

幾何学的補正問題の最適計算と精度の理論的境界

金 谷 健 一†

コンピュータビジョンやロボティクスに現れる幾何学補正問題を一般的に定式化し、補正したパラメータの不偏推定量の共分散行列の「クラメル・ラオの下界」を「フィッシャー情報行列」により特定の誤差分布を仮定することなく導出する。そして「指数型補正問題」ではパラメータの最尤推定量がその下界を第1近似において達成していることを示す。例としてステレオ視による3次元復元およびステレオデータの平面への逆投影問題に本理論を適用する。

Optimal Computation and Theoretical Accuracy Bound for Geometric Correction

KENICHI KANATANI†

The problem of geometric correction that appears in computer vision and robotics is formalized in general terms, and the *Cramer-Rao lower bound* on the covariance matrix of an unbiased estimator of the corrected parameter is derived in terms of the *Fisher information matrix* without assuming a particular noise distribution. Then, the maximum likelihood estimator is shown to attain that bound in the first order for a problem of the *exponential family*. We illustrate our theory by applying it to the problem of reconstructing 3-D structures by stereo vision and back projecting stereo data onto a planar surface.

1. はじめに

画像やセンサデータから環境モデルを構築することはコンピュータビジョンやロボティクスの中心課題のひとつであるが、画像やセンサから得られるデータには不確定さがあり、その対処が重要なテーマである。ひとつの方法が事前の知識に基づく「拘束」の利用である。たとえば固定した床または作業台表面などの既知の平面上にある特徴点をセンサで計測しても、計測した位置は誤差のため厳密にはその平面上にあるとは限らない。その場合は計測位置をその平面上へ投影して真の位置を推定するのが実際的である。このとき、次の疑問が生じる。

- 補正の精度を最大にするにはどう投影するのが最適か。
- どう補正してもある程度の誤差は避けられないが、精度の理論的な限界は何か。
- その限界を達成する方式が存在するか。そのためにはどういう条件が必要か。
- 最適な解を実用的な精度で効率良く計算する手法

が構成できるか。

本論文の目的はこれらの間に答えることである。そのために個々の問題の特殊性を捨象して一般化した「幾何学的補正問題」を考える。こうするとコンピュータビジョンやロボティクスの多くの問題が含まれることが分かる。代表的な例はステレオ視による3次元復元である。このときの拘束は「エッジ極線方程式」と呼ばれる。このような「誤差がなければ当然成立すべき条件」を持つ問題はすべて幾何学的補正問題として一般的に取り扱える。

従来は一般化すれば同一の解析が行える問題も個別の問題として研究されることが多かったが、今後の発展のためには体系化できる問題は体系化した学問としての整備が必要である^{3)~5),7),(8),10)}。本論文では個々の対象を考えるが、同様によく現れる問題として、不特定数の対象に別の対象あるいは関係をあてはめる問題がある。これは「あてはめ問題」として一般的に体系化されている⁷⁾。本論文でもそこに用いられた解析手法を用い、特定の誤差分布（たとえば正規分布）を仮定しない普遍的な結論を得る。

以下、2章で問題を一般的に記述し、3章で不偏推定量の「クラメル・ラオの下界」を導出する。4章で「指数型補正問題」では最尤推定量がその下界を第1

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

近似において達成することを示し、5章でその具体的計算法を示す。例として6章で「局所正規モデル」に、7章でステレオ視による3次元復元に、8章でステレオデータの平面への逆投影問題に本理論を適用する。

2. 幾何学的補正問題

引数 $\mathbf{u} \in \mathcal{R}^n$ の連続微分可能なスカラー関数 $F^{(k)}(\mathbf{u})$, $k = 1, \dots, L$ が与えられ (\mathcal{R}^n は n 次元実数空間を表す), 引数 \mathbf{u} の定義域 $\mathcal{U} \subset \mathcal{R}^n$ は n' 次元多様体であるとする。幾何学的補正問題とは L 個の方程式

$$F^{(k)}(\mathbf{u}) = 0, \quad k = 1, \dots, L \quad (1)$$

を満たすとは限らないデータ値 \mathbf{u} が与えられたとき、それを上式を満たすように補正する問題である。データ値 \mathbf{u} を未知の真の値 $\bar{\mathbf{u}}$ に未知の誤差が加わった確率変数と見なすと、これは次の問題として定式化できる。

【問題1】 拘束

$$F^{(k)}(\bar{\mathbf{u}}) = 0, \quad k = 1, \dots, L \quad (2)$$

を満たす $\bar{\mathbf{u}}$ をデータ値 \mathbf{u} から推定せよ。

データ値 \mathbf{u} の確率密度 $p(\mathbf{u}; \bar{\mathbf{u}})$ に関して次の仮定を置く。以下 $\nabla_{\bar{\mathbf{u}}} p = (\partial p / \partial \bar{u}_1, \dots, \partial p / \partial \bar{u}_n)^T$ と約束する (T は転置記号である)。

【仮定1】 確率密度 $p(\mathbf{u}; \bar{\mathbf{u}})$ は $\mathbf{u}, \bar{\mathbf{u}}$ に関して任意の回数だけ連続微分可能であり、すべての $\mathbf{u} \in \mathcal{U}$ に対して $p(\mathbf{u}; \bar{\mathbf{u}}) > 0$ である。 $p(\mathbf{u}; \bar{\mathbf{u}})$ の任意の式の \mathbf{u} に関する積分 $\int_{\mathcal{U}} d\mathbf{u}$ は (積分が存在すれば) $\bar{\mathbf{u}}$ に関する微分 $\nabla_{\bar{\mathbf{u}}}$ と交換できる。

