambda_{f*} = 1$ (定数) である．

次に MAP 画像をアフィン変換することにより scaled orthographic 射影画像が得られることを示す．式 (3) で与えられている B_f の特異値分解を

$$B_f = \begin{matrix} R_f & \Sigma_f & D_f \\ (2 \times 3) & (2 \times 2) & (2 \times 2) \end{matrix} \begin{matrix} (2 \times 3) \\ (2 \times 2) \\ (2 \times 3) \end{matrix}$$

($R_f^\top R_f = D_f D_f^\top = I_2$) とすると，直交行列 R_f ，対角行列 Σ_f ，行正規直交行列 D_f は既知である．このとき式 (3) により MAP モデルは

$$\mathbf{x}_{fp}^* = A_f \mathbf{X}_{fp}^* = \frac{1}{\lambda_{f*}} R_f \Sigma_f D_f \mathbf{X}_{fp}^*$$

となる．今，

$$\mathbf{X}_{fp}^* = D_f \mathbf{X}_{fp}^* = (I_2 \quad 0_2) \tilde{D}_f \mathbf{X}_{fp}^*$$

とおくと， $(I_2 \quad 0_2)$ は正射影の表現であるから， \mathbf{X}_{fp}^* は \mathbf{X}_{fp}^* を世界座標系で \tilde{D}_f 回転した後に正射影した画像となる．つまり， \mathbf{X}_{fp}^* は \mathbf{X}_{fp}^* を $\text{span } D_f^\top (= \text{span } A_f^\top)$ へ正射影した画像の基底 D_f^\top による表現となる．ここで

$$\frac{1}{\lambda_{f*}} \mathbf{X}_{fp}^* = \Sigma_f^{-1} R_f^\top \mathbf{x}_{fp}^*$$

の右辺は既知であるから，MAP 画像における画像座標 \mathbf{x}_{fp}^* は，アフィン変換 $\Sigma_f^{-1} R_f^\top$ を施すことにより，平

面 $\text{span } D_f^\top$ への scaled orthographic 射影画像における画像座標 $\frac{1}{\lambda_{f*}} \mathbf{X}_{fp}^*$ に帰着することができる．この平面 $\text{span } D_f^\top$ を仮想画像面 (virtual image plane) と呼ぶ．

3. 因子分解法

本章では MAP の因子分解法 (factorization method) について述べる．ここで 2 次元画像はカメラと立体形状の相対的な位置関係により定まるので立体が固定されカメラが運動していると仮定しても一般性を失わず，世界座標系で考えてよい．

画像座標を並べてできる計測行列を

$$W_{(2F \times P)}^* = \begin{pmatrix} W_1^* \\ \vdots \\ W_F^* \end{pmatrix}, \quad W_f^* = (\mathbf{x}_{f1}^* \quad \cdots \quad \mathbf{x}_{fp}^*)$$

とし，MAP の世界座標系における表現行列を並べてできる運動行列 (motion matrix) M ，および立体の特徴点の世界座標を並べてできる形状行列 (shape matrix) S^* を

$$M_{(2F \times 3)} = \begin{pmatrix} M_1 \\ \vdots \\ M_F \end{pmatrix}, \quad (4)$$

$$M_f = \begin{pmatrix} \mathbf{m}_f^\top \\ \mathbf{n}_f^\top \end{pmatrix} = A_f C_f, \quad (5)$$

$$S_{(3 \times P)}^* = (\mathbf{s}_1^* \quad \cdots \quad \mathbf{s}_p^*) \quad (6)$$

によって定義すると，

$$W_{(2F \times P)}^* = M_{(2F \times 3)} S_{(3 \times P)}^* \quad (7)$$

が成立する ($\text{rank } W^* \leq 3$ に注意)．

ここで M の要素 M_f からカメラ運動 C_f と奥行き λ_{f*} は

$$C_f = (\tilde{A}_f)^{-1} \tilde{M}_f, \quad \lambda_{f*} = \left(\frac{\det(B_f B_f^\top)}{\det(M_f M_f^\top)} \right)^{1/4}$$

によってただちに復元されるので， W^* を M と S^* の積に分解することができればただちにカメラ運動と立体形状を復元することができる．ところが，式 (7) が

$$W_{(2F \times P)}^* = \hat{M}_{(2F \times 3)} \hat{S}_{(3 \times P)}^* \quad (8)$$

と分解可能ならば，任意の 3×3 可逆行列 A に対して

$$W^* = (\hat{M} A) (A^{-1} \hat{S}^*) \quad (9)$$

と分解可能であるため，式 (7) の分解は一意ではなく，ユークリッド復元解に対応する分解が得られるとは限らない．しかし， W^* の任意の 1 つの分解 (8) に対し，

W^* の分解の一般解は式 (9) と表現できるため、1 つの分解 (8) を求めることさえできれば、行列 A を求めることによって式 (7) のユークリッド復元解に対応する分解を実現することができる。

ここで、ユークリッド復元解 M, S^* に対して、 $MA^{-1} = \hat{M}, \hat{S}^* = AS^*$ なる A が存在するので、 \hat{M}, \hat{S}^* はユークリッド復元解 M, S^* をアフィン変換したものがとなる。つまり \hat{M}, \hat{S}^* はアフィン復元解である。

