

## 画像から抽出される剣先情報と骨格点を用いた フェンシング競技での攻撃優先権の認識と可視化

○澤島武博 †, Sarthak Pathak ‡, Alessandro Moro §, 梅田和昇 ‡

○Takehiro SAWAHATA †, Sarthak PATHAK ‡, Alessandro MORO §,  
and Kazunori UMEDA ‡

† 中央大学理工学研究科, sawahata@sensor.mech.chuo-u.ac.jp

‡ 中央大学理工学部, § RITECS Inc.

<要約>本研究ではフェンシング競技において試合状況の可視化を目的とし、画像から抽出された剣情報と骨格点情報を用いて、試合状況を判定するシステムを提案する。高速で動く剣先を連続的にマーカレス検出し、骨格点情報と組み合わせ事前学習済み時系列ニューラルネットワークにて試合状況を判定する。実際の試合動画で実験を行い平均で98.1%の認識率を実現した。

<キーワード>フェンシング, 深層学習, スポーツ, 物体認識

### 1 序論

近年、画像処理技術の進歩に伴い、フェンシング分野における研究が活発化している。フェンシング競技は、高速で動く剣先を目で追うことや試合規則の複雑さから、初心者にとって試合状況を理解することが難しい。この問題に対処するため、試合状況の可視化に焦点を当てた研究が行われている。例えば、画像から抽出された骨格点情報を用いてフットワークを特定する研究 [1] やフェンシングの剣をマーカレスで認識しリアルタイムで表示する研究 [2, 3] が行われている。また、「Sport: Fencing Matches AI[4]」において、Alexander P. はフェンシングの試合状況分析を行い、当論文で「剣に関する情報の欠如」が試合状況を判断するための課題であると結論付けている。この課題に対処するため、J. Mo[5] は試合のオーディオ情報を分析し、剣の接触を判断するシステムの研究を行った。しかし、オーディオ情報のみに依存することは不十分であると考えられる。特に、剣の「姿勢」によって試合状況が変わるため、剣の姿勢情報の取得は重要である。この背景から、本研究では剣先の検出を含めた試合状況認識のシステム構築に焦点を当てる。

フェンシングの剣は非常に細く、剣先の動きは高速である。従来研究では、マーカースを使用した検出が一

般的であったが、試合中の選手への負担のため、実際の競技での使用は困難である。そこで、剣先に反射材テープを取り付け、赤外線カメラで追跡する「Sword tracer[6, 7]」の開発が試みられているが、この技術は絶縁テープを使用しない種目には適用できない。「Fencing Tracking and Visualization System[8]」というシステムは、24台の4K解像度カメラを使用し、それぞれが60fpsで動作するように設計されている。しかし、これらの高性能カメラを同時に運用するためには、大規模な処理能力と高度な技術的サポートが必要であり、その結果、システムの設置が困難になる可能性がある。

以上から、従来研究では、標準的な単一カメラを用いた剣検出システムは存在せず、剣の姿勢情報を活用した試合状況認識システムも開発されていない。そこで、本研究の目的を、標準的なカメラのみを使用し、選手に負担をかけずにマーカレスで剣先トラッキングを含んだ試合状況認識システムの構築とする。このシステムは画像から剣身のマスク情報と骨格点情報を取得し、それらを組み合わせ剣の所有者特定と剣先の特定を行う。特定された剣先情報を骨格点情報に統合し、深層学習に基づく学習済みネットワークにて試合状況の分類を行う。このアプローチにより、フェンシング競技における試合状況の正確な判定を目指す。

