

2 台の民生用カメラを用いた試合中のボール軌跡の 3 次元計測法

玉城 将^{*1}, 斎藤 英雄^{*2}

A method of reconstructing the trajectories of the ball in a sport match using two consumer cameras

Sho TAMAKI^{*1} and Hideo SAITO^{*2}

^{*1,2} Keio Univ. Graduate school of science and technology
3-14-1 Hiyoshi, Kohoku-ku, Yokohama-City, 223-8522 Japan

A method of reconstructing the 3D trajectories of a ball with two unsynchronized cameras is proposed. Currently, the common devices for 3D sensing in sports match scenario are time-synchronized cameras. There, however, exist several difficulties to conduct it. The method proposed in this paper requires only two known 3D coordinates and unsynchronized cameras. The key feature of the method is to use balls for camera calibration and the estimation of time-shift of between a camera and the other camera. The method is experimentally verified that it could provide accurate 3D ball trajectories.

Key Words : Sport engineering, Shape measurement, Image recognition, Table tennis, Unsynchronized camera

1. 緒 言

スポーツ場면을対象とした位置計測には様々なセンサが用いられるが、試合場면을対象とする場合、選手やボールにセンサを取り付けることが不可能な場合が多く、その場合には光学センサが唯一の計測手段である。光学センサにも複数の種類があり、位置計測に広く用いられるのは RGB カメラ（またはカラーカメラ）、赤外線カメラおよび Depth カメラ（または距離センサ）である。このうち Depth カメラは、信頼度の高い距離情報を得られる撮影範囲が狭い、あるいは解像度が RGB カメラと比較して非常に低い、さらには市販されている機器にサンプリングレートが 50 Hz を超えるものがみられないなど、スポーツ試合場面の移動物体を対象として用いることは難しい。赤外線カメラでは、対象物に再帰反射マーカを取り付ける必要があるなど、同様に試合場面では扱いが難しい。現状ではスポーツの試合場면을対象とした位置計測では RGB カメラを用いる方法で現実的あり、これまでも RGB カメラを用いて多くの研究がされている（例えば、大岡と前田⁽⁵⁾など）。本稿では、特に断りがない限りカメラとは RGB カメラのことを指す。

カメラを用いた位置計測では、2次元情報から3次元情報を復元する必要がある。これまで、多くの研究では、次のようなフローで3次元復元をしてきた。

- (1) 校正器具等を用いて既知の3次元点と画像上の2次元点を対応付けることでカメラ校正
- (2) 撮影時刻（あるいは露光）を同期した複数台のカメラで対象物体を撮影
- (3) 画像上の物体位置から三角測量によって3次元位置情報の復元

Direct Linear Transform と呼ばれる線形手法によってカメラ校正をすることから、(1) のカメラ校正法は3次元 DLT と呼ばれる。市販の3次元計測ソフトウェアで簡易に実施できることから、カメラを用いた3次元計測のデファクトスタンダードとなっている。しかし、既知の3次元点を必要とするため、試合開始前にプレー領域に校正器具などを配置する必要がある。また、試合開始前にカメラを校正できたとしても、試合中にカメラが動くと、その後、再度カメラ校正を行う必要がある。試合開始前の校正器具の配置が不可能な場合、会場の振動やカメラ

^{*1} 正員, 慶應義塾大学理工学研究科 (〒223-8522 神奈川県横浜市港北区日吉 3-14-1)

^{*2} 非会員, 慶應義塾大学理工学研究科 (〒223-8522 神奈川県横浜市港北区日吉 3-14-1)

E-mail: tamaki@hvrl.ics.keio.ac.jp

操作時にカメラが移動する可能性がある場合には、この方法を用いたカメラ校正はできない。また、複数台のカメラを用いるほとんどの場合、(2)のようにカメラの撮影時刻を同期させる。しかし、カメラの撮影時刻を同期させるにはカメラ間をケーブルで接続する必要があり、その設置が煩雑なだけでなく、試合会場によっては設置することが不可能な場合がある。GPSを用いて同期させる方法もあるが、室内競技を対象とする場合には遮蔽物の影響を受けるため十分な精度を出すことが難しい。また、そもそも民生用のカメラではシャッターを同期させる機能が無く、計測用に高価なカメラが必要となる点も場合によっては障害となる。

もし、校正器具を配置することなく、さらにシャッターを同期させる機能の無い民生用カメラを用いて3次元計測が可能となれば、多くのスポーツの試合会場で3次元計測を実施できるようになる。この技術は、単にスポーツの試合を対象とした計測機会を増やすだけでなく、これまでの実験環境に限られていたキネマティック的な解析手法と競技場面を繋ぐ役割を果たす可能性がある点にも意義がある。

