階層的 Reed Muller 符号を用いた ECOC 法による多値文書分類

情報数理応用研究

5212C015-2　荻原大陸
指導教員　後藤正幸

Multi-valued Classification of Text Data based on ECOC Approach using Hierarchical Reed Muller Code

OGIHARA Tairiku

1　研究背景・目的

近年，コンピュータネットワークの発達に伴い，電子文書が大量に扱われるようになった．これらの情報は膨大であり，人手による分類が難しいため，このような電子文書を自動処理する技術の重要性は高まる一方である．文書自動分類技術の研究は多くあるが，その中でも Support Vector Machine (SVM) や Relevance Vector Machine (RVM) などの性能の良い二値判別器による自動分類手法が提案されている [1]．他方，多値判別問題に対しては，単一の多値判別器を構築するよりも，複数の二値判別器を組み合わせて多値判別を行う方が効率的であると指摘されている [1]．複数の二値判別器の組み合わせにより多値判別を行う方法としては，1-vs-the rest 判別法や Bradley-Terry モデルを用いた手法などが最も基本的であるが [2],[3]，本研究では符号理論 [4] の枠組みを用いた Error Correcting Output Coding(以下 ECOC) 復号法に基づく多値判別法 [5] に焦点を当てる．これは，対象となるカテゴリを二値分類器数の次元で構成される空間上の“ 符号語 ”に対応させ，二値分類器の出力結果からカテゴリを推定するものである．ECOC 法は使用する符号を予め設定する必要があり，その違いは分類精度の差に繋がると考えられるが，この点に着目した研究は少ない．本研究ではまず，数多く提案されている誤り訂正符号の中で ECOC 法に相性が良いと想定される BCH 符号，Reed Muller 符号（以下 RM 符号）を用いた ECOC 法による多値分類実験を行い，RM 符号の有効性を示す．しかしながら，従来の RM 符号を用いた ECOC 法では，全てのカテゴリを二等分する二値分類器しか構成することが出来ないため，全カテゴリの部分集合に対して二値分類するような判別器を組み合わせることで，精度向上の余地がある．また，ECOC 法による多値判別では，用いる判別器数の増加と共に精度が向上する傾向があるが，RM 符号による方法ではカテゴリ数に応じて予め判別器数が決められてしまう．実用場面では分析者が柔軟に判別器数を調整し，問題に合わせてカスタマイズできる方法が望まれる．そこで本研究では，RM 符号の特性を活かしつつ，階層的に組み合わせる階層的 RM 符号を用いた ECOC 法を提案する．階層的 RM 符号は，全体の判別問題を部分問題に分解することで，従来の RM 符号では判別が難しいカテゴリの部分集合内の判別が可能になる．また，各判別器の持つ学習データ数が多い従来の RM 符号と階層的 RM 符号を組み合わせ，互いの欠点を補いつつ判別器数を増加させることで，精度が向上することを示す．階層的 RM 符号は各判別器に使用する学習データ数が半分になるという特徴を持っているため，他の符号語を組み合わせる場合に比べ，計算量の膨大な増加を抑えつつ，判別器数を増加させることが出来る．提案手法を文書分類問題に適用し，精度・計算量の面で有効性があることを示す．

2　準備

2.1　多値判別問題

K をカテゴリ数，$\mathcal{C} = \{c_1, \ldots, c_K\}$ をカテゴリ集合とする．判別問題とはカテゴリが既知の学習データを使って判別ルールを学習し，カテゴリが未知の新たな判別対象データ $\boldsymbol{x}$ に対応するカテゴリ $c \in \mathcal{C}$ を推定することである．$K \geq 3$ の場合の多値判別問題に対し，本研究では「正解カテゴリ (1)」と「不正解カテゴリ (0)」の二値に判別する二値判別器を複数組み合わせることで入力データの所属カテゴリを推定する方法を考える．本研究では二値判別器として，各カテゴリに属する確率値を出力する RVM を用いる．