このアフィン復元解をユークリッド復元解に変換するためには、アフィン復元解を表現する空間の計量を適切に定めることと同値であり、そのための条件は、カメラ基底行列 C_f が直交行列であること、つまり $C_f C_f^\top = I_3$ から導かれる計量拘束 (metric constraint)

$$M_f M_f^\top = \frac{1}{\lambda_{f*}^2} B_f B_f^\top$$

$$\iff \left(R_f^\top M_f \right) \left(R_f^\top M_f \right)^\top = \frac{1}{\lambda_{f*}^2} \Sigma_f^2 \quad (10)$$

を満たすことである。正射影の場合を除いて λ_{f*} は未知パラメータであるから、一般には計量拘束として式 (10) から λ_{f*} を消去したものをを用いる。具体的には、

$$P_f = \begin{pmatrix} p_f^\top \\ q_f^\top \end{pmatrix} = R_f^\top M_f, \quad \Sigma_f = \begin{pmatrix} p_f & 0 \\ 0 & q_f \end{pmatrix}$$

とおくと、計量拘束は

$$\begin{cases} \frac{p_f^\top p_f}{p_f^2} - \frac{q_f^\top q_f}{q_f^2} = 0 & (f = 1, \dots, F), \\ p_f^\top q_f = 0 & (f = 1, \dots, F), \\ \frac{p_1^\top p_1}{p_1^2} \left(= \frac{q_1^\top q_1}{q_1^2} \right) = 1 \end{cases} \quad (11)$$

となる。特に Scaled orthographic 射影の場合の計量拘束は、

$$\begin{cases} m_f^\top m_f - n_f^\top n_f = 0 & (f = 1, \dots, F), \\ m_f^\top n_f = 0 & (f = 1, \dots, F), \\ m_1^\top m_1 = 1 \end{cases} \quad (12)$$

となる。そして式 (7) を式 (11) が成立するように分解すればカメラ運動と立体形状のユークリッド復元解を得る。

つまり因子分解法とは \hat{M} と \hat{S}^* を求めた後に式 (11) が成立するように A を求める手法である (詳細は文献 7))。

因子分解法では W を \hat{M} と \hat{S}^* の積に分解する際と、計量拘束を満たすように A を求める際の 2 回推定を行っている。しかし、それぞれの推定が最良でも全体として最良であるとは限らない。これはあまり語られていない因子分解法の欠点である。

なお、一般にアフィン近似射影画像からは互いに鏡映対称な 2 組 (互いにネッカー反転 (Necker reversal) である、という) の復元解を得るが、この 2 組のうちどちらが真の復元解を表現しているかは点对応という情報からのみでは決定できない。そこで本稿では 2 組の復元解を同一視する。

4. 3 層パーセプトロンを用いた恒等写像学習

本章では、3 層線型パーセプトロンを用いた恒等写像学習、結合係数に制約を入れた恒等写像学習⁸⁾、および提案手法について説明する。

4.1 恒等写像学習

入力層と出力層に N 個、中間層に $H (< N)$ 個の素子を持つ 3 層線型パーセプトロンに、入力と出力がなるべく同じになるように学習 (恒等写像学習) させると、中間層には主成分分析と等価な特徴が自己組織化され¹⁾、出力値は入力値から推定されたデータの“真”の値となる。

今、この 3 層線型パーセプトロンに学習させる訓練データを

$$x_p = \begin{pmatrix} x_{p1} \\ \vdots \\ x_{pN} \end{pmatrix} \in \mathbf{R}^N \quad (p = 1, \dots, P),$$

入力ベクトル x_p に対する中間層のニューロンの出力をならべたベクトルを y_p 、3 層線型パーセプトロンの出力ベクトルを

$$\hat{x}_p = \begin{pmatrix} \hat{x}_{p1} \\ \vdots \\ \hat{x}_{pN} \end{pmatrix} \in \mathbf{R}^N \quad (p = 1, \dots, P),$$

入力層から中間層、および中間層から出力層への結合を

$$y_p = \underset{(H \times N)}{U} x_p,$$

$$\hat{x}_p = \underset{(N \times H)}{W} y_p = (w_1, \dots, w_N)^\top y_p$$

とする。このとき、入力と出力の差

$$\varepsilon_{pn} = \hat{x}_{pn} - x_{pn} \quad (p = 1, \dots, P; n = 1, \dots, N)$$

から評価関数を構成し、その関数を最小化することによって恒等写像学習を行う。すると恒等写像学習後の 3 層線型パーセプトロンにデータを入力すると、構成された評価関数から定まる基準によって推定されたデータの“真”の値が出力される。

しかし、恒等写像学習によって中間層は一意には定まらず、アフィン変換の自由度が存在する。なぜなら、任意の $H \times H$ 可逆行列 A に対して (U, W, y_p) と

$(A^{-1}U, WA, A^{-1}\mathbf{y}_p)$ は同じ性能を持つ (同じ入力に対して同じ出力を与える) からである. そのため中間層が 3 素子のパーセプトロンに

$$\mathbf{w}_p^* = \begin{pmatrix} x_{1p} \\ \vdots \\ x_{Fp} \end{pmatrix}$$

を恒等写像学習させても中間層は s_p^* そのものとはならず, 一般には s_p^* をアフィン変換したものが自己組織化される⁸⁾.

