

単眼サッカー映像におけるボールの3次元位置情報を用いた状況認識

西濃 拓郎[†] 滝口 哲也^{††} 有木 康雄^{††}

[†] 神戸大学大学院工学研究科 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

^{††} 神戸大学自然科学系先端融合研究環 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

E-mail: [†]fantakuro@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

あらまし 本稿では、サッカーの自動映像生成を目的として、高精度かつ安定にボールを追跡し、その3次元位置情報を用いて試合のイベント検出を行うシステムを提案する。追跡手法として、9次元の状態ベクトルを持つ3次元パーティクルフィルタを用いている。パーティクルフィルタを用いた追跡は局所探索であるため、一度見失うと追跡状態を回復できない。そこで本研究では、3次元の全画面探索と切り替えることでこれを改善し、さらに座標補間を行うことによって、より精度の高い追跡を行うことができるようになった。この結果、追跡精度は約15%向上し、ゴールやゴールキックといったイベントも高精度で検出できるようになった。

キーワード サッカー, ボールトラッキング, 3次元パーティクルフィルタ, 座標補間, イベント検出

Situation Recognition Using 3D Positional Information of Ball from Monocular Soccer Image Sequence

Takuro NISHINO[†], Tetsuya TAKIGUCHI^{††}, and Yasuo ARIKI^{††}

[†] Graduate School of Science and Technology, Kobe University

1-1, Rokkodai, Nada, Kobe, Hyogo, 657-8501 Japan

^{††} Organization of Advanced Science and Technology, Kobe University

1-1, Rokkodai, Nada, Kobe, Hyogo, 657-8501 Japan

E-mail: [†]fantakuro@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

Abstract In this paper, we propose a system that tracks a ball stability and accurately and detects the events of the game by using the 3D location informations for automatic soccer video production. We use , 3D particle filter with the state vector of nine dimensions in the ball tracking. As the ball tracking by particle filter is a local search, it is difficult to continue the tracking when it fails. Then, we solve this problem by switching the local search to 3D global search, and by interpolating the lost coordinates. As a result, the tracking accuracy was improved by about 15%, and the events like the goal and the goal kick can be detected with high accuracy.

Key words Soccer, ball tracking, 3D particle filter, coordinates interpolation, event detection

1. はじめに

スポーツの試合を撮影する場合や編集する場合、単に撮影するのではなく、試合の流れやイベントが分かりやすく面白い映像を撮影・編集することが要求される。これをコンピュータで自動的に行う技術として、自動映像生成技術がある。自動映像生成技術の最も簡単なものとして、デジタルカメラワークがある。これは、スポーツの試合において、フィールド全体を一つの高解像度カメラで撮影し、その映像をクリッピングし、連続的につなぎ合わせることによって新しい映像を作成する技術である。この技術によって、試合の内容が分かりやすく面白くかつ、視聴者の嗜好に即した自由な編集が可能とな

る [15]。

映像をクリッピングするためには、その映像の内容を理解し構造化する必要がある。すなわち、そのスポーツ映像におけるイベントを検出し、その種類に基づいて映像を編集する必要がある。しかし、イベントを検出するためには、ボールと選手の位置を正確に追跡しておく必要がある。追跡手法として、mean-shift [7]、カルマンフィルタ [3]、共分散トラッカ [6]、パーティクルフィルタ [1] など、多くの手法が提案されている。特に、パーティクルフィルタはオクルージョンに強いことから、ボールや選手の位置を追跡する研究において数多く用いられてきた。2次元画面上でのボール追跡にパーティクルフィルタを用いる研究 [4] があるが、この手法では、イベント

検出における重要な特徴量であるボールの3次元位置がわからないという問題がある。3次元追跡にパーティクルフィルタを用いている研究 [5] も行われているが、一度ボールを見失うと再度発見するのが困難であるという問題が解決されておらず、長時間の追跡ができないという問題がある。また、複数のカメラを用いてより正確な位置を追跡しようとする研究も行われている。本研究では、アマチュアの低コストな撮影・編集環境を想定し、単眼動画を処理対象としている。

本研究は、サッカーの試合における自動映像生成を目的として、高精度かつ安定にボールを追跡し、その3次元位置情報を用いて試合のイベント検出を行うシステムを提案する。具体的には、パーティクルフィルタを用いて、3次元フィールド座標系でボールを追跡し、スローインやゴールキックといったイベントを検出する。また、ボールを見失った際に、2次元画面座標系で全画面探索を行い、それを3次元ワールド座標系に逆投影することで、3次元追跡状態を回復するシステムと、推定結果の評価値が低い3次元座標を、時間的に前後の座標から補間するシステムを提案する。これにより、より精度の高い3次元位置情報を得ることができる。

