

Real-time pixel-level fusion of colour and thermal (RGB-T) imagery via extrinsic optical alignment: Scope for automating building inspections

M. H. Shariq and B. R. Hughes

Abstract—Building inspections and energy surveys are performed to detect and diagnose building defects, however, existing inspection methods are highly labour intensive, time consuming, and not suitable for large-scale audits. This work presents a thermal-colour fusion technique using pre-capture optical alignment that eliminates the need for compute-intensive post-processing of data. The coinciding axes of dual cameras also prevent parallax errors, which helps expand its applications to near proximity or indoor thermal inspection as well. A prototype of this design was assembled and tested, which showed a pixel-level accurate alignment of thermal and colour camera imagery with minimal computation. The research and development in this paper provides scope and feasibility to automate building inspection techniques through drone-based low-powered embedded machine learning models and identify faults in real-time for large-scale surveys of building façades.

Keywords— Building Inspection, Real-time Fault Detection, RGB-T Fusion

I. INTRODUCTION

The building sector is responsible for 40% of the overall energy consumption in the EU. Building defects, such as heat losses, moisture, and air leakages, inevitably cause inefficient space heating or cooling, which accounts considerably towards high energy consumption and associated greenhouse gas emissions. Infrared thermal imager (or thermal camera) is a non-destructive sensing device [1] that provides radiometric heat signature of targeted area viewed as a colour gradient that helps in identifying such faults. However, building energy surveying through a terrestrial-based or hand-held thermal imager poses the following limitations: (A) Inspection of a complete building is excessively time-consuming and labour-intensive [2] (B) Inspection of high-rise façades and many indoor inaccessible areas are often restricted by the physical reach of the camera (or the observer). (C) Inspection from thermal imagers alone does not provide a quantitative result. [3] Drones (quadcopters) provide solutions to remote inspections and aerial surveys in a variety of related applications [4, 5], including

building inspections [6]. While it does offer an answer to the limitations (A) and (B), integrating real-time 3D thermal model reconstruction of buildings through drones for a quantitative study (to resolve limitation (C) as well) would not be possible due to the requirement of resource-intensive hardware on-board.

A quantitative building energy audit often requires combining geometric information with thermal radiometric information [7]. This allows building energy simulation, modelling and assessment for a thorough diagnostic survey to identify, quantify and diagnose building defects and energy losses. Over the past decade, numerous studies have produced solutions to generate thermographic 3D models of building interior and exterior using various techniques.

To generate thermographic 3D models, apart from an RGB camera and thermal camera, all existing techniques required additional hardware and their associated computational requirements. Although it provides the necessary data for a quantitative inspection (hence overcoming limitation (C) above), it consequently prevents the effective use of drones for rapid and unrestricted survey of interior and high-rise façades.

To summarise, limitation (C) is dependent on (A) and (B) and vice versa. In other words, to solve all three constraints together, a minimalistic image registration technique is required to superimpose thermal data onto RGB visible imagery. This potential solution could eliminate the need for additional sensors (such as stereo range, IR depth sensors, laser scanners) to allow drone-based inspections while capable of producing 3D thermographic models in real-time.

Thankfully, monocular photogrammetry provides a time-efficient real-time solution to 3D model reconstruction through series of images captured with overlapping features [8, 9]. Thermal data can then be overlaid through various image stitching techniques [3, 10]. Ham et al. [11] developed a framework to generating 3D orthothermogram (thermographic 3D models) with superimposed RGB imagery without the use of any laser scanner or depth sensors.

Nevertheless, all image fusion techniques including feature matching, homography transformation, image mosaicking are applicable and/or accurate in those cases where the target is either planar or captured from great distance (e.g. aerial view,

Manuscript received 15 October, 2021; revised 31 January, 2022; accepted 10 February, 2022; published 16 February, 2022.

