社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.
CAS2004-95, SIP2004-138, CS2004-231 (2005-3)

[招待講演]　画像からの幾何学的推論とモデル選択

金谷　健一†

† 岡山大学工学部情報工学科

あらまし　モデル選択とはセンサーデータを観測して，それを支配する法則を推論することである．それが既知のとき，その法則に含まれるパラメータ値を最適に推定する方法が統計学でよく研究されている．しかし，モデル選択はそれを超越する問題であり，種々の「原理」が提案されているが，どれが最善だという根拠はなく，その選択はユーザに任されている．本講演の目的は，統計学で最もよく知られている赤池の AIC と Rissanen の MDL の画像処理やコンピュータビジョンへの応用を示すことである．本講演ではまず，赤池の AIC や Rissanen の MDL がそのままの形では幾何学的推論には適用できないことを指摘する．そして，それらの出発原理にさかのぼり，これを幾何学的推論に適用できる形に導出したものが筆者の提起する「幾何学的 AIC」と「幾何学的 MDL」であることを述べる．本講演ではそれらの赤池の AIC や Rissanen の MDL との本質的な違いを直観的な言葉を述べる．そして，実画像による分かりやすい例題を用いて，それらの画像処理やコンピュータビジョンにおける幾何学的推論への応用がいかに広範囲であるかを説明する．

キーワード　幾何学的当てはめ，モデル選択，AIC，MDL，画像処理，コンピュータビジョン

Invited Talk: Geometric Inference from Images and Model Selection

Kenichi KANATANI†

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

Abstract　Model selection is to infer the law that underlies observed sensor data. Once the law is known, the next step is to optimally estimate the unknown parameters contained in the governing equation, which has been well studied by statisticians. In contrast, model selection is in a sense a transcendental problem: Although many principals have been proposed in the past, there is no logical ground to assert that one is better than others, and the choice is left to the user. The purpose of this lecture is to show how Akaike's AIC and Rissanen's MDL, the two selection criteria best known in statistics, can be applied to image processing and computer vision problems. First, it is pointed out that Akaike's AIC and Rissanen's MDL cannot be applied to geometric inference in their original forms. It is shown that if we go back to their starting principles and modify their derivations so that they fit into the framework of image processing and computer vision, we obtain the geometric AIC and the geometric MDL that the author proposed. In this lecture, we describe their essential distinction from Akaike's AIC and Rissanen's MDL in intuitive terms. Using easy-to-understand real image examples, we demonstrate how extensive their applications to geometric inference in image processing and computer vision problems are.

Key words　geometric fitting, model selection, AIC, MDL, image processing, computer vision

1.　モデル選択とは何か

まず**モデル選択**とは何かを例で示そう．平面上に与えられた複数の点に直線，または２次曲線，または３次曲線，またはそれ以上の次数の曲線を当てはめたいとする．何を当てはめたらよいであろうか.

素朴なアイディアは，まず直線，２次曲線，３次曲線，... を順に当てはめ，データ点との食い違い（これを数量的に評価したものを**残差**と呼ぶ）が最も小さいものを選ぶことである．

しかし，これではうまくいかない．なぜなら，高次の曲線ほどデータ点によく当てはまり，十分高い次数の曲線を選べばすべてのデータ点を通るもの（残差が０）が得られるからである（図 1）．これは本来の目的に合わない．なぜなら，曲線を当てはめる目的は背後にある**真の曲線**を少数のしかも誤差のあるデータから推定することだからである．

一般に，誤差のあるデータから真の構造を推定するために導入する未知パラメータをもつ数式を**モデル**と呼ぶ．当然，調節するパラメータの個数（これを**自由度**と呼ぶ）が多いほどデー

