

令和元年6月24日現在

機関番号：54502

研究種目：若手研究(B)

研究期間：2014～2018

課題番号：26750296

研究課題名（和文）球技スポーツにおけるコーチング支援のための映像処理手法の開発

研究課題名（英文）Development of Video Processing Techniques to Support Coaching in Ball Sports

研究代表者

尾山 匡浩（Oyama, Tadahiro）

神戸市立工業高等専門学校・その他部局等・准教授

研究者番号：80583749

交付決定額（研究期間全体）：（直接経費） 2,800,000円

研究成果の概要（和文）：本研究は、ラグビー映像を対象とし、コーチング支援のためにシーンの自動分類と人とボールの軌跡の可視化、およびこれを組み込んだアプリケーションの開発を目的として実施した。実際にラグビーの試合を撮影した映像に対して、人手で切り出しやアノテーション作業を実施し、それらの画像に対して深層学習の手法を適用することで、シーン分類などが可能か検討した。結果としては高精度に分類を行う手法を見つけることはできなかったため、アプリケーションの完成までには至らなかった。そのため、今後も継続して研究を行い、最終的にはアプリケーションとして公開していきたい。

研究成果の学術的意義や社会的意義

本研究では、アマチュアのラグビーチームにおけるコーチング支援のために、シーンの自動分類と人とボールの軌跡の可視化、およびこれを組み込んだアプリケーションの開発に取り組んだ。研究期間内において、撮影した映像に対して、人手で切り出しやアノテーション作業を実施した他、シーン分類手法などについて検討した。今後は、撮影した映像データの整理や追加のアノテーション作業を継続して実施した後、データセットとして公開したり、解析手法の提案やアプリケーションの開発を継続して行うことで、スポーツ映像処理の研究開発に貢献していきたい。

研究成果の概要（英文）：The purpose of this research was to develop an application for the automatic scene classification and the visualization of human and ball trajectories for coaching support. I examined the possibility of scene classification by manually cutting out and annotating images of actual rugby matches and applying a deep learning method to these images. As a result, the application is not yet complete because I have not been able to find a highly classifiable method. Therefore, I would like to continue my research and make it available as an application.

研究分野：コンピュータビジョン、生体信号処理

キーワード：ラグビー コーチング支援 シーン分類 人検出

## 様式 C-19、F-19-1、Z-19、CK-19（共通）

### 1. 研究開始当初の背景

近年、球技スポーツの現場でパソコンやビデオカメラ、タブレット端末などの ICT 機器が積極的に導入されてきている。具体的には、練習などにおいて1つ1つのプレーを撮影しそのフォームを即座にチェックしたり、試合映像からスカウティングを行い戦略を立てるなど、コーチングの補助ツールとして使用されている。また、安価にこれらの ICT 機器を入手することができるようになり、プロチームだけではなくアマチュアのチームにおいても広く使われている。

試合を撮影した映像から分析を行う際には、主に以下の2つの点に着目し、コーチングを行うことが多い。

(1) セットプレー（ボールが止まったプレー、サッカーで言えばスローインやコーナーキックなど）やそれ以外のプレーなど各シーン単位に注目して、個々の選手のパフォーマンスを分析する。

(2) 各選手の移動距離やプレーエリア、ボールの軌跡などを2次元平面上に可視化したり、数値化することにより、ゲーム全体の流れや傾向の分析などを行う。

(1)の各シーン別の分析については、試合全体を通して撮影した映像から解析したいシーンを抽出する必要があるが、所望のシーンを常時即座に抽出することは難しく、早送りや停止を繰り返しながら手作業で探し出すことが一般的である。このシーン抽出の作業を補助するため代表的なソフトウェアとして、「SportsCode」というものが存在する。これを用いることで映像中の各シーンのデータベース化が行えるため、即座に希望のシーンを閲覧することができ、ゲームやプレーの分析・評価を手軽に行える。しかしながら、各シーンのデータベース構築は、自動で行われるわけではなく、使用者が試合映像を見ながら1つ1つチェックし、ラベル付けなどを行わなければならないため、その際には多大な労力が必要となってくる。そのため、映像中から自動で所望のシーンを抜き出せる機能が実現できれば、それらの労力を削減でき、使用者の負担軽減にも繋がることが予想される。

