



**University of
Nottingham**

UK | CHINA | MALAYSIA

InfraRed Thermography and 3D-Data Fusion for Architectural Heritage

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy, May 15, 2026.

Neil Alexander Sutherland

20315184

Supervised by

**Prof. Stuart Marsh
Prof. Jon Mills
Prof. Gary Priestnall
Paul Bryan
Fabio Remondino**

Signature _____

Date ____ / ____ / ____

Abstract

Comprehensive documentation is the foundation of effective conservation, repair and maintenance (CRM) practices within architectural heritage. In order to diagnose historic buildings and inform decision making, a combination of multidisciplinary surveys is fundamental to understanding a building's significance and performance. InfraRed Thermography (IRT), a non-contact, non-invasive and non-destructive imaging technique, represents an established tool for heritage conservators, allowing both qualitative and quantitative assessments of temperature to be undertaken. However, the inherent low spatial resolution of thermal imaging has led recent work to fuse thermographic and geometric data in order to generate accurate 3D representations of architectural heritage encapsulating temperature information. This thesis presents InfraRed Thermography 3D-Data Fusion (IRT-3DDF), an emerging field of research transforming traditional *two-dimensional* thermal infrared images into novel *three-dimensional* models capable of visualising thermal anomalies. To explore the use of 3D thermal models for architectural heritage, IRT-3DDF is conceptualised as a discipline, expressing the key considerations required for effective thermal data capture and fundamental procedures required for the use of thermal infrared (TIR) cameras both geometrically and radiometrically.

The originality of the presented work is encapsulated by the direct matching of infrared and visible spectrum images for data fusion processes, a previously under-researched area of IRT-3DDF due to the difficulty determining correspondences

across image modalities. This thesis introduces deep learning-based image matching, pre-trained neural networks designed for multi-view feature detection and feature matching, as a means of co-orienting multi-modal images depicting architectural heritage. Importantly, these correspondences are determined using ‘out-of-the-box’ neural networks pre-trained on vast quantities of curated RGB images, utilised outside of their expected training domain when applied to TIR images and IRT-3DDF. The outcomes of this research demonstrate that deep learning-based image matching, when used in conjunction with combined bundle block adjustments, provide a viable approach for IRT-3DDF, capable of co-registering blocks of multi-modal images across varying scales, settings, sensors and subjects. Notably, the presented methods are fully-automatic, obviating the need for: geometric pre-calibration; sensors in fixed relative orientation; manual co-registration or alignment; associated position or orientation information; or simultaneous data capture. This thesis presents an efficient, cost-effective and flexible method of IRT-3DDF that can be readily-applied for multi-modal, multi-temporal and multi-sensor fusion.

Acknowledgments

To my lead supervisor, Prof. Stuart Marsh, I thank you for giving me the opportunity to join the Geospatial Systems CDT and contribute to the remarkable research community you have fostered at the Nottingham Geospatial Institute. I remain indebted to you for your faith in taking a chance on a lowly history graduate and providing an atmosphere where I, and many others, could thrive. Your kindness, humour and wisdom has been central to all our successes.

To Prof. Jon Mills, I thank you for your endless counsel and support throughout my PhD. Notably, your investment in my research has helped me through some of the most challenging times of my PhD, for this I am truly grateful. As Director of the Geospatial Systems CDT, I thank you for building a research centre that has extended far beyond a collection of individual projects, but into a supportive and compassionate community.

To my mentor, Paul Bryan, I owe a considerable debt of gratitude. You inspired me with your words on the future of heritage documentation and H-BIM at the University of Bath, long before I knew I could pursue them! You have provided considerable backing on my journey and I thank you for your time and friendship. I hope to approach my future vocation with the same enthusiasm and kindness that you have shown me!

To Fabio Remondino and all at FBK, I thank you for your invaluable experience and

warmth hosting me in Trento. My secondments marks the most exciting chapter of my PhD, one that is only bolstered by the remarkable community you all sustain. I would also like to pay particular thanks to Luca Morelli, for his considerable knowledge, patience and rapport during our collaborations!

To Prof. Gary Priestnall, thank you for providing your considerable experience and broad perspective to this project. I have greatly appreciated the opportunities to hear your thoughts, seek your advice and share in the progress of our work!

To those at the Geospatial Systems CDT, I cannot express the joy and pride I feel having been a part of this diverse and considerate community. To Jamie Stogden, you have been the rock upon which my (and many other's) PhD's have been built, I would not be submitting this work without your unwavering support. I would also like to thank Maria Peppas, James Goodyear, Frank Atkinson and Phyllida Bailey for their assistance throughout my time at Newcastle University.

To the Nottingham Geospatial Institute and University of Nottingham more broadly, I thank you for your camaraderie throughout my time at UoN. Particularly, I would like to thank Sean Ince for his endless support, teaching and assistance, without which this work would not be possible! Also, thank you Esther Little for your support and dedication to all at the NGI, it has been central to our successes. I would also like to thank: Norma Oldfield, Chenyu Xue, Steven Lawton at The Engineering Applications Workshop, Jason Young at The Rapid Prototyping Lab, Richard Storrow and Robert Hyde at Campus Service.

To Jon Bedford and the Geospatial Survey Team at Historic England, I thank you all for your guidance and assistance throughout my PhD, much of which has been done with your Sony A7RII. This small gesture has played a huge role, a kindness I hope I can one day repay!

To my family, thank you for your unwavering love, support and assurance on this

four-year adventure! You have each picked me up when I thought I had nothing left, reminding me of our enduring mantra: with determination; success is inevitable!

Thesis Publications

The following publications were produced in the course of this PhD thesis:

- Sutherland, N., Marsh, S., Priestnall, G., Bryan, P., Mill, J. (2023). InfraRed Thermography and 3D-Data Fusion for Architectural Heritage: A Scoping Review. *Remote Sensing*, 15, 2422. DOI: 10.3390/rs15092422.
- Sutherland, N., Marsh, S., Remondino, F., Perda, G., Bryan, P., Mills, J. (2025). Geometric Calibration of Thermal Infrared Cameras: A Comparative Analysis for Photogrammetric Data Fusion. *Metrology*, 5, 43. DOI: 10.3390/metrology5030043.
- Morelli, L., Sutherland, N., Ioli, F., Vitti, A., Marsh, S., Mills, J., Bryan, P., Remondino, F. (2025). Deep Learning Multi-Modal Image Matching for the Automatic Co-Registration of Thermal InfraRed and Visible Images. *ISPRS Journal of Photogrammetry and Remote Sensing* (Under Review).
- Sutherland, N., Morelli, L., Mills, J., Bryan, P., Marsh, S., Remondino, F. (2025). Multi-Modal and Multi-Sensor Photogrammetric Data Fusion Exploiting a New Repository for Infrared Thermography Datasets. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-1/W6-2025, 235-242. DOI: 10.5194/isprs-archives-XLVIII-1-W6-2025-235-2025.

Contents

Abstract	i
Acknowledgments	iii
Thesis Publications	vi
List of Tables	xii
List of Figures	xxi
Chapter 1 Introduction	1
1.1 Thesis Aims & Objectives	5
1.2 Thesis Structure	6
Chapter 2 Conceptualisation of IRT-3DDF	9
2.1 Introduction	11
2.2 IRT-3DDF: A Definition	12
2.3 IRT-3DDF: State-of-the-Art	16
Chapter 3 Foundation of IRT-3DDF	37
3.1 Introduction	39
3.2 InfraRed Thermography: A History	42
3.3 InfraRed Thermography: Principles & Practices	45
3.4 Thermal InfraRed (TIR) Cameras	55
3.5 Thermal Data Capture	65
3.6 Conclusion	79
Chapter 4 Calibration for IRT-3DDF	80
4.1 Introduction	82
4.2 TIR Geometric Calibration: State-of-the-Art	84

4.3	Methodology	87
4.4	Results	105
4.5	Discussion	118
4.6	Conclusion	125
Chapter 5	Generation of IRT-3DDF	127
5.1	Introduction	129
5.2	Multi-Modal Image Matching: State-of-the-Art	131
5.3	Methodology	135
5.4	Results	151
5.5	Discussion	165
5.6	Conclusion	168
Chapter 6	Discussion & Future Work	170
6.1	Introduction	172
6.2	IRT-3DDF: A Justification	172
6.3	An Appraisal of 3D-Data Fusion	174
6.4	Radiometric Calibration, Correction & Validation	179
6.5	Thermal InfraRed Cameras: Resolutions, Specifications & Future Developments	187
6.6	Multi-Modal Image Matching for IRT-3DDF	189
6.7	IRT-3DDF for Architectural Heritage	198
Chapter 7	Conclusion	200
7.1	Thesis Aims & Objectives: Revisited	201
7.2	Future Work	203
	Bibliography	204
	Appendices	246
Appendix A	Buildings	247
A.1	Lenton Lodge Gatehouse, Nottinghamshire, UK	247
A.2	Castello Di Arco, Trentino, Italy	251

A.3	Dunrobin House, Warwickshire, UK	254
Appendix B	IRT-3DDF Cheatsheet	258

List of Tables

3.1	Infrared spectral regions, their wavelength properties, and common applications. * denotes bands included as part of the <i>thermal</i> or <i>thermal infrared</i> designation (Vollmer & Möllmann, 2018, p.10). . . .	47
3.2	Workswell® WIRIS Pro Specifications (Workswell s.r.o, 2020). . . .	59
3.3	Calculation of material emissivity (ε), standard deviation (σ) and difference in emissivity ($\Delta\varepsilon$) between the predominant material (Mansfield Sandstone) and secondary materials at Lenton Lodge Gatehouse. 70	
3.4	Metrics of the geodetic survey at Lenton Lodge Gatehouse using the Leica TS10 Total Station. Here, the number of measured control points, the mean standard deviation ($\mu\sigma$) for repeated marker measurements and the mean standard error (MSE) across all markers is provided.	77
4.1	Camera specifications for the Workswell WIRIS Pro (WWP & VIS), Sony α 7RII (RGB1) and Nikon D750 (RGB2) cameras.	89
4.2	Calculation of material emissivity (ε), standard deviation (σ) and difference in emissivity of materials ($\varepsilon\Delta$) used for both 2D board and 3D field calibration targets.	90
4.3	IO parameters for WWP obtained with the 2D board using MATLAB, 3D Field using Agisoft Metashape, and 3D Field using Australis. 97	
4.4	RO parameters for WWP, VIS and RGB2 using the 2D board and 3D field calibration targets.	103

4.5	Results of the combined TIR-RGB Bundle Block Adjustment for Lenton Lodge Gatehouse, separating results based on software and distortion coefficients utilised.	109
4.6	RO results for Castello Di Arco’s medieval frescoes, with the measured scale bar RMSE (mm) for each baseline (VIS-WWP or RGB2-WWP) separated into each image modality.	116
5.1	Data Acquisition statistics for both TIR and RGB data sets.	140
5.2	Combined TIR-RGB bundle block adjustment for Dunrobin House.	156
5.3	Combined TIR-RGB bundle block adjustment for Lenton Lodge Gatehouse.	156
5.4	Comparison of Original and Optimised TIR-RGB bundle block adjustment for Dunrobin House.	158
5.5	Comparison of Original and Optimised TIR-RGB bundle block adjustment for Lenton Lodge Gatehouse.	158
5.6	Ground truth infrared thermometer, corrected 2D TIR pixel (T_{2D}), and resulting 3D model (T_{3D}) temperature values for Dunrobin House and Lenton Lodge Gatehouse.	161
5.7	Duration of multi-modal image matching using exhaustive (‘brute-force’) matching with DIM and RIFT2 for Dunrobin House (4,465 image pairs).	166
6.1	Comparisons between IRT and IRT-3DDF.	173
6.2	Summary of IRT-3DDF methods strengths and shortcomings, separated by their fusion ‘space’ and datasets.	178
6.3	Thermitage’s datasets, sensors, and surveys.	192
6.4	Comparison between TIR-only image orientation checked against ‘natural’ or thermal-specific markers on Dunrobin House. Here, the standard deviation of Leica TS10 total station measurements (TST), MRE and RMSE of the oriented image block are given.	194

List of Figures

2.1	Mapping the inter-connectivity of InfraRed Thermography (IRT), 3D-Data fusion (3DDF) and Architectural Heritage research (Sutherland et al., 2023).	13
2.2	IRT-3DDF keyword co-occurrences representing 88 connected keywords within IRT-3DDF architectural heritage research (Sutherland et al., 2023). Relationships between keywords are determined by the proximity of bubbles and the link strength of joined keywords (van Eck & Waltman, 2010).	14
2.3	IRT-3DDF sensors and systems, denoting the individual sensors utilised within the research, their designation for the conceptualisation of IRT-3DDF and the configuration of technologies that can form a TTIS.	15
2.4	A TIR-only model produced by Westfeld et al. (2015), generated using 15,000 TIR images and a unmanned aerial vehicle (UAV). . .	17
2.5	A visualisation of SIFT (Lowe, 2004) (f-j) mono- and (a-e) multi-modal feature matching, demonstrating the difficulty extracting (a,b,c,d,f,g,h,i) and matching (e,j) reliable correspondence across image modalities and varying perspectives (Ham & Golparvar-Fard, 2013a). Here, (b) 236 TIR and (d) 14,946 RGB features are extracted, yet result in only 8 (incorrect) multi-modal matches.	19
2.6	Demonstration of verified deep learning features extracted with SuperGlue, matching TIR images and TLS scans as 2D images (2.5D)(Elias et al., 2023).	22

2.7	TLS scan with temperature values assigned per point (Lecomte et al., 2022).	24
2.8	A comparison between textured thermal orthomosaics generated with a high-res. FLIR SC1030 (1024×768 pix)(left) and a medium-res. FLIR SC660 (640×480 pix)(right) (Adamopoulos et al., 2020b).	25
2.9	The integration of TLS, RGB and TIR within a historic building information model (H-BIM), enabling enhanced visualisation, data management and decision making (Solla et al., 2020).	28
2.10	A 3D TIRpoint cloud, 3D RGB mesh model and resulting 3D textured thermal model using ICP for 3D-3D fusion (Maset et al., 2017).	30
2.11	A collection of IRT-3DDF TTISs comprising different image-, range- and locational-based sensors: (a) (Adán et al., 2023a), (b) (Zhu et al., 2021), (c, d) (Borrmann et al., 2014), (e) (Adán et al., 2020), (f) (Hoegner et al., 2018), (g) (Merchán et al., 2018), (h) (Wang et al., 2013).	33
2.12	An example 3D point cloud generated by a TTIS, featuring a Riegl VZ400 TLS scanner, an Nikon D90 RGB camera and a FLIR AX5 TIR camera (Adán et al., 2020).	35
3.1	William Herschel’s (1800) experiment discovering the presence of infrared radiation, utilising a prism and thermometers to determine the temperature of different electromagnetic regions (Herschel, 1800; Vollmer & Möllmann, 2018).	43
3.2	Agema Infrared Systems’ AGA Thermovision 665 (1965-1967), one of the first commercial, mass production, single-element sensor, optomechanical scanner thermal cameras (Rogalski, 2012; Vollmer & Möllmann, 2018).	44

3.3	Electromagnetic spectrum visualising regions and bands for infrared radiation, and those denoted by the term 'thermal IR' (TIR) (Vollmer & Möllmann, 2018, p.10).	46
3.4	Atmospheric transmittance of MWIR and LWIR at one nautical mile, 15.5 °C and 70% relative humidity (Usamentiaga et al., 2014). . . .	48
3.5	Plank's distribution curves implementing Wien's law of displacement, determining the maximum spectral radiance achieved by a blackbody at a given temperature (Usamentiaga et al., 2014).	51
3.6	Radiation emitted in a scene and received by an infrared sensor. . .	53
3.7	Distinctions between the different temperature values observed within a standard IRT-3DDF pipeline.	55
3.8	Uncooled VOx microbolometer FPA sensor, including microscopic image of: a microbolometer array (left) and microbolometer ROIC contact (right) (Vollmer & Möllmann, 2018).	56
3.9	(a) Workswell WIRIS Pro featuring TIR and VIS sensors; (b) Workswell WIRIS Pro configured with external battery and monitor; and (c) a sample TIR image showing the thermal signature generated by a human subject, thermal-specific survey marker and water bottle in an ambient office building (Workswell s.r.o, 2021).	58
3.10	WWP Sensor Acclimatisation comprising internal FPA sensor ('IR Core') and camera housing ('CPU') temperatures over the 60 min duration.	60
3.11	An example 50 image TIR sequence generated by the WWP beginning immediately after initialisation with automatic NUCs enabled. Here, several NUCs (represented by vertical blue lines) are undertaken at irregular intervals to compensate for the effects of sensor acclimatisation, visible in the 'brightening' of images leading up to a NUC event.	63

3.12	WWP FPA sensor acclimatisation / NUC stabilisation results using the psuedo-blackbody radiometric calibration set-up. Here, intra-image metrics (Max., Mean and Min.) are shown over the course of the 90-minute experiment generated during periodic image capture.	64
3.13	SunCalc online sun-phase tool, showing environmental conditions (sun's position, direction, azimuth and shadows) at Lenton Lodge Gatehouse on 11th February 2024 (Hoffman, 2024).	67
3.14	(a) Protmex PT6252A anemometer, (b) RS PRO RS-91 hygrometer, (c) RS PRO RS-8876 thermometer and (d) TESTO Emission Tape.	68
3.15	Sample (a,c) visible and (b,d) thermal images showing the adhered TESTO emission tape ($\varepsilon = 0.95$) for emissivity calculation of (a,b) Mansfield Sandstone and thermal targets, and for (c,d) wood and iron secondary materials.	69
3.16	A comparison of TIR images at (a) 5m, (b) 10m and (c) 15m from Lenton Lodge Gatehouse, including their GSDs (mm/pix) for the Workswell WIRIS Pro.	71
3.17	Examples of thermal-specific markers constructed for IRT-3DDF, including: (1) aluminium foil- (Dlesk et al., 2021), (2) hand-warmer- (Paziewska & Rzonca, 2022), (3) thermal/natural- (Patrucco et al., 2022b), and (4) reflective aluminium (Grechi et al., 2021) checker-board markers.	75
3.18	Design and production of the 3D-printed thermal-specific survey markers.	76
3.19	(a, b) Geodetic survey of Lenton Lodge Gatehouse, including (c) the Leica TS10 total station.	77
3.20	Example WWP images showing the designed thermal-specific targets at various GSDs (and distances): (a) 9.8 mm (7.5 m), (b) 12.4 mm (9.5 m) and (c) 15.0 mm (11.5 m).	77

4.1	Cameras used for TIR geometric calibration and IRT-3DDF methods, comprising: (a) Workswell WIRIS Pro (WWP & VIS), (b) Sony α 7RII (RGB1), and (c) Nikon D750 (RGB2).	88
4.2	Emissivity determination for the 3D test field markers, utilising TESTO emission tape ($\varepsilon= 0.95$). Here, the emissivity tape is visualised in both (a) VIS and (b) TIR image domains.	90
4.3	2D checker-board imaged in both (a) VIS and (b) TIR domains. . .	91
4.4	3D fields imaged in both VIS and TIR image domains for the (a,b) IO and (c,d) RO calibration procedures.	92
4.5	IO camera networks obtained for the (a) 2D board using MATLAB, (b) 3D field using Agisoft Metashape and (c) 3D field using Australis.	94
4.6	Gaussian radial and tangential distortion profiles generated from the obtained IO parameters with the 2D board using MATLAB (MAT), 3D field using Agisoft Metashape (AM), and 3D field using Australis (AUST). Here, the graphs represent the distortion originating at the image centre (u_o, v_o) and finishing at the maximum radial distance when using either two- (K2) or three- (K3) radial distortion coefficients are used as part of the geometric calibration.	96
4.7	Distortion models for the (a,b) 2D board using MATLAB, (c,d) 3D field using Agisoft Metashape and (e,f) 3D field using Australis, showing distortion magnitude (pix), distortion normals ($\nearrow$), principal point ($\bullet$) and image centre (320,256)($\times$).	98
4.8	RO camera networks obtained for the RGB2-WWP with (a) 2D board using MATLAB, (b) 3D field using Agisoft Metashape and (c) 3D field using Australis.	101
4.9	Lenton Lodge Gatehouse, Nottingham, UK.	106
4.10	An example WWP-RGB1 image pair generated from SuperPoint + LightGlue (Sp+Lg) in Deep-Image-Matching (DIM).	107

4.11	Combined TIR-RGB bundle block adjustment camera network for Lenton Lodge Gatehouse, showing the single strip of WWP images and two strips of RGB1 images.	108
4.12	Generated 3D models from the combined TIR-RGB bundle block adjustment of Lenton Lodge Gatehouse, textured with (a) RGB1 and (b) WWP thermal images.	110
4.13	Castello Di Arco's medieval frescoes, Trentino, Italy.	113
4.14	Castello Di Arco's medieval frescoes, featuring (a) the stereo rig featuring the WWP and RGB2 cameras, and (b) the stereo rig camera network.	113
4.15	An example of geometrically un-verified SIFT matches obtained in COLMAP between sequential mono-modal (a) RGB2 and (b) WWP images, resulting in 5,924 and 22 matches, respectively.	115
4.16	Generated textured models of Castello Di Arco's medieval frescoes, with the (a) RGB results showing the visible signs of deterioration and (b) showing temperature range visualising the most significant (cold) areas of water ingress.	117
5.1	Comparing a hand-crafted image matching approach (SIFT feature detector (Lowe, 1999), nearest neighbour matching (Riba et al., 2019) and PyDEGENSAC geometric verification (Mishkin et al., 2015)) for (a) RGB-RGB, (b) TIR-TIR, and (c) TIR-RGB image pairs, with the number of matched features being 724, 192, and 13, respectively. . .	133
5.2	Proposed 2D-2D fusion methodology utilising multi-modal image matching for IRT-3DDF.	135
5.3	(a) Dunrobin House, Warwickshire, UK; (b) Lenton Lodge Gatehouse, Nottinghamshire, UK.	137
5.4	Example image sequences for (a) TIR and (b) RGB images of Lenton Lodge Gatehouse.	139

5.5	TIR (red) and RGB (blue) camera networks for (a) Dunrobin House and (b) Lenton Lodge Gatehouse.	141
5.6	16-bit TIFF thermal images (a) before and (b) after image enhancement (histogram normalization and UnSharp Masking)	142
5.7	Deep-Image-Matching (DIM) modules and workflow.	143
5.8	Multi-modal image matching results for a TIR-RGB pair of Dunrobin House using both DL and hand-crafted matchers.	144
5.9	Multi-modal image matching results for a TIR-RGB pair of Lenton Lodge Gatehouse using both DL and hand-crafted matchers.	145
5.10	Distribution of GCPs (red) and CPs (yellow) on the 3D polygonal models of (a) Dunrobin House and (b) Lenton Lodge Gatehouse. . .	150
5.11	Generated 3D tie points for Dunrobin House, featuring the number of tie points generated for each multi-modal image matching approach. Results of the Agisoft Metashape 3D-3D fusion method have also been included for comparison.	152
5.12	Generated 3D tie points for Lenton Lodge Gatehouse, featuring the number of tie points generated for each multi-modal image matching approach. Results of the Agisoft Metashape 3D-3D fusion method have also been included for comparison.	153
5.13	(a) RGB and (b) TIR textured 3D polygonal models for Dunrobin House.	160
5.14	(a) RGB and (b) TIR textured 3D polygonal models for Lenton Lodge Gatehouse	162
5.15	Quantile thermal orthomosaic for Dunrobin House	163
5.16	Quantile thermal orthomosaic for Lenton Lodge Gatehouse	164
6.1	A comparison between (a) a 2D image and (b) a 3D model texture generated for Dunrobin House.	174

6.2	Visualisation of WWP vignetting through inter-image pixel-wise metrics, featuring inset sample images (comprising the thermos and cooler) from the TIR sequence throughout the ~ 105 min ‘dynamic’ scene experiment.	182
6.3	An example of the difference in FPN between sequential, stationary images at Castello di Arco.	183
6.4	Example WWP screenshots from Lenton Lodge Gatehouse, showing the execution of an automatic NUC outside of the programmed 3 min intervals (at the bottom of each figure, the duration between NUCs is apparent, here changing at ~ 2 min 30 secs as opposed to 3 mins).	185
6.5	Thermitage’s datasets, including: (a) Arco Castle, Italy (ARCO); (b) Lenton Lodge Gatehouse, UK (GATE); (c) Dunrobin House, UK (DUNR); and (d) Chesterton Windmill, UK (WIND).	191
6.6	Examples TIR images of the thermal-specific markers, featuring the (a) 3D test-field from Section 4.3.3 and (b) Dunrobin House.	193
6.7	An example TIR image of Dunrobin House with annotated (a) ‘natural’ and (b) thermal-specific markers.	194
6.8	An example of thermally-textured parametric models, providing the means for thermal images to be integrated into (H)BIM and asset management processes (Biswanath et al., 2023).	196
A.1	Lenton Lodge Gatehouse, also called ‘Lenton Lodge’ or simply Wolaton Park ‘Gateway’, photographed by J. Spree (data unknown). Here, the gatehouse can be seen open with Nottingham canal beneath the foreground (Lenton Times, 2006).	248
A.2	Lenton Lodge Gatehouse’s mechanical locking system, featuring (a) the underground ‘winding box’ and (b) central gate’s ornate pins (Lenton Times, 1983).	249

A.3	Comparison between the façade of Lenton Lodge Gatehouse in (a) 2007, after extensive restoration works by Check Whyte Industries, and (b) 2024, in its current state. Noticeably, the central parapet shows significant difference between the (a) recently restored and (b) heavily deteriorated stone above the traceried panels and central plaque, likely due to leaking from the roof. (This is Nottingham, 2009; Wood, 2009).	250
A.4	Images from the internal building survey of Lenton Lodge Gatehouse. Here, the rooms behind the (a) north turret, (b) central traceried panels (the other side of the circular windows) and (c) bronze plaque (behind the brick wallpaper) are shown, with visible signs of water ingress, material decay and damage identifiable.	251
A.5	Castello di Arco and its ‘Torre Renghera’ (bell tower), overlooking the Sarca Valley down to Riva del Garda (Trentino.com, 2025). . .	252
A.6	Castello di Arco depicted by (a) Matthias Burglechner (date unknown) and (b) Albrecht Dürer (1494), the latter of which now resides in the Louvre museum, Paris, France. The paintings are characterised by the large rocky ridge on top of which Castello di Arco resides, surrounded by cypresses, holm oaks, olive groves and cultivated vines (Storia dell’Arte, 2025).	253
A.7	Castello di Arco’s famed frescoes, depicting court life include: (a) ladies gesticulating whilst playing dice, (b) court ladies and gentlemen playing chess, and (c) a gentlemen and two ladies collecting roses for garlands (Storia dell’Arte, 2025).	254
A.8	Map of Lenton Lodge Gatehouse, Lenton, Nottingham, Nottinghamshire, UK.	255
A.9	Map of Castello di Arco, Trentino, Italy.	256
A.10	Map of Dunrobin House, Warwickshire, UK.	257

B.1	IRT-3DDF ‘Cheatsheet’	259
-----	---------------------------------	-----

Abbreviations

2.5D Two-and-a-half Dimension(al).

2D Two Dimension(al).

3D Three Dimension(al).

3DDF 3D-Data Fusion.

AECO Architecture, Engineering, Construction & Operations.

AI Artificial Intelligence.

ALIKE Accurate and Lightweight Keypoint Detection and Descriptor Extraction.

ALIKED A Lighter Keypoint and descriptor Extraction network via Deformable transformation.

API Application Programming Interface.

ArUco Augmented Reality University of Cordoba.

ASIFT Affine Scale Invariant Feature Transform.

ASTM American Society for Testing and Materials.

BA Bundle Adjustment.

BIM Building Information Modelling.

BRIEF Binary Robust Independent Elementary Features.

CBBA Combined Bundle Block Adjustment.

CCD Couple-Charged Device.

CFD Computational Fluid Dynamics.

CMOS Complementary Metal-Oxide-Semiconductor.

CNN Convolutional Neural Network.

CP Check Point.

CPU Central Processing Unit.

CRM Conservation, Repair and Maintenance.

CRP Close-Range Photogrammetry.

CV Computer Vision.

DCH Digital Cultural Heritage.

DeDoDe Detect, Don't Describe.

DIM Deep Image Matching.

DISK DIScrete Keypoints.

DL Deep Learning.

DN Digital Number.

DSLR Digital Single-Lens Reflex.

DUNR Dunrobin House.

EO Exterior Orientation.

FAST Features-from-Accelerated-Segment Test.

FGR Fast Global Registration.

FIJI Fiji Is Just ImageJ.

FIR Far-InfraRed.

FOV Field-of-View.

FPA Focal-Plane Array.

FPFH Fast Point Feature Histogram.

FPN Fixed-Pattern Noise.

GATE Lenton Lodge Gatehouse.

GCP Ground Control Point.

GIS Geographical Information Systems.

GNSS Global Navigation Satellite System.

GPR Ground Penetrating Radar.

GSD Ground Sampling Distance.

GUI Graphical User Interface.

H-BIM Historic Building Information Modelling.

Hg_{1-x}Cd_xTe Mercury Cadmium Telluride.

ICP Iterative Closest Point.

IHS Intensity-Hue-Saturation.

IMU Inertial Measurement Unit.

IO Interior Orientation.

IR InfraRed.

IRT InfraRed Thermography.

IRT-3DDF InfraRed Thermography 3D-Data Fusion.

Lg LightGlue.

LIBT Local Intensity Binary Transform.

LiDAR Light Detection and Ranging.

LIFT Learned Invariant Feature Transform.

LNIFT Locally Normalized Image Feature Transform.

LoD Level of Detail.

LWIR Long-Wave Infrared.

M3C2 Multiscale Model-to-Model Cloud Comparison.

MIM Maximum Index Mapping.

MRE Mean Reprojection Error.

MSE Mean Squared Error.

MVS Multi-View Stereo.

MWIR Mid-Wave Infrared.

NDT Non-Destructive Testing.

NeRF Neural Radiance Fields.

NIR Near-InfraRed.

NN Nearest Neighbour.

NRD Non-linear Radiation Distortion.

NUC Non-Uniformity Correction.

NURB Non-Uniform Rational B-spline.

ONS Office for National Statistics.

ORB Oriented FAST and rotated BRIEF.

PC Phase Congruency.

PCC Participant, Concept, Context.

RANSAC Random Sample Consensus.

RGB Red-Green-Blue.

RGB-D Red-Green-Blue-Depth.

RGB1 RGB1 - Sony α 7RII.

RGB2 RGB2 - Nikon D750 *FX*.

RIFT Radiation Invariant Feature Transform.

RMSE Root Mean Square Error InfraRed.

RO Relative Orientation.

ROIC Readout Integrated Circuit.

SfM Structure-from-Motion.

Sg SuperGlue.

SIFT Scale Invariant Feature Transform.

SLAM Simultaneous Localisation and Mapping.

SLS Selective Laser Sintering.

Sp SuperPoint.

Sp+Lg SuperPoint + LightGlue.

SRIF Scale and Rotation Invariant Feature Transform.

SURF Sped-Up Robust Features.

SWIR Short-Wave Infrared.

TIFF Tagged Image File Format.

TIR Thermal InfraRed.

TLS Terrestrial Laser Scanning.

TST Total Station Theodolite.

TTIS Three-dimensional Thermal Imaging System.

UAV Unmanned Aerial Vehicle.

UKTA UK Thermography Association.

USA United States of America.

UV Ultra-Violet.

VIS Visible.

VO_x Vanadium Oxide.

WWP Workswell WIRIS Pro.

XR eXtended Reality.

Chapter 1

Introduction

Comprehensive documentation is the foundation of effective conservation, repair and maintenance (CRM) practices for cultural heritage. The susceptibility of historic buildings, archaeological sites and cultural artefacts to natural and human disasters, coupled with ever-increasing pressures imposed by our changing climate (The Intergovernmental Panel on Climate Change (IPCC), 2022, p. 141), necessitates detailed preservation of both tangible and intangible heritage (Letellier et al., 2011). Architectural heritage, comprising all historic structures of cultural significance either protected or unprotected by regulatory statutes, represents 30% of all existing European buildings (European Statistical System (EuroStat), 2021), with $\sim 22\%$ of dwellings (Nicol et al., 2014, pp. 4-5) and $\sim 45\%$ non-domestic properties (Pout et al., 1998, p. 10) in England and Wales dated pre-1919. According to the Office for National Statistics (ONS), the age of a property is the biggest determining factor in a building's energy efficiency (Office for National Statistics (ONS), 2022), with historic buildings less resilient and ill-adapted to the challenges presented by higher temperatures, dynamic precipitation patterns and increasing pollution rates (Hao et al., 2020). To protect historic buildings, and provide spaces where occupants can experience the inherent value of our shared heritage (Whitman

et al., 2020), strategies for improving the energy efficiency and thermal comfort of vernacular architecture through retrofitting are needed, taking a holistic approach to CRM practices through low-cost, real-time and efficient methods (Adamopoulos & Rinaudo, 2019; Costa-Carrapiço et al., 2022; Historic England, 2018b; Martínez-Molina et al., 2016).

To this end, non-destructive testing (NDT) techniques, comprising a wealth of scientific methods for assessing subjects without exacting damage, represents a critical tool for heritage professionals looking to undertake thorough assessments of historic buildings and their curtilage. For architectural heritage, encapsulated by variability in building materials, practices and construction, NDT provides a means of generating rapid insights into the current state of an asset that can inform future decision-makers (El Masri & Rakha, 2020; Tejedor et al., 2022). To address the inherent complexity of architectural heritage, the collection of geometric data at a project’s inception has become a recognised pre-requisite (Denard, 2012), utilising image-, range-, or a combination of ‘reality-based’ surveys to generate accurate 3D models of an asset and provide the foundation for additional NDT investigations (De Fino et al., 2023). Specifically for built heritage, derived 3D models can allow for: rapid documentation (Oreni et al., 2014), holistic asset information management (Brumana et al., 2019), temporal analyses (Rodríguez-Gonzálvez et al., 2017), and virtual tools (Osello et al., 2018). By utilising a holistic, multi-scale and multi-disciplinary approach to NDT, conservators can make informed decisions regarding a historic building whilst maintaining its cultural value and significance (Historic England, 2020; Kioussi et al., 2013).

The analysis of temperature holds a central role in the assessment of historic buildings, with thermal data pivotal in understanding how a building performs for its occupants and how building elements behave under different environmental conditions (Historic England, 2025b). InfraRed Thermography (IRT), an established non-invasive, non-contact and non-destructive imaging technique, is capable of pro-

ducing rapid, real-time, highly-accurate and easily-interpretable measurements of a building's temperature. IRT provides 2D imagery representative of the infrared radiation emitted by objects above absolute zero (i.e., $T > 0\text{K}$), allowing for qualitative and quantitative assessments of temperature differences to be generated (Kylili et al., 2014; Usamentiaga et al., 2014). Developed out of the USA defence market in the 1950s, TIR cameras have become increasingly available and affordable since the turn of the 21st Century, enabling IRT to become a critical tool for heritage professionals and the advancement of NDT applications for architectural heritage (Corsi, 2015; Khodayar et al., 2016). Importantly, IRT allows for building diagnostics to be preventative as opposed to responsive, allowing building lifecycle costs, construction waste and climate impact to be reduced.

However, whilst IRT has demonstrated its competency as a stand-alone NDT technique, a growing body of research has emerged looking to exploit innovative photogrammetric computer vision and surveying engineering techniques to fuse complementary 2D and 3D datasets of contemporary (Adán et al., 2020; Ramón et al., 2022) and historic (Sutherland et al., 2023) buildings. Data fusion, as defined by Remondino and Rizzi (2010), is the process of combining datasets of varying source (sensor) and resolution into a unified product. The benefits of data fusion lie in the ability to use the strengths of individual data sources to complement and supplement the limitations of another, resulting in a better product than individual elements could produce. The combination of IRT with additional surveys such as ground penetrating radar (GPR)(Biscarini et al., 2020; Garrido et al., 2022), un-manned aerial vehicles (UAV)(Collaro et al., 2023; Hill et al., 2020), close-range photogrammetry (CRP)(Adamopoulos et al., 2020a; Delegou et al., 2019), terrestrial laser scanning (TLS)(Ceccarelli et al., 2018; Nardi et al., 2015), and building information modelling (BIM) or geographical information systems (GIS)(Etxepare et al., 2020; Tsilimantou et al., 2020) has demonstrated the benefit gained from integrated surveys and novel fusion methods applied to CRM practices. The combination of

geometric and radiometric data for architectural heritage bolsters their individual use, enabling interpretations of complex buildings to be made with greater objectivity (Dias et al., 2021).

InfraRed Thermography 3D-Data Fusion (IRT-3DDF), the combination of IRT, encompassing quantitative temperature data from thermal images, with accurate geometries obtained from image- or range-based surveys, allows for the generation of accurate 3D models supplemented with thermal information pivotal in the assessment of historic buildings. The rise of IRT-3DDF from the early-2000s has tracked the broader developments of heritage sciences by exploiting the growing market for TIR cameras, Structure-from-Motion (SfM) photogrammetry and computational efficiency for the generation of 3D models, encapsulating both qualitative thermal information and quantitative temperature values (Sutherland et al., 2023). Furthermore, the increasing availability, affordability and functionality of TIR cameras has come in tandem with the development of three-dimensional thermal imaging systems (TTIS), rigs comprising multiple sensors fixed on stationary, mobile or aerial platforms. These critical developments have laid the foundation for the continued development of IRT-3DDF and the growth of its subsequent applications for architectural heritage.

However, critical bottlenecks have endured from its inception by photogrammetric computer vision and geomatics communities. Firstly, the inherent low spatial resolution of infrared (IR) sensors (typically between 180×120 - 640×480 pix) means the determination of correspondences between images makes traditional SfM processes challenging at varying scales. Notably, both the geometric (i.e., distortion, narrow FOV) and radiometric (i.e., low contrast, high noise) characteristics of TIR cameras have made the extraction of reliable temperature information a challenge, leading to a lack of research where any meaningful quantitative temperature analysis can be undertaken. These challenges, coupled with effective sensors, methods and validation, provide an opportunity for IRT-3DDF to exploit novel tools generating

complete geometries of architectural heritage that can encapsulate precise thermal information. By doing so, the application of IRT-3DDF can become an established process for heritage professionals and CRM practices, tackling some of the most pressing concerns facing cultural heritage.

1.1 Thesis Aims & Objectives

Whilst current IRT-3DDF approaches are capable of creating 3D thermal models of both historic and contemporary buildings, more attention is needed to determine methods capable of handling the inherent complexity of TIR images when applied for photogrammetric 3D-data fusion (Adamopoulos et al., 2020b). With an ever-increasing market for TIR cameras, such methods need to be generalisable to different cameras, applicable to a variety of existing structures and deployable with greater automation. Furthermore, the current absence of IRT-3DDF studies correlating *known* and *observed* temperature readings offers an opportunity to assess the quality of derived 3D thermal models and the ability for reliable quantitative temperature measurements to be used for CRM practices. Therefore, the aim of this thesis is to develop a fully-automatic IRT-3DDF method capable of generating accurate and precise 3D thermal models comprising both detailed geometries and quantitative temperature measurements. To achieve this aim, this thesis shall look to apply the following objectives exploring IRT-3DDF for architectural heritage:

- **Objective 1:** Define the essential theoretical and practical requirements for the successful implementation of IRT-3DDF.
- **Objective 2:** Determine robust and rigorous methods for the geometric and radiometric calibration of TIR cameras employed within photogrammetric data fusion

- **Objective 3:** Establish a novel, cost-effective and highly-flexible data fusion method obviating sensor pre-calibration, manual co-registration or auxiliary georeferencing information.

1.2 Thesis Structure

The structure of this thesis will look to expand the aforementioned aims and objectives, with a central literature review (Conceptualisation) being followed by individual chapters addressing each objective, entitled: Foundation, Calibration and Generation. Whilst each chapter is focussed on a singular objective, the sequential structure of the thesis will allow conclusions to be carried forward to inform the following chapter. This thesis shall culminate with the application of IRT-3DDF to central case studies, applying all the findings of the thesis to real-world examples of thermal modelling. Here, the developed principles, methods and applications of the work can be examined, demonstrating the future avenues and opportunities for IRT-3DDF after the conclusion of the thesis. The following sections provide an overview of the four central chapters to outline their originality for IRT-3DDF.

Conceptualisation of IRT-3DDF

To begin the thesis, the field of IRT-3DDF is defined, classified and critiqued. This includes: (1) the origins of IRT-3DDF as a discipline and the tandem fields responsible for its conception; (2) the definition of data fusion approaches, sensors and systems within IRT-3DDF; and (3) the enduring challenges and shortcomings represented by the field. Critically, this chapter carves out the originality of the thesis, demonstrating the current under-representation of 2D-2D fusion methods and the opportunities presented by novel deep learning-based image matching models.

Foundation of IRT-3DDF

At present, radiometric accuracy of IRT-3DDF methods is largely ignored, with studies rarely assessing the accuracy of quantitative temperature measurements extracted from 2D thermal images or 3D thermal models (Adán et al., 2023b). As an understanding of InfraRed Thermography as a science has been denoted as a fundamental pillar for its effective use within building diagnostics (Balaras & Argiriou, 2002) and thermal modelling (Adán et al., 2020), this chapter address Objective 1 by demonstrating the acquisition and application of this knowledge in the context of IRT-3DDF. Expanding the scoping review by Sutherland et al. (2023), within which IRT-3DDF as a field was first introduced, this chapter presents: (1) a sequential framework that determines the radiometric performance of TIR cameras, (2) the determination of preferential environmental conditions for IRT-3DDF surveys, and (3) the required geometric and radiometric control necessary for evaluating the accuracy of 3D thermal models. Importantly, this chapter will demonstrate, using a central case study of the thesis, the need to balance photogrammetric and radiometric principles within IRT-3DDF, where optimising one variable may affect the performance of the other.

Calibration of IRT-3DDF

The determination of precise and reliable interior orientation (IO) parameters for TIR cameras is critical for their subsequent use in photogrammetric processes. In addition, the relative orientation (RO) between fixed sensors, those featured within the same camera housing or configured in a stereo system, can provide a useful constraint for several IRT-3DDF methods (Adamopoulos et al., 2020b; Alba et al., 2011; Dlesk et al., 2021). Addressing Objective 2, expanding the research paper from Sutherland et al. (2025a), this chapter evaluates the determination of accurate IO and RO parameters through a novel IRT-3DDF calibration set-up.

Importantly, the originality of this chapter resides in the comparison of intrinsics and extrinsics derived from different calibration approaches, commercial software and parameter configurations, informative for photogrammetric data fusion using modern TIR sensors and platforms.

Generation of IRT-3DDF

The task of developing a robust and efficient method for IRT-3DDF requires careful consideration for the choice of sensors, the ‘space’ in which fusion is undertaken and the scalability of methods for specific applications. Addressing Objective 3, this chapter presents the collaborative paper of Morelli et al. (2025b), demonstrating the innovative use of deep learning-based image matching for 3D thermal modelling, a previously under-represented facet of IRT-3DDF. This chapter shows the ability for infrared and visible spectrum images to be *directly* matched and oriented within a (first-of-its-kind) fully-automatic combined bundle block adjustment, removing the need for manual co-registration, auxiliary positioning data or synchronous data capture. An evaluation of the accuracy of derived 3D thermal models is undertaken, assessing both the geometric accuracy of derived models and the radiometric accuracy of derived temperature measurements. Finally, a demonstration of the value that can be extracted from 3D thermal models is given, showing the potential benefits for CRM practices applied to architectural heritage.

Chapter 2

Conceptualisation of IRT-3DDF

Contents

2.1	Introduction	11
2.2	IRT-3DDF: A Definition	12
2.3	IRT-3DDF: State-of-the-Art	16
2.3.1	TIR-Only	16
2.3.2	2D-2D Fusion	18
2.3.3	2D-3D Fusion	23
2.3.4	3D-3D Fusion	29
2.3.5	Three-dimensional Thermal Imaging Systems (TTIS) . .	32
2.3.6	IRT-3DDF & Complimentary Technologies	34

2.1 Introduction

The conception of InfraRed Thermography 3D-Data Fusion (IRT-3DDF) from the early-2000s has tracked the broader trends of Digital Cultural Heritage (DCH), exploiting the growing availability of commercial sensors, increasing functionality of structure-from-motion (SfM) photogrammetry and ever-improving efficiency of modern computing (Adán et al., 2021; Cho et al., 2015). Whilst IRT-3DDF research has grown considerably over the past ten years, critical bottlenecks have endured from its inception. Firstly, the inherent low spatial resolution of thermal images (typically between 180×120 - 640×480 pix) means the determination of correspondence between (and *across*) images makes photogrammetric processes challenging, hampering the generation of complete geometries solely from TIR images and SfM. In addition, both the geometric (i.e., distortion, narrow FOV) and radiometric (i.e., low contrast, vignetting, NUCs) characteristics of the camera have made the acquisition of accurate temperature information a constant challenge that needs to be addressed before any meaningful analysis can be undertaken. Therefore, these limitations have required methods to determine ways of generating complete geometries of built assets that can encapsulate precise thermal information, using the strengths of individual sensors to supplement the shortcomings of other tools. To uncover these methods, this chapter presents an extensive review of existing studies within IRT-3DDF. Firstly, definitions of 3D-Data Fusion and IRT-3DDF are provided to rationalise the creation of IRT-3DDF as a term and determine the delineation, intersections and classifications within the field. Subsequently, an investigation of the critical methods, technologies and practices within IRT-3DDF are explored, with a particular focus on the benefits provided by the ‘space’ in which fusion is undertaken. Finally, the merits and limitations of existing methods are determined, identifying critical research gaps and opportunities for future work.

2.2 IRT-3DDF: A Definition

As defined by Remondino and Rizzi (2010), in the context of reality-based data capture, data fusion is the process of combining datasets of varying source (sensor) and resolution into a unified product. The benefits of data fusion lie in the ability to use the strengths of individual data sources to complement and supplement the limitations of another, resulting in a better product than individual elements could produce. For architectural heritage, the fusion of geometric and radiometric data bolsters their individual use, enabling interpretations of complex buildings to be made with greater objectivity (Dias et al., 2021). Previous explorations of 3D-Data Fusion (3DDF) has looked to classify data fusion through purpose-, data- and dimension-based levels (Ramos & Remondino, 2015); outline applications and opportunities for building diagnostics (Ramón et al., 2022); and demonstrate the integration of additional NDT datasets for heritage documentation (Adamopoulos & Rinaudo, 2019).

Although terms have been used to define the creation of 3D thermal models, including ‘data integration’ (Rizzi et al., 2007), ‘3D thermal computer vision (3D-TCV)’ (Adán et al., 2021), ‘simultaneous localization and thermal mapping (SLATM)’ (Ramachandra et al., 2018) and simply ‘thermal modelling’ (Dahaghin et al., 2021), an acknowledgment of the breadth of technologies encapsulated within the research must be realised (Figure 2.2). As part of Sutherland et al. (2023)’s scoping review, IRT-3DDF was defined as “*the registration of 2D-thermographic data with high parity, geospatially derived 3D-data to provide a unified, geometrically accurate product suitable for representing architectural heritage*” (Sutherland et al., 2023, p. 5). Utilising Khalil and Tricco (2022)’s PCC format (Khalil & Tricco, 2022, p. 179), IRT-3DDF research was included based on its central criteria, including the ‘Participant’ (InfraRed Thermography), ‘Concept’ (3D-Data Fusion) and ‘Context’ (Architectural Heritage)(Figure 2.1). Therefore, for this thesis, the definition

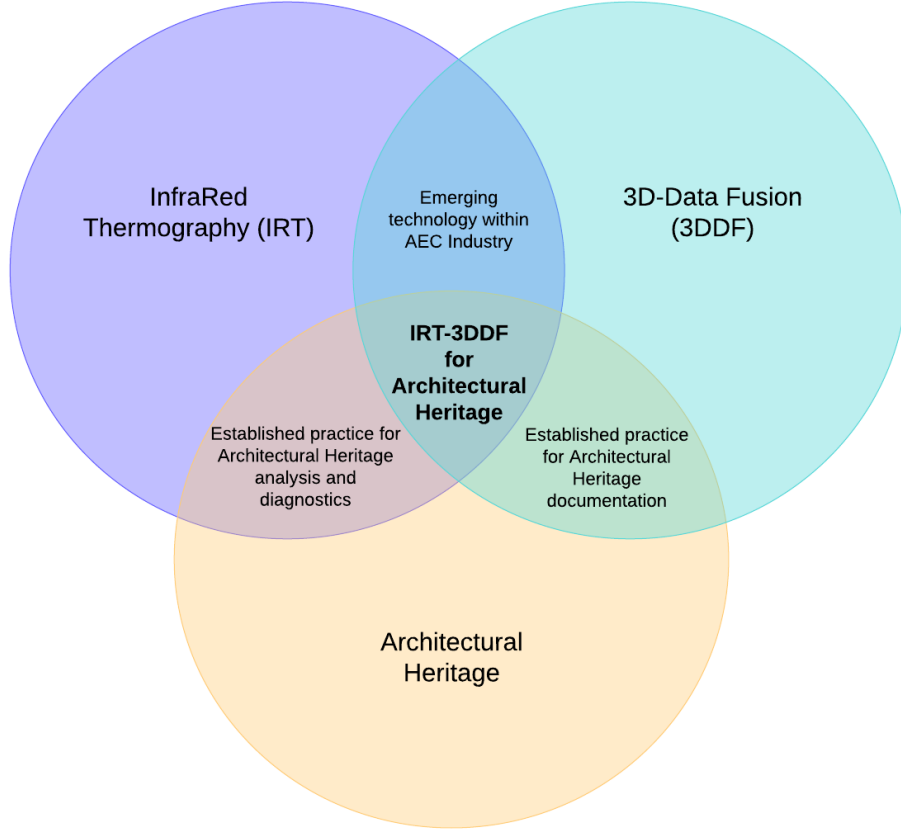


Figure 2.1: Mapping the inter-connectivity of InfraRed Thermography (IRT), 3D-Data fusion (3DDF) and Architectural Heritage research (Sutherland et al., 2023).

of IRT-3DDF shall expand the context of IRT-3DDF to include *all* methods generating 3D thermal models, capturing relevant techniques across multiple scales, subjects and sensors.

Whilst the PCC classification provides clear indication of approaches used within the research, it fails to highlight the specific sensors used for fusion, which dramatically determine the available methods. Therefore, two important questions are needed for the comprehensive classification of IRT-3DDF research: (1) in what ‘space’ (i.e., 2D image or 3D object space) is fusion taking place; and (2) what sensors are being ‘fused’ together (i.e., TIR, VIS, RGB, TLS, parametric models)(Figure 2.3)? The first question looks to identify *where* data sources are being matched, corresponded or fused, which can be in the same or different ‘space’ to the thermal information (i.e., 2D–2D, 2D–3D, or 3D–3D). The importance of the

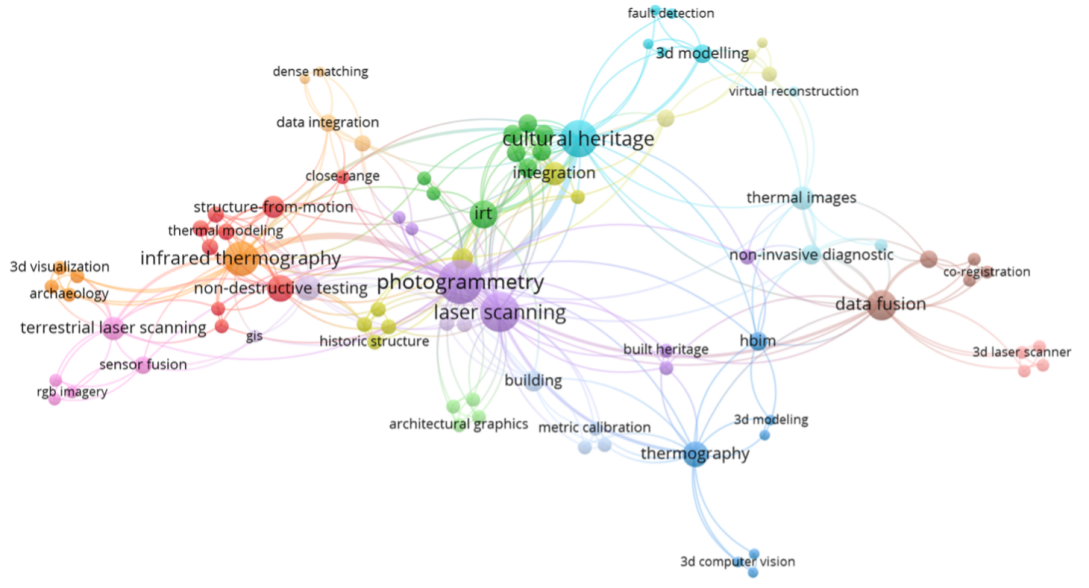


Figure 2.2: IRT-3DDF keyword co-occurrences representing 88 connected keywords within IRT-3DDF architectural heritage research (Sutherland et al., 2023). Relationships between keywords are determined by the proximity of bubbles and the link strength of joined keywords (van Eck & Waltman, 2010).

second question, denoting the ‘base’ onto which TIR is being ‘fused’, rests in the knowledge that certain methods will apply only when specific sensors are employed. For example, all TIR–VIS studies can lead with the assumption that both TIR and VIS sensors are fixed in relative orientation and that each pair of images will have the same distance (baseline) between one another. By answering these questions, IRT-3DDF can be summarised more succinctly and lead readers to the methods achievable with their given equipment.

Therefore, the following classifications denote the categories of this state-of-the-art review, breaking IRT-3DDF into the following groups based on their fusion ‘space’ and sensors:

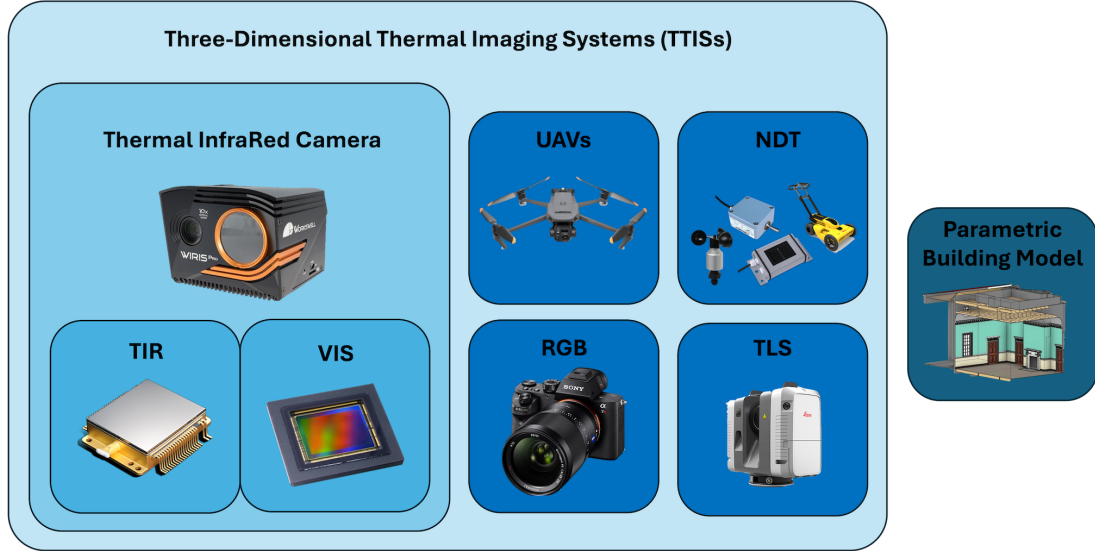


Figure 2.3: IRT-3DDF sensors and systems, denoting the individual sensors utilised within the research, their designation for the conceptualisation of IRT-3DDF and the configuration of technologies that can form a TTIS.

1. **2D–2D Fusion** – Direct matching of thermal and visible spectrum imagery in 2D image space, from which 3D models are subsequently generated and thermal information is fused using the relationship between images.
 - TIR–VIS
 - TIR–RGB
 - TIR–TLS (2.5D)
2. **2D–3D Fusion** – Thermal images are used to texture or assign temperature values to a 3D model by determining the reprojection of those images in the 3D object space.
 - TIR–VIS
 - TIR–RGB
 - TIR–TLS
 - TIR–Parametric Building Models

3. **3D–3D Fusion** – Independent 3D geometries are created separately from thermal images and a geometric ‘base’, fused together in 3D object space either by iterative closest point (ICP), common points or similar scaling methods.

- TIR–VIS
- TIR–RGB
- TIR–TLS
- TIR–Parametric Building Models

Whilst some methods may fall outside of these categories, the importance of categorising similar research enables: the comparison and contrast of common approaches, the critical analysis of decisions and outputs, and the evolution of methods from research groups and institutions. Furthermore, by assessing fringe topics related to IRT-3DDF, notably the generation of 3D models using solely thermal images (TIR-only), the configuration of TTISs, and the application of IRT-3DDF in tandem with complementary technologies, an appreciation for IRT-3DDF and the bounds it aims to push can be realised.

2.3 IRT-3DDF: State-of-the-Art

2.3.1 TIR-Only

Before exploring IRT-3DDF, we must first understand why there is a need for data fusion in the first place. This will allow us to determine the shortcoming of 3D reconstruction solely using TIR images and the nuisances of thermal image capture when tasked with generating accurate 3D thermal models. Furthermore, as TIR-only workflows are critical to other data fusion methodologies (notably 3D-3D fusion), we can gain valuable insights that help clarify additional IRT-3DDF

approaches. Firstly, it is evident that TIR-only methods are successful when surveys take into consideration the challenges presented using thermal images for SfM, designing methods that manage limitations from image capture (i.e., low spatial resolution and GSD) (Riveiro & Solla, 2016) to data processing (i.e., feature extraction and matching) (Ramachandra et al., 2018). Acknowledging the importance of methodical thermal image capture, Dlesk and Vach (2019) demonstrate a successful TIR-only method registering 474 out of 581 images with a 10 mm/pix ground sampling distance (GSD), prioritising high vertical and horizontal overlap for a two floor building. Similarly, Westfeld et al. (2015) shows how TIR-only methods can be scaled to whole buildings, using a UAV to capture 15,000 TIR images with a 26 mm/pix GSD with 90% forward and 70% side overlap (Figure 2.4). Although TIR-only approaches are unlikely to achieve the metric accuracy of other IRT-3DDF fusion methods (with Westfeld et al. (2015) producing a RMSE on CPs of 19 mm and average deviation from TLS scan of 25 mm), these considerations, when accompanied with diligent camera calibration (Chapter 4) and ground control (Section 3.5.4), enable dense 3D reconstructions using solely TIR images and SfM practices.

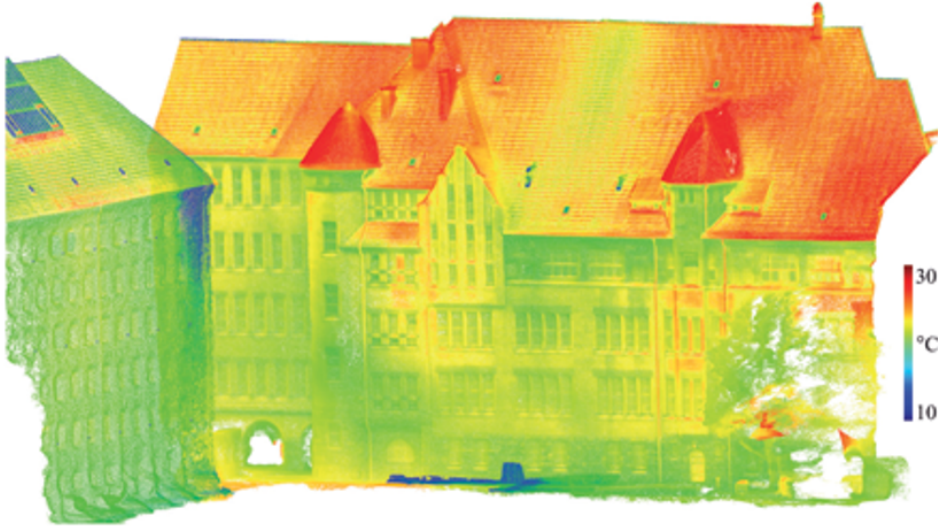


Figure 2.4: A TIR-only model produced by Westfeld et al. (2015), generated using 15,000 TIR images and a unmanned aerial vehicle (UAV).

In addition, the growth of SfM in parallel to IRT-3DDF has seen methodologies explore the ever-expanding capabilities of commercial photogrammetric software

to enable TIR-only reconstruction. Several works have demonstrated the capability of using thermal images with popular SfM software, including: Agisoft PhotoScan/Metashape (Dlesk & Vach, 2019; Paziewska & Rzonca, 2022), 3DZephyr (Clarkson et al., 2017; Maset et al., 2017) and VisualSFM (Westfeld et al., 2015). In a bid to improve the matching of TIR images for photogrammetric processes, González-Aguilera et al. (2013) utilised ASIFT to determine correspondence between thermal images for dense reconstruction and TIR-only ortho-thermogram generation. Vidas and Sridharan (2012), Prakash et al. (2006) and Ramachandra et al. (2018) test methods for monocular, stereoscopic and SLAM TIR-only modelling, respectively; outlining critical obstacles including sufficient thermal contrast, mono-modal feature matching and final model scaling. However, in the case of thermal SLAM (Ramachandra et al., 2018), when fusion is introduced (such as a supplementary RGB camera), results for real-time localisation dramatically increased, a pattern mirrored in several comparative IRT-3DDF studies (Hoegner et al., 2016; Patrucco et al., 2022b). The importance of methods exploring TIR-only 3D reconstruction cannot be understated, with critical considerations regarding survey planning, thermal data capture and image matching determining the success of photogrammetric results. Furthermore, as several methods rely on the creation of TIR-only models as initial steps for data fusion, value can be extracted by understanding how to maximise the success of these preliminary steps.

2.3.2 2D–2D Fusion

As demonstrated by TIR-only 3D reconstruction, the importance of determining reliable correspondences between images is a critical issue for IRT-3DDF; even more so when multi-modal images, those containing different electromagnetic wavelengths and sensor types, are matched together for fusion purposes. At the heart of 2D-2D fusion lies the problem of extracting reliable features for the accurate estimation of

orientation matrices representing the geometric transformation from one 2D image onto another. Although traditional feature extractors such as SIFT (Lowe, 2004), ASIFT (Yu & Morel, 2011) and SURF (Bay et al., 2006) have been comprehensively tested for TIR-TIR and TIR-VIS/RGB image matching (Ham & Golparvar-Fard, 2013a; Lagüela et al., 2012), most fail to determine sufficient reliable correspondences between images for Homography (H) or Fundamental (F) matrix determination (Figure 2.5). Weinmann et al. (2014) demonstrate one alternate approach to counter these issues, using ‘shape-based’ matching of TIR-RGB-D (depth) images and correspondence filtering for the projection of thermal information onto 3D points. For both TIR-VIS and TIR-RGB 2D-2D fusion, the determination of accurate correspondence has shaped the accuracy and automation of derived methods, inspiring new methods to address the challenges inherent to thermal images (i.e., resolution, contrast, non-uniformity).

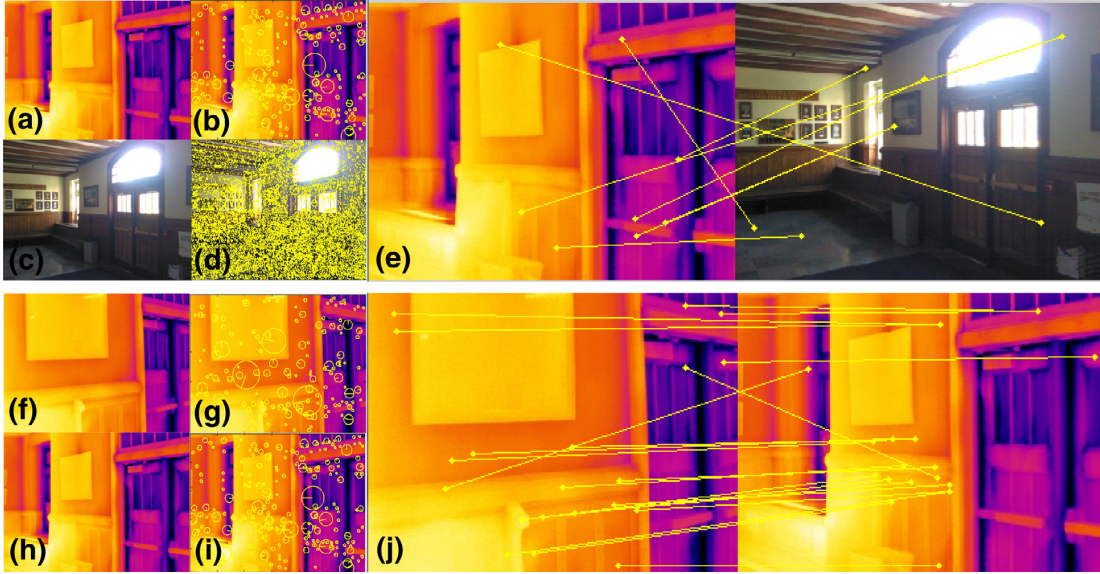


Figure 2.5: A visualisation of SIFT (Lowe, 2004) (f-j) mono- and (a-e) multi-modal feature matching, demonstrating the difficulty extracting (a,b,c,d,f,g,h,i) and matching (e,j) reliable correspondence across image modalities and varying perspectives (Ham & Golparvar-Fard, 2013a). Here, (b) 236 TIR and (d) 14,946 RGB features are extracted, yet result in only 8 (incorrect) multi-modal matches.

TIR–VIS

Common to several IRT-3DDF fusion methods, and a central principle of photogrammetry and computer vision, the use of stereo vision between fixed sensors is a helpful constraint to determine the relative pose between two cameras. This fixed relationship can either be in the form of a stereo camera rig, TTIS or a feature of the camera itself. For thermal cameras featuring both TIR and VIS sensors within the same camera housing, an assumption can be made that these sensors are in fixed relative orientation (RO) to one another, meaning pairs of images captured concurrently will have a constant distance. After determining the inability for reliable features to be automatically extracted for SfM (Figure 2.5), Ham and Golparvar-Fard (2013a)’s 2D-2D fusion method used a singular thermal camera (FLIR E60) with TIR and VIS sensors to determine the extrinsic parameters of the thermal camera by forming the essential matrix (E), mitigating the challenge of feature matching using a ‘supervised’ correspondence approach. By calculating the intrinsic and extrinsic parameters of the TIR sensor, a dense 3D reconstruction could then be obtained using MVS and co-registered with the 3D point cloud derived from the VIS sensor. Although few in number, hampered by challenges of correspondence, methods of 2D-2D fusion with TIR and VIS sensors provide a practical means for thermal modelling, using minimal equipment and supported by the stereo nature of both sensors.

TIR–RGB

Similarly to 2D-2D fusion with TIR and VIS sensors, TIR-RGB methods have looked to overcoming the challenges of correspondence by understanding the characteristics of each image and how these can be manipulated to determine better matches. Lagüela et al. (2012)’s seminal image fusion work fused TIR orthomosaics with RGB images matched using FAST (Rosten & Drummond, 2006) into IHS im-

ages containing the *Intensity* of the RGB and *Hue* and *Saturation* of TIR images. These, once matched using PhotoModeler, resulted in a 3D thermal model and ortho-rectified thermogram. Scaioni et al. (2012) and Previtali et al. (2013)’s combined TIR-RGB bundle block adjustments demonstrate rare examples of single bundles featuring both image modalities, with the former integrating a complementary ‘bi-camera’ method to fix cameras in stereo. However, both procedures are reliant on the manual identification of common points, reducing the overall automation of thermal modelling procedures. Similarly, Adamopoulos et al. (2020b), utilising manually-marked points for the projective transformation of TIR and RGB images, demonstrates the fusion of TIR, RGB and VIS sensors within the same 3D object space. Incorporating novel hand-crafted feature matchers, Lin et al. (2019b)’s four step method used RIFT (Radiation Invariant Feature Transform)(Li et al., 2020), a SIFT-like feature matcher, to match proximal RGB and TIR images prior to texture mapping. Unlike Lagüela et al. (2012) and Adamopoulos et al. (2020b), featuring manual area- or feature-based matching, this method dramatically reduces the manual tasks required in the thermal modelling process and provides greater robustness through a redundancy of detected features. The importance of utilising novel tools, adaptable for multi-modal feature matching, is a notable opportunity for IRT-3DDF.

