

# モーションキャプチャによる 舞踊の「止め」のニューラルネット検出に関する検討

## A study on neural-network detection of stops in dance using motion capture

神 優貴<sup>†</sup>      鈴木 元樹<sup>†</sup>      塩谷 浩之<sup>†</sup>  
Yuuki Jin      Genki Suzuki      Hiroyuki Shioya

### 1 はじめに

北海道を代表する一大イベントの一つによさこいソーラン祭りがある。よさこいソーラン祭りは例年6月中旬に北海道札幌市にて開催され、全国から出場チームや観光客が集まるため大きな経済効果が期待される<sup>\*1</sup>。よさこいソーランは市民参加型の集団舞踊の一つであり、各チームが毎年新しい曲に対して多様な舞踊ジャンルに基づく振付が行われている [1]。また、よさこいソーランは集団舞踊であることから踊りの振付を覚えるだけでなく、チームとしての動きの同期性も重要とされている。特に、素早い振付を一瞬だけ停止する「止め」動作のタイミングを全体で揃えることがチームとしての舞踊の上手さに紐づいている。この動作の習得には熟練者による指導が効果的であるが、新型コロナウイルスの感染拡大で対面形式での練習が難しい。このため、「止め」動作のタイミングに基づき舞踊動作を評価し、個人の練習活動を支援する技術構築が求められている。

この課題解決には、ICT技術を活用した舞踊動作のセンシングおよび解析システムの構築が有効であると考えられる。これまでに、舞踊の指導教育を目的として、モーションキャプチャ(以降、MoCap)を用いた様々な舞踊における動作の解析が種々行われてきた [2-5]。特に文献 [5]では、よさこいソーランの舞踊動作をMoCapによりデータを取得し、ニューラルネットワーク(以降、NN)による「止め」の検出が行われている。しかしながら、従来手法は速さに関する動作の特徴にのみに基づいているため、「止め」とそれに類似した一部の「止め以外」の動作を分類することが難しい。一般に「止め」動作は動きのキレに影響することから、手の加速度に着目した解析が重要であると考えられる。

そこで、本稿では、MoCapによる動作データから速さに加えて新たに加速度に関する特徴を「止め」の検出に用いる手法を提案する。速さ



図1 「止め」を行った様子

および加速度で構成される特徴を用いて階層型NNで学習することで「止め」の動作の高精度な検出を実現する。これにより、「止め」動作のタイミングに基づいた舞踊動作の評価への貢献が期待される。

### 2 MoCapによる舞踊動作データの記録

本章では、検出対象の舞踊動作およびMoCapによる動作データの取得方法について説明する。

#### 2.1 舞踊における「止め」

よさこいソーランには、踊りを上手く見せるための動作が存在し、その一つに素早い振付を一瞬だけ停止するものがある。本研究ではこれを「止め」と呼ぶ。「止め」を行った様子を図1に示す。「止め」をタイミング良く行うことで動きのキレを良く見せることができる。

一般に、振付の中でインパクトを出す瞬間に「止め」が行われる。また、集団舞踊ではチーム全体で「止め」のタイミングを揃えることで、チームの動きの同期性を高めることが可能となる。このように、「止め」はよさこいソーランにおいて非常に重要な動作である。

<sup>†</sup> 室蘭工業大学 Muroran Institute of Technology

<sup>\*1</sup> <https://app.yosakoi-soran.jp/news/view/324>



図2 PNを被験者に装着した様子

## 2.2 MoCap を用いた動作データの取得

舞踊動作の記録には、図2に示す NOITOM 社の Perception Neuron 2.0 (以降、PN)<sup>\*2</sup>を用いて行う。PNはジャイロスコープや加速度などの慣性を計測する複数の小型センサにより構成される MoCap である。

PNは多種の MoCap と比して安価であり、慣性式の MoCap であることから場所を選ばず利用可能である。「止め」を行う際は手先を停止させることに意識を向けることが多い。本研究では、図2に示すようにPNを被験者の右腕、右手首、および右手の甲に装着して「止め」を含む舞踊動作を記録する。本研究では、取得データのうち右手の甲の移動軌跡データを用いて「止め」の検出を行う。

## 3 NN を用いた「止め」の検出

本章では、NNを用いて「止め」の検出を行う。まず、PNから取得した右手の甲の移動軌跡データを用いて速さや加速度の大きさを構成される動作特徴を算出する。続いて、動作特徴を用いてNNの学習を行い、「止め」を行ったフレームを検出する。

