Vol. 44 No. SIG 17(CVIM 8) 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 Dec. 2003

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††

2 画像の特徴点を対応づける新しい手法を提案する．最大の問題は対応の大域的な整合性をどう保証するかである．従来は局所的な関係を反復によって全体に波及させたり，大域的な最適解の組合せ探索を行ったりしていたが，提案方法で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 RANSAC によって厳密に満たすべき「硬い」エピ極線拘束条件をあてはめる．またモデル選択によって画像間の関係が射影変換と見なせるかどうかを判定する．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† and Kenichi Kanatani††

We present a new method for detecting point matches between two images. The main issue is how to preserve the global consistency of individual matches. Existing methods propagate local smoothness by relaxation or do combinatorial search for an optimal solution. Our method imposes non-local constraints that should be approximately satisfied across the image. We define the “confidence” of such “soft constraints” to all potential matches. The confidence is progressively updated by “mean-field approximation”. Finally, the “hard” epipolar constraint is 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１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[14)]（LMedS）や RANSAC[2)] などの投票を組み合わ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，大きな変形をともな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．これはカメラが校正済みで「基礎行列」が既知の場合に有効であっても，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，多くの誤った対応がこの条件を満たしてしまう．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)]，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不自然な対応も多く含まれてしまう．

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

† 豊橋技術科学大学知識情報工学系
Department of Knowledge-based Information Engineering, Toyohashi University of Technology
†† 岡山大学工学部情報工学科
Department of Information Technology, Okayama University

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．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が，複雑な計算と多大な計算量を要する．またこれらは局所的な関係を反復によって全体に波及させているので，空間的に離れた特徴点に対する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ないで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．これは，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，シーンの多くの部分がほぼ平面的であると仮定し，これらの条件の充足度によって対応の整合性を評価するものである．

このとき「ほぼ$\cdots$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RANSACによって，厳密に満たすべき「硬い」エピ極線拘束条件をあてはめる．

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見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．

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

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 $\mathcal{N}$ の輝度値であり☆，$\sum_{(i,j)\in\mathcal{N}}T_p(i,j)^2=\sum_{(i,j)\in\mathcal{N}}T_q(i,j)^2=1$ と正規化すれば**正規化相関**を用いることと等しい．

基本的な処理は次のとおりである．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 $p_1,\ldots,p_N$ を，第2画像から M 個の特徴点 $q_1,\ldots,q_M$ を独立に抽出し，全 NM 個の組合せの残差 $\{J(p_\alpha,q_\beta)\}$，$\alpha=1,\ldots,N$，$\beta=1,\ldots,M$ 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_{α^*}，q_{β^*} 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すべての残差 $\{J(p_\alpha,q_\beta)\}$ 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**1対1化**」と呼ぶ．

しかし，これによって取り出した対応が正しいとは限らず，除外した対応に正しいものが含まれているかもしれない．そこですべての組合せに「確信値」を与えて，除外し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．

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

これは対応の正しさの1つの尺度であり，すべての場合の和を1に正規化すれば「確率」と見なせるが，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．$s=0$（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（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 $\{(p_\alpha,q_\beta)\}$ のうち，たかだか L（$=\min(N,M)$）個の対応しか正しくないことから，式(2)を確率と見なした（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 を決める．これは全 NM 個の組合せのうち，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意味である．

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．

☆ 実験では $w=9$ とした．

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

ただし次のように置いた．

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

式 (3) の解は，

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

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る．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

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

を満たすものを取り出し☆，$P_\lambda^{(0)}$ に関して 1 対 1 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 $\vec{r}_\mu$ とする．そしてフロー $\{\vec{r}_\mu\}$，$\mu = 1, \ldots, n_0$ 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．

$$\vec{r}_m = \sum_{\mu=1}^{n_0} \frac{P_\mu^{(0)}}{Z} \vec{r}_\mu, \qquad Z = \sum_{\mu=1}^{n_0} P_\mu^{(0)}$$

$$V = \sum_{\mu=1}^{n_0} \frac{P_\mu^{(0)}}{Z} (\vec{r}_\mu - \vec{r}_m)(\vec{r}_\mu - \vec{r}_m)^\top \tag{7}$$

