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

岡山大学工学部情報工学科

統計学者を含む他分野とのコミュニケーションを促進するため，画像の特徴点に基く幾何学的推論に「統計的方法」を用いる意味を考察する．本論文では，どういう「アンサンブル」を考えているかを明確にしなければ統計的解析を巡って混乱が生じることを指摘し，特徴点抽出の不確定性の根元をコンピュータビジョンの画像処理にさかのぼって探求する．そして，性能評価のための適切な漸近解析の方式を「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」に対して示す．最後に，漸近評価の意味を十分に考えなければ，シミュレーション実験も手法の評価も無意味になることを「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」などの統計的概念と関係づけて指摘する．

キーワード: 統計的方法, 特徴点の抽出, 漸近評価, 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

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

Kenichi Kanatani

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

In order to facilitate smooth communications with researchers in other fields including statistics, this paper investigates the meaning of "statistical methods" for geometric inference based on image feature points. We point out that statistical analysis does not make sense unless the underlying "statistical ensemble" is clearly defined. We trace back the origin of feature uncertainty to image processing operations for computer vision in general and discuss the implications of asymptotic analysis for performance evaluation in reference to "geometric fitting", "geometric model selection", the "geometric AIC", and the "geometric MDL". Referring to such statistical concepts as "nuisance parameters", the "Neyman-Scott problem", and "semiparametric models", we point out that simulation experiments for performance evaluation will lose meaning without carefully considering the assumptions involved and intended applications.

Key words: statisitcal method, feature point extraction, asymptotic evaluation, nuisance parameter, Neyman-Scott problem, semiparametric model

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さまざまの統計的方法が行われる．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 [29]．

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．画像処理の研究者同士は暗黙に了解し合えても，これを明確にしなければ統計学者など他分野の研究者とのコミュニケーションが阻害される．実際，これが元で方法論を巡って論争も起きた．本論文ではこれを明示するとともに，「そもそも幾何学的推論になぜ統計的方法が必要か」から問い直し，他分野との交流を促進して研究のさらなる進展と実際への応用を目指すものである．この分野は我が国から発信された数少ない研究成果であるが，研究面でも応用面でも欧米が一歩先行し，国内では今だ研究も応用も少ない．それでも最近関心が高まりつつあり [14, 21]，今後の発展が期待される．

†700-8530 岡山市津島中 3–1–1, 岡山大学工学部情報工学科,
Tel/Fax: (086)251-8173
E-mail: kanatani@suri.it.okayama-u.ac.jp

2. 統計的方法とは何か

2.1 統計的方法とアンサンブル

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さまざまの性質が導出される．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 [6]．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 (または母集団) 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具体的に印刷や人によって書かれるありとあらゆる形の全体であることは改めて述べるまでもない．そこにはおのずからよく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のアンサンブルを「人が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．さらに「筆者特定」(署名認証など) と「筆者不特定」とでは考えているアンサンブルが異なるが，これも自然に了解されるので明示しなくても問題は起きない．ところば画像の幾何学的推論ではこれば大きな問題となる．

2.2 外在的な不確定性と内在的な不確定性

アンサンブルの要素はある共通の性格を持つ以外は各々異なっているので，共通の性格のみを指定しても個々の要素が特定できない．これを不確定性と呼ぶことにすると，これに外在的なものと内在的 なものがある．

- 外在的な不確定性

外在的な不確定性は物理学の実験によく現れる．これは同一の対象を観測しても，観測装置の不完全さ (それ以上制御できない機器内の不純物や熱雑音や用いる物理特性の近似値など) やそれを用いる環境の変動 (気温，圧力，風，周辺物体からの輻射熱，光，設置台の微小振動など) によって結果が異なるためである．この場合のアンサンブルは理想的には一個の要素であるが，観測装置の不完全性と観測環境の変動によって異なる結果が含まれる．したがって，「得られる測定値の集合」は異なる観測過程の集合と同一視できる．理論上は観測装置の改良や観測環境の制御によってアンサンブルを小さくすることが可能であり，その極限は真値からなる一個の要素の集合 (確率分布がデルタ関数) となる．

