

画像から抽出された骨格点を用いた フェンシング競技での得点判定

澤島 武博† Alessandro Moro† Sarthak Pathak† 梅田 和昇†

†中央大学

E-mail: sawahata@sensor.mech.chuo-u.ac.jp

1 背景・目的

フェンシング競技は世界的に見ても知名度が高い反面、日本の競技人口は 6000 人弱と少なくマイナースポーツである。競技力向上を目的として競技人口増加をはかるため、日本フェンシング協会では様々な取り組みがなされている。例えば、フェンシング競技の体験者を増やす目的で自治体と協力し、体験教室が開かれている。しかし、競技を観戦する機会を作ったとしても、競技のルールが複雑で観戦しても何が凄いかわからない。一方、フェンシング競技の経験者であればルールの理解があり、感覚的に試合状況を理解できる。しかし、初めて競技を観戦する人からすると、試合状況は激しく入れ替わりそれに加え高速で移動する選手を目で追うことも難しいため、試合状況の理解が難しくなる。よって、本質的なフェンシングの面白さが伝わらないという課題がある。

近年、そのほかのスポーツでは、スポーツをよりわかりやすく、より魅力的にする目的で、野球、サッカー、ゴルフをはじめ多くの競技でボールの軌跡の可視化やプレイヤーの検出などスポーツ映像解析[1]の研究が盛んに行われている。フェンシングでも同様の研究がなされており、それは剣先を可視化[2]するものである。細い剣先は競技中高速で動く為、剣先の通った軌道が可視化されることで試合の見栄えは良くなったものの根本的なルールの理解には繋がらない。そこで、競技経験者が感覚的に理解している試合状況を可視化させることでこの課題の解決を図る。

このとき、試合状況をリアルタイムで表示させることが一つの案となる。その為に実際の試合で審判や競技経験者が試合状況を理解する為に見ている選手の姿勢や移動方向を画像から抽出し、それらで試合状況を認識することを検討する。具体的には、動画画像から抽出される骨格点情報を利用し、優先権の切り替わりである「フリーズ」の判別と得点判定を行い、攻撃優先権を可視化するシステムを提案する。

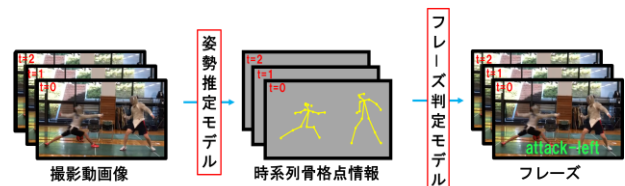


Fig.1 フリーズ判定システムの概要

2 提案手法

2.1 システム概要

提案する得点判定システムの処理の流れを Fig.1 に示す。まず競技映像から姿勢推定モデル[3]を用いて骨格点情報を取得する。次に、取得した骨格点情報にフリーズ判定モデルを適用することで得点判定を行う。

本研究のシステムは3つのフリーズを識別する。

- Simultaneous
- Attack
- Preparation-Attack

それぞれのフリーズの動画を攻撃開始時、攻撃途中、攻撃決定時の3つの瞬間(Fig.2)に分け詳しく説明する。まず、「Simultaneous」は両選手が同時に攻撃を仕掛けるフリーズでどちらの得点にもならない。Fig.2(a)を見ると真ん中の写真で両選手同じような動きをしているのが示されている。同じタイミングで攻撃を開始し攻撃途中でも同じような動きを取っている。これはお互いに攻撃優先権が発生せずどちらの得点にもならない「Simultaneous」というフリーズになる。次に、「Attack」を説明する。「Attack」は片方の選手のみが先に攻撃を仕掛けるフリーズで、同時にお互いを突いたとしても攻撃を仕掛けた選手の得点になる。Fig.2(b)の真ん中(攻撃途中)の瞬間の写真では右選手の腕部、脚部ともに相手に向かい伸びているのが分かる。最後に、「Preparation-Attack」

