

オンライン手書きデータによる記憶度推定システムの構築

苑田 翔吾^{†1} 浅井 洋樹^{†2} 山名 早人^{†3†4}

^{†1} 早稲田大学基幹理工学部 〒169-8555 東京都新宿区大久保 3-4-1

^{†2} 早稲田大学大学院基幹理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1

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

^{†4} 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: {s_sonoda,asai,yamana}@yama.info.waseda.ac.jp

あらまし 本研究では、記憶支援アプリケーションにおける記憶度をオンライン手書きデータから推定する手法を提案し、学習者の記憶を支援するシステムの構築を行う。オンライン手書きデータとは、ニンテンドーDS や iPad に代表される手書き入力可能な端末で入力された筆記時系列データのことを指す。現在、多くの記憶支援アプリケーションにおいて、学習者が記憶項目について、記憶できたか否かのテストを行うことにより、記憶度を推定している。しかし、従来の記憶テストでは、学習者の解答が正解か不正解かの 2 値の情報のみを用いて記憶度を判定しているため、段階別の記憶度評価を行うことができない。段階別の記憶度評価は、学習すべき問題を推薦するための指標などに応用可能である。そこで、本研究では記憶テストの際、オンライン手書きデータを収集し、オンライン手書きデータと記憶度の関連性について調査した。その結果、オンライン手書きデータの特徴量として、「ストロークの時間間隔」や「消しゴムモードへ遷移していた時間」などの特徴量を選択することにより、学習者が迷いながら正解した問題を精度約 85%、再現率約 70%で検出可能であることが示された。

キーワード e-ラーニング, オンライン手書きデータ

1. はじめに

物事を記憶するという行為は、子供から大人まで、誰しもの必要に迫られ経験する行為である。従って、記憶にかける時間をいかに短くできるかという課題は、限られた時間を生きる私たちにとって重要な課題の一つである。近年では、画面上に専用のペンを用いて文字を書くことができる端末が普及したため、この入力装置を有効利用した記憶支援アプリケーション¹が登場している。また、小・中・高等学校などの教育機関で手書き入力可能な手書き入力端末と記憶支援アプリケーションが採用され始め、生徒の意欲・関心の向上やテスト結果の向上など、数々の実績を上げている[1]。本稿では、教育機関への手書き入力端末普及の流れを汲み、手書き入力端末に入力された手書きデータを用いて記憶度の推定が行えるかどうかを検証する。

近年、記憶に関する研究結果を参考にしたニンテンドーDSやiPad向けの記憶支援アプリケーションが市場^{2,3}に売り出されている。これらのアプリケーションには、名称は様々だが記憶の度合いを表す指標（本稿

では記憶度と呼ぶ）が備わっており、ユーザの正答率や誤答率からどの程度記憶したかを計測し記録している。ユーザは学習度から自身の記憶度合を確認し、今後の学習計画に役立てることができる。また、記憶度からユーザが記憶していないと推測される問題を推薦するシステムを持つアプリケーションも登場している[2]。

記憶支援アプリケーションによる記憶度の計測方法はユーザテスト、アンケートの二種類に大別される。ユーザテストでは、記憶したい事象の対応付けに関して、対応前を問題、対応後を解答として分け、ユーザが問題から解答を推測し、正しい解答を入力することができるかをテストする。ユーザテストはユーザに対して解答の入力を要求することにより、ユーザの記憶を客観的に計測することができる。しかし、解答可能か不可能かの 2 値の分類となるため、偶然思い出すことができ正解した事象に関しては、次回同様の問題が出題された時、正しく解答できない可能性が高いが、偶然思い出したかどうかを判定できない。従って、完全に覚えている事象と偶然思い出せた事象に関しての評価が同一であることが問題点として挙げられる。

一方、アンケートでは、ユーザが記憶したい事象についてどの程度記憶できているかを、段階別の選択肢を用意し選ばせることによって、段階別に記憶度を評価することができる。アンケートは主観的な方法であり、必ずしも正確な記憶度が計測できるわけではない

¹ えいご漬け, “<http://www.nintendo.co.jp/ds/angj/>,” 2012.01.05 アクセス。

² DS at Nintendo,

“<http://www.nintendo.com/ds/>,” 2012.01.08 アクセス。

³ Apple - iPad 2 - See apps and games from the App Store, “<http://www.apple.com/ipad/from-the-app-store/>,” 2012.01.08 アクセス。

が、自己申告の記憶度と実際の記憶度には強い相関があることが知られている[3]。しかし、アンケートは自己評価の記憶度を別途入力する必要があるため、ユーザの負担が増大する。

このように、記憶の計測方法は一長一短であり、ユーザの記憶度を計測する有効な手段は確立されていない。記憶度は既存のアプリケーションに組み込まれている通り、学習度の指標や問題を推薦する指標に適用することができる他、教師へのフィードバックや問題集作成者へのフィードバックなどにも応用できる重要な指標である。そこで、本稿では、新しい記憶度の計測手法として、オンライン手書きデータを用いた記憶度の計測手法を提案する。

