テニスプレーにおけるスキルの定量評価*

浅野 敏郎** 吉田 智幸*** 玉野 和保**

Quantitative evaluation of tennis skills by computer vision

Toshio ASANO, Tomoyuki YOSHIDA and Kazuho TAMANO

In this paper, tennis matches are analyzed, and skills of players are evaluated quantitatively by computer vision technology. Two cameras are used to detect the 3-dimensional positions of balls and players. Subtractions of images are used to detect ball images and the prediction of ball positions by using velocity vectors of previous images enables accurate and high-speed ball detection. Players are accurately detected by elimination of white coat lines using image processing technologies. The sequential pictures of tennis experts and beginners were recorded and five scenes, which include tennis play skills, were analyzed quantitatively. The 3-dimensional trajectories of balls and players are calculated and skills between experts and beginners are compared, and the prominent differences between experts and beginners were recognized quantitatively about the trajectories of balls, movement and positions of players, height of balls and speeds of balls.

Key words: skill evaluation, stereo image, tennis, 3-dimensional position detection, image processing

1. 結 言

近年、スポーツを科学的に解析するニーズが高まっている¹⁾。撮影した映像を専門家が見て選手のプレーを検討し、スキル向上の道具として役立っている。しかし、人間による判断は感性評価であり、定量的でなく、人により評価結果が違って来る。選手へのアドバイスも、「もう少し」「そのあたり」「だいたい・・・」などのあいまいな感性用語を用いて行われており、分かりにくいという問題点があった。このため、例えば球技においては、ボールや選手についての情報を定量的に検出し、試合を定量評価する技術が求められていた。

以上のニーズに基づき、これまで画像処理技術を用いて、サッカーゲームを解析する研究が多く行われてきた。例えば、TV 中継画像からボールと選手の軌跡を抽出する研究²⁾、選手の動きの解析支援を行う研究^{3) 4)}、サッカー中継におけるカメラの自動視点切り替えによる効果的なボール追跡に関する研究⁵⁾、固定カメラ映像から 3D アニメーション表示による戦術解析支援システムの開発⁶⁾ 等が挙げられる。

一方、テニスについては、これまでスイングやフォームの解析しかなされておらず、選手やボールの位置を自動検出する研究は行われていない^{7) 8)}。本研究は、テニスの試合における選手とボールの位置を自動検出することにより、テニスのスキルを定量評価しようとするものである。定量評価することにより、次に述べる効果が期待できる。(1)スキルが定量的に解明されることにより、選手の上達が早くなる。(2)相手のプレーを定量解析することにより、より正確な作戦が立てられる。(3)テニス専門家がなくても安価にプレー解析・評価ができる。

本研究は、試合を解析し、上級者と初心者でのプレーの違いなどを定量的に評価し、練習メニューや戦術を立てる支援をするシステムの構築を目的としている。今回は、シングルの試合を対象とし、テニスコートでのボールと選手の 3 次元位置検出手法を提案する。さらに検出結果から、勝敗のポイントとなるシーンについて、そのスキルを定量的に解析する。また、上級者同士の試合と初心者同士の試合の検出結果から、スキルの違いがどこに現れているかを検証した。

本論文では、まず、ボールと選手の位置検出アルゴリズムについて説明する。つぎに、カメラパラメータと 3 次元位置の算出式について述べる。実験結果として、ボールと選手の軌道を求めた結果を示し、上級者と初心者でのプレーの違いを定量評価する。

2. 位置検出方法

2.1 ボール位置検出

カメラ画像からのボールの位置検出には差画像を使用する。2 種類のコートで撮影した原画像を図 1 に示す。(a)はベースライン側から撮影したものであり、(b)はサイドライン側からの撮影である。カメラは固定されているため、差画像を使用することにより動きのある、選手とボールが抽出される。これを図 2 に示す。上方のウィンドウの中にあるのがボールであり、左下は選手の差画像出力である。差画像の作成には、1 つ前のフレームとの差を取る。これはボールの速度が選手の移動速度より速いことを利用するもので、選手の動きの差が最小限となり、差画像で選手の抽出される部分が最も少なくなる。したがって、ボールが選手の近くにあるとき、選手をボールとして検出してしまいう誤検出を最小限に抑えている。

