

射影変換行列の最適計算によるモザイク生成

金澤 靖* 太田 直哉† 金谷 健一†

* 豊橋技術科学大学知識情報工学系 † 群馬大学工学部情報工学科

本論文では射影変換がモザイク生成を始めとする種々の画像・メディア処理において果たしている役割を述べ、画像の誤差の統計モデルに基づき、射影変換の推定の精度の理論限界を示すとともに、それを達成する最適計算の原理を示す。そして我々の公開しているプログラムによる解の精度が実際にその理論限界を達成する真に最適なアルゴリズムであることを実験的に確認する。また、互いに重なり小さい画像を用いたモザイク生成では射影変換が不安定であり、高精度の計算が要求されることを実画像例を用いて示すとともに、それが我々のプログラムによってどのように改善されるかを示す。我々のプログラムはモザイク生成のみならず、平面物体の3次元復元、航空写真や道路画像の解析、平面パターンを用いたカメラキャリブレーション等、射影変換が現れるすべての画像・メディア処理問題に有効であり、幅広い応用が期待される。

キーワード: 射影変換、モザイク生成、パノラマ画像、統計的最適化、信頼性評価、精度の理論限界

Image Mosaicing by Optimally Computed Homography

Yasushi Kanazawa*, Naoya Ohta† and Kenichi Kanatani†

*Department of Knowledge-based Information Engineering
Toyohashi University of Technology, Toyohashi, Aichi 441-8580 Japan

†Department of Computer Science
Gunma University, Kiryu, Gunma 376-8515 Japan

We first describe the role that homographies play in today's image and media applications including image mosaicing. We then present a theoretical bound on the accuracy of homography computation by introducing a statistical model of image noise and show how that bound can be attained. By doing simulation experiments, we confirm that our program is indeed a truly optimal algorithm that attains that bound. Using real image examples, we point out that instability of homography arises in image mosaicing for images with small overlaps and that high accuracy is required for homography computation. We then demonstrate how the accuracy is improved by our program. Our program is expected to play an essential role in not only image mosaicing but all image and media applications, such as 3-D reconstruction of planar objects, aerial image registration, road image analysis, and camera calibration using planar patterns, where homographies are involved.

Key words: homography, image mosaicing, panoramic image, statistical optimization, reliability evaluation, theoretical accuracy bound

謝辞: 本研究の一部は文部省科学研究費基盤研究C (2) (No. 09680352) によった。

* 441-8580 豊橋市天伯町雲雀ヶ丘 1-1 Tel: (0532)44-6888, Fax: (0532)44-6873

† 376-8515 桐生市天神町 1-5-1 群馬大学工学部情報工学科, Tel: (0277)30-1842, Fax: (0277)30-1801

E-mail: kanazawa@tutkie.tut.ac.jp, ohta@cs.gunma-u.ac.jp, kanatani@cs.gunma-u.ac.jp

1. 序論

撮像が透視変換とみなせる2台のカメラで同一の平面物体を撮影した2枚の画像間は「射影変換」と呼ばれる2次元空間の変換で結ばれる。この変換は

- 「共線」(同一直線上にある)の点を共線の点に写像し、
- 「共点」(同一交点をもつ)の直線を共点の直線に写像する

変換として特徴づけられる [7, 9]。そして、

- 「一般の位置」(どの3点も共線ではない)の4点の組が二つあれば一方を他方に写像する射影変換がただ一つ定まり、
- 「一般の位置」(どの3直線も共点ではない)の4直線の組が二つあれば一方を他方に写像する射影変換がただ一つ定まる。

特に、平面物体とそれを任意の位置から撮影した画像との対応は射影変換となる。例えば、滑走路と着陸する航空機から撮影したその画像や道路と走行する車から撮影したその画像の対応も射影変換であり、この性質がさまざまな応用に利用されている [1, 3, 4, 14, 21]。また、平面物体の2画像からの3次元復元 [11, 20] や平面パターンを用いたカメラキャリブレーション [2, 12, 13, 19] でも射影変換を計算することが最初のステップとなる。

この射影変換の応用として今日注目を集めているのは「モザイク生成」である。これは室内あるいは屋外のシーンのさまざまな方向から撮影した画像を連続的に接合し、それを提示することによって臨場感を増す技術である。接合した画像は「パノラマ画像」とも呼ばれ、今日多くの関心を集めているメディア処理技術の一つである [5, 6, 15, 17, 18, 22]。

これは原理的には古くから航空写真の接合に使われた技術と同じである。モザイク生成の対象となるのは平面とは限らないシーンであるが、シーンがカメラから十分離れていると近似的に平面とみなせる。カメラに近いシーンであってもレンズの中心を固定してカメラを回転すると(ズームは任意に変えてよい)、シーンは実質的には無限遠方にある平面物体と同一視できる。カメラが移動しても、その移動が小さければ画像の変換は射影変換で近似できる。これがモザイク生成の原理である。

