



**CZECH TECHNICAL UNIVERSITY IN PRAGUE**

---

**Faculty of Civil Engineering  
Department of Geomatics**

## **Photogrammetric processing of non-traditional data**

### **DOCTORAL THESIS**

**Ing. Adam Dlesk**

Doctoral study programme: Geodesy and Cartography  
Branch of study: Geodesy and Cartography

Doctoral thesis tutor: prof. Dr. Ing. Karel Pavelka

**Prague, 2022**

## **DECLARATION**

Ph.D. student's name: Ing. Adam Dlesk

Title of the doctoral thesis: Photogrammetric processing of non-traditional data

I hereby declare that this doctoral thesis is my own work and effort written under the guidance of the tutor prof. Dr. Ing. Karel Pavelka.  
All sources and other materials used have been quoted in the list of references.

:

In Prague on .....

.....  
signature

# Abstract

Non-traditional data are data that require a different approach than common digital images when processing using modern photogrammetric methods. By definition, as non-traditional data can be considered analogue metric images and thermal infrared images. These two case studies were researched in this doctoral thesis. For successful photogrammetric processing of the non-traditional data, innovative pre-processing and post-processing methods were designed. The methods were designed to have the possibility to achieve sufficient geometrical accuracy of the results and, in the case of the thermal infrared images, the methods were designed to have the ability to create 2D and 3D models augmented by thermal infrared information. In the doctoral thesis, the methods were described, analyzed, and tested on chosen objects. The research findings were presented in four articles in the scientific journals. This thesis is prepared as a commented set of the four articles.

**Keywords:** photogrammetry, non-traditional data, analogue metric images, thermal infrared images, structure from motion, multi-view stereo matching, pre-processing, post-processing

# Abstrakt

Netradiční data vyžadují při fotogrammetrickém zpracování moderními metodami speciální postupy, které se liší od postupů používaných při fotogrammetrickém zpracování běžných snímků pořízených digitálními kamerami. Dle definice netradičních dat je možné jako netradiční data označit analogové měřické snímky a nebo termovizní snímky. Těmito dvěma případy se věnuje předložená disertační práce. Pro fotogrammetrické zpracování těchto dat byly navrženy inovativní metody pro pre-processing a post-processing. Metody byly navrženy tak, aby výsledek fotogrammetrického zpracování měl co nejvyšší možnou přesnost a v případě zpracování termovizních snímků bylo možné výsledky využít i pro získání 2D a 3D modelů obohacených o termovizní informaci. V rámci disertační práce byly metody popsány, analyzovány a testovány na vybraných objektech. Výsledky výzkumu jsou publikované ve čtyřech článcích ve vědeckých časopisech. Disertační práce je předložena jako komentovaný soubor zmíněných článků.

**Keywords:** fotogrammetrie, netradiční data, analogové měřické snímky, termovizní snímky, structure from motion, multi-view stereo matching, pre-processing, post-processing

# Acknowledgments

Reaching this point and submitting this doctoral thesis would not be possible without constant and patient, either professional or personal support from many people. On this page, I would like to thank them and express my great appreciation.

I would like to thank my supervisor Karel Pavelka who was also my supervisor of my bachelor's and master's thesis. During the many years, he was always so supportive, helpful, and inspirational. I am grateful for his guidance.

My gratitude goes to Karel Vach and Pavel Holubec who were (and still are) mentoring me in my professional career. I appreciate so much being part of the environment that they have created. I am thankful for all the support that I received from them. During my studies, it was so important. Also, I would like to thank the rest members of the EuroGV family.

I would like to thank my family. It seems to me that without your love and support it would not be possible to reach this point. Thank you Zdeňka Dlesková, Jiří Dlesk, Markéta Englová, Tomáš Engl, and Laura Englová.

A special thank goes to Dorotea Cimburová. We started to know each other just before I was at the beginning of my Ph.D. study. I wish I could give us back all the weekends when I was studying instead of spending them with you.

Paulina Raeva was the only classmate I had. I think she was the best classmate that I could have wished for. Thank you, Paulina, for all your help, motivation, and fun.

During my studies, there were times when I was losing my motivation for further work. I would like to thank my friends who were also studying for Ph.D. It was always so motivating to talk to you. You were an inspiration to me. Thank you Martina Diaz, Žofie Cimburová, and Andres Uueni.

DOI: 10.14311/dis.fsv.2022.004



# List of Articles

The doctoral thesis is prepared as a commented set of four published articles in the scientific journals with an impact factor. All the journals are indexed in Web of Science and Scopus.

1. **Dlesk, A.**; Vach, K.; Pavelka jr., K. Structure from motion processing of analogue images captured by Rollei metric camera, digitized with various scanning resolution. *Acta Polytechnica* 60.4 (2020): 288-302.

<https://doi.org/10.14311/AP.2020.60.0288>

2. **Dlesk, A.**; Uueni, A.; Vach, K.; Pärtna, J. From analogue to digital photogrammetry: Documentation of Padise Abbey in two different time stages. *Appl. Sci.* 2020, 10, 8330.

<https://doi.org/10.3390/app10238330>

3. **Dlesk, A.**; Vach, K.; Pavelka, K. Transformations in the photogrammetric co-processing of thermal infrared images and RGB images. *Sensors* 2021, 21, 5061.

<https://doi.org/10.3390/s21155061>

4. **Dlesk, A.**; Vach, K.; Pavelka, K. Photogrammetric co-processing of thermal infrared images and RGB images. *Sensors* 2022, 22, 1655.

<https://doi.org/10.3390/s22041655>

# List of Figures

3.1	Diagram of the sharpening method. . . . .	17
3.2	Schematic workflow of the sharpening method. . . . .	18
3.3	Diagram of the reprojection method. . . . .	20
3.4	Schematic workflow of the reprojection method. . . . .	21
1	Fragments from the processing of analogue Rollei 6006 metric images (the Great America) . . . . .	140
2	Orthophoto of the north wall of the limestone quarry the Great America	141
3	Orthophoto of the south wall of the limestone quarry the Great America	142
4	Fragmetns from the processing of analogue Rollei 3003 metric images (Zikmund's chapel, Saint Vitus Cathedral in Prague) . . . . .	143
5	Fragments from the processing of analogue UMK metric images (Padise Abbey, Estonia) . . . . .	144
6	Padise Abbey: The overlay of orthophotos . . . . .	145
7	Mosaic of facade created from set of thermal infrared images and thermal profile measurement . . . . .	146
8	Orthophotos of facades created from solely close range thermal infrared images . . . . .	147
9	Orthophoto with RGT band combination and separated T band . . . . .	148
10	Processing of images by sharpening method . . . . .	149
11	Orthophotos created by sharpening method . . . . .	150
12	3D model with augmented texture created by sharpening method . . . . .	151
13	Augmented point cloud created by reprojection method . . . . .	152
14	Augmented points visible on multiple TIR images . . . . .	153
15	Example of statistics on points with multiple observation . . . . .	154

# List of Abbreviations

<b>2D</b>	Two-dimensional
<b>3D</b>	Three-dimensional
<b>BIM</b>	Building information modelling
<b>CCD</b>	Charge-coupled device
<b>CDF</b>	Cumulative distribution function
<b>CDW</b>	Close range digital workstation
<b>CMOS</b>	Complementary metal-oxide-semiconductor
<b>EXIF</b>	Exchangeable image file format
<b>FOV</b>	Field of view
<b>GCP</b>	Ground control point
<b>GIS</b>	Geographic information system
<b>HBIM</b>	Historic building information modelling
<b>IBMR</b>	Image based modelling and rendering
<b>ICP</b>	Iterative closes point
<b>IRT</b>	Infrared thermography
<b>MVS</b>	Multi-view stereo matching
<b>RGB</b>	Red, green, blue
<b>RMSE</b>	Root mean square error
<b>RPAS</b>	Remotely piloted aircraft systems
<b>SfM</b>	Structure from motion
<b>TIR</b>	Thermal infrared
<b>UMK</b>	Universal photogrammetric camera
<b>USAF</b>	U.S. Air Force

# Contents

<b>Abstract</b>	<b>3</b>
<b>Acknowledgments</b>	<b>4</b>
<b>List of Articles</b>	<b>5</b>
<b>List of Figures</b>	<b>6</b>
<b>List of Abbreviations</b>	<b>7</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Definition of non-traditional data . . . . .	2
1.2 Goals of thesis . . . . .	2
<b>2 Photogrammetric processing of analogue metric images</b>	<b>3</b>
2.1 Motivation . . . . .	4
2.2 State-of-the-art . . . . .	4
2.3 Method of the pre-processing . . . . .	5
2.4 Articles . . . . .	7
2.4.1 Article 1: Structure from motion processing of analogue images captured by Rollei metric camera, digitized with various scanning resolution . . . . .	7
2.4.2 Article 2: From analogue to digital photogrammetry: Documenta- tion of Padise Abbey in two different time stages . . . . .	9
<b>3 Photogrammetric processing of thermal infrared images</b>	<b>12</b>
3.1 Motivation . . . . .	13
3.2 State-of-the-art . . . . .	14
3.3 Methods . . . . .	15
3.3.1 Sharpening . . . . .	16
3.3.2 Reprojection . . . . .	18
3.4 The code . . . . .	21
3.5 Articles . . . . .	22
3.5.1 Article 3: Transformations in the photogrammetric co-processing of thermal infrared images and RGB images . . . . .	22
3.5.2 Article 4: Photogrammetric co-processing of thermal infrared im- ages and RGB images . . . . .	26

<i>CONTENTS</i>	9
<b>4 Conclusions</b>	<b>30</b>
4.1 Summary of thesis . . . . .	30
4.2 Fulfilments of goals . . . . .	31
4.3 Further extensibility . . . . .	32
<b>Bibliography</b>	<b>38</b>
<b>Article 1</b>	<b>39</b>
<b>Article 2</b>	<b>55</b>
<b>Article 3</b>	<b>72</b>
<b>Article 4</b>	<b>94</b>
<b>Documentation of the code</b>	<b>116</b>
<b>Image Attachments</b>	<b>139</b>

# Chapter 1

## Introduction

The thesis is prepared as a commented set of four published articles in the scientific journals.

The photogrammetry is a traditional technology used for creating a 2D and 3D object representation through the processing of images. Technology has changed dramatically over the last decades, from analogue to analytical and to digital. Nowadays, the photogrammetric processing is often carried out using a modern method called Structure from Motion (SfM) [1], [2], [3], or sometimes also called Image Based Modelling and Rendering [4]. The SfM method is based on feature detection and feature matching algorithms which automatically find corresponding tie points between the images [5]. Based on the found tie points, the relative orientation of the images is computed, and the photogrammetric model is established [6]. Common inputs for the photogrammetric processing using the SfM method are usually close-range, terrestrial, or low altitude aerial digital images captured by common digital cameras [7], [8]. SfM is a sufficient technology for computing the relative orientation of the images and sparse point cloud (tie points) generation. The subsequent steps are point cloud generation, triangulated model generation or orthophoto generation [9]. As a method for dense point cloud generation, there was developed a multi-view stereo matching (MVS) technology [10]. SfM and MVS (hereinafter SfM-MVS) processing of common digital images is already widely described in scientific literature. Images which have a different character than common digital images may require some special approaches before, during or after SfM-MVS processing. Among the different approaches there could be, for example, a pre-processing of the images, different calibration technique or post-processing of the images. Those images can be termed as “non-traditional” data.

The author of this thesis considers analogue metric images as a non-traditional

data. The processing of those data using modern SfM-MVS method requires a pre-processing method. A pre-processing method was suggested and analysed by the author. The results of the author's research were presented in two scientific journal articles. Generally, the photogrammetric processing of the analogue metric images using SfM-MVS method requires deep knowledge about the fundamentals of the classic photogrammetry. The knowledge gained was essential for the following years of the author's study, where the author of this thesis was focused more on a development of methods concerning photogrammetric processing of thermal infrared images. The thermal infrared images have such a different character compared to the common digital images that the photogrammetric processing using SfM-MVS requires a vastly different approach. Thus, thermal infrared images can be considered as non-traditional data as well. Two innovative methods of processing were designed, developed, implemented and tested. The description and testing of the methods were presented in two scientific journal articles. In this thesis, the topic of the photogrammetric processing of the thermal infrared images takes the main and major part.

## **1.1 Definition of non-traditional data**

The "non-traditional" data are images which require a different approach than the common digital images (captured by common digital camera) when processing using SfM-MVS. Among the different approaches could be considered pre-processing, different calibration techniques, or post-processing techniques. To achieve sufficient accuracy, using these techniques is crucial for the processing. The photogrammetric processing of non-traditional data without the different approaches may not be possible at all or would be possible with only low accuracy results.

## **1.2 Goals of thesis**

Among the goals of the thesis is to determine which data can be considered as non-traditional data. Then, motivation should be presented as to, why is beneficial to process the data using modern photogrammetric solutions. In order to process the non-traditional data, innovative methods shall be designed, developed, and implemented. The methods should be tested on chosen objects and the results should be analysed mainly in terms of geometrical accuracy. The presented methods should be possible to use in praxis and the following research fields.

## Chapter 2

# Photogrammetric processing of analogue metric images

In today's perspective, analogue metric images are possible to term as non-traditional data. The images have different character compared to the common digital images. In this thesis, the term analogue metric images always refer to their negatives. The negatives are stored on a glass-plate or photographic film [11]. The digital image is a digital three-band raster with red, green, and blue color representation. The genesis of the analogue metric images is different. The analogue metric images were captured by precisely manufactured metric cameras equipped by high quality calibrated lenses with stable optics [11]. The metric cameras were intended for the photogrammetry and the main topic in camera design was the accuracy. Common digital images are captured by cameras that are not intended for photogrammetry and the optics are not that stable. Even the capturing of the images for photogrammetric purposes is vastly different. Photogrammetric capturing of the images by metric cameras was a highly sophisticated work. Compared to the capturing for SfM-MVS, capturing scenario was very different when the images were intended to process by a classical terrestrial stereophotogrammetry or later by an intersection photogrammetry. The cost of capturing the analogue metric images was much higher and every image mattered. On the contrary, by digital camera, it is not a problem to capture hundreds and even thousands of digital images just for one project. The processing approach is also different. The SfM-MVS method does not have a background in classical photogrammetry, but in computer vision where the original aim was to object modelling, robotics, navigation, object recognition, etc. [12]. All these facts address that from today's perspective of SfM-MVS processing, it is possible to look at the analogue metric images as at the non-traditional data.



The main and evident difference is that analogue metric images are analogue and to process them using SfM-MVS method, it is necessary to digitize them at first. To process them by SfM-MVS, and to achieve sufficient accuracy, the digitized images must be pre-processed. A method of pre-processing the analogue metric images was suggested. The description of the method, accuracy analysis and practical use were published in two articles in the scientific peer-reviewed journals. The first article deals with the processing of analogue Rollei metric images and focuses on the description of the suggested pre-processing method and the analysis of the accuracy when the method is followed. The second article deals with the practical experiment when the analogue UMK metric images with significant historical value were processed.

## 2.1 Motivation

In the archives and repositories worldwide, there is stored an enormous number of analogue metric images with possibly very important and unique historical value [13]. Together with the general digitization of archives, the archived analogue metric images shall be digitized as well. The digitization of the image can save and protect the image from loss, or destruction. Also, the digitized images can be used for photogrammetric processing by SfM-MVS. By the processing, it is possible to get a new two-dimensional and three-dimensional significantly valuable results which were not possible to create using former methods and software [14]. Those kinds of results can be very important data for GIS and BIM, or HBIM [15]. They can be used for further visualization, temporal analysis [16], augmented reality, or virtual reality [17]. The new results can have a potentially high scientific value for architects, engineers, environmentalists, conservators, other researchers, and for public. On top of that, the archival images can capture an important object which is nowadays destroyed, restored, or changed [18]. Thus, for getting high quality results (in terms of accuracy and visual quality), it was necessary to present a clear and straight forward method of pre-processing.

## 2.2 State-of-the-art

Some articles regarding the photogrammetric processing of the analogue metric images using modern methods were presented in the literature. There is information about photogrammetric processing of analogue aerial images [19] and terrestrial images [20]. Analogue images captured by several cameras were used; Dlesk [20] processed images captured by Rollei metric camera, Poloprutský [13] processed met-

ric images captured by TAN photo-theodolite, Pavelka [18] used metric images captured by UMK and Peppia [21] processed images captured by Hasselbad camera. Maiwald [22] processed analogue non-metric images captured by multiple cameras. Among the main applications, there is the cultural heritage [18], [23], [24], urban research [25], and environmental research [26], [27]. As the prevailing photogrammetric method, SfM-MVS was used, concretely by commercial software Agisoft Metashape (successor of Agisoft Photoscan) [28], [18], [20], [13], [22], and by an open-source software MicMac [29]. All the mentioned papers show the potential of processing of the analogue images captured by variety of metric cameras for different kind of applications. But the methods of pre-processing are not described in detail and the deep analysis of the accuracy of the results is not presented. So, a pre-processing method was suggested also by the author of this thesis and the detail description of the method as well as the analysis of the accuracy (in relation to the pre-processing) was published in an article [30]. Then, the potential of the pre-processing method is shown in a practical example [14]. These two articles belong to the collection of this thesis.

## 2.3 Method of the pre-processing

According to the definition of non-traditional data, the photogrammetric processing of analogue metric images using SfM-MVS method requires a different approach. Hence, the pre-processing method was suggested by the author of this thesis.

At first, it is necessary to gather all the possible information about the origin of the analogue images. The information about images does not have to be preserved, which may complicate the pre-processing. One of the most important pieces of information is the type of the metric camera and lens that were used. Also, the calibration report of the camera with information about internal orientation parameters and with the calibrated coordinates of the camera's fiducial marks shall be gathered. Exploitable information is also a general documentation about the photogrammetric survey. The documentation can answer questions important for the context such as what was the method of photogrammetric processing, which capturing scenario was used, what was the desired accuracy, and what was the purpose or result of the original survey.

Then, the analogue images must be digitized. The digitization can be carried out using conventional or professional scanners. For digitizing, the final scanning resolution should be chosen. The choice of the scanning resolution depends on many factors. First, the quality assessment of the images should be carried out. Over time,

the analogue negatives can be deteriorated, partly destroyed, have cracks, scratches and their radiometric value can be worsened. The scanning resolution should be also estimated by the resolving power of the negative. The resolving power can be read in the documentation of the storing medium that was used. When the desired scanning resolution is estimated, the analogue negatives shall be scanned. It is important to note that most of the conventional scanners declare a much higher optic resolution than they really have [31]. The true optic resolution should be proven by testing targets (e.g., USAF1951). In general, scanning under too high resolution is time consuming and the final data volume can be enormous. The following processing is then slower and more complicated. The highest scanning resolution does not always bring the most benefits. If the colour of the result is important, every image of the dataset must be scanned using the same colour configuration.

For successful photogrammetric processing, the digitized negatives of the analogue metric images must be transformed to get the same coordinate system for each image. For the transformation, the position of fiducial marks shall be used. The calibrated coordinates of the fiducial marks are usually stated in the calibration report of the camera. If the calibrated coordinates are not known, it is possible to use the general position of the fiducial marks (from the knowledge about the metric camera). The digitized images shall be transformed to the fiducial marks using a plane transformation. Affine [13], bilinear [11] or grid-wise [32] transformation can be used. For simplicity, affine transformation is advised. The affine transformation also compensates the shrinking or sheering of the photographic film [11]. After the transformation, the images must be cropped (with respect to the origin of the coordinate system) to have the same dimension.

For the SfM-MVS processing, some initial values shall be stored to the EXIF files of the images. At least information about focal length of the used lens, physical size of the image and resolution in pixels should be written to the EXIF file. As a final step of the pre-processing, the fiducial marks and frames of the images should be masked, otherwise they could be misleading for feature matching algorithms.

Then, the pre-processed images can be processed according to the normal pipeline of SfM-MVS processing in an arbitrary photogrammetric software.

## 2.4 Articles

### 2.4.1 Article 1: Structure from motion processing of analogue images captured by Rollei metric camera, digitized with various scanning resolution

Dlesk, A.; Vach, K.; Pavelka jr., K. "Structure from motion processing of analogue images captured by Rollei metric camera, digitized with various scanning resolution." *Acta Polytechnica* 60.4 (2020): 288-302.

The aim of this article (Appendix: Article 1) was to present the suggested method of pre-processing analogue metric images and to reveal what accuracy can be expected when the suggested method is followed. For the experiment of this article, analogue images captured by Rollei 3003 metric camera were used. The images belong to the archive of EuroGV company. In 1999, the company provided a photogrammetric survey of the vaults in the Saint Vitus Cathedral in Prague. For this paper, seven images, which captured the vault of the Zikmund's Chapel, were chosen. The negatives were stored on photographic roll film with image dimensions of 36 mm x 24 mm.

The negatives were scanned by a Nikon Super Coolscan 8000 professional scanner under four different scanning resolutions – 1000 dpi, 2000 dpi, 3000 dpi and 4000 dpi. The digitized images were pre-processed according to the method presented above. Then, the images with the highest resolution were processed by a classical and original method in CDW software. In the photogrammetric model, one hundred survey points were measured on the vault. Those points were used as the reference check points in the further analysis.

Then, each group (according to the scanning resolution) of the digitized images was processed by a modern SfM-MVS method. Agisoft Photoscan was chosen as a commercial alternative and MicMac was chosen as a free open-source alternative. For each group, in both kinds of software, the absolute orientation was computed and the 100 survey points were measured on the images. The coordinates of the survey points were compared to the reference check points measured in CDW and the differences in each axis were statistically analyzed. The analysis showed a respectable accuracy on the results from both software. However, Agisoft Photoscan gave better results and the set of the differences had, in every case, a normal distribution (tested on Q-Q plots and by Shapiro-Wilk tests) which signified the reliability of the processing. On the con-

trary, MicMac also gave respectable accuracy, but the differences were not generally normally distributed. The testing showed that the best accuracy was achieved with the highest scanning resolution (4000 dpi).

Also, the results of the self-calibration were compared to each other and to the values stated in the calibration report of the used camera. It is important to note, that the Rollei metric camera has an interchangeable and focusable lens. So, the comparison to the values in calibration report was expected to show certain differences. But the comparison of the values of the self-calibration should fit each other. The comparison showed that the parameters computed by the self-calibration in Agisoft Photoscan fit each other very well while the parameters computed in MicMac differed to each other. Again, this showed the better reliability of the commercial alternative Agisoft Photoscan.

Then, using the computed photogrammetric models, the point cloud of each group (according to the scanning resolution) was generated. The point clouds were compared to the check points by Cloud/Cloud method. The differences in the coordinates of each axis were analyzed as was in the previous case. Again, the best accuracy was given by the processing of images with a 4000 dpi scanning resolution. Agisoft Photoscan provided accurate and reliable results (the set of differences had normal distribution). MicMac provided results with a significantly lower accuracy.

Overall, the most accurate results were given by processing the images with the highest scanning resolution. However, the biggest difference in accuracy occurred between 1000 dpi scanning resolution and 2000 dpi scanning resolution. With higher scanning resolution, the difference is not that significant. So, sufficient accuracy is possible to achieve even with lower scanning resolution and dealing with the data is then much faster. Different software tools gave different results. Processing in commercial Agisoft Photoscan gave more accurate and reliable results in every comparison and it can be stated that this software tool is recommended for the processing of digitized analogue metric images. On the other hand, in the case of MicMac, the accuracy and reliability were lower. However, in MicMac, it was possible to compute a relatively accurate absolute orientation. The point cloud generated by MicMac was less accurate. Considering that MicMac is a free and open-source software with a high potential of wide distribution, MicMac seems also as an option for the photogrammetric processing of the analogue metric images.

**Summary of the article:**

- Description of the pre-processing method
- Processing of the data in two types of software - commercial and open-source
- Analysis of the results when the pre-processing method was followed
- Final recommendation to scanning resolution and software

**Research findings:**

- Higher scanning resolution significantly prolongs the time of the pre-processing, dealing with those data is more complicated.
- Highest accuracy was achieved with the highest scanning resolution, but sufficient accuracy is also possible to achieve with lower scanning resolution.
- Agisoft Photoscan gave more accurate and reliable results. All sets of differences had normal distribution.

**2.4.2 Article 2: From analogue to digital photogrammetry: Documentation of Padise Abbey in two different time stages**

Dlesk, A.; Uueni, A.; Vach, K.; Pärtna, J. From analogue to digital photogrammetry: Documentation of Padise Abbey in two different time stages. *Appl. Sci.* 2020, 10, 8330.

The article (Appendix: Article 2) deals with the applied example of photogrammetric processing of the analogue metric images using SfM-MVS method. In 1991, a group from the National Design Institute of Cultural Heritage carried out a photogrammetric survey of the south wall of Padise Abbey in Estonia. For the capturing, UMK 6.5/1318 metric camera was used, and the images were processed by stereophotogrammetric method. Twelve glass-plate negatives (6 stereo-pairs) were, together with the survey documentation, stored in the archive. In 2017, the glass-plate negatives were restored and digitized to process them by SfM-MVS method. In 1991, the wall of Padise Abbey was almost destroyed and in 2017 it was already renovated. By photogrammetric processing using SfM-MVS method and generating new 2D and 3D result, it was possible to get historically valuable models of the former state of the wall before the renovation. Those results could have offered comparison of the former state to the current

renovated state. In 2017, for the comparison, the wall was also photogrammetrically surveyed by a modern equipment.

Before the pre-processing of the digitized images, all information about the photogrammetric survey was utilized. To reconstruct the survey and achieve the important information about the camera, images, control points, and capturing scenario, it was necessary to get a general knowledge about the stereophotogrammetric surveys. After the reconstruction, the images were pre-processed using the suggested method, processed by SfM-MVS, and the point cloud with the orthophoto of the wall were generated. The results were checked in terms of accuracy by all possible information that was read in the original survey documentation. From the new survey in 2017, the orthophoto and point cloud were generated as well. The results of the new survey were transformed to the same coordinate system (as results of former survey) by using former control points which were still detectable on the wall even after the years and the renovation.

Then, there were suggested methods of the spatio-temporal analysis of the results. The results from the former and new surveys were compared to each other and two methods of the comparison were proposed. The first one was the comparison in the orthophoto. The orthophotos were overlaid on each other and the former state was marked by 2D curves. Due to the marking, it is obvious which stone of the wall was preserved and which stone is the outcome of the renovation. This comparison is a result which is easily understandable and readable. The second method was the comparison of the point clouds. The point clouds were pre-processed by filter and subsampling functions and compared to each other by the Cloud/Cloud method. By analyzing the differences, it was possible to determine the preserved stones of the wall not just by visual manner but also by the exact number determination. To check the comparison and to avoid some errors from different point cloud sampling, the point cloud of the new survey (which was used as reference in the comparison) was converted to the triangulated model and the Cloud/Mesh comparison was carried out.

By this practical example, the article showed the high potential of the topic. It is important to point out that in this particular case, it was very important to read all information of the former survey documentation. For understanding the documentation, it was necessary to learn the fundamentals of the photogrammetry and get knowledge about former methods. Without all utilized information it would not be possible to get such a high-quality result. Still, today, when the photogrammetric processing is accord-

ing to the computer vision approach, for some specific applications, it is crucial to retain the fundamentals of the classic photogrammetry.

**Summary of the article:**

- Presentation of the practical example where the suggested pre-processing method was used
- Description of the process of the gathering the data from the former survey
- Generation of the 2D and 3D results with significant historical value
- Suggestion of the spatio-temporal analysis for following research fields

**Research findings:**

- The suggested pre-processing method is applicable in praxis. Using the method, it is possible to generate historically valuable models.
- For successful processing of the archival data, it is necessary to gather all possible information about the former survey.
- For successful processing of the archival data, it is necessary to have knowledge about classic photogrammetry.
- The results of the photogrammetric processing can be an input for further spatio-temporal analysis.
- It is possible to develop semi-automatic and automatic methods of detecting which part of the object is an outcome of the restoration and which part of the object is original. The idea was suggested.



## Chapter 3

# Photogrammetric processing of thermal infrared images

The thermal infrared (TIR) images have a vastly different character compared to the digital images captured by common digital cameras. The difference is given by the different character of the TIR camera. The industrial TIR cameras often have an uncooled micro-bolemeter sensor [33]. Compared to the common digital image where CMOS or CCD sensor is used [34], the physical dimension of the micro-bolometer sensor results in a much lower resolution of the TIR image [35]. The lenses of the TIR cameras are usually made of germanium (Ge) or silicon (Si) [36] and have a much lower field of view (FOV) compared to the lenses of digital cameras [37]. The main difference between the TIR image and common digital image is that in the TIR image, visual features are not detectable. This is the key negative factor for the photogrammetric processing using the SfM-MVS method. In the TIR image, there are detectable only the radiation differences on the scene [38]. Because of that, the TIR images are usually low-contrast, and the lack of features, and the feature matching algorithms of SfM may fail to find a sufficient amount of key and tie points. This results in the significantly lower accuracy of the relative orientation. Generating a point cloud is affected by the missing visual features as well. Sometimes, computing the relative orientation or point cloud generation may not be possible at all [39]. The fact, that the visible features of the scene are not detectable also complicates the transformation of the photogrammetric model. Instead of common photogrammetric targets, it is necessary to use special passive [40] or active targets [41]. To sum it up, a lack of visible features, the problem with photogrammetric targeting, the low resolution of the TIR images, and the low contrast are the main negative factors for SfM-MVS processing. To eliminate these mentioned factors and successfully process the TIR images by SfM-MVS method, a special approach must be used. Because of that, it is possible to define the TIR im-

ages as non-traditional data.

To eliminate all the mentioned negative factors, two methods of the photogrammetric co-processing of TIR images together with digital RGB images were designed by the author of this thesis. Both methods rely on the fixed system of RGB and TIR camera and for every TIR image, a corresponding RGB image must be captured. The first method was called sharpening and the second method was called reprojection. The design, development, description, implementation, and testing of the methods is the main, major, and the most innovative part of this thesis.

### 3.1 Motivation

The thermography is a science which is possible to use in all kinds of different industries [42], [43], and [44]. Among those industries, it is possible to mention civil engineering, energetics, agriculture, medical science, biology, and security. On a single thermal infrared image, it is possible to detect a thermal anomaly on the object. But, just from the thermal infrared image, it is not possible to determine the extent, dimension, area, or the exact position of the anomaly in a reference system. The problem arises when the examined object is large (e.g. library, faculty, or hospital). As was mentioned above, a thermal infrared image has usually a low resolution and the lens of the TIR camera has usually a narrow FOV. Considering the resolution, to provide a sufficient sample distance of the TIR image, it is necessary to capture the object at a relatively close distance. Capturing at close distance and the narrow FOV cause that it would be necessary to capture tens or even hundreds of thermal infrared images. Detecting the anomalies in such an amount of TIR images is difficult and when the anomaly is detected, it may not be possible to recognize even the relative location of the anomaly. This practical example demonstrates why the photogrammetric processing of TIR images is needed.

By photogrammetric processing of the TIR images, it is possible to generate 2D and 3D models (point cloud, 3D model with texture, and orthophoto). The models can be augmented by the thermal infrared information derived from the original TIR images. The thermal infrared information can be stored in the point of the point cloud or pixel of the texture or orthophoto. On those augmented results, it is easier to find the anomalies. It is possible to exactly determine their extent, position, area, or dimension and better comprehend the context of the situation. These results can be also data for advanced spatial analysis, e.g., automatic leak detection, thermal distribution in time,

etc.

## 3.2 State-of-the-art

As was shown by [45] and [46], under certain circumstances, photogrammetric processing of solely TIR images using SfM-MVS method is possible but the presented results contain imperfections visible even at the first sight. To have a universal processing method for all kinds of objects and to eliminate the mentioned negative factors, two methods of photogrammetric co-processing of TIR and RGB images were designed by the author of this thesis. The idea of the photogrammetric co-processing occurs several times in the scientific literature but the published methods are not completely universal for all kinds of objects and also do not eliminate all the mentioned negative factors. The presented methods of photogrammetric co-processing of TIR and RGB images is possible to divide into two groups.

One of the approaches of the co-processing does not require a fixed system of the TIR and RGB camera. An advantage of this approach is a larger flexibility. It gives an opportunity to use a variety of cameras and a variety of camera carriers (RPAS, Terrestrial, etc.). The capturing of the TIR and RGB images can even be carried out in different times. On the other hand, during the processing, the relative orientation of the TIR images must be computed. So, the approach does not eliminate the main negative factor and feature matching algorithms of SfM can fail. Methods with this approach were presented e.g., by [47], [48], [49], and [50].

The second approach requires a fixed system of TIR and RGB cameras, and for every TIR image, the corresponding RGB image must be captured. The main advantage of this approach is that the processing does not require the relative orientation of the TIR images and this approach is universal for every object. On the other hand, there are some difficulties. This approach requires a fixed system of the cameras. Nowadays it may not seem like a significant problem, because modern handheld industrial TIR cameras are also equipped with an RGB sensor and together it naturally creates the fixed system. Capturing the corresponding images is also a complication. During the capturing, the corresponding images must be captured with the same relative pose between each other. The synchronization between the triggers of the cameras in the fixed system may not be absolutely precise and it can cause an error especially on the moving carrier such as RPAS. Another disadvantage of this approach is that a transformation process between the corresponding images is required.

Two methods designed by the author of this thesis also belong to the category where the fixed system of TIR and RGB camera is required. The first method was called sharpening. A similar approach to the sharpening method was published by Weber [51]. His method uses the plane transformation and the method is not applicable for close-range purposes (the method was design for capturing distances more than 200 m). Which means that the method is not universal. The sharpening method uses the ray recalculation function instead which makes the method innovative compared to the method presented by Weber. The second method designed by the author of this thesis was called reprojection. In 2019, the idea of this method was already published by the author of this thesis [46]. Also in 2019, Javadnejad [52] published a similar method. Compared to the Javadnejad's method, the reprojection method contains the test of visibility which is the integral and innovative part. Without the test of visibility, it is not possible to provide the correct thermal infrared augmentation of the spatially oriented results.

### 3.3 Methods

Two methods of the photogrammetric co-processing of thermal infrared and RGB images were designed by the author. Both methods require the fixed system of TIR and RGB camera and for every TIR image a corresponding RGB image must be captured. The methods assume that the offset between the optical centers of the cameras is small and the perspective changes are minimal. In both methods, there is a spatial transformation process. The transformation process requires geometric calibration of the cameras, determination of the translation vector between the optical centers and a rotation matrix between the camera coordinate systems. Determining the required parameters was a broad topic and the description of the problem, the results, and the testing were presented in the article in the scientific journal [39]. Overall, the design of the methods was carried out mainly to eliminate these negative factors:

- Lack of features in TIR images and the impossibility of calculating the relative orientation, or the impossibility of point cloud generation by the SfM-MVS method
- Low resolution of TIR images
- Problems with non-visible photogrammetric targets

During the design, the emphasis was placed on the universality of the methods and usage for close-range purposes.

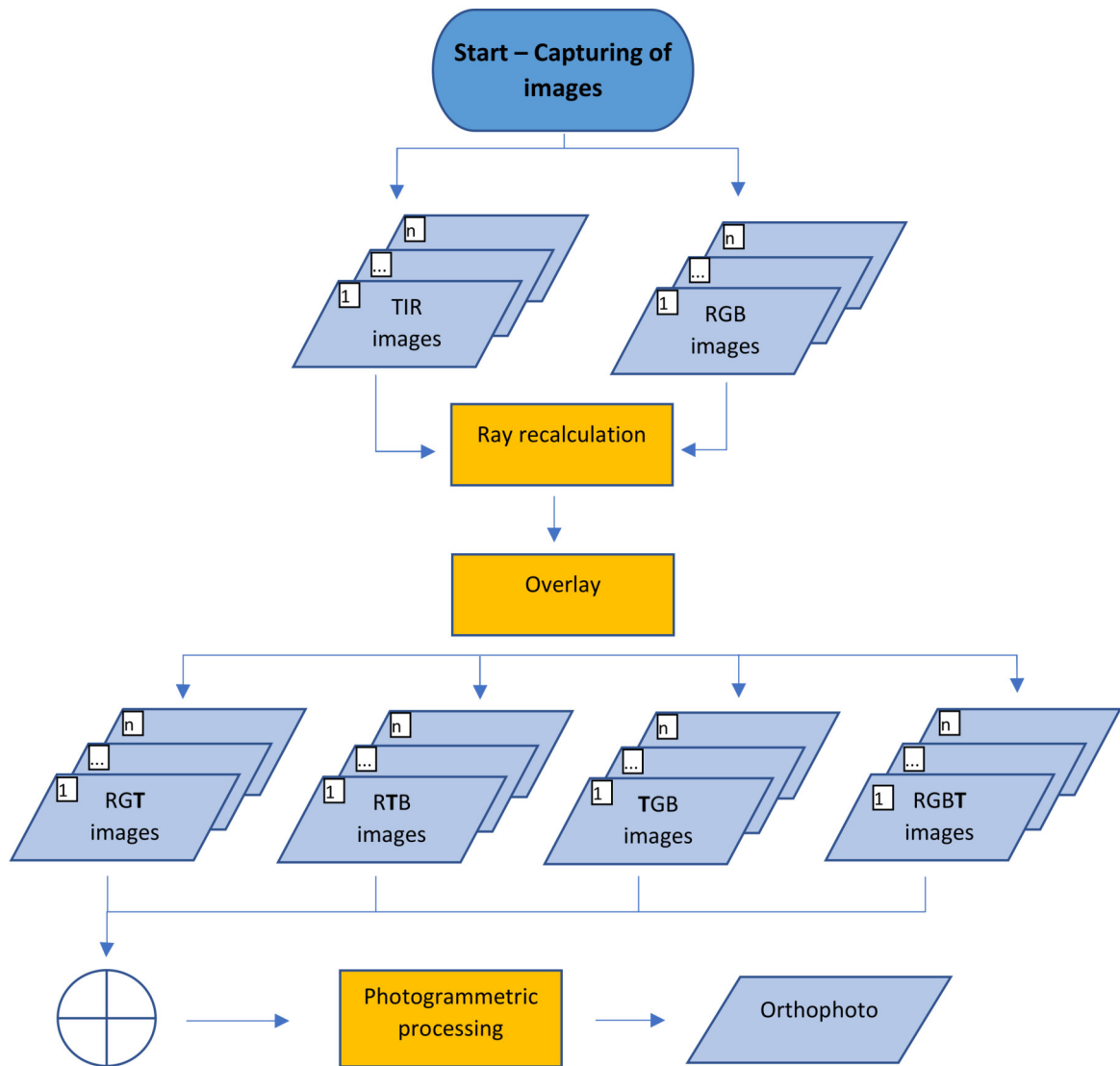
### 3.3.1 Sharpening

The first method was called sharpening, see figure 3.1 and figure 3.2. The sharpening method pre-processes the TIR images and corresponding RGB images and creates the sharpened images. The sharpened images are then processed by the SfM-MVS method according to the standard pipeline. To create the sharpened images, the ray recalculation process was designed. The ray recalculation process is the main innovative part of the method.

For the ray recalculation process, the scaled depth map of each RGB image must be created. Then, the blank raster with the same dimension as RGB image is created. Every pixel of RGB image is taken and knowing the parameters of the interior orientation and having the depth value from the depth map, the coordinates of the point on the scene in RGB camera coordinate system are calculated. Then, knowing the relative rotation and translation between the cameras, the point is transformed into the TIR camera coordinate system. Knowing the parameters of interior orientation the point is projected to TIR image. The value from the pixel of the TIR image is taken and stored to the corresponding pixel in the blank raster. This process is repeated for every pixel of the RGB image. The filled blank raster creates the remapped TIR image. According to the minimum and maximum TIR value of the all TIR images in the dataset, the pixel values of the remapped TIR image must be stretched to 8-bit, or 16-bit values by the histogram stretching function. Having the stretched remapped single-band TIR image, it is possible to swap one of the bands of the RGB image for the stretched remapped single-band TIR image and create RGT, RTB, or TGB image with e.g. 8-bit values of a band. Another option is to add the stretched remapped TIR image to RGB image and create RGBT image with e.g. 16-bit values of a band. The ray recalculation process is repeated for every image in the dataset.

Having the dataset of RGT, RTB, TGB, or RGBT images, it is possible to process the images using SfM-MVS method and create an augmented textured 3D model and augmented orthophoto. The thermal infrared information is stored in the color of every pixel of the texture, or orthophoto. In the pixel, there is stored thermal infrared information stretched to 8-bit or 16-bit values. To get the original value, the inverse histogram stretching must be applied. Sharpening method is not suited for the augmented point cloud generation, because during the generation, the color of the points is usually weighted [53]. This issue may also happen with the texture and the orthophoto but in this case the color weighting and averaging can be usually prohibited in the software.

The sharpening method was described in the more detail in the published article [54]. In the article, the question what combination of the bands (RGT, RTB, TGB or RGBT) is the most sufficient was discussed. The experiment of creating augmented texture and augmented orthophoto has been presented on the test object.



**Figure 3.1:** Diagram of the sharpening method.

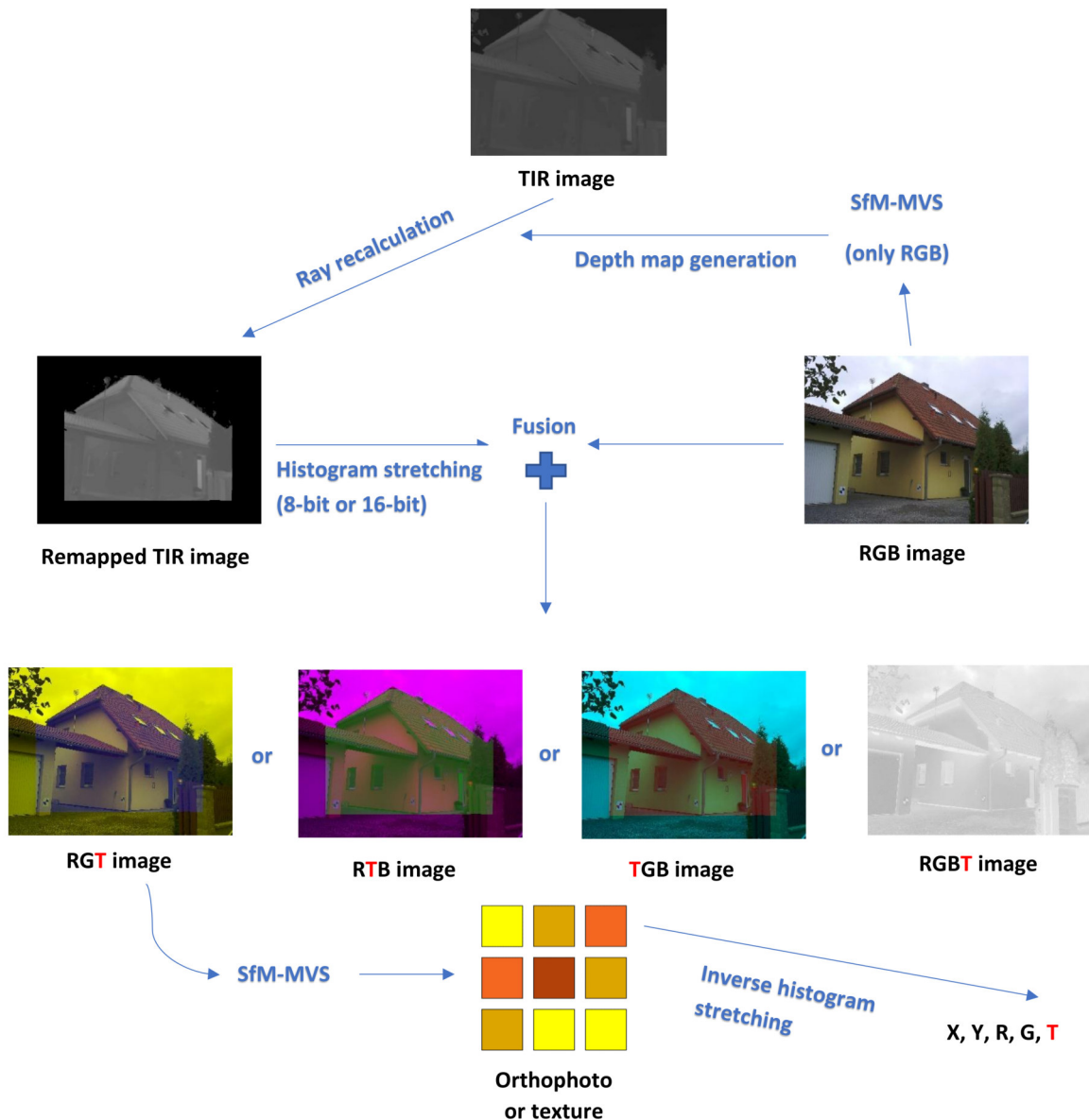


Figure 3.2: Schematic workflow of the sharpening method.

### 3.3.2 Reprojection

The reprojection method (see figure 3.3 and 3.4) was designed primarily to create a point cloud augmented by unweighted thermal infrared information. The reprojection method requires a point cloud generated by photogrammetric processing of RGB images or by laser scanning. The point cloud must be scaled and referenced in the chosen coordinate system. Together with the point cloud, parameters of the exterior orientation of the RGB images in the reference coordinate system, and scaled depth and normal maps of each RGB image are required. The depth and normal maps can be generated by SfM-MVS processing of only RGB images.

A point of the point cloud is taken. The first RGB image of the dataset is taken, and knowing the parameters of exterior and interior orientation of the RGB image, the point is transformed to RGB camera coordinate system and then reprojected to the RGB image. If the reprojected point is within the RGB image dimension, it must be examined, if the point really lies in the scene in the image. The examination is carried out using the test of visibility. The test of visibility is carried out using the corresponding depth and normal map of the RGB image. From the corresponding pixel in the depth map, the depth value is taken and compared to the depth of the point in RGB camera coordinate system. If the difference is under a certain tolerance, the point is potentially visible and then the point is tested on the normal map. The value from the corresponding pixel of normal map is taken and it is compared by cosine similarity to the normal vector stored at the point of the point cloud. If the calculated angle is under a certain tolerance, the point can be still considered as visible. Knowing the relative rotation and the relative translation between the cameras, the point is transformed into the TIR camera coordinate system and then projected to the TIR image. If the point lies within the TIR image dimension, the point is truly visible. Then, the thermal infrared value from the TIR image pixel is taken and stored to the point of the point cloud. The next RGB image of the dataset is taken, and the process is repeated. There may be several thermal infrared values assigned to a single point of the point cloud (the point may be visible on multiple images). This gives an opportunity to statistically check the normality of the set of the values, detect the outliers and eliminate wrongly assigned thermal infrared values. The resulted thermal infrared value can be then determined as a mean. The process is repeated for every point of the point cloud and then the augmented point cloud is created.

The main part of the reprojection method is the test of visibility. In the test of visibility, certain depth and angle tolerances must be set. The experiment which have shown what way the tolerances should be set was carried out and published in the article [39]. Unclarity of the setting the tolerances may be one of the disadvantages of the method. Another disadvantage is that so far, the processing time of the reprojection method is too long. For lowering the time, not all the RGB images of the dataset may be used. It is possible to make a selection of the RGB images and linearly decrease the processing time. In this case, it is important to select the right images to be sure that every point of the point cloud is visible at least in one RGB image. On the other hand, the main advantage of the method is that even the point cloud generated by the laser scanning can be used. It can increase the accuracy of the result. Also, the fact that the resulted thermal infrared value at the point can be given by the statistical calculation can be



included among its important advantages.

The detailed description of the method, the experiment for the test of visibility, and the experiment of the reprojection method were published in [39].

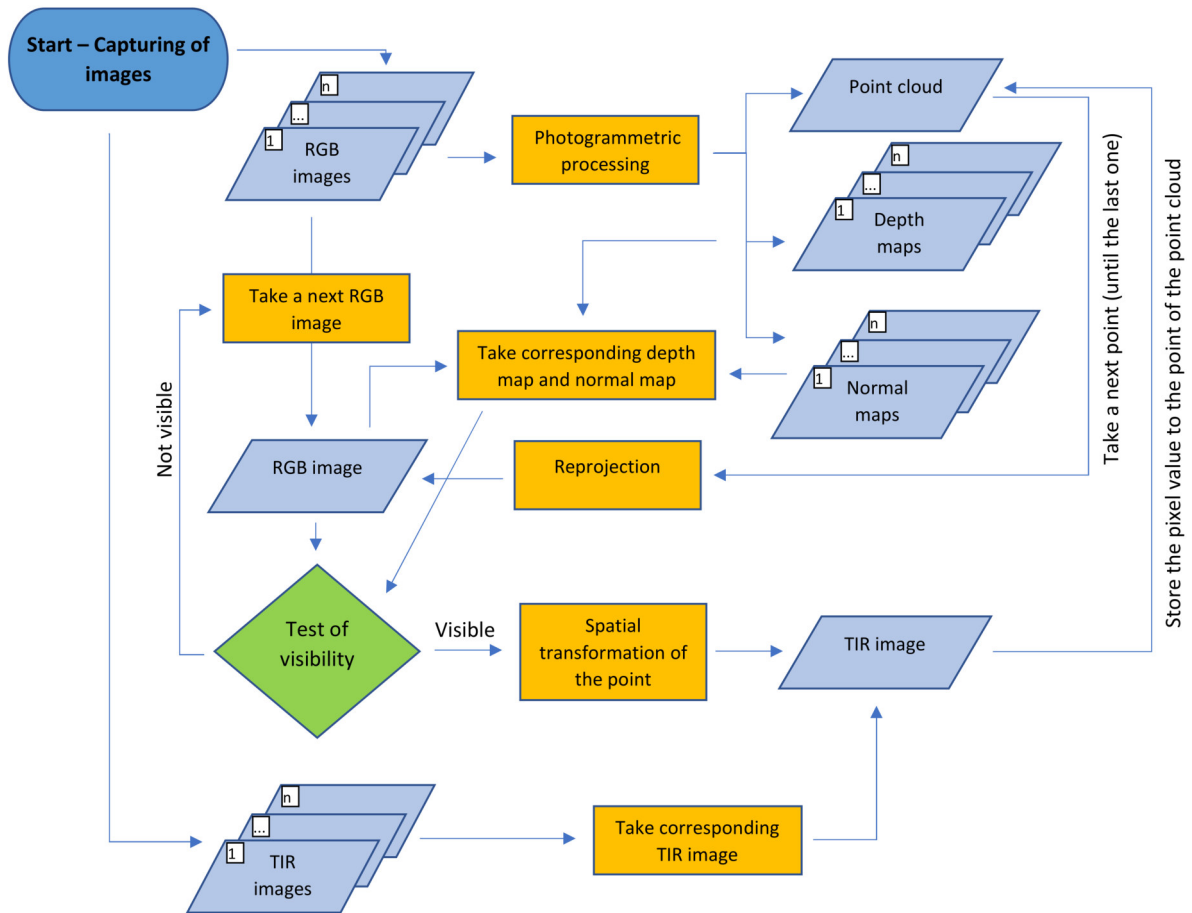
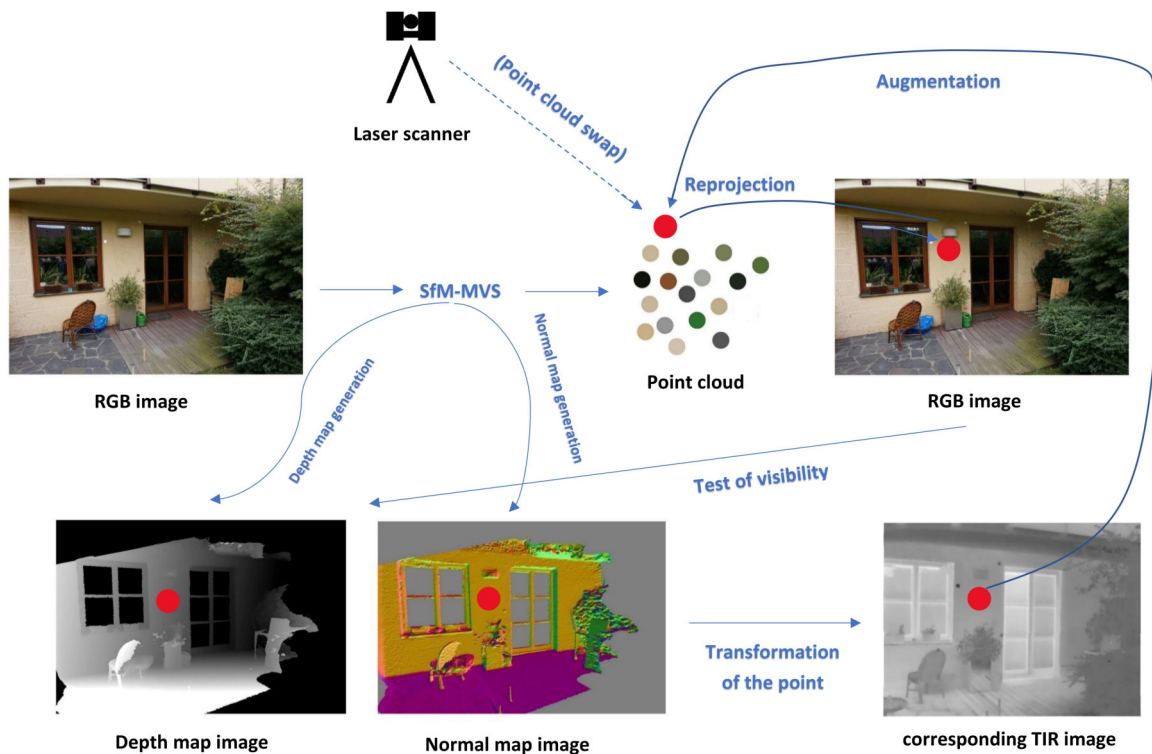


Figure 3.3: Diagram of the reprojection method.



