# JOINT GEOMETRIC CALIBRATION OF COLOR AND THERMAL CAMERAS FOR SYNCHRONIZED MULTIMODAL DATASET CREATING

V. A. Knyaz<sup>1,2\*</sup>, P.V. Moshkantsev<sup>1</sup>

<sup>1</sup> State Research Institute of Aviation System (GosNIIAS), 125319, 7, Victorenko str., Moscow, Russia – (knyaz, moshkantsev)@gosniias.ru
<sup>2</sup> Moscow Institute of Physics and Technology (MIPT), Russia

### **Technical Commission II**

KEY WORDS: calibration, accuracy, thermal imaging, multimodal data, 3D reconstruction

## **ABSTRACT:**

With increasing performance and availability of thermal cameras the number of applications using them in various purposes grows noticeable. Nowadays thermal vision is widely used in industrial control and monitoring, thermal mapping of industrial areas, surveillance and robotics which output huge amount of thermal images. This circumstance creates the necessary basis for applying deep learning which demonstrates the state-of-the-art performance for the most complicated computer vision tasks. Using different modalities for scene analysis allows to outperform results of mono-modal processing, but in case of machine learning it requires synchronized annotated multimodal dataset. The prerequisite condition for such dataset creating is geometric calibration of sensors used for image acquisition. So the purpose of the performed study was to develop a technique for joint calibration of color and long wave infra-red cameras which are to be used for collecting multimodal dataset needed for the tasks of computer vision algorithms developing and evaluating.

The paper presents the techniques for camera parameters estimation and experimental evaluation of interior orientation of color and long wave infra-red cameras for further exploiting in datasets collecting. Also the results of geometrically calibrated camera exploiting for 3D reconstruction and 3D model realistic texturing based on visible and thermal imagery are presented. They proved the effectivity of the developed techniques for collecting and augmenting synchronized multimodal imagery dataset for convolutional neural networks model training and evaluating.

## INTRODUCTION

Recent impressive progress in technical characteristics of sensors of different types (vision, near infra-red, short, middle and long wave infra-red, ultra-violet etc.) allows using these sensors for solving sophisticated tasks in wide variety of applications: surveillance and industrial infra-structures monitoring, driver assistance and robotics, 3D scene analysis and understanding. The significant part of optical 3D reconstruction and analysis applications require not only qualitative but also quantitative analysis, and the required quantitive characteristics could be provided by calibration of the acquisition devices.

Techniques for geometric calibration of thermographic cameras has been proposed for different kinds of application such as energy pollution analysis in the urban area (Conte et al., 2018), architecture, civil engineering and industry inspection (Lagüela et al., 2011), non-destructive testing (Yang et al., 2018), assessing the energy efficiency of real estates (Borrmann et al., 2012), building survey (Iwaszczuk and Stilla, 2017), material testing (Luhmann et al., 2013) and others. They allow to estimate camera interior orientation parameters for further retrieving 3D scene geometry from thermal images.

Another outcome of growing accessibility and quality of different sensors is the increase of available data which is necessary for developing and evaluating of new state-of-the-art methods and algorithms. Dealing with big datasets allows to receive higher algo-

\*Corresponding author

rithms performance due to possessing representative information accounting wide variety of possible conditions of operation. This circumstance provides the progress in machine learning by applying deep learning techniques. Now state-of-the-art deep learning models show best results in many tasks of computer vision.

The possibility of using different modalities creates a good basis for improving the performance of machine vision methods, especially for degraded vision environment. For better exploiting the advantages of multimodality, corresponding images of different spectral range should be synchronized and calibrated. But in most cases thermal and color cameras used for multimodal scene analysis has different resolution, field of view and frame rate. This circumstance requires to perform a lot of manual work on data matching, orienting, aligning and labeling and interferes the collecting synchronized multimodal datasets which are needed for successful convolutional neural networks (CNN) training.

The performed study was focused on developing a technique for joint calibration of color and long wave infra-red cameras which are to be used for collecting multimodal dataset needed for the tasks of computer vision algorithms developing and evaluating. Two color and thermal camera pairs were calibrated for multimodal imagery acquisition and further processing for synchronization and annotation. Also the technique for automated data aligning and augmenting was proposed and evaluated. CNN training on the prepared in such way dataset allowed improving CNN performance in tasks of multimodal scene segmentation, object detection and re-identification.

### 1. RELATED WORK

Using multimodal color-thermal data is necessary for a noticeable set of computer vision and industrial applications. So the problem of getting geometrically accurate thermal imagery has received scholar attention recent decades when thermal cameras were widely used for different practical and scientific researches.

