

解説



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†

金谷 健一†

1. はじめに

ロボットが原子炉内のような危険個所で自律的に作業するにはロボットに搭載したビデオカメラや距離センサなどによって環境を計測し、作業対象物や障害物の3次元記述を得る必要がある。このとき、画像や距離センサから得られるデータには本質的に不確定さがあるので、3次元形状の信頼性を最大にする最適化手法が必要となる。しかし、最適に計算しただけでは不十分である。たとえば、最適に推定した結果、対象物の位置が5m先にあるといっても、誤差が±10cm程度か、それとも±1m程度かによって対応が異なる。たとえば前者なら4.5m程度進んでから詳しい測定をし、後者なら別のセンサで計測し直すことになる。

このような信頼性評価は画像や距離データの誤差に依存する。当然、誤差が大きいほど信頼性が低下し、小さいほど向上する。したがって信頼性評価のためにはデータの誤差の大きさを知っていなければならないように思える。誤差が評価できれば、それを数値計算の丸め誤差の区間解析のように計算過程に沿って波及させて最終出力の信頼性が評価できる。しかし誤差はビデオカメラやレンジファインダなどの測定機器の特性だけでなく、照明などの測定条件や信号の処理アルゴリズム、さらにはその測定対象自身の位置や形状にも依存するので、これを予測することは困難である。ところが実は誤差の大きさを知らずに信頼性が評価できる。

なぜそのようなことが可能かといえば、その答は環境に対する事前の知識にある。たとえば対象

が平面であることが分かっていると、その上の複数の点の測定位置は誤差がなければ同一平面上にあるはずである。このような「誤差がなければ成立しているべき条件」を拘束条件と呼ぶ。画像理解の原理としてよく知られているように、拘束条件は画像に写っている内容を幾何学的に解釈するための基礎となるものであるが¹⁾、「測定値の拘束条件からのずれの程度」をデータの誤差の大きさの目安とすれば、誤差の大きさをあらかじめ知らずにデータから幾何学的な解釈を計算する過程でその信頼性も同時に評価できる。本稿では主として筆者自身の最近の研究を基にして、このことを数学的な詳細に立ち入らずに解説する。

2. 直線とコニックの当てはめ

ロボットが作業する室内環境では床、壁、天井、机など平面部分が多く、その境界には直線部分が多い。したがって画像処理によって画像上の濃淡の変化が著しい部分を画素の列(エッジ)として取り出し、ほぼ直線状のエッジに直線を当てはめることによって物体境界が指定できる。図-1の左は画像処理によって検出したエッジ画像の一部である。その1つのエッジに最適に直線を当てはめて元の濃淡画像の上に重ねたものが図-1の

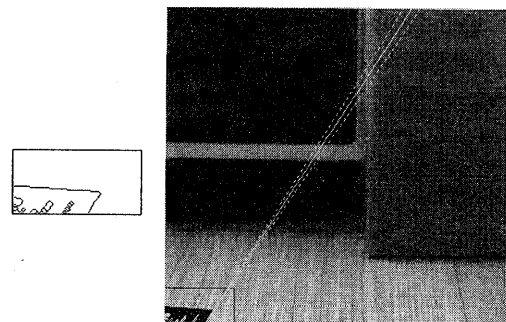


図-1 エッジ画像の一部(左)と直線当てはめの信頼性の表示(右)

† Statistics for Image Understanding: Reliability Evaluation of Geometric Interpretation from Images by Kenichi KANATANI (Department of Computer Science, Gunma University).