### 3.1 動作特徴の算出

「止め」は素早い振付を一瞬だけ停止させる動作であるため、手の速さや加速度に着目した特徴を算出する。まず、右手の甲の移動軌跡データ  $f_{i,j}$  ( $i = 1, 2, \dots, N; j \in \{x, y, z\}$   $N$ は移動軌跡の系列数) から速さおよび加速度をそれぞれ算出する。

#### <速さの算出>

PNに搭載されているセンサーは非常に高感度であるため、記録された移動軌跡データにはノイズなどの高周波成分が含まれている。速さおよび加速度の算出には移動軌跡データに対して微分が必要であるため、事前に移動軌跡対

し平滑化を適用し、ノイズ除去を行うことが望ましい。そこで、 $f_{i,j}$  に対して移動平均処理を適用する。

$$f_{i,j}^{\text{ma}} = \frac{1}{2n+1} \sum_{k=i-n}^{i+n} f_{i,j} \quad (1)$$

ただし、 $n$ は任意の自然数である。続いて、下式の微分の5点近似式により速さ  $f'_{i,j}$  を算出する。

$$f'_{i,j} = \frac{f_{i-2h,j}^{\text{ma}} - 8f_{i-h,j}^{\text{ma}} + 8f_{i+h,j}^{\text{ma}} - f_{i+2h,j}^{\text{ma}}}{12h} \quad (2)$$

ただし、 $h$ は微分の微小な幅であるが、1フレームごとの変位を求めるため本稿では  $h = 1$  とする。なお、算出した速度  $f'_{i,j}$  にも高周波成分が発生する場合、以降の加速度の算出にも微分を適用するため、速度  $f'_{i,j}$  の推移も滑らかにしたい。しかしながら、速度  $f'_{i,j}$  は正負の変化が激しい場合があり、移動平均処理を適用すると大きさの特徴を損なう恐れがある。そこで、算出した速度  $f'_{i,j}$  から以下の二乗平均平方根を算出する。

$$f_{i,j}^{\text{rms}} = \sqrt{\frac{1}{2n+1} \sum_{k=i-n}^{i+n} (f'_{i,j})^2} \quad (3)$$

続いて、3次元の速度から1次元の速さを下式より算出する。

$$f_i^{\text{abs}} = \sqrt{(f_{i,x}^{\text{rms}})^2 + (f_{i,y}^{\text{rms}})^2 + (f_{i,z}^{\text{rms}})^2} \quad (4)$$

最後に、式(4)で得られた結果に対し、式(2)を再度適用することで速さ  $f_i^{\text{abs}}$  の変化をより滑らかにし、また0から1の範囲で正規化することで速さ  $v_i^{\text{norm}}$  を算出する。この時の速さ  $v_i^{\text{norm}}$  の推移例を図3に示す。図中の四角印、丸印、三角印はそれぞれ速さが極小または0に近づくフレームである。ただし、丸印、三角印のフレームは目視で確認した「止め」を行ったフレームである。丸の「止め」は動作を停止させた直後に動作が再開されており、一方で、三角の「止め」は動作を停止させた後しばらく停止状態が続く。本研究では、それぞれ「止め1」および「止め2」として定義する。また、四角は「止め以外」と定義する。

#### <加速度の算出>

算出した速さ  $v_i^{\text{norm}}$  を用いて下式の微分の5点近似式により加速度の大きさ  $a_i$  を算出する。

$$a_i = \frac{v_{i-2h}^{\text{norm}} - 8v_{i-h}^{\text{norm}} + 8v_{i+h}^{\text{norm}} - v_{i+2h}^{\text{norm}}}{12h} \quad (5)$$

