

単眼3次元姿勢推定を用いたACL損傷リスク評価に関する一検討 A Study on ACL Injury Risk Assessment Using Monocular 3D Human Pose Estimation

宮坂 駿輔[†] 和田 直史[†] 竹沢 恵[†] 稲垣 潤[†] 井野 拓実[‡]
Shunsuke Miyasaka Naofumi Wada Megumi Takezawa Jun Inagaki Takumi Ino

1. はじめに

膝前十字靭帯 (ACL) 損傷はスポーツ現場における重大な外傷の一つであり、そのリスクを定量的に評価することは予防の観点から重要である。近年、単眼カメラ映像からの3次元姿勢推定技術が注目されており、低コストかつ簡便に骨格情報を取得できることから、臨床やスポーツ現場への応用が期待されている。一方で、推定精度や撮影条件に関する検証は十分とは言えず、実用化に向けた検討が求められる。本研究では、Drop Vertical Jump テスト [1] を対象とし、光学式3次元動作解析装置による計測結果を基準として、単眼3次元姿勢推定による膝関節角度評価の有効性を検証する。

2. Drop Vertical Jump テスト

Drop Vertical Jump テスト (以下, DVJ) は, ACL 損傷リスク検査方法の一つである。DVJでは, 30cm 台から落下・着地後, 直ちに最大垂直跳びを行い, その際の動作姿勢を評価する。近年は着地時の膝関節角度に着目した報告が散見される。一般的には, 膝屈曲角度が浅く, さらに内側に入った姿勢において, より高い ACL 損傷発症率が報告されている。また, 女性は男性と比較して2~8倍 ACL 損傷が起こりやすいとされている。

3. 光学式3次元動作解析装置

光学式3次元動作解析装置は, 人体に取り付けた反射マーカーが反射した赤外線を周囲の赤外線カメラで受光することでモーションキャプチャーを行う。本研究では, Vicon Motion System 社製の VICON を使用する。また, 床反力計を使用し, 床にかかる力の大きさ, 向きを同時に測定する。光学式3次元動作解析装置は, 測定精度が高いが非常に高価であるため, 病院やフィールドで利用することが難しく, 機器を扱えるオペレータも多くない。

4. 3次元姿勢推定

3次元姿勢推定は, 映像から人体の関節や部位の3次元空間上の位置を推定する技術である [2, 3]。図1に3次元姿勢推定の例を示す。3次元姿勢推定には, 大きく分けて複数カメラを用いる手法と単眼カメラを用いる手法が存在する。本研究では, 導入コストや設置の制約が少ない単眼カメラ映像による3次元姿勢推定を用い, その精度を VICON と比較評価する。

