3章 画像センシング技術によるチームスポーツ映像からのプレー解析

林 昌希^{†,††}, 正会員 青木 義満^{††}

キーワード：チームスポーツ映像，プレー解析，画像センシング，姿勢推定，プレー分類

1. まえがき

現在，スポーツ競技を対象としたICT技術の活用が進められている．特に，サッカーやアメリカンフットボール等，著名なプロリーグが存在し，人気の高いチームスポーツに関しては，さまざまな研究開発が行われている．欧州サッカーや北米の3大スポーツ（野球，バスケットボール，アメリカンフットボール）などを中心に非常に規模の大きな産業でもあるボールを用いたチームスポーツにおいては，実際の試合における戦術の分析と，それに基づく新たな戦術の立案・理解が非常に重要とされており，戦術解析のニーズが非常に大きい．特に，トレーニング中だけではなく，実際の試合映像に対して，選手やボールの情報をデジタル化し，統計解析するというニーズが増えている．そのような要求に対し，選手に特殊なデバイスを装着させることなく，選手やボールの位置情報を定量化することが可能である画像センシング技術に対する期待が高まっている．

例えば，カメラ映像からのサッカーのリアルタイム追跡システムとして，TRACAB¹⁾という製品が欧州サッカーや日本のJリーグで実際に使用されている．追跡されたデータは統計情報化され，各チームが対戦チームの戦術分析に使用している．また，選手ごとの移動範囲ヒートマップを表示するなど，放送映像への付加価値追加目的にも活用されている．一方で，個々の選手のプレー中の動作の認識や，フィールド上の全選手やボール情報からのプレー種類の認識といった，選手追跡の後に続くより高度な画像認識ニーズに対する実用化は，未だ限定的である．現状では，映像を見ながら必要に応じてスポーツコード²⁾などの専用ソフトウェアを用いて，手動でプレー解析に必要なタグ付けを行い，管理・解析を行っている．

筆者らは，チームスポーツの中でも，特にアメリカン

フットボールを対象とする戦術的データ解析の効率化および自動化を目的とする試合映像解析の研究を行ってきた．アメリカンフットボールは，競技の性質上，選手同士が密集する状況が多く，個々の選手を頑健に追跡し続けることは困難である．本稿では，映像全体のOptical Flowの方向や選手密集具合を特徴量として用いることで，各選手個別の追跡が達成できなくても実施可能なプレー解析手法を紹介する．また，個々の選手のアクション認識のために有用となる人物姿勢推定手法として，選手追跡と統合した人物姿勢推定のフレームワークを提案してきたので本稿では，その可能性を示す．

2. アメリカンフットボール映像からのプレー解析の研究

アメリカンフットボールは各プレーに開始終了の区切りがあり，またプレーごとに異なる陣形および移動戦術を両チームが選択する戦術的な駆け引きを行うスポーツである．手作業でプレー解析者やコーチがプレーのタグ付けを映像に行うことが現在の主流であるが，もし試合映像に自動的に各プレーの開始終了の区切りが与えられ，かつ各プレーの大まかなタイプ分類も自動化されれば，作業効率化の面で価値が高い．また，詳細な解析対象情報として，ボール保持者の変遷とその移動軌跡も推定できると価値が高い．しかし，アメリカンフットボール映像では（特にボール保持者周囲で），選手同士が遮蔽して混雑している状況が頻繁に起こるので，複数選手を個別に追跡し続けることは，最先端の複数人物追跡技術を用いても困難である．

そこで筆者らは，混雑中に選手位置をすべて追跡することは追求せず，代わりに映像全体のOptical Flowから得た大まかな人の流れ(Global Motion Feature)や選手密集度などを特徴量として用いるプレー解析手法を提案してきた．これにより，選手の完璧な追跡が困難なアメリカンフットボールにおいても，以下の三つのプレー自動解析が実現可能であることを示してきた(図1)：

- (1) ボール（およびボール保持者）軌跡推定(2.1項)³⁾
- (2) プレー時間判定（オンプレー中の範囲のみを抽出）⁴⁾
- (3) プレー分類(2.2項)⁵⁾

[†] 株式会社Recreation Lab

^{††} 慶應義塾大学 理工学部

"Play Analysis for Team Sports Video by using Image Sensing Technologies" by Masaki Hayashi (Recreation Lab, Tokyo/Faculty of Science and Engineering, Keio University, Kanagawa) and Yoshimitsu Aoki (Faculty of Science and Engineering, Keio University, Kanagawa)

