マルチモーダル生体信号情報による感情認識に関する一考察*

高橋和彦**

1. はじめに

人どうしのコミュニケーションでは感情等のノンバーバル情報が重要な役割を果たしており，人とコンピュータ，あるいはコンピュータを介した人どうしのコミュニケーションにおいても，ノンバーバル情報の利用の有効性が指摘されている．また，近年，人とロボットのインタラクションに関する研究が盛んに行われており，ロボットやコンピュータによる感情の表出や理解等も着目されてきている¹⁾²⁾．

感情認識は興味深いが困難な課題である．人は感情的な発話を約60%，表情を約70～98%の認識率で認識している³⁾．一方，コンピュータによる感情認識では，特定の個人や限定された状況・環境で収集されたデータ，あるいは意識的に表出された音声や表情が取り扱われることが多く，対象とする感情カテゴリも様々であるが，感情音声で50～60%⁴⁾⁵⁾，表情で80～90%⁶⁾⁷⁾の認識率であり，音声と表情のマルチモーダル化により更に高い認識率が得られる事⁸⁾⁹⁾も報告されている．これらは「行動的尺度」である感情に伴う人の外的反応の計測に基づく感情認識の研究であるが，この他に感情を評価・測定する手法としてアンケート等で得られるデータを多変量解析法により解析して感情の潜在因子や次元を求める「心理的尺度」に基づく方法や，感情喚起時の生理的・内的な反応を対象として心理生理学的測定を行う「生理的尺度」による方法がある．生理的尺度を用いる方法では，心理的尺度や行動的尺度を用いる場合に比べて故意，あるいは無意識の嘘や隠し立てによる影響が小さく，またその反応は各種のセンサを利用することによって定量的に計測可能であり，更に近年のウェアラブルコンピューティング技術の進展が計測を容易にしている¹⁰⁾．生理的尺度によるコンピュータでの感情認識は，脈拍や心拍，皮

図1 感情認識実験システムの概略.

Fig. 1 Experimental setup in emotion recognition.

膚電気抵抗等の自律神経系の反応に着目した研究¹¹⁾¹²⁾や脳波からの推定方法の研究開発¹³⁾¹⁴⁾¹⁵⁾等が報告されており，筆者らも知的マン・マシンインタフェースの開発の一環として生体信号からの感情認識の研究¹⁶⁾¹⁷⁾を行っている．

本研究では，脳波，脈拍，皮膚電気抵抗によるマルチモーダルな生体信号からサポートベクトルマシン (SVM)¹⁸⁾を用いて喜び，怒り，哀しみ，恐れ，リラックスの5感情を対象とする感情認識システムを提案し，感情認識システムの入力に用いるマルチモーダルな生体信号の効果について検討する．

2. 生体信号計測システム

図1に感情刺激時の生体信号を計測する実験システムの概要を示す．実験システムは，脳波，脈拍，皮膚電気抵抗を計測する3種類のセンサとデータ収集用のパーソナルコンピュータ (DELL Precision 420)，被験者に感情刺激のためのサンプルを提示するパーソナルコンピュータ (DELL Precision 420) で構成されている．

脳波 (Electroencephalogram) は図2(a)に示す簡易脳波インタフェース The CyberlinkTM (Brain Actuated Technologies 社製)¹⁹⁾によって計測される．このインタフェースは，ヘッドバンドに取り付けられた3つの乾式電

*受付：2005年1月17日 受理：2005年5月9日

**同志社大学

Doshisha University

図 2 センサー ((a) 脳波センサ, (b) 脈拍センサ, (c) 皮膚電気抵抗センサ).

Fig. 2 Overview of sensors ((a) EEG sensor, (b) Pulse sensor, (c) GSR sensor).

極により被験者の額部 (10-20 法電極位置 F_{p1} , F_{p2}) から生体信号を非侵襲で取得し, インタフェースユニットにおいて周波数分析により生体信号を脳波 (EEG), 眼電位 (EOG), 筋電位 (EMG) の各信号に分離して RS-232C 経由でパーソナルコンピュータに送信する.

