

Thermal Simultaneous Localization and Mapping

Subjects: **Robotics**

Contributor: Benjamin Ronald van Manen , Victor Sluiter , Abeje Yenehun Mersha

Thermal Simultaneous Localization and Mapping (SLAM) is a rather new research topic, with most results being published in the last decade. Research is mainly focusing on thermal odometry rather than thermal SLAM, with several research groups that succeeded in utilizing thermal cameras to compute the odometry of a vehicle. For thermal 3D-mapping, i.e., the visualization of thermal data in 3D, the most common method is to superimpose the thermal images onto a pointcloud obtained from data generated from another depth source.

hazardous environment

SLAM

computer vision

thermal

1. Thermal Feature Extraction and Matching

In a collaboration between the Royal Institute of Technology (KTH) in Sweden and FLIR Systems AB, J. Johansson et al. ^[1] evaluated the performance of different visual-based feature detectors and descriptors on thermal images. The dataset is composed of individual image pairs taken in benign structural and textured environments. Before evaluation, the images were rescaled in an 8-bit manner and their histograms were equalized. The different algorithms were tested on different image deformations, such as viewpoint change, rotation, scale, noise and downsampling. The different combinations of candidates are evaluated based on the recall, i.e., the number of correct feature matches according to RANSAC with respect to the total amount of matches. Among the floating-point descriptors, the Hessian-affine extractor in combination with the LIOP descriptor showed consistent good performance on all the different image deformations. The SURF extractor and descriptor followed behind, demonstrating particularly good invariance against Gaussian blur and Gaussian white noise. In contrast to working with visual images, binary descriptors offer a similar performance to floating-point descriptors in thermal images at a lower computational cost. Especially, the combination ORB with FREAK or BRISK provided overall competing results.

A second benchmark of feature extraction and description methods of thermal features was performed by T. Mouats et al. ^[2]. As the targeted application was estimating the odometry of an Unmanned Aerial Vehicle (UAV) using thermal images, the processing time was also considered in the evaluation. The datasets used in the benchmark consisted of a number of video sequences recorded in benign indoor and outdoor situations using a FLIR Tau2 LWIR camera. It was concluded that, in general, blob detectors such as SIFT and SURF provided lower repeatability than corner detectors like FAST and GFTT. This means that, when an object is observed from different perspectives, the likelihood is smaller that features are located at the same location on the object. On the other hand, blob features showed more distinctiveness than corner features as the comparison of their descriptor yielded

higher matching scores, i.e., more correct matches. SIFT feature extraction performed the worst in almost all aspects. Feature extractors seem not to be affected by motion blur, with the SURF blob extractor performing the best. It was observed that the descriptors were less influenced by non-uniformity noise (NUC) than expected. The effect seemed, however, to be stronger in indoor environments, where objects often have a uniform temperature, providing less contrast to the image and thus amplifying the noise visibility. The authors concluded by proposing SURF feature extraction in combination with the FREAK binary descriptor as the best combination for thermal navigation applications. The proposed combination offers a good tradeoff between good matchability as well as repeatability while also providing a satisfactory computation time for real-time navigation.

The two benchmarks presented here are performed using datasets recorded in benign environments with a uniform temperature. Both smoke and intense heat sources have a substantial impact on the environment, which may affect the feature extraction and description process as well. Furthermore, a uniform distribution of the features is desired for a more accurate estimation of the odometry [3]. Notably, in the benchmark from Mouats et al. [2], which intends to apply feature extraction and description in thermal odometry estimation, a metric capable of quantifying the feature distribution across the frame is lacking.

2. Thermal Odometry

Research from T. Mouats et al. [4] successfully implemented stereo thermal odometry on an unmanned aerial vehicle (UAV) for outdoor environments. Thermal features are obtained using the fast Hessian extractor and described with the FREAK descriptor. In their research, the authors added a comparison of different camera calibration methods. When evaluating different optimizers, it was discovered that the double dogleg algorithm provides better motion estimation with respect to other alternatives. The authors did not yet incorporate any trajectory optimization, such as loop closure detection. A global 3D map is not generated, but a local dense point cloud was created using a pair of stereo thermal images.