このとき問題になるのが、画像間の重なり小さい場合である。互いに大きく重なる画像が連続的に得られればよいが、実際の応用では画像が離散的にしか得られないことも多い。このとき重なりは1辺に沿う細長い領域、あるいは頂点付近の狭い領域となる。重なりが少ないとマッチさせる特徴点がごく少数しか得られない。

このようなときは極めて高精度の計算が要求される。これは、射影変換が8パラメータをもつ不安定な変換であり、一つのパラメータの値がわずかに変化しても、それが画像の変換としては大きな歪みとなり得るからである。

本論文の目的は、モザイク生成における射影変換のこの不安定性を実例によって示し、射影変換の高精度計算の必要性を指摘するとともに、実際に最適計算を行ってこれがどのように改善されるかを示すことである。

我々は既に射影変換の最適計算アルゴリズムを発表し [10, 16]、C++ 言語で書かれたプログラムを公開している。その計算手法は平面物体の2画像からの3次元復元のために既に用いられていたものに他ならない [8, 11]。

今日のメディア処理において、新しく提起された応用では過去に例がないという理由でしばしば素朴かつ発見的な手法が用いられ、それ以前に蓄積されたコンピュータビジョンの技術が生かされないことが多いが、モザイク生成もその一つの例である(コンピュータビジョン研究の初期においても、それ以前に蓄積された写真測量学の技術が生かされなかった)。

2. 射影変換と射影変換行列

式で表すと、射影変換は次の形に書ける。

$$x' = \frac{Ax + By + C}{Px + Qy + R}, \quad y' = \frac{Dx + Ey + F}{Px + Qy + R} \quad (1)$$

行列 \mathbf{H} とベクトル \mathbf{x}, \mathbf{x}' を

$$\mathbf{H} = \begin{pmatrix} A & B & C/f_0 \\ D & E & F/f_0 \\ P/f_0 & Q/f_0 & R/f_0^2 \end{pmatrix} \quad (2)$$

$$\mathbf{x} = \begin{pmatrix} x/f_0 \\ y/f_0 \\ 1 \end{pmatrix}, \quad \mathbf{x}' = \begin{pmatrix} x'/f_0 \\ y'/f_0 \\ 1 \end{pmatrix} \quad (3)$$

と定義すると式 (1) は次のように書ける。

$$\mathbf{x}' = Z[\mathbf{H}\mathbf{x}] \quad (4)$$

ただし $Z[\cdot]$ は z 成分を1とするスケールの正規化作用素であり、 f_0 は $x/f_0, y/f_0$ を1のオーダーにするためのスケール因子である。行列 \mathbf{H} を「射影変換行列」と呼ぶ。これは定義より定数倍の不定性があるので $\|\mathbf{H}\| = 1$ と正規化する(行列 $\mathbf{H} = (H_{ij})$ のノルムは $\|\mathbf{H}\| = \sqrt{\sum_{i,j=1}^3 H_{ij}^2}$ と定義する)。式 (4) は正規化作用素を消去した形で次のように書くこともできる。

$$\mathbf{x}' \times \mathbf{H}\mathbf{x} = \mathbf{0} \quad (5)$$

3. 射影変換行列の推定

データ点 (x_α, y_α) , (x'_α, y'_α) には不確定さがあるので、その共分散行列をそれぞれ Σ_α , Σ'_α とすると、ベクトル表示 \mathbf{x}_α , \mathbf{x}'_α の共分散行列が次のように定義される。

$$V[\mathbf{x}_\alpha] = \begin{pmatrix} \Sigma_\alpha & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{pmatrix}, \quad V[\mathbf{x}'_\alpha] = \begin{pmatrix} \Sigma'_\alpha & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{pmatrix} \quad (6)$$

通常は共分散行列 $V[\mathbf{x}_\alpha]$, $V[\mathbf{x}'_\alpha]$ は定数倍を除いて与えられるとする。すなわち未知の定数 ϵ (「ノイズレベル」と呼ぶ) が存在し、

$$V[\mathbf{x}_\alpha] = \epsilon^2 V_0[\mathbf{x}_\alpha], \quad V[\mathbf{x}'_\alpha] = \epsilon^2 V_0[\mathbf{x}'_\alpha] \quad (7)$$

と書けるとし、 $V_0[\mathbf{x}_\alpha]$, $V_0[\mathbf{x}'_\alpha]$ (「正規化共分散行列」と呼ぶ) を指定する。例えば特徴点の対応をテンプレートマッチングによって定めれば、その残差の方向依存性からこれを定めることができる。ノイズに特に方向性がない場合はデフォルト値として $V_0[\mathbf{x}_\alpha] = V_0[\mathbf{x}'_\alpha] = \text{diag}(1, 1, 0)$ を用いる ($\text{diag}(\dots)$ は \dots を対角要素とする対角行列である)。

以上のように定式化すると、ベクトル表示したデータ $\{\mathbf{x}_\alpha\}$, $\{\mathbf{x}'_\alpha\}$ の真の値を $\{\bar{\mathbf{x}}_\alpha\}$, $\{\bar{\mathbf{x}}'_\alpha\}$ とするとき、射影変換の推定は次の問題となる。

