

修士論文概要書

Master's Thesis Summary

Date of submission: 01/27/2025 (MM/DD/YYYY)

専攻名（専門分野） Department	情報理工・ 情報通信専攻	氏名 Name	福田 大翔	指導 教員 Advisor	渡辺 裕 印 Seal
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	CD 5123F103-4		
研究題目 Title	形状既知物体を利用した単眼映像からの三次元空間推定及び動体軌道推定手法の研究 Research on 3D Space Estimation and Motion Trajectory Estimation from Monocular Images Using Geometric Objects				

1. まえがき

近年、映像解析技術はスポーツや医療など多分野で活用され、特に動体軌道推定や三次元空間における物体認識に注目が集まっている。スポーツ分野では、競技によりフィールドや道具が厳密に規定されているため、各競技に特化した高精度な映像解析技術を導入している。例えば、プロスポーツ競技では ABS[1]や VAR[2]などの判定システムが実用化されているが、コストや環境面でアマチュア競技への普及は難しい。また、アマチュアスポーツ界において審判不足や誤審リスクも課題であり、安価で簡易な判定補助技術の必要性が高まっている。

そこで本研究では、単眼カメラ映像のみを使用するという制約のもと、三次元空間推定と動体軌道推定を行う手法を提案する。スポーツ競技における判定補助技術に適用することを目的とし、そのルールや特性を活用することで、推定精度の向上を目指す。野球のボールストライク判定を事例に、ストライクゾーンと三次元ボール軌道の推定を行い、既存手法との比較で有効性を示す。

2. 透視投影変換

本研究では、1 台の単眼カメラから三次元の情報を得るために透視投影変換を利用する。透視投影変換は実世界に存在する物体を二次元平面に投影する方法である[3]。\$(X,Y,Z)\$を実世界座標系の三次元座標、\$(u,v)\$を画像平面に投影された点の座標とする。また、画面中心座標を\$(c_x, c_y)\$、カメラのピクセル単位で表される焦点距離を\$(f_x, f_y)\$とする。実世界での物体の三次元座標と画像平面での二次元座標の関係はカメラの回転行列と並進ベクトル、正規化パラメータ\$s\$を用いて、式(1)で記述される。

$$s \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

3. 提案手法

3.1. 概要

本研究では、単眼カメラ映像を用いて三次元空間推定と動体軌道推定を行う手法を提案する。具体的には、野球のボールストライク判定を対象に、形状

既知物体であるボールやホームベースを活用し、高精度なボールストライク判定を可能にする。提案手法の概要を図 1 に示す。本提案は、カメラキャリブレーションブロック、ストライクゾーン推定ブロック、およびボール軌道推定ブロックの三つのブロックで構成する。カメラの位置を球審視点に設定することで、判定時に必要となるボールの形状や動きを正確に捉えることが可能となり、1 台のカメラで三次元情報を取得できる。また、三次元情報の取得に必要なカメラのパラメータは、ホームベースの大きさ情報を活用し、近似により算出する。特定の三次元空間であるストライクゾーンの推定には、人物の姿勢推定手法である BlazePose を用いて打者骨格を推定し、骨格座標を基に算出する。算出したストライクゾーンは、ホームベースの大きさ情報を用いて、再スケール化する。また、物体検出アルゴリズム YOLOv8 をベースとして作成した野球ボール専用の検出器と透視投影変換を用いて、動体軌道であるボール軌道を算出する。最後に、算出したストライクゾーンとボール軌道を用いて、ボールストライク判定を行う。

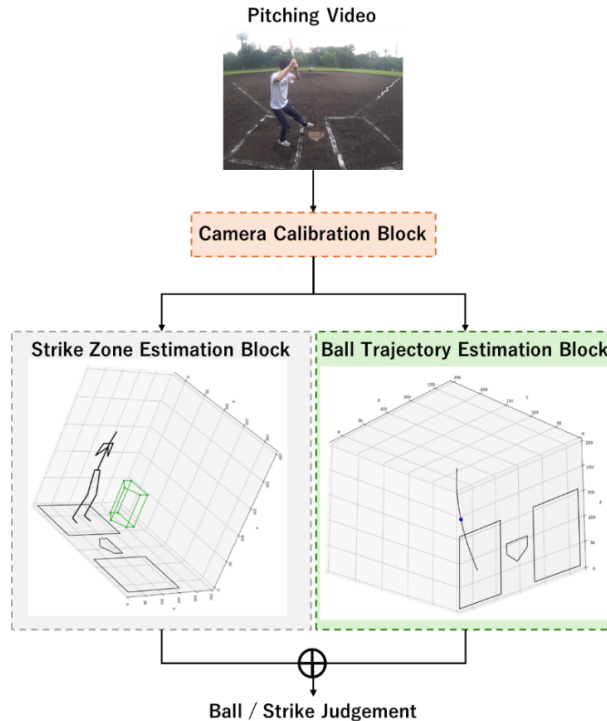


図 1 提案手法概要

3.2. カメラキャリブレーション

ホームベースの大きさ情報と画面上の大きさを基に透視投影の逆変換を用いて、カメラの各パラメータを推定する。まず、カメラの外部パラメータはカメラ位置の固定により自動的に求まる。そして、ホームベースという大きさが既知である物体を活用し、内部パラメータを高精度で近似することができる。

3.3. ストライクゾーン推定

打者依存のストライクゾーンの推定のため、Blaze Pose を用いて打者骨格を検出する。検出した座標値からストライクゾーンの上端と下端を算出する。ストライクゾーンの上端は両肩と両尻のY座標の平均値とし、下端は両膝のY座標の平均値とする。算出したストライクゾーンは、ホームベースの大きさ情報とホモグラフィ変換を使用し、再スケール化をする。

3.4. ボール軌道推定

取得した投球映像に対し、YOLOv8 をベースに作成した野球ボール専用の検出器を利用してボールを検出する。検出したボールの画面内での大きさと実際の大きさの対応関係よりボールのZ座標を推定する。推定したZ座標を基に、検出したボールのxy座標を実世界座標に変換する。これを各検出結果に対し実行することで軌道を推定する。実際のボールサイズをRBS、画面内でのボールサイズをSBSとすると、ボールのZ座標は式(2)で近似できる。ここでは、モーションブラーの影響を考慮し、画面内でのボールサイズはx軸方向を利用する。また、ボールのXY座標は式(3)、式(4)で近似できる。

$$Z \approx \frac{f_x \cdot RSB}{SBS} \quad (2)$$

$$X \approx \frac{Z \cdot (x - c_x)}{f_x} \quad (3)$$

$$Y \approx \frac{Z \cdot (y - c_y)}{f_y} \quad (4)$$

4. 実験

4.1. データセット

本実験では、球審視点により撮影された投球映像のデータセットを作成した。カメラ位置はホームベースから後ろに 300[cm]、高さ 125[cm]の位置で固定して撮影を行う。実験に用いた投球映像は様々なコースに投げ分けた 150 投球である。球速は少年野球の平均球速である約 70[km/h]を想定した。正解判定は 5 台のレーザー墨出し器及び 2 台のカメラによりボール位置を求め、ストライクか否かを判定する。

4.2. 実験結果

従来の二次元でストライクゾーンを再現する手法（従来手法 1）と、複数台カメラを用いた手法（従来手法 2）と、提案手法によるボールストライク判

定結果を分類評価指標により、評価する。ボールストライク判定を二値分類とみなし、4 種類の分類評価指標で評価した結果を表 2 に示す。表 2 より、提案手法の評価結果が、複数台カメラを使用する従来手法に近い値を示した。

表 2 ボールストライク判定評価結果

評価指標	正解率	適合率	再現率	F 値
従来手法 1	84.7	68.6	98.8	80.9
従来手法 2	99.3	98.6	100	99.3
提案手法	93.3	87.1	98.8	92.6

4.3. 提案手法による判定三次元シミュレーション

提案手法により再現したボール軌道とストライクゾーンのシミュレーションの一例を図 2 に示す。青色の球が投球されたボールを示し、黒線がボール軌道を示す。また、緑の五角柱はストライクゾーンである。図 2 より、ボール軌道とストライクゾーンを三次元で再現し、判定できていることが確認できる。

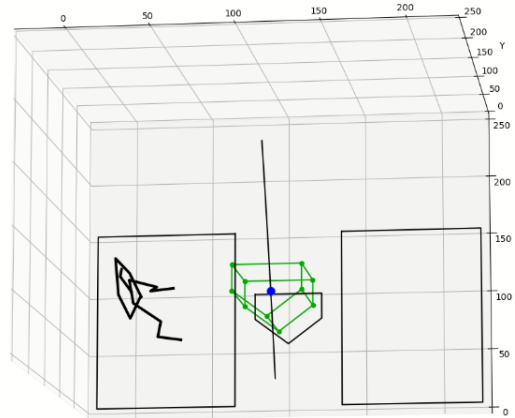


図 2 提案手法による三次元シミュレーション

5. むすび

本研究では、単眼カメラ映像を用いた形状既知物体を基準とする三次元空間推定と動体軌道推定手法を提案し、野球のボールストライク判定を対象にその有効性を検証した。提案手法は簡易なシステムでありながら、高精度な判定を実現できることを示した。また、従来手法と比較して高い判定精度を達成し、野球だけでなく他の競技や設備が限られた環境においても適用可能な手法であるといえる。

参考文献

- [1] Matthew Whitrock, “Developing MLB’s Automated Ball/Strike System (ABS),” Published in MLB Technology Blog, 2021. (最終閲覧日: 2024 年 1 月 7 日), <https://technology.mlbblogs.com/developing-mlbs-automated-ball-strike-system-abs-d4f499deff31>
- [2] 公共財団法人日本サッカー協会, “VAR,” (最終閲覧日: 2024 年 1 月 8 日), <https://www.jfa.jp/rule/var.html>
- [3] Zhengyou Zhang, “A flexible New Technique for Camera Calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 22, no.11, pp.1330-1334, Nov. 2000.

