

エピ極線幾何学による 2 画像間の密な点対応の生成

菅谷 保之* 金澤 靖† 金谷 健一*

* 岡山大学工学部情報工学科 † 豊橋技術科学大学知識情報工学系

2 画像間に密な点対応を生成する方法を述べる。まず少数の初期対応から基礎行列を計算し、それを用いてエピ極線が水平にそろるように 2 画像を平行化し、一方の画像から抽出した特徴点に対してテンプレートを水平に走査する。しかし、このような 2 次元画像上の対応探索には限界がある。これを克服するために、本論文では得られる対応に対する「シーンの 3 次元形状の自然さ」まで考慮して誤対応を除去する。まず平行化の手順を述べ、大きさの異なる複数のテンプレートをを用いる階層的探索や多数決原理、大域的整合性による疑わしい対応の除去、および仮の 3 次元復元にに基づく誤対応の除去を導入する。最後に実画像例を示す。

Generating Dense Point Matches over Two Images Using Epipolar Geometry

Yasuyuki Sugaya* Yasushi Kanazawa† Kenichi Kanatani*

*Department of Information Technology, Okayama University, Okayama 700-8530 Japan

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

Dense point matches are generated over two images by computing the fundamental matrix from a few initial matches, rectifying the two images to align epipolar lines horizontally, and horizontally sliding a template for correspondence search. However, this type of 2-D search has an inherent limitation. To overcome this, we incorporate the “naturalness of the 3-D shape of the scene” implied by the resulting matches. After stating our procedure for rectification, we describe our scheme for hierarchical search and majority voting using multi-scale templates and our technique for outlier removal using a global consistency requirement and tentatively reconstructed 3-D shapes. Finally, we show real image examples.

1. まえがき

画像から 3 次元形状を復元するには因子分解法 [11, 14] のように連続ビデオ画像上を追跡した特徴点の軌跡に基づく方法と、異なるフレーム間の特徴点を直接に対応づける方法 [9, 15] がある。本論文では後者を考える。

対応抽出の原理は、2 画像上に特徴点抽出フィルタ [17] によって独立に特徴点を抽出し、各点の近傍の類似度をテンプレートマッチング [13] で測り、類似のものを対応づけることである。そして、誤対応を排除するためにエピ極線拘束条件、射影変換、大域的整合性のような拘束条件を導入し、条件を満たさないものを排除する [16, 26]。

しかし、このように疑わしい対応を排除すると、最終的にごく少数の対応しか残らないことが多い。しかも、それが画像上の一部に集中することもある。3 次元復元は普通に対応づけた特徴点の 3 次元位置を計算し、それを頂点とする多面体にテクスチャマッピングを施して表示するので、特徴点が少ないと一部に偏ると詳細な形状が復元できない。

これを解決する代表的な方法は、得られた対応から基礎行列 [9] を計算し、第 1 画像上に新たな特徴点を定義し、計算した基礎行列が定義するエピ極線に沿って第 2 画像上の対応点を探索することである。これはステレオ画像の対応探索の原理であるが、そのエピ極線が水平であればテンプレートの走査が容易になることから、まずエピ極線が平行になるように画像を変換する方法が知られている。これは画像をあたかも平行移動したカメラから撮影したかのように変換するものであり、本論文ではこれを画像の平行化 (image rectification) と呼ぶ。

これを初めて提唱したのは Ayache ら [2, 3, 4] であり、カメラは校正済みと仮定された。その後、その簡略化 [23] や拡張 [5, 6] が提案された。一方、カメラが未校正のまま基礎行列のみから平行化する理論は Hartley [8] によって提唱され、その簡略化 [1] やパラメータの最適化 [10, 18]、三眼視への拡張 [25] が提案された。

これらは原画像を射影変換するものであり、カメラ移動の仕方によっては画像の一部が無遠に写像されることがある。これを防ぐために円筒座標 [21] や極座標 [24] に変換する方法やその拡張 [20] が提案されている。

本論文では射影変換に基づく新たな平行化の手順を

*700-8530 岡山市津島中 3-1-1, Tel. (086)254-8244, 8173
{sugaya, kanatani}@suri.it.okayama-u.ac.jp
†441-8580 豊橋市天伯町雲雀ヶ丘 1-1, (0532)44-6888
E-mail: kanazawa@tutkie.tut.ac.jp

提案する．これは原理的には従来の手法 [1, 8, 10, 18] と大差ないが，計算がより単純で，幾何学的解釈がより明確になる．

もちろん，画像を平行化しただけでは問題は解決しない．いかにしてテンプレートマッチングによって正しい対応を決定し，誤った対応を排除するかが重要な課題である．本論文では大きさの異なる複数のテンプレートを用いる階層的探索，多数決原理を導入する．さらに，得られた対応から大域的整合性を考慮して疑わしい対応を除去する．

しかし，どんなに工夫しても，このような2次元画像上の探索は本質的に不完全である．なぜなら，エピ極線拘束条件を満たし，かつその近傍が類似しているような特徴点間の誤対応は除去できないからである．これを解決するには，「そのような誤対応はそのシーンの3次元形状を考慮すると不自然である」という，3次元にまで踏み込んだ判定が不可欠である．しかし，3次元形状を計算するには正しい対応づけが必要である．本論文では反復によって，復元した3次元形状をチェックしながら誤対応を検出する方法を示す．そして，実画像実験によって，2次元画像上の処理では除去不可能な誤対応が除去されることを実証する．