脈拍 (Pulse) は図 2(b) に示すセンサにより計測される. このセンサは赤外発光ダイオードと光検出器で構成され, 被験者の耳朶に装着して耳朶における血流量の変化を赤外光の透過量の変化として検出する. 結果は AD 変換器 (インタフェース社製 PCI-3133) 経由でパーソナルコンピュータに取り込まれる.

皮膚電気抵抗 (Galvanic Skin Resistance) は図 2(c) に示すセンサにより計測される. このセンサは 2 つの電極により指先における電気抵抗の変化を電流の変化として検出するもので, 結果は AD 変換器経由でパーソナルコンピュータに取り込まれる. 電極は被験者が自然な形で接触できるようにマウスのボタン部分に装着している.

3. 感情刺激サンプルによる生体信号計測実験

生体信号計測実験では, 被験者に 3 つのセンサを装着してもらい, ビデオコンテンツを用いた感情刺激を行って生体信号を収集する. 対象とする感情は, 喜び (Joy), 怒り (Anger), 哀しみ (Sadness), 恐怖 (Fear), リラックス (Relax) の 5 感情である. 感情刺激に用いるビデオコンテンツは, テレビ放映されている 30 秒の商業映像を使用した. はじめに実験者 2 名で各感情を喚起すると思われるビデオを感情毎に 5 編ずつ収集した. 次に, 生体信号計測実験の被験者にならない人物 6 名に各ビデオを視聴してもらい, 図 3 のアンケート (# は映像番号, XX は感情カテゴリを示す) により視聴時の感情に

映像#から XX の感情を感じましたか?

- 1 感じなかった
- 2 あまり感じなかった
- 3 どちらでもない
- 4 やや感じた
- 5 強く感じた

図 3 ビデオ評価アンケート.

Fig. 3 Questionnaire for evaluating audio-video film.

表 1 感情刺激のためのビデオコンテンツ.

Tab. 1 Audio-video contents for emotion stimulation.

Joy	ファーストフードでのホームパーティの様子 仮装した芸人によるレジャー施設の紹介
Anger	煙草のボイ捨て禁止 戦争犠牲者と反戦アピール
Sadness	老人の語りかけによる葬祭施設の紹介 子供への愛情表現の重要性を訴える広告
Fear	地球温暖化による環境破壊に対する警鐘 伝染病予防ワクチンのための募金広告
Relax	川の保全を訴える環境映像 子犬達がじゃれあう映像

ついて点数付けによる評価を行い, 感情毎に上位のスコアを獲得した 2 編のビデオを生体信号計測実験時の刺激サンプル (表 1) として採用した. 計測実験においてこれらのビデオコンテンツは, HTML で記述した Web アプリケーションにより刺激提示用パーソナルコンピュータ上で Internet ExplorerTM と Real One Player 2TM を用いて被験者に提示される.

生体信号計測実験は, 照明, 音, 空調を一定とした実験室において 12 名の被験者 (日本人成人男子, 21-25 歳) を対象にして行った. 被験者には予め実験の目的と内容を説明し, データ収集の同意を得ている. また, 事前に生活パターンについての聞き取りを行い, 疲労や空腹でない健康な状態で行えるよう実験の時間帯を調整した. データの収集は被験者当たり各感情毎に 2 回ずつ行った. 図 4 に計測された生体信号の一例を示すが, 時系列波形の感情カテゴリ間での差異を見つけ出すことは難しく, また同一の感情カテゴリであっても信号の個人差は大きい. 尚, 図 5 のアンケート (# は映像番号を示す) により実験中の被験者の各感情の喚起状況についての評価もあわせて行った. 表 2 は評価結果を示したものであり, 恐れについては哀しみと認識され易い傾向があるが, 被験者はビデオコンテンツによって各感情を喚起できていたことがわかる.

図 4 生体信号の例 (上: 脳波, 中: 脈拍, 下: 皮膚電気抵抗, 左から右: 喜び, 怒り, 哀しみ, 恐れ, リラックス).