**Figure 3.4:** Schematic workflow of the reprojection method.

## 3.4 The code

The implementation of both methods was written in Python 3.7 language. The code was solely written by the author of this thesis and beside the articles, it is one of the concrete results of this thesis. Now, the code is published on the websites of the Laboratory of Photogrammetry at the Department of Geomatics <http://lfgm.fsv.cvut.cz/termo.html>. The code is released under the open-source BSD-3-clause license.

In total, 5 classes with their methods and 38 functions were developed. Several of the functions deal with the basics of photogrammetry, transformations between coordinate systems, the image processing, or format converting.

The code of the sharpening method is divided into several functions according to the steps which are presented above. The ray recalculation is carried out by `rayRecalculation()` function. The fusion of the images is carried out by `fusionRGT()`, `fusionRTB()`, `fusionTGB()`, and `fusionGeoTiFF()` (RGBT) functions. Data which are generated by SfM-MVS are post-processed by `orthophotoRecalculation8bit()` and `orthophotoRecalculation16bit()` functions (works for both orthophoto and texture of the 3D model).

The designed reprojection method is represented by a single and complex `reprojection()` function.

A detailed description of the classes and functions is presented in the documentation attached to this thesis (Appendix: Documentation of the code). In the documentation, the description, inputs, and returns of the functions are presented. The classes and functions have been separated into four different files according to their theme:

- `classes.py`
- `matrices.py`
- `transformations.py`
- `photogrammetry.py`

For the development, some of the external libraries were used. Standard package Python libraries [55] were used for mathematical operations and parsing the `.csv` files. NumPy [56] library was used for mathematical operations with vectors and matrices, Scipy library [57] for statistical operations, OpenCV [58] library was used for processing the images, and GDAL [59] library was used for reading and processing the rasters in GeoTIFF format.

## 3.5 Articles

### 3.5.1 Article 3: Transformations in the photogrammetric co-processing of thermal infrared images and RGB images

Dlesk, A.; Vach, K.; Pavelka, K. Transformations in the Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images. *Sensors* 2021, 21, 5061.

The designed sharpening and reprojection methods require a transformation process between both corresponding images of the fixed system. During the design of the methods, the question which transformation is suitable has been raised. The idea was to find a transformation which would be universal for all pairs of corresponding TIR and RGB images in the dataset. So, the process could be automatized. The research in this topic became broad and the experiments regarding the transformations

were complicated. Because of that, at first, this topic was published in the article (Appendix: Article 3).

In the beginning of the paper, the motivation was declared. Then the review of the literature was carried out. The reviewed papers were divided into two groups according to their approach. The first group does not require the fixed system of the TIR and RGB camera while the second group does. This division is significant for the transformation processes because the first approach rather uses iterative closes point (ICP) algorithm (or similar) to register data generated from TIR images and RGB images. The second approach uses the 2D plane transformation, or the 3D spatial transformation to transform data between the images. The methods suggested by the author of this thesis belong to the second group. So, in the following experiment, the plane and the spatial transformations were tested.

At first, because of the easier implementation, the 2D plane transformation was used. To have the ability to neglect the perspective changes between the cameras and use the plane transformation, images must be undistorted, and the optical centers of the cameras should be close to each other with a more or less parallel optical axis. The experiment was carried out. For the experiment, FLIR E95 handheld camera with the fixed system of TIR and RGB cameras (the offset between the optical centers of the cameras was roughly around 3 cm) was used. The passive aluminum-paper targets that are identifiable in the TIR images were placed on the test object. Then the object was captured by pairs of corresponding TIR and RGB images from different distances (2 m – 11 m). The optical axis was more-less perpendicular to the object. Then the transformation key was computed in the captured images at different distances and the test, if the single transformation key can be universally used for all the rest of the pairs, was carried out. As was shown in the experiment, it is not possible to find the universal transformation key for all distances and when this type of transformation is used, the RMSE of the residuals can be even around 7 pixels in TIR image. In the short distances there was a strong influence of the offset between the cameras. Overall, the experiment showed that even though the fixed system of cameras with a relatively short offset was used, it was not possible to find the universal transformation key. The plane transformation is not suitable for close-range purposes. In close-range capturing, the distances and the scene vary. Even in one image there can be short and long distances at the same time. Nevertheless, the plane transformation may be used for middle to far and further distances from the object where the distance from the object is more or less constant, e.g., capturing of the ground in non-urban areas from RPAS.

Then for transforming the data between corresponding TIR and RGB images, the 3D spatial transformation was tested. The disadvantage of this transformation is that several parameters of the cameras must be determined with a certain precision and sufficient accuracy. More concretely, the needed parameters are the parameters of the interior orientation of both cameras, three parameters of the relative translation between the optical centers, and three rotation angles of the relative rotation between the coordinate system of the cameras. To test the determination of the parameters, three calibration fields were designed and created. To use the ellipse operator for precise point measurement, the planar stainless calibration field with circular holes was created by a manufacturer. Then two spatial calibration fields with aluminum-paper targets were created (2-level calibration field and 3-level calibration field) by the author. Using those calibration fields, the cameras were calibrated, and the results were compared to each other.

Then, the calibration fields and the calculated corresponding parameters of interior orientation were used for the estimation of the relative pose between the cameras. To achieve a sufficiently accurate relative pose estimation, it was necessary to achieve a very good accuracy and precision in the determination of the exterior orientation of each image of the image pair. The image pairs were captured, the control points were marked on the images, the bundle adjustment was calculated, and the calculated model was scaled. Exterior orientations of the images were computed. Then, for every image pair, the exterior orientation of the TIR image was transformed to the camera coordinate system of the corresponding RGB image and the new exterior orientation parameters of the transformed TIR image give the relative translation and the relative orientation between the cameras. Because multiple image pairs were captured, there were multiple values of the relative pose estimation. The mean of the values was calculated, and the final relative translation and relative rotation between the TIR and RGB camera were determined. The best results were determined with the three-level calibration field. On the three-level calibration field, the standard deviation of the mean of the translation parameters was under 1 mm and the standard deviation of the mean of the rotation parameters was around  $0.03^\circ$ .

Then, the check of the calculated relative poses was carried out on the test object with the aluminum paper targets. Again, the image pairs were captured in different distances (2 m – 11 m). The spatial transformation with the computed relative poses was used to transform the check points between the corresponding TIR and RGB images

and the residuals were checked. The lowest RMSE gave the relative pose calculated on the three-level calibration field. In this case, for all distances, the RMSE in each axis of the residuals was around 1 pixel (maximum 1.2 pixels) in the TIR image. It was shown that spatial transformation is sufficient for the purposes of the designed methods. Also, the experiment showed that the accuracy of the relative rotation has the huge impact on the residuals. The impact is much stronger than the impact of the accuracy of the relative translation parameters.

To conclude, only 3D spatial transformation is sufficient for the sharpening and re-projection method. 2D plane transformation is not universal. The experiment showed that it is possible to achieve enough accurate relative pose for the spatial transformation. The images should be captured in stable conditions and should have the shortest possible sample distance. The targets should be equally distributed in all parts of the images. The coordinates of the targets should be determined with a high accuracy, and the targets should be spatially distributed in space (in all axes) as much as possible. If those conditions are fulfilled, it is possible to achieve a sufficient accuracy of the spatial transformation.

#### **Summary of the article:**

- Brief description of the designed methods
- Review of the related articles
- Presentation of the theoretical background of the transformation processes
- Testing of the 2D plane transformation
- Design of the calibration fields
- Geometric calibration of the TIR and RGB camera
- Determination of the relative translation and rotation of the fixed system of two cameras
- Testing of the 3D spatial transformation
- Estimation of the accuracy of the spatial transformation

#### **Research findings:**

- 2D plane transformation is not sufficient. The transformation is not universal.

- The best result of the geometric calibration and the relative pose estimation was achieved using the three-level calibration field.
- For sufficient accuracy of the spatial transformation, it is necessary to achieve relative translation and rotation with high accuracy.
- It is much more difficult to estimate a sufficiently accurate relative rotation. The relative rotation has the major influence on the final accuracy of the transformation.
- It was possible to find sufficient relative pose estimation. The testing showed that the expected error can be around 1 pixel in TIR image.
- Only 3D spatial transformation is universal and sufficient for the sharpening and reprojection method.

### **3.5.2 Article 4: Photogrammetric co-processing of thermal infrared images and RGB images**

Dlesk, A.; Vach, K.; Pavelka, K. Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images. *Sensors* 2022, 22, 1655.

This article (Appendix: Article 4) was aimed at the detailed description and testing of the sharpening and reprojection method. Few questions regarding the methods were raised and experiments which tried to answer those questions were carried out.

In the beginning, the published article presents a review of the infrared thermography (IRT) and a review of different kind of IRT applications. Then, for those applications, the motivation of creating 2D and 3D models augmented by the thermal infrared information was pointed out. The theoretical background of the IRT was presented. There was explained how the TIR image is created and why the TIR image has its own specifics. In the article, a recommended method how to correctly capture TIR image was presented. During the capturing, many external factors have an impact on the final TIR image. Those factors were mentioned.

The essential part of the sharpening method is the ray recalculation function which creates a remapped TIR image which can be overlaid with the corresponding RGB image. To create the sharpened image, the single-band remapped TIR image can be swapped for any band of the RGB image, and it is possible to create an RGT, RTB,

or TGB image. Another option is to add the remapped single-band TIR image to the RGB image and create four band RGBT image. There has been raised a question of what band combination (RGT, RTB, TGB, or RGBT) shall be used. The advantage of three band combination and 8-bit values is that the images can be processed in every photogrammetric software (they appear as common RGB images). But, one of the color band is missing and then the results of the photogrammetric processing have false colors. RGBT combination keeps the full-color information. But, not every photogrammetric software supports four band images. When the software supports four band images, most likely enables to process images with 16-bit values. Depending on the thermal sensitivity of the TIR camera and on minimum and maximum TIR values of the dataset, sometimes 8-bit values are not enough and 16-bit values are needed. Then RGBT combination is suitable to have 16-bit values. The experiment showed that by SfM processing of three band images (8-bit), it may possible to achieve a higher amount of tie points than by SfM processing of four band RGBT images (16-bit). Nevertheless, by SfM processing of RGBT images, it is possible to achieve a respectable amount of tie points as well. Because of the mentioned advantages, using RGBT (16-bit) combination is recommended. For three band images, it is not possible to give general advice on which combination (RGT, RTB, TGB) is the best. It depends on the color of the significant features of the object. The features are important for SfM-MVS processing. For example, when the significant features have a blue color, then by using RGT combination, the features would be lost and the SfM processing would result in lower amount of tie points. To determine which image combination is the best for the concrete object, it is possible to carry out an experiment that was presented in the article.

The reprojection method was described in the published article and the main focus was on the test of visibility. The test of visibility is an integral part of the reprojection method and without the test of visibility the results may be wrongly augmented. The test of visibility is carried out using the knowledge of the scene via depth and normal maps. For comparison on the depth and normal maps, the tolerances must be set. It is not possible to state how high the tolerances should be set in general (it depends on the concrete case). However, the experiment in the article gives an instruction how the tolerances can be set. The tolerance of the depth test is possible to set according to the computed RMSE on the check points of the absolute orientation of the photogrammetric model where the point cloud for augmentation was generated. Setting the tolerance for test on normal maps is more difficult. The experiment showed, that the normal vectors of the points of the point cloud can vary. To check the variance



of the normal vectors, it is possible to create a cumulative distribution function (CDF) for a group of points that should theoretically have the same normal vector. CDF can show how high the tolerance can be set and what success rate of the test can be expected. Sometimes, it is needed to have the point cloud augmented as much correct as possible, even when the cost is a significant loss of the points that the test of visibility wrongly considers as non-visible. Then, it is possible to set the tolerances very low. Sometimes, it is important to not lose points of the point cloud, even though some of the augmentations may not be correct. In this case, it is possible to use the benefit of the reprojection method and detect the wrong augmentation by statistical function.

In the published article, there was presented an experiment of using sharpening and reprojection method on the object of a family house. By the sharpening method, an augmented orthophoto and augmented texture of the 3D model was generated. By the reprojection method, a point cloud augmented by the thermal infrared information was created. The experiment showed that the augmented point cloud was negatively affected by the depth maps which were rendered on the imperfect 3D model created by processing of RGB images. In some parts of the façade of the family house, there were no features and then, in those parts, the inaccurate 3D model was generated.

#### **Summary of the article:**

- Presentation of the theoretical background of IRT
- Presentation of the correct TIR image capturing
- Detail description of the designed methods
- Discussion of what band combination of sharpened images is the most sufficient
- Presentation of experiment which shows what band combination is the best for the concrete object
- Discussion about the tolerances for the test of visibility
- Presentation of experiment which shows how the tolerances can be set
- Practical example of sharpening and reprojection method

#### **Research findings:**

- The sharpening and reprojection methods are applicable in praxis and in the following research.

- RGBT combination with 16-bit values is recommended by the author.
- In every project the tolerances of the test of visibility must set according to the concrete conditions.
- The processing time of the reprojection method is too time-consuming. In the following research, the code of the reprojection function should be rearranged in order to decrease the processing time.

# Chapter 4

## Conclusions

The topic of the photogrammetric processing of the non-traditional data was a task which required a certain creativity and innovativeness. It was not possible to treat the non-traditional data with the practice which is rooted in every professional. During the past years of research and development, the problems (mainly caused by the different character of non-traditional data) were occurring constantly. The only way to overcome the problems was to deepen the knowledge about photogrammetry, improvise and experiment.

### 4.1 Summary of thesis

In this thesis, photogrammetric processing of the non-traditional data was presented in two case studies.

The first case study was the photogrammetric processing of analogue metric images. This topic was researched during the first years of the author's study. For the successful photogrammetric processing using SfM-MVS method, a pre-processing method was designed by the author of this thesis. The method was tested on Rollei 3003 metric images. There was provided a deep accuracy analysis of the results given by two different software (commercial and open-source). The analysis showed that if the suggested pre-processing method is followed, it is possible to get results with very respectable accuracy. The pre-processing method is fully applicable in praxis. The practical example was demonstrated on the processing of UMK metric images. Through the processing by SfM-MVS, there were created historically valuable results. Using the generated 2D and 3D models, spatio-temporal analyses were suggested. Nowadays, the photogrammetric processing of the analogue metric images is much more common. That is caused by the fact that in the professional photogrammetric software,

similar pre-processing methods are being partly included. Nevertheless, the benefit of the method suggested by the author is that the pre-processed data are not determined for concrete software. The designed pre-processing method is universal. By researching this topic, the deep knowledge about classic photogrammetry was gained and could have been used for the second case study - photogrammetric processing of thermal infrared images.

The review of the literature showed that photogrammetric processing of TIR images by SfM-MVS is problematic. There was a space for developing new methods that could have improved the photogrammetric processing. Because of that, two methods of photogrammetric co-processing of TIR and RGB images were designed – sharpening and reprojection. According to the information presented in the previous chapters, both methods can be considered as innovative and universal. The design, development, coding, and testing of the methods were carried out solely by the author of this thesis. The code of the implementation is published under open-source license at the website of the Laboratory of Photogrammetry at the Department of Geomatics and can be extended by another researcher. By using the methods by other researchers, the methods will be tested on a variety of objects with different characters and then it would be possible to expect that the methods and the code will be improved. Especially, the processing time of the reprojection method should be reduced in the future. From now, it is possible to state that the methods and the code are fully applicable in praxis and in the following research. The developed methods create a significant contribution to the scientific community and are considered as the main result of the thesis.

## 4.2 Fulfilments of goals

The goals of the thesis were fulfilled. There were found two cases of non-traditional data - analogue metric images and thermal infrared images. For the photogrammetric processing, three innovative methods were designed:

- A pre-processing method for analogue metric images
- Sharpening method for processing the TIR images
- Reprojection method of processing the TIR images

All the methods were designed, developed, and implemented solely by the author of this thesis. The methods were analyzed and tested and it is possible to state that all the methods are applicable in praxis. The research of the methods was presented to

the scientific community in four articles in the scientific journals. The research findings are presented above.

### **4.3 Further extensibility**

The code of the sharpening and reprojection method is published under an open-source license at the websites of the Laboratory of Photogrammetry at the Department of Geomatics. It means that it can be used and extended by any researcher. The methods were originally designed for thermal infrared images but they can be adapted even for different datasets.

# Bibliography

- [1] S. Jiang, C. Jiang, and W. Jiang, “Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools”, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 167, pp. 230–251, 2020.
- [2] J. Iglhaut, C. Cabo, S. Puliti, L. Piermattei, J. O’Connor, and J. Rosette, “Structure from motion photogrammetry in forestry: A review”, *Current Forestry Reports*, vol. 5, no. 3, pp. 155–168, 2019.
- [3] M. W. Smith, J. L. Carrivick, and D. J. Quincey, “Structure from motion photogrammetry in physical geography”, *Progress in Physical Geography*, vol. 40, no. 2, pp. 247–275, 2016.
- [4] K. Pavelka, J. Šedina, and K. Pavelka Jr, “Knud rasmussen glacier status analysis based on historical data and moving detection using RPAS”, *Applied Sciences*, vol. 11, no. 2, p. 754, 2021.
- [5] N. Govender, “Evaluation of feature detection algorithms for structure from motion”, 2009.
- [6] V. Knyaz and S. Zheltov, “Accuracy evaluation of structure from motion surface 3D reconstruction”, in *Videometrics, Range Imaging, and Applications XIV*, SPIE, vol. 10332, 2017, pp. 200–209.
- [7] J. H. Chandler and S. Buckley, “Structure from motion (SfM) photogrammetry vs terrestrial laser scanning”, 2016.
- [8] A. Eltner and G. Sofia, “Structure from motion photogrammetric technique”, in *Developments in Earth Surface Processes*, vol. 23, Elsevier, 2020, pp. 1–24.
- [9] G. Verhoeven, “Getting computer vision airborne: Using structure from motion for accurate orthophoto production”, in *RSPSoc archaeology special interest group meeting spring 2012*, 2012, pp. 4–6.
- [10] M. J. Westoby, J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds, “‘Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications”, *Geomorphology*, vol. 179, pp. 300–314, 2012.

- [11] T. Luhmann, S. Robson, S. Kyle, and I. Harley, *Close range photogrammetry: principles, techniques and applications*. Whittles publishing Dunbeath, 2006, vol. 3.
- [12] R. I. Hartley and J. L. Mundy, "Relationship between photogrammetry and computer vision", *Integrating photogrammetric techniques with scene analysis and machine vision*, vol. 1944, pp. 92–105, 1993.
- [13] Z. Poloprutský, M. Fraštia, and M. Marčiš, "3D digital reconstruction based on archived terrestrial photographs from metric cameras", *Acta Polytechnica*, vol. 59, no. 4, pp. 384–398, 2019.
- [14] A. Dlesk, A. Uueni, K. Vach, and J. Pärtna, "From analogue to digital photogrammetry: Documentation of Padise Abbey in two different time stages", *Applied Sciences*, vol. 10, no. 23, p. 8330, 2020.
- [15] M. G. Bevilacqua, G. Caroti, I. Martínez-Espejo Zaragoza, and A. Piemonte, "Orientation of archive images on 3D digital models of painted vaults: An interesting tool for restorers", *Applied Geomatics*, vol. 10, no. 4, pp. 385–398, 2018.
- [16] C. Sevara, G. Verhoeven, M. Doneus, and E. Draganits, "Surfaces from the visual past: Recovering high-resolution terrain data from historic aerial imagery for multitemporal landscape analysis", *Journal of archaeological method and theory*, vol. 25, no. 2, pp. 611–642, 2018.
- [17] K. Pavelka Jr. and P. Raeva, "Virtual museums—the future of historical monuments documentation and visualization", *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 903–908, 2019.
- [18] K. Pavelka, J. Šedina, P. Raeva, and M. Hůlková, "Modern processing capabilities of analog data from documentation of the Great Omayyad Mosque in Aleppo, Syria, damaged in civil war", *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, pp. 561–565, 2017.
- [19] M. Cogliati, E. P. Tonelli, D. Battaglia, and M. Scaioni, "Extraction of DEMS and orthoimages from archive aerial imagery to support project planning in civil engineering", in *ISPRS Workshop on Geospatial Solutions for Structural Design, Construction and Maintenance in Training Civil Engineers and Architects, Geospace 2017*, Copernicus GmbH, vol. 4, 2017, pp. 9–16.

- [20] A. Dlesk, P. Raeva, and K. Vach, "Possibilities of processing archival photogrammetric images captured by Rollei 6006 metric camera using current method.", *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, no. 2, 2018.
- [21] M. V. Peppas, J. P. Mills, K. D. Fieber, *et al.*, "Archaeological feature detection from archive aerial photography with a SfM-MVS and image enhancement pipeline", *International Archives of Photogrammetry, Remote Sensing and Spatial Information*, 2018.
- [22] F. Maiwald, T. Vietze, D. Schneider, F. Henze, S. Münster, and F. Niebling, "Photogrammetric analysis of historical image repositories for virtual reconstruction in the field of digital humanities", *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, p. 447, 2017.
- [23] E. Housarová, K. Pavelka, and J. Šedina, "Study of Erbil Al-Qala citadel time changes by comparison of historical and contemporary image data", *European Journal of Remote Sensing*, vol. 52, no. sup1, pp. 202–208, 2019.
- [24] P. Grussenmeyer and J. Yasmine, "Photogrammetry for the preparation of archaeological excavation. a 3D restitution according to modern and archive images of beaufort castle landscape (Lebanon)", in *XXth Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS)*, 2004, pp. 809–814.
- [25] E. Nocerino, F. Menna, F. Remondino, *et al.*, "Multi-temporal analysis of landscapes and urban areas", *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 39, B4, 2012.
- [26] L. J. Vargo, B. M. Anderson, H. J. Horgan, A. N. Mackintosh, A. M. Lorrey, and M. Thornton, "Using structure from motion photogrammetry to measure past glacier changes from historic aerial photographs", *Journal of Glaciology*, vol. 63, no. 242, pp. 1105–1118, 2017.
- [27] M. Bakker and S. N. Lane, "Archival photogrammetric analysis of river–floodplain systems using Structure from Motion (SfM) methods", *Earth Surface Processes and Landforms*, vol. 42, no. 8, pp. 1274–1286, 2017.
- [28] Á. R. Miranda and J. M. V. Melón, "Recovering old stereoscopic negatives and producing digital 3D models of former appearances of historic buildings", *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, p. 601, 2017.



- [29] S. Giordano, A. Le Bris, and C. Mallet, "Toward automatic georeferencing of archival aerial photogrammetric surveys", *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 105–112, 2018.
- [30] A. Dlesk, K. Vach, and K. Pavelka Jr., "Structure from motion processing of analogue images captured by Rollei metric camera, digitized with various scanning resolution", *Acta Polytechnica*, vol. 60, no. 4, pp. 288–302, 2020.
- [31] F. S. P. Wagner, *Film scanner test-reports: Complete overview of all the models*. <https://www.filmscanner.info/en/FilmscannerTestberichte.html>, Accessed: 2020-07-15.
- [32] R. Kotowski and B. Weber, "A procedure for on-line correction of systematic errors.", *ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXV-A3, pp. 553–560, 1984.
- [33] A. Kylili, P. A. Fokaides, P. Christou, and S. A. Kalogirou, "Infrared thermography (IRT) applications for building diagnostics: A review", *Applied Energy*, vol. 134, pp. 531–549, 2014.
- [34] M. V. V. Radhakrishna, M. V. Govindh, and P. K. Veni, "A review on image processing sensor", in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1714, 2021, p. 012 055.
- [35] A. Berg, J. Ahlberg, and M. Felsberg, "Channel coded distribution field tracking for thermal infrared imagery", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 9–17.
- [36] G. ElMasry, R. ElGamal, N. Mandour, *et al.*, "Emerging thermal imaging techniques for seed quality evaluation: Principles and applications", *Food Research International*, vol. 131, p. 109 025, 2020.
- [37] *User's manual FLIR Exx series*. <https://www.flir.com/globalassets/imported-assets/document/flir-exx-series-user-manual.pdf>, Accessed: 2021-04-26.
- [38] Optris, *Basic principles of non-contact temperature measurement*. [https://www.optris.cz/tl\\_files/pdf/Downloads/Zubehoer/IR%20Basics.pdf](https://www.optris.cz/tl_files/pdf/Downloads/Zubehoer/IR%20Basics.pdf), Accessed: 2021-11-04.
- [39] A. Dlesk, K. Vach, and K. Pavelka, "Transformations in the photogrammetric co-processing of thermal infrared images and RGB images", *Sensors*, vol. 21, no. 15, p. 5061, 2021.
- [40] G. Grechi, M. Fiorucci, G. M. Marmoni, and S. Martino, "3d thermal monitoring of jointed rock masses through infrared thermography and photogrammetry", *Remote Sensing*, vol. 13, no. 5, p. 957, 2021.

- [41] T. Luhmann, J. Piechel, and T. Roelfs, “Geometric calibration of thermographic cameras”, in *Thermal Infrared Remote Sensing*, Springer, 2013, pp. 27–42.
- [42] R. A. Osornio-Rios, J. A. Antonino-Daviu, and R. de Jesus Romero-Troncoso, “Recent industrial applications of infrared thermography: A review”, *IEEE transactions on industrial informatics*, vol. 15, no. 2, pp. 615–625, 2018.
- [43] B. B. Lahiri, S. Bagavathiappan, T. Jayakumar, and J. Philip, “Medical applications of infrared thermography: A review”, *Infrared Physics & Technology*, vol. 55, no. 4, pp. 221–235, 2012.
- [44] E. Lucchi, “Applications of the infrared thermography in the energy audit of buildings: A review”, *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 3077–3090, 2018.
- [45] M. Scaioni, E. Rosina, A. L’erario, and L. Díaz-Vilariño, “Integration of infrared thermography and photogrammetric surveying of built landscape”, *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, 2017.
- [46] A. Dlesk and K. Vach, “Point cloud generation of a building from close range thermal images”, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 29–33, 2019.
- [47] Z. E. Wakeford, M. Chmielewska, M. J. Hole, J. A. Howell, and D. A. Jerram, “Combining thermal imaging with photogrammetry of an active volcano using UAV: An example from Stromboli, Italy”, *The Photogrammetric Record*, vol. 34, no. 168, pp. 445–466, 2019.
- [48] A. Sledz, J. Unger, and C. Heipke, “Thermal IR imaging: Image quality and orthophoto generation”, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives 42 (2018), Nr. 1*, vol. 42, no. 1, pp. 413–420, 2018.
- [49] E. Maset, A. Fusiello, F. Crosilla, R. Toldo, and D. Zorzetto, “Photogrammetric 3D building reconstruction from thermal images”, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, p. 25, 2017.
- [50] L. Hoegner, S. Tuttas, Y. Xu, K. Eder, and U. Stilla, “Evaluation of methods for coregistration and fusion of RPAS-based 3D point clouds and thermal infrared images”, *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 41, pp. 241–246, 2016.

- [51] I. Weber, A. Jenal, C. Kneer, and J. Bongartz, "PANTIR-a dual camera setup for precise georeferencing and mosaicing of thermal aerial images", *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 3, p. 269, 2015.
- [52] F. Javadnejad, D. T. Gillins, C. E. Parrish, and R. K. Slocum, "A photogrammetric approach to fusing natural colour and thermal infrared UAS imagery in 3D point cloud generation", *International Journal of Remote Sensing*, vol. 41, no. 1, pp. 211–237, 2020.
- [53] *Agisoft Metashape user manual*. [https://www.agisoft.com/pdf/metashape-pro\\_1\\_5\\_en.pdf](https://www.agisoft.com/pdf/metashape-pro_1_5_en.pdf), Accessed: 2021-11-04.
- [54] A. Dlesk, K. Vach, and K. Pavelka, "Photogrammetric co-processing of thermal infrared images and RGB images", *Sensors*, vol. 22, no. 4, p. 1655, 2022.
- [55] Python Core Team, *Python: A dynamic, open source programming language*, Python version 3.7, Python Software Foundation, 2019. [Online]. Available: <https://www.python.org/>.
- [56] C. R. Harris, K. J. Millman, S. J. van der Walt, *et al.*, "Array programming with NumPy", *Nature*, vol. 585, no. 7825, pp. 357–362, 2020, ISSN: 1476-4687. DOI: 10.1038/s41586-020-2649-2.
- [57] P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python", *Nature Methods*, 2020. DOI: <https://doi.org/10.1038/s41592-019-0686-2>.
- [58] G. Bradski, "The OpenCV Library", *Dr. Dobb's Journal of Software Tools*, 2000.
- [59] GDAL/OGR contributors, *GDAL/OGR geospatial data abstraction software library*, Open Source Geospatial Foundation, 2022. DOI: 10.5281/zenodo.5884351. [Online]. Available: <https://gdal.org>.

# Article 1

# STRUCTURE FROM MOTION PROCESSING OF ANALOGUE IMAGES CAPTURED BY ROLLEI METRIC CAMERA, DIGITIZED WITH VARIOUS SCANNING RESOLUTION

ADAM DLESK<sup>a,\*</sup>, KAREL VACH<sup>b</sup>, KAREL PAVELKA<sup>a</sup>

<sup>a</sup> Czech Technical University in Prague, Faculty of Civil Engineering, Department of Geomatics, Thákurova 7/2077, 166 29 Prague 6, Czech Republic

<sup>b</sup> EuroGV spol. s r.o., Opletalova 1284/37, 110 00 Prague 1, Czech Republic

\* corresponding author: [adam.dlesk@fsv.cvut.cz](mailto:adam.dlesk@fsv.cvut.cz)

**ABSTRACT.** SfM processing of archived analogue images gives an opportunity to efficiently create new and valuable 2D and 3D results. The SfM processing of digitized analogue images brings some challenges. How to digitize negatives of photogrammetric images? What scanning resolution is the most beneficial for processing? How to preprocess the digitized images to be able to process them using the SfM method? What accuracy of results is possible to expect? This paper tries to deal with all these questions.

For this paper, 7 negatives of former photogrammetric documentation of a vault were chosen. The negatives were captured by Rollei 3003 metric camera in 1999. Two pieces of software were chosen for the SfM processing. A commercial alternative Agisoft PhotoScan and free open-source alternative MicMac. The results of the SfM processing were compared to the results of an original photogrammetric method, which was used for former processing in 1999.

**KEYWORDS:** Structure from motion, Rollei metric cameras, scanning resolution, relative and absolute orientation, point cloud generation.

## 1. INTRODUCTION

There are probably many archived analogue images, which were originally used for photogrammetric documentation of different kind of objects. The *structure from motion method* (SfM) is designed primarily for processing of images captured by digital cameras. Due to the algorithms of SfM, the time cost of the photogrammetric processing has been significantly reduced and it is possible to create two-dimensional (orthophoto) and three-dimensional (point cloud, textured mesh) results very efficiently. Processing of archival analogue images (e.g. images of cultural heritage objects) using current methods brings new possibilities to create 2D and 3D results with a potential historical value. Also, there is a potential to have an opportunity to examine the documented object in different time stages. The results of the processing could be used for further work and could be new and valuable data for information systems like GIS or BIM.

The processing of analogue images using the SfM method brings some challenges. The information about images could be completely lost or prevailed only partially. In some cases, important information about the used camera and the used lens is missing. The material of the negatives of the analogue images could be deformed by time and the radiometric value of negatives could be deteriorated. There is a question, how to digitize the negatives and how to preprocess them correctly. Because the images were most likely

captured for a different photogrammetric method than SfM, there could be too few images in the dataset captured in various angles with various overlap. All these factors must be considered during the processing.

The main part of the paper is to suggest a method how to digitize, preprocess and how to process the digitized images using the SfM method. The results were compared to the results of an original method and were analysed and evaluated in terms of spatial accuracy. What accuracy is possible to achieve if the presented method is followed? For the SfM processing, *Agisoft PhotoScan* was chosen as a commercial alternative and *MicMac* as a free and open-source alternative.

As a dataset of images for this paper, analogue photogrammetric images captured by Rollei 3003 metric camera were chosen. In 1999, a Czech company EuroGV s.r.o. has carried out a photogrammetric documentation of vaults in Saint Vitus Cathedral at Prague castle in Prague. The company provided negatives of the analogue images, coordinates of measured control points, the full documentation of the project and calibration report for the camera lens. Due to all the provided data, it was possible to fully analyse the results of the SfM processing of analogue photogrammetric images.

A pilot project with processing of analogue images, captured by a Rollei metric camera, has been published [1]. Several projects of SfM processing of digitized analogue images are also presented. Processing of digitized glass plate negatives by UMK

cameras and TAN photo-theodolite was presented by [2]. In literature, *Agisoft PhotoScan* prevails as a photogrammetric software for processing analogue images. *Agisoft PhotoScan* was used also by [3] to create 3D model of historic buildings, by [4] to process analogue images captured by RolleiMetric camera and UMK camera of Omayyad mosque, by [5] to analyse historic changes of a glacier, by [6] to process archival images captured by Hasselblad H5D-60 of chosen archeological area, by [7] to process analogue non-metric images and by [8] to process analogue aerial images to create DEM and orthophoto. The open-source *MicMac* was used for processing analogue images by [9]. Mentioned papers do not analyse how various scanning resolutions affect the accuracy of the results.

### 1.1. ROLLEI METRIC

Before digital cameras, metric cameras were usually used for the photogrammetric documentation. Metric cameras are cameras with stable optics, ordinarily fix-focus lens with minimal distortion and with fiducial marks. Their parameters of interior orientation are calibrated in laboratory and are constant for a long time. The wide known metric cameras are, for example, UMK and TMK by Zeiss manufacturer and P31, P32 by Wild. [10]



FIGURE 1. Camera Rollei 6006 metric.

Rollei metric cameras are determined as a semi-metric or partial-metric cameras because the optical system is not completely stable. The parameters of interior orientation are not constant over the long time due to their interchangeable and focusable lenses. Rollei metric cameras are equipped with *réseau plate*. The *réseau plate* is placed in front of the image plane and the reference points are projected to the image during exposure (Fig. 2). The reference points are used for the image coordinate system definition.

As a sensor of the cameras, ordinary photographic roll film was used. Rollei 6006 metric and its later model Rollei 6008 metric used an image format of



FIGURE 2. Part of an image with projected reference points of *réseau plate*.

60 mm × 60 mm [11, 12]. Rollei 3003 metric used 36 mm × 24 mm format. Cameras were equipped with flash, had a wide variety of interchangeable lenses with good quality optics, were flexible, light and easy to operate. Because all the mentioned facts, Rollei metric cameras became very important in the field of architectural photogrammetry [13] and in many other fields of industry. There are many archived analogue images created in 90's of the last century.

### 1.2. SFM PROCESSING

The *structure from motion* is the process how to calculate relative orientation of the images, which are captured from different positions and different angles. *Structure from motion* processing is based on feature matching algorithm e.g. SIFT, SURF, BRISK, DAISY, HOG or GLOH and algorithm for model fitting RANSAC.

In this paper, one commercial software and one open-source software (as a free alternative) was chosen for the SfM processing of archival analogue images. As a commercial software, *Agisoft PhotoScan Professional* in version 1.4.3 developed by Agisoft LLC was chosen. *Agisoft PhotoScan* has a graphic user interface and is easy-to-use. As a free and open-source alternative, *MicMac* in version v1.0.beta13 was chosen. *MicMac* is developed by the National Institute of Geographic and Forestry Information (IGN) and French National School for Geographic Sciences (ENSG). *MicMac* does not have a graphic user interface and is operated via command line by commands and their parameters. Workflow of the SfM processing and generation of a point cloud consist of these steps:

- (1.) Tie points identification
- (2.) Relative orientation

- (3.) Absolute orientation
- (4.) Point cloud generation

In *Agisoft PhotoScan*, the tie points identification and relative orientation is obtained by a single function *Align Photos*. In *MicMac*, the tie points identification is executed by *Tapioca* command and the relative orientation is computed by command *Tapas*.

To scale the relative orientation, the model needs to be transformed to the control points. Computing the absolute orientation is obtained by function *Update* in *Agisoft PhotoScan* and by command *GCPBascule* in *MicMac*. By command *Campari* and function *Optimize Camera Alignment* the model is adjusted by bundle adjustment, and the final absolute orientation is computed.

In the absolute oriented model, it is possible to measure single discrete survey points on the images. By this, it is possible to compute coordinates of the important points or to compute coordinates of check points. The check points were used for checking the accuracy of the relative and absolute orientation of the model.

If the relative or absolute orientation is already known, it is possible to create point cloud of the object. In *Agisoft PhotoScan*, the point cloud is generated by *Build Dense Cloud* function and in *MicMac* by *C3DC* command. It is possible to export the point cloud into different formats and it can be used, for example, for creating hypsometry or other 2D or 3D models [14–17].

## 2. DATASET

For the documentation of vaults in the Saint Vitus Cathedral in Prague in 1999, EuroGV company used the intersection photogrammetry method. The results of the photogrammetric documentation were hypsographic plans of the vaults. For this paper, the vault of Zikmund Chapel was chosen.

The company used Rollei 3003 metric camera with calibrated 28 mm lens for capturing the images. As a storing medium, photographic roll film Konica Color VX400 Film with dimensions 36 mm × 24 mm convenient for capturing scenes under lower light conditions was used [18].

In total, 7 images of the vault were captured, see Fig. 3, with an approximate image scale  $M = 477$ . To achieve a sufficient accuracy in height with photogrammetric processing, it is crucial to have a proper intersection angle between the rays of observation. Because of the height of the vault, it was not possible to get a better intersection angle than  $30^\circ$ , see Fig. 4. There was a presumption that because of the too acute intersection angle, the accuracy in axis  $Z$  will be worse than accuracies in axis  $X$  and  $Y$ .

Most of the vaults were textureless and it would have not been possible to identify discrete survey points on the images. To have the ability to identify discrete survey points on the images and to reference



FIGURE 3. One of the images used for processing in this paper.

them and compute their coordinates, it was necessary to project a grid on the vault, see Fig. 5. The spacing between points of the grid was approximately 0.2 m on the vault. Although 7 images of the vault were captured, the grid was applied only on 4 images.

In the cathedral, geodetic net was created. Control points on the vaults were measured by polar method from at least two different stations of the net and the resulted position of the control points was calculated as the mean. The geodetic net was connected to the Czech coordinate reference system and height reference system.

The negatives of captured images were printed and digitized. Using manually selected control and tie points, relative orientation of the images was computed. Then, the bundle adjustment of the model was executed. From the adjusted model with the absolute orientation, it was possible to identify discrete survey points on the projected grid and compute their coordinates.

The result of the photogrammetric documentation in 1999 was a hypsometric model created from the survey points measured in software *Close-Range Digital Workstation (CDW)*.

## 3. DIGITIZING AND PREPROCESSING

The negatives of Rollei metric images were digitized on photographic film scanner Nikon Super Coolscan 8000 ED. The scanner is equipped with holder, which makes the negative flat and has the declared highest optic resolution of 4000 dpi. The true resolution testing on USAF1951 target [19] showed that even if the flatbed scanners manufacturers declare a high optical resolution (sometimes even more than 4000 dpi) the effective resolution is much lower (barely more than 2000 dpi). On the contrary, tested scanners of Nikon Super Coolscan series reach an effective resolution of more than 3500 dpi when scanning on the 4000 dpi option.

In the paper, four different scanning resolutions were chosen to compare the results to each scanning



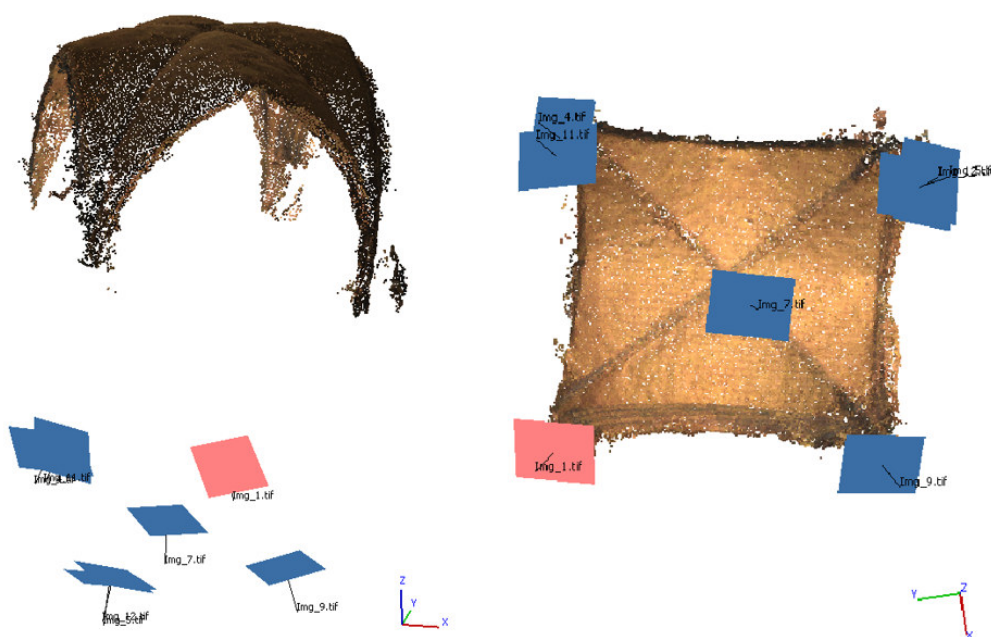


FIGURE 4. Position of images related to the vault.

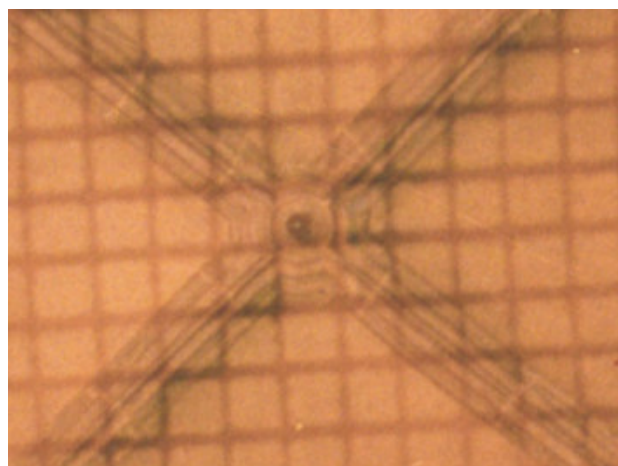


FIGURE 5. Grid projected on the vault in detail.

resolution and to determine which scanning resolution should have been chosen in this particular case.

These different scanning resolutions were chosen after the discussion about the quality of image negatives and their resolving power.

There was rough estimation by the authors of the paper, based on the resolving power and contrast of images, that advantageous scanning resolution should have been set around 2000 dpi. But since it was not possible to estimate it precisely, 4 different types of resolution were chosen:

- 1000 dpi
- 2000 dpi
- 3000 dpi
- 4000 dpi (highest possible)

Scanning resolution	Pixel size of image	Ground Sample Distance
1000 dpi	0.0254 mm	12 mm/pix
2000 dpi	0.0127 mm	6 mm/pix
3000 dpi	0.0085 mm	4 mm/pix
4000 dpi	0.0064 mm	3 mm/pix

TABLE 1. Parameters of scanned images.

The testing of SfM processing in the following chapters shows if there was a benefit (in terms of accuracy) of scanning the negatives on the highest scanning resolution or if a lower scanning resolution would be enough. *Ground sample distance* (GSD) on vault has been calculated for each scanning resolution (Tab. 1).

The scanning on the highest resolution was time consuming and many prints of dust on the digitized images were noticeable. Even though the images were scanned under the same color palette, some of the images have slightly different tone than others. This could also be caused by different aging of each negative.

For the transformation of images with reference points of réseau plate, affine, bilinear or grid-wise transformation is suitable [20]. For simplicity, affine transformation was used in our case.

The scanned images had to be transformed to have the same image coordinate system with the same origin for each image. Reference points of the réseau plate on the images were used for the transformation. Their positions were calibrated in laboratory by manufacturer and are stated in the calibration report. Average residuals of



Group of images	Average residual [mm]
1000 dpi	0.0061
2000 dpi	0.0047
3000 dpi	0.0046
4000 dpi	0.0044
CDW	0.0040

TABLE 2. Average residuals on points used for affine transformation (Image scale  $M = 477$ ).

the affine transformation on reference points are stated in the table Tab. 2.

For photogrammetric processing, the images were cropped by a rectangle. After the cropping, all the images of one group had the same image resolution.

Digitized images are missing their EXIF file and information about the image important for the SfM software. The software needs the information about the focal length of the camera and physical size of the sensor of the camera. In the case of our dataset, the information about the camera was ascertainable from the calibration report of the camera. The information was inserted into the EXIF file of each image.

To avoid potentially disturbing objects in images for photogrammetric processing (like reference points of réseau plate or other descriptive information on the images), the binary bitmaps as masks were created and implemented on each corresponding image before the start of the processing.

#### 4. PROCESSING IN CDW

The original processing of images in 1999 was carried out in software *Close-Range Digital Workstation (CDW)*. For this paper, the original processing was repeated. While original processing was done with images created by digitizing of prints of negatives, nowadays, for purposes of this paper, digitized negatives of images were processed. For processing in CDW, images with 4000 dpi scanning resolution were used. The results of the processing in CDW for the highest scanning resolution images were considered as a reference results in this paper. All the results from the SfM processing were compared to them.

For the processing in CDW, it was necessary to insert initial values of the interior orientation of the camera. The interior orientation parameters were read in the calibration report. Then, the reference points of réseau plate were semi-automatically digitized. Some of the reference points were not possible to detect, because they were in the dark or over exposed background and were not visible. The average accuracy of the transformation was 0.004 mm. Unfortunately, it was not possible to determine the type of the transformation used in the CDW software due to the missing documentation to the software.

The control points were manually marked on each image. Then, the tie points were marked as well. The tie points were chosen for their conspicuousness in images and convenient distribution. The coordinates of control points were imported to the CDW and the rough position and rough rotation was computed using *Single image* function. After the *Single image* computation, the images were computed together, and an absolute orientation has been computed. The model was adjusted using *bundle adjustment* and the parameters of accuracy were checked and can be read in Chap. 5.1. For *bundle adjustment*, the weight between the accuracy of point marking and accuracy of coordinates of control points was set the same as it was in processing in *Agisoft PhotoScan* and *MicMac*.

At the adjusted model, the coordinates of discrete survey points were measured on images. 313 survey points on the projected grid on the vault were measured. All the survey points lay on the vault and are used as check points for accuracy checking in this paper.

#### 5. SFM PROCESSING — RELATIVE AND ABSOLUTE ORIENTATION

*Agisoft PhotoScan* as a commercial alternative and *MicMac* as a free and open-source alternative were chosen for the SfM processing of images. Four groups of images (ordered by the scanning resolution) were computed separately.

It was important to know at least approximate focal length and size of the images. The processing without known focal length or with very different focal length could cause wrong relative orientation, or the computation of the relative orientation would not even be successful. Focal length was the same for all images of the dataset.

The relative orientation of all 4 groups of images was computed in both software programs. A full resolution of images was used for the tie point detection. It means that, in the function *Align Photos* in *Agisoft PhotoScan*, the parameter accuracy as set to *High* and in *MicMac*, processing parameter  $-1$  was set. All the images of all groups were successfully co-registered.

According to the Fig. 6, the highest number of tie points of the relative orientation, for both software programs, was while processing of images with the scanning resolution 3000 dpi. Most of the tie points were at the position of the projected grid on the vault.

The absolute orientation was calculated. For the transformation of the relative model, weights between the accuracy of marking of control points (in pixels) and accuracy of geodetically measured control points must be set. According to the reports from geodetic measurements, the accuracy of the control points coordinates was estimated as 0.02 m. The accuracy of the marking of control points was set as 0.0254 mm on a physical image, which corresponds to the 1 pixel in group

	Agisoft PhotoScan				MicMac				CDW
	1000 dpi	2000 dpi	3000 dpi	4000 dpi	1000 dpi	2000 dpi	3000 dpi	4000 dpi	4000 dpi
$RMSE_x$ [m]	0.0078	0.0035	0.0030	0.0026	0.0134	0.0057	0.0073	0.0102	0.0029
$RMSE_y$ [m]	0.0066	0.0054	0.0048	0.0047	0.0171	0.0046	0.0060	0.0102	0.0038
$RMSE_z$ [m]	0.0091	0.0095	0.0099	0.0062	0.0237	0.0096	0.0150	0.0149	0.0075
$RMSE_{xy}$ [m]	0.0102	0.0064	0.0056	0.0054	0.0217	0.0073	0.0094	0.0144	0.0048
$RMSE_{xyz}$ [m]	0.0137	0.0114	0.0114	0.0082	0.0321	0.0121	0.0177	0.0207	0.0089

TABLE 3. RMSE values of control points after bundle adjustment.

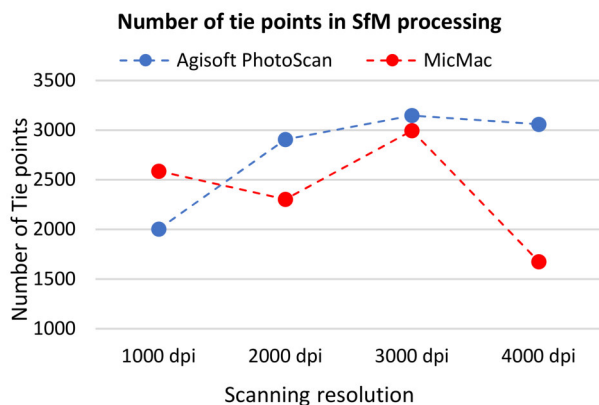
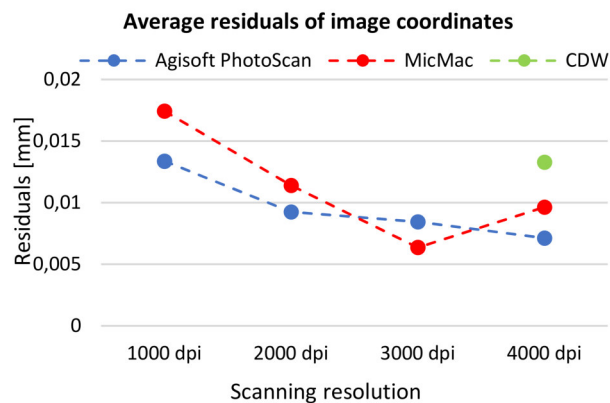


FIGURE 6. Number of tie points computed by SfM processing.

FIGURE 7. Average residuals of tie points in SfM processing (Image scale  $M = 477$ ).

of images scanned under 1000 dpi scanning resolution. The same weights were set in the processing in the *CDW*.

The bundle adjustment was computed. By that step, the positions of optical centres and rotation angles of images were computed and remained unchanged. The accuracy of the absolute orientation is given by *Root mean square error* RMSE.

According to the Tab. 3, in the case of *Agisoft PhotoScan*, there is a trend of decreasing RMSE with increasing scanning resolution. All the lowest RMSE are at 4000 dpi scanning resolution. *MicMac* gave the lowest RMSE at 3000 dpi.

After the bundle adjustment, the residuals of image coordinates of control points and tie points were checked and compared to the residuals from the *CDW* processing. In *Agisoft PhotoScan* and *MicMac*, the residuals are given in pixels and then recalculated to millimeters, see Fig. 7.

