IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 16, 2022, accepted April 7, 2022, date of publication April 14, 2022, date of current version April 22, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3167393

Drone-Enabled Multimodal Platform for Inspection of Industrial Components

PARHAM NOORALISHAHI[1], (Graduate Student Member, IEEE), FERNANDO LÓPEZ[2], (Member, IEEE), AND XAVIER P. V. MALDAGUE[1], (Senior Member, IEEE)

[1]Department of Electrical and Computer Engineering, Université Laval, Quebec, QC G1V 0A6, Canada
[2]TORNGATS Company, Quebec, QC G2E 5V9, Canada

Corresponding author: Parham Nooralishahi (parham.nooralishahi.1@ulaval.ca)

This work was supported in part by the Canada Research Chair in Multi-polar Infrared Vision (MiViM); in part by the Natural Sciences, and Engineering Research Council (NSERC) of Canada through a Discovery Grant; and in part by the ''oN Duty!'' NSERC Collaborative Research and Training Experience (CREATE) Program.

ABSTRACT In the rise of recent advancements in unmanned aerial vehicles, many studies have focused on using multi-modal platforms for remote inspection of industrial and construction sites. The acquisition of multiple data modalities assists the inspectors in acquiring comprehensive information about the targeted components. Despite the benefits of multi-modal platforms, the calibration and fusion of the obtained data modalities present many challenges that need to be addressed. Using a calibration board with geometrically known features to estimate intrinsic and extrinsic parameters and accurately align the images in thermal and visible spectral bands, is one of the main approaches to address the problem of dissimilarity of feature appearances in different spectrums.This study presents a comprehensive platform for drone-based multi-modal inspection of industrial and construction components, including three main components: 1) a sensor setup that can be used as a standalone system or a payload for a drone; 2) a multi-modal embedded system; and 3) a novel calibration board for multi-modal data fusion. The multi-modal embedded system provides the required features to record, transmit, and visualize the thermal, visible, and depth data synchronously. Additionally, the system presents a multi-modal fusion technique to form RGBD&T data containing thermal and texture information of the obtained 3D view. Moreover, this study introduces a novel self-heating calibration board that uses Thermoelectric Peltier modules to provide an identifiable and sharp pattern in thermal and visible images. The calibration board is designed with an aim also to be used as Ground Control Point (GCP) in drone surveys.

INDEX TERMS Multi-modal sensory platform, unmanned aerial vehicle, calibration board, data fusion, ground control point, thermography.

I. INTRODUCTION

During past decades, many studies investigated the use of coupled imagery sensors like visible, thermal, and depth sensors to monitor and inspect industrial and construction infrastructures, especially in Non-Destructive Inspection (NDI) applications. Different modalities represent different but often complementary types of information about the observed scene [1] which can assist companies in providing more reliable and comprehensive analysis.

The associate editor coordinating the review of this manuscript and approving it for publication was Abderrahmane Lakas.

Thanks to the recent advancements in sensing technologies and non-destructive testing methods, they provide more accurate and reliable measures. However, they also have limitations by nature. For instance, visible cameras can sense the color and textural information while not presenting any depth information and are vulnerable to low illumination conditions. Thermal cameras provide a visual presentation of thermal measurements related to the inspected scene. However, they are not capable of sensing colors.

Non-destructive testing methods also have limitations due to their intended applications. For instance, Computerized Tomography (CT) cannot distinguish between materials with similar attenuation factors [2]. Thermography is only

capable of detecting shallow or near-surface defects [2]. Also, the detection of defect's size and location is challenging in Ultrasonic testing (UT) due to the nonhomogeneous nature of the material [2].

One of the standard solutions to overcome these limitations is to use coupled sensory systems. Coupled sensors in a setup can reduce failures and ambiguities caused by sensor degradation and expand the obtained information from the environment by integrating multiple sensors. Multi-modal data acquisition strategies can be categorized into two categories: (a) collection of data modalities at different times (b) collection of data modalities with some degrees of sensor coupling and synchronization. The use of multiple data sessions or acquisition teams can lead to data inconsistency and difficulty in data management and alignment. However, multi-modal data acquisition can significantly reduce data inconsistency and process complexity.

The fusion of imagery data from different spectral bands can be instrumental in numerous applications such as industrial inspection, surveillance, monitoring, and enhanced vision. One of the essential steps in a multi-modal data fusion is an accurate and reliable spatial registration of the acquired images. Fusion techniques are often categorized based on the level of fusion of modalities [3]: (a) sub-pixel level, where the fusion is performed at a sub-pixel level using proper transformation [4], [5]; (b) pixel level, where corresponding pixels in the modalities are used for fusion purposes [6]; (c) feature level, involving extraction and selection of features that are identifiable in the modalities of interest; (d) decision level, where individual modalities are processed and reach a decision before optimally combine the decisions to yield a more robust and informed decision [7].

One of the main challenges in multi-modal registration is the dissimilarity of feature appearances in different spectral bands. Another challenge is that the obtained registration is only valid for the objects located in the two-dimensional plane covering the extracted feature points, which means that the obtained calibration configurations of multiple cameras are only valid for a specific target distance [8]. It becomes more critical in case of a low rate of discriminative features, which are commonly observed in thermal images. The most common approach to addressing these challenges is to employ a geometrically known calibration pattern identifiable in all sensory data.

One of the commonly used techniques for multi-modal data fusion is calibration-based registration. In this approach, a reference board is employed to (a) estimate the intrinsic and extrinsic parameters of the cameras, which can improve the accuracy of alignment, especially in case of the existence of radial and tangential distortions, and (b) calculate the transformation matrix for aligning images from different spectral bands [8]. The main challenge in using a calibration board is to design a board that can produce an identifiable pattern with sufficient contrast in all modalities of interest. This challenge can get more complicated in an outdoor environment where environmental conditions such as humidity, illumination, ambient temperature, wind, and other factors can significantly influence the calibration board. Moreover, specular reflection is one of the practical challenges within but not limited to passive calibration boards, especially for thermal cameras. Specular reflection, also known as mirror-like reflection, is where the reflectance of a specular surface is zero for all angles except the specular angle [9].

The processing of obtained images for image stitching or photogrammetry is a well-studied subject. Various studies addressed this challenge by employing feature matching, intensity-based registration, or other methods that predominantly suffer from alignment and registration errors and incremental drifting. One practical and widely used approach to address these issues and enhance the results, especially in drone-based applications, is Ground Control Points (GCPs). GCPs are essential tools in drone-based photogrammetry, survey, and Geographic Information System (GIS), especially when the high accuracy of absolute location is required. GCPs are ground points with known coordinates and patterns. They can be highly beneficial where the collected data is not corrected by Real-Time Kinematic (RTK) and Post-Process Kinematic (PPK) methods. GCPs should be distributed in a frequent interval all around the targeted area to ensure a GPS-grade accuracy [10]. Another area that GCPs are beneficial is when the task is to align existing terrestrial measurements where no transformation is available [10]. Also, they are essential where the number of feature points is limited to have accurate and reliable results, which are commonly observed in the drone-based thermal or multi-modal surveys and close-range inspection of industrial components.

Although Ground Control Points is a well-studied subject and many commercial solutions are available, the use of Ground Control Points for drone-based thermal and multi-modal surveys is still an open problem. In the case of a thermal survey, providing a GCP that can present a sharp and recognizable pattern is a challenge that needs to be addressed. Also, in a multi-modal survey, the employed GCP needs to generate similar and identifiable patterns in all spectral bands of interest.

This study introduces a drone-enabled multi-modal platform involving the software and hardware components for collection, transmission, recording, and fusion of multi-modal data for inspection of industrial sites. Additionally, the system presents a fusion process pipeline to generate a 3D view of the observed scene named RGBD&T data containing depth, thermal, and texture information. For camera calibration and multi-modal data fusion, a novel self-heating reference board is proposed that can generate similar patterns in thermal and visible images.

The structure of this paper is presented as follows. A brief review of related works is described in Section II. The calibration method is explained in Section III. In Section IV, the introduced multi-modal platform is described. Finally results are presented in Section V.

II. RELATED WORKS

A. MULTI-MODAL SYSTEMS FOR MONITORING AND INSPECTION

In recent years, the interest in using multiple data modalities in the non-destructive testing of industrial components has grown significantly. Combining multiple NDT techniques or sensors can improve inspection accuracy and enhance the results. Gros *et al.* investigated the use of a probe array composed of different sensors and installed on a Remotely Operated Vehicle (ROV) for inspecting off-shore platforms [11]. This study concluded that using multiple sensors provides a better assessment of defect characteristics than using individual sensors. Also, using different types of sensors can provide a more comprehensive analysis. For instance, Ultrasonic Testing (UT) is unsuitable for detecting shallow cracks, while Eddy Current (EC) can detect surface and near-surface defects. So the combination of them can assist in detecting surface and subsurface defects. Also, the use of Coordinate Measuring Machine (CMM) along with Computerized Tomography (CT) can provide the measurement in both the interior and the exterior [2]. Horn and Mayo discussed the benefits of combining UT and EC techniques for inspecting rail lines [12]. They discussed that the use of UT and EC together can improve reliability since EC provides information about the surface while UT can detect defects in the interior of the rail.

Simmen *et al.* investigated the use of multi-modal data acquisition and processing for non-destructive inspection of weld seams using passive thermography [13]. They used a visible camera to observe visible contamination like rust, moisture, and oil fractions, near-infrared, and Long Wave Infrared (LWIR) cameras to observe thermal patterns and find abnormalities. Li *et al.* presented a multi-sensory acquisition approach for inspecting Printed Circuit Boards (PCBs) using polarization and thermography techniques [14]. They stated that the use of polarization along with thermography assist in avoiding the false positives caused by uneven illumination distribution and reflection on the visible images.

