Invited Paper

Toward a target-free calibration of a multimodal structured light and thermal imaging system

Eberto Benjumea[a], Raúl Vargas[a], Rigoberto Juarez-Salazar[b], and Andres G. Marrugo[a]

[a]Facultad de Ingeniería, Universidad Tecnológica de Bolívar, Cartagena, Colombia
[b]CONAHCYT-Instituto Politécnico Nacional, CITEDI, Av. Instituto Politécnico Nacional 1310, Nueva Tijuana, Tijuana, B.C., 22435, México.

ABSTRACT

Current calibration methods for multimodal systems consisting of structured light and thermography use calibration targets with physical characteristics. However, defects in the manufacturing of these targets are common. Therefore, these methods are prone to undesired errors. We propose a calibration method for a multimodal system (a visible camera, projector, and thermal imaging camera) that does not require the construction of a physical calibration target. For this purpose, thanks to an auxiliary camera, we use a digital screen to obtain the intrinsic parameters of the camera, and a mirror to obtain the intrinsic and extrinsic parameters of the projector and the thermal imaging camera. The experimental results demonstrate that it is possible to elude the challenging task of fabricating physical targets without compromising the accuracy of the system calibration compared to conventional methods.

Keywords: Structured light, fringe projection profilometry, thermal cameras, thermography, multimodal calibration, target-free

1. INTRODUCTION

The applications of structured light techniques span industrial, cultural, scientific, and medical areas.[1–3] Despite this wide range, its combination with other imaging techniques enhances its use in other applications or allows it to obtain better results.[1,4] The modeling and calibration of cameras are vital in multimodal systems, such as in structured light systems.[5] Calibration of multimodal systems usually implies obtaining the intrinsic and extrinsic parameters of the cameras by using a common and visible calibration target in all the imaging techniques involved (thermography, micro-computed tomography, or other types of image of interest).[6] In the case of multimodal systems composed of structured light and thermography, calibration is challenging due to the incompatibility of visible cameras, projectors and thermal cameras. Recurring problems in the calibration of these multimodal systems involve thermal reflections, restricted spatial/thermal resolution and significant lens distortions in thermographic cameras.[1,7]

Calibration proposals for these systems attempt to mitigate the effects of the problems mentioned above by creating different types of calibration targets. These works usually experiment with flat targets in materials such as aluminum, printed circuit boards, insulating materials, among others.[1,7,8] They also addressed this problem by using asymmetric or symmetric circular patterns or checkerboards.[9–11] However, the fabrication of these targets presents errors, because these processes are often manually or industrially machined. Therefore, the surfaces may often not be flat, but deformed in a concave or convex manner. Similarly, the feature points may be deformed for the same reason.

In this paper, we propose a new direction in the calibration of multimodal systems (a visible camera, a projector and a thermographic camera) that does not require the construction of a physical calibration target.

Further author information: (Send correspondence to Eberto Benjumea)
Eberto Benjumea: E-mail: ebenjumea@utb.edu.co.
Raúl Vargas: E-mail: ravargas@utb.edu.co.
Rigoberto Juarez-Salazar: E-mail: rjuarez@citedi.mx.
Andres G Marrugo: E-mail: agmarrugo@utb.edu.co

Dimensional Optical Metrology and Inspection for Practical Applications XIII, edited by Kevin G. Harding, Song Zhang, Jae-Sang Hyun, Beiwen Li, Andrés G. Marrugo, Proc. of SPIE Vol. 13038, 1303808 · © 2024 SPIE · 0277-786X · doi: 10.1117/12.3013448

In this sense, with the help of an auxiliary camera, we use a digital screen to obtain the intrinsic parameters of the camera, and a mirror to obtain the intrinsic and extrinsic parameters of the projector and the thermal imaging camera. The experimental results demonstrate that it is possible to elude the challenging task of fabricating physical targets without compromising the accuracy of the system calibration compared to conventional methods.

Section 2 explains the calibration and obtaining of data in a conventional multimodal system composed of a thermal camera, projector, and visible camera. Section 3 shows and discusses the common problems in conventional calibrations. Finally, Section 4 shows our new direction in this kind of calibration.

2. CONVENTIONAL MULTIMODAL STRUCTURED LIGHT AND THERMAL IMAGING SYSTEM

The calibration has two stages of acquisition the images. In the first stage, the FPP system is calibrated with a conventional target and then the stereo system consisting of the visible camera and the thermal camera is calibrated with a visible target in both modalities, without fringe projection. In our case, we used asymmetric patterns of circles as calibration targets. The structured light system and the visible camera-thermographic camera system are calibrated by stereo triangulation considering the distortions using the pinhole model. In this way, they obtain the intrinsic and extrinsic parameters of the cameras. Generally, the origin of the world coordinate system is located in the visible camera. In this section, we discuss the phase-shifting algorithm and the stereo vision model used for 3D image reconstruction, and the conversion of phase data into metric measurements.