### 5.1. ACCURACY CHECKING ON CHECK POINTS

After the absolute orientation and bundle adjustment, 100 survey points were measured on images in *Agisoft PhotoScan* and *MicMac*. Both software programs support to measure discrete survey points. The discrete point surveying could be used for points which are very difficult or impossible to identify in subsequent point cloud. The 100 survey points were placed on the projected grid on the vault. The grid

was projected only on 4 images, so the survey points were marked on 4 images and their coordinates were computed.

It is presumed that the accuracy of the model is given by the highest possible accuracy of measurement in the given model. To check the accuracy, the measured survey points were used as check points and were compared to the reference coordinates computed in a *CDW*. The accuracy checking on measured check points is very important and should be carried out in every project to evaluate the processed absolute orientation.

For each group of the scanning resolution, differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between the corresponding check points computed in *CDW* and in *Agisoft PhotoScan* and between the corresponding points computed in *CDW* and in *MicMac* were calculated. The differences were analysed for each group of the scanning resolution (1000 dpi, 2000 dpi, 3000 dpi, 4000 dpi). Statistical values of set of differences are presented, see Tab. 4 and Tab. 6 and RMSE of check points are presented in Tab. 5 and Tab. 7.

The sets of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  have been tested if they have a normal distribution. The testing of normal distribution has been carried out by *Shapiro-Wilk* test of normality with a 95% confidence level (see Tab. 4 and Tab. 6) and by

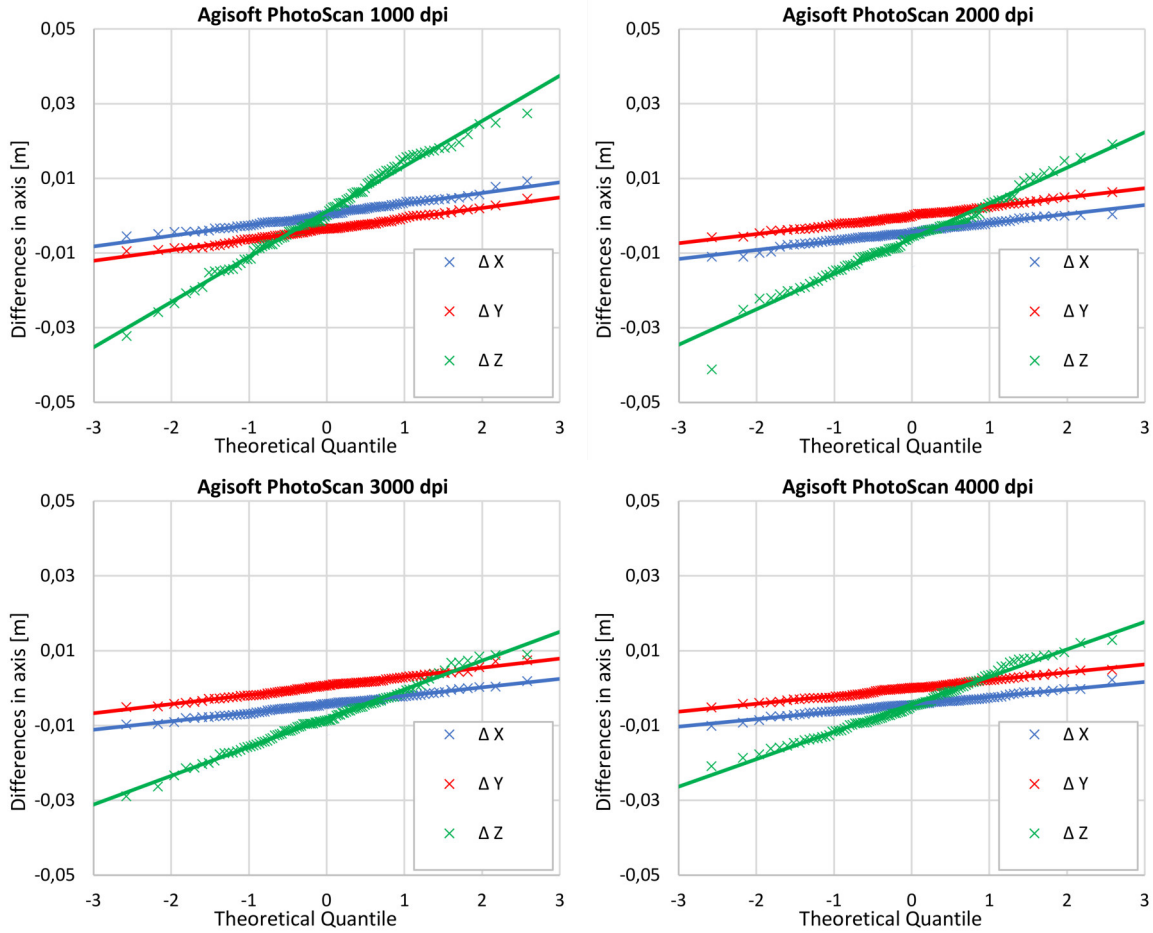


FIGURE 8. The normal  $Q - Q$  plots of values of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between check point coordinates determined by distinct point measurement in *Agisoft PhotoScan* and coordinates given by *CDW* processing.

	Agisoft PhotoScan 1000 dpi			Agisoft PhotoScan 2000 dpi			Agisoft PhotoScan 3000 dpi			Agisoft PhotoScan 4000 dpi		
	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$
Mean [m]	0.0003	-0.0036	0.0011	-0.0044	$1.11e - 6$	-0.0061	-0.0043	0.0006	-0.0080	-0.0043	$1.77e - 5$	-0.0043
Median [m]	0.0002	-0.0035	0.0001	-0.0045	$-3.3e - 5$	-0.0054	-0.0042	0.0007	-0.0086	-0.0045	0.0001	-0.0047
Std. Deviation [m]	0.0029	0.0028	0.0121	0.0024	0.0025	0.0096	0.0023	0.0024	0.0077	0.0020	0.0021	0.0074
Range [m]	0.0148	0.0141	0.0596	0.0114	0.0121	0.0602	0.0116	0.0125	0.0379	0.0123	0.0101	0.0338
Shap. - Wilk	0.291	0.763	0.696	0.238	0.939	0.161	0.965	0.754	0.746	0.552	0.835	0.436

TABLE 4. Statistical values of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between check points coordinates determined by distinct point measurement in *Agisoft PhotoScan* and coordinates given by *CDW* processing.

	Agisoft PhotoScan			
	1000 dpi	2000 dpi	3000 dpi	4000
$RMSE_x$ [m]	0.0029	0.0050	0.0049	0.0048
$RMSE_y$ [m]	0.0046	0.0024	0.0025	0.0021
$RMSE_z$ [m]	0.0121	0.0113	0.0111	0.0085
$RMSE_{xy}$ [m]	0.0054	0.0056	0.0055	0.0052
$RMSE_{xyz}$ [m]	0.0133	0.0126	0.0124	0.0100
$Max.euq.dist.$ [m]	0.0323	0.0412	0.0294	0.0218

TABLE 5. RMSE values of check points measured by distinct point measurement in *Agisoft PhotoScan* compared to the coordinates given by *CDW*.

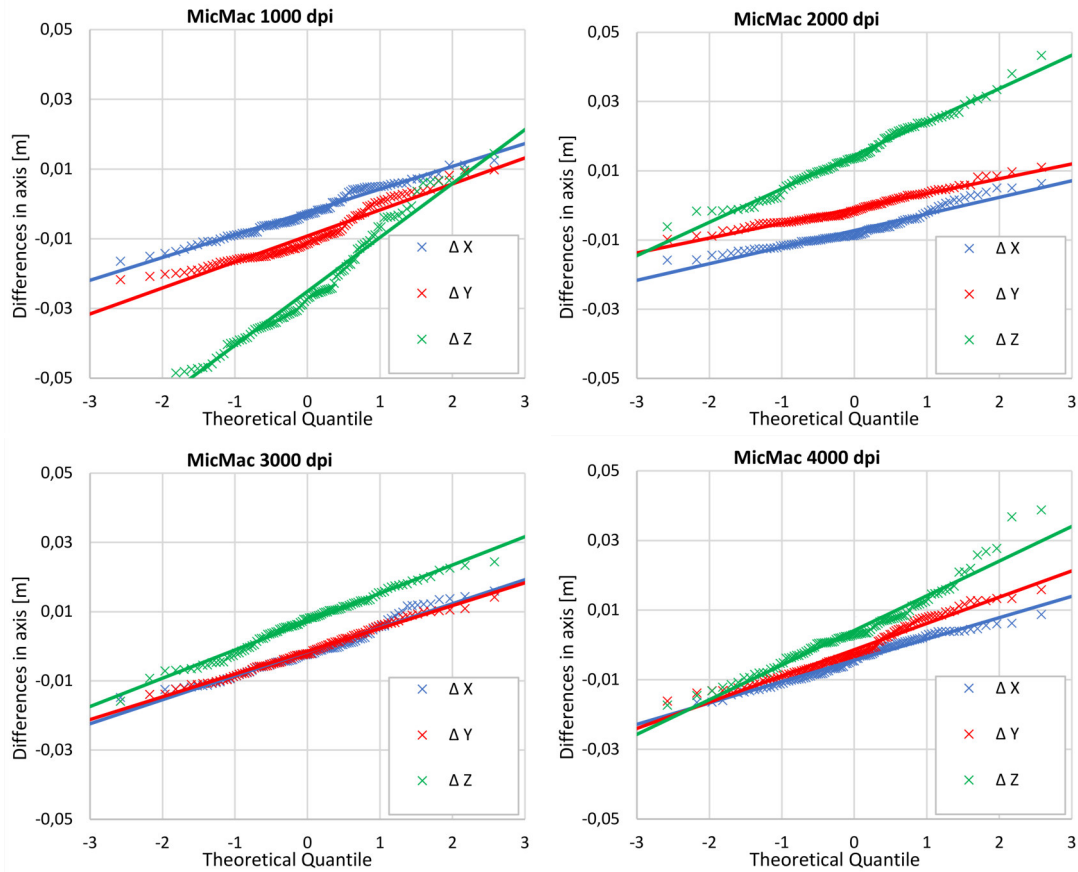


FIGURE 9. The normal  $Q - Q$  plots of values of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between check point coordinates determined by distinct point measurement in *MicMac* and coordinates given by *CDW* processing.

	MicMac 1000 dpi			MicMac 2000 dpi			MicMac 3000 dpi			MicMac 4000 dpi		
	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$
Mean [m]	-0.0023	-0.0092	-0.0251	-0.0072	-0.0009	0.0146	-0.0015	-0.0014	0.0073	-0.0044	-0.0013	0.0043
Median [m]	-0.0030	-0.0114	-0.0271	-0.0086	-0.0017	0.0137	-0.0023	-0.0021	0.0081	-0.0043	-0.0028	0.0030
Std. Deviation [m]	0.0066	0.0077	0.0157	0.0049	0.0043	0.0097	0.0071	0.0066	0.0081	0.0062	0.0077	0.0101
Range [m]	0.0289	0.0315	0.0666	0.0222	0.0208	0.0495	0.0307	0.0301	0.0403	0.0261	0.0320	0.0561
Shap. - Wilk	0.059	0.000	0.008	0.000	0.034	0.380	0.008	0.474	0.562	0.066	0.011	0.002

TABLE 6. Statistical values of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between check point coordinates determined by distinct point measurement in *MicMac* and coordinates given by *CDW* processing.

	MicMac			
	1000 dpi	2000 dpi	3000 dpi	4000 dpi
$RMSE_x$ [m]	0.0070	0.0088	0.0072	0.0076
$RMSE_y$ [m]	0.0120	0.0044	0.0067	0.0077
$RMSE_z$ [m]	0.0296	0.0174	0.0108	0.0109
$RMSE_{xy}$ [m]	0.0139	0.0098	0.0098	0.0108
$RMSE_{xyz}$ [m]	0.0327	0.0199	0.0146	0.0154
$Max.euq.dist.$ [m]	0.0534	0.0439	0.0265	0.0407

TABLE 7. RMSE values of check points measured by distinct point measurement in *MicMac* compared to the coordinates given by *CDW*.

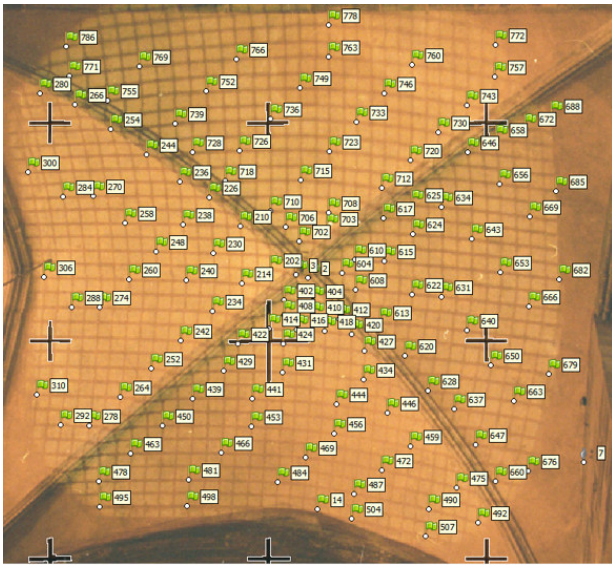


FIGURE 10. Distribution of check points on the vault.

quantile-quantile plots ( $Q - Q$  plot), see Fig. 8 and Fig. 9.

In the case of processing in *Agisoft PhotoScan*, all the datasets of differences had normal distribution. *Shapiro-Wilk* tests resulted with relatively confident  $p$ -values (lowest  $p$ -value = 0.161).  $Q-Q$  plots also show that the datasets have normal distribution even with some observed outliers. It has to be noted that the means of  $\Delta X$  are not equal to means of  $\Delta Y$  for all scanning resolutions. For scanning resolutions 2000 dpi, 3000 dpi and 4000 dpi, the  $Q - Q$  plots seem identical. The trend lines of  $\Delta X$  and trend lines for  $\Delta Y$  are more or less parallel. The trend lines of  $\Delta Y$  are shifted approximately by  $-4$  mm in  $Y$  axis on the graphs. The Tab. 5 shows that  $RMSE_x$  and  $RMSE_y$  are not equal for all four scanning resolutions. However, std. deviations of  $\Delta X$  and  $\Delta Y$  are more or less equal, see Tab. 4.  $RMSE_z$  is significantly higher as was expected by the presumption about too acute intersection angle in the  $Z$  axis. The planimetric accuracy represented by  $RMSE_{xy}$  is around 5 mm for all scanning resolutions.  $RMSE_{xyz}$  is decreasing with a higher scanning resolution. The  $RMSE_{xyz}$  is 0.01 m for the 4000 dpi scanning resolution. Maximal computed *Euclidean distance* between the check points calculated in *CDW* and in *Agisoft PhotoScan* for the 4000 dpi scanning resolution is 0.021 m. All these facts indicate that the relative and absolute orientation computed by the SfM method in *Agisoft PhotoScan* is relatively accurate and the method is applicable for processing digitized close-range analogue images captured by Rollei metric cameras. Also, the analysis showed that most of the statistical values e.g. *mean* or *range* show the best accuracy at the highest 4000 dpi scanning resolution.

In the case of *MicMac*, calculated differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between 100 points computed from *CDW* and *MicMac* does not predominantly have normal

distribution.  $P$ -values computed by *Shapiro-Wilk* test with 95 % confidence level are, in most cases, lower than 0.05. The  $Q - Q$  plots confirmed the results of *Shapiro-Wilk* testing. As it could be seen in Fig. 9, the observations do not fit the trend lines and several outliers are apparent.

$RMSE_{xyz}$  of datasets decreases over higher scanning resolution except for the 4000 dpi scanning (approximately similar to the 3000 dpi scanning).

Maximal computed *Euclidean distance* between a point calculated in *CDW* and in *MicMac* for the 3000 dpi scanning resolution is 0.027 m. In case of the 4000 dpi, the maximal *Euclidean distance* is 0.041 m due to the two significant outliers, which could be seen in  $Q - Q$  plot for 4000 dpi (Fig. 9). The analysis of results of processing in *MicMac* shows that the accuracy is lower than the accuracy in *Agisoft PhotoScan*, but even in the case of processing in *MicMac*, it was possible to get a respectable accuracy of the relative and absolute orientation.

## 6. INTERIOR ORIENTATION

Part of the SfM processing is a definition of the interior orientation of images. The interior orientation was computed by self-calibration. Considered parameters of the interior orientation in this paper are:

- Principal distance
- Principal point
- Parameters  $K_1$ ,  $K_2$  and  $K_3$  of polynomial function describing radial distortion on images

The Lens of Rollei metric cameras are interchangeable and focusable so there was a presumption that the principal distance and position of the principal point calculated by self-calibration could be different to the values stated in the calibration report. However, the values determined by the self-calibration should not differ between each other. Very similar results for all scanning resolutions were computed by self-calibration in *Agisoft PhotoScan*. However, values computed in *MicMac* are not similar to each other, see Tab. 8. The check of the computed principal distance and position of principal point showed reliability of self-calibration in *Agisoft PhotoScan*.

In this paper, radial distortion is considered as ideally symmetrical. The comparison of radial distortions calculated from the self-calibration are stated in Fig. 11 and Fig. 12 where they are compared to the radial distortion stated in calibration report.

In the case of *Agisoft PhotoScan* the radial distortion is more or less the same for all scanning resolutions and also fits to the radial distortion stated in the calibration report. Curves of the radial distortion of *MicMac* differ more. The largest difference at 15 mm from the image centre is around 0.14 mm on the image. Because the majority of the tie and check points were within 10 mm, in Fig. 11 and



	$f$ [mm]	$x_p$ [mm]	$y_p$ [mm]
Calibration report	28.76	-0.14	-0.14
PhotoScan 1000 dpi	28.80	-0.10	-0.02
PhotoScan 2000 dpi	28.85	-0.07	-0.02
PhotoScan 3000 dpi	28.81	-0.11	0.01
PhotoScan 4000 dpi	28.82	-0.06	-0.03
MicMac 1000 dpi	29.40	-0.18	0.26
MicMac 2000 dpi	28.51	0.25	-0.11
MicMac 3000 dpi	28.43	-0.16	0.12
MicMac 4000 dpi	29.51	-0.27	-0.51

TABLE 8. Values of principal distance and principal point determined by self-calibration of each processing (Image scale  $M = 477$ ).

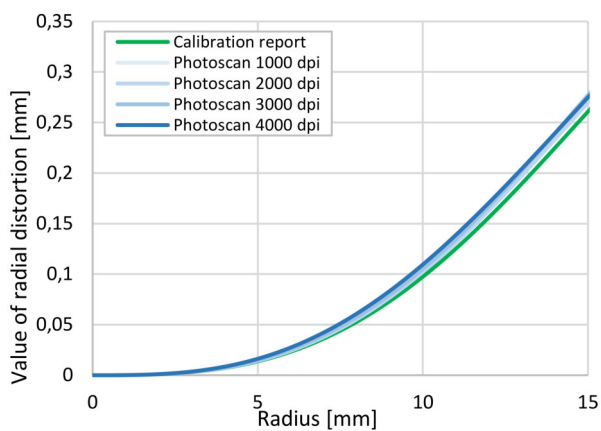


FIGURE 11. Radial distortion determined by self-calibration in *Agisoft PhotoScan* (Image scale  $M = 477$ ).

Fig. 12 the curves of radial distortion are presented just until 15 mm. In this case, it would not be relevant to present resulted radial distortion curves until the image corner.

The results of the self-calibration in *Agisoft PhotoScan* do not differ much from each other and thus seem to be more reliable.

## 7. POINT CLOUD GENERATION

The subsequent result of SfM processing is a point cloud. The point cloud is generated based on *multi-view stereo matching* or *pair-wise stereo matching* algorithms. The point cloud is obtained by *Build Dense Cloud* function in *Agisoft PhotoScan* and *C3DC* command in *MicMac*. The important parameter for the point cloud generation is the usable resolution on images (parameter *Quality* in *Agisoft PhotoScan* and *ZoomF* in *MicMac*). As a first attempt, the full resolution of images has been used for the point cloud generation. *Agisoft PhotoScan* generated the densest point cloud with the most points with full resolution parameter of images at the highest 4000 dpi scanning resolution. *MicMac* generated the point cloud with the highest number of points at

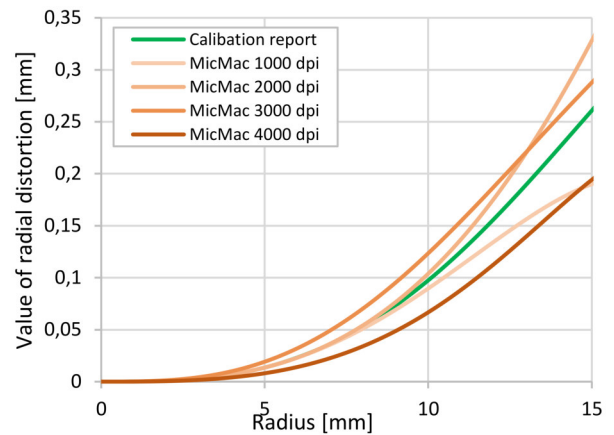


FIGURE 12. Radial distortion determined by self-calibration in *MicMac* (Image scale  $M = 477$ ).

the 4000 dpi scanning resolution with a full resolution parameter as well. But in the case of *MicMac*, the points were not equally distributed at all. There were very large parts with missing points. Similar situation occurred in the case of 2000 dpi and 3000 dpi scanning resolution.

Those point clouds would not be possible to use for further processing and modeling. To have a reasonable point cloud, which could possibly be used for further modelling and for analysis in this paper, it was necessary to set a lower used resolution as a parameter for the point cloud generation in *Agisoft PhotoScan* and in *MicMac*. A parameter, which each side of image resolution divides by 4, had to be used.

All generated point clouds were compared to the 313 survey points computed in the *CDW* by *Cloud-to-cloud* method. For each point of 313 survey points computed in *CDW*, the nearest corresponding point in the generated point cloud was found. Then, the differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  were computed between the corresponding points. The normal distribution of sets of differences were tested by *Shapiro-Wilk* test of normality with a 95 % confidence level and by *quantile-quantile plots* ( $Q - Q$  plot).

In the case of point clouds generated in *Agisoft PhotoScan*,  $p$ -values computed by *Shapiro-Wilk* test showed a sufficient confidence except for sets  $\Delta X$  and  $\Delta Y$  at the point cloud generated from the 1000 dpi scanning resolution images, see Tab. 9. The  $Q - Q$  plots show few outliers in each sets but most of the observations fit to the trend line of  $Q - Q$  plot. It has to be noted that trend lines of  $Q - Q$  plots of  $\Delta X$  and  $\Delta Y$  seem for all sets approximately equal, unlike the trend lines in  $Q - Q$  plots in relative and absolute orientation accuracy analysis (Chap. 5.1.). Also  $RMSE_x$  and  $RMSE_y$  (see Tab. 10) are almost equal, unlike Tab. 5. Std. deviations of  $\Delta X$  and  $\Delta Y$  are also similar to each other for all scanning resolution.  $RMSE$ , *std. deviation* and *range* show the increasing accuracy

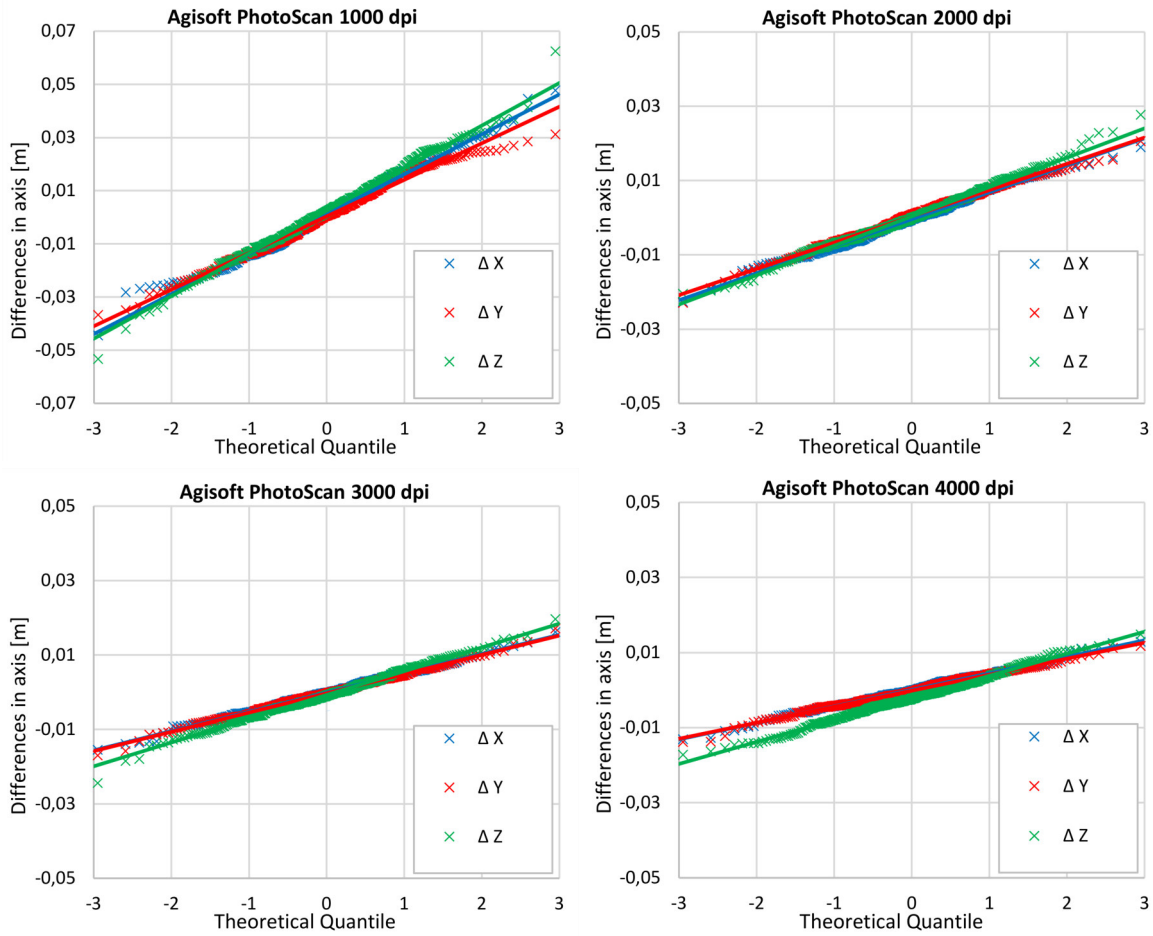


FIGURE 13. The normal  $Q - Q$  plots of values of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between point cloud generated in *Agisoft PhotoScan* and point coordinates given by *CDW*.

	Agisoft PhotoScan 1000 dpi			Agisoft PhotoScan 2000 dpi			Agisoft PhotoScan 3000 dpi			Agisoft PhotoScan 4000 dpi		
	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$
Mean [m]	0.0012	0.0003	0.0025	-0.0004	0.0003	0.0003	-0.0001	-0.0003	-0.0007	$6.6e - 5$	-0.0002	-0.0021
Median [m]	0.0019	0.0006	0.0022	-0.0003	0.0008	$-9.2e - 5$	-0.0002	-0.0004	-0.0009	-0.0001	-0.0005	-0.0022
Std. Deviation [m]	0.0151	0.0138	0.0161	0.0073	0.0071	0.0079	0.0052	0.0052	0.0064	0.0044	0.0043	0.0059
Range [m]	0.0923	0.0679	0.1157	0.0416	0.0453	0.0482	0.0317	0.0341	0.0441	0.0260	0.0256	0.0321
Shap. - Wilk	0.051	0.010	0.527	0.122	0.480	0.179	0.824	0.864	0.735	0.387	0.228	0.407

TABLE 9. Statistical values of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between point cloud generated in *Agisoft PhotoScan* and point coordinates given by *CDW*.

	Agisoft PhotoScan			
	1000 dpi	2000 dpi	3000 dpi	4000 dpi
$RMSE_x$ [m]	0.0151	0.0073	0.0052	0.0044
$RMSE_y$ [m]	0.0138	0.0071	0.0052	0.0043
$RMSE_z$ [m]	0.0162	0.0079	0.0064	0.0062
$RMSE_{xy}$ [m]	0.0205	0.0102	0.0073	0.0061
$RMSE_{xyz}$ [m]	0.0261	0.0129	0.0097	0.0087
$Max.euq.dist.$ [m]	0.0640	0.0280	0.0256	0.0191

TABLE 10. RMSE values of differences between point cloud generated in *Agisoft PhotoScan* and point coordinates given by *CDW*.

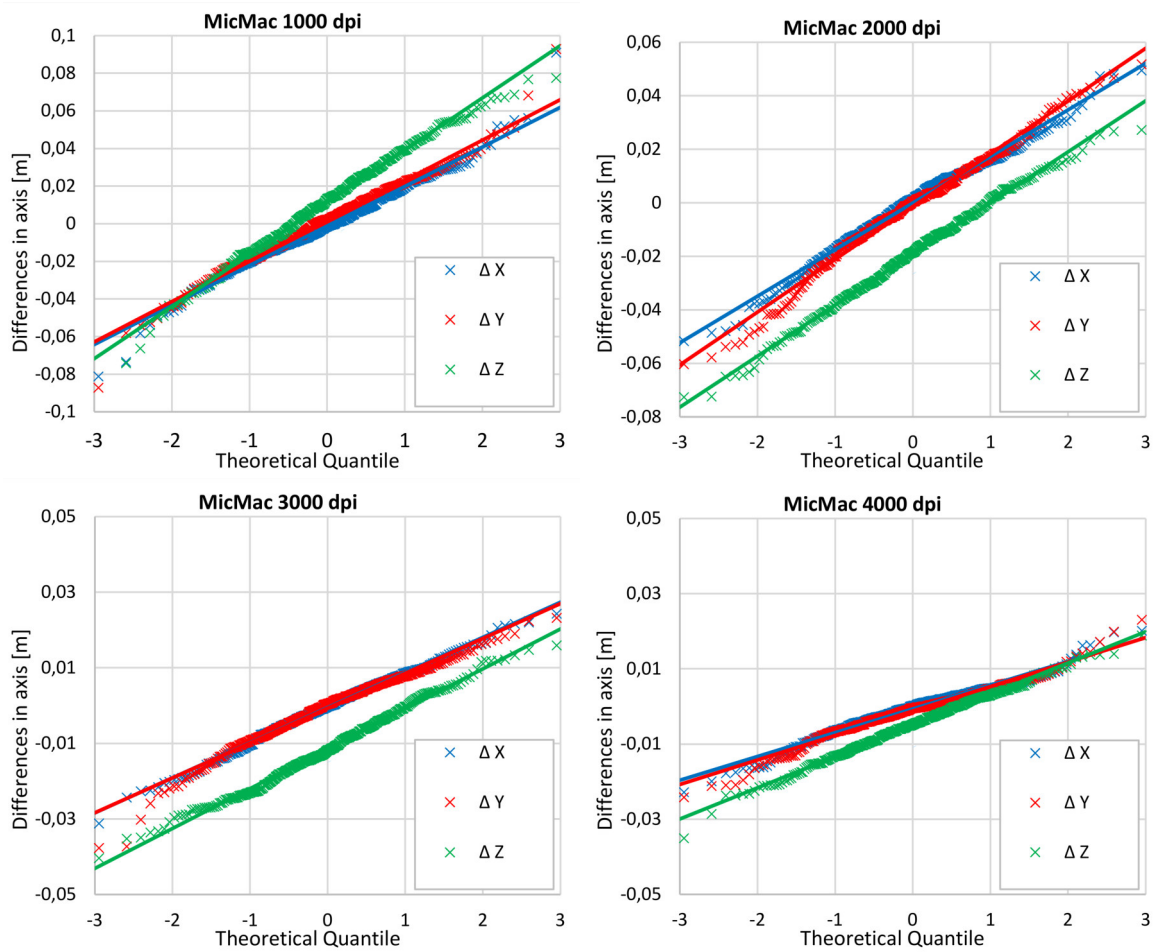


FIGURE 14. The normal  $Q - Q$  plots of values of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between point cloud generated in *MicMac* and point coordinates given by *CDW*.

	MicMac 1000 dpi			MicMac 2000 dpi			MicMac 3000 dpi			MicMac 4000 dpi		
	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$	$\Delta X$	$\Delta Y$	$\Delta Z$
Mean [m]	-0.0012	0.0015	0.0114	-0.0001	-0.0014	-0.0191	-0.0005	-0.0007	-0.0114	-0.0007	-0.0012	-0.0050
Median [m]	-0.0014	0.0021	0.0132	0.0013	0.0004	-0.0185	-0.0002	0.0005	-0.0118	0.0002	-0.0005	-0.0048
Std. Deviation [m]	0.0212	0.0216	0.0279	0.0175	0.0198	0.0191	0.0093	0.0093	0.0106	0.0064	0.0066	0.0083
Range [m]	0.1722	0.1801	0.1827	0.1012	0.1120	0.0999	0.0556	0.0609	0.0563	0.0429	0.0472	0.0539
Shap. - Wilk	0.003	0.007	0.036	0.023	0.022	0.566	0.428	0.000	0.194	0.000	0.000	0.582

TABLE 11. Statistical values of differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between point cloud generated in *MicMac* and point coordinates given by *CDW*.

	MicMac			
	1000 dpi	2000 dpi	3000 dpi	4000 dpi
$RMSE_x$ [m]	0.0212	0.0175	0.0093	0.0064
$RMSE_y$ [m]	0.0216	0.0198	0.0093	0.0067
$RMSE_z$ [m]	0.0301	0.0270	0.0155	0.0097
$RMSE_{xy}$ [m]	0.0303	0.0264	0.0132	0.0093
$RMSE_{xyz}$ [m]	0.0427	0.0378	0.0204	0.0134
$Max.euq.dist.$ [m]	0.1202	0.0845	0.0425	0.0359

TABLE 12. RMSE values of differences between point cloud generated in *MicMac* and point coordinates given by *CDW*.



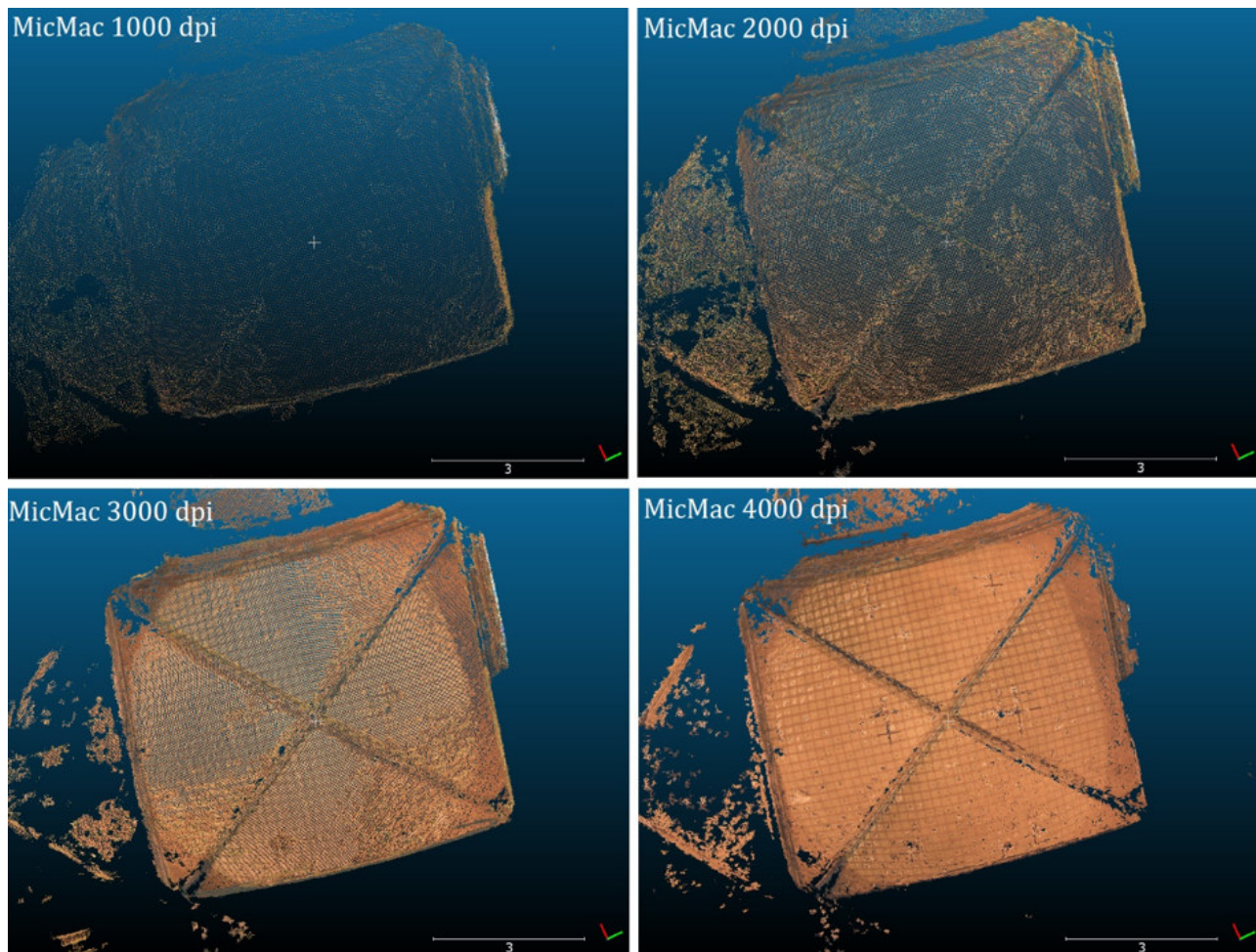


FIGURE 15. Density of point cloud at different scanning resolution (parameter ZoomF = 4).

with increasing scanning resolution, see Tab. 9 (the best results are at the 4000 dpi scanning resolution). Maximal *Euclidean distance* between corresponding points at the 4000 dpi scanning resolution is 0.019 m.

Differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between corresponding points of point clouds generated by *MicMac* and survey points in *CDW* are not, in most of the cases, normally distributed according to the *Shapiro-Wilk* test, see Tab. 11. The  $Q - Q$  plots show that most of the observations fit the trend lines but many outliers were observed. The analysis of point clouds generated by *MicMac* showed more present outliers than the point clouds generated by *Agisoft PhotoScan*.  $RMSE_x$  and  $RMSE_y$  are also more or less equal like in the case of the point clouds generated in *Agisoft PhotoScan*. Std. deviations of sets of  $\Delta X$  and  $\Delta Y$  are also very similar to each other. The maximal *Euclidean distance* is decreasing with increasing scanning resolution. The lowest maximal *Euclidean distance* 0.036 m was observed in the point cloud generated in *MicMac* using images with the 4000 dpi scanning resolution. Few other statistical parameters show the best accuracy at 4000 dpi scanning resolution, see Tab. 11 and Tab. 12.

Other results than point cloud could be textured triangulated model, orthophoto or DEM in raster form. For hypsometric model of vault, just the point cloud would be sufficient as the result. The generated point cloud can be seen at Fig. 4 or Fig. 15.

## 8. CONCLUSION

The paper shows SfM processing of digitized analogue images captured by Rollei metric camera. The images were originally used for a different method of photogrammetry. The images were originally created for the purpose of photogrammetric documentation of vault of Zikmund's chapel in Saint Vitus Cathedral. Nowadays, with the SfM processing, it is theoretically possible to get new and valuable 2D and 3D results. This paper shows the method how to digitize the analogue negatives, how to preprocess them and how to process them using the SfM method. As a software for SfM processing, *Agisoft PhotoScan* was chosen as a commercial alternative and *MicMac* as a free and open-source alternative. Then, the results were compared to the results in the *CDW*, by which the original processing was carried out. Results from the *CDW* were considered as the reference.

Before the processing, it is necessary to deal with digitizing original negatives and with preprocessing, in order to obtain accurate results from the SfM processing. Seven negatives of images were chosen for the experiment. According to the resolving power, four different scanning resolutions were chosen - 1000 dpi, 2000 dpi, 3000 dpi and 4000 dpi.

The digitized images were processed in *Agisoft PhotoScan* and in *MicMac* by the SfM method and relative and absolute orientation was computed. The relative and absolute orientation was checked on 100 check points computed in the *CDW* by the original method. The differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between corresponding check points were computed. In the case of *Agisoft PhotoScan*, the differences between check points in all axes had normal distribution, unlike differences computed in *MicMac*. According to the showed results, it is possible to conclude that *Agisoft PhotoScan* gave better results than *MicMac* in terms of accuracy. The resulted comparison showed lower *RMSE*, lower *maximal Euclidian distance*, lower *range* and lower *std. deviation*, etc. Despite that fact, *MicMac* was also able to compute relatively accurate relative and absolute orientation and the results could be sufficient for some purposes. Also, the analysis showed that with increasing scanning resolution, it is possible to get more accurate results.

Parameters of interior orientation were determined by self-calibration by processing in *Agisoft PhotoScan* and *MicMac*. Resulted parameters in *Agisoft PhotoScan* were very similar for all scanning resolutions and it shows the reliability of processing. Results of interior orientation in *MicMac* differ to each other, some of them significantly. Curves of radial distortion computed in *Agisoft PhotoScan* fit to each other and to the radial distortion stated in calibration report. Curves of radial distortion computed in *MicMac* differ by approximately 0.14 mm on image at 15 mm distance from the image center.

As one of the results of processing, point clouds were generated in both software programs for each scanning resolution. The point clouds were compared to the 313 survey points calculated in the *CDW*. For the comparison, *cloud-to-cloud* method was used and the differences  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  between the points were analysed. The analysis of accuracy showed that *Agisoft PhotoScan* is a sufficient tool for point cloud generation based on digitized analogue images. Most of the differences in all three axis were normally distributed according to the *Shapiro-Wilk* tests and to the  $Q-Q$  plots. In the case of point clouds generated in *MicMac*, most of the differences were not normally distributed according to the *Shapiro-Wilk* tests. Most of the observed values fit to the trend lines of the  $Q-Q$  plot, which could be characteristic for the fact that the observations are normally distributed, but a lot of outliers, which could cause the results of *Shapiro-Wilk* tests, were presented.

It could be stated that the chosen scanning resolution of images truly affects the accuracy of results. Almost all the results of analysis showed that with increasing scanning resolution we get better accuracy. The biggest differences were between the processing of images with a lower scanning resolution. It can be stated that scanning with the highest scanning resolution brought better results. However, one should consider that the higher resolution can be considerably more time consuming. So a compromise between the accuracy and the time should be made while dealing with a large set of images.

The experiment in this paper showed that it is possible to use the SfM method for a processing of digitized analogue images captured by Rollei metric camera and originally used for a different method of photogrammetric processing. *Agisoft PhotoScan* gave more accurate and reliable results. Differences in coordinates between corresponding check points calculated by the SfM processing and by the *CDW* were almost in all cases normally distributed. In the case of *MicMac*, most of the differences were not normally distributed, possibly because of several outliers in the observations. Nevertheless, for some kind of purposes, even the results from *MicMac* could be sufficient. With respect to the fact that *MicMac* is free, *MicMac* could be an option for SfM processing of digitized analogue images captured by Rollei metric cameras as well.

#### ACKNOWLEDGEMENTS

This project was supported by grant of the Grant Agency of the Czech Technical University in Prague, grant No. SGS20/053/OHK1/1T/11. This project is a part of internal research of EuroGV s.r.o. company.

#### REFERENCES


- [1] A. Dlesk, P. Raeva, K. Vach. Possibilities of processing archival photogrammetric images captured by Rollei 6006 metric camera using current method. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLII-2**:319 – 323, 2018. DOI:10.5194/isprs-archives-XLII-2-319-2018.
- [2] Z. Poloprutský, M. Fraštia, M. Marčíš. 3D digital reconstruction based on archived terrestrial photographs from metric cameras. *Acta Polytechnica* **59**(4):384 – 398, 2019. DOI:10.14311/AP.2019.59.0384.
- [3] Á. Rodríguez Miranda, J. M. Valle Melón. Recovering old stereoscopic negatives and producing digital 3D models of former appearances of historic buildings. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLII-2/W3**:601 – 608, 2017. DOI:10.5194/isprs-archives-XLII-2-W3-601-2017.
- [4] K. Pavelka, J. Šedina, P. Raeva, M. Hůlková. Modern processing capabilities of analog data from documentation of the Great Omayyad mosque in Aleppo, Syria, damaged in civil war. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLII-2/W5**:561 – 565, 2017. DOI:10.5194/isprs-archives-XLII-2-W5-561-2017.

- [5] L. J. Vargo, B. M. Anderson, H. J. Horgan, et al. Using structure from motion photogrammetry to measure past glacier changes from historic aerial photographs. *Journal of Glaciology* **63**(242):1105 – 1118, 2017. DOI:<https://doi.org/10.1017/jog.2017.79>.
- [6] M. V. Peppas, J. P. Mills, K. D. Fieber, et al. Archaeological feature detection from archive aerial photography with a sfm-mvs and image enhancement pipeline. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLII-2**:869 – 875, 2018. DOI:10.5194/isprs-archives-XLII-2-869-2018.
- [7] F. Maiwald, T. Vietze, D. Schneider, et al. Photogrammetric analysis of historical image repositories for virtual reconstruction in the field of digital humanities. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLII-2/W3**:447 – 452, 2017. DOI:10.5194/isprs-archives-XLII-2-W3-447-2017.
- [8] M. Cogliati, E. Tonelli, D. Battaglia, M. Scaioni. Extraction of DEMs and orthoimages from archive aerial imagery to support project planning in civil engineering. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **IV-5/W1**:9 – 16, 2017. DOI:10.5194/isprs-annals-IV-5-W1-9-2017.
- [9] S. Giordano, A. Le Bris, C. Mallet. Toward automatic georeferencing of archival aerial photogrammetric surveys. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **IV-2**:105 – 112, 2018. DOI:10.5194/isprs-annals-IV-2-105-2018.
- [10] T. Luhmann, S. Robson, S. Kyle, L. Harley. *Close Range Photogrammetry: Principles, Methods and Applications*. Whittles Publishing, 2006.
- [11] G. Suilmann. Rolleiflex 6008 metric - a new camera for industrial photogrammetry. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XXIX-Part B5**:60 – 63, 1992.
- [12] G. Pomaska. Image acquisition for digital photogrammetry using off the shelf and metric cameras. *Proceedings of the 18th International Symposium CIPA* pp. 490 – 495, 2001.
- [13] A. Streilein, H. Beyer, T. Kersten. Digital photogrammetric techniques for architectural design. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XXIX-B5**:825 – 831, 1992.
- [14] Agisoft. Agisoft Photoscan user manual, professional edition, version 1.4. [http://www.agisoft.com/pdf/photoscan-pro\\_1\\_4\\_en.pdf](http://www.agisoft.com/pdf/photoscan-pro_1_4_en.pdf). Accessed: 15 July 2020.
- [15] Micmac, apero, pastis and other beverages in a nutshell! <http://www.github.com/micmacIGN/Documentation/blob/master/DocMicMac.pdf>. Accessed: 15 July 2020.
- [16] E. Rupnik, M. Daakir, M. P. Deseilligny. Micmac—a free, open-source solution for photogrammetry. *Open Geospatial Data, Software and Standards* **2**(1):1 – 9, 2017. DOI:10.1186/s40965-017-0027-2.
- [17] J. M. Friedt. Photogrammetric 3D structure reconstruction using MicMac. [http://jmfriedt.free.fr/lm\\_sfm\\_eng.pdf](http://jmfriedt.free.fr/lm_sfm_eng.pdf). Accessed: 15 July 2020.
- [18] Technical data sheet, konica color vx400 film. <http://www.photoweb.ru/exusr/pdf/konica/vx400.pdf>. Accessed: 15 July 2020.
- [19] F. S. Patrick Wagner. Film scanner test-reports: complete overview of all the models. <http://www.filmscanner.info/en/FilmScannerTestberichte.html>. Accessed: 15 July 2020.
- [20] R. Kotowski, B. Weber. A procedure for on-line correction of systematic errors. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XXV-A3**:553 – 560, 1984.

## **Article 2**

Article

# From Analogue to Digital Photogrammetry: Documentation of Padise Abbey in Two Different Time Stages

Adam Dlesk <sup>1,\*</sup>, Andres Uueni <sup>2</sup>, Karel Vach <sup>3</sup> and Jüri Pärtna <sup>4</sup> 

<sup>1</sup> Department of Geomatics, Faculty of Civil Engineering, Czech Technical University in Prague, 16629 Prague, Czech Republic

<sup>2</sup> Cultural Heritage and Conservation Department, Estonian Academy of Arts, 10412 Tallinn, Estonia; andres.uueni@artun.ee

<sup>3</sup> EuroGV, spol. s r. o., 11000 Prague, Czech Republic; vach@eurogv.cz

<sup>4</sup> DataCap OÜ, 10416 Tallinn, Estonia; juri.partna@datacap.ee

\* Correspondence: adam.dlesk@fsv.cvut.cz

Received: 4 November 2020; Accepted: 22 November 2020; Published: 24 November 2020



**Abstract:** Exploration of historical buildings during their life cycles is one of many aspects of the cultural heritage research. For studying different conditions of historical buildings over time, it is convenient to use photogrammetric surveys carried out in different time stages. Modern photogrammetric methods give the possibility to use archived analogue photogrammetric data, re-process them, and achieve new potentially valuable results with sufficient geometric accuracy and with high visual quality. The paper presents re-processing of analogue archival photogrammetric images captured by a metric camera in 1991 and compares the results to the new photogrammetric survey with a digital camera. The object of interest is Padise Abbey in Estonia. In 1991, parts of the abbey were in poor condition and conservation and renovation works were at an early stage. Nowadays, the abbey is partly renovated, and the results of the new photogrammetric survey give an opportunity to study two completely different conditions.

**Keywords:** photogrammetry; structure from motion; analogue images; UMK metric camera; Padise Abbey

---

## 1. Introduction

Photogrammetry is a method that is used in the fields of cultural heritage and archeology to get a documentation of cultural heritage objects [1–5]. Due to the photogrammetric surveys carried out in different time stages, we can study the objects and explore two different conditions of the object.

In the past, metric analogue cameras were used for photogrammetric surveys and single-image photogrammetry and stereophotogrammetry were prevailing methods [6]. As a storing medium of photogrammetric images, glass plate negatives or photographic film were used. In the past, the metric cameras and other equipment were very expensive, and the processing required well-trained workers [7]. In 1990s, photogrammetry experienced a revolution with the arrival of digital cameras. Nowadays, close range images captured by digital cameras are often processed by effective and fast structure from motion (SfM) method. The structure from motion method has a background in computer vision science, which has different approach. The methods of computer vision were originally aimed at object recognition, navigation of autonomous vehicles, or object modelling. The internal parameters of source images for computer vision were not usually known in contrast to analogue photogrammetry, where the cameras had high optic quality with calibrated lenses and where the main topic was accuracy of a measurement [8].



Digitizing and re-processing of analogue images of photogrammetric surveys carried out during analogue era give an opportunity to get new and potentially valuable results. For example, one of the results of the re-processing of digitized analogue images could be point cloud and orthophoto of a cultural heritage object. These results can help to study how the object has changed. Re-processing of the analogue images could have high meaning in some particular cases. For example, re-processing of images was used by [9] to make a 3D restitution of destroyed and buried historical structures of the Beaufort Castle in Lebanon. Processing of digitized analogue images of the Kronentor of the Dresden Zwinger captured by non-metric cameras with mostly unknown camera geometry was presented by [10]. Archival analogue photogrammetric images of the Great Mosque in Aleppo, Syria were used and re-processed by [11]. The mosque has been damaged during the ongoing war. The analogue images were digitized and re-processed by structure from motion method and 3D results were generated. It is possible to study the former state of the mosque on those 3D results and it can help during the future renovation. By [12], the change of Erbil Al-Qala citadel was presented on comparison of historical and contemporary data.

Between 1972 and 1999, the group around Kalju Pärtna from the National Design Institute for Cultural Heritage carried out several photogrammetric surveys of important cultural heritage sites in Estonia and abroad. Photogrammetric images of the surveys contain important and high-quality records about heritage site conditions and provide valuable information for heritage researchers, owners, architects, engineers, conservators, and renovators. In 1991, one of the photogrammetric surveys was a survey of the south wall of Estonian Padise Abbey. At that time, walls of the abbey were in poor condition. Nowadays, the south wall has been renovated, so it is convenient to present the re-processing of analogue images using SfM method in this case. Twelve of the analogue images (six stereo-pairs) captured by UMK metric camera were digitized and re-processed. In 2017, new images were captured by digital camera and the images were processed by SfM method. Due to resulting point clouds and orthophotos by re-processing and by new survey, it is possible to compare the former and current state of the walls.

