

集団スポーツの位置情報データを用いた 機械学習による複数人協力プレーの識別

Identification of multi-player cooperative play by machine learning using group sports position information data

藤井 慶輔^{*1} 河原 吉伸^{*1,2} 稲葉 優希^{*3} 元安 陽一^{*4} 山本 裕二^{*5}
Keisuke Fujii Yoshinobu Kawahara Yuki Inaba Yoichi Motoyasu Yuji Yamamoto

^{*1} 理化学研究所 ^{*2} 大阪大学 ^{*3} 国立スポーツ科学センター ^{*4} 長崎国際大学 ^{*5} 名古屋大学
RIKEN Osaka University Japan Institute of Sports Sciences Nagasaki International University Nagoya University

With the development of measurement technology in recent years, tracking systems that measures position information of all players and balls in a sports game can be used, but it is not fully utilized for more advanced and tactical plays. In this research, we identified multiple cooperative play characteristics of group sports using support vector machine and deep learning using convolutional neural network combining space and temporal information, and discuss an appropriate identification method.

1. はじめに

球技などの集団スポーツにおいて、選手がチームメイトと協働しチームワークが機能する様相は、未だ自動で識別できない。近年では計測技術の発達により、トラッキングシステムと呼ばれるカメラ等により全選手とボールの位置を計測するシステムが様々なスポーツにおいて導入されている。例えばアメリカのプロバスケットボールリーグ NBA において 2013-2014 シーズンから導入した SportVU というシステムは、全 30 チームのホームアリーナに 6 台のカメラを設置し、各選手の位置情報を 25 Hz で記録して 90 秒以内にレポートをオンラインで各チームに配信できる[Lucey 14]。屋外のサッカーやラグビー等では無線システムと通信する無線センサや GPS を用いた選手位置計測が主流である[Stevens 14]。これらの測定システムから得られた情報を用いて行われる分析として、移動距離・速度などのフィットネス指標[Sampaio 15]や、攻撃選手の得点期待値 [Goldman 13]などの分析が行われている。しかし、これらの分析は選手間の相互作用を考慮せず、単に個人の位置データに基づく分析のため、チームプレーが機能する様相を捉えられていない。

一般にスポーツのチームプレーの多くは名前がつけられているが、同じ名前のプレーでも実際は相手に対応して動くため時空間的に全く同じプレーにはならず、分類することは難しい。教師なし学習におけるチームプレー分類の研究としては、再帰型ニューラルネットワーク(RNN) [Wang 16]や潜在ディリクレ配分法[Miller 17]を適用した研究があるが、選手間の相互作用を効果的に定量化したとは言えず、チームプレーが行われたかどうかの解釈ができない。そこでまずは最小限の人数で構成されたラベル可能なチームプレーに関して、教師あり学習により識別する必要があると考えた。

集団スポーツのバスケットボールにおける、攻撃者が壁となつて守備者の移動を合法的に妨害するスクリーンプレーと呼ばれる基本的なチームプレーは、上記の適切な題材である。これまでの研究では、ボール近くの限定的なプレー(オンボールスクリーン)のみを、主に選手間の距離を特徴ベクトルとしたサポート

ベクターマシン(SVM)を用いて識別した[McQueen 14]。しかし、選手は様々な場所でシュートを打つための空間を作り出すために、ボールから離れた場所においてもより多様で同時多発的なスクリーン(オフボールスクリーン)を使用し、これは自動検出するのがより困難になる。我々は以前、チームプレーを評価するため全種類のスクリーンプレーを分析した[Fujii 16]が、部分的に目視を用いたため膨大なデータに対しては自動的に識別できなかった。この研究で用いた特徴量としては、先行研究[McQueen 14]の選手間距離だけでなく、選手個人の動きや地理的情報も用いており、全種類のスクリーンプレーを抽出するにはこれらの情報が重要であると考えた。そこで本研究では、選手個人の動きと地理的情報を加えた特徴ベクトルを用いた SVM により、オンボールとオフボールのスクリーンを自動的に識別する。