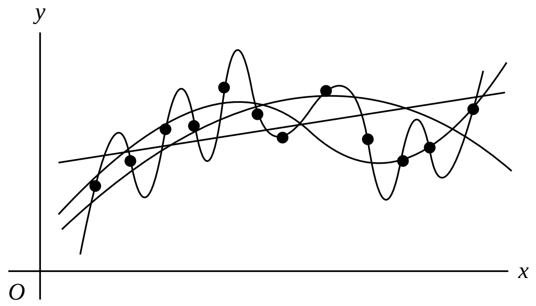


図 1　曲線の当てはめ．次数の高い曲線ほどデータ点によく当てはまる．

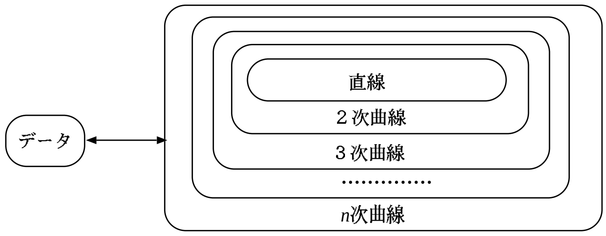


図 2　多項式曲線の包含関係．一般に大きい集合ほどデータに近い．

タによく合致する（すなわち残差が減少する）．そして，自由度が十分大きいほどモデルがデータの誤差によくフィットしてしまう．これを防いで最も適切なモデルを選ぶにはどうすればよいであろうか．これが**モデル選択**の課題である．

2.　モデル選択基準

上記の問題は次のように考えることもできる．曲線当てはめでは，直線は２次曲線の２次の係数が０になる特別の場合である．同様に，２次曲線は３次曲線の特別な場合であり，３次曲線は４次曲線の特別な場合である．すなわち，それぞれの次数の曲線の集合には**包含関係**があり，直線の集合は２次曲線の集合の部分集合であり，２次曲線の集合は３次曲線の集合の部分集合であり，どの次数の曲線の集合もより高い次数の集合の部分集合である．

一方，残差は観測データとモデルとの**距離**に相当している．残差を最小にするモデルとは観測データに最も**近い**モデルにほかならない．しかし，モデルに包含関係があると，最も近いモデルは当然ながら最も大きい集合から選ばれる（図 2）．なぜなら，部分集合に限定すると距離は増えこそすれ減ることはないからである．

このことから，部分集合の元が選ばれるためには距離だけでなく，より自由度の小さいモデルを優先する何らかの評価が必要である．これを測るのが**モデル選択規準**であり，一般に次の形をしている．

$$(\text{残差}) + (\text{自由度に対するペナルティ}) \qquad (1)$$

第１項の残差を減らそうとして高い自由度のモデルを選べば第２項が大きくなる．それに対して，両者の和を最小にするモデルを選べば両者がバランスするものが選ばれる．このようなモデル選択規準として**赤池の AIC** [1], [2]，**Schwarz の BIC** [19]，**Rissanen の MDL** [17], [18], **Mallows の Cp** [14] などいろいろなものが提案されている．しかし，どのモデルが選ばれるかは用いる規準に依存し，絶対的なものはない．

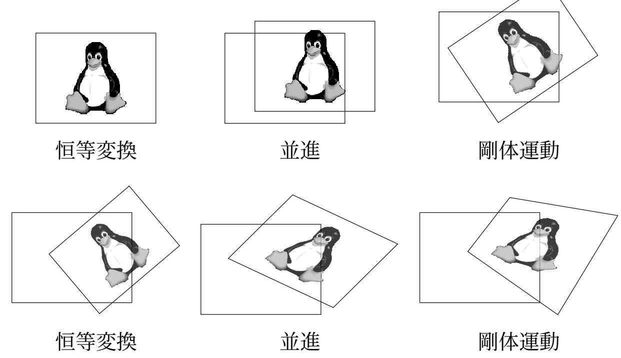


図 3　画像の変換

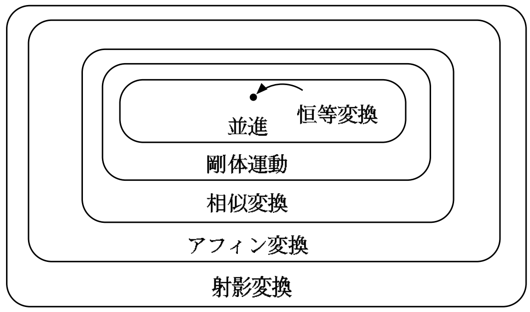


図 4　画像の変換の包含関係

3.　モデル選択の画像処理への応用

以上は統計学で研究されてきたモデル選択の考え方である．そして，これをコンピュータビジョンや画像メディア処理へ応用することができる．しかし，従来からよく認識されていなかった問題点は，統計学におけるモデル選択がそのままの形では画像を用いる幾何学的な問題に適用できないということである．これを説明する前に，まず例を挙げよう．

