

画像の3次元解釈の統計的信頼性

金 谷 健 一†

画像の3次元解釈の信頼性に関する新しい理論体系を提案する。まずエッジへの直線当てはめの信頼性を定量的に評価し、その直線の信頼性から消失点の計算の信頼性を定量的に評価する。その結果を焦点距離のキャリブレーションに応用し、推定値の信頼区間を評価する。さらに直交性当てはめの信頼性を理論的に評価し、最後にアドホックな定数を含まない幾何学的仮説の検定基準を導出する。

Statistical Reliability of 3-D Interpretation from Images

KENICHI KANATANI†

The reliability of 3-D interpretations computed from images is analyzed by employing a statistical model of image noise. First, the reliability of line fitting to edges is evaluated in terms of image noise characteristics. Then, the reliability of vanishing point estimation is deduced from the reliability of lines. The result is applied to focal length calibration, and the confidence interval of the optimal estimate is evaluated. The reliability of fitting an orthogonal frame to sensor data is also discussed. Finally, statistical criteria are derived for testing edge groupings, vanishing points, focuses of expansion, and vanishing lines.

1. はじめに

コンピュータビジョンの研究は人間の視覚認識をまねた未知の3次元シーンの“理解”を目的として米国でスタートし、特徴抽出、領域分割、対象の分類・識別などのさまざまな手法が開発された。これは基礎的研究であり、特定の応用を指向したものではなかったが、最近のロボット技術の進歩により、ロボットが作業する既知あるいは管理された空間における物体の識別、その3次元的位置と向きへの測定、ロボットの運動の計算や衝突回避の制御などの環境の“3次元計測”における実際的な応用の可能性が急速に広がった。このような計測目的の応用においてはその計算の“信頼性”が主要な課題となる。

従来、計算の信頼性は実データやシミュレーションによる実験によって経験的に評価されることが多かったが、今後の発展のためには画像のノイズの数理的モデルを導入し、理論的に信頼区間を評価したり最適性を証明したりする“理論体系”が必要となる。これは特に複数のセンサデータを統合する“センサフュージョン”において特に重要になる。なぜなら個々のデー

タの信頼性に応じた重み付けをしなければ、統合によって信頼性が向上するとは限らないからである。本論文では画像データに基づく3次元解釈の信頼性に関する新しい理論体系を提案し、具体的な例に適用してその基本的考え方を述べる。

2. 画像ノイズの統計的モデル

図1のようなカメラモデルを考える^{4)~6),9)}。以下、座標原点 O を視点、定数 f を焦点距離と呼ぶ。画像面上の点は視点とその点を通る“視線”の方向を示す単位ベクトル m (符号は任意) によって指定できる。また画像面上の直線は視点とその直線を含む平面の単位法線ベクトル n (符号は任意) によって指定できる。 m, n を N ベクトルと呼ぶ^{5),6),9)}。 N ベクトルを用いれば、公理に基づいて抽象的に定義された射影幾何学の多くの概念を具体的な“計算過程”として表現することができる^{5),6),8),9)}。そのような定式化を計算射影幾何学とも呼ぶ。

画像面上の N ベクトル m をもつ点にノイズが加わり、 N ベクトル $m' = m + \Delta m$ が観測されたとする。ただし“ノイズ”とは画像処理によって抽出された点データの不正確さをいう。以下では Δm を確率変数とみなし、その共分散行列を次のように定義する ($E[\cdot]$ は期待値)。

† 群馬大学工学部情報工学科
Department of Computer Science, Gunma University

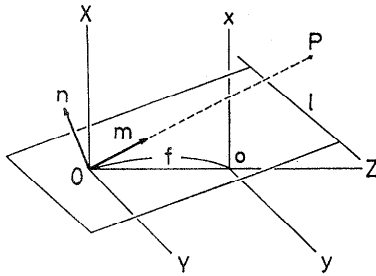


図1 カメラモデルとNベクトル
Fig. 1 Camera model and N-vectors.

$$V[m] = E[\Delta m \Delta m^T]. \quad (1)$$

ノイズは各画素に独立かつ一様に発生し、その点を中心として等方的であると仮定する。その点からの距離の平方二乗平均を ε (画素) とし、これを画素精度と呼ぶ。このとき次のことがわかる^{20), 21)} (I は単位行列)。

【命題 1】 画像の大きさが焦点距離 f に比べて小さければ、点の N ベクトルの共分散行列 $V[m]$ は次のように表せる。ただし $\bar{\varepsilon} = \varepsilon/f$ 。

$$V[m] = \frac{\bar{\varepsilon}^2}{2}(I - mm^T). \quad (2)$$

これは誤差の一つの理想化であり、実際には画像やその処理によってこのモデルからずれるかもしれない。しかし、このモデルによれば解析が容易となり、各種の最適化手法が導けるので、その結果が厳密には最適でないとしても第一近似の意味で信頼性の評価、精度の向上など有意義な応用ができる。

