

# 情報量基準による幾何学的モデルの選択

金 谷 健 一†

ロボットが画像やセンサデータから3次元環境のモデルを構築するための幾何学的推定問題を一般的に記述し、予測能力の高いモデルを恣意的な有意水準を用いずに選択する基準として「幾何学的AIC」を導出する。そして、包含関係にあるモデルに対しては、その良さがデータの誤差の大きさに関する知識なしに評価できることを示し、これを動画像の3次元解析に応用する。

## Geometric Model Selection by an Information Criterion

KENICHI KANATANI†

The problem of selecting a geometric model for robotics applications, such as building a model of the environment from image and sensor data, is expressed in general terms, and the *geometric AIC criterion* is derived as a measure for selecting a model that has high predicting capability without using any empirically adjustable significance level. We then show that the goodness of two models can be compared without any knowledge of the magnitude of the noise in the data if one model includes the other. Our method is applied to 3-D motion analysis from two views.

### 1. はじめに

ロボットが自律的に作業するには画像やセンサデータから3次元環境モデルを構築しなければならないが、画像やセンサデータには誤差がある。これに対する基本的な方法は2次元および3次元物体のモデルをあてはめることである。したがって、

- (1) 適切なモデルを選択する、
- (2) モデルを最適にあてはめる、

の2段階が必要である。これはより抽象的なレベルにもあてはまる。たとえば動画像から3次元構造を復元することは動画像に「それが3次元空間を移動する物体の投影像である」というモデルをあてはめる操作である<sup>2),3),14),18)</sup>。第2段階のモデルあてはめは「統計的推論」として数学的に定式化されているが<sup>8)~10)</sup>、第1段階のモデル選択としては経験的な推論規則がよく用いられる。これはロボットの作業空間では候補モデルが限定されていて、簡単な特徴抽出のみから容易に判定できると期待されるためであるが、十分に研究が進んではいない。本論文はこれに統計学的な基礎を与えるものである。

統計学ではモデル選択の基準として **AIC** がよく利

用される。これは最尤推定量の将来のデータに対する予測能力でモデルの良さを測ろうとする立場であり<sup>1),15)</sup>、

(データとモデルのくい違いの程度)  
+ (モデルのパラメータ数の影響)

と説明され、「データが同程度に説明できる限り、パラメータ数の少ないモデルを選ぶ」と解釈されている。しかし、これは「統計学的モデル」に対する解釈であり、「幾何学的モデル」には必ずしもあたらない。たとえば3次元空間の点集合に

- 平面をあてはめる、
- 直線をあてはめる、

という2つのモデルを比較するとする (図1)。最小二

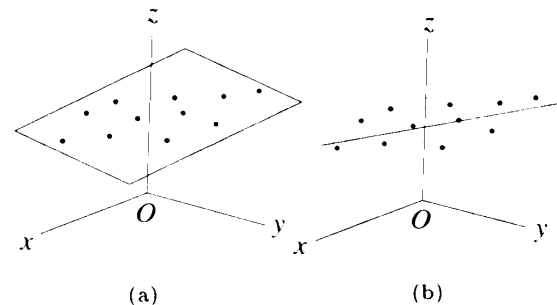


図1 空間の点集合に (a) 平面をあてはめる、(b) 直線をあてはめる。

Fig. 1 Fitting for points in space: (a) a plane is fitted; (b) a line is fitted.

† 群馬大学工学部情報工学科

Department of Computer Science, Gunma University

乗法であてはめた平面または直線からの残差をくい違いの尺度とすると、当然平面をあてはめたほうが残差がつねに少ない。一方、平面は3個のパラメータで指定できるが、直線は4個のパラメータを必要とするから<sup>6)</sup>、パラメータに関しても平面のほうが有利である。するとどんなデータでも（たとえ同一直線上にあっても）平面をあてはめるべきであろうか？

このパラドックスは、幾何学的モデルが多様体で記述されること、および攪乱母数の役割を無視したためである。本論文ではモデルのパラメータ数（自由度）は本質的な意味を持たず、その次元と余次元とが重要な役割を果たすことを示す。そして、これ考慮して、恣意的な有意水準を用いずにモデルを評価するAIC（幾何学的AICと呼ぶ）を導出する。さらに、包含関係にあるモデルに対しては、その良さがデータの誤差の大きさに関する知識なしに評価できることを示し、これを動画像の3次元解析に応用する。

## 2. 幾何学的推定

$m$ 次元空間 $\mathcal{R}^m$ に $N$ 個の点 $\mathbf{a}_1, \dots, \mathbf{a}_N$ が与えられているとする。これらはすべて $\mathcal{R}^m$ のある $m'$ 次元多様体 $\mathcal{A}$ 上に拘束されているとし、 $\mathcal{A}$ をデータ空間と呼ぶ。各点 $\mathbf{a}_\alpha$ はその真の位置 $\bar{\mathbf{a}}_\alpha$ から誤差 $\Delta\mathbf{a}_\alpha$ だけずれたものとし、

$$\mathbf{a}_\alpha = \bar{\mathbf{a}}_\alpha + \Delta\mathbf{a}_\alpha \quad (1)$$

と置く。誤差 $\Delta\mathbf{a}_\alpha$ は各 $\alpha$ について独立で、期待値 $\mathbf{0}$ 、共分散行列 $V[\mathbf{a}_\alpha]$ の正規分布に従う確率変数であるとする。各データはデータ空間 $\mathcal{A}$ に拘束されているから共分散行列 $V[\mathbf{a}_\alpha]$ は特異行列である。これは $\mathcal{A}$ の $\bar{\mathbf{a}}_\alpha$ における接空間 $T_{\bar{\mathbf{a}}_\alpha}(\mathcal{A})$ を値域とするランク $m'$ の半正値対称行列であるとする<sup>8),10)</sup>。

画像やセンサデータの精度をあらかじめ正確に推定することは実際上困難であるから、各 $V[\mathbf{a}_\alpha]$ を既知と仮定するのは現実的ではない。しかし、装置や処理アルゴリズムの特性から誤差の傾向（一様であるとか等方であるとか、ある場所や方向に出やすいなど）を定性的に予想するのは可能であることが多い。そこで各データの共分散行列 $V[\mathbf{a}_\alpha]$ が定数倍を除いて既知であると仮定する。すなわち、誤差の平均的な大きさを $\epsilon$ とし、

$$V[\mathbf{a}_\alpha] = \epsilon^2 V_0[\mathbf{a}_\alpha] \quad (2)$$

と表す。そして $\epsilon$ は未知であるが $\epsilon$ を1に正規化した行列 $V_0[\mathbf{a}_\alpha]$ は既知であるとし、 $\epsilon$ をノイズレベル、 $V_0[\mathbf{a}_\alpha]$ を正規化共分散行列と呼ぶ<sup>6)</sup>。

