

情報量基準による動画像解析の特異性の自動判定

金谷 健一[†] 塩沢 仁[†]

群馬大学工学部情報工学科

2 画像からカメラの 3 次元運動と物体の 3 次元形状を計算する動画像解析アルゴリズムは、物体が平面であったりカメラの運動が純粋回転であると破綻する。本論文では物体が平面とみなせるか、あるいはカメラの運動が純粋回転とみなせるかを画像の誤差に関する知識を用いず、さらに何らの恣意的なしきい値をも用いずに自動的に判定する新しい方法を示し、シミュレーション画像および実画像を用いた実験結果を示す。本論文で用いる基本原理は幾何学的情報量基準 (AIC) で評価した「予測能力の高い」モデルを採用することであり、ロボティクスやコンピュータビジョンの広い範囲の問題における幾何学的モデルの選択基準として新しい方法論を提起するものがある。

キーワード: 情報量基準, 仮説検定, モデル選択, 動画像解析, 平面性, 最尤推定

Automatic Singularity Test for Motion Analysis by an Information Criterion

Kenichi Kanatani[†] and Hitoshi Shiozawa[†]

Department of Computer Science
Gunma University, Kiryu, Gunma 376, Japan

The structure-from-motion algorithm for computing the 3-D camera motion and the 3-D object shape from two views fails if the object is a planar surface or the camera motion is a pure rotation. This paper presents a new scheme for automatically judging whether or not the object can be regarded as a planar surface and whether or not the camera motion can be regarded as a pure rotation without using any knowledge about the noise in the images. This judgment does not involve any arbitrarily set thresholds, either. The basic principle of our scheme is to choose a model that has “higher predicting capability” measured by the *geometric information criterion (AIC)*. The performance of our scheme is tested by real and synthetic image experiments. Our approach presents a new paradigm for geometric model selection in a wide range of problems of robotics and computer vision.

Key words: AIC, testing of hypotheses, model selection, motion analysis, planarity, maximum likelihood estimation

謝辞: 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った。

[†] 376 桐生市天神町 1-5-1 群馬大学工学部情報工学科, Tel: (0277)30-1841, Fax: (0277)30-1801,
E-mail: kanatani@cs.gunma-u.ac.jp, shiozawa@ail.cs.gunma-u.ac.jp

1. はじめに

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、ほぼ完成のレベルに洗練されている [4, 13, 17, 19]. 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と呼ばれる曲面の一種であるためである [2, 15]. 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 [11, 14, 18]. 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カメラに入射する光線に変化がなく、何らの3次元情報も得られない。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かどうかテストし(これを回転テストと呼ぶ), 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テストし(これを平面テストと呼ぶ), 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なうことが多かった。例えば、物体が平面のとき画像に誤差がなければ最小固有値0が重根になることを手掛かりにして、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?

1. しきい値の設定には画像から得られる特徴点の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、これをあらかじめ予測することは不可能に近い。
2. 仮に精度が既知であるとしても、それからわかるのは誤差の「確率分布」である。統計的検定論にしたがって、例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 $\alpha\%$ 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、このような問題はほとんど考慮されておらず、実験した例題に応じてしきい値を適当に調節する

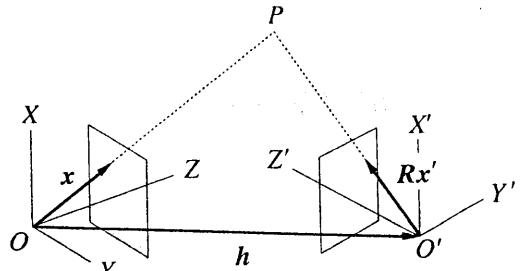


図1: カメラの運動の幾何学的関係。

ことさえ見られた。本論文では画像の精度に関する知識を用いず、さらに何らの恣意的なしきい値も用いない平面テストおよび回転テストを導出し、シミュレーション画像および実画像を用いた実験結果を示す。本論文で用いる基本原理は幾何学的情報量基準(AIC)で評価した「予測能力の高い」モデルを採用することである [6, 7, 8, 16]. これは動画像解析のみならず、ロボティクスやコンピュータビジョンの広い範囲の問題における幾何学的モデルの選択基準として新しい方法論を提起するものでもある [12].

