

深層学習による三次元姿勢推定を用いたアーチェリーの射型教示システム

A teaching system of archery shooting using 3D pose estimation by deep learning

○学 道脇 拓真 (中央大学)

Sarthak Pathak (中央大学)

正 梅田 和昇 (中央大学)

Takuma MICHIWAKI, Chuo University, michiwaki@sensor.mech.chuo-u.ac.jp

Sarthak PATHAK, Chuo University

Kazunori UMEDA, Chuo University

Archery is a sport in which it is important to repeat the same shooting form and to improve its repeatability. For this reason, it is essential to understand and master the basic shooting form. However, there is a shortage of instructors, and in university club activities, the burden of teaching is placed on the senior students. A possible solution to this problem is the use of a practice support system based on sports video analysis, but existing systems have a fixed positional relationship between the camera and the person to be measured, and it is necessary to create the same environment each time the system is used. Therefore, this study proposes a sports practice support system that uses 3D skeletal point information extracted from video images to evaluate the learner's shooting style in terms of points, eliminating the need to fix the positional relationship between the camera and the learner and allowing easy recognition of points that need improvement.

Key Words: Deep Learning, Pose Estimation, Motion Analysis, Archery

1 緒言

アーチェリーは自身の体に合った再現性の高い射型を作りあげていくスポーツであり、そのためには基本となる射型を理解し、自然と行えるようにすることが必要不可欠である。文部科学省によるとアーチェリーの競技人口は、似た競技である弓道の競技人口の約 1/10 であり、はるかに少ない [1]。そのため、指導員の不足により指導員の負担が増加しているという課題がある。この課題の解決のためにはスポーツ映像解析を用いた練習支援システムの利用が考えられる。しかし、既存のシステム [2][3] はカメラと測定対象者との位置関係が固定されており、システムを利用するために毎回同じ環境を作る必要がある。これは、公共のスポーツ施設で練習することの多いアーチェリーでは大きな課題である。また、評価できるポイントが限られていたり、フォームの悪い点に分かりづらいなどの問題もある。

そこで本研究では、深層学習による三次元姿勢推定を用いて動画画像から抽出される三次元骨格点情報を利用し、学習者の射型を複数の対象となる動作に対して点数評価を行う。そして、改善すべき点を学習者が容易に認識できる、かつカメラと学習者の位置関係を固定する必要がないアーチェリーの射型教示システムを提案する。

2 提案手法

2.1 提案手法の概要

提案するシステムの流れを Fig.1 に示す。まず矢をつがえていない素引きの練習映像から姿勢推定モデルを用いて三次元骨格点情報を取得する。次に、取得した骨格点情報に射型評価モデルを適用することで、5つの評価項目で射型の点数評価を行う。

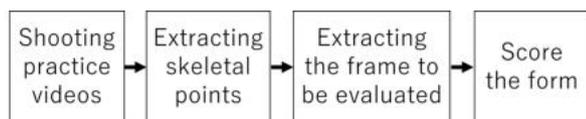


Fig.1 Flowchart of the proposed system

2.2 骨格点の抽出

骨格点の抽出には、StridedTransformer-Pose3D[4]を用いる。StridedTransformer-Pose3Dは、Transformerベースの2D pose sequenceを活用し、3D pose sequenceを生成する3D姿勢推定手法である。データセットはHuman3.6M[5]を用いる。Human3.6Mは4台のカメラで計11人の被験者を撮影した計約360万フレームの動画から成る、3D Pose Estimationの評価の際に最も標準

的に用いられるデータセットである。Fig.2に骨格点を抽出した様子を示す。

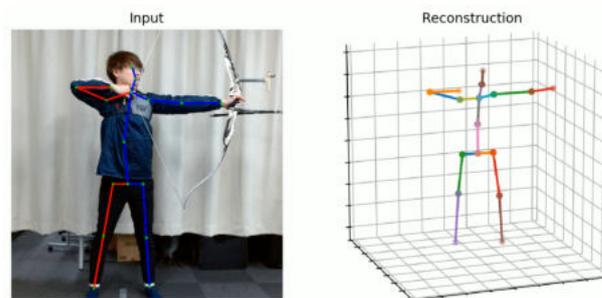


Fig.2 Skeletal point extraction with StridedTransformer-Pose3D

2.3 射型の評価

矢をつがえてから射るまでの一連の動作の内、弓を持ち上げる動作であるセットアップと、弦を引ききった状態であるフルドロの瞬間のフレームを、評価に用いるフレームとして抽出する。右肘の縦 (v) 座標が最大となるフレームをセットアップ、1つ前のフレームに対する右手首の v 座標の変化率 (n 番目のフレームでの右手首の v 座標 $- n - 1$ 番目のフレームでの右手首の v 座標 / $n - 1$ 番目のフレームでの右手首の v 座標) が 10 回以上連続で-1%以上 1%以下となった時の 10 番目のフレームをフルドロとする。

