

# Robust Real-World Crowd Pose Tracking

Das Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB ist eines der größten Institute für angewandte Forschung auf dem Gebiet der Bildgewinnung und Bildauswertung in Europa. Die Abteilung Videoauswertesysteme (VID) beschäftigt sich mit der automatischen Auswertung von Signalen bewegter bildgebender Sensorik in komplexen, ggf. nichtkooperativen Szenarien. Diese Sensorik wird beispielsweise im Aufklärungs- und Überwachungsbereich als integrierte Komponente in fliegenden, weltraumgestützten oder mobilen landgestützten Plattformen verwendet. VID entwickelt und integriert hierfür Bildauswertealgorithmen für autonome oder interaktive Systeme.

## Beschreibung

Die Analyse des Verhaltens Einzelner in Menschenmengen (engl. Crowds) oder Menschengruppen an öffentliche Plätze hat in diesem Jahr, beispielsweise durch Abstandsgebote, enorme an Bedeutung gewonnen. Dabei werden oft Körperposen als Abstraktion benutzt um die Grundrechte des Einzelnen nicht einzuschränken und die Verarbeitungsgeschwindigkeit zu beschleunigen. Das sog. Multi-Person-Pose-Tracking ist ein wichtiges Teilproblem für die Erkennung menschlicher Handlungen und das Verständnis von Szenen. Sie erfolgt in zwei Schritten: Erstens, die Schätzung der Keypoints einzelner Personen; Zweitens, der Verfolgungsschritt, wobei jeder Keypoint einer eindeutigen Person zugeordnet ist.



Abbildung 1: Posenschätzung für eine Crowd [1].

## Aufgabenstellung

Ihre Aufgabe im Rahmen dieser Abschlussarbeit ist es einen Überblick über den State of the Art Verfahren zu geben, die sich mit dem Tracking von Posen in Menschenmengen auseinandersetzen. Hierbei soll der Fokus vorrangig auf kompakten und schnellen Verfahren liegen. Anschließend entwickeln Sie auf Basis der Ergebnisse Ihrer Recherche ein eigenes Verfahren, welches in der Lage ist unter Verwendung von RGB-Kameradaten, zeitlich konsistente 3D Posen für Crowds zu generieren. Ausgangspunkt Ihrer Untersuchungen soll das in [2] referenzierte Verfahren sein, welches das Tracking nur auf Basis von Keypoints durchführt.

## Voraussetzungen

- Studienfach: Informatik, Mathematik, Angewandte Physik oder vergleichbar
- Sehr gutes Verständnis für die (theoretischen) Grundlage von Deep Learning
- Sehr gute Programmierkenntnisse (idealerweise Python)
- Erfahrung mit dem Deep Learning Framework Pytorch vorteilhaft.
- Fähigkeit zum selbstständigen Arbeiten
- Bereitschaft, sich in neue Themengebiete einzuarbeiten und Freude am Einbringen eigener Ideen

Bei Interesse senden Sie uns bitte Ihre Bewerbungsunterlagen (kurzes Anschreiben, tabellarischer Lebenslauf, Notenauszug) in elektronischer Form sowie den gewünschter Starttermin.

## Literatur

- [1] von Marcard, Timo, et al. "Recovering accurate 3d human pose in the wild using imus and a moving camera." Proceedings of the European Conference on Computer Vision (ECCV). 2018..
- [2] Snower, Michael, et al. "15 Keypoints Is All You Need." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

## Kontakt

Mickael Cormier, M. Sc.  
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung  
Fraunhoferstraße 1, 76131 Karlsruhe  
Tel.: 0721 / 6091-634  
mickael.cormier@iosb.fraunhofer.de