The overview of thermal imaging sensors for photogrammetric close-range applications and results of the geometric calibration of thermographic cameras as they are used for building inspection and material testing were given in (Luhmann et al., 2010) and (Luhmann et al., 2013). Interior orientation for a set of considered thermal cameras was estimated using two different (plane and spatial) test fields.

A calibration method (Choinowski et al., 2019) for a trifocal sensor including stereo camera and thermal imaging camera, operating in the Long Wave Infrared (LWIR) spectrum utilized for calibration a passive aluminum chessboard as a test field. After corner detection and subsequent bundle adjustment for all synchronized image triplets the remaining re-projection errors were at the sub-pixel level. Also chessboards with high contrast (as well as different markers suitable for extrinsic calibration) (Engstrom et al., 2013) were evaluated as test fields for geometric calibration thermal cameras with recommendation for indoor and outdoor scenarios. Three calibration methods for infrared cameras (Usamentiaga et al., 2018) has been studied compared to estimate the best. The performed study showed that a complete camera calibration method using a specifically designed calibration target outperforms both direct and iterative estimation of the transformation between image and world coordinates techniques.

To study potentials and limitations of Structure from Motion approach for the automated generation of thermal orthomosaics several different strategies (Conte et al., 2018) were applied for processing two thermal flights over a  $10 \ km^2$  area in Bologna city (Italy). The best results (absolute planimetric accuracy being at the level of 3 - 4 pixels) were obtained for performing camera calibration on a smaller subset of images, with a limited number of ground control points and an adaptive fitting algorithm.

A study for verifying of the geometric accuracy, repeatability and drift of thermographic cameras for further incorporating geometric information obtained by thermal cameras into the quality control processes (Lagüela et al., 2011), (Lagüela et al., 2012) has been performed using a grid of burning lamps as a test field. Evaluation of two thermographic cameras from FLIR and NEC manufacturers showed repeatability better than 1mm for all cases with geometric quality being slightly better for the NEC camera.

A geometric mask with high thermal contrast and not requiring a flood lamp (Vidas et al., 2012) has been used in automatic end-to-end system for calibrating single or multiple cameras. A clustering-based algorithm utilizing the maximally stable extremal region detector provided accurate locating calibration points on the pattern.

Several studies has been performed on estimating quality of imagebased 3D reconstruction using only calibrated thermal images (Knyaz et al., 2017), or combined color-thermal camera setup (Yang et al., 2018). A cost-effective method for 3D thermal model reconstruction based on image-based modeling (Yang et al., 2018) has been proposed for two smart phones and a low-cost thermal infrared camera. The experimental evaluation demonstrated the ability of the proposed method to effectively reconstruct a 3D thermal model.

Methods for geometric calibration of robots, equipped with a set of sensors (Borrmann et al., 2012), (Kniaz, 2017), (Dias et al., 2013) allows to reconstruct 3D model of environment containing thermal information. A robot equipped with a 3D laser scanner, a thermal camera, and a color camera has been used for creating high-precise heat distribution maps of so-called Passivhaus (or the Zero-energy) buildings (Borrmann et al., 2012). Two different calibration test fields were used for calibration of color and thermal cameras. After calibration procedure data from all the sensors are combined to model the environment precisely.

The calibration method (Dias et al., 2013) combined the utilization of special patterns for interior calibration of thermographic cameras, with the usage of a high-resolution 3D laser scanner for the extrinsic calibration, relating the cameras frames with the robot frame. The sensors setup was used in autonomous surface vehicle (ASV) ROAZ II for detection, location and tracking of human targets.

### 2. JOINT CAMERAS CALIBRATION

The purpose of joint color and thermal cameras calibration was to receive oriented aligned and synchronized color-thermal image pairs for the tasks of machine learning. Initial multimodal dataset was collected using two color-thermal camera pairs. The first sensor was FLIR ONE PRO camera for smartphones. It provides preliminary aligned and synchronized color and thermal images. The second sensor pair was color and thermal cameras mounted on DJI Mavic Pro unmanned aerial vehicle.

The collected dataset (Knyaz, 2019) contains more than 5000 pairs of color and thermal images of urban and suburb scenes gathered in different seasons, different times of day and various weather conditions.

Calibration of all cameras was carried out using the previously developed techniques for laboratory (Knyaz, 2010) and on-site calibration (Knyaz and Zheltov, 2017).