Remote inspection of industrial sites using aerial platforms is a fast-growing area of interest for different industries. In the conventional inspection of remote or hard-to-access areas, transferring the equipment and human resources is sometimes challenging and impracticable in terms of cost, time, and risk. However, the lack of direct access can cause data misinterpretation during post-processing. One of the solutions is to use multi-modal data platforms that can observe different physical properties of the site. For example, in the case of road inspection, companies spend a considerable amount of cost and time to collect and evaluate data regarding road distress [15]. They often perform inspections using digital cameras installed on a mobile platform [16]. However, as stated by Cheng *et al.*, the dark areas in visible images caused by tire marks, oil spills, or shadows, other than distress, can lead to data misinterpretation, which can be solved by using multi-modal data [17]. Also, poor illumination and shadow can adversely affect the usability of visible images. Javidi *et al.* presented a multi-modal system to address this issue [18]. They used a 3D sensor along with a visible camera to inspect road pavement. Although they reported that the system had high sensitivity to vibrations caused by passing vehicles, they could significantly improve the results compared to conventional methods.

In the case of inspection in the electric power industry, multi-modal platforms are actively used by companies and various studies. Alsafasfeh *et al.* presented a drone-based multi-sensory system containing visible and thermal cameras for detecting potential faults in photovoltaic plants [19]. Lee *et al.* coupled thermal and visible cameras for inspection of solar panels [20]. They used visible images to identify solar panel arrays and thermal images to detect faulty panels by locating highlighted or darkened spots. Hydro-Québec introduced a drone-based multi-modal platform named LineDrone for inspecting power transmission lines [21]. They used a coupled LiDAR sensor and visible camera to estimate the drone location relative to the power line. Also, Wang *et al.* presented a multi-modal system containing thermal and visible cameras installed on an unmanned helicopter for inspecting power lines [22]. They employed RGB images to detect visible defects, while thermal images were used to identify thermal faults.

One of the most studied fields of interest in multi-modal systems is the use of thermal and visible cameras as complementary sensors for industrial inspections. Liu *et al.* introduced a fast and reliable drone-enabled method to register aerial thermal and visible images obtained by an electro-optical pod [23]. They used a geometric transformation model to adjust the scale of obtained image frames. Later, they adopted a global feature extraction to extract scale and contrast invariant features uniformly. Additionally, they used an adaptive feature matching method with a kernelized correlation filter to enhance the result. Khattak *et al.* used the combination of infrared and visible cameras to present a method to provide odometry in GPS-denied degraded visual environments for drone-based monitoring, and inspection applications [24]. They employed both spectrums to extract features from regions of interest selected based on spatial entropy. Also, they used inertial sensors as a corrective measure to improve the results. Zhang & Maldague presented a thermal-visible fusion technique based on Non-Subsampled Contourlet Transform (NSCT), and compressed sensing [25]. Their experiment on NSCT-based fusion demonstrated that the employed Compressed Sensing (CS) technology could decrease the required calculations and speed up the convergence in the processing.

The use of different sensors for measuring spatial information are addressed in various studies: (a) laser range finder [26], [27], (b) visible camera [28], [29], and (c) depth sensors [30], [31]. Also, many studies specifically focused on the use of multi-modal platforms containing thermal and visible imagery sensors to generate comprehensive spatial

information for various applications like modeling buildings [30], [32], [33]. Campo *et al.* presented a method for generating sparse depth maps by using multispectral images [34]. The proposed multi-modal acquisition system contains a thermal and a visible camera with parallel fields of view. They discussed that besides the traditional approach to use thermal and visible cameras in a complementary way, it is also possible to extract 3D information. Akhloufi and Verney [35] proposed a multi-modal platform for three-dimensional thermal NDT applications. Also, they aimed to produce an augmented visualization of non-visible defects on a 3D reconstructed surface. They also presented a registration and fusion technique for multi-modal data [36]. Later, Akhloufi *et al.* [37] presented a multi-modal framework containing visible, thermal, and depth sensors. The obtained multi-modal data is aligned using the extracted features from the three modalities. Later, the thermal and visible images are added to the 3D model as texture information. Also, Daffara *et al.* employed a drone-enabled cost-effective setup containing coupled thermal-visible cameras for thermographic inspection of buildings [38]. The system provides a 3D model of a building inspected using a drone. They used a Structure from Motion (SfM) technique to reconstruct the building and then added the thermal values associated with obtained 3D points. Rangel *et al.* introduced a method for the automatic generation of 3D thermal models [39]. They explained that thermal images alone would not be sufficient for the inspections involving a large and complex environment, especially where thermal differences are minor. Furthermore, they described that using visible images in combination with thermal images can provide information about texture and enhance the accuracy of any segmentation techniques. Liu *et al.* [40] presented a registration process for multi-modal data. Also, they explained that the main issue in using multi-sensory platforms is the data fusion process because the two images may not have a linear mapping in two-dimensional space.

B. CALIBRATION BOARDS

Calibration boards for thermal cameras can be categorized into passive and active boards. Active calibration boards use a heating source to generate the required thermal contrast. On the other hand, passive boards do not require any heating source to provide a recognizable pattern. Active calibration boards can be further categorized based on the location of the heater into self-heating and external heating boards. Self-heating boards use an internal excitation source to provide the thermal contrast, such as silicon heaters, lamps, electrical current, or resistors; however, external heating boards need an external excitation source to operate. Calibration boards can also be categorized based on their shapes: squares, checkers, wire nets, circles, and spots. Figure 1 demonstrate the design samples of calibration boards.

1) PASSIVE CALIBRATION BOARDS

One of the motivations to use passive calibration boards is the need for power supply, low manufacturing cost, and the complexity of active targets despite their high performance and effectiveness [41]. Figure 2 demonstrates the used materials, patterns, and approaches for making calibration targets reviewed in this study.

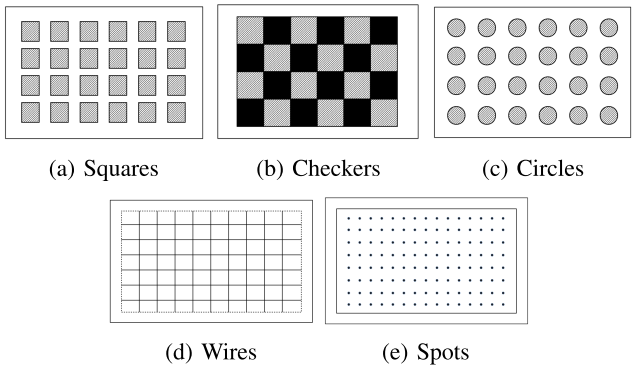

FIGURE 1. Different types of patterns used in calibration targets according to the reviewed studies.

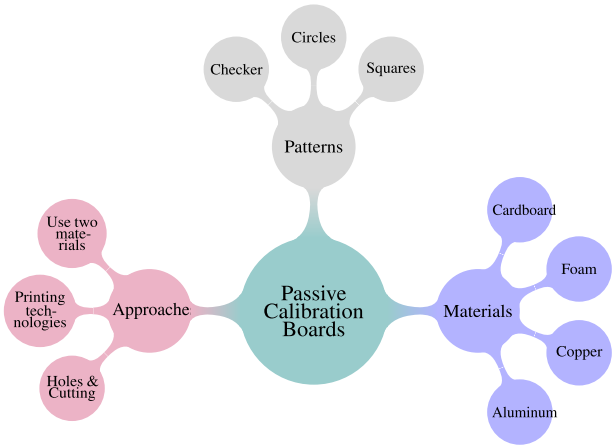

FIGURE 2. Figure shows the used materials, patterns, and approaches employed for passive calibration boards in the reviewed literature.

The most common approach in passive calibration boards is the use of two materials with different emissivities. Ursine *et al.* proposed a checkerboard made of copper squares with low emissivity installed on a background painted with a high emissivity black paint [42]. They used the board for calibration of their Thermal-Visible imaging system. Although they designed the target board for an outdoor environment, they did not address the specular reflection effect for copper plates. Liu *et al.* employed circular patterns for aligning thermal and visible cameras [43]. The pattern has made of CNC machined foam sheet and aluminum disks for the background and the pattern, respectively. The cooled disks are placed on the slots on the foam to generate the thermal contrast. Daffara *et al.* used white cardboard (emissivity is around 0.9) as the base and black painted squares made of aluminum paper (emissivity is about 0.3) for the squares, for aerial 3D thermographic inspection of buildings [38]. Also, they applied finishing (micro-roughness) to the aluminum papers

to address specular reflection. Shivakumar *et al.* designed a calibration target for their multi-modal platform by taking advantage of thermal reflectivity [44]. They used sandblasted aluminum checkers installed on a black acrylic board. Their preliminary experiment determined not to polish the surface to avoid the disturbance on RGB images. Also, they did not use any heating source because of (a) the required effort to impart sufficient heat to obtain a sharp thermal pattern, (b) the problem with the fast cooling-down time of the heated elements, and (c) the difficulty of transferring heat source during the operation.

Some studies have used printing technologies to make calibration patterns on a base panel. St-Laurent *et al.* presented a calibration rig for spatial alignment of images acquired indoor and outdoor in visible, Near-Infrared, SWIR, and LWIR wavebands from 400nm to 12μm [41]. Their experiment concluded that using black paint with high emissivity on a surface with low emissivity can generate a high contrast pattern in the spectral bands of interest. Thus they have printed a checkerboard pattern on a base aluminum panel. Harguess and Strange have employed a calibration board manufactured by 858 Graphics Inc. using a white Dibond aluminum board on which three by five asymmetric black circles are printed [45]. Their experiments have demonstrated that the acquired thermal images of the calibration target contain blurry edges caused by heat diffusion from printed squares to the aluminum panel. Campo *et al.* used a laser printer to print a checkerboard pattern on a thin aluminum paper [34]. They explained that the board suffered from specular reflection during the experiments. Also, the pattern was not sufficiently detectable due to heat diffusion, so they had to use a software solution to enhance the feature points.

Another approach to make passive calibration boards is to use a board with holes to create the pattern by taking advantage of the background scene. In this approach, the distance and the thermal difference between the base and background can effectively make the thermal contrast; however, using these boards in an uncontrolled environment can be challenging since it needs a background with low thermal texture. Vidas *et al.* created the calibration rig by cutting checker squares out of a thin opaque cardboard [46]. They placed the board in front of a backdrop with different colors and thermal profiles to generate the pattern in thermal and visible cameras.

