

ラグビー映像解析システムの開発とその応用

小林大祐^{†1} 中洲俊信^{†1} 大内一成^{†1} 青木義満^{†2}

概要：近年、スポーツ界では ICT を活用したトレーニングや戦術分析の導入が進んでおり、画像認識技術を用いた試みも行われている。筆者らは、選手の数が多く、密集や接触プレーが多いため画像による解析が困難であったラグビーを対象に、これまで人手で手間と時間をかけて行われていたタグ付け作業の省力化を目指した映像解析システムを開発した。本システムにより、一つのカメラ映像のみを入力として、ボールと選手の位置を追跡するとともに、主要なプレーの自動分類を可能にした。本システムで開発した技術は、ラグビーに限らず様々なスポーツへの活用が可能であるだけでなく、防犯カメラを用いた作業分析など産業分野への応用が期待される。

1. はじめに

ICT (Information and Communication Technology) の発展により、スポーツ界では ICT を活用したトレーニングや戦術分析の導入が進んでいる。筆者らは、2019 年ワールドカップの日本開催や、2020 年東京オリンピックでも競技種目となるため注目度が上がっているラグビーを対象に、画像認識技術を活用したラグビー映像解析システムの開発を進めている。

ラグビーは、選手の数が多いチーム 15 人と多く、接触プレーや密集プレーが頻繁に発生するため、映像による分析には技術的な課題が多く、これまで積極的に取り組まれてこなかった。選手にセンサを装着すること無く、カメラでプレーの様子を撮影するだけで選手やボールの動き、プレーの自動分類を行うことができれば、練習や試合の現場で手軽かつ詳細な戦術分析が可能になる。さらに、技術的な課題の多いラグビーを対象に画像認識の基盤技術を強化することにより、その技術を他のスポーツ、さらにはスポーツ以外の産業分野へ幅広く応用することが期待される。

2. 従来手法

チームスポーツでは、個人のログだけでなく、試合におけるチームとしてのフォーメーション、戦術を把握し、統計データに基づいた戦略立案が求められている。例えば、欧州サッカー、J リーグなどで導入されている ChyronHego 社の TRACAB[1]というトラッキングシステムは、複数台のカメラをフィールド全体が撮影できるように設置し、選手、ボールの動きを追跡しデータ化するものであるが、価格が高く、あらかじめ同システムを設置した競技場でしか利用できない。また、追跡結果やフォーメーションの特徴を用いたプレーの認識などには対応しておらず、アナリストが映像を見ながら、例えば SPORTSCODE[2]などのスポーツ専用分析ソフトを使って必要なタグを手動で入力（あるいは手動入力したデータを購入）して、戦術分析を行っているのが現状である。

本論文で対象としているラグビーについては、上述のとおり、映像解析の対象としては技術的な課題が多いため、限定的なシーン推定に取り組んだ事例[3]を除いては、これまで殆ど取り組まれてこなかった。しかしながら、トップチームでは、ICT を活用したチーム強化の試みが積極的に進められている。ラグビー日本代表チームは GPS データを積極的に活用して 2015 年ワールドカップで好成績を収めた[4]。ただし、GPS の位置精度が十分でなく、また、データを取得するためには一旦デバイスを取り外してクレードル経由で取り込む必要があるため、チーム全体の戦略分析には活用できておらず、サッカーと同様に、スポーツ専用分析ソフトでタグを手動で入力して戦術分析を行っているのが現状である。また、相手チームの GPS データは取得できないため、相手チームの分析には利用できない。これらに対し、筆者らは、センサを装着せずに映像のみからラグビーのプレーを解析し、戦術分析に役立てることが可能な映像解析システムの開発を目的としている。

3. ラグビー映像解析システム

3.1 システム概要



図 1 ラグビー映像解析システム

開発したラグビー映像解析システムのメイン画面例を図 1 に示す。入力には左側に表示されている 1 台のカメラで撮影した試合映像のみで、他のセンサは使用しない。この例では、一般的なチームでも使用されている家庭用ビデオカメラで撮影した HD 映像 (1280×720pixels, 30fps) を用

