

サッカー放送映像からの特定選手追跡の提案とその評価 Evaluation of Specific Player Tracking from Soccer Broadcast

中村 達也[†] 中島 克人[†]
Tatsuya Nakamura Katsuto Nakajima

1. はじめに

サッカー放送映像から特定選手のプレーに注目した要約映像を作成するには、映像内の同じユニフォームを着る多数の選手の中から特定選手のみを識別し、追跡する必要があるが、サッカーのフィールド上に小さく映る選手を区別する事は大変挑戦的な課題である。

本稿ではこの課題に対して、近年画像認識タスクで高い性能が報告されている深層学習に基づく人や背番号の認識技術と団体競技の選手追跡等で有効と思われる粒子フィルタ[1]を組み合わせる手法を提案し、評価する。

2. 提案手法

2.1 処理の流れ

全体の処理フローを図 1 に示す。まず、入力映像から 1 フレーム取得し、前処理としてフィールド抽出を行う事で選手追跡に不要な領域を削除した後、検出処理、識別処理、追跡処理、探索処理からなる特定選手追跡アルゴリズムを適用する。追跡に成功したフレームは要約映像のために蓄積する。図 2 はその詳細を示し、以下に各処理を詳述する。

2.2 前処理(フィールド抽出)

フィールドとそれ以外の領域で分割し、観客等の識別や追跡の障害となる画像領域を除外する。IFAB の規定[2]では、フィールドの色は緑色でなければならないとされているため、緑色の色相値付近である 80~160 を閾値とした 2 値画像を求め、ノイズ除去のため収縮・膨張したものをマスクとしてフィールドと思われる領域を抽出する。

2.3 検出処理

2.3.1 選手検出

識別や追跡を精度良く行うために正確な選手の位置情報を使用したい。そこで、深層学習を用いた一般物体検出器で人物(=選手)の位置と大きさを求める。今回は検出精度と速度の面から COCO dataset で事前学習された YOLOv3[3]の出力する人物検出矩形を使用する。なお、フィールド上に小さく映る選手を追跡する事が目的であるため、入力フレームの 1/100 以上の面積を有する検出矩形は排除している。

2.3.2 チーム判別

人物検出矩形に対し、予め登録しておいたユニフォームの RGB 値に基づき、追跡対象の特性選手が属するチームと思われる矩形のみをこの後の追跡対象候補とし、処理時間の削減を行う。判別は、登録した RGB 値を中心に一定の範囲(今回は $\pm 6\%$)にある画素数が人物検出矩形内に閾値以上の比率かどうかで行う。厳密には、映像内での観測さ