- 内在的な不確定性

内在的な不確定性は，観測は完全であるのにそもそも各個体が異なるためのものである「統計的処理」という場合はこちらを意味することが多い．例えば手書き文字認識では同じ文字でも異なる筆者により，または同じ筆者でも異なる気分で書いたものを対象にし，通常は文字読み取り装置は正しいと仮定する．医学，薬学では同じ病名でもそれ以外 (年齢，性別，職業，病歴，体力差，健康状態，その他) にさまざまな個体差ある患者の集合を対象にし，通常は医者の診断や種々の検査結果は正しいと仮定する．物理データでも，各個体がミクロに異なる状態のもの (熱ゆらぎ，乱流，原子崩壊など) は，観測装置は正しいと仮定しても結果は異なる．例えば気象データの不確定性は温度計や圧力計の不確定性によるものではなく，現象に内在するゆらぎである．このため，観測装置や観測環境を制御しても不確定性を減少させることができず，データ数を増やすことによって統計的に不確定性を減少させるしかない (大数の法則)．

それでは画像の幾何学的推論で統計的方法を用いるときのアンサンブルは何であり，その不確定性はどういう性格のものであろうか．

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は，同じ画像をデータとしても顔や人物などの画像認識とは性格が異なる．画像認識の背後にあるアンサンブルは，例えば異なる人物の異なる姿勢や異なる照明条件のもとで撮影された画像の集合である．

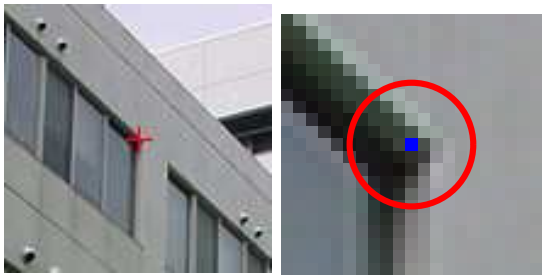
それに対して，幾何学的推論では **1** 枚 (または **1** 組) の画像を対象とする．例えば建物の画像から幾つかの特徴点 (孤立点，頂点，直線の交点など) を抽出し，それらが同一直線上にあるか，もしあればそれらを通る直線の方程式は何か，その直線にはどの程度の不確定性があるか，などを推定するのが幾何学的推論 (の最も単純な例) である．

これになぜ統計的方法が用いられるかというと，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，それからのずれが著しい点はそのようなずれの確率は低いから，それはそもそもその直線上にはなかったなどと推論する．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だからである．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 [7] や **SUSAN**[26] と呼ばれる作用素を始め，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 [4, 22, 25]．また，一枚の画像ではなく，連続するビデオフレーム上を移動する特徴点を追跡するアルゴリズムとしてよく知られているものに **Kanade-Lucas-Tomasi** 法がある [28]．しかし，例えば Harris 作用素によって，いま解析している画像中の，いま注目している建物のある頂点の位置を検出したとすると，結果は一通りである (図 1)．何



(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．

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．

よく考えると，「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．なぜ初めから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，それらが本質的に不可能な処理だからである．例えば抽出したい点を「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」と客観的に指定すれば，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，建物の角とか窓の位置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れる．このため，特徴点抽出は本質的にヒューリスティクスである．

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環境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現れ方が大きく変わる．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環境や反射特性など) がすべてわかれば画像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．このため，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．したがって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およびその予測に基づく処理アルゴリズムの集合 と対応する．きる．そして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．したがって，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．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．

4. 特徴点位置の統計的モデル

4.1 特徴点の共分散行列

前節の解釈に基づいて解析を進めるにはいくつかの仮定が必要である．第一は，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は不偏であることである．

特徴点の抽出の性能は画像の性質に依存する．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．したがって，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，濃淡の変化が明瞭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．すなわち，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定性的な性質は定義できてもその絶対的な大きさを定めることはできない．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，およ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．

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ように設計されている．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]．これは正規化共分散行列を単位行列に選んでよいことを意味する．

異なる特徴点はその周りの濃淡変化が互いに関係がないことが多く，その不確定性はほぼ独立とみなせる．しかし，連続した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 [27]．

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．これも「抽出アルゴリズム」の一種である．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およびその周りの濃淡変化は画像処理による結果と類似している．これは一般に画像処理アルゴリズムが人間の視覚認識を模倣していることを考えると当然であろう．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 [18]．

