

深層学習を用いて卓球放送映像から獲得する ボール軌道と戦術知表現

○ 林 勲 (関西大学大学院) 馮 楊蘊 (関西大学大学院) 入江 穂乃香 (関西大学大学院)

Tactics Knowledge Representation Based on Ball Trajectory Acquired
from Broadcast Video of Table Tennis by Deep Learning

*I. Hayashi (Kansai University), Y. Feng (Kansai University), H. Irie (Kansai University)

Abstract— Recently, we have been developing a system that automatically acquires tactics and strategies of the table tennis match from broadcast video. In this system, the input and output data are constructed by automatically extracting the ball position and the player position from the broadcast video. During the match, an algorithm removes noise and estimates the ball and player position. In this paper, we introduce the motion tracking system. In the motion tracking system, the ball trajectory and player position are automatically extracted from the 30fps broadcast video and converted into two-dimensional coordinates. The ball trajectory is estimated by preprocessing with the white blob extraction process and RGB extraction process, and then the player's skeleton position is estimated by CenterNet of deep learning (DNN). The position of the ball hidden at the body is estimated using the Kalman filter and the bicubic interpolation method. Finally, we discuss the future image of this system which acquires the table tennis strategy and makes the strategy visible to directors and coaches using the if-then rule by fuzzy ensemble learning.

Index terms— Table Tennis, Video Analysis, Ball Trajectory, Tactics and Strategy Acquisition, Deep Learning

1 はじめに

最近、スポーツ界への人工知能の応用が切望されている [1, 2]. 従来のスポーツ技能の研究では、技能スキルは単機能とメタ機能の階層構造から構成されるとして、身体的構造モデルや骨格構造モデルを用いて、内部モデル [3] を議論してきた。しかし、最近では、内部モデルに深層学習やデータサイエンスの AI モデルを用いてスポーツ技能の評価や動作解析、戦略分析等の研究が行われている [1, 4, 5].

卓球は、勝負の作戦を論理的に組み立てるスポーツといわれ、戦術と戦略は定式化されている。我々が研究を進めているのは、視覚化可能なデータサイエンスの卓球 AI モデルである。その目的としては、卓球トップ選手へのサポートツールや一般競技者への健康増進ツールとして、「AI コーチ」の実現を目指す [6-11]。本論文では、この AI 卓球の一例として、放送映像から選手の戦術と戦略を獲得する卓球戦略獲得システムについて議論する [11-13]。システムの実装には、映像抽出モジュールと戦略獲得モジュール、戦略表示モジュールが必要となる。特に、映像抽出モジュールでは、放送映像から対戦シーンのみを抽出し、深層学習の CenterNet [?] を用いて、選手の身体位置とボール軌道、ラリー軌道等を自動抽出する。また、戦略獲得モジュールでは、そのラリー軌道等から戦術戦略の知識を獲得する。一方、監督やコーチによる選手へのコーチングが勝負や技能改善を大きく左右するため [14]、監督やコーチへの戦略表示モジュールも必要となる。ここでは、2016 年のリオ・デ・ジャネイロオリンピック大会の卓球女子シングルの放送映像を用いてラリー軌道を抽出し、特定の選手の規則化された戦術と戦略を獲得する。

将来的には、アンサンブル学習型ファジィクラスタリングモデル [15] により、戦略を if-then ルール [16] として獲得することで、戦略を視覚化することも可能となる。この戦略表示機能の将来像についても議論する。

2 卓球戦略獲得システム

卓球 [17] とは、競技者が卓球台を挟んで向かい合い、プラスチック製のボールをラケットで打ち合って得点を競う競技である。卓球台の上面は、長さ 2.74m、幅 1.525m の長方形であり、地面より 76cm の高さで台は水平に置かれる。ネットは卓球台の長辺の等分の位置に台から 15.25cm の高さに垂直に張られる。ボールは直径 40mm のプラスチック製であり白色で無光沢のものが使用される。ボールはワンバウンドで相手コートに返球する。試合形式は 2001 年 9 月 1 日より変更され、現在は、各ゲーム 11 点先取の 7 ゲーム制 (4 ゲーム先取) か 5 ゲーム制 (3 ゲーム先取)、あるいは、3 ゲーム制 (2 ゲーム先取) で行われる。ただし、10 対 10 以後は、先に 2 点差を付けた方が 1 ゲーム先取となる。サービスは 2 本交代であるが、10 対 10 以降は 1 本交代となる。

