

コンピュータビジョンのためのくりこみ法

金 谷 健 一†

画像データを解析して3次元情報を抽出する場合に、通常は多数のデータに最小二乗法を適用して、ノイズに対するロバスト性を増している。本論文ではまずコンピュータビジョンに典型的に現れる最小二乗法では、その重みを最適に選んでも解に統計的な偏差があることを指摘する。そして、統計解析によって偏差を除去する「くりこみ法」と呼ぶ手法を定式化する。これを消失点や出現点の推定、コニックの当てはめ、3次元運動解析に適用した例を示し、画像の誤差についての情報がなくても精度が向上することを確認する。

Renormalization for Computer Vision

KENICHI KANATANI†

In extracting 3-D information from image data, it is customary to apply least-squares optimization to a large number of data to increase robustness. We first point out that the least-squares solution for typical problems of computer vision is "statistically biased" in the presence of noise. We then formulate a scheme called "renormalization", which removes the bias by automatically adjusting to image noise. In order to observe its effectiveness, the renormalization procedure is applied to estimation of vanishing points and focuses of expansion, conic fitting to image data, and 3-D motion analysis.

1. はじめに

画像データから3次元情報を抽出することは、ロボット技術の進歩につれて今日ますます重要になっている。解析の目的が物体形状あるいは物体やロボット自身の位置の計測であれば、その精度が重要な問題となる。通常画像データには誤差があるので、パラメータ数より多くのデータから最小二乗法のような最適化によってロバストな解を求めるのが普通である。このとき、データおよびそれから計算した値の「信頼性」を正しく重みづけなければ信頼性を向上させることはできない。このような問題に対して、画像データの誤差を数理的にモデル化し、画像データから計算した値の信頼性を共分散行列の形で定量的に評価する理論体系が提案され¹⁷⁾、カメラキャリブレーションの計算の最適化^{14), 20), 21)}や画像データの幾何学的推論の統計的基礎づけ^{11), 18), 19)}に応用されている。本論文ではこれをさらに発展させる。

コンピュータビジョンにおいてはパラメータの比、すなわち定数倍を除いた値を定める問題がきわめて多い。そのような問題の最小二乗法はほとんどの場合、データから作られる行列の固有値問題に帰着する。本

論文ではまず、そのような解には統計的な偏差が存在することを指摘する。不偏推定値を得るには、計算した解から理論的に導かれる偏差を差し引けばよいが、そのためには画像の誤差の絶対量を推定しなければならない。実際の画像では誤差はカメラ系や処理アルゴリズムだけでなく、照明条件や画像中の物体の性質にも依存するので、これをあらかじめ推定しておくことは困難である。

そこで本論文では画像の誤差を推定しながら偏差を除去する「くりこみ法」と呼ぶ手法を定式化する。そして消失点や出現点の推定、コニックの当てはめ、3次元運動解析に適用した例を示し、画像の誤差に関する情報がなくても精度が向上することを確認する。

2. 最小二乗推定と最適重み

コンピュータビジョンにおいてはパラメータ u_1, \dots, u_n の比を定める問題がよく生じる。そのような場合、定数倍の不定性を除去するために $u_1^2 + \dots + u_n^2 = 1$ のような正規化条件をつけるのが普通である。パラメータ u_1, \dots, u_n が線形拘束条件

$$a_{1\alpha}u_1 + \dots + a_{n\alpha}u_n = 0 \quad (1)$$

を満たすとする。各 $a_{i\alpha}, \alpha = 1, \dots, N$ は画像データから計算される量である。画像に誤差があればすべての拘束条件が同時に満たされるとは限らないので、次の最小二乗法を考える。

† 群馬大学工学部情報工学科
Department of Computer Science, Gunma University

$$\sum_{\alpha=1}^N W_{\alpha}(a_{1\alpha}u_1 + \dots + a_{n\alpha}u_n)^2 \rightarrow \min \quad (2)$$

ただし W_{α} は α 番目のデータの信頼性に応じた正の重みである。 n 次元行列 $\mathbf{A}=(A_{ij})$ を

$$A_{ij} = \sum_{\alpha=1}^N W_{\alpha} a_{i\alpha} a_{j\alpha} \quad (3)$$

と定義し、 n 次元ベクトル $\mathbf{u}=(u_i)$ を用いれば、問題は次のように書ける。

$$(\mathbf{u}, \mathbf{A}\mathbf{u}) \rightarrow \min, \quad \|\mathbf{u}\|=1 \quad (4)$$

ただし (\cdot, \cdot) でベクトルの内積を、 $\|\cdot\|$ でベクトルのノルムを表す。この解はよく知られたように、行列 \mathbf{A} の最小固有値に対する単位固有ベクトル \mathbf{n}_n で与えられる⁹⁾。各 $a_{i\alpha}$ は画像データから計算されるので、誤差を含んでいる。誤差がないときの真の値にバーをつけて、 $a_{i\alpha} = \bar{a}_{i\alpha} + \Delta a_{i\alpha}$ とおく。まず、式(2)の最小二乗法の重み W_{α} をどのように選ぶべきかを考察する。各 u_i が厳密な値 \bar{u}_i のとき

$$\varepsilon_{\alpha} = a_{1\alpha}\bar{u}_1 + \dots + a_{n\alpha}\bar{u}_n \quad (5)$$

と置く。厳密な $\bar{a}_{i\alpha}$ に対しては $\sum_{i=1}^n \bar{a}_{i\alpha}\bar{u}_i = 0$ であるから、 $\varepsilon_{\alpha} = \sum_{i=1}^n \Delta a_{i\alpha}\bar{u}_i$ となる。誤差がランダムであれば $\Delta a_{i\alpha}$ は確率変数とみなせる。これに対する期待値を $E[\cdot]$ で表す。 $E[\Delta a_{i\alpha}] = 0$ であれば $E[\varepsilon_{\alpha}] = 0$ である。分散は次のように評価される。

$$\sigma_{\alpha}^2 = E[\varepsilon_{\alpha}^2] = \sum_{i,j=1}^n E[\Delta a_{i\alpha}\Delta a_{j\alpha}]\bar{u}_i\bar{u}_j \quad (6)$$

$E[\Delta a_{i\alpha}\Delta a_{j\alpha}]$ を (ij) 要素とする行列を $V[\mathbf{a}_{\alpha}]$ とおけば

$$\sigma_{\alpha}^2 = (\bar{\mathbf{u}}, V[\mathbf{a}_{\alpha}]\bar{\mathbf{u}}) \quad (7)$$

