Thermal 3D Mapping of Building Façades

Dorit Borrmann, Jan Elseberg, and Andreas Nüchter

Abstract Never before in history were humans as dependant on energy as we are today. But the natural ressources are limited and a waste of energy has drastic influences on the environment. In their Action Plan for Energy Efficiency [6] the European Commission estimates that the largest and cost-effictive energy savings potential lies in residential ($\approx 27\%$) and commercial ($\approx 30\%$) buildings. To eliminate heat and air conditioning losses in buildings and factories heat and air leaks need to be localized and identified. Imagine the availability of a complete 3D model of every building that architects can use to analyze the heat insulation of buildings and to identify necessary modifications. In these 3D models temperature peaks are not only detectable but also their extent is visible. A robot equiped with a 3D laser scanner, a thermal camera, and a color camera constitutes the basis for our approach. The data from all three sensors and from different locations are joined into one high-precise 3D model that shows the heat distribution. This paper describes the setup of the hardware and the methods applied to create the 3D model, including the automatic co-calibration of the sensors. Challenges unique to the task of thermal mapping of outdoor environments are discussed.

1 Introduction

Recently a lot of work has been done to capture and reconstruct the world around us. Think of a technology that captures not only the precise 3D repre-

Dorit Borrmann, Jan Elseberg, and Andreas Nüchter
Jacobs University Bremen gGmbH, Campus Ring 1, 28759 Bremen,
e-mail: d.borrmann|j.elseberg|a.nuechter@jacobs-university.de
This work was supported by SEE-ERA.NET project ThermalMapper under the project number ERA 14/01 [14]. We gratefully acknowledge the work of Diana Babiac who collected the photos and performed the image based modeling.

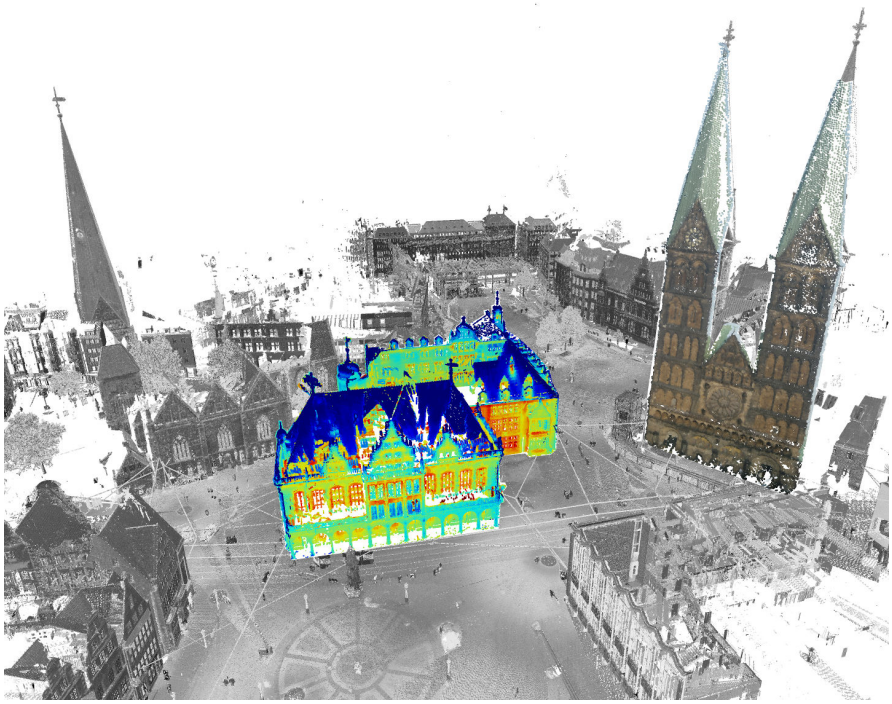

Fig. 1 Laser scan with reflectance, color and thermal information.

