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

A Research on the Time Sequential Estimation of Emotions of e-Learning Users by the Face and Biometric Information

黒川 智哉[†], 正会員 野須 潔[†], 山崎 清之[†]
Tomoya Kurokawa[†], Kiyoshi Nosu[†] and Kiyoyuki Yamazaki[†]

Abstract There has been a lot of researches on the estimation / characterization of human emotions by using communication channels such as facial expressions. However, most of this research has focused on extracting facial features for some specific emotions at specific situations because the difficulty of the general characterization. We have developed a system that can characterize an emotion of an e-Learning user by analyzing his/her facial expression and biometric signals. The criteria used to classify the eight emotions were based upon a time sequential subjective evaluation of emotions as well as a time sequential analysis of facial expressions and biometric signals. The average coincidence ratio between these discriminated emotions by using the criteria of emotion diagnosis and the time sequential subjectively evaluated emotions for ten e-Learning examinees was 71%. When only the facial expressions were observed, the coincidence ratio was 66%. This suggests the multi-modal emotion diagnosis is effectiveness for estimating an e-Learning user's emotions.

キーワード：顔画像解析, 表情解析, 生体信号, 心理状態推定, e-Learning

1. ま え が き

顔は重要なコミュニケーションメディアといわれており、われわれは顔の表情から相手の心理状態を推定し、読取っている。顔の画像から人の心理状態を推定する研究は進められている^{1)~6)}。しかし、これらの研究は意識的に作った表情を対象としたものや、静止画像間の比較によるものが多く、推定する心理状態も「怒り」「悲しみ」といった基本心理状態への分類を対象としている。また、音声と表情などマルチモーダルによる心理状態推定の研究も行われているが、表情と生体信号から複数の心理状態推定を行った報告は少ない。

本研究の目標は、電子教材使用状態に対応した心理状態推定基準データをあらかじめ用意し、学習者の顔画像および生体情報の変化から心理状態を推定することである。本研究の特徴は以下の点にある。

- (1) e-Learning学習者の自然な使用状況における表情、生体信号の時系列変化を基に心理状態推定基準(テンプレート)を作成している。
- (2) e-Learning使用状況に基づいた四つのSD (Semantic Differential) 心理状態対に対応した心理状態推定基準

を作成。

- (3) テンプレートマッチング法とマハラノビスの距離を用いたアルゴリズム。

具体的には従来研究から、心理状態判断に有効であると考えられる特徴点【右唇・左唇・右内眉・右外眉・左内眉・左外眉】の6種と、【脈拍数・呼吸数・皮膚温度】の3種の生体信号の心理状態推定基準データを求めた。

この基準データを基に、顔画像および生体信号から時々刻々と変化する心理状態を推定した。顔画像および生体信号を単独で用いた場合と、併用した場合の推定精度を比較した。また、時系列変化データをリアルタイムで処理し、心理状態を推定するシステムの実現性も検討した。

2章では、心理状態推定基準データの抽出手順を述べる。3章および4章では、各々顔画像および生体信号の解析手順を述べる。5章では、顔画像および生体信号の心理状態推定の基準値の算出結果を述べる。6章では、心理状態推定基準(テンプレート)とマハラノビスの距離による心理状態推定結果を述べる。7章ではリアルタイム心理状態推定の可能性を検討し、8章では本研究で得られた事項をまとめる。

2. 心理状態推定基準(テンプレート)の抽出手順

心理状態推定基準データの抽出を以下の手順で行った。

- (1) 被験者は実際に提供されているe-Learning学習用教材コンテンツを用いて学習を行う。教材の内容は情報

2006年9月13日受付, 2007年8月28日最終受付, 2007年9月13日採録

[†] 東海大学 開発工学部 感性デザイン学科

(〒410-0395 静岡県沼津市西野317, TEL 055-968-1211)

[†] School of High Technology for Human Welfare, Tokai University
(Nishino 317, Numazu, 410-0395, Japan)

図1 測定中の様子
Experiment scenery.

表1 学習者評価心理状態
Evaluated psychological states.