5. 膝関節角度の算出方法

本研究では, DVJ における膝関節角度を評価対象とする。計測環境を図2に示す。図2のように, VICON を用

[†]北海道科学大学大学院 工学研究科 情報工学専攻

[‡]北海道科学大学大学院 保健医療学研究科 リハビリテーション科学専攻

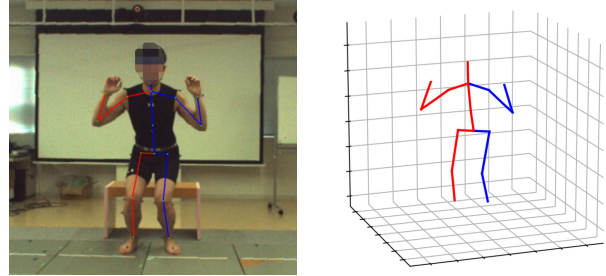


図1: 3次元姿勢推定結果の例

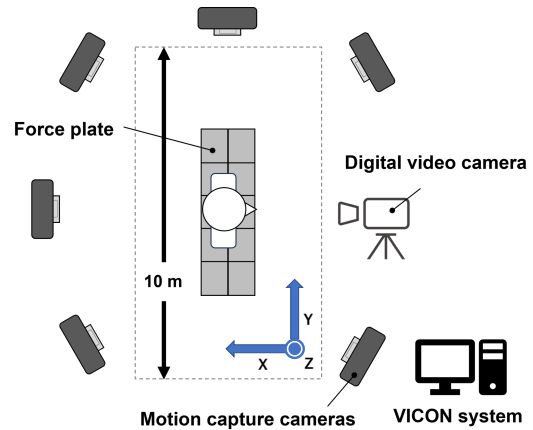


図2: VICON による計測環境

いて DVJ の動作解析を行うと同時に, 被験者の正面に配置した高速度カメラで映像を撮影する。撮影した映像に対して3次元姿勢推定手法を適用し, 各関節の3次元座標 (XYZ 座標) を求める。膝関節角度は, 文献 [4] を参考に, 股関節, 膝関節, 足首の3点の座標から算出する。膝関節を原点 O とし, 股関節方向のベクトルを $\mathbf{v}_1 = (x_1, y_1, z_1)$, 足首方向のベクトルを $\mathbf{v}_2 = (x_2, y_2, z_2)$ とすると, 膝関節角度 θ は次式で定義される。

$$\theta = \cos^{-1} \left(\frac{-\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\| \|\mathbf{v}_2\|} \right) \quad (1)$$

6. 評価実験

6.1. 実験条件

被験者は18歳から22歳の男女23名 (男性10名, 女性13名) であり, 各被験者が複数回試行した合計66試行のデータを使用した。入力映像には, 図3左のオリジナル映像に加えて, 図3右のように人物領域をクロップして拡大した映像を使用した。ここでは, YOLOv8 を用いて人物を検出し, 全てのフレームで全身が含まれるよう人物領域を決定した。いずれも解像度は 1280×720 画素, フレームレートは 120fps である。使用した3次元姿勢推定手法は, Transformer ベースの代表的な手法である MHFormer [5],

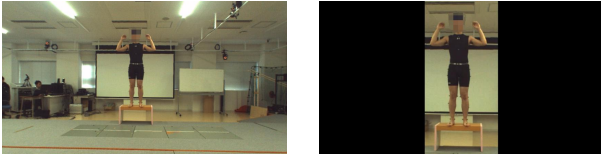


図3: オリジナル映像(左)とクロップ後の映像(右)

表1: 膝関節角度推定誤差 (RMSE \pm std.)

Method	Overall	Male	Female
MHF	21.6 \pm 9.4	28.1 \pm 7.3	16.6 \pm 7.6
MHF (with crop)	14.8 \pm 8.4	20.8 \pm 9.2	10.2 \pm 3.2
MAGF	30.9 \pm 8.0	35.1 \pm 8.2	27.7 \pm 6.1
MAGF (with crop)	26.9 \pm 8.4	29.9 \pm 8.9	24.5 \pm 7.3
MixSTE	20.2 \pm 7.6	25.9 \pm 6.1	15.7 \pm 5.2
MixSTE (with crop)	16.6 \pm 8.2	22.1 \pm 8.3	12.4 \pm 5.1

MotionAGFormer[6], MixSTE[7]の3手法を用いた。3次元姿勢推定で求めた膝関節角度をVICONを基準とした二乗平均平方根誤差(RMSE)により評価した。誤差の評価区間は、床反力計の計測データに基づき、着地直後からジャンプ直前までのフレームを対象とした。

6.2. 実験結果

各手法における膝関節角度推定誤差の平均および標準偏差を表1に示す。全ての手法において、クロップ後の映像の方がRMSEが小さくなり、前処理の有効性が確認された。また、性別ごとの比較では、すべての条件で女性の方が男性よりもRMSEが小さくなる傾向が見られた。

図4は、MHFormer (with crop)におけるRMSEごとの被験者数のヒストグラムである。図4より、女性は10度前後に集中している一方、男性はばらつきが大きく、30度を超える被験者も見られた。

図5は、男女別のRMSEとVICONによる膝関節角度の時間変化を示している。破線はRMSE、実線は膝関節角度、赤線は女性、青線は男性を示す。横軸は動作進行度(評価区間を0~100%に正規化したもの)を表す。図5より、RMSEは中盤(屈曲時)で小さく、動作の開始と終了時(伸展時)で大きくなる傾向が男女共通して確認された。

6.3. 考察

表1に示すように、すべての手法において、人物を中心にクロップした前処理映像を用いた条件の方が、オリジナル映像よりもRMSEが低かった。これは、人物が画面中央に大きく映ることで、モデルが被写体の姿勢情報をより正確に捉えやすくなったことが要因と考えられる。

