

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
kazokada@sfsu.edu

平成 21 年 9 月 4 日

ミーンシフトは、ロバストな統計的データ解析手法で、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。この手法は特に統計学上で理論的である反面、実装が比較的簡易で応用範囲が広いことが特徴として挙げられる。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。

1 はじめに

ミーンシフト (mean shift) は、統計学においてよく知られるカーネル密度推定 (kernel density estimation) [49] を用いたロバストなデータ解析手法で、福永と Hostetler [24] 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 [9] を皮切りに注目され、特に Comaniciu らによる一般的で正確な定式化 [14, 15] 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 (density function) の極値探索問題について最急勾配法 (gradient ascent method) 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、また情報統合をより頑強に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 [14, 15, 22, 27, 50, 8, 41, 53]、ビデオ上の物体トラッキング [16, 12, 13,

67, 18, 69, 48, 1, 70]、クラスタ解析 [24, 23, 9, 25, 44]、医用画像解析 [57, 56, 36, 38, 18, 31, 69, 53, 39] テクスチャ解析 [25, 29]、モーション解析 [59, 19, 54, 52]、ハフ変換・情報統合 (information fusion) [9, 13, 8, 73]、物体認識 [61, 60]、マルチメディア [44]、コンピュータグラフィックス [45]、インテリジェント交通システム [67, 73] 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。主な理論的研究結果の報告には、カーネル回帰分析とロバスト M 推定との関連 [15]、界関数最適化との関連 [21, 50]、EM アルゴリズムとの関連 [4, 6]、 k 平均クラスタ解析との関連 [9]、half-quadratic 最適化との関連 [71]、可変カーネル幅への拡張 [17, 11, 13]、スケールスペースへの拡張 [10, 35]、事後確率推定への拡張 [40, 61]、動的メーンシフトへの拡張 [3, 72, 26, 4, 5, 7]、非線形多様体空間への拡張 [59, 54, 53, 51, 52, 55]、メドイドシフト (medoidshift) への拡張 [47, 60]、などが挙げられる。

本稿はこのメーンシフト法の総合的な理解に向けて、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、メ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

本節では、メーンシフト法の基本アルゴリズムを、まず簡単に説明する。

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。本稿で紹介するメーンシフト (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 $S = \{x_i | i = 1, \dots, n\}$ 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 $f(x \in X)$ 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 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 $\nabla_x f(x) = 0$ 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 $x_0 \in X$ 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 $s(x_j)$ とすると、初期値 x_0 から次の手続きを反復する

一般的極大値探索法:

以下の手続きを収束するまで繰り返す: $j = 0, 1, 2, \dots$

(1) x_j におけるステップの計算: $s(x_j)$

(2) 現在位置の更新: $x_{j+1} \leftarrow x_j + s(x_j), j \leftarrow j + 1$

一般に極大値探索法は、この反復の繰り返しにより x_0, x_1, \dots, x_j が逐次的に収束し、かつその収束値 $f(x_j)$ が最頻値となるように、繰り返しステップ $s(x_j)$ を目的関数 $f(x)$ 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の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 [24, 23] 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 $S = \{x_i\}$ が与えられたとする。次に任意の点 μ を中心として半径 h の円領域 $T_h(\mu)$ を考え、この円領域内にあるサンプル点の平均値 $m(\mu)$ を計算する。福永のミーンシフトは、現在位置 x におけるこの平均値 $m(x)$ と x との差分ベクトルを繰り返しステップ $s(x) = m(x) - x$ 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すことで、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 [23]:

以下の手続きを収束するまで繰り返す: $j = 0, 1, 2, \dots$

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

ここで $|T_h(x)|$ は 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 $T_h(\mu)$ を次のような関数 $K_f(x; \mu, h)$ として数式で表す。

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

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。

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

ここでは、 n がサンプル S 中のサンプル点の総数をさし、数列和の範囲が 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 [9] 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 $K(x; \mu, h)$ で、また非負値の重み関数を $w(x) : S \rightarrow (0, \infty)$ で表すと、サンプル x_i の重み付き平均が次のように定義できる。

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

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 [9]:

以下の手続きを収束するまで繰り返す: $j = 0, 1, 2, \dots$

(1) 式 (3) による重み付き平均値 $m(x_j)$ の計算

(2) 現在位置の更新: $x_{j+1} \leftarrow m(x_j), j \leftarrow j + 1$

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 $f(x)$ のグラジエント (空間偏導関数ベクトル) $\nabla_x f(x)$ とステップ幅を表す正定数 γ の積として定義する。

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、適当な値に手動で設定するのが煩雑また困難になる。これに比べ、ミーンシ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

最急勾配法:

以下の手続きを収束するまで繰り返す: $j = 0, 1, 2, \dots$

- (1) x_j でのステップの計算: $s(x_j) = \gamma \nabla_x f(x_j)$
- (2) 現在位置の更新: $x_{j+1} \leftarrow x_j + s(x_j), j \leftarrow j + 1$

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シフトの利点として、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 ミーンシ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シフトの最も基本的で汎用な応用問題である。一般的にクラスタ解析とは、与えられたサンプル集合 $S = \{x_i \in X\}$ 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。ミーンシフトを用いた手法は基本的に次の手順に従う。

ミーンシフトクラスタリング [9, 14]:

- (1) n 個の初期値 $T = \{y_i \in X\} \leftarrow S$ の設定
- (2) 各初期値 y_i からミーンシフトを実行
- (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 y_i を収束先の最頻値点のインデックス c でラベル付けする

この手法を要約すると、まず各サンプル点を初期値としてミーンシフトを実行し、その結果同じ点に収束したサンプルの部分集合を一つのクラスタとみなすことで、サンプル集合のグループ分けをする。このため、ミーンシフトを用いたクラスタ解析におけるクラスタ数 C は、3.1 節に後述するカーネル密度関数の最頻値の数 (確率分布中のピークの数) に対応するため、データから

直接決まる。具体的には、上記の手順において、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 $x_i \leftarrow y_i$ にミーンシフトの収束点に基づくラベル付け $x_i \leftarrow c$ 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、ミーンシ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シフトの繰り返しに影響されない。この方式は、前述した Cheng [9] 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シ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。