と書ける。 $V[\mathbf{a}_{\alpha}]$ はベクトルデータ $\mathbf{a}_{\alpha}=(a_{i\alpha})$ の「共分散行列」である。各データ \mathbf{a}_{α} の誤差が互いに独立であれば、統計量 $x_{\alpha} = \varepsilon_{\alpha}/\sigma_{\alpha}$ の確率密度は標準正規分布 $e^{-\sum_{\alpha=1}^N x_{\alpha}^2}/\sqrt{2\pi}$ で近似できる。これを最大にするように $\bar{\mathbf{u}}$ を推定することにすれば（一種の「最尤推定」であるが、統計学でよく議論されているものとはやや異なる）、これは $\sum_{\alpha=1}^N \varepsilon_{\alpha}^2/\sigma_{\alpha}^2 \rightarrow \min$ を意味する。これは式(3)で $W_{\alpha} \propto 1/\sigma_{\alpha}^2$ 、すなわち

$$W_{\alpha} = \frac{\text{const.}}{(\bar{\mathbf{u}}, V[\mathbf{a}_{\alpha}]\bar{\mathbf{u}})} \quad (8)$$

と選ぶことに相当する。これを「最適重み」と呼ぼう。

しかし、これを計算しようとする、次の二つの問題が生じる。第一は $V[\mathbf{a}_{\alpha}]$ の評価である。各 $a_{i\alpha}$ の

誤差に何らかの仮定を設けて（すなわち誤差の「モデル」を導入して）、誤差の確率分布が記述できれば理論的には評価できる。しかし画像の誤差の挙動をあらかじめ定量的に予測することは一般には困難である。そこで、画像の誤差の大きさを特徴づける基本的なパラメータ（たとえば画素の位置ずれの標準偏差）を ε （厳密なデータに対して $\varepsilon=0$ ）とし、 $V[\mathbf{a}_{\alpha}]$ を ε のべき級数

$$V[\mathbf{a}_{\alpha}] = \varepsilon^2 V_0[\mathbf{a}_{\alpha}] + \varepsilon^4 V_2[\mathbf{a}_{\alpha}] + \dots \quad (9)$$

に展開し（通常は奇数べきの期待値は0）、4次以上の項を無視して $V_0[\mathbf{a}_{\alpha}]$ を用いる。こうすれば ε^2 は未知であっても、重みの定数倍は解に影響しないので考える必要がない。また誤差のない場合（ $\varepsilon=0$ ）も含めることができる。

第二に、最適重みは真の解 $\bar{\mathbf{u}}$ を含んでいるので直接に計算できないことである。これを解決するには、定数重みによる推定値を用いて最適重みを近似し、推定値を改良し、必要ならそれを反復すればよいように思える。しかし、次に示すように、そのような方法では解に偏差が生じる。

3. 最小二乗解の不偏推定

誤差によって行列 \mathbf{A} が $\mathbf{A} = \bar{\mathbf{A}} + \Delta\mathbf{A}$ に摂動すると、その固有値と固有ベクトルがどう摂動するかを示すのがよく知られた「摂動定理」⁹⁾である。 $\bar{\mathbf{A}}$ の固有値を $\bar{\lambda}_1 \geq \dots \geq \bar{\lambda}_{n-1} > \bar{\lambda}_n (=0)$ とし、対応する単位ベクトルのつくる直交系を $\{\bar{\mathbf{u}}_1, \dots, \bar{\mathbf{u}}_{n-1}, \bar{\mathbf{u}}_n\}$ とすると $\bar{\mathbf{A}} + \Delta\mathbf{A}$ の最小固有値に対する単位固有ベクトル $\hat{\mathbf{u}}_n + \Delta\mathbf{u}_n$ に対して次の関係が成り立つ^{9),17)}。

$$\Delta\mathbf{u}_n = - \sum_{i=1}^{n-1} \frac{(\bar{\mathbf{u}}_i, \Delta\mathbf{A}\bar{\mathbf{u}}_n)}{\bar{\lambda}_i} \bar{\mathbf{u}}_i + O(\Delta\mathbf{A})^2 \quad (10)$$

ただし $O(\Delta\mathbf{A})^2$ は $\Delta\mathbf{A}$ の各要素の二乗あるいはそれ以上のオーダーの項である。一般には $E[\Delta\mathbf{A}]$ は0ではないので、解 \mathbf{u}_n は $\bar{\mathbf{u}}_n$ の不偏推定値ではなく、 $O(E[\Delta\mathbf{A}])$ だけの偏差が存在する。

偏差がわかれば、それを差し引くと不偏推定値が得られる。すなわち \mathbf{u}_n の代わりに $\mathbf{u}_n - E[\Delta\mathbf{u}_n]$ とすればよい。しかし $E[\Delta\mathbf{u}_n]$ の計算には $E[\Delta\mathbf{A}]$ および真の値 $\bar{\lambda}_i, \bar{\mathbf{u}}_i, i=1, \dots, n-1$ が必要であり、これを知ることはできない。一方、もし $E[\Delta\mathbf{A}]$ が評価できれば

$$\hat{\mathbf{A}} = \mathbf{A} - E[\Delta\mathbf{A}] \quad (11)$$

と置き、その最小固有値に対する単位固有ベクトルを $\hat{\mathbf{u}}_n$ とすると、式(10)の摂動定理から $E[\hat{\mathbf{u}}_n] = \bar{\mathbf{u}}_n +$

$O(\Delta A)^2$ であることがわかる. 式(3)より

$$\Delta A = \sum_{\alpha=1}^N W_{\alpha}(\bar{\mathbf{a}}_{\alpha} + \Delta \mathbf{a}_{\alpha})(\bar{\mathbf{a}}_{\alpha} + \Delta \mathbf{a}_{\alpha})^T - \bar{\mathbf{a}}_{\alpha} \bar{\mathbf{a}}_{\alpha}^T \quad (12)$$

であるから, 期待値をとれば

$$E[\Delta A] = \sum_{\alpha=1}^N W_{\alpha} E[\Delta \mathbf{a}_{\alpha} \Delta \mathbf{a}_{\alpha}^T] = \sum_{\alpha}^N W_{\alpha} V[\mathbf{a}_{\alpha}] \quad (13)$$

