

隠れマルコフモデルを用いたテニスにおけるラリー系列からのパターン発見 Pattern Recognition for Tennis Tactics using Hidden Markov Model from Rally Series

宮原 捺希*¹ 手塚 太郎*² 中内 靖*³
Natsuki Miyahara Taro Tezuka Yasushi Nakauchi

*¹ 筑波大学図書館情報メディア研究科
Graduate School of Library Information and Media Studies, University of Tsukuba

*² 筑波大学図書館情報メディア系
Faculty of Library Information and Media Science, University of Tsukuba

*³ 筑波大学システム情報系
Faculty of Engineering, Information and Systems, University of Tsukuba

In this paper, we propose pattern recognition for tennis tactics using ball trajectory data from motion capture system. The purpose of the study is to adapt machine learning in order to implement feature extraction of rallies in tennis game using positions of ball bounce. We modeled this task as time-series data statistical modeling based on the Hidden Markov Model. We also conducted experiments and we verified the dispersion of the mixture component and the centroid, corresponding to four types of tennis court area division. Moreover, we implemented feature extraction of rally according to the initial state probability and the state transition probability.

1. はじめに

近年、スポーツにおいてデータ分析が普及しており、データを利用したコーチングや、フィジカル面や技術面にとどまらない戦略企画を行うことができるようになってきている。例えば、バレーボールの試合では、監督がタブレットを手に試合を分析したものをもとにアドバイスをしている姿がよく見受けられる[1]。データ分析が普及する以前のコーチングは口頭のみによるものが主流だったため、指導者の偏見や、選手と指導者間の価値観の違いにより、選手が疑問や反論を生むことも多かったと考えられる。しかしながら、客観的なデータに基づいてコーチングができるならば、双方にとって納得した上で、より効率的なコーチングが可能になるものと考えられる。

種々のスポーツがあるなかで、本研究ではテニスを対象に研究を行うこととする。既存のテニスの分析システムとして、Hawk-Eye[2]や PlaySight[3]などが挙げられる。それらのシステムは、ボールのバウンド位置の検出や、ショットの種類(サーブ、フォアハンド、バックハンド etc.)の識別およびラベリングを行うものであり、従来のテニスの分析において手作業で行われてきたものを自動化させ、効率化させる。現時点ではそれらのデータは容易に得られるものではないが、我々が日常で利用しているスマートフォンなどのカメラでスポーツの動作のラベリングを自動的に行う研究[4][5]もされており、将来スポーツのデータが身近に感じられるようになることが期待できる。そのようなときに得られたデータをどのように活用するかが今後の課題となってくる。本研究では、テニスのデータの利用方法を考え、テニスのモーションキャプチャシステム PlaySight のバウンド位置データに機械学習を適用し、選手のプレーのパターン抽出を行う。

2. 研究目的

2.1 テニスと機械学習

機械学習を適用したテニスの分析として、選手のランキングや試合データから試合の勝敗を予測する研究がある[6][7][8]。これらのデータ利用と分析方法は、試合の結果のみを予測するといった観戦目的での分析方法であり、試合の内容や選手のプレーの特徴といった観点では分析は行われていない。そのため、本研究では選手の戦術や判断に直接関わるバウンド位置を利用し、過去の試合データから得失点それぞれの傾向にあるパターンを分類することを最終的な目的とする。得失点パターンが抽出されることにより、選手の得意・不得意が把握できる。それによりその後の選手の競技力向上や苦手克服に繋げることを本研究の狙いとする。

2.2 バウンド位置を利用したデータ分析

図1は Hawk-Eye のデータ提示の一例である。これはポイントの一発目に打たれるサーブのバウンド位置を、4つのエリアごとの確率で提示されたものである。このデータは直観的に、コート奥側にいる選手がサーブを打つ際にどのエリアをどれくらいの確率で狙ってくるのかというデータを数値化して提示している。

