L. Teixeira, F. Maffra, M. Moos, and M. Chli, "Vi-rpe: Visual-inertial relative pose estimation for aerial vehicles," *IEEE Robotics and Automation Letters (RA-L)*, vol. 3, no. 4, pp. 2770–2777, 2018: argued that a collaborative relative pose estimation is required to enable MAV flights close to structures that may have insufficient texture near or on the target. The proposed method uses a master and slave architecture, where the slave MAV is positioned close to the target structure, and the master MAV acting as a reference to the slave is positioned further away. The master is retrofitted with a known constellation of LED markers in order for the slave MAV to estimate the relative pose.

J. G. Morrison, D. Gálvez-López, and G. Sibley, "Moarslam: Multiple operator augmented rslam," in *Distributed Autonomous Robotic Systems*, pp. 119–132, Springer, 2016 A collaborative SLAM system is the work of , where the system is designed to work with multiple hand-held devices equipped with either a monocular camera and IMU or stereo camera pair, along with a mobile phone or autonomous robot for network and computing capabilities. Each agent performs full SLAM on board, and a centralized server is used for storing and sharing maps between agents. However, this system exhibits few other collaborative features, for one it does not perform a global optimization of the maps created by different agents, nor are the map optimization work load shared amongst other agents or server.

I. Deutsch, M. Liu, and R. Siegwart, "A framework for multi-robot pose graph slam," in *IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pp. 567–572, June 2016 a system that can combine different pose graph based SLAM systems running on the agents is presented. Additionally, the centralized server informs agents about updates in their respective pose graphs. However, the global map and sub-maps of other agents are not shared amongst the agents.

T. Cieslewski and D. Scaramuzza, "Efficient decentralized visual place recognition using a distributed inverted index," *IEEE Robotics and Automation Letters (RA-L)*, vol. 2, pp. 640–647, April 2017: A rapid exploration method for a MAV in an unknown environment is devised. The core contribution of this work lies in maintaining a constant velocity while exploring an unknown environment, thereby reducing the amount of wasted energy hovering in place. During outdoor experiments, however, they experienced frequent visual odometry failures due to lighting conditions and sudden quadrotor attitude changes.

E. Palazzolo and C. Stachniss, "Information-driven autonomous exploration for a vision-based mav," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, p. 59, 2017: proposed an information-driven path planning algorithm for a MAV through an unknown environment. The MAV is assumed to be carrying a forward-facing camera to perform object mapping, and has a 3D bounding box around the target object of interest as a prior. This method is validated in simulation only.

D. Zou and P. Tan, "CoSLAM: Collaborative visual slam in dynamic environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 354–366, 2013: CoSLAM obtains multiple monocular camera streams as input, and groups cameras using a place recognition module. The disadvantage of this system are, first, all cameras need to be synchronized. Secondly, to initialize the SLAM systems all cameras must observe the same scene. Lastly, the processing of all data requires a graphics processing unit (GPU) to run at real time.

P. Schmuck and M. Chli, "CCM-SLAM: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams," *Journal of Field Robotics (JFR)*, 2018: Collaborative centralized multi-agent SLAM (CCM-SLAM) The system was designed to fuse data streams from multiple monocular cameras on board multiple MAVs. Each MAV runs a visual odometry system for autonomy, and offloads the computationally intensive tasks back to the central server. All information between the MAVs and the central server are shared, for example if an agent visits an area of the environment previously visited by another agent, previous experiences are retrieved from the server. Finally, the server detects and removes redundant data without compromising the robustness of the estimation.

A. Cunningham, V. Indelman, and F. Dellaert, "DDF-SAM 2.0: Consistent distributed smoothing and mapping," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5220–5227, IEEE, 2013: is an example of decentralized SLAM system, where an anti-factor is introduced in order to remove and replace redundant information in a graph SLAM problem.

T. Cieslewski and D. Scaramuzza, "Efficient decentralized visual place recognition using a distributed inverted index," *IEEE Robotics and Automation Letters (RA-L)*, vol. 2, pp. 640–647, April 2017: proposes a method that requires a similar amount of data exchange between agents as a centralized approach without precluding any matches. This is achieved by pre-assigning visual bag-of-words vocabulary to different robots, and obtaining a candidate selection to choose which robot to send the full query. Both works were validated in simulation only.