$P_\mu^{(0)}$ 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．

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

ただし，$(\vec{a}, \vec{b})$ はベクトル $\vec{a}$，$\vec{b}$ の内積を表す．これにより個々のフロー $\vec{r}_\lambda$ 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 の範囲内に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

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

を満たすものを取り出し，$P_\lambda^{(0)} P_\lambda^{(1)}$ に関して 1 対 1 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) を結ぶとする（$\mu = 1, \ldots, n_1$）．これらを次の 3 次元ベクトルで表す．

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

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

$$\boldsymbol{x}'_\mu = Z[\boldsymbol{H}\boldsymbol{x}_\mu] \tag{11}$$

$\boldsymbol{H}$ 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け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 $\{\boldsymbol{x}_\mu\}$，$\{\boldsymbol{x}'_\mu\}$ を射影変換を満たす組 $\{\bar{\boldsymbol{x}}_\mu\}$，$\{\bar{\boldsymbol{x}}'_\mu\}$ で近似して，これを未知数と見なし，確信値 $P_\mu^{(0)} P_\mu^{(1)}$ で重みづけた式

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

を制約条件

$$\bar{\boldsymbol{x}}'_\mu = Z[\boldsymbol{H}\bar{\boldsymbol{x}}_\mu], \quad \mu = 1, ..., n_1 \tag{13}$$

のもとで $\boldsymbol{H}$，$\{\bar{\boldsymbol{x}}_\mu\}$，$\{\bar{\boldsymbol{x}}'_\mu\}$ について最小化すればよい．この解は**くりこみ法**☆☆☆[8)] を修正することにより

☆ 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．

☆☆ 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．

☆☆☆ 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/

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 $\boldsymbol{H}$ 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．

$$D_\lambda^H = \|\boldsymbol{x}'_\lambda - Z[\boldsymbol{H}\boldsymbol{x}_\lambda]\|^2 \tag{14}$$

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次のように定義する．

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

減衰定数 t は局所相関の場合と同様に，条件

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

から定める．ただし次のように置いた．

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

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，たと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．まず空間相関や大域的整合性の場合と同様に，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

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

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１対１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 $\boldsymbol{x}_\mu$，$\boldsymbol{x}'_\mu$，$\mu = 1, \ldots, n_2$ で表す．

$\{\boldsymbol{x}_\mu, \boldsymbol{x}'_\mu\}$ 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]．

$$(\boldsymbol{x}_\mu, \boldsymbol{F}\boldsymbol{x}'_\mu) = 0 \tag{19}$$

$\boldsymbol{F}$ は**基礎行列**と呼ばれるランク 2 の特異行列である[4]．式 (19) は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 $S_m = 0$，$\boldsymbol{F}_m = \boldsymbol{O}$ 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 $\boldsymbol{F}$ を計算する．式 (19) より基礎行列 $\boldsymbol{F}$ 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 $\boldsymbol{F}$ 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 $\boldsymbol{F}$ 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（$\boldsymbol{P}_{\mathbf{k}} = \mathrm{diag}(1,1,0)$ と置く）．

$$D_\mu^F = \frac{(\boldsymbol{x}_\mu, \boldsymbol{F}\boldsymbol{x}'_\mu)^2}{\|\boldsymbol{P}_{\mathbf{k}}\boldsymbol{F}^\top\boldsymbol{x}_\mu\|^2 + \|\boldsymbol{P}_{\mathbf{k}}\boldsymbol{F}\boldsymbol{x}'_\mu\|^2} \tag{20}$$

上式の f_0^2 倍は点 $\boldsymbol{x}_\mu$，$\boldsymbol{x}'_\mu$ 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) 次式を満たす対応を取り出し，それらの全確信値 $P_\mu^{(0)}P_\mu^{(1)}P_\mu^{(2)}$ の和を S とする．

$$D_\mu^F \leq \frac{2d^2}{f_0^2} \tag{21}$$

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) $S > S_m$ なら $S_m \leftarrow S$，$\boldsymbol{F}_m \leftarrow \boldsymbol{F}$ と更新する．