$\bar{\mathbf{u}} \in \mathcal{U}$ であるから、 $\bar{\mathbf{u}} \notin \mathcal{U}$ に対しては $p(\mathbf{u}; \bar{\mathbf{u}})$ はどう定義してもよい。そこで、多様体 \mathcal{U} の $\bar{\mathbf{u}}$ における接空間を $T_{\bar{\mathbf{u}}}(\mathcal{U}) \subset \mathcal{R}^n$ とするとき、微小な任意の $\Delta \mathbf{u} \in T_{\bar{\mathbf{u}}}(\mathcal{U})^\perp$ の1次近似において $p(\mathbf{u}; \bar{\mathbf{u}} + \Delta \bar{\mathbf{u}}) = p(\mathbf{u}; \bar{\mathbf{u}})$ が成り立つように拡張する (記号 \perp は直交補空間を表す)。これは次の仮定と同値である。

【仮定2】 $\nabla_{\bar{\mathbf{u}}} p \in T_{\bar{\mathbf{u}}}(\mathcal{U})$ 。

拘束 (1) の L 個の方程式は \mathbf{u} の関数として代数的に従属でもよい。 r 個のみが独立であれば実質的には r 個の拘束しかなく、 \mathbf{u} は多様体 \mathcal{U} のある $n' - r$ 次元 (すなわち余次元 r) の部分多様体に拘束される。真の値 $\bar{\mathbf{u}} \in \mathcal{S}$ は \mathcal{S} の特異点ではないと仮定し、 r を式 (1) のランクと呼ぶ。厳密に述べると次のようになる。

【仮定3】 拘束 (1) は $\bar{\mathbf{u}}$ の近傍で余次元 r を持つ \mathcal{U} の部分多様体 \mathcal{S} を定義する。

以下 L 次元ベクトル $(v^{(1)}, \dots, v^{(L)})^T$ を $(v^{(k)})$ と略記する。データ値 \mathbf{u} を $\bar{\mathbf{u}} + \Delta \mathbf{u}$ と書けば、式 (2)

より $\Delta \mathbf{u}$ の1次近似において

$$F^{(k)}(\bar{\mathbf{u}} + \Delta \mathbf{u}) = (\nabla \bar{F}^{(1)}, \dots, \nabla \bar{F}^{(L)})^T \Delta \mathbf{u} \quad (3)$$

である。 $\nabla \bar{F}^{(k)}$ は $(\partial F^{(k)} / \partial u_1, \dots, \partial F^{(k)} / \partial u_n)^T$ の $\mathbf{u} = \bar{\mathbf{u}}$ での値を表し、 $(\nabla \bar{F}^{(1)}, \dots, \nabla \bar{F}^{(L)})$ は $\nabla \bar{F}^{(1)}, \dots, \nabla \bar{F}^{(L)}$ を順に列とする L 次元行列である。制約 $\mathbf{u} \in \mathcal{U}$ より1次近似において $\Delta \mathbf{u} \in T_{\bar{\mathbf{u}}}(\mathcal{U})$ であるから、式 (3) は $T_{\bar{\mathbf{u}}}(\mathcal{U})$ から L 次元空間 \mathcal{R}^L のある線形部分空間

$$\mathcal{L}_{\bar{\mathbf{u}}} = \{(\nabla \bar{F}^{(k)}, \Delta \mathbf{u}) \mid \Delta \mathbf{u} \in T_{\bar{\mathbf{u}}}(\mathcal{U})\} \quad (4)$$

への線形写像を定義する。 (\cdot, \cdot) はベクトルの内積である。接空間 $T_{\bar{\mathbf{u}}}(\mathcal{U}) \subset \mathcal{R}^n$ への射影行列を $\mathbf{P}_{\bar{\mathbf{u}}}^{\mathcal{U}}$ とすると、任意の $\Delta \epsilon \in \mathcal{R}^n$ に対して $\mathbf{P}_{\bar{\mathbf{u}}}^{\mathcal{U}} \Delta \epsilon \in T_{\bar{\mathbf{u}}}(\mathcal{U})$ であるから、部分空間 $\mathcal{L}_{\bar{\mathbf{u}}}$ の次元は $\{\mathbf{P}_{\bar{\mathbf{u}}}^{\mathcal{U}} \nabla \bar{F}^{(k)}\}$, $k = 1, \dots, L$ のランク (線形独立なもの数) に等しい。したがって $\mathcal{L}_{\bar{\mathbf{u}}}$ の次元はたかだか r であるが、次のことを仮定する。

【仮定4】 線形部分空間 $\mathcal{L}_{\bar{\mathbf{u}}}$ の次元は拘束のランク r に等しい。

これは数学的には式 (1) の L 個の方程式のそれぞれが \mathcal{U} の余次元1の部分多様体を定義し、それらが $\mathbf{u} = \bar{\mathbf{u}}$ において互いに横断的に交わっていることを表す (図1)。このため L 個の式 (1) をたとえばひとつの式 $\sum_{k=1}^L F^{(k)}(\mathbf{u})^2 = 0$ に置き換えることは許されない。こうするとランクは r であるが、 $\mathcal{L}_{\bar{\mathbf{u}}} = \{\mathbf{0}\}$ になってしまうからである。部分空間 $\mathcal{L}_{\bar{\mathbf{u}}}$ の次元が式 (1) のランク r より小さいとき $\bar{\mathbf{u}}$ を特異値と呼び、それ以外のものを通常値と呼ぶ。仮定4は真の値 $\bar{\mathbf{u}}$ が特異値ではないことを意味する。式 (4) の線形部分空間 $\mathcal{L}_{\bar{\mathbf{u}}}$ は L 次元空間中の \mathbf{u} によってパラメータ化された多様体 $\mathcal{F} = \{(F^{(k)}(\mathbf{u}) \in \mathcal{R}^L \mid \mathbf{u} \in \mathcal{U})$ の $\mathbf{u} = \bar{\mathbf{u}}$ における接空間としても定義できる。しかし問題によっては $\mathbf{u} = \bar{\mathbf{u}}$ が写像 $(F^{(k)}(\mathbf{u})) : \mathcal{U} \rightarrow \mathcal{R}^L$ の特異点になる