† 群馬大学工学部情報工学科

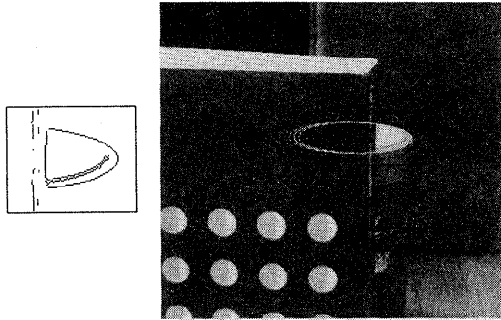


図-2 エッジ画像の一部（左）とコニック当てはめの信頼性の表示（右）

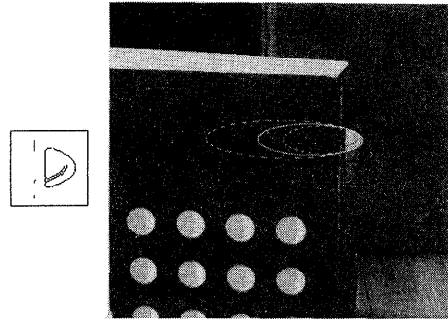


図-3 エッジ画像の一部（左）とコニック当てはめの信頼性の表示（右）

右の白い実線であり、白い破線はこの当てはめの信頼性を表すものである。その意味は次のとおりである。

平面上の直線は2個のパラメータで指定できる。したがって直線は2次元パラメータ空間の1点として指定できる。これが確率的に変動すれば、その2次元空間上に確率分布を定義する。これを真の位置が座標原点となるように変数変換して2次元正規分布で近似し、この空間を変動が最も大きい方向の1次元空間とその直交補空間とに直和分解する。そして分布を前者の空間上の1次元正規分布で近似し、その標準偏差に相当するパラメータの定義する直線を図示したものが図-1の右の破線である。以下、このような信頼性を表す図形の対を標準変位と呼ぶ。1次元分布で近似してよいことは、その方向の標準偏差に比べて、それと直交する方向の標準偏差が無視できるほど小さいことによる。これは理論解析によって証明できる。

この手法は多変量解析やパタン認識と同じであるが、多変量解析やパタン認識と根本的に異なるのは、パラメータ空間には1サンプル点しか得られないことである。したがってサンプル分散・共分散を計算することができない。それに代わるものとして統計学で攪乱母数を持つ不偏推定のクラメル・ラオの下界³⁾に相当するものを用いる。当てはめに最尤推定を用いれば、画像の誤差の高次の項を除いてその下界が達成できることが証明される⁷⁾。

この計算では各エッジ点の位置の誤差の大きさについては未知としているが、その代わりに「エッジは誤差がなければ同一直線上にある」という拘束条件を用いている。この計算は誤差を下から

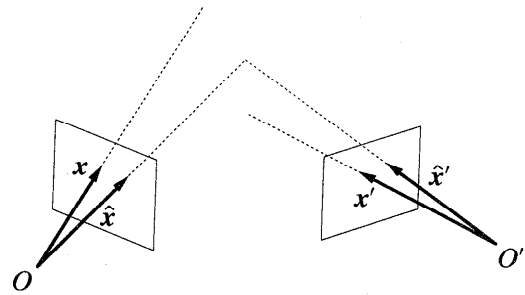


図-4 視線が交わるように対応点を補正する

押さえる、すなわち下界を求めることに相当する。したがって計算した標準変位は「このエッジの真の位置には本質的にこの程度の不確実性があり、これ以上良い当てはめはできない」ということを示している。

図-1から想像できるように、直線状のエッジに当てはめた直線のエッジの部分での不確実性はごく小さい。しかし、画像理解の1つの手法としてよく知られているように、空間で平行な直線の投影は画像上で延長すると消失点と呼ばれる1点で交わり、その位置を知るとそれらの直線の空間での方向が計算できる¹⁾。そのような応用ではエッジに当てはめた直線の延長線の信頼性を評価する必要がある。たとえばカメラの焦点距離やその位置のキャリブレーションにこれを応用すれば、計算した焦点距離や位置の信頼性が定量的に評価できるし、その信頼性を最大にする方式を設計することもできる²⁾。

画像上で直線に次いで重要な特徴はコニック（円錐曲線）である。これは楕円、放物線、双曲線の総称であるが、実際に重要なのは楕円である。その理由は人工物には球形あるいは円形の部分が多く、それらが画像上に楕円として現れるた

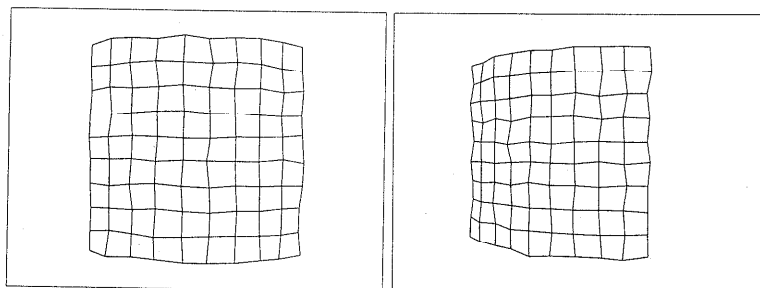


図-5 ステレオ画像のシミュレーション

めである。特に原子炉などではメータやバルブやハンドルのような円形物体が多いので、これらを識別することが重要となる。また、画像理解で知られているように、画像上に抽出したコニックから、「それが円の投影像である」という知識を用いてその円の空間中の位置を計算することもできる⁷⁾。