2.2　Relevance Vector Machine

RVM[2] は，回帰および分類問題を解くために提案された疎なカーネルベースのベイズ流学習手法であり，Silva らによって文書問題に適用されている [6]．優れた分類性能を持つ Support Vector Machine (SVM) の特性の多くを引き継ぎながら，確率モデルとして解釈できる点が最大の特徴である．入力ベクトルを $\boldsymbol{x}$，カテゴリラベルを $c \in \{c_1, c_2\}$，N 個のトレーニング文書セットを $\{\boldsymbol{x}_n, t_n\}_{n=1}^{N}$ とすると ($t_n \in \{c_1, c_2\}$)，$K = 2$ の場合の分類モデルは，$\boldsymbol{x}$ がカテゴリ $c_k (k = 1, 2)$ に判別される確率をロジスティック回帰関数を用い，

$$p(c = c_k \mid \boldsymbol{x}) = \frac{1}{1 + \exp(-f^k_{RVM}(\boldsymbol{x}))}, \tag{1}$$

$$f^k_{RVM}(\boldsymbol{x}) = \sum_{i=1, t_i = c_k}^{N} w_i K(\boldsymbol{x}, \boldsymbol{x_i}), \tag{2}$$

と表現できる．ただし，$w_i \sim N(0, \alpha_i^{-1})$ である．式 (2) における $K(.,.)$ はカーネル関数であり，入力された 2 つのデータ点を高次元空間上に写像し，内積を計算したものである．w_i は重み付けのパラメータであり，α の事後確率最大化により α_i^{-1} は推定されるが，その結果ほとんどの $\alpha_i \to \infty$ すなわち，w_i が 0 となる．w_i が 0 でないものを Relevance Vector (RV) と呼び，これらを用いて決定関数 $f_{RVM}(x)$ を構成する．RVM は高い汎化能力を持ち，出力が確率値で与えられる等，多くの利点を持っている．1 つの RVM で K 値判別を行う際には，K 個の線形モデルを組み合わせる確率的な方法を用い，α_i^{-1} は 2 カテゴリの場合と同じように計算する．しかし，この方法は学習にかかる計算量が 2 カテゴリ RVM の K^3 倍となってしまう点が不利である [1].

3　従来手法

複数の二値判別器 r $(r = 1, \ldots, R)$ の組み合わせにより多値判別器を構成する方法は，多くの有効な手法が提

案されている．本節では，その代表的な従来手法として，1-vs-the rest 判別法と最も一般的な ECOC 法である Exhaustive 符号を用いる方法について述べる．また，1-vs-the rest 判別と，ECOC 法に符号理論の枠組みを用いた BCH 符号と Reed Muller 符号を用いた判別性能を比較し，Reed Muller 符号の有効性を示す．

3.1 1-vs-the rest 多値判別法 [7]

1-vs-the rest 多値判別法では，全ての $k = 1, 2, ..., K$ に対し，判別対象カテゴリ c_k とそれ以外のカテゴリに分ける二値判別器を作成する．入力 $\boldsymbol{x}$ に対する K 個の判別器の出力を $\boldsymbol{H} = (H_{c_1}, H_{c_2}, ..., H_{c_K})$ とすると，

$$\hat{c} = \mathrm{argmax}_{c_k} H_{c_k}, \tag{3}$$

とするカテゴリ $\hat{c}$ に入力データ $\boldsymbol{x}$ を判別する．

ここで，判別器とカテゴリの関係を $\{1, 0\}$ の二値を要素とする $K \times R$ 行列 $\boldsymbol{W}$ で表わす．$\boldsymbol{W}$ の各列は各判別器 r が各カテゴリをどのように分類するかを表現し，1 に対応するカテゴリと 0 に対応するカテゴリを二分するものとする．このとき，$\boldsymbol{W}$ における各行が各カテゴリに与えられた符号語 $\boldsymbol{W}_{c_i}$ を意味する．図 1 は $K = 5$ における 1-vs-the rest 多値判別法の符号語構成である．