2. 提案システム

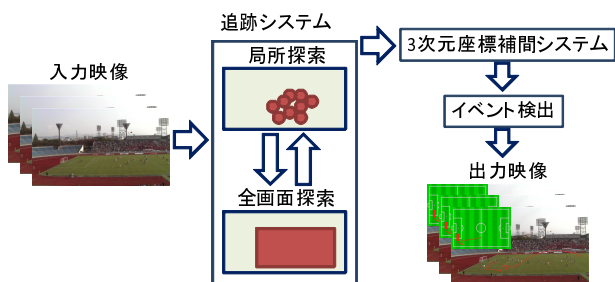


図1 提案システムの概略

図1に提案システムの概略を示す。まず、追跡システム部では3次元パーティクルフィルタを用いた局所探索と、テンプレートマッチングを用いた全画面探索を切り替えることによってボールの追跡を行う。得られた座標を3次元座標補間システムによって補間する。さらに、補間後の3次元位置情報を用いてイベントの検出を行い、フィールドの上面図を付加した映像を出力する。

次に、システムの各部の処理内容について述べる。

3. 局所探索

3.1 アルゴリズム

ボールの追跡においては、その不規則な動きやオクルージョンが問題となる。そのような状況でもロバストな追跡ができる手法として、パーティクルフィルタ (Condensation 法) が提案され、その有効性が報告されている [1]。これは、状態量と尤度を持つ多数の粒子に

よって確率分布を離散的に近似し、確率モデルを伝搬させることによって、安定した精度でボールの追跡を可能にする手法である。

時刻 t における追跡対象の状態量を x_t とし、同時刻の画像特徴を z_t とする。また、それぞれの時系列を $X_T = \{x_1, \dots, x_T\}$, $Z_T = \{z_1, \dots, z_T\}$ とする。アルゴリズムの概略を図2に示す。

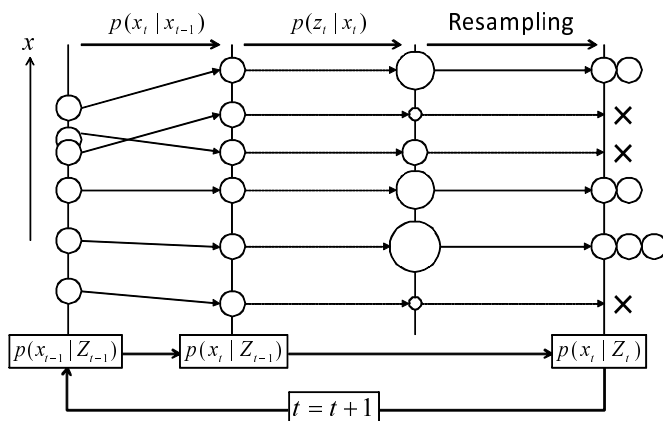


図2 パーティクルフィルタのアルゴリズム

時刻 $t-1$ の粒子の分布 $p(x_{t-1}|Z_{t-1})$ に対して、遷移確率 $p(x_t|x_{t-1})$ をもって状態を更新する。得られた予測粒子の分布 $p(x_t|Z_{t-1})$ の時刻 t におけるデータへの適合度、つまり粒子の尤度 $p(z_t|x_t)$ を求め、全体の尤度で規格化してリサンプリングする。得られた粒子の分布 $p(x_t|Z_t)$ が、時刻 t におけるフィルタ粒子の分布となる。パーティクルフィルタでは各粒子の出現率を等しくしているので、リサンプリングの作業が必要となる。尤度が低いと図2のように死滅してしまう可能性が高くなり、尤度が高いと分裂して仲間を増やす。この操作を繰り返す処理がパーティクルフィルタである。

3.2 状態空間

本システムでは、ボールの複雑な3次元運動を推定することが目的であるため、サッカー場に固定した座標系 (以下、ワールド座標系) におけるボールの3次元位置、速度および加速度からなる9次元の空間を状態空間としている。また、パーティクルフィルタの粒子の総数を N 個とし、時刻 t における第 n 番目の粒子の状態ベクトル $x^{(n)}(t)$ を式 (1) のように定義する。以降、誤解のない範囲内で、 (t) または上付きの (n) を省略することがある。

$$\begin{aligned} x^{(n)} &= [p_x^{(n)}, p_y^{(n)}, p_z^{(n)}, v_x^{(n)}, v_y^{(n)}, v_z^{(n)}, a_x^{(n)}, a_y^{(n)}, a_z^{(n)}]^T \\ &= [(\mathbf{p}^{(n)})^T, (\mathbf{v}^{(n)})^T, (\mathbf{a}^{(n)})^T]^T \end{aligned} \quad (1)$$

ここで、 $p_{x,y,z}^{(n)}$, $v_{x,y,z}^{(n)}$ および $a_{x,y,z}^{(n)}$ は、それぞれワールド座標系における位置、速度、加速度成分である。