2. 動画像からの3次元復元の原理

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、撮像面は画像面 $Z = 1$ への透視変換とみなせる。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 画像座標系をとると、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 (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1)。 $\{\mathbf{h}, \mathbf{R}\}$ を運動パラメータと呼ぶ。運動後のカメラに関する $X'Y'Z'$ 座標系をとり、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。空間中のある特徴点が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}' で表される位置に観測されるとすると、画像に誤差がなければよく知られた次のエッジ極線方程式が成立する [1, 2, 5].

$$(\mathbf{x}, \mathbf{G}\mathbf{x}') = 0 \quad (1)$$

行列 \mathbf{G} は次のように定義され、基本行列と呼ばれる。

$$\mathbf{G} = \mathbf{h} \times \mathbf{R} \quad (2)$$

ただし、ベクトル \mathbf{a}, \mathbf{b} の内積を (\mathbf{a}, \mathbf{b}) で表し、ベクトル $\mathbf{a} = (a_i)$ と行列 $\mathbf{U} = (U_{ij})$ の外積 $\mathbf{a} \times \mathbf{U}$ はその (ij) 要素が $\sum_{k,l=1}^3 \epsilon_{ikl} a_k U_{lj}$ の行列と定義する¹。エ

¹ ϵ_{ijk} はエディンガトンのイプシロンであり、 (iji) が (123) の偶順列なら1、奇順列なら-1、それ以外は0をとる。

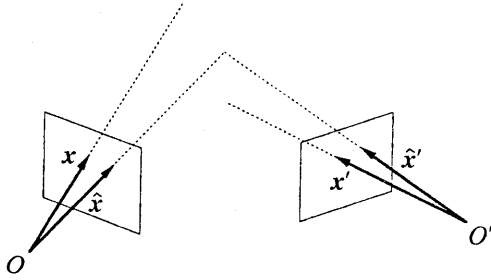


図 2: 視線が交わるように補正する.

ビ極線方程式 (1) はベクトル x, x' の定める視線が空間中で交わる必要十分条件である。しかし画像に誤差があれば、例えば正しい運動パラメータ $\{h, R\}$ に対しても x, x' が式 (1) のを満たすとは限らない。すなわち、ベクトル x, x' の定める視線が空間中で交わるとは限らない (図 2)。したがって、視線が交わるようにベクトル x, x' を補正しなければならない。

3. 一般モデルの残差平方和

N 組の対応点 $(x_\alpha, y_\alpha), (x'_\alpha, y'_\alpha), \alpha = 1, \dots, N$ が与えられたとし、これをベクトル x_α, x'_α で表して

$$x_\alpha = \bar{x}_\alpha + \Delta x_\alpha, \quad x'_\alpha = \bar{x}'_\alpha + \Delta x'_\alpha \quad (3)$$

と置く。 $\bar{x}_\alpha, \bar{x}'_\alpha$ は画像に誤差がないときに観測されるはずの真の位置である。誤差 $\Delta x_\alpha, \Delta x'_\alpha$ はそれぞれ独立で期待値 0 , 共分散行列 $V[x_\alpha], V[x'_\alpha]$ の正規分布に従う確率変数であるとする²。画像処理によって得られる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が、誤差の傾向 (一様であるとか等方であるとか、ある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。そこで共分散行列 $V[x_\alpha], V[x'_\alpha]$ が定数倍を除いて既知であると仮定する。すなわち、ある未知の定数 ϵ があって

$$V[x_\alpha] = \epsilon^2 V_0[x_\alpha], \quad V[x'_\alpha] = \epsilon^2 V_0[x'_\alpha] \quad (4)$$

と書け、 $V_0[x_\alpha], V_0[x'_\alpha]$ は既知であるとする。 ϵ をノイズレベルと呼び、 $V_0[x_\alpha], V_0[x'_\alpha]$ を正規化共分散行列と呼ぶ [5]。

画像に誤差があると各対応点 x_α, x'_α は誤差を含んでいるので、それらがエビ極線方程式 (1) を満たすように補正しなければならない。これには無数の解が存在するが、その中で確率的に最も確からしいものを選ぶ。それには正規化共分散行列 $V_0[x_\alpha], V_0[x'_\alpha]$ に関するマハラノビス距離の二乗和

$$J = \sum_{\alpha=1}^N (x_\alpha - \hat{x}_\alpha, V_0[x_\alpha]^{-1} (x_\alpha - \hat{x}_\alpha))$$

²ベクトル x_α, x'_α の第 3 成分は共に定数 1 であるから、共分散行列 $V[x_\alpha], V[x'_\alpha]$ 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。

$$+ \sum_{\alpha=1}^N (x'_\alpha - \hat{x}'_\alpha, V_0[x'_\alpha]^{-1} (x'_\alpha - \hat{x}'_\alpha)) \quad (5)$$