4.2 二乗誤差を最小とする学習

通常の恒等写像学習では,

$$\begin{aligned} E &= \frac{1}{2} \sum_{p=1}^P \|\hat{\mathbf{x}}_p - \mathbf{x}_p\|^2 \\ &= \frac{1}{2} \sum_{p=1}^P \sum_{n=1}^N (\hat{x}_{pn} - x_{pn})^2 = \frac{1}{2} \sum_{p=1}^P \sum_{n=1}^N \varepsilon_{pn}^2 \end{aligned}$$

を最小とするように結合係数が学習される. ここで

$$\frac{\partial E}{\partial \mathbf{w}_n} = \sum_{p=1}^P \varepsilon_{pn} \mathbf{y}_p,$$

$$\frac{\partial E}{\partial U} = \sum_{p=1}^P \sum_{n=1}^N \varepsilon_{pn} \mathbf{w}_n \mathbf{x}_p^\top$$

により, 結合係数の学習則は, 学習係数 α を用いて,

$$\Delta \mathbf{w}_n = -\alpha \sum_{p=1}^P (\hat{x}_{pn} - x_{pn}) \mathbf{y}_p,$$

$$\Delta U = -\alpha \sum_{p=1}^P \sum_{n=1}^N (\hat{x}_{pn} - x_{pn}) \mathbf{w}_n \mathbf{x}_p^\top$$

となる.

4.3 M 推定を利用した学習

M 推定は最もよく利用されるロバスト推定法の 1 つであり, 二乗誤差の最小化

$$\min \sum_{p=1}^P \sum_{n=1}^N \frac{1}{2} \varepsilon_{pn}^2$$

の代わりに, それを変形した評価基準を用いる. なぜなら, 二乗誤差は例外値の影響を大きく受けるという欠点があるため, その例外値からの寄与を制御したいからである. そこで, $\frac{1}{2}x^2$ の代わりに, $x=0$ で唯一の最小値を持ち $x>0$ で単調増加であるような正定値の偶関数 $\rho(x)$ を誤差関数として用い,

$$\min \sum_{p=1}^P \sum_{n=1}^N \rho(\varepsilon_{pn})$$

を最小化させる. ここで二乗誤差は $\rho(x) = \frac{1}{2}x^2$ の場

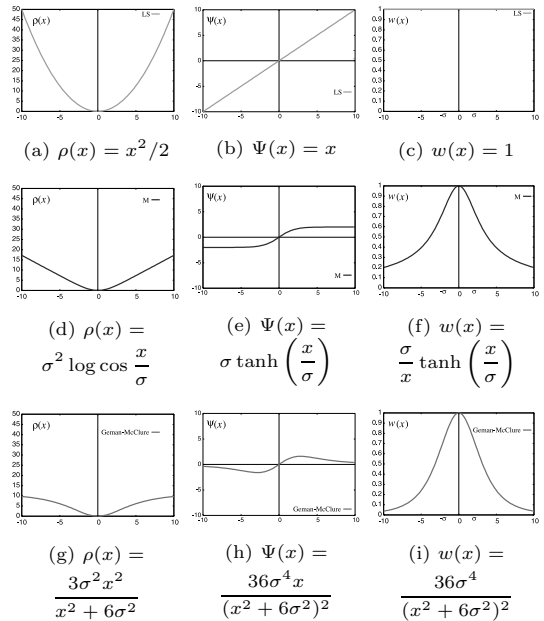


図 3 評価関数

Fig. 3 Evaluation function.

合であり, 二乗誤差最小化は M 推定の一様であるといえることができる.

さて, 評価関数 ρ において, 各データの影響力は ρ の微分である影響関数 (influence function)

$$\Psi(x) = \frac{\partial \rho(x)}{\partial x}$$

によって与えられる. 二乗誤差に対する ρ および Ψ は, 図 3 の (a), (b) のようになり, データの影響力はモデルからのずれに正比例して大きくなるのが分かる.

これと, M 推定で代表的に用いられる 2 つの評価関数を比較する. このとき, 二乗誤差の場合を 1 としたときの各データの重みは

$$w(x) = \frac{\Psi(x)}{x}$$

によって表現でき, これを重み関数 (weight function) と呼ぶ.

さて M 推定で代表的に用いられる 1 つ目の評価関数は Ψ が双曲線正接関数 \tanh となる対数双曲線余弦関数 $\log \cosh$ である. 本稿では, そのマクローリー展開が $\frac{1}{2}x^2 + o(x^2)$ となるように定数倍した

本稿では, 最小二乗推定の誤差関数を $\frac{1}{2}x^2$ としたので, 重み関数は $w(x) = \frac{\Psi(x)}{x}$ と定義されるが, 最小二乗推定の誤差関数を x^2 とすると重み関数は $w(x) = \frac{\Psi(x)}{2x}$ と定義される.

$$\begin{aligned} \rho(x) &= \sigma^2 \log \cosh \frac{x}{\sigma} \\ &= \frac{1}{2}x^2 + \frac{1}{12\sigma^2}x^4 + o(x^4) \quad (\sigma \neq 0) \end{aligned}$$

