

大域的な整合性を保証するロバストな画像の対応づけ

金澤 靖[†]

金谷 健一[‡]

[†] 豊橋技術科学大学知識情報工学系 [‡] 岡山大学工学部情報工学科

2 画像の特徴点を対応づける新しい手法を提案する．このとき対応の大域的な整合性をどう保証するかが最大の問題となる．従来は対応の局所的な相互関係を反復によって全体に波及させたり，大域的な最適解を組合せ探索していたが，提案方法は全体的にほぼ成立すべき「柔らかい」拘束条件を導入し，その充足の度合いを測る「確信値」をすべての対応候補に定義する．そして，それを「平均場近似」によって段階的に更新し，最終的に確信値の高い対応に RANSAC によって，厳密に満たすべき「硬い」エピ極線拘束条件を当てはめる．またモデル選択によって画像間の関係が射影変換とみなせるかどうかも判定する．これらを実画像を用いて実験し，本手法がカメラの回転やズーム変化にロバストであることを示す．

Robust Image Matching Preserving Global Consistency

Yasushi Kanazawa[†]

Kenichi Kanatani[‡]

[†]Department of Knowledge-based Information Engineering
Toyohashi University of Technology, Toyohashi, Aichi 441-8580 Japan

[‡]Department of Information Technology, Okayama University, Okayama 700-8530 Japan

We present a new method for detecting point matches between two images. The biggest issue is how to preserve the global consistency of individual matches. Many existing methods gradually propagate local smoothness by relaxation or do combinatorial search for an optimal solution. Our method imposes various non-local constraints that should be approximately satisfied as “soft” constraints by defining their “confidence” to all potential matches. The confidence is progressively updated by “mean-field approximation”. In the end, the “hard” epipolar constraint that should be exactly satisfied is strictly imposed by RANSAC. We also introduce a model selection procedure for testing if the image mapping can be regarded as a homography. Using real images, we demonstrate that our method is robust to camera rotations and zooming changes.

1. まえがき

画像間の対応を定めることはコンピュータビジョンの基本処理の一つである．これには連続ビデオ画像から隣接フレームごとに対応を追跡する場合と，異なる画像間の対応を直接的に探索する場合に分けられる．ここでは後者を考える．

基本となるのはテンプレートマッチングによる局所相関の探索である．すなわち，2 画像から特徴抽出フィルタ [3, 15] によって特徴点を抽出し，相関の高い 2 点を対応させればよい．しかし，これは画像間が並進のみの場合しか有効ではなく，視差があると対応する点の見え方が変化する．特にカメラの回転やズーム変化があると局所的な相関が著しく低下し，誤対応が増加する．このため何らかの拘束条件を導入して対応の妥当性を判定しなければならない．

シーンが平面あるいは遠景であれば，2 画像は「射影変換」[4] で結ばれるから，この拘束条件に最小メジアン法 (LMedS)[14] や RANSAC[2] などの投票を組合せてロバストな対応づけができる．著者らはテンプレートマッチングの残差を解析し [7]，投票によ

ってテンプレートを段階的に変形させる「段階的マッチング」[10] を提案し，大きな変形を伴う画像間もロバストな対応づけができることを示した．

しかし一般のシーンでは，成立すべき拘束条件は「エピ極線方程式」[4] しか存在しない．従来からこれを利用した投票法が提案されているが [1, 4, 17]，カメラが校正済みで「基礎行列」が既知の場合は効果的であっても，基礎行列が未知の場合はこれは非常に弱い拘束条件であり，多くの誤まった対応がこの条件を満たしてしまう．このため，画像の局所相関と基礎行列に対する投票のみでは互いに矛盾する極めて不自然に対応が多く含まれる．

これを解決するには，得られる対応が「全体的に自然である」という何らかの大域的整合性を保証しなければならない．しかし，これを拘束条件の形に記述するのが難しい．この問題に対して従来から，各特徴点に種々の属性を定義し，属性間の類似度を最大にする反復や組合せ探索が試みられている．しかし解は各行各列に 1 が 1 個以内の 0,1 から成る置換行列であり，困難な整数計画法となる．そこでこれを避けるために，置換行列の実数行列による近似 [12, 16]，テンソル投票 [11]，距離変換と多重解像度探索の組合せ [13]，グラフラベルの弛緩法 [17] などの反復手法が提案されている．しかしいずれも複雑な計算と多大な計算量を要する．またほとんどは局所的な相互

[†]441-8580 豊橋市天伯町雲雀ヶ丘 1-1, 豊橋技術科学大学知識情報工学系, Tel: (0532)44-6888, Fax: (0532)44-6873
E-mail: kanazawa@tutkie.tut.ac.jp

[‡]700-8530 岡山市津島中 3-1-1, 岡山大学工学部情報工学科, Tel/Fax: (086)251-8173
E-mail: kanatani@suri.it.okayama-u.ac.jp

関係を反復によって全体に波及させているので、空間的に離れた複数の特徴点に対する大域的な整合性を直接に反映させることができない。

本論文ではこのような最適化や組合せ探索を用いずに大域的な整合性を保証する対応づけアルゴリズムを提案する。その基本原理は、シーンの3次元形状はそれほど変則的ではなく対応関係はほぼ連続的であり、またシーンの多くの部分がほぼ平面的であると仮定して、このような大域的な性質の充足度によって整合性を評価することである。

このとき「ほぼ...である」という「柔らかい」拘束条件をどう扱うかが課題となる。本論文ではその充足度を測る「確信値」を全対応候補に定義し、この値の低い対応も最後まで保持する。そして、確信値の高い対応から大域的な性質を推定し、各対応との適合性を測って段階的に確信値を更新する。最後に確信値の高い対応にRANSACによって、厳密に満たすべき「硬い」エッジ線拘束条件を当てはめる。推定した大域的な性質と個々の対応との適合性を計算することは物理学の多体問題で用いる「平均場近似」と類似の考え方である。