射型評価には、左右の肩、手首、右肘の計5点の骨格点座標を使用する。提案する射型評価手法は大きく二つに分けられる。一つ目の手法は、3ないし4個の骨格点座標のなす角を2つの方向ベクトルより算出し、しきい値処理により点数評価を行う。

二つ目の手法は、2個の骨格点の v 座標を比較し、しきい値処理によって点数評価を行う。以上の二つの評価手法を組み合わせ、5つのポイントのスコアを0から100の11段階で算出する。

評価対象とする5つの項目は以下の通りである。Fig.3に(1)~(5)の評価手法の模式図を示す。なお、Fig.3a~3b, ならびに3c~3eが、それぞれ前述のセットアップ、フルドロに対応する。

- (1) セットアップ時、引き分け(弓を持ち上げる際に少し弦を引いておくこと)ができていないか
- (2) セットアップ時、両肩と右肘・右手首が平行であるか

- (3) フルドロー時、両肩、両手首が平行であるか
- (4) フルドロー時、体の軸が傾いていないか
- (5) フルドロー時、右肘が下がっていないか

(1)~(5)のそれぞれの点数評価手法について説明する。

(1)については、セットアップ時の右肘、右肩、左肩の骨格点座標から右肩から左肩、右肩から右肘の方向ベクトルを求め、ベクトル間の角度 θ を計算し、 110° に近いほど高得点となるようにスコアを定義する。

(2)については、セットアップ時の右手首、右肘、右肩、左肩の骨格点座標から右肩から左肩、右肘から右手首の方向ベクトルを求め、ベクトル間の右手首と右肘、右肩と左肩を通る2直線のなす角 θ を計算し、 0° に近いほど高得点となるようにスコアを定義する。

(3)については、フルドロー時の右手首、左手首、右肩、左肩の骨格点座標右肩から左肩、右手首から左手首の方向ベクトルを求め、右手首と左手首、右肩と左肩を通る2直線のなす角 θ を計算し、 0° に近いほど高得点となるようにスコアを定義する。

(4)については、フルドロー時の左肩と右肩の v 座標の差 Δv を求め、その差が0に近いほど高得点となるようにスコアを定義する。

(5)については、フルドロー時の右手首と右肘の v 座標の差 Δv を求め、その差が0に近いほど高得点となるようにスコアを定義する。

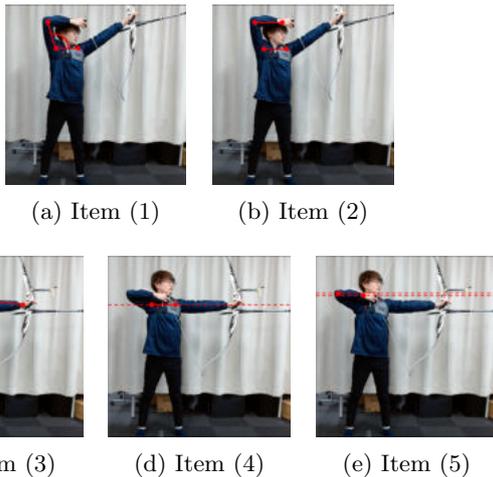


Fig.3 Evaluation items

3 提案手法の評価実験

3.1 実験概要

経験者と初心者の被験者各10名ずつ計20名が矢をつがえずに弓を引く動作である素引きを行う様子を、的的方向を基準として 45° 、 90° 、 135° の角度から距離2mの位置で撮影した。また、初心者は事前に経験者の射型動画を見て正しい射型を理解したのち、撮影を行った。すべての射手は右利きで、この実験では66インチ(弓の長さ)-18ポンド(弦を引く時に必要な力)のベアボウを使用した。撮影した動画を射型評価モデルに適用してスコアを算出し、経験者と初心者の得られたスコアの平均を比較した。そして、経験者が動画を見て採点したスコアと本実験で得られた平均スコアの相関を調べた。