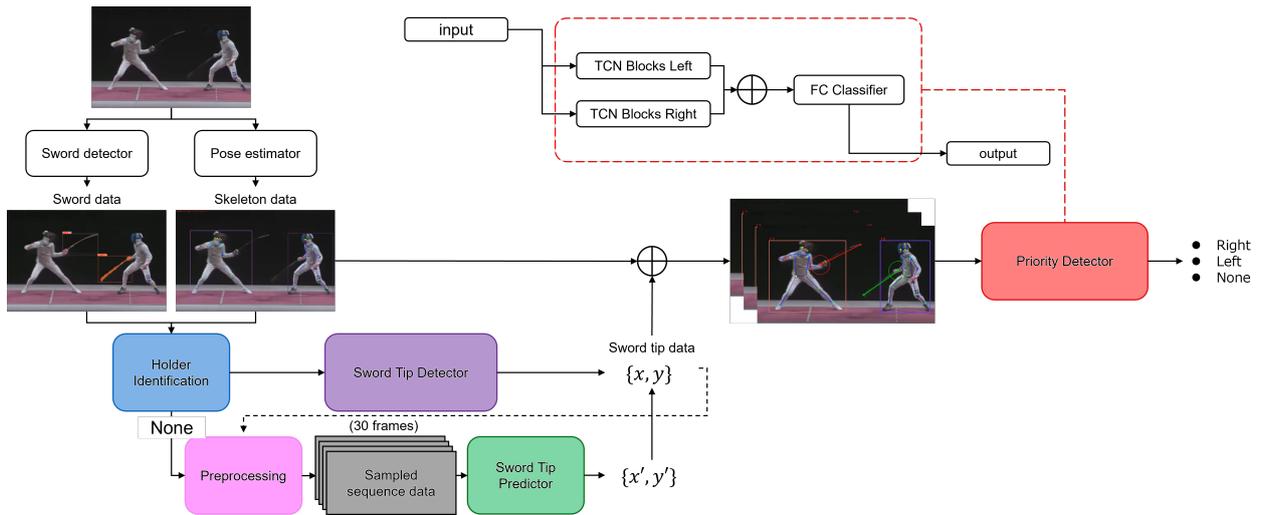


図 1 剣先検出を含んだフェンシング競技での攻撃優先権判定手法のフローチャート

## 2 試合状況の推定

### 2.1 概要

提案システムは、画像のみを用いてフェンシング競技における攻撃優先権の判定を行う。提案システムのフローチャートを図1に示す。システムは剣先の検出と攻撃優先権の判定から構成される。まず、剣先座標の取得を行う。画像から骨格点情報と剣全体と手首を含めたマスク情報を使用し剣先を特定する。剣は高速で動くため、図2のようにブレによりカメラに映らなくなる場合がある、この場合にも過去のデータから未検出フレームを予測し補間を行い、連続的な剣先の検出を行う。特定された剣先を骨格点の一部とした骨格点情報を攻撃優先権判別システムに入力し攻撃優先権をフレームごとに判定する。

「攻撃優先権」とは、フェンシングにおいて先に攻撃を仕掛けた選手に与えられる権利を指す。この概念は、得点を決定する重要な要素である。具体的には、攻撃優先権を持たない選手が相手の攻撃を防御に成功した場合、その選手に攻撃優先権が移譲される。防御には、相手の剣を叩いて攻撃を阻止する動作が含まれる。このような状況を正確に判断するためには、剣の姿勢情報が不可欠である。さらに、剣同士の直接的な接触がなくとも、剣を相手に向ける行動だけで攻撃優先権が交代することがあるため、剣先の位置を含む情報で推定することが求められる。そこで、本研究では、「攻撃優先権」は以下のように分類し、画像のみを用いて、攻撃優先権を上述の基準に基づき推定する。

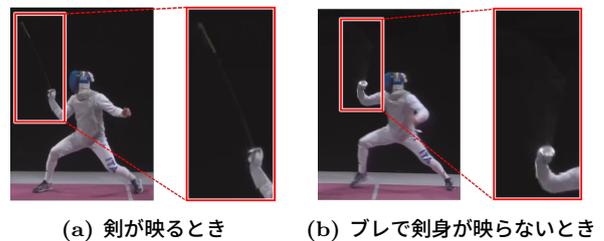


図 2 剣の映り方

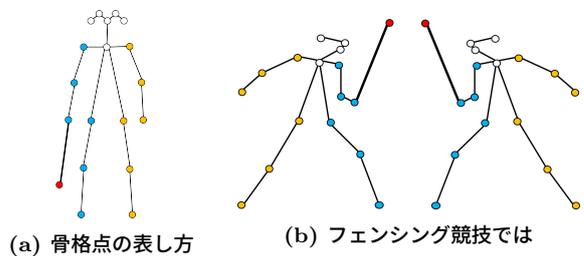


図 3 骨格点の入れ替え

- **right**: 右選手が攻撃優先権を保持している。
- **left**: 左選手が攻撃優先権を保持している。
- **none**: どちらの選手も保持していない。

### 2.2 剣先の検出

#### 2.2.1 骨格点と剣身の取得

骨格点情報の取得には VitPose[9] を用いて画像から骨格点を推定する。検出した骨格点は図 3a のように剣を持っている側を常に右側と認識させる。左手で剣を持っている場合は骨格点を左右入れ替える処理を行っている。フェンシング競技が始まる直前は両選手とも図 3b のように構えておりこの時の、相手の骨格

点に近い側を剣を持っている側の骨格点として判断させている。

剣全体の情報取得には YOLACT[10, 11] を使用する。処理速度の速いリアルタイムインスタンスセグメンテーションを採用し図4にあるように剣全体と手首までを含めた部分を剣全体として学習データを作成し、200枚のデータセットで学習を行い検出器を作成した。

