

論文

機械学習による位置認識システムと連携した自動追尾システムの開発

林 勇 樹*
追手門学院大学

Design and Implementation of an Automatic Tracking System Utilizing Machine Learning for Position Recognition

Yuki HAYASHI
Otemon Gakuin University

Abstract

In recent years, the importance of video recording and motion analysis has been growing in various sports, including competitive swimming. However, the number of support staff able to attend international or major competitions is typically limited, forcing one individual to handle both camera operation and performance analysis in many cases. To address this constraint, we developed a system that integrates a machine learning-based athlete position recognition module with a motorized pan-tilt camera head. Our approach runs recognition and control processes on separate PCs, exchanging coordinate information in real time. By calibrating between the pool coordinate system and the camera coordinate system, and employing multiple predefined reference points, the system can operate reliably without program modifications—even when the camera is placed off-center relative to the pool. Field tests conducted at four different venues, both in Japan and overseas, confirmed that specifying a target lane alone is sufficient for long-duration tracking, requiring minimal manual intervention by analysis staff. Although future work on zoom control and multi-camera collaboration is anticipated, the results of this research demonstrate a fundamental technology that can reduce the labor needed for video recording and improve the efficiency of performance analysis.

キーワード：水泳競技, 自動追尾撮影, 位置認識, 電動パン・チルト制御, スポーツ映像解析, カメラキャリブレーション

Keywords : Swimming, Automated tracking, Position recognition, Motorized pan-tilt control, Sports video analysis, Camera calibration

*責任著者：林 勇樹, 追手門学院大学, yu-hayashi@otemon.ac.jp

Corresponding Author : Hayashi Yuki, Otemon Gakuin University, yu-hayashi@otemon.ac.jp

研究背景・目的

近年、競技スポーツにおける映像解析の重要性はますます高まっている。試合中の選手の動きやフォームを映像として記録し、それをもとにパフォーマンスを分析することで、戦術立案や個別指導の質を向上させることが期待できるからである。しかし、国際大会や国内の主要大会など、実際の試合会場に帯同できる分析担当者（アナリスト）の人数は限られている。国際大会においてはADカードの制約から、分析スタッフを1～2名程度しか派遣できないケースが多い。こうした少人数体制で試合に臨む際、アナリストには選手のレース撮影だけでなく、撮影した映像をもとにした分析・次のレースへの戦略検討など、多岐にわたる業務が求められる。しかし、実際の試合は午前に予選、午後決勝といった形でスケジュールが立て込んでおり、長時間にわたり撮影作業に専念しなければならない。撮影に人手を取られている間は、分析やデータ整理といった次の作業に着手することが難しく、結果的に分析作業の開始が遅れてしまう。

このような背景から、撮影作業の自動化がスポーツの現場で求められている。パン・チルト機構を備えたカメラ制御システムと、機械学習によるリアルタイムな選手位置認識AIとを連携させることで、撮影によって時間的に拘束されなくなれば、アナリストは他の業務（データ分析やフィードバック作業など）により多くの時間を割くことが可能となる。すでに、林ら（2023）が選手の位置情報を検出するAIシステムの開発を進めているが、こうした認識結果を実際にカメラシステムへ入力し、リアルタイムにパン・チルト角度を制御して映像を自動収録するシステムの構築は報告されていない¹⁾。

本論文では、これまで開発・検討してきた選手位置認識AIからの出力を特定のフォーマットで受け取り、その座標情報を用いてカメラの向き（パン・チルト角度）を自動制御するシステムの設計と運用事例について報告する。本システムは「プライマリシステム（選手位置認識システム）」と「セカンダリシステム（カメラ制御システム）」という形で機能（モジュール）を分割するので、本研究ではプライマリシステムからの指示の送信と、それに基づくセカンダリシステムのコントロールを行う部分の開発を行う。実際の設計上どのような点に留意し、どんな課題が生じたのかを示すとともに、将来的にアナリストの少人数体制を補完するための有望な方法

としての可能性を検討する。

関連研究・事例

スポーツ競技における映像解析や自動追尾の研究は、近年バスケットボールやサッカーなどを中心に数多く報告されている。Okuma et al. (2013) は、自己学習 (Self-Learning) フレームワークを導入し、潜在SVM・変形可能部品モデルを用いて少量のラベル付きデータから自動的に追加ラベルを取得することで、スポーツ映像中の選手検出精度を20%以上向上させた²⁾。Arbués-Sangüesa et al. (2019) は、単一カメラ映像を対象に、深層学習による選手の姿勢（ポーズ）推定とVGG由来のセマンティック特徴を融合することで、バスケットボール選手の検出・追跡精度を向上させる手法を提案し、MOTAなどの指標に基づく評価結果を示している³⁾。このように、深層学習や機械学習を活用して、複雑な放送映像からプレイヤーを正確に抽出する技術の開発が行われている。また、Citraro et al. (2020) は、移動かつ未校正のカメラ映像から、フィールドラインおよび選手の位置情報を使ってカメラの外部パラメータ（姿勢や焦点距離）をリアルタイムで推定するアプローチを提案している⁴⁾。完全畳み込み型の深層学習アーキテクチャを用い、サッカー、バスケットボール、バレーボールなど複数のシーンに適用可能である点を示した。