Fig. 4 Example of bio-potential signals (top: EEG, middle: Pulse, bottom: GSR, left to right: joy, anger, sadness, fear, and relax).

表 2 感情刺激の評価結果 [%].

Tab. 2 Emotion recognition results by human subjects [%].

In \ Out	Joy	Anger	Sadness	Fear	Relax	No emotion
Joy	83.4	8.3	0	0	0	8.3
Anger	8.3	75.0	0	0	0	16.7
Sadness	0	0	91.7	0	0	8.3
Fear	0	0	41.7	58.3	8.3	0
Relax	25.0	0	0	0	75.0	0

映像#からどの感情を感じましたか？

- ・ 喜び
- ・ 怒り
- ・ 哀しみ
- ・ 恐れ
- ・ リラックス
- ・ 分からない

図 5 感情評価アンケート.

Fig. 5 Questionnaire for evaluating emotion.

コンピュータによる感情認識を行うため, 収集した生体信号から特徴ベクトルを構成する. 本研究では, Affective Computing の先行研究¹¹⁾ を参考に, 以下の統計量の特徴量として使用した.

$$\mu_X = \frac{1}{T} \sum_{t=1}^T X(t) \quad (1)$$

$$\sigma_X = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (X(t) - \mu_X)^2} \quad (2)$$

$$\delta_X = \frac{1}{T-1} \sum_{t=1}^{T-1} |X(t+1) - X(t)| \quad (3)$$

$$\bar{\delta}_X = \frac{\delta_X}{\sigma_X} \quad (4)$$

$$\gamma_X = \frac{1}{T-2} \sum_{t=1}^{T-2} |X(t+2) - X(t)| \quad (5)$$

$$\bar{\gamma}_X = \frac{\gamma_X}{\sigma_X} \quad (6)$$

ここで, μ_X は原信号の平均, σ_X は原信号の標準偏差, δ_X , γ_X はそれぞれ 1 次及び 2 次精度での原信号のグラジェント, $\bar{\delta}_X$, $\bar{\gamma}_X$ は正規化された信号のグラジェントを表し, X は計測された生体信号, t はサンプリング数, T はサンプリング総数である. 尚, 特徴量の計算の前に, ノイズを除去するためデジタルフィルタによるローパスフィルタ処理 (遮断周波数 脳波:30Hz, 脈拍及び皮膚電気抵抗:10Hz) を行っている. これらの特徴量から次の特徴ベクトル $\mathbf{x}$ を定義する.

$$\mathbf{x}^T = [\mu_e \quad \sigma_e \quad \delta_e \quad \bar{\delta}_e \quad \gamma_e \quad \bar{\gamma}_e \\ \mu_p \quad \sigma_p \quad \delta_p \quad \bar{\delta}_p \quad \gamma_p \quad \bar{\gamma}_p]$$

$$[\mu_s \quad \sigma_s \quad \delta_s \quad \bar{\delta}_s \quad \gamma_s \quad \bar{\gamma}_s] \quad (7)$$

ここで e は脳波, p は脈拍, s は皮膚電気抵抗を示す.