### *1.1. The Importance of Padise Abbey, Historical Background*

Padise Area was an important trade settlement, e.g., the location of the stronghold suggests connections with the sea and a harbor site. Padise Stronghold, or the Old Stronghold (Vanalinnamägi) as it is popularly known, is a site of the Middle Iron Age (7th–8th centuries) hill fort.

There is the Padise Abbey (Estonian: Padise klooster) close to the Old Stronghold. The Abbey was a former Cistercian monastery in Padise in Harju County, Estonia, settled in 1310 by the dispossessed monks of Dünamünde Abbey in Latvia. It was converted into a fortress after its dissolution in 1559 and later used as a country house until 1766. The ruins are a museum now.

The interest of monks from Dünamünde Abbey in the present Daugavgrīva near Riga in Padise was first documented in 1283 in a letter from King Eric V of Denmark regarding the acquisition of land for the construction of a Cistercian monastery, but almost certainly they had had a presence there for several decades previously as a part of the Christianization of the territories of Estonia newly conquered by the Teutonic Knights. In 1305, Dünamünde Abbey was appropriated by the Teutonic Knights and the monks dispossessed. Then, King Eric VI of Denmark gave them permission to build a fortified monastery in Padise, where they moved in 1310, although construction of the stone buildings did not begin until 1317. The new monastery was made subordinate to Stolpe Abbey in Pomerania in 1319.

By 1343, at the time of the St. George's Night Uprising, when it was still only partly built, the monastery was burnt down and rebuilding began after 1370.

By about 1400, the monastery had acquired extensive estates in Estonia and also in southern Finland and throughout the 15th century, it enjoyed a period of great prosperity and influence as one of the most important spiritual centers of Estonia.

During the Livonian War, in 1559, the monks were ejected, and the buildings and estates were confiscated. The monastery became a fortress, which the Swedes duly took in 1561, Russians took in 1576, and Swedes took again in 1580.

In 1622, King Gustavus Adolphus of Sweden gave the estates of the former Padise Abbey (Figure 1) to Thomas Ramm and it remained in the possession of his family until 1919 [13–19].



**Figure 1.** Padise Abbey in 2017.

### 1.2. UMK Camera

For the photogrammetric survey, Kalju Pärtna and his team chose a metric camera UMK 6.5/1318 with a super wide lens SuperLamegon PI 5.6/64. From the photogrammetric perspective, for re-processing of the analogue images, it was important to study the used camera.

UMK cameras (Universal Messungs Kammer) were metric cameras with high-quality lenses manufactured by Carl Zeiss Company in Jena, formerly East Germany. As a first camera of the series, UMK 10/1318 with a wide-angle lens Lamegon 8/100A was designed. The designation of UMK 10/1318 refers to the focal length of the used lens (10 cm) and to the dimension of the storing medium (13 × 18 cm). As a universal UMK 10/1318 was not ideal for some photogrammetric applications, Carl Zeiss Company came with other types of UMK cameras. There were manufactured cameras UMK 20/1318 and UMK 30/1318 with longer focal lengths but also camera UMK 6.5/1318 with a super wide-angle lens SuperLamegon PI 5.6/64.

The lenses of cameras were focusable in steps, except of Super-Lamegon, which had a fixed-focus set to 8 m. UMK cameras were equipped with a shutter and adjustable aperture. The camera was possible to rotate by  $\pm 15^\circ$  and  $\pm 30^\circ$  upwards or downwards.

The construction of cameras was designed to have long-term steady parameters of interior orientation with low distortion. The parameters of interior orientation were usually calibrated in laboratories.

As a storing medium, glass plate negatives with dimension 13 × 18 cm (useful dimension was 120 × 166 mm) or roll-film 19-cm wide were used. The glass plate negatives were possible to use in portrait orientation or landscape orientation. A frame with fiducial marks that were projected to the

image was placed in front of the images. It was possible to make marks on the frame (e.g., calibrated camera constant of camera, position of an image in stereo-pair) [20].

## 2. Original Survey

In 1991, K. Pärtna and his team from the National Design Institute for Cultural Heritage carried out a photogrammetric survey of the south wall of the Padise Abbey. As a photogrammetric method he chose stereophotogrammetry. Using a metric camera UMK 6.5/1318, he captured six stereo-pairs. As well as the glass plate negatives, we found only some of the documentation to the photogrammetric survey. From that documentation, we were able to read only partially how the original survey was carried out. Unfortunately, the calibration report of the camera was not found.

### 2.1. Photogrammetric Survey

From the available sketch of the original photogrammetric survey, we were able to read that K. Pärtna and his team captured images in stereo-pairs. The capturing scenario is apparent from the sketch (see Figure 2). Three stereo-pairs (A1–B1, A2–B2, A3–B3) were captured in one line and one stereo-pair (A5–B5) was shifted by 1 m to the line. One stereo-pair (A4–B4) was captured from an unknown further distance (according to the sketch, approximately 16 m). Stereo-pair A1\_15–B1\_15 (missing in Figure 2) was captured at exactly the same place as stereo-pair A1–B1 but the camera was rotated by  $15^\circ$  upwards. The distances between the images of the same stereo-pair were measured and were reported in an available document. There were markers M1–M6 on the south wall, signaled by crosses carved to the stones of the south wall of the abbey. The distances between those markers were measured most likely using a long tape. The result of the photogrammetric survey was a drawing of stones of the wall.

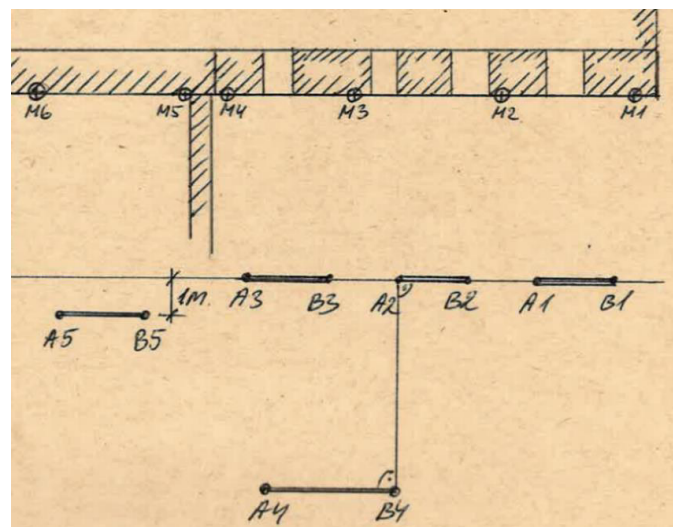


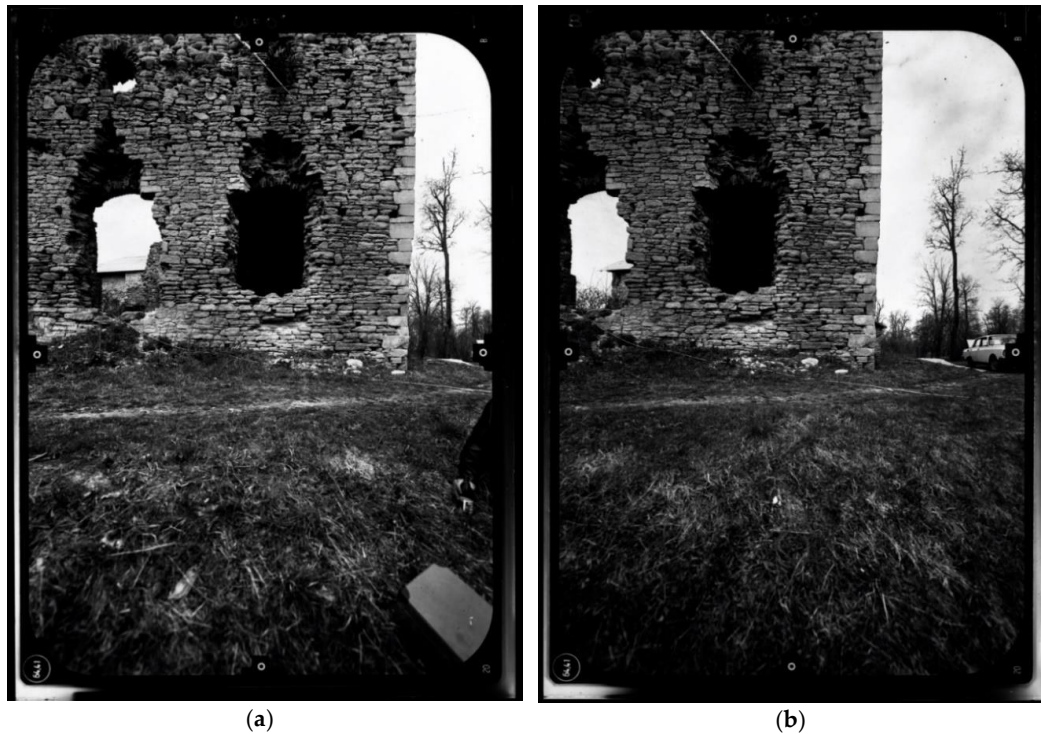
Figure 2. Sketch of original photogrammetric survey in 1991.

### 2.2. Captured Images

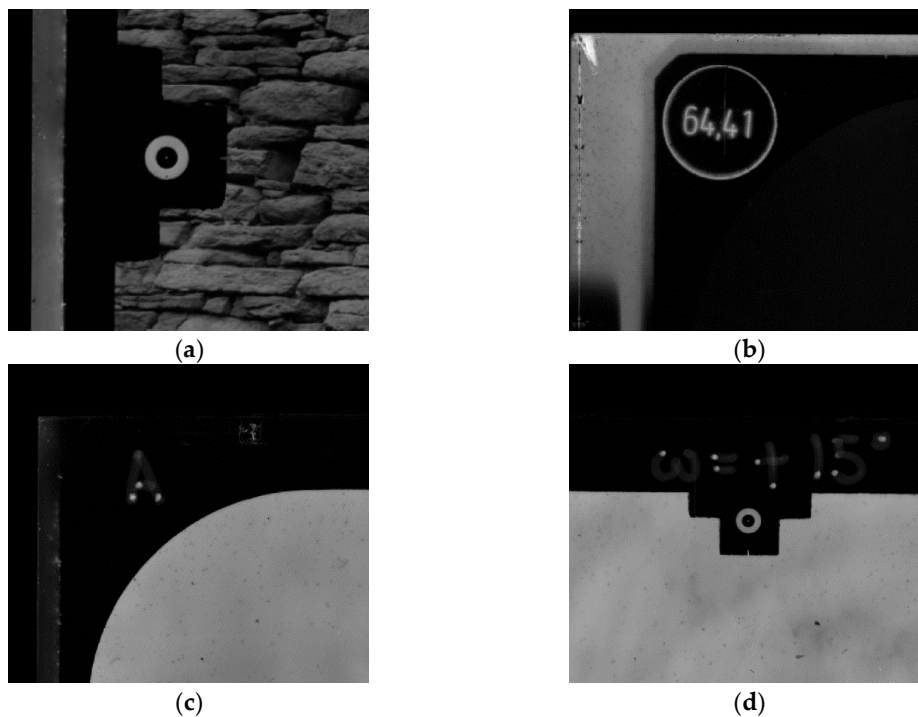
Black and white images captured by UMK 6.5/1318 camera were stored on glass plate negatives (Figure 3). The used frame had four fiducial marks with calibrated coordinates. We also found information about calibrated camera constant of the used lens in mm on the frame. The used lens had camera constant  $c = 64.41$  mm. On the frame, there were also found marks which signify the left/right position of an image in a stereo-pair (left image marked by “A” and right image marked by “B”). Another mark on the frame was a rotation angle  $\omega$  (Figure 4). Most of the images were marked as  $\omega = 0$ , so it means that zenith angle of the optical axis was  $z = 90^\circ$ . Images of stereo-pair A1\_15–B1\_15 were marked as  $\omega = +15^\circ$ . Images of this stereo-pair were captured in exactly the same



position as the stereo-pair A1–B1 (Figure 2) but they were rotated up by  $15^\circ$  so the zenith angle was  $z = 75^\circ$ . Stereo-pairs (A1–B1, A1\_15–B1\_15, A2–B2, A3–B3) were captured approximately 8 m from the wall so their approximate scale was  $M = 1:124$ .



**Figure 3.** Images of one stereo-pair of the south wall of Padise Abbey: (a) Left image A1; (b) Right image B1.



**Figure 4.** Details of the frame: (a) one of the four fiducial marks; (b) camera constant in mm written on the frame; (c) mark “A” indicates that the image is the left image of a stereo-pair; (d) mark “ $\omega = +15^\circ$ ” indicates that the axis of the lens is rotated by  $15^\circ$  upwards.

### 3. Re-Processing of Original Images

To compare different conditions of the south wall in 1991 to the current condition, we have digitized and re-processed glass plate negatives of the original photogrammetric survey and created results such as a point cloud and an orthophoto. Those results were possible to compare with results from the new photogrammetric survey carried out in 2017.

#### 3.1. Data

All 12 images of photogrammetric survey carried out in 1991 were available for digitizing and re-processing. Most of the glass plates evinced some kind of mechanical decay. Most common decays were scratches and cracks. Overall, the negatives were assessed as they were in relatively good condition without any major imperfections.

For creating a complete picture of the original survey, it was necessary to acquire every available documentation. Despite the glass plate negatives, we had the sketch of the scene with six marked points on the south wall and measured distances between the marked points. We also had a report with measured distances between each image of each stereo-pair. For re-processing and following an accuracy check, we wanted to use all information about the original survey.

#### 3.2. Digitization

The glass plate negatives were digitized on a flatbed scanner Epson Perfection V700 Photo. The glass plate negatives were scanned in 1200 dpi and 2400 dpi scanning resolution. After the comparison between the different scanning resolutions, we decided to work with 2400 dpi scanning resolution. We assumed that scanning under 1200 dpi scanning resolution was not sufficient because geometric value of the negatives was assessed as higher. Images scanned under 2400 dpi were demanding in terms of data size. Even these data were not easy to operate during pre-processing due to the size and scanning glass plate negatives under a higher scanning resolution than 2400 dpi would cause difficulties. Additionally, scanning under too high of a scanning resolution does not necessary increase geometric value of the digitized images. After digitization, knowing approximate image scale  $M = 1:124$ , we were able to calculate a ground sample distance (GSD), so 1 pixel in-image is equivalent to approximate 1.3 mm on the wall.

#### 3.3. Pre-Processing

The digitized images had to be transformed to achieve the same position of the origin of the image coordinate system for all images. Fortunately, we had available a report from the original survey with calibrated coordinates of each fiducial mark. The coordinates were used as the destination coordinates in affine transformation. For transformation, affine transformation was considered as sufficient according to [21].

After the transformation, the images had to be cropped by rectangular region to obtain the same size in pixels for each image. Transformation and cropping were carried out in software Autodesk Raster Design. The final resolution of the images was  $11,717 \times 16,441$  pixels.

#### 3.4. Processing

For re-processing of the digitized glass plate negatives, we used structure from motion method and as a software tool, we chose commercial software Agisoft Metashape [22] in version 1.6.1. There are some freeware and opensource alternatives, e.g., MicMac [23]. By [24], it was shown that MicMac is suitable for processing digitized analogue photogrammetric images. There is also presented a comparison of the results between Agisoft Photoscan (predecessor of Agisoft Metashape) and MicMac. Based on the conclusions of the comparison, we chose Agisoft Metashape for our re-processing. The used workflow of processing images in Agisoft Metashape consisted of these steps:

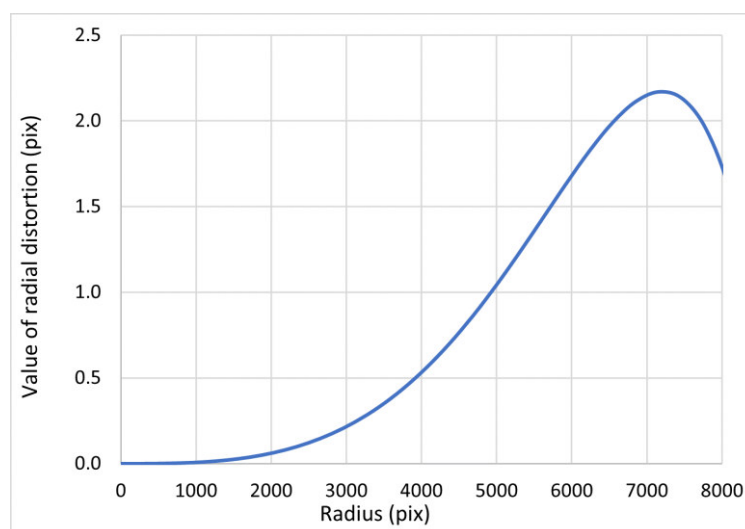
- Relative orientation
- Scaling
- Refinement of orientation
- Point cloud generation
- Textured 3D model generation
- Orthophoto generation

At first, after uploading images to the software, it was necessary to insert focal length of the images and pixel size in mm in the image. In the case of digital images, this information is ordinarily stored in an EXIF file and the software reads the values automatically. The digitized images lack this information. The camera constant was written on the frame of the image and the physical size of 1 pixel in the image in millimeters, we were able to calculate from the known scanning resolution. We did not want to have the frame in the processing, so to cover the frame, a mask was created and applied on every image.

The relative orientation was carried out by function “Align Photos” in Agisoft Metashape. During this process, the algorithm finds corresponding tie points between the images. Based on the tie points, the relative position and rotation of the images is found. The root-mean-square (RMS) error of reprojection of the tie points of the relative orientation was 0.9 pix. After the relative orientation, the signalized markers on the wall were found in the images. The relative orientation was scaled using the longest measured distance on the wall. The rest of measured distances were left as check distances for an accuracy check. Then refinement of relative orientation was carried out by function “Optimize Cameras”.

After these steps, the parameters of interior orientation were calculated by self-calibration. The considered parameters of the interior orientation were camera constant, coordinates of the principal point, and coefficients of the radial distortion,  $k_1$ ,  $k_2$ , and  $k_3$ . The maximum of the radial distortion (Figure 5) is around 2 pix, which confirms the fact mentioned in chapter 1.2 that lenses of UMK cameras were designed to have low distortion. Unfortunately, due to the missing original calibration report, we could not compare all the computed parameters of the interior orientation to the calibrated parameters. The resulted parameters are presented in Table 1.

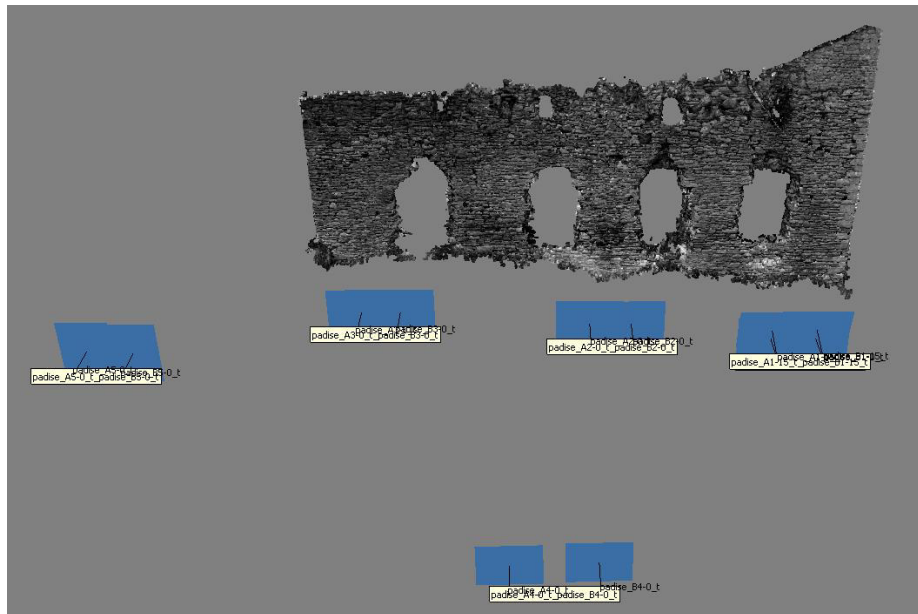
Using “Build Dense Cloud” function, a colored point cloud was generated. For point cloud generation, parameter “Medium” was used, so each size of the images was divided by 4. These down-sampled images were enough to create a sufficiently dense point cloud (Figure 6) for the next steps.



**Figure 5.** Radial distortion for camera used in re-processing determined by self-calibration in Agisoft Metashape.

**Table 1.** Resulted parameters of interior orientation (camera constant  $c$  and position of principal point  $pp_x$  and  $pp_y$ ) by self-calibration. The calculated values are compared to the input values.

Parameter	Input Value	Calculated Value	Difference
$c$ (mm)	64.41	64.22	0.19
$pp_x$ (mm)	Unknown	0.047	x
$pp_y$ (mm)	Unknown	0.052	x



**Figure 6.** Resulted point cloud of the south wall created by re-processing with relative position of captured images.

As a next step, a textured 3D model and orthophoto with resolution 0.002 m/pix were created. The resolution of the orthophoto was chosen according to the approximate GSD.

### 3.5. Accuracy Assessment

Accuracy checking is an integral part of photogrammetric processing. We have used all the available information that we found to provide a proper accuracy check. In the original survey, four distances were measured for scaling. We used the longest one for scaling and adjustment of our relative orientation. The remaining three distances we used as check distances and we calculated them in our photogrammetric model. The differences between measured value and calculated value are presented in Table 2. On check distances M1–M2 and M2–M3, there is a difference 0.005 m. This denotes a proper scaling of relative orientation. Check distance M5–M6 has a difference around 0.02 m. This distance is visible only on stereo-pair A5–B5, and markers M5 and M6 are placed out of the south wall.

**Table 2.** Differences between measured distances and calculated distances from photogrammetric model.

		Measured d (m)	Calculated d (m)	Difference (m)
Control distance	M3–M4	8.955	8.955	0.000
	M1–M2	7.172	7.177	0.005
Check distances	M2–M3	8.780	8.785	0.005
	M5–M6	7.780	7.758	0.022

In the original processing, distances between each image of each stereo-pair were measured. We used this information for an accuracy check. We calculated the distances from computed exterior orientation of images and compared them to the measured values from the original survey (see Table 3). The difference between images of stereo-pair A4–B4 is the largest (0.021 m). This stereo-pair captures the south wall from a farther distance (approx. 19 m farther) and the tie points are not well-distributed in the images, which could cause worse accuracy. The remaining differences do not exceed 0.007 m.

**Table 3.** Differences between measured distances between images of each stereo-pair and distances calculated in photogrammetric model.

Stereo-Pair	Measured d (m)	Calculated d (m)	Difference (m)
A1–B1	2.060	2.066	0.006
A1_15–B1_15	2.060	2.067	0.007
A2–B2	1.849	1.852	0.003
A3–B3	1.676	1.678	0.002
A4–B4	3.444	3.423	0.021
A5–B5	1.923	1.919	0.004

All the presented distances had direction more or less parallel to the X axis. The results show the accuracy of scaling the photogrammetric model in one axis. To have a proper accuracy check, we should check the distances in different directions (Y axis and Z axis). Since we did not have measured distances from the original measurement in different directions, it was not possible.

#### 4. New Survey and Processing

Since the survey in 1991, Padise Abbey has been partly renovated. The renovation of the south wall of the abbey, in which we were interested, is already in an advanced stage and the state of the wall is very different on the first sight. After deeper inspection, it was obvious that during the renovation, original stones of the wall were preserved. It was even possible to find the carved markers (M1–M4) from the original survey. In 2017, we carried out a new photogrammetric survey. The new images were processed using structure from motion method and point cloud and orthophoto was created to have the possibility to compare the different state and condition of south wall of the abbey.

For capturing the images, we used a digital non-metric camera Nikon d610 with Nikon AF-S DX Nikkor 35 mm f/1.8G lens and we captured 30 images with resolution  $6016 \times 4016$  pixels from different positions at different angles. The images were captured 18–39 m from the object with average GSD 4.8 mm/pix.

Structure from motion method was used to process the images and Agisoft Metashape was chosen as a software for SfM processing. After computation of relative orientation, the model was georeferenced to have the same coordinate system as we had in the photogrammetric model made by re-processing. As a control and check points, we used the markers (M1–M4) from the original survey, which were visible and identifiable also in new digital images. We also used significant corners of the stones (P1–P6) of the wall which were possible to identify in images captured in 1991 and in the new images captured in 2017 (see, Figure 7). The coordinates of the markers and significant points we measured in the photogrammetric model made by re-processing. In total, four control points equally distributed in-model were used for georeferencing.

Then, refinement of the orientation was carried out. After the refinement, we measured six check points in images, and we compared their calculated coordinates to the coordinates calculated in the photogrammetric model of re-processing. In Table 4, there are presented differences in coordinates, Euclidean distance, and root mean square errors (RMSE) for each axis. The largest differences are in Y axis, which is due to the configuration of images. RMSE in X and Z axis (plane XZ is approximately parallel to the south wall) are under 0.01 m. In our case, the RMSE values represent only the accuracy of georeferencing of the new survey model to the model of the re-processing. The accuracy of the



model of the new survey itself is represented by reprojection error in pixels on control and check points (Table 4). None of the reprojection error exceed 1 pixel. RMS reprojection error of the tie points was 0.4 pix.

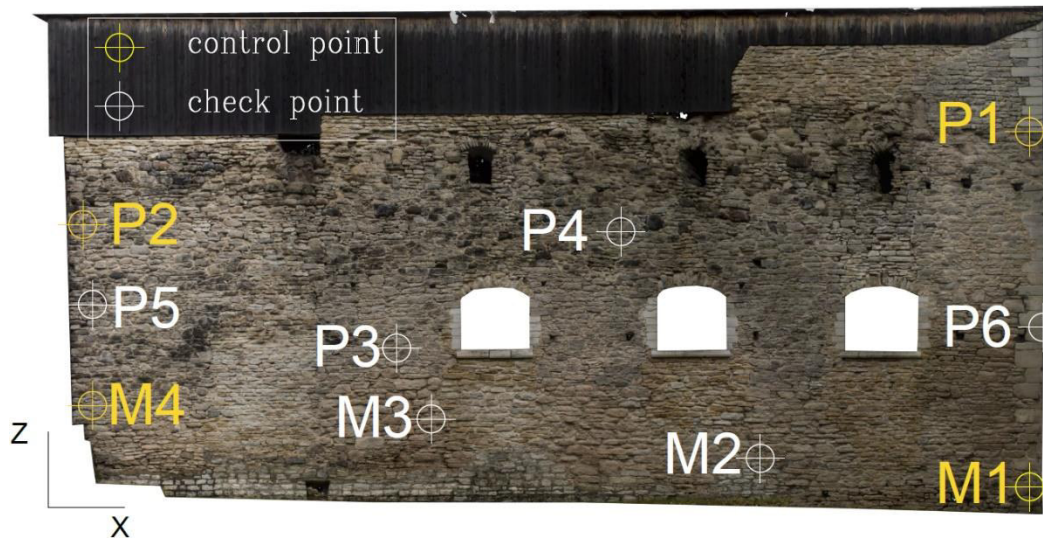


Figure 7. Distribution of control points (yellow) and check points (white).

Table 4. Differences on control and check points between measured coordinates from re-processing and adjusted coordinates from new survey.

		X (m)	Y (m)	Z (m)	Euq. Dist. (m)	Reprojection Error (pix)
Control points	M1	−0.001	0.001	−0.001	0.002	0.8
	M4	0.003	−0.003	−0.001	0.004	0.3
	P1	−0.001	−0.001	0.001	0.002	0.4
	P2	−0.002	0.003	0.002	0.004	0.3
	RMSE	0.002	0.002	0.001	0.003	
Check points	M2	0.003	−0.005	0.005	0.007	0.4
	M3	0.007	−0.028	0.011	0.031	0.3
	P3	0.004	−0.033	0.013	0.036	0.6
	P4	0.002	−0.007	0.009	0.012	0.7
	P5	0.001	−0.010	0.003	0.011	0.4
	P6	−0.006	0.001	0.008	0.010	0.6
	RMSE	0.004	0.019	0.009	0.0209	

After the accuracy check on check points, point cloud, 3D model, and orthophoto with 0.005 m/pix resolution were generated. The resolution of the orthophoto was chosen according to the average GSD (4.8 mm/pix) and according to the fact that RMS reprojection error and reprojection errors on control points and check points do not exceed 1 pixel.

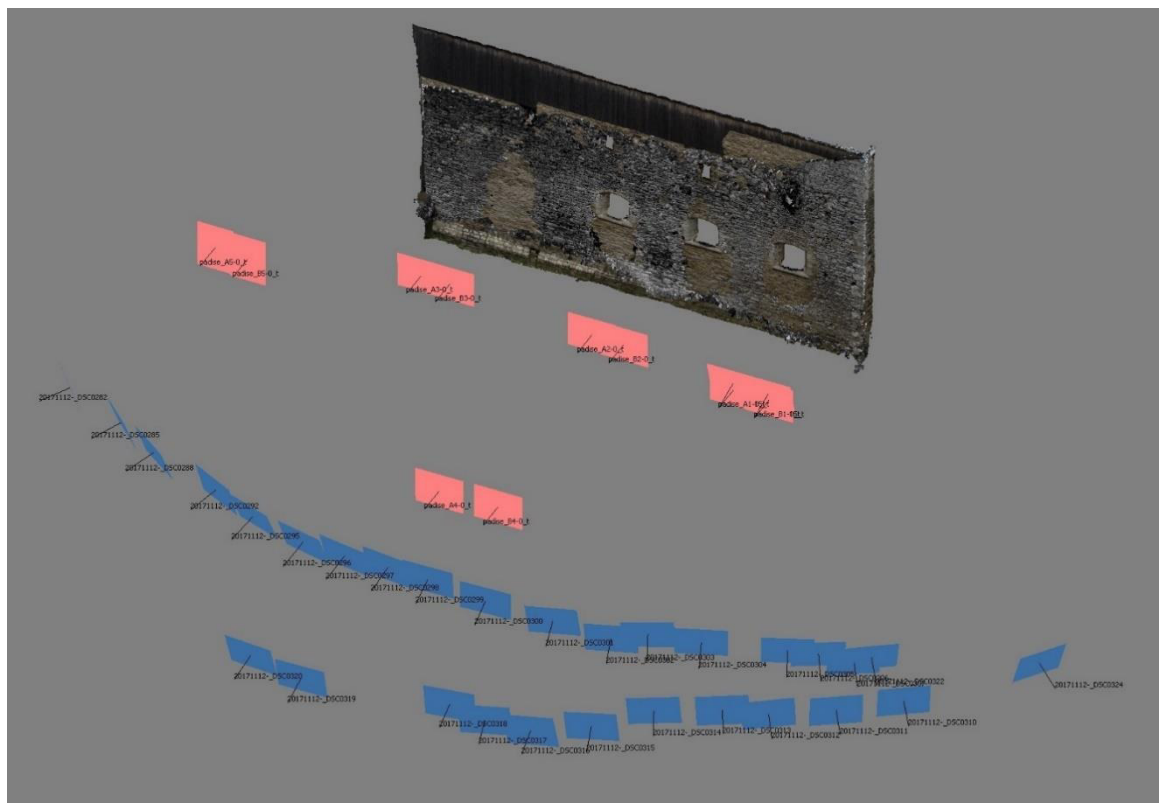
### 5. Comparison

The main and important difference between both processing are the datasets. The original images are digitized analogue images captured by a precisely manufactured metric camera and the images of the new survey are images captured by a non-metric camera, which is not designed for photogrammetric purposes. Due to the high cost of the images, just 12 images stored on glass plates were captured during the original survey. The average overlap on the object was approx. 5 images. During the survey in 2017, we captured 30 images and almost all objects are seeable on every image. During the new survey, we captured the images with larger GSD (avg. 4.8 mm/pix). This GSD is sufficient for determining if a stone was preserved or an outcome of the renovation. For processing

digitized analogue and digital images, we used the same method, and we processed the data in the same software. Even though both datasets were inherently different, we achieved respectable reprojection error (in both cases, RMS reprojection error is under 1 pix.). The comparison of datasets is presented in Table 5. The merged point clouds and difference in capturing scenario is presented in Figure 8.

**Table 5.** Comparison of datasets of the re-processing and the new survey.

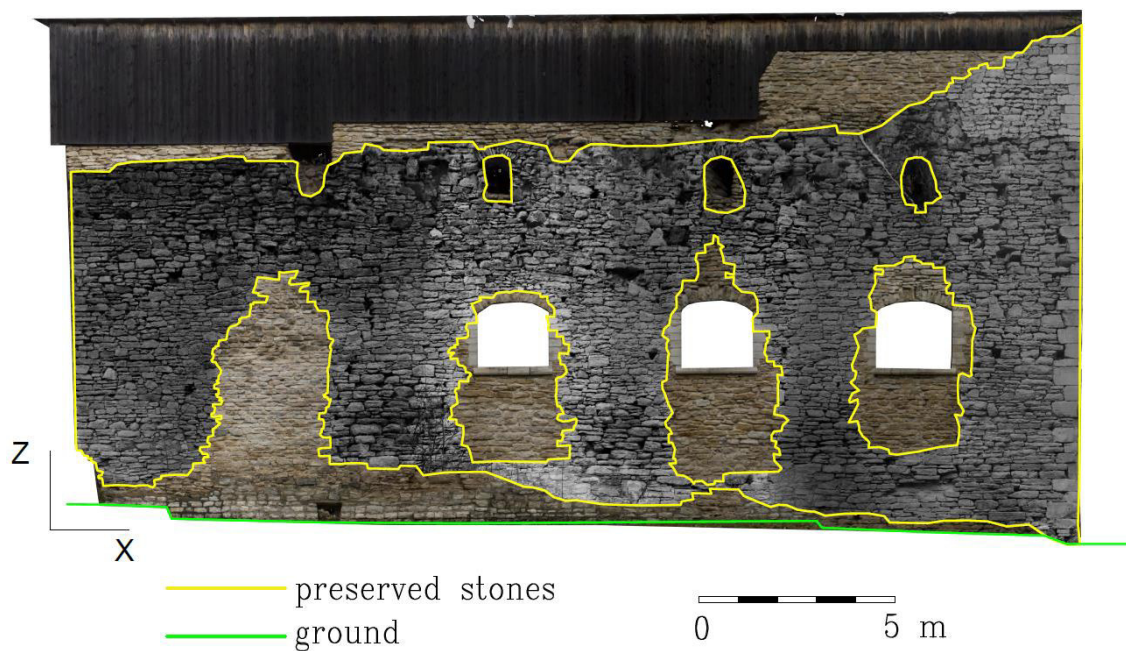
		Re-Processing	New Survey
Images	Num. of images	12	30
	Distance from the object (m)	8–19	18–39
	F 35 mm eqv. (mm)	12.6	35
	Sensor size (mm)	130 × 180	36 × 24
	Resolution (pix)	11,717 × 16,441	6016 × 4016
	Approx. GSD (mm/pix)	1.3	4.8
Processing	Num. of tie points	12,691	21,159
	RMS reprojection error (pix)	0.9	0.4



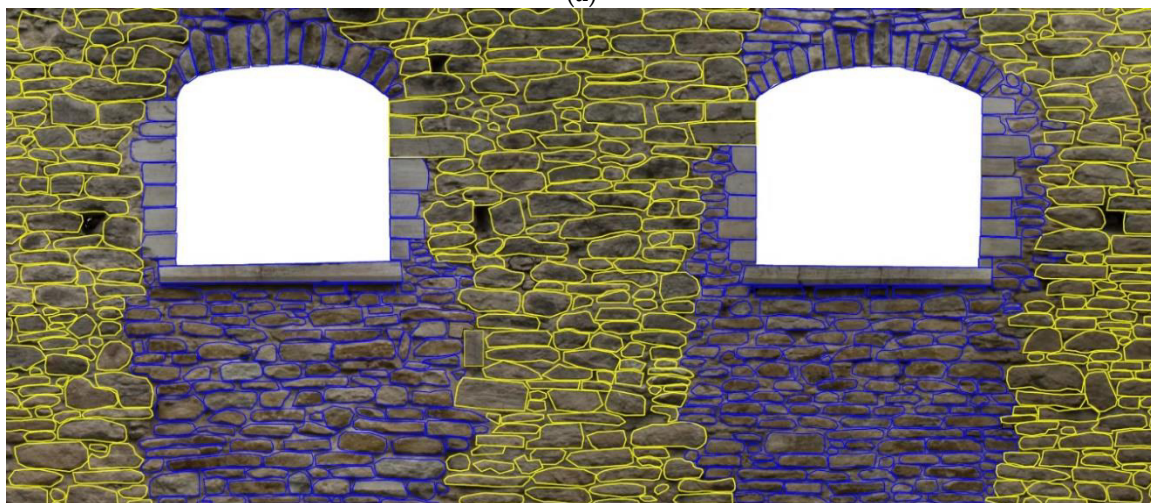
**Figure 8.** Merged point clouds of the re-processing and the new survey. On the figure, it is possible to observe different capturing scenarios (the highlighted images belong to re-processing).

For comparison of different states of Padise Wall in 1991 and 2017, we generated a textured 3D model, point cloud, and orthophoto. The different states of the south wall are possible to examine and explore in these results. The easiest way to compare the state of the wall is orthophoto comparison. The orthophotos were overlaid on each other (Figure 9a) and the overlay was exported to .pdf format. This result is easily transportable, readable, and understandable for architects, conservators, restorators, and even for the wider public. On the overlay, it is easily possible to determine which stones of the wall are preserved after the renovation and which stones are the outcome of the renovation. From the

overlay, it is possible to create manually a drawing of stones of the wall with semantic information (Figure 9b). The result could be a data for GIS and BIM systems [25].



(a)



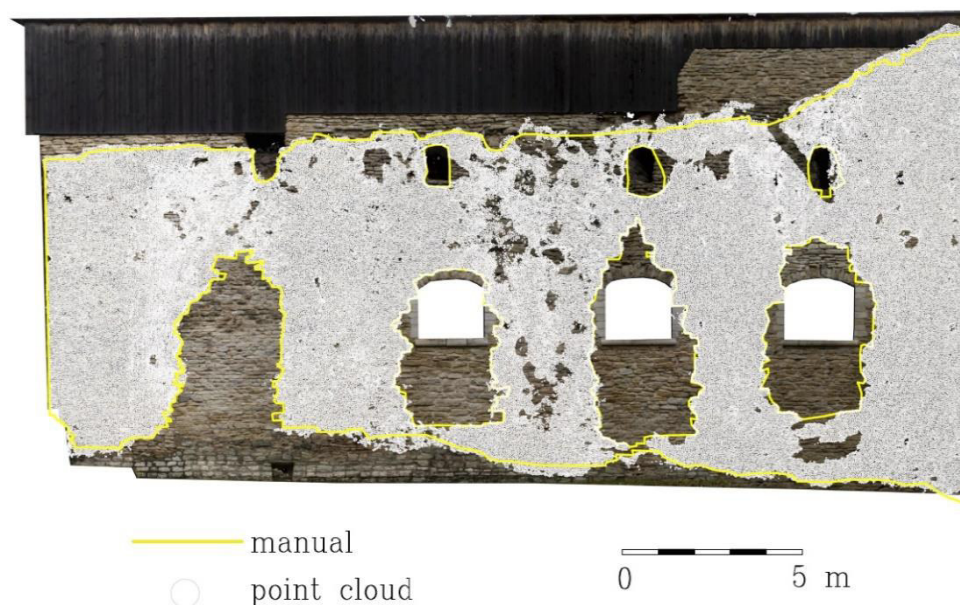
(b)

**Figure 9.** Comparison between states in 1991 and 2017 of the south wall of Padise Abbey: (a) an overlay of cropped orthophoto generated by re-processing (grey colors) and by new survey; (b) a drawing of stones. Yellow color indicates preserved stones, blue color indicates new stones after renovation.

On the other hand, point cloud is not a result intended for general public. Even though tools like plas.io, HexaLab, Potree, and Sketchfab bring the point clouds closer to the wider audience, it is necessary to have certain knowledge and tools to operate and interpret point clouds. Despite that fact, point cloud can offer an interesting comparison with the potential of gathering the semantic data automatically and semi-automatically. The point clouds generated by re-processing and by new survey were compared to each other by Cloud/Cloud distance method in CloudCompare [26] software. According to [27], both point clouds had to be pre-processed before Cloud/Cloud comparison. Both point clouds were filtered using statistical outlier removal function (SOR filter in CloudCompare).



After the filtering, the point cloud of the new survey had a resolution of 1 point per ca. 2 cm<sup>2</sup> and the point cloud from re-processing had 1 point per ca. 0.6 cm<sup>2</sup>. To use the point cloud of the new survey as the reference, we had to subsample the point cloud of the re-processing to have lower resolution (final resolution was 1 point per 3 cm<sup>2</sup>). From the result of comparison, it is possible to easily examine the changes between the states in 1991 and 2017, not only by visible matter but also by determination of differences. The comparison showed that 76.2% of points have a lower difference than 0.05 m. In respect to the accuracy of georeferencing of the model of the new survey, we considered these points as points belonging to the preserved stones. The rest of the points were considered as points belonging to the outcome of the renovation. The points with a difference under the threshold 0.05 m were overlaid to the orthophoto and the position of points were compared to the manually created boundary from the orthophoto (see Figure 10).



**Figure 10.** Overlay of the point cloud where differences of Cloud/Cloud comparison were under chosen threshold of 0.05 m and of the orthophoto with the manually created boundary of preserved stones.

The final chosen resolution of the reference point cloud (new survey point cloud) after filtering can have an impact on the result of the Cloud/Cloud comparison. To avoid that, we also created a mesh from the point cloud of the new survey. The mesh was used as reference in Cloud/Mesh comparison with point cloud of re-processing. A figure of 76.5% of the differences were under the chosen threshold of 0.05 m and we considered them also as points belonging to the preserved stones. The overlay of the points with orthophoto and manually created boundary were carried out.

By eye, both overlays seem identical (for that reason, we present only overlay with Cloud/Cloud comparison in Figure 10). Both match with the boundary, which verifies the 0.05 m threshold (especially around the windows). The percentage of points under the threshold is very similar so we observed that even Cloud/Cloud comparison was relevant.

For distinction between preserved and new stones, it was sufficient to use presented method. The topic of finding the semantic information from multitemporal data and Cloud/Cloud and Cloud/Mesh comparison encourages future development of advanced semi-automatic methods. By [28] were presented methods of change detection on depth maps of orthophotos generated based on the images captured in different time stages. One of the methods was comparing histograms of the depth maps and another method was based on subtracting of depth maps of orthophotos between each other. Methods for structural changes on cultural heritage objects were presented by [29]. The point cloud of the cultural heritage object was by automatic methods enriched by the information about the difference

to the laser scanning point cloud, difference to the plane generated by RANSAC algorithm, luminance, curvature, and roughness.

From the point clouds of the re-processing and the new survey, textured 3D models were created. The textured 3D model is possible to upload to one of the internet services or export to the .pdf format and present the results to the wider public. The main importance of textured 3D models we see with the development and expansion of virtual reality (VR) technologies [30].

## 6. Conclusions

Due to the photogrammetric surveys carried out in different time stages, it is possible to create 2D and 3D results which help to study, examine, and explore different states and different conditions of cultural heritage objects. Those multitemporal results are comparable and could be valuable for architects, restorators, and conservators.

Capturing of the images for photogrammetric processing during analogue era was a highly sophisticated work with a high-quality metric camera. Every capture of an image must be precisely planned. The number of the images was a topic. The storing medium was expensive, and the photogrammetric processing took much longer with higher number of images. On the other side, nowadays, using a (sometimes low-cost) non-metric digital camera, it is possible to capture many images without any additional cost. The preparation is thinner. Processing of the images is much faster using structure from motion method and it is possible to gain a variety of two-dimensional and three-dimensional results. Using modern structure from motion method, it is also possible to re-process digitized analogue images. Although the datasets are inherently different, to achieve respectable accuracy, the pipeline of processing could be similar for digital and analogue datasets. Even though it is possible to process the digitized analogue data using modern methods, which were designed originally by computer vision science, for analogue data restoration, digitization, and successful processing, it is essential to keep the knowledge about fundamentals of the analogue photogrammetry.

In 1991, K. Pärtna carried out photogrammetric survey of the south wall of Padise Abbey. Using an analogue metric camera UMK 6.5/1318, he and his team captured 12 images in six stereo-pairs. After more than 25 years, the glass plate negatives and partial documentation were found, and the negatives were digitized and re-processed using SfM method. The relative orientation was computed, scaled, adjusted, and the accuracy was checked on measured distances between markers and measured distances between each image of each stereo-pair. Then, the textured 3D model, point cloud, and orthophoto were generated.

The new photogrammetric survey with a digital camera was carried out in 2017 after the renovation of the south wall of Padise Abbey. The images were processed using SfM method as well. The model was georeferenced to the model made by re-processing to obtain the same coordinate system. As control points and check points, we used markers carved into stones in 1991 and significant features of stones of the south wall. The RMSE of differences in coordinates in X axis was 0.004 m and in Z axis 0.009 m (XZ plane is parallel to the wall). The results of the processing were also a textured 3D model, point cloud, and orthophoto.

The different states of Padise Wall between 1991 and 2017 are possible to explore and present on comparison of results made by re-processing of images and new survey. In our opinion, an overlay of two orthophotos offers the easiest comparison of different states. The overlay is an easily understandable result, could be exported to the widely used formats, could be a source for CAD drawing, and could be a data for GIS and BIM systems. On the other hand, operating point cloud requires some knowledge and proper software. As was shown, using Cloud/Cloud and Cloud/Mesh comparison, it is possible to determine differences concerning which stones of the wall are preserved, and which stones of the wall are an outcome of renovation. Differences under chosen threshold (0.05 m) were considered as points belonging to the preserved stones and an overlay with orthophoto was carried out. The overlay showed that the points fit to the manually created boundary. The semi-automatic or automatic gathering of semantic data from Cloud/Cloud or Cloud/Mesh comparison is going to be a topic of the future

research. From the point clouds were created textured 3D models. The textured 3D models are going to be a more and more important result with development of VR technologies.

The created results can be supported by other modern survey measurement methods like laser scanning, GNSS and total station measurement, ground penetrating radar measurements, etc. Additionally, further testing and comparing different open software tools can be additional work. All the methods offer different results with different backgrounds. By merging them and comparing them, we can get a variety of 2D and 3D results. From those results, it is possible to derive manually, semi-automatically, or automatically important semantic information, which can significantly help to better understand the cultural heritage objects.

**Author Contributions:** Conceptualization, K.V. and J.P.; methodology, A.D. and A.U.; technical advisory and supervision, K.V. and J.P.; data processing, A.D. and A.U.; investigation, K.V. and J.P.; writing—original draft preparation, A.D. and A.U.; writing—review and editing, A.D. and A.U.; visualization, A.D.; project administration, A.U. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project was funded by Estonian Ministry of Culture (Project no. 7-8/79) and by a grant of the Grant Agency of the Czech Technical University in Prague, grant No. SGS20/053/OHK1/1T/11.

**Acknowledgments:** This research is partly research of EuroGV spol., s r. o., Archaeovision LLC, Estonian Academy of Arts and Department of Geomatics at the Czech Technical University in Prague.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Yastikli, N. Documentation of cultural heritage using digital photogrammetry and laser scanning. *J. Cult. Herit.* **2007**, *8*, 423–427. [\[CrossRef\]](#)
2. Barrile, V.; Bilotta, G.; Lamari, D.; Meduri, G.M. Comparison between techniques for generating 3D models of cultural heritage. *Recent Adv. Mech. Mechatron. Civ. Chem. Ind. Eng. Math. Comput. Sci. Eng. Ser.* **2015**, *49*, 140–145.
3. Giuliano, M.G. Cultural Heritage: An example of graphical documentation with automated photogrammetric systems. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2014**, *XL-5*, 251–255. [\[CrossRef\]](#)
4. Ippoliti, E.; Meschini, A.; Sicuranza, F. Structure from motion systems for architectural heritage. A survey of the internal loggia courtyard of Palazzo del Capitani Ascoli Piceno, Italy. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2015**, *XL-5/W4*, 53–60. [\[CrossRef\]](#)
5. Faltýnová, M.; Raeva, P.; Poloprutský, Z.; Matoušková, E.; Housarová, E. Complex analysis and documentation of historical buildings using new geomatics methods. *Civ. Eng. J.* **2016**, *4*, 1–8.
6. Kraus, K. *Photogrammetry*; Dümmler: Bonn, Germany, 1993; Volume 2.
7. Albertz, J.; Wiedemann, A. *From Analogue to Digital Close-Range Photogrammetry*; First Turkish-German Joint Geodetic Days; Istanbul Technical University: Istanbul, Turkey, 1995; pp. 245–253.
8. Hartley, R.I.; Mundy, J.L. Relationship between photogrammetry and computer vision. In Proceedings of the SPIE 1944, Integrating Photogrammetry Techniques with Scene Analysis and Machine Vision, Orlando, FL, USA, 11–16 April 1993; Volume 1944, pp. 92–105.
9. Grussenmeyer, P.; Yasmine, J. Photogrammetry for the preparation of archaeological excavation. A 3D restitution according to modern and archive images of Beaufort Castle landscape (Lebanon). In Proceedings of the XXth Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS), Istanbul, Turkey, 12–23 July 2004; pp. 809–814.
10. Maiwald, F.; Vietze, T.; Schneider, D.; Henze, F.; Münster, S.; Niebling, F. Photogrammetric analysis of historical image repositories for virtual reconstruction in the field of digital humanities. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2017**, *XLII-2/W3*, 447–452. [\[CrossRef\]](#)
11. Pavelka, K.; Šedina, J.; Raeva, P.; Hůlková, M. Modern processing capabilities of analog data from documentation of the Great Omayyad Mosque in Aleppo, Syria, damaged in civil war. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2017**, *XLII-2/W5*, 561–565. [\[CrossRef\]](#)
12. Housarová, E.; Pavelka, K.; Šedina, J. Study of Erbil Al-Qala citadel time changes by comparison of historical and contemporary image data. *Eur. J. Remote Sens.* **2019**, *52*, 202–208. [\[CrossRef\]](#)

13. Padise Muuseumi Sihtasutus. Available online: <http://www.padiseklooster.ee> (accessed on 15 May 2018).
14. Visit Harju. Available online: <http://www.visitharju.ee> (accessed on 15 May 2018).
15. Digital Archive, Estonian Churches. Available online: <http://kirikud.muinas.ee/?page=1&subpage=1681b> (accessed on 15 May 2018).
16. Kultuurimälestiste Riiklik Register. Padise Kloostri Varemend Vallikraaviga. Available online: <https://register.muinas.ee/public.php?menuID=monument&action=view&id=2921> (accessed on 15 May 2018).
17. Padise Virtual Tour. Available online: <http://www.loodeeesti.ee/padise/padiseklooster/padisekloosterh.html> (accessed on 15 May 2018).
18. Estonia 360, Padise Abbey Aeropanoram. Available online: [http://www.estonia360.ee/padise\\_klooster/](http://www.estonia360.ee/padise_klooster/) (accessed on 15 May 2018).
19. Eesti mõisaportaál. Padise klooster ja mõis. Available online: <http://www.mois.ee/harju/padise.shtml> (accessed on 15 May 2018).
20. Luhmann, T.; Robson, S.; Kyle, S.; Harley, I. *Close Range Photogrammetry*; Principles, techniques and applications; Whittles Publishing: Scotland, UK, 2006; Chap. 3.3.1.3; pp. 140–141.
21. Poloprutský, Z.; Fraštia, M.; Marčíš, M. 3D digital reconstruction based on archived terrestrial photographs from metric cameras. *Acta Polytechnica* **2019**, *59.4*, 384–398. [[CrossRef](#)]
22. Agisoft Metashape User Manual. Available online: [https://www.agisoft.com/pdf/metashape-pro\\_1\\_5\\_en.pdf](https://www.agisoft.com/pdf/metashape-pro_1_5_en.pdf) (accessed on 18 October 2020).
23. MicMac—A Free Open-Source Photogrammetric Suite. Available online: <https://micmac.engg.eu/index.php/Accueil> (accessed on 3 November 2020).
24. Dlesk, A.; Vach, K.; Pavelka, K., Jr. Structure from motion processing of analogue images captured by Rollei metric camera, digitized with various scanning resolution. *Acta Polytechnica* **2020**, *60.4*, 288–302. [[CrossRef](#)]
25. Reina Ortiz, M.; Yang, C.; Weigert, A.; Dhanda, A.; Min, A.; Gyi, M.; Su, S.; Fai, S.; Santana Quintero, M. Integrating heterogeneous datasets in HBIM of decorated surfaces. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2019**, *XLII-2/W15*, 981–988. [[CrossRef](#)]
26. CloudCompare [GPL software]. Available online: <http://www.cloudcompare.org/> (accessed on 3 November 2020).
27. Antova, G. Application of areal change detection methods using point clouds data. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *221*, 1–8. [[CrossRef](#)]
28. Peteler, F.; Gattet, E.; Bromblet, P.; Guillon, O.; Vallet, J.M.; De Luca, L. Analyzing the evolution of deterioration patterns: A first step of an image-based approach for comparing multitemporal data sets. *Digit. Herit.* **2015**, *2*, 113–116.
29. Nespeca, R.; De Luca, L. Analysis, thematic maps and data mining from point cloud to ontology for software development. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2016**, *XLI-B5*, 347–354. [[CrossRef](#)]
30. Pavelka, K., Jr.; Raeva, P. Virtual museums—The future of historical monuments documentation and visualization. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2019**, *XLII-2/W15*, 903–908. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

# Article 3



## Article

# Transformations in the Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images

Adam Dlesk<sup>1,\*</sup>, Karel Vach<sup>2</sup> and Karel Pavelka<sup>1</sup> 

<sup>1</sup> Department of Geomatics, Faculty of Civil Engineering, Czech Technical University in Prague, 16629 Prague, Czech Republic; pavelka@fsv.cvut.cz

<sup>2</sup> EuroGV, Spol. s. r. o., 11000 Prague, Czech Republic; vach@eurogv.cz

\* Correspondence: adam.dlesk@fsv.cvut.cz

**Abstract:** The photogrammetric processing of thermal infrared (TIR) images deals with several difficulties. TIR images ordinarily have low-resolution and the contrast of the images is very low. These factors strongly complicate the photogrammetric processing, especially when a modern structure from motion method is used. These factors can be avoided by a certain co-processing method of TIR and RGB images. Two of the solutions of co-processing were suggested by the authors and are presented in this article. Each solution requires a different type of transformation—plane transformation and spatial transformation. Both types of transformations are discussed in this paper. On the experiments that were performed, there are presented requirements, advantages, disadvantages, and results of the transformations. Both methods are evaluated mainly in terms of accuracy. The transformations are presented on suggested methods, but they can be easily applied to different kinds of methods of co-processing of TIR and RGB images.

