

# デジタルペン筆記における学習状況把握のための 削除・被削除ストロークの検出

前田 祐樹<sup>1</sup> 三浦 元喜<sup>2</sup> 小林 祐作<sup>3</sup>

**概要:** デジタルペンによって学習者の筆記活動を収集すると、他の学習者の参考になる生の筆記データをプロジェクタに投影したり、活動内容を分析したりすることが可能になる。しかし、限られた授業時間内で、教師がすべての学習者の学習状況を把握することは難しい。そこで我々は、とくに学習者の筆記における、削除・修正箇所とその時点における状況（いつ、どのような筆記を行っているときに、間違いに気づいて修正したのか？）を教師が着目したり把握したりしやすくするため、サポートベクターマシン(SVM)を用いて削除・修正箇所を自動的に抽出する方法を開発した。実際の筆記データを収集し、削除箇所特有の筆記特徴量を計算したうえで、SVMを用いて学習・判別した場合の精度について検討した。その結果、削除コマンドについては最適な条件下において約8割、被削除ストロークについては84%の認識精度が得られることを確認した。

**キーワード:** オンライン筆記分類, 学習状況の把握, SVM

## Detecting Erase Strokes for Understanding Learning States from Online Handwritten Notes

MAEDA YUUKI<sup>1</sup> MIURA MOTOKI<sup>2</sup> KOBAYASHI YUSAKU<sup>3</sup>

**Abstract:** Collecting of student notes by digital pen enables teachers to share the student notes to the class by a projected screen, and analyze the notes to enhance group learning. However, it is difficult for teachers to grasp all students' learning states in the lecture time. Therefore, we focus on the deleted strokes that represent student mistakes. We collected freestyle handwritten student notes, and manually labeled to the notes. Based on the labeled notes, we extracted features for the erase symbols and the deleted strokes. We have tested support vector machine (SVM) techniques to classify the erase symbols and deleted strokes from usual handwritten notes. We confirmed the proposed method can correctly classify 80% of erase symbols and 84% of deleted strokes under the best conditions.

**Keywords:** Online Handwritten Note Classification, Grasping Learning States, SVM

### 1. はじめに

教育の情報化が浸透し、一般的な教室においてタブレットやノート PC, スマートフォンなどのデジタル機器を学習者が日常的に使用することが増えている。学習者がネットワークに接続されたデジタル機器を使用すると、

教師が学習者の回答や考えを授業時間中に適宜収集し、全体的な学生の回答状況を知ることができる。こうした学習者の回答や考えを収集して、類似しているものをまとめて表示することで、学習者は他人の考えと自分の考えを対比させたり、関連付けたりすることができるようになる。

タブレットやノート PC, スマートフォンなどのデジタル機器は、ディスプレイに教材や講義資料を表示することができたり、教師からのフィードバックを提示することができるため有用性が高い。一方、こうした多機能なデ

<sup>1</sup> 九州工業大学 工学部 総合システム工学科

<sup>2</sup> 九州工業大学 基礎科学研究系

<sup>3</sup> 九州工業大学 工学部 電気電子工学科

デジタル機器やデバイスは、起動や入力のための操作が複雑になりがちである。そのため、学習者が操作に慣れるための時間を必要としたり、学習には直接関係ない知識を得る必要があったりする。

学習者の操作に対する負荷や負担を最小限に抑える方法として、デジタルペンによって学習状況を収集する方法が提案されている [1], [2], [3]。デジタルペンにはいくつかの方式があるが、アノト方式のデジタルペンは特殊なドットパターンが埋め込まれた専用紙を、ペン先のカメラが読み取って筆記を電子化する方式を採用しており、一般的な「紙」と「ペン」のみで学習者の活動（回答や考え）を収集できるという特徴がある。アノト方式のデジタルペンには通常、ディスプレイは付属していないため、教師が学習者に個別にフィードバックを返すことはできないが、集約した筆記をプロジェクトで投影することによって学習者に他者の回答や考えを見せることができる。

デジタルペンによって学習者の筆記を逐次集約して、プロジェクトで提示する利点として、(1) 学習者の間違いを含む、生の回答を提示できること、(2) 集約した筆記を計算機で分析・判定して、リアルタイムに採点結果や類似性を求め、フィードバックできることの2つが挙げられる。また、(3) 詳細な学習者の作業ログが残るので、事後の分析が行えることも利点のひとつである。しかし、デジタルペンの利点や学習効果をさらに高めるためには、(3) の分析を事後ではなく、できるだけ授業時間内に行うことが望ましいと我々は考えている。例えば、学習者が授業時間内に他の学習者のミスを知ることによって、学習者に与えられた問題のミスするポイントを学び、自分がそのミスをしていないかの確認ができる。そこで我々は、教師が学習者個別の作業筆記から学習者の詳細な状況を把握する時間を短縮し、授業時間内にも行えるようにするためのアプローチのひとつとして、機械学習技術を適用して筆記の分類と解析を行うことを考えた。本稿ではその最初の取り組みとして、我々はサポートベクターマシン (SVM) 技術を使用し、学習者の手書き筆記ノートにおける削除・修正箇所を自動的に抽出するための方法と、適用事例について議論する。我々がなぜ削除・修正箇所の抽出に着目したのかについて最初に説明する。その理由としては、教師が学習者の活動のうち、削除や修正を抽出して着目しやすくすることで、学習者の活動経緯や当時の状況（いつ、どのような筆記を行っているときに、間違いに気づいて修正したのか？）を推測しやすくする効果があると考えたからである。さらに、学習者のデジタルペンで書かれたもの全てを映すと一つ一つが小さくなり、小さな文字を読む必要がある。しかし、削除・被削除ストロークを検出し色を変化させることでより着目しやすくなる。

