

投影歪みを考慮した姿勢推定技術によるサッカーシュートフォーム 分析システム

—骨盤中心座標および画角補正によるフォーム比較—

森 島 園[†] 中村 悠大[‡] 小川 英人[‡] 渡辺 裕[‡]

[†] 早稲田大学基幹理工学部 〒169-0072 東京都新宿区大久保 3-14-9

[‡] 早稲田大学大学院基幹理工学研究科 〒169-0072 東京都新宿区大久保 3-14-9

E-mail: [†] s_morishima.10@akane.waseda.jp

あらまし サッカーのシュート技術向上には、熟練者と一般選手の動作差異の客観的な評価が重要である。本研究では、MediaPipe Pose World Landmarks を用いた三次元姿勢推定に対し、骨盤中心座標への変換および画角補正を適用し、フォームを直感的に比較可能な分析システムを提案する。また、将来的に MediaPipe Pose World Landmarks の三次元推定誤差を低減するモデルを構築するための前提として、正面(0°)を基準に斜め 30°・60°から撮影した場合の推定誤差を評価した。

キーワード: 動作解析, 三次元姿勢推定, MediaPipe, 画角補正

Soccer Shooting Form Analysis System Using Pose Estimation Considering Projection Distortion

—Form Comparison Using Pelvis-Centered Coordinates and Viewpoint Correction—

Sono Morishima[†] Yuta Nakamura[‡] and Eito Ogawa[‡] and Hiroshi Watanabe[‡]

[†] School of Fundamental Science and Engineering, Waseda University 3-14-9 Okubo, Shinjuku-ku, Tokyo 169-0072, Japan

[‡] Graduate School of Fundamental Science and Engineering, Waseda University 3-14-9 Okubo, Shinjuku-ku, Tokyo 169-0072, Japan

E-mail: [†] s_morishima.10@akane.waseda.jp

Abstract To improve soccer shooting technique, it is necessary to objectively evaluate the differences in movement between skilled players and beginners. In this study, we propose a shooting form analysis system based on MediaPipe Pose World Landmarks. By applying a coordinate transformation centered on the pelvis and viewpoint correction, this system enables an intuitive comparison of shooting forms. Furthermore, as a preliminary study toward developing a model to reduce the 3D pose estimation error of MediaPipe Pose World Landmarks, we evaluate the estimation error at camera viewpoints of 30° and 60°.

Keyword Motion Analysis, 3D Pose Estimation, MediaPipe, Field of View Correction

1. はじめに

近年、スポーツ分野においてカメラ映像を用いた動作解析技術の導入が進んでいる。特にサッカーのシュートフォーム解析は、選手の技術向上や怪我の予防において重要視されている。高精度な動作解析にはマルチカメラ環境や光学式モーションキャプチャシステムが用いられてきたが、これらは大規模な設備投資が必要であり、一般的な練習環境での導入は困難である。

これに対し、単一の汎用カメラ映像から物体検出モデル (YOLO 等) を用いてボールを検知し、姿勢推定

ライブラリ (MediaPipe Pose 等) を用いて骨格情報を抽出する簡易的なアプローチに注目した。しかし、2D 平面座標を用いた解析では、カメラの撮影位置や高さなどの画角が異なる動画間を比較する際、撮影角度に起因する投影歪みの影響を強く受ける。これにより、算出される関節角度や身体部位間の距離が実際の身体情報と乖離し、異なる環境で撮影された熟練者と一般選手のフォームを正確に比較・評価できないという深刻な課題があった。

本研究では、この課題を解決するため、MediaPipe

Pose World Landmarks (3D) を用いて、身体基準座標系への正規化、および正面視点への画角補正を導入したシュートフォーム比較システムを提案する。また、提案手法を用いて撮影角度を意図的に変化させた実験を行い、MediaPipe World Landmarks の 3 次元推定自体が持つ画角依存の内的誤差特性について検証・考察した結果を報告する。さらに、この MediaPipe World Landmarks の内的誤差が本システムに与える影響と本システムが実践的な示唆を与えられるかを検証する。

