

## 動作解析に向けた3次元人物姿勢推定手法の提案

大野 祐汰<sup>†</sup> Prima Oky Dicky Ardiansyah<sup>†</sup> 伊藤 久祥<sup>†</sup>岩手県立大学ソフトウェア情報学部<sup>†</sup>

## 1. はじめに

従来、非接触型センサを用いて人物の3次元姿勢を推定する研究は多くから行われており、リハビリテーションやスポーツ科学における個人の動作解析に利用されている。深度センサを搭載した Kinect v2<sup>1)</sup>が開発されており、複数人の3次元姿勢を実時間で推定できる。近年、画像センシングのみによる手法が提案されており、推定人数と推定可能な範囲の改善が行われている。

本研究では、ステレオカメラと OpenPose ライブラリ<sup>2)</sup>を組み合わせ、複数人の3次元姿勢を同時に推定する手法を提案し、その性能を検証するとともに、今後の発展として動作解析に向けた検討を行う。

## 2. 既存の人物3次元姿勢推定手法

## 2.1. Kinect v2

Kinect v2は深度センサを用いて、センサ前方にある物体までの距離を計測し、得られる深度情報を分析することで、人物の3次元姿勢を推定している。しかし、現行の仕様では、センサから4.5mまでしか人物の3次元姿勢を推定できず、同時推定できる人数は6人までに限られ、より広範囲かつ大人数の3次元姿勢を推定できる手法が求められている。

## 3. 関連研究

## 3.1. OpenPose ライブラリ

OpenPose ライブラリ (以下、OpenPose と呼ぶ) は畳み込みニューラルネットワークを用いて、人物画像から身体関節および顔の特徴点と、これらの特徴点の関連性を検出し、2次元姿勢を推定している。図1は、OpenPoseの実行例を示す。この例では、肩や肘など18点の特徴点を推定している。

## 4. 提案する3次元姿勢推定手法

本研究では、OpenPoseを利用して、左右の人物画像からそれぞれ推定した1対の2次元姿勢をもと

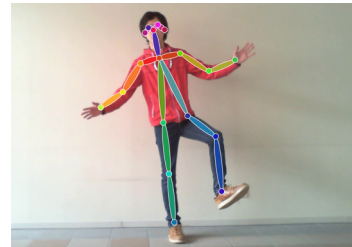


図1 OpenPoseの実行例

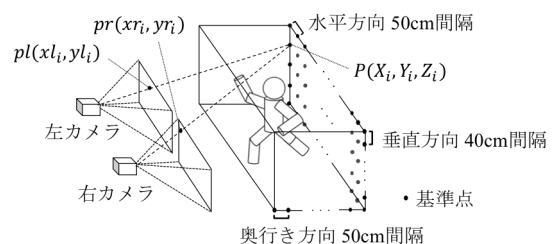


図2 基準点の配置の概略図

に、3次元姿勢を推定する。

## 4.1. 2次元姿勢推定の改善

人物画像に映る人物の大きさやその姿勢により、OpenPoseを利用して推定する2次元姿勢に誤差が生じる問題がある。提案手法では、片方の人物画像のみに対して、OpenPoseで2次元姿勢の推定を行い、得られる特徴点の2次元座標に対するもう片方の人物画像上での対応点を、テンプレートマッチングで求める。ただし、人物を撮影する2台のカメラに対して、事前にステレオキャリブレーションを行うものとする。

## 4.2. 3次元姿勢推定の改善

上記で推定した1対の2次元姿勢をもとに、三角測量の原理で3次元姿勢を推定する。このとき、カメラの解像度やレンズの歪みの影響で、正確な3次元姿勢が推定できないことが判明したため、本研究では、複数の基準点によるカメラの3次元キャリブレーションを行う。

図2は、基準点の配置の概略図を示す。ここで、基準点におけるワールド座標 $P_i(X_i, Y_i, Z_i)$ と、三角測量の原理で算出するワールド座標 $p_i(x_i, y_i, z_i)$ との間に

$$P_i = A \cdot p_i + \xi \quad (1)$$

が成り立つ。ここで、 $A$ は $m \times n$ の変換行列、 $\xi$ は残

3D human pose estimation for motion analysis  
Yuuta OHNO<sup>†</sup>, PRIMA Oky Dicky Ardiansyah<sup>†</sup>, Hisayoshi ITO<sup>†</sup>

<sup>†</sup>Faculty of Software and Information Science, Iwate Prefectural University

表1 奥行き位置ごとの基準点の平均誤差

距離 (cm)	点数	平均誤差 (cm)	距離 (cm)	点数	平均誤差 (cm)
1000	77	12.43	550	77	3.95
950	77	10.13	500	54	3.38
900	77	11.33	450	54	2.94
850	77	8.36	400	42	2.36
800	77	6.38	350	35	1.99
750	77	6.30	300	25	2.00
700	77	6.20	250	15	2.20
650	77	4.72	200	9	2.20
600	77	4.35			