2024 年度

早稲田大学大学院基幹理工学研究科情報理工・情報通信専攻 修士論文

形状既知物体を利用した単眼映像からの三次元空間推定
及び動体軌道推定手法の研究

Research on 3D Space Estimation and Motion Trajectory
Estimation from Monocular Images Using Geometric Objects

福田 大翔

(5123F103-4)

提出日：2025.1.27

指導教員：渡辺 裕 教授

研究指導名：オーディオビジュアル情報処理研究

目次

第1章 序論.....	1
1.1 研究背景	1
1.2 研究目的	1
1.3 本論文の構成	2
第2章 関連研究	3
2.1 まえがき	3
2.2 BlazePose	3
2.3 YOLOv8.....	4
2.4 透視投影変換	4
2.5 野球中継映像を用いる手法	5
2.6 複数台カメラを用いる手法	5
2.7 むすび.....	5
第3章 提案手法	6
3.1 まえがき	6
3.2 提案手法の概要	6
3.3 カメラキャリブレーション	7
3.4 ストライクゾーン推定	7
3.4.1 ストライクゾーン算出	7
3.4.2 ホモグラフィ変換を用いた再スケール化パラメータ算出	8
3.5 ボール軌道推定	10
3.5.1 検出器作成	10
3.5.2 三次元ボール軌道算出	11
3.6 むすび.....	11
第4章 実験結果と考察	12
4.1 まえがき	12
4.2 データセット	12
4.3 評価指標	14
4.4 ボール軌道とストライクゾーンの合成.....	15
4.5 ボールストライク判定	17
4.6 ストライクゾーンの予測精度	17
4.7 変化球の軌道再現	17
4.8 考察.....	20
4.9 むすび.....	20

第 5 章 結論と今後の展望	21
5.1 結論.....	21
5.2 今後の展望	21
謝辞.....	22
参考文献.....	23
表一覧.....	25
図一覧.....	26
研究業績.....	27

第 1 章 序論

1.1 研究背景

近年、映像解析技術はスポーツ競技や監視などで広く応用されており[1]，特に動体軌道推定や三次元空間における物体認識の分野で注目されている．その中でも，スポーツ競技は，フィールドや使用する道具の大きさや形状が厳密に規定されており，競技ルールに基づいて試合や判定が行われる．そのため，正確な動体情報の取得が競技の公正性を担保する重要な要素となる．例えば，野球ではボールがストライクゾーンを通過したかどうかを判定する「ボールストライク判定」，サッカーではゴールラインをボールが完全に越えたかを判定する「ゴール判定」が存在する．これらの場面では，正確な判定を行うために三次元空間での動体軌道推定が不可欠となる．

現在，プロ競技の現場では，正確な判定を補助するために専用の計測装置や複数台のカメラを用いた高度なシステムが導入されている．その代表例として，野球の Automated Ball and Strike System (ABS)[2]やサッカーの Video Assistant Referee (VAR)[3]が挙げられる．これらのシステムは，高精度な判定を可能にする一方で，導入コストが非常に高く，専用機材や設置環境に依存するため，特定の会場や大会でのみ使用されている．ABS や VAR のような高度なシステムがプロ向けに普及する一方で，依然としてコストや環境面での制約が多く，アマチュア競技への導入は進んでいないのが現状である．

また，アマチュア競技においては，審判員不足という深刻な問題がある．審判資格や競技経験を持たない人物やボランティアが審判を務めるケースが多く，このような状況では，誤審の発生リスクが高まり，競技の公平性を損なう恐れがある．この問題を解決するためには，安価かつ簡易に導入できる判定補助技術の開発が求められている．

1.2 研究目的

そこで本研究では，単眼カメラ映像のみを使用するという制約のもと，三次元空間推定と動体軌道推定を行う手法を提案する．スポーツ競技における判定補助技術に適用することを目的とし，そのルールや特性を活用することで，推定精度の向上を目指す．本研究では，野球のボールストライク判定を事例に，単眼カメラから得られる映像を用いて，ストライクゾーンの推定とボールの三次元軌道の推定を行う．また，提案手法のボールストライク判定精度を評価し，その結果を既存手法と比較することで，提案手法の適用可能性と利便性を示す．

1.3 本論文の構成

本論文の構成を以下に示す.

第 1 章は本章であり, 本論文の研究背景, 研究目的について述べる.

第 2 章では本研究で用いる関連研究について述べる.

第 3 章では本研究で提案する三次元空間推定および動体軌道推定手法について述べる.

第 4 章では実験の結果及び考察について述べる

第 5 章では本論文の結論及び今後の展望について述べる.

第2章 関連研究

2.1 まえがき

本章では，姿勢推定モデルである **BlazePose**，物体検出モデルである **YOLOv8**，三次元座標位置決定のための透視投影変換，従来のボール軌道推定手法について説明する．

2.2 **BlazePose**

BlazePose[4]は姿勢推定モデルの一種である．**BlazePose** を利用することで動画像内の人物の骨格情報を推定することができる．検出される特徴点の数は 33 点であり，それぞれの特徴点の三次元座標を取得可能である．本研究で利用する **BlazePose** により検出した三次元骨格情報の出力画像を図 2.1 に示す．

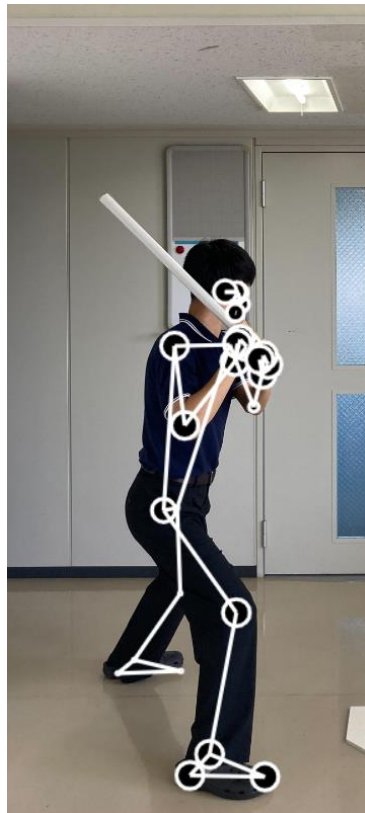


図 2.1 **BlazePose** 出力画像例

BlazePose では，図 2.1 のような画像に加えて，検出された特徴点の **XYZ** 座標と信頼度が出力される．なお，**XYZ** 座標の原点は腰の中心であり，信頼度は 0 から 1 の範囲の値をと

る．信頼度は1に近いほど推定が正確であり，検出に失敗した特徴点の XYZ 座標と信頼度は0となる．また，二次元骨格情報の取得やリアルタイムでの処理も可能である．

2.3 YOLOv8

YOLO (You Only Look Once) [5]は画像や動画内の物体をリアルタイムで検出するための，高速かつ効率的なディープラーニングベースの物体検出アルゴリズムである．本研究で用いる YOLOv8[6]は，以前の YOLO モデルよりも処理速度が非常に速く，リアルタイムでの物体検出が可能である．検出された物体は，フレームごとに X 座標の最大値(x_{max})・最小値(x_{min})，Y 座標の最大値(y_{max})・最小値(y_{min})，信頼度が出力される．なお，XY 座標の原点は画面左上の角であり，信頼度は0から1の範囲の値をとる．

2.4 透視投影変換

透視投影変換は実世界に存在する物体を二次元平面に投影する方法である[7]． (X, Y, Z) を実世界座標系の三次元座標， (u, v) を画像平面に投影された点の座標とする．カメラの内部パラメータ行列 A は，画像中心座標を (c_x, c_y) ，カメラのピクセル単位で表される焦点距離を f_x, f_y とすると式(1)で表される．

$$A = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

カメラの外部パラメータ行列 $[R|t]$ は，カメラの回転行列と並進ベクトルを用いて式(2)で表される．

$$[R|t] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \quad (2)$$

式(1)，式(2)より，実世界での物体の三次元座標と画像平面での二次元座標の関係は正規化パラメータ s を用いて，式(3)のようになる．

$$s \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3)$$

2.5 野球中継映像を用いる手法

野球中継映像を利用して投球されたボールの軌道を推定する手法が存在する[8][9][10]. この手法では、投手のリリース位置と捕手の捕球位置を基準に、フレーム内で取得したボールの位置情報を用いて三次元空間上でのボール軌道を推定する.

しかし、この手法にはいくつかの課題がある. 特に、野球中継映像を基に軌道を推定している点が挙げられる. 通常、野球中継映像を撮影するカメラはバックスクリーン横に設置され、約 100m 離れた位置から望遠レンズで撮影することが求められる. このような条件下で、遠距離からボールの動きを正確に捉えられる高性能なカメラを一般的に個人が保有しているとは考えにくい.

そのため、この手法を一般利用に適用するには、コストや機材の点で大きなハードルがあるといえる.