2. 関連研究

本研究では、単一の RGB 画像から身体の主要な 33 点の関節ランドマークを抽出可能な MediaPipe Pose World Landmarks [1] を導入する。本モデルは、画面内のピクセル座標を出力する通常の 2D 姿勢推定とは異なり、骨盤中心を仮想的な空間原点(0, 0, 0)と見立てた、メートル単位の実空間 3 次元座標(x, y, z)を予測出力する特性を持つ。本モデルは、ヒートマップと回帰を組み合わせた学習により精度を維持しつつ、推論時には回帰エンコーダーのみを使用する[2]。これにより、カメラ視点が傾いている場合や、キック動作における脚の前後運動のような奥行き方向の運動に対しても、入力画像内の 2D 特徴パターンの歪みから、AI モデルが内部的にパースペクティブ変化を反転推論し、疑似的な 3D 構造の復元を可能としている。本研究ではこの 3D ランドマークを採用することで、設置画角の幾何学的な投影歪みの影響を低減することを試みた。また、サッカーボールの検出において、物体検出モデルである YOLO [3]を導入した。本モデルは、物体検出を単一の回帰問題として再定義し、1 回のネットワーク処理だけで画像内の全物体の位置とクラス確率を同時に予測するシステムである。従来の複雑なパイプラインを排除したことで、標準モデルで 45fps、軽量版で 155fps という圧倒的なリアルタイム処理スピードを達成している。

3. 提案手法

図 3.1 に本システムの全体イメージ、図 3.2 に処理の全体フローを示す。本システムは、入力された動画からボールおよび身体の 3 次元骨格情報を抽出し、骨格を正面化した上で、熟練者との幾何学的な比較に基づいた自然言語フィードバックを自動生成する。



図 3.1 システム全体イメージ

Figure 3.1 Overview of the Proposed System



図 3.2 本システムの全体フロー

Figure 3.2 Overall Workflow of the Proposed System

3.1 インパクトフレーム推定

YOLO によって、検出されたボールの軌跡および移動速度の変化から、ボールが足に衝突した瞬間であるインパクトフレームを自動推定する。しかし、ボールが明確に確認できない場合など自動推定が不正確な時のために、手動でインパクトフレームを探し、設定することができる機能も含めた。これらのどちらかの手法で設定したインパクトフレームを基準とし、熟練者の動画と一般選手の動画における時間的な対応関係を動的時間伸縮法 (DTW: Dynamic Time Warping) によって同期させることで、動作のタイミングが異なる 2 つの映像間での正確な比較を可能とした。

3.2 骨盤中心座標への正規化

一般選手と熟練者の比較において、無機質な数値データの提示のみに留まると、専門知識を持たない一般選手が自身の課題を直感的に把握し、次の練習へ活かすことが困難である。そこで本システムでは、熟練者のモデルデータと一般選手のデータを幾何学的に比較し、差分が特定の閾値を超えた部位を自動抽出した上で、大規模言語モデルやルールベースのテンプレートを用いて、具体的かつ実践的な自然言語によるアドバイスへと変換する機能を実装した。この自然言語フィ

ードバックを破綻なく機能させるためには、カメラと被験者との距離や、画面内での立ち位置といった外的な変動要因を排除し、純粋な身体運動の差分のみを抽出する必要がある。そこで、本システムにおいて使用した MediaPipe Pose World Landmarks は、特性として骨盤中心を基準とした 3 次元座標を出力する。しかし、サッカーのキックのような高速かつ大きな動的運動においては、推論ノイズにより、各フレームにおける原点位置に微小な揺らぎが生じた。そこで、MediaPipe Pose World Landmarks が提供する 3 次元座標に対し、各フレームで左右の股関節の中点を算出し、この点を空間全体の絶対的な原点 (0, 0, 0) としてすべての関節座標を平行移動する骨盤中心座標への正規化を施した。これにより、世界座標系から身体基準座標系への移行が達成され、骨盤中心から見た各関節の相対位置での厳密な比較が可能となった。また、利き足が異なる場合の対応として、ユーザーが自らの動画と、比較する見本動画について、利き足を設定した上で、それらのキック脚が異なる場合は、ユーザーの動画から抽出された骨格のみを x 軸反転処理することにより、常にキック足を右側として統一する。

