Research Article

Multimodal Registration and Fusion for 3D Thermal Imaging

Moulay A. Akhloufi[1] and Benjamin Verney[2,3]

[1]*Electronics Engineering, Universidad Técnica Federico Santa Maria (UTFSM), 239-0123 Valparaíso, Chile*
[2]*Computer Vision and Systems Lab., Electrical and Computer Engineering, Quebec City, QC, Canada G1V 0A6*
[3]*Laboratoire Electronique, Informatique et Image, Bourgogne University, 71200 Le Creusot, France*

Correspondence should be addressed to Moulay A. Akhloufi; akhloufi@gel.ulaval.ca

3D vision is an area of computer vision that has attracted a lot of research interest and has been widely studied. In recent years we witness an increasing interest from the industrial community. This interest is driven by the recent advances in 3D technologies, which enable high precision measurements at an affordable cost. With 3D vision techniques we can conduct advanced manufactured parts inspections and metrology analysis. However, we are not able to detect subsurface defects. This kind of detection is achieved by other techniques, like infrared thermography. In this work, we present a new registration framework for 3D and thermal infrared multimodal fusion. The resulting fused data can be used for advanced 3D inspection in Nondestructive Testing and Evaluation (NDT&E) applications. The fusion permits the simultaneous visible surface and subsurface inspections to be conducted in the same process. Experimental tests were conducted with different materials. The obtained results are promising and show how these new techniques can be used efficiently in a combined NDT&E-Metrology analysis of manufactured parts, in areas such as aerospace and automotive.

1. Introduction

Quality control of manufactured parts is becoming increasingly critical in many areas. This increase is driven by the need for higher quality and more robust products, especially in areas where safety is of critical importance, such as in the aerospace and automotive industries. Over the years, many techniques have been developed for products inspection. The research community has developed new algorithms and techniques and the industry has developed and commercialized new technologies for sensing and capturing the signals used in these inspections.

The inspection of an object can be conducted in order to detect flaws at the surface or subsurface levels. At a surface level, 2D and 3D vision can be used, while at a subsurface level more sophisticated techniques like thermography are needed. Nondestructive Testing and Evaluation (NDT&E) techniques are often used when we want to preserve the integrity of the object during the inspection [1]. In this paper, we are interested in the following vision-based nondestructive approaches: 3D computer vision and infrared thermal imaging.

3D vision is an area of computer vision that has attracted increasing interest in recent years from both the research community and the industry. 3D inspection, metrology, CAD matching, geometric dimensioning, and tolerancing (GD&T) are widely used in industrial applications in areas such as aerospace, automotive, manufacturing, mining, and energy [2–30]. 3D vision captures and measures three-dimensional visible surface features. It can perform inspections at a very high accuracy ($\sim$50–100 μm in controlled conditions). The captured data of an object can be compared to its CAD model using 3D pattern matching techniques in order to estimate the deviation of the manufactured part from its ideal model and detect defective areas. However, 3D vision cannot detect subsurface defects. In order to achieve this last goal, different NDT&E techniques are used [1, 2, 31]: ultrasonic testing, Eddy-Current, thermography, and so forth.

Thermography is a very popular NDT&E technique that can be used for image subsurface flaws of different materials [1, 2, 31]. This increasing popularity is mainly due to the decrease in infrared thermal camera prices, the increase in infrared sensors resolution, and the new image processing techniques specifically developed for this type of application.
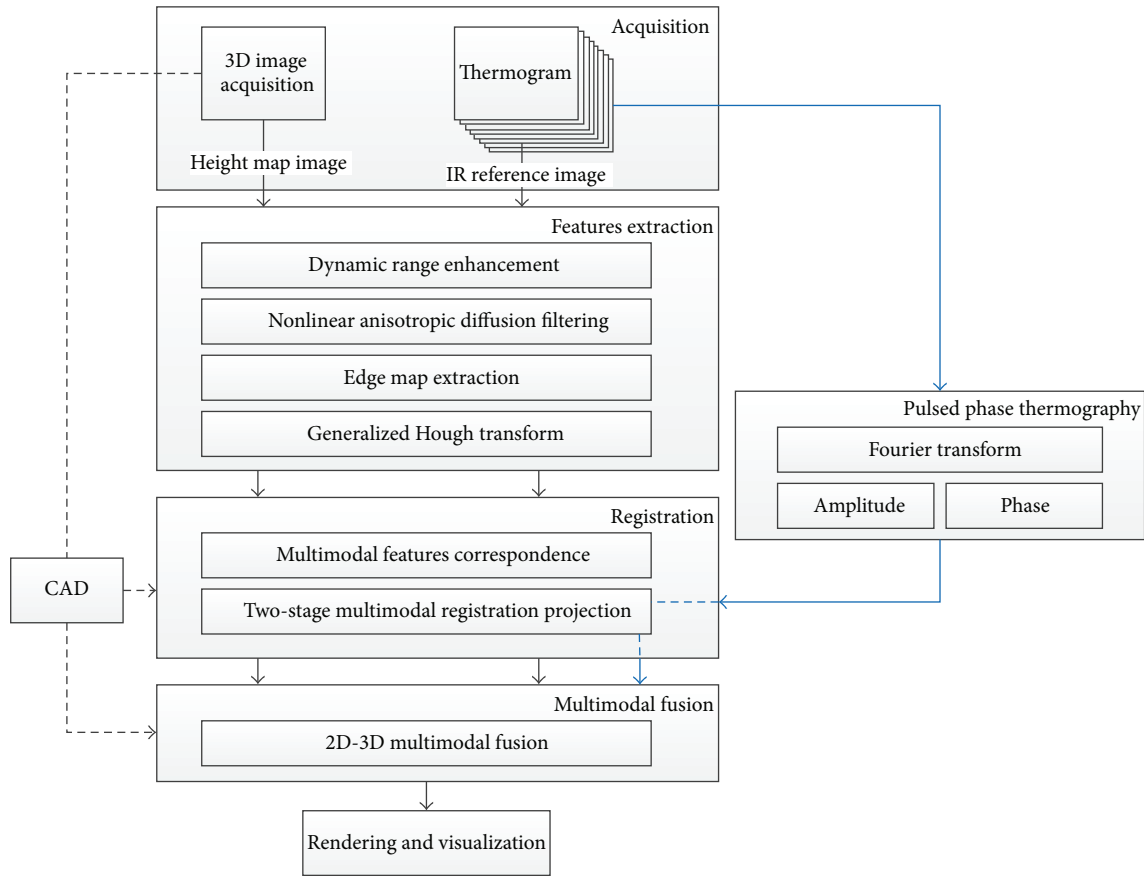
FIGURE 1: Architecture of the multimodal registration and fusion framework.

These two modalities are used separately in many inspection applications. The two approaches can be used to inspect the same object and lead to a separate analysis for the same inspection and quality control task. Combining these modalities can be very useful in inspection applications and NDT&E [32, 33] in aerospace, automotive, and other industries.

This work presents a new framework for multimodal fusion of 3D data and infrared thermal images for Nondestructive Testing and Evaluation. The proposed framework permits the extraction of nonvisible subsurface defects from infrared thermograms and their mapping over the 3D model of the inspected part. Features detection, multimodal registration, and multimodal fusion algorithms are proposed in order to achieve this goal. The resulting fused image is then processed and rendered for 3D visualization.

2. Proposed Framework

The proposed framework enables the fusion of 3D captured data and the 2D thermal infrared images. A new multistage registration algorithm is presented. It uses corresponding detected features to estimate the projective mapping in order to align the 2D infrared and the 3D data prior to fusion. In order to build the final fused model, several steps are necessary. Figure 1 gives an overview of the algorithmic steps involved in this framework. They can be divided into the following sections: data acquisition, pulsed phase thermography processing, features extraction, multimodal registration, and fusion.

(1) Data acquisition:

    (a) 3D images are captured using a 3D camera. The resulting 3D data are fused prior to point cloud meshing and rendering in order to build the final 3D model of the digitized object.

