

戦略的創造研究推進事業 CREST

研究領域

「人間と調和した創造的協働を実現する
知的情報処理システムの構築」

研究課題

「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」

研究終了報告書

研究期間 2014年10月～2020年3月

研究代表者：佐藤 洋一
(東京大学生産技術研究所 教授)

§ 1. 研究実施の概要

(1) 実施概要

少子高齢化が進む我が国において、医療・福祉などの諸分野における活動を維持・発展させていくために、人間と情報システムの創造的協働による支援技術への期待が高まっている。高齢化とともに医療サービスへの需要が急増する中、如何にして限られた医療従事者により質の高い医療サービスを維持するかが喫緊の課題であり、また、障害者を含む高齢者等の社会的弱者が増えていく中、これらの人々の社会生活への参加を支えるバリアフリー化の実現も重要な課題となっている。人と情報システムとの創造的協働により、これらの課題の解決を図るためには、グループとして活動する人々のそれぞれが何に注意を向け、他者とのインタラクションの中でどのように行動しているのかを詳細に把握するための技術が欠かせない。このような背景の下、本研究課題では、複数の人物が利用するウェアラブルカメラや装着型アイトラッカーなどのデバイスを分散型センサとして活用することにより、グループとして活動する人々の注視・行動を計測し理解する技術(集合視技術と呼ぶ)の開発に取り組んだ。さらに、開発された集合視のコア技術を応用し、手術室での協働支援、街中での安全ナビゲーション支援という2種類のシナリオに対して、プロトタイプシステムを構築し、さまざまな応用における集合視技術の有効性を示すことを目標とした。

研究体制としては、この目標の達成に向けて、コンピュータビジョン、HCI、拡張現実感の各分野で実績を有する東京大学(研究代表者:佐藤洋一)、慶應義塾大学(杉本麻樹)、CMU(Kris M. Kitani)、Max Plank Institute for Informatics(Andreas Bulling)、大阪大学(最終年度から東京大学、菅野裕介)の海外機関を含む5つの研究グループにより研究チームを組織し、参画する若手研究者の人材交流も含め、グループ間で密に連携を取りながら以下の各研究項目の実施に当たった(カッコ内は担当グループ)。

研究項目 I. 集合視による計測技術の開発

- ・長時間注視計測のためのウェアラブルデバイスの開発 (MPI、阪大(東大2)、慶大)
- ・一人称視点映像における二人称注視推定 (MPI、阪大(東大2))
- ・空間座標にマッピングした注視情報解析のための位置・姿勢推定 (慶大)
- ・一人称視点映像における人物対応付け (東大)

研究項目 II. グループの注視・行動解析

- ・一人称・二人称視点映像からの行動認識 (東大、MPI、CMU、阪大(東大2))
- ・グループの行動における動作・反応ペアの検出 (東大)
- ・グループの行動予測 (CMU、東大)

研究項目 III. 集合視が実現するライフインベーション

- ・手術室での協働支援のための集合視の解析・可視化 (慶大、東大)
- ・街機能情報の共有による安全ナビゲーション支援 (CMU、東大)

その結果、集合視コア技術の開発および集合視を活用した支援システムの構築の両方に関して十分な成果を挙げ、当初の研究開発目標を達成することができた。その中でも特に、一人称視点映像からの逆強化学習によるオンライン行動予測、自己運動特徴を手掛かりとした一人称視点映像における人物対応付け、二人称視点からのアピアランスベース視線推定、タスクに依存した視線移動を考慮した一人称視点映像からの注視推定、長時間の一人称視点映像の閲覧支援について、顕著な成果を得ることができた。これらの成果の多くは当該分野におけるトップ国際会議やジャーナルに採択され、コンピュータビジョン分野のトップ国際会議である ICCV や、HCI 分野のトップ国際会議である CHI と UIST での受賞を含め合計 14 件の賞を得ている。また、本プロジェクトからの情報発信と集合視技術に関するコミュニティ形成のために、コンピュータビジョン、拡張現実感、HCI の各分野における主要国際会議に併催する形で 9 件のワークショップやチュートリアルを実施するなど、学会でのアウトリーチ活動にも積極的に取り組んだ。

