

人間の動作系列からの感情・情動推定手法

加賀谷 拓† 羽倉 淳‡ 藤田 八ミド‡

岩手県立大学ソフトウェア情報学部ソフトウェア情報学科

1. はじめに

今後の社会では人間の感情を認識し介護支援やサービスを行うロボットなどのシステムへの要求が高まることが予想される。人間の感情を認識することで、指示待ちのシステムではなく、自律的に様々な状況に合わせた行動をとることが可能となる。そのため、これまでも顔表情や音声、身体動作、生体信号などを用いて多くの感情認識手法が提案されてきた[1][2][3][4]。この中でも、特に身体動作からの認識は利用者と対面しない状況下において動作者を観察することが出来るため、介護支援・サービスの観点で重要であると考えられる。

一方で身体動作からの感情認識の研究[5]においては、定型的一連の動作や、身体の最終的な姿勢そのものがテンプレートとして定義されており、テンプレートと認識対象となる動作とのマッチングにより認識を行っている。そのため、テンプレートと完全に一致しない動作については認識が困難である。そこで、本研究では感情を含む身体動作から動作の感情を表す特徴を様々な視点から抽出する。その特徴に基づきテンプレートを作成することで感情を伴う動作の一般化を行う。これにより、構築したテンプレートに基づく感情表現を抽出することで、システム利用者の感情を推測するシステムの構築を試みる。

2. 感情と動作

本研究での感情と動作の定義を以下に示す。まず感情の定義としては、プルチックの基本感情モデルを用いている。プルチックは感情の立体モデルを提唱した人物で、日常見られる感情はいくつかの因子に分類することが出来ると述べている。このことから、感情を強度ごとに分類したものを以下に示す。

表 1. 8 基本感情の強度差による分類

感情レベル	1	2	3	4	5
喜び	平穏	爽快	幸福	喜び	歡喜
受容			合一	受容	崇拜
恐れ	臆病	懸念	心配	恐れ	恐怖
驚き			混乱	驚き	驚嘆
悲しみ		陰気	落胆	悲しみ	悲嘆
嫌悪	面倒	うんざり	嫌気	嫌悪	憎悪
怒り		当惑	苛立ち	怒り	激怒
期待		構え	予期	期待	願望

感情を伴った動作は、その強度ごとに異なる。一般的に強度が強いほど動作は大きく・早くなるのがわかっていて、これは、身体の動作により不安定になった心理状態を緩和しようとするためであると言われている。また、舞踊表記譜法を構築したラバン[6]は人間の動作を「時間・空間・エネルギー」の尺度で分類できることを示し、舞踊における感情表現の手法を構築した。このことを以下の表 2 に示す。

表 2. ラバンの動作特徴による分類

	攻撃形態	陶酔形態
時間特徴	急変化, 断続的	緩慢, 連続的
空間特徴	直線的, 収斂	曲線的, 発散
エネルギー特徴	エネルギー大	エネルギー小

上記を元に本研究での動作の定義として、動作部位が動き始めてから元の位置に戻るまでを一つの動作とする。また、同様の動作を繰り返し行っている場合は、繰り返し終了し動き始めの位置に戻るまでを一つの動作とみなすこととする。

3. 動作特徴抽出によるテンプレート作成

本手法では、感情を含む動作の特徴テンプレートにより観測対象となる人間の行動系列から感情行動を抽出することで、感情の推定を行う。そのためにまず、表 1 の各感情を示す動作を役者に演じさせる。その際に、動作を特徴付ける身体の特徴点の動きを取得し、動作データとする。動作データから 3.1 で定義する特徴空間の各特徴量を人手にて抽出し、これを特定の感情行動の各特徴点に対するテンプレートとする。全ての特徴点に対するテンプレートマッチングの結果に基づき、感情動作の推定を行う。図 1 に概念図を示し、本節ではその詳細を述べる。