【問題 1】 真の値 $\{\bar{\mathbf{x}}_\alpha\}$, $\{\bar{\mathbf{x}}'_\alpha\}$ に対して拘束条件

$$\bar{\mathbf{x}}'_\alpha \times \mathbf{H} \bar{\mathbf{x}}_\alpha = \mathbf{0} \quad (8)$$

を満たすノルム 1 の行列 \mathbf{H} を誤差のあるデータ $\{\mathbf{x}_\alpha\}$, $\{\mathbf{x}'_\alpha\}$ から推定せよ。

よく用いられる方法は次の関数を最小にする最小二乗法 (「代数的距離最小化法」) である。

$$J_{LS} = \sum_{\alpha=1}^N \|\mathbf{x}'_\alpha \times \mathbf{H} \mathbf{x}_\alpha\|^2 \quad (9)$$

これは未知数 \mathbf{H} の 2 次式であるため、解はその係数行列の最小固有値に対する固有ベクトルとして容易に求まる。この簡便さのために射影変換行列を用いるほとんどすべてのシステムでこの方法が用いられてきた。しかし、この方法では解に統計的な偏差が生じて精度が低下し、この偏差は重みを導入して $\sum_{\alpha=1}^N W_\alpha \|\mathbf{x}'_\alpha \times \mathbf{H} \mathbf{x}_\alpha\|^2$ としても除去できないことが指摘されている [8, 10, 16]。

4. 精度の理論限界

データに誤差がある限り真の値は得られない。その結果どのような方法を用いてもそれ以上の精度の達成

が不可能な限界が存在する。これは理論的に導出できる。射影変換行列 \mathbf{H} の真の値を $\bar{\mathbf{H}}$ とし、何らかの方法で計算した推定値を $\hat{\mathbf{H}}$ とする。その精度を次の「共分散テンソル」で評価する。

$$\mathcal{V}[\hat{\mathbf{H}}] = E[\mathcal{P}((\hat{\mathbf{H}} - \bar{\mathbf{H}}) \otimes (\hat{\mathbf{H}} - \bar{\mathbf{H}}))\mathcal{P}^\top] \quad (10)$$

$E[\cdot]$ は期待値を表す。演算 \otimes は「テンソル積」であり、行列 $\mathbf{A} = (A_{ij})$, $\mathbf{B} = (B_{ij})$ に対して $\mathbf{A} \otimes \mathbf{B}$ は $(ijkl)$ 要素が $A_{ij}B_{kl}$ のテンソルである。またテンソル $\mathcal{P} = (P_{ijkl})$, $\mathcal{T} = (T_{ijkl})$ に対して $\mathcal{P}\mathcal{T}\mathcal{P}^\top$ は $(ijkl)$ 要素が $\sum_{m,n,p,q=1}^3 P_{ijmn}P_{klpq}T_{mnpq}$ のテンソルである。式 (10) 中の $\mathcal{P} = (P_{ijkl})$ は次のように定義された「射影テンソル」である。

$$P_{ijkl} = \delta_{ik}\delta_{jl} - \bar{H}_{ij}\bar{H}_{kl} \quad (11)$$

ただし δ_{ij} は「クロネッカデルタ」であり、 $i = j$ のとき 1、そうでないとき 0 をとる。射影変換行列はノルム 1 に正規化されているから、これは 9 次元パラメータ空間の 8 次元単位球面 S^8 上の点とみなせる。テンソル \mathcal{P} は変位 $\hat{\mathbf{H}} - \bar{\mathbf{H}}$ を S^8 の $\bar{\mathbf{H}}$ での接空間 $T_{\bar{\mathbf{H}}}(S^8)$ へ射影するものである。推定値 $\hat{\mathbf{H}}$ の「平方平均二乗誤差」(root-mean-square error) を次のように定義する。

$$\text{rms}[\hat{\mathbf{H}}] = \sqrt{E[\|\mathcal{P}(\hat{\mathbf{H}} - \bar{\mathbf{H}})\|^2]} \quad (12)$$

射影テンソル \mathcal{P} の定義より、取り得る値の範囲は $0 \leq \text{rms}[\hat{\mathbf{H}}] \leq 1$ である。

統計的最適化理論 [8] から共分散テンソル $\mathcal{V}[\hat{\mathbf{H}}]$ の下界が次のように得られる。

$$\mathcal{V}[\hat{\mathbf{H}}] \succ \epsilon^2 \left(\sum_{\alpha=1}^N \sum_{k,l=1}^3 \bar{W}_\alpha^{(kl)} (\mathbf{e}^{(k)} \times \bar{\mathbf{x}}'_\alpha) \otimes \bar{\mathbf{x}}_\alpha \otimes (\mathbf{e}^{(l)} \times \bar{\mathbf{x}}'_\alpha) \otimes \bar{\mathbf{x}}_\alpha \right)_8^- \quad (13)$$

$$\bar{W}_\alpha = \left(\bar{\mathbf{x}}'_\alpha \times \bar{\mathbf{H}} V_0[\mathbf{x}_\alpha] \bar{\mathbf{H}}^\top \times \bar{\mathbf{x}}'_\alpha + (\bar{\mathbf{H}} \bar{\mathbf{x}}_\alpha) \times V_0[\mathbf{x}'_\alpha] \times (\bar{\mathbf{H}} \bar{\mathbf{x}}_\alpha) \right)_2^- \quad (14)$$