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 licence (CC BY <http://creativecommons.org/licenses/by/4.0/>). Unrestricted use (including commercial), distribution and reproduction is permitted provided that credit is given to the original author(s) of the work, including a URI or hyperlink, this public license and a

copyright notice. This article has been subject to single-blind peer review by a minimum of two reviewers. Digital Object Identifier <https://doi.org/10.55066/proc-icec.2021.30>

M. H. Shariq is a 3rd year PhD student and Research Assistant at the department of Mechanical Engineering, University of Hull, Hull HU6 7RX (email: H.Shariq@hull.ac.uk)

B. R. Hughes is the Head of Department of Engineering at University of Hull, Hull HU6 7RX (email: B.R.Hughes@hull.ac.uk).

normal view of façades, etc.). Photogrammetry for buildings (especially indoors) contain closer ranged varying depths of geometry, i.e. a non-planar parallax, which cannot be accurately aligned through stereo-based homography transformations or feature detection.

Handheld FLIR thermal cameras [12] comprises a ‘Multi Spectral Dynamic Imaging’ (MSX) technology to merge its visible camera result with its corresponding thermal image [13]. A few examples of an edge representation of the visible data overlaid on the radiometric thermal image are shown in **Figure 1** [14]. As highlighted using white circles, the edge images are seen inconsistent and do not appear to be accurately aligned with their respective thermal version. This is due to the same limitation (fusion of non-planar and/or close ranged target with no additional hardware) as discussed above.

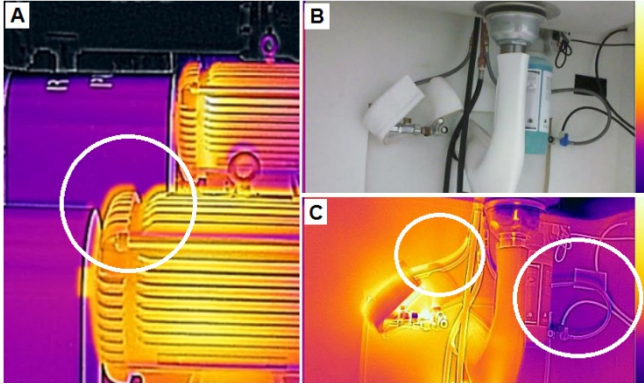


Figure 1. (A) Offset of thermal imagery against visible imagery, and (C) double pipes seen due to misplaced thermal fusion, which can be compared with (B) its visible / RGB image

Hence, superimposing radiometric thermal data with pixel-level accuracy on the photogrammetric model, without the use of any additional equipment, is yet to be unravelled and can certainly provide a possibility for a drone-based solution to real-time thermographic 3D reconstruction of buildings, which can be used for a detailed quantitative inspection of buildings.

II. BACKGROUND

A pinhole camera model [15] is examined for the RGB camera as well as the thermal camera in this paper. The general camera projection matrix $\mathbf{P}$ is given by:

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} \mathbf{R} \\ \mathbf{t} \end{bmatrix}, \text{ with } \mathbf{K} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \text{ where,} \quad (1)$$

$\mathbf{R}$ is the 3x3 rotation matrix representing camera orientation (extrinsic)

$\mathbf{t}$ is the transformation vector for representing camera position (extrinsic)

$\mathbf{K}$ is the intrinsic parameters matrix that includes:

f_x, f_y , the focal length,

c_x, c_y , the principal point (optical centre),

s , the skew coefficient

An example of a stereo camera configuration in the cartesian coordinate system is shown in **Figure 2** [16]. The left camera (C_1) and the right camera (C_2) both target the point P , which is projected as p_1 and p_2 on their respective image planes as shown. A near-distant and/or non-planar target scene affects

the accuracy of homography transformation and fusion results, which this paper aims to resolve.

