

# Optimization of Human Body Pose Estimation

## Optimierung der Schätzung menschlicher Körperposen

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 Bildauswertelgorithmen für autonome oder interaktive Systeme.

### Beschreibung

Das Schätzen menschlicher Körperposen auf Basis von 2D RGB-Bilddaten hat in den vergangenen Jahren zunehmend an Bedeutung gewonnen. Ein prominentes Beispiel ist die COCO Keypoint-Challenge, die seit einigen Jahren regelmäßig durchgeführt wird und bei der sich Forscherteams aus der ganzen Welt messen. Existierende Verfahren stehen jedoch noch vor diversen Herausforderungen. Eine davon ist die Frage, wie man mit entarteten Posenschätzungen umgeht oder sie direkt vermeidet. Abbildung 1 zeigt ein Beispiel für eine solche entartete Pose. Das Wissen über menschliche Körperproportionen und diversen physikalischen Eigenschaften einer Abbildung der Situation auf einen Bildsensor können dabei verwendet werden um weniger stark entartete Schätzungen zu generieren.



Abbildung 1: Entartete Posenschätzung

Existierende Verfahren stehen jedoch noch vor diversen Herausforderungen. Eine davon ist die Frage, wie man mit entarteten Posenschätzungen umgeht oder sie direkt vermeidet. Abbildung 1 zeigt ein Beispiel für eine solche entartete Pose. Das Wissen über menschliche Körperproportionen und diversen physikalischen Eigenschaften einer Abbildung der Situation auf einen Bildsensor können dabei verwendet werden um weniger stark entartete Schätzungen zu generieren.

### Aufgabenstellung

In einem ersten Schritt erweitern Sie einen bereits existierenden Datensatz um weitere Beispiel. Dieser Datensatz bildet anschließend die Grundlage für Ihre Untersuchungen und Experimente. Dabei umfasst Ihre Arbeit unter anderem Aufgaben wie die mathematische Definition einer menschlichen Pose (basierend auf COCO-Schema), die mathematische Formulierung des zugrundeliegenden Optimierungsproblems und das Durchführen von Transfer Learning auf den von Ihnen gewonnenen Daten. Abschließend evaluieren Sie Ihr Verfahren gegenüber State-of-the-Art-Verfahren hinsichtlich Leistungsfähigkeit und Geschwindigkeit.

### Voraussetzungen

- Studienfach: Informatik, Mathematik, Elektrotechnik, Angewandte Physik oder vergleichbar
- Gute Programmierkenntnisse (idealerweise Python oder C++)
- Grundlegende Kenntnisse im Themengebiet Deep Learning
- Erfahrung im Umgang mit Deep Learning Frameworks wie Tensorflow oder Keras
- Fähigkeit zum selbstständigen Arbeiten
- Bereitschaft, sich in neue Themengebiete einzuarbeiten und Freude am Einbringen eigener Ideen

### Quellen & Literatur

[1] Cao et al.: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017*

### Kontakt

Thomas Golda, M. Sc.  
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung  
Fraunhoferstraße 1, 76131 Karlsruhe  
Tel.: 0721 / 6091-631  
thomas.golda@iosb.fraunhofer.de