本論文ではまた、モデル選択によって画像間の変換が射影変換とみなせるかどうかを判定する。そして実画像実験によって、本手法がカメラの回転やズーム変化にロバストであることを示す。

2. テンプレート マッチング

本論文では第1画像の特徴点 p と第2画像の特徴点 q の局所相関を次の残差(平方和)で測る。

$$J(p, q) = \sum_{(i, j) \in \mathcal{N}} |T_p(i, j) - T_q(i, j)|^2 \quad (1)$$

ここに $T_p(i, j)$, $T_q(i, j)$ はそれぞれ点 p, q を中心とする $w \times w$ テンプレート \mathcal{N} の輝度値である¹。

基本的な処理は次の通りである。特徴抽出フィルタ[3, 15]によって第1画像から N 個の特徴点 p_1, \dots, p_N を、第2画像から M 個の特徴点 q_1, \dots, q_M を独立に抽出し、全 NM 個の組合せの残差 $\{J(p_\alpha, q_\beta)\}$, $\alpha = 1, \dots, N, \beta = 1, \dots, M$ を $N \times M$ の表にまとめる。そして表中の最小値 $J(p_{\alpha^*}, q_{\beta^*})$ を探して点 $p_{\alpha^*}, q_{\beta^*}$ を対応させる。 $J(p_{\alpha^*}, q_{\beta^*})$ を含む行と列を削除した $(N-1) \times (M-1)$ の表に対して同様の処理を繰り返せば、最終的に $L = \min(N, M)$ 個の対応が得られる。この計算はあらかじめ全ての残差 $\{J(p_\alpha, q_\beta)\}$ をソートしておけば効率的に計算できる。以上の操作を「残差 J に関する一対一化」と呼ぶ。

しかしこれによって取り出された対応が必ず正しいとは限らず、また除外された対応に正しいものも含まれているかもしれない。そこで本論文ではすべ

¹実験では $w = 9$ とした。

での組合せに「確信値」を与えて、除外された対応でも確信値が高ければ次の段階で復活させる。そして確信値を更新してこの処理を反復する。

以下、対応の候補として第1画像の点 p と第2画像の点 q を選んだ組合せを (p, q) と略記する。

3. 局所相関に関する確信値

対応候補 (p, q) は残差 $J(p, q)$ が小さいほど正しい可能性が高いから、残差に関する確信値を次のように定義する。

$$P = e^{-sJ(p, q)} \quad (2)$$

これは対応の正しさの一つの尺度である。すべての場合の和を1に正規化すれば「確率」とみなせるが、本論文では上限が1となるように正規化したものを「確信値」と呼んで区別する。

式(2)は統計物理学ではGibbs分布と呼ばれ、 $s = 1/kT$ (k はボルツマン定数)と置いて T を温度と呼んでいる。 $s = 0$ (すなわち温度 $T = \infty$)なら残差 $J(p, q)$ (物理学ではエネルギー)に無関係に $P = 1$ となる。 s が増えれば(温度 T が下がれば)大きな残差(エネルギー)に対する確信値は急激に減衰する。

このとき減衰定数 s (すなわち温度 T)をどう定めるかが問題となる。本論文では、全 NM 個の組合せ $\{(p_\alpha, q_\beta)\}$ の内、高々 $L (= \min(N, M))$ 個の対応しか正しくないことから、式(2)を確率とみなした(すなわち和を1に正規化した)残差の期待値が残差の小さい L 個の平均値と等しくなるように s を決定する。これは全 NM 個の組合せの内、実質的には残差の小さい L 個程度を考えればよいということを表している。

全 NM 個の組合せの残差 $J(p_\alpha, q_\beta)$ を昇順にソートした λ 番目を残差を J_λ と書けば、この条件は次式のように書ける。

$$\sum_{\lambda=1}^{NM} J_\lambda \frac{e^{-sJ_\lambda}}{Z} = \bar{J} \quad (3)$$

ただし次のように置いた。

$$Z = \sum_{\lambda=1}^{NM} e^{-sJ_\lambda}, \quad \bar{J} = \frac{1}{L} \sum_{\lambda=1}^L J_\lambda \quad (4)$$

式(3)の解は、

$$\Phi(s) = \sum_{\lambda=1}^{NM} (J_\lambda - \bar{J}) e^{-sJ_\lambda} \quad (5)$$

と置くと式(3)が $\Phi(s) = 0$ と書けるから、この解をニュートン法で求めればよい。得られた s を用いて式(2)で定義した λ 番目の対応の確信値を $P_\lambda^{(0)}$ と書く。

4. 空間相関に関する確信値

2 画像を重ねて対応する点を結んだベクトル (始点は第 1 画像, 終点は第 2 画像とする) を (オプティカル) フローと呼ぶ. シーンの 3 次元形状が極端に変則的なものでなければ, 各フローの方向と大きさはある狭い範囲に分布していると考えられる. この分布を次のように推定する.

まず全 NM 個の組合せから

$$P_\lambda^{(0)} > e^{-k^2/2} \quad (6)$$

を満たすものを取り出し², $P_\lambda^{(0)}$ に関して一対一化したものを仮の対応とする. n_0 個の仮の対応が得られたとし, μ 番目の対応のフローを \vec{r}_μ とする. そしてフロー $\{\vec{r}_\mu\}$, $\mu = 1, \dots, n_0$ の確信値 $P_\mu^{(0)}$ を重みとする平均 \vec{r}_m と共分散行列 V を次のように計算する.

$$\begin{aligned} \vec{r}_m &= \sum_{\mu=1}^{n_0} \frac{P_\mu^{(0)}}{Z} \vec{r}_\mu, & Z &= \sum_{\mu=1}^{n_0} P_\mu^{(0)} \\ V &= \sum_{\lambda=1}^{n_0} \frac{P_\mu^{(0)}}{Z} (\vec{r}_\mu - \vec{r}_m)(\vec{r}_\mu - \vec{r}_m)^\top \end{aligned} \quad (7)$$

$P_\mu^{(0)}$ を重みとしているので, これらは実質的には残差が小さく正しい可能性の高い少数の対応によって定まる.