しかし、非同期のカメラによる3次元計測(図1)は簡単ではない。画像から3次元空間の点の位置を求めるには、最低でも2視点からその点を撮影する必要がある¹。その点が移動する場合、非同期のカメラでは原理的に正確な3次元位置が算出できない。これまで、この問題の解決を試みた研究はいくつかみられる。Zhou and Tao⁽⁸⁾は、事前にカメラ校正をした2台のカメラによって直線的に運動する物体を2つの時刻で撮影し、合計4画像における物体の位置からカメラ間の時刻ずれを推定する方法を提案している。野口・加藤⁽⁴⁾は、2台のカメラによって撮影したゆっくり移動するマーカの画像上の軌跡を求め、その軌跡からカメラ校正および時刻ずれの推定を同時に行う方法を提案している。以上は、カメラ間の時刻ずれを求めることをベースとする方法であり、清水ら⁽⁷⁾の方法のように、時刻ずれを求めず、画像上のマーカの軌跡と2台のカメラの幾何学的な位置関係のみを用いて3次元計測を行うものもある。以上全ての方法は、撮影前に何らかの器具を用いてカメラを校正する、あるいは時刻ずれを推定することを仮定している。つまり、校正後にはカメラを移動できないという制約がある。従来の手法と同様、会場の振動やカメラ操作時にカメラが移動する可能性がある場合には、これらの方法を用いることはできない。

そこで本研究では、計測対象である飛翔中のボールを用いることで、その3次元軌跡と同時にカメラ校正や時間差を推定する方法を提案する。従来の研究のように撮影前にカメラ校正をしておく必要がないため、ボールが画像に映り込む時間以外であれば自由にカメラを移動できる。本手法は野口・加藤⁽⁴⁾の方法をベースとした方法であるが、マーカではなくボールを用いるため、高速で曲線的な動きを扱う点、シーン中の他の物体への遮蔽等を扱う点が異なる。また、本稿では特にボールの運動に焦点を当てているが、本質的にはあらゆる物体の運動を対象とできる方法である。

本稿の構成は以下の通りである。第2章では本稿で紹介する3次元計測の基礎となる射影幾何をまとめる。第3章では画像上の対応点を用いたカメラ校正法を紹介する。第4章では、非同期カメラを用いた3次元計測法の詳細を述べる。第5章では、実画像を用いて本稿で提案する手法を用いたボール軌跡の復元を実施し、その実用性を評価する。最後に、第6章で本稿をまとめる。

2. 射影幾何

本章では、まず本稿で用いる記法を導入する。次に本稿で紹介する方法の基礎となる射影幾何に関する理論について最小限の事項を述べる。なお、Hartley and Zissermanの文献⁽²⁾などに、詳しい情報がまとめられている。

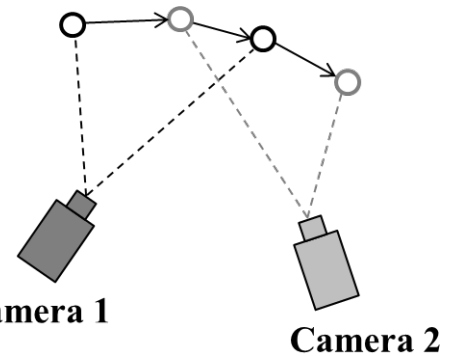


Figure 1. Unsynchronized cameras. Moving ball is not imaged simultaneously from two view point because the exposure of them is not synchronized.

¹ 三角測量によって計測する場合。他にも3次元計測の技術はあるが、移動する物体の3次元位置は三角測量によって計測されるのが一般的である。

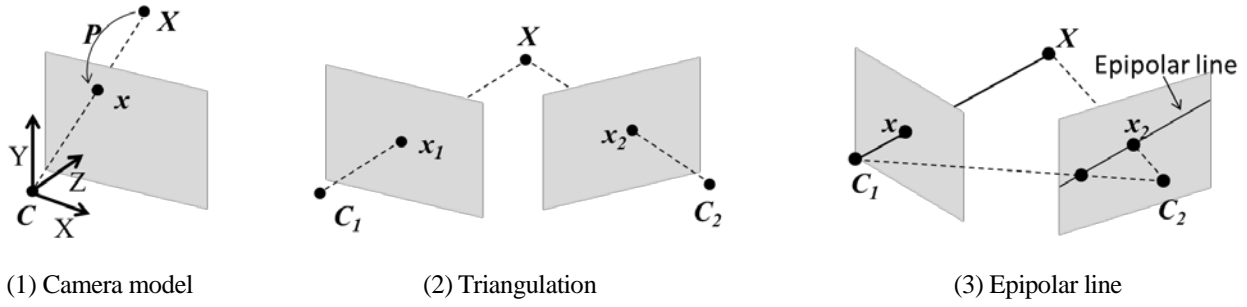


Figure 2. Projective geometry

2・1 記法

列ベクトルと行列を **bold** 体で表記する．特に単位ベクトルと零ベクトルを，それぞれ \mathbf{I} , $\mathbf{0}$ と書く．空間中の点 \mathbf{p} をその座標ベクトルで表すとき，それが斉次座標²であるなら $\tilde{\mathbf{p}}$ のように表記する．

2・2 カメラモデル

実際のカメラを透視投影によってモデル化したものを透視投影カメラモデルと呼ぶ．カメラを用いた3次元情報の復元に最も頻繁に使われ，これによって多くのカメラとレンズの組み合わせを近似できる．レンズ歪みについては透視投影による投影位置にレンズ歪みによる偏位が加わったものとしてモデル化できる．本研究でも透視投影カメラモデルを用いることとし，ここでは透視投影カメラモデルについて説明する．