### 2.2.2 剣先の特定

検出された骨格点と剣身のマスク情報を使用して剣先の特定を行う。まず剣を持っている手首からしきい値内にある剣情報をすべて取得し、その中でも一番信頼度の高いものを当選手が持つ剣と認識する。次に、その特定されたマスク情報の手首から遠いマスクの座標を剣先として認識する。これにより両選手それぞれの剣先を認識している。

### 2.2.3 補間

剣身の検出や剣先を特定できなかった場合は過去の剣先の座標情報から剣先位置を推定し、補間を行う。推定ネットワークには Temporal Convolutional Networks (TCN)[12] を使用し、手首の座標を原点とした過去 10 フレーム分の剣先の軌道データを使用して未検出フレームの剣先推定を行う。

### 2.3 攻撃優先権の判定

優先権判定器は取得された骨格点情報と剣情報を特徴量とし、攻撃優先権の推定を行う。ネットワークの構造は図2に示す。剣先の情報を含んだ骨格点情報をそれぞれ重みを共有しない2つのTCNベースのネットワークに入力し、時系列の骨格点情報を特徴量に落とし込み、2つの特徴量をコンカチネートし、全結合層の分類器により出力を得る構造となっている。骨格点に関しては、後ろ腰の位置を原点としたそれぞれの選手の座標に対応した相対座標となっている。

## 3 実験

### 3.1 剣先検出の実験

#### 実験条件

本実験では、フェンシングの試合動画から剣先位置を特定するシステムの評価を行った。事前学習モデル

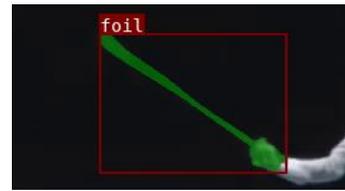


図4 剣と手首のマスク方法



図5 測定環境：V120-Trio



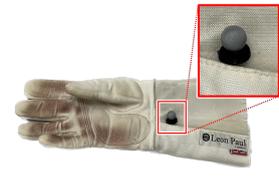
(a) 剣（通常時）



(b) 剣（マーカ装着時）



(c) 組み立て方法



(d) グローブへの装着方法

図6 マーカの装着（真値推定用）

にて検証を行う。学習データはモーションキャプチャにて取得されたデータは10フレーム分を1つのデータとして3600データである。前と後ろを向いている状況と剣の持ち手を左右と変えている各900個のデータを4種類で合計3600データとなっている。剣先には3Dプリンタで部品(図6c)を作成し、図6bにあるように反射マーカを装着した。手首部分にはマジックテープにてマーカを固定した(図6d)。実験検証データに関しては実際の試合動画からランダムに選んだ70フレームを計5セットを使用し、フェンシング経験者である我々が剣先の位置を手動で特定した。これらの位置を本実験における剣先の真値とした。システムは20ピクセル以内の誤差を正解と判定するようにした。

### 考察

結果を表1に示すように、平均で88.7%の精度で剣先を検出することに成功した。検出結果を図7に示す。図7aは検出と剣先の特定制とも成功している。図7bに関しては、右の選手の剣先検出が成功している、一方で左の選手は検出がうまくいっておらず手首にNoneの文字が出ているが予測された値により補間が行われている。図7cに関しては剣同士が近く、剣同士の区別がうまくいっておらず右選手の剣を自分の件として認識してしまっている。図7dに関しては、選手同士が重なっており、背後の選手の骨格点が適正な位置を推定できていないため剣先検出も誤った座標を出力している。今後は、ノイズに対応したシステムへの改善やフェンシング競技用に骨格点検出器を再学習させることで精度向上を図る。今回学習や実験では奥に人が映っておらず背景が比較的簡単なもので研究を行った。今後はより背景が複雑な場合など様々な場面に対応できるように汎用性の向上も目指す。

### 3.2 攻撃優先権判定の実験

#### 実験条件

本実験は、実際の試合状況の動画を使用し行われた。学習データの攻撃優先権はフェンシング経験者である我々が競技規則に基づき2688フレームにてアノテーションを行った。また検証データとしてそれぞれのフレームに著者が競技規則に基づき試合状況のアノテーションを行い、それらを真値としてすべてのシステムの精度評価を行った。合計500フレームの動画で実験を行った。

#### 実験条件

結果を図8に示す。認識率は平均で98.2%となった。図9には攻撃優先権を表示するデモを示した。画面上部に映っているものが攻撃優先権を示している。左選手が攻めているときは図9のように右向きに赤の矢印が表示される。右の選手が攻めている場合は左向きの緑の矢印が表示される。また、どちらの選手も攻撃優先権を保持していないときは何も表示されない。一貫して高精度出力を得ることが出来た。骨格点が未検出であった場合も同じようなデータが学習データに含まれていたため剣先検出の実験とは違わず安定した結果を得ることができた。