一方、複数のセンサや視点を統合するアプローチとして、Guo et al. (2024) がLiDARと複数カメラを使ったプレイヤートラッキングおよび3次元ポーズ推定システムを構築し、その結果を元にVR/ARプラットフォーム上で試合をリアルタイムに可視化する手法を報告している⁵⁾。こうした高度なセンサ融合は、自動カメラ制御を含む新たな競技映像サービスの可能性を示す事例といえる。また、複数視点からの自動追尾という観点では、高橋ら（2014）がサッカーボールを機械学習で検出し、実空間上に投影することで高精度なボール追跡を行う方法を示している⁶⁾。今井ら（2018）や劉ら（2022）は、それぞれサッカーやバスケットボールにおけるプレー認識やファウル検出の自動化を試みており、選手およびボールの位置情報を多様な機械学習手法と組み合わせることで、イベント解析の自動化が可能であることを示した^{7,8)}。渡辺（2018）や清水ら（2019）は、スポーツ領域へのAI応用が進んでいる背景を概観し、トラッキング技術の高度化や自動映像収録に向けた研究開発動向が

今後ますます重要となると指摘している^{9,10)}。

こうした先行研究の多くは、認識（検出・追跡）部分と撮影制御部分を一体化したシステムとして扱う例が多い。しかし水泳のように、水面の反射環境が会場ごとに異なり、一時的に選手が潜水してオクルージョンが発生するケースが頻繁に起こる競技では、認識アルゴリズムの細かな調整が必要となる可能性が高い。本研究では、認識システムと制御システムを分離し、座標情報を最小限の形式でやりとりする設計を採ることで、認識部分を改良しても制御システム自体はほとんど変更せずに済む構成を目指している。また、Citraro et al. (2020) のようにカメラ姿勢を都度リアルタイム推定する手法もあるが、本研究ではプール座標とカメラ座標の対応付けを事前のキャリブレーション処理によって辞書化し、短時間の現地作業だけで容易に座標変換が適用できるワークフローを構築する⁴⁾。こうした方針は、競技会場ごとに照明条件や撮影位置が大きく変わる水泳競技の現場運用において、システムのセットアップ負荷を抑えられるという利点がある。

システム設計

本研究で構築したシステムは、大きく分けて「AI ベースの選手位置認識モジュール」と「カメラ制御モジュール」の2つに分割される。まず、位置認識のための AI は林ら（2023）の先行研究で開発された学習モデルを用い、映像中の選手がプール上のどの位置にいるかをリアルタイムで検出する¹⁾。この位置認識モジュールは独立した PC 上で稼働し、そこで推定された座標情報をローカルネットワーク経由でカメラ制御用 PC へ送信する。カメラ制御用 PC は、受け取った座標情報をもとにモーター付きの雲台（パン・チルト機構）に指令を出し、撮

影対象にカメラを向け続ける仕組みを担っている。こうした二層構成を採用する理由は、認識側と制御側を切り離すことで、それぞれを最適化しやすくする点にある。たとえば、位置認識モジュールは水面反射や多数のスイマーの混在といった環境要因に合わせて学習モデルを改良したり、推論パラメータを微調整したりといった実装変更を行うことが多い。一方、カメラ制御側はハードウェアの特性や動作速度を考慮しつつ、座標→パン・チルト角度への変換ロジックを安定化させる必要がある。両者を同一のプログラムにまとめてしまうと、どちらか一方に変更が加わるたびに全体を修正しなければならないケースが生じやすい。そこで、最小限の通信プロトコル（「いまどこを撮影したいか」という座標データと、「現在雲台が向いているパン・チルト角度」の情報）に絞って連携させることで、開発・運用上の柔軟性を確保している。

本システムの運用において、図1のように事前のキャリブレーションによって「プール（または撮影フィールド）の座標系」と「カメラのパン・チルト角度」を対応付けたキャリブレーションデータを作成することが必須となる。具体的には、プールの四隅と、各レーンの中央部、プール内の中心部分など、目印となる座標を特定（図1左）し、そのポイントが中心となる位置に合わせてカメラ制御用 PC 側でパン・チルトの値を校正（図1右）する。これにより、認識モジュールが“プール座標”を出力する仕様はそのままに、カメラ制御用 PC でその値を正しく変換して「どう動けば目的通り撮影できるか」を計算できるようになる。こうした座標系の対応付けを一度行ってしまえば、大会中通じて同じセットアップのまま運用することが可能となる。実際の試合では、午前の予選から午後の決勝に至るまで長時間