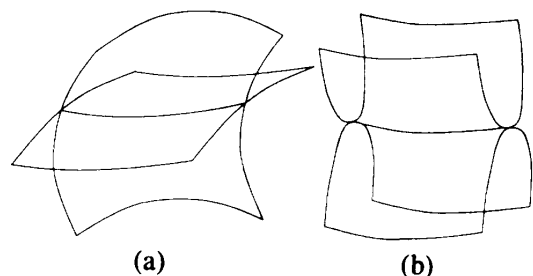


図1 (a) 横断的な交わり。 (b) 非横断的な交わり
Fig. 1 (a) Transversal intersection. (b) Non-transversal intersection.

ことがある. このとき式(4)の \bar{u} をデータ値 u で置き換えると, 同様に定義される部分空間 \mathcal{L}_u の次元は r より大きくなる. そのような問題は退化していると呼ぶ.

3. 不偏補正の精度の限界

n 次元確率変数 \bar{l} を

$$\bar{l} = \nabla_{\bar{u}} \log p \quad (5)$$

と定義し, スコアと呼ぶ⁸⁾. \bar{u} に関する (フィッシャー) 情報行列を次式で定義する. $E[\cdot]$ は同時密度分布 $p(u, \bar{u})$ に関する期待値である.

$$J = E[\bar{l}\bar{l}^T] \quad (6)$$

仮定2より $\bar{l} \in T_{\bar{u}}(U)$ であるが, u が U のすべてにわたるとき \bar{l} は接空間 $T_{\bar{u}}(U)$ のあらゆる方向をとり得ると仮定する. これは次の仮定と等価である.

【仮定5】 情報行列 J は $T_{\bar{u}}(U)$ を値域とする n 次元半正値対称行列である.

(ム-ア・ベンローズ) 一般逆行列を $(\cdot)^-$ で表す. (kl) 要素が $\bar{W}^{(kl)}$ である L 次元行列 \bar{W} を $(\bar{W}^{(kl)})$ と略記し,

$$(\bar{W}^{(kl)}) = ((\nabla \bar{F}^{(k)}, J^- \nabla \bar{F}^{(l)}))^- \quad (7)$$

と定義する. 右辺は (kl) 要素が $(\nabla \bar{F}^{(k)}, J^- \nabla \bar{F}^{(l)})$ の L 次元行列の一般逆行列の略記である. n 次元行列 N, S を

$$N = \sum_{k,l=1}^L \bar{W}^{(kl)} (J^- \nabla \bar{F}^{(k)}) (J^- \nabla \bar{F}^{(l)})^T \quad (8)$$

$$S = J^- - N \quad (9)$$

と定義する.

真の値 \bar{u} のある推定量を $\hat{u} = \hat{u}(u)$ とする. これは \bar{u} の満たすべき制約を満たしていることを要求する.

【仮定6】 任意の $u \in U$ に対して $\hat{u}(u) \in S$.

推定量は一般にはその期待値が真の値に一致するとき「不偏」であるというが, これを \hat{u} にあてはめることはできない. 多様体 S が「曲がってれば」 \hat{u} の期待値が S にあるとは限らないからである. そこで不偏性を次式で定義し, \hat{u} がこれを満たす不偏推定量であるとする.

$$P_{\bar{u}}^S E[\hat{u} - \bar{u}] = 0 \quad (10)$$

$P_{\bar{u}}^S$ は接空間 $T_{\bar{u}}(S)$ への射影行列である (図2). \hat{u} の分布は多様体 S 上に十分局在し, $E[\hat{u} - \bar{u}]$ は十分小さいとする. そして推定量 \hat{u} の共分散行列を次の行列と同一視する.

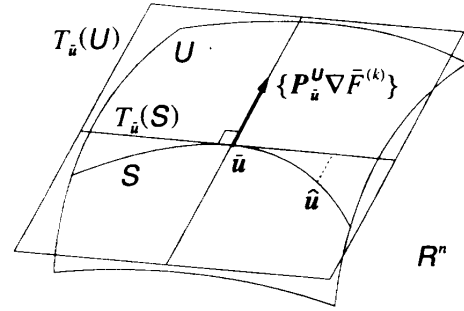


図2 部分多様体 S の接空間 $T_{\bar{u}}(S)$ とその $T_{\bar{u}}(U)$ における直交補空間. 推定量 \hat{u} は S に拘束されている.

Fig. 2 Tangent space $T_{\bar{u}}(S)$ to submanifold S and its orthogonal complement in $T_{\bar{u}}(U)$. The estimator \hat{u} is constrained to be in S .

$$V[\hat{u}] = P_{\bar{u}}^S E[(\hat{u} - \bar{u})(\hat{u} - \bar{u})^T] P_{\bar{u}}^S \quad (11)$$

対称行列 A, B に対して $u \succeq B$ とは $A - B$ が半正値であることと約束すると, 次の定理が成立する⁹⁾.

【定理1】 $V[\hat{u}] \succeq S$.

(証明の道筋) S が $T_{\bar{u}}(S)$ を値域とする n 次元半正値対称行列であることは S の定義から容易に証明できる. \bar{u} の U 上の任意の変分に対して式(10)が不変であることから変分原理により $E[P_{\bar{u}}^S(\hat{u} - \bar{u})(S\bar{l})^T] = S$ が得られる. これから $V[\hat{u}], S$ を対角ブロックに, S を非対角ブロックに持つ $2n$ 次元対称行列が半正値であることが証明できる. この行列は $V[\hat{u}] - S, S$ からなる対角ブロック行列に合同変換されることから, $V[\hat{u}] - S$ が半正値であることが分かる. □

問題1は母数 \bar{u} を持つ分布のサンプル u から \bar{u} を推定する問題でもあり, 定理1はクラメル・ラオの不等式に相当するので⁸⁾, S をクラメル・ラオの下界と呼ぶ.