これを収束するまで反復すると☆☆，全確信値の和 S_m が最大になる基礎行列 $\boldsymbol{F}_m$ が求まる．これを用いて全 NM 個の組合せに対して $\boldsymbol{F}_m$ 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) を満たすものを選ぶ．それらの中から式 (18) を満たすものを取り出し，$P_\lambda^{(0)}P_\lambda^{(1)}P_\lambda^{(2)}$ に関して１対１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にあれば 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，射影変換を満たす対応はエピ極線方程式も満たすので，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

☆ 実験では $d = 3$ を用いた．

☆☆ 実験では 100 回連続して更新されないことを収束条件とした．

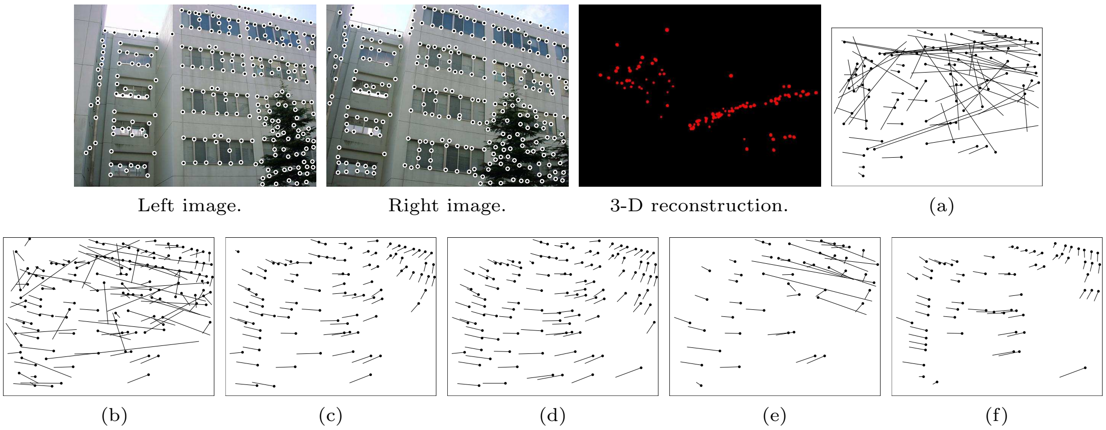

図 1 上段：入力画像と復元した 3 次元形状（上から見た図）．(a) 局所相関による対応．下段：(b) 空間相関を考慮した対応．(c) 大域的整合性を考慮した対応．(d) 最終的な対応．(e) 初期対応 (a) から RANSAC で求めた対応．(f) Zhang らの方法による対応

Fig. 1 Upper row: Input images and 3-D reconstruction. (a) Initial matches based on local correlations. Bottom row: (b) Matches with spatial consistency incorporated. (c) Matches with global smoothness added. (d) Final matches with the epipolar constraint imposed. (e) Matches obtained from (a) by direct RANSAC. (f) The method of Zhang et al.

のが**幾何学的モデル選択**であり，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 $\boldsymbol{H}$ と基礎行列 $\boldsymbol{F}$ を最適にあ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]．

$$\text{G-AIC}^H = J^H + 2(2n+8)\epsilon^2$$
$$\text{G-AIC}^F = J^F + 2(3n+7)\epsilon^2 \qquad (22)$$

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うに推定できる[5]．

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

そして $\text{G-AIC}^H < \text{G-AIC}^F$ のとき画像間の関係は射影変換であると見なす．

8. 実画像実験

図 1 の上段に 2 つの実画像と，それぞれから Harris 作用素[3] で抽出した 300 個の特徴点を示す．図 1(a) 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である．この例では正規化相関を用いた．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る．図 1(b) 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(c) 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(d) 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(a) の初期対応に直接に RANSAC を適用して得られた対応を図 1(e) に示す．図 1(d) と比べて誤った対応がかなり残されている．図 1(f) は Zhang ら[17] の方法による結果である☆．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