    (b) Thermograms (sequence of thermal infrared images) are captured using an infrared camera. Active thermography is used. A heat source is part of the acquisition setup and is used in order to heat the inspected part's surface while the sequence of infrared images is captured.

(2) Pulsed phase thermography (PPT) processing: frequency domain analysis is conducted in order to visualize deeper defects. These defects become visible at different frequencies (the frequency is correlated with the depth of the detected defect).

(3) Features extraction: multimodal corresponding features are extracted and used for computing the registration transform between the 2D infrared images and the 3D data space. Several image processing steps are necessary in order to extract these features.

(4) Multimodal registration and fusion: the extracted features are used in the estimation of the multimodal registration. The features are matched across the modalities and a piecewise image transform is computed. The obtained transform permits the projection of the infrared images into the 3D space of the object.

The resulting fused data are rendered and visualized using three-dimensional computer graphics techniques.

3. 3D Computer Vision

A wide range of 3D techniques and technologies are available today (contact and noncontact techniques, optical, projection based, phase shifting, interferometry, etc.). 3D vision is among them and is an important area of research in computer vision [2–30, 32].

In recent years, technological advancements in 3D vision technology have made this field widely adopted by the industry. In general, 3D vision approaches can be divided into passive and active techniques.

Passive techniques compute the three-dimensional information from the acquired image data without external active sources. Stereovision [7] is among them and is widely used when the accuracy of the captured data is less important or falls within the stereo limited precision (fire tracking [8], face modeling [9], vision guided robotics [10], etc.). Other passive approaches include shape from shading [15], shape from defocus [16], photometric stereo [17], and structure from motion [18].

Active techniques use an external active source that projects a specific waveform onto the surface of the object. The camera captures the projected waveform and constructs the 3D model of the object. This category includes laser triangulation [11], time of flight [19], phase shifting [12, 13, 20], moiré interferometry [21], and structured light projection [22]. These techniques can achieve higher precision than their passive counterpart.

The accuracy achieved by the active 3D techniques made them increasingly attractive to the industry [23]. Active techniques are also very popular in gaming and entertainment [24–26]. Laser triangulation based systems are the most popular in metrology, CAD matching, and 3D inspection applications. A wide range of products are available commercially and each year new technologies emerge with an increase in precision, flexibility, and robustness. A 3D laser scanner will be used for the capture of 3D data in this work.

A 3D laser scanner system works as follows [27, 28]: a laser point or line is projected over the object surface and the reflected light is measured by a camera. The displacement of the laser in the camera optical plane is proportional to the object height. The laser line forms a plane that intersects the object and captures only a linear slice of it. In order to capture the full 3D image of the object, it is necessary to move the object or the camera.

3.1. 3D Reconstruction. An active 3D laser triangulation scanner is composed of a laser source (point or line), a camera (sensor), and the necessary optical lenses. Given the geometrical properties of such a system, we can establish the relation between the height of the object z and the corresponding displacement p in the sensor relative to its position at a predefined zero height. The 3D coordinates of the object are computed using the following equations [27, 28]:

$$z = \frac{df}{p + f \tan(\theta)},$$
$$x = z \tan(\theta),$$

(1)

where p is the position of the laser image on the sensor, θ is the laser deflection angle, d is the baseline between the lens and the laser source, and f is the focal length of the camera.

The error σ_z in the estimation of z is inversely proportional to the camera baseline d and the focal length f of the camera lens and is proportional to the square of the distance [27, 28]:

$$\sigma_z \approx \frac{z^2}{fd} \sigma_p,$$

(2)

where σ_p is the error in the estimation of position p.

We can achieve higher accuracies by carefully choosing the right sensor resolution, the operational distance, and the working field of view. The laser point can be rotated to scan the object slice along the x-axis or a laser line can be used to get similar results.

3.2. Calibration. In order to compute the position and height of the object surface, a calibration of the triangulation system must be performed. Different calibration techniques can be used in order to compute the geometrical intrinsic and extrinsic parameters (laser line/plan orientation relative to the camera sensor and camera sensor parameters). Intrinsic camera parameters can be extracted using standard camera calibration techniques [29]. For 3D calibration, a high precision calibration tool can be used in order to extract a look-up table (LUT) and estimate the parametric equation coefficients relating the 3D position to the change in the image coordinates induced by the height of the object above the zero calibration plane [30]. A procedure using a planar object positioned at different heights (X_i, Y_i, Z_i) along the z-axis and the corresponding projections in the image plan (x_i, y_i) can be used for the calibration of the 3D system.

4. Active Infrared Thermography

In the last decade we have seen an increase of interest in thermal infrared imaging. The decrease in infrared camera prices and the improvement in their quality increased the importance of the research in this area. Infrared imaging is widely used in defense, security [34], and NDT [1, 2, 31] applications. In recent years, new applications of infrared imaging are emerging: fire propagation study, face recognition, biometrics, biomedical imaging [35], and so forth.

The infrared spectrum spans the wavelength range between $0.74\,\mu$m and $1000\,\mu$m in the electromagnetic spectrum [36]. The most used spectral bands in infrared imaging
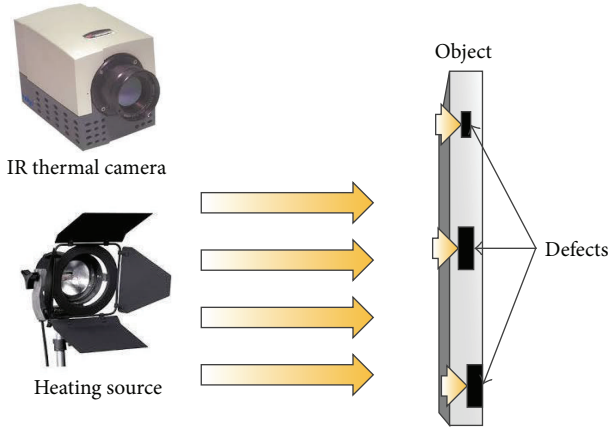
FIGURE 2: Optical pulsed thermography (reflection mode).

FIGURE 3: Absolute thermal contrast: healthy part (solid line), containing defect (dashed line), and absolute thermal contrast (blue line) [43].

are as follows: Near IR (0.74–1 μm), Short-wave IR (1–3 μm), Mid-wave IR (3–5 μm), and Long-wave IR (8–14 μm). Near and Short-wave IR are active infrared spectrums and capture the reflectivity of incident light on the surface, while Mid-wave and Long-wave IR are the thermal infrared spectrums and capture the emitted thermal radiation from the object. These last two thermal spectrums are widely used for imaging nonvisible defects in NDT&E applications.

Infrared thermography is used in NDT&E and allows the visualization of heat patterns on an object or a scene. The theoretical and experimental aspects of this modality were largely studied [37–46].

Thermography can be divided into two modes: passive and active thermography. In passive mode, the captured thermal radiation results from the natural thermal contrast present in the object. A network of veins in a human hand or face creates local thermal contrasts that can be detected using a thermal camera without any additional stimulation. In the active mode, an external source is used in order to stimulate the surface of the object and produce a thermal contrast. This last mode is widely used in NDT&E applications and permits the detection of different type of subsurface flaws: cracks, voids, delaminations, and so forth. In active thermography, different approaches were developed in order to extract the subsurface defects, for example, pulsed thermography, lock-in thermography, and vibrothermography.

In this work we are interested in pulsed thermography and its variant pulsed phase thermography. Figure 2 shows a typical setup for pulsed and pulsed phase thermography in reflection mode where the source and the camera are in front of the inspected part (in transmission mode, the camera is in front of the part and the source is in the opposite side behind the part).