2.1 Phase-shifting algorithm

Phase-shifting algorithms are robust and can achieve pixel-wise phase measurement with high resolution and accuracy.[12] Within an N-step phase-shifting algorithm featuring uniform phase increments, the representation of the k^{th} fringe image [12] is

$$I = I' + I'' \ \cos(\phi + 2k\pi/N) \ , \tag{1}$$

where I' is the average intensity, I'' represents to the modulation of the intensity, and ϕ corresponds to the phase under consideration. Where ϕ can be obtained using the following equation

$$\phi = -\tan^{-1}\left[\frac{I_k \sin(2k\pi/N)}{I_k \cos(2k\pi/N)}\right] \ . \tag{2}$$

The arctangent function produces a phase with 2π discontinuities. Phase-unwrapping algorithms facilitate the resolution of these discontinuities, determining the optimal value κ that denotes the precise number of 2π increments to be applied at each data point.

$$\Phi = \phi + 2\pi \times \kappa \ , \tag{3}$$

where Φ represents the continuous unwrapped phase.

2.1.1 Stereo vision model

Based on a pinhole camera model, the stereo vision model establishes correspondence between the capture and projection systems and the world coordinate system. This enables the establishment of a relationship between a point in world coordinates $[x^w, y^w, z^w]^T$ and its projection onto the camera's image plane at coordinates $[u^c, v^c]^T$,[13,14] as

$$s^c[u^c, v^c 1]^T = \mathbf{A^c}[\mathbf{R}^c, \ \mathbf{t}^c][x^w, y^w, z^w 1]^T \ , \tag{4}$$

where s^c is a scaling factor; $\mathbf{A}^c$ represents the camera intrinsic matrix including the focal lengths and central point coordinates; and the tuple $[\mathbf{R}^c \ \mathbf{t}^c]$ is a rigid transformation responsible for mapping the world coordinate system onto the camera coordinate system.

Treating the projector as an inverse camera, the point $[x^w, y^w, z^w, 1]^T$ undergoes a similar reprojection onto the projection image plane, yielding coordinates $[u^p, v^p]^T$, governed by the following equation.

$$s^p[u^p, v^p, 1]^T = \mathbf{A^p}[\mathbf{R^p}, \ \mathbf{t^p}][x^w, y^w, z^w, 1]^T \ , \tag{5}$$

where s^p is a scaling factor, $\mathbf{A}^p$ is the projector intrinsic matrix, and $[\mathbf{R}^p,\ \mathbf{t}^p]$ represents a rigid transform from the world coordinate system to the projector coordinate system.

Analogously for the thermal camera, we can model its behavior by the expression:

$$s^t[u^t, v^t, 1]^T = \mathbf{A^t}[\mathbf{R^t},\ \mathbf{t^t}][x^w, y^w, z^w, 1]^T \ , \tag{6}$$

where once again s^t is a scaling factor, $\mathbf{A}^t$ is the thermal camera intrinsic matrix, and $[\mathbf{R}^t,\ \mathbf{t}^t]$ represents a rigid transform from the world coordinate system to the thermal camera coordinate system.

Lens distortions may occur in optical systems and cause measurement inaccuracies, which may affect the accuracy of the multimodal system. Therefore, it is imperative to consider these distortions. Typically, the lens distortion model employed in such systems is represented by the following mathematical expression:

$$\begin{bmatrix}\bar{u}_d \\ \bar{v}_d\end{bmatrix} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6)\begin{bmatrix}\bar{u} \\ \bar{v}\end{bmatrix} + \begin{bmatrix}2p_1\bar{u}\bar{v} + p_2(r^2 + 2\bar{u}^2) \\ 2p_2\bar{u}\bar{v} + p_1(r^2 + 2\bar{v}^2)\end{bmatrix} \ , \tag{7}$$

with

$$r^2 = \bar{u}^2 + \bar{v}^2 \ , \tag{8}$$

and

$$[\bar{u}, \bar{v}, 1]^T = \mathbf{A}^{-1}[u, v, 1]^T \ , \tag{9}$$

where k_1, k_2, k_3 are the coefficients of radial distortion, and p_1, p_2 are the coefficients of tangential distortion, $[\bar{u}_d, \bar{v}_d]^T$ are the normalized coordinates of the distorted points, $[\bar{u}, \bar{v}]^T$ are the free-distortion normalized coordinates, and the superscript $(.)^{-1}$ denotes the inverse matrix operation.

