

二種類の SVM を用いたオンライン類似数式文字識別

糟谷 勇児[†] 山名 早人^{††, †††}

[†] 早稲田大学大学院理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1

^{††} 早稲田大学理工学術院 〒169-8555 東京都新宿区大久保 3-4-1

^{†††} 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

E-mail: [†] kasuya-u@yama.info.waseda.ac.jp, ^{††} yamana@waseda.jp

あらまし 数式を効率よくコンピュータに入力する方法として、手書きの数式を認識し、コンピュータで利用可能な形式に変換するシステムの開発が望まれている。しかし、数式中の文字や記号の認識は一般の文字と異なり r と γ 、 w と ω のような類似文字が多く存在するため困難である。そこで本稿では 2 種類の SVM (Support Vector Machine) を使い分け、類似の文字を精度よく識別する手法を提案する。二種類の SVM とは(1)画像を入力として用いる SVM と(2)ペンの位置の時系列情報を入力とする SVMGDTW(SVM Gaussian Dynamic Time Warping)である。本手法により r と γ は 86.7%、 w と ω は 85.6% で判別できることを確認した。また本手法を DTW とテンプレートマッチングからなる認識手法と併用したところ、数字・英字・ギリシア文字・数式記号からなるサンプル(全 106 文字)に対して 91.1% の認識率を得た。

キーワード 数式認識、オンライン手書き文字認識、SVM

Recognition of Similar Character Pairs with Two Types of SVMs for Online Mathematical Expression Recognition

Yuji KASUYA[†] Hayato YAMANA^{††, †††}

[†] Graduate School of Science and Engineering, Waseda University Okubo3-4-1, Shinjyuku-ku, Tokyo, 169-8555 Japan

^{††} Science and Engineering Waseda, University Okubo3-4-1, Shinjyuku-ku, Tokyo, 169-8555 Japan

^{†††} National Institute of Informatics Hitotsubashi2-1-2, Chiyoda-ku, Tokyo, 101-8430 Japan

E-mail: [†] kasuya-u@yama.info.waseda.ac.jp, ^{††} yamana@waseda.jp

Abstract Mathematical expression recognition systems which recognize mathematical expressions and translates them into digital data usable in computer is needed. However characters and symbols in mathematical expressions are sometimes similar and difficult to discriminate. This paper proposes a method to recognize similar character pairs with two types of SVM (Support Vector Machine). One is normal SVM which uses images of handwriting as input; the other is SVMGDTW which uses sequences of pen position. With the proposed method, “ γ ” and “ r ” are discriminated with a recognition rate of 86.7% , “ ω ” and “ w ” with 85.6%. Besides, the proposed method combined with DTW and template matching recognizes numbers, alphabets, Greeks and mathematical symbols (105 characters in total) with 91.1% recognition rate.

Keyword Mathematical expression recognition, Online handwriting character recognition, SVM

1. はじめに

数式を効率よくコンピュータに入力する方法として、手書きで書いた数式を認識し、コンピュータ上で使用可能な形式に変換するシステム(以下、数式認識システム)の開発が望まれている。しかし、数式は分数や累乗などの上下にも広がる構造を持つため、画像からの認識では文字の 1 つ 1 つを切り出すことが困難である。そこで文字の切り出しが比較的容易に行える、ペンの位置座標の時系列情報(オンライン情報あるいはストローク情報という)を利用した数式認識システムが研究されている[1-4]。筆者らはオンライン数式認識システムの作成を目指して、デジタルペンを用いて数

式サンプルの採取を行っている[5-6]。本稿では採取したサンプルを用い、数式に使用される文字や記号(以下、数式文字)1文字を認識する手法について述べる。

数式文字は数字、英字、ギリシア文字、数式記号からなる。数式文字の認識と通常の文字認識が異なる点は、 v と ν (ニュー)、 K と κ (カッパ)、 r と γ (ガンマ)、 w と ω (オメガ)のように、類似した文字が多数存在することである。この問題に対してヒューリスティックに基づいた方法[3]や文字の書き方に制限を持たせる方法[2]が提案されているが、類似の文字を統一的に扱う方法は提案されていない。より多数の文字を扱ったり、ユーザの筆跡に基づいた判別を行ったりす