4. 最尤推定による補正

補正問題1の最尤推定とはデータ値 u が与えられたとき, それを確率密度に代入した $p(u; \bar{u})$ を \bar{u} の関数と見た尤度を最大にする \bar{u} を選ぶことである. その解 \hat{u} を \bar{u} の最尤推定量と呼ぶ. 「変数」と特定の「値」を区別するために \bar{u} を真の値とし, それを未知変数と見たものを x と書く. 最尤推定量を求めるには, 尤度の対数をとって符号を変えたものの $x \in U$ に関する最小化

$$J = -\log p(u; x) \rightarrow \min \quad (12)$$

を次の拘束のもとで実行すればよい.

$$F^{(k)}(x) = 0 \quad (13)$$

n 次元行列 \bar{L} を次式で定義する.

$$\bar{L} = -\nabla_{\mathbf{x}}^2 \log p|\mathbf{x}=\bar{\mathbf{u}} \quad (14)$$

$\nabla_{\mathbf{x}}^2 f$ はヘッセ行列 ((ij) 要素が $\partial^2 f/\partial x_i \partial x_j$ の行列) である. $\mathbf{x} = \bar{\mathbf{u}} + \Delta\mathbf{x}$ と置き, 最尤推定量 $\hat{\mathbf{u}}$ は真の値 $\bar{\mathbf{u}}$ の近くにあると仮定して式 (12) の J を真の値 $\bar{\mathbf{u}}$ の近傍で展開する.

$$J = \bar{c} - (\bar{L}, \Delta\mathbf{x}) + \frac{1}{2}(\Delta\mathbf{x}, \bar{L}\Delta\mathbf{x}) + \dots \quad (15)$$

\bar{c} は J の $\mathbf{x} = \bar{\mathbf{u}}$ での値であり, \dots は $\Delta\mathbf{x}$ の 3 次以上の項を表す. $\bar{\mathbf{u}}$ の近傍で J を最小にする \mathbf{x} の存在と一意性を保証するために次の仮定を置く.

【仮定 7】 任意の $\mathbf{u} \in \mathcal{U}$ に対して \bar{L} は接空間 $T_{\bar{\mathbf{u}}}(\mathcal{U})$ を値域とする半正値対称行列である.

式 (13) から $\Delta\mathbf{x}$ の拘束が 1 次近似において

$$(\nabla \bar{F}^{(k)}, \Delta\mathbf{x}) = 0 \quad (16)$$

と書ける. 制約 $\mathbf{x} \in \mathcal{U}$ より $\Delta\mathbf{x}$ は 1 次近似において制約 $\Delta\mathbf{x} \in T_{\bar{\mathbf{u}}}(\mathcal{U})$ を受ける. 拘束 (16) のもとで式 (15) の J を最小にする $\Delta\mathbf{x}$ が一意的に定まり, \mathbf{u} の最尤推定量 $\hat{\mathbf{u}} = \bar{\mathbf{u}} + \Delta\mathbf{x}$ が次式で表せる⁹⁾.

$$\hat{\mathbf{u}} = \bar{\mathbf{u}} + \mathbf{S}'\bar{\mathbf{l}} \quad (17)$$

\mathbf{S}' は次式で定義した n 次元行列である.

$$\mathbf{S}' = \bar{L}^{-1} - \mathbf{N}' \quad (18)$$

$$\mathbf{N}' = \sum_{k,l=1}^L \bar{W}'^{(kl)} (\bar{L}^{-1} \nabla \bar{F}^{(k)}) (\bar{L}^{-1} \nabla \bar{F}^{(l)})^T \quad (19)$$

$$(\bar{W}'^{(kl)}) = ((\nabla \bar{F}^{(k)}, \bar{L}^{-1} \nabla \bar{F}^{(l)})^{-1}) \quad (20)$$

定義より $\mathbf{N}'\bar{L}\mathbf{N}' = \mathbf{N}'$, $\mathbf{S}'\bar{L}\mathbf{S}' = \mathbf{S}'$ であることが導かれ, \mathbf{S}' は $T_{\bar{\mathbf{u}}}(\mathcal{S})$ を値域とする n 次元半正値対称行列であることが示される⁹⁾. したがって最尤推定量 $\hat{\mathbf{u}}$ の共分散行列が

$$V[\hat{\mathbf{u}}] = E[\mathbf{S}'\bar{\mathbf{u}}\bar{\mathbf{u}}^T\mathbf{S}'] \quad (21)$$

と表せる. 次の特殊な場合を考える.

【仮定 8】 \bar{L} はデータ値 \mathbf{u} に依存しない.

これは確率密度 $p(\mathbf{u}; \bar{\mathbf{u}})$ が次の形に表されることを意味する.

$$p(\mathbf{u}; \bar{\mathbf{u}}) = C(\bar{\mathbf{u}}) \exp[(\mathbf{f}(\mathbf{u}), \bar{\mathbf{u}}) + g(\mathbf{u})] \quad (22)$$

$\mathbf{f}(\mathbf{u})$ は n 次元関数, $C(\mathbf{u})$, $g(\mathbf{u})$ はスカラ関数である. これは指数型分布と呼ばれ, 通常現れる分布はほとんどこの型である. そのような分布に対する問題 1 を指数型補正問題と呼ぶ. このとき $\bar{L} = \mathbf{J}$ であるから $\mathbf{S}' = \mathbf{S}$ となり, 式 (17) より次の命題を得る.

【命題 1】 指数型補正問題の最尤推定量 $\hat{\mathbf{u}}$ は第 1 近

似において不偏推定量である.

