

疾走動作の段階的改善を目的とした教師なし学習による類似動作検出法の検討 Study on the Method of Detecting Similar Motion with Unsupervised Learning for Phased Improvement of Running Motion

徐 燦振[†]
Chanjin Seo

佐羽内 雅人[†]
Masato Sabanai

小方 博之[‡]
Hiroyuki Ogata

大谷 淳[†]
Jun Ohya

1. はじめに

近年、画像処理分野において機械学習を用いて人間の運動動作を評価するシステムが注目を集めている。具体的には、(1)人間の関節情報といった運動の動作情報を入力とし、(2)出力として予め付けられた運動の動作の客観的な点数になるように動作に関する機械学習を行い、(3)その学習結果を用いて、評価したい運動の動作に対して利用者に点数をフィードバックするような流れで研究が行われている。本研究では、このようなシステムを自動動作評価システムと呼ぶ。このようなシステムの従来方法として、Pirsiavashら[1]の研究が挙げられる。これらの方法を用いると、動作に対する点数を出力することで動作の良し悪しを客観的に理解できる。特に[1]では動作の点数の低い箇所を全て指摘することで動作の改善をさせる手法を提案している。

しかし、指摘された多数の箇所を同時に改善することは一般的に困難である。このような問題点を解決するためには多数の改善箇所を全て指摘するよりも、現在評価された運動動作よりも 1 ランク評価が高くなるような少数の改善箇所を指摘した方がより効果的であると考えられる。そして、この操作を繰り返すことで徐々に動作をランクアップし、最終的には点数の低い箇所を全て改善するようなシステムの方が現実に即していると考えられる。本論文ではこのような段階的動作改善を目指すための一歩として、教師なし学習によって汎用的な特徴量が抽出できると仮定し、それを用いて類似する運動動作の検出を行うことを検討する。教師なし学習を行う理由として、従来方法では動作評価のためのスコアリングを重視しているため特定の特徴量しか得られないが、事前知識を極力使用しないで汎用的な特徴量を得ることができ、かつこの特徴量から従来方法では得られない特徴を用いた動作改善を促すことができると考えられる。そして得られる特徴量を解析することで類似する動作の集合を得ることができ、これを用いることで本研究の目指す 1 ランク上の動作改善に必要な改善箇所をフィードバックができるのではないかと期待している。また、類似する動作を検出することで、検出された集合内において 1 ランク上の動作を目指した動作改善を促すことで段階的改善が円滑に進められると検討している。

ゆえに、本論文では上述の特徴量から類似する動作を検出するために、まず教師なし学習手法の一つである Autoencoder を使用して生成モデルを得る。次に、得られた生成モデルの情報に対してクラスタ分析を行い、得られるクラスタの特徴を 2 章で述べる学習済みモデルにおける入力寄与度を用いて考察していく。また運動動作の評価の題材に、短距離走における疾走動作に着目し、本論文の提案手法から得られた動作の集合について考察していく。

[†] 早稲田大学創造理工学部総合機械工学科

[‡] 成蹊大学理工学部システムデザイン学科

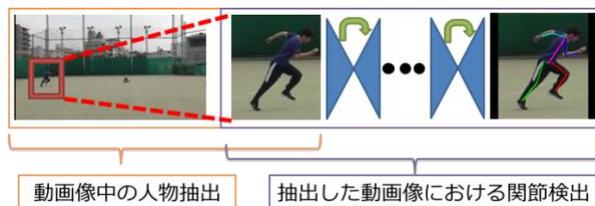


図 1 疾走動作の関節検出手法の概要図

2. 提案手法

2.1 疾走動作の関節検出手法と特徴量抽出手法

前処理として疾走動作の動画像から人物を抽出し、関節を 16 点検出する。動画像中の人物の抽出には YOLOv3[3] を、関節の検出には Stacked Hourglass Network[4] を用いた。

$$\bar{q}(j) = \frac{\bar{p}(j) - \bar{p}(0)}{|\bar{p}(j) - \bar{p}(0)|} \quad (1)$$