In pulsed thermography a pulse is sent by an external source and the thermal infrared camera captures the sequence of thermal infrared images (thermogram). The flow of energy stimulating the surface of the object will dissipate in the surface of the inspected object and in the presence of a defect having a different thermophysical property compared with the object, we can obtain a measurable thermal contrast
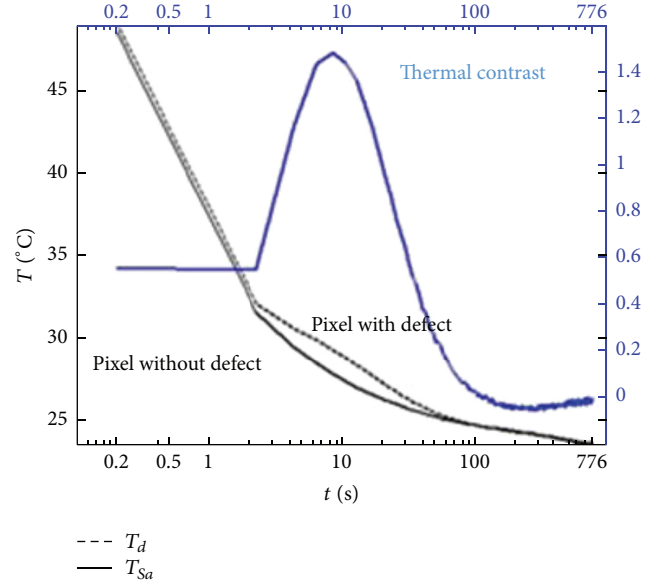
$\Delta T(t)$. The time of appearance of a subsurface defect in the thermogram is proportional to its depth (Figure 3). Consider

$$\Delta T(t) = T_d(t) - T_{Sa}(t), \tag{3}$$

where $T_d(t)$ and $T_{Sa}(t)$ are, respectively, the temperature in the presence and absence of defects.

Pulsed phase thermography (PPT) is an extension to pulsed thermography. PPT is more efficient in the detection of a large number of defects especially those located deeper within the part. Pulsed forms of the thermal wave (left side of Figure 4: square and thermal decay) can be approximated by a sum of sinusoids with frequencies ranging from 0 to ∞. In PPT, the thermogram is processed in the frequency domain using Fourier transform.

The discrete Fourier transform (DFT) can be used to extract amplitude and phase information from pulsed thermograph data. The DFT can be written as [43, 44]

$$F_n = \Delta t \sum_{k=0}^{N-1} T(k\Delta t) \exp^{(-j2\pi nk/N)} = \text{Re}_n + \text{Im}_n, \tag{4}$$

where j is the imaginary number ($j^2 = -1$), n designates the frequency increment ($n = 0, 1, \ldots N$), Δt is the sampling interval, and Re and Im are the real and the imaginary parts of the transform, respectively.

After decomposition in real and imaginary part of the complex transform, we can estimate the amplitude and the phase [44]:

$$A_n = \sqrt{\text{Re}_n^2 + \text{Im}_n^2},$$

$$\phi_n = \tan^{-1}\left(\frac{\text{Im}_n}{\text{Re}_n}\right). \tag{5}$$
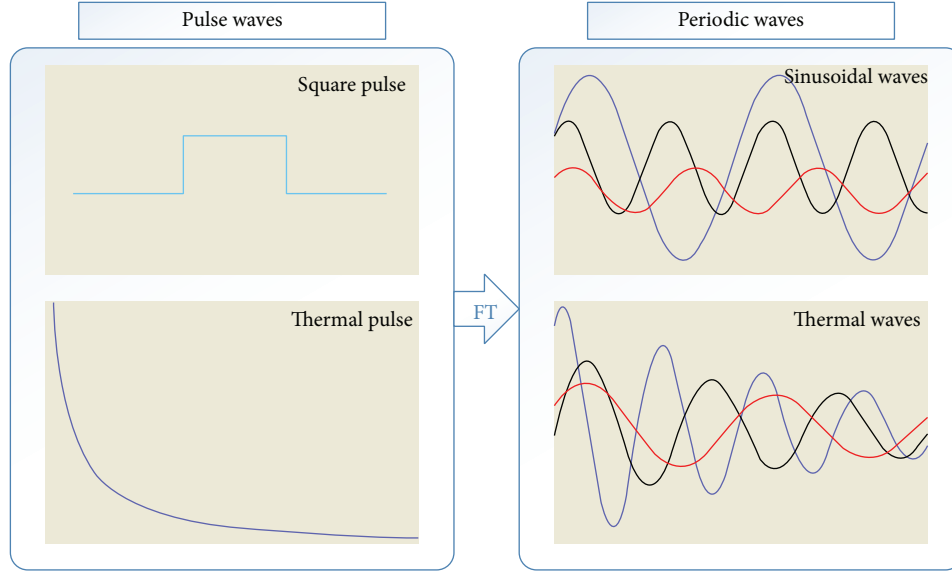
FIGURE 4: Periodic thermal wave decomposition.

Subsurface defects are the most important data to visualize using thermal infrared. Using pulsed phase thermography (PPT) we can extract deeper defects than conventional techniques. Additionally, if the thermal characteristic of the inspected part is known, we can estimate the defects depth using the computed phase of the Fourier transform of the thermal image.

5. Features Detection

The multimodal registration between 2D infrared images and 3D needs the extraction of corresponding features in the two modalities. This step is achieved using the Hough transform [32, 33, 47, 48] of edges computed using preprocessed 3D height map and infrared images.

5.1. Preprocessing. The quality of infrared images is influenced by the amount of thermal noise affecting the sensor. Height map images can also present some noisy data. The noise can decrease the performance of the feature detection and render it less robust. Image enhancement filters help reduce the noise effect in these images.

The 3D height map and the infrared phasegram images have a very low contrast. This makes it difficult to extract the features of interest. The first step consists in enhancing the dynamic range of these images using contrast stretching [49, 50]. Consider

$$
I'(\vec{x}) = \left[\frac{\left(I'_{\max}(\vec{x}) - I'_{\min}(\vec{x})\right)}{P_{\max} - P_{\min}}\left(I(\vec{x}) - P_{\min}\right) + I'_{\min}(\vec{x})\right], \tag{6}
$$

where $I(\vec{x})$ represents the original image; $I'(\vec{x})$ is the resulting image; $I'_{\min}(\vec{x})$ and $I'_{\max}(\vec{x})$ are, respectively, the minimum and maximum output intensity value; $P_{\min}$ and $P_{\max}$ are the intensity corresponding, respectively, to the chosen minimum and maximum percentile of the cumulative input image histogram.

The second enhancement step consists in filtering the images using nonlinear anisotropic diffusion filters [51]. This step leads to lower noise images and help in the extraction of cleaner edge maps prior to Hough transform processing. These filters are used for edge preserving image enhancement. The proposed diffusion process encourages intraregion smoothing while inhibiting interregion smoothing. The mathematical framework of the filter is given by the equations below:

$$
\frac{\partial}{\partial t}I'(\vec{x},t) = \nabla \cdot \left(c(\vec{x},t)\,\nabla I'(\vec{x},t)\right), \tag{7}
$$

where $I'(\vec{x},t)$ is the original image; $\vec{x}$ represents the image axes (i.e., (x, y)); t refers to the iteration step.

$c(\vec{x},t)$ is the *diffusion function* given by

$$
c(\vec{x},t) = \exp\left(-\left(\frac{\left|\nabla I'(\vec{x},t)\right|}{\kappa}\right)^{2}\right), \tag{8}
$$

where κ is the *diffusion constant*.

5.2. Edge Map. Feature detection is carried out on edge map images of the different modalities. The features can be important key points of the object or special removable features used during the experiments. The features must be visible in the different modalities. This step involves the use of Canny edge detector [52–54]. This edge detector is an operator that uses a multistage algorithm to detect a wide range of edges in images. It has good localization and uniqueness properties and has low error rate edge detection [52, 53]. This detector operates in six steps [54]. This edge detector shows a very good performance when dealing with

noisy images such as thermal infrared images. Its filtering and spurious edge elimination reduces noise and selects the most interesting edges for the detection process. This filtering enhances the performance and robustness of Hough transform for feature extraction.

5.3. Feature Extraction. Hough transform is a very popular technique in computer vision. It is commonly used to detect and match shapes in images. This technique is based on the original Hough transform algorithm used for detecting an object having a simple analytic equation describing its boundary. Hough transform can be used to detect lines, circles, or ellipses. A general version was developed and is used to detect arbitrary shapes. GHT (generalized Hough transform) [47, 48] supports object boundaries of arbitrary nonanalytic shapes. Instead of using a parametric equation, GHT uses a look-up table to define the relationship between the boundary positions and orientations and the Hough space parameters. In order to detect an object, a model of this object is learned and look-up table values are computed offline.