sentation of buidlings but at the same time thermal information. These models enable architects and construction engineers to inspect existing buildings, to run simulations of heat and air flow, and to use the gained information to increase the energy efficiency of these buildings. Thermal imaging is state of the art in recording energy related issues, while terrestrial laser scanning has been used for years to create 3D models. The combination of these two yield a 3D model that contains precise temperature information including the dimensions of heat and air leaks. In [17] an evaluation on the impact of 3D models onto the user's ab ability to interpret thermal information concludes that a good representation of a 3D model can help the user, e.g., to locate heat sources or to detect thermal bridges. To solve the remaining problem of identifying heat sources an optical camera is added to our setup. Laser scan data acquired at different positions is combined into complete models of the environment by use of registration algorithms from the geodesy and robotics community. This paper presents our approach towards the creation of complete 3D thermal models of building façades. Further information including an extensive review of related work can be found on the project webpage [14].

2 Background and State of the Art

To assess the energy efficiency of houses thermal cameras are commonly used. These cameras measure temperatures precisely, but return only 2D images of the environment and therefore the loss of energy can only be roughly quantified. Images are projections to 2D. From a sequence of images it is in principle possible to perform a 3D reconstruction. These approaches are called bundle adjustment or structure from motion (SFM), if the focus lies on solving simultaneous localization and mapping (SLAM), i.e., on recovering the 3D structure of the environment and the sensors poses (position and orientation). Since reliable solutions to image based 3D reconstruction for thermal images have not been presented yet, we use the emerging technology of terrestrial laser scanning. Laser scanning methods are well established in the surveying community and in robotics. Terrestrial 3D laser scanning systems yield precise 3D point clouds. Scanning from different poses enables one to digitize complete indoor environments and to resolve occlusions. Registration algorithms from the geodesy and robotics community are available to automatically align scan views from different poses [4, 16].

In related work Högner and Stilla present a modified van as surveying vehicle for acquiring thermal images in urban environments [8]. However, in the focus are image-based techniques like SFM. Prakash et al. present stereo imaging using thermal cameras, but focus on small scale applications [13]. Iwaszczuk et al. suggest an approach to map terrrestrial and airborne infrared images onto existing building models [9]. The model is textured by extracting polygonal parts from the image and mapping those onto the model using standardized masked correlation.

Only a little work has been done for combining 3D scanners and thermal cameras. Carbelles at al. present a methodology to exhaustively record data related to a World Heritage Monument using terrestrial laser scanning, close range photogrammetry and thermal imagery [5]. They use four different sensors for data acquisition: a reflectorless total station, a terrestrial laser rangescanning sensor, a digital photo camera and a thermal camera and use a total of eight natural control points with the help of the total station to relate the geometry between different sensors. In [12] a first automatic approach for multispectral texture mapping is presented. Their method is based on the extraction of a depth map in the form of an image from the model geometry whose pixels maintain an exact correspondence with the vertices of the 3D model. Afterwards the registration with the chosen texture is done which is based on maximization of mutual information. Alba et al. combine the data from a terrestrial laser scan with images acquired by a bi-camera system, i.e., a system consisting of an optical camera and a thermal camera [1]. The fusion of the two data sets is achieved by using some control points that are measured manually with both optical camera and laser scanner. 3D environment mapping using 3D scanners on mobile robots are subject of research [16, 11].

Up to our knowledge automatic modeling using 3D scanning and thermal imaging has not been done yet.

3 Advanced Mutual Calibration between the 3D Sensor and the Cameras

3.1 Experimental Setup and Data Acquisition

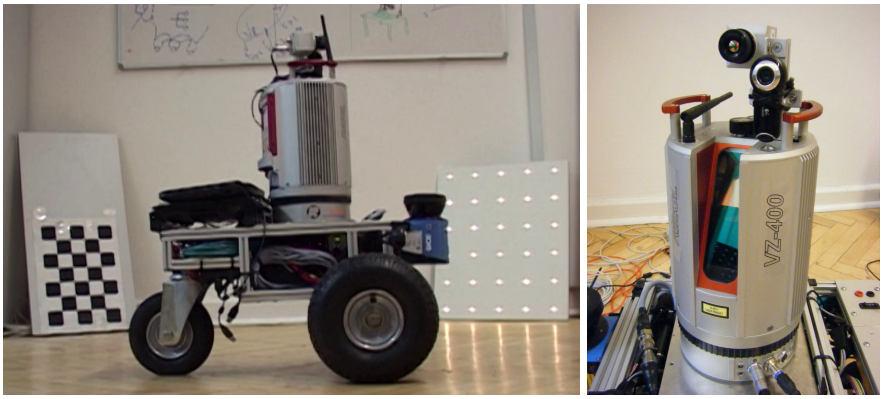

Fig. 2 The robot Irma3D, equiped with a 3D laser scanner, a thermal camera and a webcam. In the background the calibration patterns for the optical camera (left) and the thermal camera (right) are displayed.