差画像からボールを検出するために、前のフレームでの移動量をもとに予測ウィンドウを設ける。図 3 において、2 フレーム前のボールの座標 P_2 と 1 フレーム前のボールの座標 P_1 をもとに x 座標と y 座標それぞれの移動量を求める。求めた移動量を 1 フレーム前のボール座標に加算し算出された座標位置を、現在のフレームでのボールの予測位置点 P_0 とする。その点を中心にウィンドウを設けボールの座標抽出を行う。予測ウィンドウは、始めに、 2×6 (画素)の大きさで設置する。これでボールが検出できない場合、ウィンドウを拡大していく。

Fig.1 Original images

* 原稿受付 平成 16 年 3 月 26 日

** 正会員 広島工業大学 (広島市佐伯区三宅 2-1-1)

*** 広島工業大学 (現 日立ハイブリッドネットワーク株式会社; 横浜市戸塚区戸塚町 180)

Fig.2 Detection of a ball by subtraction of sequential images

Fig.3 Prediction of a ball position

この予測ウィンドウでは、ウィンドウ内に少しでもボールが入ると、ウィンドウ内に入っている部分のみをボールとして検出する。そのため、このままでは正確なボールの重心位置を検出することができない。そこで、予測ウィンドウで検出した座標を中心として、再びウィンドウを設ける。これにより、ウィンドウ内に完全な形でボールが納まり、正確なボールの重心位置を検出することが可能となる。

2.2 選手の位置検出

選手の位置は足位置と定義する。足位置検出方法は、まず、選手の体の重心位置を求める。理由として、直接足の検出を行うのは困難であるが、体を検出することは比較的容易に行うことができるためである。体より上に足が存在することは無いので、体の重心位置より下に検出用ウィンドウを設置し足の検出を行う。以下に、具体的な検出方法を述べる。

サイドライン側にカメラを設置した場合を例とする。今回は、シングルスを対象としたので、左右のコートに一人ずつ選手が存在する。処理は左右別々に行う。

試合は、サーブから始まる。そのため、試合開始直後は左右どちらのコートにいる選手もベースライン付近にいる。そこで1, 2フレーム目はベースライン付近に処理ウィンドウを設置する。テニスでは選手は白のウェアを着ていることが多いので、原画像（濃淡画像）を適当な閾値で2値化処理すれば選手が抽出できる。しかし、2値化結果は図4(a)に示すように、選手と白線の両方を抽出する。コートの白線の消去、及び、選手の体が切れている場合それらをつなぐために、縮小・拡大を行い、最も面積の大きなものを選手と定義し、体の重心位置を求める。この結果を図4(b)に示す。テニスの動きでは足が体より上に来ることは無いので、重心位置より下に図5に示すように、処理ウィンドウを設置する。設置したウィンドウで、投影処理を行い、以下の式で示す位置を選手の足位置座標 (x_a, y_a) とする。

