Black Sea Journal of Engineering and Science
doi: 10.34248/bsengineering.1578563

Open Access Journal
e-ISSN: 2619 – 8991

Review
Volume 8 - Issue 2: 558-568 / March 2025

THERMAL SLAM: HARNESSING TEMPERATURE VARIATIONS TO ENHANCE OBJECT DETECTION AND TRACKING PERFORMANCE

Fathia SALEM[1]*, Osman Serdar GEDİK[1]

[1]*Ankara Yıldırım Beyazıt University, Faculty of Engineering, Department of Computer Engineering, 06460, Ankara, Türkiye*

Abstract: Thermal Simultaneous Localization and Mapping (SLAM) is a burgeoning field that collects robotics, computer vision, and thermal imaging. In this paper, we tried to present a thorough review of recent advancements in thermal SLAM, with a focus on its role in enhancing object detection and tracking. For better performance in low light, resistance to obstructions, and accuracy in bad weather, thermal SLAM systems work better with visual-based SLAM systems because they use changes in temperature in the environment. The review paper explains the fundamental principles of SLAM, including sensor technologies, data fusion techniques, and mapping algorithms. It then explores the methodologies used for object detection and tracking within the Thermal SLAM framework, encompassing classical approaches and deep learning techniques tailored for thermal imagery analysis. Additionally, the paper discusses challenges and limitations specific to thermal SLAM, such as thermal drift, sensor noise, and calibration issues, while also identifying potential areas for future research. The paper provides a comprehensive survey of applications that utilize thermal SLAM for object detection and tracking across various domains, including autonomous navigation, surveillance, search and rescue operations, and environmental monitoring. It synthesizes case studies and experimental results from relevant literature to demonstrate the effectiveness and practicality of thermal SLAM in complex scenarios where traditional visual-based methods struggle. Overall, this review emphasizes the role of thermal SLAM in advancing autonomous systems and enabling robust object detection and tracking in challenging environments. Examining recent developments, challenges, and applications, it sheds light on the progress made in this field.

Keywords: SLAM, Thermal camera, Object detection, Robust localization, Object tracking

***Corresponding author:** Ankara Yıldırım Beyazıt University, Faculty of Engineering, Department of Computer Engineering, 06460, Ankara, Türkiye
E mail: fafyalmnfy2014@gmail.com (F. SALEM)
Fathia SALEM https://orcid.org/0009-0000-9332-8125
O Serdar GEDİK https://orcid.org/0000-0002-1863-8614

Received: November 3, 2024
Accepted: January 28, 2025
Published: March 15, 2025

Cite as: Salem F, Gedik OS. 2025. Thermal slam: Harnessing temperature variations to enhance object detection and tracking performance, BSJ Eng Sci, 8(2): 558-568.

1. Introduction

Thermal Simultaneous Localization and Mapping (SLAM) is a developing discipline that integrates robotics, computer vision, and thermal imaging. This study presents a thorough examination of the latest developments in thermal SLAM, specifically highlighting its contribution to improving object detection and tracking. Thermal SLAM systems are more effective than visual-based SLAM systems in low light conditions, when there are obstructions, and during bad weather. This is because thermal SLAM systems rely on detecting changes in temperature in the surroundings.

In the field of robotics and autonomous systems, one of the most important tasks is using a technique known as Simultaneous Localization and Mapping (SLAM) that accomplishes estimation of robot's position in a generated map. The essential aim of this system is to offer assistance to a mobile agent in the process of navigating through an unfamiliar environment. The agent is able to simultaneously determine its location and then create a map to cover the adjacent "surrounding" area thanks to this capability. It is essential for autonomous systems to have an accurate estimation of the location and the surrounding map because it serves as a forerunner to a variety of other robotic operations, including navigation, exploration, and manipulation. The front end and the back end are the two fundamental components that make up a conventional structural localization and mapping architecture. The front end is tasked with acquiring sensor input and converting it into a more appropriate format for inference. Conversely, the back end utilizes the data obtained from the front end to calculate the agent's states. Additionally, the back end plays a vital role in generating an environment representation and optimizing the agent states that are consistent across the entire globe (Cadena et al., 2016; Li et al., 2019). Most front-ends used in SLAM systems primarily depend on either range sensor (such as depth or LIDAR sensors) or vision sensors (RGB cameras) to detect the adjacent environment. Noteworthy examples are ORB-SLAM as explored in the research by Mur-Artal et al. (2015) which employs RGB sensors, and LOAM by Zhang and Singh (2014) which utilizes LIDAR sensors as their respective SLAM front-ends. While both range-based and vision-based SLAM systems generally perform effectively across different applications, their success significantly relies on ideal visibility circumstances. Nevertheless, the use of current sensors that are based on range, as well as vision for SLAM estimates, becomes challenging in the presence of unfavorable lighting conditions or airborne particulates like dust, soot, and smoke. For example, it is commonly recognized that RGB cameras are unable to operate in low light conditions; in contrast, cameras based on depth can be affected by bright illumination and glare (Debeunne and Vivet, 2020; Khattak et al., 2019b). Similar visual challenges arise when RGB or other sensors operate in environments with airborne particles (such as depth and LIDAR), as discussed by Bijelic et al. (2020) dense fog, or mist. Thermal imaging cameras, in contrast, are unaffected by most airborne pollutants and illumination conditions, as explored in the research by Brunner and Peynot (2014). The Long-Wave Infrared (LWIR) data released by nearby objects is captured by these cameras. The cameras that utilize properties of thermal imaging have distinct benefits that make them a practical alternative for SLAM applications in areas with limited visual information. However, the development of a complete thermal SLAM system poses a series of difficult obstacles. The key problem is to abstract or encode temperature data in a way that best supports the graph optimization process. The difficulty of this task stems from the fact that cameras utilizing thermal properties record the temperature distribution of the surroundings rather than their visual appearance and shape. The problem is exacerbated by the re-measurement accuracy of 8-bit thermal data, which reduces contrast and complicates the application of conventional feature identification and data binding techniques.

In addition, these types of thermal cameras require periodic halting of a process that takes between 0.5 and 1 second to perform Non-Uniformity Correction (NUC), often referred to as Flat Field Correction (FFC) (Mouats et al., 2015; Delaune et al., 2019). During this procedure, the sensor is exposed to a consistent temperature to calculate the correction values for fixed- pattern noise. Thermal cameras pose several additional challenges, including restricted resolution, deteriorating signal-to-noise ratio, and diminished contrast over time. These concerns suggest that the traditional approaches used for other optical sensors are not suitable for the usual front-end abstraction process in thermal SLAM systems, as explained by Wang et al. (2017). Thermal SLAM utilizes temperature fluctuations in the surroundings and enhances conventional visual-based SLAM systems by providing additional benefits, including better detection and tracking in low-light situations, resistance to obstructions, and improved accuracy in unfavorable weather conditions.

