

画像計測技術を用いた卓球パフォーマンス分析の自動化

玉城 将

目 次

要約	1 -
1. 緒言	2 -
2. ボールの 3 次元軌跡計測法	4 -
2.1 ボール検出	4 -
2.2 時間的および幾何学的カメラ校正	5 -
2.2.1 ボールの抽出	5 -
2.2.2 ボールの 2 次元軌跡の復元	5 -
2.2.3 時刻ずれおよびカメラ外部パラメータの推定	5 -
2.3 ボールの 3 次元軌跡復元	6 -
2.3.1 ボールの抽出	7 -
2.3.2 衝突時刻の推定	8 -
2.3.3 軌跡の復元	8 -
3. 実験	9 -
3.1 実験条件	9 -
3.2 実験結果	11 -
3.2.1 カメラ校正前のボール検出	11 -
3.2.2 時間的および幾何学的カメラ校正	12 -
3.2.3 カメラ校正後のボール検出	12 -
3.2.4 ボールの 3 次元軌跡の復元	13 -
3.2.5 ボールの 3 次元軌跡からの情報推定	13 -
4. 実用に向けたシステム設計	14 -
5. まとめ	15 -
参考文献	15 -

画像計測技術を用いた卓球パフォーマンス分析の自動化

玉城 将 斎藤英雄 吉田和人 山田耕司 池袋晴彦

要約

近年、プレー中の膨大な情報を用いた選手のパフォーマンスに関する統計的な分析、いわゆるパフォーマンス分析がスポーツ競技現場で利用され始めている。この流れは、コンピュータソフトウェアによって簡易に情報を記録および分析できるようになったことに起因しており、今後も情報処理技術の発展に伴ってその重要性は高まると考えられる。これまでに著者らは Graphic User Interface ソフトウェアを用いた卓球のパフォーマンス分析法を提案している。2012 年に開催されたロンドン五輪における日本代表チームの支援にこのソフトウェアが用いられるなど、パフォーマンス分析の利用には一定の進展がみられる。その一方で、従来の手入力に依存したシステムでは情報入力に要する膨大な人的コストが問題となり、多くの試合を分析対象とすることは難しい。この問題の解決には、画像計測技術を用いた情報収集の自動化が不可欠である。

そこで本研究では、卓球パフォーマンス分析自動化を目的とした卓球ボールの 3 次元軌跡復元法を提案する。競技会で行われる試合であっても対応できるよう、本研究では 3 次元位置が既知の点群の設置や、ケーブル等を用いたカメラ間の時間同期は行わない等の制約を設ける。その制約下で、精度良く且つ頑健に卓球ボールの 3 次元位置を計測する方法を明らかとし、さらにボールの 3 次元軌跡復元を自動化するシステム開発が本研究の課題となる。

本研究ではまず、事前にカメラの位置関係を求める必要がなく、且つカメラを同期させる必要が無いボールの 3 次元軌跡計測法を明らかにした。本手法は(1)ボール候補領域の検出、(2)カメラ校正、(3)ボール軌跡復元の 3 つの処理から構成される。(1)ボール候補領域の検出:画像上のボールの形状と色に基づいて 1 画像から複数のボール候補領域を検出する。(2)カメラ校正:次に複数フレーム間のボールの動きに基づき、ボール候補から真のボールを抽出し、抽出されたボールの位置情報に基づいてカメラ間の時刻ずれおよび位置関係を算出する。(3)ボール軌跡復元:推定されたカメラ間の時刻ずれおよび位置関係を満たし、且つ 3 次元空間中でボールらしい動きをするボール候補のみ抽出する。抽出されたボール候補から衝突時点で分割した複数の 3 次元軌跡を復元し、最後に軌跡を接続することで 1 ラリーの軌跡が復元される。1 ラリーの軌跡を 1 つのスプライン曲線で復元すると、ボールと他の物体との衝突が含まれる場合に精度よく復元できない問題が生じるため、複数の軌跡を接続することで 1 ラリーの軌跡を表現する。実際の卓球選手の試合を対象として本手法の計測精度を評価した。試合開始後の 10 ラリーを用いてカメラ校正を行った結果、カメラに設定した真の時刻ずれが -3 ms であったのに対し、推定された時刻ずれは -3.04 ms であり、本手法の推定精度が非常に高いことが示された。また、カメラ間の光軸に関する誤差が 1.07 度、距離の誤差は 0.03 m であり、カメラ間の幾何学的な位置関係も高い精度で求められた。このように高精度でカメラ校正が可能であるため、再投影誤差は 1.31 pixel、実長換算で約 3.2 mm 程度であり、卓球選手の戦術的なパフォーマンスに関する情報収集に十分実用可能な計測精度と考えられた。以上のように、本研究によって非同期カメラを用いた場合のボール軌跡の復元に必要な全ての工程を含む包括的な枠組みが提案された。

次に、ボール軌跡の復元を簡易に行うシステムを設計した。ボール軌跡の復元は事前に撮影したプレーを対象として行うケースと、目の前で行われているプレーを対象に行うケースが存在し、その両方のワークフローに対応する必要がある。特に後者に対応するには、計算コストの高いボール候補の検出を複数台の PC で分散処理する必要がある。そこで、ボール候補検出、ボールを用いたカメラ校正、ボールの軌跡復元をそれぞれ独立したプログラムとし、ボール候補検出によって得られた情報はデータベースシステムに格納されて処理終了を待たずとも他のプログラムからアクセスできるよう設計した。このように 3 つのプログラムに分割することは、システムのメンテナンス性を高めることにも有効であると考えられる。

以上のように、本研究では卓球の試合を対象としてパフォーマンス分析を行う上で重要となるボール軌跡の復元を簡易に行うためのアルゴリズム、および実用システムの開発を進めた。今後、本研究成果に基づき多くの試合を対象として卓球のパフォーマンス分析が実施され、卓球一流選手の試合の時空間的な理解が促進されることが期待される。

代表者所属:慶應義塾大学大学院

1. 緒言

近年、スポーツ試合中の客観的情報に基づいた試合の分析、いわゆるパフォーマンス分析が盛んに行われている。パフォーマンス分析は膨大なデータに基づく統計値を用いることを特徴としており、客観的証拠によってコーチらは主観的な分析結果を最適化できる点が、従来、選手の主観あるいはコーチや同じチームの選手ら第三者の観察のみによって行われてきた分析方法とは異なる。この流れは、コンピュータソフトウェア開発が身近になり、研究者自身でソフトウェア開発を行うケースが増えたこと (Lees, 2003; Nevil et al., 2008)、さらにはパフォーマンス分析に必要な機能を搭載したコンピュータ・ソフトウェアが市販されたことが関係している (玉城・宮地, 2012)。例えば、O'Donoghue and Ingram (2001) は自ら開発したコンピュータソフトウェアを用いてテニスの一流選手による 37,229 ラリーを対象にデータの入力および分析を行うことで、性別およびコートサーフェイスがテニス選手の戦術に与える影響を定量的に示している。現在、パフォーマンス分析は様々な種目を対象として行われ、スポーツ科学における重要な課題の一つとなっている。その一連の取り組みについては Nevil et al. (2008) が詳しくまとめている。

卓球のパフォーマンス分析に関する研究は、例えば Lanzoni et al. (2011) による評価指標の分類が挙げられる。Lanzoni et al. は卓球を対象とした先行研究を包括的に調査することにより、卓球のパフォーマンス分析における評価指標を選手の属性、技術、戦術、用具、プレー条件、心理、およびルールに分類している。また、Lanzoni et al. (2013) はその分類に基づいてアジア選手とヨーロッパ選手のパフォーマンスの違いを分析することで、アジア選手はヨーロッパ選手と比較して攻撃的な打法を多く用いること、両者の打球コースに違いがあること等を指摘している。Zhang et al. (2013) は打球番号注 1) に着目してラリーを 3 つの局面に分類し、各局面の選手のパフォーマンスを評価している。このパフォーマンス分析法は、世界トップの競技力を誇る中国ナショナルチームで用いられている。著者らは Graphic User Interface を備えたソフトウェアを開発し、卓球の試合場面における即時的なパフォーマンス分析を実現した。2012 年に開催されたロンドン五輪における日本代表チームの支援にこのソフトウェアが用いられた。

このように、パフォーマンス分析の利用には一定の進展がみられる。しかし、これまでに報告されたパフォーマンス分析法は全て、情報収集に膨大な時間を要するという問題がある、分析時間が膨大である場合、パフォーマンス分析を実施することで競技力向上の時間効率を下げる可能性がある。パフォーマンス分析の研究においてこの点が検討されることはほとんど無いが、競技力強化を目的とした現場において、それをいかに効率良く行うかは重要な課題であり、分析に要する時間はパフォーマンス分析の利用を決定付ける重要な要素の一つであると考えられる。記録に要する時間を少なくする一つの方法は、記録する情報を少なくすることである。しかし、得られる分析結果は抽象的となり、具体的にどの点に選手の課題があるかを検討することが難しくなる。Lanzoni et al. (2011) は卓球の技術に関する評価指標として打法、フォアハンドかバックハンドか、フットワークの種類、打球位置、選手の移動方向、戦術に関する評価指標として、得点打・失点打・中立打 (得点打でも失点打でも無い打球) の種類、打法の組み合わせ、打球コース、試合時間を挙げている。また、吉田 (2007) は卓球のラリーにおける攻防を時間、空間および回転の 3 つの要素に分けられると指摘しており、Lanzoni et al. の挙げた指標に加えて球速、打球間隔、ボールの回転も卓球における戦術および技術を評価する重要な指標であると考えられる。しかし、いずれの情報も手動で短時間且つ正確に記録するのは困難である。