TIR-TLS (2.5D)

With the application of existing feature matchers, and the advent of new deep-learning approaches, the matching of TIR images with other geometric products can become more prominent for IRT-3DDF (Morelli et al., 2024b). González-Aguilera et al. (2012) and Hoegner et al. (2018)’s research matched correspondence in TIR image with 2D images generated from TLS data (2.5D), demonstrating the capability of some operators to find correspondence between multi-sensor datasets. Capitalising on the growth of learning-based image matching, Elias et al. (2023)’s SLAM

fusion method utilises SuperGlue (Sarlin et al., 2019) for matching TIR-TLS (2.5D) images, using space resection to adhere thermal information on a mobile TLS point cloud (Figure 2.6). This represents one of the first learning-based feature matching methods for IRT-3DDF and demonstrates the possibilities of expanding commercial SfM and hand-crafted feature matchers for thermal modelling. To summarise, 2D-2D fusion within IRT-3DDF has relied on accurate correspondence between TIR images and an additional data source for the generation of accurate geometries; which, with the growth of configurable learning-based tools, could become more prevalent and automated for multi-modal modelling.

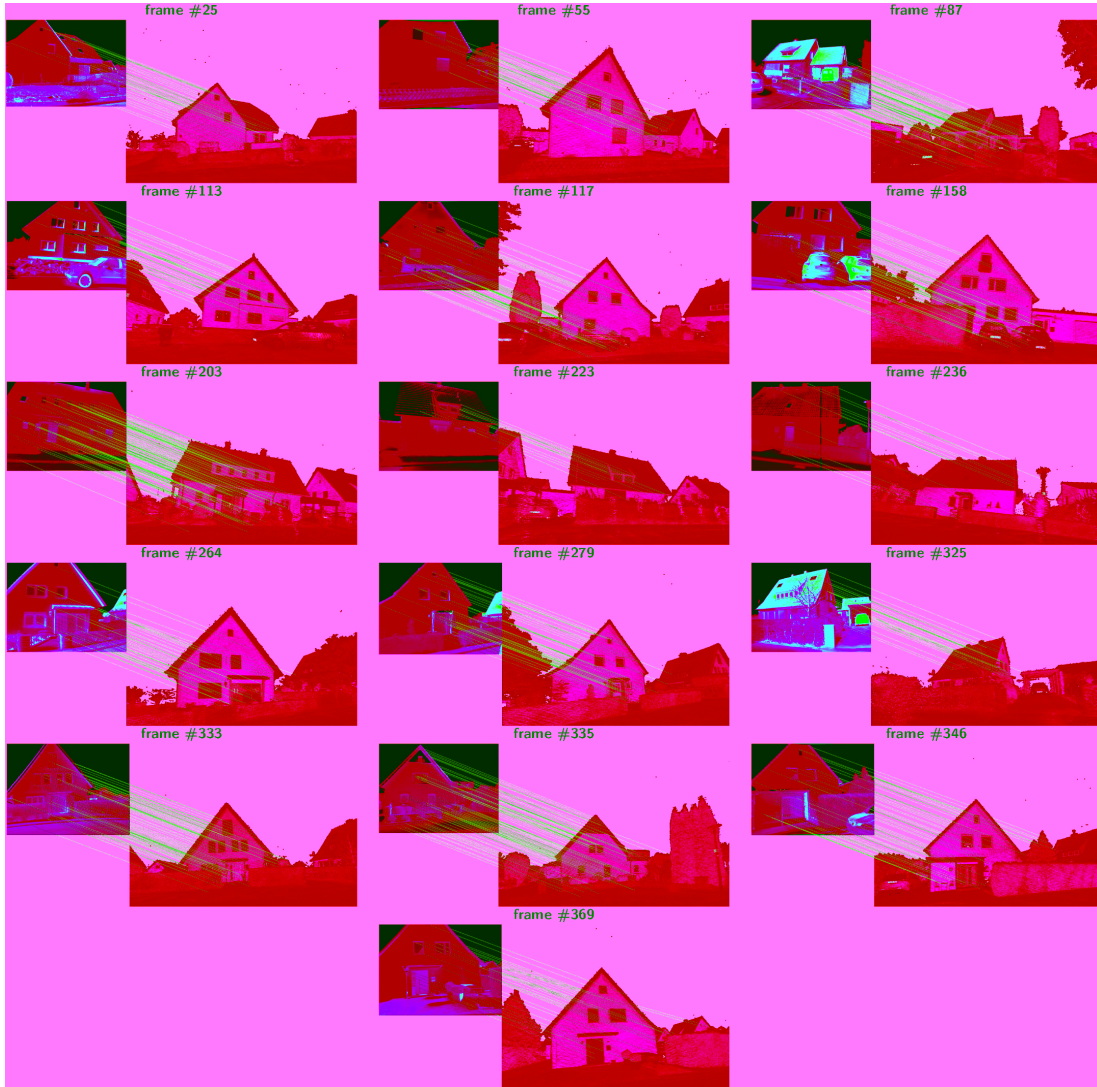


Figure 2.6: Demonstration of verified deep learning features extracted with SuperGlue, matching TIR images and TLS scans as 2D images (2.5D)(Elias et al., 2023).

2.3.3 2D–3D Fusion

The most popular IRT-3DDF approach, 2D-3D fusion looks to re-project thermal images onto 3D geometries from varying sensors and platforms. As opposed to other fusion methods, 2D-3D fusion looks to mitigate the issues of determining correspondence between, and across, modalities by using photogrammetric principles of perspective projection to assign thermal information from 2D images to 3D points. 2D-3D fusion methods have largely exploited stereo rigs, as introduced in Alba et al. (2011)’s seminal work, fixing sensors in relative orientation from which transformation parameters can be calculated. However, notably for IRT, which requires specific conditions for meaningful data to be captured, 2D-3D methods have also looked to capture datasets independently, allowing for thermal images to be projected onto pre-existing point clouds or building models generated at different times. Not only does this allow data to be captured under optimal conditions, but has enabled the comparison of: multiple thermal data acquisitions (Barbieri et al., 2023; Lewińska & Maciuk, 2020), varying TIR sensor resolutions (Adamopoulos et al., 2020b) and thermal texturing strategies (Hoegner & Stilla, 2007). 2D-3D fusion relies on the accurate determination of relative orientation between sensors and pose estimation of TIR images, using independent 3D geometries for the accurate representation of a building and its structure.

TIR–VIS

As previously highlighted, thermal cameras featuring both TIR and VIS sensors within the same camera housing enable a fixed distance between the two sensors to be adopted as a constraint. For IRT-3DDF, this allows the projection of thermal images onto 3D products generated from the proximal VIS sensor, applying a transformation key from one image to another. Dlesk et al. (2021) and Dlesk et al. (2022) present two methods for IRT-3DDF using a singular FLIR E95 camera:

a ‘Sharpening’ method and a ‘Reprojection’ method. Notably, the ‘reprojection’ method requires the geometric calibration of both sensors and estimation of relative pose between TIR and VIS sensors to undertake a spatial transformation. Both Lecomte et al. (2022) and Macher and Landes (2022) use the relative pose of TIR and VIS sensors within the same camera to colourise a TLS scan with temperature information, obtaining the RO of the TIR images through rigorous stereo calibration procedures (Figure 2.7). Such methods demonstrate the benefit of using the relationship between proximal sensors and the necessity of geometric calibration for IRT-3DDF.

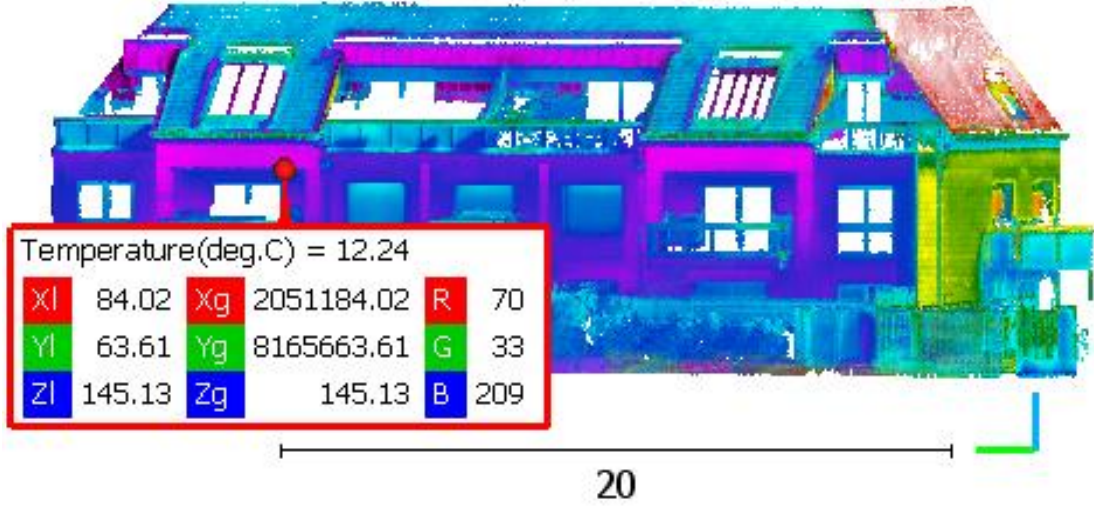


Figure 2.7: TLS scan with temperature values assigned per point (Lecomte et al., 2022).

To assess the use of a VIS sensor as a-priori information for the TIR sensor, Patrucco et al. (2022b) compare three SfM-based methods for thermal modelling. This compared: (1) a solely VIS SfM pipeline, (2) a solely TIR SfM pipeline, and (3) a fusion pipeline where the absolute EO parameters of (1) were used as initial approximates for TIR poses, subsequently optimised with GCPs. When applied to three heritage case studies of varying complexity, fusion methods were preferable over TIR-only SfM. This presents an important application for the VIS sensor as a means of providing initial estimation of camera poses for the accompanying TIR sensor. Although many studies discount the use of the VIS sensor due to its poorer

resolution compared to DSLRs, and inability to operate under preferable IRT conditions (i.e., darkness), opportunities exist to use it for initial approximation, as a fixed constraint and model quality metric.

To this end, demonstrating the fusion of multiple TIR, VIS and RGB sensors, Adamopoulos et al. (2020b) present a method using both the VIS sensor from a thermal camera in tandem with an additional DSLR. Here, to overcome the challenges presented by Dlesk et al. (2021), a transformation key between TIR-VIS pairs are determined for each acquisition strip as opposed to a universal key. After processing the DSLR RGB and VIS images in a combined bundle block adjustment, with a dense point cloud generated via MVS, VIS images are replaced with their corresponding TIR pair for texturing. This method shows not only a robustness for both medium-res. (640×480 pix) and high-res. (1024×768 pix) thermal cameras (Figure 2.8), but the usefulness of the VIS sensor for initial pose estimation. The use of a thermal camera’s VIS sensor can provide helpful information for the orientation of accompanying TIR images, and future research should analyse the benefit this sensor can provide for evolving IRT-3DDF.

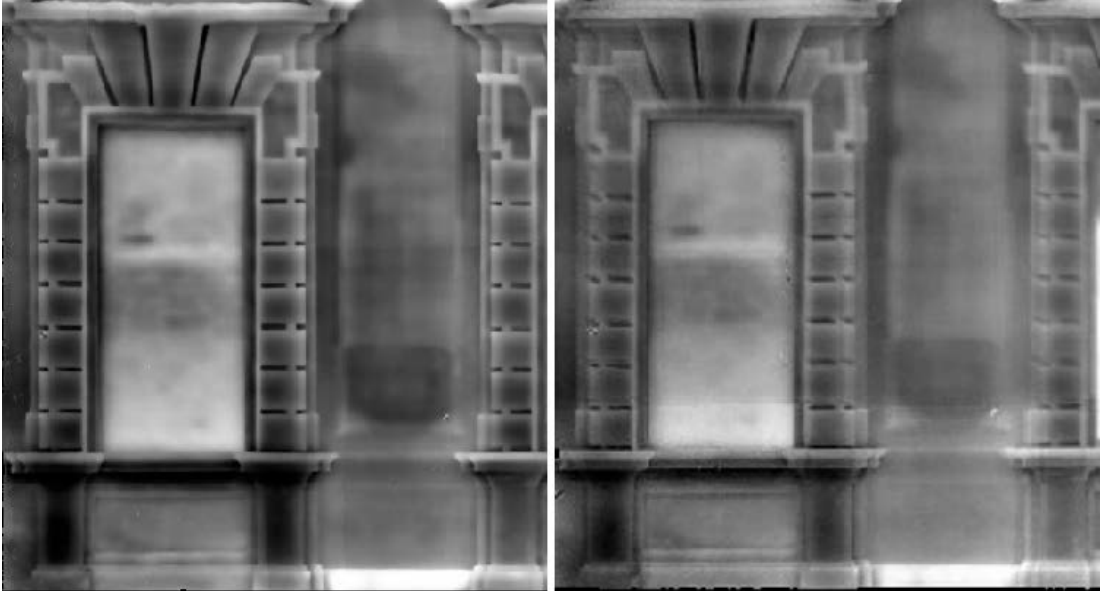


Figure 2.8: A comparison between textured thermal orthomosaics generated with a high-res. FLIR SC1030 (1024×768 pix)(left) and a medium-res. FLIR SC660 (640×480 pix)(right) (Adamopoulos et al., 2020b).

TIR–RGB

When high resolution RGB cameras are used in place of low resolution VIS sensors, model geometries are built exploiting both fixed and flexible TIR-RGB set ups. Javadnejad et al. (2020)’s fixed TIR-RGB work compared 2D (affine) and 3D (spatial) transformations to assign temperature values to RGB point clouds; determining the latter, whilst providing a more rigorous transformation, was outperformed due to the difficulties of achieving accurate TIR geometric calibration. Similarly, Yang et al. (2018) exploited the fixed relative orientation between two iPhone SE’s and a FLIR ONE to texture an RGB model using low-cost mobile sensors. Furthermore, where no fixed orientation is present, the extraction of corresponding features has enabled the texturing of existing RGB geometries. Lin et al. (2019b)’s critical four-step method, using RIFT to determine correspondence in TIR-RGB pairs, subsequently uses mono-plotting and image space resection to texture the RGB point cloud. This, combined with the final global pose refinement, provides an advantageous texturing method using independently acquired TIR and RGB images with high geometric and radiometric accuracy. The continued development of TIR-RGB methods exploiting independent cameras, not benefitting from the constraints of stereo sensors or fixed cameras, must look to plan surveys where photogrammetric practices are tailored towards the nuisances of TIR images and find innovative ways of determining correspondence through local features or visible GCP markers (Patrucco et al., 2020).

TIR–TLS

To generate rapid and dense geometries for thermal modelling, 3D TLS data has been used as a popular ‘wire frame’ for the adherence of 2D TIR images. The seminal work of 2D-3D fusion by Cabrelles et al. (2009) on the Petra World Heritage site demonstrates thermal texturing using space resection and known GCPs to assign

TIR images. Where no supplementary GCPs are used, 2D-3D fusion methods have either: (1) produced unscaled textured models from mobile platforms (Borrmann et al., 2014); (2) required manual identification of common points (Adamopoulos et al., 2021; Costanzo et al., 2014; Lagüela et al., 2011b; Lecomte et al., 2022; Macher & Landes, 2022; Maierhofer et al., 2011; Mileto et al., 2015) or linear features (Lagüela et al., 2013; Zhu et al., 2021) between 2D TIR images and 3D TLS scans to determine EO parameters; or (3) used image metadata (i.e. GNSS, IMU to align TIR images and TLS scans (Hess et al., 2015; Napolitano et al., 2019). Like many previous categories, the manual identification of correspondence between TIR and TLS data, or inclusion of GCPs, have hampered the automation of IRT-3DDF methods and requires the transformation of TLS data into more manageable forms for data fusion.

In a bid to overcome the manual processes involved in 2D-3D fusion, three-dimensional thermal imaging systems (TTISs), both custom and commercial, have allowed data to be registered when all data is collected within the same acquisition. Originating from Alba et al. (2011)’s ‘bi-camera’ study, comprising TIR-RGB stereo system to help align data to a TLS scan, IRT-3DDF studies have looked to increase the accuracy of thermal texturing using RGB data from both fixed and free poses. Adán et al. (2020)’s MoPAD system showcases how combining sensors on a mobile platform can generate dense thermal point clouds for energy monitoring and thermal comfort assessment (although building and computation dependent). Similarly, (Adán et al., 2021) calibrates a TTIS (inc., TLS: Riegl 3D-VZ400, TIR: FLIR A65, RGB: Nikon D90) for the projective transformation of TIR images onto a TLS scan, generating $360^{\circ} \times 100^{\circ}$ dense point clouds of a church nave. In summary, although many 2D-3D fusion methods using TIR-TLS data are constrained by the need for common points or features, often difficult to detect and match automatically, the generation of dense geometries can be the starting point for additional processing, data manipulation and value extraction (Patrucco et al., 2022a; Solla et al., 2020)

(Figure 2.9).

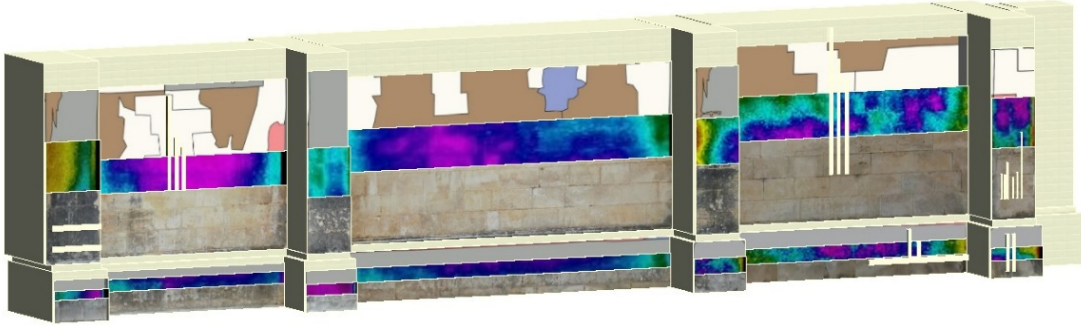


Figure 2.9: The integration of TLS, RGB and TIR within a historic building information model (H-BIM), enabling enhanced visualisation, data management and decision making (Solla et al., 2020).

TIR-Parametric Building Models

The generation of parametric building models, often originating from Scan-to-BIM data, provides a logical progression for 2D-3D fusion, with choices over building Level of Detail (LoD) and texturing strategy configurable to specific building types and purposes. Championed by the research of the Technical University Munich (TUM), Hoegner and Stilla (2007)’s TIR-TLS texturing work has provided the foundation for their future IRT-3DDF approaches, building image sequences into 2D polygonal texture masks applied through homography to 3D building models. Refinements of these methods have: (1) used accompanying GNSS coordinates for pre-orientation of TIR images (Hoegner & Stilla, 2018); (2) extracted line, edge and gradient features between TIR images and parametric models (Hoegner et al., 2016; Iwaszczuk & Stilla, 2017); (3) included the parametric model in the TIR orientation bundle adjustment to extract thermal textures (Hoegner & Stilla, 2015, 2016, 2018); and (4) projected textures from TIR-derived point clouds (Biswanath et al., 2023). The texturing of 3D parametric models can provide a simplified geometry for the visualisation of thermal information, modelled to the necessary LoD for the defined building model, and become a repository for building diagnostics and management (Patrucco et al., 2022b).

2.3.4 3D–3D Fusion

As demonstrated in Section 2.3.1, the ability to generate TIR-only 3D thermal models has become possible in the last ten years due to advances in SfM photogrammetric software, even with the low spatial resolution of TIR images. However, as proven by the research in this chapter, data fusion provides the benefits of a better geometry whilst maintaining the important temperature data for IRT-3DDF. Therefore, methods exploring 3D-3D fusion have looked to take preliminary 3D point clouds or 3D textured models generated solely from TIR images and fuse them with better geometries derived from additional sensors. 3D-3D fusion has looked to exploit novel SfM technologies, accompanying positional data and tandem sensors to provide more accurate methods of implementing IRT-3DDF.

TIR–VIS

Common across all TIR-VIS studies, whether it be 2D-2D, 2D-3D or 3D-3D, is the convenience of capturing both datasets at the same time knowing their relative orientation remains the same across all image pairs. However, the lack of TIR-VIS research for 3D-3D fusion (at present there are no explicit examples using both sensors within the same camera housing) points to several important considerations. Firstly, the VIS sensor found on most terrestrial and aerial thermal cameras tends to feature lower quality RGB sensors, meaning additional data sources are considered for the geometry as opposed to the convenience of the integrated sensor. Paziewska and Rzonca (2022)’s comparative study of UAV ‘base’ geometries for subsequent fusion demonstrates the benefit that integrating supplementary (non-VIS sensors) can have for accurate IRT-3DDF. Furthermore, as highlighted in Section 3.5, the acquisition of data at different times for different sensors can maximise their individual quality and mitigate their limitations. Although the use of TIR-VIS for 3D-3D fusion has been limited, opportunities could arise from the integrated use of

the sensors relative orientation in addition to other sensors or its use as a quality metric (Adamopoulos et al., 2020b).

TIR-RGB

The co-registration of 3D point clouds from independent sensors, as opposed to TIR-VIS, has demonstrated the benefits of capturing data from separate surveys and transferring thermal information onto accurate 3D geometries. Several methods have demonstrated the ease of TIR-RGB implementation, with the manual selection of common points a simple means of co-registering 3D point clouds. Both Scaioni et al. (2017) and Wakeford et al. (2019)’s marker-based research used Agisoft PhotoScan to manually mark common points in both TIR and RGB images to co-register 3D point clouds in the same reference system. Dahaghin et al. (2021) used GCPs extracted from Google Earth to undertake a Helmert transformation of UAV-acquired point clouds, using additional CPs to determine an absolute distance of < 0.3 m. The ease of implementation, using semi-automatic SfM processes for the co-registration of independent point clouds, has shown 3D-3D fusion to be a simple and scalable method for the generation of terrestrial and aerial 3D thermal models.



Figure 2.10: A 3D TIRpoint cloud, 3D RGB mesh model and resulting 3D textured thermal model using ICP for 3D-3D fusion (Maset et al., 2017).

Whilst the use of existing or extracted GCPs is a common procedure for IRT-3DDF, methods have demonstrated that 3D-3D fusion can be undertaken based on the

assumption that IRT and RGB point clouds are in close proximity, with Iterative Closest Point (ICP) able to align / optimize point clouds based on similar geometries (Figure 2.10)(Maset et al., 2017; Truong et al., 2017). Guilbert et al. (2020) and Grechi et al. (2021)’s 3D-3D fusion studies, monitoring coastal landslides and rock formations, respectively, used the Multiscale Model-to-Model Cloud Comparison (M3C2) algorithm (Lague et al., 2013) to refine point cloud co-registration, chosen for its robustness with varying point densities, local residuals and noise. Hoegner and Stilla (2016) compared various IRT-3DDF methods including 3D-3D fusion using ICP and GNSS/IMU data as initial parameters, finding superior results when compared to fusion in 2D image space. The benefits of ICP lie in the ability to register point clouds captured within the same acquisition (i.e., UAVs, TTISs), registering point clouds in close proximity and appropriately scaled.

However, it must be noted that most co-registration or ICP methods listed above perform successfully due to the initialisation information, either in the form of scaled point clouds or proximity prior to fusion. To address this, Lin et al. (2019b) used a four-step fusion method involving an initial coarse registration between unscaled TIR and RGB point clouds, using Fast Point Feature Histogram (FPFH)(Rusu et al., 2009) and Fast Global Registration (FGR)(Zhou et al., 2016) to determine co-registration. This process outperformed RANSAC (Fischler & Bolles, 1981) for the co-registration of 3D products and provided a coarse alignment prior to additional 2D-2D and 2D-3D fusion steps. To summarise, the ability for 3D products to be generated solely from TIR images has been proven to be a legitimate procedure within IRT-3DDF, one that can be exploited when initial position, scale and control information is provided. Although no TIR-TLS methods are currently defined, TIR-RGB fusion shows that when independent model generation is optimised through detailed survey and suitable control networks, the co-registration or transfer of thermal information can be easily implemented.

TIR-Parametric Models

As parametric models represent 3D geometries built to a specific LoD, largely from acquired TLS or Scan-to-BIM data, their registration with 3D thermal models presents a suitable method for texturing 3D products. Hoegner and Stilla (2018) present a 3D-3D fusion method coarsely registering 3D thermal point clouds with a building model, using recorded GNSS from a TTIS as a-priori information for the refinement of the TIR image poses. Similarly, Biswanath et al. (2023), texturing an LoD3 building model from a 3D TIR point cloud obtained through different minimisation approaches, used nearest neighbour to adhere ‘texels’ (texture pixels) to each building façade. Much like the research carried out by Hoegner and Stilla (2016) (from the same research group), Biswanath et al. (2023) found promising results over traditional 2D-3D texturing methods. These studies prove that the 3D geometry used for IRT-3DDF can be varied; which, whilst representing different LoDs, still provide a sufficient base for thermal modelling.

2.3.5 Three-dimensional Thermal Imaging Systems (TTIS)

Three-dimensional Thermal Imaging Systems (TTIS’), defined as platforms comprising multiple sensors on fixed or ‘dynamic’ units, have been a tandem development within IRT-3DDF, allowing for the capture of a wealth of datasets in both stationary and mobile forms (Figure 2.11). TTIS’ typically feature a thermal camera (often featuring both TIR and VIS sensors), a LiDAR range-based sensor and GNSS/IMU units, all fixed on a singular platform. Much like Alba et al. (2011)’s seminal ‘bi-camera’ study, the ability to fix sensors and determine their relative poses can enable the registration of all sensors in the same coordinate system and transfer data from one space to another. This system enables the capture of huge quantities of data from one single location, reducing the time for overall data capture and providing greater context through the combination of multiple datasets

corresponding to the same acquisition site (Adán et al., 2020). Furthermore, these systems can be transferred to terrestrial (Adán et al., 2023a) and aerial (Hoegner et al., 2016; Iwaszczuk & Stilla, 2017) mobile mapping platforms for greater access, repetition, mobility and scale (Hoegner & Stilla, 2018).

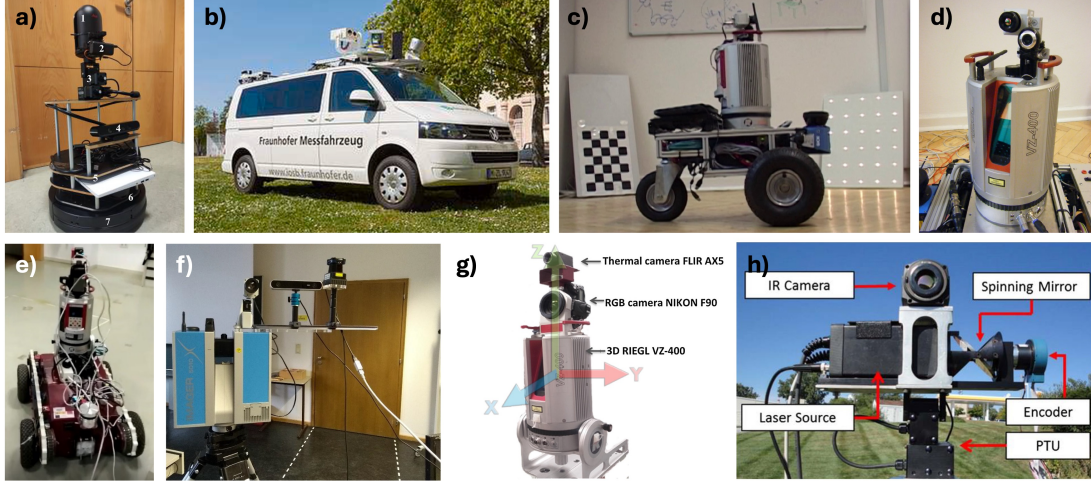


Figure 2.11: A collection of IRT-3DDF TTISs comprising different image-, range- and locational-based sensors: (a) (Adán et al., 2023a), (b) (Zhu et al., 2021), (c, d) (Borrmann et al., 2014), (e) (Adán et al., 2020), (f) (Hoegner et al., 2018), (g) (Merchán et al., 2018), (h) (Wang et al., 2013).

Evolving their own TTIS', termed MoPAD and MoPAD2 (Mobile Platform for Autonomous Digitization), Adán et al. (2020) and Adán et al. (2023b) demonstrate the effectiveness of configured systems for the generation of 3D *temporal* models for the AECO industry and energy monitoring campaigns within historic buildings. Similarly, work from Biswanath et al. (2023), Hoegner and Stilla (2018), and Zhu et al. (2021) use a mobile mapping system called MODISSA (Mobile Distributed Situation Awareness) comprising Jenoptik IRTCM 640 thermal camera and two Velodyne LiDAR for various IRT-3DDF methods, demonstrating the versatility of TTIS' and the methods they enable. Borrmann et al. (2014)'s thermal SLAM robot named Irma3D generated 3D thermal models of indoor scenes with a TLS, RGB and TIR cameras, resulting in the automatic detection of thermally-significant regions and volumetric measurements.

Although the benefits of TTIS' lie in the mobility of the platform, vast data capture

and completeness of surveys, some significant limitations exist. Such platforms typically have lower quality IR sensors; which, when coupled with the inability to navigate to areas of interest as easily as a hand-held camera, can lead to poorer TIR images. Furthermore, the processing of the data generated by a TTIS can be computationally expensive and does not provide the flexibility of capturing certain datasets when conditions are optimal. This, coupled with the significant price of TTIS' and the platforms to carry them (e.g., drones, vehicles or helicopters) means their adoption may be limited, especially in industries such as DCH. The merits of TTIS' lie in the ability to capture vast amounts of thermal information and the combination of sensors for optimal geometric and radiometric data (Figure 2.12).

2.3.6 IRT-3DDF & Complimentary Technologies

In addition to the plethora of approaches within IRT-3DDF, complimentary technologies have enabled data fusion methods to address specific applications, analyse obtained results and evolve the wealth of topics covered within IRT-3DDF. In a bid to extract value from the geometric and radiometric data contained within a thermal model, BIM has been a central tool for information management and data exploration (Hoegner & Stilla, 2018). Both Barbieri et al. (2023) and Brumana et al. (2018b) utilise historic BIM (H-BIM) for the analysis of IRT-3DDF results, exploiting active thermography for multi-temporal analysis and NURBs for cloister vaults element modelling, respectively. Although no specific tools for the integration of IRT (or indeed IRT-3DDF) exist within commercial BIM software (Adán et al., 2020), the growth of parallel energy- and building-related fields keep this an active research area (Cho et al., 2015; Ramón et al., 2022).

The integration of live environmental sensor data within historic buildings, captured through surface thermometers or hygrometers, remains an active research field, with building (Adán et al., 2020) and environmental (Grechi et al., 2021)

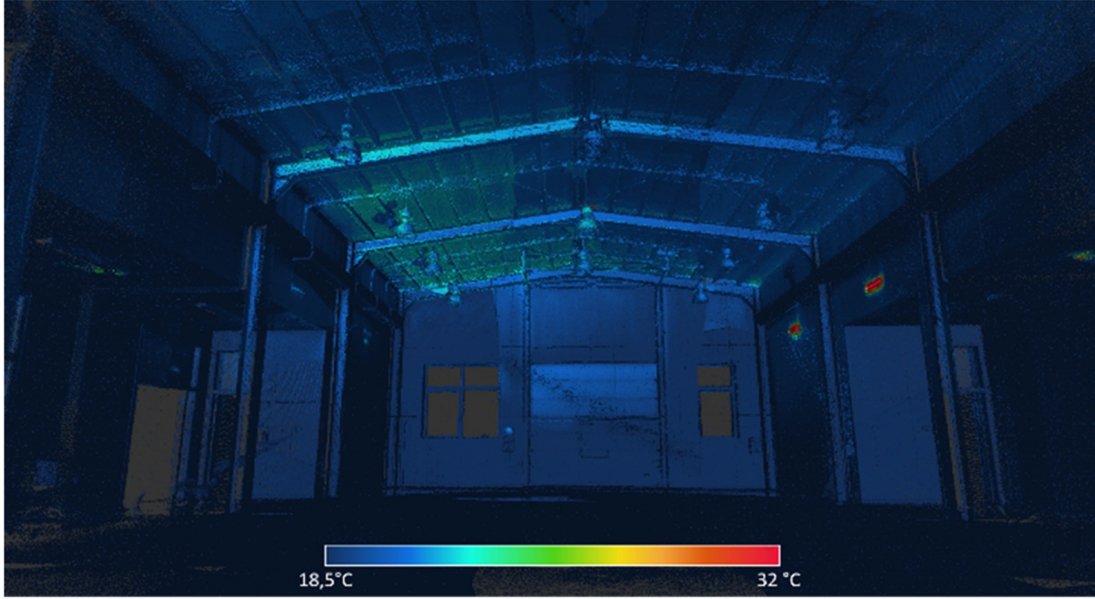


Figure 2.12: An example 3D point cloud generated by a TTIS, featuring a Riegl VZ400 TLS scanner, an Nikon D90 RGB camera and a FLIR AX5 TIR camera (Adán et al., 2020).

monitoring systems able to complement and correct IRT-3DDF products. Ham and Golparvar-Fard (2013b)’s Energy Performance Augmented Reality (EPAR) model combines their 2D-2D fusion method with computational fluid dynamics (CFD) to compare ‘actual’ and ‘supposed’ thermal models using EnergyPlus. Similarly, Adán et al. (2021)’s work compares, contrasts and combines IRT-3DDF with live environmental monitoring system sensor data, allowing for anomalies to be detected and deterioration to be monitored. Utilising distinct element modelling, Napolitano et al. (2019) combined TIR and TLS to map existing crack widths, structural loading patterns and settlement displacement of plastered masonry. The combination of sensors and models allows for the assessment of temperature change over time, local vs. global temperature discrepancies, and qualitative vs. quantitative assessments of temperature changes to be interpreted. Much like BIM, environmental sensor networks within IRT-3DDF provides additional information for CRM practices, assessing both current states and future possibilities.

The development of cameras and complimentary technologies has not been limited to thermal cameras, with several hardware advances enabling new methods of

IRT-3DDF to be explored and expanded. Most notably, the rise of UAV's within DCH has presented a critical tool in the arsenal of a heritage professional, allowing complete assessments of various subjects as well as architectural heritage. Wakeford et al. (2019) present the first 3D reconstruction of an active volcano, using 3D-3D fusion from a UAV to accurately monitor changes. Both Dahaghin et al. (2019) and Hou et al. (2022) assess building roofs using UAV's through similar fusion methods, with the latter testing the optimal conditions for thermal image capture including: camera altitudes, distinct camera angles, flight path design and building style. The use of UAV's has become an established tool within heritage sciences, and represents a crucial tool for the complete documentation and modelling of architectural heritage within IRT-3DDF.

Finally, both the expanse of platforms and cameras within IRT-3DDF have seen thermal modelling branch into tandem fields of computer vision using novel robots and sensors. Yang et al. (2018) used an iPhone SE and FLIR ONE (FLIR's attachable thermal camera unit)(320×240 pix) to demonstrate the capability of low-cost image-based modelling using affordable sensors. Rangel and Soldan (2014), Vidas et al. (2013), and Weinmann et al. (2014) used RGB-D cameras for 3D reconstruction, the latter using a four-step calibration and registration pipeline finishing with Efficient Perspective- η -Point (EP η P). Ramachandra et al. (2018) presents SLATM (Simultaneous Localisation and Thermal Mapping) for IRT-3DDF, showing promising results for fusion methods over using thermal cameras in isolation (Vidas & Sridharan, 2012). The studies presented here demonstrate IRT-3DDF may no longer be hamstrung by issues of correspondence and localisation, proving the capabilities of emerging tools for the extraction of value from IRT.

Chapter 3

Foundation of IRT-3DDF

Contents

3.1	Introduction	39
3.1.1	Foundation: Aims & Objectives	39
3.1.2	Chapter Structure	40
3.1.3	Scope, Assumptions & Limitations	41
3.2	InfraRed Thermography: A History	42
3.3	InfraRed Thermography: Principles & Practices	45
3.3.1	Infrared Radiation & the Electromagnetic Spectrum	46
3.3.2	Infrared Radiation & the Laws of Thermodynamics	48
3.3.3	Infrared Radiation & Temperature Measurement	52
3.4	Thermal InfraRed (TIR) Cameras	55
3.4.1	Workswell WIRIS Pro	57
3.4.2	Radiometric Calibration, Correction & Validation	58
3.5	Thermal Data Capture	65
3.5.1	Environmental Conditions	65
3.5.2	Photogrammetry: GSD, Image Overlap & Incident Angles	70
3.5.3	Thermography Approach	72
3.5.4	Survey Control & Thermal Targets	74
3.6	Conclusion	79

3.1 Introduction

Incorporating accurate temperature measurements within IRT-3DDF is fundamental for its effective use within CRM practices. The novelty of IRT-3DDF lies not solely in the ability to observe thermal phenomena in three dimensions, but to quantify these effects for further interpretation and analysis. Critically, the transformation of infrared radiation into temperature values is the by-product of internal (i.e., FPA sensor, lens optics, NUCs) and external (i.e., environmental conditions, material emissivity, subject distance) dynamics acting upon TIR cameras, features that need to be accounted for if IRT-3DDF is to be executed successfully. To date, considerations for radiometry within IRT-3DDF have been largely neglected, meaning methods citing high accuracy and precision of derived temperature values remain unsubstantiated (Adán et al., 2023b). Furthermore, to extract the inherent value of TIR images for IRT-3DDF, a balance between radiometric and photogrammetric principles is fundamental, with image overlap, incident angle and GSD having a significant effect on the overall accuracy of derived 2D images and 3D models. Consequently, for IRT-3DDF to be meaningfully undertaken, more than superficial modelling and visualisation, an understanding of electromagnetic radiation, TIR cameras and photogrammetry is paramount.

3.1.1 Foundation: Aims & Objectives

Therefore, addressing Objective 1 of the thesis, this chapter will define the foundational theoretical and practical knowledge for successful implementations of IRT-3DDF, achieved through a sequential workflow of procedures that address radiometric and geometric accuracy. Providing the first comprehensive assessment of protocols central to IRT-3DDF, this chapter will provide the necessary context for the execution of fusion methods in Chapters 4 and 5, identifying how geo-

metric and radiometric accuracy is ensured. As demonstrated by tandem fields that have advocated for application-specific procedures to ensure radiometric accuracy (Brooke, 2018; Dugdale, 2016), these procedures will be camera-, scene-, and subject-independent, allowing for procedures to be implemented across terrestrial and aerial IRT-3DDF use cases (Aragon et al., 2020; Wan et al., 2024). Whilst each procedure can be readily improved with greater access to calibration blackbodies, FPA sensor monitors and field radiometers; at its heart, this chapter provides the means to understand the nature of infrared radiation, the performance of TIR cameras and the considerations necessary for purposeful IRT-3DDF. This chapter will:

1. Outline the fundamental principles of infrared radiation, thermodynamics and temperature measurement in the context of IRT-3DDF
2. Develop procedures for assessing the performance of TIR cameras, including: the accuracy of derived temperature values, the required time for FPA sensor acclimatisation, and the reliability and periodicity of shutter-based non-uniformity corrections (NUCs).
3. Specify and justify required methodological decisions for IRT-3DDF, including photogrammetric, thermographic and subject-specific considerations.

3.1.2 Chapter Structure

This chapter begins with the history (Section 3.2) and theory (Section 3.3) of InfraRed Thermography, critical to the understanding IRT-3DDF and the notation used throughout subsequent chapters. Section 3.4 outlines the functioning of TIR cameras and presents a pseudo calibration set-up for assessing radiometric camera calibration, FPA sensor acclimatisation and shutter-based NUCs. Finally, Section 3.5 presents the critical procedures used for the execution of subsequent

IRT-3DDF methods, including: environmental conditions, photogrammetric survey control and radiometric validation. These are undertaken through the case study of Lenton Lodge Gatehouse (Appendix A.1), presenting the methodological justifications made in Chapters 4 and 5.

3.1.3 Scope, Assumptions & Limitations

For the successful execution of the aims and objectives, several assumptions must be acknowledged. Firstly, although the terms *infrared* (IR), *thermal* and *thermal infrared* (TIR) are used interchangeably throughout the thesis, all refer to the same bands of electromagnetic radiation denoted in Section 3.3.1. Secondly, the TIR camera utilised for this thesis, a Workswell WIRIS Pro, was employed as it was readily available for the execution of the thesis. Although the WWP provides integrated access to FPA sensor temperature, simultaneous image acquisition and sensor monitoring was not available. Therefore, sensor acclimatisation experiments were undertaken without accompanying imagery (and the effects this might elicit). Thirdly, no access to a calibration blackbody, climatic chamber or surface thermometer could be acquired, with calibration set-up controlled within practical limitations and reference materials measured using an infrared thermometer. Finally, no radiometers or surface thermometers were utilised for validation, although thermal-specific markers were employed in conjunction with an infrared thermometers. Critically, it is important to state what this chapter *can* and *cannot* achieve due to limitations in resources; however, it is these very limitations that will allow us to form methods that can be refined for future cameras and use cases of IRT-3DDF for architectural heritage.

3.2 InfraRed Thermography: A History

Although the discovery of infrared radiation is rightfully credited to William Herschel, it was Emilie du Chatelet's 'Dissertation sur la nature et la propagation du feu' (1744) that first outlined the possibility of different colours of light having different 'heating powers' (du Châtelet, 1744; Vollmer & Möllmann, 2018). It was not until 1800 that William Herschel, a royal astronomer to King George III of England responsible for the discovery of Uranus, discovered the varying heating properties of different colours in the 'thermometrical (electromagnetic) spectrum' (Lucchi, 2018). Herschel, looking to find suitable optical filters and lenses for telescopes, devised an experiment to determine the heat of separate colours exiting a prism (Figure 3.1). This demonstrated an increase in heat from violet to red light, but most notably that heat continued past the visible light to an area of invisible light Herschel termed 'calorific rays' (Herschel, 1800). Herschel's experiments determined the existence of what we now know as infrared radiation, confirming: (1) the difference and similarities between light and 'dark heat'; (2) the ability for this heat to be quantitatively measured using a thermometer; and (3) the transmission of such rays was different among materials (Barriera et al., 2012; Corsi, 2015; Maldague, 2001).

Several decades later, Herschel's discovery would be cemented by confirming the presence of infrared radiation and its links with temperature measurement. Thomas Seebeck (Seebeck, 1826) would discover the thermoelectric effect using an antimony-copper pair, demonstrating that a change in voltage could be obtained proportional that of changing temperature (ΔT), starting the development of temperature sensors with greater accuracy and sensitivity (Rogalski, 2012; Vollmer & Möllmann, 2018). In 1830, Leopoldo Nobili would employ the 'Seebeck Effect' to create the first thermocouple, an electrical thermometer capable of detecting infrared radiation (Nobili, 1830; Vollmer & Möllmann, 2018). In collaboration with Macedonio Melloni, Nobili-Melloni (1933) would develop the first 'thermophile infrared detec-

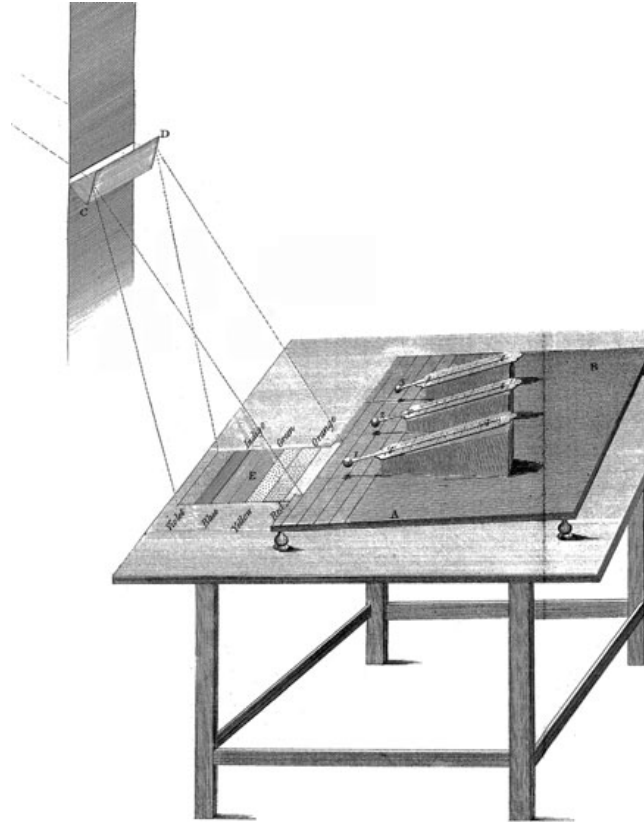


Figure 3.1: William Herschel's (1800) experiment discovering the presence of infrared radiation, utilising a prism and thermometers to determine the temperature of different electromagnetic regions (Herschel, 1800; Vollmer & Möllmann, 2018).

tor', combining ten thermocouples in a series circuit. This achieved a temperature sensitivity forty times greater than existing thermometers, capable of detecting the thermal signature of a person 10 meters away (Barr, 1962; Rogalski, 2012; Vollmer & Möllmann, 2018). Soon after, Edmond Becquerel (Becquerel, 1839) would discover the photovoltaic effect, John Herschel (Herschel, 1840) would produce the first thermograph and Samuel Pierpont Langley (Langley, 1880) would discover the bolometer. Importantly, the development of the bolometer, a material that changes its electrical conductivity when heat by infrared radiation, opened the door for discoveries and applications across greater distances (Vollmer & Möllmann, 2018).

Throughout the 20th Century, developments of military applications in WW1 and WW2 lead to the established use of photoconductive detectors to identify enemy positions, detect incoming vessels and see in the dark of night (Figure 3.2). During

this period, the development of pyroelectric detectors and cryogenic cooling expanded the range of materials capable of detecting infrared radiation, notably lead selenide (PbSe), lead telluride (PbTe), indium antimonide (InSb) or mercury cadmium telluride ($\text{Hg}_{1-x}\text{Cd}_x\text{Te}$) (Khodayar et al., 2016; Lawson et al., 1959; Vollmer & Möllmann, 2018). It was only in the early 1960s and 1970s that commercial infrared cameras would become available from the likes of Inframetrics and Agema Infrared Systems (now merged within industry-leading Teledyne FLIR®), powered by the introduction of charge-coupled device (CCD) sensors and the boom of commercial digital cameras (Boyle & Smith, 1970; Maldague, 2001; Vollmer & Möllmann, 2018). The introduction of the Focal-Plane Array (FPA) in the 1980s, a bidimensional sensor built upon the principles of the CCD, enabled increases in resolution for infrared detectors utilising both cooled and uncooled thermal cameras (Khodayar et al., 2016; Maldague, 2001).



Figure 3.2: Agema Infrared Systems’ AGA Thermovision 665 (1965-1967), one of the first commercial, mass production, single-element sensor, optomechanical scanner thermal cameras (Rogalski, 2012; Vollmer & Möllmann, 2018).

Today, the microbolometer has become the foundation for low-cost commercial in-

frared detectors, enabling thermal cameras to: (1) become smaller, light and more portable; (2) reduce pixel size and increase sensor resolution; (3) reduce manufacturer and consumer costs; (4) increase accuracy and sensitivity across different wavelengths and temperatures; and (5) provide greater operability and functionality for users within a wealth of industries (Lucchi, 2018). Thermal cameras are now central tools for: building and industrial inspection; agricultural, veterinary and medical science; military, surveillance and security; and gas, fire and fault detection (Gade & Moeslund, 2014). It is evident that the future developments of infrared technologies will continue to be driven largely by the military sector, still responsible for $\sim 80\%$ of the IRT market (Corsi, 2015; Maldague, 2001). This future will look to expand the applications of IRT (i.e., automotive, smart buildings and environmental control), develop sensors with increasing stand-alone and integrated capabilities (i.e., quantum detectors, silicon microbolometers, integrated sensors), and extract greater value from infrared thermography (Corsi, 2015; Khodayar et al., 2016).

3.3 InfraRed Thermography: Principles & Practices

To understand the processes central to IRT, an appreciation for the very radiation it is tasked with detecting, and the relationship with temperature it encapsulates, is necessary. Infrared Thermography, also termed *thermal imaging* or simply *thermography*, represents the scientific discipline dedicated to the detection and transformation of thermal infrared radiation (*thermos-*) into a signal capable of representing the distribution of temperature (*-graphy*) (Ibarra-Castanedo et al., 2013; Maldague, 2001; Vollmer & Möllmann, 2018). The science of IRT looks to define the relationship between the emitted radiation from an object and the resulting

radiation observed by a detector, accounting for the numerous influences that can impact the transmittance and attenuation of infrared radiation. Only by doing so can accurate temperature measurements be derived and displayed.

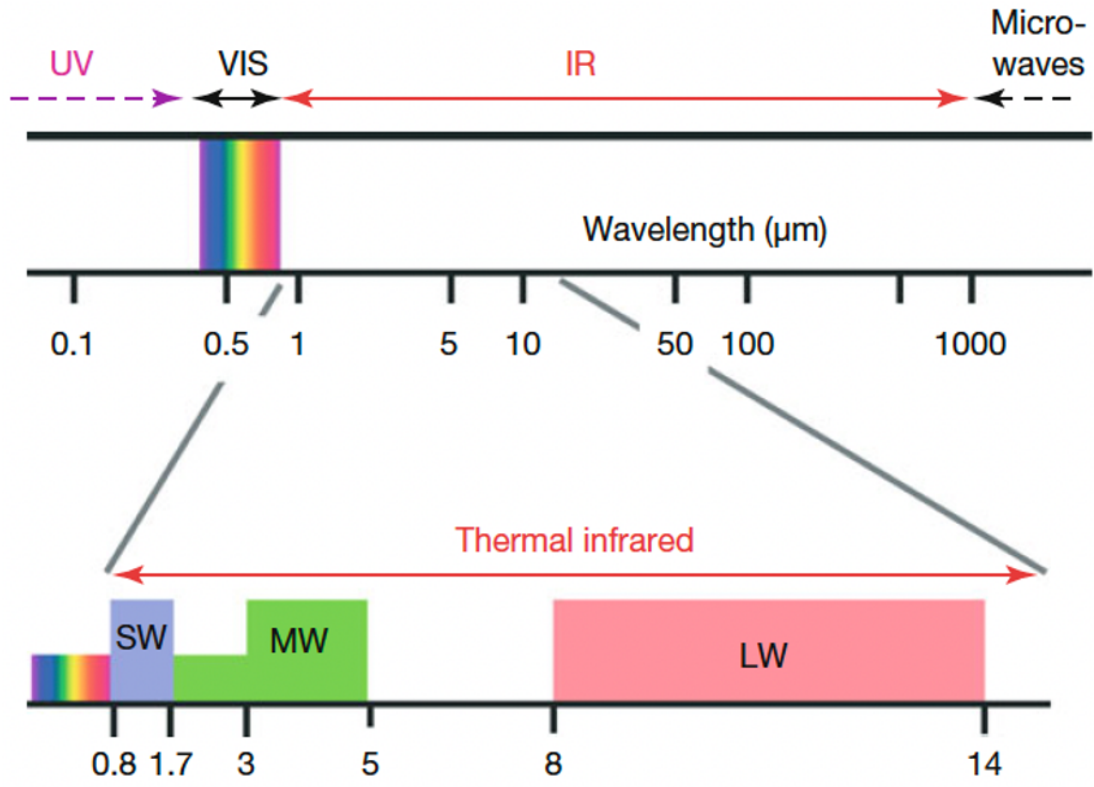


Figure 3.3: Electromagnetic spectrum visualising regions and bands for infrared radiation, and those denoted by the term 'thermal IR' (TIR) (Vollmer & Möllmann, 2018, p.10).

3.3.1 Infrared Radiation & the Electromagnetic Spectrum

Infrared radiation is radiation emitted as electromagnetic waves (or photons) by all objects with a temperature (T) above absolute zero ($T > 0 \text{ K}$ or $-273.15 \text{ }^\circ\text{C}$). The wavelength and intensity of the infrared radiation are determined by the temperature of the object, with higher temperatures resulting in a shorter wavelength of infrared radiation. Comprising part of the electromagnetic spectrum, sandwiched between the visible and microwave wavelength regions, infrared radiation spans wavelengths (λ) of $\sim 0.7\text{--}1000 \text{ }\mu\text{m}$ (Figure 3.3). Importantly, infrared radiation is

absorbed and emitted by both molecules (mass) and photons (mass-less), meaning the transportation of energy can take place with or without matter (Gade & Moeslund, 2014; Maldague, 2001). This is fundamental to understand how incident radiation is received by an infrared sensor in non-contact inspection processes, with infrared radiation propagating in straight lines much like visible light (Çengel et al., 2019; Vollmer & Möllmann, 2018).

Region	Wavelengths (μm)	Properties	Applications
NIR	0.7 – 1.4	Weak absorption by water High reflectance from vegetation	Vegetation health monitoring Night vision Fibre optics
SWIR*	1.4 – 3	Absorbed by water Penetrates haze	Water content analysis Geology Surveillance Food inspection
MWIR*	3 – 5	Detects hot objects Absorbed by gases	Thermal imaging Missile tracking Gas leak detection
–	5 – 7	Strong atmospheric absorptivity	Infrared Spectroscopy Gas Detection
LWIR*	7 – 14	High atmospheric transmittance	Thermal imaging Temperature measurement Night vision
FIR	14 – 1000	Strong interaction with molecules	Astronomy Biomedical imaging Climate studies

Table 3.1: Infrared spectral regions, their wavelength properties, and common applications. * denotes bands included as part of the *thermal* or *thermal infrared* designation (Vollmer & Möllmann, 2018, p.10).

Although defined regions and bands differ across (and within) scientific disciplines, the infrared portion of the electromagnetic spectrum can be broadly divided into five regions: Near-Infrared (NIR), Short-Wave InfraRed (SWIR), Medium-Wave InfraRed (MWIR), Long-Wave InfraRed (LWIR) and Far-Infrared (FIR). Table 3.1 shows the regions of infrared radiation, their associated bands within the electromagnetic spectrum and their applications. The importance of these bands lies in their unique abilities to absorb and transmit infrared radiation, with wavelengths in the 7-14 μm range, those most commonly utilised for thermal cameras, capable of transmitting the most radiation through the atmosphere and therefore best suited

for ambient temperature inspection ($T > 400$ °C)(Usamentiaga et al., 2014; Vollmer & Möllmann, 2018). In addition, unlike wavelengths in the 5-7 μm range, the absorption of infrared radiation by CO_2 or H_2O molecules in the atmosphere is lowest for LWIR at such temperatures, allowing for the maximum amount of radiation to be detectable for infrared sensors (Figure 3.4).

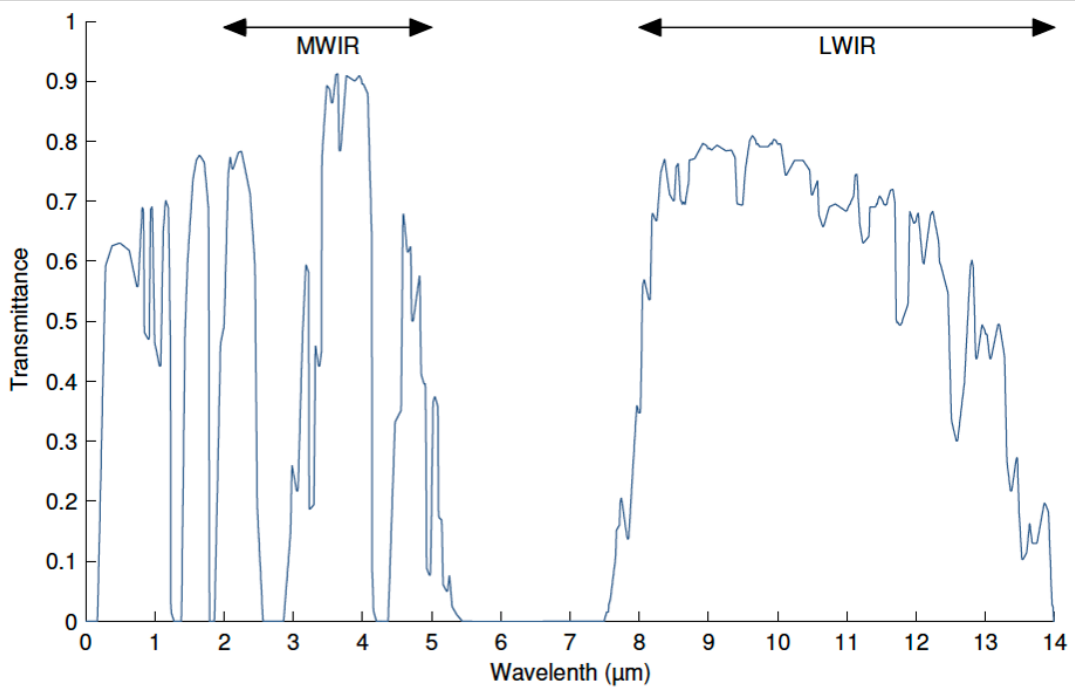


Figure 3.4: Atmospheric transmittance of MWIR and LWIR at one nautical mile, 15.5 °C and 70% relative humidity (Usamentiaga et al., 2014).

3.3.2 Infrared Radiation & the Laws of Thermodynamics

Whilst the discovery of infrared radiation by Herschel illuminated the relationship between the electromagnetic spectrum and temperature, confirmation of his principles would later be defined by several laws central to the theory of thermodynamics. As previously stated, every object above absolute zero will emit infrared radiation. However, the total proportion of energy either absorbed (α), reflected (ρ) or transmitted (τ) at a given wavelength, as defined by the law of energy conservation, must

total the incident radiation (Barriera et al., 2012):

$$\alpha(\lambda) + \tau(\lambda) + \rho(\lambda) = 1 \quad (3.1)$$

However, the proportion of energy absorbed by an object and subsequently emitted is defined as the emissivity (ε), a value that differs amongst materials due to their inherent properties. Emissivity, ranging from 0 to 1, expresses the ability for an object to emit infrared radiation (Lagüela et al., 2016). Encapsulated by Kirchhoff's Law (Kirchhoff, 1860), at thermal equilibrium, the proportion of incident radiation absorbed and subsequently emitted are equal, meaning the spectral absorptivity (α) and emissivity (ε) are the same at any given wavelength or temperature (Equation 3.2). Simply, this denotes that good 'absorbers' of temperature are also good 'radiators' of temperature and visa-versa (Corsi, 2015).

$$\alpha(\lambda) = \varepsilon(\lambda) \quad (3.2)$$

Therefore, by embracing this relationship, Equation 3.1 can be re-written as Equation 3.3.

$$\varepsilon(\lambda) + \tau(\lambda) + \rho(\lambda) = 1 \quad (3.3)$$

When all incident (absorbed) radiation is subsequently emitted (i.e., $\varepsilon = 1$), this object is known as a blackbody, a diffuse emitter that reflects and transmits no radiation regardless of wavelength, temperature or direction (Barriera et al., 2012; Usamentiaga et al., 2014). However, a blackbody represents an idealised object, with real objects emitting, reflecting and transmitting portions of the incident radiation (Vollmer & Möllmann, 2018). As demonstrated in Equation 3.1, when subject to incident radiation (ϕ_i), an object is defined by three parameters: absorptivity (α), reflectivity (ρ) and transmissivity (τ). These parameters specify the proportion of total energy either absorbed within, reflected off, or transmitted through an object; which, by virtue, also specifies the ratio of that incident radiation compared

to that of a perfect blackbody (Riveiro et al., 2016). These parameters, wavelength, temperature and orientation-dependent, determine the amount of energy leaving an object, ultimately dictating the amount of infrared radiation detectable by an infrared detector (Maldague, 2001). For opaque materials, transmissivity is discarded ($\tau=0$), with reflectivity becoming:

$$\rho(\lambda) = 1 - \varepsilon \quad (3.4)$$

Building the foundations for the relationship between spectral radiance and temperature, Planck's Law (Planck, 1901) defined the distribution of the temperature of a blackbody and the intensity of infrared radiation at different wavelengths (Gade & Moeslund, 2014), defined as :

$$I(\lambda, T) = \frac{2\pi hc^2}{\lambda^5 \left(e^{\frac{hc}{\lambda k_B T}} - 1 \right)}, \quad (3.5)$$

where $I(\lambda, T)$ represents the intensity of infrared radiation as a given wavelength and temperature, c the speed of light, h Planck's constant and k_b Boltzmann's constant. Equation 3.5 proves objects with higher temperatures emit most of their infrared radiation in lower spectral ranges (3-4 μm at $\sim 475^\circ\text{C}$) (Maldague, 2001).

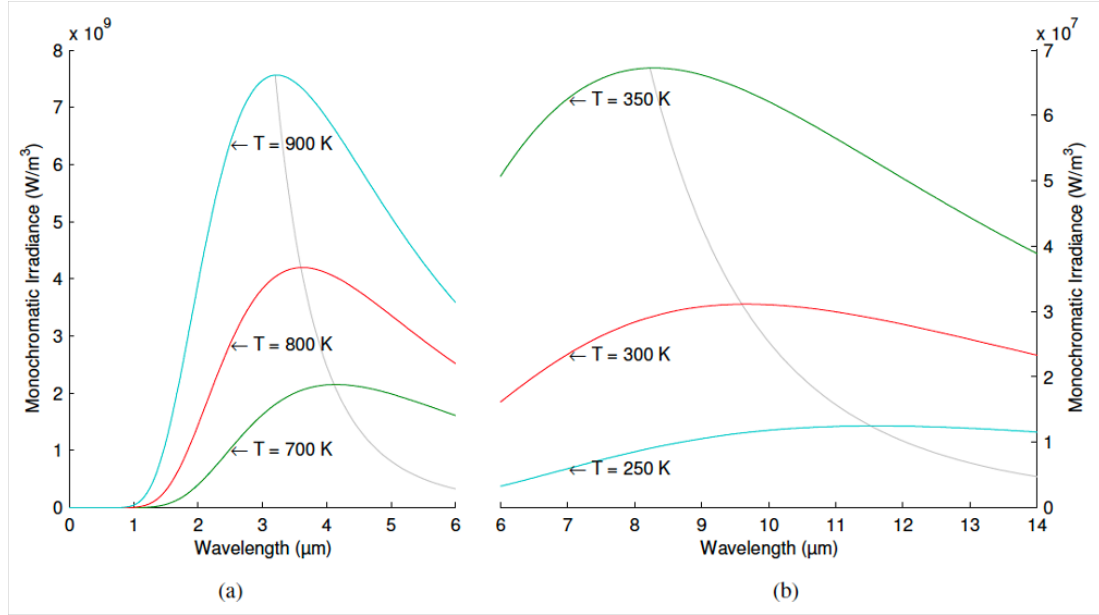


Figure 3.5: Planck's distribution curves implementing Wien's law of displacement, determining the maximum spectral radiance achieved by a blackbody at a given temperature (Usamentiaga et al., 2014).

Planck's Law of distribution, when combined with Wien's Law of displacement (Wien, 1897), which determined the derivative of Planck's distributions across all temperatures, allows for the maximum radiance at any given temperature and wavelength to be calculated (Figure 3.5). This stipulates that temperature and wavelength are inversely proportional, with maximum radiance being obtained at lower temperatures by higher wavelengths (Corsi, 2015). However, real objects are not idealised blackbodies or greybodies (consistent across wavelengths), with characteristics of the incident radiation (i.e., spectral intensity, angle, wavelength) and observed object (surface roughness, planarity, chemical composition) determining the amount of energy that is reflected and transmitted (Barriera et al., 2012). Therefore, the emissivity of a material can be denoted as a fraction of either the emitted radiation and the total incident radiation, or the proportion of emitted radiation against that of a blackbody at the same temperature (Equation 3.6). The latter, known as Stefan-Boltzmann's Law (Boltzmann, 1884; Stefan, 1879), defines

the emissivity of a grey body at a given wavelength:

$$\varepsilon(\lambda) = W_{rb} / W_{bb} \quad (3.6)$$

where W_{rb} represents the emitted radiation of a real (grey) body as a proportion of a blackbody (W_{bb}) (Garrido et al., 2018). Here, a blackbody is defined as:

$$W_{bb} = \sigma T^4 \quad (3.7)$$

with T as the absolute temperature and σ the Stefan-Boltzmann constant (Çengel et al., 2019). Therefore, to represent the total emitted radiation of an object, as opposed to a black body, the adaptation of Stefan-Boltzmann's equation for real objects integrating the emissivity of a grey body (assuming ε is constant across all wavelengths):

$$W_{rb} = \varepsilon \sigma T^4 \quad (3.8)$$

The implementation of these principles is central to the practice of IRT, informing the spectral ranges suitable for specific applications, the required characteristics of utilised infrared sensors, and the compensation needed for material emissivity.

3.3.3 Infrared Radiation & Temperature Measurement

As defined by the laws of thermodynamics, the accurate determination of temperature depends on accounting for the properties of, and between, the subject and the sensor detecting infrared radiation. Therefore, the identification and calculation of parameters acting within a scene are vital for the effective implementation of IRT. Figure 3.6 demonstrates the relationship between an observed object, the dynamics acting within the scene and the resulting radiation received by the infrared camera. Importantly, not all emitted radiation from an object is received by the infrared sensor, with the total observed radiation (W_{tot}) composed of three sources:

emission by the object (limestone)(E_{lime}), emission by proximal reflected radiation (E_{refl}), and emission by the atmosphere (E_{atm})(Garrido et al., 2020b; Maldague, 2001; Usamentiaga et al., 2014):

$$W_{tot} = E_{lime} + E_{refl} + E_{atm} \quad (3.9)$$

Firstly, the emission of a real object (E_{obj}), as seen in Equation 3.8, is dependent on its absolute temperature and emissivity. This represents the most critical value for IRT and must be calculated for accurate quantitative (as opposed to ‘apparent’) temperature measurements to be obtained (Lagüela et al., 2016). However, the total radiation emitted by the object is subject to the transmittance of the atmosphere (τ_{atm}), a product of the attenuation coefficient, distance and relative humidity of the scene (Aubrecht et al., 2016; Usamentiaga et al., 2014):

$$E_{lime} = \varepsilon_{lime} \cdot \tau_{atm} \cdot \sigma \cdot (T_{lime})^4 \quad (3.10)$$

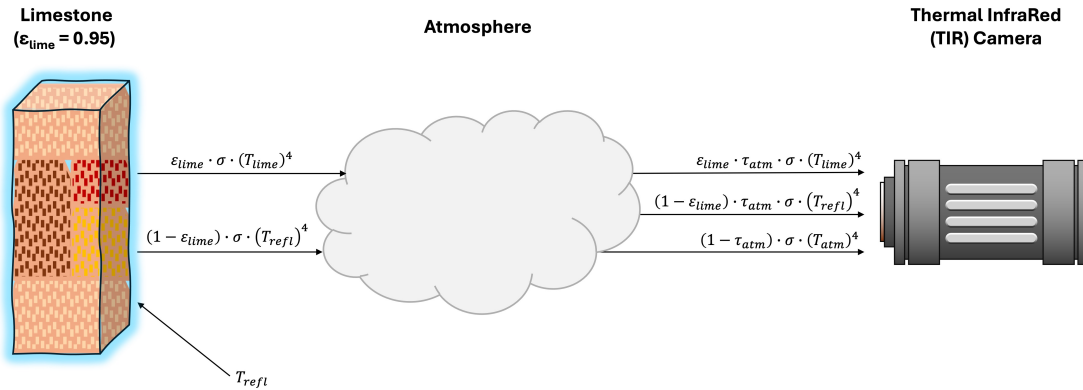


Figure 3.6: Radiation emitted in a scene and received by an infrared sensor.

