

学習者の記憶の度合とオンライン手書きデータの関係性の調査

苑田 翔吾^{†1} 浅井 洋樹^{†1†2} 山名 早人^{†3†4}

^{†1} 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

^{†2} 早稲田大学メディアネットワークセンター 〒169-8050 東京都新宿区戸塚町 1-104

^{†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 本稿では、著者らの先行研究を基に記憶支援アプリケーションにおける記憶度とオンライン手書きデータ特徴量の検証を行う。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。先行研究では、記憶のテストの際、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。本稿ではさらにこれらの関係性を調査し、記憶支援アプリケーションにフィードバックを行うシステムの構築を目指す。

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、誰しもが必要に迫られ経験する行為である。従って、記憶にかか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推測し記録している。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、記憶したい事象を問題と解答として分け、ユーザは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、ユーザの記憶を正確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。従って、完全に覚えている事象と偶然思い出し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

¹ えいご漬け, “<http://www.nintendo.co.jp/ds/angj/>,” 2013. 01.08 アクセス。

² Nintendo DS - Official Site, “<http://www.nintendo.com/ds/>,” 2013.01.08 アクセス。

³ Apple - iPad - From the App Store, “<http://www.apple.com/ipad/from-the-app-store/>,” 2013.01.08 アクセス。

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。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。

本稿は先行研究として、著者らが過去に執筆した論文[4]を参照する。先行研究では、記憶に関連するオンライン手書きデータの特徴量を選別するために、7名の被験者のオンライン手書きデータを解析し、特徴量を決定づけている。しかし、7名という人数は統計的处理を行うに当たって十分な人数ではない。また、漢字の書き取り問題という特定のタスクのみを取り上げているため、他のタスクでも同様の結論を導けるかが疑問である。そこで、本稿では先行研究の結論を再度検証すると同時に、漢字の書き問題と同時に漢字の読み問題を被験者に解かせることで、幅広いタスクで記憶度を推定可能であるかどうかを確かめる。

本稿では次の構成をとる。まず2節で先行研究について述べる。次に3節で検証実験の目的と検証実験の方法を述べた後、4節で検証実験の結果と検証実験の考察を述べる。最後に5節でまとめる。

2. 先行研究

先行研究[4]では、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査する実験を行った。学習者が、記憶度が高いと判断した場合と記憶度が低いと判断した場合の両者のオンライン手書きデータ間で検定を行い、有意な差が表れる特徴量を明らかにした。

2.1. 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[5]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[6]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7名にテストを依頼した。

2.2. 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書かせる。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1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2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3に示した。