3.3 状態遷移モデル

ボールの運動はその高さによってつぎの3つの状態に分類できる．それぞれについて，時刻 $t - \alpha$ から t へと粒子の状態を遷移させるフィルタを定式化する．

(1) 空中を飛んでいるとき

ボールは理想的な放物線運動を行うと仮定する．ただし I および O はそれぞれ 3×3 の単位行列および零行列である．

$$\phi_{high}(x) = \begin{bmatrix} I & \alpha I & (\alpha^2/2)I \\ O & I & \alpha I \\ O & O & I \end{bmatrix} x \quad (2)$$

(2) 地上を転がっているとき

ボールには摩擦力がかかり，等加速度直線運動を行うと仮定する．ただし動摩擦係数を μ ，重力加速度を g とする．

$$\phi_{ground}(x) = x + \left[0, \dots, 0, \frac{[-v_x, -v_y]}{\sqrt{v_x^2 + v_y^2}} \mu g, 0 \right]^T \quad (3)$$

(3) 空中から地上へ落ちた瞬間

ボールは地上でバウンドする．ただし反発係数を e とする．

$$\phi_{bound}(x) = [p_x, p_y, p_z, v_x, v_y, -ev_z, a_x, a_y, a_z]^T \quad (4)$$

さらに，状態ベクトルの各成分が取り得る値にリミット ϕ_{limit} を設けることによって，ボールの位置が場外に出るなどといった追跡対象外の状態や，物理的にあり得ない速度の粒子が生ずることを防いでいる．

以上の4つのフィルタとガウスノイズ ω によって，遷移モデルを次の式 (5) のように定式化する．なお \oplus は排他的論理和， \circ は合成関数を表わす．

$$x(t) = ((\phi_{high} \oplus \phi_{ground} \oplus \phi_{bound}) \circ \phi_{limit})x(t - \alpha) + \omega(t - \alpha) \quad (5)$$

3.4 尤度評価

状態ベクトル x の3次元位置成分 p を2次元画面座標に投影し，その座標を中心とした領域とテンプレート画像との正規化相互相関値を尤度として計算する．

3.4.1 透視投影変換

座標投影モデルとして，ピンホールモデルによる透視投影変換を用いる．投影変換のためには，カメラキャリブレーションが必要となる．

キャリブレーションの手法として Tsai の手法 [2] を用いている．Tsai のキャリブレーション手法とは，図3のように多数の世界座標が既知の点 (X_w, Y_w, Z_w) と，

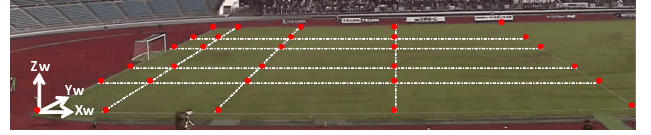


図3 ワールド座標の既知な点

それぞれの点に対応する画面座標 (x_f, y_f) の組を与えたときに，外部パラメータとして回転行列 R (未知数3個) と平行移動ベクトル T (未知数3個)，内部パラメータとして焦点距離 f ，レンズの歪曲収差係数 k ，スケール要素 s_x, s_y ，画像原点 (C_x, C_y) の12個のパラメータを非線形最適化によって求める手法である．サッカーのフィールドは，一般的な大きさが規定されており，フィールド内のペナルティエリアやゴールエリアなどのラインの長さも定められているため，ワールド座標が既知な点を取りやすいという利点をもっている．

キャリブレーションによって得られたパラメータと以下の式を用いて，ワールド座標 (X_w, Y_w, Z_w) を画面座標 (x_f, y_f) に変換する座標変換を行う．

まず，外部パラメータの R と T を用いてカメラ座標 (X_c, Y_c, Z_c) を求める．

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = R \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + T \quad (6)$$

次に，焦点距離 f を用いて，理想的な画面座標 (x_u, y_u) を求める．

$$x_u = f \frac{X_c}{Z_c} \quad (7)$$

$$y_u = f \frac{Y_c}{Z_c} \quad (8)$$

レンズの半径方向の歪みを考慮した座標 (x_d, y_d) は，

$$x_u = x_d(1 + kr^2) \quad (9)$$

$$y_u = y_d(1 + kr^2) \quad (10)$$

$$r = \sqrt{x_d^2 + y_d^2} \quad (11)$$

最後に，撮影機器に固有の値である x, y 方向の CCD 素子の間隔 d_x と d_y を利用して，画面座標 (x_f, y_f) を導出する．

$$x_f = \frac{s_x x_d}{d_x} + C_x \quad (12)$$

$$y_f = \frac{s_y y_d}{d_y} + C_y \quad (13)$$

3.4.2 正規化相互相関

粒子の3次元座標を透視投影変換して2次元座標を求め，その座標を中心とした領域とテンプレート画像との間で正規化相互相関値を求め，これを粒子の尤度とする．探索領域内の座標 (x, y) における輝度値を $I(x, y)$ と表し，同様にテンプレート画像内の座標 (x, y) における輝