カメラの向きをいろいろに変えて遠景を撮影した複数の画像を張り合わせて視野に広い画像を作ることを**画像モザイク生成**といい，得られる画像を**パノラマ画像**と呼ぶ [11], [12]．ただし，元の画像のままでは張り合わせることができない．これはシーンの同じ部分を撮影しても，カメラの向きが変わるとその画像が何らかの変換を受けるからである．その変換としては**恒等変換**（まったく同一），**並進**（平行移動），**剛体変換**（回転と平行移動），**相似変換**（回転と平行移動とスケール変化），**アフィン変換**（相似変換に斜めに歪みをを加えたもの）などいろいろなものが考えられる（図 3）．これらはすべて**射影変換**とよばれる変換の部分群である [6]．

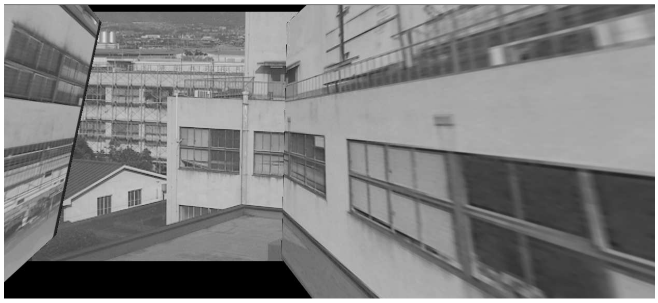
このとき異なる画像を，その共通に写っている部分が最もよく合うように張り合わせるにはどの変換を加えればよいであろうか．ここで注目すべきことは，候補となる変換群が包含関係をもつことである（図 4）．例えば並進は剛体変換の特殊な場合（回転がない）であり，剛体変換は相似変換の特殊な場合（スケール変化がない）である．したがって，最も一般な射影変換

白点が重なるように張り合わせる．

モデル選択なし　　　　　　　　モデル選択あり

図 5　パノラマ画像の生成（その 1）

白点が重なるように張り合わせる．



無限遠点を回り込んだ変換が生じる．

モデル選択による張り合わせ．

図 6　パノラマ画像の生成（その 2）

を用いればすべての場合が含まれ，これを用いればそれ以外を考える必要がないように思える．実際，実用化されているパノラマ画像生成法はほとんどすべて射影変換に基づいている．

しかし，重なり部分が非常に少なく，かつ対応関係にかなりの誤差があるときに問題が生じる．変換が包含関係の外側にいくほど変換のパラメータ数（すなわち自由度）が増え，射影変換は８個のパラメータを持っている．そして，その一つをわずかに変えても変換された画像に大きな変化が生じる．一方，例えば剛体変換は３自由度であり，パラメータが多少変化しても位置と向きが変わるだけで，それ以外の画像の変形は生じない．

実際，図 5 の上段の２画像を図中の白点が一致するように両者を射影変換して張り合わせると，下段左のようになる．確かに白点の位置はほとんどぴったり重なるが，それ以外の部分の重なりが悪い．これに対し，筆者が提案する**幾何学的 AIC** [7] によって最も適切なモデル（この場合は変換）を選ぶと，相似変換が選ばれる．それに従って白点が最もよく一致するような相似変換で張り合わせると，下段右のようになる．全体がよく一致していることがわかる [11]．

図 6 はさらに極端な場合である．上段の図中の白点が一致するように射影変換すると，中段のようになる．これは右画像の中央付近が射影変換によって無限遠に写像され，その先が反対側から現れている．射影変換はこのようにパラメータの選び方によっては無限遠を含む写像となる [6]．これに幾何学的 AIC を適用すると，この場合は並進が選ばれ，白点が最もよく一致する並進は下段のようになる．

4.　幾何学的 AIC とは何か

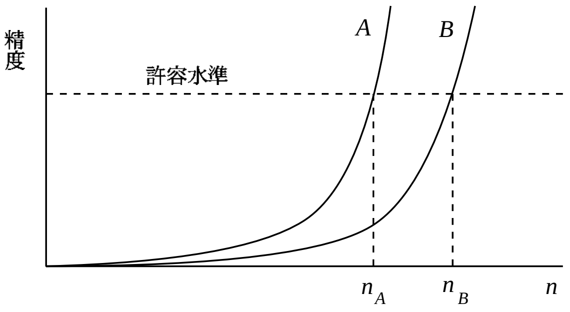
先に述べた幾何学的 AIC とは何であろうか．これは統計学で知られている赤池の AIC（An Information Criterion または Akaike Information Criterion の略）と違うのであろうか．これを完全に説明しようとすると数学的な理論に深入りしてしまうので，ここでは直観的なイメージを描くことに留める．