2.2 3D Reconstruction

Reconstruction process employs a triangulation approach involving the intersection of a plane line through Equations (4) and (5).[15] To reconstruct each scene in three dimensions, phase-shifted binary fringe patterns are projected, and the wrapped phase is estimated using gray coding for phase unwrapping. The continuous phase is then converted into projector coordinates.

$$[u^p, v^p] = \frac{P}{2\pi}[\phi_u, \phi_v] \ , \tag{10}$$

Here, $[\phi_u, \phi_v]$ represents the continuous phases acquired in the u- and v-directions, and P stands for the pitch of the projected fringes. Following this, normalization of the camera and projector coordinates took place, with adjustments made for distortions. Ultimately, the determination of 3D coordinates is achieved through triangulation, employing equations (4) and (5).

2.3 Obtaining temperature as texture

The points of the 3D reconstruction in camera coordinates (world coordinates) are transformed to the thermal camera coordinate system, using the matrix of extrinsic parameters between these cameras. Eq. 11 shows this process.

$$[x^t, y^t, z^t, 1]^T = [\mathbf{R_c^t},\ \mathbf{t_c^t}][x^c, y^c, z^c, 1]^T \ , \tag{11}$$

Where $[x^t, y^t, z^t, 1]^T$ are the points of the 3D reconstruction in coordinates of the thermal camera, $[\mathbf{R_c^t},\ \mathbf{t_c^t}]$ is a matrix that transforms the points from the coordinate system of the visible camera to the coordinate system of the thermal camera, and $[x^c, y^c, z^c, 1]^T$ are the points of the 3D reconstruction in coordinates of the visible camera.

These points are projected to the thermal camera sensor using the intrinsic matrix of the thermal camera. Distortions must be taken into account. These pixels have sub-pixel resolution. Then, we interpolate them cubically with the temperature of the nearest pixels to obtain the temperature of each point of the 3D reconstruction.

3. PROBLEMS IN CONVENTIONAL CALIBRATIONS

We implemented the multimodal system shown in the figure 1. It consists of an Epson Powerlite W39 projector with 1280×800 pixels, a IDS UI-3880CP-C-HQ Rev.2 (AB00855) color camera with 2048×1600 pixels, and an Arduino Uno for synchronization. The thermal camera is an Infiray P2 camera with 256×192 pixels, and thermal resolution of 0.04 $^{\circ}C$. The calibration process involves the utilization of 18-step phase-shifted fringe patterns and 6 gray-coded binary patterns for extracting phase maps in horizontal and vertical orientations. During the 3D reconstruction stage, the absolute phase per camera pixel is obtained from the projected vertical fringe patterns to derive metric values.

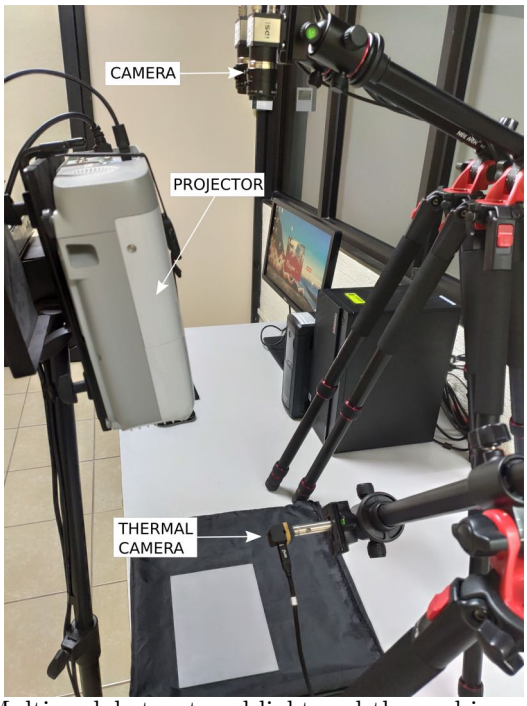


Figure 1: Multimodal structured light and thermal imaging system.