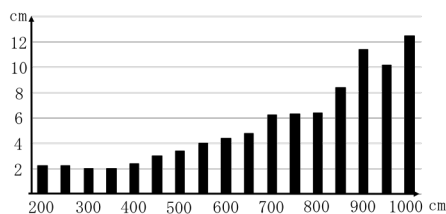


図3 奥行き位置ごとの基準点の平均誤差

差である。なお、 $A$ は、次式より、最小自乗法で求める。

$$\operatorname{argmin} \sum_{i=1}^N \|A \cdot p_i - P_i\|^2 \quad (2)$$

最後に、 $A$ を利用して人物の3次元姿勢を補正する。

### 5. 検証

実験により、提案手法の有用性を検証する。はじめに、解像度  $1280\text{px} \times 720\text{px}$  の2台のUSBウェブカメラを  $60\text{cm}$  の間隔で配置し、ステレオキャリブレーションを行う。次に、カメラ視野内に、水平方向  $50\text{cm}$ 、垂直方向  $40\text{cm}$ 、奥行き方向  $50\text{cm}$  の間隔で基準点を配置し、式 (2) より、変換行列を求める。最後に、複数人を同時に撮影し、その3次元姿勢の推定を行う。

表1および図3は、1004基準点の奥行き位置ごとの平均誤差を示す。カメラから距離が離れるほど、誤差が増大する傾向がみられ、基準点全体の平均誤差は  $6.30\text{cm}$  であった。

図4は、カメラから  $6\text{m}$  離れた地点にいる被験者7名を撮影し、3次元姿勢を推定した結果を示す。最も左に映る人物の右肩や顔の特徴点などにおいて、誤差が大きいが、全体的に実際の姿勢と同様な結果が得られることが分かった。また、提案手法を用いることで、より広範囲において人物と物体の動作を同時に分析できる。図5は、バレーボール選手の3次元姿勢推定の例を示す。

### 6. おわりに

本研究では、ステレオカメラと OpenPose を用いた複数人の3次元姿勢推定手法を提案した。実験を

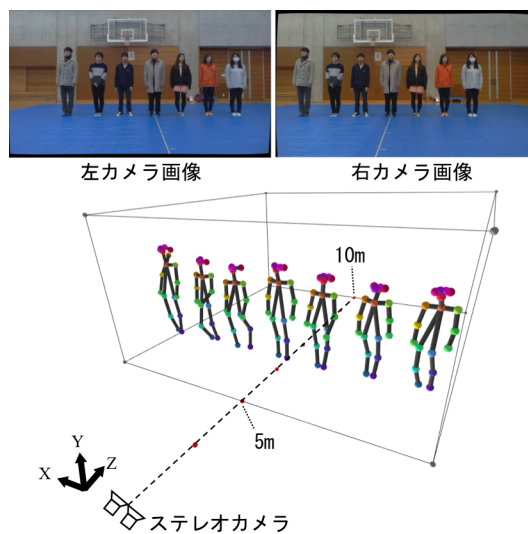


図4 被験者7名の3次元姿勢の推定結果

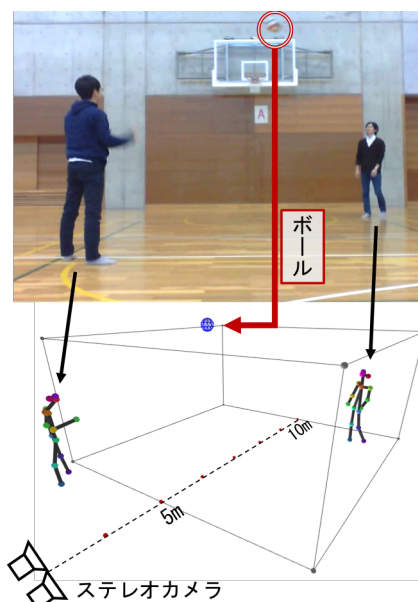


図5 バレーボール選手の3次元姿勢推定の例

通して、提案手法が Kinect v2 より広範囲かつ大人数を同時に推定可能であることが分かった。今後は、3次元姿勢の推定精度向上のため、カメラの解像度や画質の向上、テンプレートマッチングの精度の改善を行う。また、動作解析への応用に向けてさらなる検討を行う。

### 参考文献

- 1) Kinect for Windows, <https://developer.microsoft.com/ja-jp/windows/kinect> (2018/1/10 閲覧)
- 2) Cao, Z., Simon, T., Wei, S.E., Sheikh, Y., "Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields", CVPR, pp.7291-7299, (2017).