

表情・動作情報とタッチパネル操作情報の統合による 学習状態の推定

須藤 優介[†] 山名 早人^{‡§}

[†] 早稲田大学基幹理工学研究科 〒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, {sutou, yamana}@yama.info.waseda.ac.jp

あらまし 効率的な学習指導を行うために、学習状態を機械的に推測する研究が行われている。従来、学習状態推定の研究分野では、デスクトップ環境を前提とした研究が広く行われてきた。一方で、教育現場ではタブレットデバイスの導入を積極的に進めており、今後タブレットデバイスを用いた学習状態の推定は重要度を増していくと考えられる。そこで本研究では、タブレットデバイスから得られる情報を用いた学習状態の推定手法を提案する。学習者の表情・動作と、タッチパネル操作情報を特徴量として、問題を解くときに悩んだか否か、自信を持って答えたか否かを推定する。また、学習者の表情や動作が安定して取得できない状況を考慮した特徴量の重み付けを行い、解答時に悩んだか否か、自信があったか否かの推定について、それぞれ平均精度 75.3%、66.7%、平均再現率 74.8%、60.1%を達成した。

キーワード ユーザインタフェース, 感情推定, 学習支援, 端末操作情報, 画像処理

1. はじめに

効率的な学習指導を行うためには、教師が生徒の学習状態を理解し、学習状態に応じて指導方針を決定していくことが望ましい。しかし、一斉授業や家庭学習といった学習環境では、生徒一人ひとりの学習状態を教師が把握することは難しい。このような問題を解決するために、電子機器を用いて自動的に学習状態を推定するような研究が行われている [1]-[7]。

従来、e-Learning や一斉授業で使用される電子機器としては、デスクトップ型の PC が主流であった。そのため、学習状態推定の研究分野でも、デスクトップ型の PC で学習を行うユーザを対象とした研究が広く行われてきた。これらの研究では、生体情報、表情、動作、端末操作など、学習者の様々な情報を基に感情推定を行っている。学習者の感情を推定することにより、理解度合はどの程度か、教材が適切か否か、今後どのように教育していくかといった事の指標とすることができる。しかしながら、

- ・ デスクトップ PC を用いた授業自体が少ない
- ・ 特殊な装置を用いて生体情報を取得するなど、一般的でない学習環境を要する研究が多い

という問題があり、一般的な学習環境への適用は難しかった。

一方で、近年、教育現場ではタブレットデバイスの導入が進んでおり、総務省が展開する「フューチャースクール推進事業 [8]」の中でも、すべての児童・生徒に対し 1 台ずつタブレットデバイスを配布することを目指している。教育現場のタブレットデバイス導入を受けて、タブレットデバイスを使用する学習者の学習

状態推定は今後重要度を増していくと考えられるが、タブレットデバイスを用いて学習状態を推定するような研究は、我々の知る限り存在しない。

そこで本研究では、タブレットデバイスから得られる情報のみを基に学習状態の推定を行う手法を提案する。本手法では、既存研究で学習状態推定への有用性が示されている「学習者の表情・動作」と「端末操作情報」をタブレットデバイスから取得し、それらを統合して学習状態を推定する。さらに、表情・動作情報を安定して取得することは困難であるという前提のもと、表情・動作情報取得状態に応じて、学習状態推定時の特徴量を選択する。学習環境としては、英語の四択問題に解答する環境を想定し、「悩んでいるか否か」、「自信を持って答えたか否か」の 2 つに関して推定を行う。

本稿は以下の構成をとる。まず 2 節では関連研究をまとめ、3 節で提案手法について説明する。さらに 4 節で実験・評価について述べ、5 節でまとめを行う。

2. 関連研究

タブレットデバイスを用いた学習状態推定の研究は、我々の知る限り存在しない。本節では、デスクトップ PC で学習する学習者を対象とした、既存の学習状態推定に関する研究について述べる。

2.1. 表情・動作・生体情報に基づく学習状態推定

学習者の表情、動作、生体情報を基に感情推定を行った研究について、表 1 にまとめた。これらの研究では、表情・動作、生体情報といった情報が、学習状態推定に有用な情報であることが示されている。

しかし、表情や動作に基づく学習状態推定の研究で

は、表情や動作が取得できない状態では推定が行えないという問題がある。繁田ら[2]の研究では、初期フレームにおいて顔の特徴点を手動で設定するにも関わらず、フェイストラッキング中に特徴点が5画素以上ずれる割合が58~89%になるという結果が出ている。

また、生体情報を取得する研究では、生体情報取得のために学習者が特殊な装置を装着する必要がある、学習者の負担が大きくなる上に、一般的な学習環境下では適用するのが難しい。