オンライン手書きデータとは、手書き入力端末で得られた時系列情報を含む筆跡データのことを指す。一方、画像に代表される時系列情報を含まない筆跡データをオフライン手書きデータと呼ぶことで、本稿では両者を呼び分けている。オンライン手書きデータは、ペンの座標や筆圧、筆記速度などの詳細な筆跡データが含まれるため、オフライン手書きデータと比較してより多くの情報を持つ。近年では手書き入力端末の登場、普及により、専用ペンと手書き入力端末を用いて入力されるオンライン手書きデータが身近なものとなりつつある。

本稿ではユーザテストの問題点である完全に記憶している事象と偶然思い出せた事象について、オンライン手書きデータから両者を区別することを目指す。そのために、まずは記憶度を反映しているオンライン手書き特徴量を調査する。その後、定めたオンライン手書き特徴量から完全に記憶している事象と偶然思い出せた事象を判別するシステム構築を目指す。本提案システムによって、従来の記憶テストの手軽さでアンケートのような段階別評価を行うことが可能となる。

本稿では次の構成をとる。まず2節で関連研究について述べ、3節で記憶度とオンライン手書き特徴量の関係性について調査する。次に4節で記憶度判別の提案手法と検証実験を述べた後、最後に5節でまとめる。

2. 関連研究

オンライン手書きデータから記憶度を推定する研究は我々の知る限り存在しない。本節では既存の記憶支援アプリケーションの記憶度測定に関する考察やオンライン手書きデータの特徴量選択について述べる。

2.1. 記憶支援アプリケーションでの記憶度推定法

本節では、寺澤[4]らのマイクロステップ計測技術[5]による記憶の習得度の計測に関する研究を例に挙げ、既存の記憶支援アプリケーションの記憶度測定法について述べる。

マイクロステップ計測技術とは、学習スケジュールや学習者からシステムへのフィードバックを基に、測定困難であった微細な学習効果の積み重ねを測定する技術である。テスト兼学習法として、自己評定と客観テストを実装しており、これら2つの学習法から学習の到達度（記憶度）を決定している。2011年岡山大学大学院の山本[3]らはマイクロステップ計測技術を実装した任天堂DS用英単語学習ソフトを利用して、顕在記憶と潜在記憶の違いを学習の到達度から調査した。心理学では、その情報を獲得したときのこと（文脈）を意識的に思い出さなくても利用可能な記憶のことを、潜在記憶と呼ぶ。反対に、覚えた時の文脈を意識的に思い出さなくては使えない記憶のことを顕在記憶と呼ぶ。山本らは研究の結果、顕在記憶に起因する学習の効果と潜在記憶に起因する学習の効果の蓄積パターンが異なるという結論を得た。また、自己評定と客観テストの結果が強い正の相関を示し、自己評定値が客観的な記憶度を測定する指標として、妥当性が高いことが述べられている。

2.1.1 節、2.1.2 節では、山本らの実験で用いられた「やればできる! THE マイクロステップ技術で覚える英単語[7]」の記憶の計測法である自己評定と客観テストについて述べる。記憶の対応付け元の語句を問題、記憶の対応付け後の語句を解答と定義し、今後これらの語を用いる。

2.1.1. 自己評定

自己評定とは、記憶の対応付けに関して段階別な自己評価を行うことである。「やればできる! THE マイクロステップ技術で覚える英単語[7]」における、自己評定の手順を図1に示す。また、この時のニンテンドーDSの画面を図2に示す。

Step1. 問題が表示され、学習者は問題から解答を推測する。
Step2. 学習者が解答表示ボタンを押すと、画面が遷移する。
Step3. 問題と解答が表示され、学習者が自己評価を入力する。

図1 自己評定の手順

英単語	broad				
broad	bro:d 1: 広い 2: 広範囲に及ぶ 3: 大まかな				
意味表示	<table border="1"><tr><td>3 良い</td><td>2 もう少し</td></tr><tr><td>1 だめ</td><td>0 全くだめ</td></tr></table>	3 良い	2 もう少し	1 だめ	0 全くだめ
3 良い	2 もう少し				
1 だめ	0 全くだめ				

図2 自己評定時のニンテンドーDSの画面
([7]を基に作成)

2.1.2. 客観テスト

客観テストとは、出題された問題に解答することである。「やればできる！THE マイクロステップ技術で覚える英単語[7]」における、客観テストの手順を図3に示す。また、この時のニンテンドーDSの画面を図4に示す。

- Step1. 問題の一覧が表示され、学習者は回答する問題を選択する。
Step2. 問題を選択すると、問題に対する解答の候補が複数表示される。
Step3. 解答の候補から学習者自身が記憶している解答を選択する。記憶していない場合は、「分からない」ボタンを押す。