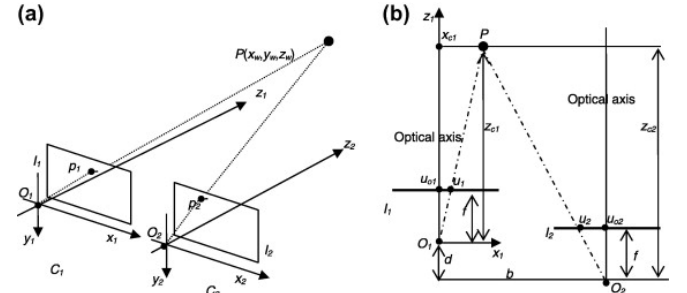


Figure 2. (a) Stereo camera arrangement, (b) optical axes of both cameras from top view [16]

III. METHODS

This paper focusses on eliminating the parallax errors by physically aligning the projections of visible (RGB) and thermal camera. Following sub-sections describe the fundamental process.

The Hybrid RGB-T Module:

The idea is to geometrically align the optical axes of the two cameras with the help of a 45° angled one-way mirror-glass. This is represented in **Figure 3** where the thermal camera reflects off this mirror to see the target (scene) while the RGB camera sees through the same mirror to capture the very target through the same coinciding optical axis. In this paper, this apparatus is referred to as the ‘hybrid RGB-T module’.

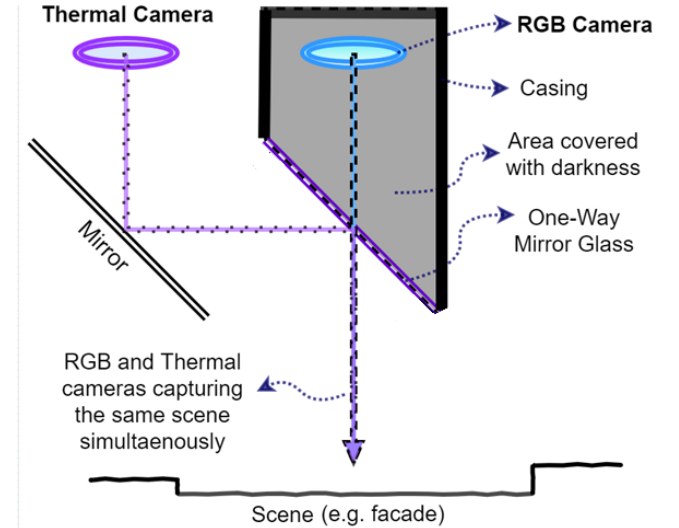


Figure 3. Proposed hybrid RGB-T module

Computation After Extrinsic Alignment:

Since the cameras capture nearly perpendicular from the scene, the skew and lens distortions can be corrected easily through MATLAB. Hence, the skew coefficient, s , becomes zero. Also, as the optical axes of both cameras coincides during capture, the extrinsic parameters of both cameras can be expressed as:

$$\begin{bmatrix} \mathbf{R} \\ \mathbf{t} \end{bmatrix}_{RGB} = \begin{bmatrix} \mathbf{R} \\ \mathbf{t} \end{bmatrix}_{Thermal} \quad (2)$$

Therefore, the RGB-T camera projection matrix $\mathbf{P}$ can be defined as:

$$\mathbf{P} = \begin{bmatrix} \mathbf{R} \\ \mathbf{t} \end{bmatrix} (\mathbf{K}_R + \mathbf{T}_{Th}^{-1} \mathbf{K}_{Th}) \quad (3)$$

where $\mathbf{K}_R$ and $\mathbf{K}_{Th}$ are the intrinsic matrix of RGB camera and thermal camera respectively, and $\mathbf{T}^{-1}$ is the inverse affine transformation of thermal camera from general pose to origin.

The Proposed Process:

The proposed process of capturing and superimposing images from both cameras with a pixel-level accuracy using the hybrid RGB-T module is shown in **Figure 4**.