2.6 複数台のカメラを用いる手法

複数台のカメラを使用して投球されたボールの軌道を推定する手法が存在する[11][12]. この手法では、通常 3 台のカメラを設置し、各カメラで取得したボールの位置情報を統合することで三次元的なボール軌道を推定する. カメラごとに異なる視点からの情報を活用することで、ボールの動きを高精度で捉えることが可能であり、特にプロフェッショナルな試合環境では有用性が高いとされる.

しかしながら、この手法にはいくつかの課題がある. まず、複数台のカメラを使用するため、導入コストが高額になる点が挙げられる. これにより、草野球や少年野球などのアマチュア競技や、設備が限られた施設では導入が困難となる可能性が高い. また、カメラの時間同期や位置関係のキャリブレーションも場合によって必要となり、システムの設置や運用が複雑化する点も大きな課題である. 具体的には、カメラを適切な位置に配置し、それぞれの視野を重複させるためには精密な計算と作業が求められるほか、設置場所の選定が限られることもある.

さらに、運用時にも、複数のカメラが同時に正常に動作することが前提となるため、故障や誤差が生じた場合には、データの正確性に影響を与える可能性がある. このような技術的および運用上の制約から、一般利用や低コスト環境への適用には高いハードルがあるといえる.

2.7 むすび

本章では、本研究に関連する姿勢推定技術である **BlazePose**, 物体検出技術である **YOLOv8**, 三次元座標位置決定のための透視投影変換、従来のボール軌道推定手法について述べた.

第3章 提案手法

3.1 まえがき

本章では単眼映像を利用し，形状既知の物体を基準とした三次元空間推定および動体軌道推定を行う手法を提案する．この手法は，規定されたフィールドや道具を活用することで，高精度な推定を実現する．具体例として野球のボールストライク判定を対象とし，その有効性を検証する．提案手法は，姿勢推定アルゴリズムを用いたストライクゾーンの推定と，物体検出モデルを用いたボール検出を組み合わせることで，ストライクゾーンとボール軌道の三次元推定を行う．この際，野球における形状既知情報を活用することで，単眼映像のみから高精度な推定を実現する．

3.2 提案手法の概要

本研究では野球における形状既知情報を活用し，1台のカメラ映像を基にボールストライク判定を行う手法を提案し，有効性を確認する．提案手法の概要を図3.1に示す．提案手法はカメラキャリブレーションブロック，ストライクゾーン推定ブロック，ボール軌道推定ブロックの3ブロックで構成される．各ブロックの詳細は次節以降に述べる．また，本研究ではカメラのレンズ位置を原点として座標系を設定する．本研究の座標設定を図3.2に示す．

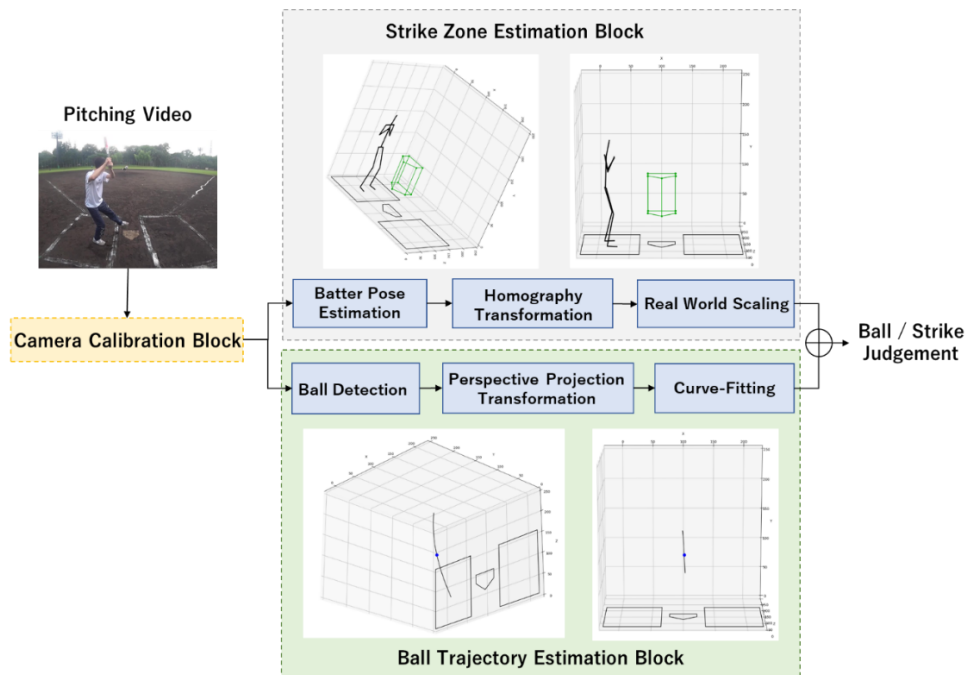


図 3.1 提案手法概要

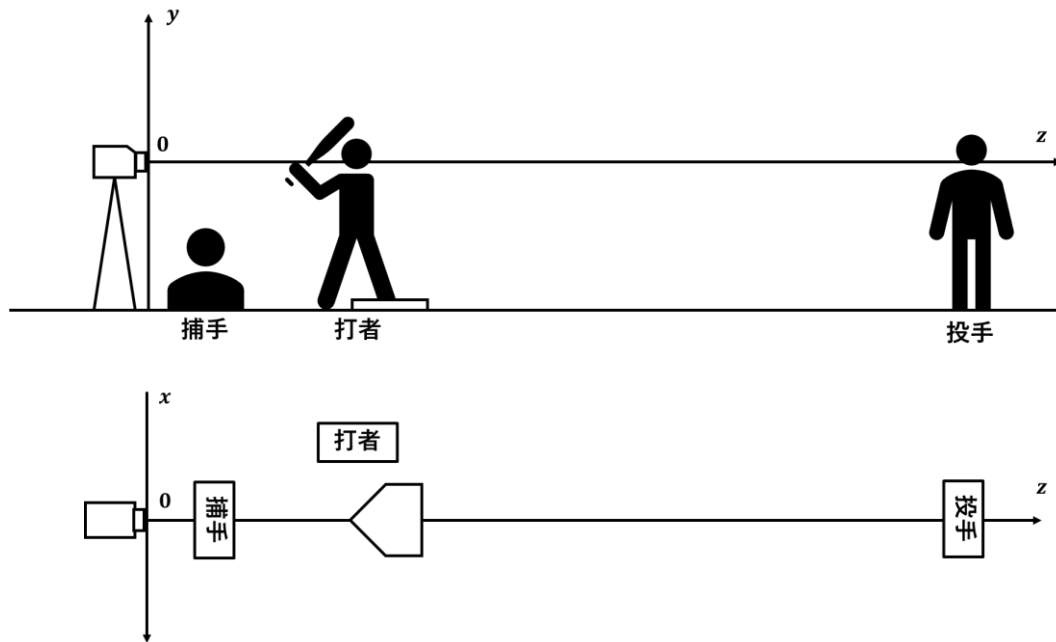


図 3.2 座標系設定

3.3 カメラキャリブレーション

単眼映像から三次元情報を正確に取得するためには、カメラのパラメータを算出する必要がある。本研究では、形状既知情報として野球場に設置されたホームベースを利用し、簡易なカメラキャリブレーションを行う。ホームベースは位置と大きさが既知であり、これを基に画面内の座標と実世界の座標を対応付けることで、透視投影の逆変換によりカメラの内部パラメータが近似できる。また、カメラの外部パラメータはカメラ位置を固定とすることで求まる。

3.4 ストライクゾーン推定

3.4.1 三次元ストライクゾーン算出

ストライクゾーンは打者の骨格に基づいて決定され、打者ごとに異なる[13]。そのため、打者の打撃姿勢に応じてストライクゾーンを推定する必要がある。本研究では、打者依存のストライクゾーンを推定するために、姿勢推定アルゴリズムである **BlazePose** を用いて打者の骨格を検出する。

検出した骨格の座標値を基に、ストライクゾーンの上端と下端を算出する。上端は両肩と両尻の Y 座標の平均値、下端は両膝の Y 座標の平均値として定義する。また、**BlazePose** により検出された三次元座標を活用し、ストライクゾーンの上端と下端をそれぞれ式(4)および式(5)で表す。

$$upper = \frac{y_{left_{shoulder}} + y_{right_{shoulder}} + y_{left_{hip}} + y_{right_{hip}}}{4} \quad (4)$$

$$bottom = \frac{y_{left_{knee}} + y_{right_{knee}}}{2} \quad (5)$$

3.4.2 ホモグラフィ変換を用いた再スケール化パラメータ算出

BlazePose のような三次元姿勢推定は相対スケールであるため、実際のスケールとは異なる場合がある．この課題を解決するため、本研究ではホームベースの大きさを基準とし、打者の足幅を実世界スケールで推定することで、ホモグラフィ変換を用いて再スケール化を行う．

まず、図 3.3 に示したように、画面上におけるホームベースの 5 つの角座標を基に、実世界と同じ比率を再現するホモグラフィ行列を算出する．このホモグラフィ行列を用いて、打者の両足の二次元位置座標を変換する．変換の際に使用する打者とホームベースの座標値を図 3.4 に示す．変換後の二次元両足間距離は式(6)で、三次元両足間距離は式(7)で表される．また、変換後の二次元両足間距離と実世界でのホームベースの大きさ情報 *LenBase* を用いて、三次元両足間距離を再スケール化するパラメータは式(8)で近似される．