を最小にする $\hat{x}_\alpha, \hat{x}'_\alpha$ を選ばばよい [5]。ここに $(\cdot)^{-1}$ は (Moore-Penrose) 一般逆行列³である [9]。上式の最小値は運動パラメータ $\{h, R\}$ の関数であるから、これを $J[h, R]$ と書く。これは次のようになる [9]。

$$J[h, R] = \sum_{\alpha=1}^N \frac{(x_\alpha, Gx'_\alpha)^2}{(x'_\alpha, G^T V_0[x_\alpha] G x'_\alpha) + (x_\alpha, G V_0[x'_\alpha] G^T x_\alpha)} \quad (6)$$

確率的に最も確からしい運動パラメータ $\{h, R\}$ はこれを最小にするものである。上式から直ちに並進 h のスケールが不定であることがわかる⁴。そこで並進を $\|h\| = 1$ と正規化する ($\|\cdot\|$ はベクトルのノルムを表す)。

仮定した誤差モデルによれば、式 (6) を最小にする運動パラメータ $\{\hat{h}, \hat{R}\}$ およびそれに対応する修正位置 $\hat{x}_\alpha, \hat{x}'_\alpha$ は $\{h, R\}, \bar{x}_\alpha, \bar{x}'_\alpha$ の最尤推定量にほかならない。その解はラメル・ラオの下界と呼ばれる精度の理論限界を第 1 近似において達成することが証明される [10]。上式の最小化はくりこみ法と呼ぶ反復法によって精度よく効率的に計算できる [3, 4, 5]。同時に計算した $\{\hat{h}, \hat{R}\}$ の共分散行列も計算できるので、得られた 3 次元形状の信頼性を定量的に評価し、それを可視的に表示することもできる [4]。

式 (6) の最小値 $J[\hat{h}, \hat{R}]$ を \hat{j} と書き、残差平方和と呼ぶ。このとき、 \hat{j}/ϵ^2 が第 1 近似において自由度 $N - 5$ の χ^2 分布に従うことが示される [9]。したがって二乗ノイズレベルの不偏推定量が次のように得られる。

$$\hat{\epsilon}^2 = \frac{\hat{j}}{N - 5} \quad (7)$$

4. 平面モデルの残差平方和

物体面が平面 Π である場合を考える。平面 Π 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 n とし、原点 O から平面 Π までの距離 (ベクトル n の方向に正) を d とする。位置ベクトルを r とすると、平面 Π の方程式は $(n, r) = d$ と書ける。以下 $\{n, d\}$ 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 x, x' で表される点が平面 Π 上にある必要十分条件は次式が成り立つことである [2, 5, 10, 11, 14, 18]。

$$x' \times Ax = 0 \quad (8)$$

³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。

⁴これは速く大きい物体に対するカメラの大きい運動と近く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。

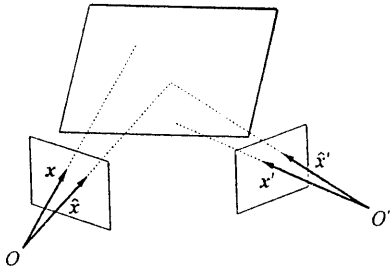


図 3: 視線が平面上で交わるように補正する。

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

$$A = R^T(hn^T - dI) \quad (9)$$

これは x, x' の定める視線がちょうど平面 Π で交わる必要十分条件であり、エッジ極線方程式 (1) よりも強い条件である。しかし、画像に誤差があれば式 (8) は厳密には成立しない。そこで視線が平面 Π 上で交わるように x, x' を補正しなければならない (図 3)。

N 組の対応点 $x_\alpha, x'_\alpha, \alpha = 1, \dots, N$ が与えられたとき、平面パラメータ $\{n, d\}$ および運動パラメータ $\{h, R\}$ を固定すると、 $\bar{x}_\alpha, \bar{x}'_\alpha$ の最尤推定量 $\hat{x}_\alpha, \hat{x}'_\alpha$ はそれらが式 (8) を満たすという拘束のもとで式 (5) の J を最小にするものである。その最小値は $\{n, d\}$ および $\{h, R\}$ の関数であるから、これを $J[n, d, h, R]$ と書く。これは次の形に書ける [9]。

$$J[n, d, h, R] = \sum_{\alpha=1}^N (x'_\alpha \times Ax_\alpha, W_\alpha(x'_\alpha \times Ax_\alpha)) \quad (10)$$

$$W_\alpha = \left(x'_\alpha \times AV_0[x_\alpha]A^T \times x'_\alpha + (Ax_\alpha) \times V_0[x'_\alpha] \times (Ax_\alpha) \right)_2^- \quad (11)$$