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

となり，画像の変換は射影変換とは見なせない．求めた基礎行列から金谷・三島[9] の方法で 3 次元復元を行い，それを真上から見た形状を図 1 の上段右に示す．

図 2(a)，(b) は遠景の例であり，図 2(a)～(f) がそれぞれ図 1(a)～(f) に対応する結果である．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．また，撮影過程でカメラが多少回転し，さらにシーンに類似したテクスチャが多いため，図 2(e) のように直接に RANSAC を適用しても誤対応を十分取り除けない．Zhang らの方法でもやはり図 2(f) のように得られる対応数が少ない．この画像対に対する

☆ `http://www-sop.inria.fr/robotvis/personnel/zzhang/softwares.html` のプログラムを使用した．

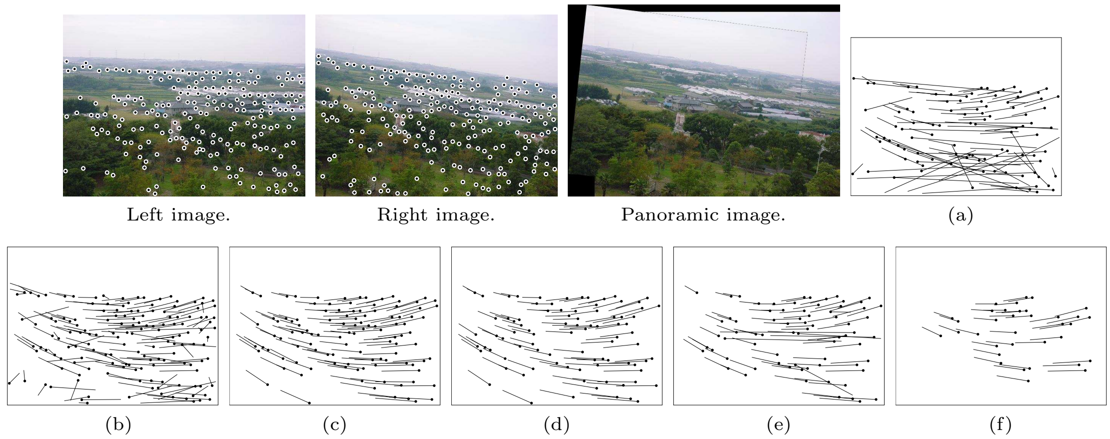

図 2 上段：入力画像と生成したパノラマ画像．(a)～(f) は図 1(a)～(f) に対応する結果
Fig. 2 Upper row: Input images and the generated panoramic image. The results (a)～(f) correspond to Fig. 1(a)～(f).

