論　文

画像からの幾何学的推論はどういう統計的モデルに基づくのか

金谷　健一†

For Geometric Inference from Images, What Kind of Statistical Model Is Necessary?

Kenichi KANATANI†

あらまし　統計学を含む他分野との相互理解を目的として，画像の特徴点に基づく幾何学的推論に「統計的方法」を用いる意味を考察する．まずコンピュータビジョンにおける画像処理の本質と，それに由来する特徴点位置の不確定性の根元を探求し,「幾何学的当てはめ」,「幾何学的モデル選択」や「幾何学的 AIC」,「幾何学的 MDL」の意味を述べる．そして漸近評価の解釈を「撹乱母数」,「ネイマン・スコット問題」,「セミパラメトリックモデル」と関連させ，統計的手法を用いるには幾何学的推論に特有な性質を十分考慮する必要があることを指摘する．

キーワード　統計的方法，特徴点の抽出，漸近評価，ネイマン・スコット問題，セミパラメトリックモデル

1. まえがき

コンピュータビジョンにおける画像からの推論には種々の統計的方法が用いられる．最近は，従来机上の空論と思われていた統計的学習理論を今日の進歩した計算機能力によって実際問題に応用する試みもある [27]．一方，筆者は画像から抽出した点や直線などの幾何学的データに内在する関係の推定を目的として，精度を最大にする最適化手法を導いたり理論限界を導出するのに統計的方法を用いた [8], [9]．

しかし両者は同じ「統計的」という言葉を用いても，その意味がかなり異なる．これを明確にしなければ画像処理の研究者間はもちろん，統計学などの他分野の研究者との間で方法論を巡る論争が起きがちである．

画像からの幾何学的推論は我が国から発信された数少ない研究分野であるが，現在では欧米の研究が先行している．それでも最近は我が国でも関心が高まりつつある [14], [20]．本論文は新しい手法を提示するものではなく,「そもそも幾何学的推論になぜ統計的方法が必要か」を問い直し，研究者間の幅広い連携によって理論研究の更なる進展を促そうとするものである．

† 岡山大学工学部情報工学科，岡山市
Department of Information Technology, Okayama University, Okayama-shi, 700-8530 Japan

2. 統計的方法とは何か

数学や物理学の普通の問題は確定的であり，公理や基礎方程式から様々な性質が導出される．コンピュータビジョンにおいてもカメラの投影モデルを基にして画像から 3 次元復元を行う緻密な理論が展開されている [7]．それに対して，観測したデータの性質を記述するのではなく,「それらはある集合からランダムに選ばれたサンプル（標本）である」とみなし，その集合の性質を推論するのが統計的方法である．その集合は「アンサンブル（母集団）」と呼ばれ，実在する対象（例えば国内の全住民）のこともあるが，仮想的な可能性の全体であることも多い．

統計的方法を用いるとき，普通はどういうアンサンブルを問題にしているかは自明で明示されないことが多い．例えば文字認識では各文字が印刷や手で書かれるあらゆる形の全体であることは述べるまでもない．そこにはおのずから現れやすい形と現れにくい形があり，自然に確率分布が定義される．

もちろん手書き文字認識では対象を「手で書いた文字の集合」に限定しなければならないが，それが「手書き文字」という言葉から自然に理解される．更に「筆者特定」（署名認証など）と「筆者不特定」とでは対象とするアンサンブルが異なるが，これも自然に了解される．ところが画像の幾何学的推論ではこれが大きな問題となる．このことはこれまではっきり認識さ

れていなかった．これを具体的に指摘することが本論文の目的である．

3. 幾何学的推論とは何か

3.1 幾何学的推論のアンサンブル

筆者らが研究してきた幾何学的推論では１枚（または１組）の画像を対象とする．例えば建物の画像から頂点や窓の角などの「特徴点」をいくつか抽出し，それらが同一直線上にあるか，もしあればそれらを通る直線の方程式は何か，その直線にはどの程度の不確定性があるか，などを推定するのが幾何学的推論（の最も単純な例）である．

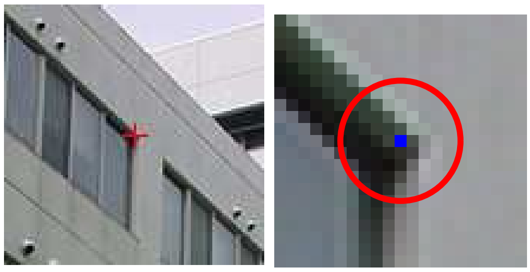
これに統計的方法が用いられる理由は，抽出した特徴点に不確定性があるからである．そして，その不確定性を考慮して最も信頼できる直線の位置を推定し，ずれが著しい点はそのようなずれの確率は低いからそもそもその直線上にはなかったと推論する．このとき，背後にあるアンサンブルは何であろうか．

