

# バドミントン選手の移動軌跡と骨格情報に基づく ショット情報推定

田中 直樹<sup>\*1</sup>宍戸 英彦<sup>\*2</sup>吹田 真士<sup>\*3</sup>西島 壮<sup>\*4</sup>亀田 能成<sup>\*2</sup>北原 格<sup>\*2</sup>

**Abstract** --- ショット情報はバドミントン競技のパフォーマンス分析にとって重要な基礎データであるが、手動でのショット情報の収集に多大な時間と労力を要するという問題が存在する。シャトルトラッキングによってショット情報の自動推定を試みた事例があるが、高速に移動するシャトルの追跡の難しさや、シャトルが画角外に見切れるなどの課題がある。本研究では、試合映像から選手の移動軌跡と骨格情報を推定し、推定した情報を元に深層学習を用いてヒットタイミングを推定する。また、ヒットタイミング周辺の選手の移動軌跡に基づき、戦術分析に有用な打点の地上投影位置を推定する。実証実験を通じて、f-measure が 95.9%の精度でヒットタイミングを推定し、RMSE が 0.54[m]の精度で打点の地上投影位置を推定できることを確認した。

**Keywords:** ショット情報, 移動軌跡, 骨格情報, 映像分析, バドミントン

## 1 はじめに

IT技術の発展に伴い、競技スポーツ分野においてデータ分析が重要視されている[1,2]。また、映像や画像を用いた分析は、情報を視覚的に確認できることから特に注目が集まっている。バドミントン競技では、選手のフォームチェックや戦術分析を通じた競技力向上を目的として、選手の身体情報 [3,4]やショット情報[5-7]の分析が盛んに研究されている。バドミントン競技は複数のラリーと呼ばれる球の打ち合いで構成され、各ラリーには様々なショットが含まれる。そのため、得失点に繋がるラリー展開を分析できるショット情報は、重要な基礎データとされている。

一般に、バドミントン競技では Sportscode[8]というソフトウェアを用いて、映像からショット情報を分析している[9]。Sportscode では、様々なスポーツ競技シーンをデータベース化することで、選手のパフォーマンスを分析可能である。一方でデータ収集は、選手がショットを放ったタイミングや得点パターンなどのデータを手動で入力する必要がある。バドミントンでは1時間を超える試合もあるため、手動入力によるデータ収集には多くの時間と労力がかかることが課題である。そこで本研究では、試合映像からショット情報を自動的に推定することで、データ収集の課題解消を目的とする。図1に示すように、データ収集を自動化することで、データ分析担当者がより多くの時間と労力を分析に費やすことができる。

シャトル追跡処理によってショット情報の自動推定を試みた研究が報告されている[5,6]。シャトルを追跡することで、シャトル位置および軌跡が取得可能となり各ショットの情報を推定できる。しかし、バドミントン競技の映像記録は、1台の固定カメラでコート撮影するのが一般的であるため、シャトルが映像から見切れてしまうフレームが存在する。また、ショット直後のシャトルは非常に高速に移動するため、ビデオレートの撮影では追跡が難しい状況が発生する。後者の課題は、複数の高解像度・高フレームレートのカメラで同期撮影が可能であれば、解消が可能であるが、大会などの競技現場でそのような計測を実施するのは困難である。

シャトルの情報をいわずに、選手がショットを放ったタイミングにおけるフォームの特徴に基づき、選手の骨格情報からショット情報の推定を試みた研究[7]が報告されている。しかし、バドミントン競技において想定されるフォームはシャトルの配球位置により無数に存在するため、すべてのフォームに対しショット情報を推定するまでには至っていない。よって骨格情報のみを利用する手法には、無数に存在するフォームすべてに対応できないという課題が存在する。