るためには、筆跡からの学習に基づく類似文字の判別手法が必要である。

そこで本稿では 2 種類の SVM（画像を入力として用いる SVM とオンライン情報を入力とする SVMGDTW）を使い分け、類似の文字を精度よく認識する手法を提案する。

以下、2 節ではオンライン手書き数式認識における本研究の位置づけについて、3 節では SVM について、4 節では提案手法について、5 節では提案手法を検証するための実験についてそれぞれ述べ、6 節でまとめを行う。

2. オンライン手書き数式認識

オンライン手書き数式認識はペン先の位置の時系列情報から認識を行い、コンピュータで扱うことができる形式に変換する作業である。本節ではオンライン手書き数式認識について軽く触れ、本研究の位置づけについて述べる。

一般にオンライン手書き数式認識は 3 つのフェーズからなる[4]。すなわち、

- (1)文字の切り出し
- (2)切り出した文字の認識
- (3)数式構造の構成

である(図 1)。(1)では手書きの数式を一文字単位に分解する。(2)では(1)で切り出した文字をそれぞれ認識し文字を判別する。(3)では(2)の結果から分数や累乗、上付き下付きの数式などの数式の構造を判別する。

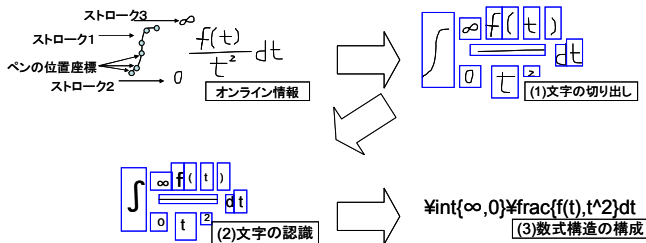


図 1 数式認識の流れ

本研究は(2)の文字の認識に関するものである。文字の認識の結果は数式構造の基礎となるだけでなく、認識結果をフィードバックすることで、より正確に文字の切り出しを行うことを可能にする。このため、数式文字を精度よく認識する手法を研究することは数式認識システムを構築する上で最重要の課題である。

既存の研究では、数式文字の認識を行う手法として、隠れマルコフモデルを用いた方法[1]、2次元ワープ法を用いた方法[3]などが提案されている。しかし、[1]の手法は英字で 89.2%、ギリシア文字で 89.7%と必ずしも高い精度は出ていない。[3]の手法では英字、数字、ギリシア文字、数式記号(全 122 文字)に対して 96.6%の精度であるが、テストに用いているサンプルは 3 人分のみである。一般に、人数を増やした場合、認識率

は低下すると考えられる。このように数式文字の認識は英数字や漢字の認識と比較して精度が低い傾向にある。この原因として数式には類似の文字(γ と r 、 κ と K 等)が多いことが挙げられる。そこで特に類似の文字を判別する方法として[2]では、入力ストロークの開始点、終点、屈曲点の前後でのストローク情報の変化などから認識を行うという手法が用いられている。しかし、このようなヒューリスティックに基づく方法は

- ・ システムの設計者が認識対象ごとに対処法を考えなければならないため、類似文字が多数ある場合は対応が難しい
 - ・ あらかじめ設定された対処しか行わないため、ユーザ個人の筆跡に適応させた認識ができない
- 等の問題点がある。そこで本研究では類似の文字を統一的に扱う手法として SVM を使用した認識手法を提案する。SVM は、サンプルから学習を行うため、サンプルさえあれば自動で学習・認識ができ、ユーザ個人の筆跡に対応した認識が可能である。

3. SVM

SVM (Support Vector Machine) は有効な 2 クラス分類の手法として、文字認識だけではなく、自然言語処理や、バイオインフォマティクスなど様々な分野で使用されている。本章では SVM について簡単に説明する。詳しくは[7]などを参照されたい。