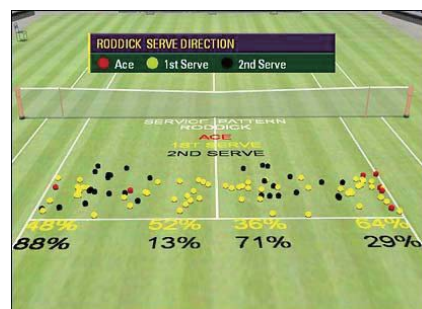


図1. Hawk-Eyeによるサーブのバウンド位置分析

このデータ提示方法は、サーブという一つのショットに対しての分析であり、その後複数のショットで構成されるラリーに対しての

分析は行われていない。既存のシステムにおいてもすべてショットにおけるバウンド位置が取得できるので、それらを利用しショットの組み立てを考慮した分析手法を提案することが本研究の目的である。

3. 隠れマルコフモデル(HMM)によるモデリング

前節で述べた通り本研究の目的となる、テニスのラリーによって繰り返される複数のバウンドを対象とする分析手法を考えた。本研究で扱うデータは、テニスのモーションキャプチャシステム PlaySight より取得している。PlaySight はテニスコートの周りに設置されたトラッキングカメラにより、選手とボールの位置座標を 50FPS で取得するものである。

本研究では時系列の統計モデリングを行う隠れマルコフモデル(HMM)を適用している。HMM は、出力ベクトル(観測変数)と出力ベクトルの確率分布によって推定された状態(予測変数)が、状態遷移確率をもって接続されたものとして定義される[9]。系列が出力される確率は、初期状態確率と状態遷移確率を掛け合わせることににより与えられる。HMM はマルコフ性を持ち、一つの状態は直前の状態にのみ依存する。HMM の学習は、複数の系列データによって行われる。

本研究で用いる学習データは試合中のラリーから観測される時系列のバウンド位置データである。例えば図2のようにすべてのバウンド位置(青色点)が観測されたとすると、HMM が表現するのは各状態(混合要素)に対応するガウス分布とその間の遷移確率である。初期状態確率と状態遷移確率を掛け合わせることににより、あるラリーの出力確率を数値的に示すことができる。学習された HMM において系列を表現するときは各状態の重心(橙色点)が用いられる。

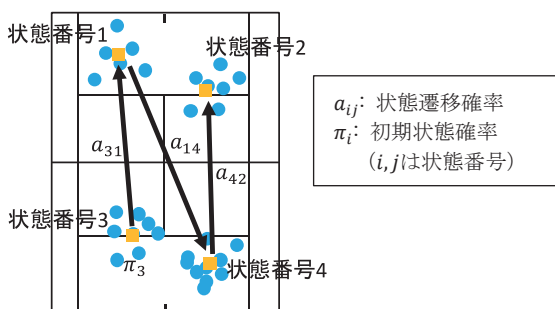


図2. HMM の構成

4. 実装

PlaySight より2人の男子選手の試合のデータを取得した。本研究では2試合分のラリーの系列データに HMM を適用した。ラリーのデータというのは、各ポイントにおいてサーブが有効であるときのバウンドの系列データを表す。つまり、ダブルフォルトとなったものは扱わない。今回扱ったデータは系列数(ラリー数)が214、合計バウンド数は681となった。本研究ではサーブのバウンドを除いたもので実験を行った。その理由としては、サーブは手前のサービスエリアという枠にバウンドさせなければならないという規則があり、そのために観測されるバウンドがサービスエリア付近に集中し、予測されるエリアの重心の位置が偏ってしまうという問題がある。そのため、HMM での初期状態はサーブの次のショットであるリターンのバウンドとなる。2試合においてすべてのバウンド位置が図4の青点のように観測されたものに対して、状態数を4, 6, 8, 12に指定しモデリングを行った。状態数をこの4つに指定した理由としては、テニスコートのエリアに

は左右を区別するフォアサイド・バックサイドの概念があり、上陣と下陣をそれらのエリアに分割するとコート全体の分割数が4(図3左の上陣)、さらにセンター(真ん中の概念)を考慮すると6(図3右の上陣)、そしてそれぞれに前後の概念を加えて分割すると8(図3左の下陣)、12(図3右の下陣)となるからである。図4のように分割された箇所に HMM の状態と重心が分散される形が望ましい。