を用いることにする．このとき，

$$\Psi(x) = \sigma \tanh \frac{x}{\sigma}, \quad w(x) = \frac{\sigma}{x} \tanh \frac{x}{\sigma}$$

である．

M 推定で代表的に用いられる 2 つ目の評価関数は Geman-McClure の ρ 関数である．そのマクローリン展開が $\frac{1}{2}x^2 + \frac{1}{12\sigma^2}x^4 + o(x^4)$ となるように定数倍した

$$\rho(x) = \frac{3\sigma^2 x^2}{x^2 + 6\sigma^2}$$

を用いることにする．このとき，

$$\Psi(x) = \frac{36\sigma^4 x}{(x^2 + 6\sigma^2)^2}, \quad w(x) = \frac{36\sigma^4}{(x^2 + 6\sigma^2)^2}$$

である．

図 3 の (d), (e), (f) に双曲線正接関数の ρ, Ψ および w を, (g), (h), (i) に Geman-McClure の ρ 関数の場合の ρ, Ψ および w を示す．データがモデルからあまりずれていない場合, データの影響力はモデルからのずれに正比例して大きくなるが, データがある程度モデルから離れるとその影響力は, 双曲線正接関数の場合はほぼ一定の値と影響力が抑えられており, Geman-McClure の ρ 関数の場合は影響力が 0 となっていることが分かる．一般に, Geman-McClure の ρ 関数には例外値の影響力が 0 となり, 例外値の影響をほぼ完全に除去できるという利点があるものの, 初期値が例外値に適合した場合, 逆に正常値を例外値と判定して正常値の影響力を 0 としてしまうので誤った解に収束しやすくなるという欠点を持つ．

いずれの評価関数を用いるにせよ,

$$E = \sum_{p=1}^P \sum_{n=1}^N \rho(\varepsilon_{pn})$$

を最小とするように結合係数が学習される．

ここで評価関数はパラメータ $\sigma (\neq 0)$ を含むが, σ は, 誤差 ε_{pn} の分布に従って

1. $\varepsilon^{\text{med}} = \text{med} |\varepsilon_{pn}|$ を計算,
2. 標準偏差の推定値を次の式で計算¹⁹⁾:

$$\hat{\sigma} = 1.4826 \left(1 + \frac{5}{PN-1} \right) \varepsilon^{\text{med}},$$
3. $\sigma = \hat{\sigma}$ とする,

のように決めるものとする．いったん, 評価関数を定めると,

$$\frac{\partial E}{\partial \mathbf{w}_n} = \sum_{p=1}^P \Psi(\varepsilon_{pn}) \mathbf{y}_p,$$

$$\frac{\partial E}{\partial \mathbf{U}} = \sum_{p=1}^P \sum_{n=1}^N \Psi(\varepsilon_{pn}) \mathbf{w}_n \mathbf{x}_p^\top$$

により, 結合係数の学習則は, 学習係数 α を用いて,

$$\Delta \mathbf{w}_n = -\alpha \sum_{p=1}^P \Psi(\varepsilon_{pn}) \mathbf{y}_p,$$

$$\Delta \mathbf{U} = -\alpha \sum_{p=1}^P \sum_{n=1}^N \Psi(\varepsilon_{pn}) \mathbf{w}_n \mathbf{x}_p^\top$$

と表現される．ここで, 大域収束性を保証するために結合係数の学習則を最急降下法により与えたが, ニュートン法に基づく学習則を用いると高速に動作することが期待できる．

4.4 結合係数に制約を入れた恒等写像学習

～提案手法～

主成分分析の表現が特異値分解で与えられることから, 中間層が 3 素子で構成される 3 層パーセプトロンにより w_p^* を恒等写像学習させると中間層には s_p^* と等価な特徴が自己組織化される (図 4) のであった．

もちろん, 中間層の表現によらず恒等写像学習から入力データの評価基準を満たす最良の推定値が得られるので恒等写像学習をデータ補正の目的のためだけに用い, その後に通常の因子分解法を適用してもよい．しかし恒等写像学習によるデータを補正の際に過剰に適合してしまい, 因子分解法の計量拘束を満たすために最良の推定をしたとしても全体として最良でない推定となることがある．そこで本稿では, 中間層にユークリッド復元解が自己組織化されるように結合係数に因子分解法における計量拘束に対応する制約を入れることによって, この 2 つの推定を同時に行う．具体的には, 計量拘束 (12) を反映させるために, エネルギー関数 E に

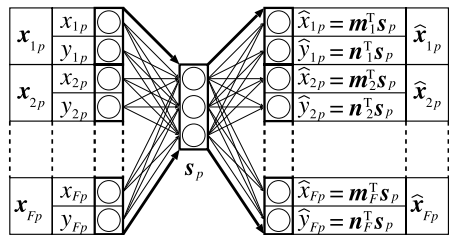


図 4 因子分解法と恒等写像学習
Fig. 4 Factorization and autoassociative learning.

$$E_1 = \frac{1}{2} \sum_{f=1}^F \|m_f^\top m_f - n_f^\top n_f\|^2,$$

$$E_2 = \frac{1}{2} \sum_{f=1}^F \|m_f^\top n_f\|^2$$

という評価基準を加えた,

$$E = \sum_{p=1}^P [\rho(\hat{x}_{pf} - x_{pf}) + \rho(\hat{y}_{pf} - y_{pf})] \\ + \beta(E_1 + E_2)$$