ミーンシ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手法は初期値集合 T をサンプル集合 S で置き換えること ($T = S$) により求まり、福永らのクラスタ解析はこの方式をとった。最近になって、動的ミーンシフト (dynamic mean shift) [3, 72, 26] やガウス拡散ミーンシフト (Gaussian blurring mean shift) [4, 5, 7] により、この手法が収束の高速化に向けて再認識されている。これらの拡張法については、5.4 節で触れる。

2.6 他のクラスタ解析法との比較

ミーンシ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。言い換えると、特徴空間中でのクラスタ形状が標準的な楕円型以外の場合でも、ミーンシフトによるクラスタ解析は効果的である。

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シフトクラスタ解析法として解釈できる [9]。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より頑強にしたものとして、ミーンシ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、ミーンシフト法は、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は注意を要する。3.4 節でカーネル幅の自動推定法について述べる。

3 ミーンシフトの理論

本節では、ミーンシ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 [14, 15] 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) [49] は、Parzen windows [42, 20] 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 $S = \{x_i | i = 1, \dots, n\}$ 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。

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

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 $H = hI$ と簡略化して考える。厳密な d 次元のカーネルは有界でコンパクトな台を持つ対称性を持った関数として定義され、[64] 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。基本概念の理解のため簡単に言えば、カーネル関数は対称軸中心から関数値が急激に等速で減少する関数又は局所的な注目領域を持つ重み分布と考えてよい。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 $K(\mathbf{x}; x_i, h) = K(\frac{\mathbf{x}-x_i}{h})$ 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 $\|\mathbf{x}\|^2 = \sum_{l=1}^d |x_l|^2$ 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、

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

ガウス関数による正規分布カーネル K_g や、

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

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、ミーンシ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シフト

ここでミーンシフトの公式 (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 $g(x) = -k'(x)$ で置き換えて上式を次のように変形する。

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

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

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

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

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

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

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

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。

$$\nabla_x f_K(x) = \frac{2}{h^2} f_G(x) s_G(x) \quad (17)$$

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。

$$s_G(x) = \frac{1}{2} h^2 \frac{\nabla_x f_K(x)}{f_G(x)} \quad (18)$$

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 $\nabla_x f_K(x)$ 以外はすべて正值のスカラー項な

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 $\nabla_x f_K(x) = 0$ 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 [9, 15]。

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_f と異なる K_e によって定義され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、底が e である指数関数の導関数はそれ自身であるため $g(x) = 0.5k(x)$ 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、式 (18) より、ステップ幅 γ を変数項 $\frac{h^2}{2f_G(x)}$ 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 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 [15] にある定理がもっとも一般的に知られている。

ミーンシフトの収束定理 [15, 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 $\{x_j\}_{j=0,1,\dots}$ と対応する密度推定値の数列 $\{f_K(x_j)\}_{j=0,1,\dots}$ が収束し、かつ密度推定値が単調に増加すること $f_K(x_j) \leq f_K(x_{j+1}) \forall j$ 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。証明の詳細については [15] と Li らによる最近の報告 [32] を参照されたい。

また実装上では、繰り返し計算を終了させるための収束判定が必要となるが、ミーンシフトベクタのノルムが一定の閾値以下になるかどうか検定する手法が一般的である。

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。カーネル幅の値の変化は、3.1 節で述べたカーネル密度推定の結果に大きく影響する。カーネル幅の値が小さい場合、対応する分布関数には多数の最頻値が存在するが、値を増加させるにつれ最頻値の数は単調に減少し、最終的にはもとのデータに関わらず一つの最頻値しか持たないユニモーダルな分布に収束してしまう。ミーンシフト法は、基本的に与えられたデータにより定義されるカーネル密度関数の最頻値点を探索する手法であるので、その目的関数の極値の数と極値点自体を変えてしまうカーネル幅の設定には当然注意を要する。

統計学的手法では、未知の密度関数 $f(x)$ とそのカーネル推定 $\hat{f}(x)$ 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 [49, 64, 46, 58]。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 [17] 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 [37] 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 [34] の基礎である Lindeberg の自動スケール決定法 (automatic scale selection) [33] 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 [23], 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 [11, 35, 37, 38]。

カーネル幅の自動推定の応用研究も若干報告されている。若原と小倉 [63] は、カーネル幅をサンプルデータから直接求め、カーネル幅を順に増大させる階層的クラスタ解析 (hierarchical clustering analysis) を提案した。その結果は手書き文字のクラスタ解析に応用されている。Wang と Suter [65] は、間違っただけのピーク (false peak) にミーンシフトが収束する条件を、カーネル幅の範囲とし

て解析的に導出した。この結果と、ミーンシフトによる最小値探索 (mean shift valley algorithm) を用いて、より安定な画像の領域分割アルゴリズムが提案されている。

これらの研究結果に関わらず、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。ここで与えられたデータ $S = \{x_i\}$ を、確率密度関数 $p(X)$ に従う確率変数 X から抽出した独立で同じ確率分布に従う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。

簡単な例を使って概念的に説明すると、系列 $\{3, 4, 5, 8000\}$ の平均値は 2003 であるが、最終項の 8000 を外値であるとみなすと、実際の平均値は 4 であると考えられる。ここで、任意値の近傍のみを考慮する局所的な重み付き平均計算を導入し、4 の近傍における計算結果を考える。この場合、重み分布の注目領域の幅が大きすぎれば、平均値と同様の結果となるが、注目領域の幅を 4000 以下に減らせば、実際の値 4 を推定できる。カーネル法は、この平均値計算のような統計値推定から外値の影響を減らすため、局所的な注目領域をもつカーネル関数を導入する手法と理解してよい。

ここで、もとの分布 $p(X)$ のモード位置をパラメタ y として、尤度分布 (likelihood distribution) を $p(x_1, \dots, x_n | y)$ 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 [15, 50]。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 [15]。

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 [13, 8] を提案した。この手法は、各データ点 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 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 [28, 62] を用いたものがある [21, 50]。Fashing と Tomashi [21] 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 [40]。

最近になって、高速ガウス変換を応用したミーンシフトの高速化が Yang ら [68] により、正規分布カーネル K_g を用いる場合のミーンシフト法と 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 [4, 6] により、また half-quadratic 最適化の枠組みでのミーンシフトの理論的解析とその高速化が Yuan と Li [71] により報告されている。