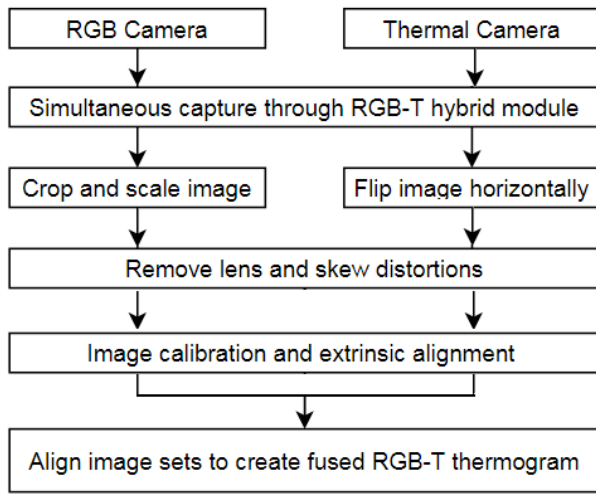


Figure 4. Process of proposed method to capture and superimpose RGB-T data

IV. DESIGN AND EXPERIMENTATION

For the purpose of proof-of-concept, a prototype of the hybrid RGB-T module is designed as a CAD model and then 3D printed. A thermal camera and a smartphone camera are used to superimpose both imagery using a one-way mirror installed onto the module. The design, fabrication, list of equipment and experimental setup are detailed in the following sub-sections.

Design Specification and Fabrication:

A simplified version of the proposed RGB-T hybrid module is designed to hold a single one-way mirror and the camera, while is engineered to be mounted onto the lens of a thermal camera. Apart from the hybrid module, the following equipment was used for the experiment:

- 1) FLIR T640 Thermal Camera
- 2) Visible Camera (LG G Pro 2 smartphone was used for this experiment)
- 3) One-way Mirror
- 4) Dark coated plate
- 5) Tripod stand (for the thermal camera)
- 6) Calibration chart (A3)

The one-way mirror was created using a low carbon clear glass and placing a one-way silver reflective film on one side. The technical specifications and comparison of the thermal camera [17] and the smartphone camera are detailed in **Table 1**.

Table 1. Technical specifications of thermal camera and RGB smartphone camera

Sensor/Spec	Thermal Camera	RGB Camera
Make	FLIR	LG
Model	T640	G Pro 2
FOV	45°	45°
Resolution	640 x 480	1920 x 1080
Framerate	25 FPS	60 FPS
Range	-40°C - 150°C	-
NEDT	< 30mK	-
Weight	1.3kg	0.172kg

The arrangement of the RGB-T module equipment (1-4) is shown in **Figure 5**.

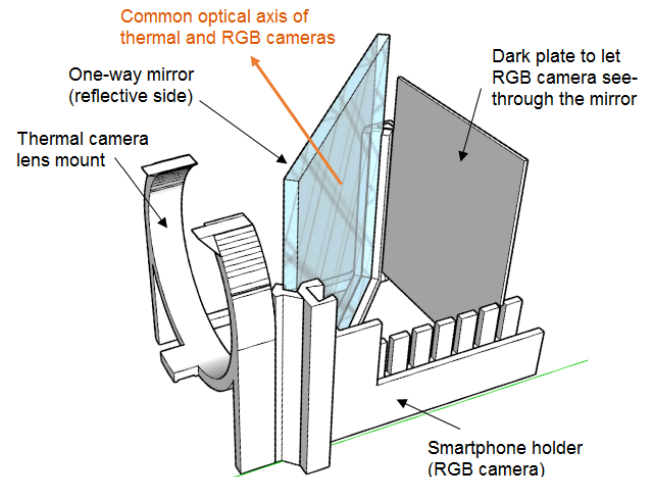


Figure 5. CAD model showing all components of the proposed RGB-T hybrid module

A 3D printed RGB-T hybrid module, which was printed using PLA plastic (0.4mm nozzle diameter) on a 'Raise 3D Pro2 Plus' [18] 3D Printer, is shown in **Figure 6**.