卓球は、試合でサービスから 3 球目までに得点した場合、その試合の勝率が 7 割程度となる。そのため、選手は、特にサービスに注力し、サーブ時には、あらかじめパターン化している攻撃戦略が有利となるようなサービスを実行する。特に、選手は過去の試合の戦歴と相手選手の弱点を記憶しており、最適な攻撃パターンとサービスの種類を学習している。したがって、ここでは、サービスから 3 球目までの攻撃パターンの戦略の獲得のみに特化する。また、試合会場では、基本的には三脚を使用するカメラによる映像撮影は許容されていない。使用を許可されているのは、選手が持参するカメラと放送用カメラのみである。そこで、戦略の獲得は放送映像を原動画とする。Fig.1 に卓球戦略獲得システムの概要を示す。将来的に実現するシステムでは、試合中に選手が打ったボール軌道とサーブ回転、選手位置を自動抽出し、ボール色と同じ他の白色を画像から抜き出し、卓球台上のボール位置の座標を獲得して、仮想データ (バーチャルデータ) を発生する

ファジアンサンブル学習により、3球目攻撃の戦略をif-thenルールで獲得する。このファジールールによる「戦略知識の可視化」により、選手やコーチへの教示が可能となる。本システムは、映像抽出モジュールと戦略獲得モジュール、戦略表示モジュールの3つのモジュールから構成されるが、現在は映像抽出モジュールの一部のみが完成している。

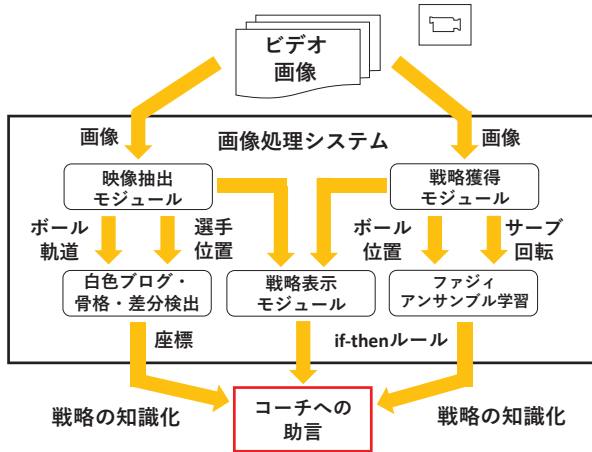


Fig. 1: Concept of Table Tennis Strategy Acquisition System

3 映像抽出モジュール

映像抽出モジュールは、30fpsの放送映像からボール軌道とサーブ回転、選手位置を自動抽出し、2次元座標に変換する。映像抽出モジュールは、Python3.7.6とOpenCV4.2.0により構築した。処理過程をFig.2に示す。ボール軌道は、放送映像を各ラリーごとに自動分割し、その後、白色ボールのための白色映像検出過程とフレーム間差分検出過程により推定される。

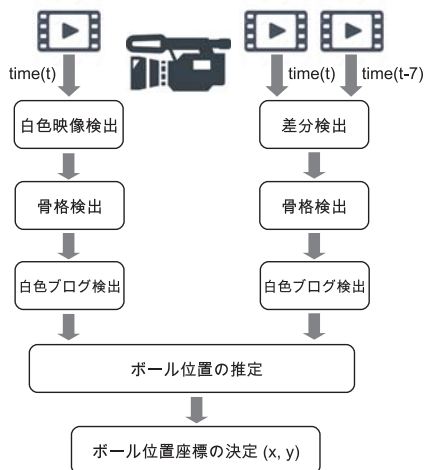


Fig. 2: Process of Video Extraction Module

白色映像検出過程では、まず、RGBのそれぞれのカラースケールをしきい値によって変更し画像を変形する。次に、CenterNetによって選手の骨格位置を学習から推定した。Fig.3にCenterNetによって獲得した選手の骨格情報を加えた放送映像の一例 [18] を示す。

