

佐藤 洋一

東京大学生産技術研究所  
教授

## 集合視による注視・行動解析に基づくライフイノベーション創出

### § 1. 研究成果の概要

本プロジェクトでは、複数ウェアラブルデバイスを分散型センサとして用いる集合視により、さまざまな空間・時間のスケールで起こるグループの注視・行動を計測し理解する技術の開発を目指している。2018 年度は主に以下の研究課題に取り組んだ。

#### 1. 一人称視点映像における人物位置予測

一人称視点映像中の人物の将来位置を予測する新たな問題に取り組む。具体的には、胸部に装着したウェアラブルカメラ映像中に映る人物の短いクリップが与えられた時に、当該人物が近い将来撮影者の視野のどの地点に映るかを予測する問題を考える。本研究では、一人称視点映像特有の手掛かりを生かしながら対象人物の位置履歴、姿勢及び映像中の自己運動の 3 つの手がかりの時系列を入力とする畳み込みニューラルネットワークを用いた予測手法を提案し(図 1)、独自に収集した歩行映像データセットにおいてその有効性を実証した。

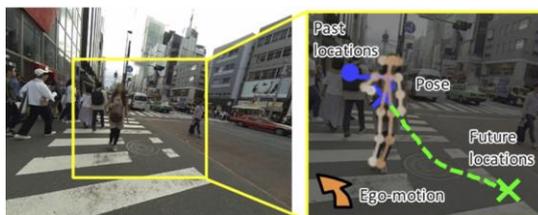


図1:ウェアラブルカメラ映像からの歩行者の人物位置予測

#### 2. 一人称視点映像からの撮影者の三次元姿勢推定技術

ウェアラブルカメラを装着した撮影者の三次元姿勢を、一人称視点映像のみから推定する技術の開発に取り組む。本研究では、与えられた映像から姿勢を推定するために、Imitation Learning

の機械学習フレームワークを用いた。本手法では、モーションキャプチャーにより取得した姿勢情報を、物理シミュレーション内において再現するポリシー群の作成を行い。ポリシー群が作成した実演履歴を利用して、映像からの姿勢推定を可能にした。さらに本研究では、シミュレーション内で学習したポリシーを実世界に適用するためのドメイン適用手法も提案している。

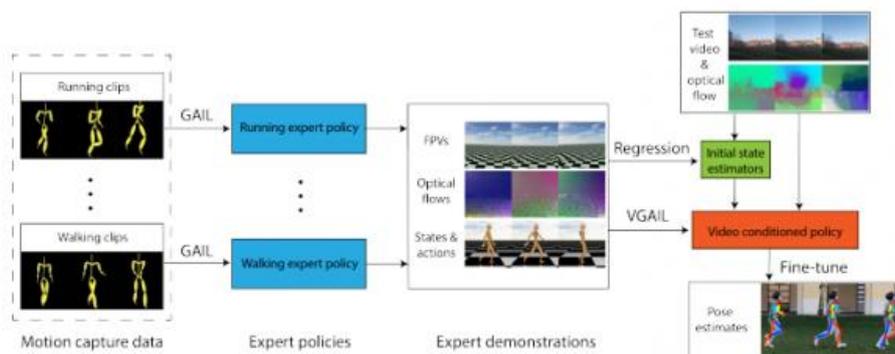


図2: 提案手法の概要図

### 3. アピランスの学習に基づく視線方向検出技術と学習用データセット構築

映像に映る人物の見え情報からその人物の視線を推定するアピランスベース視線推定を大きく発展させるため、過去の研究で作成した様々な環境で人物の視線を記録した20万枚を超えるデータセットに対し、手動の補正も併用しながら顔特徴点の正確なアノテーションを拡充したほか、ネットワーク構造の違いや視線範囲、照明環境、解像度などの違いによる影響を詳細に分析した。学習に基づく視線推定アプローチの学術的な基盤として、今後の深層学習に基づく視線推定の研究や、人物行動解析などの応用にも大きく寄与するものである。

#### 【代表的な原著論文】

1. Takuma Yagi, Karttikeya Mangalam, Ryo Yonetani, and Yoichi Sato, “Future Person Localization in First-Person Videos”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018), pp. 7593–7602, 2018.6
2. Ye Yuan, Kris Kitani. “3D Ego-Pose Estimation via Imitation Learning”, European Conference on Computer Vision (ECCV) 2018.9
3. Xucong Zhang, Yusuke Sugano, Mario Fritz, Andreas Bulling, “MPIIGaze: Real-World Dataset and Deep Appearance-Based Gaze Estimation”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 41, Issue 1, pp. 162–175, 2018.

## § 2. 研究実施体制

### (1) 東大グループ

- ① 研究代表者: 佐藤 洋一 (東京大学生産技術研究所 教授)
- ② 研究項目
  - ・ 複数ウェアラブルカメラ映像の統合によるグループの注意・行動解析

### (2) 慶大グループ

- ① 主たる共同研究者: 杉本 麻樹 (慶應義塾大学理工学部 准教授)
- ② 研究項目
  - ・ 協働支援のための時空間を考慮した視線情報の解析と可視化

### (3) MPI グループ

- ① 主たる共同研究者: Andreas Bulling (Max Planck Institute for Informatics, Perceptual User Interfaces Group, Independent Research Group Leader)
- ② 研究項目
  - ・ 大規模集合視センシングのための注視推定・映像解析

### (4) CMU グループ

- ① 主たる共同研究者: Kris M. Kitani (Carnegie Mellon University, The Robotics Institute, Assistant Research Professor)
- ② 研究項目
  - ・ 集合視による人物行動理解と支援

### (5) 阪大グループ

- ① 主たる共同研究者: 菅野 裕介 (大阪大学大学院情報科学研究科 准教授)
- ② 研究項目
  - ・ 大規模集合視センシングのための注視推定・映像解析