3.3 骨格の画角補正

本システムでは、データに不慣れな選手であっても画面を一瞥するだけでフォーム上の幾何学的な修正点を直感的に視認・確認できるよう、熟練者のモデルデータと一般選手の骨格構造を同一画面上へ直接重ね合わせて表示する視覚的フィードバック手法を開発した。骨盤中心座標系を基準として身体の向きを算出し、三次元空間上で回転行列を適用することで、骨格全体をカメラの正面方向へ回転させる正面化（画角補正）処理を行う。本処理により、斜め 30° や 60° といった異なる方向から撮影された動画であっても、身体の向きを統一した共通座標系へ変換される、熟練者と一般選手の骨格構造を直接重ね合わせて比較することが可能となった。

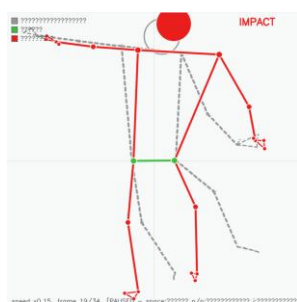


図 3.3 画角補正後の骨格重ね合わせ表示例

Figure 3.3 Example of Overlaying Corrected Skeletons

図 3.3 に、2021-22 シーズンのパリ・サンジェルマンのトレーニング映像（プロの練習を観る，2022）[4]における、2 選手に対して骨格データをそれぞれ割り当てて同一空間上に投影した様子を示す。実線がユーザー、点線が見本動画の骨格である。これにより、指導者や選手本人が画面を一瞥するだけで、熟練者モデルに近づけるためには、「上半身に対して脚を傾けるべき」、「上体を起こし、垂直にすべき」といったフォーム上の幾何学的な修正点を、直感的に視認できるようになった。また、重ね合わせた骨格は、再生速度を変えたり、停止した状態で前後のフレームに戻ったり、進んだりすることができるため、インパクト時以外の情報も得ることができる。

4. 実験

4.1 MediaPipe World Landmarks の内的誤差特性

提案手法の画角補正および MediaPipe Pose World Landmarks の 3 次元推論精度を評価する実験を行う。数理解幾何学的な座標回転において関節角度等の特徴量が不変であることは既知であるが、本モデルの 3D 座標は単一 2D 画像から AI が z 軸を疑似推論した予測値であり、視点変化による内的バイアスの影響を受ける。本実験では、将来的なキャリブレーションフィルタの実装に向け、正面に対して 30°、60° の方向から撮影した場合における、MediaPipe Pose World Landmarks が有する推定誤差の影響を評価する。実験では、比較的単純な動きをした二人の被験者に対してそれぞれ、三脚を正面 (0°)、斜め 30°、60° に設置し、同期撮影した動画群を使用する。評価指標には、正面を真値 (GT) としたときの各関節角度の平均誤差 (deg)、上半身の傾きの平均誤差 (deg)、骨盤の回旋率の平均誤差 (deg) および足首速度の平均誤差 (m/s) を用いる。実験結果を表 1 に示す。

表 4.1 被験者 1 の場合の誤差

Table 4.1 Errors for Subject 1

特徴量	30°		60°	
	補正	なし	なし	あり
左膝角度		11.68	15.20	15.20
右膝角度		10.83	9.52	9.52
左腰角度		8.66	11.65	11.65
右腰角度		8.07	11.75	11.75
左足首角度		29.42	39.99	39.99
右足首角度		14.32	16.38	16.38
上半身傾き		3.32	4.15	4.15
骨盤角度		4.75	4.94	4.94
肩角度		3.90	4.17	4.17
骨盤回旋率		7.74	14.19	14.19
右足速度		0.30	0.35	0.37
左足速度		0.22	0.43	0.51