図3 客観テストの手順

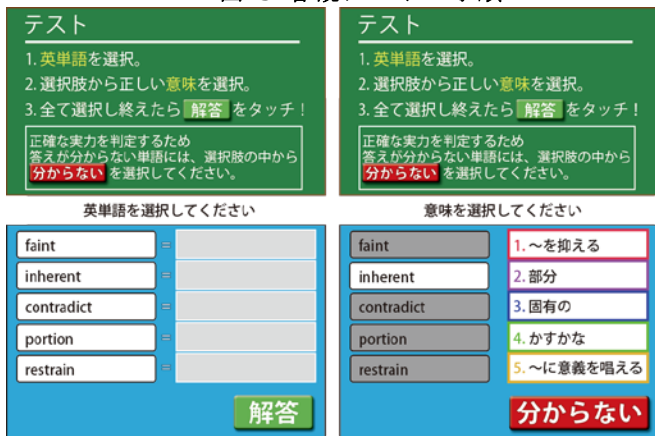


図4 客観テスト時のニンテンドーDSの画面
([7]を基に作成)

2.2. オンライン手書きデータの特徴量

様々な目的を実現するために、オンライン手書きデータに関する研究が既になされてきた。本節では、研究の目的とそれを実現するためのオンライン手書きデータの特徴量について述べる。

2.2.1. 認知的負荷とオンライン手書きデータの関係

Kun Yu ら[8]は認知的負荷とオンライン手書きデータの関係性について研究を行った。認知的負荷とは、タスクを達成するために脳から要求された精神的な資源の度合いを表現したものである。脳には一定限度量の注意力しかなく、その限度を超えてしまうと情報の処理能力が低下してしまう。従って、過剰なタスクは書き手に大きな認知的負担をかけ、手書き文字に影響を及ぼす。Kun Yu らは認知的負荷を意図的にかけ、その時に収集したオンライン手書きデータを調査した。Kun Yu らが収集したオンライン手書きデータ特徴量は以下の通りである。尚、ストロークとは、筆記者が書いた線、あるいは線を書くことを指す。

- ストロークの最高筆圧
- ストロークの最低筆圧

- ストロークの平均筆圧
- ストロークの最高速度
- ストロークの最低速度
- ストロークの平均速度
- ストロークの長さ

これらの7つの特徴量と認知的負荷の関係性を調査した結果、認知的負荷が「ストロークの最高筆圧」「ストロークの最低速度」に大きく影響を及ぼすことがわかった。

Kun Yu らの研究結果から、本稿も同様にストロークの筆圧、ストロークの速度が重要なオンライン手書きデータ特徴量となると考え、調査対象とした。

2.2.2. 書写技能に基づく筆跡に現れる個人性の抽出

中村ら[9]はオンライン手書きデータから個人性を抽出する研究を行った。

中村らは、筆跡に個人性が現れる主な要因は書字学習の過程で生じた書字習慣の差であると考え、書写技能からオンライン手書き特徴量の選出を行った。書写技能の分類を図5に示す。

また、中村らは図5の用筆法に関して、始筆、送筆、終筆の定義が明確ではないことから、ストローク全体の始めの20%を始筆部、終わりの20%を終筆部、残りの中間を送筆部と定義した。これを図6に示す。



図5 書写技能の分類 ([9] を基に作成)

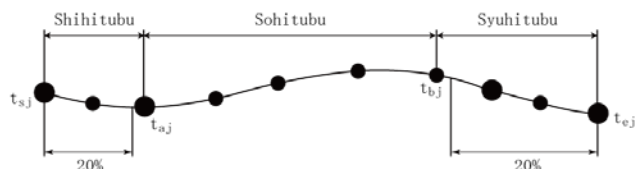


図6 ストロークの分解 ([9]を基に作成)

中村らは解析対象として、「チキユウ」という文字を筆記させることを1セットとし、4人の筆記者からそれぞれ50セット収集したデータセットを用いた。こ

れにより、以下に示す結果が得られた。

- 各々の特徴量間に個人間の差があり、何らかの個人性を表していること。
- 特性によって個人内でのばらつきに差があること、また、ばらつきが個人間で差がある特性があること。
- 筆者に依存せず相関が強い特性値の組み合わせと、筆者に依存して相関の強い組み合わせがあること。

中村らの研究結果から、オンライン手書きデータ特徴量は個人間で差が生じるということがわかる。このため、本稿においても、個人差を考慮した上でシステムを構築する。

3. 記憶度とオンライン手書きデータの関連性調査

記憶度を推定するシステムを構築するために、学習者の記憶度に関連するオンライン手書きデータの特徴量を調べる実験を行った。学習者が、記憶度が低いと判断した場合と記憶度が高いと判断した場合の両者のオンライン手書きデータの間で検定を行い、有意な差が表れるかどうかを検証する。