コニックに対しても直線の当てはめと同様な解析ができる。当然、楕円状のエッジ全体が得られれば信頼性の高い当てはめができるが、実際の画像では部分的なエッジしか得られないことが多い。図-2、図-3の左は画像処理によって抽出したエッジ画像の一部である。その楕円部分のエッジに最適に当てはめた楕円とその標準変位をもとの濃淡画像に重ねたものが図-2、図-3の右である。これから分かるように、エッジ部分が短くなると信頼性が急速に低下する。

ここではくりこみ法⁸⁾⁹⁾と呼ぶ方法で最適な当てはめを行っている。これは最尤推定を計算する非線形最適化問題を計算が簡単な最小二乗法に置き換えて解き、この置き換えによって生じる解の統計的な偏差を理論的に解析し、誤差の大きさを推定しながら反復によってその偏差を除去する方法である。これによって誤差の高次の項を除いて最尤推定と同程度の精度が達成できることが示される⁷⁾。

3. ステレオ視による3次元復元

ステレオ視（立体視）とは2台のカメラで撮影した画像から三角測量の原理で3次元形状を計算する方法であり、移動ロボットや自律走行車などのセンサとして最も基本的な手段の1つである。このとき画像処理によって第1画像のどの部分が第2画像のどの部分に対応するかを検出するステ



図-6 図-5から復元した形状と各格子点の信頼性

レオマッチングが必要となり、その探索の方針や順序がいろいろ考えられる。たとえば第1画像の局所的な画像を切り出しては第2画像上に重ねながら走査し、何らかの基準で一致の程度が最大になる場所を選ぶとする。こうして得られた対応点の視線の方向をカメラモデルから計算し、その交点を求めればその点の空間中の位置が計算できるはずである。

しかし実際には誤差のために2つの視線が1点で交わるとは限らない。2つの画像上の対応点の画像座標が与えられたとき、それらの定義する視線が1点で交わるための必要十分条件を式で表したものはエピ極線方程式と呼ばれているが、画像処理で抽出した対応点はこれを満たしているとは限らない。そこでこのエピ極線方程式を拘束条件とすれば、直線やコニックと同様な信頼性解析ができる⁹⁾。

具体的には対応点をその定義する視線が交わるように最適に補正し、その補正量から画像の誤差の大きさ（の下限）を推定する（図-4）。この補正は、「最適」の基準として最尤推定を用いれば、誤差が正規分布に従うときに、パラメータ空間で分布の共分散行列の一般逆行列を計量とする距離

