ミーンシフトの原理と応用

岡田和典

サンフランシスコ州立大学 計算機科学科
E-mail: kazokada@sfsu.edu

あらまし ミーンシフトは、カーネル密度推定を用いるロバストなデータ解析手法で、福永らによって提唱されたのち最近の Cheng や Comaniciu の定式による広範なビジョン問題への応用成功例が知られている。本稿では、ミーンシフト法の基本原理およびその一般的な特徴と利点を順を追って解説した上で、最近の理論的拡張および画像の領域分割やビデオ上の物体追跡などの実際のビジョン応用例について概説する。

キーワード ミーンシフト, カーネル密度推定, 最頻値探索, クラスタリング, 領域分割, 物体追跡

Mean Shift: from Theory to Application

Kazunori Okada

Computer Science Department, San Francisco State University
1600 Holloway Ave. San Francisco, CA 94132, USA
E-mail: kazokada@sfsu.edu

Abstract Mean shift is a popular robust framework for statistical data analysis using kernel density estimation, originally proposed by Fukunaga and Hostetler in 70's. Recently, due to the work by Cheng and Comaniciu, this method has been re-discovered and successfully applied to a wide range of vision applications. This article provides a comprehensive overview of the basic theory and applications of mean shift, highlighting its practical and theoretical advantages, recent theoretical extensions, as well as vision applications such as image segmentation and object tracking.

Key words Mean shift, Kernel density estimation, Mode seeking, Clustering, Segmentation, Tracking

1. はじめに

ミーンシフト (mean shift) は、カーネル密度推定 [45] を用いるロバストなデータ解析手法で、福永と Hostetler [22] によって '75 年に提唱されたのち長く忘れられていたが、最近になって '95 年に発表された Cheng の論文 [7] を皮切りに注目さ

れ、特に Comaniciu らによる一般的で精確な定式化 [12], [13] を使った、広範なビジョン問題への応用成功例で知られている。この手法は、密度関数の極値探索問題について最急勾配法よりロバストで効率的な解法を与え、また一般的なクラスタ解析、空間最尤推定、また情報統合を頑強化する基礎技術でもある。この手法の応用としては、画像のセ

グメンテーションとエッジ保存の画像の平滑化 [4], [12], [13], [20], [25], [38], [46], [47]、ビデオ上の物体トラッキング [1], [10], [11], [14], [16], [44], [55], [57], [58]、クラスタ解析 [7], [21]~[23], [41]、医用画像解析 [16], [29], [33], [35], [36], [47]~[49], [57] テクスチャ解析 [23], [27]、モーション解析 [17]、ハフ変換・情報統合 [4], [7], [11], [61]、物体認識 [51]、マルチメディア [41]、コンピュータグラフィックス [42]、インテリジェント交通システム [55], [61] などへの応用例が報告されている。また理論的にも活気のある研究がされており、最近にも様々な理論的考察や新規な拡張が盛んに発表されている (カーネル回帰分析とロバスト M 推定との関連 [13]、界関数最適化との関連 [19], [46]、EM アルゴリズムとの関連 [5], [6]、 k 平均クラスタ解析との関連 [7]、half-quadratic 最適化との関連 [59]、可変カーネル幅への拡張 [9], [11], [15]、スケールスペースへの拡張 [8], [32]、事後確率推定への拡張 [37]、動的ミーンシフトへの拡張 [3], [24], [60]、Grassmann manifold への拡張 [47])。本稿はこのミーンシフト法のチュートリアル参考資料として、その理論と応用の両点からの包括的な解説を目的とする。以下、ミーンシフト法の基本原理と理論およびその一般的な特徴と利点を順を追って説明した上で、最近の理論的拡張および実際の応用例の主なものについて概説する。

2. ミーンシフトの原理

2.1 最頻値探索問題

最頻値探索問題 (mode seeking problem) は、データ解析における重要な定式化手法のひとつで、ビジョン・画像処理や機械学習などにおける多くの問題にも適用されている。本稿で紹介するミーンシフト (mean shift) は、この最頻値探索問題の効率的でロバストな解法である。

d 次元空間 $X \subset \mathbb{R}^d$ 上におけるベクトル点の集合 $S = \{\mathbf{x}_i | i = 1, \dots, n\}$ が観測や実験結果のデータまたは統計的サンプルとして与えられたとする。これらの点の集まりがどのように分布しているかを示す関数 $f(\mathbf{x} \in X)$ は、一般に密度関数 (density

function) と呼ばれる。この密度関数は空間上の任意の位置 $\mathbf{x}$ におけるサンプル点の密度を表し、関数値が高いところの周辺にはサンプル点がたくさん集まり、低いところにはサンプル点が少ないことを示す。ここでいう最頻値 (mode) は、この密度関数の極大値 (局所的な最大値) として定義され、対応する極大値点は空間上で点の密度が局所的に最も高いところ、つまり点がローカルに最も凝集している位置を指し、偏微分方程式 $\nabla_{\mathbf{x}} f(\mathbf{x}) = 0$ の解に対応する。統計学的には、密度関数を離散化したものは度数分布として、密度関数の極大値と極大値点はこの度数分布のピークにおける度数と対応する事象として解釈できる。また与えられた密度関数を確率分布のモデルとみなせば、密度関数の極大値点は最も確率の高い事象と理解できることから、この極大値が最頻値と呼ばれる。

最頻値探索は、このような密度関数の極大値探索の問題で、任意の初期値 $\mathbf{x}_0 \in X$ の近傍における密度関数の極大値を求める。この問題の解法には、繰り返しアルゴリズムを用いるのが一般的で、 j 番目の繰り返しステップを $\mathbf{s}(\mathbf{x}_j)$ とすると、初期値 $\mathbf{x}_0$ から次の手続きを反復する

一般的極大値探索法:

(1) $\mathbf{x}_j$ におけるステップの計算: $\mathbf{s}(\mathbf{x}_j)$

(2) 現在位置の更新: $\mathbf{x}_{j+1} \leftarrow \mathbf{x}_j + \mathbf{s}(\mathbf{x}_j), j \leftarrow j+1$

一般に極大値探索法は、この反復の繰り返しにより $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_j$ が逐次的に収束し、かつその収束値 $f(\mathbf{x}_j)$ が最頻値となるように、繰り返しステップ $\mathbf{s}(\mathbf{x}_j)$ を目的関数 $f(\mathbf{x})$ やサンプル点 S を引数とするベクトル関数として定義する。このように定義される手法には、ミーンシフトの他にも後述する最急勾配法 (gradient ascent; 2.4 節) や共役勾配法また EM アルゴリズムなども含まれる。これらの手法は局所解におちいると大局的最大値が求まらないという欠点があるものの、理論的な汎用性や計算上の簡易性により非常に多くの分野で使われている。

ミーンシフトは、与えられたサンプル集合 S で定義されるカーネル密度関数 (カーネル関数の和

で定義する平滑化された密度関数)における極大値探索法であるが、この理論的詳細は3.1節で触れる。

2.2 直感的なミーンシフトの定式

ミーンシフトを最初に提案した福永らによる定義[21],[22]はミーンシフトの手法を直感的に理解するのに適した例であるので、二次元の場合についてまず簡単に説明する。サンプル点の集合 $S = \{\mathbf{x}_i\}$ が与えられたとする。次に任意の点 $\boldsymbol{\mu}$ を中心として半径 h の円領域 $T_h(\boldsymbol{\mu})$ を考え、この円領域内にあるサンプル点の平均値 $\mathbf{m}(\boldsymbol{\mu})$ を計算する。福永のミーンシフトは、現在位置 $\mathbf{x}$ におけるこの平均値 $\mathbf{m}(\mathbf{x})$ と $\mathbf{x}$ との差分ベクトルを繰り返しステップ $\mathbf{s}(\mathbf{x}) = \mathbf{m}(\mathbf{x}) - \mathbf{x}$ として定義し、1) 円領域内で平均を計算し 2) そこに現在位置を更新する手続きを反復繰り返しすることで、初期値近傍の最頻値点に収束させる。

福永のミーンシフト[21]:

- (1) 平均値の計算: $\mathbf{m}(\mathbf{x}_j) = \frac{1}{|T_h(\mathbf{x}_j)|} \sum_{\mathbf{x} \in T_h(\mathbf{x}_j)} \mathbf{x}_i$
- (2) 現在位置の更新: $\mathbf{x}_{j+1} \leftarrow \mathbf{m}(\mathbf{x}_j), j \leftarrow j+1$

ここで $|T_h(\mathbf{x})|$ は $\mathbf{x}$ を中心とする円領域内にあるサンプル点の数を示す。

次に、円領域 $T_h(\boldsymbol{\mu})$ を次のような関数 $K_f(\mathbf{x}; \boldsymbol{\mu}, h)$ として数式で表す。

$$K_f(\mathbf{x}; \boldsymbol{\mu}, h) = \begin{cases} 1 & \text{if } \|\mathbf{x} - \boldsymbol{\mu}\| \leq h \\ 0 & \text{if } \|\mathbf{x} - \boldsymbol{\mu}\| > h \end{cases} \quad (1)$$

これは詳しく後述するカーネル関数のひとつで、閉領域内の均一分布を表すフラットカーネルと呼ばれる。定数 h は一般的にカーネル幅 (bandwidth) と呼ばれる。福永によるミーンシフト繰り返しステップは、このカーネルによる重み付き平均を使って次のように再定義できる。

$$\mathbf{s}(\mathbf{x}) = \mathbf{m}(\mathbf{x}) - \mathbf{x} = \frac{\sum_{i=1}^n K_f(\mathbf{x}_i; \mathbf{x}, h) \mathbf{x}_i}{\sum_{i=1}^n K_f(\mathbf{x}_i; \mathbf{x}, h)} - \mathbf{x} \quad (2)$$

ここでは数列和の範囲がサンプル S 全体に及ぶ事に注意する。福永のミーンシフトの平均値計算を式(2)で置き換えても同じ最頻値に収束する。

ミーンシフトの名前の由来は、平均値(ミーン)へ現在位置を更新(シフト)する繰り返し手続き

による。任意のパラメタは円領域の半径 h のみとなり、高次元の場合でも、同じ半径の球体を考えれば容易に拡張できる。

2.3 一般的なミーンシフトの定式

Cheng [7] は、前節で述べた福永の定義によるミーンシフトを次の三点について拡張した。

- (1) 一般的なカーネル関数への拡張
- (2) サンプル集合に対する重み関数の導入
- (3) サンプル集合を固定し任意の初期値からの探索

第三点については、2.5節でミーンシフトのクラスタリング応用における問題として後述する。フラットカーネル以外のガウス関数なども含む一般的カーネルを $K(\mathbf{x}; \boldsymbol{\mu}, h)$ で、また非負値の重み関数を $w(\mathbf{x}): S \rightarrow (0, \infty)$ で表すと、サンプル $\mathbf{x}_i$ の重み付き平均が次のように定義できる。

$$\mathbf{m}(\mathbf{x}) = \frac{\sum_{i=1}^n K(\mathbf{x}_i; \mathbf{x}, h) w(\mathbf{x}_i) \mathbf{x}_i}{\sum_{i=1}^n K(\mathbf{x}_i; \mathbf{x}, h) w(\mathbf{x}_i)} \quad (3)$$

この式を用いて、福永の定義を拡張したミーンシフトの一般化定式が求まる。

Cheng のミーンシフト [7]:

- (1) 式(3)による重み付き平均値 $\mathbf{m}(\mathbf{x}_j)$ の計算
 - (2) 現在位置の更新: $\mathbf{x}_{j+1} \leftarrow \mathbf{m}(\mathbf{x}_j), j \leftarrow j+1$
- 導入された重み関数以外は明らかに同一の定式となる。この重み関数は5.2節で後述するミーンシフトのスケールスペースへの拡張や、6.2節で概説する動画像におけるトラッキングへの応用において重要な役割を果たす。また一般的なカーネル関数の定義と性質については3.1節で説明する。

2.4 最急勾配法との比較

最急勾配法は、数多くある極値探索問題の解法の中で最も明解で汎用性の高い手法で、繰り返しステップを、ステップ方向を表す目的関数 $f(\mathbf{x})$ のグラジエント(空間偏導関数ベクトル) $\nabla_{\mathbf{x}} f(\mathbf{x})$ とステップ幅を表す定数 γ の積として定義する。

最急勾配法:

- (1) $\mathbf{x}_j$ でのステップの計算: $\mathbf{s}(\mathbf{x}_j) = \gamma \nabla_{\mathbf{x}} f(\mathbf{x}_j)$
 - (2) 現在位置の更新: $\mathbf{x}_{j+1} \leftarrow \mathbf{x}_j + \mathbf{s}(\mathbf{x}_j), j \leftarrow j+1$
- γ が正值の場合には極大値(山の頂点; 最急上昇法)に負値の場合には極小値(谷の底; 最急降下法)

に収束できる。実用的にはこの γ の値の設定が問題で、大きすぎると振動や発散を起こし、また小さすぎると収束時間が増加するため、適当な値に手で設定するのが煩雑また困難になる。これに比べ、ミーンスフトにはステップ幅定数がなく、ステップ幅が可変で自動的に決まる最急勾配法として解釈できる (3.3 節)。このため最急勾配法と比べたミーンスフトの利点として、振動や発散することがないための確実な収束と効率的な高速収束の点が挙げられる。定式の簡略化については、新たにカーネル幅の定数パラメタが導入されるため実用簡易性の面での向上は大きくない。カーネル幅の自動推定法については 3.4 節で述べる。

2.5 ミーンスフトを用いたクラスタ解析

クラスタ解析は、前節で定義したミーンスフトの最も基本的で汎用な応用問題で、与えられたサンプル集合 $S = \{\mathbf{x}_i \in X\}$ を、任意数 C 個の相互背反で完全な部分集合 $S_1, \dots, S_C$ にグループ分けするデータ解析を指し、画像の領域分割やパターン認識やデータマイニング等の様々な分野で重要な手法である。ミーンスフトを用いた手法は基本的に次の手順に従う。

ミーンスフトクラスタリング [7], [12]:

- (1) n 個の初期値 $T = \{\mathbf{y}_i \in X\} \leftarrow S$ の設定
- (2) 各初期値 $\mathbf{y}_i$ からミーンスフトを実行
- (3) 目的関数の任意数 C 個の最頻値点を求める
- (4) 各初期値 $\mathbf{y}_i$ を収束先の最頻値点のインデックス c でラベル付ける

初期値集合 T にサンプル集合 S を代入 ($T \leftarrow S$) すれば、総てのサンプル点 $\mathbf{x}_i \leftarrow \mathbf{y}_i$ にミーンスフトの収束点に基づくラベル付け $\mathbf{x}_i \leftarrow c$ をすることになり、結果としてデータを同ラベル c を持つサンプル点の集合 S_c にグループ分けできる。クラスタ中心推定法 (2 & 3 ステップ) には、ふつう初期値集合 T とサンプル集合 S を分けて扱うので、ミーンスフトの目的関数が固定して処理され、目的関数の最頻値点がミーンスフトの繰り返しに影響されない。この方式は、前述した Cheng [7] の拡張の第三点として、 S の部分集合や S と違う集合を初期値集合 T とした場合のクラスタ解析を

も可能とする。また異なった初期値からのミーンスフトは計算上お互いに独立なため、並列処理による高速化が可能となる利点がある。ミーンスフトの繰り返し毎にサンプル集合 S を動的に変化させる拡張法もあるが、これには 5.4 節で触れる。