2) ACTIVE CALIBRATION BOARDS

Passive calibration boards often provide a low-contrast pattern. Also, the automatic detection of the calibration pattern can be very challenging in scenarios where the thermal image is noisy, or the surrounding objects have intensive thermal contrast. Active calibration boards are mainly categorized into self-heating and externally-heated boards. Figure 3 demonstrates different types of active calibration boards and the common heating sources based on the reviewed literature.

The primary motivation to use the self-heating approach is to generate high contrast patterns that can be used in outdoor environments despite environmental conditions. Different studies have focused on using various heating sources in their design, such as lamps, silicon heaters, liquid crystal or LED displays, thermostatic heaters, resistors, and passing electrical current.

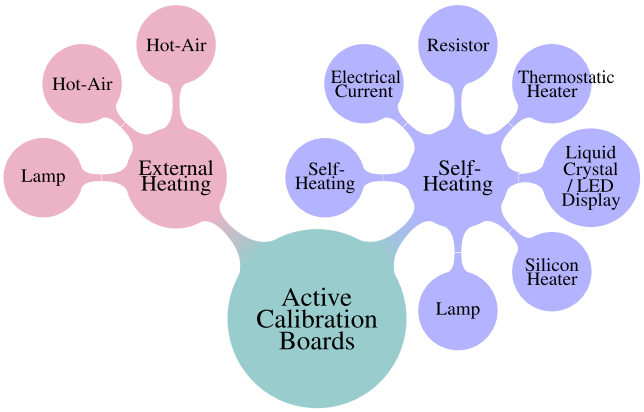

FIGURE 3. Figure shows the categorization of active calibration boards and the used heating sources based on reviewed literature.

Lamps are commonly used in calibration boards since they appear in thermal and visible cameras. Rankin *et al.* presented a board made of a black metal frame with white plastic inserts. They taped miniature light lamps to the back of installed inserts [47]. Also, Yang *et al.* presented a calibration board made of a black plastic board with 25 holes for aligning thermal and visible cameras [29]. They installed twenty-five miniature lamps on the back of the holes in the board. The heat and light emitted by installed bulbs created a pattern recognizable in thermal and visible cameras. Lagüela *et al.* have designed a target board for calibrating thermal cameras, made of a wooden plank with 64 mounted burning lamps [48]. Ellmauthaler *et al.* proposed a new technique for thermal camera calibration using a 9×9 target board consisting of 81 miniature light bulbs [49].

Some studies use silicon heaters as a heating source for the calibration board. St-Laurent *et al.* designed a calibration board for geometric calibration of visible and MWIR/LWIR thermal cameras [8]. The board is made of a plastic plate with a grid of holes, installed on a thin silicon rubber heater. A high-emissivity black paint was applied on the silicon heater while the board was painted white. The heater generates heat while the plastic plate remains at room temperature. Acampora *et al.* presented a method for 3D texturing of thermal images on CAD models [50]. They designed a board made of an aluminum sheet and the checkerboard pattern created by vinyl film cuts for calibration purposes. Their experiment concluded that an unheated board could not generate a sharp, high contrast pattern. Thus, they attached a flexible silicone heater to heat the board temporarily before each experiment.

Some studies passed an electrical current through the board to generate the required thermal contrast. Rasmussen *et al.* have studied the fusion of thermal and visible image streams

for search & rescue purposes [51]. This study presented a calibration target containing a board and a grid net of wires placed on the board to create the pattern. A small electrical current is passed through the wires to warm them up, which successfully generates the patterns. The wire joints were later used as feature points to measure calibration parameters. Also, Liu *et al.* used nickel-chromium heat-resisting wires placed on an acrylic board and connected the wires to a 30V/5A power source to heat them [43].

Besides the mentioned techniques, other studies have chosen different innovative methods to make self-heating boards. Starr and Lattimer used a cold metal grid placed in front of an LCD computer monitor as the target board for mapping indoor environment with low visibility using a stereo thermal camera system [52]. In another study, a small thermostatic heater was installed on the setup as a heat source [53]. Also, Gschwandtner *et al.* investigated using small resistors at the center of black squares on a black & white checkerboard instead of LED lamps [54].

External-heating calibration board needs an external heat source to provide the required thermal contrast. Peric *et al.* investigated using two types of boards as an initial step for the fusion of thermal and visible images [55]. They used polished aluminum as a base and black opaque insulating tape for the first board. Since the pattern was not fully detectable, they used an external heater to provide sharp corners. The second board is built from cardboard for the base and aluminum foil for checkers. This board is placed on a LED display to achieve the required thermal contrast. They have concluded that the second board is suitable for indoor use since the slight radiation received by the LED display provides needed thermal contrast. The heater used in the first board generates solid thermal contrast, which lets the board be used in uncontrolled environments. Kong *et al.* presented a novel calibration board to calibrate stereo thermal cameras for their autonomous landing system [56]. They placed black squares on a mirror heated with a heater. The unreflectivity of mirrors in the thermal spectrum combined with the pasted squares generates a sharp pattern in thermal images.

Some studies used hot airflow as the external heating source for their calibration boards. Ng *et al.* presented a plastic board with a metal grid of wires installed on it as the calibration target [57]. They used a heat gun to heat the wires and generate a pattern recognizable in thermal images. Another study has created the board by miling the checkerboard patterns on a circuit board with high-emissivity for registration of thermal stereo imaging system used for surface reconstruction of water waves [58]. They used a hairdryer before data acquisition to heat the board.

Lamps also are used as an external heat source in some studies. Saponaro *et al.* presented a calibration board designed for thermal stereo camera systems [59]. They printed the checkerboard pattern on a sheet placed on a glazed finish ceramic tile to retain the heat. Also, they used a 250W flood lamp to heat the board. Based on the presented results in this study, the board could not provide strong enough thermal contrast. Therefore, they had to employ contrast enhancement and non-uniformity correction techniques to get a reliable result from the corner detection algorithm. Liu *et al.* also used lamps for heating the target board in one of their investigated methods [43]. They used a simple checkerboard heated using a lamp to generate the thermal contrast. The thermal contrast is generated due to white and black squares' different heat absorption rates.

FIGURE 4. The top photo is a view of the first calibration board, and the bottom photo is the thermal view of the same board. FLIR A700 uncooled thermal camera is used for data collection.

III. CALIBRATION OF MULTI-MODAL PLATFORM

One of the common issues in the calibration of multi-modal imaging systems containing thermal and visible images is to provide a sharp pattern identifiable in both thermal and visible footage. This study presents two reference boards with a checkerboard pattern for calibrating and aligning thermal and visible cameras.

A. FIRST CALIBRATION BOARD

The first reference board is made of a stainless steel plate (type 301) with 0.54 emissivity (ϵ), and the checkerboard patterns are black Vinyl 3M electrical tapes with around 0.95 emissivity. The board size is 150mm $\times$ 125mm and the size of black squares is 20mm $\times$ 15mm making a 5 $\times$ 7 checkerboard pattern. Due to the use of two different materials with different colors, the board provides a sharp distinguishable pattern in both modalities without a need for any heating source. This study mainly utilizes this board for camera calibration and registering thermal and visible images in an indoor environment. Figure 4 presents the thermal and visible views of the first calibration board.

B. SECOND CALIBRATION BOARD

The second board is a novel self-heating calibration board designed with an eye to be also used as GCP in a drone-based thermographic survey. The board employs Thermoelectric cooling/heating (TEC) modules, also known as Peltier plate modules, to generate patterns visible in thermal images. TECs operate based on the Peltier effect. This effect generates a thermal difference by transferring heat between two electrical junctions. The module generates an electrical current when a voltage is applied across the joined conductors. The heat is transferred from one junction to the other when the current flows through both conductors. Using Thermoelectric Coolers has many benefits: (a) low-maintenance cost as it does not have any moving part; (b) ability to cool/heat below or above ambient temperature; (c) the module can be employed as a heating or cooling source; (d) precise temperature control; (e) high reliability and extensive life-time due to their solid-state construction; and (f) no electrical or acoustical noise. Additionally, they are inexpensive, lightweight, and efficient, making them a suitable choice for the intended application [60].

Although the minimum recommended size for a calibration board is 5 × 4, this study uses a prototype 3 × 3 version of the checkerboard board with a size of 122mm × 121mm as proof of principle to investigate the feasibility of using the proposed method and whether the adopted approach can provide identifiable and sharp patterns in visible and thermal images. The board is printed using Acrylonitrile Butadiene Styrene (ABS) material, a thermoplastic polymer with low production cost. ABS material has an approximate emissivity of 0.92 [61] and the melting point of $200°C$. The experiment conducted by Morgan *et al.* using a thermal camera shows that the emissivity of this type of material does not change for observation angles less than 40 degrees. However, it drops off significantly for angles more than 70 degrees [61]. The setup includes four slots for placing four TEC1-12706 modules with the size of 40mm × 40mm in fixed positions to make the checkerboard pattern, as shown in Figure 5. The specification of the TEC module is listed in Table 1. Since wires generate heat while electrical current flows through them, they can disturb the pattern in thermal images; therefore, multiple internal wiring pathways are placed inside the board to hide the wires properly. Additionally, electrical fans are added under TEC modules to help the modules during operation. During the experiment, the board is powered using an adjustable DC power supply for profiling purposes; however, the board can be powered with a Li-Po battery for an outdoor environment.

The introduced board operates in two modes: cooling mode and heating mode. These modes are based on the TEC module's capabilities to perform cooling and heating of one side based on the direction of the flown current. The TEC modules are set to transfer heat from the front side to the rear side in the cooling mode. In this case, the module shows a lower temperature than the board's base providing the required thermal contrast. On the other hand, TEC modules are set to transfer heat from the rear side to the front side in the heating mode. In this mode, the board's front side gets warmer than the base causing the needed thermal contrast. In both modes, the amount of thermal contrast is determined based on the magnitude of applied voltage.

