

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視・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, System Scientist)
- ② 研究項目
 - ・集合視による人物行動理解と支援

§ 2. 研究実施の概要

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。2015年度は主に以下の研究課題に取り組んだ。

1. 深層学習に基づく詳細な手動作の認識¹⁾

料理など詳細な手動作を必要とする活動を記録した一人称視点映像が与えられた際に、手動作の種類と把持している物体の種類の両方を認識可能な深層学習のアーキテクチャを新たに開発した。提案アーキテクチャは映像の見えを学習する **appearance stream** と映像中の動きを学習する **motion stream** から構成されており、特に前者では手領域の切り出しや物体位置の検出といった、ウェアラブルカメラ映像解析において重要な技術が組み込まれている。実験の結果、物体認識、動作認識の両タスクにおいて既存手法を大幅に上回る性能を確認した(図 1)。

2. アイウェア型デバイスを用いたユーザの内部状態推定²⁾

ユーザがどのような内部状態で視線を注視対象に向けているのかを把握することができれば、集合視の枠組みにおいて幾何学的な視線の集中・分散とユーザの状態を加味した行動解析が可能となる。本研究では、アイウェア型デバイスに装着した光センサ群から得られる視線の周辺情報を機械学習することにより、ユーザの表情識別や、日常行動中における内部状態の推定が可能であることを示した(図 2)。

3. インタラクションにおける動作・反応の認識³⁾

インタラクションを行う二者それぞれが装着したウェアラブルカメラを用いて、「指差し」や「頷き」、「物の受け渡し」といった動作・反応を認識する手法を提案した。本研究では、「頷き」などの頭部運動を含む動作・反応は自身が装着したウェアラブルカメラから得られる「一人称視点」映像によく現れる一方で、「指差し」のような手動作の認識には、対話相手のカメラに映る自身の「二人称視点」情報が有効であることを示した。また、1000以上の映像からなる大規模なインタラクションデータセットを公開した(図 3)。

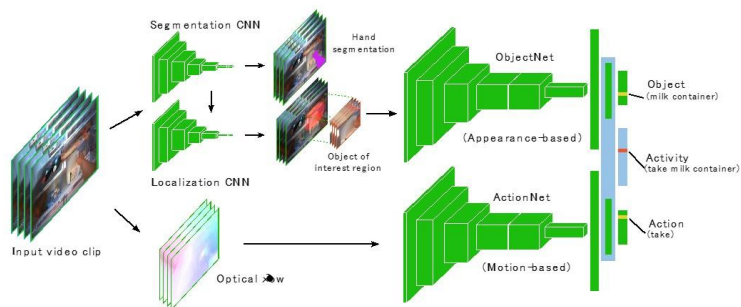


図1: 深層学習に基づく詳細な手動作の認識

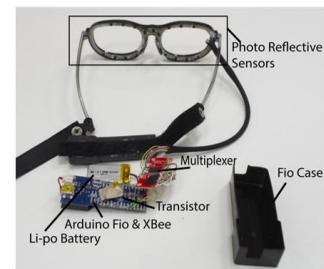


図2: アイウェア型デバイスを用いた内部状態推定



図3: インタラクションにおける動作・反応の認識

- 1) Minghuang Ma, Kris M. Kitani, “Going Deeper into First-Person Activity Recognition”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- 2) Katsutoshi Masai, Yuta Sugiura, Masa Ogata, Kai Kunze, Masahiko Inami, Maki Sugimoto, “Facial Expression Recognition in Daily Life by Embedded Photo Reflective Sensors on Smart Eyewear”, International Conference on Intelligent User Interfaces (IUI), 2016
- 3) Ryo Yonetani, Kris M. Kitani, Yoichi Sato, “Recognizing Micro-Actions and Reactions from Paired Egocentric Videos”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.