ここではクラス A に属する n_A 個のサンプル $x_i (i=1,2,\dots,n_A)$ とクラス B に属する n_B 個のサンプル $x_i (i=n_A+1,\dots,n=n_A+n_B)$ から、入力 x が A と B のどちらのクラスに属するかを SVM を用いて判別する方法について述べる。ただしここで A を正のクラス、B を負のクラスとし、 $y_i=1 (1 \leq i \leq n_A)$ 、 $y_i=-1 (n_A+1 \leq i \leq n)$ となる定数 y_i を導入する。SVM は

$$f(x) = \sum_{i=1}^n \lambda_i y_i K(x_i, x) + b \quad (1)$$

を計算し、 $f(x) \geq 0$ ならば入力 x をクラス A、 $f(x) < 0$ ならば入力 x をクラス B として判別する。

式(1)の λ_i はラグランジュ乗数であり、サンプルからの学習により決定する変数である。 λ_i の値は以下の式(2)を満たしつつ式(3)最大化させることにより決定する。

$$\sum_{i=1}^n \lambda_i y_i = 0 \quad (2)$$

$$\sum_{i=1}^n \lambda_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \lambda_i \lambda_j K(x_i \cdot x_j) \quad (3)$$

式(1)の b は閾値であり、サポートベクトル x_+ ($\lambda_i \neq 0$ となる x_i) と x_+ に対応する y_i の値 y_+ から

$$b = y_+ - \sum_{i=1}^n y_i \lambda_i K(x_i, x_+) \quad (4)$$

のように自動的に決定することができる。あるいは、

テストサンプルに対して認識を行い認識率から試行錯誤的に決定することも可能である。

式(1)(3)の $K(u,v)$ はカーネル関数と呼ばれる関数である。入力 u,v がベクトルであるとき、カーネル関数としては一般にガウシアンカーネル

$$K(u,v) = \exp\left(-\frac{|u-v|^2}{\delta^2}\right) \quad (5)$$

がよく用いられる。ただし、 $|u-v|$ は u から v までのユークリッド距離であり、 δ は拡散の大きさを表す定数である。また δ は使用者が適宜決めるパラメータである。 u,v がベクトルではなく時系列情報の場合、Gaussian Dynamic Time Warping Kernel (GDTWK) と呼ばれるカーネル関数が提案されている[8]。ここで $D_{DTW}(u,v)$ を DTW (Dynamic Time Warping) により求めた距離とする。DTW とは時系列情報を伸縮を許して対応付け、動的計画法により最短の対応を計算する距離の測定方法である。GDTWK では

$$K(u,v) = \exp\left(-\frac{D_{DTW}(u,v)}{\delta^2}\right) \quad (6)$$

によりカーネル関数を定義する。このカーネル関数を用いた SVM である SVMGDTW は時系列の判別、特にオンライン文字認識に有効であることが分かっている。

4. 提案手法

本節では提案手法である SVM を用いた類似数式文字の認識手法について述べる。提案手法では、既存の文字認識手法を用いた認識を最初に行い、次に相互に誤認識が起こりやすい文字のペアについてのみ SVM を用いて詳細な判別を行う。既存の文字認識手法として、本稿では DTW(Dynamic Time Warping)とテンプレートマッチングを組み合わせた認識法を用いる。SVM を用いた判別には、画像を入力とする SVM(SVM2D と呼ぶ)と角度の時系列情報を入力とする SVMGDTW を使い分ける。以下 4.1 では文字認識手法の全体について説明し、4.2 では DTW、4.3 ではテンプレートマッチング、4.4 では SVM を用いた類似文字の判別について詳しく説明する。

4.1 文字認識手法の全体

ここでは文字認識手法の全体の流れについて図 2 を用いて示す。以下、①～⑤までの順序で認識の流れについて説明する。

- ① 入力としてペンの位置座標を時系列でサンプリングしたものを用いる
- ② 入力は点の数が位置や人によって偏りがある。このため点の位置を、点の間のユークリッド距離が均等になるように並べなおす。
- ③ DTW を用いて認識を行う。DTW を用いて入力をデータベースと照合し、最も類似性が高い 2 種類