最後に、非線形空間へのミーンシフトの理論的拡張が報告されている。Subbarao ら [59, 54, 53, 51, 52, 55] による Riemannian manifolds へのミーンシフトの拡張や、Sheikh ら [47] によるミーンシフトを一般的な距離空間 (metric space) に拡張したメドイドシフト (medoidshift) などが興味深い。これらの非線形ミーンシフトについては、5 節にて触れる。

4 ミーンシフトの利点と欠点

以上説明した基本的なミーンシ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。

ここに欠点としてあげたミーンシ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 [53, 48, 6, 71]。

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直接計算せずに、最頻値が探索できる。
- (c) 最勾配法にあるステップ幅 γ パラメタを、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、大局的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。

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 $h(x)$ 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 [17, 11, 13] 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 x を中心に置いたカーネルの x_i データ各点における n 個の関数値 $K(x_i)$ の和としてでなく、 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 x における関数値 $K(x)$ の和としてみなし、空間上に配置したカーネル幅 $h(x)$ 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 [11]。

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 (information fusion) [13, 39] 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 (scale space) [66, 30] 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等方平滑化などへの応用も知られている [43, 2]。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の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 [10] による提案と、岡田ら [35] による非等方カーネルへの拡張が知られている。ガウス関数による画像 $I(x)$ の平滑化には、正規分布カーネル K_g による入力の畳み込み積分で行う。

$$L(x) = \int_{\Omega} I(x') K_g(x; x', h) dx' \quad (19)$$

ただし Ω は画素値すべての集合をさす。この積分式を離散化したもの

$$L(x) = \sum_{i=1}^n I(x_i) K_g(x; x_i, h) \quad (20)$$

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 $I(x)$ を重み $w(x)$ 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値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。

この拡張は、ミーンシフトがカーネル密度関数以外の任意の正值関数に適用出来ることを指すため理論的に重要で、4 節で挙げた、カーネル和で表せられる目的関数にしか適用できないという欠点 (a) への対処としても理解できる。

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 [40] は、推定変数 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、ベ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。[40] 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。

Vik ら [61] は、入力データ y に統計的アピアランスモデルを当てはめる問題を、最大事後確率推定問題として定義し、ミーンシフトと似た繰り返しアルゴリズムがその解法として導出されることを報告した。具体的には、物体認識問題で良く使われる統計的アピアランスモデルとして、因子分析 (factor analysis) における潜在変数 (latent variable) x 上に定義される、ノンパラメトリックなカーネル密度推定 $p(x)$ をまず導入した。次に事後確率分布 $p(x|y)$ を解析的に求め、それを直接最大化する x をミーンシフトを用いて求めた。この手法は複数のカメラ視点から捉えた物体画像を用いた物体認識問題に応用されている。

5.4 動的ミーンスフトへの拡張

2.5 節で解説したミーンスフトクラスタリング法において、データ各点からミーンスフトを行い任意 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、ミーンス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、ミーンスフトの繰り返し毎に動的に変化するので、動的ミーンスフト [72, 26] 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 [24]。この手法では、ミーンスフトの繰り返し毎に、 S 中の各サンプル点 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。

最近になって、同様な手法が Carreira-Perpinan によりガウス拡散ミーンスフト (Gaussian blurring mean-shfit) [5, 7] として報告されている。この報告では、一般的なミーンスフトが一次収束であるのに対し、ガウス拡散ミーンスフトは二次収束できることが示されている。

5.5 非線形空間への拡張

これまでに扱ってきたミーンスフトはユークリッド空間中に定義されているので、非線形空間に分布するデータ点に直接応用することはできない。例えば、球表面上に分布するデータ点の平均値は、同じ球面上には常に求まらない。このことからミーンスフトの結果はこの球面上に拘束されず、よってミーンスフトを使って球面上のデータの最頻値点が求まらないことが分かる。

Subbarao, Tuzel と Meer の一連の研究 [59, 54, 53, 51, 52, 55] は、微分幾何学を導入して、より一般的な曲面上でのミーンスフトの拡張を報告した。ある曲面について、その局所座標 (coordinate charts) 間の座標変換 (transition map) の総てが解析関数として表すことが出来る場合、その曲面は解析多様体 (analytic manifold) と呼ばれる。Subbarao らは、データ空間がこの解析多様体で表現できる場合において、その空間上におけるミーンスフトを導出した。具体的には、リー群 (matrix Lie group) [59]、グラスマン多様体 (Grassmann manifolds) [54]、正値対称行列による多様体 (symmetric positive definite matrices) [53]、基本行列による多様体 (essential manifolds) [52] 上での結果がそれぞれ報告された。さらに、これらの様々の多様体のタイプを含む、より一般的なリーマン多様

体 (general Riemannian manifolds) [55] 上でのメーンシフトの定義が示されている。

これらの非線形空間上でのメーンシフトは、非ベクトル空間上でのデータ解析問題、例えばモーション解析 [59, 52]、拡散テンソル MRI 画像の平滑化手法 [53] などに応用できる。

5.6 メドイドシフトへの拡張

メドイドシフト (medoidshift) [47] は、最近提案された最頻値探索アルゴリズムで、ユークリッド距離以外のより一般的な距離空間へのメーンシフトの拡張として報告された。ここでいうメドイド (medoid) は、他のデータ点への距離の平均が最小になるようなデータ点として定義され、k-平均クラスタ解析の拡張である、k-メドイドクラスタ解析にも見られる。メドイドシフト法は、局所的グラジエントを、平均値を用いずに重み付きメドイドを使って近似するので、データ間の距離が存在しているだけで平均値が定義されていない空間上でも適応可能である。このことから、5.5 節で触れた解析的多様体へのメーンシフトの拡張と同様に、メーンシフトの非線形空間へのより柔軟な拡張を提供する。同じ理由から、この手法の収束点は常に与えられたデータ点の一つとして拘束される。このため、収束条件の設定や繰り返し計算の必要がなくなり、メーンシフトに比べより簡易な実装が可能となるが、Sheikh ら [47] の報告では、画像の領域分割応用において、メーンシフトに比べ計算量が多くなることが欠点として指摘された。