1.1. Non-Uniformity Correction (NUC)

Non-Uniformity Correction (NUC) is an essential preprocessing step in thermal imaging that compensates for the inherent inconsistencies in the response of thermal detectors Wu et al (2023). These inconsistencies can degrade image quality, which in turn affects the accuracy of subsequent tasks like localization and mapping in systems such as SLAM. Thermal cameras are susceptible to fixed pattern noise (FPN), a phenomenon caused by variations in the sensitivity of individual sensor elements. These variations arise due to manufacturing imperfections or external environmental factors, leading to a non-uniform thermal response across the camera's array of detectors.

To address these issues, NUC typically involves a brief pause in image capture, during which the thermal camera analyzes a uniform temperature source. This process allows the camera to adjust for sensitivity discrepancies, thereby improving the uniformity of the thermal images. However, this adjustment process can result in temporary data interruptions, halting image capture for a short period. As a result, real-time applications such as thermal SLAM, which require continuous data acquisition, are challenged by these interruptions.

The NUC process, if not properly executed, can introduce artifacts into the thermal images, such as blurred edges, a low signal-to-noise ratio, and inadequate texture representation. These artifacts negatively impact the clarity of thermal images and, consequently, the accuracy of localization and mapping tasks. To address these challenges, this study proposes a novel scene-based NUC method that is integrated with a monocular thermal SLAM system. This method not only reduces the data interruptions typically caused by NUC but also improves the overall quality of thermal images by taking advantage of real-time processing capabilities inherent in modern SLAM technologies. By incorporating advanced denoising algorithms along with optimized NUC strategies into our MonoThermal-SLAM framework, this approach aims to expand the use of thermal cameras in a variety of challenging environments. This integration ensures high-quality spatial localization and mapping, even in conditions that previously hindered real-time thermal imaging applications.

1.2. FFC (Flat Field Correction)

Flat Field Correction (FFC) is an essential preprocessing technique applied to thermal images to compensate for non-uniformities in sensor response, which can be exacerbated in challenging environments such as fire incidents as explored in the research by van Manen et al. (2023). The performance of thermal cameras can degrade due to various factors, including temperature variations and environmental conditions. In this study, FFC is particularly crucial given that inconsistent image quality may hinder feature extraction and pose estimation accuracy.

In the context of FirebotSLAM, a proactive triggering mechanism for FFC was employed based on several criteria: changes in temperature detected by the camera, the vehicle's turning rate, anticipated turns during navigation, and elapsed time since the last correction.

This adaptive approach ensures that corrections are performed at critical moments when inaccuracies are most likely to occur.

It was observed that timely application of FFC significantly contributes to maintaining consistent image quality throughout data acquisition phases during fire scenarios. The integration of FFC into FirebotSLAM enables better feature tracking by mitigating noise and enhancing overall imaging reliability. However, it is important to note that while effective during periods with stable environmental conditions or gradual changes in scene temperature, rapid fluctuations—such as those caused by dynamic fire sources—may still challenge traditional correction methodologies.

The authors highlight that further optimization of FFC routines could enhance system resilience against harsh thermal gradients encountered within smoke-filled environments while improving the robustness of extracted features for Odometry calculations.

1.3. LWIR (Long-Wave Infrared)

(Improving SLAM with Thermal Imagery)

Recent advancements in Simultaneous Localization and Mapping (SLAM) have encountered significant challenges when applied to Long-Wave Infrared (LWIR) imagery. Traditional visual SLAM methods often struggle in environments characterized by poor visibility or substantial illumination changes, such as those found at night or in adverse weather conditions. Keil et al. (2024) address these challenges by proposing a novel approach that integrates learned feature descriptors into existing Bag of Words (BoW) localization frameworks.

The authors highlight that thermal imagery experiences dramatic temperature-driven appearance changes from day to night, which complicates feature extraction and recolonization tasks within SLAM systems. They reveal that conventional feature descriptors such as ORB are not robust enough to maintain consistency across these diurnal variations, leading to failures in place recognition.

To mitigate these issues, the study employs Gluestick—a learning-based method for feature extraction and matching—to enhance robustness against illumination changes. The authors introduce an extensive dataset collected with FLIR Boson thermal cameras over 24-hour cycles, providing valuable resources for training and testing their proposed methodology.

The results show that their integrated approach significantly improves local tracking performance and successfully enables relocalization between day and night imagery—achieving high recall rates compared to traditional methods like ORB-SLAM3. This work represents a significant step towards achieving all-day autonomy for robotic systems utilizing LWIR cameras, ultimately paving the way for more effective long-term mapping solutions in diverse environments.

2. Simultaneous Localization and Mapping (SLAM)

Mapping is a crucial mechanism employed by mobile robots to create maps of the locations they operate in. These maps are utilized to determine their relative location within the environment, facilitating effective path planning (localization). The creation of Extended Kalman Filter SLAM (EKF-SLAM) technology is where SLAM got its start, as described by Leonard and Durrant-Whyte (1991). Initial efforts in integrating SLAM involved a wide range of sensors, such as LIDAR, ultrasonic, inertial sensors, and GPS. Montemerlo (2002) proposed FastSLAM, a hybrid approach that combines Particle Filter and Extended Kalman Filter algorithms. Later, the same team introduced an improved version called FastSLAM 2.0, as explored in the research by Montemerlo et al. (2003). Dellaert and Kaess (2006) made a significant contribution to the field by introducing Square Root Smoothing and Mapping is a crucial mechanism employed by mobile robots to create maps of the locations they operate in. These maps are utilized to determine their relative location within the environment, facilitating effective path planning (localization). The creation of Extended Kalman Filter SLAM (EKF-SLAM) technology is where SLAM got its start. Initial efforts in integrating SLAM involved a wide range of sensors, such as LIDAR, ultrasonic, inertial sensors, and GPS. Montemerlo (2002) proposed FastSLAM, a hybrid approach that combines Particle Filter and Extended Kalman Filter algorithms. Later, the same team introduced an improved version called FastSLAM 2.0, as explored in the research by Montemerlo et al. (2003). Dellaert and Kaess (2006) made a significant contribution to the field by introducing Square Root Smoothing and Mapping (SAM), a technique that uses square root smoothing to improve the optimization of the SLAM problem, resulting in improved mapping efficiency. Kim et al. (2007) introduced a method called Unscented FastSLAM (UFastSLAM) that utilizes unscented transformation. This method has been shown to have improved resilience and accuracy when compared to FastSLAM 2.0. The field of SLAM has recently experienced a significant increase in interest in camera-based systems to decrease the weight and complexity of the systems. Camera-based SLAM systems that rely exclusively on visual input are usually known as visual SLAM (vSLAM), as noted by Taketomi et al. (2017). Significantly, the field of literature has witnessed the development of low-computation methods specifically designed for unmanned aerial vehicles (UAVs) that have limited onboard resources. Among these methods, visual Odometry stands out as a common Simultaneous Localization and Mapping (SLAM) application for small UAVs. Initial inquiries examined the incorporation of both Long Wave Infrared (LWIR) and visible spectra to tackle difficulties presented by environmental elements like fog or smoke. Maddern and