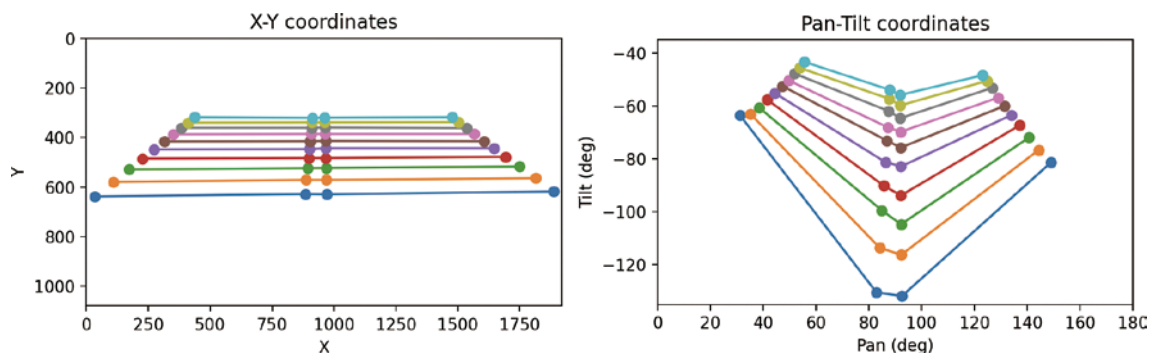


図1：カメラ座標系とパン・チルト角度の対応イメージ

にわたって撮影を行う場面が多い。本システムを使用すると、あらかじめ追尾したいレーンを指定しておくだけで、認識モジュールと制御モジュールの連携により、カメラが自動的にパン・チルトを調整し続ける。アナリストはカメラ操作に手をかけずに済むため、その分の時間を映像の分析や次レースの作戦立案に充てられる。以上のように「AIによる位置推定」「パン・チルト制御」「両者をつなぐ座標変換とネットワーク通信」を柱とすることで、シンプルかつ拡張性の高い自動撮影システムを構成している。

実装の詳細

本研究で開発したシステムは、図2のようにAIによる認識システムとカメラ・雲台側の制御プログラムとを別々のPCで動作させる二層構成を採用しており、両者の間をTCPソケット通信によって連携させている。

ハードウェアはGPUを搭載したPCを認識側に用い、YOLOベースのモデルを用いて、図3のように競泳選手の位置と移動軌跡を推定する。一方、カメラ制御側のPCはEdelkrone Head PLUS v2に対するAPIコマンドを発行し、パンおよびチルト角度をリアルタイムに操作する役割を担う。以下では、実際に使用したプログラムのうち、認識側プログラムと制御側プログラムを中心に、それぞれの実装上の特徴や通信の手順について順を追って述べる。

選手認識側（プライマリシステム）プログラム

本プログラムでは、筆者が先行研究で訓練・開発した重み学習済みモデルを用いて水泳競技の映像から選手を検出し、リアルタイムに追跡する。プログラム冒頭では、リスト1のようにセカンダリシステム側とのソケット接続を確立する。

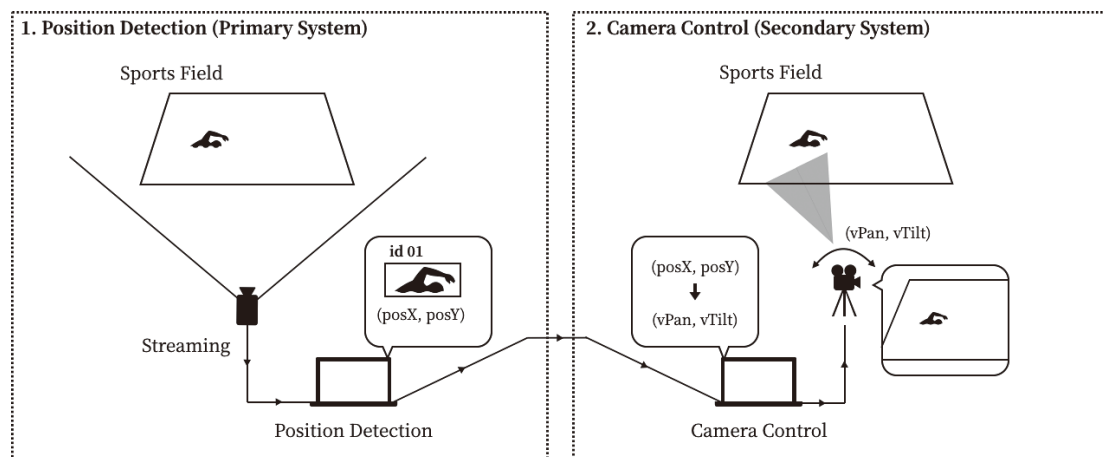


図2：プライマリ・セカンダリシステム連携の模式図



図3：認識システムによる泳者位置と軌跡同定の様子

リスト1：プライマリシステムとセカンダリシステムとの通信

```
def start_server():
    server_socket = socket.socket(socket.AF_INET,
    socket.SOCK_STREAM)
    server_socket.connect((client_ip, client_port))
    print(f"Connected to client {client_ip}:{client_port}")
    ... (中略) ...
    return server_socket, receive_thread
```