Secondly, as demonstrated in Equation 3.4, energy not emitted or transmitted must be accounted for by reflectivity (E_{refl}), radiation received by the infrared sensor as a product of the temperature, atmospheric transmittance and reflectivity of proximal

surroundings (Aubrecht et al., 2016; Usamentiaga et al., 2014):

$$E_{refl} = (1 - \varepsilon_{lime}) \cdot \tau_{atm} \cdot \sigma \cdot (T_{refl})^4 \quad (3.11)$$

Lastly, although capable of absorbing radiation from the emitted object, the atmosphere between the object and infrared sensor is similarly capable of emitting the radiation it receives:

$$E_{atm} = (1 - \tau_{atm}) \cdot \sigma \cdot (T_{refl})^4 \quad (3.12)$$

As a result, by combining Equations 3.9, 3.10, 3.11, and 3.12, we are able to determine a unified formula for temperature measurement utilising the key parameters: (1) the emissivity of the object under investigation, (2) the reflected temperature in the scene, (3) the temperature of the atmosphere, and (4) the transmittance of the atmosphere:

$$T_{obj} = \sqrt[4]{\frac{W_{tot} - (1 - \varepsilon_{obj}) \cdot \tau_{atm} \cdot \sigma \cdot (T_{refl})^4 - (1 - \tau_{atm}) \cdot \sigma \cdot (T_{atm})^4}{\varepsilon_{obj} \cdot \tau_{atm} \cdot \sigma}} \quad (3.13)$$

By determining accurate values for these parameters, the calculation of accurate temperature measurements using IRT can be realised.

IRT-3DDF & Temperature Measurements

Whilst the principles of IRT remain constant for IRT-3DDF; critically, the *representation* of temperature takes several forms. By processing 2D TIR images as part of a 3D modelling pipeline, temperature values are observed, transformed and amalgamated throughout several processes, often combining multiple temperature values into a singular value. Figure 3.7 displays the distinctions between different component temperatures within IRT-3DDF, demonstrating the transformation of real-world temperature values into 2D/3D digital temperature values. To distinguish between these different temperature values within future thermal modelling

methods, the following terms are adopted for all subsequent IRT-3DDF workflows:

- T_{real} : The true (theoretical) temperature of a specimen.
- T_{obs} : The observed temperature of the specimen, achieved using a blackbody, infrared thermometer or similar measuring device. This value is still dependent on the observation, with parameters affecting the accuracy of the reading against T_{real} .
- T_{2D} : The recorded temperature within the 2D thermal image. This is used to represent either an individual pixel value or the image in its entirety.
- T_{3D} : The temperature encapsulated within the generated 3D thermal model. Similarly, this term is used to specify individual point or mesh values as well as the model in its entirety.

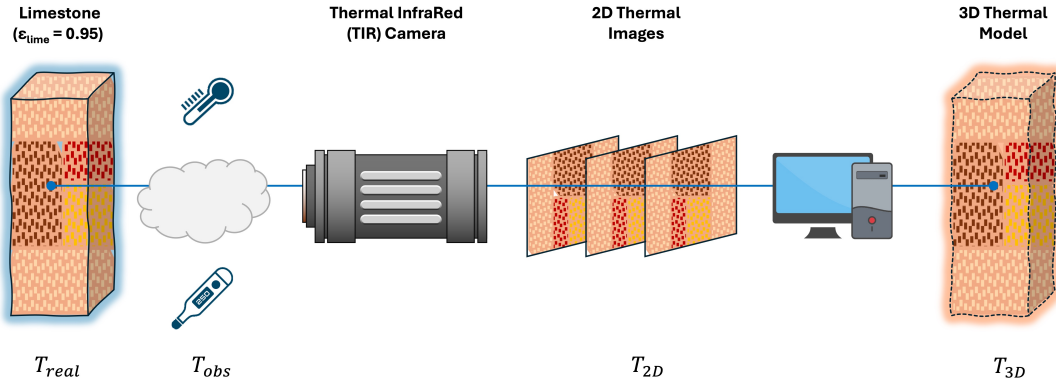


Figure 3.7: Distinctions between the different temperature values observed within a standard IRT-3DDF pipeline.

3.4 Thermal InfraRed (TIR) Cameras

The development of the microbolometer represents the foundation upon which the modern IRT industry has been built, accounting for more than 95% of current commercial thermal infrared cameras as of 2010 (Corsi, 2015; Rogalski, 2012; Vollmer &

Möllmann, 2018). Comprising individual bolometric detector elements (pixels) as a two-dimensional array (sensor), the resolution of a thermal camera is dependent on the overall size of an FPA sensor, the size of each individual microbolometer (e.g., $17 \times 17 \mu\text{m}$) and the fill factor between pixels (typically $>80\%$), all of which are improving drastically with newer camera models (Figure 3.8)(Vollmer & Möllmann, 2018). Typically ranging from 180×120 pix to 640×480 pix, sensor resolution has continually been denoted as the biggest barrier to successful IRT-3DDF, largely due to the difference in identifiable features *across* different modalities (Sutherland et al., 2023). Although data fusion methods have been undertaken with remarkably low resolution infrared sensors (notably Jones (2001)’s 8 mm 16×16 pix sensor and Javadnejad et al. (2020)’s 80×60 pix SenseFly Albris sensor), a larger sensor provides a greater GSD, preferential for effective IRT-3DDF (Adamopoulos et al., 2020b). Ultimately, the greater the sensor size the more likely successful fusion can be achieved, with trade-offs between camera price, thermal sensitivity and system cooling demands necessary considerations (Lagüela et al., 2016; Vollmer & Möllmann, 2018).

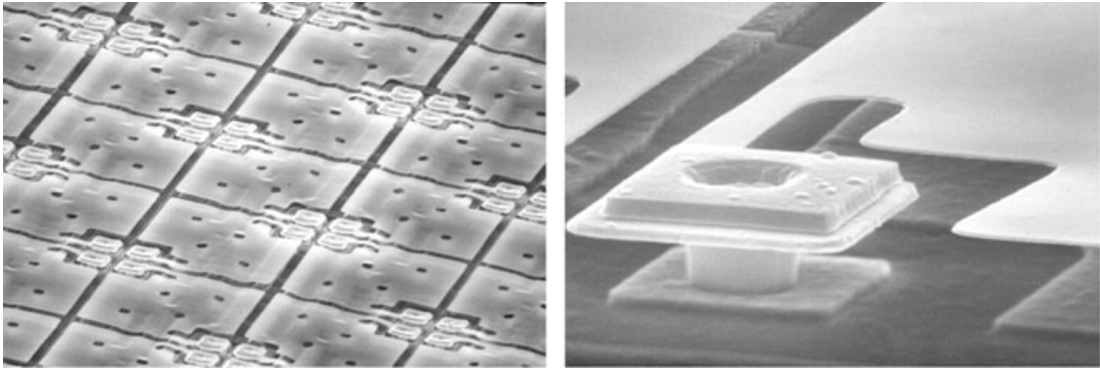


Figure 3.8: Uncooled VOx microbolometer FPA sensor, including microscopic image of: a microbolometer array (left) and microbolometer ROIC contact (right) (Vollmer & Möllmann, 2018).

3.4.1 Workswell WIRIS Pro

For the execution of this thesis, a Workswell® WIRIS Pro (WWP) UAV thermal imaging system was used, offering a vanadium oxide (VOx) microbolometer FPA sensor with 640×512 pix resolution and thermal sensitivity of $0.05\text{ }^{\circ}\text{C}$. The Workswell WIRIS Pro features both a thermal sensor and visible sensor (1920×1080 pix) fixed within the same camera housing, allowing for thermal-visible image pairs to be captured concurrently during each imaging survey. Although the Workswell WIRIS Pro features two sensors, WWP is used as the designation for the thermal sensor within the camera. To differentiate between visible spectrum images derived from independent cameras, as opposed to images coming from sensors housed within the Workswell WIRIS Pro; the term VIS is adopted for the visible spectrum sensor housed within the Workswell WIRIS Pro thermal camera (see Figure 2.3). Originally designed as a UAV camera, the Workswell WIRIS Pro has been modified to work terrestrially. This has included: the use of an external battery, the attachment of a digital monitor and keyboard, and the mounting of the camera on a traditional camera tripod. As suggested by Workswell s.r.o., yearly servicing of the Workswell WIRIS Pro is fundamental to maintain the accuracy of temperature measurements. The thermal camera was serviced in August 2023 before the execution of the data capture in Chapters 4 and 5, which included radiometric calibration, replacement of the camera's motherboard and lens cleaning.

The Workswell WIRIS Pro allows for a fully-customisable set-up that accounts for environmental conditions, observed temperatures and desired outputs. Firstly, the environmental conditions of the scene were set for every survey to achieve accurate temperature measurements, including material emissivity, reflected temperature, atmospheric temperature and distance (Section 3.3). Secondly, although the Workswell WIRIS Pro allows a wealth of file formats to be generated, 16-bit TIFF images encoding raw temperature values as a digital numbers (DN) were employed exclu-

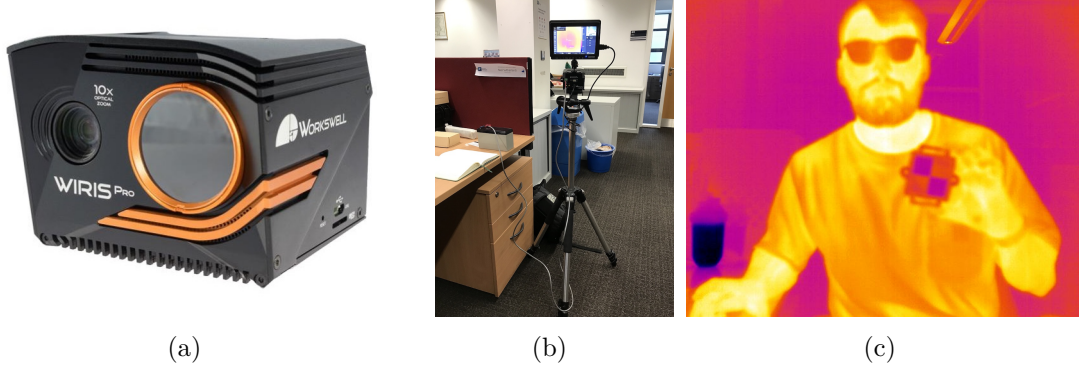


Figure 3.9: (a) Workswell WIRIS Pro featuring TIR and VIS sensors; (b) Workswell WIRIS Pro configured with external battery and monitor; and (c) a sample TIR image showing the thermal signature generated by a human subject, thermal-specific survey marker and water bottle in an ambient office building (Workswell s.r.o, 2021).

sively. Finally, to ensure radiometric consistency, the Workswell WIRIS Pro allows user's to access: the FPA sensor and CPU temperatures, NUCs and their periodicity, and the temperature range of the environment (for IR wavelength sensitivity configuration). Table 3.2 summarises the key properties of the Workswell WIRIS Pro, whilst Figure 3.9(b) shows the camera system after modifications and Figure 3.9(c) shows sample images obtained from the WWP.

3.4.2 Radiometric Calibration, Correction & Validation

At present, radiometric calibration, correction and validation remains an overlooked component of IRT-3DDF, resulting in methods that regularly ignore the radiometric accuracy of derived thermal models (Daffara et al., 2020; Merchán et al., 2018). Whilst access to a blackbody calibrator (Lin et al., 2018b), proximal FPA sensor temperature (Wan et al., 2021), and field radiometers (Kusnierek & Korsæth, 2014) is essential for radiometric calibration, correction and validation, respectively; the need to undertake investigations that can assess *relative* consistency and reliability of TIR cameras if *absolute* accuracy and precision is similarly valid (Adán et al., 2021; Kelly et al., 2019). Therefore, in order to evaluate the performance of the WWP, radiometric calibration, correction and validation experiments were under-

Specifications:	Values:
TIR Sensor	
FPA Sensor Size	10.88×8.71 cm
Pixel Size	17.00 μm
Resolution	640×512 pix
Nominal Focal Length	13.00 mm
Field-of-View (FOV)	45°×36°
Lens Optics (Material)	Germanium (Ge)
IR Detector	Uncooled VOx microbolometer
Spectral Range	7.5–13.5 μm
Temperature Sensitivity	0.05 °C
Temperature Accuracy	±2 % or ±2 °C
Temperature Ranges	–25 to +150 °C –40 to +550 °C
Frame Rate	30 Hz or <9 Hz
VIS Sensor	
CMOS Sensor Size	5.23×2.94 mm
Pixel Size	2.72 μm
Resolution	1920×1080 pix
Nominal Focal Length	3.50 mm
Power Supply, Weight, & Dimensions	
Input Voltage	9–36 VDC
Power	12 W
Weight	~430 grams
Dimensions	83×85×68 mm
Housing material	Aluminium
Operating temperature range	–15 to +50 °C

Table 3.2: Workswell® WIRIS Pro Specifications (Workswell s.r.o, 2020).

taken prior to IRT-3DDF to determine: (1) the time required for the FPA sensor to acclimatise; (2) the accuracy of derived temperature measurements; and (3) the stability and reliability of shutter-based non-uniformity corrections (NUCs).

FPA Sensor Acclimatisation

To stabilise an FPA sensor, allowing it to come to temperature after being turned on, a period of acclimatisation is required; with the duration, typically ranging from 15-60 mins, heavily camera- and scene-dependent (Berni et al., 2009; Döpper

et al., 2020; Kelly et al., 2019; Wan et al., 2021). Whilst FPA sensor temperature is not always accessible as a metric within commercial TIR cameras, assumptions of FPA sensor temperature can be made by monitoring the camera body temperature (Wan et al., 2021), the response of integrated shutter-based NUCs (Aubrecht et al., 2016; Kelly et al., 2019), or intra- and inter-image changes (e.g., fixed-pattern noise (FPN)(Hardie et al., 2000)). Although the Workswell WIRIS Pro provides proximal temperatures for both the FPA sensor and CPU, this is only available as supplementary information and cannot be extracted during image acquisition (Wan et al., 2021). As a result, values for both temperatures were manually plotted to assess sensor acclimatisation. Secondly, although the sensor temperature cannot be viewed during image acquisition, by undertaking regular shutter-based NUCs (which *are* computed utilising the sensor temperature) as part of a periodic image capture test (see Section 3.4.2), we can assess the extent of internal corrections which are being undertaken *based* on the FPA sensor temperature.

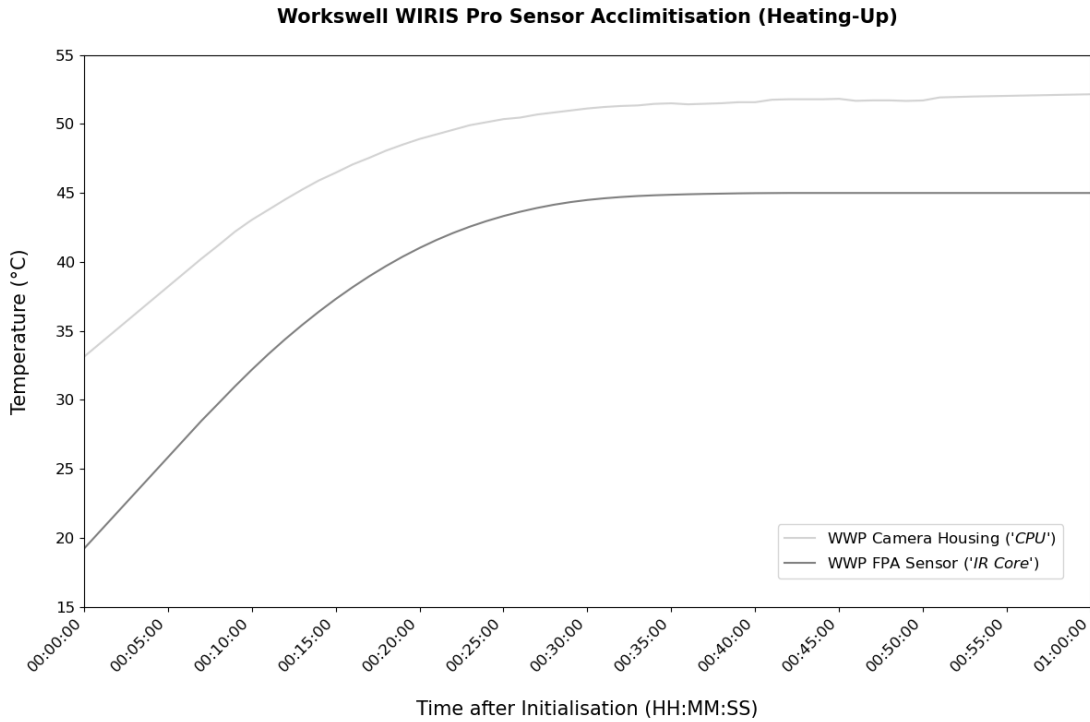


Figure 3.10: WWP Sensor Acclimatisation comprising internal FPA sensor ('IR Core') and camera housing ('CPU') temperatures over the 60 min duration.

Results of a 60-minute WWP sensor acclimatisation experiment can be seen in Figure 3.10, showing the listed internal temperature of the FPA sensor (termed ‘IR Core’) and camera housing (termed ‘CPU’) from initialisation (turning on the Workswell WIRIS Pro). In line with similar findings for UAV thermal cameras (and that of the WWP specifically)(Wan et al., 2021), the FPA sensor and camera housing reach a stable temperature around 35 mins after initialisation, determining the required time for the sensor to heat-up during subsequent IRT-3DDF methods. However, these results cannot be viewed in isolation, with several considerations needed to wholly understand: (1) whether meaningful assumptions on the FPA sensor can be made indirectly (as sensor monitoring and periodic image capture cannot be done concurrently), (2) how FPA sensor temperature is integrated into internal NUC processes (Wan et al., 2021), and (3) how environmental conditions affect the acclimatisation and stability of the FPA sensor for ‘real-world’ thermal imaging surveys)(Kelly et al., 2019; Kusnierek & Korsath, 2014).

Radiometric Camera Calibration

The introduction of the microbolometer has allowed TIR cameras to significantly reduce overall camera unit size (Olbrycht & Więcek, 2015); however, this has come at the cost of internal cooling systems and closer processor-based camera electronics, all of which generate infrared radiation and alter the internal temperature of the camera during operation (Acorsi et al., 2020; Budzier & Gerlach, 2015). Existing radiometric calibration approaches have looked to address the dynamic response of FPA sensors by: (1) varying blackbody temperature to simulate changing environmental conditions (Aragon et al., 2020; Lin et al., 2017; Nugent et al., 2013; Tempelhahn et al., 2016); (2) comparing linear and exponential models to map temperature variation over time (Budzier & Gerlach, 2015; Cao & Tisse, 2013; Gonzales et al., 2019; Nugent et al., 2014); and (3) contrasting calibration results derived in controlled lab- or field-based experiments (Gazzano et al., 2024; Kelly

et al., 2019; Kusnierek & Korsæth, 2014). To undertake radiometric calibration, correction and validation experiments, a pseudo blackbody-like calibration set-up was configured, observing a uniform specimen (TESTO emission tape: $\varepsilon=0.95$) in conditions with constraints on: atmospheric temperature and relative humidity; lighting, reflections and irradiance; camera distance, motion and incidence angle (Aragon et al., 2020; Wan et al., 2021). Whilst temperature accuracy cannot be determined to the same degree as a traditional blackbody, an infrared thermometer (RS PRO RS-8876: temp. accuracy ± 1 °C, temp. res. 0.1 °C) was used to assess temperature accuracy of T_{2D} against T_{obs} over time.

Shutter-based Non-Uniformity Corrections (NUCs)

The susceptibility for FPA sensors to increase in temperature over time has led manufacturers to develop processes that correct the differences in gain and offset values of each microbolometer during operation. To mitigate this phenomenon, many modern commercial cameras feature a non-uniformity correction, commonly termed NUCs, where individual microbolometer values are corrected against a mechanical shutter with homogenous temperature (Döpfer et al., 2020; Nugent et al., 2014). Although the implementation of shutter-based NUCs varies amongst thermal cameras, with some researchers configuring their own shutters (inc., filters (Olbrycht et al., 2012), semi-transparent- (Olbrycht & Więcek, 2015) or external shutters (Virtue et al., 2021)), most are in-built, inaccessible (w.r.t., manufacturer’s implementation, generated shutter images or shutter temperature) and provide basic configuration (e.g., duration between NUCs (periodicity) and manual/automatic implementation)(Figure 3.11).

To assess the reliability and periodicity of NUCs for the WWP, the radiometric calibration set-up was used to determine the ability for NUCs to return drifted T_{2D} values back to known T_{obs} values. Periodic image capture was undertaken using

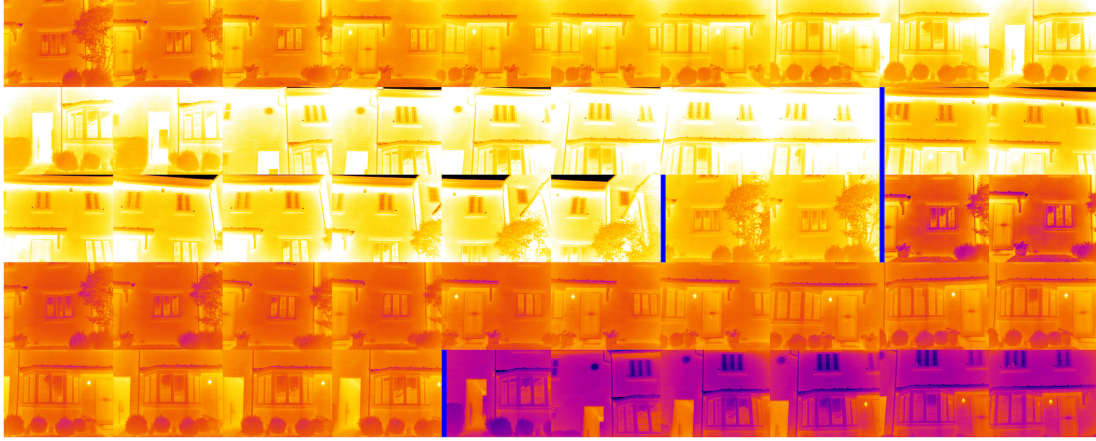


Figure 3.11: An example 50 image TIR sequence generated by the WWP beginning immediately after initialisation with automatic NUCs enabled. Here, several NUCs (represented by vertical blue lines) are undertaken at irregular intervals to compensate for the effects of sensor acclimatisation, visible in the ‘brightening’ of images leading up to a NUC event.

the WWP in a 90-minute window, with images taken every 15 seconds. NUCs were implemented automatically every 10 mins starting at 2 mins (after the WWP’s internal monitors deem the camera to have finished ‘heating up’). The following intra-image metrics, those calculated from a singular image of the TIR sequence, were plotted over the duration of the test: intra-image mean (II_{mean}), intra-image minimum (II_{min}) and intra-image maximum (II_{max}) (Yuan & Hua, 2022). To determine the reliability of NUCs, recorded IR thermometer measurements (T_{obs}) were taken periodically throughout the experiments, allowing for the assessment of the radiometric calibration of the WWP and the ability for NUCs to return drifted values back within range of a known value.

The results of the radiometric camera calibration and NUC experiments can be seen in Figure 3.12, visualising the NUC intervals throughout periodic image capture. Notably, we can see the importance of the NUCs upon initialisation, responding to the acclimatisation of the FPA sensor. Although Wan et al. (2021) determined the uncertainty of NUCs to be considerable for the WWP under dynamic environmental conditions, it is clear that under controlled conditions the NUCs are successful at

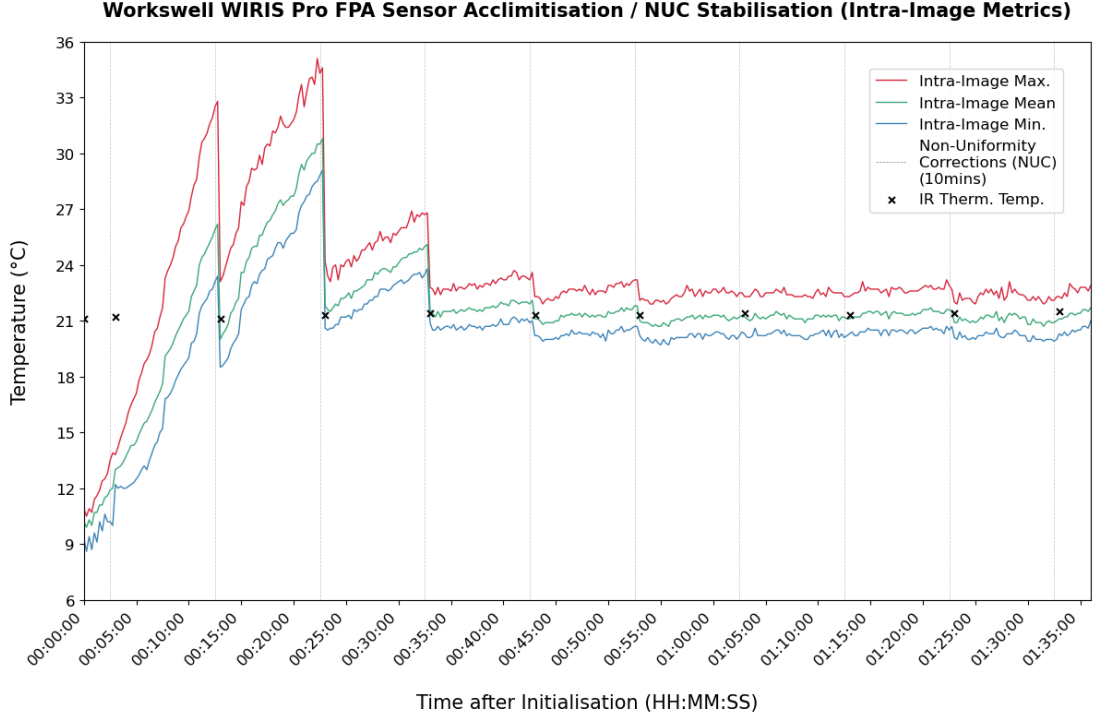


Figure 3.12: WWP FPA sensor acclimatisation / NUC stabilisation results using the psuedo-blackbody radiometric calibration set-up. Here, intra-image metrics (Max., Mean and Min.) are shown over the course of the 90-minute experiment generated during periodic image capture.

returning T_{2D} values back to T_{obs} upon each NUC iteration, with a mean absolute difference ($\mu_{|\Delta|}$) of $0.4\text{ }^{\circ}\text{C}$ ($\sigma=0.2\text{ }^{\circ}\text{C}$) for NUCs after 15 mins. Importantly, in line with the results of the Section 3.4.2 (~ 35 mins) and those of Wan et al. (2021) (60 mins), under periodic image capture the WWP does not provide stable temperature readings before ~ 52 mins, suggesting that periodic image capture increases the time required for the FPA sensor to stabilise and provide accurate temperature measurements (surprisingly, the WWP CPU suggested the heating up period had finished after 2 mins, after which the NUCs began). In summary, the radiometric calibration, correction and validation experiments suggest that, during operation, we can rely on the WWP’s factory *calibration* for accurate T_{2D} values; however, we must not rely on factory *correction*, meaning an initial period for FPA sensor acclimatisation and the implementation of regular NUCs fundamental for consistent TIR surveys (Wan et al., 2024).

3.5 Thermal Data Capture

For IRT-3DDF to be successful, a balancing act between IRT and photogrammetric practices must be undertaken. Whilst the stability of TIR cameras, impact of changing environmental conditions and choice of thermography approach are common challenges within IRT; when coupled with the necessity for high image overlap, simultaneous multi-modal data capture and geometric survey control, the process of IRT-3DDF becomes increasingly complex. Although standards, guidance and publications exist for IRT (ASTM International, 2015; British Standards Institute, 2017; Colyer & Arnolds, 2019; Garrido et al., 2020a) and close-range photogrammetry (Historic England, 2017, 2024; Stylianidis & Remondino, 2016; Waldhäusl & Ogleby, 1994) (often adopted within IRT-3DDF research (Napolitano et al., 2019; Patrucco et al., 2022a)), best practices for the capture of TIR images for 3D modelling are yet to be clearly defined. The complexity of thermal data capture for IRT-3DDF comes from the fact that choices to optimise one variable may hinder the performance of another, meaning decisions must be made acknowledging the purpose of the survey and desired outputs. Accompanied by experiments for multi-modal data captured at Lenton Lodge Gatehouse (Chapters 4 and 5), this section determines where considerations for thermal data capture are needed to address the interconnectivity of IRT and 3DDF practices, subsequently summarised as an IRT-3DDF ‘cheatsheet’ available in Appendix B.

3.5.1 Environmental Conditions

The ability to mitigate the effects of dynamic environmental conditions acting upon TIR cameras is critical for consistent temperature measurements within multi-image and multi-sequence surveys (Mesas-Carrascosa et al., 2018; Vollmer & Möllmann, 2018). As previously stated in Section 3.3, several sources of interference need to be

calculated for accurate temperature measurements; which, when left unaccounted, can lead to several degrees of error (Adán et al., 2023b; Kelly et al., 2019; Wan et al., 2021). Firstly, the most important source of external radiation that must be considered is the sun, which significantly affects the collection of thermal images. Solar radiation provides not only a source of radiant flux that will change the thermal equilibrium of the subject, but result in thermal shock to the TIR camera, generating inaccurate and unstable TIR image sequences (Elfarkh et al., 2023; Vollmer & Möllmann, 2018). Similarly, prior weather conditions (i.e., rain, snow, humidity, surface moisture) will impact a building and may not be representative of its normal conditions.

Critically for UAV research, the influence of ambient air temperature, wind speed, atmospheric attenuation and flight duration on accurate IR radiation measurements requires careful determination (Kusnieriek & Korsæth, 2014; Malbêteau et al., 2021; Wan et al., 2024). Kelly et al. (2019) and Wan et al. (2021) demonstrate that in controlled laboratory conditions, increases in ambient temperature and wind speed resulted in pronounced shifts from true temperature values obtained from different aerial and terrestrial TIR cameras. In addition, an ‘attenuation coefficient’ representing the absorptivity of molecules within the atmosphere (i.e., water vapour) requires inclusion for reliable temperature measurements (Acorsi et al., 2020; Dugdale et al., 2019). Conceptualised in the Beer-Lambert law (Lambert, 1760), this is commonly achieved through an indirect calculation utilising the distance from the subject, ambient temperature and relative humidity (Lagüela et al., 2016). With most UAV cameras capable of introducing these parameters pre- and post- TIR image capture, the need for sensors capable of collecting known values for distance, air temperature and relative humidity are vital for effective IRT-3DDF (Aubrecht et al., 2016).

To mitigate environmental conditions, procedures were undertaken at Lenton Lodge Gatehouse before, during and after the TIR data capture. Firstly, to determine an

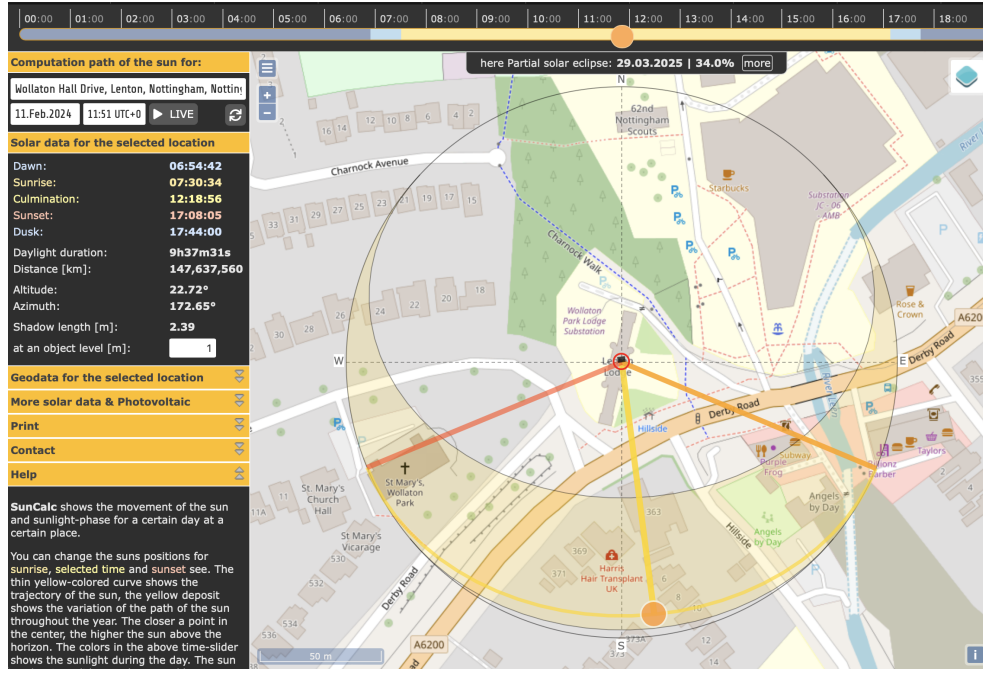


Figure 3.13: SunCalc online sun-phase tool, showing environmental conditions (sun's position, direction, azimuth and shadows) at Lenton Lodge Gatehouse on 11th February 2024 (Hoffman, 2024).

optimal time for TIR image capture, SunCalc, an online sunlight-phase tracker, was used to assess solar radiation in relation to the building's orientation, including: sunrise, culmination, sunset and shadows (Hoffman, 2024). For Lenton Lodge Gatehouse, facing East-South-East, data acquisition between 09:30-12:00 provided optimal conditions to maximise morning 'baking' of the structure and to minimise shadows on the façade (Figure 3.13). Secondly, weather was checked 24-, 4- and 1-hour before each survey to determine any previous weather conditions that might affect the IR radiation being detected. Thirdly, an anemometer (Protmex PT6252A) was used upon arrival at the site to determine accurate wind speed measurements for the site, including wind direction. Lastly, ambient air temperature was recorded using a hygrometer (RS PRO RS-91), with the air temperature and relative humidity used as initial input for the WWP (Figure 3.14) (Mesas-Carrascosa et al., 2018).



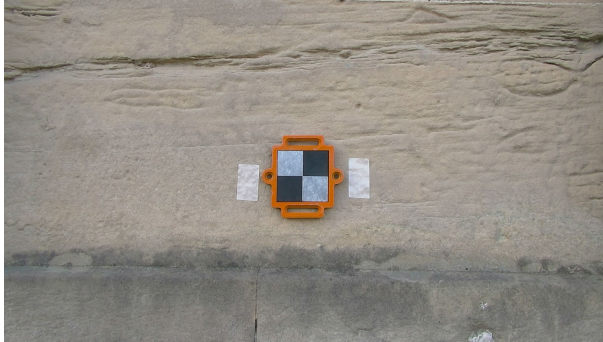
Figure 3.14: (a) Protmex PT6252A anemometer, (b) RS PRO RS-91 hygrometer, (c) RS PRO RS-8876 thermometer and (d) TESTO Emission Tape.

Emissivity Determination

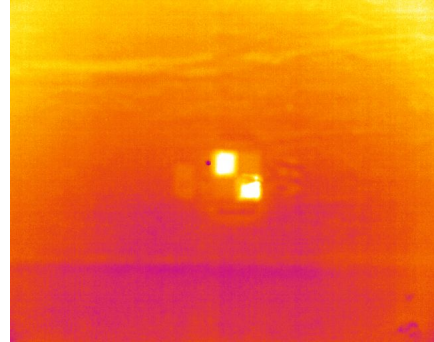
Emissivity represents a critical value for IRT, defined as the single-biggest influence on accurate temperature measurements (Aubrecht et al., 2016). Therefore, the determination of accurate emissivity values (ε) for materials within a scene are fundamental if quantitative temperature measurements are required for IRT-3DDF. Both Adán et al. (2021) and Previtali et al. (2012) included emissivity values within their IRT-3DDF pipelines, the latter undertaking laboratory experiments to determine the emissivity of both angular and linear tiles. On the other hand, Costanzo et al. (2014), looking for qualitative temperature differences of thermal anomalies, set a mean emissivity value based on suspected material values. However, as observed by Fryskowska-Skibniewska et al. (2022), the emissivity of an object is dependent on a range of factors, including the observation angle, temperature, wavelength, and the material's inherent properties (i.e., surface roughness, oxidation); meaning a-priori emissivity values are not sufficient and should be confirmed in-situ (Lagüela et al., 2016).

To determine emissivity values (ε) for the predominant building materials at Lenton Lodge Gatehouse, the ASTM standard 'Practice for Measuring and Compensating

for Emissivity Using Infrared Imaging Radiometers' (ASTM E1933-14) was used (ASTM International, 2022). Modifying ASTM E1933-14's non-contact thermometer method (NTM), TIR images of the desired specimen and a material with a known emissivity, termed a surface-modifying material (such as TESTO SE & Co. emission tape, where $\varepsilon=0.95$) are acquired (Figure 3.14)(Mineo & Pappalardo, 2021). Subsequently, ThermoLab, Workswell's freely available post-processing software, was used to retroactively alter emissivity values until the temperature of the desired specimens (sandstone, wood and metal) matched that of the values noted for the surface-modifying tape. The emissivity value for the predominant material at Lenton Lodge Gatehouse was then set within the WWP. Results of the emissivity calculations at Lenton Lodge Gatehouse can be seen in Table 3.3, with sample TIR and VIS images showing the adherence of the TESTO emission tape seen in Figure 3.15.



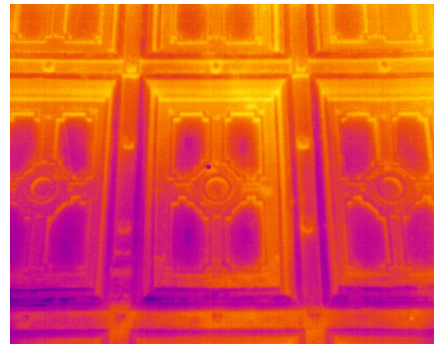
(a) VIS



(b) TIR



(c) VIS



(d) TIR

Figure 3.15: Sample (a,c) visible and (b,d) thermal images showing the adhered TESTO emission tape ($\varepsilon = 0.95$) for emissivity calculation of (a,b) Mansfield Sandstone and thermal targets, and for (c,d) wood and iron secondary materials.

Material	ε	σ	$\Delta\varepsilon$
TESTO emission tape	0.95	–	0.25
Mansfield Sandstone	0.70	0.03	–
Wood	0.90	0.14	0.20
Cast-Iron	0.83	0.08	0.15
Marker (Rubber)	0.91	0.03	0.21
Marker (Aluminium)	0.37	0.04	0.53

Table 3.3: Calculation of material emissivity (ε), standard deviation (σ) and difference in emissivity ($\Delta\varepsilon$) between the predominant material (Mansfield Sandstone) and secondary materials at Lenton Lodge Gatehouse.

3.5.2 Photogrammetry: GSD, Image Overlap & Incident Angles

The fundamental practices for effective SfM (Luhmann et al., 2006; Stylianidis & Remondino, 2016) are not reserved solely for visible spectrum imagery, with considerations for ground sampling distance (GSD), image overlap and incidence angles critical for IRT-3DDF. Firstly, the GSD of TIR images, the real-world distance represented by an individual pixel within an image (e.g., mm/pix), plays a central role in the accuracy of orientation (Dlesk et al., 2021), the extraction of reliable features (Döpper et al., 2020; Lewińska & Maciuk, 2020), the deviation of temperature values (Wan et al., 2021), and degree of atmospheric attenuation (Dugdale et al., 2019) (Figure 3.16). By understanding the required accuracy of derived 3D models, appropriate GSDs (and by extension, capture distances) can be determined for each modality within IRT-3DDF.

Secondly, image overlap and incidence angles represent critical values for IRT-3DDF, increasing the opportunities for feature matching (Lin et al., 2019b) and effective thermal texturing (Kelly et al., 2019). Whilst largely case dependent, research has suggested horizontal and vertical image overlap of 50% (Lagüela et al., 2011b) – 90% (Patrucco et al., 2022b) for IRT-3DDF. Adamopoulos et al. (2021) shows 90% horizontal overlap and 70% vertical overlap of imaging strips perpen-

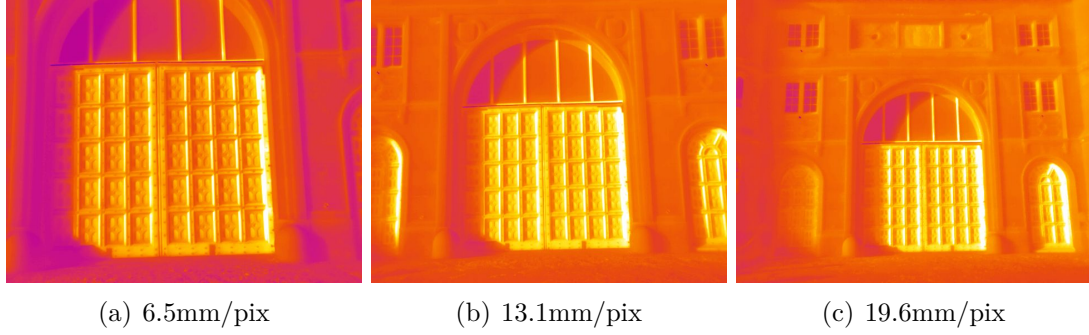


Figure 3.16: A comparison of TIR images at (a) 5m, (b) 10m and (c) 15m from Lenton Lodge Gatehouse, including their GSDs (mm/pix) for the Workswell WIRIS Pro.

dicular to the building facade is desirable for TIR image sequences, with additional oblique images beneficial for successful orientation. Dabetwar et al. (2022), comparing image overlap and orientation angles for UAV thermal modelling determined: (1) a high accuracy against a reference cloud when 90% horizontal overlap is used compared to 80% and 70%, (2) greater tie point multiplicity (ratio of project 2D image points to 3D tie points) across different incidence angles, and (3) that overlap plays a more significant role than incidence angle for model accuracy. In similar vein, Hou et al. (2022) finds a 30° incidence angle, the angle between the normal ray of the camera and the normal ray of the object, superior to 45° not only for reconstruction, but critical for accurate temperature projection (Vollmer & Möllmann, 2018, p. 35-39).

Finally, the number of TIR images captured, the integration of additional surveys and the time taken for each to be undertaken will have substantial effects on the results for IRT-3DDF. To meet the requirements of the scene, acknowledging the need for sufficient overlap, IRT-3DDF research shows the collection of hundreds of TIR images to be a common practice (e.g., 581 for a two-storey building (10 mm/pix GSD)(Dlesk & Vach, 2019), 568 for a complex parabolic arch (36 mm/pix avg. GSD)(Patrucco et al., 2022b), 308 for acute-angle imaging of restricted façades (15-53 mm/pix GSD)(Patrucco et al., 2022a)). However, this comes with a likely

increase in sensor temperature and the introduction of radiometric errors for longer acquisitions (Mesas-Carrascosa et al., 2018). In summary, all photogrammetric practices optimising geometric accuracy must be viewed in relation to the IRT practices ensuring radiometric consistency, with the former having considerable effect on the latter.

For Lenton Lodge Gatehouse, the following practices were implemented to maximise the use of TIR images geometrically and radiometrically. TIR images were captured from known distances for the comparison of different GSDs, with each acquisition strip being entered into the Workswell WIRIS Pro prior to image capture. Additionally, 90% horizontal and 70% vertical overlap were prioritised for TIR image sequences, with Equation 3.14 determining the required spacing between consecutive images:

$$z = 2x \tan\left(\frac{v}{2}\right)\left(\frac{1-y}{100}\right) \quad (3.14)$$

where: v is the horizontal FOV of the camera ($^{\circ}$); y the desired overlap of sequential images (%); x is the distance away from target (m); and z is the resulting distance between concurrent camera positions (m). Furthermore, images were taken perpendicular to the façade to minimise temperature discrepancies and provide suitable images for texturing. The Workswell WIRIS Pro, configured on a tripod and set to take images automatically every 12 seconds, allowed for the reduction of movement from the user.

3.5.3 Thermography Approach

The exploitation of different thermography approaches, the use of passive or active thermography during image capture, provides an additional variable for the analysis of architectural heritage, enabling IRT-3DDF to observe changes in historic buildings and gain additional insights for CRM practices. Passive thermography is

defined as the measurement of radiation without any artificial heat stimulus, capturing the natural thermal behaviour of a structure (Usamentiaga et al., 2014). The influence of solar radiation, considered a source of passive thermography due to its variability (Lagüela et al., 2016), has allowed passive IRT-3DDF research to use the sun as a means of generating thermal contrast either by: investigating historic buildings during sunrise or sunset (Patrucco et al., 2022a); undertaking assessments on cloudy days to minimise shadows (Barbieri et al., 2023; Lagüela et al., 2013); or repeated surveys seasonally to determine change under different solar conditions (Grechi et al., 2021), all acknowledging that this comes at the risk of changing TIR images over time (Scaioni et al., 2017).

In contrast, active thermography involves the artificial heating (or cooling) of an object to stimulate materials and excite thermal signatures (Lagüela et al., 2016). For IRT-3DDF, Brumana et al. (2018a) pre-heat an entire room within the Magio Grasselli Palace to help identify the structural geometry of brick vault intrados, with active thermography helping generate discernible thermal signatures. Similarly, Maierhofer et al. (2011) heat a stone column with two 500W halogen lamps, fusing TIR images with a TLS mesh to correlate temperature and delamination. In contrast, Barbieri et al. (2023) captured TIR images of water treatment performance on historic brickwork, subject to targeted wetting (cooling) over a 75-minute period. Importantly, approaches incorporating active thermography must consider how heating/cooling may affect the fabric of a historic building and whether such interventions remain non-destructive (Sutherland et al., 2023).

Notably, the choice of thermography approach will impact the availability of methods within IRT-3DDF. For example, to undertake TIR capture at night, common for energy audit surveys (Lucchi, 2018), the capture of accompanying VIS imagery will be redundant (Borrmann et al., 2014). For Lenton Lodge Gatehouse, passive thermography was used, with no additional heating of the structures utilised to generate a discernible thermal signature (apart from solar radiation). However, for

both case studies in Chapter 5, different timings were chosen to compare and contrast approaches for IRT-3DDF, with Lenton Lodge Gatehouse captured during the day and Dunrobin House captured at night. This allows for different fusion methods to be applied and for the extraction of meaningful conclusions to be assessed at different times.

3.5.4 Survey Control & Thermal Targets

Metric Survey Control

Establishing control for geospatial surveys is central for effective scaling, positioning and evaluation of derived 3D products (Historic England, 2024). Whilst traditional geodetic surveys measure reflective, black-and-white or ‘natural feature’ targets using a theodolite or total station, IRT-3DDF has required the construction of thermal-specific markers: observable across different imaging modalities; identifiable in inherently low resolution TIR images; and extractable in different perspectives, conditions and locations. This has led to approaches that construct targets using materials with contrasting emissivity values that can generate clear markers; which, after determining their 2D image coordinates across all domains, can be used to assign control and validate the metric quality of resulting models (Figure 3.17). The importance of geodetic surveys undertaken with thermal-specific markers is critical not only for estimating the overall geometric accuracy of 3D modelling approaches, but central to IRT-3DDF methods implementing projective (Adamopoulos et al., 2020b), Helmert (Patrucco et al., 2022b) or datum (Dahaghin et al., 2021) transformations.

Whilst current approaches for IRT-3DDF survey markers have been wholly acceptable, several considerations have yet to be addressed in their entirety, notably: (1) the durability of markers for long-term IRT-3DDF research; (2) the performance

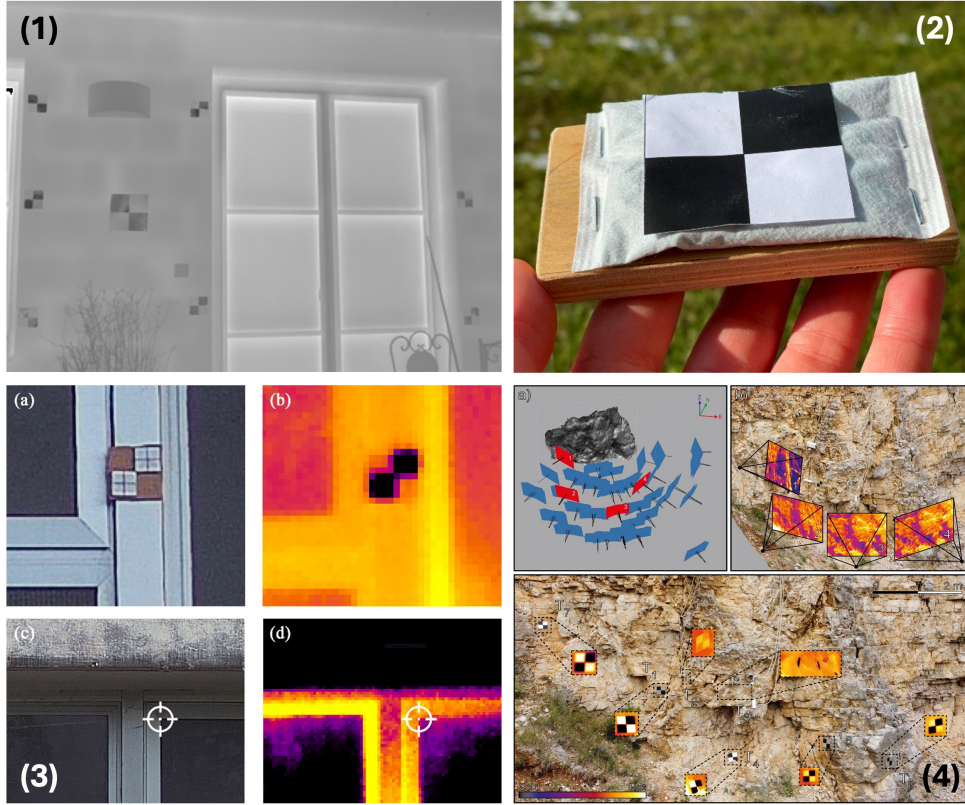


Figure 3.17: Examples of thermal-specific markers constructed for IRT-3DDF, including: (1) aluminium foil- (Dlesk et al., 2021), (2) hand-warmer- (Paziewska & Rzonca, 2022), (3) thermal/natural- (Patrucco et al., 2022b), and (4) reflective aluminium (Grechi et al., 2021) checkerboard markers.

of targets under different environmental conditions and observation angles, and (3) the relationship between target size and GSD. Therefore, to generate a thermal-specific marker capable of withstanding various prolonged changing environmental conditions, selective laser sintering (SLS), a high-precision 3D printing technique fusing powdered nylon resin by a guided laser, was employed for the creation of thermal target baseplates. In addition, milled aluminium and rubber squares were used to generate a crosshair identifiable in both thermal and visible spectrums. Thermal-specific markers were constructed with dimensions of 60×60 mm, in line with suggested dimensions (Historic England, 2024) and importantly a $GSD \sim 2.5 \times$ (23.5 mm GSD at 18 m) to $\sim 7.5 \times$ (7.6 mm GSD at 6 m) greater than that of the WWP. For Lenton Lodge Gatehouse, eight markers were adhered to the façade of Lenton Lodge using Rothbucher Systeme RMSK FIX, a survey-grade and heritage-

suitable adhesive that leaves no trace once removed and does no damage to building fabric. Targets were positioned maximising coverage, visibility and access for installation, whilst minimising obstructions, survey data voids and possible damage to historic fabric (Figure 3.18).

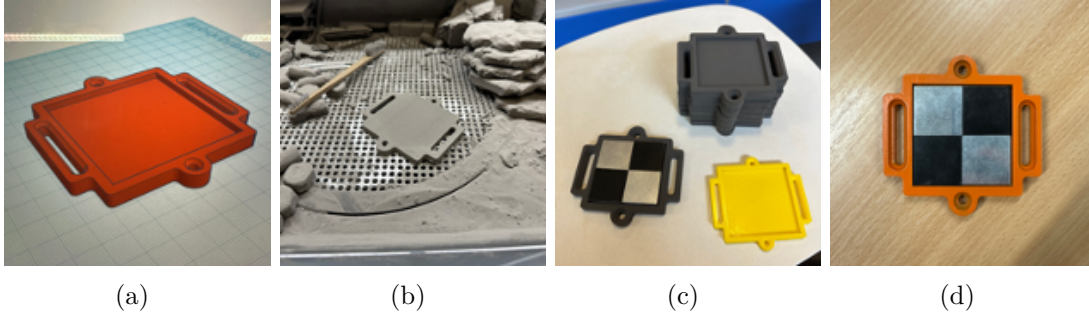


Figure 3.18: Design and production of the 3D-printed thermal-specific survey markers.

For the derived geodetic survey, a Leica TS10 total station theodolite (TST)(angular precision: 0.1 mgon; non-prism distance accuracy (0 m-500 m): 2 mm + 2 ppm) was used to measure markers following the specifications outlined in (Figure 3.19) (Historic England, 2024). Of the eight thermal-specific markers and additional ten natural points measured, half of each were used as GCPs and CPs. Measurement performance was determined by the repeated observation of markers and natural points from three primary control locations, with the derived coordinates being the mean of the easting (X), northing (Y) and height (Z) of the three observations. Table 3.4 presents the results of the control network for all points captured at Lenton Lodge Gatehouse, with Figure 3.20 demonstrating the extraction of the thermal-specific markers across GSDs.



Figure 3.19: (a, b) Geodetic survey of Lenton Lodge Gatehouse, including (c) the Leica TS10 total station.

Markers	No. Control Points	$\mu\sigma$ (mm)	MSE (mm)
Thermal-specific	8	0.48	0.27
Natural	10	1.07	0.62
Overall	18	0.77	0.44

Table 3.4: Metrics of the geodetic survey at Lenton Lodge Gatehouse using the Leica TS10 Total Station. Here, the number of measured control points, the mean standard deviation ($\mu\sigma$) for repeated marker measurements and the mean standard error (MSE) across all markers is provided.

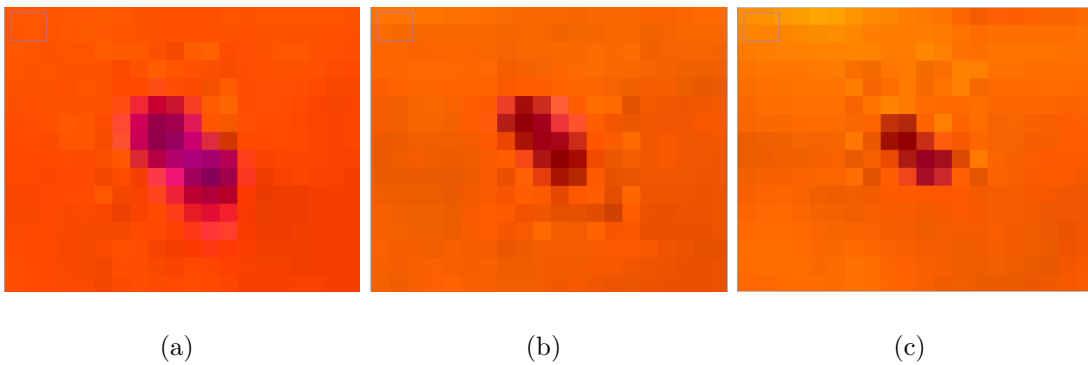


Figure 3.20: Example WWP images showing the designed thermal-specific targets at various GSDs (and distances): (a) 9.8 mm (7.5 m), (b) 12.4 mm (9.5 m) and (c) 15.0 mm (11.5 m).

Radiometric Survey Control

An established network of known temperature measurements is critical for validation within IRT-3DDF, allowing 2D pixel- (T_{2D}), 3D model- (T_{3D}) and real-world (T_{obs}) temperature values to be compared during thermal modelling. As previously stated, the barrier to effective TIR radiometric calibration, correction and validation is in the access to accurate and precise temperature measurements (Kelly et al., 2019); which, when not present, have resulted in methods that often ignore radiometric accuracy of derived thermal models. For IRT-3DDF, temperature validation has commonly been achieved using temperature loggers (Dugdale, 2016) or thermometers, with Lagüela et al. (2011b) & Lagüela et al. (2013) and Lin et al. (2019b) & Paziewska and Rzonca (2022) correlating 3D models with known temperature measurements from contact (TESTO 720) and infrared (non-contact)(RAYNGER MX4 & Sonel KT-165) thermometers, respectively. Similarly, Adán et al. (2021) orchestrated a network of environmental sensors to provide the requisite data for environmental monitoring, ground truthing and comfort assessments.

The novelty and originality of the thermal-specific markers outlined in Section 3.5.4, designed to be low-cost, easy-to-manufacture and highly configurable, is that their presence allows for several open challenges within IRT-3DDF to be solved, notably: (1) their use as configurable calibration markers for lab- and field-based geometric calibration procedures, (2) their use as survey markers for photogrammetric scaling, referencing and validation, and (3) their use as ground truth for radiometric calibration, correction and verification. To determine radiometric ground truth (and reflected temperature for Section 3.4.1), temperature measurements were undertaken on the thermal-specific markers at Lenton Lodge Gatehouse using the RS PRO RS-8876 thermometer (temp. res.: 0.1 °C; temp. acc.: ± 1 °C)(Figure 3.14), with a distance to spot ratio of 1:20 and the ability for emissivity to be set during measurement. Due to their visibility and consistent temperature throughout TIR

image sequences, temperature readings were taken on the rubber elements of the thermal survey targets. In addition, two temperature readings, one on each side of the building, were taken on the predominant building material, perpendicular and ~ 1 m from the façade.

3.6 Conclusion

A foundational understanding of IRT and photogrammetric principles is central to IRT-3DDF, without which superficial 3D products are derived with little geometric or radiometric accuracy. The novelty of this chapter has been to demonstrate the acquisition and application of such knowledge, outlining the critical decisions that have been made for successful IRT-3DDF in subsequent chapters. Critically for IRT-3DDF, a balance must be sought between maximising radiometric consistency whilst implementing the necessary procedures central to structure-from-motion photogrammetry, with methodological choices for one discipline often having considerable impact on the effectiveness of the other. In addition, the IRT ‘Cheatsheet’ provided in Appendix B shares a useful framework for those looking to undertake IRT-3DDF surveys, building on the established use and sharing of guidance within DCH. The justification and results of the decisions in this chapter will be discussed in subsequent methods, highlighting the benefits and limitations that come from implementing each fundamental practice.

Chapter 4

Calibration for IRT-3DDF

Contents

4.1	Introduction	82
4.1.1	Calibration: Aims & Objectives	83
4.1.2	Chapter Structure	84
4.2	TIR Geometric Calibration: State-of-the-Art	84
4.3	Methodology	87
4.3.1	Cameras	87
4.3.2	Calibration Set-Up	88
4.3.3	Interior Orientation	92
4.3.4	Relative Orientation	100
4.4	Results	105
4.4.1	Lenton Lodge Gatehouse	105
4.4.2	Castello Di Arco	112
4.5	Discussion	118
4.5.1	TIR Geometric Calibration	118
4.5.2	Calibration for IRT-3DDF	123
4.6	Conclusion	125

4.1 Introduction

The determination of precise and reliable interior orientation (IO) parameters for TIR cameras is critical for their subsequent use in photogrammetric processes. Whilst geometric calibration represents a largely routine procedure for cameras sensing in the visible spectrum (Remondino & Fraser, 2006), where 2D boards or 3D coded test fields are commonly used, the process presents several challenges when undertaken for TIR cameras. Notably, these comprise: (1) the inherent low resolution of commercial TIR sensors, (2) the construction of accurate calibration targets with sufficient difference between material emissivities (ε), and (3) the imaging of control points with high contrast and low noise (Usamentiaga et al., 2017). In parallel, the estimation of the relative orientation (RO) between sensors has played a crucial role in the development of IRT-3DDF methods using TIR cameras (Section 2). By establishing a fixed RO between independent cameras (Alba et al., 2011), or exploiting tandem sensors within the same camera housing (Ham & Golparvar-Fard, 2013a), the mapping of TIR-RGB/VIS image pairs can be realigned through effective calibration. Importantly, whilst the standard error of image coordinate residuals (ElSheikh et al., 2023) and individual calibration parameters (Luhmann et al., 2016) have become established precision metrics for estimating the performance of geometric camera calibration, these cannot be used as a proxy for accuracy within subsequent photogrammetric applications (Fraser, 2013). Therefore, the appraisal of calibration methods requires emphasis to be placed on the *use* of IO and RO parameters (Hastedt & Luhmann, 2015; Luhmann et al., 2013), particularly when specific challenges and novel applications within IRT-3DDF rely on accurate geometric calibration (Dlesk et al., 2021).

4.1.1 Calibration: Aims & Objectives

Adapted from the metrological study presented by Sutherland et al. (2025a), this chapter addresses Objective 2 of the thesis, determining robust and rigorous methods for the geometric calibration of TIR cameras employed within photogrammetric data fusion. Notably, the performance of different calibration approaches applied to two IRT-3DDF methods is assessed to determine the importance of pre-calibration for thermal modelling of architectural and cultural heritage. The methods employed within this chapter not only demonstrate the origins of a novel IRT-3DDF method expanded in Chapter 5, but demonstrate the mitigation of several challenges well-established within IRT-3DDF. Finally, as a wealth of photogrammetric CV software capable of performing geometric calibration now exists, a first-of-its-kind comparative analysis of popular commercial software benchmarked against dedicated metrology software will indicate the level of precision, control and information being sacrificed in order to achieve greater procedural automation (Duran & Atik, 2021; Remondino & Fraser, 2006). Therefore, the research reported in this chapter will:

1. Compare 2D board and 3D field calibration processes, detailing production, observation and calculation of IO and RO parameters using three commercial software packages: MathWorks' MATLAB, Agisoft Metashape and Photometrix's Australis;
2. Assess the accuracy of derived IO parameters applied to IRT-3DDF using a combined TIR-RGB bundle block adjustment for a historic building façade (2D-2D fusion);
3. Assess the accuracy of derived RO parameters applied to IRT-3DDF using a relative pose implementation on medieval frescoes (2D-3D fusion).

4.1.2 Chapter Structure

The chapter is structured as follows: Section 4.2 provides a brief state-of-the-art of TIR geometric calibration and its applications for IRT-3DDF; Section 4.3 undertakes a comparative analysis of TIR geometric calibration using different calibration targets, software and distortion parameters; Section 4.4.1 applies IO parameters to a novel 2D-2D fusion method, combining TIR-RGB bundle block adjustment; Section 4.4.2 applies RO parameters to a 2D-3D fusion relative pose implementation; and Section 4.5 reviews the performance of both 2D board and 3D field calibration processes, providing suggestions for future adoption, refinement and applications.

4.2 TIR Geometric Calibration: State-of-the-Art

The introduction of error in geometric calibration for TIR cameras is largely attributed to two components: the construction of accurate calibration targets and the extraction of control points in the derived thermal images (Usamentiaga et al., 2018). Originating from the CV community, the adoption of 2D calibration boards have become the predominant approaches for TIR geometric calibration, made using low emissivity materials ($\varepsilon < 0.5$; e.g., aluminium foil, copper, stainless steel) contrasted with high emissivity materials ($\varepsilon > 0.8$; e.g., rubber, wood, velvet) (Fryskowska-Skibniewska et al., 2022). 2D calibration boards have been utilised to calibrate infrared sensors across different spectral wavelengths (Kubota et al., 2021; St-Laurent et al., 2016; Swamidoss et al., 2021), for passive and active thermography (Luhmann et al., 2016; Rankin et al., 2013; Sun et al., 2022), and for indoor and outdoor applications (Bison et al., 2012; Harguess & Strange, 2014; Motayyeb et al., 2023). However, the overall accuracy of existing 2D calibration boards, often manually constructed to generate suitable calibration patterns, requires new methods for geometric calibration tailored to TIR cameras.

In a bid to improve the quality of calibration boards, several works have looked to reduce human intervention by employing superior manufacturing processes for accurate board construction. Usamentiaga et al. (2017)’s critical study demonstrates the use of DiBond[®], an aluminium composite material, and UV-printing to produce a cost-efficient, easy-to-manufacture and accurate 2D calibration board. Similarly, ElSheikh et al. (2023)’s comprehensive literature review analysed the best materials, heating techniques and associated errors of existing geometric calibration methods for TIR cameras, using a stainless steel (SS304) base plate with UV-printing and laser-engraving for calibration boards. Zhang et al. (2023a) demonstrated the suitability of high-tolerance 3D-printing to improve accuracy, constructing 3D-printed elements (PLA and aluminium plate) configurable into a 2D checker-board pattern. Roshan et al. (2024) introduced high-tolerance aluminium ArUco (Garrido-Jurado et al., 2014) components into a traditional checker-board, mitigating the established issues of partial pattern observation and extraction for 2D calibration boards.

Although these approaches provide considerable improvements in target construction, 2D calibration boards remain prone to the projective coupling of interior (IO) and exterior (EO) orientation parameters (Luhmann et al., 2016), with higher correlations between parameters a likely bi-product of poorer network geometry imposed by low resolution TIR cameras. Conversely, effective 3D calibration approaches reduce the propensity for correlations between IO and EO parameters, leading to greater overall accuracy for subsequent photogrammetric applications (Remondino & Fraser, 2006). Much like 2D boards, 3D calibration fields have been constructed using various materials and pattern configurations for TIR geometric calibration, with Westfeld et al. (2015) creating a 3D test field using velour and silver foil and Yastikli and Guler (2013) using an iron calibration rig with 77 plastic markers of contrasting depths. In a bid to calibrate thermal stereo set-ups, Rangel and Soldan (2014) and Adán et al. (2017) used reflective markers for the calibration of thermal

cameras fixed with a depth camera and laser scanner, respectively. Both Luhmann et al. (2013) (2D board and 3D field) and Dlesk et al. (2021) (2D board, 2-level- and 3-level 3D fields) compared multiple calibration approaches, determining 3D fields as superior for subsequent 3D thermal modelling. However, opportunities are still present to determine the performance of new 3D test fields for: the calculation of IO parameters, their comparison against existing 2D board approaches, and their application for IRT-3DDF.

Finally, the exploitation of fixed stereo systems within IRT-3DDF has enabled calibration approaches to determine the RO between sensors and cameras for data fusion applications. By assuming a fixed RO between multiple cameras or sensors, such information can be used as an a-priori constraint for sensor fusion (Patrucco et al., 2020), to determine a universal transformation key between image pairs (Dlesk et al., 2021), or as a means of additional sensor integration (Adamopoulos et al., 2020b). For sensors in fixed relative orientation within the same camera housing, Lecomte et al. (2022) and Macher and Landes (2022) used MATLAB’s Stereo Camera Calibrator tool to estimate the relative pose of TIR and RGB cameras, the latter comparing results to known camera dimensions. Fixing several cameras on a stereo rig, Lerma et al. (2010) present a distance-constrained self-calibration approach, obtaining more precise IO and EO parameters for a FLIR B4 thermal camera when integrating known baseline constraints. Ursine et al. (2012) assessed the accuracy of RO parameters using a 2D board, determining on-site calibration as a suitable approach under favourable conditions. Notably, the development of reliable RO calibration methods and the application of derived parameters for novel IRT-3DDF methods remains overlooked and warrants further investigation.

4.3 Methodology

The following section outlines the processes involved in the geometric calibration of *all* cameras used in subsequent IRT-3DDF methods. This includes: (1) the construction of both 2D board and 3D field calibration targets; (2) the calculation of IO parameters for a combined TIR-RGB bundle block adjustment; and (3) the determination of RO parameters for a relative pose implementation. Importantly, TIR and RGB cameras are not fixed in RO in Section 4.4.1, meaning imagery for the calibration processes and IRT-3DDF method were acquired asynchronously. For Section 4.4.2, independent cameras were fixed on a stereo rig to determine the IO and RO parameters for eventual IRT-3DDF methods.

4.3.1 Cameras

For the capture of TIR images, the Workswell[®] WIRIS Pro (Section 3.4.1), featuring both thermal (640×512 pix) and visible (1920×1080 pix) spectrum sensors within the same camera housing, was employed (Figure 4.1). Although the Workswell WIRIS Pro features both an infrared and visible sensor, WWP is used as the designation for the infrared sensor within the camera. To differentiate between visible spectrum images derived from the sensor housed within the Workswell WIRIS Pro, as opposed to images captured from independent digital cameras; the term VIS is adopted for the visible spectrum sensor housed within the Workswell WIRIS Pro thermal camera, whilst the term RGB is used to denote separate digital cameras. For RGB images, two different digital cameras were used for each data fusion method listed in Section 4.1.1, including a Sony $\alpha 7$ R II (7952×5304 pix)(RGB1) and Nikon D750 (6016×4016 pix)(RGB2). Geometric specifications for each camera are listed in Table 4.1.