真の位置 $\bar{\mathbf{a}}_\alpha$ 、 $\alpha = 1, \dots, N$ はデータ空間 $\mathcal{A}$ のある部分多様体 $\mathcal{S}$ の上にあるとし、 $\mathcal{S}$ をモデルと呼

ぶ<sup>7),11)</sup>。次の問題を考える。

**【問題1】** データ $\{\mathbf{a}_\alpha\}$ からモデル $\mathcal{S}$ を最適に推定せよ。

## 3. モデルの残差平方和

モデル $\mathcal{S}$ は $n$ 次元ベクトル $\mathbf{u}$ をパラメータとして含むとする。 $\mathbf{u}$ の定義域は $\mathcal{R}^n$ のある $n'$ 次元部分多様体 $\mathcal{U}$ であるとし、これをパラメータ空間と呼ぶ。そして $\mathcal{U}$ の次元 $n'$ をモデル $\mathcal{S}$ の自由度と呼ぶ。また、多様体 $\mathcal{S}$ の次元が $p$ であるとき、 $r = m' - p$ と置き、 $p$ 、 $r$ をそれぞれモデル $\mathcal{S}$ の次元、余次元と呼ぶ<sup>8),10)</sup>。最適性の基準として最尤推定を採用すると、問題1は次のようになる。あるモデル $\mathcal{S}$ を仮定すると、観測したデータ $\{\mathbf{a}_\alpha\}$ は $\mathcal{S}$ 上にあるとは限らない。そこで各 $\mathbf{a}_\alpha$ を $\mathcal{S}$ 上の点 $\hat{\mathbf{a}}_\alpha$ に射影する（図2）。このような射影の仕方は無数に存在するが、その中で最尤推定の意味で最も確からしいものは、正規化共分散行列 $V_0[\mathbf{a}_\alpha]$ に関するマハラノビス距離の二乗和

$$S(\{\hat{\mathbf{a}}_\alpha\}; \mathcal{S}) = \sum_{\alpha=1}^N (\mathbf{a}_\alpha - \hat{\mathbf{a}}_\alpha, V_0[\mathbf{a}_\alpha]^{-1} (\mathbf{a}_\alpha - \hat{\mathbf{a}}_\alpha)) \quad (3)$$

を最小にする $\hat{\mathbf{a}}_\alpha$ である<sup>6)</sup>。ただし、本論文ではベクトル $\mathbf{a}$ 、 $\mathbf{b}$ の内積を $(\mathbf{a}, \mathbf{b})$ と表し、 $(\cdot)^{-1}$ は（ムーア・ペンローズ）一般逆行列（スペクトル分解し、零でない固有値をその逆数に置き換えたもの）を表す<sup>11)</sup>。式(3)の最小値はモデル $\mathcal{S}$ の関数であるから、それを $S(\mathcal{S})$ と書く。問題1の最尤推定解はこれを最小にするものである<sup>8)</sup>。その解を $\hat{\mathcal{S}}$ と書き、 $S(\hat{\mathcal{S}})$ の最小値 $S(\hat{\mathcal{S}})$ を残差平方和と呼ぶ。以上の関係を式で書くと次のようになる。

$$S(\hat{\mathcal{S}}) = \min_{\mathbf{u} \in \mathcal{U}} S(\mathcal{S}) = \min_{\mathbf{u} \in \mathcal{U}} \min_{\hat{\mathbf{a}}_\alpha \in \mathcal{S}} S(\{\hat{\mathbf{a}}_\alpha\}; \mathcal{S}) \quad (4)$$

このように $S(\hat{\mathcal{S}})$ を定義すると、 $S(\hat{\mathcal{S}})/\epsilon^2$ は第1近似において自由度 $rN - n'$ の $\chi^2$ 分布に従う確率変数

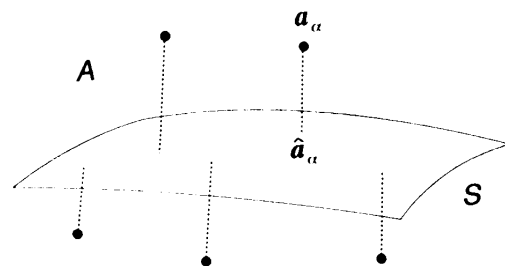


図2 データ $\mathbf{a}_\alpha$ を多様体 $\mathcal{S}$ 上に射影する。

Fig. 2 Datum  $\mathbf{a}_\alpha$  is projected onto manifold  $\mathcal{S}$ .

であることが示される<sup>11)</sup>。したがって二乗ノイズレベル  $\epsilon^2$  の不偏推定量が次のように得られる。

$$\hat{\epsilon}^2 = \frac{S(\hat{S})}{rN - n'} \quad (5)$$

#### 4. 幾何学的 AIC

モデル  $S$  の「良さ」の評価となるのは、まず観測データ  $\{a_\alpha\}$  を良く説明していることである。これは残差平方和  $S(\hat{S})$  が小さいことを意味する。しかし、自由度  $n'$  が大きいほど  $S(\hat{S})$  は小さくなり、特に  $rN$  個以上のパラメータを持つ多様体は一般にすべてのデータ点を通るように定めることができるので、残差平方和  $S(\hat{S})$  は 0 となる。このような人為的なモデルは良いモデルとはいえない。なぜなら、たまたま観測した「現在のデータ」 $\{a_\alpha\}$  を説明しても、それ以外のデータに対して意味がある保証がないからである。

このことからモデルの「良さ」をその「予測能力」で測るのが自然である<sup>1),15)</sup>。「現在のデータ」 $\{a_\alpha\}$  と同じ確率分布に従う独立な「将来のデータ」を  $\{a_\alpha^*\}$  とするとき、現在のデータ  $\{a_\alpha\}$  から計算した最尤推定解  $\{\hat{a}_\alpha\} \in \hat{S}$  の将来のデータ  $\{a_\alpha^*\}$  に関する残差平方和を

$$S^*(\hat{S}) = \sum_{\alpha=1}^N (a_\alpha^* - \hat{a}_\alpha, V_0[a_\alpha]^{-1}(a_\alpha^* - \hat{a}_\alpha)) \quad (6)$$