これらが求まれば, 全 NM 個の組合せに対して空間相関に関する確信値を次のように定義する.

$$P_\lambda^{(1)} = e^{-(\vec{r}_\lambda - \vec{r}_m, V^{-1}(\vec{r}_\lambda - \vec{r}_m))} \quad (8)$$

ただし, (\vec{a}, \vec{b}) はベクトル \vec{a} , \vec{b} の内積を表わす. これにより個々のフロー \vec{r}_λ は, 推定した平均フロー \vec{r}_m との差が推定した共分散行列 V の範囲内であれば高い確信値が与えられ, 大きさや方向が全体の傾向と著しくずれるものには低い値が与えられる.

このように大域的な相関を, 推定した場との相互作用によって評価する方法は統計物理学の多体問題に用いる「平均場近似」と類似の考え方である.

5. 大域的整合性に関する確信値

次に, シーンの大部分はほぼ平面的であるか, あるいはかなり遠方にあると仮定する. これは画像間の変換がおおまかには射影変換で近似できることを意味する. これに対する確信値を定義するために, 空間相関の場合と同様にして仮の対応を選び直す. そのために全 NM 個の組合せから

$$P_\lambda^{(0)} P_\lambda^{(1)} > e^{-2k^2/2} \quad (9)$$

²これは正規分布の k シグマ区間を選ぶことに対応する. 実験では $k = 3$ を用いた. こうすると確信値の定義から L 個程度の対応が選ばれる.

を満たすものを取り出し, $P_\lambda^{(0)} P_\lambda^{(1)}$ に関して一対一化したものを新たな仮の対応とする. n_1 個の仮の対応が得られたとし, μ 番目の対応が第 1 画像の点 (x_μ, y_μ) と第 2 画像の点 (x'_μ, y'_μ) を結ぶものであるとする ($\mu = 1, \dots, n_1$). これらを次の 3 次元ベクトルで表す.

$$\mathbf{x}_\mu = \begin{pmatrix} x_\mu/f_0 \\ y_\mu/f_0 \\ 1 \end{pmatrix}, \quad \mathbf{x}'_\mu = \begin{pmatrix} x'_\mu/f_0 \\ y'_\mu/f_0 \\ 1 \end{pmatrix} \quad (10)$$

ただし f_0 は任意の定数である³. もしシーンが平面または十分遠方であれば両者は次の形の射影変換で結ばれる [4, 5].

$$\mathbf{x}'_\mu = Z[\mathbf{H}\mathbf{x}_\mu] \quad (11)$$

\mathbf{H} は射影変換行列と呼ばれる正則行列であり, $Z[\cdot]$ はベクトルの第 3 成分を 1 とする正規化を表す.

シーンが平面でなく遠方にもなればこれは成立しないが, 仮の対応はそれに近いと仮定して射影変換で近似する. それには点 $\{\mathbf{x}_\mu\}$, $\{\mathbf{x}'_\mu\}$ を射影変換を満たす組 $\{\bar{\mathbf{x}}_\mu\}$, $\{\bar{\mathbf{x}}'_\mu\}$ で近似して, これを未知数とみなし, 確信値 $P_\mu^{(0)} P_\mu^{(1)}$ で重みづけた式

$$J = \sum_{\mu=1}^{n_1} P_\mu^{(0)} P_\mu^{(1)} (\|\mathbf{x}_\mu - \bar{\mathbf{x}}_\mu\|^2 + \|\mathbf{x}'_\mu - \bar{\mathbf{x}}'_\mu\|^2) \quad (12)$$

を制約条件

$$\bar{\mathbf{x}}'_\mu = Z[\mathbf{H}\bar{\mathbf{x}}_\mu], \quad \mu = 1, \dots, n_1 \quad (13)$$

のもとで \mathbf{H} , $\{\bar{\mathbf{x}}_\mu\}$, $\{\bar{\mathbf{x}}'_\mu\}$ について最小化すればよい. この解はくりこみ法⁴[8] を修正することにより簡単に計算できる. $P_\mu^{(0)} P_\mu^{(1)}$ を重みとしているので, 解は実質的には, 残差が小さく大きさや方向のそろった少数の対応によって定まる.

このようにして射影変換行列 \mathbf{H} が求まれば, 全 NM 個の組合せに対して, 推定した射影変換からのずれを次のように評価する.

$$D_\lambda^H = \|\mathbf{x}'_\lambda - Z[\mathbf{H}\mathbf{x}_\lambda]\|^2 \quad (14)$$

ただし \mathbf{x}_λ , \mathbf{x}'_λ は λ 番目の対応の 2 点を表わすベクトルである. これを用いて局所相関の場合と同様に, 大域的整合性に関する確信値を Gibbs 分布によって次のように定義する.

$$P_\lambda^{(2)} = e^{-tD_\lambda^H} \quad (15)$$

³実験では数値計算の安定化のためにほぼ画像サイズに等しい $f_0 = 600$ としている.

⁴以下に C++ プログラムが公開されている.
<http://www.suri.it.okayama-u.ac.jp/>

減衰定数 t は局所相関の場合と同様に，条件

$$\frac{1}{Z} \sum_{\lambda=1}^{NM} D_{\lambda}^H e^{-tD_{\lambda}^H} = \bar{D}^H \quad (16)$$

から定める．ただし次のように置いた．

$$Z = \sum_{\lambda=1}^{NM} e^{-tD_{\lambda}^H}, \quad \bar{D}^H = \frac{1}{L} \sum_{\lambda=1}^L D_{\lambda}^H \quad (17)$$

式(16)の解は，式(5)において J_{λ} を D_{λ}^H に置き換えてニュートン法で解けばよい．

これにより，例えば画像の上部では視差が小さく下部では大きいようなシーンでは上部の大きいフローや下部の小さいフローの確信値が抑えられ，大域的に整合性のとれたフローに高い確信値が与えられる．これも一種の平均場近似といえる．