表 1 表情・動作・生体情報に基づく学習状態推定の研究

年	著者	使用特徴量	推定項目
'07	黒川ら[1]	・ 眉毛・唇の動き ・ 脈拍数 ・ 呼吸数 ・ 皮膚温度	・ 理解している/していない ・ 飽きている/興味がある ・ 疲れている/集中している ・ 難しい/簡単
'09	繁田ら[2]	・ 眉毛の動き ・ 唇の動き	・ 理解している/していない ・ 飽きている/興味がある ・ 疲れている/集中している ・ 難しい/簡単
'10	中村ら[3]	・ 顔の動き・傾き ・ 凝視時間 ・ つぶやき時間	・ 主観的難易度
'12	Szafirら[4]	・ 脳波	・ 集中度

2.2. 端末操作情報に基づく学習状態推定

端末操作情報を用いて学習状態推定を行った中村ら[5]の研究について、表 2 にまとめた。中村らの研究では、操作時間間隔を基に、学習者の行き詰まり状態を約 80%の再現率で検出している。

しかし、学習者自身の操作で演習問題に対するヒントを自由に表示・閲覧できるような特殊な環境下で実験を行っており、一般的な学習環境下での有用性は示されていない。

表 2 端末操作情報に基づく学習状態推定の研究

年	著者	使用特徴量	推定項目
'02	中村ら[5]	・ 操作時間間隔	・ 行き詰まり状態

2.3. 表情・動作と端末情報に基づく学習状態推定

表情・動作情報と端末操作情報を両方用いた学習状態推定の研究について、表 3 にまとめた。Kapoorら[6]の研究では、興味があるか否かについて約 86%の精度、中村ら[7]の研究では、主観的難易度について約 85%の精度で推定に成功している。しかし、Kapoorらの手法では赤外線カメラと圧力センサ付きの椅子、中村らの研究ではステレオカメラを使用するなど、両者ともに特殊な装置を用いて学習者の情報を取得している。これらの装置はいずれも一般的な学習環境では導入が難しい。

表 3 表情・動作と端末情報に基づく学習状態推定の研究

年	著者	使用特徴量	推定項目
'05	Kapoorら[6]	・ 眉毛・目・口の動き ・ 頭の動き ・ 姿勢 ・ 問題のレベル・進行状況	・ 興味があるか否か
'07	中村ら[7]	・ 表情 ・ 操作時間間隔	・ 主観的難易度

2.4. 関連研究のまとめ

本節では、既存の学習状態推定に関する研究として、

- ・ 表情・動作や生体情報を使用した研究
- ・ 端末操作情報を使用した研究
- ・ 表情・動作と端末操作情報を組み合わせた研究を紹介した。

いずれも高い精度での推定に成功しているが、

- ・ 特殊な装置を用いている
- ・ 特定の学習用ソフトウェアでしか実験を行っていない
- ・ 表情・動作を取得できない環境では推定を行えない

といった問題が残されている。

3. 提案手法

本節では、提案手法について詳細に述べる。提案手法では、タブレットデバイスを用いた学習状態推定という今までに無い試みを行うと同時に、既存研究の問題点解決を図る。具体的には、以下の様な要件を満たす学習状態推定システムを構築する。

- ・ 特殊な装置を用いない
- ・ 学習環境（教材、問題）に対する制限が弱い
- ・ 表情・動作が安定して取得できないことを考慮した学習状態推定

3.1. 対象とする学習環境・学習状態

提案手法では、タブレット端末を用いて英語の四択問題に解答するような環境を対象とし、解答時に「悩んだか否か」、「自信を持って答えたか否か」の推定を行う。

選択式の問題は、センター試験でも採用されるなど広く使用されており、学習教材・問題としては一般的なものであると考えられる。また、選択式の問題では解答の途中経過が見られず、学習者の理解度を解答結果から推測することが難しい。これらの理由から、選択式問題は、学習状態の推定が重要な意味を持つ学習環境であると考えた。また、推定する学習状態を、「悩んだか否か」、「自信を持って答えたか否か」としたのは、教師が生徒の理解度を把握する上で有用な判断材料になると考えたためである。

実際に実験を行った時の様子と、タブレットデバイス上の解答画面を図 1 に示す。図 1(a)のように、タブレットデバイスは、学習者の顔が内側カメラに写る位置に設置し、角度をつけた状態で固定する。

学習者、教材に多少の制限をかけることになるものの、タブレットデバイス 1 台で構築可能な環境であり、

- ・ 特殊な装置を用いない
 - ・ 学習環境（教材、問題）に対する制限が弱い
- という要件は満たしていると考えられる。

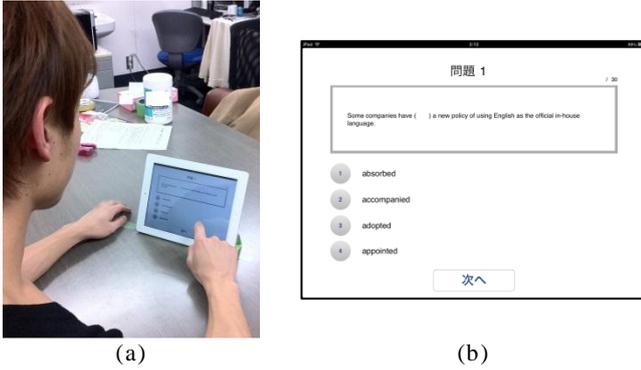


図 1 (a)解答中の様子 (b)実際の解答画面

3.2. 表情・動作情報の解析

表情・動作情報の取得には、Active Shape Model[9]を用いる。モデルベースの手法は、最新の感情推定研究[10]でも有用性が示されており、カメラ以外の特殊な機器を必要としないため、提案手法での採用を決定した。なお、提案手法では、Yao Wei らが開発したASMLibrary[11]を使用する。