本研究では、単眼の試合映像からシャトルの情報をいれることなく、ショット情報の推定を目指す。ここで、本研究で推定対象とするショットの情報は、分析への応用可能性の観点から、選手がショットを放ったタイミング(ヒットタイミング)とラケットとシャトルが衝突した三次元位置をコート上に鉛直投影した位置(打点の地上投影位置)の二つとする。

\*1 筑波大学 大学院システム情報工学研究群

\*2 筑波大学 計算科学研究センター

\*3 筑波大学 体育系

\*4 東京都立大学 人間健康科学研究科

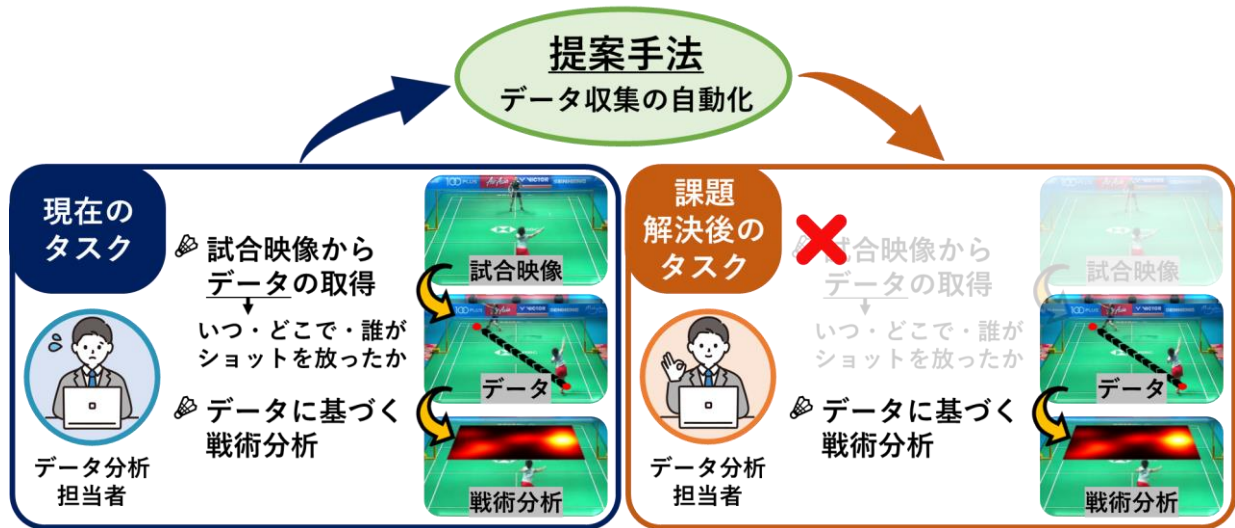


図1 データ収集の自動化によるデータ分析の効率化

Fig.1 Improving data collection efficiency by automating data collection

バドミントン試合映像において常に画角内に写る、選手の情報に基づきショット情報を推定する。3節で後述するように、映像から得られる選手の情報の中からショット情報の推定に効果的に利用できる特徴量として、選手の移動軌跡と骨格情報を設定する。

本研究の貢献は以下の3点である。

1. シャトルの写らないフレームを含む単眼試合映像からのショット情報の推定
2. ショット情報の推定に効果的に利用可能な特徴量の獲得
3. 様々なフォームから放たれるショット情報の推定

## 2 関連研究

### 2.1 人物姿勢推定

Vicon や OptiTrack のような三次元トラッキングセンサーは物体の姿勢を推定するために、身体にマーカーを装着する必要がある。このような手法は姿勢を高精度で推定可能であるが、選手の動きに制限がかかることから実際の試合への適用が困難である。

