

時系列パターンの時間長正規化を伴う カーネル最小分類誤り学習法

山田 浩嗣[†] 谷 陵真[†] 渡辺 秀行[‡] 大崎 美穂[†] 片桐 滋[†]
[†]同志社大学 [‡]ATR

1. はじめに

幾何マージンやカーネル法を活用するパターン分類器学習法の一つにカーネル最小分類誤り (KMCE: Kernel Minimum Classification Error) 学習法^[1]がある。これまで、固定次元パターンの分類におけるその有用性は示されてきたが、カーネルに伴うグラム行列の半正定値性を保証できない可変長パターン分類への適用は必ずしも十分に検証されていない。本研究では、この可変長性の問題を解決するため、最小歪基準に基づいて原可変長パターンの長さを正規化した上で KMCE 学習を適用する手法の効果を検証する。

2. カーネル最小分類誤り学習法

入力パターン \mathbf{x} を J 個のクラス $\{C_j\}_{j=1}^J$ のうちの 1 つに分類する問題が与えられたとする。KMCE 学習法では、カーネルに基づく高次元空間 (特徴空間 H) において識別関数は次式となる。

$$g_j(\mathbf{x}; \Lambda) = \sum_{m=1}^M \tau_{m,j} K(\mathbf{x}, \mathbf{p}_m) \quad (1)$$

ここで $g_j(\mathbf{x}; \Lambda)$ は C_j に対する識別関数であり、 \mathbf{x} が C_j に帰属する程度を示す。また、 Λ は学習対象の分類器パラメータ集合であり、学習標本集合 $\{\mathbf{x}_n\}_{n=1}^N$ を集約したプロトタイプ集合 $\{\mathbf{p}_m\}_{m=1}^M$ とカーネル重み $\tau_j = \{\tau_{m,j}\}_{m=1}^M$ から成る。なお、 $K(\cdot, \cdot)$ はカーネルであり、ガウスカーネルのような場合、そのカーネル幅は Λ に含まれる。

KMCE 学習法では、特徴空間 H での幾何マージン r_H を誤分類尺度 $D_y(\mathbf{x}; \Lambda)$ としている。

$$D_y(\mathbf{x}; \Lambda) = -\frac{\tau_y^T \mathbf{k}(\mathbf{x}) - \tau_i^T \mathbf{k}(\mathbf{x})}{\sqrt{(\tau_y - \tau_i)^T K (\tau_y - \tau_i)}} = -r_H \quad (2)$$

ここで \mathbf{x} が所属するクラスを C_y 、最も誤りやすいクラスを C_i としている。また、 $\mathbf{k}(\mathbf{x})$ は経験的カーネルマップ、 \mathbf{K} はグラム行列と呼ばれ、次式で表される。

$$\mathbf{k}(\mathbf{x}) = [K(\mathbf{x}, \mathbf{p}_1), K(\mathbf{x}, \mathbf{p}_2), \dots, K(\mathbf{x}, \mathbf{p}_M)] \quad (3)$$

$$\mathbf{K} = \begin{bmatrix} K(\mathbf{p}_1, \mathbf{p}_1) & \dots & K(\mathbf{p}_1, \mathbf{p}_M) \\ \vdots & \ddots & \vdots \\ K(\mathbf{p}_M, \mathbf{p}_1) & \dots & K(\mathbf{p}_M, \mathbf{p}_M) \end{bmatrix} \quad (4)$$

KMCE 学習法では損失関数として次式の平滑化分類誤り数損失を採用している。

$$\ell(D_y(\mathbf{x}; \Lambda)) = \frac{1}{1 + \exp(-\alpha D_y(\mathbf{x}; \Lambda))} \quad (\alpha > 0) \quad (5)$$

ここで、 α は損失関数の平滑度を調整するハイパーパラメータである。学習パラメータ集合 Λ の最適化は、学習

標本 $\{(\mathbf{x}_n, y_n)\}_{n=1}^N$ から計算される次式の経験的平均損失の最小化を通して行われる。

$$L(\Lambda) = \frac{1}{N} \sum_{n=1}^N \ell(D_{y_n}(\mathbf{x}_n; \Lambda)) \quad (6)$$

3. 時間長正規化の手法

最小歪基準に基づいて可変長パターンの長さの正規化を行うため、特に分節化 K 平均法^[2]を採用する。元々本手法は、複数の可変長パターンを分節化を伴ってクラスタリングするための手法であるが、ここでは単一のパターンを対象に、分節化と各分節内におけるクラスタリングのみを行うように利用する。

4. 評価実験

可変長パターンを音声特徴 (フレーム) 系列とする単語分類実験を通して評価した。音声データは ETL-WD-I を利用した。分節化 K 平均法による正規化長は 38 フレーム (ベクトルあたり 988 次元) とした。評価方法としては、男女各 3 名ずつの 492 単語を学習用標本として、残りの男女各 7 名ずつの 492 単語を試験用標本とした Hold-Out 法を用いた。また、カーネルはガウスカーネルを用いた。

初期のプロトタイプには、長さの正規化を経て得た学習用パターンに K -平均法を適用して、各単語 2 つずつを準備した。また、今回は重みのみの学習とする。

学習用	99.59%
試験用	60.03%

表に学習用と試験用のデータに対する分類率を示す。表から、少なくとも学習はほぼ適切に行われている様子がわかる。しかし、試験用データに対する分類率は大幅に低く、正規化パターンの高次元性あるいは (正規化に伴う) 情報欠損に因る過学習が生じた可能性がある。

5. おわりに

今後は、異なるフレーム数への正規化などを通して、試験用データに対する精度低下の原因を明らかにし、また重み係数のみならずプロトタイプをも学習する効果の検証を進めたい。

謝辞: 本研究の一部は科研費 (JP26280063) および文科省 H26 年度私学戦略的研究基盤形成支援事業「ドライバ・イン・ザ・ループ」の支援を受けて行われた。また、ETL-WD-I データベースは NII 音声資源コンソーシアムからの提供を受けた。

参考文献

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