仮にボールの 3 次元軌跡が画像から自動的に復元することができれば、ボールの回転を除く多くの情報が自動的に得られ、この問題の解決に繋がる。しかし、撮影環境やコストの制約上、ボール軌跡の計測はどのようなケースでも行えるものではない。ボールを対象とした 3 次元計測では、露光を同期させた複数のカメラを用いてボールを撮影し、三角測量によって撮影されたボールの 3 次元位置を復元し、最後に復元されたボール位置の軌跡を補間する方法が一般的である。しかし、この方法では外部同期に対応するカメラおよび同期信号発生器といった特殊な撮影機器が必要となり、さらにそれらをケーブル等で接続する必要がある。また、計測される 3 次元位置の精度を高めるためには対象物体が移動する領域に 3 次元位置が既知のマーカーを置くなどして事前にカメラ間の幾何学的位置関係を算出すること (以下「幾何学的カメラ校正」という) が必要である。スポーツの競技会や練習場では、撮影環境やコストの制約上このような装置を用いることができない場面は多い。

安徳ら (2013) は、カメラ 1 台のみを用いてテニスボールの画像上の軌跡と選手の画像上の位置を計測し、選手がフォアハンドで打球したか、バックハンドで打球したかを記録している。北原と内田 (2013) は露光を同期していない 2 台のカメラを用いてテニスボールのバウンド位置や選手の位置をコートの 2 次元座標で

計測している。カメラは複数台用いているが、競技現場における利用の簡易さを考慮してカメラの露光は同期させていない。これらの研究では、システム利用の手間を少なくするために 2 次元情報のみを計測しており、ボールの 3 次元位置は計測していない。しかし、カメラの光軸がコート平面と平行に近くなるほど計測される位置の精度が低下する点は原理上避けられない。また、ボールに回転が加わることによる曲線的な軌道もスポーツにおいて戦術的あるいは技術的に重要な要素であるが、2 次元の情報からは解析できない。プレー中の情報を記録する技術として 3 次元的な軌跡を計測できるものが求められる。

高野橋ら(2007)は、カメラの露光時間を長くすることで得られるボールの残像、すなわち画像上の短いボール軌跡を 3 次元空間に逆投影し、それらの共通部分を 3 次元形状として復元している。カメラ間の露光が非同期の場合であっても、画像上の軌跡が得られれば 3 次元空間中の交線を求めることで 3 次元軌跡が得られることを示した。しかし、連続的な軌跡を復元することには成功しているが、薄く撮像されたボールの残像を頑健に検出することは非常に難しい問題であり、有効な方法は提案されていない。実際の試合映像を対象としてこの方法を用いることは難しい。

清水ら(2009)は、エピポーラ線と画像上の軌跡の交点を求めることによってボールの 3 次元位置を計測している。非同期カメラ間の時刻ずれが未知の状態、さらにカメラがパン・チルトする場合であっても 3 次元軌跡を復元できることを示している。しかし、軌跡とエピポーラ線が平行に近くなる場合には計測精度が低下する問題がある。卓球やテニスのようなネット型スポーツを撮影する場合には、選手とボールとの重なりを少なくするためにカメラはコートのサイドライン側に水平に並べられることが多いと考えられるが、その場合にはボール軌跡とエピポーラ線は頻繁に平行になる。計測精度の観点からはエピポーラ線と軌跡の交点を用いる方法は適切ではないと考えられる。

Matsumoto et al.(2010)は、複数の非同期アフィンカメラによって撮影された移動体の軌跡が周波数領域においては対応関係を保持する点に着目し、周波数領域において 3 次元軌跡を復元することで、非同期カメラから移動体の 3 次元軌跡を復元している。しかし、対象物とカメラの距離が短く、アフィンカメラで近似することが適切ではない場面は多い。各務ら(2012)は非同期射影カメラであっても光軸を平行にしたステレオカメラであればアフィンカメラと同様の理論を適用できること、さらには同期カメラと比較して高密度な 3 次元形状を復元できることを示している。しかし、2 台のカメラを用いて移動体の 3 次元位置を計測する場合にはカメラの光軸を直角に配置した時に最も精度が高くなるため、光軸を平行に複数のカメラを配置するという制約は計測精度の観点からは望ましくない。

Zhou and Tao(2003)は、各カメラの外部パラメータが既知の状態では移動体を撮影し、各視点で撮影された移動体の移動距離の比率から時刻ずれを推定し、推定された時刻ずれおよびオプティカルフローを用いて実際には撮影されていない時刻の画像を生成している。ボールを検出することができれば、この方法を用いてカメラ間の時刻ずれを推定すること(以下「時間的カメラ校正」という)ができる可能性がある。しかし、ボールスポーツを対象とする場合、フレーム間の移動体の動きを等速直線運動に近似することは、実際の移動体の動きとは異なる可能性がある。また、この方法で扱うのは時間的カメラ校正のみであり、幾何学的なカメラ校正は別の方法で実施する必要がある。

野口と加藤(2006)は、2 台の非同期カメラで撮影したマーカーの軌跡を線形に復元し、カメラ間の時刻ずれの初期値からカメラ間の対応点を大マーカーに推定し、推定された対応点によって計算された基礎行列の誤差に基づき時刻ずれを最適化することで時間的カメラ校正と幾何学的カメラ校正を同時に行う方法を提案している。また、Ying and Sato(2007)は、野口と加藤の方法が異なるフレームレートのカメラを組み合わせる場合にも適用可能であることを示している。この方法では、時間的カメラ校正と幾何学的カメラ校正を同時に扱う行うため、マーカーをボールに置き換えることでボールの 3 次元計測に応用できる可能性がある。ただし、フレーム間の移動体の動きを線形近似していることから、フレームレートに対して移動体の動きが十分に遅くなければならないという制約がある。この近似はスポーツプレー中のボールのように高速で曲線的に移動する物体を扱う場合には妥当ではないと考えられる。また、マーカーを用いていることから、ボールを対象とした場合に生じる誤検出等の問題には対応していない。ボールを対象とした場合、これらの点が問題になる可能性がある。

このように、非同期カメラによって簡易に 3 次元計測する方法を提案する研究はいくつかみられるが、従来手法をそのまま用いるだけではボールの 3 次元軌跡を復元することは難しい。また、実際にボールの 3 次元軌跡を復元しようとする場合にはカメラ校正だけではなく、ボールの検出から軌跡復元方までを含む包括的手法を検討する必要がある。

そこで本研究では、非同期カメラを用いた新たなボールの 3 次元計測法を提案する。まず、野口と加藤の提案する非同期カメラの時間的および幾何学的カメラ校正法に修正を加えることで、ボールを用いた高精度な時間的および幾何学的カメラ校正が可能であることを示す。次に、非同期カメラを用いた場合、ボールの検出は各カメラで独立して行われることが多いが、提案手法ではカメラ間の時刻ずれおよび位置関係を高い精度で求められるため、この結果を用いることで頑健にボールを検出する方法を示す。以上のように、本論文における提案手法は、非同期カメラを用いた場合のボール軌跡の復元に必要な全ての工程を含む包括的な枠組みである。提案手法によって時間同期されたカメラが不要になり、且つ計測対象のボールを用いたカメラ校正が可能となることでボウリングスポーツにおける 3 次元計測が簡易化されることが期待できる。本論文では特に卓球での計測に焦点を当てているが、テニスやバレーボール等、プレー中の多くの時間でボールが空中を移動しているスポーツであれば提案手法は適用可能である。

2. ボールの 3 次元軌跡計測法

本章では画像からボールを検出し、ボールの 3 次元軌跡を復元するまでの手法の詳細を説明する。本手法によって復元される軌跡は 1 ラリーのものであり、ラリーの開始および終了時刻が既知であることを前提とする。まず、画像から形状および色に基づいてボール候補を検出する。もし時間的および幾何学的にカメラが未校正である場合、画像上の動きに基づいてボールを抽出し、抽出されたボールの位置情報を用いたカメラ校正を行う。カメラ校正後、ボール候補から 3 次元空間中でボールらしい動きをするものを抽出し、ボールの 3 次元軌跡を復元する。カメラ内部パラメータ、レンズ歪み係数は OpenCV ライブラリや Matlab に実装されている Zhang (2000) の方法を用いる等によって簡易に推定できるため、本手法ではこれらは入力値として扱う。また、時刻ずれの初期値も入力値となっているが、これは撮影された映像をコマ送りするなどして推定できる数フレームの誤差を含む時刻ずれである。コート四隅の画像上の座標は手動で記録することとし、鉛直方向の算出、画素と実長の比率の算出、バウンド検出、および打球コースの記録に用いる。以降、各処理の詳細を説明する。