2. 平行化の手順

本論文の平行化の手順は次の通りである (図1)．

1. 2画像から少数 (8組以上) の特徴点の対応を与える．
2. その対応から基礎行列を計算する．
3. その基礎行列から各画像のエピ極点の方向を計算する．
4. 各画像をエピ極点が水平方向に位置するように回転する．
5. 各画像に射影変換を施して，エピ極点を水平方向の無限遠に写像する．
6. 2画像の対応する特徴点が同じ高さになるように第2画像を射影変換する．

以下，各ステップの手順の詳細を述べる．

2.1 対応の入力

入力 is 自動対応づけ処理 [16, 26] で得られる対応から信頼性があると思われるものを選んでよいし，手動で与えてもよい．これは応用に依存して決めることとして，本システムでは単に与えられる入力とする．

2.2 基礎行列の計算

少数の対応からでも高精度に基礎行列を求める必要があるため，統計的に最適な手法であるくりこみ

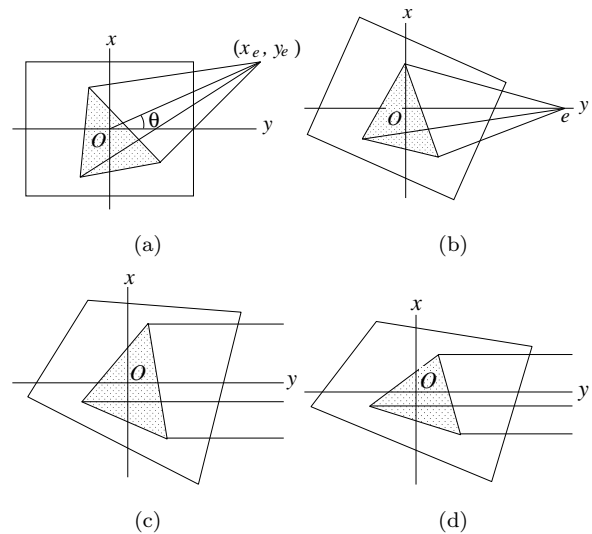


図1: 画像の回転 ((a)→(b))，エピ極線の平行化 ((b)→(c))，エピ極線の高さの調節 ((c)→(d))．

法¹[19]を用いる．

このとき画像の座標系を指定する必要がある．本システムではフレームの中心を原点 (0, 0) とし，上方を x 軸，右方を y 軸とし， z 軸を奥行き方向にとる (単位は画素)．そして z 軸を光軸 (レンズの中心軸) とみなし，画像からある距離 f_0 (単位は画素) だけ離れた位置に視点 (カメラのレンズ中心) があるとみなす²．

注1: 基礎行列は座標系に依存しているため，例えば画像フレームの左上を原点にとると値が変わる．他人の実験結果と比較するときは座標系に向きやスケールに注意が必要である．

注2: ここで仮定するカメラモデルは仮想的であり，実際のカメラがそうである必要はない．実際のカメラでは光軸が別の位置 (光軸点) を通り，画素に対応する受光素子の配列間隔の縦横比 (アスペクト比) が1でなかったり，縦横方向間の角度 (歪み角) が直角でなかったりする．また視点までの距離 (焦点距離) も f_0 とは限らない．しかし，画像を解析する限り，これらはどう仮定してもよい．それに対して，3次元形状を復元するには，実際のカメラの物理的な構造を正しく計測 (カメラ校正) しなければ，復元した形状は物理的な形状と一致しない．

2.3 エピ極点の計算

計算した基礎行列を F とするとき， F^T ， F の固有値0の単位固有ベクトルをそれぞれ e ， e' とする．固有ベクトルの符号は不定であるから，第3成分が非負になるように符号を選ぶ (第3成分が0のときの符号は任意でよい)．

¹以下にプログラムを公開している．

<http://www.img.tutkie.tut.ac.jp/>

²原理的には任意にとってよい．本システムでは $f_0 = 600$ と置いている．

ベクトル e, e' の始点を視点に置いたとき、これらの指す画像上の位置がエピ極点と呼ぶ。

注 3: データに誤差があっても、基礎行列 F は $\det F = 0$ であるように計算されるので、 F^T も F も固有値 0 をもつ。実際の計算では半正値対称行列 $FF^T, F^T F$ の最小固有値に対応する固有ベクトルをそれぞれ e, e' とする。

注 4: よく知られているように、第 1 画像の視点から見ると第 2 画像の視点（第 1 画像の座標系に関して）ベクトル e の方向にあり、第 2 画像の視点から見ると第 1 画像の視点（第 2 画像の座標系に関して）ベクトル e' の方向にある。ただし、これらの“方向”は仮定した仮想カメラモデルに対するものである。その実際の物理的な方向を知るにはカメラ校正が必要である。ここでは 3 次元情報は利用しないので、その必要はない。

2.4 画像の回転

ベクトル $e = (e_1, e_2, e_3)^T$ から第 1 画像上のエピ極点 (x_e, y_e) が $e_3 \neq 0$ のとき次のように定まる。

$$x_e = f_0 \frac{e_1}{e_3}, \quad y_e = f_0 \frac{e_2}{e_3} \quad (1)$$

これが y 軸上に載るように画像を原点の周りに次のように回転する。これによって各点 (x, y) が次の点 (\tilde{x}, \tilde{y}) に写像される。

$$\tilde{x} = x \cos \theta - y \sin \theta, \quad \tilde{y} = x \sin \theta + y \cos \theta \quad (2)$$

ただし、 θ は原点と (x_e, y_e) を結ぶ線分から y 軸（正の方向または負の方向）までの角度である。これによりエピ極点が $(0, e)$ に移る。ここに

$$e = x_e \sin \theta + y_e \cos \theta \quad (3)$$

である。第 2 画像についても同様に回転する。

注 5: 角度 θ は次のように計算するとよい。

$$\theta = \begin{cases} \tan^{-1}(e_1/e_2) & |e_2| \geq |e_1| \\ \pi/2 - \tan^{-1}(e_2/e_1) & |e_1| > |e_2| \end{cases} \quad (4)$$