ここに $(\cdot)_r^-$ はランクを r に拘束した一般逆行列⁵を表す [9]。平面パラメータ $\{n, d\}$ および運動パラメータ $\{h, R\}$ の最尤推定量 $\{\hat{n}, \hat{d}\}, \{\hat{h}, \hat{R}\}$ は式 (10) を最小にするものである。この最小化はくりこみ法と呼ぶ反復法によって精度よく効率的に計算できる [11]。式 (10) の最小値 $J[\hat{n}, \hat{d}, \hat{h}, \hat{R}]$ を \hat{J}_Π と書き、平面モデルの残差平方和と呼ぶ。平面モデルが正しいとすれば \hat{J}_Π/ϵ^2 が第 1 近似において自由度 $2N-8$ の χ^2 分布に従うことが示される [9]。したがって二乗ノイズレベルの不偏推定量が次のように得られる。

$$\hat{\epsilon}^2 = \frac{\hat{J}_\Pi}{2N-8} \quad (12)$$

5. 回転モデルの残差平方和

⁵ スペクトル分解し、大きい r 個の固有値のみを残し、それ以外の固有値を 0 に置き換えて一般逆行列を計算したもの。

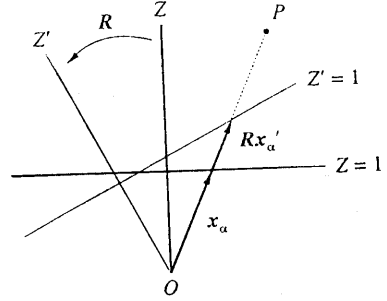


図 4: カメラの純粋回転。

カメラの運動が純粋回転である必要十分条件は

$$x \times Rx' = 0 \quad (13)$$

となる (図 4)。 N 組の対応点 $x_\alpha, x'_\alpha, \alpha = 1, \dots, N$ が与えられたとき、回転 R を固定すると、 $\bar{x}_\alpha, \bar{x}'_\alpha$ の最尤推定量 $\hat{x}_\alpha, \hat{x}'_\alpha$ はそれらが式 (13) を満たすという拘束のもとで式 (5) の J を最小にするものである。その最小値は R の関数であるから、これを $J[R]$ と書く。これは次の形に書ける [9]。

$$J[R] = \sum_{\alpha=1}^N (x_\alpha \times Rx'_\alpha, W_\alpha(x_\alpha \times Rx'_\alpha)) \quad (14)$$

$$W_\alpha = \left((Rx'_\alpha) \times V_0[x_\alpha] \times (Rx'_\alpha) + x_\alpha \times RV_0[x'_\alpha]R^T \times x_\alpha \right)_2^- \quad (15)$$

回転 R の最尤推定量 \hat{R} は式 (14) を最小にするものである。式 (14) の最小値 $J[\hat{R}]$ を \hat{J}_R と書き、回転モデルの残差平方和と呼ぶ。回転モデルが正しいとすれば \hat{J}_R/ϵ^2 が第 1 近似において自由度 $2N-3$ の χ^2 分布に従うことが示される [9]。したがって二乗ノイズレベルの不偏推定量が次のように得られる。

$$\hat{\epsilon}^2 = \frac{\hat{J}_R}{2N-3} \quad (16)$$

6. 動画像の情報量基準

対象が平面であってもなくても、カメラの運動が純粋回転であってもなくても、抽象化すると運動解析は次のように統一的に記述できる。対応点 x, x' は一つの 6次元ベクトル $x \oplus x'$ とみなせる。ベクトル x, x' の第 3 成分は共に 1 であるから、ベクトル $x \oplus x'$ は 6次元空間 \mathcal{R}^6 の 4次元アフィン部分空間

$$\mathcal{X} = \{(x, y, 1, x', y', 1)^T | x, y, x', y' \in \mathcal{R}\} \quad (17)$$

に拘束されている。これをデータ空間と呼ぶ。 N 組の対応点 x_α, x'_α を与えることは、4次元データ空間

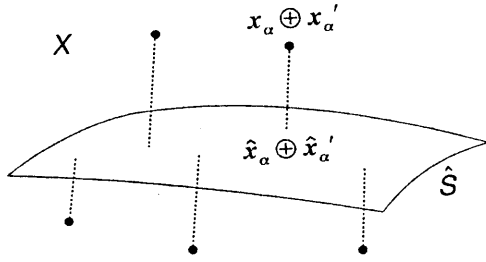


図5: データ点を最適に当てはめた多様体 \hat{S} 上に最適に射影する。

\mathcal{X} に N 個の点 $x_\alpha \oplus x'_\alpha$ を与えることと等価である。対象またはカメラの運動に何らかの拘束条件が存在して、それが n 次元ベクトル u をパラメータとする方程式

$$F^{(k)}(x, x'; u) = 0, \quad k = 1, \dots, L \quad (18)$$

で表せるとする。これらの L 個の方程式は x, x' の方程式として代数的に独立でなくてもよい⁶。独立なもの個数 r をこの拘束のランクと呼ぶ [10]。このとき式 (18) はデータ空間 \mathcal{X} に余次元 r の多様体 S を定義する。これを (幾何学的) モデルと呼ぶ。そして、この多様体を指定するパラメータ u の定義域 U (\mathcal{R}^n のある多様体であるとする) の次元 n' をモデル S の自由度と呼ぶ。