3. 直線当てはめの信頼性

通常の画像処理ではハフ変換やエッジ作用素などを適用して直線部分を検出し、しきい値処理、細線化などにより画素の列としてのエッジを求め、最小二乗法などによって直線を当てはめて、交点の計算や消失点の推定などの処理に進む。エッジに誤差がないときの直線の N ベクトルを n とし、誤差を含んだ画素に当てはめた直線の N ベクトルを n' とすると、直線当てはめの信頼性は共分散行列

$$V[n] = E[\Delta n \Delta n^T] \quad (3)$$

によって評価できる。命題1から次の命題が得られる²⁰⁾。ただし γ はエッジ密度 (単位画素長さ当たりのエッジ画素数) である。

【命題 2】 長さ w 、方向 u のエッジに当てはめた直線の N ベクトル m の共分散行列は次のように表せる。ただし、 $\kappa = \varepsilon^2/\gamma$ であり、 m_G はエッジの中点の

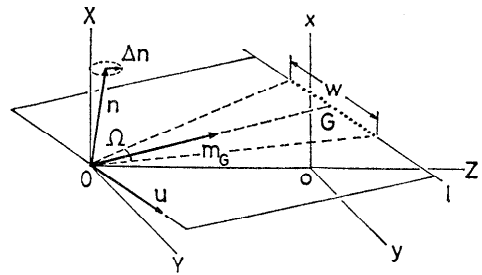


図2 エッジへの直線当てはめ
Fig. 2 Line fitting to an edge.

N ベクトルである。

$$V[n] = \frac{6\kappa}{w^3}uu^T + \frac{\kappa}{2f^2w}m_G m_G^T. \quad (4)$$

定数 κ を解像度定数と呼ぶ。上式がよい近似であることは数値実験でも確かめることができる²¹⁾。これから次のことがわかる (図 2)。

- N ベクトル n の誤差はほとんど常に方向 u に生じる。
- 当てはめの誤差は解像度定数 κ に対してはほぼ $\kappa^{1/2}$ に比例する (したがって画素精度 ε に比例し、エッジ密度 γ に対しては $\gamma^{-1/2}$ に比例)。
- 当てはめの誤差はエッジの長さ w に対しては $w^{-3/2}$ に比例する。

ここでは直線がエッジの画素列に最小二乗法で当てはめられると仮定しているが、単に両端点を結んだとすれば、その N ベクトル n の共分散行列 $V[n]$ は次のようになる²¹⁾。

$$V[n] \approx \left(\frac{\varepsilon}{w}\right)^2 uu^T. \quad (5)$$

4. 消失点の推定

空間中で平行な直線群の投影像は画像面で共点 (延長すれば一点で交わる) であり、その共通の交点が消失点である。消失点の N ベクトルは直線の 3 次元的方向を示している^{4) - 6), 9)}、消失点の計算は画像の 3 次元解釈の最も基本的な計算の一つである^{2), 16), 18)}。

いま $\{n_\alpha\}$, $\alpha=1, \dots, N$ を空間中の平行線の投影像の N ベクトル、 m をそれらの消失点の N ベクトルとする。仮想的な中心線 l_G の N ベクトルを n_G とし、 n_G と m の両方に直交する N ベクトル m_G をもつ点 P_G を消失点 P の共役点と呼ぶ^{5), 6), 9)} (図 3)。各直線 l_α はそれぞれ長さ w_α のエッジに当てはめたものとし、各エッジの中心点 G_α の N ベクトルを m_{G_α} とする。ベクトル n_G, n_α のなす角 ϕ_α を直線 l_α

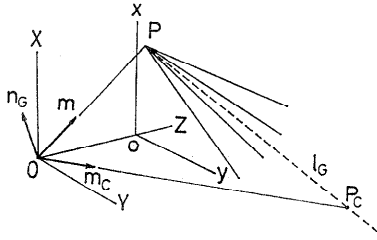


図3 中心線 l_G と消失点 P の共役点 P_G
 Fig. 3 The center line l_G and the conjugate point P_G to the vanishing point P .

の(中心線 l_G からの) 偏角と呼び、ベクトル $m_{c\alpha}$, m のなす角 θ_α を消失点 P の α 番目のエッジの中心点からの視差と呼ぶ。消失点の N ベクトルの共分散行列は次のように表せる^{20), 21)}。

$$V[m] \approx \frac{6\kappa m c m^T}{\sum_{\alpha=1}^N w_\alpha^2 \sin^2 \phi_\alpha / \sin^2 \theta_\alpha} \quad (6)$$

これから次のことがわかる。

- 消失点の誤差は中心線に沿って生じやすい。
- 消失点の誤差は解像度定数に対してほぼ $\kappa^{1/2}$ に比例する。
- 消失点の誤差は各エッジの長さ w_α に対してほぼ $w_\alpha^{-3/2}$ に比例する。
- 消失点の誤差は各直線の偏角 ϕ_α に対してほぼ $1/\sin \phi_\alpha$ に比例する。
- 消失点の誤差は各エッジの中心からの視差 θ_α に対してほぼ $\sin \theta_\alpha$ に比例する。

5. 焦点距離のキャリブレーション

画像を N ベクトルによって解析するときは図1のカメラモデルを仮定しているが、このカメラモデルは3次元解釈を問題にしない限り、例えば消失点の推定などでは任意に仮定してよい。しかし空間中での長さや角度などの3次元解釈を問題にするときには、仮定したカメラモデルを実際のカメラと一致させるキャリブレーションが必要となる。特に重要なのが焦点距離 f であり、これは次のようにして定められる^{4) - 6), 10), 12), 13)}。