こうすると (e_1, e_2) が半平面 $e_1 + e_2 \geq 0$ にあれば y 軸上の正の方向に ($e > 0$)、 $e_1 + e_2 < 0$ にあれば y 軸上の負の方向に ($e < 0$) 写像される。

注 6: 上記の計算では $e_3 \neq 0$ と仮定しているが、 $e_3 = 0$ ならエピ極点は原点から見てベクトル $(e_1, e_2)^T$ の方向にある無限遠点となり、 $e = \pm\infty$ となる。しかし、式 (4) のよって θ を計算して式 (2) の変換を行えば、 e_3 が 0 かどうかに問わず画像の変換ができる。

注 7: $(x_e, y_e) = (0, 0)$ のときは式 (4) の θ の計算ができない。しかし、本論文ではエピ極点は画像中にないと仮定する。したがって $(x_e, y_e) \neq (0, 0)$ であり、 $e \neq 0$ である。第 2 画像についても同様である。この仮定については後により詳しく考察する。

2.5 エピ極線の平行化

よく知られているように、一方の画像上の点 p の他方の画像上の対応位置 p' は点 p のエピ極線上にあり、すべてのエピ極線はその画像のエピ極点を通る。したがって、エピ極点を無限遠に写像する射影変換を施せば、すべてのエピ極線が平行になる。そこで次の射影変換によってエピ極点 $(0, e)$ を無限遠点 $(0, \pm\infty)$ に写像する。

$$\hat{x} = \frac{\tilde{x}}{1 - \tilde{y}/e}, \quad \hat{y} = \frac{\tilde{y}}{1 - \tilde{y}/e} \quad (5)$$

第 2 画像についても同様の射影変換を施す。

注 8: 射影変換は 8 パラメータをもつ 1 次分数式で表せるが [12]、次の条件を課すと式 (5) の形となる。

1. 点 $(0, e)$ を無限遠点 $(0, \pm\infty)$ に写像する。
2. 原点および x 軸上の点は動かさない。
3. y 軸の伸縮の割合が原点の近傍で 1 である。

注 9: $e_3 = 0$ のとき、すなわち $e = \pm\infty$ のときは式 (5) の分母の \tilde{y}/e を 0 とすればよいから、この変換は e_3 が 0 かどうかに問わず実行できる ($e \neq 0$ と仮定していることに注意。↔注 7)。

2.6 高さの調節

以上の変換によって各画像上のエピ極線は水平になるので、対応するエピ極線の高さが同じになるように、次の射影変換によって第 2 画像を上下に伸縮する。

$$\hat{x}' = \frac{a\hat{x} + b}{c\hat{x}' + 1}, \quad \hat{y}' = \frac{a\hat{y}'}{c\hat{x}' + 1} \quad (6)$$

係数 a, b, c は次のように定める。入力として与えた第 1 画像の N 個の特徴点 $\{(x_\alpha, y_\alpha)\}$ とそれに対応する第 2 画像の特徴点 $\{(x'_\alpha, y'_\alpha)\}$ に前項までに変換を施した位置を $\{(\hat{x}_\alpha, \hat{y}_\alpha)\}, \{(\hat{x}'_\alpha, \hat{y}'_\alpha)\}$ とし、変換後の高さなるべくそろうように、すなわち

$$\hat{x}_\alpha \approx \frac{a\hat{x}'_\alpha + b}{c\hat{x}'_\alpha + 1}, \quad \alpha = 1, \dots, N \quad (7)$$

となるように a, b, c を定める。これには分母を払った最小二乗法

$$J = \sum_{\alpha=1}^N \left(a\hat{x}'_\alpha + b - \hat{x}_\alpha(c\hat{x}'_\alpha + 1) \right)^2 \rightarrow \min \quad (8)$$

を用いる。 a, b, c で微分して 0 と置くと、正規方程式

$$\begin{pmatrix} \sum_{\alpha=1}^N \hat{x}'_\alpha{}^2 & \sum_{\alpha=1}^N \hat{x}'_\alpha & -\sum_{\alpha=1}^N \hat{x}_\alpha \hat{x}'_\alpha{}^2 \\ \sum_{\alpha=1}^N \hat{x}'_\alpha & N & -\sum_{\alpha=1}^N \hat{x}_\alpha \hat{x}'_\alpha \\ -\sum_{\alpha=1}^N \hat{x}_\alpha \hat{x}'_\alpha{}^2 & -\sum_{\alpha=1}^N \hat{x}_\alpha \hat{x}'_\alpha & \sum_{\alpha=1}^N \hat{x}_\alpha{}^2 \hat{x}'_\alpha{}^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{\alpha=1}^N \hat{x}_\alpha \hat{x}'_\alpha \\ \sum_{\alpha=1}^N \hat{x}_\alpha \\ -\sum_{\alpha=1}^N \hat{x}_\alpha{}^2 \hat{x}'_\alpha \end{pmatrix} \quad (9)$$

を得る．これを解いて a, b, c が定まる．入力した特徴点位置が正しければ式 (7) が等号で成立するはずであるが，特徴点位置には少なくとも整数化の誤差があるので，一般には等号は成立しない．そこで，点 $(\hat{x}'_\alpha, \hat{y}'_\alpha)$ の変換後の位置を点 $(\bar{x}'_\alpha, \bar{y}'_\alpha)$ とするとき，平行化の精度を次の値 h で評価する．

$$h = \sqrt{\frac{1}{N} \sum_{\alpha=1}^N (\hat{x}'_\alpha - \bar{x}'_\alpha)^2} \quad (10)$$