ただし、1フレームごとの変位を求めるため、本稿では  $h = 1$  とする。その後、式(5)で得られた

\*2 <https://www.noitom.com/>

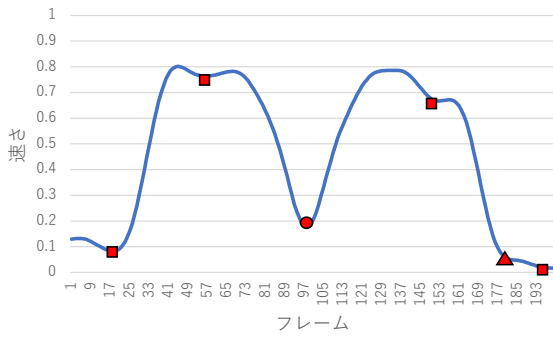


図3 速度の推移例

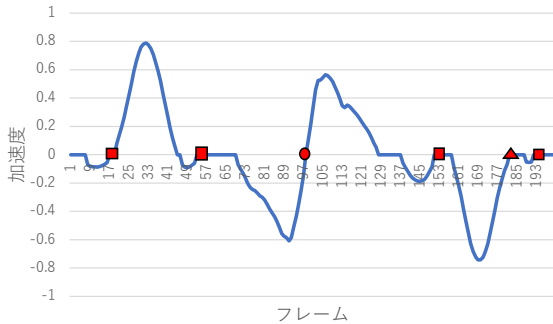


図4 加速度の推移例

結果に対して、 $-1$ から $1$ の範囲で正規化することで $a_i^{\text{norm}}$ を算出する。しかしながら、時系列データであるため $a_i^{\text{norm}}$ が厳密に $0$ になるフレームが存在するとは限らない。そこで、 $a_i^{\text{norm}}$ に対して式(6)を適用し、加速度 $a_i^{\text{fil}}$ を算出する。

$$a_i^{\text{fil}} = \begin{cases} 0 & -0.05 \leq x \leq 0.05 \\ a_i^{\text{norm}} & \text{else} \end{cases} \quad (6)$$

$a_i^{\text{fil}}$ を算出することで、速度が極小となるフレームと $0$ に近づくフレームを同時に算出可能とする。この時の加速度の推移例を図4に示す。図中の四角、丸、三角は加速度が負から $0$ となるフレームであり、それぞれ図3の速度が極小または $0$ に近づくフレームと対応している。

#### <学習用データセットの作成>

速度および加速度を用いてNN学習用データを作成する。図3および図4の四角、丸、三角となる瞬間のフレーム番号を $t$ とおく。このとき、 $v_i^{\text{norm}}$ と $a_i^{\text{fil}}$ を用いて、「止め1」および「止め2」の瞬間の前後のフレームに関して特徴の総和を、式(7)~(10)により算出することで4素子用の学習データ $\mathbf{r}_t = [V_{\text{back}}(t), V_{\text{forward}}(t), A_{\text{back}}(t), A_{\text{forward}}(t)]^T$ を取得する。

$$V_{\text{back}}(t) = \sum_{k=t-25}^t v_k^{\text{norm}} \quad (7)$$

$$V_{\text{forward}}(t) = \sum_{k=t}^{t+25} v_k^{\text{norm}} \quad (8)$$

$$A_{\text{back}}(t) = \sum_{k=t-25}^t a_k^{\text{fil}} \quad (9)$$

$$A_{\text{forward}}(t) = \sum_{k=t}^{t+25} a_k^{\text{fil}} \quad (10)$$

### 3.2 ニューラルネットによる「止め」の検出

本節では、前節で算出した動作特徴を用いて順伝搬型NNにより「止め」を検出方法について述べる。提案手法で用いるNNは、階層型ニューラルネットで入力と中間および出力層から構成される。中間層は1層とする。前節で算出したフレームの動作特徴をベクトルとして $\mathbf{r}_t$ を入力に用い、出力では検出結果の判定を示すクラスの数 $c=1, \dots, C$ を対応させる。入力・中間・出力層のノード数をそれぞれ $M, L, C$ とする。 $x_0 = 1$ によるバイアスを設定する。動作特徴となるベクトル $\mathbf{r}_t$ を $(x_1, \dots, x_M)$ に対応させる。 $l$ 番目の中間層のノードにおける以下の伝搬式が得られる。

$$h_l(\mathbf{r}_t) = h\left(\sum_{m=0}^M w_{l,m}^1 x_m\right) \quad (11)$$

ただし $w_{m,l}^1$ は入力層と中間層間の重みとする。活性化関数 $h(\cdot)$ として本実験ではシグモイド関数を用いる。バイアスを含む中間素子の出力と中間層と出力層間の重み $w_{l,c}^2$ との線形和によって $y_c(\mathbf{r}_t)$ が、出力層における $c$ 番目の出力となる。

$$y_c(\mathbf{r}_t) = \sum_{l=0}^L w_{c,l}^2 h_l(\mathbf{r}_t) \quad (12)$$

さらに、出力層の全出力を用いたソフトマックス関数で変換することで、クラスを示す確率値として以下の $s_c$ が得られる。

$$s_c = \frac{\exp\{y_c(\mathbf{r}_t)\}}{\sum_{c'=1}^C \exp\{y_{c'}(\mathbf{r}_t)\}} \quad (13)$$