Vedaldi と Soatto [47, 60] は、ユークリッド距離を用いた場合、Sheikh らの報告に反して、メドイドシフトはメーンシフトより計算量が少ないことを示した。またメドイドシフトの欠点として、総てのモードを常に検出できないことを示した。この欠点を、ユークリッドメドイドシフトの効率性を生かしながら解消するために、ユークリッドメドイドシフトの非線形拡張であるクイックシフトが提案された。この非線形拡張には、サポートベクタマシンなどに用いられる Mercerl kernel によるカーネルトリックが使われている。また高速化のためには、貪欲法 (greedy algorithm) が用いられた。Vedaldi と Soatto により、この手法のシーンのカテゴリ認識への応用が報告されている。

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、均一 (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 $I(x)$ の画素が M 個あるとし、これを $(x_1, f_1), \dots, (x_m, f_m), \dots, (x_M, f_M)$ のサンプルベクトル集合として考える。ここでベクトル f_m は画素 x_m における画素値 $f_m = I(x_m)$ を表す。入力画像がグレースケールである場合には、 $\{f_m\}$ 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、 $\{f_m\}$ が三次元のベクトル点集合となる。ここでは、画素 x_m における特徴量 y_m を対応する画素値で設定する $y_m = f_m$ 。ミーンシフトによる領域分割は、この特徴量サンプル $T = \{y_m\}$ 、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 x_m のラベル L_m を出力として、次の手順で行う。

ミーンシフトの領域分割:

- (1) M 個の y_m からミーンシフトを実行
- (2) 結果 M 個の収束点 z_m を得る
- (3) 距離が h 以下の近傍の点 z_m をグループ化
- (4) 結果 $\{z_m\}$ を C 個のクラスタ $\{\mathcal{C}_c\}$ に分割
- (5) 各画素 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 $L_m = c | z_m \in \mathcal{C}_c$

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 [14, 15] 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。このため、距離が大きく離れているが偶然に輝度値が似ている画素を同領域に分割してしまうことを防ぐことができる。この方式は、画素 x_m における特徴量 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 $y_m = (x_m, f_m)$ 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、輝度値の三次元と空間変数の二次元を合わせた五次元となる。

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 [14, 15]。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 z_m を用いて、 z_m の画素値を、対応する初期値画素 x_m における画素値 f_m として代入する。各クラス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 [22] やステレオ視によるレンジデータ [27, 50, 8] 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 [41] が提案されている。

6.2 動画像におけるトラッキング

Comaniciu らによって提案されたミーンシフトトラッキング [16, 12] 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ト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 [16, 12, 13, 18, 69, 67, 70, 48, 1]。

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 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 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 $x_{k+1,0}$ を、ミーンシフト法の初期値 y_0 と設定し、このミーンシフトでの収束点 y_j をターゲット点 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 x_k を中心に抽出するカラー画素値のヒストグラム q_u と最も似たカラーヒストグラム $p_u(y)$ を供給する、 $k+1$ 番フレームでの位置 y_j を、 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シフトが収束するので、実時間の物体追跡が可能となる。またミーンシ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 $p_u(y)$ と 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 $\rho(p_u(y), q_u) = \sum_{u=1}^U \sqrt{p_u(y)q_u}$ を最大化する 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 [16] の導出に従いテー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。

$$\begin{aligned} \rho(p_u(y), q_u) \\ \approx \frac{1}{2} \sum_{u=1}^U \sqrt{p_u(y_0)q_u} + \frac{C_h}{2} \sum_{i=1}^n w_i k \left(\left\| \frac{y-x_i}{h} \right\|^2 \right) \end{aligned} \quad (21)$$

ここで y_0 、 x_i 、 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、及びミーンシフト変数を示す。また $k(x)$ 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 y に依存しないので、 $\rho(p_u(y), q_u)$ 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 のミーンシ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。

$$w_i = \sum_{u=1}^U \delta(b(x_i) - u) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (22)$$

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

本節で説明した基本的なミーンシ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 [13]、AdaBoost とミーンシフトを併用したアンサンブル物体追跡 [1]、ミーンシフトを使った高速な大域的な最大値探索法 [48]、非対称のカーネルに拡張したミーンシフト物体追跡 [70] 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。ミーンシフトの研究発信地のひ

ミーンスフトの物体追跡 [各フレーム k で]:

- (1) 初期値 $y_{k,0}$ を前物体位置 $y_{k-1,0}$ で設定
- (2) y_0 において $p_u(y_0), q_u, \rho(p_u(y_0), q_u)$ の計算
- (3) y_0 近傍の候補各点で w_i を式 (22) より計算
- (4) 式 (3) でミーンスフト y_1 を計算
- (5) y_1 において $\rho(p_u(y_1), q_u)$ を計算
- (6) $\rho(p_u(y_1), q_u) < \rho(p_u(y_0), q_u)$ の場合 $y_1 \leftarrow \frac{1}{2}(y_0 + y_1)$ とする
- (7) $y_0 \leftarrow y_1$ とし y の収束まで (4)-(6) を繰返

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 [57] や、一般 CT 画像における体内臓器の三次元領域分割 [56] 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 [36, 38] 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 [39] 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 [31] 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像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 [18, 69]。また 5.5 節で概説した非線形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 [53] 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良好に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、前述したミーンシ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 [25] 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 を併用したテキスト解析による、画像における文書部位の推定法 [29] も報告された。

その他のビジョン問題への応用として、ミーンシフトによる motion field 表現を使ったカメラモーションの認識の研究 [19] 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 [61] 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、ミーンシフトクラスタ解析を応用した、映画の種類の認識 [44] や、CG の研究として、ミーンシ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 [45]、またインテリジェント交通システム開発における、ミーンシ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 [67] や、可変カーネル幅ミーンシ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 [73] などが報告されている。

7 むすび

本稿はミーンシフト法の基礎理論の解説とその応用について述べた。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、計算するミーンシフトベクトルを逐次にメモリ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シ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、各ミーンシ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シ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、ミーンシ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、ガウスクERNEL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シ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 [74] で、B.Georgescu 氏のミーンシ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。MatlabTM 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シフトクラスタ解析のコードなどが MatlabTM のサイト [75](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 [76] から入手可能である。最後に、D.Comaniciu 氏のホームページ [77] 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。

