

ラバン特徴と Autoencoder を用いた ヒューマノイドロボットの身体動作生成に関する研究

Motion Generation for Humanoid Robots Using Laban's Features and Autoencoder

加藤 瑛樹^{*1} 加藤 昇平^{*1*2}
Hideki Kato Shohei Kato

^{*1}名古屋工業大学 大学院工学研究科情報工学専攻

Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

^{*2}名古屋工業大学 情報科学フロンティア研究院

Frontier Research Institute for Information Science, Nagoya Institute of Technology

Recently, emotion expression of robot has been considered as a very important factor during communication between robots and human. As ways of emotion expression, there is not only verbal information but also nonverbal information such as facial expressions and body motions. Nonverbal information has a greater influence on emotional transmission than verbal information. In this paper, we focus on body motions, which is one of way for a robot to express its emotion. We propose an estimation method and a body motion generation function that can help a robot to express the appropriate emotion by using body motion features based on Laban theory.

1. はじめに

近年、人とコミュニケーションをとる様々なロボットが開発されており、これらのロボットは認知症予防や見守りロボットなど、今後の社会での活躍が期待されている。

コミュニケーションとは互いの意志や感情を伝達することである。その際、言語による情報だけではなく顔の表情や身振りなど言葉によらない非言語情報も重要な要素である。また、人は擬人視を行った対象から感情を感じることで、その対象に対して愛着や親近感を抱きやすくなることが知られている。親近感を抱きやすくなることで、人のロボットに対する不安感を軽減し、介護などの場面において物理的なサポートだけでなく、心理的なサポートも行う事が可能になる。本研究では、人と同じ動作が可能であり、人と同じ生活環境で重要な役割を担うとされるヒューマノイドロボットを対象として、ロボットが身体動作を用いて適切な感情を表出する機能の実現を目標としている。そして、このためには以下のような機能が必要になると考えられる。

1. 自己の身体動作が相手に与える感情を自覚できる
2. ある感情を表出する身体動作を自律的に生成できる
3. 表出すべき感情を適切に表出できる

本稿ではラバン身体動作表出理論 [Laban 88] (以下、「ラバン理論」) に基づいた身体動作の特徴 (以下、「ラバン特徴」) を定義し、1. の機能としてラバン特徴と機械学習手法の 1 種である Random Forests [Breiman 01] を用いた感情推定手法を提案し、2. の機能としてラバン特徴とニューラルネットワークの 1 種である Autoencoder [Geoffrey 06] を用いた身体動作生成に関する研究を行った。

2. ラバン身体動作表出理論

本節ではラバン理論 [Laban 88] に基づいた身体動作の特徴であるラバン特徴の抽出手法を説明する。

表 1: ラバン特徴量

ラバン特徴	特徴	特徴量の抽出法
<i>Space</i>	身体動作の 方向的偏り具合	四肢先端の単位運動ベクトル 顔単位方向ベクトルの 内積の重み付き合計
<i>Time</i>	身体動作の速さ	全関節の角速度の重み付き合計
<i>Weight</i>	身体動作の力強さ	全関節のトルクの重み付き合計
<i>Inclination_x</i>	姿勢の 前方向への偏り	重心位置と支持脚中心位置の 前後方向の差
<i>Inclination_y</i>	姿勢の 左右方向への偏り	重心位置と支持脚中心位置の 左右方向の差
<i>Height</i>	姿勢の高さ	重心位置と基本重心位置との 上下方向の差
<i>Area</i>	姿勢の 水平面上の広がり	四肢先端が水平面上に作る 四角形の面積

2.1 ラバン理論

ラバン理論 [Laban 88] とは動作体の心理状態と身体運動の相関を規定する舞踊学の理論である。Darwing が提唱した「動物と身体動作の構造に関する理論」を受け継ぎ、Rudlf von Laban を中心とするドイツの表現主義舞踊の創作者らが構築した。ラバン理論の概念は数理的かつ具体的であるため、理工学との親和性が高いと言われている。

2.2 ラバン特徴

本研究では、ラバン特徴量の抽出手法として増田ら [増田 11] が提案した抽出手法を使用する。表 1 に増田らが提案した抽出式で得られるラバン特徴・特徴量の概要・特徴量抽出方法の概要を示す。増田らはヒューマンフォームロボット (HFR) KHR-2HV の全身動作からラバン特徴量を得る抽出式を提案し、ラバン特徴量と喜怒哀楽の 4 感情に高い相関があることを示した。時刻 t における各ラバン特徴量は、以下に提示する抽出式を用いて算出される。

