カメラキャリブレーションの数理

——最良格子パタンはどのようにして導かれたか——

その2　画像処理と統計的信頼性

金谷　健一

1　はじめに

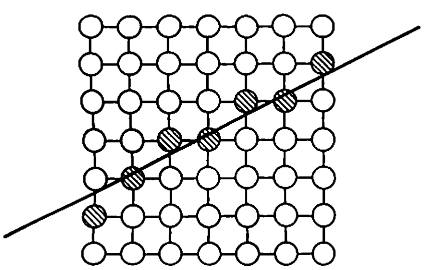
前回はカメラによる撮像の原理，画像からの3次元解釈における焦点距離の果たす役割，消失点の幾何学的な意味，および焦点距離キャリブレーションの原理について述べた．そこでも述べたように，ビデオカメラから入力された画像は，コンピュータの中では**ディジタル画像**として表現される．これは格子状に配置された多数の画素（例えば512×512個）からなり，コンピュータの内部では各画素に対するメモリがそれぞれあって，モノクロ画像ではその点での濃淡値，カラー画像では赤(R)，緑(G)，青(B)の三原色の各成分の強度が数値データとして記憶されている．

前回は格子パタンを用いるキャリブレーションについて述べたが，画像を用いる限り，ディジタル画像の処理による誤差が避けられない．今回は画像処理から求まる消失点の信頼性を考慮すれば，パタンの最適な配置が理論的に定まることを示す．

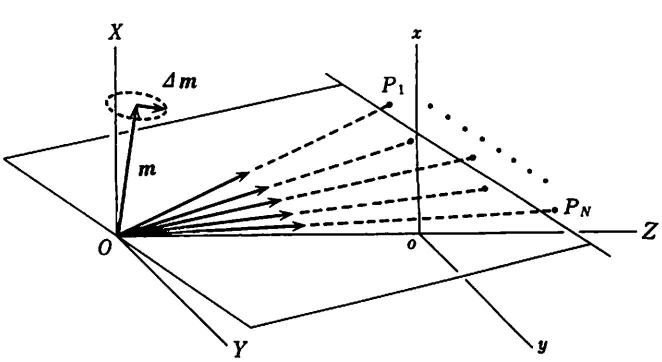
2　エッジへの直線当てはめ

画像を解析して3次元環境を理解することをめざすコンピュータビジョンの基本となる画像処理の一つは，濃淡画像を入力し，2値化，エッジ検出，細線化などの処理によって直線や曲線のみからなる**エッジ画像**に変換することである．**エッジ**とは画像上の線や曲線を表す一続きの画素の列のことであり，コンピュータのメモリ内部では**チェーン符号**と呼ばれるデータとして記憶されているのが普通である．これは始点の画素の番号と，それに続く画素を順々にたどる経路を符号化したものである．

濃淡画像をエッジ画像に変換するいろいろなアルゴリズムが開発されているが，エッジが離散的な画素からなっていることに加えて，照明やピントが不完全で画像が完全に正確でないことのため，直線部分を表すエッジでも完全に一直線上にあるわけではない（図1(a)）．そこで，正しい直線の位置を推定するために，



(a)　エッジに直線を当てはめる．



(b)　直線当てはめの原理

図1

エッジに直線を当てはめる．その原理をわかりやすく説明するために，前回説明した透視変換のモデルを考える．ただし，このモデルは直線当てはめのための便宜的なものであり，カメラモデルに一致している必要はない．したがって，焦点距離 f も適当な値を仮定してよい．

視点 O から各画素に向かう単位ベクトルを $\boldsymbol{m}_\alpha$, $\alpha=1,\cdots,N$ とする．これを各画素の**Nベクトル**と呼ぶ．求めたいものは当てはめた直線 l であるが，直線 l の方程式を直接に計算するのではなく，直線 l と視点 O とで定まる平面の単位法線ベクトル $\boldsymbol{n}$ を求めてもよい(図1(b))．$\boldsymbol{n}$ を直線 l の**Nベクトル**と呼ぶ．図1(b)からわかるように，直線 l が正しく当てはまっていれば $(\boldsymbol{m}_\alpha, \boldsymbol{n})=0$, $\alpha=1,\cdots,N$ である（$(\cdot,\cdot)$ はベクトルの内積）．しかし，誤差があるので厳密には成り立たない．そこで次の**最小二乗法**を用いる．

$$\sum_{\alpha=1}^{N}(\boldsymbol{m}_\alpha, \boldsymbol{n})^2 \to \min \tag{1}$$