start_server() 関数内では、セカンダリシステム側の IP アドレスおよびポート番号を指定し、認識側 PC が制御側 PC に対して TCP クライアントとしての役割を果たす。さらに別スレッドとして制御 PC からの指令を受け取るようにする。制御 PC は撮影画角のリセットや追尾レーン番号の変更を任意のメッセージを介して伝えてくるため、対応するトリガーが来れば追尾状態を初期化するなどの操作を行う。

また、選手位置の推論と追跡は、リスト2のように行われる。

リスト2：カスタムトラッカーアルゴリズムを指定したYOLOの認識

```
model.track(frame, persist=True, classes=[1],
    tracker="custom_botsort.yaml")
```

本システムでは、YOLO モデルに BoT-SORT をベースにしたトラッキングアルゴリズムを組み込んでおり、映像フレームごとに各選手の位置と ID を持続的に割り当てる。事前に学習した検出のクラス ID のうち、[0] が泳者頭部、[1] が泳者全体に相当するため、運用においてはクラス 1 に限定した検出を行う。検出された位置情報をもとに、「どのレーンの泳者であるか」を判定する過程（図4）では、事前に読み込んだ座標変換行列を利用してレーン領域を逆透視変換し、映像上（図5）での多角形を求める。フレームごとに選手がどの領域に含まれるかを調べ、対象レーンに入っているオブジェクトを「現在の追尾 ID」として更新する。

また、オクルージョンの発生により、一定期間ターゲットを見失った場合、2通りの現象が想定される。1つ目はターゲットが一時的に隠れている場合であり、トラッキングアルゴリズムにより同じ ID で再度認識される可能性がある。もうひとつはトラッキングアルゴリズム上で同じ ID だと同定されなかった場合で、この場合は同じ ID で再度認識される可能性はない。そのような制約から脱するため、一定フレーム数連続で選手を検知できなかった場合に再検索へ移行するためのしきい値を導入した。最終的に認識された選手のカメラ座標系における (x,y) 座標は、JSON 形式にシリアル化し、リスト3のように制御 PC へ送信される。

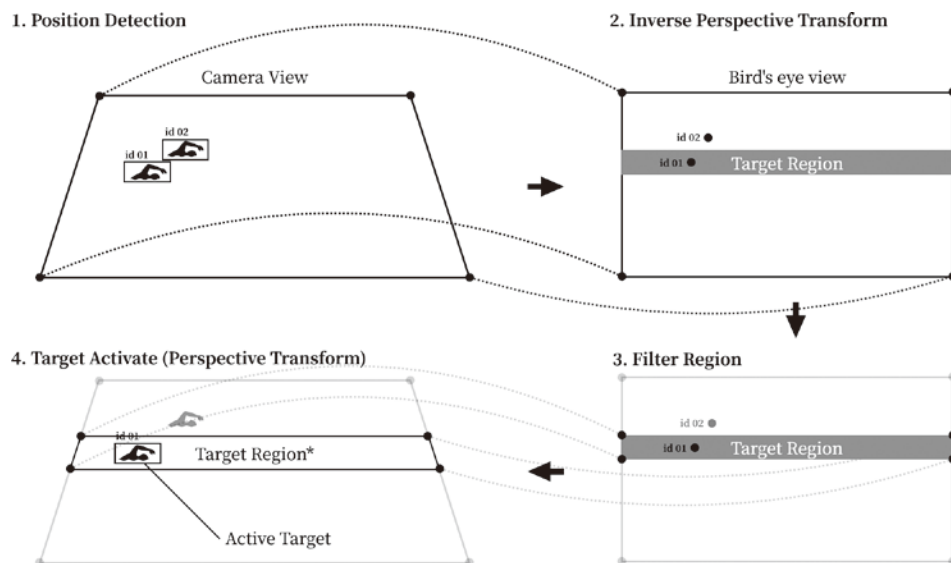


図4：座標変換行列によるターゲット選定の流れ



図5：実際の映像におけるターゲット領域の選定

リスト3：制御PCへの認識位置データの送信プログラム

```
pos_data = {
    'pos': (int(x+margin), int(y+margin)),
    'lane': currenttracklane
}
msg = json.dumps(pos_data)
server_socket.sendall(msg.encode())
```