Active Shape Modelにより、動画像上の顔領域から図 2 のように 68 点の特徴点を得ることができる。特徴点は、動画像の左上の点を原点として、画像右方向を x 軸正方向、下方向を y 軸正方向とした xy 平面上の座標で表現される。このように取得した特徴点を基に、以下の特徴量を抽出する。

- A) 眉毛の動き
- B) 口の動き
- C) 目の開閉の動き
- D) 頭の動き (鼻の座標の動き)
- E) 画面からの距離 (検出した顔の大きさ)
- F) 顔の傾き
- G) 特徴点検出に成功したフレームの割合 (68 点すべての特徴点に成功した状態を成功とする)

特徴量の選定は既存研究を基に行った。表情の変化は、眉毛・目・口の動きによってなされ、既存手法でも表情の変化を取得するためにこれらの特徴量として使用している。提案手法では、眉毛・目・口の動きを、特定の特徴点同士の距離の変化で表現する。例として、眉毛の動きの取得方法を以下に示す。

- step1. 動画の N フレーム目で、目の上部の点と、眉毛上の 3 点を繋いだ 3 本の線分の長さを合計する (左右それぞれについて行う)
- step2. step1 で算出した値を、検出した顔の大きさと正規化する (左右それぞれについて行う)
- step3. $N+1$ フレームでも、step1, step2 と同じ操作を行う
- step4. step2 の算出結果と、step3 の算出結果の差分を取る (左右それぞれについて行う)

- step5. 左右それぞれの差分を合計する
- step6. 全フレームに対して step1~step5 の処理を行い、step5 の算出結果の総和をとる
- step7. 処理したフレーム数で step6 の結果を割る (正規化)

口の動き、目の動きに関しても、上記の処理手順と同様にして取得する。

動作情報としては、頭の動き、画面からの距離、頭の傾き、特徴点に成功したフレーム数を取得する。前述の眉毛や目の動きと同じように、頭の動きや画面からの距離、顔の傾きも、前のフレームとの差分の総和により表現する。また、学習者の動きが大きいとトラッキングに失敗することが多いため、特徴点検出に成功したフレームの割合を、動きの大きさを表す特徴量として使用している。

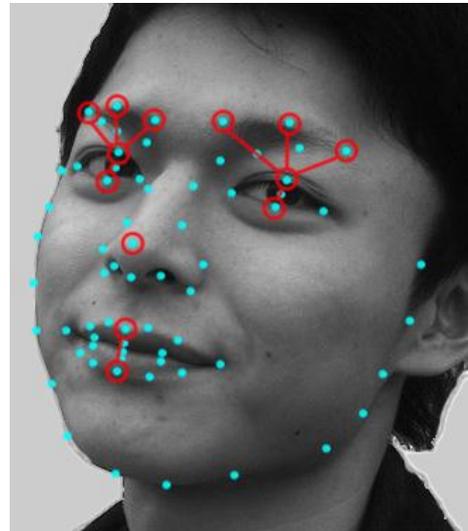


図 2 Active Shape Model によるアライメント

3.3. 端末操作情報の解析

提案手法では、端末操作情報として、以下の特徴量を抽出する。

- a) 答えがあっているか否か
- b) 問題に答えるのににかかった時間
- c) 答えを変更した回数
- d) 最後に答えを選んでから次の問題へ行くまでの時間
- e) 最後の答えを選択した時のタップ時間 (画面に触れている時間)
- f) タップしていた時間の合計

端末操作情報としては、操作時間間隔が学習状態推定に有用であることが既存研究で示されている[5][7]。本研究では、英語の四択問題に解答する環境を対象として学習状態の推定を行うため、操作時間間隔を細分化し、「問題に答えるのににかかった時間」と「最後に答えを選んでから次の問題へ行くまでの時間」を取得している。また、答えの変更回数も、「悩んだか否か」の

推定に有効な特徴量であると判断し、抽出を行なっている。

さらに、本研究ではタッチデバイスで操作情報を取得するため、タッチデバイス特有の情報として、画面に触れている時間の長さを取得している。

3.4. 機械学習を用いた学習状態の推定

提案手法では、Multiple Kernel Learning(MKL)[12]を用いて、表情・動作情報と端末操作情報を統合した学習状態推定を行う。MKLを使用することにより、各特徴量に対する最適な重み付けを学習して、学習状態を推定することが可能になる。

また、提案手法では、表情・動作情報が安定して取得できないことを前提として学習状態推定を行う。具体的には、動画像処理において特徴点検出に成功しているフレームの割合が一定以下だった場合、端末操作情報のみを用いた推定に切り替えることで、推定精度・再現率の向上を図る。なお、提案手法では、経験的に、特徴点検出に成功しているフレームが全フレームの3割未満だった場合に、端末操作情報のみを用いた推定に切り替える。