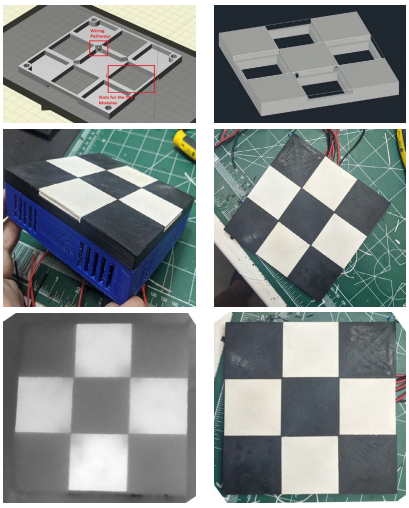

FIGURE 5. The top row photos from left to right are respectively the bottom and top view of the board's design. The middle row photos are photos of the manufactured board. The white squares are TEC modules installed on the setup, and the black squares are part of the setup made of ABS material. Also, the bottom row from left to right shows the acquired thermal and visible images of the board in heating mode.

TABLE 1. The specification of TEC1-12706 module.

Dimensions	$40 \times 40 \times 3.75$ mm
Working Current (I_{max})	4.3 - 4.6 A (rated at 12V)
Internal Resistance	$2.1 \sim 2.4\Omega$
Maximum Voltage (V_{max})	15 V
Maximum Temp. Difference (T_{max})	$59°C \leq$
Working Temperature	$-55°C \sim 83°C$

C. GEOMETRIC CALIBRATION

Accurate knowledge of projection parameters is a preliminary step for quantitative geometric analysis of imagery data in a wide range of applications. Calibration algorithms can be used to recover different camera parameters: (a) intrinsic camera parameters including focal length, the position of principal point, and the camera's scale; (b) the parameters related to non-linear lens distortion; and (c) extrinsic camera parameters including camera's rotation and translation in the coordinate system. Some calibration methods use multiple views of a known pattern with unknown position and orientation in space [62].

The first step includes the mechanical alignment of the sensors and the alignment of both modalities. Since this study uses convergent optical axes configuration, the alignment of thermal and visible images is valid for a specific distance. The registration error (δ) can be estimated as explained in Equation 1 [8], where f is focal length, l_{pix} is the pixel size, d_c is the distance between sensors principal points. D^* is the

optimal distance with no registration error, and D is the target distance. Based on the equation, if the optical axes are placed in parallel ($D^* = \infty$), any scene suffers registration error. The perfect registration of involved modalities at a certain distance can be achieved by aligning the sensors mechanically in a way that the selected feature points from the calibration board can be observed at the principal points of both sensors.

$$\delta = \frac{f.d_c}{l_{pix}} \left(\frac{1}{D} - \frac{1}{D^*} \right) \tag{1}$$

The next step is to detect the checkerboard patterns in thermal and visible images to calibrate the cameras and align both modalities. To do so, the method presented by Geiger *et al.* in [63] is employed for detecting the checkerboard pattern. The method detects the corner points of the pattern with sub-pixel accuracy. Later, Zhang's calibration method [64] is employed to estimate the camera parameters using the first calibration board. This method uses multiple views of the calibration board to estimate the camera parameters. Next, the thermal and visible images are aligned by matching the extracted feature points and estimating the homography matrix. The estimated homography matrix is recorded in the system to be used during the operation for fusing the image modalities.

FIGURE 6. The left photo is the 3D model of FLIR A700 thermal camera and the right photo is the 3D model of Intel RealSense D435i. The models are obtained from STEP models provided by the manufacturers.

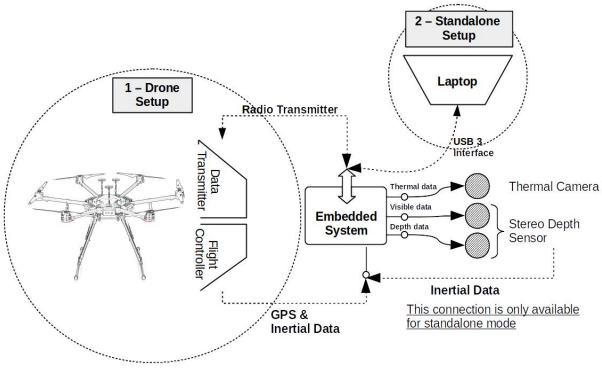

FIGURE 7. The figure presents the system's abstract design. The inertial data provided by the stereo depth sensor is only used in standalone mode.

IV. MULTI-MODAL PLATFORM FOR INSPECTION OF INDUSTRIAL COMPONENTS

This study introduces a multi-modal acquisition and processing platform for inspecting industrial components using Unmanned Aerial Vehicles. The setup is designed to attach the sensors to the aerial platform and stabilize them against the drone's vibrations while operating. The platform includes thermal, visible, and stereo depth cameras and communicates directly with the flight controller to obtain the drone's inertial and GPS data. Also, embedded software is developed for calibration, acquisition, transmission, and fusion of multi-modal data. The data fusion method presented in this study generates RGBD&T data frames that contain the thermal, visible, and depth information of the observed scene and the drone's telemetry data. Later, the RGBD&T data is projected to the system coordinate to form a point cloud of the observed scene.

A. MULTI-MODAL HARDWARE SETUP

The multi-modal system includes a setup designed to hold all hardware components steady while damping vibrations caused by the drone's rotors. Additionally, the system has an embedded system connected to the drone and attached sensors. It acquires multi-sensory data from the platform and GPS & inertial data from the drone. Also, it uses the drone's data transmitter to communicate with Ground Control Station to provide a live view of the observed scene and assist the operator in executing commands remotely. Table 2 shows the specification of sensors used in this system, and Table 3 presents the configuration of the employed embedded board. In this study, a FLIR A700 thermal camera and an Intel RealSense D435i stereo depth sensor containing visible and near-infrared cameras are employed for collecting multi-modal data, as shown in Figure 6.

The developed system can work in two primary modes: drone mode and standalone mode. The system can be used as a handheld setup in standalone mode, letting users move around and observe the scene using the sensors. Due to the lack of inertial sensors in the standalone mode, the Inertial Measurement Unit (IMU) inside the stereo depth sensor is adopted to obtain inertial data. In drone mode, the system is installed on an aerial platform. Also, the inertial sensors of the drone's flight controller are used for obtaining telemetry data. Moreover, the system communicates with the user using the data transmitter embedded in the DJI M600 Pro drone. Figure 7 demonstrates the view of the system's abstract design. In this system, the thermal camera connects to the embedded system using a PoE Ethernet interface, and the stereo depth sensor uses USB 3 interface for transmitting data.

The introduced setup in this study is designed and 3D printed specifically for DJI Matrice 600 Pro. The setup can carry all the sensors and the embedded system while preventing the vibration caused by the drone from affecting the acquired imagery data. Figure 8 presents the sensor setup. The setup includes a mounting component made of a flexible material that attaches the sensors to the drone's structure and shields the sensors from vibrations.

TABLE 2. The specification of employed sensors in the multi-modal system.

	Configuration	Sensor Technology	Resolution	Sensitivity	Acquisition Rate	FOV	Focus	Interface	Range
Thermal Sensor	FLIR A700 Image Streaming	Uncooled Microbolometer	640 × 480	< 30 mK	30 Hz	14° × 10°	Motorized Focus	M12 to RJ45F	−20 to 120°, 0 to 650°, 300 to 2000°
Visible Sensor	Intel RealSense D435i	Rolling Shutter	1920 × 1080	–	30 Hz	69° × 42°	Auto	USB-C 3.1 Gen 1	–
Depth Sensor		Stereoscopic	1280 × 720	<2% at 2 m	<90 Hz	87° × 58°	–		0.3 – 3.0 m

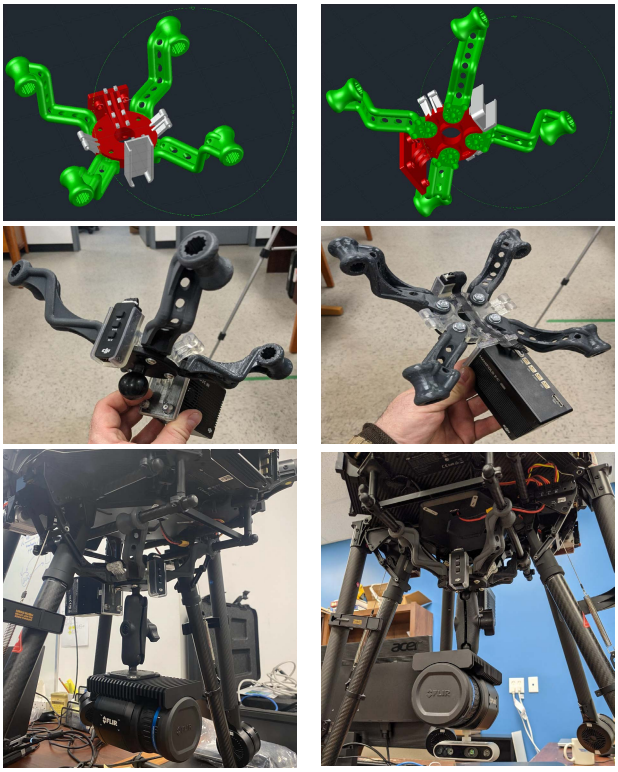

FIGURE 8. The setup design is presented in the first row. The second row is a view of the printed setup, and the last row is the setup installed on a DJI Matrice 600 pro.

B. MULTI-MODAL EMBEDDED SOFTWARE SYSTEM

In this study, an embedded system is developed on top of Robot Operating System (ROS) using C++ and Python programming languages that handles the calibration, acquisition, transmission, storage, and configuration process.[1] The system's components take advantage of the ROS communication graph model for internal communication. A comprehensive ROS driver node is developed for FLIR thermal cameras using a PoE Ethernet interface that provides two-way communication between the thermal camera and the computer.[2] As presented in Figure 9, the system has three ROS nodes communicating with the sensors and flight controller. The data streams received from the sensors are collected by

[1]The source code for the embedded system is available at https://github.com/parham/lemanchot-dc

[2]The source code for FLIR PoE thermal camera's ROS driver is available at https://github.com/parham/ros_flir_spinnaker

TABLE 3. Embedded system (DJI Manifold 2-C)'s configuration.