6. エピ極線拘束条件の投票

最後にエピ極線拘束条件を厳密に満足する組合せを取り出す．そのためにまず空間相関や大域的整合性と同様にして，これまでに定まった確信値から仮の対応を選び直す．すなわち全 NM 個の組合せから

$$P_{\lambda}^{(0)} P_{\lambda}^{(1)} P_{\lambda}^{(2)} > e^{-3k^2/2} \quad (18)$$

を満足するものを取り出し，全確信値 $P_{\lambda}^{(0)} P_{\lambda}^{(1)} P_{\lambda}^{(2)}$ に関して一対一化したものを新たな仮の対応とする． n_2 個の仮の対応が得られたとし， μ 番目の対応を式(10)のベクトル $x_{\mu}, x'_{\mu}, \mu=1, \dots, n_2$ で表す．

$\{x_{\mu}, x'_{\mu}\}$ が正しい対応であれば次のエピ極線方程式が満たされる [4]．

$$(x_{\mu}, Fx'_{\mu}) = 0 \quad (19)$$

F は基礎行列と呼ばれるランク 2 の特異行列である [4]．上式は局所相関や空間相関や大域的整合性とは異なり，必ず満たされなければならない「硬い」拘束条件である．そこで RANSAC [2, 4] を用いてこれが満たされるもののみを取り出す．そのために初期値を $S_m = 0, F_m = O$ とし，次の操作を反復する．

1. n_2 個の仮の対応からランダムに 8 個を選ぶ．
2. 選んだ 8 個の対応から基礎行列 F を計算する．式(19)から基礎行列 F のスケールは不定であるから，8 個のエピ極線方程式から得られる F に関する連立 1 次方程式を解けばよい．
3. n_2 個の仮の対応の各々について，求めた F の定めるエピ極線拘束条件からのずれを次のように測る．

$$D_{\mu}^F = \frac{(x_{\mu}, Fx'_{\mu})^2}{\|P_k F^T x_{\mu}\|^2 + \|P_k F x'_{\mu}\|^2} \quad (20)$$

ただし $P_k = \text{diag}(1, 1, 0)$ と置いた．上式の f_0^2 倍は点 x_{μ}, x'_{μ} からそれらを通るべきエピ極線までの (高次の項を無視した) 距離の二乗の和に等しい [4, 5]．

4. 次式を満す対応を取り出し，それらの全確信値 $P_{\mu}^{(0)} P_{\mu}^{(1)} P_{\mu}^{(2)}$ の和を S とする．

$$D_{\mu}^F \leq \frac{2d^2}{f_0^2} \quad (21)$$

ただし d (単位は画素) は要求精度を指定する定数である⁵．

5. $S > S_m$ なら $S_m \leftarrow S, F_m \leftarrow F$ と更新する．

これを収束するまで反復すると⁶，全確信値の和 S_m が最大になる基礎行列 F_m が求まる．これを用いて全 NM 個の組合せに対して F_m の定めるエピ極線拘束条件からのずれ D_{λ}^F を式(20)と同じ形で評価し，式(21)を満すものを選ぶ．それらの中から式(18)を満すものを取り出し， $P_{\lambda}^{(0)} P_{\lambda}^{(1)} P_{\lambda}^{(2)}$ に関して一対一化したものを最終的な対応とする．

7. 画像の変換のモデル選択

先に述べたように，シーンが厳密に平面あるいは十分遠方であれば 2 画像は射影変換で結ばれる．このとき基礎行列の計算は退化し，シーンの 3 次元形状は復元できない [4, 5]．反面，2 画像が射影変換で結ばれればすべての画素の対応が自動的に定まる．このため画像間の関係が射影変換かどうかを知ることは実際問題で非常に重要となる．

素朴な方法は射影変換(11)に対する残差とエピ極線方程式(19)に対する残差を比較することであるが，射影変換はエピ極線方程式より強い拘束条件であり，射影変換を満す対応はエピ極線方程式も満すので，射影変換に対する残差はエピ極線方程式に対する残差より必ず大きくなる．したがって残差と拘束条件の強さのバランスを考慮する必要がある．これを評価するのがモデル選択であり，幾何学的当てはめに対する代表的な規準が幾何学的 AIC[6]である．

最終対応が n 個得られたとき，それらに射影変換行列 H と基礎行列 F を最適に当てはめた残差をそれぞれ J^H, J^F とすると，それぞれに対する幾何学的 AIC は次のようになる [5]．

$$\begin{aligned} \text{G-AIC}^H &= J^H + 2(2n + 8)\epsilon^2 \\ \text{G-AIC}^F &= J^F + 2(3n + 7)\epsilon^2 \end{aligned} \quad (22)$$