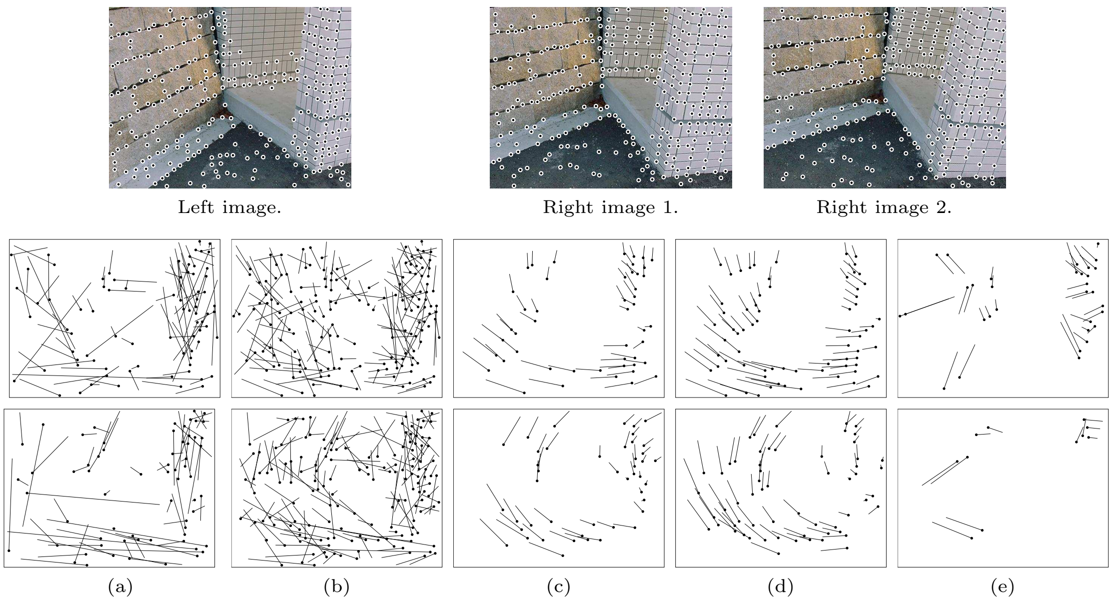

図 3 上段：入力画像．右画像 1，2 はそれぞれ約 5°，10° 回転している．中段：左画像と右画像 1 を用いた結果．下段：左画像と右画像 2 を用いた結果．(a)～(e) はそれぞれ図 1(a)～(d)，(f) に対応する結果
Fig. 3 Upper row: Input images. The right images 1 and 2 are rotated approximately by 5° and 10°, respectively. Middle row: Results using the left image and the right image 1. Bottom row: Results using the left image and the right image 2. The results (a)～(e) correspond to Fig. 1(a)～(d) and (f).

幾何学的 AIC はそれぞれ

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

であり，画像間の変換が射影変換であると判定された．求めた射影変換行列から生成したパノラマ画像を図 2 の上段に示す．

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認する実験を行った．図 3 の上段の画像は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難しい．右画像 1，2 は異なる位置でカメラをそれぞれ約 5 度および 10 度回転して得た画像である．中段は左画像と右画像 1 を用いた対応，下段は左画像と右画像 2 を用いた対応を示す．それぞれ (a)～(e) は図 1(a)～(c)，(f) に対応する結果である．

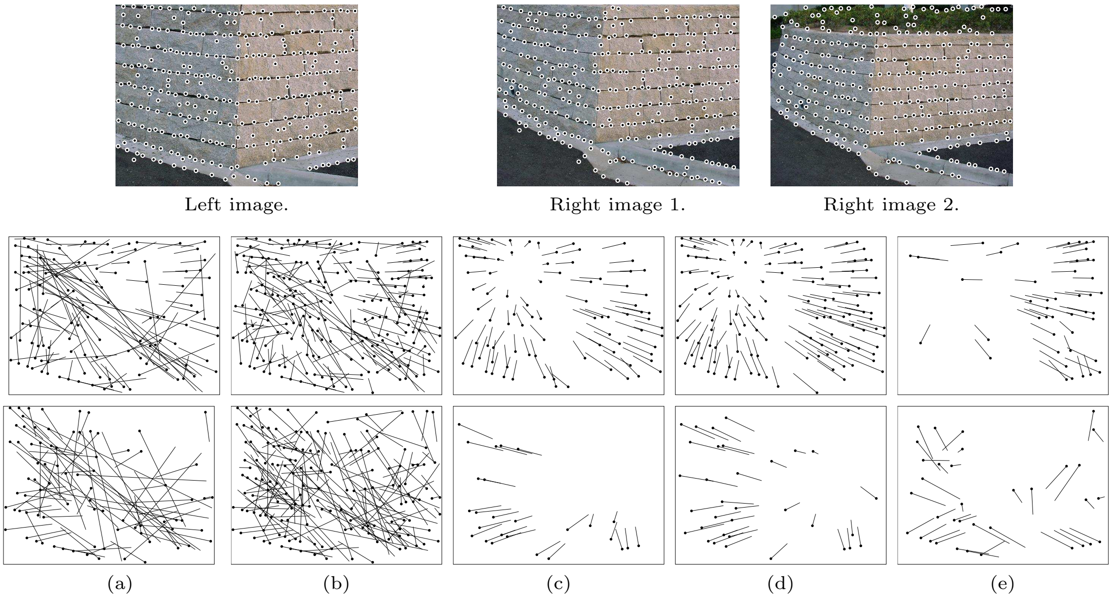

図 4 上段：入力画像．右画像 1，2 はそれぞれ約 80%，65%縮小している．中段：左画像と右画像 1 を用いた結果．下段：左画像と右画像 2 を用いた結果．(a)∼(e) はそれぞれ図 3(a)∼(e) に対応する結果

Fig. 4 Upper row: Imput images. The right images 1 and 2 are zoomed out approximately by 80% and 65%, respectively. Middle row: Results using the left image and the right image 1. Bottom row: Results using the left image and the right image 2. The results (a)∼(e) correspond to Fig. 3(a)∼(e).