の文字を出力する。これを第一候補、第二候補とする。ただし、データベースには学習サンプルを②と同じ均等化を行い、式(7)で計算される角度情報に変換し、あらかじめ登録しておく。

- ④ データベース中の③の第一候補、第二候補の 2 種類の文字と入力をテンプレートマッチングで比較する。テンプレートマッチングによる類似度が最も高かったサンプルに対応する文字を④の結果とする。ただし、第一候補と第二候補の間で DTW の距離が大きく離れている場合 (実験 3 では 100 以上とした) 第一候補をそのまま認識結果とする。入力と比較する画像はあらかじめ 2 値画像化してデータベースに登録しておく。
- ⑤ ④の認識結果が SVM2D、SVMGDTW を行う文字としてリストに登録されていれば、それぞれの SVM で認識を行い、その結果を最終結果とする。登録されていなければ④の結果を最終結果とする。SVM2D、SVMGDTW により判別を行う文字は事前に実験を行い、効果が高かった文字をあらかじめ設計者がリストに登録しておく。

以上のように①～④までの既存の認識手法の結果に、類似文字がある場合のみ⑤で SVM を使用することで、類似文字を高い精度で認識しつつ、使用する SVM を最小限に抑え、効率の良い認識を行うことができる。

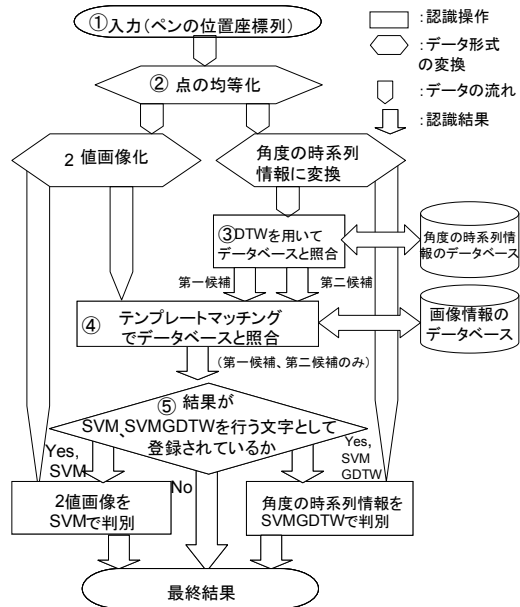


図 2 提案手法の流れ

4.2 DTW

DTW は二つの時系列の間の距離 (正確には距離の公理を満たさないため距離ではない) を動的計画法によって計算する手法である。本稿では DTW への時系列の入力としてペンが (x_1, y_1) から (x_2, y_2) に移動したときの角度

$$\theta = \arctan((y_2 - y_1)/(x_2 - x_1)) \quad (7)$$

をサンプリングしたものを用いる。ただし角度と角度の距離 (DTW に使用するローカルな距離) は、比較する二つの角度を θ_1, θ_2 とすると、 θ_1 と θ_2 の間の角の内、角度の小さいほうの角度とする。DTW による文字認識では、学習サンプルと入力を DTW で比較し、距離の値が最小となるサンプルに対応する文字を認識結果とする。

4.3 テンプレートマッチング

本稿で用いるテンプレートマッチングは、オンライン情報を 2 値画像に変換し、2 値画像同士の類似度を計算する手法である。2 値画像の作成方法を以下の図 3 用いて説明する。まず入力ストロークの外接長方形を求め、外接長方形を $m \times n$ ピクセルに分割する。次に全てのピクセルについて、線が通っているピクセルの値を 1 に、そうでないピクセルの値を 0 にする。