3.2 実験環境

練習環境を想定し、シューティングラインを設置した。被験者がシューティングライン上で素引きを行う様子をFig.4の環境で撮影した。実験環境の模式図をFig.5に示す。



Fig.4
Experimental environment

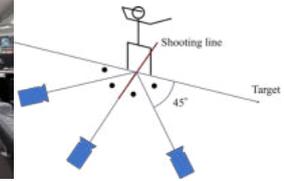


Fig.5 Schematic diagram

3.3 実験結果と考察

本実験により得られた経験者と初心者各10名のそれぞれの平均スコアをTable 1に示す。全体の平均スコアは全てのカメラ角度において初心者より経験者の方が高得点となっていることが分かる。それぞれの評価項目について見ていくと、カメラ角度が 135° の際に、評価項目(1)、(2)のスコアが他の角度のスコアより大きく下がる場合があることが分かる。この理由は、カメラ角度が 135° の場合にセットアップ時の右腕周りの骨格点の誤検出が発生しやすいからであると考えられる。また、全てのカメラ角度において評価項目(1)、(3)、(4)のスコアの経験者と初心者の差が小さいことが分かる。これは、本実験で使用したベアボウが18ポンドと一般的に中級者が試合で使う弓の半分ほどの力で引ける弓であり初心者でも簡単に引けるため、評価項目(1)、(3)、(4)で評価対象としている身体が弓の強さに耐えきれていない際になりやすい悪い射型にならなかったためであると考えられる。

次に、Table 2に経験者が採点したスコアと本実験で得られた平均スコアの相関係数を示す。また、Fig.6~8に、それぞれの散布図を示す。Table 2、Fig.6~8より、経験者に関しては、相関係数はカメラ角度 90° のときに最も大きい。 45° 、 135° で相関係数が小さくなっている理由は、 45° 、 135° のときに右腕周りから外れてしまったり、前腕と上腕が重なってしまうことで骨格点の誤検出がまれに発生するためであると考えられる。初心者に関しては、全てのカメラ角度において正の相関はあるものの、相関係数は小さいことが分かった。この理由は、経験者が動画を見た際に分かる明らかに悪い点が本手法の評価対象である5つの評価項目に含まれていない場合が多かったからであると考えられる。例としてFig.9とFig.10を比較すると、経験者の射型であるFig.9は正しい位置でアンカリング(ドロイングで持ってきた引き手を顎の下に固定すること)しているのに対し、初心者の射型であるFig.10ではかなりオーバードローとなっており、経験者であれば目で悪い点が把握できる。

本実験を通して、提案手法ではカメラ角度が変化しても初心者より経験者の方が高得点となっており、また経験者が採点したスコアと得られたスコアとは、カメラ角度によらず正の相関があることを確認した。しかし、骨格点の誤検出や、初心者を対象とした評価項目の不足などにより、相関が小さくなることも確認した。

Table 1 Comparison of average scores between experienced persons and beginners (%)

	angle	(1)	(2)	(3)	(4)	(5)	average
experienced	45°	100	55	88	86	71	80
	90°	92	77	73	87	80	82.4
	135°	76	47	68	92	69	70.4
beginners	45°	100	17	71	81	26	59
	90°	92	49	79	82	45	69.4
	135°	80	29	80	75	47	61.6

Table 2 Correlation coefficients between scores scored by experienced persons and calculated scores

	45°	90°	135°
experienced	0.42	0.89	0.35
beginners	0.43	0.24	0.31
experienced persons and beginners	0.72	0.60	0.48