となる. しかし前節で指摘したように, 式(13)を正確に評価することは困難である. そこでやはり(9)の展開を用いて

$$\mathbf{B} = \sum_{\alpha=1}^N W_{\alpha} V_0[\mathbf{a}_{\alpha}] \quad (14)$$

とおき,

$$\hat{\mathbf{A}} = \mathbf{A} - c\mathbf{B} \quad (15)$$

と定義する. そして $E[\hat{\mathbf{A}}] = \bar{\mathbf{A}}$ となるように定数 c を定める. あるいはある零でない定数 k があって $E[\hat{\mathbf{A}}] = k\bar{\mathbf{A}}$ であってもよい (定数倍しても固有ベクトルは同じ). しかし等号を厳密に成立させることはできない. 一方 $E[\hat{\mathbf{A}}] \propto \bar{\mathbf{A}}$ であれば, 真の解 $\bar{\mathbf{u}}_n$ に対して

$$E[(\bar{\mathbf{u}}_n, \hat{\mathbf{A}}\bar{\mathbf{u}}_n)] = (\bar{\mathbf{u}}_n, E[\hat{\mathbf{A}}]\bar{\mathbf{u}}_n) \propto (\bar{\mathbf{u}}_n, \bar{\mathbf{A}}\bar{\mathbf{u}}_n) = 0 \quad (16)$$

となる. そこで $(\bar{\mathbf{u}}_n, \hat{\mathbf{A}}\bar{\mathbf{u}}_n) = (\bar{\mathbf{u}}_n, \mathbf{A}\bar{\mathbf{u}}_n) - c(\bar{\mathbf{u}}_n, \mathbf{B}\bar{\mathbf{u}}_n) = 0$ となるように定数 c を選ぶ. そのような $\bar{\mathbf{u}}_n$ と c とが一意的でなければ c が最小になるように選ぶ. このことは

$$\min_{\substack{(\mathbf{u}, \mathbf{A}\mathbf{u}) \\ (\mathbf{u}, \mathbf{B}\mathbf{u})}} \quad (17)$$

を意味する. よく知られているように, この解は「一般固有値問題」

$$\mathbf{A}\mathbf{u} = c\mathbf{B}\mathbf{u} \quad (18)$$

の最小一般固有値 c に対する単位一般固有ベクトル $\bar{\mathbf{u}}_n$ によって与えられる. これもよく知られているように, (18)の解は \mathbf{B} (正値対称行列) の「スペクトル分解⁹⁾」を $\mathbf{B} = \sum_{\alpha=1}^N \sigma_{\alpha} \mathbf{r}_{\alpha} \mathbf{r}_{\alpha}^T$ ($\{\mathbf{r}_{\alpha}\}$ は \mathbf{B} の各固有値 σ_{α} に対する単位固有ベクトルのつくる正規直交系) とし, $\mathbf{T} = \sum_{\alpha=1}^N \mathbf{r}_{\alpha} \mathbf{r}_{\alpha}^T / \sqrt{\sigma_{\alpha}}$ とおけば, 普通の固有値問題 $\hat{\mathbf{A}}\bar{\mathbf{u}} = c\bar{\mathbf{u}}$ に帰着する. ただし $\hat{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^T$ (対称行列), $\mathbf{u} = \mathbf{T}\bar{\mathbf{u}}$ である.

4. くりこみ法

3章で指摘したように, 最適重み W_{α} は真の解 $\bar{\mathbf{u}}_n$ を含んでいるので, 推定値 $\hat{\mathbf{u}}_n$ で代用し, 上に述べた

一般固有値問題を解き, その解 $\hat{\mathbf{u}}_n'$ を新たな推定値として反復すればよい. しかし, その代わりに一般固有値問題を反復によって解いて, その反復の過程で最適重みを更新することを考える. 行列 $\hat{\mathbf{A}} = \mathbf{A} - c\mathbf{B}$ の最小固有値 λ_n に対する単位固有ベクトルを \mathbf{u}_n とすると, $(\mathbf{u}_n, \hat{\mathbf{A}}\mathbf{u}_n) \neq 0$ なら

$$\hat{\mathbf{A}}' = \hat{\mathbf{A}} - \frac{(\mathbf{u}_n, \hat{\mathbf{A}}\mathbf{u}_n)}{(\mathbf{u}_n, \mathbf{B}\mathbf{u}_n)} \mathbf{B} \quad (19)$$

と置けば $(\mathbf{u}_n, \hat{\mathbf{A}}'\mathbf{u}_n) = 0$ となる. $(\mathbf{u}_n, \hat{\mathbf{A}}\mathbf{u}_n) = \lambda_n$ に注意し, 最適重みの更新を含めれば, この反復は次のように記述できる.

1. $c=0, W_{\alpha}=1$ と置く.
2. 次の行列 \mathbf{A}, \mathbf{B} を計算する.

$$\mathbf{A} = \sum_{\alpha} W_{\alpha} \mathbf{a}_{\alpha} \mathbf{a}_{\alpha}^T, \quad \mathbf{B} = \sum_{\alpha} W_{\alpha} V_0[\mathbf{a}_{\alpha}] \quad (20)$$

3. 行列

$$\hat{\mathbf{A}} = \mathbf{A} - c\mathbf{B} \quad (21)$$

の最小固有値を λ_n とし, 対応する単位固有ベクトルを \mathbf{u}_n とする. λ_n が十分 0 に近ければ, \mathbf{u}_n を返して終了する.

4. 定数 c と重み W_{α} を次のように更新し, ステップ 2 へ戻る.

$$c \leftarrow c + \frac{\lambda_n}{(\mathbf{u}_n, \mathbf{B}\mathbf{u}_n)} \quad (22)$$

$$W_{\alpha} \leftarrow \frac{1}{(\mathbf{u}_n, V_0[\mathbf{a}_{\alpha}]\mathbf{u}_n)} \quad (23)$$

現在の推定値 $\hat{\mathbf{A}}$ の最小固有値 λ_n に対する単位固有ベクトルが \mathbf{u}_n のとき, 重みの変化を考えなければ, 行列 $\hat{\mathbf{A}}$ は次のステップで

$$\hat{\mathbf{A}}' = \hat{\mathbf{A}} - \frac{\lambda_n}{(\mathbf{u}_n, \mathbf{B}\mathbf{u}_n)} \mathbf{B} \quad (24)$$

に更新される. 摂動定理により $\hat{\mathbf{A}}'$ の最小固有値 λ_n' は

$$\lambda_n' = \lambda_n - \left(\mathbf{u}_n, \left(\frac{\lambda_n}{(\mathbf{u}_n, \mathbf{B}\mathbf{u}_n)} \mathbf{B} \right) \mathbf{u}_n \right) + O\left(\left(\frac{\lambda_n}{(\mathbf{u}_n, \mathbf{B}\mathbf{u}_n)} \mathbf{B} \right)^2 \right) = O(\lambda_n^2) \quad (25)$$

となる. これは λ_n が 0 に “2次収束” し, ニュートン法によって方程式の解を求めるのと同様に 3~4 回の反復で十分であることを意味する. 重みの更新を含めると, 厳密な解析は難しい. しかし, 多くの実験によれば収束は速く 3~4 回の反復で十分である.