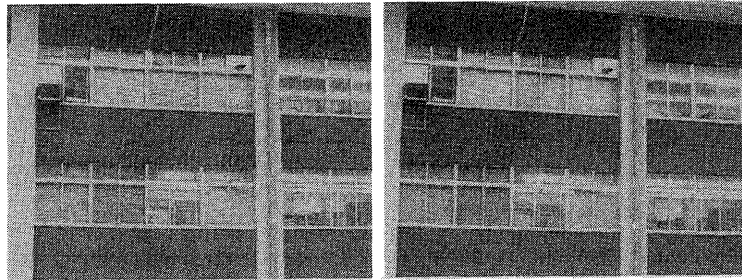


図-7 ステレオ実画像

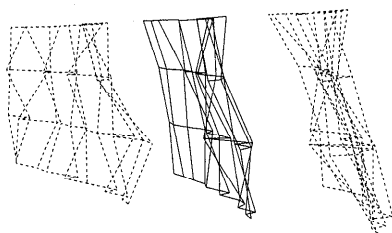


図-8 図-7から復元した形状とその信頼性

(マハラノビス距離)による最短経路を計算することに相当する⁵⁾。このような補正を単に**最適補正**と呼ぶ。

図-5は実際のカメラによる撮影に似たスケールで乱数誤差を加えて作成したステレオ画像のシミュレーションである。図-6は復元した3次元形状を横から見たものであり、破線は真の位置を示している。各格子点の周りの楕円体はその真の位置の信頼範囲を表すものであり、誤差の分布を正規分布で近似したときの各方向の標準偏差を表すものである。これを**標準信頼領域**と呼ぶ。図から分かるように、奥行き方向の誤差がそれに直交する方向の誤差に比べてきわめて大きい。計算では画像に加えた誤差の大きさに関する情報は用いず、エピ極線方程式のみを用いているので、エピ極線方程式をたまたま満たす対応点には誤差がないと判定される。したがって、図-6は「避けることのできない誤差の下界」を表している。

図-7は建物の壁面をとったステレオ画像である。左右の画像の窓枠から抽出した格子を特徴として復元した3次元形状を図-8に示す。破線は図-6に示したような各格子点の標準信頼領域を包絡する格子である。移動ロボットや自律走行車では従来は物体の位置を最適に推定することのみに主眼が置かれていたが、それだけでは不十分であり、推定した位置の信頼性評価が欠かせないこ

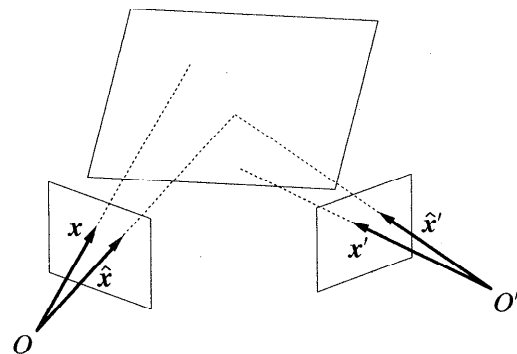


図-9 視線が平面上で交わるように対応点を補正する

とがこの例からも分かる。

この例で奥行き方向の信頼性が低いのは、まず第1に建物までの距離に比べて2台のカメラを結ぶ線分(基線)の長さが短いためである(約1/16)。さらに2台のカメラの位置関係と各カメラの焦点距離のキャリブレーションの精度が十分でないためでもある。拘束条件として用いるエピ極線方程式には2台のカメラの位置関係と各カメラの焦点距離が含まれているが、**拘束条件が満たされない要因を誤差と定義している**ので、キャリブレーション誤差まで取り込んだ評価となっているのである。

この例では特徴格子の各点を別々に3次元復元しており、真の形状に関しては何の知識も仮定していない。しかし「図-7は平面物体の画像である」という知識があれば、それを拘束条件としてより厳しい信頼性評価を行うことが可能となる⁶⁾。このことは図-4で、対応点の定義する視線が単に個々に交わるだけでなく「すべての対応点に対して共通の平面上で交わる」ように補正することに相当する(図-9)。この条件からその共通の平面をくりこみ法によって最適に推定することができる⁶⁾。

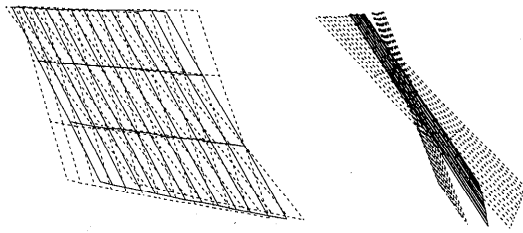


図-10 平面性の拘束条件を用いて図-7から復元した形状とその信頼性

図-10は図-7から平面性の拘束条件を用いて復元した形状とその信頼性を表示したものであり、異なった角度から眺めたものを2つ示す。実線は復元した格子であり、破線は標準変位である。これは直線当てはめと同様に、3自由度を持つ平面のパラメータ空間から最も変動が生じやすい1次元部分空間を抽出して計算したものである。図-8と比較すると、平面であるという知識によって信頼性が向上していることが分かる。

4. 動画像の3次元解析

ステレオ視では2台のカメラの位置関係が既知でなければならない。しかし多数の対応点が得られれば、カメラの位置関係と各点の3次元位置の両方を(スケールを除いて)計算によって求めることができる。したがって、カメラを自由に移動しながら撮影した画像からカメラの運動と対象物体の3次元形状の両方が計算できる。これは画像理解の基本的な方法論の1つであり、**動画像の3次元解析**と呼ばれる¹⁾。