The setup for simultaneous acquisition of 3D laser scan data, thermal, and optical images is the robot Irma3D (cf. Fig. 2). Irma3D is built of a Volksbot RT-3 chassis. Its main sensor is a Riegl VZ-400 laser scanner from terrestrial laser scanning. Two cameras are mounted on top of the scanner. The Logitech QuickCam Pro 9000 webcam has a video resolution of 1600×1200. The optris PI160 thermal camera has an image resolution of 160×120 pixels and a thermal resolution of 0.1° C. It acquires images at a frame rate of 120 Hz and with an accuracy of 2° C. The laser scanner acquires data with a field of view of $360^\circ \times 100^\circ$. To achieve the full horizontal field of view the scanner head rotates around the vertical scanner axis when acquiring the data. We take advantage of this feature when acquiring image data. Since the cameras are mounted on top of the scanner, they are also rotated. We acquire 9 images with each camera during one scanning process to cover the full 360°.

3.2 Data processing procedure

After acquiring the 3D data it has to be merged with the color information. This processing consists of five steps that will be explained in this section.

3.2.1 Intrinsic calibration of thermal and optical cameras

Each sensor perceives the world in its own local coordinate system. To join the perceived information we need the specific parameters of these coordinate systems. Each camera has unique parameters that define how a point (X, Y, Z) in world coordinates is projected onto the image plane. Given the focal length (f_x, f_y) of the camera and the camera center (c_x, c_y) image coordinates (x, y) are calculated as:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X/Z \\ Y/Z \\ 1 \end{bmatrix}. \tag{1}$$

Given the radial distortion coefficients k_1, k_2, k_3 and the tangential distortion coefficients p_1, p_2 and $r = \sqrt{x^2 + y^2}$ the corrected image points (x_c, y_c) are calculated as

$$\begin{pmatrix} x_c \\ y_c \end{pmatrix} = \begin{pmatrix} x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 y + p_2(r^2 + 2x^2) \\ y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2y^2) + 2p_2 x \end{pmatrix} \tag{2}$$

To determine the parameters of optical cameras chessboard patterns are commonly used because the corners are reliably detectable in the images. A number of images showing a chessboard pattern with known number and size of squares are recorded. In each image the internal corners of the pattern are detected and the known distance between those in world coordinates allows to formulate equations (1) and (2) as a non-linear least squares problem and solve for the calibration parameters.

[10] have explored the calibration procedure using different types of thermal cameras. Generally an object with a unique pattern having distinct targets is used which eases labeling and increases accuracy of the calibration process. The points are actively or passively heated. In case of passive heating different material causes the the pattern to show up. For low resolution thermal cameras a chessboard pattern is error-prone even after heating it with an infrared lamp. For pixels that cover the edge of the squares the temperature is averaged over the black and white parts thus blurring the edges. Instead a pattern with clearly defined heat sources such as small lightbulbs is suggested as it shows up nicely in thermal images. Fig. 2 shows our pattern in the background. It is composed of 30 tiny 12 Volt lamps, each with a glassbulb diameter of 4 mm. The overall size of the board is 500 mm (width) $\times$

570 mm (height). Identifying the heat sources in the image enables us to perform intrinsic calibration in the same way as for optical cameras. To detect the light bulbs in the thermal image a thresholding procedure is applied to create a binary image showing regions of high temperature. A further thresholding step discards effectively all regions that are too large or too small. If the remaining number of regions is equal to the number of lightbulbs in the pattern the regions are sorted according to the pattern. To calculate the exact center of the features with sub-pixel accuracy, the mean is calculated by weighing all the pixels in the region by their temperature.