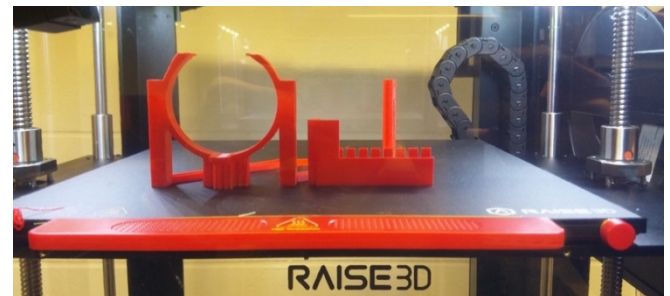


Figure 6. 3D printed model of the proposed hybrid module

Experimental Setup and Calibration:

Figure 7 shows the arrangement of the experimental setup. The thermal camera is mounted on a tripod. The hybrid RGB-T (3D printed) module is mounted on the thermal camera lens that holds the one-way mirror and an RGB camera (a smartphone for this experiment).

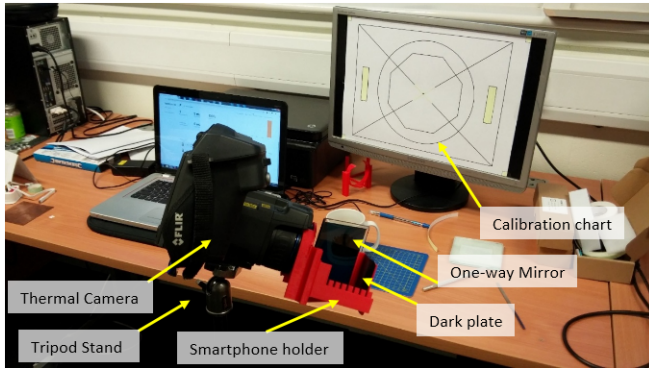


Figure 7. Experimental setup

A calibration chart of A3 size was used (as seen in Figure 7) that contains marks made using a florescent glow-in-the-dark tape (outlined using visible black rectangle), which can be seen through both the thermal camera and RGB camera. This helps in the calibration process to accurately align the two corresponding (concurrently captured) images.

The captured images are required to be scaled, cropped and re-positioned, while any lens or skew distortion needs to be removed on both camera images before aligning together.

However, although this technique required less equipment and was easier to set-up, the image filtration process was manual and time consuming. Other techniques of calibration (such as IR LED bulbs on a grid, or fluorescent light on a checkboard) shall be used for a better accuracy in calibration with less manual post-processing.

V. RESULTS AND DISCUSSION

Two sets of preliminary results captured using the RGB-T hybrid module are shown in **Figure 8** (capturing a mug with hot water) and **Figure 9** (capturing a 3D printer with heat-bed and nozzle temperature maintained at 75°C and 220°C respectively).

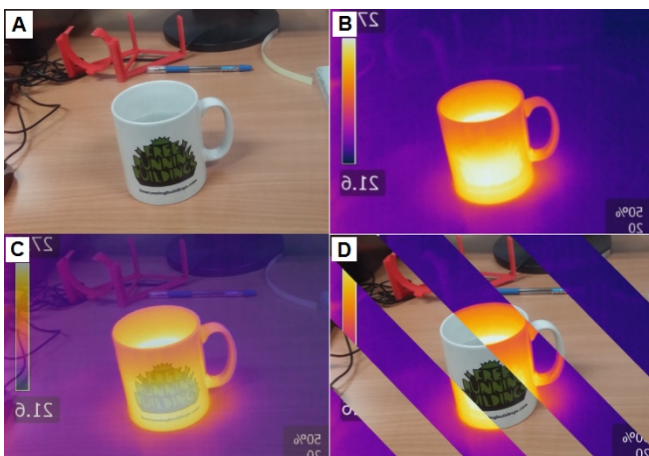


Figure 8. RGB-T results capturing a mug with hot water showing (A) raw RGB image, (B) raw thermal image, (C) fused RGB-T result and (D) overlaid images with stripped cut-out