2.6 他のクラスタ解析法との比較

ミーンスフトによるクラスタ解析は、データの分布の関数モデルの仮定を必要としないノンパラメトリックなクラスタ解析手法の一つである。このため、EM アルゴリズムを使った混合ガウス分布関数あてはめに代表される、パラメトリックなクラスタ解析に比べると、より拘束の少ないデータへの適用が可能となる。また多くのクラスタ解析問題に良く使われる k -平均クラスタ解析法は、カーネル幅 h を無限大にした場合のミーンスフトクラスタ解析法として解釈できる [7]。このため、有限幅のカーネルを導入することにより、データ中の外値 (Outlier) からの影響が少なくなることから、 k -平均解析法を外値ノイズに対して頑強化したものとして、ミーンスフト法を理解できる。最後に、一般のクラスタ解析法はふつうクラスタ数 C の既知を前提とするが、ミーンスフト法は、自動的にクラスタ数を決定し、 C 値の仮定を必要としないことが重要な利点の一つである。しかし、カーネル幅 h の値がクラスタ数も含めた解析結果を大きく影響するので、この定数の値決めに注意を要する。3.4 節でカーネル幅の自動推定法について述べる。

3. ミーンスフトの理論

本節では、ミーンスフトの基本理論について主に Comaniciu の研究に沿って解説する。詳しくは [12], [13] を参照されたい。

3.1 カーネル密度推定

カーネル密度推定法 (kernel density estimation) [45] は、Parzen windows [18], [39] と呼ばれ、最もポピュラーであるノンパラメトリックな密度関数の推定法である。 d 次元空間 $X \subset \mathbb{R}^d$ 上におけるデータ点の集合 $S = \{\mathbf{x}_i | i = 1, \dots, n\}$ が与えられたとする。ついでカーネル関数 (kernel) K

とそのカーネル幅 $h > 0$ が与えられれば、多変数のカーネル密度推定はデータ点を中心に置いた幅 h のカーネル集合の和として定義される。

$$f(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n K(\mathbf{x}; \mathbf{x}_i, h) \quad (4)$$

一般的には多次元空間におけるカーネル形状は $d \times d$ の対称正定値行列 H によって拘束されるが、以下では解説上 H を単位行列に比例するもの $H = h\mathbf{I}$ と簡略化して考える。厳密な d 次元のカーネルは有界で compact な support を持つ対称性を持った関数として定義され、[53] に詳しいが、ここでは特に球対称なカーネル関数の定義を考える。まず空間の任意点とサンプル点の距離とカーネル幅の比の関数 $K(\mathbf{x}; \mathbf{x}_i, h) = K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right)$ を考える。ここで球対称のカーネル関数は、プロファイル (profile) と呼ばれる、 $\mathbf{x} \geq 0$ 上の有界で区分的に連続 (piecewise continuous) な非負値の関数 $k(\mathbf{x}) : [0, \infty) \rightarrow \mathbb{R}$ を用いて定義できる。

$$K(\mathbf{x}) = k(\|\mathbf{x}\|^2) \quad (5)$$

ここで $\|\mathbf{x}\|$ はベクトルのノルム $\|\mathbf{x}\|^2 = \sum_{i=1}^d |x_i|^2$ を表す。このような球対象のカーネルの例として、次のプロファイル k_g で定義される、

$$k_g(\mathbf{x}) = A_g \exp\left(-\frac{1}{2}\mathbf{x}\right) \quad (6)$$

ガウス関数による正規分布カーネル K_g や、

$$K_g(\mathbf{x}) = A_g \exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right) \quad (7)$$

また以下のプロファイル k_e

$$k_e(\mathbf{x}) = \begin{cases} A_e(1 - \mathbf{x}) & 0 \leq \mathbf{x} \leq 1 \\ 0 & \mathbf{x} > 1 \end{cases} \quad (8)$$

で定義される Epanechnikov カーネル K_e

$$K_e(\mathbf{x}) = \begin{cases} A_e(1 - \|\mathbf{x}\|^2) & \|\mathbf{x}\| \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

がある。ここで、式 (5) にあるプロファイル関数 k を用いて、式 (4) にあるカーネル K を用いたカー

ネル密度推定を再定義できる。

$$f_K(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (10)$$

A_g と A_e はカーネルを確率分布とみなすための正規化項で $\int_{\mathbb{R}^d} K(\mathbf{x}) d\mathbf{x} = 1$ を保証するが、ミーンスフトの公式にカーネルは比の形で現れるのでこのような定数積項は無視してよく、計算上の簡略化のためふつう $A_g = A_e = 1$ とおく。またカーネルの球対称性から定義中の空間変数 $\mathbf{x}$ とサンプル点 $\mathbf{x}_i$ が互換となる。

$$K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) = K\left(\frac{\mathbf{x}_i - \mathbf{x}}{h}\right) \quad (11)$$

このため 2. 節で用いた、空間変数 $\mathbf{x}$ を中心とするカーネルのサンプル点 $\mathbf{x}_i$ 上の関数値 $K(\mathbf{x}_i; \mathbf{x}, h)$ と、本節の $\mathbf{x}_i$ を中心とするカーネルの $\mathbf{x}$ における値 $K(\mathbf{x}; \mathbf{x}_i, h)$ は同値となる事に注意する。

3.2 カーネル密度の勾配推定とミーンスフト

ここでミーンスフトの公式 (2) をカーネル密度関数の勾配推定 (gradient estimator) から導出する。カーネル密度関数の勾配推定は、微分作用素の線形性によりカーネル密度推定の微分として式 (10) から定義される。

$$\nabla_{\mathbf{x}} f_K(\mathbf{x}) = \frac{2}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x} - \mathbf{x}_i) k' \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \quad (12)$$

このプロファイルの空間微分 k' を新たな関数 $g(\mathbf{x}) = -k'(\mathbf{x})$ で置き換えて上式を次のように変形する。

$$\begin{aligned} \nabla_{\mathbf{x}} f_K(\mathbf{x}) &= \frac{2}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}) g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \\ &= \left[\frac{2}{nh^{d+2}} \sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \right] \times \\ &\quad \left[\frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \mathbf{x}_i}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)} - \mathbf{x} \right] \end{aligned} \quad (13)$$

ここで関数 $g(\mathbf{x})$ をプロファイルとして定義されるカーネル関数 $G(\mathbf{x})$ を式 (5) と同様に定義する。

$$G(\mathbf{x}) = g(\|\mathbf{x}\|^2) \quad (14)$$

式(13)における積項の左辺は、このカーネル $G(\mathbf{x})$ で定義した $\mathbf{x}$ での密度推定 $f_G(\mathbf{x})$ に比例する。

$$f_G(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) \quad (15)$$

また同項の右辺は、前述した式(2)と同型の、カーネル $G(\mathbf{x})$ を用いたミーンシフトである。

$$\mathbf{s}_G(\mathbf{x}) = \mathbf{m}_G(\mathbf{x}) - \mathbf{x} = \frac{\sum_{i=1}^n G\left(\frac{\mathbf{x}-\mathbf{x}_i}{h}\right) \mathbf{x}_i}{\sum_{i=1}^n G\left(\frac{\mathbf{x}-\mathbf{x}_i}{h}\right)} - \mathbf{x} \quad (16)$$

式(15)と式(16)を式(13)に代入すると、カーネル K によるカーネル密度関数の勾配推定は次のように簡略化できる。

$$\nabla_{\mathbf{x}} f_K(\mathbf{x}) = \frac{2}{h^2} f_G(\mathbf{x}) \mathbf{s}_G(\mathbf{x}) \quad (17)$$

この式から、カーネル G によるミーンシフトは、 K による密度勾配推定と G による密度推定の比とカーネル幅 h の関数として導出される。

$$\mathbf{s}_G(\mathbf{x}) = \frac{1}{2} h^2 \frac{\nabla_{\mathbf{x}} f_K(\mathbf{x})}{f_G(\mathbf{x})} \quad (18)$$

3.3 ミーンシフトの性質と収束証明

前節で導いたミーンシフトの公式(18)から様々な理論的考察が得られる。まず式(18)の右辺は密度勾配推定項 $\nabla_{\mathbf{x}} f_K(\mathbf{x})$ 以外はすべて正值のスカラ項なので、 G によるミーンシフトのベクトル方向は K による密度関数の勾配方向に一致することが解る。

これからミーンシフトによる最頻値探索は、ミーンシフトカーネル G に対応するカーネル K で定義する密度関数の最頻値点 $\nabla_{\mathbf{x}} f_K(\mathbf{x}) = \mathbf{0}$ に、最急勾配法の要領で収束することが解る。このようなカーネル K を、そのプロファイルの微分で定義されるカーネル G のシャドウカーネル (shadow kernel) と呼ぶ[7], [13]。