最後に、算出した再スケール化パラメータを図 3.5 に示す．式(4)および式(5)で得られたストライクゾーンの上端および下端に乗算することで、より精度の高いストライクゾーンを再構築することが可能となる．

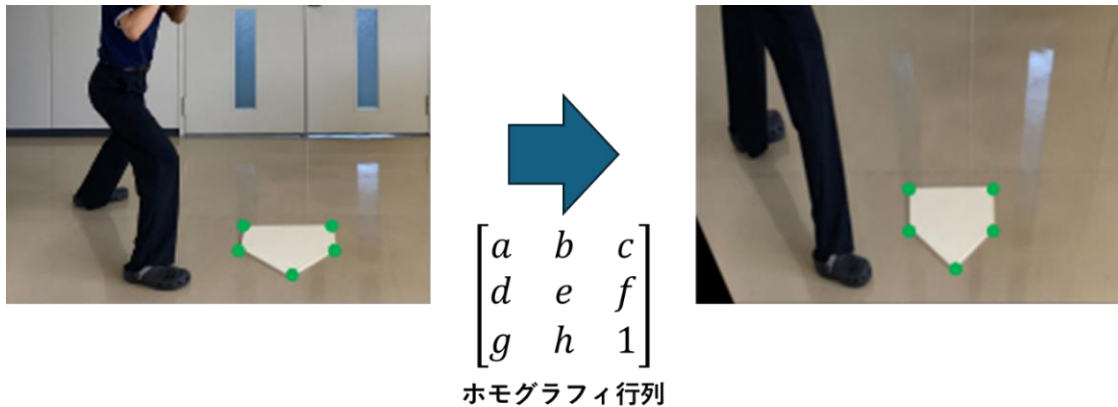


図 3.3 ホームベースの大きさ情報を用いたホモグラフィ変換

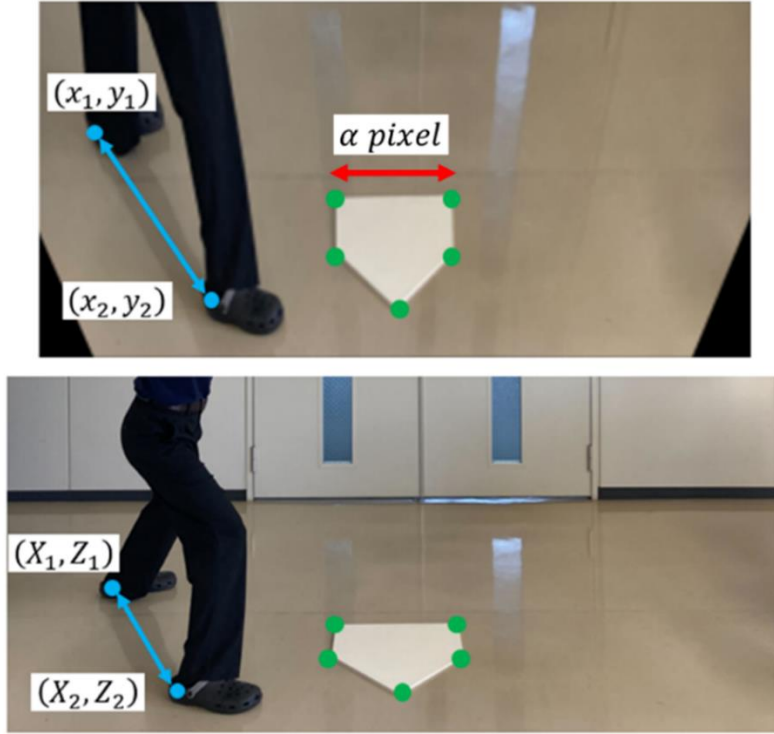


図 3.4 ホームベースの大きさ情報を用いた足幅長算出
(2024 年 情報処理学会第 86 回全国大会にて発表)

$$Distance_{2D} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (6)$$

$$Distance_{3D} = \sqrt{(X_1 - X_2)^2 + (Z_1 - Z_2)^2} \quad (7)$$

$$Q \approx \frac{LenBase \cdot Distance_{2D}}{\alpha Distance_{3D}} \quad (8)$$

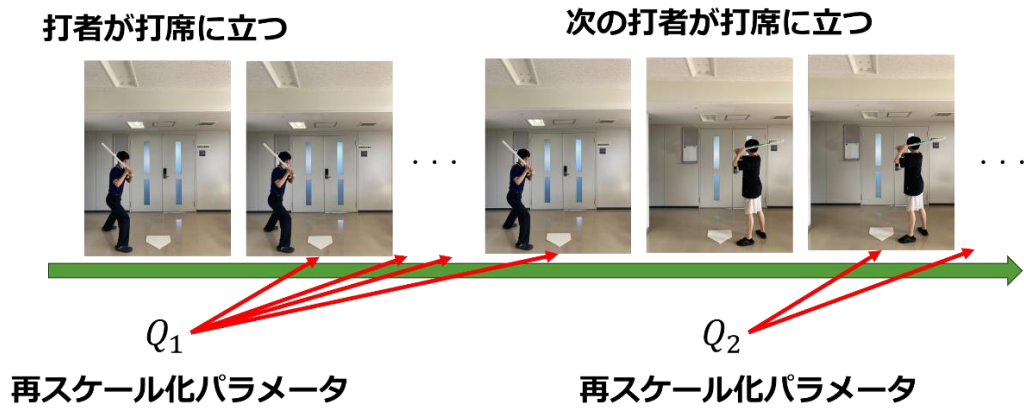


図 3.5 再スケール化パラメータの使用例

3.5 ボール軌道推定

3.5.1 検出器作成

投球されたボールは、画面上で非常に小さいため、汎用的な YOLOv8 モデルでは検出が困難な場合がある。そこで、本研究では野球ボール専用の検出器を作成する。具体的には、様々な大きさの野球ボールを撮影した画像をデータセットとして収集し、それを用いて YOLOv8 モデルをトレーニングする。ボールのサイズごとのデータセットの数を表 3.1 に示す。データセットのラベル付けには、自動ラベル付けが可能な Autodistill[14]を用いる。Autodistill は大規模な教師ありデータセットを用いずに、高性能な機械学習モデルを自動的に構築・精緻化するための手法である。学習は 100 エポックで実施し、検出器作成の全体的なプロセスを図 3.6 に示す。

表 3.1 学習データセットの内訳

ボールサイズ	データセット数
大	100
中	100
小	100

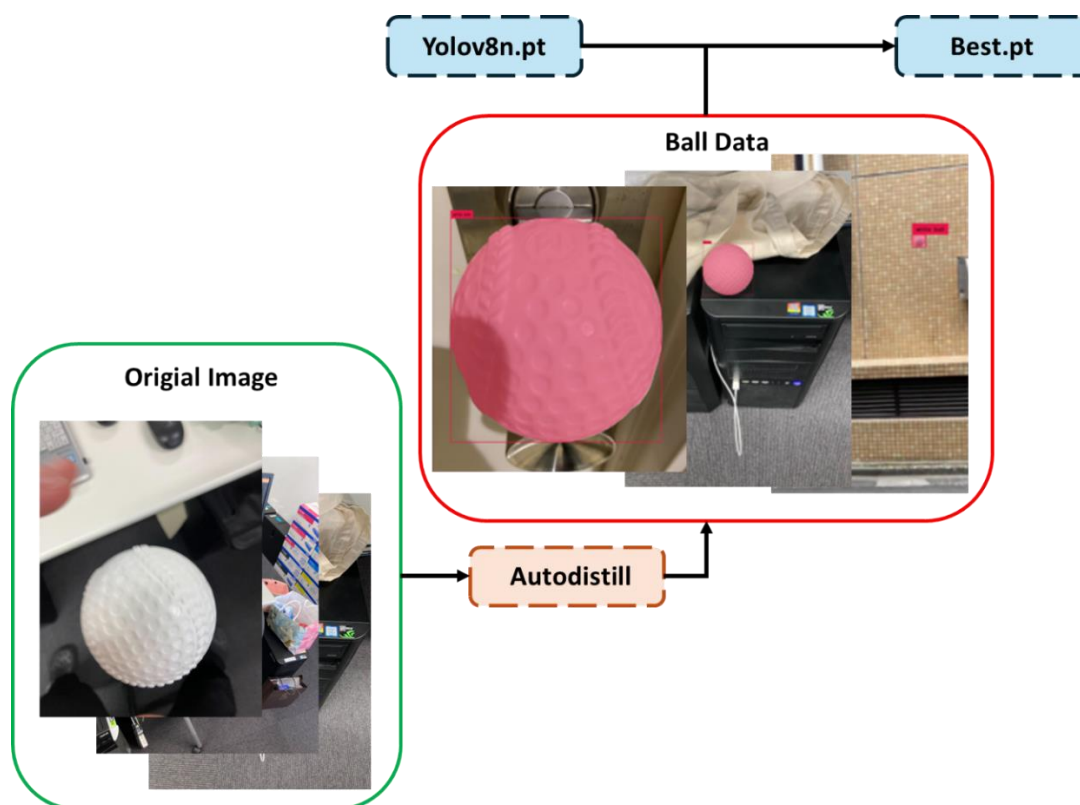


図 3.6 検出器作成の概要

3.5.2 三次元ボール軌道算出

取得した投球映像に対し，野球ボール専用に設計した検出器を用いてボールを検出する．検出されたボールの画面上の大きさと実際のボールサイズとの対応関係を利用し，ボールの Z 座標を推定する．この Z 座標を基に，画面上で検出されたボールの xy 座標を実世界座標に変換する．この処理を各フレームで繰り返すことで，ボールの三次元軌道を推定する．