Vidas (2012) introduced a technique for integrating 8-bit thermal and RGB images to facilitate UAV navigation. Their research uncovered notable daily fluctuations in the visible spectrum, which differed from the thermal spectrum's steady but less pronounced features. The integration of the combined spectrum resulted in enhanced outcomes compared to algorithms that rely exclusively on visual or thermal frames. Poujol et al. (2016) showed that significant improvements in performance can be achieved by integrating visual and thermal spectra in traditional visual Odometry methods. Their research utilized two image fusion methods: monochrome threshold-based image fusion (Rasmussen et al., 2009) and monocular visual Odometry, as explored in the research by Geiger et al. (2011). These techniques were applied to data gathered from an electric car navigating through an urban setting. The experimental results highlighted the enhanced value of fused images, which resulted in more resilient solutions. In a preliminary assessment, Brunner and Peynot (2010) examined the integration of optical and thermal cameras for determining position in settings filled with smoke or dust, with a specific emphasis on autonomous ground vehicles (AGVs). Although visual scans were not sufficient for determining relative movements, thermal imaging was able to offer estimates, albeit with less precision. Additionally, Brunner et al. (2013) suggested a technique to improve a Visual SLAM (VSLAM) algorithm by integrating LWIR and normal spectra. The SLAM architecture described by Chen et al. (2017) utilizes both thermal and visual data to create a color map of situations with low lighting conditions. One of the most widely approaches is Graph-based SLAM that used for Simultaneous Localization and Mapping (Grisetti et al., 2010), where the robot's trajectory and the environment map are represented as a graph as shown in Saputra et al (2020). In this method, robot poses are modeled as nodes, and the spatial relationships between them are represented as edges based on sensor measurements and motion constraints. Optimization techniques are applied to adjust the node positions, minimizing errors and refining both the robot's path and the map. This approach effectively handles large-scale environments and loop closures, ensuring accurate, consistent mapping and localization (Hoshi et al., 2024). In the study by Mouats et al. (2014), a method called multispectral stereo Odometry was presented. This method combines optical and thermal sensors and is specifically designed for ground vehicles. Khattak et al. (2019a) combined radiometric sensors, the FLIR Tau2, and a visual camera to facilitate the movement of a tiny quadrotor in poorly illuminated indoor situations contaminated with dust. The use of an Intel NUC-i7 computer for onboard processing, along with thermal frames, allowed for effective feature selection and Extended Kalman Filtering to estimate drone Odometry. This ensured consistent performance even in situations with poor visibility.

SLAM is vital for autonomous navigation and exploration, but current solutions struggle with performance consistency due to a lack of diverse, high-quality datasets and robust evaluation metrics. One major issue is the absence of high-quality datasets that cover diverse all-weather conditions and provide a reliable metric for assessing robustness. This limitation significantly restricts the scalability and generalizability of SLAM technologies, impacting their development, validation, and deployment. To address this Zhao et al. (2023), introduce SubT-MRS, a challenging dataset designed to enhance SLAM in all-weather environments. It includes over 30 diverse scenes, multimodal sensors (LiDAR, cameras, IMU, thermal), and various robot locomotion (aerial, legged, wheeled). We also propose new robustness metrics, offering valuable insights for advancing SLAM research.

The SLAM algorithm operates in an iterative fashion, where the robot continuously updates its position estimate and refines the map based on new sensor readings and previously acquired information. The overall architecture of the SLAM algorithm is shown in Figure 1, Tourani et al (2022). Building upon the foundational concepts of SLAM, we now move to the specific algorithmic procedures that enable simultaneous localization and mapping.

2.1. SLAM Specific Algorithmic Procedures Initialization

Sensor Data Acquisition: The robot gathers sensory information from its surroundings using sensors such as cameras, LiDAR, or sonar.

Initial Pose Estimation: The robot's initial position and orientation are estimated based on sensor data and prior knowledge.

2.1.1. Motion Prediction

Odometry: The robot estimates its current position and orientation based on its previous position and the movement commands it has executed.

Motion Model: A mathematical model is used to predict the robot's motion and account for potential errors in its movement.

2.1.2. Observation Update

Feature Extraction: Distinctive features are extracted from the sensor data, such as edges, corners, or planes.

Feature Matching: Corresponding features are identified between consecutive sensor readings to track changes in the environment.

Map Update: The robot updates its map of the environment based on the observed features and their locations relative to the robot's estimated position.

2.1.3. Data Association

Feature Tracking: The robot tracks the features it has previously observed to maintain consistency in the map.

Loop Closure Detection: If the robot revisits a previously mapped area, it detects and corrects accumulated errors in its position estimates.

2.1.4. Map Optimization

Pose Graph Optimization: The robot refines its position estimates and the map using optimization techniques to minimize errors and inconsistencies.

Bundle Adjustment: A more advanced optimization technique that jointly refines the robot's poses and the positions of landmarks in the map.

2.1.5. Output

Localization: The robot's current position and orientation are determined.

Mapping: A map of the environment is generated, which can be used for navigation, path planning, and other tasks.

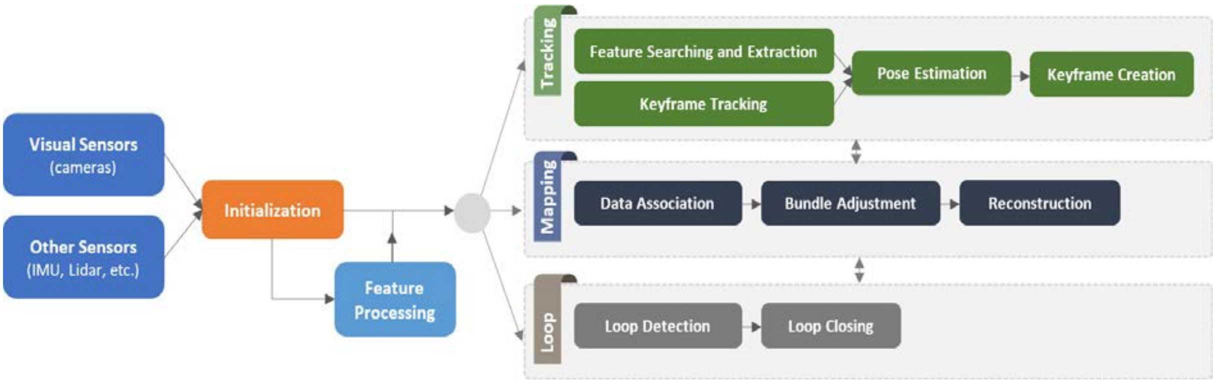

Figure 1. The flowchart of a standard visual SLAM approach, Tourani et al (2022).