表 4.2 被験者 2 の場合の誤差

Table 4.2 Errors for Subject 2

特徴量	30°		60°	
	なし	あり	なし	あり
補正				
左膝角度	11.82	11.82	14.76	14.76
右膝角度	6.26	6.26	10.80	10.80
左腰角度	5.36	5.36	10.18	10.18
右腰角度	7.51	7.51	15.24	15.24
左足首角度	12.42	12.42	15.77	15.77
右足首角度	13.69	13.69	16.18	16.18
上半身傾き	3.53	3.53	5.33	5.33
骨盤角度	4.58	4.58	5.94	5.94
肩角度	5.57	5.57	10.09	10.09
骨盤回旋率	5.89	5.89	9.53	9.53
左足速度	0.26	0.18	0.50	0.57
右足速度	0.21	0.24	0.47	0.47

表 4.1, 表 4.2 より, 画角補正の有無に関わらず, 30° および 60° のいずれの条件下でも, 関節角度の平均誤差は, 膝, 腰が約 10 deg 前後, 足首が約 20deg 前後の値を示した. また, 上半身の傾き, 骨盤・肩角度, 骨盤回旋率はいずれも, 約 5deg 前後の平均誤差を示した. よって, これらの項目においては, MediaPipe Pose World Landmarks 自体の深度推論バイアスは被験者や撮影角度 (30°・60°) に関わらず常に同程度に発生することが確認できた. また, 画角補正ありとなしの場合で値が変化しなかったことから, 誤差の主因が画角補正による影響ではなく, MediaPipe Pose World Landmarks 自体の深度推論バイアスであることが確認でき, 深度推論バイアスによる誤差の程度が明らかになった. 一方で, 足首のふり速度に関しては, 撮影角度が大きくなるほど誤差が有意に増加した. これは, キック動作の高速移動に伴うモーションブルー等が AI の深度予測を不安定化させたと考えられ, これが速度特徴量における精度の悪化に繋がったと言える.

4.2 本システムによるフォーム比較の適用評価

4.1 節では画角補正および姿勢推定自体の内的な推定誤差を定量的に検証したが, これは提案システムの基盤をなす精度の観点にとどまる. 本節では, 提案システムを実際の選手ペアに適用し, 生成される自然言語フィードバックが実際のキックフォームと整合し, 指導現場において実践的な示唆を与えられるかを検証する. さらに, 4.1 節で検証した MediaPipe World Landmarks の深度推論バイアスによる誤差が本システムに与える影響を考察した.

4.2.1 適用例

本手法によるフィードバック生成の具体的な適用例として, 2021-22 シーズンのパリ・サンジェルマン

のトレーニング映像 (プロの練習を観る, 2022) [4] から抽出した, 当時既にワールドクラスの熟練者であったキリアン・エムバペ選手とプロ 3 年目の若手であるエリック・ジュニオール・ディナ・エビンペ選手のシュートフォームを対象に, 本システムを適用した. 本実験では, エムバペ選手の蹴り方を模範フォーム (以下, 見本と呼ぶ) とし, エビンペ選手 (以下, ユーザーと呼ぶ) のフォームとの差分を検証した. 両者がインパクトを迎えたフレームにおける計 15 項目の特徴量をシステムが自動抽出したところ, 6 項目 (軸足の左右の開き, キック足の膝の曲げ角度, キック足の振り出し角度, 軸足側の股関節角度, 軸足とキック足の間隔, キック足の振り抜きスピード) は見本との差分が閾値以内であり, 「見本とほぼ一致している」と判定された. 一方, 残る 9 項目については閾値を超える差分が検出され, 以下のようなフィードバックが自動生成された. 生成結果を, キック動作における身体部位のまとめ (軸足まわり・キック足まわり・体幹の回転) ごとに整理して示す.

① 軸足まわり

軸足の踏み込み位置 (前後) は, 見本が $z=-0.47\text{m}$, ユーザーが $z=-0.17\text{m}$ であり, 約 30cm の差が検出された. この差分から, システムは「軸足が見本より体の中心に対して前に出すぎている」ことを検出し, 「もう少し後ろに置きましょう」というフィードバックを生成した. 軸足の高さ (股関節外転具合) についても, 体の中心から踵までの距離が見本で約 0.50m, ユーザーで約 0.58m と約 8cm の差があり, 「体の中心からかかとまでの直線距離が長い」ことが指摘され, 「脚を地面に対して寝かせるようにしてください」という指摘がなされた. 軸足の膝の曲げ角度 (見本 161.1° / ユーザー 147.0°) についても, 「軸足の膝が見本より曲がりすぎている. もう少し伸ばして体を安定させましょう」という指摘が生成された. また, 軸足のつま先の向き (見本 -174.3° / ユーザー -127.2°) についても閾値を超える約 47° の差分が検出され, 「軸足のつま先が見本より内側を向いている. 少し外側に向けましょう」という指摘がなされた.