【命題 3】 空間中で互いに直交する2直線の消失点の仮の焦点距離 f に対する N ベクトルを m, m' とすると、真の焦点距離 f は次のように与えられる。

$$f = f \sqrt{\frac{m_1 m'_1 + m_2 m'_2}{m_3 m'_3}} \quad (7)$$

これから次の手順が考えられる。まず空間中に配置

した直交格子パターンを撮像する。そして格子像に直線を当てはめて、仮の焦点距離 f を用いて2方向の消失点の N ベクトルを計算する。最後に式(7)により真の焦点距離 f を計算する。

二つの消失点はそれぞれ N 本および N' 本の直線の交点であるとし、それぞれの直線は長さ w_α および $w'_{\alpha'}$ のエッジに当てはめたものとする。それぞれの中心線は画像原点で交わるとし、各直線の偏角を ϕ_α および $\phi'_{\alpha'}$ とする。また二つの消失点の画像原点からの視差をそれぞれ θ および θ' とする。上述の手順で計算した焦点距離の分散 $V[f]$ は次のように表せることが導ける^{10), 13), 22)}。

$$V[f] \approx \frac{3\kappa f^2}{2} \left(\frac{1/\cos^2 \theta}{\sum_{\alpha=1}^N w_\alpha^2 \sin^2 \phi_\alpha} + \frac{1/\cos^2 \theta'}{\sum_{\alpha'=1}^{N'} w'_{\alpha'}^2 \sin^2 \phi'_{\alpha'}} \right) \quad (8)$$

図4のように配置した正方格子パターンを考える。視点 O から格子パターンの中心までの距離を r とする。格子は2組の N (奇数) 本の直線からなっているとす。パターンの大きさを l とし、各格子の大きさを d とする。 l/r が小さければ式(8)を最小にする θ, θ' は次のように与えられる²²⁾。

$$\theta = \theta' = \sin^{-1} \sqrt{\frac{3 + \sqrt{33}}{12}} \approx 58.6^\circ \quad (9)$$

このとき画像上では格子は画像の中心において次の角度で交差する。

$$\gamma = \pi - \cos^{-1} \frac{\sqrt{33} - 5}{2} \approx 111.9^\circ \quad (10)$$

直観的にいえば、焦点距離はパターンに奥行きがなければ定まらないので、式(8)の $1/\cos^2 \theta, 1/\cos^2 \theta'$ が小さくなるようにパターンをカメラに対してなるべく大きい角度で傾けなければならない。しかし傾け過ぎると画像中のエッジが短くなって直線当てはめの精度が低下したり、エッジ間の間隔が狭くなって消失点の計

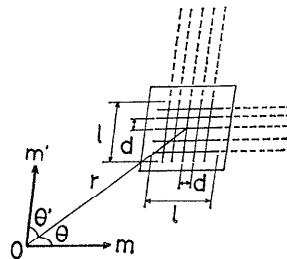


図4 格子パターンの3次元配置
 Fig. 4 3-D configuration of the grid pattern.

算の精度が低下する。これらの競合する要因がちょうどバランスするとき式(9)が最小値をとる。

一般に測定の精度を上げるには測定を繰り返してその平均をとればよいが、単純に平均をとると、ひとつでも大きくずれたデータがあれば結果が大きく左右される。したがって信頼性の高いデータには大きく、信頼性の低いデータには小さく重み付けした加重平均が必要である。パタンを空間中でいろいろに配置して N 枚の画像から焦点距離を $f_\alpha, \alpha=1, \dots, N$ を計算したとき、分散を最小にする意味で最適な加重平均は次のように与えられることが証明できる²²⁾。

$$f = \sum_{\alpha=1}^N W_\alpha f_\alpha, W_\alpha = \frac{1}{V[f_\alpha]} \bigg/ \sum_{\beta=1}^N \frac{1}{V[f_\beta]} \quad (11)$$

このようにして計算した最適推定値 f の分散は次のようになる²²⁾。

$$V[f] = 1 \bigg/ \sum_{\alpha=1}^N \frac{1}{V[f_\alpha]} \quad (12)$$

これから正規分布による近似を用いて、推定値 f の $(100-a)\%$ 信頼区間が次のように与えられる。

$$[f - \lambda_a \sqrt{V[f]}, f + \lambda_a \sqrt{V[f]}]. \quad (13)$$

ただし λ_a は正規分布の $a\%$ 点である。しかし、これを計算するには解像度定数 κ を知る必要がある。これをあらかじめ予測することが難しければ、 t 分布による区間推定を行えばよい。 t 分布による $(100-a)\%$ 信頼区間は次のようになる。

$$\left[f - t_{a, N-1} \frac{s}{\sqrt{N-1}}, f + t_{a, N-1} \frac{s}{\sqrt{N-1}} \right]. \quad (14)$$

ただし $s = \sqrt{\sum_{\alpha=1}^N W_\alpha (f_\alpha - f)^2}$ であり、 $t_{a, N}$ は自由度 N の t 分布の $a\%$ 点である。こうすれば解像度定数 κ は任意に仮定してよい。

[例 1] 図5のパタンを空間中のいろいろな位置、い

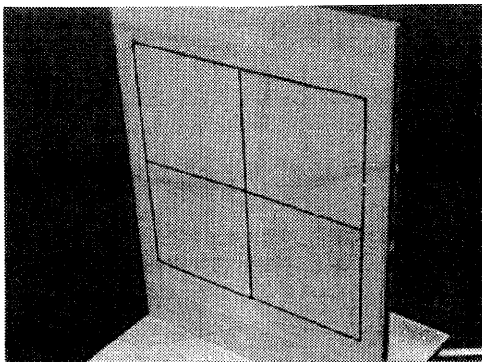


図5 格子パタンの実画像

Fig. 5 A real image of a grid pattern.

