

20XX 年度卒業論文

低照度環境下の 2 次元人物姿勢推定における構造的
一貫性を用いた誤推定骨格の除去

中部大学 工学部 情報工学科

EX00000 情報太郎

20XX 年 2 月

指導教授 宇佐美 裕康

[架空の記入例・提出不可]

1. はじめに

2 次元人物姿勢推定は、行動認識や作業支援へ人体構造を与える基盤技術である。低照度では局所コントラストが低下し、関節信頼度が高いまま肢長比の崩れた骨格を出力することがある。本研究では既存推定器を固定し、関節信頼度と骨格構造のー貫性から誤推定骨格を除外して、採用骨格の精度と採用率の関係を調べる。

対象は単一人物を含む切出し画像とし、提案処理は関節座標を補正せず、骨格を後段へ渡すか否かだけを判定する。したがって目的は全フレームの位置精度向上ではなく、許容する採用率の下で誤推定の流入を抑えることである。なお、以下の研究設定、データ、実験条件、数値および結果はすべて架空である。

2. 関連研究

HRNet は高解像度表現と複数解像度の反復融合により、高精度な関節位置推定を実現した [1]。ViTPose は Vision Transformer と軽量の復号器による単純な構成でも高い性能を示した [2]。本研究はこれらと異なり、推定器に依存しない出力判定を追加して、撮影環境の劣化時に後段へ渡す骨格を選別する。

3. 提案・実施内容

提案処理の流れを図 1 に示す。固定した姿勢推定器から $J = 17$ 個の関節座標 $\hat{\mathbf{p}}_j$ と信頼度 c_j を取得し、人物単位の骨格品質スコア q を算出する。

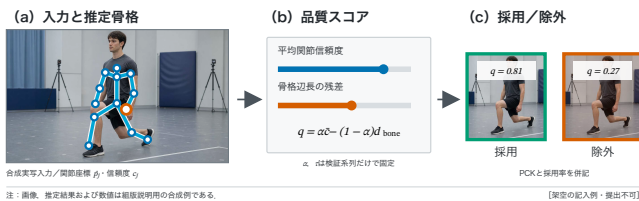


図 1: 誤推定骨格を除外する処理の流れ

平均関節信頼度を $\bar{c}$ 、胴体長で正規化した骨格辺 e の長さを $\tilde{\ell}_e$ 、学習系列におけるその中央値を μ_e とする。 $\mathcal{E}$ は COCO 骨格の 16 辺とし、胴体長には左右肩中点と左右腰中点の距離を用いる。信頼度 0.1 未満の関節は欠損として $c_j = 0$ とし、胴体長が画像高の 5% 未満なら正規化不能として除外する。骨格辺集合 $\mathcal{E}$ に対し、

$$d_{\text{bone}} = \frac{1}{|\mathcal{E}|} \sum_{e \in \mathcal{E}} |\tilde{\ell}_e - \mu_e|, \quad (1)$$

$$q = \alpha \bar{c} - (1 - \alpha) d_{\text{bone}}$$

と定義する。式 (1) の $q < \tau$ を除外し、 α と τ は検証系列だけで決める。

μ_e は学習系列だけから算出し、検証系列では α と τ の組を 0.05 刻みで探索する。評価前に推定器、 μ_e 、 α 、 τ を固定し、評価結果を見た再調整は行わない。図 1 の判定はフレーム単位であり、前後フレームの推定結果は用いない。

4. 実験・評価

20–40 lx の室内で 12 名から各 1 系列 200 フレーム、計 2,400 フレームを収録し、各フレームの 1 名に 17 関節を付与した。人物単位で学習、検証、評価へ 4 系列ずつ分け、同一人物と隣接フレームの混入を防いだ。照度 20, 30, 40 lx と直立、歩行、着座が各集合へ含まれるよう割り当てた。COCO 事前学習済み HRNet-W32 を 256×192 画素入力で固定し、除外なし、平均信頼度、提案スコアを比較した。検証系列で採用率 80% 以上の PCK@0.2 を最大