Space は身体動作の方向的な偏りの大きさの程度を示す特徴量であり、四肢先端の単位運動ベクトル (運動ベクトルの単位ベクトル)、および顔単位方向ベクトルの内積の重みつき合計と定義する。値が大きい場合は身体動作は単一方向に向かっており、値が小さい場合は身体動作は散逸方向に向かっている。

$Space(t)$ は下式で算出される．

$$Space(t) = \sum_{m=0}^3 \sum_{n=m+1}^4 a_{mn}(\vec{r}_m(t) \cdot \vec{r}_n(t)) \quad (1)$$

ここで， m, n は顔および四肢先端の識別番号を示す． \vec{r}_m は顔の単位方向ベクトル ($m=0$ の場合)，あるいは四肢先端 m の単位運動ベクトル ($m=1,2,3,4$ の場合) を示す． a_{mn} は重み係数であり， m, n に関連する部位の視覚的質量と被注目度を考慮し定義者が設定する．

$Time$ は身体動作の速さの程度を示す特徴量であり，全関節の角速度の重みつき合計と定義する．値が大きい場合は身体動作は素早く，値が小さい場合は身体動作は遅い． $Time(t)$ は下式で算出される．

$$Time(t) = \sum_i b_i \dot{\theta}_i(t) \quad (2)$$

ここで， i は関節番号を示し $\dot{\theta}_i$ は関節 i の角速度を示す． b_i は重み係数であり，関節 i に関連する部位の視覚的質量と被注目度を考慮し定義者が設定する．

$Weight$ は身体動作の力強さの程度を示す特徴量であり，全関節のトルクの重みつき合計と定義する．値が大きい場合は身体動作は力強く，値が小さい場合は身体動作は弱々しい． $Weight(t)$ は下式で算出される．

$$Weight(t) = \sum_i c_i \tau_i(t) \quad (3)$$

ここで， i は関節番号を示し τ_i は関節 i のモータのトルクを示す． c_i は重み係数であり，関節 i に関連する部位の視覚的質量と被注目度を考慮し定義者が設定する．

$Inclination$ は姿勢の偏りの程度を表す特徴量であり，重心位置と支持脚中心位置の差と定義する． $Inclination$ は前後 (x 方向)，左右 (y 方向) の両成分についてそれぞれ求められる．それぞれを $Inclination_x$ ， $Inclination_y$ とし， $Inclination_x$ が大きい場合は姿勢は前寄りであり， $Inclination_x$ が小さい場合は姿勢は後ろ寄りである．また $Inclination_y$ が大きい場合は姿勢は左寄りであり， $Inclination_y$ が小さい場合は姿勢は右寄りである． $Inclination_x(t)$ ， $Inclination_y(t)$ は下式で算出される．

$$Inclination_x(t) = cog^x(t) - col^x(t) \quad (4)$$

$$Inclination_y(t) = cog^y(t) - col^y(t) \quad (5)$$

ここで， cog^x は HFR の重心の x 座標を， cog^y は HFR の重心の y 座標をそれぞれ示し． col^x は支持脚中心の x 座標を， col^y は支持脚中心の y 座標をそれぞれ示す．

$Height$ は姿勢の高さの程度を表す特徴量であり，重心位置と基本重心位置（直立状態での重心位置）との上下方向との差と定義する．値が大きい場合は姿勢は上寄りであり，値が小さい場合は姿勢は下寄りである． $Height(t)$ は下式で算出される．

$$Height(t) = cog^z(t) - cog_0^z \quad (6)$$

ここで， cog^z は重心の z 成分を示し cog_0^z は基本姿勢（直立）における重心の z 成分を示す．

$Area$ は姿勢の水平面上の広がりを表す特徴量であり，四肢先端が水平面上に作る四角形の免責と定義する．値が大きい場

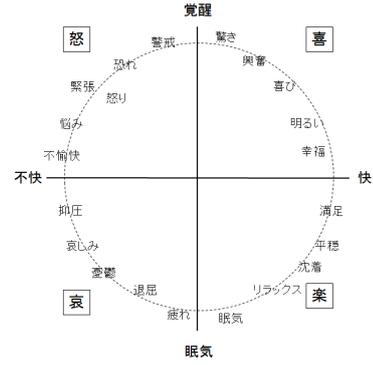


図 1: ラッセルの感情円環モデルと喜怒哀楽

合は姿勢は広く，値が小さい場合は姿勢は狭い． $Area(t)$ は下式で算出される．

$$Area(t) = \sum_{m=1}^3 \sum_{n=m+1}^4 \frac{1}{2} [Pos_m^x(t)Pos_n^y(t) + Pos_m^y(t)Pos_n^x(t)] \quad (7)$$