For example, suppose that we know the shape and orientation of the desired feature. We can specify an arbitrary reference point (x_{ref}, y_{ref}) within the feature, with respect to which the shape (*i.e.,* the distance r and angle β of normal lines drawn from the boundary to this reference point) of the feature is defined. Our look-up table (i.e., *R-table*) will consist of these distance and direction pairs, indexed by the orientation Ω of the boundary.

The Hough transform space is now defined in terms of the possible positions of the shape in the image, that is, the possible ranges of (x_{ref}, y_{ref}). In other words, the transformation is defined by [32, 33, 48]

$$x_{ref} = x + r \cos(\beta),$$
$$y_{ref} = y + r \sin(\beta) \tag{9}$$

(the r and β values are derived from the *R-table* for particular known orientations Ω). If the orientation of the desired feature is unknown, this procedure is complicated by the fact that we must extend the accumulator by incorporating an extra parameter to account for changes in orientation. In our case, the features are of circular form, limited scale changes and we do not need to account for orientation variation. This helps to reduce the scale search in the Hough space and makes the procedure fast and robust.

6. Multimodal Registration and Fusion

The registration between the different modalities needs the use of corresponding extracted features in order to compute the transform that projects infrared images into the 3D space. Registration between 2D and 3D data is a complex task in a general configuration. This type of registration has been used in medical imaging and remote sensing [55, 56]. In medical imaging, the 3D modality is constructed using multiple projections. This type of 3D reconstruction reduces the complexity of the registration between 2D and 3D. In remote sensing the large distance between the image

sensor and the captured data reduces the registration errors. In this work, a two-stage process was developed in order to reduce the distortions caused by the 2D-3D registration process. Furthermore, we will use a 3D height map during the registration process. This 3D height map image is a 2D representation of the 3D data, thus reducing the complexity of the registration process.

6.1. Features Correspondence Estimation. The feature extraction can cause a different number of features to be extracted in 2D (f_i) and 3D (F_j). We need to find the best matching features $\langle f_i, F_j \rangle$ in order to extract the optimal transformation T between the two modalities and eliminate outliers:

$$\langle f_i, F_j \rangle^* = \arg\min_{f_i, F_j} \left(T \langle f_i, F_j \rangle \right). \tag{10}$$

To make the procedure more robust, we developed a new matching strategy based on the use of a set of neighboring feature triplets instead of individual features. The objective is to solve for the best correspondence between feature triplets in 2D and 3D. The approach is iterative and the algorithmic steps are given in Figure 5. In this procedure, we use the 2D coordinates of the height map (and we ignore the z coordinate). We are interested in a spatial registration of the 2D image with the projected 3D image. The 3D coordinates are recovered later when the fused data are projected in 3D space.

Given a set of 3D features triplets F_j ($j = A_1, B_1, C_1$), we select a set of 2D features triplets f_i ($i = A_2, B_2, C_2$) and compute the center of gravity (COG) of the 3D (Cog$_{3D}$) and 2D (Cog$_{2D}$) set of the N features ($N = 3$):

$$Cog_{3D} = \frac{\sum_{ABC_1} F_j}{N};$$
$$Cog_{2D} = \frac{\sum_{ABC_2} f_i}{N}. \tag{11}$$

Once the centers of gravity computed, we align the two corresponding coordinates (we move the 2D COG coordinates Cog$_{2D}$ and align them with the 3D COG coordinates Cog$_{3D}$):

$$Cog_{3D} \equiv Cog_{2D}. \tag{12}$$

From this new reference position we solve for the best affine transformation between different feature correspondences. The best transformation is the one that minimizes the Euclidian distance d between the features coordinates resulting from the affine registrations of the 2D features triplet with the 3D features triplet:

$$A^* = \arg\min_{A_k} \left(d \left(F_j - A f_i \right) \right). \tag{13}$$

The obtained transformation A^* establishes the best feature triplets correspondence and serves as an initialization for the next feature correspondences. If multiple points are located in a certain area, the triplet configuration is not unique and we have to identify the best triplet possible using
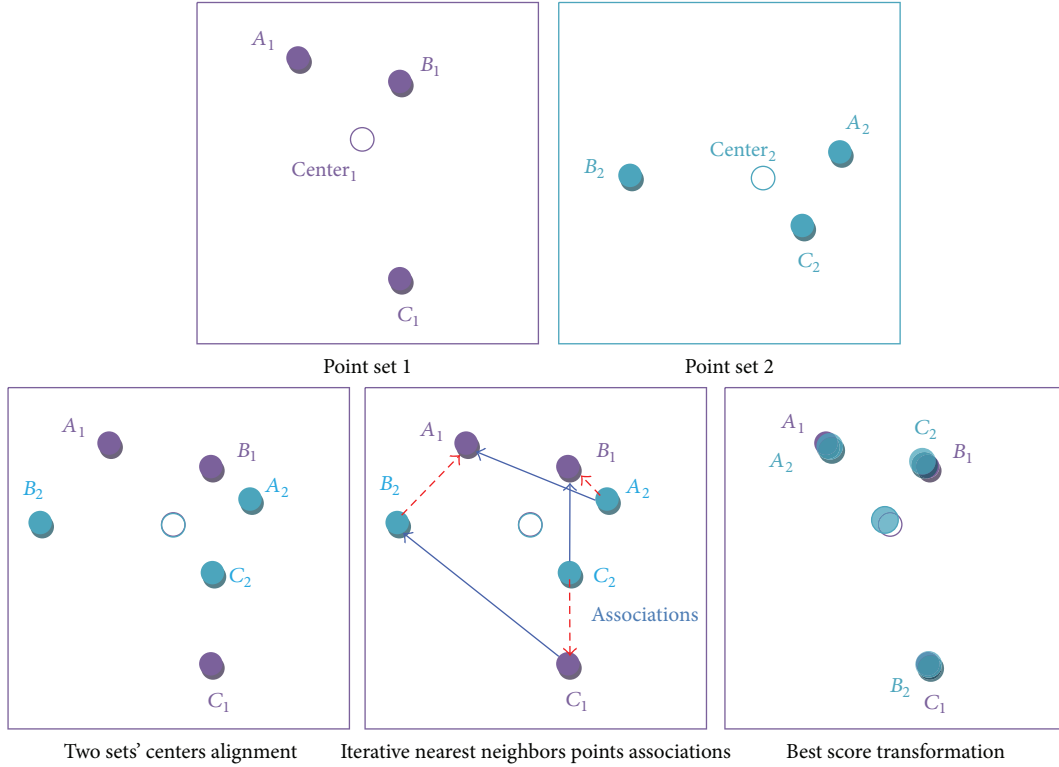
FIGURE 5: Features correspondence steps: the iterative process solves for the best correspondence that minimizes the distance between feature triplets. An example of the iterative step is given: the red arrows established the first association and computed the distance score. This procedure is repeated until the best association is found (arrows in blue). This last association is the one that minimized the distance between the feature triplets under an affine transform (best score).

a computed score. The score is the distance computed above (the lower the distance, the better the correspondence). The best score (lower interfeatures distance) is the one that globally minimizes the transformations between the sets of selected points. This procedure is iterative and tries to solve for the best correspondence between the feature triplets in the image.

This features matching strategy gives a robust correspondence estimation and eliminates outliers present in the two modalities. However, in the presence of a large number of points, it can be time consuming. The use of information available about the experimental setup can help reduce the processing time by initializing the search using a relative orientation between the 2D and 3D cameras. Additionally, if the first selected features give a successful correspondence, their alignment matrix will help accelerate subsequent steps.