カメラ制御側（セカンダリシステム）プログラム

カメラ制御側のスクリプトは、認識結果を受信しながら Edelkrone Head PLUS v2 のパン・チルト角を制御する役割を担う。本研究で用いた機器は Edelkrone 社が定義する API ドキュメントによってネットワーク経由でパン・チルト角度のコントロールが可能である。パン・チルト角度のコントロールは、REST API により回転方向と -1 ～ 1 の規格化角速度を与えることで実現する。目的のパン・チルト角度を設定するコントロールも可能であるが、時々刻々と変化する選手位置を追従するためには、常時機器に対して移動指示を速度パラメータで渡し続ける必要がある。

セカンダリシステム側が待受状態に入り、サーバ（認識側プログラム）から送られてくる座標情報を解析する。上述の JSON 形式の文字列を読み込んだうえで目的位置を格納するグローバル変数に座標が保管され、適宜更新される。Head PLUS v2 への制御コマンドは、本プログラムがローカルホストに対して REST API リクエストを送る形で実装する。指定したパン・チルト角速度（角度ではなく角速度である点に注意）を POST リクエストとして送信し、電動雲台のヘッド部分が動く構

造になっている。なお、現在のパン・チルト角度を知るために Head PLUS v2 からのレスポンスを待機する必要がある。この制御ループは `main()` 関数の中で一定周期（0.1 秒ごと）に実行されており、リスト4のように目標角と現在角との偏差に基づく移動角速度を算出している。

リスト4：目標角と現在の角度との偏差から次の速度を決定するコード

```
while True:
    status = get_device_status()
    current_pan = status['data']['readings']['headPan']
    current_tilt = status['data']['readings']['headTilt']
    aim_pan, aim_tilt = xy_to_pan_tilt(aim_x, aim_y)
    pan_speed = tendclamp(aim_pan - current_pan)
    tilt_speed = clamp((aim_tilt - current_tilt) / 5, -1, 1)
    move_to_position(-pan_speed, tilt_speed)
```

ここでリスト4中の `xy_to_pan_tilt()` という座標変換関数は、事前に定義された辞書型から「レーン番号とプール幅に応じた (x,y) 座標とパン・チルト角度の対応」を計算する関数である。

この座標変換において注意すべき点として、図6に示す回転角の差が挙げられる。フィールドを直進するターゲットを中央付近から撮影するケースでは、フィールドの左右端に選手がいる場合の距離と、フィールドの中央付近に選手がいる場合では距離が異なるため、選手の移動する画素数に対応するパン角度は非線形な値を設

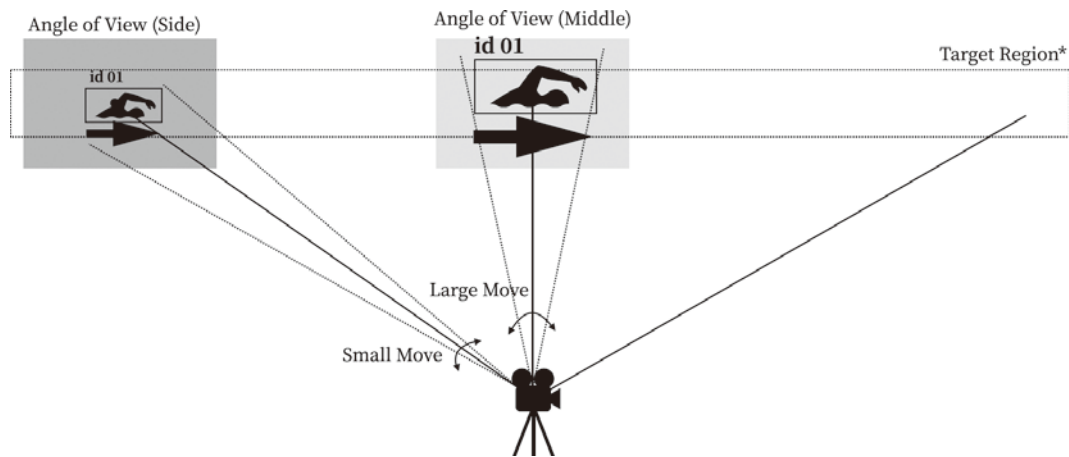


図6：フィールド左右端と中央部分での撮影対象選手までの距離差による回転角の違い

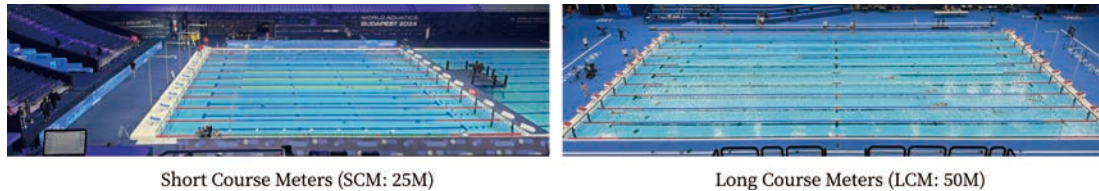


図7：会場毎に異なるフィールドサイズや撮影位置の中央からのずれ

定する必要がある。これにより、人間が撮影するような滑らかな移動が実現される。本プログラム内では、カメラの物理可動域を制限しており、実際に計算した目標角がこの範囲を逸脱しそうな場合はクリップ処理を行い、安全を確保している。さらに、現場の状況に応じて「別のレーンの選手の追尾に切り替えたい」「撮影開始地点を逆側の端点に変更したい」といった画角リセット操作も必要になるため、適宜キーボードからの入力を監視し、それに応じて目標座標を強制的に更新するようにした。