図1 アメリカンフットボール映像の解析

以降では、これらのうち(1)と(3)のプレー解析手法について紹介する。

2.1 ボール軌跡推定

本項ではボール軌跡推定³⁾の手法を紹介する。処理手順としては、最初にプレー開始位置の推定とプレー終了位置の推定を行い、ボール保持者軌跡の開始位置と終了位置を求める。その後、開始・終了位置を両端とする1プレー中において、選手検出結果を入力として、以下の3種類の「選手矩形のボール保持者尤度」を用いて各フレームでのボール保持者を選手検出結果から選択し、その結果を全フレーム分結合した結果をボール保持者軌跡(ボールの軌跡およびボール保持者軌跡)とする。

- (1) 被アプローチ尤度(2.1.3項)
- (2) プレー終了位置への距離尤度(2.1.4項)
- (3) チーム識別尤度(2.1.5項)

また、パス中でボールが検出できたフレームではボールの位置をボール保持者軌跡として採用する。以降、この三つの尤度について述べ、2.1.6項にてボール保持者推定処理の説明を行う。

2.1.1 プレー開始位置推定

アメリカンフットボールでは初期陣形を組み、ボールがセットされている場所周辺は両チームの選手が毎回同じように密集した画像が観測される。そこで、筆者らはこのボール開始位置周辺領域の検出器を、Histograms of Oriented Gradients (HOG) 特徴量とAdaboost識別器を用いた物体領域検出器として学習し、各プレーの画像中での

開始位置(初期陣形)を検出する。

2.1.2 プレー終了位置推定

アメリカンフットボールのプレー終了位置はボールを保持している攻撃側の選手に、防御側のチームが多数集まり、その選手を最終的につかまえることのできた位置である。これを画像特徴量に置き換えようとする、選手全体のオプティカルフロー場の進行方向が一番集積した位置に相当する。そこで、画像を入力としてフィールド上の各位置で推定するプレー終了位置の尤度として「進行方向集中度」および「選手密集度」の2つを定義し、これら二つの尤度の重みつき和が最大となる位置をプレー終了位置とする。

進行方向集中度は、各選手の動き方向をその進行方向上に配置する各グリッド上に蓄積し、その全選手分の和を計算した結果値である。複数の選手が目指している方向と場所ほど進行方向集中度の集積度合いが高くなるので、プレー終了位置を含めて複数の選手が向かっている位置の値が高くなる性質を持つ尤度量である。一方、選手密集度は、まずHOG+Adaboostによる1フレームでの選手検出器結果を用いて選手矩形の3人分程度の大きさの窓によるスライディングウィンドウ処理を行い、各窓内の人数(検出矩形とのオーバーラップ割合)に比例した尤度を窓中心位置に割当て、これを各フレームでの選手密集度とする。

2.1.3 被アプローチ尤度

アメリカンフットボールでは、ディフェンスチームの誰かがオフェンスチームのボール保持者を捕まえた時点で1プレーが終了するため、ボール保持者周辺に存在する複数の

ディフェンスチーム選手がボール保持者に接近していく。したがって、ディフェンスチーム選手の進行方向ベクトルの指す先の直線が一番集中する交差領域がボール保持者の現在位置である可能性が高く、これを表現した「被アプローチ尤度」をボール保持者位置の尤度として使用する。被アプローチ尤度は、各選手のオプティカルフローの方向を選手群の進行方向と見なし、各位置でのフロー方向前方の直線上のすべての座標に均一に尤度を加えた2次元フィールド座標上のマップとして表現する。これにより、複数の選手が進んでいる直線方向が重なっている位置ほど尤度が高くなるフィールドマップが計算される。

2.1.4 プレー終了位置への距離尤度

選手検出位置と2.1.2項で求めた終了位置の距離を用いてボール保持者らしさの尤度を算出する。大きなパスが行われ遠くの別の選手にボール保持者が切替わらない限りは、他の尤度が高い選手のうち、プレー終了位置に最も近い選手がボール保持者である可能性が高いことが、この尤度を用いる理由である。

2.1.5 チーム識別尤度

各選手検出枠に対し、(ボールを保持しているはずの) オフェンスチームらしさの見える尤度を、チーム識別尤度として使用する。選手検出器の学習データにチームのオフェンス/ディフェンス側の2クラスのラベルも付与しておき、検出枠内の色ヒストグラムを特徴ベクトルとしたSVM (Support Vector Machines) を学習しておく。そしてテスト動画の各選手検出枠に対して、SVMでオフェンスチームクラスのスコアを推定することで、そのスコア値をチーム識別尤度として選手検出領域に割当てて。

