

Hosting laboratory : Laboratoire Systèmes et Applications des Technologies de l'Information et de l'Energie (SATIE)
Université Paris Sud Bâtiment 660, rue Noetzlin, Gif-sur-Yvette
Doctoral advisors : S. Le Hégarat-Masclé, E. Aldea
{sylvie.le-hegarat/emanuel.aldea}@u-psud.fr



Egocentric video registration for collaborative localization

Scientific project Our group is involved in an European project focused on improving the safety and security of mass gatherings in complex scenarios and widespread urban environments. In such situations, the crowd is often inhomogeneous in terms of density and motion dynamics. In addition, pedestrians are distributed over large areas. The surveillance over the course of an event is challenging for Law Enforcement Agencies (LEAs) even without disturbances. An up-to-date overview is often missing, and fast reactions in case of accidents, offensive behaviour or crowd densities that trap people and endanger lives (often referred to as “panic situations”) are difficult, while pre-emptive measures seem impossible.

Crowd analysis is an emerging topic in computer vision, being closely related to image processing, machine learning and multiple object tracking. Advances in this field benefit to a variety of problems, which may be more methodological, such as graph modelling for collective motion analysis, or more practical, such as improving the security of a specific location.

For a significant number of problems related to direct risk assessment (i.e. detection of aggressive behaviour, detection of phenomena that indicate danger for the crowd such as extreme density, tracking and identification of individuals who may represent a threat to others), mobile cameras can provide valuable information. Pose estimation/calibration of stationary and mobile cameras in particular for distributed camera networks is an established field of research ([1, 2, 3]), but urban areas raise specific difficulties (large distances, difficult interest point matching due to homogenous areas or repetitive structures, significant amounts of dynamic blobs representing groups of people without a clear background). To overcome these issues, alternative ways to allow highly accurate camera pose estimation (geo-registration) must be found. This challenge represents the objective of the proposed PhD scholarship.

Preliminary work on visual SLAM based calibration has been carried out by Fraunhofer IOSB [4], with promising results on the calibration of “mid-scale” camera networks with non-overlapping fields of view. The scientific aim of the PhD is to extend this approach to a large scale urban environment, by estimating the pose of a moving camera (UAV or wearable camera) in order to relate elements of interest to a larger context, which is available only in the wide field of vision of static cameras. Analysis of egocentric video is the foundation of augmented reality [5], and it has also developed lately as a prerequisite for re-identification in surveillance, or for collaborative target localization [6]. In our context however, the work will be focused on the highly accurate registration of a mobile camera within a static camera network or within a previously constructed cartography. Previous results indicate that exploiting additional sensors injects helpful information for the pose estimation [7, 8]. Based on previous work performed on GPS filtering in urban environments [9, 10] and on inertial-vision data fusion for navigation [11], we intend to rely as well on additional data sources (GPS and inertial sensors) in order to constrain and solve the pose estimation problem for wearable or UAV cameras.

Beside the collaboration opportunities with Fraunhofer IOSB, the PhD project is expected to capitalize on our group research experience in static camera pose estimation in large urban scenes [12], embedded systems and data fusion.

Prerequisites Research Master degree with strong background in computer vision and related fields.

Skills A good level in C++ programming. Data fusion. Experience with sensors, data recording, and generally any previous experience acquired in robotics related projects or robotics clubs will be appreciated.

- [1] J. C. SanMiguel, C. Micheloni, K. Shoop, G. L. Foresti, and A. Cavallaro, “Self-reconfigurable smart camera networks,” *Computer*, vol. 47, no. 5, pp. 67–73, 2014.
- [2] D. Devarajan, Z. Cheng, and R. J. Radke, “Calibrating distributed camera networks,” *Proceedings of the IEEE*, vol. 96, no. 10, pp. 1625–1639, 2008.
- [3] C. Ding, B. Song, A. Morye, J. A. Farrell, and A. K. Roy-Chowdhury, “Collaborative sensing in a distributed ptz camera network,” *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3282–3295, 2012.

- [4] T. Pollok and E. Monari, "A visual slam-based approach for calibration of distributed camera networks," in 13th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2016, Colorado Springs, CO, USA, August 23-26, 2016, 2016, pp. 429–437.
- [5] R. Castle, G. Klein, and D. W. Murray, "Video-rate localization in multiple maps for wearable augmented reality," in Wearable Computers, 2008. ISWC 2008. 12th IEEE International Symposium on. IEEE, 2008, pp. 15–22.
- [6] Y. Wang and A. Cavallaro, "Prioritized target tracking with active collaborative cameras," in Advanced Video and Signal Based Surveillance (AVSS), 2016 13th IEEE International Conference on. IEEE, 2016, pp. 131–137.
- [7] T. Teixeira, D. Jung, and A. Savvides, "Tasking networked cctv cameras and mobile phones to identify and localize multiple people," in Proceedings of the 12th ACM international conference on Ubiquitous computing. ACM, 2010, pp. 213–222.
- [8] Y. Goldman and I. Shimshoni, Robust epipolar geometry estimation using noisy pose priors. Technion-Israel Institute of Technology, Faculty of Computer Science, 2014.
- [9] S. Zair, S. Le Hégarat-Masclé, and E. Seignez, "A-contrario modeling for robust localization using raw gnss data," IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 5, pp. 1354–1367, 2016.
- [10] S. Zair, S. Le Hégarat Masclé, and E. Seignez, "Outlier detection in gnss pseudo-range/doppler measurements for robust localization," Sensors, vol. 16, no. 4, p. 580, 2016.
- [11] N. Zarrouati, E. Aldea, and P. Rouchon, "So(3)-invariant asymptotic observers for dense depth field estimation based on visual data and known camera motion," in American Control Conference 2012, Montreal, 2012, pp. 4116 – 4123.
- [12] N. Pellicano, E. Aldea, and S. L. Hégarat-Masclé, "Robust wide baseline pose estimation from video," in 23rd International Conference on Pattern Recognition, ICPR 2016, Cancún, Mexico, December 4-8, 2016, 2016, pp. 3820–3825.