一方でこの問題を動作認識として考えると、近年注目の畳み込みニューラルネットワーク(CNN)等の深層学習は特徴抽出を自動で行い、ヒト個人の動きを詳細に識別することが可能になってきている[Simonyan 14]。複数人の動作に関しては、観客の拍手[Conigliaro 15]を識別する研究等があるが、選手間の協力プレーでは、より相互作用が複雑なため未だ確認されていない。そこで比較対象として、近年動作認識で優れた成績を出した時空間情報を融合した CNN [Feichtenhofer 16]を用いて、我々はスクリーンプレーを識別したが、その識別性能は SVM より明らかに低かったため、4章の考察にて簡単に議論することと定める。本研究の目的は、集団スポーツに特徴的な複数人の協力プレーに関して適切な識別手法について議論することにある。

2. 方法

分析では、2015年にSTATS SportVUシステムで記録されたある国際試合の選手・ボール位置データ(毎秒 25 フレーム)を使用した。分析したプレー時間は 94 分で、2 チームの合計得点は 316 点であった。協力プレーの自動識別は先行研究[McQueen 14]に従いデータ分割、特徴抽出、および分類の 3 つのステップで行った。

連絡先: 藤井 慶輔, 理化学研究所・革新知能統合研究センター, 大阪府吹田市古江台 6-2-3, keisuke198619@gmail.com

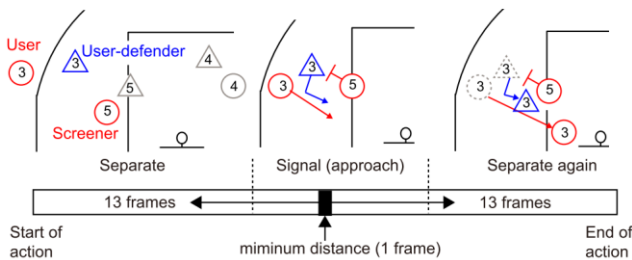


図1 スクリーン、アクション、シグナルに関する説明

2.1 データ分割

データ分割では、図1に示されるスクリーナーと2人のユーザー候補(壁となる/壁を使う攻撃選手)およびユーザー守備選手からなる、セグメント化された1829個の位置データセット(アクション)を自動検出した。まず全ての攻撃選手をスクリーナー候補として、各スクリーナー候補に対して最も近い3人の守備選手と2人の攻撃選手をそれぞれユーザー守備選手候補とユーザー候補と定義した。そして次の条件を満たした場合、スクリーンが起こる可能性のあるシグナルと定義した: (1)スクリーナー候補とユーザー守備選手候補が1.2m以内で接近し、(2)ユーザー守備選手は、ユーザー候補にとって最も近い選手である。シグナルの時間幅は、スクリーナー候補とユーザー守備選手候補が最小距離となった前後13フレームとして定義された(図1)。また、信号が極端に短い場合(3フレーム、0.12s未満)は分析から除外し、時間的に隣接する同一アクションは融合した。

全てのアクションについて、競技経験者はスクリーンであるか否かのラベリングを行った。後に示すように性能を高めるため、2つのプレー(オン/オフボールスクリーン)を事前に分類した。オフボールスクリーンは、ボールに事前に関与しないスクリーンを示し、スクリーナー及びユーザー候補者が図1の最小距離の時刻でボールを保持していない場合に定義した(逆の場合はオンボールスクリーンと定義した)。モーションキャプチャデータを用いた事前の分析[Fujii 16]では、オンボールスクリーンとオフボールスクリーンの97.2%(140/144)がアクションの一部であることが示されている。しかし、すべてのアクションの41.2%(140/340)しか実際にはスクリーンではなかった。本研究における実際のゲームでのデータセットでは、この確率はさらに24.7%(451/1829)に減少した。そこで次に、以下の特徴ベクトルを抽出し、SVMによる分類を行った。

2.2 特徴抽出

特徴抽出では、各選手の移動距離と地理情報を先行研究の特徴ベクトル[McQueen 14]に加えた。これらの変数は誤識別率に基づいて選択した。移動距離については、アクションの開始から終了までの間の4人の選手(スクリーナー、ユーザー守備選手、ユーザー候補1と2)の距離を使用した。地理情報としては、