5. 信頼範囲と問題点

\hat{A} の最小固有値はくりこみ法が収束したとき0であり、その固有ベクトル \mathbf{u}_n が解となる。残りの固有値を $\lambda_1, \dots, \lambda_{n-1}$ とし、対応する単位固有ベクトルの正規直交系を $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ とすると、解 \mathbf{u}_n の共分散行列は

$$V[\mathbf{u}_n] = \varepsilon^2 \sum_{i=1}^{n-1} \frac{\mathbf{u}_i \mathbf{u}_i^T}{\lambda_i} \quad (26)$$

となることが証明できる (文献17), 34) の議論を拡張すればよいが、詳細省略)。したがって、解 \mathbf{u}_n の信頼範囲を各軸 \mathbf{u}_i 方向の半径が $\varepsilon/\sqrt{\lambda_i}$ の楕円体と見積ることができる。さらに、単に信頼範囲を示すだけでなく²²⁾、この信頼範囲を最小にするようにシステムを最適に設計したり²⁴⁾、推定した \mathbf{u} が理論的に予想される値 $\bar{\mathbf{u}}$ の誤差を含んだサンプル値とみなせるかどうかを、 \mathbf{u} が $\bar{\mathbf{u}}$ の一定の信頼水準を表す楕円体の中に入るかどうかで検定できる。これによって幾何学的な構造に関する仮説を定量的な χ^2 検定に帰着させることができる¹⁸⁾。

共分散行列 $V_0[\mathbf{a}_\alpha]$ は一般に \mathbf{a}_α の関数である。これまでの議論ではそれを真の値 $\bar{\mathbf{a}}_\alpha$ で評価したものと仮定しているが、これをデータ $\mathbf{a}_\alpha = \bar{\mathbf{a}}_\alpha + \Delta\mathbf{a}_\alpha$ で置き換えた $V_0[\mathbf{a}_\alpha]$ を用いると $O(\varepsilon^2)$ の誤差を生じる。したがって、最適適みにも $O(\varepsilon^2)$ の誤差を生じるが、重みを除けば \mathbf{A} に対する誤差は $O(\varepsilon^4)$ である。

一方、問題によっては特定の \mathbf{u}_n と $\bar{\mathbf{a}}_\alpha$ について $(\mathbf{u}_n, V_0[\mathbf{a}_\alpha]\mathbf{u}_n) = 0$ となり、したがって $W_\alpha = \infty$ になることもある。これは誤差 $\Delta\mathbf{a}_\alpha$ が \mathbf{a}_α のある真線形部分空間に拘束され、 \mathbf{u}_n がそれと直交するときに生じる。原因は式(9)の展開で第1近似をとったためであり、その場合は ε^4 の項まで展開する。

以上の議論では、 \mathbf{a}_α の共分散行列 $V[\mathbf{a}_\alpha]$ を ε で展開した各項がわかっているとしたが、これは自明なことではない。たとえば各 \mathbf{a}_α が画像上の特徴点の複雑な関数であれば、特徴点の誤差の挙動から $V_0[\mathbf{a}_\alpha]$ を計算するには複雑な解析を要する。データが画像上のエッジに当てはめた直線であれば解析はさらに複雑になる。

6. 消失点、出現点の推定

よく知られているように、空間中の平行線の投影像を画像上で延長すると、共通の“消失点”で交わり、その消失点の位置は空間中での平行線の3次元方向を

表す^{9), 17)}。したがって、画像上で消失点を検出することは画像の3次元解釈の手がかりとなる。とくにカメラキャリブレーションではこれが重要な役割を果たす^{2), 3), 6), 20), 21), 24), 35)~37)}。

しかし、画像に誤差があれば平行線の投影像は厳密に一点では交わらないので最小二乗法で推定する。各直線は画像処理によって検出したエッジ画素に最小二乗法であてはめられる。当てはめの精度は、各エッジ画素の統計的挙動をモデル化して理論的に導くことができる^{17), 18), 33), 34)}。このような解析によれば、消失点推定のくりこみ法は次の形をとる³⁴⁾。ただし入力エッジに当てはめた N 本の直線の“ N ベクトル^{9), 10), 17)}” $\{\mathbf{n}_\alpha\}$, $\alpha=1, \dots, N$, 出力は消失点の N ベクトル \mathbf{m} である。以下、本論文での定数 f は透視変換の“焦点距離^{9), 10)}”である。

1. $c=0$, $W_\alpha=1$, $\alpha=1, \dots, N$ と置く。
2. “不偏モーメント行列”

$$\hat{N} = \sum_{\alpha=1}^N W_\alpha \left(\mathbf{n}_\alpha \mathbf{n}_\alpha^T - c \left(\mathbf{u}_\alpha \mathbf{u}_\alpha^T + \frac{w_\alpha^2}{12f^2} \mathbf{m}_{G\alpha} \mathbf{m}_{G\alpha}^T \right) \right) \quad (27)$$

の最小固有値 λ_m に対する単位固有ベクトル \mathbf{m} を計算する。 λ が十分0に近ければ、 \mathbf{m} を返して終了する。

3. c , W_α を次のように更新し、ステップ2へ戻る。

$$c \leftarrow c + \frac{\lambda_m}{\sum_{\alpha=1}^N W_\alpha ((\mathbf{u}_\alpha, \mathbf{m})^2 + \frac{w_\alpha^2}{12f^2} (\mathbf{m}_{G\alpha}, \mathbf{m})^2)} \quad (28)$$

$$W_\alpha \leftarrow \frac{1}{(\mathbf{u}_\alpha, \mathbf{m})^2 + \frac{w_\alpha^2}{12f^2} (\mathbf{m}_{G\alpha}, \mathbf{m})^2} \quad (29)$$

