

H-008

バスケットボール放送映像におけるコート特徴点を利用したカメラの動き推定

Camera Motion Estimation Using Feature Points in Basketball Broadcast Video

田中 大晴[†]
Taisei Tanaka黒木 修隆[†]
Nobutaka Kuroki沼 昌宏[†]
Masahiro Numa

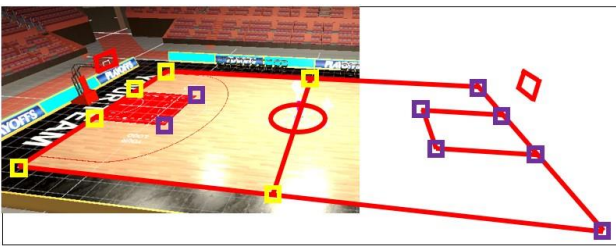
1. はじめに

画像処理によるスポーツ解析の取り組みは進んでおり、大規模な国際大会では複数台のカメラで選手やボールを計測することが少なくない。一方、一つの放送映像からスポーツ解析を行う技術があれば、録画済みの試合を解析できることや、小規模な地方大会でも利用できるなど、コスト面で非常に有益である。

2. 従来手法

放送用カメラは計測用のカメラと異なり、カメラマンがいつパン・チルト・ズームの操作を行うかわからない。すなわち、カメラの姿勢パラメータと内部パラメータが動的に変化する。大石ら [1] はサッカーの試合映像においてこれらのパラメータを推定するため、コート特徴点を YOLOv4 で追跡する手法を提案した。例えば図 1 のようにコートの特徴点を 4 つ以上検出することにより、カメラの動きを推定することが可能になる。

しかし、サッカーの場合は大きなコートを広い画角で放送するシーンが多いのに対して、バスケットボールの場合はコートが小さく、図 2 のように局所的にズームするシーンが多い。この場合、特徴点不足によってカメラパラメータの推定、特に距離とズーム倍率の推定が難しくなる。



画面内のコート特徴点からカメラの動きを推定

図 1 大石らの手法によるカメラの姿勢推定



図 2 カメラの距離とズーム倍率の推定が難しくなる例

3. 提案手法

3.1 コートとゴールの特徴点検出

提案手法ではコートおよびゴール上の特徴点を図 3 のように定義する。これらの特徴点を YOLOv7 によってバウンディングボックス $\mathbf{b} = (x, y, w, h)$ の形で検出する。そして、中心座標を $\mathbf{m} = (x, y)$ とする。少なくとも 4 個以上の特徴点 \mathbf{m}_i を検出し、それらをカメラパラメータの計算に用いる。

3.2 検出結果を用いたカメラパラメータの算出

カメラパラメータの算出の際には、Zhang の手法 [2] を応用する。これは、ワールド座標 $\mathbf{w} = (x, y, z)$ から射影変換 f によって計算されるピクセル座標と、実際に計測されたピクセル座標 \mathbf{m} の誤差がなるべく小さくなるようなカメラパラメータ \mathbf{p} を

$$\bar{\mathbf{p}} = \operatorname{argmin}_{\mathbf{p}} \sum_{i=1}^n \|f(\mathbf{p}, \mathbf{w}_i) - \mathbf{m}_i\| \quad (1)$$

のように求める手法である。このような手法は Bundle Adjustment と呼ばれる。YOLOv7 で検出した 4 個以上の特徴点について \mathbf{w}_i と \mathbf{m}_i を求め、それらを式(1)に代入することでカメラパラメータ \mathbf{p} を求める。 \mathbf{p} の中にはカメラの姿勢を表す回転行列 \mathbf{R} と平行移動ベクトル \mathbf{T} の他、焦点距離 d (ズーム倍率に連動する) も含まれている。一般に d が可変のカメラでは光軸 (視線) 方向の距離の推定が難しくなる。

3.3 競技場内のカメラ位置推定

カメラの設置位置とは、図 4 に示すようにワールド座標系におけるカメラ座標系の原点 $(0,0,0)$ のことである。よって設置位置 $Q = (Q_x, Q_y, Q_z)$ は、

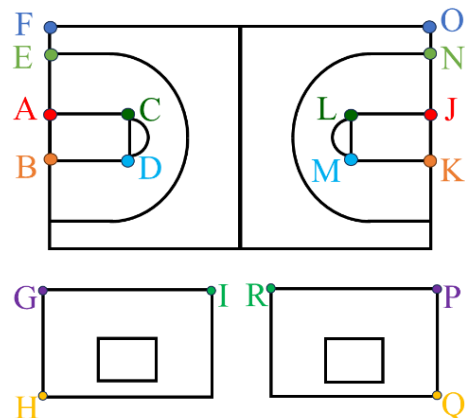


図 3 特徴点の定義

[†] 神戸大学大学院工学研究科

Graduate School of Engineering, Kobe University

$$\begin{bmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{bmatrix} = [R|T]^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (2)$$