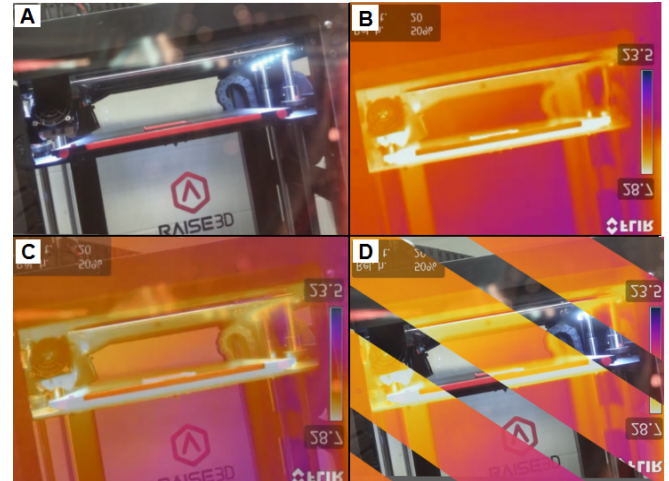


Figure 9. RGB-T results capturing a heated 3D printer showing (A) raw RGB image, (B) raw thermal image, (C) fused RGB-T result and (D) overlaid images with stripped cut-out

Both figures contain raw pair of images captured from both cameras (i.e. A and B) as well as fused images shown as partial cut-outs of thermal and RGB strips (i.e. D), and a semi-transparent RGB data overlaid on its corresponding thermal image (i.e. C).

Thanks to the coinciding optical axes, even the raw images align with a pixel-level accuracy as seen from the results. In this experiment, the Field of View (FOV) of both cameras were identical, however, in the case where they are not, lens distortion needs to be corrected before alignment. It can also be seen the thermal image is flipped horizontally. This is due to the mirror, which can easily be resolved using a dual mirror system as initially proposed in Figure 3.

VI. CONCLUSION AND FUTURE SCOPE

This paper addresses the limitations of stereo-based parallax errors and inaccuracies while fusing images that are non-planar and/or near-distant that often prevents such data to be used for machine learning. An RGB-T hybrid module is proposed, which was designed and fabricated, that aligns the optical axes of two cameras (in this case: a thermal and smartphone camera) through a one-way mirror. Results show the pair of corresponding raw images align with a pixel-level accuracy despite capturing a near-distant and non-planar target. This delivers a proof-of-concept of a radiometric fused output using a real-time alignment capability with no need for image stitching or feature detection to obtain an accurately fused thermal-visible (RGB-T) result.

This work does not only prove to help in real-time qualitative building inspection for fault detection, but also provides a massive scope in large-scale real-time drone-based automatic fault-detection of building façade. This is illustrated as a flow chart in **Figure 10** where the dotted line represents a non-real-time flow to prepare transfer learning prior to actual inspection and the solid lines are real-time flow of data transfer or actions that are performed by the embedded system on the drone.

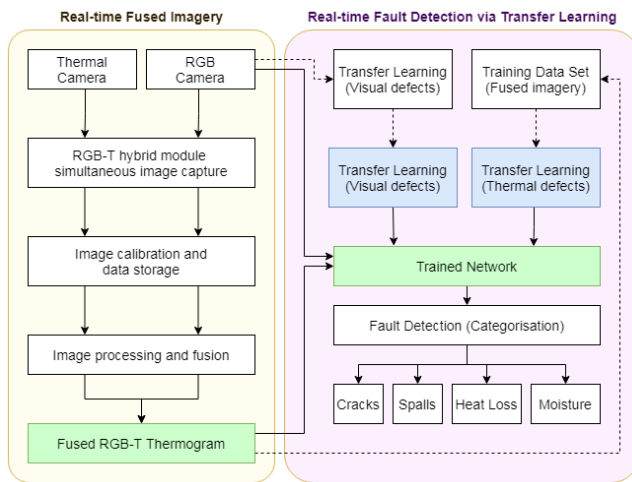


Figure 10. Flowchart showing scope of this work and future prospect using this framework