したがって、平方平均二乗誤差の理論的下界が次のように得られる。

$$\text{rms}[\hat{\mathbf{H}}] \geq \sqrt{\text{tr}\mathcal{V}[\hat{\mathbf{H}}]} \quad (15)$$

ただし、テンソル $\mathcal{T} = (T_{ijkl})$ の「トレース」 $\text{tr}\mathcal{T} = \sum_{k,l=1}^3 T_{klkl}$ と定義する。

テンソル \mathcal{T} , \mathcal{S} に対して $\mathcal{T} \succ \mathcal{S}$ は $\mathcal{T} - \mathcal{S}$ が半正値テンソル (固有値がすべて非負のテンソル) であ

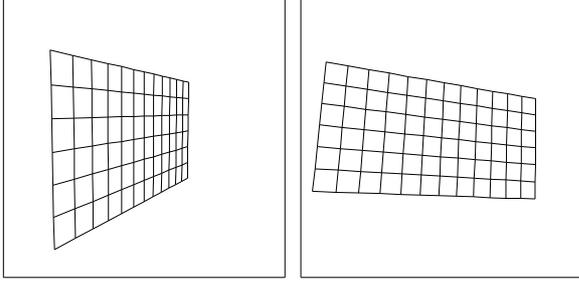


図 1: 格子パタンの射影変換。

ることを意味する。また $(\cdot)_r^-$ はランクを r に射影した「一般逆テンソル」であり、ベクトル $\mathbf{a} = (a_i)$ と行列 $\mathbf{A} = (A_{ij})$ の積 $\mathbf{a} \times \mathbf{A} \times \mathbf{a}$ は (ij) 要素が $\sum_{k,l,m,n=1}^3 \varepsilon_{ikl} \varepsilon_{jmn} a_k a_m A_{ln}$ の対称行列を表す。ただし ε_{ijk} はエディンガトンのイプシロンであり、 (ijk) が (123) の偶置換のとき 1、奇置換のとき -1、その他で 0 をとる。ベクトル $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \mathbf{e}^{(3)}$ を次のように定義した。

$$\mathbf{e}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (16)$$

テンソル \mathcal{T} と行列 \mathbf{A} に対して $\mathcal{T}\mathbf{A} = \lambda\mathbf{A}$ となるとき、行列 \mathbf{A} はテンソル \mathcal{T} の「固有値」 λ の「固有行列」であるという。テンソル $\mathcal{T} = (T_{ijkl})$ と行列 $\mathbf{A} = (A_{ij})$ の積 $\mathcal{T}\mathbf{A}$ は (ij) 要素が $\sum_{k,l=1}^3 T_{ijkl} A_{kl}$ の行列である。テンソル \mathcal{T} の固有値と固有行列は \mathcal{T} を 9×9 行列と同一視して固有値を計算し、9 次元固有ベクトルを 3×3 行列と同一視すればよい [8]。

テンソル $\mathcal{T} = (T_{ijkl})$ は $T_{ijkl} = T_{klij}$ のとき「対称テンソル」と呼ぶ。対称テンソルは 9 個の非負の固有値をもち、対応する固有行列は互いに直交するノルム 1 の行列に選べる。ただし、行列 $\mathbf{A} = (A_{ij})$ と $\mathbf{B} = (B_{ij})$ の内積を $(\mathbf{A}; \mathbf{B}) = \sum_{i,j=1}^3 A_{ij} B_{ij}$ と定義する。固有値がすべて非負のとき、テンソルは「半正値」であるという。

半正値対称テンソル \mathcal{T} の固有値を $\lambda_1 \geq \dots \geq \lambda_9$ 、対応するノルム 1 の固有行列を $\mathbf{U}_1, \dots, \mathbf{U}_9$ とするとき、 $\lambda_r > 0$ であればランク r の一般逆テンソル \mathcal{T}_r^- は次のように計算される。

$$\mathcal{T}_r^- = \sum_{i=1}^r \frac{\mathbf{U}_i \otimes \mathbf{U}_i}{\lambda_i} \quad (17)$$