モーメント行列 $\boldsymbol{M}$ を

$$\boldsymbol{M}=\sum_{\alpha=1}^{N}\boldsymbol{m}_\alpha\boldsymbol{m}_\alpha^{\mathrm{T}} \tag{2}$$

と定義すると（$^{\mathrm{T}}$ はベクトルの転置），

$$\sum_{\alpha=1}^{N}(\boldsymbol{m}_\alpha, \boldsymbol{n})^2=(\boldsymbol{n}, \boldsymbol{M}\boldsymbol{n}) \tag{3}$$

と書ける．これは単位ベクトル $\boldsymbol{n}$ に関する2次形式であり，線形代数学でよく知られているように，これを最小にする $\boldsymbol{n}$ はモーメント行列 $\boldsymbol{M}$ の最小固有値に対する単位固有ベクトルで与えられる．

問題は，そのようにして当てはめた直線の信頼性である．これを定量的に評価するには，誤差の数理的モデルを導入して，それに基づいた統計的手法を用いる．**誤差のモデル**というのは，**誤差はどのように生じるのか**という前提条件を定めることである．誤差の原因は普通多くの要因が複雑に絡み合っているので，それを精密に解析することはなかなか難しい．しかし，誤差にはもともと予測不可能なランダムな部分があるので，あまり詳しく調べても意味がない．それよりも，誤差の大まかな性質に基づいて誤差の影響を減らす工夫をするのが実際的である．

もちろん誤差の細かい性質まで考慮すればするほどより効果的になるだろうが，細かく解析しようとすると，データや装置や処理に依存する多くのモデルパラメータを推定しなければならないので，そのための予備実験を何回も行なう必要がある．しかもその結果は，特定の場合にしか当てはまらないので一般性がない．したがって，モデルは**単純さとその効果とのバランス**で決めなければならない．このような考察を必要とするのが“工学”の研究の特色でもある．

そこで単純化して，エッジを構成する各画素の直線上の正しい位置（画素と画素の間でもよい）からのずれは各画素それぞれ独立にどの方向にも等しい確率で発生し，その大きさの標準偏差が ϵ 画素であるとする．この ϵ を**画素精度**と呼ぶ．画像が精密であれば誤差は離散化のためのみであり，ϵ はほぼ1画素程度であるが，画像が不完全であればその程度に応じて2〜3画素とか4〜5画素などと見積らなければならない．

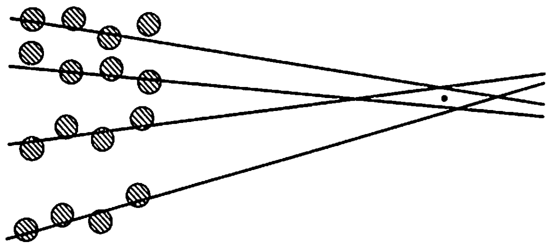
このような誤差のモデルを立てると，確率論の計算によって，当てはめた直線のNベクトルの統計的挙動が計算できる．これを**共分散行列**という量によって表すことができるが，その計算には固有値問題の**摂動定理**が必要となるので，詳細は省略する．結果として得られる式を検討すると，次のことがわかる(図1(b))．

- 当てはめた直線は，正しい位置にある直線と，ほぼエッジの中点で交わっている確率が極めて大きい．エッジ全体から離れるようにずれる確率は無視できるほど小さい．
- 当てはめた直線のずれは画像の精度が高いほど，またエッジを構成する画素が密であるほど小さく，画素精度 ϵ にほぼ比例し，エッジ密度(＝単位長さ当たりの画素の数)を ρ とすると $\rho^{-1/2}$ にほぼ比例する．
- 当てはめた直線のずれは，エッジの長さが長いほど小さく，エッジの長さを w とすると $w^{-3/2}$ にほぼ比例する．