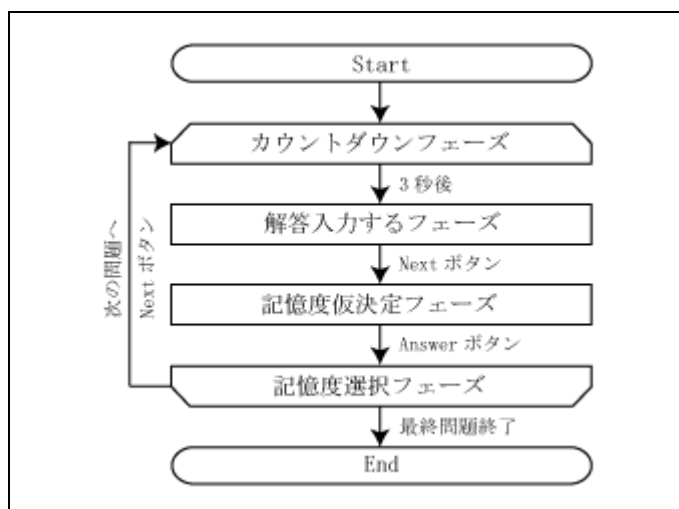


図 1 実験のフローチャート

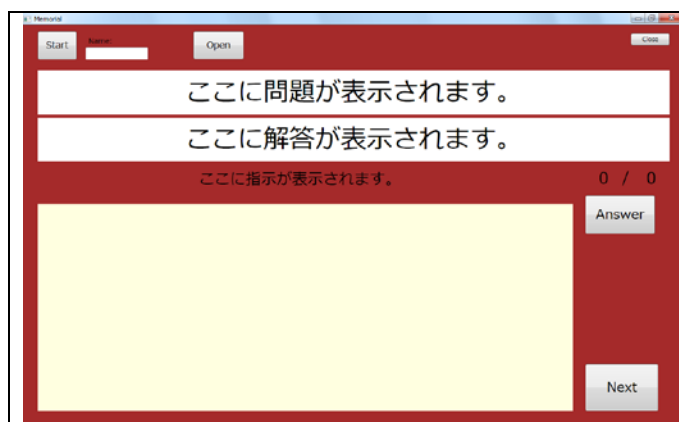


図 2 アプリケーション起動時

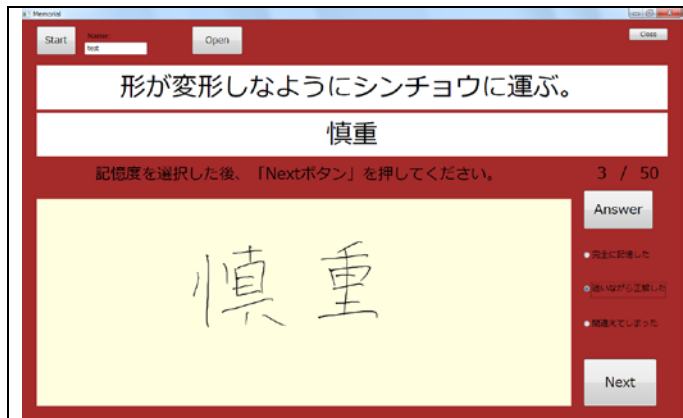


図 3 記憶度選択時

図 1 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。

2.2.1. 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、被験者は次の問題への準備時間を十分に確保することができる。

2.2.2. 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。解答を入力し終わったら、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。

2.2.3. 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。

2.2.4. 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の一つを選択し、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。

2.2.5. 記憶度の選択肢

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。本稿では、問題の解答の際に迷いが生じたか否か、すなわち「完全に記憶している」と「迷いながら正解した」の選択を記憶の度合いの正解と定義する。

2.2.6. 出題する問題

本実験では、財団法人日本漢字能力検定協会⁴が実施する漢字検定の書き取り問題を出題した。漢字検定準 1 級から 5 問、2 級から 10 問、準 2 級から 5 問の計 20 問を被験者に解答させた。

2.3. オンライン手書きデータ特徴量

本項では、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 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 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。

2.3.1. 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。

⁴財団法人 日本漢字能力検定協会,”
<http://www.kanken.or.jp/index.php>”, 2013.01.08 アクセス。

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。

2.3.2. 計算によって求められるオンライン手書き特微量

2.3.2 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微量は無数に存在するが、本稿では関連研究[7][8]と筆者の推測を基に、以下のオンライン手書き特微量を設定した。

- ストロークの時間に関する特微量
 - ・ 書き始めるまでの時間 (IniStrT)
 - ・ ストロークの時間間隔の総和 (StrIntT)
- ストロークの筆圧に関する特微量
 - ・ 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微量
 - ・ 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微量
 - ・ 消去したストロークの本数 (CntEraseStr)
 - ・ 消しゴムモードに遷移した回数 (CntEraseMode)
 - ・ 消しゴムモードに遷移していた時間 (EraseT)

各特微量の省略語括弧内で示した。移行の表や図では、各特微量を省略語で示す。

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(X_{i1}, Y_{i1}), (X_{i2}, Y_{i2}), ..., (X_{ii}, Y_{ii})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。

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

2.4. 実験結果

2.4.1. 検定による特微量選択

「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微量の相違を確認するために、各特微量において被験者ごとにマンホイットニーのU検定を行った。

有意水準を5%としたとき、以下の特微量において有意差が見られた。

- 書き始めるまでの時間
- ストロークの時間間隔の総和
- 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

3. 検証実験の目的と検証実験方法

2節の先行研究より、漢字の書き取り問題において、学習者の記憶度は「ストロークの時間に関する特微量」と「消去ストロークに関する特微量」に関係性があることが明らかとなった。本稿では、先行研究の結果をより汎用性が高く信憑性が高いものとするために、以下の実験を行う。

3.1. 検証実験の目的

先行研究の問題点を解消するために、実験によって明らかにしたい項目と理由を以下に示す。

- 十分な被験者人数の下、学習者の記憶度とオンライン手書きデータの特微量の関係性の確認を行う。

先行研究では7人の被験者で学習者の記憶度とオンライン手書きデータの特微量の関係性を調査したが、7人という人数は統計的処理を行う上で十分な人数であるとは言いがたい。そこで、本稿では被験者の人数を増やし、先行研究で明らかとなったオンライン手書きデータ特微量と記憶度の関係性を再度確認する。