図2(1)に透視投影カメラモデルによる撮像を示した．3次元空間での座標 \mathbf{X} とその画像上に投影された座標 \mathbf{x} との関係は，射影カメラ行列 \mathbf{P} によって次式のように記述される．

$$\tilde{\mathbf{x}} = \mathbf{P}\tilde{\mathbf{X}} \quad (1)$$

$$\begin{aligned} \mathbf{P} &= \mathbf{K}[\mathbf{R} | \mathbf{t}] \\ &= \begin{bmatrix} a_x & s & x_0 \\ & a_y & y_0 \\ & & 1 \end{bmatrix} \mathbf{R}[\mathbf{I} | -\tilde{\mathbf{C}}] \end{aligned} \quad (2)$$

ただし， a_x , a_y はそれぞれ画素単位で表した x , y 方向の焦点距離を表し， s は画像の縦軸と横軸の間の非直行歪み， (x_0, y_0) はカメラ光軸と画像平面との交点を表す．以上の5つのパラメータはカメラ内部パラメータと呼ばれる，カメラの位置・姿勢によらないカメラ固有の定数であり，上三角行列 \mathbf{K} に集約される．これに対し，カメラの位置 \mathbf{C} (あるいは並進 \mathbf{t}) と姿勢 \mathbf{R} を外部パラメータと呼ぶ．

透視投影カメラモデルは，3次元空間座標，画像上の座標，そして光学中心が全て直線上にあるという理想化されたモデルであり，実際にはレンズ歪みを考慮する必要がある．特に，レンズの焦点距離が小さい場合にはその影響は大きい．レンズ歪みに関しては Brown⁽¹⁾ により提案された，放射方向歪みと接戦方向歪みを考慮したモデルが広く利用されている．ここで，カメラ内部パラメータの影響を受ける前の画像座標系 (つまり $\mathbf{K}^{-1}\mathbf{x}$) を用いてレンズ歪みモデルを定式化する．レンズ歪み補正前の画像座標成分を (x', y') ，画像中心からの距離を r ，放射歪み係数を Dr_1, Dr_2, Dr_3 ，接戦方向歪み係数を Dt_1, Dt_2 とすると，レンズ歪みのモデルは次式のようになる．

$$\begin{aligned} x &= x' + x'(Dr_1r^2 + Dr_2r^4 + Dr_3r^6) + Dt_1(r^2 + 2x'^2) + 2Dt_2x'y' \\ y &= y' + y'(Dr_1r^2 + Dr_2r^4 + Dr_3r^6) + Dt_2(r^2 + 2y'^2) + 2Dt_1x'y' \end{aligned} \quad (3)$$

² 空間の次元を1つ上げ，これらの比によって空間を表現するような座標の取り方．例えば2次元平面上の点 $\mathbf{x} = (x, y)^T$ を3つの実数の組 $(x_1, x_2, x_3)^T$ によって表す．ただし， $x = x_1/x_3$, $y = x_2/x_3$ を満たし， x_1, x_2, x_3 のいずれか1つは必ず0ではない．

2・3 3次元形状の復元

2台のカメラのカメラ内部パラメータ、レンズ歪み、およびカメラ外部パラメータが既知であれば、対応点の3次元位置を復元できる (図2 (2)). 3次元空間中の点と片方の画像上の点の斉次座標をそれぞれ $[X, Y, Z, 1]^T$, $[x, y, 1]^T$ とし、次のような射影カメラを考える.

$$\begin{aligned} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} &= P \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \end{aligned} \quad (4)$$

式 (4) より、次の2つの拘束式が得られる.

$$\begin{bmatrix} p_{31}x - p_{11} & p_{33}x - p_{12} & p_{33}x - p_{13} \\ p_{31}y - p_{21} & p_{32}y - p_{22} & p_{33}y - p_{23} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_{14} - p_{34}x \\ p_{24} - p_{34}y \end{bmatrix} \quad (5)$$

同様にもう一方の画像上の点から2つの拘束式が得られるため、合計4つの拘束式が得られる. 未知数3に対して拘束式が4つであり、3次元空間中の座標は求まる. 最小二乗法で線形に求めることもできるが、反復解法によって得られる画像上の点と光学中心を通る直線との距離が最小となる点が最尤解となる.

