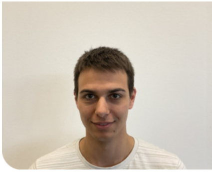
A Privacy-Preserving Multi-Camera People Tracking System

Graduate

Jan Fecker



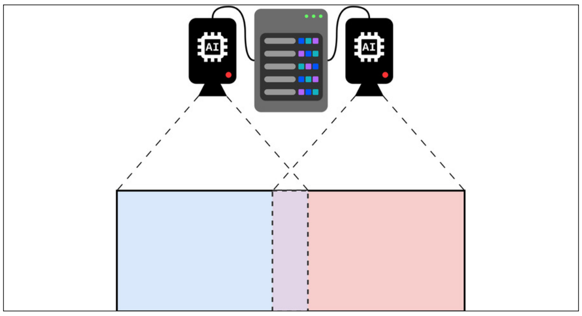
Yanick Thalmann

Initial Situation: The ability to detect, track, and consistently re-identify individuals across multiple cameras is a fundamental requirement for intelligent systems used in applications such as surveillance, service robotics, and indoor monitoring. In multi-camera environments, individuals frequently move between camera views with varying perspectives, illumination conditions, and levels of occlusion, making long-term identity preservation challenging. Many existing systems address these challenges through centralized processing architectures that transmit raw video streams to a central server. While such approaches can achieve high tracking accuracy, they introduce substantial privacy risks, high bandwidth requirements, and limited scalability, particularly in resource-constrained or privacy-sensitive deployments.

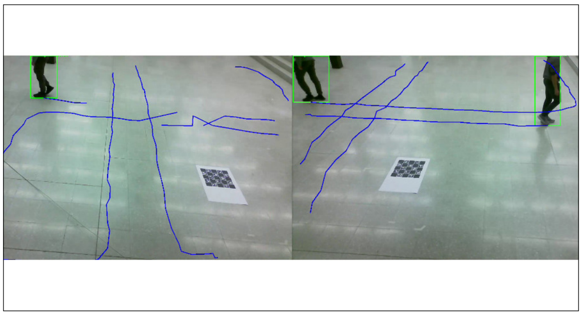
Approach: The objective of this thesis is to design, implement, and evaluate a privacy-preserving multi-camera person tracking and re-identification system that eliminates the transmission of raw visual data while enabling cross-camera trajectory fusion. Person detection, appearance-based re-identification, and short-term tracking are executed locally on distributed camera nodes using pre-trained deep learning models suitable for real-time edge deployment. All cameras are geometrically calibrated to a common reference plane using a ChArUco-based calibration procedure, allowing detections and trajectories to be projected into a shared ground-plane coordinate system via homography estimation. Only anonymized metadata, including re-identification embeddings, timestamps, and ground-plane coordinates, is transmitted to a central server over a lightweight, loss-tolerant UDP communication protocol. On the server side, incoming tracklets are temporally synchronized and associated across cameras using a combination of geometric proximity and appearance similarity, formulated as a linear assignment problem and solved using the Hungarian algorithm.

Conclusion: The experimental evaluation shows that cross-camera trajectory fusion operates reliably when detections are consistently available and camera-to-server synchronization is stable. In the evaluated scenarios, trajectories are successfully associated across cameras, enabling coherent movement analysis. Performance is primarily affected in transition regions where detection reliability decreases as individuals move between camera-covered areas, resulting in fragmented tracklets. The system was not evaluated in densely crowded scenes; therefore, its behavior under frequent path crossings is not assessed within the scope of this work. The use of a UDP-based communication protocol introduces an inherent trade-off, as metadata packets may be lost under increased system load or network variability, even when local detection and tracking operate correctly. Runtime and network traffic scale with the number of tracked individuals due to higher metadata throughput; however, movement analysis and region-based statistics remain functional under moderate load. Overall, the results indicate that the proposed system provides reliable privacy-preserving multi-camera fusion under typical operating conditions.

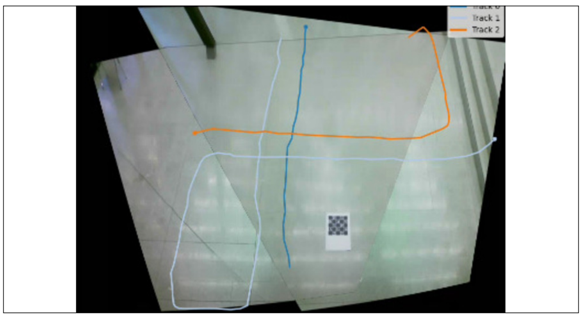
System overview illustrating two camera fields of view and the associated server.
Own presentment



Side-by-side camera configuration where detection and tracking are executed locally on each camera before fusion.
Own presentment



Two-dimensional tracking output after fusion for two individuals during a path-crossing event.
Own presentment



Advisors
Prof. Dr. Martin Weisenhorn, Marc Benz

Co-Examiner
Jonas Schmid, Stettbacher Signal Processing

Subject Area
Image Processing and Computer Vision