実際のボールサイズを RBS ，画面内でのボールサイズを SBS とすると， Z 座標は式(9)で近似できる．また，投球されたボールは Y 軸方向への変化が大きいため，モーションブラーがバウンディングボックスに与える影響を軽減するために，画面内でのボールサイズは x 軸方向のデータを用いる．ボールの XY 座標は式(10)，式(11)で近似できる．

さらに，フレームごとに得られた XYZ 座標を連結し，曲線近似を行うことで，三次元空間におけるボールの軌道を推定する．

$$Z \approx \frac{f_x \cdot RSB}{SBS} \quad (9)$$

$$X \approx \frac{Z \cdot (x - c_x)}{f_x} \quad (10)$$

$$Y \approx \frac{Z \cdot (y - c_y)}{f_y} \quad (11)$$

3.6 むすび

本章では，単眼映像を用いてストライクゾーン推定およびボール軌道推定を行う提案手法について述べた．提案手法は，形状既知情報を活用することで，1台の単眼カメラ映像映像のみから三次元空間推定および動体軌道推定が可能となる．

第4章 実験結果と考察

4.1 まえがき

本章では，提案手法の有効性を評価するために行った実験の概要，結果，およびその考察について述べる．実験では，単眼映像を用いて取得したストライクゾーン推定結果およびボール軌道推定結果を，複数台カメラを用いた従来手法の結果と比較し，提案手法のボールストライク判定精度を検証する．また 1 台のカメラ映像を基に，形状既知情報を活用した三次元空間推定及び動体軌道推定が可能であるかを三次元空間の推定結果により示す．

4.2 データセット

本研究では，まず球審視点で撮影された投球映像のデータセットを作成する．撮影に使用したカメラは，ホームベース後方 300 cm，地上高 125 cm の位置に固定して設置する．実験で用いた投球映像は，異なるコースに投げ分けた 150 球で構成されている．正解判定は，5 台のレーザー墨出し器と横視点カメラを用いて実施する．実験時の機器配置を図 4.1 に示す．また，正解判定の決定に使用した映像の例を図 4.2 に示す．なお，投球の球速は少年野球の平均球速である約 70 km/h を想定している．

実験における正解判定は，投手、打者、および捕手の 3 視点から得られる判定を基に決定した．実験に使用した投球の正解判定結果を表 4.1 に示す．

表 4.1 実験に利用した投球データ正解判定

判定	投球数
ストライク	70
ボール	80

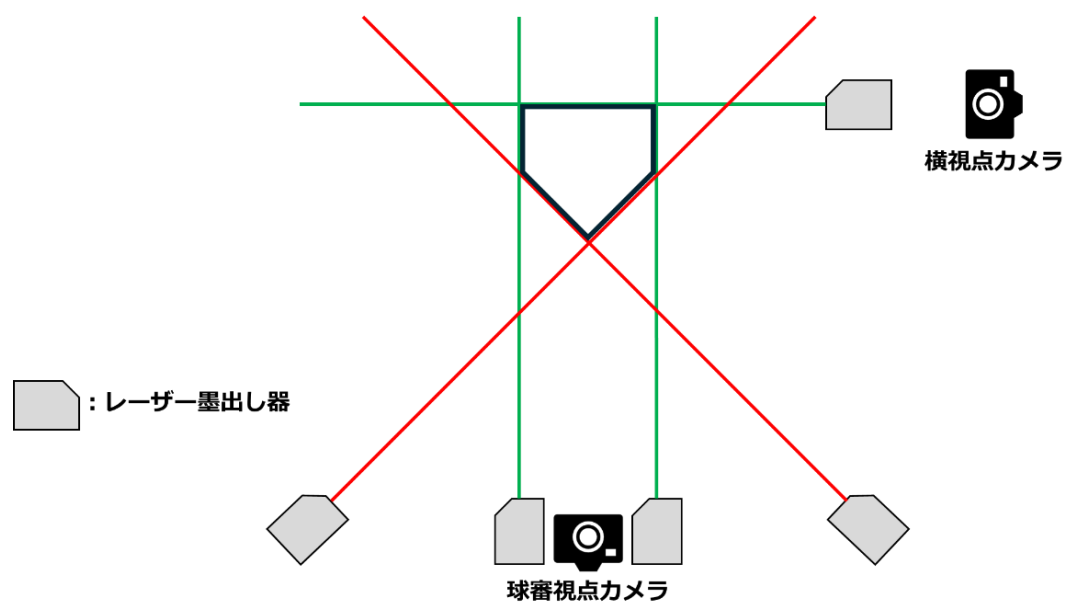


図 4.1 実験時の機器配置

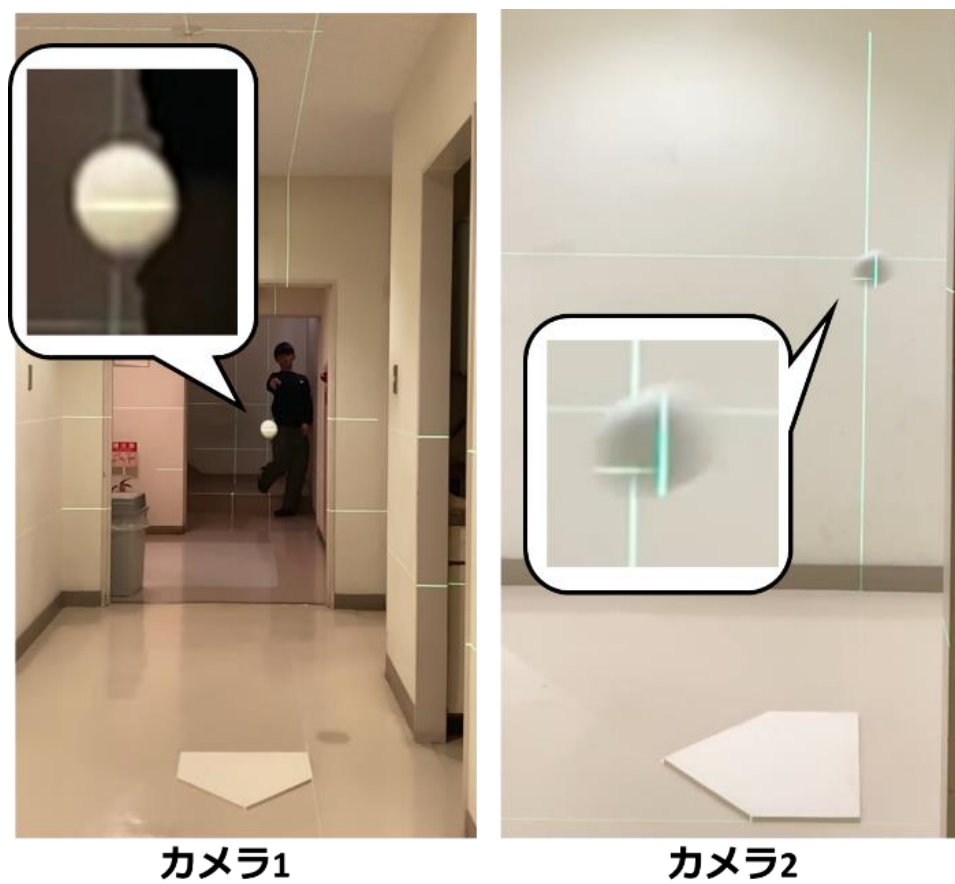


図 4.2 正解判定決定時使用映像例
(2024 年 スポーツ情報処理研究会にて発表)

4.3 評価指標

クラス分類の評価指標である，正解率（Accuracy），適合率（Precision），再現率（Recall），F 値（F-measure）について述べる．本実験では，ストライク判定を二値分類とみなし，4 つの分類評価指標で評価する．二値分類における予測値と真値の関係を表す混同行列を図 4.3 に示す．

正解	Positive	Negative
	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)
		予測値
		Positive Negative

図 4.3 二値分類の混同行列

正解率（Accuracy）は全データの中で正しく分類予測ができた割合を示し，式(12)で表される．

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

適合率（Precision）は Positive と予測されたデータのうち，実際に Positive であった割合を示し，式(13)で表される．

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

再現率（Recall）は正解が Positive であるデータのうち，正しく Positive と分類予測ができた割合を示し，式(14)で表せる．

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

F 値 (F-measure) は適合率と再現率の調和平均である．分類精度の評価に用いることがあり，式(15)で表される．

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (15)$$

4.4 ボール軌道とストライクゾーンの合成

取得したボール軌道とストライクゾーンの三次元座標情報を統合し，アニメーション表示することで，視覚的に判定結果を確認できるシステムを構築した．図 4.4(a)に捕手視点から見たボール軌道とストライクゾーンの様子を，図 4.4(b)に横方向から見た様子を，図 4.4(c)に俯瞰図をそれぞれ示す．