Figure 4.1: Cameras used for TIR geometric calibration and IRT-3DDF methods, comprising: (a) Workswell WIRIS Pro (WWP & VIS), (b) Sony α 7RII (RGB1), and (c) Nikon D750 (RGB2).

4.3.2 Calibration Set-Up

The design of effective TIR geometric calibration targets reduce human error in construction, generate high-contrast images through different material emissivity values, and maintain contrast throughout the calibration imaging. Although 2D calibration boards have become the established standard for IRT-3DDF, forming the comparative basis within this section, this work aims to introduce a novel, low-cost and multi-purpose 3D field that is: precise in manufacturing and construction, easy to assemble and configure, and provides clear markers in both image modalities. Where calibration methods for IRT-3DDF differ from other TIR approaches is that targets must be observable in both infrared and visible spectral ranges, thereby allowing homologous points to be identified across image modalities for the estimation of RO parameters. Therefore, construction and observation processes must appreciate the duality of the required calibration target, creating targets capable of calibrating multiple sensors precisely and simultaneously.

For all calibration targets, ASTM E1933-14 was used to determine emissivity values of each target material (Section 3.5.1)(ASTM International, 2022). For each calibration target, eight perpendicular images were captured at varying temperatures, with each image constituting four points where a surface-modifying material was

Workswell WIRIS Pro infrared sensor (WWP)	
Resolution	640×512 pix
Sensor Size (FPA)	10.88×8.71 mm
Pixel Size	17.00 μm
Nominal Focal Length	13.00 mm
Workswell WIRIS Pro visible sensor (VIS)	
Resolution	1920×1080 pix
Sensor Size (CMOS)	5.23×2.94 mm
Pixel Size	2.72 μm
Nominal Focal Length	3.50 mm
Sony α7RII (RGB1)	
Resolution	7952×5304 pix
Sensor Size (CMOS)	35.90×24.00 mm
Pixel Size	4.50 μm
Nominal Focal Length	35.00 mm
Nikon D750 (RGB2)	
Resolution	6016×4016 pix
Sensor Size (CMOS)	35.90×24.00 mm
Pixel Size	5.95 μm
Nominal Focal Length	50.00 mm

Table 4.1: Camera specifications for the Workswell WIRIS Pro (WWP & VIS), Sony α 7RII (RGB1) and Nikon D750 (RGB2) cameras.

adhered (TESTO emission tape)($\varepsilon=0.95$). The determination of emissivity values for both 2D board and 3D field targets can be seen in Figure 4.2 and Table 4.2. For all 2D and 3D approaches, a temperature-controlled water bath was utilised to submerge the 2D board and 3D field markers to excite thermal contrast. Components were submerged in the water bath for 90 secs at 40 °C prior to positioning, constituting a passive thermography approach once image capture was undertaken (ElSheikh et al., 2023; Usamentiaga et al., 2017).

2D Board

For the 2D calibration target, a checker-board was designed using DiBond[®], an aluminium composite material, and an industrial signage UV-printer (Figure 4.3)(Us-

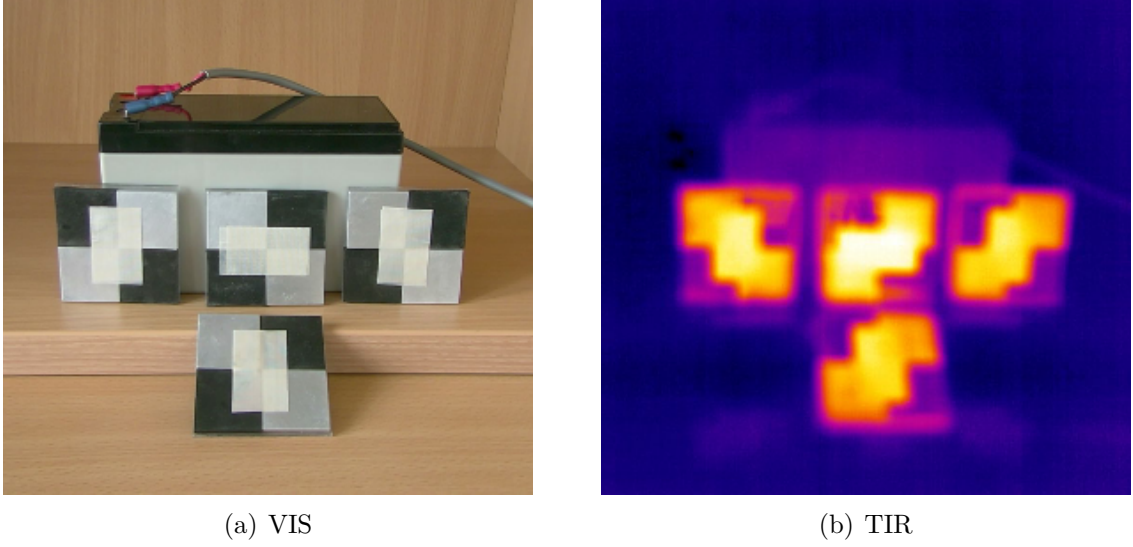


Figure 4.2: Emissivity determination for the 3D test field markers, utilising TESTO emission tape ($\varepsilon = 0.95$). Here, the emissivity tape is visualised in both (a) VIS and (b) TIR image domains.

	Material	ε	σ	$\Delta\varepsilon$
2D Board	DiBond®	0.62	0.04	0.20
	UV-printing	0.82	0.04	
3D Field	Aluminium	0.20	0.03	0.65
	Rubber	0.85	0.05	

Table 4.2: Calculation of material emissivity (ε), standard deviation (σ) and difference in emissivity of materials ($\varepsilon\Delta$) used for both 2D board and 3D field calibration targets.

amentiaga et al., 2017). DiBond, comprising two aluminium sheets sandwiching a polymer core, provides several benefits over previously listed materials as it is: (1) perfectly planar, (2) a good heat conductor, (3) inexpensive, (4) available in several material finishes, (5) able to be cut to any size, and (6) appropriate for high-resolution UV-printing (i.e., <600 dpi). An A3 piece of silver, brushed-finish DiBond was used, printed on with black ink using two passes of a SwissQ Nyala 4 UV-printer (1350 dpi, 53.1 dots/mm). As both checker-board (ElSheikh et al., 2023) and circle-board (Usamentiaga et al., 2017) patterns have been deemed suitable for high-precision calibration targets, a 5×8 (35×35 mm) checker-board pattern was printed.

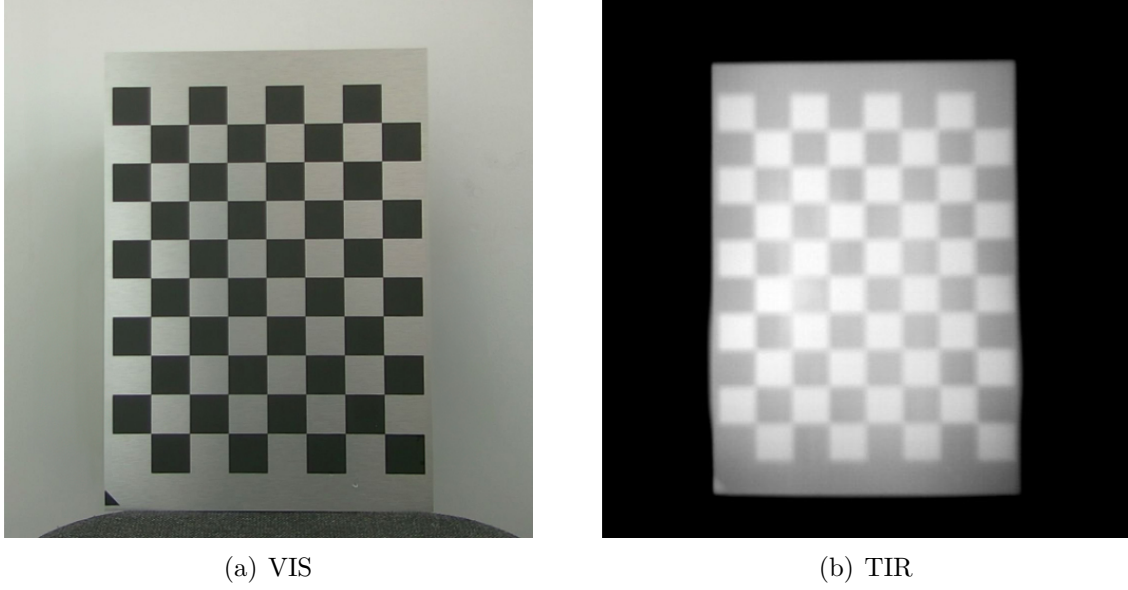


Figure 4.3: 2D checker-board imaged in both (a) VIS and (b) TIR domains.

3D Field

For the 3D calibration targets, two separate multi-level test fields were configured using thermal-specific markers (Section 3.5.4). Importantly, these multi-level 3D test fields were designed to be low-cost, easy-to-manufacture and highly configurable to encourage future adoption. These provide a crosshair generated by materials with different colours and emissivity values (Table 4.2), clearly identifiable in both image modalities. 16 markers provided a suitable configuration that covered both image formats and could be positioned after active heating. Figure 4.4 shows the two 3D calibration targets developed for the determination of IO and RO parameters in Section 4.4.1 & 4.4.2, respectively. For the metric evaluation of the IO values in Section 4.4.1, each marker was measured using a Leica TS10 total station theodolite (angular accuracy: 0.1 mgon; distance accuracy (non-prism): 2 mm + 2 ppm), although 3D world coordinates were not included within the self-calibrating bundle adjustment. For Section 4.4.2, a series of scale bars were measured between markers to provide scaling and reference for the relative pose estimation.

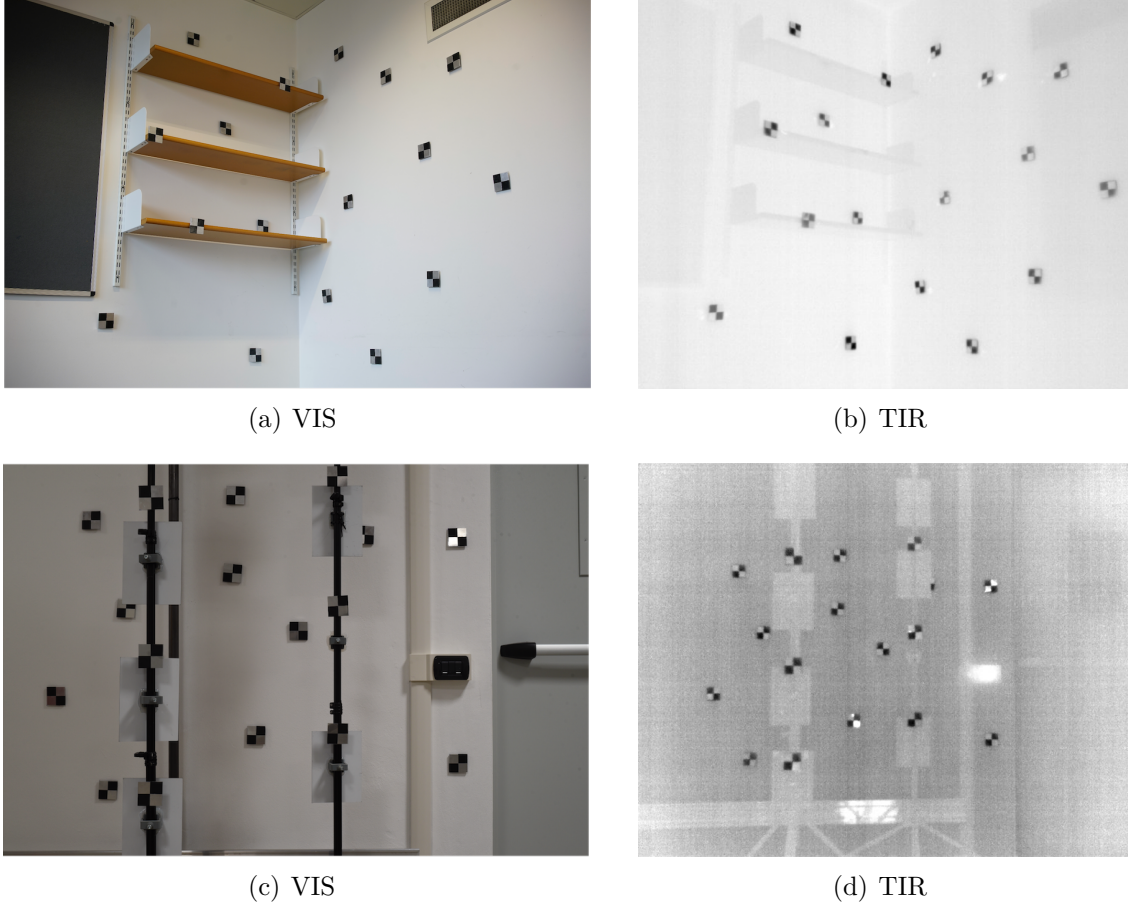


Figure 4.4: 3D fields imaged in both VIS and TIR image domains for the (a,b) IO and (c,d) RO calibration procedures.

4.3.3 Interior Orientation

For TIR geometric calibration, the WWP can be defined as a standard photogrammetric camera following the Brown model (Brown, 1971), merely observing a different spectral range (7-14 μm). However, as the configuration of the WWP's germanium lens and FPA sensor is optimised for radiometry as opposed to geometry (Luhmann et al., 2016), the determination of accurate distortion parameters is paramount. Therefore, to test the importance of the distortion coefficients, parameters were determined for each calibration target with either two- or three radial distortion coefficients (k_1, k_2, k_3), in addition to tangential distortion coefficients (p_1, p_2). After initial investigations, the FPA sensor was deemed perpendicular, meaning no differential scaling (b_1) or non-orthogonality (b_2) values were included as required

calibration parameters (Fraser, 2013). Finally, the optics of the WWP were fixed during manufacturing, with the camera focussed at infinity and unchanged during each calibration procedure (Workswell s.r.o, 2021). As an appropriate camera network geometry is fundamental for effective calibration using both 2D and 3D calibration targets (Luhmann et al., 2006), images were taken with: high convergence, orthogonal roll ($\pm 90^\circ$), varying depths along the principal axis, and the entirety of the target within the image format (Figure 4.5)(Remondino & Fraser, 2006).

For the calculation of IO parameters, three separate commercial software packages were employed for geometric calibration. Firstly, MathWorks' MATLAB, a popular software for CV camera calibration, was used for the geometric calibration using the 2D board. Secondly, for the 3D field calibration target, Agisoft Metashape, an industry-standard software widely adopted for SfM and MVS workflows, was used to assess the feasibility of proposed methods to work with commonly-employed photogrammetric software. Finally, to assess the validity of proposed 3D field methods for high-precision metrological applications, Photometrix's Australis was also used to undertake calibration processes. Although both 3D field software are capable of undertaking calibration using marker-less feature-based self-calibration (Perda et al., 2024), only 2D marker coordinates were used within the calibration pipelines. For both software, the same 2D image coordinates, 3D world coordinates and accompanying constraints were used for implementation. 2D image coordinates of the test field markers were user-defined in Agisoft Metashape and weighted at 0.33 pix. Similarly, 3D marker coordinates, used solely for metric evaluation, were measured with a Leica TS10 total station, with a weight of 2 mm applied to account for propagated error values.

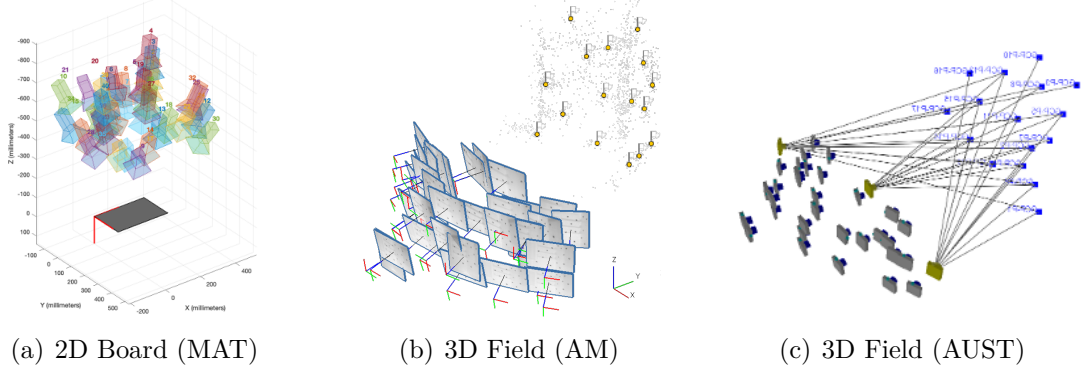


Figure 4.5: IO camera networks obtained for the (a) 2D board using MATLAB, (b) 3D field using Agisoft Metashape and (c) 3D field using Australis.

2D Board (MATLAB)

For the 2D board, MATLAB’s Camera Calibrator app (2024b) was used for the calibration of the WWP, using Zhang’s algorithm (Zhang, 2000). This method involves the capture of a 2D planar target from which detected points can be compared to known metric pattern dimensions. Geometric calibration is achieved through a closed-form solution, followed by a non-linear refinement based on the maximum likelihood criterion. Of the 42 images captured for calibration, two were discarded due to partial pattern recognition not meeting a nominal minimum threshold (Min-CornerMetric=0.3), with the remaining 40 being used to calculate IO parameters and accompanying errors. The 2D board distortion profiles (Figure 4.6) and distortion models (Figure 4.7) can be seen below, with the resulting IO parameters compared in Table 4.3.

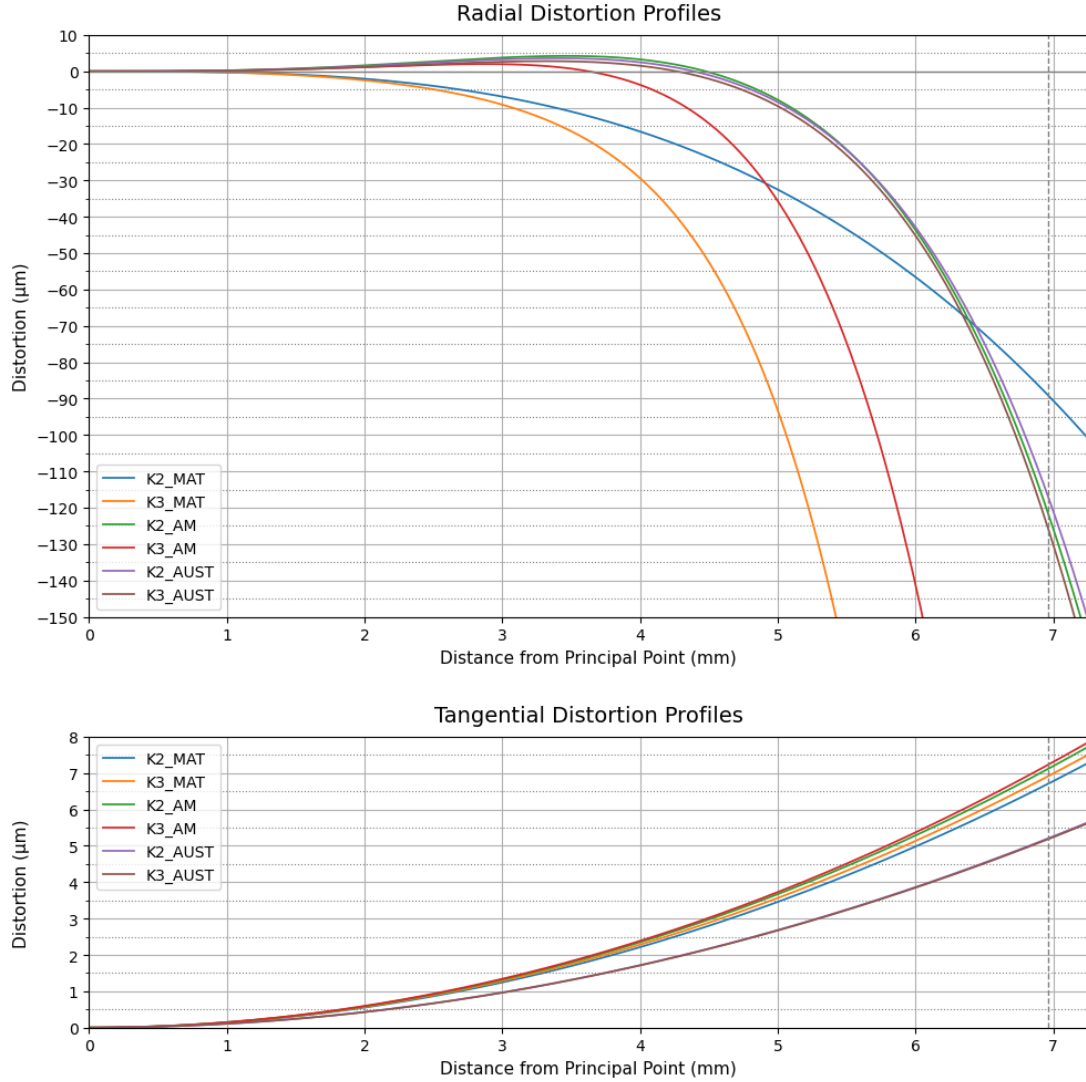
3D Field (Agisoft Metashape)

As the Metashape GUI does not allow for purely marker-based image orientation, several procedures were implemented to undertake a self-calibrating bundle adjustment prioritising marker 2D coordinates for the 3D field. Firstly, nominal values for the focal length and pixel size were provided, with redundant calibration pa-

rameters fixed at zero (i.e., b_1 , b_2 , extraneous k coefficients). Secondly, images were oriented in a standard photogrammetric workflow, including weighted 2D marker coordinates as part of the initial bundle adjustment. Once all images had been oriented, all tie points were removed (apart from one to allow for optimisation) and cameras were optimised in an additional self-calibrating bundle adjustment. This allowed for the markers to be the predominant information for the bundle adjustment as opposed to ambiguous tie points generated from local features. The 3D field distortion profiles (Figure 4.6) and distortion models (Figure 4.7) can be seen below, with the resulting IO parameters compared in Table 4.3.

3D Field (Australis)

Finally, Photometrix’s Australis was included to provide a robust metrological solution for calculating the IO parameters of the WWP. Australis’ workflow was implemented with a series of spatial resections and bundle adjustments undertaken to orient all images using the weighted 2D marker coordinates. Once completed, an additional free network self-calibrating bundle adjustment was undertaken to refine the initial IO parameters. Of the 37 images used for calibration, all featured a minimum of eight intersecting rays per marker, with the convergence limit and number of bundle adjustment iterations set at 0.005 and 12, respectively. The results generated by Australis for the 3D field, including distortion profiles (Figure 4.6), distortion models (Figure 4.7) and IO values (Table 4.3) can be seen below.



The maximum radial distance from the principal point during calibration is represented by ----- ($r = 6.96\text{mm}$)

Figure 4.6: Gaussian radial and tangential distortion profiles generated from the obtained IO parameters with the 2D board using MATLAB (MAT), 3D field using Agisoft Metashape (AM), and 3D field using Australis (AUST). Here, the graphs represent the distortion originating at the image centre (u_o, v_o) and finishing at the maximum radial distance when using either two- (K2) or three- (K3) radial distortion coefficients are used as part of the geometric calibration.

Coefficient	Intrinsics	2D Board (MATLAB)		3D Field (Agisoft Metashape)		3D Field (Australis)	
		Value	σ	Value	σ	Value	σ
k_2	f (mm)	13.021	0.032	13.038	0.015	13.034	0.014
	c_x (pix)	311.219	1.821	312.585	0.626	312.926	0.740
	c_y (pix)	255.558	1.928	257.755	0.614	257.921	0.721
	k_1	0.044	0.009	-0.043	0.004	-0.039	0.003
	k_2	0.005	0.054	0.368	0.014	0.347	0.013
	p_1	-0.000	0.001	-0.001	0.000	0.000	0.000
	p_2	0.001	0.001	0.000	0.000	-0.000	0.000
	MRE (pix)	0.50		0.16		0.12	
k_3	S_O (mm)	3.20		3.25		2.06	
	f (mm)	13.017	0.032	13.033	0.015	13.030	0.014
	c_x (pix)	311.232	1.823	312.541	0.623	312.908	0.730
	c_y (pix)	255.482	1.926	257.741	0.612	257.940	0.711
	k_1	0.055	0.017	-0.031	0.008	-0.031	0.008
	k_2	-0.177	0.247	0.247	0.078	0.270	0.073
	k_3	0.791	1.047	0.350	0.221	0.214	0.204
	p_1	-0.000	0.001	-0.001	0.000	0.000	0.000
	p_2	0.001	0.001	0.000	0.000	-0.000	0.000
	MRE (pix)	0.50		0.16		0.12	
	S_O (mm)	3.65		3.25		2.08	

Table 4.3: IO parameters for WWP obtained with the 2D board using MATLAB, 3D Field using Agisoft Metashape, and 3D Field using Australis.

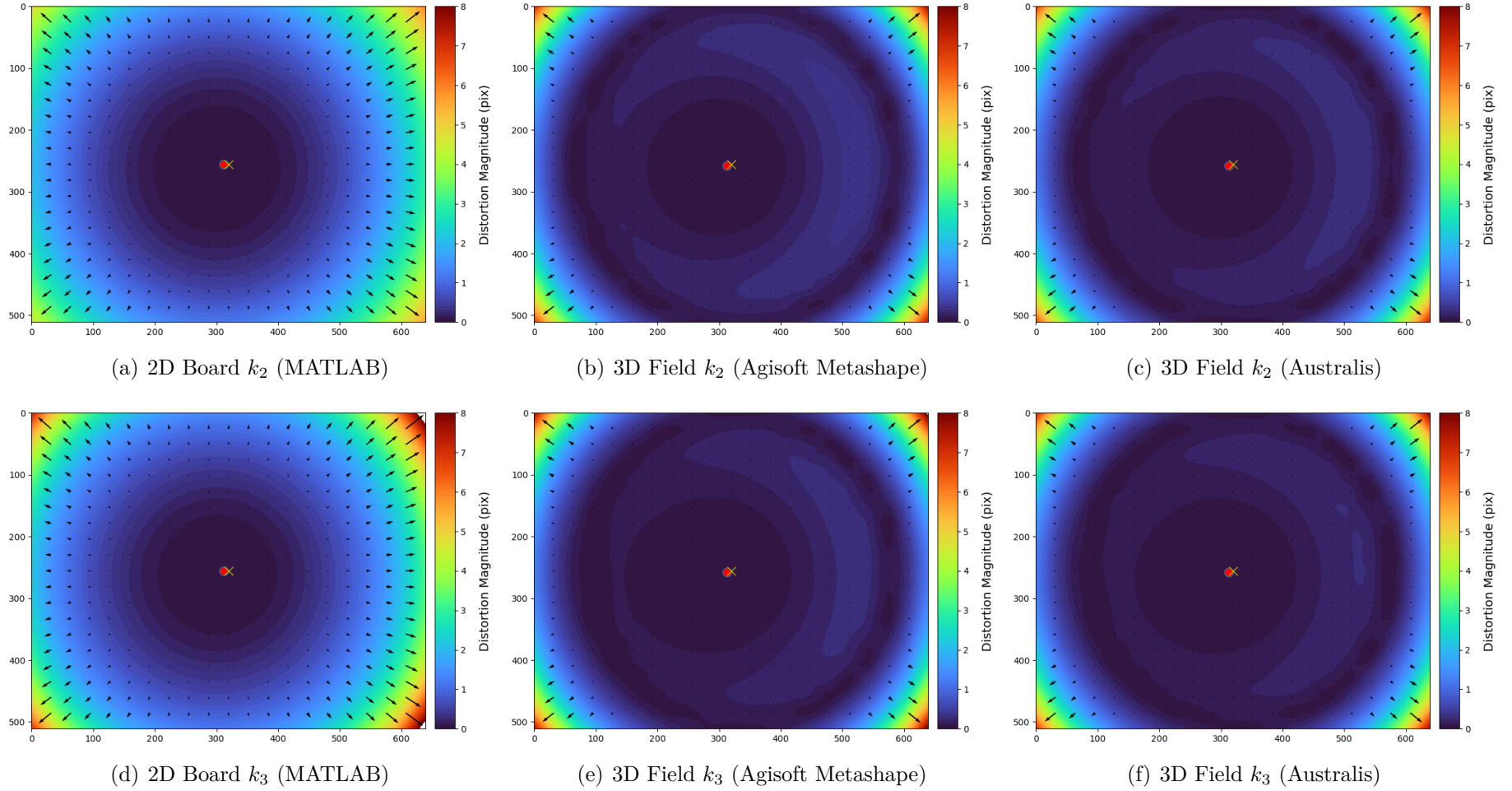


Figure 4.7: Distortion models for the (a,b) 2D board using MATLAB, (c,d) 3D field using Agisoft Metashape and (e,f) 3D field using Australis, showing distortion magnitude (pix), distortion normals ($\nearrow$), principal point ($\bullet$) and image centre (320,256)($\times$).

To assess the performance of each calibration approach, the standard error of IO parameters (σ), Mean Reprojection Error (MRE)(pix) in 2D image space and standard error (S_O)(mm) of camera poses in 3D object space are used (Luhmann et al., 2016). The results in Table 4.3 show notable differences between the 2D board and 3D field calibration approaches. Firstly, it must be noted that the scale of the 2D board (A3-sized) and resulting camera networks are significantly smaller than the 3D calibration targets ($\sim 4\times$), with the derived MRE and S_O heavily scale-dependent. MATLAB calculates a focal length $\sim 13 \mu\text{m}$ smaller than the 3D field approaches, which show comparable focal lengths across all four distortion coefficient tests (all within $8 \mu\text{m}$). This is likely a product of insufficient camera-to-object distance variation in the 2D board network geometry, with the planar target providing less depth in feature points and individual camera positions (Fraser, 2013). In addition, the higher intrinsic standard errors are largely due to the ambiguous extraction of checker-board corners (MRE=0.50 pix) propagating into greater IO imprecision. Most significantly, the results generated by the 2D board show significantly different distortion models from that of the 3D field calibrations (Figure 4.7), the culmination of higher IO standard errors, poor feature point extraction and the difficulty filling the image format effectively whilst maintaining preferable network geometry.

In contrast, the observable similarities in the IO values, distortion models and MRE from both Agisoft Metashape and Australis are expected, utilising the same 2D image coordinates for their respective bundle adjustments and obtaining marginally different radial and tangential profiles (Figure 4.6). Both software packages determined similar IO values across calibration tests, capturing the convex germanium lens geometry equating for a distortion of ~ 0.25 pix on the right of the image format. Australis obtains the best MRE and S_O results across both k_2 and k_3 coefficients. Notably, Australis provides greater control over the self-calibrating bundle adjustment (inc., iterations, convergence limit, minimum number of rays per marker);

which, when coupled with native access to the covariance matrix, allowed for results to be refined for greater IO precision and metric accuracy. However, the validity of these metrics must not be viewed in isolation for the assessment of calibration processes, but taken into account when intrinsics are applied independently in subsequent photogrammetric applications.

4.3.4 Relative Orientation

The estimation of the relative pose between fixed stereo cameras determines the translation and rotation vectors necessary for mapping one image onto another. For IRT-3DDF, this relationship is of particular importance when there are no means of determining the poses of TIR images independently, often due to low contrast or noisy images providing few discernible features for photogrammetric image matching. Furthermore, if TIR and RGB images are to be captured concurrently, this fixed relationship can help mitigate sub-optimal conditions for each image modality during image acquisition. For the comparison of 2D board and 3D field calibration targets, two separate baselines were assessed. Firstly, as the Workswell WIRIS Pro features both an infrared sensor and a visible sensor within the same camera housing, the baseline between both internal sensors was determined. However, as the resolution of the VIS sensor is sub-optimal for detailed modelling (1920×1080 pix)(Section 2.3.3), a stereo rig fixing the WWP and RGB2 was assembled to acquire additional data for Section 4.4.2. For each calibration set-up, as with the calculation of the IO parameters, baselines were determined with two- or three radial distortion coefficients to assess their importance in determining precise and reliable RO parameters for IRT-3DDF.

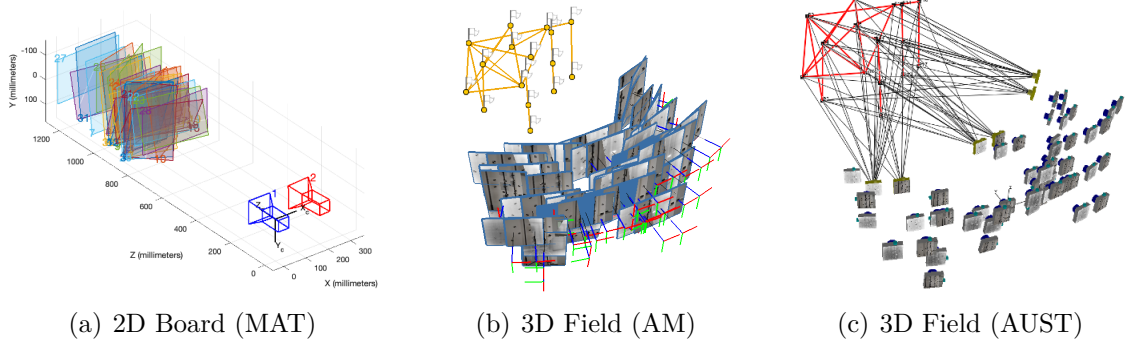


Figure 4.8: RO camera networks obtained for the RGB2-WWP with (a) 2D board using MATLAB, (b) 3D field using Agisoft Metashape and (c) 3D field using Australis.

2D Board (MATLAB)

MATLAB’s Stereo Camera Calibrator App (2024a) was used to compute the 2D board calibration. Firstly, IO values for the Workswell WIRIS Pro sensors (WWP & VIS) and RGB2 were re-generated using the same processes outlined in Section 4.3.3. These values are subsequently used as fixed values for the relative pose estimation. MATLAB’s Stereo Camera Calibrator is limited by its inability to calibrate two cameras of different resolution, meaning that imagery and associated IO values (i.e., f_x , f_y , c_x , c_y) for VIS and RGB2 had to be resampled to that of the WWP to accommodate stereo calibration (ultimately rescaled upon completion). For the VIS-WWP camera and RGB2-WWP stereo rig, 37 and 20 image pairs were processed, respectively. The estimated relative translation and rotation values for each stereo pair can be seen in Table 4.4, showing the pose of the Workswell WIRIS Pro thermal sensor with respect to the VIS/RGB2 sensor.

3D Field (Agisoft Metashape)

For the relative pose estimation using Agisoft Metashape, using a different 3D field than that used for the calculation of the initial camera parameters (Figure 4.4), IO values for all sensors and cameras were re-calculated using the approach outlined

in Section 4.3.3. Derived IO parameters were subsequently used as fixed values for Agisoft Metashape’s Multi-Camera System for the VIS-WWP sensors and RGB2-WWP stereo rig. Much like the IO method, ambiguous tie points were removed after initial orientation, with an additional self-calibrating bundle adjustment undertaken to refine RO values. Results were then scaled using a collection of scale bars allowing error metrics to be derived. For the VIS-WWP camera and RGB2-WWP stereo rig, 37 and 31 image pairs were processed, respectively. Results for both baseline tests can be seen in Table 4.4, similarly showing the mapping of the WWP onto VIS/RGB2.

3D Field (Australis)

For the calculation of RO parameters in Australis, the initial IO values for all sensors were obtained through the processes outlined in Section 4.3.3. Subsequently, image pairs were imported into Australis for stereo calibration with the same 2D marker coordinates and constraints as those used in Agisoft Metashape. A self-calibrating bundle adjustment was then undertaken to obtain the extrinsic parameters of the cameras, with scale bars used to scale the scene, derive the RO parameters and calculate error metrics. Results for both baseline tests in Australis can be seen in Table 4.4.

Coefficient	Baseline	Parameters	2D Board (MATLAB)		3D Field (Agisoft Metashape)		3D Field (Australis)	
			Value	σ	Value	σ	Value	σ
k_2	VIS-WWP	X (mm)	-41.530	0.191	-39.738	0.396	-40.271	2.691
		Y (mm)	-1.992	0.149	0.409	0.363	-0.148	2.845
		Z (mm)	0.309	0.114	1.142	0.219	6.124	1.456
		ω ($^\circ$)	-0.080	0.013	-0.092	0.008	-0.175	0.073
		ϕ ($^\circ$)	0.579	0.016	0.717	0.009	0.701	0.071
		κ ($^\circ$)	-0.080	0.010	0.001	0.005	0.000	0.021
		MRE (pix)		0.76		0.55		0.22
		S_O (mm)		4.48		3.20		2.32
	RGB2-WWP	X (mm)	-184.464	0.331	-183.659	0.629	-183.084	1.493
		Y (mm)	-0.216	0.241	1.501	0.578	1.242	2.023
		Z (mm)	6.925	0.252	9.737	0.501	11.154	1.276
		ω ($^\circ$)	0.917	0.015	0.586	0.015	0.424	0.080
		ϕ ($^\circ$)	3.644	0.020	3.243	0.015	3.208	0.066
		κ ($^\circ$)	-0.046	0.015	-0.147	0.014	-0.175	0.296
		MRE (pix)		0.87		1.10		0.14
		S_O (mm)		5.26		2.17		0.77
k_3	VIS-WWP	X (mm)	-41.503	0.191	-39.721	0.395	-40.540	2.889
		Y (mm)	-1.990	0.149	0.406	0.362	-0.614	2.895
		Z (mm)	0.371	0.114	1.272	0.219	-1.338	1.094
		ω ($^\circ$)	-0.080	0.013	-0.092	0.008	0.025	0.075
		ϕ ($^\circ$)	0.579	0.016	0.713	0.009	0.715	0.072
		κ ($^\circ$)	-0.080	0.010	0.001	0.005	-0.007	0.019
		MRE (pix)		0.76		0.55		0.29
		S_O (mm)		4.48		3.20		3.54
	RGB2-WWP	X (mm)	-184.446	0.331	-183.644	0.629	-183.143	1.495
		Y (mm)	-0.215	0.241	1.490	0.578	1.064	2.143
		Z (mm)	7.017	0.252	9.340	0.502	11.132	1.263
		ω ($^\circ$)	0.917	0.014	0.565	0.015	0.418	0.085
		ϕ ($^\circ$)	3.644	0.020	3.267	0.015	3.213	0.068
		κ ($^\circ$)	-0.046	0.015	-0.146	0.014	-0.154	0.311
		MRE (pix)		0.87		1.10		0.14
		S_O (mm)		5.26		2.11		0.76

Table 4.4: RO parameters for WWP, VIS and RGB2 using the 2D board and 3D field calibration targets.

From the results in Table 4.4, an indication of the overall performance of the derived RO parameters can be determined, with the translation (X , Y , Z) and rotation (ω , ϕ , κ) of the TIR sensor with respect to the reference sensor (VIS or RGB2). Here, XYZ are aligned with the image coordinate system (x , y), where X is aligned with x pointing right, Y is aligned with y pointing down, and Z points forwards from the perspective centre. To gain some understanding of the performance of relative pose estimation, the baselines (horizontal displacement X in camera space) between the fixed sensors was utilised. According to Workswell s.r.o., the baseline between the centre of both lenses within the Workswell WIRIS Pro is 40.7 mm. Similarly, the measured baseline between the lenses of the RGB2 and WWP stereo rig was 184 mm, with a depth offset of 10 mm.

All results achieve a reasonable estimation for the baseline (X) between the VIS-WWP and RGB2-WWP sensors, with MRE mostly <1 pix and $3D S_O < 6$ mm for all tests. Expectedly, a greater proportional range is present within the translation values and S_O for the VIS-WWP across each software package and distortion coefficients than that of the RGB2-WWP. This is largely due to the shorter baseline and poorer resolution of the VIS sensor with respect to RGB2. However, the RO results for both 3D field tests show markedly similar results for the stereo rig baseline, although the σ for the Australis are captured within the RO results are notably higher than both MATLAB and Agisoft Metashape across all results. Whilst the derived RO parameters and error metrics suggest adequate performance for all calibration processes, their application for IRT-3DDF is needed to determine the value of calibration processes for specific data fusion- and heritage-specific applications.

4.4 Results

The following section outlines the application of the derived IO and RO parameters for the documentation of architectural and cultural heritage using IRT-3DDF. Section 4.4.1 executes a combined TIR-RGB bundle block adjustment using the IO parameters obtained in Section 4.3.3 as fixed camera intrinsics. This case study introduces critical tools expanded in Chapter 5, with results demonstrating initial experiments and future avenues utilising multi-modal image matching to generate correspondences across image modalities. Section 4.4.2, a 2D-3D fusion relative pose implementation, applies the RO parameters from Section 4.3.4 to derive a textured 3D thermal model of medieval wall paintings. Notably, whilst insights regarding radiometric calibration, correction, and validation of TIR cameras are integrated from Chapter 3, the focus of this chapter remains on geometric calibration and resulting network geometry.

4.4.1 Lenton Lodge Gatehouse

For the evaluation of the IO parameters obtained in Section 4.3.3, Lenton Lodge Gatehouse, a 19th century Grade II* listed property in Nottingham, UK, was used as a case study (Appendix A.1). Facing East-South-East, thermal images were acquired in January 2024 between 10:30-11:00 AM using passive thermography, maximising the baking of the structure from solar radiation and minimising shadows on the façade. 20 WWP images were captured in a single strip with a GSD of 9.81 mm, whilst 25 RGB1 images were captured in two strips with GSDs of 1.03 mm and 1.29 mm (Figure 4.9). Importantly, imagery from the WWP and RGB1 were captured independently with no fixed relative orientation, meaning no consistent baseline between the cameras can be assumed. Additionally, a network of thermal-specific survey markers and natural points was acquired for subsequent

referencing, scaling and quality evaluation (Section 3.5.4). Furthermore, a 45 min acclimatisation period was undertaken for the WWP’s FPA sensor and CPU to generate stable temperature readings in accordance with findings from Section 3.4.2. Finally, NUCs were undertaken at 3 min intervals during the survey to maintain consistent temperature values over the duration of the TIR sequence (Section 3.4.2).



Figure 4.9: Lenton Lodge Gatehouse, Nottingham, UK.

With the proliferation of deep-learning approaches for both mono- (Jin et al., 2020) and multi-modal (Velesaca et al., 2024) image matching, the possibility of using ‘out-of-the-box’ neural networks for photogrammetric data fusion is becoming a reality (Elias et al., 2023). Therefore, to explore multi-modal image matching between TIR and RGB images, Deep-Image-Matching (DIM), an open-source Python library designed for robust multi-view image matching, was utilised (Section 5.3.3)(Morelli et al., 2024a). To facilitate effective multi-modal matching, WWP images underwent a histogram normalisation and UnSharp Masking to enhance detectable features. Additionally, RGB images were downsampled from their original resolution (7952×5304 pix) to $\frac{1}{8}$ th size (994×663 pix) to more closely mirror the resolution of the WWP images. For all WWP and RGB images, 8,000 keypoints were detected

using a SuperPoint (DeTone et al., 2018) feature extractor and LightGlue feature matcher (Lindemberger et al., 2023) (Sp+Lg)(Figure 4.10). Utilising DIM, all 45 images were matched exhaustively, resulting in a possible 990 image pairs across the dataset.

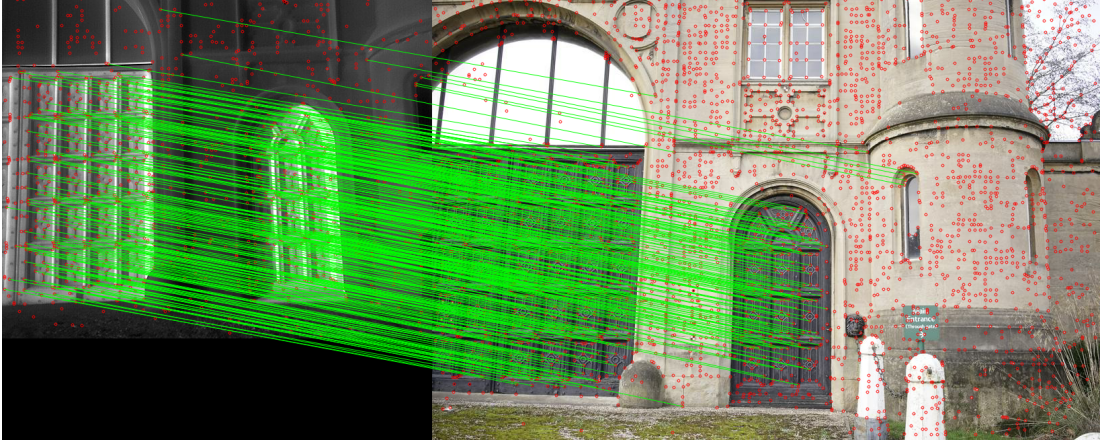


Figure 4.10: An example WWP-RGB1 image pair generated from SuperPoint + LightGlue (Sp+Lg) in Deep-Image-Matching (DIM).

For the combined TIR-RGB bundle block adjustment, the generated database from DIM was imported into COLMAP for image orientation (Schonberger & Frahm, 2016). Here, each block of WWP and RGB1 images were assigned a distinct camera model, using the IO parameters from the TIR geometric calibration in Section 4.3.3. To keep the focus of the experiments on the performance of the TIR geometric calibration (independent variable), intrinsics for RGB1 were calculated through self-calibration using the 25 images of Lenton Lodge Gatehouse and held constant for each calibration subset. Importantly, IO values for the RGB1 focal length and principal point were divided by eight to match the resolution of the downsampled images provided to DIM. For each calibration set-up, an independent bundle adjustment was undertaken, fixing the IO values for both the WWP and RGB1 to assess the accuracy of the geometric calibration. Finally, to determine the possibility of providing no initial IO values for a combined TIR-RGB bundle block adjustment, a self-calibration approach using two radial distortion coefficients was undertaken whereby camera parameters were refined during the bundle adjustment. For all

cases, derived sparse 3D tie points and camera poses were imported into Agisoft Metashape for metric evaluation (Figure 4.11).



Figure 4.11: Combined TIR-RGB bundle block adjustment camera network for Lenton Lodge Gatehouse, showing the single strip of WWP images and two strips of RGB1 images.

Results for the combined TIR-RGB bundle block adjustment are listed in Table 4.5, separating RGB1 and WWP image blocks into their respective modalities for metric evaluation. For the WWP blocks, Root-Mean Squared Error (RMSE) on the CPs is separated into horizontal (X), depth (Y), vertical (Z) components, with the $\text{RMSE}_{\text{WWP}} \text{ TOT}$ representing the total RMSE of results in 3D space. Additionally, the RMSE of the RGB blocks ($\text{RMSE}_{\text{RGB1}}$) and the MRE of the WWP blocks (MRE_{WWP}) are included for a comprehensive assessment of the combined TIR-RGB bundle block adjustment and camera intrinsics. Firstly, the results of the WWP image blocks show notable differences between calibration approaches, with the $\text{RMSE}_{\text{WWP}} \text{ TOT}$ obtained from both 2D board calibration tests roughly twice that of the 3D field results. Calculating a focal length ranging from 13-

Parameters		RMSE _{RGB1} (mm)	MRE _{WWP} (pix)	RMSE _{WWP} (mm)			
				X	Y	Z	TOT
Self-Calibration	k_2	2.86	0.32	4.02	16.31	9.03	19.07
MATLAB	k_2	2.59	0.40	3.88	14.11	10.55	18.04
	k_3	2.60	0.31	3.98	12.39	8.63	15.62
Agisoft	k_2	2.44	0.27	2.72	5.74	6.12	8.82
Metashape	k_3	2.47	0.28	2.45	6.91	5.69	9.28
Australis	k_2	2.32	0.29	3.05	5.44	6.35	8.89
	k_3	2.19	0.29	3.06	5.60	6.28	8.95

Table 4.5: Results of the combined TIR-RGB Bundle Block Adjustment for Lenton Lodge Gatehouse, separating results based on software and distortion coefficients utilised.

16 μm less than that obtained from the 3D field methods (Table 4.3), coupled with an inability to capture the unique distortion of the WWP’s germanium lens (Figure 4.7), the 2D board shows significant error in the depth (Y) component of the RMSE_{WWP}. Consequently, the results obtained from the the 2D board calibration targets are similar to that of COLMAP where no prior calibration is undertaken and nominal IO values are corrected via ‘on-the-job’ self-calibration. The poorer MRE_{WWP} and RMSE_{WWP} of the 2D board using MATLAB is likely due to the scale of the calibration network geometry being insufficient for upscaling to the 3D scene and the projective coupling of IO and EO values rendering IO values scene-dependent.

In contrast, the results of the 3D fields using both Agisoft Metashape and Australis show positive results, obtaining similar MRE_{WWP} values for the combined TIR-RGB bundle block adjustment and all RMSE_{WWP} TOT values in line with the GSD of the WWP block (GSD=9.81 mm). Agisoft Metashape achieved the best RMSE_{WWP} TOT results with two radial distortion coefficients, with Australis achieving comparable results with the best RMSE_{RGB1} when three radial distortion coefficients were used. Interestingly, the MRE_{WWP} and RMSE_{WWP} results obtained using either two- or three radial distortion coefficients show little discernible difference; which,



(a)



(b)

Figure 4.12: Generated 3D models from the combined TIR-RGB bundle block adjustment of Lenton Lodge Gatehouse, textured with (a) RGB1 and (b) WWP thermal images.

when combined with the similar distortion models and profiles (Figures 4.6 and 4.7), suggests that for germanium lenses where no complex distortion is present, either two- or three distortion coefficients could be used for effective TIR geometric calibration.

Furthermore, as each combined TIR-RGB bundle block adjustment featured the same RGB1 IO values and multi-modal correspondence, all RGB blocks present similar $\text{RMSE}_{\text{RGB1}}$ of <3 mm, with mono-modal matches (those *between* RGB images) from Sp+Lg allowing for successful orientation of the downsampled RGB block. However, a strong correlation is also apparent between the WWP ($\text{RMSE}_{\text{WWP TOT}}$) and RGB1 ($\text{RMSE}_{\text{RGB1}}$) image blocks ($r=0.86$), suggesting the influence of the WWP (and by virtue the fixed IO values) on multi-modal image orientation. Similarly, the MRE_{WWP} across all tests shows notable correlation, closely mirroring the $\text{RMSE}_{\text{RGB1}}$ and RMSE_{WWP} values. This reiterates the suitability for image- and object space metrics to be utilised in tandem for the assessment of the combined TIR-RGB bundle block adjustment and camera intrinsics (Sledz & Heipke, 2022).

The results obtained with COLMAP, undertaking simultaneous IO and EO via ‘on-the-job’ self-calibration, proves the benefit that can be provided by effective geometric calibration for sensor fusion pipelines (Hastedt & Luhmann, 2015). Given the nominal focal length and image centre as starting IO values, the results of the combined TIR-RGB bundle block adjustment perform worse than if successful pre-calibration is undertaken, suggesting the benefit that can be gained from calibration using multi-modal correspondence. Notably, $\text{RMSE}_{\text{WWP TOT}}$ were the same for when nominal values (f_x, f_y, u_o, v_o) or full IO values from the 3D field $(f_x, f_y, c_x, c_y, k_1, k_2, p_1, p_2)$ were refined, with the bundle adjustment deriving the same IO values for both the WWP and RGB1 regardless of the initial starting point. The resulting 3D models, generated from the RGB1 images and textured with WWP thermal images, can be seen in Figure 4.12.

4.4.2 Castello Di Arco

For the evaluation of the RO parameters, Castello Di Arco, located in the Sarca Valley of Trentino, northern Italy, was chosen as a case study (Appendix A.2). In a bid to reduce water ingress and plaster delamination caused by the castle's rock wall foundations, a space heater has been utilised to actively heat the façades after heavy rain and reduce the impact of the moisture on the frescoes (Figure 4.13). Whilst active IRT provides established NDT techniques for cultural heritage wall paintings (Williams et al., 2022), to determine the extent of the deterioration and assess the current conservation practices, IRT-3DDF was employed to localise visible signs of delamination and discolouration with quantitative temperature differences. Therefore, to assess the impact of current conservation practices, the imaging survey was undertaken after active thermography, with the space heater removed just prior to data acquisition. The survey was undertaken in October 2024 after successive days of dry weather (avg. temp. 13 °C). The fresco corner, facing North-East, was subject to solar radiation from 06:52 (sunrise) until the commencement of the survey (10:45-11:30 AM).

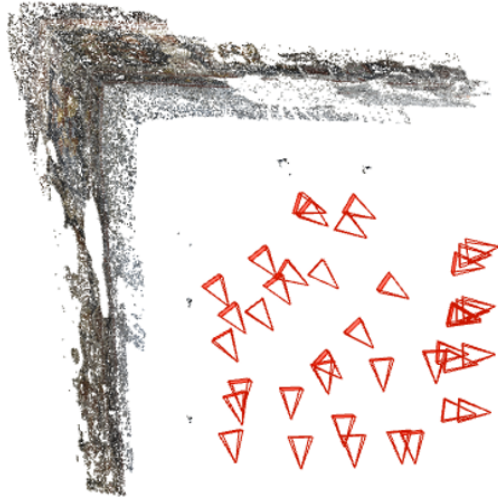
To utilise the findings generated in Section 4.3.4, stereo results were used from both the in-built thermal camera and purpose-built stereo rig baselines. For both baselines (VIS-WWP and RGB2-WWP), 40 stereo pairs were captured of the fresco corner and adjoining walls (Figure 4.13). Captured at a distance of ~ 2.5 m away from the frescoes, the GSD for the WWP, VIS and RGB2 sensors were 3.3 mm, 2.0 mm and 0.3 mm, respectively (Figure 4.14). As previously stated, the benefit of utilising known RO parameters is essential in cases where no discernible thermal contrast can be identified or generated through active thermography. For Castello Di Arco, presenting a temperature range of ~ 3 °C, the derived thermal images presented low contrast and high noise, meaning few suitable correspondence could be used for feature-based matching (Figure 4.15).



Figure 4.13: Castello Di Arco's medieval frescoes, Trentino, Italy.



(a)



(b)

Figure 4.14: Castello Di Arco's medieval frescoes, featuring (a) the stereo rig featuring the WWP and RGB2 cameras, and (b) the stereo rig camera network.

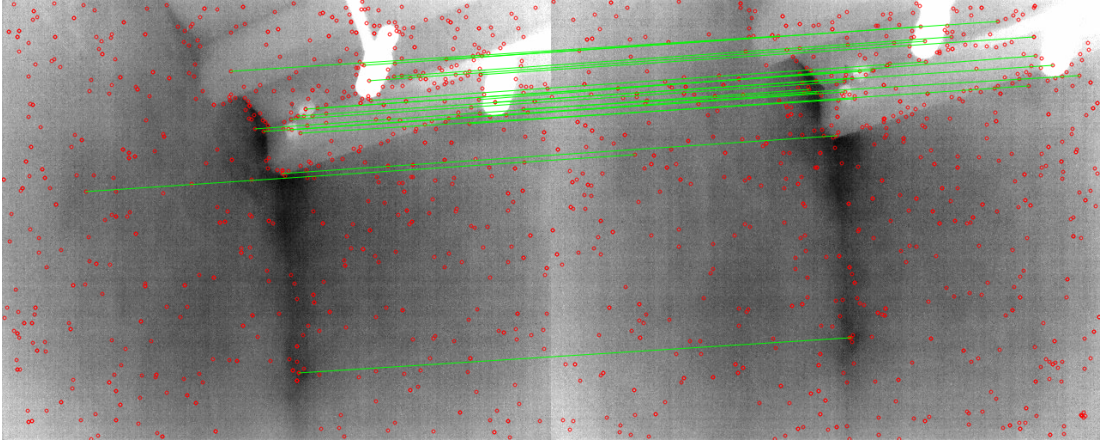
Initial tests proved the inability for WWP images to be oriented mono-modally, necessitating RO parameters to a reference camera/sensor for effective modelling. To mitigate these challenges, a network of the thermal-specific markers, actively heated

and positioned in front of the walls (Figure 4.13), provided manual tie points to aid orientation of the WWP block and generate multi-modal correspondence (Bison et al., 2012). To generate control for the resulting models, measured scale bars were taken between positioned markers, with two used as control and two reserved for validation. To undertake the modelling, Agisoft Metashape’s Multi-Camera System was utilised, with VIS/RGB2 set as the ‘primary’ camera (geometry) and the WWP set as the ‘secondary’ camera (texture). IO values were fixed from the initial calibration process (Section 4.3.4), with the X , Y , Z translations and ω , ϕ , κ rotations input as initial constraints for the bundle adjustment. In addition, 2D image coordinates (markers and tie points) were both weighted at 1.0 pix and 1.5 pix for the RGB2-WWP and VIS-WWP baselines, respectively. Using a 2D-3D fusion approach, a 3D thermal model was generated, using the ‘primary’ image dataset as the base geometry, subsequently textured using the ‘secondary’ TIR image dataset (Figure 4.16).

Table 4.6 shows the results from each baseline calibration test specifying the distortion co-efficients used for the WWP during calibration (k_2 or k_3). The RMSE on the measured scale bars are separated for each modality to provide an assessment of the image orientation in each imaging domain. Additionally, as with the results reported in Section 4.4.1, a self-calibration process has been undertaken where no initial IO or RO parameters were provided, with IO, RO and EO parameters obtained simultaneously via a self-calibrating bundle adjustment. For the VIS-WWP baseline, the success of the obtained RO parameters is evident, with all approaches able to orient the 40 stereo pairs successfully. For the sensors within the Workswell WIRIS Pro, calibration results are comparable across all calibration targets, obtaining similar $\text{RMSE}_{\text{VIS-WWP}}$ results for the reference VIS sensor ranging from 1.5-2.7 mm. Interestingly, the 2D board obtained the best $\text{RMSE}_{\text{VIS-WWP}}$ value amongst VIS blocks, showing the suitability of 2D boards for standard CV geometric calibration. Importantly, all results (apart from MATLAB) provide a better



(a)



(b)

Figure 4.15: An example of geometrically un-verified SIFT matches obtained in COLMAP between sequential mono-modal (a) RGB2 and (b) WWP images, resulting in 5,924 and 22 matches, respectively.

$\text{RMSE}_{\text{VIS-WWP}}$ for the WWP image blocks than that of purely self-calibration, suggesting the superiority of pre-calibration for sensor systems with short baselines.

For the stereo rig (RGB2-WWP), all approaches achieve a similar $\text{RMSE}_{\text{RGB2-WWP}}$ for the RGB2 image blocks, showing the suitability of each calibration approach (across different distortion coefficients) to obtain camera intrinsics for traditional digital cameras. However, notable differences for the 2D board calibration target using MATLAB are identifiable, with the IO and RO parameters leading to a markedly higher $\text{RMSE}_{\text{RGB2-WWP}}$ for the WWP block of ~ 9 mm. This is likely due to the different FOV between the fixed cameras during RO calibration, with the different focal lengths of the sensors (50 mm vs. 13 mm) making it more difficult

to fill both image formats with the 2D board during simultaneous image capture (Ursine et al., 2012). However, we see comparable results from both 3D field calibration approaches, obtaining a $\text{RMSE}_{\text{RGB2-WWP}}$ of <2 mm for the WWP image block. Significantly, Australis was able to obtain the lowest MRE, S_O and RMSE for the WWP blocks in RO calibration methods across both distortion coefficient subsets. Much like the findings of Section 4.4.1, the results from both 3D calibration methods demonstrate the superiority of the 3D field (and accompanying software) to obtain precise and reliable IO values that are image invariant, scene-independent, and metrically-accurate for photogrammetric reconstruction.

Parameters		RMSE VIS-WWP (mm)		RMSE RGB2-WWP (mm)	
		VIS	WWP	RGB2	WWP
Self-Calibration	k_2	2.43	4.30	1.46	1.42
	k_3	1.49	5.04	2.20	8.72
MATLAB	k_2	1.51	5.01	2.17	8.82
	k_3	2.72	4.12	1.26	1.87
Agisoft Metashape	k_2	2.74	4.24	1.26	1.89
	k_3	2.25	3.83	1.26	1.58
Australis	k_2	2.28	3.95	1.26	1.65
	k_3				

Table 4.6: RO results for Castello Di Arco’s medieval frescoes, with the measured scale bar RMSE (mm) for each baseline (VIS-WWP or RGB2-WWP) separated into each image modality.

As self-calibration was used as a benchmark to determine the ability for IO, RO and EO parameters to be obtained simultaneously in the same bundle adjustment, it is important to understand the performance of these results when compared to the RO pre-calibration processes. Firstly, the block RMSE values from the VIS-WWP baseline were favourable, establishing intrinsics in line with pre-calibrated values and an estimated baseline (X) of -41.2 mm (~ 0.5 mm from the reference value). However, the results of the RGB2-WWP baseline, obtaining a better $\text{RMSE}_{\text{RGB2-WWP}}$ for the WWP block than any calibration approach, generates a markedly different distortion model for the WWP (notably c_x , c_y and p_1 , p_2) and translation values

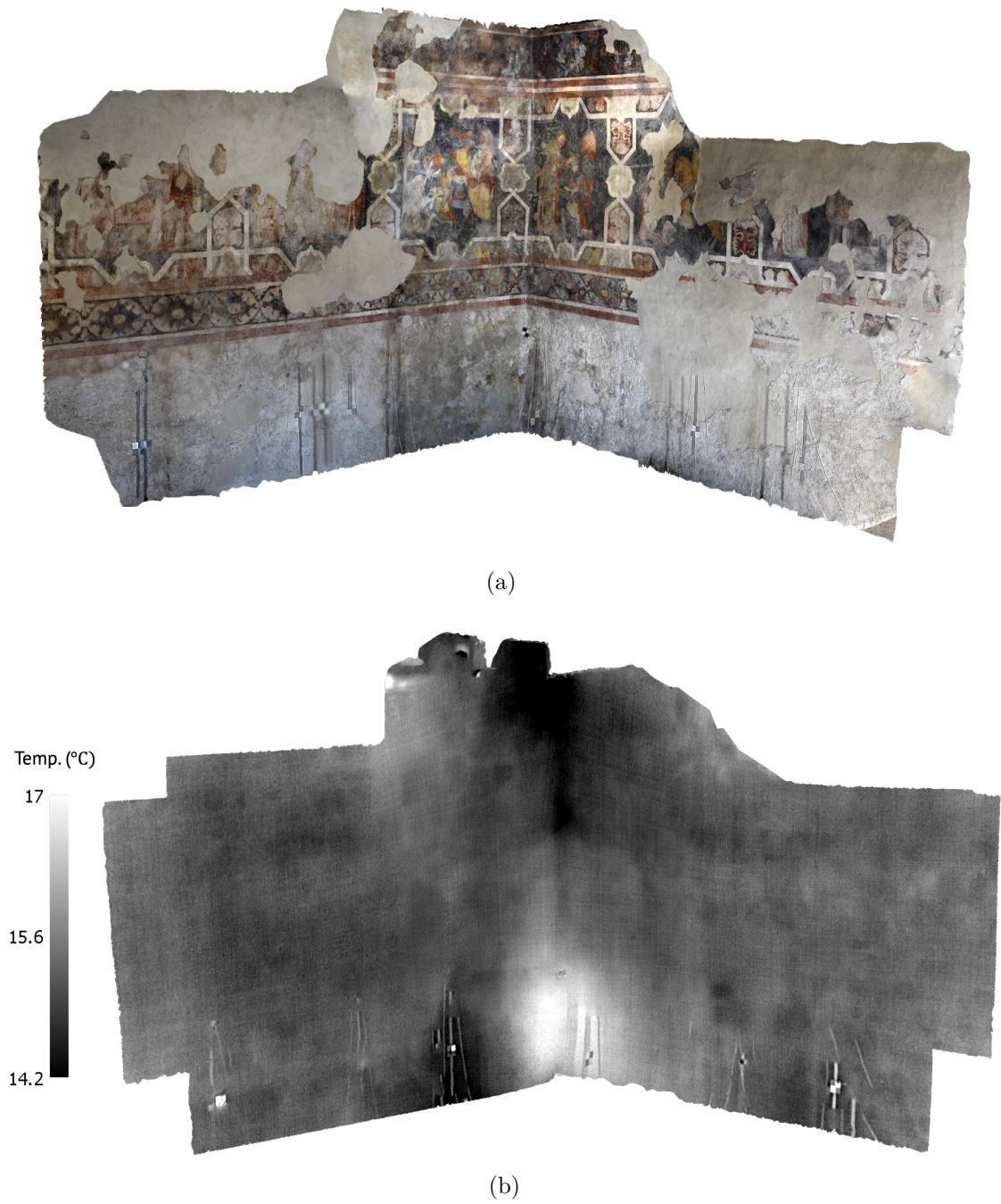


Figure 4.16: Generated textured models of Castello Di Arco's medieval frescoes, with the (a) RGB results showing the visible signs of deterioration and (b) showing temperature range visualising the most significant (cold) areas of water ingress.

($X=-187.8$ mm, $Z=-0.8$ mm). This suggests at the susceptibility for ‘on-the-job’ self-calibration to find *any* scene-dependent solution for IO and RO parameters that fit with the extracted tie points and included markers, especially when fewer than 200 tie points exist in the WWP block. This reiterates the importance of known IO and RO parameters for scenes where self-calibration is more impressionable, with parameters (and logical constraints) a necessary feature when insufficient correspondence is present or when modelled sensors present unique distortion profiles.

4.5 Discussion

4.5.1 TIR Geometric Calibration

Calibration Targets

As historical methods of TIR geometric calibration have been overly reliant on the manual construction of 2D boards, the ambition of this chapter was to determine a novel 3D test field design, preferable for photogrammetric camera calibration, which could be compared against established contemporary approaches. Critically, the success of all TIR geometric calibration approaches is in the ability for control points to be extracted with high contrast and low noise, with the greatest contributing factor being the emissivity of the materials used. As seen in Table 4.2, the difference in emissivity between the DiBond and UV-printing ink was notably small ($\Delta\epsilon=0.20$). This made it difficult to excite thermal contrast and extract accurate checker-board corners, resulting in uncertainty propagating into imprecise IO values and larger error values (i.e., IO σ , MRE, S_O)(Table 4.3). Although scale, network geometry and calibration set-up still play a role, material emissivity is likely the largest contributing factor for the poor calibration and IRT-3DDF results obtained with the 2D board, demonstrating the importance of material choice and

emissivity calculation as part of effective TIR geometric calibration.

In contrast, the emissivity difference for the 3D test field provided well-defined crosshairs with high contrast throughout the entirety of the imaging surveys ($\Delta\epsilon=0.65$). Highly-configurable in terms of quantity, geometry, scale and setting, the novelty of the proposed custom markers is characterised by their suitability as: (1) crosshair targets for 3D test fields; (2) additional tie points for photogrammetric multi-modal image matching; (3) GCPs and CPs for photogrammetric evaluation; and (4) known points for quantitative temperature verification. However, the proposed method is still limited by the manual identification of markers within the imaging surveys as opposed to fully-automatic feature detection. Roshan et al. (2024)’s 2D ChArUco checker-board provides a logical progression of the current thermal markers, with milled aluminium coded markers a novel approach for both future calibration and survey applications.

However, although coded markers (Roshan et al., 2024); distinctive shapes (Zhang et al., 2022); speckles (Lagüela et al., 2011a); actively heated elements (Sun et al., 2022) and mono- or multi-modal local features (Perda et al., 2024) may provide avenues for future exploration, the ability for these to be observed by commercial thermal cameras with inherent low resolution, common for TTISs incorporating multiple image- and range-based sensors (Adán et al., 2021), remains an open challenge. Finally, whilst the use of active thermography provided the necessary contrast for effective control point identification in both calibration targets, the considerable limitations of existing and alternate active thermography approaches include: the difficulty constructing electrical targets, the inaccuracy of control point extraction, the inability to take devices into the field, and the need for repeated heating cycles (Lagüela et al., 2011a; Zhang et al., 2023a).

