

## 1. はじめに

2次元画像を対象とした人体の3次元姿勢推定では、人体部位である2次元関節位置とその奥行きを推定する。Zhouらの提案した3次元姿勢推定手法[1]では、様々な環境下で撮影された画像からの3次元姿勢推定が可能である。しかし、オクルージョンが発生した部位の特徴を捉えることは難しいため、奥行き推定に失敗することがある。そこで本研究では、3次元ベクトル場推定モジュールを導入することで、関節間の各方向に対する特徴を抽出し、3次元姿勢推定の精度向上を目指す。

## 2. 従来手法

Zhouらの手法[1]が提案したネットワーク構造を図1に示す。入力画像から畳み込み層により求めた特徴を、2次元姿勢推定モジュールに入力する。2次元姿勢推定モジュールは、各関節の2次元ヒートマップを推定する。次に、2次元姿勢推定モジュールの出力および中間特徴をdepth回帰モジュールに入力し、depthを推定する。depthに対する損失関数には、人の骨の長さをもとにしたGeometric Lossを用いている。

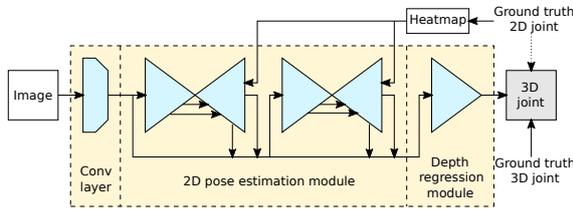


図1：文献[1]のネットワーク構造

## 3. 提案手法

提案手法は、3次元ベクトル場推定モジュールを導入することで、オクルージョンが発生した際の姿勢推定精度を向上する手法を提案する。

### 3.1. ネットワーク構造

提案手法のネットワーク構造を図2に示す。文献[1]のネットワーク構造に、3次元ベクトル場推定モジュールを追加する。3次元ベクトル場推定モジュールは、2次元姿勢推定モジュールの出力値とDepth回帰モジュールから獲得した中間特徴を入力とする。3次元ベクトル場推定モジュールは、Deconvolutionによる3次元拡張およびMax PoolingとUpsamplingを用い、3次元ベクトル場を獲得する。

### 3.2. 損失関数

2次元姿勢推定モジュールの損失関数には平均二乗誤差を使用し、depth回帰モジュールの損失関数には正解座標データが2次元座標データるときGeometric Loss、3次元座標データるとき平均二乗誤差を使用する。3次元ベクトル場推定モジュールの損失関数には、3次元ベクトル場の推定結果と正解とのL2ノルム二乗を用いる。3次元ベクトル場の正解には単位ベクトル $\mathbf{v} = (\mathbf{x}_{j_2} - \mathbf{x}_{j_1}) / \|\mathbf{x}_{j_2} - \mathbf{x}_{j_1}\|_2$ を用いる。ここで、 $\mathbf{x}_{j_1}, \mathbf{x}_{j_2}$ は関節 $j_1, j_2$ の3次元関節座標である。3次元ベクトル場は座標 $\mathbf{p}$ が $0 \leq \mathbf{v} \cdot (\mathbf{p} - \mathbf{x}_{j_1}) \leq l_c$ かつ $|\mathbf{v}_{\perp}^x \cdot (\mathbf{p} - \mathbf{x}_{j_1})| \leq \sigma_y$ 、かつ $|\mathbf{v}_{\perp}^z \cdot (\mathbf{p} - \mathbf{x}_{j_1})| \leq \sigma_z$ であることを生成する条件とする。このとき、 $\sigma_y, \sigma_z$ はそれぞれy, z方向の正解3次元ベクトル場の幅を示し、 $l_c = \|\mathbf{x}_{j_2} - \mathbf{x}_{j_1}\|$ は関節間の長さを示す。

### 3.3. 学習方法

学習は3ステージに分けて行う。図2に示すように、ステージ1では2次元姿勢推定モジュールを2次元座標データで学習する。ステージ2, 3では、全モジュールを2次元および3次元座標データで学習する。Depth回帰モジュールの損失関数として、ステージ2は平均二乗誤差のみ、ステージ3は平均二乗誤差とGeometric Lossの両方を使用する。

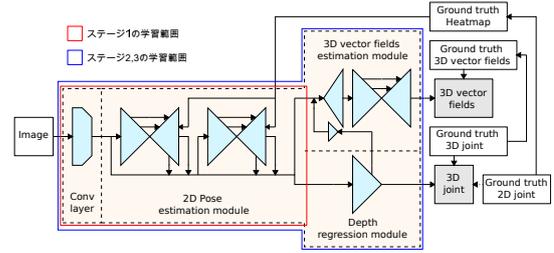


図2：提案手法のネットワーク構造

## 4. 評価実験

提案手法の有効性を示すため、評価実験を行う。

### 4.1. 実験概要

2次元座標データとしてMPIIデータセット、3次元座標データとしてHuman3.6Mデータセットを用いる。2次元姿勢推定の精度をPCKh@0.5、3次元姿勢推定の精度をMPJPEで評価する。PCKh@0.5は正解関節位置と推定関節位置の誤差が頭部サイズを基準とした範囲の50%以内に含まれている割合、MPJPEは3次元空間における推定関節位置と正解関節位置の誤差の平均値である。

### 4.2. 実験結果

各ステージにおける2次元姿勢推定及び3次元姿勢推定の精度を表1に示す。提案手法のPCKh@0.5は、従来手法と比べ同程度の精度であるが、MPJPEはステージ2以降で誤差が減少している。

表1：各ステージにおける精度

		ステージ1	ステージ2	ステージ3
従来手法	PCKh@0.5	84.2	91.7	92.0
	MPJPE	-	63.9	63.2
提案手法	PCKh@0.5	84.2	91.7	91.9
	MPJPE	-	63.0	61.8

従来手法と提案手法による3次元姿勢、3次元ベクトル場の各方向の成分の推定の例を図3に示す。3次元ベクトル場は各方向の成分が強いほど値が高くなり、色が赤くなる。図3(a)の3次元ベクトル場の各方向の成分の推定例では、胴体から脚部にかけてのY方向成分、腕部のZ方向成分で値が高くなっている。図3(b)に示す従来手法では、胴体によって隠れている左腕部の推定に失敗している。一方、提案手法は正解3次元関節位置との差が従来手法に比べ少ない。以上より、3次元ベクトル場を導入することで、オクルージョンが生じた部位の3次元姿勢推定の精度向上に寄与しているといえる。

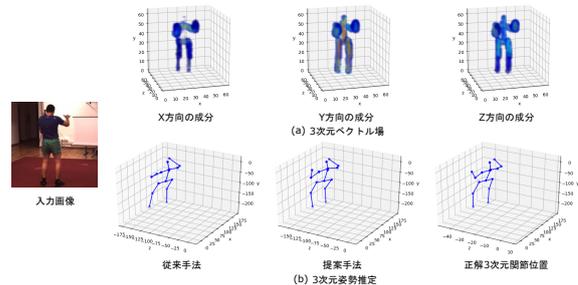


図3：3次元姿勢推定の例

## 5. おわりに

本研究では、3次元ベクトル場を導入したCNNによる3次元姿勢推定手法を提案し、3次元姿勢推定の精度向上を実現した。今後の課題として、3次元ベクトル場推定モジュールの見直しによる精度向上が挙げられる。

### 参考文献

- [1] X. Zhou, et al., "Towards 3D Human Pose Estimation in the Wild: A Weakly-Supervised Approach", ICCV, 2017.