化し、 $\alpha = 0.65$ 、 $\tau = 0.42$ を得た。PCK@0.2 は胴体長で正規化し、採用骨格だけで算出する一方、採用率を必ず併記した。系列ごとに指標を求め、4 系列を等重みで平均した結果を表 1 に示す。

表 1: 低照度評価系列における誤推定骨格の除外結果 (架空値)

除外方法	PCK@0.2 [%]	採用率 [%]
なし	73.1	100.0
平均信頼度	78.2	86.4
提案スコア	82.1	81.7

提案スコアは評価 800 フレーム中 654 フレームを採用した。系列別 PCK@0.2 は 79.8–84.3%、採用率は 77.5–85.0% であり、平均値が単一系列だけで生じたものではなかった。除外 146 フレームの目視分類では、遮蔽または動きぼけを伴う末端関節の誤推定が 91、正しいが学習時の辺長分布から外れた姿勢が 24、人物切出しのずれ等が 31 であった。

5. 考察

表 1 より、提案スコアは除外なしに対して PCK@0.2 を 8.9 ポイント高めたが、採用率は 18.3 ポイント低下した。平均信頼度による除外と比べても PCK@0.2 は 3.9 ポイント高く、骨格辺長が高信頼度の誤推定を補足した可能性がある。一方、24 フレームの有効姿勢を除外しており、学習分布から外れる屈曲姿勢を誤りとみなす偽除外が残る。また、関節間比率を保ったまま骨格全体が平行移動する誤推定、胴体の遮蔽、人物切出しの失敗は式 (1) だけでは検出にくい。本手法は位置を修正せず、難しいフレームを捨てて精度と採用率を交換するため、欠損を許容できない用途にはそのまま適用できない。単一室内・固定照度・単一推定器での評価であり、カメラ、人物属性、屋外光、時系列判定、後段タスクへの一般化は未検証である。

6. おわりに

関節信頼度と正規化骨格辺長を統合し、低照度映像の誤推定骨格を除外した。信頼できる骨格だけを後段へ渡す構成には有用性があるが、実運用には複数カメラでの再評価と欠損骨格を扱う設計が必要である。

参考文献

- [1] K. Sun, B. Xiao, D. Liu, and J. Wang, “Deep high-resolution representation learning for human pose estimation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5693–5703.
- [2] Y. Xu, J. Zhang, Q. Zhang, and D. Tao, “ViTPose: Simple vision transformer baselines for human pose estimation,” in *Advances in Neural Information Processing Systems*, vol. 35, 2022, pp. 38571–38584.

研究業績

[架空の記入例・提出不可]

1. 情報太郎, 宇佐美 裕康, 「低照度映像の 2 次元姿勢誤推定除去」, 令和 X 年度 電気・電子・情報関係学会東海支部連合大会, A1-1, 20XX 年 9 月。口頭発表 (査読なし)。
2. 情報太郎, 宇佐美 裕康, 「構造指標による人物骨格品質評価」, 情報処理学会 第 XX 回全国大会, 1Z-01, 20XX 年 3 月。口頭発表 (査読なし)。

目 次

第 1 章	はじめに	1
1.1	研究背景	1
1.2	研究目的	2
1.3	本研究の貢献	2
1.4	本論文の構成	2
第 2 章	背景・関連研究	3
2.1	2 次元人物姿勢推定	3
2.2	低照度環境における誤推定	4
2.3	本研究の位置づけ	4
第 3 章	問題設定	5
3.1	対象と仮定	5
3.2	記号と定式化	5
3.3	評価課題	6
第 4 章	提案手法	7
4.1	全体構成	7
4.2	骨格品質スコア	7
4.3	実装条件	8
第 5 章	実験・評価	9
5.1	データと評価手順	9
5.2	比較条件と評価指標	9
5.3	結果	10
5.4	考察	10
第 6 章	おわりに	11
6.1	本研究のまとめ	11
6.2	今後の課題	12
	生成 AI 等の利用に関する申告	13
	謝辞	14
	参考文献	15

第 1 章 はじめに

[架空の記入例・提出不可]