スクリーンが起こった領域(図2a)、4選手とボールの距離、3選手の角度(図2b)を使用した。スクリーンが起こった領域は、図1の最短距離時のスクリーナー位置を元に9つの領域に分類した(図2a)。図1の最小距離時の4人に関する選手の位置を使用して選手ボールの距離を計算した。3選手の角度は、図2bに示すユーザー(2人の候補)、スクリーナー、ユーザー守備選手の角度として定義した。次に先行研究で用いられた選手間距離[McQueen 14]を修正し、9つの距離を選択した(スクリーナーとユーザー守備選手、スクリーナーと2人のユーザー、ユーザー守備選手と2人のユーザー、リングと4人の距離)。この9種類の距離ごとに5つの特徴ベクトル(最小値、図1前半・後半の振幅(最大-最小値)、前半・後半の平均値)を計算した[McQueen 14]。以上より、計65次元の特徴ベクトルを使用した。分析では、移動距離情報(12次元)と地理情報(8次元)を加えたり除いたりすることにより、識別性能(図3)をテストした。

2.3 識別

識別にはガウスカーネルを用いたソフトマージンSVMを使用した。訓練セットは1218アクション(オンボール:496、オフボール:722アクション)を用い、それとは別にテストセットを611アクション(オンボール:249、オフボール:362アクション)用意して、10分割交差検証を使用した。異なる乱数で5回識別し、誤識別率の中央値に基づいて制約パラメータを決定した。

以前に報告された識別器[McQueen 14]と性能を比較するために、ROC曲線に基づくAUC(Area Under the Curve)、再現率および適合率を計算した。ROC曲線は、真陽性率の累積分布関数を偽陽性率に対してプロットすることによって作成される。これを用いて識別性能を評価するため、ROC曲線の下面積(AUC)を計算した。再現率は、真陽性と真陰性の和に対する真陽性の比(真陽性率に等しい)として定義され、適合率は、真陽性と真陰性の和に対する偽陽性の比率として定義された。再現率-適合率のトレードオフ曲線も累積分布関数を用いてプロットした。適合率-再現率のトレードオフを評価するために、Fスコアを $(2 \times \text{適合率} \times \text{再現率}) / (\text{適合率} + \text{再現率})$ として計算した。

3. 結果

まずROC曲線(5回のうち性能が中央値にあたるもの)を図3に示す。すべての情報を用いて得られたオンボール(図3a)およびオフボール(図3c)スクリーンのAUC中央値は、それぞれ0.907および0.824であった。オンボールスクリーンは、オフボールスクリーンと比較してパターン数が限られているため、オフボールスクリーン識別器よりも正確で、以前のオンボールスクリーン識別器[McQueen 14](0.80-0.85)よりも正確であった。

特に、オンボールスクリーンのAUC中央値は、地理学的情報なしでは0.898であったのに対し、移動距離情報なしでは0.867よりも減少した。対照的に、オフボールスクリーンの中央値AUCは、移動距離情報なしの0.814と比較して、地理情報なしで0.775以上減少した。

再現率と適合率との間のトレードオフは、図3bおよび図3dに示した。全ての情報を用いた識別器では、ボール(図3b)およびオフボールスクリーン(図3d)のFスコアはそれぞれ0.793および0.615であった。前者に関しては、先行研究の限定的なオンボールスクリーン識別器[McQueen 14](Fスコア:0.810)に近かった。