これは特徴点の不確定性がなぜ生じるのかという問題に帰着する．特徴点に不確定性がなければ統計的方法は無意味である．特徴点の不確定性を考えることは，その可能な位置の集合を考え，現在の位置をそれからランダムに選ばれたサンプルとみなすことである．しかし「可能な位置の集合」とは何であろうか．

3.2 特徴点抽出の不確定性

特徴点の抽出には Harris 作用素 [6] や SUSAN [24] を始め，多くのアルゴリズムが提案され，互いに性能が比較されている [3], [21], [23]．また，連続ビデオフレーム上で特徴点を追跡する Kanade-Lucas-Tomasi 法 [26] もよく知られている．しかし，例えば Harris 作用素によって，今解析している画像中の，今注目している建物のある窓の角の位置を検出したとすると，結果は一通りである (図 1)．何度適用しても画像に何の外乱も入らないし，Harris 作用素の内部状態（判定のしきい値など）も変化しないから同じ結果しか得られない．したがって「可能な位置の集合」はその１点のみである．それ以外の位置がなぜ「可能」なのか．どうすればそれ以外の位置が選ばれるのであろうか．この疑問に対して満足できる解釈がこれまで与えられていなかった．

「抽出した点以外に可能な位置が存在するはずだ」と考える論拠は「抽出した点が必ずしも正しい位置ではない」という事実である．しかし，正しい位置でないなら，なぜその点を抽出したのか．なぜはじめから



(a) (b)

図 1 (a) 建物画像のある特徴点，(b) その拡大図と特徴点の不確定性

Fig. 1 (a) A feature point in an image of a building. (b) Its enlargement and the uncertainty of the feature location.

正しい位置を抽出しなかったのか．その答は「正しい位置が抽出できない」からである．なぜできないのか．

3.3 コンピュータビジョンの画像処理

種々の特徴点抽出アルゴリズムが存在する理由は，これが本質的に不可能な処理だからである．もし抽出したい点が「画像の濃淡値がこれこれの意味で最も激しく変化している点」のように指定されているなら，そのアルゴリズムは唯一に定まる（計算手順にバリエーションが存在しても出力は同じ）．しかし，コンピュータビジョンで抽出したいのは濃淡値で記述できる「画像の性質」ではなく，建物の角のようなそこに写っている「3 次元対象の性質」である．一般に 3 次元対象の性質をどういう画像の性質に対応させるかは不明であり，その対応のさせ方の数だけ処理アルゴリズムが生まれる．この意味で特徴点抽出は本質的にヒューリスティクスである．この事実は従来はっきり意識されず，客観的な処理と錯覚されることが多かった．

抽出したい 3 次元対象やその特徴を指定しても，その画像上での見え方はその 3 次元形状の詳細や撮影方向や照明や対象の光反射特性などの多くの要因に依存し，それらのわずかな変化でも画像上での見え方が変わる．理論的には，そのようなシーンの 3 次元情報（3 次元形状や撮影方向や照明や反射特性など）がすべてわかれば画像上での見え方が定まり，したがってその対象の位置を厳密に決定できる．しかし，画像からそのようなシーンの 3 次元情報を引き出すことがコンピュータビジョンの目的であるから，コンピュータビジョンのための画像処理の段階ではそれらは未知であり，何らかの予測に基づくほかはない．このため，異

なる予測の仕方だけ異なる処理アルゴリズムが存在する．現在の画像に対してある予測は当たっているかもしれないが別の予測は誤っているかもしれない．

このことから，特徴点の「可能な位置の集合」は「抽出のための予測の集合」及びその予測に基づく「処理アルゴリズムの集合」に対応させるのが妥当であるが，従来はこのような解釈がなされることはなかった．この解釈によれば，現在の位置はその集合からサンプルしたアルゴリズムによって得られたものとみなせる．そうすれば，そのアルゴリズムを用いる限り何度抽出しても同じ位置しか得られないことの説明がつく．別の位置を得るためには別のアルゴリズムをサンプルしなければならないからである．

4. 特徴点位置の統計的モデル

4.1 特徴点の共分散行列

前章の解釈に基づいて解析を進めるには，考えられるあらゆる予測に基づく抽出の「平均」が真に抽出したい点であること，すなわち処理アルゴリズム全体が「不偏」であることを仮定する必要がある．そうでなければアルゴリズムを工夫する意味がない．

特徴点抽出の性能は画像に依存する．抽出したい点の周りの濃淡値にあまり変化がなければ，どういうアルゴリズムを適用してもよい結果は得られないであろう．その結果，あらゆるアルゴリズムで抽出した位置は広く分散するであろう．一方，濃淡の変化がめいりょうな部分ではどのアルゴリズムを適用してもほぼ正確な位置が得られるであろう．その結果，あらゆるアルゴリズムで抽出した位置は狭い範囲に集中しているであろう．このことから抽出したい点の近傍の濃淡値の変化に応じて，あらゆるアルゴリズムで抽出されるだろう位置の分布を予測する「共分散行列」が定義される．