1.1 研究背景

人物姿勢推定は、画像中の肩、肘、膝等の関節位置を求め、人物の動作を骨格として表現する技術である。高解像度表現を維持する HRNet[1] や Transformer を用いる ViTPose[2] により、明瞭な人物画像に対する推定精度は向上している。一方、監視映像や夜間の行動計測では、照度不足によるコントラスト低下、動きぼけ、人物間の重なりが同時に生じる。このとき、存在しない位置へ関節が集まった骨格や、左右の手足が入れ替わった骨格が出力され、人数計測や動作分類等の後段処理へ誤りが伝播する。

姿勢推定器を対象環境ごとに再学習する方法は有力であるが、十分な低照度画像と関節アノテーションを準備できない場面もある。また、推定器が出力する関節信頼度だけで骨格を選別すると、一部の関節が高信頼度であっても人体として不自然な骨格を採用する場合がある。そこで、既存の推定器を変更せず、出力された骨格の構造的ー貫性を利用して誤推定を除外する後処理を検討する。

1.2 研究目的

本研究の目的は、低照度の単眼 RGB 動画を対象として、関節信頼度と骨格辺長の一貫性から誤推定骨格を判定する方法を構築し、精度と採用率の両面から有効性を評価することである。ここで誤推定の除外とは、関節座標を補正することではなく、後段処理へ渡す骨格を選択することを指す。したがって、採用した骨格の位置精度が高くても、多数の骨格を捨てる方法は実用上十分とは限らない。本研究では、検証系列だけで判定閾値を決め、未使用の評価系列において PCK@0.2 と採用率を同時に報告する。

1.3 本研究の貢献

本例で想定する貢献は、次の 3 点である。

1. 関節信頼度と正規化骨格辺長を統合した、推定器非依存の骨格品質スコアを定義した。
2. 動画系列単位で学習・検証・評価を分離し、低照度および動きぼけ条件に対する評価手順を示した。
3. 除外による PCK@0.2 の変化だけでなく採用率も併記し、後段処理へ与える欠損とのトレードオフを整理した。

これらは卒業論文の記述例として設定した架空の貢献であり、実在する研究成果ではない。

1.4 本論文の構成

第 2 章では人物姿勢推定と誤推定除外に関する技術的背景を述べ、第 3 章で入力、出力および評価課題を定義する。第 4 章では骨格品質スコアと判定処理を示し、第 5 章で架空データを用いた比較結果を報告する。第 6 章では得られた知見と限界をまとめ、今後の課題を示す。

第 2 章 背景・関連研究

[架空の記入例・提出不可]

2.1 2 次元人物姿勢推定

2 次元人物姿勢推定では，画像上の人物ごとに複数の関節座標を求め，所定の接続関係に従って骨格を構成する．代表的な構成には，人物領域を検出した後で個人ごとに関節を推定する top-down 方式と，画像全体から関節候補を検出して人物単位に対応づける bottom-up 方式がある．Part Affinity Fields[3] は，関節候補に加えて関節間の結びつきを表す場を推定し，複数人物の骨格を構成する bottom-up 方式である．

HRNet[1] は，解像度の異なる表現を並列に保持しながら情報交換することで，位置情報を必要とする姿勢推定に高解像度表現を利用する．ViTPose[2] は，Vision Transformer を基盤とした簡潔な姿勢推定の構成を示している．本研究は，これらの推定器の内部構造を変更するものではない．推定器から人物ごとの関節座標と関節信頼度を取得することだけを前提とし，出力後の骨格を処理対象とする．

2.2 低照度環境における誤推定

低照度画像では，センサ雑音とコントラスト低下により関節周辺の局所特徴が不明瞭になる．さらに，露光時間の増加による動きぼけや，暗部での人物領域検出の欠落が重なると，関節位置だけでなく人物と関節の対応も不安定になる．その結果，身体の一部だけが別人物へ接続された骨格，胴体に対して手足が過度に長い骨格，複数関節が同一点付近へ縮退した骨格が生じ得る．

関節信頼度の平均に閾値を設ける方法は簡便であるが，各関節の局所的な確からしさしか利用しない．一方，人体骨格は関節間の接続関係をもち，同一人物内の正規化辺長は一定の範囲に収まると考えられる．そこで本例では，関節信頼度を局所的指標，骨格辺長の基準からの逸脱を構造的指標として分け，両者を統合して骨格単位の利用可否を判定する．