2・4 基礎行列とエピポーラ線

3次元空間中の点 \mathbf{X} を2視点から撮影し、それぞれの画像上の座標を $\mathbf{x}_1, \mathbf{x}_2$, その斉次座標を $\tilde{\mathbf{x}}, \tilde{\mathbf{x}}$ とおくと、これらの間には次式が成り立つ.

$$\tilde{\mathbf{x}}_2^T \mathbf{F} \tilde{\mathbf{x}}_1 \quad (6)$$

この行列 \mathbf{F} は基礎行列と呼ばれ、カメラ内部パラメータ、および2台のカメラの相対的な位置や姿勢の情報を含んでいる. 第1の画像において点 \mathbf{x}_1 が得られたとすると、次式によって第2の画像にエピポーラ線と呼ばれる1つの直線 \mathbf{l} を投影できる (図2 (3)).

$$\tilde{\mathbf{l}} = \mathbf{F} \tilde{\mathbf{x}}_1 \quad (7)$$

エピポーラ線とは、第1画像の点 \mathbf{x}_1 に対応する第2画像上の全ての点の集合である. もしカメラ内部パラメータが既知の場合、基礎行列からはカメラの6自由度の運動のうちで、並進のスケールを除く回転3自由度と並進2自由度が求められる. そのため、カメラ内部パラメータおよび基礎行列が求められれば、スケールの不定性を含んだ3次元形状を復元できる. 並進のスケールについては、実長が既知の線分などから求めることができる.

基礎行列は 3×3 であるが、定数倍の不定性があり、且つ階数が2であることから自由度は7である. 2画像におおける1組の点の対応から式(1)により1つの拘束が得られるため、7点の対応がわかれば基礎行列が求まる. ただし、量子化誤差や対応点を記録する際の誤差の影響を抑えるために8点以上の対応点から頑健に基礎行列を求める方が良い結果が得られる. ここで、点と直線の距離を算出する関数を $d(\mathbf{x}, \mathbf{l})$ と表すと、基礎行列 \mathbf{F} を推定するため、あるいは評価するための誤差関数は、次式の画像上での幾何学的誤差がよく用いられる.

$$E(\mathbf{F}) = \frac{1}{N} \sum_i^N \left\{ d(\mathbf{x}_{1,i}, \mathbf{F} \tilde{\mathbf{x}}_{2,i})^2 + d(\mathbf{x}_{2,i}, \mathbf{F}^T \tilde{\mathbf{x}}_{1,i})^2 \right\} \quad (8)$$

3. カメラ校正

カメラ校正とは各カメラのカメラ内部パラメータ、外部パラメータ、レンズ歪みを求めることである. ただし、レンズ歪みについては、用いるレンズや焦点距離によってはその影響が無視できる場合もある. スポーツの試合場面を対象とした研究では、3次元座標が既知の点 (以下「コントロールポイント」という) を6点以上用いて、各カメラで射影カメラ行列を算出することが多い. しかし、プレー領域にコントロールポイントが少ない場合に

は、校正器具等を用いてコントロールポイントを設定する必要がある。このことは、3次元計測の実施を難しくする。ここでは、主に画像上の対応点を用いたカメラ校正の方法を紹介する。そのフローは以下の通りである。

- (1) 各カメラのカメラ内部パラメータおよびレンズ歪み係数を算出
- (2) 各カメラでプレー領域を撮影
- (3) 対応点から基礎行列を算出
- (4) 基礎行列から並進のスケールを除くカメラ外部パラメータを抽出
- (5) 3次元座標が既知の複数の点から並進のスケールを算出

このフローでは、最小2つのコントロールポイントでスケールの不定性なくカメラ校正ができる。スポーツのプレー領域では、例えばコートラインの交点など、2点間の実長が既知の図形が映り込むことは多い。その場合、撮影前にコントロールポイントを配置しておく必要はない。(4)の基礎行列算出には、できるだけ多くの対応点が必要になる。例えば、プレー中の移動物体(選手やボールなど)を用いることで多数の対応点を記録することは可能である。なお、基礎行列を算出するためには対応点全てが1つの平面上にあってはいけないという制約に注意が必要である。厳密に平面上になくても、それに近い場合には正しい基礎行列は求まらない。

最後に、以上のカメラ校正に関する各処理のコンポーネント化が進んでいるため、プログラムの実装が容易である点についても言及する。(1)と(3)はC, C++, Python, Javaから利用できるライブラリであるOpenCV⁽³⁾の関数を呼び出すだけで実装できる。あるいは、MatlabであればComplete Camera Calibration Toolboxに(1), Computer Vision System Toolboxに(3)を行う関数がある。(4)については、Hartley and Zissermanの文献⁽²⁾に詳しく説明されている通り、基礎的な行列演算を行うだけであり、高度なプログラミングスキルは必要ない。

4. 非同期カメラを用いたボール軌跡復元法

第1章で述べた通り、非同期カメラによる3次元計測に関してはいくつかの先行研究がみられる。まず、ここまでに述べた射影幾何の知識を用いて、主要な2つの先行研究の概要を述べ、続いて提案手法について述べる。

4・1 先行研究

清水ら⁽⁷⁾は、2台の非同期カメラである運動物体を撮影し、一方の画像における軌跡と、他方のカメラにおける物体位置 p_1 から求めたエピポーラ線との交点を p_1 の対応点 p_2 とすることでリアルタイムに3次元位置を計測する方法を提案している。この方法は、カメラ間の時刻ずれが一定ではない場合にも対応できる点、一方のカメラがパン・チルトする場合には運動物体の動きを2次曲面で近似し、他方の画像上の物体位置と3次元空間中の点を結んだ線分との交点を求めることで3次元位置を計測できる点に特長がある。しかし、エピポーラ線と軌跡が平行に近くなると正しく交点を求められないことや、片方のノイズが直接的に3次元位置の誤差に影響を与えることなどの問題もみられる。また、カメラ校正は事前に行われていることを仮定しており、軌跡復元とは別の枠組みでカメラ校正の方法を考える必要がある。