ただし ϵ は各特徴点の位置の誤差の大きさを表す定数であり，エピ極線方程式に対する残差から次のよ

⁵実験では $d = 3$ を用いた．

⁶実験では 100 回連続して更新が起こらないことを収束条件とした．

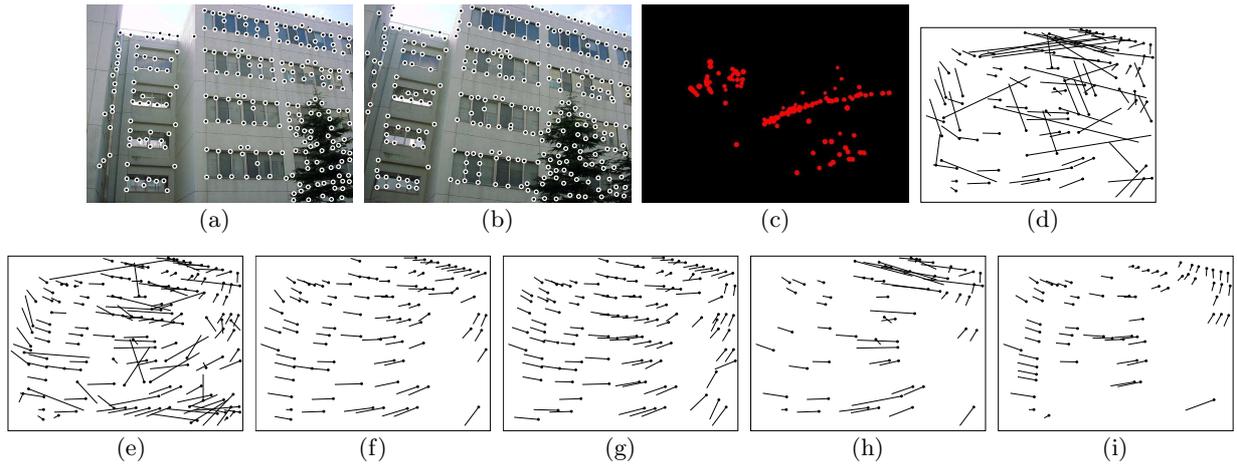


図 1: (a), (b) 入力画像と抽出した特徴点 . (c) 最終的な対応から復元した 3 次元形状 (上から見た図) . (d) 局所相関による対応 . (e) 空間相関を考慮した対応 . (f) 大域的整合性を考慮した対応 . (g) 最終的な対応 . (h) 初期対応 (d) から RANSAC で求めた対応 . (i) Zhang らの方法 [17] による対応 .

うに推定できる [5] .

$$\hat{\epsilon}^2 = \frac{J^F}{n-7} \quad (23)$$

そして $G-AIC^H < G-AIC^F$ のとき画像間の関係は射影変換であるとみなす .

8. 実画像実験

図 1(a), (b) は建物を撮影した 2 つの画像から Harris 作用素 [3] でそれぞれ独立に抽出した 300 個の特徴点をマークしたものである . 図 1(d) は局所相関のみによる対応 (空間相関を推定する仮の対応) のオプティカルフローを示す . この画像には周期的なパターンが多く、局所相関のみでは多くの誤対応が残っている .

図 1(e) は空間相関を考慮した対応 (大域的整合性を推定する仮の対応) であり、図 1(f) は大域的整合性を考慮した対応 (エピ極線拘束条件の投票のための仮の対応) である . 図 1(g) はさらに RANSAC を適用して得られた最終対応である . 拘束条件を追加することによって次第に精度が向上している .

比較のため、図 1(d) の初期対応に直接に RANSAC を適用して得られた対応を図 1(h) に示す . 図 1(g) と比べて誤った対応がかなり残されている . 図 1(i) は Zhang らの方法 [17] による結果である⁷ . 本手法に比べて得られる対応の数が少ない .

本手法では初期に多くの誤対応が存在しても、次の段階でそれらの確信値が低下し、逆に棄却された正しい対応の確信値が増大し、確信値の順序が次第に入れ替わって最終的にほとんど正しい対応が得られている .

この画像対では射影変換とエピ極線方程式に対する幾何学的 AIC は

$$G-AIC^F = 0.0039, \quad G-AIC^H = 0.0231$$

となり、画像の変換は射影変換とはみなせない . 求めた基礎行列から文献 [9] の方法で 3 次元復元を行ない、それを真上から見た形状を図 1(c) に示す .

図 2(a), (b) は石柱の基礎部分を撮影した 2 画像から抽出した特徴点である . 図 2(d), (e), (f), (g) はそれぞれ局所相関のみから定めた対応、空間相関を考慮した対応、大域的整合性を考慮した対応、最終的な対応である . この画像は大部分がほぼ様なテクスチャからなり、初期対応にはほとんど画像サイズに等しい長いフローが多く存在する . しかし空間相関を考慮した段階でそのような長いフローは取り除かれている .

図 2(h) は図 2(d) の初期対応に直接に RANSAC を適用して得られた対応であり、長いフローが残されている (これらもエピ極線方程式を満たす) . 図 2(i) は Zhang らの方法 [17] による結果であり、得られる対応数が少ない .

この画像対に対する幾何学的 AIC はそれぞれ

$$G-AIC^F = 0.0041, \quad G-AIC^H = 0.0101$$

であり、図 1 の例と同様に画像間の変換は射影変換とみなせない . 求めた基礎行列から 3 次元復元を行なって真上から見た形状を図 2(c) に示す .

図 3(a), (b) は遠景を撮影した 2 画像から抽出した特徴点である . 図 3(d), (e), (f), (g) はそれぞれ局所相関のみから定めた対応、空間相関を考慮した対応、大域的整合性を考慮した対応、最終的な対応である . この場合は大域的整合性を考慮した段階でほぼ正しい対応が得られている .

⁷<http://www-sop.inria.fr/robotvis/personnel/zhang/software.html> のプログラムを使用した .