^{†1} (株)東芝
Toshiba Corporation
^{†2} 慶應義塾大学
Keio University

いた。本システムで実現した各機能の概要は、以下の通りである。

- **選手／ボールの移動軌跡記録機能**

入力映像から両チームの選手（チームも識別）とボールを検出して追跡し、その位置を推定して図1右側の仮想二次元フィールド上にマッピングすることにより、プレー中の両チーム選手とボールの動きを連続的に記録することを可能にした。また、ボール保持チームの推定も行い、どちらのチームが攻撃しているかも合わせて記録する。

- **プレー推定による自動タグ付け機能**

入力映像からその時点その時点のプレー内容を推定し、推定結果の尤度が高い場合は自動でタグとして記録する機能を開発した。プレー推定処理は毎秒1回行い、推定結果と尤度は、画面左下に上位2候補を表示して確認できるようにした。尚、必要とされるタグは、チームの強化方針、アナリスト個人によって異なる場合があることが事前のヒヤリングで確認されたため、ユーザ毎のカスタマイズを容易にすることが今後求められる。

- **タグに基づくプレー検索機能**

自動タグ付け機能によりつけられたタグを手がかりに、所望のプレーを検索することを可能にした。図2は「ラインアウト」のシーンを検索した結果の一例である。これまではタグ付け作業を手動で行うか、手動入力したデータを購入するしかなかったため、試合の現場で試合中にタグを使った分析を行うことは困難であった（現場で手動タグ付けを行うことは負荷が高く、作業中は他の分析ができなくなるため、試合後に作業することが多い）が、これを自動化することで手間を省くだけでなく、例えばハーフタイムに前半の気になるプレーを重点的に確認し、後半の戦略に活用することなども可能になる。



図2 タグに基づくプレー検索

- **ボール移動軌跡に基づく手書き検索機能**

選手／ボールの移動軌跡記録機能により、ボール保持チームを推定した上でボール移動軌跡を記録する。この移動軌跡から、図3右上のように「赤チームがこのあたりをこの方向に攻めたシーン」のような直感的な検索を手書きのクエリーで行うことを可能にした。赤チームが自陣右側付近を矢印の方向に攻めたシーンとして、図3右下のサムネ

イルに示す3つのシーンが検索された例を示している。尚、この機能は、手書きタブレット向けの手書きスケッチによる資料検索技術[5]を活用した。

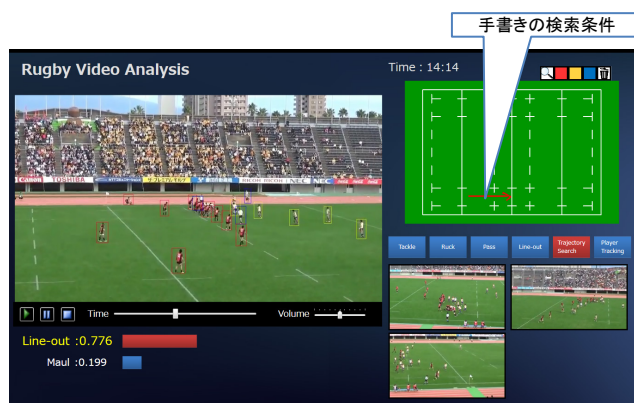


図3 ボール移動軌跡に基づく手書き検索

- **走行距離・スピード推定機能**

選手の検出と追跡、およびその位置推定において、現時点では個人の識別までは行っていないが、着目したい選手を選択した場合に、その選手を追跡している間は、二次元フィールド上の推定位置のフレーム間の差分をもとに、走行距離と現在のスピードを推定する。図4にその一例を示す。画面左下に着目した選手の走行距離とスピードが表示されている。



図4 走行距離・スピード推定

3.2 これまでの課題

上述の機能の実現に際して、一つのカメラ映像のみを入力として、特徴量設計方式によるボール検出と、ディープラーニング方式による選手検出を行うハイブリッド型映像解析技術を開発し、ボールと選手を精度良く検出可能にした。また、それらの動きを高精度に追跡し、移動軌跡を仮想二次元フィールド上にマッピングすることで、動きを俯瞰して分析可能にした。さらに、自動的なプレー分類をディープラーニングによって行い、これまで人手で行われていた主要プレーのタグ付け作業の省力化を可能にした[6]。