とする。これは確率変数であるから期待値をとり、

$$I(S) = E^*[E[S^*(\hat{S})]] \quad (7)$$

と置く。これが小さいほど良いモデルであると見なせる。ただし  $E[\cdot]$ ,  $E^*[\cdot]$  はそれぞれ現在のデータ  $\{a_\alpha\}$  および将来のデータ  $\{a_\alpha^*\}$  に関する期待値である。  $I(S)$  をモデル  $S$  の期待残差と呼ぶ。この期待残差の不偏推定量が第 1 近似において次の形で与えられる。

$$AIC(S) = S(\hat{S}) + 2(pN + n')\epsilon^2 \quad (8)$$

これをモデル  $S$  の幾何学的情報量基準（幾何学的 AIC）と呼ぶ。

（導出）数学的な厳密さは文献 7), 11) に譲るが、式 (8) を得る道筋は直観的には次のとおりである。式 (3), (6) の右辺の形を象徴的にそれぞれ  $S_{a_\alpha, \hat{a}}$ ,  $S_{a_\alpha^*, \hat{a}}$  などと書く。  $\{a_\alpha\}$ ,  $\{a_\alpha^*\}$  は互いに独立であるから、

$$I(S) = E^*[E[S_{a_\alpha^*, \hat{a}}]] = E^*[S_{a_\alpha^*, \hat{a}}] + E[S_{\hat{a}, \hat{a}}] \quad (9)$$

と書ける。  $\{a_\alpha\}$ ,  $\{a_\alpha^*\}$  は同一分布を持つから、上式

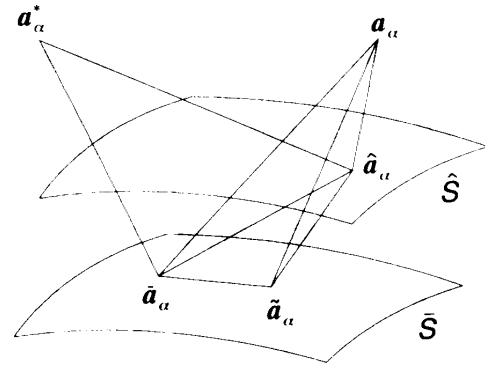


図3 期待残差の評価。  
Fig. 3 Evaluation of expected residual.

の右辺第 1 項は  $E[S_{a_\alpha, \hat{a}}]$  に等しい。データ  $a_\alpha$  を  $\hat{S}$  に最適に投影した点が  $\hat{a}_\alpha$  である。データ  $a_\alpha$  を  $\bar{S}$  に最適に投影した点を  $\tilde{a}_\alpha$  とすると (図 3),

$$E[S_{a_\alpha, \hat{a}}] = E[S_{a_\alpha, \hat{a}}] + E[S_{\hat{a}, \hat{a}}] + E[S_{\tilde{a}_\alpha, \hat{a}}] \quad (10)$$

と書ける。式 (9) の右辺第 2 項は次のように書ける。

$$E[S_{\tilde{a}_\alpha, \hat{a}}] = E[S_{a_\alpha, \hat{a}}] + E[S_{\hat{a}, \hat{a}}] \quad (11)$$

$S_{a_\alpha, \hat{a}} = S(\hat{S})$  であるから次式を得る。

$$I(S) = E[S(\hat{S})] + 2(E[S_{\tilde{a}_\alpha, \hat{a}}] + E[S_{\hat{a}, \hat{a}}]) \quad (12)$$

$S_{\tilde{a}_\alpha, \hat{a}}$  は真のモデル  $\bar{S}$  内での変動の残差平方和であり、その  $1/\epsilon^2$  倍は第 1 近似においてモデル  $\bar{S}$  の次元  $p$  に等しい自由度の  $\chi^2$  分布に従う確率変数の  $N$  個の独立な和となる。したがって  $S_{\tilde{a}_\alpha, \hat{a}}$  の期待値は  $\epsilon^2 Np$  である。一方、 $S_{\hat{a}, \hat{a}}$  は最尤推定によるモデルのあてはめの残差平方和であり、これは誤差が小さいときクラメル・ラオの下界と呼ばれる理論限界を達成する<sup>7),8)</sup>。この近似において  $S_{\tilde{a}_\alpha, \hat{a}}/\epsilon^2$  はモデルの自由度  $n'$  に等しい自由度の  $\chi^2$  分布に従うので  $S_{\tilde{a}_\alpha, \hat{a}}$  の期待値は  $\epsilon^2 n'$  である。以上より式 (8) を得る。 □

統計学で用いられる AIC はデータ数無限大の「漸近解析」によって「大数の法則」と「中心極限定理」から導かれるが<sup>1),15)</sup>、式 (8) は誤差が小さいときの「摂動解析」として導出される点が対照的である。統計学のことはばを用いると、真の位置  $\{\bar{a}_\alpha\}$  が「攪乱母数」にあたり、データの個数  $N$  だけ存在する。一方、多様体  $S$  を指定する  $n'$  個のパラメータが「構造母数」にあたるが、これはデータの確率密度関数に陽には含まれていない。含まれているのは攪乱母数  $\{\bar{a}_\alpha\}$  のみであり、構造母数は間接的に攪乱母数  $\{\bar{a}_\alpha\}$  の相互の関係を記述している。画像理解やロボティクスで見られる推定問題のほとんどすべてがこの形である<sup>8)</sup>。式 (8) から次のことが分かる。