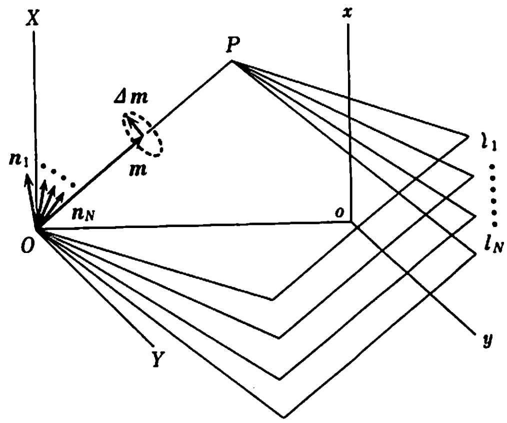
3　消失点の推定の信頼性

次に，当てはめた直線の消失点を計算しなければならない．しかし，各直線の当てはめが必ずしも正確ではないから，同一点で交わるはずの直線も一点で交わっているとは限らない(図2(a))．したがって，最も適切な点を正しい交点として推定しなければならない．これも透視変換のモデルを考えるとわかりやすい．やはり，このモデルも直線当てはめの場合と同じく便宜的なものであり，焦点距離 f も仮の値でよい．

当てはめた直線のNベクトルを $\boldsymbol{n}_\alpha$, $\alpha=1,\cdots,M$ とする．これらの共通の交点のNベクトルを $\boldsymbol{m}$ とす



（a） 消失点の推定



（b） 交点推定の原理

図 2

ると，図 2（b）からわかるように，誤差がなければ $(\boldsymbol{m}_\alpha, \boldsymbol{n})=0$, $\alpha=1,\cdots,M$ である．しかし，誤差があるので厳密には成り立たない．そこで次の**最小二乗法**を用いる．

$$\sum_{\alpha=0}^{M} W_\alpha(\boldsymbol{m}_\alpha, \boldsymbol{n})^2 \to \min \tag{4}$$

ただし W_α は当てはめた各直線の寄与の度合いを示す重みである．例えば長いエッジに当てはめた直線は短いエッジに当てはめた直線より信頼性が高いと考えられるから，より大きい重みを割り当てるべきである．統計学の理論を用いれば，当てはめた直線の信頼性を表す共分散行列から最適な重みを理論的に導くことができるが，詳細は省略する．

やはり**モーメント行列** $\boldsymbol{N}$ を

$$\boldsymbol{N}=\sum_{\alpha=1}^{M} \boldsymbol{n}_0 \boldsymbol{n}_\alpha^{\mathrm{T}} \tag{5}$$

と定義すると，

$$\sum_{\alpha=1}^{M}(\boldsymbol{n}_\alpha, \boldsymbol{m})^2=(\boldsymbol{m}, \boldsymbol{N}\boldsymbol{m}) \tag{6}$$

と書ける．したがって直線当てはめの場合と同様に，これを最小にする $\boldsymbol{m}$ はモーメント行列 $\boldsymbol{N}$ の最小固有値に対する単位固有ベクトルで与えられる．

少し脇道にそれるが，上の議論から誤差に関する考察を除けば，直線当てはめと交点推定の問題が数学的には**同じ形式**で記述されていることがわかる．実はこれは**射影幾何学**における**双対性**を反映したものである．上に定義した点と直線の“N ベクトル”は画像面を **2 次元射影空間**とみなしたときの点と直線の**同次座標**（を正規化したもの）にほかならない．射影幾何学では点と直線に関する正しい命題が得られたら，その命題中の「点」と「直線」を，「同一直線上にある（共線である）」と「同一交点をもつ（共点である）」を互いに交換しても正しい（**双対**な）命題が得られる．そして，同次座標で記述すれば，双対な式が同一の形式に書ける．

確かにこれは数学的には美しい結果であるが，誤差を考慮すると双対性がくずれる．それは，点の誤差と直線の誤差とは誤差の要因が異なるためである．射影幾何学は“誤差”や“計算のあふれ”などの現実的な要素を完全に無視し，公理から導いた純粋な抽象理論であり，それが“数学理論”としての価値を高めている．一方，工学的応用のためにはまさにそのような現実的な要素こそが中心的な課題であり，それをどう解決するかで“工学理論”としての価値が定まる．このことにも，数学者が指向する方向と工学者が指向する方向との乖離が見られる．

