

ニューラルネットワークを用いた操作平面による直感的なロボットの動作生成

Operation generation of the intuitive robot by the operation plane using the neural network

伊藤 惇貴*¹
Junki Ito

有沢 怜士那*²
Reona Arisawa

加納 政芳*²
Masayoshi Kanoh

中村 剛士*³
Tsuyoshi Nakamura

小松 孝徳*⁴
Takanori Komatsu

*¹中京大学大学院 情報科学研究科
Graduate School of Computer and Cognitive Sciences, Chukyo University

*²中京大学 情報理工学部
School of Information Science and Technology, Chukyo University

*³名古屋工業大学大学院 工学研究科
Graduate School of Engineering, Nagoya Institute of Technology

*⁴信州大学 ファイバーナノテク国際若手研究者育成拠点
IYREC, Shinshu University

In recent years, since research and development of humanoid robots have been carried out, they have been familiar to us. Humanoid robots can move like human beings, because their appearance is similar to human beings. However, humanoid robots generally have a lot of DOFs, so creating/editing their motion not only is difficult, but takes a considerable amount of time. In this paper, we propose a motion-editing plane using onomatopoeias which refer to words that mimic the appearance or sound of things to enable richly realistic expressions. To create the plane, first, features of some known motions having an onomatopoeia are extracted by P-type Fourier descriptors. Second, similarity relationship between the features is learned by using a five-layer auto-associative neural network. Finally, the third layer, which has two units, of neural network is used as a motion-editing plane. Using this plane, even people who are not known to robotics can edit motion of humanoid robots intuitively.

1. はじめに

近年、人型ロボットの研究開発が進み、人型ロボットはより身近な存在になりつつある。人型ロボットは、容姿が人に似ているため、より人間に近いモーションが求められている。しかし、人型ロボットのモーション編集作業は自由度が多く手間がかかるため、容易な作業ではない。そこで、オノマトペを用いてモーションの編集作業を行う手法について考える。

オノマトペは、「擬音語」「擬態語」「擬声語」の総称であり、物体の音や響き、状態などを感覚的に表現したものであるため、一般語彙に比べて臨場感にあふれた繊細な表現や絶妙な差異を表現することができる。このことから、オノマトペの利用によって、ロボットのモーション編集作業の感覚的理解の補助、認知的負担の軽減が可能になると考える。

本稿では、オノマトペを用いてモーション編集するための操作平面システムを提案する(図1)。まず、ロボットモーションの基本波形をオノマトペの音素特徴に基づいた属性ベクトル¹⁾によって変化させる。次に、変形された波形の特徴を、P形フーリエ記述子^{2),3)}を用いて抽出する。その後、抽出した波形特徴を用いて、ニューラルネットワークを恒等写像学習させることで、オノマトペ間の類似関係を学習させて、ロボットの操作平面を生成する。使用者が操作平面上の任意の座標点を選択すると、ニューラルネットワークによりP形フーリエ記述子の特徴量が抽出されるので、これを逆フーリエ変換することでロボットのモーション波形を復元し、ロボットを動作させ

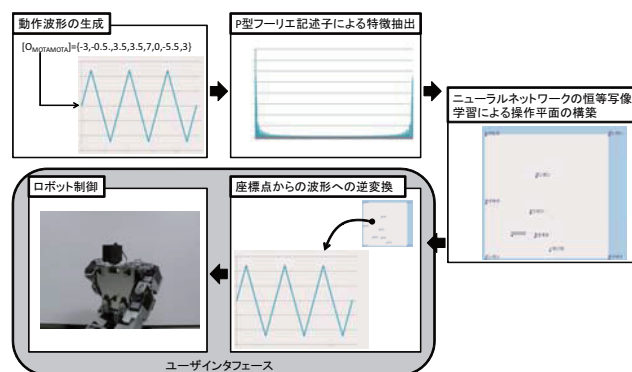


図 1: システムの概要

ることができる。

2. オノマトペの数値化と波形変換

2.1. オノマトペの数値化

本稿では、文献1)と同様にXYXY型のオノマトペを対象とする。

まず、オノマトペを構成する母音と子音に対して、「強さ」「硬さ」「湿度」「滑らかさ」「丸さ」「弾性」「速さ」「温かさ」の8属性を割り当てる。ここで、各属性の値は[-2, 2]の範囲で設定される。たとえば、子音のKについては、以下のように定義される。

$$K = \{2, 2, 1, 0, 0, 0, 2, -1\} \quad (1)$$

連絡先: 伊藤惇貴, 中京大学大学院情報科学研究科, 住所:愛知県豊田市床立 101 mail:h11202m@st.chukyo-u.ac.jp
連絡先: 加納政芳, 中京大学情報理工学部, 住所:愛知県豊田市床立 101 mail:mkanoh@sist.chukyo-u.ac.jp