**Keywords:** photogrammetry; thermal infrared image; thermal camera calibration; relative pose estimation; close-range photogrammetry



**Citation:** Dlesk, A.; Vach, K.; Pavelka, K. Transformations in the Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images. *Sensors* **2021**, *21*, 5061. <https://doi.org/10.3390/s21155061>

Academic Editor: Gregorij Kurillo

Received: 5 May 2021  
Accepted: 24 July 2021  
Published: 26 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Photogrammetric processing of close-range thermal infrared (TIR) images deals with several challenging difficulties. The resolution of TIR images is much lower compared to the resolution of images captured by common digital RGB cameras. The field of view (FOV) of thermal infrared cameras is generally much narrower. When the task is to inspect a large object by a TIR camera, usually many images must be captured. The high amount of images greatly complicates the thermal inspection of the object. Easing the inspection would include creating the super-resolution image [1] or creating, e.g., point cloud, a 3D textured model, or the orthophoto generated by photogrammetric processing using the structure from motion (SfM) method.

The SfM processing of TIR images is complicated by several factors. In addition to aforementioned issue of low-resolution of TIR images, the contrast of the TIR images depends on the variance of the thermal radiation of the scene. Where the scene is in a more or less similar temperature environment, e.g., interior of the room, building façade, etc., the TIR image can have very low contrast and the image lacks features. This is a crucial complication for processing using the structure from motion (SfM) method. The algorithms of SfM are unable to find a sufficient amount of key and tie points (see Figure 1a), which significantly complicates the calculation of relative orientation, reduces the accuracy, or even makes the computation impossible.

Another factor that makes photogrammetric processing of TIR images challenging is that the conventional way of photogrammetric targeting is not usable, because the targets are not detectable on TIR images (ordinarily, there is no difference in thermal radiation on targets). These factors complicate the scaling or georeferencing of the results of the SfM processing.



**Figure 1.** TIR and RGB image of the same scene: (a) feature matching algorithm failed to find any key point or tie point; (b) feature matching algorithm found sufficient amount of tie points (blue dots).

All mentioned factors must be taken into account. To eliminate these negative factors, several methods of photogrammetric processing of TIR images and RGB images together were presented. The motivation of the co-processing is presented in the example (Figure 1). In this example, it is apparent that the TIR image lacks any features so the SfM algorithm for feature matching failed to find any tie points. On the other hand, the RGB image of the same scene has identifiable features; the algorithm was successful in feature matching, resulting in a sufficient number of tie points on the RGB image. The main idea of the co-processing is to use the features with the quality of the RGB image and the thermal information from the TIR image.

Photogrammetric co-processing of TIR and RGB images was presented in several prior research contributions. The reported methods can be divided into two categories. The first category of methods does not require capturing the corresponding RGB image to every TIR image in the same configuration during the capturing. The scene can be captured by a TIR camera and RGB camera separately, even at different times. The second category requires a fixed system of TIR and RGB cameras and for every TIR image, an RGB image must be captured. In both categories, the results of the photogrammetric processing of TIR images are 2D and 3D models (orthophotos, point clouds, and textured models), which are augmented by information which is derived from the original TIR image. These results can help understand better the thermal behavior of objects of interest or localize the thermal anomalies in the reference coordinate system and determine their scope. This could be efficiently used in the different areas of research.

The following methods belong to the first category. A method presented by [2] consists of the processing of RGB and TIR images separately. Based on the images, the point clouds were obtained from the SfM method. Both of the generated point clouds were co-registered based on known position of the images gathered from GNSS. Then, for following 3D model generation, the author used the geometric quality from the processing of RGB images, and the texture was applied from TIR images. A similar idea was presented by [3]. RGB images were processed by the SfM method and the digital model of surface was generated. Then the group of TIR images was processed by the SfM method as well. When the final orthophoto of the scene was generated, the TIR images were applied on generated DSM (created from RGB images). Reference [4] presented a method where the scene was captured by TIR images and RGB images separately. Each group was then processed separately by the SfM method and the point clouds from each group were generated. Then both point clouds were co-registered by the ICP algorithm and the point cloud from RGB images was used for DSM generation. The texture from TIR images was then applied to

the DSM and the result was the 3D textured model with the thermal infrared information. Reference [5] presented a method of generating a point cloud from TIR images captured by a TIR camera and a point cloud from RGB images captured by a DSLR camera. Images were processed separately using the SfM method; from both groups of images, colored point clouds were generated. Then the point clouds were co-registered using the ICP method. Then every point of the RGB point cloud was augmented by thermal information, which was derived from the nearest point from the TIR point cloud. The nearest points were found by the nearest neighbor search function. The result was point cloud augmented by thermal information.

Several methods belong to the category that requires a fixed system of cameras and capturing corresponding images. Developing methods in this category is supported by the fact that modern handheld industrial thermal infrared cameras (which are suited for close-range inspection of the objects) are equipped with both thermal infrared sensors and digital RGB sensors. The cameras are usually lightweight, easily operable, and have the potential of being used for photogrammetric purposes. Those cameras are not the only option—another way could be to create one's own fixed system of TIR camera and RGB camera. A paper by [6] belongs in this category of methods. Different co-registration methods were presented, including methods of co-registering corresponding images based on features or based on 2D line segments, methods of point cloud co-registering (each point cloud was generated from TIR and RGB images separately), by the iterative closest point (ICP) algorithm, and by corresponding segmented planes. By [7], a PANTIR dual camera setup that consisted of a panchromatic and TIR camera carried on the helicopter was described. Both cameras were almost parallel to each other. The registration between corresponding images was carried out using plane transformation. The author mentions the fact that, in the case of a non-parallel optical axis, the images must then be taken at constant altitude. Moreover, before the fusion of the images, both images must be undistorted. Then the corresponding RGB and TIR images are fused to the PANTIR images and the process of photogrammetric processing using SfM continues. A method presented by [8] was using a modern industrial thermal infrared camera FLIR T1030sc equipped with a TIR sensor and an RGB sensor in a fixed configuration was presented. For every TIR image, an RGB image was captured. Then both sensors were geometrically calibrated in order to undistort the images. The transformation key of the projective transformation was computed and was used for all corresponding image pairs. Then the scene was captured by a FLIR T1030sc camera and by a DSLR RGB camera. The images captured by the DSLR camera were processed using the SfM method together with RGB images captured by a FLIR T1030sc camera. The result of the processing was a DSM. Then the RGB images captured by the FLIR T1030sc camera were replaced by the corresponding undistorted and geometrically corrected TIR images. Only TIR images were used for the final generation of a texture for the 3D model. The final result of the process was an orthophoto augmented by thermal information. By [9], a method of photogrammetric co-processing of TIR images and RGB images was presented. The corresponding images (RGB and TIR images) were co-registered either by plane transformation (concretely by affine transformation) or by 3D spatial transformation. Overall, the important topic of this category of papers is the transformation between the images captured by cameras in a fixed configuration. One approach is to use 2D plane transformation and the second approach is to use 3D spatial transformation. The plane transformation was used by [7], concretely [8] used projective transformation and [9] used affine transformation. Spatial transformation was used by [9] as well. Previously, authors of this article suggested two methods of co-processing the TIR and RGB images, which were captured by the fixed camera system [10]. Both methods use SfM processing. The first method was called Sharpening (according to the pan-sharpening method); for this method, plane transformation was designed. The result of this method is mainly augmented orthophoto. The second method is called Reprojection because it vastly depends on the accurate reprojection process. Our method uses spatial transformation and the result of the method is a point cloud.



The focus of this article is the transformation between the images captured by the fixed camera system. This article describes the process of the transformations, different requirements of transformation, expected accuracy of transformation processes, results, and potential usability for TIR and RGB image co-processing. The paper discusses the plane transformation, which was designed for the Sharpening method, and spatial transformation, which was designed for the Reprojection method. Using the plane transformation requires finding a universal transformation key between the TIR image and the corresponding RGB image. In this paper, an experiment examines if the transformation key is dependent on the scene, and how the scene influences the accuracy of the transformation). In the literature, the images are captured in most of the cases from larger distance, often-times from UAV, e.g., [2–4,6,7,9,11–13]. However, interest from the authors of this article involves a close-range application. Thus, capturing for the experiments was done at close range (from 2 to 11 m, approximately). In those distances, the scene changed a lot during the capturing. Spatial transformation required accurate relative poses and relative orientation between the TIR and RGB cameras mounted on a fixed rig and determination of the internal orientation of the cameras. This paper presents a process of geometric calibration of the cameras, determination of interior orientation parameters, and determination of the relative orientation and pose. Then the calculated parameters were tested on an experiment and the accuracy of the spatial transformation was evaluated.

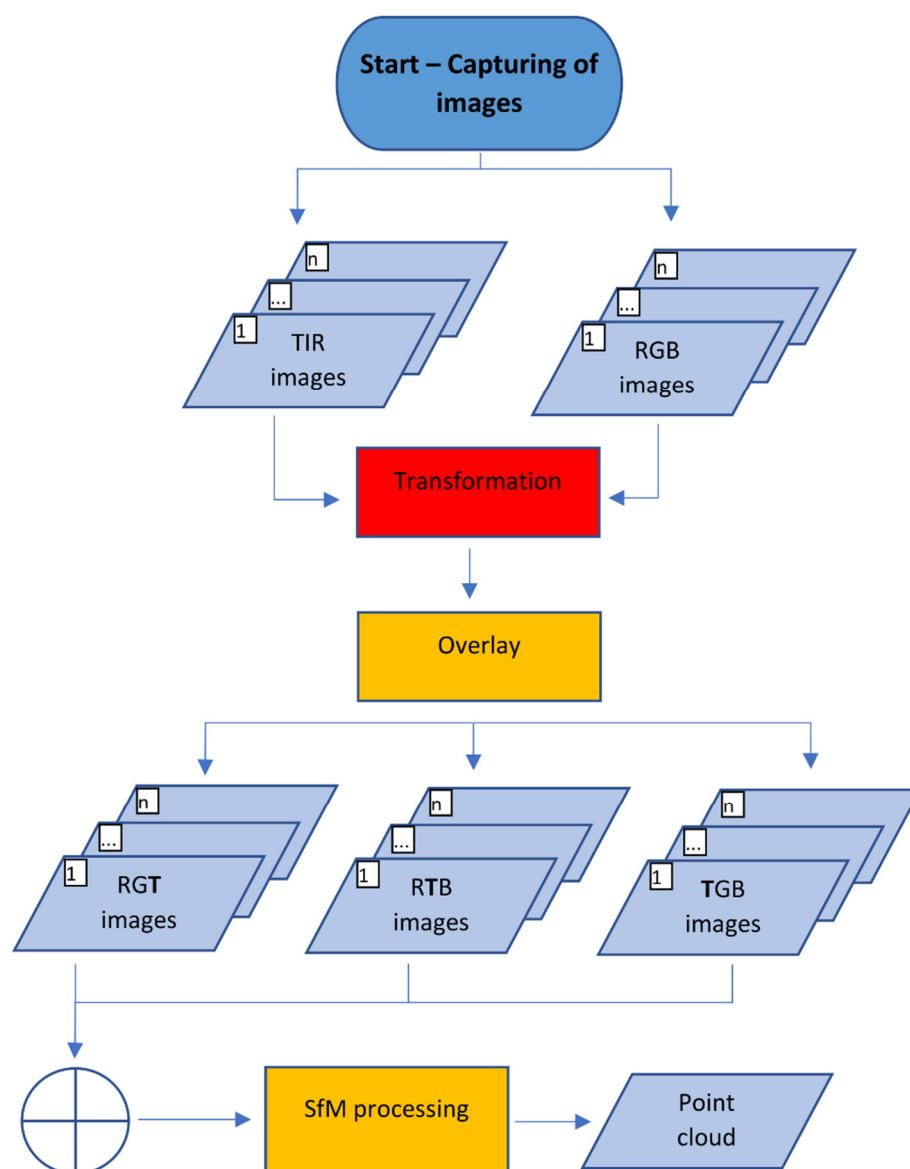
Both suggested methods (Sharpening and Reprojection), which are, for context, briefly presented in this article as well, are in development, and the results of this article are needed for final implementation. Even though the transformations are presented on the two suggested methods, the knowledge and the results of the experiments can be used in a variety of methods of photogrammetric co-processing of TIR and RGB images.

### 1.1. Suggested Methods

Two methods of photogrammetric co-processing of TIR images, together with corresponding RGB images, were suggested by the authors of this article. These methods are called Sharpening and Reprojection. In both suggested methods, there is the presumption of having a fixed relative pose between the TIR camera and RGB camera, and for every TIR image having captured a corresponding RGB image. Both methods were designed for photogrammetric processing by the structure from motion method. Both methods were previously presented by the authors of this article [10]. The methods are still in development and both rely on correct transformation processes. Results of this article are needed to finish the development and implementation.

#### 1.1.1. Sharpening and Plane Transformation

For the Sharpening method (see Figure 2), during the capture of the TIR images, for every TIR image, a corresponding RGB image was captured. The capturing scenario is the same as suggested for processing by the SfM method. Before the processing, the TIR image was transformed by a plane transformation to the RGB image to create an overlay of the images. The same transformation key was then used for all other image pairs. Presumptions for sufficient plane transformations are that both cameras are placed close next to each other, both cameras have more or less parallel optical axes, and images of the captured cameras are undistorted. To create the overlay, one of the channels of the RGB image was skipped and replaced by values from the thermal infrared image, which were recalculated by a function to 8-bit values (0–255) or 16-bit values. By this process (in case of skipping the blue channel of the RGB image), a sharpened false-colored RGT image was created, where visible information from the RGB camera (channels RG) and thermal infrared information from the TIR image (channel T) was partly stored. An RTB or TGB combination could be analogously created, or even 4-channel RGBT images. This process was carried out for every image pair of TIR and RGB images.



**Figure 2.** Diagram of the Sharpening method.

The group of sharpened images was then processed by SfM using the standard workflow, the relative and absolute orientation was calculated and then a point cloud, a 3D model, and an orthophoto were created. The result of this method is orthophoto augmented by thermal infrared information. The orthophoto has false colors (e.g., RGT colors). For every pixel of the orthophoto, the thermal information was calculated back from the T value in the pixel using the inverse previously used function. The thermal information was then stored in the pixel.

For the transformation in the Sharpening method, a plane transformation [14] was designed. Concretely, affine transformation and projective transformation were considered. As stated below, both transformations were tested and gave more or less similar results. Testing of the affine transformation provided better results; that is why it is presented and used for the following experiments.

One way how to determine the coordinate transformation between pixels in an RGB image  $P_{RGB}$  and pixels in a TIR image  $P_{TIR}$  is the plane transformation. In this article, affine transformation is used; see Equation (1).

$$P_{TIR} = \begin{bmatrix} m_X \cos \alpha & -m_Y \sin(\alpha + \beta) \\ m_X \sin \alpha & m_Y \cos(\alpha + \beta) \end{bmatrix} P_{RGB} + T \quad (1)$$

For purposes of the Sharpening method, on the selected image pairs, it is necessary to find a transformation key, respectively, parameters of affine transformation rotation angle  $\alpha$ , shearing angle  $\beta$ , scaling factor  $m_X$ ,  $m_Y$  and translation  $T$  between the images. The same transformation key is then applied to the rest of the image pairs. Presumptions for having sufficient results of the transformation include having both cameras next to each other and having their optical axes parallel, as much as is possible. The accuracy of the transformation is also dependent on the scene, on the images, and the distance to the scene. The testing of the influence of the distance is carried out in this paper. Then the results and sufficiency of affine transformation are discussed.

### 1.1.2. Reprojection and Spatial Transformation

In the Reprojection method (see diagram Figure 3), there is the presumption that, during the capturing scenario, the relative pose of the TIR camera and RGB camera is fixed and a corresponding RGB image is captured for every TIR image. During the capture, the optical state of the cameras must not be changed. Then, only RGB images are photogrammetrically processed by the SfM method and the point cloud is generated. Then every point of the point cloud is reprojected to the automatically, semi-automatically, or manually chosen set of RGB images, where the point of the point cloud is potentially visible. Then it is necessary to carry out “test of visibility” for every image, to check if the point of the point cloud is truly visible on the RGB image. It is possible to carry out the visibility test, for example, using depth and the normal map of the tested RGB image from the processing, using the SfM method. The depth and normal map are rendered, based on the 3D model, which is generated based on the dense point cloud from RGB images. If the test of visibility, based on the corresponding depth map and normal map, proves that the point is visible on the RGB image, the reprojected point on the RGB image is transformed by a spatial transformation to the TIR image. The thermal infrared information can be read from the pixel on the TIR image (where the point was transformed) and the thermal infrared information can be added to the point in the point cloud. Then it is possible to create a colored point cloud with thermal infrared information.

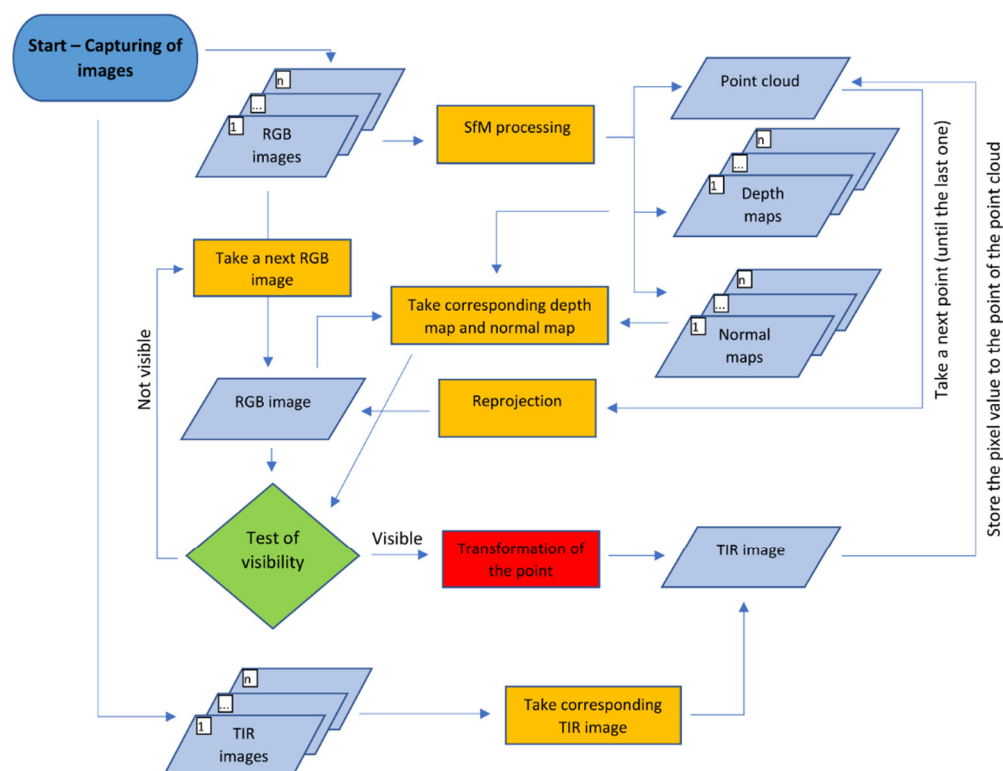


Figure 3. Diagram of the Reprojection method.

As the transformation for the Reprojection method, spatial transformation was designed [14]. Having a point  $P_w$  in world coordinates, and knowing the parameters of the exterior orientation of the RGB image, i.e., rotation matrix  $R_{RGB}$  consisting of rotation angles  $\omega, \varphi, \kappa$ , and translation  $T_{RGB}$  position of the camera in world coordinates, consisting of  $X_{RGB}, Y_{RGB}, Z_{RGB}$ , it is possible to transform the point and get the coordinates of the point in the RGB camera coordinate system  $P_{C_{RGB}}$ .

$$R = R_\omega R_\varphi R_\kappa, \quad T_{RGB} = \begin{bmatrix} X_{RGB} \\ Y_{RGB} \\ Z_{RGB} \end{bmatrix} \quad (2)$$

$$P_{C_{RGB}} = R_{RGB}^{-1}(P_w - T_{RGB}) \quad (3)$$

Knowing the rotation matrix  $R_{RGB \rightarrow TIR}$  and translation  $T_{RGB \rightarrow TIR}$  directly between the RGB camera coordinate system and the TIR camera coordinate system, it is possible to transform the point  $P_{C_{RGB}}$  to the TIR camera coordinate system  $P_{C_{TIR}}$ .

$$R_{RGB \rightarrow TIR} = R_{TIR}^{-1} R_{RGB} \quad (4)$$

$$T_{RGB \rightarrow TIR} = R_{RGB}^{-1}(T_{TIR} - T_{RGB}) \quad (5)$$

$$P_{C_{TIR}} = R_{RGB \rightarrow TIR}(P_{C_{RGB}} - T_{RGB \rightarrow TIR}) \quad (6)$$

Then, from the TIR camera coordinate system, knowing the calibration matrix of the TIR camera  $K_{TIR}$ , it is possible to reproject the point  $P_{C_{TIR}}$  to the thermal infrared image and get the position on the image  $P_{TIR}$ .

$$K_{RGB} = \begin{bmatrix} c & s & p_x + \Delta x \\ 0 & c(1+m) & p_y + \Delta y \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

$$P_{TIR} = K_{TIR} P_{C_{TIR}} \quad (8)$$

Calibration matrix  $K_{TIR}$  consists of camera constant  $c$ ,  $p_x$  and  $p_y$  are the coordinates of the principal point of the camera,  $m$  is the scale difference, and  $s$  is sheer. Parameters  $\Delta x$  and  $\Delta y$  are the calculated values of radial distortion for the concrete pixel.

This process requires the determination of the rotation matrix  $R_{RGB \rightarrow TIR}$  between the RGB and TIR camera coordinate systems, the translation  $T_{RGB \rightarrow TIR}$ , and calibration matrix  $K_{TIR}$ . The rotation matrix  $R_{RGB \rightarrow TIR}$  is calculated by Equation (4) and can then be decomposed to differences in rotation for each axis between the camera coordinate systems  $\Delta\omega, \Delta\varphi$ , and  $\Delta\kappa$ . The translation vector consists of  $\Delta X, \Delta Y$ , and  $\Delta Z$ . The considered parameters of the calibration matrix are  $c, p_x, p_y$ , and coefficients of radial distortion  $k_1$  and  $k_2$ . The determination of those parameters is problematic because it requires a process of geometric calibration of both the TIR camera and RGB camera, and then processing of TIR and RGB images together with highly accurate determination of the exterior orientation of the image pairs. The determination of the relative pose and process of geometric calibration is discussed in this paper in the following experiments.

## 1.2. Handheld Thermal Infrared Cameras

For the capture of close-range thermal infrared images in common industrial applications, it is suitable to use handheld TIR cameras. The handheld thermal infrared cameras are usually light and durable. Modern cameras are also equipped with digital RGB cameras. It is possible to use the cameras separately or for every captured TIR image, the corresponding RGB image. The main reason as to why the TIR cameras are equipped with RGB cameras is to achieve better orientation or better visual impressions during the thermal infrared inspection. The software for inspection of thermal infrared images usually includes the functions of overlaying the TIR images and corresponding RGB images. The

software also has functions of gathering the edges from digital RGB images, projecting the edges to the TIR images for better orientation in TIR images. These features help to better understand the results of TIR imaging and lead to better inspections overall. However, the fact that handheld TIR cameras are equipped with RGB sensors is applicable for purposes of photogrammetry, which helps the suggested methods by authors who require a fixed system of RGB and TIR cameras.

According to Table 1, the resolution of TIR images on handheld cameras starts from  $320 \times 240$  pixels to  $640 \times 480$  pixels. The resolution of digital cameras is according to the public technical specifications (5 megapixels). The 5 MP resolution of RGB cameras is a standard among handheld TIR cameras, but it is below standard compared to the professional digital cameras, which are commonly used in photogrammetry. Nevertheless, the 5 MP resolution of the digital camera promises to achieve better and more accurate results of photogrammetric processing.

**Table 1.** Comparison of parameters of handheld thermal infrared cameras [15,16].

Camera	TIR Camera Resolution (pix)	RGB Camera Resolution (pix)
FLIR E75	$320 \times 240$	5 MP
FLIR E85	$384 \times 288$	5 MP
FLIR E95	$464 \times 348$	5 MP
FLIR E96	$640 \times 480$	5 MP
Fluke TiS60+	$320 \times 240$	5 MP
Fluke Ti480 PRO	$640 \times 480$	5 MP
Fluke TiX501	$640 \times 480$	5 MP

The FLIR E95 thermal infrared camera, which was used for the experiments in this paper, is equipped with an uncooled microbolometer thermal sensor with dimension  $7.888 \times 5.916$  mm and resolution  $464 \times 348$  pix. The pixel pitch of the sensor was  $17 \mu\text{m}$ . The resolution of the TIR image, compared to the other presented handheld TIR cameras, is in the mid-range category. FLIR E95 is possibly equipped with different lenses. The 29 mm lens has the longest focal length and the narrowest field of view (FOV)  $14^\circ \times 10^\circ$ . The 10 mm lens has the widest FOV  $42^\circ \times 32^\circ$ , but the FOV is still narrow compared to the common digital cameras. This factor also complicates the capturing of the TIR images for photogrammetric processing. Only the 10 mm lens was used for the experiments of this paper.

The lenses of the TIR camera of FLIR E95 are focusable with a focus ring. The digital camera of FLIR E95 has a resolution of 5MP, specifically  $2592 \times 1944$  pixels. The focal length of the lens of the RGB camera is 3.29 mm and the field of view is  $53^\circ \times 41^\circ$ , which is wider than the FOV of the 10 mm lens on the TIR camera. The focus of the lens of the digital RGB camera is fixed, according to the manual [15]. A summary of the parameters of the TIR camera and RGB camera on FLIR E95 is presented in Table 2.

**Table 2.** Parameters of TIR camera and RGB camera of FLIR E95.

Parameter	TIR Camera	RGB Camera
Focal length (mm)	10	3.29
Sensor size (mm)	$7.89 \times 5.92$	$3.67 \times 2.74$
Resolution (pix)	$464 \times 348$ pix	$2592 \times 1944$
Size of a pixel (mm)	0.017 mm	0.0014 mm
FOV ( $^\circ$ )	$42^\circ \times 32^\circ$	$53^\circ \times 41^\circ$

## 2. Affine Transformation for Sharpening Method

Testing—using the plane transformation for the Sharpening method—was conducted. For the testing, the FLIR E95 camera was used. FLIR E95 is a TIR camera equipped with a digital RGB camera, and for every TIR image, it is possible to capture a corresponding

RGB image. Both cameras are naturally in a fixed position to each other. The goal of the testing was to examine if there was a universal transformation key that could be applied to all image-pairs of the dataset or if there was a dependency on the distance, which could complicate the process of transformation in the Sharpening method, and how high the influence of the different distance was. Before the testing, some of the presumptions were done. The first presumption was that both cameras are close to each other (the RGB camera is approximately 3 cm above the TIR camera). Both cameras have, by eye, parallel optical axis. Both cameras have fixed focusing during the capturing. Moreover, the images must be undistorted before the transformation process. Pilot projects of geometric calibration show that both cameras are practically undistorted. The process of undistortion is probably carried out by internal processes of the FLIR E95 camera because even the camera itself has overlay functions. The radial distortion of the FLIR E95 camera is described more in Section 3.1.

As a test object, a façade of a family house with a plane structure was chosen. On the façade, cross aluminum paper photogrammetric targets were placed and equally distributed. According to the previously mentioned literature, classic photogrammetric targets are not often used for capturing TIR images. To achieve visibility of the targets, instead of a color difference, there must be a difference in thermal radiation, which is created by using a different emissivity of the two used materials. Instead of a black color, aluminum material was used to achieve dark color on the thermal image. Aluminum foil is a material with a low emissivity (emissivity  $\epsilon = 0.04$  [17]), so instead of emitting the thermal radiation itself, aluminum reflects the thermal radiation of the surrounding ambiance. As a white color, a paper was used on the target, which has high emissivity and mainly emits thermal radiation ( $\epsilon = 0.93$  [17]). In praxis, having a cross aluminum paper target placed on the façade, the center of the target is visible on the TIR image because the aluminum part reflects thermal radiation of the cold of the sky (dark color in our case), and paper (bright color) emits thermal radiation—significantly warmer itself (see Figure 4). Similar targets were used in [5]. However, in some cases, even black and white printed photogrammetric targets could be visible on TIR images. While the paper has high emissivity, printed ink has lower emissivity ( $\epsilon = 0.83$  [18]). The same effect is achieved by having aluminum paper cross targets; it is possible to achieve by ink-paper targets, but with significantly lower emissivity difference.



**Figure 4.** Image captured from 4 m distance. In the image are recognizable cross aluminum paper targets.

After the photogrammetric targeting, 10 stations (from 2 m from the façade to 11 m) were made. Those short distances were chosen because our goal was to focus on close-range TIR photogrammetry. From each station, the TIR image and the corresponding RGB image were captured. During imaging, the focus of the TIR camera was fixed to the 7 m



distance. The RGB camera was fixed focus. All images were captured perpendicular to the façade, so all of the photogrammetric targets in the images were at a more or less same distance to the image. An image example is presented in Figure 4. Approximate sample distances of the captured image pairs are presented in Table 3.

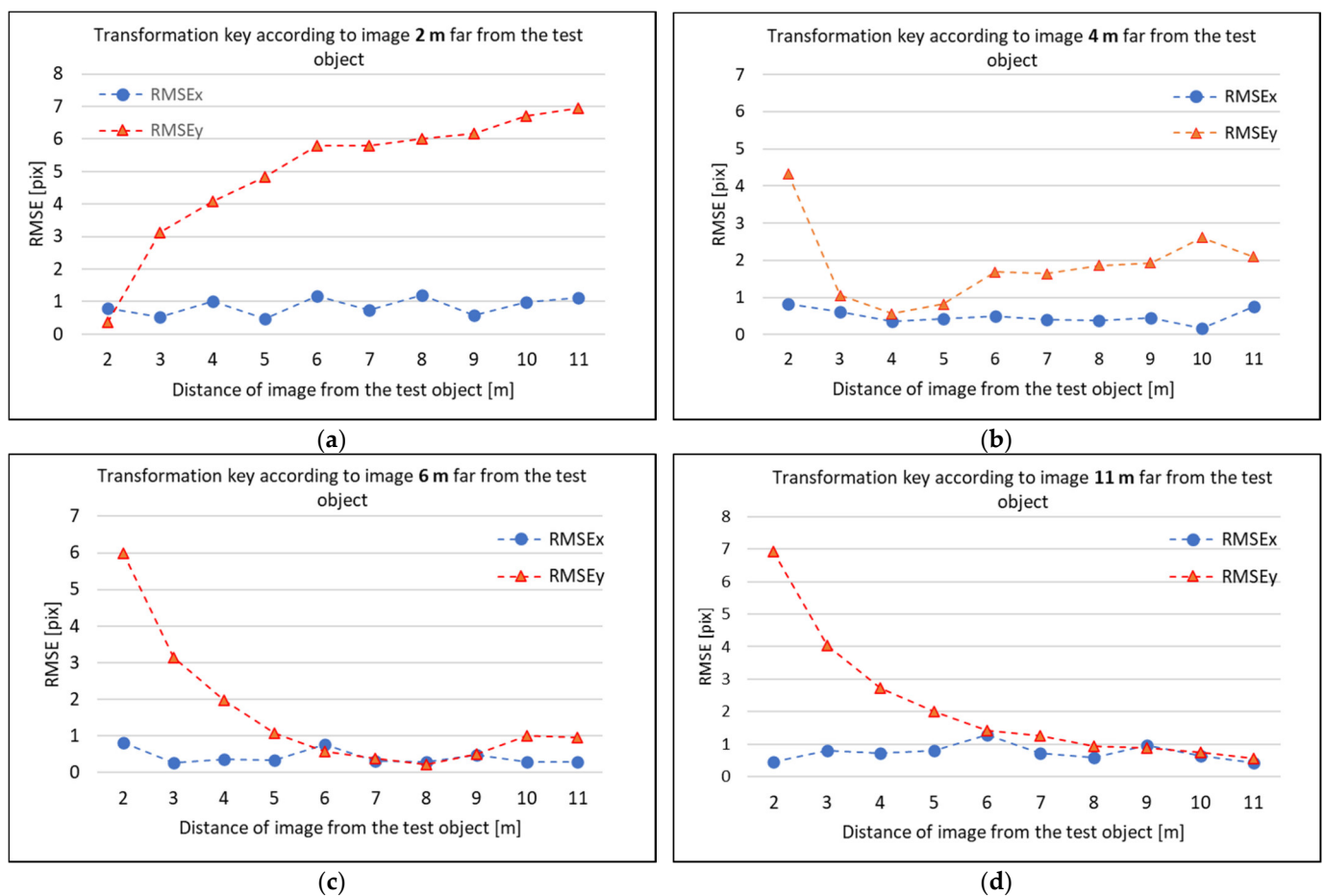
**Table 3.** Approximate sample distances for images captured from different distances.

Distance from the Object		2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m	11 m
Sample distance (mm)	TIR image	3.4	5.1	6.8	8.5	10.2	11.9	13.6	15.3	17.0	18.7
	RGB image	0.9	1.3	1.7	2.2	2.6	3.0	3.4	3.9	4.3	4.7

The affine transformation was chosen. The projective transformation was tested and the results were more or less the same or slightly worse. Thus, the affine transformation was chosen. For the affine transformation, a transformation key between the TIR image and RGB image of one image pair was calculated, i.e., tested to see if the transformation key calculated on one chosen pair was universal and sufficient enough for all other image pairs of the dataset. At first, the transformation key calculated on the image pair captured from the station 2 m from the test object was used. The transformation itself resulted in 0.4 pixel RMS. Then the same transformation key was used for the other pairs captured from other stations (3–11 m). For all distances, the transformed TIR images were overlaid on the corresponding RGB images and the residuals were measured on the photogrammetric targets on TIR images. From the residuals, RMSE for each axis on the TIR images were calculated. The X-axis of the images went from the image center to the right and Y-axis from the center image perpendicularly downwards. The resulting RMSEs are presented in Figure 5a, and are presented separately for each axis (RMSE<sub>x</sub>, RMSE<sub>y</sub>). The same procedure was repeated again, but the transformation key was computed in image pairs captured from station 4 m from the test object (Figure 5b), then in image pairs, captured 6 m from the test object (Figure 5c) and 11 m (Figure 5d). All of the presented RMSE pixel values are related to the TIR image pixels. Approximately 1 pixel in the TIR image was equal to 4 pixels in the RGB image.

From the presented figures, it is noticeable that there is no universal transformation key for all of the image pairs. Moreover, in all figures, it is possible to read that RMSE<sub>x</sub> is much lower and stable over all of the distances compared to the RMSE<sub>y</sub>. RMSE<sub>y</sub> differs, and there is a certain trend for all cases. The fact that RMSE<sub>y</sub> is higher could be caused by the relative position between the TIR camera and the RGB camera. The RGB camera is placed around 3 cm above the TIR camera in the same direction, as is Y-axis in the images (vertical direction). In the X-axis direction in the image (horizontal direction), the cameras are, by eye, placed in the same position. Moreover, the optical axis of the cameras is, likely, not ideally parallel, which can make an influence as well.

The observed trend in RMSE<sub>y</sub> could be readable. In the case of Figure 5a, the transformation key was computed in the image pair 2 m from the façade. At this distance, the RMSE<sub>y</sub> is low, 0.4 pixel (on TIR image), and equal to the RMS of the affine transformation. However, when the distance from the test object increases, the RMSE<sub>y</sub> increases as well. The RMSE<sub>y</sub> increases more rapidly when the distance is shorter, and when the distance from the façade is longer, the RMSE<sub>y</sub> increases slower. One could observe the same in the case of the transformation key computed in image pair 11 m from the façade. The lowest RMSE<sub>x</sub> and RMSE<sub>y</sub> are at an 11 m distance (Figure 5d). With a decreasing distance, the RMSE<sub>y</sub> increases, and when the distance is shorter, the RMSE<sub>y</sub> increases more rapidly. In the case of the transformation key, computed according to the 4 m image pair (Figure 5b) and 6 m image pair (Figure 5c), the RMSE<sub>y</sub> decreased to the distance where the transformation key was computed and then increased again with the longer distance.

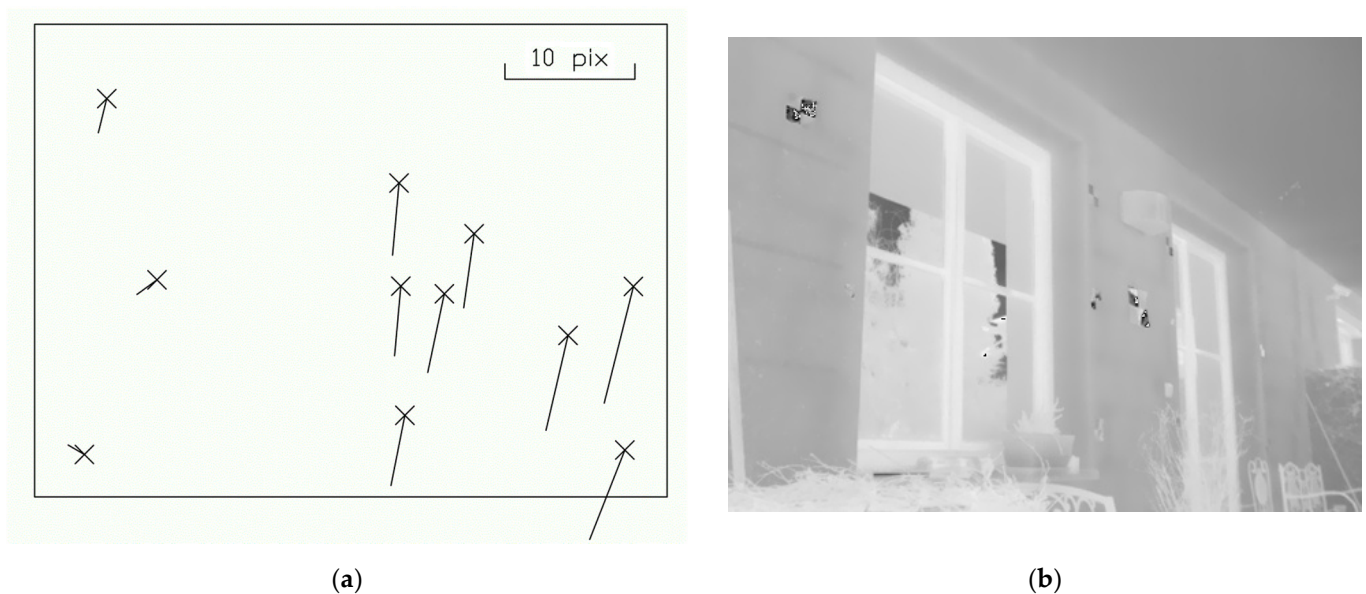


**Figure 5.** RMSE<sub>x</sub> and RMSE<sub>y</sub> of residuals on the TIR image transformed to the RGB image using affine transformation: (a) transformation key computed on 2 m image-pair; (b) transformation key computed on 4 m image-pair; (c) transformation key computed on 6 m image-pair; (d) transformation key computed on 11 m image-pair.

The results showed that it was not possible to find a universal transformation key that could be used for the plane transformation between the TIR and RGB image and the coordinate transformation was dependent on the distance from the object. For close-range photogrammetry, the plane transformation does not seem to be usable. The plane transformation does not particularly seem to be used for very close-range photogrammetry, with distances shorter than 5 m. At close-range photogrammetry, there are also captured images from different convergent angles, and the object structure is very often spatial. Thus, in one image, there could be an object, at the same time, in 2 m distances and 11 m distances. This was tested in one image pair (see Figure 6). The transformation key, which was computed in the 2 m image pair, was used. One could observe that the scene closer to the image (approximately 2 m) shows significantly lower residuals. When the scene is further (approximately 7 m), the residuals rise. All of the main residuals have the same direction.

Nevertheless, the plane transformation could still be used as a proper coordinate transformation. For example, taking the images from the more or less constant distance, especially from a further distance, could still be a reasonable option. One of the scenarios could be capturing the image pairs of the ground with no significant spatial structure (considered a plane, especially from a further distance) from UAV with the same flying altitude during capturing. This should be tested in the following experiments. However, for close-range photogrammetry, the plane transformation is not the right option.





**Figure 6.** Residuals of the transformed TIR image to the corresponding RGB image using the transformation key computed on image pair 2 m from the façade: (a) residuals on the targets (the scale of the residuals were increased); (b) thermal infrared image.

### 3. Spatial Transformation for Reprojection Method

The second suggested method, “Reprojection”, is designed to use a spatial transformation between the RGB image and the corresponding TIR image. The spatial transformation requires knowing the interior orientation of both RGB and TIR cameras and the relative pose between the cameras in the fixed system. At first, the interior orientation was determined by calibration and then the calculated parameters of interior orientation were used for relative pose estimation. The relative pose estimation was carried out by computing the photogrammetric model of the set of image pairs and determining the relative pose between the image pairs.

#### 3.1. Camera Calibration

There are two types of calibration of the thermal infrared camera—radiometric calibration and geometric calibration. The radiometric calibration was carried out by a contractor. The geometric camera calibration of the thermal infrared camera is presented for purposes of this article. Via geometric calibration are the estimated parameters of the interior orientation of the camera, which are necessary for the calculations of the suggested method. As the considered parameters of interior orientation, there were camera constant  $c$ , the position of the principal point  $p_x$  and  $p_y$ , and coefficients of the radial distortion  $k_1$ ,  $k_2$ . Tangential distortion was not considered, the shear was universally considered as 1, and the scale difference  $m$  was considered as 0.

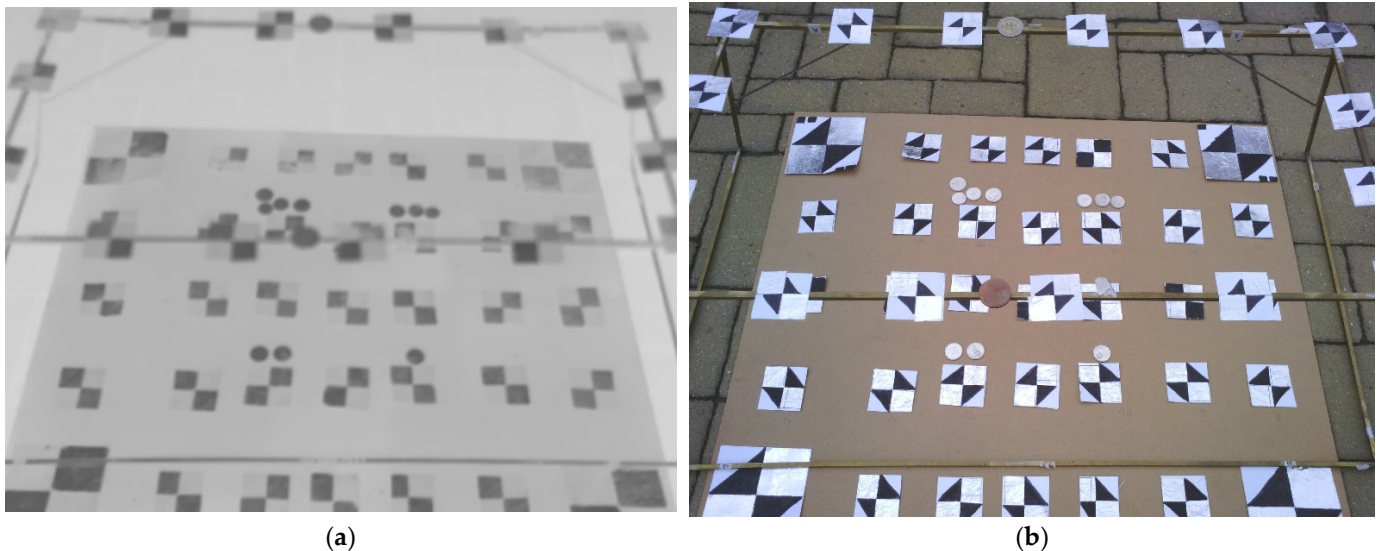
The parameters were determined by a camera calibration process using conventional photogrammetry. For the camera calibration, three calibration fields were designed and created.

- 2-level spatial calibration field;
- 3-level spatial calibration field;
- Plane calibration field.

Standard calibration fields are not usable in purposes of calibration of the thermal infrared camera, so calibration fields with visible targets in the TIR image must have been designed.

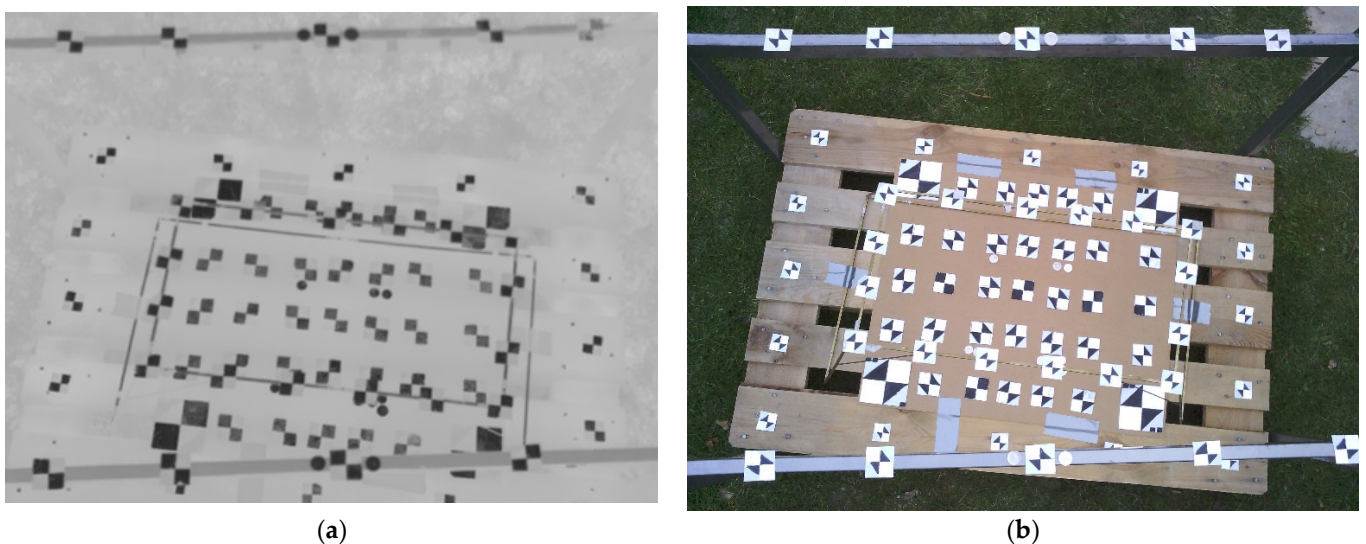
The 2-level spatial calibration field (Figure 7) was designed according to the suggestions for the spatial calibration field [19]. The calibration field consisted of a plane desk with approximate dimensions (700 × 600 mm) where were placed 35 cross aluminum

paper cross targets. A frame was placed on the desk, where another 16 targets were placed. Together, the 2-level spatial calibration field was created. The targets of the calibration field were measured by a total station, from one station, and there were measured distances on the field. The accuracy of the coordinates of the measured targets was determined as 3 mm and the accuracy of the measured distance was determined as 1 mm.



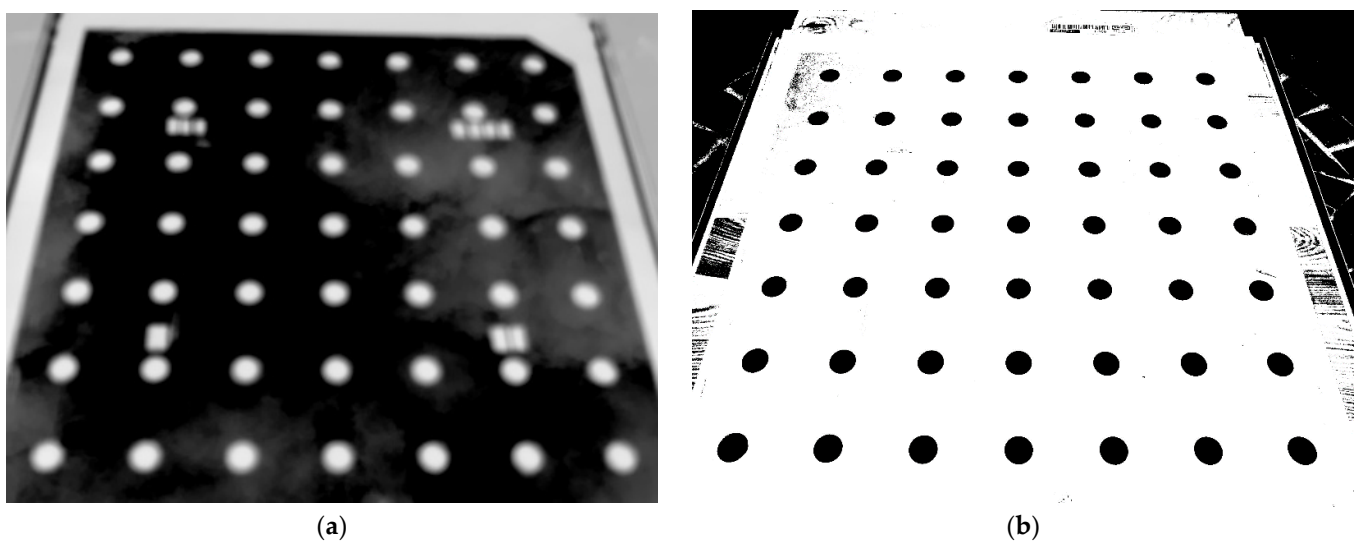
**Figure 7.** Images for camera calibration using the 2-level calibration field: (a) TIR image; (b) RGB image.

As a second calibration field, a 3-level calibration field was created (see Figure 8), on the same principle as the 2-level calibration field created. The approximate dimension of the 3-level calibration field was  $1200 \times 800 \times 73$  mm. The reason for creating the 3-level calibration field was to test if there was improvement in terms of accuracy, and if there was more 3D structure on the calibration field. The base was a plane desk, where the same targets were placed, similar to the 2-level calibration field. Over the desk, a first construction frame was placed, and then over the first construction frame, the second construction frame was placed. On both frames, targets were placed. Targets were measured by total stations and 12 distances were measured by a measuring tape. Both spatial calibration fields were homemade creations. Again, the accuracy of the coordinates of the measured targets was determined as 3 mm and the accuracy of the measured distance was determined as 1 mm.



