

対象物抽出が不正確な場合における局所的特徴量を用いたパターン分類

鈴木 泉

長岡技術科学大学

1 はじめに

パターン分類のプロセスは、対象の観測、前処理、特徴抽出、分類の4つに大別され、対象の観測で重要なプロセスが対象物抽出である。例えば、画像分類であれば2次元の部分領域、テキストの分類であれば文字列（1次元）の部分領域を抽出する。対象の抽出が不正確であると分類精度は大きく低下する。しかし、不正確な対象の抽出に影響を受けずに、SVMなど既存の手法を使って分類することが出来れば、その応用範囲は広い。本稿では、局所的特徴量を用いることによる、multiclassの学習データ(TD)が与えられていない場合のデータ合成方法と、それをを用いた対象物抽出が不正確な場合の分類手法を提案する。

2 解決すべき課題

問題とするクラスを $W = \{w_1, w_2, \dots, w_K\}$ とし、特徴空間を $S = e_1 \times e_2 \times \dots \times e_M$ とする。対象を抽出するプロセスは、観測された情報の部分集合の1つを求めることと解釈できる。よって集合の包含関係より、対象の抽出が不正確な場合とは、以下の3つのケースのいずれか、または複数が起きるものと定義される[1]。

Multiclass case (MC):異なる複数クラスの対象を含む

Incomplete case (IC):対象の一部が欠ける

Other-class case (OC):クラス W 以外の対象を含む

MC と IC において所望される出力は以下のとおりである。単一のクラスの場合はそのクラスに分類し、MC の場合は対象のサイズ比に応じて、**図1** のように出力する。OC は1個のクラスとして扱い、MC に帰着させる。

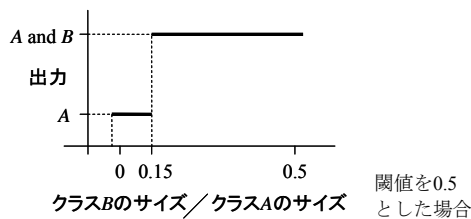


図1 Multiclass case(MC)における出力

問題を容易にするため、以下の2条件を仮定。

条件1. クラス $w_{1,2}$ の対象を任意に選び、その中のクラス w_k , $k = 1, 2$ の対象を抽出すると、 $p(\mathbf{x}|w_k)$ に従う。

(つまり、multiclass において、各クラスは独立)

条件2. クラス w_k の incomplete object は $p(\mathbf{x}|w_k)$ に従う

Pattern Classification Using Local features for Improper Object Extraction

Izumi Suzuki, Nagaoka University of Technology

3 既存の手法

Multi-label classification とは、TD と分類結果いずれも複数のクラスへの所属を許容する分類法であり、single class の TD の他に multiclass の TD が必要である。[2,3]. Bayesian spam filtering [4] など、統計的文書分類 (Bag-of-words 法、または BoW 法) の多くは、multiclass の TD は与えられなくとも MC に対応できる。例えば文書の言語判定において、ある言語 w_1 の TD は、様々なタイプの下位カテゴリに分けられる。例えば、TD は科学技術、歴史、文学、法律といったジャンルの何れかであるとする。仮に分類しようとする対象の一部が科学技術、残りが歴史の文書であったとしても、言語 w_1 と分類されることが期待できる。上記のようにクラス w_1 の TD の集合とクラス w_2 の TD の集合を合併した集合が multiclass $w_{1,2}$ の TD の集合となるためには、特徴量の値域は2値、つまり特徴空間は $\{0, 1\}^M$ と解釈する必要がある[1]。その場合の単純ベイズ識別器は以下のように特徴付けられる。

クラス w_k の TD を $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N$ とする。特徴量 e_m の尤度関数 $p(\mathbf{x}_m|w_k)$ を決定するのは $\mathbf{x}_m = 1$ における関数値 $p(\mathbf{x}_m = 1|w_k)$ のみであり (e_m が出現する確率と呼ぶ)、後述の式(3)における演算 \oplus を用い、次式で定義される。

$$p(\mathbf{x}_m = 1|w_k) = e_m(\mathbf{x}^1) \oplus e_m(\mathbf{x}^2) \oplus \dots \oplus e_m(\mathbf{x}^N) \quad (1)$$