② キック足まわり

キック足首の角度 (見本 135.7° / ユーザー 103.4°) については, 「キック足首が見本より曲がっている. もう少し足首を伸ばしてボールを捉えましょう」という指摘がなされた. 一方, キック足の膝の曲げ角度 (見本 154.9° / ユーザー 153.7°) およびキック足の振り出し角度 (股関節, 見本 107.8° / ユーザー 114.6°) はいずれも差分が閾値内であり, 「見本とほぼ一致している」と判定された.

③ 体幹の回転

腰と肩のひねり量（見本 31.3° / ユーザー 16.7°）、骨盤の回転角度（見本 91.2° / ユーザー 77.1°）、肩の回転角度（見本 96.5° / ユーザー 80.5°）はいずれも見本を下回っており、それぞれ「腰と肩のひねりが見本より小さい。もっと腰を回してひねりを作りましょう」、「骨盤の開きが見本より足りない。もっと腰を回転させましょう」、「肩の開きが見本より足りない。振り抜きに合わせて自然に開きましょう」という指摘がなされた。また、上体の前後の傾き（見本 177.9° / ユーザー 168.8°）についても、「上体が見本より前に突っ込みすぎている。少し起こしましょう」という指摘が生成された。

④ その他

軸足とキック足の間隔（スタンス幅、見本 0.35m / ユーザー 0.31m）およびキック足の振り抜きスピード（インパクト時点、見本 4.62m/s / ユーザー 5.37m/s）は、いずれも差分が閾値内であり、「見本とほぼ一致している」と判定された。

①から④より、ユーザーのフォームにおける課題は、軸足の踏み込み位置が前すぎる点や、それに伴う体の開き不足、およびつま先の向きの不一致といった「軸足の設置・安定性」に加え、上体の突っ込みや肩・腰のひねり不足といった「体幹の回旋・姿勢」に集中していることが示唆された。また、軸足の膝が曲がりすぎていることで踏み込みが安定せず、結果としてキック時の上体の前傾やひねりの不足を招いている可能性が高い。一方で、キック足の振り抜き角度やスタンス幅が見本と一致している点はポジティブな要素であるが、キック足首が曲がりすぎてボールを捉える際のインパクト面が不安定になっていることも改善の余地がある。複数の特徴量が、軸足の安定性不足とそれに起因する体幹の回旋不足という共通の課題を裏付けており、今後は特に軸足の適切な踏み込み位置と膝の角度、およびインパクト時の上体の起こしを意識することで、より効率的なパワー伝達が可能になると考えられる。

4.2.2 検証結果

生成されたフィードバックの妥当性、および3次元推定の誤差が本システムの実用性に与える影響について評価を行った。生成されたフィードバックの方向性（「伸ばす／曲げる」「離す／近づける」等）とその根拠となる特徴量が、実際の動画の動きと一致しているかを確認したところ、検証対象とした計15件のフィードバックのうち13件で、方向性と差分の符号が一致していることを確認した。一方、実際の動画とフィードバック結果を比較した際に正確でないと思われた項目は、上体の前後の傾きの特徴量（見本 177.9° / ユー

ザー 168.8°）とキック足の振り抜きスピードであった。これらの点については4.2.3で考察する。

4.2.3 考察

本実験の結果は、(1)内部整合性の高さ、(2)姿勢推定誤差の局所的な混入、(3)今後の課題の3点に整理できる。

(1) 内部整合性の高さ

本実験で使用した動画においては、内部整合性の検証において高い一致率が得られ、閾値ベースの判定ロジックが幾何学的な差分を正しく自然言語表現へ変換できていることが確認された。そのため、本システムは、4.1節で検証したMediaPipe World Landmarksの深度推論バイアスによる誤差の影響を受けなかったことが分かる。これは、4.1節の実験結果から得られた3次元推定の誤差は被験者や撮影角度（30°・60°）に関わらず常に同程度に発生するという特性より、見本データとユーザーデータの双方において同様の幾何学的バイアスが同等に生じているため、同一システム内で相対的な比較を行うプロセスにおいてこれらの誤差の影響が軽減されたためである。