脇道が長くなったが，前節に導入した誤差のモデルと当てはめた直線の共分散行列を用いれば，このように推定した消失点の信頼性も統計的に評価することができる．これも共分散行列という量によって表せるが，やはり固有値問題の**摂動定理**を必要とするので，詳細は省略する．結果として得られた式を検討すると，次のことがわかる．

- 消失点の正しい位置からのずれは，消失点を定める直線群の中心線に沿ってエッジから遠ざかる，あるいは近づくように発生する確率が極めて大きく，それと直交する方向にずれる確率は無視できるほど小さい．
- 推定は平均的には正しい位置から直線群の中心線

に沿って遠ざかる方にやや偏る．

●消失点のずれは，各エッジに当てはめた直線の誤差にほぼ比例する（したがって，画素精度および各エッジの密度の $-1/2$ 乗，各エッジの長さの $-3/2$ 乗にほぼ比例する）．

●消失点のずれは，消失点を定める直線の間隔が狭いほど大きい．

●消失点のずれは，直線を当てはめたエッジから消失点が遠くにあるほど大きい．

第2点の偏りについては，**くりこみ法**と呼ぶ手法によって補正して不偏推定値を得ることができるが，詳細は省略する．

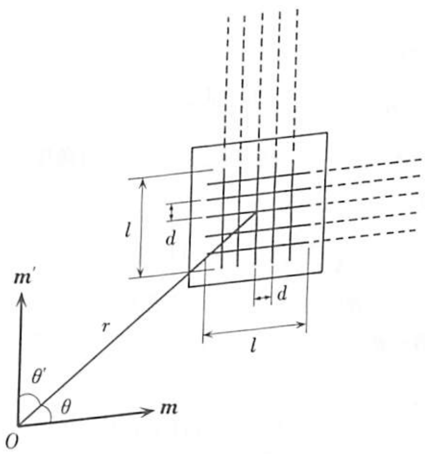
4 テストパタンの最適配置

消失点の誤差の統計的挙動がわかれば，前回の解説で示した式から計算される焦点距離 f の信頼性，具体的には f の**分散**が計算できる．“分散”というと，普通は多数のデータ（例えば多くの学生の試験の得点）があるとき，各データの平均値からの差の二乗の平均として定義するが，ここでの意味はやや違って，“データ”が一つもなくても分散が定義できる．これは確率論に慣れている人にとっては何でもないが，慣れていない人はとまどうかもしれない．

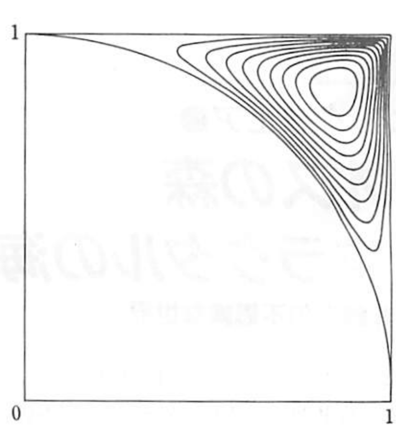
ポイントは**誤差のモデル**を仮定しているためである．誤差はこのモデルにしたがってランダムに発生すると仮定しているから，誤差がある値をとれば f がある値に定まり，誤差が別の値をとれば f は別の値となる．そこで誤差の値のすべての可能性（無限にあってもよい）を考えると，それに対する f の可能なすべての値の集合が得られる．そのような想像上の集合を**アンサンブル**と呼ぶ．ここでいう“分散”は，そのような仮想的なアンサンブルの各々の要素にその確率を与えて定まる f の分散のことである．これはちょうど二つのさいころを振ってある目の組み合わせが出る確率を「各さいころは互いに独立で，どの目の出る確率も等しい」という（仮想的な）“モデル”から計算するようなものである．どのような目の組み合わせがどの割合で出て，その平均，分散がどうなるかが，さいころを一度も振らずに（アンサンブルを考えて）計算できる．

このことから，今観測している画像には誤差がどのように入っているかまったくわからないが，たとえどんな誤差が入っていても，計算した焦点距離は真の値を中心として幅が標準偏差（＝分散の平方根）程度の範囲におさまっていると推論できる．したがって，真の値は計算値を中心とするその程度の範囲内にあると考えてよい．これが統計学でよく知られた**信頼区間**を計算する原理である．このため，“分散”が計算の信頼性を測る基本的な尺度（大きいほど信頼性が低い）となる．