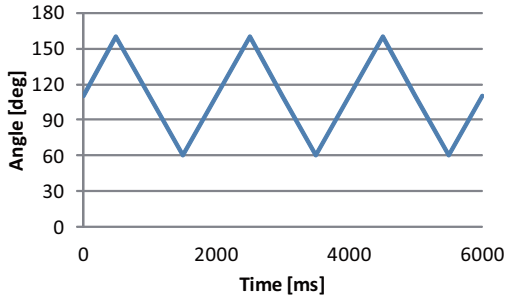


図 2: 基本波形

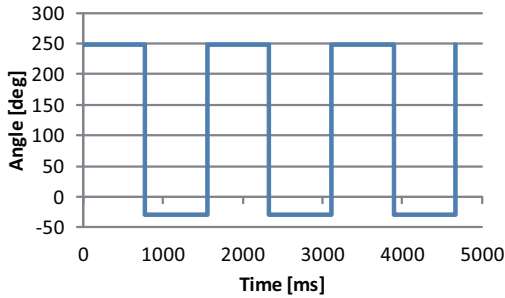


図 3: 「ガシガシ」の波形

同様に、全ての子音と母音に対して属性ベクトルを与え、それらを組み合わせることでオノマトペの属性を決定する。たとえば、「カサカサ」というオノマトペの場合、 $[K(\text{子音})][A(\text{母音})][S][A][K][A][S][A]$ という 4 つの子音と 4 つの母音で構成される。ここで、オノマトペ $XYXY$ の各音素を、母音 $X^{(v)}$, $Y^{(v)}$ および、子音 $X^{(c)}$, $Y^{(c)}$ と表記したとき、オノマトペ O の i 番目の属性値 O_i は以下の式で計算できる¹⁾。

$$O_i = 2X_i^{(c)} + X_i^{(v)} + \frac{(2Y_i^{(c)} + Y_i^{(v)})}{2} \quad (2)$$

ただし、 $Z_i^{(s)}$ は、 $Z \in \{X, Y\}$ の $s \in \{v = \text{母音}, c = \text{子音}\}$ の i 番目の属性値を表す。

2.2. 波形の生成

式 (2) で求められる 8 次元ベクトルの値に従い、基本波形 (図 2) の振幅および周期を変更することでオノマトペ O の動作波形を生成する。

たとえば、「ガシガシ」、「ウロウロ」の波形例を図 3、図 4 にそれぞれ示す。波形変換の詳細については文献 1) を参照されたい。

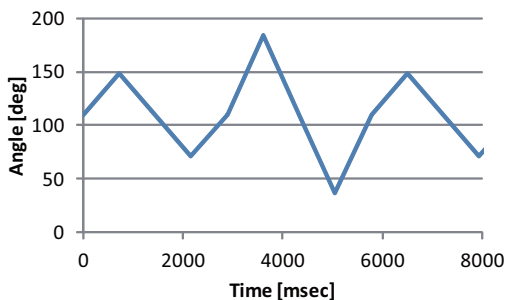


図 4: 「ウロウロ」の波形

3. 波形特徴の抽出

オノマトペの波形特徴を抽出するために P 形フーリエ記述子を使用する。P 形フーリエ記述子とは、全曲率関数を指数関数の指数部にもつ複素関数を被展開関数とする離散フーリエ変換である。

P 形フーリエ記述子には、(1) 原曲線の端点と再生曲線の端点が常に一致する、(2) 曲線の平行移動、拡大、縮小に関して不変である、(3) 再生曲線は再生次数の増大に伴って原曲線に視覚的に近づく、(4) ごく少数の低域成分に原曲線のパターンクラス情報が含まれており特徴パラメータとして優れている、といった特徴がある。本稿では、これらの特徴のうち、(3) と (4) に注目し、ロボットのモーション波形の特徴量として P 形フーリエ記述子を用いることを考える。