	01	02	03	04	05
W_{c_1}	1	0	0	0	0
W_{c_2}	0	1	0	0	0
W_{c_3}	0	0	1	0	0
W_{c_4}	0	0	0	1	0
W_{c_5}	0	0	0	0	1

図 1.1-vs-the rest 多値判別法の符号構成 (K=5)

3.2 ECOC 復号法に基づく多値判別法

ECOC(誤り訂正符号) とは，情報系列に対して機械で処理しやすい形で冗長性を付加し，誤り訂正能力を追加することで信頼の向上を図る技術である．多少の雑音が混入しても訂正が可能であるこの技術を多値判別問題に適用する事で，二値判別器の組合せで高精度な多値判別が期待できる．

3.2.1 Exhaustive 符号による多値判別法

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
W_{c_1}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
W_{c_2}	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
W_{c_3}	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1
W_{c_4}	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1
W_{c_5}	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0

図 2.Exhaustive 符号における符号構成 (K=5)

図 2 は Dietterich and Bakiri[5] によって提案された，ECOC 法の一つである Exhaustive 符号の $K = 5$ における符号構成である．例えば 3 つ目の判別器は入力データを $\{ c_1, c_4 \}$ と $\{ c_2, c_3, c_5 \}$ に分ける事を示している．判別フェーズでは，新たな入力 $\boldsymbol{x}$ に対する，各判別器の出力である確率値を $\boldsymbol{G} = (G_{c_1}, G_{c_2}, ..., G_{c_K})$ とした時，符号語 $\boldsymbol{W}_{c_i}$ の r 番目の判別器の値 $W_{c_i \cdot r}$ が 1 の場合は G_r，0 の場合は $1 - G_r$ をかけあわせ，以下の式で判別を行う．

$$\hat{c} = \mathrm{argmax}_{c_i} \prod_{r=1}^{R} G_r^{W_{c_i \cdot r}} (1 - G_r)^{1 - W_{c_i \cdot r}}. \tag{4}$$

この手法は，いくつかの判別器に誤りが生じても，他の判別器によってその誤りを訂正することが出来るため，判別精度が高いことが知られている．

3.2.2 BCH 符号による多値判別法

符号語間のハミング距離が大きく，誤り訂正の性能が良い符号語として BCH 符号がある．本研究ではこの多値分類問題への適用可能性についての検討を行う．いま，符号長 n，情報ビット数 k，誤り訂正可ビット数 t で定義される BCH 符号を，(n, k, t)BCH 符号と表す．この時，最小ハミング距離 d_{min} は $d_{min} = 2t + 1$ を満たす．このような性質を持つ BCH 符号を用いて符号語の集合を生成し，ECOC 復号法に適用する．(15, 5, 3)BCH 符号を用いて生成した $K = 8$ の場合の符号構成を図 3 に示す．

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
W_{c_1}	1	0	1	0	0	1	1	0	1	1	1	0	0	0	0
W_{c_2}	0	1	0	1	0	0	1	1	0	1	1	1	0	0	0
W_{c_3}	0	0	1	0	1	0	0	1	1	0	1	1	1	0	0
W_{c_4}	0	0	0	1	0	1	0	0	1	1	0	1	1	1	0
W_{c_5}	0	0	0	0	1	0	1	0	0	1	1	0	1	1	1
W_{c_6}	1	1	1	1	0	1	0	1	1	0	0	1	0	0	0
W_{c_7}	1	0	0	0	1	1	1	1	0	1	0	1	1	0	0
W_{c_8}	1	0	1	1	0	0	1	0	0	0	1	1	1	1	0

図 3.BCH 符号における符号構成 (K=8)