キャリブレーションおよび運用のポイント

本システムを運用するうえでは、まず撮影会場ごとに「プール外形・レーンごとに対応するパン・チルト角度」を定量化した辞書データ（キャリブレーション）を作成する作業が必要となる。コード中では **Left, Middle, Ghost-Middle, Right** の4つのキーポイントで各レーン番号に対応する値を連想配列として持たせている。この設計により、競技会場のレイアウトやカメラの設置位置が異なる場合（図7）でも、キャリブレーションファイルを差し替えるだけでシステムを再利用できるよう設計

した。

小括

認識プログラム（プライマリシステム）と制御プログラム（セカンダリシステム）を分割したうえで、TCPを介した座標データ送受信によって連携する本研究の実装は、比較的シンプルな構成を維持しつつ、柔軟性を高めるといふ狙いがある。認識モデルの更新や学習パラメータの変更が発生しても、制御側のパラメータを大きく修正する必要はなく、逆に雲台やカメラを別のメーカー製品に切り替える場合でも、通信インタフェースさえ保てば認識側の実装を流用できる。この汎用性は、本システムを運用環境に適合させやすくしている利点である。

実運用事例と課題

本研究で開発した雲台コントロールシステムは、国外2箇所・国内2箇所の計4つの異なる環境で試験的に運用され、それぞれで安定した撮影を行うことができた。いずれの場合も、撮影ポジションやプール長が同一ではなかったが、既述のフィールド辞書（キャリブレーション）

ションデータ）を作成するだけで、基本的なシステム構成を変えることなく対応できた。実際に運用してみた際に特に有効だったのは、カメラを設置してシステムを起動してしまえば、操作者が行う作業は「どのレーンを撮影ターゲットにするか」を指定するだけで済むことである。たとえば競泳の場合、レースでは1つのレーンに1人しか入らないため、「特定のレーンにいる選手」を認識して、そこにカメラを向け続けるだけで自動追尾撮影が成立する。ボール競技のように対象物が高速で移動するスポーツでは、カメラがどの程度スピードについていけるのか、あるいはプレイのどの部分を撮影対象とするのかといった別の設計課題が出てくるかもしれないが、本研究が対象とした競泳種目においては、そうした問題は起きず、運用時には撮影の自動化が非常に円滑に機能した。

また、実運用の中での注意点は、撮影位置の制約が会場ごとに異なるため、常にプールの「ほぼ中央」から撮影できるとは限らないという事実である。国際大会のように複数国のアナリストが一箇所で撮影・分析を行う環

境（図8）では、会場設計上の理由で観客席や壁に遮られ、カメラを正面中央に置けない場合も多い。このとき、水平方向（パン）の左右両端が会場によって偏ってしまう場合がある。本研究では、図9のようにカメラが正面を向いた状態を“仮の中心”として扱うために、フィールド左右端と中央を示す **Left, Middle, Right** というキャリブレーションポイントに加えて、**Ghost-Middle** という新たな基準点を追加する実装を行った。これは、実際のプール中央（理想的な長軸方向のセンター）とは異なる「カメラが実際に真つすぐ向いたときの位置」を別の参照点として扱うことで、左右のパン角やチルト角を合わせこむ制御をよりスムーズに行う手法である。例えば、本来のプール中央に合わせるだけだと、カメラ軌道が左右対称になってしまい目的の映像が撮影できなかったが、ゴーストミドルを設定することでカメラが見かけ上の正面を軸に動くようになり、キャリブレーションデータの修正のみで適切にフィールドをカバーした撮影ができた。