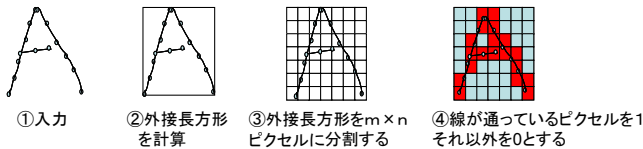


図 3 2 値画像の作成法

テンプレートマッチングでは、図 3 の方法で作成した 2 値画像 u, v の左から i 番目、上から j 番目のピクセルの値をそれぞれ u_{ij}, v_{ij} とすると

$$D_{TM}(u, v) = \sum_{ij} (u_{ij} - v_{ij})^2 \quad (8)$$

により二つの画像の距離を求める。テンプレートマッチングによる認識では入力と距離の値が最小の学習サンプルに対応する文字を認識結果とする。

4.4 SVM を用いた類似文字判別提案手法

提案手法では 2 種類の SVM、すなわち SVM2D と SVMGDTW を対象となる文字によって使い分ける。

SVM2D は 4.3 の 2 値画像を入力として用いる。このため、カーネル関数を以下の式(9)で定義する。

$$K(u, v) = \exp\left(-\frac{D_{TM}(u, v)}{\delta^2}\right) \quad (9)$$

学習は学習サンプルから(2)(3)式に基づいて行う。SVM2D は画像から判別を行うため、線の長さなどの 2 次元的な差異に基づいた判別が可能である。

SVMGDTW は(7)式の角度の時系列を入力に用いる。すなわち、4.3 で説明した DTW を行い(6)式のカーネル関数を計算する。SVMGDTW は筆跡の時系列を入力とするため、書き順や線の曲がり具合などの時系列的な差異に基づいた判別が可能である。

提案手法では誤認識が起りやすい文字のペアにたいして SVM2D と SVMGDTW を、実際に適用してみ、認識精度が高い方の SVM を使用する。各文字のペアに対して、SVM2D と SVMGDTW のどちらが有効

かは実験 1 で確認する。また閾値 b については(4)式を用いて決定するだけでなく、ユーザや文脈に合わせて動的に変更する。例えば、 γ を余り使用しないユーザの場合であれば閾値を調整することで、 r の方に偏った認識を行うことができる。このような閾値の変更による認識率の変化については実験 2 で確認する。

5. 実験

本節では、提案手法である SVM を用いた数式文字認識に関して実験を行い、その有効性を検証する。

実験 1 では、誤認識が起りやすい文字のペア 10 組について DTW、テンプレートマッチング、SVM2D、SVMGDTW の 4 つの認識機で個々に認識を行い、認識率を測定し比較する。実験 1 により SVM2D および SVMGDTW が類似文字の判別にどの程度有効であるかを検証する。

実験 2 では、SVM の閾値 b を $-1 \sim 1$ まで 0.1 刻みで変更した場合、判別する二つの文字のそれぞれの認識率がどのように変わるのかを調査する。実験 2 により、4.3 で述べた閾値の動的な変更が認識に効果的であるかを検証する。

実験 3 では、4.5 で提案した認識手法について、SVM を使用する場合と使用しない場合両方の認識率を測定し比較する。実験 3 により、提案手法による類似文字判別が、文字認識全体にどの程度効果があるのかを検証する。ただし、実験では[5]で採取したサンプル各文字 60 個 (60 人分) を用いた。サンプルはペンテルの airpen を用いて、毎秒 20 回の間隔でサンプリングされたものである。

以降、実験 1, 2, 3 について、それぞれ 5.1、5.2、5.3 で述べる。

5.1 実験 1

本実験では類似文字のペア 10 組、(a,q)(g,y)(h,n)(i,j)(p,D)(\gamma,r)(\kappa,K)(v,v)(\tau,t)(\omega,w) を認識し認識率を測定した。認識手法として DTW、テンプレートマッチング、SVM2D、SVMGDTW の 4 つを用いた。判別を行う文字のペアは事前に DTW とテンプレートマッチングによる認識を行って、誤認識が多いものを選んだ。なお(6)(9)式の δ は SVM2D では $\delta = 2, 3, 4, 5$ 、SVMGDTW では $\delta = 8, 9, 10, 11$ としてそれぞれの場合について実験を行った。2 値画像のサイズは縦横 10×10 とした。認識率はクロスバリデーション法を用いて測定した。ここでは、60 個のサンプルを 20 個ずつの 3 組にわけ、その内 2 組を学習用に、残り 1 組をテスト用に使用した。これを学習用とテスト用の組み合わせを変えて全 3 回行い、認識率を測定した。表 1 に認識結果を示す。ただし表 1 では、スペースの関係からテンプレートマッチングを TM と表記した。