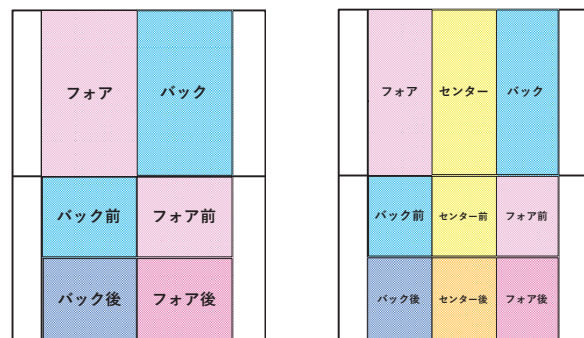


図3. テニスコートエリア分けの例

5. 結果と考察

図4に状態数4, 6, 8, 12を設定してモデリングした結果を示す。観測バウンド位置を青色点で示し、予測されたエリアの重心を橙色点で示した。初期状態であるリターンのバウンドは全て下陣で観測されるようになっている。

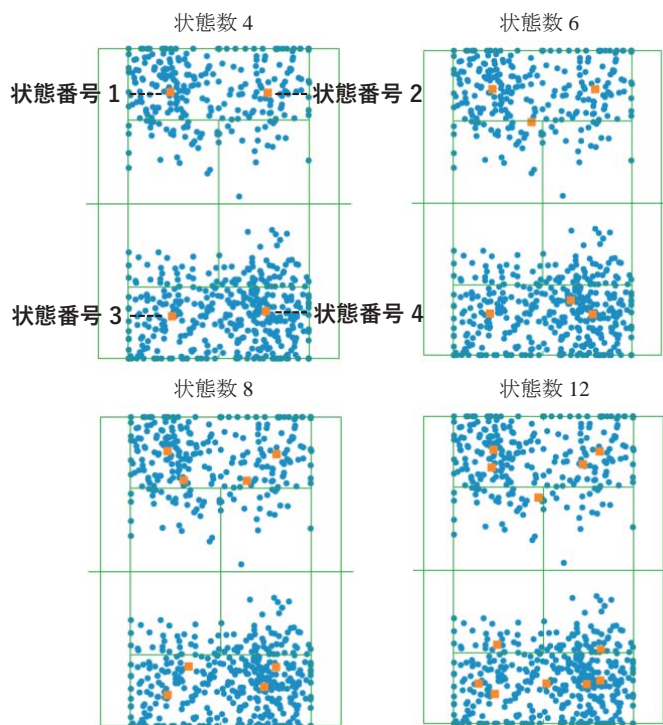


図4. 観測データと予測変数(状態と重心)

状態数4では両陣ともフォアサイド・バックサイドに重心が分散されている。状態数6では、上陣でセンターを含めた3つのエリアに分散されているように見えるが、下陣では重心がフォアサイド側に2つ密集している。選手によっては、センターを狙うという意識はあまり持たずに、単純に左右に打ち分けることを意識するというプレースタイルも多い。そのような状況であるとセンター

にバウンドが集まらず、左右のみにバウンドが集中することも考えられる。

前後のエリアを区別する状態が学習されることを期待して実行した状態数 8 と状態数 12 では、前後に重心が分散されず、陣の後方に重心が偏ってしまっている。考えられる理由としては、テニスの試合において、打球はよりネットより深い深めのエリアを狙ったほうが有利であるため、それを意識したプレースタイルであるほど、陣の後方にバウンドが密集し、観測されたデータからは前後のエリアが区別されにくくなることが考えられる。

状態数 4 の HMM の初期状態確率 π_i を図 5 に示す。この HMM における初期状態はリターンのバウンドであり、コートの下陣にバウンドするよう設定されている。つまりこの状況における初期状態は重心 3 と 4 (図 4 左上)となる。状態 3 では初期状態確率が 36.6%、状態 4 では 62.0%となっており、この試合においてリターンはフォアサイドを狙われる傾向が強いことがいえる。

図 6 に状態遷移確率 a_{ij} を示す。左側の表はテニスコートの上陣にある状態 1, 2 から 3, 4 への遷移確率、右側の表は下陣にある状態 3, 4 から 1, 2 への遷移確率を示したものである。

i	3	4
π_i	0.3657	0.6203

図 5. 初期状態確率 π_i (i は状態番号)