CenterNetへの入力はmp4形式の映像動画であり、出力は17点の2次元の骨格位置座標とそれらの骨格領域を決定する2次元座標の2点である。したがって、2名の選手から、出力全体の変数は、 $17 \times 2 \times 2 + 2 \times 2 \times 2 = 76$ 個となる。CenterNetによる選手骨格位置の推定後、その結果を用いて、選手のラケットとテーピング、ゼッケン文字、靴等の白色を消去し、白色ボールを検出した。



Fig. 3: Broadcast Video with Body Coordinates Added by CenterNet

一方、フレーム間差分検出過程では、まず、現フレームの画像と7フレーム前の画像の間でRGBのカラースケールの差分を計算した。次に、白色映像検出過程のCenterNetを用いた選手の骨格位置から、選手のラケットとテーピング、ゼッケン文字、靴等の白色プログ部分を消去し、骨格領域以外の白色プログ部分を検出した。最後に、次フレームのボール位置を推定した。ボール位置の推定は、その領域を半径80pixelの円内で、かつ、卓球台に10pixel分縮小した内部範囲の白色プログ部分(ボール候補群)に限定して抽出した。ただし、検出が不能の場合には、1フレームごとに半径を20pixelずつ拡大した、最後に、カルマンフィルタやバイキューブ補間法を用いて、ボールが身体等に隠れている場合(オクルージョン)のボール位置の補間を行い、ボール軌道の座標(x, y)を推定した。

Fig.4に、白色プログ抽出過程とフレーム間差分抽出過程によるボール位置の推定イメージの例を示す。国際卓球連盟(ITTF)から2016年のリオ・デ・ジャネイロオリンピック大会の決勝トーナメント戦での卓球試合の放送映像の提供を受けた [18]。対象とした放送映像は、2016年のリオ・デ・ジャネイロオリンピック大会卓球女子シングルス準決勝、日本の福原愛選手と中国の李晓霞選手の試合映像である。第4ゲームの福原0-4李のあとの第5ラリーの映像である。Fig.4の最下部にその放送映像のフレーム画像を示す。手前側の福原選手が奥側の李選手に対して、フォア側のボールに反応して打つ瞬間の画像である。ボールは白色であるが、映像では赤色にしている。最上段の図は、白色映像検出過程でのCenterNetにより選手の骨格の検出と白色プログ部分を検出したフレーム画像の結果である。選手の骨格と白色ボールだけでなく、五輪マークの白色、卓球台の白枠、ネット上部の白色部分、映像上の得点枠の白色等、多くの白色プログ部分が検出されていることがわかる。中間図は、白色映像検出過程の映像にフレーム間差分検出過程のCenterNetを適用することに

より、再度、選手の骨格情報を検出している。また、白色ブログ領域から、五輪マーク、卓球台の枠、ネット上部部分、映像上の得点枠が消去され、ボールの白色と選手の骨格のみが検出されている。これらの画像処理により白色ボールの軌道位置が検出できている。ボール軌道の検出後、映像での長さ 2740mm、幅 1525mm の卓球台 (台形) をアフィン変換によって長さ 341.5pixel、幅 155pixel の長方形に変換し、ボール軌道の位置座標を推定する。



Fig. 4: Images of Video Extraction Module

Fig.4 の映像では、次のようなラリーの経過となっている。

- (1) 福原選手が、卓球台の左端 (x :小, y :小) から李選手のフォア中間領域 (x :小, y :中) へサービスを打つ。
- (2) 李選手が福原選手のフォア側手前領域 (x :大, y :小) に返球する。
- (3) 福原選手が李選手のフォア奥領域 (x :小, y :大) にスマッシュを打つ。
- (4) スマッシュ以降のボール軌道は福原選手の身体によってオクルージョンが発生する。