3.2.3 Reed Muller 符号による多値判別法

Reed Muller 符号 (RM 符号) は同じ誤り訂正能力の BCH 符号と比べると一般的に効率は悪いとされるが，符号構成が構造的に美しく，符号語の列における 1 と 0 の数の比が全て 1：1 であるという特徴を持つ．そのため，RM 符号を ECOC 法による多値分類に適用した場合,「正解カテゴリに属するデータ (1)」と「正解カテゴリに属さないデータ (0)」の比が等しい学習データを用いて二値判別器を構成する事ができる．ECOC 法による多値分類では，各カテゴリ間のハミング距離が大きいことに加え，各判別器における「1 に属するカテゴリの学習データ数」と「0 に属するカテゴリの学習データ数」の比が等しいことが高精度な分類に繋がる事が知られているため，各カテゴリの学習データ数が等しい場合には，RM 符号は有効に働くと考えられる．Reed Muller 符号を用いて生成した $K = 8$ の場合の符号語行列を図 4 に示す．

	01	02	03	04	05	06	07
W_{c_1}	0	0	1	0	1	1	0
W_{c_2}	1	0	0	0	0	1	1
W_{c_3}	0	1	0	0	1	0	1
W_{c_4}	1	1	1	0	0	0	0
W_{c_5}	0	0	1	1	0	0	1
W_{c_6}	1	0	0	1	1	0	0
W_{c_7}	0	1	0	1	0	1	0
W_{c_8}	1	1	1	1	1	1	1

図 4.RM 符号における符号構成 (K=8)

3.3 予備実験

符号理論の分野で示されている有効な符号を ECOC 法に適用した場合の性能を評価した研究は少ない．そこで，BCH 符号，RM 符号を適用した多値判別法の特徴を比較するため，新聞記事を用いた分類実験を行い，分類精度の評価を行った．実験には，読売新聞 2005 年の 8 カテゴリ (政治・経済・スポーツ・社会・文化・生活．犯罪事件・科学) の記事を使用した．すべての記事は 1 カテゴリのみに属し，カテゴリの重複はない．学習データを 1 カテゴリ 100 件，200 件，300 件として，それぞれ 3 回繰り返し，その平均値によって評価を行う．テストデータは一

律 500 件とする．特徴量としては単語頻度を使い，文書頻度 5 以上の単語によって特徴量空間を構成する．実験の結果を図 5 に示す．

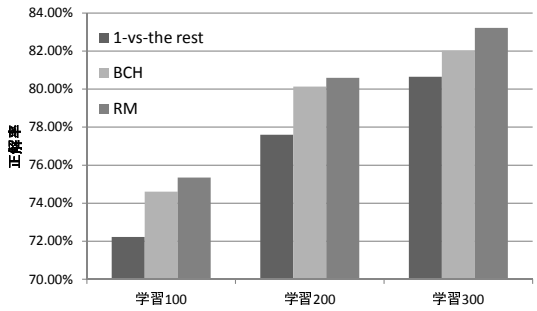

図 5. 符号語間の分類精度比較

表 1. それぞれ符号語における判別器数比較

判別手法	1-vs-the rest	BCH 符号	Reed Muller 符号
判別器数	8	15	7

図 5 より，全ての学習データ数に対して，RM 符号を適用した多値判別法が最も精度が高く，次いで BCH 符号を適用した多値判別法，1-vs-the rest 多値判別法であることがわかる．3.2 節でも述べたように,「正解カテゴリに属するデータ (1)」と「正解カテゴリに属さないデータ (0)」の比が等しい符号語構成を行うことが出来ているため，高精度な分類が可能であると考えられる．また，表 1 より，使用した判別器の数も RM 符号が最も少なく，計算量の面からも RM 符号の有効性が示せたと言える．このように，RM 符号は学習データの偏りを生じることなく，ECOC 法に適した有効な符号語を作成することが出来ることが分かる．