- 第1心理状態対： 簡単－ 難しい
第2心理状態対： 面白い－ つまらない
第3心理状態対： 理解できた－ 理解できなかった
第4心理状態対： 集中している－ 飽きている

通信に関するもので文章やアニメーションも含んだものである。この教材を5分間、画面に向かってマウスを操作し、自習形式で学習してもらった。

- (2) 学習中は顔画像をディスプレイ上部より、また学習端末のディスプレイを被験者の後方より、合計2台の家庭用ビデオカメラを用いて撮影した。同時に被験者の脈拍数、呼吸数、皮膚温度を生体信号測定装置で測定した。図1はその様子を示したものである。なお、被験者に追跡位置検出用のマーカの装着は行わなかった。また、学習者ごとに学習前の1分間、何も提示せず学習者が平静状態であった状態も測定し、この1分間の平均値をその学習者の基準となる平静時のデータとした。
- (3) 測定終了後直ちに、被験者に撮影した学習端末の画面を提示しながら、30秒ごとに心理状態を回答してもらった。評価心理状態は表1に示す心理状態対で示した。これらの項目について「1」～「3」の3段階で評価してもらい、「1」を表1の左側の心理状態、「2」を中間状態、「3」を右側の心理状態とした。選択された心理状態を、評価が「2」の場合を除き、その時点の代表複合心理状態とした。したがって、ある時点では最大四つの代表心理状態が検出される。

以上の述べた手順で5分間の調査結果から30秒ごとの心理状態の評価データを取得した。同時に取得した表情・生体信号と照らし合わせて、心理状態推移とその傾向や動きの解析を行った。顔画像および生体信号の解析手順は3章、4章で述べる。

3. 顔画像解析手順

顔画像解析手順を以下に示す。

- (1) 顔画像情報はe-Learning用端末のディスプレイ上部に設置した家庭用ビデオカメラで撮影し、パソコンに動画ファイルとして取込み、画像解析を行った。画面の分解能は640×480画素である。特徴点の追跡は、特徴点付近の画面フレーム間の相関を計算して追跡を行う2次元画像解析ソフト（デジモ社製2D-PTV）により行った。
- (2) 顔画像から心理状態の推定用の基準点および特徴点として以下の8点を使用した。

A：基準点

- 0-1) 頭頂 0-2) あご

B：特徴点

- 1-1) 右唇 1-2) 左唇
2-1) 右内眉 2-2) 左内眉
3-1) 右外眉 3-2) 左外眉

- (3) これら8カ所の点座標の2次元移動軌跡追跡を15フレーム/秒、1点の追跡範囲を20画素平方に設定し、手で初期位置設定を行い、その後自動で追跡を行った。図2は顔特徴点追跡画面の一例である。

- (4) こうして得られた2次元座標系データは640×480画素における絶対座標で提供されるが、そのままでは頭のぶれや、画面内の顔の占める割合が異なった場合などの状況により、異なった数値が与えられることになってしまう。そういったノイズを可能な限り除去するため、【頭頂－あご】の距離を1辺の長さとし、顔領域を内包する正方形を図3のように作成し、各特徴点の座標の相対座標化を行った。

また、各特徴点の相対座標を算出は、図3の左頂点を原点とし、図4のように顔特徴点の相対座標を表した。

こうして得られた各特徴点の相対座標に対し、基準とな

図2 顔特徴点の処理画像
Image processing of feature points.

図3 交点座標算出概念図
Intersection point.

図4 相対座標算出概念図
Calculation of relative coordinates.

図5 右唇点変化率測定例

Example of a measured change rate of a facial point (right corner of a mouth).

る平静時の位置状態との差を求め、基準線分長(【頭頂-あご】の距離)で除算することにより、相対座標内での特徴点ごとの変化率を求めた。

図5は、右唇の特徴点位置変化率の測定例を示す。

4. 生体信号解析手順

生体信号は、学習者への装着しやすさとその有効性⁷⁾より、【脈拍数・呼吸数・皮膚温度】の3種類を用いた。それぞれ脈拍数は左人差し指に脈派ピックアップを、呼吸数は右鼻下に呼吸プローブを、皮膚温度は左中指にサーミスタ温度プローブを設置し、生体信号測定装置(日本光電製LEG-1000)にて一括で測定した。それぞれの生体信号取得設定値、測定データ範囲を表2に示す。

これらにより得られた生体信号データを10秒間隔でサンプリングし、基準とした平静時の各生体信号からの変化率を求めた。10秒間隔を選択したのは、ノイズを消去した状

表2 生体信号取得設定と測定データ範囲

Biometric measurement set-up and the ranges of measured data.

項目	HI CUT	LO CUT	本検討の実測データ範囲
脈拍数	50Hz	0.3Hz(0.5s)	59~87(回/分)
皮膚温度	100Hz	0.05Hz(3s)	28~35(°C)
呼吸数	0.5Hz	—	4~26(回/分)

HI CUT:高域遮断周波数

LO CUT:低域遮断周波数, () 内は時定数

10秒毎の脈拍数データ

図6 脈拍数の変化率測定例

Example of a biometric signal data.