野口・加藤⁽⁴⁾は、2台の非同期カメラでゆっくり移動するマーカを撮影し、カメラ間の時刻ずれとカメラ校正を同時に行う方法を提案している。フレーム間のマーカの運動を直線で近似することで、カメラの時刻ずれ t から一方の画像上のマーカと対応するもう一方の画像上のマーカ位置を推定でき、さらにはその対応点を用いて基礎行列が計算できる点に着目し、時刻ずれ t に関する最適化計算によって最適な時刻ずれと基礎行列を求めている。カメラ間の時間差が一定であることを仮定するが、他の先行研究⁽⁶⁾⁽⁸⁾でも時刻ずれが一定であると仮定するものは多く、且つ実画像を用いた結果でよい精度を示していることから、この仮定に問題はないと考えられる。また、清水ら⁽⁷⁾の方法でみられたような、エピポーラ線と軌跡が平行に近くなることが計測精度に与える影響はない。ただし、フレーム間のマーカの動きを直線で近似していることから、フレームレートと比較してマーカ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。本研究では、試合中のボールをマーカの代わりとして用いることを考えており、そのためには、フレーム間の運動物体の動きが曲線的である場合にも対応できる必要がある。また、試合中のボールは他の物体に隠れることも多く、その対応が必要となる。

4・2 提案手法

本手法では、3次元空間中を移動するボールを2台の非同期カメラによって撮影し、カメラ校正、カメラ間の時刻ずれの推定、および3次元計測を同時に行う。そのフロー(図3)は以下の通りである。

- (1) カメラ内部パラメータ K_1, K_2 およびレンズ歪み係数 $Dr_1, Dr_2, Dr_3, Dt_1, Dt_2$ の算出

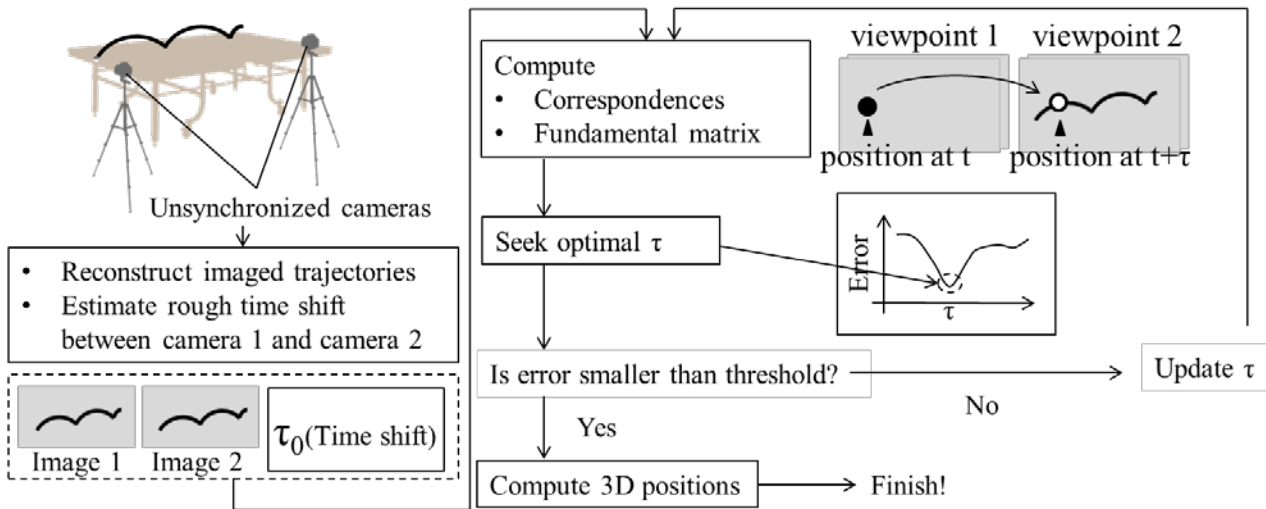


Figure 3. Flowchart of proposing method

- (2) 大まかな時刻ずれ τ_0 の算出
- (3) 画像上のボール位置 p_1, p_2 の記録
- (4) 画像上のボール軌跡の復元
- (5) 時刻ずれと基礎行列の推定
- (6) 基礎行列と内部パラメータによるカメラ外部パラメータの推定
- (7) 実長が既知の線分によるスケールパラメータの推定
- (8) ボールの3次元座標の算出

本稿で紹介する方法において手作業が必要となるのは (2), (3), (7) であるが、いずれも技術的には自動化が可能な処理であるため、本質的には全て自動化が可能な枠組みである。(1), (6) の方法については、第3章で紹介した通りである。以降、(4), (5), (7) について詳しく述べる。