$$\begin{aligned} x_a &= (x_1 + x_2) / 2 \\ y_a &= y_2 \end{aligned} \quad (1)$$

(a) Binary image (b) Elimination of coat lines

Fig.4 Detection of a player

Fig.5 Detection of foot position

ここで、 y 方向については両足が分離できない場合もあるため、足ウィンドウの中で投影処理を行い、最も下の位置を足位置 y 座標としている。

3. 3次元位置算出方法

3.1 カメラパラメータの算出

カメラパラメータの算出方法として、まず、カメラ画像から3次元位置が既知である6点を基準点とすると、式(2)が成立する¹⁰⁾。

$$A \cdot C = R \quad (2)$$

ここで、 C は、11個の要素からなるカメラパラメータ行列である。但し、

$$A = \begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -X_1x_{c1} & -Y_1x_{c1} & -Z_1x_{c1} \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -X_1y_{c1} & -Y_1y_{c1} & -Z_1y_{c1} \\ X_2 & Y_2 & Z_2 & 1 & 0 & 0 & 0 & 0 & -X_2x_{c2} & -Y_2x_{c2} & -Z_2x_{c2} \\ 0 & 0 & 0 & 0 & X_2 & Y_2 & Z_2 & 1 & -X_2y_{c2} & -Y_2y_{c2} & -Z_2y_{c2} \\ X_3 & Y_3 & Z_3 & 1 & 0 & 0 & 0 & 0 & -X_3x_{c3} & -Y_3x_{c3} & -Z_3x_{c3} \\ 0 & 0 & 0 & 0 & X_3 & Y_3 & Z_3 & 1 & -X_3y_{c3} & -Y_3y_{c3} & -Z_3y_{c3} \\ X_4 & Y_4 & Z_4 & 1 & 0 & 0 & 0 & 0 & -X_4x_{c4} & -Y_4x_{c4} & -Z_4x_{c4} \\ 0 & 0 & 0 & 0 & X_4 & Y_4 & Z_4 & 1 & -X_4y_{c4} & -Y_4y_{c4} & -Z_4y_{c4} \\ X_5 & Y_5 & Z_5 & 1 & 0 & 0 & 0 & 0 & -X_5x_{c5} & -Y_5x_{c5} & -Z_5x_{c5} \\ 0 & 0 & 0 & 0 & X_5 & Y_5 & Z_5 & 1 & -X_5y_{c5} & -Y_5y_{c5} & -Z_5y_{c5} \\ X_6 & Y_6 & Z_6 & 1 & 0 & 0 & 0 & 0 & -X_6x_{c6} & -Y_6x_{c6} & -Z_6x_{c6} \\ 0 & 0 & 0 & 0 & X_6 & Y_6 & Z_6 & 1 & -X_6y_{c6} & -Y_6y_{c6} & -Z_6y_{c6} \end{pmatrix}$$

$$C = [C_{11} \ C_{12} \ C_{13} \ C_{14} \ C_{21} \ C_{22} \ C_{23} \ C_{24} \ C_{31} \ C_{32} \ C_{33}]^t$$

$$R = [x_{c1} \ y_{c1} \ x_{c2} \ y_{c2} \ x_{c3} \ y_{c3} \ x_{c4} \ y_{c4} \ x_{c5} \ y_{c5} \ x_{c6} \ y_{c6}]^t$$

(X_i, Y_i, Z_i) : 3次元位置
 (x_{ci}, y_{ci}) : カメラ座標 $i = 1 \sim 6$
 t は転置を意味する。

式(2)よりカメラパラメータ C を求める式は,

$$C = (A^T A)^{-1} A^T R \quad (3)$$

となる。もう1台のカメラについても同様にして、カメラパラメータを求める。

今回実験に使用した動画での基準点6点は、図6に示すとおりである。A点を(0, 0, 0)と置き、これを基点として、残りの点をテニスコートの実測値に沿って与える。B～点までの値は、B(10.97, 0, 0), C(9.60, 2.00, 5.49), D(1.37, 2.00, 18.28), E(10.97, 0, 23.77), F(0, 0, 23.77)である。C, D点は、事前に立てておいた2mのスタンドの頂点とした。

3.2 3次元位置の算出

3.1で求めたカメラパラメータを用いて、3次元位置を算出する。以下にその式を示す。

$$\begin{aligned} (C_{11} - C_{31}x_c)X + (C_{12} - C_{32}x_c)Y + (C_{13} - C_{33}x_c)Z &= x_c - C_{14} \\ (C_{21} - C_{31}x_c)X + (C_{22} - C_{32}x_c)Y + (C_{23} - C_{33}x_c)Z &= y_c - C_{24} \\ (P_{11} - P_{31}x_p)X + (P_{12} - P_{32}x_p)Y + (P_{13} - P_{33}x_p)Z &= x_p - P_{14} \end{aligned} \quad (4)$$

但し、 C : 左カメラのカメラパラメータ

P : 右カメラのカメラパラメータ

(x_c, y_c) : 左カメラのカメラ座標

(x_p, y_p) : 右カメラのカメラ座標

以上の連立方程式から任意のカメラ座標に対する3次元位置 (X, Y, Z) が求まる。

4. 実験

4.1 実験方法

今回は、シングルの試合を実験の対象とし、2台のカメラを用いた。ボールがぶれて撮影されないように、シャッタースピードは1/10000に設定した。上級者同士の試合は大学のコートで撮影したため、カメラの設置位置の制限からベースライン方向から撮影した。初心者同士の試合は、市営のコートで撮影しカメラ位置の制限がなかったため、サイドライン側から撮影した。撮影したビデオテープからパソコン上に30フレーム/秒の720×480画素のAVIファイルとして動画をキャプチャーした。2台のカメラからの動画の同期は、サーブで手からボールが離れた瞬間やボールが地面についた瞬間などを基準とした。作成した動画から、ボールと選手の位置座標を検出し、式(4)より3次元位置の算出を行った。