2.3 本研究の位置づけ

Part Affinity Fields[3] が関節候補から人物骨格を組み立てる際の対応を扱うのに対し，本研究は既に構成された骨格を採用するか除外するかを判定する．また，HRNet[1] や ViTPose[2] のような姿勢推定モデルを低照度データで再学習するのではなく，モデルの出力だけで完結する後処理を対象とする．このため導入は容易である一方，除外された骨格を復元できず，姿勢推定器自体の位置誤差も改善しない．本例では，この適用範囲を越えて性能を主張せず，採用骨格の精度と欠損量を同時に評価する．

本章の低照度条件，想定する誤推定および研究上の差分は，論文構成を示すための架空設定である．

第 3 章 問題設定

[架空の記入例・提出不可]

3.1 対象と仮定

入力は，固定カメラで撮影した低照度の単眼 RGB 動画とする．各フレームに既存の 2 次元姿勢推定器を適用し，検出人物ごとに J 個の関節座標と信頼度を得る．カメラ校正値，深度画像および時間的に前後するフレームは，骨格の採否判定には用いない．姿勢推定器の学習および人物領域の検出は本研究の対象外とし，少なくとも胴体長を定義する関節が出力されることを仮定する．

低照度条件では人物そのものが未検出となる場合もあるが，本研究が判定できるのは推定器から出力された骨格だけである．したがって，未検出人物の回復と，出力済み骨格に含まれる誤推定の除外を区別する．また，評価時に正解関節が与えられるが，品質スコアの算出時には正解を使用しない．

3.2 記号と定式化

フレーム t で検出された人物 n の関節 j について，推定座標を $\hat{p}_{tnj} \in \mathbb{R}^2$ ，信頼度を $c_{tnj} \in [0, 1]$ とする．骨格辺集合を $\mathcal{E}$ とし，胴体長で正規化した辺 $(j, k) \in \mathcal{E}$ の長さを ℓ_{tnjk}

とする．学習系列から求めた辺長の基準値を μ_{jk} とし，評価系列の情報は基準値の算出に用いない．

骨格選別器 g は，関節座標と信頼度を入力として採用フラグ $a_{tn} \in \{0, 1\}$ を返す．

$$a_{tn} = g(\{\hat{\mathbf{p}}_{tnj}, c_{tnj}\}_{j=1}^J, \{\mu_{jk}\}_{(j,k) \in \mathcal{E}}). \quad (3.1)$$

式 (3.1) において， $a_{tn} = 1$ なら骨格を後段処理へ渡し， $a_{tn} = 0$ なら誤推定候補として除外する．目標は，正しい骨格をできる限り残しながら，大きな関節位置誤差または人体構造からの逸脱をもつ骨格を除外することである．第 4 章では， g を関節信頼度と正規化辺長の逸脱から構成する．

3.3 評価課題

正解関節座標を $\mathbf{p}_{tnj}$ ，人物の基準長を s_{tn} とする．関節 j が正しい条件を

$$\|\hat{\mathbf{p}}_{tnj} - \mathbf{p}_{tnj}\|_2 \leq 0.2s_{tn} \quad (3.2)$$

とし，採用骨格に含まれる可視関節について式 (3.2) を満たす割合を PCK@0.2 として集計する．同時に，全検出骨格数に対する $a_{tn} = 1$ の骨格数の割合を採用率とする．前者は残した骨格の位置精度，後者は除外による情報欠損を表すため，どちらか一方だけでは選別器を評価しない．

判定閾値は検証系列で固定し，通常照明，低照度および動きぼけを含む評価系列へ同じ値を適用する．主要比較は，除外なし，平均関節信頼度による除外，構造的ー貫性を統合した提案法の 3 条件とする．本章の入力条件，記号の運用，閾値決定および評価課題は，すべて架空の記入例である．

第 4 章 提案手法

[架空の記入例・提出不可]