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あるのは「処理アルゴリズム」のアンサンブルであり，正規性，独立性，不偏性，相関などの統計的な概念や仮定は処理アルゴリズムに対して定義される．

しかし，従来からこれが「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」の代わりに単に「画像のノイズ (誤差) モデル」と呼ぶ習慣がある．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った．

もちろん鮮明な画像ほどどのアルゴリズム (「人間」を含めて) 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．そして「抽出したいもの」と「実際に計算していること」が乖離している以上，どんな鮮明な画像でも処理には本質的な不確定性が介入する．

(a) (b)

図 2: (a) 室内シーン．(b) そのエッジ画像

画像から抽出するのは点とは限らず，むしろ歴史的には物体の境界を検出することから研究が始まった．これはエッジ検出と呼ばれる．その目的は「物体の境界」の抽出であるが，実際の処理は画像の濃淡値の変化を抽出している (図 2)．これも本質的にヒューリスティクスであり，無数の研究にもかかわらず未だ決定版はなく，将来もないと予想される．

例えば影の境界は画像の濃淡の境界であるが，最近これを物体の境界でないとして排除する方法が研究されている．これをカラー画像のスペクトルから判定しても，最終的に知りたいのはスペクトルではなく，それが「異なる物体の境界」かどうかである．しかし「異なる物体」の定義は主観的あるいは応用依存である以上 (人物認識では顔と胴は同じ物体であるが，顔認識では異なる物体とみなす)，根本的な解決は存在しないと思われる．

このようにコンピュータビジョンでは画像が鮮明かどうかとは無関係に，あらゆる処理に不確定性が介在する．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．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る．これは画像自体のノイズとは区別しなければならない．またこれは，文字認識の筆者による変動や顔認識の姿勢や照明条件による変動のような「認識対象のアンサンブル」とも区別しなければならない．

5. 漸近解析とは何か

5.1 通常の場合

既に述べたように，統計的推定はアンサンブルの性質を「それから抽出されたサンプルから推定する」ことであるが，推定のために用いるアンサンブルに関する何らかの仮定をモデル (または模型) と呼ぶ．

物理実験のように不確定性が外在的なときは観測条件を制御して精度の向上を図ることができるが，不

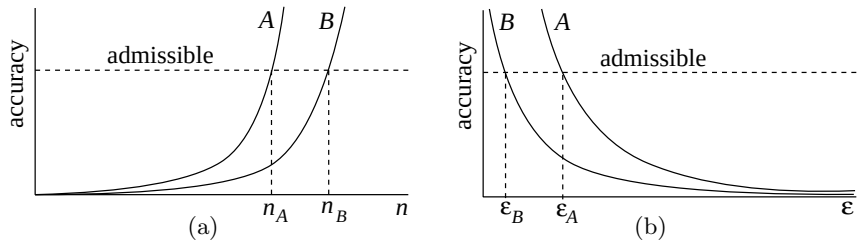


図 3: (a) 通常の統計的推論では実験回数 $n \to \infty$ で急速に精度が向上することが望ましい．なぜなら，より少ない実験回数で許容精度を達成できるからである．(b) 幾何学的推論ではノイズレベル $\epsilon \to 0$ で急速に精度が向上することが望ましい．なぜなら，より大きな不確定性があっても許容精度が達成できるからである．

確定性が対象に内在的なときに精度を上げるには実験を多数回行ない，仮定したモデルに基く統計的推論を行なうしかない．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))．

5.2 幾何学的推定の場合

特徴点に基づく幾何学的推定では，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る．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，およ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．

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．すなわち実験回数は $n = 1$ に限定されている．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

前節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

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

を満たすとき，パラメータ (母数)$\boldsymbol{u}$ を推定すること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 と見なせばよい (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 [17]．また遠方シーンや平面シーンを撮影した 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尤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，最尤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．

(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．

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 $\boldsymbol{u}$ を最尤推定し，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 $\boldsymbol{u}$ にその最尤推定量 $\hat{\boldsymbol{u}}$ 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, 2] 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なうと次の幾何学的 **AIC** が得られる [10]．