The vector  $v_e^i = (x_p, y_p, m_x, m_y, a_0, ..., a_5)^T$  of estimated interior orientation parameters for calibration includes coordinates of principal point  $x_p, y_p$ , image scales  $m_x, m_y$  and additional parameters  $a_0, ..., a_5$  accounting non-linear distortion in camera model (co-linearity equations). Preliminary calibration includes estimation of cameras interior and exterior orientation.

The calibration procedure uses the basic model of central projection for relation between point of a 3D scene and corresponding point of its image – the collinearity equation, expressing the condition that the point of the scene G, the center of the projection of O and the image of this point g lie on one straight line:

$$\mathbf{X}_G = \mathbf{X}_0 - \lambda \mathbf{A}^T \cdot (\mathbf{x}_g - \mathbf{x}_p) \tag{1}$$

Here

 $\mathbf{X}_0 = (X_0, Y_0, Z_0)$  – coordinates of the center of the projection,  $\mathbf{X}_G = (X, Y, Z)$  – scene point coordinates,

 $\mathbf{x}_g = (x, y, -f)$  – the corresponding coordinates of the scene point in the image,

A – coordinate system transformation matrix,

 $\mathbf{x}_p$  – coordinates of the main point of the snapshot,  $\lambda$  – scale factor.

For distortions accounting additional terms are introduced into the collinearity equations  $\Delta x$  and  $\Delta y$ , describing various distorting factors. Then the real (distorted)  $x_d$ ,  $y_d$  coordinates of a point in the image are defined as:

$$x_d = x + \Delta x; \tag{2}$$

$$y_d = y + \Delta y; \tag{3}$$

The practice of photogrammetric measurements have proved that the good description for nonlinear distortion is the following model (Beyer, 1992):

$$\Delta x = a_0 \cdot y_a + x_a (a_1 r^2 + a_2 r^4 + a_3 r^6) + a_4 (r^2 + 2x_a^2) + 2a_5 x_a y_a; \Delta y = a_0 \cdot x_a + y_a (a_1 r^2 + a_2 r^4 + a_3 r^6) + a_5 (r^2 + 2y_a^2) + 2a_4 x_a y_a;$$

with  $r^{2} = x_{a}^{2} + y_{a}^{2}$ 

## Here

 $x_a, y_a$  – coordinates of a point on the image,  $a_0, ..., a_5$  – camera interior orientation parameters:  $a_0$  – coefficient of affine distortion;  $a_1, a_2, a_3$  – coefficients of radial distortion;  $a_4, a_5$  – coefficients of tangential distortion.

The vector  $v_e^l = (x_p, y_p, m_x, m_y, a_0, ..., a_5)^T$  of estimated parameters for test field calibration includes coordinates of principal point, image scales and additional parameters correspondingly, spatial coordinates of reference points being known by independent precise measurements.

## 2.1 FLIR ONE PRO camera calibration

FLIR ONE PRO thermal camera (Figure 1) acquires synchronized visible images of  $1440 \times 1080$  pixels and thermal images of  $640 \times 480$  pixels. Thermal sensor works in  $8 - 14 \mu m$  spectral range and it has pixel size of  $12 \mu m$ .



Figure 1. FLIR ONE PRO camera.

Main technical characteristics of FLIR ONE Pro are given in Table 1.

Color and thermal sensors of FLIR ONE PRO are aligned to acquire aligned images. The camera generates output images in different modes: color, thermal gray, thermal rainbow, thermal ironbow and some others.

Table 1. FLIR ONE PRO camera specification

Thermal sensor	Pixel size 12 $\mu m$
Spectral range	$8-14 \ \mu m$
Thermal resolution	160x120
Visual resolution	1440x1080
HFOV / VFOV	55 °/ 43 °
Frame rate	8.7 Hz
Focus	Fixed 15 cm – Infinity
Scene dynamic range	-20 °C to 400 °C
Accuracy	$\pm$ 3 °C
Thermal sensitivity (MRTD)	$150 \ mK$
Operating temperature	0 – 35 °C
Size	68x34x14 mm

FLIR ONE PRO camera provides MSX (or Muti Spectral Dynamic imaging) technology which makes use of the high contrast visual image to emboss key aspect of a scene on top of full thermal image. This feature helps to easily identify the position of heat patterns in the captured image of a scene (Figure 2), but as Figure 2 demonstrates, it could not provide accurate matching of corresponding features.



Figure 2. MSX image of the test field

Laboratory joint color and thermal cameras calibration was performed using special test field (Figure 3) allowing to obtain a required contrast for thermal images.