度値を $T(x, y)$ としたとき、相関値 R は次式 (14) のように表わされる。ただし、 \bar{I} と \bar{T} はそれぞれ探索領域とテンプレート画像の平均輝度値とする。

$$R = \frac{\sum_{i,j} \{I(i, j) - \bar{I}\} \cdot \{T(i, j) - \bar{T}\}}{\sqrt{\sum_{i,j} \{I(i, j) - \bar{I}\}^2 \cdot \sum_{i,j} \{T(i, j) - \bar{T}\}^2}} \quad (14)$$

なお、本研究で対象とした画像データはカラー画像であるため、探索領域、テンプレート画像を赤、緑、青の3色に分解し、それぞれのカラー画像に対して式 (14) を適用する。それぞれのカラー画像の相関値を R_R, R_G, R_B とすると、式 (15) のように最終的な相関値 R' を定義する。

$$R' = R_R \times R_G \times R_B \quad (15)$$

ただし、いずれかのカラー画像の相関値が負である場合、その他のカラー画像の相関値に関わらず、 $R' = 0$ とする。これは、2つのカラー画像の相関値が低くなって負となったとき、その積によって R' が正の値になってしまうのを防ぐためである。

4. 全画面探索

パーティクルフィルタは粒子の存在する領域のみを探索する局所探索であり、短時間であればボールのオクルージョンに対しても頑健な追跡が可能である。しかし、一度ボールを見失ってしまった場合には、粒子がその位置に止まり続けてしまうため、再び追跡を始めるのが困難になるという欠点をもっている。本システムではこの問題に対して、テンプレートマッチングによる全画面探索を行うことによって改善する。

全画面探索では、一定の尤度を超えるボールの対象領域を検出した場合、その領域を物ボール領域とみなす。複数の対象領域の候補が検出された場合には、その中でもっとも尤度の高い領域をボール領域とする。次に、その領域の中心座標をボールの画面座標として、次のフレームからは3次元局所探索に切り替える。もし、一定の尤度を超える領域を検出できなかった場合には、次のフレームにおいても全画面探索を継続する。

4.1 動的探索領域

探索領域が狭すぎるとボールが領域外に存在した場合に検出できないが、逆に広すぎると別の物体をボールと誤認識してしまう確率が高くなり、処理時間も長くなってしまう。このトレードオフの関係における最適な探索領域を決定することが重要となる。本研究では、動的に探索領域を決定するための特徴量として、ボールを見失った時点での速度を利用する。図4は時刻 t でボールを見失った場合に、時刻 $t+k$ での探索領域を表わしている。

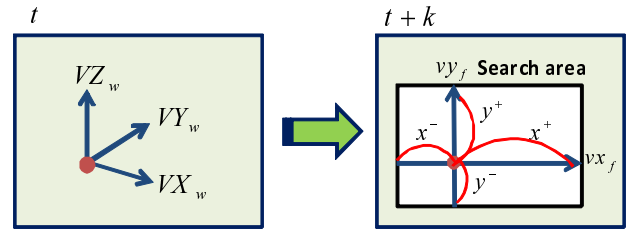


図4 探索領域の動的確保

具体的な探索領域の算出方法を述べる。まず、ワールド座標系と画面座標系における速度をそれぞれワールド速度、画面速度とする。前述の透視投影変換を用いて、ワールド速度 (VX_w, VY_w, VZ_w) を画面速度 (vx_f, vy_f) に変換する。ベースとなる探索領域を $30k \times 30k$ 画素とし、この領域を時刻 t での画面速度 (vx_f, vy_f) の割合に応じて変化させる。図4中の x^+, y^+, x^-, y^- は次式で定義する。ただし、 vx_f, vy_f の正の方向をそれぞれ x^+, y^+ とし、負の方向を x^-, y^- とする。

$$x^+ = \frac{|vx_f|}{vx_f + vy_f} (20k) \quad (16)$$

$$y^+ = \frac{|vy_f|}{vx_f + vy_f} (20k) \quad (17)$$

$$x^- = \frac{|vx_f|}{vx_f + vy_f} (10k) \quad (18)$$

$$y^- = \frac{|vy_f|}{vx_f + vy_f} (10k) \quad (19)$$

最後に、 x 座標は x^+ から x^- まで、 y 座標は y^+ から y^- までの領域を探索領域とする。

4.2 逆透視投影変換

全画面探索によってボールを検出できた場合に、画面座標が得られる。次のフレームから3次元局所探索に切り替えるためには、この2次元画面座標を3次元ワールド座標へと変換しなければならない。これは、前述の透視投影変換の逆変換であるといえる。逆透視投影変換では画面座標 (x_f, y_f) をワールド座標 (X_w, Y_w, Z_w) に変換するため、そのままでは2変数から3変数への変換になってしまう。そこで本研究では、見失ったフレームと再発見したフレームではボールの高さが変化していないと仮定して、式 (13) から (6) を逆に計算することでワールド座標を求めている。