Fig.4 における福原選手と李選手のラリーの映像に対して、手入力でマーキングした真値のボール軌道と本モジュールにより推定したボール軌道を Fig.5 に示す。上図が x 方向のボール軌道、下図が y 方向のボール軌道である。ただし、放送映像での卓球台位置の左下端を原点、横方向を x 方向、奥行方向を y 方向としている。したがって、上図では縦軸の原点が卓球台の x 方向の左端を表し、下図では縦軸の原点が卓球台の y 方向の手前側 (卓球台の左下端) を表す。また、赤色の実線が真値のボール軌道、青色の実線が推定のボール軌道、緑色の破線が卓球台の範囲を表し、横軸がフレーム番号の時間軸を表す。

Table 1: Average Errors between Correct Trajectory and Estimated Trajectory

フレーム帯	単位 : mm		
	誤差 (x)	誤差 (y)	誤差 (x,y)
45089~45107	10.98	20.43	24.57
45108~45122	30.61	29.16	43.04
45123~45134	865.83	456.12	980.60
45089~45134	334.13	184.71	384.81

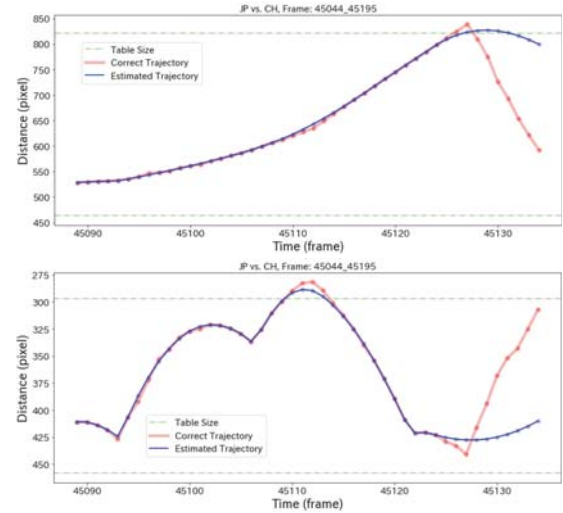


Fig. 5: Estimated Ball Trajectory by Video Extraction Module

映像中のボール軌道は、 x 方向では卓球台の左端から右側方向へ移動するので、Fig.5 の上図では、ボール軌道は増加傾向となっている。スマッシュ以前のボールの推定軌道は真値のボール軌道とほぼ一致しているが、45127 フレーム以後の推定軌道はオクルージョンの理由により正しく推定されておらず、オクルージョンの解消は今後の課題である。 y 方向では、福原選手の自コートでのサービスのボール着弾点と李選手のコートでのボール着弾点が、それぞれ、45093 フレームと 45107 フレーム付近で推測できている。また、李選手がフレーム時刻 45112 付近で返球し、福原選手がフレーム時刻 45122 付近の着弾点のボールに対してフレーム時刻 45127 付近で返球している。ただし、フレーム時刻 45127 以後のボール軌道はオクルージョンの理由により正しく推定されていない。

手入力によるボール軌道の真値の座標位置と本モジュールによるボール軌道の推定の座標位置との誤差を表 1 に示し、卓球台上のボール軌道とラリー軌道を Fig.6 に示す。表 1 では、ボールの着弾点ごとに全フレームを分割し、着弾点の x 方向と y 方向、2次元距離の誤差を示した。最後に、全フレームの誤差も示した。45123~45134 のフレーム帯で、福原選手が打ったボールが自身の身体でオクルージョンされ、それぞれの誤差が大きくなっている。しかし、45122 フレームまでの誤差は、10mm から 45mm 程度であり、本モジュールが真値のボール軌道を精度良く推定していることを示している。

一方、Fig.6 では、上図にフレームごとの真値と推定

のボール軌道を示し、下図にボールの着弾点から生成したラリー軌道を示す。オクルージョンにより、ボール軌道とラリー軌道の最終時には、真値と推定値で大きく異なっているが、それまでの軌道はほぼ一致していることから、本モジュールにより、放送映像からボール軌道とラリー軌道を精度良く獲得できていることがわかる。

