

シーン検索のためのサッカー中継画像の解析 Analyzing Video Image Sequences for Scene Retrieval in Soccer Games

山田 明人[†]
Akihito Yamada

白井 良明[†]
Yoshiaki Shirai

三浦 純[†]
Jun Miura

1. はじめに

近年、我々が利用できる映像情報の量は著しく増加している。放送、医療、教育などさまざまな分野での映像情報の利用は、今後ますます多くなることが予想され、膨大な映像情報の中から必要な情報を効率よく検索することが必要になる。本論文では、サッカー中継画像から視聴者の求めるシーンを検索することを考え、そのために必要な、カメラの姿勢・ズームのオンライン推定および、選手とボールの検出・追跡と位置推定について述べる。さらに、それらの処理結果を利用したシーン認識の例を述べる。

2. カメラの姿勢・ズームのオンライン推定

中継画像ではカメラの姿勢・ズームがフレームごとに変わるので、そのオンライン推定が必要となる。フィールド上にはフレーム間の対応づけに用いることのできる特徴点が少ない。フィールド上の白い直線を用いて推定する手法 [1] があるが、直線の情報が少ないところには適用できない。ここでは、グラウンド上の白線をすべて用いて推定する。

図 1 において、世界座標系 (X, Y, Z) 上の位置 (X_{cp}, Y_{cp}, Z_{cp}) にカメラが固定され、カメラ座標系 (X_c, Y_c, Z_c) において、 X_c 軸まわりの回転(チルト)、 Y_c 軸まわりの回転(パン)、および焦点距離の 3 つのパラメータがフレームごとに変化する。

カメラ位置は不変なので最初に一度だけ求めておく。画面上の静止物体の位置と世界座標系での位置の対応が 6 組以上得られれば、カメラ位置を求めることができる。ここでは、ゴールポストの角やグラウンド上の白線の交点を対応点として、手動で対応を与えカメラ位置を求める。

パンとチルトによる回転の座標変換行列を R とすると、カメラ座標系と世界座標系の関係は

$$[X_c, Y_c, Z_c]^t = R[X - X_{cp}, Y - Y_{cp}, Z - Z_{cp}]^t \quad (1)$$

となる(t は転置)。さらに、焦点距離を f とすると画像座標 (x, y) とカメラ座標との間に以下の関係が成り立つ。

$$X_c = Z_c(x/f), \quad Y_c = Z_c(y/f). \quad (2)$$

白線はグラウンド上にあるので、式 (1)(2) を代入して得られる (x, y) と (X, Y) の関係式を用いて画像上の白線と公式ルールで定められたラインとの対応を取る。

まず、画像上より選手以外の白色領域をライン候補領域として抽出し、世界座標系に投影する。投影された各ピクセルに一番近いモデルの座標との距離がしきい値(現在は $1[m]$) 以内であれば、対応を取る(図 2 参照)。前フレームでのパラメータを初期値として山登りを行って、対応点が一番多くなるパラメータを求める。推定されたパラメータを用いて画像列を合成してモザイク画像を生成した例を図 3 に示す。

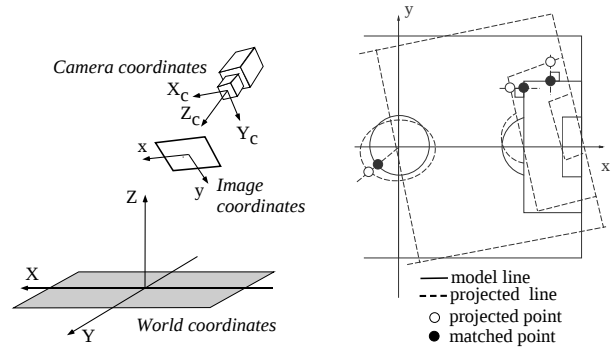


図 1: 座標間の関係

図 2: 対応点の決定法

3. 選手の検出と位置推定

色情報を用いて選手の検出と追跡を行う [2]。選手の追跡では隠蔽が問題となる。テンプレートマッチングを用いる方法 [3, 4] では、異なるチーム間の隠蔽のみしか考慮していない。ここでは、画像中の選手の位置関係も用いて選手間の前後関係を推定することにより、同じチームの選手同士の隠蔽にも対処している。選手の領域が求められれば最下部を足元としてグラウンド上の選手の位置を計算する。

4. ボールの検出と位置推定

ボールは小さくかつ高速で運動するため、形状や色のみで発見することは難しい。そこで、領域の大きさや形状に基づき複数のボール候補領域を生成し、一つの候補に絞られるまで追跡を行う [2]。

単眼でのボールの位置推定においては、ボールがグラウンド上にあればその位置は計算できるが、空中にあるときには計算できない。ボールの運動の物理法則を考慮して飛んでいるときにも 3 次元位置を推定する手法 [2, 5] があるが、ボールの進行方向がカメラの光軸方向に近いときには推定の信頼性が低くなる。そこで、ボールがバウンドした点とボールが選手と重なる点でボールの軌跡をセグメントに分割し、それぞれのセグメントについてボールが飛んでいるか転がっているかを、軌跡が放物線と直線のいずれのモデルに当てはまるかで判定する。ボールがグラウンド上にあるとき(直線の場合)のみその位置を推定する。モデルの当てはめは以下の手順で行う。