m, n は四肢先端の識別番号を示し， Pos_m^x ， Pos_m^y は四肢先端の x 座標と y 座標をそれぞれ示す．

抽出式 (1) から (7) を用いて抽出された数値を特徴ごとに正規化し，これを身体動作におけるラバン特徴量の時系列データとする．

短時間の身体動作を対象とした場合，身体動作時間全体のラバン特徴量の平均をその身体動作のラバン特徴量とみなすものとする．例えば短時間身体動作の $Space$ は下式で算出される．

$$Space = \frac{1}{N} \int Space^*(t) dt \quad (8)$$

ここで， N は動作時間長を， T は単位時間長を示し， $Space^*(t)$ は $Space$ 全体に対して正規化した $Space(t)$ を示す．他 6 種類のラバン特徴についても (8) 式と同様に短時間身体動作のラバン特徴量を算出する．

3. 予備実験

身体動作が表出する感情を取得する目的で，印象評価実験を行った．

3.1 ラッセルの円環モデル

ラッセルの円環モデル [Russell 80] では図 1 のように快 - 不快，覚醒 - 眠気をそれぞれ軸とする 2 次元平面上に様々な感情が配置される．この 2 次元平面を 2 軸で分割した各領域の代表的な感情が喜怒哀楽となる．本研究では，身体動作が表出する感情としてラッセルの円環モデルを参考に喜怒哀楽の 4 感情を扱う．

3.2 印象評価実験

被験者は 20 代の男女 15 名，身体動作は HFR である KHR-2HV の 8[s] 間の全身動作を 40 種類用意した．図 2 に KHR-2HV の外観を，図 3 に印象評価実験に用いた回答用紙を示す．被験者は観察した動作から，喜怒哀楽それぞれの感情がどの程度感じられるかを推定し図 3 の回答用紙に回答する．その感情を全く感じなければ 0 に印をつけ，その感情を強く感じるほど 1 に近い位置に印をつける．回答の印の位置に応じて 0

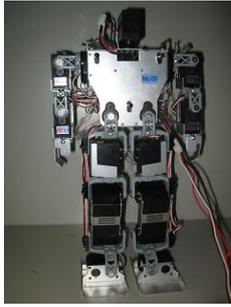


図 2: KHR-2HV の外観

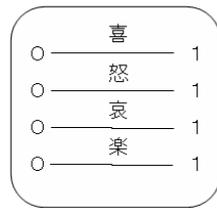


図 3: 回答用紙

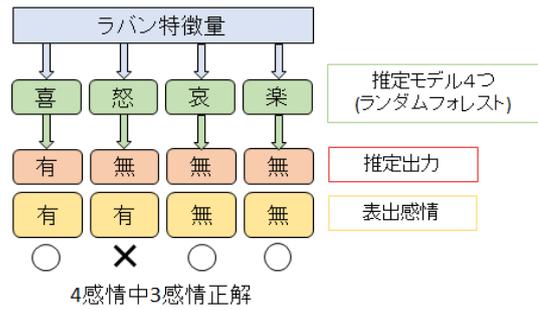


図 4: 感情推定モデル

表 2: 各感情を表出する動作の個数

α	喜	怒	哀	楽	複数感情
100	14(0)	7(0)	11(0)	8(0)	0
90	15(1)	8(1)	11(0)	9(1)	3
80	16(2)	9(2)	11(0)	10(2)	6
70	17(3)	9(2)	11(0)	11(3)	8
60	21(7)	11(4)	13(2)	13(5)	18
50	21(7)	12(5)	13(2)	14(6)	20

~100 までの評価値に線形変換し、ある動作について 15 人のそれぞれの感情の評価値の平均を「感情評価値」と定義する。回答を喜怒哀楽のどれか 1 つとしなかったのは、身体動作が複数の感情も表出し得る場合があると考えたためである。

4. 感情推定

ラバン特徴と Random forests[Breiman 01] を用いて身体動作が表出する感情を推定する感情推定実験を行った。

4.1 身体動作の表出感情

予備実験で得られた感情評価値を用いて身体動作の表出感情を決定する。身体動作の表出感情は喜怒哀楽のどれか 1 種類だけではなく、複数の感情を表出する場合も考えられる。そこで本実験では、身体動作の表出感情を高々 2 種類とし、採択される感情は以下の 2 種類とした。

1. 「主感情」
その身体動作の感情評価値で最大となる感情
2. 「副感情」
その身体動作の 2 番目に大きい感情評価値であり、その値が主感情の感情評価値の $\alpha\%$ より大きい感情