態で特徴点を捉えるのにもっとも適していたからである。図6は、脈拍数の変化率測定例を示す。

5. 心理状態推定基準値の算出

図7に示す手順で、男子大学生10名の被験者により心理状態推定基準(テンプレート)を求めた。3章、4章で述べた手順で得た顔画像および生体信号変化を30秒間隔で平均した時系列データと、表1で述べた【第1対:簡単-難しい、第2対:面白い-つまらない、第3対:理解できた-理解できなかった、第4対:集中している-飽きている】の代表心理状態の時系列変化に対応させた。

第1心理状態対~第4心理状態対は独立していると仮定した。例えば、ある時点の複合心理状態が【難しい・面白い・理解できた・集中している】と主観評価された場合、その時点の顔画像および生体信号データは【難しい】・【面白い】・【理解できた】・【集中している】の4種類の心理状態情報を同時に所持していることとした。

このようにして、8種類の心理状態ごとに対応する顔特徴点および生体信号変化率のデータを蓄積し、顔特徴点6種と生体信号3種の9種のデータごとに集計し、10名のデータから各心理状態に対する変化率の平均とその標準偏差、サンプル数を求め、これらの値を心理状態推定の基準値とした。図8は参考として、心理状態推定の基準値を平均値の95%信頼区間として求めた場合の例(第1心理状態対:右唇)である。また、全基準値を付表1に示す。

6. マハラノビスの距離による心理状態推定

図9に示す手順で、心理状態推定を行った。5章で述べた基準値(テンプレート)を用いて心理状態を推定するアルゴリズムには、マハラノビスの距離⁸⁾を採用した。マハラノビスの距離は標本分布が正規分布であると仮定して標本の

図7 心理状態推定基準値の算出手順
Calculation procedure for the criteria of psychological state estimation.

図8 心理状態推定基準95%信頼区間(第1対:右唇)
Criteria of position change rates for a right corner of a mouth.

表3 代表心理状態と推定心理状態の一致率
Coincidence rates between the representative psychological states and the estimated psychological states.

心理状態\データ	顔画像	生体信号	併用
簡単	75%	53%	81%
難しい	71%	43%	57%
面白い	0%	0%	0%
つまらない	60%	67%	73%
理解できた	64%	51%	69%
理解できない	75%	50%	75%
集中している	56%	44%	56%
あきっている	67%	54%	71%
総合一致率	67%	53%	71%

平均との誤差を標準偏差で割ったものであり、1次元だけでなく多次元に拡張することができる。以下にその一般式を示す。

$$D(\xi - \mu)^2 = (\xi - \mu)^t S^{-1} (\xi - \mu)$$

$$D(\xi - \mu)^2: \text{マハラノビスの距離}, \quad S: \text{共分散行列}$$

ξ : 標本値ベクトル, μ : 母集団の平均ベクトル

5名の被験者男子大学生が参加した心理状態推定アルゴリズムの評価実験を行い、被験者毎に顔画像および生体信

号を30秒間隔で平均した時系列データを得た。心理状態ごとに顔画像のみの場合の S は12次元行列、生体信号のみの場合は3次元行列、顔画像と生体信号を併用した場合は15次元行列として、マハラノビスの距離を求めた。こうして得られたマハラノビスの距離を第1心理状態対～第4心理状態対各々について、短かい距離の心理状態を抽出し、その時点の四つの心理状態の組合せからなる複合推定心理状態とした。この複合推定心理状態と、2章で述べた心理状態主観評価方法で得られたデータを照らし合わせ、各心理状

図9 代表心理状態一致率の算出手順

Calculation procedure for coincidence rates of psychological state estimation.

態対に対して一致度合いを算出した。また、心理状態主観評価で、心理状態対のいずれでもない評価(主観評価値2, 中間状態)された場合、時系列のデータから該当した心理状態対を除外して計算を行った。

表3は、心理状態ごとの一致率を示している。また、全一致心理状態数と評価に対する全サンプル数の比である総合一致率も示した。ここで、「面白い」の心理状態の主観評価サンプル数は、他の心理状態に比べて極端に少なく、一つしかなかった。より娯楽的な要素のある教材を使った場合、「面白い」の心理状態も相当量のサンプルが得られ、他の心理状態並みの一致率が得られると思われる。「面白い」の心理状態を除けば、表3が示す通り、顔画像および生体信号の個々のデータのみで推定を試みた場合より、それらを併用した場合の方がより推定精度が高く、時系列に沿った心理状態推定において有用であるということを検証することができた。

