

戦略的国際共同研究プログラム(SICORP)

日本－イスラエル共同研究

終了報告書 概要

1. 研究課題名：「多様なカメラを活用した群衆行動の変化検出」
2. 研究期間：2017年1月～2020年3月
3. 主な参加研究者名：
日本側チーム

	氏名	役職	所属	研究分担
研究代表者	佐藤 洋一	教授	東京大学生産技術研究所	研究統括および基盤技術開発
研究参加者	樋口 啓太	特任講師	東京大学生産技術研究所	基盤技術開発
研究参加者	米谷 竜	助教	東京大学生産技術研究所	基盤技術開発
研究期間中の全参加研究者数			6名	

相手側チーム

	氏名	役職	所属	研究分担
研究代表者	Michael Werman	Professor	School of Computer Science and Engineering, The Hebrew University of Jerusalem	研究統括および基盤技術開発
主たる共同研究者	Shmuel Peleg	Professor	School of Computer Science and Engineering, The Hebrew University of Jerusalem	基盤技術開発
研究期間中の全参加研究者数			8名	

4. 国際共同研究の概要

都市部における犯罪やテロ、事故、自然災害など、我々の日常生活を脅かす様々な事案の防止および解決のために、街中に広く設置されている防犯カメラや、急速に普及が進む車載カメラ、さらに今後普及が見込まれるウェアラブルカメラなどから得られる映像データの利活用への期待が高まっている。このような背景の下、本研究では、防犯カメラやウェアラブルカメラといった多様なカメラから得られる映像データからの人々の行動解析に資する基盤技術の開発に取り組んだ。具体的には、複数カメラからと IoT センサの統合による人物追跡技術、プライバシーを保護した分散映像からのモデル学習技術による重要箇所発見、Human-in-the-loop の考え方に基づく映像閲覧支援として、高速映像再生時の映像ぶれ抑制技術、作業空間とカメラ位置の 3 次元可視化による複数自己視点映像閲覧支援技術、音と映像のマルチモーダル解析による対象人物音声強調技術などについて成果を得た。

5. 国際共同研究の成果

5-1 国際共同研究の学術成果および実施内容

複数映像にわたる人物追跡:

グループ行動解析の基盤技術として、複数カメラ映像中での特定人物追跡のために、映像中の人物の見えに加えて、スマートフォンなどに搭載される加速度センサから取得される

動き情報を用いることによりカメラの観察方向や人物の見えの変動に頑健な人物追跡手法を開発した。また、映像中に映る人物の動きを手掛かりとして、カメラの位置・向きが大きく異なる条件下においても、高精度に多視点カメラ間自動幾何校正を実現する手法を開発した。

長時間映像を用いた日常生活のルーチン解析:

街中で記録された映像から人の行動を解析するにあたっては、映像に映りこむ人物のプライバシーを適切に保護する必要がある。そこで、本共同研究では、プライバシー保護分野で良く知られている差分プライバシーと呼ばれる考えに基づき、doubly-permuted homomorphic encryption (DPHE) と呼ぶ疎な高次元データからのモデル学習に適した手法を開発し、プライバシー情報を保護したまま、複数映像中に共通して映り込む対象の自動発見を実現した。

“Human-in-the-loop” に基づく人物・イベント解析:

人が長時間映像を閲覧し内容を理解する際の支援という観点から、映像閲覧支援技術の研究開発に取り組んだ。ウェアラブルカメラにより得られる自己視点映像では視点位置が常に変化するため、高速再生時の極端な映像ぶれや、複数自己視点映像間での相対位置関係の把握の困難さが問題となる。これに対し、映像のフレームの適切なサンプリングによる高速再生時の映像ぶれ抑制手法、ならびに、環境と複数カメラの 3 次元位置関係可視化による映像閲覧支援手法を開発した。

5-2 国際共同研究による相乗効果

多様なカメラ映像データから人々の行動モデルの学習と変化検出を行うための基盤技術の構築に向け、ウェアラブルカメラ映像を用いた人物行動センシングに研究実績を有する東大チームと、サーベイランス映像解析で研究実績を有するヘブライ大学チームが協力して研究開発に取り組んだ。これにより、目標達成のための重要課題を適切に設定し、双方の強みを生かした形で各研究項目に取り組むことができた。

5-3 国際共同研究成果の波及効果と今後の展望

本共同研究を通じ、日本側の若手研究者が、コンピュータビジョン分野でトップレベル実績を有するイスラエル側の研究者と人的ネットワークを構築することが出来たことは、若手研究者の人材育成の観点から大きな成果と考えている。多様なカメラ映像からの人物行動センシングは、さまざまな応用の場面において今後更に重要性を増すことが予想されるため、今回構築することができた人的ネットワークは日・イスラエル双方の研究チームの研究に大きく寄与することが期待される。

Strategic International Collaborative Research Program (SICORP)
Japan—Israel Joint Research Program
Executive Summary of Final Report

1. Project title : 「Learning and Detecting Changes in Population Behavior From Video」
2. Research period : January 2017 ~ March 2020
3. Main participants :
Japan-side

	Name	Title	Affiliation	Role in the research project
PI	Yoichi Sato	Professor	Institute of Industrial Science, The University of Tokyo	Project management and R&D of each WP
Collaborator	Keita Higuchi	Project Lecturer	Institute of Industrial Science, The University of Tokyo	R&D of each WP
Collaborator	Ryo Yonetani	Research Associate	Institute of Industrial Science, The University of Tokyo	R&D of each WP
Total number of participants throughout the research period:				6

Partner-side

	Name	Title	Affiliation	Role in the research project
PI	Michael Werman	Professor	School of Computer Science and Engineering, The Hebrew University of Jerusalem	Project management and R&D of each WP
Co-PI	Shmuel Peleg	Professor	School of Computer Science and Engineering, The Hebrew University of Jerusalem	R&D of each WP
Total number of participants throughout the research period:				8

4. Summary of the international joint research

In a wide variety of applications, there are high expectations for the utilization of video data obtained from various types of cameras such as surveillance cameras, in-vehicle cameras, and wearable cameras. In this research, we have developed fundamental technologies for analyzing human behaviors from multiple videos captured by various cameras. In particular, we have developed 1) a person tracking method by multiple cameras and body-worn IMU sensors based on motion signatures, 2) a privacy-preserving framework for learning visual models by leveraging distributed private image data, 3) several methods for efficient video browsing such as oscillation suppression of fast-forwarded videos, and browsing group first-person videos with 3D visualization.