基本的な拘束条件はステレオ視の場合と同じ「エピ極線方程式」である。ただ、ステレオ視では2台のカメラの位置関係を指定するパラメータ(運動パラメータ)が既知であったのに対し、今度はそれらもまた未知パラメータと見なす。この問題に対しても誤差の大きさを未知数として最尤推定に相当する最適解がくりこみ法によって計算できるとともに、その信頼性も評価できる。

ただし、くりこみ法は未知パラメータに関して線形な拘束条件に対する計算法であり、エピ極線方程式は運動パラメータに関して線形ではないので直接に適用することができない。一方、運動パラメータを適当に組み合わせたものを新しい変数とすると、それらについては線形な式となる。それらを要素とする行列は**基本行列**と呼ばれてい

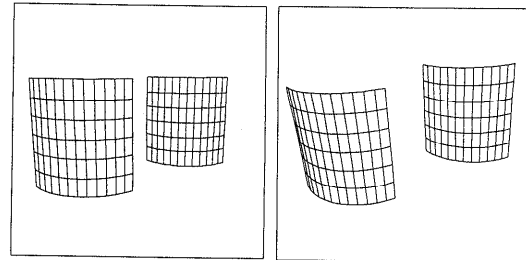


図-11 動画像のシミュレーション

る¹⁾。したがって基本行列を未知数とすれば、くりこみ法によって最適な推定値が計算できる。このような手法を**線形化手法**と呼ぶ。

しかし線形化手法で計算した基本行列は、それからもとの運動パラメータが定まるための必要十分条件(分解可能条件)を満たしているとは限らない。したがって分解可能条件を拘束条件とし、これを満たすように基本行列を補正してからもとの運動パラメータに分解する。同時に計算した運動パラメータの信頼性も評価する。これを最適に行うための誤差の挙動に関する情報は基本行列を計算するくりこみ法の過程から得られる。このようにして運動パラメータが得られれば、以下ステレオ視の場合と同様にして3次元復元の信頼性が計算できる¹⁾。

図-11は実際のカメラによる撮影に似たスケールで作成した動画像のシミュレーションである。この各格子点に独立に乱数誤差を加え、カメラの並進と回転を計算する。並進も回転もそれぞれ3個のパラメータで指定できるから、解はそれぞれ3次元パラメータ空間の1点となる。乱数誤差をいろいろ変えてこれを繰り返す、3次元空間にプロットしたのが図-12、図-13、図-14であり、左が並進、右が回転である。図-12はくりこみ法を用いないで(エピ極線方程式の二乗和を最小にする)最小二乗法を用いたものであり、図-13はくりこみ法を用いたものである。両図とも基本行列の補正は行っていない。図-14はくりこみ法と基本行列の最適補正を併用した結果である。見やすくするために真の値を中心とする箱を表示し、真の値の周りにクラメル・ラオの下界から導かれる標準信頼領域を図示している。図-12、図-13、図-14を通して対応する箱および楕円体の実寸は同一である。