のように求められる。

通常、放送用カメラは位置が固定されているので、 Q が試合中に変化することは考えにくい。そこで提案手法では図6のようにコートおよび右面から推定された Q の分布を基にカメラの光軸の交点を求める。具体的には、半径 15cm 以内の密度が最も高い点 Q_{max} を求め、カメラの設置位置とする。

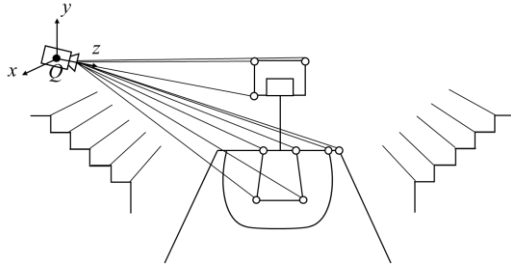


図4 競技場内のカメラ位置

表1 コートとゴールの特徴点の検出精度

	適合率	再現率	F 値		適合率	再現率	F 値
A	0.949	1.000	0.974	J	1.000	1.000	1.000
B	1.000	0.746	0.855	K	1.000	0.969	0.984
C	0.928	1.000	0.963	L	0.985	1.000	0.992
D	1.000	0.907	0.951	M	1.000	0.939	0.969
E	1.000	0.786	0.880	N	0.987	0.867	0.923
F	0.999	1.000	0.999	O	0.736	1.000	0.848
G	0.990	1.000	0.995	P	0.985	1.000	0.992
H	0.925	0.933	0.929	Q	0.993	1.000	0.996
I	0.992	1.000	0.968	R	0.972	1.000	0.986
平均					0.970	0.956	0.962

4. 評価実験

4.1 実験内容

約 11 秒の試合映像から YOLOv7 を用いて特徴点 A-R の追跡を行った。そして全フレームにおいてカメラパラメータ R , T , および d を算出し、最後にカメラの設置位置 Q_{max} を推定した。

4.2 実験結果と考察

コートとゴールの特徴点の検出結果の一例を図5に示す。また、それらの検出精度を表1に示す。適合率は平均 97.0%、再現率は平均 95.6%であった。適合率がどの特徴点も高く、誤検出が少ない点はカメラ位置推定において望ましい結果である。

推定されたカメラ位置 Q の分布 (x - y 面) を図6に示す。 Q が光軸 (視線) 方向に前後しているのは、ズーム中の距離の推定が難しいためである。この問題に対処するために交点の座標を

求めた結果、 $Q_{max} = (-28.9, 1.13, 9.72)$ となった。すなわち、このカメラはセンターサークルより後方に 28.9m、高さ 9.72m の位置にあると推定される。この値は会場の規模から考えて妥当であるものの、真値との比較については今後の課題である。

5. まとめ

本研究では、バスケットボール放送映像における競技場内のカメラ位置、角度および内部パラメータの推定を目的として、コートおよびゴール上の特徴点を YOLOv7 によって検出する手法を提案した。また、推定が難しいとされる距離とズーム倍率の算出については、左右に向けたカメラの光軸の交点を求める手法を提案した。特徴点の検出については、適合率 97.0%、再現率 95.6% の精度が得られた。また、カメラ位置については従来手法よりも信頼度の高い座標を得ることが可能になった。今後は得られたカメラパラメータを利用してワールド座標系の選手の動きを正確に算出することで、ゲーム分析を支援する予定である。

6. 参考文献

- [1] 大石一哉, 玉衛淳輝, 黒木修隆, 平野健介, 沼 昌宏, “サッカー中継映像における YOLOv4 を用いたフィールド特徴点と選手の同時検出,” 電気情報通信学会 技術研究報告, vol. 121, no. 423, MVE2021-49, pp. 55-59, 2022 年 3 月.
- [2] Zhengyou Zhang. A flexible new technique for camera calibration. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(11):1330-1334, 2000.



図5 特徴点の検出結果の一例

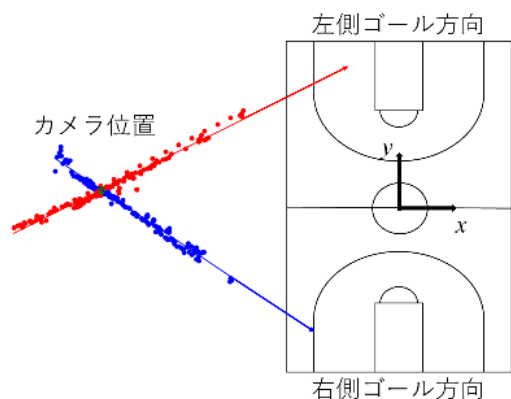


図6 カメラ位置の座標の推定結果