ただし w_α は α 番目のエッジの長さ (単位は画素) であり、 \mathbf{u}_α , $\mathbf{m}_{G\alpha}$ はそれぞれその方向および中心点の N ベクトルである。

物体あるいはカメラが空間中を平行移動すると、物体の複数の特徴点の画像上での軌跡は延長すると共通の“出現点”で交わり、その画像上の位置は物体あるいはカメラの3次元並進方向を表す^{9), 17)}。その幾何学的解釈は消失点の場合と同一であり、出現点を計算することは動画像の3次元解釈の手がかりとなる。またカメラキャリブレーションでも重要な役割を果たす²²⁾。

出現点推定は複数の直線の共通の交点の計算という

意味で消失点推定と形式的に同じであるが、誤差の構造が異なる。消失点の場合は直線が“エッジ”に当てはめられるのに対し、出現点の場合は異なる時刻の画像上の対応する“特徴点”を結んで得られるからである。理論的な解析によれば、出現点推定のくりこみ法は次の形をとる³⁴⁾。ただし、入力は N 本の軌跡の N ベクトル $\{\mathbf{n}_\alpha\}$, $\alpha=1, \dots, N$, 出力は出現点の N ベクトル \mathbf{m} である。

1. $c=0$, $W_\alpha=1$, $\alpha=1, \dots, N$ と置く。
2. “不偏モーメント行列”

$$\hat{N} = \sum_{\alpha=1}^N W_\alpha \left(\mathbf{n}_\alpha \mathbf{n}_\alpha^T - c \left(\frac{\mathbf{u}_\alpha \mathbf{u}_\alpha^T}{1 - (\mathbf{m}_{1\alpha}, \mathbf{m}_{2\alpha})} + \frac{\mathbf{m}_{G\alpha} \mathbf{m}_{G\alpha}^T}{1 + (\mathbf{m}_{1\alpha}, \mathbf{m}_{2\alpha})} \right) \right) \quad (30)$$

の最小固有値 λ_m に対する単位固有ベクトル \mathbf{m} を計算する。 λ が十分 0 に近ければ、 \mathbf{m} を返して終了する。

3. c , W_α を次のように更新し、ステップ 2 へ戻る。

$c \leftarrow c$

$$c \leftarrow c + \frac{\lambda_m}{\sum_{\alpha=1}^N W_\alpha \left(\frac{(\mathbf{u}_\alpha, \mathbf{m})^2}{1 - (\mathbf{m}_{1\alpha}, \mathbf{m}_{2\alpha})} + \frac{(\mathbf{m}_{G\alpha}, \mathbf{m})^2}{1 + (\mathbf{m}_{1\alpha}, \mathbf{m}_{2\alpha})} \right)} \quad (31)$$

$$W_\alpha \leftarrow \frac{1}{\frac{(\mathbf{u}_\alpha, \mathbf{m})^2}{1 - (\mathbf{m}_{1\alpha}, \mathbf{m}_{2\alpha})} + \frac{(\mathbf{m}_{G\alpha}, \mathbf{m})^2}{1 + (\mathbf{m}_{1\alpha}, \mathbf{m}_{2\alpha})}} \quad (32)$$

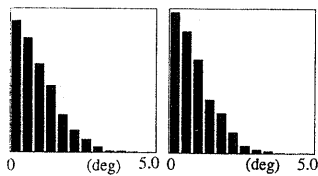
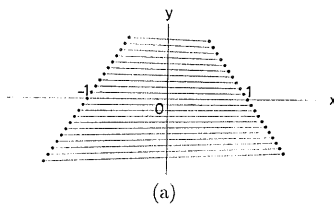


図 1 出現点推定の誤差
(a) 軌跡の例。(b) くりこみ法を用いない場合。
(c) くりこみ法を用いる場合。

Fig. 1 Focus of expansion estimation.
(a) simulated trajectories; (b) without renormalization; (c) with renormalization.

ただし、 $\mathbf{m}_{1\alpha}$, $\mathbf{m}_{2\alpha}$ は α 番目の特徴点の 2 時刻での N ベクトルであり、 \mathbf{u}_α , $\mathbf{m}_{G\alpha}$ はそれぞれそれらを結ぶ軌跡の方向および 2 特徴点の midpoint の N ベクトルである。

例 1 図 1 の (a) の軌跡を考え、各端点の座標に正規乱数によって独立に誤差を与え、毎回異なる乱数を用いて出現点の N ベクトルと真の N ベクトルとのなす角度のヒストグラムを作る。(b) は (最適重み) 最小二乗法による結果、(c) はくりこみ法による結果である。偏差がやや減少し、精度が改善されている。

7. コニック当てはめ

多くの工業製品には円形部分や球形部分があり、それらは画像上では楕円、一般には“コニック”に投影される。このため画像中のコニックを検出することによって物体の認識や分類の手がかりが得られる。また、対応するコニックの空間中の位置や向きを計算することもできる^{5), 12), 17), 23), 27), 29), 30)}。しかし、そのためには画像上に検出されたエッジ画素の列に正確にコニックを当てはめなければならない。そのための技法がいろいろ発表されているが^{1), 4), 26), 28), 31)}、最適重みの議論はなされても基本的には最小二乗法であり、統計的偏差が存在する⁷⁾。この偏差はくりこみ法によって除去できる⁸⁾。

各エッジ画素の誤差が等方かつ独立で微小とすれば、摂動定理に基づくテンソル解析^{7), 8), 17)}によりくりこみ法は次の形をとる。入力はコニックを当てはめる画素の N ベクトル $\{\mathbf{m}_\alpha\}$, $\alpha=1, \dots, N$, 出力はコニックを表示する対称行列 \mathbf{Q} である¹²⁾ (\mathbf{I} は単位行列)。また $m_{\alpha(i)}$ はベクトル \mathbf{m}_α の第 i 成分であり、 δ_{ij} はクロネッカのデルタである。