4.1 全体構成

本研究では，姿勢推定器自体を再学習せず，推定された関節座標と信頼度から誤推定骨格を除外する．図 4.1 に処理の全体像を示す．入力フレームから既存の姿勢推定器を用いて骨格を求めた後，関節信頼度と骨格辺長の構造的ー貫性を統合し，後段処理へ渡す骨格を選択する．この構成により，HRNet[1] や ViTPose[2] 等の推定器を変更せずに適用できる．

4.2 骨格品質スコア

人物ごとに J 個の関節を推定し，関節 j の座標を $\hat{\mathbf{p}}_j \in \mathbb{R}^2$ ，信頼度を $c_j \in [0, 1]$ とする．平均関節信頼度を $\bar{c} = J^{-1} \sum_{j=1}^J c_j$ とする．また，骨格辺集合を $\mathcal{E}$ とし，各辺長を人物の胴体長で正規化した後，学習系列から求めた基準辺長との差の平均を d_{bone} とする．骨格品質スコア q を

$$q = \alpha \bar{c} - (1 - \alpha) d_{\text{bone}} \quad (4.1)$$

と定義する．式 (4.1) の第 1 項は関節単位の確からしさ，第 2 項は人体構造からの逸脱を

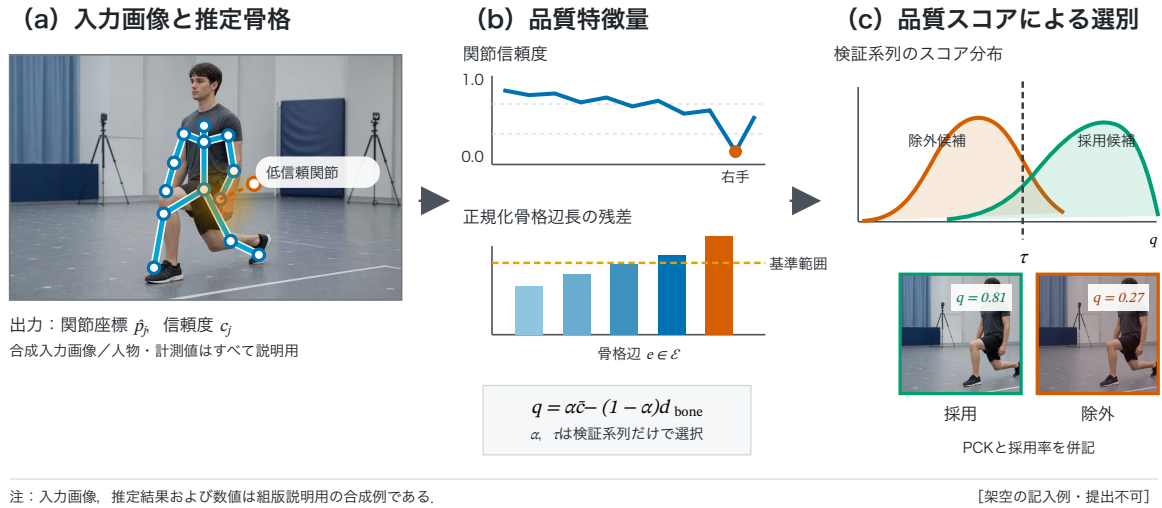


図 4.1: 構造的ー貫性を用いた誤推定骨格除外の処理フロー

表す． $q < \tau$ となる骨格を誤推定候補として除外する．この判定は，Part Affinity Fields で関節間の対応を扱う方法 [3] とは異なり，既に構成された骨格の利用可否を判定する後処理である．

4.3 実装条件

記入例では，姿勢推定器を HRNet-W32，入力解像度を 256×192 画素とした．重み α と閾値 τ は検証系列の PCK@0.2 と採用率の関係から決め，評価系列には触れない．基準辺長は学習系列だけから計算し，左右反転時は関節番号も対応づける．なお，モデル，解像度，データ数およびパラメータは説明用の架空設定である．

第 5 章 実験・評価

[架空の記入例・提出不可]

5.1 データと評価手順

説明用に設定した架空の室内動画 12 系列，2,400 フレームを用いた．同じ撮影系列の隣接フレームが異なる分割へ混入しないよう，動画系列単位で学習 4 系列，検証 4 系列，評価 4 系列へ分けた．評価系列には，通常照明，低照度，動きぼけの 3 条件を含めた．図 5.1 に示すように，検証系列で閾値を決定した後，評価系列ではパラメータを固定して指標を集計した．