REFERENCES

- [1] Cai, W., Wen, X., Wang, S., Wang, L.: A real-time detection method of building energy efficiency based on image processing. *J. Vis. Commun. Image Represent.* (2019). <https://doi.org/10.1016/j.jvcir.2019.02.032>
- [2] Son, H., Lee, S., Kim, C.: Automated 3D Model Reconstruction to Support Energy-efficiency. In: *Procedia Engineering* (2016)
- [3] Lagüela, S., Armesto, J., Arias, P., Herráez, J.: Automation of thermographic 3D modelling through image fusion and image matching techniques. *Autom. Constr.* (2012). <https://doi.org/10.1016/j.autcon.2012.05.011>
- [4] Seo, J., Duque, L., Wacker, J.: Drone-enabled bridge inspection methodology and application. *Autom. Constr.* (2018). <https://doi.org/10.1016/j.autcon.2018.06.006>
- [5] Entrop, A.G., Vasenev, A.: Infrared drones in the construction industry: Designing a protocol for building thermography procedures. In: *Energy Procedia* (2017)
- [6] Rakha, T., Gorodetsky, A.: Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones, (2018)
- [7] González-Jorge, H., Lagüela, S., Krelling, P., Armesto, J., Martínez-Sánchez, J.: Single image rectification of thermal images for geometric studies in façade inspections. *Infrared Phys. Technol.* (2012). <https://doi.org/10.1016/j.infrared.2012.05.003>
- [8] Pizzoli, M., Forster, C., Scaramuzza, D.: REMODE: Probabilistic, monocular dense reconstruction in real time. In: *Proceedings - IEEE International Conference on Robotics and Automation* (2014)
- [9] Schöps, T., Sattler, T., Häne, C., Pollefeys, M.: Large-scale outdoor 3D reconstruction on a mobile device. *Comput. Vis. Image Underst.* (2017). <https://doi.org/10.1016/j.cviu.2016.09.007>
- [10] Luhmann, T., Piechel, J., Roelfs, T.: Geometric calibration of thermographic cameras. In: *Remote Sensing and Digital Image Processing* (2013)
- [11] Ham, Y., Golparvar-Fard, M.: An automated vision-based method for rapid 3D energy performance modeling of existing buildings using thermal and digital imagery. *Adv. Eng. Informatics.* (2013). <https://doi.org/10.1016/j.aei.2013.03.005>
- [12] FLIR Systems Inc, <https://www.flir.co.uk/browse/industrial/handheld-thermal-cameras/>
- [13] Jaspers, M.E.H., Carrière, M.E., Meij-de Vries, A., Klaessens, J.H.G.M., van Zuijlen, P.P.M.: The FLIR ONE thermal imager for the assessment of burn wounds: Reliability and validity study. *Burns.* (2017). <https://doi.org/10.1016/j.burns.2017.04.006>
- [14] Multi-Spectral Dynamic Imaging (FLIR MSX), <https://www.flir.co.uk/discover/professional-tools/what-is-msx/>
- [15] Hartley, R., Zisserman, A., Richard, H., Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*, 2nd Edition. *Clim. Chang.* 2013 - *Phys. Sci. Basis.* (2003). <https://doi.org/10.1017/CBO9781107415324.004>
- [16] Chen, W., Cao, L., Wu, Z., Huang, S.: A stereo range computation method using thermal infrared and visible cameras. *Infrared Phys. Technol.* 62, 115–123 (2014). <https://doi.org/10.1016/j.infrared.2013.11.003>
- [17] FLIR T640 Thermal Camera for Predictive Maintenance | FLIR Systems, <https://www.flir.co.uk/products/t640/>
- [18] Top Quality Large Format 3D Printer | \$5999 | Raise3D, <https://www.raise3d.com/pro2-plus/191-199, 2010>.