3.1. 実験環境

オンライン手書きデータの収集には、ワコム社の Cintiq 12WX[10]を使用した。CINTIQ 12WX は、最大解像度 1280×800 ドットで表示、筆圧 1,024 段階で測定可能である。実験に用いるアプリケーションは Microsoft 社の Windows Presentation Foundation で開発を行った。また、実験は著者の研究室で行い、情報理工学を専攻する著者の研究室メンバー6 名にテストを依頼した。

3.2. 実験方法

実験においては、漢字の書き取り問題を出題し、被験者に解答をタブレット端末上にかかせる。その後、出題した漢字に対する記憶度を4つの選択肢の中から選択させる。このときの被験者が選択した記憶度とタブレット端末上に書いたオンライン手書きデータ特徴量の関係性を調査する。

実験のフローチャートを図7に示す。また、起動時のアプリケーションのスクリーンショットを図8に示す。問題に対する解答が表示され、被験者が記憶度を選択している時のアプリケーションのスクリーンショットを図9に示した。

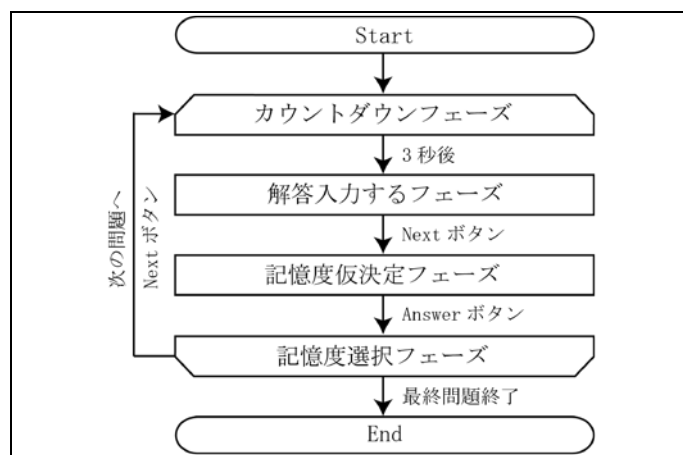


図 7 実験のフローチャート

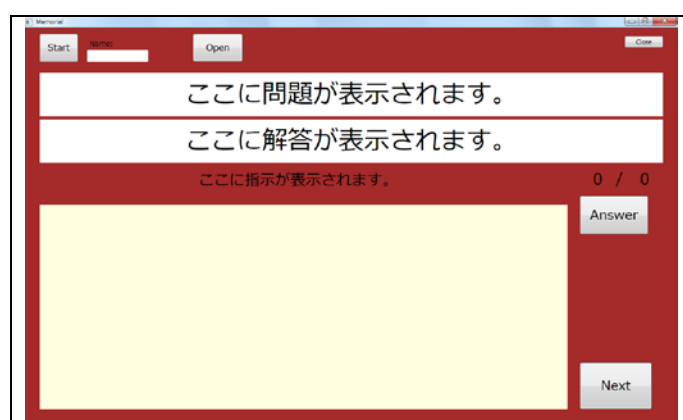


図 8 アプリケーション起動時

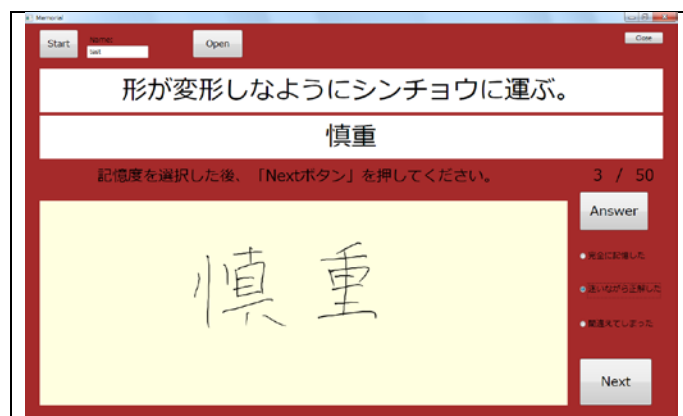


図 9 記憶度選択時

図7で示した通り、実験用アプリケーションは以下の4つのフェーズを持つ。

1. カウントダウンフェーズ
2. 解答入力フェーズ
3. 記憶度仮決定フェーズ
4. 記憶度選択フェーズ

これらのフェーズについて順に説明していく。

(1) カウントダウンフェーズ

カウントダウンフェーズは問題が出題される前に3秒間のカウントダウンを行って、プレイヤーに問題を解答する準備時間を与えるフェーズである。問題が出題

される 3 秒前からカウントダウンが始まり、0 秒になると、問題が表示される。カウントダウンにより、前フェーズでボタンを押した直後に問題が表示されることがなくなるため、各問いに対して公平性を高めることができる。

(2) 解答入力フェーズ

解答入力フェーズは、問題が画面に表示され、被験者が問題に対する解答を入力するフェーズである。解答を入力し終えたら、画面右下の Next ボタンを押すことによって、次のフェーズへ遷移することができる。問題に対する解答が何もわからなかった場合、何も記入せずに Next ボタンを押して、次のフェーズへ遷移するように求めた。また、解答を入力するための時間制限は設定せず、ストロークを消しゴムモードで消すことによって、何度も書き直すことができる。