画像や映像から物体の姿勢が推定できる手法も存在する[10,11]。この手法には主にトップダウンとボトムアップという二つのアプローチがある。トップダウンは、はじめに物体検出手法を用いて人物領域を検出したのち、各検出領域内で人物の関節点を検出する。ボトムアップは、物体検出手法を用いずに画像中のすべての関節点を検出したのち、人の骨格が形成されるように関節点をつなぎ合わせる。トップダウンは一般に高精度であるが、人物数の増加に伴う処理時間の増加が著しいことが知られている。しかし、本研究で用いるバドミントンの試合映像に写るのは、固定の少数人数の選手であり、上述した悪影響を受けにくい。よって本研究ではトップ

ダウンの人物姿勢推定手法を採用する。

画像からの人物姿勢推定手法は、様々なスポーツに適用されている。例えば、Dejanらはスキージャンプの映像からジャンパーとスキー板の姿勢を推定し、ジャンプ中のジャンパーの姿勢を評価するシステムを提案した[12]。野中らはラグビー映像中のタックルシーンにおいて選手の姿勢を推定し、脳震盪の原因となり得る危険なタックルを判定するシステムを提案した[13]。Kaustubhらは卓球映像から選手の姿勢を推定し、時系列情報を入力とする深層学習を用いて、11種類の打法分類を99%以上の精度で達成した[14]。このように、スポーツ競技における選手の姿勢情報は、パフォーマンスを評価する上で重要な情報である。よって本研究におけるショット情報の推定においても、選手の姿勢情報を利用する。

### 2.2 バドミントン競技のショット情報に基づく分析

バドミントンの戦術分析においてショットの情報は重要である。Weiらはヒットタイミングにおける選手の姿勢情報から、ショットを6種類(5種類+その他)に分類した[15]。また、Weiらは各ショットの使用率に基づき試合の状況を分析した。さらに分析結果を実際の試合結果と比較し、ショット情報から得られる試合状況分析の結果と実際の試合結果が同等であることを確認した。本研究で推定対象とするショット情報は、バドミントンの戦術分析への応用可能性を考慮し選定する。

### 2.3 ショット情報の推定

ショット情報の推定手法としてシャトルの追跡を試みた研究が報告されている。Chenらは画像差分法を用いて画像中のシャトル位置を検出し、様々なショットにおけるシャトルの軌跡を比較した[5]。Nyanらは深層学習を用いたTracknet[16]や多項式カーブフィッティングをシ

ャトルの軌跡に適用し、ショット情報の推定を試みた[6]. これらの研究では、シャトルが画角内で常時鮮明に観察されていることが前提とされている. しかし、バドミントン競技映像は 1 台の固定カメラを用いて選手を中心に撮影するのが一般的であるため、シャトルが映像から見切れるフレームが存在し、その結果シャトル追跡が困難となる状況が存在する. そこで、シャトルを追跡せずにショット情報の推定を実現する手法が提案された[7]. ヒットタイミングにおける選手のフォームに表れる特徴を解明することで、選手の骨格情報からヒットタイミングを特定し、2 選手の打点を時系列に結ぶことでショットの方向を推定している. しかし、バドミントン競技において想定されるフォームはシャトルの配球位置により無数に存在するため、全てのフォームに対してヒットタイミングを特定し、ショット情報を推定するまでには至っていない.

### 3 ショット情報の推定に利用する特徴量

本研究では単眼映像から 2 人の選手の二次元骨格情報(1 人あたり 17 関節点)を取得し、深層学習モデルに入力することでショット情報を推定する. しかし、得られた骨格情報すべてを深層学習モデルに入力する場合、以下の点が懸念される.

- 人物骨格推定手法の関節推定誤差がショット情報推定に悪影響を及ぼす
- 入力次元数の多さが処理コストや過学習が起きる可能性を高める

そこで本研究では、得られた骨格情報の中から選手位置(両足首の midpoint)と手首位置(利き手首の腰からの相対位置)を深層学習モデルの入力情報とする. これらの特徴量は後述するバドミントンの試合特性を考慮して選択された. また、ヒットタイミングとの関連性を可視化することで特徴量としての妥当性を評価した.