- 他タスクについて、学習者の記憶度とオンライン手書きデータ特徴量の関係性を確認する。

先行研究では、漢字の書き取り問題にタスクを限定しているため、漢字の書き取り問題以外の他のタスクでも同様に、学習者の記憶度とオンライン手書きデータ特徴量の関係性が言えるかどうかを確認していない。そこで、本稿では漢字の読み問題タスクを被験者に解かせ、検証を行う。

- 同様の記憶度の漢字を書いた場合における各オンライン手書きデータ特徴量のばらつき度合いを見る。

同様の記憶度の漢字を書いた場合でも、オンライン手書きデータ特徴量に少なからず差異はあるはずである。しかし、先行研究ではオンライン手書きデータ特徴量のばらつきの度合いを考慮せず、記憶の度合いを判断していたため、記憶の度合いの精度が落ちてしまっていた可能性がある。そこで、本稿では「完全に記憶している」問題を被験者に複数問解かせ、オンライン手書きデータ特徴量のばらつき度合いを観察する。

3.2. 検証実験の方法

専門分野が多様な大学生男性 15 名、女性 15 名の計 30 名に実験を依頼した。その他、漢字出題のシステムやオンライン手書きデータの収集方法は 2 節の先行研究と同様である。

3.1 項で述べた目的を達成するために、以下の手順に従って実験を行った。

1. チュートリアル
2. 易しい漢字の書き取り問題（8 級）の確認
3. 易しい漢字の書き取り問題（8 級）の解答
4. 難しい漢字の書き取り問題（2 級）の解答
5. 易しい漢字の読み問題（8 級）の確認
6. 易しい漢字の書き取り問題（8 級）の解答
7. 難しい漢字の書き取り問題（準 1 級）の解答
8. 記憶の度合いに関するアンケート

以上の手順について順に説明する。尚、括弧内は漢字検定のレベルを示している。

3.2.1. チュートリアル

被験者が、実験に用いるタブレット端末の使い方と漢字の解答システムの操作方法を十分に理解できるように、実験のチュートリアルを用意した。実験のチュートリアルは計 4 問で構成されており、漢字の書き取り問題が 2 問、漢字の読み問題が 2 問出題される。チュートリアルにより、全ての被験者が実験の環境を理解した上で、以降の問題を解答することが可能となる。

3.2.2. 簡単な漢字の書き・読み問題の確認と解答

同様の記憶度の漢字を書いた場合における各オ

ンライン手書きデータ特徴量のばらつき度合いを見るために、被験者に漢字検定 8 級相当の漢字の易しい漢字の書き・読み問題を各 10 題ずつ出題し、解答させた。この時出題される漢字は、被験者が「完全に記憶している」漢字であると想定しているため、被験者が完全に記憶していない漢字が出題されてしまった場合は、意図したデータを取得することができない。これを防ぐため、被験者が易しい漢字の書き・読み問題を解答する前に、易しい漢字の書き・読みの全問題と全解答を予め被験者に見せ、完全に記憶しているかどうかを確認させた。

3.2.3. 難しい漢字の書き・読み問題の解答

難しい漢字の書き・読みの問題を各々 50 問ずつ出題し、被験者に解答させた。漢字の書き取り問題には、漢字検定 2 級相当の問題を出題し、漢字の読み問題には、漢字検定準 1 級相当の問題を出題した。各々問題を解答した後、問題に正答できた場合、被験者に「完全に記憶している」か「迷いながら正解した」のどちらかを記憶度に応じて選択させた。

3.2.4. 記憶の度合いに関するアンケート

被験者に「完全に記憶している」と「迷いながら正解した」の判断基準の要素を選択させ、選択した判断基準の要素の並び替えを行わせた。判断基準の項目は以下の 8 種類である。

- A) 解答を思い出すまで（解答の 1 画目を書き始めるまで）に要した時間
- B) 解答を記入し始めてから記入し終わるまでに要した時間
- C) 消しゴムを使用したか否か
- D) 解答を記入するときの筆記速度
- E) 解答を記入するときの筆圧
- F) 表面的な動きには出ていないが、頭の中で「完全に記憶していない」と思いながら解答を記入した。
- G) 初めに思い浮かんだ解答が正答とは異なる解答だった。
- H) その他（理由も記載して下さい。）