スポーツ映像から自動的にシーンを抽出する取り組みは、現在までに多数報告されている。

[1]では、動画像だけではなく音響信号を組み合わせるハイライト映像を作成する技術が報告されていたり、それ以外にも映像情報のみを用いてサッカーのゴールシーンを抽出する研究も行われている[2]。しかし、それ以外の個別シーン（例えば、コーナーキックやスローインなど）についてはあまり検討されていない。そのため、ゴールシーン以外の個別のシーンなどについても抽出する技術が確立できれば、コーチング支援のツールとして役立てることができる。

一方、(2)に述べた選手やボールの動きの可視化を実現するためには、映像中から常に各選手やボールの動きを把握し追跡するとともに、複数台のカメラの映像から選手やボールの座標を算出する必要がある。「TRACAB」というシステムは、16台の専用カメラをスタジアムに設置することで自動的に選手の走行距離などを計測することが可能となっており、2010年の南アフリカW杯などで実際に使用されている。また、[3]では8台のカメラから高精度に選手を追跡できる報告も行われている。このように自動追跡のシステムは、商品化されているものも存在するが、非常に多くのカメラをスタジアムに設置しなければならず、限られた場所でのみ使用できないといった問題がある。またシステム自体も非常に高価で、アマチュアレベルのチームが採用することは困難である。そのため、比較的小数台(1~3台)のカメラを用いた追跡技術の確立が望まれている。このような少数台のカメラ映像を用いた選手やボールの追跡に関する研究は、現在までにいくつか行われている[4,5]が、実用化までには至っていない。

上述のように試合映像からシーンを抽出したり、選手を追跡したりする試みは各方面で行われているが、その対象となっている球技のほとんどはサッカーである。この理由として、競技人口が多く人気スポーツであること、また身体的な接触が少ない競技であることなどが挙げられる。一方で、ラグビー競技においても盛んに ICT 機器を用いたコーチングが行われているのを各所で目にしていく。しかしながら、ラグビー自体は、身体的な接触も多く競技ルールも複雑で、選手やボールの遮蔽が多発することもあり、なかなか研究対象となっていないのが現状であり、これらのコーチングツールの開発が望まれている。

### 2. 研究の目的

本研究では球技としてあまり着目されていないラグビーを対象とし、撮影した試合の動画像から自動的に各シーンを抽出したり、選手・ボールを自動で追跡する手法を確立することを目的とする。研究期間内に実現するべきこととして、以下の事柄が挙げられる。

- (a) 試合映像から自動的にシーン(例：スクラム、ラインアウト、モールなど)を抽出する手法の確立と実際のシステムの構築
- (b) 試合映像から自動的にボールや選手を検出・追跡する手法の確立
- (c) (b)で得られた情報を元に、選手の移動した軌跡を2次元平面上に描き視覚化するシステムの構築
- (d) (a)や(c)で開発したシステムを組み込んだアプリケーションの開発

### 3. 研究の方法

本研究では、ラグビーの試合を撮影した動画像から(1)スクラムやモールなどの各シーンを自動で抽出すること、(2)選手やボールの動きを試合全体を通して追跡することで選手の動きを可視化することを目的としている。以下、(1)と(2)を実現するための方法を説明する。

### (1) ラグビー動画の撮影

シーンの分類や選手の追跡を実施するためには、実際のプレー映像を利用する必要がある。しかしながら、サッカーと異なり、ラグビーのデータセットは公開されているものは少ない。そこで、様々なシチュエーション(土や芝などのグラウンドの種類による違い、ユニフォームの違いなど)において動画を計測する。本研究では、アマチュア向けのコーチングツール開発を目的としているため、一般の家庭用ビデオカメラを利用して撮影を行う。なお、本研究のシーン分類で対象とするのは、図1に示すように左からスクラム・ラインアウト・モール・ラックおよびその他の5つとする。画像サイズは1440×880、フレームレートは30fpsである。



図1：対象とするシーンの例

### (2) シーンの抽出

シーンの抽出を行うために、近年画像認識の分野で非常に多くの成果を挙げている深層学習の1つである畳み込みニューラルネットワーク(Convolutional Neural Network：CNN)を利用する。CNNは、画像の局所的な特徴抽出を行う畳み込み層と、特徴を局所ごとにまとめるプーリング層、最終的な出力を担う全結合層を組み合わせたニューラルネットワークの1つである。この畳み込みとプーリングにより特徴抽出と位置や回転に対する不変性を獲得することができる。図2に動画からのシーン抽出の流れを示す。以下、図中の各パートについて説明する。まず動画から各フレームに分割することで、カラーの静止画を得る。この画像に対して、スクラムなどの各プレーが映っているフレームのみを集め、これら画像を224×224画素にリサイズし、CNNへの入力画像として使用する。その後、CNNにより特徴抽出を行い、各シーンの推定を行う。