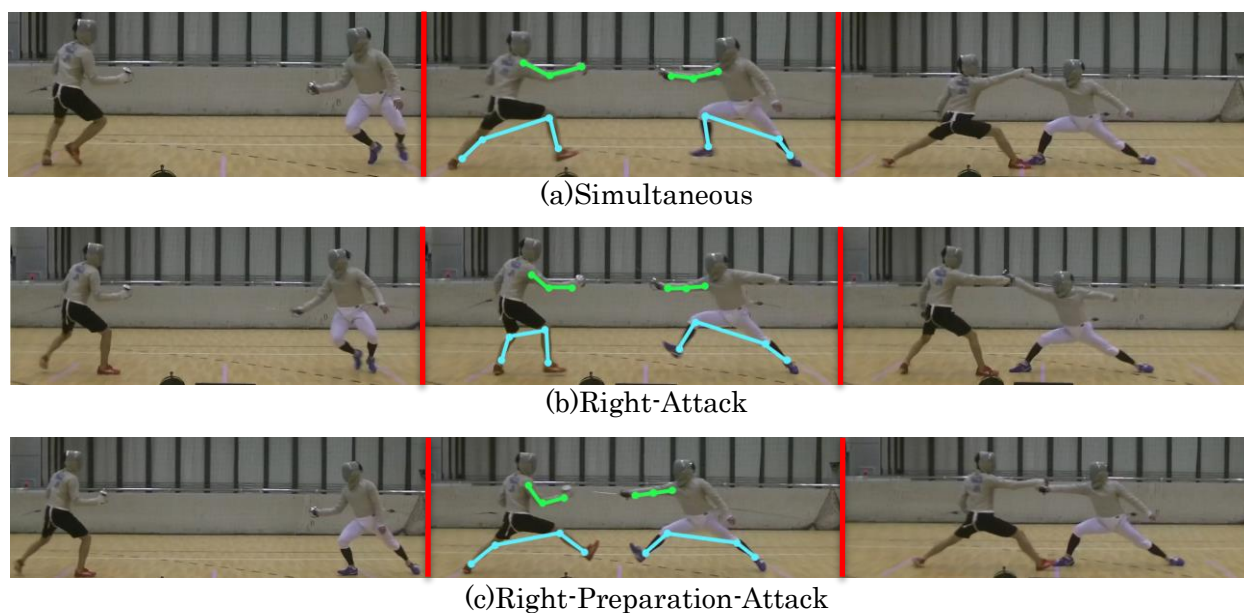


Fig.2 フレーズ

は同時に攻撃を開始したものの片方の選手の攻撃準備動作中に先に腕を伸ばし攻撃を先に仕掛けることで攻撃優先権を得るフレーズである。Fig.2(c)の真ん中の写真通り、足の伸び具合は変わらないものの、腕の伸び具合を比較すると左選手よりも右選手の方が先に腕を伸ばしているのが分かる。これも「Attack」のフレーズ同様、同時に突いたとしても攻撃を仕掛けた選手の得点になる。さらに、「Attack, Preparation-Attack」については左右の選手の識別も行うため、合計のフレーズ数は以下の5つとなる。

- Simultaneous
- Attack-right
- Preparation-Attack-right
- Attack-left
- Preparation-Attack-left

2.2 姿勢推定モデル

画像からの骨格点抽出には、Lightweight OpenPose[3]を利用する。OpenPoseは、画像中にある人物の二次元骨格点情報を検出するシステムで、OpenPoseのネットワークデザインと後処理の最適化により、従来のOpenPoseよりも処理を軽くしたものがLightweight OpenPoseである。

OpenPoseは3D モーションキャプチャシステムに及ばないものの、RGB カメラのみで映像から2D 骨格点情報を抽出できる。また、屋外での利用が可能

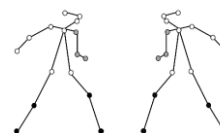


Fig.3 使用する骨格点

である。

2.3 学習を利用しない手法

学習を利用しない手法では、「移動値」と「腕の伸び値」を判定材料とし、フレーズの判定を行う。それぞれの値を選手間で差を取り、時系列での振る舞いによってフレーズを判定する。この手法は深層学習手との比較に用いる。

・移動値

両足の膝、足首の4点(Fig.3:黒)を使用する。移動値はFig.4にあるように前フレームとの差 $\delta_n = \sqrt{(x_{curr} - x_{prev})^2 + (y_{curr} - y_{prev})^2}$ の4点の合計 $\delta_{n:all:foot} = \delta_9 + \delta_{10} + \delta_{12} + \delta_{13}$ とする。