(3) 記憶度仮決定フェーズ

記憶度仮決定フェーズは被験者が記憶度を考えるために用意されたフェーズである。解答が表示された状態で記憶度を選択・決定してしまうと、自身の記憶度を高く評価してしまう傾向がある。これを防止するために、記憶度仮決定フェーズを設け、解答が表示される前に一度自身の記憶を考えるフェーズを設けた。

(4) 記憶度選択フェーズ

記憶度選択フェーズは問題と解答が表示された状態で記憶度を選択・決定するフェーズである。画面右のラジオボタンから合致する記憶度合いを一つ選択し、Next ボタンを押すことで決定され、次のフェーズへ遷移する。

被験者は以下の選択肢から問題に対する記憶度に合致するものを選択する。

- 正解：完全に記憶している
- 正解：迷いながら正解した
- 不正解：間違えてしまった

選択肢は問題に対する自身の解答が正解であったかあるいは不正解であったかによって 2 つに大別される。

本実験では、財団法人日本漢字能力検定協会が実施する漢字検定の書き取り問題を出題した。筆者が「被験者である大学生が迷いながらも正解できる難易度である」と想定した漢字が出題されている漢字検定準 2 級の漢字書き取り問題を選択した。漢字検定準 2 級の過去問題から計 50 問を被験者に解答させた。

3.3. オンライン手書きデータ特徴量

本節では、実験で考察するオンライン手書きデータの特徴量について述べる。3.3.1 節では入力デバイスから直接得られるオンライン手書きデータを述べ、3.3.2 節では得られた情報から算出可能な特徴量について述

べる。

3.3.1. デバイスから得られるオンライン手書きデータ

ワコム社の CINTIQ 12WX で得られるオンライン手書きデータの中で、研究に用いたものは、

- X 座標
- Y 座標
- 筆圧
- ペンの入力状態
- ペンのモード
- 経過時間

以上の 6 つである。

X 座標、Y 座標、筆圧はその時点での各々ペンの場所、筆圧を表す。筆圧は 1024 段階で取得可能であり、0 から 1 までの小数値で表される。

ペンの入力状態は、ペンがタブレットに接しているか否かを指す。ワコム社の CINTIQ 12WX では、ペンがタブレットに触れていない状態でも、ある程度ペンとタブレットが近ければ、各種パラメータが取得可能となる。ただし、このときの筆圧は 0 となる。

ペンのモードは、ストロークモードと消しゴムモードの 2 種類の状態を指す。ストロークモードはペンの先でタブレットに入力した場合であり、キャンバスに線を書くことができる。消しゴムモードはペンの裏側でタブレットに入力した場合であり、キャンバスに書かれているストロークを消すことができる。ペンの向きを変えるだけでモードを変更することができるため、ボタンを押さずに 2 つのモードを行き来可能である。経過時間は問題が表示されてから経過した時間を意味する。

3.3.2. 計算によって求められるオンライン手書き特徴量

3.3.1 節で述べたオンライン手書きデータから計算によって求められる特徴量は無数に存在するが、本稿では 2.2.1 節の関連研究と筆者の推測を基に、以下のオンライン手書き特徴量を設定した。

- ストロークの時間に関する特徴量
 - ・ 書き始めるまでの時間 (IniStrT)
 - ・ ストロークの時間間隔 (StrIntT)
- ストロークの筆圧に関する特徴量
 - ・ ストロークの最高筆圧 (MaxPF)
 - ・ ストロークの最低筆圧 (MinPF)
 - ・ ストロークの平均筆圧 (AvePF)
 - ・ ストロークの筆圧の標準偏差 (SDPF)
- ストロークの移動平均速度に関する特徴量
 - ・ ストロークの最高移動平均速度 (MaxV)
 - ・ ストロークの最低移動平均速度 (MinV)
 - ・ ストロークの平均移動平均速度 (AveV)

- ストロークの移動平均速度の標準偏差 (SDV)
- 消しゴムモードに関する特徴量
 - 消去したストロークの本数 (EraseT)
 - 消しゴムモードに遷移した回数 (CntEraseStr)
 - 消しゴムモードに遷移していた時間 (CntEraseMode)

各特徴量の省略語括弧内で示した．移行の表や図では，各特徴量を省略語で示す．

ストロークの時間間隔は，ストロークとストロークの間の時間の総和であり，最初のストローク開始から最後のストロークの終了までの総時間から，ストローク時間の総和時間を引いたものである．

ストロークの移動平均とは，ストローク中の連続する5つの標本座標点に関する平均速度を求めたものである．すなわち，ストロークの移動平均は，ストローク中の連続する5つの標本座標点における距離の総和と，ストローク中の連続する5つの標本座標点における時間の差の商で定義される．移動平均の定義で扱う連続する標本座標点の数を変えながら移動平均を求めた結果，ストローク中の標本座標点の数や算出された移動平均の値から，移動平均の定義で扱う連続する標本座標点は5つが最適であると考え，定めた．ただし，ストローク中の標本座標点が5つに満たない場合は，2つから4つまでのいずれかで移動速度を求めた．移動平均の式を式(1)に示した．