4. 実験・評価

提案手法の評価を行うため、評価実験を行った。本節では、実験の方法と結果、および評価について述べる。

4.1. データ取得環境

本実験におけるデータ取得環境について、表4にまとめた。被験者には、タブレットデバイス上で全ての問題に解答してもらった後、各問題で悩んだか否か、自信を持って答えたか否かを問うアンケートに答えてもらった。

表4 データ取得環境

被験者	大学生 25 名 (男 11 : 女 14, 文系 21 : 理系 4)
問題	英語の 4 択問題 10 問 (センター試験レベル 20 問と中学英語レベル 10 問) 参考: 図 1(b)
制限時間	30 分
デバイス 設置環境	角度をつけて固定 参考: 図 1(a)
その他	・一度解答すると前の問題には戻れない ・キャプチャを行うことは事前に伝えている ・「カメラに写る位置から動かないように」など、学習者の動きに対して制限は設けていない

4.2. 実験結果

4.2.1. 取得データの傾向

我々が取得したデータの全体的な傾向について述

べる。

表5に、アンケート集計結果についてまとめた。25人の被験者に30問の問題を解いてもらったため、全750の回答を得ている。なお、本実験では、「まったく自信がなかった」、「あまり自信がなかった」という回答を「自信がなかった」としてまとめ、「少し自信があった」、「自信があった」という回答を「自信があった」としてまとめている。同様に、「とても悩んだ」、「少し悩んだ」を「悩んだ」、「あまり悩まなかった」、「まったく悩まなかった」を「悩まなかった」としてまとめている。

表5 アンケート集計結果

	自信がなかった	自信があった	合計
悩んだ	252	100	352
悩まなかった	28	370	398
合計	280	470	750

また、当該問題中3割以上のフレームで顔特徴点取得に成功していたのべ問題数の割合は72%、当該問題中、1フレームも顔特徴点取得に成功していなかったのべ問題数の割合は19.2%であり、特徴点の取得が安定して行えないことを示す結果となった。

4.2.2. 各手法との比較結果

提案手法が有効に機能しているか調べるために、以下の3パターンとの比較を行った。

- ① 全データ(表情・動作, 端末操作情報のすべて)を使用し, MKLを用いて分類する(表情・動作情報の取得状態に応じた使用特徴量の切り替えを行わない)

- ② 端末操作情報のみを利用し, MKLを用いて分類する

- ③ 全データ(表情・動作, 端末操作情報のすべて)を使用し, 特徴量の重み付けをせずに分類する
比較実験では, 25人分のデータを5人ずつの5セットに分割し, leave-one-out cross-validationにより, Precision, Recall, Accuracyを算出している。なお, 本実験では, 正解セットとしてアンケート結果を使用する。

以下, 図3, 図4に, 各手法を用いて各学習状態の推定を行った時のPrecision, Recall, 図5に, PrecisionとRecallで大きな差が見られなかった3つの手法のAccuracyを示す。

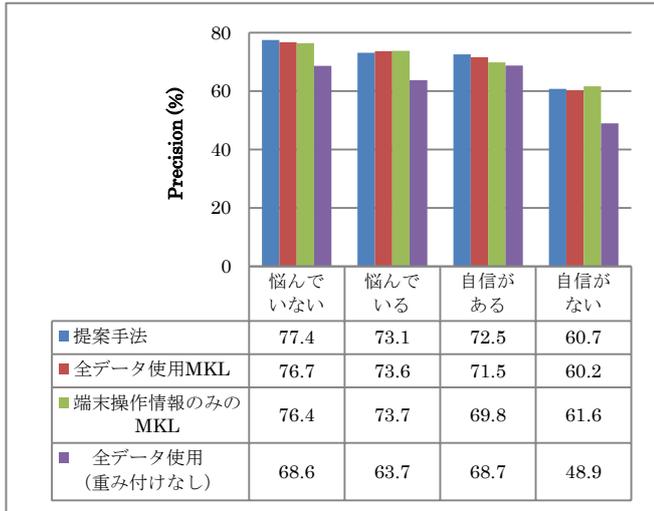


図 3 実験結果 (Precision)

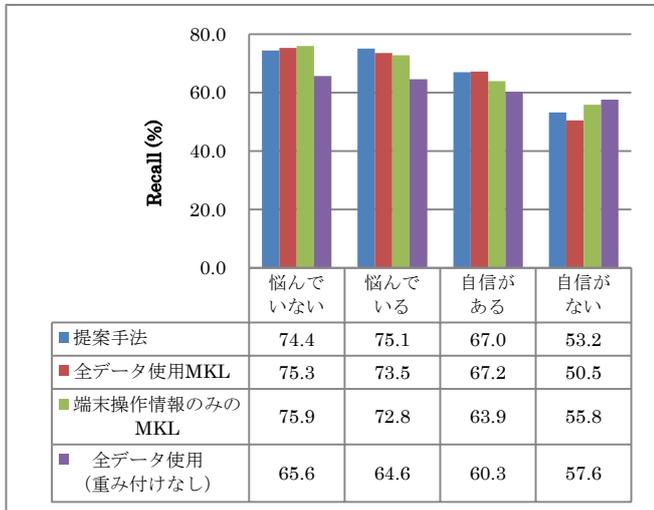


図 4 実験結果 (Recall)