性別による比較では、表1に示す通り、すべての条件で女性の方が男性よりもRMSEが小さく、推定精度が高い傾向が確認された。さらに、図5において、動作開始直後(0%付近)および終了直前(90%付近)では、男女のVICONによる膝角度はほぼ一致しているにもかかわらず、推定誤差には15~20度の差が生じており、女性の方が一貫して高い推定精度を示した。性別による違いの背景には様々な要因が考えられるが、その詳細な検討は今後の課題である。

また、図5では、RMSEの時間変化とVICONによる膝

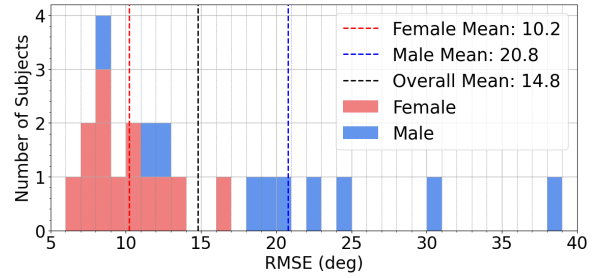


図4: MHFormer (with crop)における各被験者のRMSE

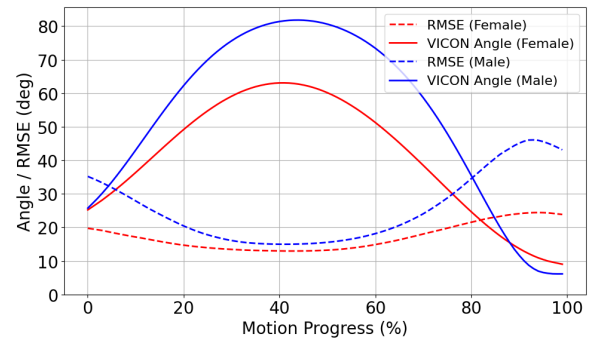


図5: VICONによる膝関節角度とRMSEの時間変化

関節角度の平均値を重ねて描画した。その結果、膝が大きく屈曲する中盤では誤差が小さく、伸展している動作の開始直後および終了直前では誤差が大きくなる傾向が、男女共通で確認された。これは、屈曲時には関節の視覚的変化が大きく、モデルが角度を捉えやすいのに対し、伸展時には関節形状の変化が少なく、推定が難しくなるためと推察される。

7. おわりに

本研究では、単眼3次元姿勢推定手法を用いて、Drop Vertical Jumpテストにおける膝関節角度の推定精度を検証した。その結果、人物を中心にクロップする前処理が推定精度の向上に有効であることが確認できた。また、女性の方が男性よりも誤差が小さくなる傾向や、膝屈曲時の方が膝伸展時よりも誤差が小さくなる傾向も確認された。今後は、データ数を増やして男女間の推定精度の差についてさらに検証を行う。

参考文献

- [1] F. R. Noyes, et al.: "The Drop-Jump Screening Test Difference in Lower Limb Control By Gender and Effect of Neuromuscular Training in Female Athletes" The American Journal of Sports Medicine, 33(2), (2005).
- [2] M. M. Azam, et al.: "A Survey on 3D Egocentric Human Pose Estimation" arXiv:2403.17893, (2024).
- [3] C. Zheng, et al.: "Deep Learning-Based Human Pose Estimation: A Survey" arXiv:2012.13392, (2023).
- [4] T. E. Hewett, et al.: "Biomechanical Measures of Neuromuscular Control and Valgus Loading of the Knee Predict Anterior Cruciate Ligament Injury Risk in Female Athletes," The American Journal of Sports Medicine, 33(4), (2005).
- [5] W. Li, et al.: "MHFormer: Multi-Hypothesis Transformer for 3D Human Pose Estimation," CVPR, (2022).
- [6] S. Mehraban, et al.: "MotionAGFormer: Enhancing 3D Human Pose Estimation With a Transformer-GCNFormer Network," WACV, (2024).
- [7] J. Zhang, et al.: "MixSTE: Seq2seq Mixed Spatio-Temporal Encoder for 3D Human Pose Estimation in Video," CVPR, (2022).