クラメル・ラオの下界を達成する推定量を有効推定量と呼ぶ. 式 (21) 中の \mathbf{S}' に式 (9) を代入して変形すると右辺は \mathbf{S} となることが分かる⁹⁾. したがって次の命題を得る.

【命題 2】 指数型補正問題の最尤推定量 $\hat{\mathbf{u}}$ は第 1 近似において有効推定量である.

式 (17) は関数 J を引数の真の値の周りで 2 次関数に近似し, 式 (13) を線形近似して得られる. 命題 1, 2 に「第 1 近似において」とあるのはこの意味であり, 誤差が小さいときの第 1 近似としての表現である.

5. 最尤推定量の計算

最尤推定量 $\hat{\mathbf{u}}$ を厳密に計算するには非線形最適化の数値探索が必要である. 以下, 最尤推定量 $\hat{\mathbf{u}}$ の良い近似値を簡単な計算で求める公式を導出する. まず, 拘束がなければ最尤推定量は \mathbf{u} 自身であるとする. すなわち次の仮定を置く.

【仮定 9】 $\nabla \mathbf{x} p(\mathbf{u}; \mathbf{x})|_{\mathbf{x}=\mathbf{u}} = \mathbf{0}$

最小化 (12) の解はデータ値 \mathbf{u} の近くにあると仮定して $\mathbf{x} = \mathbf{u} + \Delta\mathbf{x}$ を式 (12) の J に代入し, \mathbf{u} の近傍で展開する.

$$J = c + \frac{1}{2}(\Delta\mathbf{x}, \mathbf{L}\Delta\mathbf{x}) + \dots \quad (23)$$

c は J の $\mathbf{x} = \mathbf{u}$ での値であり, \mathbf{L} は次の m 次元行列である.

$$\mathbf{L} = -\nabla_{\mathbf{x}}^2 \log p|\mathbf{x}=\mathbf{u} \quad (24)$$

\mathbf{u} の近傍で J を最小にする \mathbf{x} の存在と一意性を保証するために次の仮定を置く.

【仮定 10】 任意の $\mathbf{u} \in \mathcal{U}$ に対して \mathbf{L} は接空間 $T_{\mathbf{u}}(\mathcal{U})$ を値域とする半正値対称行列である.

式 (13) は 1 次近似において

$$(\nabla F^{(k)}, \Delta\mathbf{x}) = -F^{(k)} \quad (25)$$

となる. $F^{(k)}$, $\nabla F^{(k)}$ はそれぞれ $F^{(k)}(\mathbf{u})$, $(\partial F^{(k)}/\partial u_1, \dots, \partial F^{(k)}/\partial u_n)^T$ の略である. 制約 $\mathbf{u} \in \mathcal{U}$ より, $\Delta\mathbf{x}$ は 1 次近似において制約 $\Delta\mathbf{x} \in T_{\mathbf{u}}(\mathcal{U})$ を受ける. 拘束 (25) に対するラグランジュ乗数を $\lambda^{(k)}$ として高次の項を無視すれば, 仮定 7 より式 (23) の J を最小にする $\Delta\mathbf{x}$ が次のように定まる.

$$\Delta\mathbf{x} = \sum_{k=1}^L \lambda^{(k)} \mathbf{L}^{-1} \nabla F^{(k)} \quad (26)$$

$\lambda = (\lambda^{(k)})$ と置く. L 次元行列 $\mathbf{V} = (V^{(kl)})$ を

$$V^{(kl)} = ((\nabla F^{(k)}, \mathbf{L}^{-1} \nabla F^{(l)})) \quad (27)$$

と定義し、式 (26) を式 (25) に代入して次式を得る。

$$V\lambda = -(F^{(k)}) \tag{28}$$

ここで計算上の問題が生じる。式 (13) の L 個の方程式が代数的に独立ではなく r 個だけが独立であっても、データ値で定義した式 (27) の行列 V のランクは r より大きくなることがある。そのような問題は退化していると呼ぶ。このとき $u \rightarrow \hat{u}$ の極限で V の小さい正の固有値のいくつかが 0 に収束するので、誤差が小さいと方程式 (28) が悪条件となる。これを避けるために式 (28) の両辺を V の大きい r 個の固有値に対応する固有空間に射影し、式 (25) の L 個の方程式のうちから実質的に r 個のみを用いる。式 (25) は式 (13) の 1 次近似であり、1 次近似の範囲では差は無視できる。そこで

$$(W^{(kl)}) = ((\nabla F^{(k)}, L^{-1}\nabla F^{(l)})_r)^{-} \tag{29}$$

と置く。 $(\cdot)_r^{-}$ はランクを r に強制した (大きい順に r 個の固有値のみを残し、残りの固有値を 0 に置き換えて計算した) 一般逆行列を表す。これから次の命題を得る。

【命題 3】 u の最尤推定量 \hat{u} は第 1 近似において次式で与えられる。

$$\hat{u} = u - L^{-1} \sum_{k,l=1}^L W^{(kl)} F^{(k)} \nabla F^{(l)} \tag{30}$$

式 (26) から分かるように補正量 Δx は多様体 U のデータ値 u における接空間 $T_u(U)$ の元であるから (図 3)、式 (30) の補正で \hat{u} が厳密に $\hat{u} \in U$ であるとは限らない。そのときは \hat{u} に高次の補正を加えて多様体 U 上に引き戻す。引き戻した点が厳密に部分多様体 S 上になれば、その点をデータ値と見なし

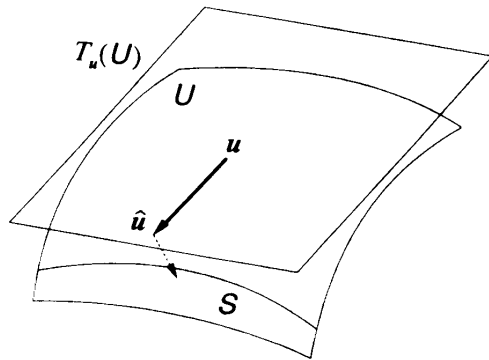


図 3 拘束を完全に満たすように第 1 近似の補正にさらに高次の補正を加える。

Fig. 3 Higher order correction is added to the first order correction so that the constrained is satisfied exactly.