Interior Orientation

The derived IO parameters, obtained by comparing calibration targets, the number of distortion coefficients and software packages, point to several considerations for TIR geometric calibration. As previously highlighted, the quality of the 2D board drastically affect the derived intrinsics, incapable of capturing the distortion of the germanium lens and producing significantly larger IO σ , MRE and S_O values (Table 4.3). In addition, whilst an initial set of 20 images were captured for interior orientation using MATLAB, an additional 20 (totalling 40 images) were needed to refine IO values within expected limits. Although MRE and S_O errors obtained during calibration would indicate agreeable performance of the 2D board, these metrics cannot be viewed in isolation, with the comparison and application of IO values in Section 4.4.1 demonstrating the inability for the 2D board to achieve precise TIR camera intrinsics for scene-independent photogrammetric use. As a result, even though 2D boards are established as a method for TIR geometric calibration, the susceptibility of projective coupling from suboptimal calibration network geometry, the tendency for derived IO values to be image-variant, and the observed scene-dependency of calibration processes means that the use of 2D boards should be limited for metrically-accurate IRT-3DDF.

In contrast, the 3D field, presenting a strong configuration of multi-level points, was able to capture the unique distortion of the WWP optics with either two- or three radial distortion coefficients. The acknowledged benefits of a 3D calibration field, including the ability for the image format to be properly filled, greater scale and depth information to be present, and the high convergence between point observations, allowed for more precise and reliable estimates of IO values. The comparable performance of both Agisoft Metashape and Australis software packages, both in terms of distortion profiles (Figure 4.6) and distortion models (Figure 4.7), is likely due to the identical camera network geometry, 2D marker coordinates

and associated constraints. However, for these results to be achieved, the inability for the Agisoft Metashape GUI to run a purely marker-based bundle adjustment required the removal of ambiguous tie points in the ‘textureless’ calibration target after initial alignment and for an additional optimising bundle adjustment to be undertaken. Therefore, when deciding between software packages, considerations between automation and accuracy must be made in accordance with the camera network for TIR geometric calibration, with Australis allowing for greater access to the self-calibrating bundle adjustment and accompanying parameter correlations. Finally, regarding choice of distortion coefficients or software package, there is little in the IO or RO calibration performance metrics or IRT-3DDF results that would suggest a preference for two- or three radial distortion coefficients; however, this will largely be TIR camera and calibration dependent.

For the self-calibration results in Table 4.5, it is apparent the influence that imprecise intrinsics can have for multi-modal image matching. The obtained principal point for the self-calibration presented a difference from calibrated IO values in c_x and c_y of ~ 10 pix and 4 pix, respectively; which, even though capable of determining a similar distortion profile to the 3D field approaches, resulted in expectedly higher $\text{RMSE}_{\text{TIR TOT}}$ (Clarke et al., 1998). As discussed in Section 4.4.1, the resulting IO values from the combined TIR-RGB bundle block adjustment self-calibration were practically the same when either nominal- (simple pinhole camera model), full- (OpenCV camera model) or pre-calibrated IO values were refined (Schonberger & Frahm, 2016). This demonstrates a considerable opportunity to understand the influence of bundle iterations and initial intrinsics on multi-modal image orientation. Therefore, the exploration of on-site self-calibration (Senn et al., 2022), utilising varying initial camera intrinsics and constraints, is a logical progression for Section 4.4.1 (Hastedt & Luhmann, 2015).

Finally, the difference in TIR intrinsics obtained for Sections 4.3.3 and 4.3.4 must be noted, stressing the need for TIR cameras to be regularly calibrated prior to pho-

togrammetric use (IO and RO calibration experiments were undertaken six months apart). For example, an observable offset in the principal point and tangential distortion coefficients was present for WWP IO values used in the Sections 4.4.1 & 4.4.2. This phenomena could be caused by mechanical stress or component instability within the Workswell WIRIS Pro, a possible side effect of rotational movements during calibration or camera handling and transportation (Luhmann et al., 2016). Furthermore, the influence of sensor temperature and ambient conditions during geometric calibration can be seen to have influence on derived TIR images, causing an observable increase in temperature values within the obtained calibration datasets (Figure 4.4). As sensor, camera housing and environmental temperatures have an established impact on resulting camera intrinsics (Daakir et al., 2019) and radiometric accuracy (Wan et al., 2021), the analysis of this effect on FPA sensors (detecting the very radiation that will be causing instability), warrants future investigation.

Relative Orientation

The results of the relative pose estimations in Sections 4.3.4 and 4.4.2 demonstrate the success of the proposed calibration approaches and importance of pre-calibrated IO and RO parameters for IRT-3DDF. Firstly, the results for the 2D board provided an agreeable estimation of the relative pose for both baselines, notably for the VIS-WWP baseline where the focal lengths and FOVs of the VIS-WWP image pairs made filling the image format more achievable. However, the results from both 3D field calibration approaches demonstrate the superiority of a strong camera network and multi-level calibration target for TIR relative pose estimation, with both Agisoft Metashape and Australis achieving successful results when IO and RO parameters were applied (notably for the RGB2-WWP stereo rig <2 mm for all image blocks). Importantly, both Sections 4.3.4 and 4.4.2 show the capability for RO parameters to be generated across varying baselines and sensors, critical for implementation

within IRT-3DDF where a vast array of resolutions and TTISs.

The importance of pre-calibration in Section 4.4.2 is evident, where self-calibration was highly unsuccessful in obtaining simultaneous IO, RO and EO parameters in a low contrast setting. Even though IO values for VIS and RGB2 were in line with pre-calibrated values, both VIS-WWP and RGB2-WWP results failed to provide reliable IO values for the WWP, with VIS-WWP results determining a focal length $\sim 100 \mu\text{m}$ from pre-calibrated values and RGB2-WWP results achieving a principal point (282, 243) significantly off from the image centre (320, 256). Clearly, Agisoft Metashape’s Multi-Camera System calculates IO values for the ‘primary’ camera as a priority, fitting the ‘secondary’ camera within the remaining constraints. In addition, altering the weights given to WWP tie points and user-defined coordinates of the custom markers, imperative for generating multi-modal correspondence, provided notable differences in obtained IO and RO parameters, likely due to the extremely low number of tie points in the WWP image block ($\text{No.}_{\text{WWP}} < 200$). Therefore, additional investigations assessing the stability of relative pose estimations for IRT-3DDF via ‘on-the-job’ self-calibration are needed, notably the inclusion of: (1) initial estimates and constraints for IO parameters; (2) initial estimates and constraints for RO parameters (most likely including horizontal displacement in X) (Lerma et al., 2010); and (3) greater TIR mono- and TIR-RGB multi-modal tie points (in a similar vein to Section 4.4.1).

4.5.2 Calibration for IRT-3DDF

Whilst the use of IRT as part of a broader catalogue of NDT techniques has long been established for architectural heritage (Tejedor et al., 2022), the presented examples of IRT-3DDF reveal opportunities where 2D imaging techniques can embrace synergies with 3D modelling practices (Adán et al., 2021). The ability to generate rapid, real-time and repeatable thermal models creates opportunities for the

assessment of architectural heritage through the comparison of 3D thermal models over time, the localisation of thermal anomalies and the quantification of material degradation. Similarly, the results from Castello di Arco, with known and observable discrepancies in the temperature of the medieval frescoes, point to possible interventions and areas for investigation. Firstly, it is evident the use of a space heater to alleviate the effects of the water seeping through the wall has been insufficient to address a pressing issue in the upper corner, with a temperature difference of ~ 1.5 °C from the fresco average still apparent. The use of IRT-3DDF helps demonstrate the extent of the temperature difference, the proximity of affected areas in relation to the space heater and identifies additional areas on the wall where conservation work should be considered. However, whilst the focus of this work has been on geometric calibration, for additional insights to be extracted, future emphasis on radiometric calibration, correction and validation is needed for the implementation of effective IRT-3DDF (Adán et al., 2023b).

Notably, the results from Section 4.4.1 also demonstrate the success of a novel deep-learning based multi-modal image matching method for thermal modelling, representing a real opportunity for the development of IRT-3DDF by exploiting ‘out-of-the-box’ neural networks not trained on TIR images. This section confirmed: (1) the ability for feature correspondence to be determined across image modalities, (2) the strength of these points for simultaneous image orientation as part of a combined TIR-RGB bundle block adjustment, and (3) the increased automation of IRT-3DDF without a reliance on known survey points or manual registration. Similarly, the results in Section 4.4.2 represent a critical method for IRT-3DDF where insufficient temperature differences means little thermal contrast is visible within a given scene. The ability for scenes to exploit proximal sensors and reliable calibration procedures represents a critical opportunity for indoor thermal modelling, where the generation of contrast is a likely obstacle. However, the manoeuvrability and construction of the RGB2-WWP stereo rig provides considerable limitations

in terms of data capture and synchronisation, likely exacerbated when the scene increases in scale and complexity. The findings from Section 4.4 demonstrate the benefit of using a high-resolution digital camera to generate the ‘base’ geometry for IRT-3DDF, with most commercial thermal cameras (and accompanying visible sensors) developed for building inspection as opposed to photogrammetry. Therefore, the logical evolution of this research should focus on combining the complementary methods of Section 4.4 for the establishment of TIR geometric calibration as a fundamental component of IRT-3DDF. This would further develop the novelty of multi-modal correspondence for orienting TIR and RGB images whilst utilising the proximal sensors within thermal cameras, helping mitigate environments where both mono- and multi-modal TIR correspondence may not be determined.

4.6 Conclusion

The determination of reliable IO and RO parameters for TIR cameras is critical for their subsequent use in photogrammetric processes. The performance of calibration procedures cannot be used as a proxy for their success in eventual 3D reconstruction, notably for data fusion methods utilising sensors of varying resolutions and modality. Addressing Objective 2, this chapter has comparatively evaluated existing 2D- and novel 3D calibration targets, detailing construction, calibration and applications for IRT-3DDF. Results suggest the superiority of a 3D field for the estimation of TIR intrinsics and provide insights regarding appropriate camera parameters, software packages and future refinements. In addition, the estimation of the relative pose between two baselines was determined to assess the accuracy of image orientation utilising in-built sensors from a singular thermal camera or a constructed stereo rig. The application of IO and RO parameters to two IRT-3DDF methods demonstrates the opportunities to be gained for NDT within architectural heritage exploiting thermal modelling. The first method, proposing a novel combined TIR-

RGB bundle block adjustment utilising deep-learning multi-modal correspondence, presenting an automatic data fusion method where simultaneous image orientation can be achieved without a reliance on GCPs or manual registration. This demonstrates the capability of using neural networks generalised outside of their training domain, a facet that can be investigated further to determine its durability across varying sensors, settings and scenarios (Chapter 5). The second method, utilising a fixed stereo rig for the generation of a textured 3D thermal model, provides a workflow demonstrating the importance of proximal sensors and effective RO parameters when low thermal contrast and few TIR correspondences can be determined. Future work will look to combine developed methods to exploit the higher resolution and flexibility of commercial digital cameras whilst utilising fixed sensors within a thermal camera for instances when TIR images require additional support for successful image orientation.

Chapter 5

Generation of IRT-3DDF

Contents

5.1	Introduction	129
5.1.1	Generation: Aims and Objectives	130
5.1.2	Chapter Structure	131
5.2	Multi-Modal Image Matching: State-of-the-Art	131
5.3	Methodology	135
5.3.1	Case Studies	137
5.3.2	Step 1: Image Acquisition and Pre-Processing	138
5.3.3	Step 2: Image Matching and Orientation	141
5.3.4	Step 3 and 4: Image Processing and Optimised Bundle Adjustment	147
5.3.5	Step 5: Thermal Modelling	148
5.3.6	Step 6: Evaluation	149
5.4	Results	151
5.4.1	Camera Pose Metrics	154
5.4.2	Temperature Metrics	160
5.4.3	Modelling	162
5.5	Discussion	165
5.5.1	Multi-Modal Image Matching for IRT-3DDF	165
5.5.2	2D-2D Fusion for IRT-3DDF	167
5.6	Conclusion	168

5.1 Introduction

Whilst IRT-3DDF research has grown over the past decade, marked by the ever-increasing availability, affordability and functionality of TIR cameras, critical bottlenecks have endured from its conception. Firstly, the inherent low spatial resolution and small imaging format of commercial TIR cameras, comparatively to modern DSLR or mirrorless cameras, means the determination of correspondences for photogrammetric processes remains challenging *between* different sensors. Secondly, as most TIR cameras do not provide any direct positioning data, the orientation of resultant imagery is not provided for incorporation into subsequent modelling methods. As seen in Chapter 2, IRT-3DDF research has relied on several broad approaches: the co-registration of independently oriented TIR and RGB image blocks using GCPs (Dahaghin et al., 2021; Scaioni et al., 2017) or iterative closest point (ICP) algorithms (Maset et al., 2017; Truong et al., 2017); the use of sensors in fixed relative orientation to determine TIR poses (Dlesk et al., 2021); or the fusion of TIR images with additional geometric (Adán et al., 2021) or parametric (Hoegner & Stilla, 2016) 3D models. However, the effectiveness of these methods is often restricted by: the manual collection of topographic measurements and identification of GCP projections in the imagery; a manual rough co-registration between the RGB and TIR blocks before ICP; the inability for surveys to be undertaken independently to optimise image capture for IRT (e.g., RGB acquisition during the day and TIR acquisition during the night); or, not benefiting from complementary observations from the geometric data in both TIR and RGB image blocks (e.g., multi-modal tie points that can improve the final orientation accuracy).

To overcome these limitations, fully automatic methods for IRT-3DDF must be developed to co-register TIR and RGB images that can benefit from the radiometric information of the former *and* the geometric quality of the latter. Importantly, IRT-3DDF methods exploiting combined TIR-RGB bundle block ad-

justments, where automatic correspondences *between* TIR and RGB images are required, have been under-researched and remain limited (Previtali et al., 2013). Multi-modal image matching, the process of establishing these correspondences between images acquired with different types of sensors, allow the direct and simultaneous orientation of both TIR and RGB imagery. Feature matching across domains (i.e., TIR-RGB) relies on detecting, describing, and matching features accounting for non-linear radial distortion (NRD) and markedly different image resolutions between image sources. Whilst challenging, the benefits of determining significant quantities of automatic multi-modal features presents a considerable benefit over current, manual or semi-automatic processes, allowing for thermal models to be generated with greater ease, automation and application.

5.1.1 Generation: Aims and Objectives

Expanding the results of Section 4.4.1, and the collaborative paper of Morelli et al. (2025b), this chapter presents a novel 2D-2D fusion approach for IRT-3DDF, simultaneously orienting TIR and RGB images in a combined TIR-RGB bundle block adjustment. The suitability of deep-learning (DL) multi-modal image matching for photogrammetric IRT-3DDF is assessed by comparing the performance of: (1) state-of-the-art deep-learning multi-modal matchers; (2) traditional, termed ‘hand-crafted’, feature matchers; and (3) an existing, popular, IRT-3DDF approach. Significantly, the proposed 2D-2D fusion pipeline offers an automatic and efficient workflow that eliminates the need for GCP-based co-registration and supports the use of thermal imaging for: daily and seasonally conditions, terrestrial and aerial use cases, and civil and heritage applications. This chapter will:

- Develop a novel 2D-2D fusion pipeline that enables the automatic co-registration of TIR and RGB image blocks within a combined bundle block adjustment, eliminating the need for GCPs, fixed relative sensor orientation or simultane-

ous data capture for IRT-3DDF;

- Evaluate the performance of the pre-trained DL neural networks for robust multi-modal image matching between TIR and RGB images, demonstrating that DL-based matchers, trained to be invariant to significant radiometric and geometric distortions not associated with multi-modal sensors, can generalise beyond their original mono-modal training domain;
- Provide an adaptable and scalable IRT-3DDF pipeline suitable for varying sensor formats, subjects and applications.

5.1.2 Chapter Structure

This chapter is structured as follows: Section 5.2 outlines the state-of-the-art in hand-crafted and deep-learning multi-modal image matching, including approaches specific to IRT-3DDF; Section 5.3 outlines the implemented methodology including multi-modal image matching, photogrammetric processing and thermal modelling; Section 5.4 compares the resulting camera poses and generated 3D models from two case studies on contemporary and historical building façades; and Section 5.5 discusses the development of the proposed workflow, its merits over existing approaches and future improvements.

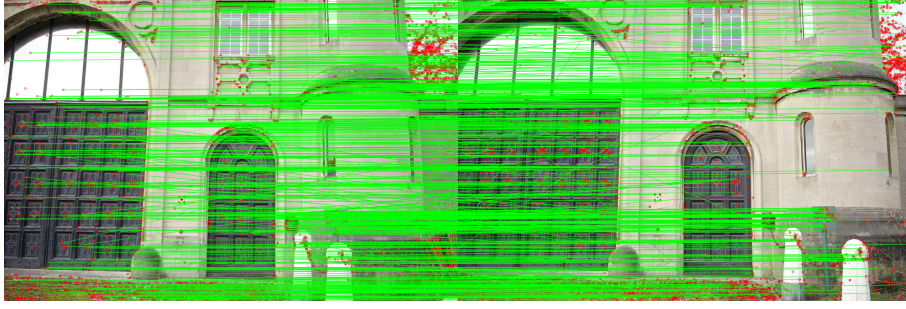
5.2 Multi-Modal Image Matching: State-of-the-Art

Image matching is a fundamental component of photogrammetry and computer vision, with traditional, termed ‘hand-crafted’, methods providing well-established algorithms for image registration, 3D reconstruction, object detection and local-

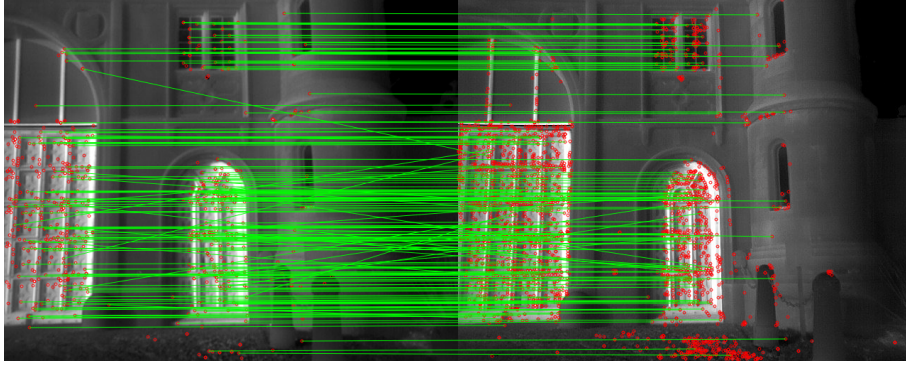
isation (Remondino et al., 2014). Notably, SIFT (Scale Invariant Feature Transform)(Lowe, 1999), despite its limited invariance to illumination and affine transformations, marked a significant milestone as the first detector and descriptor demonstrating sufficient invariance for a broad spectrum of applications (Lowe, 2004). Even today, SIFT remains the benchmark feature matcher among non-DL-based approaches (Jin et al., 2020), together with subsequent hand-crafted feature-based approaches such as SURF (Speeded-Up Robust Features)(Bay et al., 2006), ORB (Oriented FAST and rotated BRIEF)(Rublee et al., 2011) and FAST (Features from Accelerated Segment Test)(Rosten & Drummond, 2006).

However, whilst these algorithms are adequate for matching RGB-RGB or TIR-TIR image pairs (e.g., SIFT (Dahaghin et al., 2019), ASIFT (González-Aguilera et al., 2013) or FAST (Lagüela et al., 2012)), both intensity (detection) and gradient (description) information are sensitive to non-linear radiation distortion (NRD) *between* TIR and RGB images (Figure 5.1)(Saleem et al., 2017). This has led to the development of several hand-crafted multi-modal image matchers specifically designed to be insensitive to NRD and applicable across multiple image modalities. Among these, RIFT (Radiation Invariant Feature Transform)(Li et al., 2020) provides a SIFT-like multi-modal image matcher using frequency-domain information (phase congruency)(PC) and maximum index mapping (MIM) to detect and describe keypoints. Similarly, SRIF (Scale and Rotation Invariant Feature Transform)(Li et al., 2023a) integrates: (1) a pyramid scale space strategy to project key points within scale space, enabling scale invariance for each image, and (2) a local intensity binary transform (LIBT) to provide greater structural information for feature description. Whilst these hand-crafted multi-modal image matchers provide suitable thermal-specific methods for IRT-3DDF (Lin et al., 2019b), their ability to provide robust multi-view matches across challenging scales, poses and baselines requires assessment (Sledz & Heipke, 2022).

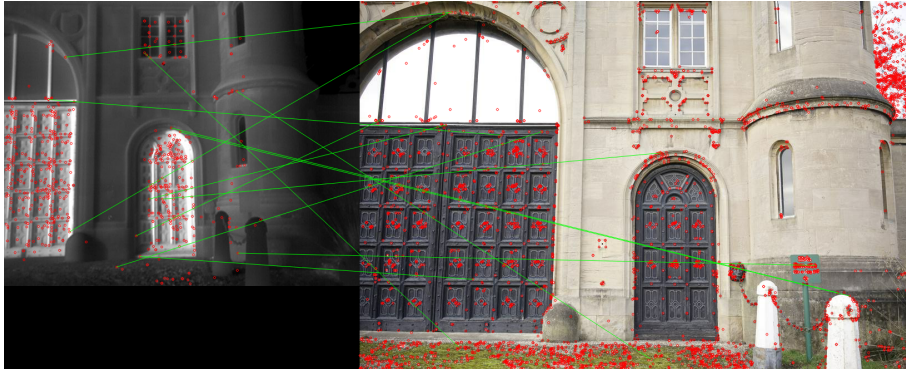
In recent years, the proliferation of DL approaches for image matching has sought



(a) RGB-RGB (724)



(b) TIR-TIR (192)



(c) TIR-RGB (13)

Figure 5.1: Comparing a hand-crafted image matching approach (SIFT feature detector (Lowe, 1999), nearest neighbour matching (Riba et al., 2019) and PyDEGENSAC geometric verification (Mishkin et al., 2015)) for (a) RGB-RGB, (b) TIR-TIR, and (c) TIR-RGB image pairs, with the number of matched features being 724, 192, and 13, respectively.

to overcome the limitations of hand-crafted algorithms by using CNNs to build new detectors and descriptors, commencing with the seminal work of LIFT (Yi et al., 2016), followed by other end-to-end DL approaches such as LF-Net (Ono et al., 2018), SuperPoint (DeTone et al., 2018), and ASLFeat (Luo et al., 2020). In addition, feature matching, after local feature detection, can be trained in isolation

using graph neural networks as a substitute to nearest neighbour matching and epipolar geometry verification, with SuperGlue (Sarlin et al., 2019) and LightGlue (Lindemberger et al., 2023) presenting robust and versatile matching solutions for RGB images. Conceptually, any combination of hand-crafted and learning-based components (i.e., detector, descriptor, matcher) is feasible for the determination of correspondences (Jin et al., 2020). However, these networks are pre-trained on extensive mono-modal datasets with varying illumination, perspective changes and wide baselines (Barroso-Laguna et al., 2019; Mishchuk et al., 2017; Verdie et al., 2014); which, when used for multi-modal image matching, comprising images with inherent geometric, radiometric and spectral differences, are applied beyond their original mono-modal training domain.

For multi-modal image matching, the ability to overcome these challenges has justified their exploration for thermal-visible image matching (Velesaca et al., 2024). PhotoMatch (Ruiz De Oña et al., 2023), a multi-modal image matching software, demonstrates the application of ‘out-of-the-box’ DL image matching methods with ‘detect-and-describe’ approaches like D2-Net (Dusmanu et al., 2019) and R2D2 (Revaud et al., 2019), both superior to hand-crafted methods in six multi-modal scenarios (inc., scanning electron microscopy, infrared, LiDAR, aerial, and magnetic resonance images). Similarly, Chang et al. (2024) presents a novel lightweight network built to exploit local and global feature information for infrared-visible image matching, achieving competitive results with RIFT, R2D2, SuperPoint+LightGlue (Sp+Lg) and LoFTR (Sun et al., 2021). These cases demonstrate the success of ‘out-of-the-box’ DL feature detectors and matchers for 2D scenarios; however, the adaptability of DL neural networks for 2D-2D fusion remains open for investigation.

The use of 2D-2D fusion approaches, previously discussed in Section 2.3.2, point to several conclusions for multi-modal image matching: (1) hand-crafted features, although successful mono-modally (Maset et al., 2017), are unsuitable for comprehensive multi-modal image matching (Ham & Golparvar-Fard, 2013a; Hoegner et

al., 2018); (2) image orientation through combined bundle block adjustments, utilising multi-modal correspondence, are suitable, performed either manually (Scaioni et al., 2017), semi- (Previtali et al., 2013) or fully-automatically (Elias et al., 2023); and (3) DL approaches are the logical evolution of 2D-2D fusion. With increasing access, availability and applications for DL image matchers, the ability for multi-sensor correspondence to be used within single bundle adjustment has become well-established, utilised for 2.5D TLS- (Morelli et al., 2024b), historical- (Morelli et al., 2022) and day-and-night (Morelli et al., 2025a) images. However, the development of methods specific to TIR images, appreciating the unique characteristics of infrared radiation, specific use cases and desired applications, is required for the complete assessment of multi-modal image matching for IRT-3DDF.

5.3 Methodology

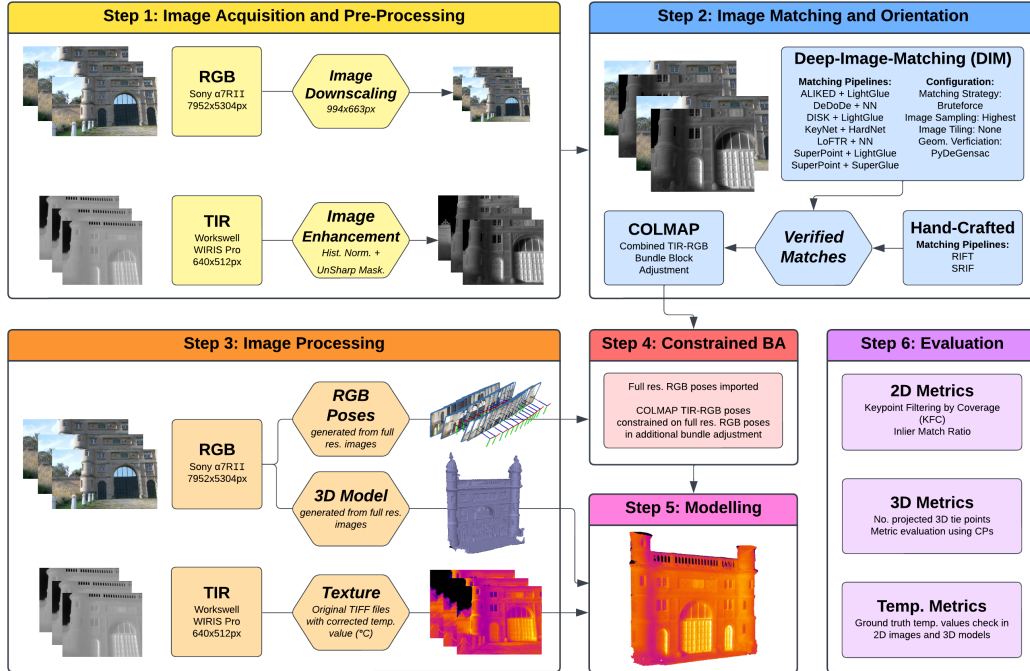


Figure 5.2: Proposed 2D-2D fusion methodology utilising multi-modal image matching for IRT-3DDF.

The outlined methodology introduces DL-based multi-modal image matching for the generation of IRT-3DDF. Importantly, this proposed method co-registers both image blocks through shared observations of local features in mono-modal (i.e., TIR-TIR, RGB-RGB) *and* multi-modal (i.e., TIR-RGB) image pairs. Whilst traditional methods rely on the co-registration of independent TIR and RGB blocks using GCPs/ICP, fixed relative orientation between sensors, or GNSS/IMU data, the proposed method is fully automatic. In addition, an optimised modelling and texturing method is presented to exploit the benefits of both individual datasets for 3D thermal modelling. Importantly, the proposed IRT-3DDF methods enable asynchronous acquisition of TIR and RGB datasets, particularly advantageous in settings where optimal conditions for data collection (e.g., lighting, temperature and desired application) do not coincide, allowing additional TIR surveys to be collected and compared over time or under different experimental conditions. The proposed workflow, outlined in Figure 5.2, comprises the following steps:

1. **Data Acquisition and Pre-Processing:** Asynchronous imaging surveys using separate TIR and RGB sensors are undertaken, followed by pre-processing to enhance image features prior to multi-modal image matching.
2. **Image Matching and Orientation:** Exhaustive matching is performed between TIR and RGB blocks using a collection of hand-crafted and pre-trained DL-based approaches, subsequently oriented in a combined bundle block adjustment.
3. **Image Processing:** Full-resolution RGB images are used to generate camera poses and 3D polygonal model through a classical photogrammetric workflow. In addition, 16-bit TIR images, encoding accurate temperature values, are prepared for eventual 3D model texturing.
4. **Optimised Bundle Adjustment:** Full-resolution RGB poses are imported as reference poses, with an additional bundle adjustment undertaken con-

straining the hand-crafted and learning-based TIR-RGB blocks on the full-resolution RGB poses.

5. **Modelling:** Integration of optimised TIR-RGB blocks, 3D polygonal model and temperature corrected TIR images for 3D thermal modelling.
6. **Evaluation:** Evaluation of hand-crafted and DL multi-modal tie points (2D Metrics), pre- and post- optimisation TIR-RGB blocks in object space (3D Metrics) and derived 3D thermal model temperature values (Temp. Metrics).

5.3.1 Case Studies

To evaluate DL-based methods for multi-modal image matching, two contrasting architectural case studies were utilised. Firstly, to assess the implementation of IRT-3DDF for energy efficiency and retrofitting applications, a 21st Century contemporary dwelling in Warwickshire, UK, named Dunrobin House (DUNR) was selected (Appendix A.3). Secondly, Lenton Lodge Gatehouse (GATE) was used as a heritage case study in Nottingham, UK (Appendix A.1). This provides two case studies of contrasting scale, materials and age for the assessment of the proposed IRT-3DDF workflow. Both buildings can be seen in Figure 5.3.



(a) DUNR



(b) GATE

Figure 5.3: (a) Dunrobin House, Warwickshire, UK; (b) Lenton Lodge Gatehouse, Nottinghamshire, UK.

5.3.2 Step 1: Image Acquisition and Pre-Processing

Cameras

For the execution of the proposed method, the same cameras from Chapter 4 were employed: the Workswell WIRIS Pro (WWP) and Sony α 7R II (RGB1) (Section 4.3.1). Although the WWP features both TIR and VIS sensors, the limited resolution of the VIS sensor (1920×1080 pix) required an alternative camera for the effective generation of the base geometry. Therefore, RGB1, with a full-frame sensor and a resolution of 7952×5304 pix, was used to capture higher quality RGB images. As with Section 4.4.1, the WWP and RGB1 were not fixed in relative orientation, meaning the proposed method cannot assume a constant baseline between the two cameras. Prior to image capture, both the WWP and RGB1 were geometrically calibrated utilising the method in Section 4.3.3), obtaining the following IO values: f , c_x , c_y , k_1 , k_2 , p_1 , p_2 . In addition, implementing the findings of Section 3.4.2, the WWP underwent 75-minute sensor acclimatisation period to generate stable temperature readings (increased from 52 minutes to account for colder environmental conditions), with NUCs set at 3-minute intervals and periodic image capture undertaken every 12 seconds.

Data Acquisition

In line with the recommendations of Section 3.5, as well as protocols for effective multi-modal image matching, the following considerations were made for both thermal and visible data capture. For both WWP datasets, passive thermography was used, with ambient air temperature and reflected temperature recorded using the RS PRO RS-91 hygrometer and RS PRO RS-8876 infrared thermometer, input into the WWP prior to image capture. In addition, emissivity values (ε) for the predominant building material in each case study were determined and set within

the WWP. For Lenton Lodge Gatehouse, facing East-South-East, data acquisition between 09:30-12:00 presented conditions that would maximise morning ‘baking’ of the structure and excite a thermal signature. In contrast, TIR images of Dunrobin House were acquired at night (22:18-22:25) to determine a thermal signature of the building when no internal or external stimuli were present (representing preferable conditions for building energy audits)(Sadhukhan et al., 2020). To mitigate the low resolution of the TIR images, strips were captured with 90% horizontal and 70% vertical overlap (Figure 5.4). For both structures, a network of thermal-specific markers (GCPs and CPs) was acquired to validate the accuracy of the proposed pipeline (Section 3.5.4). For Dunrobin house, a fixed TIR GSD of 9.8 mm was used for all images, whilst for Lenton Lodge Gatehouse strips at different distance were captured, with GSD ranging between 9.8 mm and 26.2 mm. Details of both surveys can be seen in Table 5.1, with the resulting camera networks visualised in Figure 5.5.



(a) TIR



(b) RGB

Figure 5.4: Example image sequences for (a) TIR and (b) RGB images of Lenton Lodge Gatehouse.

Image Enhancement

The ‘appearance’ of TIR images differs significantly from images in the visible spectrum, with TIR images often characterised by low contrast, high noise and blurred details. The dynamic range of thermal images is often determined by areas of

Statistics	Case Study	
	Dunrobin House	Lenton Lodge Gatehouse
No. TIR Images	45	90
TIR res. (pix)	640×512	640×512
Min. TIR GSD (mm)	9.8	9.8
Max. TIR GSD (mm)		26.2
No. RGB Images	50	55
Full RGB res. (pix)	7952×5304	7952×5304
Full RGB Min. GSD (mm)	1.0	1.0
Full RGB Max. GSD (mm)	1.3	2.6
Downsampled RGB res. (pix)	994×663	994×663
Downsampled Min. RGB GSD (mm)	8.0	8.0
Downsampled Max. RGB GSD (mm)	10.4	20.8
No. Image Pairs (Bruteforce)	4,465	10,440
No. GCPs	3	4
No. CPs	3	4

Table 5.1: Data Acquisition statistics for both TIR and RGB data sets.

significant temperature change (i.e., sea, sky), reducing the visibility of a subject presenting relatively smaller temperature changes. Therefore, thermal image enhancement has become an fundamental pre-processing step for image registration, object detection and 3D reconstruction (Akcaay & Avsar, 2017; Conte et al., 2018; Ribeiro-Gomes et al., 2017; Zhang et al., 2023b). Unlike Section 4.4.2, the underlying assumption of the proposed method, given its reliance on multi-modal matches, is the presence of sufficient texture in both the TIR and RGB datasets, successfully determining local features in the images that are repeatable, unique and distinctive.

To improve the number of detectable features between TIR and RGB images prior to image matching, the WWP’s 16-bit TIFF images were imported into FIJI (FIJI Is Just ImageJ)(Schindelin et al., 2012), independently contrast stretched and enhanced with an UnSharp Masking filter, significantly increasing the number of matches extracted by both hand-crafted and DL-based matching strategies (Figure 5.6). Subsequently, TIR images were converted from 16-bit to 8-bit files for compatibility with subsequent image matching processes. To facilitate effective

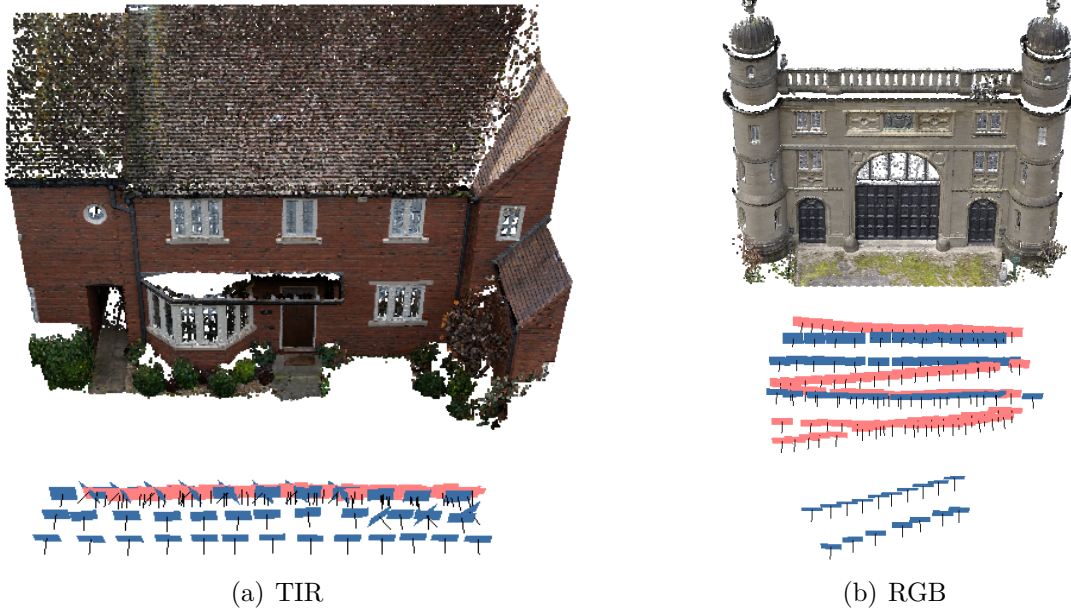


Figure 5.5: TIR (red) and RGB (blue) camera networks for (a) Dunrobin House and (b) Lenton Lodge Gatehouse.

multi-modal matching, RGB images were downsampled from their original resolution (7952×5304 pix) to $\frac{1}{8}$ of their original size (994×663 pix) (Adamopoulos et al., 2021) (notably, none of the tested DL approaches found sufficient key points by directly matching TIR and full-resolution RGB images before resampling, showing a limited scale invariance of the DL local features tested)(Marelli et al., 2023). Importantly, the original 16-bit TIFF images, encoding temperature values as DN_s, were re-introduced in subsequent texturing steps to include accurate temperature measurements for the 3D thermal models.

5.3.3 Step 2: Image Matching and Orientation

Deep-Image-Matching

Deep-Image-Matching (DIM), a flexible, open-source Python library designed for robust multi-view image matching, was used for the implementation of the multi-modal image matching (Morelli et al., 2024a). DIM enables front-end access to

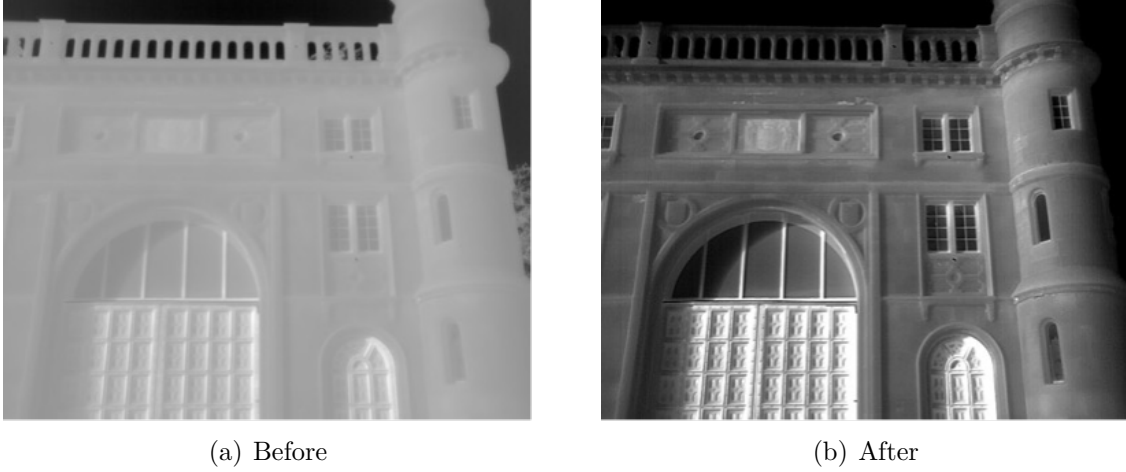


Figure 5.6: 16-bit TIFF thermal images (a) before and (b) after image enhancement (histogram normalization and UnSharp Masking)

a wide range of state-of-the-art computer vision algorithms, providing the essential foundation for bundle adjustments and accurate photogrammetric reconstruction. DIM employs a modular workflow designed for effective multi-view matching, including pipeline selection, image pair selection (matching strategy) and image pre-processing (Figure 5.7). Importantly, DIM does not perform either the bundle adjustment or 3D scene reconstruction itself, but exports matches in formats compatible with various photogrammetric software. DIM supports various combinations of local features and matching algorithms, spanning traditional hand-crafted-to recent state-of-the-art learning-based approaches, including: SIFT (Lowe, 1999), ORB (Rublee et al., 2011), ALIKE (Zhao et al., 2021), ALIKED (Zhao et al., 2023), DeDoDe (Edstedt et al., 2023a), DISK (Tyszkiewicz et al., 2020), Key.Net (Barroso-Laguna et al., 2019)+HardNet8 (Pultar, 2020), XFeat (Potje et al., 2024), LightGlue (Lindenberger et al., 2023), LoFTR (Sun et al., 2021), SE2-LoFTR (Böckman & Kahl, 2022), RoMA (Edstedt et al., 2023b), SuperPoint (DeTone et al., 2018) and SuperGlue (Sarlin et al., 2019).

For the comprehensive assessment of pre-trained ‘out-of-the-box’ image matchers for IRT-3DDF, various hand-crafted and DL-based matchers were tested to compare their respective performances. The following DL features and matchers were con-

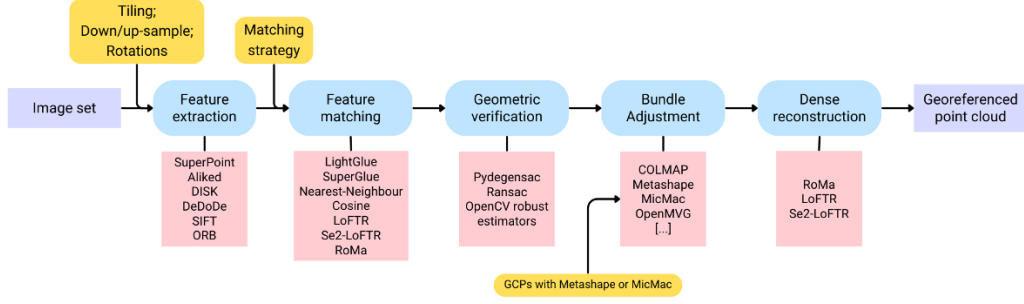


Figure 5.7: Deep-Image-Matching (DIM) modules and workflow.

sidered: (1) SuperPoint (Sp) features combined with SuperGlue (Sg) matcher; (2) SuperPoint, ALIKED and DISK features combined with LightGlue (Lg) matcher; (3) DeDoDe features and KeyNet detector with HardNet descriptor, combined with nearest neighbour (NN) matching. In addition, the comparison was extended to include the use of LoFTR, a semi-dense method with a featureless approach designed to operate on image pairs. As a hand-crafted method for mono-modal matching, SIFT+NN matching was tested, which is considered the benchmark among traditional matching algorithms (Jin et al., 2020). The DIM ‘bruteforce’ matching strategy was used to exhaustively match all images in the combined TIR-RGB image blocks. This resulted in 4,465 and 10,440 pairs for Dunrobin House and Lenton Lodge Gatehouse, respectively (Table 5.1). For each pipeline, a maximum of 8,000 keypoints was extracted from each image, a quantity commonly used in local features and matchers accuracy assessment (Jin et al., 2020). Additionally, DIM settings remained consistent across each pipeline, including: no tiling, highest quality image sampling, local features and matchers parameters, and DEGENSAC (Chum et al., 2005) geometric verification error threshold (4 pix). The mono-modal and multi-modal tie points extracted in DIM are directly written to an SQLite database, ready for running through COLMAP’s (Schonberger & Frahm, 2016) incremental reconstruction pipeline.

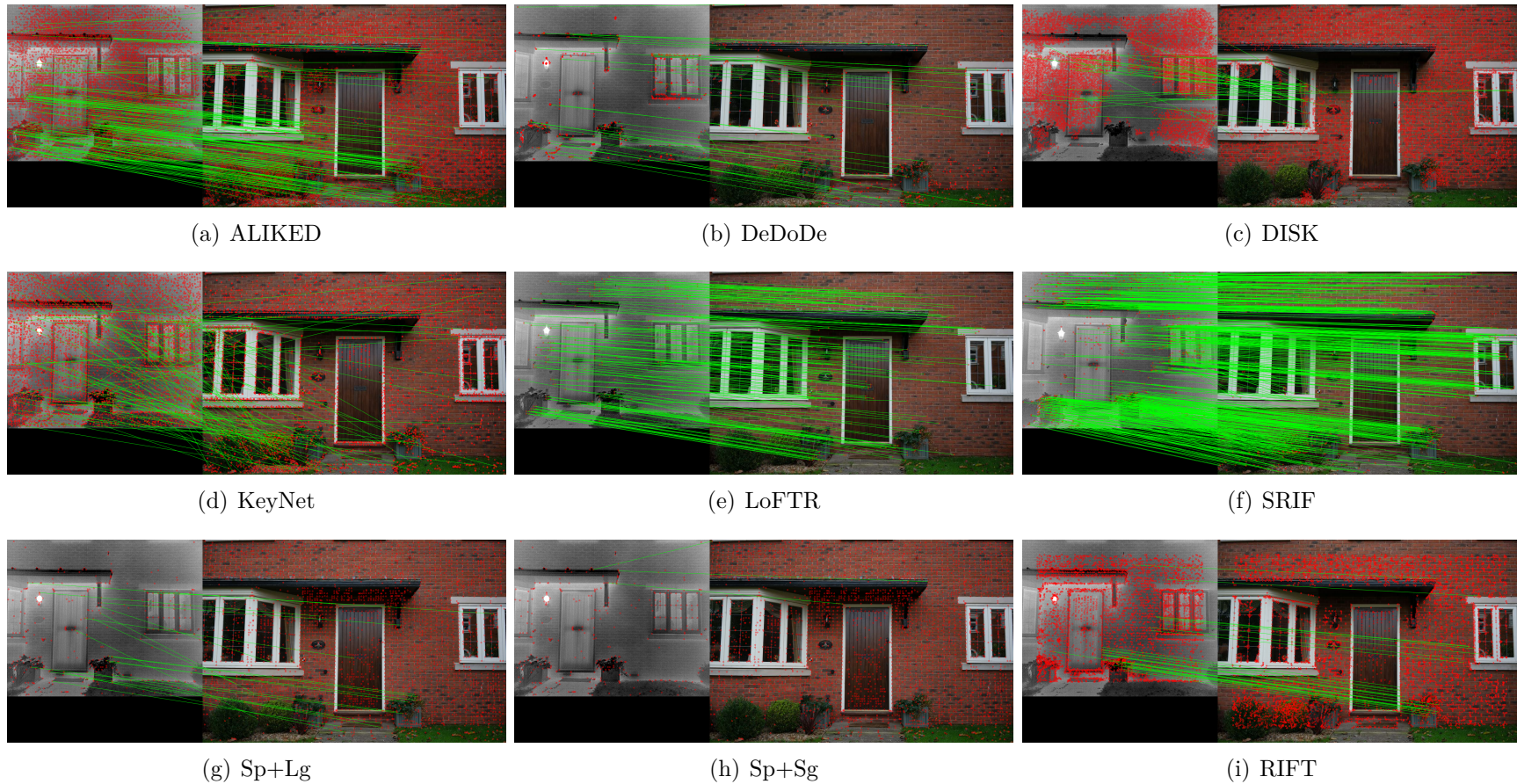


Figure 5.8: Multi-modal image matching results for a TIR-RGB pair of Dunrobin House using both DL and hand-crafted matchers.

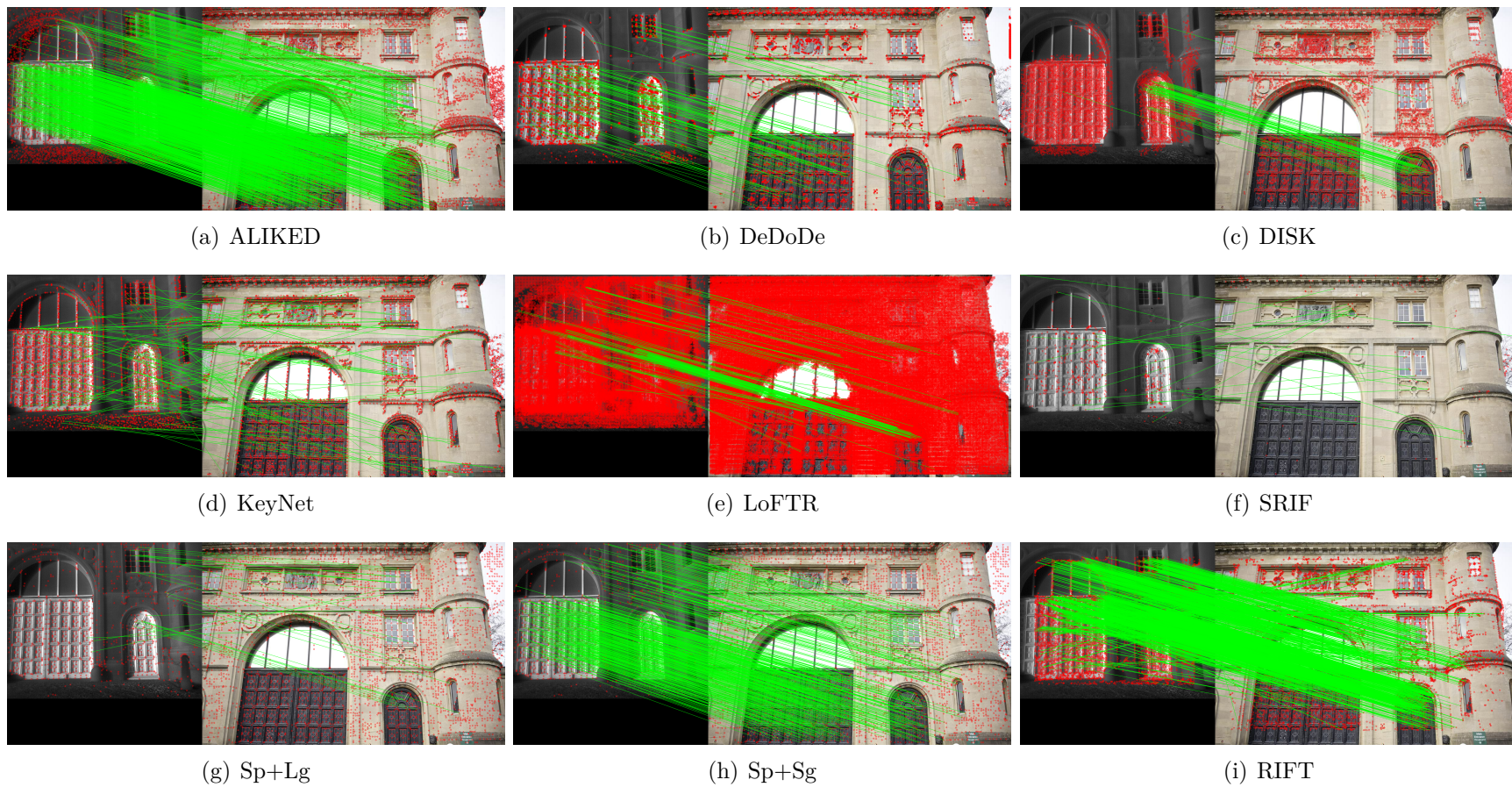


Figure 5.9: Multi-modal image matching results for a TIR-RGB pair of Lenton Lodge Gatehouse using both DL and hand-crafted matchers.

Hand-Crafted Multi-Modal Processing (RIFT)

To provide a benchmark for comparison of DL matching algorithms, two hand-crafted matchers specifically designed for multi-modal images were chosen: RIFT2 and SRIF. RIFT2 (Li et al., 2023b) reduces the cost of the RIFT’s (Li et al., 2020) original MIM convolution sequence rings for rotation invariance by introducing a dominant index into MIM, reducing processing time. Initial experiments were undertaken on RIFT’s three main parameters: (1) the number of convolution orientations of the log-Gabor filter (N_O); (2) the number of scales at which the log-Gabor filter is generated (N_S); and (3) the patch size of the feature descriptor (J). As iterations of these values produced similar results across several TIR-RGB pairs, and as no values within any DL approaches were pre-tuned, initial values of 6 (N_O), 4 (N_S) and 96 (J) were used. All features and geometrically-verified matches were exported from MATLAB and subsequently imported into COLMAP.

In addition, SRIF refines previous multi-modal image matchers like RIFT and LNIFT (Locally Normalized Image Feature Transform)(Li et al., 2022) by providing scale and orientation invariance across images. Specifically, SRIF introduces two components: (1) a pyramid scale space strategy to project key points across images in scale space; and (2) LIBT for feature description to provide greater structural information across and within images. Feature descriptor, much like RIFT, is subsequently completed in a SIFT-like vector fashion. For SRIF, stock configurations settings were implemented for the patch size ($J=128$), grid size ($N_{\text{grid}}=8$) and histogram bins ($N_{\text{bin}}=4$) with geometric verification set at 4 pix. Similarly, RIFT features, generated in DIM, were exported in a database compatible with COLMAP.

Combined TIR-RGB Bundle Block Adjustment

The geometrically verified mono-modal and multi-modal tie points were imported into COLMAP for the combined TIR-RGB bundle block adjustment (CBBA). For each subset of TIR and RGB images, a distinct camera model was assigned, using initial values for self-calibration (Section 5.3.2). Images were incrementally oriented using the resection-intersection approach to generate sparse 3D point clouds for each matching procedure. At this stage, KeyNet and RIFT were deemed to have too few suitable multi-modal matches for image orientation, with TIR and RGB images unable to be oriented conjointly. Resulting tie points with their camera poses, 2D and 3D coordinates were then exported using the Bundler format (Snavely et al., 2006) to Agisoft Metashape (Agisoft Metashape, 2024) for: referencing using GCPs, the optimised bundle adjustment process, thermal texturing and metric evaluation against CPs.

5.3.4 Step 3 and 4: Image Processing and Optimised Bundle Adjustment

Currently, the proposed method as seen in Section 4.4.1, orienting TIR and RGB images primed for multi-modal image matching (Section 5.3.2), does not exploit the full-resolution RGB1 images for the extraction of highest quality tie points (7953×5304 vs. 994×663 pix). Therefore, to implement a data fusion method that maximises the potential of both datasets, RGB camera poses were estimated using the original full-resolution RGB images in a separate bundle adjustment. For each multi-modal matching method, full resolution RGB poses were introduced as ‘known reference’, constrained ($\sigma_{xyz}=1$ mm; $\sigma_{\omega\phi\kappa}=0.1$ deg) and optimised in an additional TIR-RGB bundle block adjustment in Agisoft Metashape. This refines the position and orientation of the TIR-RGB blocks, allowing the TIR image blocks

to benefit from the information of the full-resolution RGB images. The benefits of re-introducing the high-resolution images back into the proposed optimisation include: (1) generating features with greater information, distinction and quantity; (2) providing high-quality matches for both TIR and RGB pose refinement; and (3) using higher-resolution imagery for the generation of an accurate 3D polygonal model through MVS. Similarly, ‘enhanced’ thermal images, utilised to improve multi-modal image matching, have lost the encoded temperature values during contrast stretching (Section 5.3.2). Therefore, the original 16-bit TIFF files, encoding the accurate temperature information of both datasets, were re-introduced for texturing (Adán et al., 2021; Patrucco et al., 2020; Previtali et al., 2012).

5.3.5 Step 5: Thermal Modelling

Thermal Texturing

Once the poses of the TIR-RGB blocks have been optimised using the full resolution RGB images, the process of thermal texturing can begin. The refined camera poses, polygonal model generated from full resolution RGB images and corrected TIR images were combined to create a 3D textured thermal model. To build the thermal textures, Agisoft Metashape’s ‘mosaic’ texturing approach was used, commonly employed for thermal 3D modelling and orthomosaicking (Grechi et al., 2021; Malbêteau et al., 2021; Zhao et al., 2018). Here, greater statistical weight is given to central image pixel values (deemed ‘low frequency’ information) and images aligned with façade normals (‘high frequency’ information), meaning pixels perpendicular to the facades, representing the angle of incidence of emitted infrared radiation, assign their temperature values (Agisoft Metashape, 2024). The final product is a geometrically- and radiometrically-accurate 3D thermal model.

5.3.6 Step 6: Evaluation

As highlighted in Jin et al. (2020) (and that of Chapter 4), evaluation of SfM applications, be it mono- or multi-modal, should occur downstream of the entire SfM pipeline, not relying solely on image-specific metrics (e.g., no. features, no. matches, no. tie points). Therefore, whilst Morelli et al. (2025b) expands the performance of each learning-based image matcher through specific image metrics, the efficacy of multi-modal image matching for IRT-3DDF is evaluated through two approaches: (1) camera pose- and (2) temperature metrics. This allows for focus to remain on the two critical components for IRT-3DDF: (1) determining accurate camera poses for orientation, modelling and texturing; and (2) deriving accurate temperature measurements for quantitative analysis, respectively. Additionally, as the comparison of IRT-3DDF methods has proven helpful in demonstrating the strengths and weaknesses of individual methods (Adamopoulos et al., 2020b; Hoegner et al., 2016; Javadnejad et al., 2020; Patrucco et al., 2022b), a common 3D-3D fusion method, co-orienting individual TIR and RGB image blocks utilising GCPs manually-identified in both image modalities, was undertaken for comparison with the proposed 2D-2D fusion method.

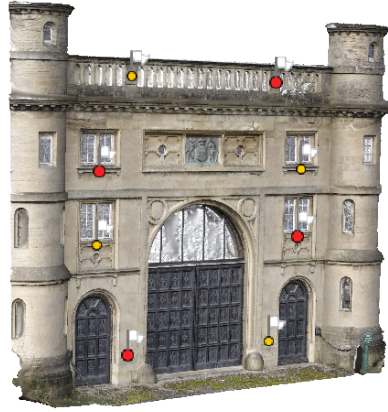
Camera Pose Metrics

For the evaluation of both case studies, the oriented TIR-RGB blocks and 3D tie points were imported into Agisoft Metashape to triangulate GCPs and CPs (Figure 5.10). This includes the assessments of the *original* combined TIR-RGB bundle block adjustment (Section 5.3.3) and the block *optimised* with the full resolution RGB images (Section 5.3.4). As highlighted in Table 5.1, GCPs and CPs were split for both datasets, with an even distribution of targets across both building façades. To guarantee the best accuracy in positioning, the projections of the targets in RGB images were marked using the full resolution images, subsequently

downscaled to match the $\frac{1}{8}^{\text{th}}$ resolution of the images processed in Section 5.3.3. The RMSE on the CPs was employed to find the error of each individual modality (i.e., RMSE_{TIR} , RMSE_{RGB}) and determine the orientation performance of both modalities separately. Finally, the inlier match ratio, representing the ratio between successfully triangulated tie points and the number of raw matches exported from COLMAP, was included to provide a proven image-specific metric for multi-modal image matching (Morelli et al., 2022).



(a) DUNR



(b) GATE

Figure 5.10: Distribution of GCPs (red) and CPs (yellow) on the 3D polygonal models of (a) Dunrobin House and (b) Lenton Lodge Gatehouse.

Temperature Metrics

As the validation of derived temperature values within 3D thermal models remains a critical metric for IRT-3DDF (Adán et al., 2023b; Lin et al., 2019b), a comparison between obtained infrared thermometer- (T_{obs}), 2D pixel- ($T_{2\text{D}}$) and 3D texel ($T_{3\text{D}}$) temperature values was undertaken (Sections 3.3.3). For each dataset, multiple temperature readings were observed using the RS PRO RS-8876 infrared thermometer (Section 3.5.4). Temperature readings were taken at the beginning of both surveys adjacent to visible survey markers (GCPs/CPs), enabling the temperature of the predominant building material to be assessed. In addition, 2D pixel

values were extracted from the first instance of the thermal-specific marker in the imaging surveys where the marker is perpendicular with the WWP’s optical axis. This enables the comparison of singular pixel values in 2D images to those blended during the texturing process in the derived 3D model (Lin et al., 2019b).

5.4 Results

The following section assess the accuracy of the proposed methodology 2D-2D fusion method with different hand-crafted and DL-based matchers across the two datasets. This evaluation is undertaken using both camera pose and temperature metrics, with the former used to characterise image matching and orientation performance, and the latter comparing the accuracy of derived temperature measurements. The obtained results are compared with an additional thermal modelling method, a 3D-3D fusion method co-registering independently-generated TIR and RGB point clouds using GCPs, to assess the performance of the proposed pipeline against a common IRT-3DDF approach. Firstly, it must be noted that not all approaches were able to orient the TIR-RGB blocks for both datasets, meaning they were excluded from the evaluation using the camera poses and temperature metrics. This included KeyNet+HardNet and SIFT and SRIF, which failed due to insufficient matches across modalities (Figure 5.8 and 5.9). Notably, although SRIF detects a good number of matches on Dunrobin House (in contrast, the matches extracted for Lenton Lodge Gatehouse are very limited), they are not triangulated successfully in COLMAP, preventing the resection of TIR images on the RGB block.

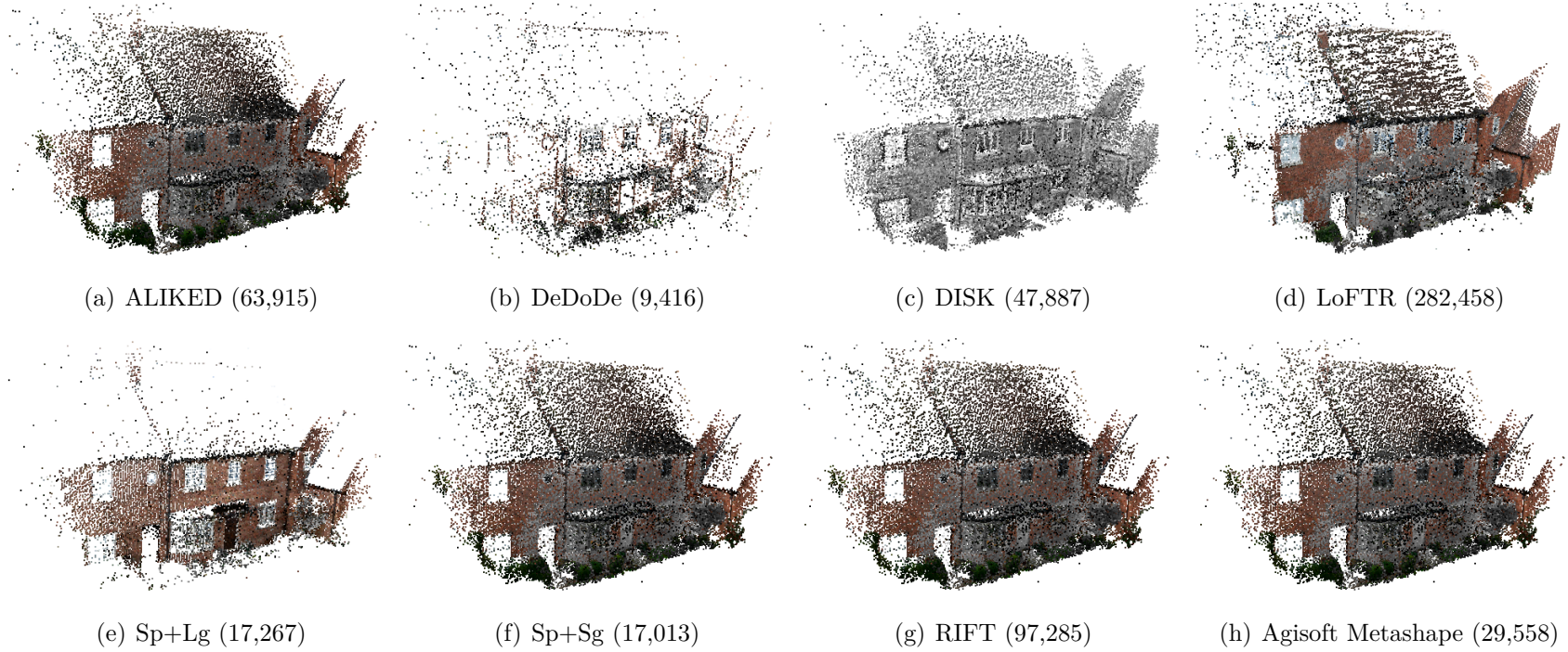


Figure 5.11: Generated 3D tie points for Dunrobin House, featuring the number of tie points generated for each multi-modal image matching approach. Results of the Agisoft Metashape 3D-3D fusion method have also been included for comparison.

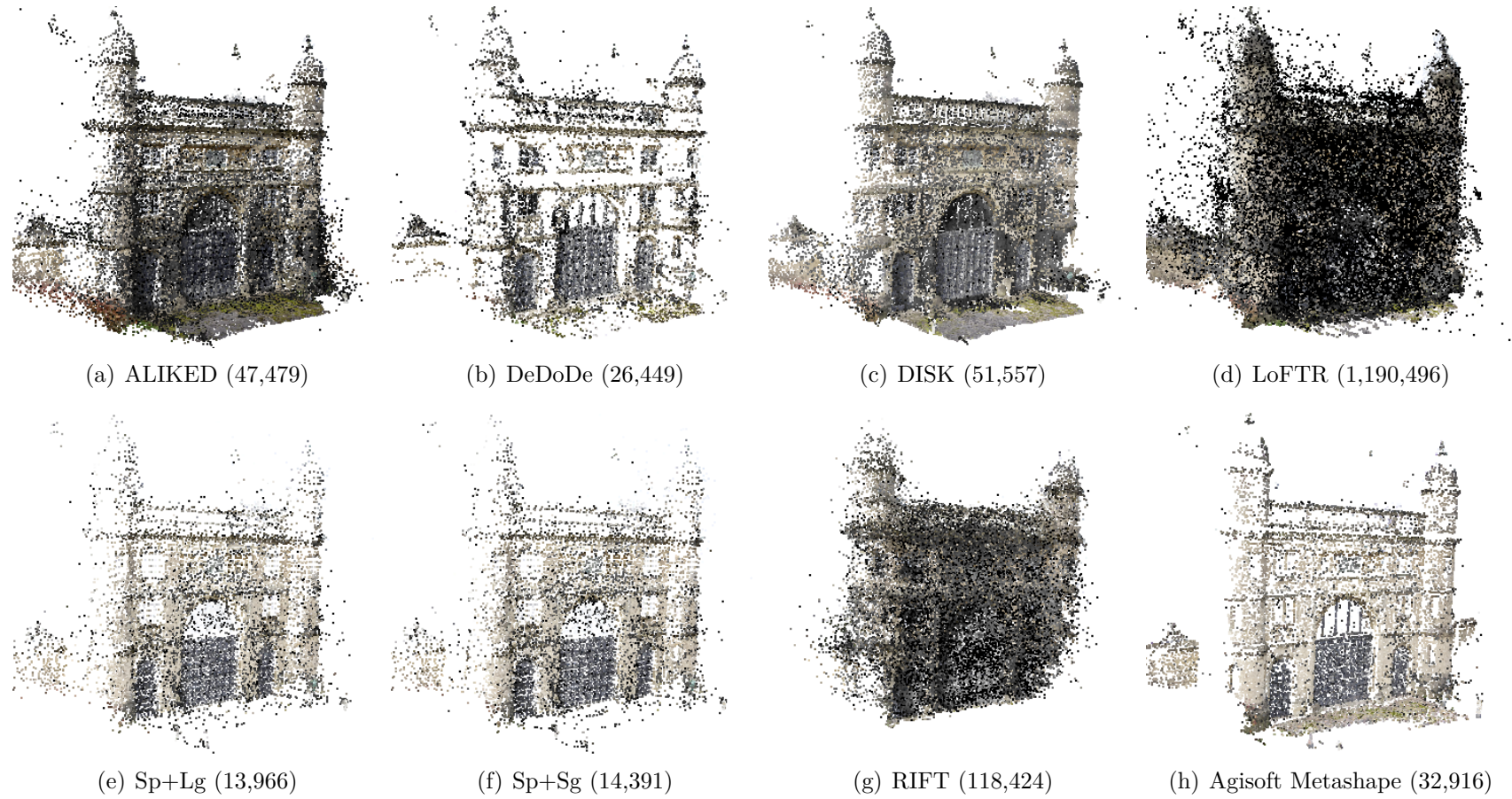


Figure 5.12: Generated 3D tie points for Lenton Lodge Gatehouse, featuring the number of tie points generated for each multi-modal image matching approach. Results of the Agisoft Metashape 3D-3D fusion method have also been included for comparison.

