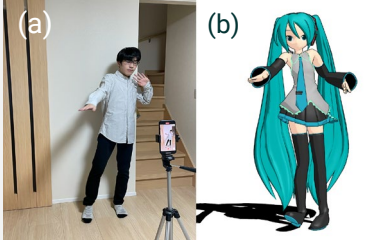


# スマートフォンカメラを用いた複数人物の3D姿勢推定とその応用

創価高等学校 2年 東 翔生

## 1. 概要

スマートフォンカメラやWebカメラを利用してコンピュータにリアルタイムで画像を取り込む。その画像から複数の人物姿勢の推定及び描画を行った。



▲ 図1 推定の様子 (a) スマートフォンの映像をコンピュータに取り込み、(b) フリーソフトMikuMikuMoving(MMM)[1,2]のプラグイン機能を用いて描画している。

## 2. 背景

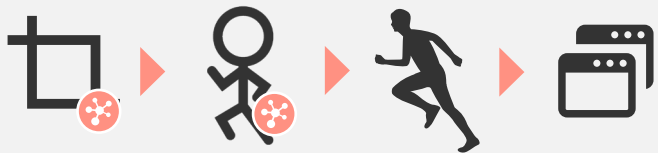
VR(Virtual Reality) / AR(Augmented Reality)技術が急速に発展する中、人物の動作を3Dでコンピュータに読み込むトラッキング技術の需要が高まっている。このトラッキングを実現するにはマーカーを用いたトラッカーやKinectなどの機材が必要である。しかしながら、これらの手法には

- ◆ 導入のコスト
- ◆ 装着の負担

などの問題がある。本研究では単眼のRGB画像から複数人物の姿勢を推定することで、これらの問題を解決し、よりトラッキング技術を身近にすることを目的とする。

## 3. 実装方法

実装は主に 画像切り出し、推論、変換、描画 に分かれる。



### 切り出し

- ・ IVCamから映像を取得
- ・ Faster R-CNN[3]で切り出し

### 推論

- ・ Moonら[4]の手法をベースに推論
- ・ 各人物についてRoot-jointの深度と各関節の相対位置を推定

### 変換

- ・ 推定した関節から各肢の3次元ベクトルを計算
- ・ ベクトルをオイラー角に変換

### 描画

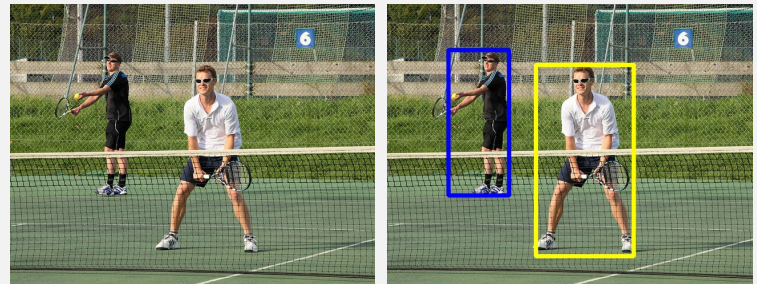
- ・ オイラー角をクォータニオンに変換してMMMで描画

▲ 図2 実装のパイプライン スマートフォン等から取得した画像は図の手順に沿って切り出し、推論、変換、描画が行われる。

## 謝辞

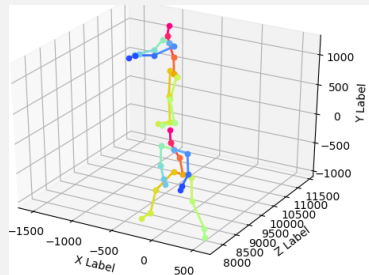
本研究を行うにあたり、方針に関してご指導を賜った電気通信大学プログラミング教室担当講師の呉先生に感謝いたします。また、同じく講師の岡崎先生には画像処理及び深層学習についてご指導、武田先生には研究にご助言賜りまして感謝いたします。最後に、本研究を進めるにあたり精神的・経済的な支援をくれた家族、多くの激励をくださった友人、高校の先生に御礼申し上げます。

## 4. 結果と考察

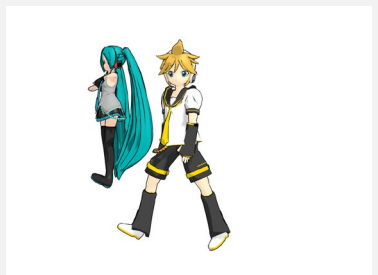


(a) 処理に用いる画像[5]

(b) 人物の切り出し結果



(c) 3D姿勢推定結果



(d) ボーンに変換後、描画

▲ 図3 過程と結果 (a) RGB画像1枚から (b) すべての人物を切り出す。(c) その後、それぞれの人物について3Dの姿勢推定を実行。(d) 推定した関節をボーンに変換しMMMで描画する。

## 成果

- ・ 1枚の画像から複数人の姿勢推定ができた
- ・ 推定した姿勢をボーンとして描画できた

## 課題点

- ・ 関節位置の単純なプロットだけでは姿勢の再現精度に限界がある
- ・ 時間的な整合性が保たれていない

## 5. まとめと展望

### まとめ

精度に向上の余地はあるが、特別な機材を使わずにトラッキングを行うことができ、目的を達成した。

### 今後の展望

- ・ 可動域の制限による、表現力と解剖学的整合性の向上を目指す
- ・ 時間軸情報を考慮して高速・高精度化を図る

## 参考文献

- [1] あにまさ「初音ミクVer2.pmx, 鏡音レン.pmx」Vocaloid Promotion Video Project. <https://sites.google.com/view/vpvp/>. (最終閲覧 2022/2/5)
- [2] mogg「MikuMikuMoving」<https://sites.google.com/site/mikumikumoving/>. (最終閲覧 2022/2/5)
- [3] Shaoqing Renら「Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks」NeurIPS. 2015.
- [4] Gyeongsik Moonら「Camera Distance-aware Top-down Approach for 3D Multi-person Pose Estimation from a Single RGB Image」ICCV. 2019.
- [5] Tsung-Yi Linら「Microsoft COCO: Common Objects in Context」ECCV. 2014.