A. Richardson, J. Strom, and E. Olson, "AprilCal: Assisted and repeatable camera calibration," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, November 2013: AprilCal guides users to collect a high-quality camera images inorder to perform camera intrinsics calibration for a monocular camera.

J. Rebello, A. Das, and S. L. Waslander, "Autonomous active calibration of a dynamic camera cluster using next-best-view," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017: Calibrates the extrinsics between a stereo camera pair, where one camera is rigidly mounted and the other camera is actuated by a gimbal. The key contribution of this work is the next-best-view selection and using it to reduce the total amount of image data for a good extrinsic calibration.

R. Bajcsy, "Active perception," 1988 Actuating a camera intentionally is also known as active vision. Active vision is by no means new, one of the earliest work can be traced back to [11] where active vision is defined as a problem of an intelligent data acquisition process.

M. W. Achtelik, S. Weiss, M. Chli, F. Dellaerty, and R. Siegwart, "Collaborative stereo," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2242–2248, IEEE, 2011 demonstrated that two MAVs equipped with onboard monocular cameras and IMUs can form a flexible stereo rig. This was achieved by performing feature correspondence in the overlapping field of view between the MAVs to estimate the relative pose.

D. Zou and P. Tan, "CoSLAM: Collaborative visual slam in dynamic environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 354–366, 2013 multiple handheld cameras, but cameras are all synchronized making it impractical for robotic applications, where the input of each camera should be computed asynchronoously in order to cope with missing data and delays. Additionally, it is assumed that all cameras observe the same scene at the start.

A. Cunningham, V. Indelman, and F. Dellaert, "DDF-SAM 2.0: Consistent distributed smoothing and mapping," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5220–5227, IEEE, 2013: presents a fully decentralized SLAM system where each robot maintains a consistent augmented local map that combines local and neighbourhood information, but the system has only been validated in simulation.

References

[1] L. Teixeira, F. Maffra, M. Moos, and M. Chli, "Vi-rpe: Visual-inertial relative pose estimation for aerial vehicles," *IEEE Robotics and Automation Letters (RA-L)*, vol. 3, no. 4, pp. 2770–2777, 2018.

[2] J. G. Morrison, D. Gálvez-López, and G. Sibley, "Moarslam: Multiple operator augmented rslam," in *Distributed Autonomous Robotic Systems*, pp. 119–132, Springer, 2016.

[3] I. Deutsch, M. Liu, and R. Siegwart, "A framework for multi-robot pose graph slam," in *IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pp. 567–572, June 2016.

[4] T. Cieslewski and D. Scaramuzza, "Efficient decentralized visual place recognition using a distributed inverted index," *IEEE Robotics and Automation Letters (RA-L)*, vol. 2, pp. 640–647, April 2017.

[5] E. Palazzolo and C. Stachniss, "Information-driven autonomous exploration for a vision-based mav," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, p. 59, 2017.

[6] D. Zou and P. Tan, "CoSLAM: Collaborative visual slam in dynamic environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 354–366, 2013.

[7] P. Schmuck and M. Chli, "CCM-SLAM: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams," *Journal of Field Robotics (JFR)*, 2018.

[8] A. Cunningham, V. Indelman, and F. Dellaert, "DDF-SAM 2.0: Consistent distributed smoothing and mapping," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5220–5227, IEEE, 2013.

[9] A. Richardson, J. Strom, and E. Olson, "AprilCal: Assisted and repeatable camera calibration," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, November 2013.

[10] J. Rebello, A. Das, and S. L. Waslander, "Autonomous active calibration of a dynamic camera cluster using next-best-view," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017.

[11] R. Bajcsy, "Active perception," 1988.

[12] M. W. Achtelik, S. Weiss, M. Chli, F. Dellaerty, and R. Siegwart, "Collaborative stereo," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2242–2248, IEEE, 2011.