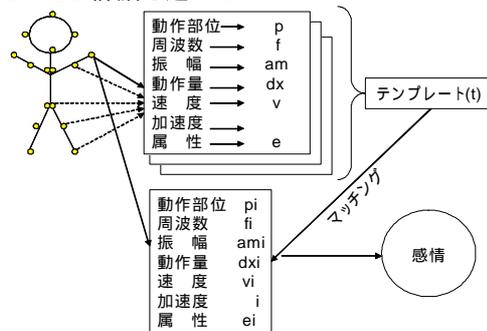


図 1. システムの概念図

3.1 演技からのテンプレート作成

感情を含む動作の特徴空間を構成する各特長を表 2 の動作特徴の分類を用いて、以下のように定義する：

1. 動作部位 (p)・・・動作ポイント名
 2. 周波数 (f)・・・一動作の基本周波数
 3. 振幅 (am)・・・一動作の最大振幅
 4. 動作量 (dx)・・・x,y,z それぞれのフレーム間距離の差
 5. 速度 (v)・・・各フレームの動作量 / 時間 = 速度, x,y,z のベクトルの向き
 6. 加速度 (a)・・・各フレームの速度の微分
 7. 属性 (e)・・・動作の発生した感情と強度
- 4・5・6 の特徴はフレーム数分の時系列データを保持する。これによって得られたデータの集合を一つのテンプレート(1)とし、テンプレートの集合を(2)とする。また、それぞれのテンプレートは(3)に示すデータを保持する。各大文字が表しているのは抽出された定数である。

$$t_i = \langle p, f, am, dx, v, a, e \rangle \quad (1)$$

$$T = \{t_i \mid i = 0, \dots, N\} \quad (2)$$

ここで、4 から 6 の特徴は、フレーム数に対応する時系列データも同時に保持するものとする。

役者に表 1 の感情毎に演じさせて得た動作データから特徴空間の各特徴量を上記の定義に基づき抽出する。

これらは、上記の各特徴の定義で示した数値であり、動作データから自動的に計算されるようアルゴリズム化することができる。これらをテンプレートデータベース内に保持し、感情を推定したい対象者の行動とのマッチングの際に用いる。

3.2 Segmental CDP 法に基づくマッチング

Segmental CDP 法[7]は動的計画法 (DP) を基に、時系列データとのマッチングを可能にしたもので、ユーザが設定したセグメント長との時系列の入力データに対して類似度を算出することでマッチングを行う。これにより、テンプレートと入力されたデータの特徴の時系列データの各特徴毎に類似部分を検出する。

3.3 感情認識手法

検出の結果、それぞれの特徴毎に類似度の高かったフレームに対し類似レベルとして評価を与える。特徴毎の評価後、それぞれの評価値を加算し類似レベル合計とする。この類似レベルがある一定の閾値以下であれば、マッチングしたテンプレートが示す感情状態であると認識する。この手法では、対応する感情が複数見つかる場合が考えられる。その場合にはそれぞれの感情が同様に発生しているものとする。

4. 演技に基づく感情抽出実験

提案手法により感情抽出が行えることを確認するため、特定の動作に限定し実験を行った。

4.1 テンプレート作成

テンプレートを作成する際に用いる動作の取得にはモーションキャプチャシステム (Motion Analysis EVA HiRES Ver.7.0) [8]を用いた。また、演技者は感情を観衆に伝える方法を学んでいる岩手県立大学演劇部員である。演技者は表 1 にある 8 感情 5 段階の感情をそれぞれ起立状態で 10 秒間の演技で得られたデータを利用する。

取得したデータを観察した結果、恐れ、驚き、悲しみ、嫌悪、怒りの強度 3 以下において、類似した動作が検出された。それは、他の動作を行わず、特定の動作部位のみが反復性・周期性を持つ比較的小さな動作を行うというものであった。テンプレートの構築にはまずこの動作から行うこととした。

動作データからは上記に示した 7 特徴を抽出する。今回テンプレートを構築する際、怒りの強度 3「苛立ち」で取得したデータを用いることとした。動作部位は右手人差し指の先端を表す R-Index-Top とし、属性は表 1 から Ikari-3 と示す。以下に実際に取得したデータからテンプレートを作成した例を示す。

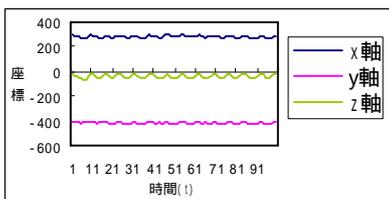


図 2. テンプレート作成時に用いた動作データ (x 軸と z 軸に周波数と振幅の小さい反復動作が見られる)