α 番目の特徴点 p_α の共分散行列を $V[p_\alpha]$ とする．上の議論より不確定性の絶対的な大きさを定めることはできず，定義できるのは定性的な性質のみである．例えば，その点の周りの濃淡がどの方向にも同じ程度に変化している場合はほぼ等方性であり，代表的な等確率線（「誤差楕円」）はほぼ円であろう (図 1(b))．一方，物体境界上の点はどんなアルゴリズムを適用しても同じ境界上の他の点との区別が難しいから，その点の共分散行列はその境界上に沿う細長い誤差楕円をもつであろう．

そこで不確定性の絶対的な大きさを未知定数 ϵ とし，共分散行列 $V[p_\alpha]$ を

$$V[p_\alpha] = \epsilon^2 V_0[p_\alpha] \tag{1}$$

と書く．$V_0[p_\alpha]$ は特徴点間の不確定性の相対的な比較，及びその点の周りの各方向の不確定性の相対的な比較を記述するものであり,「正規化共分散行列」と呼ぶ [8], [9]．それに対して未知定数 ϵ を「ノイズレベル」と呼ぶ [8], [9]．ただし，式 (1) を「特徴点 p_α の共分散行列」と呼ぶとしても，それは点 p_α の性質ではなく，点 p_α の近傍に適用される「抽出アルゴリズム全体の性質」である．このことはこれまではっきり意識されていなかった．

4.2 特徴点抽出アルゴリズムの特性

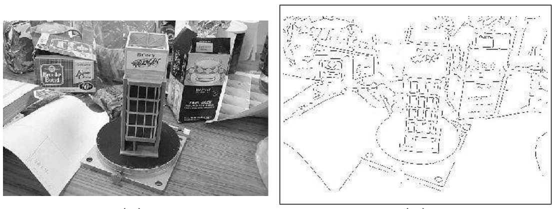
ほとんどの特徴点抽出アルゴリズムは濃淡がほぼ一様な領域や物体の境界線上の点は出力せず，その点の周りの濃淡の変化を計算し，どの方向にも変化が大きい点を出力する [3], [6], [21], [23], [24]．したがって特徴点抽出アルゴリズムを用いる限り，抽出した特徴点の共分散行列はほぼ等方性と仮定して問題がないことが実験的にも確認されている [13]．

異なる特徴点はその周りの濃淡変化に関係がないことが多く，各々の不確定性はほぼ独立とみなせる．しかし，ビデオフレーム上で特徴点を追跡するとフレーム間に強い相関があることが観測されている [25]．

応用によっては人間がディスプレイ上でマウスを操作して特徴点を抽出することもある．経験的には人間は孤立点や交点のような「わかりやすい」点を選ぶ傾向にあり，そのような点の周りの濃淡がどの方向にもほぼ同じ程度に変化している [13]．このため人間が抽出する特徴点は画像処理による結果と類似している．一般に画像処理アルゴリズムは人間の視覚認識を模倣しているから，これは当然であろう．一方，複数の画像上で対応点を選ぶ場合は微細な部分で強い相関があることが報告されている [17]．

4.3 画像処理とコンピュータビジョン

以上のように，幾何学的推論の背景に「処理アルゴリズム」のアンサンブルがあり，正規性，独立性，不偏性，相関などの統計的な概念は処理アルゴリズムの集合に対して定義されるというのが本論文の主張である．しかし従来，これが「画像の性質」と混同されてしばしば誤解や混乱が生じた．その最大の要因は特徴点の位置の不確定性を安易に「画像のノイズ」あるいは「画像の誤差」と呼ぶ習慣である．特に式 (1) の仮定を「特徴点位置の不確定性のモデル」の代わりに単



(a) (b)

図 2 (a) 室内シーン，(b) そのエッジ画像
Fig. 2 (a) An indoor scene. (b) Detected edges.

に「画像のノイズ（誤差）モデル」と呼ぶ習慣のためと思われる．加えて ϵ を「ノイズレベル」と呼んだことも誤解を招いた一因であろう．

もちろん鮮明な画像ほどどのアルゴリズムもより望ましい結果を出力するから，画像のよさとアルゴリズムの性能とは切り離せない．しかし，コンピュータビジョンの目的は画像自体の解析ではなく，そこに写っている対象の性質の解析である．例えば物体の境界を抽出する「エッジ検出」も，実際は画像の濃淡値の変化を抽出している (図 2)．最終的には理想的な検出手段が見つかると考える研究者もいるが，本論文の考察からは，これもヒューリスティクスである以上どんなに研究しても決定版は得られないと考えられる．

このように「抽出したいもの」と「実際に計算していること」が乖離 (かいり) している以上，どんな鮮明な画像でも処理に本質的な不確定性が介入する．したがって処理結果は「統計的」に判定しなければならないというのが本論文の立場である．そして，その背後にあるのは（本質的に不完全な）処理アルゴリズムのアンサンブルであり，これは画像の不鮮明とは区別しなければならないと主張している．しかし従来から，「誤差」が各画素の輝度値のランダム変動と同一視され，３次元特徴の抽出アルゴリズムであるにもかかわらず，各画素の輝度値にランダム誤差を与えたテスト画像で評価されることが多かった．