複素平面上の点 $z(j) = x + iy$ (i は虚数単位) について、間隔 δ で等間隔な $n + 1$ 個の点列 $z(0), \dots, z(n)$ を考える。隣り合う点 $z(j), z(j + 1)$ を結んでできる折線図形の P 表現 $\omega(j)$ は、以下の式で求められる。

$$\omega(j) = \frac{z(j+1) - z(j)}{\delta} \quad (j = 0, \dots, n-1) \quad (3)$$

このとき折線図形の P 形フーリエ記述子 $c(k)$ は P 表現に対する離散フーリエ変換として、以下の式で計算できる。

$$c(k) = \frac{1}{n} \sum_{j=0}^{n-1} \omega(j) \exp(-2\pi i \frac{jk}{n}) \quad (4)$$

ここで、以下の式で表される $\hat{c}(k)$ を考える。

$$\hat{c}(k) = \begin{cases} c(k) & (k = 0, 1, \dots, n/2) \\ c(n+k) & (k = -n/2 + 1, \dots, -1) \end{cases} \quad (5)$$

P 形フーリエ記述子は低周波成分によって図形の概形を表現することができる。そこで本稿では、 $|k| \leq N$ なる低周波成分のみを用いた $\hat{c}_N(k)$ をモーション波形の特徴量として利用する。

$$\hat{c}_N(k) = \begin{cases} \hat{c}(k) & (-N \leq k \leq N) \\ 0 & (\text{otherwise}) \end{cases} \quad (6)$$

これをニューラルネットワークの入力信号および教師信号として用いることで、恒等写像学習を行う。

4. ニューラルネットワークによる操作平面の構築

前節の方法で得られる P 形フーリエ記述子による波形特徴量をニューラルネットワークに入力信号・教師信号として与えることで恒等写像学習を行い、操作平面を構築する。恒等写像学習のニューラルネットワークは 5 層で構成され、第 3 層のユニット数は入出力のユニット数よりも少ない構造を持つ (図 5)。本稿では、第 3 層のユニット数を 2 つとして 2 次元特徴空間を生成し、これを操作平面として利用する。

5. 操作平面からのモーション波形抽出

モーション波形の抽出は、操作平面でマウスクリックされた座標 (x, y) をニューラルネットワークの第 3 層に入力し、第 5 層から特徴量 $c_N(k)$ を取り出すことで行われる。

$|k| \leq N$ なる低周波成分のみを用いた $\omega_N(j)$ は折線図形の N 次の P 表現といい、以下の式で表される。

$$\omega_N(j) = \sum_{k=-N}^N \hat{c}_N(k) \exp(-2\pi i \frac{jk}{n}) \quad (7)$$

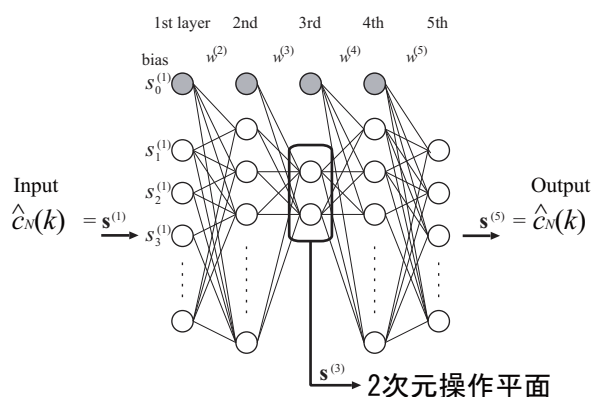


図 5: 恒等写像学習ネットワーク

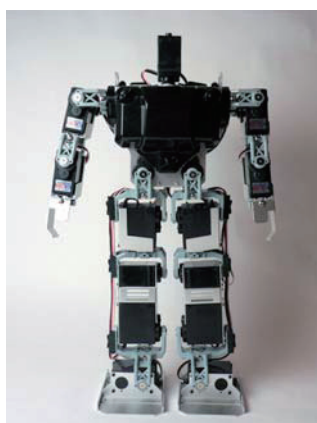


図 6: KHR-2HV

これを逆フーリエ変換することによって得られる曲線を、 N 次の再生曲線と呼ぶ。

$$z_N(j) = z_N(0) + \delta \sum_{r=1}^{j-1} \omega_N(r) \quad (8)$$

本稿では、 N 次の再生曲線によってロボットのモーションを生成する。

6. 評価実験

6.1. 使用するロボット

本評価実験では、近藤科学製の KHR-2HV(図 6) を用いた。KHR-2HV は 17 自由度を持つ。本稿では、KHR-2HV の腕を前後に動作させて歩行を模擬し、提案手法で歩行動作を直感的に操作できるかについて評価する。