Fig.6 Calibration points for the 3D conversion

実験では、上級者同士の動画からスキルがよく現れているシーンを3種、初心者同士の動画からは2種、合計5種のシーンを定量評価したが、本論文では、それぞれ1種類ずつ紹介する。本論文で紹介する動画のフレーム数は、上級者同士の動画は199フレーム、初心者同士の動画は397フレームからなる。

4.2 実験結果

算出したボールと選手の3次元位置を図7, 図8に示す。(a)は、横方向Xに対する奥行きZを表したグラフである。(b)は、奥行きZに対する高さYを表したグラフである。図7は上級者同士、図8は初心者同士の試合である。これ以降、必要に応じて図7をシーン1、図8をシーン2と呼ぶ。

まず、シーン1について動画と比較すると、図7(a)より、ボールと選手の軌道に大きな誤りは無く、幅方向、奥行き方向は良好な結果であると言える。また、23m付近でボールの軌道が途切れているのは、ボールが奥選手の服と重なり検出できなかったことが原因である。図7(b)より、17m付近でボールの軌道に大きなブレが生じている。この原因としては、ボールを打ち返すシーンでは、ラケットや人をボールとして誤検出してしまうためである。選手に関しては、片方のカメラでは検出できる足が、影の映りなどにより、もう片方のカメラでは検出できないことによる対応点の不一致から生じる計算による誤差が見られた。

シーン2について動画と比較すると、図8(a), (b)ともに大きな誤差は無く、良好な結果が得られた。これは、サイドライン側にカメラを設置したことにより、テニスコートを横長に大きく撮影することが可能となり、それが結果に表れたと考えられる。

表1は、画面上でボール位置を手動検出したものと自動検出したものの誤差である。サイドライン側から撮影したシーン2は、ベースライン側から撮影したシーン1に比べ、幅方向Xに対する誤差が3.5cm増加しているものの、奥行きZに関しては、6.8cmの減少が見られた。また、X, Y, Zそれぞれの誤差で大きなばらつきが無いことから、サイドライン側からの撮影が有効であると言える。

5. スキルの解析

5.1 勝敗の解析

まず、それぞれのシーンでの勝敗の原因となったプレーを実験結果から検証する。図9(a)は、シーン1で時間的に変化するボールと選手の奥行き方向の位置を、(b)は幅方向の位置をグラフ化したものであり、(c)はボールの高さを示している。このシーンは、奥選手が打ったボールが手前選手の左横を抜けて、手前選手が打ち返すことができず、奥選手が勝利したシーンである。図9(b)を見ると、160フレームあたりで一瞬1m程、手前選手がボールと反対の右方向へ動いたことがわかる。その後、ボールの方向に動くことができず、選手の動きが止まっていることが図9(b)から分かる。これが要因である。

シーン2は、何度かラリーが続いた後、まず手前選手が短めのショットで攻撃を仕掛けたが、奥選手がそれをコート手前に返し、手前選手は、それを上手く返すことができなかった奥選手が勝利した映像である。図8(a)より、ラリーはほぼ同じコースで打ち合い、勝敗を決めたショットは、ラリーのコースから外した、選手から遠い場所に打ったことが分かる。

Table 1 Error of ball position detection (unit: cm)

	X	Y	Z
Scene 1 (View from end line)	2.0	2.4	9.3
Scene 2 (View from side line)	5.5	2.7	2.5

Fig.7 Detection results of 3-dimensional positions (Experts)

5.2 スキルの検証

5.2.1 選手のポジション

テニス経験者は、サーブに対してベースライン際に対応する。図7(a)より、手前選手のサーブに対して奥選手はベースライン際にポジションを取り、サーブに対応している。しかし、図8(a)より、初心者の奥選手はベースラインより前にポジション取りをしていることが分かる。ポジションが前過ぎると、ボールを打ち難い高い位置で打たなければならなくなり、フォームが崩れ、ミスショットの原因となる。例えば、図8(b)より、奥選手(身長167cm)は、約145cmと高い位置で打っている。このことから、この選手がポジション取りに失敗していることが判断できる。上級者は、リターン後もベースライン際にポジションを取り、ボールの動向を見守って行動に移っていることが分かる。ベースライン際にポジションを取ることは、打球がアウトかどうかを見定めることができるなどの利点がある。また、ボレーに出るときは、しっかりと前に出ている。それに対して、初心者の奥選手はその後も立っている位置は変わらず、打ち難い高い位置でボールを打ち返している。手前選手は初心者であるが、軟式テニス経験者であり、ベースライン際にポジションを取っている。

