

OpenPoseによるバスケットボール投入予測

Prediction of Basketball Free Throw Scoring by OpenPose

中井 真人^{*1}

Masato NAKAI

角田 善彦^{*1}

Yosihiko TSUNODA

孫 財東^{*1}

Caidong SUN

村越 英樹^{*1}

Hideki MURAKOSHI

林 久志^{*1}

Hisashi HAYASHI

網代 剛^{*1}

Tsuyoshi AZIRO

^{*1}産業技術大学院大学 産業技術研究科

School of Industrial Technology, Advanced Institute of Industrial Technology

We aim for application developments in several fields such as healthcare and robot control. When considering to achieve high quality performance in such applications, it is often necessary to have large amount of accurate and various human posture data. OpenPose of CMU can realize simultaneous posture recognition of multiple people in real time by simple Web camera. We developed two application systems utilizing OpenPose. The first one is a basketball free throw prediction model with human posture data recognized by OpenPose. The second one is a healthcare application which aims to detect brain stroke tendency. In this article, we focus on what we achieved through our first application development. We accomplished to build a high quality motion prediction model with OpenPose data. We consider it opens a way for other research and application developments which are related to human posture and motion.

1. はじめに

我々のチームでは、ヘルスケアやロボット制御分野などを念頭に置き、それら分野での応用開発に取り組んでいる。それらの性能向上を考える場合、精度の高い大量で多様な姿勢データが必要とされる場合が多い。昨年夏に CVPR2017 で発表された米国カーネギーメロン大 (CMU) の OpenPose [Cao 17] [CMU-OP GitHub17] は、簡単な Web カメラを使って図 1 の様にリアルタイムで多人数の姿勢が同時に認識できる。その動画は SNS 等で拡散されかなりの反響を呼んだ。OpenPose の応用研究としてリアルタイムの姿勢データを使ってバスケットボールの投入予測と脳梗塞初期診断判定 (FAST [Takan] [NCVC]) システムを構築した。その結果、投入予測モデルでは先行研究と比較してかなり優位であることが分った。理由としては 2 つあり、従来は全体の姿勢データを時系列で取得するにはモーションキャプチャーの様なセンサーを取り付ける必要がありデータ取得が煩雑であった。今ひとつはデータ取得が煩雑であるためサンプリングが極めて少なく回帰分析等の統計モデルを使った分析ができなかった。一方 OpenPose は Web カメラで動画を撮るだけで容易に多くの姿勢データが取得でき統計モデルを使った分析が可能になった。さらに統計モデルを使った投入予測では有意な精度が得られたので OpenPose が認識するリアルタイムの姿勢データは、統計モデルに耐え得る精度を持つことが判明した。

2. 目的

OpenPose は 2017 年 3 月にステレオ (複眼) カメラによる 3 次元捕捉 [CMU-OP GitHub17] に対応しているが、本研究では導入の容易さや実績を重視して単眼 Web カメラを用いた OpenPose 二次元位置認識環境を構築した (本稿 4.1 参照)。OpenPose の 1 つの特長は、リアルタイムで多人数の姿

勢を同時に認識する点である。これを実現するため深層学習を使ったトップダウンの位置認識や PAF (Part Affinity Fields) [Cao 17] による部位の所属認識を採用した独特の工夫をしている。その結果 OpenPose の動画は見た目には精度と汎化性能は驚くべきものがある。しかし一般的に精度と汎化性能はトレードオフの関係があり両立は難しいとされている。そのため OpenPose の性能を確かめるべく OpenPose の姿勢認識データを使った予測モデルを構築して、その予測精度で OpenPose の性能を判断することにした。予測モデルは動画が撮り易く結果が明瞭なバスケットボールフリースローの動作から命中を予測するものとした。手続きとしてはフリースローの動画を OpenPose に通すと我々の環境では、人体姿勢部位検出の引数指定にもよるが 1 秒間に 5 ~ 20 フレーム程度の分解能で姿勢データが出力される。フレームデータを繋いで一旦時系列データとし、そこから特徴量を抽出する。命中予測は特徴量と命中結果とでロジステック回帰で行った。



図 1: OpenPose

3. 先行研究

スポーツ動作の解析は従来はビデオ画像を静止画に分割して姿勢位置を抽出するものが大多数であった [奥山 90] [松原 90]。一方姿勢データを直接取得する方法として体の部位にセンサーやモーションキャプチャーを取り付けて観測した。しかしデータ取得が煩雑なので回帰の様な統計モデルが使えず上