Figure 3. Sample images used for joint calibration.

A plane metal plate with 49 accurate holes was used as a test field for calibration. For calibration of thermal camera different conditions of acquisition (positive and negative thermal contrast of the test field relatively to a background) were evaluated. The results of estimation of interior orientation parameters of FLIR ONE PRO cameras are presented in Table 2.

Table 2. Results of joint FLIR ONE PRO cameras calibration

Co	olor camera	The	rmal camera
$m_{x_c}$	0.01210001	$m_{x_t}$	0.01184560
$m_{y_c}$	0.01240495	$m_{y_t}$	0.01221366
$x_{p_c}$	301.0574	$x_{p_t}$	318.9313
$y_{p_c}$	140.8023	$y_{p_t}$	132.4496
$a_{0_c}$	-0.02317030	$a_{0_t}$	-0.01824143
$a_{1_c}$	0.00148937	$a_{1_t}$	0.00153720
$a_{2_c}$	0.00000517	$a_{2t}$	-0.00007570
$a_{3c}$	-0.00000178	$a_{3t}$	-0.00000178
$a_{4c}$	-0.00196251	$a_{4t}$	-0.00146680
$a_{5c}$	-0.00059759	$a_{5t}$	-0.00079429

The residuals on reference points after bundle adjustment procedure were at the level of 2.3 mm for the working space of  $2m \times 2m \times 2m$  as for positive thermal contrast as for negative thermal contrast.

## 2.2 UAV cameras calibration

The second color-thermal pair of cameras are mounted on DJI Mavic Pro (Figure 4) unmanned aerial vehicle (UAV). It is equipped with high resolution 12 Mpix CMOS camera with 80 degrees field of view (FOV) lens. Long wave infra-red ( $8-14\mu m$ ) camera MH-SM567-6 is additionally mounted on DJI Mavic Pro quad-copter. The MH-SM567-6 camera is of uncooled amorphous silicon FPA type with the resolution of  $640 \times 480$  pixels.



Figure 4. Color and thermal cameras installed on DJI Mavic quadcopter.

Main technical characteristics of MH-SM567-6 are given in Table 3.

For on-site calibration a modern building having regular structures (Figure 5) was used a test field for the calibration. Several reference distances were measured independently by laser range finder to provide metric characteristics.

For automated detection and identification of reference points an original algorithms were developed for reference points of circular shape (laboratory test field) and for "corner" reference points (on-site test field).

The vector of estimated parameters

$$\begin{aligned} \psi_e = & (x_p, y_p, m_x, m_y, a_0, ..., a_5, \\ & x_{1r}, y_{1r}, z_{1r}, ..., x_{nr}, y_{nr}, z_{nr})^T \end{aligned}$$

### Table 3. DJI Mavic Pro cameras specification

Thermal sensor	
Туре	Uncooled FPA
FPA format	640*480
Pixel pitch	$17 \mu m$
Sensitivity	$\leq 150 \ mK$ at $f/1.0 \ 300 K$
Frame rate	50Hz
Spectral range	$8-14\mu m$
Lens	12.8 mm
Color sensor	
Sensor	Type: 1/2.3" CMOS
Effective Pixels:	12.35 million
Lens FOV:	78.8°
Focal Length:	28 mm
Aperture:	f/2.2
Distortion:	$\leq 1.5\%$
Focus Range:	$0.5m~to~\infty$



Figure 5. Testfield used for on-site camera calibration.

for on-site calibration includes interior orientation parameters  $v_e^i$  and spatial coordinates of detected reference points  $X_r$  being estimating along with the parameters of interior orientation. Several reference distances between reference points were measured independently by a laser rangefinder. They were used to determine accurate scale of the test field.

The resulting residuals at the reference points after bundle adjustment were at the level of 60 mm. This accuracy is quite enough for the task of scene 3D reconstruction and photorealistic texturing for creating multimodal dataset.

## 3. APPLYING CALIBRATION RESULTS FOR DATASET CREATION

## 3.1 Scene 3D model reconstruction

To provide synchronization of collected color and thermal images the technique based on scene 3D reconstruction (Knyaz, 2019) was applied. Both color and thermal image sequences were used for scene 3D reconstruction by structure from motion technique.

The algorithm for scene 3D model reconstruction (Knyaz and Zheltov, 2017) includes the following steps. Firstly, a set of corresponding points are detected in a given image sequence. Then initial stereo pair of images with a sufficient stereo basis and overlapping of 60-70% is chosen, and its relative orientation is estimated using detected corresponding points. The spatial coordinates in the basis coordinate system for corresponding points visible in these images are calculated using determined relative orientation parameters of the images.