ただし α はパラメータであり、100~50 まで 10 刻みで設定した。 $\alpha = 100$ の場合は主感情だけが表出感情となる。

以上のように身体動作が表出している感情を決定した結果、各感情を表出する動作の個数は表 2 のようになった。ただし () 内の値は、その感情を副感情として表出する動作の個数を表す。

表 2 を見ると哀の副感情は $\alpha = 60$ 以下から増加するのに対して、喜・怒・楽の副感情は $\alpha = 90$ から増加していることがわかる。このことから哀の副感情は身体動作からは感じられにくく、喜・怒・楽の副感情は哀の副感情と比べて身体動作から感じられやすいことが考えられる。

4.2 Random Forests

Random Forests[Breiman 01] は識別性能の低い決定木を弱学習機として多数用意し、それらの出力を統合して高い予測性

能を得ることができる機械学習手法の 1 種である。特徴としてノイズに頑健で高速な識別が可能であり、高次元・低サンプルに強い点があげられる。

以下に Random Forests のアルゴリズムを示す。

1. データセット D から N 個のブートストラップサンプル B_1, B_2, \dots, B_N を作成する。
2. 作成した各サンプル $B_k (k = 1, 2, \dots, N)$ をトレーニングデータとして、 N 本の決定木 $T_k (k = 1, 2, \dots, N)$ を作成する。
3. 指定したノード数に達するまで、以下の方法で決定木のノードを作成する。
 - (a) トレーニングデータ B_k の説明変数 M 個のうち、 m 個をランダムに選択する。 $m = \sqrt{M}$ が多く用いられる。
 - (b) 選択された m 個の説明変数のうち、トレーニングデータを最もよく分類するものと、そのときの閾値を用いてノードの分割関数を決定する。
4. データセット D のある 1 つのデータ $d_i (i = 1, 2, \dots, D)$ を、 d_i を学習していない決定木の群に対し入力として与え、それぞれの決定木の出力を統合し推定出力を求める。ただし、データ d_i をトレーニングデータとして学習していない決定木は N 本の決定木のうち約 $\frac{1}{3}$ である。
5. データセット D に対し求めた推定出力から正答率を算出し、これを全体の正答率とする。

4.3 感情推定実験

本実験の感情推定モデルの概要図を図 4 に示す。ラバン特徴から喜怒哀楽のどれか 1 つを推定するのではなく、喜怒哀楽それぞれについて「その感情を表出しているか否か」を Random Forests を用いて推定する。このことにより、身体動作が複数感情を表出する場合も考慮することができる。説明変数を動作時間全体の平均をとったラバン特徴量 7 項目、目的変数を喜怒哀楽それぞれについて「その感情を表出しているか否か」のラベルとした 4 種類のデータセットを作成して、それぞれ Random Forests を学習させる。

4.4 実験結果

正答率を表 3 に示す。完全正答は喜怒哀楽それぞれの推定モデルの出力が正答となった割合を表す。 $\alpha = 100$ の場合は高い正答率が得られたが、表出感情として感情を複数採択することになる $\alpha \leq 90$ からは喜については正答率の向上が、怒・哀

表 3: 正答率 (単位: %)

α	喜	怒	哀	楽	完全正答
100	87.15	91.85	100	95.03	83.68
90	89.43	87.53	100	90.05	74.28
80	94.00	86.70	100	85.60	69.85
70	88.00	86.95	100	92.50	71.93
60	92.05	84.93	92.50	95.00	69.71
50	92.13	79.90	92.50	88.13	63.28

・楽については正答率の低下が確認された．正答率の低下の原因としては，主感情と副感情を同等に扱ってしまったことが考えられる．

5. 身体動作生成

身体動作からは単位時間ごとのラバン特徴量が得られる．ラバン特徴量は感情と相関するため，ラバン特徴量を加工することで，表出感情を変化させることができる．しかし，7次元のラバン特徴を同時に加工することは困難であるため，次元圧縮手法の1つである Autoencoder[Geoffrey 06] を用いてラバン特徴を3次元に圧縮する．圧縮されたラバン特徴を加工した後，KHR-2HV の各関節角度を算出することで身体動作を生成する．