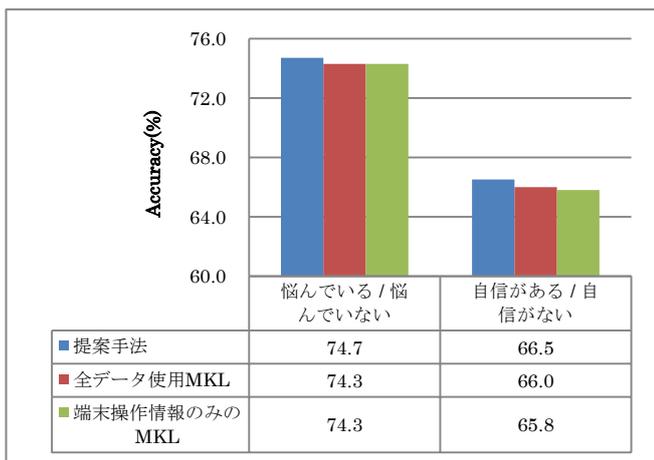
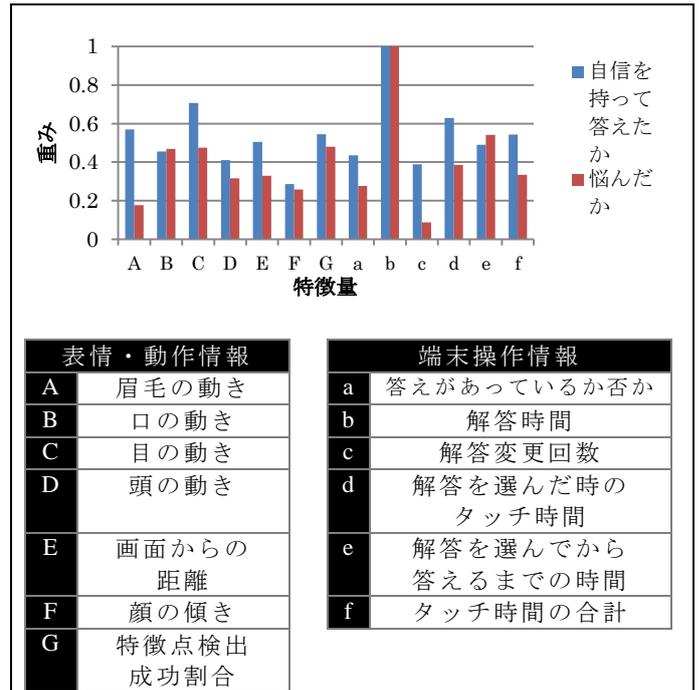


図 5 実験結果 (Accuracy)

また、MKLによる各特徴量に対する重み付けの結果を、図 6 に示す。



表情・動作情報		端末操作情報	
A	眉毛の動き	a	答えがあっているか否か
B	口の動き	b	解答時間
C	目の動き	c	解答変更回数
D	頭の動き	d	解答を選んだ時のタッチ時間
E	画面からの距離	e	解答を選んだから答えるまでの時間
F	顔の傾き	f	タッチ時間の合計
G	特徴点検出成功割合		

図 6 各特徴量に対する重み付けの結果

4.2.3. 個人差に関する実験の結果

図 3, 図 4, 図 5 を見ると、提案手法と、

- ・ 全データを使用し、単純に MKL を用いて分類する (表情・動作情報の取得状態に応じた使用特徴量の切り替えを行わない) 手法
- ・ 端末操作情報のみを利用し、MKL を用いて分類する手法

の実験結果には、大きな差が見られなかった。実際に、マン・ホイットニーの U 検定を行った結果、統計的に有意な差は確認されなかった。

2010 年中村ら[3]の研究では、学習中に現れる表情の変化には個人差があるという前提の下、学習者毎に識別器を作ることで高い精度を実現している。そこで我々も、表情情報が有効なユーザとそうでないユーザが存在する可能性を考え、各個人に対する各手法の Accuracy を比較した。悩んでいるか否かについて図 7、自信があるか否かについて図 8 に示す。

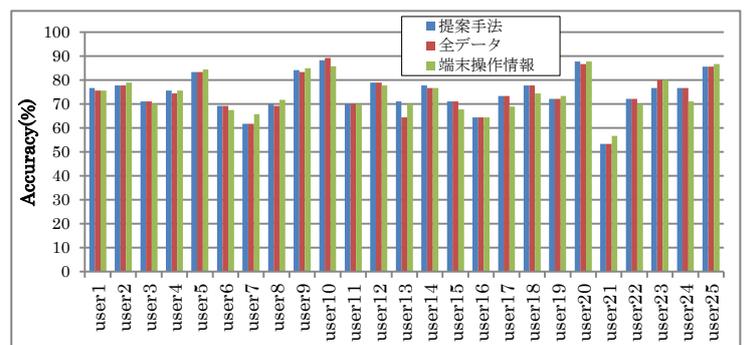


図 7 学習者毎の Accuracy (悩んでいるか否か)