5. 射影変換行列の最適推定

統計的最適化理論 [8] によると、式 (13) の下界を第 1 近似において、すなわち $O(\epsilon^4)$ の項を除いて達成す

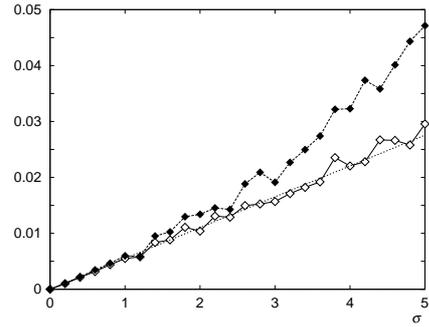


図 2: 平方平均二乗誤差。実線: 最適計算、破線: 最小二乗法、点線: 理論的下界。

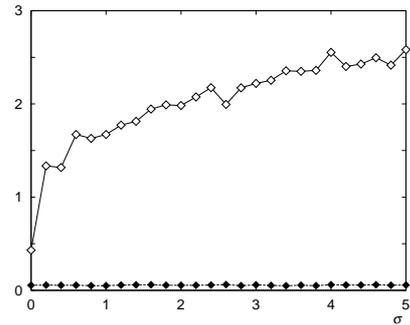


図 3: 計算時間。実線: 最適計算、破線: 最小二乗法、点線: 理論的下界。

るには二乗「マハラノビス距離」

$$J = \sum_{\alpha=1}^N (\mathbf{x}_\alpha - \bar{\mathbf{x}}_\alpha, V_0[\mathbf{x}_\alpha]_2 (\mathbf{x}_\alpha - \bar{\mathbf{x}}_\alpha)) + \sum_{\alpha=1}^N (\mathbf{x}'_\alpha - \bar{\mathbf{x}}'_\alpha, V_0[\mathbf{x}'_\alpha]_2 (\mathbf{x}'_\alpha - \bar{\mathbf{x}}'_\alpha)) \quad (18)$$

を式 (8) の拘束条件のもとに最小化すればよい (「最尤推定」)。ただし、ベクトル \mathbf{a}, \mathbf{b} の内積を (\mathbf{a}, \mathbf{b}) と書く。式 (8) を線形近似し、ラグランジュ乗数を導入して拘束条件を消去すると、問題は次式の最小化となる。

$$J = \sum_{\alpha=1}^N (\mathbf{x}'_\alpha \times \mathbf{H}\mathbf{x}_\alpha, \mathbf{W}_\alpha (\mathbf{x}'_\alpha \times \mathbf{H}\mathbf{x}_\alpha)) \quad (19)$$

$$\mathbf{W}_\alpha = \left(\mathbf{x}'_\alpha \times \mathbf{H}V_0[\mathbf{x}_\alpha]\mathbf{H}^\top \times \mathbf{x}'_\alpha + (\mathbf{H}\mathbf{x}_\alpha \times V_0[\mathbf{x}'_\alpha] \times (\mathbf{H}\mathbf{x}_\alpha)) \right)_2^- \quad (20)$$

この「残差」(J の最小値) を \hat{J} とすると第 1 近似において \hat{J}/ϵ^2 が自由度 $2(N-4)$ の χ^2 分布に従うことが知られている [8]。したがって、二乗ノイズレベル ϵ^2 の不偏推定量が次のように得られる。