注 10： 射影変換は 8 パラメータをもつ 1 次分数式で表せるが [12]，次の条件を課すと式 (6) の形となる．

1. 水平線は水平線に写像される．
2. y 軸上の点は y 軸上の点に射影される．
3. 原点の近傍で x 軸と y 軸の伸縮の割合が等しい．

注 11： 上記の方法は従来の方法 [1, 8, 10, 18] とは異なる．Hartley [8] は両画像のエピ極線を水平にそろえる画像の射影変換が基礎行列 F から解析的に定まることを示し，従来の方法はこれに従っている．したがって，上記のような個々の特徴点位置を用いる調節は不要である．しかし，Hartley の方法で定まる射影変換は一意的ではなく，3 自由度が残る．実際，変換した第 2 画像に

- 水平に平行移動，
- 水平に伸縮，
- y 軸を固定して x 軸を傾ける線形変換（せん断）

を加えてもエピ極線は不変である．従来の方法では不定の 3 自由度を定めるために，個々の特徴点位置を用いた種々の最適化を施している．上記の方法はその不定の 3 自由度を

1. 第 1 画像の原点は不変，
2. 第 2 画像の原点の水平位置は不変，
3. 両画像の原点の近傍でアスペクト（縦横比）と直交関係が保たれる

ように定めたことに等しい．従来の方法では最適化の意味が明確でなく，変換画像がどのようなものかは変換してみないとわからないのに対して，本論文の方法では最適化計算が不要であるだけでなく，変換画像の性格が明確であり，かつ妥当であると考えられる．

注 12： 高さの調節は本来は式 (10) の h （あるいは h^2 ）を最小にする最適化を行うべきであり，そのためのくりこみ法も構成できるが，ここでは簡単のために式 (8) の最小二乗法を用いている．これは伸縮が一様 ($c = 0$) の場合には最適であり，実際問題では c はごく小さいのでこれで十分である．

3. テンプレートマッチング

3.1 画像の変換

前節の変換をすべて合成し，ベクトルと行列で表すと，第 1 画像については次のように書ける．

$$\hat{x} = Z[GRx], \quad x = \begin{pmatrix} x/f_0 \\ y/f_0 \\ 1 \end{pmatrix}, \quad \hat{x} = \begin{pmatrix} \hat{x}/f_0 \\ \hat{y}/f_0 \\ 1 \end{pmatrix} \quad (11)$$

$$R = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\gamma & 1 \end{pmatrix} \quad (12)$$

$$\gamma = \frac{e_3}{e_1 \sin \theta + e_2 \cos \theta} \quad (13)$$

ここに， $Z[\cdot]$ は z 成分を 1 とする正規化作用素であり， θ は式 (4) で計算する角度である．第 2 画像については次のように書ける．

$$\bar{x}' = Z[KG'R'x']$$

$$x' = \begin{pmatrix} x'/f_0 \\ y'/f_0 \\ 1 \end{pmatrix}, \quad \bar{x}' = \begin{pmatrix} \hat{x}'/f_0 \\ \hat{y}'/f_0 \\ 1 \end{pmatrix} \quad (14)$$

$$K = \begin{pmatrix} a & 0 & b/f_0 \\ 0 & a & 0 \\ cf_0 & 0 & 1 \end{pmatrix} \quad (15)$$

ただし， G', R' は式 (12) の G, R を第 2 画像について同様に定義したものである．

注 13： 実際に画像を変換するには，変換後の画像フレームを二つ用意し，各々のフレームの各画素の位置をベクトル \hat{x}, \bar{x}' で表し，逆変換

$$x = Z[R^\top G^{-1} \hat{x}], \quad x' = Z[R'^\top G'^{-1} K^{-1} \bar{x}'] \quad (16)$$

によって定まるベクトル x, x' に対応する第 1 画像，第 2 画像の画素の輝度値（カラー画像では R, G, B 値）に対応するフレームにコピーすればよい．

注 14： 式 (13) の γ は f_0/e のことであるが，式 (1), (3) を使わずに元の $e = (e_1, e_2, e_3)^\top$ に戻ってこのように計算すると e_3 が 0 かどうかに関らず計算できる．

注 15： 注 7 で述べたように，エピ極点は画像中ないと仮定している．もしエピ極点が画像中があれば，原点からエピ極点に近づく部分が無限遠に写像され，エピ極点を過ぎた部分が反対側の無限遠方向から戻ってくる．その場合は円筒座標や極座標による処理で対処できる [20, 24, 21]．しかし，このようなことが起こるのはカメラをほぼ前後に動かす場合であり，そのようにカメラを移動すると 3 次元復元の精度が低くなる [24]．このため実用上は重要でないと考えられるので，本システムではカメラはほぼ左右に移動するものとし，どちらかの画像でエピ極点が原点からある距離 L 以内であれば警告メッセージを出して終了する³．その判定条件は第 1 画像について

$$|\gamma| \geq \frac{f_0}{L} \quad (17)$$

であり，第 2 画像についても同様である．

注 16： 変換画像フレームの各画素に式 (16) の変換を適用した原画像上の位置は一般には画素の間に来るので，本システムでは周囲の画素の輝度値を双 1 次補間している．

³本システムでは L は画像サイズにとっている．

3.2 多重スケールテンプレートマッチング

新たな対応を生成するには、まず Harris 作用素⁴[7, 17]で第1原画像上に指定した個数の特徴点を発生させ、各々を式(11)によって変換後の位置に写像する。そして、その各々に対して変換後の第2画像上の同じ高さの水平線に沿ってテンプレートマッチングで対応点を探索する。最後にその位置を式(16)の逆写像によって第2原画像上に戻す。