5. 座標補間

サッカーの試合において、ボールは不規則な動きをするためオクルージョンや、画面外への移動などが頻繁に起こる。このため、追跡に失敗する区間が存在する。しかし、イベント検出を行うためには、ボールの高精度な3次元座標が必要不可欠である。そこで本研究では、追跡に失敗した区間、もしくは追跡できているが尤度の低い区間に対して、運動モデルを仮定して補間する。以後

この補間の対象となる区間のことを、未知の区間と呼ぶことにする。

5.1 短区間の補間

未知の区間が1秒にも満たないごく短い時間である場合には、摩擦力や重力、抵抗力等による加速度は無視できるため、この区間において物体は等速直線運動をしていると仮定できる。

したがって、未知の区間が時刻 $t+1$ から $t+\alpha-1$ であるとすると、 x, y, z 軸方向の速度 $\hat{v}_x, \hat{v}_y, \hat{v}_z$ はそれぞれ、次式のように求められる。

$$\hat{v}_x = \frac{p_x(t+\alpha) - p_x t}{\alpha} \quad (20)$$

$$\hat{v}_y = \frac{p_y(t+\alpha) - p_y t}{\alpha} \quad (21)$$

$$\hat{v}_z = \frac{p_z(t+\alpha) - p_z t}{\alpha} \quad (22)$$

以上より、未知の区間の座標は次式のように求められる。ただし $0 \leq k < \alpha$ である。

$$p_x(t+k) = p_x t + \hat{v}_x k \quad (23)$$

$$p_y(t+k) = p_y t + \hat{v}_y k \quad (24)$$

$$p_z(t+k) = p_z t + \hat{v}_z k \quad (25)$$

5.2 長区間の補間

この未知の区間は図5のようにボールが急激に飛翔したことによって観客席が背景となり、尤度が極端に下がってしまうことが原因で起こることが多い。



図5 長区間見失ってしまう場合の例

この図はゴールキックからのボールの軌跡を描画しているが、黄色の軌跡のところでは背景が観客席となってしまう、尤度が下がっている。この未知の区間における補間は、 (x, y) 座標と z 座標に分けて補間する。

まず、 (x, y) 座標の補間について考える。観客席が背景になるほど高く上がったボールは、ワールド座標系 (x, y) 平面において、ほぼ等速直線運動を行う。そのため、この補間は式(23)と式(24)と同様に扱える。

次に z 座標の補間について考える。長時間観客席が背景になってしまうのは、ボールが図5のようにワールド座標系 x 軸方向に対してほぼ平行に飛ぶときである。 x 軸に対して平行に放物線運動を行うボールは、観客席が

背景になるとき、すなわち見失う時の高さ、再発見するときの高さが等しくなる。

従って、未知の区間において z 座標は運動モデルとして鉛直投射運動と仮定できる。すなわち、未知の区間が時刻 $t+1$ から $t+\alpha-1$ であるとすると、 z 軸方向の速度 v_z と重力加速度 g を用いて未知の区間の z 座標 $p_z(t+k)$ は次式(26)のように定式化できる。ただし $0 \leq k < \alpha$ である。

$$p_z(t+k) = p_z t + v_z t k - \frac{1}{2} g k^2 \quad (26)$$

6. イベント検出

本研究では、各フレームにおいて、状態ベクトル $x(t)$ と図6で与えたフィールドの絶対座標の2つを利用して、イベントのルールを作成するとともに、そのルールを用いてイベントを検出する。取り扱ったイベントは、セットプレーと呼ばれるスローイン、ゴールキック、それとゴールである。なお、今回はボールの追跡結果のみを利用してイベントの検出を行った。

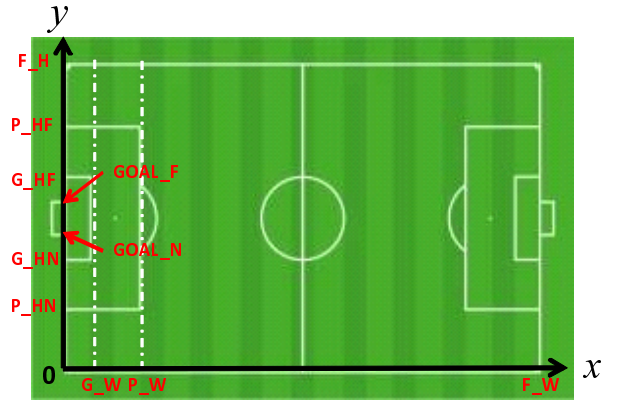


図6 フィールド座標

6.1 IF ~ THEN ルール

式(1)で定義される状態ベクトル $x(t)$ から得られる p_x と p_y に論理演算を施すことによって、表1に示す各種イベントの生起を判定するためのルールを作成する。