統計学でいう**モデル**とは観測したデータを説明する数式のことであり，原因となる要因と結果として観測されるデータとの因果関係を記述するものである．それに対して，画像を用いる幾何学的問題ではデータはある高い次元の空間の点集合とみなされ，**モデル**とはその点集合の満たす幾何学的な関係のことである．

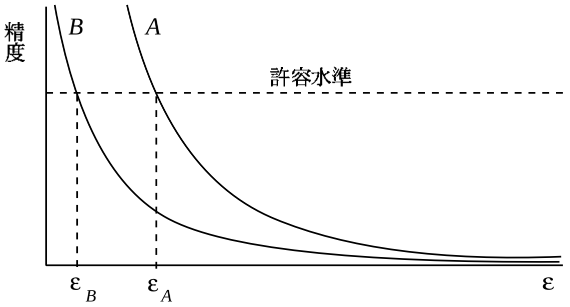
このときモデルの性質としてパラメータ数で表される自由度だけでなく，その**次元**も重要になる．例えば「点集合が同一平面上にある」は一つの幾何学的関係である．そして，平面の自由度は３であり，次元は２である．これを反映して，幾何学的 AIC のペナルティにはモデルの自由度だけでなく，その次元も含まれる．それに対して赤池の AIC ではペナルティは式 (1) のように自由度のみである．また残差の計算法も赤池の AIC と幾何学的 AIC とでは異なる．

さらに本質的な違いは，赤池の AIC が**観測データ数が非常に多い場合の近似式**（これを**漸近評価**という）であるのに対して，幾何学的 AIC は**データの誤差が少ない場合の評価**である点である．

それでは，このように違うものをなぜ同じ **AIC** という名称で呼ぶのであろうか．それは両者の出発点となる原理が共に，**カルバック・ライブラー情報量**（または**ダイバージェンス**）と呼ばれる量（俗に**エントロピー**とも呼ばれる）に基づいているからである．すなわち，原理が同じで，異なるのは適用する問



観測データ数 N が大きいほど急速に精度が向上することが望ましい．



誤差 ϵ が小さいほど急速に精度が向上することが望ましい．

図 7　二通りの漸近評価

題の形式と漸近評価の方法である．

原理が同じでも結果が異なるのは，カルバック・ライブラー情報量は未知の構造に対して定義されるので，それを計算することができないからである．そこで赤池は，観測データが大きいときに成立する**大数の法則**や**中心極限定理**によってそれを推定することによって AIC を導いた．それに対して，幾何学的 AIC は誤差が小さいときはデータは第 1 近似としてその真の値に近いということによってカルバック・ライブラー情報量を近似的に評価するものである（大数の法則や中心極限定理は用いていない）[7].

5.　漸近評価の意味

統計学でなぜ観測データ数の多い場合の漸近評価を問題にするのであろうか．それは，観測データが多いほど因果関係が推定しやすくはなるが，観測データを増やすには多くの実験や調査が必要であり，それにコストがかかるからである．したがって，観測データを増やせばより急速に精度が向上する方法があれば，これは許容精度を達成するのに必要な実験や調査の回数がより少なくて済むので，より望ましいといえる（図 7 上段）．このため，統計学における推定方法の性能は観測データ数の多い場合の漸近評価によって比較される．

一方，画像を用いる幾何学的推論では，通常データ数は初めから固定されており，データに多少の誤差があっても正しい推論ができることが望ましい．したがって，誤差が少なくなるほど精度がより急速に向上する方法は，ある精度を達成するのにより大きい誤差まで許容されるので，より望ましいといえる（図 7 下段）．このことから，幾何学的推論の性能は誤差が小さい場合の漸近評価によって比較するのが妥当である．

6.　幾何学的 MDL

統計学でよく知られたモデル選択基準には赤池の AIC の他に Rissanen の MDL [17], [18]（Minimum Description Length の略）がある．これはデータとモデルを指定するのに必要最低限の符号の列の長さ（これを**記述長**と呼ぶ）を最小とするモデルを採用するという原理に基づいている．