ストローク i が N 個の標本座標点 $(Xi1, Yi1), (Xi2, Yi2), \dots, (Xii, Yii)$ を持つとき，ストローク i の k 番目の移動平均 $MovAveV_{ik}$ を式(1)で表せる．

$$MovAveV_{ik} = \frac{\sum_{l=k}^{k+4} \sqrt{(X_{i(l+1)} - X_{il})^2 + (Y_{i(l+1)} - Y_{il})^2}}{T_{i(k+4)} - T_{ik}} \quad (1)$$

ただし， T_{ik} はストローク i の k 番目の標本座標点が得られた時のタイムスタンプとする．

また， $MovAveV_{ik}$ から求められるストローク i の平均移動平均速度 $AveV_i$ は，ストローク i で定義された移動平均の数を M とすると，式(2)で表せる．

$$AveV_i = \frac{\sum_{k=1}^M MovAveV_{ik}}{M} \quad (2)$$

3.4. 実験結果

3.4.1. 解析するデータ対象

被験者が各選択肢を選んだ回数を表1にまとめた．不正解の選択肢である「間違えてしまった」を選択した場合のオンライン手書きデータについては，本節での解析の対象としなかった．本稿の目的は，正解した問題に関しての記憶度を推定し，記憶度が低いと判断された問題を選出する点にある．従って，本稿の目的と逸れるため解析の対象から除外した．

表1 被験者が各選択肢を選んだ回数

被験者 選択肢	A	B	C	D	E	F	G
完全に記憶している(回)	14	6	13	9	6	7	7
迷いながら正解した(回)	3	3	1	3	2	5	4
間違えてしまった(回)	3	11	6	8	12	8	7

3.4.2. 検定による特徴量選択

「完全に記憶している」と「迷いながら正解した」の間のオンライン手書きデータの各特徴量の相違を確認するために，各特徴量においてマンホイットニーのU検定を行った．被験者毎に調査を行った理由は，2.2.2項で述べた通り，各個人によって筆跡が異なるからである．被験者毎に検定した検定の結果を表2に示す．このとき，有意水準0.05を下回った値に関して，下線を引いた．

表2 各特徴量におけるマンホイットニーのU検定結果

各特徴量	被験者	A	B	C	D	E	F
IniStrT		1.01	2.00	2.03	7.32	2.52	3.06
		E-02	E-02	E-01	E-02	E-02	E-03
StrIntT		2.20	3.40	3.02	1.60	4.76	2.16
		E-02	E-04	E-03	E-05	E-02	E-02
MaxPF		9.05	3.66	7.11	3.62	7.41	2.26
		E-01	E-02	E-01	E-01	E-01	E-01
AvePF		7.20	6.94	2.19	1.78	1.84	8.18
		E-01	E-01	E-01	E-01	E-01	E-01
SDPF		1.82	4.66	6.36	7.90	5.05	2.16
		E-01	E-05	E-02	E-04	E-01	E-02
MinV		5.49	1.51	1.25	1.25	3.03	3.47
		E-01	E-01	E-01	E-02	E-02	E-01
AveV		4.97	9.43	4.80	1.26	9.79	1.35
		E-01	E-04	E-01	E-01	E-01	E-02
SDV		9.68	5.87	6.70	1.00	2.95	4.42
		E-01	E-01	E-03	E+00	E-01	E-01
CntEraseMode		1.26	NA	1.43	2.37	1.95	6.45
		E-01	NA	E-04	E-07	E-02	E-03
CntEraseStr		1.81	NA	1.98	3.29	3.35	5.81
		E-01	NA	E-04	E-07	E-02	E-03
EraseT		1.27	NA	2.29	2.38	3.03	5.31
		E-01	NA	E-04	E-07	E-02	E-03

特徴量の定義上，あるいはデバイスの都合上，不適切な特徴量がいくつかあった．「ストロークの最低筆圧」は，たいていの場合において0に近い非常に小さい値となってしまう，特徴量として意味を成さない値となってしまったため除外した．

また，「ストロークの最高移動平均速度」は，デバイスの都合上，全く同様のタイムスタンプにも関わらず，座標が複数点存在する場合があります，この場合の移動平均速度が無限大となってしまう場合があったため除外した．

表2の結果と実際に得られたデータより，被験者に依存せず高い有意差を持つ特徴量を選出した．選出した特徴量は以下の通りである．

- 書き始めるまでの時間
- ストロークの時間間隔
- 消去したストロークの本数
- 消しゴムモードに遷移した回数

● 消しゴムモードに遷移していた時間
従って、これらの特徴量を「完全に記憶している」と「迷いながら正解した」の分類を行うことができる特徴量と考えた。

4. 記憶度判別の提案手法と検証実験

4.1. 提案手法

3.4.2 項で得られた各特徴量を基に、オンライン手書きデータから「完全に記憶している」と「迷いながら正解した」の分類を行う。本稿では、2 値分類問題の解決法として Support Vector Machine (SVM) を選択した。