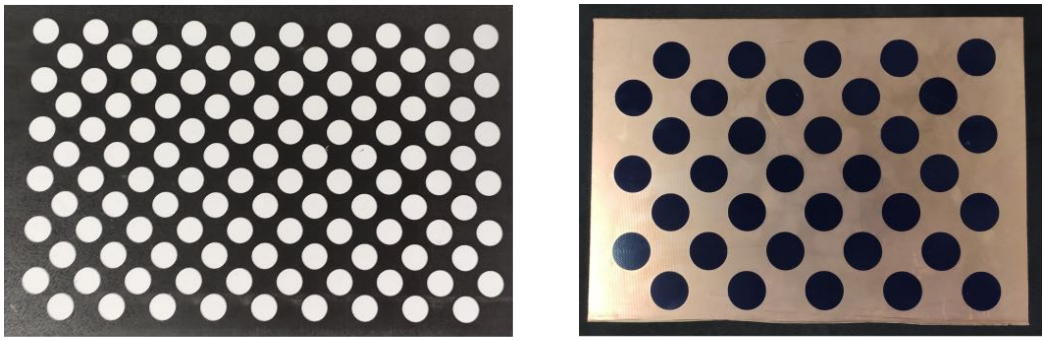
The calibration process was lengthy due to the use of two stages of image acquisition. We used the target shown in Fig. 2*a* in the calibration of the structured light system. We obtained 0.46 as the mean reprojection error in that process. In the calibration of the system composed of the visible camera and the thermal camera, we used the target shown in Fig. 2*b*. In this stage, the mean reprojection error was 0.32.

The design of the thermal target shown in Fig. 2*b* presented a set of problems. Given the resolution of the thermal camera sensor, detectors in software such as MATLAB did not detect the circles because they were represented by too few pixels (5 or 7 pixels). Therefore, it was necessary to decrease the number of circles to 35. A higher number of circles implies the almost total discard of images making calibration impossible.

This last calibration required more than 60 images for each camera. Attempts with 30 images were unsuccessful. This situation was evident at the moment of superimposing the temperature as texture in the 3D reconstructions. Fig. illustrates the problem. After calibration, we thermally stimulated wooden dices with an electric plate (See Fig. 3*a*). We projected the fringes and acquired them along with the respective thermal image of the scene. The 3D reconstruction is shown in Fig. 3*b*. However, the texture did not fit exactly on dice. We can see how it is displaced with respect to the objects (See Fig. 3*c*). This situation can be generated, given the small graphic resolution of the thermal camera and the number of points in the target (35 feature points).

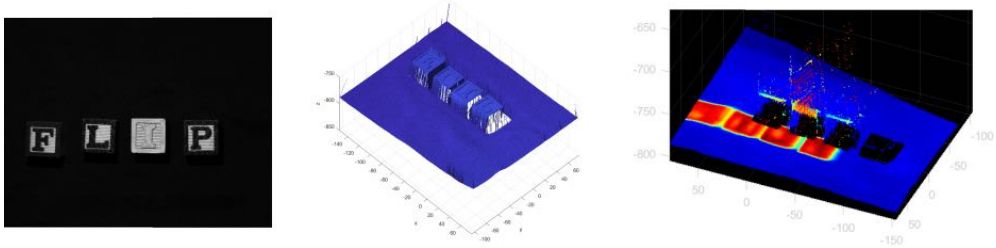
4. NEW DIRECTION IN MULTIMODAL CALIBRATION

Our approach propose a target-free calibration methodology for a multimodal thermal imaging and structured light system. Initially, we calibrated a stereo system composed of two visible cameras (main camera and auxiliary



(a) (b)

Figure 2: Targets used in calibration: (a) Structured light target, and (b) Target of the visible-thermal camera system.



(a) (b) (c)

Figure 3: Problem in calibration with 30 images.

camera) using a target presented on an LCD screen. Then, using an uncalibrated projector, we projected vertical and horizontal fringe patterns onto the back of a mirror. Images of the mirror with the patterns were acquired using the two visible cameras and a thermal camera. Then, we searched for point correspondences between the visible cameras via phase. Thus, we can reconstruct the mirror using stereo triangulation. We then define some feature points on the reconstructed mirror surface to calibrate the main camera and projector system. Then, using the mirror corners, we estimated the homography between the points on the main camera and the thermal camera. In this manner, we interpolate the position of the feature points in the mirror plane on the thermographic camera sensor and calibrate the main camera and thermal camera system. This methodology would circumvent the problematic feature detection in the thermal camera sensor. It would also avoid problems caused by manufacturing defects in targets.

5. CONCLUSION

In this paper we explored a conventional way of calibrating a multimodal structured light and thermal imaging system. We also discussed the problems frequently encountered in this type of approach. And we presented a new way for calibrating a multimodal structured light and thermal imaging system. This direction could circumvent the frequent problems caused by targets in low-cost optical systems. Our approach is more elaborate than traditional ones, but promises to reduce the impact of errors of target fabrication and feature detection in low-resolution thermal camera sensors. Also, the problems of feature detection in challenging situations due to thermal reflections would be avoided.