2.1.6 ボール保持者軌跡の推定

最初に、ボール初期位置をボール保持者軌跡のスタート地点でのボール保持者位置として設定する。そして各フレームにおいて、3種のボール保持者尤度の重み付き線形和が最大となる位置の選手をボール保持者として推定する。この処理を終了フレームまでの全フレームについて行うことで、最終的に各フレームで選択された検出結果の集合としてボール保持者軌跡が推定される。またパスプレーの際には、パス中にボールそのものを追跡できていれば、そのボール位置をボール保持者軌跡として使用することができるので、その間は人の位置はボール保持者として採用しない。パスが行われたプレー(パスプレー)、ボール保持者が走ったプレー(ランプレー)のプレー種類ごとに、各尤度の重み係数の値を個別に調整することで、よりうまく軌跡を推定できることを理由に、ランプレーとパスプレーでは予備実験から得られたプレー種類個別の重み係数の値を使用している。処理の詳細は文献³⁾を参照されたい。

2.2 プレー分類

アメリカンフットボールの攻撃プレーの種類としては、開始後に前方遠目の選手のパスを行う「パスプレー」と、パ

スは行わず開始位置から走って距離を稼ぐ「ランプレー」に大きく分類できる。そこで筆者らは、左/右のどちら側の領域において、ランプレー/パスプレーのどちらで攻めたかの計4種類の攻撃プレー種類を動画から分類する手法を提案した⁵⁾。

まずプレー種類分類の問題設定としては、プレー開始時刻とプレー終了時刻を均等に6セクションに分割した上で、セクションごとに算出した入力特徴量をRandom Forestsを用いて4プレー種類クラスのいずれかに分類を行うものとした。1プレー全体に対するプレー種類クラスの推定を行うのではなく、1プレー内の短い各セクションでプレー種類を独立に判定するという点に注意されたい。これにより、動画中の短い6セクションそれぞれでのプレー種類クラスが予測されるので、より細かい攻撃チームのプレースタイル理解が可能となる。

各セクションで算出する特徴ベクトルとしては、該当セクションの毎フレームにおいて、(2.1.1項と2.1.2項の手法で求める) プレー開始/プレー終了位置に対する各チームの選手位置の距離を、一つの列ベクトルとして結合したものをを用いる。「1プレー中では、選手のプレー種類変化は初期陣形位置を基準とした相対的座標系上での変化として表現される」という仮定に基づき本特徴量を採用した。これにより、実験で用いた試合映像に対して、平均10~20秒程度の長さで構成される各セクションのプレー種類を、80%程度の認識率で識別可能となった。

3. 選手追跡に統合された、位置合わせ窓内画像特徴を用いたチームスポーツ選手の姿勢推定

筆者らは、旧来からの人物追跡システムに統合した形で計算コストも低くシンプルに行うことができるチームスポーツ映像向けの新たな人物関節位置推定手法としてLabel-Grid Classifier⁶⁾およびPoselets-Regressor⁷⁾を提案し、これらを用いて下半身の関節位置を推定する手法⁶⁾と上半身の背骨姿勢および上半身の体向きを推定する手法⁷⁾をそれぞれ提案した(図2)。以降、3.1項で研究背景と旧来手法について述べる。3.2項で筆者らの主提案である関節位置推定の仕組みである、Label-Grid Classifier/Poselets-Regressorについて概観する。その後、下半身姿勢推定手法を3.3項で、上半身姿勢推定手法を3.4項でそれぞれ個別に述べる。3節で述べる内容については、すでに文献⁸⁾でまとめ報告を发表済みであるので、そちらも参考にされたい。

3.1 研究背景と旧来の人物姿勢推定手法

スタジアム映像からの選手軌跡の追跡は、欧米のプロサッカーリーグやプロバスケットボールリーグ等ですでに商用化されている。これらのシステムから得られる選手のフィールド座標系上での2次元軌跡情報は、走行距離ランキングや選手ごとの移動範囲ヒートマップなどの位置統計

図2 チームスポーツ映像での追跡に統合された姿勢推定

情報の算出にも使用されている。しかし、「シュート」、「パス」などの各アクションの認識は自動化されておらず、映像を見ながら各アクションのタグ付けを手動で行っているのが現状であり、映像中に映る選手のアクションの詳細な自動解析までには至っていない。