1. $c=0$, $\mathbf{Q} = -\mathbf{I}/\sqrt{3}$ と置く。

2. $W_\alpha = 1/\|\mathbf{Q}\mathbf{m}_\alpha\|^2$ と置き、“モーメント行列” $\mathbf{M} = (M_{ij})$ と“不偏モーメントテンソル” $\hat{\mathbf{M}} = (\hat{M}_{ijkl})$ を次のように計算する。

$$M_{ij} = \sum_{\alpha=1}^N W_\alpha m_{\alpha(i)} m_{\alpha(j)} \quad (33)$$

$$\hat{M}_{ijkl} = \sum_{\alpha=1}^N (W_\alpha m_{\alpha(i)} m_{\alpha(j)} m_{\alpha(k)} m_{\alpha(l)} - \frac{c}{2} (\delta_{ij} m_{\alpha(k)} m_{\alpha(l)} + \delta_{ik} m_{\alpha(j)} m_{\alpha(l)} + \delta_{il} m_{\alpha(k)} m_{\alpha(j)} + \delta_{jk} m_{\alpha(i)} m_{\alpha(l)} + \delta_{jl} m_{\alpha(i)} m_{\alpha(k)} + \delta_{kl} m_{\alpha(i)} m_{\alpha(j)})) \quad (34)$$

3. テンソル \hat{M} の最小固有値に対する単位固有行列を Q とし, 最小固有値を λ_m とする. λ_m が十分 0 に近ければ, Q を返して終了する.
4. 定数 c を次のように更新し, ステップ 2 へ戻る.

$$c \leftarrow c + \frac{\lambda_m}{\text{tr } Q \text{tr}(MQ) + 2\text{tr}(MQ^2)} \quad (35)$$

ただし, テンソル \hat{M} の “単位固有行列” とは $\hat{M}Q = \lambda Q$ となる $\|Q\|=1$ の対称行列のことをいい, λ をその “固有値” という. ここに $\hat{M}Q$ はその (ij) 要素が $\sum_{k,l=1}^3 \hat{M}_{ijkl} Q_{kl}$ である対称行列であり, $\|Q\|^2 = \sum_{i,j=1}^3 Q_{ij}^2$ である. テンソルの固有値, 固有行列はテンソルと行列をそれぞれ 6次元行列, 6次元ベクトルとみなし, 固有値, 固有ベクトルを計算することによって容易に求まる^{7),8)}.

例 2 図 2 の (a) のように楕円の上半分に分点をと, その x, y 座標に独立に正規乱数によって誤差を加え, コニックを当てはめる. これを毎回異なる

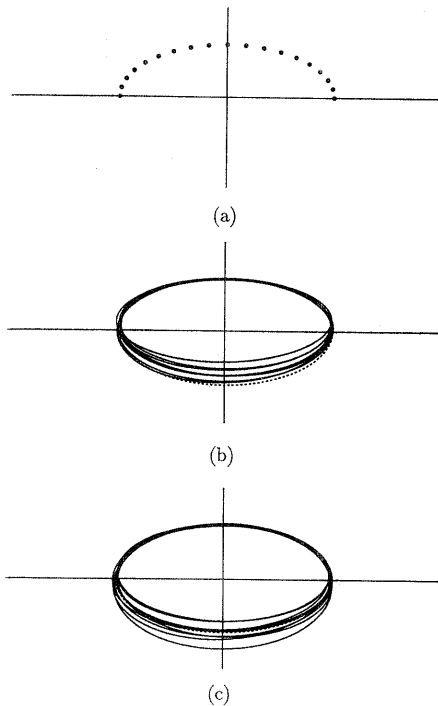


図 2 コニック当てはめ
(a) 楕円の上半分上の分点. (b) くりこみ法を用いない場合. (c) くりこみ法を用いる場合.

Fig. 2 Conic fitting.

(a) points on an upper-half ellipse; (b) without renormalization; (c) with renormalization;

乱数を用いて繰り返す. (b) は単純な (最適重み) 最小二乗法で当てはめた例であり (10 個), 点線が真のコニックである. 解が扁平な方向に偏っていることがわかる. (c) はくりこみ法を用いて当てはめた例であり, 偏差が除去され, 真の値の周りに分布していることがわかる.

8. 3次元運動解析

物体またはカメラが空間中で剛体運動するとき, 物体の特徴点の運動前後の画像間での対応が 8 組以上求まれば, 物体またはカメラの 3次元運動と各特徴点までの距離が計算できる^{13),15),17)}. しかし, これまで研究されたのは基本的には最小二乗法であり³⁰⁾, 解には統計的偏差が存在する^{16),17),25)}. この偏差はくりこみ法によって除去できる.

画像上の各特徴点の誤差が等方かつ独立で微小とすれば, 摂動定理に基づくテンソル解析^{16),17),32)}により, くりこみ法は次の形をとる³²⁾. 入力の特徴点の運動前後の N ベクトル $\{m_\alpha\}, \{m'_\alpha\}$, $\alpha=1, \dots, N$, 出力はカメラの運動を指定する回転行列 R , 並進ベクトル h (単位ベクトル), および運動前後のカメラのレンズの中心から各特徴点までの距離 $\{r_\alpha\}, \{r'_\alpha\}$, $\alpha=1, \dots, N$ である.

1. $c=0, W_\alpha=1, \alpha=1, \dots, N$ と置く.
2. “モーメント行列” $M=(M_{ij}), M'=(M'_{ij})$ と “不偏モーメントテンソル” $\hat{M}=(\hat{M}_{ijkl})$ を次のように計算する.

$$M_{ij} = \sum_{\alpha=1}^N W_\alpha m_{\alpha(i)} m_{\alpha(j)},$$

$$M'_{ij} = \sum_{\alpha=1}^N W_\alpha m'_{\alpha(i)} m'_{\alpha(j)} \quad (36)$$

$$\hat{M}_{ijkl} = \sum_{\alpha=1}^N W_\alpha (m_{\alpha(i)} m_{\alpha(j)} m'_{\alpha(k)} m'_{\alpha(l)} - c(\delta_{ji} m_{\alpha(i)} m_{\alpha(k)} + \delta_{ik} m_{\alpha(j)} m'_{\alpha(l)})), \quad (37)$$

3. テンソル \hat{M} の最小固有値に対する $\|G\|=\sqrt{2}$ の固有行列を G とし, 最小固有値を λ_m とする. λ_m が十分 0 に近ければ, G を返して終了する.
4. 定数 c と重み W_α を次のように更新し, ステップ 2 へ戻る.

$$c \leftarrow c + \frac{\lambda_m}{\text{tr}(G^T M G) + \text{tr}(G M' G^T)} \quad (38)$$

$$W_\alpha \leftarrow \frac{1}{\|G m'_\alpha\|^2 + \|G^T m_\alpha\|^2} \quad (39)$$