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

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**[23, 24] 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離散値に量子化し，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]．

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

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尤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．赤池の AIC では式 (3) ではなく，それを ϵ^2 で割った，すなわち式 (3) の正規化共分散行列 $V_0[p_\alpha]$ を式 (1) の共分散行列 $V[p_\alpha]$ に置き換えたものを用いている (したがって J は対数尤度の -2 倍に等しい)．式 (4) を ϵ^2 で割ると $\hat{J}/\epsilon^2 + 2(Nd + n) + O(\epsilon^2)$ となり，$Nd + n$ は各特徴点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 (対数尤度の -2 倍)+2(未知パラメータの総数) の形となる．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．幾何学的推定ではアンサンブルは仮想的であるから，実際には 1 個のサンプルしか観測できない．しかし仮に毎回異なる位置を繰り返してサンプルできるとすると，n 個のサンプル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，このような仮想的に推定は式 (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し，その分布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たりする試みもある [16]．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サンプルを n 回繰り返したときの $n \to \infty$ の漸近解析に相当しているので，通常の統計的推測における実験回数 $n \to \infty$ の漸近的性質が幾何学的推定において $\epsilon \to 0$ の摂動解析に現れるのである．その結果，n に対する漸近評価 $\cdots + O(1/\sqrt{n^k})$ が ϵ に対する漸近評価 $\cdots + O(\epsilon^k)$ として現れる．

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．

例えば文字認識では文字パタンのアンサンブルを考え，そ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 とは別問題であり，学習の統計的解析は $n \to \infty$ の挙動を対象とする．$N \to \infty$ ではない．

しかし特徴点に基く幾何学的推論ではアンサンブルが仮想的で現実には 1 回しか実験できないことから，1 回の実験データの次元と混同する誤解が生じている．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，文字認識や顔画像認識において一つの文字や顔画像を 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，仮定した幾何学的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を N とするとき，$N \to \infty$ の挙動を議論する研究がある [19]．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．それにもかかわらずこのような議論が行なわれる背景には，統計学において類似の問題が存在するからである．例えば遠方の地面に長い棒状の構造物が横たわっているとする．これに向かってある方向にレーザ光を照射し，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．レーザ光は原理的に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 (一般に方向に依存する) 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 [19]．このような問題はネイマン・スコット問題 [20] と呼ばれ，誤差変量 (errors-in-variables) モデル [5] も類似の構造を持っている．これに対する最適な推定法として，どの方向に照射するかの確率分布の存在を仮定し，実際に照射した方向はそこからのサンプルであるとみなす解析が行なわれる．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る [3]．しかし，各レーザ照射をランダムな試行とすると，一般に試行のたびに誤差の特性が変化することになる．このような病的な状況は heteroscedastic と呼ばれる [15]．

7.3 幾何学的推定とセミパラメトリックモデル

上述の問題は，画像に特徴点抽出処理を行な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．これに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, 21]．また，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 とランダム誤差の標準偏差 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 [19] (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公平な比較ができない．

8. まとめ

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統計学や他分野の研究者とのコミュニケーションが困難

で，それに起因する誤解や論争が生じやすい．本論文では画像の研究者間では暗黙に了解 (時には誤解) されている背景を明示し，今後の研究の発展および他分野とのコミュニケーションの促進を試みた．

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」と「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て考察した．本論文の主な指摘は次の通りである．

- 「統計的方法」とは観測データの性質を調べるのでなく，それらがある「アンサンブル」からランダムに抽出されたサンプルであるとみなし，仮定したモデルに基いてそのアンサンブルの性質を推論することである．
- どういうアンサンブルを考えているかを明確にしなければ，統計的解析を巡って混乱が生じる．
- 特徴点の位置の不確定性は，特徴点抽出アルゴリズムの不完全性を反映したものであり，その不完全性はコンピュータビジョンの宿命であって避けられない．
- 特徴点に基く推論で考えるアンサンブルは「特徴点が抽出され得る可能な位置の集合」である．これは異なる予測に基づく「特徴点の抽出アルゴリズムの集合」と対応する．
- 画像上に N 個の特徴点を抽出すれば，考えているアンサンブルは $2N$ 次元ベクトルの仮想的な値の集合であり，観測できるのはそれからの 1 個のサンプルである．
- このような場合の推定の性能評価はノイズレベル $\epsilon \to 0$ に対する漸近的解析によるのが適切である．
- その漸近評価は仮想的な実験を n 回繰り返すときの $n \to \infty$ の漸近評価と実質的に等価であり，したがって通常の統計解析の $n \to \infty$ の性質が幾何学的推論の $\epsilon \to 0$ に対する性質として現れる．
- 特徴点数 $N \to \infty$ の漸近評価はその意味を十分に考えなければ，シミュレーション実験も手法の評価も無意味なものになる．