2.1 ボール検出

本手法では空中を移動するボールを検出対象とするため、まずはフレーム間差分によって移動体を検出する。フレーム間差分とは、時刻 t における画像と時刻 $t - s$, $t + s$ の画像との差分を求め、その論理積をとることで時刻 t における移動体を検出する方法である。図 1 に 60 fps で撮影された卓球の打球シーンを差分間隔 3 でフレーム間差分した場合に検出される移動体の例を示した。

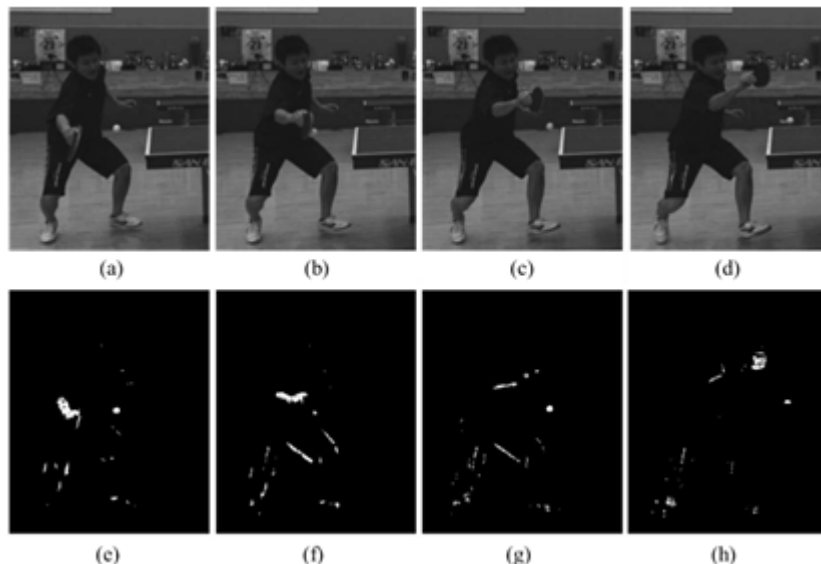


図 1. (a)-(d)は 60fps で撮影された映像の各フレーム、(e)-(h)はフレーム間差分の結果

多くの場合、ボールは他の物体と比較して速く移動しているため、差分間隔を短く設定することで、検出される他の移動体を少なくできる。次に、検出された移動体から色、円形度および面積がボールと類似している領域を抽出する。この処理で必要となる閾値は以下の 3 つである。

- ・ 輝度の下限 [0-1.0]: L_I
- ・ 円形度の下限 [0-1.0]: L_C
- ・ 面積の下限 [pixel^2]: L_A
- ・ 面積の上限 [pixel^2]: U_A

なお、領域の周囲長が l 、面積が S である領域の円形度 C は以下の式で定義される。

$$C = 4\pi S / l^2 \quad (5)$$

ここまでの方法によって検出される物体がボール候補である。1 フレームの画像のみからボールを検出するため、ボールだけを正確に検出することは非常に難しく、多くの誤検出が含まれることが想定される。そのため、次にボール候補から正しいボールを抽出する必要があるが、時間的および幾何学的カメラ校正の前後で異なる方法を用いるため、次節以降で詳しく説明する。

2.2 時間的および幾何学的カメラ校正

2.2.1 ボールの抽出

ボール候補から、複数フレームに渡って空中を移動するボールと類似した動きをするものを検出する。ボールの動きに関する条件は、移動速度、隣合う移動速度の差、そして隣り合う移動方向間の角度によって定める。すなわち、この処理で必要となる閾値は以下の 4 つである。

- ・ 移動速度の上限 [m/s]: U_V
- ・ 移動速度の下限 [m/s]: L_V
- ・ 連続した移動距離の変化率の上限 [%]: U_D
- ・ 連続した 2 つの移動ベクトル間の上限 [rad]: U_R

ただし、 U_V と L_V は卓球台のコーナーから求められた画素と実長の比率を用いて、それぞれ画像上の速度 U_V と L_V に変換して用いる。ここで、 i 番目のフレームにおけるボール候補の位置を $\mathbf{p}_i$ と表記し、 $\mathbf{p}_i$ から $\mathbf{p}_{i+1}$ まで線形に移動したと仮定した場合の速度ベクトルを $\mathbf{v}_{i,i+1}$ と表記する。 $\mathbf{v}_{i,i+1}$ が U_V 以下且つ L_V 以上である場合、 $\mathbf{p}_i$ と $\mathbf{p}_{i+1}$ は異なるフレームにおける同じボールである可能性があると考え。続いて $\mathbf{p}_{i+2}$ は、 $\mathbf{v}_{i+1,i+2}$ が U_V 以下且つ L_V 以上であることに加え、 $\mathbf{v}_{i,i+1}$ に対する $\mathbf{v}_{i+1,i+2}$ の変化率が U_D 以下、且つ成す角が U_R 以下であるという条件も併せて満たした場合、 $\mathbf{p}_i$ 、 $\mathbf{p}_{i+1}$ と同じボールである可能性があると考え。以降、 $\mathbf{p}_{i+4}$ まで連続して 5 フレーム前述の条件を満たすボール候補を最終的に正しいボールであると判断する。

このようにボール候補に対して強い制約を課したとしても、ボールに類似した領域がボールらしい動きをした場合には誤検出が生じる。このような誤検出の混入は本手法におけるボール抽出の限界であると同時に、ボールの検出を行う他の多くの方法に当てはまる限界でもある (Huang et al., 2008; 石井ら, 2007; Yu et al., 2004; Chen and Zhang, 2006)。現状では画像から正しいボールだけを検出することは困難であり、以降の処理では少数の誤検出が含まれることを仮定する。

2.2.2 ボールの 2 次元軌跡の復元

ここまでに得られたボール位置から、ボールの画像上の軌跡を復元する。精度が低下することを避けるため、ボール位置を連続して記録できた区間のみ軌跡を推定する。軌跡の復元は、3 次スプライン補間を用いる。スプライン補間では、ボールバウンドのような急激な運動の変化を含む軌跡は正確に復元できない。しかし、以降の処理では誤検出への対応を行うため、バウンド前後の軌跡が正しく復元されない場合でもカメラ校正の結果にほとんど影響は与えない。

2.2.3 時刻ずれおよびカメラ外部パラメータの推定

前述したように、時刻ずれの初期値 τ_0 は既に得られている。ここで、カメラ i における j 番目のフレームの時刻を t_j^i と表記する。 τ_0 が正しい場合にはカメラ 1 における時刻 t_j^1 と、カメラ 2 における時刻 $t_j^2 - \tau_0$ は同じ時刻となる。そして、カメラ 2 において画像上のボール軌跡が得られている期間であれば、 τ_0 に基づき、カメラ 1 で検出されたボールがカメラ 2 のどこに撮像されるかを推定できる。同様に、カメラ 1 において画像上のボール軌跡が得られている期間であれば、カメラ 2 で検出されたボールがカメラ 1 のどこに撮像されるかを推定できる。このように画像上のボール軌跡と時刻ずれによって、画像間に多数の対応点が得られるため、基礎行列 $\mathbf{F}_0$ が計算できる。基礎行列はカメラ間の位置関係とそれぞれのカメラ内部パラメータで定まる 3×3 の行列であり、カメラ内部パラメータが既知である場合には基礎行列から一方のカメラに対するもう一方のカメラの位置姿勢が求められる。ただし、 τ_0 が数フレームの誤差を含むものであるため、 $\mathbf{F}_0$ は正確ではない。

ここで、時刻ずれと画像上の軌跡によって算出した p_i^1 のカメラ 2 における対応点を p_i^2 、 p の斉次座標を $\tilde{p}$ 、点と直線の距離を算出する関数を d 、 N 個の数値の中央値を算出する関数を Med_i^N と表記すると、基礎行列 F の誤差関数 E は次式のように定義できる。

$$E(F) = \text{Med}_i^N \left(d \left(p_j^1, F p_j^{2'} \right) + d \left(p_j^{2'}, F^T p_j^1 \right) \right). \quad (6)$$

$F p_j^{2'}$ は $p_j^{2'}$ と対応するカメラ1におけるエピポーラ線、 $F^T p_j^1$ は p_j^1 と対応するカメラ2におけるエピポーラ線である。数式(6)はエピポーラ線と対応する点との距離の中央値であり、基礎行列が正しい値に近づくほど小さくなる。本手法では、時刻ずれに含まれると想定される誤差の範囲内に渡って固定ステップで基礎行列の誤差を求め、最小の値を返した時刻ずれを最適な時刻ずれとする。カメラの時刻ずれの値の変化によって基礎行列の誤差がどのように推移するかを図 2 に示した。

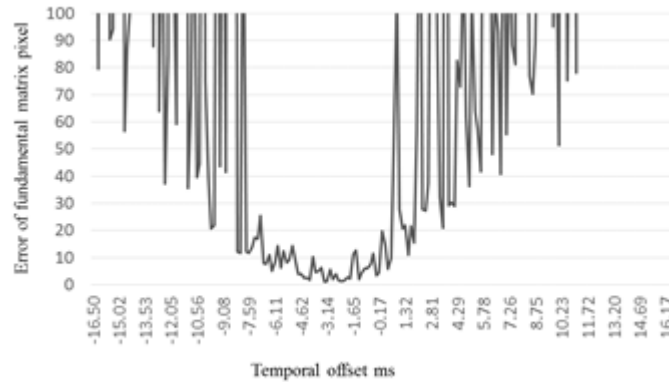


図 2. 基礎行列の誤差. 真値は-3.0 ms.