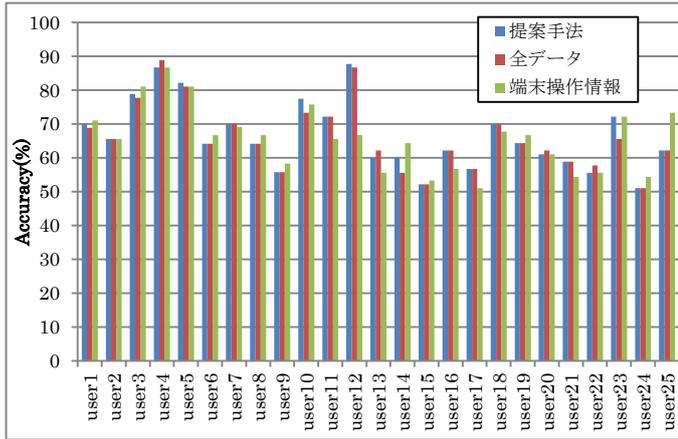


図 8 学習者毎の Accuracy (自信があるか否か)

実験結果を見ると、「悩んでいるか否か」に関しては、各手法の Accuracy が個人の中で大きく異なるケースは少ないが、「自信があるか否か」では、表情情報を使ったほうが Accuracy が高くなるユーザ、低くなるユーザが存在していることが分かった。

4.2.4. 傾向の似たユーザで学習を行った実験の結果

傾向の似た被験者で学習した結果を用いることで、精度の向上を図れるのではないかと考え、実験を行った。

実験では、被験者 20 名を、

- ・ 全てのデータを使用した場合のほうが Accuracy が高かったグループ
- ・ 端末操作情報のみを用いた場合のほうが Accuracy が高かったグループ

の 2 つのグループに分けた。全データを使用した場合の方が Accuracy が高いグループは、表情・動作に学習状態が表出しやすいグループであると言える。なお、全 25 名の被験者のうち、全てのデータを用いたときの結果と端末操作情報のみを用いたときの結果に大きな差が見られない被験者 5 名分のデータは除いた。上記の 2 つのグループに分けた後、さらにグループ内で 5 名ずつの 2 グループに分ける。そして、片方のグループを学習セット、残りのグループをテストセットとして推定を行う。その後、テストセットとして使用したグループで学習し、学習セットとして使用したグループをテストセットとして推定を行う。なお、提案手法では「全てのデータを使用した場合のほうが Accuracy が高かったグループ」のデータで実験を行った。こうして、各手法 10 名分のデータに対し学習状態の推定を実行した。結果を図 9 に示す。

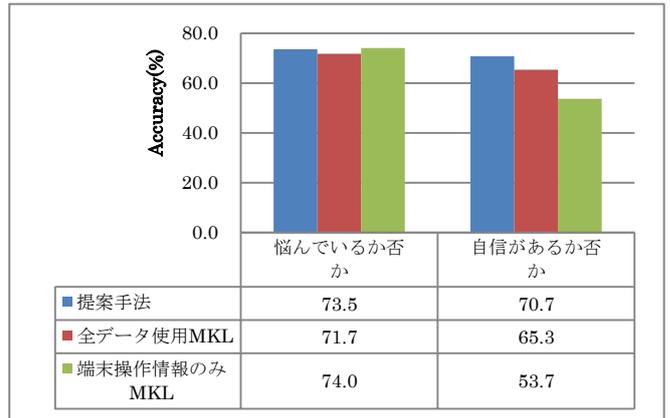


図 9 傾向の似たグループで学習を行った時の実験結果 (Accuracy)

実験結果から、「自信があるか否か」に関しては、グルーピングを行うことで提案手法の Accuracy が大きく向上しているのが見て取れる。一方で、「悩んでいるか否か」に関してはグルーピングの効果は見られなかった。

この結果から、

- ・ 「悩んでいるか否か」に関しては、表情・動作情報に学習状態が表出しにくく、傾向の似たユーザでグルーピングを行なっても効果が少ない
- ・ 「自信があるか否か」に関しては、表情・動作情報に学習状態が表出しやすいユーザ、しにくいユーザが混在しており、傾向の似たユーザのグルーピングが有効である

ということが分かった。

4.2.5. 学習状態の表出が弱いデータを除去した実験の結果

本研究では、正解セットとしてアンケートの結果を用いている。アンケートでは、各問題に対して以下のような質問が与えられ、被験者は自分が問題に解答した時のことを思い出してこの質問に答える。