Then for the next image in the sequence an exterior orientation is performed using the image coordinates of the detected points and the spatial coordinates of the points found at the previous step. The spatial coordinates of new detected points visible in the current image and in the previous image are calculated using the determined parameters of the exterior orientation. The new calculated 3D points are added to the 3D model. Also spatial coordinates of previously calculated points are re-calculated using corrected exterior orientation parameters. Then the procedure of the exterior orientating of a next image and adding new points in the 3D model is repeated.

After determining the exterior orientation for all images, a more detailed model is generated using the correlation algorithm for correspondence problem solving. Some reference distances between the distinctive features of the surveyed scene are measured independently along with the image acquisition. They allow estimating the real scale of the reconstructed 3D model.



Figure 6. Scene 3D model created using color and thermal images

Such procedure is performed as for RGB as for LWIR image sequences resulting in two 3D models (RGB-based 3D model and LWIR-based 3D model), which are determined in its own system of coordinates.

Samples of resulting 3D models built from LWIR and RGB image sequences are shown in Figure 6 .

## 3.2 Camera orientation

The technique of scene 3D model reconstruction provides external orientation parameters for all images used for 3D model generation. These parameters are determined in the coordinate system related to the resulting 3D model as for RGB-based as for LWIR-based 3D models. So it is needed to transform them in one common coordinate system for performing accurate 2D labelling and establishing correspondence between color and thermal images.

To find the parameters of coordinate system transformation the following technique was used. Firstly 3D coordinates of feature

descriptors were found using image coordinates of the descriptor and image orientation parameters. This procedure results in two 3D coordinate arrays corresponding to RGB-based and to LWIRbased scene 3D models. The parameters of system coordinates transformation from LWIR-based 3D model to RGD-based 3D model were found as least mean squares solution for the distances between corresponding points.



Figure 7. Image orientation and synchronization

Non-linear minimization of the observation (re-projection) errors by bundle adjustment (Knyaz, 2010) is used to determine camera exterior orientation for images not used in 3D reconstruction procedure. Redundant number of reference points is used in bundle adjustment procedure of minimizing the squared re-projection error for the 2D image coordinates of reference points  $x_{ij}$  in an image j as a function of the unknown image exterior orientation parameters (R, X) and known reference 3D point coordinates  $p_i$ using non-linear least squares. For the projection equations:

$$x_{ij} = f(p_i, R_j, X_j), \tag{4}$$

the iteratively minimized re-projection errors are

$$E = \sum_{i,j} \left( \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial R} \Delta R + \frac{\partial f}{\partial X} \Delta X - r_{ij} \right), \quad (5)$$

where  $r_{ij} = x_{ij} - \hat{x}_{ij}$  is the current residual vector (2D error in the predicted position) and the partial derivatives are with respect to the unknown orientation parameters (camera rotation and translation).

After this procedure scene 3D model and image exterior orientation in this 3D scene are determined thus providing the conditions for automatic image labeling for objects located in the 3D scene. Also it allows to extend the dataset by creating new synthetic images of the scene by a virtual camera.

The multimodal dataset generated and augmented using this technique was used for CNN training for the tasks of object detection and object re-identification (Knyaz, 2019), (Kniaz and Knyaz, 2019). The evaluation of CNN trained on the created multimodal dataset showed improving of the CNN performance for considered tasks.

#### 4. CONCLUSION

The techniques for laboratory and on-site joint geometric calibration of color and long wave infra-red cameras were developed. For laboratory calibration plane test field providing the possibility of acquiring thermal images of as positive as negative contrast is used. The results of calibration in different conditions showed that interior orientation parameters do not depend significantly on contrast sign. A modern building having regular structures served as a test field for on-site calibration, and its character features (window corners) being used as reference points.

The results of joint color and long wave infra-red cameras calibration were used for processing of multimodal image dataset collected for the developing and evaluating convolutional neural network models. The developed technique for based on joint geometric calibration allows aligning, synchronizing and annotating multimodal images in automated mode.

Exploiting of the generated multimodal imagery dataset for CNN models training allows to improve deep learning performance for multimodal object recognition and re-identification.

### ACKNOWLEDGEMENTS

The reported study was funded by Russian Foundation for Basic Research (RFBR) according to the project No 17-29-04509.