対象 \mathbf{x}^0 の出力は通常の単純ベイズ識別器と同様:

$$\text{Class}(\mathbf{x}^0) = \arg \max_{w_k \in W} \prod_{m=1}^M p(\mathbf{x}_m^0 | w_k) \quad (2)$$

また、Tf-idf に代表される、特徴量の**重み付け**、また、分類対象で出現した特徴量に加重する特徴量の**制約**[1]もしばしば行われる。

テンプレート・マッチングは MC のみならず、OC にも対応できる。しかしテンプレート・マッチングが可能であるのは、「顔」や「道路標識」といったバリエーションの少ない対象物である。それ以外の、例えば「自動車」あるいは「歩行者」といった対象物に対しては、画像を特徴ベクトル化して情報量を減らした上で、パターン分類の手法を用いる必要がある[5]。

最後に、IC については、BoW 法においては文書が非常に短い場合を除き、対応可能である。また、OC は閾値を設定することで検出する機会が多いが、フィルターなどのように、1)問題とするクラスと 2)それ以外の全てのクラスで対象クラスが構成される場合がある。

4 解決方法

TD が与えられていない multiclass を学習させるための一方法は、条件1を活用して multiclass の TD を合成す

ることである。以下の条件を満たす特徴量を用いる。

条件 3. x^1, x^2 を対象, $M(x^1, x^2)$ を x^1 と x^2 から成る対象全体とする。対象 x の m 番目の特徴量を $e_m(x) \in [0, 1]$ と表わす。各特徴量 $e_m, m = 1, 2, \dots, M$ が

$$e_m(x^{1,2}) = e_m(x^1) \oplus e_m(x^2) \text{ for any } x^{1,2} \in M(x^1, x^2) \quad (4)$$

を満たすこと。ただし, \oplus は2項演算でなくても良いが交換, 結合則を満たす必要がある。(max 演算など)

例えば, 式(5)で定義される BoW 法における**相対頻度** $e_m(x) = f_m(x) / N_m(x)$ とその**加重平均**は, 式(4)を満たす。

$$e_m(x^1) \oplus e_m(x^2) = \{N_m(x^1)e_m(x^1) + N_m(x^2)e_m(x^2)\} / \{N_m(x^1) + N_m(x^2)\}$$

ただし $N_m(x) = \sum_m f_m(x), f_m(x)$ は特徴量の出現回数。(5)

また, 画像識別における局所の特徴量を用いた Bag-of-keypoints 法 (BoK 法) においても, 式(5)と同様の相対頻度と加重平均は条件 3 を近似的に満たす。

次に, IC を考慮する場合は, 条件2より, クラス w_1 と w_2 が $r_1 : r_2$ の場合は, $r_1 : r_2$ の加重平均を用いる。OC は以下のように扱う。OC は具体的にクラス $\{w_{K+1}, w_{K+2}, \dots, w_L\}$ から成り, それぞれの TD が与えられているとする。これらのクラスは分離可能である必要は無い。共起する可能性のあるクラス同士は TD を合成し multiclass を作る。すべての single-class と multiclass の TD を集合としての合併により, OC の TD を生成する。

最後に分類プロセスは, 複数のクラスをまとめ分類結果を絞り込む方法など, 様々なバリエーションが考えられる。本稿では紙面の都合上割愛する。

5 特徴空間の拡張

BoW 法では, 特徴量 e_m の尤度関数は $[0, 1]$ 上の連続関数ではなく, 2 値 $\{0, 1\}$ 上の関数であり, 式(1)で個々の TD の特徴量の相対頻度の加重平均を求めている。通常通り特徴空間を $[0, 1]^M$ とするには, 推定された分布 $p(x | w_k), k = 1, 2$ に対し,

$$p(z | w_{1,2}) = \int_z = ax+(1-a)y p(x | w_1) p(y | w_2) dz, 0 \leq a \leq 1 \quad (6)$$

によって確率変数を合成する。つまり, 尤度関数値のベクトル $p(x_m = 1 | w_k), m = 1, 2, \dots, M$ が確定値ではなく確率分布に従うと解釈し, 確率分布の「加重平均」を求めると拡張する[1]。

合成された TD で作成されたベイズ識別器は, MC に対処できる BoW 法の拡張となっている。すなわち, 核関数の合成で尤度関数を推定する場合に限れば, 式(6)に従って作った分布と, 合成された TD から推定された分布は同一である。