3.2.2 Extrinsic calibration – cameras and laser scanner

After calculating the internal parameters of the cameras we need to align the camera images with the scanner coordinate system, i.e., extrinsic calibration. The three rotation and three translation parameters are known as the extrinsic camera parameters and they are unique to every camera view. Once all the points are in the camera coordinate system the projection to the image can be defined up to an factor s using equation [3]:

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3}$$

Suppose there are n images of the calibration pattern and m planar points on the pattern considering the distortions as independent and identically distributed noise than the maximum likelihood estimate of the transformation between the scanner and camera coordinate system is obtained by minimizing

$$\sum_{i=1}^{n} \sum_{j=1}^{m} ||\mathbf{p}_{ij} - \hat{\mathbf{p}}(\mathbf{A}, \mathbf{D}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{P}_j)||^2 \tag{4}$$

where $\mathbf{A}$ is the intrinsic matrix, $\mathbf{R}_i$ the rotation matrix, $\mathbf{t}_i$ the translation vector, and $\mathbf{D}$ the distortion parameters. $\hat{\mathbf{p}}(\mathbf{A}, \mathbf{D}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{P}_j)$ defines the projection of point $\mathbf{P}_j$ in image i, according to equation (3) and (2). This approach assumes that we have a number of points that are identifiable in both the laser scan and the image. For this purpose we attach the calibration pattern onto a board. For the optical camera this is a printed chessboard pattern and for the thermal camera light bulbs arranged in a regular grid pattern. The position of the points of these patterns are known. Algorithm 1 detects the points in a laser scan.

Algorithm 1 Calibration pattern detection in a laser scan

Require: point cloud, specification of calibration pattern
1: discard points outside the area of the expected board
2: find the most prominent plane using RANSAC
3: project a generated plane model into the center of the detected plane
4: use ICP to fit the plane model to the data points
5: **return** position of the lightbulbs according to ICP result

3.2.3 3D to 2D projection and color mapping

During the data acquisition phase laser scans and images are acquired simultaneously. After determining the relations between scanner and cameras in the calibration step this relation is used directly to color the point cloud according to the images.

3.2.4 Projection/occlusion/resolution errors

Due to the different sensor poses and fields of view the sensors perceive different parts of the world. An area that is visible for one sensor might be occluded for the other sensor. When mapping the color information to the point cloud this causes wrong correspondences and therefore faulty colored points. This impact is increased by the low resolution of the thermal camera. With only 120 by 160 pixels per image each pixel corresponds to many 3D points seen by the laser scanner leading to errors at jump edges. Consequently small calibration inaccuracies have a large impact on the results. To solve this problem we take advantage of the fact that if a point belongs to an object there will be more points on that object. We take all points that are are projected onto one pixel and its neighboring pixels. The points are clustered depending on their distance to the scanner. A heuristic based on distance to the 3D scanner and size of the cluster determines effectively which points are considered and enhanced with color information. This removes also some correct color information but the improvement prevails.

3.2.5 Scan Registration

Laser scans acquired at different positions are registered into one common coordinate system using 6D SLAM from *The 3D Toolkit* (3DTK) [2]. The complete model of the environment can be inspected in the viewer from 3DTK enhanced with either reflectance values, thermal data or color from photos (cf. Fig. 1). Switching between the different views enables the user to detect sources of wasted energy and to locate them clearly in the more realistic optical view.

4 Experimental results

The algorithms were tested on three outdoor data sets each recorded with the robot Irma3D.

Data set `Jacobs Day` consists of laser scans, photos, and thermal images recorded at 13 different positions in the early afternoon. Fig. 3 shows part of the point cloud with color information (top) and thermal information (middle). It strikes immediately that the building to the left appears significantly warmer than the building in the back. Even diffuse sunlight on a cloudy day distorts the measurements in a way that a meaningful analysis becomes impossible.

After the conclusions drawn from data set `Jacobs Day` a second data set `Jacobs Night` was recorded at night consisting of laser scans and thermal images from 15 different positions. The result is shown in Fig. 3 (bottom). It becomes clear that the correct temperature values allow for an analysis of the data.