本研究では、選手位置の特性がより顕著に表れるバドミントンシングルスを対象とする. シングルスでは通常、自陣コート中央付近をベースとし、自身がショットを放った後は、ベースに戻り相手選手の打球に備える移動特性がある. そのため、図 2(上)に示すように選手位置の時系列変化はヒットタイミングと関連がある. さらに、ヒットタイミングにおける選手位置は選手がショットを放つために移動した位置である. よって、ヒットタイミング周辺の選手位置と打点の地上投影位置は近距離にある. ラケットを保持する選手の利き手は、ヒットタイミングに最も身体中心から離れるという姿勢特性がある. よって、図 2(下)に示すように利き手首の腰からの相対位置もヒットタイミングと関連がある.

以上より、これらの特徴量はヒットタイミング周辺の時系列変化に特徴があり、ヒットタイミングと打点の地上投影位置を推定するために有用である.

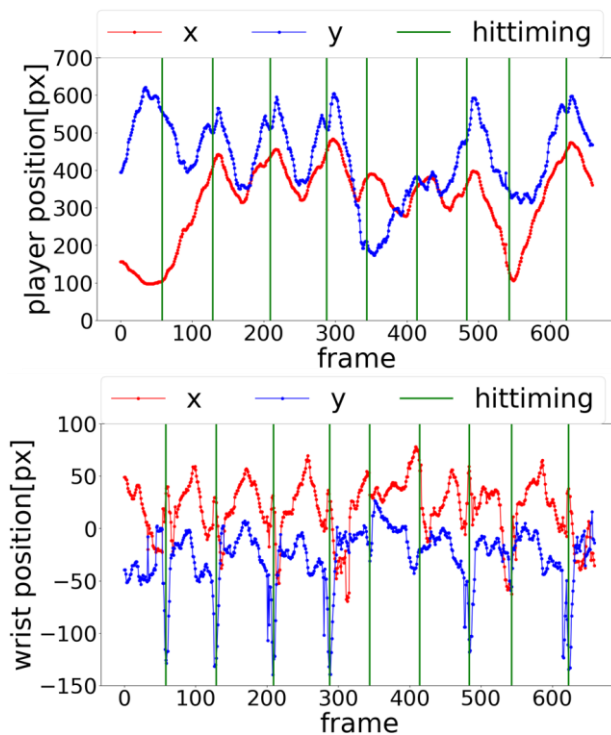


図2 選手位置(上)と手首位置(下)それぞれの時系列変化とヒットタイミングの関連性

Fig.2 Relationship between time-series changes in player position (upper) and wrist position (lower), respectively, and hit timing

### 4 ショット情報の推定

本節では、図 3 に示す提案モデルについて述べる. 単眼試合映像から既存の人物骨格推定手法[11]を用いて、選手の骨格を推定する. 次に前述したショット情報の推定に有用な特徴量を取得し、ヒットタイミングと打点の地上投影位置を推定する.

#### 4.1 ヒットタイミング推定

3 節で得られた選手位置と手首位置の時系列データを、ヒットタイミングを推定する深層学習モデルの入力情報とする. 選手位置はショットを放つ選手と受ける選手の 2 人分を利用し、手首位置はショットを放つ選手 1 人分のデータを利用する. 図 4 に示すように、これらの時系列データにステップ  $s$ 、長さ  $w$  のスライディングウィンドウを適用し、部分時系列データ群を生成する. モデルは各時刻の部分時系列データを入力とする LSTM[17]を使用する. 出力は入力された部分時系列データの時間内にヒットタイミングが含まれる確率である. 正解データとして、各部分時系列データに対し 0(ヒットタイミングを含まない)か 1(ヒットタイミングを含む)のラベルを付与する. モデルの出力と正解データの誤差が小さくなるように深層学習モデルを学習する.

