

SVM を用いたテクスチャ分類における 局所高次モーメントカーネルと従来カーネルの比較

亀山 啓輔[†]

[†] 筑波大学 大学院システム情報工学研究科 コンピュータサイエンス専攻
Keisuke.Kameyama@cs.tsukuba.ac.jp

1 はじめに

画像テクスチャの特徴量としてテクスチャの局所的な高次モーメント (LHOM) や高次モーメントスペクトル (LHOMS) がしばしば用いられる。利用するモーメントの次数を上げることで、高次の特徴量を抽出することができる反面、特徴空間の次元が高くなる問題がある。これに対し、Support Vector Machine (SVM) [1] は、高次元の特徴量空間における特徴量の内積のみを評価するカーネル法を採用することで計算量の増大を回避することができ、実際に高次モーメント (HOM) 特徴量の内積として定義された HOM カーネルを用いた SVM により、効率的に高次モーメント特徴量を利用したテクスチャ分類が可能であることが報告されている [2]。

本研究では、テクスチャ分類に際して用いるカーネル関数として従来より用いられてきた一般的なカーネルと、LHOM(S) カーネルを用いた場合の比較を行い、LHOM(S) カーネルを利用した SVM の有効性を示す。

2 局所高次モーメントスペクトルとカーネル関数

信号 $s(t)$ を \mathbf{R}^2 上の実数値信号であるとする。信号 s の局所 n 次モーメントスペクトル (LHOMS) は、位置 $t = \mathbf{x}$ 近傍における窓関数 $w(t)$ を用いた局所フーリエ変換 $S_w(\omega, \mathbf{x}) = \int s(t + \mathbf{x})w(t) \exp(-j\omega^T t) dt$ を用いて、

$$M_{s,w,n}(\omega_1, \omega_2, \dots, \omega_{n-1}, \mathbf{x}) = S_w^* \left(\sum_{k=1}^{n-1} \omega_k, \mathbf{x} \right) \prod_{k=1}^{n-1} S_w(\omega_k, \mathbf{x}), \quad (1)$$

と求められる。この特徴量に対応する LHOM(S) カーネル関数は、内積により

$$K_{w,n}(s, v; \mathbf{x}, \mathbf{y}) = \langle M_{s,w,n}(\mathbf{x}), M_{v,w,n}(\mathbf{y}) \rangle = (2\pi)^{2(n-1)} \int \left[\int s_w(\mathbf{z}, \mathbf{x}) v_w(\mathbf{z} + \boldsymbol{\tau}, \mathbf{y}) d\mathbf{z} \right]^n d\boldsymbol{\tau} \quad (2)$$

と求められる。ここで、 $s_w(t, \mathbf{x}) = w(t)s(t + \mathbf{x})$ である。この関数は次数 n が増大してもその計算量はほぼ

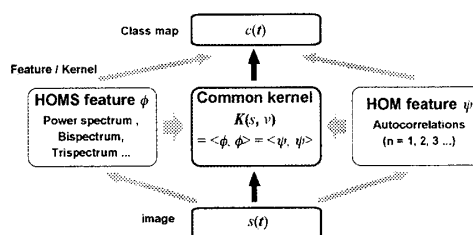


図 1: LHOM(S) カーネルを用いた高次モーメント、高次モーメントスペクトルを統合した認識方式

一定であるという特徴があり、高次の特徴量を用いる上で有効である。また、以下の系により、局所パワースペクトル ($M_{s,w,2}$), 局所バイスペクトル ($M_{s,w,3}$) などの LHOMS 特徴量を用いる場合と、局所高次自己相関関数 ($m_{s,w,2}, m_{s,w,3}, \dots$) などの LHOM 特徴量を用いる場合とに対応するカーネル関数が比例関係にあることを示しており、カーネル法による場合には、両特徴量を利用したパターン認識方式が統一的に扱うことが可能となることを意味する (図 1)。

系 1 画像信号 $s(t)$ の n 次 LHOMS および LHOM をそれぞれ $M_{s,w,n}$ と $m_{s,w,n}$ とおくと、

$$\langle M_{s,w,n}, M_{v,w,n} \rangle = c \langle m_{s,w,n}, m_{v,w,n} \rangle$$

が成立する。ここで、 $c \in \mathbf{R}$, $c \neq 0$ である。 [2]。

3 テクスチャ分類実験

3.1 実験条件

SVM を用いたテクスチャ分類実験では、従来より知られている多項式カーネル

$$K(s, v) = (\langle s, v \rangle + 1)^d \quad (d = 1, 2, \dots) \quad (3)$$

や RBF カーネル

$$K(s, v) = \exp\left(-\frac{\|s - v\|^2}{\gamma^2}\right). \quad (\gamma \in \mathbf{R}) \quad (4)$$

と LHOM(S) カーネルを用いた場合と比較した。実験はいずれも 2 類分類問題とし、 32×32 画素の局所画像を入力とした。各クラスとも元画像のランダムな位置

[‡] Local Higher-Order Moment Kernel and Conventional Kernels in Texture Classification by SVM

[†] Keisuke Kameyama, Grad. School, SIE, Univ. Tsukuba

表 1: カーネル関数による学習可能性, サポートベクトル数, 認識率比較. (a) 人工テクスチャ, (b) 自然テクスチャの場合. “SM”はソフトマージン法を表す.