4 提案手法

4.1 提案手法の概要

従来の RM 符号を用いた手法では，カテゴリ数に対応したパラメータを持つ RM 符号を用いたのみであり，判別器数が事前に設定されてしまう．しかしながら，カテゴリ数が事前に設定されている場合，ECOC 法では判別器数を増加させることで精度の向上が見込めることが知られており，分析者がより柔軟にこれらを調節できることが望ましい．また，従来の RM 符号を用いた手法では全てのカテゴリを二分する二値分類器に関して，データの偏りを考慮しただけであり，カテゴリの部分集合を二分する判別を行うことは想定していない．例えば，図 4 の 3 列目の判別器は入力データを $\{ c_1, c_4, c_5, c_8 \}$ と $\{ c_2, c_3, c_6, c_7 \}$ に分ける事を示している．しかし，$\{ c_1, c_4, c_5, c_8 \}$ 内における二値判別は行われておらず，これらを更に判別することで精度向上の余地がある．例として，経済，社会，政治，スポーツの 4 つのカテゴリを持つ新聞記事を二分する事を考える．従来手法では，全体を二分する判別器のみを用いているため,{経済・社会} と {政治・スポーツ} の様に全体を二分することしか出来ず，スポーツのように統計的特徴が他カテゴリと大きく異なるカテゴリが判別の性能を悪化させている可能性がある．この問題を解決するために，4 カテゴリのうち {経済} と {スポーツ} を二分するような部分問題に分解することで，統計的に異なる特徴を持つカテゴリの影響を軽減することが期待できる．そこで，まず本研究では，RM 符号の特徴を保ちつつ判別器数を増加させるために RM 符号を階層的に用い，複数の小さな RM 符号によって一つの二値判別問題を複数の部分問題に分解する方法を示す．部分問題に分解することにより，各判別器に使用できる学習データ数は少なくなるが，従来の RM 符号では判別が難しかったカテゴリの部分集合内の判別が可能になる．また，詳細な判別が難しいが各判別器の持つ学習データ数が多い従来の RM 符号と，階層的 RM 符号を組み合わせることで，互いの欠点を補い，精度の向上の余地がある．さらに，階層的 RM 符号では，各判別器に使用する学習データ数が半減する特徴から，全体の計算量の増加を抑えつつ，判別器数を増加することができる．この様に，本研究では，階層的 RM 符号を作成し，従来の RM 符号と組み合わせることで，計算量の増加を抑えつつ，精度が向上する符号語の作成法を提案する．

4.2 符号語作成アルゴリズム

提案手法の符号語作成アルゴリズムは「正解カテゴリに属するデータ (1)」と「正解カテゴリに属さないデータ (0)」の比が等しい学習データを用いて各判別器を学習でき，増加させる各判別器の計算量を抑えつつ判別器数を増やすことができる．そのため，提案手法は現実的な計算量で，高精度な分類が期待できる．提案手法では，RM 符号を階層的に用いて小さな RM 符号を複数作成し，二値判別問題を部分問題に分解する．具体的には，前述した $\{ c_1, c_4, c_5, c_8 \}$ と $\{ c_2, c_3, c_6, c_7 \}$ に分ける二値判別問題を $\{ c_1, c_8 \}$ と $\{ c_4, c_5 \}$ を分ける部分問題と $\{ c_2, c_7 \}$ と $\{ c_3, c_6 \}$ に分ける部分問題等に分解することで，カテゴリの部分集合内の二値判別を可能とする．そして，複数の小さな RM 符号を組み合わせ，階層的 RM 符号を構成し，従来の RM 符号と組み合わせる．提案手法の特徴として，部分問題では，対象とするカテゴリが従来の RM 符号の二値判別問題が持つ 8 カテゴリから，4 カテゴリとなり，各判別器に使用する学習データ数が半減するため，一般的な符号語を組み合わせた場合よりも計算量の大幅な削減が見込める．以下では，$\boldsymbol{W}$ の中で判別に用いないカテゴリが存在するため，これを $*$ と表記するものとする．以下の 7 つの Step を用いて階層的 RM 符号と従来の RM 符号を組み合わせる事で提案手法の符号語を作成する．