表1 実画像によるキャリブレーションの例
Table 1 Calibration data obtained from real images.

| α | f_α | $V[f_\alpha]$ | W_α | θ_α | θ'_α |
|----------|------------|---------------|------------|-----------------|------------------|
| 1 | 1027.0 | 119.1 | 0.021 | 84.6 | 63.7 |
| 2 | 1136.0 | 28.4 | 0.083 | 75.7 | 61.0 |
| 3 | 1149.9 | 18.6 | 0.126 | 69.9 | 60.4 |
| 4 | 1146.4 | 13.5 | 0.174 | 58.3 | 59.4 |
| 5 | 1092.4 | 21.4 | 0.110 | 42.3 | 58.3 |
| 6 | 1167.6 | 55.3 | 0.042 | 80.7 | 74.6 |
| 7 | 1147.3 | 19.3 | 0.122 | 68.8 | 73.9 |
| 8 | 1142.2 | 17.9 | 0.131 | 54.0 | 73.4 |
| 9 | 1092.5 | 20.0 | 0.117 | 45.7 | 72.8 |
| 10 | 1056.9 | 31.4 | 0.075 | 35.6 | 71.8 |

ろいろな向きに置き、エッジを検出し、直線を当てらめ、消失点を推定する。表1にそれから推定した焦点距離を示す。また、解像度定数を $\kappa=1$ として理論的に計算した分散 $V[f_\alpha]$ およびそれから定まる最適重み W_α 、消失点の視差 θ, θ' も示す。最初および6番目のものは格子の一方が画像面に対してほぼ平行な場合であり、このような配置では信頼性が非常に低いことがわかる。4番目の例(図5の場合)が最適な配置に最も近い。このとき格子は画像上で角度 112° で交差している。式(11)による最適推定値は $f = 1125.0$ (画素)であり、式(14)による 95% 信頼区間は $[1099.6, 1150.4]$ である。□

6. 直交性の当てはめ

ロボットの室内作業環境では物体や壁や床などに直交する辺が多く、ロボットの環境理解や運動制御の重要な手がかりとなる。しかし、直交する3方向を別々に(例えば消失点を計算して)計測しても、誤差のためそれらが厳密に直交しているとは限らない。したがって、それらに厳密な直交系を当てはめる必要が生じる。いま単位ベクトル $\{m_i\}, i=1, 2, 3$ が計測によって得られたとし、これに正規直交系 $\{r_i\}, i=1, 2, 3$ を次のように当てはめる問題を考える。

$$\sum_{i=1}^3 W_i \|r_i - m_i\|^2 \rightarrow \min. \quad (15)$$

重み W_i はデータ m_i の信頼性に応じて定めるべきであり、 m_i の共分散行列が $V[m_i]$ であれば、 $W_i \propto 1/\text{tr } V[m_i]$ とするのが妥当である。

式(15)の解は“特異値分解”による方法、“極分解”による方法、および“四元数表示”による方法によって解析的に計算できる^{5), 6), 9), 19)}。こうして計算した $\{r_i\}$ を列とする回転行列 R を最良回転行列と呼ぶことにする。

$\{m_i\}$ が初めから正規直交系であれば当然 $\{r_i\}$ は $\{m_i\}$ に一致するが、もし誤差があれば最良回転行列 R は別の回転行列 R' に摂動する。しかし R も R' も回転行列であるから、その差はある軸 l (単位ベクトル) の周りのある角度 ΔQ の回転である。そこで $\Delta l = \Delta Q l$ と置き、回転行列 R の共分散行列を

$$V[R] = E[\Delta l \Delta l^T] \tag{16}$$

と定義する。 $V[R]$ の最大固有値に対する固有ベクトルが誤差の最も生じやすい回転軸であり、その固有値が誤差の回転角の二乗平均を表している。この共分散行列は理論的に次のように評価できる¹⁹⁾ ((...) はベクトルの内積)。

【命題 4】 共分散行列 $V[m_i]$ をもつ単位ベクトル $\{m_i\}$, $i=1, 2, 3$ に当てはめた最良回転行列 R の共分散行列は次のように与えられる。

$$V[R] = \sum_{i,j=1}^3 \frac{\sum_{k=1}^3 W_k^2 (r_i \times r_k, V[m_k] (r_j \times r_k))}{(1-W_i)(1-W_j)} r_i r_j^T \tag{17}$$

【例 2】 図 6 (a) は直方体の実画像であり、図 6 (b) はそのエッジ画像である。3 方向の辺の消失点を計算すると、それらの 3 次元方向 m_1, m_2, m_3 がそれぞれ次のように推定できる (X 軸は上, Y 軸は右に正)。

$$\begin{pmatrix} 0.244 \\ -0.792 \\ 0.559 \end{pmatrix}, \begin{pmatrix} 0.300 \\ 0.636 \\ 0.711 \end{pmatrix}, \begin{pmatrix} -0.914 \\ 0.030 \\ 0.405 \end{pmatrix}$$

それぞれの理論的共分散行列 $V[m_1], V[m_2], V[m_3]$ は式(6)によって与えられる。これから計算した最良回転行列 R は次のようになる。

$$\begin{pmatrix} 0.239 & 0.320 & -0.917 \\ -0.780 & 0.626 & 0.015 \\ 0.578 & 0.712 & 0.399 \end{pmatrix}$$