カメラに設定した時刻ずれの真値-3.0 ms に近いほど基礎行列の誤差が小さくなることが分かる。しかし、基礎行列算出の際の不安定さから基礎行列の誤差は激しく値が上下している。局所解が非常に多く、勾配等を用いた可変ステップの最小値探索法では大域的な最小値に至らない場合が多い。実用上は局所解を避けることが非常に重要となるため、本手法では局所解に陥ることを避けるため、設定した時間幅の中で必要な精度で時刻ずれを推定できる方法として固定ステップで最小値を探索することとした。

最後に、推定された最適な時刻ずれによって推定される対応点から基礎行列を求める。基礎行列は RANSAC によって外れ値を除外した全ての対応点からランダムに 8 点を選択して算出する。ただし、プレー中のボールを用いる場合には対応点が様に分布しないため、抽出される 8 点の組み合わせによっては正しく基礎行列が算出されない場合がある。この不安定さに対応するため、対応点の抽出と基礎行列の算出を N_i 回繰り返す。算出された全ての基礎行列からカメラ外部パラメータおよび再投影誤差を求め、再投影誤差が最小のものを選択する。

提案手法における本節で説明した部分は野口と加藤(2006)の方法に基づいているが、卓球ボールを対象とできるよう修正したため、幾つかの相違点がある。まず、画像上の軌跡を復元する方法が異なる。野口と加藤はフレームレートに対してゆっくり動かすマーカーを対象としているため、画像上の軌跡は線形補間を用いているが、本手法で対象とする卓球ボールは高速且つ曲線的に移動するため、スプライン補間を用いている。また、基礎行列の誤差の評価方法も異なる。野口と加藤は検出が容易なマーカーを対象としているため、基礎行列の誤差はエピポーラ線と対応点の距離の平均値を用いている。しかし、本手法で対象とする卓球ボールは、様々な理由により誤検出が含まれる可能性があり、また、バウンドを含む区間の軌跡が正確に復元できないため、平均では外れ値の影響を大きく請けすぎる可能性がある。そのため、本手法では外れ値が含まれていたとしても誤差分布の代表値を正確に反映できる中央値を用いており、数式(6)は誤検出の有無に関わらず正しい基礎行列を与えた際に最小値を返すことが期待できる。

2.3 ボールの 3 次元軌跡復元

2.3.1 ボールの抽出

基礎行列 $\mathbf{F}$ を用いて計算できるエピポーラ線を用いてもう一方の画像におけるボール候補から正しいボールを抽出する(図 3). 必要となる閾値は, 以下の 2 つである.

- ・ エピポーラ線とボールの距離の上限 [m]: U_E
- ・ 軌跡を支持すると考えるボール候補の距離の上限 [m]: U_S

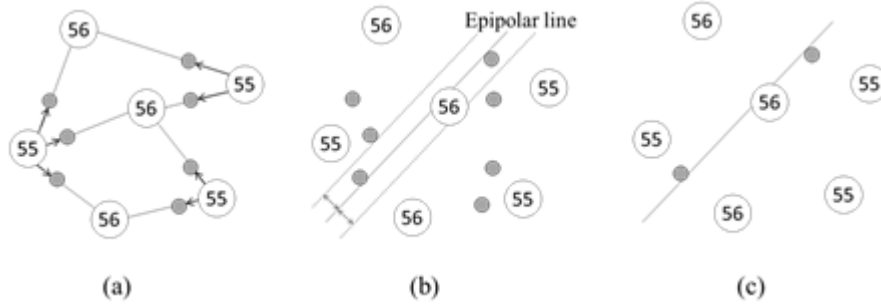


図 3. エピポーラ線に基づくボール抽出. 白い円がボール候補, その番号がボール候補が検出されたフレーム番号を示す. 灰色の円は前後フレームのボール候補から推定されたボール候補を示す. (a) 距離が閾値以下の 2 点の間を線形に移動したと近似して $55+\tau$ (時刻ずれ) 時点のボール位置を推定. (b) もう一方の画像で検出されたボール位置からエピポーラ線を求め, (c) エピポーラ線に近いボール候補のみ抽出.

ただし, U_E は画像上の距離 U_e に変換して用いることとする. まず, 時刻 $t_j^1 + \tau$ の直前および直後にカメラ 1 において検出されたボール候補を対象とし, 2 点間の距離がボールがフレーム間で移動する距離の最大値, すなわち U_v とフレーム間時刻との積を下回る組み合わせを抽出する(図 3(a)). 抽出された組み合わせの間を等速で線形に移動していると近似することで $p_{t_j^1 + \tau}^1$ の候補を推定する(図 3(a) の円). 次に, カメラ 2 の時刻 t_j^2 におけるボール候補位置 $p_{t_j^2}^2$ に対応するエピポーラ線 $\mathbf{F}p_{t_j^2}^2$ を求める. エピポーラ線との距離が U_e 以内の点を探索し(図 3(b)), 探索されたボール候補, ボール候補を推定する際に用いられた隣り合う 2 点, エピポーラ線と対応するもう一方の画像上のボール候補を対応付けて抽出する. これによって, 多くの誤検出が候補から除外される(図 3(c)). なお, 後述するように, ここで線形近似によって推定されたボール位置は正しいボールを抽出する目的のみに用いられるものであり, 最終的なボールの測定精度には影響を与えない.

次に, エピポーラ線によるボール抽出によって得られた全ての対応点から, 3 次元位置を計算する. 誤った組み合わせが存在する場合, 3 次元空間であれば正しいボール位置から大きく離れると考えられる. そこで, Yan et al. (2006) の方法を 3 次元に拡張し, 3 次元空間中のボール候補に対して軌跡モデルをフィッティングし, 正しいボールを抽出する. まずボールの運動を等加速度運動モデルに近似し, ボール候補に対してモデルをフィッティングする. この近似は短時間のみ成り立つものであるため, 時間領域における幅 h_w の走査窓を用いることで, フィッティングを h_w フレーム毎に行う. h_w はボールの運動を等加速度運動に近似できる程度の期間であり, 3 以上の整数とする. ステップ幅は $h_w/3$ とし, フレームの $2/3$ は前後の走査窓との重複させる. まず, 3 つの連続したフレームにおいて, ボール間の距離がボールがフレーム間で移動する距離の最大値を越えないボール候補 p_1, p_2, p_3 を選択する. 各フレーム時刻を t_1, t_2, t_3 , t_1 と t_2 および t_2 と t_3 の差分をそれぞれ Δt_{21} と Δt_{32} , 加速度を a , t_1 におけるボール速度を v_1 とすると, 等加速度運動モデルは以下の数式によって求められる.

$$v_1 = \frac{p_2 - p_1}{\Delta t_{21}} - \frac{\Delta t_{21} a}{2}. \quad (7)$$

$$a = 2 \frac{\Delta t_{21}(p_3 - p_2) - \Delta t_{32}(p_2 - p_1)}{\Delta t_{21} \Delta t_{32} (\Delta t_{21} + \Delta t_{32})} \quad (8)$$

ただし, 隣り合ったフレームのみから求められたモデルは不正確であることが多い. そのため, 求めたモデルから h_w フレームの期間の軌跡を復元し, その軌跡との距離が U_S 以内のボール候補の中から, 時刻が最小, 最大, 中央の 3 つの候補を選択して再度モデルを求める. この軌跡の最適化処理を, 選択される 3 点

が変化しなくなるまで、または、復元された軌跡と各ボール候補との誤差の総和が増加し始めるまで反復する(図 4)。



図 4. 軌跡モデルの最適化. 図中の黒い円は検出されたボール候補, 赤い円は軌跡モデルより推定されたボール位置であり, (a)は最適化前, (b)は3回最適化を繰り返した後の軌跡モデルを示している.

時間的に近い3点によって推定されたモデル(図 4(a))と比較して, 最適化処理を繰り返すことで時間が離れた3点によって推定されたモデル(図 4(b))は検出されたボール位置(図中の黒丸)との距離が近い. このように, 最適化処理を繰り返すことでより確からしいモデルが求められる. 以上の処理を, h_w 間の全ての候補を対象として行い, 誤差が最小となる軌跡モデルを選択する. このようにして復元された軌跡との距離が U_s 以内のボール候補と対応する画像上のボール位置を正しいボールとして抽出する. 抽出されたボール位置からは, スプライン補間によって各画像上のボール軌跡を復元する. 最後に, カメラ間の時刻ずれと軌跡によって各ボールの対応点を推定し, 三角測量によってボールの3次元位置を算出する.

2.3.2 衝突時刻の推定

数式(7)を用いることで, 前節で求められた各ボール位置における速度を算出する. このボール速度より, ボールと他の物体との衝突が生じた時刻を推定する. 本手法ではボールの軌跡を多項式で近似するが, その場合, 衝突のような急激な運動状態の変化を含む軌跡は正確に復元できない. そのため, ここで推定された衝突で区切ってボール軌跡を復元することで復元される軌跡の精度を高めることとする.