5.4.1 Camera Pose Metrics

The results of the camera pose metrics are separated into two sub-sections, addressing: (1) the *original* combined TIR-RGB bundle block adjustments (Section 5.3.3), and (2) the *optimised* bundle adjustment re-introducing full resolution RGB poses (Section 5.3.4). For both buildings, the RMSE on CPs is split into the individual image modalities: RMSE_{TIR} and RMSE_{RGB} . Furthermore, the number of 3D tie points, the inlier match ratio, and MRE for the *entire* block are provided as additional image-specific metrics. The results of the camera pose metrics are summarised in Tables 5.2, 5.3 5.4 & 5.5. The 3D tie points of each matching variant, triangulated during the incremental orientation of the multi-modal blocks, are shown in Figures 5.11 and 5.12.

Combined TIR-RGB Bundle Block Adjustment

For Dunrobin House, the results of the camera pose metrics for the multi-modal image matchers and CBBA show favourable results, both in terms of orientation and accuracy measures. All DL matchers achieve an inlier match ratio $>72\%$, meaning $\sim 7/10$ matched 2D features result in 3D tie points. Across the learning-based approaches, RMSE_{RGB} ranges from 1.6-3.2 mm, an optimal result considering the GSD of the downsampled RGB images (8.0-10.4 mm). This reiterates the suitability of DL matchers for traditional RGB imagery, despite their low resolution or the presence of other modalities as part of the BA (Jin et al., 2020; Ma et al., 2021; Morelli et al., 2024a; Ruiz De Oña et al., 2023). Expectedly, RMSE_{TIR} , with a GSD for TIR images of 9.8 mm, is an order of magnitude higher, ranging from 5.7-14.1 mm. LoFTR, as a semi dense matcher, expectedly produced more tie points than any other approach ($\sim 4\times$) and an adequate RMSE_{TIR} of 13 mm. The results from Dunrobin House demonstrate that with a favourable camera network and appropriate GSD, DL feature matching algorithms can handle multi-modal

image matching for close-range building façades. The best of the DL approaches, SuperPoint with both SuperGlue or LightGlue, produces similar results in terms of MRE (0.3 pix) and RMSE. As LightGlue is an evolution of SuperGlue, designed to improve computational efficiency, the only noticeable difference between them is the execution time (Table 5.7).

Similarly, the results of the CBBA for Lenton Lodge Gatehouse show the success of the 2D-2D fusion pipeline. Importantly, Lenton Lodge Gatehouse’s camera network captured images over increasing distances, with resulting GSDs ranging for both TIR and downsampled RGB images from 9.8-26.2 mm and 8.0-20.8 mm, respectively. For all DL approaches, combined TIR-RGB bundle block adjustments achieve RMSE_{RGB} between 6.3-34.2 mm, and $\text{RMSE}_{\text{TIR}} \sim < 20$ mm. ALIKED, DISK and Sp+Lg performed favourably, with a RMSE_{TIR} of 10.8, 15.5 and 15.2 mm, respectively. Interestingly, ALIKED and DISK produce comparatively more tie points than other DL approaches, a pattern observable across both datasets and supported by their remarkable high match ratio (91% and 92% for Lenton Lodge Gatehouse). This measure, as well as the density of matched features across individual images, points to the strength of these approaches as multi-modal image matchers, even without being trained on such data (Morelli et al., 2025b). For Lenton Lodge Gatehouse, LoFTR exhibits a behaviour distinct from the previous dataset, producing a particularly noisy semi-dense point cloud and high RMSE_{RGB} (34.2 mm). This drop in performance is also reflected in its very low inlier match ratio (21%). However, LoFTR, being a semi-dense matcher, triangulates a substantially higher number of 3D tie points compared to the other methods.

CBBA	Local Feature	Matcher	No. 3D Tie Points	Inlier Match Ratio	MRE (pix)	RMSE _{TIR} (mm)	RMSE _{RGB} (mm)
COLMAP	ALIKED	LightGlue	64,296	0.83	1.0	14.1	1.9
	DeDoDe	Nearest Neighbour	9,524	0.75	0.7	8.8	3.1
	DISK	LightGlue	47,964	0.75	1.2	11.1	3.2
	SuperPoint	LightGlue	17,295	0.84	0.3	8.4	1.7
	SuperPoint	SuperGlue	17,040	0.84	0.3	5.7	1.7
	LoFTR		282,504	0.72	0.7	13.5	1.6
	RIFT2		46,800	0.22	0.9	7.7	13.3

Table 5.2: Combined TIR-RGB bundle block adjustment for Dunrobin House.

CBBA	Local Feature	Matcher	No. 3D Tie Points	Inlier Match Ratio	MRE (pix)	RMSE _{TIR} (mm)	RMSE _{RGB} (mm)
COLMAP	ALIKED	LightGlue	47,882	0.91	0.4	10.8	6.3
	DeDoDe	Nearest Neighbour	27,175	0.74	0.4	21.4	4.7
	DISK	LightGlue	51,852	0.92	0.5	15.5	7.2
	SuperPoint	LightGlue	13,973	0.84	0.6	15.2	10.0
	SuperPoint	SuperGlue	14,409	0.84	0.6	17.0	9.9
	LoFTR		1,194,197	0.21	0.6	13.1	34.2
	RIFT2		118,463	0.50	0.8	22.3	36.6

Table 5.3: Combined TIR-RGB bundle block adjustment for Lenton Lodge Gatehouse.

RIFT2, the only reference hand-crafted matcher to successfully oriented the TIR-
RGB blocks, presents markedly different results to the learning-based matching
approaches. For Dunrobin House, whilst RIFT2 exhibits a RMSE_{TIR} in line with
the DL-based approaches, reporting a RMSE of 7.7 mm and a MRE of 0.9 pix,
only 22% of the 2D matches are used as effective 3D tie points, suggesting the
geometrically verified matches are not successful within traditional SfM workflows.
In addition, its RMSE_{RGB} is $10\times$ higher than that of other methods, an image block
likely richer in features with respect the TIR block. This is mirrored in Lenton Lodge
Gatehouse, similar achieving the worst RMSE_{RGB} (36.6 mm) and RMSE_{TIR} (22.3
mm). Upon inspection, RIFT2 presents a markedly uneven tie points distribution
on RGB pairs; which, when combined with the need for initialisation using the RGB
images for a successful BA, may have led to a poor block orientation. Interestingly,
RIFT2 required specific conditions for initialisation to successfully co-orient the
multi-modal block, including: (1) an additional geometric verification with epipolar
geometry and error threshold for inlier matches (4.0 pix to 1.5 pix); and (2) the RGB
block was initialised first during the combined TIR-
RGB bundle block adjustment, mitigating its complete automation. Overall, TIR-
RGB image orientation results are preferable for the learning-based methods. However,
the performance of 2D-2D fusion methods must be assessed through the optimised BA
processes, determining their effectiveness utilising all the value present with each
modality.

Optimised TIR- RGB Bundle Block Adjustment

This section evaluates the effectiveness of the proposed optimisation method for
reintroducing the full resolution RGB images as constraints for an additional, con-
strained BA (Section 5.3.4). The results for Dunrobin House and Lenton Lodge
Gatehouse are presented in Tables 5.4 and 5.5, respectively. The original results
with downsampled RGB images (‘Original’) are compared to those obtained with
full-resolution RGB images (‘Optimised’).

Local Features	Matcher	Original		Optimised	
		RMSE _{TIR} (mm)	RMSE _{RGB} (mm)	RMSE _{TIR} (mm)	RMSE _{RGB} (mm)
ALIKED	LightGlue	14.1	1.9	12.5	1.7
DeDoDe	Nearest Neighbour	8.8	3.1	9.8	1.9
DISK	LightGlue	11.1	3.2	10.5	1.4
SuperPoint	LightGlue	8.4	1.7	11.3	1.4
SuperPoint	SuperGlue	5.7	1.7	8.6	1.2
LoFTR		13.5	1.5	16.8	3.7
RIFT2		7.7	13.3	9.8	3.5
Agisoft Metashape (Downsampled Res.)		12.9	1.3	-	-
Agisoft Metashape (Full Res.)		-	-	12.9	1.2

Table 5.4: Comparison of Original and Optimised TIR-RGB bundle block adjustment for Dunrobin House.

Local Features	Matcher	Original		Optimised	
		RMSE _{TIR} (mm)	RMSE _{RGB} (mm)	RMSE _{TIR} (mm)	RMSE _{RGB} (mm)
ALIKED	LightGlue	10.8	6.3	11.6	3.7
DeDoDe	Nearest Neighbour	21.4	4.7	19.1	2.7
DISK	LightGlue	15.5	7.2	10.3	2.0
SuperPoint	LightGlue	15.2	10.0	13.7	2.6
SuperPoint	SuperGlue	17.0	9.9	15.2	2.3
LoFTR		13.1	34.2	11.6	8.3
RIFT2		36.6	22.3	29.5	19.3
Agisoft Metashape (Downsampled Res.)		15.3	2.0	-	-
Agisoft Metashape (Full Res.)		-	-	15.3	2.0

Table 5.5: Comparison of Original and Optimised TIR-RGB bundle block adjustment for Lenton Lodge Gatehouse.

For Dunrobin House, the process of reintroducing the detailed information from full resolution RGB images has negligible effect on the RMSE_{TIR} . Whilst improving the RGB block could theoretically improve the orientation of the TIR block, it is likely that the maximum achievable accuracy with the multi-modal features has been reached at this GSD. However, significant improvements in the RMSE_{RGB} are observed, decreasing from an average of 3.8 mm for the downsampled RGB images to 2.1 mm with the full-resolution RGB images. This brings results in line with the reference 3D-3D fusion results from Agisoft Metashape processed at full resolution. Importantly, these results ensure that the maximum geometric accuracy is maintained for the RGB block, on which the construction of the 3D model will be based. LoFTR is the only method that does not improve with this procedure, likely due to the quantity of semi-dense points and influence of the TIR block, increasing the MRE to ~ 8 pix. RIFT2 also shows substantial improvement, although its RMSE_{TIR} remains higher than that of the other methods. Finally, all the results of the feature + matcher learning-based algorithms achieve a better RMSE_{TIR} than that of the reference 3D-3D fusion method, suggesting the superiority of the proposed 2D-2D fusion method for IRT-3DDF.

Similarly, the results from Lenton Lodge Gatehouse largely confirm the trends observed for Dunrobin House, with all RMSE_{RGB} results improving when constrained on the full resolution RGB images. Although the RMSE_{RGB} does not always reach the 2 mm achieved by Agisoft Metashape, it remains within a range of 2.0-3.7 mm for all feature + matcher DL-base approaches. Additionally, all approaches improve their RMSE_{TIR} (apart from ALIKED), demonstrating the benefit that can be gained, possibly at greater GSDs, when introducing the high resolution RGB images. For LoFTR, which shows similarly noisy results for Dunrobin House, the RMSE_{RGB} improves from 34 mm to 8 mm; however, 8 mm is still considerably higher compared to the other methods. Finally, RIFT2 shows minimal change between the RMSE_{TIR} and RMSE_{RGB} values when constrained on high resolution RGB poses.

This marks a significant outlier in the results across all constrained BA processes, likely a results of poor matches as part of both original and optimised BAs. Overall, the result from Lenton Lodge Gatehouse show promise for the proposed 2D-2D fusion method, with the implemented optimisation allowing results to be refined after initial, low-resolution, multi-modal orientation.



Figure 5.13: (a) RGB and (b) TIR textured 3D polygonal models for Dunrobin House.

5.4.2 Temperature Metrics

After the modelling process was undertaken (Section 5.3.5), generating textured 3D thermal models of both Dunrobin House (Figure 5.13) and Lenton Lodge Gatehouse (Figure 5.14), temperature results were extracted from both models for analysis. The following section reports the comparative results between observed ground truth (T_{obs}), 2D pixel ($T_{2\text{D}}$) and 3D texel ($T_{3\text{D}}$) temperature values for both datasets. As stated in Section 5.3.6, 2D pixel values were extracted from the first instance of the marker in the imaging surveys where the marker is perpendicular to the optical axis, with meter recordings taken adjacent to the building GCP/CP markers on the predominant building material.

To compare Dunrobin House temperature values obtained in situ, in 2D TIR images

Dataset	GCP	Temperature (°C)		
		Ground Truth (T_{obs})	2D Pixel (T_{2D})	3D Model (T_{3D})
Dunrobin House	P4	2.3	2.0	1.9
	P5	2.4	1.9	1.9
Lenton Lodge Gatehouse	P1	1.3	1.1	2.5
	P5	1.9	2.2	3.9

Table 5.6: Ground truth infrared thermometer, corrected 2D TIR pixel (T_{2D}), and resulting 3D model (T_{3D}) temperature values for Dunrobin House and Lenton Lodge Gatehouse.

and in 3D models, two separate temperature readings were taken on P4 (bottom right marker) and P5 (bottom central marker adjacent to bay window)(Figure 5.10). Results for the 3D model temperature values fall in line with the observed ground truth and 2D pixel values extracted from the 16-bit TIFF images (Table 5.6), with a difference of ~ 0.5 °C across all three measures. As the Dunrobin House case study presented stable environmental conditions throughout the 9-minute survey, temperature values between sequential or repeat images remained stable throughout the TIR strips. Furthermore, similar results are obtained for the T_{2D} and T_{3D} values, suggesting the texturing methods success in assigning values using images in line with façade normals. This demonstrates the success of the proposed modelling and texturing method, pointing towards the suitability of the proposed method for energy and infrastructure inspection applications (Garrido et al., 2018).

For Lenton Lodge Gatehouse, results in Table 5.6 show differences between ground truth, 2D pixel and 3D model values. Ground truth values for P1 and P5, both featured at the bottom of the facades, show differences between 1-2 °C. This is likely due to the increasing temperature of the façade as the sun moved onto the building (Figure 3.13) and exposed areas previously shadowed to sunlight during passive thermography. This resulted in temperature values for corresponding points increasing significantly during the 22-minute survey, meaning the model texture was built using increasing values within the TIR sequence. This is similarly confirmed by

extracted T_{2D} values of both the thermal-specific markers and Mansfield sandstone, increasing in temperature from the first instance. However, T_{2D} values are in line with ground truth T_{obs} values for both case studies, demonstrating: the ability for IR thermometers to be used for in-situ measurements within IRT-3DDF, previously overlooked and under-utilised; the accuracy of derived temperature measurements when known environmental and subject-specific values are included within TIR cameras prior to TIR imaging surveys (i.e., ambient temperature, relative humidity, material emissivity); and the importance of prioritising perpendicular images within IRT-3DDF texturing approaches.

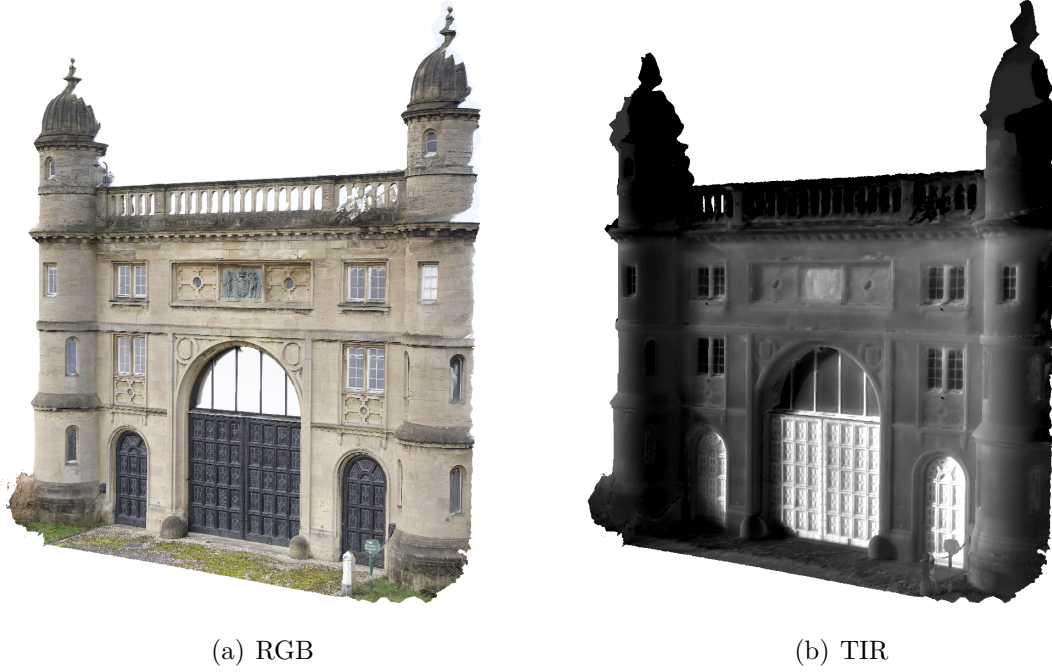


Figure 5.14: (a) RGB and (b) TIR textured 3D polygonal models for Lenton Lodge Gatehouse

5.4.3 Modelling

The following section outlines the final processes of the proposed IRT-3DDF pipeline and the application of derived 3D products for building diagnostics. As the success of thermal texturing largely depends on the registration between 3D models

and image texture (Iwaszczuk & Stilla, 2017), the two best performing multi-modal constrained pipelines from Section 5.4.1 were selected for Dunrobin and Lenton Lodge Gatehouse, namely SuperPoint+LightGlue and DISK+LightGlue, respectively. These methods were used to generate thermally-textured 3D models (Figures 5.13 & 5.14) and thermal orthomosaics (Figures 5.15 & 5.16) to aid the identification of anomalies across both façades. The interpretation of the derived 3D thermal models demonstrates that the achieved level of accuracy for the proposed IRT-3DDF pipeline is sufficient for civil and heritage applications.



Figure 5.15: Quantile thermal orthomosaic for Dunrobin House

Upon inspection of the Dunrobin House 3D model and orthomosaic, several thermal phenomena can be observed. As the thermal survey was undertaken at night (22:18-22:25), visible differences in internal and external building temperature are apparent, resulting in the thermal bridging beneath the roofline, windows and front bay window. Additionally, although TIR images of the roof are limited in the TIR imaging survey, areas of deteriorated loft insulation are identifiable, confirmed in the 2D images and by an internal build inspection. This justifies TIR as a tool for identifying ‘invisible’ phenomena in infrastructure inspections, aided by the ability

to measurements of quantity and location (Kylili et al., 2014). Temperature values range from -2.5 to 5.5 °C, with the main brick façade averaging 1.8 °C, in line with the atmospheric temperature on the night of data acquisition.

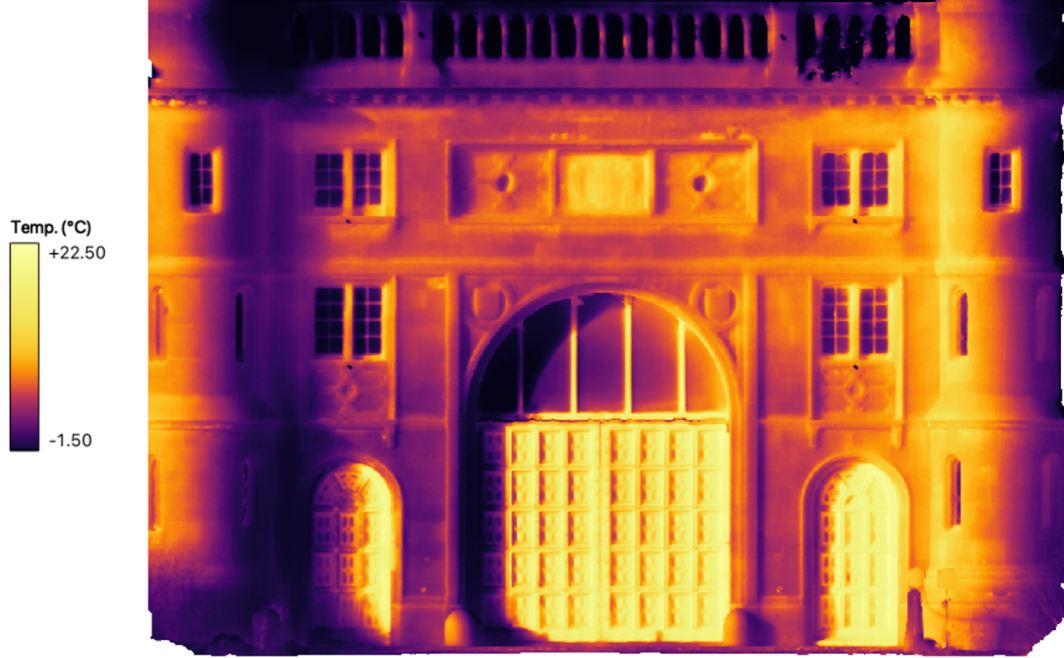


Figure 5.16: Quantile thermal orthomosaic for Lenton Lodge Gatehouse

For Lenton Lodge Gatehouse, the RGB texture, TIR texture and TIR orthomosaic show visible signs of deterioration of the Mansfield sandstone, evident above the decorative traceried panels and bronze plaque underneath the parapet of the façade. Interestingly, this was replaced in the early 2000s due to water damage and corrosion from the lead roofing, suggesting this problem continues to affect the structure. For example, the shrubbery visible in the RGB domain (right turret when facing facade)(Figure 5.14), protruding defaced sandstone visible in the 3D geometry and temperature difference visible in the TIR domain (Figure 5.14) demonstrate the probable return of previous water ingress atop the parapet, corroborated by internal building surveys and hygrometer readings. Additionally, we are able to identify additional areas of deteriorating stone, either defacing or suffering from biological colonisation. Finally, due to the passive thermography employed to generate thermal contrast during the imaging survey, we see the different emissivity

and conductivity of materials featured on the façade, with the wood and iron central gate gaining significant temperature during the course of the survey and presenting significantly increased temperature values (20-40 °C).

5.5 Discussion

5.5.1 Multi-Modal Image Matching for IRT-3DDF

As previously stated, the results of the camera pose metrics, despite varying levels of accuracy, demonstrate the overwhelming success of DL-based methods to orient TIR and RGB blocks within a combined bundle block adjustment, marking a significant junction for IRT-3DDF and the core originality of this thesis. This validates the continued exploration of 2D-2D fusion methods, a facet of IRT-3DDF (and by extension other 3D data fusion protocols)(Morelli et al., 2022) which has been previously under-researched and hampered by the limitations of hand-crafted image matchers. The results in Tables 5.2 & 5.3 demonstrate the following: (1) the success of learning-based feature + matchers to translate matched keypoints into viable 3D tie points (all >74%); (2) the ability for RGB orientation to be highly accurate even with the influence of multi-modal correspondence; and (3) the diversity of multi-modal features (e.g., linear, corner, blob or regions) being detected irrespective of NRD differences (Jiang et al., 2021; Morelli et al., 2025b). Additionally, the computational time required for such processes, seen in Table 5.7, points to the feasibility of DL approaches at varying scales. However, the necessity for RGB images to be downsampled for successful matching with TIR images highlights a limited scale invariance of the DL-based methods. This limitation could potentially be addressed by automating feature detection across an image pyramid at multiple scales, similar to the approach employed by SIFT. In summary, continued testing of DL-based approaches on new multi-modal datasets, development of multi-modal-

specific evaluation criteria, and comparison with established hand-crafted matchers is required to cement the success of 2D-2D fusion for learning-based multi-modal image matching (Bellavia et al., 2022; Chang et al., 2024; Jiang et al., 2021).

BA	Detector	Descriptor	Matcher	Matching Time (min)
Matching		ALIKED	LightGlue	9.4
		DeDoDe	Nearest Neighbour	6.7
		DISK	LightGlue	52.2
		LoFTR	LightGlue	134.1
		SuperPoint	LightGlue	19.3
		SuperPoint	SuperGlue	152.7
MATLAB		RIFT2		>3 days

Table 5.7: Duration of multi-modal image matching using exhaustive (‘bruteforce’) matching with DIM and RIFT2 for Dunrobin House (4,465 image pairs).

Furthermore, the failures of several of the multi-modal matchers during and after the combined bundle block adjustments point to several key considerations for future implementation. It must be noted that SIFT, SRIF and KeyNet+HardNet were unable to co-orient both image blocks during the CBBA, only able to orient individual modalities depending on the initialisation of the bundle adjustment (i.e., starting with either TIR-TIR or RGB-RGB matches). Upon inspection, the quantity and quality of multi-modal matches for these approaches were insufficient for effective image orientation. Additionally, LoFTR, the sole semi-dense matcher, although a viable option for multi-modal image matching, produced RMSE_{RGB} values three to four times higher than other learning-based methods, heavily dependent on the subject and available texture (as seen from the drop of inlier match ratio between Dunrobin House and Lenton Lodge Gatehouse). Future work should assess the benefit of utilising feature + matcher approaches as opposed to a semi-dense matcher for certain multi-modal applications. Finally, the poor performance of RIFT2 and SRIF, both state-of-the-art hand-crafted multi-modal matchers, suggests at the requirements for multi-modal matchers tailored towards 3D reconstruction as opposed to comparatively less complex registrations or transformations. Only available as a

MATLAB implementation, RIFT2 also lacks any built-in parallelization, resulting in the matching time being unreasonably long and unsuitable for large datasets (Chang et al., 2024). Primarily used for aerial and satellite datasets, only used for superficial pairing operations within IRT-3DDF (Lin et al., 2019b), additional tests and tuning is required to assess RIFT2 for multi-view terrestrial applications.

5.5.2 2D-2D Fusion for IRT-3DDF

The results of the proposed 2D-2D fusion method successfully achieved the intended goal: the co-registration of RGB and TIR blocks utilising multi-modal matches. The proposed workflow demonstrates that co-orienting TIR and RGB images in a combined bundle block adjustment is a fitting approach for IRT-3DDF, mitigating the long-standing issues presented by hand-crafted multi-modal image matching. Importantly, this is done without the need for: GCPs or a posteriori co-registration; a fixed baseline between independent sensors (range-, image-based or TTISs); or simultaneous data capture between TIR images and base geometries. Notably, this proposed method is fully automatic; which, when compared with the reference 3D-3D method using Agisoft Metashape, required 200 and 479 manually-annotated 2D marker coordinates for Dunrobin House and Lenton Lodge Gatehouse, respectively. Regarding the base geometry, the current method imports the 3D model for texturing from the independent RGB block used to generate the RGB poses for the constrained optimised BA. As these poses are not *exactly* the same as those leaving the BA (constrained $\sigma_{xyz}=1$ mm; $\sigma_{\omega\phi\kappa}=0.1$ deg), additional tests are needed to compare the geometries derived from both sets of poses (independent and ‘optimised’) to see whether the proposed method can be refined.

The importance of the proposed method is in the orientation of the TIR image block; which, when derived to a degree of accuracy, can then be used to texture a wealth of 3D base geometries (image (Zhu et al., 2021), range- (Biswanath et al.,

2023) or parametric model (Hoegner & Stilla, 2007) based). Although correct 16-bit TIFF TIR images were re-introduced solely for texturing (Section 5.3.5), we must not assume accurate T_{2D} values will translate to accurate T_{3D} values, with temperature changes *across* images still present within the TIR sequences (Iwaszczuk & Stilla, 2017; Lin et al., 2019b). For Lenton Lodge Gatehouse, over the course of a 22-minute survey, temperature values for re-visited points show significant changes which will inadvertently affect the resulting texture. For points in Table 5.6, differences of 0.7 °C and 2.5 °C are observed between the first and last observation of P1 (shadowed area) and P5 (illuminated area), respectively. Similarly, inconsistencies in the quantile orthomosaic for Dunrobin House (see. Fig. 5.15) demonstrates the shortcomings of the current texturing method to reduce discrepancies between lower and upper strips, presenting a ~ 0.5 °C difference between the lower and upper half of the façade. Currently, the method presents a fully automated texturing procedure exploiting weighted pixel values and image perpendicularity; however, opportunities to refine this method are plenty, including: statistical texturing strategies (Lagüela et al., 2012; Malbêteau et al., 2021), multi-modal texture fusion (Adamopoulos et al., 2021; Vidas et al., 2013), temperature representation (i.e., point vs. mesh) (Lin et al., 2018a) and TIR image enhancement (Luhmann et al., 2016; Zhang et al., 2023b).

5.6 Conclusion

This chapter contributes to the growing field of IRT-3DDF, demonstrating the application of pre-trained neural networks for multi-modal image matching methods generating 3D thermal models. Importantly, these neural networks, trained on vast datasets of RGB imagery with significant radiometric and geometric variations, show that they can generalise outside their training domain to accommodate multi-modal image matching. Although 2D-2D fusion methods remain under-utilised for

IRT-3DDF, this chapter shows the effective use of multi-modal image matching to automatically co-register TIR and RGB imagery in a combined bundle adjustment. This is compared to established hand-crafted methods, demonstrating the superior performance of DL in managing the significant radiometric differences between TIR and RGB images. In addition, the effectiveness of the proposed method is evaluated through both 2D and 3D metrics, whilst assessing the processes ability to generate accurate temperature measurements within derived models. Notably for IRT-3DDF, the proposed methodology demonstrates the possibility for asynchronous and repeated surveys to be undertaken, allowing images to be generated at optimal times for each modality and capturing specific conditions for varying applications. Finally, this chapter demonstrates a fusion pipeline without a reliance on GCPs, fixed relative orientation between sensors or positioning information from GNSS/IMU, drastically increasing its efficiency and applicability. The future developments of this work should look to expand the 2D-2D fusion procedures by demonstrating the versatility of the proposed pipeline across new sensors, subjects and scales.

Chapter 6

Discussion & Future Work

Contents

6.1	Introduction	172
6.2	IRT-3DDF: A Justification	172
6.3	An Appraisal of 3D-Data Fusion	174
6.3.1	Sensor Configurations	175
6.3.2	Fusion Space	177
6.4	Radiometric Calibration, Correction & Validation	179
6.4.1	Radiometric Calibration	181
6.4.2	Radiometric Correction & Validation	183
6.4.3	Radiometry vs. Photogrammetry: A Balancing Act . . .	186
6.5	Thermal InfraRed Cameras: Resolutions, Specifications & Future Developments	187
6.6	Multi-Modal Image Matching for IRT-3DDF	189
6.6.1	IRT-3DDF: Training, Tuning & Testing	190
6.6.2	Multi-Modal & Multi-Sensor Calibration	192
6.6.3	Thermal Modelling: Poses, Geometries & Textures . . .	195
6.7	IRT-3DDF for Architectural Heritage	198

6.1 Introduction

The emergence of InfraRed Thermography 3D-Data Fusion, a discipline, at its heart, dedicated to overcoming the inherent complexity combining datasets of varying source and scale, reflects the broader trends within photogrammetry, computer vision and geomatics. With an ever-increasing spectrum of sensors and platforms capable of representing the real-world, all providing greater functionality, flexibility and degrees of automation, the ability (and, by extension, the *need*) to develop new methods for data fusion will continue to grow. Similarly, the adoption of IRT-3DDF for architectural heritage mirrors the growth seen in similar fields within heritage science (e.g., eXtended Reality (XR), Artificial Intelligence (AI), H-BIM), taking advantage of novel methods from the AECO industry and tailoring solutions to advance CRM practices (Kennedy et al., 2024). Importantly, the innovative disciplines that become established tools within digital cultural heritage are those that refine, diversify and enrich CRM practices, providing value above existing, time-tested approaches. Therefore, to assess the current value of IRT-3DDF, and the opportunities it holds as a dedicated discipline, an appraisal of its merits, limitations and future must be undertaken. In a bid to contextualise the previous research within the broader scope of IRT-3DDF, the following sections expand the findings from this thesis, pointing to critical observations, opportunities and critical recommendations for future work within the field.

6.2 IRT-3DDF: A Justification

The classification of IRT-3DDF as a field has looked to answer two fundamental questions: (1) what value do we achieve through data fusion that TIR images alone cannot provide; and (2) what benefit does *3D*-data provide over traditional *2D* images. Firstly, it could be argued that with the continued increase in sensor res-

olution within the commercial TIR camera market (Corsi, 2015), the production of 3D thermal models using solely TIR images may soon become sufficiently accurate and automated (both in terms of geometry and radiometry) to render data fusion obsolete. However, even if this were to be the case (although unlikely at the current rate of academic and commercial development), IRT-3DDF stills provides marked benefits that cannot be achieved utilising TIR images alone (Table 6.1). Although we previously defined data fusion as a process that uses “the strengths of individual data sources to complement and supplement the limitations of another” [p. 3], fusion allows for the combination of datasets *regardless* of their strengths or limitations, meaning the value of an integrated output often exceeds the value coming from an individual input. In the context of IRT-3DDF, we have seen the benefit of fusing range-, image- and TTIS-based sensors not only to improve geometric accuracy, but to interpret a 3D thermal model from multiple perspectives and modalities. The value provided by data fusion lies in the coherent representation of a complex asset; which, when applied to architectural heritage, lends itself to the inherently multi-disciplinary nature of heritage documentation.

IRT	IRT-3DDF
Mono-disciplinary 2D dataset	Multi-disciplinary 3D datasets
Analysis on 2D images	Analysis on 2D images & 3D models
Selective visualisation & integration	Enhanced visualisation & integration
2D area	3D volume
Proximal & Localised	Contextual & Globalised
Highly-interpretable	Nuanced interpretation
Rapid, real-time data capture	Pre- & post-processing required

Table 6.1: Comparisons between IRT and IRT-3DDF.

Secondly, as IRT represents an established and highly-effective NDT technique in its own right, an appraisal of 3D-data fusion is justifiable. The interpretation of 2D images is a proven approach for identifying thermal anomalies, detecting invisible phenomena and visualising temperature (Garrido et al., 2022), requiring little in the way of post-processing for meaningful insights to be extracted (Brooke, 2018). How-

ever, it is evident that IRT-3DDF does not seek to be a *replacement* for IRT, but a viable alternative or accessory capable of utilising novel technologies when appropriate. Where IRT-3DDF has been successfully applied, past superficial ‘visualisation’, ‘method development’ or ‘data integration’ (Sutherland et al., 2023), data fusion has enabled: virtual interactions with thermal point clouds within XR (Pérez et al., 2024); the representation of TIR images using Neural Radiance Fields (NeRF)(Lin et al., 2024); and, the amendment of TIR images within BIM (Natephra et al., 2017). The novelty of IRT-3DDF lies in the ability to observe inherently complex phenomena in three dimensions, allowing components such as proximity, volume and orientation to become relevant factors in building diagnostics (Figure 6.1). As a result, to cement the future of IRT-3DDF, greater distinction must be sought between the individual merits of IRT and 3D-data fusion tailored to specific NDT use cases and applications.



(a) 2D Image



(b) 3D Model

Figure 6.1: A comparison between (a) a 2D image and (b) a 3D model texture generated for Dunrobin House.

6.3 An Appraisal of 3D-Data Fusion

From its inception, IRT-3DDF has fallen within broad categories, represented by the formulaic structure of 2D-2D, 2D-3D and 3D-3D data fusion in Chapter 2. Pur-

posefully, each of the three categories of IRT-3DDF have been undertaken as part of this thesis, including: 2D-2D fusion (Sections 4.4.1 & 5.3), 2D-3D fusion (Section 4.4.2), and 3D-3D fusion (Section 5.3.6). The benefit of including each approach as part of this work lies not in their direct comparison, but in the evaluation of their use in specific cases. Whilst this thesis has focussed on the development of a novel 2D-2D fusion method, it is important to stress this does not provide a *perfect* method, but demonstrates the implementation of a method *best suited* for the given data and applications. When IRT-3DDF is successful, achieving purposeful data fusion to extract meaningful insights, an understanding for the constraints imposed by the scene, equipment and applications is expressed. Whether this be a fixed orientation between sensors, the requirement for surveys to be undertaken at night or the need for future H-BIM integration, adapting methods to accommodate specific conditions is a critical focus for IRT-3DDF. By doing so, the products generated by IRT-3DDF (not only the resultant 3D thermal models, but the accompanying camera poses, rectified images and derived temperature values), can be repurposed and adapted to tackle additional challenges.

6.3.1 Sensor Configurations

In order to evaluate the diversity within IRT-3DDF, Table 6.2, presenting a summary of our defined data fusion approaches designated by sensors and fusion ‘space’, points to some of the merits and critical considerations when choosing an appropriate 3D-data fusion method. Viewing each sensor configuration (represented by each row), it is apparent that as the ‘quality’ of the additional data source increases (from a comparatively low-res. VIS sensor up to an ‘intelligent’ parametric model), so does the complexity of fusion and texturing. Not only does this progression require fusion methods to determine correspondences with: images of significantly greater resolution (i.e., RGB), vastly different sensors (i.e., TLS) and completely

different digital products (i.e., parametric building models); this comes with significant time, expense and experience costs. Although the VIS sensor is unsuited for detailed base geometry creation, the benefits of TIR-VIS, utilising fixed sensors coming from a singular TIR camera or TTIS, are not to be understated, allowing for rapid and simultaneous data capture exploiting fixed baseline constraints. Notably, the fusion of multi-modal sensors from tandem sensors, be it thermal (Güidi et al., 2025) or hyperspectral (Altieri et al., 2022), is now accommodated by commercial SfM software such as Agisoft Metashape and RealityScan (formerly RealityCapture)(Epic Games, 2025), demonstrating the utility of pre-fixed sensors as a means of 3D-data fusion.

In contrast, what TIR-RGB, TIR-TLS or TIR-Parametric Model approaches lose in functionality, they gain in flexibility; all capable of providing asynchronous data capture, ease of repeated surveys and rich geometries for effective modelling. Excitingly, TIR-RGB fusion, combining datasets of comparable source, is being continually advanced, with the ubiquity of scale invariant learning-based image matching models providing great promise across *all* multi-view modelling disciplines (Jiang et al., 2021; Jin et al., 2020; Ma et al., 2021). In addition, the development of TTISs integrating TIR sensors *within* a TLS body (e.g., Leica BLK360, Zoller+Fröhlich 5010C+‘T-Cam’)(Fang et al., 2022) means the assumption of a fixed RO between sensors can now be employed between sensors of different source. Furthermore, developments from the AECO industry are now enabling the texturing of (H)BIM parametric models with photogrammetric results to become a common plug-in to BIM software (Brumana et al., 2018a; Martín-Lerones et al., 2021), increasing the automation of some processes which have previously hampered the adoption of parametric models for conservation practices (Solla et al., 2020). In summary, the benefits of having highly-accurate and detailed 3D geometries (not solely for thermal modelling, but for holistic asset information management (Historic England, 2018a)) must be the focus for future IRT-3DDF methods, exploiting sensor config-

urations for the extraction of maximal interpretative and generative value.

6.3.2 Fusion Space

By inspecting the fusion ‘spaces’ of IRT-3DDF (represented by the columns of Table 6.2), we can similarly determine the shared obstacles and opportunities of different approaches, notably the preparation of data *prior* to fusion. Opposed to 2D-3D and 3D-3D fusion, where 3D products are generated in advance of the fusion method, 2D-2D fuses images immediately and directly, executing IRT-3DDF pipelines where images can be readily used and transformed into thermal models (similarly for thermal SLAM)(Ramachandra et al., 2018). The novelty of the proposed method in Chapter 5 lies in the ability for IRT-3DDF workflows to be orchestrated with *complete* automation, inputting 2D multi-modal images in one end and generating a 3D thermal model at the other. Whilst the proposed pipeline required the transfer of derived camera networks through multiple software (e.g., DIM, COLMAP, Metashape), the execution of these processes within a singular program or API represents a realistic target for future refinement, capable of being tailored to sensor configurations, matching models and specific applications.

		Fusion ‘Space’		
		2D-2D	2D-3D	3D-3D
	General	✓ (Comparative) ease matching common formats ✗ Challenge of determining correspondence ✗ Relies on high contrast, low noise	✓ Ease of repeated TIR surveys ✓ Ability to texture existing models ✗ More challenging 2D-3D correspondence	✓ Strength of ICP algorithms ✗ Often requires manual alignment or GCPs ✗ Poorer quality of 3D TIR data
TIR-VIS	✓ Able to employ fixed RO constraint ✓ (Comparatively) cheaper (single camera / UAV) ✓ Simultaneous data capture	✓ Closer res. makes image matching easier ✓ Similar perspectives of resulting images ✗ Greater GSD reduces model detail	✓ Rapid & vast data capture using UAVs ✗ Lower res. VIS sensor not preferable for modelling ✗ Requires optimal conditions for both modalities	✓ Proximal sensors aid initial co-registration ✓ Ease of initial modelling using SfM ✗ Fusing two lower-quality 3D models
TIR-RGB	✓ Exploiting highest res. imaging data ✓ Asynchronous data capture ✗ Often processing intensive	✓ Proven ability of combined bundle block adjustments ✗ Different res. makes image matching challenging ✗ Difficultly calibrating & manoeuvring stereo rigs	✓ Suitable for aerial & terrestrial cases ✗ Requires scaling / referencing information ✗ Invariance of feature matching approaches	✓ Strong geometry generated from RGB ✓ Ease of initial modelling using SfM ✗ Requires scaling / referencing information
TIR-TLS	✓ High-res. dense point clouds for ‘base geometry’ ✗ Largely terrestrial use ✗ Expensive hardware	✓ Possible using TLS 2.5D images ✓ Challenge of transforming TLS from 3D to 2.5D ✗ Range of methods need expanding	✓ Combines best datasets in preferred formats ✓ Use of TTISs for vast data collection ✗ Difficultly determining common features automatically	✓ Dense cloud from TLS provides best 3D ‘base’ ✗ Uncommon fusing dense TLS & sparse TIR clouds ✗ TIR GSD makes assigning temperature inefficient
TIR-Parametric Models	✓ (H)BIM Integration ✗ Compromises between LoD & texture quality ✗ Requires user expertise, processing time, costly software	–	✓ Current BIM libraries capable of texturing models ✓ Beginning for simulation & diagnostics ✗ Challenging fusion approaches	✓ Future area of exploration ✓ No research fusing parametric models and TIR clouds ✗ Would require significant data manipulation

Table 6.2: Summary of IRT-3DDF methods strengths and shortcomings, seperated by their fusion ‘space’ and datasets.

However, whilst generating novel 2D-2D methods may aptly address the challenges posed by processing TIR images photogrammetrically, we must appreciate that the generation of 3D models within DCH is well-established, and the determination of methods capable of exploiting this available data is preferable. By achieving sensor fusion using previously-generated base models separate from TIR data capture (notably 2D-3D fusion), numerous repeated TIR surveys can be undertaken and subsequently integrated, allowing for temporality (i.e., daily, nightly, seasonally) to become a component for IRT-3DDF. Additionally, this opens opportunities for the comparison of individual surveys, the simulation of change and the statistical analysis of derived temperature measurements. In summary, the appraisal of 3D-data fusion points to the following conclusions: (1) the future of IRT-3DDF may lie in tandem fields, adapting tools to handle the inherent complexity of TIR images; (2) new sensors will continue to dictate new methods, with the capabilities of devices determining the flexibility of new fusion approaches; and (3) methods *must* be developed tailored to specific applications, without which 3D thermal models cannot be purposefully employed.

6.4 Radiometric Calibration, Correction & Validation

To reiterate a central mantra of this thesis: a foundational knowledge of infrared radiation is fundamental to successful IRT-3DDF. The effectiveness of IRT for architectural heritage, and by extension IRT-3DDF, is largely determined by the experience of the professional undertaking TIR imaging surveys, with the planning of thermal data capture dictating the derived outcomes. Only by understanding the nature, transmittance and observation of infrared radiation can best practices be implemented and refined for specific investigations. Importantly, as a plethora

of camera- (i.e., FPA sensor temp., NUCs, FPN), scene- (i.e., ambient temp., wind, solar radiation) and subject- (i.e., material emissivity, thermography approach, condition) specific phenomena affect both the radiation *emitted* by a body and the radiation *received* by the TIR camera, the failure to acknowledge and account for such sources of error limits methods hoping to undertake quantitative thermography to purely qualitative results. Therefore, for IRT-3DDF to pursue its ambition of providing more than novel visualisations, extracting meaningful quantitative analysis of derived temperature measurements, greater transparency around possible sources of radiometric error and means of validating derived radiometric values is paramount.

Firstly, an important distinction must be made regarding the term ‘radiometric calibration’ for IRT-3DDF, with this thesis using three separate terms to address radiometry, namely: calibration, correction and validation. Whilst the term ‘radiometric calibration’ can represent the rigorous procedure of establishing gain and offset values against that of a calibration blackbody (Riveiro & Solla, 2016; Vollmer & Möllmann, 2018)(often undertaken by the camera manufacturer during quality checks and frequently used as a proxy for the accuracy of derived temperature values within IRT-3DDF research)(Dlesk et al., 2022; Guilbert et al., 2020), the term has also been used to include procedures where image sequences are *corrected* for sensor dynamics (Ribeiro-Gomes et al., 2017), environmental conditions (Grechi et al., 2021) or enhanced visualisation (Lewińska & Maciuk, 2020). However, many of these amendments do not involve the integration of known temperature references, meaning the use of the term ‘calibration’ is ambiguous and misleading. In addition, much like *calibration* seeks to determine the accuracy of recorded T_{obs} and T_{2D} measurements, *validation* must be undertaken to correlate T_{obs} and derived T_{3D} model- values, allowing results generated from IRT-3DDF to state their accuracy and precision comprehensively. As radiometric accuracy is denoted as a core component of IRT-3DDF, the purposeful inclusion of these terms is to differenti-

ate methodological steps where temperature measurements are *calibrated* to known values, *corrected* through intra- or inter-image post-processing, or *validated* against known ground truth. Only by differentiating between procedures can we isolate current shortcomings and demonstrate future opportunities for accurate thermal modelling.

6.4.1 Radiometric Calibration

At present, the lack of means to calibrate, correct and validate temperature measurements throughout IRT-3DDF workflows presents a critical barrier, one that has forced research (including the methods proposed here) to find provisional or pragmatic means of assessing radiometric accuracy (Kraaijenbrink et al., 2018; Pestana et al., 2019). However, although inherently heuristic in nature, these solutions are not to be dismissed, providing the opportunity to develop methods that can assess *relative* consistency and reliability if not capable of determining *absolute* accuracy and precision. Whilst access to a blackbody calibrator is necessary for radiometric calibration, access to proximal FPA sensor temperature dictates sensor acclimatisation, and access to field radiometers provide the necessary ground truth for validation, the development of amendable methods capable of integrating such data when (or *if*) it becomes available is an achievable standard for IRT-3DDF. Furthermore, the evident sequential nature of radiometric calibration, correction and validation procedures, requiring pre-, during and post-processing steps to mitigate radiometric errors, allows conclusions to be carried forward to build an overall picture of radiometric accuracy within implemented methods. The interconnectivity of radiometric calibration, correction and validation within IRT-3DDF requires future methods to document all decisions taken to ensure radiometric accuracy; which, through transparent and justified methods, including those without the necessary access to hardware, can begin to generate 3D thermal models with greater

confidence and reliability.

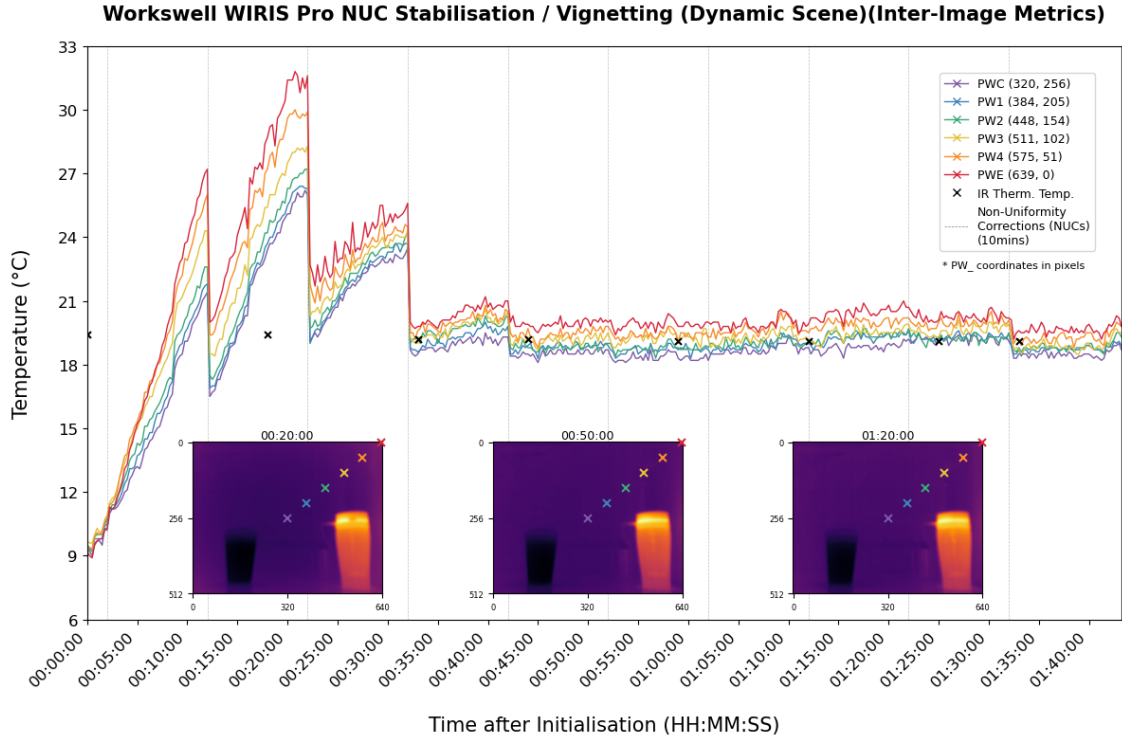


Figure 6.2: Visualisation of WWP vignetting through inter-image pixel-wise metrics, featuring inset sample images (comprising the thermos and cooler) from the TIR sequence throughout the ~ 105 min ‘dynamic’ scene experiment.

Assessing the proposed methods, whilst the pseudo calibration set-up provided the necessary environment to evaluate the radiometric consistency of temperature measurements, FPA sensor acclimatisation and NUCs (all of which did not *require* absolute temperature for assessment), the inability to observe multiple known temperature measurements within controlled conditions prevented: (1) the calculation of camera-specific gain and offset values for image correction; (2) the simulation of changing environmental conditions (inc., wind, atmospheric temperature, relative humidity)(Figure 6.2)(Wan et al., 2021); (3) the evaluation of trade-offs between NUC increments and image consistency (Cao & Tisse, 2013; Lagüela et al., 2016); and (4) the assessment of TIR FPN including vignetting (Hardie et al., 2000), striping (Lin et al., 2018b) or dead pixels (Budzier & Gerlach, 2015)(Figure 6.3). In addition, the inability to observe and extract FPA sensor temperatures *during*

image capture significantly reduces the methods from which resulting TIR can be corrected and verified (Wan et al., 2021). Therefore, as access to such equipment is evidently lacking throughout IRT-3DDF (Adán et al., 2023b; Lin et al., 2017; Malmivirta et al., 2019), greater transparency on the confidence of derived temperature measurements needs to be brought forward (Treier et al., 2024, p. 723); not to criticise methods for ignoring radiometric accuracy, but so results can be viewed as either relative or absolute (Döpper et al., 2020).

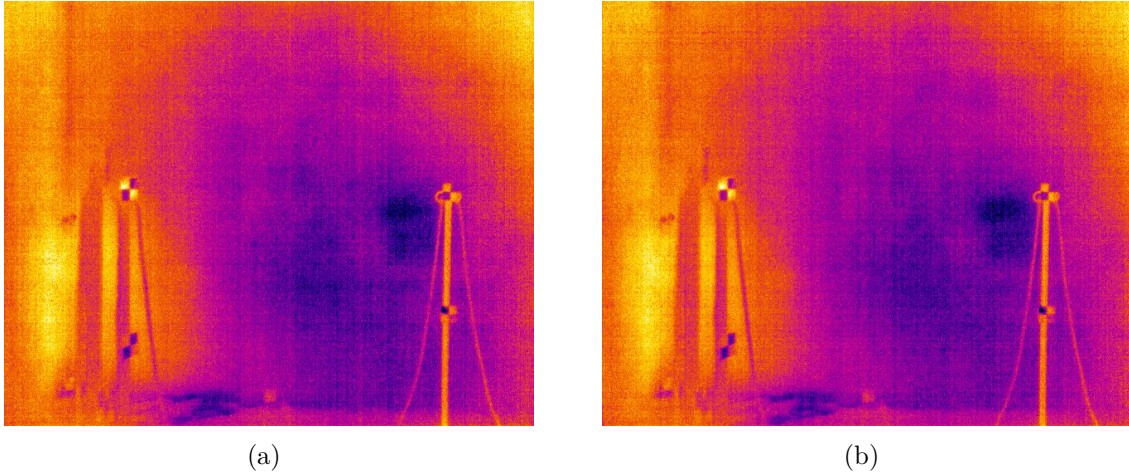


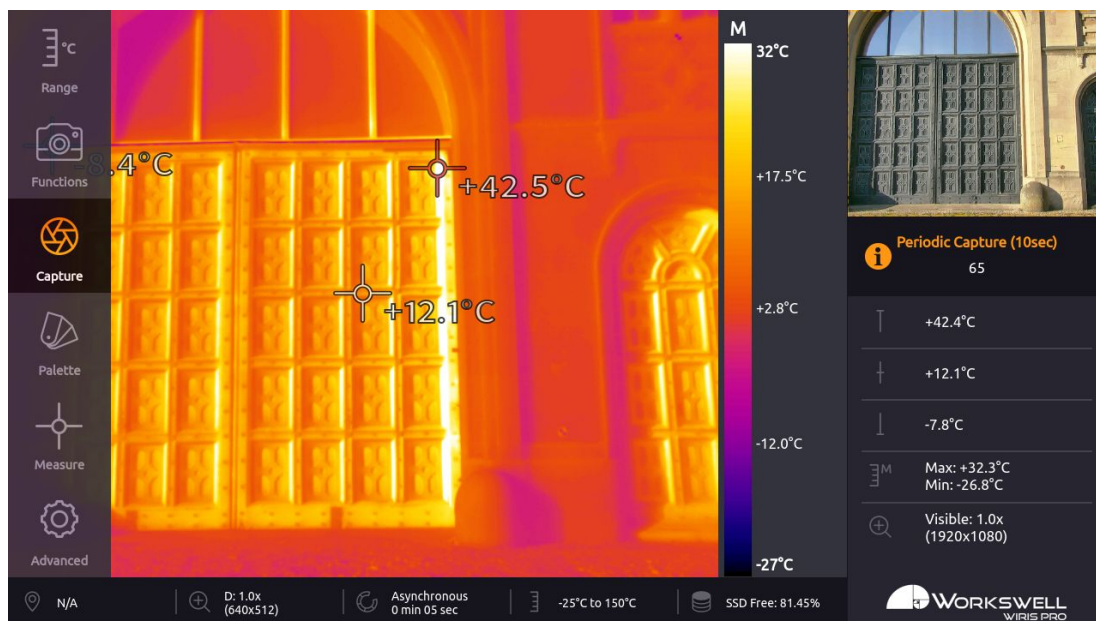
Figure 6.3: An example of the difference in FPN between sequential, stationary images at Castello di Arco.

6.4.2 Radiometric Correction & Validation

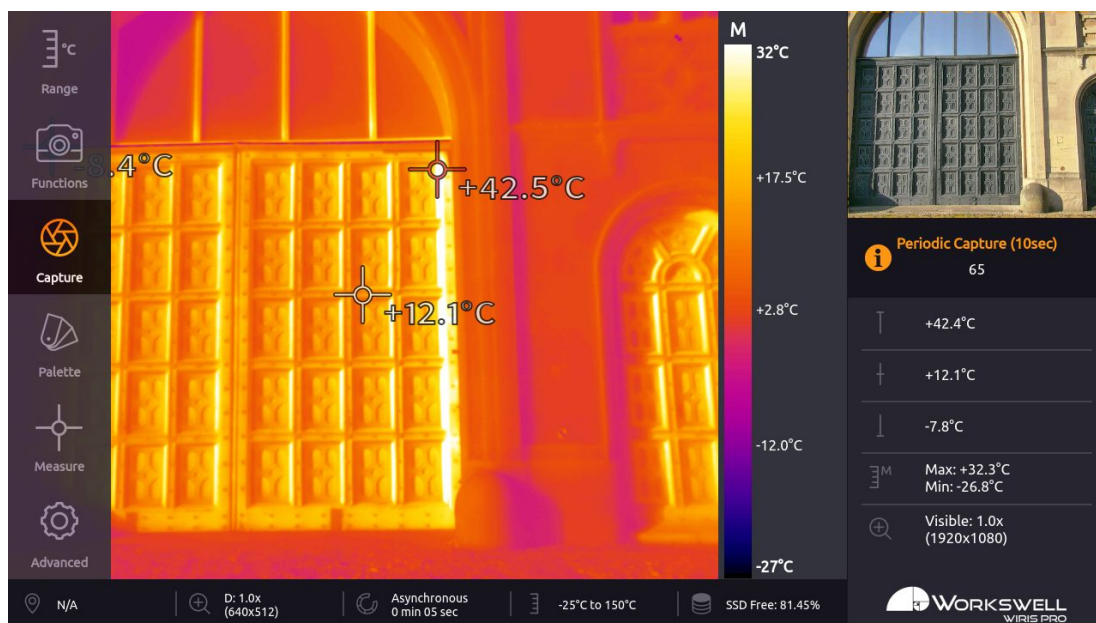
However, even if radiometric calibration was to be undertaken with all the preferred hardware, a critical fact must still be acknowledged: environmental conditions during an IRT survey cannot be wholly accounted for or mitigated. Whilst several laboratory experiments have looked to simulate ‘live’ surveys using controlled climatic chambers, fans and external camera shields (Kelly et al., 2019; Lin et al., 2018b; Wan et al., 2021) (including this work, where a ‘dynamic’ scene of varying temperatures was used to determine radiometric calibration accuracy, FPA sensor acclimatisation, NUC periodicity)(Figure 6.2), these results do not translate to TIR surveys undertaking real data capture. For Lenton Lodge Gatehouse, where TIR

surveys were undertaken with ambient temperature just above freezing, a 75 min acclimatisation period was needed to bring the FPA sensor to a stable temperature, ~ 40 mins longer than that calculated during pre-calibration. In addition, whilst incremental NUCs were set at 3 min intervals, the WWP undertook several additional automatic NUCs in response to the camera’s internal thermometers and environmental conditions (Figure 6.4). Therefore, even though pre-calibration plays a significant role in determining the stability of each individual TIR camera, greater focus must be placed on *correction* methods capable of integrating known scene- and subject-specific environmental data in the post-processing of derived TIR sequences (Dugdale et al., 2019; Wan et al., 2024).

To this end, the need for proximal environmental and subject sensors, recording the critical information for correcting and validating derived TIR images, represents the greatest opportunity for IRT-3DDF, with TIR image sequences requiring accurate data representative of the atmospheric attenuation within the scene (Section 3.3.3)(Dugdale, 2016). At present, such data is collected (and has been collected for the proposed work) using hygrometers, anemometers, barometer and infrared thermometers (Mesas-Carrascosa et al., 2018), ultimately limited in their scope, frequency and accuracy for IRT-3DDF. However, the ability to automatically generate and integrate this information for the correction of TIR images, in line with the changes present within dynamic environmental conditions, represents a logical progression of current practices (Lin et al., 2017). DJI’s Dock 3, with integrated wind speed, ambient temp. and humidity sensors, demonstrates one of the first commercial solutions to this issue, beginning the process of including site-specific values within TIR image generation (DJI, 2025). Furthermore, the ability to orchestrate a extensive network of known subject temperatures, across the exterior and interior of a building requires further investigation, with the correlation of multiple derived T_{obs} , $T_{2\text{D}}$ and $T_{3\text{D}}$ values fundamental for comprehensive assessments of radiometric accuracy (Adán et al., 2021). In summary, if scene-specific calibration,



(a)



(b)

Figure 6.4: Example WWP screenshots from Lenton Lodge Gatehouse, showing the execution of an automatic NUC outside of the programmed 3 min intervals (at the bottom of each figure, the duration between NUCs is apparent, here changing at ~ 2 min 30 secs as opposed to 3 mins).

correction and validation procedures are to be implemented into future IRT-3DDF methods, this has to be initiated with the integration of known temperature measurements, including FPA sensor temperature-, environmental- and subject-specific sensors (Abolt et al., 2018; Mesas-Carrascosa et al., 2018; Szostak et al., 2023; Wang et al., 2023; Wang et al., 2025).

6.4.3 Radiometry vs. Photogrammetry: A Balancing Act

Finally, for IRT-3DDF data capture to be cohesive, including specific IRT, photogrammetry, geospatial engineering and building surveying practices, we must accept that a ‘balancing act’ will always be required to undertake these disciplines simultaneously. This balancing act is evident in the duration required to undertake surveys, with the need for sufficient image overlap for photogrammetric processing increasing the possibilities of thermal drift between TIR images. It is apparent in the need for stable building conditions for building energy audits, even though this comes at the cost of reducing contrast in TIR images. It is seen in the geometric calibration of TIR cameras, requiring novel test fields that serve multi-modal calibration but reduce the automation of control point detection. The purpose of identifying this problem is not to deem it unsolvable, but to stress the need for decisions to be justified within IRT-3DDF, demonstrating the considerations taken to achieve specific goals. By doing so, new methods can be developed informed on the rationale presented by preceding research, critical for the advancement of IRT-3DDF as an academic discipline and tool for CRM practices.

In this context, the difficulty balancing radiometric and photogrammetric principles within the presented work identifies several areas for future IRT-3DDF research to investigate. Firstly, whilst the use of passive thermography for all TIR surveys was chosen as a means of generating stable TIR sequences (although this was undoubtedly complicated by the emerging solar radiation at Lenton Lodge Gatehouse), the

need for contrast within TIR images may be aided by the *appropriate* use of active, pulsed, or lock-in thermography (Ibarra-Castanedo et al., 2013). For example, the thermal contrast at Castello di Arco rendered multi-modal image matching redundant, even though apparent temperature differences were present across the frescoes. Exciting thermal contrast may not only improve the number of extractable features (<200 tie points within the TIR block), but may point to new areas of deterioration. In addition, although the proposed work used Agisoft Metashape’s own ‘mosaic’ texturing approach, the ability to control texturing strategy (Biswanath et al., 2023; Hoegner et al., 2007), enhancement (Zhang et al., 2023b) and visualisation (Luhmann et al., 2013; Vidas et al., 2013) is a critical venture for IRT-3DDF, with the temperature metrics in Table 5.4.2 pointing to the ambiguity of procedures unsuited to TIR images. To conclude, the novelty of IRT-3DDF, residing in the transformation of temperature into observable, quantifiable products, must be placed at the heart of future methods, acknowledging the challenges, compromises and opportunities that come from handling temperature photogrammetrically.

6.5 Thermal InfraRed Cameras: Resolutions, Specifications & Future Developments

The development of commercial TIR cameras (and by extension FPA sensors) has been central to the advancement of IRT-3DDF, enabling TIR images to be used photogrammetrically in processes common for visible spectrum imagery. Critically, the use of vanadium oxide (VOx) for thermal detectors, the predominant material for commercial microbolometers in FPA sensors, is enabling: greater sensitivity to noise equivalent temperature differences (0.039 K for VOx); reductions in sensor material and production costs; greater applicability for ambient temperature environments (<350 °C); fewer demands on system cooling; and (most importantly) the continued

reduction in pixel size (Lagüela et al., 2016; Rogalski, 2012; Vollmer & Möllmann, 2018). Although the likes of quantum detectors and silicon microbolometers are regularly discussed as replacements for VOx (Corsi, 2015; Gade & Moeslund, 2014), the central features of new TIR cameras for IRT-3DDF will be threefold: (1) improvements in image resolution (both spatial and thermal), (2) reductions in overall cost, and (3) improvements in internal cooling (i.e., the minituration of liquid nitrogen cyro-cooling)(Khodayar et al., 2016). However, in practice, the greater *accessibility* of features *within* TIR cameras, including native access to: FPA sensor and CPU temperature; NUCs and their associated correction images; and amendable raw temperature files are desired features that could have considerable impact on the speed at which IRT-3DDF methods can progress.

Importantly, the ever-increasing resolution of TIR cameras remains the biggest enabler of successful data fusion methods, allowing greater information to be used for feature matching, thermal modelling and quantitative analysis. For detailed thermographic surveys of buildings, the UK Thermography Association (UKTA) recommends a thermal camera with a sensor $>640 \times 480$ pix (Colyer & Arnolds, 2019; Fox et al., 2014). For IRT-3DDF, Adamopoulos et al. (2020b) demonstrate the effectiveness of ‘mid-range’ TIR cameras by validating their SfM pipeline, developed with a FLIR T1030c (1024×768 pix), with a FLIR SC660 (640×480 pix) thermal camera, resulting in similar projection errors to their higher resolution TIR camera (2-3 cm as opposed to 1 cm). Although thermal models have been generated with remarkably low-resolution TIR cameras (Borrmann et al., 2013; Javadnejad et al., 2020; Jones, 2001), mid-range thermal cameras, those with a pixel count in excess of 300,000 pixels, provide the greatest cost-benefit for data-fusion methodologies. However, as with the trade-offs seen in Section 3.5, the choice of TIR camera must be made acknowledging the desired: IRT-3DDF method, additional integrated sensors, geometric and radiometric accuracy, coverage and mobility, and camera budget.

Finally, the growth of the commercial TIR camera market has been shared by the

emergence of Three-dimensional Thermal Imaging Systems (TTISs), previously defined as “platforms comprising multiple sensors on fixed or ‘dynamic’ units” [p. 32]. The introduction of this term as part of this thesis has looked to highlight the possibilities exploiting hardware where multiple sensors are housed on a single platform, providing constraints that historically had to be configured manually (Alba et al., 2011). As the availability of terrestrial laser scanners with integrated TIR sensors becomes more common for IRT-3DDF (inc., Leica’s BLK360 G1 160×120 pix 71°×56° panoramic TIR camera (Adán et al., 2023a) and Zoller+Fröhlich’s 382×288 pix modular T-Cam (Fang et al., 2022)), it is apparent that the integration of low-resolution TIR sensors within commercial platforms may lead to greater automation and application in the coming years. In addition, the prevalence of UAVs with in-built thermal cameras (i.e., FLIR SIRAS, Autel Evo MAX T4 and DJI’s T series) or those capable of taking thermal payloads (i.e., Workswell’s catalogue or DJI’s Zenmuse series) is allowing for vast and complete data capture of both buildings (Paziewska & Rzonca, 2022) and landscapes (Wakeford et al., 2019). Therefore, attention must be paid to the future of TTISs, assessing their utility as a single hardware choice and their complete automation for IRT-3DDF.

6.6 Multi-Modal Image Matching for IRT-3DDF

Regardless of their classification, two central issues have hampered the development of fusion methods *across* IRT-3DDF: the inherent low resolution of TIR sensors and the inability to extract correspondences suitable for multi-modal data fusion. Excitingly, instead of finding ways around these issues, either by reducing the automation of derived methods or reducing the extractable value from acquired data, the proposed method breaks this trend by confronting both challenges openly. The exploration of multi-modal image matching, using DL-based models far from their intended domain, operation or application, points not only to the success of its

adoption for IRT-3DDF, but the merits of looking to tandem fields to solve established challenges. Notably, this adoption is done in the absence of preferential implementation for learning-based approaches, in a space where training and tuning have become ubiquitous. Much like the proposed methods, demonstrating the initial exploration of a tool that overcomes the central challenges faced by *all* IRT-3DDF research, future methods need to investigate solutions irrespective of initial or ideal implementation, with tandem fields offering opportunities to adopt procedures that can be subsequently refined and improved. As a result, the derived method can be readily-applied to varying multi-modal, multi-temporal and multi-sensor fusion methods, whether it be for visible, thermal or other multi-spectral dataset.