図 4 は別の画像例である．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．右画像 1，2 は異なる位置でカメラをそれぞれ約 80%および 65%にズームした画像である．図 3 と同様に計算した結果を中段および下段に示す．

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 Zhang らの方法を用いても正しい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 768 MB であり，OS には Debian GNU/Linux を用いた．計算時間は上記の実験例で平均約 23 秒かかった．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いえる．

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 1 対 1 化して仮の対応を選び，それから全体的な傾向を推定する．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更新し，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」エピ極線方程式をあてはめる．

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，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関係を反復によって全体に波及させる方法が主であったが，本方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．ただし，初期に与える特徴点数が少なかったり，局所的に偏在したりしていると大域的な推定ができないという問題は残る．

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21世紀COEプログラム「インテリジェント ヒューマン センシング」，科学研究費基盤研究C（2）（No.15500113），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った.

参 考 文 献

1) Beardsley, P., Torr, P. and Zisserman, A.: 3D model acquisition from extended image sequences, *Proc. 4th Euro. Conf. Comput. Vision*, Cambridge, U.K., Vol.2, pp.683–695 (1996).

2) Fischler, M.A. and Bolles, R.C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Comm. ACM*, Vol.24, No.6, pp.381–395 (1981).

3) Harris, C. and Stephens, M.: A combined corner and edge detector, *Proc. 4th Alvey Vision Conf.*, Manchester, U.K., pp.147–151 (1988).

4) Hartley, R. and Zisserman, A.: *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U.K. (2000).

5) Kanatani, K.: *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，Vol.37, No.6, pp.1073–1080 (1996).

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，電子情報通信学会論文誌A, Vol.J86-A, No.12, pp.1520–1509 (2003).

8) Kanatani, K., Ohta, N. and Kanazawa, Y.: Optimal homography computation with a reliability measure, *IEICE Trans. Inf. & Sys.*, Vol.E83-D, No.7, pp.1369–1374 (2000).

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，Vol.42, No.SIG 6 (CVIM 2), pp.1–8 (2001).

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，電子情報通信学会論文誌D-II, Vol.J86-D-II, No.6, pp.816–824 (2003).

11) Lee, M.-S., Medioni, G. and Mordohai, P.: Inference of segmented overlapping surfaces from binocular stereo, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol.24, No.6, pp.824–837 (2002).

12) Maciel, J. and Costeira, J.: A global solution to sparse correspondence problems, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol.25, No.2, pp.187–199 (2003).

13) Olson, C.F.: Maximum-likelihood image matching, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol.24, No.6, pp.853–857 (2002).

14) Rousseeuw, P.J. and Leroy, A.M.: *Robust Regression and Outlier Detection*, Wiley, New York (1987).

15) Smith, S.M. and Brady, J.M.: SUSAN—A new approach to low level image processing, *Int. J. Comput. Vision*, Vol.23, No.1, pp.45–78 (1997).

16) van Wyk, M.A., Durrani, T.S. and van Wyk, B.J.: A RKHS interpolator-based graph matching algorithm, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol.24, No.7, pp.988–995 (2002).

17) Zhang, Z., Deriche, R., Faugeras, O. and Luong, Q.-T.: A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry, *Artif. Intell.*, Vol.78, pp.87–119 (1995).

(平成15年3月17日受付)
(平成15年9月 9 日採録)

(担当編集委員 岡谷 貴之)

金澤 靖（正会員）
1985年豊橋技術科学大学工学部情報工学課程卒業．1987年同大学院修士課程修了．博士（工学）．富士電機（株），群馬高専講師を経て，現在豊橋技術科学大学知識情報工学系助教授.

金谷 健一（正会員）
1972年東京大学工学部計数工学科（数理工学）卒業．1979年同大学院博士課程修了．工学博士．群馬大学工学部情報工学科教授を経て，現在岡山大学工学部情報工学科教授．IEEEフェロー.

☆ 作成したプログラムを以下に公開している．
http://www.img.tutkie.tut.ac.jp/