表 1 から SVM を用いることにより、(h,n)を除く全て

の組み合わせに対して認識率が上昇したことが分かる。特に(y,g)は SVMGDTW で 6.6%、(γ,r)は SVMGDTW で 6.7%、(ω,w)は SVM2D で 5.9%、(τ,t)は SVMGDTW で 5.0%の認識率の向上が見られた。その他(i,j)では SVMGDTW で 100%の識別が可能になった。これらの文字に対して SVM による認識は有用であるといえる。

表 1 類似文字の認識結果[%]

認識対象	DTW	TM	SVMGDTW					SVM2D				
			8	9	10	11	2	3	4	5		
aq	90.8	97.5	90.0	90.0	89.2	89.2	96.7	98.3	99.2	99.2		
pD	83.3	92.5	71.7	72.5	73.3	76.7	96.7	96.7	96.7	96.7		
nh	89.2	95.8	71.7	71.7	73.3	75.8	93.0	95.0	95.8	95.8		
j	99.2	90.0	100	100	99.2	99.2	90.0	90.8	92.5	93.3		
yg	91.7	74.2	96.7	98.3	97.5	97.5	80.8	83.3	85.0	81.7		
γr	80.0	75.8	85.8	86.7	85.0	85.0	81.7	82.5	81.7	83.3		
ωw	76.3	79.7	82.2	83.1	83.1	83.9	81.4	83.9	83.1	85.6		
τt	90.8	88.2	95.8	95.0	94.1	94.1	91.6	93.3	92.4	91.6		
κK	58.3	52.5	58.3	57.5	58.3	58.3	54.2	59.2	58.3	55.8		
νv	61.7	55.8	58.3	59.1	58.3	61.6	63.3	62.5	60.0	60.8		

ただし、(κ,K)(ν,v)に関しては SVM を用いても若干の上昇は得られたものの、認識率は低いままであった。そこで、認識に失敗したサンプルを図示してみたところ、サンプルの筆者が κ,K と ν,v をうまく書き分けられていないことが分かった。図 3、図 4 に κ と K、ν と v のサンプルの内、認識に失敗した最初の 5 つを例として示す。図 3 から、κ のサンプルが K と見分けがつかず、逆に K のサンプルのほうが曲線的で κ らしく見えることもあることが分かる。また図 4 から、ν と v はほとんど同じで見分けがつかず、書き分けができていないことが分かる。よって現状では、これらの文字の認識は難しく、正確な認識を行うためには、ギリシア文字と英字の書き分けができる人から改めてサンプルを採取する必要がある。

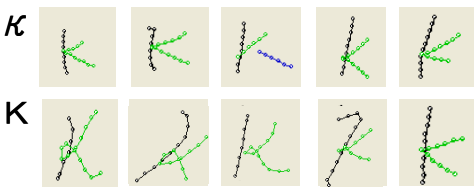


図 3 κ (上) と K (下) の例



図 4 ν (上) と v (下) の例

5.2 実験 2

本実験では SVM の閾値 b を変化させて、判別する文字それぞれの認識率がどのように変化するかを測定した。ここでは実験 1 で精度が大きく向上した、(y,g)

(γ,r)(τ,t)(ω,w) を例にとり実験を行った。認識は SVMGDTW を用いて行い、閾値 b を -1 から 1 まで 0.1 刻みで変化させ、実験 1 と同様のクロスバリデーションで認識率を測定した。表 2 に、それぞれの文字のペアを閾値を変えて認識したときの認識率を示す。また図 5 に (γ,r) の認識率をグラフにしたものを示す。ただし、表 2、図 5 の X 認識率は文字 X のサンプルを X として認識できた割合を示す。図 5 の横軸は閾値の値を、縦軸は認識率を表し、r γ 識別率は r と γ を正しく識別できた割合を示す。

表 2 閾値を変化させた場合の認識率の変化[%]