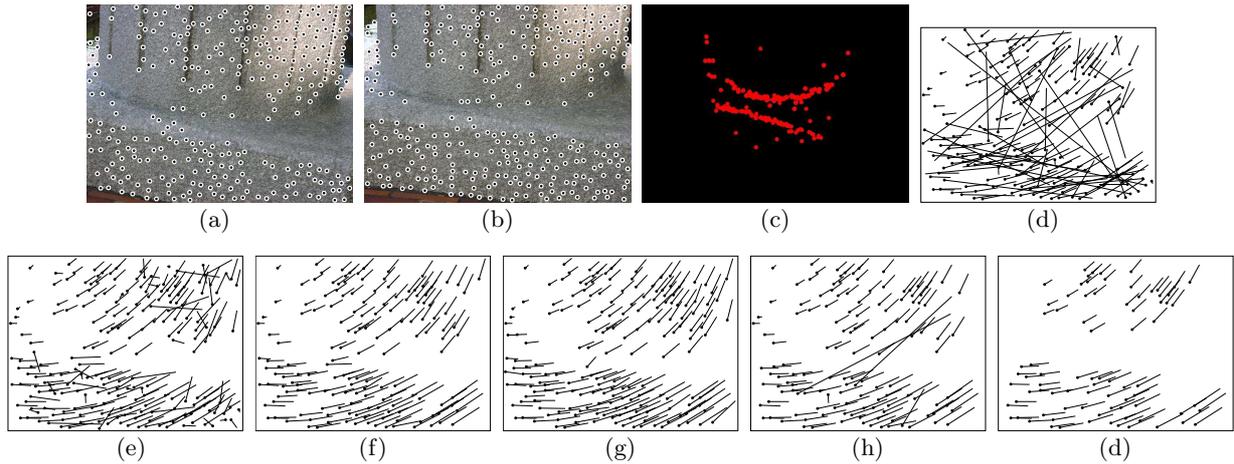


図 2: (a), (b) 入力画像と抽出した特徴点 . (c) 最終的な対応から復元した 3 次元形状 (上から見た図) . (d) 局所相関による対応 . (e) 空間相関を考慮した対応 . (f) 大域的整合性を考慮した対応 . (g) 最終的な対応 . (h) 初期対応 (d) から RANSAC で求めた対応 . (i) Zhang らの方法 [17] による対応 .

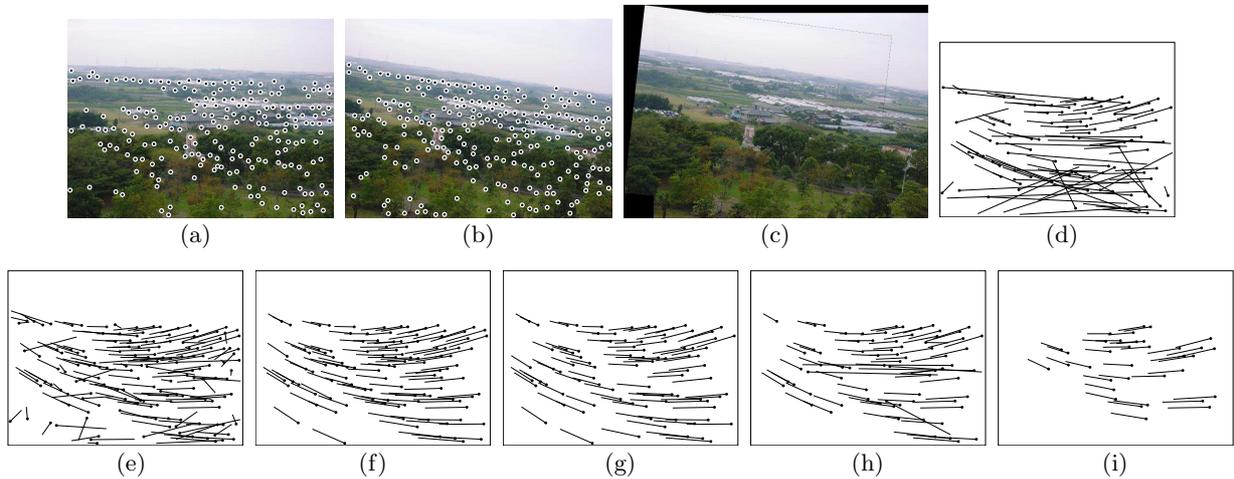


図 3: (a), (b) 入力画像と抽出した特徴点 . (c) 最終的な対応から生成したパノラマ画像 . (d) 局所相関による対応 . (e) 空間相関を考慮した対応 . (f) 大域的整合性を考慮した対応 . (g) 最終的な対応 . (h) 初期対応 (d) から RANSAC で求めた対応 . (i) Zhang らの方法 [17] による対応 .

図 3(h) は図 3(d) の初期対応に直接に RANSAC を適用して得られた対応である . 撮影過程でカメラが多少回転し , またシーンに類似したテクスチャが多いため , 直接的な RANSAC では誤対応を十分取り除けていない . 図 3(i) は Zhang らの方法 [17] による結果であり , やはり得られる対応数が少ない .

この画像対に対する幾何学的 AIC はそれぞれ

$$G\text{-AIC}^F = 0.0092, \quad G\text{-AIC}^H = 0.0088$$

であり , 画像間の変換が射影変換であると判定された . 図 3(c) に求めた射影変換行列から生成したパノラマ画像を示す .

次に , カメラの回転やズーム変化に対して本手法がどの程度ロバストかを確かめる実験を行なった .

図 4(a), (b) が用いた入力画像である . 画像内の大

部分がよく似たテクスチャの周期的パターンであり , 局所相関のみで対応を決定することは非常に難しい . 図 4(c), (d) はそれぞれ図 4(b) のシーンに対してカメラを 5 度および 10 度だけ回転して得られた画像である . その下の図の上段は図 4(a), (b) 用いた対応 , 中段は図 4(a), (c) を用いた対応 , 下段は図 4(a), (d) を用いた対応を示す . それぞれ (e)~(i) は図 1, 図 3 の (d)~(h) に対応する結果である .

図 5(a), (b) は別の入力画像であり , これも局所相関のみで対応を決定することは困難である . 図 5(c), (d) はそれぞれ図 5(b) のシーンを 80% および 65% にズームした画像である . その下の図は得られた結果を図 4 と同様に配置したものである .