N 組の対応点 x_α, x'_α が与えられたとき、 N 点 $x \oplus x'_\alpha \in \mathcal{X}$ を含むようなモデル多様体 S は誤差のため一般には存在しない。そこでパラメータ $u \in U$ を調節して多様体 S を N 点 $x \oplus x'_\alpha \in \mathcal{X}$ に (最尤推定の意味で) 最適に当てはめる。そして各 $x_\alpha \oplus x'_\alpha$ を最適に当てはめた多様体 \hat{S} 上へ (最尤推定の意味で) 最適に射影点する (図5)。最適な射影点を $\hat{x}_\alpha \oplus \hat{x}'_\alpha$ とすると、定義より $\hat{S}, \hat{x}_\alpha, \hat{x}'_\alpha$ はそれぞれ S, x_α, x'_α の最尤推定量である [10]。このとき $\hat{x}_\alpha, \hat{x}'_\alpha$ を式 (5) に代入して得られる残差平方和を \hat{j} とすると、 \hat{j}/ϵ^2 は第1近似において自由度 $rN - n'$ の χ^2 分布に従う確率変数であることが示される⁸ [9]。したがって二乗ノイズレベルの不偏推定量が次のように得られる。

$$\hat{\epsilon}^2 = \frac{\hat{j}}{rN - n'} \quad (19)$$

このようなモデルの「よさ」はその「予測能力」で測ることができる [16]。すなわち、現在のデータ $x_\alpha,$

⁶ただし各々がデータ空間 \mathcal{X} で互いに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 [9, 10]。

⁷考えている空間 (この場合は \mathcal{X}) の次元と多様体の次元との差。この場合、余次元 r の多様体は $4 - r$ 次元である。

⁸したがって、モデル S の自由度 n' が大きいほど残差平方和 \hat{j} は小さくなる。特に rN 個以上のパラメータをも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 \hat{j} は0となる。

x'_α と同じ確率分布に従う独立な「将来のデータ」を x^*_α, x'^*_α とするとき、現在のデータ x_α, x'_α から計算した最尤推定量 $\hat{x}_\alpha, \hat{x}'_\alpha$ の将来のデータ x^*_α, x'^*_α に関する残差平方和を

$$\hat{J}^* = \sum_{\alpha=1}^N (x^*_\alpha - \hat{x}_\alpha, V_0[x_\alpha]^{-1} (x^*_\alpha - \hat{x}_\alpha)) + \sum_{\alpha=1}^N (x'^*_\alpha - \hat{x}'_\alpha, V_0[x'_\alpha]^{-1} (x'^*_\alpha - \hat{x}'_\alpha)) \quad (20)$$

とすると、

$$I(S) = E^*[E[\hat{J}^*]] \quad (21)$$

が小さいほどよいモデルであるときみなせる [6, 7, 8, 16]。ただし $E[\cdot], E^*[\cdot]$ はそれぞれ現在のデータ x_α, x'_α および将来のデータ x^*_α, x'^*_α 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次の形で与えられる⁹ [7]。

$$AIC(S) = \hat{j} + 2(pN + n')\epsilon^2 \quad (22)$$

ただし $p = 4 - r$ はモデル S を定義する多様体の次元 (これを単にモデル S の次元と呼ぶ) である。式 (22) をモデル S の幾何学的情報量基準 (幾何学的 AIC) と呼ぶ [6, 7, 8]。

7. 情報量基準によるモデル選択

二つのモデル S_1 と S_2 が与えられたとし、それぞれの次元, 余次元, 自由度を p_1, r_1, n'_1 および p_2, r_2, n'_2 とする。そして S_1 は弱いモデルであって一般に正しいと考えられるものであり、 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する。このことを次のように書く。

$$S_2 \subset S_1 \quad (23)$$

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 \hat{j}_1 および \hat{j}_2 とすると、当然データの如何にかかわらず

$$\hat{j}_2 \geq \hat{j}_1 \quad (24)$$

である。このことは単に残差平方和の大小によってはモデルのよさが評価できないことを意味する。

モデル S_1 のもとでは二乗ノイズレベル ϵ^2 が式 (19) によって推定できる。それを代入すると、モデル S_1 および S_2 の幾何学的 AIC が次のように評価できる。

$$AIC(S_1) = \hat{j}_1 + \frac{2(p_1N + n'_1)}{r_1N - n'_1} \hat{j}_1$$

$$AIC(S_2) = \hat{j}_2 + \frac{2(p_2N + n'_2)}{r_1N - n'_1} \hat{j}_1 \quad (25)$$