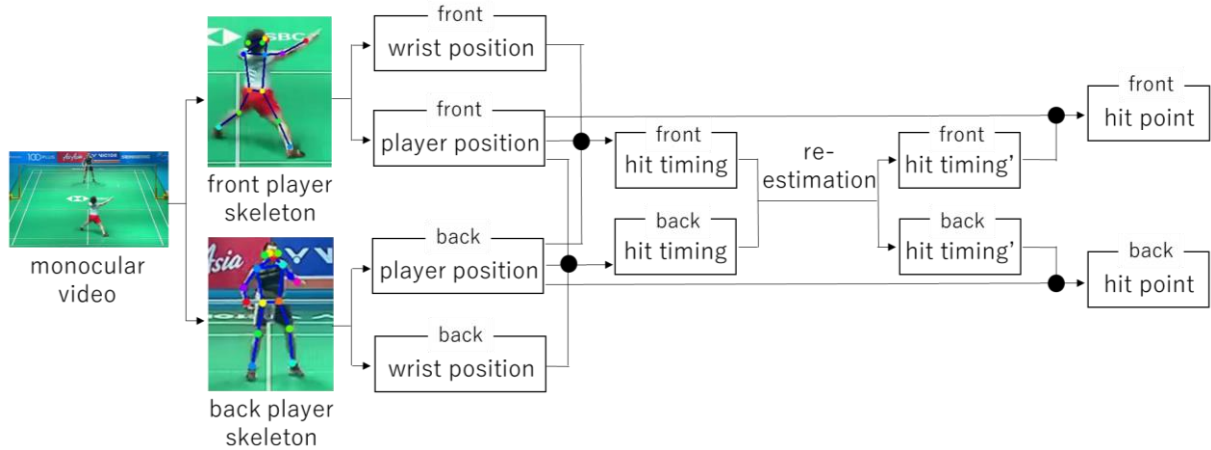


図3 単眼試合映像からヒットタイミングと打点の地上投影位置を推定する処理の流れ

Fig.3 Process flow for estimating hit timing and horizontal position of the hit point from monocular game video

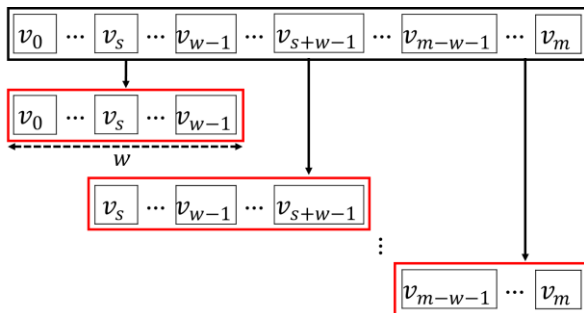


図4 スライディングウィンドウを適用することによる部分時系列データ群の生成

Fig.4 Generating a set of partial time series data by applying a sliding window

さらに推定精度向上のため、2人の選手(選手X, 選手Y)の推定結果に対し以下の条件を適用する。これは2人の選手が必ず交互にショットを放つというバドミントンの競技特性を利用した条件である。

- 選手Xのヒットタイミングが連続している場合、連続する二つのヒットタイミング間で、選手Yの出力確率が最大のフレームを選手Yのヒットタイミングとして再推定する。

上記の再推定の後、ヒットタイミングの推定結果と真値を比較することで、推定精度を評価する。

#### 4.2 打点の地上投影位置推定

打点の地上投影位置推定には、3節で得た選手位置と4.1節で得られたヒットタイミングを利用する。各ヒットタイミングを中心とする長さ $w$ の部分時系列データを抽出する。推定モデルは、各ヒットタイミングの部分時系列データを入力とするLSTMを用いて、打点の地上投影位置(二次元座標)を出力する。正解データは、試合映像から目視で確認した打点の地上投影位置とし、モデルの出力と正解データの誤差が小さくなるようにモデルを学習する。