5. 漸近解析とは何か

「統計的推定」はアンサンブルの性質をそれから抽出したサンプルに基づいて推定することであるが，推定のためのアンサンブルに関する何らかの仮定を「モデル（模型）」と呼ぶ．データの不確定性が観測条件の不備によるならそれを改善して精度を上げることができるが，不確定性が対象に内在的なときは実験を多数回行い，仮定したモデルに基づく統計的推論を行う

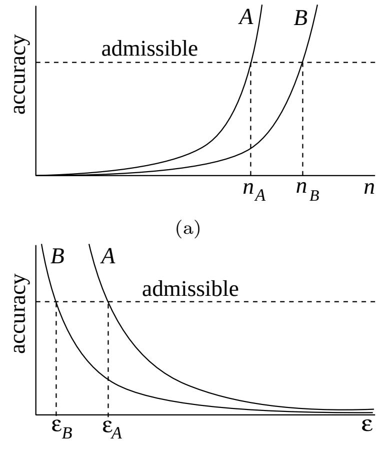


図 3 (a) 通常の統計的推論では実験回数 $n \to \infty$ で急速に精度が向上することが望ましい．なぜなら，より少ない実験回数で許容精度を達成できるからである．(b) 幾何学的推論ではノイズレベル $\epsilon \to 0$ で急速に精度が向上することが望ましい．なぜなら，より大きな不確定性があっても許容精度が達成できるからである

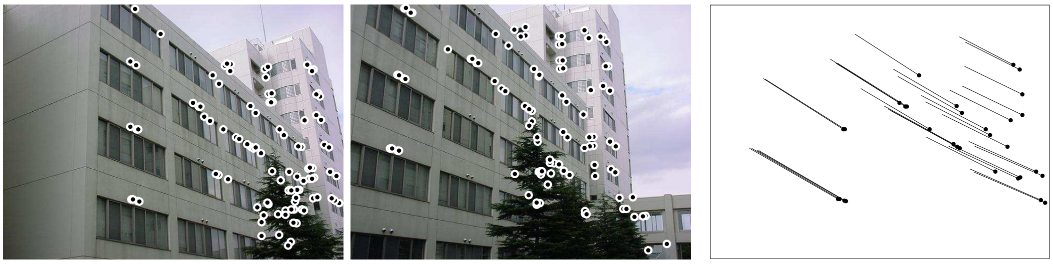
Fig. 3 (a) For the standard statistical estimation, it is desired that the accuracy increases rapidly as the number of experiments $n \to \infty$, because admissible accuracy can be reached with a smaller number of experiments. (b) For geometric inference, it is desired that the accuracy increases rapidly as the noise level $\epsilon \to 0$, because admissible accuracy can be reached in the presence of larger uncertainty.

しかない．しかし通常，実験にはコストが伴い，また可能な回数が限定されていることも多い．

これを考慮して，推定性能の評価として実験回数 n を限りなく多くとるときの精度の増加速度で評価する「漸近解析」がよく行われる．これは，実験回数 $n \to \infty$ で急速に精度が向上する推定法はそうでない方法に比べてより少ない実験回数で許容精度を達成できるからである（図 3(a)）．

一方，特徴点に基づく幾何学的推定では前述のように，背後にあるアンサンブルは他のアルゴリズムを適用したら得られるであろう特徴点の仮想的な位置の全体であり，推定の目的は特徴点があらゆる可能な位置をとるときの解釈の期待値（＝真の解釈と仮定する）を推定すること，及びその推定の真の解釈からのずれをその仮想的アルゴリズムのアンサンブル平均の意味でなるべく小さくすることであると考えられる．

しかしアンサンブルは仮想的であり，特定のアルゴリズムを用いる限り他の値を取り出すことができない．



(a) (b)

図 4 (a) 異なる位置から撮影した建物の画像から抽出した特徴点．(b) それらのうち，対応する位置を線分で結んだ「オプティカルフロー」(黒丸が左画像に対応)．各対応の端点は 4 次元空間の 1 点と同一視できる

Fig. 4 (a) Two images of a building and extracted feature points. (b) "Optical flow" consisting of segments connecting corresponding feature points (black dots correspond to the positions in the left image). The two endpoints can be identified with a point in a four-dimensional space.

