

[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 $\hat{p}_j$ 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。

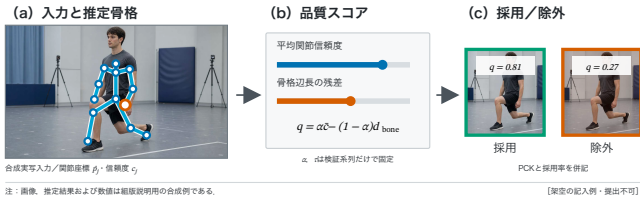


図 1: 誤推定骨格を除外する処理の流れ

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 $\tilde{e}$ 、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、

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

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

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]

- 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）。
- 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）。