7.3 実装コード例

ここで MatlabTM を使って、カラー画像の領域分割へ向けた、ミーンシフトアルゴリズムの簡単な実装例を挙げる。特徴空間には三次元カラー画素値を用いた。カーネルにはガウス関数を用い、効率的な計算のため、カーネル計算時には標準偏差の三倍の範囲内のデータ点のみを使った。次に近傍にある収束点をグループ化し、それぞれグループの収束点平均値を再代入したもの (NewIMG_G) を領域分割結果として出力する。テスト用の入力画像に用いたレナ画像は [78] からダウンロードできる。もとの画像は 512x512 画素だが、実行時間を早くするため 128x128 に縮小してから領域分割を行った。関数 Meanshift_3D_Loop は、画像中の総ての画素を初期値としてミーンシフトを実行する。また関数 Groupmodes_3D は、カーネル幅で定義される近傍にある収束点をグループ化しそれぞれに整数のレベルをあてる。

プログラム ミーンシフトを使ったカラー画像領域分割

```
function [NewIMG_G,nPeaks] = MeanShiftSeg
% 画像入力
Input = imread('LenaRGB.tif'); % 512x512 カラー画像
Input = imresize(Input,0.25); % 画像を4分の1に縮小

% パラメタの設定
SIZE = size(Input); % 入力画像サイズ
IM = double(Input); % データタイプの変換
N = length(Input(:))/3; % 画素数
IR = squeeze(IM(:,:,1)); % 赤色画像
IG = squeeze(IM(:,:,2)); % 緑色画像
IB = squeeze(IM(:,:,3)); % 青色画像
SAMPLES(:,1) = IR(:); % 1次元ベクトル化
SAMPLES(:,2) = IG(:); % 1次元ベクトル化
SAMPLES(:,3) = IB(:); % 1次元ベクトル化
scaleX      = 10; % カーネル幅
maxIt       = 200; % 最大繰り返し数
thStop      = 0.01; % ミーンシフト収束定数
thDist      = 1; % グルーピング収束定数

% ミーンシフト
tic;
[p_conv_arr] = Meanshift_3D_Loop(SAMPLES,...
                                SAMPLES,...
                                scaleX,...
                                maxIt,...
                                thStop);
toc;

% 収束点のグループ化
tic;
[nPeaks,myLabel] = Groupmodes_3D(p_conv_arr,...
                                scaleX,...
```

```

thDist);

toc;

% 収束点グループの平均値計算
ms_val=zeros(nPeaks,3); % 収束点バッファ
ms_cnt=zeros(nPeaks,1); % 収束点数
ms_ave=zeros(nPeaks,3); % 収束点平均
for i=1:SIZE(1)
    for j=1:SIZE(2)
        n = j+(i-1)*SIZE(2);
        ms_val(myLabel(n),:) = ms_val(myLabel(n),:) ...
                                + p_conv_arr(n,:);
        ms_cnt(myLabel(n),:) = ms_cnt(myLabel(n),:) ...
                                + 1;
    end
end
for k = 1:nPeaks
    ms_ave(k,:) = ms_val(k,:)./ms_cnt(k);
end

% 領域分割結果の出力画像
NewIMG_G = zeros(SIZE);
for i=1:SIZE(1)
    for j=1:SIZE(2)
        n = j+(i-1)*SIZE(2);
        NewIMG_G(j,i,:)=ms_ave(myLabel(n),:);
    end
end
NewIMG_G = uint8(NewIMG_G);

% 入力画像の3次元スカッタダイアグラム
figure(1);
IS = ones(N,1);
scatter3(IR(:),IG(:),IB(:),IS,[0 0 0]);

% 入力画像

```

```

figure(2);
imagesc(Input);
daspect([1,1,1]);
axis off;

% 領域分割結果の出力画像
figure(3);
imagesc(NewIMG_G);
daspect([1,1,1]);
axis off;

function[p_conv_arr] = Meanshift_3D_Loop(INITARR,...
                                         SAMPLES,...
                                         scaleX,...
                                         maxIt,...
                                         thStop)

% 入力のチェックと設定
[M tmp1] = size(INITARR); % 初期値
[N tmp2] = size(SAMPLES); % データサンプル
if tmp1 ~= 3, warning('Mx3,2次元行列ではありません'); end
if tmp2 ~= 3, warning('Nx3,2次元行列ではありません'); end

% パラメタの設定
p_conv_arr = zeros(N,3);
scaleX2 = scaleX*scaleX;

for p=1:M,
    % 初期値の設定
    valX = INITARR(p,1);
    valY = INITARR(p,2);
    valZ = INITARR(p,3);

    % k: ミーンシフトの繰り返しループ
    for k=1:maxIt
        mySumX = 0;

```

```

mySumY = 0;
mySumZ = 0;
myW = 0;

% n: カーネル和計算のためのデータサンプルのループ
for n=1:N,
    valIX = SAMPLES(n,1);
    valIY = SAMPLES(n,2);
    valIZ = SAMPLES(n,3);
    nDist2 = ((valX-valIX)*(valX-valIX)...
              + (valY-valIY)*(valY-valIY)...
              + (valZ-valIZ)*(valZ-valIZ))...
              / scaleX2;
    if nDist2 < 9, % 標準偏差の3倍の範囲
        w = exp(-0.5*nDist2);
        mySumX = mySumX + w*valIX;
        mySumY = mySumY + w*valIY;
        mySumZ = mySumZ + w*valIZ;
        myW     = myW     + w;
    end
end %for n

% 平均値の計算
newValX = mySumX/myW;
newValY = mySumY/myW;
newValZ = mySumZ/myW;

% ミーンシフトベクトルのノルムの2乗
testVal = (valX-newValX)*(valX-newValX)...
           + (valY-newValY)*(valY-newValY)...
           + (valZ-newValZ)*(valZ-newValZ);

% 繰り返し変数の更新
valX = newValX;
valY = newValY;
valZ = newValZ;