このシステムにより，打者ごとに異なるストライクゾーンを正確に定義し，取得したボール軌道に基づいてストライク・ボール判定をリアルタイムに行うことが可能となる．

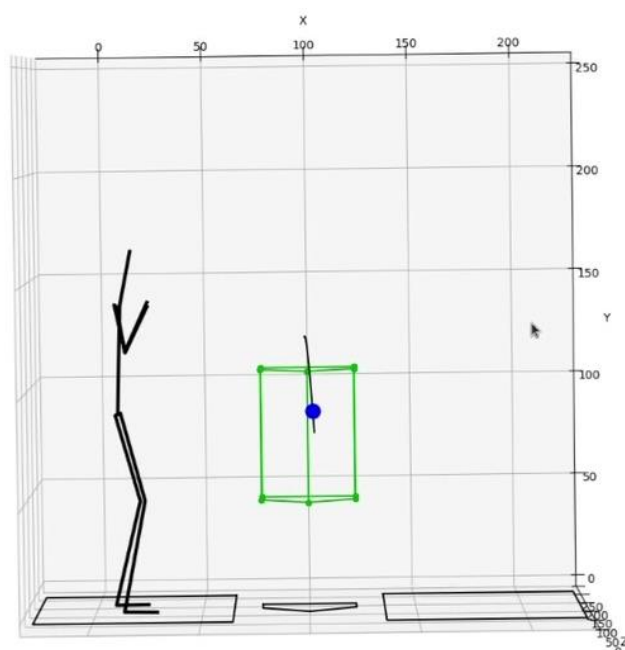


図 4.4 (a) 合成した三次元空間 (捕手視点)

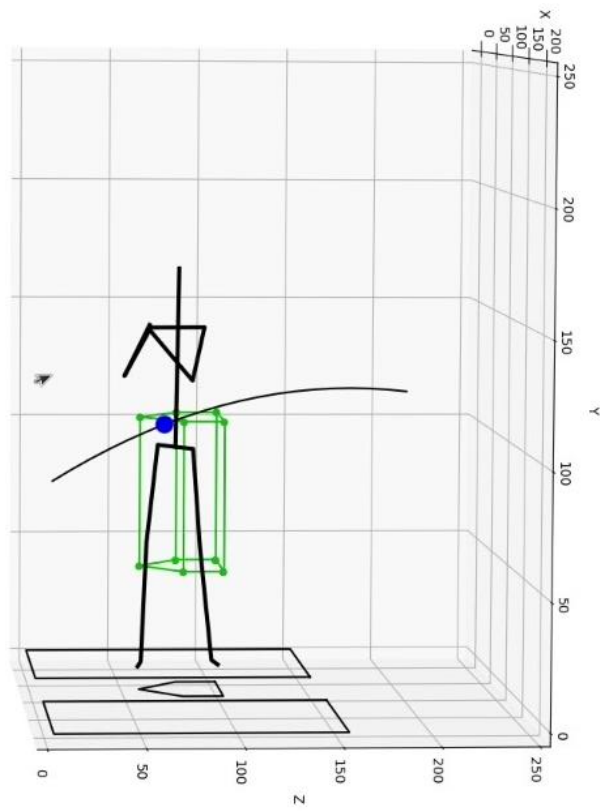


図 4.4 (b) 合成した三次元空間（横）

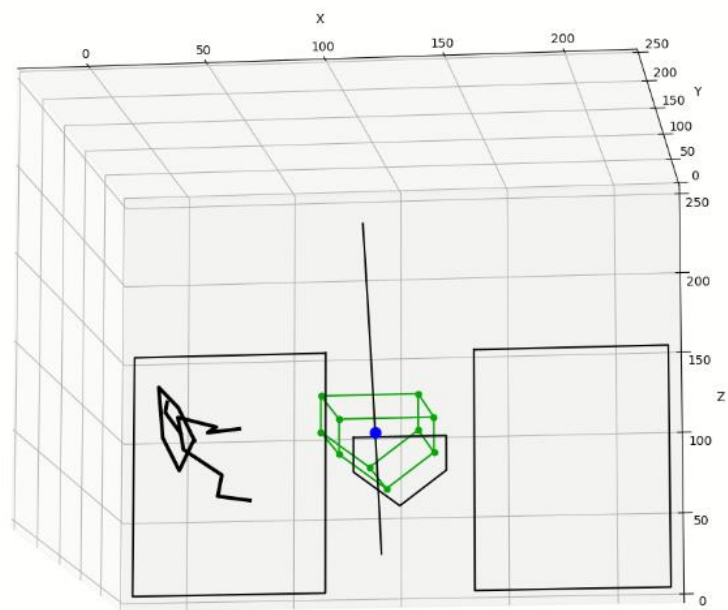


図 4.4 (c) 合成した三次元空間（俯瞰図）

4.5 ボールストライク判定

従来の二次元でストライクゾーンを再現しボールストライク判定した結果と、複数台カメラを用いてボールストライク判定した結果と、提案手法による三次元ボールストライク判定した結果を分類評価指標により、評価する。従来手法と提案手法によるボールストライク判定を分類評価指標で評価した結果を表 4.2 に示す。また、複数台カメラを用いてボールストライク判定をする従来手法 2 を到達目標としたときの、従来手法 1 と提案手法の到達率を図 4.5 に示す。

表 4.2 分類評価結果

評価指標	正解率 (Accuracy)	適合率 (Precision)	再現率 (Recall)	F 値 (F-measure)
従来手法 1	84.7	68.6	<u>98.8</u>	80.9
従来手法 2	99.3	98.6	100	99.3
提案手法	<u>93.3</u>	<u>87.1</u>	<u>98.8</u>	<u>92.6</u>

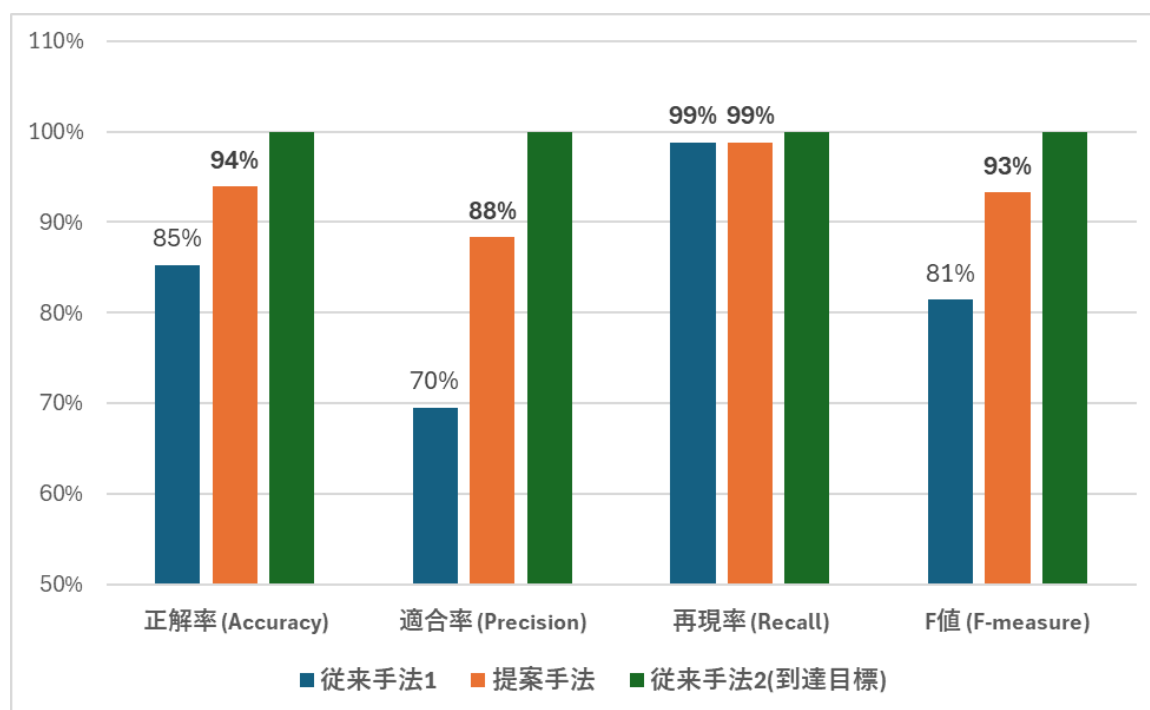


図 4.5 ボールストライク判定手法別到達率

4.6 ストライクゾーンの予測精度

BlazePose による三次元姿勢推定結果をそのまま使用し、ストライクゾーンを算出した場合と、ホームベースの大きさを基に、再スケール化パラメータを算出し、ストライクゾーンを実世界スケール化させた場合（提案手法）のストライクゾーン予測精度を比較した結果を表 4.3 に示す。実験に使用した 20 の打撃姿勢を図 4.6 に示す。