もうひとつの理由として、学習者が筆記を公開されるときに恥ずかしい気持ちを軽減することができる可能性が

挙げられる。アノト方式のデジタルペンはペン先がボールペンで作られており、また通常は消しゴム機能が付属していないため、間違いを消すということが原理的に行いにくい。そこで、学習者が事前に間違いに気づいて削除しなかった部分を自動的に抽出して、プロジェクトに投影するときに非表示にしたり、薄い色で表示することができれば、計算ミスや記述の間違いを含む「生の筆記」が他者に公開されることの心理的な負担が多少なりとも軽減できるのではないかと考えたからである。

本研究の目的は、文章のほかに数式や図形を含むような、学習者のノート筆記データから、SVMを利用して削除箇所（被削除ストローク）を検出することである。ここで、被削除ストロークは、学習者が取り消し線や塗りつぶしの線によって、明示的に削除したい意志を表明した筆記部分とする。この削除意志を表明するための取り消し線や塗りつぶし線のことを、本研究では「削除コマンド」と呼ぶことにする。削除コマンドでも被削除ストロークでもない筆記については、以下の論文では「通常の筆記」と表現する。

## 2. 削除・修正箇所検出機能実現にあたっての考え方

削除・修正箇所の検出機能を実現するにあたって、我々は以下の項目を重視することにした。

- デジタルペンによる筆記集約手法の特性として、学習者に事前に操作方法について説明したり準備期間を設ける必要がないことが挙げられる。このことから、削除コマンドについても事前に説明したり強制するのではなく、学習者が自然に削除目的で行った筆記について、システムの側で柔軟に対応すべきであると考えている。
- 上記にも関連するが、削除コマンドの定義は必要に応じて教師が簡単に更新・修正できることが望ましい。
- デジタルペンをを用いた筆記集約活動には、文字や文章だけでなく、図や数式が含まれる可能性がある。あらかじめ入力エリアを規定したり、入力可能な文字種を限定したりするのではなく、学習者が通常行っているノート筆記について削除・修正箇所を検出できることが望ましい。

第1の項目を実現するために、我々は学習者が自然に記述した筆記データを収集したうえで、それぞれの筆記について(1) 削除コマンド、(2) 被削除ストロークの分類を手作業で行い、提案手法の性能を検証する際の正解データとした。

第2の項目は、今後筆記データの特性が異なる場合に備えて、教師がサポートベクター分類器のモデルを簡単に修正できるようなインターフェースを設けることによって実現する。

第3の項目は、第1、第2の項目とも関係するが、入力

される可能性がある筆記データを制限せず、どのような形式の筆記であってもシステムが対応できるようにする必要があるので、図や数式が含まれるような筆記データを収集し、検証に用いることにした。

### 3. 提案手法

本研究の目的は、学習者の自由筆記データ（通常の筆記）から、削除コマンドと被削除ストロークを抽出・分類することである。ここで、被削除ストロークと通常の筆記の違いについて考えてみると、被削除ストロークは最初に筆記された時点では、通常の筆記として書かれているため、被削除ストロークと通常の筆記の違いはほとんどない。被削除ストロークを特徴づけるためには、削除コマンドがあらかじめ決定している必要がある。したがって、我々はまず、すべての筆記データから削除コマンドのみを抽出する。その後、抽出した削除コマンドの情報を利用して、被削除ストロークおよび、通常の筆記の特徴量を計算し、被削除ストロークを抽出するという手順をとる。以下では、この手順について、削除コマンドを抽出する「フェーズ1」と、被削除ストロークを抽出する「フェーズ2」に分けて説明する。

#### 3.1 フェーズ1：削除コマンドの抽出

今回我々は、学習者が数式や図形を含めて回答した筆記を収集し、そのなかに含まれていた学習者が自由な発想で記入した削除コマンドを分析した。その結果、(1) 2本以上の長い横線（図1）、(2) 大きなバツ印（図2）、(3) 塗りつぶし線（図3）の3つが代表的な削除コマンドであることがわかった。

これらの代表的な削除コマンドを特徴付ける指標として、我々は筆記データの各ストローク（ペンで書き始めてから、ペン先を離すまでのひと続きからなる筆記）に対して、以下の特徴量を計算することにした。

- ストロークの長さ。2つ以上の長い横線や、バツ印、塗りつぶしのストロークの長さは、通常の筆記（文字）と比べて大きい。ただし、図形や、数式における答えの下に書く「下線」との区別は難しい。
- ストロークを包含する矩形領域 (BoundingBox) の幅と高さ。2つ以上の長い横線、バツ印および塗りつぶし線における、包含矩形領域の幅と高さは、通常の筆記よりも大きい。（表 A-2 の「width」および「height」）
- ストロークの複雑さ（折れ点の数）。塗りつぶし線は、通常の筆記よりも多くの折れ点を含んでいる。ここで、折れ点の数は後述する Ramer-Douglas-Peucker アルゴリズムを用いて計算した。（表 A-2 の「ramer」）
- 対象となるストロークよりも前に書かれたストロークとの交差の回数。削除コマンドのストロークは通常、以前に書いた筆記に重ねられることが多い。したがっ