このときテンプレートのサイズが問題となる。両画像で見え方が大域的にほぼ同じ領域ではテンプレートが大きいほうがロバストであるが、そうでなければテンプレートを小さくする必要がある。しかし、どちらかは事前には不明である。

そこで、本システムでは 33×33 , 17×17 , 9×9 , 5×5 , 3×3 の5種類のテンプレートを用いる。そして、探索する前に両画像にそれぞれ 17×17 , 9×9 , 5×5 , 3×3 , 0×0 (何もしない) のガウス平滑化フィルターを施しておく。標準偏差はそれぞれ 8, 4, 2, 0.5, 0 である。対応を定める方法として、本システムは次の二通りを考える。

階層的探索 $s = 16$ から出発して、 $(2s+1) \times (2s+1)$ テンプレートが最もよくマッチする位置を定め、次にその内部を $(s+1) \times (s+1)$ テンプレートで探索する。これが最小となる位置が大きいテンプレートの中心から s 以上であれば「対応点なし」と判定する。そうでなければ s を半分にし、 $s = 1$ となるまで同様の探索を行う。

多数決原理 各々のテンプレートが最もよくマッチする水平位置をソートして $y_1 \leq \dots \leq y_5$ とする。区間 $[y_1, y_3]$, $[y_2, y_4]$, $[y_3, y_5]$ の中の最も狭い区間の幅が4画素以下なら、その区間に入る4個の値の平均を対応位置とし、そうでないときは「対応点なし」と判定する。

多くの実画像実験を行っても、どちらかがよいとは断定できなかったので、本システムではオプションによってどちらも選択できるようにしている。

注 17: テンプレート内の画像の比較は通常は残差平方和

$$RSS = \sum_{(x,y) \in T} (I_1(x,y) - I_2(x',y'))^2 \quad (18)$$

で行う。ただし、 T はテンプレート領域であり、 I_1, I_2 はそれぞれ第1, 第2画像の輝度値である。また (x', y') はテンプレート領域の画素 (x, y) と比較される第2画像の画素である。第1画像と第2画像とで照明環境が変化している場合は、直接に比較してもよくマッチしない。この照明による輝度値の変化を打ち消すには、比較する領域のそれ

ぞれについて輝度値の平均が0, 分散が1となるように正規化しておけばよい。これは式(18)の残差平方和の代わりに正規化相関を用いることに等価である。しかし、輝度値を正規化すると、本来マッチすべきでない領域がマッチすることも生じ、対応づけ能力が低下する。そこで、本システムでは輝度値の正規化を行うか行わないかはユーザーのオプションとしている。

注 18: 原画像上の特徴点を第1変換画像に写像した位置は一般に画素間にあり、実数座標をもつ。本システムではこれを整数化せず、第2変換画像上のその実数の高さの水平線を探している。画像は双1次補間によって連続空間とみなしている。

注 19: ガウス平滑化は行っても行わなくてもほとんどの場合は変化はないが、物体境界上の特徴点の背景に見え方が2画像で異なる場合に、その相違を緩和して対応しやすくする効果も見られた。

注 20: 平滑化は画像フレームの境界付近に対しても部分ガウス関数を適用して、全画素に適用している。またテンプレートが画像フレームからはみ出す場合も、残差をフレーム内部の画素のみについて計算して、1画素当たりに換算している。このため画像フレームの境界の影響は生じない。

3.3 サブ画素補正

平行化したエッジ極線が厳密に正しければ対応点はその上にあるはずであるが、もともとの特徴点は画素単位でしか指定されないの、エッジ極線も厳密ではない。さらに水平方向にも画素単位でしか探索しないので、精度が限定される。そこで、上記の手順で定めた対応点の周りにテンプレートをサブ画素精度で移動してより詳細な探索を行う。

本システムではその点から上下左右に微小距離 h だけ離れた9個所からテンプレート残差が最小の位置を探し、その点から h を半分にして同様な操作を行い、 h が十分小さくなるまでこれを反復している。 h の初期値には式(10)で定まる値を用いて、平行化の精度に連動させている。

3.4 大域的整合性

上述の対応づけには誤対応が含まれることがある。これは、変換後の両フレームを重ねて、対応点間を線分で結んだ“オプティカルフロー”を表示するとよくわかる。すべてのフローは水平線分となるが、少数の際だって長い線分が含まれることがある。そのような対応は3次元シーンとして不自然であるから除去する。

本システムでは、画像を平行化に用いた初期対応のフローの長さの平均 μ と標準偏差 σ を計算し、フローの長さが平均 μ から標準偏差 σ の2倍の区間 $[\mu - 2\sigma, \mu + 2\sigma]$ になければ誤対応の可能性があると判定して排除している。