$$\hat{\epsilon}^2 = \frac{\hat{J}}{2(N-4)} \quad (21)$$

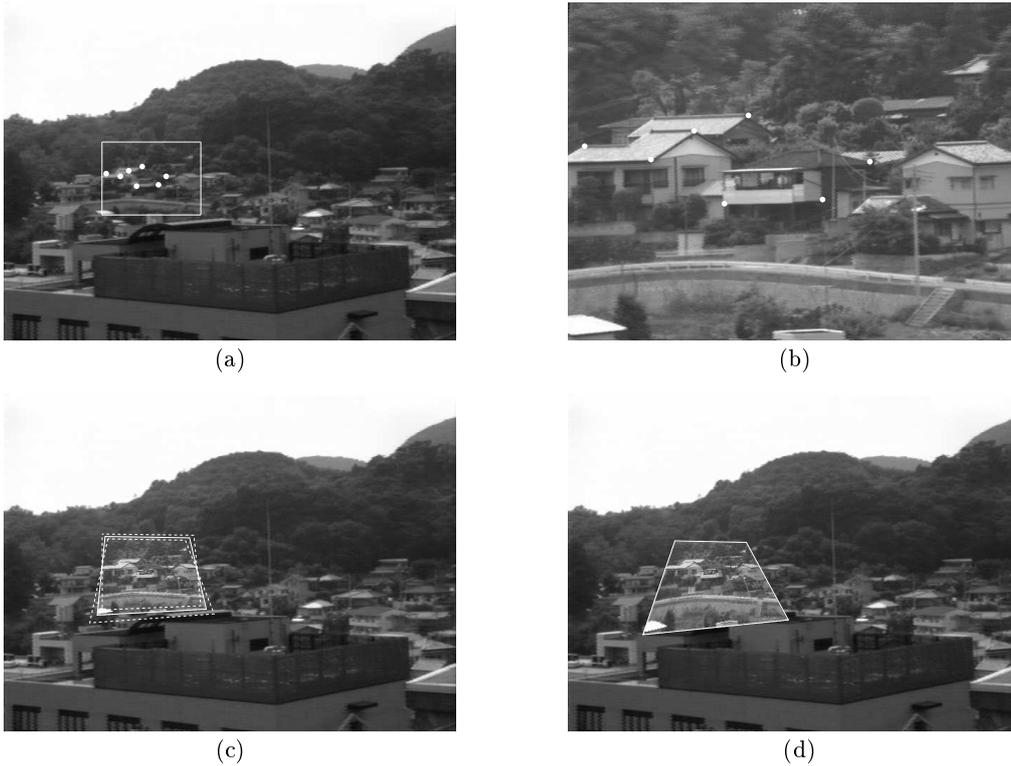


図 4: (a) 遠方の実画像シーン。(b) 図 (a) の白枠部分に対応するズーム画像。(c) 最適計算した射影変換による写像 (実線の枠内) と標準偏位による境界線 (破線)。(d) 最小二乗法による結果。

得られる解 $\hat{\mathbf{H}}$ が最適であるから、その共分散テンソル $\mathcal{V}[\hat{\mathbf{H}}]$ は式 (13) の右辺の真の値 \mathbf{H} , $\{\bar{\mathbf{x}}_\alpha, \bar{\mathbf{x}}'_\alpha\}$ に計算値 $\hat{\mathbf{H}}$ とデータ $\{\mathbf{x}_\alpha, \mathbf{x}'_\alpha\}$ を代入することによって評価できる。また二乗ノイズレベル ϵ^{-2} には式 (21) の推定値を用いる。

得られた共分散テンソル $\mathcal{V}[\hat{\mathbf{H}}]$ の最大固有値 λ_{\max} に対する固有行列 \mathbf{U}_{\max} は \mathbf{H} の 9 次元パラメータ空間で誤差の最も生じやすい方向を表し、その固有値 λ_{\max} が分散を表す。そこで推定値 $\hat{\mathbf{H}}$ をその方向に標準偏差に相当する $\pm \sqrt{\lambda_{\max}}$ だけ摂動させ、単位行列に正規化し、得られる二つの行列 $\{\mathbf{H}^{(+)}, \mathbf{H}^{(-)}\}$ を「標準偏位」と呼ぶ [16] ($N[\cdot]$ はノルムを 1 とする正規化を表す)。

$$\begin{aligned} \mathbf{H}^{(+)} &= N[\hat{\mathbf{H}} + \sqrt{\lambda_{\max}} \mathbf{U}_{\max}] \\ \mathbf{H}^{(-)} &= N[\hat{\mathbf{H}} - \sqrt{\lambda_{\max}} \mathbf{U}_{\max}] \end{aligned} \quad (22)$$

6. 精度と計算時間

最適解を求めるには式 (19) を適当なライブラリプログラムを用いて最小化すればよいが、我々は「くりこみ法」とよぶ手法を用いたプログラムを C++ 言語で書き、これを公開している¹。これは計算の簡単な式 (9) の最小二乗法を適用し、それによって生じる偏

差を理論的に予想し、それを差し引いて補正するという手続きを数回反復するものである。このような方法で最適解に到達できることが一般的に解析されている [8]。そこで、まず我々のプログラムによる解が実際に理論的な下界を達成していることを確認する。

図 1 は互いに射影変換で対応する格子パターンである。各頂点を特徴点とし、その x , y 座標に独立に期待値 0、標準偏差 σ (画素) の正規乱数を加えて射影変換行列を計算する。図 2 は横軸に σ にとり、平方平均二乗誤差 (式 (12) で期待値 $E[\cdot]$ をサンプル平均値 $\sum_{k=1}^n (\cdot)/n$ に置き換えたもの) を各 σ に対してプロットしたものである ($n = 50$)。実線は我々のプログラムによる最適計算であり、破線は式 (9) の最小二乗法、破線は式 (15) の理論的下界である。図 3 はその計算の平均計算時間をプロットしたものである。ただし、計算には Sun Ultra-1 (Solaris 2.5.1) ワークステーションを用いた。これから最小二乗法は精度が低く、我々のプログラムは精度が理論的下界に実際に到達し、もはや精度向上の余地がないことが確認できる。ただし、その代わりに最小二乗法に比べて約 25 倍の計算時間がかかっている²。

² 図 1 は文献 [10] と同じものであるが、平方平均二乗誤差の定義が文献 [10] とやや異なるので図 2 は文献 [10] の結果とやや異なっている。また図 3 は文献 [10] の結果より計算時間が短縮され、プログラムがより効率化されている。