5.2 比較条件と評価指標

比較条件は，骨格を除外しない条件，平均関節信頼度だけで除外する条件，式 (4.1) の提案スコアで除外する条件の 3 つとした．採用した骨格の関節位置精度を PCK@0.2 で評価し，これと併せて全検出骨格に対する採用率を報告した．PCK のみを最大化すると難しいフレームを多数除外できるため，採用率を必ず同時に確認する．

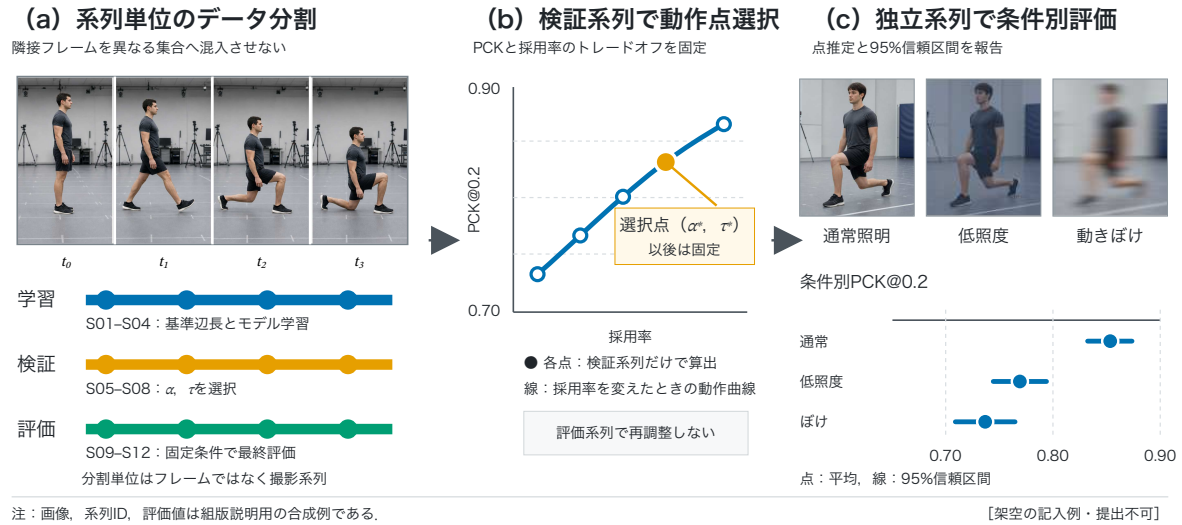


図 5.1: データ分割と評価手順. 系列間の情報漏洩を避けて評価する

表 5.1: 低照度条件における誤推定骨格除外の結果 (架空値)

除外方法	PCK@0.2 [%]	採用率 [%]
なし	73.1	100.0
平均関節信頼度	78.2	86.4
提案スコア	82.1	81.7

5.3 結果

低照度条件の架空結果を表 5.1 に示す. 提案スコアは, 除外なしと比べて採用骨格の PCK@0.2 が 8.9 ポイント高かった一方, 採用率は 18.3 ポイント低下した. 平均関節信頼度と比べると, PCK@0.2 は 3.9 ポイント高く, 採用率は 4.7 ポイント低かった. これは, 骨格辺長の項が信頼度だけでは除外できない構造的な誤推定にも反応した可能性を示す.

5.4 考察

表 5.1 は, 提案スコアが低精度骨格を優先して除外した可能性を示すが, 姿勢推定器自体の改善を意味しない. 誤除外, 見逃し, 採用率および後段性能を併せて閾値を選ぶ.