Weight	∼ 205 g	**Dimension**	91 × 61 × 35 mm
Processor	Intel Core i7-8550U	**Memory**	8GB 64 bit, DDR4 2400 MHz
Power	5 – 60 W	**Power Supply**	15.2 - 27.0 V

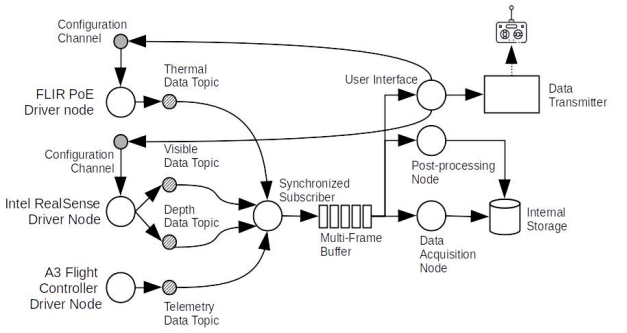

FIGURE 9. The figure presents the abstract design of the ROS-based embedded software presented in this paper. The system provides three driver nodes for thermal, depth, and visible cameras as well as the drone's flight controller. The system subscribes to the exposed data topics for multi-modal data acquisition, transmission, and fusion.

a synchronized subscriber, ensuring the collection of data frames with the same timestamp in a fixed data rate.

Moreover, a user-friendly graphical user interface is developed to visualize the multi-modal data in standalone and drone mode. The control options in the user interface are only available in standalone mode. The system uses the customizable SDK ports existing in the drone's remote controller for the drone mode. Figure 10 shows two views of the system in the presented operating modes.

In addition to the acquisition, transmission, and recording features, the system also includes two processing steps for ensuring data reliability and quality: motion blurriness detection and image normalization. For the detection of motion blurriness, an algorithm presented in [65] is adopted that flags the visible images affected by motion blurriness which can assist inspectors in the post-processing step. This method calculates a blurriness metric that later can be used to flag the images as blurry with a customizable threshold defined before each mission.

For illumination normalization of the consecutive images, a dense histogram adjustment technique [66] along with a floating window with a fixed length is adopted that uses a reference image to match the histogram of the collected images. In the floating window approach, the first frame is considered the reference frame for each iteration, and the

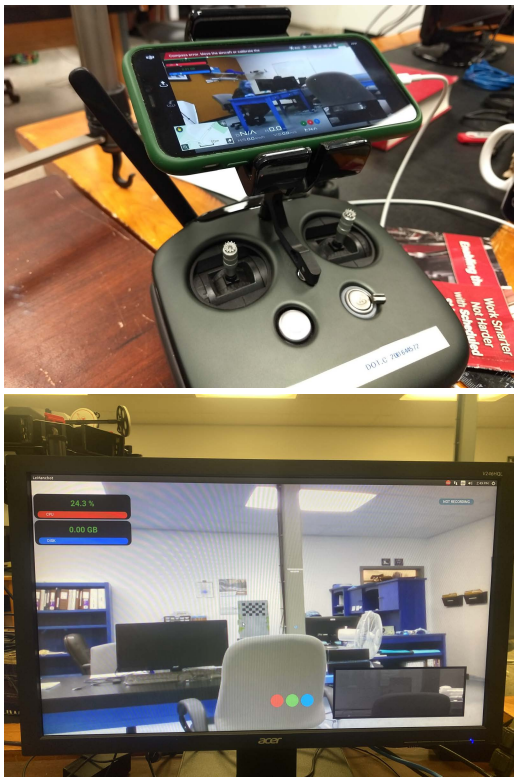

FIGURE 10. The top photo shows a view of the system's user interface in drone mode, and the bottom photo is the system's user interface in standalone mode.

following frames are normalized based on the determined reference frame.

C. MULTI-MODAL DATA FUSION

The system introduced in this study provides three types of image modalities: visible, thermal, and depth images. It acquires depth and visible images from an Intel RealSense D435i sensor which features a stereo infrared, an RGB camera, and an infrared laser projector. The sensor assesses the depth information of the observed scene by taking advantage of the weak pattern projected on the surrounding objects. Also, the system uses a FLIR A700 uncooled thermal camera to obtain thermal images. The system collects these modalities synchronously using independent processes. In addition to the acquisition, storage, and transmission, the system presents a data fusion technique to provide RGBD&T data presenting a 3D view of the observed scene containing texture and thermal information.

The presented fusion technique uses two independent steps to fuse the modalities: (a) registration of depth map and visible image, and (b) registration of visible and thermal images. Intel RealSense D435i provides a built-in feature to align the visible and depth images in real-time. So, the study mainly focuses on the registration of thermal and visible images involving two main challenges: (a) thermal images often do not contain enough reliable features using standard feature extraction techniques; (b) it is essential to find

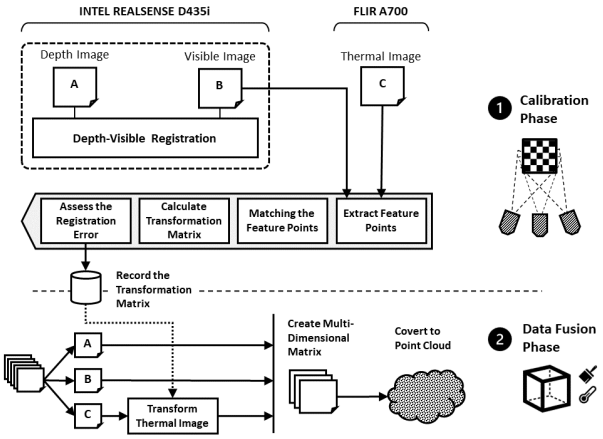

FIGURE 11. The figure presents the proposed process pipeline for multi-modal data fusion. The system uses feature points of the reference board to calculate transformation matrix aligning all modalities during initialization step. Later, the thermal, depth, and visible images were fused to form a RGBD&T data which later is used to generate multi-modal point cloud.

distinguishable and matchable features in both modalities. Therefore, this study employs a calibration-based registration method to address these challenges. Figure 11 demonstrates the proposed process pipeline for fusion of image modalities.

Since the sensors are relatively in a fixed position close to each other, thermal and visible images can be fused using a transformation matrix. The system uses a calibration board to provide a sharp and identifiable pattern visible in both modalities. During the calibration step, the feature points of employed calibration boards are extracted as explained in Section III-C. In this step, the feature points are used to estimate the transformation matrix for aligning the thermal and visible images during the operation. After aligning thermal, depth, and visible images, they are concatenated to provide a multi-dimensional matrix referred to as RGBD&T data that represents the pixels' depth, color, and thermal values.

As explained before, the registration of thermal and visible images is reliable for the distance for which the cameras were calibrated. Therefore, a post-processing step filters the pixels with a depth value out of a defined range. The range of acceptable depth is determined during the calibration based on the inspection requirements. This step decreases possible misalignment and removes the noisy data obtained by the depth sensor. Finally, the RGBD&T data is projected to provide a point cloud presenting the scene using Equation 2, where D is the depth map, p_x and p_y are the coordinates of the principal points, f_x and f_y are the focal lengths of image.

$$\begin{aligned} &\forall x \in \{0, \ldots, D_w\}, \quad \forall y \in \{0, \ldots, D_h\}, \\ &d_{x,y} = \frac{D(x,y)}{1000} : \\ &P(x,y,z) = d_{x,y} \times \left[\frac{x - p_x}{f_x}, \quad \frac{y - p_y}{f_y}, \quad 1\right]^{-1} \end{aligned} \tag{2}$$

V. RESULTS & DISCUSSIONS

In this section, the obtained results of the conducted experiment are presented. Firstly, the introduced multi-modal fusion technique is assessed. Later, the use of the presented reference board is evaluated in an indoor environment. Also, the feasibility of using the reference board as GCP for drone surveys is investigated.

A. ASSESSMENT OF MULTI-MODAL FUSION TECHNIQUE

In this study, the system acquires the three modalities synchronously with a customizable rate. Later, the system uses the transformation matrix obtained during the calibration to form the RGBD&T data. For assessing the fusion method, an experimental piping system used in the oil & gas industry was selected as a use case to demonstrate the system applicability and performance. During the experiment, the multi-sensory platform was moved through the pipeline in a 2.5 meters distance to obtain multi-modal data. Figure 12 presents the sample result of the obtained data during the data acquisition. As shown in this figure, the thermal view of the point cloud contains points with no values due to the different fields of view of the thermal and depth sensors. Also, the minor distortions in the reconstructed point cloud are caused by the sensor's internal depth estimation error.

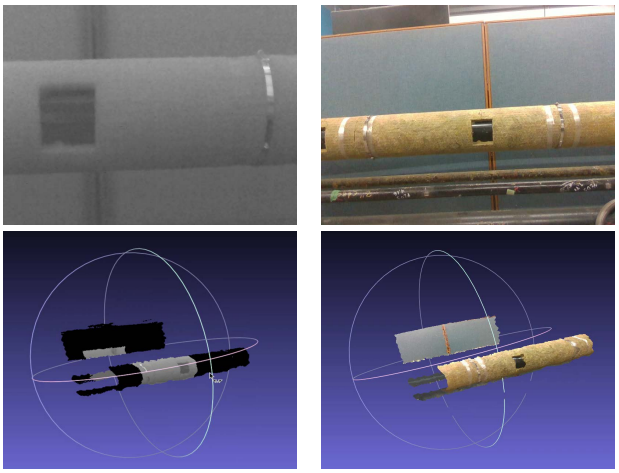

FIGURE 12. The top left, and right photos are the acquired thermal and visible images. The bottom left, and right photos are respectively the 3D model presenting thermal and texture values.