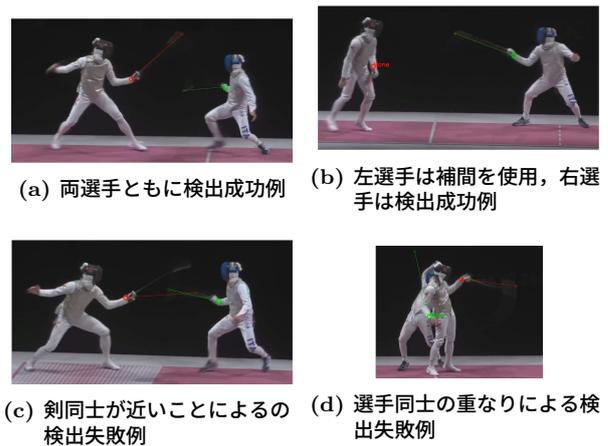


図7 剣先検出実験の結果

表1 剣検出検出実験の結果

Scene	Player	Accuracy[%]
1	Right	100.0
	Left	100.0
2	Right	86.7
	Left	93.3
3	Right	80.0
	Left	81.7
4	Right	100.0
	Left	100.0
5	Right	70.0
	Left	75.0

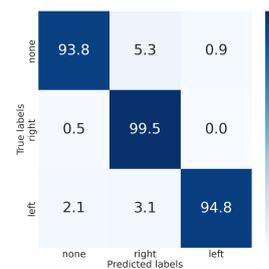


図8 混合行列：攻撃優先権判定実験の結果

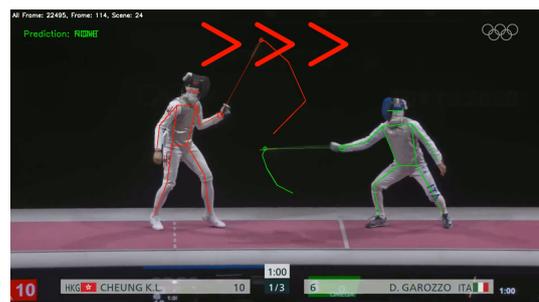


図9 攻撃優先権可視化のデモ

## 4 結論

本研究では、画像のみを用いたフェンシング競技での試合状況の認識システムの構築を行った。今まで検出が難しかった剣先の検出を行い、骨格点情報と組み合わせ攻撃優先権を判別する手法を提案し、実際の試合動画を使用した実験ではその有効性を検証した。今後の展望としては、処理速度の向上、動画から攻撃優先権出力まで一貫した end-to-end モデルの構築を目指す。

## 参考文献

- [1] K. Zhu, A. Wong, and J. McPhee, “FenceNet: Fine-grained Footwork Recognition in Fencing,” in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jun. 2022, pp. 3588–3597.
- [2] F. Malawski, “Real-Time First Person Perspective Tracking and Feedback System for Weapon Practice Support in Fencing,” Jan. 2018.
- [3] F. Malawski, “Immersive Feedback in Fencing Training using Mixed Reality,” *Computer Science*, vol. 23, no. 1, Art. no. 1, Mar. 2022.
- [4] Pageaud. A (2019, September 14), “Sport: Fencing Matches AI. Kaggle,” <https://www.kaggle.com/datasets/alexpgd/sport-fencing-matches-ai> / Accessed July. 10, 2023.
- [5] J. Mo, “Allez Go: Computer Vision and Audio Analysis for AI Fencing Referees,” *Journal of Student Research*, vol. 11, Nov. 2022.
- [6] M. Takahashi *et al*, “Sword tracer: visualization of sword trajectories in fencing,” in ACM SIGGRAPH 2018 Talks, Vancouver British Columbia Canada: ACM, Aug. 2018, pp. 1–2.
- [7] M. Takahashi *et al*, “Real-time visualization of sword trajectories in fencing matches,” *Multimed Tools Appl*, vol. 79, no. 35, pp. 26411–26425, Sep. 2020.
- [8] Y. Hanai *et al*, “Fencing tracking and visualization system,” in SIGGRAPH Asia 2021 Real-Time Live!, in SA ’21. New York, NY, USA: Association for Computing Machinery, Dec. 2021, p. 1.
- [9] Y. Xu *et al*, “ViTPose: Simple vision transformer baselines for human pose estimation,” arXiv:2204.12484, 2022
- [10] D. Bolya *et al*, “YOLACT: Real-time instance segmentation,” *Proc. ICCV*, pp. 9156–9165,, Oct. 2019.
- [11] D. Bolya *et al*, “YOLACT++ Better Real-Time Instance Segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 2, pp. 1108–1121, Feb. 2022.
- [12] S. Bai, J. Z. Kolter, and V. Koltun, “An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling.” arXiv, Apr. 19, 2018. doi: 10.48550/arXiv.1803.01271.