**Figure 8.** Images for camera calibration using the 3-level calibration field: (a) TIR image (b) RGB image.

As a third calibration field, a plane calibration field was created, see Figure 9. In [18], the authors showed that, by using circular targets, it is possible to achieve better results compared to the chessboard plane calibration field. The plane calibration field was created professionally by the manufacturer. The manufacturer advised using stainless steel to have the highest possibility of a firm (not banded) plane field. The calibration field was a  $350 \times 350$  mm desk, 1 mm thick, from stainless steel; in the desk, there were cut-out circles with a diameter of 15 mm. The circles were cut out with a 0.1 mm tolerance of position. The dimensions of the circles were chosen to have enough pixels in the TIR images, if imaging was from a 30 cm distance for the ellipse operator or gravity operator function. The desk was placed on a flat wooden desk. The main idea behind cut-out circles from the calibration field was to have the same effect that we had with cross aluminum paper targets. Stainless steel is a material with low emissivity and the desk was placed on a wooden flat desk, so the holes had high emissivity; then, it was possible to detect the circles on TIR images.



**Figure 9.** Images for camera calibration using plane calibration field: (a) TIR image; (b) RGB image.

Capturing of images was carried out according to all of the suggested rules by [14,20,21]. On the TIR camera of FLIR E95, it is not possible to change aperture or shutter speed, so it was not possible to operate with focus depth. The TIR camera focused on the 7 m distance using a focusing ring, and the focus was fixed. The influence of changes in the focus to the interior orientation parameters was published in [22]. The RGB camera has a fixed-focus lens. Fixed-focus lenses rely on a very large depth of field. The depth of field of the RGB camera was suspiciously too large. The RGB images were sharp even in very short distances. Theoretically, there is post-processing on RGB cameras to eliminate blurring at such short distances. That could be a potential problem for calibrating the camera on the plane field, where images were captured very close to the plane field (40 cm). However, we had to presume that both cameras had a fixed optical system during the capturing.

For spatial calibration fields, eight images were captured. Four different oblique images with variant intersection angles and four perpendicular images were captured. The roll angle of the camera was rotated by  $\pm 90^\circ$  to determine the accurate principal point. Each side of the image was fully covered by targets at least once to have the ability to describe the distortion of the camera. In the case of the plane field, pilot experiments showed that eight images are not a sufficient number, probably because of coupling between the parameters of interior orientation and exterior orientation. The experiments showed that adding more convergent images with variant roll angles give more reasonable and stable results. Because of that, the plane field was captured by 16 convergent images with different roll angles and four perpendicular images with different roll angles.

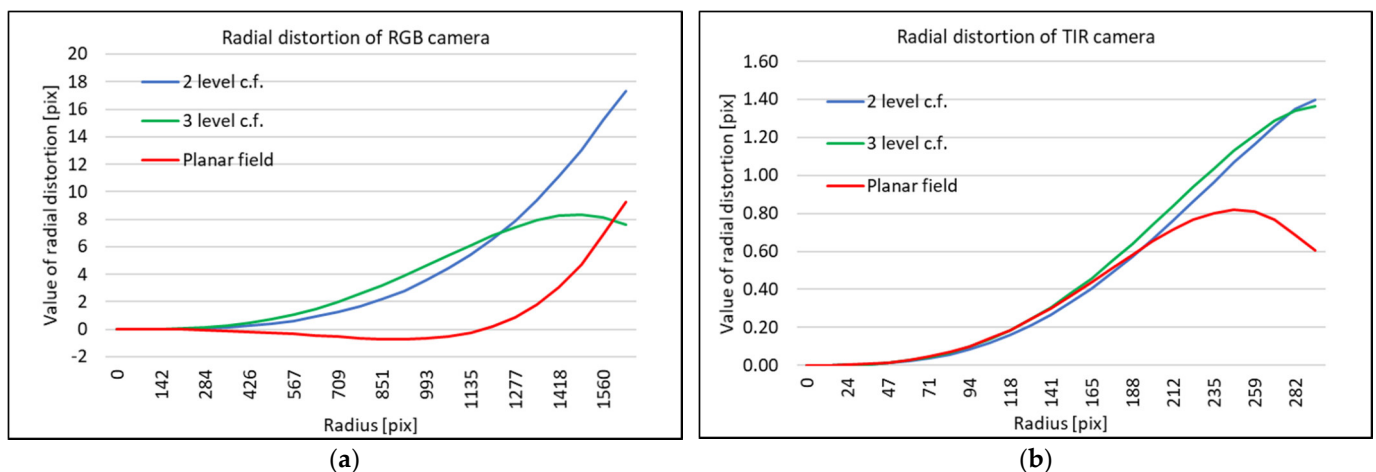


In the case of spatial fields, the targets were measured in images manually. The images were sharp enough to mark targets in the images with 1 pixel accuracy. Targets in images that belonged to the plane calibration field were measured by ellipse or gravity operators. At first, the ellipse operator did not work well in RGB images during the digitization of targets. Thus, the RGB images had to be pre-processed. Contrast of the images had to significantly increase and saturation of the images decrease. This process created more or less binary (black and white) images and it was then possible to use the ellipse operator. The pre-process also solved a problem of a visible thickness (1 mm) of the plane field in convergent images. Spatial calibration fields were large enough to capture the images with a certain distance, where the TIR images focused on the 7 m distances were still sharp. On the other hand, for proper distribution of the points, the plane field had to be captured from a closer distance so the TIR images were noticeably blurred. In the case of blurred circles, it was not possible to use the ellipse operator; the gravity operator was used instead. The different ways of digitizing the points in images led to setting different weights (in the following bundle adjustment process).

The calibration process was computed in the photogrammetric tool Close Range Digital Workstation (CDW). The resulting parameters are presented in Table 4 and Figure 10.

**Table 4.** The resulting interior orientations of the TIR camera and the RGB camera of FLIR E95, determined on different calibration fields.

		c (pix)	Px (pix)	Py (pix)
TIR camera	2-level calibration field	592.8	−2.7	1.0
	3-level calibration field	593.5	−3.3	1.4
	Plane calibration field	591.8	−2.6	2.1
RGB camera	2-level calibration field	2483.5	−26.5	25.3
	3-level calibration field	2481.4	−23.4	27.1
	Plane calibration field	2472.3	−29.1	31.8



**Figure 10.** Radial distortion determined by camera calibration using different calibration fields: (a) RGB camera (b) TIR camera.

Computed parameters of the interior orientation are very similar in the case of the 2-level calibration field and 3-level calibration field. On the other hand, there is a certain difference between computed principal points based on plane calibration and both multi-level calibration fields.

According to the relatively low values of radial distortion (Figure 10) for both cameras, there is a presumption that the images of both cameras are undistorted by a process in the FLIR E95 camera. Figure 10 shows the highest value of radial distortion around 18 pix in the far image corner; in the case of the TIR camera, the highest value is around 1.4 pix.

One could observe (Figure 10a)) different trends of distortion in comparison to the plane calibration field and on both multi-level calibration fields. The difference between the curves is increasing, and in the image width from the center (1296 pix), the difference in the radial distortion value is around 6 pixels. In the case of the TIR camera, the radial distortion is computed on the plane calibration field and both multi-level calibration fields seemed more or less the same from the image center to the image width (232 pix). In the far image corners (232 pix and further from the center), the curves of the radial distance differ by approximately 0.5 pix. In total, absolute values of the radial distortion are low. The resulting radial distortion confirms that the images could be considered undistorted even for plane transformation.

The advantage of the plane and 2-level calibration fields is that both fields are easily portable and can be used for on-site calibration, before and after the capturing images, to check if the optical system of the cameras remained fixed during the capturing.

### 3.2. Estimation of the Relative Pose

The relative pose was determined by a rotation matrix consisting of differences of rotation in three axes  $\Delta\omega$ ,  $\Delta\varphi$ , and  $\Delta\kappa$ , and by translation parameters among  $\Delta X$ ,  $\Delta Y$ , and  $\Delta Z$ . The translation parameters define the position of the TIR camera in the RGB camera coordinate system, when the X-axis is from the optical center of the RGB camera pointing to the right, Y-axis is from the optical center down, and Z-axis is the optical axis of the RGB camera going towards the scene. The parameters of relative orientation define the rotation of the TIR camera in the RGB camera coordinate system. For relative pose determination, all three calibration fields were used again. Each calibration field was captured by eight convergent image pairs in different angles in the fixed configuration using camera FLIR E95. Images were captured in a way to have images covered by targets on the calibration fields, as most as possible. This requirement was complicated by the different fields of view of both cameras. For each calibration field, eight image pairs were processed separately by the conventional photogrammetry process in CDW. For each group of eight images, the interior orientation was used according to the used calibration field. Thus, for example, the resulting interior orientation by calibration, on the plane calibration field, was used also for bundle adjustment when processing the eight image pairs, which also captured the plane calibration field. This is important to note because interior orientation correlates with exterior orientation. According to [23], in particular, principal point estimation from camera calibration strongly correlates with the exterior orientation. The exterior orientations of the images of the eight image pairs were estimated by bundle adjustment. Parameters and results of the bundle adjustment are presented in Table 5. For every image pair, using Equation (4), rotation matrices were computed. The rotation matrices were decomposed to the three angles  $\Delta\omega$ ,  $\Delta\varphi$ , and  $\Delta\kappa$  determining the relative rotation between the TIR and RGB camera. Using Equation (5), the relative translation of the cameras was computed. Therefore, for every calibration field, eight relative poses were computed, i.e., eight time parameters  $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$ ,  $\Delta\omega$ ,  $\Delta\varphi$ , and  $\Delta\kappa$ . From eight values of each parameter, the mean and its standard deviation were calculated. The mean was considered the final value. The resulting parameters of the relative pose are presented in Table 6. Standard deviations of the means are presented in Table 7.

### 3.3. Verification of Calculated Parameters

An experiment was carried out to test the calculated parameters of the relative pose. A façade of a family house was chosen as a test object for the experiment. Aluminum paper cross targets were placed on the façade. The coordinates of the targets were measured by a total station from the single station. FLIR E95 was used to capture the images in the experiment. For every RGB image, TIR image was also captured. In total, there were 35 image pairs captured. Among the 35 image pairs, there were also 10 image pairs that were captured perpendicular to the façade, at 10 different distances from the façade (2–11 m), the same as it was when testing the plane transformation mentioned in the

previous chapter. The sample distances of the image pairs are approximately the same as stated in Table 3. The capturing scenario is presented in Figure 11.

**Table 5.** Parameters and results of the bundle adjustment. (The software that was used, where the bundle adjustment was computed, uses different angle combinations: azimuth, tilt, swing).

		Distance from the Object (m)	Sample Distance (mm)	Avg. Residuum on Image Coordinates X	Avg. Residuum on Image Coordinates Y	$\sigma_X$ (mm)	$\sigma_Y$ (mm)	$\sigma_Z$ (mm)	$\sigma_\alpha$ (mm)	$\sigma_\nu$ (mm)	$\sigma_\kappa$ (mm)
2-level calibration field	TIR	0.9	1.5	0.28	0.29	0.89	0.67	0.95	0.057	0.060	0.046
	RGB	0.9	0.4	0.96	1.11	0.90	0.70	0.95	0.058	0.063	0.047
3-level calibration field	TIR	1.5	2.6	0.24	0.23	1.13	0.82	1.27	0.048	0.049	0.040
	RGB	1.5	0.6	0.88	1.00	1.12	0.82	1.26	0.048	0.050	0.040
Plane calibration field	TIR	0.4	0.7	0.13	0.24	0.23	0.26	0.26	0.036	0.062	0.061
	RGB	0.4	0.2	0.72	0.91	0.33	0.31	0.32	0.047	0.078	0.078

**Table 6.** The resulting parameters  $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$ ,  $\Delta\omega$ ,  $\Delta\varphi$ , and  $\Delta\kappa$ .

	$\Delta X$ (m)	$\Delta Y$ (m)	$\Delta Z$ (m)	$\Delta\omega$ (°)	$\Delta\varphi$ (°)	$\Delta\kappa$ (°)
2 level calibration field	−0.0001	−0.0246	−0.0045	−0.800	0.078	−0.009
3 level calibration field	−0.0002	−0.0248	−0.0065	−0.833	−0.061	−0.007
Plane calibration field	−0.0002	−0.0252	−0.0061	−0.663	0.163	−0.020

**Table 7.** The standard deviation of the mean of parameters  $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$ ,  $\Delta\omega$ ,  $\Delta\varphi$ ,  $\Delta\kappa$ .

	$\sigma_{\Delta X}$ (m)	$\sigma_{\Delta Y}$ (m)	$\sigma_{\Delta Z}$ (m)	$\sigma_{\Delta\omega}$ (°)	$\sigma_{\Delta\varphi}$ (°)	$\sigma_{\Delta\kappa}$ (°)
2 level calibration field	0.0008	0.0005	0.0004	0.031	0.047	0.027
3 level calibration field	0.0007	0.0008	0.0004	0.036	0.029	0.016
Plane calibration field	0.0006	0.0005	0.0001	0.058	0.065	0.041

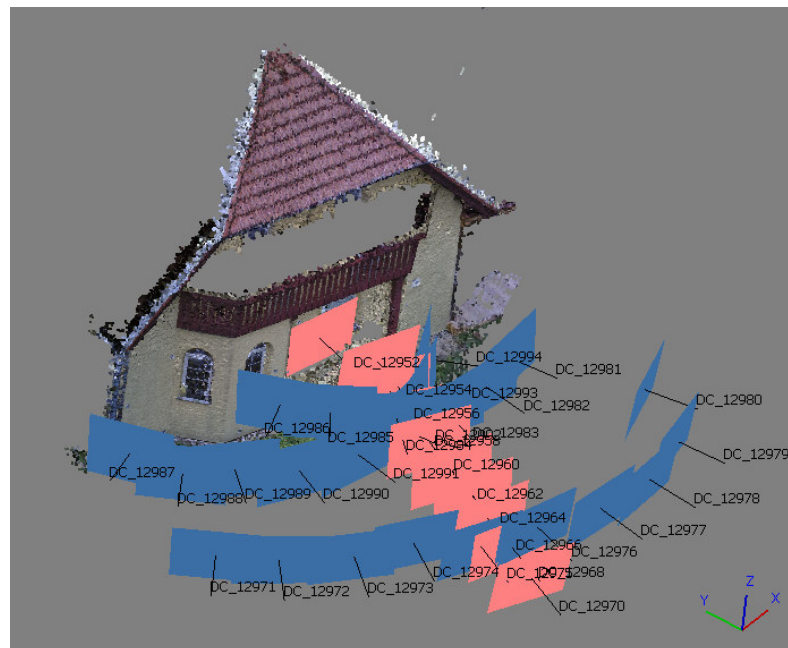
First, only RGB images were processed solely by the SfM method using commercial software Agisoft Metashape. The images were processed in three different projects. For each project, before the processing, there were set parameters of the interior orientation of the cameras according to the used calibration field. All of the inputted values of interior orientation were fixed and were not changed during further processing. After the relative orientation of the images, the geo-referencing on the targets was carried out. The total RMSE of the georeferencing was around 5 mm for all three groups of interior orientation. Parameters of exterior orientation ( $X$ ,  $Y$ ,  $Z$ ,  $\omega$ ,  $\varphi$ ,  $\kappa$ ) for the 10 perpendicular images were used for testing the computed relative poses.

For each image, of the 10 perpendicular images, the measured photogrammetric targets on the façades were transformed to the RGB camera coordinate system. Then, using the parameters of the relative pose ( $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$ ,  $\Delta\omega$ ,  $\Delta\varphi$ ,  $\Delta\kappa$ ), the points were transformed to the TIR camera coordinate system of the corresponding TIR image. From the TIR camera coordinate system, the points were reprojected to the corresponding TIR image using calculated parameters of the interior orientation. The affiliation to the chosen calibration field was preserved. For example, when a plane calibration field was used for camera calibration and for calculation of parameters of relative poses between the cameras, then the interior orientation was also used for calibration input before processing in Agisoft Metashape. After exterior orientation of RGB images, the relative pose estimated on the plane calibration field was used to transform the point between the camera coordinate systems and, again, the interior orientation calibrated on the plane calibration field was used for reprojection to the TIR image. This was highly important in the case of coupling the parameters of the interior orientation and exterior orientation and relative pose, respectively. The reprojected point to the TIR image was compared to the position of the photogrammetric target in the TIR image, and the residuals were calculated. From the residuals, RMSE<sub>x</sub> and RMSE<sub>y</sub> were calculated. The process was repeated for all 10 perpendicular image pairs. The resulting

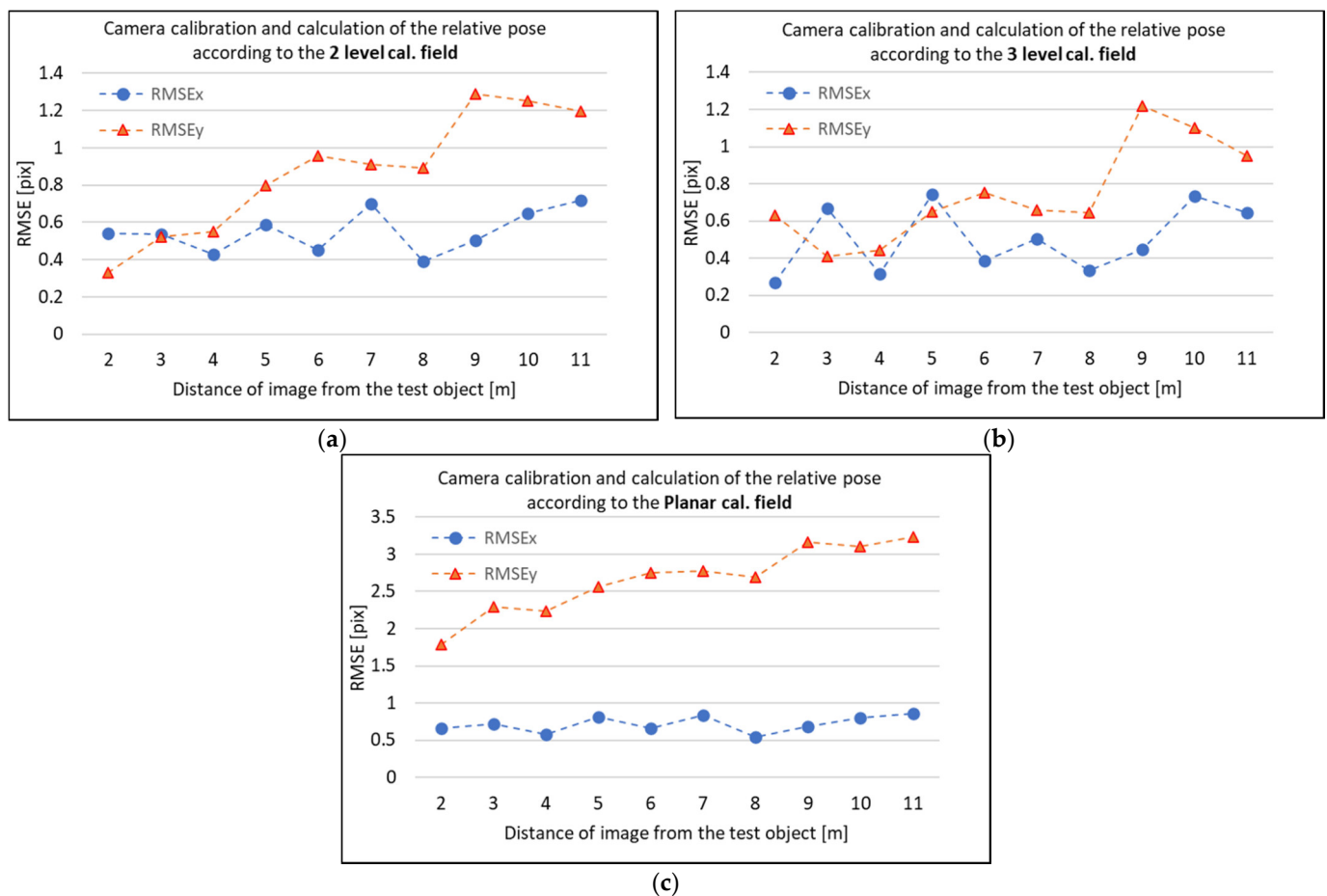
RMSE<sub>x</sub> and RMSE<sub>y</sub> are presented in Figure 12. All of the presented RMSE pixel values are related to the TIR image pixel.

According to Figure 12, in the case of the 2-level and 3-level calibration fields, one could observe that the RMSE<sub>x</sub> and RMSE<sub>y</sub> are both mostly under the 1-pixel in the corresponding TIR image. The RMSE<sub>y</sub> is a bit higher, and at some of the distances, RMSE<sub>y</sub> exceeds 1 pix. It is important to mention that, even though the resulting RMSEs are respectable, the TIR image has significantly low-resolution compared to the RGB image, and even 1 pixel does matter.

RMSE<sub>y</sub> can be influenced by the Y coordinate of the principal point of the TIR camera and also by the Y coordinate of the principal point of the RGB camera (which influences the exterior orientation angle  $\omega_{RGB}$ ). RMSE<sub>y</sub> is also influenced slightly by parameter  $\Delta Y$  of the relative pose and by  $\Delta\omega$ . On one TIR image, all residuals were more or less constant with the same direction. The testing showed that estimation of the translation between the cameras does not have such influence, but the estimation of the relative rotation influences the result strongly. The difference of  $0.05^\circ$  in  $\Delta\omega$  can cause the difference of 0.5 pix in the TIR image. According to Table 7,  $\Delta\omega$  was calculated with  $\sigma_{\Delta\omega}$   $0.03^\circ$ – $0.06^\circ$ . Estimation of the  $\Delta\omega$  angle can then have a high influence. There are insufficient results in the case of the plane calibration field (Figure 12c). However, the result is still better compared to the plane transformation and does not differ with too many overall distances (1.8 pix for 2 m and 3.2 pix for 11 m). From Figure 12c, it is (slightly) possible to observe a trend in the case of the plane calibration field, but there is not enough data to make a proper statement. In the case of the plane calibration field, insufficient results may have their core in the insufficient interior orientation determination, as was previously discussed.



**Figure 11.** Capturing scenario of the experiment; 10 perpendicular images are highlighted.



**Figure 12.** RMSE<sub>x</sub> and RMSE<sub>y</sub> of residuals (on TIR images) of spatial transformation using different result parameters  $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$ ,  $\Delta\omega$ ,  $\Delta\varphi$ ,  $\Delta\kappa$ : (a) parameters calculated on the 2-level calibration field; (b) parameters calculated on the 3-level calibration field; (c) parameters calculated on the plane calibration field.

#### 4. Conclusions

To obtain accurate 2D and 3D models, enriched with information derived from the original thermal infrared image, it seems necessary to come up with certain methods of photogrammetric co-processing of RGB and TIR images. Two of the methods were previously suggested by the authors and they are also briefly presented in this article. Both methods are designed for close-range photogrammetry purposes. For the first method, Sharpening, the plane transformation was designed. For the second method, the spatial transformation was designed. Determining and testing the transformations were the main goals of this article. For purposes of this paper, a handheld TIR camera, FLIR E95, equipped by an RGB camera, was used. Cameras like this offer, naturally, a fixed system of TIR and RGB cameras. This advantage could be used for photogrammetric co-processing.

For plane transformation (affine transformation, in our case), it was necessary to use undistorted images captured by cameras relatively close to each other, with the same direction of their optical axes. Then the transformation key for the affine transformation was computed. An experiment was carried out to see if there was a possibility of finding a universal transformation key, which could be used for all other image pairs of the dataset. The universal transformation key would let us avoid the process of computing a unique transformation key for every image pair, which would make the Sharpening method difficult and significantly less automatic. However, the experiment showed that, in the case of our dataset, it was not possible to find a universal transformation key. The resulting RMSE<sub>y</sub> was dependent on the distance of the object to the image. This was probably caused by the shift of the TIR and RGB camera (the RGB camera was physically placed 3 cm above



the TIR camera). According to the results of the experiment, there is potential to use the plane transformation for purposes of photogrammetry, where there are longer distances to the object, and where the distance of the capturing is more or less constant. One of the applications could be to use UAVs from constant mid-range distances, capturing the ground where surface changes are not significant. However, the article is focused on close-range distances, and at close-range photogrammetry, there are expected substantial differences in the distance on the scene during the capturing. For close-range photogrammetry, plane transformation does not seem sufficient.

Spatial transformation is designed for the Reprojection method. Spatial transformation also requires a fixed system of TIR and RGB cameras. The advantage of spatial transformation is that the process does not require undistorted images. Spatial transformation requires geometric calibration of both cameras and estimation of the relative pose between TIR and RGB cameras. The relative pose consists of relative translation with parameters  $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$ , and relative rotation with parameters  $\Delta\omega$ ,  $\Delta\phi$ , and  $\Delta\kappa$ . The process of geometric calibration and estimation of the relative pose is described in this paper. The experiments showed that, during camera calibration, it is very important to decouple parameters of the interior and exterior orientation. For estimation of the relative pose, it is important to use decoupled interior orientation parameters, and accurately and precisely determine the exterior orientation of the images, of the image pairs. This requires a short imaging distance (larger sample distance), proper intersection angles, and accurate control points on the scene. In the provided experiments, the relative pose was estimated on three different calibration fields. The best results showed estimation on the 3-level calibration field. The standard deviation of the mean of the calculated translation parameter was less than 1 mm and the rotation parameter less than  $0.04^\circ$ . The calculated parameters were tested on 10 different image pairs with all of them captured perpendicular to the façade. The exterior orientation of 10 RGB images of the image pair was determined by the SfM method when processed together with several other images of the façade. The testing showed that the distance does not have a major influence on the resulting RMSE in the TIR image. RMS was oscillating around 0.5 pix and RMSEy was a bit higher, 0.8 pix, but, twice, RMSEy little exceeded the 1 pixel value. RMSE 1 pixel is still respectable, but a pixel in the TIR image has a larger sampling distance due to lower resolution of the TIR image compared to the RGB image.

Overall, the testing of spatial transformation showed better reliability. For close-range purposes, it is not sufficient to use plane transformation and it is advisable to use spatial transformation instead. If the spatial transformation was also used for the Sharpening method, the method would be an alternative to the Reprojection method. Based on the conclusions of this article, the Sharpening method must be redesigned and the spatial transformation must be used. The authors believe that the way on “how” will be via spatial transformation implementation being original. Detailed descriptions of processes, of both methods and results of photogrammetric co-processing using those methods, will be presented in a following article.

**Author Contributions:** Conceptualization, A.D., K.V., and K.P.; methodology, A.D.; technical advisory and supervision, K.V. and K.P.; data processing, A.D.; investigation, K.V. and K.P.; writing—original draft preparation, A.D.; writing—review and editing, A.D.; visualization, A.D.; project administration, K.V. and K.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Grant Agency of the Czech Technical University in Prague, grant no. SGS21/054/OHK1/1T/11.

**Acknowledgments:** This research is partially from the Department of Geomatics at the Czech Technical University in Prague and partially internal research of EuroGV spol., s r. o.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Capel, D.; Zisserman, A. Computer vision applied to super resolution. *IEEE Signal Process. Mag.* **2003**, *20*, 75–86. [[CrossRef](#)]
2. Wakeford, Z.E.; Chmielewska, M.; Hole, M.J.; Howell, J.A.; Jerram, D.A. Combining thermal imaging with photogrammetry of an active volcano using UAV: An example from Stromboli, Italy. *Photogramm. Rec.* **2019**, *34*, 445–466. [[CrossRef](#)]
3. Sledz, A.; Unger, J.; Heipke, C. Thermal IR imaging: Image quality and orthophoto generation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *42*, 413–420. [[CrossRef](#)]
4. Maset, E.; Fusiello, A.; Crosilla, F.; Toldo, R.; Zorzetto, D. Photogrammetric 3D building reconstruction from thermal images. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *4*, 25–32. [[CrossRef](#)]
5. Grechi, G.; Fiorucci, M.; Marmoni, G.M.; Martino, S. 3D Thermal Monitoring of Jointed Rock Masses through Infrared Thermography and Photogrammetry. *Remote Sens.* **2021**, *13*, 957. [[CrossRef](#)]
6. Hoegner, L.; Tuttas, S.; Xu, Y.; Eder, K.; Stilla, U. Evaluation of methods for coregistration and fusion of rpas-based 3d point clouds and thermal infrared images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 241–246. [[CrossRef](#)]
7. Weber, I.; Jenal, A.; Kneer, C.; Bongartz, J. PANTIR—a dual camera setup for precise georeferencing and mosaicing of thermal aerial images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 269–272. [[CrossRef](#)]
8. Adamopoulos, E.; Volinia, M.; Giroto, M.; Rinaudo, F. Three-Dimensional Thermal Mapping from IRT Images for Rapid Architectural Heritage NDT. *Buildings* **2020**, *10*, 187. [[CrossRef](#)]
9. Javadnejad, F.; Gillins, D.T.; Parrish, C.E.; Slocum, R.K. A photogrammetric approach to fusing natural colour and thermal infrared UAS imagery in 3D point cloud generation. *Int. J. Remote Sens.* **2020**, *41*, 211–237. [[CrossRef](#)]
10. Dlesk, A.; Vach, K. Point Cloud Generation of a Building from Close Range Thermal Images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 29–33. [[CrossRef](#)]
11. Acorsi, M.G.; Gimenez, L.M.; Martello, M. Assessing the Performance of a Low-Cost Thermal Camera in Proximal and Aerial Conditions. *Remote Sens.* **2020**, *12*, 3591. [[CrossRef](#)]
12. Šedina, J.; Housarová, E.; Raeva, P. Using RPAS for the detection of archaeological objects using multispectral and thermal imaging. *Eur. J. Remote Sens.* **2019**, *52*, 182–191. [[CrossRef](#)]
13. Řehák, M.; Pavelka, K. Using of uav for photogrammetry and thermal imaging. In Proceedings of the 33rd Asian Conference on Remote Sensing, ACRS2012, CD-Proceedings, Pattaya, Thailand, 26–30 November 2012.
14. Luhmann, T.; Robson, S.; Kyle, S.; Harley, I. *Close Range Photogrammetry. Principles, Techniques and Applications*; Whittles Publishing: Scotland, UK, 2011; pp. 448–458.
15. User’s Manual FLIR Exx Series. Available online: <https://www.flir.com/globalassets/imported-assets/document/flir-exx-series-user-manual.pdf> (accessed on 26 April 2021).
16. Fluke: Thermal Cameras. Available online: <https://www.fluke.com/en-in/products/thermal-cameras> (accessed on 26 April 2021).
17. The Engineering ToolBox: Emissivity Coefficient Materials. Available online: [https://www.engineeringtoolbox.com/emissivity-coefficients-d\\_447.html](https://www.engineeringtoolbox.com/emissivity-coefficients-d_447.html) (accessed on 7 June 2021).
18. Usamentiaga, R.; Garcia, D.F.; Ibarra-Castanedo, C.; Maldague, X. Highly accurate geometric calibration for infrared cameras using inexpensive calibration targets. *Measurement* **2017**, *112*, 105–116. [[CrossRef](#)]
19. Remondino, F.; Fraser, C. Digital camera calibration methods: Considerations and comparisons. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *36*, 266–272.
20. Luhmann, T.; Piechel, J.; Roelfs, T. Geometric calibration of thermographic cameras. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2010**, *38*, 411–416.
21. Förstner, W.; Wrobel, B.P. *Photogrammetric Computer Vision*; Springer International Publishing: Cham, Switzerland, 2016; pp. 696–707.
22. Sanz-Ablanedo, E.; Chandler, J.H.; Wackrow, R. Parameterising internal camera geometry with focusing distance. *Photogramm. Rec.* **2012**, *27*, 210–226. [[CrossRef](#)]
23. Clarke, T.A.; Wang, X.; Fryer, J.G. The principal point and CCD cameras. *Photogramm. Rec.* **1998**, *16*, 293–312. [[CrossRef](#)]

# Article 4

## Article

# Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images

Adam Dlesk<sup>1,\*</sup>, Karel Vach<sup>2</sup> and Karel Pavelka<sup>1</sup> 

<sup>1</sup> Department of Geomatics, Faculty of Civil Engineering, Czech Technical University in Prague, 16636 Prague, Czech Republic; pavelka@fsv.cvut.cz

<sup>2</sup> EuroGV, spol. s r. o., 11000 Prague, Czech Republic; vach@eurogv.cz

\* Correspondence: adam.dlesk@fsv.cvut.cz

**Abstract:** In some applications of thermography, spatial orientation of the thermal infrared information can be desirable. By the photogrammetric processing of thermal infrared (TIR) images, it is possible to create 2D and 3D results augmented by thermal infrared information. On the augmented 2D and 3D results, it is possible to locate thermal occurrences in the coordinate system and to determine their scale, length, area or volume. However, photogrammetric processing of TIR images is difficult due to negative factors which are caused by the natural character of TIR images. Among the negative factors are the lower resolution of TIR images compared to RGB images and lack of visible features on the TIR images. To eliminate these negative factors, two methods of photogrammetric co-processing of TIR and RGB images were designed. Both methods require a fixed system of TIR and RGB cameras and for each TIR image a corresponding RGB image must be captured. One of the methods was termed sharpening and the result of this method is mainly an augmented orthophoto, and an augmented texture of the 3D model. The second method was termed reprojection and the result of this method is a point cloud augmented by thermal infrared information. The details of the designed methods, as well as the experiments related to the methods, are presented in this article.

**Keywords:** photogrammetry; thermal infrared image; close-range photogrammetry; thermography; augmented orthophoto; augmented point cloud



**Citation:** Dlesk, A.; Vach, K.; Pavelka, K. Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images. *Sensors* **2022**, *22*, 1655. <https://doi.org/10.3390/s22041655>

Academic Editor: Nikolaos Doulamis

Received: 16 December 2021

Accepted: 17 February 2022

Published: 20 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

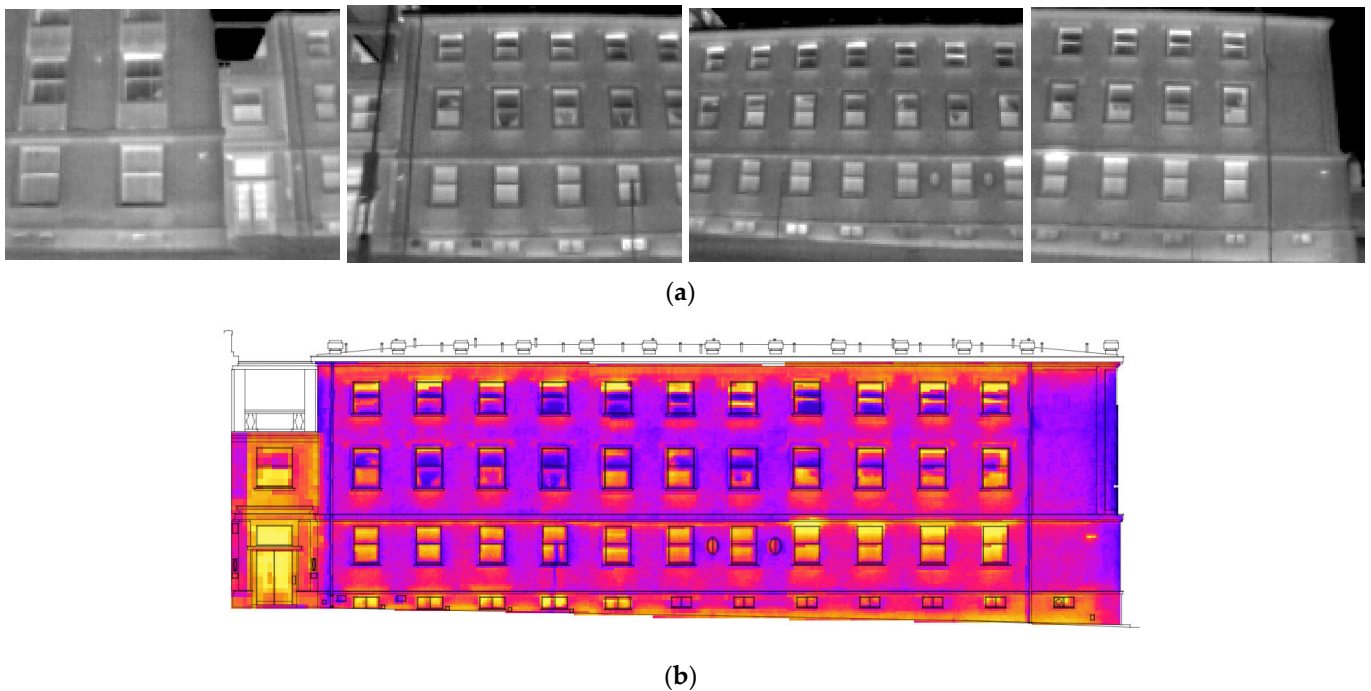
## 1. Introduction

Infrared thermography (IRT) is a non-contact, non-destructive, and non-invasive technique [1] which can be carried out using measuring devices that are able to receive radiation (of a certain wavelength) from an object which has a temperature higher than absolute zero [2]. Thermal infrared cameras are devices which are equipped with a sensor which can receive the thermal infrared radiation of the most external layer of the surface of the object and then create a thermal infrared (TIR) image, also called a thermogram. The TIR image is a matrix where a value of the intensity of the TIR radiation, or the intensity converted to the surface temperature, is stored in each element of the matrix. Then, on the TIR image, it is possible to analyze the temperature of the captured object [3] or the temperature distribution, differences, anomalies, maximum, minimum, etc. The TIR image can also be called as an IRT image [4], but when referring to photogrammetry, the term TIR image is more often used [5–8]. As is described in published studies, thermography can be used in a wide variety of applications in many different industries.

The main goal of this article is the development of methods aimed to building diagnostics, but in the literature, there are many published applications of thermography. A review of published articles concerned with IRT in the building industry was presented by [3]. In the building industry, relevant applications include surface diagnostics [9], material investigation [10], and energy monitoring [11]. In the building industry, IRT has also been used for object detection on a building [12] and for archaeological purposes [13]. IRT is

also widely used in other industries, such as energetics [14], agriculture [15], security [16], zoology [17], seed [18] and plant [19] monitoring, and environmental protection [20]. There is also extensive use of IRT in the medical industry. A review of published articles in this industry was presented by [21].

Most of the applications referred to treat TIR images as a simple 2D image [9,13,17–19]. This approach is sufficient for many purposes. However, for some applications, further spatial information, such as dimension, orientation, or the area of studied thermal occurrences, may be necessary, e.g., [22]. Adding spatial information to the TIR image can increase the value of the thermal investigation. When the studied object is large (e.g., buildings, urban areas, agriculture fields, glaciers, volcanos), and knowing that TIR cameras usually have low resolution and a narrow field of view, it is necessary to capture the object with tens or hundreds of TIR images. Thermal investigation of the object then becomes enormously complicated or even impossible. When TIR images are processed photogrammetrically, the results of the processing can be an orthophoto, a mosaic (2D) (see Figure 1), a point cloud or a textured 3D model (3D). All the spatial results can be augmented by thermal infrared information derived from the original TIR images. The results are optionally scaled and transformed to the selected coordinate system.



**Figure 1.** (a) TIR images of a building façade. The thermal investigation of the façade on single TIR images is confusing; (b) A mosaic created from the single TIR images. The mosaic helps in carrying out improved investigation of the leaks on the façade.

The structure from motion (SfM) method is a prevalent photogrammetric method [23]. Photogrammetric processing of TIR images can also be carried out using the SfM method [24–26]. Using these results, it is possible to better understand the studied object, as they enable analysis in three dimensions, automatic or semi-automatic detection of thermal anomalies, the gathering of spatially oriented semantic information, etc.

The photogrammetric processing of TIR images involves several challenging difficulties [27]. TIR images have low resolution compared to the common RGB images [7]. TIR images lack visible features or can even be featureless (if there are no temperature changes in the object scene), which significantly complicates, or makes impossible, SfM processing [28]. Algorithms for the SfM method are not able to identify a sufficient number of key or tie points on TIR images, and, as a result, the co-registration of TIR images can fail (see Figure 2).





**Figure 2.** TIR and RGB image of the same scene: (a) Feature matching algorithm failed to find any key points or tie points; (b) Feature matching algorithm found a sufficient number of tie points (blue dots) [28].

Another issue is that commonly used photogrammetric targets are not visible on TIR images and, therefore, it is not possible to scale or transform the results. To address the negative factors referred to, several methods of photogrammetric co-processing of TIR and RGB images have been presented in the literature. A review of these methods has been provided in [28]. Generally, for the processing of RGB images, geometric accuracy is used and the results are augmented by corresponding information from TIR images. Overall, the presented methods can be divided into two groups according to their approach. The first approach does not require a fixed system of RGB and TIR cameras and capturing the RGB and TIR images can be carried out completely separately, even at different times. The second approach requires a fixed system of RGB and TIR cameras and, for every captured TIR image, the corresponding RGB image must be captured. The fact that modern hand-held thermal infrared cameras for industrial applications usually have both RGB and TIR sensors, and the camera itself creates a natural fixed system, helps in the development of methods of this approach.

In [24], the authors proposed two methods. Both methods belong to the category in which a fixed system of cameras is required. The first proposed method is called the “sharpening method” and the second method is called the “reprojection method”. The aim of this article is to build on the research published in previous articles [24,28], to describe the sharpening and reprojection methods in detail, and to highlight their specific characteristics. The methods are discussed, and the advantages and disadvantages are pointed out.

A similar approach which is based on the sharpening method was also published by [29]. However, only a plane transformation was used there and, for close-range purposes, plane transformation is not sufficient. Instead of plane transformation, a ray recalculation method has been developed and used. The ray recalculation method is the major improvement for the sharpening method. A similar approach to that of the reprojection method was used also by [6]. The main difference between the approaches is the “test of visibility”, which was designed for the reprojection method. The test of visibility is a key and necessary part of the reprojection method. The test of visibility is described and related experiments are presented in this paper.

#### *Thermal Infrared Image*

This section briefly describes the theoretical background of the process of creating a TIR image. The information on the theoretical background is mainly drawn from [3,30–33].

Every object which has a temperature above absolute zero ( $-273.15\text{ }^{\circ}\text{C}$ ) emits electromagnetic radiation from its surface. The emitted radiation passes through the ambient atmosphere and can be captured by a device which has a thermal infrared sensor, e.g., a thermal infrared camera. The device captures the radiation in a certain wavelength range and records the intensity of the radiation. Given the parameters of the object and the surrounding ambience, the device can calculate the surface temperature.

The theory of thermography is based on a theoretical object called the black body. The black body is an ideal object which absorbs all incoming radiation and emits all its energy from its surface. According to Planck's radiation law, the spectral specific radiation  $M_{0\lambda}$  of a black body can be described. In Equation (1),  $c$  is the speed of light,  $h$  is Planck's constant,  $k$  is the Boltzmann constant,  $T$  is the temperature in kelvins,  $C_1 = 3.74 \times 10^{-16}\text{ W m}^2$  and  $C_2 = 1.44 \times 10^{-2}\text{ K m}$ .

$$M_{0\lambda} = \frac{2\pi hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} = \frac{C_1}{\lambda^5} \frac{1}{e^{\frac{C_2}{\lambda T}} - 1} \quad (1)$$

The spectral specific radiation is dependent on the wavelength for different temperatures. The dependency is presented in [34]. In [34], it is possible to observe that when the temperature of a black body increases, the maximum of the spectral specific radiation is in shorter wavelengths. This phenomenon is described by Wein's displacement law  $\lambda_{max}$ , where  $b = 2898\text{ }\mu\text{m} \cdot \text{K}$  and  $T$  is the temperature in kelvins.

$$\lambda_{max} = \frac{b}{T} \quad (2)$$

By integrating the formula for spectral specific radiation over all wavelengths from 0 to  $\infty$ , the intensity of the radiation  $M_0$  is obtained. According to the Stefan–Boltzmann law, the intensity of the radiation is equal to the fourth power of its absolute temperature. In non-contact measurements, this law is used for the determination of the surface temperature from the thermal radiation.

$$M_0 = \sigma T^4, \quad \sigma = 5.67 \times 10^{-8}\text{ Wm}^{-2}\text{K}^{-4} \quad (3)$$

Real objects, unlike a black body, do not emit all their energy. The spectral specific radiation for these objects is given by Equation (4):

$$M_\lambda = \varepsilon_\lambda M_{0\lambda} \quad (4)$$

where  $\varepsilon_\lambda$  is the spectral emissivity of the object. The value of the spectral emissivity is between 0 and 1 (the spectral emissivity of a black body is equal to 1). The spectral emissivity is dependent on the wavelength, temperature, material and the angle of the observation. However, there are some objects for which it is possible to consider the emissivity as a constant (in a certain wavelength range). In the literature, these objects are called grey bodies.

If there is radiation towards the object (e.g., the Sun) and the object has a temperature higher than absolute zero, the object:

- absorbs the energy from the source (absorption  $\alpha$ ),
- reflects the energy from the source (reflectance  $\rho$ ),
- the energy from the source transmits through the object (transmissivity  $\tau$ ),
- emits the absorbed energy (emissivity  $\varepsilon$ ).

Then  $\varepsilon + \rho + \tau = 1$  and, because most bodies do not transmit in the infrared part of the electromagnetic spectrum, then  $\varepsilon + \rho = 1$ . The emissivity is a key parameter in non-contact temperature measurement. The emissivity of object materials is usually stated in published tables, many publicly available on the internet. When the emissivity is known, the reflectance is given by  $\rho = 1 - \varepsilon$ .

In the case of non-contact temperature measurement, a device with a thermal infrared sensor is placed in front of the examined object. The object has a temperature  $T_{OBJ}$  and the



object is made from material with emissivity  $\varepsilon_{OBJ}$ . The object emits its own energy through radiant flux  $\varepsilon \dot{I}_{OBJ}$ . It goes through the atmosphere with the transmissivity  $\tau_{ATM}$ , so the radiant flux  $\varepsilon \tau \dot{I}_{OBJ}$  goes to the device. All surrounding objects emit the radiation to the examined objects and the examined object reflects the radiation. Afterwards, it goes to the atmosphere and the radiant flux entering the device is  $(1 - \varepsilon) \tau \dot{I}_{REF}$ . The atmosphere also has its own energy and the radiant flux of the atmosphere is  $(1 - \tau) \dot{I}_{ATM}$ . The summary of these three fluxes entering the device is called the equation of thermography and, using the knowledge referred to above, it is possible to determine the surface temperature of the examined object. Here, it is important to note that determining parameters, such as the emissivity of the object, the reflected temperature, and the transmissivity and temperature of the atmosphere are critically important for surface temperature measurement; insufficient description of these parameters could provide the main source of errors in thermography.

$$\dot{I}_{TOTAL} = \varepsilon \tau \dot{I}_{OBJ} + (1 - \varepsilon) \tau \dot{I}_{REF} + (1 - \tau) \dot{I}_{ATM} \quad (5)$$

$$\dot{I}_{OBJ} = \frac{1}{\varepsilon \tau} \dot{I}_{TOTAL} - \frac{1 - \varepsilon}{\varepsilon} \dot{I}_{REF} - \frac{1 - \tau}{\varepsilon \tau} \dot{I}_{ATM} \quad (6)$$

One of the devices which can capture thermal infrared radiation is a thermal infrared camera. TIR cameras are categorized according to their detector and range of wavelengths. TIR cameras which are used in common industrial applications are usually cameras with a microbolometer sensor [35] as a thermal detector with a wide range of wavelength detection (approx. 2.5  $\mu\text{m}$ –20  $\mu\text{m}$ ). The optics of TIR cameras are made from germanium [36]. The germanium lens steers the radiation to the sensor, which is a grid with resolution much lower than that of sensors of common RGB digital cameras. The resolution of the grid of the TIR cameras is usually from  $160 \times 120$  individual sensors to  $640 \times 480$  individual sensors [28]. Knowing the basic laws of thermography, the software of the TIR cameras recalculates the captured radiation to determine the temperature (possibly in  $^{\circ}\text{C}$ ).

## 2. Methods

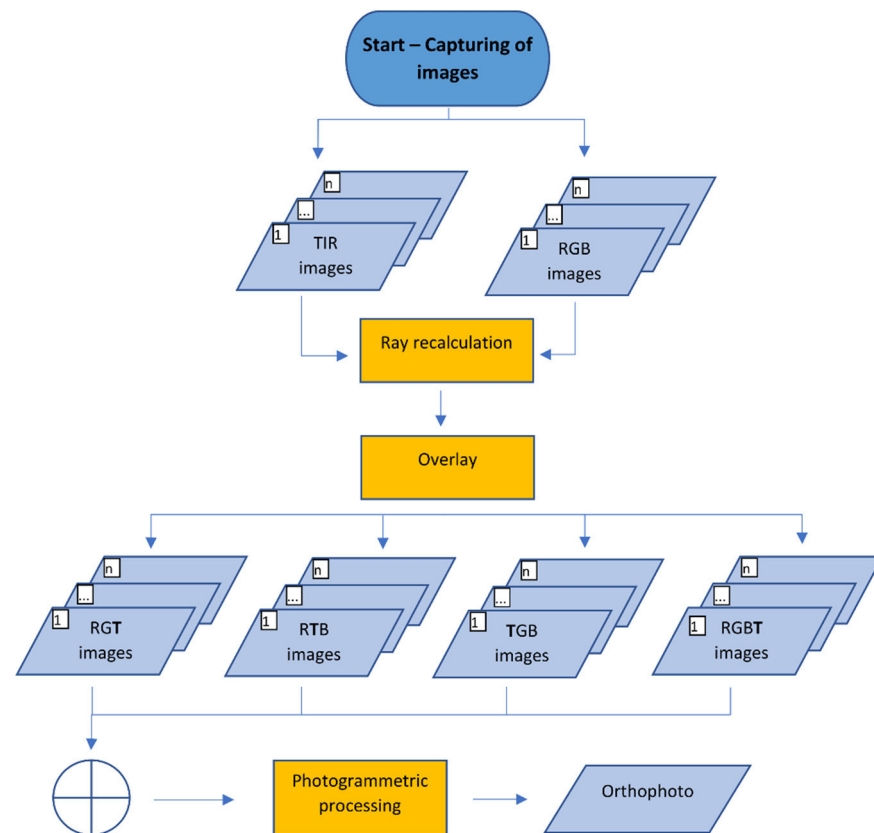
Two methods were proposed in [24,28]. In this section, the methods are described in detail.

### 2.1. Sharpening Method

The sharpening method was designed mainly to create an orthophoto or a texture augmented with thermal infrared information derived from original TIR images. The sharpening method is named after pansharpening but, in the sharpening method, panchromatic images are not used. TIR images are sharpened by the visual features from RGB images captured by common digital cameras. The sharpening method requires a fixed system of TIR and RGB cameras. For capturing the TIR image, it is necessary to capture its corresponding RGB image. In the sharpening method, the images are processed using the SfM method. So, when the images are captured, it is necessary to follow basic capturing instructions for the SfM method. The set of TIR images, and the set of the corresponding RGB images, are the input for the sharpening method. In Figure 3, the workflow of the sharpening method is presented.

At first, thermal infrared values (in  $^{\circ}\text{C}$ ) in TIR image pixels must be recalculated using a histogram stretching function to 8-bit (0–255) or 16-bit values (0–65,535). The minimum and the maximum of the thermal infrared values should be set the same for all TIR images of the dataset. The minimum and maximum values for the histogram stretching function can be set as actual extremes of the dataset or can be set as the minimum and maximum of the TIR camera thermal range. A large range between the minimum and the maximum can mean that the range of 8-bit values would not be enough (this depends on the TIR camera used). For example, TIR camera FLIR E95 has a temperature range from  $-20^{\circ}\text{C}$  to  $120^{\circ}\text{C}$  [37]. Recalculating the values to 8-bit values would mean a step of  $0.55^{\circ}\text{C}$ . This step is not enough for many applications especially when considering that the FLIR E95

camera's thermal sensitivity is around  $0.04\text{ }^{\circ}\text{C}$ . On the other hand, 16-bit values would mean a step of  $0.002\text{ }^{\circ}\text{C}$ . This is less than the thermal sensitivity of the example camera and, thus, there is no need to use a larger range of values than 16-bit.