An experiment was conducted using the introduced unheated calibration board in an indoor environment to assess the registration process. To do so, the board was placed at three different distances in the scene while the thermal and visible images were collected using the multi-modal platform. Later, two modalities were aligned, as explained in this study. The registration process is evaluated using five metrics: (a) Root-Mean-Square Error (RMSE), (b) Peak Signal-to-Noise Ratio (PSNR), (c) Structural Similarity Index (SSIM), (d) Feature-based similarity index (FSIM), and (e) Spectral angle mapper (SAM). In this study, the implementations of the metrics provided by Müller are employed for data assessment [67].

RMSE presents the square root of Mean Squared Error (MSE) as shown in Equation 3, where y^V and y^T respectively present visible and thermal images; N is the number of pixels in the images. PSNR can be defined as the ratio between the signal's maximum value and the power of possible distortion as shown in Equation 4, where MAX_y is the maximum possible value of the signal. SSIM is a perceptual metric simply demonstrating the perceptual differences between images. FSIM is a metric to determine the similarity of two images based on two criteria: Phase Congruency (PC) and Gradient Magnitude (GM) [68]. Finally, SAM is a metric comparing two spectra using a spectral angle error [69]. Table 4 presents the result of thermal and visible registration in three different distances. The results show that the system can align the data modalities successfully using the board placed at different distances.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i^V - y_i^T)^2} \tag{3}$$

$$PSNR = 20\ log_{10}\left(\frac{MAX_y}{RMSE}\right) \tag{4}$$

TABLE 4. The result of thermal and visible registration is explained in this table.

Distance (cm)	RMSE	PSNR	SSIM	FSIM	SAM
100	0.0163	35.73	0.803	0.33	13.764
210	0.0156	37.163	0.915	0.44	29.183
320	0.0163	35.734	0.803	0.33	13.76

B. EVALUATION OF THE REFERENCE BOARD IN AN INDOOR ENVIRONMENT

In this experiment, the prototyped reference board is evaluated in a controlled environment to assess its functionality. The generated temperature differences and the board's power consumption are also investigated. The reference board was used without cooling fans during the experiment and connected directly to an adjustable DC power supply. A FLIR T650sc uncooled thermal camera is placed at a 3 meters distance to acquire thermal data.

In order to investigate the board's power consumption and thermal contrast, multiple input voltages were tested. Figure 13 shows the thermal views of the board with different voltages and obtained thermal contrast. The input voltage was set manually in the power supply, and the temperature differences and drawn current were measured. The feasibility of using the board as a portable reference board powered by a Lithium polymer battery was also assessed by applying relatively low voltage. Table 5 presents the results of this evaluation. The results demonstrate that the board can generate a sharp and matchable pattern in the thermal spectrum using different voltages. Also, it is determined that the amount of proper input voltage for the board is based on environmental

TABLE 5. The table demonstrates the measured power consumption and estimated temperature differences of the board. * ΔT presents the temperature difference between white and black squares in the board measured by thermal camera.

Voltage (V)	Current (mA)	$\Delta T(^{\circ}C)$*
0.5	–	3
1.0	300	4.9
1.5	750	10.52
2.0	810	16.1

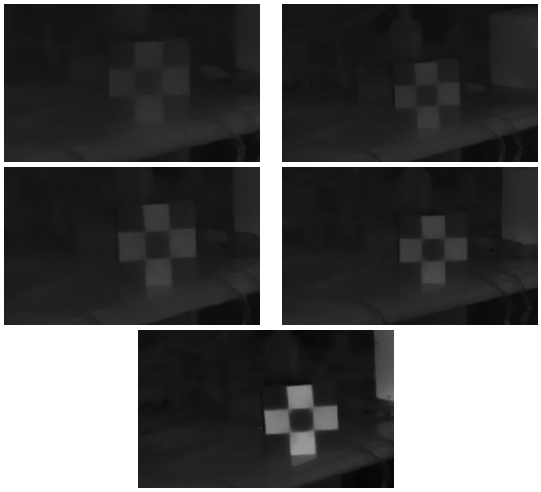

FIGURE 13. The figure demonstrates the thermal images collected while the board uses different voltages. The left photo from the top row is the board when unplugged.

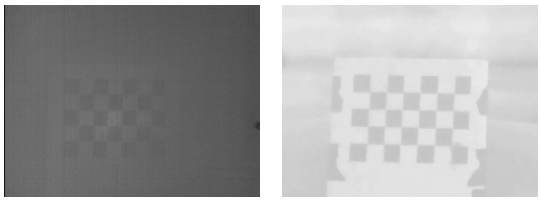

FIGURE 14. The left photo is the thermal image collected by Teledyne Calibir GXM640, and the right photo is the thermal image acquired by FLIR A700. Both cameras were used with default settings without any improvement or optimization.

conditions like temperature and surrounding objects and the required thermal contrast. Since thermal cameras have different sensitivity, the thermal contrast needed for the board to make the thermal pattern visible can vary.

To investigate the effect of the camera's sensitivity on the visibility of the checkerboard pattern in thermal images, two uncooled thermal cameras, including FLIR A700 and Teledyne Calibir GXM640, were used to acquire thermal images of the introduced unheated calibration board. Figure 14 shows the collected thermal images. The results demonstrate the difference in the pattern's visibility with different cameras. Also, it emphasizes the necessity and applicability of the introduced self-heating reference board with adjustable thermal contrast.

C. OUTDOOR ASSESSMENT OF THE BOARD AS GCP

This experiment investigated the applicability of using the reference board as a Ground Control Point for drone survey applications. To do so, the board was placed on the ground in an outdoor environment on a sunny day. The environmental conditions of the experiment are listed in Table 6. The board was connected to an adjustable DC power supply for self-heating purposes. Also, a moving cart was used for transporting the equipment. A DJI M300 drone equipped with a Zenmuse H20T camera was employed for acquiring thermal and visible images as shown in Figure 15. To analyze the board's power consumption, the power supply was adjusted in different input voltages to measure the current drawn by the board for self-heating. Figure 16 demonstrates the voltage-current curve.

During the experiment, the drone was hovering above the board at various altitudes from $5-50$ meters while collecting thermal and visible images. Samples of collected data is presented in Figure 17. The experiment intends to determine the thermal visibility of the board in different altitudes.

TABLE 6. The environment conditions of the outdoor assessment of the introduced calibration board.

Temperature	22 c	**UV**	low
Humidity	64%	**Wind**	10 km/h

FIGURE 15. A photo of the conducted experiment using the DJI M300 drone is presented while hovering above the board.

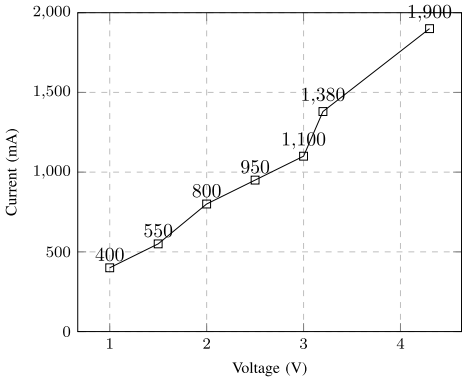

FIGURE 16. The figure presents the voltage-current curve of the reference board. It shows the consumed electrical current in candidate input voltages. The increase in voltage causes the increase in temperature differences. The measures can help determining the required electrical power for specific inspection based on environmental conditions and inspection's requirements.

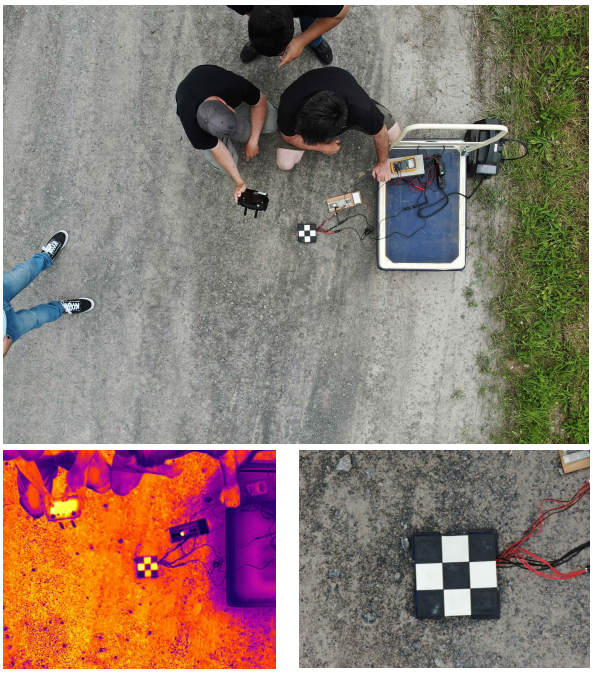

FIGURE 17. The top photo shows a view of the introduced board, captured by a drone flying in approximately 6 meters. The bottom left, and right are respectively the board's thermal and zoomed visible view with the same experimental parameters.

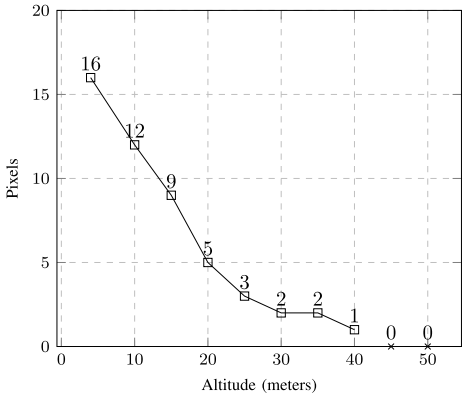

FIGURE 18. The figure shows the measured pixel width of the board's squares at different altitudes. At 45 and 50 meters, the squares were not recognizable. The number of pixels presenting the board is directly related to the board size. So by increasing the board size, the number of pixels would increase. The experiment mainly presents the board's ability to provide sharp and recognizable patterns that can be identifiable in different altitudes in thermal and visible modalities.

The results presented in Figure 18 show that the thermal pattern formed by the board was identifiable till 40 meters altitude. However, based on our experiments, the most suitable altitudes for the alignment of thermal images were 15 meters and lower.