6.2. 3D Thermal Infrared Registration.
The 2D-3D corresponding features are used in a two-stage registration estimation approach. The first stage uses the homogeneous coordinates of corresponding features in the 3D height map and the 2D infrared image in order to solve for the perspective projective mapping [57, 58]. This projection can give an ideal solution when we deal with planar objects; however distortions arise when objects with complex shapes are used. In order to reduce these distortions, we solve for the best transformation in two steps. In the first step, we select the most external corresponding features $\langle f_i^c, F_j^c \rangle$ (close to the external contours of the object), so that the projection matrix T_1 is given by minimizing for

$$F_j^c = T_1 f_i^c. \tag{14}$$

In the second stage, a projection matrix T_2 is estimated using all of the corresponding features. We get T_2 by minimizing for

$$F_j = T_2 f_i. \tag{15}$$

The final projection matrix T is given by the best transformation minimizing a weighted Euclidian distance between the corresponding features. Consider

$$T = \arg\min_{T_{1,2}} \left(\sum w_e d \left(T_{1,2}\left(f_i\right), T_{1,2}\left(F_j\right)\right)\right) \tag{16}$$

with w_e being the weights chosen in order to minimize the distortion for the features located along the principal edges of the object.

A second strategy is used in order to reduce the distortion of internal parts of the object. A piecewise linear transformation is used. In this transformation, affine projections are estimated and applied to small triangular regions of the image [59].

(1) Divide a set of N control features into corresponding triangle meshes $[f_{ijk}; F_{ijk}]$ (Delaunay triangulation is the most commonly used technique for this division):

$$M = \left\{ \left[f_{ijk}; F_{ijk} \right] \mid ijk = 0 \cdots N \right\}. \tag{17}$$

(2) Using the three corresponding vertices $\{i, j, k\}$, compute the affine mapping A_{ijk} between the corresponding meshes:

$$F_{ijk} = A_{ijk} f_{ijk}. \tag{18}$$

2D-3D multimodal registration is a complex process that can lead to high distortions during the projection mapping. The proposed approach improves the accuracy of features positioning by first projecting the image as close as possible to its final location. Then a local piecewise transform refines the mapping of the remaining data.

6.3. Multimodal Fusion. The registration process extracts the transformations necessary to align the infrared images into the 3D space. After the alignment of the images, a fusion process is conducted.

Given a 2D infrared image I_{2D} (original, amplitude, or phase image) and the 3D height map image I_{3D}, the 2D image is projected using the process described above. Let T_A be a transform representing the registration process that projects the image coordinates (x, y) into the new coordinates (x', y'):

$$I'_{2D}\left(x', y'\right) = T_A\left(I_{2D}\left(x, y\right)\right). \tag{19}$$

The fused 3D height map is given by fusing the height map image I_{3D} with the projected 2D infrared image I'_{2D}. Consider

$$I'_{3D}\left(u, v\right) = F\left(I_{3D}\left(u, v\right), I'_{2D}\left(x', y'\right)\right), \tag{20}$$

where (u, v) are the 2D coordinates of the height map matrix corresponding to the projected 2D coordinates (x', y').

The fusion operator F can take different functions depending on the visualization/analysis objective. In this work, we fuse a set of infrared images (original images, phasegram, etc.) i ($i = 1 \cdots N$) with the 3D data for flaws visualization and analysis:

$$F_1 \longrightarrow I'_{3D}\left(u, v, \alpha_{3D}\right) = \left[I'_{2D}\left(x', y', \alpha_{2D}\right)\right]_i, \tag{21}$$

where α_{2D} and α_{3D} are the transparency coefficients.

The fusion can be carried out using 3D scanned data or its corresponding CAD model. The resulting fused 3D height map is then rendered for visualization in 3D space.

7. Experimental Results

3D image acquisition was conducted using a laser triangulation scanning system from Sick [14] and a motorized linear axis from Zaber [60]. The 3D system has a height accuracy of 0.2 mm. Infrared acquisition was conducted using FLIR Phoenix thermal camera [61] operating in the MWIR spectrum of 2.5–5 μm with a resolution of 640×512 pixels. The thermal active setup used halogen projectors as a heating source. In order to improve the features detection performance, we used removable aluminum features. These features are of circular shape (two concentric circles) and have 1 mm width. These characteristics make them visible in the two modalities (3D and thermal infrared). The algorithms were implemented using C++ and Matlab. The 3D computer graphics algorithms (modelling, rendering, and visualization) were implemented using OpenGL and GLUT [62, 63].

The proposed framework permits the multimodal fusion of the captured images. The resulting fused data is a multimodal fused height map image. In order to visualize the results in 3D space, computer graphics modeling and rendering techniques are performed. Geometric primitives are created using height map 3D data converted into 3D point clouds. In this step triangular meshes are generated. The resulting 3D mesh proceeds through a modelling step that places the resulting 3D information in the 3D world coordinate system. The original fused height map contains the texture data used in the final rendering of the resulting object.

The final 3D object can be used to visualize the defects and other metrological information at different 3D viewpoints.

Tests were conducted using different types of material (aluminum, steel, carbon, and fiber glass composites). Figure 6 shows the main results of the steps involved in the proposed framework.

Figure 7 provides some examples of the multimodal fusion results with different parts.

Figure 8 shows the possibility of visualizing the 3D thermogram at different viewpoints (some defects are best viewed at specific angles).

Figure 9 shows the evolution of the 3D thermogram, and we can see that the subsurface defects become visible over time depending on their depth.

Figure 10 shows an example of the estimated positions of the defects overlaid on the 3D data. Samples with known 2D defect positions were used in benchmarking the estimation of the computed positions of the defects. The mean position error was around 2.1 mm. While the precision can be sufficient in many NDT&E inspections where only the approximate positions of the defects are needed and their visualization, it can be further improved using different approaches such as

(1) higher resolution thermal infrared cameras;

(2) better image quality using high quality thermal sensors and cameras with lower noise;

(3) more features used during the registration and fusion (when available in the inspected object);

(4) local registration and fusion by using a scanning strategy (e.g., a multimodal robotic NDT&E system as illustrated in Figure 11).
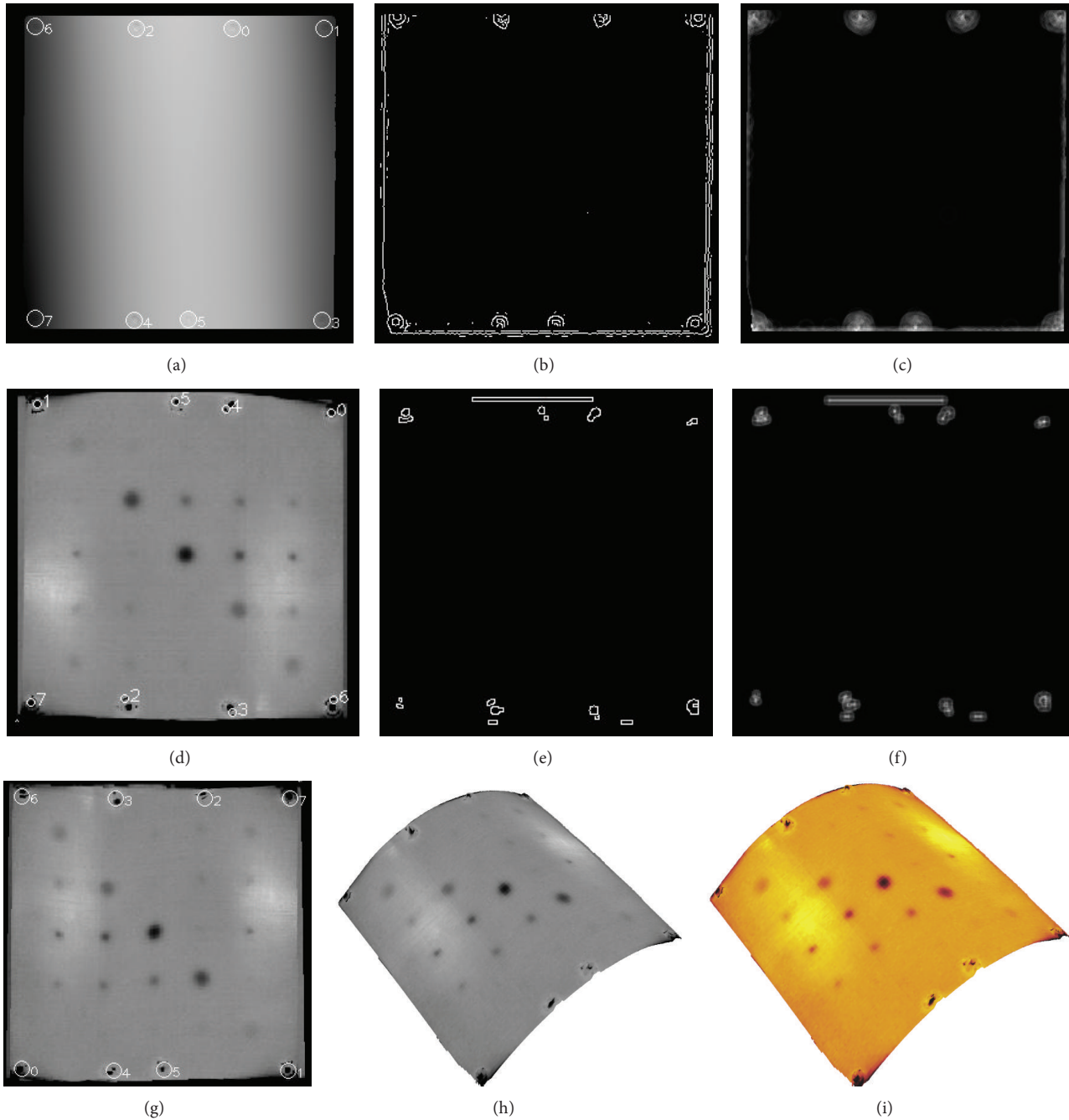
FIGURE 6: Steps of the multimodal fusion framework: (a) 3D image with the detected circular features; (b) the edge map of the 3D filtered image; (c) the Hough space used to detect the 3D image features; (d) thermal infrared image with the detected circular features; (e) the edge map of the filtered infrared phase image; (f) the Hough space used to detect the infrared image features; (g) the registered infrared image projected using the registration transform prior to warping in the 3D space; (h) rendering of the multimodal fused IR-3D data; (i) color rendered data.