スローインはボールの y 座標がフィールドより小さい ($p_y < 0$)、もしくは大きい ($p_y > F_H$) 場合に生起したと判定する。

ゴールキックは、ボールの y 座標がゴール以外の位置にありフィールド外に出た場合 ($(p_x < 0) \cap ((0 < p_y < GOAL_N) \cup (GOAL_F < p_y < F_H))$)、もしくは、ボールが長時間ゴールライン上に置かれれば ($(p_x, p_y) = (G_W, G_HN \cup G_HF)$)、その時点で生起したと判定する。

ゴールは、ボールの y 座標がゴールの間にありフィールドの外に出た場合 ($(p_x < 0) \cap (GOAL_N < p_y <$

表 1 イベントの種類

Event	Parameter
スローイン	TH
ゴールキック	GK
ゴール	GL

$GOAL_F$)), またはボールがフィールドの中央に置かれれば $((p_x, p_y) = (\frac{F_W}{2}, \frac{F_H}{2}))$ 生じたと判定する.

以下に判定ルールを示す. (p_x, p_y) はボールの 3 次元ワールド座標の x, y 座標を表わしている. なお本研究では, フィールドの左半分のみ写した映像を実験として用いているため, イベントも左半分で起こるもののみを対象として検出している.

$$TH = (p_y < 0) \cup (p_y > F_H)$$

$$GK = ((p_x < 0) \cap ((0 < p_y < GOAL_N) \cup (GOAL_F < p_y < F_H)))$$

$$\cup ((p_x, p_y) = (G_W, G_HN \cup G_HF))$$

$$GL = ((p_x < 0) \cap (GOAL_N < p_y < GOAL_F))$$

$$\cup ((p_x, p_y) = (\frac{F_W}{2}, \frac{F_H}{2}))$$

7. 評価実験

本章では, 実映像を用いた実験を行い, その結果を示し, 考察を行う.

7.1 実験環境

視点固定映像は, XVID で圧縮された 1280×720 画素, フレームレート 30fps のサッカー映像を利用した, 内容は第 83 回全国高校サッカー選手権大会京都府大会決勝, 桂高等学校 (以下白チーム) 対京都朝鮮高級学校 (以下赤チーム) の試合を撮影したものである. 図 7 のように, フィールドの左半分が画面の下半分に写っている映像となっている.



図 7 実験に使った映像

テンプレート画像は以下の図 8 を用いた.

ボールはほとんど形状変化しないため, 画像は 1 枚のみ用いている.

また, キャリブレーションによって得られたカメラパラメータは以下の表 2 の通りである.



図 8 テンプレート画像

表 2 カメラパラメータ

外部パラメータ			内部パラメータ	
R	R_x	93.291679[deg]	f	3.155433[m]
	R_y	5.392349[deg]	k	3.609726e-02
	R_z	-0.296688[deg]	s_x	1.0
T	T_x	-28.381686[m]	s_y	1.0
	T_y	12.246007[m]	C_x	640[pixel]
	T_z	71.529384[m]	C_y	360[pixel]

7.2 追跡実験の結果

追跡実験は, 元映像から, 平均 330 フレームのクリッピング映像を 12 本抽出して, それぞれに対してパーティクルフィルタのみで追跡を行ったものと, 提案手法として全画面探索と座標補間を追加したものの比較実験を行った. なお, パーティクルフィルタの初期値は手動でワールド座標を与えた.

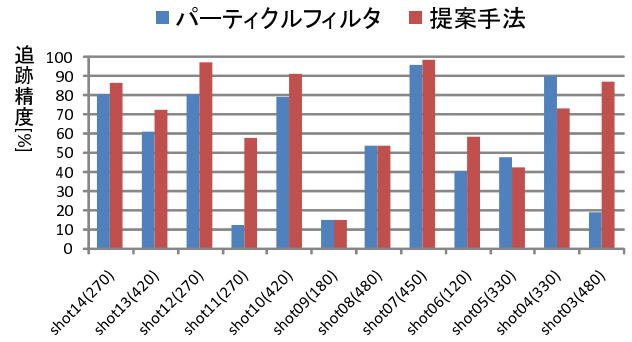


図 9 追跡精度

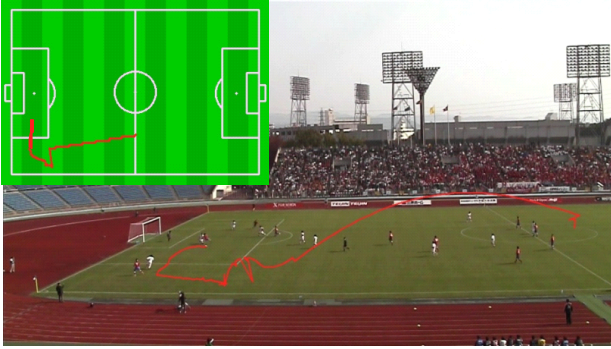
実験結果を図 9 に示す. shot の括弧内の数字はフレーム数である. 比較実験の結果, 提案手法により平均で約 15 % の追跡精度の向上が見られた. shot11 と shot03 に関しては, パーティクルフィルタのみでは完全に見失ってしまい, 追跡精度が 20 % 以下だったものが大幅に向上しているのがわかる.

shot09, shot06, shot05 のように両方共に精度が低いものは, パーティクルフィルタによる追跡の途中でボールを見失ったにもかかわらず, その時の尤度が閾値より大きかったために, 全画面探索に切り替わらなかったことが原因である.