SVM には、「完全に記憶している」と「迷いながら正解した」時の 3.4.2 項で得られた各特徴量を特徴ベクトルとして、学習器を作成しておく。作成された学習器を使って、同被験者のオンライン手書きデータから、「完全に記憶している」あるいは「迷いながら正解した」を判別する。

SVM の入力する特徴量には、オンライン手書きされた入力から 3.4.2 項で得られた各特徴量を算出した値を用いる。提案手法のフローチャートを図 10 に示す。

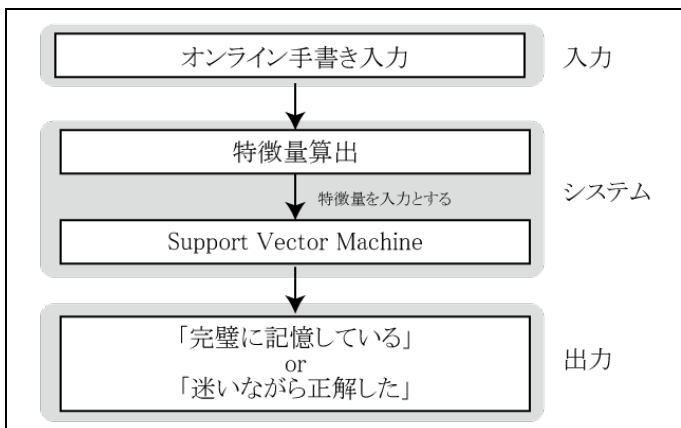


図 10 提案手法のフローチャート

4.2. 検証実験

提案手法である SVM による判定の精度を確認するために、3.4 節で得られた「完全に記憶している」「迷いながら正解した」と回答されたオンライン手書きデータを対象に、3.4.2 項で得られた 5 つの特徴量を使って交差検定を行った。SVM と交差検定には R[11] を用いた。4.2.1 節では実験方法、4.2.2 節では実験結果について述べる。

4.2.1. 実験方法

R の e1071 パッケージ[12]に含まれる svm 関数を用いて、提案手法の精度測定を行った。各被験者のデータの内、学習者が「完全に記憶している」「迷いながら正解した」と選択した問題を交差検定対象データとし

た。交差検定対象データから任意の問題を 1 問選択し、残りの交差検定対象データで学習を行う。学習後、先ほど選択した任意の問題 1 問に関して「完全に記憶している」「迷いながら正解した」の選択肢を推測する。この一連の操作を、各交差検定対象データにおいて 1 回ずつ選択されるまで、すなわち、交差検定対象データの数だけ繰り返す。このときに使用した SVM のパラメータである GAMMA と cost については、被験者毎にグリッドサーチによって算出した。グリッドサーチとは、書く派ラメータを少しずつ増減させ、網羅的に最適なパラメータを探索する手法である。本稿の場合は、gamma 値と cost 値を増減させ、交差検定の精度が最も高くなる値を採用した。

4.2.2. 実験結果

「迷いながら正解した」の SVM による精度と再現率を表 3 に示す。

表 3 「迷いながら正解した」における SVM の精度と再現率

被験者	正解		不正解		「迷いながら正解した」精度	「迷いながら正解した」再現率
	完全に記憶している	迷いながら正解した	完全に記憶している	迷いながら正解した		
A	8	7	3	1	88%	70%
B	16	12	2	1	92%	86%
C	21	4	5	1	80%	44%
D	34	6	3	0	100%	67%
E	29	0	4	0	0%	0%
F	22	6	3	3	67%	67%

表 3 より、被験者 E を除いて、「迷いながら正解した」において、精度約 85%、再現率約 70%で予測することができた。

被験者 E の特徴量からは、全問題において「完全に記憶している」と予測されている。これは特徴量が 5 つに対して「迷いながら正解した」問題数が 4 つしか存在しなかったことから、うまく学習が行えなかったために精度が低くなったと考えられる。

4.2.3. 記憶支援システムへの応用の考察

本提案手法を記憶支援ソフトへ応用する場合、既存のシステムでは問題に正解したことから「記憶している」と処理されていた事象に関して、本提案手法では「迷いながら正解した」と処理することにより、学習者へより詳細なフィードバックを行うことが可能となる。例えば、近いうちに再度問題を出題させることで再学習に要する時間を減らすことが可能である。また、既に記憶している事象に対して「記憶していない」と予測される不正解が少ない、すなわち「迷いながら正解した」における精度が高いことから、学習者が無駄な学習を行う可能性が低いことが言える。