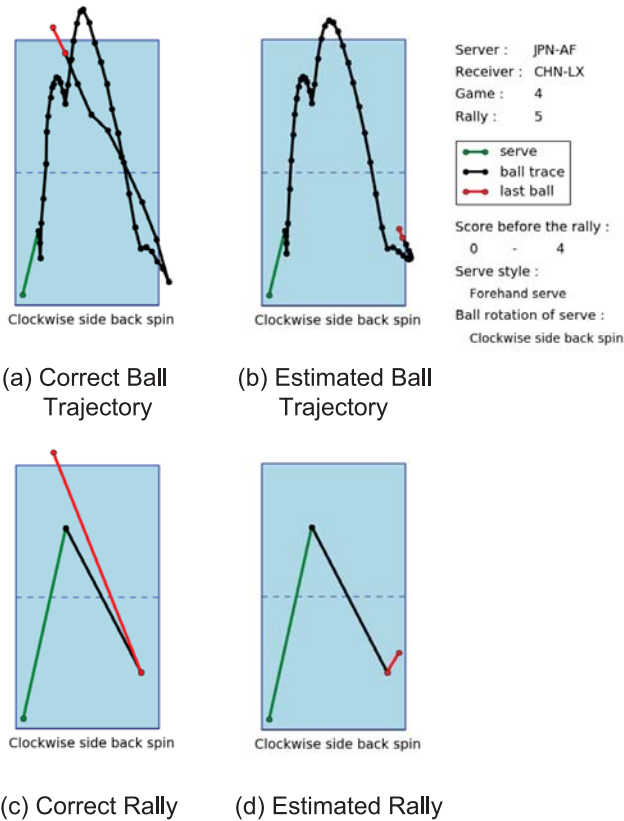


Fig. 6: Ball Trajectory and Rally

4 戦術知の表現

卓球戦略獲得システムでは、サービスからの3球目攻撃パターンを抽出し、その戦術戦略を獲得する。ただし、本システムは現在構築中のため、ここでは、ボール軌道から推定した着弾位置によるラリー軌道の戦略獲得の一部のみを説明する。なお、完成版では、ボール軌道の着弾位置だけでなく、ボール軌道全体やボール回転、身体位置、スコア、サーブ種類等から戦略を獲得する。Fig.7に映像抽出モジュールによる放送映像のボール軌道の例を示す。また、Fig.8にボール着弾点からなるラリー軌道の例を示す。これらのラリー軌道は、2016年のリオ・デ・ジャネイロオリンピック大会卓球女子シングルス準決勝、日本の福原愛選手と中国の李暁霞選手の第2ゲームの試合映像から抽出されたものである。特に、Fig.8では、李暁霞選手がサービス権を得た場合の3球目攻撃パターンを抽出している。上部に得点のラリー軌道を示し、下部に失点のラリー軌道を示す。このように、ボール着弾点の全てのラリー軌道から特定条件に合致したラリー軌道のみを表示することもできる。なお、ラリー軌道の抽出は、放送映像を取得した後、約1時間程度で処理が可能である。