6.6.1 IRT-3DDF: Training, Tuning & Testing

However, whilst the success of the proposed method is in the application of neural networks *without* any training or tuning, this is not to dismiss the future opportunities that could be presented by preparing thermal-specific DL-based image matchers, a practice that has facilitated considerable advances in remote sensing, medical or computer vision fields (Jiang et al., 2021). For training and tuning of multi-modal neural networks to become a reality, vast and diverse datasets of thermal images are required, representative of a wealth of subjects and scenes. For example, Superpoint was a pre-trained on MagicPoint’s synthetic shape dataset to teach it to detect geometric features such as corners and edges ($\sim 100,000$ greyscale shapes)(DeTone et al., 2018), subsequently fine-tuned using Microsoft’s comprehensive ‘Common Objects in Context’ (COCO) dataset of $\sim 300,000$ images ($\sim 200,000$ segmented and labelled)(Lin et al., 2014). However, even though the collection of such data would require extensive curation and annotation, and the ability to compile the necessary number of images for training may not be achieved (even with data augmentation)((Velesaca et al., 2024)), bringing thermal datasets of architectural heritage

together represents a sensible endeavour for IRT-3DDF nonetheless.



(a) ARCO



(b) GATE



(c) DUNR



(d) WIND

Figure 6.5: Thermitage’s datasets, including: (a) Arco Castle, Italy (ARCO); (b) Lenton Lodge Gatehouse, UK (GATE); (c) Dunrobin House, UK (DUNR); and (d) Chesterton Windmill, UK (WIND).

To this end, in order to begin the progress of building such a collection, this PhD has introduced *Thermitage*, a curated repository of multi-modal, multi-temporal and multi-sensor thermal images tailored towards architectural heritage (Sutherland et al., 2025b). The ambition for *Thermitage*, sharing an initial collection of IRT-3DDF datasets collected during this PhD, is to provide a collaborative repository for the generation of training and validation data for learning-based benchmarking (Figure 6.5 & Table 6.3). In addition, the applications of *Thermitage* include: (1) the development of new fusion methods utilising *Thermitage*’s datasets; (2) the use of *Thermitage* to validate or compare derived methods across different case studies; and (3) the re-purposing of derived *Thermitage*’s 3D thermal products for additional

investigations. Although the quantity of images needed for deep learning-based training and tuning is not currently facilitated by Thermitage, the ability to develop machine learning applications is more than achievable with the presented datasets, including the likes of image segmentation (Lin et al., 2019a), subject detection (Motayyeb et al., 2023) and NeRF (Lin et al., 2024). In summary, as learning-based tools have been critical for the development of the proposed novel IRT-3DDF method, the facilitation of additional methods using common approaches is an exciting venture for the future of thermal modelling.

Datasets	Sensors			Surveys		
	TIR	VIS	RGB	UAV	TLS	SLAM
ARCO	✓	✓	✓	—	—	✓
GATE	✓✓	✓✓	✓✓	—	✓	—
DUNR	✓✓	✓✓	✓✓	✓	—	—
WIND	✓	✓	✓✓	✓✓	—	—

Table 6.3: Thermitage’s datasets, sensors, and surveys.

6.6.2 Multi-Modal & Multi-Sensor Calibration

The importance of accurate intrinsics for IRT-3DDF, emphasised by the inclusion of Objective 2 dedicated to calculating the interior- and relative orientation of IRT-3DDF sensors, has been placed at the forefront of the proposed research, understanding the significant role of camera calibration in a range of data fusion methods. When achieved through self-calibration, we have observed the failure for accurate camera parameters to be obtained (Section 4.4.1 & 4.4.2), reinforcing the necessity for TIR cameras (featuring highly unique germanium lenses) to be calibrated prior to photogrammetric data fusion. In addition, we have found that even when pre-calibrated IO values are assigned at the initialisation of a bundle adjustment, values are refined to *absorb* resulting error rather than *account* for it. Whilst the prospects of on-site self-calibration are appealing, and were deemed successful in Chapter 5, this thesis reiterates the opportunities that can come from precise

IO and RO calibration, notably the inclusion of tandem sensors within IRT-3DDF pipelines. The intersection between geometric camera calibration and IRT-3DDF must be expanded in future work, notably: (1) the execution of bundle adjustments with *constrained* IO and RO parameters (not possible in COLMAP or Metashape, but achievable in Australis); (2) the processing of TIR, VIS *and* RGB images in a singular bundle adjustment with / without RO constraints; and (3) the inclusion of multi-modal correspondence within geometric camera calibration set-ups.

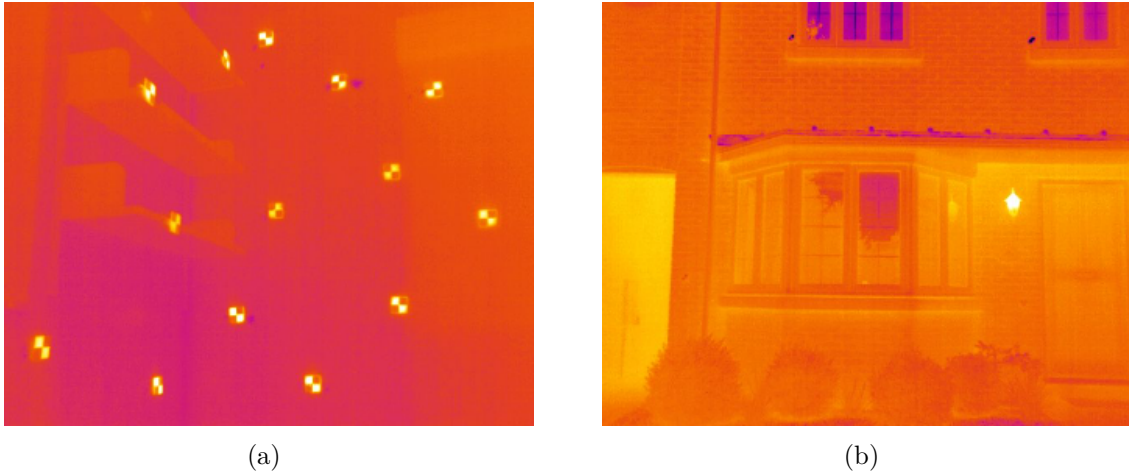


Figure 6.6: Examples TIR images of the thermal-specific markers, featuring the (a) 3D test-field from Section 4.3.3 and (b) Dunrobin House.

To facilitate the geometric calibration for TIR cameras, this work designed configurable thermal-specific markers, a significant achievement of the proposed work over existing superficial and cursory solutions (Section 3.5.4). As previously noted, the novelty of the proposed markers, designed to be low-cost, easy-to-manufacture and highly configurable, is in: (1) their use as configurable calibration markers for lab- and field-based geometric calibration procedures, (2) their use as survey markers for photogrammetric scaling, referencing and validation, and (3) their use as ground truth for radiometric calibration, correction and verification (Figure 6.6). To demonstrate this value, Figure 6.7 & Table 6.4 show the geometric accuracy of a TIR-only bundle adjustment of DUnrobin House checked against either the proposed thermal-specific markers and ‘natural’ markers, showing the variance in 2D-

(MRE) and 3D- (RMSE) space. Both the accuracy of geodetic survey ($\mu\sigma$) and the accuracy of image orientation (RMSE) highlight the ambiguity that can come from the use of natural markers within IRT-3DDF, typically where materials of different emissivities meet, such as windows, doors or corners. Whilst future iterations will look to improve: the automation of marker extraction, the performance of markers in low contrast settings, and the robustness of design and manufacturing, this product represents an innovative solution that can be readily-adopted for thermal surveying.

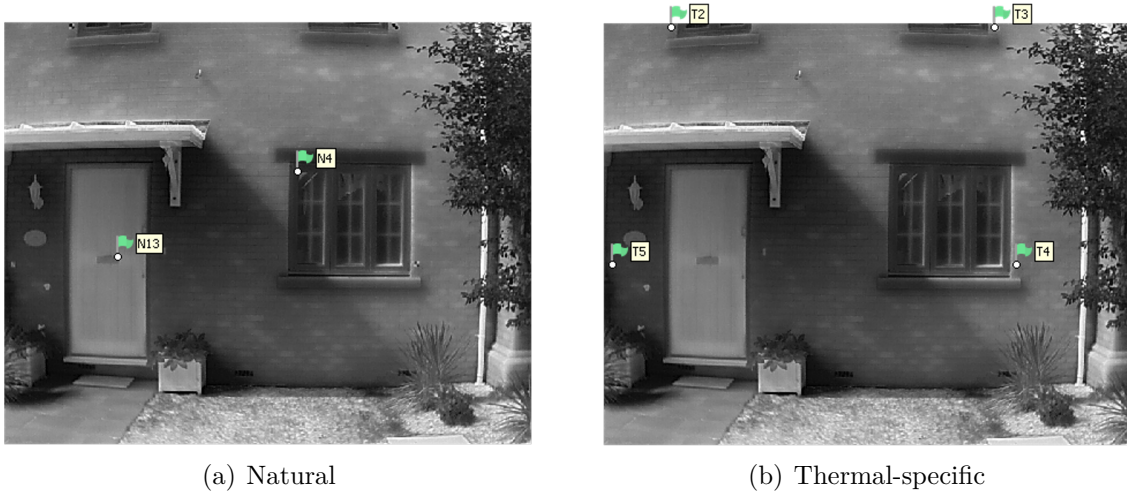


Figure 6.7: An example TIR image of Dunrobin House with annotated (a) ‘natural’ and (b) thermal-specific markers.

Markers	No. CPs	TST $\mu\sigma$ (mm)	MRE (pix)	RMSE (mm)
Thermal-specific	12	0.50	0.28	5.60
Natural	11	1.02	0.39	34.26

Table 6.4: Comparison between TIR-only image orientation checked against ‘natural’ or thermal-specific markers on Dunrobin House. Here, the standard deviation of Leica TS10 total station measurements (TST), MRE and RMSE of the oriented image block are given.

6.6.3 Thermal Modelling: Poses, Geometries & Textures

Finally, it is important to stress that the proposed 2D-2D fusion workflow provides considerable opportunities for future refinement, with the derived thermal models representing the introduction of a novel IRT-3DDF pipeline rather a definitive conclusion. As the proposed method combines several sequential processing steps, the siloing of each procedure in Section 5.3 can enable individual amendments to be identified for the improvement of the workflow in its entirety. Firstly, the central novelty, achievement and *priority* of the proposed IRT-3DDF pipeline was in the determination of highly-accurate multi-modal camera poses, without which subsequent thermal modelling steps could not be undertaken successfully. For the best-performing image matchers across Lenton Lodge Gatehouse and Dunrobin House, TIR blocks achieved a RMSE of ~ 10 mm (Table 5.5), exploiting the full value of both TIR and RGB images in the optimised bundle adjustments. Whilst future evolutions of the proposed method will look to refine base geometry modelling and thermal texturing (and even re-purpose derived camera poses in supplementary IRT-3DDF methods)(Figure 6.8), these cannot be undertaken without the utmost confidence in the derived TIR camera poses. As the starting point for the numerous procedures involved in thermal modelling, considerable focus should be placed on the calculation of accurate camera poses within an IRT-3DDF pipeline, including methods to inspect, refine and validate resulting camera networks (Iwaszczuk & Stilla, 2017; Lin et al., 2019b).

By extension, the use of these cameras poses for dense reconstruction highlights several nuisances, with the ‘base’ geometry not only providing the means for *3D* modelling, but a gateway to further analysis. Firstly, it must be noted the presented works, all undertaken terrestrially, were limited to façade modelling only, with upscaling to full buildings an exciting and necessary test for the proposed method (Sutherland et al., 2025b). Currently, the base geometry is obtained in

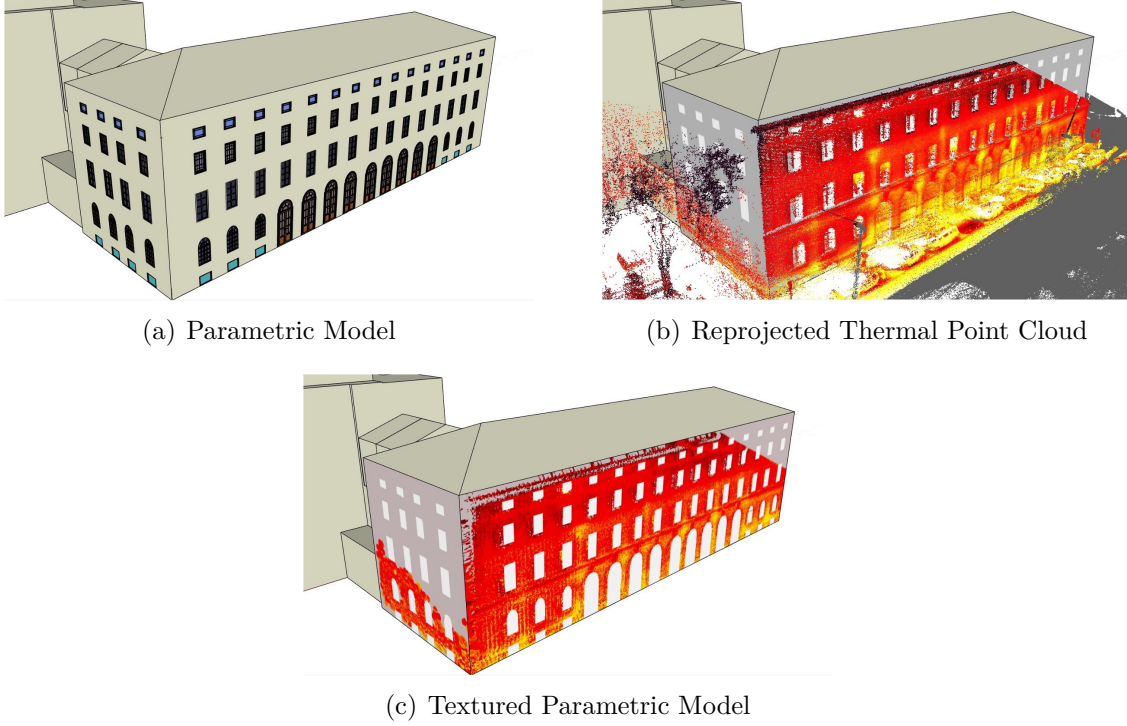


Figure 6.8: An example of thermally-textured parametric models, providing the means for thermal images to be integrated into (H)BIM and asset management processes (Biswanath et al., 2023).

a separate SfM and MVS pipeline, imported into the same 3D space as the derived camera poses. Although several methods have looked to compare derived data fusion geometries to determine their deviation from known ground truth (typically high-resolution RGB- or TLS models)(Adamopoulos et al., 2020b; Patrucco et al., 2022b), this would have been redundant for the proposed method as the derived base geometry did not encapsulate any information from the TIR image blocks. Therefore, to optimise the proposed pipeline, the use of the derived RGB poses for the base geometry (those obtained *in tandem* with the TIR image block) and the comparison of derived 3D models against a reference geometry warrants investigation, removing a tangential operation and progressing towards complete automation.

It seems appropriate to conclude the critique of the proposed pipeline by evaluating texturing, which perfectly epitomises the complexity of 3D thermal models

and the interconnectivity of radiometric and photogrammetric practices. Whether used for simple visualisations or the transformation of ortho-rectified images into additional IRT-3DDF products, a coherent texture not only represents the observed asset precisely, but provides accurate *quantitative* temperature measurements of the scene. Champions of texturing approaches across IRT-3DDF, Technical University of Munich (TUM) have demonstrated (through exhaustive case studies) the need for texturing methods to be placed at the forefront of data fusion pipelines, custom-built to facilitate particular sensors and scenes (Hoegner et al., 2007; Hoegner & Stilla, 2016; Zhu et al., 2021). At present, automatic texturing approaches, those found in commercial SfM software, are unsuited for highly-accurate thermal modelling, regardless of texturing / mosaicking approach (Lin et al., 2018a; Malbêteau et al., 2021). Furthermore, if (as previously suggested) scene-specific drift correction methods are to be implemented for IRT-3DDF, the effect these will have on resulting textures needs to be comprehensively tested against known ground truth (Mesas-Carrascosa et al., 2018).

However, even though the development of *quantitative* thermal textures is preferable, the generation of thermal textures capable of aiding *qualitative* interpretation must also be appreciated. Specifically, considerable opportunities reside in the fashioning of textures *primed* for novel visualisations, including the use of: enhanced (Zhang et al., 2023b), augmented (Lagüela et al., 2012), hybrid (Vidas et al., 2013), segmented (Scaioni et al., 2012), thresholded (Weinmann et al., 2014), stacked (Solla et al., 2020) and simulated (Natephra et al., 2017) thermal textures. In addition, the use of qualitative thermal textures for the production of 2D orthomosaics or rectified thermograms, often preferred as a means of visualisation by heritage conservators (Historic England, 2025b), suggest the suitability of IRT-3DDF products amended for more practical purposes. Whilst this thesis has stopped short of developing a fully-tailored texturing procedure, the evolution of this work must look to calculate and validate radiometric texturing approaches against known in-situ

temperature measurements, with the accurate amalgamation of 2D pixel- into 3D texel values a considerable ongoing challenge (Biswanath et al., 2023).

6.7 IRT-3DDF for Architectural Heritage

The purpose of CRM practices within architectural heritage has long been to prioritise *proactive* maintenance over *reactive* repairs (Forster & Kayan, 2009). This philosophy, central to NDT techniques such as IRT, has enabled new technologies to become established tools for heritage documentation (i.e., H-BIM), demonstrating their inherent value when employed for specific challenges (i.e., interoperability, lifecycle management, simulation). To understand how IRT-3DDF might emulate this trend, it is pivotal to understand not only the value it provides in line with this stated philosophy, but how it can provide value addressing the pressing challenges facing architectural heritage. Whilst the ambition for IRT-3DDF is to inform energy efficiency, retrofitting and thermal comfort, having been openly discussed both in this thesis and among those championing the field (Adán et al., 2020), it is only through purposeful case studies that assessments of its value can be undertaken. Furthermore, it is evident that a critical barrier for novel technologies such as H-BIM has been their effective demonstration and operation, with modern technologies cautiously adopted and remaining secondary to established traditional techniques. Therefore, the ability for IRT-3DDF to generate products capable of being used by (and tailored towards) conservators, heritage professionals and decision-makers needed to be investigated. If IRT-3DDF is to become more than an academic discipline, instilling meaningful insights as a critical tool in the arsenal of conservators, the value of IRT-3DDF should feature at the centre of future methods, demonstrating and evaluating its benefit aiding architectural heritage.

The differentiation between IRT-3DDF in the AECO industry and in architectural

heritage rests on one essential factor: the importance of significance within heritage. At the heart of architectural heritage lies the significance that a building holds and the need for both its tangible and intangible heritage to be preserved. This significance is not just encapsulated within an asset’s history, but the influence this history has had on a building’s current fabric and function. The proposed work affirms that a detailed knowledge of a historic building and its significance are fundamental for successful IRT-3DDF, emphasising the importance of scale, style, setting on available methods (as shown by the inclusion of Appendix A as a means of providing such knowledge). When studies have exploited IRT-3DDF with a deeper understanding of the architectural heritage at hand, data fusion methods are developed that maximise the benefits of individual datasets and enhance CRM practices (Sutherland et al., 2023). Whilst the challenge of this work is in the development of a generalisable method of IRT-3DDF, we must acknowledge that significance must be held central to each case of architectural heritage. By developing and applying methods with significance in mind, IRT-3DDF can become an innovative approach for monitoring architectural heritage.

Chapter 7

Conclusion

As the challenges facing architectural heritage mount, the combination of multi-disciplinary surveys will continue to be the bedrock of effective conservation, repair and maintenance practices. In a bid to develop a novel tool in the arsenal of heritage conservators, this thesis has introduced InfraRed Thermography 3D-Data Fusion (IRT-3DDF), a field generating geometrically and radiometrically accurate 3D thermal models representative of historic buildings. The aim of this thesis, to develop a fully-automatic IRT-3DDF method for 3D thermal modelling, has been achieved through the introduction of learning-based multi-modal image matching for sensor fusion, applying pre-trained neural networks outside of their intended training domain. By determining multi-modal correspondences across image modalities, subsequently included as part of a combined bundle block adjustment, TIR images were co-oriented in fully-automatic processes across varying scales, settings, sensors and subjects. Importantly, this is realised without: sensor pre-calibration or fixed RO constraints; manual alignment, co-registration or ICP; known positioning, orientation or reference data; or simultaneous data capture. As a result, the derived method can be readily-applied to different multi-modal, multi-temporal and multi-sensor use cases, expanding the value that can be extracted from these novel

3D products.

7.1 Thesis Aims & Objectives: Revisited

Whilst the core originality of this thesis is in the development of a novel 2D-2D fusion method; critically, this work addresses several under-represented and under-researched facets for effective IRT-3DDF. Revisiting the central objectives of this research (Section 1.1), this thesis set out to achieve the following objectives:

- **Objective 1:** Define the essential theoretical and practical requirements for the successful implementation of IRT-3DDF.
- **Objective 2:** Determine robust and rigorous methods for the geometric and radiometric calibration of TIR cameras employed within photogrammetric data fusion
- **Objective 3:** Establish a novel, cost-effective and highly-flexible data fusion method obviating sensor pre-calibration, manual co-registration or auxiliary georeferencing information.

7.1.1 Objective 1

Firstly, through the conceptualisation of IRT-3DDF as a discipline, denoting the various synergies between configured sensors and fusion ‘spaces’, this thesis has outlined the decisions that need to be made regarding methodological assumptions, shortcomings and mitigations. This has allowed several data fusion pipelines within the thesis to maximise the inherent quality of individual sensors, implementing fusion methods that strategically combine datasets in accordance with known limitations and constraints. In addition, an appreciation for IRT as a science has been

placed at the centre of all proposed experiments, surveys and methods, fundamental to understanding the detection of infrared radiation, the functions of TIR cameras and the extraction of reliable temperature measurements. Finally, by providing a ‘cheatsheet’ for those ready to adopt IRT-3DDF, this thesis outlines critical areas that need to be considered for the implementation of successful data fusion, notably the trade-offs between radiometric and photogrammetric processes.

7.1.2 Objective 2

Secondly, in a bid to evaluate the importance of interior- (IO) and relative- (RO) orientation parameters for successful sensor fusion, this thesis has presented a first-of-its-kind comparative analysis of geometric calibration approaches for TIR cameras. To undertake this comprehensive evaluation, this work presented a novel 3D test field composed of custom thermal-specific markers, capable of providing: (1) crosshair targets for geometric calibration; (2) additional tie points for photogrammetric multi-modal image matching; (3) GCPs and CPs for photogrammetric evaluation; and (4) known points for quantitative temperature verification. In addition, this work has developed a procedural workflow for calibrating, correcting and validating temperature information from commercial TIR cameras, a process that can be readily improved with greater access to calibration blackbodies, FPA sensor temperature and field radiometers.

7.1.3 Objective 3

Finally, to develop a novel, cost-effective and highly-flexible method for IRT-3DDF, this thesis presented the first example of 2D-2D fusion using learning-based multi-modal image matching within IRT-3DDF, overcoming the long-standing issues of low spatial resolution and poor correspondences when matching thermal and visible

spectrum images. Importantly, this was done using ‘out-of-the-box’ neural networks without training or tuning, significantly reducing the preparation and complexity of fusion processes. The derived methods are not only fully-automatic, but provide a workflow that can be readily-applied for various sensor configurations, architectural styles and environmental constraints. As a result, 3D thermal models encapsulating both a detailed base geometry and quantitative temperature values are generated, providing a novel product for the analysis of a building’s overall performance.

7.2 Future Work

With ever-increasing resolutions of commercial TIR cameras, and the growing market for mobile and aerial TTISs, the continued development of IRT-3DDF must look to determine new fusion methods tailored towards specific sensors and scenarios. For the evolution of the proposed work, the generalisability of the novel IRT-3DDF method must be tested across different datasets, with a seamless interface representing a realistic development for future adoption and automation. In addition, the integration of *known* temperature measurements within radiometric calibration, correction and validation procedures will add critical rigour to the quantitative analysis of derived 3D thermal models, increasing the confidence and reliability with which derived products can be applied. Finally, an evaluation of the inherent *value* of current 3D thermal models for CRM practices, as well as an understanding of the additional resources that can be extracted from resulting models, represents a necessary progression for future research.

Bibliography

- Abolt, C., Caldwell, T., Wolaver, B., & Pai, H. (2018). Unmanned Aerial Vehicle-based Monitoring of Groundwater Inputs to Surface Waters using an Economical Thermal Infrared Camera. *Optical Engineering*, 57(5)(053113). DOI: 10.1117/1.OE.57.5.053113
- Acorsi, M. G., Gimenez, L. M., & Martello, M. (2020). Assessing the Performance of a Low-Cost Thermal Camera in Proximal and Aerial Conditions. *Remote Sensing*, 12(3591). DOI: 10.3390/rs12213591
- Adamopoulos, E., Colombero, C., Comina, C., Rinaudo, F., Volinia, M., Girotto, M., & Ardissono, L. (2021). Integrating Multiband Photogrammetry, Scanning, and GPR for Built Heritage Surveys: The Façades of Castello del Valentino. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, VIII-M-1-2021, 1–8. DOI: 10.5194/isprs-annals-VIII-M-1-2021-1-2021
- Adamopoulos, E., & Rinaudo, F. (2019). 3D Interpretation and Fusion of Multidisciplinary Data for Heritage Science: A Review. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W15, 17–24. DOI: 10.5194/isprs-archives-XLII-2-W15-17-2019
- Adamopoulos, E., Rinaudo, F., Volinia, M., & Girotto, M. (2020a). Multispectral Sensing and Data Integration for the Study of Heritage Architecture. *Engineering Proceedings*, 2(64). DOI: 10.3390/ecsa-7-08198

- Adamopoulos, E., Volinia, M., Girotto, M., & Rinaudo, F. (2020b). Three-Dimensional Thermal Mapping from IRT Images for Rapid Architectural Heritage NDT. *Buildings*, *10*(187). DOI: 10.3390/buildings10100187
- Adán, A., Prado, T., Prieto, S. A., & Quintana, B. (2017). Fusion of Thermal Imagery and LiDAR Data for Generating TBIM Models. *IEEE SENSORS 2017*, 1–3. DOI: 10.1109/ICSENS.2017.8234261
- Adán, A., López-Rey, A., & Ramón, A. (2023a). Robot for Thermal Monitoring of Buildings. *Automation in Construction*, *154*(105009). DOI: 10.1016/j.autcon.2023.105009
- Adán, A., Pérez, V., Ramón, A., & Castilla, F. J. (2023b). Correction of Temperature from Infrared Cameras for More Precise As-Is 3D Thermal Models of Buildings. *Applied Sciences*, *13*(6779). DOI: 10.3390/app13116779
- Adán, A., Pérez, V., Vivancos, J. L., Aparicio-Fernández, C., & Prieto, S. A. (2021). Proposing 3D Thermal Technology for Heritage Building Energy Monitoring. *Remote Sensing*, *13*(1537). DOI: 10.3390/rs13081537
- Adán, A., Quintana, B., García Aguilar, J., Pérez, V., & Castilla, F. J. (2020). Towards the Use of 3D Thermal Models in Constructions. *Sustainability*, *12*(8521). DOI: 10.3390/su12208521
- Agisoft Metashape. (2024). Agisoft Metashape Professional. Retrieved September 22, 2025, from Available from: <http://www.agisoft.com>
- Akcay, O., & Avsar, E. O. (2017). The Effect of Image Enhancement Methods During Feature Detection and Matching of Thermal Images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-1/W1*, 575–578. DOI: 10.5194/isprs-archives-XLII-1-W1-575-2017
- Alba, M. I., Barazzetti, L., Scaioni, M., Rosina, E., & Previtali, M. (2011). Mapping Infrared Data on Terrestrial Laser Scanning 3D Models of Buildings. *Remote Sensing*, *3*(9), 1847–1870. DOI: 10.3390/rs3091847

- Altieri, G., Maffia, A., Pastore, V., Amato, M., & Celano, G. (2022). Use of High-Resolution Multispectral UAVs to Calculate Projected Ground Area in *Corylus avellana* L. Tree Orchard. *Sensors*, *22*(7103). DOI: 10.3390/s22197103
- Aragon, B., Johansen, K., Parkes, S., Malbeteau, Y., Al-Mashharawi, S., Al-Amoudi, T., Andrade, C. F., Turner, D., Lucieer, A., & McCabe, M. F. (2020). A Calibration Procedure for Field and UAV-Based Uncooled Thermal Infrared Instruments. *Sensors*, *20*(3316). DOI: 10.3390/s20113316
- Ashurst, J., & Dimes, F. G. (Eds.). (1998). *Conservation of Building and Decorative Stone* (1st ed.). Butterworth-Heinemann.
- ASTM International. (2015). Practice for Location of Wet Insulation in Roofing Systems Using Infrared Imaging. DOI: 10.1520/C1153-10R15
- ASTM International. (2022). Practice for Measuring and Compensating for Emissivity Using Infrared Imaging Radiometers. DOI: 10.1520/E1933-14R18
- Aubrecht, D. M., Helliker, B. R., Goulden, M. L., Roberts, D. A., Still, C. J., & Richardson, A. D. (2016). Continuous, Long-Term, High-Frequency Thermal Imaging of Vegetation: Uncertainties and Recommended Best Practices. *Agricultural and Forest Meteorology*, *228-229*, 315–326. DOI: 10.1016/j.agrformet.2016.07.017
- Balaras, C., & Argiriou, A. (2002). Infrared Thermography for Building Diagnostics. *Energy and Buildings*, *34*(2), 171–183. DOI: 10.1016/S0378-7788(01)00105-0
- Barbieri, E., Franzoni, E., Lambertini, A., Pizzigatti, C., Trevisiol, F., & Bitelli, G. (2023). 3D Data Management and Thermographic Studies as a Knowledge Base for the Conservation of a Rationalist Architecture. *The Future of Heritage Science and Technologies Materials Science*, 3–15. DOI: 10.1007/978-3-031-17594-7_1
- Barr, E. (1962). The Infrared Pioneers—II. Macedonio Melloni. *Infrared Physics*, *2*(2), 67–74. DOI: 10.1016/0020-0891(62)90023-4

- Barriera, E., de Freitas, V., Delgrado, J., & Ramos, N. (2012). *Thermography Applications in the Study of Buildings Hygrothermal Behaviour, Infrared Thermography* (R. Prakash, Ed.). INTECH Open Access Publisher.
- Barroso-Laguna, A., Riba, E., Ponsa, D., & Mikolajczyk, K. (2019). Key.Net: Key-point Detection by Handcrafted and Learned CNN Filters. DOI: 10.48550/ARXIV.1904.00889
- Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded Up Robust Features. In A. Leonardis, H. Bischof, & A. Pinz (Eds.), *Computer Vision – ECCV 2006* (pp. 404–417). Springer-Verlag. DOI: 10.1007/11744023_32
- Becquerel, A. (1839). Recherches sur les Effets de la Radiation Chimique de la Lumiere Solaire au Moyen des Courants Electriques. *Comptes Rendus de L’Academie des Sciences*, 9, 145–149.
- Bellavia, F., Morelli, L., Menna, F., & Remondino, F. (2022). Image Orientation with a Hybrid Pipeline Robust to Rotations and Wide-Baselines. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-2/W1-2022, 73–80. DOI: 10.5194/isprs-archives-XLVI-2-W1-2022-73-2022
- Berni, J., Zarco-Tejada, P., Suarez, L., & Fereres, E. (2009). Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring From an Unmanned Aerial Vehicle. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3), 722–738. DOI: 10.1109/tgrs.2008.2010457
- Biscarini, C., Catapano, I., Cavalagli, N., Ludeno, G., Pepe, F., & Ubertini, F. (2020). UAV Photogrammetry, Infrared Thermography and GPR for Enhancing Structural and Material Degradation Evaluation of the Roman Masonry Bridge of Ponte Lucano in Italy. *NDT & E International*, 115(102287). DOI: 10.1016/j.ndteint.2020.102287
- Bison, P., Bortolin, A., Cadelano, G., Ferrarini, G., Furlan, K., & Grinzato, E. (2012). Geometrical Correction and Photogrammetric Approach in Thermo-

- graphic Inspection of Buildings. *11th International Conference on Quantitative InfraRed Thermography (QIRT 2012)*, 1–9. DOI: 10.21611/qirt.2012.285
- Biswanath, M. K., Hoegner, L., & Stilla, U. (2023). Thermal Mapping from Point Clouds to 3D Building Model Facades. *Remote Sensing*, 15(4830). DOI: 10.3390/rs15194830
- Bökman, G., & Kahl, F. (2022). A Case for Using Rotation Invariant Features in State of the Art Feature Matchers. *Computer Vision and Pattern Recognition (CVPR 2022)*. DOI: 10.48550/arXiv.2204.10144
- Boltzmann, L. (1884). Ableitung des Stefan’schen Gesetzes, betreffend die Abhängigkeit der Wärmestrahlung von der Temperatur aus der electromagnetischen Lichttheorie. *Annalen der Physik*, 258(6), 291–294. DOI: 10.1002/andp.18842580616
- Borrmann, D., Elseberg, J., & Nüchter, A. (2013). Thermal 3D Mapping of Building Façades. In S. Lee, H. Cho, K.-J. Yoon, & J. Lee (Eds.), *Intelligent Autonomous Systems 12* (pp. 173–182). Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-33926-4_16
- Borrmann, D., Nüchter, A., Dakulović, M., Maurović, I., Petrović, I., Osmanković, D., & Velagić, J. (2014). A Mobile Robot Based System for Fully Automated Thermal 3D Mapping. *Advanced Engineering Informatics*, 28, 425–440. DOI: 10.1016/j.aei.2014.06.002
- Boyle, W. S., & Smith, G. E. (1970). Charge Coupled Semiconductor Devices. *Bell System Technical Journal*, 49(4), 587–593. DOI: 10.1002/j.1538-7305.1970.tb01790.x
- British Standards Institute. (2017). ISO 10880:2017: Non-destructive Testing - Infrared Thermographic Testing - General Principles. Retrieved April 24, 2025, from Available from: [https://www.iso.org/standard/61881.html#:~:text=ISO%2010880%3A2017%20provides%20general,%2Ddestructive%20testing%20\(NDT\)](https://www.iso.org/standard/61881.html#:~:text=ISO%2010880%3A2017%20provides%20general,%2Ddestructive%20testing%20(NDT))

- Brooke, C. (2018). Thermal Imaging for the Archaeological Investigation of Historic Buildings. *Remote Sensing*, 10(1401). DOI: 10.3390/rs10091401
- Brown, D. C. (1971). Close-Range Camera Calibration. *Photogrammetric Engineering*, 37(8), 855–866.
- Brumana, R., Condoleo, P., Grimoldi, A., Banfi, F., Landi, A. G., & Previtali, M. (2018a). HR LOD based HBIM to Detect Influences on Geometry and Shape by Stereotomic Construction Techniques of Brick Vaults. *Applied Geomatics*, 10(4), 529–543. DOI: 10.1007/s12518-018-0209-3
- Brumana, R., Condoleo, P., Grimoldi, A., Landi, A. G., Attico, D., Turrina, A., Banfi, F., & Previtali, M. (2018b). HBIM Feeding Open Access Vault Inventory Through GeoDB HUB. In E. Fink, R. Brumana, A. Doulamis, P. Patias, M. Ioannides, & J. Martins (Eds.), *Lecture Notes in Computer Science* (pp. 27–38). Springer Cham. DOI: 10.1007/978-3-030-01762-0_3
- Brumana, R., Ioannides, M., & Previtali, M. (2019). Holistic Heritage Building Information Modelling (HHBIM): From Nodes to Hub Networking, Vocabularies and Repositories. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W11, 309–316. DOI: 10.5194/isprs-archives-XLII-2-W11-309-2019
- Budzier, H., & Gerlach, G. (2015). Calibration of Uncooled Thermal Infrared Cameras. *Journal of Sensors and Sensor Systems*, 4, 187–197. DOI: 10.5194/jsss-4-187-2015
- Cabrelles, M., Galcerá, S., Navarro, S., Lerma, J. L., Akasheh, T., & Haddad, N. (2009). Integration of 3D Laser Scanning, Photogrammetry and Thermography to Record Architectural Monuments. *Proceedings of the 22nd International CIPA Symposium*, 23–30.
- Cao, Y., & Tisse, C.-L. (2013). Shutterless Solution for Simultaneous Focal Plane Array Temperature Estimation and Nonuniformity Correction in Uncooled Long-Wave Infrared Camera. *Applied Optics*, 52(25)(6266). DOI: 10.1364/ao.52.006266

- Ceccarelli, S., Guarneri, M., Ferri De Collibus, M., Francucci, M., Ciaffi, M., & Danielis, A. (2018). Laser Scanners for High-Quality 3D and IR Imaging in Cultural Heritage Monitoring and Documentation. *Journal of Imaging*, 4(130). DOI: 10.3390/jimaging4110130
- Çengel, Y. A., Boles, M. A., & Kanoğlu, M. (2019). *Thermodynamics: An Engineering Approach* (9th ed.). McGraw Hill Education.
- Chang, R., Yang, C., Zhang, H., Xie, H., Zhou, C., Pan, A., & Yang, Y. (2024). Speed Matters, A Robust Infrared and Visible Image Matching Method at Real-Time Speed. *Journal of Real-Time Image Processing*, 21(15). DOI: 10.1007/s11554-023-01395-4
- Cho, Y. K., Ham, Y., & Golpavar-Fard, M. (2015). 3D As-Is Building Energy Modeling and Diagnostics: A Review of the State-of-the-Art. *Advanced Engineering Informatics*, 29, 184–195. DOI: 10.1016/j.aei.2015.03.004
- Chum, O., Werner, T., & Matas, J. (2005). Two-View Geometry Estimation Unaffected by a Dominant Plane. *Computer Vision and Pattern Recognition (CVPR 2005)*, 772–779. DOI: 10.1109/CVPR.2005.354
- Clark, R. (2006). House for Sale – or a Tourist Hotspot? *The Telegraph*. Retrieved January 4, 2024, from Available from: <https://www.telegraph.co.uk/finance/property/advice/propertymarket/3346948/House-for-sale-or-a-tourist-hotspot.html>
- Clarke, T. A., Wang, X., & Fryer, J. G. (1998). The Principal Point and CCD Cameras. *The Photogrammetric Record*, 16(92), 293–312. DOI: 10.1111/0031-868X.00127
- Clarkson, G., Luo, S., & Fuentes, R. (2017). Thermal 3D Modelling. *Proceedings of 34th International Symposium on Automation and Robotics in Construction (ISARC 2017)*, 493–499. DOI: 10.22260/ISARC2017/0068
- Collaro, C., Enríquez-Muñoz, C., López, A., Enríquez, C., & Jurado, J. M. (2023). Detection of Landscape Features with Visible and Thermal Imaging at the

- Castle of Puerta Arenas. *Archaeological and Anthropological Sciences*, 15(152). DOI: 10.1007/s12520-023-01831-3
- Colyer, M., & Arnolds, B. (2019). Building Thermography: Non-Destructive Building Testing. Retrieved February 4, 2024, from Available from: <https://ukta.org/wp-content/uploads/2023/02/building-thermography-ukta.pdf>
- Commercial News Media. (2016). Lenton Lodge Sold to University of Nottingham. *Commercial News Media*. Retrieved January 4, 2024, from Available from: <https://www.commercialnewsmedia.com/archives/43481>
- Conte, P., Girelli, V. A., & Mandanici, E. (2018). Structure from Motion for Aerial Thermal Imagery at City Scale: Pre-Processing, Camera Calibration, Accuracy Assessment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 320–333. DOI: 10.1016/j.isprsjprs.2018.10.002
- Corsi, C. (2015). New Frontiers for Infrared. *Opto-Electronics Review*, 23(1). DOI: 10.1515/oere-2015-0015
- Costa-Carrapiço, I., González, J. N., Raslan, R., & Sánchez-Guevara, C. (2022). Understanding the Challenges of Determining Thermal Comfort in Vernacular Dwellings: A Meta-Analysis. *Journal of Cultural Heritage*, 58, 57–73. DOI: 10.1016/j.culher.2022.09.019
- Costanzo, A., Minasi, M., Casula, G., Musacchio, M., & Buongiorno, M. (2014). Combined Use of Terrestrial Laser Scanning and IR Thermography Applied to a Historical Building. *Sensors*, 15(1), 194–213. DOI: 10.3390/s150100194
- Daakir, M., Zhou, Y., Pierrot Deseilligny, M., Thom, C., Martin, O., & Rupnik, E. (2019). Improvement of Photogrammetric Accuracy by Modeling and Correcting the Thermal Effect on Camera Calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 148, 142–155. DOI: 10.1016/j.isprsjprs.2018.12.012
- Dabetwar, S., Kulkarni, N. N., Angelosanti, M., Niezrecki, C., & Sabato, A. (2022). Sensitivity Analysis of Unmanned Aerial Vehicle-borne 3D Point Cloud Re-

- construction from Infrared Images. *Journal of Building Engineering*, 58(105070). DOI: 10.1016/j.jobbe.2022.105070
- Daffara, C., Muradore, R., Piccinelli, N., Gaburro, N., de Rubeis, T., & Ambrosini, D. (2020). A Cost-Effective System for Aerial 3D Thermography of Buildings. *Journal of Imaging*, 6(76). DOI: 10.3390/JIMAGING6080076
- Dahaghin, M., Samadzadegan, F., & Dadras Javan, F. (2019). 3D Thermal Mapping of Building Roofs Based on Fusion of Thermal and Visible Point Clouds in UAV Imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W18, 271–277. DOI: 10.5194/isprs-archives-XLII-4-W18-271-2019
- Dahaghin, M., Samadzadegan, F., & Dadras Javan, F. (2021). Precise 3D Extraction of Building Roofs by Fusion of UAV-based Thermal and Visible Images. *International Journal of Remote Sensing*, 42(18), 7002–7030. DOI: 10.1080/01431161.2021.1951875
- De Fino, M., Galantucci, R. A., & Fatiguso, F. (2023). Condition Assessment of Heritage Buildings via Photogrammetry: A Scoping Review from the Perspective of Decision Makers. *Heritage*, 6, 7031–7067. DOI: 10.3390/heritage6110367
- Delegou, E. T., Mourgi, G., Tsilimantou, E., Ioannidis, C., & Moropoulou, A. (2019). A Multidisciplinary Approach for Historic Buildings Diagnosis: The Case Study of the Kaisariani Monastery. *Heritage*, 2, 1211–1232. DOI: 10.3390/heritage2020079
- Denard, H. (2012). A New Introduction to the London Charter. In D. Bentkowska-Kafel & H. Denard (Eds.), *Paradata and Transparency in Virtual Heritage* (1st ed., pp. 57–71). Routledge.
- DeTone, D., Malisiewicz, T., & Rabinovich, A. (2018). SuperPoint: Self-Supervised Interest Point Detection and Description. *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 337–349. DOI: 10.1109/CVPRW.2018.00060

- Dias, I. S., Flores-Colen, I., & Silva, A. (2021). Critical analysis about emerging technologies for Building's façade inspection. *Buildings*, 11(2), 1–19. DOI: 10.3390/buildings11020053
- DJI. (2025). DJI Dock 3: Specs. Retrieved October 25, 2025, from Available from: <https://enterprise.dji.com/dock-3/specs>
- Dlesk, A., & Vach, K. (2019). Point Cloud Generation of a Building from Close Range Thermal Images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-5/W2, 29–33. DOI: 10.5194/isprs-archives-XLII-5-W2-29-2019
- Dlesk, A., Vach, K., & Pavelka, K. (2021). Transformations in the Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images. *Sensors*, 21(5061). DOI: 10.3390/s21155061
- Dlesk, A., Vach, K., & Pavelka, K. (2022). Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images. *Sensors*, 22(1655). DOI: 10.3390/s22041655
- Döpfer, V., Gränzig, T., Kleinschmit, B., & Förster, M. (2020). Challenges in UAS-Based TIR Imagery Processing: Image Alignment and Uncertainty Quantification. *Remote Sensing*, 12(1552). DOI: 10.3390/rs12101552
- du Châtelet, G.-É. L. T. d. B. (1744). *Dissertation sur la Nature et la Propagation du Feu*. Prault fils. DOI: [ark:/12148/cb30363541t](https://doi.org/10.3390/12148/cb30363541t)
- Dugdale, S. J. (2016). A Practitioner's Guide to Thermal Infrared Remote Sensing of Rivers and Streams: Recent Advances, Precautions and Considerations. *WIREs Water*, 3, 251–268. DOI: 10.1002/wat2.1135
- Dugdale, S. J., Kelleher, C. A., Malcolm, I. A., Caldwell, S., & Hannah, D. M. (2019). Assessing the Potential of Drone-based Thermal Infrared Imagery for Quantifying River Temperature Heterogeneity. *Hydrological Processes*, 33(7), 1152–1163. DOI: 10.1002/hyp.13395
- Duran, Z., & Atik, M. E. (2021). Accuracy Comparison of Interior Orientation Parameters from Different Photogrammetric Software and Direct Linear Trans-

- formation Method. *International Journal of Engineering and Geosciences*, 6(2), 74–80. DOI: 10.26833/ijeg.691696
- Dusmanu, M., Rocco, I., Pajdla, T., Pollefeys, M., Sivic, J., Torii, A., & Sattler, T. (2019). D2-Net: A Trainable CNN for Joint Detection and Description of Local Features [arXiv:1905.03561 [cs]]. *Computer Vision and Pattern Recognition (CVPR 2019)*. DOI: 10.48550/arXiv.1905.03561
- Edstedt, J., Bökman, G., Wadenbäck, M., & Felsberg, M. (2023a). DeDoDe: Detect, Don’t Describe – Describe, Don’t Detect for Local Feature Matching. *Computer Vision and Pattern Recognition (3DV 2024)*. DOI: 10.48550/ARXIV.2308.08479
- Edstedt, J., Sun, Q., Bökman, G., Wadenbäck, M., & Felsberg, M. (2023b). RoMa: Robust Dense Feature Matching. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.2305.15404
- El Masri, Y., & Rakha, T. (2020). A Scoping Review of Non-Destructive Testing (NDT) Techniques in Building Performance Diagnostic Inspections. *Construction and Building Materials*, 265(120542). DOI: 10.1016/j.conbuildmat.2020.120542
- Elfarkh, J., Johansen, K., Angulo, V., Camargo, O. L., & McCabe, M. F. (2023). Quantifying Within-Flight Variation in Land Surface Temperature from a UAV-Based Thermal Infrared Camera. *Drones*, 7(617). DOI: 10.3390/drones7100617
- Elias, M., Weitkamp, A., & Eltner, A. (2023). Multi-Modal Image Matching to Colorize a SLAM-based Point Cloud with Arbitrary Data from a Thermal Camera. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 9(100041). DOI: 10.1016/j.ophoto.2023.100041
- ElSheikh, A., Abu-Nabah, B. A., Hamdan, M. O., & Tian, G.-Y. (2023). Infrared Camera Geometric Calibration: A Review and a Precise Thermal Radiation Checkerboard Target. *Sensors*, 23(3479). DOI: 10.3390/s23073479
- Epic Games. (2025). RealityScan. Retrieved October 20, 2025, from Available from: <https://www.epicgames.com/realityscan>

- Ettxepare, L., Leon, I., Sagarna, M., Lizundia, I., & Uranga, E. J. (2020). Advanced Intervention Protocol in the Energy Rehabilitation of Heritage Buildings: A Miñones Barracks Case Study. *Sustainability*, *12*(6270). DOI: 10.3390/SU12156270
- European Statistical System (EuroStat). (2021). European Statistical System Population and Housing Census 2021. Retrieved February 17, 2025, from Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_and_housing_census_2021_%E2%80%93_overview
- Fang, J., Zhang, Y., Zhang, Y., Guo, H., & Sun, Z. (2022). High-Definition Survey of Architectural Heritage Fusing Multisensors—The Case of Beamless Hall at Linggu Temple in Nanjing, China. *Sensors*, *22*(3369), 1–16. DOI: 10.3390/s22093369
- Fischler, M. A., & Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Graphics and Image Processing*, *24*(6), 381–395. DOI: 10.1145/358669.358692
- Forster, A. M., & Kayan, B. (2009). Maintenance for Historic Buildings: A Current Perspective. *Structural Survey*, *27*(3), 210–229. DOI: 10.1108/02630800910971347
- Fox, M., Coley, D., Goodhew, S., & De Wilde, P. (2014). Thermography methodologies for detecting energy related building defects. *Renewable and Sustainable Energy Reviews*, *40*, 296–310. DOI: 10.1016/j.rser.2014.07.188
- Fraser, C. (2013). Automatic Camera Calibration in Close Range Photogrammetry. *Photogrammetric Engineering & Remote Sensing*, *79*(4), 381–388. DOI: 10.14358/PERS.79.4.381
- Fryskowska-Skibniewska, A., Delis, P., Kedzierski, M., & Matusiak, D. (2022). The Conception of Test Fields for Fast Geometric Calibration of the FLIR VUE PRO Thermal Camera for Low-Cost UAV Applications. *Sensors*, *22*(2468). DOI: 10.3390/s22072468

- Gade, R., & Moeslund, T. B. (2014). Thermal Cameras and Applications: A Survey. *Machine Vision and Applications*, 25(1), 245–262. DOI: 10.1007/s00138-013-0570-5
- GardaSee. (2025). The Castle of Arco. Retrieved October 1, 2025, from Available from: <https://www.garda-see.com/en/themes/culture-nature/castles/castle-of-arco>
- Garrido, I., Lagüela, S., Otero, R., & Arias, P. (2020a). Thermographic Methodologies used in Infrastructure Inspection: A Review: Post-Processing Procedures. *Applied Energy*, 266(114857). DOI: 10.1016/j.apenergy.2020.114857
- Garrido, I., Lagüela, S., & Arias, P. (2018). Infrared Thermography’s Application to Infrastructure Inspections. *Infrastructures*, 3(35). DOI: 10.3390/infrastructures3030035
- Garrido, I., Lagüela, S., Sfarra, S., & Arias, P. (2020b). Development of Thermal Principles for the Automation of the Thermographic Monitoring of Cultural Heritage. *Sensors*, 20(3392). DOI: 10.3390/s20123392
- Garrido, I., Solla, M., Lagüela, S., & Rasol, M. (2022). Review of InfraRed Thermography and Ground-Penetrating Radar Applications for Building Assessment. *Advances in Civil Engineering*, 2022(5229911). DOI: 10.1155/2022/5229911
- Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F., & Marín-Jiménez, M. (2014). Automatic Generation and Detection of Highly Reliable Fiducial Markers under Occlusion. *Pattern Recognition*, 47(6), 2280–2292. DOI: 10.1016/j.patcog.2014.01.005
- Gazzano, O., Chambon, M., Ferrec, Y., & Druart, G. (2024). Long-Term Radiometric Stability of Uncooled and Shutterless Microbolometer-Based Infrared Cameras. *Sensors*, 24(6387). DOI: 10.3390/s24196387
- Gonzales, E. M., Pando, M. A., Whelan, M., Fratta, D., Ramos, L. F., & Aguilar, R. (2019). Structural Evaluation and Preliminary Analysis of the St. Jerome Hall of the Church of the Nativity in Bethlehem. *11th International Confer-*

- ence on Structural Analysis of Historical Constructions (SAHC)*, 18, 1063–1072. DOI: 10.1007/978-3-319-99441-3_115
- González-Aguilera, D., Lagüela, S., Rodríguez-Gonzálvez, P., & Hernández-López, D. (2013). Image-based Thermographic Modeling for Assessing Energy Efficiency of Buildings Façades. *Energy and Buildings*, 6, 29–36. DOI: 10.1016/j.enbuild.2013.05.040
- González-Aguilera, D., Rodríguez-Gonzálvez, P., Armesto, J., & Lagüela, S. (2012). Novel Approach to 3D Thermography and Energy Efficiency Evaluation. *Energy and Buildings*, 54, 436–443. DOI: 10.1016/j.enbuild.2012.07.023
- Grechi, G., Fiorucci, M., Marmoni, G. M., & Martino, S. (2021). 3D Thermal Monitoring of Jointed Rock Masses Through Infrared Thermography and Photogrammetry. *Remote Sensing*, 13(957). DOI: 10.3390/rs13050957
- Güidi, N. E., Di Filippo, A., & Barba, S. (2025). High-Resolution 3D Thermal Mapping: From Dual-Sensor Calibration to Thermally Enriched Point Clouds. *Applied Sciences*, 15(10491). DOI: 10.3390/app151910491
- Guilbert, V., Antoine, R., Heinkelé, C., Maquaire, O., Costa, S., Gout, C., Davidson, R., Sorin, J.-L., Beaucamp, B., & Fauchard, C. (2020). Fusion of Thermal and Visible Point Clouds: Application to the Vaches Noires Landslide, Normandy, France. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2020, 227–232. DOI: 10.5194/isprs-archives-XLIII-B2-2020-227-2020
- Ham, Y., & Golparvar-Fard, M. (2013a). An Automated Vision-Based Method for Rapid 3D Energy Performance Modeling of Existing Buildings Using Thermal and Digital Imagery. *Advanced Engineering Informatics*, 27(3), 395–409. DOI: 10.1016/j.aei.2013.03.005
- Ham, Y., & Golparvar-Fard, M. (2013b). EPAR: Energy Performance Augmented Reality Models for Identification of Building Energy Performance Deviations Between Actual Measurements and Simulation Results. *Energy and Buildings*, 63, 15–28. DOI: 10.1016/j.enbuild.2013.02.054

- Hao, L., Herrera-Avellanosa, D., Del Pero, C., & Troi, A. (2020). What are the Implications of Climate Change for Retrofitted Historic Buildings? A Literature Review. *Sustainability*, 12(7557). DOI: 10.3390/su12187557
- Hardie, R. C., Hayat, M. M., Armstrong, E., & Yasuda, B. (2000). Scene-based Nonuniformity Correction with Video Sequences and Registration. *Applied Optics*, 39(8), 1241–1250. DOI: 10.1364/AO.39.001241
- Harguess, J., & Strange, S. (2014). Infrared Stereo Calibration for Unmanned Ground Vehicle Navigation. *The International Society for Optical Engineering 9084*, 9084 (90840S). DOI: 10.1117/12.2053537
- Hastedt, H., & Luhmann, T. (2015). Investigations on the Quality of the Interior Orientation and its Impact in Object Space for UAV Photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1/W4, 321–328. DOI: 10.5194/isprsarchives-XL-1-W4-321-2015
- Herschel, J. F. W. (1840). I. On the Chemical Action of the Rays of the Solar Spectrum on Preparations of Silver and Other Substances, both Metallic and Non-Metallic, and on Some Photographic Processes. *Philosophical Transactions of the Royal Society of London*, 130, 1–59. DOI: 10.1098/rstl.1840.0002
- Herschel, W. (1800). XIV. Experiments on the Refrangibility of the Invisible Rays of the Sun. *Philosophical Transactions of the Royal Society of London*, 90, 284–292. DOI: 10.1098/rstl.1800.0015
- Hess, M., Vanoni, D., Petrovic, V., & Kuester, F. (2015). High-Resolution Thermal Imaging Methodology for Non-Destructive Evaluation of Historic Structures. *Infrared Physics & Technology*, 73, 219–225. DOI: 10.1016/j.infrared.2015.09.020
- Hill, A. C., Laugier, E. J., & Casana, J. (2020). Archaeological Remote Sensing Using Multi-Temporal, Drone-Acquired Thermal and Near Infrared (NIR) Imagery: A Case Study at the Enfield Shaker Village, New Hampshire. *Remote Sensing*, 12(690). DOI: 10.3390/rs12040690

- Historic England. (2017). *Photogrammetric Applications for Cultural Heritage Guidance for Good Practice* (Guidance HEAG066). Historic England. York, United Kingdom.
- Historic England. (2018a). *3D Laser Scanning for Heritage* (Guidance HEAG155). Historic England. York, United Kingdom.
- Historic England. (2018b). *Energy Efficiency and Historic Buildings: How to Improve Energy Efficiency* (Guidance HEAG094). Historic England. Swindon, United Kingdom.
- Historic England. (2020). *Energy Efficiency and Traditional Homes: Historic England (Advice Note 14)* (Guidance HEAG295; HEAN014). Historic England. Swindon, United Kingdom.
- Historic England. (2024). *Geospatial Survey Specifications for Cultural Heritage* (Guidance HEAG317). Historic England. York, United Kingdom.
- Historic England. (2025a). Lenton Lodge (Former Gateway to Wollaton Park) and Attached Bollards. Retrieved August 16, 2023, from Available from: <https://historicengland.org.uk/listing/the-list/list-entry/1270829?section=official-list-entry>
- Historic England. (2025b). *Understanding the Environmental Performance of Historic Buildings for Conservation* (Guidance HEAG329). Historic England. Swindon, United Kingdom.
- Hoegner, L., Abmayr, T., Tosic, D., Turzer, S., & Stilla, U. (2018). Fusion of TLS and RGB Point Clouds with TIR Images for Indoor Mobile Mapping. *Proceedings of the 2018 International Conference on Quantitative InfraRed Thermography (QIRT 2018)*, 341–349. DOI: 10.21611/qirt.2018.019
- Hoegner, L., Kumke, H., Schwarz, A., Meng, L., & Stilla, U. (2007). Strategies for texturing building models with low resolution infrared image sequences. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 6.

- Hoegner, L., & Stilla, U. (2015). Building Facade Object Detection from terrestrial Thermal Infrared Image Sequences Combining Different Views. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, II-3/W4*, 55–62. DOI: 10.5194/isprsannals-II-3-W4-55-2015
- Hoegner, L., & Stilla, U. (2016). Automatic 3D Reconstruction and Texture Extraction for 3D Building Models from Thermal Infrared Image Sequences. *Proceedings of the 2016 International Conference on Quantitative InfraRed Thermography (QIRT 2016)*, 322–331. DOI: 10.21611/qirt.2016.042
- Hoegner, L., & Stilla, U. (2018). Mobile Thermal Mapping for Matching of Infrared Images with 3D Building Models and 3D Point Clouds. *Quantitative InfraRed Thermography Journal*, 15(2), 252–270. DOI: 10.1080/17686733.2018.1455129
- Hoegner, L., Tuttas, S., Xu, Y., Eder, K., & Stilla, U. (2016). Evaluation of Methods for Coregistration and Fusion of RPAS-Based 3D Point Clouds and Thermal Infrared Images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B3*, 241–246. DOI: 10.5194/isprsarchives-XLI-B3-241-2016
- Hoegner, L., & Stilla, U. (2007). Texture Extraction for Building Models from IR Sequences of Urban Areas. *Proceedings of the 2007 Urban Remote Sensing Joint Event*, 1–6. DOI: 10.1109/URS.2007.371812
- Hoffman, T. (2024). SunCalc: Sunrise, Sunset, Shadow Length, Solar Eclipse. Retrieved January 15, 2024, from Available from: <https://www.suncalc.org/>
- Hou, Y., Chen, M., Volk, R., & Soibelman, L. (2022). Investigation on Performance of RGB Point Cloud and Thermal Information Data Fusion for 3D Building Thermal Map Modeling Using Aerial Images under Different Experimental Conditions. *Journal of Building Engineering*, 45(103380). DOI: 10.1016/j.jobe.2021.103380
- Ibarra-Castanedo, C., Tarpani, J. R., & Maldague, X. P. V. (2013). Nondestructive Testing with Thermography. *European Journal of Physics*, 34(6), S91–S109. DOI: 10.1088/0143-0807/34/6/S91

- Italy This Way. (2025). Arco, Italy - Town in a Picturesque Setting Near Lake Garda. Retrieved October 1, 2025, from Available from: <https://www.italythisway.com/places/articles/arco-history.php>
- Iwaszczuk, D., & Stilla, U. (2017). Camera Pose Refinement by Matching Uncertain 3D Building Models with Thermal Infrared Image Sequences for High Quality Texture Extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 132, 33–47. DOI: 10.1016/j.isprsjprs.2017.08.006
- Javadnejad, F., Gillins, D. T., Parrish, C. E., & Slocum, R. K. (2020). A Photogrammetric Approach to Fusing Natural Colour and Thermal Infrared UAS Imagery in 3D Point Cloud Generation. *International Journal of Remote Sensing*, 41(1), 211–237. DOI: 10.1080/01431161.2019.1641241
- Jiang, X., Ma, J., Xiao, G., Shao, Z., & Guo, X. (2021). A Review of Multimodal Image Matching: Methods and Applications. *Information Fusion*, 73, 22–71. DOI: 10.1016/j.inffus.2021.02.012
- Jin, Y., Mishkin, D., Mishchuk, A., Matas, J., Fua, P., Yi, K. M., & Trulls, E. (2020). Image Matching Across Wide Baselines: From Paper to Practice. *International Journal of Computer Vision*, 129, 517–547. DOI: 10.1007/s11263-020-01385-0
- Jones, G. (2001). A Novel Approach for Surveillance Using Visual and Thermal Images. *Proceedings of the DERA/IEEE Workshop Intelligent Sensor Processing*, 2001, 1–9. DOI: 10.1049/ic:20010104
- Kelly, J., Kljun, N., Olsson, P.-O., Mihai, L., Liljeblad, B., Weslien, P., Klemedtsson, L., & Eklundh, L. (2019). Challenges and Best Practices for Deriving Temperature Data from an Uncalibrated UAV Thermal Infrared Camera. *Remote Sens.*, 11(567). DOI: 10.3390/rs11050567
- Kennedy, C. J., Penman, M., Watkinson, D., Emmerson, N., Thickett, D., Bosché, F., Forster, A. M., Grau-Bové, J., & Cassar, M. (2024). Beyond Heritage Science: A Review. *Heritage*, 7(3), 1510–1538. DOI: 10.3390/heritage7030073

- Khalil, H., & Tricco, A. C. (2022). Differentiating Between Mapping Reviews and Scoping Reviews in the Evidence Synthesis Ecosystem. *Journal of Clinical Epidemiology*, 149, 175–182. DOI: 10.1016/j.jclinepi.2022.05.012
- Khodayar, F., Sojasi, S., & Maldague, X. (2016). Infrared Thermography and NDT: 2050 Horizon. *Quantitative InfraRed Thermography Journal*, 13(2), 210–231. DOI: 10.1080/17686733.2016.1200265
- Kioussi, A., Karoglou, M., Labropoulos, K., Bakolas, A., & Moropoulou, A. (2013). Integrated Documentation Protocols Enabling Decision Making in Cultural Heritage Protection. *Journal of Cultural Heritage*, 14S, e141–e146. DOI: 10.1016/j.culher.2013.01.007
- Kirchhoff, G. (1860). I. On the Relation Between the Radiating and Absorbing Powers of Different Bodies for Light and Heat. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 20(130), 1–21. DOI: 10.1080/14786446008642901
- Kraaijenbrink, P. D. A., Shea, J. M., Litt, M., Steiner, J. F., Treichler, D., Koch, I., & Immerzeel, W. W. (2018). Mapping Surface Temperatures on a Debris-Covered Glacier With an Unmanned Aerial Vehicle. *Frontiers in Earth Science*, 6. DOI: 10.3389/feart.2018.00064
- Kubota, Y., Ke, Y., Hayakawa, T., Moko, Y., & Ishikawa, M. (2021). Optimal Material Search for Infrared Markers under Non-Heating and Heating Conditions. *Sensors*, 21(6527). DOI: 10.3390/s21196527
- Kusnierek, K., & Korsæth, A. (2014). Challenges in Using an Analog Uncooled Microbolometer Thermal Camera to Measure Crop Temperature. *International Journal of Agricultural & Biological Engineering*, 7(4), 60–74. DOI: 10.3965/j.ijabe.20140704.007
- Kylili, A., Fokaides, P. A., Christou, P., & Kalogirou, S. A. (2014). Infrared Thermography (IRT) Applications for Building Diagnostics: A Review. *Applied Energy*, 134, 531–549. DOI: 10.1016/j.apenergy.2014.08.005

- Lague, D., Brodu, N., & Leroux, J. (2013). Accurate 3D Comparison of Complex Topography with Terrestrial Laser Scanner: Application to the Rangitikei Canyon (NZ). *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 10–26. DOI: 10.1016/j.isprsjprs.2013.04.009
- Lagüela, S., Armesto, J., Arias, P., & Herráez, J. (2012). Automation of Thermographic 3D Modelling Through Image Fusion and Image Matching Techniques. *Automation in Construction*, 27, 24–31. DOI: 10.1016/j.autcon.2012.05.011
- Lagüela, S., Díaz-Vilariño, L., Martínez, J., & Armesto, J. (2013). Automatic Thermographic and RGB Texture of As-Built BIM for Energy Rehabilitation Purposes. *Automation in Construction*, 31, 230–240. DOI: 10.1016/j.autcon.2012.12.013
- Lagüela, S., Díaz-Vilariño, L., & Roca, D. (2016). Infrared Thermography: Fundamentals and Applications. In B. Riveiro & M. Solla (Eds.), *Non-Destructive Techniques for the Evaluation of Structures and Infrastructure* (1st ed.). CRC Press.
- Lagüela, S., González-Jorge, H., Armesto, J., & Arias, P. (2011a). Calibration and Verification of Thermographic Cameras for Geometric Measurements. *Infrared Physics & Technology*, 54(2), 92–99. DOI: 10.1016/j.infrared.2011.01.002
- Lagüela, S., Martínez, J., Armesto, J., & Arias, P. (2011b). Energy Efficiency Studies Through 3D Laser Scanning and Thermographic Technologies. *Energy and Buildings*, 43, 1216–1221. DOI: 10.1016/j.enbuild.2010.12.031
- Lambert, J. H. (1760). *Photometria, sive De mensura et gradibus luminis, colorum et umbrae*. Sumptibus Viduae Eberhardi Klett. DOI: 10.3931/E-RARA-9488
- Langley, S. P. (1880). The Bolometer and Radiant Energy. *Proceedings of the American Academy of Arts and Sciences*, 16(342). DOI: 10.2307/25138616

- Lawson, W., Nielsen, S., Putley, E., & Young, A. (1959). Preparation and Properties of HgTe and Mixed Crystals of HgTe-CdTe. *Journal of Physics and Chemistry of Solids*, 9(3-4), 325–329. DOI: 10.1016/0022-3697(59)90110-6
- Lecomte, V., Macher, H., & Landes, T. (2022). Combination of Thermal Infrared Images and Laserscanning Data for 3D Thermal Point Cloud Generation on Buildings and Trees. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-2/W1-2022, 129–136. DOI: 10.5194/isprs-archives-XLVIII-2-W1-2022-129-2022
- Lenton Times. (1983). Lenton Lodge – The Inside Story (Issue 26). *The Lenton Listener*. Retrieved August 16, 2023, from Available from: https://www.lentontimes.co.uk/images/gallery/derby_road/derby_rd_listener_26.htm
- Lenton Times. (2006). Lenton Times Back Issues - A Brief Synopsis of the Main Articles - Issue No. 24 October 2006 (24th ed.). *Lenton Times*. Retrieved August 16, 2023, from Available from: https://www.lentontimes.co.uk/back_issues/issue_24/issue_24_back.htm
- Lerma, J. L., Navarro, S., Cabrelles, M., & Seguí, A. E. (2010). Camera Calibration with Baseline Distance Constraints. *The Photogrammetric Record*, 25(130), 140–158. DOI: 10.1111/j.1477-9730.2010.00579.x
- Letellier, R., Schmid, W., LeBlanc, F., Eppich, R., & Cabbi, A. (2011). *Recording, Documentation, and Information Management for the Conservation of Heritage Places: Guiding Principles*. Donhead Publishing.
- Lewińska, P., & Maciuk, K. (2020). Thermal and Spatial Data Integration for Recreating Rebuilding Stages of Wooden and Masonry Buildings. *Photogrammetric Record*, 35(171), 402–425. DOI: 10.1111/phor.12332
- Li, J., Hu, Q., & Ai, M. (2020). RIFT: Multi-Modal Image Matching Based on Radiation-Variation Insensitive Feature Transform. *IEEE Transactions on Image Processing*, 29, 3296–3310. DOI: 10.1109/TIP.2019.2959244

- Li, J., Hu, Q., & Zhang, Y. (2023a). Multimodal Image Matching: A scale-Invariant Algorithm and an Open Dataset. *ISPRS Journal of Photogrammetry and Remote Sensing*, 204, 77–88. DOI: 10.1016/j.isprsjprs.2023.08.010
- Li, J., Shi, P., Hu, Q., & Zhang, Y. (2023b). RIFT2: Speeding-up RIFT with A New Rotation-Invariance Technique. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.2303.00319
- Li, J., Xu, W., Shi, P., Zhang, Y., & Hu, Q. (2022). LNIFT: Locally Normalized Image for Rotation Invariant Multimodal Feature Matching. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–14. DOI: 10.1109/TGRS.2022.3165940
- Lin, D., Dong, Z., Zhang, X., & Maas, H.-G. (2019a). Unsupervised Window Extraction from Photogrammetric Point Clouds with Thermal Attributes. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W5, 45–51. DOI: 10.5194/isprs-annals-IV-2-W5-45-2019
- Lin, D., Jarzabek-Rychard, M., Schneider, D., & Maas, H.-G. (2018a). Thermal Texture Selection and Correction for Building Facade Inspection Based on Thermal Radiant Characteristics. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2, 585–591. DOI: 10.5194/isprs-archives-XLII-2-585-2018
- Lin, D., Maas, H.-G., Westfeld, P., Budzier, H., & Gerlach, G. (2018b). An Advanced Radiometric Calibration Approach for Uncooled Thermal Cameras. *The Photogrammetric Record*, 33(161), 30–48. DOI: 10.1111/phor.12216
- Lin, D., Westfeld, P., & Maas, H.-G. (2017). Shutter-less Temperature-Dependent Correction for Uncooled Thermal Camera under Fast Changing FPA Temperature. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-1/W1, 619–625. DOI: 10.5194/isprs-archives-XLII-1-W1-619-2017
- Lin, D., Jarzabek-Rychard, M., Tong, X., & Maas, H.-G. (2019b). Fusion of Thermal Imagery with Point Clouds for Building Façade Thermal Attribute Mapping.