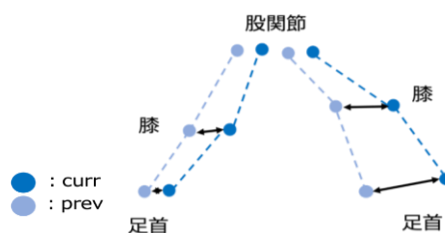


Fig.4 移動値

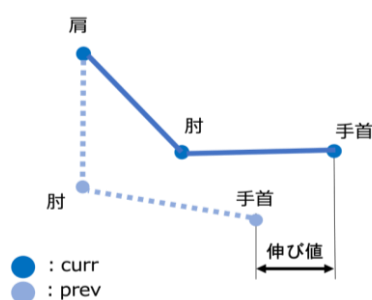


Fig.5 腕の伸び値

・腕の伸び値

剣を持っている側の肩、肘、手首の3点(Fig.3:灰色)を利用する。伸び値の算出には Fig.5 にあるように肩の2次元座標($x_{shoulder}$, $y_{shoulder}$)と手首の2次元座標(x_{wrist} , y_{wrist})を利用する。伸びの初期値 $X_0 = |x_{0:wrist} - x_{0:shoulder}|$ を求め、各フレーム n での差 $\delta_{n:arm} = X_n - X_0$ を腕の伸び値とする。

・しきい値の設定

以上より求めた移動値 $\delta_{n:foot}$ と腕の伸び値 $\delta_{n:arm}$ の2選手間での差を取る。 $\delta_0 - \delta_1 > 0$ となれば0の選手が多く動いている。また、腕が多く伸びていると判断できる。これらの値の時系列でみるために最小二乗法により点群の傾きを算出する。この点群の傾きが正になれば0の選手が先に手を伸ばし、先に多く移動したと認識が可能になる。またその逆も同様である。

2.4 深層学習手法

深層学習ベースのフレーズ判定モデルには Long Short-Term Memory(LSTM) [5][6]を用いる。また、ネットワークの概要を Fig.6 に示す。入力は2.3 で利用した7点、2選手分の時系列骨格点情報を用いる。出力は5フレーズ分の0~4のクラス名である。

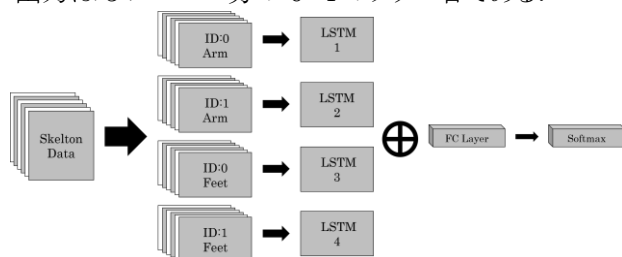


Fig.6 ネットワークの概要

2.4.1 判定器の作成

フレーズ判定器は、入力に時系列情報を扱うためネットワークには LSTM を使い、データセットで学習を行い作成する。

LSTM は、長期的な依存関係を学習することので

きる、Recurrent neural network(RNN)の一種である。RNN とは時系列データにおいて出力側から入力側へのフィードバックを持つ、ある時刻の出力が次の時刻の出力になる構造を持つニューラルネットワークである。よって過去の情報を引き継ぎながら時系列を考慮した認識などを実現できるネットワークになる。シンプルな RNN の処理は線形変換と活性化関数のみで、演算結果そのものが次の入力に引き継がれる。それに対して LSTM は Cell State, Input Gate, Output Gate, Forget Gate を持ち、それらが長期記憶を保持し、状況に応じて選択的に記憶して選択的に次の時刻に記憶を伝え選択的に出力する。そのため時系列が長くなった時も長期の記憶の保持を可能にしている。また勾配爆発、消滅などが起こりにくい為、学習も安定して行われる。

2.4.2 データセットの作成

データセットは2選手の時系列骨格点情報である。競技が開始されてから得点が決まるまでの動画を1フレーズに対し20回撮影した。これらの動画に姿勢推定モデルを適用することで、時系列の骨格点情報を得る。合計100個のデータセットで学習を行う。今回は学習データの拡張は行っていない。