を用いて学習則を導出した．ここで， β は，M 推定による誤差と制約条件とのバランスを調節するパラメータである．本稿の場合， β の値を経験的に決めなければならないという欠点があるが，2 つの推定を同時に行うことができるという利点がある．このとき，結合係数の更新式は，以下ようになる．

$$\Delta m_f = -\alpha \sum_{p=1}^P \Psi(\hat{x}_{pf} - x_{pf}) s_p \\ - \alpha \beta \{ (m_f^\top m_f - n_f^\top n_f) m_f + (m_f^\top n_f) n_f \} \\ \Delta n_f = -\alpha \sum_{p=1}^P \Psi(\hat{y}_{pf} - y_{pf}) s_p \\ - \alpha \beta \{ (n_f^\top n_f - m_f^\top m_f) n_f + (m_f^\top n_f) m_f \}$$

となる ($\beta = 0$ の場合は結合係数に制約がない場合の学習)．

なお，実際の学習には慣性定数 γ を用いて

$$\Delta w = \Delta w^{\text{current}} + \gamma \Delta w^{\text{old}}$$

(w は m_f または n_f) のように慣性項を付け加えてある．

4.5 頑健な恒等写像学習後の処理

M 推定を用いた恒等写像学習が終了した後， $|\varepsilon_{pn}| > \sigma$ を満たすものを例外値と判定し，正常値のみを用いて最小二乗基準により恒等写像学習⁸⁾を行ったものを最終的な復元解とする．

5. 人工データによる実験

本章では人工データを用いて提案手法を評価する．データ生成に用いた射影は透視射影である．

5.1 データ生成と学習におけるパラメータ

100 個の特徴点 $\{s_p\}_{p=1}^{100}$ は 1 辺が 200 ピクセルの立方体の一様分布から生成した．また 10 個のカメラ基底行列は， $C_1 = I_3$ とし， $C_2 \sim C_{10}$ はオイラー角表示の 3 つの角度が一様分布に従うように生成した．さらに 10 個の奥行きパラメータは， $\lambda_{1*} = 1$ ， $\lambda_{2*} = 2$

とし， $\lambda_{3*} \sim \lambda_{10,*}$ は閉区間 $[1, 2]$ 上の一様分布から生成した．

これら C_f ， λ_{f*} ， s_p からカメラ座標を $X_{fp} = \lambda_{f*}^{-1} C_f s_p$ と定め，透視射影により画像座標を計算した．計算された画像座標を通常値と例外値 2 つのグループに分け，通常値のグループのデータに標準偏差が 1 ピクセルとなるガウスノイズを添加し，例外値のグループのデータに閉区間 $[25, 35]$ ， $[-35, -25]$ の和集合である区間 $[25, 35] \cup [-35, -25]$ 上の一様分布から生成されるノイズを添加した．

なお，学習定数 $\alpha = 1 \times 10^{-7}$ ，制約項の重み $\beta = 1000$ ，慣性定数 $\gamma = 0.9$ とした．そして学習回数は 1000 回である．

5.2 復元精度の評価基準

形状復元誤差はフロベニウスノルムでの相対誤差で評価：

$$\| \|S^{*\text{true}} - S^{*\text{estimated}}\| \|S^{*\text{true}}\|^{-1} \times 100(\%).$$

奥行き復元誤差は相対誤差で評価：

$$\frac{\lambda_{f*}^{\text{estimated}} - \lambda_{f*}^{\text{true}}}{\lambda_{f*}^{\text{true}}} \times 100(\%).$$

運動復元誤差は真の k_f と推定値とのなす角度で評価：

$$\cos^{-1} \| (k_f^{\text{estimated}})^\top k_f^{\text{true}} \| \quad (\text{degree}).$$

5.3 実験結果

例外値の割合を増やしながら，最小二乗推定 (LS)⁸⁾，対数双曲線余弦関数に基づく M 推定 (M)，対数双曲線余弦関数に基づく M 推定後に正常値と推定されたデータのみを用いた再推定 (M-wLS)，例外値が正常値か最初から既知の場合に最小二乗推定 (known) による復元誤差を比較した (図 5，図 6，図 7)．

ここで Geman-McClure の ρ 関数に基づく M 推定の結果は省略した．というのも Geman-McClure の ρ 関数に基づく M 推定の結果は対数双曲線余弦関数に基づく M 推定の結果に比べて少し劣ったからである．一般に Geman-McClure の ρ 関数に基づく M 推定は，誤差の大きい例外値の影響力が 0 となり，例外値の影響をほぼ完全に除去できるという利点があるものの，初期値が例外値に適合した場合，逆に正常値を例外値と判定してしまい，例外値に引きずられる

この例外値が未知の場合の最小二乗推定の結果は，因子分解法の 2 段階の推定を一度に行っているため，バッチ処理の因子分解法²¹⁾とは厳密には異なるが，同等の結果と考えて差し支えない．

という欠点を持つ．本実験において Geman-McClure の ρ 関数に基づく M 推定の結果が対数双曲線余弦関数に基づく M 推定の結果に比べて悪くなった原因は Geman-McClure の ρ 関数に基づく M 推定における初期値の与え方の難しさに起因すると考えられる．