- ISPRS Journal of Photogrammetry and Remote Sensing*, 151, 162–175. DOI: 10.1016/j.isprsjprs.2019.03.010
- Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2014). Microsoft COCO: Common Objects in Context [Version Number: 3]. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.1405.0312
- Lin, Y. Y., Pan, X.-Y., Fridovich-Keil, S., & Wetzstein, G. (2024). ThermalNeRF: Thermal Radiance Fields. *International Conference of Computational Photography*. DOI: 10.48550/ARXIV.2407.15337
- Lindenberger, P., Sarlin, P.-E., & Pollefeys, M. (2023). LightGlue: Local Feature Matching at Light Speed. *IEEE/CVF International Conference on Computer Vision (ICCV)*, 17581–17592. DOI: 10.1109/ICCV51070.2023.01616
- Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), 91–110. DOI: 10.1023/B:VISI.0000029664.99615.94
- Lowe, D. (1999). Object Recognition from Local Scale-Invariant Features. *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 2, 1150–1157. DOI: 10.1109/ICCV.1999.790410
- Lucchi, E. (2018). Applications of the Infrared Thermography in the Energy Audit of Buildings: A Review. *Renewable and Sustainable Energy Reviews*, 82, 3077–3090. DOI: 10.1016/j.rser.2017.10.031
- Luhmann, T., Fraser, C., & Maas, H.-G. (2016). Sensor Modelling and Camera Calibration for Close-Range Photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 37–46. DOI: 10.1016/j.isprsjprs.2015.10.006
- Luhmann, T., Piechel, J., & Roelfs, T. (2013). Geometric Calibration of Thermographic Cameras. In C. Kuenzer & S. Dech (Eds.), *Thermal Infrared Remote Sensing* (pp. 27–42). Springer. DOI: 10.1007/978-94-007-6639-6_2
- Luhmann, T., Robson, S., & Kyle, S. (Eds.). (2006). *Close-Range Photogrammetry: Principles, Methods and Applications* (2nd). Whittles Publishing.

- Luo, Z., Zhou, L., Bai, X., Chen, H., Zhang, J., Yao, Y., Li, S., Fang, T., & Quan, L. (2020). ASLFeat: Learning Local Features of Accurate Shape and Localization [Version Number: 2]. *Computer Vision and Pattern Recognition (CVPR 2020)*. DOI: 10.48550/ARXIV.2003.10071
- Ma, J., Jiang, X., Fan, A., Jiang, J., & Yan, J. (2021). Image Matching from Hand-crafted to Deep Features: A Survey. *International Journal of Computer Vision*, 129(1), 23–79. DOI: 10.1007/s11263-020-01359-2
- Macher, H., & Landes, T. (2022). Combining TIR Images and Point Clouds for Urban Scenes Modelling. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2022*, 425–431. DOI: 10.5194/isprs-archives-XLIII-B2-2022-425-2022
- Maierhofer, C., Krankenhagen, R., Röllig, M., Schlichting, J., Schiller, M., Seidl, T., Mecke, R., Kalisch, U., Hennen, C., & Meinhardt, J. (2011). Investigating Historic Masonry Structures with a Combination of Active Thermography and 3D Laser Scanner. *Quantitative Infrared Thermography Journal*, 8(1), 115–118. DOI: 10.3166/qirt.8.115-118
- Malbêteau, Y., Johansen, K., Aragon, B., Al-Mashhawari, S. K., & McCabe, M. F. (2021). Overcoming the Challenges of Thermal Infrared Orthomosaics Using a Swath-Based Approach to Correct for Dynamic Temperature and Wind Effects. *Remote Sensing*, 13(3255). DOI: 10.3390/rs13163255
- Maldague, X. (2001). *Theory and Practice of Infrared Technology for Nondestructive Testing*. Wiley.
- Malmivirta, T., Hamberg, J., Lagerspetz, E., Li, X., Peltonen, E., Flores, H., & Nurmi, P. (2019). Hot or Not? Robust and Accurate Continuous Thermal Imaging on FLIR cameras. *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 1–9. DOI: 10.1109/PERCOM.2019.8767423
- Marelli, D., Morelli, L., Farella, E. M., Bianco, S., Ciocca, G., & Remondino, F. (2023). ENRICH: Multi-purposE dataset for beNchmaRking In Computer

- vision and pHotogrammetry. *ISPRS Journal of Photogramm. and Remote Sens.*, 198, 84–98. DOI: 10.1016/j.isprsjprs.2023.03.002
- Martínez-Molina, A., Tort-Ausina, I., Cho, S., & Vivancos, J. L. (2016). Energy Efficiency and Thermal Comfort in Historic Buildings: A Review. *Renewable and Sustainable Energy Reviews*, 61, 70–85. DOI: 10.1016/j.rser.2016.03.018
- Martín-Lerones, P., Olmedo, D., López-Vidal, A., Gómez-García-bermejo, J., & Zalama, E. (2021). BIM Supported Surveying and Imaging Combination for Heritage Conservation. *Remote Sensing*, 13(1584), 1–11. DOI: 10.3390/rs13081584
- Maset, E., Fusiello, A., Crosilla, F., Toldo, R., & Zorzetto, D. (2017). Photogrammetric 3D Building Reconstruction from Thermal Images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W3, 25–32. DOI: 10.5194/isprs-annals-IV-2-W3-25-2017
- Merchán, P., Merchán, M. J., Salamanca, S., & Adán, A. (2018). Application of Multisensory Technology for Resolution of Problems in the Field of Research and Preservation of Cultural Heritage. *Lecture Notes in Computer Science*, 10754, 32–47. DOI: 10.1007/978-3-319-75789-6_3
- Mesas-Carrascosa, F.-J., Pérez-Porras, F., Meroño De Larriva, J., Mena Frau, C., Agüera-Vega, F., Carvajal-Ramírez, F., Martínez-Carricondo, P., & García-Ferrer, A. (2018). Drift Correction of Lightweight Microbolometer Thermal Sensors On-Board Unmanned Aerial Vehicles. *Remote Sensing*, 10(615). DOI: 10.3390/rs10040615
- Mileto, C., Vegas, F., & Lerma, J. L. (2015). Multidisciplinary Studies, Crossreading and Transversal Use of Thermography: The Castle of Monzón (Huesca) as a Case Study. *Proceedings of the Defensive Architecture of the Mediterranean—XV to XVIII Centuries*, 2, 405–412. DOI: 10.4995/fortmed2015.2015.1786

- Mineo, S., & Pappalardo, G. (2021). Rock Emissivity Measurement for Infrared Thermography Engineering Geological Applications. *Applied Sciences*, 11(3773). DOI: 10.3390/app11093773
- Mishchuk, A., Mishkin, D., Radenovic, F., & Matas, J. (2017). Working hHard to Know Your Neighbor’s Margins: Local Descriptor Learning Loss. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.1705.10872
- Mishkin, D., Matas, J., & Perdoch, M. (2015). MODS: Fast and Robust Method for Two-View Matching. *Computer Vision and Image Understanding*. DOI: <http://dx.doi.org/10.1016/j.cviu.2015.08.005>
- Morelli, L., Bellavia, F., Menna, F., & Remondino, F. (2022). Photogrammetry Now and Then: From Hand-Crafted to Deep-Learning Tie Points. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-2/W1-2022, 163–170. DOI: 10.5194/isprs-archives-XLVIII-2-W1-2022-163-2022
- Morelli, L., Ioli, F., Maiwald, F., Mazzacca, G., Menna, F., & Remondino, F. (2024a). Deep-Image-Matching: A Toolbox for Multiview Image Matching of Complex Scenarios. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-2/W4-2024, 309–316. DOI: 10.5194/isprs-archives-XLVIII-2-W4-2024-309-2024
- Morelli, L., Perda, G., Ioli, F., Trybała, P., Sterpin, A., Rigon, S., Sutherland, N., Medici, M., Remondino, F., & Vitti, A. (2024b). Co-Registering Laser Scanning Point Clouds and Photogrammetric Images with Deep Learning Multi-Modal Matching. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-2/W8-2024, 335–342. DOI: 10.5194/isprs-archives-XLVIII-2-W8-2024-335-2024
- Morelli, L., Perda, G., Malek, S., Remondino, F., & Legat, K. (2025a). Night and Day Aerial Photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-1/W4-2025, 81–88. DOI: 10.5194/isprs-archives-XLVIII-1-W4-2025-81-2025

- Morelli, L., Sutherland, N., Remondino, F., & Mills, J. P. (2025b). Deep Learning Multi-Modal Image Matching for the Automatic Co-Registration of Thermal InfraRed and Visible Images. *ISPRS Journal of Photogrammetry and Remote Sensing*, (Under Review).
- Motayyeb, S., Samadzedegan, F., Dadrass Javan, F., & Hosseinpour, H. (2023). Fusion of UAV-based Infrared and Visible Images for Thermal Leakage Map Generation of Building Façades. *Heliyon*, 9(e14551). DOI: 10.1016/j.heliyon.2023.e14551
- Napolitano, R., Hess, M., & Glisic, B. (2019). Integrating Non-Destructive Testing, Laser Scanning, and Numerical Modeling for Damage Assessment: The Room of the Elements. *Heritage*, 2, 151–168. DOI: 10.3390/heritage2010012
- Nardi, I., Sfarra, S., Ambrosini, D., Dominici, D., & Rosciano, E. (2015). Complementarity of Terrestrial Laser Scanning and IR Thermography for the Diagnosis of Cultural Heritage: The Case of Pacentro Castle. *Proceedings of the 13th International Workshop on Advanced Infrared Technology & Applications*, 263–266.
- Natephra, W., Motamedi, A., Yabuki, N., & Fukuda, T. (2017). Integrating 4D Thermal Information with BIM for Building Envelope Thermal Performance Analysis and Thermal Comfort Evaluation in Naturally Ventilated Environments. *Building and Environment*, 124, 194–208. DOI: 10.1016/j.buildenv.2017.08.004
- Nicol, S., Beer, C., & Scott, C. (2014). *The Age and Construction of English Homes: A Guide to Ageing the English Housing Stock*. IHS BRE Press.
- Nobili, L. (1830). Beschreibung eines Thermo-Multiplicators oder Elektrischen Thermoskops. *Annalen der Physik*, 96(10), 245–252. DOI: 10.1002/andp.18300961004
- Nugent, P. W., Shaw, J. A., & Pust, N. J. (2013). Correcting for Focal-Plane-Array Temperature Dependence in Microbolometer Infrared Cameras Lacking Thermal Stabilization. *Optical Engineering*, 52(6)(061304). DOI: 10.1117/1.oe.52.6.061304

- Nugent, P. W., Shaw, J. A., & Pust, N. J. (2014). Radiometric Calibration of Infrared Imagers Using an Internal Shutter as an Equivalent External Blackbody. *Optical Engineering*, 53(12)(123106). DOI: 10.1117/1.oe.53.12.123106
- Office for National Statistics (ONS). (2022). Age of the property is the biggest single factor in energy efficiency of homes. Retrieved January 6, 2023, from Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles>
- Olbrycht, R., & Więcek, B. (2015). New Approach to Thermal Drift Correction in Microbolometer Thermal Cameras. *Quantitative InfraRed Thermography Journal*, 12(2), 184–195. DOI: 10.1080/17686733.2015.1055675
- Olbrycht, R., Więcek, B., & Mey, G. D. (2012). Thermal Drift Compensation Method for Microbolometer Thermal Cameras. *Applied Optics*, 51(11), 1788–1794. DOI: 10.1364/AO.51.001788
- Ono, Y., Trulls, E., Fua, P., & Yi, K. M. (2018). LF-Net: Learning Local Features from Images [Version Number: 2]. *Computer Vision and Pattern Recognition (NIPS 2018)*. DOI: 10.48550/ARXIV.1805.09662
- Oreni, D., Brumana, R., Della Torre, S., Banfi, F., Barazzetti, L., & Previtali, M. (2014). Survey Turned into HBIM: The Restoration and the Work Involved Concerning the Basilica di Collemaggio After the Earthquake (L'Aquila). *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5, 267–273. DOI: 10.5194/isprsannals-II-5-267-2014
- Osello, A., Lucibello, G., & Morgagni, F. (2018). HBIM and Virtual Tools: A New Chance to Preserve Architectural Heritage. *Buildings*, 8(1). DOI: 10.3390/buildings8010012
- Patrucco, G., Cortese, G., Tonolo, G. F., & Spanò, A. (2020). Thermal and Optical Data Fusion Supporting Built Heritage Analyses. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B3-2020, 619–626. DOI: 10.5194/isprs-archives-XLIII-B3-2020-619-2020

- Patrucco, G., Gómez, A., Adineh, A., Rahrig, M., & Lerma, J. L. (2022a). 3D Data Fusion for Historical Analyses of Heritage Buildings Using Thermal Images: The Palacio de Colomina as a Case Study. *Remote Sensing*, *14*(5699). DOI: 10.3390/rs14225699
- Patrucco, G., Tonolo, F. G., Sammartano, G., & Spanò, A. (2022b). SFM-based 3D Reconstruction of Heritage Assets Using UAV Thermal Images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B1-2022*, 399–406. DOI: 10.5194/isprs-archives-XLIII-B1-2022-399-2022
- Paziewska, J., & Rzonca, A. (2022). Integration of Thermal and RGB Data Obtained by Means of a Drone for Interdisciplinary Inventory. *Energies*, *15*(4971). DOI: 10.3390/en15144971
- Perda, G., Morelli, L., Remondino, F., Fraser, C., & Luhmann, T. (2024). Analyzing Target-, Handcrafted- and Learning-Based Methods for Automated 3D Measurement and Modelling. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLVIII-2/W7-2024*, 105–112. DOI: 10.5194/isprs-archives-XLVIII-2-W7-2024-105-2024
- Pérez, E., Merchán, P., Espacio, A., & Salamanca, S. (2024). Fusion of Thermal Point Cloud Series of Buildings for Inspection in Virtual Reality. *Buildings*, *14*(2127). DOI: 10.3390/buildings14072127
- Pestana, S., Chickadel, C. C., Harpold, A., Kostadinov, T. S., Pai, H., Tyler, S., Webster, C., & Lundquist, J. D. (2019). Bias Correction of Airborne Thermal Infrared Observations Over Forests Using Melting Snow. *Water Resources Research*, *55*(12), 11331–11343. DOI: 10.1029/2019wr025699
- Pevsner, N., & Williamson, E. (1979). *The Buildings of England: Nottinghamshire* (2nd ed.). Yale University Press.
- Planck, M. (1901). Ueber das Gesetz der Energieverteilung im Normalspectrum. *Annalen der Physik*, *309*(3), 553–563. DOI: 10.1002/andp.19013090310

- Potje, G., Cadar, F., Araujo, A., Martins, R., & Nascimento, E. R. (2024). XFeat: Accelerated Features for Lightweight Image Matching. *Computer Vision and Pattern Recognition (CVPR 2024)*. DOI: 10.48550/ARXIV.2404.19174
- Pout, C. H., Davidson, P., & Moss, S. A. (1998). *Non-Domestic Building Energy Fact File (BR 339)* (1st ed.). IHS BRE Press.
- Prakash, S., Lee, P. Y., & Caelli, T. (2006). 3D Mapping of Surface Temperature Using Thermal Stereo. *Proceedings of The 9th International Conference on Control, Automation, Robotics and Vision*, 1–4. DOI: 10.1109/I-CARCV.2006.345342
- Previtali, M., Barazzetti, L., Redaelli, V., Scaioni, M., & Rosina, E. (2013). Rigorous Procedure for Mapping Thermal Infrared Images on Three-Dimensional Models of Building Façades. *Journal of Applied Remote Sensing*, 7(1)(073503). DOI: 10.1117/1.JRS.7.073503
- Previtali, M., Erba, S., Rosina, E., Redaelli, V., Scaioni, M., & Barazzetti, L. (2012). Generation of a GIS-based Environment for Infrared Thermography Analysis of Buildings. *Proceedings of SPIE, 8511 (85110U)*. DOI: 10.1117/12.930050
- Pultar, M. (2020). Improving the HardNet Descriptor. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.2007.09699
- Ramachandra, B., Nawathe, P., Monroe, J., Han, K., Ham, Y., & Vatsavai, R. R. (2018). Real-Time Energy Audit of Built Environments: Simultaneous Localization and Thermal Mapping. *Journal of Infrastructure Systems*, 24(3)(04018013). DOI: 10.1061/(ASCE)IS.1943-555X.0000431
- Ramón, A., Adán, A., & Javier Castilla, F. (2022). Thermal Point Clouds of Buildings: A Review. *Energy and Buildings*, 274. DOI: 10.1016/j.enbuild.2022.112425
- Ramos, M. M., & Remondino, F. (2015). Data Fusion in Cultural Heritage: A Review. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5/W7, 359–363. DOI: 10.5194/isprsarchives-XL-5-W7-359-2015

- Rangel, J., & Soldan, S. (2014). 3D Thermal Imaging: Fusion of Thermography and Depth Cameras. *12th International Conference on Quantitative InfraRed Thermography*, 1–10. DOI: 10.21611/qirt.2014.035
- Rankin, A., Huertas, A., Matthies, L., Bajracharya, M., Assad, C., Brennan, S., Bellutta, P., & Sherwin, G. W. (2013). Unmanned Ground Vehicle Perception Using Thermal Infrared Cameras. *The International Society for Optical Engineering 8045, 8045 (804503)*. DOI: 10.1117/12.884349
- Remondino, F., & Fraser, C. (2006). Digital Camera Calibration Methods: Considerations and Comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 266–272. DOI: 10.3929/ETHZ-B-000158067
- Remondino, F., & Rizzi, A. (2010). Reality-based 3D Documentation of Natural and Cultural Heritage Sites: Techniques, Problems, and Examples. *Applied Geomatics*, 2, 85–100. DOI: 10.1007/s12518-010-0025-x
- Remondino, F., Spera, M. G., Nocerino, E., Menna, F., & Nex, F. (2014). State of the Art in High Density Image Matching. *Photogrammetric Record*, 29(146), 144–166. DOI: 10.1111/phor.12063
- Revaud, J., Weinzaepfel, P., De Souza, C., Pion, N., Csurka, G., Cabon, Y., & Humenberger, M. (2019). R2D2: Repeatable and Reliable Detector and Descriptor. DOI: 10.48550/arXiv.1906.06195
- Riba, E., Mishkin, D., Ponsa, D., Rublee, E., & Bradski, G. (2019). Kornia: An Open Source Differentiable Computer Vision Library for PyTorch. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.1910.02190
- Ribeiro-Gomes, K., Hernández-López, D., Ortega, J., Ballesteros, R., Poblete, T., & Moreno, M. (2017). Uncooled Thermal Camera Calibration and Optimization of the Photogrammetry Process for UAV Applications in Agriculture. *Sensors*, 17(2173). DOI: 10.3390/s17102173
- Riveiro, B., Lourenço, P. B., Oliveira, D. V., González-Jorge, H., & Arias, P. (2016). Automatic Morphologic Analysis of Quasi-Periodic Masonry Walls from Li-

- DAR. *Computer-Aided Civil and Infrastructure Engineering*, 31(4), 305–319.
DOI: 10.1111/mice.12145
- Riveiro, B., & Solla, M. (Eds.). (2016). *Non-Destructive Techniques for the Evaluation of Structures and Infrastructure* (1st ed.). CRC Press.
- Rizzi, A., Voltolini, F., Girardi, S., Gonzo, L., & Remondino, F. (2007). Digital Preservation, Documentation and Analysis of Paintings, Monuments and Large Cultural Heritage with Infrared Technology, Digital Cameras and Range Sensors. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1–6.
- Rodríguez-Gonzálvez, P., Muñoz-Nieto, A. L., DelPozo, S., Sanchez-Aparicio, L. J., Gonzalez-Aguilera, D., Micoli, L., Barsanti, S. G., Guidi, G., Mills, J., Fieber, K., Haynes, I., & Hejmanowska, B. (2017). 4D Reconstruction and Visualization of Cultural Heritage: Analyzing our Legacy Through Time. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W3, 609–616. DOI: 10.5194/isprs-archives-XLII-2-W3-609-2017
- Rogalski, A. (2012). History of Infrared Detectors. *Opto-Electronics Review*, 20(3), 279–308. DOI: 10.2478/s11772-012-0037-7
- Roshan, M. C., Isaksson, M., & Pranata, A. (2024). A Geometric Calibration Method for Thermal Cameras using a ChArUco Board. *Infrared Physics & Technology*, 138(105219). DOI: 10.1016/j.infrared.2024.105219
- Rosten, E., & Drummond, T. (2006). Machine Learning for High-Speed Corner Detection. In A. Leonardis, H. Bischof, & A. Pinz (Eds.), *Computer Vision – ECCV 2006* (pp. 430–443). Springer-Verlag. DOI: 10.1007/11744023_34
- Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An Efficient Alternative to SIFT or SURF. *2011 International Conference on Computer Vision*, 2564–2571. DOI: 10.1109/ICCV.2011.6126544
- Ruiz De Oña, E., Barbero-García, I., González-Aguilera, D., Remondino, F., Rodríguez-Gonzálvez, P., & Hernández-López, D. (2023). PhotoMatch: An Open-Source

- Tool for Multi-View and Multi-Modal Feature-Based Image Matching. *Applied Sciences*, 13(5467). DOI: 10.3390/app13095467
- Rusu, R. B., Blodow, N., & Beetz, M. (2009). Fast Point Feature Histograms (FPFH) for 3D Registration. *Proceedings of the 2009 IEEE International Conference on Robotics and Automation*, 3212–3217. DOI: 10.1109/ROBOT.2009.5152473
- Sadhukhan, D., Peri, S., Sugunaraj, N., Biswas, A., Selvaraj, D. F., Koiner, K., Rosener, A., Dunlevy, M., Goveas, N., Flynn, D., & Ranganathan, P. (2020). Estimating Surface Temperature from Thermal Imagery of Buildings for Accurate Thermal Transmittance (U-value): A Machine Learning Perspective. *Journal of Building Engineering*, 32(101637). DOI: 10.1016/j.jobbe.2020.101637
- Saleem, S., Bais, A., Sablatnig, R., Ahmad, A., & Naseer, N. (2017). Feature Points for Multisensor Images. *Computers & Electrical Engineering*, 62, 511–523. DOI: 10.1016/j.compeleceng.2017.04.032
- Sarlin, P.-E., DeTone, D., Malisiewicz, T., & Rabinovich, A. (2019). SuperGlue: Learning Feature Matching with Graph Neural Networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2020*, 4938–4947. DOI: 10.1109/CVPR42600.2020.00498
- Scaioni, M., Rosina, E., L’erario, A., & Díaz-Vilariño, L. (2017). Integration of Infrared Thermography and Photogrammetric Surveying of Built Landscape. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-5/W1, 153–160. DOI: 10.5194/isprs-Archives-XLII-5-W1-153-2017
- Scaioni, M., Rosina, E., Barazzetti, L., Previtali, M., & Redaelli, V. (2012). High-Resolution Texturing of Building Facades with Thermal Images. *Proceedings of SPIE*, 8354, 1–10. DOI: 10.1117/12.920613
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji:

- An Open-Source Platform for Biological-Image Analysis. *Nature Methods*, 9(7), 676–682. DOI: 10.1038/nmeth.2019
- Schonberger, J. L., & Frahm, J.-M. (2016). Structure-from-Motion Revisited. *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 4104–4113. DOI: 10.1109/CVPR.2016.445
- Seebeck, T. J. (1826). Ueber die Magnetische Polarisation der Metalle und Erze durch Temperaturdifferenz. *Annalen der Physik*, 82(3), 253–286. DOI: 10.1002/andp.18260820302
- Senn, J. A., Mills, J., Walsh, C. L., Addy, S., & Peppas, M. V. (2022). On-Site Geometric Calibration of RPAS Mounted Sensors for SfM Photogrammetric Geomorphological Surveys. *Earth Surface Processes and Landforms*, 47(6), 1615–1634. DOI: 10.1002/esp.5338
- Sledz, A., & Heipke, C. (2022). Joint Bundle Adjustment of Thermal Infra-Red and Optical Images Based on Multimodal Matching. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B1-2022*, 157–165. DOI: 10.5194/isprs-archives-XLIII-B1-2022-157 - 2022
- Snaveely, N., Seitz, S. M., & Szeliski, R. (2006). Photo Tourism: Exploring Photo Collections in 3D. *ACM Transactions on Graphics*, 25(3), 835–846. DOI: 10.1145/1141911.1141964
- Solla, M., Gonçalves, L. M., Gonçalves, G., Francisco, C., Puente, I., Providência, P., Gaspar, F., & Rodrigues, H. (2020). A Building Information Modeling Approach to Integrate Geomatic Data for the Documentation and Preservation of Cultural Heritage. *Remote Sensing*, 12(4028). DOI: 10.3390/rs12244028
- Stefan, J. (1879). Über die Beziehung zwischen der Wärmestrahlung und der Temperatur. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Classe*, 79, 391–428.
- St-Laurent, L., Mikhnevich, M., Bubel, A., & Prévost, D. (2016). Passive Calibration Board for Alignment of VIS-NIR, SWIR and LWIR Images. *13th Inter-*

- national Conference on Quantitative InfraRed Thermography (QIRT 2016)*, 777–784. DOI: 10.21611/qirt.2016.124
- Storia dell’Arte. (2025). Il Castello di Arco. Retrieved October 1, 2025, from Available from: https://www.progettostoriadellarte.it/2019/09/15/il-castello-di-arco/#google_vignette
- Stylianidis, E., & Remondino, F. (Eds.). (2016). *3D Recording, Documentation and Management of Cultural Heritage*. Whittles Publishing.
- Sun, J., Shen, Z., Wang, Y., Bao, H., & Zhou, X. (2021). LoFTR: Detector-Free Local Feature Matching with Transformers. DOI: 10.48550/arXiv.2104.00680
- Sun, S., Wei, W., Yuan, X., & Zhou, R. (2022). Research on Calibration Methods of Long-Wave Infrared Camera and Visible Camera. *Journal of Sensors*, 2022(8667606). DOI: 10.1155/2022/8667606
- Sutherland, N., Marsh, S., Priestnall, G., Bryan, P., & Mills, J. (2023). InfraRed Thermography and 3D-Data Fusion for Architectural Heritage: A Scoping Review. *Remote Sensing*, 15(2422). DOI: 10.3390/rs15092422
- Sutherland, N., Marsh, S., Remondino, F., Perda, G., Bryan, P., & Mills, J. (2025a). Geometric Calibration of Thermal Infrared Cameras: A Comparative Analysis for Photogrammetric Data Fusion. *Metrology*, 5(43). DOI: 10.3390/metrology5030043
- Sutherland, N., Morelli, L., Mills, J., Bryan, P., Marsh, S., & Remondino, F. (2025b). Multi-Modal and Multi-Sensor Photogrammetric Data Fusion Exploiting a New Repository for Infrared Thermography Datasets. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (Under Review).
- Swamidoss, I. N., Bin Amro, A., & Sayadi, S. (2021). Systematic Approach for Thermal Imaging Camera Calibration for Machine Vision Applications. *Optik*, 247(168039). DOI: 10.1016/j.ijleo.2021.168039

- Szostak, R., Zimnoch, M., Wachniew, P., & Jasek-Kamińska, A. (2023). Self-Calibration of UAV Thermal Imagery Using Gradient Descent Algorithm. *Drones*, 7(11), 683. DOI: 10.3390/drones7110683
- Tejedor, B., Lucchi, E., Bienvenido-Huertas, D., & Nardi, I. (2022). Non-Destructive Techniques (NDT) for the Diagnosis of Heritage Buildings: Traditional Procedures and Futures Perspectives. *Energy and Buildings*, 263(112029). DOI: 10.1016/j.enbuild.2022.112029
- Tempelhahn, A., Budzier, H., Krause, V., & Gerlach, G. (2016). Shutter-less Calibration of Uncooled Infrared Cameras. *Journal of Sensors and Sensor Systems*, 5, 9–16. DOI: 10.5194/jsss-5-9-2016
- The Intergovernmental Panel on Climate Change (IPCC). (2022). *Global Warming of 1.5°C: IPCC Special Report on Impacts of Global Warming of 1.5°C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (1st ed.). Cambridge University Press.
- This is Nottingham. (2009). Global Fire takes Lenton Lodge. *Nottingham Post*. Retrieved October 1, 2025, from Available from: <https://web.archive.org/web/20141021014820/http://www.nottinghampost.com/Global-takes-Lenton-Lodge/story-12259488-detail/story.html>
- Treier, S., Herrera, J. M., Hund, A., Kirchgessner, N., Aasen, H., Walter, A., & Roth, L. (2024). Improving Drone-Based Uncalibrated Estimates of Wheat Canopy Temperature in Plot Experiments by Accounting for Confounding Factors in a Multi-View Analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 218, 721–741. DOI: 10.1016/j.isprsjprs.2024.09.015
- Trentino.com. (2025). Castello di Arco. Retrieved October 1, 2025, from Available from: <https://www.trentino.com/en/highlights/castles/castello-di-arco/>
- Truong, T. P., Yamaguchi, M., Mori, S., Nozick, V., & Saito, H. (2017). Registration of RGB and Thermal Point Clouds Generated by Structure From Motion.

- Proceedings of the 2017 IEEE International Conference on Computer Vision Workshops*, 419–427. DOI: 10.1109/ICCVW.2017.57
- Tsilimantou, E., Delegou, E. T., Nikitakos, I. A., Ioannidis, C., & Moropoulou, A. (2020). GIS and BIM as Integrated Digital Environments for Modeling and Monitoring of Historic Buildings. *Applied Sciences*, 10(1078). DOI: 10.3390/app10031078
- Tyszkiewicz, M. J., Fua, P., & Trulls, E. (2020). DISK: Learning Local Features with Policy Gradient. DOI: 10.48550/ARXIV.2006.13566
- University of Southampton Archaeology. (2025). Mansfield Stone - White. Retrieved August 16, 2023, from Available from: https://archaeologydataservice.ac.uk/archives/view/stones_ahrb_2005/cfm/Public/details/RockDetails.cfm?RockCode=MANSWH&BCnt=5&Choice=1
- Ursine, W., Calado, F., Teixeira, G., Diniz, H., Silvino, S., & De Andrade, R. (2012). Thermal / Visible Autonomous Stereo Visio System Calibration Methodology for Non-controlled Environments. *11th International Conference on Quantitative InfraRed Thermography (QIRT 2012)*, 1–11. DOI: 10.21611/qirt.2012.261
- Usamentiaga, R., Garcia, D., Ibarra-Castanedo, C., & Maldague, X. (2017). Highly Accurate Geometric Calibration for Infrared Cameras using Inexpensive Calibration Targets. *Measurement*, 112, 105–116. DOI: 10.1016/j.measurement.2017.08.027
- Usamentiaga, R., Ibarra-Castanedo, C., & Maldague, X. (2018). Comparison and Evaluation of Geometric Calibration Methods for Infrared Cameras to Perform Metric Measurements on a Plane. *Applied Optics*, 57(18), 1–10. DOI: 10.1364/ao.57.0000d1
- Usamentiaga, R., Venegas, P., Guerediaga, J., Vega, L., Molleda, J., & Bulnes, F. (2014). Infrared Thermography for Temperature Measurement and Non-Destructive Testing. *Sensors*, 14, 12305–12348. DOI: 10.3390/s140712305

- van Eck, N. J., & Waltman, L. (2010). Software Survey: VOSviewer, A Computer Program for Bibliometric Mapping. *Scientometrics*, 84(2), 523–538. DOI: 10.1007/s11192-009-0146-3
- Velesaca, H. O., Bastidas, G., Rouhani, M., & Sappa, A. D. (2024). Multimodal Image Registration Techniques: A Comprehensive Survey. *Multimedia Tools and Applications*, 83, 63919–63947. DOI: 10.1007/s11042-023-17991-2
- Verdie, Y., Yi, K. M., Fua, P., & Lepetit, V. (2014). TILDE: A Temporally Invariant Learned DETector. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.1411.4568
- Vidas, S., Moghadam, P., & Bosse, M. (2013). 3D Thermal Mapping of Building Interiors Using an RGB-D and Thermal Camera. *Proceedings of the 2013 IEEE International Conference on Robotics and Automation*, 2311–2318. DOI: 10.1109/ICRA.2013.6630890
- Vidas, S., & Sridharan, S. (2012). Hand-held Monocular SLAM in Thermal-Infrared. *Proceedings of The 12th International Conference on Control Automation Robotics & Vision (ICARCV), Th35.3*, 859–864. DOI: 10.1109/ICARCV.2012.6485270
- Virtue, J., Turner, D., Williams, G., Zeliadt, S., McCabe, M., & Lucieer, A. (2021). Thermal Sensor Calibration for Unmanned Aerial Systems Using an External Heated Shutter. *Drones*, 5(119). DOI: 10.3390/drones5040119
- Vollmer, M., & Möllmann, K.-P. (2018). *Infrared Thermal Imaging: Fundamentals, Research and Applications* (2nd). Wiley-VCH Verlag GmbH & Co. KGaA.
- Wakeford, Z. E., Chmielewska, M., Hole, M. J., Howell, J. A., & Jerram, D. A. (2019). Combining Thermal Imaging with Photogrammetry of an Active Volcano Using UAV: An Example from Stromboli, Italy. *The Photogramm. Record*, 34(168), 445–466. DOI: 10.1111/phor.12301
- Waldhäusl, P., & Ogleby, C. (1994). 3 x 3 Rules for Simple Photogrammetric Documentation of Architecture. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 30, 426–429. Retrieved

- September 22, 2025, from Available from: https://www.cipaheritage__documentation.org/wp-content/uploads/2017/02/CIPA__3x3_rules__20131018.pdf
- Wan, Q., Brede, B., Smigaj, M., & Kooistra, L. (2021). Factors Influencing Temperature Measurements from Miniaturized Thermal Infrared (TIR) Cameras: A Laboratory-Based Approach. *Sensors*, *21*(8466). DOI: 10.3390/s21248466
- Wan, Q., Smigaj, M., Brede, B., & Kooistra, L. (2024). Optimizing UAV-based Uncooled Thermal Cameras in Field Conditions for Precision Agriculture. *International Journal of Applied Earth Observation and Geoinformation*, *134*(104184). DOI: 10.1016/j.jag.2024.104184
- Wang, C., Cho, Y. K., & Gai, M. (2013). As-Is 3D Thermal Modeling for Existing Building Envelopes Using a Hybrid LIDAR System. *ASCE Journal of Computing in Civil Engineering*, *27*(6), 645–656. DOI: 10.1061/(ASCE)CP.1943-5487.0000273
- Wang, Z., Zhou, J., Ma, J., Wang, Y., Liu, S., Ding, L., Tang, W., Pakezhamu, N., & Meng, L. (2023). Removing Temperature Drift and Temporal Variation in Thermal Infrared Images of a UAV Uncooled Thermal Infrared Imager. *ISPRS Journal of Photogrammetry and Remote Sensing*, *203*, 392–411. DOI: 10.1016/j.isprsjprs.2023.08.011
- Wang, Z., Zhou, J., Tang, W., Zhou, X., Zhou, B., Göttsche, F.-M., & Liu, S. (2025). Temporal Normalization of UAV Thermal Infrared Data From Long-Duration Flights. *IEEE Transactions on Geoscience and Remote Sensing*, *63*. DOI: 10.1109/TGRS.2025.3553563
- Weinmann, M., Leitloff, J., Hoegner, L., Jutzi, B., Stilla, U., & Hinz, S. (2014). Thermal 3D Mapping for Object Detection in Dynamic Scenes. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *II-1*, 53–60. DOI: 10.5194/isprsannals-II-1-53-2014

- Westfeld, P., Mader, D., & Maas, H.-G. (2015). Generation of TIR-attributed 3D Point Clouds from UAV-based Thermal Imagery. *Photogrammetrie - Fernerkundung - Geoinformation*, 5, 381–393. DOI: 10.1127/1432-8364/2015/0274
- Whitman, C. J., Prizeman, O., Walker, P., & Gwilliam, J. A. (2020). Heritage Retrofit and Cultural Empathy: A Discussion of Challenges Regarding the Energy Performance of Historic UK Timber-framed Dwellings. *International Journal of Building Pathology and Adaptation*, 38(2). DOI: 10.1108/IJBPA-02-2019-0023
- Wien, W. (1897). XXX. On the Division of Energy in the Emission-Spectrum of a Black Body. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 43(262), 214–220. DOI: 10.1080/14786449708620983
- Williams, J., Corvaro, F., Vignola, J., Turo, D., Marchetti, B., & Vitali, M. (2022). Application of Non-Invasive Active Infrared Thermography for Delamination Detection in Fresco. *International Journal of Thermal Sciences*, 171(107185). DOI: 10.1016/j.ijthermalsci.2021.107185
- Wood, G. (2009). Lenton Lodge Gatehouse, Derby Road, Nottingham (1823, Grade II listed). Available from: <https://www.flickr.com/photos/gwoodward/3254504564/in/photostream/>
- Workswell s.r.o. (2020). Workswell WIRIS Pro User Manual [Version: 201215 EN]. Retrieved January 13, 2024, from Available from: <https://visionaerial.com/wp-content/uploads/wwp-user-manual-en-201215.pdf>
- Workswell s.r.o. (2021). WIRIS Pro: Industrial Drone Thermal Camera. Retrieved July 10, 2023, from Available from: <https://workswell-thermal-camera.com/workswell-wiris-pro-thermal-inspection-camera-with-ultrazoom/>
- Yang, M.-D., Su, T.-C., & Lin, H.-Y. (2018). Fusion of Infrared Thermal Image and Visible Image for 3D Thermal Model Reconstruction Using Smartphone Sensors. *Sensors*, 18(2003). DOI: 10.3390/s18072003
- Yastikli, N., & Guler, E. (2013). Performance Evaluation of Thermographic Cameras for Photogrammetric Measurements. *The International Archives of the Pho-*

- togrammetry, Remote Sensing and Spatial Information Sciences, XL-1/W1*, 383–387. DOI: 10.5194/isprsarchives-XL-1-W1-383-2013
- Yi, K. M., Trulls, E., Lepetit, V., & Fua, P. (2016). LIFT: Learned Invariant Feature Transform [Version Number: 2]. *Computer Vision and Pattern Recognition (ECCV 2016)*. DOI: 10.48550/ARXIV.1603.09114
- Yu, G., & Morel, J.-M. (2011). ASIFT: An Algorithm for Fully Affine Invariant Comparison. *Image Processing On Line, 1*, 11–38. DOI: 10.5201/ipol.2011.my-asift
- Yuan, W., & Hua, W. (2022). A Case Study of Vignetting Nonuniformity in UAV-Based Uncooled Thermal Cameras. *Drones, 6*(394). DOI: 10.3390/drones6120394
- Zhang, C., Zou, Y., Dimyadi, J., & Chang, R. (2023a). Thermal-textured BIM Generation for Building Energy Audit with UAV Image Fusion and Histogram-based Enhancement. *Energy and Buildings, 301*(113710). DOI: 10.1016/j.enbuild.2023.113710
- Zhang, D., Liu, Y., Zhao, Y., Liang, J., Sun, B., & Chu, S. (2023b). Algorithm Research on Detail and Contrast Enhancement of High Dynamic Infrared Images. *Applied Sciences, 13*(12649). DOI: 10.3390/app132312649
- Zhang, J., Zhang, S., Peng, G., Zhang, H., & Wang, D. (2022). 3DRadar2ThermalCalib: Accurate Extrinsic Calibration between a 3D mmWave Radar and a Thermal Camera Using a Spherical-Trihedral. *25th International Conference on Intelligent Transportation Systems (ITSC)*, 2744–2749. DOI: 10.1109/ITSC55140.2022.9922522
- Zhang, Z. (2000). A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(11), 1330–1334. DOI: 10.1109/34.888718
- Zhao, T., Niu, H., Anderson, A., Chen, Y., & Viers, J. (2018). A Detailed Study on Accuracy of Uncooled Thermal Cameras by Exploring the Data Collection Workflow. In J. A. Thomasson, M. McKee, & R. J. Moorhead (Eds.),

- Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping III* (p. 13). SPIE. DOI: 10.1117/12.2305217
- Zhao, X., Wu, X., Chen, W., Chen, P. C. Y., Xu, Q., & Li, Z. (2023). ALIKED: A Lighter Keypoint and Descriptor Extraction Network via Deformable Transformation [Version Number: 2]. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.2304.03608
- Zhao, X., Wu, X., Miao, J., Chen, W., Chen, P. C. Y., & Li, Z. (2021). ALIKE: Accurate and Lightweight Keypoint Detection and Descriptor Extraction. *Computer Vision and Pattern Recognition (CVPR)*. DOI: 10.48550/ARXIV.2112.02906
- Zhou, Q.-Y., Park, J., & Koltun, V. (2016). Fast Global Registration. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer Vision and Pattern Recognition (ECCV 2016)* (pp. 766–782). DOI: 10.1007/978-3-319-46475-6_47
- Zhu, J., Xu, Y., Ye, Z., Hoegner, L., & Stilla, U. (2021). Fusion of Urban 3D Point Clouds with Thermal Attributes using MLS Data and TIR Image Sequences. *Infrared Physics & Technology*, 113(103622). DOI: 10.1016/j.infrared.2020.103622

Appendices

Appendix A

Buildings

A.1 Lenton Lodge Gatehouse, Nottinghamshire, UK

Designed by Sir Jeffry Wyattville between 1820-1825, Lenton Lodge Gatehouse, the historical entrance to Wollaton Hall, represents one of the finest examples of Elizabethan Revival in Britain (Figure A.1)(Pevsner & Williamson, 1979). After acquiring additional pastures South-East of Wollaton Hall, expanding the boundary walls of the 500-acre parkland out to Derby Road and Nottingham Canal (Figure A.8), Henry Willoughby, 6th Lord Middleton sought to commission a grand entrance to the park mimicking the style of the main house (Lenton Times, 1983). Built from ‘white’ Mansfield sandstone, a medium-grained sandy dolomitic stone quarried locally since the Medieval period (also used for Nottingham Castle and Newark Town Hall), the structure features a central carriage arch built from thick ashlar blocks surrounded by two domed corner turrets (Ashurst & Dimes, 1998; University of Southampton Archaeology, 2025). Flanked by two screen walls, moulded plinths and extended by cast-iron bollards, the central gatehouse features a three-



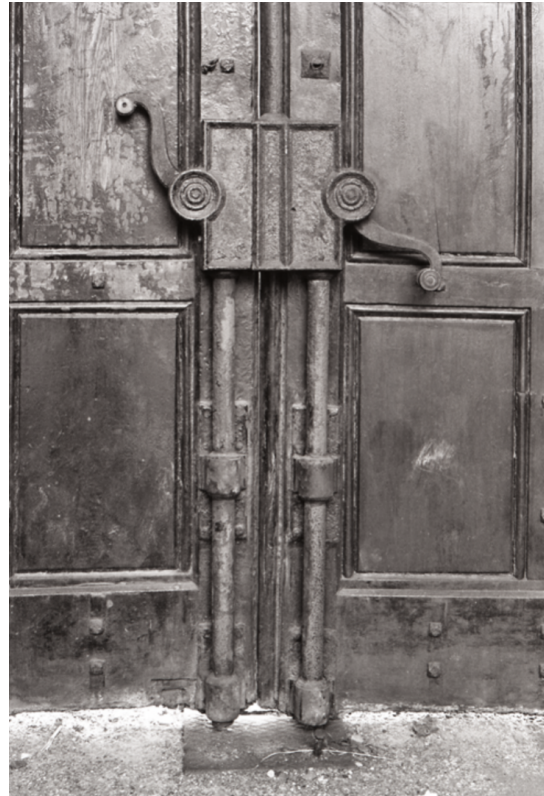
Figure A.1: Lenton Lodge Gatehouse, also called ‘Lenton Lodge’ or simply Wollaton Park ‘Gateway’, photographed by J. Spree (data unknown). Here, the gatehouse can be seen open with Nottingham canal beneath the foreground (Lenton Times, 2006).

storey structure with framed wooden gates and internal coffered ashlar barrel vault (Pevsner & Williamson, 1979). Importantly, the original central gate, featuring an underground iron opening mechanism, survives as a one-of-a-kind design (Figure A.2). A copper coat of arms plaque, located under the central parapet, shows the original seal of the Middleton family and remains a central feature on the building’s façade (Figure A.3)(Historic England, 2025a).

After being sold by Michael Willoughby, 11th Lord Middleton for £200,000 in 1924, along with large portions of parkland and Wollaton Hall itself, Lenton Lodge Gatehouse was acquired by several owners through long-term leases and purchase agreements. Throughout the early 20th Century, the building fell into a state of disrepair, separated from Wollaton Hall by large areas of housing development in the 1920s and the closure of the canal in 1937. To protect the structure (with suggestions that the building be demolished or filled with concrete to preserve the deteriorating



(a)



(b)

Figure A.2: Lenton Lodge Gatehouse's mechanical locking system, featuring (a) the underground 'winding box' and (b) central gate's ornate pins (Lenton Times, 1983).

frame), the building was listed by English Heritage in 1952, receiving a Grade II* status which remains today (ref. 1270829)(Historic England, 2025a; Lenton Times, 1983). In the 1981, the property was purchased by Lawrie Williamson on a 99-year lease, renovating one of the gatehouse apartments into a family dwelling and design studio, subsequently taken over by several private owners for residential and commercial purposes into the 21st Century. In 2016, after several decades of repair and eventual decay (Clark, 2006), including the 'award winning' restoration work undertaken by Chek Whyte (Figure A.3), the property was sold to the University of Nottingham who currently maintain and operate within the property (Commercial News Media, 2016).

Currently, Lenton Lodge Gatehouse shows visible signs of material degradation, no



(a)



(b)

Figure A.3: Comparison between the façade of Lenton Lodge Gatehouse in (a) 2007, after extensive restoration works by Check Whyte Industries, and (b) 2024, in its current state. Noticeably, the central parapet shows significant difference between the (a) recently restored and (b) heavily deteriorated stone above the traceried panels and central plaque, likely due to leaking from the roof. (This is Nottingham, 2009; Wood, 2009).

longer in regular use by the university or maintained within a broader conservation plan. On the exterior of the building, extensive areas of stone spalling are apparent under the parapet cornice and turrets, resulting in discolouration, biological colonisation and areas of exposed sub-stone. This is likely a product of leaking from the lead roof, a problem that has plagued the structure for several decades (the presence of shrubbery growing out of the parapet also indicates a substantial drain issue)(Figure A.3)(Lenton Times, 1983). Worryingly, this issue has become exponentially worse during the four years of this research, meaning essential work is needed to stabilise the material loss of historic fabric and large-scale interventions. In addition, the same issues seem to be affecting the traceried panels and bronze plaque at the top of the façade, identifiable from both external and internal inspections (Figure A.4).



Figure A.4: Images from the internal building survey of Lenton Lodge Gatehouse. Here, the rooms behind the (a) north turret, (b) central traceried panels (the other side of the circular windows) and (c) bronze plaque (behind the brick wallpaper) are shown, with visible signs of water ingress, material decay and damage identifiable.

A.2 Castello Di Arco, Trentino, Italy

Castello di Arco, a historic fortified castle that has its roots atop a rocky cliff of Sarca Valley since ~ 1000 AD, bears a rich history of occupation and conflict. Overlooking the Sarca river terminating at Riva del Garda, within the broader region of Trentino-



Figure A.5: Castello di Arco and its ‘Torre Renghera’ (bell tower), overlooking the Sarca Valley down to Riva del Garda (Trentino.com, 2025).

Alto Adige in northern Italy (Figure A.9), the first known structures were built atop the mountainous landscape in the Roman era, with the name ‘Arco’ likely originating from the Latin (‘Arx’) or Occitan (‘Arc-’) terms meaning ‘fortress’ or ‘rock tower’, respectively (Italy This Way, 2025) (the ‘Torre Renghera’ is considered the first structure to have been built around 500 AD)(GardaSee, 2025). The greater fortress is believed to have been built by the Arco commune themselves as a strategic defensive fortress, later granted to the Arco family in 1124 AD. The Counts of Arco, holding the castle as a fortification from the 10th to 17th Centuries, withstanding several military expulsions and re-acquisitions, held the castle until its seizure by Emperor Leopold I in 1680. In 1703, after decades of deterioration and occupation by local bandits, Castello di Arco was sieged by French troops under the command of General Vendôme during the War of the Spanish Succession, later ordered for demolition (GardaSee, 2025).

Bought by the Servizio Beni Culturali of the Provincia Autonoma di Trento (Au-



(a)



(b)

Figure A.6: Castello di Arco depicted by (a) Matthias Burglechner (date unknown) and (b) Albrecht Dürer (1494), the latter of which now resides in the Louvre museum, Paris, France. The paintings are characterised by the large rocky ridge on top of which Castello di Arco resides, surrounded by cypresses, holm oaks, olive groves and cultivated vines (Storia dell'Arte, 2025).

onomous Province of Trento) in 1982, extensive restoration works were undertaken between 1986-2003 to repair the 23,000 m² site and open Castello di Arco to the public, including the stabilisation of the central tower, the transformation of surviving rooms and greens into exhibition spaces, and the discovery of the famed Sala Degli Fresco in a previously rubble-filled room. The painted frescoes, covering four walls of a small salon, mark the most significant part of Castello di Arco, depicting various scenes of court life and chivalry in the 14th Century. This includes: scenes depicting ladies and gentlemen playing chess and dice games; St. George defeating the dragon, a knight accompanying two dames picking roses from a garden for garlands; and a gentlemen assisting a lady on horseback (Figure A.7)(Storia dell'Arte, 2025; Trentino.com, 2025). However, the passage of time has caused severe deterioration of the frescoes, with the majority of the paintings now faded and flaking from the stone wall. This decay has been accelerated by the water ingress and moisture from the North-West rock wall of the Sala, with few strategies able to reduce the influence of impeding moisture. At present, the museum utilises a space heater in

a bid to reduce the humidity within the room; however, as evident in the results in Section 4.4.2, this only benefits a very small area the frescoes.



Figure A.7: Castello di Arco’s famed frescoes, depicting court life include: (a) ladies gesticulating whilst playing dice, (b) court ladies and gentlemen playing chess, and (c) a gentlemen and two ladies collecting roses for garlands (Storia dell’Arte, 2025).

A.3 Dunrobin House, Warwickshire, UK

Dunrobin House, a 25-year-old contemporary dwelling in Warwickshire , provides an example of common 21st Century housing in the UK, featuring building materials, architectural features and building conditions commensurate with large portions of domestic properties (Figure A.10). Whilst the house does not represent ‘architectural heritage’, the building provides a suitable case study for the investigation of IRT-3DDF for energy efficiency, thermal comfort and energy retrofitting strategies. Upon initial TIR imaging surveys, corroborated by internal building inspections, areas of deteriorated loft insulation were identified at on the front of the roof, likely due to nesting of birds during the winter. This was later confirmed in both 2D TIR images and 3D thermal models in Section 5.4.3.

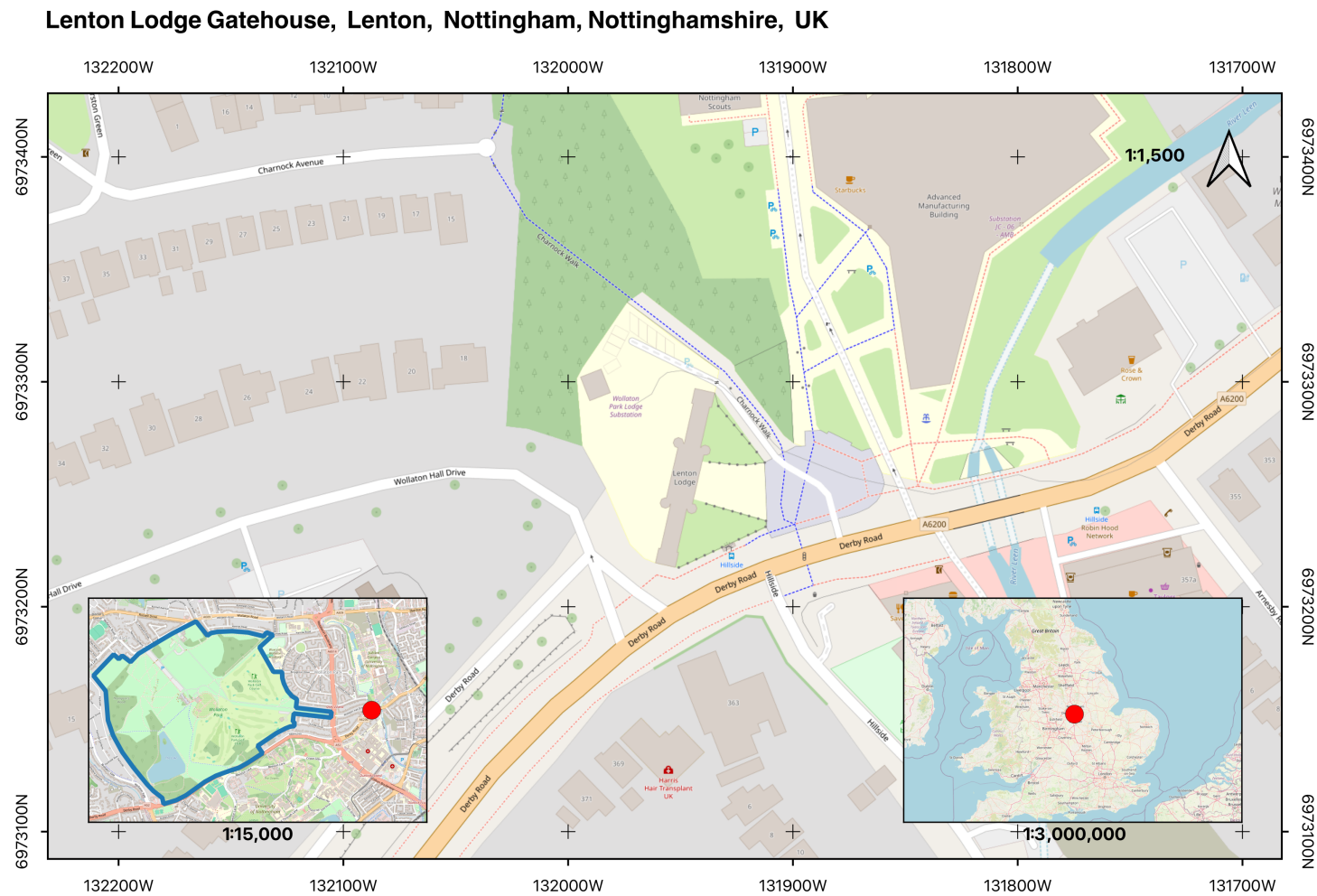


Figure A.8: Map of Lenton Lodge Gatehouse, Lenton, Nottingham, Nottinghamshire, UK.

Castello di Arco, Sarca Valley, Trentino, Italy

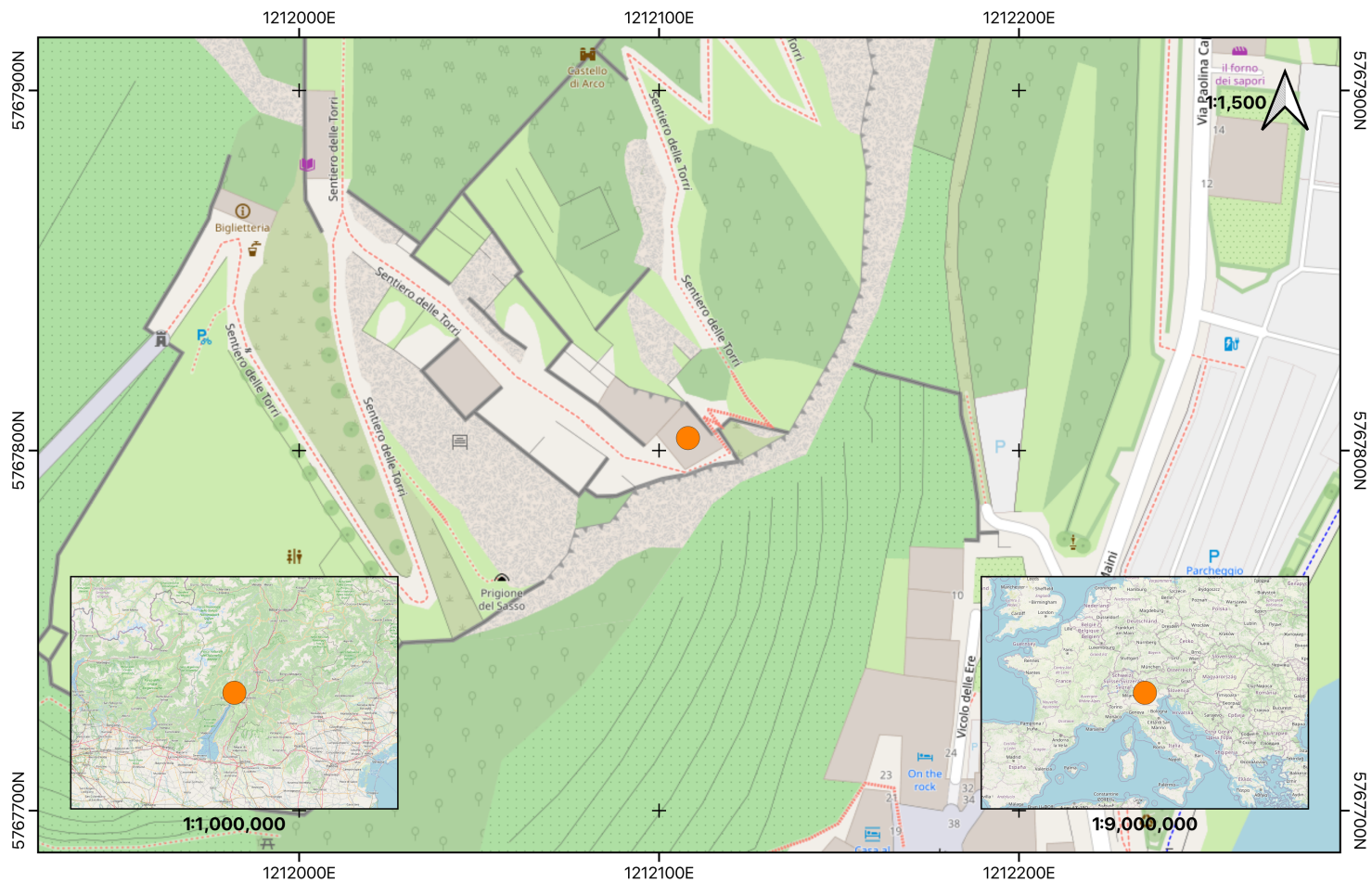


Figure A.9: Map of Castello di Arco, Trentino, Italy.

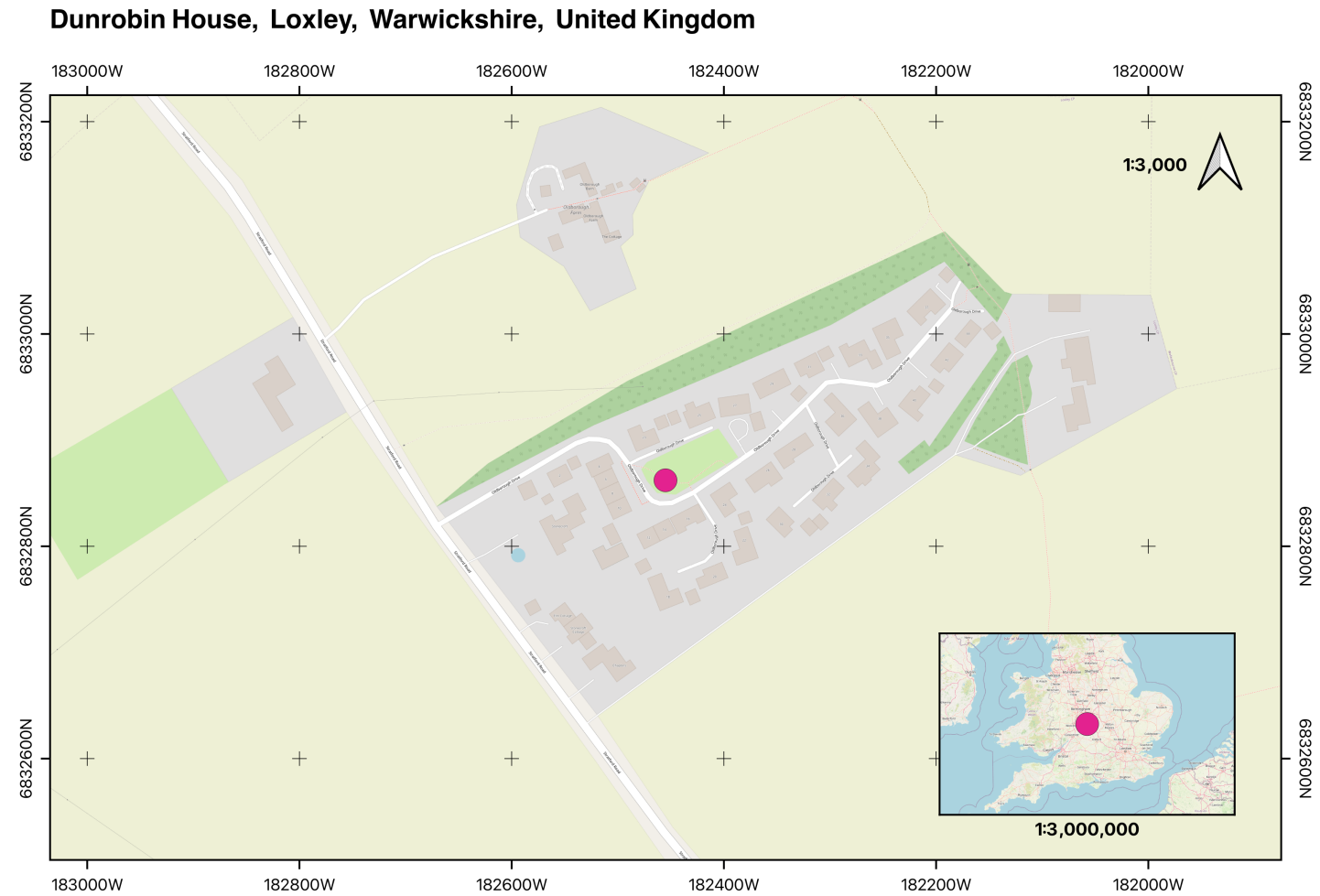


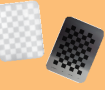
Figure A.10: Map of Dunrobin House, Warwickshire, UK.

Appendix B

IRT-3DDF Cheatsheet

Photogrammetric Thermal Image Capture - 3x3 Rules

Inspired by the 3x3 Rules presented by Peter Waldhäusl and Cliff Ogleby at the 1988 CIPA Symposium, this guide provides helpful insights for the capture of thermal data with the ambition of processing images photogrammetrically. This guide indicates critical considerations for the operation of a suitable thermal camera, the design of comprehensive surveys for data acquisition and the processing of thermal images.

	Thermal Cameras	Data Acquisition	Data Processing
General	Thermal Camera Specifications - The choice of IR camera is critical for the success of thermal modelling - Important camera features include: - FPA Sensor Size - Resolution (>640x480) - Wider FOV camera desirable - Thermal Sensitivity (<0.1°C) - Spectral Range (~ 8µm to 14µm) - Accuracy (~ +/- 2°C) - RAW temp. image formats provided - Integrated VIS sensor desirable 	Survey Conditions - Conditions during data capture critical to generate and apply effective thermal images - Survey notes should including: - Previous weather conditions (2hrs and 24hrs) - Wind speed - Atmospheric temperature - Relative humidity - Sun position and angle - Presence of shadows and occlusions - Any changes during survey 	General Photogrammetry - Inherent low resolution and low contrast makes traditional image matching difficult - Capture images perpendicular to subject - Add perspective images of full facades for an overview and improved alignment - Determine a suitable matching strategy and define a purpose for each dataset (RGB = geometry, TIR = temperature values) 
Geometric	Geometric Camera Calibration - Understanding the internal geometry of an IR camera is fundamental for photogrammetric processes - IR cameras with FPA sensors will follow the Brown Model ($f_x, f_y, c_x, c_y, k_1, k_2, p_1, p_2$) - Geometric camera calibration should be undertaken prior to photogrammetric processes, tailored for each sensor within the thermal camera 	Ground Control - Determining suitable ground control is a critical part of photogrammetric networks - Survey targets should be visible for both TIR and RGB cameras - This requires materials with different emissivity and reflectivity values - Suggested materials include: - Aluminium - Rubber - Copper - Woods - PLA 	Photogrammetric Layout / Coverage - Determine desired thermal GSD for specific purpose and data output - High overlap best for thermal images (< 80%) v = FOV of the camera (horizontal) (in degrees) x = distance away from target y = desired % overlap of captured images z = distance between concurrent camera positions $z = 2 - \frac{xy}{50} \tan\left(\frac{v}{2}\right)$
Radiometric	Radiometric Camera Calibration - Thermal cameras should be radiometrically calibrated with a blackbody every 1-2 years to obtain accurate temperature measurements (check with camera manufacturer) - Quantify how long the sensor takes to heat up and provide stable temperature measurements before data capture (typically 20-60 minutes) - Monitor camera housing, CPU and sensor temperatures during image capture (which may influence readings) 	Radiometric Consistency - Quality thermal images are produced by maximising image contrast and exploiting prime survey conditions - Choose between ACTIVE (artificial heating of subject) or PASSIVE (natural state) thermography - Set thermal camera emissivity value (ϵ) for dominant material in scene - Determine suitable time intervals for Non-Uniformity Corrections (NUCs) (a means of correcting camera temperature 'drifting') (camera dependent) 	Radiometric Correction - Post-processing image enhancing can be beneficial for matching processes to encourage more matches between TIR and RGB images (i.e. CLAHE, Unsharp Mask) - A network of known temperature measurements vital for the correction and validation of thermal images in post-processing - Utilise known temperature measurements to correct thermal image intensity histograms 

Created by Neil Sutherland, Geospatial Systems CDT, University of Nottingham
(neil.sutherland@nottingham.ac.uk)

Figure B.1: IRT-3DDF 'Cheatsheet'