**Figure 3.** Diagram of the sharpening method.

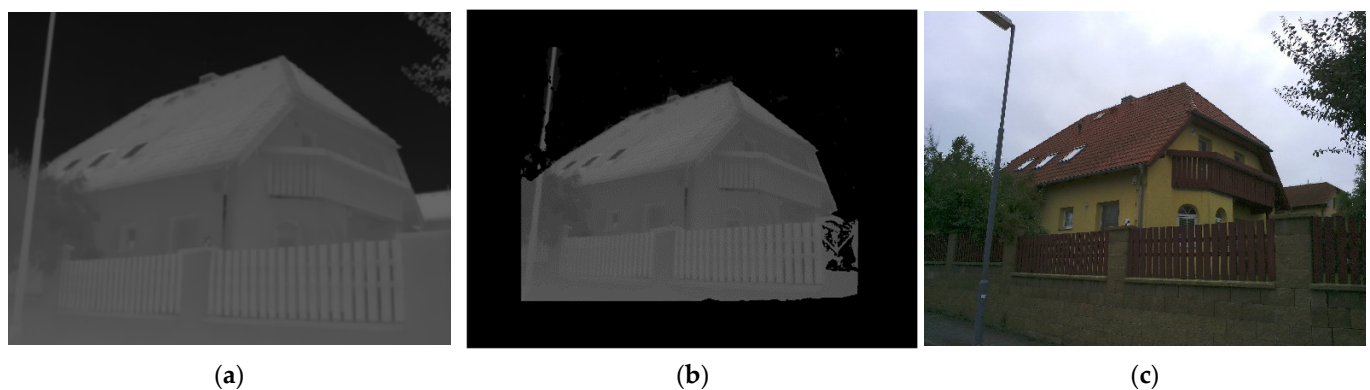
To overlay the TIR and RGB image and to create a sharpened image, it is necessary to transform the values of the TIR image to the RGB image. A first idea was to transform the information using plane transformation, such as affine or projective transformation. Previously published experiments [28] showed that any plane transformation is not sufficient for close-range purposes. Plane transformation is only sufficient when the images are more-or-less perpendicular to the object, captured over a longer and constant distance from the object, and when the depth of the images is more-or-less constant. However, the authors' interest is in terrestrial close-range photogrammetry, where there are expected short distances and many changes in the depth of the scene. Because previously published experiments [28] showed that plane transformation is insufficient, a ray recalculation method has been designed.

The ray recalculation method requires parameters of the interior orientation of the TIR and RGB cameras, the relative rotation and translation between the RGB camera and the TIR camera of the fixed system, and depth maps of the RGB images. To obtain depth maps of the RGB images, the RGB images must be preprocessed solely by any photogrammetric method, possibly the SfM method. When the 3D triangulated model or the point cloud is generated and scaled, the depth maps are rendered onto one of the generated results.

Having all the required inputs, the first RGB image of the dataset is taken. Then a blank raster with the same dimensions as the RGB image taken is created. The goal is to remap the corresponding TIR image to the blank raster image and to fill the blank pixels with the thermal infrared information from the TIR image. The process of the remapping is described below.

For the first pixel of the RGB image, a ray from the optical center of the RGB image through the undistorted position of the pixel is computed. A point which lies on the scene,

and which is projected to the pixel of the RGB image, must be found on the ray. The ray itself is infinite, but knowing the depth information derived from the corresponding pixel of the depth map, the 3D coordinates of the point on the scene in the RGB image camera coordinate system is possible to compute using the basic similarity of triangles. The coordinates of the point are possible to transform to the TIR camera coordinate system (the rotation and translation between the RGB camera and the TIR camera for a fixed system is known). The point is then reprojected from the TIR camera coordinate system to the TIR image (the internal orientation of the TIR camera is also known). The pixel value of the thermal infrared (the TIR image is already recalculated to 8-bit or 16-bit values) is stored to the pixel of the blank raster (the same position as the examined pixel of the RGB image). The process is then repeated for every pixel of the RGB image and the remapped TIR image is created (see Figure 4).



**Figure 4.** (a) Original TIR image; (b) TIR image remapped by a ray recalculation function with the same resolution as RGB image; (c) Original RGB image; the remapped TIR image and RGB image can be overlaid and merged.

Then, the remapped TIR image with 8-bit values can replace one of the channels of the RGB image (with 8-bit channels) and create the sharpened image. For example, from the RGB image, the blue channel is replaced by the remapped TIR image which creates the RGT image (see Figure 5a), where visible features of the scene (channels, R and G) and thermal infrared information (in channel T) are partly stored. This could be applied analogously and the RTB or TGB sharpened image can be created (see Figure 5b,c). The sharpened image should be stored in an image format without compression so as not to change the data in the T channel. It is possible to process these 3-channel images in any photogrammetric software because the images have the same structure as common RGB images. In advanced photogrammetric software e.g., Agisoft Metashape [38], there is support of multi-channel images with, e.g., 16-bit values of a band. Then, it is possible to use the remapped TIR image with 16-bit values and to add the channel to the RGB image, which must also be recalculated to 16-bit. Then the RGBT (Figure 5d) image, with full visible features (channels RGB) and thermal infrared information (channel T), is created. Every image of the dataset should be sharpened under the same scheme (e.g., RGT). The fact that the visible information from RGB images is present in all combinations (at least partly) helps the algorithms to better find the key and tie points and eliminates the problem with photogrammetric targeting. In this case, standard photogrammetric targets can be used and are easily identifiable in the images. The question of which combination is the most sufficient to use for SfM processing has been raised and, to answer the question, an experiment was performed (see Section 3.1).



**Figure 5.** Sharpened images: (a) RGT combination; (b) RTB combination; (c) RGT combination; (d) RGBT combination.

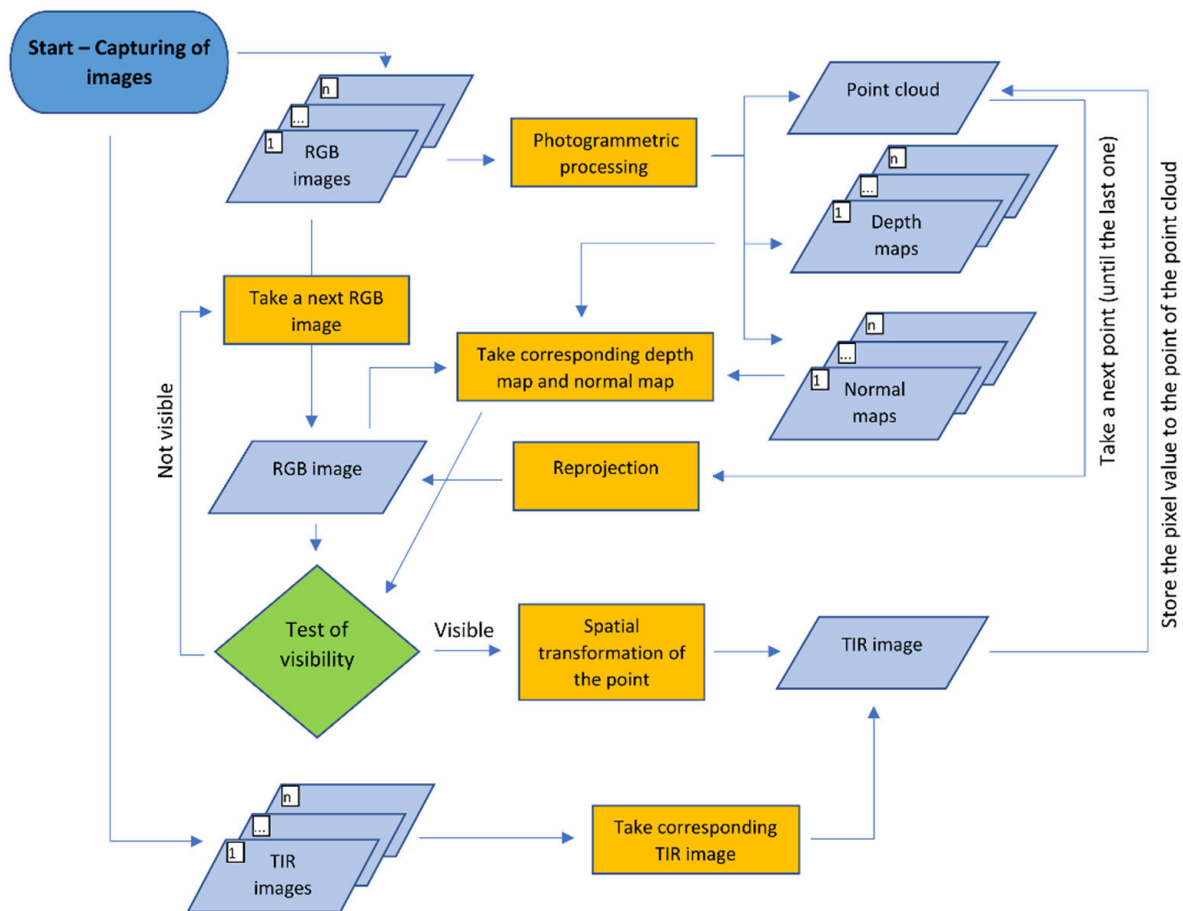
When the dataset of sharpened images is ready, the process continues with standard photogrammetric processing with the SfM method. The results of the processing can be a colored point cloud, a textured 3D model or an orthophoto. The color of the point cloud is usually weighted in some way, meaning that we would not have the original values from the TIR image. When creating texture for the 3D model or the orthophoto, usually the software has an option to prevent averaging or weighting of the colors. Because of this, the sharpening method is more suitable for an augmented orthophoto or textured 3D model. In the case of an orthophoto and texture of the 3D model, it is possible to recalculate the values of the T channel back to the original thermal infrared information. By adding the original thermal infrared information to the pixels, the spatial results augmented by the thermal infrared information are created. For a point cloud, augmented by the thermal infrared information, it is more convenient to use the reprojection method.

## 2.2. Reprojection Method

The reprojection method was designed primarily to create a point cloud augmented with thermal infrared information which is derived directly from a TIR image. The sets of TIR images and corresponding RGB images captured by the fixed camera system are the inputs of the reprojection method. The images should be captured according to the common suggested capturing scenarios for SfM. Relative rotation and translation between the RGB and TIR camera coordinate systems, and the parameters of the interior orientation of both cameras gathered by geometric camera calibration, must be known. In Figure 6, the workflow of the reprojection method is shown.

The first step of the reprojection method is SfM processing of the RGB images only. This is an essential advantage of this method, because the TIR images do not enter SfM processing at all; thus, all the negative factors referred to above are naturally eliminated. After SfM processing, the resulting photogrammetric model is scaled or transformed by control points. After scaling, the point cloud is generated. For the subsequent steps of the reprojection method, it is necessary to generate the depth map and normal map of each processed image. The depth maps and normal maps can be rendered on the point cloud, or on the 3D triangulated model generated based on the point cloud.

Then every point of the point cloud should be augmented by thermal infrared information. The algorithm takes a first point of the point cloud and reprojects the point to the camera coordinate system of the first taken RGB image from the set of RGB images which were used in SfM processing (the external orientation parameters of the images are known). Knowing the parameters of the interior orientation of the RGB camera, the point is transformed from the camera coordinate system to the RGB image. Even though the point is possible to reproject to the RGB image, this does not mean that the point belongs to the scene in the image (the point can be physically behind the image scene). To check if the point is visible in the image, it is necessary to carry out the test of visibility. The test of visibility requires information about the scene. The information about the scene is given by the depth map and the normal map of the image. The correct test of visibility is the key part of the reprojection method.



**Figure 6.** Diagram of the reprojection method.

For the test of visibility in the reprojection method, four tasks have been designed. The first task is to check if the projection of the point of the point cloud is within the RGB image dimension. Knowing the exterior and interior orientation of the RGB image, the point is simply reprojected to the RGB image and the coordinates in the images are read and compared. If the reprojected point is within the image dimension, then the depth test is carried out. For the depth test, it is necessary to create a depth map of the RGB image. The position of the reprojected point of the point cloud to the RGB image is used also for the depth map image and the depth value from the corresponding pixel is read. Then the test compares the Z-coordinate of the point transformed to the RGB camera coordinate system to the depth value from the pixel. If the difference is within a certain tolerance, the point is still considered as potentially visible. The question is how to set the tolerance. The depth test is easy to implement, and it is very efficient for points which are not clearly visible; however, the reliability of the test decreases when the distance differences are closer to the tolerance. In close-range applications, the tolerance will have an important effect when deciding the visibility of the point (especially parts behind the corners and on edges of structures, such as windows and door jambs). Because of this uncertainty, a normal test as a third task of the test of visibility should be added. The third task of the test of visibility is the normal test. For the normal test, the normal maps of RGB images must be rendered. In each pixel of the normal map the normal vector converted to RGB values is stored. For the normal test, it is necessary to also have a normal vector of each point of the point cloud. In common photogrammetric software, it is possible to store the normal vector of each point to the point cloud. The laser scanning results also provide a normal vector of each point. When the normal vector is not present in the point cloud, it is possible to calculate it in other software, such as Cloud Compare [39]. The normal vector of the point taken of the point cloud must be transformed to the camera coordinate system of the RGB image. If the



point of the point cloud is reprojected to the RGB image within the image dimension, then the image coordinates are used, and the pixel value of the normal map is read. The normal vector of the point is then compared to the values read in the normal map. The vectors are compared according to cosine similarity. If the angle between the two 3-dimensional normal vectors is below a certain tolerance, the point of the point cloud is still considered as visible, and the visibility test continues with the last task. Theoretically, the normal test can be removed, but it requires a very accurate point cloud and very accurate parameters of exterior orientation of the RGB images. As a fourth task, the point of the point cloud is transformed from the camera coordinate system of the RGB image to the camera coordinate system of the TIR image and the point is reprojected to the TIR image. If the reprojected point is within dimensions of the TIR image, the point is considered as visible, and the process of the reprojection method continues for other points of the point cloud, or for other images.

If the test of visibility proves that the point is truly visible on the RGB image, then, knowing the relative rotation and translation between the RGB and TIR cameras, the point is transformed from the RGB camera coordinate system to the TIR camera coordinate system of the corresponding TIR image. From the TIR camera coordinate system, the point is projected to the TIR image. The value from the pixel where the point was projected is read and stored to the point of the point cloud. Then the point is reprojected to the second RGB image of the RGB images set and so on. If the point is visible in more than one image, more thermal infrared values to the point of the point cloud are stored. The final thermal infrared value is then calculated as a mean of the set of values; there is also the possibility to statistically detect outliers and remove them from the set of values.

The advantage of this method is that, in the process, it is possible to replace the point cloud generated from RGB images with a laser scanning point cloud referenced in the same coordinate system. A point cloud from a laser scanner can have higher accuracy or density compared to the point cloud generated from RGB images. Then, the process remains unchanged and the result of the reprojection method is augmented by a laser scanning point cloud. When a laser scanning point cloud is used for the reprojection method, the depth and normal maps should be rendered on a 3D triangulated model, because the model is continuous. The main disadvantage of the reprojection method is that the computation of the method (when there are a lot of points of the point cloud and many images) takes many hours. To linearly reduce the processing time, it is not necessary to use all the RGB images in the set, but it is enough to use only selected images. For example, to process the point cloud of a building façade, it is not necessary to use all 30 images from which the point cloud is created, but it is sufficient to manually select just 3 or 4 images, if it is certain that each point is visible in at least one image. However, this requires manual selection input which precludes automatization of the processing.

### 3. Experiments

In Section 2, several questions were raised regarding the presented methodologies. To answer these questions, experiments were carried out and are presented in this chapter. For the experiments, TIR camera FLIR E95 was chosen. FLIR E95 is a camera equipped with a thermal infrared sensor and an RGB sensor and a fixed camera system is created which is required for the processing using sharpening and reprojection methods. The general parameters of FLIR E95 cameras are detailed in Table 1. Both cameras of FLIR E95 were geometrically calibrated using a three-level spatial calibration field. The relative translation and rotation between the RGB and TIR camera coordinate systems was also computed on the three-level spatial calibration field. The result of geometrical calibration and the relative translation and rotation is presented in Tables 2 and 3. The process of gathering the resulting parameters, as well as the impact of the accuracy of the resulting parameters to the transformation processes of the suggested methods, is presented in [28]. The TIR camera was radiometrically calibrated by the manufacturer.

**Table 1.** Parameters of TIR camera and RGB camera of FLIR E95 [28].

Parameter	TIR Camera	RGB Camera
Focal length (mm)	10	3.29
Sensor size (mm)	7.89 × 5.92	3.67 × 2.74
Resolution (pix)	464 × 348 pix	2592 × 1944
Size of a pixel (mm)	0.017 mm	0.0014 mm
FOV (°)	42° × 32°	53° × 41°

**Table 2.** Resulting parameters of the interior orientation of TIR camera and RGB camera of FLIR E95, determined on a three-level calibration field [28].

		c [pix]	Px [pix]	Py [pix]
TIR camera	3 level cal. field	593.5	−3.3	1.4
RGB camera	3 level cal. field	2481.4	−23.4	27.1

**Table 3.** Resulting parameters of relative translation ( $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$ ) and relative rotation ( $\Delta\omega$ ,  $\Delta\phi$ , and  $\Delta\kappa$ ) [28].

	$\Delta X$ [m]	$\Delta Y$ [m]	$\Delta Z$ [m]	$\Delta\omega$ [°]	$\Delta\phi$ [°]	$\Delta\kappa$ [°]
3 level cal. field	−0.0002	−0.0248	−0.0065	−0.833	−0.061	−0.007

### 3.1. Sharpening Method—Test of Image Band Combinations

The question of which combination of image bands is the most suitable for SfM processing has been raised. An experiment was carried out. Three image samples were chosen, with each image sample capturing a different part of the object which was a family house. Initially, the RGB images of three samples were processed solely in Agisoft Metashape. The RGB images were masked to have exactly the same computational area as the masked images of different combinations (i.e., RGT, RTG, TGB, RGBT). The images were co-registered using their full resolution. Then the point cloud was generated. The number of tie points (TP) and the number of points in the point cloud (PC) were noted. The process was repeated in the same way for all four combination of image bands (also masked), and the numbers of TP and PC were compared as percentages (see Table 4) to the numbers resulting from the RGB image processing. As a last step, the TIR images were also converted to .bmp image format and processed, and the numbers of results were also compared. In Agisoft Metashape, it is possible to set a key point or tie point limit for co-registration. In our case, no parameter limit was set, so, in Table 4 the maximum possible tie points are provided.

**Table 4.** The number of tie points and points of the point cloud from processing of samples of different band combination in comparison to the numbers of processing of RGB images.

		RGT	RTG	TGB	RGBT	TIR
Sample 1 (25 images)	TP	96%	65%	82%	77%	7%
	PC	98%	81%	93%	88%	4%
Sample 2 (10 images)	TP	100%	89%	76%	99%	9%
	PC	98%	84%	99%	81%	6%
Sample 3 (18 images)	TP	95%	75%	84%	84%	12%
	PC	97%	92%	84%	89%	3%

According to Table 4, it is possible to see that the image band combination RGT gave the best results. Most of the numbers are almost the same as the results given by processing the RGB images. The remaining combinations gave numbers around 80% of the numbers for RGB image processing. The RGBT combination compared to the RGT combination gave



the worse results, but the advantage of a lesser temperature step and full colour information still prevailed there. The processing of TIR images showed a very low number of tie points and points of the point cloud. It must be noted that the resolution of TIR images is lower than the resolution of RGB images so that the comparison is not fair. The numbers are still presented to illustrate that some co-processing with RGB images is necessary. Even though all TIR images were co-registered, the number of tie points was very low. The point cloud generated based on TIR images was too thin and insufficient even on first impression.

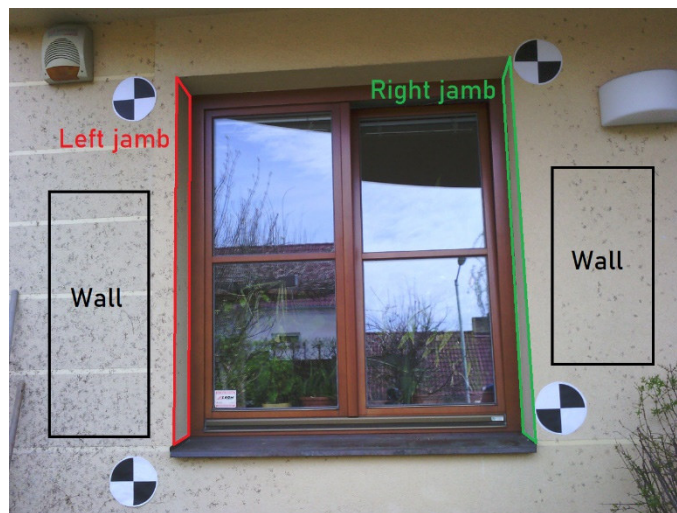
### 3.2. Test of the Visibility–Experiment

The aim of this experiment was to determine what is the best way to set the tolerances of the test of the visibility. For this experiment, part of the building façade was chosen as a test scene. The scene was captured by thermal infrared camera FLIR E95 and for each corresponding RGB image, the TIR image was captured. The methods of this article are focused on close-range photogrammetry, so the façade was captured at a distance of 2–4 m. The average sample distance of the images was approximately 1.3 mm/pix. Overall, 24 images were captured. On the façade, four photogrammetric targets were placed and used as control points. The targets were measured by a total station just from one station. The images were processed using the SfM method with the commercial photogrammetric software Agisoft Metashape. For image co-registration, full image resolution was used. The average tie point residuum in the images was 0.93 pix. The photogrammetric model was transformed to the targets and the RMSE<sub>xyz</sub> on control points was 0.001 m (approx. 1 pixel). After georeferencing, the point cloud and 3D textured model were generated. Again, the full resolution of the images was used for point cloud generation. The experiment was performed on the points of the point cloud. The depth maps and normal maps have been rendered on a 3D model.

The tolerances in the depth test could have been set according to the accuracy indicators of the photogrammetric model, e.g., sample distance, RMSE values of the control or check points. In this experiment, the tolerances were set according to the RMSE<sub>xyz</sub> on the control points. The RMSE<sub>xyz</sub> on the control points was 0.001 m. The tolerances have been set as RMSE<sub>xyz</sub> multiplied by 5, 10 and 15. The tolerances for the experiment were 0.005 m, 0.010 m, 0.015 m. It is more difficult to decide the angle tolerance in a normal test. The normal vectors usually have higher variance, especially at surfaces of the model, which are either textureless or not ideally covered with sufficient image overlap or by the images. It is possible to set the tolerance by a rough estimate or by deeper examination of the variance of the normal vectors of the point cloud. To examine the variance of the normal vectors on the point cloud, three groups of the point cloud were manually created. Only points which represented a plain façade wall (W) belonged to the first group. Only points which represented left window jamb (LB) and right window jamb (RB) (see Figure 7a) belonged to the other two groups. All three parts were considered as planar and similar normal vectors were expected at all points of each group. In each group of points, an average normal vector, as a mean of the  $x$ ,  $y$ ,  $z$  components of the normal vector of all points of the group, was computed. Then, every normal vector of every point of the group was compared to the average normal vector by cosine similarity and the angle between the vectors was calculated. The distribution of the set of computed angles is presented in a cumulative distribution graph in Figure 8.

According to Figure 8, the normal vector of the points varied. On the part of the point cloud which represents the wall, the variance of the normal vector was the lowest and the angle tolerance for the normal test could be set between 20°–30° to cover most of the points. A more problematic situation was on the jambs. These parts did not have ideal intersection angles in the images, so the variance of the vectors was higher. To cover around 80% of the points, the angle tolerance should be set between 50°–70°. This wide-angle tolerance can lead to wrong indications if the point is visible or not. For practical application, it is more important to ensure correct indication of the visibility even when a certain percentage of

points is lost, rather than to indicate the point as visible when it is not. That would lead to incorrect thermal augmentation of the points and the final result would be misleading.



(a)

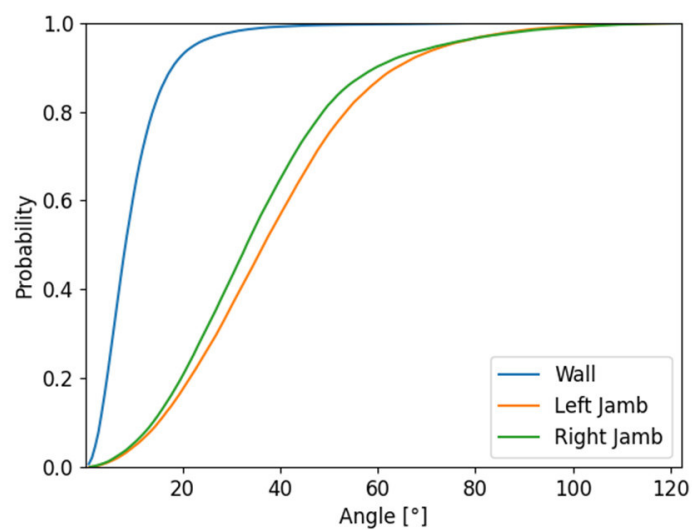


(b)



(c)

**Figure 7.** (a) Centre image. In the figure, the positions of the points of the three groups of point cloud (wall, left jamb, right jamb) are shown; (b) Left image; (c) Right image.



**Figure 8.** Cumulative distribution graph of set of angle differences.

For the following experiment, three images of the dataset were chosen. The left image captures the window from the left side, the center image is perpendicular to the façade and the right image captures the scene from the right side. For each image, points of the groups

were reprojected to the images and the visibility test with various tolerances was carried out. The results of the experiment of the visibility test are presented in Table 5. In the table, when the group of points is visible in the image is noted. Then, the depth tests and normal tests were carried out with various tolerances. According to the  $RMSE_{xyz}$  on the control points on the photogrammetric model, 5 mm, 10 mm and 15 mm tolerances were set for the depth test. According to the examination of the variances of the normal vectors of the point cloud, 25°, 40° and 60° tolerances for the normal test were set.

**Table 5.** Testing of tolerances for the depth test and the normal test.

Reality			Depth Test (Tol = 0.005)	Depth Test (Tol = 0.01)	Depth Test (Tol = 0.015)	Normal Test (Tol = 25°)	Normal Test (Tol = 40°)	Normal Test (Tol = 60°)	Depth Test (Tol = 0.01) and Normal Test (Tol = 40°)
Left Image 13,233	W	Visible	100%	100%	100%	98%	99%	100%	99%
	LJ	Not Visible	1%	3%	4%	3%	10%	27%	1%
	RJ	Visible	84%	96%	99%	43%	68%	87%	66%
Centre image 13,239	W	Visible	100%	100%	100%	95%	99%	100%	99%
	LJ	Visible	17%	33%	49%	22%	48%	76%	16%
	RJ	Visible	6%	13%	21%	25%	50%	78%	7%
Right image 13,217	W	Visible	72%	100%	100%	96%	99%	100%	99%
	LJ	Visible	86%	95%	97%	48%	72%	88%	70%
	RJ	Not Visible	0%	1%	2%	3%	10%	25%	0%

From Table 5, it is possible to observe high confidence when the wall group of points was reprojected. Both tests for all images show that they were very successful even using the lowest tolerances.

The depth test showed the worst results when testing the wall group of points reprojected in the right image with the lowest tolerance (72%). The image captures the group of points at a very acute angle. The capturing angle has an effect on the results of the test of visibility, especially for the depth test. This effect is possible to observe at the left jamb and right jamb points in the center image. Even though all the points of the left jamb and right jamb group are visible, the depth test was able to recognize the visibility only in 49% of points on the left jamb when the largest 15 mm tolerance was set. The right jamb is captured at an even more acute angle, so the depth test recognized the visibility only in 21% of points. When the points of the left jamb were tested in the right image (the right image axis is approximately perpendicular to the plane of the left jamb), the depth test was successful in 97% of points, and when the points of the right jamb were tested in the left image (also perpendicular angle), the depth test was successful in 99% (using the 15 mm tolerance) of points. The testing showed that when the image captures the group of points with normal vector towards the camera, the depth test is confident even with the middle tolerance of 10 mm.

The normal test looks independently at the angle of capture of the object. According to the results, the normal test is dependent of the variance of the normal vector in a group of points. The results correspond to the cumulative distribution function in Figure 8. For example, the normal test on points of the left jamb with a tolerance of 40° showed 48% in the center image and 72% for the right image, respectively. According to the cumulative distribution function, the expected value was 56%. Because of the large variance of the normal vector on jambs, the normal test in some cases falsely determined points not visible as visible. With the 60° tolerance, the normal test wrongly determined 27% of the points on the left jamb in the left image and 25% of the points on the right jamb in the right image. Fortunately, in these cases, the depth test is reliable in detecting not visible points, and, when both tests were applied together, there was only 1% of points wrongly determined as visible, and 0% on the right image, respectively.

To conclude, it is not straightforward to set the proper tolerances for the depth and normal tests. It is not possible to give general advice. The tolerances depend on each project, on its purpose and the demands placed. It is not possible to set a general rule.

For some of the projects it is possible to set low tolerances for the test. In this case, a lot of points of the point cloud could be lost. However, in some of the applications, where it is important to obtain accurate temperature data, it is a better solution to lose a certain amount of points rather than to have the full recalculated point cloud with potentially wrong temperature values. It is advisable to set lower tolerances when there is a larger dataset of testing images.

### 3.3. The Sharpening and Reprojection Method Used on the Test Object

To test the proposed sharpening and reprojection methods, an experiment was carried out. As a testing object, a family house was chosen (see Figure 9). The family house is inhabited, so parts of the building are covered by vegetation and other regular objects. The family house facade is isolated, so there were not expected to be temperature changes. The family house is colored in one color and the facades are textureless. Due to these two facts, there was presumed to be a difficulty with processing of images using the SfM method. However, the object was chosen because it represents a common real object. On the object black and white targets were placed for scaling and orientation of the results. The targets were measured by a total station.



**Figure 9.** The test object.

#### 3.3.1. Capturing the Images

The process of capturing the images was carried out according to the suggestion of the local distributor of the FLIR E95 camera. At first, a day with cloudy weather and stable air conditions was chosen. As input to the camera, multiple parameters were measured and set in the TIR camera. The air temperature and humidity were measured using a calibrated thermometer and hygrometer. The reflected ambient temperature is a very important parameter. Crumpled aluminum foil was placed on the object and the average temperature of the foil was measured by the FLIR E95 camera and used as the reflected ambient temperature. The reflected ambient temperature was determined at 3.1 °C and the main source of the reflectance was the cloudy sky. Because the weather was constant during capturing of the images, the reflected ambient temperature was expected also to be constant. This was confirmed by measurement at the end of image capturing when the reflected ambient temperature was measured at 3.0 °C. After parameter setting, the object was captured by the FLIR E95 camera in different positions and angles to cover every part of the object in multiple images. For every TIR image, a corresponding RGB



image was captured. In total, 187 image pairs of corresponding images were captured. The average sample resolution of the RGB image was 3.2 mm/pix. On the object, five surface temperature check points were established and the surface temperature of the façade was measured using a calibrated touch thermometer. The temperatures were compared to the temperatures in the captured TIR images. The values did not exceed 0.9 °C which is below the accuracy of reading the temperature of the FLIR E95 camera signifying that all parameters were set correctly.

### 3.3.2. Processing Using the Sharpening Method

The sharpening method was used to create TIR orthophotos of the facades of the objects with the temperature value stored in each pixel of the orthophotos.

Initially, only RGB images were processed in the Agisoft Metashape software. For the relative orientation computing, the full resolution of the images was used. The average residuum of tie points in the images was 1.3 pix. To scale the model, the relative orientation was referenced using 10 control points. The  $RMSE_{xyz}$  of the referencing was 6.5 mm. After referencing, the point cloud was generated. It was obvious that some parts of the building with very low texture were not ideally described by the point cloud. To partially improve the result, lower filtering was used when the point cloud was generated. From the generated point cloud, the 3D triangulated model was generated. The model was deformed in only some of the textureless parts. On the model, the scaled depth maps of the images were generated. It was necessary to prepare the following data:

- The set of RGB and TIR images;
- Parameters of the interior orientation of the RGB camera and the TIR camera;
- Relative translation and rotation between the RGB camera and the TIR camera;
- Depth maps for the RGB images.

The group of TIR images, the corresponding RGB images and the corresponding depth maps were prepared for the process of ray recalculation. Based on the created depth maps and the known parameters of the interior orientation of both cameras and the known relative translation and rotation between the cameras, 4-channel RGBT images were created. In the fourth channel, temperature information recalculated using the histogram stretching function to 65,536 values (16-bit) was stored. The recalculated images were processed by Agisoft Metashape. First, the relative orientation was computed. In some pixels of the RGBT image there was a “no data” value in the fourth channel. Those parts were covered by a mask and then there were no calculated tie and key points. The model of the relative orientation was referenced using five control points (equally distributed on the object). Five check points were used for the accuracy check. The average tie point residuum in the images was 1.9 pix. The  $RMSE_{xyz}$  of the referencing was 0.6 pixels, which signifies relatively good accuracy of the relative orientation. However, the  $RMSE_{xyz}$  on check points was 17 mm. This error was probably caused because measurement of the coordinates of the targets by total station was carried out from multiple stations. After the accuracy check, the 3D textured model was created and, using the model, four orthophotos (one for each façade) were generated. To achieve unchanged temperature values, every averaging of the pixel values (colors) was disabled during orthophoto generation. The fourth channel of the RGBT orthophoto was separated and recalculated back to the temperature values, using the inverse histogram stretching function, and TIR orthophotos were created (see Figure 10). The TIR orthophotos were visualized in QGIS [40]. Some parts of the façade, which were covered by some other object, were cropped on the orthophotos.



**Figure 10.** The resulting orthophotos augmented by thermal information.

### 3.3.3. Processing Using Reprojection Method

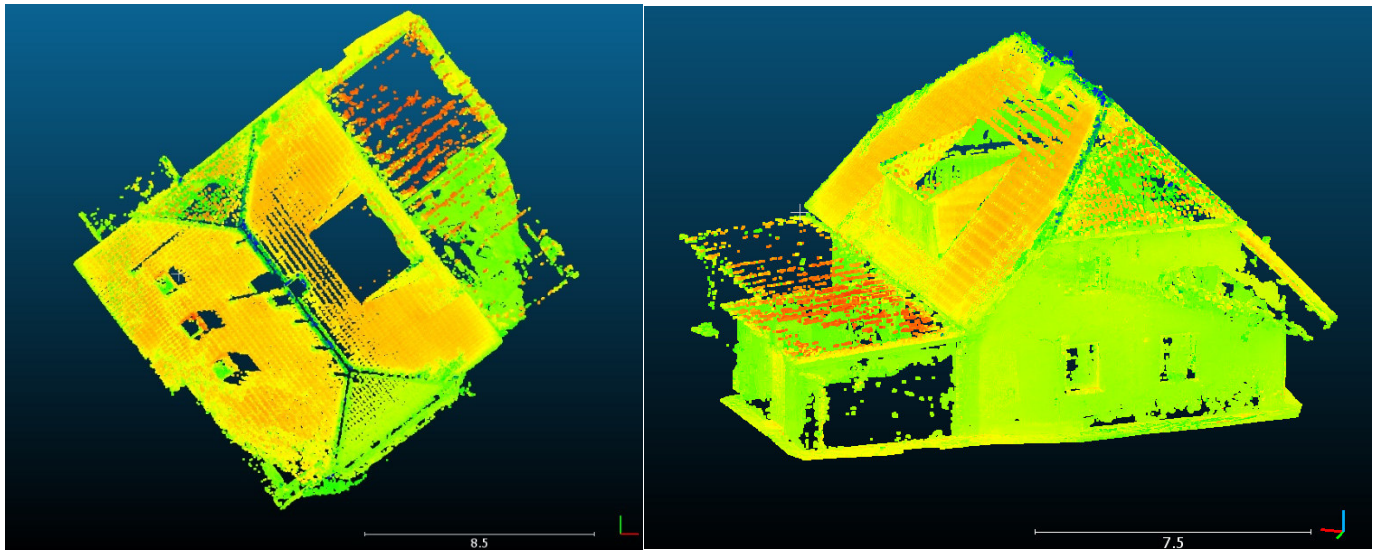
The reprojection method was used to create a point cloud of the object where a temperature value is stored in points of the point cloud. The reprojection method can use a point cloud generated by photogrammetric processing or a point cloud acquired by laser scanner. For this experiment we chose a point cloud acquired by laser scanner.

The testing object was laser scanned using a Leica BLK360 laser scanner. The point cloud from the laser scanner was processed in a Cyclone Register 360 and the point cloud was referenced to the seven control points. The  $RMSE_{xyz}$  of the referencing on the control points was 0.012 m. It was then necessary to process the RGB images to obtain their absolute exterior orientation. The model from RGB images was already computed for the sharpening method and the details are presented in Chapter 3.2.2. From that model the scaled depth maps and normal maps were created. The following data were prepared:

- Set of RGB and TIR images
- Parameters of the interior orientation of the RGB camera and the TIR camera
- Relative translation and rotation between the RGB camera and the TIR camera
- Referenced point cloud acquired by laser scanner
- Exterior orientation of RGB images
- Depth maps of RGB images
- Normal maps of RGB images

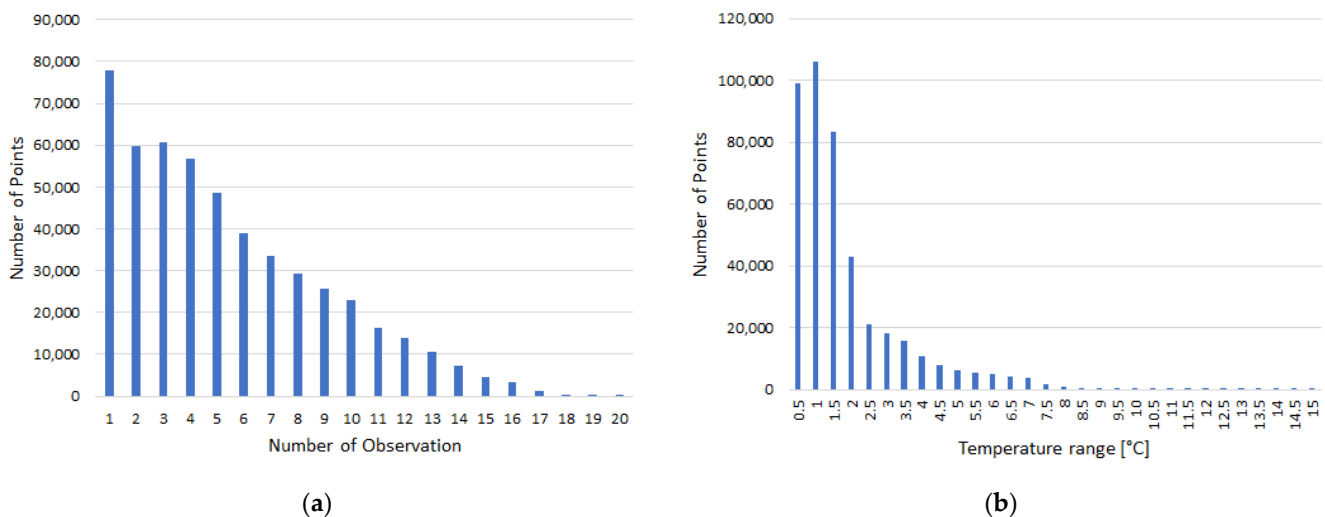
To reduce the processing time of the reprojection method, the point cloud was subsampled to 0.02 m distance between the points, so the point cloud had 1.2 million points in the end. Moreover, only half of the images (94 RGB images) were chosen from the RGB image dataset. According to the previous experiment, where the tolerances were discussed, the tolerance for the depth test was set at 0.025 m ( $RMSE_{xyz}$  on control points in photogrammetric model multiplied by five). The tolerance for the normal test was set at  $25^\circ$ . The tolerances could have been set low, because the reprojection method was carried out using a substantial number of images. Then the reprojection process was started. The processing took 12.3 h. The processing time was too high even for a significantly subsampled point cloud. The final resulting point cloud augmented by the temperature information had 511,622 points, which was 42% of the inputted point cloud (see Figure 11). The points were removed either because of low tolerances, or because of imperfections of the 3D model which was used for depth and normal map rendering. Some parts of the model (where the

object did not have any texture) were visibly wrong mainly because of a lack of point cloud points from the photogrammetric processing.



**Figure 11.** The resulting augmented point cloud visualized in Cloud Compare.

The augmented point cloud was analyzed. From the total 511,622 augmented points, most of the points (15%) were visible in just one TIR image. According to Figure 12a, 85% of points were visible in more than one image (of a set of 94 images), so, in these cases, the temperature information was determined as a mean of the multiple values. In the case of multiple values, the ranges between the minimum temperature value and the maximum temperature value for each point were analysed. According to Figure 12b, 92% of the observed ranges were under 4 °C, which corresponds to the temperature measurement accuracy of the TIR camera used ( $\pm 2$  °C) [37]. The remaining points can be considered as points with potentially incorrectly assigned temperature, but the high range could be caused by just one outlier.



**Figure 12.** (a) Histogram of number of observation values on points of the point cloud. (b) Histogram of temperature range values on points of the point cloud.

#### 4. Conclusions

The aim of this article was to present and describe the proposed sharpening and re-projection methods in detail. Both implemented methods were also tested by experiment



(see Section 3.3). Each method has its problematic questions which have been raised. Experiments were carried to try to answer these questions.

The sharpening method is easy to implement, fast, and suitable mainly for augmented orthophoto and texture of 3D model generation. In the case of the sharpening method, the question of which image band combination is most suitable to use has been raised. According to the experiment described in Section 3.1, the best combination appears to be the RGT combination. This combination obtained slightly fewer tie points compared to the processing of RGB images. When the color information of the result is not needed and when the difference between the minimum and maximum temperature in the dataset is not too large, the RGT combination appears to be a good option. However, the RGT 8-bit combination may not be most suitable for many applications. In the case of a larger difference between the minimum and maximum temperature in the dataset, the 8-bit values may not be enough and a 16-bit RGBT combination should be used. This combination also obtained a respectable number of tie points compared to the RGB images. The RGBT combination also has an advantage in that the full-color information is stored in the resulting augmented orthophoto.

The reprojection method was designed for augmented point cloud generation. Compared to the sharpening method, the processing time is (so far) too high. Among the advantages is that the method can also process laser scanning data (the sharpening method relies on a point cloud generated by photogrammetry). In the reprojection method, the test of visibility is an innovative element. The test of visibility is the key process in the reprojection method. Without the proper visibility test, the final augmentation of the point cloud would be incorrect or misleading. There are many questions regarding the visibility test. The main question is how to set the tolerances in the depth test and normal test. An experiment on the visibility test was carried out and is presented in Section 3.2. The experiment did not provide general rules on how to set the tolerances but gave ideas on how the tolerances can be determined. The tolerances depend on the concrete case. Sometimes it is necessary to set low tolerances, to lose a significant number of points, and to have the certainty that every point will be augmented by correct temperature information. Sometimes it can be more important to preserve every point of the point cloud even when the augmented temperature information may be wrong. In this case, it is possible to use the advantage of the reprojection method, that every point can be augmented by multiple values, and then the incorrect values can be determined and eliminated by outlier detection.

This article sought to present the methods and to raise the main questions regarding the methods. In future research, more extensive testing of the methods on selected objects with different characteristics will be performed.

**Author Contributions:** Conceptualization, A.D., K.V. and K.P.; methodology, A.D.; software, A.D.; technical advisory and supervision, K.V. and K.P.; data processing, A.D.; investigation, K.V. and K.P.; writing—original draft preparation, A.D.; writing—review and editing, A.D.; visualization, A.D.; project administration, K.V. and K.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Institutional grant CTU in Prague, SGS2022 nr. OHK1-077/22.

**Acknowledgments:** This research is partly research of the Department of Geomatics at the Czech Technical University in Prague and partly internal research of EuroGV spol., s r. o.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Usamentiaga, R.; Venegas, P.; Guerediaga, J.; Vega, L.; Molleda, J.; Bulnes, F.G. Infrared Thermography for Temperature Measurement and Non-Destructive Testing. *Sensors* **2014**, *14*, 12305–12348. [[CrossRef](#)]
2. Gade, R.; Moeslund, T.B. Thermal cameras and applications: A survey. *Mach. Vis. Appl.* **2014**, *25*, 245–262. [[CrossRef](#)]

3. Kylili, A.; Fokaides, P.A.; Christou, P.; Kalogirou, S.A. Infrared thermography (IRT) applications for building diagnostics: A review. *Appl. Energy* **2014**, *134*, 531–549. [[CrossRef](#)]
4. Adamopoulos, E.; Volinia, M.; Giroto, M.; Rinaudo, F. Three-Dimensional Thermal Mapping from IRT Images for Rapid Architectural Heritage NDT. *Buildings* **2020**, *10*, 187. [[CrossRef](#)]
5. Hoegner, L.; Tuttas, S.; Xu, Y.; Eder, K.; Stilla, U. Evaluation of methods for coregistration and fusion of rps-based 3d point clouds and thermal infrared images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B3*, 241–246. [[CrossRef](#)]
6. Javadnejad, F.; Gillins, D.T.; Parrish, C.E.; Slocum, R.K. A photogrammetric approach to fusing natural colour and thermal infrared UAS imagery in 3D point cloud generation. *Int. J. Remote Sens.* **2020**, *41/1*, 211–237. [[CrossRef](#)]
7. Sledz, A.; Unger, J.; Heipke, C. Thermal IR imaging: Image quality and orthophoto generation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-1*, 413–420. [[CrossRef](#)]
8. Wakeford, Z.E.; Chmielewska, M.; Hole, M.J.; Howell, J.A.; Jerram, D.A. Combining thermal imaging with photogrammetry of an active volcano using UAV: An example from Stromboli, Italy. *Photogramm. Rec.* **2019**, *34*, 445–466. [[CrossRef](#)]
9. Balaras, C.A.; Argiriou, A.A. Infrared thermography for building diagnostics. *Energy Build.* **2002**, *34*, 171–183. [[CrossRef](#)]
10. Lehmann, B.; Ghazi Wakili, K.; Frank, T.; Vera Collado, B.; Tanner, C. Effects of individual climatic parameters on the infrared thermography of buildings. *Appl. Energy* **2013**, *110*, 29–43. [[CrossRef](#)]
11. Adán, A.; Pérez, V.; Vivancos, J.-L.; Aparicio-Fernández, C.; Prieto, S.A. Proposing 3D Thermal Technology for Heritage Building Energy Monitoring. *Remote Sens.* **2021**, *13*, 1537. [[CrossRef](#)]
12. Hoegner, L.; Stilla, U. Building facade object detection from terrestrial thermal infrared image sequences combining different views. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *II-3/W4*, 55–62. [[CrossRef](#)]
13. Brooke, C. Thermal Imaging for the Archaeological Investigation of Historic Buildings. *Remote Sens.* **2018**, *10*, 1401. [[CrossRef](#)]
14. Denio, H. Aerial solar thermography and condition monitoring of photovoltaic systems. In Proceedings of the 2012 38th IEEE Photovoltaic Specialists Conference, Austin, TX, USA, 3–8 June 2012; pp. 613–618.
15. Vadivambal, R.; Jayas, D.S. Applications of Thermal Imaging in Agriculture and Food Industry—A Review. *Food Bioprocess Technol.* **2011**, *4*, 186–199. [[CrossRef](#)]
16. Li, J.; Gong, W.; Li, W.; Liu, W. Robust pedestrian detection in thermal infrared imagery using the wavelet transform. *Infrared Phys. Technol.* **2010**, *53*, 267–273. [[CrossRef](#)]
17. Speakman, J.R.; Ward, S. Infrared thermography: Principles and applications. *Zoology-Jena* **1998**, *101*, 224–232.
18. ElMasry, G.; ElGamal, R.; Mandour, N.; Gou, P.; Al-Rejaie, S.; Belin, E.; Rousseau, D. Emerging thermal imaging techniques for seed quality evaluation: Principles and applications. *Food Res. Int.* **2020**, *131*, 109025. [[CrossRef](#)]
19. Costa, J.M.; Grant, O.M.; Chaves, M.M. Thermography to explore plant-environment interactions. *J. Exp. Bot.* **2013**, *64*, 3937–3949. [[CrossRef](#)]
20. Domazetović, F.; Šiljeg, A.; Marić, I.; Faričić, J.; Vassilakis, E.; Panđa, L. Automated Coastline Extraction Using the Very High Resolution WorldView (WV) Satellite Imagery and Developed Coastline Extraction Tool (CET). *Appl. Sci.* **2021**, *11*, 9482. [[CrossRef](#)]
21. Lahiri, B.B.; Bagavathiappan, S.; Jayakumar, T.; Philip, J. Medical applications of infrared thermography: A review. *Infrared Phys. Technol.* **2012**, *55*, 221–235. [[CrossRef](#)]
22. Lewis, A.; Hilley, G.E.; Lewicki, J.L. Integrated thermal infrared imaging and structure-from-motion photogrammetry to map apparent temperature and radiant hydrothermal heat flux at Mammoth Mountain, CA, USA. *J. Volcanol. Geotherm. Res.* **2015**, *303*, 16–24. [[CrossRef](#)]
23. Poloprutský, Z. Parametric modelling for HBIM: Design of window library for rural building. *Civ. Eng. J.* **2019**, *4*, 620–630. [[CrossRef](#)]
24. Dlesk, A.; Vach, K. Point Cloud Generation of a Building from Close Range Thermal Images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-5/W2*, 29–33. [[CrossRef](#)]
25. Scaioni, M.; Rosina, E.; L’Erario, A.; Díaz-Vilariño, L. Integration of infrared thermography and photogrammetric surveying of built landscape. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-5/W1*, 153–160.
26. Maset, E.; Fusiello, A.; Crosilla, F.; Toldo, R.; Zorzetto, D. Photogrammetric 3D building reconstruction from thermal images. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *4*, 25–32. [[CrossRef](#)]
27. Dlesk, A.; Vach, K.; Holubec, P. Usage of photogrammetric processing of thermal images for civil engineers. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-5*, 99–103. [[CrossRef](#)]
28. Dlesk, A.; Vach, K.; Pavelka, K. Transformations in the Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images. *Sensors* **2021**, *21*, 5061. [[CrossRef](#)] [[PubMed](#)]
29. Weber, I.; Jenal, A.; Kneer, C.; Bongartz, J. PANTIR—a dual camera setup for precise georeferencing and mosaicing of thermal aerial images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *XL-3/W2*, 269–272. [[CrossRef](#)]
30. Kadlec, K. Teoretické základy bezdotykového měření teploty (část 1). *Automa* **2014**, *2*, 13–15.
31. Kadlec, K. Teoretické základy bezdotykového měření teploty (část 2). *Automa* **2014**, *4*, 24–26.
32. Fokaides, P.A.; Kalogirou, S.A. Application of infrared thermography for the determination of the overall heat transfer coefficient (U-Value) in building envelopes. *Appl. Energy* **2011**, *88*, 4358–4365. [[CrossRef](#)]
33. Optris—Basic Principles of Non-Contact Temperature Measurement. Available online: [https://www.optris.cz/tl\\_files/pdf/Downloads/Zubehoer/IR%20Basics.pdf](https://www.optris.cz/tl_files/pdf/Downloads/Zubehoer/IR%20Basics.pdf) (accessed on 4 November 2021).

34. Edelman, G.J.; Hoveling, R.J.M.; Roos, M.; Leeuwen, T.G.; Aalders, M.C.G. Infrared Imaging of the Crime Scene: Possibilities and Pitfalls. *J. Forensic Sci.* **2013**, *58*, 1156–1162. [[CrossRef](#)] [[PubMed](#)]
35. Elßner, M. Vacuum qualitz evaluation for uncooled micro bolometer thermal imager sensors. *Microelectron. Reliab.* **2014**, *54*, 1758–1763. [[CrossRef](#)]
36. Luhmann, T.; Piechel, J.; Roelfs, T. Geometric calibration of thermographic cameras. *Therm. Infrared Remote Sens. Remote Sens. Digit. Image Processing* **2013**, *17*, 27–42.
37. User's Manual FLIR Exx Series. Available online: <https://www.flir.com/globalassets/imported-assets/document/flir-exx-series-user-manual.pdf> (accessed on 26 April 2021).
38. Agisoft Metashape User Manual. Available online: [https://www.agisoft.com/pdf/metashape-pro\\_1\\_5\\_en.pdf](https://www.agisoft.com/pdf/metashape-pro_1_5_en.pdf) (accessed on 4 November 2021).
39. Cloud Compare [GPL Software]. Available online: <http://www.cloudcompare.org/> (accessed on 3 November 2020).
40. QGIS Development Team, QGIS Geographic Information System. Open Source Geospatial Foundation. Available online: <https://qgis.org/en/site/> (accessed on 4 November 2021).

# Documentation of the code

# **Documentation of the code: Sharpening and Reprojection method**

author: Ing. Adam Dlesk

Department of Geomatics, Faculty of the Civil Engineering, Czech Technical  
University in Prague

thesis: Photogrammetric processing of non-traditional data



## **Acknowledgement**

This code is a part of the thesis “Photogrammetric processing of non-traditional data”. The code was solely written by Adam Dlesk who is the author of the thesis. The code is released under BSD-3-Clause. The code requires some of the external libraries released under different licenses. There were used Python standard package libraries, GDAL library, OpenCV library, NumPy and Scipy library.