しかし，パラメータ値が実数をとるモデル（例えば実数係数の多項式）を実数データに当てはめる場合，実数を符号化すると無限大の長さになる．実数は 10 進法で表しても，2 進法でも，その他どう表しても一般には無限小数となることから，これは明らかである．そこで Rissanen は実数軸上に離散的な点列をとり，最も近い値で代表させた（これを**量子化**と呼ぶ）．この点列の間隔（**量子化幅**）を大きくとるほどモデルの記述長は短くなるが，その分，モデルの記述の精度が悪くなり，それを用いてデータを符号化すると，その記述長が増えてしまう（情報理論によれば，データはそれを発生させるメカニズムがよく知られているほど短く記述できる）．

そこで Rissanen はその全体の記述長が最小になるように量子化幅を定め，その記述長をそのモデルの善し悪しの評価基準とした（**2 段階符号化**）．しかし，その最小の記述長を厳密に計算するのは複雑過ぎて，実際上不可能である．そこで Rissanen は，データ列の長さが十分大きいときの近似式（漸近評価）を用いて，それを近似的に計算した．これが Rissanen の MDL である．

前節の考え方によれば，赤池の AIC と同様に，Rissanen の MDL もそのままでは画像の幾何学的な推論には適用できない．なぜなら，対象とする問題の形が異なり，また漸近評価の方向が異なるからである．そこで筆者は，画像の幾何学的な推論問題に対して，Rissanen と同じ最小記述長を求めるという方針で，誤差が小さい場合の近似的な評価を行なった．その結果として得られる基準が**幾何学的 MDL** [7] である．

7.　モデル選択基準の比較

いろいろな問題に幾何学的 AIC と幾何学的 MDL の両方を適用して比較してみると，幾何学的な退化に関して差が大きいことがわかった [9]．この**退化**というのは，モデルが特殊な場合になることである．これはある係数が 0 になるなどして自由度が低下する場合と，例えば空間中の点列に平面を当てはめる問題で，その平面が直線になるように次元が低下する場合とがある．

このとき幾何学的 AIC は，真のモデルが退化していなければほとんどの場合に退化していないと判定し，真のモデルが退化していれば大きい確率で退化していると判定するが，真のモデルが退化していてもある確率で退化していないと判定する．それに対して幾何学的 MDL では，真のモデルが退化していればほとんどの場合に退化していると判定し，真のモデルが退化していなけば大きい確率で退化していないと判定するが，真のモデルが退化していなくてもある確率で退化している判定する．

要するに，幾何学的 AIC は十分な証拠がない限り退化していると判定しないのに対して，幾何学的 MDL では退化していな

いものを退化していると判定しがちである．この相違は式 (1) のペナルティの項の違いに起因する．一般に幾何学的 MDL のほうが幾何学的 AIC より複雑な（自由度や次元が大きい）モデルに対するペナルティが大きい．

以上より，結局モデル選択基準として何がよいかという絶対的な基準はなく，応用の目的によって適切なものを選ぶしかないことがわかる．例えば，退化が生じると困り，退化でないものを退化と判定しても特に問題が生じない場合は，より安全側に判定する幾何学的 MDL が適している．一方，破綻が生じるなどのどうしても必要な場合以外は退化の判定をなるべく避けたい場合は幾何学的 AIC が適している．

モデル選択の目的は，モデルの善し悪しを比較する規準を導入することであるが，その基準の善し悪しを比較する規準（モデル選択規準の選択規準）は存在しない．例えば，AIC の出発原理はカルバック・ライブラー情報量であるが，これによってモデルを判定すればよいという根拠はない．同様に，MDL もその出発原理は最小記述長であるが，記述長の短いモデルを用いれば都合がよいという根拠はない．これらの出発原理を比較する規準を導入しても，その規準の善し悪しをどう判定するのかという問題に直面する．結局はどの時点かで，妥当と思われる原理を選ぶしかないと思われる．

8.　モデル選択の応用

筆者はこれまでに幾何学的 AIC や幾何学的 MDL を，画像から抽出したデータの幾何学的構造を判定するさまざまな問題に適用し，その有効性を確認している．例えば，3 節に紹介したパノラマ画像の生成以外に次のような応用がある．