すなわち実験回数は $n=1$ に限定される．このとき推定の性能をどう評価すればよいであろうか．

明らかに推定の目的は少ない実験回数で許容精度を達成することでなく，処理の不確定性が大きくても精度が高いことである．とすると，特徴点位置の不確定性が小さいと仮定した場合の精度の増加速度を評価すればよい．具体的には式 (1) 中のノイズレベル ϵ が小さくした場合を解析すればよい．なぜなら，ノイズレベル $\epsilon \to 0$ で急速に精度が向上する推定法はそうでない方法に比べてより不確定な特徴点抽出処理に対しても許容精度が達成できるからである（図 3(b)）．

6. 幾何学的推論の漸近評価

前章の考察を具体例で示す．

6.1 幾何学的当てはめ

画像から抽出した N 個の特徴点 $p_1, ..., p_N$ の真の位置を $\bar{p}_1, ..., \bar{p}_N$ とする．これらがある u 次元ベクトル $\boldsymbol{u}$ でパラメータ化された拘束条件

$$\boldsymbol{F}(\bar{p}_\alpha, \boldsymbol{u}) = \boldsymbol{0}, \qquad \alpha = 1, ..., N \tag{2}$$

を満たすとき，パラメータ（母数）$\boldsymbol{u}$ を推定する問題が「幾何学的当てはめ」である [9]．拘束条件 (2) を「(幾何学的) モデル」と呼ぶ．

典型的な問題は指定した N 点の真の位置がある直線あるいはあるパラメータ化された曲線 (円や楕円など) の上にあるとき，その直線あるいは曲線を求めることであるが，そのまま複数の画像間の拘束にも拡張できる [9]．例えば第 1 画像の点 (x_α, y_α) が第 2 画像の点 (x'_α, y'_α) に対応するとき $(x_\alpha, y_\alpha, x'_\alpha, y'_\alpha)$ を 4 次元空間の特徴点 p_α とみなせばよい (図 4)．カメラの撮像を透視投影とすれば，拘束条件 (2) に相当するのは「エピ極線方程式」と呼ばれる関係であり，パラメータ $\boldsymbol{u}$ として 2 画像を撮影したカメラの相対位置関係を定める「基礎行列」を求める問題となる [16]．また遠方や平面を撮影した 2 画像では対応する特徴点は「(平面) 射影変換」で結ばれ，パラメータ $\boldsymbol{u}$ に対応する「射影変換行列」を求める問題となる [12]．

特徴点 p_α の共分散行列を式 (1) のように書き，特徴点の不確定性を正規分布とみなすと，N 個の特徴点の可能な位置のアンサンブル上での「最ゆう推定」は正規化共分散行列 $V_0[p_\alpha]$ に関する「マハラノビス距離」の 2 乗和

$$J = \sum_{\alpha=1}^{N} (p_\alpha - \bar{p}_\alpha, V_0[p_\alpha]^{-1}(p_\alpha - \bar{p}_\alpha)) \tag{3}$$

を $\{\bar{p}_\alpha\}$ と $\boldsymbol{u}$ について拘束条件 (2) のもとで最小化すればよい．ただし各 p_α をベクトルとみなし，$(\cdot, \cdot)$ はベクトルの内積を表す．

ノイズレベル ϵ が小さいと仮定して ϵ に関して摂動解析を行なうと，式 (3) を最小化する解 $\hat{\boldsymbol{u}}$ の共分散行列 $V[\hat{\boldsymbol{u}}]$ は $\epsilon \to 0$ で $\boldsymbol{O}$ に収束し（「一致性」），かつ $V[\hat{\boldsymbol{u}}]$ は $O(\epsilon^4)$ の項を除いて精度の理論限界を達成すること（「漸近有効性」）が証明できる [9]．このことから，最ゆう推定を用いれば許容精度を達成するのに他の方法に比べて特徴点の不確定性が大きくてもよいことが結論される．

6.2 幾何学的モデル選択

幾何学的当てはめではモデル (2) は既知であり，そ

のパラメータ $\boldsymbol{u}$ を推定するものであるが，拘束条件に複数のモデル $\boldsymbol{F}_1(\bar{p}_\alpha, \boldsymbol{u}_1) = \boldsymbol{0}$, $\boldsymbol{F}_2(\bar{p}_\alpha, \boldsymbol{u}_2) = \boldsymbol{0}$, ... がある場合，どれが妥当かを判定するのが「（幾何学的）モデル選択」である [9].

これには各々のモデルを仮定してはパラメータ $\boldsymbol{u}$ を最ゆう推定し，式 (3) の最小値（「残差 (平方和)」）$\hat{J}$ の小さいモデルを採用すればよさそうであるが，式 (3) の $\boldsymbol{u}$ にその最ゆう推定量 $\hat{\boldsymbol{u}}$ を代入すると，$\hat{\boldsymbol{u}}$ は J を最小にするように定めているので異なる拘束条件を公平に比較できない．そこでその残差の偏りを是正したものを用いることが考えられる．これが赤池の AIC [1] の思想であり，その導出の出発点は「カルバック情報量」である．そこでノイズレベル ϵ が小さいと仮定して ϵ に関して摂動解析を行うと次の「幾何学的 AIC」が得られる [10].