The feasibility of using the board in drone surveys was also investigated in both modes. The drone was set to hover at 5 meters for this experiment while the board was set to heating mode. Later, the same experiment was conducted in cooling mode. Figure 19 shows samples of conducted experiment. The obtained results demonstrate that the board provides

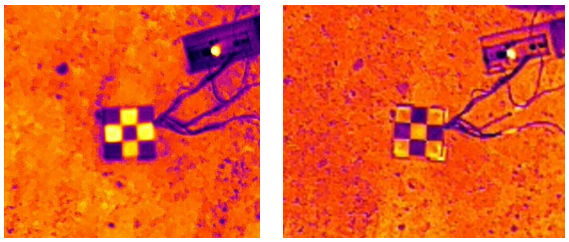

FIGURE 19. The right photo is the board in heating mode, and the left photo is the board in cooling mode.

a more distinguishable pattern in heating mode. Based on the conducted investigation, the board's mode should be determined based on the environment's condition, such as temperature and surrounding thermal pattern.

VI. CONCLUSION

In this study, a multi-modal platform is presented to inspect industrial components. The system can acquire the modalities of interest synchronously. Also, the system includes a customizable user interface that assists inspectors in monitoring the system status and the drone's internal state. Additionally, a fusion technique is presented, providing RGBD&T data containing thermal and texture information for the collected 3D models. Moreover, a novel self-heating reference board is proposed for system calibration. The board can provide sharp and identifiable patterns visible in thermal and visible modalities. In this study, an experiment was conducted to assess the performance of the proposed multi-modal fusion method. The results demonstrated that the method could fuse the modalities with significantly low registration error. Moreover, the novel reference board is evaluated for indoor and outdoor applications. Finally, the feasibility of using the board as GCP is investigated. The results demonstrate its capabilities for drone surveys.

ACKNOWLEDGMENT

Special thanks to TORNGATS Company, especially Frédéric Lajoie and Pierre Delisle, for their support in testing and manufacturing of the parts.

REFERENCES

[1] I. R. Farah, W. Boulila, K. S. Ettabaa, and M. Ben Ahmed, "Multiapproach system based on fusion of multispectral images for land-cover classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4153–4161, Dec. 2008.

[2] P. Singh, Y. Wu, R. Kaucic, J. Chen, and F. Little, "Multimodal industrial inspection and analysis," *J. Comput. Inf. Sci. Eng.*, vol. 7, no. 1, pp. 102–107, Mar. 2007.

[3] P. Gamba, F. Dell'Acqua, and B. V. Dasarathy, "Urban remote sensing using multiple data sets: Past, present, and future," *Inf. Fusion*, vol. 6, no. 4, pp. 319–326, Dec. 2005.

[4] I. R. Farah and M. B. Ahmed, "Towards an intelligent multi-sensor satellite image analysis based on blind source separation using multi-source image fusion," *Int. J. Remote Sens.*, vol. 31, no. 1, pp. 13–38, Jan. 2010.

[5] S. Delalieux, P. J. Zarco-Tejada, L. Tits, M. Á. J. Bello, D. S. Intrigliolo, and B. Somers, "Unmixing-based fusion of hyperspatial and hyperspectral airborne imagery for early detection of vegetation stress," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2571–2582, Jun. 2014.

[6] J. Marcello, A. Medina, and F. Eugenio, "Evaluation of spatial and spectral effectiveness of pixel-level fusion techniques," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 432–436, May 2013.
[7] W. Li, S. Prasad, and J. E. Fowler, "Decision fusion in kernel-induced spaces for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3399–3411, Jun. 2014.
[8] L. St-Laurent, D. Prévost, and X. Maldague, "Fast and accurate calibration-based thermal/colour sensors registration," in *Proc. Int. Conf. Quant. Infr. Thermography*, 2010, Paper. QIRT2010-126.
[9] D. K. Edwards. (2015). *Radiation transfer between specular and imperfectly diffuse surfaces*. Accessed: Dec. 12, 2021. Begell House Inc. ed., New York, NY, USA: doi: 10.1615/hedhme.a.000207.
[10] E. Loosli, "Ground control points: How many do you need (and when are checkpoints enough)," Wingtra, Giesshübelstrasse, Zurich, Switzerland Tech. Rep., Mar. 2021.
[11] X. E. Gros, P. Strachan, and D. Lowden, "Fusion of multiprobe NDT data for ROV inspection," in *Proc. Challenges Our Changing Global Environ. Conf. (OCEANS MTS/IEEE)*, Oct. 1995, pp. 2046–2050.
[12] D. Horn and W. R. Mayo, "NDE reliability gains from combining eddy-current and ultrasonic testing," *NDT E Int.*, vol. 33, no. 6, pp. 351–362, Sep. 2000.
[13] K. Simmen *et al.*, "Multimodal image acquisition and processing system for in-line quality monitoring of welding processes,"in *Proc. 3rd Young Weld. Professionals Int. Conf.*, Halle Saale, Germany, Aug. 2017, pp. 16–18.
[14] M. Li, N. Yao, S. Liu, S. Li, Y. Zhao, and S. G. Kong, "Multisensor image fusion for automated detection of defects in printed circuit boards," *IEEE Sensors J.*, vol. 21, no. 20, pp. 23390–23399, Oct. 2021.
[15] S.-J. Yu, S. R. Sukumar, A. F. Koschan, D. L. Page, and M. A. Abidi, "3D reconstruction of road surfaces using an integrated multi-sensory approach," *Opt. Lasers Eng.*, vol. 45, no. 7, pp. 808–818, 2007.
[16] K. Ishikawa, J.-I. Takiguchi, Y. Amano, and T. Hashizume, "A mobile mapping system for road data capture based on 3D road model," in *Proc. IEEE Int. Conf. Control Appl.*, Oct. 2006, pp. 43–50.
[17] H. D. Cheng, J. Chen, C. Glazier, and Y. Hu, "Novel approach to pavement cracking detection based on fuzzy set theory," *J. Comput. Civil Eng.*, vol. 13, no. 4, pp. 270–280, 1999.
[18] B. Javidi, D. Kim, and S. E.-Sayed Kishk, "A laser-based 3D data acquisition system for the analysis of pavement distress and roughness," Connecticut Transp. Inst., School Eng., Univ. Connecticut, Tech. Rep. JHR 04-300, 2004.
[19] M. Alsafasfeh, I. Abdel-Qader, B. Bazuin, Q. Alsafasfeh, and W. Su, "Unsupervised fault detection and analysis for large photovoltaic systems using drones and machine vision," *Energies*, vol. 11, no. 9, pp. 1–18, 2018.
[20] S. Lee, K. E. An, B. D. Jeon, K. Y. Cho, S. J. Lee, and D. Seo, "Detecting faulty solar panels based on thermal image processing," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2018, pp. 1–2.
[21] F. Miralles, P. Hamelin, G. Lambert, S. Lavoie, N. Pouliot, M. Montfrond, and S. Montambault, "LineDrone technology: Landing an unmanned aerial vehicle on a power line," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 6545–6552.
[22] B. Wang, X. Chen, Q. Wang, L. Liu, H. Zhang, and B. Li, "Power line inspection with a flying robot," in *Proc. 1st Int. Conf. Appl. Robot. Power Ind. (CARPI)*, Oct. 2010, pp. 1–6.
[23] X. Liu, Y. Ai, B. Tian, and D. Cao, "Robust and fast registration of infrared and visible images for electro-optical pod," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1335–1344, Feb. 2019.
[24] S. Khattak, C. Papachristos, and K. Alexis, "Visual-thermal landmarks and inertial fusion for navigation in degraded visual environments," in *Proc. IEEE Aerosp. Conf.*, Mar. 2019, pp. 1–9.
[25] Q. Zhang and X. Maldague, "An infrared-visible image fusion scheme based on NSCT and compressed sensing," *Proc. SPIE*, vol. 9474, May 2015, Art. no. 94740Y.
[26] K. Nagatani, K. Otake, and K. Yoshida, "Three-dimensional thermography mapping for mobile rescue robots," in *Field and Service Robotics*. Cham, Switzerland: Springer, 2014, pp. 49–63.
[27] D. Borrmann, H. Afzal, J. Elseberg, and A. Nüchter, "Mutual calibration for 3D thermal mapping," *IFAC Proc. Volumes*, vol. 45, no. 22, pp. 605–610, 2012.
[28] P. Aksenov, I. Clark, D. Grant, A. Inman, L. Vartikovski, and J.-C. Nebel, "3D thermography for quantification of heat generation resulting from inflammation," in *Proc. 3D Modeling Symp.*, Princeton, NJ, USA: Citeseer, 2003, pp. 1–11.
[29] R. Yang, W. Yang, Y. Chen, and X. Wu, "Geometric calibration of IR camera using trinocular vision," *J. Lightw. Technol.*, vol. 29, no. 24, pp. 3797–3803, Dec. 15, 2011.
[30] S. Vidas, P. Moghadam, and M. Bosse, "3D thermal mapping of building interiors using an RGB-D and thermal camera," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 2311–2318.
[31] K. Skala, T. Lipić, I. Sović, L. Gjenero, and I. Grubić, "4d thermal imaging system for medical applications," *Periodicum biologorum*, vol. 113, no. 4, pp. 407–416, 2011.
[32] Y. Ham and M. Golparvar-Fard, "Rapid 3D energy performance modeling of existing buildings using thermal and digital imagery," in *Proc. Construct. Res. Congr.*, May 2012, pp. 991–1000.
[33] D. Iwaszczuk, L. Hoegner, and U. Stilla, "Matching of 3D building models with IR images for texture extraction," in *Proc. Joint Urban Remote Sens. Event*, Apr. 2011, pp. 25–28.
[34] F. B. Campo, F. L. Ruiz, and A. D. Sappa, "Multimodal stereo vision system: 3D data extraction and algorithm evaluation," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 5, pp. 437–446, Sep. 2012.
[35] M. A. Akhloufi and B. Verney, "Fusion framework for 3D inspection and thermal NDT," SAE Tech. Paper, Warrendale, PA, USA, Tech. Rep. 2013-01-2171, 2013.
[36] M. A. Akhloufi and B. Verney, "Multimodal registration and fusion for 3D thermal imaging," *Math. Problems Eng.*, vol. 2015, pp. 1–14, Jan. 2015.
[37] M. A. Akhloufi, Y. Guyon, C.-I. Castanedo, and A. Bendada, "Three-dimensional thermography for non-destructive testing and evaluation," *Quant. Infr. Thermography J.*, vol. 14, no. 1, pp. 79–106, Jan. 2017.
[38] C. Daffara, R. Muradore, N. Piccinelli, N. Gaburro, T. de Rubeis, and D. Ambrosini, "A cost-effective system for aerial 3D thermography of buildings," *J. Imag.*, vol. 6, no. 8, p. 76, Aug. 2020.
[39] J. Rangel and S. Soldan, "3D thermal imaging: Fusion of thermography and depth cameras," in *Proc. Int. Conf. Quant. Infr. Thermography*, 2014, pp. 1–10.
[40] H. Liu, S.-H. Lee, and J. S. Chahl, "Registration of multispectral 3D points for plant inspection," *Precis. Agricult.*, vol. 19, no. 3, pp. 513–536, Jun. 2018.
[41] L. St-Laurent, M. Mikhnevich, A. Bubel, and D. Prévost, "Passive calibration board for alignment of VIS-NIR, SWIR and LWIR images," *Quant. Infr. Thermography J.*, vol. 14, no. 2, pp. 193–205, Jul. 2017.
[42] W. Ursine, F. Calado, G. Teixeira, H. Diniz, S. Silvino, and R. D. Andrade, "Thermal/visible autonomous stereo visio system calibration methodology for non-controlled environments," in *Proc. Int. Conf. Quant. Infr. Thermography*, 2012, pp. 1–10.
[43] R. Liu, H. Zhang, and S. Scherer, "Multiple methods of geometric calibration of thermal camera and a method of extracting thermal calibration feature points," The Robot. Inst., Carnegie Mellon Univ., Tech. Rep., 2018. Organization: The Robotics Institute, Carnegie Mellon University Address: Forbes Ave, Pittsburgh, PA, United States
[44] S. S. Shivakumar, N. Rodrigues, A. Zhou, I. D. Miller, V. Kumar, and C. J. Taylor, "PST900: RGB-thermal calibration, dataset and segmentation network," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 9441–9447.
[45] J. Harguess and S. Strange, "Infrared stereo calibration for unmanned ground vehicle navigation," *Proc. SPIE*, vol. 9084, Jun. 2014, Art. no. 90840S.
[46] S. Vidas, R. Lakemond, S. Denman, C. Fookes, S. Sridharan, and T. Wark, "A mask-based approach for the geometric calibration of thermal-infrared cameras," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 6, pp. 1625–1635, Jun. 2012.
[47] A. Rankin, A. Huertas, L. Matthies, M. Bajracharya, C. Assad, S. Brennan, P. Bellutta, and G. W. Sherwin, "Unmanned ground vehicle perception using thermal infrared cameras," *Proc. SPIE*, vol. 8045, May 2011, Art. no. 804503.
[48] S. Lagüela, H. González-Jorge, J. Armesto, and P. Arias, "Calibration and verification of thermographic cameras for geometric measurements," *Infr. Phys. Technol.*, vol. 54, no. 2, pp. 92–99, Mar. 2011.
[49] A. Ellmauthaler, E. A. B. da Silva, C. L. Pagliari, J. N. Gois, and S. R. Neves, "A novel iterative calibration approach for thermal infrared cameras," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 2182–2186.
[50] L. Acampora, F. D. Filippis, A. Martucci, and L. Sorgi, "3D reconstruction of thermal images," in *Proc. 26th Aerosp. Test. Seminar*, 2011, pp. 263–277.
[51] N. D. Rasmussen, B. S. Morse, M. A. Goodrich, and D. Eggett, "Fused visible and infrared video for use in wilderness search and rescue," in *Proc. Workshop Appl. Comput. Vis. (WACV)*, Dec. 2009, pp. 1–8.
[52] J. W. Starr and B. Y. Lattimer, "A comparison of IR stereo vision and LIDAR for use in fire environments," in *Proc. IEEE Sensors*, Oct. 2012, pp. 1–4.