ミーンシフトの収束点は一般的には同じカーネルによる密度推定の最頻値ではないことに注意する。例えば、福永の定式に使われたフラットカーネル K_f のシャドウは式(9)で定義したEpanechnikovカーネル K_e であるので、福永のミーンシフトは K_e による密度関数の最頻値に収束する。

ただし、式(7)で定義した正規分布カーネル K_g については $g(\mathbf{x}) = 0.5k(\mathbf{x})$ となるので、定数項を無視してミーンシフトカーネルとそのシャドウが同一である($K_e = G_e$)とみなしてよい。このことから、実装の簡易上正規分布カーネルが良く使われる。

またミーンシフトは、ステップ幅 γ を変数項 $\frac{h^2}{2f_G(\mathbf{x})}$ とおいた最急勾配法とみなせる。具体的には、カーネル幅 h の二乗に比例し、現在点 $\mathbf{x}$ における G での密度推定値に反比例するように、ステップ幅を自動設定する最急降下法としてミーンシフトが理解できる。これから、 h が大きいくほどステップ幅も大きくなり、密度 f_G の小さいサンプルのまばらなところではステップ幅が大きくなり、密度が高くサンプルのあつまる最頻値に近い地点ではステップ幅が小さくなることから、効率的で振動のない収束が得られることが推測できる。

ミーンシフトの収束の条件と証明は[13]にある定理がもっとも一般的で精確である。

ミーンシフトの収束定理 [13, Theorem1]:

カーネル K が、凸状(convex)で単調減少(monotonically decreasing)するプロファイル k を持つならば、 K をシャドウとするカーネル G を用いたミーンシフト式(16)による最頻値探索は、 K による密度関数(10)上の初期値近傍の最頻値点に必ず収束する。

具体的には、ミーンシフトのステップ列 $\{\mathbf{x}_j\}_{j=0,1,\dots}$ と対応する密度推定値の数列 $\{f_K(\mathbf{x}_j)\}_{j=0,1,\dots}$ が収束し、かつ密度推定値が単調に増加すること $f_K(\mathbf{x}_j) \leq f_K(\mathbf{x}_{j+1}) \forall j$ を証明できることから、最頻値点への収束が保証される。本稿で示すカーネル K_e と K_g も含めて、上記の収束条件を充たすどのカーネル関数にも収束が保証される。

3.4 カーネル幅の推定

カーネル幅 h の値の設定は、カーネル K の種類の設定に比べて、ミーンシフトアルゴリズムの結果により大きく影響する重要な未解決問題である。統計学的手法では、未知の密度関数 $f(\mathbf{x})$ とそのカーネル推定 $\hat{f}(\mathbf{x})$ の誤差を最小化するよ

うな最適カーネル幅を求める手法が知られている [43], [45], [50], [53]。これらは、ISE(integrated squared error), MISE(mean ISE), AMISE(asymptotic MISE) のような推定誤差関数の最小値化により、推定の偏り (bias) と分散 (variance) 間で最もバランスの取れた最適解を得るものである。またカーネル幅で正規化されたミーンシフトベクトルの長さを最大化するカーネル幅を求める手法が Comaniciu [15] により提案されている、またこの手法の高次微分や非等方カーネル幅行列 H への拡張は岡田ら [34] によって報告された。これらの手法は最近注目される SIFT 法 [31] の基礎である Lindeberg の自動スケール決定法 (automatic scale selection) [30] と関連している。安定性 (stability) を指標としたカーネル幅推定法もいくつか提案されている。クラスタ解析の結果の安定性を最適化するもの [21], また目的密度関数中の注目するピークにガウス関数をあてはめ、結果が最も安定となるカーネル幅を選ぶ手法が提案されている [9], [32], [34], [35]。基本的には、問題の設定次第で最適性の定義自体が変化するので、汎用の自動カーネル幅推定法の導出は難しい問題である。このため問題によっては、グラフィックスを使ったユーザインタフェースやシステムの他部位からの出力を使った値決めも実装に向けて考えられる。

3.5 理論的比較

ミーンシフトのようなカーネルを用いたデータ処理は、データ中の外値 (outlier) ノイズに頑強なロバスト推定法の基盤となる。与えられたデータ $S = \{\mathbf{x}_i\}$ を、確率密度関数 $p(X)$ に従う確率変数 X から抽出した IID ランダムベクトル変数とみなすと、式 (10) にあるカーネル密度推定は、この確率密度関数の近似推定 $\hat{p}(X)$ と考えられる。一般的にこのような確率分布に最頻値が複数存在する場合、これをマルチモーダル分布と呼ぶ。データ中に外値ノイズが多く存在する場合、もとの分布 $p(X)$ がユニモーダル (最頻値が一つのみ) でも、その推定 $\hat{p}(X)$ がマルチモーダルとなる。このようなデータから、ノイズに頑強にもとの分布

の統計量を求める手法をロバスト推定と呼ぶ。

ここで、もとの分布 $p(X)$ のモード位置をパラメタ $\mathbf{y}$ として、尤度分布 (likelihood distribution) を $p(\mathbf{x}_1, \dots, \mathbf{x}_n | \mathbf{y})$ とおくと、これを用いた最尤推定 (MLE: maximum likelihood estimation) は、サンプル S によるカーネル密度推定関数の直接最大化として定義できる [13], [46]。このため、ミーンシフトによる密度推定の極値解析は、上記の空間最尤推定問題の枠組みで理解できる。

また一般的なロバスト推定の手法として知られる M 推定法は、原点が最小値で単量増加する対称非負値の関数の和を目的関数として最適化を行う手法である。この目的関数は、カーネル密度関数の符号を変えたもの一致することから、ミーンシフトを用いた上記の空間最尤推定は M 推定法の枠組みでも理解できる [13]。

ここでマルチモーダルな分布 $\hat{p}(X)$ 自体をデータと考え、その極大値中最大のものを選ぶ手法を voting algorithm と呼ぶ。Comaniciu と Meer はミーンシフトを応用したより一般的な統計的ロバスト情報統合法 [4], [11] を提案した。この手法は、各データ点 $\mathbf{x}_i$ と対応する局所カーネル幅 h_i を、統計的に独立な空間点パラメタ $\mathbf{y}$ の推定とその uncertainty とみなし、後述する可変カーネル幅ミーンシフトによるモード追跡を利用して、分布極大値中最も重要なものを計算する。これから、分布形状の仮定がない場合においても、データから得られるマルチモーダルな分布 $\hat{p}(X)$ による、ロバストな空間最尤推定が可能となる。

ミーンシフト公式の理論的導出と裏づけには、3.2 節で述べた目的関数の空間微分によるものの他に、界関数最適化 (variational bound optimization) の理論 [26], [52] を用いたものがある [19], [46]。Fashing と Tomashi [19] はカーネル密度による目的関数の二次下界関数 (quadratic lower bound) の最適化の枠組みを使ってミーンシフト法を考察し、ミーンシフト法と二次微分ヘシアンを用いる Newton 法が同値であることを示した。また後述するミーンシフト法の最大事後確率推定への拡張にも、この界関数最適化の枠組みが使われている

る [37]。

最近になって、高速ガウス変換を応用したミーンスフトの高速化が Yang ら [56] により、正規分布カーネル K_g を用いる場合のミーンスフト法と EM アルゴリズムの理論的同値性とその収束の高速化が、Carreira-Perpian [5], [6] により、また half-quadratic 最適化の枠組みでのミーンスフトの理論的解析とその高速化が Yuan と Li [59] により、最後に、非線形の Grassmann manifold へのミーンスフトの拡張とその拡散テンソル画像解析への応用が Subbarao と Meer [47] により報告されている。