(2) 顕著な成果

<優れた基礎研究としての成果>

1. 一人称視点映像からのオンライン行動予測

概要:

ウェアラブルカメラを装着した人物が記録した一人称視点映像から、「どこへ行き、何をするか」といった装着者の行動を予測する技術を開発した。本技術は 1) 時空間的に遠く離れた場所・時刻での行動を予測可能、2) 移動先だけでなく、そこでの行動の内容(例えば、特定のものを手に取る)も予測可能、という点で既存の行動予測に関する技術とは大きく異なる。本研究成果は国際的に非常に高く評価され、コンピュータビジョン分野における最難関国際会議 ICCV にオーラル発表論文として採択されるとともに、Best Paper Honorable Mention を受賞した。

2. 一人称視点映像における人物対応付け

概要:

一人称視点映像中で特徴的に観察される自己運動特徴を手掛かりに、複数のウェアラブルカメラ映像が与えられた際に、どのカメラに誰が映っているかを特定する人物対応付け技術を他に先駆けて開発した。本研究成果は、提案手法におけるアプローチの新規性と、既存の顔検出・認識では上手く扱うことが難しいような場合においても人物対応付けが可能であるという有効性が高く評価され、コンピュータビジョン分野における最難関国際会議 CVPR に採択された。

3. 一人称視点映像からの注視推定

概要:

一人称視点映像から撮影者の注視位置を推定する問題について、撮影者が行っているタスクに特有の視線移動パターンを考慮した推定手法を考案し、ベンチマークデータセットで従来手法を大幅に超える推定精度を実現した。心理学分野で広く知られていた視線の動きは実行しているタスクに依存するというトップダウン的な側面を、世界で初めて視線推定手法に取り入れたものとして、その新規性・有効性が高く評価され、コンピュータビジョン分野の最難関国際会議 ECCV においてオーラル発表論文(採択率 2.4%)として採択された。

<科学技術イノベーションに大きく寄与する成果>

1. 一人称視点映像閲覧のための EgoScanning

概要:

長時間にわたって撮影された一人称視点映像を効率的に閲覧するためのインタフェース: EgoScanning を開発した。提案インタフェースは、閲覧者の関心・目的に応じてビデオの再生速度が適応的に制御される点を特長としており、閲覧者はこのインタフェースを用いることで所望の箇所を見落とすことなく、高速に一人称視点映像を閲覧することが可能になる。本研究成果はアプローチの新規性と有効性が高く評価され、HCI 分野の最難関国際会議 CHI 2017 に採択された。さらに、Egoscanning を実際の外科手術の現場での支援に応用したビデオ閲覧支援システムを構築し、その有効性を確認した (ACM IUI 2019)。当該技術は実用化を見据えて特許出願済み(特願 2016-241446、PCT/JP2017/043080)であり、映像閲覧を伴う各種応用システム開発への活用が期待される。

2. アピアランスベース視線推定

概要:

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アランスベース視線推定を大きく発展させ、実際の応用環境に近い条件で有効な技術を開発した。具体的には、多様な環境下で収集された大規模な視線推定用データセットを構築し、そのデータセットを用いた深層学習によって高い視線推定性能を実現した (IEEE CVPR 2015 で採択。拡張版が IEEE TPAMI 2018 に採択)。また、さまざまな条件下でのアピアランスベース視線推定の利用について詳細な評価実験を行った結果を HCI 分野のトップ会議 CHI 2019 で報告するとともに、アピアランスベース視線推定のライブラリを OpenGaze Toolkit として広く公開した。

<代表的な論文>

1. Nicholas Rhinehart and Kris M. Kitani, “First-Person Forecasting with Online Inverse Reinforcement Learning,” in Proc. IEEE International Conference on Computer Vision (ICCV 2017), pp. 3716–3725, October 2017.
2. Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling, “Appearance-based Gaze Estimation in the Wild,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 4511–4520, June 2015.
3. Keita Higuchi, Ryo Yonetani, and Yoichi Sato, “Egoscanning: quickly scanning first-person videos with egocentric elastic timelines,” Conference on Human Factors in Computing Systems (CHI), pp. 6536–6546, May 2017.