て、交差の回数が通常の筆記よりも多くなる。（表 A-2 の「cross」）

- 対象となるストロークと、その前後のストロークとの、以下の4つの数値 (a,b,c,d) からなるベクトル量のコサイン類似度。(a) 最小二乗法によって計算したストロークの傾き、(b) 主成分分析により計算したストロークの直線性、(c) ストロークを包含する矩形領域 (BoundingBox) の幅および (d) 高さ。この値は、2本以上の長い横線およびバツ印を識別するために導入した。（表 A-2 の「erase」）

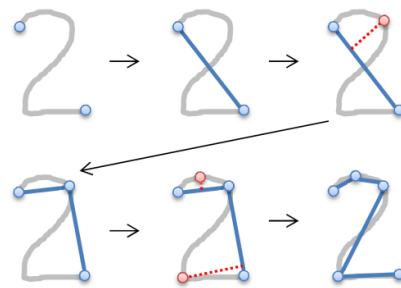


図4 Ramer-Douglas-Peucker アルゴリズム

ここで、折れ点の数を計算するための Ramer-Douglas-Peucker アルゴリズム [4]（以下、我々は Ramer の方法と呼ぶ）について説明する。Ramer の方法では、まず、すべてのストロークの開始点と終了点を特徴点（図4）として取得する。直線距離がしきい値（図4、右上）よりも大きかった場合には、隣接する特徴点間の直線から最も離れた点が特徴点として選択する。それ以上の特徴点を選択されなくなるまで、この選択は、再帰的に行う。我々はしきい値として、ストロークを包含する矩形領域の高さと幅のうち、大きい方の「5分の1」の長さを用いた。

#### 3.2 フェーズ2：被削除ストロークの抽出

被削除ストロークの抽出にあたっては、フェーズ1と同様に、筆記データの各ストローク（ペンで書き始めてから、ペン先を離すまでのひと続きからなる筆記）に対して、以下の特徴量を計算することにした。

- 対象ストローク筆記後に書かれた削除コマンドのうち、対象ストロークの矩形領域と一部でも重なりをもった削除コマンドの数。（表 A-3 の「oc」）
- 上記における「重なりをもった削除コマンド」において、対象ストロークと重なった部分の面積の合計。
- 対象ストローク後に書かれた、削除コマンドのストロークと交差する点の数。（表 A-3 の「crossa」）
- 上記の「交差する削除コマンド」が書かれた時刻と、対象ストロークが書かれた時刻の差。「交差する削除コマンド」が複数存在する場合は、その最大値。（表 A-3 の「otd」）

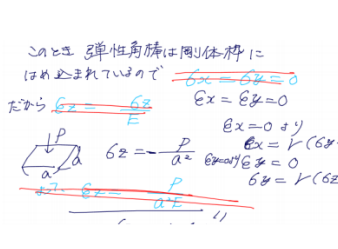


図 1 2本以上の長い横線

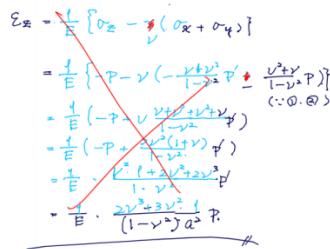


図 2 バツ印

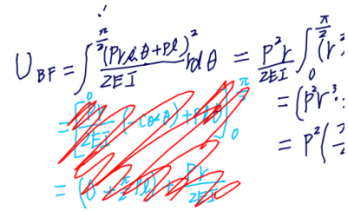


図 3 塗り潰し

これらの特徴値はフェーズ1によって決定された、削除コマンドで計算されることに注意する。重なり判定については、ストロークや削除コマンドの矩形領域 (BoundingBox) を使用した。すなわち、2つの矩形領域がすこしでも重なりをもっていた場合に、「重なっている」という判定を行っている。

#### 4. 実験

我々は材料力学の講義を対象として、40名の学習者の手書き筆記データを収集した。学習者にはアノト方式のデジタルペンを用いて、演習問題を解いてもらった。このとき、デジタルペンの筆記は削除できないこと、および、削除が必要な場合は取り消し線や塗りつぶし線など、自由な方法で削除筆記をしかまわらないことを説明した。実験にあたっては、具体的な取り消し線や塗りつぶし線などの削除方法については指示を行っていない。

講義の後、我々はすべての筆記データに対して、各自の判断に基づき、手作業によって削除コマンドと被削除ストロークに色付けを行った。図7に、色付けの結果を適用した学習者筆記を示す。赤と青の色は、それぞれ、削除コマンドと被削除ストロークを示す。この色情報を用いて、我々は正解となるラベル付きのデータを生成した。ただし、一部の筆記について、筆記データに含まれる時刻情報および時系列情報が壊れていたため、それらのデータは分析対象から除外した。