4. 検証実験の結果と検証実験の考察

4.1 項では、本稿で使用するオンライン手書きデータ特徴量を説明し、4.2 項以降は 3.1 項で述べた目的に即して各々説明していく。

4.1. オンライン手書きデータ特徴量

2 節の先行研究から、ストロークの時間とストロークの消去に関する特徴量を以下のように定義した。

- 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の最大値（MaxIntT）

● 消しゴムモードに関する特徴量

- ・ 消しゴムモードに遷移していた時間 (EraseT)

先行研究ではストロークの時間間隔の総和を特徴量として定義していたが、ストロークの時間間隔の総和は解答する文字によって有利不利がある。そこで、本稿ではストロークの時間間隔の最大値を新たな特徴量として定義した。ストロークの時間間隔の最大値とは、前のストロークを終えてから次のストロークを始めるまでの時間の中で最も長い時間である。

4.2. 記憶判別の検証

2 節の先行研究と同様に、「完全に記憶している」「迷いながら正解した」と回答された難しい漢字の書き取り問題のオンライン手書きデータを対象に、4.1 項で挙げた 3 つの特徴量を用いて SVM による交差検定を行う。SVM と交差検定には R[9]を用いた。

4.2.1. 検証方法

R の e1071 パッケージ[10]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。被験者の任意の交差検定データ 1 問を取り除いた後、残りの被験者自身の交差検定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、各パ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. 検証結果

実験の結果、「迷いながら正解した」と答えた問題が 2 問以下だった 7 人の被験者については、学習器の学習が十分に行えないため、値の算出から除外した。従って、残りの 23 人の各被験者において、「迷いながら正解した」に関する精度、再現率、適合率を算出した。各被験者における精度、再現率、適合率の度数分布図を図 4 に示す。

全被験者の平均値は精度約 80%，再現率約 51%，適合率約 83%となった。

先行研究と比較して、精度がやや低下し、再現率が大幅に低下してしまった。これは先行研究の被験者は研究室から募った被験者であることが最も大きな要因だと考えられる。研究室の被験者は本研究に対す

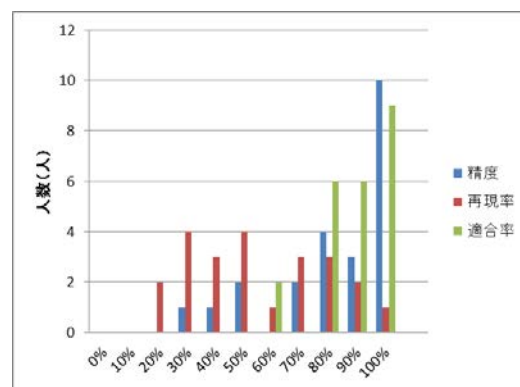


図 4 漢字書き取り問題における各被験者の精度、再現率、適合率の度数分布図

る理解があり、バイアスがかかってしまった可能性が高い。従って、本研究結果の値が真の精度、再現率、適合率を表していると考えられる。

4.3. 漢字の読み問題に関する記憶度と特徴量の関係性の検証

4.2 項と同様の検証実験を漢字の読み問題を対象に行い、漢字の読み問題においても漢字の書き取り問題と同様に、記憶度の判別が可能であるかどうかを検証する。

4.3.1. 検証方法

検定対象データが漢字の読み問題になる点を除いて、4.2.1 の検証方法に準ずる。

4.3.2. 検証結果

実験の結果、4.2.2 と同様に、1 人の被験者を除外した後に、被験者ごとに「迷いながら正解した」に関する精度、再現率、適合率を算出した。各被験者における精度、再現率、適合率の度数分布図を図 5 に示す。

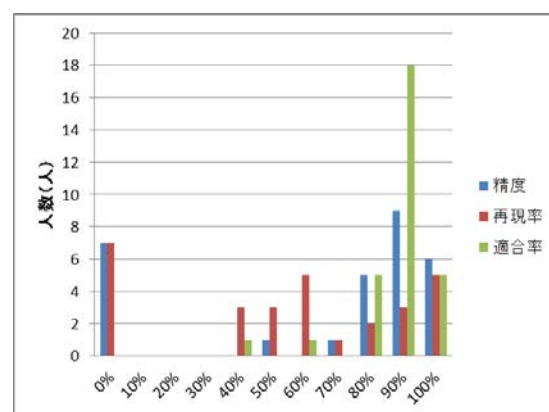


図 5 漢字読み問題における各被験者の精度、再現率、適合率の度数分布図

全被験者の平均値は精度約 64%，再現率約 52%，適合率約 83%となった。精度と再現率が 0%となってしまう被験者が 7 人いた。これは「完全に記憶している」と「迷いながら正解した」における特徴量の差異が十分になく、学習器が予測を行う際、全て「完全に