6 検証実験

文書分類 (言語判定) において以下の 2 つの実験を実施した。特徴量はいずれも byte 単位の 2-gram 列。使用した文書は英語(E), ドイツ語(G), フランス語(F), スペイン語(S), イタリア語(I), ポルトガル語(P)のいずれも科学技術関連の新聞記事。各文書はスペースを含め 450 文字程度。特徴量の重み付けはせず, 特徴量の制約のみ行った。

実験 1 IC + MC の検証

対象クラスは w_1 :E, w_2 :G。分類対象は w_1 と w_2 の比が 0, 0.1, 0.2, ..., 1 の 11 通りについて各 10 個。 w_1 と w_2 の TD はそれぞれ 15 文書を用意。クラス $w_{1,2}$ に対しては, 式(5)によって w_1 と w_2 の比率が 7:3, 5:5, 3:7 のデータをそれぞれ 15 個合成。学習後に尤度関数を重ね合わせた。考察で触れるが, 事前確率は $P(w_1) = P(w_2), P(w_{1,2}) / P(w_1) = 1.07$ とした。出力を $w_{1,2} = 1, w_1 = w_2 = 0$ として求めた 10 個の対象の平均値を図 2 (a) に示す。

実験 2 IC + OC の検証

対象クラスは w_1 :E。OC は $w_0 = \{w_2, w_3, \dots, w_6\} = \{G, F, S, I, P\}$ 。分類対象は w_1 と w_0 の比 w_1/w_0 が 0, 0.1, 0.2, ..., 0.5 の 6 通りについて各 10 個。TD は $w_1, w_2, w_3, \dots, w_6$ からそれぞれ 15 文書を用意。クラス $w_{1,0}$ に対しては, 式(5)によって w_1 と $w_i, i = 2, 3, \dots, 6$ の比率が 7:3, 5:5 のデータをそれぞれ 15 個合成。学習後に尤度関数を重ね合わせた。事前確率は $P(w_{1,0}) / P(w_1) = 1.02$ とした。出力を $w_0 = 1, w_1 = 0$ として求めた 10 個の対象の平均値を図 2 (b) に示す。実験 1, 2 とも, 図 1 のような理想通りでは無いものの, 各ケースへの対応が可能であることが確認される。

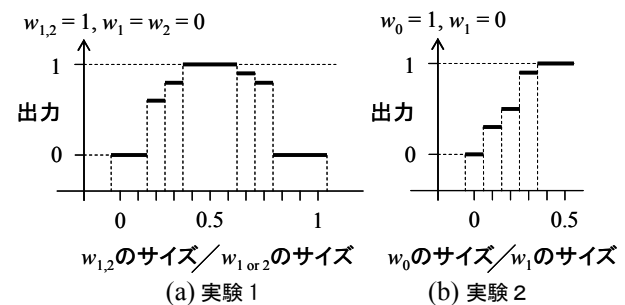


図 2 各実験のサイズ比毎の 10 対象の平均値

7 考察および結論

考察: 合成や合併によって尤度関数は生成することが出来るが, 事前確率は求めることが出来ない。これらを等確率としてしまうと著しく精度が低下する場合がある。よって本手法ではベイズ識別器よりは, SVM などの decision machine の利用が適している。一方で, class status value が求められるベイズ識別器には, クラスに属す度合いを数値化することが出来る長所がある。

結論: 局所の特徴量を用いることで, パターン分類において対象の抽出が不正確な場合においても, 通常の, つまり抽出が正確な場合のパターン分類法に帰着させることが可能であり, 画像分類への応用が期待される。

参考文献

- [1] I. Suzuki, Pattern Classification Using Bag-of-Keypoints for Improper Object Extraction, CINTI 2013, Budapest
- [2] G.Tsoumakas and I.Katakis, Multi-label classification: an overview, International Journal of Data Warehouse & Mining, 3(3), pp.1-13, 2007.
- [3] G. Madjarov, D. Kocev, D. Gjorgjevikj, and S. Džeroski, An extensive experimental comparison of methods for multi-label learning, Pattern Recognition, Vol. 45(9), 2012.
- [4] M. Sahami, S. Dumais, D. Heckerman and E. Horvitz, "A Bayesian approach to filtering junk e-mail", AAAI'98 Workshop on Learning for Text Categorization, 1998.
- [5] D. A. Forsyth and J. Ponce, Computer Vision: A Modern Approach, 2e, Prentice Hall, 2011.