あてはめた方向と m_1, m_2, m_3 の差はそれぞれ 1.35° ,

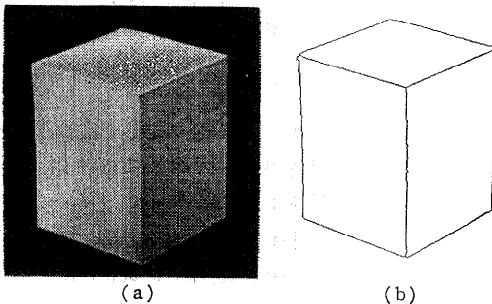


図 6 (a) 直方体の実画像 (b) そのエッジ画像
Fig. 6 (a) A real image of a rectangular box.
(b) Detected edges.

$1.25^\circ, 0.97^\circ$ である。式(17)によってこの最良回転行列の理論的共分散行列を計算すれば、この当てはめの(未知の)正しい位置からのずれの回転角 ΔQ の標準偏差は 0.49° であると推定できる。 □

7. 幾何学的仮説の検定

画像中の 3 次元物体を認識するための標準的手法として、画像上に検出した物体像のエッジに対して、データベースに登録してある複数の線画モデルのそれぞれを重ね合わせて、モデルのエッジと物体像のエッジとが最もよくマッチするモデルを選ぶ方法がよく知られている^{20, 15)} (図 7)。モデルのエッジと物体のエッジとのマッチの尺度は従来は発見的に定義されることが多かったが、次の命題より $(n', V[n]^{-1}n')$ を用いるのが統計的に合理的である。ただし $V[n]^{-1}$ は $V[n]$ の“一般逆行列”であり、 $\chi_{2,2}^2$ は自由度 2 の χ^2 分布の $\alpha\%$ 点である。

【命題 5】 N ベクトル n' が共分散行列 $V[n]$ の確率変数 n の実現値であるという仮説は、次式が成り立てば有意水準 $\alpha\%$ で棄却される。

$$(n', V[n]^{-1}n') > \chi_{2,2}^2 \tag{18}$$

本来はつながっているべきエッジが切れ切れに検出されることがある (図 8)。これらを統合してよいかを考える。単純な方法としては、隣り合うエッジの向きや位置のずれが適当に設定したしきい値より小さければ一本のエッジに置き換えればよい^{11, 23)}。しかし、これを盲目的に行うと矛盾を生じる^{7), 11)}。これを避けるには、エッジが統合できるという仮説を立てて直線を当てはめて、各エッジのその直線からのずれがその

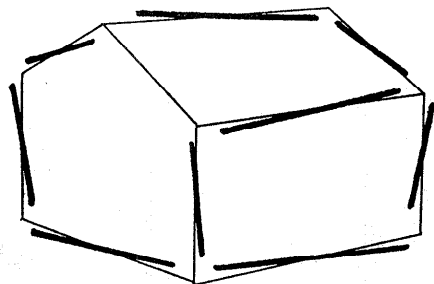


図 7 モデルマッチング
Fig. 7 Model matching.

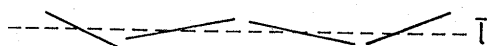


図 8 エッジを統合してよいか
Fig. 8 How should we group segments together?

仮説を棄却できるほど大きいかを次の命題によって検定すればよい^{7),11)}.

【命題 6】 第 α 番目のエッジの N ベクトル, その中点の N ベクトル, その方向ベクトル, その長さをそれぞれ $\mathbf{n}_\alpha, \mathbf{m}_{G\alpha}, \mathbf{u}_\alpha$ (単位ベクトル), w_α (画素) とする ($\alpha=1, \dots, N$). 当てはめた直線の N ベクトル $\bar{\mathbf{n}}$ が次の不等式を満たせば, すべてのエッジが同一直線上にあるという仮説は有意水準 $(100-a)\%$ で棄却される.

$$\sum_{\alpha=1}^N \left(\frac{w_\alpha^3}{6} (\bar{\mathbf{n}}, \mathbf{u}_\alpha)^2 + 2f^2 w_\alpha (\bar{\mathbf{n}}, \mathbf{m}_{G\alpha})^2 \right) > \kappa \chi_{a, 2N}^2. \quad (19)$$

消失点は画像の3次元解釈の重要な手がかりであるが, 画像の誤差のため平行線の投影像が一点で交わるとは限らない (図 9). 画像上に検出されたエッジが共点であるかどうかは, その延長線の互いの交点の不一致の程度を適当に設定したしきい値と比較して判断すればよいようであるが^{4),17),18)}, そのしきい値を固定すれば不合理が生じる^{7),11)}. この場合もエッジは共点であるという仮説を立てて, 各エッジのずれがその仮説を棄却できるほど大きいか検定すればよい^{7),11)}, エッジの統合の場合と同様にして次の検定基準を得る¹⁴⁾.