いずれの場合も , カメラの回転やズームの変化が大きいと直接に RANSAC を適用しても正しい対応

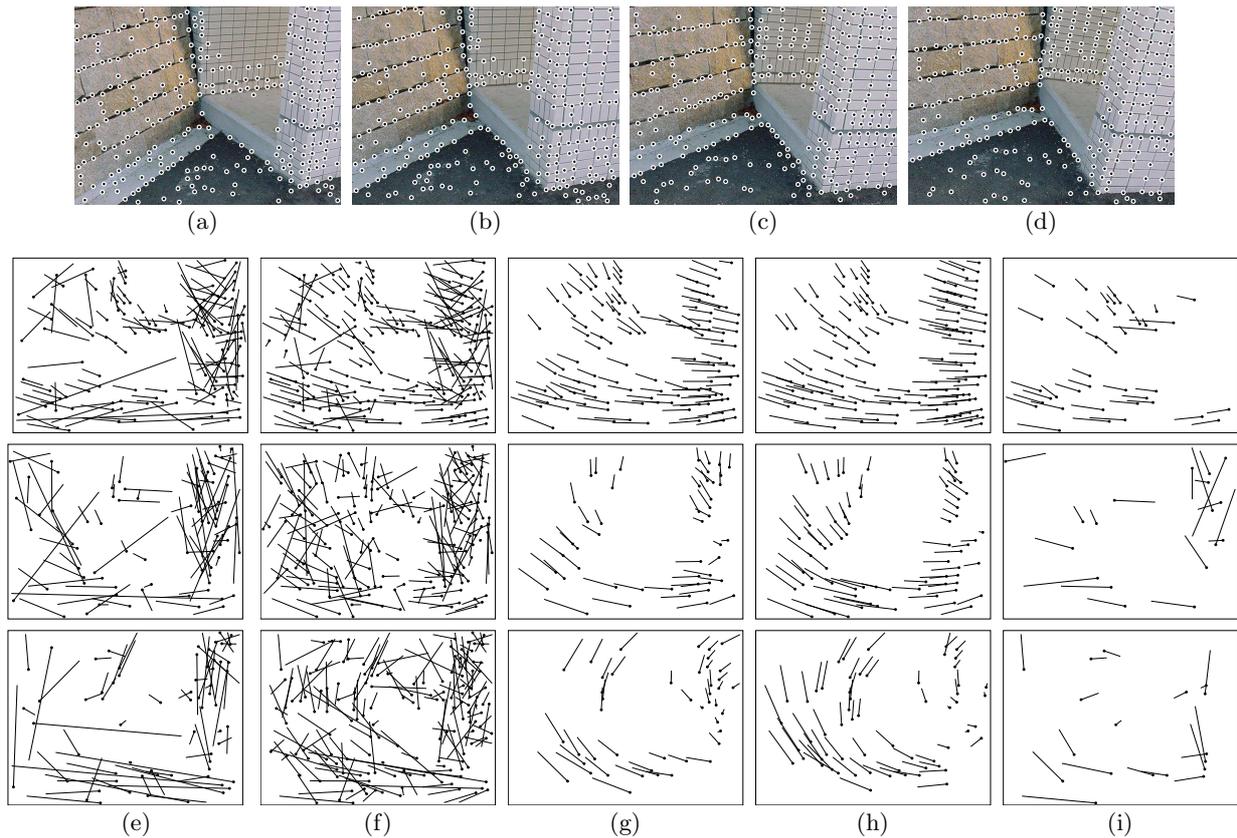


図 4: 画像の回転の影響: (a) 左画像 . (b) 右画像 . (c) 右画像 (b) を 5° 回転した画像 . (d) 右画像 (b) を 10° 回転した画像 . 以下, 上段は (a) と (b) を用いた結果, 中段は (a) と (c) を用いた結果, 下段は (a) と (d) を用いた結果: (e) 局所相関による対応 . (f) 空間相関を考慮した対応 . (g) 大域的整合性を考慮した対応 . (h) 最終的な対応 . (i) 初期対応 (e) から RANSAC で求めた対応 .

がほとんど得られない . しかし本手法では得られる対応数は減るものの, 正しい対応が得られている .

用いた計算機は Pentium III 700MHz, 主メモリ 768MB であり, OS には Debian GNU/Linux を用いた . 計算時間は上記の実験例の平均で約 23 秒かった . 直接に RANSAC を行うと約 14 秒で済むが, 得られる対応の精度とロバスト性を考慮すれば計算コストは妥当であると言える .

9. まとめ

本論文で提案する対応づけアルゴリズムは, ほぼ満たすべき「柔らかい」拘束条件を段階的に課すことによって対応の大域的整合性を保証するものである . まずテンプレートマッチングの残差から出発して, その確信値をすべての対応候補に対して計算する . 次に確信値の高いものを一対一化して仮の対応を選び, それから全体的傾向を推定する . そして, その推定した傾向からすべての対応候補の確信値を計算し直し, 次の段階に進む . 最後に RANSAC によって, 厳密に満たすべき「硬い」エッジ線方程式を当てはめる .

確信値は正規分布や Gibbs 分布を用いて上限を 1 に

正規化し, 分散は正しい対応数の上限 $L = \min(N, M)$ を介して同じしきい値でほぼ同じ対応数が得られるように設定している . このため異なる拘束条件も同一の基準で比較できる .

従来は隣接する対応の滑らかさなどの局所的な相互関係を反復によって徐々に全体に波及させる方法が主であったが, 本手法では仮の対応から推定した全体の傾向に基づいて確信値を定めるので, 大域的な性質を直接に反映させることができる . これは物理学の多体問題で用いる「平均場近似」と類似の考え方である .

本手法で得られる対応はすべてが正しいとは保証されないが, それぞれの対応にその確信値が評価されているので, これを以後の統計処理や判定条件に反映させることができる .

また本論文ではモデル選択を導入して, シーンが平面または遠景かどうかを判定する方法を示した . そして, 実画像実験によって本手法がカメラの回転やズーム変化にロバストであることを確認した .

謝辞: 本研究の一部は文部科学省科学研究費基盤研究 C(2) (No. 13680432), テレコム先端技術センター, 栢森情報科学財団の助成による .