もしMicrosoft社のKinectのように各選手のボーン姿勢がチームスポーツ映像から推定できれば、そのボーン姿勢情報を入力として用いるだけで詳細なアクション認識が容易に実現できる。また、推定された詳細な姿勢変化データから「1アクションの中でもどういった違いがあるのか」の解析まで行うことができ、「背骨の角度の違い」など、姿勢そのものの違いを捉える直感的な認識処理も容易に実現可能である。しかしKinectのようにRGB-Dデータを入力とすることは、広いスタジアムを撮影する必要があるチームスポーツ映像では厳しいので、RGB画像のみで人物姿勢推定を実現する必要があると言える。

これまでのRGB画像を入力とする人物姿勢推定技術は、Pictorial Structures⁹⁾と呼ばれるパーツ検出器の空間配置モデルを人物画像にあてはめる研究が主になされてきた。Pictorial Structuresを用いた人物姿勢推定手法の代表例であるFlexible-Mixtures-of-Parts (FMP)¹⁰⁾などのモデルが検出を行いやすいような「(主に正面向きで)四肢のパーツがすべてカメラに映っている姿勢」を主に含むデータセット(例：Leedsデータセット¹¹⁾、FLICデータセット¹²⁾)を対象として研究されてきた。このため、サッカーなどのチームスポーツ動画では頻出である「ゴール方向に横向きに走っていて、カメラ位置から見て奥側の多くの部位が遮蔽して見えない姿勢」や「回転する動きのアクション中に多くのパーツが遮蔽する姿勢」には、FMPなどの旧来のパーツ検出ベースの手法では対応しづらいので、これらの姿勢も含めてす

べての姿勢が推定することが可能な手法がチームスポーツ映像向けには必要であった。そこで、横向きでもパーツ遮蔽時にも対応できる手法を提案するに至った。

3.2 Label-Grid Classifier/ Poselets-Regressor:基準関節位置が位置合わせ済みのHOG特徴を用いた相対的関節位置推定

Label-GridとPoselets-Regressorは共に、顔検出と顔認識のペアのごとく、人物を追跡した時点で後半の認識に用いる窓内の画像パターンの位置を確定させる、位置合わせ窓内の密な画像特徴(具体的にはHOG特徴量)を用いた関節位置の推定手法である。

まず前半の処理では、図2左側のように、人物追跡や頭部追跡手法を用いて、対象人物の基準関節位置(骨盤中心もしくは頭部中心)を推定する。後半の処理では、基準関節位置を原点とするローカル座標系で窓原点をその基準関節位置に位置合わせした窓領域から計算したHOG特徴ベクトルを入力とし、Random Forestsによって「基準関節位置に対する相対的な目標関節位置」を推定する。この仕組みを識別器で行う場合、Label-Grid Classifierと最初名付け⁶⁾、この仕組みを回帰で行う場合をPoselets-Regressor⁷⁾と名付けた。直感的に理解しやすいように言い換えておくと、人検出でよく用いられるHOG特徴量を、窓内の局所座標系の原点がある関節に位置合わせされたHOGパターンのみをRandom Forestsに学習させる関節位置推定手法であるとも言える。

3.3 下半身姿勢推定：Label-Grid Classifierによる足の4関節位置の推定

Label-Grid Classifierの論文⁶⁾では、追跡によりまず骨盤位置が求まり、骨盤を中心とする人物領域のHOGベクトルからRandom Forestsを用いて「骨盤中心に対する足の関節グリッド位置」をグリッド位置識別問題として推定

した(図2②-a)。

本手法により、これまでパーツモデルを用いた姿勢推定手法が不得手としていたパーツ間遮蔽の多い姿勢も含めて、(HOGの位置合わせさえ良ければ)すべての姿勢パターンでの下半身関節位置を推定できるようになったことが本手法の貢献である。チームスポーツ動画では、ゴール方向に移動するので横向きに選手が映ることが大半であるが、旧来手法が苦手とするこれらの「横向きに映る際に奥側パーツの遮蔽が多い時」もすべて関節位置推定が可能となったことを実験で示した。

3.4 上半身姿勢推定: Poselets-Regressorを用いた背骨姿勢推定と、背骨角度で条件付けした上半身方向推定器

Poselets-Regressorの論文⁷⁾では、頭部検出を用いたTracking-by-Detectionにより頭部中心位置を追跡したのち、頭部中心を基準位置とする上半身領域のHOGを入力としてRegression Forestsにより「頭部中心に対する骨盤中心の位置」を推定した(図2②-b)。これにより頭部中心と骨盤中心をつないだ直線を、おおよその「背骨姿勢(上半身の傾き)」とみなすことができる。