図から分かるように、最小二乗解には統計的な

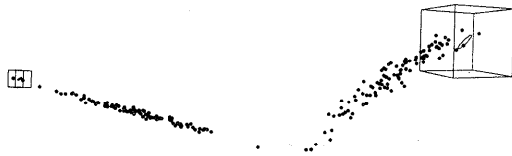


図-12 最小二乗法による並進（左）と回転（右）の誤差

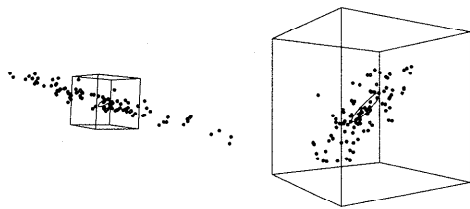


図-13 くりこみ法による並進（左）と回転（右）の誤差

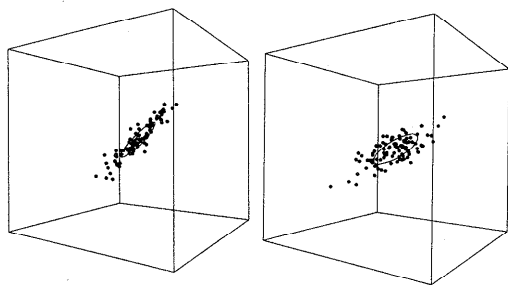


図-14 くりこみ法と基本行列の最適補正による並進（左）と回転（右）の誤差

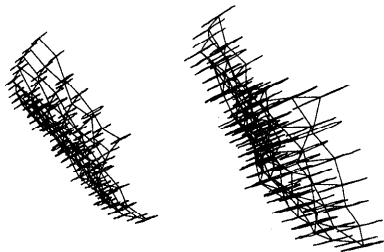


図-15 復元した3次元形状と各格子点の標準信頼領域

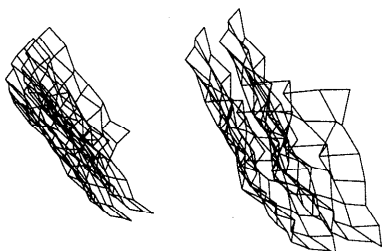


図-16 復元した3次元形状の信頼性

偏差が存在し、くりこみ法によってそれが取り除かれ、さらに基本行列の最適補正によって精度が向上している。さらに図-14 から分かるように、この結果は理論的な下界をほぼ達成しており、これ以上の改良は不可能である。従来、線形化手法は精度が低いとして排除されることが多かったが、分解可能条件を拘束条件とする最適補正と併用すれば精度の理論限界に到達できることが分かる。

図-15 は復元した3次元形状を横から見た一例であり、各格子点の周りにその標準信頼領域を図示したものであるが、ステレオ視の場合と同様に奥行き方向に細長い。見やすくするために各標準信頼領域を包絡する2つの格子を図示したのが図-16である。このように単に形状を最適に3次元復元できるだけでなく、その信頼性まで定量的に評価できる。しかもその計算には画像の誤差の大きさに関する情報は用いていない。

5. オプティカルフローの3次元解析

前章に示した解析を実画像に対して行おうとすると、第1画像のどの点が第2画像のどの点に対応するかを知らなければならない。これはステレオ視の場合と同じであるが、ステレオ視では運動パラメータが既知であるため、第1画像の各点に対して第2画像上でエピ極線方程式を満たす点の集合を計算し、その近傍のみを探索すればよい。そのような点集合はエピ極線と呼ばれる直線となり、すべてのエピ極線は延長するとエピ極点と呼ばれる1点で交わる¹⁾。カメラを左右に水平に平行移動した配置ではエピ極線は（“無限遠点”で交わる）水平線となり、「対応する点は同じ高さにある」ことを意味する。しかし動画画像解析では運動パラメータが未知であるから対応点探索の手掛かりが少なく、いろいろな工夫が考えられているがなかなか困難な問題である。

しかしカメラの移動が微小であれば対応点はもとの点の近傍にあり、画像処理による探索が容易である。このようにして得られた対応点間を結んだものは画像の各点の“速度場”と見なせる。これをオプティカルフローと呼ぶ。ただし探索処理は容易でも、ある点の近傍にその点と同じ濃淡値を持つ領域があればその中のどの点に進んだのか分からないというアパチャー問題と呼ばれる不定

性があり、信頼性は一般に低い。またステレオ視による3次元復元の信頼性は基線長が短いほど低いことから分かるように、オプティカルフローからの3次元復元はそれほど高い精度が期待できない。しかしオプティカルフローは画像全体に密に得られるので、多数の画素に対する最適化を行えば精度が向上すると期待できる。そのためには検出したフローの精度を適切に評価し、それに応じた重みを適切に選んだ最適手法が必要となる。

解析の基本は有限運動の場合と同じであり、「エピ極線方程式」が拘束条件となる。ただしこの場合は瞬間的な並進速度と回転速度が運動パラメータとなり、有限運動の場合の瞬間運動の極限をとって得られる。この場合も運動パラメータを組み合わせたフロー行列を未知数とする線形化手法を用い、くりこみ法でこれを計算した後、分解可能条件を満たすようにフロー行列を最適に補正してから運動パラメータを計算する¹¹⁾。

図-17は実際のカメラによる撮影に似たスケールで作成したオプティカルフローのシミュレーションである。このフローに独立に乱数誤差を加え、カメラの並進速度と回転速度を計算する。それぞれ3個のパラメータで指定できるから、解は