5. Outcomes of the international joint research

5-1 Scientific outputs and implemented activities of the joint research

Tracking people across multiple cameras: Tracking a target person across different cameras

is one of the core elements for analyzing human behaviors from various types of videos. In this work, we developed a method for tracking a target person across multiple videos based on the similarity of motion signatures from the videos and the IMU sensors carried by the person. The method is shown to be robust against the variation of the person's appearance and camera viewing directions. Also, we developed a multi-view camera calibration method that can handle widely different camera positions and orientations by using people trajectories.

Discovering routines of our daily life: Privacy preservation is an essential issue for human behavior analysis from various types of videos recorded in public spaces. To this end, we developed a privacy-preserving method for visual recognition by leveraging distributed private image data. Our method utilizes a homomorphic cryptosystem with newly introduced efficient encryption scheme called doubly-permuted homomorphic encryption (DPHE), which is optimized for sparse high dimensional data. Our experimental evaluation on several public datasets shows that the proposed approach achieves comparable performance against state-of-the-art visual recognition methods while preserving privacy and significantly outperforms other privacy-preserving methods.

Human-in-the-loop video event analysis:

As much as we would like to have fully automatic analysis, it is inevitable to have people in the loop for decision making. The display of information to people has two challenges. (i) videos are usually very long, (ii) videos recording the same events are often taken from very different viewpoints, making it difficult for a person to comprehend the events by observing the videos. To overcome the first problem, we developed an adaptive frame sampling method to suppress the camera shake in a fast-forwarded egocentric video captured by a wearable camera. For the second problem, we proposed a novel user interface that effectively use the 3D visualization of camera poses in the environment to assist users to understand complex group activities from multiple egocentric videos.

5-2 Synergistic effects of the joint research

Toward the development of fundamental technologies for understanding population behaviors from multiple videos, the University of Tokyo team and the Hebrew University of Jerusalem team worked closely by utilizing their strengths in human behavior analysis with wearable cameras and surveillance video analysis, respectively. The collaboration helped to identify critical technical challenges and carry out each research successfully by leveraging each group's strength.

5-3 Scientific, industrial or societal impacts/effects of the outputs

Through the joint research, young researchers in the University of Tokyo team were able to build a human network with internationally well-known computer vision researchers in the Hebrew University of Jerusalem team. This is one of the important accomplishments of this joint research in terms of human resource development.

国際共同研究における主要な研究成果リスト

1. 論文発表等

*原著論文 (相手側研究チームとの共著論文)

・査読有り：発表件数：計 0 件
該当無し

・査読無し：発表件数：計 0 件
該当無し

*原著論文 (相手側研究チームを含まない日本側研究チームの論文)：発表件数：計 3 件
・査読有り：発表件数：計 3 件

1. Ryo Yonetani, Vishnu Boddeti, Kris Kitani, and Yoichi Sato, "Privacy-Preserving Visual Learning Using Doubly Permuted Homomorphic Encryption", in Proc. *IEEE International Conference on Computer Vision (ICCV 2017)*, pp. 2059-2069, 2017. DOI: 10.1109/ICCV.2017.225
2. Yuki Sugita, Keita Higuchi, Ryo Yonetani, Rie Kamikubo, and Yoichi Sato, "Browsing Group First-Person Videos with 3D Visualization," In Proc. *ACM International Conference on Interactive Surfaces and Spaces (ISS '18)*, pp.55-60, 2018. DOI: 10.1145/3279778.3279783
3. Keita Higuchi, Hiroki Tsuchida, Eshed Ohn-Bar, Yoichi Sato, and Kris Kitani, "Learning Context-Dependent Personal Preferences for Adaptive Recommendation," to appear in *ACM Transactions on Interactive Intelligent Systems*, 2019. DOI: 10.1145/3359755

・査読無し：発表件数：計 0 件
該当なし

*その他の著作物 (相手側研究チームとの共著総説、書籍など)：発表件数：計 0 件
該当なし

*その他の著作物 (相手側研究チームを含まない日本側研究チームの総説、書籍など)：発表件数：計 0 件
該当なし

2. 学会発表

*口頭発表 (相手側研究チームとの連名発表)

発表件数：計 0 件 (うち招待講演：0 件)

*口頭発表 (相手側研究チームを含まない日本側研究チームの発表)

発表件数：計 4 件 (うち招待講演：0 件)

*ポスター発表 (相手側研究チームとの連名発表)

発表件数：計 0 件

*ポスター発表 (相手側研究チームを含まない日本側研究チームの発表)

発表件数：計 2 件

3. 主催したワークショップ・セミナー・シンポジウム等の開催

該当無し

4. 研究交流の実績 (主要な実績)

【合同ミーティング】

- ・ 2017年10月16日～18日の日程で、東京大学生産技術研究所において全体プロジェクトミーティングを実施した
- ・ オンライン会議システムを利用した研究ミーティングを適宜実施した

5. 特許出願

研究期間累積出願件数：0件

6. 受賞・新聞報道等

該当無し

7. その他

該当無し