Borges et al. [5] utilized a monocular thermal camera and semi-dense optical flow estimation to determine the odometry of a road vehicle. The scale of the estimated path is subsequently obtained from inertial data and road plane segmentation. The researchers proactively triggered the camera's Flat Field Correction (FFC) against Non-Uniform Noise (NUC) to prevent the loss of the image stream during critical moments such as turns. The optimal FFC triggering is determined according to temperature changes, the current vehicle turning rate, the upcoming turns and the time elapsed since the last correction. The method proposed by the authors is specifically tailored to the application of autonomous road vehicles. The algorithm namely relies on road plane segmentation to estimate the scale factor of the movement as well as a trajectory known a priori to trigger Flat Field Corrections.

The benefits of using the raw radiometric data from the thermal camera instead of rescaled 8-bit images were investigated by S. Khattak et al. [6]. The authors compared the performance of two monocular visual-inertial odometry systems operating on 8-bit rescaled thermal images, OKVIS and ROVIO, as well as two monocular thermal-inertial odometry algorithms, ROTIO and KTIO. The authors first explored the use of radiometric data by modifying ROVIO into ROTIO, which extracts FAST corner features in 8-bit rescaled images. The keypoints are

then tracked across the frames in a semi-direct manner using the radiometric data. Improving on ROTIO, KTIO also detects features in the radiometric data according to their gradient value. The obtained points of interest are subsequently tracked in a semi-direct manner using the gradient image of the radiometric data. In contrast to rescaled 8-bit images, the authors noted that radiometric data are unaffected by contrast changes due to sudden temperature changes. Additionally, KTIO showed significantly higher operational robustness. The algorithm demonstrated accurate performance in situations where other visual–inertial and thermal–inertial odometry methods had already diverged during tests on a UAV in a dusty underground mine. Due to its absence in feature repeatability, the feature extraction used in KTIO is unsuitable for conventional feature-based loop closure detection methods, such as bag of visual words. The authors, therefore, recommend investigating more direct loop closure detection methods [7].

M.R.U. Saputra et al. [8] designed DeepTIO, a monocular thermal–inertial odometry algorithm utilizing a deep-learning approach. The optical flow from the thermal camera is estimated using the FlowNet-Simple network. In an improved version, an additional network incorporates "hallucinated" details into the thermal images. The hallucination network is trained from visual images from a multi-spectral dataset. The algorithm was tested, among other locations, in a smoke-filled room at a firefighters training facility. No fire source, however, was visible in the dataset. DeepTIO outperformed other thermal–inertial odometry methods, but the authors noted that the network lacks flexibility with respect to frame rate changes. Furthermore, it was observed that the hallucinated details did not provide any significant additional accuracy in the odometry estimation.

J. Jiang et al. [9] modified the visual deep-learning optical flow estimation algorithm RAFT for thermal images, ThermalRAFT. To decrease the image noise, the thermal images are pre-processed using a Singular Value Decomposition (SVD) that decomposes the image in independent linear components. The highest singular value is set to zero, dimming the image but also diminishing the noise. The remaining singular values are reallocated in ascending order, further diminishing the noise while also enhancing the details in low-contrast image regions. Gradient features similar to DSO are extracted from the thermal images before being tracked using the proposed ThermalRAFT algorithm. Potential loop closures are detected using the DBoW2 library by describing the extracted gradient features using the BRIEF descriptor.

Y. Wang et al. [10] and W. Chen et al. [11] extracted edge features using a Difference of Gaussians (DoG) filter instead of point features. The DoG filter's bandpass filter behavior makes it more robust against image noise and low-contrast regions. Y. Wang et al. implemented the DoG edge filter in ETIO, a monocular thermal–inertial odometry algorithm for visually degraded environments that tracks the edge features using IMU-aided Kanade–Lucas–Tomasi feature tracking. ETIO was tested in outdoor settings, a mine dataset and an indoor environment filled with artificial smoke. Similarly, Chen et al. implemented the DoG edge filter into EIL-Simultaneous Localization and Mapping (SLAM), a thermal–LiDAR edge-based SLAM algorithm. EIL-SLAM combines the estimated thermal and LiDAR odometry to increase the depth and scale estimation. A visual–thermal–LiDAR place recognition algorithm is utilized to detect loop closures for global bundle adjustment. The algorithm was tested in an urban scene during day- and nighttime.