4. ミーンスフトの利点と欠点

以上説明した基本的なミーンスフト法とそのクラスタ解析法への応用についての利点と欠点をまとめる。

利点: (a) 最頻値探索において振動や発散のない繰り返し収束が数理的に保証されている。(b) 目的関数であるカーネル密度関数の値を計算せずに最頻値が探索できる。(c) 最急勾配法にあるステップ幅 γ パラメータを手動で値決めする必要がない。(d) 最急勾配法に比べ自動可変ステップ幅により効率的な収束が可能となる。(e) マルチモーダル分布における頑強な空間的最尤推定が可能となる。(f) 汎用のノンパラメトリックなクラスタ解析を提供する。(g) 任意のデータに対してクラスタ数 C が自動的に決まる。(h) k -平均クラスタ解析法よりデータの外値ノイズに対してより頑強である。(i) クラスタ解析において並列処理による高速化が可能である。

欠点: (a) カーネル密度関数のようなカーネル和で定義される目的関数にしか適用できない。(b) 高次元空間では密度関数推定自体が困難となるので適用が難しくなる。(c) データ解析結果がカーネル幅 h に大きく影響されるのでその推定に注意を要する。(d) 単純繰り返し法であるため大局の最大値は常にもとまらない。(e) 繰り返し法以外の解法がある場合それ比べ計算時間が大きく増加する。

ここに欠点としてあげたミーンスフトの性質についても、最近になってその改良に向けて様々な提案がなされてきている [6], [44], [47], [59]。

5. 理論的拡張

5.1 可変カーネル幅ミーンスフトへの拡張

2. 節で解説したミーンスフトの基本アルゴリズムは、カーネル幅 h を定数として扱った。Comaniciu はこのカーネル幅を空間変数の関数 $h(\mathbf{x})$ とみなし、より柔軟で正確なカーネル密度推定を可能とする、可変カーネル幅ミーンスフト [9], [11], [15] を提案した。この手法は、カーネルの球対称性を利用 (3.1 節) してミーンスフト公式中のカーネル和を、現在点 $\mathbf{x}$ を中心に置いたカーネルの $\mathbf{x}_i$ データ各点における n 個の関数値 $K(\mathbf{x}_i)$ の和としてでなく、 $\mathbf{x}_i$ データ各点を中心に置かれた n 個のカーネルの現在点 $\mathbf{x}$ における関数値 $K(\mathbf{x})$ の和としてみなし、空間上に配置したカーネル幅 $h(\mathbf{x})$ を各データ点近傍の局所統計量をもとに導出する。具体的には、まずミーンスフトクラスタ解析を行い、3.4 節で概説したガウス関数のあてはめ安定性をもとに、解析結果の各クラスタそれぞれに最安定カーネル幅を求め、ついでデータ各点に、対応するクラスタでの幅推定値を代入することで、 $\{h_i | i = 1, \dots, n\}$ を求める。可変カーネル幅ミーンスフト公式は、 $\{h_i\}$ を用いた可変幅カーネル密度推定の空間微分から 3.2 節と同様に導出できる [9]。この拡張は、voting algorithm の一種である、前述した統計的ロバスト情報統合 [11], [36] の基盤ともなる。

5.2 スケールスペースへの拡張

スケールスペース [28], [54] は、入力画像を初期値とする拡散方程式の解析解を用いた画像表現・処理の理論で、エッジを保存する画像の非平方平滑化などへの応用も知られている [2], [40]。この方程式解は、カーネル幅が連続に単調増加するガウス関数での入力平滑化として考えられるので、広く知られる Gaussian pyramid 表現を連続化したものとも理解できる。このスケールスペースへのミーンスフトの拡張は、Collins [8] による

提案と、岡田ら [32] による非等方カーネルへの拡張が知られている。ガウス関数による平滑化には、正規分布カーネル K_g による入力 I の畳み込み積分で行う。この積分式を離散化したものと式 (10) にあるカーネル密度推定の類似性に注目し、その空間微分から 3.2 節と同様にミーンシフトを導出すると、入力画像 $I(\mathbf{x})$ を重み $w(\mathbf{x})$ と置いた、式 (3) にある Cheng の重み付きミーンシフト式が求まる。画像値はふつう正值なので、全体の和が 1 となるように正規化する前処理を行えば、ガウス平滑化した画像の輝度地分布における極値探索法である、スケールスペースミーンシフトが定義される。この方式は、ミーンシフト法の画像関数への直接適用を可能とし、またガウス平滑化を施した画像輝度地分布のピークを、実際に平滑化を施さずに求めることができるので、実装における計算量の面での利点がある。

5.3 最大事後確率推定への拡張

3.5 節ではミーンシフト法による空間最尤推定への応用について触れた。岡田ら [37] は、推定変数 $\mathbf{x}$ についての事前確率分布 (prior distribution) を導入し、ミーンシフト法の最大事後確率推定 (MAP: maximum a-posteriori estimation) への拡張である、事前確率に拘束されたミーンシフト法 (prior-constrained mean shift) を提案した。この方式は、ペイズの公式に基づき、事前確率分布とカーネル密度推定で近似した尤度関数 (likelihood distribution) から事後確率分布 (posterior distribution) を導き、その最大化を計る。[37] では、正規分布カーネル K_g の場合について、3.5 節で触れた界関数最適化 (variational bound optimization) 手法を用いて、 K_g を正に取る引き込み型事前分布と、負に取る押し出し型事前分布の両方を導入したミーンシフト公式が提案された。これにより、ミーンシフト法の理論的利点を保持したまま、ユーザや統合システム他部位の出力により指定する空間地点へ、インタラクティブにミーンシフトの収束を引き込んだり、遠ざけたりすることが可能となる。

5.4 動的ミーンシフトへの拡張

2.5 節で解説したミーンシフトクラスタリング法において、データ各点からミーンシフトを行い任意 C 個のクラスタ中心を推定する過程で、初期値集合 T と与えられたサンプル集合 S をまったく同値とみなして、ミーンシフトの繰り返し毎にサンプル集合自体をシフトさせる方法がある。この手法は、 S のカーネル密度関数として定義される目的関数自体が、ミーンシフトの繰り返し毎に動的に変化するので、動的ミーンシフト [24], [60] や CAMSHIFT [3] と呼ばれ、福永によるオリジナルな定式もこの方式による [22]。この手法では、ミーンシフトの繰り返し毎に、 S 中の各サンプル点 $\mathbf{x}_i$ に式 (3) によるシフトを一回のみ適用して、サンプル集合全体を一括して処理する。この手法は比較的効率的な収束が得られることが知られているが、結果の収束点がもとの目的関数の最頻値ではなくなることや、また実装時におけるメモリ使用量が増大する欠点がある。

6. ミーンシフトの応用

6.1 画像のセグメンテーションと平滑化

画像の領域分割・セグメンテーションはミーンシフト法の最も重要な応用問題のひとつで、任意の入力画像を、あらかじめ決めた特徴量 (feature) に基づいて、homogeneous な小領域に分割する手法である。基本的なミーンシフトによる画像の領域分割は、2.5 節で解説したミーンシフトのクラスタ解析の手法を用いて、輝度値と空間変数の両方を考慮した領域分割を可能とする。

まず一般的な輝度値のみを用いた領域分割法について説明する。与えられた入力画像 $I(\mathbf{x})$ の画素が M 個あるとし、これを $(\mathbf{x}_1, \mathbf{f}_1), \dots, (\mathbf{x}_m, \mathbf{f}_m), \dots, (\mathbf{x}_M, \mathbf{f}_M)$ のサンプルベクトル集合として考える。ここでベクトル $\mathbf{f}_m$ は画素 $\mathbf{x}_m$ における画素値 $\mathbf{f}_m = I(\mathbf{x}_m)$ を表す。入力画像がグレースケールである場合には、 $\{\mathbf{f}_m\}$ は一次元の点集合だが、カラー画像の場合は、RGB や LUV などの適当なカラー空間を適用して、 $\{\mathbf{f}_m\}$ が三次元のベクトル点集合となる。ここでは、画

素 $\mathbf{x}_m$ における特徴量 $\mathbf{y}_m$ を対応する画素値で設定する $\mathbf{y}_m = \mathbf{f}_m$ 。ミーンスフトによる領域分割は、この特徴量サンプル $T = \{\mathbf{y}_m\}$ 、カーネル関数の種類、任意のカーネル幅 h を入力とし、画素 $\mathbf{x}_m$ のラベル L_m を出力として、次の手順で行う。