て補正を反復する¹¹⁾。

6. 局所正規モデル

取扱いが容易かつ現実的なモデルはデータ値 u を正規分布に従う確率変数と見なすことであるが、正規分布は無限遠にまで正の確率を持つものに対して、ここでは u が \mathcal{R}^n の n' 次元多様体 U に限定されている。そこで誤差は小さいと仮定し、各 $u \in U$ の分布は $\hat{u} \in U$ の近傍に集中していて接空間 $T_{\hat{u}}(U)$ での正規分布と同一視できるとする。このような分布を局所正規分布と呼ぶ (図 4)。このとき確率密度は次式で表せる。

$$p(u; \hat{u}) = C e^{-(u-\hat{u}, V[u]^{-1}(u-\hat{u})) / 2} \tag{31}$$

C は正規化定数であり、 u の共分散行列 $V[u]$ は次式で定義される。

$$V[u] = P_u^S E[(u - \hat{u})(u - \hat{u})^T] P_u^S \tag{32}$$

その値域は $T_{\hat{u}}(S)$ であるとする。文中で導入した仮定は満たされ、この補正問題は指数型である。そして式 (31) より $L = \bar{L} = J = V[u]^{-1}$ であり、式 (12) の最小化は次の最小化と等価である。

$$(x - u, V[u]^{-1}(x - u)) \rightarrow \min \tag{33}$$

これはデータ位置 u を式 (13) の定める多様体 S 上で u からのマハラノビス距離 ($V[u]$ の一般逆行列を計量とする距離) が最小の点に補正することに相当する。このとき式 (27), (29), (30) より最適な補正は次式で表せる。

$$\hat{u} = u - V[u] \sum_{k,l=1}^L W^{(kl)} F^{(k)} \nabla F^{(l)} \tag{34}$$

$$(W^{(kl)}) = ((\nabla F^{(k)}, V[u] \nabla F^{(l)})_r)^{-} \tag{35}$$

推定量 \hat{u} の共分散行列は式 (21) と同様の計算により次式となる¹¹⁾。

$$V[\hat{u}] = \hat{V}[\hat{u}] - \sum_{k,l=1}^L \hat{W}^{(kl)} (\hat{V}[\hat{u}] \nabla \hat{F}^{(k)}) (\hat{V}[\hat{u}] \nabla \hat{F}^{(l)})^T \tag{36}$$

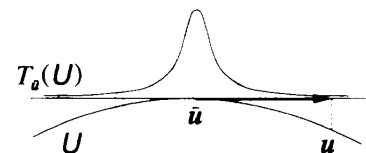


図 4 局所正規分布
Fig. 4 Local Gaussian distribution.

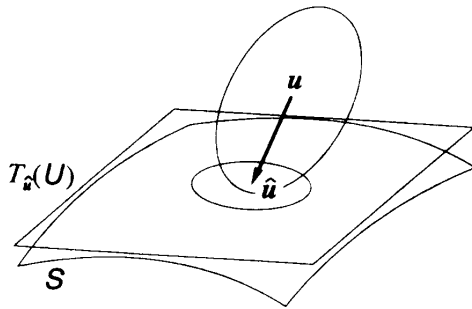


図5 データ値 u はそれを中心とする楕円体と S との接点に補正される。

Fig. 5 Datum u is corrected to the tangent point of an ellipsoid centered at u to S .

$\nabla \hat{F}^{(k)}$ は $\nabla F^{(k)}$ を \hat{u} で評価して得られる値であり、 $\hat{V}[u]$ は $V[u]$ を値域が $T_{\hat{u}}(U)$ となるように射影した行列である¹¹⁾。 $\hat{W}^{(kl)}$ は式 (35) で $V[u]$, $\nabla F^{(k)}$ をそれぞれ $\hat{V}[u]$, $\nabla \hat{F}^{(k)}$ で置き換えた値である。変数 u に拘束がない場合を考えると、式 (33) の等マハラノビス距離面は n 次元空間の楕円体を表す。式 (33) は u がそれを中心とする楕円体と S との接点に補正されることを意味する (図5)。

7. ステレオ視の最適補正

カメラのレンズの中心を原点とし、光軸を Z 軸とする XYZ 座標系をとる。原点 O と撮像面との距離を長さの単位とすれば、撮像は画像面 $Z=1$ への透視変換と見なせる。画像面上で光軸の位置を画像原点とし、 X 軸、 Y 軸に平行な xy 画像座標軸をとると、画像座標 (x, y) の点は XYZ 座標系に関してベクトル $\mathbf{x} = (x, y, 1)^T$ で表される。カメラを XYZ 座標系に関して原点 O の周りに行列 \mathbf{R} だけ回転し、ベクトル \mathbf{h} だけ平行移動した位置に第2のカメラを置いたステレオ視を考える。 $\{\mathbf{h}, \mathbf{R}\}$ を運動パラメータと呼ぶ。第2のカメラに関する $X'Y'Z'$ 座標をとり、同様に $x'y'$ 画像座標系を定義すると、画像座標 (x', y') の点は $X'Y'Z'$ 座標系に関してベクトル $\mathbf{x}' = (x', y', 1)^T$ で表される。 $x = \bar{x} + \Delta x$, $y = \bar{y} + \Delta y$ と書くと、誤差 $(\Delta x, \Delta y)$ が平均 $(0, 0)$ 、分散・共分散が $\sigma_x^2, \sigma_y^2, \gamma_{xy}$ の正規分布に従えば、 \mathbf{x} の共分散行列は次式となる。

$$V[\mathbf{x}] = V[\mathbf{x}'] = \begin{pmatrix} \sigma_x^2 & \gamma_{xy} & 0 \\ \gamma_{xy} & \sigma_y^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (37)$$