8. Conclusion

This work presents a new framework for multimodal registration and fusion of 3D and thermal infrared images for Nondestructive Testing and Evaluation (NDT&E) applications. The proposed architecture leads to a 3D thermal imaging system which can be used in combined NDT&E and 3D inspection applications. With this system, we can perform various tasks such as

(1) visualizing 3D information at different viewpoints;

(2) conducting geometric dimensional metrology analysis;

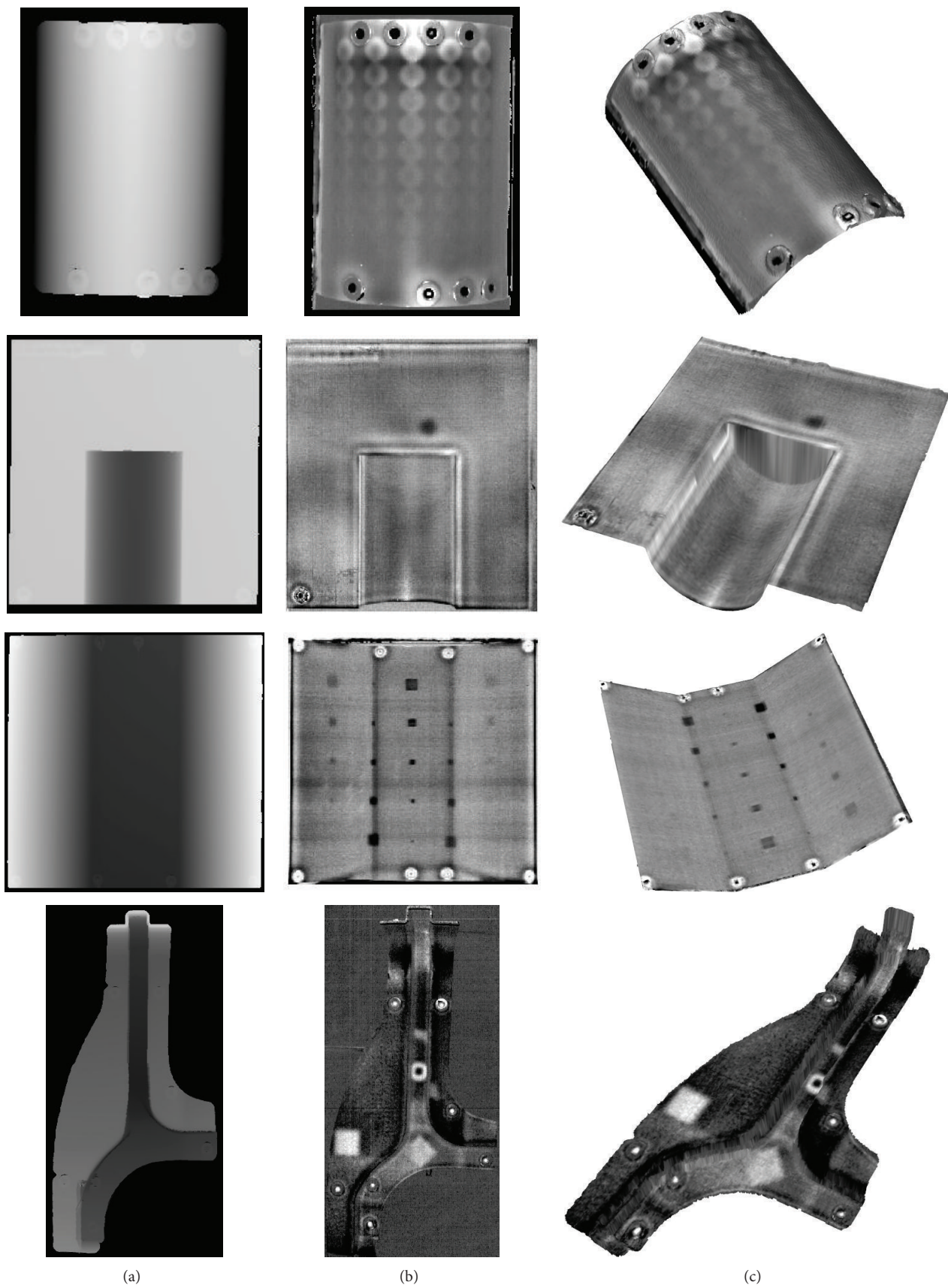(3) performing 3D quality control;