Step0) K に対応したパラメータを持つ符号長 M の RM 符号を作成し，符号語 $\tilde{\boldsymbol{W}}_{c_i}$ の m 番目 ($m = 1, 2, ..., M$) の判別器の値を $\tilde{W}_{c_i \cdot m}$ とする．

Step1) $m = 1$ とする．

Step2) $\tilde{W}_{c_i \cdot m} = 1$ のカテゴリを $C_m^1 = \{c \in C \mid \tilde{W}_{c_i \cdot m} = 1\}$，$\tilde{W}_{c_i \cdot m} = 0$ のカテゴリを $C_m^0 = \{c \in C \mid \tilde{W}_{c_i \cdot m} = 0\}$ とする．

Step3) C_m^1 に対して，C_m^1 のカテゴリ数に対応したパラメータを持つ小さな RM 符号を作成する．C_m^0 に対するカテゴリは判別せず，それらのカテゴリの場所には $*$ を挿入する．

Step4) 逆に C_m^0 に対して，C_m^0 のカテゴリ数に対応したパラメータを持つ小さな RM 符号を作成する．C_m^1 に対するカテゴリは判別せず，それらのカテゴリの場所には $*$ を挿入する．

Step5) $m = m + 1$ とし，$m \leq M$ であれば，Step3 へ戻る．さもなくば，得られた $m = 1, 2, ..., M$ に対する小さな RM 符号を列方向に並べ，階層的 RM 符号を構築する．

Step6) K に対応したパラメータを持つ step0 の RM 符号と階層的 RM 符号を組み合わせる．

□

	01	02	03	04	05	06	07	08	09	10	11	12	⋯	47	48	49
Wc_1	0	0	1	0	1	1	0	*	*	*	1	1	⋯	1	1	0
Wc_2	1	0	0	0	0	1	1	1	1	0	*	*	⋯	*	*	*
Wc_3	0	1	0	0	1	0	1	*	*	*	0	1	⋯	*	*	*
Wc_4	1	1	1	0	0	0	0	0	1	1	*	*	⋯	0	1	1
Wc_5	0	0	1	1	0	0	1	*	*	*	0	0	⋯	*	*	*
Wc_6	1	0	0	1	1	0	0	0	0	0	*	*	⋯	0	0	0
Wc_7	0	1	0	1	0	1	0	*	*	*	1	0	⋯	1	0	1
Wc_8	1	1	1	1	1	1	1	1	0	1	*	*	⋯	*	*	*

図 6. 提案手法の符号語構成

図 6 は $K = 8$ の場合の提案手法によって作成された符号語である．各判別器に与えられる学習データ内の「正解カテゴリに属するデータ (1)」と「正解カテゴリに属さないデータ (0)」の比は等しいまま，従来法に比べ大幅に判別器数が増加している．また，階層的 RM 符号の各判別器が持つカテゴリ数が半減し，計算量が削減されているため，全体の計算量を抑えながら判別器数を増加させることが出来る．

5 実験・考察

提案手法の有効性を検討するために，新聞記事を用いて分類実験を行い，分類精度及び計算量の評価を行った．実験条件は 4.5 節と同様とした．比較対象として，従来の RM 符号を用いた判別器構成，RM 符号に 1-vs-the rest 判別法を組み合わせたものを用いた．実験結果を図 7 に示す．

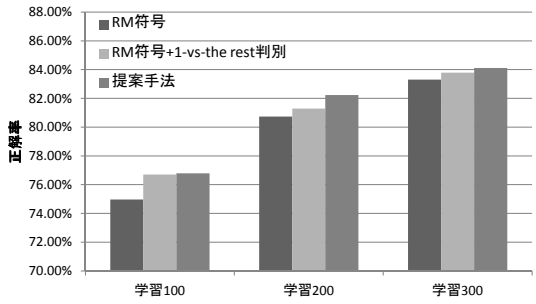

図 7. 符号語間の分類精度比較

図 7 より，全ての学習データ数において，提案手法が比較手法よりも精度が高いことが分かる．これは，RM 符号の特徴を保ったまま，部分集合内の二値判別が可能になるように判別器を増加させたこと，また従来の RM 符号と階層的な RM 符号を組み合わせることで，お互いの欠点を補う判別が可能になったためである．