表A-1に、ラベル付けされた筆記データについて、シート毎の通常筆記数 (normal)、削除コマンド数 (erase)、被削除ストローク数 (deleted) を示す。シートIDは、講義日 (最初の桁) とペンID (最後の2桁) で構成した。なお、学習者が1日の講義で記入した筆記は、A4サイズの用紙 (シート) 1枚以内に収まっている。

これらのデータを用いて、我々はまず、削除コマンドを抽出するフェーズ1の分類性能を検証した。付録に掲げた、表A-2は、SVMのモデル構築 (学習) を行う上で、通常筆記と被削除ストロークを合わせたストローク (cls0) の数 (num) および、削除コマンド (cls1) の数 (num) と、上記で説明したフェーズ1のための特徴量の基本的な統計量 (ただし、シートの数が多いので、3分の1のみを掲載した) を示している。また、図5は表A-2に掲載できなかった

た残りのシートの筆記を含む、それぞれのストロークの特徴量の平均値を比較するためのグラフである。シート126および130には削除コマンドが含まれていなかったため、cls1の行は省略した。「erase」は、3.1のコサイン類似度を表す。「cross」と「ramer」もそれぞれ交差数、折れ点の数を表す。

我々はJava版のLIBSVM[5]を用いて、各シートに対するクロスバリデーションと、SVMにおける最適なカーネルパラメータの探索を行うグリッドサーチによって分類性能を評価した。表A-4は、フェーズ1の分類の結果を示す。削除コマンドを構成するストロークの総数が、355であり、通常筆記+被削除ストロークの総数よりも少ないので、クロスバリデーションにおける対象シート以外に含まれる、すべての削除コマンドストロークをモデル構築に利用した。通常筆記+被削除ストロークについては、クロスバリデーションにおける対象シート以外の筆記から、ランダムに抽出して、モデル構築に利用した。

表A-4には、各シートを抜き出して行ったクロスバリデーションにおいて、モデルを構築する際のグリッドサーチを行い、最適なF値となるカーネルパラメータ結果を求めたデータについて、すべてのシートに対しての平均値と標準偏差を示している。なお、「通常筆記+被削除」の数字は、ランダム抽出する際のストローク数を示す。Avg(F) およびSD(F) は、F値の平均値と標準偏差を示す。F値は、精度と再現率から計算される。Cとγは、グリッドサーチで求めたRBFカーネルパラメータのCとγの最適値である。「通常筆記+被削除」ストロークの数が増加するに従い、認識率が向上した (図8)。これは、様々なバリエーションの通常筆記および被削除ストロークが学習に利用されることで、削除コマンドの抽出に成功する可能性が高まったことを示している。この結果から、カーネルパラメータが適切に選択された場合、削除コマンドの70~80%を正しく分類することができることが確認された。

つぎに、被削除ストロークを抽出するフェーズ2について述べる。付録に掲げた表A-3は、SVMのモデル構築 (学習) を行う上で、通常筆記 (cls0) と被削除ストローク (cls2) の数 (num) と、フェーズ2のための特徴量の基本的な統計量 (3分の1のみ) を示している。また、図6は表A-3に掲載できなかった残りのシートの筆記を含む、それぞれの