認識対象	閾値	-0.5	-0.3	-0.1	0	0.1	0.3	0.5
		γ,g	γ 認識率	15	18	43.3	98.3	100
	g 認識率	100	100	100	96.7	35	0	0
γ,r	γ 認識率	0	5	53.3	90	98.3	100	100
	r 認識率	100	98.3	93.3	83.3	63.3	30	10
τ,t	τ 認識率	39	49	89.8	91.7	98.3	98.3	100
	t 認識率	100	100	96.7	96.6	73.3	45.1	20
ω,w	ω 認識率	100	98.3	90	75	45	25	1.6
	w 認識率	3	44.5	82.8	93.1	98.3	100	100

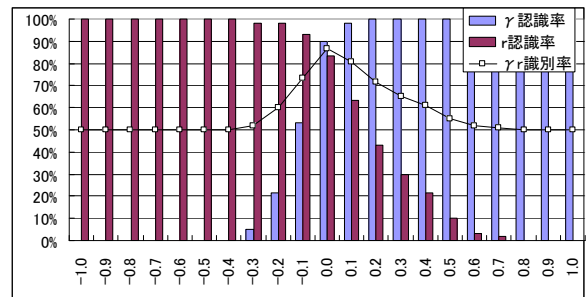


図 5 γ と r の識別の認識率と閾値の関係

図 5 から、(γ,r) では最大の認識率を示す閾値は b=0 であることが分かる。また、b=-0.2 とすれば、r をほとんど認識しつつ、γ をある程度認識することができる。さらに b=-1 とすれば、書かれた文字に関わらず r と認識することになる。これを利用すると、ユーザの文字の使用頻度に合わせて閾値を調整することで、ユーザに合わせた認識が可能となる。例えば、γ を全く使用しないユーザに関しては b=-1 とし、半々で使用するユーザに関しては b=0 とすればよい。加えて文脈に合わせてシステムが動的に閾値を変更することにより、文脈に即した判別が可能となる。例えば、同じ式の中で γ が使用されていたら、次に書かれた文字も r より γ である可能性が高い。この場合 b の値を少し増加させることにより、γ と認識され易くすることができる。これらのことは表 2 から、他の文字の判別においても同様であることが分かる。

実験 2 の結果から、閾値を動的に変更することでユーザや文脈に合わせた認識が効果的にできることが分かる。

5.3 実験3

本実験では、4.1の認識手法において、SVMによる類似文字認識の有無で、全体としてどの程度認識率が変化するかを測定した。認識対象とした文字を表3に示す。認識対象の文字は英字(46字)、数字(10字)、ギリシア文字(23字)、数式記号(27字)の全106字である。ただし、英字のO,oは0と、P,V,S,Zはそれぞれの小文字p,v,s,zと見分けがつかないため使用しなかった。また1と同じ書き方のl、9と同じ書き方のq、×と同じ書き方のxは使用しなかった。これらの文字は文字認識での判別が難しいため、一般に書き方に制限を設けたり、他の文字との大きさの差で判断を行ったりといった方法で対処が行われる。ギリシア文字、数式記号は使用頻度を基準にそれぞれ23文字、27文字を選び認識対象とした。サンプルは各文字60個(60人分)を用い、その内40個(40人分)を学習サンプル、20個(20人分)をテストサンプルとした。認識する文字の中でSVM2D、SVMGDTWを使用して認識を行うペアを表4に示す。SVMを使用する文字はDTWとテンプレートマッチングによる認識で相互に誤認識が起りやすく、かつSVMを使用することで認識精度が向上するものを選んだ。SVMのパラメータは実験1、2と同様のクロスバリデーションを行って、最も認識率が高いときの値とした。SVM2DとSVMGDTWのどちらを使用するかも同様に認識率が高いほうとした。また2値画像のサイズは縦横8×7とした。

本実験の結果を表5示す。表5から提案手法を使用しない場合と比べて、提案手法を使用した場合、認識率が1.46%改善したことが分かる。これは誤認識の内、15.4%が改善されたことを意味する。本実験から一部の誤認識が起りやすいペアに対して、提案手法を適用することで、全体的な認識率の向上が可能であることが分かる。