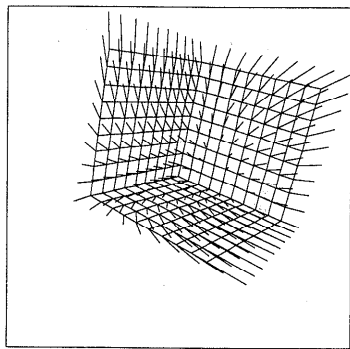


図-17 オプティカルフローのシミュレーション

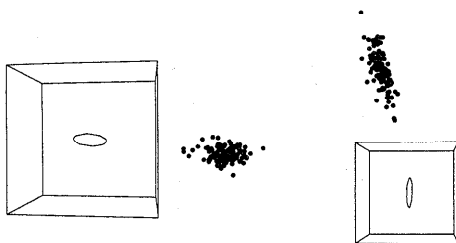


図-18 最小二乗法による並進速度(左)と回転速度(右)の誤差

それぞれ3次元パラメータ空間の1点となる。乱数誤差をいろいろ変えてこれを繰り返し、3次元空間にプロットしたのが図-18、図-19、図-20であり、左が並進速度、右が回転速度である。図-18はくりこみ法を用いないで(エピ極線方程式の二乗和を最小にする)最小二乗法を用いたものであり、図-19はくりこみ法を用いたものである。両図ともフロー行列の補正は行っていない。図-20はくりこみ法とフロー行列の最適補正を併用した結果である。有限運動の場合と同様に真の値を中心とする箱を表示し、真の値の周りにクラメル・ラオの下界から導かれる標準信頼領域を図示している。図-18、図-19、図-20を通して対応する箱および楕円体の実寸は同一である。

有限運動の場合と同様に、最小二乗解には統計的な偏差が存在し、くりこみ法によってそれが取り除かれ、さらにフロー行列の最適補正によって精度が向上していることが分かる。図-20から分かるようにこの結果は理論的な下界をほぼ達成しており、これ以上の改良は不可能である。図-21は復元した3次元形状を横から見た一例であり、各格子点の周りにその標準信頼領域(この場合は線分に縮退する)を図示している。見やすくするために各標準信頼領域を包絡する2つの格子を図示したのが図-22である。

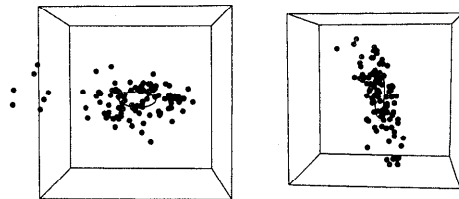


図-19 くりこみ法による並進速度(左)と回転速度(右)の誤差

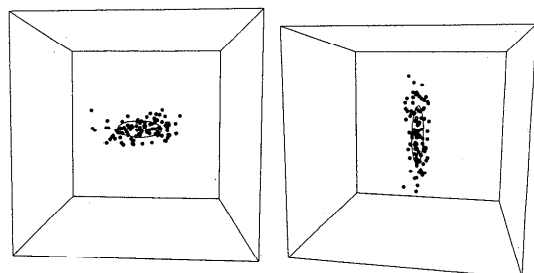


図-20 くりこみ法とフロー行列の最適補正による並進速度(左)と回転速度(右)の誤差

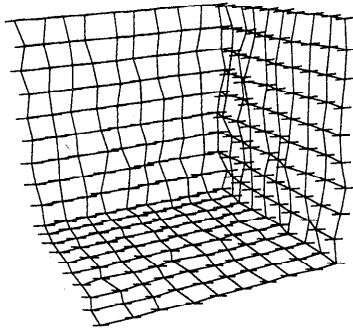


図-21 復元した3次元形状と各格子点の標準信頼領域

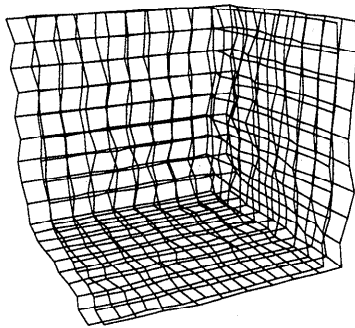


図-22 復元した3次元形状の信頼性

6. おわりに

本稿ではロボットが画像や距離センサによって得たデータから3次元環境モデルの構築を行う場合に、幾何学的解釈の基礎となる「拘束条件」を利用し、誤差の大きさをあらかじめ知ることなく最適解を求め、その過程でその信頼性も同時に評価できることを、主として筆者自身の最近の研究を基にして解説した。ここには紹介しなかったが、オブティカルフローに対しても「勾配方程式」と呼ばれる拘束条件に基づいた統計的な最適な検出法を構成し、得られたフローの信頼性を評価することもできる⁷⁾。また、画像だけでなく距離センサから得られる距離データに対しても同様な解析ができる¹⁰⁾。