# Table of Contents

1. Classes .....	4
1.1. classes.py .....	4
2. Functions .....	8
2.1. matrices.py .....	8
createRotationMatrix .....	8
createRotationMatrix2 .....	8
createKameraMatrix .....	8
omega_phi_kappaFromR .....	9
omega_phi_kappaFromR .....	9
getRbetweenCameras .....	9
quaternionFromR .....	9
2.2. transformation.py .....	11
world2camera .....	11
camera2world .....	11
camera2imagePix .....	11
camera2imageM.....	12
image2camera .....	12
imagePix2imageM.....	12
azimuth_tilt_swing2omega_phi_kappa .....	13
azimuth_tilt_swing2omega_phi_kappa .....	13
rgbCameraSys2Termo .....	13
csv2GeoTIFF.....	14
csv2RawGeoTIFF.....	14
jpg2GeoTIFF.....	14
csv2jpg.....	14
orthophotoRecalculation8bit .....	15
orthophotoRecalculation16bit .....	15
2.3. photogrammetry.py .....	16
getDistortionX .....	16
getDistortionY.....	16
createConverter .....	16
createDepthMapConverter .....	16
convertImgUsingConverter .....	17
convertArrayUsingConverter.....	17

convertGeoTIFFUsingConverter .....	17
fusionGeoTIFF.....	18
fusionRGT .....	18
fusionRTB.....	18
fusionTGB .....	18
visibilityTestDepth.....	19
visibilityTestNormalVector .....	19
rayRecalculationBmp.....	19
rayRecalculationGeoTiff .....	20
reprojection.....	20

# 1. Classes

## 1.1. classes.py

---

class **Camera**

The class represents position and rotation of the common camera.

### Methods

`__init__`(pos\_x, pos\_y, pos\_z, omega, phi, kappa, name)

Constructor method.

#### Parameters

- `pos_x` (float) – Position of the optical center of the camera on the X axis.
- `pos_y` (float) – Position of the optical center of the camera on the Y axis.
- `pos_z` (float) – Position of the optical center of the camera on the Z axis.
- `omega` (float) – Omega parameter of the rotation matrix of the camera.
- `phi` (float) – Phi parameter of the rotation matrix of the camera.
- `kappa` (float) – Kappa parameter of the rotation matrix of the camera.
- `name` (string) – Label of the camera.

`show()`

Prints the parameters of the position and rotation of the camera in the console.

class Camera.**Calibration**

The class represents a set of values which define geometric parameters of the camera, its camera's lens and camera's sensor. The parameters are considered as photogrammetric camera calibration.

### Methods

`__init__`(c, px, py, K1, K2, pixel\_size, im\_dim\_x, im\_dim\_y, m, s)

Constructor method.

#### Parameters

- `c` (float) – Constant of the camera in pixels.

- **px** (float) – Position of the principal point of the camera on the X axis in pixels.
- **py** (float) – Position of the principal point of the camera on the Y axis in pixels.
- **k1** (float) – The first coefficient of the radial distortion of Brown's distortion model
- **k2** (float) – The second coefficient of the radial distortion of Brown's distortion model
- **pixel\_size** (float) – Physical pixel pitch between in millimeters.
- **m** (float) – Scale difference in X and Y axis.
- **s** (float) – Skew coefficient.

**show()**

Prints the parameters of the calibration in the console.

class **ThermalCamera**

The class represents position and rotation of the camera in the camera coordinate system of the second camera in the fixed system.

#### Methods

**\_\_init\_\_**(dx, dy, dz, d\_omega, d\_phi, d\_kappa)

Constructor method.

#### Parameters

- **dx** (float) – Position of the optical center of the camera on the X axis in the camera coordinate system of the second camera in the fixed system.
- **dy** (float) – Position of the optical center of the camera on the Y axis in the camera coordinate system of the second camera in the fixed system.
- **dz** (float) – Position of the optical center of the camera on the Z axis in the camera coordinate system of the second camera in the fixed system.
- **d\_omega** (float) – Omega parameter of the rotation matrix of the camera. Related to the relative rotation of the camera coordinate systems on in the fixed system.
- **d\_phi** (float) – Phi parameter of the rotation matrix of the camera. Related to the relative rotation of the camera coordinate systems on in the fixed system.
- **d\_kappa** (float) – Kappa parameter of the rotation matrix of the camera. Related to the relative rotation of the camera coordinate systems on in the fixed system.

**show()**

Prints the parameters of the relative position and relative rotation of the camera in the console.

class Camera.**Calibration**

The class represents a set of values which define geometric parameters of the camera, its camera's lens and camera's sensor. The parameters are considered as photogrammetric camera calibration.

### Methods

**`__init__`**(*c, px, py, K1, K2, pixel\_size, im\_dim\_x, im\_dim\_y, m, s*)

Constructor method.

#### Parameters

- **`c`** (float) – Constant of the camera in pixels.
- **`px`** (float) – Position of the principal point of the camera on the X axis in pixels.
- **`py`** (float) – Position of the principal point of the camera on the Y axis in pixels.
- **`K1`** (float) – The first coefficient of the radial distortion of Brown's distortion model
- **`K2`** (float) – The second coefficient of the radial distortion of Brown's distortion model
- **`pixel_size`** (float) – Physical pixel pitch between in millimeters.
- **`m`** (float) – Scale difference in X and Y axis.
- **`s`** (float) – Skew coefficient.

**`show()`**

Prints the parameters of the calibration in the console.

class **Point2D**

The class represents a position of a point in the common two-dimensional coordinate system.

### Methods

**`__init__`**(*point\_x, point\_y*)

Constructor method.

#### Parameters

- **`point_x`** (float) – Position of a point in the common two-dimensional coordinate system in X axis.
- **`point_y`** (float) – Position of a point in the common two-dimensional coordinate system in Y axis.

**`show()`**

Prints the position of the point in image.

---

class **Point3D**

The class represents a position of a point in the common three dimensional coordinate system.

#### Methods

**`__init__`**(point\_x, point\_y, point\_z , R, G, B, nx, ny, nz)

Constructor method.

#### Parameters

- **point\_x** (float) – Position of the point on the X axis.
- **point\_y** (float) – Position of the point on the Y axis.
- **point\_z** (float) – Position of the point on the Z axis.
- **R** (int) – Red color value in the RGB color model.
- **G** (int) – Green color value in the RGB color model.
- **B**(int)– Blue color value in the RGB color model.
- **nx** (float) – X element of normal vector of the point.
- **ny** (float) – Y element of normal vector of the point.
- **nz** (float) – Z element of normal vector of the point.

**`show()`**

Prints the parameters of the point.

# 2. Functions

## 2.1. matrices.py

Dependencies: NumPy, math

### Functions:

---

**createRotationMatrix**(*omega*, *phi*, *kappa*)

Creates rotation matrix from the input parameters (angle scheme – omega, phi, kappa).

#### Parameters

- **omega** (float) – Rotation around X axis.
- **phi** (float) – Rotation around Y axis.
- **kappa** (float) – Rotation around Z axis.

#### Returns

Rotation matrix.

---

**createRotationMatrix2**(*yaw*, *pitch*, *roll*)

Creates rotation matrix from the input parameters (angle scheme – yaw, pitch, roll).

#### Parameters

- **yaw** (float) – Yaw in the navigation coordinate system.
- **pitch** (float) – Pitch in the navigation coordinate system.
- **roll** (float) – Roll in the navigation coordinate system.

#### Returns

Rotation matrix.

---

**createKameraMatrix**(*c*, *px*, *py*, *pixel\_size*)

Creates camera matrix *K*.

#### Parameters

- **c** (float) – Constant of the camera
- **px** (float) – Position of the principal point of the camera on the X axis.



- `py` (float) – Position of the principal point of the camera on the Y axis.
- `pixel_size` (float) – Physical pixel pitch between in millimeters.

**Returns**

Camera matrix.

---

`omega_phi_kappaFromR` (R)

Derives rotation angles from the rotation matrix.

**Parameters**

- `R` (array) – Rotation matrix.

**Returns**

Angles omega, phi, kappa.

---

`omega_phi_kappaFromR` (R)

Another solution for derivation of rotation angles from the rotation matrix.

**Parameters**

- `R` (array) – Rotation matrix.

**Returns**

Angles omega, phi, kappa

---

`getRbetweenCameras` (R1, R2)

Creates rotation matrix between the two cameras where are given their rotation matrix in the reference coordinate system.

**Parameters**

- `R1` (array) – Rotation matrix of camera\_1 .
- `R2` (array) – Rotation matrix of camera\_2.

**Returns**

Rotation matrix between the camera\_1 and camera\_2.

---

`quaternionFromR` (R)

Derives quaternion components from the rotation matrix.

**Parameters**

- **R** (array) – Rotation matrix.

**Returns**

Quaternion components.

---

## 2.2. transformation.py

Dependencies: NumPy, OpenCV, GDAL, math, csv, matrices.py, photogrammetry.py

### Functions

---

**world2camera**(*point*, *camera*)

Transforms point from the world coordinate system to the camera coordinate system.

#### Parameters

- **point** ([Point3D](#)) – Point which shall be transformed.
- **camera** ([Camera](#)) – Camera coordinate system defined by the camera where the point shall be transformed.

#### Returns

[Point3D](#) with transformed coordinates.

---

**camera2world**(*point*, *camera*)

Transforms point from the camera coordinate system to the world coordinate system.

#### Parameters

- **point** ([Point3D](#)) – Point which shall be transformed.
- **camera** ([Camera](#)) – Camera coordinate system defined by the camera where the point shall be transformed.

#### Returns

[Point3D](#) with transformed coordinates.

---

**camera2imagePix**(*point*, *camera*)

Transforms point from the camera coordinate system to the image coordinate system.

#### Parameters

- **point** ([Point3D](#)) – Point which shall be transformed.
- **camera** ([Camera](#)) – Image coordinate system defined by the camera where the point shall be transformed.

#### Returns

[Point2D](#) with transformed coordinates (in pixels).

---

**camera2imageM**(*point*, *camera*)

Transforms point from the camera coordinate system to the image coordinate system.

#### Parameters

- **point** ([Point3D](#)) – Point which shall be transformed.
- **camera** ([Camera](#)) – Image coordinate system defined by the camera where the point shall be transformed.

#### Returns

[Point2D](#) with transformed coordinates (in millimeters).

---

**image2camera**(*im\_x*, *im\_y*, *im\_z*, *K*, *pix\_in\_mm*, *dim\_x*, *dim\_y*)

Transforms point from image coordinate system to the camera coordinate system.

#### Parameters

- **im\_x** (float) – Position of point in image coordinate system on the X axis.
- **im\_y** (float) – Position of point in image coordinate system on the Y axis.
- **im\_z** (float) – Position of point in image coordinate system on the Z axis.
- **K** (array) – Calibration matrix of the camera.
- **pix\_in\_mm** (float) – Physical dimension of the pixel pitch in millimeters.
- **dim\_x** (float) – Resolution of the image in X axis.
- **dim\_y** (float) – Resolution of the image in Y axis.

#### Returns

Position of the point in the camera coordinate system.

---

**imagePix2imageM**(*point2D*, *camera*)

Transforms point in image coordinates in pixels to image coordinates in millimeters and adds the Z element.

#### Parameters

- **point** ([Point2D](#)) – Point in image coordinates in pixels.
- **camera** ([Camera](#)) – Camera.

#### Returns

Position of the point in the image coordinate system with the Z element.

---

`azimuth_tilt_swing2omega_phi_kappa` (*azimuth*, *tilt*, *swing*)

Transforms the angles between different angle schemes.

**Parameters**

- `azimuth` (float) – Azimuth value.
- `tilt` (float) – Tilt value.
- `kappa` (float) – Kappa value.

**Returns**

Angles in omega, phi, kappa scheme.

---

`azimuth_tilt_swing2omega_phi_kappa` (*azimuth*, *tilt*, *swing*)

Transforms the angles between different angle schemes.

**Parameters**

- `azimuth` (float) – Azimuth value.
- `tilt` (float) – Tilt value.
- `kappa` (float) – Kappa value.

**Returns**

Angles in omega, phi, kappa scheme.

---

`rgbCameraSys2Termo` (*point*, *tir\_camera*)

Transforms point from the one camera coordinate system to the image coordinate system of another camera. The function was developed to transform the point between the cameras as fixed system - from RGB camera coordinate system to TIR image.

**Parameters**

- `point` (`Point3D`) – Position of the point in camera coordinate system.
- `tir_camera` (`ThermalCamera`) – Second camera, where the point shall be transformed.

**Returns**

Position of the point in the image (in TIR image).

---

**csv2GeoTIFF**(*path, min, max, nrows, ncols*)

Convert TIR images captured by FLIRE95 TIR camera stored as .csv to GeoTIFF. The values are recalculated to 16-bit values.

#### Parameters

- **path** (string) – Path to csv file.
- **min** (float) – Minimum value on the image for histogram stretching function.
- **max** (float) – Maximum value in the image for histogram stretching function.
- **nrows** (int) – Number of rows of the image.
- **ncols** (int) – Number of columns of the image.

#### Returns

Image in GeoTIFF with 16-bit values.

---

**csv2RawGeoTIFF**(*path, min, max, nrows, ncols*)

Convert TIR images captured by FLIRE95 TIR camera stored as .csv to GeoTIFF. The values are not recalculated to 16-bit values.

#### Parameters

- **path** (string) – Path to .csv file.
- **nrows** (int) – Number of rows of the image.
- **ncols** (int) – Number of columns of the image.

#### Returns

Image in GeoTIFF 16-bit values.

---

**jpg2GeoTIFF**(*path*)

Convert image in .jpg format to GeoTIFF.

#### Parameters

- **path** (string) – Path to .jpg file.

#### Returns

Image in GeoTIFF.

---

**csv2jpg**(*path, min, max, nrows, ncols*)

Convert TIR images captured by FLIRE95 TIR camera stored as .csv to .jpg. The values are recalculated to 8-bit values.

#### Parameters

- **path** (string) – Path to .csv file.
- **min** (float) – Minimum value on the image for histogram stretching function.
- **max** (float) – Maximum value in the image for histogram stretching function.
- **nrows** (int) – Number of rows of the image.
- **ncols** (int) – Number of columns of the image.

#### Returns

Image in .jpg format with 8-bit values.

---

**orthophotoRecalculation8bit**(*path, min, max, n\_band*)

Recalculates the orthophoto from 8-bit values to raw values according to the minimum and maximum of the histogram stretching function.

#### Parameters

- **path** (string) – Path to the orthophoto.
- **min** (float) – Minimum value on the image for histogram stretching function.
- **max** (float) – Maximum value in the image for histogram stretching function.
- **n\_band** (int) – Number of the band with thermal values.

#### Returns

Recalculated orthophoto in GeoTIFF.

---

**orthophotoRecalculation16bit**(*path, min, max, n\_band*)

Recalculates the orthophoto from 16-bit values to raw values according to the minimum and maximum of the histogram stretching function.

#### Parameters

- **path** (string) – Path to the orthophoto.
- **min** (float) – Minimum value on the image for histogram stretching function.
- **max** (float) – Maximum value in the image for histogram stretching function.
- **n\_band** (int) – Number of the band with thermal values.

#### Returns

Recalculated orthophoto in GeoTIFF.

---



## 2.3. photogrammetry.py

Dependencies: NumPy, OpenCV, GDAL, math, transformations.py, matrices.py

### Functions

---

**getDistortionX**(*point\_2D, camera*)

Calculates distortion value on X axis on image.

#### Parameters

- **point\_2D** (*Point2D*) – Position on the image in pixels.
- **camera** (*Camera*) – Camera.

#### Returns

Distortion value.

---

**getDistortionY**(*point\_2D, camera*)

Calculates distortion value on Y axis on image.

#### Parameters

- **point\_2D** (*Point2D*) – Position on the image in pixels.
- **camera** (*Camera*) – Camera.

#### Returns

Distortion value.

---

**createConverter**(*rgb, thermal\_camera, distance*)

Creates converter for the TIR image only according to the set distance.

#### Parameters

- **rgb** (*Camera*) – RGB camera.
- **thermal\_camera** (*Camera*) – TIR Camera.
- **distance** (*float*) – Distance between cameras and the scene.

#### Returns

The converter (array).

---

**createDepthMapConverter**(*rgb, thermal\_camera, depth\_map*)

Creates converter for the TIR image based on the depth map of the RGB image.

#### Parameters

- `rgb` (`Camera`) – RGB camera.
- `thermal_camera` (`Camera`) – TIR Camera.
- `depth_map` (`string`) – Path to the depth map.

**Returns**

The converter (array).

---

`convertImgUsingConverter` (*img\_path*, *converter*)

Uses converter to recalculate the TIR image in .bmp format.

**Parameters**

- `img_path` (`string`) – Path to TIR image.
- `converter` (`array`) – Converter which is used for recalculation of the TIR image.

**Returns**

The recalculated TIR image in .bmp format.

---

`convertArrayUsingConverter` (*a\_array*, *converter*)

Uses converter to recalculate the the array.

**Parameters**

- `a_array` (`array`) – Array which shall be recalculated.
- `converter` (`array`) – Converter which is used for recalculation of the array.

**Returns**

The recalculated array.

---

`convertGeoTIFFUsingConverter` (*path*, *converter*)

Uses converter to recalculate the TIR image in GeoTIFF format.

**Parameters**

- `path` (`string`) – Path to TIR image.
- `converter` (`array`) – Converter which is used for recalculation of the TIR image.

**Returns**

The recalculated TIR image in GeoTIFF.

---

**fusionGeoTIFF** (*rgb\_path, thermal\_path*)

Fuses the RGB image and TIR image (both must be in GeoTIFF) to RGBT image.

**Parameters**

- **rgb\_path** (string) – Path to RGB image.
- **thermal\_path** (string) – Path to TIR image.

**Returns**

The fused RGBT image.

---

**fusionRGT** (*rgb\_path, thermal\_path*)

Fuses the RGB image and TIR image, both in .bmp.

**Parameters**

- **rgb\_path** (string) – Path to RGB image.
- **thermal\_path** (string) – Path to TIR image.

**Returns**

The fused RGT image.

---

**fusionRTB** (*rgb\_path, thermal\_path*)

Fuses the RGB image and TIR image, both in .bmp.

**Parameters**

- **rgb\_path** (string) – Path to RGB image.
- **thermal\_path** (string) – Path to TIR image.

**Returns**

The fused RTB image.

---

**fusionTGB** (*rgb\_path, thermal\_path*)

Fuses the RGB image and TIR image, both in .bmp.

**Parameters**

- **rgb\_path** (string) – Path to RGB image.
- **thermal\_path** (string) – Path to TIR image.

**Returns**

The fused TGB image.

---

**visibilityTestDepth**(*depth\_map, point, camera, tol*)

According to the depth map and set tolerance the function determines if the point is potentially visible.

#### Parameters

- **depth\_path** (string) – Path to the depthmap image.
- **point** (Point3D) – Coordinates of the point in the world coordinate system.
- **camera** (Camera) – Camera.
- **tol** (float) – The tolerance.

#### Returns

Returns True/False value.

---

**visibilityTestNormalVector**(*normal\_map, point, camera, tol*)

According to the normal map and set tolerance the function determines if the point is potentially visible.

#### Parameters

- **normal\_path** (string) – Path to the normal map image.
- **point** (Point3D) – Coordinates of the point in the world coordinate system.
- **camera** (Camera) – Camera.
- **tol** (float) – The tolerance.

#### Returns

Returns True/False value.

---

**rayRecalculationBmp**(*path, rgb\_camera, tir\_camera*)

Ray recalculation function which is a core of the sharpening method. This concrete function works only with images which are in .bmp format.

#### Parameters

- **path** (string) – Path to the TIR images in .bmp format. In the same directory depth maps in .tif format are required.
- **rgb\_camera** (Camera) – RGB camera of the fixed system. The calibration of the camera is required.
- **tir\_camera** (Camera) – TIR camera of the fixed system. The calibration of the camera is required. The relative rotation and translation is required.

#### Returns

Recalculated images in .bmp format.

---

**rayRecalculationGeoTiff** (*path*, *rgb\_camera*, *tir\_camera*)

Ray recalculation function which is a core of the sharpening method. This concrete function works only with images which are in .tiff format.

#### Parameters

- **path** (*string*) – Path to the TIR images in .tiff format. In the same directory depth maps in .tif format are required.
- **rgb\_camera** (*Camera*) – RGB camera of the fixed system. The calibration of the camera is required.
- **tir\_camera** (*Camera*) – TIR camera of the fixed system. The calibration of the camera is required. The relative rotation and translation is required.

#### Returns

Recalculated images in .tiff format.

---

**reprojection** (*path\_pc*, *path\_eo*, *path\_images*, *rgb\_camera*, *tir\_camera*, *depth\_tol*, *normal\_tol*)

The reprojection method. For further information, please see the thesis.

#### Parameters

- **path\_pc** (*string*) – Path to the point cloud which is the main input for the reprojection method. As the delimiter, a single space is expected. The information shall be in this order: X, Y, Z, R, G, B, n\_x, n\_y, n\_z (elements of the normal vector).
- **path\_eo** (*string*) – Path to the .txt file where is stored information about external orientation of the RGB images. As the delimiter, “;” is expected. Lines which starts with “#” are ignored (e.g. header). The information shall be in this order: Name, X, Y, Z, omega, phi, kappa.
- **path\_images** (*string*) – Path to the .txt file where are stored depth map and normal map images. The images must have the same name as the RGB image with a suffix *\_depth\_map*, or *\_normals* respectively.
- **rgb\_camera** (*Camera*) – RGB camera of the fixed system. The calibration of the camera is required.
- **tir\_camera** (*Camera*) – TIR camera of the fixed system. The calibration of the camera is required. The relative rotation and translation is required.
- **depth\_tol** (*float*) – The tolerance for the depth test of test of visibility.
- **normal\_tol** (*float*) – The tolerance for the normal test of test of visibility.

## Returns

The point cloud augmented with the TIR information. The point cloud is stored with suffix `_reprojection` in the same path as original point cloud. In the same path is stored also file with the statistic information for each point. For each point is presented number of valid observations on images, standard deviation, range, and result of the normality test by Shapiro Wilk test.

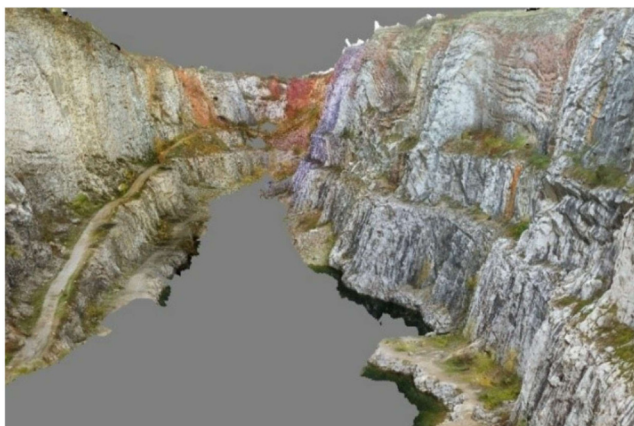
# Image Attachments



# FRAGMENTS FROM THE PROCESSING OF ANALOGUE ROLLEI 6006 METRIC IMAGES (THE GREAT AMERICA)



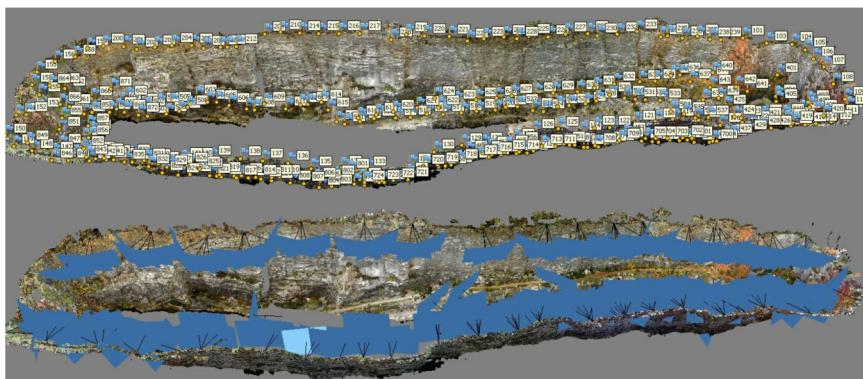
a)



b)



c)



d)

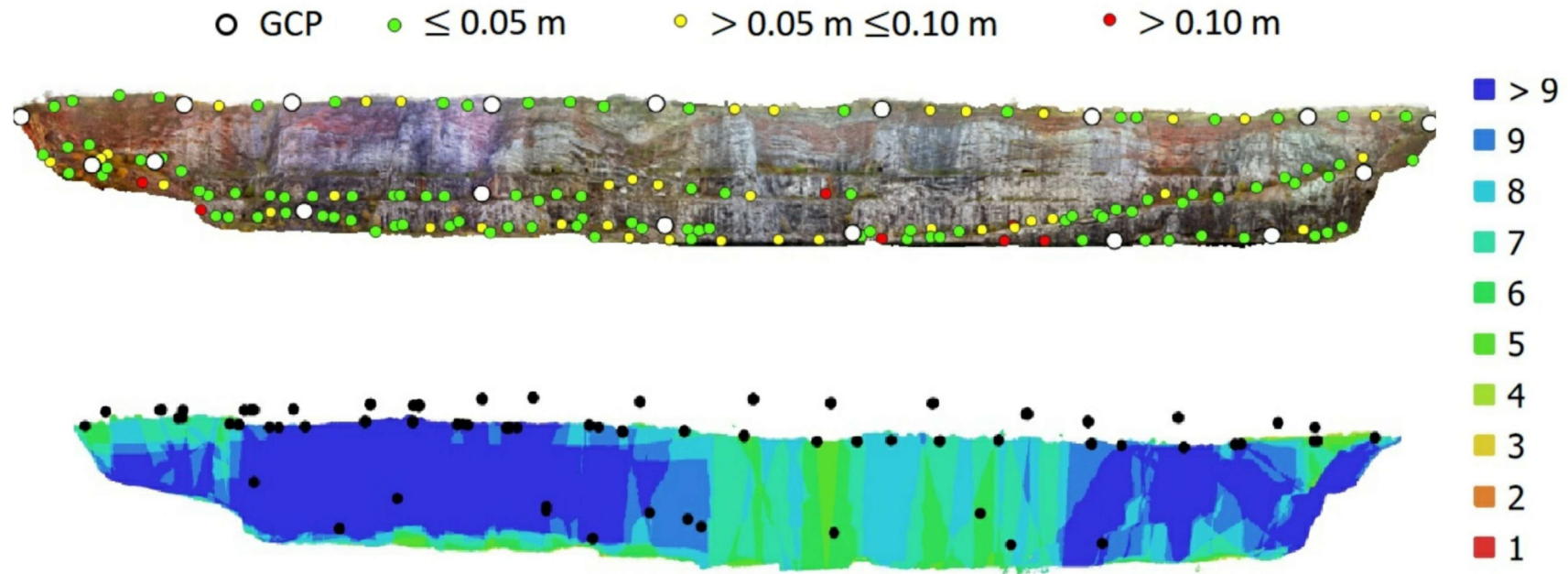
## DESCRIPTION

Fragments from the processing of the analogue metric images of the limestone quarry the Great America. The images were captured by Rollei 6006 metric camera. The analogue images were pre-processed and then processed in Agisoft Photoscan. In total, 170 images were processed.

- a) Digitizing of analogue metric images.
- b) Created textured 3D model.
- c) A print of Rollei 6006 metric image.
- d) The point cloud with the control and check points (on top). The point cloud with camera positions (below).

Created by Adam Dlesk  
9. 8. 2018

# ORTHOPHOTO OF THE NORTH WALL OF THE LIMESTONE QUARRY THE GREAT AMERICA



## DESCRIPTION

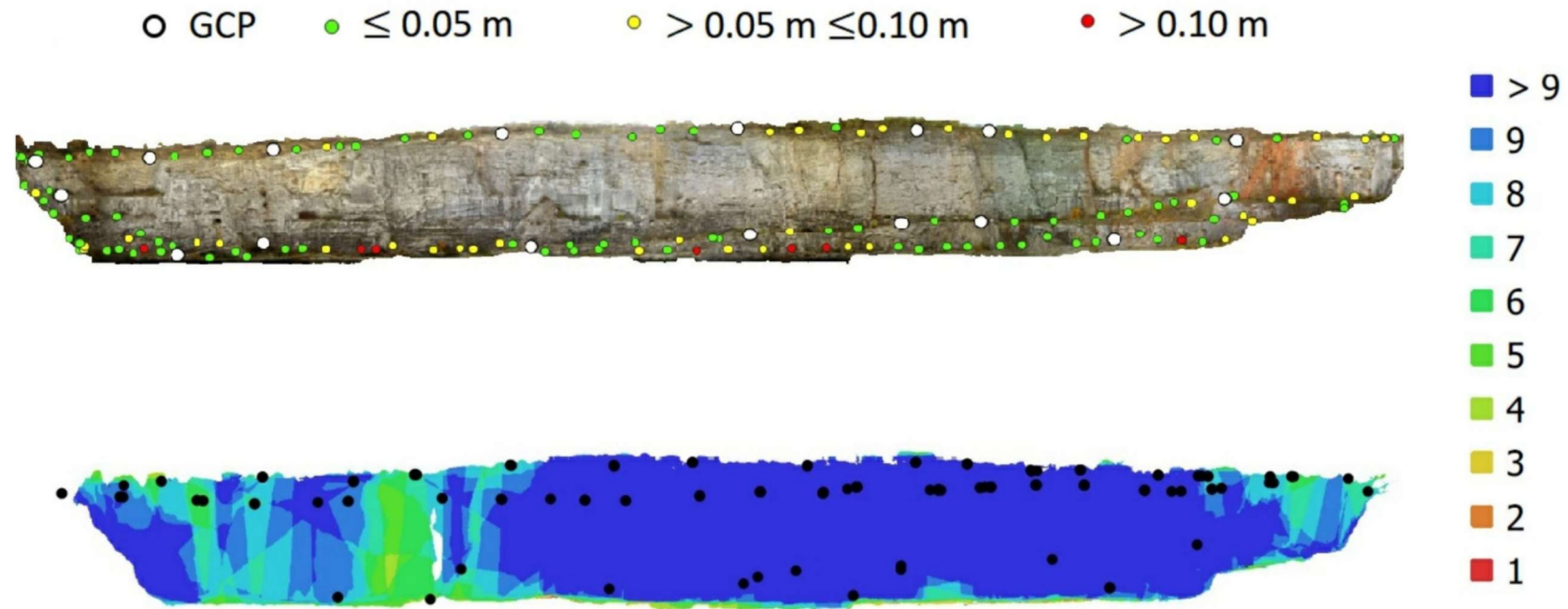
The orthophoto was created by photogrammetric processing of 170 analogue metric images using modern methods. The images were captured by Rollei 6006 metric camera.

For the analysis, there were used 35 control points and 293 check points. In the upper image, it is possible to see the distribution of the check points and by the colour of the point, it is possible to approximately determine the measured RMSE.

In the lower image, it is possible to see the image overlap of the north wall. It is possible to observe that the RMSE of check points correlate with the number of overlapping images.

Created by Adam Dlesk  
1. 6. 2018

# ORTHOPHOTO OF THE SOUTH WALL OF THE LIMESTONE QUARRY THE GREAT AMERICA



## DESCRIPTION

The orthophoto was created by photogrammetric processing of 170 analogue metric images using modern methods. The images were captured by Rollei 6006 metric camera.

For the analysis, there were used 35 control points and 293 check points. In the upper image, it is possible to see the distribution of the check points and by the colour of the point, it is possible to approximately determine the measured RMSE.

In the lower image, it is possible to see the image overlap of the south wall. It is possible to observe that the RMSE of check points correlate with the number of overlapping images.

Created by Adam Dlesk  
1. 6. 2018





# FRAGMENTS FROM THE PROCESSING OF ANALOGUE ROLLEI 3003 METRIC IMAGES (ZIKMUND'S CHAPEL, SAINT VITUS CATHEDRAL IN PRAGUE)



a)



b)



c)

## DESCRIPTION

Fragments from the processing of the analogue metric images of the Zikmund's chapel, Saint Vitus Cathedral. The images were captured by Rollei 3003 metric camera. This project was used for accuracy analysis. For the analysis, 100 hundred check points were used.

- a) Set of Rollei 3003 metric images.
- b) A computed point cloud with the position of the images.
- c) Distribution of the check points.

# FRAGMENTS FROM THE PROCESSING OF ANALOGUE UMK METRIC IMAGES (PADISE ABBEY, ESTONIA)



a)



b)

## DESCRIPTION

Fragments from the processing of the analogue metric images of the Padise Abbey in Estonia. The images were captured by UMK 6.5/1318 metric camera. The project is a practical case. By new processing of archival data is possible to get valuable results.

- a) Padise abbey in 1991.
- b) Padise abbey after restoration in 2017.
- c) Control point on the metric image captured in 1991.
- d) Control point is still detectable even on the new image captured in 2017.
- e) Overlap of the images captured in 1991 (generated in Agisoft Metashape).
- f) Point cloud generated based on images captured in 1991 using modern method (processed in Agisoft Metashape).



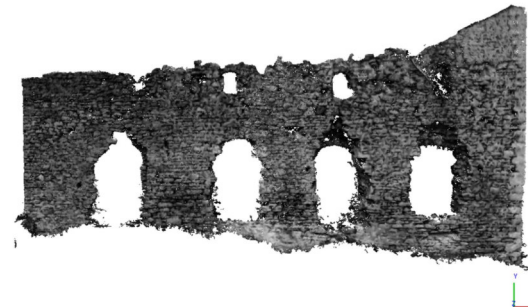
c)



d)



e)



f)

Created by Adam Dlesk  
12. 11. 2020



Laboratory of Photogrammetry  
Department of Geomatics  
FCE, CTU in Prague



## PADISE ABBEY: THE OVERLAY OF ORTHOPHOTOS



### DESCRIPTION

The overlay represents the south wall of the Estonian Padise Abbey. In past years the wall was renovated.

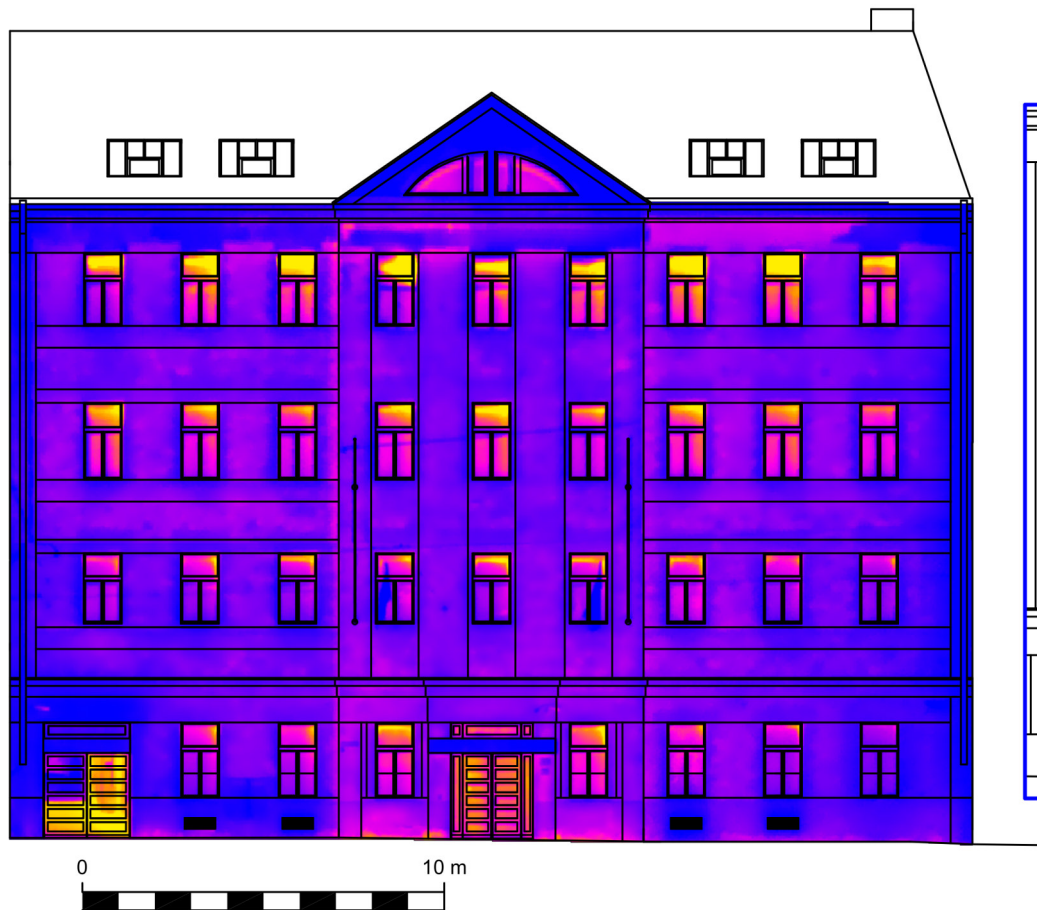
Using the suggested pre-processing and processing method, the analogue metric images captured in 1991 by UMK 6.5/1318 were processed and the orthophoto was generated.

In 2017, there was carried out a new survey with the digital camera. The images were processed and the orthophoto was generated as well.

By comparing the two orthophotos which both represent different time stage of the abbey, it is possible to detect what was the outcome of the restoration and what part of the wall is the preserved part.

Created by Adam Dlesk  
13. 11. 2020

# MOSAIC OF FACADE CREATED FROM A SET OF THERMAL INFRARED IMAGES, THERMAL PROFILE MEASUREMENT



Temperature [°C]

- < -6.0
- -4.0
- -2.0
- 0.0
- > 2.0

$\epsilon = 0.9$

## DESCRIPTION

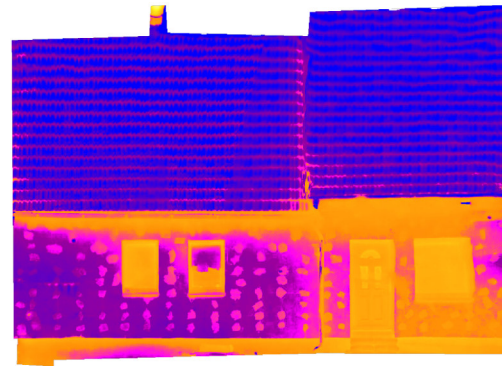
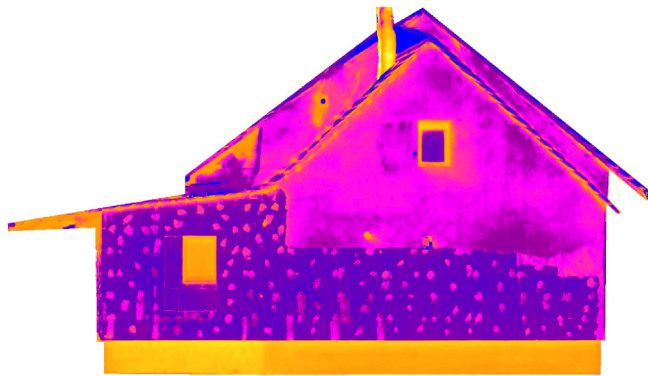
The mosaic was created by 6 terrestrial TIR images. The images were transformed into CAD drawing by projective transformation.

There is also presented a thermal profile measurement. On the profile, it is possible to analyze the thermal leaks of the windows.

Created by Adam Dlesk  
13. 6. 2019



# ORTHOPHOTOS OF FACADES CREATED FROM CLOSE RANGE THERMAL INFRARED IMAGES



Temperature [° C]



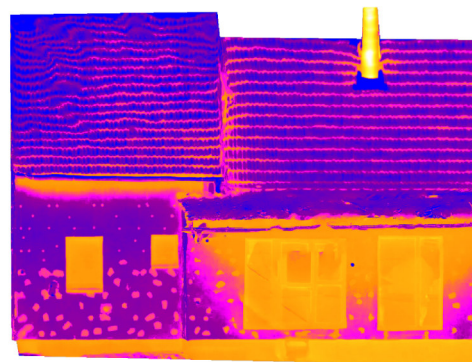
$\epsilon = 0.9$

## DESCRIPTION

The orthophotos of the facades have been created by photogrammetric processing of close-range thermal infrared images (solely). Each pixel of the orthophoto is augmented by temperature information derived from the original thermal infrared images.

This building was chosen to have a particular case where it was possible to process just thermal infrared images by SfM-MVS method.

Created by Adam Dlesk  
13. 6. 2019



## ORTHOPHOTO WITH RGT BAND COMBINATION (LEFT) AND SEPARATED T BAND (RIGHT)



### DESCRIPTION

The orthophoto was created by the sharpening method.

On the orthophoto (left), it is possible to notice the seams between the images of the orthomosaic. It is a good sign. During photogrammetric processing, the colours were not weighted, or averaged and then the R and G bands are affected by the different lighting of the RGB images.

On the right, there is a separated T band (of the same orthophoto) without the seams. It signifies that the thermal infrared measuring was carried out correctly under stable conditions.

Created by Adam Dlesk  
13. 6. 2019



Laboratory of Photogrammetry  
Department of Geomatics  
FCE, CTU in Prague

## PROCESSING OF IMAGES BY SHARPENING METHOD



a)



b)

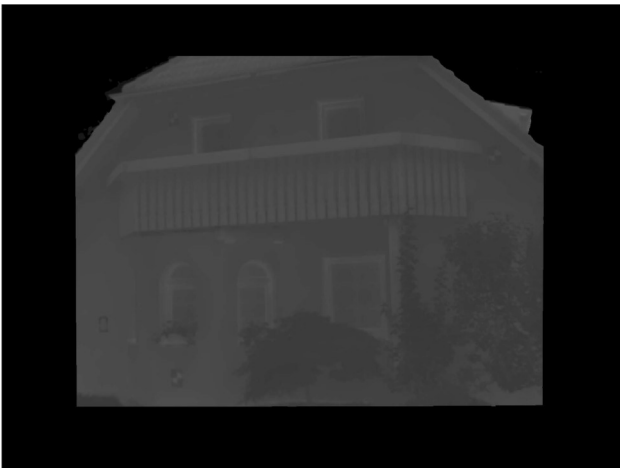


c)

### DESCRIPTION

Processing of images by the sharpening method.

- a) Original RGB image
- b) Depth map of the image generated as an outcome of SfM-MVS processing
- c) Original TIR image (captured in the fixed system together with RGB image)
- d) Remapped TIR image using ray recalculation function
- e) Fused RGT image



d)



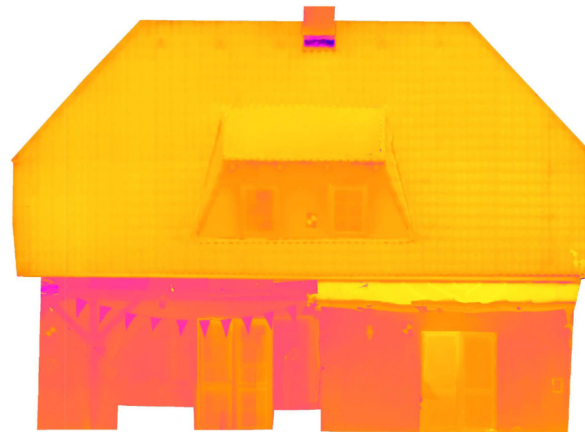
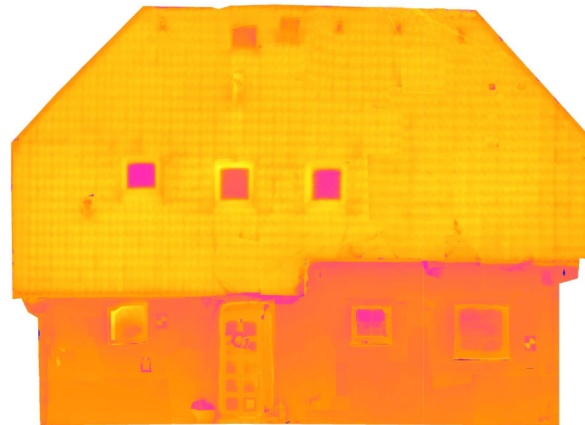
e)

Created by Adam Dlesk  
11. 3. 2022



# ORTHOPHOTOS CREATED BY SHARPENING METHOD

Temperature [°C]



## DESCRIPTION

$\epsilon = 0.9$

The orthophotos were created by the sharpening method. In total, 187 images were captured for the processing.

Some parts of the facades were covered by common objects and those parts were cut off from the orthophotos.

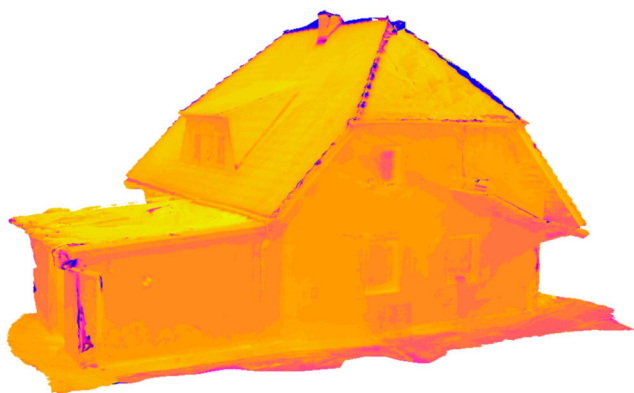
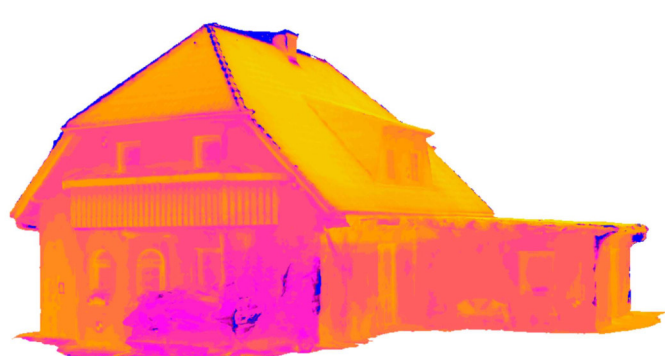
The main problem was that the facade of the building was textureless (common case). Still, the processing by the sharpening method was successful and the augmented orthophotos were generated with respectable quality and accuracy.

Created by Adam Dlesk  
22. 10. 2021



Laboratory of Photogrammetry  
Department of Geomatics  
FCE, CTU in Prague

## 3D MODEL WITH AUGMENTED TEXTURE CREATED BY SHARPENING METHOD



Temperature [°C]

< 7

10

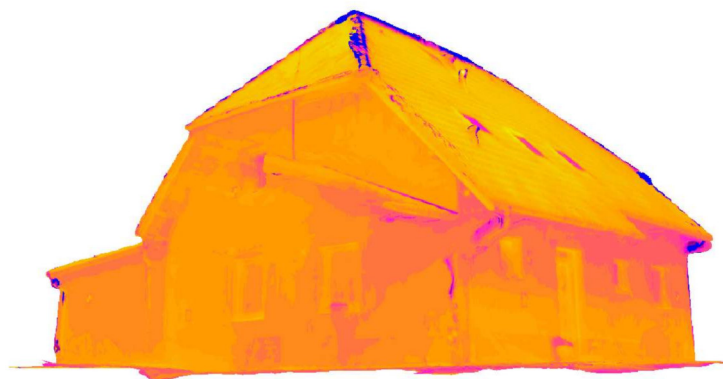
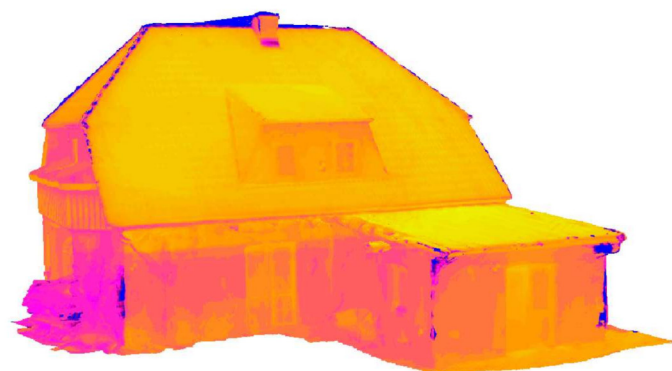
11

15

22

> 25

$\varepsilon = 0.9$



### DESCRIPTION

The augmented texture of the model was created by the sharpening method.

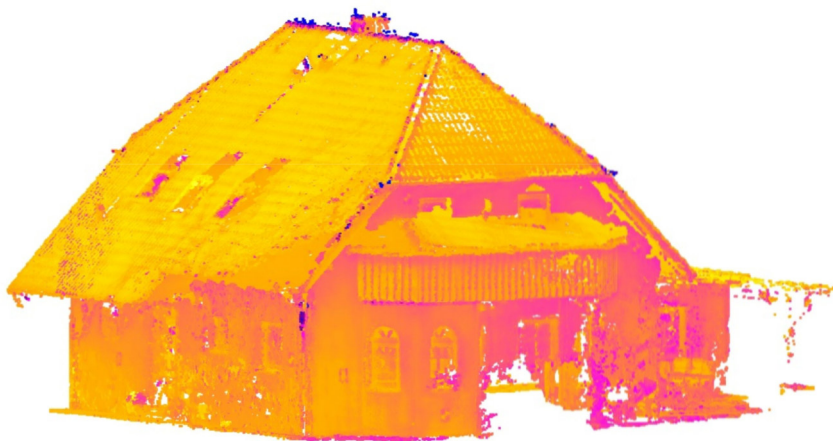
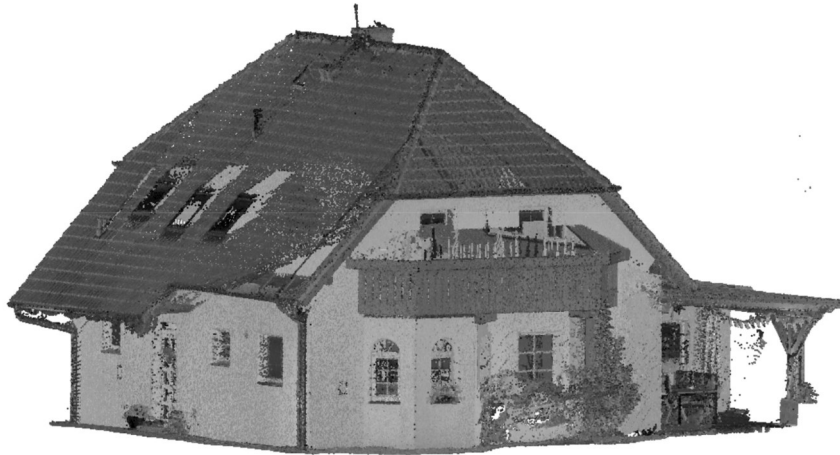
The augmented texture can be used in the following fields which like virtual reality or augmented reality.

Created by Adam Dlesk  
12. 2. 2022



Laboratory of Photogrammetry  
Department of Geomatics  
FCE, CTU in Prague

# AUGMENTED POINT CLOUD CREATED BY REPROJECTION METHOD



## DESCRIPTION

The augmented point cloud was created by reprojected method (lower image).

For the reprojection method, the point cloud created by Leica BLK laser scanner was used. The point cloud was subsampled and then had 1,294,217 points (image on the top).

The augmented point cloud has 40% points of the original number. The rest of the points did not pass the selected tolerances of the test of visibility. Some of the points were visible on multiple TIR images and temperature value of the point was calculated as a mean (lower image).

Temperature [°C]



$\epsilon = 0.9$





# AUGMENTED POINTS VISIBLE ON MULTIPLE TIR IMAGES



## DESCRIPTION

The augmented point cloud was generated by reprojection method.

The fact that some of the points are visible on multiple TIR images enables statistical analysis of the augmentation. The final temperature value is a mean of the multiple values.

On top, there are highlighted points (purple) that were visible on 4 and more TIR images. It was 61% of points of the augmented point cloud.

Below, there are highlighted points (purple) that were visible on 8 and more TIR images. It was 21% of points of the augmented point cloud.





# EXAMPLE OF STATISTICS ON POINTS WITH MULTIPLE OBSERVATION



## DESCRIPTION

The augmented point cloud was generated by reprojection method.

Example of the possible statistics on the augmented points which were visible on multiple TIR images.

On the top, there are highlighted points (purple) whose temperature range was lower than 1°C. It was 55% of the points of the augmented point cloud.

On the bottom, there are highlighted points (purple) whose temperature range was lower than 0.5°C. It was 35% of points of the augmented point cloud.