6.2. 評価用操作平面の構築

本稿では、ロボットの腕を前後方向に動作させるための操作平面を構築する。操作平面の構築に利用するオノマトペを以下の手続きで選択した。

まず、20 個のオノマトペを付与した波形を生成し実際にロボットで動作させた。つぎに、それぞれの動作が、付与されたオノマトペを表現できているかアンケートを取った。アンケートは、5 段階リッカートスケールで行った。アンケート結果より、平均点 3.5 点以上の 10 個の波形を選び出し、それらの波形特徴 $\hat{c}_N(k)$ を用いて、操作平面を構築した。



図 7: 構築された操作平面と出力例

構築された操作平面を図 7 に示す。同図中には、「サクサク」、「ウロウロ」および、「サクサク」と「ウロウロ」の中間をクリックした際の波形例も同時に示した。この結果から、「サクサク」と「ウロウロ」の中間からそれぞれの特徴を持つ波形が得られていることが見て取れる。

6.3. 評価方法

被験者は構築した操作平面を用いてロボットを操作し、再生された動作に対して評価を行った。評価するオノマトペは操作平面の学習に用いた 10 のオノマトペ、および「ガンガン」と「サクサク」と「ウロウロ」のそれぞれの中間点から生成されたオノマトペとした。評価手法には 5 段階リッカートスケールを用いた。

また、ユーザビリティ評価指標の 1 つである SUS(System Usability Scale)⁵⁾ を用いて本システムのユーザビリティについて評価した。SUS とは、10 の設問を用いてユーザビリティを簡易的に調査する手法である。表 1 に SUS の設問項目を示す。SUS は、各質問に対する評点 (1~5 点) から下記の式で算出される。

$$SUS = 2.5 \times \sum_{i=1}^5 \{(x_{2i-1} - 1) + (5 - x_{2i})\} \quad (9)$$

ここで、 x_i は設問 i に対する評点である。SUS は、0 から 100 ポイントの値を取り、値が大きいくほど評価対象のシステムは高いユーザビリティを有していると判断できる。

以上の実験を 20 名に対して行った。

7. 結果・考察

学習に用いた 10 のオノマトペの評価結果を図 8 に、2 つのオノマトペの中間から生成された動作についての結果を図 9 に示す。

まず、学習に用いた 10 のオノマトペの操作平面への写像結果について見ると、それぞれの評点は、事前アンケート時よりも低下したのも見られたが、事前アンケート結果との間に有意な差は見られなかった。このことから、P 形フーリエ記述子および恒等写像学習による動作波形特徴量の抽出が適切に行えたと考えられる。

操作平面上の 2 つのオノマトペ間から生成された動作については、3.0 ポイント以上の評点を得ており、学習に用いたオノマトペとほぼ同等の結果が得られたが、学習用のオノマトペと比較して被験者の間で評価のばらつきが認められた。このことから、操作平面から得られる波形は、個々によって評価がわかれることが示唆された。

表 1: SUS の設問項目

1	このシステムをしばしば使いたいと思う
2	このシステムは不必要なほど複雑であると感じた
3	このシステムは容易に使えと思った
4	このシステムを使うのに技術専門家のサポートがいるかもしれない
5	このシステムにあるさまざまな機能がよくまとまっていると感じた
6	このシステムでは、一貫性のないところが多くあったとおもった
7	たいいていのユーザは、このシステムの使用方法について、素早く学べるだろう
8	このシステムはとても扱いにくいと思った
9	このシステムを使うのに自信があると感じた
10	このシステムを使い始める前に多くのことを学ぶ必要があった

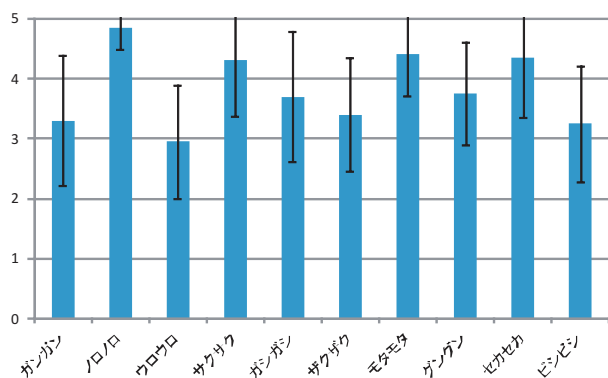


図 8: 学習に用いたオノマトペに対する評価結果

つぎに、動作平面上的疎な部分（「ガンガン」と「ガシガシ」の中間点、および「ザクザク」と「ノロノロ」の中間点）から生成した動作と、密な部分（「ビシビシ」と「ウロウロ」の中間点）から生成した動作との間で評点に差異があるか調査したところ、これらの3つの動作間に有意な差異は認められなかった。このことから、操作平面上的疎密に関係なく、ユーザの印象に比較的即した動作が生成できる可能性が示唆された。