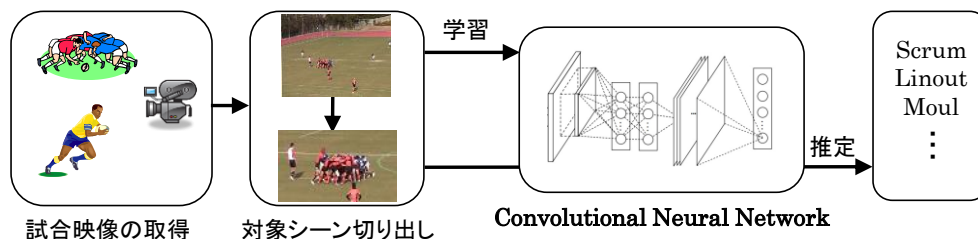


図2：映像からのシーン推定の流れ

### (3) 人とボールの検出

人とボールの検出を行うために図3に示す流れで処理を行う。まず、切り出されたフレームに対して深層学習の1つの手法として提案されているYOLOv2を用いてボールと人の検出を試みる。また、同時に白線の情報から手動でコーナー箇所の特定制を行う。これらの情報をもとにHomography 行列を利用して選手とボールを2次元平面にマッピングする。

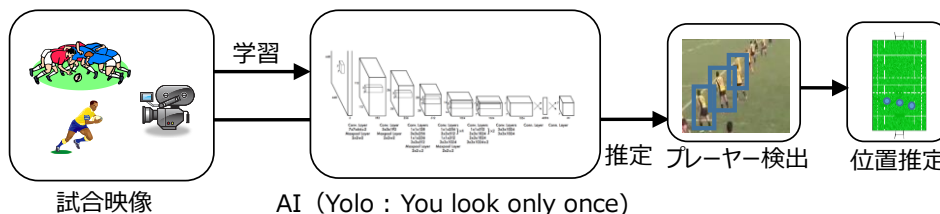


図3：映像からの人検出と位置推定の流れ

## 4. 研究成果

本研究の成果に関して、方法と照らし合わせて以下の項目ごとに説明する。

### (1) ラグビー動画の撮影

これまでに主に著者の所属するチームを対象として、約50試合以上の動画撮影を行い、その中の10試合を対象として、手作業でフレームの切り出しとシーンの抽出、アノテーションを実施した。本研究はこれらのデータを利用して以降の検証を実施した。

### (2) シーンの抽出

CNNを用いたシーン抽出に関して、以下の実験①と実験②を行い、シーンの推定が可能か評価した。

実験①：5種類(スクラム・ラインアウト・ラック・モール・その他)のシーンを分類

実験②：3種類(スクラム・ラインアウト・その他)のシーンを分類

シーンごとに学習用として 200 枚、評価用として 50 枚のそれぞれ異なるシチュエーションの画像を用意し、それぞれの実験を行った際の推定精度を表 1 に示す。これを見ると、5 種類のシーン分類の場合には平均で 42%程度の分類精度であり、3 種類に限った場合でも 78%程度の精度にとどまっている。このことから、現状では動画からのシーンの分類を実行するソフトウェアの開発は困難であることがわかる。また、誤分類した結果について見てみると、5 種類分類ではスクラム・モール・ラックを互いに誤認識していることが多く、3 種類ではスクラムとラインアウトをそれぞれ間違っ認識していることが多かった。

表 1: 各シーンの分類精度[%]

	Scrum	Lineout	Maul	Lack	Other
5 scene	36.0	46.0	38.0	40.0	50.0
3 scene	78.0	76.0	80.0		

### (3) 人とボールの検出および 2 次元平面への可視化

予め学習済みの Yolo のモデルに対して、切り出した人とボールの画像を用いて再学習を行い、動画像に対してそれぞれの検出と平面上へのプロットが可能か検討した。出力した動画の 1 シーンを結果として図 4 に示す。左が人検出後の結果、右が射影変換を行った結果となっている。これより、人やボールの検出に関しては可能であったが、写っているサイズが小さかったり遮蔽が起こったりすると、誤検出や未検出が多発する結果となった。