3 比較実験

2. で述べた二つの提案手法の精度評価を行うために、学習を利用しない手法と深層手法で比較実験を行う。合計5フレーズで判定を行いそれらの正答率で評価した。

3.1 実験環境

撮影場所は中央大学多摩キャンパス第一体育館3階アリーナ A コートである。

試合環境を想定し、手前にどちらが触れたかを表示させる審判器を選手に隠れることがないようにピスト前方に設置した。Fig.7 の環境で撮影を行なった。

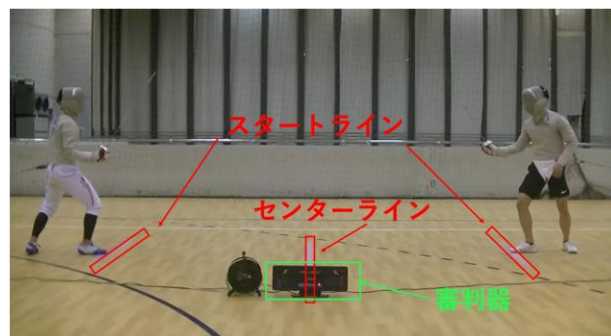


Fig.7 実験環境

選手はスタートラインを踏まないように構え、スタートライン間は4メートル(ルールで決められている)中央にはセンターラインを設置した。動画は審判のはじめの合図からトゥシェ(突き)の判定が審判器に表示されるまでの間のものである。約1秒程度のものになっている。

3.2 実験方法

被験者は3人とした。2人を選び5フレーズの撮影を2回行った。次に、この2人を入れ替える。また、同じように再度撮影を行う。それを全ての組み合わせで行う、1フレーズにつき合計で12動画を撮影し全ての動画は60個である。これら60個の画像に学習を利用しない手法、深層学習用いた手法の二つの認識率を評価した。

3.3 実験結果

実験結果を表1に示す。

学習を用いない手法の平均認識率は48.3%、深層学習は80.0%の認識率となった。2つの手法で認識率に31.7%もの差がある。骨格点の誤認識や未検出が原因である。大きく外れた値に大きく影響を受けていた。一方データの欠陥があるものも学習していた深層学習手法ではこれらの問題は判定結果に影響しないため学習を利用しない手法よりも認識率が高くなった。また、骨格点の変化の仕方、例えば途中からいきなり腕が伸びるのと初めからゆっくりと腕が伸びる。この差は最小二乗誤差を用いた学習を用いない手法では差が出にくく、これらの差を見分けられたのが深層学習であった。以上により深層学習手法の認識率が高くなったと考える。

Table1 実験結果

	0	1	2	3	4	平均
学習を用いない手法	8.3	58.3	66.7	50.0	58.3	48.3
深層学習手法	75.0	75.0	83.3	83.3	83.3	80.0

4 結論

本研究では画像から抽出される骨格点情報を用いたフェンシング競技での得点判定システムを構築した。実験では、深層学習手法の有効性を確認した。今後は深層学習手法の改良、制度向上、処理速度の改善、フレーズの多様性、撮影範囲の拡大などを行い、実際の試合環境の中でも利用可能な得点判定システムを構築する。

参考文献

[1] H.-C. Shih, ‘A Survey of Content-Aware Video Analysis for Sports’, IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 5, pp. 1212–1231, May 2018, doi: 10.1109/TCSVT.2017.2655624.

[2] M. Takahashi, S. Yokozawa, H. Mitsumine, T. Itsuki, M. Naoe, and S. Funaki, ‘Sword tracer: visualization of sword trajectories in fencing’, in ACM SIGGRAPH 2018 Talks, Vancouver British Columbia Canada, Aug. 2018, pp. 1–2. doi: 10.1145/3214745.3214770.

[3] 国際フェンシング連盟競技規則(t)日本語版 発行: 日本フェンシング協会 t.9 定義

[4] D. Osokin, ‘Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose’, arXiv:1811.12004 [cs], Nov. 2018, Accessed: Apr. 21, 2022. [Online]. Available: <http://arxiv.org/abs/1811.12004>

[5] J. C. Núñez, R. Cabido, J. J. Pantrigo, A. S. Montemayor, and J. F. Vélez, ‘Convolutional Neural Networks and Long Short-Term Memory for skeleton-based human activity and hand gesture recognition’, Pattern Recognition, vol. 76, pp. 80–94, Apr. 2018, doi: 10.1016/j.patcog.2017.10.033.

[6] M. B. Shaikh and D. Chai, ‘RGB-D Data-Based Action Recognition: A Review’, Sensors, vol. 21, no. 12, Art. no. 12, Jan. 2021, doi: 10.3390/s21124246.

[7] ‘スポーツ向けビデオカメラ GY-TC100’, JVC ブランド製品情報. <https://www.jvc.com/jp/pro/video/lineup/gy-tc100/> (accessed Apr. 21, 2022).

[8] ‘FIE公認審判器 ファベロ 一体型 LED表示方式 | 東京フェンシング商会 | フェンシング用品通販. ALLSTAR 製品他. ’. <https://www.tf-fencing.co.jp/shop/catalog/118/FV831/> (accessed Apr. 21, 2022).