

# 時系列パターンのための カーネル最小分類誤り学習法

大越 俊<sup>†</sup> 山田 浩嗣<sup>†</sup> 渡辺 秀行<sup>††</sup> 片桐 滋<sup>†</sup> 大崎 美穂<sup>†</sup>  
<sup>†</sup>同志社大学 <sup>††</sup>ATR

## 1. はじめに

カーネルを用いたパターン分類法の一つ、カーネル最小分類誤り(Kernel Minimum Classification Error: KMCE)学習法<sup>[1]</sup>は、元々固定次元パターンの分類を想定していた。一方、パターン間距離の計算に非線形時間軸伸縮を避けていく可変長パターンの分類においては、カーネルに伴うグラム行列の半正定値性を保証できず、KMCE法の適用が困難であった。この問題の解決を目指し、可変長パターンから切り出した固定次元ベクトルパターンに対してカーネル写像を適用し、その出力を状態遷移モデルによって統合する新しいKMCE法を提案する。

## 2. 時系列パターンのための KMCE 学習法

固定次元ベクトル(フレームベクトルと呼ぶ)の系列から成る入力可変長パターン $\mathbf{X}$ を $J$ 個のクラスの一つに分類する問題を考える。提案手法における識別関数 $G_j(\mathbf{X}; \mathbf{\Lambda})$ を、可変長パターン中のフレーム $\mathbf{x}_t$ における識別関数値 $g_{\varphi_{j,t} \theta_{j,t}}(\mathbf{x}_t)$ (後述)の動的時間軸伸縮(DTW)法による最大累積値として以下のように定義する。

$$G_j(\mathbf{X}; \mathbf{\Lambda}) = \frac{1}{T} \sum_{t=1}^T g_{\varphi_{j,t} \theta_{j,t}}(\mathbf{x}_t) \quad (1)$$

ここで、 $G_j(\mathbf{X}; \mathbf{\Lambda})$ は $\mathbf{X}$ のクラス $C_j$ に対する帰属度を表し、 $\mathbf{\Lambda}$ は分類器パラメータ集合である。最大累積値をもたらす最適な経路 $\{(\varphi_{j,1} \theta_{j,1}); \dots; (\varphi_{j,T} \theta_{j,T})\}$ はDTWによって求められる。また、 $g_{\varphi_{j,t} \theta_{j,t}}(\mathbf{x}_t)$ は

$$g_{\varphi_{j,t} \theta_{j,t}}(\mathbf{x}_t; \mathbf{\Lambda}) = \sum_{m=1}^M \tau_{m, \varphi_{j,t} \theta_{j,t}} K(\mathbf{x}_t, \mathbf{p}_m) \quad (2)$$

と定義する。ここで $\{\mathbf{p}_m\}_{m=1}^M$ はフレームベクトルをの局所的分布を代表するプロトタイプベクトルの集合であり、 $\{\tau_{m, \varphi_{j,t} \theta_{j,t}}\}_{m=1}^M$ は学習によって調整されるカーネル重みである。 $K(\cdot, \cdot)$ はカーネルである。

本手法では分類は式(1)を用いて行うが、KMCE法の学習はフレームベクトルを対象にして行う。 $\mathbf{X}$ が所属するクラスを $C_y$ 、最も誤り易いクラスを $C_{y^*}$ とすると、フレームインデックス $t$ における誤分類尺度はカーネルによって写像された空間 $H$ 上での幾何マージン $r_H$ の正負を反転したものととして次のように定義する。ただし $(\varphi_{j,t} \theta_{j,t}) = (j, t)$ である。

$$D_{y,t}(\mathbf{x}_t; \mathbf{\Lambda}) = -\frac{\tau_{y,t}^T k(\mathbf{x}_t) - \tau_{y^*,t}^T k(\mathbf{x}_t)}{\sqrt{(\tau_{y,t} - \tau_{y^*,t})^T \mathbf{K} (\tau_{y,t} - \tau_{y^*,t})}} = -r_H \quad (3)$$

ここで $k(\mathbf{x})$ と $\mathbf{K}$ は、それぞれ次のような経験的カーネルマツ

ブとグラム行列である。

$$k(\mathbf{x}) = [K(\mathbf{x}, \mathbf{p}_1) \dots K(\mathbf{x}, \mathbf{p}_M)]^T \quad (4)$$

$$\mathbf{K} = \begin{bmatrix} K(\mathbf{p}_1, \mathbf{p}_1) & \dots & K(\mathbf{p}_1, \mathbf{p}_M) \\ \vdots & \ddots & \vdots \\ K(\mathbf{p}_M, \mathbf{p}_1) & \dots & K(\mathbf{p}_M, \mathbf{p}_M) \end{bmatrix} \quad (5)$$

KMCE学習法は、勾配法などの手法を用いて次の経験的平均損失を最小にする $\mathbf{\Lambda}$ を探索する。

$$L(\mathbf{\Lambda}) = \frac{1}{N} \sum_{n=1}^N \sum_{t=1}^T \ell(D_{y,t}(\mathbf{x}_t; \mathbf{\Lambda})) \quad (6)$$

ここで $\ell(D_{y,t}(\mathbf{x}_t; \mathbf{\Lambda}))$ は

$$\ell(D_{y,t}(\mathbf{x}_t; \mathbf{\Lambda})) = \frac{1}{1 + \exp(-\alpha D_{y,t}(\mathbf{x}_t; \mathbf{\Lambda}))} \quad (\alpha > 0) \quad (7)$$

で表される平滑化分類誤り損失であり、 $\alpha$ は損失関数の平滑度を制御するハイパーパラメータである。

## 3. 評価実験

提案法の動作確認のため、単純な状態遷移モデルを用いた可変長の単語音声パターンの分類実験を行った。音声データはETL-WD-Iを用いた。男女各6名の492単語を学習標本、男女各7名の492単語を試験用標本とした(Hold-Out法)。各音素クラス用に、1状態・2プロトタイプ(/状態)型の状態遷移モデルを用いた。

KMCE法の実行では、セグメンタル $K$ 平均法<sup>[2]</sup>でプロトタイプを初期化し、カーネル重みのみの学習を行った。学習結果を表1に示す。学習によって学習用データのみならず試験用データに対する分類精度も向上し、提案法が期待通りに動作していることが確認できる。

	学習前	学習後
学習用	2.85%	56.40%
試験用	3.48%	59.34%

表1 認識率(分類率)

謝辞:ETL-WD-IデータセットはNII音声資源コンソーシアムからの提供を受けた。また本研究の一部は、科研費(18H03266)および文科省平成26年度私学戦略的研究基盤形成支援事業「進化適応型自動車運転支援システム「ドライバ・イン・ザ・ループ」」の支援をうけて行われた。

## 参考文献

- [1] 渡辺秀行, 他: 信学技報 PRMU, Dec. 2010.
- [2] Juang B.-H., et al.: IEEE Trans. ASSP, Vol. 38, No. 9 pp. 1639-1641, 1990.