$$\text{G-AIC} = \hat{J} + 2(Nd+n)\epsilon^2 + O(\epsilon^4) \tag{4}$$

ここに d は拘束条件 (2) がデータの空間に定義する多様体の次元である．これが入ることが赤池の AIC との見かけ上の違いであり，N 個の特徴点の位置の不確定性を反映している．

一方，Rissanen の MDL [22] はデータとパラメータを符号化して得られる最小の記述長をモデルのよさとするものである．Rissanen にならって実数パラメータを量子化し，その量子化幅を全体の記述長が最小になるように定め，ノイズレベル ϵ が小さいと仮定して ϵ に関して摂動解析を行なうと次の「幾何学的 MDL」が得られる [10].

$$\text{G-MDL} = \hat{J} - (Nd+n)\epsilon^2 \log\left(\frac{\epsilon}{L}\right)^2 + O(\epsilon^2) \tag{5}$$

ここに L はある基準長であり，データの大きさに対する比が $O(1)$ となるように選ぶ（特徴点データの場合は画像のサイズにとればよい）．これを厳密に定めるにはデータがデータ空間のどの部分に現れやすいかという事前分布が必要になるが，通常はオーダが同じである限りモデル選択にほとんど影響を与えない [10].

6.3　等価的な統計的解釈

幾何学的推論の漸近評価の仕方が通常の統計的推論と正反対であるにもかかわらず，前 2 節の結果は通常の統計学の結果に類似している．通常の統計的推定では実験回数を n とすると，標準的な条件のもとで最ゆう推定量の共分散行列は $1/n \to 0$ で $\boldsymbol{O}$ に収束し（「一致性」），$O(1/n^2)$ の項を除いて「クラメル・ラオの下界」を達成すること（「漸近有効性」）が知られている．すなわち $1/\sqrt{n}$ が幾何学的推定におけるノイズレベル ϵ と同じ役割を果たしている．

この対応はモデル選択でも同様である．式 (4) を ϵ^2 で割ると $\hat{J}/\epsilon^2 + 2(Nd+n) + O(\epsilon^2)$ となり，$\hat{J}/\epsilon^2$ は対数ゆう度の -2 倍であり，$Nd+n$ は各特徴点の真の位置を含めた未知パラメータの総数であるから，高次の項を除くと赤池の AIC と同じ（対数ゆう度の -2 倍)+2(未知パラメータの総数) となる．また式 (5) でも ϵ を $1/\sqrt{n}$ に置き換えると Rissanen の MDL と同じ形になる．

これは次のように解釈できる．特定のアルゴリズムを用いる限り 1 個の特徴点位置しか抽出できないが，仮に毎回異なるアルゴリズムを適用して n 個の位置を抽出したとすると，これから真の位置を最適に推定するには正規分布モデルのもとではサンプル平均をとればよい．サンプル平均の共分散行列はもとの共分散行列の $1/n$ になるから，このような仮想的な推定は式 (1) においてノイズレベル ϵ を $1/\sqrt{n}$ 倍することと等価である．

実際，特徴点抽出アルゴリズムの内部パラメータ（判定のしきい値など）をランダムに変動させては特徴点を抽出することによって擬似的にアルゴリズムのアンサンブルを生成し，得られる特徴点位置の共分散行列を計算したり，平均をとって抽出精度を上げる試みもある [4]（各画素の輝度値にランダム誤差を加えるのではないことに注意）．このような繰り返し全体は一つの抽出処理とみなせるから，実質的に高い精度の処理を実現したといえる．

要するに $\epsilon \to 0$ の摂動解析は仮想的に n 回サンプルしたときの $n \to \infty$ の漸近解析に相当するので，n に関する漸近評価 $\cdots + O(1/\sqrt{n^k})$ が ϵ に関する漸近評価 $\cdots + O(\epsilon^k)$ として現れるのである．このような解析は筆者が従来から行ってきたが，その背後にあるアンサンブルの解釈を明確にしてはいなかった．

7.　撹乱母数とセミパラメトリックモデル

7.1　何に対する漸近性か

通常の統計的推定の $n \to \infty$ の漸近解析の n は「実験回数」であり,「試行回数」,「観測回数」,「サンプル数」とも呼ばれる．これを増やすほど考えているアンサンブルの性質がより明らかになることは直観的に理解しやすい．

しかし n は「データ数」と呼ばれることもあるので

誤解を招きやすい．例えば 1 回の実験データが 100 次元ベクトルのときデータ数は 100 ではなく $n = 1$ である．なぜなら，考えているのは 100 次元ベクトルのアンサンブルから選ばれた 1 サンプルだからである．

文字パターンのアンサンブルを考える文字認識の学習では望ましい応答が得られるための学習回数 n を問題にする．これは各文字を表現する特徴ベクトルの次元 N とは別問題である．学習の統計的解析は $n \to \infty$ の挙動を対象し，$N \to \infty$ ではない．