3. Thermal SLAM

Thermal SLAM is an extension of traditional SLAM that uses thermal images or infrared sensors for both localization and mapping as shown in Figure 2, by Shin and Kim (2019) making it especially effective in environments where visibility is poor, such as smoke-filled areas, complete darkness, or obstacles that block traditional optical sensors. While Thermal SLAM leverages different sensor modalities, it is fundamentally based on the same core principles as conventional SLAM as we mentioned before:

Localization: This refers to determining the robot's position and orientation in the environment. In SLAM, localization is achieved by comparing the sensor data (in this case, thermal data) with the evolving map being built. The robot's location is continuously updated as it moves through the environment, helping it maintain an accurate representation of its surroundings.

Feature Mapping: In both traditional and thermal SLAM, the process of mapping involves detecting and tracking key features in the environment, such as temperature gradients, heat sources, or other thermal signatures. These features—often distinct in thermal images—are used to create and refine a 2D or 3D map, allowing the robot to both understand and navigate the environment.

Pose Estimation: Pose estimation involves calculating the robot's position and orientation within the map. This is critical for localization as it enables the robot to track its movements in space, updating its pose in real time as it encounters new thermal features. Accurate pose estimation ensures that the robot can navigate and maintain awareness of its surroundings, even in challenging conditions where optical data may be insufficient.

By relying on thermal data, Thermal SLAM overcomes the limitations of conventional SLAM in low-visibility environments while still maintaining the essential framework of localization, mapping, and pose estimation. Thermal SLAM is a recently developed field of research, with most discoveries published in the past ten years.

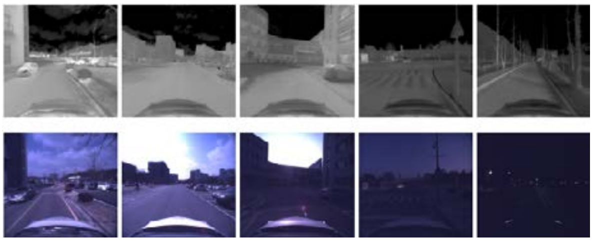

Figure 2. Set of Thermal-infrared (firstly) and RGB images (secondly), in contrast, RGB images, thermal images can capture overlook the time of captured (day or night) Shin and Kim (2019).

Thermal SLAM Research primarily concentrates on thermal Odometry, which involves using thermal cameras to calculate vehicle Odometry. However, thermal mapping, specifically the representation of thermal data in three dimensions, often requires overlaying thermal images onto point clouds created from other depth sources. A thermal-infrared SLAM system was first presented by Shin and Kim (2019) in Figure 3 who used full radiometric 14-bit raw data and LIDAR observations to estimate motion in six Degrees of Freedom (DoF). The experimental results demonstrated the advantages of the 14-bit system, overcoming restrictions related to the re-scaling procedure and showing increased resistance to data loss. Later, Khattak et al. (2020) suggested a thermal/inertial system that uses complete radiometric data to estimate Odometry. They emphasized that this system is resistant to data loss caused by sudden changes resulting from Automatic Gain Control (AGC) re- scaling. Although previous research has shown positive outcomes, SLAM methods still require significant processing resources and often necessitate the use of high-resolution thermal images. Several of the previously mentioned studies depend on high-resolution thermal cameras like the FLIR Tau2, which involve significant expenses. Additionally, the incorporation of a small yet robust onboard computer system contributes to the financial costs, as well as the need to take into account factors like space, weight, and power usage.

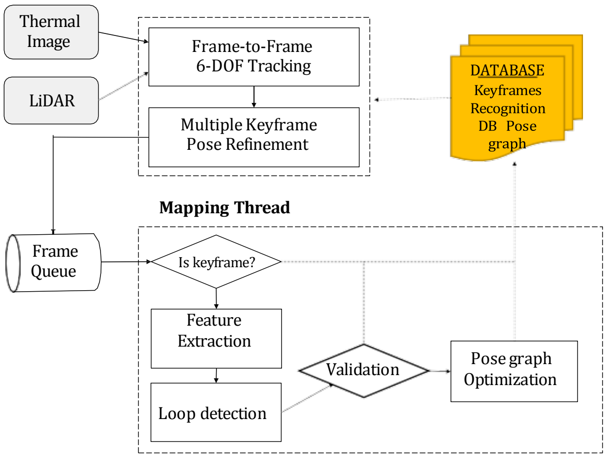

Figure 3. Thermal-Infrared SLAM Using Sparse Depth Information, Shin and Kim (2019).

3.1. Thermal Feature Extraction and Matching

Johansson et al. (2016) assessed how well several visual-based feature detectors and descriptors performed on thermal images using a dataset of image pairs taken in various textural and structural contexts. The dataset included pairs of images captured in different structural and textured environments. The images were initially resized using an 8-bit method, followed by histogram equalization prior to evaluation. The methods were tested under various image deformations, including changes in viewpoint and noise. The evaluation criteria were determined by considering recall, which measures the proportion of correct feature matches identified by RANSAC out of the total number of matches. The combination of the Hessian-affine extractor with the LIOP descriptor consistently demonstrated robust performance in various image deformations among floating-point descriptors. The SURF extractor and descriptor exhibited excellent performance, particularly in terms of their robustness against Gaussian blur and Gaussian white noise. Binary descriptors, when used in thermal imaging, provided performance comparable to floating-point descriptors, albeit with reduced computational expense. Combining binary descriptors like ORB with FREAK or BRISK produced competitive outcomes. Mouats et al. (2018) conducted another benchmark that specifically examined feature extraction and description algorithms for thermal features. This benchmark considered processing time, as it aimed at predicting UAV Odometry using thermal images. The datasets included video sequences captured in safe indoor and outdoor environments with a FLIR Tau2 LWIR camera. Overall, blob detectors like SIFT and SURF showed lower repeatability than corner detectors such as FAST and GFTT. Nevertheless, the characteristics of the blob displayed a more pronounced uniqueness, leading to elevated matching scores. The performance of SIFT feature extraction was subpar in various aspects. Surprisingly, the presence of motion blur had only a modest impact on feature extractors, and the SURF blob extractor had the highest performance. Although descriptors were relatively unaffected by Non-Uniformity Correction (NUC) as expected, it was observed that noise visibility was amplified in indoor situations with uniform temperatures. The authors suggested employing the SURF feature extraction technique in conjunction with the FREAK binary descriptor for thermal navigation applications. This combination provides a well-balanced compromise between match ability, repeatability, and real-time computing. The benchmarks were performed using datasets collected in controlled situations with consistent temperatures. However, environments with smoke or severe heat sources might have a major impact on the processes of extracting and describing features. Moreover, achieving a homogeneous dispersion of characteristics is essential for precise Odometry estimation. It is worth mentioning that the benchmark developed by Mouats et al. (2018) does not include a measurement for assessing the distribution of features across frames, which is crucial for accurately estimating thermal Odometry.