さて，図の横軸は例外値の割合であり，縦軸は誤差を表している．グラフは 100 回の実験の平均値であり，エラーバーは平均値から誤差の標準偏差の分だけ上下に伸びている．そして図中の LS, M, M-wLS, known についてはすでに述べたとおりである．

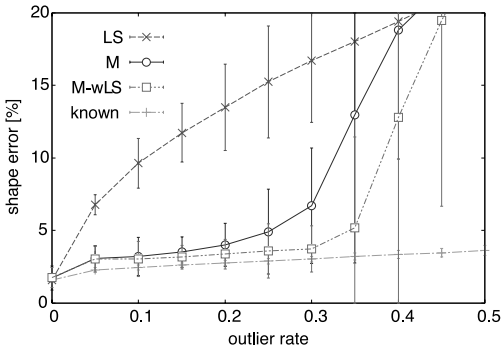


図 5 例外値の割合と形状復元誤差 (透視射影)
Fig. 5 Shape error vs. outlier rate.

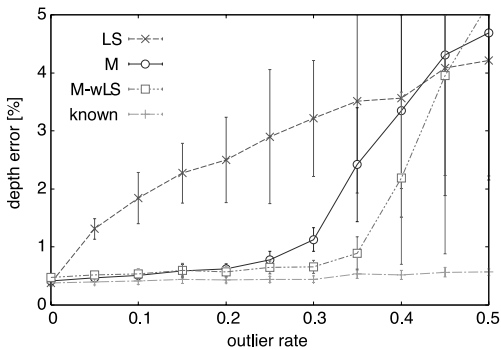


図 6 例外値の割合と奥行き復元誤差 (透視射影)
Fig. 6 Depth error vs. outlier rate.

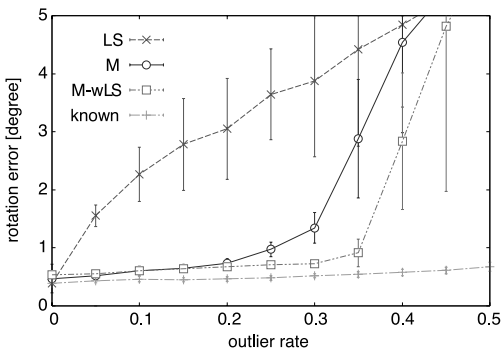


図 7 例外値の割合と運動復元誤差 (透視射影)
Fig. 7 Motion error vs. outlier rate.

図 5 ~ 図 7 により，対数双曲線余弦関数に基づく M 推定後に正常値と推定されたデータのみを用いた再推定 (M-wLS) すると，例外値が 30% ~ 35% 程度までなら，ほぼ完全に例外値の影響を除去できていることが分かる．そして，例外値が 35% を超えると誤差が急激に大きくなることから，例外値が 35% を超えると例外値の影響を除去できていない，つまり提案手法の breakdown point は例外値が 30% ~ 35% のところにあることが分かる．

6. 実画像による実験

本章では，提案手法を実画像に適用し，最小二乗基準と比較した．

6.1 データ生成

被写体は，図 8 に示すような本の上に垂直に立てられたビデオケースである．カメラの内部パラメータはアスペクト比 1:1 で，歪みがなく，光軸と画像面の交点は画像中心であるとした．焦点距離は 35 mm カメラ相当で約 149 mm とした．実験は 100 枚の画像を用い，各画像は 320 × 240 ピクセルである．特徴点の追跡には Kanade-Lucas-Tomasi トラッカ¹⁸⁾ (KLT トラッカ) を用いた．

図 8 の左段は上から第 1, 第 50, 第 100 入力画像であり，右段は左段の画像に対する KLT トラッカによる追跡結果から 100 枚の画像すべてにおいて追跡

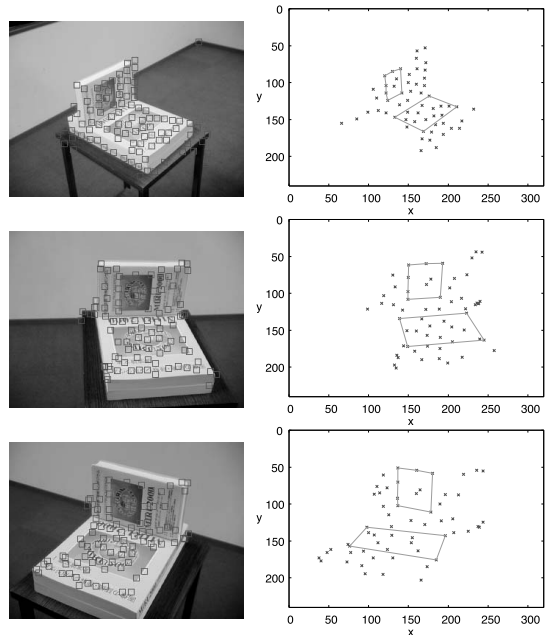


図 8 (上段から) 第 1, 50, 100 入力画像
Fig. 8 The 1-st, 50-th and 100-th input images.