ミーンスフトの領域分割:

- (1) M 個の $\mathbf{y}_m$ からミーンスフトを実行
- (2) 結果 M 個の収束点 $\mathbf{z}_m$ を得る
- (3) 距離が h 以下の近傍の点 $\mathbf{z}_m$ をグループ化
- (4) 結果 $\{\mathbf{z}_m\}$ を C 個のクラスタ $\{\mathcal{C}_c\}$ に分割
- (5) 各画素 $\mathbf{x}_m$ について収束先のクラスタのインデックス c でラベル付けをする $L_m = c | \mathbf{z}_m \in \mathcal{C}_c$

このアルゴリズムは画像の次元数に関わらず、輝度値の似ている小領域に分割できるが、距離が大きく離れている画素でも、偶然やノイズによって輝度値が似ていれば、同領域にラベル付けがされるので、結果としてラベル分布が空間的に homogeneous でなくなる傾向がある。このことを考慮し、Comaniciu と Meer [12], [13] は、輝度値と空間変数の合成空間を特徴空間とすることで、空間と輝度値の両方に似通った領域への分割を可能とする手法を提案した。この方式は、画素 $\mathbf{x}_m$ における特徴量 $\mathbf{y}_m$ を、その画素空間変数と、対応する画素値ベクトルの合成ベクトル $\mathbf{y}_m = (\mathbf{x}_m, \mathbf{f}_m)$ とみなす。輝度値 (range) と空間変数 (space) の両方に対するカーネル幅 (h_r, h_s) を準備すれば、上記のアルゴリズムをそのまま適用してよい。二次元カラー画像を入力とする場合には、この方式の特徴ベクトルは五次元となる。

この領域分割アルゴリズムを部分的に利用して、エッジを保存した画像の平滑化のアルゴリズムが定義できる [12], [13]。具体的には、ステップ (2) で求まる特徴量空間内での収束点 $\mathbf{z}_m$ を用いて、 $\mathbf{z}_m$ の画素値を、対応する初期値画素 $\mathbf{x}_m$ における画素値 $\mathbf{f}_m$ として代入する。各クラスタに属する画素の集合 (basin of attraction) にその最頻値の画素値を上書きすることとなり、ガウス平滑化などと比べて、エッジの平滑化をせずに輝度値分布の平滑化を行うことができる。

一般の画像以外にも、リモートセンシン

グ画像 [20] やステレオ視によるレンジデータ [4], [25], [46] の領域分割への応用成功例も報告されている。最近ではこの基本方式の拡張として、Morse 理論を用いた密度分布のトポロジカルな分割に基づく階層的な画像領域分割法 [38] が提案されている。

6.2 動画におけるトラッキング

Comaniciu らによって CVPR'00 で提案されたミーンスフトトラッキング [10], [14] は、動画上の物体追跡・トラッキング問題での実時間ロバスタアルゴリズムである。この手法は、物体追跡問題の解法における最も知られた手法の一つで、多くの追跡問題への応用が報告されている [1], [10], [11], [14], [16], [44], [55], [57], [58]。

物体追跡問題は、動画上の k 番目のフレームにおける物体の現在位置を $\mathbf{x}_k$ とすると、次の $k+1$ 番目のフレームにおいて、同物体の対応するターゲット点 $\mathbf{x}_{k+1}$ を推定し、これを動画上で繰り返す問題である。本手法は、 $k+1$ 番目のフレーム上で、前フレームでの物体位置である $\mathbf{x}_{k+1,0}$ を、ミーンスフト法の初期値 $\mathbf{y}_0$ と設定し、このミーンスフトでの収束点 $\mathbf{y}_j$ をターゲット点 $\mathbf{x}_{k+1}$ の推定とみなすことを基本とする。これにより、隣接するフレーム間において、 k 番フレームでの物体位置 $\mathbf{x}_k$ を中心に抽出するカラー画素値のヒストグラム $\mathbf{q}_u$ と最も似たカラーヒストグラム $\mathbf{p}_u(\mathbf{y})$ を供給する、 $k+1$ 番フレームでの位置 $\mathbf{y}_j$ を、 $\mathbf{x}_k$ の近傍で探索する。実装上、カーネル幅を適当に設定すれば、数回の繰り返しでミーンスフトが収束するので、実時間の物体追跡が可能となる。またミーンスフト法の頑強性により、画像中のノイズによる影響が少ない追跡も可能となる。

理論的には、両フレームにおける U 個の領域からなるカラーヒストグラム $\mathbf{p}_u(\mathbf{y})$ と $\mathbf{q}_u$ をカーネル関数で平滑化したものを考え、この分布の類似度を表す Bhattacharyya 係数 $\rho(\mathbf{p}_u(\mathbf{y}), \mathbf{q}_u) = \sum_{u=1}^U \sqrt{\mathbf{p}_u(\mathbf{y}) \mathbf{q}_u}$ を最大化する $\mathbf{y}$ を考える。Bhattacharyya 係数は理論的に Bayes エラーと関連し metric の性質を持つ類似度関数として知られている。ここで [14] の導出に従いテラー展開によ

る近似を施すと、Bhattacharyya 係数は次の式として表される。

$$\rho(\mathbf{p}_u(\mathbf{y}), \mathbf{q}_u) \approx \frac{1}{2} \sum_{u=1}^U \sqrt{\mathbf{p}_u(\mathbf{y}_0), \mathbf{q}_u} + \frac{C_h}{2} \sum_{i=1}^n w_i k \left(\left\| \frac{\mathbf{y} - \mathbf{x}_i}{h} \right\|^2 \right) \quad (19)$$

ここで $\mathbf{y}_0$, $\mathbf{x}_i$, $\mathbf{y}$ は $k+1$ フレームにおける初期地点、その近傍のターゲット候補点、及びミーンスフト変数を示す。また $k(\mathbf{x})$ は幅 h のカーネルプロファイルで、 C_h は定数項である。上式の右辺第一項は変数 $\mathbf{y}$ に依存しないので、 $\rho(\mathbf{p}_u(\mathbf{y}), \mathbf{q}_u)$ の最大化には、右辺の第二項のみを考慮すればよい。この重み付きカーネル和関数は、2.3 節で説明した式 (3) による Cheng のミーンスフト公式で最適化する。ここで正值の重み関数 w_i は次のように定義される。

$$w_i = \sum_{u=1}^U \delta(b(\mathbf{x}_i) - u) \sqrt{\frac{\mathbf{q}_u}{\mathbf{p}_u(\mathbf{y}_0)}} \quad (20)$$

ただし δ はデルタ関数、 $b(\mathbf{x})$ は点 $\mathbf{x}$ での輝度値を指す。以上のアルゴリズムの手順を次に示す。

ミーンスフトの物体追跡 [各フレーム k で]:

- (1) 初期値 $\mathbf{y}_{k,0}$ を前物体位置 $\mathbf{y}_{k-1,0}$ で設定
- (2) $\mathbf{y}_0$ において $\mathbf{p}_u(\mathbf{y}_0), \mathbf{q}_u, \rho(\mathbf{p}_u(\mathbf{y}_0), \mathbf{q}_u)$ の計算
- (3) $\mathbf{y}_0$ 近傍の候補各点で w_i を式 (20) より計算
- (4) 式 (3) でミーンスフト $\mathbf{y}_1$ を計算
- (5) $\mathbf{y}_1$ において $\rho(\mathbf{p}_u(\mathbf{y}_1), \mathbf{q}_u)$ を計算
- (6) $\rho(\mathbf{p}_u(\mathbf{y}_1), \mathbf{q}_u) < \rho(\mathbf{p}_u(\mathbf{y}_0), \mathbf{q}_u)$ の場合 $\mathbf{y}_1 \leftarrow \frac{1}{2}(\mathbf{y}_0 + \mathbf{y}_1)$ とする
- (7) $\mathbf{y}_0 \leftarrow \mathbf{y}_1$ とし $\mathbf{y}$ の収束まで (4)-(6) を繰返