- 画像中の対象の対称性の判定 [3]
- ２画像に３次元形状を復元する十分な視差が含まれているかの判定 [5], [13]
- ステレオ画像中の対象が平面または無限遠方かの判定 [10]
- 折れ線データを一つの線分で近似するか２次曲線で近似するかの判定 [4]
- 動画像中の動きがシーン全体の動きか，あるいは独立に移動する物体が存在があるかの判定 [16]
- 四辺形が正方形や長方形などの規則図形のどれに一番近いかの判定 [20]
- カメラを移動しながら撮影してビデオ画像からのカメラの移動とズーム変化の判定 [15]
- カメラを移動しながら撮影してビデオ画像からの背景と独立に移動する物体を抽出する問題 [8]

これらはほんの一例であり，画像処理やコンピュータビジョンのさまざまな問題の解決に役立つと期待される．ただし，前節に指摘したように，モデル選択を用いれば常に有益な結果が得られるわけではない．モデル選択を活かすには，システム設計者が問題を十分に理解し，それに基づいて適切な規準を選び，適切な方法で適用することが必要である．

9.　ま と め

本講演ではモデル選択規準に関する解釈とその論理の解明，およびその動画像理解への応用という，理論と応用の両面について述べた．そして，画像による幾何学的推論に適したモデル選択規準の考え方を示し，その応用の広がりを述べた．

繰り返しなるが，モデル選択は絶対的なものではなく，また使えば必ず有益であるというものでもない．重要なことは，どういう応用にどういう規準をどのように使えばよいかという判断である．本講演はその判断の基礎を与えることを目的としている．これによりモデル選択の有用な応用範囲がいっそう広がることを期待している．

謝辞． 長年に渡って本研究にいろいろな形で参加頂いた豊橋技術科学大学の金澤靖助教授，岡山大学の菅谷保之助手,（株）朋栄の松永力博士，群馬大学の太田直哉教授に感謝します．

文　　献

[1] H. Akaike, A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, **16**-6, 716–723 (1974).

[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**, 5–11 (1976).

[3] K. Kanatani, Comments on “Symmetry as a Continuous Feature”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**-3, 246–247 (1997).

[4] K. Kanatani, Comments on “Nonparametric Segmentation of Curves into Various Representations”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**-12, 1391–1392 (1997).

[5] 金谷健一, 自己評価を伴うアクティブビジョン, 日本ロボット学会誌, **15**-2, 268–274 (1997).

[6] 金谷健一,「形状ＣＡＤと図形の数学」, 共立出版 (1998).

[7] 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, 1385–1393 (2001).

[8] K. Kanatani, Motion segmentation by subspace separation: Model selection and reliability evaluation, *International Journal of Image and Graphics*, **2**-2, 179–197 (2002).

[9] 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, **J85-D-II**-9, 1497–1499 (2002).

[10] Y. Kanazawa and K. Kanatani, Infinity and planarity test for stereo vision, *IEICE Transactions on Information and Systems*, **E80-D**-8, 774–779 (1997).

[11] 金澤 靖, 金谷 健一, 幾何学的 AIC による画像モザイク生成の安定化, 電子情報通信学会論文誌, **J83-A**-6, 686–693 (2000).

[12] 金澤靖, 金谷健一, パノラマ画像の作り方—イメージモザイキングのための射影変換—, 電子通信情報学会誌, **83**-12, 944–946 (2000).

[13] 金澤 靖, 金谷 健一, 大域的な整合性を保証するロバストな画像の対応づけ, 情報処理学会論文誌: コンピュータビジョンとイメージメディア, **44**-SIG 17, 70–77 (2003).

[14] C. L. Mallows, Somme comments on C_p, *Technometrics*, **15**-4, 661–675 (1973).

[15] 松永 力, 金谷 健一, 平面パタンを用いる移動カメラの校正：最適計算，信頼性評価，および幾何学的 AIC による安定化, 電子情報通信学会論文誌 A, **J83-A**-6, 694–701 (2000).

[16] N. Ohta and K. Kanatani, Moving object detection from optical flow without empirical thresholds, *IEICE Transactions on Information and Systems*, **E81-D**-2, 243–245 (1998).

[17] J. Rissanen, Universal coding, information, prediction and estimation, *IEEE Trans. Inform. Theory*, **30**-4, 629–636 (1984).

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

[19] G. Schwarz, Estimating the dimension of a model, *The Annals of Statistics*, **6**-2, 461–464 (1978).

[20] I. Triono, N. Ohta and K. Kanatani, Automatic recognition of regular figures by geometric AIC, *IEICE Transactions on Information and Systems*, **E81-D**-2, 246–248 (1998).