【命題 7】 第 α 番目のエッジの N ベクトル, その中点の N ベクトル, その長さをそれぞれ $\mathbf{n}_\alpha, \mathbf{m}_{G\alpha}, w_\alpha$ (画素) とする ($\alpha=1, \dots, N$). 次の不等式が満たされれば, N ベクトル \mathbf{m} の点が共通の交点であるという仮説は有意水準 $(100-a)\%$ で棄却される.

$$\sum_{\alpha=1}^N w_\alpha^3 \left(1 - \frac{|\mathbf{m}, \mathbf{n}_\alpha \mathbf{m}_{G\alpha}|^2}{1 - (\mathbf{m}, \mathbf{m}_{G\alpha})^2} \right) > 6\kappa \chi_{a, 2N}^2. \quad (20)$$

物体が空間中を並進する, あるいは静止した物体に

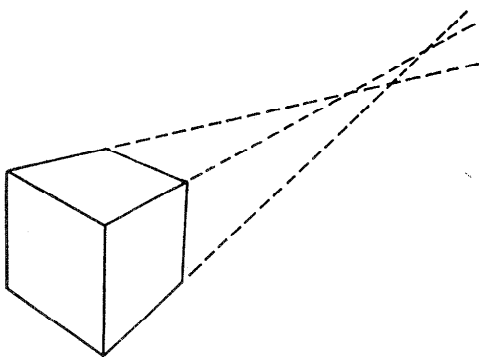


図 9 エッジは共点とみなせるか
Fig. 9 How can we judge the concurrency of edge segments?

対してカメラが並進するとき, 対応する特徴点の画像上での軌跡は共点であり, その共通の交点 (出現点) の N ベクトルが空間中での3次元方向を示す^{5),6),9)}. これに対しても消失点とほとんど同じ議論ができる. 異なるのは, 消失点がエッジに当てはめた直線の交点であるのに対して, 出現点は対応する特徴点を結んで得られる軌跡の交点であり, その誤差は特徴点検出の誤差に起因する点である. このとき上の命題に対応する検定の基準は次のようになる¹⁴⁾.

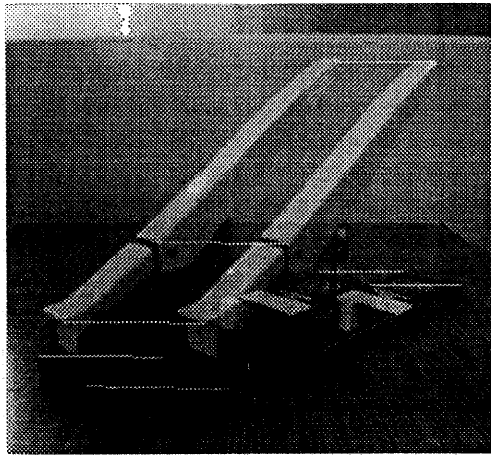
【命題 8】 第 α 番目の軌跡の N ベクトル, 対応する特徴点の中点の N ベクトル, 対応する特徴点の間隔をそれぞれ $\mathbf{n}_\alpha, \mathbf{m}_{G\alpha}, w_\alpha$ (画素) とする ($\alpha=1, \dots, N$). 次の不等式が満たされれば, N ベクトル \mathbf{m} の点が共通の交点であるという仮説は有意水準 $(100-a)\%$ で棄却される.

$$\sum_{\alpha=1}^N w_\alpha^2 \left(1 - \frac{|\mathbf{m}, \mathbf{n}_\alpha \mathbf{m}_{G\alpha}|^2}{1 - (\mathbf{m}, \mathbf{m}_{G\alpha})^2} \right) > \epsilon^2 \chi_{a, 2N}^2. \quad (21)$$

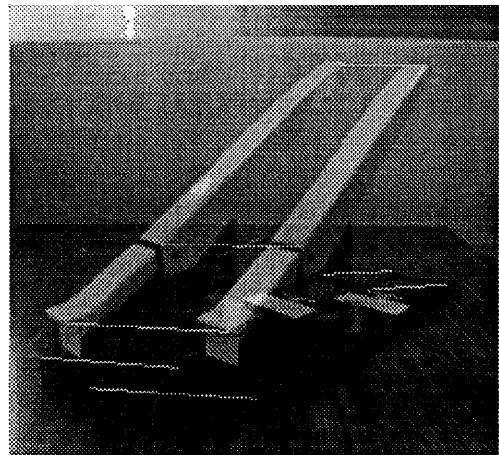
【例 3】 図 10 は移動物体の画像の重ね合わせである. (a)は並進のみ, (b)は並進と微小回転, (c)は並進とそれより大きい回転を施している. 特徴点を7個選び, 7本の軌跡を定義し, それらが共点であるという仮説を立てる. この仮説が受理されるか棄却されるかは各特徴点の検出精度 ϵ (画素) と有意水準に依存する. 例えば $\epsilon \leq 2.5$ なら (a) では有意水準 95% で仮説が受理され, (b), (c) では棄却される. $2.5 < \epsilon \leq 7$ なら (a), (b) では有意水準 95% で仮説が受理され, (c) では棄却される. また $\epsilon > 7$ なら (a), (b), (c) 共に有意水準 95% で仮説が受理される. □

空間中の N 組の平行組群の消失点を $\{P_\alpha\}$, $\alpha=1, \dots, N$ とする. もしそれらの平行線が空間中ですべて水平であるとすれば, それらの消失は共点, すなわちすべて同一直線 (“水平線”, 一般には消失線) の上にある. しかし消失点を画像から計算すれば, 誤差のために消失点が共線であるとは限らない (図 11). しかも消失点は計算で定義されるので, 原画像中に位置するとは限らず, 無限遠方にあるかも知れない. この場合も消失点が共線であるという仮説を立て, 各消失を定義するエッジのずれがその仮説を棄却できるほど大きいか検定すればよい¹⁴⁾.