5.2.2 打球の高さ

Fig.8 Detection results of 3-dimensional positions (Beginners)

図8(b)より初心者同士の試合はラリーでボールの高さが3.5mを超えるような場合があるなど、打球が高いことが分かる。特にサーブではその特徴が顕著に表れている。図7(b)より、上級者のサーブはネットぎりぎりを通り直線的な軌道を描いているのに対し、初心者のサーブは図8(b)を見ると、ネットから2mも高いところを通り、軌道も緩やかな弧を描く打ち易いものとなっている。

5.2.3 打球のスピード

図10は打球のスピードを計算したものである。それぞれボールを打った瞬間から2フレーム後の距離をもとに時速を算出したものである。但し、上級者のショット1の時速に関しては、打った瞬間が検出できなかったため検出できたところから2フレーム後の距離をもとに時速を算出した。図10よりまず目に付くのが、サーブの速度の違いである。上級者と初心者では、およそ100km/hと大きな差がある。サーブの軌道の違いについてはすでに述べたが、サーブは最も分かりやすく上級者と初心者の違いが出る部分であると考えられる。また、ショットに関しては、初心者は上級者に比べスピードが遅く、どのショットもほとんど速度は変わらない。一方、上級者は相手の速いショット(ショット1:125km/h)をボレーで勢いを殺してスピードの遅いボールを返す(ショット2:35km/h)といった緩急をつけた試合を行っている。

(a) Z direction move of players and a ball

(b) X direction move of players and a ball

(c) Ball height

Fig.9 Quantitative evaluation of a play

6. 結 言

2台のカメラでコート全体を撮影したステレオ画像からボールと選手の3次元位置を検出した。そこから、テニスの試合の定量評価を行い、試合の敗因と上級者と初心者のプレーの違いを検証した。

(a) Expert

(b) Beginner

Fig.10 Comparison of ball speeds

- (1) ボール位置予測、白線の除去などを行い、テニスコートにおけるボールと選手の位置を検出する手法を開発した。
- (2) 上級者と初級者の試合の動画から、スキルと関係すると考えられるシーンを5種選び、ボールと選手の軌道を算出した。
- (3) 上級者と初心者の、(i) ボールと選手の軌跡、(ii) 選手のポジション、(iii) 打球の高さ、(iv) ボールのスピードを解析し、スキルを定量評価した。

今後は、動画からの上記スキルの自動抽出、実時間解析システムの構築を目指す。

参 考 文 献

- 1) 宮原秀夫, 鳴尾文司, 越後富夫, 片山宗臣, 宇野洋二: ITとスポーツ特別小特集, 電子情報通信学会論文誌 D-II, **J85-D-II**, 1 (2002)1.
- 2) 吉田貴夫, 小沢真治: サッカー中継のシーン解析, 第2回画像センシングシンポジウム講演論文集(1996)183.
- 3) 瀧岡志, 長谷川純一: チームスポーツにおける集団行動解析のための特徴量とその応用, 電子情報通信学会論文誌 D-II, **J81-D-II**, 8 (1998)1802.
- 4) 浅野敏郎, 宮下雄一郎, 平松久佳, 玉野和保: Fuzzy-SOMを用いた自律行動体の行動様式の類似性, 精密工学会誌, **68**, 9 (2002)1190.
- 5) 松本圭介, 須藤智, 斎藤英雄, 小沢真治: サッカー放込における視点選択のための多視点画像の統合によるボール追跡, 電気学会論文誌 C, **121-C**, 10 (2001)1530.
- 6) 西田絵美子, 久本賢, ラシキア・ジョージ: サッカー戦術解析支援システムの開発, 第8回画像センシングシンポジウム講演論文集(2002)203.
- 7) 山和淳司, 大谷淳, 石井健一郎: 隠れマルコフモデルを用いた動画画像からの人物の行動認識, 電子情報通信学会論文誌 D-II, **J76-D-II**, 12 (1993)2556.
- 8) Joo Kooi TAN, Masako KOGA, Seiji ISHIKAWA: A Human Recovery System Based on Multiple Video Cameras, 第4回画像センシングシンポジウム講演論文集(1998)127.
- 9) 宮森恒: 映像と音響情報の協調による内容検索のためのテニス動作自動注釈付け, 電子情報通信学会論文誌 D-II, **J86-D-II**, 4 (2003)511.
- 10) 井口征士, 佐藤宏介: 三次元画像計測, 昭晃堂 (1990)91.