ボール速度の鉛直成分, および隣合うボール速度ベクトル間の角度の変位と手動で記録した衝突を図 5 に示した.

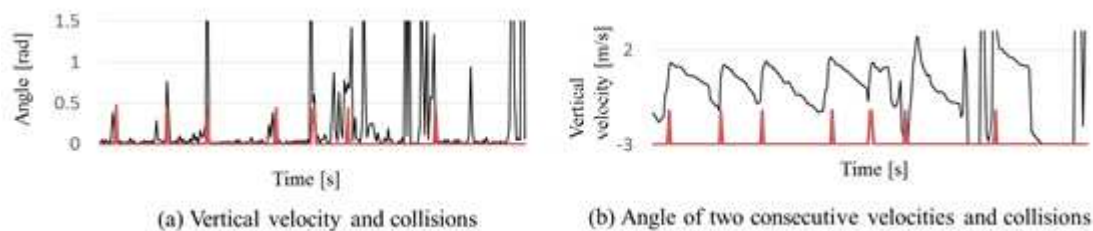


図 5. 衝突時のボールの運動状態の変化. 赤いパルスが衝突発生時点を示している.

ボールの衝突時には, ボール速度の鉛直成分の符号が負から正に変化する, あるいは隣り合う速度ベクトル間の角度が大きくなる. そこで, 以下の閾値を用いることで衝突を検出する.

- ・ 隣合う速度ベクトル間の角度の上限 [rad]: U_p

計測されたボール速度は計測誤差が含まれるボール位置から求めたものであるため, 衝突が発生していない場合にも衝突が発生している場合と類似した変化を示す場合がある. その場合には実際には衝突が起こっていない時刻で軌跡が分割されるが, 以降の処理で軌跡同士の融合を行うため最終的に得られる軌跡には影響を与えない.

2.3.3 軌跡の復元

最後にボールの軌跡を復元する. 前節で推定された衝突時刻によって軌跡を区切って復元し, それらを接続することで1つの軌跡を復元する.

まず, 衝突間の軌跡を復元する. 隣り合う2つの衝突間のボール位置から3次スプライン補間によって軌跡を復元し, サンプル間隔が等しくなるようリサンプリングする. リサンプリングされたボール位置を Savitzky-Gollay フィルタ (Savitzky, 1964) によって平滑化し, 再度, 3 次スプラインを用いて軌跡を復元する. 非同期カメラによって計測されたボール位置の時間間隔は均一ではないが, これによって一般的な平滑化手法を適用し, 計測ノイズを除去できる.

次に、複数の軌跡を接続する。ここでの軌跡の接続とは、ラリーの開始時刻から終了時刻まで必ず 1 つの軌跡からボール位置が推定されるように、各軌跡に対応する期間 $[t_s, t_e]$ を与えることである。ここで、時系列的に前の軌跡を T 、後の軌跡を T' 、 T の開始および終了時刻をそれぞれ t_s 、 t_e 、 T' の開始および終了時刻をそれぞれ t'_s 、 t'_e と表記する。時系列的に最も前の軌跡の開始時刻はラリー開始時刻、後の軌跡の終了時刻はラリー終了時刻が与えられるため、ここでは t_e と t'_s を求める方法を考える。隣り合う 2 つの軌跡は前節で推定された衝突で区切られており、基本的には図 6(a) のような関係にあると考えられる。

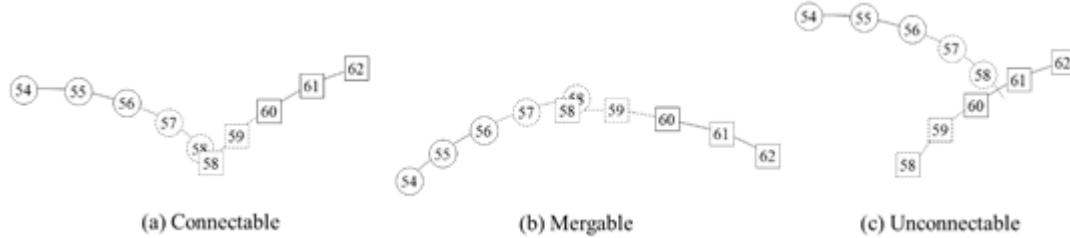


図 6. 異なる 2 つの軌跡間に成り立つ関係。図中の数字はボールが検出されたフレーム番号を示している。2 つの軌跡はフレーム番号 58 の時点で最も近づく。

しかし、計測された軌跡の中には、誤って検出された衝突によって複数の軌跡に分割されている場合が考えられる。例えば図 6(b) の 2 つの軌跡は、衝突時刻におけるボール位置は近いが、衝突時点におけるボールの進行方向がほとんど同じである。このような軌跡は、1 つの軌跡として再度軌跡モデルを計算すべきである。また、ボールではない物体の軌跡が混入している場合には図 6(c) のように、衝突時刻におけるボールの距離が離れている。このように 2 つの軌跡の間には、接続可能、接続可能且つ融合可能、接続不可の 3 つの関係があり、このいずれかを判定する必要がある。そこで、以下の閾値を用いて、軌跡間の関係を定義する。

- ・ 衝突時刻におけるボール位置の間の距離の上限 [m]: U_I
- ・ 衝突時刻におけるボール速度間の角度の上限 [rad]: U_M
- ・ 軌跡間の時間の上限 [s]: U_T

ここで、 T と T' を時系列的に互いに近づく方向に延長し、2 つの軌跡モデルから推定されるボール位置が最も近づく時刻を t_i と表記する。 T の復元時に用いられたボール位置の時刻のうち最も新しい時刻と T' の復元時に用いられたボール位置の時刻のうち最も古い時刻の差が U_T 以上、または、 T と T' の t_i における距離が U_I よりも大きい値の場合には接続不可であると判定する。 U_I 以下の場合、次に 2 つの軌跡モデルを用いて t_i 前後におけるボール速度ベクトルを求め、その角度が U_M より大きい値の場合には接続可能、 U_M 以下の場合には融合可能と判定する。接続可能な場合、 t_e と t'_s には t_i が設定される。融合可能な場合、 T と T' を復元する際に用いられたボール位置から再度軌跡モデルを計算し、新たに計算された軌跡の開始時刻は t_s 、終了時刻は次の軌跡と接続時に求められる。接続不可の場合、 t_e および t'_s には T の復元時に用いられたボール位置の時刻のうち最も新しい時刻と T' の復元時に用いられたボール位置の時刻のうち最も古い時刻の平均値を与えることとする。軌跡の接続の流れは以下の通りである。

- Step 1. 全軌跡間で接続可能な組み合わせを探索する
- Step 2. 接続可能な組み合わせの中で最も期間が長くなる組み合わせを抽出する
- Step 3. 既に抽出された軌跡との間に時系列的な矛盾が生じる軌跡を除外する
- Step 4. 未抽出の接続可能な軌跡の組み合わせの中で期間が最も長くなる組み合わせを抽出する
- Step 5. Step 3-4 を抽出される軌跡が無くなるまで繰り返す
- Step 6. 融合可能な組み合わせを融合する

以上の結果得られる軌跡の組み合わせが、本手法の出力として得られる 1 ラリーの軌跡である。

3. 実験

提案手法によって精度よく時間的および幾何学的カメラ校正ができること、そしてボールの 3 次元軌跡が復元できることを検証するために実験を行った。

3.1 実験条件

図 7 に実験環境を示した。



図 7. 実験環境

卓球経験者による2試合を対象とし、それぞれ撮影するカメラ、カメラと卓球台の距離、選手を変更した。以降、対象とした2試合をそれぞれ試合1、試合2と呼ぶ。試合1はGrasshopper3(Point Grey社)を2台用いて撮影した。カメラ間には3msの時刻ずれを設定し、解像度は1920×1080とした。カメラは図7(a)に示したように設置し、図8(a)、図8(b)に示したような画像が得られた。