ある特徴点が第1, 第2のカメラ座標系に関してそれぞれ \mathbf{x}, \mathbf{x}' で表される位置に観測されるとする。基本行列を $\mathbf{G} = \mathbf{h} \times \mathbf{R}$ (右辺は \mathbf{h} と \mathbf{R} の各列との外積

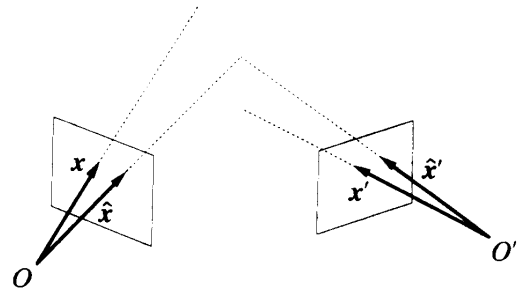


図6 視線が交わるように補正する。

Fig. 6 Corresponding points are corrected so that their lines of sight meet in the scene.

を列とする3次元行列を表す)と定義すると、画像に誤差がなければエビ極線方程式^{1)~3),15),17)}

$$(\mathbf{x}, \mathbf{G}\mathbf{x}') = 0 \quad (38)$$

が成立する。画像に誤差があり、対応する特徴点の位置を画像間の局所的相関から計算したとすると、 \mathbf{x}, \mathbf{x}' はエビ極線方程式を満たすとは限らない。言い換えると \mathbf{x}, \mathbf{x}' の定める「視線」が空間中で交わらない (図6)。特徴点の3次元位置を定めるためには視線が交わるように \mathbf{x}, \mathbf{x}' を補正しなければならないが、これは次の補正問題となる。

【問題2】 拘束

$$(\bar{\mathbf{x}}, \mathbf{G}\bar{\mathbf{x}}') = 0 \quad (39)$$

を満たす $\bar{\mathbf{x}}, \bar{\mathbf{x}}'$ をデータ \mathbf{x}, \mathbf{x}' から推定せよ。

これは \mathbf{x}, \mathbf{x}' の直和 $\mathbf{u} = \mathbf{x} \oplus \mathbf{x}'$ を考えると問題1に帰着し、 \mathbf{x}, \mathbf{x}' の最適補正は式 (34) より次式となる。

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{x} - \frac{(\mathbf{x}, \mathbf{G}\mathbf{x}')V[\mathbf{x}]\mathbf{G}\mathbf{x}'}{(\mathbf{x}', \mathbf{G}^T V[\mathbf{x}]\mathbf{G}\mathbf{x}') + (\mathbf{x}, \mathbf{G}V[\mathbf{x}]\mathbf{G}^T \mathbf{x})} \\ \hat{\mathbf{x}}' &= \mathbf{x}' - \frac{(\mathbf{x}, \mathbf{G}\mathbf{x}')V[\mathbf{x}']\mathbf{G}^T \mathbf{x}}{(\mathbf{x}', \mathbf{G}^T V[\mathbf{x}]\mathbf{G}\mathbf{x}') + (\mathbf{x}, \mathbf{G}V[\mathbf{x}]\mathbf{G}^T \mathbf{x})} \end{aligned} \quad (40)$$

その結果 $\hat{\mathbf{x}}, \hat{\mathbf{x}}'$ の定める視線は一点 \mathbf{r} で交わり、その位置が計算できるとともに、共分散行列 $V[\mathbf{r}]$ も計算できる¹⁴⁾。

8. 平面への最適逆投影

ステレオ視によれば3次元位置が一意的に復元できる。一方、復元した点がある既知の平面上にあるという拘束があれば、画像上のその点の定める視線をその平面との交点に逆投影すればよいから、1枚の画像で十分である。しかし、ステレオ画像が得られているときは、この拘束を生かしてより信頼性の高い復元ができる。これは次のように定式化できる。平面の単位法線ベクトルを第1のカメラに関する XYZ 座標系に

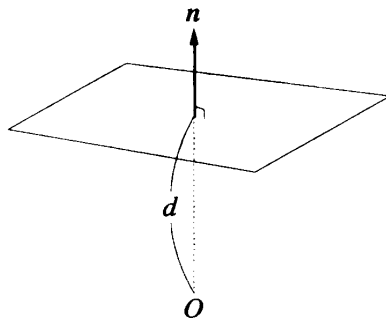


図7 平面パラメータ
Fig. 7 Surface parameters.

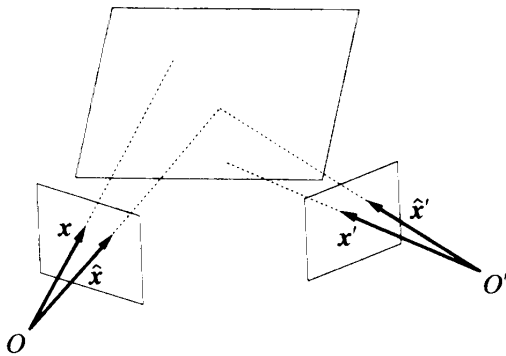


図8 視線が平面上で交わるように補正する。
Fig. 8 Corresponding points are corrected so that their lines of sight meet on the plane.