しかし、これまでのシステムでは、選手検出とボール検出を別の方式で行っていたことと、フィールド上にマッ

ングするための座標変換処理において射影変換行列の更新に繰り返しの処理が必要だったため、実際の現場で活用するには処理時間が課題だった。選手とボールを安定して追跡するためには、毎秒5フレーム以上のフレームレートが必要であることがわかり、リアルタイム処理を実現するためには1フレームあたり0.2秒以内での処理が求められる。そこで、ディープラーニングで同時に選手とボールを検出できるようにするとともに、カメラの撮影に現場で許容可能な制約条件を導入して計算量を削減することで処理全体の高速化を検討した。

3.3 ディープラーニングによる選手とボールの同時検出

これまで、ディープラーニングによる物体検出手法の一つである Faster R-CNN(Region Based Convolutional Neural Networks)[7]を用いて選手検出を行い、ボールは特徴量設計方式による検出を行っていた。さまざまな体格差や姿勢の変化がある選手の検出には、それらに対応可能なディープラーニングを用い、大きさ、色、形状が一定のボールはその検出に適した特徴量を設計することでそれぞれ精度良く検出していました。しかし、この方式では別々の処理が必要になるため、これを統合することによる処理量削減を検討した。

今回導入した、選手・密集領域およびボールを同時に検出するネットワークの構造を図5に示す。ボールは映像中で小さく映るため、従来のFaster R-CNNのネットワーク構造のままではCNN最終層の出力は解像度が低くなるため検出精度が低下する。そこで、CNNの複数の浅い層の解像度の高い出力も結合して対象候補領域の推定、クラス識別および矩形回帰を行うことで、小さい対象物体に対する検出精度の向上を図った。その結果、完全に見えているボールの検出精度（適合率）を従来手法（特徴量設計方式）の66.7%から89.1%に大きく改善した。

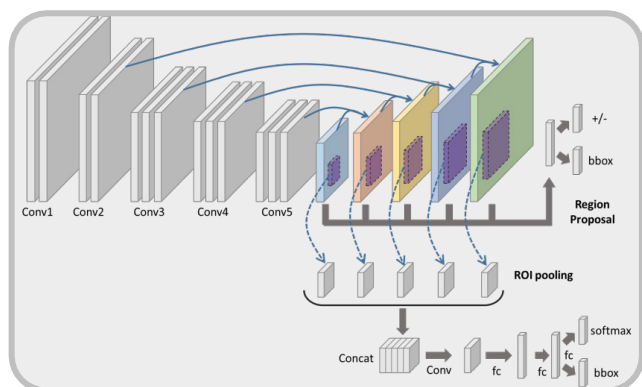


図5 選手・密集領域とボールを精度良く同時検出するネットワーク構造

3.4 仮想2次元フィールド座標変換処理の高速化

本システムでは、一つのカメラ映像のみを解析対象としているため、検出した選手とボールの仮想2次元フィールド

ドへのマッピングには、カメラの動きに対応する必要がある。入力映像と仮想2次元フィールドの対応関係を推定する際、従来は、カメラの動きがズームを含めて自由度の高いものであったため、繰り返し計算が必要となり処理時間がかかっていた。そこで、現場で許容可能な制約として、ズーム無しで左右の横振りのみで映像を撮影する条件を導入し、繰り返し計算を不要とする方式にすることで処理の高速化を図った。具体的には、フィールド内のラインの角度を推定し、図6のように奥側と手前側のライン交点（4点）とそれらに対する射影変換行列を事前に複数箇所登録しておくことで、実際の処理時にはラインの角度を推定するだけで射影変換行列を取得できるようにした。

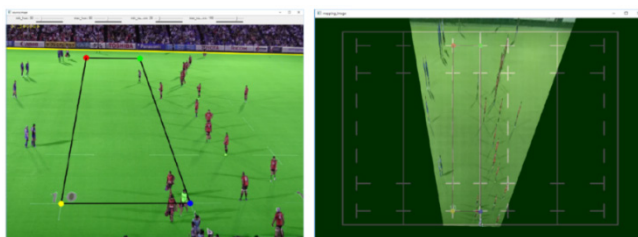


図6 入力映像と仮想2次元フィールド上の対応する4点の対応付け例

選手とボールの同時検出と合わせて、従来手法と処理時間の比較を行った結果を表1に示す。従来手法における2次元マッピングは、カメラの移動が少ない場合は表1にある通り0.3秒程度で動作するが、移動が大きい場合は誤差最小化のための繰り返し処理回数が増えるため1秒以上かかる場合もあった。これに対し提案手法では、カメラの移動量に関わらず0.035秒で処理可能となり、大幅な高速化を実現した。さらに、選手・密集領域検出とボール検出を同時に行うと共に、実行速度とメモリ効率に優れている深層学習フレームワークであるMXNet[8]で実装することで、検出処理全体の処理時間をほぼ半減した。これらの処理の高速化により、全体で毎秒5フレーム以上で安定して処理可能となり、実際の試合の現場で活用できる目処が立った。