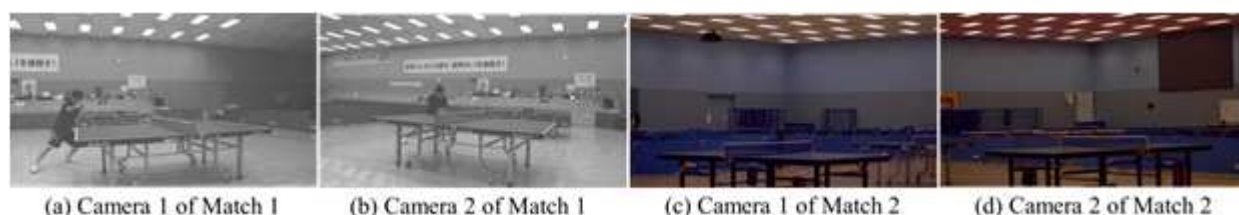


図 8. 撮影画像のサンプル。画像の明るさは視認性向上のため実際とは異なる。

試合2はLumix GH3(Panasonic社)を2台用いて撮影した。カメラ間の時刻ずれはカメラの仕様上調整できず、解像度は1920×1080とした。カメラは図7(b)に示したように設置し、図8(c)、図8(d)に示したような画像が得られた。両試合とも、ボールの検出はダウンサンプリングした960×540の画像から行い、検出後の座標値をダウンサンプリング前の座標値に変換した。ボールの検出率を評価するため、試合1からは35ラリー、試合2からは10ラリーをランダムに抽出し、ボール位置を手動で記録した。画像上のボール半径は最大で15 pixel程度であったため、手動で記録した位置から15 pixelよりも離れた位置で検出されたボール候補を誤検出と判定した。時間的および幾何学的カメラ校正は始めの10ラリーの検出結果を用いて行う。時刻ずれ探索時のステップは0.0165 msとした。時間的および幾何学的カメラ校正の精度は、試合1の結果より評価した。幾何学的カメラ校正の精度を評価するため、校正点計18点を用いた方法で外部パラメータを求め、最後にバンドル調整によって最適化した。バンドル調整にはLourakis and Argyros (2009)の公開するライブラリSBAを利用した。バンドル調整後の再投影誤差は0.21であった。時間的および幾何学的カメラ校正の結果については、野口と加藤(2006)の方法と提案手法との間で比較することとする。なお、試合2の撮影ではカメラ間の時刻ずれが未知であり、さらに従来手法によるカメラ外部パラメータの計測はされなかった。そのため、試合2の実験結果はボールの検出率を評価する目的のみで用いることとした。

提案手法における各閾値は複数の値を入力し、最も良い結果が得られる値を選択した。試合1を対象とした場合の閾値はそれぞれ L_I : 0.24, L_C : 0.3, L_A : 13, U_A : 531, U_V : 25, L_V : 1, U_D : 0.3, U_R : 0.524, U_E : 0.025, U_S : 0.15, U_P : 0.209, U_I : 0.015, U_M : 0.785, U_T : 1であり、試合2を対象とした場合の閾値のうち試合1の場合と異なるものは L_I : 0.12のみであった。本実験では試合1の画像の方が試合2の画像よりも輝度が高かったため、 L_I は試合1よりも小さい値となった。 L_C , L_A , U_A は解像度によって最適な値が変わるが、本実験では2試合とも同じ解像度で撮影したため同じ値であった。 U_D は画角によって最適な値が変わるが、図8に示したように全ての画角で類似した画角であったため、同じ値であった。なお、2視点撮影において卓球ボールと選手の重なりを少なくし、且つ光軸が垂直に近くなることを考慮した場合には図7に示したカメラ配置が最適であり、本実験ではこれ以外の条件については検討しなかった。 U_E , U_S は測定精度によって最適な値が若干変わるが、解像度が同じであれば同程度の精度で測定できるため本実験では同じ値であった。 U_V , L_V , U_R , U_P , U_I , U_M , U_T は卓球を対象とする限り同じ値を設定できるパラメータである。これら7つの閾値については、今後、多くの試合を対象としてボール軌跡を復元し、最適な値を定めることができ

れば定数として扱うことができると考えられる。

3.2 実験結果

3.2.1 カメラ校正前のボール検出

対象とした全ラリーを合計したフレーム数、ボールの存在するフレーム数を表 1, ボール候補検出の結果を表 2 に示した。

表 1. 対象フレーム数. #frames は全てのフレーム数, #ball-frames はボールの存在するフレーム数を示している

	Match 1		Match 2	
	View 1	View 2	View 1	View 2
#frames	3422	3379	1412	1397
#ball-frames	3312	3239	1199	1251

表 2. ボール候補検出の結果. Detected はボール候補に真のボールが含まれていた割合, Missed は含まれていなかった割合を示している

	Match 1		Match 2	
Indicator	View 1	View 2	View 1	View 2
Detected [%]	96.0	92.2	92.4	91.3
Missed [%]	4.0	7.8	7.6	8.7

ボール候補が検出されなかった原因は, 選手との重なりが 66.2% (298/601), ネットアセンブリとの重なりが 19.8% (119/601), ボールと類似した色の背景との重なりが 11.3% (68/601), 画像上での静止が 2.7% (16/601) であった. 画像上のボールは, カメラの光軸方向に移動している期間や投げ上げたボールが放物線の頂点付近で非常に遅くなった場合など, 画像上でほぼ静止する. その際に, フレーム間差分ではボールが背景と判定されて検出されないことが生じる. 未検出となる最も大きい要因は選手やネットといったプレー領域内の他の物体との重なりによるものであった. また, 多くのフレームでは正しいボールに加えて誤検出が含まれていた. 誤検出の多くは選手の移動領域であった. その中から 5 フレーム連続してボールらしい動きをするものを抽出した場合の検出結果を表 3 に示した.

表 3. 画像上の動きに基づくボール抽出の結果. TP は真陽性, FP は偽陽性, TN は真陰性, FN は偽陰性の割合を示す. Detection rate 3D はボールが存在するフレームに対する 3 次元座標を算出できる TP の割合を示す.

	Match 1		Match 2	
Indicator	View 1	View 2	View 1	View 2
TP [%]	82.7	70.7	62.2	64.9
FP [%]	0.3	0.1	0.1	1.6
TN [%]	3.2	4.1	15.1	10.5
FN [%]	13.8	25.1	22.6	23.1
Detection rate [%]	85.5	73.8	73.2	72.4
Detection rate 3D [%]	61.6	66.4	48.7	46.1
Precision [%]	99.7	99.9	99.8	97.6

抽出されたボールの 97% 以上は正しいボールであった. このことから, 画像上のボールの動きに基づいた抽出法は, ボールを用いたカメラ校正を目的とした場合には有効と言える. この抽出法がボール軌跡の復元を目的とした場合にも有効かを検討するため, もう一方の視点における対応点が推定できるもの, すなわち

3次元位置を計算できるものだけを抽出した結果を表3に「Detection rate 3D」として示した。この結果は、ボール軌跡を復元するためのボール抽出法としては検出率の低さが問題になる可能性を示している。提案手法では、この問題に対応するために、時間的および幾何学的カメラ校正の後に再度ボール抽出法を行う。その結果については後に述べる。

3.2.2 時間的および幾何学的カメラ校正

カメラ校正の結果を表4に示した。

表4. 時間的および幾何学的カメラ校正の結果

	Nogucni & Kato	Proposed method
Error of temporal offset [ms]	1.95	0.04
Error of optical axis [degree]	3.26	1.07
Error of camera center [m]	0.27	0.03
Reprojection error [pixel]	2.71	1.31

野口と加藤(2006)の方法をそのまま用いた場合には真値とは1.95 ms異なる時刻ずれが推定されたが、提案手法における修正を加えることで、ほぼ真値と同じ時刻ずれを推定できている。提案手法における基礎行列の誤差評価が正確であり、そのため大域的な正しい時刻ずれを探索できたと考えられる。再投影誤差は、試合撮影前にチェッカーボードを14の異なる位置および姿勢で撮影した。提案手法ではカメラの光軸と位置が正確に推定できており、そのことから再投影誤差も小さい値であった。この再投影誤差は実長だと約3.2mmであり、高い精度でカメラ校正ができていることが分かる。以上の結果より、スポーツプレー中のボールを用いて時間的および幾何学的カメラ校正を行う場合、提案手法が有効であると考えられる。

3.2.3 カメラ校正後のボール検出

時間的および幾何学的カメラ校正によって得られたカメラ間の時刻ずれ、およびカメラ外部パラメータを用いたボール検出の結果を表5に示した。

表5. カメラ間の時刻ずれおよび幾何学的位置関係を利用したボール抽出の結果。TPは真陽性、FPは偽陽性、TNは真陰性、FNは偽陰性の割合を示す。Detection rate 3Dはボールが存在するフレームに対する3次元座標を算出できるTPの割合を示す。

	Match 1		Match 2	
Indicator	View 1	View 2	View 1	View 2
TP [%]	85.4	85.0	66.7	69.8
FP [%]	0.8	1.9	3.9	2.1
TN [%]	3.2	4.1	15.1	10.5
FN [%]	10.6	9.0	14.3	17.6
Detection rate [%]	88.2	88.6	78.6	77.9
Detection rate 3D [%]	79.9	85.2	72.4	73.9
Precision [%]	99.1	97.8	94.5	97.0

検出されるボールの割合は試合1では約88%、試合2では約78%であった。このうち、3次元位置を計算できるものだけを抽出した結果(Detection rate 3D)を見ると、表3の結果と比較して18-27%多くのボールを抽出できている。このように、本手法によって時刻ずれやカメラ外部パラメータが既知とできることにより、1視点におけるボールの動きだけを利用する場合と比較して頑健性を高められることが示された。

ここで、試合1では約20%、試合2では約27%のボールが抽出されなかった要因を検討する。まず、10%前後のボールは未検出であったことが挙げられる。その理由については「3.2.1 カメラ校正前のボー