FIGURE 7: Examples of multimodal fusion results (from (a) to (c): 3D image, thermal IR image, and multimodal fusion and rendering).
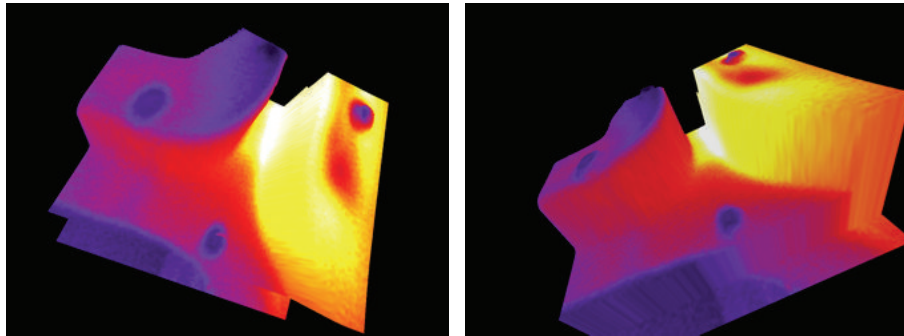
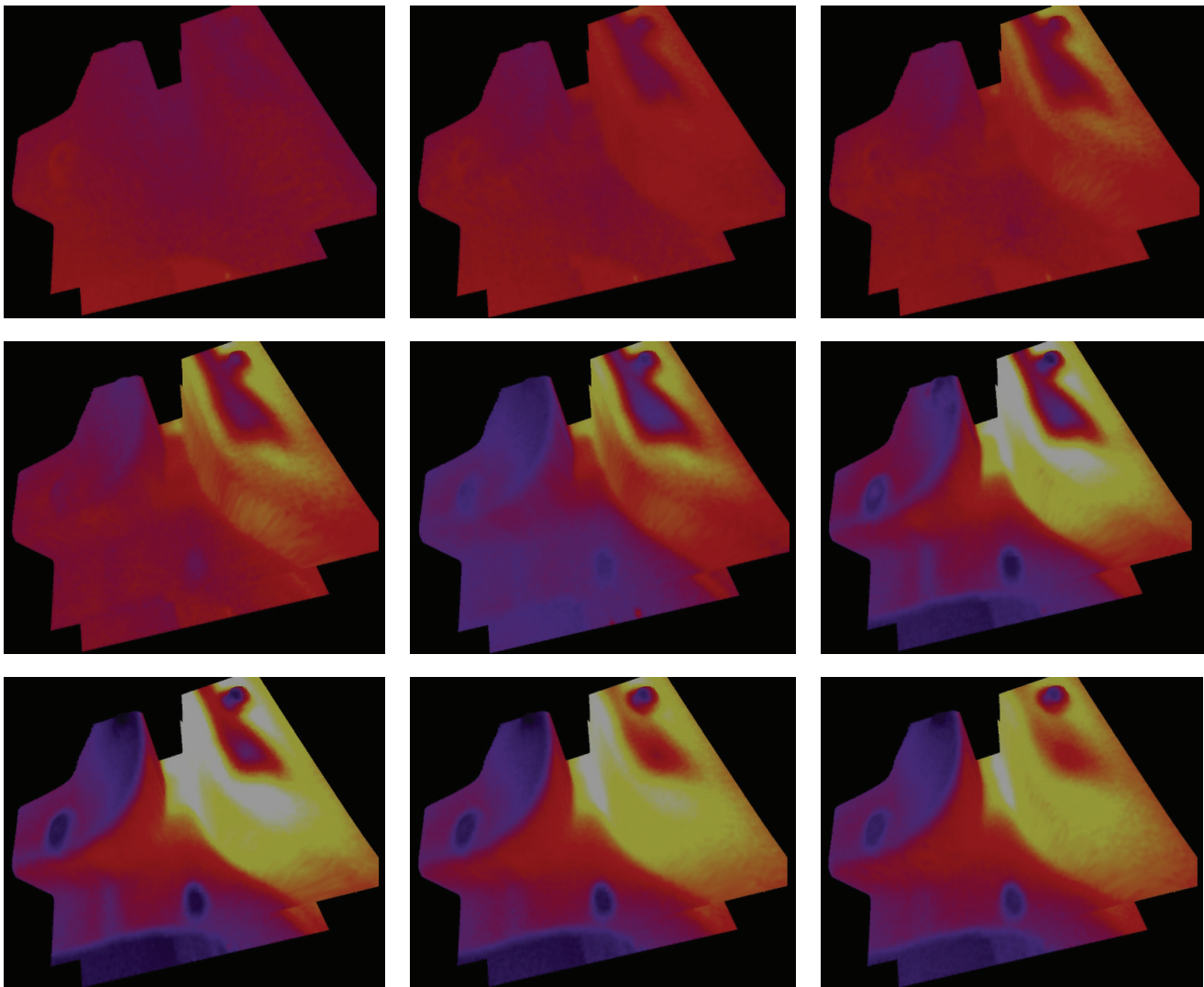FIGURE 8: Two 3D views of the same inspected object.

FIGURE 9: Visualization example: with evolution of the 3D thermogram over time, the defects become visible depending on their depth.

(4) inspecting, evaluating, and quantifying subsurface defects by using an appropriate mode (e.g., thermography).

The approach operates on each modality in order to extract salient features and compute the registration transform.

A two-stage multimodal registration is proposed and permits obtaining higher performances. The estimated registration projection is used for multimodal fusion of 3D height map data and infrared thermal images.

The proposed techniques were successfully tested using 3D vision and active infrared thermography. They can be

Figure 10: Examples of estimated position of the defects overlaid on the 3D data.
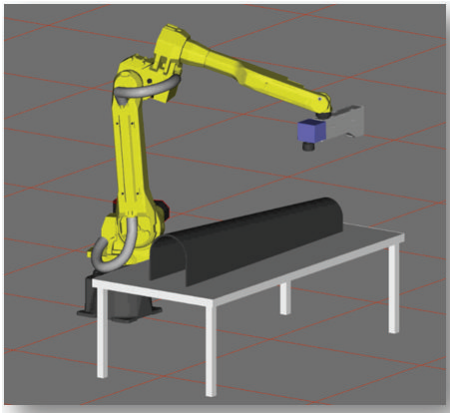

Figure 11: Simulation environment of a multimodal robotic NDT&E system.

easily extended to take into account other NDT&E or inspection modalities. The proposed framework is built as a modular architecture and is easily adaptable to different 3D systems, infrared imaging spectrums, different imaging modalities, and other NDT&E techniques. Also, with the knowledge of thermal properties of the inspected parts, we can estimate the defects depth and visualize them overlaid in three-dimensional space using this framework. Future work includes the development of an automatic flaw detection algorithm operating on pulsed phase thermography images, an estimation of extracted defects depth with their visualization in 3D space, and the integration of the developed framework in a robotic path planning and control system for large parts inspection.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] X. P. V. Maldague, "Introduction to NDT by active infrared thermography," *Materials Evaluation*, vol. 60, no. 9, pp. 1060–1073, 2002.

[2] NDT Resource Center, *NDT Course Material: Ultrasonic Testing, Eddy Current Testing, Infrared/Thermal Testing (IRT)*, 2015, https://www.nde-ed.org/EducationResources/CommunityCollege/communitycollege.htm.

[3] E. Trucco and A. Verri, *Introductory Techniques for 3D Computer Vision*, Prentice-Hall, 1998.

[4] S. Zhang, *Handbook of 3D Machine Vision: Optical Metrology and Imaging*, Series in Optics and Optoelectronics, Taylor & Francis, New York, NY, USA, 2013.

[5] C. Wöhler, *3D Computer Vision Efficient Methods and Applications*, X Media Publishing, 2nd edition, 2013.

[6] B. Cyganek and J. P. Siebert, *An Introduction to 3D Computer Vision Techniques and Algorithms*, Wiley-Blackwell, 2009.

[7] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1–3, pp. 7–42, 2002.

[8] M. A. Akhloufi and L. Rossi, "Three-dimensional tracking for efficient fire fighting in complex situations," in *Visual Information Processing XVIII*, vol. 7341 of *Proceedings of SPIE*, p. 12, April 2009.

[9] N. Uchida, T. Shibahara, T. Aoki, H. Nakajima, and K. Kobayashi, "3D face recognition using passive stereo vision," in *Proceedings of the IEEE International Conference on Image Processing 2005 (ICIP '05)*, pp. 950–953, September 2005.

[10] Z. Kowalczuk and D. Wesierski, "Vision guided robot gripping systems," in *Automation and Robotics*, J. M. Ramos Arreguin, Ed., InTech, 2008, http://www.intechopen.com/books/automation_and_robotics/vision_guided_robot_gripping_systems.

[11] J. G. D. M. França, M. A. Gazziro, A. N. Ide, and J. H. Saito, "A 3D scanning system based on laser triangulation and variable field of view," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '05)*, vol. 1, pp. 425–428, September 2005.

[12] S. Zhang, D. van der Weide, and J. H. Oliver, "Superfast phase-shifting method for 3-D shape measurement," *Optics Express*, vol. 18, no. 9, pp. 9684–9689, 2010.

[13] Y. Wang and S. Zhang, "Superfast multifrequency phase-shifting technique with optimal pulse width modulation," *Optics Express*, vol. 19, no. 6, pp. 5149–5155, 2011.

[14] Sick, "IVP 3D," 2015, http://www.sick.com/group/en/home/products/product_portfolio/vision/pages/high-end_cameras.aspx.

[15] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah, "Shape from shading: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp. 690–706, 1999.

[16] P. Favaro and S. Soatto, "A geometric approach to shape from defocus," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 406–417, 2005.