この背骨姿勢推定の応用として⁷⁾では、背骨姿勢の傾き角度範囲を5分割し、その範囲ごとに8方向上半身向き推定器を個別に学習する方向推定手法を提案した。Poselets-Regressorと同じく8方向推定器の入力も頭部中心位置を原点とした上半身のHOG特徴を用いる。例えば、背骨傾きが30°から60°の範囲に入る上半身HOGパターンのみを入力として期待する8方向推定器(Random Forests)を学習する。これにより、学習時にRandom Forestsが少ないHOGパターンから強い特徴を選択しやすくなり、かつ背骨が大きく傾いている時も体向きを推定できる、スポーツ映像に適した体方向推定手法を提案することができた。

4. むすびと展望

チームスポーツ、特にアメリカンフットボール映像を対象とした、画像センシング技術によるプレー解析および姿勢推定に関する研究事例について紹介した。個々の選手の動きのみならず、全体の選手の動きを特徴量とするプレー解析手法は、アメリカンフットボール以外のチームスポーツに対しても有効と考えており、サッカーやラグビー等、他の競技への適用を検討している。また、プレー中の選手の姿勢情報の推定結果は、個々の選手についての詳細なプレーの認識や技能評価への活用が大いに期待されており、

チームスポーツのみならず、テニス他の競技への展開を視野に入れた研究を進めている。以上、これらの技術を現場のニーズを取り込みながらさらに発展させ、よりインテリジェントでかつ現場の戦術解析に実際に役に立つスポーツ映像解析技術を実現していきたい。(2016年5月31日受付)

〔文 献〕

- 1) <http://chyronhego.com/sports-data/tracab>
- 2) <http://sportscode.jp/>
- 3) 蔵野隼二, 山本大樹, 林昌希, 大島京子, 田藤雅基, 青木義満: “アメリカンフットボール映像におけるプレー開始・終了位置と選手動作情報を用いたボール軌跡抽出”, 精密工学会誌 (Jan. 2015)
- 4) J. Kurano, M. Hayashi, M. Tanabiki, J. Furuyama and Y. Aoki: "Automatic Match Analysis in Sports Videos Using Bayesian Network and Random Forest", 2016 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP 2016)
- 5) 蔵野隼二, 林昌希, 古山純子, 田藤雅基, 青木義満: “チームスポーツ映像におけるプレー開始・終了位置算出結果を用いた自動プレー分類”, ビジョン技術の実利用ワークショップ (Dec. 2015)
- 6) M. Hayashi, K. Ohshima, M. Tanabiki and Y. Aoki: "Lower Body Pose Estimation in Team Sports Videos using Label-Grid Classifier integrated with Tracking-by-Detection", IPSJ Transactions on Computer Vision and Applications (CVA) (Feb. 2015)
- 7) M. Hayashi, K. Ohshima, M. Tanabiki and Y. Aoki: "Upper Body Pose Estimation for Team Sports Videos Using a Poselet-Regressor of Spine Pose and Body Orientation Classifiers Conditioned by the Spine Angle Prior", IPSJ Transactions on Computer Vision and Applications (CVA) (Oct. 2015)
- 8) M. Hayashi and Y. Aoki: "Human Body Pose Estimation for Team Sport Videos with Poselets-Regressor and Head Detector", CVIM研究会 (May 2016)
- 9) P.F. Felzenszwalb and D.P. Huttenlocher: "Pictorial structures for object recognition", IJCV (2015)
- 10) Y. Yang and D. Ramanan: "Articulated pose estimation with flexible mixtures-of-parts", CVPR (2011)
- 11) Leeds Sports Pose Dataset, <http://www.comp.leeds.ac.uk/mat4saj/lsp.html>
- 12) Frames Labeled in Cinema (FLIC) dataset, <http://bensapp.github.io/flic-dataset.html>

はやし まさき (株) Recreation Lab 代表取締役。
画像認識を用いた人物姿勢推定と行動認識が主な研究
興味。博士(工学)。

あおき よしまさ 慶應義塾大学理工学部電子工学科
准教授。画像計測・画像認識、特に人物を対象とした
画像センシング技術に関する研究に従事。博士(工学)。
正会員。