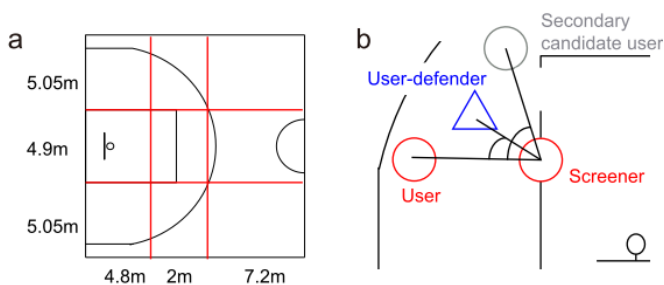


図2 (a)スクリーンの領域。(b)3選手の角度。

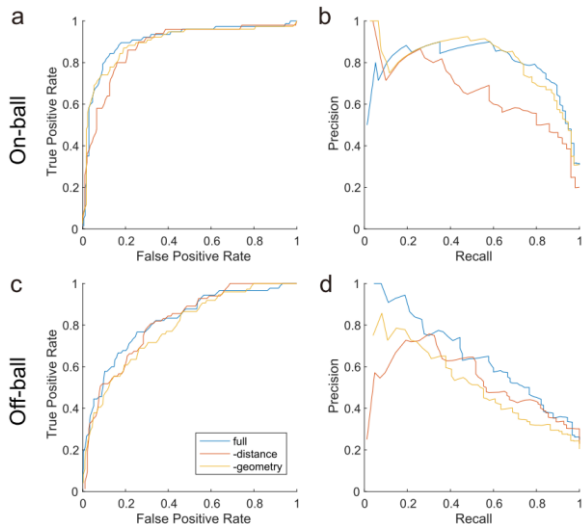


図 3: 移動距離および地理的情報がある場合とない場合のボール(a, b)・オフボール(c, d)スクリーンの ROC 曲線(a, c)および再現率-適合率トレードオフ(b, d)。

次に図 4 に偽陰性(見逃し)および偽陽性(誤警報)による誤識別の具体例をそれぞれ図示した。例えば、いくつかの偽陰性のアクションでは、スクリーンが行われたように見えたが、ユーザー守備選手はスクリーンを避け、比較的長い距離を移動した(図 4a)。また誤警報の一部では、スクリーナーはビデオでみると全くスクリーンを行っていないアクションが存在した(図 4b)。このような場合は今回のような 1 選手につき 1 座標点のデータのみから真のスクリーンと区別することは不可能である。両方の誤識別の例では、特徴ベクトルの移動距離が関与していた。しかし全体としては、この移動距離情報は識別器の性能の改善に貢献した(図 3)。

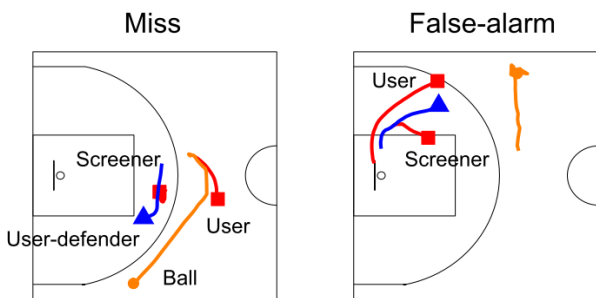


図 4: 偽陰性(見逃し)および偽陽性(誤識別)の例。各記号は、アクションの開始時のプレイヤーとボールの位置を示す。実線は、プレイヤーとボールの 2 秒の移動軌跡を示す。

4. 考察

本研究では、SVM を用いて集団スポーツの協力プレーを識別した。入力する特徴ベクトルとして、選手の移動距離、地理学的情報、選手、リング、ボールの距離の情報を用いて、スクリーンプレーと呼ばれる協力プレーを検出した。特に、関連する選手の移動距離およびボールやコートを含む地理学的情報が、識別器の性能を向上させた。従来から用いられてきた選手間距離[McQueen 14]に追加されたこの特徴ベクトルは、ボール周辺に限らず、様々な場所、形式のスクリーンプレーを認識し、5 人の攻撃選手のチーム攻撃過程を反映しているため、チーム攻撃のための重要な情報になる(頻度としては 10 分間の攻撃で、

オフボールスクリーンが 26.4 回、ボールスクリーンが 21.5 回記録されるほど重要なプレーである)。この協力プレーではボールを保持する重要な選手だけでなく、ボールから離れた攻撃選手がチームメートをフリーにするために守備選手を妨害するため、重要な役割を果たす。今後大規模なデータセットを用いることにより、オン/オフボールスクリーンが発生する状況等を分類することによって、チーム攻撃の特性を分類できること等が期待される。