### REFERENCES

Beyer, H., 1992. Advances in characterization and calibration of digital imaging systems. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XXIX, pp. 545–555.

Borrmann, D., Afzal, H., Elseberg, J. and Nüchter, A., 2012. Mutual calibration for 3d thermal mapping. *IFAC Proceedings Volumes* 45(22), pp. 605 – 610. 10th IFAC Symposium on Robot Control.

Choinowski, A., Dahlke, D., Ernst, I., Pless, S. and Rettig, I., 2019. Automatic calibration and co-registration for a stereo camera system and a thermal imaging sensor using a chessboard. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-2/W13, pp. 1631–1635.

Conte, P., Girelli, V. A. and 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, pp. 320 – 333.

Dias, A., Brás, C., Martins, A., Almeida, J. and Silva, E., 2013. Thermographic and visible spectrum camera calibration for marine robotic target detection. In: *2013 OCEANS - San Diego*, pp. 1–5.

Engstrom, P., Larsson, H. and Rydell, J., 2013. Geometric calibration of thermal cameras. In: G. W. Kamerman, O. K. Steinvall, G. J. Bishop and J. D. Gonglewski (eds), *Electro-Optical Remote Sensing, Photonic Technologies, and Applications VII; and Military Applications in Hyperspectral Imaging and High Spatial Resolution Sensing*, Vol. 8897, International Society for Optics and Photonics, SPIE, pp. 104 – 111.

Iwaszczuk, D. and 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, pp. 33 – 47.

Kniaz, V. V., 2017. A photogrammetric technique for generation of an accurate multispectral optical flow dataset. In: F. Remondino and M. R. Shortis (eds), *Videometrics, Range Imaging, and Applications XIV*, Vol. 10332, International Society for Optics and Photonics, SPIE, pp. 104 – 114.

Kniaz, V. V. and Knyaz, V. A., 2019. Chapter 6 - multispectral person re-identification using gan for color-to-thermal image translation. In: M. Y. Yang, B. Rosenhahn and V. Murino (eds), *Multimodal Scene Understanding*, Academic Press, pp. 135 – 158.

Knyaz, V., 2019. Multimodal data fusion for object recognition. *Proc. SPIE 11059, Multimodal Sensing: Technologies and Applications* 11059, pp. 198 – 209.

Knyaz, V. A., 2010. Multi-media projector – single camera photogrammetric system for fast 3d reconstruction. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVIII-5, pp. 343–348.

Knyaz, V. A. and Zheltov, S. Y., 2017. Accuracy evaluation of structure from motion surface 3d reconstruction. *Proc. SPIE 10332, Videometrics, Range Imaging, and Applications XIV* 10332, pp. 103320P.

Knyaz, V. A., Vygolov, O., Kniaz, V. V., Vizilter, Y., Gorbatsevich, V., Luhmann, T. and Conen, N., 2017. Deep learning of convolutional auto-encoder for image matching and 3d object reconstruction in the infrared range. 2017 IEEE International Conference on Computer Vision Workshops pp. 2155 – 2164.

Lagüela, S., Armesto, J., Arias, P. and Herráez, J., 2012. Automation of thermographic 3d modelling through image fusion and image matching techniques. *Automation in Construction* 27, pp. 24 – 31.

Lagüela, S., González-Jorge, H., Armesto, J. and Arias, P., 2011. Calibration and verification of thermographic cameras for geometric measurements. *Infrared Physics & Technology* 54(2), pp. 92 – 99.

Luhmann, T., Ohm, J., Piechel, J. and Roelfs, T., 2010. Geometric calibration of thermographic cameras. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVIII-5, pp. 411–416.

Luhmann, T., Piechel, J. and Roelfs, T., 2013. *Geometric Calibration of Thermographic Cameras*. Springer Netherlands, Dordrecht, pp. 27–42.

Usamentiaga, R., Ibarra-Castanedo, C. and Maldague, X., 2018. Comparison and evaluation of geometric calibration methods for infrared cameras to perform metric measurements on a plane. *Appl. Opt.* 57(18), pp. D1–D10.

Vidas, S., Lakemond, R., Denman, S., Fookes, C., Sridharan, S. and Wark, T., 2012. A mask-based approach for the geometric calibration of thermal-infrared cameras. *IEEE Transactions on Instrumentation and Measurement* 61(6), pp. 1625–1635.

Yang, M.-D., Su, T.-C. and Lin, H.-Y., 2018. Fusion of infrared thermal image and visible image for 3d thermal model reconstruction using smartphone sensors. *Sensors*.