¹ <http://www.ail.cs.gunma-u.ac.jp/Labo/research.html>

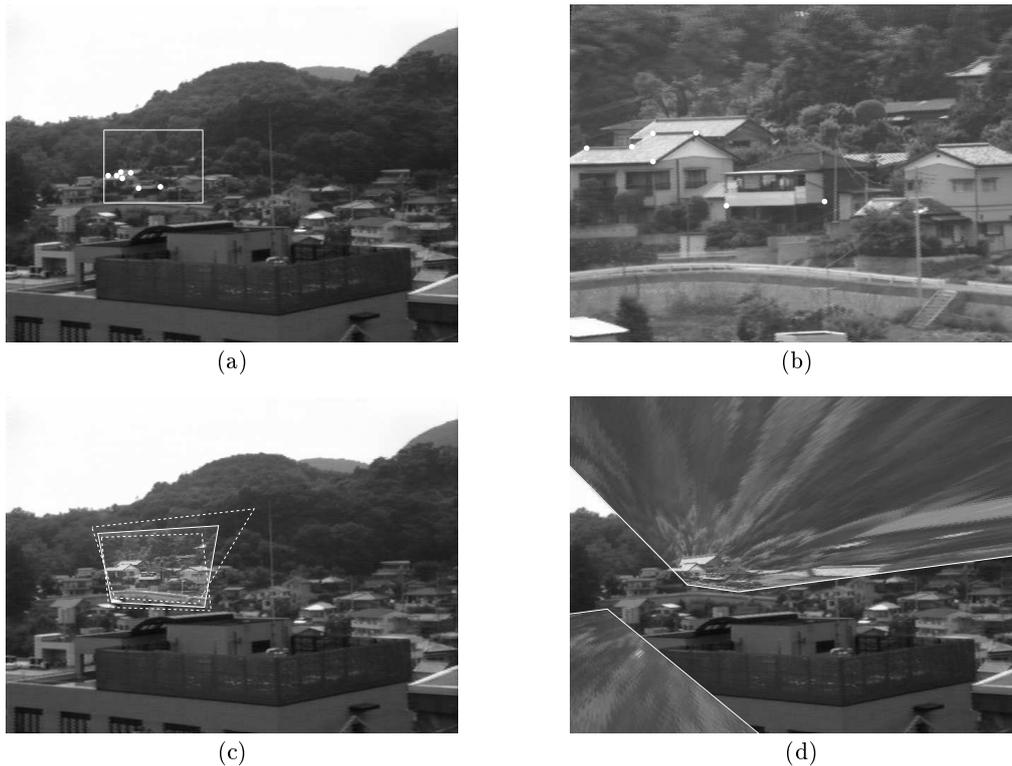


図 5: (a) 遠方の実画像シーン。(b) 図 (a) の白枠部分に対応するズーム画像。(c) 最適計算した射影変換による写像 (実線の枠内) と標準偏位による境界線 (破線)。(d) 最小二乗法による結果。

7. 射影変換によるモザイク生成

序論で指摘したように、射影変換は自由度が多いので H のわずかの誤差で変換に大きな歪みが生じる可能性がある。これは重なりが少ないモザイク生成において得られる特徴点が少ないときに顕著となるので、射影変換行列の計算の精度が重大な影響を与える。

図 4(a) は遠方のシーンである。図 4(b) はそのズーム画像であり、図 4(a) 中の白枠部分に対応している。図中にマークした特徴点を選び、射影変換行列を最適計算し、図 4(b) をそれによって図 4(a) 中に写像したものが図 4(c) である。実線が写像した画像の境界線であり、破線は標準偏位による写像の境界線である。図 4(d) は最小二乗法による結果であり (標準偏位は定義されない)、精度が低下していることがわかる。

図 5 は同じ画像で特徴点がより少ない場合である。図 5(c) の標準偏位は右上部分に不確定性が大きいことを示している。実際、図 5(d) の最小二乗法では右上部分が無限遠点に達し、左下に回り込んでいる。このように、わずかな誤差が大きな影響を与えることがわかる。

図 6(a), (b) は屋外シーンであり、図中にマークした特徴点を用いてモザイク生成したものが図 6(c) である。図 6(d) は最小二乗法による結果である。図 7 は同

じ画像で特徴点の取り方を変えた場合であり、わずかな誤差が画像の大きな変動をもたらしている。したがって射影変換行列の計算の精度が極めて重要であり、理論的に最適なアルゴリズムを使用することが必要である。我々のプログラムはそれに答えるものであると思われる。

8. まとめ

本論文では射影変換がモザイク生成を始めとする種々の画像・メディア処理において果たしている役割を述べ、画像の誤差の統計モデルに基き、射影変換の推定の精度の理論限界を示すと同時に、それを達成する最適計算の原理を示した。そして我々の公開しているプログラムによる解の精度が実際にその理論限界を達成する真に最適なアルゴリズムであることを実験的に確認した。また、互いに重なり小さい画像を用いたモザイク生成では射影変換が不安定であり、高精度の計算が要求されることを実画像例を用いて示すと同時に、それが我々のプログラムによってどのように改善されるかを示した。

我々のプログラムはモザイク生成のみならず、平面物体の3次元復元、航空写真や道路画像の解析、平面パターンを用いたカメラキャリブレーション等、射影変換が現れるのすべての画像・メディア処理問題に有効であり、幅広い応用が期待される。