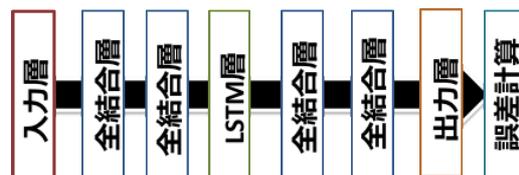


図 2 Autoencoder によるネットワーク構造

次に、前処理で検出した関節情報から疾走動作の汎用的な特徴量を事前知識なしで抽出するために教師なし学習を行う。まず頭部の関節座標を基準として式(1)のような正規化を行う。ここで $\bar{p}(j)$ は関節 j の座標で、 $j=0$ は頭部である。

そして、各関節の $\bar{q}(j)$ を入力として Autoencoder による学習を行う。ネットワーク構造を図 2 に示す。入力層と出力層はそれぞれ 30 次元、全結合層は各 100 次元である。また、中間層では 50 次元にデータを圧縮している。本手法では Autoencoder の中間層を全結合層ではなく、LSTM 層としているのが特徴である。疾走動作は時間に対して連続的な動きであるので LSTM 層を用いることで時系列の要素を取り入れた特徴量を得ることができると考えられる。

2.2 入力寄与度算出アルゴリズム

深層学習において入力出力にどのような影響を及ぼすかを分析することは、入力のどの要素が学習モデルの出

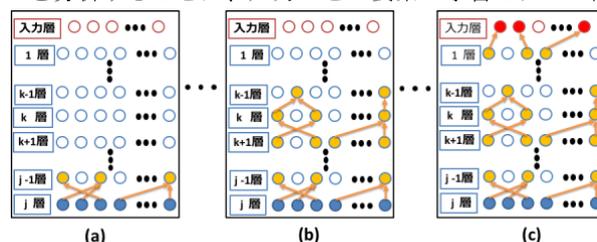


図 3 入力寄与度検出方法の概要

力に関わるのかを検証する上で重要である。本節では、図 2 のモデルで得た学習モデルを用いてそのモデルにおける入力された各関節の入力寄与度を計算する手法について説明する。はじめにニューラルネットワークの最深の j 層の各ノードの値の構成に最も寄与している $j-1$ 層のノードを見つける。次に見つけたノードに 1 点を付け、その他のノードには 0 点を付ける。図 3(a) のように $j-1$ 層全てのノードの点数付けを行った後、 $j-1$ 層において 0 点だったノードを除いたノードに接続された $j-2$ 層のノードで同様の点数付けを行う。入力層に到達するまで図 2(b) のように同様な操作を繰り返す。その結果、図 3(c) のように入力層における寄与ノードが得られる。この方法より、生成されたモデル内にて入力された疾走動作においてどの関節が最もその動きに寄与しているかを観察する。

3. 計測実験と結果

疾走動作を計測するにあたり、二十歳前半の健常な男性被験者 6 人に 15m を全力で走ってもらった。被験者 6 人の内、2 人は陸上経験者であり、4 人は未経験者である。15m 全体が写るようにカメラを図 4 の位置に配置した環境で実験を行った。被験者の疾走動作の撮影には 60fps のビデオカメラを用いた。また、被験者には一人あたり 2~4 回走ってもらい、計 19 本の疾走動作の撮影映像を取得し、これを 2 章で述べた手法で解析した。LSTM 層で出力されるデータを階層的クラスタリングした結果、大まかに 4 つのクラスタが存在することがわかった。

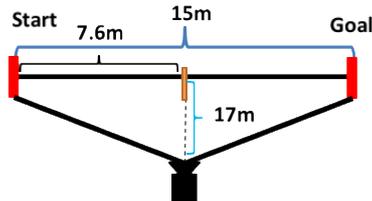


図 4 実験環境の概要図

4. 考察

3 章で得た 4 つのクラスタについて、2 章で述べた入力寄与度を、鈴木らの疾走動作に関する評価方法[2]と比較することで考察していく。各クラスタにおける各関節の入力寄与度について寄与が最も高い項目を図 5 に示す。疾走動作を鈴木らの評価方法で点数付けし、クラスタ別に分布を示したものを表 1 に掲げる。まず、図 5 によればクラスタ 0 の被験者は上肢と左下肢の y 軸方向の入力寄与が大きく、