最後に、SUSの結果について考察する。図10に各設問の評点を示す。SUSは、66.4(±14.9)であり、本システムは実用に足る可能性があると考えられる。アンケート後の口頭インタビューにおいて、「このシステムを使う場面を想定しにくい」との意見が多く見られ、これがSUSの評点を下げた原因と考えられる。その一方で、各設問の回答結果を見ると、被験者はこのシステムを「使用方法を学ぶことが簡単」であり「容易に使うことができる」と評価している。これらのことから、本システムの用途をより明確にできれば、すなわち、ロボットのモーション編集という作業が使用者にとってより身近な作業であれば、より高い評価が得られたものと推察される。

8. おわりに

本稿では、ニューラルネットワークを用いた操作平面によるロボットの直感的な動作生成システムを開発した。操作平面から出力された波形を用いてロボットを動作させることにより

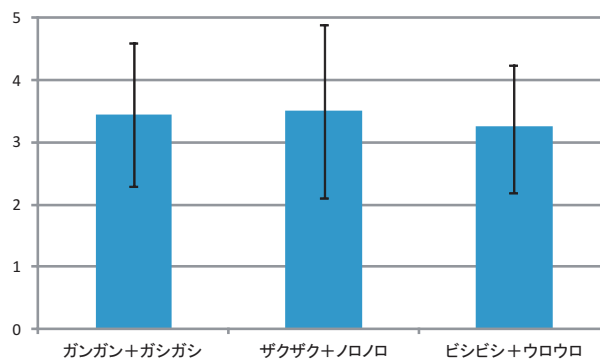


図 9: 中間領域から作成した波形に対する評価結果

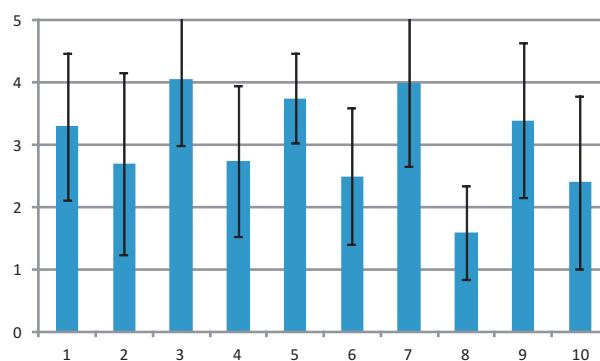


図 10: 評価項目ごとの評点

ロボットの動作編集に馴染みのない人でも容易にロボットモーションの編集を行うことができ、モーション編集の認知的負担を軽減できたのではないかと考える。

今回のシステム構築では、歩く動作の腕の前後方向の動きに着目したが、腕の前後方向の動きだけでは可動範囲や可動数の少なさのため、オノマトペの持つ直感的な印象を十分に表現できていない可能性もある。今後、足や肘などの他の関節も同時に操作することによって、より直感的な操作が可能となるシステムの開発を進めていく。

参考文献

- 1) 小松孝徳, 秋山広美: ユーザの直感的表現を支援するオノマトペ表現システム, 電子情報通信学論文誌, Vol.J92-A, No.11, pp.752-763, 2009.
- 2) 上坂吉則: 開曲線にも適応できる新しいフーリエ記述子, 電子情報通信学論文誌, Vol.J67-A, No.3, pp.166-173, 1984.
- 3) 加納政芳, 加藤昇平, 伊藤英則: 判別の難易度に基づく類似箇所検出の高速化, 情報処理学会論文誌, Vol.42, No.11, pp.2689-2698, 2001.
- 4) 岩田彰: デジタル信号処理, コロナ社, 1995.
- 5) J. Brooke: SUS: A “quick and dirty” usability scale, P. W. Jordan, B. Thomas, B. A. Weerdmeester and McClelland (eds.), Usability Evaluation in Industry, pp. 189-194, Taylor & Francis, London, UK, 1996.