表1 1フレームあたりの処理時間（秒）

	従来手法	提案手法
選手・密集領域検出	0.230	0.113
ボール検出	0.064	
選手追跡	0.215	0.015
2次元マッピング	0.332	0.035
その他	0.020	0.020
計	0.861	0.183

4. 他のスポーツへの応用

1章に述べたとおり、ラグビーは試合に出場する選手の数が1チーム15人と多く、接触プレーや密集プレーが頻繁

に発生するため、画像による分析は技術的にハードルが高いとされてきたが、今回開発したラグビー映像解析システムによりそれを可能にした。ラグビーを対象に精度良く解析可能な基盤技術を開発することにより、他のスポーツへの活用も期待される。

ラグビー同様にフィールド上の球技であるサッカーを対象に、本システムで開発した各要素技術を適用してみた。選手とボールの検出に際して、サッカー選手の新たな教示は行わずにそのまま適用した結果の一例を図7に示す。

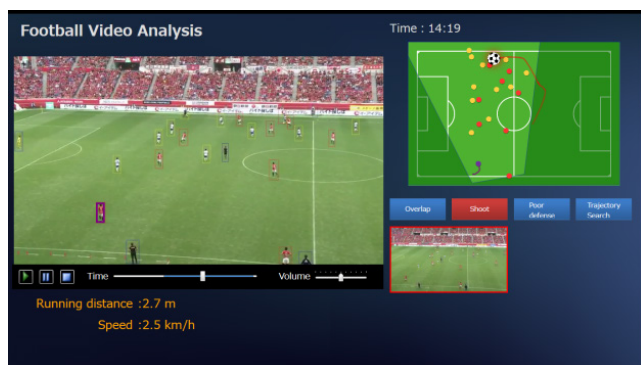


図 7 サッカー映像への応用例

サッカー映像に対しても精度良くボールと選手を検出・追跡し、位置推定できていることがわかる。上述の通り、サッカーでは複数カメラを用いた映像解析が導入され始めているが、本システムは1台のカメラで同様の映像解析が実現できるため、サッカーをはじめ、さまざまなスポーツへの応用が期待される。

5. まとめ

ラグビーを対象にした映像解析の基盤技術として、ディープラーニングを用いた選手とボールの同時検出、および仮想二次元フィールド座標変換処理の高速化に取り組んだ。その結果、仮想二次元フィールド上でのプレー分析とタグ付け作業の省力化が可能なラグビー映像解析システムの現場適用に目処をつけた。ラグビー向けに強化したこれらの基盤技術は、サッカーなどの他のスポーツはもとより、店舗や工場内の業務効率化など、スポーツ以外の産業分野への幅広い活用を検討していく。

参考文献

- [1] ChyronHego TRACAB, <http://chyronhego.com/tracab/>.
- [2] SPORTSCODE, <http://sportscode.jp/>.
- [3] 尾山匡浩, “ラグビー映像を対象としたシーンの自動推定,” VIEW2016 ビジョン技術の実利用ワークショップ, pp.128-131, 2016.
- [4] 廣瀬俊朗, “GPS データの効果,” 情報処理, Vol.57, No.9, pp.822-823, 2016.
- [5] 中洲俊信, 山地雄士, 柴田智行, 井本和範, “手書きスケッチによる直感的な資料検索システム,” ヒューマンインタフェース学会論文誌, Vol.18, No.3, pp.141-152, 2016.
- [6] 小林大祐 他, “ラグビー映像解析による戦術分析の効率化,”

第 23 回画像センシングシンポジウム (SSII2017), DS2-03, 2017.

- [7] S. Ren, et al., Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, In arXiv:1506.01497, 2015.
- [8] MXNet, <https://mxnet.incubator.apache.org/>.