(2) 姿勢推定誤差の混入事例

一方で、上体の傾きについて、図3.2に示した骨格と実際の動画には大きな違いは見られないが、測定値が見本もユーザーも非常に大きい値になってしまった。これは、同じ特徴量計算のプログラムを利用している表4.1、表4.2では、膨大な値になることなく、正常に測定できている。よって、動作の複雑化によるMediaPipe World Landmarksの深度推論の限界であると考えられる。実際のシュート動作におけるインパクトの瞬間は、表4.1や表4.2の単純な運動とは異なり、単なる上体の前傾だけでなく、肩のひねりや腕の大きな振りが同時に高速で行われる。単一の2D画像から3D空間を疑似復元するMediaPipe World Landmarksの特性上、この激しい上体のひねりによる2D外観の変化を、AIモデルが前後の傾き（奥行き方向の変化）として誤認・混同して推論してしまった可能性が高い。また、速度系の値については、4.1節で述べたように、撮影角度が大きくなるほど誤差が有意に増加するため、信頼度の高い数値でないと分かる。そのため、本実験において、正常に見える値になったが、他動画での比較・検証が必要である。さらに、画質の粗い動画に対しては、MediaPipe World Landmarksによる骨格抽出が難しく、壊れてしまう場合があった。

(3) 今後の課題

上記のような姿勢推定が不安定になりやすい場面への対応として以下の3点があげられる。

- 上体のひねりと前後の傾きを分離するため、胸椎

- のひねり角も加味した座標変換を行う。
- 画質が悪く骨格抽出が不可能なフレームについて、前後数フレームとの角度変化量を監視し、1フレームだけ極端に不連続な値が出た場合は自動的にフラグを立て、前後フレームの値で補間する。
- 本システムを実践環境への適用させるために、改善のための具体的な練習メニューの提示までを行う。

5. まとめ

本稿では、汎用的な単一カメラ映像から、撮影画角の影響を抑制してサッカーのシュートフォームを客観的に比較・評価するシステムを提案した。骨盤中心を原点とする身体座標系への変換と回転行列による正面視点補正を導入することで、熟練者と一般選手のフォームを幾何学的に重ね合わせ、数値および直感的なフィードバックを提示することを可能とした。また、画角変化に対する精度の検証実験により、MediaPipe Pose World Landmarks を用いた解析では、AI モデル固有の一定の内的推論誤差が、 30° および 60° の撮影画角で定常的に存在するという重要な知見を得た。さらに、実際のシステム適用を通じて、フィードバックは実際の動画に対しておおよそ妥当であり、MediaPipe Pose World Landmarks の内的誤差による影響は受けていないことが確認できた。

今後の展望としては、実戦環境への適用として、本システムを様々な場面に対応できるように、より頑健なシステムとなることを目標とする。さらに、この AI 由来の定常的な内的誤差を事前に数理モデル化し、撮影角度に応じてキャリブレーションを施すフィルターアルゴリズムを導入することで、さらなる解析精度の極限的な向上を目指す。

文 献

- [1] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chan, J. Lee, W.-T. Chang, W. Hua, M. Georg, M. G. Yong, and M. Grundmann : “MediaPipe: A Framework for Building Perception Pipelines”, Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops, (June. 2019)
- [2] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann : “BlazePose: On-device Real-time Body Pose Tracking”, Proc. 4th CVPR Workshop on Fourth Workshop on Computer Vision for AR/VR, (June. 2020)
- [3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi : “You Only Look Once: Unified, Real-Time Object Detection”, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp.779-788, (June.

2016)

- [4] プロの練習を観る:"パリスサンジェルマン シュート練習 2021.12~2022.8 PSG: Shooting Practice", <https://www.youtube.com/watch?v=0G112gkwV1k> (参照日：2026 年 7 月 13 日)