3.2. Thermal Image Processing

Over the past few years, techniques in image processing have been developed to address pattern noise in thermal imagery (Lu, 2020; Chen et al., 2021). Nevertheless, several conventional methods for reducing noise in thermal images fail to acknowledge the significant amount of time required for processing and encounter difficulties in successfully addressing the fundamental issues related to the measurement of light intensity in resized thermal images. To reduce the effects of Automatic Gain Control (AGC) operations on photometric changes, Mouats et al. (2015) suggested modifying the AGC threshold to minimize fluctuations in illumination. However, the core problem of photometric change remains unresolved, as this strategy merely postpones the eventual alteration in illumination. Brunner and Peynot (2010) and Vidas and Sridharan (2012) proposed a technique for histogram normalization of 14-bit thermal images to provide consistent illumination. In contrast, Papachristos et al. (2018) applied fixed interval re-scaling in familiar settings. Nevertheless, both approaches strongly depend on temperature priors that are peculiar to specific conditions, which restricts their practical usefulness. Recently, the use of deep neural networks in thermal image processing approaches has demonstrated potential. However, there is a lack of typical image processing methods specifically designed to handle noise and AGC concerns simultaneously in rescaled low-contrast thermal images.

3.3. Feature Association on Thermal Images

Traditional feature-matching techniques, such as BRIEF (Calonder et al., 2010; Zhao et al., 2020), have proven useful in odometry and SLAM (Simultaneous Localization and Mapping) studies. Lucas and Kanade (1981) assessed various feature descriptors on rescaled images using thermal properties, which showed inferior matching ability compared to visible images. The difference in results might be ascribed to intrinsic challenges in thermal image quality, such as noise. Learning-based feature association methods have become increasingly popular in certain SLAM systems, demonstrating comparable performance to traditional techniques and often surpassing them in difficult circumstances. GCNv2 and DXS (Mouats et al, 2018; Tang et al., 2019) employ deep neural networks (DNNs) to identify features, extract them, and generate descriptors, with trials demonstrating improved resilience in environments with less texture. Zhao et al. (2020) created a method for monocular visual odometry by utilizing deep optical flow to compare consecutive frames. It is important to emphasize that the methods stated above are mostly intended for images visible to the naked eye and may not be directly suitable for infrared photographs. In the context of Super Thermal, Lu and Lu (2021) introduced a DNN that can find characteristics and calculate descriptors simultaneously. This study showcased the benefits of using a deep model for matching thermal features. However, this network has not been investigated within a comprehensive SLAM system, and its computational expense has not been discussed.

3.4. Thermal Odometry

The process of obtaining a high-accuracy estimate of rotational speed in SLAM systems that rely on thermal imaging is challenging because these cameras collect heat distribution instead of visual appearance and geometry. However, there have been attempts to create thermal Odometry, though these have mainly focused on short distances or have not achieved the same level of performance as RGB-based Odometry. Mouats et al. (2015) employed a Fast-Hessian feature extractor to compute the distance for thermal Odometry in UAV tracking, while Borges and Vidas (2016) developed a practical thermal Odometry system that includes an automated technique for scheduling the Nonuniformity Correction (NUC) operation. However, this system is limited to outdoor situations because it relies on road lane estimation to calculate scale. Recent progress in thermal Odometry has been made by combining it with other types of sensors. Delaune et al. (2019) used thermal and inertial sensors to track UAVs employing an Extended Kalman Filter (EKF) algorithm. They showed that this approach was effective in different illumination conditions by using FAST and KLT trackers. Similarly, Khattak et al. (2019a) devised a thermal-inertial Odometry system for monitoring UAVs. They utilized a direct technique based on key-frames to reduce radiometric errors between consecutive frames using raw radiometric data. Although there have been advancements in thermal-inertial Odometry, there is currently no published research on thermal-inertial SLAM. Vidas and Sridharan (2012) developed a portable thermal SLAM system that uses FAST-based feature tracking and bundle adjustment-based optimization. Nevertheless, it lacks a loop closure module, which is crucial in modern SLAM frameworks for producing coherent trajectories and maps. Moreover, the absence of a sensor not influenced by the environment, such as an Inertial Measurement Unit (IMU), complicates accurate estimation in diverse contexts. Shin and Kim (2019) introduced a feature-based LIDAR-thermal SLAM method that combines thermal data with sparse range measurements from LIDAR to enhance scale estimates. However, using their system in an autonomous driving environment, which often involves greater temperature variations than indoor settings, raises questions about its universal applicability.

3.5. Thermal Mapping

Vidas and Sridharan (2012) created a monocular SLAM system specifically designed for thermal cameras that are carried by hand. The program identifies corner features using the GFTT and FAST detectors and then tracks them using sparse Lucas-Kanade optical flow. To analyze the images with thermal properties, the original 14-bit thermal data is transformed into 8-bit data within a predetermined range centered on the average value between the lowest and maximum 14-bit intensity. Homography motion estimation utilizes matched SURF features to continue tracking following a Flat Field Correction (FFC). Local path refinements are conducted on every frame between two key-frames, and a new metric is introduced based on five factors to determine the selection of a new key-frame. Nevertheless, the technique frequently falls short of achieving a re-projection error below 1.5 for extended durations, especially during pure rotations. The authors did not disclose information about the accuracy of the resulting 3D thermal map. Shin and Kim (2019) integrated depth data from LiDAR and thermal data from a LWIR camera to generate a thermal map. A limited number of LiDAR points are transformed and mapped onto the 14-bit thermal image. These points are then monitored and traced in the next frame using a direct method. The loop closure, extracted through the bag of visual words technique, involves obtaining ORB features from the 8-bit transformed image. However, this method may lead to incorrect detections in settings with limited intricate patterns. To address this problem, supplementary geometric verification is introduced. The authors assert that they have created a visual SLAM system capable of accurately calculating a trajectory and generating a 3D thermal map throughout the day, regardless of outdoor illumination conditions. (Van Manen et al., 2023) developed Chameleon, a stereo thermal-inertial SLAM system based on EKF-SLAM. Pose estimation depends on tracking up to thirty landmarks retrieved via SIFT feature extraction and description. Experiments conducted in both cold and heated conditions showed that relying solely on thermal-inertial data for localization is not feasible due to the lack of distinct differences in thermal images, especially in cold environments. While no specific margin of error is provided, the algorithm claims to approximate location within a few meters of a firefighting training facility where an active fire is present. Nevertheless, the limited mapping method cannot generate a clear and understandable representation of the building, and moisture formation on the lens in hot conditions negatively impacts the accuracy of determining position, resulting in a high dependence on inertial data. When only a thermal map is needed, it is typically preferable to overlay thermal data onto a 3D map created by a SLAM technique utilizing more widely used sensors. This is accomplished by acquiring the position and depth information via sensors like LiDAR, RGB cameras, or depth cameras. Subsequently, thermal data is incorporated into the 3D map by projecting map points onto thermal images. However, in environments filled with smoke or under fire conditions, secondary sensors may lose effectiveness, negatively impacting the accuracy and reliability of thermal mapping algorithms.

