

RNN を利用した動作解析手法の検討

A Study on Motion Analysis Method using Recurrent Neural Network

酒井 正人^{*1}, 越智洋司^{*2}

Masato SAKAI^{*1}, Youji OCHI^{*2}

^{*1}近畿大学大学院 総合理工学研究科

^{*1}Graduate School of Science and Engineering, Kindai University

^{*2}近畿大学 理工学部

^{*2}Faculty of Science and Engineering, Kindai University

Email: ochi@ele.kindai.ac.jp

あらまし：機械学習を利用し動作の識別を自動的に行うような動作解析手法では、動作の特徴量を考慮する必要がない反面、学習内容のブラックボックス化による教育支援への弊害という欠点がある。

本研究では、モーションセンサデバイスである Kinect により得られる関節座標データと、時系列データの特徴量を自動的に学習することができる RNN (Recurrent Neural Network) によって行われる動作識別の結果を利用して、学習した動作の推測・可視化できるかの検証を目的とする。

キーワード：学習結果、動作識別、機械学習、特徴量、Kinect, RNN

1. はじめに

人の動作を指導する際、そのアドバイスは抽象的になることが少なくない。動作解析によって動作を定量的に評価しアドバイスを行えば、初心者動作習得を容易化できると考えられる。動作識別をする際は、学習支援に繋がる動作モデルを考え、そのモデルを表現する特徴量を定義する必要がある⁽¹⁾。動作認識の精度はそのモデルや特徴量に対応可能なものに限定される。そのため、識別対象の動作を変更あるいは増やすと、その度に特徴量の抽出方法を見直す必要性があり動作識別が容易ではない。

そこで本研究は、機械学習の一種であり時系列データから特徴量を自動的に学習可能な RNN (Recurrent Neural Network) に着目し、Depth センサ搭載型カメラデバイスの Kinect により得られる人体関節座標データに対して RNN を適用し、リアルタイム識別を可能とするシステムを試作した。我々はラジオ体操の 5 パターンを対象に RNN を適用し、高い識別精度からその有効性を示したが⁽²⁾、学習支援の観点では、学習内容がブラックボックス化されるため、識別判断の基準となる動作特徴量が把握できず、学習内容が動作指導に利用できないなどの弊害を持つ欠点がある。RNN の学習内容の可視化を行うツールを用いて RNN の隠れ層の反応を見る手法⁽³⁾があるが、このツールが対応するのはテキストデータのみで座標データは対応できない。本発表では、本システムの動作識別結果を分析して、学習者の誤り行動を解析できるかを検証する。

2. 動作識別手法

提案手法は、動作データの取得 (2.1)、入力・出力データの構築 (2.2)、RNN による学習 (2.3)、の 3 つの手順で構成する。以下に各手順の詳細を記す。

2.1 動作データの取得

動作データの取得手順を以下に記す。「動作データ」とは、Kinect から単位フレーム (30fps) あたりに送られる人体 25 関節の各 3 次元 (x,y,z) 座標値 (3×25=75 種類の座標値)、その時のフレーム番号、教師ラベルをまとめた時系列データを指す。

手順 (1) 高さ 115cm の所に Kinect を設置し、Kinect から 230cm 離れた場所に被験者が立つ。

手順 (2) 識別対象の全種類の動作データを 1 ファイルにまとめて取得するために、任意の時間毎に動作と教師ラベルを共に変更しながら順番に動作データを取得する。「教師ラベル」とは、RNN で動作パターンを識別するための指標である。例えば、動作 1 種類目の時は教師ラベルが「1」、動作 2 種類目の時は教師ラベルが「2」というような数値で表される。

手順 (3) 取得した動作データを csv 形式で保存する。

2.2 入力・出力データの構築

入力・出力データの構築手順を以下に記す。ここでいう「入力データ」とは、動作データ上の動作種類毎に一定フレーム数で分割していき、フレーム番号と教師ラベルの列を取り除いた関節座標のみのデータ群を指し、単位入力データ (学習用データ 1 個) は分割フレーム数分の時系列データとなる。また、「出力データ」とは、単位入力データに対する教師ラベルの値 1 つのみのデータを指す。

手順 (1) 2.1 節で保存した動作データを読み込む。

手順 (2) 入力・出力データの構造となるように整形し、入力・出力データ群にまとめて構築する。

2.3 RNN による学習

動作学習手順を以下に記す。尚、本研究では RNN において長い時系列データの扱いが可能な LSTM (Long Short-Term Memory) という手法を利用する。

手順 (1) 2.3 節で構築した入力・出力データを用いて RNN で動作学習を行う。
手順 (2) 学習が完了したら、学習結果を保存する。
上記の学習モデルを図 1 に示す。