連絡先: 産業技術大学院大学 産業技術研究科 創造技術専攻
b1617mn@aait.ac.jp b1613yt@aait.ac.jp

級者と初心者の姿勢の差を論じるのが殆どであった [小郷 92] [安松谷 12]。一方バスケットボール試合ではシュート、ボール奪取等の回数データが試合毎に大量に記録されているので統計モデルを使った勝敗予測モデルが可能だった [齊藤 13]。スポーツ動作の統計モデルとしては隠れマルコフ [Stamp 12] を使った稲邑の姿勢抽象化モデル [稲邑 03] があるが、これは転移学習を目的としている。時系列の予測モデル [北川 93] は一般に定常波と非定常波に分かれ前者は ARIMA モデル、後者はカルマンフィルター [有本 77] が使われる事が多い。しかしこれらのモデルは逐次的に観測された説明変数と目的変数を使う。一方バスケットボールの投入予測は一連の姿勢データから最後に目的変数が得られるので特殊な時系列モデルとなる。この様なモデルとしては LSTM による一連の文章から肯定的か否定的かを予測する感情分析 [Yogatama 17] [Socher 12] や株価の動向から短期的に上昇下降を予測する SCW [Wang 14] がある。しかし何れも大量のデータを必要とする。バスケットの投入予測モデルは特殊な時系列モデルであり、少量のサンプリングであったので、時系列予測ではなく投入前後の姿勢とその差分から特徴量を抽出したロジステック回帰 [Bishop 06] を採用した。

4. OpenPose とは

安価な Web カメラで撮った動画をリアルタイムで多数人数を同時に姿勢認識するモデルで、昨年夏の CVPR2017 でカーネギーメロン大学 (CMU) が開発し学術公開したものである。図 2(左端) の様に 17 の関節点が繋がった 2 次元の姿勢を認識し、その結果をファイルに出力する。図 2(中央, 右端) の様に OpenPose は指先と顔も同様に認識する。

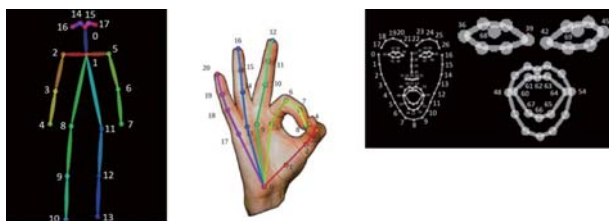


図 2: Analysis Point

4.1 OpenPose の動作方法

本研究では、次のような構成の PC サーバ上で OpenPose を稼働させた。インストールは CMU OpenPose オンラインリポジトリ [CMU-OP GitHub17] の “Installation, Reinstallation and Uninstallation” から、特に以下のハードウェアやソフトウェアに関して推奨の仕様 (CPU コア数やメモリ容量) やバージョンを満たしていることを確認して行う。

【ハードウェア】

- CPU: AMD Ryzen 7 1800X (コア数 8)
- メインメモリ 16GB
- GPU: NVidia GeForce GTX 1080 ti (GPU メモリ: 11GB)

【ソフトウェア】

- OS: Ubuntu 14.04 LTS
- CUDA バージョン: 8.0
- cuDNN バージョン: 5.1 for CUDA8.0

4.2 OpenPose 出力データ

本稿 2 章に記した通り、本研究での OpenPose は 2 次元の位置認識なので姿勢、指先、顔の位置データは何れも x 値, y 値, 信頼度の 3 値で表現される。OpenPose は隠れた位置の部位まで推測するので 0~1 の信頼度がある。また多人数の姿勢を同時に認識できるので、1 フレームのファイルには左側の人物から人数分出力される。OpenPose は図 2 左端の人体姿勢 17 点の捕捉を基本とし、加えて図 2 中央および右端のように指先や顔のデータ出力は引数で指定できる。姿勢データは 1 フレーム毎に基本 17 点、指先、顔の同期をとって別ファイルに出力される。出力データ分解能は、計算環境の性能や出力部位の引数指定により変わるが、我々の環境では 1 秒間に 5~20 フレーム程度であった。

5. バスケット投入予測の手法

5.1 OpenPose の姿勢データ

本研究では OpenPose の姿勢データのみを使用した (図 2 左端 17 関節点データ)。51 人のバスケットボールフリースロー動作の動画を撮り、OpenPose で解析した姿勢認識データをダウンロードした。

5.2 時系列への変換

ダウンロードしたフレームファイル群を読み込み 17 点の関節を時系列に変換した。

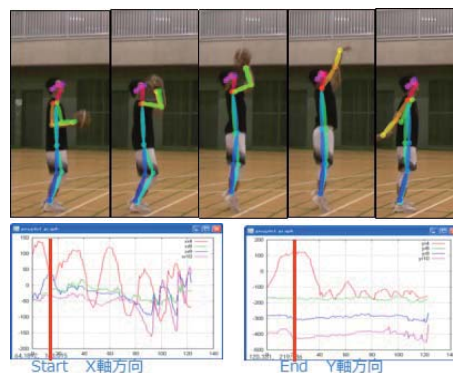


図 3: Time Series

51 人の全員は膝を曲げてから投入していたので、投入開始は膝を最大に屈折した時点とし、投入終了は手が最高点に達する時点とし、約 3~5 秒間の時系列に揃えた。また各関節点は首関節からの相対距離とした。