【問 XX】で悩んだか否か

- とても悩んだ
- 少し悩んだ
- あまり悩まなかった
- まったく悩まなかった

【問 XX】で自信を持って答えたか否か

- まったく自信がなかった
- あまり自信がなかった
- 少し自信があった
- かなり自信があった

しかし、アンケート結果は被験者の記憶や性格に左右されてしまい、必ずしも正確なデータが取れるとは限らない。特に、「少し悩んだ」、「あまり悩まなかった」、「あまり自信がなかった」、「少し自信があった」といった中間に位置する回答は、被験者の正確な感情では無い可能性や、被験者の感情が表情や端末操作情報に表出しにくい可能性が考えられる。

そこで、アンケートへの回答が「少し悩んだ」、「あまり悩まなかった」、「あまり自信がなかった」、「少し自信があった」のいずれかに当てはまるデータをテストデータから除くことで、システムの精度が向上するか否かを調査した。結果を図 10 に示す。なお、アンケートへの回答が「少し悩んだ」、「あまり悩まなかった」、「あまり自信がなかった」、「少し自信があった」のいずれかに当てはまるデータを除いた結果、

悩んでいるか否か : 750 データ → 351 データ
 自信があるか否か : 750 データ → 384 データ
 となり、データ数は約半分となった。

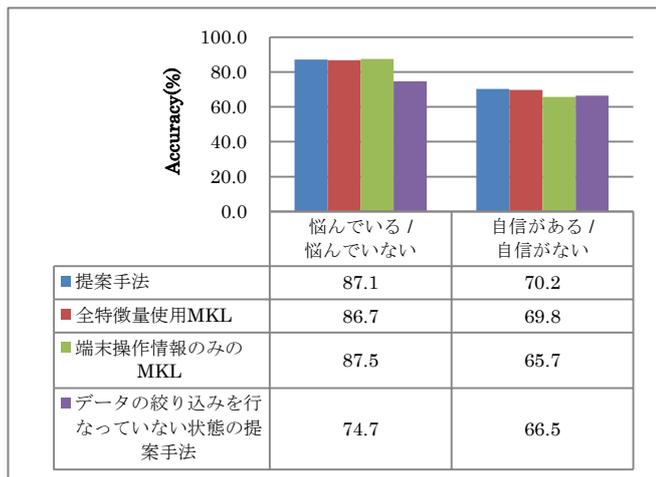


図 10 テストデータを限定したときの実験結果 (Accuracy)

実験結果から、「悩んでいるか否か」、「自信があるか否か」ともに、提案手法の Accuracy が向上していることが分かる。「悩んでいるか否か」の Accuracy に関しては、データの絞り込みを行っていない場合との比較で、マン・ホイットニーの U 検定において有意水準 5% での優位性が示された。

4.3. 評価

提案手法におけるポイントは、以下の 4 点である。

- ① タブレットデバイスを用いた学習状態の推定
- ② 特殊な装置を用いない
- ③ 学習環境 (教材, 問題) に対する制限が弱い
- ④ 表情・動作が安定して取得できないことを考慮した学習状態推定

したがって、今回の実験で評価すべきは、「タブレットデバイスのみを用いた、特殊な学習環境下でない

学習状態の推定で、十分な精度・再現率を達成出来るか」、「表情・動作が取得できないことを考慮した手法により、精度・再現率が向上するか」の二点であると考える。以下で、この二つの観点から提案手法の評価を行う。

4.3.1. タブレットデバイスのみを用いた学習状態推定に関する評価

既存研究では、特殊な学習環境下において、高い精度で学習者の感情推定をおこなっている。例えば、表情・動作情報を用いた研究としては、2010 年中村ら [3] の研究で、ステレオカメラを使用し、個人に特化した識別器を生成した環境下で、学習者が「難しい」と感じている状態を 84.6% の精度で推定している。2002 年中村ら [5] の研究では、独自の学習ソフト上で、操作時間間隔のみから、約 8 割の再現率で行き詰まり状態を検知している。

図 3, 図 4 を見ると、提案手法における精度・再現率は、既存手法と比べると低い値をとっている。しかしながら、「悩んでいない」状態に関しては、精度 77.4%、再現率 74.4% という値をとっており、他の学習状態に関する結果を見ても、システムの品質としては低くないと考えられる。提案手法では、タッチデバイス特有の情報を使用すると同時に、MKL により各特徴量に最適な重みを学習している。環境を制限しないにも関わらず高い精度・再現率を実現できたのは、これらが効率的に機能したからであると考えられる。実際に、図 6 の「自信を持って答えたか否か」推定時の重み付け結果を見ると、「解答時間」、「眉毛の動き」など、既存手法で有用性が明らかになっている特徴量に次いで、タッチデバイス特有の特徴量である「解答を選んだ時のタッチ時間」が高く重み付けされている。また、図 3, 図 4, 図 5 を見ると、大半の項目で、重み付けなしの推定結果を提案手法の推定結果が上回っている。この結果は MKL による特徴量の最適な重みの学習が機能していることを示している。