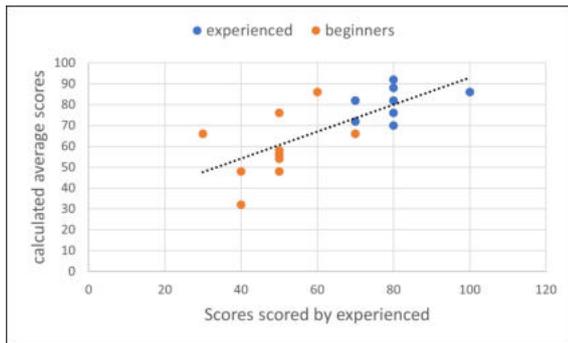


Fig.6 Experienced persons and beginners, 45 °

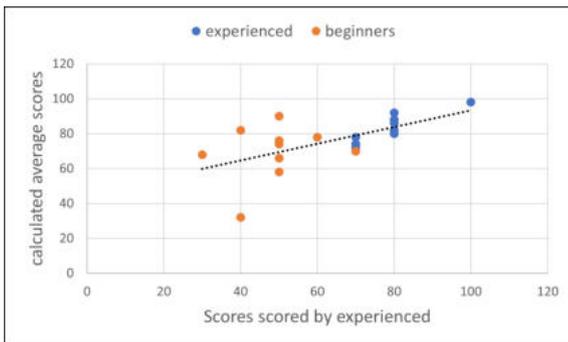


Fig.7 Experienced persons and beginners, 90 °

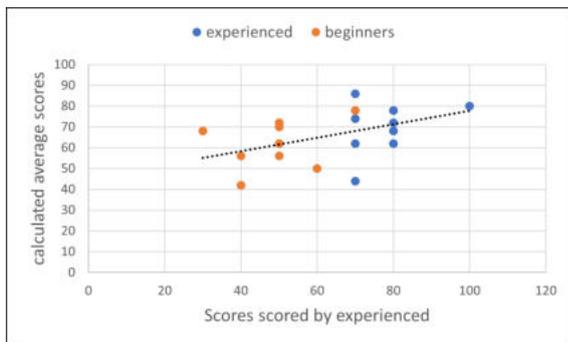


Fig.8 Experienced persons and beginners, 135 °



Fig.9 Experienced form

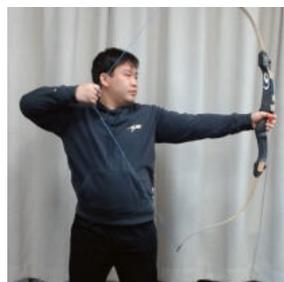


Fig.10 Beginners' form

4 結言

本研究では、画像から抽出される三次元骨格点情報を用いたアーチェリーの学習支援システムを構築した。具体的には、骨格点の y 座標の変化から評価に用いる 2 つのフレームを抽出し、関

節の角度と骨格点の y 座標の比較によってスコアを算出する手法を提案した。実験では、カメラと評価対象者の位置関係が変化しても初心者より経験者の方が高得点となっており、また経験者が採点したスコアと得られたスコアはカメラ角度が変化しても正の相関となっており、本システムの有効性を確認できた。

今後の課題としては、経験者の採点との相関をより大きくすること、対象者の練習効率の向上が考えられる。射型評価の精度向上のためには、評価対象ごとのしきい値の改善や、骨格点の誤検出が発生しやすい範囲を把握する必要があると考えられる。また実験を通して、弓の強さによって重視すべき評価対象が異なることが分かったので、様々な状況に対応できるよう評価対象の追加を考えている。対象者の練習効率の向上のためには、現在の射型と理想の射型の三次元空間上での描写や、それぞれの評価対象のスコアに対応し、改善に必要な動作及びトレーニングを教示するシステムの追加を考えている。

参考文献

- [1] 文部科学省, “諸外国および国内におけるスポーツ振興施策等に関する調査研究”, スポーツ政策調査研究 報告書, p.477, 2010.
- [2] 河瀬裕美, 松塚茜, “音と画像処理を用いたリアルタイムスポーツ学習支援システム”, 情報処理学会インタラクティブ論文誌, no.3, p.465-468, 2011.
- [3] Jonathan Then Sien Phang, King Hann Lim, Basil Andy Lease; Dar Hung Chiam, “Deep Learning Pose Estimation for Kinematics Measurement in Archery,” Int. Conf. on Green Energy, Computing and Sustainable Technology, pp.298-302, 2022.
- [4] Wenhao Li, Hong Liu, Runwei Ding, Mengyuan Liu, Pichao Wang, Wenming Yang, “Exploiting Temporal Contexts with Strided Transformer for 3D Human Pose Estimation,” IEEE Transactions on Multimedia, doi: 10.1109/TMM.2022.3141231, 2021.
- [5] Catalin Ionescu, Dragos Papava, Vlad Olaru, Cristian Sminchisescu, “Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 7, pp. 1325-1339, 2013.