- データ数  $N$  が大きいとき、モデル  $S$  の自由度

$n'$  は幾何学的 AIC にほとんど影響を及ぼさない。

- あてはめる多様体の次元  $p$  が本質的な役割を果たす。
- データ数  $N$  が幾何学的 AIC の第 2 項に現れている。

### 5. 残差平方和の解析的表現

モデル  $S$  が次の形の拘束条件で定義されるとする。

$$F^{(k)}(\mathbf{a}, \mathbf{u}) = 0, \quad k = 1, \dots, L \quad (13)$$

これら  $L$  個の方程式は  $\mathbf{a} \in \mathcal{A}$  の方程式として代数的に独立でなくてもよい。独立なもの個数  $r$  をこの拘束条件のランクと呼ぶ<sup>8)</sup>。ただし  $L$  個の方程式の各々がデータ空間  $\mathcal{A}$  で横断的に交わる余次元 1 の多様体を定義することが必要である<sup>8),10),11)</sup>。このとき式 (13) はデータ空間  $\mathcal{A}$  に余次元  $r$  の多様体  $S$  を定義する。各データ  $\mathbf{a}_\alpha$  を式 (3) で定義するマハラノビス距離の二乗和  $S(\{\hat{\mathbf{a}}_\alpha\}; S)$  が最小になるように  $S$  上に射影した点  $\hat{\mathbf{a}}_\alpha$  は第 1 近似において次のように表される<sup>10)</sup>。

$$\hat{\mathbf{a}}_\alpha = \mathbf{a}_\alpha - V_0[\mathbf{a}_\alpha] \sum_{k,l=1}^L W^{(kl)} F_\alpha^{(k)} \nabla_{\mathbf{a}} F_\alpha^{(l)} \quad (14)$$

$$(W_\alpha^{(kl)}) = ((\nabla_{\mathbf{a}} F_\alpha^{(k)}, V_0[\mathbf{a}_\alpha] \nabla_{\mathbf{a}} F_\alpha^{(l)}))_r^- \quad (15)$$

$(\cdot)_r^-$  はランクを  $r$  に拘束した (すなわちスペクトル分解し、固有値を大きいほうから  $r$  個だけ残して残りを 0 とおいて計算した) 一般逆行列を表す<sup>11)</sup>。式 (15) は  $(\nabla_{\mathbf{a}} F_\alpha^{(k)}, V_0[\mathbf{a}_\alpha] \nabla_{\mathbf{a}} F_\alpha^{(l)})$  を  $(kl)$  要素とする行列のランクを  $r$  に拘束した一般逆行列の  $(kl)$  要素が  $W_\alpha^{(kl)}$  に等しいことを表す略記である。 $F_\alpha^{(k)}$ ,  $\nabla_{\mathbf{a}} F_\alpha^{(k)}$  はそれぞれ  $F^{(k)}(\mathbf{a}_\alpha, \mathbf{u})$ ,  $\nabla_{\mathbf{a}} F^{(k)}(\mathbf{a}_\alpha, \mathbf{u})$  の略記であり、 $\nabla_{\mathbf{a}}(\cdot)$  は  $\partial(\cdot)/\partial a_1, \dots, \partial(\cdot)/\partial a_n$  を要素とする列ベクトルである。幾何学的には式 (14) はデータ  $\mathbf{a}_\alpha$  がデータ空間  $\mathcal{A}$  でそれを中心とする楕円体  $(\mathbf{a} - \mathbf{a}_\alpha, V_0[\mathbf{a}_\alpha]^{-1}(\mathbf{a} - \mathbf{a}_\alpha)) = \text{const.}$  と多様体  $S$  との接点  $\hat{\mathbf{a}}_\alpha$  へ射影されることを示す (図 4)。

式 (14) を式 (3) に代入すると第 1 近似において次のようになる<sup>8)</sup>。

$$S(S) = \sum_{\alpha=1}^N \sum_{k,l=1}^L W_\alpha^{(kl)} F_\alpha^{(k)} F_\alpha^{(l)} \quad (16)$$

これは  $\mathbf{u}$  の関数であり、これを最小にする値  $\hat{\mathbf{u}} \in \mathcal{U}$  の定義する多様体  $\hat{S}$  が  $S$  の最尤推定解である。それを式 (8) に代入して幾何学的 AIC の具体的な表現が得られる。

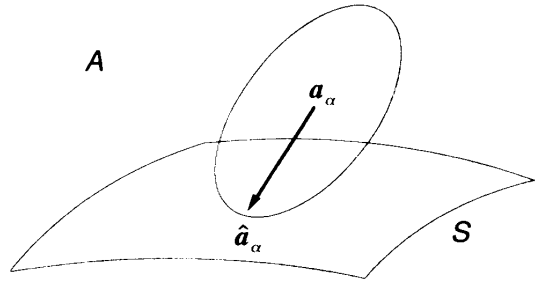


図 4 データ  $\mathbf{a}_\alpha$  は楕円体  $(\mathbf{a} - \mathbf{a}_\alpha, V_0[\mathbf{a}_\alpha]^{-1}(\mathbf{a} - \mathbf{a}_\alpha)) = \text{const.}$  と多様体  $S$  との接点  $\hat{\mathbf{a}}_\alpha$  へ射影される。

Fig. 4 Datum  $\mathbf{a}_\alpha$  is projected onto the tangent point  $\hat{\mathbf{a}}_\alpha$  of ellipsoid  $(\mathbf{a} - \mathbf{a}_\alpha, V_0[\mathbf{a}_\alpha]^{-1}(\mathbf{a} - \mathbf{a}_\alpha)) = \text{constant}$  to manifold  $S$ .

### 6. 包含関係にある幾何学的モデルの選択基準

式 (8) によって幾何学的 AIC を計算するにはノイズレベル  $\epsilon$  が既知でなければならないが、先に指摘したようにこれは現実的ではない。しかし、「個々のモデルの評価」ではなく、包含関係にあるモデルの選択を考えればこれを避けることができる。 $S_1$ ,  $S_2$  をそれぞれ次元, 余次元, 自由度が  $p_1, r_1, n'_1$  および  $p_2, r_2, n'_2$  のモデルとする。そして  $S_2$  はモデル  $S_1$  にさらに制約を加えて得られた強いモデルであるとし、これを次のように表す。

$$S_2 \succ S_1 \quad (17)$$

モデル  $S_1$  および  $S_2$  のもとでの残差平方和をそれぞれ  $S(\hat{S}_1)$ ,  $S(\hat{S}_2)$  とすると、データの如何にかかわらず  $S(\hat{S}_2) \geq S(\hat{S}_1)$  である。

モデル  $S_1$  のもとでは二乗ノイズレベル  $\epsilon^2$  が式 (5) によって推定できる。それを代入するとモデル  $S_1, S_2$  の幾何学的 AIC が次のように評価できる。

$$\begin{aligned} AIC(S_1) &= S(\hat{S}_1) + \frac{2(p_1 N + n'_1)}{r_1 N - n'_1} S(\hat{S}_1) \\ AIC(S_2) &= S(\hat{S}_2) + \frac{2(p_2 N + n'_2)}{r_1 N - n'_1} S(\hat{S}_1) \end{aligned} \quad (18)$$