(a)



(b)



(c)



(d)

図 6: (a), (b) 屋外シーン。(c) 最適計算した射影変換によるモザイク生成。(d) 最小二乗法による結果。

参考文献

- [1] A. Criminisi, I. Reid and A. Zisserman, Duality, rigidity and planar parallax, *Proc. 5th Euro. Conf. Comput. Vision*, June 1998, Freiburg, Germany, Vol. 2, pp. 846–861.
- [2] D. Q. Huynh, Calibration of a structured light system: A projective approach, *Proc. IEEE Conf. Comput. Vision Patt. Recog.*, June 1997, Puerto Rico, pp. 225–230.
- [3] M. Irani and P. Anandan, Parallax geometry of pairs of points for 3D scene analysis, *Proc. 4th Euro. Conf. Comput. Vision*, April 1996, Cambridge, U.K., Vol. 1, pp. 17–30.
- [4] M. Irani, P. Anandan and D. Weinshall, From reference frames to reference planes: Multi-view parallax geometry and applications, *Proc. 5th Euro. Conf. Comput. Vision*, June 1998, Freiburg, Germany, Vol. 2, pp. 829–845.
- [5] 星野准一, 増田功, 観測・推定画像間の誤差最小化によるパノラマ動画の生成, 電子情報通信学会論文誌, **J81-D-II-11** (1998), 2565–2575.
- [6] 星野准一, 増田功, 局所的な輝度変動にロバストなパノラマ画像生成法, 電子情報通信学会論文誌, **J82-D-II-2** (1999), 222–229.
- [7] 金谷健一, 「画像理解 — 3次元認識の数理 —」, 森北出版, 1990.
- [8] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam, 1996.
- [9] 金谷健一, 「形状CADと図形の数学」, 共立出版, 1998.
- [10] K. Kanatani, Optimal homography computation with a reliability measure, *Proc. IAPR Workshop on Machine Vision Applications*, November 1998, Makuhari, Japan, pp. 426–429.
- [11] K. Kanatani and S. Takeda, 3-D motion analysis of a planar surface by renormalization, *IEICE Trans. Inf. & Syst.* **E78-D-8** (1995), 1074–1079.
- [12] 蚊野浩, 金出武雄, 任意のカメラ配置におけるステレオ視とステレオカメラ校正, 電子情報通信学会論文誌 D-



(a)



(b)



(c)



(d)

図 7: (a), (b) 屋外シーン。(c) 最適計算した射影変換によるモザイク生成。(d) 最小二乗法による結果。

- II, **J79-D-II-11** (1996), 1810-1818.
- [13] 蚊野浩, 木村茂, 田中昌也, 金出武雄, ビデオレートステレオマシンにおけるカメラ幾何補正の実現, 日本ロボット学会誌, **16-4** (1998), 527-532.
- [14] 野口卓, 奥富正敏, ステレオ画像からの道路平面に対する射影変換行列の導出, 情報処理学会コンピュータビジョンとイメージメディア研究会, 97-CVIM-108-4 (1997), 23-30.
- [15] H. S. Sawhney, S. Hsu and R. Kumar, Robust video mosaicing through topology inference and local to global alignment, *Proc. 5th Euro. Conf. Comput. Vision*, June 1998, Freiburg, Germany, Vol. 2, pp. 103-119.
- [16] 清水慶行, 太田直哉, 金谷健一, 信頼性評価を備えた最適な射影変換の計算プログラム, 情報処理学会コンピュータビジョンとイメージメディア研究会, 98-CVIM-111-5 (1998-5), 33-40.
- [17] R. Szeliski and H.-U. Shum, Creating full view panoramic image mosaics and environment maps, *Proc. SIGGRAPH'97*, August 1997, Los Angeles, CA, U.S.A., pp. 251-258.
- [18] 谷口行信, 阿久津明人, 外村佳伸, PanoramaExcerpts: パノラマ画像の自動生成・レイアウトによる映像一覧, 電子情報通信学会誌, Vol. J82-D-II, No. 3 (1999), 390-398.
- [19] B. Triggs, Autocalibration from planar surfaces, *Proc. 5th Euro. Conf. Comput. Vision*, June 1998, Freiburg, Germany, Vol. 1, pp. 89-105.
- [20] T. Viéville, C. Zeller and L. Robert, Using collineations to compute motion and structure in an uncalibrated image sequences, *Int. J. Comput. Vision*, **20**(3), 1996, 213-242.
- [21] T. Werner, T. Pajdla and V. Hlaváč, Efficient 3-D scene visualization by image extrapolation, *Proc. 5th Euro. Conf. Comput. Vision*, June 1998, Freiburg, Germany, Vol. 2, pp. 382-396.
- [22] I. Zoghlami, O. Faugeras and R. Deriche, Using geometric corners to build a 2D mosaic from a set of images, *Proc. IEEE Conf. Comput. Vision Patt. Recog.*, June 1997, Puerto Rico, pp. 420-425.