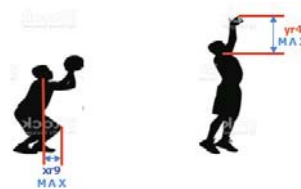


図 4: Start and Last Figure

5.3 ロジステック回帰モデルの採用

ボールが籠に入ったか入らなかったかの 2 値の教師データなので、ロジステック回帰モデルを採用した。2 値判定モデルでは定評のある Xgboost [Chen 16] も検討したが、分析結果の

出力は命中に寄与する特徴量が不明瞭なので採用しなかった。また非線形の2値判定モデルのSVM[Platt 99]があるが、これは判定確率を計算しないので採用しなかった。

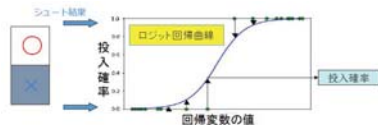


図 5: Logistic Regression

5.4 特徴量の作成

ロジスティック回帰モデルに投入する特徴量を17点の関節点の投入開始時の値、投入終了時の値、その差分とその速度とした。22の特徴量を投入したが、有意な特徴量として図7に示すものが残った。

6. 実験結果

6.1 精度

51人を命中予測の高い順に並べて、命中した実績の累計の図を作成すると図6のパレート図が得られた。これで示された精度ARは41%で有効な判別モデルが構築できていることが示された。(ARは完全モデルと無反応モデルで囲まれた凸部分の割合)

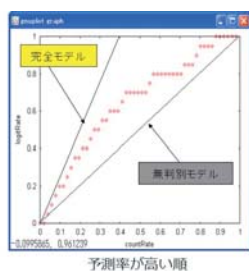


図 6: AR Accurate

6.2 有意な特徴量の検定

投入された特徴量で有意なものが正しくモデルを説明できているか検討した。図7は回帰の符号より、命中予測が高くなる特徴量の方向である。これによると開始時は腕を前から後ろへ一旦引いて反動を付けて投擲するオーバースローと投球時には膝の伸張と腕の振り上げによりボールに力を伝達している事が判明した。これにより特徴量の解釈でもこのモデルの有効性を確認した。

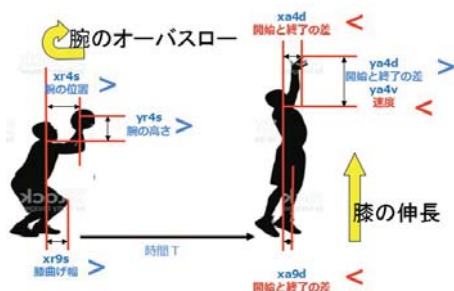


図 7: Feature Direction for Maximum Precision

6.3 診断と矯正

投入予測モデルはレベル診断システムに使えることが判明した。OpenPoseはリアルタイムで姿勢を認識するので図8に示す様に被験者の投入姿勢から直ちに投入予測を計算してレベル診断ができる。



図 8: Diagnosis System

また図9に示す様に有意な特徴量について初心者と上級者とを比較するとかなりの相違があることが判明した。この様に有意な特徴量を比較することで投入姿勢を矯正できることも分った。

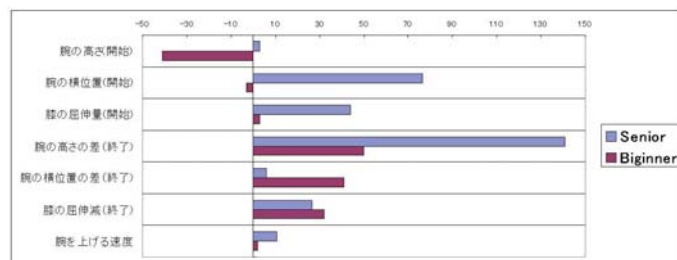


図 9: Diagnosis by Comparison

7. まとめと今後

一連の動作から投げられたボールの命中を予測する難しいモデルで期待された精度が得られたので、OpenPoseのリアルタイムの姿勢認識精度は極めて高いと判断した。OpenPoseは容易に姿勢データを得られることから、様々な場面で大量で高精度の姿勢データを取得することができる。図10の様に運転手や自転車の人と同時に観測でき、自動運転等の危険回避モデルに応用できると考えられる。



図 10: Prediction for Auto Driving

また競技での姿勢データが得られることで、図 11 の様に敵対的に学習するロボット [Fu 17] の初期学習としての逆強化学習 [Wulfmeier 14] に適用が考えられる。