3. 試作システム

3.1 リアルタイム動作識別システム

2.3 節で保存した学習結果を利用して新たなデータに対してリアルタイムに動作識別を可能とするシステムを開発した。「新たなデータ」とは、2.1, 2.2 節と同様にして新たに構築した入力・出力データを指し、認識率などを求める際は学習の際に使用した動作データに含まれない被験者から取得する。ユーザインタフェースを図 2 に示す。識別結果としては、識別したい動作種類の数（教師ラベルの数）に対応する確率の値を、Softmax 関数によりリアルタイムに表示する。図 2 は動作 2 についての出力の例であり、この場合は新たな入力データが動作 2 のデータに一番近いと機械が判断していることになる。

3.2 認識分布を利用した動作解析

一般的な機械学習の研究ではその認識精度が重要となるが、学習者をユーザとする場合、誤認識の要因には学習者の動きに誤った動きが含まれていることが想定できる。つまり、学習支援を対象とする場合は、分類器の精度の他に学習者に起因する精度低下の要因を分析することが重要となる。学習支援システムとしては、その要因を明らかにすることで学習者への指導が可能となる。

本研究では、認識結果における認識結果の分布に着目する。図 3 は、ラジオ体操の 5 パターンについて学習したシステムにおいて、動作 4 の動きに対して、正しく認識できた場合と誤った場合の認識分布である。誤って認識した場合には動作 1 ならびに動作 3 として認識した割合が高くなっており、これは各動作の類似性が関与していると考えられる。

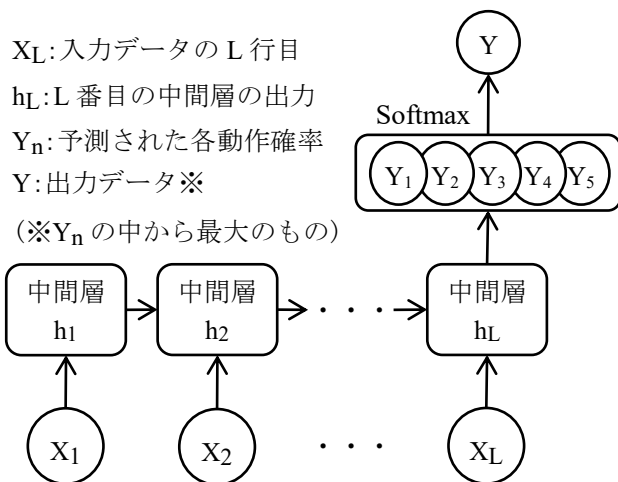


図 1 RNN (LSTM) による動作学習モデル



図 2 リアルタイム動作識別システム

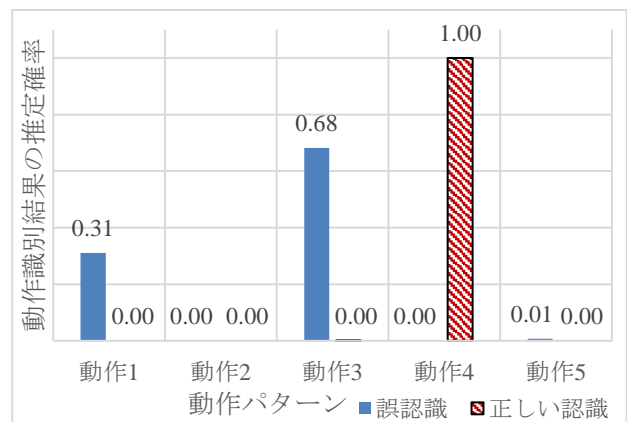


図 3 認識パターンの分布

4. おわりに

本研究では Kinect により得られる関節座標データと RNN を利用した動作識別とその結果を利用した動作解析について述べた。今後の課題として、機械学習内部の重みやデータに対する反応を見るための手段を考える必要があることや、動作識別後に行う定量的な動作指導の方法や学習者の動作特徴の可視化手法を考案する必要がある。

謝辞

本研究は JSPS 科研費 JP17K01098 の助成を受けた。

参考文献

- (1) 宮城諒, 平石広典: “機械学習を利用した人間動作検出センサのための行動検出”, 情報処理学会第 77 回全国大会, Vol.4, pp.313-314 (2015)
- (2) 酒井正人, 越智洋司, Kinect と機械学習を利用したラジオ体操動作識別, 2017 年度教育システム情報学会学生研究発表会 (2018)
- (3) Yao Ming, Shaozu Cao, Ruixiang Zhang, Zhen Li, Yuanzhe Chen, Yangqiu Song and Huamin Qu: “Understanding Hidden Memories of Recurrent Neural Networks”, arXiv, cs.CL, 1710.10777v1 (2017)