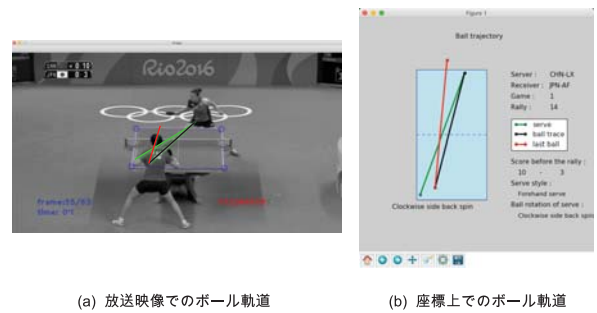


Fig. 7: Ball Trajectory by Video Extraction Module

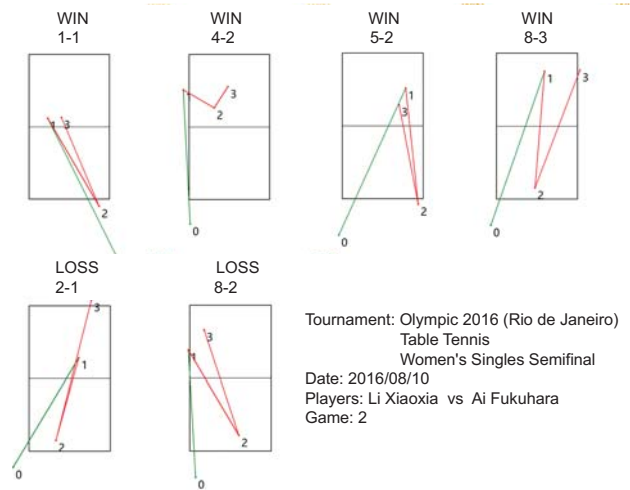


Fig. 8: Display of Ball Trajectories

Fig.9に戦略獲得モジュールでのボール着弾点による戦術と戦略の獲得方法を示す。例えば、いま対戦している相手等の同一選手が、第1攻撃パターン、第2攻撃パターン、第3攻撃パターンのボール着弾点で攻撃しているとする。これらの攻撃パターンは、そのラリー軌道は類似しているが、まずは過去の戦術戦略データベースと照合される。過去の戦術戦略データベースでは、ボール着弾点の全てのラリー軌道を記録保存しており、逐次、AIモデルを用いて同類のラリー軌道の一つのクラスとしてクラスタリングしている。それぞれのラリー軌道はデータベースの各クラスとの類似度を計算する。ラリー軌道が同じクラスを参照する場合、攻撃パターンとして認定され、クラス内のラリー軌道の平均値である代表軌道が決定される。ここでは、この代表軌道を戦術と定義する。Fig.9では、第2ゲームの最初の3本のラリー軌道が、データベースの同じクラスを参照しており、第1攻撃パターンとして認定されている。第1攻撃パターンの代表軌道は“ミドル攻撃”であるので、相手選手はまずはミドル攻撃の戦術で攻めていると推定する。その後、同様に、第2攻撃パターンをチキータ攻撃の戦術と推定し、第3攻撃パターンをクロス攻撃の戦術と推定する。ここでは、これらの戦術の連続形体を戦略と定義する。したがって、第2ゲームでは、ミドル攻撃 → チキータ攻撃 → クロス攻撃 [19] の3種類の戦術の連続形体から構成される戦略が獲得されている。