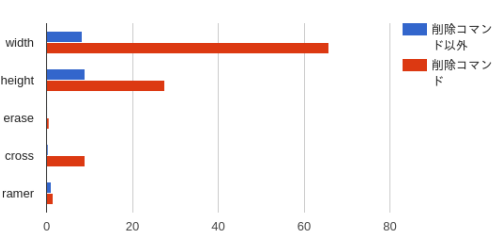


図 5 「フェーズ 1」 削除コマンド以外と削除コマンドの比較

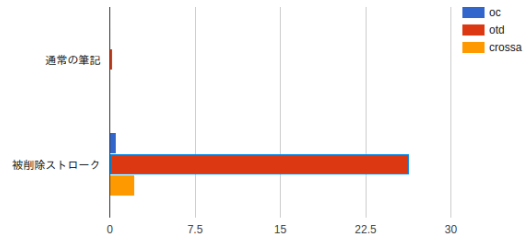


図 6 「フェーズ 2」 通常の筆記と被削除ストロークの比較

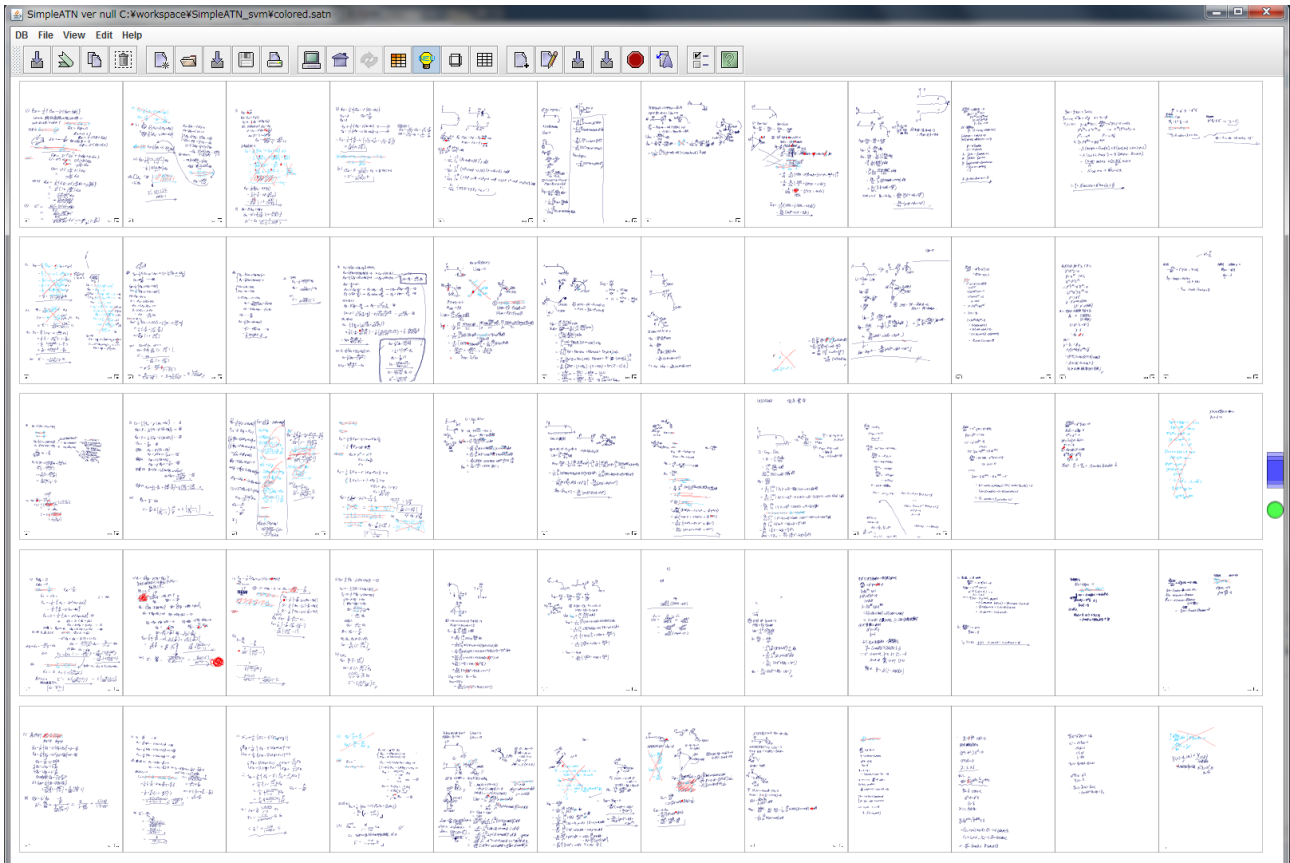


図 7 学習者によって書かれたノート

ストロークの特徴量の平均値を比較するためのグラフである。「oc」、「otd」と「crossa」はそれぞれ、削除コマンドとの重なり回数、時刻差、および交差点の数を表す。表 A-3 から、通常筆記 (cls0) と被削除ストローク (cls2) の特徴量について、平均値に差があることが確認できた。

フェーズ 2 についても、フェーズ 1 と同様、各シートを対象としたクロスバリデーションを実行した。また、各シートに対する最適なカーネルパラメータをグリッドサーチで求めた。結果を表 A-5 に示す。また、図 9 に、モデルの学習に用いたストローク数と、F 値の関係を示す。なお、フェーズ 2 においては、削除コマンドはモデルの学習および筆記の分類から除外している。モデル構築に用いた通常筆記と被削除ストロークは、対象シートを除いたすべてのシートから無作為に選択した。結果として、モデル構

築に用いた通常筆記および削除コマンドの数は、認識精度に影響を与えないことがわかった。

なお、フェーズ 2 の実験では、我々はすべて削除コマンドが適切にフェーズ 1 で検出したことを想定している。そのため、実際の状況ではフェーズ 2 の認識率は表 A-5 に示す値よりも悪くなると考えられる。

## 5. 関連研究

SVM によって手書き文字を分類・認識する研究はこれまでに様々な試みがなされている。Ahmad らは文字認識に SVM 技術を適用し、IRONOFF と UNIPEN データセットを用いた評価を行った [6]。その結果は、SVM の認識率が文字レベルにおいては有意に優れていることが示された。Bahlmann らは動的時間伸縮法 (Dynamic Time Warping)

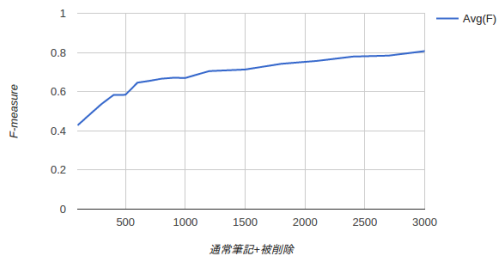


図 8 「フェーズ 1」 F 値と通常筆記+被削除の数の関係

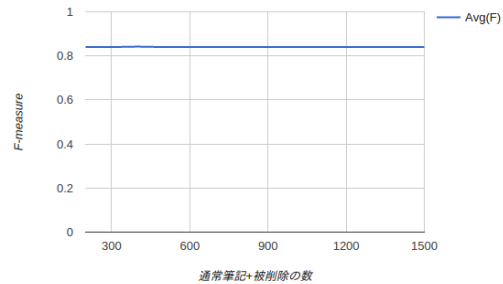


図 9 「フェーズ 2」 F 値と通常筆記+被削除の数の関係

と SVM を組み合わせた、SVM カーネル構築法による文字分類手法を提案している [7]。Huang らは隠れマルコフモデル (HMM) と SVM を組み合わせた文字認識アプローチを提案している [8]。彼らは複数の HMM を特徴抽出器として使い、SVM を使って UNIPEN データセットの文字を分類した。HMM を特徴抽出器として利用したことによって、特徴ベクトルの次元を下げることに成功している。また、Huang らのグループでは、マルチレイヤフォワードネットワークと HMM を組み合わせることにより、手書き記号認識のための高速特徴選択モデル [9] も発表している。