一方で近年注目の深層学習、特に高い動作認識性能を示す時空間情報を融合した CNN[Feichtenhofer 16]に xy データから復元した動画データを入力して今回の集団動作を識別したが、トレーニングセットのエラー率が 20%より高く、明らかに今回の SVM(テストデータに対してオンボール: 18.1%、オフボール: 22.9%)より低い性能を示した。この CNN で用いた時間情報とはオプティカルフローから得た動きの輪郭で、空間情報は画像の見目の特徴を示すが、これらは xy 位置データを用いれば容易に求められると考えられる。本研究の課題で求められる特徴量としては、既知の複数点位置の時空間的関係性から、全体の(協力プレーという)機能を識別する能力と言える。CNN を用いた個人の動作認識においては、複数の関節点から動きの種類を識別できるが、空間的な構造を無視したとしても、カーネル法を用いた動的モード分解[Kawahara 16]等を用いて分類が可能である。本研究では各選手が物理的に接続されていないので、時々刻々と変化する拘束条件で位置関係が変化する、つまり時間と空間で異なる特徴の相互作用が関与するため識別が難しいのかもしれない。以上の困難な点はその他のゴール型集団競技(サッカーやラグビー等)、あるいは野球やバレーボール等のチームプレーに関しても同様である可能性が高く、今後はこの点を明らかにしていく必要がある。

応用例として、本研究により提案されたシステムは、ボールから離れた攻撃選手の動き(オフボールの動き)を視覚化することができるので、目立ちにくいオフボールの動きとその寄与に焦点を当てるようコーチと選手を動機付けられる可能性がある。その結果、チーム攻撃の質を間接的に向上させることが期待できる。発展として、得点やスクリーンに対する守備[Fujii 16]、あるいはスクリーンの質の評価などがこのシステムに加わると、より有益になると考えられる。

謝辞

本研究は科研費特別研究員奨励費 26-407 の助成を受けて実施したものです。データ提供等において、STATS LCC. の Charlie Rohlf 様に感謝申し上げます。

参考文献

[Conigliaro 15] Conigliaro, D., et al. The s-hock dataset: Analyzing crowds at the stadium. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2015). pp. 2039-2047.

[Fujii 16] Fujii, K., et al. Resilient help to switch and overlap hierarchical subsystems in a small human group. *Scientific reports* 6 (2016).

[Feichtenhofer 16] Feichtenhofer, C., Pinz, A. & Zisserman, A. Convolutional two-stream network fusion for video action recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2016) pp. 1933-1941.

-
- [Kawahara 16] Kawahara, Y. Dynamic Mode Decomposition with Reproducing Kernels for Koopman Spectral Analysis. *Advances in neural information processing systems*. (2016). pp. 911-919
- [Lucey 14] Lucey, P., et al. How to get an open shot: analyzing team movement in basketball using tracking data. *MIT Sloan Sports Analytics Conference* (2014).
- [Miller 17] Miller, A.C., & Bornn, L. Possession Sketches: Mapping NBA Strategies. *MIT Sloan Sports Analytics Conference* (2017).
- [Stevens 14] Stevens, TGA., et al. Measuring acceleration and deceleration in soccer-specific movements using a local position measurement (LPM) system. *International journal of sports physiology and performance* 9.3 (2014) pp.446-456.
- [Sampaio 15] Sampaio, J., et al. Exploring game performance in the national basketball association using player tracking data. *PLoS One* 10.7 (2015).
- [McQueen 14] McQueen, A., Wiens, J., & Guttag, J. Automatically recognizing on-ball screens. *MIT Sloan Sports Analytics Conference*. (2014).
- [Goldman 13] Goldman, M., & Rao, JM. Live by the Three, Die by the Three? The Price of Risk in the NBA. *MIT Sloan Sports Analytics Conference*. (2013).
- [Simonyan 14] Simonyan, K., and Zisserman, A. Two-stream convolutional networks for action recognition in videos. *Advances in neural information processing systems*. (2014). pp. 568-576
- [Wang 16] Wang, KC., & Zemel, R. Classifying NBA offensive plays using neural networks. *MIT Sloan Sports Analytics Conference*. (2016).