式 (6), (7) の期待残差の定義より、次の量は 2 つのモデルのあてはまりの良さの比を表す。

$$K = \sqrt{\frac{AIC(S_2)}{AIC(S_1)}} \quad (19)$$

$K < 1$  であればモデル  $S_1$  を強いモデル  $S_2$  に置き換えることによって予測能力が高まると期待される。

### 7. モデル選択と仮説検定

幾何学的 AIC によるモデル選択は統計学の仮説検定の否定的な意味とは対照的である。仮説検定論によれば、 $S(\hat{S}_2)/\epsilon^2$  が第 1 近似において自由度  $r_2 N - n'_2$

の  $\chi^2$  分布に従うから、「モデル  $S_2$  が正しい」という仮説は

$$\frac{S(\hat{S}_2)}{\epsilon^2} > \chi_{r_2 N - n'_2, a}^2 \quad (20)$$

のとき有意水準  $a\%$  で棄却される。 $\chi_{q, a}^2$  は自由度  $q$  の  $\chi^2$  分布の上側  $a\%$  点である。この検定にはノイズレベル  $\epsilon$  が既知でなければならないが、式 (5) で  $\epsilon^2$  を推定すれば式 (20) は次のように書ける。

$$\frac{S(\hat{S}_2)}{S(\hat{S}_1)} > \frac{\chi_{r_2 N - n'_2, a}^2}{r_1 N - n'_1} \quad (21)$$

左辺は対数尤度比に等しいから、この形は尤度比検定と呼ばれるものである。この検定は、式 (21) が成立すれば  $(100 - a)\%$  の信頼水準で「モデルが  $S_2$  であることに確信が持てない」、したがって「強い仮説（帰無仮説）を採用する積極的な理由がない」という意味である。しかし誤差がある限り信頼水準を上げれば（有意水準を下げれば）どんな仮説も必ず棄却されるものであり、「どちらのモデルがより妥当であるか」という問に答えるものではない。一方、幾何学的 AIC によるモデル選択は肯定的意味を持ち、「弱いモデルが真だとすると、それから予測できることは強いモデルからはさらに良く予測できる」、したがって「強いモデルを採用する積極的な理由がある」という意味である。このとき恣意的な有意水準を何ら必要としない。

## 8. 動画像の 3 次元解析

### 8.1 特異性の判定基準

2 画像の対応点からカメラの 3 次元運動と物体の 3 次元形状とを計算する運動解析アルゴリズムは多くの研究者によって研究された<sup>5), 17)</sup>。しかしそのアルゴリズムは物体が平面のとき破綻するので、別のアルゴリズムを用いなければならない。これは平面が 3 次元解析に特異性をもたらす臨界面であるためである<sup>3), 14)</sup>。平面の解析アルゴリズムもまた多くの研究者によって研究されている<sup>12), 13), 16)</sup>。しかし、平面の場合も非平面の場合もカメラの並進は 0 でないと仮定している。並進が 0 のとき、すなわちカメラの運動が純粋回転であるときは何らの 3 次元情報も得られない<sup>3)</sup>。したがって、2 画像から 3 次元解析をするときは

- (1) まず並進が 0 かどうか判定し、カメラの運動が純粋回転であればメッセージを出して終了し、
- (2) 次に物体が平面かどうか判定し、平面であれば平面解析アルゴリズムを適用し、
- (3) そうでなければ一般の 3 次元解析アルゴリズムを適用する、

という手順を踏まなければならない。しかし画像には誤差があるので、並進が 0 でも物体が平面でも一般の 3 次元解析アルゴリズムを適用すると何らかの（信頼性のない）解が得られる。これに対して従来は適当な判定基準と恣意的なしきい値を用いて判定を行うことが多かった。たとえば、物体が平面のとき画像に誤差がなければ最小固有値 0 が重根になることに着目して、最小固有値と 2 番目の固有値の差がある定数以内なら平面と判定する、などである。しかし、そのしきい値をどう定めたらよいのであろうか。

- しきい値の設定には特徴点検出の精度に関する知識が不可欠である。なぜなら、精度が低くければしきい値を大きく、精度が高いと小さくとらなければならないからである。しかし、精度はその画像の撮影条件やその物体の形状や位置や向きに依存し、画像ごとに異なるので、あらかじめ予測することは不可能に近い。
- 仮に精度が既知でも、それから分かるのは誤差の「確率分布」である。仮説検定論にしたがって、たとえば平面物体が平面でないと判定される確率（有意水準）が  $a\%$  以下になるように決めるとしても、その有意水準は何 % にすべきなのであろうか。その定め方次第で判定結果は変化する。

従来はこの問題はほとんど考慮されておらず、中には実験例に応じてしきい値を調節したのではないかと疑われるものもある。幾何学的 AIC を用いると、この問題は次のように解決する。

### 8.2 3 次元解析のモデル

カメラのレンズの中心を原点  $O$  とし、その光軸を  $Z$  軸とする  $XYZ$  座標系をとる。原点  $O$  と撮像面との距離を長さの単位にとれば、撮像は画像面  $Z = 1$  への透視変換と見なせる。画像面上で光軸の位置を画像原点とし、 $X$  軸、 $Y$  軸に平行な  $xy$  画像座標軸をとると、画像座標  $(x, y)$  の点は  $XYZ$  座標系に関してベクトル  $\mathbf{x} = (x, y, 1)^T$  で表される<sup>6)</sup> ( $T$  は転置を表す)。このカメラを  $XYZ$  座標系に関して原点  $O$  の周りに回転行列  $\mathbf{R}$  だけ回転し、ベクトル  $\mathbf{h}$  だけ平行移動した位置に移動する運動を考える (図 5)。 $\{\mathbf{h}, \mathbf{R}\}$  を運動パラメータと呼ぶ<sup>6)</sup>。運動後のカメラに関する  $X'Y'Z'$  座標をとり、同様に  $x'y'$  画像座標系を定義すると、画像座標  $(x', y')$  の点は  $X'Y'Z'$  座標系に関してベクトル  $\mathbf{x}' = (x', y', 1)^T$  で表される。運動前後のカメラ座標系に関してそれぞれベクトル  $\mathbf{x}, \mathbf{x}'$  で表される点が空間中の同一の特徴点の投影像である必要十分条件は次のエピ極線方程式が成り立つことである<sup>3), 6)</sup>。