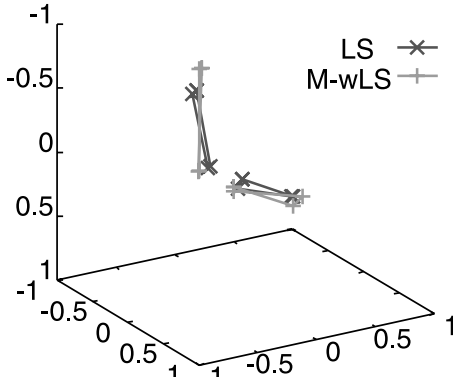


図9 復元された2平面

Fig.9 Reconstructed two planes.

された特徴点選り出したものである(全画像にわたって追跡された特徴点は59点). 復元には, これら特徴点を用いた. なお, 右段の特徴点を結んで作られている領域は, 本の上面に含まれる長方形とビデオケースの前面に含まれる長方形である. ビデオケースの前面に含まれる長方形の4点は全画像にわたってうまく追跡されているが, 本の上面に含まれる長方形の4点のうち右上と左下の2点は, 第50入力画像ではうまく追跡できているが第100入力画像では追跡に失敗している.

復元形状(図9)を評価するために, 本の面とビデオケースの面のなす角度(実測では 90°)を図8に描かれている2つの長方形それぞれの周および内部に含まれる点に対し, 最小二乗基準で平面をあてはめ, それらのなす角度によって計算した. 復元形状を評価する際に平面のなす角度で評価するのは, ユークリッド復元解を評価するための計量に関する量の中で, 平面のなす角度はグローバルスケールパラメータに影響されず, 多数の点から計算されるので個々の点の復元誤差に影響されにくい量だからである.

また, 実画像実験における制約項の重みを $\beta = 100$, 慣性定数を $\gamma = 0.9$ とした. 残りのパラメータについては, 最小二乗推定(LS)の場合, 学習定数は $\alpha = 1 \times 10^{-5}$, 学習回数は 1×10^5 回とし, 対数双曲線余弦関数に基づくM推定(M)の場合, 学習定数は $\alpha = 2.5 \times 10^{-4}$, 学習回数は 1×10^5 回とし, 対数双曲線余弦関数に基づくM推定後に正常値と推定されたデータのみを用いた再推定(M-wLS)の場合, M推定の段階では学習定数は $\alpha = 2.5 \times 10^{-4}$, 学習回数は 1×10^5 回であり, 再推定時では学習定数は $\alpha = 2.5 \times 10^{-4}$, 学習回数は 1×10^4 回とした.

その結果, 最小二乗推定(LS)では 110.0° (差

20.0°), 対数双曲線余弦関数に基づくM推定(M)では 88.7° (差 1.3°), 対数双曲線余弦関数に基づくM推定後に正常値と推定されたデータのみを用いた再推定(M-wLS)では 88.7° (差 1.3°)となり, 提案手法は実画像に対しても良好に動作し, 平面のなす角度に着目した場合, 最小二乗基準(通常の因子分解法)よりも良い結果が得られた. ここで, 対数双曲線余弦関数に基づくM推定とその後正常値と推定されたデータのみを用いた再推定の結果が変わらなかったのは, 例外値が対数双曲線余弦関数に基づくM推定の段階でほぼ取り除かれたからであると考えられる.

7. おわりに

本稿では, 恒等写像学習を用いた計量アフィン射影画像からの運動と形状のロバストな復元アルゴリズムを提案した. 提案アルゴリズムを人工データにより検証した結果, 例外値が30~35%程度含まれていても遜色なく復元できていることが分かる. また, 実画像を用いた実験においても, 最小二乗基準(通常の因子分解法)よりも良好な結果が得られることが確かめられた.

現在, 例外値の除去に利用している統計量はメジアンのみであるが, 他の統計量, もしくはデータの分布そのものを用いることにより, 例外判定を精密に行う工夫も行いたい.

謝辞 実画像データを提供してくださった産業技術総合研究所蔵田武志氏に感謝します.

参考文献

- Baldi, P. and Hornik, K.: Neural networks and principal component analysis, *Neural Networks*, Vol.2, pp.53-58 (1989).
- Black, M.J. and Rangarajan, A.: On the unification of line process, outlier detection and robust statistics with applications in early vision, *International Journal on Computer Vision*, Vol.25, No.19, pp.57-92 (1996).
- Christy, S. and Horaud, R.: Euclidean reconstruction: from paraperspective to perspective, *Proc. 4th European Conf. on Computer Vision*, Cambridge, UK, Vol.2, pp.129-140 (Apr.1996).
- Croux, C. and Filzmoser, P.: Robust factorization of a data matrix, *COMPSTAT, Proc. Computational Statistics*, Payne, R. and Green, P. (Eds), pp.245-249, Physica-Verlag, Heidelberg (1998).
- Fischler, M.A. and Bolles, R.C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and auto-