表 4.3 ストライクゾーン推定の誤差評価

	ストライクゾーン上端 [cm]	ストライクゾーン下端 [cm]
BlazePose	15.75	6.43
提案手法	7.35	3.24

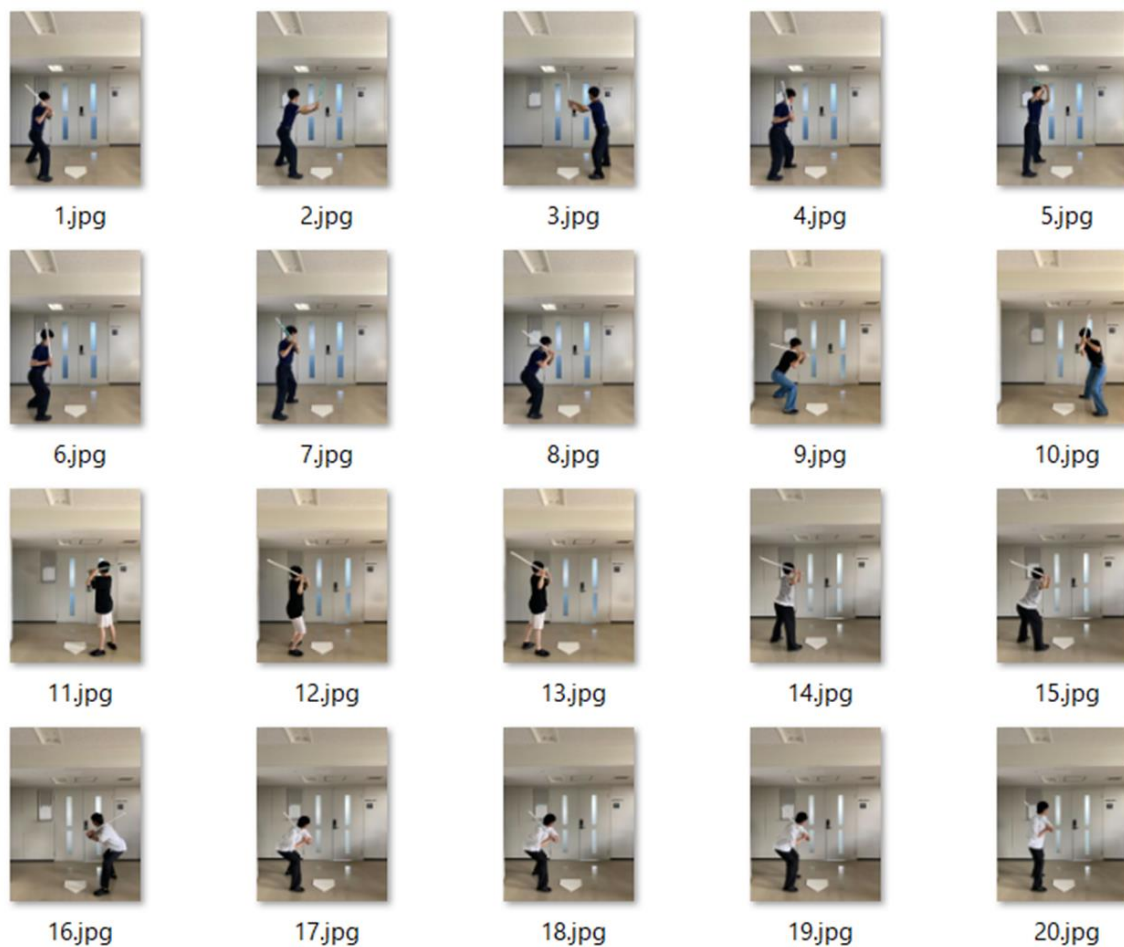


図 4.6 ストライクゾーン推定精度比較実験用データ

4.7 変化球の軌道再現

変化球を投じた際に，推定した三次元ボール軌道にボールの軌道変化が反映されているかを検証する．投球した変化球はスライダー，カーブ，スプリット，シュートである．提案手法により推定した変化球のボール軌道を図 4.7 に示す．

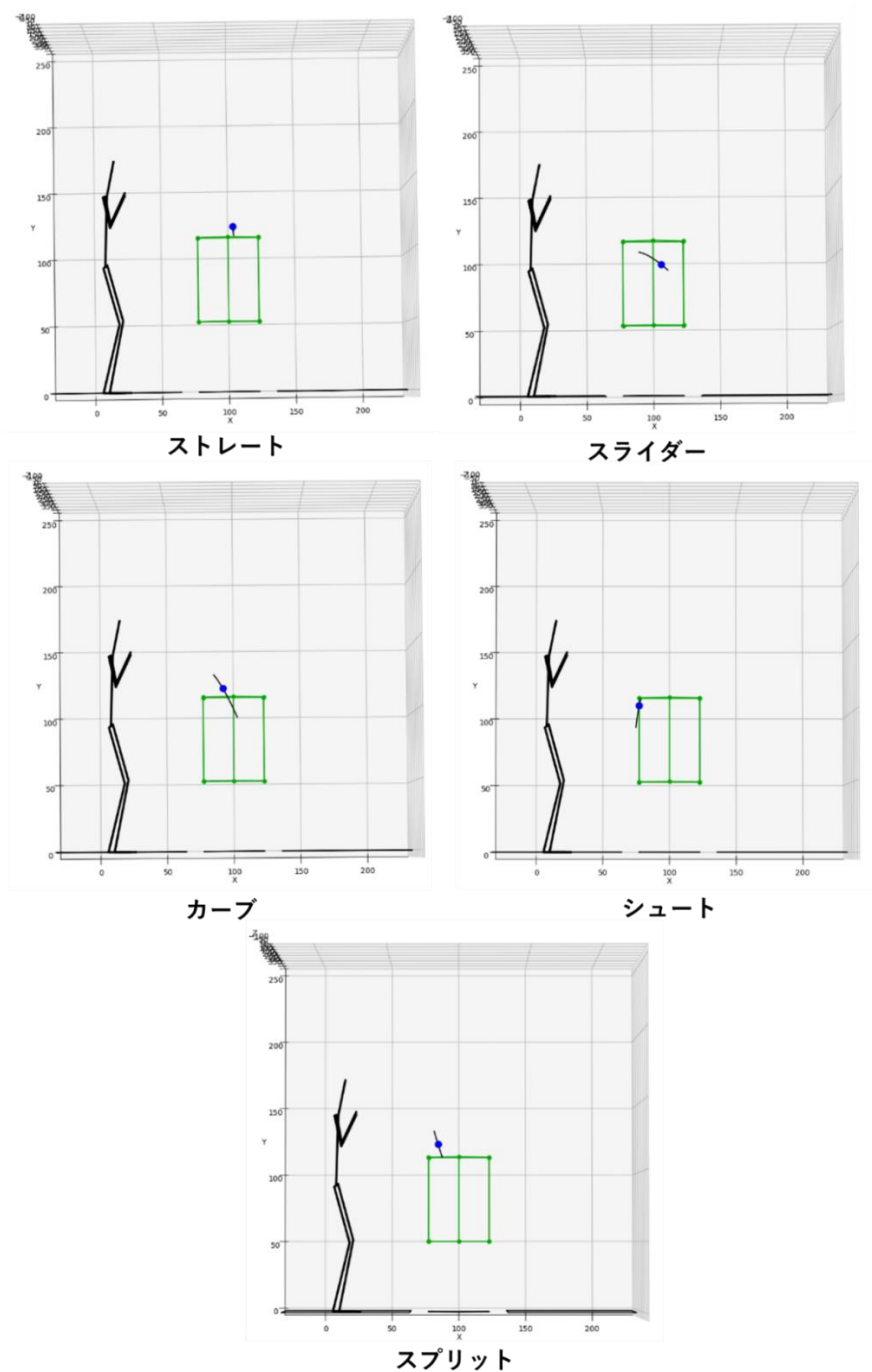


図 4.7 投球球種別の三次元ボール軌道

4.8 考察

図 4.4(a), 図 4.4(b), 図 4.4(c)に示す結果から, ストライクゾーンおよびボール軌道が三次元的に再現できていることが確認できる. したがって, 形状既知情報を活用することで, 1 台のカメラ映像から三次元空間推定と動体軌道推定が可能であるといえる.

表 4.2 より, ボールストライク判定において, 提案手法は 150 球の投球データに対して正解率 93.3%を達成した. これは, 複数台カメラを用いた従来手法 2 の正解率 99.3%に近い値であり, 単眼映像を用いて高い精度を維持できていることを示した. また, 従来手法 1 (正解率 84.7%) と比較すると約 8%の精度向上が見られた. 適合率 87.1%, 再現率 98.8%, F 値 92.6%という結果から, 提案手法は誤判定を抑えつつ見逃しの少ない判定を実現できており, 形状既知物体を活用した三次元空間推定と動体軌道推定が有効に機能していると考えられる.

表 4.3 の結果から, BlazePose によって取得した骨格情報の値から直接ストライクゾーンを推定した場合, 上端 15.75cm, 下端 6.43cm の誤差が生じた. しかし, ホームベースを基準に再スケール化を行うことで, それぞれ 7.35cm および 3.24cm に誤差を半減させることに成功した. これにより, 形状既知物体を利用することによってスケールの不整合問題を解消し, 高精度なストライクゾーン推定を実現できることを示した. また, ストライクゾーンの推定は打者の体格や姿勢に応じて動的に調整されるため, 提案手法は異なる打者に対しても高い汎用性を持つことが確認された.

図 4.7 は, 捕手視点から見たボール軌道を示している. ストレートやスライダー, カーブ, スプリット, シュートなど, さまざまな球種に対してその特徴的な軌道を正確に再現している. 特に, スライダーやカーブなど横方向の変化を伴う球種に対しても, その軌道を高精度に推定できていることを確認した. この結果は, 形状既知情報を活用した単眼映像からの動体軌道推定により, ボール軌道を高精度で再現できている可能性を示している.

4.9 むすび

本章では, ボールストライク判定結果及び三次元空間と動体軌道の推定結果について述べた. ボールストライク判定結果として, 1 台のカメラからボールストライク判定を行う提案手法が, 2 台のカメラを使用する従来手法に近い精度を示すことが確認できた. また, 形状既知情報を活用することで, 1 台のカメラ映像から三次元空間推定と動体軌道推定が可能であることを示した.