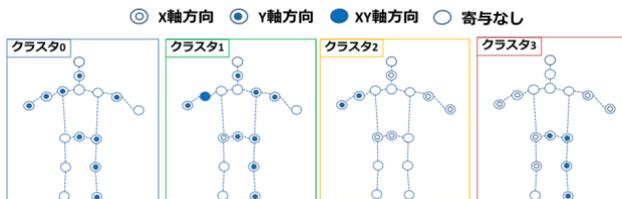


図 5 クラスタ間の入力寄与の違い

他クラスタよりも上下動が大きく効率の悪い疾走をしていると考えられる。実際、表 1 を見るとクラスタ 0 の被験者は鈴木らの評価法でも得点が低い。次に、クラスタ 1 は評価点数としては 4 点と 8 点付近に 2 つのピークをなしている。このことからクラスタ 1 では鈴木らの評価で疾走動作

表 1 各クラスタと疾走動作の点数評価との関連表

	0点	1点	2点	3点	4点	5点	6点	7点	8点	9点	10点
クラスタ0	17%	0%	25%	41%	15%	0%	0%	0%	0%	0%	0%
クラスタ1	0%	0%	2%	11%	20%	12%	12%	5%	20%	3%	10%
クラスタ2	0%	0%	0%	13%	3%	29%	4%	9%	5%	3%	30%
クラスタ3	0%	0%	0%	2%	10%	25%	51%	3%	6%	0%	0%

の良いものと悪いものが複合していると考えられる。クラスタ 2 においても点数が同様に分散しており、クラスタ 1 のような違いがあると考えられる。クラスタ 1 の低い点数の動作はクラスタ 0 で見られた入力寄与に近いことから、低得点の理由としてクラスタ 0 と同様のことがいえる。だが、クラスタ 2 の低得点の動作は x 軸方向の動きが多いことから疾走の推進力があると考えられるが、鈴木らの評価法では点数が低い。このことから点数の低い入力寄与はクラスタ 0 のようなパターンだけではないと考えられる。クラスタ 1 と 2 の高いパターンについても入力寄与の類似点も見られるが違いも確認できる。これは疾走タイプの違いまたは疾走の癖の違いが関係しているのではないかと考えられる。クラスタ 3 については表 1 より鈴木らの評価において点数が中間的な被験者が集まっている。このクラスタでは x 軸方向の入力寄与は多く見られることから、疾走ための推進力はあると考えられる。高得点の動作に見られるような入力寄与は少ないものの、今後上手く動作改善を促すことで疾走動作の評価を上げられる可能性は高いと考えられる。

以上のことから、教師なし学習で得た分類は鈴木らの評価法からも理解できるという点で妥当性を有すると考えられる。また、鈴木らの評価では低得点のものでも、提案手法では複数のクラスタに分かれたことから、疾走動作のデータを増やしていくことで、1 次元的な評価法ではわからないような疾走動作のタイプや癖を検出し、きめ細かなアドバイスができるようになると考えられる。

5. おわりに

本論文では段階的な動作改善に必要な疾走動作に関する汎用的な特徴量を、中間層を LSTM 層に置き換えた Autoencoder のネットワーク構造を用いることで得ることができた。この特徴量についてクラスタ分析を行なった結果、大まかに 4 つのクラスタが見られた。入力寄与度の概念を導入して各クラスタの疾走動作の傾向を理解する方法を提案し、鈴木らの評価法との比較によって、分類の妥当性を確認した。

謝辞

本研究結果の一部は早稲田大学重点領域研究機構ヒューマンパフォーマンス研究所の活動の一環として行われたものである。

参考文献

- [1] H. Pirsiavash, C. Vondrick, and A. Torralba, "Assessing the Quality of Actions", European Conference on Computer Vision (ECCV), pp.556-571, 2014.
- [2] 鈴木 康介, 友添 秀則, 吉永 武史, 梶 将徳, 平山 公紀, "疾走動作の観察的動作評価法に関する研究", 体育科教育学研究, Vol.32, No.1, pp.1-20, 2016.
- [3] J. Redmon, and A. Farhadi, "YOLOv3: An Incremental Improvement", arXiv:1804.02767v1, 2018.
- [4] A. Newell, K. Yang, and J. Deng "Stacked Hourglass Networks for Human Pose Estimation", arXiv:1603.06937v2, 2016.