⁴以下にプログラムを公開している。
<http://www.img.tutkie.tut.ac.jp/programs/>

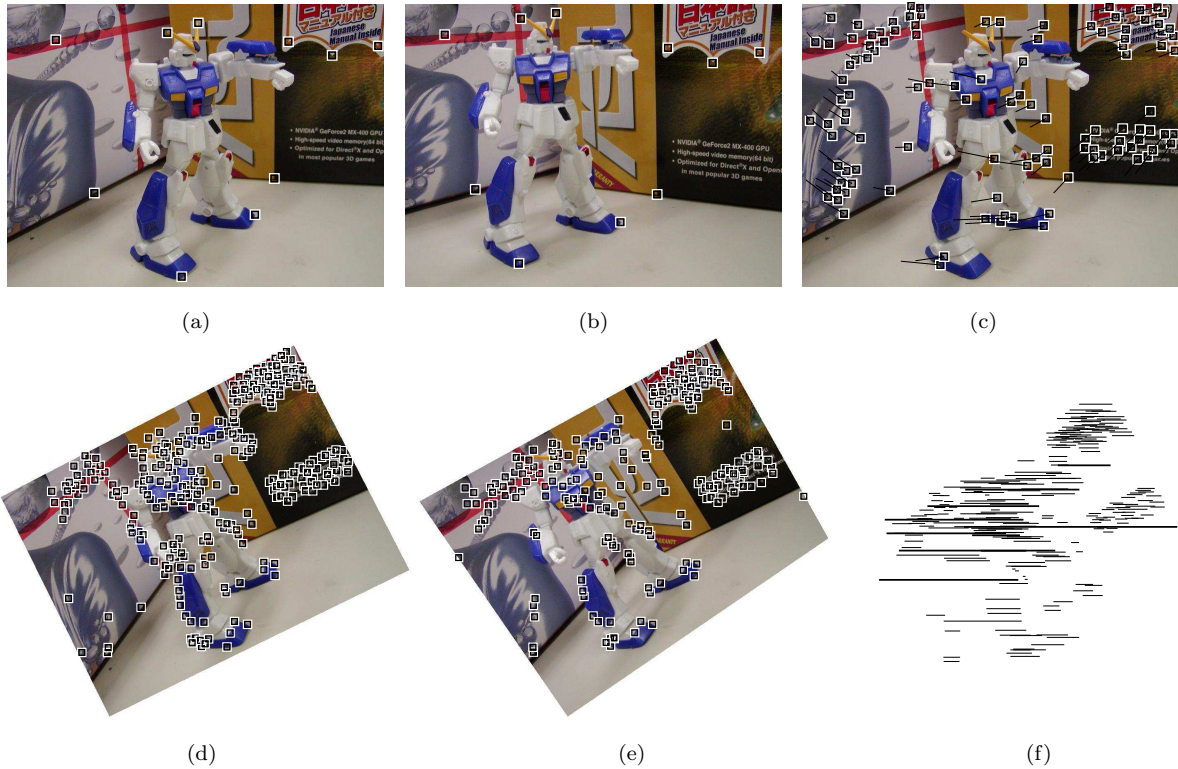


図 2: (a), (b) 入力画像と入力した対応点 . (c) 自動対応づけで得られる対応 (オプティカルフロー) . (d), (e) 第 1 画像上から検出した特徴点を平行化した画像上で探索した結果 . (f) その対応のオプティカルフロー . 太線は大域的整合性によって除去されたフロー .

4. 3次元情報による誤対応除去

以上の処理によって高精度な対応が得られるが、それでも誤対応が入り込むことがある。そのような誤対応は

1. エピ極線拘束条件を正しく満たし、
2. かつその近傍の見え方が同じ、

というものである。これは例えば物体境界上の T 型交点のように、物理的には異なる点であるのに画像上では同じに見えるもので、2次元画像処理のみによって除去することは原理的に不可能である。

しかし、3次元情報を利用すればこれが検出できる。得られた対応から3次元復元を行なうと、対応位置の誤りは奥行き方向に大きな変動を生じ、負の奥行き（カメラの背後の位置）が計算されることがある。正の奥行きであっても、特徴点を頂点とするシーンの多面体表示を行なうと、誤った特徴点が著しい刺（とげ）状の突起として現れる。そこで、そのような刺を生じる特徴点を誤対応として除去する。具体的な手順は次の通りである。

まず、それまでに得られている対応から基礎行列を計算し、各特徴点の3次元位置を計算する。これには金谷・三島 [15] の方法を用いる。そして、どちらかのカメラからみて奥行きが負となる特徴点を除

去する。残った特徴点から改めて基礎行列をし直し、再び3次元復元を行い、負の奥行きが生じるか調べ、これをすべての奥行きが正となるまで反復する。

次に、第1画像の特徴点を頂点とするドロネー三角形網を定義し、第1画像上を基準として、各点 p の離散ラプラシアン $L(p)$ を次のように定義する。

$$L(p) = \frac{(p \text{ の奥行き}) - (p \text{ の接続点の奥行き})}{(p \text{ の接続点の平面距離})} \quad (19)$$

ただし点 p の「接続点」とは点 p とドロネー三角形網の辺で結ばれている点のことであり、その辺が「接続辺」である。また「平面距離」とは奥行きに直交する方向の距離であり、2点 $(X, Y, Z), (X', Y', Z')$ に対して $\sqrt{(X - X')^2 + (Y - Y')^2}$ のことである。そして、しきい値 θ を定め⁵、

$$|L(p)| > \theta \quad (20)$$

かつ p が極値であるものを刺と判定する。ただし「極値」とは

$$(p \text{ の奥行き}) > (p \text{ の接続点の奥行き})$$

または

$$(p \text{ の奥行き}) < (p \text{ の接続点の奥行き})$$

⁵本システムではしきい値は $\theta = 3$ としている。

となる点 p のことである．このような刺を生じる対応を除去し，残った対応から基礎行列を計算し直す．そして，再び3次元復元を行い，刺が生じなくなるまでこれを反復する．