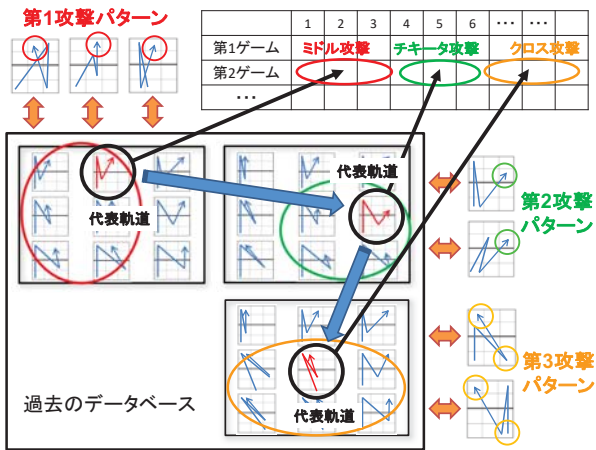


Fig. 9: Acquisition of Tactics and Strategy

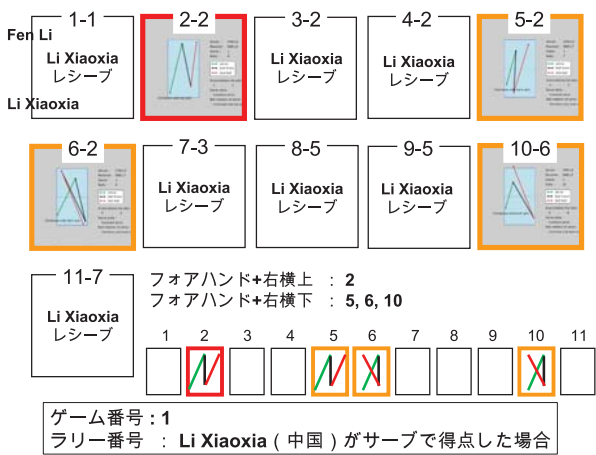


Fig. 10: Tactics of First Game

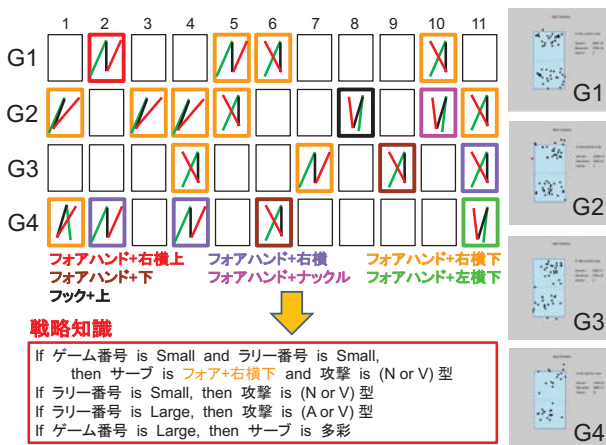


Fig. 11: Strategy of All Games

具体的な事例として、リオ・デ・ジャネイロオリンピック大会の決勝トーナメント戦の試合における合計372プレー機会中の6,862ラリーを解析した。ここでは、2016年8月8日の女子卓球シングルス3回戦、中国の李曉霞 (Li Xiaoxia) 選手とスウェーデンの Li Fen 選手の試合映像を例として、李選手のサービスからの3

球目攻撃の戦略の分析の結果を紹介する。Fig.10に第1ゲーム(セット)での李選手のサービス種類とボール着弾点から得られたラリー軌跡を示す。ただし、李選手がサービスを出し得点したラリー軌道のみを描画し、サービスの種類により色分けされている。また、下部にラリー軌道をN型とA型で記述し、サービスの種類を色別に識別した。

第2ゲームと第3ゲームも同様に解析し、全ゲームの全ラリーの戦略知識をFig.11に示す。行にゲーム番号、列にラリー番号を表した表形式に、サーブで得点した場合のラリー軌道とサーブ種類を示した。また、右側には、ボール着弾点の位置をゲームごとに示した。これらの解析結果から、李選手の3球目攻撃には一定の攻撃パターンの戦略がみられた。

5 おわりに

本論文では、放送映像から身体位置やボール軌道、ラリー軌道等を自動抽出し、戦術戦略の知識を自動獲得する卓球戦略獲得システムについて議論した。映像抽出モジュールでは、深層学習と白色ログ抽出アルゴリズムにより、高精度なボール軌道とラリー軌道を獲得することを示した。また、戦略獲得モジュールでは、ボール軌道とボール回転の推定から戦術戦略の知識を自動獲得する本システムの将来像を示した。

今後、本システムをより高精度にするため、ボール軌道のオクルージョン処理とボール回転認識の新たな手法を提案する必要がある。

6 謝辞

本研究を推進するにあたり、順天堂大学の吉田和人先生、名桜大学の玉城将先生、University of LjubljanaのProf.Miran Kondric、関西大学の竹中要一先生と学生の関谷侑希君、国際卓球連盟、日本卓球協会、国立スポーツ科学センター、オムロン株式会社研究開発センターの皆様には、多大なご助言を頂きました。この場を借りて厚く御礼を申し上げます。

なお、本研究の一部は、科学研究費補助金基盤研究(C)一般「バーチャルデータ発生型ファジィバグを用いた放送映像の卓球戦略獲得ボードの開発」(No.20K11981, 2020年~2025年)の助成を得た。

参考文献

- 1) 清水, 清田: スポーツ競技と AI, 人工知能, Vol.34, No.4, pp.492-496 (2019)
- 2) 青木: スポーツ競技を対象とした画像 AI 技術, 人工知能, Vol.34, No.4, pp.497-502 (2019)
- 3) 川人: 小脳内部モデルとモザイクの計算論的神経科学, 人工知能, Vol.17, No.6, pp.720-731 (2002)
- 4) I.Hayashi, M.Fujii, T.Maeda, J.Leveille, T.Tasaka: Extraction of Knowledge from the Topographic Attentive Mapping Network and its Application in Skill Analysis of Table Tennis, *Journal of Human Kinetics*, Vol.55, pp.39-54, DOI:10.1515/hukin-2017-0005 (2017).