The experiments on the campus of Jacobs University Bremen showed that reliable temperature measurements are not possible during the day due to the sunlight. However, by collecting data at night the advantage of the photo camera information is mostly lost. To overcome this issue we suggest to collect only data with the laser scanner and the thermal camera at night and combine them with the presented framework. During daylight an independent data set of only photos is recorded. Dataset `Bremen` consists of laser scans and thermal images recorded at night at 11 different positions. In addition to that 143 photos were taken independently during the day. Bundler [15] is used to reconstruct the environment from these photos. The Patch-based Multi-view Stereo Software (PMVS) [7] is used to produce a dense reconstruction of the bundler results. The photos cover only a part of the environment mapped with thermal information, namely the market place. As bundler does not preserve scale of the environment the resulting model was semi-automatically scaled and registered to the point cloud from the laser scanner. The resulting point cloud is shown in Fig. 1. The Town Hall in the center is enhanced with thermal information while the cathedral is part of the result from bundler. A video showing the entire scene with thermal information is available under `http://youtu.be/TPoCebERysc`.

5 Conclusions

We presented a method for multi-modal 3D mapping of building façades by use of a 3D laser scanner, a thermal camera, and a photo camera. We identified problems and suggested solutions distinct to outdoor scenarios. In future work we plan to further develop the suggested methods, i.e., include

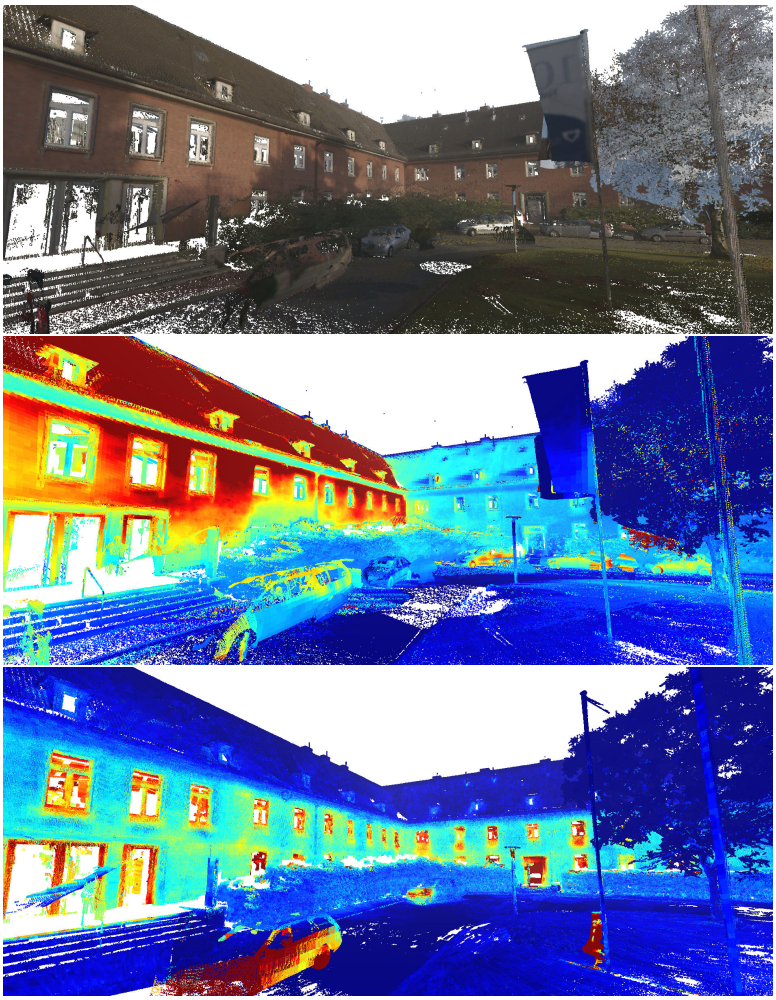

Fig. 3 Colored scans. Top: `Jacobs Day` with color information. Middle: `Jacobs Day` with thermal information. Bottom: `Jacobs Night` with thermal information.