ル検出」で述べたように、選手、ネットおよび類似した背景色との重なりであった。次に、一方のカメラには映っていない、あるいは両方のカメラにおける隣り合うフレームに映っていないボールが多かったことが挙げられる。手動で記録したボール、すなわち映像に映る全てのボールを対象として、時刻ずれおよびエピソード線を用いたボール抽出を行った結果、試合 1 のカメラ 1 では 7.9%、カメラ 2 では 6.7%、試合 2 のカメラ 1 では 9.3%、カメラ 2 では 12.5% のボールが抽出できなかった。これらのことが、抽出されないボールがあった要因と考えられる。カメラの台数が 2 台だと、どのようにカメラを配置しても選手やネットとボールの重なりを防ぐことを防ぐのは困難であり、一般的な競技場面では背景に類似した色が映り込むことを完全に防ぐことはできない。また、計測精度を考慮すると 2 台のカメラ光軸は垂直に近く配置することになるため、そうすると 2 視点間の視差は大きくなり、選手とのオクルージョン等によってボールが一方のカメラには映らない、あるいは両方のカメラにおける隣り合うフレームに映らないことがある程度のフレーム数生じることは避けられない。本報告書では 2 台のカメラを用いた場合のカメラ校正法について述べたが、1 つのカメラと各カメラの間のカメラ校正として実施することで、3 台以上のカメラを用いる場合でも提案手法を適用できる。今後、計算できる 3 次元位置の割合を高めるためには、撮影する視点を増やし、少なくとも 2 視点にボールが映るようにする必要がある。

3.2.4 ボールの 3 次元軌跡の復元

復元された軌跡の例を図 9 に示した。

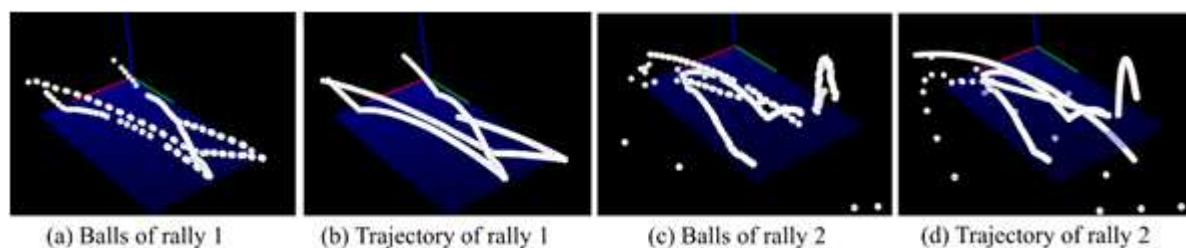


図 9. ボール軌跡の復元. (a) は (b) を復元する際に、(c) は (d) を復元する際に用いたボール位置。

図 9(b) は軌跡の復元が成功したラリーの一例である。復元に用いたボール位置(図 9(a))では図中右上のボールバウンド前後でボール位置を計算できていないが、前後の軌跡から綺麗に補間されている。復元された軌跡が検出されたボールの何割を包含できていたかを推定した。軌跡から求められたボール位置をカメラ 1 とカメラ 2 に再投影した結果、検出されたボール位置から 15 pixel 以内であった割合は試合 1 ではそれぞれ 98.3%、98.6%、試合 2 ではそれぞれ 93.5%、93.8%であった。このことから、撮影視点を増やし、検出できるボールの割合を高めることで、ラリー全体のボール軌跡は問題なく復元されることが考えられる。次に、復元された軌跡がラリー全体の軌跡の何割を復元できていたかを推定した。軌跡から求められたボール位置をカメラ 1 とカメラ 2 に再投影した結果、手動で記録したボール位置から 15 pixel 以内であった割合は試合 1 ではそれぞれ 93.9%、95.6%、試合 2 ではそれぞれ 80.6%、81.8%であった。表 5 のボール抽出率と比較して高い割合のボール位置を復元できている。これは、図 9(b)に見られたように軌跡の接続が精度良く行われることで、未検出期間が補間されたためである。前述した通り、様々な理由によってボールが未検出になることはある。本手法は、未検出の前後の軌跡が復元できている場合、この問題に対応できるものであると考えられる。ただし、全てのラリーにおいて問題なく軌跡が復元できたわけではない。図 9(d) は復元後の軌跡に誤りがみられたラリーの一例である。図中右下のバウンド直後のボールが検出できなかったため、卓球台を貫通する軌跡が推定されている。また、復元に用いたボール(図 9(c))に誤検出が含まれていたため、誤検出を含む期間において実際とは大きく異なる軌跡が復元されている。提案手法における 3 次元空間でのボール抽出は、時間領域に走査窓を設定し、各走査窓の中で最も長くボールに類似した動きをした物体を正しいボールと考えるものである。この方法では、ボールが未検出になる期間がある場合、選手の足などの他の物体がボールと判断される。表 5 に示したように本実験では最終的に残ったボールに 1-4% の誤検出が含まれていた。正しい軌跡を復元するためにも、撮影視点を増やすなどによってボールの未検出を減らす必要がある。

3.2.5 ボールの 3 次元軌跡からの情報推定

復元されたボール軌跡からパフォーマンス分析に役立つ情報をどれだけ抽出できるかを評価する必要がある。今回は、特に重要となる「バウンド」と「打球」の検出を試みた。バウンドはボールと卓球台との距離が

閾値以下になる期間において最も卓球台との距離がボール半径に近づいた時点として検出し、打球はボール速度のサイドライン方向成分の符号の反転および変化率に基づいて検出した。その結果、バウンドの検出率が97.2%, 打球の検出率が65.0%, 打球の計測位置の誤差が 2.9 ± 3.4 cm, バウンドの計測位置の誤差が 9.5 ± 20.2 cmであった。バウンドが高い確率で検出される一方、打球の検出率が低く、計測精度も両者で大きい差がみられる。この結果は、本実験で用いたカメラが2台であったことが影響している。撮影視点が2つの場合には、どのようにカメラを配置したとしてもボールと選手との重なりが発生し、特に打球位置付近で多く生じる。この結果より、やはり本手法を用いてパフォーマンス分析における情報収集の自動化を目指す場合、撮影視点を3視点以上にする必要があると考えられる。

4. 実用に向けたシステム設計

本研究で提案するボール軌跡復元法を簡易に且つ効率的に実施するためのシステムを設計する。ボール軌跡の復元は事前に録画した試合を対象に行う「オフライン処理」と、目の前で行われている試合を対象として撮影と同時に行う「オンライン処理」の2通りがある。図10にオフライン処理の場合とオンライン処理の場合とのワークフローを示した。

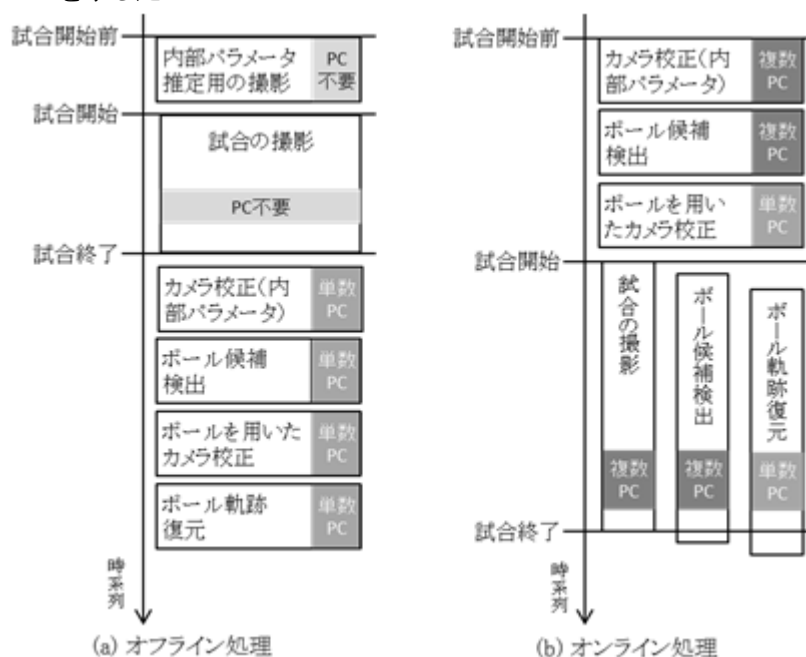


図 10. ボール軌跡復元のワークフロー. (a) オフライン処理と(b) オンライン処理とで必要とされる機能が異なる. オンライン処理における「複数 PC」と書かれた処理は撮影視点の数と同数の PC が必要とされることを示している.