1. 画面系列上のボールの位置をモザイク画像上に投影する。
2. 軌跡の曲率を用いて検出したバウンド点と、ボールと選手とが重なっている点で軌跡をセグメントに分割する(図 4(a) 参照)。
3. 分割された各セグメントについて、直線か放物線のモデルを当てはめる。当てはめには AIC[6] を用いる。図 4(b) でセグメント 1,4 は直線, 2,3 は放物線となっている。

[†] 大阪大学工学研究科電子制御機械工学専攻



図 3: オンラインで推定したカメラパラメータを用いたモザイク画像の生成

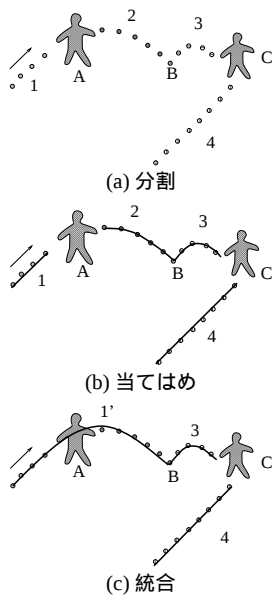
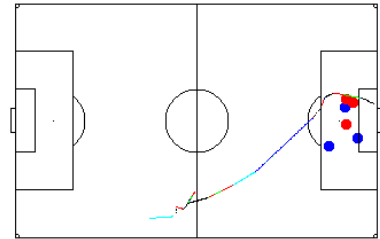


図 4: 運動モデルの当てはめ



(a) ボールがゴールのそばを通った例



(b) センタリングが上がった例

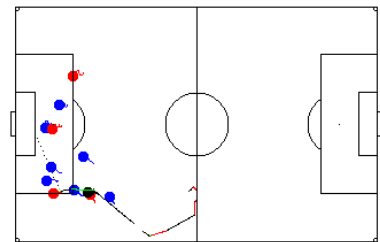


図 5: 認識した惜しいシーン．右図は選手の位置とボールの軌跡を示す．

4. 前後のセグメントが統合可能かどうかを AIC で判定する．図 4(c) では，図 4(b) のセグメント 1 と 2 が一本の放物線 1' に統合されている．

5. シーンの認識

ボールと選手の位置推定の結果からシーンを認識するために，各種のシーンを定義する．例えば「惜しいシーン」では「ボールを持った攻撃チーム選手がペナルティエリア内に入った」，「ボールを持っていなくても攻撃チーム選手二人以上がペナルティエリア内にいる」といった，画像情報から抽出した選手とボールの位置に基づいて判定することのできる定義を与え，それらを用いてシーンを認識する．図 5 に惜しいシーンとして認識された例を示す．その他，ゴールシーンの認識やボール支配率の計算なども行っている．

6. おわりに

本論文では，サッカー中継画像から興味深いシーンを自動検索するために必要な画像認識技術と検索の例について述べた．グラウンド上の白線をすべて利用することにより広範囲の視野に対してカメラの姿勢・ズームを推定した．さらに，選手とボールの発見と位置推定を行っ

た．ボールについては AIC を用いてボールが飛んでいるか転がっているかを判定し，それをもとに位置を求めた．選手とボールの位置関係に基づいてシーンを定義し，それを用いて興味深いシーンを自動抽出した．さまざまな日照条件への対応や処理の高速化，興味深いシーンの定義の自動生成などが今後の課題である．

参考文献

- [1] H. Kim and K.S. Hong. "Soccer Video Mosaicing using Self-Calibration and Line Tracking", Int. Conf. on Pattern Recognition, pp. 592-596, 2000.
- [2] Y. Ohno, J. Miura, and Y. Shirai. "Tracking Players and Estimation of the 3D Position of a Ball in Soccer Games", Int. Conf. on Pattern Recognition, pp. 145-148, 2000.
- [3] Y. Seo, S. Choi, H. Kim and K. S. Hong. "Where Are the Ball and Players? Soccer Game Analysis with Color-Based Tracking and Image Mosaick", Int. Conf. on Image Analysis and Processing, pp. 196-203, 1997.
- [4] S.S. Intille and A.F. Bobick. "Visual Tracking Using Closed-Worlds", MIT Media Laboratory, Perceptual Computing Section Technical Report No. 294, Nov. 1994.
- [5] T. Kim, Y. Seo and K.-S. Hong. "Physics-based 3D Position Analysis of a Soccer Ball from Monocular Image Sequences", Int. Conf. on Computer Vision, pp. 721-726, 1998.
- [6] 坂元，石黒，北川. "情報量統計学"，共立出版，1983.