4・2・1 画像上のボール軌跡の復元

画像からボール位置を n フレーム連続して記録できた区間のみ3次スプライン曲線によって軌跡を復元する。復元される軌跡は時刻を入力とし画像上の位置を出力とする時間関数である。ただし、記録されたボール位置には記録する際の誤差が含まれるため、事前に Savitzky-Golay フィルタ等によって平滑化しておく。軌跡の復元に使用した連続したフレームの開始時刻 t_{start} 、終了時刻 t_{end} を軌跡 l と対応付けて記録する。すなわち、以上の処理により、軌跡はボール位置が n フレーム連続して記録できた複数の区間が $\{l_i, t_{start,i}, t_{end,i}\}$ という形式で記録される。ただし、添字の i は区間の番号を示す。

4・2・2 時刻ずれと基礎行列の推定

各画像上の軌跡が時間関数の形で得られているため、仮に正しいカメラ間の時刻ずれが既知であれば、一方の画像で撮影されたボール位置から、第2の画像における対応点が求められる。そうすると、2画像間の対応点が多数得られるため、基礎行列が推定できる。しかし、ここまでで得られている時刻ずれ τ_0 は数フレームの誤差を含む時刻ずれであるため、対応点と基礎行列は正しい値にはならない。そこで、野口・加藤⁽⁴⁾の方法を参考にし、基礎行列と時刻ずれの誤差最小化計算を行う。ここで、各視点の画像上のボール位置をそれぞれ p_1, p_2 、その対応点を c_1, c_2 、 p_1 と c_1 、 p_2 と c_2 のペアによって求めた基礎行列を F 、式(8)によって求めた基礎行列の誤差を E と表すと、誤差最小化のアルゴリズムは以下の通りである。ただし、1回目の処理の τ_j には τ_0 が代入される。

Step 1 時刻ずれ τ_j を用いて点 p_1, p_2 の対応点 c_1, c_2 を求める。

Step 2 点 p_1, p_2 とその対応点 c_1, c_2 より基礎行列 F を推定する。

Step 3 式(8)を最小化するような時刻ずれ τ_{j+1} とその時の誤差 E を求める。

Step 4 E が閾値以上の場合、 $j = j+1$ としてStep 1に戻る。それ以外の場合には処理を終了する。

Step1では、画像上の軌跡の開始時刻と終了時刻の間の区間のみ対応点を求める。例えば、第1の画像上の点 $p_{1,i}$ の対応点を復元する時、 $p_{1,i}$ の時刻 $t_{1,i}$ に関して、 $t_{start} + \tau < t_{1,i} < t_{end} + \tau$ である軌跡が第2の画像上で復元され

ていない場合には、対応点は求めない。軌跡の外の区間でも対応点を求めることは可能だが、信頼性が落ちるため、時刻ずれと基礎行列の推定では用いないこととした。

4・2・3 ボールの3次元座標の算出

既に各画像上の軌跡、正確な時刻ずれが求められている。これらを使い、まず、各視点で記録できたボール位置 p_1, p_2 の対応点 c_1, c_2 を求める。この時、前節で説明した時刻ずれと基礎行列の推定とは異なり、軌跡の外の区間についても、直前あるいは直後いずれかの軌跡を用いて対応点を推定したい。時系列的に最初の軌跡の開始時刻よりも前、あるいは最後の軌跡の終了時刻よりも後の場合には、それぞれ、最初の軌跡、最後の軌跡を用いて推定する。ある軌跡の終了時刻と、他の軌跡の開始時刻の間の時刻の場合、どちらの軌跡を用いるかは、2つの軌跡の最も近づく時刻を基準とし、どちら側にあるかで決定する(図5)。例えば、ある画像に軌跡 $T_1\{l_i, 1000, 300\}$ と軌跡 $T_2\{l_{i+1}, 450, 600\}$ があり、2つの軌跡が時刻350で最も近づく場合、 $(-\infty, 350]$ は軌跡1、 $(350, \infty)$ は軌跡2によって位置を推定する。以上で画像間の対応点、およびカメラ内部パラメータ、カメラ外部パラメータ、レンズ歪み係数が求められているため、式(5)の拘束を用いてボールの3次元座標を算出する。

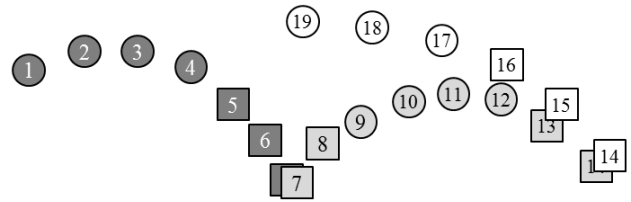


Figure 5. Determining switching points of trajectories. Circles are recorded ball. Squares are estimated ball from a trajectory model. The color of ball represents which trajectory they obey. The number of balls are the time.