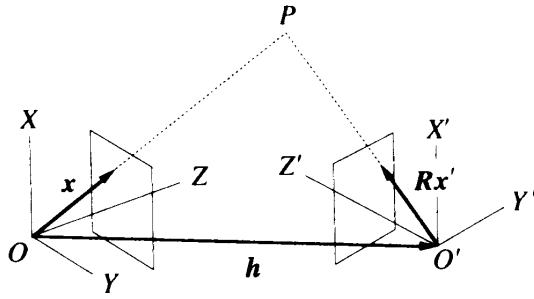


図5 カメラの運動の幾何学的関係。  
Fig. 5 Geometry of camera motion.

$$(\mathbf{x}, \mathbf{G}\mathbf{x}') = 0 \quad (22)$$

行列  $\mathbf{G}$  は次のように定義され、基本行列と呼ばれる<sup>3),6)</sup>。

$$\mathbf{G} = \mathbf{h} \times \mathbf{R} \quad (23)$$

ただし右辺はベクトル  $\mathbf{h}$  と行列  $\mathbf{R}$  の各列とのベクトル積を列とする行列を表す。式 (22) は並進  $\mathbf{h}$  を任意に定数倍しても成立する。これは遠くの大きい物体に対するカメラの大きい運動と近くの小さい物体に対するカメラの小さい運動とが画像のみからは区別できないという事実に対応している。そこで並進を  $\|\mathbf{h}\| = 1$  と正規化する ( $\|\cdot\|$  はベクトルのノルムを表す)。エッジ検出方程式 (22) は次元 3, 余次元 1, 自由度 5 のモデル  $S$  を定義する (並進  $\mathbf{h}$  は単位ベクトルとしているから自由度 2 であり, 回転  $\mathbf{R}$  は 3 自由度を持つ)。

$N$  組の対応点  $\mathbf{x}_\alpha, \mathbf{x}'_\alpha, \alpha = 1, \dots, N$  は独立でそれぞれ期待値  $\bar{\mathbf{x}}_\alpha, \bar{\mathbf{x}}'_\alpha$ , 共分散行列  $V[\mathbf{x}_\alpha], V[\mathbf{x}'_\alpha]$  の正規分布に従うとする。ベクトル  $\mathbf{x}_\alpha, \mathbf{x}'_\alpha$  の第 3 成分はともに定数 1 であるから,  $V[\mathbf{x}_\alpha], V[\mathbf{x}'_\alpha]$  はそれぞれ  $XY$  面,  $X'Y'$  面を値域とするランク 2 の特異行列である。これを次のようにノイズレベルと正規化共分散行列に分解する。

$$V[\mathbf{x}_\alpha] = \epsilon^2 V_0[\mathbf{x}_\alpha], \quad V[\mathbf{x}'_\alpha] = \epsilon^2 V_0[\mathbf{x}'_\alpha] \quad (24)$$

このとき運動パラメータ  $\{\mathbf{h}, \mathbf{R}\}$  の最尤推定量は「くりこみ法」<sup>4),6)</sup> によって精度良く効率的に計算できる<sup>5)</sup>。対応する最尤推定解を  $\hat{S}$  とすると, 残差平方和  $S(\hat{S})$  は次のように表せる。

$$S(\hat{S}) = \sum_{\alpha=1}^N \frac{(\mathbf{x}_\alpha, \hat{\mathbf{G}}\mathbf{x}'_\alpha)^2}{(\mathbf{x}'_\alpha, \hat{\mathbf{G}}^\top V_0[\mathbf{x}_\alpha] \hat{\mathbf{G}}\mathbf{x}_\alpha) + (\mathbf{x}_\alpha, \hat{\mathbf{G}} V_0[\mathbf{x}'_\alpha] \hat{\mathbf{G}}^\top \mathbf{x}'_\alpha)} \quad (25)$$

ただし  $\hat{\mathbf{G}}$  は式 (23) の基本行列において  $\{\mathbf{h}, \mathbf{R}\}$  を  $\{\hat{\mathbf{h}}, \hat{\mathbf{R}}\}$  で置き換えたものである。

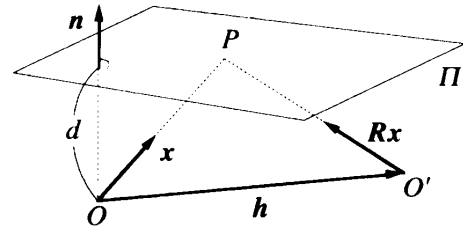


図6 平面物体の観測。  
Fig. 6 Observation of a planar surface.

### 8.3 平面性の判定

物体面が平面  $\Pi$  である場合を考える (図 6)。平面  $\Pi$  の単位法線ベクトルを運動前のカメラに関する  $XYZ$  座標系に関して  $\mathbf{n}$  とし, 原点  $O$  から平面  $\Pi$  までの距離 (ベクトル  $\mathbf{n}$  の方向に正) を  $d$  とする。位置ベクトルを  $\mathbf{r}$  とすると, 平面  $\Pi$  の方程式は  $(\mathbf{n}, \mathbf{r}) = d$  と書ける。 $\{\mathbf{n}, d\}$  を平面パラメータと呼ぶ。運動前後のカメラ座標系についてそれぞれ  $\mathbf{x}, \mathbf{x}'$  で表される点が平面  $\Pi$  上にある必要十分条件は次式が成り立つことである<sup>3),6),8),12),13),16)</sup>。

$$\mathbf{x}' \times \mathbf{A}\mathbf{x} = 0 \quad (26)$$

ただし  $\mathbf{A}$  は次のように定義した 3 次元行列である ( $\mathbf{I}$  は単位行列を表す)。

$$\mathbf{A} = \mathbf{R}^\top (\mathbf{h}\mathbf{n}^\top - d\mathbf{I}) \quad (27)$$