これらの SVM によって手書き文字を認識する手法は、本研究における SVM による手書き筆記の分類を行い削除コマンドと被削除ストロークを識別する手法とアプローチは類似している。しかし、IRONOFF や UNIPEN データセットは、枠のなかに 1 文字ずつ書かれた手書き文字データセットであり、文字単位での認識にとどまっている。我々は数式や図形に加えて、文字単位の識別がされていない文章を含むような手書き筆記データについて、SVM による分類を試みている点が異なる。

## 6. まとめと今後の課題

本稿では、自由形式の手書きのノートから削除コマンドと被削除ストロークを分類するための特徴量を提案した。我々は、学習者のノートと特徴量の有効性をテストした。削除コマンドの検出のために、我々は、カーネルパラメータが適切に選択された場合で、学習に用いる通常筆記と被削除ストロークのバリエーションが十分である場合は削除コマンドの分類性能が高まることを確認した。また、被削除ストロークの抽出については、削除コマンドが正しく認識されていれば、被削除ストロークの 84% が正しく分類されることを確認した。

今後の課題として、フェーズ 1 のための特徴量を再度検討する必要があると考えている。その理由として、削除コマンドの検出率は、被削除ストロークの検出率に直接影響を与えるからである。また、提案手法を用いた、教師に提示するシステムや、学習者筆記をプロジェクタ投影する際に配慮する機能を追加し、その効果について検証を行っていく。

## 参考文献

- [1] 今井順一, 山本大輔, 小松川浩. デジタルペンを活用したリメディア教育での授業デザイン. *メディア教育研究*, Vol. 5, No. 1, pp. 57–66, 2008.
- [2] Nagai Masahiro and Kitazawa Takeshi. Understanding the Features of Digital Pen Use in Initial Introductory Lessons. In *Proceedings of the 18th International Conference on Computers in Education. Putrajaya, Malaysia: Asia-Pacific Society for Computers in Education. Retrieved August, 2013.*
- [3] Motoki Miura, Taro Sugihara, and Susumu Kunifuji. Improvement of Digital Pen Learning System for Daily Use in Classrooms. *Educational Technology Research*, Vol. 34, pp. 49–57, October 2011.
- [4] Urs Ramer. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, Vol. 1, No. 3, pp. 244–256, 1972.
- [5] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, Vol. 2, No. 3, p. 27, 2011.
- [6] Abdul Rahim Ahmad, M Khalia, Christian Viard-Gaudin, and Emilie Poisson. Online Handwriting Recognition using Support Vector Machine. In *TENCON 2004. 2004 IEEE Region 10 Conference*, pp. 311–314. IEEE, 2004.
- [7] Claus Bahlmann, Bernard Haasdonk, and Hans Burkhardt. Online handwriting recognition with support vector machines—a kernel approach. In *Frontiers in handwriting recognition, 2002. proceedings. eighth international workshop on*, pp. 49–54. IEEE, 2002.
- [8] B.Q. Huang, C.J. Du, Y.B. Zhang, and M.-T. Kechadi. A Hybrid HMM-SVM Method for Online Handwriting Symbol Recognition. In *Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on*, Vol. 1, pp. 887–891, Oct 2006.
- [9] B.Q. Huang and M.-T. Kechadi. A Fast Feature Selection Model for Online Handwriting Symbol Recognition. In *Machine Learning and Applications, 2006. ICMLA '06. 5th International Conference on*, pp. 251–257, Dec 2006.

**謝辞** 本研究の一部は JSPS 科研費 (課題番号 15K00485) および公益財団法人電気通信普及財団の支援によるものです。

## 付 録

表 A.1 各シートでのストローク数

sheet	normal	erase	deleted	sheet	normal	erase	deleted	sheet	normal	erase	deleted
109	533	11	49	209	399	4	12	242	581	0	0
111	518	10	374	211	584	16	101	243	492	2	3
113	630	10	64	213	418	18	37	318	435	0	0
118	604	6	43	214	496	17	13	319	433	10	5
119	384	16	13	218	894	18	45	322	250	2	17
122	435	15	133	219	630	0	0	323	320	0	0
123	532	8	25	222	716	3	6	326	274	3	3
126	433	0	0	223	563	0	0	327	306	0	0
127	624	11	39	226	540	4	15	328	329	0	0
128	555	2	34	228	824	9	150	329	264	2	4
129	358	22	196	229	467	4	14	330	355	0	0
130	388	0	0	230	323	0	0	331	483	5	2
131	515	16	307	231	501	14	79	332	163	2	3
132	411	10	69	232	137	0	0	334	368	2	23
134	381	1	2	234	627	9	148	335	171	0	0
135	348	2	62	235	579	17	65	336	145	4	29
136	901	1	2	236	951	6	43	338	225	0	0
138	330	20	207	238	791	3	38	339	39	4	248
142	365	4	16	239	523	0	0	340	293	2	20
143	386	5	62	240	482	2	5	343	92	3	77