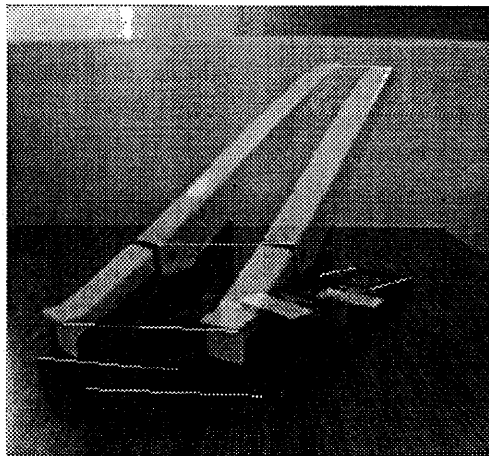
仮説に基づいて推定した消失線 \bar{l} の N ベクトルを $\bar{\mathbf{n}}$ とし, \bar{l} 上で各消失点 P_α に最も近い点を \bar{P}_α とする. 消失点 P_α の N ベクトルを \mathbf{m}_α とすると, 点 \bar{P}_α の N ベクトル $\bar{\mathbf{m}}_\alpha$ は次のように与えられる^{7),11)}. ただし $\mathbf{k} = (0, 0, 1)^T$ である.



(a) 並進
(a) A pure translation.



(b) 並進と微小回転
(b) A translation and a small rotation.



(c) 並進と回転
(c) A translation and a rotation.

図 10 移動物体の重ね合わせ画像
Fig. 10 Superimposed images of a moving object.

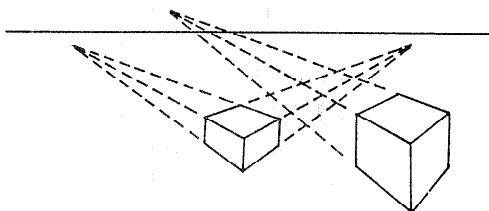


図 11 消失点は共線とみなせるか
Fig. 11 How can we judge the collinearity of vanishing points?

$$\bar{m}_\alpha = \pm N \left[\mathbf{k} - (\bar{\mathbf{n}}, \mathbf{k}) \bar{\mathbf{n}} - \frac{|\bar{\mathbf{n}}, \mathbf{m}_\alpha, \mathbf{k}|}{(\mathbf{m}_\alpha, \mathbf{k})} \bar{\mathbf{n}} \times \mathbf{k} \right]. \quad (22)$$

点 \bar{P}_α を正しい消失点とみなして、消失点の検定と同様にすれば、次の検定基準を得る¹⁴⁾。

【命題 9】 検定すべき消失点 P_α の N ベクトルを \mathbf{m}_α とする ($\alpha=1, \dots, N$)。消失点 P_α を定義する第 β 番目のエッジの N ベクトル, その中点の N ベクトル, その長さをそれぞれ $\mathbf{n}_\beta^{(\alpha)}, \mathbf{m}_{G\beta}^{(\alpha)}, \omega_\beta^{(\alpha)}$ (画素) とする ($\beta=1, \dots, N^{(\alpha)}$)。次の不等式が成立すれば、すべての消失点が共線であるという仮説は有意水準 $(100-a)\%$ で棄却される $\left(n = \sum_{\alpha=1}^N N^{(\alpha)} \right)$ 。

$$\sum_{\alpha=1}^N \sum_{\beta=1}^{N^{(\alpha)}} \omega_\beta^{(\alpha)3} \left(1 - \frac{|\bar{\mathbf{m}}_\alpha, \mathbf{n}_\beta^{(\alpha)}, \mathbf{m}_{G\beta}^{(\alpha)}|^2}{1 - (\bar{\mathbf{m}}_\alpha, \mathbf{m}_{G\beta}^{(\alpha)})^2} \right) > 6\kappa\chi_{a, 2n}^2. \quad (23)$$

ここに示した判定基準は微小誤差，正規分布，第一近似など多くの仮定に基づいているため，大きなずれに対しては厳密な確率としての定量的意味はなくなるが，形式的にはどんなずれに対しても適用できる。しかし，実際の場面では統計的要因以外も考慮して，例えばエッジ統合の例では原画像の濃淡値やエッジ強度，さらにはシーンに関する知識（直線からなる室内シーンである，等）を用いれば，より現実的な判定ができると思われる。また，画像の精度を表す定数 ϵ, κ を原画像の濃淡値の様子によっては場所ごとに変えたほうがよいこともありうる。他の基準についても同様であり，いろいろな変形や拡張が可能であろう。しかし，本章で強調したいのは基本的な考え方である。

従来の研究では，判定の必要が生じる度に“適当な小さい定数”を導入し，それと比較するということがよく行われる。そのような方法が実際的であるとしても，一般性を欠くのでその場限りとなり，普遍的な指針とはなりえない。これに対して本章で示した基準はアドホックな定数を一切含んでいない。実際にはいくつかの定数を適当に調節する必要があるかもしれないが，それらがすべて幾何学的な，あるいは（厳密な確率の意味はないとしても）統計的な意味をもっている点が重要である。