次に計算量に関しての実験結果を図 8 に示す．

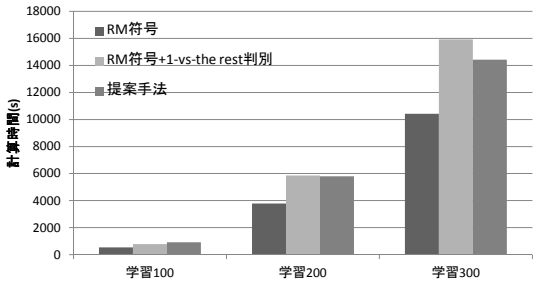

図 8. 符号語間の計算量比較

表 2. それぞれ符号語における判別器数比較

判別手法	RM 符号	RM 符号+1-vs-the rest	提案手法
判別器数	7	15	49

図 8 及び表 2 より，提案手法は計算量の増加を抑えながら判別器数を大幅に増加させることが出来ていることがわかる．これは，各判別器が持つカテゴリ数を半減し，学習データ数を強制的に半分にしたことで，学習データ数が多い場合に計算量が膨大に増加するという RVM の欠点を補うことが出来たためと考えられえる．また，学習データが少ない場合には，各判別器の学習データの削減の効果が薄く，判別器数の増加による影響の方が大きいため，比較手法よりも提案手法の計算量が増加しているが，その差は微小である．

以上の議論に加えて，ECOC 法が持つメリットとして並列処理が可能であることが挙げられる．並列処理における計算時間では，判別に使用する複数の二値判別器の中で，計算時間が最も長い判別器がボトルネックとなる．従来の RM 符号に使用されるいずれかの判別器の計算時間は，階層的 RM 符号と 1-vs-the rest 判別に用いられる全ての判別器よりも長いため，全手法において学習にかかる時間は同じである．すなわち，提案手法は並列処理によって計算時間を増加させることなく，分類精度を向上させることが可能であると言える．

6 まとめと今後の課題

本研究では，複数カテゴリの文書分類問題を対象とし，RM 符号の特性を保ちつつ，判別器数を増加させることで計算量を抑えつつ精度を向上させる符号語の構成法を提案した．また，実際の新聞記事の分類問題に適用し，評価実験によって有効性を示した．今後の課題として，q 元の RM 符号を用い，0,1,*の数のバランスを考慮した手法の提案や，カテゴリ数に制約がない場合の手法の提案が挙げられる．また，各カテゴリによってデータの偏りがある条件における有効性の検証も必要である．

参考文献

[1] C. M. Bishop. “Pattern Recognition and Machine Learning,” *Springer-Verlag,* 2006.

[2] M. E. Tipping. “Sparse Bayesian Learning and the Relevance Vector Machine,” *Journal of Machine Learning Research,* 211-244, 2001.

[3] S. Ikeda, “Combining binary machines for multi-class: Statistical model and parameter estimation,” *The Institute of Statistical Mathematics,* 58, 157-166, 2010

[4] 平澤茂一，西島利尚，“符号理論入門,” 培風館，東京，1999.

[5] T. G. Dietterich and G.Bakiri. “Solving Multi-class Learning Problems via Error-Correcting Output Codes,” *Journal of Articial Intelligence Research,* vol.2, pp. 263-286, Jan.1995.

[6] C. Silva and B. Ribeiro. “Scaling Text Classification with Relevance Vector Machines,” *IEEE International Conference on Systems, Man, and Cybernetics,* 4186-4191,2006.

[7] T. K. Huang and R. C. Weng, C. J. Lin, “Generalized Bradley-Terry models and multi-class probability estimates,” *Journal of Machine Learning Research,* Vol.7, pp.85-115, 2006.