4. Thermal Inertial-SLAM Method

Visual SLAM's perform better in good weather conditions. However, in low visibility, foggy, stormy, blizzardy, or dark environments with poor visibility, such as darkness at night, enclosed environments, and tunnels, Visual SLAM's have difficulty reading the surrounding data due to their equipment (usually LIDAR).

TI-SLAM: Thermal-Inertial SLAM with Probabilistic Neural Networks for Adverse Conditions Saputra et al. (2021) introduced an approach called TI-SLAM that integrates neural network-based sensor abstraction with probabilistic pose graph optimization to improve pose accuracy in visibility-limited situations, such as darkness, smoke, or dust.

Neural Sensor Abstraction: The front-end uses Mixture Density Networks (MDN) to convert raw thermal and inertial data into a probabilistic format, optimizing the interpretation process for more accurate mapping and localization. TI- SLAM's neural network uses a ResNet 50 model to transform raw thermal data into distinctive 128-dimensional global descriptors. This structure is adapted for processing thermal images, which typically contain less feature variation than RGB images. For training, the network exploits triple loss with samples obtained from the Bag of Traced Words (BoTW) algorithm applied to RGB images. This method enables efficient training by focusing on generating high-quality outputs.

Robust Pose Graph Optimization: TI-SLAM is designed to perform under difficult conditions. This is achieved with thermal imaging and guidance that captures infrared radiation, which is unaffected by visible light conditions, as explored in the research by Khattak et al. (2019a). Thermal-inertial SLAM offers advantages compared to traditional SLAM methods due to its use of thermal and inertial data, robustness to environmental changes, reduced dependency on feature subtraction, robust predictions using Probabilistic Neural Networks (PNNs), graph-based optimization for consistent matching, and adaptability to different sensor configurations. Thermal and inertial data complement visual information, providing robustness to illumination changes and accurate motion estimation. Unlike traditional SLAM, it relies less on feature extraction algorithms using differential thermal signatures for matching. PNNs provide robust predictions by accounting for uncertainties in sensor measurements and motion estimates. Graph-based optimization reduces variance by providing globally consistent maps and trajectories. Furthermore, its modular nature allows for easy integration of additional sensors or adjustments to the sensor setup, increasing adaptability to various robotic platforms and application requirements.

Since the TI-SLAM algorithm is older compared to today's algorithms, we had to use specific versions of the additional packages needed for this algorithm to function, or versions of the packages that are still in use today. When we updated some packages, we encountered different errors in the algorithm, making the process more challenging than anticipated. Once we managed to run the TI-SLAM algorithm, the second major issue was the large size of the datasets used. The algorithm was able to reduce this dataset from 14 bits to 8 bits, but even this was taking up too much space in our virtual machine. Moreover, the size of these datasets caused the SLAM algorithm to generate output for an extended period. Additionally, to place these datasets in the correct file locations and provide them as input to the algorithm, the configuration file had to be configured separately for each dataset. Constantly updating the contents of this configuration file for each dataset was also a time-consuming problem for us. The last and biggest challenge we faced while running and testing this algorithm was that the machines we used were not equipped with sufficient hardware, such as RAM (Random Access Memory), required for large datasets. Although we ran this algorithm on our virtual machine with a high RAM size of 16 GB, we were unable to obtain output for many datasets. Due to the algorithm's large number of requirements for these datasets, our virtual machines were unable to meet this need, resulting in a "Process Killed" response for most of the large datasets. Despite all these problems, we tested our TI-SLAM algorithm on some datasets and achieved successful results. Thermal and inertial data complement visual information, providing robustness to illumination changes and accurate motion estimation. Unlike traditional SLAM, it relies less on feature extraction algorithms, using differential thermal signatures for matching. Probabilistic Neural Networks provide robust predictions by accounting for uncertainties in sensor measurements and motion estimates. Graph-based optimization reduces variance by providing globally consistent maps and trajectories. Furthermore, its modular nature allows for easy integration of additional sensors or adjustments to the sensor setup, increasing adaptability to various robotic platforms and application requirements.

In summary, Thermal Inertial SLAM stands out for its ability to use thermal and inertial data, robust prediction using PNNs, and graph-based optimization, providing a robust and adaptable solution for mapping and positioning tasks in harsh environments.

5. Conclusion

Thermal SLAM, a new field, may improve object detection and tracking in challenging situations (Taketomi et al., 2017). Thermal SLAM leverages temperature fluctuations for robust operation in low light, adverse weather, and occlusions (Brunner and Peynot, 2014). Although Thermal SLAM seems promising, the low resolution of thermal cameras, signal-to-noise ratio deterioration, and low contrast in thermal images present distinct obstacles (Delaune et al., 2019). To overcome these constraints, the implementation of Thermal SLAM requires specialized feature extraction, matching, and image processing (Wang et al., 2017). For short-range applications, thermal Odometry has seen improvements (Lucas and Kanade, 1981; Lu and Lu, 2021). Thermal-inertial Odometry enhances performance (Borges and Vidas, 2016; Delaune et al., 2019), however, a complete thermal-inertial SLAM system with robust loop closure is still not achievable. Fusing thermal data with 3D maps from LiDAR is common in thermal mapping. While effective, this strategy may be limited in scenarios involving smoke or fire damage to secondary sensors (Maddern and Vidas, 2012).

Future studies should focus on developing processing methods for low-resolution, low-contrast thermal imagery, including feature extraction and matching techniques. Deep learning methods for Thermal SLAM are promising, potential for robust and dependable object recognition and tracking in complicated circumstances. In conclusion, Thermal SLAM could transform autonomous systems in visually challenged contexts (Papachristos et al., 2018). Additionally, approaches for thermal mapping that are independent of secondary sensors for smoke-filled environments should be investigated (Taketomi et al., 2017). Thermal SLAM has the potential to revolutionize autonomous navigation, search and rescue, and environmental monitoring by overcoming existing challenges and exploring new research areas. The future of Thermal SLAM is bright, and continuous research will unveil its full potential. Finally, Table 1 shows a summary of the works reviewed in this article.