謝辞: 本論文の内容は筆者と理化学研究所の甘利俊一博士，フランスの Gérald Battail 博士，米国 Rutgers 大学の Peter Meer 教授，群馬大学の太田直哉助教授，東北大学の岡谷貴之講師，ATR の木下敬介博士との議論の中から生まれたものである．本研究の一部は文部科学省科学研究費基盤研究 C (2) (No. 13680432)，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．

参考文献

[1] H. Akaike, A new look at the statistical model identification, IEEE Trans. Autom. Control, **16**-6 (1977-12), 716–723.

[2] 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, 153 (1976-3), 5–11.

[3] 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, **6**-2 (1996-6), 96–109.

[4] F. Chabat, G. Z. Yang and D. M. Hansell, A corner orientaion detector, **Image Vision Comput.**, **17**-10 (1999-8), 761–769.

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

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

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

[8] 金谷健一, 「空間データの数理—3 次元コンピューティングに向けて—」, 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, 電子情報通信学会論文誌 A, **J84-A**-11 (2001), 1385–1393.

[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, 電子情報通信学会論文誌 D-II, **J85-D-II**-9 (2002), 1497–1499.

[12] K. Kanatani, N. Ohta and Y. Kanazawa, Optimal homography computation with a reliability measure, *IEICE Trans. Inf. & Sys.*, **E83-D**-7 (2000), 1369–1374.

[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?電子情報通信学会論文誌 A, **J85-A**-2 (2002), 231–239.

[14] 栗原祐介, 太田直哉, 撹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, 情報処理学会研究報告, 2002-CVIM-135-14 (2002-11), to appear.

[15] Y. Leedan and P. Meer, Heteroscedastic regression in computer vision: Problems with bilinear constraint, *Int. J. Comput. Vision.*, **37**-2 (2000-6), 127–150.

[16] P. Meer, 私信, 2002.

[17] 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.

[18] D. D. Morris, K. Kanatani and T. Kanade, Gauge fixing for accurate 3D estimation, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, December 2001, Kauai, Hawaii, U.S.A., Vol. 2, pp. 343–350.

[19] 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, 電子情報通信学会論文誌 A, **J83-A**-1 (2000), 83–95.

[20] J. Neyman and E. L. Scott, Consistent estimates based on partially consistent observations, *Econometrica*, **16**-1 (1948-1), 1–32.

[21] 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, 2002 年 7–8 月，名古屋, pp. 335–340.

[22] D. Reisfeld, H. Wolfson and Y. Yeshurun, Contex-free attentional operators: The generalized symmtery transform, *Int. J. Comput. Vision*, **14**-2 (1995-3), 119–130.

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

[24] J. Rissanen, Fisher information and stochastic complexity, *IEEE Trans. Inf. Theory*, **42**-1 (1996-1), 40–47.

[25] C. Schmid, R. Mohr and C. Bauckhage, Evaluation of interest point detectors, **Int J. Comput. Vision**, **37**-2 (2000-6), 151–172.

[26] S. M. Smith and J. M. Brady, SUSAN—A new approach to low level image processing, *Int. J. Comput. Vision*, **23**-1 (1997-5), 45–78.

[27] 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, 情報処理学会研究報告, 2002-CVIM-133-24 (2002-5), 177–184.

[28] 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/`.

[29] 上田修功, ベイズ学習 [I]–[IV], 電子情報通信学会誌, **85**-4 (2002), 265–271; **85**-6 (2002), 421–426; **85**-7 (2002), 504–509. **85**-8 (2002), 633–638.