式 (26) は次元 2, 余次元 2, 自由度 8 のモデル  $S_\Pi$  を定義する。平面パラメータ  $\{\mathbf{n}, d\}$  と運動パラメータ  $\{\mathbf{h}, \mathbf{R}\}$  の最尤推定量を  $\{\hat{\mathbf{n}}, \hat{d}\}, \{\hat{\mathbf{h}}, \hat{\mathbf{R}}\}$  とする。これらは「くりこみ法」<sup>4),6)</sup> によって精度良く効率的に計算できる<sup>12)</sup>。対応する最尤推定解を  $\hat{S}_\Pi$  とすると残差平方和は次のように表せる。

$$S(\hat{S}_\Pi) = \sum_{\alpha=1}^N (\mathbf{x}'_\alpha \times \hat{\mathbf{A}}\mathbf{x}_\alpha, \hat{\mathbf{W}}_\alpha (\mathbf{x}'_\alpha \times \hat{\mathbf{A}}\mathbf{x}_\alpha)) \quad (28)$$

$$\hat{\mathbf{W}}_\alpha = \left( \mathbf{x}'_\alpha \times \hat{\mathbf{A}}V_0[\mathbf{x}_\alpha] \hat{\mathbf{A}}^\top \times \mathbf{x}'_\alpha + (\hat{\mathbf{A}}\mathbf{x}_\alpha \times V_0[\mathbf{x}'_\alpha] \times \hat{\mathbf{A}}\mathbf{x}_\alpha) \right)_2^{-1} \quad (29)$$

ただし  $\hat{\mathbf{A}}$  は式 (27) で  $\{\mathbf{n}, d\}, \{\mathbf{h}, \mathbf{R}\}$  をそれぞれ  $\{\hat{\mathbf{n}}, \hat{d}\}, \{\hat{\mathbf{h}}, \hat{\mathbf{R}}\}$  に置き換えたものであり, ベクトル  $\mathbf{a}$  と行列  $\mathbf{U}$  に対して  $\mathbf{a} \times \mathbf{U} \times \mathbf{a}$  は  $(\mathbf{a} \times \mathbf{U})(\mathbf{a} \times \mathbf{I})^\top$  の略記である。 $S_\Pi \succ S$  であるから,  $K_\Pi$  を次のように定義すると,  $K_\Pi > 1$  であれば物体形状は平面であると判定される。

$$K_\Pi = \sqrt{\frac{N-5}{7N+5} \left( \frac{S(\hat{S}_\Pi)}{S(\hat{S})} + \frac{4N+16}{N-5} \right)} \quad (30)$$

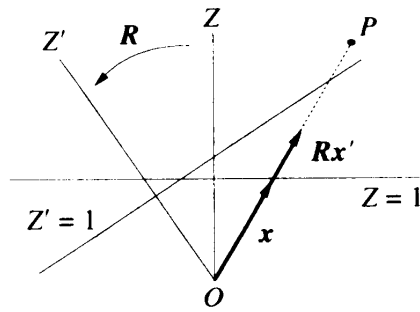


図7 カメラの純粋回転.  
Fig. 7 Pure camera rotation.

8.4 回転性の判定

カメラの運動が純粋回転である必要十分条件は

$$\mathbf{x} \times \mathbf{R}\mathbf{x}' = \mathbf{0} \quad (31)$$

となる<sup>3)</sup> (図7)。これは次元2, 余次元2, 自由度3のモデル \$S\_R\$ を定義する。回転 \$R\$ の最尤推定量 \$\hat{R}\$ の近似値は特異値分解によって容易に計算できる<sup>3)</sup>。\$\hat{R}\$ に対応する最尤推定解を \$\hat{S}\_R\$ とすると、残差平方和 \$S(\hat{S}\_R)\$ は次のように表せる。

$$S(\hat{S}_R) = \sum_{\alpha=1}^N (\mathbf{x}_\alpha \times \hat{R}\mathbf{x}'_\alpha, \hat{W}_\alpha (\mathbf{x}_\alpha \times \hat{R}\mathbf{x}'_\alpha)) \quad (32)$$

$$\hat{W}_\alpha = (\mathbf{x}_\alpha \times \hat{R}V_0[\mathbf{x}'_\alpha] \hat{R}^\top \times \mathbf{x}_\alpha + (\hat{R}\mathbf{x}'_\alpha) \times V_0[\mathbf{x}_\alpha] \times (\hat{R}\mathbf{x}'_\alpha))^{-2} \quad (33)$$

\$S\_R \succ S\$ であるから、\$K\_R\$ を次のように定義すると、\$K\_R > 1\$ であればカメラの運動は純粋回転であると判定される。

$$K_R = \sqrt{\frac{N-5}{7N+5} \left( \frac{S(\hat{S}_R)}{S(\hat{S})} + \frac{4N+6}{N-5} \right)} \quad (34)$$

8.5 実験例

空間中に角度 \$\pi - \theta\$ をなす2枚の平面格子を異なる視点から撮影したシミュレーション画像を生成する。画像は \$512 \times 512\$ 画素、焦点距離は \$f = 600\$ (画素) と想定する。画像上の各格子点の \$x, y\$ 座標に独立に期待値0, 標準偏差 \$\sigma\$ (画素) の正規乱数を誤差として加えたものを特徴点とし、3次元復元を行う。図8は各 \$\theta\$ に対して毎回異なる誤差を用いて100回計算し、平面と判定される割合をいくつかの \$\sigma\$ の値に対して示したものである。このように判定は誤差に自動的に適応し、\$\sigma = 1\$ のときは \$\theta = 22^\circ\$ のとき平面と判定される割合がほぼ50%となる。図9はそのとき平面と判定される一例であり、図10はそのとき平面と判定されない一例である。図11はそれぞれから復元し

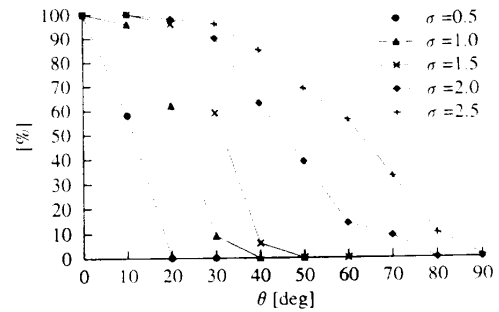


図8 平面と判定される割合.  
Fig. 8 Percentage of the instances judged to be planar.

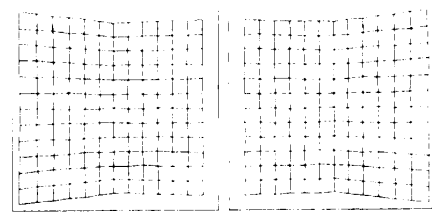


図9 平面と判定される例 (\$\sigma = 1, \theta = 22^\circ, K\_R = 1.43, K\_\Pi = 0.87\$).  
Fig. 9 An instance for which the object is judged to be planar (\$\sigma = 1, \theta = 22^\circ, K\_R = 1.43, K\_\Pi = 0.87\$).