ACKNOWLEDGMENTS

This work has been funded by Universidad Tecnológica de Bolívar (UTB). Raúl Vargas thanks Universidad Tecnológica de Bolívar (UTB) for a PhD scholarships. Also, Eberto Benjumea thanks MinCiencias and Sistema General de Regalías (Programa de Becas de Excelencia) for a PhD scholarship.

REFERENCES

[1] An, Y. and Zhang, S., "High-resolution, real-time simultaneous 3D surface geometry and temperature measurement," *Optics Express* **24**, 14552–14563 (June 2016). Publisher: Optica Publishing Group.

[2] Juarez-Salazar, R., Rodriguez-Reveles, G. A., Esquivel-Hernandez, S., and Diaz-Ramirez, V. H., "Three-dimensional spatial point computation in fringe projection profilometry," *Optics and Lasers in Engineering* **164** (May 2023).

[3] Meza, J., Contreras-Ortiz, S. H., Perez, L. A. R., and Marrugo, A. G., "Three-dimensional multimodal medical imaging system based on freehand ultrasound and structured light," *Optical Engineering* **60**, 054106 (May 2021). Publisher: SPIE.

[4] Jablonski, R. Y., Osnes, C. A., Khambay, B. S., Nattress, B. R., and Keeling, A. J., "An in-vitro study to assess the feasibility, validity and precision of capturing oncology facial defects with multimodal image fusion," *The Surgeon* **16**, 265–270 (Oct. 2018).

[5] Juarez-Salazar, R., Zheng, J., and Diaz-Ramirez, V. H., "Distorted pinhole camera modeling and calibration," *Applied Optics* **59**, 11310–11318 (Dec. 2020). Publisher: Optica Publishing Group.

[6] McClatchy, D. M., Rizzo, E. J., Meganck, J., Kempner, J., Vicory, J., Wells, W. A., Paulsen, K. D., and Pogue, B. W., "Calibration and analysis of a multimodal micro-CT and structured light imaging system for the evaluation of excised breast tissue," *Physics in Medicine & Biology* **62**, 8983 (Nov. 2017). Publisher: IOP Publishing.

[7] Landmann, M., Heist, S., Dietrich, P., Lutzke, P., Gebhart, I., Templin, J., Kühmstedt, P., Tünnermann, A., and Notni, G., "High-speed 3D thermography," *Optics and Lasers in Engineering* **121**, 448–455 (Oct. 2019).

[8] Barone, S., Paoli, A., and Razionale, A. V., "Assessment of chronic wounds by three-dimensional optical imaging based on integrating geometrical, chromatic, and thermal data," *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine* **225**, 181–193 (Feb. 2011).

[9] ElSheikh, A., Abu-Nabah, B. A., Hamdan, M. O., and Tian, G.-Y., "Infrared Camera Geometric Calibration: A Review and a Precise Thermal Radiation Checkerboard Target," *Sensors* **23**, 3479 (Jan. 2023). Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.

[10] Lagüela, S., González-Jorge, H., Armesto, J., and Arias, P., "Calibration and verification of thermographic cameras for geometric measurements," *Infrared Physics & Technology* **54**, 92–99 (Mar. 2011).

[11] Luhmann, T., Piechel, J., and Roelfs, T., "Geometric Calibration of Thermographic Cameras," in [*Thermal Infrared Remote Sensing: Sensors, Methods, Applications*], Kuenzer, C. and Dech, S., eds., *Remote Sensing and Digital Image Processing*, 27–42, Springer Netherlands, Dordrecht (2013).

[12] Zuo, C., Feng, S., Huang, L., Tao, T., Yin, W., and Chen, Q., "Phase shifting algorithms for fringe projection profilometry: A review," *Optics and Lasers in Engineering* **109**, 23–59 (2018).

[13] Zhang, S. and Huang, P. S., "Novel method for structured light system calibration," *Optical Engineering* **45**(8), 083601–083601 (2006).

[14] Li, B., Karpinsky, N., and Zhang, S., "Novel calibration method for structured-light system with an out-of-focus projector," *Applied Optics* **53**(16), 3415–3426 (2014).

[15] Meza, J., Vargas, R., Romero, L. A., Zhang, S., and Marrugo, A. G., "What is the best triangulation approach for a structured light system?," *Proc. SPIE* **11397**, 113970D (2020).