以上説明した基本的なミーンスフト物体追跡アルゴリズムの性能向上に向けて、すでに多くの研究結果が報告されている。主なものには、複数の追跡推定結果をロバストに情報統合することによるモーション解析精度の向上 [11]、AdaBoost とミーンスフトを併用したアンサンブル物体追跡 [1]、ミーンスフトを使った高速な大域的最大値探索法 [44]、非対称のカーネルに拡張したミーンスフト物体追跡 [58] などがある。

6.3 医用画像解析

医用画像解析はビジョン研究の重要な応用分野

のひとつで、ビジョン研究初期の頃から盛んな分野であるが、最近の三次元医用画像技術の進歩に伴い、より高度な情報処理が必要となってきた。ミーンスフトの研究発信地のひとつであるシーメンズプリンストン研究所が、医用画像解析応用研究を行っていたこともあり、早くからこの分野でのミーンスフトの応用が盛んとなった。

コントラストを向上させた CT 画像における血管の認識 [49] や、一般 CT 画像における体内臓器の三次元領域分割 [48] に向けて、初期地点から円状に走査線を定義し、その走査線上の輝度値分布が大きく変化する地点の位置決め、ミーンスフトを応用した手法が報告された。また肺がんの早期発見に向けた胸部 CT 画像によるスクリーニングの開発に向けて、スケールスペースに拡張したミーンスフトを応用した、肺がんの陰影である肺結節のロバスト解析法 [33], [35] が提案された。さらに全身 PET-CT 画像における病巣変化の定量解析に向けて、可変カーネル幅ミーンスフトによるロバスト情報統合を応用した、三次元対応点の推定法 [36] や、皮膚がんの診断に向けた、ミーンスフトによる体後背部二次元画像上のほくろの自動解析法 [29] が報告されている。

臨床医用画像解析問題以外にも、時系列に撮像された顕微鏡動画上でのがんやその他の細胞の追跡にミーンスフトが応用されている [16], [57]。また analytic manifold へ拡張されたミーンスフトを応用した、統合的な脳研究と診断に向けて重要な拡散テンソル MRI 画像の平滑化手法 [47] も興味深い。

6.4 その他

本節で述べた、画像領域分割、物体追跡、医用画像解析のような、ミーンスフトの応用として比較的よく知られているもの以外にも、広範なビジョン問題への応用成功例が報告されている。テキスト解析に向けて、前述したミーンスフトセグメンテーションを、高次元特徴ベクトル (texton) へ応用 [23] した報告がある。この手法は、LSH (locality sensitive hashing) を用いた高次元空間での計算量の爆発への対応を提案した。ま

た前述した CAMSHIFT と SVM を併用したテクスチャ解析による、画像における文書部位の推定法 [27] も報告された。その他のビジョン問題への応用として、ミーンスフトによる motion field 表現を使ったカメラモーションの認識の研究 [17] やノンパラメトリックな appearance model を用いた物体認識と向き推定研究 [51] が報告されている。

より一般的な工学問題への応用例に次のようなものがある。マルチメディアの研究として、ミーンスフトクラスタ解析を応用した、映画の類型の認識 [41] や、CG の研究として、ミーンスフト領域分割を応用した、画像の概要をまとめる低解像度画像の自動生成である画像 retargeting の開発 [42]、またインテリジェント交通システム開発における、ミーンスフト物体追跡を応用した、車載の night-vision カメラ動画中の歩行者の追跡 [55] や、可変カーネル幅ミーンスフトによる統計的ロバスト情報統合を応用した、追い抜き車両の自動認識 [61] などが報告されている。

7. む す び

本稿はミーンスフト法の基礎理論の解説とその応用の概説について述べた。本研究分野では、現在活発な研究がなされており、最近の発表論文の数もかなり多く、これから更に色々な分野への応用開発が期待できる。

7.1 実装上のヒント

実装上の大きな問題の一つに計算量の問題がある。多くの応用問題の拘束として、実時間処理が求められたり、計算量の減少が好ましい場合が多々ある。本稿で解説した理論的な解決法とは別に、より効率的なシステムの構築に向けて、適当な近似的処理をソフトウェアの実装の段階で導入することは重要である。

そのようなアイデアのひとつとして、計算するミーンスフトベクトルを逐次にメモリテーブル保存する方式が考えられる。複数のサンプル点からのミーンスフトを適用する場合に、まず特徴空間を適当なサイズの幅で離散化し、各ミーンスフトベクトルの計算ごとに、繰り返し変数の現在位置

にあたるテーブルエントリを参照し、もしエントリがすでに存在すれば使用し、なければ書き込むことを繰り返す。近傍の初期値からのミーンスフト収束は同じようなパスを取る可能性が高いので、メモリの使用量を抑えながら離散化誤差を少なくするよう幅を上手く設定すれば、ミーンスフトベクトル計算を減らすことによる高速化が得られる。この手法は特徴空間が低次元で局所的な場合に有効である。また、ガウスカーネルのような空間連続なカーネルを使用する場合、理論的には定義した空間変数領域全体でのカーネル和計算が必要となるが、カーネル幅の定数倍の領域での和の計算をすることにより、近似的に正しいミーンスフトの高速計算が可能となる。

7.2 実装例の公開サイト

実装例として、Rutger 大学の CAIP 研究室のサイト [62] で、B.Georgescu 氏のミーンスフトによるセグメンテーションとクラスタリングの C++ コードが公開されている。Matlab 用のコードもインターネット上で入手可能で、B.Finkston 氏の基本的なミーンスフトクラスタ解析のコードなどが Matlab のサイト [63](キーワード”mean shift”でサーチする)で公開されている。Intel 社が提供している OpenCV は、コンピュータビジョン用のまとまったオブジェクトライブラリで、CAMSHIFT [3] をもちいた物体追跡の実装例が含まれている。このライブラリは [64] から入手可能である。最後に、D.Comaniciu 氏のホームページ [65] には、本稿で解説した多数の論文の他にも、実験結果のビデオ等が公開されている。

参考文献

- [1] S. Avidan. Ensemble tracking. *IEEE Trans. Pattern Anal. and Machine Intell.*, 29(2):261–271, Feb. 2007.
- [2] M. J. Black, G. Sapiro, D. Marimont, and D. Heeger. Robust anisotropic diffusion. *IEEE Trans. Image Processing*, 7(3):421–432, 1998.
- [3] G. R. Bradski. Computer vision face tracking for use in a perceptual user interface. *Intel Technology Journal*, 2nd Quarter, 1998.
- [4] Haifeng C. and P. Meer. Robust fusion of uncertain information. *IEEE Trans. Systems Man Cybernetics Part. B*, 35(3):578–586, June 2005.