- 5) 林, 藤井, 田阪, 王, 前田: TAM ネットワークによる卓球技能の知識獲得, 第 23 回人工知能学会全国大会論文集, 1K1-OS8-3 (2009)
- 6) 玉城, 斎藤, 吉田, 山田, 尾崎: 卓球のパフォーマンス分析とビジョン技術, *ViEW* ビジョン技術の実利用ワークショップ講演論文集, Paper ID: No.ROMBUNNO.OS3-O4 (2012)
- 7) 吉田, 山田, 玉城, 内藤, 加賀: 卓球のワールドクラスの試合におけるラリーの特徴 – ラリー中の打球回数に着目して –, *コーチ学研究*, Vol.28, No.1, pp.65-74 (2014)
- 8) 玉城, 斎藤, 吉田: ボールマークと軌跡に基づく卓球ボールの 3 次元回転計測, *電子情報通信学会論文誌*, Vol.J98-D, No.4, pp.639-650 (2015)
- 9) K.Muelling, A.Boularias, B.Mohler, B.Scholkopf, J.Peters: Learning strategies in table tennis using inverse reinforcement learning, *Biological Cybernetics*, Vol.108, pp.603-619 (2014).
- 10) 林, 入江, 関谷, 中山, 八瀬: データサイエンスから見る AI 卓球の可能性, *統計*, 6 月号, pp.23-28 (2020)
- 11) 林, 関谷, 入江, 荻野: 放送映像からの卓球競技の軌道追跡, 第 36 回ファジィシステムシンポジウム講演論文集, pp.263-268 (2020)
- 12) I.Hayashi, M.Ogino, H.Irie, S.Tamaki, K.Yoshida, M.Kondric: AI Coach: Learning Table Tennis Strategy Rules from Video, *Proc. of the 16th International Table Tennis Federation (ITTF) Sports Science Congress (ITTF-SSC2019)*, p.76 (2019).
- 13) I.Hayashi, M.Ogino, H.Irie, S.Tamaki, K.Yoshida, M.Kondric: Development of Image Processing System to Realize Table Tennis Strategy Board, *Proc. of Japan Table Tennis Association Sports Science and Medicine Committee International Meeting 2018 (JTTA-SSMC2018)*, p.21 (2018).
- 14) X.Zhou, D.Wang, P.Krähenbühl: Objects as Points, *Arxiv Sanity Preserver*, No.arXiv:1904.07850 (2019).
- 15) 入江, 林: 正誤バーチャルデータの発生による pdi-Bagging の特性評価, 第 29 回インテリジェント・システム・シンポジウム講演論文集, Paper ID: No.A3-3 (2019)
- 16) 入江, 林: 台形型メンバシップ関数による学習型ファジィ推論の設計評価, *知能と情報*, Vol.31, No.6, pp.908-917 (2019)
- 17) 日本卓球協会: 卓球の基本的なルール, <https://jtta.or.jp/rule/>, 2021 年 11 月 1 日現在 (2021)
- 18) 国際卓球連盟: itTV, <https://www.ittf.com/itTV/>, 2021 年 11 月 1 日現在 (2021)
- 19) 日本卓球協会: 観戦ガイド, <https://jtta.or.jp/guide/>, 2021 年 11 月 1 日現在 (2021)

[連絡先]

林 勲 関西大学大学院 総合情報学研究科
〒 569-1095 大阪府高槻市霊仙寺町 2-1-1
tel. 072-690-2448
fax. 072-690-2491
e.mail ihaya@kansai-u.ac.jp