```

```

        % 収束条件のチェック
        if testVal < thStop*thStop*scaleX2
            break;
        end

    end %for k

    % 繰り返しの収束点を出力変数に格納
    p_conv_arr(p,1) = valX;
    p_conv_arr(p,2) = valY;
    p_conv_arr(p,3) = valZ;

end % for p

function[nPeaks,myLabel] = Groupmodes_3D(p_conv_arr,...
                                         scaleX,...
                                         thDist)

% 入力のチェックと設定
[N, tmp] = size(p_conv_arr);
if tmp ~= 3, warning(' 入力が 3 次元行列でない'); end

% パラメタの設定
scaleX2 = scaleX*scaleX;

% 出力バッファの初期化
nPeaks      = 1; % モード数のカウンタ
myLabel     = zeros(1,N); % 出力モードのレーベル
myLabel(1)  = nPeaks;

% 近傍収束点のグループ化
for i=2:N
    valX = p_conv_arr(i,1);
    valY = p_conv_arr(i,2);
    valZ = p_conv_arr(i,3);

```

```

for j=1:i-1
    % 収束点 I と J の距離計算
    dist = (valX-p_conv_arr(j,1))*(valX-p_conv_arr(j,1))...
           + (valY-p_conv_arr(j,2))*(valY-p_conv_arr(j,2))...
           + (valZ-p_conv_arr(j,3))*(valZ-p_conv_arr(j,3));

    % 近傍条件のチェック
    if dist < thDist*scaleX2
        myLabel(i) = myLabel(j);
        break;
    end
end %for j

% 新たなレーベルの更新
if myLabel(i)==0
    nPeaks = nPeaks+1;
    myLabel(i) = nPeaks;
end

end %for i

```

図1と図2に上記のコードの実行結果を示す。カーネル幅には10、最大繰り返し数は200、ミーンシフト収束定数には0.01、グループ化近傍定数には1を用いた。2.0GHzのCPUと1GBメモリを用いたインテルPC上で、ウィンドウズXP™上のMatlab7.0R14™を使って実験した結果、約11分の実行時間がかかった。出力として、12個のモード($C=12$)が検出された。

このコードをもとに、実装上の高速化、様々なカーネル関数やカーネル幅の領域分割結果への影響の評価ができる。また、空間と輝度値の五次元合成特徴空間、エッジを保存した平滑化、非等方空間カーネル幅等への拡張も簡単に出来る。是非試して欲しい。

参考文献

- [1] S. Avidan: "Ensemble tracking," IEEE Trans. Pattern Anal. and Machine Intell., Vol. 29, No. 2, pp. 261–271 (2007)