関して n とし、原点 O から平面までの距離 (n の方向に正と約束する) を d とする。 $\{n, d\}$ を平面パラメータと呼ぶ (図7)。第1, 第2のカメラ座標系についてそれぞれ x, x' で表される点がこの平面上にある必要十分条件は

$$x' \times Ax = 0, \quad A = R^T(hn^T - dI) \quad (41)$$

となる。しかし画像に誤差があれば式 (41) は厳密には成立しない。式 (41) は x, x' の定める視線がちょうどその平面上で交わる必要十分条件であり、エッジ極線方程式 (38) よりも強い条件である。そこで視線がその平面上で交わるように x, x' を補正する (図8)。これは次の補正問題となる。

【問題3】 拘束

$$\bar{x}' \times A\bar{x} = 0 \quad (42)$$

を満たす \bar{x}, \bar{x}' をデータ x, x' から推定せよ。

前節と同じ条件のもとで最適補正は次式となる。

$$\begin{aligned} \hat{x} &= x - (x' \times AV[x])^T W(x' \times Ax) \\ \hat{x}' &= x' + ((Ax) \times V[x'])^T W(x' \times Ax) \end{aligned} \quad (43)$$

W は次式で定義する (ベクトル v と行列 T に対して $v \times T \times v$ は $(v \times T)(v \times I)^T$ の略記である)。

$$W = \left(x' \times AV[x]A^T \times x' + (Ax) \times V[x'] \times (Ax) \right)_2^{-1} \quad (44)$$

この結果、復元点の3次元位置 r が計算できるとともに、共分散行列 $V[r]$ も計算できる¹¹⁾。

9. まとめ

7, 8章では本論文の一般論の意味を示す例として理論を機械的に適用したが、装置や処理の特性を反映した誤差分布を調べるなど、実際問題としての課題は多く残されており、本理論を骨組みとして問題独自の要因を考慮すればより有益な結果が得られる。また、8章の応用で平面の位置を既知とするのは非現実的に見えるが、これはそれ自体として用いるのではなく、種々の推定問題と組み合わせることによって威力を発揮する。現実の環境では壁や床や天井のように平面部分が多いから、平面の復元は実際問題として重要である。ステレオ視を用いると平面上の特徴点を観測すれば各点の3次元位置が一意的に定まるが、それに平面をあてはめるのではなく、画像データから平面が直接に推定できる¹³⁾。またカメラを任意に移動させながら平面物体を撮影しても、得られた動画像から平面とカメラの運動パラメータとが推定できる¹²⁾。平面の位置が分かると、観測した特徴点を本論文の方法でその平面に逆投影して3次元モデルが構成される。このときの平面の推定は、その逆投影によって信頼性が最大になるように構成したものである。

この問題を抽象化するとさらに応用が広がる。たとえば一般形状の物体を撮影した動画像から得られる点対応あるいはオプティカルフローからカメラの運動と物体の形状が復元できるが、誤差を考慮した統計的に最適解は複雑な非線形最適化問題の解となる。このときパラメータのある関数を新しい変数と見なせば問題が単純化される。これは線形化手法と呼ばれているが^{2),3)}、変換した変数ももとのパラメータに分解できるための分解可能条件を満たすとは限らない。従来は分解可能条件を満たす適当な近似値に置き換えてから分解が行われていたが、その結果は精度の低い解となるので、線形化手法は信頼性がないとされることが多かった^{3),17)}。しかし、線形化手法で計算した値の共分散行列に基づいて分解可能条件を満たすように最適に補正すれば、線形化手法の精度は直接的最適化と同程度であることが分かる^{6),10),16)}。

このように画像やセンサのデータから3次元環境モデルを構築する広い範囲の問題に対して、本理論に

よって解の最適な推定やその信頼性を定量的に評価するための理論的基礎が与えられる。

参考文献

- 1) Faugeras, O.: *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA (1993).
- 2) 金谷健一: 画像理解—3次元認識の数理—, 森北出版 (1990).
- 3) Kanatani, K.: *Geometric Computation for Machine Vision*, Oxford University Press, Oxford (1993).
- 4) 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).
- 5) 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 (1994).
- 6) Kanatani, K.: Renormalization for Motion Analysis: Statistically Optimal Algorithm, *IEICE Trans. Infor. Sys.*, Vol.E77-D, No.11, pp.1233-1239 (1994).
- 7) 金谷健一: 当てはめ問題の最適推定と精度の理論的限界, 情報処理学会論文誌, Vol.36, No.8, pp.1865-1873 (1995).
- 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) 金谷健一: 幾何学的補正問題の精度の理論限界と最適計算法, 情報処理学会研究報告, 95-CV-94-5 (1995-5).
- 10) 金谷健一: 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, 情報処理, Vol.37, No.1, pp.52-60 (1996).
- 11) Kanatani, K.: *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam (1996).
- 12) Kanatani, K. and Takeda, S.: 3-D Motion Analysis of a Planar Surface by Renormalization, *IEICE Trans. Inf. & Syst.*, Vol.E78-D, pp.1074-1075 (1995).
- 13) 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).
- 14) 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).
- 15) Maybank, S.: *Theory of Reconstruction from Image Motion*, Springer, Berlin (1993).
- 16) Ohta, N. and Kanatani, K.: Optimal Structure-from-Motion Algorithm for Optical Flow, *IEICE Trans. Inf. & Syst.*, Vol.E78-D, No.12, pp.1559-1566 (1995).
- 17) Weng, J., Huang, T.S. and Ahuja, N.: *Motion and Structure from Image Sequences*, Springer, Berlin (1993).

(平成7年8月1日受付)

(平成8年1月10日採録)

金谷 健一 (正会員)



1947年岡山県生。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。現在群馬大学工学部情報工学科教授。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, 1987), “Group-Theoretical Methods in Image Understanding” (Springer, 1990), 「画像理解」(森北出版, 1990), “Geometric Computation for Machine Vision” (Oxford University Press, 1993), 「空間データの数理」(朝倉書店, 1995), “Statistical Optimization for Geometric Computation” (Elsevier Science, 1996)。1987年情報処理学会論文賞受賞。