5. 実画像を用いた実験

卓球の試合を対象として提案手法によるボールの3次元軌跡復元を行った。本章では、まず実験条件および実験結果を説明し、次に実験結果から本手法の精度および制限について考察する。

5・1 実験条件

卓球の男子トップ選手による1試合を撮影し、試合中からランダムに10本のラリーを抽出した。本実験では、抽出されたラリー中のボールを対象として3次元軌跡を復元することとした。撮影には民生用カメラ Lumix GH3 (Panasonic) を使用し、解像度 1920×1080 、フレームレート 60 fps を設定し、カメラ間でシャッターの同期はしなかった。カメラ間の撮影時刻のずれは、撮影された映像をコマ送りで再生し、目視で約 106 フレームであると推定し、これを初期値とした。

軌跡の復元は提案手法および野口・加藤の方法を用いて行う。ただし、野口・加藤の方法では移動体(本実験の場合にはボール)が未検出の場合の処理が無いため、そのまま実施すると著しく精度が低くなる可能性がある。そのため、未検出の場合の対応は提案手法と統一し、フレーム間のボールの動きをスプライン曲線で近似するか、それとも直線で近似するか、という部分のみに着目し、両者を比較することとした。なお、両手法とも反復的な誤差最小化を用いて近似解を求めるものであり、常に同じ解が求まる保証はない。そのため、1ラリーに対して計20回ずつ軌跡を復元し、その安定度を評価する。また、基礎行列の誤差の収束判定に用いる閾値は両手法ともに5とする。これは、ボール位置の記録時に1.5 pixel程度の誤差が含まれていることを仮定した値であり、経験的にこの値が設定された。

5・2 結果

ラリー毎の再投影誤差の平均値および標準偏差を図6、ラリー毎の時刻ずれの分布を図7、復元されたボール軌跡を図8に示した。カメラ間の時刻ずれは提案手法、野口・加藤の方法ともに 107.10 ± 0.02 ms であった。再投影誤差は提案手法では 7.50 ± 8.09 pixel、野口・加藤の方法では 7.96 ± 8.9 pixel であり、特に誤差の大きい2本の試技を除いた場合の再投影誤差は 4.02 ± 2.70 pixel、野口・加藤の方法では 4.27 ± 3.14 pixel であった。

5・3 考察

時刻ずれは 107.06 – 107.12 フレームの範囲内で推定されており、真の刻ずれはこの範囲内にある可能性が高い。そのため、時刻ずれの残差は最大で 0.06 フレーム、すなわち 1 ms であり、正確な時刻ずれが推定できていると考えられる。

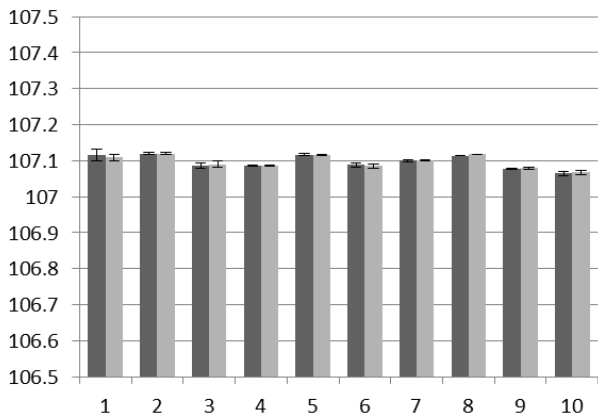


Figure 6. Estimated timeshift of each trials [ms]. The numbers in horizontal axis denotes the number of trial.

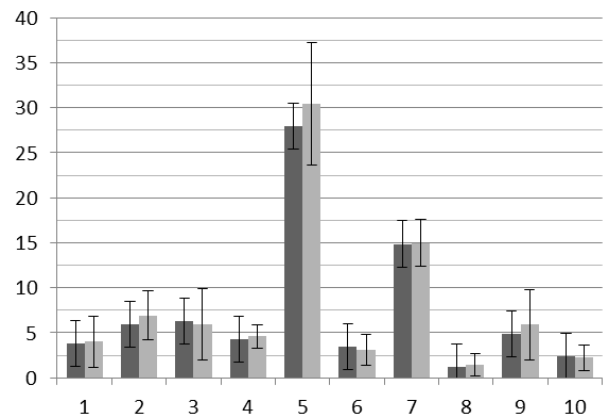
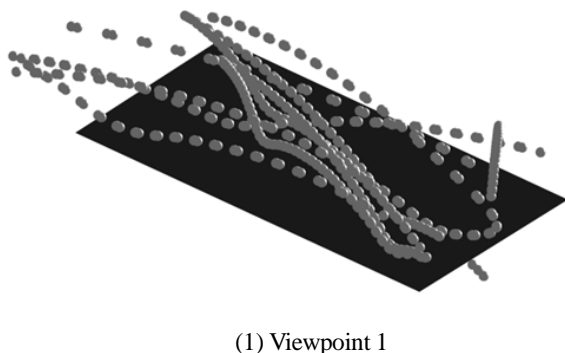
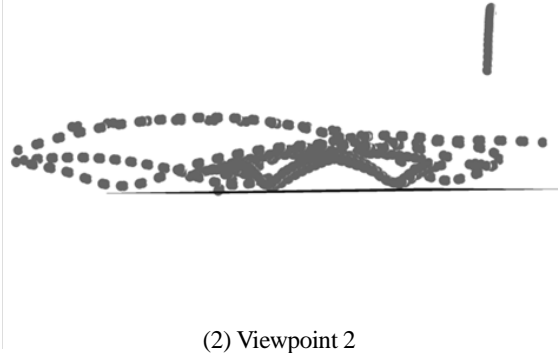


Figure 7. Reprojection error of each trials [pixel]. The numbers in horizontal axis denotes the number of trial.