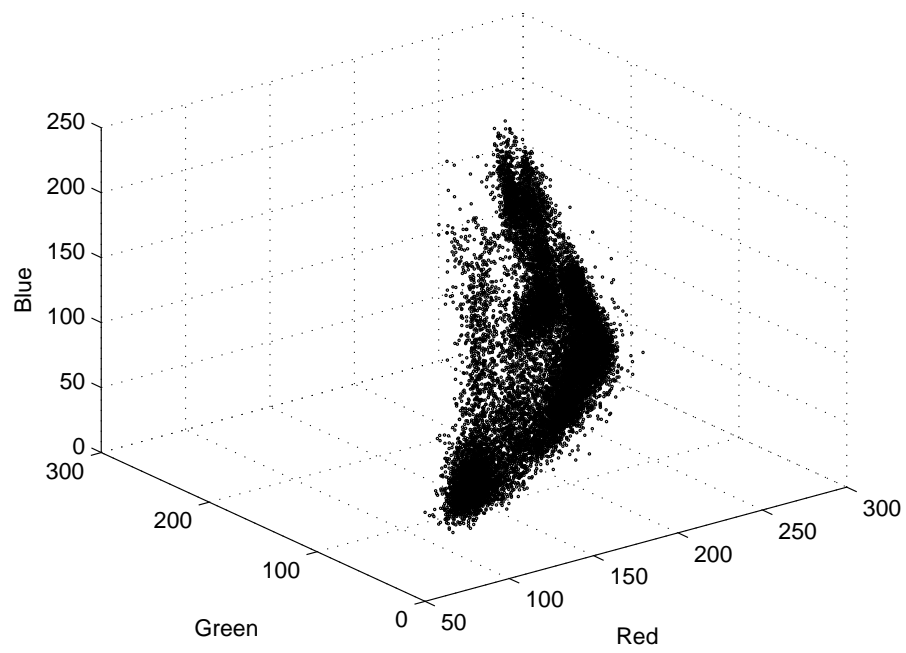


図 1: 入力カラー画像のスキッタダイアグラム。

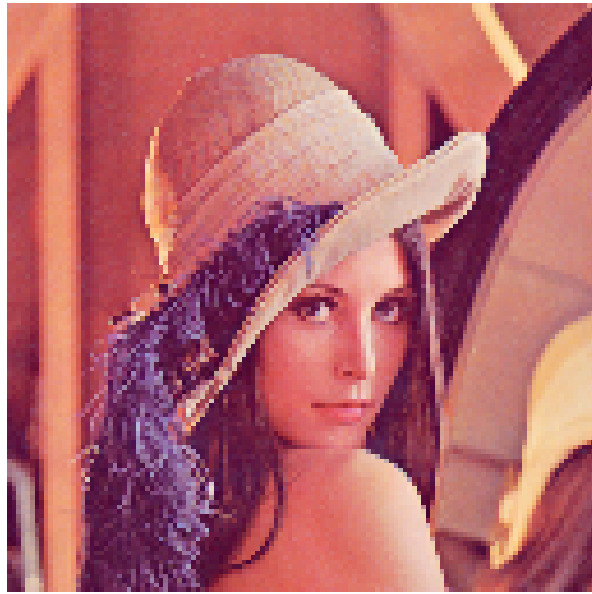


図 2: 領域分割の結果。入力画像(上)、ミーンシフトによる領域分割結果(下)。

- [2] M. J. Black, G. Sapiro, D. Marimont, and D. Heeger: “Robust anisotropic diffusion,” *IEEE Trans. Image Processing*, Vol. 7, No. 3, pp. 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] M. A. Carreira-Perpinan: “Acceleration strategies for gaussian mean-shift image segmentation,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2006)
- [5] M. A. Carreira-Perpinan: “Fast nonparametric clustering with gaussian blurring mean-shift,” *Proc. Int. Conf. Machine Learning* (2006)
- [6] M. A. Carreira-Perpinan: “Gaussian mean-shift is an EM algorithm,” *IEEE Trans. Pattern Anal. and Machine Intell.*, Vol. 29, No. 5, pp. 767–776 (2007)
- [7] M. A. Carreira-Perpinan: “Generalised blurring mean-shift algorithms for nonparametric clustering,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2008)
- [8] H. Chen and P. Meer: “Robust fusion of uncertain information,” *IEEE Trans. Systems Man Cybernetics Part. B*, Vol. 35, No. 3, pp. 578–586 (2005)
- [9] Y. Cheng: “Mean shift, mode seeking, and clustering,” *IEEE Trans. Pattern Anal. and Machine Intell.*, Vol. 17, No. 8, pp. 790–799 (1995)
- [10] R. T. Collins: “Mean-shift blob tracking through scale space,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Vol. II, pp. 234–240 (2003)
- [11] D. Comaniciu: “An algorithm for data-driven bandwidth selection,” *IEEE Trans. Pattern Anal. and Machine Intell.*, Vol. 25, No. 2, pp. 281–288 (2003)
- [12] D. Comaniciu: “Kernel-based object tracking,” *IEEE Trans. Pattern Anal. and Machine Intell.*, Vol. 25, No. 5, pp. 564–575 (2003)
- [13] D. Comaniciu: “Nonparametric information fusion for motion estimation,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Vol. I, pp. 59–66 (2003)
- [14] D. Comaniciu and P. Meer: “Mean shift analysis and applications,” *Proc. IEEE Int. Conf. Computer Vision*, pp. 1197–1203 (1999)
- [15] D. Comaniciu and P. Meer: “Mean shift: A robust approach toward feature space analysis,” *IEEE Trans. Pattern Anal. and Machine Intell.*, Vol. 24, No. 5, pp. 603–619 (2002)
- [16] D. Comaniciu, V. Ramesh, and P. Meer: “Real-time tracking of non-rigid objects using mean shift,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 142–149 (2000)
- [17] D. Comaniciu, V. Ramesh, and P. Meer: “The variable bandwidth mean shift and data-driven scale selection,” *Proc. IEEE Int. Conf. Computer Vision*, Vol. I, pp. 438–445 (2001)
- [18] 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*, Vol. 24, No. 6, pp. 697–711 (2005)

- [19] 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*, Vol. 8, No. 2, pp. 323–340 (2006)
- [20] R. O. Duda and P. E. Hart: *Pattern Classification and Scene Analysis*, John Wiley & Sons, Inc., (1973)
- [21] M. Fashing and C. Tomasi: “Mean shift is a bound optimization,” *IEEE Trans. Pattern Anal. and Machine Intell.*, Vol. 27, No. 3, pp. 471–474 (2005)
- [22] L. Friedman, N. Netanyahu, and M. Shoshany. “Mean shift-based clustering of remotely sensed data,” *Proc. IEEE Int. Geoscience and Remote Sensing Symp.*, pp. 3432–3434 (2003)
- [23] K. Fukunaga: *Introduction to Statistical Pattern Recognition*, Academic Press, (1990)
- [24] K. Fukunaga and L. Hostetler: “The estimation of the gradient of a density function, with applications in pattern recognition,” *IEEE Trans. Information Theory*, Vol. 21, No. 1, pp. 32–40 (1975)
- [25] B. Georgescu, I. Shimshoni, and P. Meer: “Mean shift based clustering in high-dimension: a texture classification example,” *Proc. IEEE Int. Conf. Computer Vision*, Vol. I, pp. 456–463 (2003)
- [26] K. Inoue and K. Urahama: “Hierarchically distributed dynamic mean shift,” *Proc. IEEE Int. Conf. Image Processing*, Vol. I, pp. 269–272 (2007)
- [27] A. Jagmohan, M. Singh, and N. Ahuja: “Dense Stereo Matching Using Kernel Maximum Likelihood Estimation,” *Proc. IARP Int. Conf. Pattern Recognition*, Vol. III, pp. 28–31 (2004)
- [28] T. Jebara and A. Pentland: “Maximum conditional likelihood via bound maximization and the CEM algorithm,” *Neural Information Processing Systems*, Vol. 11, pp. 494–500 (1998)
- [29] 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.*, Vol. 25, No. 12, pp. 1631–1639 (2003)
- [30] J. J. Koenderink: “The structure of images,” *Biol. Cybern.*, Vol. 50, pp. 363–370 (1984)
- [31] 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*, Vol. 52, No. 11, pp. 1966–1969 (2005)
- [32] X. Li, Z. Hu, and F. Wu: “A note on the convergence of the mean shift,” *Pattern Recognition*, Vol. 40, No. 6, pp. 1756–1762 (2007)
- [33] T. Lindeberg: “Feature detection with automatic scale selection,” *Int. J. Computer Vision*, Vol. 30, No. 2, pp. 79–116 (1998)
- [34] D. Lowe: “Distinctive image features from scale-invariant keypoints,” *Int. J. Computer Vision*, Vol. 60, No. 2, pp. 91–110 (2004)