“基本行列” G が求まれば, “特異値分解” によって

“運動パラメータ” $\{R, t\}$ が計算できる^{13), 15), 17), 38)}.
また、運動前後の各特徴点までの距離 $\{r_a\}, \{r_a'\}$ も計算できる。

ただし、カメラがある特徴点に向かって移動する場合はその特徴点の重みが発散するので、 ϵ^4 の項まで展開する（詳細省略）。これは、カメラの移動によってもその特徴点の視差が生じないことによる。一般に物体方向にカメラ移動が移動すると、生じる視差が少なく、3次元復元の精度が低下する。

例3 図3の(a), (b)は空間中の立方体領域に特徴点をランダムに選んだときの、カメラの移動前後のシミュレーション画像である。運動前後の各特徴点の x, y 座標に独立に正規乱数によって誤差を加え、運動を解析する。これを毎回異なる乱数を用いて

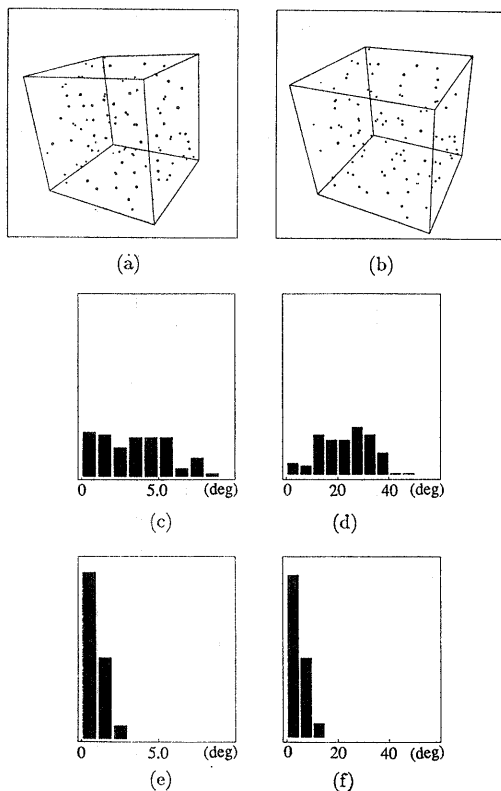


図3 3次元運動解析

(a), (b) 運動前後の画像。(c), (d) くりこみ法を用いない場合の回転と並進の誤差。(e), (f) くりこみ法を用いる場合

Fig. 3 3-D motion analysis.

(a), (b) images before and after motion; (c), (d) errors in rotation and translation without renormalization; (e), (f) with renormalization.

繰り返す。(c), (d)はくりこみ法を用いない場合（最適重みは用いる）の回転と並進の誤差のヒストグラムである。ただし回転の誤差は回転角の差の絶対値、並進の誤差は並進方向のくい違いの角度で測る。明らかに統計的偏差が存在する。(e), (f)はくりこみ法を適用した場合であり、偏差が減少して精度が向上している。

9. まとめ

画像データを解析するコンピュータビジョンの多くの問題において、解は最小二乗法によってロバストに求まり、計算も単純である。本論文では、もともとなる画像データに誤差があれば、最小二乗の重みを最適に選んでも解には一般に統計的偏差が存在することを指摘し、誤差の大きさを最小二乗の残差から推定しながら反復によって偏差を除去する「くりこみ法」と呼ぶ手法を一般的な形式に記述し、さまざまな理論的考察を行った。そして消失点や出現点の推定、コニックの当てはめ、3次元運動解析に適用した例を示し、画像の誤差についての情報がなくても精度が向上することを確認した。

本論文に示した応用例からもわかるように、くりこみ法を適用するには解の偏差が“理論式”として求まっていなければならない。コンピュータビジョンの多くの問題は複雑な非線形構造をもっているため、画像の誤差モデルから解の誤差の挙動を解析的に記述するためには、誤差モデルの理想化、ベクトル解析、テンソル解析、漸近展開、摂動定理、線形近似等の解析手法が要求される。実際その理論解析に99%の努力が費やされるといってもよい。

このような理論解析のためには何段階かの理想化と近似が必要となる。その結果として、装置の特殊性に起因する歪みのような理論化が困難な誤差を無視することになり、記述された誤差の挙動は実際の誤差と厳密には一致しないこともある。しかし、近似的にも誤差が記述できて、それを用いて精度が改善されれば、それは工学的成果であろう。また、それは仮定した性質をもつ誤差を補正したことを意味するから、精度が向上しない場合は、理論化で無視した特殊な誤差の存在を示唆している。このことから、まず一般的と考えられる誤差を理論解析して、くりこみ法で精度の改善を計るのは妥当なアプローチであると思われる。