注 21： 金谷・三島 [15] の方法では基礎行列から2画像を撮影したカメラの焦点距離を計算し，それに基づいて3次元復元を行なう．しかし，対応に誤りが含まれていると計算される基礎行列が誤差を含み，計算した焦点距離が虚数になることがある（途中の平方根をとる式が負になる）．これを防ぐ処理も提案されているが [22]，ここでは適当な焦点距離の近似値を用いればよい．正しくない焦点距離を用いると復元形状が歪むが，それは空間全体に射影変換を施した，射影復元と呼ばれるものであり，定性的な遠近は再現できる．ここでの3次元復元は正しい形状を計算することではなく，負の奥行きや刺を検出するのが目的であるから，このような射影復元で十分である．

5. 実画像例

図 2(a), (b) は2枚の入力画像であり，図 2(c) は両画像から独立に Harris 作用素によって300個の特徴点を抽出し，金澤・金谷の方法⁶ [16] で自動的対応づけを行なったものである（第1画像上に得られた特徴点とそのオプティカルフローを示す）．これによって109点の対応が得られたが，背景の平坦な部分に局在する傾向がある．

図 2(d), (e) は図 2(a), (b) に記入した10個の特徴点を手動で入力し，それによって平行化した画像である．図 2(d) は，第1原画像上に Harris 作用素によって300個の特徴点を抽出し，それを写像した位置を記入している．そして，3.2節に述べた多重テンプレートマッチング（階層的探索）によって図 2(e) の画像上を走査し，得られた225点の対応位置を図 2(e) 中に記入している．この対応のオプティカルフローが図 2(f) である．11本の太線は3.3節の大域的整合性によって誤対応と判定された対応である．

図 3(a) はこの結果から得られた対応から仮の3次元復元形状である．これから負の奥行きや刺を除去して16点を除去し（この場合は負の奥行きは生じていない），最終的に得られた対応を図 3(b), (c) に示す．合計198個の対応が新たに得られた．図 2(c) の自動対応づけに比較して得られる対応数が多く，孤立した物体部分からかなり対応が得られている．図 3(d) はその3次元復元形状である．図 3(e), (f) はテンプレートマッチングによる対応から最終的に除去された対応である．

これ以外にもさまざまな画像について実験し，最終的に除去された特徴点を観察すると，次のように

分類されることがわかった．

1. 第1画像に指定した特徴点に対応する点が第2画像中に存在しない場合．これは，その点がフレームに外に出ている場合と，手前の物体の背後に隠れている場合とがある．
2. 画像中にほぼ周期的に繰り返すテクスチャがあるとき，誤った周期に対応する点．
3. 平行化したテクスチャパターンがたまたま類似している点．
4. 手前の物体の境界と背景上で生じる見掛けの特徴点の対応．
5. シーンの奥行きに大きな変動があり，正しい対応であるのに大域的整合性に反する，あるいは刺であるとみなされて除去された対応．

上記1~4.のような誤対応が生じることは以前から知られており，3次元復元を行う場合の障害となっている．しかし，いずれの場合もエッジ極線拘束条件を正しく満たしているため，これらを2次元画像上の探索のみで除去することは困難である．これらを除去するには，ここで行ったように，その対応の意味する3次元構造まで調べなければならない．

一方，シーンの3次元形状の「自然さ」を要求すると，上記5.のように，正しい対応でも除去されるものがある．これはトレードオフであり，やむを得ない面もあるが，これをどう減らすかが今後の課題である．

6. まとめ

本論文では，2画像上に密な点对応を生成するシステムを作成し⁷，その手順を示した．その基本は，少数の初期対応から基礎行列を計算し，エッジ極線が水平にそろうように2画像を平行化し，一方の画像から抽出した特徴点に対して，テンプレートマッチングによって対応を探索するものである．本論文では，精度向上のために大きさの異なる複数のテンプレートによる階層的探索や多数決原理を用い，大域的整合性によって疑わしい対応を除去する手法を導入した．

本論文の平行化法は従来の Hartley[8] の方法に基づくものとは異なり，最適化計算が不要であり，変換画像の性格が明確かつ妥当であると考えられる．

しかし，このような2次元画像上の探索には本質的に不完全である．本論文では，エッジ極線拘束条件を満たして近傍が似ている誤対応を除去するために，3次元形状まで考慮した誤対応除去法を示した．そ