[53] Z. Yu, S. Lincheng, Z. Dianle, Z. Daibing, and Y. Chengping, "Camera calibration of thermal-infrared stereo vision system," in *Proc. 4th Int. Conf. Intell. Syst. Design Eng. Appl.*, Nov. 2013, pp. 197–201.
[54] M. Gschwandtner, R. Kwitt, A. Uhl, and W. Pree, "Infrared camera calibration for dense depth map construction," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2011, pp. 857–862.
[55] D. Peric, V. Lukic, M. Spanovic, R. Sekulic, and J. Kocic, "Geometric calibration of multi-sensor image fusion system with thermal infrared and low-light camera," *Proc. SPIE*, vol. 9250, Oct. 2014, Art. no. 92500O.
[56] W. Kong, D. Zhang, X. Wang, Z. Xian, and J. Zhang, "Autonomous landing of an UAV with a ground-based actuated infrared stereo vision system," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 2963–2970.
[57] Y. H. Ng and R. Du, "Acquisition of 3D surface temperature distribution of a car body," in *Proc. IEEE Int. Conf. Inf. Acquisition*, Jun. 2005, p. 5.
[58] V. Hilsenstein, "Surface reconstruction of water waves using thermographic stereo imaging," in *Image and Vision Computing New Zealand*, vol. 2. Princeton, NJ, USA: Citeseer, 2005.
[59] P. Saponaro, S. Sorensen, S. Rhein, and C. Kambhamettu, "Improving calibration of thermal stereo cameras using heated calibration board," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 4718–4722.
[60] M. R. Atta, "Thermoelectric cooling," in *Bringing Thermoelectricity into Reality*, Intechopen, ed., Jul. 2018, doi: 10.5772/intechopen.75791.
[61] R. V. Morgan, R. S. Reid, A. M. Baker, B. Lucero, and J. D. Bernardin, "Emissivity measurements of additively manufactured materials," Los Alamos Nat. Lab. (LANL), Los Alamos, NM, USA, Tech. Rep. LA-UR-7-20513, 2017.
[62] W. Burger, "Zhang's camera calibration algorithm: In-depth tutorial and implementation," in *Proc. HGB*, 2016, pp. 1–6.
[63] A. Geiger, F. Moosmann, O. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 3936–3943.
[64] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Dec. 2000.
[65] R. Liu, Z. Li, and J. Jia, "Image partial blur detection and classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
[66] H. Niu, Q. Lu, and C. Wang, "Color correction based on histogram matching and polynomial regression for image stitching," in *Proc. IEEE 3rd Int. Conf. Image, Vis. Comput. (ICIVC)*, Jun. 2018, pp. 257–261.
[67] M. U. Müller, N. Ekhtiari, R. M. Almeida, and C. Rieke, "Super-resolution of multispectral satellite images using convolutional neural networks," 2020, *arXiv:2002.00580*.
[68] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
[69] R. H. Yuhas, A. F. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, vol. 1, 1992, pp. 147–149.

PARHAM NOORALISHAHI (Graduate Student Member, IEEE) received the master's degree in computer science (artificial intelligence). He is currently pursuing the Ph.D. degree in electrical engineering with Université Laval, working on developing drone-enabled techniques for the inspection of large and complex industrial components using multimodal data processing. He is also a Researcher with a demonstrated history of working in the telecommunication industry and industrial inspection and in-depth expertise in robotics & drones, embedded systems, advanced computer vision, and machine learning techniques. He is a Specialist in embedded and intelligent vision systems. During his bachelor's degree, he was involved in designing and developing the controlling and monitoring systems for a fixed-wing drone for search and rescue purposes. Also, during his master's degree, he worked extensively on machine learning and computer vision techniques for robotic and soft computing applications.

FERNANDO LÓPEZ (Member, IEEE) received the Ph.D. degree in mechanical engineering, in 2014. He worked as a Postdoctoral Researcher at Université Laval, conducting research projects with various industrial partners, mainly in aerial IR thermography (IRT) inspection, energy efficiency, and robotic IRT for the non-destructive testing and evaluation (NDT&E) of aerospace components. He is currently the Director of Research and Development at TORNGATS, leading several research and development initiatives on advanced NDT&E methods. He is a Senior Scientist with over 12 years of experience in industry and research in infrared (IR) imaging, advanced NDT&E of materials, applied heat transfer, and signal processing. His scientific contributions include more than 20 publications in peer-reviewed journals and international conferences. He is a member of the Standard Council Canada ISO/TC 135/SC 8 on Thermographic Testing. He was a recipient of several academic and research awards, including the 2015 CAPES Doctoral Thesis Award in Engineering, the 2015 UFSC Honorable Mention Award, Emergent Leaders of the Americas Award from the Ministry of Foreign Affairs and International Trade of Canada, and the Best Presentation Award from 7th International Workshop Advances in Signal Processing for NDE of Materials. He is also the Chair of the Program Committee of the CREATE NSERC Innovative Program on NDT.

XAVIER P. V. MALDAGUE (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Université Laval, Quebec City, QC, Canada, in 1982, 1984, and 1989, respectively. He has been a Full Professor with the Department of Electrical and Computing Engineering, Université Laval, since 1989, where he was the Head of the Department, from 2003 to 2008 and 2018. He has trained over 50 graduate students (M.Sc. and Ph.D.) and has more than 300 publications. His research interests include infrared thermography, nondestructive evaluation (NDE) techniques, and vision/digital systems for industrial inspection. He is an Honorary Fellow of the Indian Society of Nondestructive Testing. He is also a fellow of the Canadian Engineering Institute, the American Society of Nondestructive Testing, and the Alexander von Humbolt Foundation, Germany. He holds the Tier 1 Canada Research Chair in Infrared Vision. He has been the Chair of the Quantitative Infrared Thermography (QIRT) Council, since 2004.