- [35] K. Okada, D. Comaniciu, N. Dalal, and A. Krishnan: "A robust algorithm for characterizing anisotropic local structures," Proc. European Conf. Computer Vision, Vol. I, pp. 549–561 (2004)
- [36] K. Okada, D. Comaniciu, and A. Krishnan: "Robust 3D segmentation of pulmonary nodules in multislice CT images," Proc. Int. Conf. Medical Imaging Computing and Computer-Assisted Intervention, Vol. II, pp. 881–889 (2004)
- [37] K. Okada, D. Comaniciu, and A. Krishnan: "Scale selection for anisotropic scale-space," Proc. IEEE Conf. Computer Vision and Pattern Recognition, Vol. I, pp. 594–601 (2004)
- [38] K. Okada, D. Comaniciu, and A. Krishnan: "Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT," IEEE Trans. Medical Imaging, Vol. 24, No. 3, pp. 409–423 (2005)
- [39] K. Okada and X. Huang: "Robust click-point linking: Matching visually dissimilar local regions," Proc. IEEE Int. Workshop Beyond Multiview Geometry (2007)
- [40] K. Okada, M. Singh, and V. Ramesh: "Prior-constrained scale-space mean shift," Proc. British Machine Vision Conf., Vol. II, pp. 829–838 (2006)
- [41] S. Paris and F. Durand: "A topological approach to hierarchical segmentation using mean shift," IEEE Conf. Computer Vision and Pattern Recognition, (2007)
- [42] E. Parzen: "On estimation of a probability density and mode," Annals of Mathematical Statistics, Vol. 33, pp. 1065–1076 (1962)
- [43] P. Perona and J. Malik: "Scale-space and edge detection using anisotropic diffusion," IEEE Trans. Pattern Anal. and Machine Intell., Vol. 12, No. 7, pp. 629–639 (1990)
- [44] Z. Rasheed, Y. Sheikh, and M. Shah: "On the use of computable features for film classification," IEEE Trans. Circuits and Systems for Video Technology, Vol. 15, No. 1, pp. 52–64 (2005)
- [45] V. Setlur, T. Lechner, M. Nienhaus, and B. Gooch: "Retargeting images and video for preserving information saliency," IEEE Computer Graphics and Applications, Vol. 27, No. 5, pp. 80–88 (2007)
- [46] S. J. Sheather and M. C. Jones: "A reliable data-based bandwidth selection method for kernel density estimation," J. Royal Statistical Soc. Series B, Vol. 53, pp. 683–690 (1991)
- [47] Y. A. Sheikh, E.A. Khan, and T. Kanade: "Mode-seeking by medoidshifts," Proc. IEEE Int. Conf. Computer Vision, Vol. I, pp. 1–8 (2007)
- [48] 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, Vol. 16, No. 5, pp. 1457–1469 (2007)
- [49] B. W. Silverman: *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, (1986)

- [50] M. Singh, H. Arora, and N. Ahuja: “A robust probabilistic estimation framework for parametric image models,” Proc. European Conf. Computer Vision, Vol. I, pp. 508–522 (2004)
- [51] R. Subbarao, Y. Genc, and P. Meer: “Nonlinear mean shift for robust pose estimation,” Proc. IEEE Workshop on Applications of Computer Vision, (2007)
- [52] R. Subbarao, Y. Genc, and P. Meer: “Robust unambiguous parameterization of the essential manifold,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, (2008)
- [53] R. Subbarao and P. Meer: “Discontinuity preserving filtering over analytic manifolds,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, (2007)
- [54] R. Subbarao and P. Meer: “Nonlinear mean shift for clustering over analytic manifolds,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, pages Vol. I, pp. 1168–1175 (2006)
- [55] R. Subbarao and P. Meer: “Nonlinear mean shift over riemannian manifolds,” Int. J. Computer Vision (2009)
- [56] H. Tek, M. Bergholdt, D. Comaniciu, and J. Williams: “Segmentation of 3D medical structures using robust ray propagation,” Proc. Int. Conf. Medical Imaging Computing and Computer-Assisted Intervention, Vol. I, pp. 572–579 (2002)
- [57] H. Tek, D. Comaniciu, and J. Williams: “Vessel detection by mean shift based ray propagation,” Proc. IEEE Int. Workshop Mathematical Methods in Biomedical Image Analysis (2001)
- [58] B. A. Turlach: “Bandwidth selection in kernel density estimation: A review,” Technical report, Louvain-la-Neuve, Belgium, Institut de Statistique (1993)
- [59] O. Tuzel, R. Subbarao, and P. Meer: “Simultaneous multiple 3D motion estimation via mode finding on lie groups,” Proc. IEEE Int. Conf. Computer Vision, Vol. I, pp. 18–25 (2005)
- [60] A. Vedaldi and S. Soatto: “Quick shift and kernel methods for mode seeking,” Proc. European Conf. Computer Vision, Vol. IV, pp. 705–718 (2008)
- [61] 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., Vol. 29, No. 5, pp. 901–905 (2007)
- [62] 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)
- [63] T. Wakahara and K. Ogura: “Extended mean shift in handwriting clustering,” Proc. IARP Int. Conf. Pattern Recognition, Vol. I, pp. 384–388 (1998)
- [64] M. P. Wand and M. C. Jones: *Kernel Smoothing*, Chapman & Hall, (1995)
- [65] H. Wang and D. Suter: “False-peaks-avoiding mean shift method for unsupervised peak-valley sliding image segmentation,” Proc. Int. Conf. Digital Image Computing: Techniques and Applications, pp. 581–590 (2003)

- [66] A. Witkin: "Scale-space filtering," Proc. Int. Joint. Conf. Artificial Intell., pp. 1019–1021 (1983)
- [67] F. Xu, X. Liu, and K. Fujimura: "Pedestrian detection and tracking with night vision," IEEE Trans. Intelligent Transportation Systems, Vol. 6, No. 1, pp. 63–71 (2005)
- [68] C. Yang, R. Duraiswami, N. Gumerov, and L. Davis: "Improved fast gauss transform and efficient kernel density estimation," Proc. IEEE Int. Conf. Computer Vision, pp. 464–471, 2003.
- [69] 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, Vol. 53. No. 11, pp. 2405–2414 (2006)
- [70] A. Yilmaz: "Object tracking by asymmetric kernel mean shift with automatic scale and orientation selection," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1–6 (2007)
- [71] X. Yuan and S. Z. Li: "Half quadratic analysis for mean shift: with extension to a sequential data mode-seeking method," Proc. IEEE Int. Conf. Computer Vision, (2007)
- [72] K. Zhang, J. T. Kwok, and M. Tang: "Accelerated convergence using dynamic mean shift," Proc. European Conf. Computer Vision, pp. 257–268 (2006)
- [73] Y. Zhu, D. Comaniciu, M. Pellkofer, and T. Koehler: "Reliable detection of overtaking vehicles using robust information fusion," IEEE Trans. Intelligent Transportation Systems, Vol. 7, No. 4, pp. 401–414 (2006)
- [74] <http://www.caip.rutgers.edu/riul/research/code.html>.
- [75] <http://www.mathworks.com/matlabcentral/>.
- [76] <http://www.intel.com/technology/computing/opencv/>.
- [77] <http://www.caip.rutgers.edu/comanici/>.
- [78] <http://en.wikipedia.org/wiki/Lenna>.