⁹この証明には当てはめ最尤推定量が第1近似においてクラメル・ラオの下界と呼ばれる精度の理論限界を達成するという事実を用いる [9, 10]。

もし $AIC(S_2) < AIC(S_1)$ 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。その条件は次のように書ける。

$$\frac{\hat{J}_2}{\hat{J}_1} < 1 + \frac{2(p_1 - p_2)N + 2(n'_1 - n'_2)}{r_1 N - n'_1} \quad (26)$$

このモデル選択は統計学の仮説検定の否定的な意味とは逆であり、「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よく予測できる」という積極的な意味である [6, 12]。このとき、ノイズの強さを知る必要がないだけでなく、仮説検定に用いるような恣意的な有意水準を必要としない。

8. 平面テストと回転テスト

一般の運動解析のモデル S は式 (1) のエビ極線方程式の定める \mathcal{X} の余次元 1 の (3次元) 多様体である。このモデルは $\{h, \mathbf{R}\}$ がパラメータ \mathbf{u} に相当するので、自由度 5 のモデルである¹⁰。このモデルの幾何学的 AIC は次のようになる。

$$AIC(S) = \hat{J} + (6N + 10)\epsilon^2 \quad (27)$$

一方、物体が平面であるというモデル S_{Π} は式 (8) によって定義される。式 (8) の 3 個の成分方程式は代数的に従属であり、ランクは 2 である。したがってモデル S_{Π} は \mathcal{X} の余次元 2 の (2次元) 多様体である。これを指定する平面パラメータと運動パラメータの自由度は 8 である¹¹。モデル S_{Π} の幾何学的 AIC は次のようになる。

$$AIC(S_{\Pi}) = \hat{J}_{\Pi} + (4N + 16)\epsilon^2 \quad (28)$$

$S_{\Pi} \succ S$ であるから、次の条件が満たされるときに物体形状は平面であると判定される [8]。

$$\frac{\hat{J}_{\Pi}}{\hat{J}} < 3 + \frac{4}{N-5} \quad (29)$$

カメラの運動が純粹回転であるというモデル S_R は式 (13) によって定義される。式 (13) の 3 個の成分方程式は代数的に従属であり、ランクは 2 である。したがってモデル S_R は \mathcal{X} の余次元 2 の (2次元) 多様体である。これを指定するパラメータは回転 \mathbf{R} であるから、このモデルの自由度は 3 である。モデル S_R の幾何学的 AIC は次のようになる。

$$AIC(S_R) = \hat{J}_R + (4N + 6)\epsilon^2 \quad (30)$$

$S_R \succ S$ であるから、次の条件が満たされるときにカメラの運動は純粹回転であると判定される [8]。

$$\frac{\hat{J}_R}{\hat{J}} < 3 + \frac{14}{N-5} \quad (31)$$

¹⁰ 並進は単位ベクトルであるから 2 自由度をもち、回転に 3 自由度ある。

¹¹ \mathbf{n} は単位ベクトルであるから自由度が 2 であり、距離 d に 1 自由度ある。そして運動パラメータが 5 自由度をもつ。

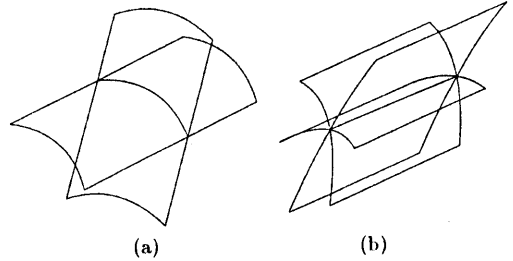


図 6: (a) 平面モデルを二つの一般モデルの共通部分とする。(b) 回転モデルは無数の一般モデルの共通部分となる。

9. 3次元解釈の多義性と特異性

よく知られているように、一般の 3次元解析アルゴリズムでは 2 組の解が得られ、「物体が運動前後のカメラの前方にある」という条件を課して初めて解が一意的に定まる [2]。これは異なる運動パラメータが同一の 3次元多様体 $S \subset \mathcal{X}$ を定義するからであり、モデル S それ自身は一般に一意的に定まる。したがって残差平方和 \hat{J} は一般に一意的に定まる。