謝辞 本研究に関して有益なご討論をいただいた都立大学の田川憲男助手、数値実験を行っていただいた

群馬大学大学院生の浦沢康二氏(現沖電気), 岩崎利夫氏(現東芝), および武田佐知男氏に感謝いたします。

参 考 文 献

- 1) Bookstein, F.L.: Fitting Conic Sections to Scattered Data, *Comput. Graphics Image Process.*, Vol. 9, No. 1, pp. 56-71 (1979).
- 2) Caprile, B. and Torre, V.: Using Vanishing Points for Camera Calibration, *Int. J. Comput. Vision*, Vol. 4, pp. 127-140 (1990).
- 3) Echigo, T.: A Camera Calibration Technique Using Three Sets of Parallel Lines, *Machine Vision Appl.*, Vol. 3, pp. 159-167 (1990).
- 4) Ellis, T., Abboot, A. and Brillault, B.: Ellipse Detection and Matching with Uncertainty, *Image Vision Comput.*, Vol. 10, pp. 271-276 (1992).
- 5) Forsyth, D., Mundy, J.L., Zisserman, A., Coelho, C., Heller, A. and Rothwell, C.: Invariant Descriptors for 3-D Object Recognition and Pose, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, pp. 971-991 (1991).
- 6) Grosky, W. I. and Tamburino, L. A.: A Unified Approach to the Linear Camera Calibration Problem, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, pp. 663-671 (1990).
- 7) 岩崎利夫, 金谷健一: コニックあてはめの統計的偏差と不偏推定法, 情報処理学会研究会報告, 92-CV-77 (1992).
- 8) 岩崎利夫, 丸山 保, 金谷健一: くりこみ法によるコニック当てはめ, 情報処理学会研究会報告, 92-CV-79 (1992).
- 9) 金谷健一: 画像理解/3次元認識の数理, 森北出版 (1990).
- 10) Kanatani, K.: Computational Projective Geometry, *CVGIP: Image Understanding*, Vol. 54, No. 3, pp. 333-348 (1991).
- 11) Kanatani, K.: Hypothesizing and Testing Geometric Properties of Image Data, *CVGIP: Image Understanding*, Vol. 54, No. 3, pp. 349-357 (1991).
- 12) 金谷健一: コニックの計算論と直交性解釈, 情報処理学会論文誌, Vol. 33, No. 6, pp. 817-824 (1992).
- 13) 金谷健一: 3次元運動のコンピュータシミュレーション, 光学, Vol. 21, No. 10, pp. 650-656 (1992).
- 14) Kanatani, K.: Statistical Analysis of Focal-length Calibration Using Vanishing Points, *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 6, pp. 767-775 (1992).
- 15) 金谷健一: 運動の3次元コンピュータシミュレーション, 情報処理学会研究会報告, 93-CV-81 (1993).
- 16) Kanatani, K.: Unbiased Estimation and Statistical Analysis of 3-D Rigid Motion from Two Views, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 15, No. 1, pp. 37-50 (1993).
- 17) Kanatani, K.: *Geometric Computation for Machine Vision*, Oxford Univ. Press, Oxford (1993).
- 18) 金谷健一: 画像の3次元解釈の統計的信頼性, 情報処理学会論文誌, Vol. 34, No. 10, pp. 2062-2070 (1993).
- 19) 金谷健一, 川島孝夫: 画像データの幾何学的構造の仮説検定, 情報処理学会論文誌, Vol. 31, No. 6, pp. 879-886 (1990).
- 20) Kanatani, K. and Onodera, Y.: Anatomy of Camera Calibration Using Vanishing Points, *IEICE Trans. Infor. Sys.*, Vol. 74, No. 10, pp. 3369-3378 (1991).
- 21) 金谷健一, 浦沢康二: 消失点とカメラキャリブレーションの統計的信頼性, 日本ロボット学会誌, Vol. 9, No. 7, pp. 813-820 (1991).
- 22) 金谷健一, 浦沢康二: 画像原点推定のためのカメラキャリブレーション, 情報処理学会研究会報告, 93-CV-81 (1993).
- 23) 劉 武, 金谷健一: コニックの3次元解釈とその応用, 情報処理学会研究会報告, 92-CV-76 (1992).
- 24) 丸山 保, 金谷健一: カメラの焦点距離の最適キャリブレーションシステム, 情報処理学会研究会報告, 93-CV-82 (1993).
- 25) 森尻智昭, 金谷健一: 3次元運動の不偏推定法と統計的誤差解析, 情報処理学会研究会報告, 92-CV-77 (1992).
- 26) Porrill, J.: Fitting Ellipses and Predicting Confidence Envelopes Using a Bias Corrected Kalman Filter, *Image Vision Comput.*, Vol. 8, No. 1, pp. 37-41 (1990).
- 27) Rothwell, C. A., Zisserman, A., Marinos, C. I., Forsyth, D. A. and Mundy, J. L.: Relative Motion and Pose from Arbitrary Plane Curves, *Image Vision Comput.*, Vol. 10, No. 4, pp. 250-262 (1992).
- 28) Safaei-Rad, R., Tchoukanov, I., Benhabib, B. and Smith, K. C.: Accurate Parameter Estimation of Quadratic Curves from Grey-level Images, *CVGIP: Image Understanding*, Vol. 54, No. 2, pp. 259-274 (1991).
- 29) Safaei-Rad, S., Tchoukanov, I., Smith, K. C., and Benhabib, B.: Constraints on Quadratic-curved Features under Perspective Projection, *Image Vision Comput.*, Vol. 10, No. 8, pp. 532-548 (1992).
- 30) Safaei-Rad, R., Tchoukanov, I., Smith, K. C. and Benhabib, B.: Three-dimensional Location Estimation of Circular Features for Machine Vision, *IEEE Trans. Robotics Automation*, Vol. 8, No. 5, pp. 624-640 (1992).

- 31) Sampson, P.D.: Fitting Conic Sections to "Very Scattered" Data: An Iterative Refinement of the Bookstein Algorithm, *Comput. Graphics Image Process.*, Vol. 18, No. 1, pp. 97-108 (1982).
- 32) 武田佐知男, 金谷健一: くりこみ法による3次元運動の解析, 情報処理学会研究報告, 93-CV-82 (1993).
- 33) 浦沢康二, 金谷健一: 幾何学的計算の統計解析: I. 基礎理論, 情報処理学会研究報告, 92-CV-77 (1992).
- 34) 浦沢康二, 金谷健一: 幾何学的計算の統計解析: II. エッジ, 消失点, 出現点, 情報処理学会研究報告, 92-CV-78 (1992).
- 35) 浦沢康二, 金谷健一: 焦点距離キャリブレーションの統計的信頼性, 情報処理学会研究報告, 92-CV-79 (1992).
- 36) Wang, C.-C.: Extrinsic Calibration of a Vision Sensor Mounted on a Robot, *IEEE Trans. Robotics Automation*, Vol. 8, No. 2, pp. 161-175 (1992).
- 37) Wang, L.-L. and Tsai, W.-H.: Camera Calibration by Vanishing Lines for 3-D Computer Vision, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, No. 4, pp. 370-376 (1991).
- 38) Weng, J., Huang, T.S. and Ahuja, N.:

Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation, *IEEE Trans. Patt. Anal. Machine Intell.*, Vol. 11, pp. 451-467 (1989).

(平成5年6月24日受付)

(平成5年10月14日採録)



金谷 健一 (正会員)

1947年岡山県生。1969~1970年米国 Case Western Reserve 大学留学。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。同年群馬大学工学部情報工学科助手。1983年同助教授。1988年同教授。米国Maryland大学, デンマークCopenhagen 大学, 英国 Oxford 大学客員研究員, 著書「線形代数」(講談社), "Group-Theoretical Methods in Image Understanding" (Springer), 「画像理解」(森北出版), "Geometric Computation for Machine Vision" (Oxford Univ. Press). 1987年情報処理学会論文賞受賞。