本提案手法には、「完全に記憶している」あるいは「迷いながら正解した」のラベル付けされたオンライ

ン手書きデータが学習者ごとに必要である。従って、記憶支援システムを使い始めた直後は判定を行うことができないため、まずは記憶したい事象に関して「完全に記憶している」あるいは「迷いながら正解した」かのラベル付けを行う。その後、一定数のオンライン手書きデータを収集した後に、自動で判定を行っていく。そして、「迷いながら正解した」と判定された問題に関して出題の頻度を上げることで、記憶の支援を図ることが可能である。

4.3. 追加実験

4.2 節で使用したデータは約 30 問であり、本当に学習が行えたのか、また時間を置いて取得した新たなデータに対して十分な精度、再現率が出すことができるかどうか不安が残る。そこで、本節では、4.2 節で作成した学習器が未知のデータへ適応可能であるかどうかを検証する。

4.3.1. 実験方法

4.2 節で得られた 1 人の被験者のデータから学習器を作成し、新たに収集するオンライン手書きデータから被験者が選択した「完全に記憶している」と「迷いながら正解した」の選択肢を、作成した学習器で分類する。新たなオンライン手書きデータを得る為に、4.2 節における被験者 F に、100 問の漢字書き取りテストを解答させ、「完全に記憶している」と選択したデータ 48 問と、「迷いながら正解した」と選択したデータ 27 問を得た。そして、4.2 節と同様に解析を行った。

4.3.2. 実験結果

「迷いながら正解した」における SVM による精度と再現率を表 4 に示す。

表 4 「迷いながら正解した」における SVM の精度と再現率 2

予測	正解		不正解		「迷いながら正解した」精度	「迷いながら正解した」の再現率
	完全に記憶している	迷いながら正解した	完全に記憶している	迷いながら正解した		
F	44	20	7	4	83%	74%

表 4 より、約 85%の精度と約 70%の再現率を確認することができ、4.2 節で作成した学習器が学習を行っていることがわかる。従って、4.2 節で作成した学習器が学習を行えたこと、また時間を置いて取得した新たなデータへ適応可能であることが示された。

5. まとめ

本稿では、記憶支援アプリケーションテストにおける従来の「正解」を、「完全に記憶している」「迷いながら正解した」の 2 値に区別することにより、より詳細なフィードバックを行える可能性を示した。また、Support Vector Machine を用いることにより、「完全に記憶している」「迷いながら正解した」の 2 値の分類を、精度約 85%、再現率約 70%で行えることを示した。今

後の課題としては、4.3.24.2.2 項で定義した検出率を向上させるために、特徴量選択の見直し、また提案手法の改良が考えられる。また、漢字の書き取り問題以外への適用についても検証していきたい。

参 考 文 献

- [1] 文部科学省「教育の情報化に関する手引き」について,“http://www.mext.go.jp/a_menu/shotou/zyouho/1259413.htm”, 2011.12.4 アクセス.
- [2] 新技術を実装した NINTENDO DS 用英単語学習ソフトを使い,“麻布高校生対象に長期学習実験を実施”, http://www.okayama-u.ac.jp/up_load_files/soumu-pdf/press-090826-1.pdf, 2011.12.4 アクセス.
- [3] 山本浩司, 寺澤孝文, “任天堂 DS を用いた長期的な英単語学習の効果の検証”, 2011 年度日本認知科学会第 28 回大会, 2011.
- [4] Terasawa Takafumi official web site,<http://ed-www.okayama-u.ac.jp/~shinri/terasawa/research001.html>,2011.12.25アクセス.
- [5] 寺澤孝文, 太田信夫, 吉田哲也, “マイクロステップ計測法による英単語学習の個人差の測定”, p.275, 風間書房, 2007.
- [6] 藤田哲也, “絶対役立つ教養の心理学 一人生を有意義に過ごすために一”, p.40, ミネルヴァ書房, 2009.
- [7] やればできる! THE マイクロステップ技術で覚える英単語,“http://www.d3p.co.jp/s_ds/s_ds_019.html”, 2011.12.25アクセス.
- [8] Kun Yu, Julien Epps, Fang Chen, “Cognitive Load Evaluation of Handwriting Using Stroke-level Features”, IUI ‘11 Proceedings of the 16th international conference on Intelligent user interfaces, pp.423-426, 2011.
- [9] Yoshikazu NAKAMURA and Jun’ichi TOYODA, members, “An Extraction of Individual Handwriting Characteristics Based on Calligraphic Skills”, The Journal of the Institute of Electronics, Information, and Communication Engineers, Vol.J77-D-II, No.3, pp.510-518, 1994.
- [10] Wacom | Cintiq 12WX,“<http://wacom.jp/jp/product/s/cintiq/cintiq12/spec/>”,2012.01.05アクセス.
- [11] The R Project for Statistical Computing,“<http://www.r-project.org/>”,2011.01.06アクセス.
- [12] CRAN - Package e1071,“<http://cran.r-project.org/web/packages/e1071/index.html>”,2011.01.06アクセス.