- mated cartography, *Comm. ACM*, Vol.24, No.6, pp.381-395 (1981).
- 6) 藤木 淳, 蔵田武志: アフィンモデルにおける因子分解法の数理とその検証, 信学技報, PRMU97-22, pp.167-174 (May 1997).
 - 7) 藤木 淳: 点対応を用いた複数の 2 次元画像からの 3 次元形状復元—因子分解法の数理, 統計数理, Vol.49, No.1, pp.77-107 (2001).
 - 8) 藤木 淳, 栗田多喜夫: 恒等写像学習を用いた scaled orthographic 射影画像列からの運動と形状の復元, 信学技報, PRMU2002-222, pp.67-70 (Feb. 2003).
 - 9) Huynh, D.Q. and Heyden, A.: Outlier detection in video sequences under affine projection, *Proc. CVPR*, Kauai, Hawaii, U.S.A., Vol.1, pp.695-701 (Dec. 2001).
 - 10) 河田敬義: アフィン幾何・射影幾何, 岩波書店, 東京 (1976).
 - 11) 金谷健一: 因子分解法を用いない因子分解法: 平行投影から透視変換へ, 信学技報, PRMU98-26, pp.1-8 (June 1998).
 - 12) Ke, Q. and Kanade, T.: Robust L_1 norm factorization in the presence of outliers and missing data by alternative convex programming, *Proc. CVPR*, pp.739-746 (2005).
 - 13) 蔵田武志, 藤木 淳, 興梠正克, 坂上勝彦: ライブ映像からの 3 次元構造と運動の復元のための高速でロバストなアプローチ, 信学論 (D-II), Vol.J84-D-II, No.12, pp.2515-2524 (2001).
 - 14) Mundy, J.L. and Zisserman, A. (Eds): *Geometric invariance in computer vision*, MIT Press (1992).
 - 15) Poelman, C.J. and Kanade, T.: A paraperspective factorization method for shape and motion recovery, *IEEE Trans. PAMI*, Vol.19, No.3, pp.206-218 (1997).
 - 16) Quan, L.: Self-calibration of an affine camera from multiple views, *International Journal on Computer Vision*, Vol.19, No.1, pp.93-105 (1996).
 - 17) Rousseeuw, R.J. and Leroy, A.M.: *Robust Regression and Outlier Detection*, John Wiley & Sons, NY (1987).
 - 18) Shi, J. and Tomasi, C.: Good Features to Track, *Proc. Computer Vision and Pattern Recognition '94*, pp.593-600, Seattle, WA, USA (June 1994).
 - 19) 島井博行, 栗田多喜夫, 梅山伸二, 田中 勝, 三島健稔: ロバスト統計に基づいた適応的な背景推定法, 信学論 (D-II), Vol.J86-D-II, No.6, pp.796-806 (2003).
 - 20) Sugaya, Y. and Kanatani, K.: Outlier removal for motion tracking by subspace separation, *IEICE Trans. INF. & SYST.*, Vol.E-86-D, No.6, pp.1095-1102 (2003).
 - 21) Tomasi, C. and Kanade, T.: Shape and motion from image streams under orthography: A factorization method, *International Journal on Computer Vision*, Vol.9, No.2, pp.137-154 (Nov. 1992).
 - 22) Torr, P.H.S. and Zisserman, A.: MLESAC: A new robust estimator with application to estimating image geometry, *Journal of Computer Vision and Image Understanding*, vol.78, No.1, pp.138-156 (2000).
 - 23) Torre, F.D. and Black, M.J.: A framework for robust subspace learning, *International Journal on Computer Vision*, Vol.54, No.1/2/3, pp.117-142 (2003).
 - 24) Triggs, B.: Matching constraints and the joint image, *Proc. 5th International Conference on Computer Vision*, Cambridge, MA, USA, pp.338-343 (June 1995).
 - 25) Xu, L. and Yuille, A.: Robust principal component analysis by self-organizing rules based on statistical physics approach, *IEEE Trans. on Neural Networks*, Vol.6, No.1, pp.131-143 (1995).
 - 26) Yokoya, N., Takemura, H. and Hwang, K.: A factorization method using 3-d linear combination for shape and motion recovery, *Proc. 14th International Conference on Pattern Recognition*, Brisbane, Australia, pp.959-963 (Aug. 1998).

(平成 17 年 9 月 21 日受付)

(平成 18 年 3 月 20 日採録)

(担当編集委員 坂野 鋭)



藤木 淳

1993 年東京大学工学部計数工学科卒業。1995 年同大学大学院工学系研究科修士課程修了。同年電子技術総合研究所入所。2001 年(独)産業技術総合研究所に改組, 現在に至る。コンピュータビジョンの研究に従事。



高橋 隆史 (正会員)

1994年筑波大学第三学群情報学類卒業。1999年同大学大学院博士課程工学研究科修了。同年日本学術振興会特別研究員(PD)。2001年龍谷大学理工学部数理情報学科助手。2005

年同大学講師，現在に至る。博士(工学)。ニューラルネットワーク，パターン認識の研究に従事。電子情報通信学会，日本神経回路学会，IEEE，INNS 各会員。



栗田多喜夫 (正会員)

1958年生。1981年名古屋工業大学工学部電子工学科卒業。同年電子技術総合研究所入所。1990～1991年カナダ国立科学研究協議会(NRC)招聘研究員。現在(独)産業技術総

合研究所脳神経情報研究部門副研究部門長。工学博士。統計的パターン認識および生体模倣型ビジョンの研究に従事。電子情報通信学会，日本神経回路学会，人工知能学会，行動計量学会，日本顔学会，IEEE CS 各会員。