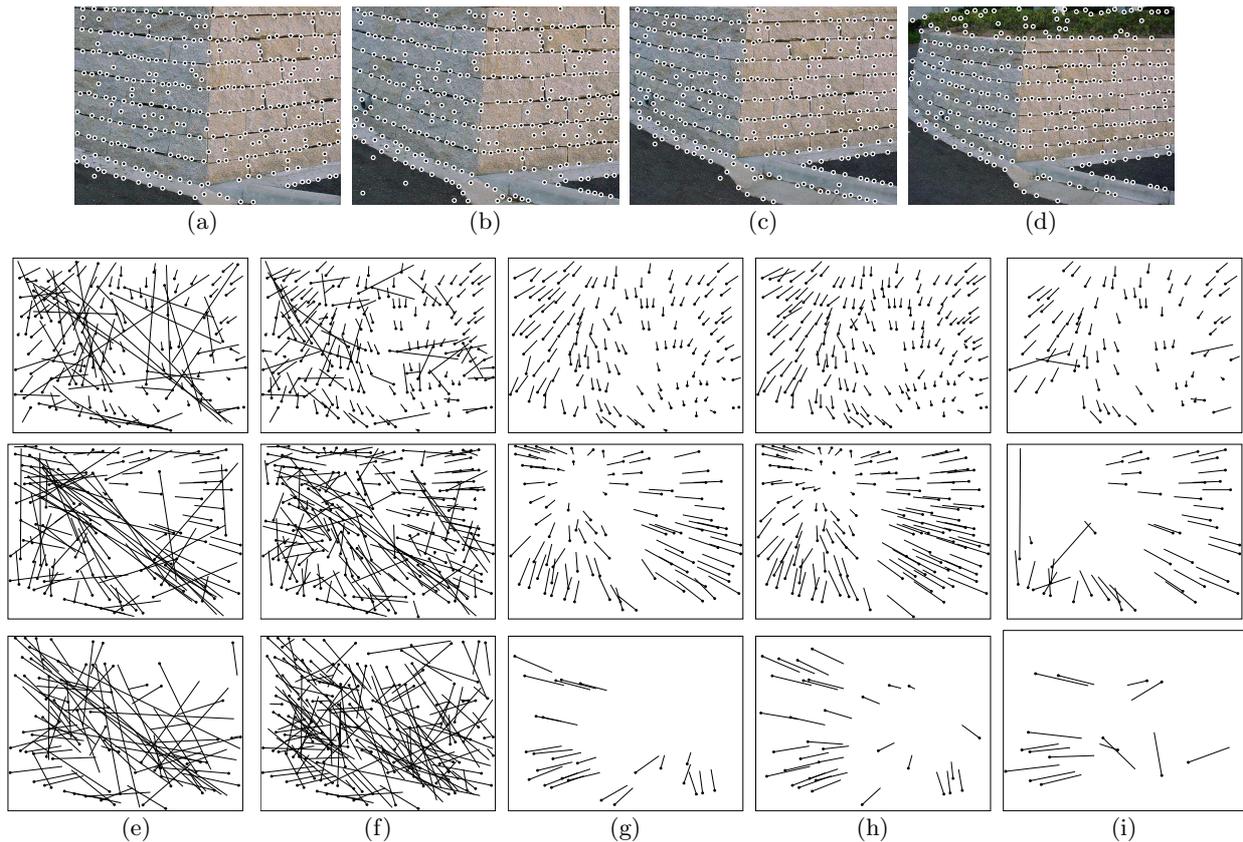


図 5: ズーム変化の効果: (a) 左画像. (b) 右画像. (c) 右画像 (b) を 80%縮小した画像. (d) 右画像 (b) を 65%縮小した画像. 以下, 上段は (a) と (b) を用いた結果, 中段は (a) と (c) を用いた結果, 下段は (a) と (d) を用いた結果: (e) 局所相関による対応. (f) 空間相関を考慮した対応. (g) 大域的整合性を考慮した対応. (h) 最終的に得られた対応. (i) 初期対応 (e) から RANSAC で求めた対応.

参考文献

- [1] P. Beardsley, P. Torr and A. Zisserman, 3D model acquisition from extended image sequences, *Proc. 4th Euro. Conf. Comput. Vision*, April 1996, Cambridge, U.K., Vol. 2, pp. 683–695.
- [2] M. A. Fischler and R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Comm. ACM*, **24-6** (1981), 381–395.
- [3] C. Harris and M. Stephens, A combined corner and edge detector, *Proc. 4th Alvey Vision Conf.*, August 1988, Manchester, U.K., pp. 147–151.
- [4] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U.K., 2000.
- [5] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam, the Netherlands, 1996.
- [6] 金谷健一, 情報量基準による幾何学的モデル選択, 情報処理学会論文誌, **37-6** (1996), 1073–1080.
- [7] 金谷健一, 金澤靖, テンプレートマッチングによる対応探索の自動しきい値設定法, 情報処理学会研究報告 2002-CVIM-132-4 (2002-3), 23–30.
- [8] K. Kanatani, N. Ohta and Y. Kanazawa, Optimal homography computation with a reliability measure, *IEICE Trans. Inf. & Sys.*, **E83-D-7** (2000), 1369–1374.
- [9] 金谷健一, 三島等, 未校正カメラによる 2 画像からの 3 次元復元とその信頼性評価, 情報処理学会論文誌: コンピュータビジョンとイメージメディア, **42-SIG 6** (2001), 1–8.
- [10] 金澤靖, 金谷健一, 段階的マッチングによる画像モザイク生成, 情報処理学会研究報告 2002-CVIM-132-5 (2002-3), 31–38.
- [11] M.-S. Lee, G. Medioni and P. Mordohai, Inference of segmented overlapping surfaces from binocular stereo, *IEEE Trans. Patt. Anal. Mach. Intell.*, **24-6** (2002), 824–837.
- [12] J. Maciel and J. Costeira, Robust point correspondence by concave minimization, *Image Vision Comput.*, **20-9/10** (2002-8), 683–690.
- [13] C. F. Olson, Maximum-likelihood image matching, *IEEE Trans. Patt. Anal. Mach. Intell.*, **24-6** (2002), 853–857.
- [14] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, Wiley, New York, 1987.
- [15] S. M. Smith and J. M. Brady, SUSAN—A new approach to low level image processing, *Int. J. Comput. Vision*, **23-1** (1997), 45–78.
- [16] M. A. van Wyk, T. S. Durrani and B. J. van Wyk, A RKHS interpolator-based graph matching algorithm, *IEEE Trans. Patt. Anal. Mach. Intell.*, **24-7** (2002), 988–995.
- [17] Z. Zhang, R. Deriche, O. Faugeras and Q.-T. Luong, A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry, *Artif. Intell.*, **78** (1995), 87–119.