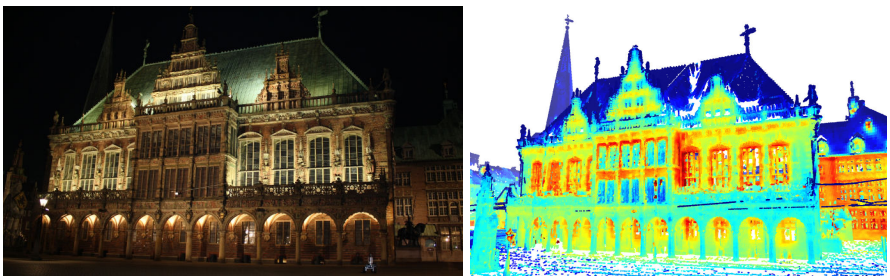

Fig. 4 Irma3D in front of the Town Hall in Bremen while recording data at night (left). Resulting 3D thermal map of `Bremen` (right).

tree detection and develop automatic registration of data recorded at different times to obtain a multi-modal-spatio-temporal model of the environment.

References

1. M. I. Alba, L. Barazzetti, M. Scaioni, E. Rosina, and M. Previtali. Mapping infrared data on terrestrial laser scanning 3D models of buildings. *Remote Sensing*, 3(9):1847–1870, 2011.
2. A. Nüchter et al. 3DTK – The 3D Toolkit. `http://slam6d.sourceforge.net/`, 2011.
3. G. Bradski and A. Kaehler. *Learning OpenCV, Computer Vision with OpenCV library*. O'Reilly Media, 1st edition, 2008.
4. C. Brenner, C. Dold, and N. Ripperda. Coarse orientation of terrestrial laser scans in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(1):4–18, 2008.
5. M. Cabrelles, S. Galcera, S. Navarro, J. L. Lerma, T. Akasheh, and N. Haddad. Integration of 3D laser scanning, photogrammetry and thermography to record architectural monuments. In *Proceedings of the 22nd CIPA Symposium*, Kyoto, Japan, 2009.
6. Commission of the European Communities. Addressing the challenge of energy efficiency through Information and Communication Technologies. Communications to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2008) 241 final, May 2008.
7. Y. Furukawa and J. Ponce. Accurate, dense, and robust multi-view stereopsis. *IEEE PAMI*, 32(8):1362–1376, 2010.
8. L. Högner and U. Stilla. Texture extraction for building models from ir sequences of urban areas. In *Proc. of 2007 Joint Event: URBAN / URS*, 2007.
9. D. Iwaszczuk, L. Hoegner, and U. Stilla. Matching of 3D building models with IR images for texture extraction. In *Joint Urban Remote Sensing Event*, Munich, Germany, 2011.
10. T. Luhmann, J. Piechel, J. Ohm, and T. Roelfs. Geometric calibration of thermographic cameras. In *Intern. Archives of Photogrammetry, Remote Sensing and Spatial Information*, vol. 38 part 5, Newcastle upon Tyne, UK, 2010. Commission V Symposium.
11. A. Nüchter. *3D Robotic Mapping: The Simultaneous Localization and Mapping Problem with Six Degrees of Freedom*. No. 52 in Tracts in Advanced Robotics. Springer, 2009.
12. A. Pelagottia, A. Del Mastio, F. Uccheddu, and F. Remondino. Automated multispectral texture mapping of 3d models. In *EUSIPCO 2009*, Glasgow, Scotland, 2009.
13. S. Prakash, L. Y. Pei, and T. Caelli. 3d mapping of surface temperature using thermal stereo. In *Proc. ICARCV*, 2006.
14. Project Webpage. Project ThermalMapper. `http://www.faculty.jacobs-university.de/anuechter/thermalmapper.html`, 2011.
15. N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring image collections in 3d. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006)*, 2006.
16. H. Surmann, A. Nüchter, and J. Hertzberg. An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments. *Journal Robotics and Autonomous Systems (JRAS)*, 45(3–4):181–198, 2003.
17. J. Wardlaw, M. Gryka, F. Wanner, G. Brostow, and J. Kautz. A new approach to thermal imaging visualisation. EngD Group Project, University College London, 2010.