オフライン処理(図 10(a))では、ボール軌跡の復元時には全ての試合映像がデータとして存在するため、全ての処理を1台のPCで実行できること、さらに試合映像の収集ではカメラ内部パラメータを計算するための映像を撮影する他は単にカメラを設置して録画ボタンを押下するのみであり、従来の試合映像撮影と作業内容が大きく変わらないことなど、作業に要するコストと手間が小さく、多くのケースで利用できる点にメリットがある。一方、オンライン処理は、ボール候補にかかる計算コストが高いためカメラの台数と同数のPCを用意する必要があり、試合映像撮影時にカメラのみならずPCを設置できる必要があるなど、オフライン処理と比較すると実施が難しいケースが多いと考えられる。しかしそのような環境が構築できた場合、プレーの時空間情報に関する分析結果の即時フィードバック、あるいはデータに基づく練習等、プレー時の内観とデータのすり合わせがしやすくなる点にメリットがあると考えられ、データを用いた競技力向上の有用性がさらに高まる可能性は高い。一長一短があり、利用環境や目的に適したものを選択する必要がある。

これらのワークフローいずれにも対応できるようシステムを設計した。本研究で提案されたボール軌跡復元法は(a)ボール候補検出, (b)ボールを用いたカメラ校正, (c)ボール軌跡の復元を行う3つのプログラムから構成される。各プログラムのフローチャートを図11に示した。

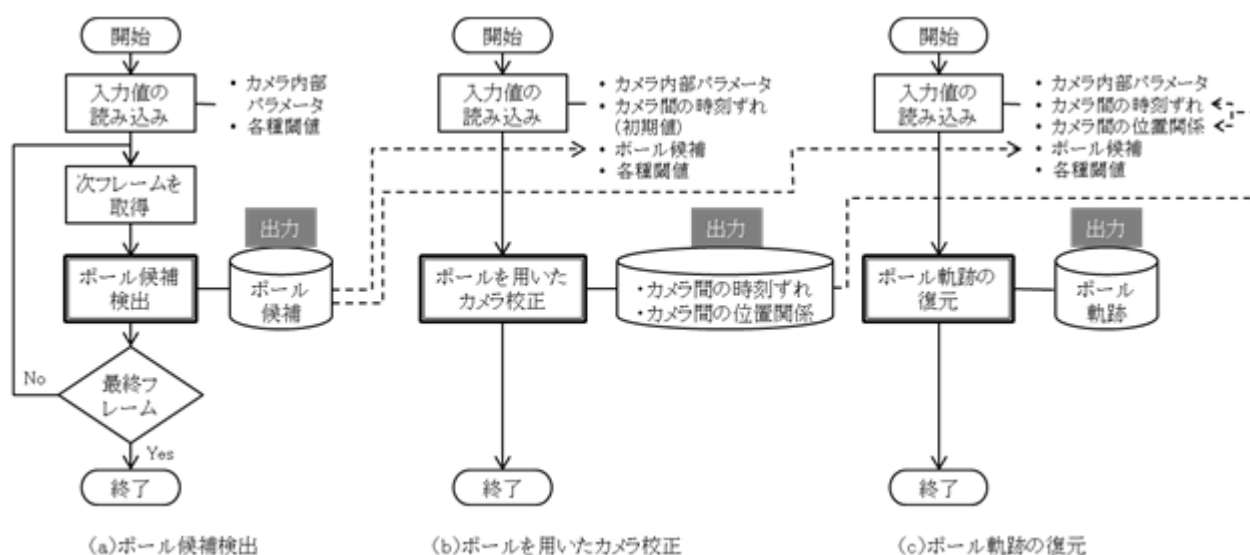


図 11. 各処理のフローチャート. 太線で囲まれた処理が本報告書で説明したアルゴリズムと対応する. 矩形は処理, 菱形は分岐, 円柱はデータベースを示す. 各処理の終了を待たずともデータベースにはアクセス可能とする.

ボール候補検出(図 11(a))はオフライン処理の場合には映像データから次々に読み込まれるフレーム画像を対象として, オンライン処理の場合にはカメラが次々にキャプチャするフレーム画像を対象として実行される. 検出されたボール候補はデータベースシステムに保存され, 即座に他のプログラムからアクセスできるようになる. オフライン処理では全ての処理を1つのPCで行うが, オンライン処理ではボール候補検出は複数台のPCで行うことになる. 複数台での処理が必要になるのはボール候補検出のみであるため(図 10), カメラの台数-1 台のPCにはボール候補検出用のプログラムのみインストールしておき, いずれか1台のPCに残りのプログラムをインストールしておくこととなる. プログラムの分割は, 2つのワークフローに対応できるようになることに加え, メンテナンス性の向上にもつながる. 各処理は完全に独立しており, いずれかのアルゴリズムを変更した際, 変更を加えていない部分と対応するプログラムはそのまま利用できる.

以上は, 本手法を実用システムとして実装するための大まかな設計であり, 最適なプログラム間の通信方法や利用するデータベースシステムは未検討の状態である. 実際に動作するシステムの実装を進め, 本設計の妥当性を検証することが今後の課題である.

5. まとめ

本研究では, 卓球パフォーマンス分析自動化を目的とした卓球ボールの3次元軌跡復元法を提案した. 提案手法は, 事前にカメラの位置関係を求める必要がなく, 且つカメラを同期させる必要が無いため, 競技会会場等でも利用しやすい点に特長があった. 本手法は(1)ボール候補領域の検出, (2)カメラ校正, (3)ボール軌跡復元の3つの処理に大別され, いずれも実際の卓球選手の試合を対象とした実験によって, その妥当性が検証された. ボールの軌跡復元法の妥当性検証後, ボールの軌跡復元全体を簡易に行うシステムが設計された. ボール検出, ボールを用いたカメラ校正, ボールの軌跡復元をそれぞれ独立したプログラムとし, システムのメンテナンス性を高めるとともに, オンライン処理およびオフライン処理にも対応できるよう設計された. 以上のように, 本研究では卓球の試合を対象としてパフォーマンス分析を行う上で重要となるボール軌跡の復元を簡易に行うためのアルゴリズム, および実用システムの開発が進められた. 今後, この研究成果に基づいたシステム開発を完成させ, 多くの試合を対象として卓球のパフォーマンス分析を実施することで, 卓球一流選手の試合の時空間的な理解を促進させることが課題となる.

参考文献

安徳知博, 鹿嶋雅之, 佐藤公則, 渡邊睦(2013) 動画画像処理によるテニスプレー自動認識・記録に関する研究, 情報処理学会研究報告.CVIM, 2013, 14:1.

R. Hartley and A. Zisserman (2003) Multiple view geometry in computer vision, Cambridge Univ Press: 279.

- S. Hu, F. Zaojun, X. De, and T. Min: (2013) Trajectory Prediction of Spinning Ball Based on Fuzzy Filtering and Local Modeling for Robotic Ping-Pong Player, *Instrumentation and Measurement*, 62(11): 2890.
- Y. Huang, J. Llach, and C. Zhang (2008) A Method of Small Object Detection and Tracking Based on Particle Filters, *Proc. 19th International Conference on Pattern Recognition*: 1.
- 各務友美, 坂上文彦, 佐藤淳, 石丸和寿, 今西勝之: 非同期マルチカメラによる高密度3次元復元, 画像の認識・理解シンポジウム(MIRU2012).
- 北原駿, 内田理 (2013) テニス競技映像からのプレー自動分析手法の提案, 映像情報メディア学会技術報告, 37(35): 51.
- 楠堀誠司, 吉田和人 (2011) 卓球ゲームにおける打法や状況毎の打球速度, 日本体育学会大会予稿集, 62: 239.
- H. Matsumoto, J. Sato, F. Sakaue (2010) Multiview constraints in frequency space and camera calibration from unsynchronized images, *Proc. Computer Vision and Pattern Recognition*: 1601.
- 石井規弘, 北原格, 亀田能成, 大田友一 (2007) 2視点からの映像を利用したサッカーボール追跡手法, 画像の認識・理解シンポジウム(MIRU2007).
- M. I. A. Lourakis, A. A. Argyros (2009) SBA: A Software Package for Generic Sparse Bundle Adjustment, *ACM Trans. Math. Software*, 36(1): 1.
- 野口真身, 加藤丈和 (2006) マーカの軌跡を用いた非同期カメラの幾何学的・時間的キャリブレーション, 情報処理学会論文誌.CVIM, 47(5): 59.
- A. Savitzky and M. J. E. Golay (1964) Smoothing and differentiation of data by simplified least squares procedures, *Analytical Chemistry*, 36(8): 1627.
- 清水彰一, 藤吉弘亘, 長坂保典, 高橋友一, 岩堀祐之 (2009) 複数の非同期パン・チルトカメラを用いた 3 次元位置推定法, 精密工学会誌, 75(2): 284.
- 高野橋健太, 眞鍋佳嗣, 安室喜弘, 千原國宏 (2005) 残像を利用したボールの 3 次元軌跡の計測, 情報処理学会研究報告.CVIM, 2005(112): 139.
- C. Wei, and Z. Yu-Jin (2006) Tracking Ball and Players with Applications to Highlight Ranking of Broadcasting Table Tennis Video, *Proc. Computational Engineering in Systems Applications*: 1896.
- F. Yan, A. Kostin, W. Christmas, and J. Kittler (2006) A Novel Data Association Algorithm for Object Tracking in Clutter with Application to Tennis Video Analysis, *Proc. Computer Vision and Pattern Recognition*: 634.
- P. Ying, J. Sato (2007) Computing Epipolar Geometry from Unsynchronized Cameras, *Proc. Image Analysis and Processing*: 475.
- X. Yu, C. H. Sim, J. R. Wang, and L. F. Cheong (2004) A Trajectory-Based Ball Detection and Tracking Algorithm in Broadcast Tennis Video, *Proc. International Conference on Image Processing*,: 1049.
- Z. Zhang (2000) A flexible new technique for camera calibration, *Pattern Analysis and Machine Intelligence*, 22(11): 1330.
- C. Zhou, H. Tao (2003) Dynamic depth recovery from unsynchronized video streams, *Proc. Computer Vision and Pattern Recognition*, 2: II-351.