表 3. 作成したテンプレート例

動作部位	R-Index-Top
周波数	実部 -92.8958..... 虚部 266.03.....
振幅	10.8618
動作量	x1(1,2)=-1.7456
速度	x1(1,2)=45.9759, x
加速度	1379.277
属性	Ikari, 3

4.2 感情認識実験

ここでは上記のテンプレートの有効性を示すため、新たに取得した動作データとのマッチングを行った。今回の実験では、テンプレート化したデータと同様の動作を行った以下のデータを用意した。

それぞれ図 3 は膝の上を右手人差し指で繰り返し叩く動作で、類似性が高くなると考えられる。図 4 は腕を上げる動作で、類似性は無いと考えられる。まずはこの二つのそれぞれ異なる動作について評価する。

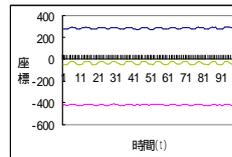


図 3. 類似動作

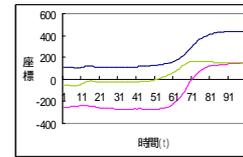


図 4. 非類似動作

5. 実験結果と考察

図 3 の動作においては、動作部位・周波数・動作量の特徴において、類似度が高くなった (表 4)。速度の類似度では、以上の値が算出されたが、この値の類似度は低いため、マッチングしているとは考え難い。加速度に關しても同様のことが言えるが、今回は閾値を低めに設定したため、マッチングはしていない。また、図 4 の動作はテンプレートとは全く違う動作であるため、類似度はどの項目も最低値となっている (表 5)。

表 4. 類似動作

表 5. 非類似動作

	類似度	類似レベル
動作部位	R-Index-Top	1
周波数	224.448776	1
動作量	16.503613	1
速度	6821.604004	2
加速度	6139447.5	3
向き	y	1
類似レベル合計		9

	類似度	類似レベル
動作部位	R-Index-Top	1
周波数	100000000	3
動作量	100000000	3
速度	100000000	3
加速度	100000000	3
向き	z	3
類似レベル合計		16

以上の結果から今回の実験では、テンプレートに存在する動作の類似動作ではその動作の持つ感情を認識することが可能であることが確認できた。また、全く意味の異なる動作特徴の似ていない非類似動作は異なる動作として認識された。

6. まとめ

本論文では人間の連動作から、動作者がどのような感情であるかを認識することを目的とし、テンプレートに登録されている動作と類似した動作の感情を認識することが出来た。また、特徴の類似していない動作については、異なる感情として認識することができた。このことから本研究の有効性が示された。しかし、データによっては動作特徴は類似しているが、動作の持つ感情の意味が異なるものも存在することが予想される。この問題を解決するためには、より多くのデータに対し実験をし、問題点を明確にしていく必要がある。今後は上記の問題点の解決と、目的である連動作から時系列で変化していく各感情の認識を目指す。

参考文献

- [1] 岩崎晴美 他 4 名: 顔の表情抽出, 計算科学研究センター研究報告(2000).
- [2] ドナ・エリクソン, 大西真由子: 感情音声の認識に関する男女差, 岐阜市立女子短期大学研究紀要(2004).
- [3] Kenji Amaya, Armin Bruderlin, Tom Calvert: Emotion from Motion, Proceedings of the conference on Graphics interface(1996).
- [4] 津田良一, 竹川仁志: 心拍変動による精神作業ストレスの定量解析, 九州東海大学紀要応用情報学部 3 号工学部第 29 号(2002).
- [5] 八村広三郎: 舞のデジタルアーカイブ, 立命館大学(2002)
- [6] Ann Hutchinson: Labanotation, Theatre ArtsBooks, 1977
- [7] 岩淵 晃: 曲内類似性及び曲間非類似性を利用した曲境界の自動検出, 岩手県立大学ソフトウェア情報学部卒業論文要旨集(2004).
- [8] Motion Analysis Corporation: EVA Version 7.0 Reference Manual, 住商エレクトロニクス (2002).