このように、あらかじめ定義した複数のキャリブレーション

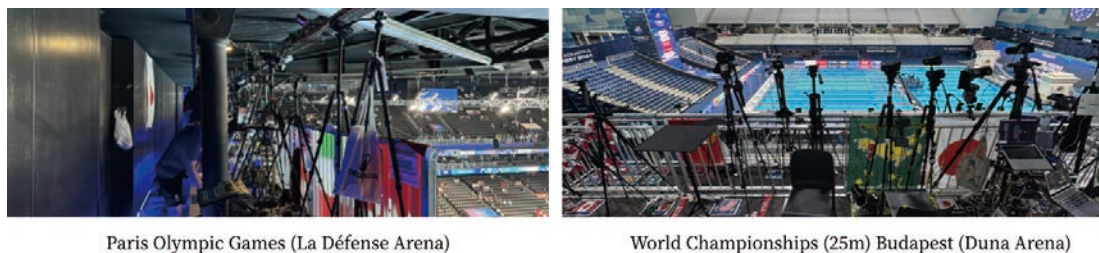


図8：会場ごとの撮影可能場所の配置の違い

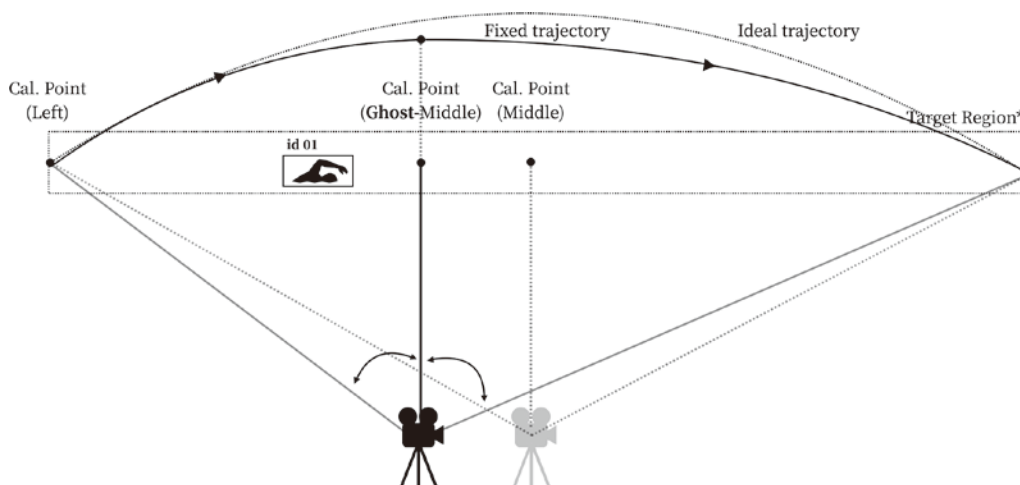


図9：撮影場所がフィールド中央部からずれてしまう場合の模式図

ションポイントを環境に応じて設定する構成は、国外・国内のさまざまな撮影環境に対応するうえで非常に役立った。撮影者は会場の配置によって補正作業が必要になるものの、基本的には手順に沿ってプール上の目標地点とその際のカメラコントロールパラメータ（パン・チルト角）をデータ化するだけで雲台の自動制御を実現できるため、運用の手間や学習コストを最小限に抑えられる。設定が終わると、人手をほとんどかけずに長時間追尾を維持できることから、実際の競技会場での撮影省力化に大きく貢献できるシステムであると感じられた。

考察

本研究で開発したシステムは、機械学習による位置推定結果をパン・チルト制御に反映することで、水泳競技における自動撮影を実現しているが、ここで扱っていない大きな要素として「ズーム制御」がある。今回の実装では、あえてズームの操作を加えず、常に同じ画角・焦点距離で撮影する構成を採用した。これは水平方向（パン）および若干の垂直方向（チルト）を中心とした動きがほとんどで、近距離で撮影するような大きなズーム操作を行わなくとも十分に映像が取得できる前提のシステム設計としたためである。実際、国内外の複数会場でテストを行った結果、1レースに同時に複数選手が泳いでいても、先頭集団を捉える程度の比較的広い画角であれば撮り逃しのリスクは大きくなかった。

しかし、今後より汎用的なスポーツ自動撮影システムを目指す場合や、複数台のカメラを連携させて「特定レーンのみを詳細に追う」「トップと後続集団を別々に収録する」といった運用を行う際には、ズーム制御を含むシステム拡張が求められる。特に、あるカメラが極端にズームインして対象を捉えようとする場合、単にパンとチルトを合わせるだけではターゲットがフレーム外に出てしまう可能性が高まる。そこで、映像から実際に選手が十分な大きさで映っているかどうかを再確認し、必要に応じてズームレベルや画角を自動調整するフィードバック制御を行うことが理想的である。この機能を実装するには、現在の映像を再度解析し、「ターゲットのバウンディングボックスが画面中央付近かつ適切なサイズで映っているか」といった評価を行う仕組みが必要となる。こうした設計は開発コストが高くなるうえ、撮影の場面により連続的な再推定が求められるため、さらなる研究課題として残されている。

また、本システムが実際に撮影した映像データに対し、どこまで高精度な動作解析が可能になるかという点も今後を検討するうえで重要なポイントである。単にフィールド全体を俯瞰で収録しているだけ（図2左 プライマリシステム）では、各選手のフォームや詳細な身体動作を定量的に評価するには解像度不足となりがちだが、本研究の自動追尾によってより近い位置からの映像を取得できるのであれば、ストローク解析やターン動作の評価といった高度な解析を行うための基礎データとして活用できる可能性が高まる。ズームを使わずとも、選手に追従するカメラワーク自体が解像度向上の一助になるからである。