例外的なのは誤差がないときの 4次元データ空間 \mathcal{X} の N 点 $\bar{x}_\alpha \oplus \bar{x}'_\alpha$, $\alpha = 1, \dots, N$ が特殊な配置に退化して、それらを通るモデル多様体 S が複数存在する場合である。このとき x_α , $\alpha = 1, \dots, N$ は 3次元空間中の特殊な曲面上にあり、臨界面面と呼ばれる [2, 15]。代表的な臨界面面は平面である。これは次のように解釈できる。一般モデル S は \mathcal{X} の 3次元多様体であるが、平面モデル S_{Π} は \mathcal{X} の 2次元多様体である。そして、平面モデル S_{Π} は二つの一般モデルの共通部分となるので、一般モデル S の解が一意的に定まらない (図 6(a))。したがって、平面アルゴリズムを用いても複数の解が得られ、「平面が運動前後のカメラの前方にある」という条件を課しても最大 2 組の解が得られる [2]。しかしモデル S_{Π} それ自身は一般に一意的に定まる。したがって残差平方和 \hat{J}_{Π} は特徴点が一般の位置にあれば一意的に定まる。

一方、回転モデル S_R はやはり \mathcal{X} の 2次元多様体であるが、これは無数の一般モデルの共通部分となっている (図 6(b))。実際、次の関係が成立している。

$$\{x \oplus x' | x \times R x' = 0\} = \bigcap_{h \in \mathcal{R}^3} \{x \oplus x' | (x, G x') = 0\} \quad (32)$$

このことから、カメラが純粹回転であれば何らの 3次元解釈もできないことがわかる。

10. 実験

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 512×512 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 σ (画素) の正規乱数を誤差として加えたも

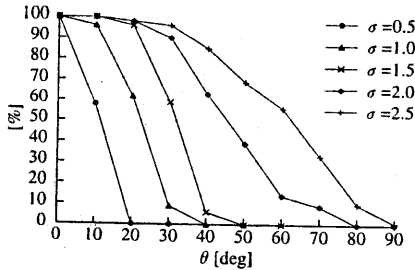


図 7: 平面と判定される割合.

のを特徴点とし、3次元復元を行なう。図7は横軸に角度 θ をとり、縦軸にそのときの平面と判定される割合を $\sigma = 0.5 \sim 3.0$ の場合に図示したものである。これから $\sigma = 1$ (画素), $\theta = 22^\circ$ のとき平面と判定される割合がほぼ50%となるのがわかるが、このとき平面と判定される例を図8(a)に、それから復元した3次元形状を図8(b)に示す。比較のために真の形状と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てある。図9は平面と判定されない例である。このように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。

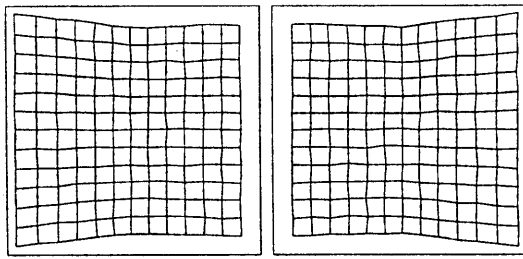
図10に実画像例を示す。図10(上段)のようにマウスで特徴点を選ぶと、これらは同一平面上にはないと判定される。特徴点を人為的に選んでいるため誤差が少なく、その結果、判定が自動的に厳しくなるからである。各画素の座標に期待値0、標準偏差3画素の正規乱数を加えて整数値に丸めた位置を図10(下段)に示す。この場合は平面であると判定される。このことは特徴点の位置が非平面構造を復元するには不適切であり、3画素程度の誤差でもはや平面としての復元しかできないことを意味する。回転テストについても同様な結果が得られる。

11. まとめ

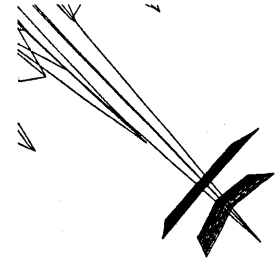
本論文では2画像から3次元形状復元を行なう場合に、物体が平面とみなせるかどうか、およびカメラの運動が純粋回転であるかどうかを画像の誤差や3次元復元の精度に関する知識を用いず、さらに何らの恣意的なしきい値をも用いずに自動的に判定する新しい方法を述べ、シミュレーション画像および実画像を用いた実験結果を示した。本論文で用いた基本原理は幾何学的情報量基準(AIC)で評価した予測能力の高いモデルを採用することであり、ロボティクスやコンピュータビジョンの広い範囲の問題における幾何学的モデルの選択基準として新しい方法論を提起するものである[12]。