- [5] M. A. Carreira-Perpinan. Acceleration strategies for gaussian mean-shift image segmentation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2006.
- [6] M. A. Carreira-Perpinan. Gaussian mean-shift is an EM algorithm. *IEEE Trans. Pattern Anal. and Machine Intell.*, 29(5):767–776, 2007.
- [7] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. and Machine Intell.*, 17(8):790–799, 1995.
- [8] R. T. Collins. Mean-shift blob tracking through scale space. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages II:234–240, 2003.
- [9] D. Comaniciu. An algorithm for data-driven bandwidth selection. *IEEE Trans. Pattern Anal. and Machine Intell.*, 25(2):281–288, 2003.
- [10] D. Comaniciu. Kernel-based object tracking. *IEEE Trans. Pattern Anal. and Machine Intell.*, 25(5):564–575, 2003.
- [11] D. Comaniciu. Nonparametric information fusion for motion estimation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages I:59–66, 2003.
- [12] D. Comaniciu and P. Meer. Mean shift analysis and applications. In *Proc. IEEE Int. Conf. Computer Vision*, pages 1197–1203, 1999.
- [13] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. and Machine Intell.*, 24(5):603–619, 2002.
- [14] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 142–149, 2000.
- [15] D. Comaniciu, V. Ramesh, and P. Meer. The variable bandwidth mean shift and data-driven scale selection. In *Proc. IEEE Int. Conf. Computer Vision*, pages I:438–445, 2001.
- [16] O. Debeir, P. Van Ham, R. Kiss, and C. Decaestecker. Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes. *IEEE Trans. Medical Imaging*, 24(6):697 – 711, 2005.
- [17] L. Y. Duan, J.S. Jin, T. Qi, and C.-S. Xu. Nonparametric motion characterization for robust classification of camera motion patterns. *IEEE Trans. Multimedia*, 8(2):323–340, April 2006.
- [18] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, Inc., New York, 1973.
- [19] M. Fashing and C. Tomasi. Mean shift is a bound optimization. *IEEE Trans. Pattern Anal. and Machine Intell.*, 27(3):471–474, 2005.
- [20] L. Friedman, N. Netanyahu, and M. Shoshany. Mean shift-based clustering of remotely sensed data. In *IEEE Int. Geoscience and Remote Sensing Symp.*, pages 3432–3434, 2003.
- [21] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, San Diego, 1990.
- [22] K. Fukunaga and L. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Information Theory*, 21(1):32–40, 1975.
- [23] B. Georgescu, I. Shimshoni, and P. Meer. Mean shift based clustering in high-dimension: a texture classification example. In *Proc. IEEE Int. Conf. Computer Vision*, pages I:456–463, 2003.
- [24] K. Inoue and K. Urahama. Hierarchically distributed dynamic mean shift. In *Proc. IEEE Int. Conf. Image Processing*, pages I:269–272, 2007.
- [25] A. Jagmohan, M. Singh, and N. Ahuja. Dense Stereo Matching Using Kernel Maximum Likelihood Estimation. In *Proc. IARP Int. Conf. Pattern Recognition*, pages III:28–31, 2004.
- [26] T. Jebara and A. Pentland. Maximum conditional likelihood via bound maximization and the CEM algorithm. In *Neural Information Processing Systems 11*, pages 494–500, 1998.
- [27] K. I. Kim, K. Jun, and J. H. Kim. Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm. *IEEE Trans. Pattern Anal. and Machine Intell.*, 25(12):1631 – 1639, 2003.
- [28] J. J. Koenderink. The structure of images. *Biol. Cybern.*, 50:363–370, 1984.
- [29] T.K. Lee, M.S. Atkins, M.A. King, S. Lau, and D.I. McLean. Counting moles automatically from back images. *IEEE Trans. Biomedical Engineering*, 52(11):1966–1969, Nov. 2005.
- [30] T. Lindeberg. Feature detection with automatic scale selection. *Int. J. Computer Vision*, 30(2):79–116, 1998.
- [31] D. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Computer Vision*, 60(2):91–110, 2004.
- [32] K. Okada, D. Comaniciu, N. Dalal, and A. Krishnan. A robust algorithm for characterizing anisotropic local structures. In *Proc. European Conf. Computer Vision*, pages I:549–561, 2004.
- [33] K. Okada, D. Comaniciu, and A. Krishnan. Robust 3D segmentation of pulmonary nodules in multislice CT images. In *Proc. Int. Conf. Medical Imaging Computing and Computer-Assisted Intervention*, pages II:881–889, 2004.
- [34] K. Okada, D. Comaniciu, and A. Krishnan. Scale selection for anisotropic scale-space. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages I:594–601, 2004.

- [35] K. Okada, D. Comaniciu, and A. Krishnan. Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT. *IEEE Trans. Medical Imaging*, 24(3):409–423, 2005.
- [36] K. Okada and X. Huang. Robust click-point linking: Matching visually dissimilar local regions. In *IEEE Int. Workshop Beyond Multiview Geometry (CVPR07)*, 2007.
- [37] K. Okada, M. Singh, and V. Ramesh. Prior-constrained scale-space mean shift. In *Proc. British Machine Vision Conf.*, pages II:829–838, 2006.
- [38] S. Paris and F. Durand. A topological approach to hierarchical segmentation using mean shift. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 17–22 June 2007.
- [39] E. Parzen. On estimation of a probability density and mode. *Annals of Mathematical Statistics*, 33:1065–1076, 1962.
- [40] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. and Machine Intell.*, 12(7):629–639, 1990.
- [41] Z. Rasheed, Y. Sheikh, and M. Shah. On the use of computable features for film classification. *IEEE Trans. Circuits and Systems for Video Technology*, 15(1):52–64, Jan. 2005.
- [42] V. Setlur, T. Lechner, M. Nienhaus, and B. Gooch. Retargeting images and video for preserving information saliency. *IEEE Computer Graphics and Applications*, 27(5):80–88, Sept.–Oct. 2007.
- [43] S. J. Sheather and M. C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *J. Royal Statistical Soc. Series B*, 53:683–690, 1991.
- [44] C. Shen, M. J. Brooks, and A. van den Hengel. Fast global kernel density mode seeking: Applications to localization and tracking. *IEEE Trans. Image Processing*, 16(5):1457–1469, May 2007.
- [45] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, 1986.
- [46] M. Singh, H. Arora, and N. Ahuja. A robust probabilistic estimation framework for parametric image models. In *Proc. European Conf. Computer Vision*, pages I:508–522, 2004.
- [47] R. Subbarao and P. Meer. Discontinuity preserving filtering over analytic manifolds. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 17–22 June 2007.
- [48] H. Tek, M. Bergholdt, D. Comaniciu, and J. Williams. Segmentation of 3D medical structures using robust ray propagation. In *Proc. Int. Conf. Medical Imaging Computing and Computer-Assisted Intervention*, pages I:572–579, 2002.
- [49] H. Tek, D. Comaniciu, and J. Williams. Vessel detection by mean shift based ray propagation. In *IEEE Int. Workshop Mathematical Methods in Biomedical Image Analysis (CVPR01)*, 2001.
- [50] B. A. Turlach. Bandwidth selection in kernel density estimation: A review. Technical report, Louvain-la-Neuve, Belgium, Institut de Statistique, 1993.
- [51] T. Vik, F. Heitz, and P. Charbonnier. Robust pose estimation and recognition using non-gaussian modeling of appearance subspaces. *IEEE Trans. Pattern Anal. and Machine Intell.*, 29(5):901–905, May 2007.
- [52] M. J. Wainwright and M. I. Jordan. Graphical models, exponential families, and variational inference. Technical Report 649, Department of Statistics, University of California, Berkeley, 2003.
- [53] M. P. Wand and M. C. Jones. *Kernel Smoothing*. Chapman & Hall, 1995.
- [54] A. Witkin. Scale-space filtering. In *Int. Joint. Conf. Artificial Intell.*, pages 1019–1021, Karlsruhe, 1983.
- [55] F. Xu, X. Liu, and K. Fujimura. Pedestrian detection and tracking with night vision. *IEEE Trans. Intelligent Transportation Systems*, 6(1):63–71, March 2005.
- [56] C. Yang, R. Duraiswami, N. Gumerov, and L. Davis. Improved fast gauss transform and efficient kernel density estimation. In *Proc. IEEE Int. Conf. Computer Vision*, pages 464–471, 2003.
- [57] X. Yang, H. Li, and X. Zhou. Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and kalman filter in time-lapse microscopy. *IEEE Trans. Circuits and Systems*, 53(11):2405–2414, 2006.
- [58] A. Yilmaz. Object tracking by asymmetric kernel mean shift with automatic scale and orientation selection. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 1–6, 17–22 June 2007.
- [59] X. Yuan and S. Z. Li. Half quadratic analysis for mean shift: with extension to a sequential data mode-seeking method. In *Proc. IEEE Int. Conf. Computer Vision*, 14–21 Oct. 2007.
- [60] K. Zhang, J. T. Kwok, and M. Tang. Accelerated convergence using dynamic mean shift. In *Proc. European Conf. Computer Vision*, pages 257–268, 2006.
- [61] Y. Zhu, D. Comaniciu, M. Pellkofer, and T. Koehler. Reliable detection of overtaking vehicles using robust information fusion. *IEEE Trans. Intelligent Transportation Systems*, 7(4):401–414, Dec. 2006.
- [62] <http://www.caip.rutgers.edu/riul/research/code.html>.
- [63] <http://www.mathworks.com/matlabcentral/>.
- [64] <http://www.intel.com/technology/computing/opencv/>.
- [65] <http://www.caip.rutgers.edu/comanici/>.