## 5 実証実験

### 5.1 実験環境の設定

実験に用いたデータには、世界ランク上位選手による3本のシングルス映像(計5セット)を用いる。各セットをVideo No.1~No.5としたとき、No.1~No.4はすべて同一の女子選手らによる試合映像であり、No.5は男子選手らによる試合映像である。ヒットタイミングや打点の地上投影位置を推定するモデルにおいて、No.1~No.3を訓練データ、No.4をテストデータとする。No.5は5.2節で後述するようにモデルの汎用性を確認するためのデータとして使用する。

選手の骨格推定にはHRNet[11]を用いる。HRNetは二次元画像中の人物の骨格形状を深層学習によって推定するアルゴリズムであり、画像内に複数の人が混在する場合でも推定可能である。推定アルゴリズム中のスライディングウィンドウで使用する各パラメータは $s=5$ ,  $w=40$ とした。 $s$ はデータ数とデータ間の類似度に影響することから、データの量と質を考慮し設定した。 $w$ は各部分時系列データが二つ以上のヒットタイミングを含まない範囲で、より大きな値とした。

### 5.2 ヒットタイミング推定の実験

本研究では、推定されたヒットタイミングは打点の地上投影位置推定に利用される。打点の地上投影位置推定のモデルには、ヒットタイミングを中心とした数十フレームの選手位置が入力される。よって、ヒットタイミング推定結果における数フレームの誤差が与える影響は小さく、一定の誤差( $E$ フレーム)は許容できる。本実験では、モデルの入力時間長( $w$ )の半分の誤差までであれば、モデルの入力時間内にヒットタイミングの真値が含まれることを考慮し、 $E=w/2=20$ フレーム(約0.67秒)を許容誤差とした。つまり、真値から20フレーム以内の推定結果を予想正解と定義した。

表 1 様々な入力パターンに対するコート手前の選手のヒットタイミング推定結果

Table 1 Hit timing estimation results for front court player for various input patterns

input	recall(%)	precision(%)	f_measure(%)
$pos_f$	79.3	82.1	80.7
$pos_b$	76.0	79.3	77.6
$wri_f$	91.7	94.1	92.9
$pos_f, pos_b$	80.2	78.9	79.5
$pos_f, wri_f$	90.1	91.6	90.8
$pos_b, wri_f$	95.9	94.3	95.1
$pos_f, pos_b, wri_f$	93.4	<b>96.6</b>	95.0
$pos_f, pos_b, wri_f + re-est.$	<b>96.7</b>	95.1	<b>95.9</b>

表 2 訓練データに含まれない手前コートの選手のヒットタイミング推定結果

Table 2 Hit timing estimation results for front court player not included in the training data

input	recall(%)	precision(%)	f_measure(%)
$pos_f, pos_b, wri_f + re-est.$	94.3	95.6	94.9

評価指標として, recall, precision, f-measure を使用する. 表 1 に手前コートの選手におけるヒットタイミングの推定結果を示す. ヒットタイミング推定に用いた 3 種類の入力データを  $pos_f$  (手前コートの選手位置),  $pos_b$  (奥コートの選手位置),  $wri_f$  (手前コートの選手の手首位置) と定義する. これらのデータを組み合わせ入力した結果をそれぞれ示す. 4.1 節で述べた再推定の処理は  $+re-est.$  と定義する. 本研究の最終的な提案モデルである表 1 最下段のモデルの推定精度は f-measure が 95.9% であった.

さらに, モデルの汎用性を評価するため訓練データに含まれない選手の試合データ (Video No.5) においても最終的な提案モデルを適用した. Video No.4 と同様の条件で実験した結果を表 2 に示す. 異なる選手においても高精度にヒットタイミングを推定できていることから, 提案モデルが選手全体に共通する特徴を学習できていることを確認した. したがって, 提案モデルは選手ごとの特性やフォームの違いにかかわらず, ヒットタイミングを推定可能であることがわかる.