参考文献

- [1] 金谷健一, 画像理解—3次元認識の数理—, 森北出版, 1990.
- [2] K. Kanatani, *Geometric Computation for Machine Vision*, Oxford University Press, Oxford, 1993.
- [3] 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, **35-2** (1994), 201-209.
- [4] K. Kanatani, Renormalization for motion analysis: Statistically optimal algorithm, *IEICE Trans. Infor. & Sys.*, **E77-D-11** (1994), 1233-1239.
- [5] 金谷健一, 「空間データの数理—3次元コンピュータビジョンに向けて—」, 朝倉書店, 1995.
- [6] 金谷健一, 幾何学的モデルの選択基準について, 情報処理学会研究報告, 95-MSP-2 (1995-7), 27-32.
- [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 (1995-7), 19-24.
- [8] 金谷健一, 情報量基準による幾何学的推論, 情報処理学会研究報告, 95-AI-101 (1995-9), 1-6.
- [9] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Lecture Note, Department of Computer Science, Gunma University, Second Edition, June, 1995.
- [10] 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, **36-8** (1995), 1865-1873.
- [11] K. Kanatani and S. Takeda, 3-D motion analysis of a planar surface by renormalization, *IEICE Trans. Inf. & Syst.*, **E78-D-8** (1995), 1074-1079.
- [12] 金澤靖, 金谷健一, 情報量基準によるステレオ視の平面自動判定, 情報処理学会研究報告, 95-CV-97 (1995-11).
- [13] H. C. Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections, *Nature*, **293-10** (1981), 133-135.
- [14] H. C. Longuet-Higgins, The reconstruction of a plane surface from two perspective projections, *Proc. Roy. Soc. Lond.*, **B227** (1986), 399-410.
- [15] S. Maybank, *Theory of Reconstruction from Image Motion*, Springer, Berlin, 1993.
- [16] 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.
- [17] R. Y. Tsai and T. S. Huang, Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces, *IEEE Trans. Patt. Anal. Mach. Intell.*, **6-1** (1984), 13-27.
- [18] J. Weng, N. Ahuja and T. S. Huang, Motion and structure from point correspondences with error estimation: Planar surfaces, *IEEE Trans. Sig. Proc.*, **39-12** (1991), 2691-2717.
- [19] J. Weng, N. Ahuja and T. S. Huang, Optimal motion and structure estimation, *IEEE Trans. Patt. Anal. Mach. Intell.*, **15-9** (1993), 864-884.

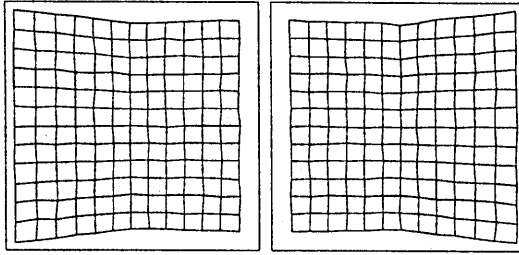


(a)



(b)

図 8: (a) 平面と判定される例. (b) 復元した3次元形状.



(a)



(b)

図 9: (a) 平面と判定されない例. (b) 復元した3次元形状.

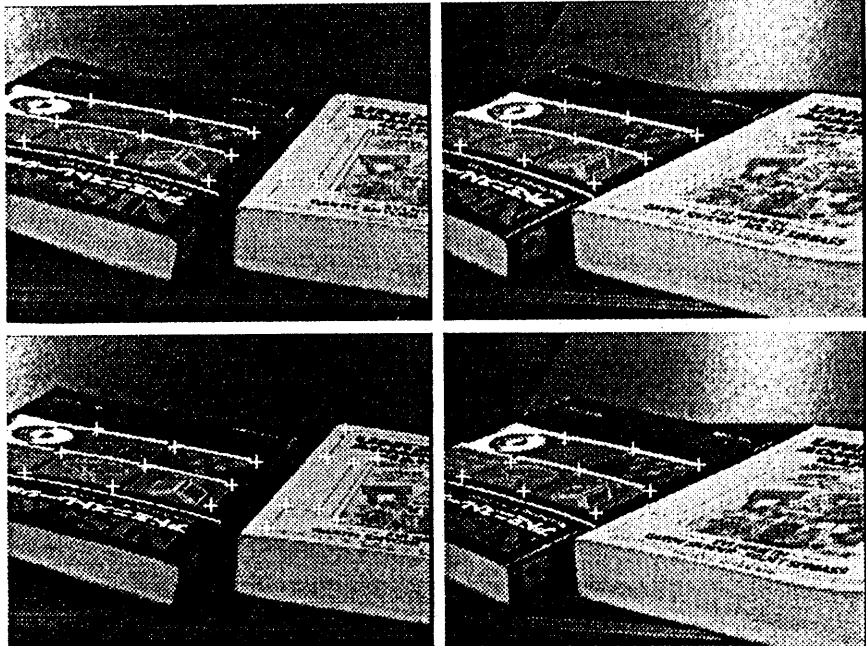


図 10: (上段) 特徴点が同一平面上にはないと判定される実画像例. (下段) 特徴点が同一平面上にあると判定される実画像例.