Table 1. A summary of the works (methods)reviewed in this article

Method	Year	Authors	Limitations	Key Features
Thermal-Infrared SLAM	2019	(Shin and Kim, 2019)	Computationally expensive	Requires high-resolution thermal cameras and powerful computers.
Thermal/Inertial SLAM	2019	(Khattak et al., 2019a)	Computationally expensive	Uses full radiometric data for odometry.
Monocular thermal SLAM	2020	(Bijelic et al., 2020)	Lacks loop closure for consistent trajectories	Not suitable for large-scale environments or globally consistent maps.
Feature-based LiDAR-thermal SLAM.	2019	(Shin and Kim, 2019)	Requires additional LiDAR sensor, which may not be suitable for all environments.	Enhances thermal data with sparse range measurements from LiDAR for scale estimation.
Chameleon (Stereo thermal-inertial SLAM).	2023	(van Manen et al., 2023)	Limited effectiveness in low- contrast environments.	Relies heavily on inertial data, and potentially inaccurate thermal maps.
Visual and thermal Sensor Fusion (VSLAM+Thermal).	2012	(Maddern and Vidas, 2012)	Requires robust image fusion algorithms, and may be computationally expensive.	Combines visual and thermal spectra for navigation.
Multi-spectral Stereo Odometry.	2016	(Borges and Vidas, 2016)	Integrates optical and thermal sensors for ground vehicles.	Requires stereo camera setup (thermal and visual).
Thermal Radiometric SLAM with NUC.	2020	(Khattak et al., 2020)	Requires high-resolution thermal camera and powerful computer.	Integrates high-resolution thermal camera, powerful computer, and radiometric sensors.
Super thermal (Deep Learning).	2018	(Mouats et al., 2018)	Limited research on integration within a complete SLAM system, and computational cost needs evaluation.	Deep neural network for thermal feature matching.
Graph-Based Thermal- Inertial SLAM with Probabilistic Neural Networks.	2020	(Saputra et al., 2020)	May require significant computational resources for neural network operations.	Combines graph- based SLAM with probabilistic neural networks for enhancement. thermal-inertial navigation.
Graph-based SLAM.	2010	Grisetti et al., 2010)	Tutorial nature may limit depth in specific applications like thermal-inertial SLAM.	Provides a comprehensive overview of graph-based SLAM, a foundational technique in robotics
Only Look Once, Mining Distinctive Landmarks from ConvNet for Visual Place Recognition.	2017	(Chen et al., 2017)	Focused on visual place recognition, may not directly address thermal-inertial SLAM challenges.	Introduces a method extract distinctive landmarks from Convolutional Neural Networks for place recognition.
SubT-MRS Dataset: Pushing SLAM Towards All weather Environments.	2023	(Zhao et al., 2023)	Dataset-focused, may not provide a complete SLAM system.	Presents a dataset designed to challenge and improve SLAM systems in all-weather environments, potentially including thermal data.

Author Contributions

The percentages of the authors' contributions are presented below. All authors reviewed and approved the final version of the manuscript.

	F.S.	O.S.G.
C	50	50
D	50	50
S	50	50
L	50	50
W	50	50
CR	50	50
SR	50	50

C=Concept, D= design, S= supervision, L= literature search, W= writing, CR= critical review, SR= submission and revision.

Conflict of Interest

The authors declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because of there was no study on animals or humans.

References

Bijelic M, Gruber T, Mannan F, Kraus F, Ritter W, Dietmayer K, Heide F. 2020. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 14-19, Seattle, WA, USA, pp: 11682-11692.

Borges PVK, Vidas S. 2016. Practical infrared visual odometry. IEEE Trans Intell Transp Syst, 17(8): 2205-2213.

Brunner C, Peynot T. 2010. Visual metrics for the evaluation of sensor data quality in outdoor perception. In: Proceedings of the 10th Performance Metrics for Intelligent Systems Workshop, September 28-30, Baltimore, MD, USA, pp: 55-61.

Brunner C, Peynot T. 2014. Perception quality evaluation with visual and infrared cameras in challenging environmental conditions. In: Experimental Robotics: The 12th International Symposium on Experimental Robotics, March 17- 20, Tokyo, Japan, pp: 231-240.

Brunner C, Peynot T, Vidal-Calleja T, Underwood J. 2013. Selective combination of visual and thermal imaging for resilient localization in adverse conditions: Day and night, smoke and fire. J Field Robot, 30(4): 641-666.

Cadena C, Carlone L, Carrillo H, Latif Y, Scaramuzza D, Neira J, Leonard JJ. 2016. Past, present, and future of Simultaneous Localization and Mapping: Toward the robust-perception age. IEEE Trans Robot, 32(6): 1309-1332.

Calonder M, Lepetit V, Strecha C, Fua P. 2010. Brief: Binary robust independent elementary features. In: Proceedings of the Computer Vision– ECCV 2010:11th European Conference on Computer Vision. Springer- Verlag, September 5- 11, Heraklion, Crete, Greece, pp: 778-792.

Chen X, Liu L, Zhang J, Shao W. 2021. Infrared image denoising based on the variance-stabilizing transform and the dual-domain filter. Digit Signal Process, pp: 113-128.

Chen Z, Maffra F, Chli M. 2017. Only look once, mining distinctive landmarks from convnet for visual place recognition. In: Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), September 24-28, Vancouver, BC, Canada, pp: 6435-6441.

Debeunne C, Vivet D. 2020. A review of Visual-LiDAR fusion based Simultaneous Localization and Mapping. Sensors, 20(7): 20-68.

Delaune J, Hewitt R, Lytle L, Sorice C, Thakker R, Matthies L. 2019. Thermal-inertial odometry for autonomous flight throughout the night. In: Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), November 4-8, Macau, China, pp: 1090- 1097.

Dellaert F, Kaess M. 2006. Square Root SLAM: Simultaneous Localization and Mapping via square root information smoothing. Int J Robot Res, 25(12): 1181-1203.

Geiger A, Ziegler J, Stiller C. 2011. Stereoscan: Dense 3D reconstruction in real-time. In: Proceedings of the 2011 IEEE Intelligent Vehicles Symposium (IV), June 5-9, Baden-Baden, Germany, pp: 963-968.

Grisetti G, Kümmerle R, Stachniss C, Burgard W. 2010. A tutorial on graph-based SLAM. IEEE Intell Transp Syst Mag, 2(4): 31-43.

Hoshi M, Hara Y, Nakamura S. 2024. Graph-based SLAM using wall detection and floor plan constraints without loop closure. ROBOMECH J, 11(1), 18-32.

Johansson J, Solli M, Maki A. 2016. An evaluation of local feature detectors and descriptors for infrared images. In: Proceedings of the Computer Vision–ECCV 2016 Workshops: October 8-10, Amsterdam, Netherlands, Proceedings, Part III, pp: 142.

Keil C, Gupta A, Kaveti P, Singh H. 2024. Towards Long-Term SLAM on Thermal Imagery. arXiv preprint arXiv: 2403.19885.

Khattak S, Papachristos C, Alexis K. 2019a. Keyframe-based direct thermal–inertial odometry. In: Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), May 20-24, Montreal, QC, Canada, pp: 6197-6203.

Khattak S, Papachristos C, Alexis K. 2019b. Visual- thermal landmarks and inertial fusion for navigation in degraded visual environments. In: Proceedings of the 2019 IEEE Aerospace Conference, March 2-9, Big Sky, MT, USA, pp: 1-9.