今回、評価対象から除外した心理状態対のいずれにも該当しない状態は、マハラノビスの距離の増大、または距離変化測度の増大として現れると予想される。この状態を除外することは、図9の手順において、マハラノビスの距離の演算結果を利用したフィルタリングにより実現される。

7. リアルタイム推定の可能性の検討

リアルタイム推定を行うには以下の演算処理をリアルタ

イムで行う必要がある。

- (1) 顔画像特徴点の追跡
- (2) 顔画像特徴点の位置変化率の演算
- (3) 生体信号の検出
- (4) 生体信号の変化率の演算
- (5) テンプレートマッチング法による心理状態推定演算

(2)～(5)は、本研究においてリアルタイム演算処理が実現されている。(1)は、15フレーム/秒の演算速度の画像処理を実現している。しかし、実験で撮影した画像の特徴点追跡を自動で行った時、手動で指定した追跡位置からずれることがあった。今回の対象映像を用いた場合、自動追跡にて、5画素以上ずれる割合は58%～89%の範囲であった。すべての特徴点の自動追跡と手動追跡の一致率が90%近くなるよう、顔撮影時の照明条件や人物背景のセットなどの工夫、また自動追跡データに基づく位置変化率が設定値を超えた場合に、自動で初期値に戻るなどの追跡外れ対策アルゴリズム導入が必要になる。リアルタイム推定が可能になれば、学習を提示する電子教材を心理状態の変化に応じて、変更することが可能となる。

8. む す び

e-Learning学習者の心理状態評価に適した心理状態推定基準(テンプレート)を用いて時間的に変化する心理状態を推定する方法を開発した。得られた結果を要約すると以下のようになる。

- (1) 顔画像による心理状態推定精度は67%であった。
- (2) 生体信号による心理状態推定精度は53%であった。
- (3) 顔画像と生体信号を併用しての心理状態推定精度は、71%に向上した。

これらの値は、【驚き・納得・困惑・面白い】といった一般的な心理状態に対して、時系列の各時点でもっとも強かった物を代表心理状態として推定した同様の実験で得られた推定精度53%より大幅に改善されている⁹⁾。これは、推定する代表心理状態を一般的な心理状態ではなくe-Learningに対応した心理状態としたこと、および複数のSD (Semantic Differential) 心理状態対に対して、各時点の主観評価を用いて、心理状態推定基準(テンプレート)を作成したことによると考えられる。

今後、リアルタイムでの顔画像軌跡追跡の精度向上、心理状態対のいずれでもない状態の判別方法も含んだ心理状態推定の精度の向上、これらの検討を生かしたリアルタイム心理状態推定システムの開発を行う予定である。

[文 献]

- 1) P. Ekman, W.V. Friesen: "Facial Action Coding System: A Technique for Measurement of Facial Movement", Consulting Psychologists Press Palo Alto, California (1978)
- 2) 吉川, 桜井ほか: "目・口領域のオプティカルフロー推定による「笑い」認識の試み(顔とジェスチャの認識)", 信学技報HIP, ヒューマン情報処理, 103, 455, pp.79-84 (Nov. 2003)

付表1 心理状態推定基準値
Criteria of psychological state estimation.