With the exception of DeepTIO [8], all thermal odometry algorithms have only been tested in benign environments. Although the performance of DeepTIO [8] was evaluated in real smoke, no fire source was present in the vicinity. The effect of the fire source itself on the odometry estimation could, therefore, not be analyzed. Finally, a global trajectory optimization step, for example, in the form of bundle adjustment, has not been implemented in any of the current thermal odometry algorithms.

3. Thermal Mapping

S. Vidas et al. [12] developed a monocular SLAM system for hand-held thermal cameras. The algorithm detects corner features using the GFTT and FAST detectors and subsequently tracks them using sparse Lucas–Kanade optical flow. In order to process the thermal images, the raw 14-bit thermal data are converted to 8-bit from a fixed range centered around the mean value between the lowest and the highest 14-bit intensity. Matched SURF features are used in a homography motion estimation in order to resume tracking after an FFC. Only local trajectory optimizations are performed on all frames between two keyframes. A novel metric based on five criteria to select a new keyframe is additionally proposed. The authors claim that a reprojection error less than 1.5 can be considered as a good motion estimate. In most sequences, the proposed algorithm could not attain this value for longer periods. Being a semi-direct approach, the system notably failed during pure rotations. No information on the quality of the obtained 3D thermal map was presented by the authors.

Y. Shin et al. [13] combined depth data from a LiDAR and thermal data from a LWIR camera to create a thermal map. A sparse set of LiDAR points are projected into the 14-bit thermal image and are subsequently tracked in the following frame using a direct approach. Potential loop closures are detected using the bag of visual words approach by extracting ORB features from 8-bit images converted from a predefined temperature range. The authors note that this approach can result in a large amount of false positives in environments with a lack of detailed textures. An additional geometric verification is, therefore, presented to try to mitigate this phenomenon. The authors claim to have developed an all-day visual SLAM system as it is capable of computing an accurate trajectory and 3D thermal map irrespective of the outdoor lighting conditions.

E. Emilsson et al. [14] developed Chameleon, a stereo thermal–inertial SLAM system based on EKF-SLAM. The pose is estimated by tracking a maximum of thirty landmarks extracted using SIFT feature extraction and description. Experiments were conducted in both cold and heated environments, with the authors mentioning that positioning using only thermal–inertial data is not feasible in the former due to a lack of contrast in the thermal images. Although no quantitative error is provided, it is claimed that the algorithm estimates the position within a few meters inside a firefighting training facility where an active fire is present. The sparse mapping fails to provide a comprehensible map of the building. Finally, the authors note that, because of a temperature difference between the environment and the enclosed box in which the cameras are mounted, condensation was forming over time, leading to degraded tracking and mapping results.

When a thermal map is the only desired result, superimposing thermal data onto a 3D map generated from a SLAM algorithm utilizing a more common sensor is often preferred. Both the setup's location and the depth data

are then obtained from, for example, LiDAR measurements [\[15\]\[16\]](#) or from an RGB camera [\[17\]\[18\]](#) or from a depth camera [\[19\]](#). Thermal information is then added to the 3D map by projecting the map points into the thermal images.

In benign environments, the construction of a thermal map is often aided by a second more reliable source of depth information. In a smoke-filled environment, these secondary sensors are, nevertheless, rendered useless. The thermal–inertial SLAM algorithm from E. Emilsson et al. [\[14\]](#), tested in a fiery environment, delivers a very limited map, and the system lacks global optimization capabilities. Finally, Chameleon suffers from lens condensation in fiery environments. Therefore, pose estimation heavily relied on inertial data for a significant part of the trajectory.