記憶している」と予測してしまったためである。

4.2.2 の漢字の書き取り問題の検証結果と比較すると、漢字の読み問題において、精度が大きく減少してしまった。この要因を探るために、3.2.4 の記憶の度合いに関するアンケート結果を考察する。

4.3.3. アンケート結果

3.2.4 の記憶の度合いに関するアンケート結果を表 1 に示す。表 1 内で記載されている数値は全被験者数である 30 人を最大値とした累積の数値である。

表 1 記憶の度合いに関するアンケート結果

順位	漢字の書き取り問題		漢字の読み問題	
	回答	人数	回答	人数
1 番 目	A	14	A	14
	F	10	F	13
	B	1	B	0
2 番 目	A	21	A	22
	F	13	F	20
	B	12	B	5
3 番 目	A	25	A	26
	F	20	F	23
	B	13	B	10

漢字の書き取り問題は、高い順位に A（解答を思い出すまで（解答の 1 画目を書き始めるまで）に要した時間。）が選ばれる傾向にあり、漢字の読み問題は、漢字の書き取り問題と比較して、高い順位に F（表面的な動きには出ていないが、頭の中で「完全に記憶していない」と思いながら解答を記入した。）と回答をする人が多いことが、表 1 からわかる。A に関しては特徴量に反映され、効果的な学習が行えるが、F に関してはデバイスから情報を取得することはできない。そのため、漢字の読み問題においては、精度が低下したと考えられる。

以上より、漢字の書き取り問題から漢字の読み問題へタスクを変更すると、予測の精度が低下するという検証結果とその理由が得られた。すなわち、記憶の度合いがオンライン手書きデータ特徴量に反映されやすいタスクと反映されづらいタスクがあることが明らかとなった。

4.4. 同様の記憶度による特徴量のばらつきの検証

同一の記憶度によるオンラインデータ手書き特徴量のばらつきを検証するために、易しい漢字の書き取り問題と易しい漢字の読み問題を被験者に解かせた結果を検証する。また、これらの易しい問題から得られたオンライン手書きデータ特徴量と、難しい漢字の書き取り問題と難しい漢字の読み問題の「完全に記憶している」項目のオンライン手書きデータ特徴量を比較する。

4.4.1. 検証方法

漢字の問題の解答に関する各オンライン手書きデータ特徴量から、4.1 項のオンライン手書きデータ特徴量に関する平均と標準偏差を被験者ごとに求め、さらに被験者ごとの数値から、被験者全体の平均と標準偏差を求め、検証する。

4.4.2. 検証結果

求めた結果を表 2 に示す。ただし、表 2 内の数値の単位は秒とし、各特徴量の接頭の文字 Ave、Std はそれぞれ平均、標準偏差を意味する。まず、各被験者各問題の IniStrT、MaxIntT、EraseT に関して、各々平均値、標準偏差値を求める。そして、これらの値を AveIniStrT、StdIniStrT、AveMaxIntT、StdMaxIntT、AveEraseT、StdEraseT とおくと、被験者ごとに値が算出可能なので、30 人分データが作成可能である。これらの 30 人分のデータに対して、平均と標準偏差を算出し、まとめたものが表 2 である。また、漢字書き取り問題と漢字読み問題の双方において、各被験者の書き始めるまでの平均時間の度数分布図を図 6 と図 7 に示す。

表 2 難易度によるオンライン手書きデータ特徴量の差異（秒）

	Ave IniStrT	Std IniStrT	Ave MaxIntT	Std MaxIntT	Ave EraseT	Std EraseT
易・書問題の平均	1.50	0.50	0.48	0.26	0.13	0.33
難・書問題の平均	2.59	1.40	0.88	0.92	0.08	0.17
易・読問題の平均	1.10	0.22	0.35	0.14	0.12	0.35
難・読問題の平均	1.89	0.86	0.46	0.50	0.12	0.37
易・書問題の標準偏差	0.40	0.88	0.13	0.25	0.08	0.17
難・書問題の標準偏差	0.92	1.26	0.36	0.56	0.16	0.29
易・読問題の標準偏差	0.17	0.11	0.12	0.18	0.05	0.16
難・読問題の標準偏差	0.56	0.65	0.21	0.56	0.10	0.25