Khattak S, Papachristos C, Alexis K. 2020. Keyframe- based thermal–inertial odometry. J Field Robot, 37(4): 552-579.

Kim C, Sakthivel R, Chung WK. 2007. Unscented FastSLAM: A robust algorithm for the simulataneous Localization and Mapping problem. In: Proceedings of the 2007 IEEE International Conference on Robotics and Automation, April 10-14, Rome, Italy, pp: 2439-2445.

Leonard JJ, Durrant-Whyte HF. 1991. Mobile robot localization by tracking geometric beacons. In: Proceedings of the 1991 IEEE International Conference on Robotics and Automation, April 9- 11, Sacramento, CA, USA, pp: 376-382.

Li J, Li Z, Feng Y, Liu Y, Shi G. 2019. Development of a human–robot hybrid intelligent system based on brain teleoperation and deep learning SLAM. IEEE Trans Autom Sci Eng, 16(4): 1664-1674.

Lu C. 2020. Stripe non-uniformity correction of infrared images using parameter estimation. Infr Phys Tech, 107(1), pp: 1-7.

Lu Y, Lu G. 2021. Superthermal: Matching thermal as visible through thermal feature exploration. IEEE Robot Autom Lett, 6(2): 2690-2697.

Lucas BD, Kanade T. 1981. An iterative image registration technique with an application to stereo vision. In: Proceedings of the IJCAI'81: 7th International Joint Conference on Artificial Intelligence, August 24-28, Vancouver, BC, Canada, pp: 674-679.

Maddern W, and Vidas S. 2012. Towards robust night and day place recognition using visible and thermal imaging. In:

Proceedings of the RSS 2012 Workshop: Beyond laser and vision: Alternative sensing techniques for robotic perception, July 9, Sydney, Australia, pp: 1-8.

Montemerlo M. 2002. FastSLAM: A factored solution to the Simultaneous Localization and Mapping problem. In: Proceedings of AAAI02, July 28- August 1, Edmonton, Canada, pp: 593-598.

Montemerlo M, Thrun S, Koller D, Wegbreit B. 2003. FastSLAM 2.0: An improved particle filtering algorithm for Simultaneous Localization and Mapping that provably converges. In: Proceedings of the IJCAI, August 9-15, Acapulco, Mexico, pp: 1151-1156.

Mouats T, Aouf N, Chermak L, Richardson MA. 2015. Thermal stereo odometry for UAVs. IEEE Sens J, 15(11): 6335-6347.

Mouats T, Aouf N, Nam D, Vidas S. 2018. Performance evaluation of feature detectors and descriptors beyond the visible. J Intell Robot Syst, 92: 33-63.

Mouats T, Aouf N, Sappa AD, Aguilera C, Toledo R. 2014. Multispectral stereo odometry. IEEE Trans Intell Transp Syst, 16(3): 1210-1224.

Mur-Artal R, Montiel JMM, Tardos JD. 2015. ORB- SLAM: A versatile and accurate monocular SLAM system. IEEE Trans Robot, 31(5): 1147-1163.

Papachristos C, Mascarich F, Alexis K. 2018. Thermal-inertial localization for autonomous navigation of aerial robots through obscurants. In: Proceedings of the 2018 International Conference on Unmanned Aircraft Systems (ICUAS), June 12-15, Dallas, TX, USA, pp: 395-401.

Poujol J, Aguilera CA, Danos E, Vintimilla BX, Toledo R, Sappa AD. 2016. A visible-thermal fusion-based monocular visual odometry. In: Proceedings of Robot 2015 Second Iberian Robotics Conference: Advances in Robotics, November 19-21, Lisbon, Portugal, pp: 213-225.

Rasmussen N. D, Morse B. S, Goodrich M. A, Eggett D. 2009. Fused visible and infrared video for use in wilderness search and rescue. In: Proceedings of the 2009 Workshop on Applications of Computer Vision (WACV), December 7-8, Snowbird, UT, USA, pp: 1-8.

Saputra MRU, De Gusmao PP, Lu CX, Almalioglu Y, Rosa S, Chen C, Trigoni N. 2020. Deeptio: A deep thermal-inertial odometry with visual hallucination. IEEE Robot Autom Lett, 5(2): 1672- 1679.

Saputra MRU, Lu CX, de Gusmao PPB, Wang B, Markham A, Trigoni N. 2021. Graph-based thermal-inertial SLAM with probabilistic neural networks. IEEE Trans Robot, 38(3): 1875-1893.

Shin YS, Kim A. 2019. Sparse depth enhanced direct thermal-infrared SLAM beyond the visible spectrum. IEEE Robot Autom Lett, 4(3): 2918- 2925.

Taketomi T, Uchiyama H, Ikeda S. 2017. Visual SLAM algorithms: A survey from 2010 to 2016. IPSJ Trans Comput Vis Appl, 9: 1-11.

Tang J, Ericson L, Folkesson J, Jensfelt P. 2019. GCNv2: Efficient correspondence prediction for real-time SLAM. IEEE Robot Autom Lett, 4(4): 3505-3512.

Tourani A, Bavle H, Sanchez-Lopez JL, Voos H. 2022. Visual slam: What are the current trends and what to expect? Sensors, 22(23): 9297-9326.

van Manen BR, Sluiter V, Mersha AY. 2023. FirebotSLAM: Thermal SLAM to increase situational awareness in smoke-filled environments. Sensors, 23(17): 7611-7636.

Vidas S, Sridharan S. 2012. Hand-held monocular SLAM in thermal-infrared. In: Proceedings of the 2012 12th International Conference on Control Automation Robotics and Vision (ICARCV), December 5-7, Guangzhou, China, pp: 500-506.

Wang S, Clark R, Wen H, Trigoni N. 2017. Deepvo: Towards end-to-end visual odometry with deep recurrent convolutional neural networks. In: Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), May 29-June 3, Singapore, pp: 2043-2050.

Wu Y, Wang L, Zhang L, Bai Y, Cai Y, Wang S, Li Y. 2023. Improving autonomous detection in dynamic environments with robust monocular thermal SLAM system. ISPRS J Phot Remote Sens, 203: 265-284.

Zhang J, and Singh S. 2014. LOAM: Lidar odometry and mapping in real-time. In: Proceedings of Robotics: Science and Systems (RSS), July 12-16, Berkeley, CA, USA, pp: 109-117.

Zhao S, Singh D, Sun H, Jiang R, Gao Y, Wu T, Xu J. 2023. Subt-mrs: A subterranean, multi-robot, multi-spectral and multi-degraded dataset for robust SLAM. arXiv preprint arXiv:2307.07607.

Zhao S, Wang P, Zhang H, Fang Z, Scherer S. 2020. Tp-tio: A robust thermal-inertial odometry with deep thermal point. In: Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 25-29, Las Vegas, NV, USA, pp: 10321-10328.