References

1. Johansson, J.; Solli, M.; Maki, A. An evaluation of local feature detectors and descriptors for infrared images. In Proceedings of the Computer Vision—ECCV 2016 Workshops, Amsterdam, The Netherlands, 8–10 and 15–16 October 2016; Part III. pp. 711–723.
2. Mouats, T.; Aouf, N.; Nam, D.; Vidas, S. Performance evaluation of feature detectors and descriptors beyond the visible. *J. Intell. Robot. Syst.* 2018, 92, 33–63.
3. Nguyen, H.; Lee, S. Orthogonality Index Based Optimal Feature Selection for Visual Odometry. *IEEE Access* 2019, 7, 62284–62299.
4. Mouats, T.; Aouf, N.; Chermak, L.; Richardson, M.A. Thermal stereo odometry for UAVs. *Sensors* 2015, 15, 6335–6347.
5. Borges, P.V.K.; Vidas, S. Practical infrared visual odometry. *IEEE Trans. Intell. Transp. Syst.* 2016, 17, 2205–2213.
6. Khattak, S.; Mascarich, F.; Dang, T.; Papachristos, C.; Alexis, K. Robust thermal-inertial localization for aerial robots: A case for direct methods. In Proceedings of the International Conference on Unmanned Aircraft Systems (ICUAS), Atlanta, GA, USA, 11–14 June 2019.
7. Khattak, S.; Papachristos, C.; Alexis, K. Keyframe-based direct thermal-inertial odometry. In Proceedings of the International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019.
8. Saputra, M.R.U.; de Gusmao, P.B.; Lu, C.X.; Almalioglu, Y. DeepTIO: A deep thermal-inertial odometry with visual hallucination. *IEEE Robot. Autom. Lett.* 2020, 5, 1672–1679.
9. Jiang, J.; Chen, X.; Dai, W.; Gao, Z.; Zhang, Y. Thermal-Inertial SLAM for the Environment with Challenging Illumination. *IEEE Robot. Autom. Lett.* 2022, 7, 8767–8774.
10. Wang, Y.; Chen, H.; Liu, Y.; Zhang, S. Edge-Based Monocular Thermal-Inertial Odometry in Visually Degraded Environments. *IEEE Robot. Autom. Lett.* 2023, 8, 2078–2085.

11. Chen, W.; Wang, Y.; Chen, H.; Liu, Y. EIL-SLAM: Depth-enhanced edge-based infrared-LiDAR SLAM. *J. Field Robot.* 2021, 39, 117–130.
12. Vidas, S.; Sridharan, S. Hand-held monocular SLAM in thermal-infrared. In *Proceedings of the 2012 12th International Conference on Control Automation Robotics & Vision (ICARCV)*, Guangzhou, China, 5–7 December 2013.
13. Shin, Y.; Kim, A. Sparse depth enhanced direct thermal-infrared SLAM beyond the visible spectrum. *IEEE Robot. Autom. Lett.* 2019, 4, 2918–2925.
14. Emilsson, E.; Rydell, J. Chameleon on fire—Thermal infrared indoor positioning. In *Proceedings of the 2014 IEEE/ION Position, Location and Navigation Symposium—PLANS 2014*, Monterey, CA, USA, 5–8 May 2014.
15. Borrmann, D.; Nüchter, A.; Dakulović, M.; Maurović, I.; Petrović, I.; Osmanković, D.; Velagić, J. The project thermalmapper—thermal 3d mapping of indoor environments for saving energy. *IFAC Proc. Vol.* 2012, 45, 31–38.
16. De Pazzi, D.; Pertile, M.; Chiodini, S. 3D Radiometric Mapping by Means of LiDAR SLAM and Thermal Camera Data Fusion. *Sensors* 2022, 22, 8512.
17. Ham, Y.; Golparvar-Fard, M. An automated vision-based method for rapid 3D energy performance modeling of existing buildings using thermal and digital imagery. *Adv. Eng. Inform.* 2013, 27, 295–409.
18. Yamaguchi, M.; Saito, H.; Yachida, S. Application of LSD-SLAM for visualization temperature in wide-area environment. In *Proceedings of the International Conference on Computer Vision Theory and Applications*, Porto, Portugal, 27 February–1 March 2017; SciTePress: Setubal, Portugal, 2017; Volume 5, pp. 216–223.
19. Chen, X.; Dai, W.; Jiang, J.; He, B.; Zhang, Y. Thermal-Depth Odometry in Challenging Illumination Conditions. *IEEE Robot. Autom. Lett.* 2023, 8, 8767–8774.

Retrieved from <https://encyclopedia.pub/entry/history/show/111991>