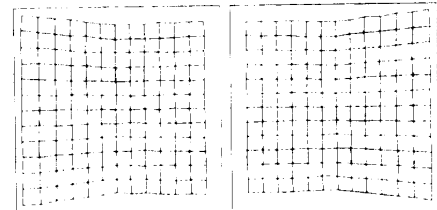


図10 平面と判定されない例 (\$\sigma = 1, \theta = 22^\circ, K\_R = 2.34, K\_\Pi = 1.13\$).  
Fig. 10 An instance for which the object is judged to be non-planar (\$\sigma = 1, \theta = 22^\circ, K\_R = 2.34, K\_\Pi = 1.13\$).

た3次元形状をある位置から見たものである。一般モデルによる復元と平面モデルによる復元とを重ねて図示し、真の形状を点線で示している。このように画像としてはほんのわずかの違いでも、平面と判定されたものは非平面としての復元はほとんど意味がなく、逆に非平面と判定されたものはほぼ真の形状に近い形状が復元されている。

9. まとめ

本論文ではロボットが画像やセンサデータから3次元環境のモデルを構築するための幾何学的推定問題を一般的に記述し、予測能力の高いモデルを恣意的な有意水準を用いないで選択する基準として「幾何学的AIC」を導出した。そして、包含関係にあるモデルに

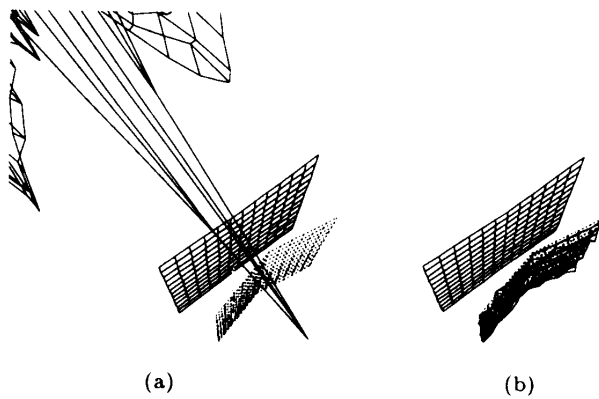


図 11 3次元形状 (点線は真の形状). (a) 図9からの復元. (b) 図10からの復元.

Fig. 11 Reconstructed 3-D shapes (the true shape is drawn in dashed lines). (a) Reconstruction from Fig. 9; (b) Reconstruction from Fig. 10.

対しては、その良さがデータの誤差の大きさに関する知識なしに評価できることを示し、動画画像解析において対象が平面であるか、あるいはカメラの運動が純粹回転であるかを自動的に判定する方式を示した。本論文はコンピュータビジョンやロボティクスの広い範囲の問題に応用が可能であり、機械による認識・理解の新しい方法論を提起するものである。

謝辞 本研究の発端となった討論をしていただいた理化学研究所の甘利俊一氏、および実験に参加していただいた日立製作所の塩沢仁氏に感謝します。本研究の一部は文部省科学研究費一般研究B (No.07458067) および大川情報通信基金からの助成による。

### 参考文献

- 1) Akaike, H.: A New Look at the Statistical Model Identification, *IEEE Trans. Automation Control*, Vol.19, pp.176-723 (1974).
- 2) Faugeras, O.: *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA (1993).
- 3) Kanatani, K.: *Geometric Computation for Machine Vision*, Oxford University Press, Oxford (1993).
- 4) 金谷健一: コンピュータビジョンのためのくりこみ法, *情報処理学会論文誌*, Vol.35, No.2, pp.201-209 (1994).
- 5) Kanatani, K.: Renormalization for Motion Analysis: Statistically Optimal Algorithm, *IEICE Trans. Inf. & Syst.*, Vol.E77-D, No.11, pp.1233-1239 (1994).
- 6) 金谷健一: 空間データの数理—3次元コンピュータリングに向けて—, 朝倉書店 (1995).

- 7) 金谷健一: 幾何学的モデル選択の情報量基準, *電子情報通信学会技術研究報告*, IT95-16, pp.19-24 (1995).
- 8) 金谷健一: 当てはめ問題の最適推定と精度の理論限界, *情報処理学会論文誌*, Vol.36, No.8, pp.1865-1873 (1995).
- 9) 金谷健一: 画像理解のための統計学: 画像の幾何学的解釈の信頼性評価, *情報処理*, Vol.37, No.1, pp.52-60 (1996).
- 10) 金谷健一: 幾何学的補正問題の精度の理論限界と最適計算法, *情報処理学会論文誌*, Vol.37, No.3, pp.363-370 (1996).
- 11) Kanatani, K.: *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam (1996).
- 12) Kanatani, K. and Takeda, S.: 3-D Motion Analysis of a Planar Surface by Renormalization, *IEICE Trans. Inf. & Syst.*, Vol.E78-D, No.8, pp.1074-1079 (1995).
- 13) Longuet-Higgins, H.C.: The Reconstruction of a Plane Surface from Two Perspective Projections, *Proc. Roy. Soc. Lond.*, Ser.B, Vol.227, pp.399-410 (1986).
- 14) Maybank, S.: *Theory of Reconstruction from Image Motion*, Springer, Berlin (1993).
- 15) 坂本慶行, 石黒真木夫, 北川源四郎: 情報量統計学, 共立出版 (1983).
- 16) Weng, J., Ahuja, N. and Huang, T.S.: Motion and Structure from Point Correspondences with Error Estimation: Planar Surfaces, *IEEE Trans. Sig. Proc.*, Vol.39, No.12, pp.2691-2717 (1991).
- 17) Weng, J., Ahuja, N. and Huang, T.S.: Optimal Motion and Structure Estimation, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol.15, No.9, pp.864-884 (1993).
- 18) Weng, J., Huang, T.S. and Ahuja, N.: *Motion and Structure from Image Sequences*, Springer, Berlin (1993).

(平成7年9月18日受付)

(平成8年3月12日採録)

### 金谷 健一 (正会員)



1947年岡山県生。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。現在群馬大学工学部情報工学科教授。米国Maryland大学, デンマークCopenhagen大学, 英国Oxford大学客員研究員。