[17] C. Hernandez, G. Vogiatzis, and R. Cipolla, "Multiview photometric stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 3, pp. 548–554, 2008.

[18] F. Dellaert, S. M. Seitz, C. E. Thorpe, and S. Thrun, "Structure from motion without correspondence," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '2000)*, vol. 2, pp. 557–564, June 2000.

[19] L. Li, "Time-of-flight camera—an introduction," Technical White Paper SLOA190B, Texas Instruments, Dallas, Tex, USA, 2014, http://www.ti.com/lit/wp/sloa190b/sloa190b.pdf.

[20] X. Han, *Three-dimensional shape measurement based on the phase shifting and stereovision methods [Ph.D. thesis]*, State University of New York at Stony Brook, Stony Brook, NY, USA, 2010.

[21] C. Li, Z. Liu, H. Xie, and D. Wu, "Novel 3D SEM Moiré method for micro height measurement," *Optics Express*, vol. 21, no. 13, pp. 15734–15746, 2013.

[22] J. Geng, "Structured-light 3D surface imaging: a tutorial," *Advances in Optics and Photonics*, vol. 3, no. 2, pp. 128–160, 2011.

[23] H. Pottmann, S. Leopoldseder, and M. Hofer, "Simultaneous registration of multiple views of a 3D object," in *Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, part 3A, Commission III, pp. 265–270, 2002.

[24] Microsoft Kinect, 2015, http://www.microsoft.com/en-us/kinectforwindows/.

[25] Asus Xtion PRO, 2015, http://www.asus.com/Multimedia/Xtion_PRO_LIVE/.

[26] Intel Real Sense, 2015, http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-3d-camera.html.

[27] J.-A. Beraldin, F. Blais, L. Cournoyer, G. Godin, and M. Rioux, *Active 3D Sensing*, QUADERNI 10, NRC Canada, Centro di Ricerche Informatiche per i Beni Culturali, Scuola Normale Superiore di Pisa, 2000, http://www1.cs.columbia.edu/~allen/PHOTOPAPERS/beraldin.pdf.

[28] J.-A. Beraldin, F. Blais, L. Cournoyer, G. Godin, M. Rioux, and J. Taylor, "Active 3D sensing," in *Proceedings of the e-Way into the Four Dimensions of Cultural Heritage Congress*, NRC 46495, Vienna, Austria, April 2003.

[29] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[30] A. Picón, A. Bereciartua, J. A. Gutiérrez, and J. Pérez, "3D high precision tube bevel measurement using laser based rotating scanner," in *Proceedings of the IEEE Conference on Emerging Technologies and Factory Automation (ETFA '06)*, pp. 1190–1197, September 2006.

[31] NDT Resource Center, "Introduction to NDT," 2015, https://www.nde-ed.org/GeneralResources/IntroToNDT/GenIntroNDT.htm.

[32] M. A. Akhloufi and B. Verney, "Fusion framework for 3D inspection and thermal NDT," SAE Technical Paper 2013-01-2171, 2013.

[33] M. A. Akhloufi, B. Verney, and O. Aubreton, "Multimodal fusion system for NDT and Metrology," in *Proceedings of the 12th International Conference on Quantitative InfraRed Thermography (QIRT '14)*, pp. 7–11, Bordeaux, France, 2014.

[34] M. A. Akhloufi, C. Porcher, and A. Bendada, "Fusion of thermal infrared and visible spectrum for robust pedestrian tracking," in *Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications XI*, vol. 9076 of *Proceedings of SPIE*, Baltimore, Md, USA, May 2014.

[35] E. F. J. Ring and K. Ammer, "Infrared thermal imaging in medicine," *Physiological Measurement*, vol. 33, no. 3, pp. R33–R46, 2012.

[36] MIVIM, "Electromagnetic Spectrum," 2015, http://mivim.gel.ulaval.ca/dynamique/index.php?idD=58&%20Lang=1.

[37] M. Vollmer and K.-P. Möllmann, *Infrared Thermal Imaging: Fundamentals, Research and Applications*, Wiley-VCH, Weinheim, Germany, 2010.

[38] C. Meola, Ed., *Infrared Thermography Recent Advances and Future Trends*, Bentham eBooks, 2012.

[39] G. C. Holst, *Common Sense Approach to Thermal Imaging*, SPIE Optical Engineering Press, Bellingham, Wash, USA, 2000.

[40] X. Maldague, *Theory and Practice of Infrared Technology for Non Destructive Testing*, John-Wiley & Sons, 2001.

[41] C. Meola and G. M. Carlomagno, "Recent advances in the use of infrared thermography," *Measurement Science and Technology*, vol. 15, no. 9, pp. R27–R58, 2004.

[42] X. Maldague and P. O. Moore, Eds., *Nondestructive Handbook, Infrared and Thermal Testing*, vol. 3, ASNT Press, Columbus, Ohio, USA, 3rd edition, 2001.

[43] C. Ibarra-Castanedo, *Quantitative subsurface defect evaluation by pulsed phase thermography: depth retrieval with the phase [Ph.D. thesis]*, Laval University, 2005, http://theses.ulaval.ca/archimede/fichiers/23016/23016.html.

[44] X. Maldague and S. Marinetti, "Pulse phase infrared thermography," *Journal of Applied Physics*, vol. 79, no. 5, pp. 2694–2698, 1996.

[45] L. D. Favro, X. Han, Y. Wang, P. K. Kuo, and R. L. Thomas, "Pulse-echo thermal wave imaging," in *Review of Progress in Quantitative Nondestructive Evaluation*, D. O. Thompson and D. E. Chimenti, Eds., vol. 14, pp. 425–429, Springer, New York, NY, USA, 1995.

[46] R. L. Thomas, J. J. Pouch, Y. H. Wong, L. D. Favro, P. K. Kuo, and A. Rosencwaig, "Subsurface flaw detection in metals by photoacoustic microscopy," *Journal of Applied Physics*, vol. 51, no. 2, pp. 1152–1156, 1980.

[47] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, Upper Saddle River, NJ, USA, 3rd edition, 2006.

[48] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[49] I. T. Young, J. J. Gerbrands, and L. J. van Vliet, *Fundamentals of Image Processing*, Ver. 2.3, Delft University of Technology, 1995–2007, http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/TUDELFT/FIP2_3.pdf.

[50] E. R. Davies, *Computer and Machine Vision: Theory, Algorithms, Practicalities*, Academic Press, 4th edition, 2012.

[51] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[52] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[53] B. Green, "Canny Edge Detection Tutorial," 2002, http://dasl .mem.drexel.edu/alumni/bGreen/www.pages.drexel.edu/_ weg22/can_tut.html.

[54] A. Jepson, "Edge Detection," March 2015, http://www.cs.toronto .edu/~jepson/csc420/notes/edgeDetection.pdf .

[55] S. Yang, I. Atmosukarto, J. Franklin, J. F. Brinkley, D. Suciu, and L. G. Shapiro, "A model of multimodal fusion for medical applications," in *Multimedia Content Access: Algorithms and Systems III*, vol. 7255 of *Proceedings of SPIE*, San Jose, Calif, USA, January 2009.

[56] J. Le Moigne, N. S. Netanyahu, and R. D. Eastman, *Image Registration for Remote Sensing*, Cambridge University Press, 1st edition, 2011.

[57] G. Wolberg, *Digital Image Warping*, IEEE Computer Society Press, 1990.

[58] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics—Principles and Practice*, Addison-Wesley, 1990.

[59] A. Goshtasby, "Piecewise linear mapping functions for image registration," *Pattern Recognition*, vol. 19, no. 6, pp. 459–466, 1986.

[60] Zaber, "Linear translational stages," 2015, http://www.zaber.com/ products/product_group.php?group=T-LSR.

[61] FLIR, "Phoenix MWIR thermal camera," 2015, http://www.flir .com/legacy/view/?id=51542.

[62] S. Guha, *Computer Graphics Through OpenGL: From Theory to Experiments*, Chapman and Hall/CRC, 2010.

[63] GLUT, "The OpenGL Utility Toolkit," 2015, https://www.opengl .org/resources/libraries/glut/.