§ 2. 研究実施体制

(1) 研究チームの体制について

①「東大」グループ

研究代表者： 佐藤 洋一(東京大学生産技術研究所 教授)

研究項目：

- 一人称視点映像における人物対応付け
- 一人称・二人称視点映像からの行動認識
- 手術室での協働支援のための集合視の解析・可視化
- 街機能情報の共有による安全ナビゲーション支援

②「慶大」グループ

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)

研究項目：

- 長時間注視計測のためのウェアラブルデバイス設計
- 空間座標にマッピングした注視情報解析のための位置・姿勢推定
- 手術室での協働支援のための集合視の解析・可視化

③「MPI」グループ

主たる共同研究者： Andreas Bulling (Max Planck Institute for Informatics, Perceptual User Interfaces Group, Independent Research Group Leader)

研究項目：

- 長時間注視計測のためのウェアラブルデバイス設計
- 一人称視点映像における二人称注視推定
- 一人称・二人称視点映像からの行動認識

④「CMU」グループ

主たる共同研究者： Kris M. Kitani (Carnegie Mellon University, The Robotics Institute, Assistant Research Professor)

研究項目：

- 一人称・二人称視点映像からの行動認識
- グループの行動予測
- 街機能情報の共有による安全ナビゲーション支援

⑤「阪大(東大2)」グループ(最終年度から東京大学)

主たる共同研究者： 菅野 裕介(東京大学生産技術研究所 准教授)

研究項目：

- 長時間注視計測のためのウェアラブルデバイス設計
- 一人称視点映像における二人称注視推定
- 一人称・二人称視点映像からの行動認識

(2) 国内外の研究者や産業界等との連携によるネットワーク形成の状況について

ウェアラブルカメラを含む多様なカメラを用いてグループ行動を解析するという異なる観点のもと、JST SICORP(研究領域「レジリエントな社会のためのICT」)において東大グループがヘブライ大学との共同研究を開始した。また、視覚障害者のためのナビゲーションシステムに関して、CMU グループは日本 IBM の研究者と緊密に連携しながら各要素技術の開発とアプリケーションシステムの構築を進めている。

更に手術室での協働支援のための集合視の解析・可視化に関して、CREST チーム間連携として無影灯カメラと一人称視点カメラを連携させることによる手術映像記録の研究に発展しており、慶應義塾大学病院形成外科との連携による多視点の映像の記録・利活用の研究に繋がっている。また、機械学習を援用した手術映像の記録・解析は、慶應義塾大学病院 SIP AI ホスピタル事業の一部の研究項目ともなっている。

これらの取り組みに加え、本プロジェクトからの情報発信と集合視技術に関する研究者コミュニティの形成のために、コンピュータビジョン、AR/MR、HCI の各分野における主要国際会議に併設の国際ワークショップと独立のセミナーを開催してきている。また、本 CREST における研究の重要なコア技術である、行動認識技術と逆強化学習技術に関するチュートリアル講演を国際会議にて行っている。具体的な会議とセミナー・チュートリアル名は以下の通り。

- Workshop on EyeWear Computing (UbiComp 2016 併設)
- Workshop on Egocentric Perception, Interaction, and Computing (ECCV 2016 併設)
- IEEE 4th Workshop on Egocentric Vision (CVPR 2016 併設)
- Workshop on Human Behavior Analysis and Visualization for Collective Visual Sensing (ISMAR 2016 併設)
- EyeWear Computing - Augmenting the Human with Head-mounted Wearable Assistants (Dagstuhl Seminar, January 2016)
- International Workshop on Human Activity Analysis with Highly-Diverse Cameras (WACV 2017 併設、山口高平チームと共同で開催)
- The 2nd International Workshop on Egocentric Perception, Interaction and Computing (ICCV2017 併設)
- Workshop on Anticipating Human Behavior (ECCV2018 併設)
- Tutorial on Human Activity Recognition (ECCV2018 併設)
- Tutorial on Inverse Reinforcement Learning for Computer Vision (ECCV2018 併設)