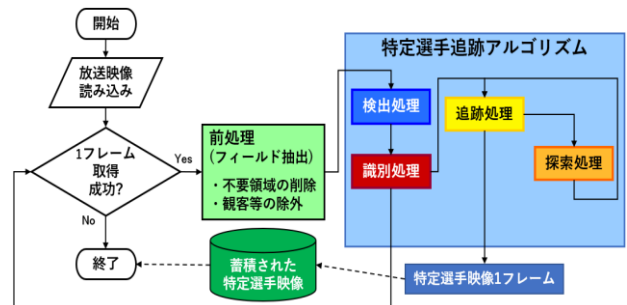


図 1 全体の処理フロー

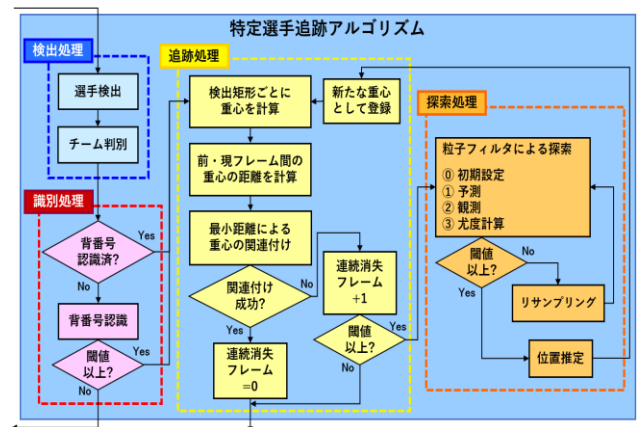


図 2 特定選手追跡アルゴリズムの処理フロー

れるユニフォーム色の変動に対処するため、[4]で提案の通り、登録した判別色は、当該矩形の内外の明度変動を反映して動的変動させる。

同チームと見なす選手の矩形を以後、**選手矩形**と称する。

2.4 識別処理(背番号認識)

深層学習を用いた背番号認識手法が提案されている[5]が、学習に十分なデータセットが公開されておらず、その人手作成も多大な労力を要する。そこで本研究では、ハウスナンバーの公開データセットである SVHM[6]を学習した数字認識モデルを使用する。このモデルは、放送映像内の小さくかつ選手の姿勢やモーションブラー等によって変形した背番号と同様に不鮮明な多くのハウスナンバーで学習しているため、識別したい背番号を 1~2 桁に絞った場合、良好な精度で放送映像内の背番号が認識可能なことが分かっている[7]。本研究では、識別したい背番号は 1 種類に限定した上で、信頼度が 0.85 以上を識別成功と定めた。識別成功時、後に行う探索処理のために[4]に基づいて追跡色としてユニフォームと頭部の色の登録を行う。

2.5 追跡処理

2.5.1 重心追跡アルゴリズム

背番号が識別出来たフレームを基点に追跡処理を行う。手順を以下に示す。

[†]東京電機大学 Tokyo Denki University

Step1 選手矩形群の重心を求める

選手矩形の重心を左上と右下の座標から求め、それぞれ一意の ID を割り当てて管理する。背番号識別で発見した特定選手の矩形には特別にフラグを立て、他の味方選手と区別可能にしておく。

Step2 前・現フレーム間で矩形同士の関連付けを行う

選手検出処理は常に実行されるため、毎フレーム重心を求めることができる。フレームレートが 30 fps 程度であれば、1 フレーム間での選手の移動距離は大きくないことを前提に、前後のフレーム間で重心位置の近い選手矩形同士を関連づけ、前フレームの ID を後フレームに伝承する。

追跡を行う中で検出された選手が増える場合がある。上記で、関連付けられず、かつ、ID が付与されていない選手矩形を追跡対象として新規登録する。登録作業は Step1 と同様である。関連付けに失敗し、かつ、ID が割り当て済みの場合、その選手矩形はフレームから消失した状態とみなし、連続消失フレーム数を 1 増やす。関連付けに成功した選手矩形は連続消失フレーム数を 0 とする。連続消失フレーム数が閾値である 30 を超えた場合、その選手矩形を追跡対象から削除する。もし、削除された選手矩形が特定選手だった場合、追跡処理を終了し、探索処理に入る。

2.5.2 逆追跡アルゴリズム

サッカー放送映像のあるショットの間、しばらく特定選手が見つからず、背番号が識別できて初めて見つかるという事がしばしばある。本研究の目的はオフラインで要約映像を作成することであるため、特定選手が識別できたフレームから遡って、即ち、時間的に逆方向に追跡を行うこととした。前述のように全ての選手矩形を追跡・管理するのはこの逆追跡を可能にするためである。

逆追跡は次のように行う。特定選手の識別ができない間は、検出処理で獲得された選手矩形群を蓄積していく。その後、識別に成功したら、そのフレームを基点に蓄積した選手矩形群を用いて前項と同様の方法で追跡を行う。蓄積した選手矩形群が無くなる、もしくは、逆追跡中の特定選手が追跡対象から削除された場合、識別成功フレームに戻り、重心追跡アルゴリズムを開始する。