⁶以下にプログラムを公開している．
<http://www.img.tutkie.tut.ac.jp/programs/>

⁷作成したプログラムソースの一部を下記に公開している．
<http://www.suri.it.okayama-u.ac.jp/>

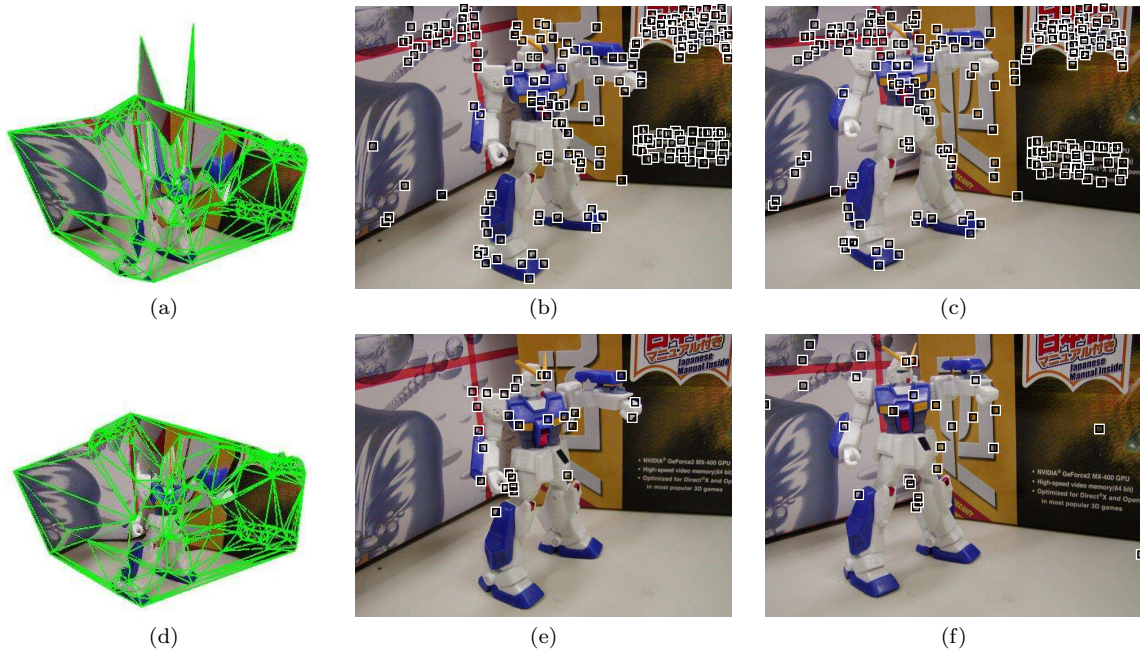


図 3: (a) 図 2 のテンプレートマッチングで得られた対応から復元した仮の 3 次元復元形状 . (b), (c) 3 次元形状を考慮して最終的に得られる対応点 . (d) その 3 次元復元形状 . (e), (f) 除去された対応点 .

して、実画像例によってその有効性を確認するとともに、問題点を考察した .

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. 15500113) の助成による .

参考文献

- [1] K. A. Al-Shalfan, J. G. B. Haigh and S. S. Ipson, Direct algorithm for rectifying pairs of uncalibrated images, *Electronics Lett.*, **36**-5 (2000-3), 419–420.
- [2] N. Ayache, *Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception*, MIT Press, Cambridge, MA, U.S.A., 1991.
- [3] N. Ayache and C. Hansen, Rectification of images for binocular and trinocular stereo vision, *Proc. 9th Int. Conf. Pattern Recog.*, November 1988, Rome, Italy, pp. 11–16.
- [4] N. Ayache and F. Lustman, Trinocular stereo vision for robotics, *IEEE Trans. Pattern Anal. Mach. Intell.*, **13**-1 (1991-1), 73–85.
- [5] A. Fusiello, E. Trucco and A. Verri, Rectification with unconstrained stereo geometry, *Proc. British Mach. Vision Conf.*, September 1997, Essex, U.K., pp. 400–409.
- [6] A. Fusiello, E. Trucco and A. Verri, A compact algorithm for rectification of stereo pairs, *Mach. Vision Appl.*, **12**-1 (2000-6), 16–22.
- [7] C. Harris and M. Stephens, A combined corner and edge detector, *Proc. 4th Alvey Vision Conf.*, Manchester, U.K., August 1988, pp. 147–151.
- [8] R. Hartley, Theory and practice of projective rectification, *Int. J. Comput. Vision*, **35**-2 (1999), 115–127.
- [9] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U.K., 2000.
- [10] F. Isgrò and E. Trucco, Projective rectification without epipolar geometry, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, June 1999, Fort Collins, CO, U.S.A., Vol. 1, pp. 94–99.
- [11] 金出武雄, コンラッド・ポールマン, 森田俊彦, 因子分解法による物体形状とカメラ運動の復元, 電子情報通信学会論文誌 D-II, **J74-D-II-8** (1993-8), 1497–1505.
- [12] 金谷健一, 「形状 CAD と図形の数学」, 共立出版, 1998.
- [13] 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, **J86-A-12** (2003-12), 1502–1509.
- [14] 金谷健一, 菅谷保之, 因子分解法の完全レシピ, 電子情報通信学会技術報告, PRMU2003-118, (2003-10), 19–24.
- [15] 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** (2001-6), 1–8.
- [16] 金澤靖, 金谷健一, 大域的な整合性を保証するロバストな画像の対応づけ, 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, **44-SIG 17** (2003-12), 70–77.
- [17] 金澤靖, 金谷健一, コンピュータビジョンのための画像の特徴点抽出, 電子情報通信学会誌, **87**-12 (2004-12), 1043–1048.
- [18] C. Loop and Z. Zhang, Computing rectifying homographies for stereo vision, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, June 1999, Fort Collins, CO, U.S.A., Vol. 1, pp. 125–131.
- [19] 三島等, 金谷健一, 基礎行列の最適計算とその信頼性評価, 情報処理学会研究報告, 99-CVIM-118-10 (1999-9), 67–74.
- [20] D. Oram, Rectification for any epipolar geometry, *Proc. British Mach. Vision Conf.*, September 2001, London, U.K., pp. 653–662.
- [21] S. Roy, J. Meunier and I. J. Cox, Cylindrical rectification to minimize epipolar distortion, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, June 1997, Puerto Rico, pp. 393–399.
- [22] 高橋 茂雄, 中辻 敦忠, 金谷健一, 未校正 2 画像からの 3 次元復元のための焦点距離計算の安定化, 情報処理学会研究報告, 2003-CVIM-141-12 (2003-11), 79–86.
- [23] D. V. Papadimitriou and T. J. Dennis, Epipolar line estimation and rectification for stereo image pairs, *IEEE Trans. Image Process.*, **5**-4 (1996-4), 672–679.
- [24] M. Pollefeys, R. Koch and L. Van Gool, A simple and efficient rectification method for general motion, *Proc. 7th Int. Conf. Comput. Vision*, September 1999, Kerkyra, Greece, Vol. 1, pp. 496–501.
- [25] C. Sun, Uncalibrated three-view image rectification, *Image Vision Comput.*, **21**-3 (2003-3), 259–269.
- [26] 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.