表 A.2 フェーズ 1 の検証に利用した基本的な統計 (全シートのうち約 3 分の 1 を掲載)

sheet	cls	num	width		height		erase		cross		ramer	
			Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD
109	0	582	12.7	11.6	9.7	7.0	0.04	0.19	0.26	0.62	1.12	1.50
109	1	11	138.0	98.3	11.7	6.6	0.73	0.45	5.09	3.06	0.73	1.54
111	0	892	8.5	15.2	6.9	5.7	0.02	0.15	0.24	0.69	0.85	1.12
111	1	10	72.9	55.9	107.6	96.3	0.60	0.49	8.20	4.79	1.70	2.19
113	0	694	6.4	9.2	6.3	4.8	0.04	0.19	0.17	0.59	0.90	1.18
113	1	10	34.5	27.7	12.7	9.0	0.40	0.49	10.00	9.38	3.70	3.26
118	0	647	9.9	16.2	8.5	6.9	0.04	0.19	0.16	0.48	1.00	1.35
118	1	6	132.5	59.4	6.8	1.7	1.00	0.00	6.83	2.03	0.00	0.00
119	0	397	10.1	7.6	10.9	8.6	0.05	0.21	0.18	0.54	0.81	1.13
119	1	16	29.5	5.5	23.6	2.3	0.56	0.50	2.94	1.52	0.13	0.48
122	0	568	10.5	22.1	10.4	15.2	0.04	0.19	0.24	0.69	0.88	1.22
122	1	15	63.4	62.5	36.5	25.9	1.00	0.00	3.40	2.58	0.33	1.25
123	0	557	8.6	11.1	7.5	5.8	0.05	0.21	0.21	0.61	0.88	1.15
123	1	8	18.4	7.6	13.0	6.5	0.38	0.48	4.13	1.83	0.50	1.00
126	0	433	9.4	16.0	9.8	7.7	0.03	0.18	0.17	0.53	1.03	1.20
127	0	663	9.2	12.9	9.1	7.7	0.03	0.18	0.31	0.65	0.89	1.26
127	1	11	30.3	32.1	31.4	22.3	0.00	0.00	47.73	69.74	11.73	12.00
128	0	589	9.2	13.6	7.2	5.1	0.04	0.20	0.14	0.45	0.88	1.17
128	1	2	202.8	14.5	10.9	1.9	1.00	0.00	9.50	5.50	0.00	0.00
129	0	554	9.6	19.7	9.1	6.7	0.04	0.20	0.17	0.61	0.98	1.20
129	1	22	33.5	54.9	43.3	67.8	0.23	0.42	9.95	7.34	2.05	1.94
130	0	388	7.5	14.2	7.2	5.6	0.03	0.18	0.19	0.53	0.83	1.09
131	0	822	10.1	11.8	11.7	38.0	0.03	0.17	0.18	0.54	1.03	1.38
131	1	16	77.2	56.3	54.6	73.7	0.50	0.50	8.81	8.22	1.25	1.92
132	0	480	10.1	18.0	8.4	8.7	0.06	0.23	0.17	0.73	0.89	1.32
132	1	10	70.9	75.5	19.5	15.0	0.40	0.49	9.20	4.33	3.10	3.27
134	0	383	10.9	19.3	10.7	7.6	0.03	0.17	0.19	0.72	0.78	1.05
134	1	1	7.8	0.0	5.3	0.0	0.00	0.00	4.00	0.00	0.00	0.00
135	0	410	10.0	15.9	8.5	7.7	0.04	0.20	0.13	0.58	0.77	1.12
135	1	2	237.2	68.2	5.5	1.4	1.00	0.00	13.50	10.50	0.00	0.00
136	0	903	10.1	24.3	10.6	28.7	0.04	0.20	0.31	1.36	1.11	1.26
136	1	1	9.7	0.0	7.8	0.0	0.00	0.00	17.00	0.00	11.00	0.00
138	0	537	9.7	16.0	8.6	8.5	0.05	0.21	0.24	1.03	0.74	1.12
138	1	20	106.2	76.8	19.6	23.8	0.50	0.50	4.05	2.52	0.00	0.00
142	0	381	8.2	6.4	9.1	7.3	0.05	0.22	0.11	0.36	0.86	1.25
142	1	4	29.1	6.6	14.0	1.5	0.75	0.43	3.00	1.87	0.00	0.00
143	0	448	8.1	4.2	8.0	6.5	0.03	0.18	0.23	1.02	1.26	1.48
143	1	5	86.9	81.6	26.8	47.7	0.60	0.49	5.00	3.29	1.60	3.20