第 6 章 おわりに

[架空の記入例・提出不可]

6.1 本研究のまとめ

本研究では、低照度の単眼 RGB 動画から得られた 2 次元人物骨格を対象として、誤推定骨格を後段処理の前に除外する方法を検討した。既存の姿勢推定器は再学習せず、人物ごとの平均関節信頼度と、胴体長で正規化した骨格辺長の基準からの逸脱を統合して骨格品質スコアを定義した。さらに、動画系列単位で学習・検証・評価を分け、検証系列で決定した閾値を固定して評価系列へ適用した。

架空の低照度条件では、提案スコアにより採用骨格の PCK@0.2 が除外なしより 8.9 ポイント、平均関節信頼度だけによる除外より 3.9 ポイント高くなった。一方、採用率は除外なしより 18.3 ポイント低下した。この結果は、構造的に不自然な骨格を優先的に除外できた可能性を示すが、姿勢推定器自体の精度向上を意味しない。また、精度の向上と引き換えに骨格の欠損が増えるため、後段処理が許容できる採用率を考慮して閾値を選ぶ必要がある。

以上より、本例の範囲では、関節信頼度と骨格構造を併用することが出力骨格の品質管理に有用である可能性を確認した。ただし、使用した動画、数値および結論は組版と論述のための架空例であり、実在する研究成果ではない。

6.2 今後の課題

第 1 に，時間情報を利用していないため，単一フレームでは低品質でも前後フレームから補間可能な骨格を除外する場合がある．骨格軌跡の速度と加速度を用いた判定を追加し，誤除外と見逃しの変化を確認する必要がある．第 2 に，基準辺長を単一の撮影環境から求めており，体格，撮影距離および視点の違いに対する頑健性は未評価である．人物ごとの尺度推定と複数環境での外部評価が必要となる．第 3 に，採用率だけでなく，動作分類や人数計測等の後段タスクの性能を直接測り，用途ごとに閾値を決定する必要がある．

実研究へ展開する場合は，撮影対象者の同意とデータ管理手順を明示し，照度，遮蔽，動きおよび人物属性ごとの失敗例を報告する．これらを確認した上で，単純な後処理として導入できる範囲と，再学習が必要な範囲を切り分けることが課題である．

生成 AI 等の利用に関する申告

[架空の記入例・提出不可]

本論文の作成にあたり，著者は OpenAI ChatGPT（GPT-5，2026 年 1 月 15 日から 2 月 20 日まで）を，研究課題の論点整理，Python 実装の補助，コードのリファクタリング，章構成の検討，文章表現の推敲および批判的確認に使用した．利用範囲は，第 4 章の前処理コード，第 5 章の集計コードおよび全章の文章表現である．生成 AI の出力は研究上の根拠としてそのまま採用せず，実験設計，ソースコード，モデル出力，数値結果，図表，主張，引用文献および最終稿を，著者が原資料，一次文献および実行ログと照合した．本架空例では，個人情報，認証情報および未公開研究データを外部 AI サービスへ入力していない．生成 AI を著者として扱わず，本論文の正確性，研究公正，独創性および最終内容については著者が責任を負う．

AI 利用の範囲と人間による検証責任を分けて記録する申告形式は，宇佐美ら [4] の開示例を参考にした．

謝辞

[架空の記入例・提出不可]

本研究を進めるにあたり，研究方針の検討から実験結果の解釈，論文の推敲まで終始ご指導いただいた指導教員に深く感謝申し上げます。また，実験条件と評価方法について有益なご意見をいただいた宇佐美研究室の皆様に感謝いたします。

参考文献

- [1] K. Sun, B. Xiao, D. Liu, and J. Wang, “Deep high-resolution representation learning for human pose estimation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5693–5703.
- [2] Y. Xu, J. Zhang, Q. Zhang, and D. Tao, “ViTPose: Simple vision transformer baselines for human pose estimation,” in *Advances in Neural Information Processing Systems*, vol. 35, 2022, pp. 38571–38584.
- [3] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “Realtime multi-person 2D pose estimation using part affinity fields,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7291–7299.
- [4] H. Usami, K. Hara, A. Tsuboi, and N. Matsuda, “LLM judges have dark current: A psychometric datasheet for LLM-as-a-judge evaluation,” *arXiv preprint arXiv:2606.15610v1*, 2026.

低照度環境下の 2 次元人物姿勢推定における構造的ー貫性を用い
た誤推定骨格の除去

EX00000 情報太郎

(中部大学工学部情報工学科)

20XX 年 2 月