第5章 結論と今後の展望

5.1 結論

本研究では、形状既知物体を利用した単眼映像からの三次元空間推定および動体軌道推定手法を提案し、野球競技におけるストライクゾーン推定およびボールストライク判定を対象としてその有効性を検証した。提案手法の特徴は、形状既知物体であるホームベースやボールを活用して、単眼映像のみで正確な三次元空間と動体軌道を取得する点にある。これにより、従来の複数台カメラを用いた手法と比較して、低コストかつ簡易に導入可能な技術としての利便性を高めている。

実験では、BlazePose を用いて打者の骨格情報を取得し、ホームベースを基準にストライクゾーンを再スケール化することで、ストライクゾーン推定の誤差を半減することに成功した。また、150 球の投球データを用いたボールストライク判定においては、正解率 93.3%、適合率 87.1%、再現率 98.8%、F 値 92.6%という高い精度を達成した。さらに、変化球の軌道再現においても、ストレートやスライダー、カーブなど、複雑な球種の軌道を正確に再現することができ、単眼映像を用いた動体軌道推定手法として十分な実用性を持つことが示された。

これらの結果から、本研究で提案した形状既知物体を利用した単眼映像からの三次元空間推定および動体軌道推定手法は、低コストで導入可能な判定補助技術として有効であり、特にアマチュア競技や簡易な設備環境での利用に適した技術であることが示された。今後、本手法を他の競技や異なる環境に応用することで、さらなる汎用性の高い技術として発展させることが期待される。

5.2 今後の展望

本研究の提案手法は単眼映像を基にしているため、機器や設置条件が整っていない環境でも適用可能な汎用性が特徴である。今後は、他のスポーツ（例：バドミントンやバレーボール）や動体追跡が必要なタスクへの適用を進め、応用範囲を広げることが期待される。また、専用の計測器を用いた測定結果との比較による定量的評価や、リアルタイム処理の実現、動体検出精度のさらなる向上についても検討を進める。

謝辞

本研究を進めるにあたり，熱心にご指導いただいた渡辺裕教授に深く感謝申し上げます．

また，新たな視点を与えてくださり，困難に直面した際には助言をくださった研究室の皆様には厚く御礼申し上げます．

さらに，長年にわたり支えてくださり，励ましや経済的支援を惜しまず与えてくれた家族に，心より感謝いたします．

参考文献

- [1] 森 享宏, “スポーツ競技の映像技術との関わり”, 映像メディア学会誌 vol.69, no.4, pp. 309-312, Nov. 2015.
- [2] Matthew Whitrock, “Developing MLB’s Automated Ball/Strike System (ABS),” Published in MLB Technology Blog, 2021. (最終閲覧日 : 2024 年 1 月 7 日) ,
<https://technology.mlblogs.com/developing-mlbs-automated-ball-strike-system-abs-d4f499deff31>
- [3] 公共財団法人日本サッカー協会, “VAR,” (最終閲覧日 : 2024 年 1 月 8 日) ,
<https://www.jfa.jp/rule/var.html>
- [4] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang and M. Grundmann, “BlazePose: On-device Real-time Body Pose tracking,” CVPR Workshop on Computer Vision for Augmented and Virtual Reality, Seattle, Jun. 2020.
- [5] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” Computer Vision and Pattern Recognition (CVPR), pp.779-788, Jun. 2015.
- [6] Glenn Jocher, “ultralytics/yolov8,” (最終閲覧日 : 2024 年 10 月 10 日) ,
<https://github.com/ultralytics/ultralytics>
- [7] Zhengyou Zhang, “A flexible New Technique for Camera Calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 22, no.11, pp.1330-1334, Nov. 2000.
- [8] Y. Shimano, Y. Kuwano, M. Takahashi, and H. Kaneko, “Pitch Location Estimation from Broadcast Baseball Video”, NHK Science & Technology Research Laboratories, ITE Winter Annual Convention, 21B-4, Dec. 2023.
- [9] M. Takahashi, M. Fujii and N. Yagi, “Automatic Pitch Type Recognition from Baseball Broadcast Videos,” Tenth IEEE International Symposium on Multimedia, pp. 15-22, Dec. 2008.
- [10] W. Chu¹, C. Wang¹ and J. Wu, “Extraction of Baseball Trajectory and Physics-Based Validation for Single-View Baseball Video Sequences,” IEEE International Conference on Multimedia and Expo (ICME), pp. 1813-1816, Jul. 2006.

- [11] S. Miyata, H. Saito, K. Takahashi, D. Mikami, M. Isogawa and H. Kimata, “Ball 3D Trajectory Reconstruction without Preliminary Temporal and Geometrical Camera Calibration,” IEEE Conference on Computer Vision and Pattern Recognition Workshops, vol.205, no.40, pp. 164-169, Jul. 2017.
- [12] 齊藤, 井口, : “野球におけるストライク・ボールの自動判定システム”, 電子情報通信学会ソサイティ大会, pp.158, Mar. 2007.
- [13] 日本野球規則委員会, “野球規則集”, 2019, (最終閲覧日 : 2024 年 10 月 10 日) , <http://ru.ishibb.com/wp/wp-content/uploads/2019/03/3149527570a5c8c6b99ba438bb04c30f.pdf>
- [14] R. T. Hsu et al., “Autodistill: A framework for automated data distillation,” Autodistill Documentation, online, 2024. (最終閲覧日 : 2024 年 12 月 20 日), <https://autodistill.github.io/>

表一覧

表 3.1	学習データセットの内訳	10
表 4.1	実験に利用した投球データ正解判定	12
表 4.2	分類評価結果	17
表 4.3	ストライクゾーン推定の誤差評価	18

図一覧

図 2.1	BlazePose 出力画像例.....	3
図 3.1	提案手法概要図	6
図 3.2	座標系設定.....	7
図 3.3	ホームベースの大きさ情報を用いたホモグラフィ変換.....	8
図 3.4	ホームベースの大きさ情報を用いた足幅長算出.....	9
図 3.5	再スケール化パラメータの使用例	9
図 3.6	検出器作成の概要	10
図 4.1	実験時の機器配置	13
図 4.2	正解判定決定時使用映像例	13
図 4.3	二値分類の混同行列	14
図 4.4 (a)	合成した三次元空間（捕手）	15
図 4.4 (b)	合成した三次元空間（横）	16
図 4.4 (c)	合成した三次元空間（上空）	16
図 4.5	ボールストライク判定手法別到達率	17
図 4.6	ストライクゾーン推定精度比較実験用データ	18
図 4.7	投球球種別の三次元ボール軌道	19

研究業績

- [1] 福田大翔, 中島聖, 渡辺裕, “骨格絶対長を考慮した三次元姿勢推定手法の検討”, 2023 年 電子情報通信学会総合大会, D-12-67, pp.103, Mar. 2023.
- [2] 中島聖, 福田大翔, 渡辺裕, “骨格の左右対称性を考慮した 3 次元姿勢推定法の検討”, 情報処理学会第 120 回オーディオビジュアル複合情報処理研究会, Vol.2023-AVM-120, No.4, pp.1-2, Mar. 2023.
- [3] H. Fukuta, K. Nakayama, and H. Watanabe, “Three-Dimensional Baseball Strike Judgement by Monocular Video from Umpire’s Viewpoint,” IEEE Global Conference on Consumer Electronics, OS-AIP, pp.160-163, Oct. 2023.
- [4] K. Nakayama, H. Fukuta, and H. Watanabe, “Point Cloud Denoising and Outlier Detection with Local Geometric Structure by Dynamic Graph CNN,” IEEE Global Conference on Consumer Electronics, OS-AIP, pp.156-159, Oct. 2023.
- [5] 福田大翔, 杉山秀治, 渡辺裕, “ホモグラフィ変換に基づくホームベースの大きさ情報を活用したストライクゾーン取得精度の向上”, 情報処理学会全国大会, 7U-02, Mar. 2024.
- [6] 杉山秀治, 福田大翔, 渡辺裕, “アピアランスベースの視線推定における実現可能性を考慮した精度向上”, 情報処理学会全国大会, 5S-04, Mar. 2024.
- [7] 福田大翔, 足立翔平, 渡辺裕, “野球におけるボールストライク自動判定手法の提案”, スポーツ情報処理研究会, vol.48, no.14, pp.20-23, Mar. 2024.
- [8] 足立翔平, 福田大翔, 渡辺裕, “野球試合映像からの高速な全投球映像抽出手法の検討”, スポーツ情報処理研究会, vol.48, no.14, pp.16-19, Mar. 2024.
- [9] 福田大翔, 中山光典, 杉山秀治, 渡辺裕, “自動ボールストライク判定におけるフロートラッキングを用いた処理量削減法”, 映像情報メディア学会年次大会, 11A-5, Aug. 2024.
- [10] 杉山秀治, 福田大翔, 中山光典, 渡辺裕, “Class Activation Mapping を用いた視線推定手法の評価および有効性の確認”, 映像情報メディア学会年次大会, 11A-4, Aug. 2024.

- [11] 中山光典, 杉山秀治, 福田大翔, 渡辺裕, “主成分分析を用いたエッジの移動抑制による点群平滑化の改善手法”, 映像情報メディア学会年次大会, 13B-5, Aug. 2024.