表 A.3 フェーズ 2 の検証に利用した基本的な統計 (全シートのうち約 3 分の 1 を掲載)

sheet	cls	num	Avg(oc)	SD(oc)	Avg(otd)	SD(otd)	Avg(crossa)	SD(crossa)
109	0	533	0.00	0.00	0.00	0.00	0.00	0.00
109	2	49	0.53	0.86	12.02	18.98	1.14	1.51
111	0	518	0.00	0.06	0.95	15.23	0.00	0.00
111	2	374	1.17	0.69	321.60	192.17	0.21	0.75
113	0	630	0.00	0.06	0.27	4.77	0.00	0.09
113	2	64	0.38	0.52	26.22	38.74	1.48	2.02
118	0	604	0.00	0.00	0.00	0.00	0.00	0.00
118	2	43	0.16	0.37	2.37	6.63	0.95	1.35
119	0	384	0.00	0.00	0.00	0.00	0.00	0.00
119	2	13	1.54	1.01	14.00	8.43	3.08	4.16
122	0	435	0.00	0.10	0.06	1.20	0.00	0.00
122	2	133	0.86	0.81	21.08	21.13	0.38	0.76
123	0	532	0.00	0.00	0.00	0.00	0.00	0.00
123	2	25	0.36	0.48	6.84	11.50	1.32	1.57
127	0	624	0.00	0.00	0.00	0.00	0.00	0.00
127	2	39	0.97	0.80	7.49	7.08	6.69	6.09
128	0	555	0.00	0.00	0.00	0.00	0.00	0.00
128	2	34	0.50	0.61	76.09	85.86	0.56	0.69
129	0	358	0.01	0.12	1.14	15.25	0.00	0.00
129	2	196	1.40	0.96	99.95	75.31	1.07	2.48
131	0	515	0.01	0.09	0.36	4.55	0.00	0.09
131	2	307	0.70	0.62	34.52	38.46	0.42	1.06
132	0	411	0.00	0.05	0.01	0.20	0.00	0.05
132	2	69	0.78	0.41	21.19	17.98	1.30	2.37
134	0	381	0.00	0.00	0.00	0.00	0.00	0.00
134	2	2	0.00	0.00	0.00	0.00	2.00	2.00
135	0	348	0.00	0.00	0.00	0.00	0.00	0.00
135	2	62	0.10	0.30	2.79	9.28	0.44	0.66
136	0	901	0.00	0.00	0.00	0.00	0.00	0.00
136	2	2	0.00	0.00	0.00	0.00	8.50	0.50
138	0	330	0.00	0.00	0.00	0.00	0.00	0.00
138	2	207	0.59	0.81	13.43	20.77	0.38	0.66
142	0	365	0.00	0.00	0.00	0.00	0.00	0.05
142	2	16	0.69	0.68	11.44	12.32	0.69	0.92
143	0	386	0.00	0.00	0.00	0.00	0.00	0.00
143	2	62	0.71	0.45	39.95	27.82	0.40	1.30

表 A.4 フェーズ 1 のグリッド検索結果

通常筆記+被削除	Avg(F)	SD(F)	Avg(logC)	SD(logC)	Avg(log $\gamma$ )	SD(log $\gamma$ )
100	0.428	0.198	10.80	4.00	-13.1	1.73
200	0.483	0.207	10.20	3.76	-13.0	1.73
300	0.537	0.219	10.20	3.89	-13.4	1.78
400	0.583	0.215	8.00	4.88	-12.6	2.11
500	0.584	0.223	7.10	4.65	-11.9	2.93
600	0.646	0.202	7.00	6.62	-12.2	3.49
700	0.655	0.216	6.65	5.67	-12.4	2.91
800	0.666	0.229	6.55	5.28	-12.6	2.34
900	0.671	0.233	6.45	5.16	-12.2	3.89
1000	0.670	0.227	7.70	5.51	-13.1	1.87
1200	0.705	0.213	6.50	5.53	-12.2	2.36
1500	0.713	0.205	6.55	5.50	-12.7	2.59
1800	0.742	0.212	5.50	5.86	-11.6	3.32
2100	0.757	0.193	5.55	5.97	-11.7	3.05
2400	0.779	0.196	6.25	5.89	-11.8	2.99
2700	0.786	0.168	5.31	5.07	-12.0	3.03
3000	0.807	0.181	5.25	5.46	-11.4	2.89

表 A.5 フェーズ 2 のグリッド検索結果

通常筆記+被削除	Avg(F)	SD(F)	Avg(logC)	SD(logC)	Avg(log $\gamma$ )	SD(log $\gamma$ )
200	0.840	0.127	-5.00	0.00	-2.38	2.26
300	0.840	0.127	-5.00	0.00	-2.42	2.33
400	0.842	0.126	-4.78	1.47	-2.64	2.81
500	0.840	0.127	-4.87	0.88	-2.64	2.57
600	0.840	0.127	-5.00	0.00	-2.56	2.45
700	0.840	0.127	-4.60	2.65	-2.82	3.06
800	0.840	0.127	-4.87	0.88	-2.73	2.69
900	0.840	0.127	-5.00	0.00	-2.73	2.72
1000	0.840	0.127	-4.64	2.36	-3.13	3.02
1200	0.840	0.127	-4.64	2.36	-3.22	3.19
1500	0.840	0.127	-4.60	2.65	-3.36	3.38