さらに、図 10 を見ると、「とても悩んでいる / 全く悩んでいない」状態の推定に関しては、87.1% という高い値を記録している。

以上の結果から、環境を限定している既存手法ほどの精度は出ないものの、タッチデバイスから得られる情報のみで高い品質の学習状態推定を行えることが分かった。

4.3.2. 表情・動作が安定して取得できないことを考慮した手法に関する評価

図 3, 図 4, 図 5 を見てみると、提案手法、全データを使用し、単純に MKL を用いて分類する手法、端末操作情報のみを利用する手法の精度・再現率・Accuracy の間に大きな差は見られない。しかしながら、

傾向の似たユーザでグルーピングを行った時の結果である図 9 を見ると、「自信があるか否か」の推定において、端末操作情報のみを用いた場合と提案手法の Accuracy の間には大きな差があり、統計的にも有意な差であることが確認できた。

「悩んでいるか否か」の推定では、表情・動作情報が大きな意味を持たなかったが、「自信があるか否か」の推定では、表情・動作情報が有効に働くユーザが存在していた。提案手法によってうまく取得できていない表情・動作情報を除去すると同時に、表情・動作情報が有効なユーザのみに対し表情・動作情報を用いることで、精度の向上を図ることに成功したと言える。

提案手法と、「全データ（表情・動作、端末操作情報のすべて）を使用し、MKL を用いて分類する（表情・動作情報の取得状態に応じた使用特徴量の切り替えを行わない）手法」の間には、どの実験においても統計的に有意な差は見られなかった。しかし、「自信があるか否か」の推定においては、学習状態が表出する特徴量の個人差が大きく、より細かなグルーピングを行うことで、提案手法の有効性がより高まる可能性が考えられる。

5. まとめ

本稿では、タブレットデバイスから得られる表情・動作情報と端末操作情報を統合し、学習状態の推定を実現する手法を提案した。タブレット端末特有の情報の利用、MKL の使用、表情・動作取得状態に応じた特徴量選択により、一般的な環境下でも高い精度で推定をおこなえることを示した。

今後の課題としては、より環境依存度の低い手法の提案と精度の向上が挙げられる。選択問題・英語問題以外の学習環境や、閲覧主体型の学習環境に対しても提案手法を適用し、改善を加え、より学習環境に対し柔軟に対応できるシステムを目指したい。また、傾向の似たユーザのグルーピングを自動的に行うなどし、事前データ無しにグルーピング・精度向上ができるシステムの構築にも挑戦していきたい。

参 考 文 献

- [1] 黒川 智哉, 野須 潔, 山崎 清之, “顔画像および生体信号のテンプレートをを用いた e-Learning 学習者の心理状態時系列変化推定の検討”, 映像情報メディア学会誌, Vol.61, No.12, pp.1779-1784, 2007.
- [2] 繁田 亜友子, 小池 武士, 黒川 智哉, 野須 潔, “e-Learning 学習者の顔画像解析による感情推定アルゴリズムの研究”, 電気学会論文誌 C, Vol.129, No.10, pp.1881-1888, 2009.
- [3] 中村和晃, 角所考, 村上正行, 美濃導彦, “e-learning における学習者の顔動作観測に基づく主観的難易度の推定”, 電子情報通信学会論文誌 D, Vol. J93-D, No.5, 2010.
- [4] D. Szafir and B Mutlu, “Pay attention! : designing adaptive agents that monitor and improve user engagement”, CHI 2012, pp.11-20, 2012.
- [5] 中村喜宏, 赤松則夫, 桑原恒夫, 玉城幹介, “操作時間間隔の変動に着目した CAI 学習の行き詰まり検知方法”, 電子情報通信学会論文誌, Vol. J85-D-I, No.1, pp.79-90, 2002.
- [6] A. Kapoor and R. W. Picard, “Multimodal affect recognition in learning environments”, MULTIMEDIA 2005, pp.677-682, 2005.
- [7] 中村 和晃, 角所 考, 美濃 導彦, “e-learning 中の学習者の端末操作と顔画像から推定される主観的難易度を用いた教材評価の試み”, 情報処理学会研究報告 第 5 回 CMS 研究会, pp.41-48, 2007.
- [8] 総務省, “フューチャースクール推進事業”, http://www.soumu.go.jp/main_sosiki/joho_tsusin/kyouiku_joho-ka/future_school.html (2013.1.5 アクセス)
- [9] T.F. Cootes, C.J. Taylor, D. Cooper, and J. Graham, “Active Shape Models -Their Training and Application”, Computer Vision and Image Understanding, Vol. 61, No. 1, pp. 38-59, 1995.
- [10] D. McDuff, A. Karlson, A. Kapoor, A. Roseway and M. Czerwinski, “AffectAura: an intelligent system for emotional memory”, CHI 2012, pp.849-858, 2012.
- [11] “ASMLibrary”, <http://code.google.com/p/asmlibrary/> (2013.1.11 アクセス)
- [12] F.R. Bach, G.R.G. Lanckriet and M.I. Jordan, “Multiple kernel learning, conic duality, and the SMO algorithm”, International Conference on Machine Learning, p.6, 2004.