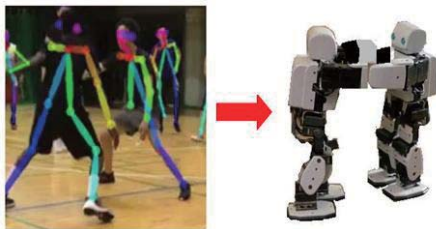


図 11: Inverse Reinforcement Learning

本稿研究で得られた OpenPose 姿勢データ解析の知見を活用し、FAST [Takano] [NCVC] と呼ばれる一般人が行う脳梗塞兆候判定を情報デバイスが検出・実行するシステムの試作も行った。この試作では図 12 の様に OpenPose の指先と顔の認識も使うことで正しく判定できることが分かった。この取り組みにより、OpenPose 技術のヘルスケア分野への応用可能性の一端を開いた。この詳細については別の機会に報告する予定である。

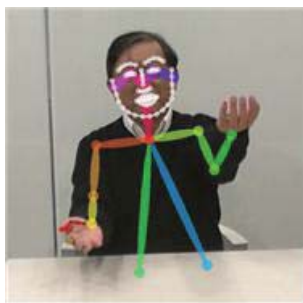


図 12: FAST

謝辞

東京都立産業技術高等専門学校のバスケットボール部にはフリースローの動画撮影に協力頂き感謝申し上げます。

参考文献

- [Cao 17] Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, CVPR2017
- [CMU-OP GitHub17] CMU-Perceptual-Computing-Lab openpose, <https://github.com/CMU-Perceptual-Computing-Lab/openpose>
- [Takano] Junji Takano: t-PA - a 'Magic Bullet' For Treating Stroke, <https://www.pyroenergen.com/articles13/ischemic-stroke-detection.htm>
- [NCVC] 国立研究開発法人国立循環器病研究センター: [103] 脳梗塞が起こったら - FAST を覚えよう, <http://www.ncvc.go.jp/cvinfo/pamphlet/brain/pamph103.html>
- [奥山 90] 奥山秀雄: コンピューター動作分析システム (APAS) を用いたバスケットボールのシュート分析 (1990), 日本体育学会第 41 回大会
- [松原 90] 松原孝 猪木原孝二 川上雅之 [他]: バスケットボールゲームのシュートについて (1990), 岡山理科大学紀要
- [小郷 92] 小郷克敏 錦井利臣 小澤雄二 中川保敬: 連続動作時における主運動の視覚情報獲得時期 (1992), 日本体育学会大会号
- [安松谷 12] 安松谷 亮宏 曾我真人 瀧寛和: バスケットボールにおけるシュートフォームの学習支援環境の構築 (2012), 情報処理学会シンポジウム論文集
- [斉藤 13] 斉藤百合子 倉石平: バスケットボール競技におけるゲーム分析の指標に関する事例研究 (2013), 早稲田大学修士論文
- [Stamp 12] Mark Stamp: A Revealing Introduction to Hidden Markov Models (2012)
- [稲邑 03] 谷江博昭 稲邑哲也 中村仁彦: 統計的情報量を用いた原始シンボル空間の構成とそれに基づく原始シンボル系列の記号化 (2003), 第 21 回ロボット学会学術講演会
- [北川 93] 北川源四郎: 時系列解析プログラミング - FORTRAN77 (1993), 岩波コンピュータサイエンス
- [有本 77] 有本卓: カルマン・フィルター (1977), 産業図書システム・サイエンス・シリーズ
- [Socher 12] Richard Socher, Brody Huval, Christopher D. Manning, Andrew Y. Ng: Semantic Compositionality through Recursive Matrix-Vector Spaces, EMNLP2012
- [Yogatama 17] Dani Yogatama, Phil Blunsom, Chris Dyer, Edward Grefenstette, Wang Ling: Learning to Compose Words into Sentences with Reinforcement Learning, ICLR2017
- [Wang 14] Jialei Wang, Peilin Zhao, Steven C.H.: Exact Soft Confidence-Weighted Learning, ICML2014
- [Bishop 06] Christopher M. Bishop: Pattern Recognition and Machine Learning, Springer (2006)
- [Chen 16] Tianqi Chen, Carlos Guestrin, Christopher: XGBoost: A Scalable Tree Boosting System, Arxiv2016
- [Platt 99] John C. Platt: Fast Training of Support Vector Machines using Sequential Minimal Optimization (1999)
- [Fu 17] Justin Fu, Katie Luo, Sergey Levine: Learning Robust Rewards with Adversarial Inverse Reinforcement Learning, Arxiv2017
- [Wulfmeier 14] Markus Wulfmeier, Peter Ondruška, Ingmar Posner: Maximum Entropy Deep Inverse Reinforcement Learning, NIPS2014