本研究で取り組んだ「パン・チルト制御＋選手位置のAI推定」という枠組みは実運用ベースでは十分な成果を示したものの、ズーム制御や映像へのフィードバック制御など、機能を拡張するうえでの技術的課題が残されている。その一方で、シンプルな設計を心がけたことによる、現場での運用負荷の軽いシステムが開発できたことは本研究の大きな収穫である。前述のように、本システムは認識（プライマリシステム）と制御（セカンダリシステム）を分けて設計実装しているため、映像解析システムとの組み合わせによって、その他のスポーツ競技へと横展開を図ることが期待できると考えられる。

まとめ

本研究では、水泳競技における選手位置認識AIと電動雲台（パン・チルト）の制御システムを独立したモジュールとして構築し、それらをリアルタイムに連携させる方法を提案した。認識モジュールと制御モジュールを切り分け、プール座標系とカメラ座標系の対応付けをキャリブレーションによって事前に確立する設計としたことで、撮影環境（国内外の施設やレーン配置など）が変わっても、システム全体を大きく作り替えることなく運用することが可能となった。実際に複数の会場で試験を行ったところ、異なる撮影位置でも、事前設定を行うだけでスムーズに対象レーンや選手を追尾し続けられることが確認でき、分析スタッフの負担を大幅に削減できる点が明らかとなった。

本研究の独自性は、単一カメラでの選手位置認識をリアルタイムに行いながら、パン・チルト制御側への最小限のデータ送信のみでコントロールを実現する仕組みにある。この構成は認識モデルの更新や雲台等ハードウェア

ア変更に対して柔軟性が高く、追加のモデルや機材を組み合わせる余地を残している。また Ghost-Middle を含む複数の基準点を補完的に用いる座標変換手法により、撮影場所がプールの中央から大きくずれている場合でも選手を追従できた。

一方で、本研究ではズーム制御や映像内容の再評価といった高度なフィードバックは扱っておらず、追尾対象がフレーム内にどの程度大きく映るかなどの細部は今後の課題として残る。本研究のシステムをベースに複数台連携やズーム導入を進めれば、選手に合わせた適切なズームを動的に設定するなど、より汎用的な自動撮影の実現が十分に期待できる。本研究が示した運用上の拡張性は、水泳競技のみならず他種目への展開にも有用であり、今後のスポーツ映像解析や競技支援システム開発に寄与すると考えられる。

謝辞

本研究は、追手門学院大学スポーツ研究センターによる機材貸与と、公益財団法人ヤマハ発動機スポーツ振興財団（スポーツチャレンジ研究助成（基本）16期（2022）/18期（2024））からの助成金の支援を受けて実施したものである。

利益相反

本研究において申告すべき利益相反は存在しない。

参考・引用文献

- 1) 林勇樹, 仲井志文, 今井元太, 松井健 (2023) 機械学習を用いた競泳レース分析システムの開発. *日本水泳・水中運動学会 2023 年次大会論文集*, 76-81.
- 2) Okuma, K., Lowe, D. G., Little, J. J (2013) Self-learning for player localization in sports video. *arXiv preprint arXiv:1307.7198*.
- 3) Arbués-Sangüesa, A., Ballester, C., Haro, G (2019) Single-camera basketball tracker through pose and semantic feature fusion. *arXiv preprint arXiv:1906.02042*.
- 4) Citraro, L., Márquez-Neila, P., Savare, S., Jayaram, V., Dubout, C., et al. (2020). Real-time camera pose estimation for sports fields. *Machine Vision and Applications*, 31 (3), 16.
- 5) Guo, W., Pan, Z., Xi, Z., Tuerxun, A., Feng, et al.

(2024) Sports Analysis and VR Viewing System Based on Player Tracking and Pose Estimation with Multimodal and Multiview Sensors. *arXiv preprint arXiv:2405.01112*.

- 6) 高橋正樹, 中村俊之, 三科智之 (2014) 機械学習を利用した複数視点映像からのサッカーボール追跡. 一般社団法人 映像情報メディア学会 映像情報メディア学会技術報告 38.51, pp. 5-8.
- 7) 今井友揮, 内山彰, 馬込卓弥, 東野輝夫 (2018) サッカートラッキングデータを用いた機械学習に基づくプレー認識手法の提案. *研究報告ユビキタスコンピューティングシステム*, 2018 (47), 1-8.
- 8) 劉靖, 北濱幹士, 山田光穂, 星野祐子 (2022) バスケットボールファウルデータセットとモデル認識. *第20回日本データベース学会年次大会*, 144-18.
- 9) 渡辺裕 (2018) スポーツ情報処理の研究開発動向. *映像情報メディア学会誌*, 72 (6), 905-908.
- 10) 清水千弘, 清田陽司 (2019) エディトリアル: 特集「スポーツ競技と AI」. *人工知能*, 34 (4), 492-496.