4. SVM による感情認識実験

4-1. 感情認識システム

個人独立型の感情認識を目的とするため, 収集した複数の被験者の生体信号データを用いた学習により感情認識システムを構築する. 感情認識システムはパターン認識用 SVM で構成される. SVM は, 入力となる対象の特徴ベクトルをカーネル関数により高次元の特徴空間に写像し, その特徴空間において 2 つのクラスを分類する最適な判別関数 (超平面) をマージンと呼ばれる学習サンプルの境界と超平面との距離を最大化する二次計画問題の解として求め, この判別関数を用いてデータの分類を行う. 従って SVM は, 基本的には注目するクラスと他のクラスを分離するという 2 クラス分類問題を解く手法であるため, 多クラス分類問題へ適用するためには複数の SVM を組み合わせて認識器を構成する必要がある, 種々の方式が提案されている. 本研究では, One-vs-All 型²⁰⁾ の構成法により感情認識システムを構築する. 具体的には, 図 6 に示すように, 対象とする感情とそれ以外の感情に分類する SVM を各感情毎に構成して学習を行う. 認識時には全ての SVM に同時に特徴ベクトルを入力し, 最大値選択機において各 SVM からの判別関数の出力を比較し, 出力値の大きい SVM の感情カテゴリを最終的な認識結果として選択する. 本実験における SVM のカーネル関数とパラメータの設定は, 種々の組み合わせ実験を行って学習サンプルについての認識結果が 100% となるように試行錯誤的に決定している. 本実験では, カーネル関数をガウス関数 $K(z_1, z_2) = \exp(-\lambda \|z_1 - z_2\|)$ (λ は学習サンプルの分散で決まるパラメータ) とした. マージンに関するパラメータは大きくすることで学習データに対する認識率が向上するため, 大きな値を取ることが推奨されているが, 本実験では 10^{10} に設定した. 学習及び認識の評価は Leave-one-out 法²¹⁾ により行った.

4-2. 感情認識実験結果

表 3, 4, 5 は 1 つのセンサ情報のみで特徴ベクトル (6 次元) を構成して学習と認識を行った場合, 表 6, 7, 8 は 2 種類のセンサ情報を組み合わせて特徴ベクトル (12 次元) を構成して学習と認識を行った場合, 表 9 は全てのセンサ情報を用いて特徴ベクトル (18 次元) を構成して学習と認識を行った場合の結果である.

図 6 SVM による感情認識システム

Fig. 6 Emotion recognition system using SVM.

表 3 認識結果 (脳波) [%].

Tab. 3 Confusion matrix (EEG) [%].

In \ Out	Joy	Anger	Sadness	Fear	Relax
Joy	58.3	16.7	16.7	8.3	0
Anger	25.0	16.7	8.3	16.7	33.3
Sadness	33.3	0	8.3	8.3	50.0
Fear	16.7	0	25.0	41.6	16.7
Relax	16.7	0	25.0	8.3	50.0

表 4 認識結果 (脈拍) [%].

Tab. 4 Confusion matrix (pulse) [%].

In \ Out	Joy	Anger	Sadness	Fear	Relax
Joy	16.7	8.3	50.0	8.3	16.7
Anger	25.0	16.7	8.3	16.7	33.3
Sadness	16.7	0	58.3	0	25.0
Fear	25.0	0	41.6	16.7	16.7
Relax	16.7	0	50.0	8.3	25.0

表 5 認識結果 (皮膚電気抵抗) [%].

Tab. 5 Confusion matrix (GSR) [%].

In \ Out	Joy	Anger	Sadness	Fear	Relax
Joy	0	33.3	16.7	33.3	16.7
Anger	0	33.3	25.0	25.0	16.7
Sadness	8.3	33.4	25.0	25.0	8.3
Fear	8.3	25.0	41.7	8.3	16.7
Relax	8.3	25.0	33.4	25.0	8.3

● 1 センサの場合: 脳波によって得られる認識率は平均 35% (表の対角成分の平均値) であり, 怒りと哀しみでの認識率が低い. 脳波の周波数成分による 4 感情の認識を行った既報¹⁶⁾ と同程度の認識率である. 脈拍や皮膚電気抵抗では怒りと哀しみの認識率が脳波よりも比較的高いが誤認識の割合も大きく, これらのセンサ情報のみでの感情認識は難しい.

● 2 センサの場合: 脳波と組み合わせることで脈拍ま

表 6 認識結果 (脳波, 脈拍) [%].

Tab. 6 Confusion matrix (EEG, pulse) [%].

In \ Out	Joy	Anger	Sadness	Fear	Relax
Joy	33.3	16.7	25.0	16.7	8.3
Anger	25.0	33.3	25.0	0	16.7
Sadness	33.3	0	8.4	25.0	33.3
Fear	25.0	8.3	8.3	41.7	16.7
Relax	33.3	16.7	25.0	0	25.0

表 7 認識結果 (脳波, 皮膚電気抵抗) [%].

Tab. 7 Confusion matrix (EEG, GSR) [%].