2.6 探索処理

他選手やテロップ等とのオクルージョンやフレームアウト等で特定選手の追跡が中断した場合、粒子フィルタで探索を行う。

まず、最後に更新された特定選手の重心周辺(今回は 15×15 pixel の範囲、以下同様)に粒子を散布する(50 個)。次に、特定選手の次フレームでの位置を予測し、粒子を移動させる。今回は運動モデルにはランダムウォークのみを採用している。最後に、各粒子の周辺領域(10×10 pixel)内の追跡色の存在率を尤度として評価する。この際も照度変動等に対処すべく、[4]に基づく手法で尤度を求める。

尤度が上限閾値(0.80)を超える粒子が存在する場合、尤度が下限閾値(0.30)を超える粒子群の尤度による重み付き平均から選手の存在位置を推定し、その位置を新たな特定選手の重心として追跡処理を行う。存在しない場合、リサンプリングを行い、探索を継続する。一定回数継続するごとに粒子数を増やし、散布する範囲を広げることで想定外の位置への移動に対応する。

3. 性能評価

実際の放送映像上の白色ユニフォームのチームの背番号 19 番の選手に提案手法を適用し、性能評価を行った。解像度は 1280×720 pixel で、適用時間は試合開始から 10 分間である。映像には何も手を加えないため、フィールドを俯瞰したシーン以外が含まれる。プレーに注目した映像を作成するためにオプティカルフローを用いたショット切替り検出を行った。ショット切替り時には、特定選手が未識別の状態にし、追跡と探索処理に使用する変数は全て初期化する。

評価指標のための分類の定義を表 1、フレームごとに分類評価した結果を表 2 に示す。

表 1 評価指標のための分類の定義

True Positive (TP)	False Positive (FP)
フラグが立っている選手が特定選手である	フラグが立っている選手が特定選手ではない
False Negative (FN)	True Negative (TN)
フレーム内に特定選手がいる時に識別or探索処理	フレーム内に特定選手がいない時に識別or探索処理

表 2 性能評価

手法	適合率(%)	再現率(%)	F 値(%)
逆追跡無し	96.4	37.7	54.2
逆追跡有り	92.9	50.4	65.3

逆追跡の採用により、再現率が大幅に向上していることが分かる。再現率の更なる向上には、人が行っているような、背番号以外のスパイクや髪の色、ポジションなどの他の情報を利用する必要があり、これは今後の課題である。

適合率に関しては、非常に高い値を得ており、特定選手の追跡に本手法が有効である事を示していると言える。

4. まとめ

本稿では、サッカー放送映像から特定選手のみを識別し追跡するために、深層学習に基づく人や背番号の認識技術と粒子フィルタを組合せた手法を提案した。評価の結果、ほとんど誤追跡すること無く、追跡すべきフレームの内、約半分の追跡に成功した。今後は、上記の再現率向上のための改良に加え、選手の密集等の追跡が困難な状況への対応を行う。

参考文献

- [1] M. Isard, et al., "CONDENSATION - Conditional Density Propagation for Visual Tracking," In *Int J Computer Vision* (1998).
- [2] IFAB, "International Football Association Board | IFAB," <http://www.theifab.com/laws> (2019-06-21 参照).
- [3] J. Redmon, et al., "Yolov3: An Incremental Improvement," In *arXiv:1804.02767* (2018).
- [4] 中村 達也ら, 他, "サッカー放送映像からの特定選手追跡のための動的尤度を用いた粒子フィルタの提案," *映像情報メディア学会年次大会* (2018).
- [5] G. Li, et al., "Jersey Number Recognition with Semi-Supervised Spatial Transformer Network," In *CVPR Workshops* (2018).
- [6] Y. Netzer, et al., "Reading Digits in Natural Images with Unsupervised Feature Learning," In *NIPS Workshops* (2011).
- [7] 中村 達也ら, 他, "サッカー放送映像からの特定選手追跡のための背番号認識," *電子情報通信学会総合大会* (2019).