(a) Synthetic texture

Kernel Type	Trained?	SV number (ratio to training set)	Test rate
Polynomial ($d=1, \dots, 5$)	No	-	-
RBF ($\gamma=10$)	Yes	39 (97.5%)	80%
LHOM($n=2$) ($\sigma=8$)	No	-	-
LHOM($n=3$) ($\sigma=8$)	Yes	2 (5%)	83%
LHOM($n=4$) ($\sigma=8$)	Yes	3 (7.5%)	82%
LHOM($n=5$) ($\sigma=8$)	Yes	8 (20%)	84%

(b) Natural texture

Kernel Type	Trained?	SV number (No SM / Use SM)	Test rate (No SM / Use SM)
Polynomial ($d=1, \dots, 5$)	Partly (at $d=1$ only)	40 / 40	50% / 50%
RBF ($\gamma=10$)	Yes	40 / 40	65% / 65%
LHOM($n=2$) ($\sigma=8$)	Yes	5 / 40	69% / 79%
LHOM($n=3$) ($\sigma=8$)	Yes	27 / 40	73% / 75%
LHOM($n=4$) ($\sigma=8$)	Yes	30 / 35	75% / 75%
LHOM($n=5$) ($\sigma=8$)	Yes	37 / 40	67% / 69%

から切り出した 20 の学習データを用いた. LHOM(S) カーネル関数は式 (2) を離散化した

$$K_{w,n}(\mathbf{s}, \mathbf{v}; \mathbf{x}, \mathbf{y}) = \sum_i \left[\sum_j s_w(\mathbf{z}_j, \mathbf{x}) v_w(\mathbf{z}_j + \boldsymbol{\tau}_i, \mathbf{y}) \right]^n \quad (5)$$

を用い, 窓関数として等方性 $\Sigma = \sigma^2 I$ を仮定したガウス関数 $w(\mathbf{t}, \Sigma) = 2\pi^{-1} |\Sigma|^{-1/2} \exp(-\frac{1}{2} \mathbf{t}^T \Sigma^{-1} \mathbf{t})$ を用いた. 次数 n と窓幅 σ はそれぞれ ($n=2, 3, 4, 5$) と ($\sigma=4, 8, 12$) について試みた.

学習した SVM をテストする際には図 2(a) に示すような 2 クラスの組み合わせ画像に対する分類正答率を評価した.

3.2 実験 1: 人工テクスチャ

本実験では, 計算機で生成した 2 種類のテクスチャ

$$s_A(\mathbf{x}) = \sin(\omega_0^T \mathbf{x}) + \sin(2\omega_0^T \mathbf{x}) + \sin(3\omega_0^T \mathbf{x})$$

$$s_B(\mathbf{x}) = \sin(\omega_0^T \mathbf{x}) + \sin(2\omega_0^T \mathbf{x} + \phi) + \sin(3\omega_0^T \mathbf{x})$$

の分類を試みた. ここで, 基本空間周波数ベクトルは, $\omega_0 = [\pi/6, \pi/6]^T$ である. 両テクスチャの相違点は, s_B の 2 倍周波成分の位相シフト $\phi (= \pi/2)$ のみであるため 3 次以上の特徴量が分類に有効となる.

表 1(a) に示すように, RBF 及び 3,4,5 次の LHOM(S) カーネルを用いた場合に学習可能であった. LHOM(S)

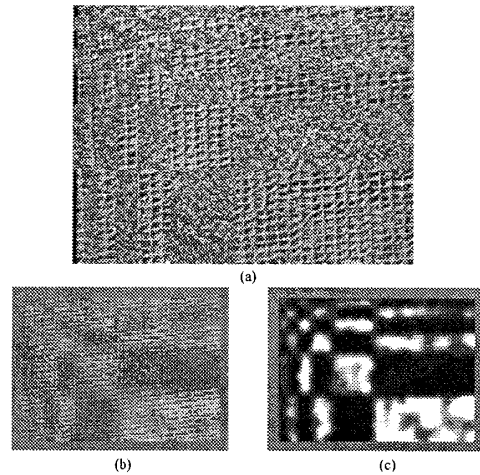


図 2: (a) 自然テクスチャの評価用組合せ. 分類結果: (b) RBF カーネル使用, (c) LHOM(S) カーネル使用

カーネルを用いた場合に認識率が向上しているのは, ソフトマージンを用いない場合にも低く抑えられているサポートベクトル数からも分かるように, 特徴空間内でのクラス分離度の高さによるものと考えられる.

3.3 実験 2: 自然テクスチャ

Vision Texture データ [3] より, 布地のテクスチャ “fabric17” と “fabric18” を選び, 輝度データとして用いた.

表 1(b) から, 実験 1 と同様の傾向を見て取ることができる. 図 2 に示した SVM の線形出力マップからは, LHOM(S) カーネルの場合に対比して, RBF カーネルを用いた場合にはクラス間のコントラストが著しく低いことが分かる. このことは, LHOM(S) カーネルを用いた場合のクラスの高さと分類器としての安定性の高さを示すと考えられる.

4 結論

SVM を用いたテクスチャ分類に際して用いるカーネル関数として, 多項式カーネルや RBF カーネルと, 局所高次モーメント (LHOM(S)) カーネルを用いた場合の比較を行い, 分類正答率, SVM 応答の安定性, サポートベクトル数等から LHOM(S) カーネルを利用した場合の有効性を示した. 今後の課題としては窓関数の形状決定の指針を示していくことなどが考えられる.

参考文献

- [1] V. N. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed. Springer, 2000.
- [2] K. Kameyama and K. Taga, *Proc. IJCNN2004*, vol. 4, 2004, pp. 3009–3014.
- [3] MIT Vision and Modeling Group, “Vision texture,” 1995.