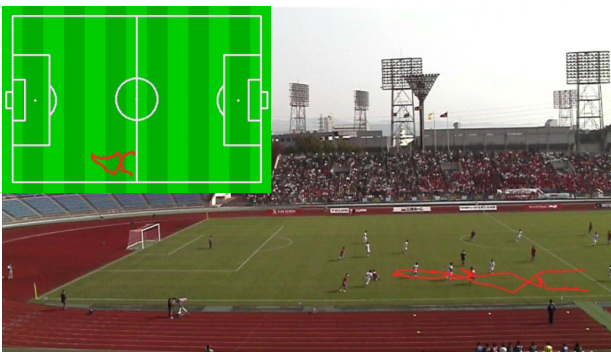
また, shot05 と shot04 のように提案手法の精度が下がってしまったものは, 共にスローインが含まれている映像であり, このときボールは選手とのオクルージョンによって見えなくなっている. 提案手法では, スローインの動作中全画面探索を行うため, 局所探索よりも精度が下がっている.

結果として試合の状況によって結果が左右されてるこ

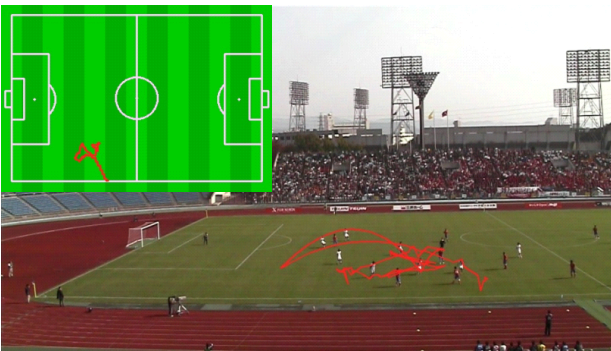
とが分かった．



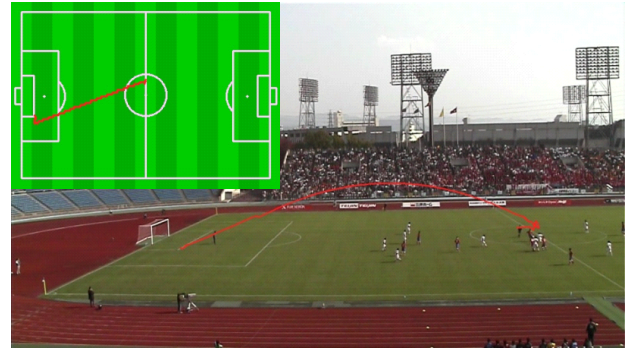
(a)



(b)



(c)



(d)

図 10 ボールの追跡結果

るため、実際の実験時間は 10 分強である．IF ~ THEN ルールに基づいてイベントを検出し、適合率と再現率を導出した．ただし、それぞれの計算式は式 (27)、(28) の通りである．

$$\text{適合率} = \frac{\text{正解したイベント数}}{\text{検出したイベント数}} \quad (27)$$

$$\text{再現率} = \frac{\text{正解したイベント数}}{\text{実際に生起しているイベント数}} \quad (28)$$

表 3 イベント検出結果

イベント	適合率	再現率
スローイン	54%	70%
ゴールキック	100%	100%
ゴール	100%	100%

実験結果を表 3 に示す．ここで、適合率及び再現率の計算に当たっては、本システムにより検出されたイベントの区間が、実試合のイベント区間と重複していれば正解とした．あるイベントに対して、比較的短時間で同判定が反復された場合は誤りとみなしていない．

ゴールキックとゴールに関しては高い精度の検出結果が得られた．これらのイベントはボールの位置が特徴的で、その継続時間も長いからであると考えられる．

一方、スローインの適合率が低いのは、タッチライン際でのボールの競り合いや、フィールドの奥側での攻防などによるボールの誤追跡が原因である．再現率の低下も同じ理由である．

8. 結 論

8.1 ま と め

本論文では、サッカーの試合における自動映像生成を目的として、ボールの 3 次元追跡手法を提案し、獲得した 3 次元位置情報を用いてイベント検出を行うことでその有効性を示した．