同様に，特徴点に基づく幾何学的推論では，例えば画像から 50 個の特徴点を抽出すれば，それらの x, y 座標からなる 100 次元ベクトルが 1 個のデータである．濃淡値などの他の情報を用いない限り，その 100 次元ベクトルがその画像の特徴ベクトルであり，文字認識において一つの文字を 100 次元ベクトルで表すのと同じである．そして，これをある 100 次元ベクトルのアンサンブルから選ばれたサンプルとみなし，仮定したモデルからそのアンサンブルの性質を推定するのが統計的方法である．

7.2 ネイマン・スコット問題

ところが背後にあるアンサンブルを明確にせず，抽出する特徴点の個数 $N \to \infty$ の挙動を議論する研究がある [18]．その画像がその特徴点位置で代表されているので，N を変えると別の画像になる．にもかかわらずこのような議論が行われるのは，統計学で類似の問題が存在するからであろう．例えば遠方の棒状の構造物に向かってレーザ光を照射し，その反射時間からその構造物の位置を推定する問題を考えよう．レーザ光は任意の方向に何度でも照射できるとするが，その照射方向には誤差が入るとする．そして，なるべく少ない照射回数でなるべく高い推定精度を達成したいとする．このような推定の精度は照射回数を n とするとき $n \to \infty$ の挙動で評価するのが自然である．

背後にあるアンサンブルはあらゆる照射方向に対する反射時間である．普通は各照射の厳密な方向には関心がなく，その構造物の位置のみに関心があるので，前者（照射方向の真値）を「撹乱母数」，後者（構造物の位置）を「構造母数」と呼ぶ [18]．このとき構造母数は一定数であるが，撹乱母数は照射実験の繰り返しに応じて限りなく増加する．このような変則的な問題は「ネイマン・スコット問題」[19] と呼ばれ，普通は拘束条件が式 (2) のような陰関数となるから「誤差変量 (errors-in-variables) モデル」[5] でもある．また変数変換を行うとデータごとに誤差特性が変化するから heteroscedastic でもある [15]．

これに対して照射方向の「確率分布」を仮定し，各照射方向はそこからのサンプルであるとみなす解析が行われる．このような定式化は「セミパラメトリックモデル」と呼ばれ，最適解の探求は最適な「推定関数」の探求に帰着する [2], [20]．

7.3 幾何学的推定とセミパラメトリックモデル

本論文の考察からは，上述の問題は画像に特徴点抽出処理を行って得た特徴点位置によって画像を特徴づけ，その幾何学的構造を解析する問題と明らかに異なる．したがってセミパラメトリックモデルを適用して望ましい結果が得られるかは検討を要する課題である [14], [20]．その際にどういうアンサンブル（またはアンサンブルのアンサンブル）を考えているのか，特に特徴点の数を増やすことをどう解釈するのかを明確にする必要がある．これが本論文の結論である．

これは単に概念上の問題ではなく，実際のシミュレーション実験の仕方に影響を与える．シミュレーションでは特徴点数も加えるランダム誤差も任意に調節できるが，ある特徴点数 N とノイズレベル ϵ に対してアルゴリズム A がアルゴリズム B より優れるとしても別の N と ϵ に対しては逆のことがあり得る．N と ϵ の変化のさせ方として，ϵ を固定して N を変える [18]（より少ない特徴点でも精度が出る）のと，N を固定して ϵ を変える [11]（より不確定な処理でも精度が出る）のが考えられるが，どちらで評価すべきかを指定しなければ正当な比較が不可能となる．

8. む す び

コンピュータビジョンのための画像処理には特殊な事情が多く，「統計的」という言葉を用いてもその意味があいまいなことが多く，それに起因する誤解や論争が生じやすい．本論文の目的は，その背景を明示することによって画像の研究者間だけでなく，統計学を含んだ幅広い分野の研究者との連携によって理論研究の更なる進展を促そうとするものである．

まずコンピュータビジョンにおける画像処理の本質と，それに由来する特徴点位置の不確定性の根元を探求し，「幾何学的当てはめ」，「幾何学的モデル選択」や「幾何学的 AIC」，「幾何学的 MDL」の意味を述べた．そして漸近評価の解釈を「撹乱母数」，「ネイマン・スコット問題」，「セミパラメトリックモデル」などの概念と関係づけ，統計的手法を用いるには幾何学的推論に特有な性質を考慮する必要があることを指摘した．

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. 11500113)，テレコム先端技術研究支援センター，栢森情報科学振興財団の助成によった．

文　献

[1] 赤池弘次, “情報量基準 AIC とは何か—その意味と将来への展望, ” 数理科学, vol.153, pp.5–11, March 1976.

[2] 甘利俊一, 川鍋元明, “線形関係の推定—最小 2 乗法は最良であるのか?,” 応用数理, vol.6, no.2, pp.96–109, June 1996.

[3] F. Chabat, G. Z. Yang, and D. M. Hansell, A corner orientaion detector, “Image Vision Comput.,” vol.17, no.10, pp.761–769, Aug. 1999.