In \ Out	Joy	Anger	Sadness	Fear	Relax
Joy	33.3	8.3	16.7	16.7	25.0
Anger	25.0	25.0	25.0	8.3	16.7
Sadness	25.0	8.3	25.0	16.7	25.0
Fear	8.3	16.7	8.3	41.7	25.0
Relax	16.7	25.0	16.7	8.3	33.3

表 8 認識結果 (脈拍, 皮膚電気抵抗) [%].

Tab. 8 Confusion matrix (pulse, GSR) [%].

In \ Out	Joy	Anger	Sadness	Fear	Relax
Joy	25.0	25.0	25.0	8.3	16.7
Anger	58.4	8.3	8.3	16.7	8.3
Sadness	25.0	8.3	16.7	25.0	25.0
Fear	50.0	8.3	16.7	16.7	8.3
Relax	50.0	8.3	25.0	0	16.7

表 9 認識結果 (脳波, 脈拍, 皮膚電気抵抗) [%].

Tab. 9 Confusion matrix (EEG, pulse, GSR) [%].

In \ Out	Joy	Anger	Sadness	Fear	Relax
Joy	58.4	8.3	16.7	8.3	8.3
Anger	0	49.9	16.7	16.7	16.7
Sadness	25.0	0	33.3	16.7	25.0
Fear	8.3	16.7	16.7	41.6	16.7
Relax	0	25.0	25.0	16.7	33.3

たは皮膚電気抵抗のみの場合より認識率が改善されており, 脳波と他のセンサ情報との併用が有効であることがわかるが, 誤認が増加してしまう感情カテゴリも一部存在する.

- 3 センサの場合: 平均 43% の認識率であり, 最も高い認識率が得られている. 全てのセンサ情報を使うことで怒り, 哀しみの認識率が脳波のみの場合より改善されているが, リラックスでは認識率の低下が見られる.

全体的に認識率が低い原因として, センサや特徴抽出の問題の他に, 個人独立型の学習と認識を行っている影響が大きく, 個人毎に認識器を構成すれば認識率は改善されるものと予測される¹⁷⁾. 対象の感情カテゴリ数を減らし, 正の感情である喜び, 負の感情である怒り, 普通

表 10 認識結果 (3 感情: 脳波, 脈拍, 皮膚電気抵抗) [%].

Tab. 10 Confusion matrix (3 emotions: EEG, pulse, GSR) [%].

In \ Out	Joy	Anger	Relax
Joy	66.7	25.0	8.3
Anger	0	66.7	33.3
Relax	0	33.3	66.7

であるリラックスの 3 感情のみのデータを対象として全てのセンサ情報を使用した場合について学習・認識実験を行った結果を表 10 に示す²²⁾. 3 感情の認識では平均 67% の認識率が得られている. 5 感情と 3 感情では学習・認識に使用するデータ数が異なるため結果の直接的な比較は困難であるが, 表 9 より本実験で構成した感情認識手法では哀しみ, 恐れ, リラックス間での誤認識の傾向が大きいことがわかり, そのため 5 感情の実験ではこれらの感情に対する認識率が低下するため全体の認識率が低くなってしまふと考えられる.

以上より, 全てのセンサ情報を用いた 5 感情認識で平均認識率は 43% であるが, マルチモーダルセンサ情報は感情認識にとって有用であるといえる.

5. まとめ

本研究では, マルチモーダルな生体信号情報からの感情認識方法を提案し, その特徴について検討した. 脳波, 脈拍, 皮膚電気抵抗を計測するセンサによる生体信号計測システムを構築し, ビデオコンテンツによる感情刺激時の生体信号データを収集した. サポートベクトルマシンによる感情認識システムを構成し, 収集したデータによる学習・認識実験を行った結果, 喜び, 怒り, 哀しみ, 恐れ, リラックスの 5 感情について全てのセンサ情報を用いた場合に平均認識率 43% が得られ, マルチモーダルセンサ情報の感情認識における有用性を示した.