この理論の発展として興味深いのは、拘束条件が事前の知識によっては特定できず、複数の可能性がある場合である。このときは、モデルの「良さ」をその予測能力で評価する幾何学的情報量基準(AIC)によって、拘束条件の選択基準が得られる⁷⁾。たとえば、点列を直線、あるいはコニッ

クと見なしてよいか、物体を平面と見なしてよいか、動画像系列が十分な3次元情報を含んでいるか、などが誤差の大きさに関する情報なしに判定できる。このとき、人為的に設定するようないしきい値は何も用いない。

本稿で例として採りあげた問題は従来それぞれ画像やロボットの研究者が独立に研究してきた。しかし、見かけは異なっても抽象化すれば数学的な構造は同一である。今後の発展のためには、「抽象化すれば同一の数学的解析が行える問題」は体系化した学問として整備することが必要である⁹⁾。このような仕事は工学者自身が行わなければならない。なぜなら、抽象化して現れる構造は数学としては目新しいものではなく、数学者が改めて研究する性格のものではないからである。本稿で解説したものは“新しい数学理論”ではなく、“既存の数学に新しい工学的観点を与えた”ものである。

謝辞 本稿の基となった筆者の研究に対する有益なコメントをいただいた東京大学の甘利俊一教授、杉原厚吉教授、出口光一郎助教授、群馬大学の太田直哉助手、群馬工業高専の金澤靖助手および実験に加わった筆者の研究室の学生諸君に感謝します。

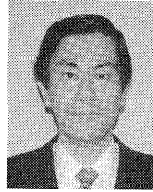
参考文献

- 1) Kanatani, K.: Geometric Computation for Machine Vision, Oxford University Press, Oxford (1993).
- 2) 金谷健一: 画像の3次元解釈の統計的信頼性, 情報処理学会論文誌, Vol. 34, No. 10, pp. 2062-2070 (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, Vol. 35, No. 2, pp. 201-209 (1993).
- 4) Kanatani, K.: Renormalization for Motion Analysis: Statistically Optimal Algorithm, IEICE Trans. Inf. & Syst., Vol. E77-D, No. 11, pp. 1233-1239 (1994).
- 5) 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, 東京 (1995).
- 6) 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).
- 7) Kanatani, K.: Statistical Optimization for Geometric Computation: Theory and Practice, Elsevier Science, Amsterdam (1996), to appear.
- 8) Kanazawa, Y. and Kanatani, K.: Direct Reconstruction of Planar Surfaces by Stereo Vision, IEICE Trans. Inf. & Syst., Vol. E78-D,

No. 7, pp. 917-922 (1995).

- 9) Kanazawa, Y. and Kanatani, K.: Reliability of 3-D Reconstruction by Stereo Vision, IEICE Trans. Inf. & Syst., Vol. E78-D, No. 10, pp. 1301-1306 (1995).
- 10) Kanazawa, Y. and Kanatani, K.: Reliability of Fitting a Plane to Range Data, IEICE Trans. Inf. & Syst., Vol. E78-D, No. 12 (1995), to appear.
- 11) Ohta, N. and Kanatani, K.: Optimal Structure from Motion Algorithm for Optical Flow, IEICE Trans. Inf. & Syst., Vol. E78-D, No. 12 (1995), to appear.

(平成7年2月24日受付)



金谷 健一 (正会員)

1947年生。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。同年群馬大学工学部情報工学科助手。現在同教授。著書「線形代数」(1987)(講談社),「Group-Theoretical Methods in Image Understanding」(1990)(Springer),「画像理解」(1990)(森北出版),「Geometric Computation for Machine Vision」(1993)(Oxford University Press),「空間データの数理」(1995)(朝倉書店),「Statistical Optimization for Geometric Computation」(1996)(Elsevier Science)。1987年本会論文賞。