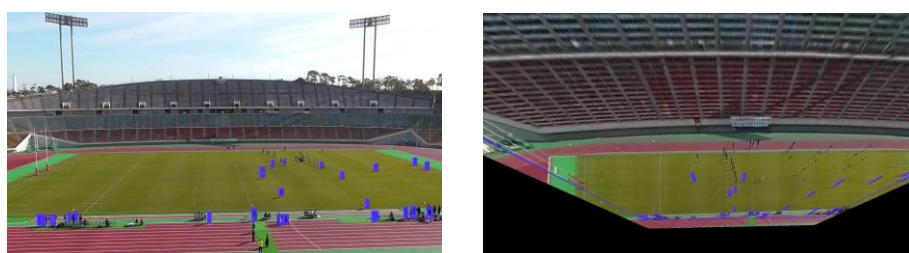


図 4: 人検出と射影変換実行後の結果

また、このモデルに対して、新たにスクラムとラインアウトの画像を学習用として再学習を行い、同様に人検出と同時にシーンの検出が可能か検討を行った。その結果の一例を図 5 に示す。左側の 2 枚がスクラムとラインアウトの箇所を検出できた画像であり、右側の 2 枚は誤認識の例となっている。このように動画中からシーンの対象となる箇所を特定することも可能な場合があったが、高い精度で検出することはできなかった。



図 5: 人とシーン推定実行後の結果

本研究では、ラグビー映像を対象とし、コーチング支援のためにシーンの自動分類と人とボールの軌跡の可視化、およびこれを組み込んだアプリケーションの開発を目的として実施した。実施当初は、ラグビーを対象とした上述のようなシステムは、この世に存在していなかったが、現在では特定のチーム用のソフトウェアとして開発が行われている報告がある[6, 7]。本研究では、特定チームではなく一般の競技者用に公開することを目的として研究開発を行ってきたが、完成までには至らなかった。今後は、撮影した映像データの整理や追加のアノテーション作業を継続して実施し、可能であればデータセットとして公開していきたい。これによって、より多くの研究者などによる開発が加速すると考えられる。また、著者による研究も継続し、公開できるレベルにまでなればアプリケーションとして公開していきたいと考えている。

### <引用文献>

- [1] 宮内ら, “テキスト・音声・画像の協調的処理による放送型スポーツ映像におけるハイライト検出とインデクシング”, 信学論 D-II, pp. 1692-1700, 2002
- [2] 山本ら, “サッカー映像のシーン自動解析の研究”, 信学会 PRMU 研究会, pp. 73-78, 2005
- [3] H. B. Shitrit et al., “Tracking Multiple People under Global Appearance Constraints”, ICCV 2011
- [4] 片岡ら, “単眼カメラを用いたサッカー戦術解析のための複数選手とボールの追跡”, 画像電子学会誌, 2012
- [5] T. D’ Orazio et al., “A review of vision-based systems for soccer video analysis”, Pattern Recognition, Vol. 43, pp. 2911-2926, 2010

- [6] 小林 他, “ラグビー映像解析による戦術分析の効率化”, SSII2017, 2017  
[7] 大内 他, “ラグビー映像解析システムの開発”, 信学論 B, Vol. J100-B, No. 12, pp. 941-951, 2017

#### 5. 主な発表論文等

[学会発表] (計 5 件)

- [1] 尾山匡浩、中尾大輝、“ラグビー動画からのシーン抽出方法に関する一検討”、動的画像処理実利用化ワークショップ(DIA)、2015  
[2] 尾山匡浩、“ラグビー映像を対象としたシーンの自動推定”、ビジョン技術の実利用ワークショップ(ViEW)、2016  
[3] 尾山匡浩、“CNNを用いたラグビー映像のシーン推定”、映像情報メディア学会 2016 年冬季大会、2016  
[4] 尾山匡浩、“ラグビーにおけるコーチング支援のための映像解析に関する研究”、日本フットボール学会 16th Congress、2018  
[5] 尾山匡浩、“ラグビー競技におけるコーチング支援のための映像解析技術に関する研究”、日本コーチング学会第 30 回学会大会、2019