表3 実験3で用いた文字の種類

英字 46 字	abcdefghijklmnopqrstuvwxyz ABCDEFGHIJKLMNOPQRSTUVWXYZ
数字 10 字	1234567890
ギリシア文字 23 字	$\alpha \beta \gamma \delta \eta \zeta \theta \kappa \lambda \mu \nu \xi \pi$ $\rho \tau \phi \psi \omega \Gamma \Delta \Pi \Sigma \Omega$
数式記号 27 字	+ - ± ∓ ÷ × < > ≤ ≥ = ≠ ∫ ∫ ∂ ∇ () {} [] √ ∞ % →

表4 SVM2D,SVMGDTWを使用した文字のペア

SVM2Dを使用したペア	(a,9)(j,J)(p,D),(l,)])(δ,8)(η,n)(ω,w)
SVMGDTWを使用したペア	(y,g)(γ,r)(κ,K)(τ,t)

表5 提案手法使用前後での認識率と誤認識率[%]

	認識率	誤認識率
提案手法使用前	89.6	10.4
提案手法使用后	91.1	8.9

6. 終わりに

本稿では2種類のSVM(画像を入力として用いるSVM2Dとペンの位置の時系列を入力とするSVMGDTW)を使い分け、類似の文字を精度よく認識する手法を提案した。また提案手法に対して、筆者らが[5]で採取したサンプルを対象として検証を行った。

実験1ではωとwをSVM2Dを用いて85.6%で認識し、γとrをSVMGDTWを用いて86.7%で認識するなど、SVMが類似文字の判別に有用であることを示した。

実験2ではSVMの閾値を変更することにより、ユーザの文字の使用頻度や文脈に合わせた認識が可能であることを確認した。

実験3ではDTWとテンプレートマッチングからなる認識手法に、提案手法を適用した結果、適用前に比べて誤認識の15.4%を改善し、91.1%の認識率で数式文字106文字を認識可能であることを確認した。

ただし、κとK、νとvなどはサンプルを採取する段階で文字の書き分けができていないことも明らかになった。これに対しては、ギリシア文字を書き慣れた人からサンプルを採取しなおしたり、書き方に自然な限定を付たりなどして対処を行いたい。またpとP、sとSなど本質的に書き方の変わらない文字については認識を保留している。これらの文字については、文字認識の段階での判別は難しいため、数式を入力する段階で、ユーザがうまく書き分けることが出来るようなインターフェースが必要となる。今後はこのようなユーザにとってもシステムにとっても都合の良い数式認識インターフェースを提案していきたい。

文 献

- [1] U. Garain and B. B. Chaudhuri, "Recognition of Online Handwritten Mathematical Expression," IEEE trans. on syst., man, and cybern. Part B, Vol.34, NO.6, pp.2366-2376, Dec.2004.
- [2] 青島史郎,鈴木隆広,森健策,末長康仁,"実時間手書きストローク解析による数式入力システム,"信学論(D-II), Vol.183, No.5, pp.1232-1245, May.2000.
- [3] 田畑耕一,福田亮治,鈴木昌和,"2次元ワーブを併用したオンライン英数字・数学記号認識,"信学技報 PRMU, Vol.100, No.701, Mar.2001
- [4] K. Chan and D. Yeung, "Mathematical expression recognition: a survey," IJDAR, Vol.3, No.1, pp.3-15, Aug.2000
- [5] 糟谷勇児,山名早人,"デジタルペンを用いた数式サンプル採取システムと採取文字に対する認識実験,"信学技報 PRMU, Vol.105, No.374, pp.7-12, Oct.2005
- [6] 数式サンプルデータベース
<http://www.yama.info.waseda.ac.jp/~kasuya-u/>
- [7] N. Cristianini and J. Shawe-Taylor, サポートベクターマシン入門, 大北剛(訳),(社)共立出版,東京,2005
- [8] C. Bahlmann,B. Haasdonk and H. Burkhardt,"On-line Handwriting Recognition with Support Vector Machines-A Kernel Approach," IEEE proc. of the IWFHR'02, Aug.2002