NNの学習では、前節で作成した学習用データセットを用いて交差エントロピー誤差を最小化する各層の重みとバイアスを決定する。最終的に、学習を行ったNNに対して学習用データ形式と同様のテストデータを入力することで、「止め1」、「止め2」、および「止め以外」の分類が行われる。

表1 学習用データセット数の内訳

	訓練	テスト
「止め1」	6	5
「止め2」	6	5
「止め以外」	11	13

表2 正解率

	PM(%)	CM1(%)	CM2(%)
全体	<b>93.5</b>	82.6	<b>93.5</b>
「止め1」	<b>100.0</b>	75.0	91.7
「止め2」	<b>100.0</b>	81.7	<b>100.0</b>
「止め以外」	87.1	88.5	<b>91.7</b>

## 4 実験

本実験では、舞踊の経験8年の熟練者1名および舞踊未経験の初心者5名の動作データから、前章に基づき学習用データセット(表1)を構築する。ただし、今回は $n=5$ とする。提案手法のNNは入力素子数4, 中間層1層, 中間素子数100, 出力素子数3で構成されている。ただし、学習率は0.01とし、活性化関数はシグモイド関数を用いた。提案手法のパラメータは汎化性能が最大となる際の値を採用した。

提案手法(PM)の有効性を検証するため、速さのデータのみを用いて分類する従来手法(CM1)[4]と、提案手法と同様の特徴量を代表的な機械学習手法のSVM[6]で学習する手法(CM2)との精度を比較する。ただし、本実験では、データセット数が少ないため、各手法の分類精度を2分割交差検定により算出する。CM1とCM2のパラメータは正解率が一番良い時の値を採用した。従来手法のNNは入力素子数51, 中間層1層, 中間素子数100, 出力素子数3で構成されている。ただし、入力素子数は提案手法と異なるが、学習に用いるデータの内訳は提案手法と同様である。SVMではRBFカーネルを利用する。ただし、学習に用いるデータセットは提案手法と同様である。評価指標は正解率を用いた。

全ての手法の分類結果を表2に示す。表2より、提案手法の「止め」の正解率が他の手法よりも高いことが示されている。具体的に、まず、PMはCM1と比較することで加速度に基づく特徴量を導入することの有効性が示された。特に、「止め1」と「止め2」の正解率が大きく向上していることが確認できる。この結果から、提案手法では従来手法と比較して大幅な精度向上に成功したと言える。

続いて、PMとCM2を比較すると全体の正解率が同等であることが示された。また、PMの方が「止め1」の分類精度が高く、CM2の方が「止め以外」の分類精度が高いことが明らかとなった。「止め」の検出という観点から評価すると、「止め1」の正解率の高い提案手法の方がSVMよりも「止め」の検出に有効であると言える。以上により、提案手法の有効性が示された。

## 5 まとめ

本稿では、モーションキャプチャによる舞踊の「止め」をNNにより検出する手法を提案した。速さと加速度に基づく特徴量を用いてNNを学習することで高精度に「止め」の検出を行った。実験により、提案手法の有効性が示された。

## 参考文献

- [1] 佐美靖, 森谷きよし, “YOSAKOI ソーラン祭り参加による運動量増加と体力の向上,” 日本生気象学会雑誌, 2005, vol. 42, no. 4, pp. 145-157.
- [2] 飯野友里恵, 森谷友昭, 高橋時市郎, “ストリートダンス動作の分析とダンス指導への応用(映像表現フォーラム),” 映像情報メディア学会技術報告, 2011, vol. 35, no. 14, pp. 49-52.
- [3] 田中祐典, 齋藤剛, “モーションキャプチャを用いたダンス上達支援システムの開発,” 第75回全国大会講演論文集, 2013, no. 1, pp. 225-226.
- [4] 柴田傑, 玉本英夫, 海賀孝明, 横山洋之, “身体動作の3次元計測によるリアルタイム舞踊学習支援システム,” 日本バーチャルリアリティ学会論文誌, vol. 17, no.4, pp. 353-360, 2012.
- [5] 神 優貴, 柴田 傑, 塩谷 浩之, “慣性式 MoCap による舞踊における「止め」の検出に関する検討,” 情報処理学会北海道シンポジウム 2020 予稿集, 2020.
- [6] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine learning*, vol. 20, no. 3, pp. 273-297, 1995.