追跡手法として既に有効性の示されているパーティク

図 10 は追跡結果である．2 次元画面上でボールの軌跡を描いたもので、背景が観客席になっても問題なくボールの軌跡を描けていることがわかる．図中の左上の図は 3 次元ワールド座標の x, y 座標を用いて作成した上面図である．空中のボールが線形補間によって真っ直ぐ飛んでいるのがわかる（図 10 の (d)）．

7.3 イベント検出の実験結果

実験は、上記の試合の前半 20 分間を対象とした．ただし、半分近くはフィールドの左半分で試合が行われてい

ルフィルタに関して、2次元画面座標ではなく3次元ワールド座標で状態ベクトルを設定することによって、より高い精度で状況認識ができると考えた。また、パーティクルフィルタの欠点である、再発見が難しいという点を、全画面探索を行うことによって改善し、さらに、3次元ワールド座標において尤度の低かった座標を補間することで、追跡精度を向上させた。これによってボールの追跡精度に約15%の向上が見られた。

また、追跡結果の座標を利用して論理演算を行うことによりスローイン、ゴールキック、ゴールの3つのイベント検出を行った。今回は実験を1例のみで行ったため、定量評価は十分とは言えないが、提案した追跡手法から各種イベントを検出できることが示された。

8.2 今後の課題

追跡手法に関しての課題は、まず、現状では初期位置を手動で与えているため、これを自動化するシステムの構築があげられる。さらに、物理パラメータや閾値を動的に最適化するシステム、透視投影変換の際に必要なフィールドの既知の座標を自動的に検出するシステムなども考えられる。

また、ボールと同時に選手の動きを追跡することにより、さらに高度なイベント、たとえばオフサイドやパス、ドリブルなどのイベントを検出することができるように考えると考えられる。

イベント検出の手法については、IF~THENルールでの検出は主観に大きく依存するため、統計的な手法などを用いて学習を行えるようなシステムの構築を目指す。

文 献

- [1] 樋口知之, “粒子フィルタ”, 電子情報通信学会誌, vol.88, No.12, pp. 989-994, Decm 2005.
- [2] ROGER Y. TSAL, “A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Camera and Lenses”, IEEE JOURNAL OF ROBOTICS AND AUTOMATION, Vol.RA-3, No.4, AUGUST 1987,
- [3] 片山 徹, “応用カルマンフィルタ”, 朝倉書店, 2000.
- [4] 矢野一樹, “パーティクルフィルタを用いた物体追跡と自動映像生成への応用”, 神戸大学大学院 修士論文, 2008.
- [5] Toshihiko Misu, Atsushi Matsui, Masahide Naemura, Mahito Fujii, Nobuyuki Yagi, “Distributed Particle Filtering for Multiocular Soccer-ball Tracking”, IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 937-940, 2007.
- [6] Helio Palaio, Jorge Batista, “Multi-Objects Tracking Using an Adaptive Transition Model Particle Filter with Region Covariance Data Association”, IEEE International Conference on Pattern Recognition, 2008.
- [7] 立花 智哉, 山下淳, 金子透, “Haar-like 特徴を使った顔検出と Mean-Shift トラッカによる複数視点人物追跡システム”, 映像情報メディア学会技術報告, Vol.31, No.14, pp. 71-74, 2007.
- [8] J. Ren, J. Orwell, G.A. Jones, M. Xu, “A general

framework for 3D soccer ball estimation and tracking”, IEEE International Conference on Image Processing, pp. 1935-1938, 2004.

- [9] Yi Zhang, Hanqing Lu, Changsheng Xu, “Collaborate Ball and Player Trajectory Extraction in Broadcast Soccer Video”, IEEE International Conference on Pattern Recognition, 2008.
- [10] 石井規弘, 北原格, 亀田能成, 大田友一, “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), 1035-1040, 2007.
- [11] Xinguo Yu, Xin yan, Yiqun Li, “Tool-Aided Semantics Acquisition for Live Soccer Video”, IEEE International Conference on Multimedia and Expo, pp. 893-896, 2008.
- [12] 三須俊彦, 高橋正樹, 蓼沼眞, 八木伸行, “サッカー映像のフォーメーション解析に基づく実時間イベント検出”, FIT2005, LI-003, pp. 141-144, 2005.
- [13] 三須俊彦, 苗村昌秀, 藤井真人, 八木伸行, “局所的な選手人数分布に基づくサッカーイベント判別法”, FIT2007, H-062, pp. 149-152, 2007.
- [14] 窪田進太郎, “TVサッカー映像のオブジェクト情報に基づいた個人視聴向き内容解説”, 神戸大学大学院 修士論文, 2007.
- [15] 熊野雅仁, 有木康雄, 塚田清志, “ボールと選手に着目したデジタルカメラワークの実現法-デジタルシューティングによるサッカー解説映像生成システムに向けて-”, 映像情報メディア学会誌, Vol59, No.2, pp. 271-278, 2005