$i \setminus j$	3	4
1	0.3443	0.6557
2	0.5494	0.4501

$i \setminus j$	1	2
3	0.5137	0.4272
4	0.6263	0.3522

図 6. 状態遷移確率 a_{ij} (i, j は状態番号)

上陣のフォア側である状態 1 に関しては状態 4、つまりクロス方向である下陣のフォア側への遷移確率が 65.6%となっている。そして下陣のフォア側である状態 4 に関しても、状態 1 つまりクロス方向である上陣のフォア側への遷移確率が 62.6%となっている。この結果から、選手は自分のフォアに来たら相手のフォアに返すというフォアのクロスラリーを行う傾向が強いことが読み取れる。そしてほぼ同じ確率でストレート方向に展開する傾向にあることが読み取れる。

状態 2 と状態 3 は上陣と下陣それぞれのバック側を表しているが、左右の打ち分けの確率にあまり偏りが見られない。そのため、選手は自分のバック側に球が打たれると、相手の左右に同じくらいの割合で打ち分けをし、バックでクロスラリーをする傾向は弱いとも言える。両選手の違いを比較するとすれば、上陣からはバック側からクロス方向に遷移する確率のほうが高い一方で、下陣からはバック側からストレート方向に遷移する確率が高いことが読み取れる。

6. まとめと今後の課題

本研究では、テニスのモーションキャプチャシステムから取得されるボールのバウンド位置データに、隠れマルコフモデルを適用し、テニスのラリーのパターン分析を行った。

推定するパラメータの一つである状態がどのように分散するかを考察した。テニスコートを分割する上で基本の概念となるフォア・センター・バックと前後を加味したエリアが区別されることを期待し、それに対応した状態数を指定し出力結果の比較を行った。状態数 4 の出力は左右に重心が分散されたが、センターと前後を含めたエリア分けに対応する状態数の出力は、期待通りに分散されなかった。これは観測データの元になる選手のプレ

ースタイルに依存することが考えられ、コートを広く利用した配球をするプレーであれば分散されたエリアが検出されることも考えられる。しかし、どれくらい広くテニスコートを使って配球をしているかの指標として扱える見込みがある。今後は図 3 に示したエリア分けのパターンと、出力された重心を用いて、選手がどのような配球を行う傾向があるのかを検証する手法を考えることが課題である。

また本研究でモデリングした HMM ではサーブおよびポイントの終わり(アウト・ネット・エース)をパラメータとして扱っていない。今回扱ったラリーのバウンドと連動させることで、ポイントの序盤、中盤のラリー、終盤のすべてを加味したパターンを分析し、提案システムのさらなる改善を図っていく予定である。

謝辞

本研究は、JSPS 科研費 JP16K00228, JP16H02904 の助成を受けたものである。

参考文献

- [1] 鈴木淳平: バレーボールのゲーム分析—世界一決定戦での得失点の経緯とゲーム展開の研究, 駒澤大学総合教育研究部紀要(4), 2010.
- [2] Hawk-Eye., <https://www.hawkeyeinnovations.com> (2018-02-28 参照)
- [3] PlaySight., <https://www.playsight.com> (2018-02-29 参照)
- [4] Draschkowitz L., Draschkowitz C., Hlavacs H.: Predicting Shot Success for Table Tennis Using Video Analysis and Machine Learning, INTETAIN, Social Informatics and Telecommunications Engineering, 2014.
- [5] Chen, HT., Tsai, WJ., Lee, SY., Yu, JY.: Ball tracking and 3d trajectory approximation with applications to tactics analysis from single-camera volleyball sequences. Multimedia Tools and Applications 60(3), 2012.
- [6] Sipko, M.: Machine Learning for the Prediction of Professional Tennis Matches, Imperial College London, 2015.
- [7] Ma, S., Liu, C., and Tan, Y.: Winning matches in Grand Slam men's singles: an analysis of player performance-related variables from 1991 to 2008. Journal of sports sciences, 31(11), 2013.
- [8] O'Malley, JA.: Probability Formulas and Statistical Analysis in Tennis. Journal of Quantitative Analysis in Sports, 4(2), 2008.
- [9] 徳田恵一: 隠れマルコフモデルによる 音声認識と音声合成, IPSJ Magazine, 情報処理学会, 2004.