今格子パタンを図3(a)のように置いたとする．格子パタンは互いに直交する長さ l，幅 d の各方向に N（奇数）本の線分からなっている．格子の中心が光軸上の視点 O から距離 r のところにあり，格子の線分はそれぞれ光軸から角度 θ, θ' の方向に伸びている．このとき，焦点距離 f の分散 $V[f]$ は次のようになるこ



(a) 格子パタンの配置



(b) パタンの向きと計算の信頼性の関係を表す等高線

図3

とが導ける（ϵ, ρ はそれぞれ２節に述べた画素精度，エッジ密度である）．

$$V[f]=\frac{18\epsilon^2 r^5}{N(N^2-1)\rho f^3 d^2 l^3(1-\cos^2\theta-\cos^2\theta')^2} \times\left(\frac{\sin\theta}{\cos^2\theta}+\frac{\sin\theta'}{\cos^2\theta'}\right) \tag{7}$$

三角関数の公式 $\cos^2\theta=1-\sin^2\theta$, $\cos^2\theta'=1-\sin^2\theta'$ を代入すると，上式は $\sin\theta$, $\sin\theta'$ の関数とみなせる．$x=\sin\theta$, $y=\sin\theta'$ と置き，

$$F(x,y)=\frac{1}{x^2+y^2-1}\left(\frac{1}{1-x^2}+\frac{1}{1-y^2}\right) \tag{8}$$

と定義すると，式（７）の右辺は（定数）$\times F(x,y)$ と書ける．関数 $F(x,y)$ の定義域は $0<x<1$, $0<y<1$, $x^2+y^2>1$ である．境界で ∞ となるので，見やすいように $1/F(x,y)$ の等高線を示したものが図３（ｂ）である．これを見ると，確かにあるところで $F(x,y)$ が最小になっている．具体的に計算すると，最小値を与える θ, θ' は次の値となる．

$$\theta=\theta'=\sin^{-1}\sqrt{\frac{\sqrt{33}+3}{12}}=58.6106038\cdots° \tag{9}$$

このときパタンは画像面に対して次の角度 β をなしている．

$$\beta=\cos^{-1}\sqrt{\frac{\sqrt{33}-3}{6}}=47.4422040\cdots° \tag{10}$$

このように置いた格子パタンを撮影すると，画像の中央で格子線は次の角度 γ で交わっているように見える．

$$\gamma=\pi-\cos^{-1}\frac{\sqrt{33}-5}{2}=111.856381\cdots° \tag{11}$$

前回にも指摘したように，焦点距離が定まる幾何学的原理は奥行きと投影像の大きさとの関係にあるので，パタンの傾きが大きいほど精度のよい推定が行なえる．しかし画像処理の観点からは，パタンが傾くと傾いた方向にエッジが短くなって直線当てはめの精度が低下し，それと直交する方向の直線間の間隔が狭くなって，消失点の推定の精度が低下する．これらの互いに相反する要因がちょうどつり合う最適の傾きが上に示した値である．これが測定によって定まるパラメータなどにはよらない**数学的に普遍的な値として定ま**るのは大変興味のあることである．

これで格子パタンの最適配置が得られ，そのように配置すれば最適なキャリブレーションができることがわかった．これですべてが解決され，もう考えることはないであろうか．普通，工学の問題ではそういうことはなく，必ずといっていいほど“次の問題”が控えており，これがまた工学を研究する楽しみでもある．これは次回のテーマである．

参考文献

1) 甘利俊一，金谷健一：「線形代数」，講談社，1987.
2) K. Kanatani: Statistical analysis of focal-length calibration using vanishing points, *IEEE Transactions on Robotics and Automation*, 8-6 (1992), 767-775.
3) 金谷健一，浦沢康二：消失点とカメラキャリブレーションの統計的信頼性，日本ロボット学会誌，9-7(1991)，813-820.
4) 髙木幹雄，下田陽久監修：「画像解析ハンドブック」，東大出版会，1991.
5) 鳥脇純一郎：「画像理解のためのディジタル画像処理」，Ｉ，II，昭晃堂，1988.

（かなたに・けんいち，群馬大学・工学部）