データ属性\心理状態		簡単	難しい	面白い	つまらない	理解できた	理解できない	集中	飽きている
顔画面	右唇	x 変移平均	0.14	0.16	0.14	0.14	0.12	0.16	0.16
		座 標準偏差	0.14	0.15	0.13	0.1	0.13	0.16	0.15
		標 サンプル数	28	33	31	27	41	13	31
		y 変移平均	0.02	-0.03	0.03	-0.04	0.02	-0.02	0.03
		座 標準偏差	0.15	0.12	0.17	0.12	0.1	0.09	0.22
		標 サンプル数	28	33	31	27	41	13	31
	左唇	x 変移平均	0.15	0.16	0.09	0.1	0.13	0.06	0.11
		座 標準偏差	0.22	0.16	0.19	0.16	0.13	0.19	0.17
		標 サンプル数	28	33	31	27	41	13	31
		y 変移平均	0.05	0	0.02	-0.01	0.01	-0.03	0.04
		座 標準偏差	0.12	0.06	0.15	0.11	0.09	0.11	0.13
		標 サンプル数	28	33	31	27	41	13	31
	右内眉	x 変移平均	0.19	0.12	0.14	0.13	0.15	0.18	0.14
		座 標準偏差	0.28	0.23	0.22	0.19	0.17	0.18	0.25
		標 サンプル数	28	33	31	27	41	13	31
		y 変移平均	0.04	0.03	0.03	0.05	0.03	0.05	0.05
		座 標準偏差	0.03	0.04	0.06	0.03	0.05	0.07	0.16
		標 サンプル数	28	33	31	27	41	13	31
	左内眉	x 変移平均	0.17	0.1	0.19	0.19	0.19	0.18	0.19
		座 標準偏差	0.36	0.17	0.29	0.28	0.27	0.14	0.37
		標 サンプル数	28	33	31	27	41	13	31
		y 変移平均	0.05	0.04	0.03	0.04	0.04	0.06	0.04
		座 標準偏差	0.03	0.03	0.05	0.03	0.11	0.08	0.07
		標 サンプル数	28	33	31	27	41	13	31
	右外眉	x 変移平均	0.08	0.07	0.08	0.07	0.09	0.09	0.08
		座 標準偏差	0.08	0.06	0.09	0.09	0.09	0.08	0.09
		標 サンプル数	28	33	31	27	41	13	31
		y 変移平均	0.06	0.04	0.03	0.04	0.03	0.05	0.04
		座 標準偏差	0.03	0.03	0.05	0.04	0.04	0.05	0.05
		標 サンプル数	28	33	31	27	41	13	31
	左外眉	x 変移平均	0.01	0.03	0.02	0.04	0.04	0.04	0.02
		座 標準偏差	0.1	0.06	0.08	0.07	0.09	0.09	0.09
		標 サンプル数	28	33	31	27	41	13	31
		y 変移平均	0.05	0.05	0.04	0.05	0.05	0.05	0.05
		座 標準偏差	0.04	0.03	0.05	0.03	0.04	0.04	0.06
		標 サンプル数	28	33	31	27	41	13	31
生体信号	脈拍数	変移平均	-0.013	0.035	0.003	0.008	-0.015	0.025	-0.003
		標準偏差	0.03	0.05	0.05	0.03	0.02	0.08	0.06
		サンプル数	28	33	31	27	41	13	31
	呼吸数	変移平均	-0.027	-0.005	0.016	0.021	-0.002	0.035	0.014
		標準偏差	0.04	0.03	0.08	0.04	0.03	0.04	0.07
		サンプル数	28	33	31	27	41	13	31
	皮膚温	変移平均	0	0.013	0.001	0.015	0.003	0.022	0.001
		標準偏差	0.04	0.06	0.04	0.07	0.08	0.05	0.06
		サンプル数	28	33	31	27	41	13	31

- Y.I. Tian, T. Kanade, J.F. Cohn: "Recognizing Action Units for Facial Expression Analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, 23, 2, pp.97-115 (Feb. 2001)
- 長野, 任ほか: "顔特徴を用いた表情の認識 (福祉と言語処理, 一般)", 信学技報WIT, 福祉情報工学, 103, 116, pp.61-66 (June 2003)
- 佐々木, 目加田ほか: "パーツの特徴に基づく動きパターンを用いた表情分類法に関する検討", 映像学技報, 24, 60 pp.33-38 (Oct. 2000)
- Y. Yoshitomi et al.: "Effect of Sensor Fusion Recognition of Emotional States Using Voice", Face action Thermal of Image of Face, IEEE International Workshop of Robot and Human Interactive Communication 2000, pp.178-183
- J.L. Andreassi: "心理生理学 ヒトの行動と生理的反応", ナカニシヤ出版, 東京 (1985)
- 涌井良幸, 涌井貞美: "図解でわかる多変量解析", 第六章, 日本実業出版社 (2001)
- 黒川, 奥田, 野須: "顔画像および生体信号によるe-Learning学習者の心理状態時系列変化推定の一検討", 日本人間工学会誌, 42, 6, (Dec. 2006)

黒川 智哉 2005年, 東海大学開発工学部情報通信工学科卒業。2007年, 同研究科修士課程修了。NTTソフトにてソフトウェア開発に従事。

野須 潔 1977年, 東京大学大学院工学系研究科博士課程修了, 1993年, 筑波大学大学院経営政策学研究科修士課程修了。1977年, 日本電信電話公社 (現NTT) 入社。2001年, 東海大学開発工学部。2004年より, 同大学開発工学部感性デザイン学科教授。ヒューマンインタフェース, 情報環境のユーザビリティ評価の研究に従事。工学博士, 経営学修士。正会員。

山崎 清之 1989年, 東京大学大学院医学系研究科博士課程修了。同年, 東海大学開発技術研究所講師。1995年, 同大学医用生体工学科助教授, 2004年より, 同大学開発工学部医用生体工学科教授。医用生体工学, 特に脳波, ERPの人間工学的応用に関する研究に従事。医学博士。