5.1 Autoencoder

Autoencoder[Geoffrey 06] はニューラルネットによる次元圧縮手法の1種である．入力層と出力層のユニット数が同じであり，入力データを教師データとして用いて学習を行う．そのため，入出力層のユニット数より少ないユニット数となる隠れ層において，圧縮された入力データが得られる．本研究では予備実験で扱った40種類の動作を対象に，単位時間ごとに得られる7次元のラバン特徴量を3次元に圧縮した．隠れ層の3つのユニットの出力をそれぞれ u, v, w 軸にとり，圧縮されたラバン特徴の軌跡を3次元空間上にプロットする．この空間の全体図を図5(a)に， uv 平面， vw 平面， uw 平面をそれぞれ図5(b)～図5(d)に示す．初期姿勢は全ての動作で同じであるため，軌跡の初期地点は同じとなる．軌跡の初期地点は P ，原点は O とする．また，各身体動作の表出感情を主感情として，喜・怒・哀・楽の順に水色・赤色・青色・緑色と色分けを行った．凡例を図5(a)だけに示すが，図5(b)～図5(d)でも同様である．

5.2 考察

ラッセルの円環モデルを参考に図5(b)～図5(d)に関して考察を行う．快・不快方向に関しては，快の感情に属する喜・楽の軌跡が全体的に存在しており，不快の感情に属する怒・哀はそれぞれ偏りが確認できる．覚醒・眠気方向に関しては，覚醒の感情に属する喜・怒は初期地点 P から w 成分が大きい方向に偏っており，眠気の感情に属する哀・楽は w 成分が比較的小さい事が確認できる．以上の事から，この空間においてプロットされた軌跡は感情ごとにある程度特徴的な偏りがあることが考えられる．この偏りを考慮してある身体動作のラバン特徴を加工することで，ある感情を意図的に付加した身体動作が生成できると考えられる．

6. おわりに

本研究ではロボットが身体動作を用いて適切な感情を表出する機能の実現のため，ラバン特徴を用いた感情推定手法と身

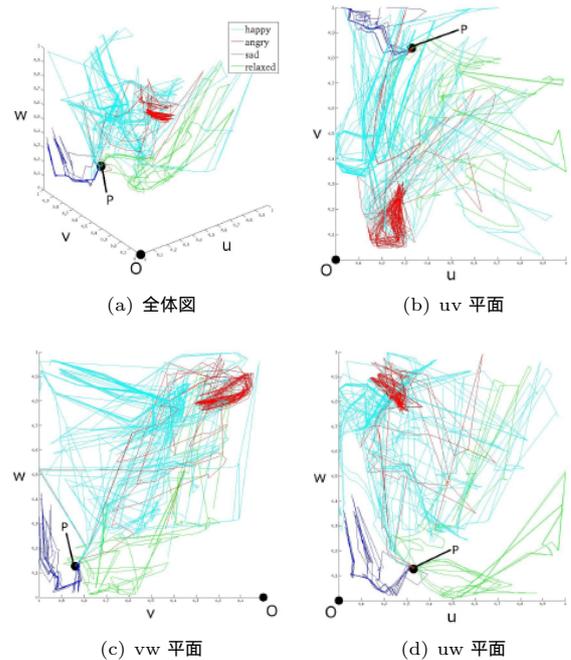


図 5: 圧縮されたラバン特徴

体動作生成に関する研究を行った．感情推定手法では，身体動作が1つの感情のみ表出する場合において高い正答率が得られたが，複数の感情を表出し得る場合においては正答率の低下が確認された．原因としては主感情と副感情を同等に扱ってしまったことが考えられる．これは推定モデルの学習時に主感情と副感情の重み付けを行うことで改善できると考えている．身体動作生成に関する研究では，3次元に圧縮されたラバン特徴の軌跡に特徴的な偏りがあることが確認できた．この特徴を考慮してラバン特徴を加工することで感情を変化させた身体動作を生成できると考えている．今後の課題はラバン特徴から動作体の各関節角度を算出する手法を検討することである．

参考文献

- [増田 11] 増田 恵, 加藤 昇平, 伊藤 英則: ラバン理論に基づいたヒューマンフォームロボット身体動作の動作特徴抽出と表出感情推定, 日本感性工学会論文誌, Vol.10, No.2, p.295-303, 2011
- [Laban 88] Laban, R. V.: *mastery of Movement*, Princeton Book Co. Pub. 1988
- [Russell 80] James A. Russell. A circumplex model of affect. In *Journal of Personality and Social Psychology*, pp. 1161-1178, 1980.
- [Breiman 01] Breiman L. : *Random forests*, *Machine Learning*, 45, pp.5-32, 2001
- [Geoffrey 06] Geoffrey E. Hinton, R. Salakhutdinov, *Reducing the Dimensionality of Data with Neural Networks*, *Science* 313(5786), pp.504-507