[4] K. Cho, P. Meer, J. Cabrera, “Performance assessment through bootstrap,” IEEE Trans. Pattern Anal. & Mach. Intell., vol.19, no.11, pp.1185–1198, Nov. 1997.

[5] W. A. Fuller, Measurement Error Models, Wiley, New York, 1987.

[6] C. Harris and M. Stephens, “A combined corner and edge detector,” Proc. 4th Alvey Vision Conf., pp.147–151, Manchester, U.K., Aug. 1988.

[7] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, Cambridge, U.K., 2000.

[8] 金谷健一, 空間データの数, 朝倉書店, 1995.

[9] K. Kanatani, Statistical Optimization for Geometric Computation: Theory and Practice, Elsevier, Amsterdam, the Netherlands, 1996.

[10] 金谷健一, “幾何学的当てはめにおけるモデル選択,” 信学論 (A), vol.J84-A, no.11, pp.1385–1393, Nov. 2001.

[11] 金谷健一, 松永　力, “幾何学的 AIC と幾何学的 MDL の退化検出性能の比較,” 信学論 (D-II), vol.J85-D-II, no.9, pp.1497–1499, Sept. 2002.

[12] K. Kanatani, N. Ohta, and Y. Kanazawa, “Optimal homography computation with a reliability measure,” IEICE Trans. Inf. & Syst., vol.E83-D, no.7, pp.1369–1374, July 2000.

[13] 金澤　靖, 金谷健一, “画像の特徴点に共分散行列は本当に必要か,” 信学論 (A), vol.J85-A, no.2, pp.231–239, Feb. 2002.

[14] 栗原祐介, 太田直哉, “攪乱母数を含まない推定方式によるオプティカルフローからの形状復元,” 情処学研報, 2002-CVIM-135-14, pp.87–94, Nov. 2002.

[15] Y. Leedan and P. Meer, “Heteroscedastic regression in computer vision: Problems with bilinear constraint,” Int. J. Comput. Vision., vol.37, no.2, pp.127–150, June 2000.

[16] 三島　等, 金谷健一, “基礎行列の最適計算とその信頼性評価,” 情処学研報, 99-CVIM-118-10, pp.67–74, Sept. 1999.

[17] D. D. Morris, K. Kanatani, and T. Kanade, “Gauge fixing for accurate 3D estimation,” Proc. IEEE Conf. Comput. Vision Pattern Recog., Kauai, Hawaii, U.S.A., vol.2, pp.343–350, Dec. 2001.

[18] 長尾淳平, 韓　太舜, “かく乱母数を含む場合の MDL 基準の構築と空間図形モデル推定問題への応用,” 信学論 (A), vol.J83-A, no.1, pp.83–95, Jan. 2000.

[19] J. Neyman and E. L. Scott, “Consistent estimates based on partially consistent observations,” Econometrica, vol.16, no.1, pp.1–32, Jan. 1948.

[20] 岡谷貴之, 出口光一郎, “画像からのカメラの姿勢・3 次元形状復元における推定精度の限界について,” 第 6 回画像の認識・理解シンポジウム講演論文集, pp.335–340, 名古屋, July/Aug. 2002.

[21] D. Reisfeld, H. Wolfson, and Y. Yeshurun, “Contex-free attentional operators: The generalized symmetry transform,” Int. J. Comput. Vision, vol.14, no.2, pp.119–130, March 1995.

[22] J. Rissanen, Stochastic Complexity in Statistical Inquiry, World Scientific, Singapore, 1989.

[23] C. Schmid, R. Mohr, and C. Bauckhage, “Evaluation of interest point detectors,” Int. J. Comput. Vision, vol.37, no.2, pp.151–172, June 2000.

[24] S. M. Smith and J. M. Brady, “SUSAN—A new approach to low level image processing,” Int. J. Comput. Vision, vol.23, no.1, pp.45–78, May 1997.

[25] 菅谷保之, 金谷健一, “部分空間分離法による特徴点追跡のアウトライア除去,” 情処学研報, 2002-CVIM-133-24, pp.177–184, May 2002.

[26] C. Tomasi and T. Kanade, “Detection and Tracking of Point Features,” CMU Tech. Rep. CMU-CS-91-132, April 1991;
http://vision.stanford.edu/~birch/klt

[27] 上田修功, ベイズ学習 [I]–[IV], 信学誌, vol.85, no.4, pp.265–271, April 2002; vol.85, no.6, pp.421–426, June 2002; vol.85, no.7, pp.504–509, July 2002; vol.85, no.8, pp.633–638, Aug. 2002.

（平成 14 年 10 月 15 日受付，15 年 1 月 22 日再受付）

金谷　健一　（正員）

1972 東大・工・計数卒．1979 同大大学院博士課程了．工博．群馬大学工学部情報工学科教授を経て，現在，岡山大学工学部情報工学科教授．IEEE フェロー．