(1) Viewpoint 1



(2) Viewpoint 2

Figure 8. Reconstructed ball trajectory. The black rectangle denotes a table. The grey spheres denote balls.

再投影誤差は、提案手法の方がやや小さい値であった。卓球ボールを対象とした際、フレーム間の軌跡を曲線で近似することでの精度が大きくなるものと推察できる。多くのラリーで0-6 pixelの範囲内となっている。本実験で撮影された映像において6 pixelは実長換算で約12 mmである。この誤差には、ボール位置を記録する際の誤差と復元されたボール3次元位置の誤差の両方が含まれている。仮に、全てボール3次元位置の誤差であると考えた場合でも、例えば1200 mm以上飛翔するボールの平均速度を測定する場合等であれば、かなり正確な値が求められると考えられる。復元された軌跡の形状(図8)からもその正確さが理解できる。ただし、2本のラリーは再投影誤差が非常に大きかった。時刻ずれは他のラリーと同程度の値が推定されていることから、正しい基礎行列を求められなかったことが原因であると推察される。正しい基礎行列を求められなかった原因は不明であり、安定性の向上のためにさらなる解析が必要である。基礎行列が正しく求められないことが起こり得ることは、本手法の原理上の限界である。既に述べた通り、対応点が1つの平面上、あるいはそれに近い場合には正しい基礎行列は求まらない。卓球の場合だと、例えばラリーがサービスミスで終了した場合、極端に同じコースのみを用いてプレーした場合など、ボールの軌跡が全て1つの平面上にある、あるいはそれに近いケースがあり、その場合には基礎行列が正しく求められない。これについては、再投影誤差が非常に大きい場合には他のラリーで求められた基礎行列を再利用する、あるいはカメラが固定されている期間の複数のラリーのボール位置を用いて基礎行列を用いるなどの回避方法を併せて考える必要がある。

今回、12 mm程度の残差が残された。実験室で計測する場合のように1 mm以下の精度が求められる場合、この精度では問題となる。本実験の結果からは、時刻ずれはある程度正確に求められており、カメラが非同期であることは、大きな精度劣化につながっていないことが推察される。すなわち、さらに精度を高めるには、画像上のボール位置の精度を向上させる、およびカメラの台数を増やすことが考えられる。カメラの台数は、今回は2

台のみのケースを考えたが、基礎行列を多視点幾何に一般化した概念を用いることで3台以上に拡張可能である。ボール位置は、マーカを用いないためどうしても精度が低くなるため問題となる。ボールの形状あるいは色の特徴から、ボールの中心座標をサブピクセル精度で得ることができれば、この問題に対応できる可能性がある。これについては、さらなる検討が必要である。

6. まとめ

本研究では、飛翔中のボールを用いることで、その3次元軌跡と同時にカメラ校正や時間差を推定する方法を紹介した。従来の研究のように撮影前にカメラ校正しておく必要がないため、ボールが画像に映り込む時間以外であれば自由にカメラを移動できる点に特長があった。実験より、本手法が正確に時刻ずれを推定できること、およびボールの3次元軌跡を求められることが検証された。今後、ボール検出時の精度向上や、カメラの台数を増やすことにより、さらに精度を向上させることが課題となる。

謝 辞

本研究の一部は、(独)情報通信研究機構の委託研究「革新的な三次元映像技術による超臨場感コミュニケーション技術の研究開発」の補助により行われたものである。

文 献

- (1) Brown, D. C., "Close-range camera calibration". Photogrammetric engineering, Vol. 37, No. 8, pp. 855-866.
- (2) Hartley, R., Zisserman, A., "Multiple view geometry in computer vision", Cambridge Univ Press, (2004).
- (3) Itseez, "OpenCV", <http://opencv.org/>.
- (4) 野口真身, 加藤丈和, "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, Vol. 47, No. 5 (2006), pp. 59-68.
- (5) 大岡昌平, 前田正登, "野球の試合における投球されたボールの軌道", シンポジウム:スポーツ・アンド・ヒューマン・ダイナミクス講演論文集, (2012), pp. 330-335.
- (6) Seki, A., Okutomi, M., "Simultaneous Optimization of Structure and Motion in Dynamic Scenes Using Unsynchronized Stereo Cameras", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, (2007), pp. 1-8.
- (7) 清水彰一, 藤吉弘亘, 長坂保典, 高橋友一, 岩堀祐之, "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, Vol. 75, No. 2 (2009), pp. 284-289.
- (8) Zhou, C., Tao, H., "Dynamic depth recovery from unsynchronized video streams", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (2003), pp. II-351-8 vol.2.