### 5.3 打点の地上投影位置推定の実験

打点の地上投影位置推定にはヒットタイミングが利用されるため, ヒットタイミングの推定精度が打点の地上投影位置の推定精度に影響を及ぼす. 本実験では, 打点の地上投影位置の推定モデル単体の推定精度を評価するため, ヒットタイミングには真値を利用する. 出力される二次元座標は, 図 5 に示すように x 軸をネットと平行に, y 軸をネットと垂直にとり, コート左上の黄色の点を原点とするコート座標上の点とする. 実際のバドミントン

シングルス自陣コートのサイズ (5.18[m] × 6.70[m]) を考慮し, 出力座標は  $x=[0\sim 5.18]$ ,  $y=[0\sim 6.70]$  とする. 評価指標として RMSE (二乗平均平方根誤差) を使用する. テストデータ全体に対するモデルの出力と真値の RMSE は 0.54[m] であった.

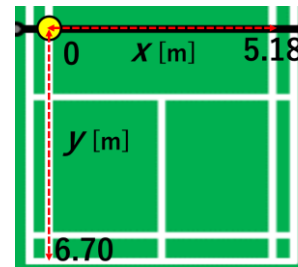


図 5 打点の地上投影位置を出力するコート座標

Fig.5 Court coordinates to output the horizontal position of hit point

## 6 戦術分析への応用

本節では, 5.2 節, 5.3 節で推定したショット情報に基づく戦術分析を実施し, 推定した情報の分析への応用可能性について述べる. 分析の例として, コート半面を九つ (3×3) の領域に分割する分析手法の一つを示す [18,19]. この分析では, 特定の領域からのショット方向の割合に基づく各選手の特徴を明らかにする. ショット方向は 2 人の選手の打点の地上投影位置を時系列に接続することで推定できる. ショット方向の割合は各選手の配球の特徴を示す指標として戦術分析に利用できる.

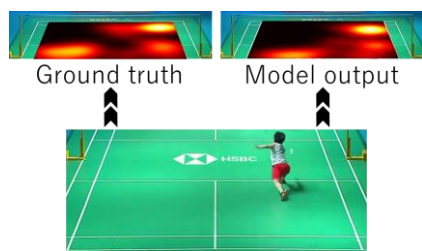


図 6 コート右前の領域からのショット方向の割合

Fig.6 Shot direction percentage from the right front area of the court

また、ショット方向の割合をヒートマップとして描画することで、選手の特徴を可視化する。図 6 に手前コートの選手のコート右前の領域からのショット方向の割合を、ヒートマップで可視化した結果を示す。推定した情報と真値それぞれから得られたショット方向の割合を比較すると、二つのヒートマップに大きな差がないことがわかる。以上より、本手法で推定したヒットタイミングと打点の地上投影位置は、真値に基づく分析と同等の分析結果を導けることを確認した。

## 7 おわりに

本研究では、選手の移動軌跡と骨格情報に基づくショット情報の推定手法を提案した。ヒットタイミングに特徴を持つ移動軌跡と骨格情報の時系列データを入力とする深層学習モデルを用いてヒットタイミングを推定した。さらに、ヒットタイミング周辺の選手の移動軌跡から打点の地上投影位置を推定した。実証実験の結果、手前コートの選手のヒットタイミングの推定精度 (f-measure) は 95.9% であり、打点の地上投影位置の推定精度は RMSE が 0.54[m] であった。さらに、推定した情報と真値それぞれに基づく戦術分析を定性的に評価し、推定した情報から真値と同等の戦術分析ができることを確認した。