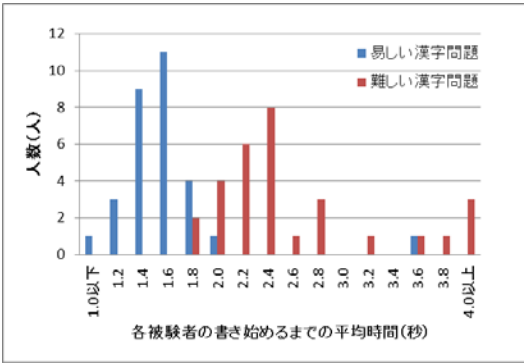


図 6 漢字書き取り問題において各被験者の書き始めるまでの平均時間の度数分布図

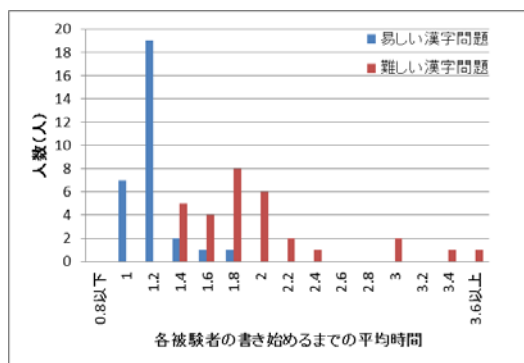


図 7 漢字読み問題において各被験者の書き始めるまでの平均時間の度数分布図

表 2 の数値は全て、被験者が「完全に記憶している」と判断した際に得られたオンライン手書きデータ特徴量であるが、難しい漢字の書き取り問題の「書き始めるまでの時間」や「ストロークの時間間隔の最大値」の値が、易しい漢字の書き取り問題のそれらの値よりも大きな数値をとることがわかる。また、図 6 と図 7 の度数分布図からも、このことが確認できる。特に「易しい漢字の書き問題の書き始めるまでの時間の平均」と「難しい漢字の書き取り問題の書き始めるまでの時間の平均」は、図 6 から差異があることが明らかである。

このような差異が生まれた理由は、易しい問題を解いた際に選択した「完全に記憶している」問題と難しい問題を解いた際に選択した「完全に記憶している」問題とでは、記憶の度合が異なる可能性が高いと言える。さらに、被験者が記憶の度合が異なることに気づいていなかったか気づいていないかによって 2 通りの理由が考えられる。

- 被験者は記憶の度合に気づいていたが、「完全に記憶している」と「迷いながら正解した」のどちらかを選択しなければならないので「完全に記憶している」を選択した。
- 被験者は記憶の度合の変化に気づいておらず、難しい問題については、「完全に記憶している」と選択する基準を無意識に下げてしまった。

実験を通して以上の知見を得られた。

5. まとめ

本稿では、学習者の記憶の度合とオンライン手書きデータ特徴量の関係性について被験者実験を通して検証した。検証の結果、「問題における正答を『完全に記憶している』『迷いながら正解した』の 2 値に区別することが、漢字の書き取り問題において精度約 80%、再現率約 51%、漢字の読み問題において精度約 64%、再現率約 52%で判別可能であること」「タスクによって、オンライン手書きデータ特徴量への記憶の度合の反映度が異なること」「被験者が『完全に記憶している』と

判断した場合でも、オンライン手書きデータ特徴量には大きな差異があること」が明らかとなった。

今後の課題としては、本稿を応用した記憶支援アプリケーションの開発とそのアプリケーションに用いるための効率的な機械学習が行える機械学習器の設計・開発が挙げられる。

謝辞

本研究の一部は、早稲田大学特定課題研究（2012）によるものである。

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”, 2012.12.4 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, 2012.12.4 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] 苑田翔吾, 浅井洋樹, 山名早人, “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”, DEIM2012, C8-4, 2012.
- [5] Wacom | Cintiq 12WX, “<http://cintiq.jp/12wx/>”, 2013.01.05 アクセス。
- [6] Windows Presentation Foundation, “<http://msdn.microsoft.com/en-us/library/ms754130.aspx>”, 2013.01.05 アクセス。
- [7] 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.
- [8] Yoshikazu NAKAMURA and Jun’ichi TOYODA, , “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.
- [9] The R Project for Statistical Computing, “<http://www.r-project.org/>”, 2011.01.06 アクセス。
- [10] CRAN - Package e1071, “<http://cran.r-project.org/web/packages/e1071/index.html>”, 2013.01.06 アクセス。