参考文献

- 1) F. Hara: Artificial Emotion of Face Robot through Learning in Communicative Interactions with Human, Proceedings of 13th IEEE International Workshop on Robot and Human Interactive Communication, 7-15, 2004.
- 2) R. E. Kaliouby and P. Robinson: Mind Reading Machines - Automated Interface of Cognitive Mental States from Video, Proceedings of 2004 IEEE International Conference on Systems, Man, and Cybernetics, 682-688, 2004.
- 3) R. W. Picard: "AFFECTIVE COMPUTING," The

- MIT Press, 1997.
- 4) F. Dellaert, T. Polzin, and A. Waibel: Recognizing Emotion in Speech, Proceedings of Fourth IEEE International Conference on Spoken Language Processing, Vol. 3, 1970-1973, 1996.
 - 5) J. Nicolson, K. Takahashi, and R. Nakatsu: Emotion Recognizing in Speech Using Neural Networks, Neural Computing and Applications, Vol. 9, No. 4, 290-296, 2000.
 - 6) I.A. Essa and A. P. Pentland: Facial Expression Recognition using a Dynamic Model and Motion Energy, Proceedings of the 5th International Conference on Computer Vision, 360-367, 1995.
 - 7) Y. Yacoob and L. Davis: Recognizing Human Facial Expressions from Log Image Sequences Using Optical Flow, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 6, 636-642, 1996.
 - 8) L. S. Chen and T. S. Huang: Emotional Expressions in Audiovisual Human Computer Interaction, Proceedings of 2000 IEEE International Conference on Multimedia and Expo 2000, Vol. 1, 423-426, 2000.
 - 9) L. C. DeSilva: Audiovisual Emotion Recognition, Proceedings of 2004 IEEE International Conference on Systems, Man, and Cybernetics, 649-654, 2004.
 - 10) K. Sasaki: Sensor Network And Its Application To Wearable Healthcare Systems, Proceedings of the 2003 JSME-IIP/ASME-ISPS Joint Conference on Micromechatronics for Information and Precision Equipment, 147-150, 2003.
 - 11) R. W. Picard, E. Vyzas, and J. Healey: Toward Machine Emotional Intelligence: Analysis of Affective Physiological State, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 10, 1175-1191, 2001.
 - 12) A. Faro, D. Giordano, and S. Mineo: A Fuzzy Processor Based Distributed System for Implementing Affective Human-Computer Interfaces, Proceedings of 2004 IEEE International Conference on Systems, Man, and Cybernetics, 2693-2698, 2004.
 - 13) K. Ishino and M. Hagiwara: A Feeling Estimation System Using a Simple Electroencephalograph, Proceedings of 2003 IEEE International Conference on Systems, Man, and Cybernetics, 4204-4209, 2003.
 - 14) 脳機能研究所: <http://www.bfl.co.jp/>
 - 15) ヒトセンシング株式会社: <http://www.h-sensing.co.jp/>
 - 16) 高橋和彦: SVM による EEG からの感情識別に関する一考察, 日本人間工学会誌, Vol. 39, No. 2, 90-92, 2003.
 - 17) K. Takahashi: Comparison of Emotion Recognition Methods from Bio-Potential Signals, 日本人間工学会誌, Vol. 40, No. 2, 90-98, 2004.
 - 18) V. Vapnik: "The Nature of Statistical Learning Theory," Springer, 1995.
 - 19) A. Junker: "BRAIN-BODY ACTUATED SYSTEM," United States Patent, 5474082, 1995.
 - 20) O. Chapelle, P. Haffner, and V. Vapnik: Support Vector Machines for Histogram-based Image Classification, IEEE Transactions on Neural Networks, Vol. 10, No. 5, 1055-1064, 1999.
 - 21) 石井健一郎, 上田修功, 前田英作, 村瀬洋: "わかりやすいパターン認識," オーム社, 1998.
 - 22) K. Takahashi: Remarks on SVM-Based Emotion Recognition from Multi-Modal Bio-Potential Signals, Proceedings of 13th IEEE International Workshop on Robot and Human Interactive Communication, 95-100, 2004.