## 参考文献

[1] Pascual J.Figueroa, Neucimer J.Leite, Ricardo M.L.Barros, (2006) "Tracking soccer players aiming their kinematical motion analysis," Computer Vision and Image Understanding Vol.101, pp.122-135.  
 [2] S. Messelodi, C.M.Modena, V.Ropele, S.Marcon, M.Sgro, (2019) "A Low-Cost Computer Vision System for Real-Time Tennis Analysis," ICIAP, pp.106-116.  
 [3] Kokum Weeratunga, Anuja Dharmaratne, Khoo Boon How, (2017) "Application of computer vision and vector space model for tactical movement classification in badminton," IEEE Computer Society Conference on CVPR Workshops, vol.2017, pp.132-138.  
 [4] David Alder, Paul R.Ford, Joe Causer, A.Mark Williams, (2014) "The coupling between gaze behavior and opponent kinematics during anticipation of

badminton shots," Human Movement Science, Vol.37, pp.167- 179, Oct.

[5] Chen Bingqi, Wang Zhiqiang, (2007) "A Statistical Method for Analysis of Technical Data of a Badminton Match Based on 2-D Seriate Images," Tsinghua Science and Technology ISSN 1007-0214 14/19 pp.594-601, Vol.12, No.5.

[6] Nyan-Ping Ju, Dung-Ru Yu, Tsi-Ui Ik, Wen-Chin Peng, (2020) "Trajectory-based Badminton Shots Detection," ICPAI, pp.64-71.

[7] Yui Yoshikawa, Hidehiko Shishido, Masashi Suita, Yoshinari Kameda, Itaru Kitahara, (2021) "Shot detection using skeleton position in badminton videos," SPIE Proceedings, Vol.11766, IWAIT.

[8] Hudl Sportcode, <https://www.hudl.com/products/sportcode>

[9] 平野 加奈子, 飯塚 太郎, 鳥賀陽 真未子, (2017) "リオデジャネイロオリンピックに向けたバドミントン日本代表に対する映像支援," Sports Science in Elite Athlete Support Vol.2, pp.1-9.

[10] Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh, (2019) "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," IEEE Transactions on Pattern Analysis and Machine Intelligence, 14 pages.

[11] K. Sun, B. Xiao, D. Liu, J. Wang, (2019) "Deep High-Resolution Representation Learning for Human Pose Estimation," CVPR, pp.5693-5703.

[12] Dejan Stepec, Danijel Skocaj, (2022) "Video-Based Ski Jump Style Scoring from Pose Trajectory," IEEE/CVF Winter Conference on Applications of Computer Vision Workshops, pp.682-690.

[13] Naoki Nonaka, Ryo Fujiwara, Monami Nishio, Hidetaka Murakami, Takuya Tajima, Mutsuo Yamada, Akira Maeda, Jun Seita, (2022) "End-to-End High-Risk Tackle Detection System for Rugby," IEEE/CVF Conference on CVPR Workshops, pp.3549-3558.

[14] Kaustubh Milind Kulkarni, Sucheth Shenoy, (2021) "Table Tennis Stroke Recognition Using Two-Dimensional Human Pose Estimation," IEEE/CVF Conference on CVPR Workshops, pp.4571-4579.

[15] Wei-Ta Chu, Samuel Situmeang, (2017) "Badminton Video Analysis based on Spatiotemporal and Stroke Features," ICMR, pp.448-451.

[16] Y.-C. Huang, I.-N. Liao, C.-H. Chen, T.-U. Ik, and W.-C. Peng, (2019) "Tracknet: A deep learning network for tracking high-speed and tiny objects in sports applications," CAVA in AVSS.

[17] Sepp Hochreiter; Jürgen Schmidhuber, (1997) "Long short-term memory," Neural Computation 9 (8): 1735-1780.

[18] F. Josue, M. F. Abdullah, I. Zulkapri et al., (2020) "Movement pattern in term of court coverage among top international male and female badminton players during BWF World Championships 2013," Jurnal Sains Sukan & Pendidikan Jasmani, vol. 9, no. 1, pp. 9-14.

[19] Careelmont S, (2013) "Badminton shot classification in compressed video with baseline angled camera," Master Thesis, University of Ghent.