8. ま と め

本論文では画像の3次元解釈の信頼性に関する新しい理論体系を提案し，具体的な例についてその基本的な考え方を示した。まず画像処理によって検出したエッジへの直線当てはめの信頼性を画素レベルのノイズ特性により定量的に評価し，次にその直線の信頼性から消失点の計算の信頼性を定量的に評価した。そして，その結果を焦点距離のキャリブレーションへ応用し，推定値の信頼区間を評価した。さらに直交性当てはめの信頼性を理論的に評価し，最後にエッジ統合，消失点，出現点，消失線に関するアドホックな定数を含まない仮説検定の基準を統計学に基づいて導出した。

本研究の一部は文部省科学研究費一般研究C (No. 05808032) による。

参 考 文 献

- 1) Andress, K. M. and Kak, A. C.: Evidence Accumulation & Flow of Control in a Hierarchical Spatial Reasoning System, *AI Magazine*, Vol. 9, No. 2, pp. 75-94 (1988).
- 2) Brillault-O'Mahony, B.: New Method for Vanishing Point Detection, *CVGIP: Image Understanding*, Vol. 54, No. 2, pp. 289-300 (1991).
- 3) Brooks, R.A.: *Model-Based Computer Vision*, UMI Research Press, Ann Arbor, MI (1984).
- 4) Kanatani, K.: *Group-Theoretical Methods in Image Understanding*, Springer, Berlin, F.R.G. (1990).
- 5) 金谷健一: 画像理解/3次元認識の数理, 森北出版, 東京 (1990).
- 6) Kanatani, K.: Computational Projective Geometry, *CVGIP: Image Understanding*, Vol. 54, No. 3, pp. 333-348 (1991).
- 7) Kanatani, K.: Hypothesizing and Testing Geometric Properties of Image Data, *CVGIP: Image Understanding*, Vol. 54, No. 3, pp. 349-357 (1991).
- 8) 金谷健一: コニックの計算論と直交性解釈, 情報処理学会論文誌, Vol. 33, No. 6, pp. 817-824 (1992).
- 9) Kanatani, K.: *Geometric Computation for Machine Vision*, Oxford Univ. Press, Oxford, U.K. (1993).
- 10) Kanatani, K.: Statistical Analysis of Focal Length Calibration Using Vanishing Points, *IEEE Trans. Robotics Automation*, Vol. 8, No. 6, pp. 767-775 (1992).
- 11) 金谷健一, 川島孝夫: 画像データの幾何学的構造の仮説検定, 情報処理学会論文誌, Vol. 31, No. 6, pp. 879-886 (1990).
- 12) Kanatani, K. and Onodera, Y.: Anatomy of Camera Calibration Using Vanishing Points, *IEICE Trans.*, Vol. 74, No. 10, pp. 3369-3378 (1991).
- 13) 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 Vol. 9, No. 7, pp. 813-820 (1991).
- 14) 金谷健一, 浦沢康二: 幾何学的仮説の統計的検定, 情報処理学会研究報告, 92-CV-80 (1992. 11).
- 15) Lowe, D.G.: *Perceptual Organization and Visual Recognition*, Kluwer Academics, Boston, MA, U.S.A. (1985).
- 16) Magee, M. J. and Aggarwal, J. K.: Determining Vanishing Points from Perspective Images, *Comput. Vision Gr. Image Process.*, Vol. 26, pp. 256-267 (1984).
- 17) Mulgaonkar, P. G., Shapiro, L. G. and Haralick, R. M.: Shape from Perspective: A Rule-Based Approach, *Comput. Vision Gr. Image Process.*, Vol. 36, pp. 298-320 (1986).
- 18) 田中俊江, 川島孝夫, 金谷健一: 平行性仮説による多面体の3次元形状復元, 電子情報通信学会論文誌 D-II, Vol. J-72-D-II, No. 4, pp. 517-525 (1989).

- 19) 辻 弘行, 金谷健一: 3次元回転行列の最良推定とその応用, 情報処理学会研究報告, 92-CV-76 (1992.1).
- 20) 浦沢康二, 金谷健一: 幾何学的計算の統計解析 I: 基礎理論, 情報処理学会研究報告, 92-CV-77 (1992.3).
- 21) 浦沢康二, 金谷健一: 幾何学的計算の統計解析 II: エッジ, 消失点, 出現点, 情報処理学会研究報告, 92-CV-78 (1992.5).
- 22) 浦沢康二, 金谷健一: 焦点距離キャリブレーションの統計的信頼性, 情報処理学会研究報告, 92-CV-79 (1992.9).
- 23) Weiss, R. and Boldt, M.: Geometric Grouping Applied to Straight Lines, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, Miami Beach, FL, U. S. A., pp. 656-661 (June 1986).
- 24) Weiss, R., Nakatani, H. and Riseman, E. M.: An Error Analysis for Surface Orientation from Vanishing Points, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, pp. 1179-1185 (1990).

(平成4年12月14日受付)

(平成5年3月11日採録)



金谷 健一 (正会員)

1947年岡山県生. 1969~1970年米 Case Western Reserve 大学留学. 1972年東京大学工学部計数工学科卒業. 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助手. 1983年同助教授. 1988年同教授. 米国 Maryland 大学, デンマーク Copenhagen 大学, 英国 Oxford 大学客員研究員. 著書「線形代数」(講談社), “Group-Theoretical Methods in Image Understanding” (Springer-Verlag), 「画像理解」(森北出版), “Geometric Computation for Machine Vision”, (Oxford Univ. Press.) 1987年情報処理学会論文賞受賞.