



**HAL**  
open science

# Geospatial recording and point cloud classification of heritage buildings

Arnadi Dhestaratri Murdiyoso

► **To cite this version:**

Arnadi Dhestaratri Murdiyoso. Geospatial recording and point cloud classification of heritage buildings. Modeling and Simulation. Université de Strasbourg, 2020. English. NNT : 2020STRAD007 . tel-02953668

**HAL Id: tel-02953668**

**<https://theses.hal.science/tel-02953668>**

Submitted on 30 Sep 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*École doctorale mathématiques, sciences de l'information et de l'ingénieur*

Laboratoire ICube, Équipe TRIO UMR 7357 INSA de Strasbourg

# THÈSE

présentée par

**Arnadi Dhestaratri MURTIYOSO**

soutenue publiquement le 9 juillet 2020

pour obtenir le grade de **Docteur de l'Université de Strasbourg**

Discipline : Sciences de l'ingénieur

Specialité : Topographie et Géomatique

---

## Relevé 3D et classification de nuages de points du patrimoine bâti

---

### THÈSE DIRIGÉE PAR

M. Pierre GRUSSENMEYER      Professeur des universités, INSA de Strasbourg, France

### RAPPORTEURS EXTERNES

M. Andreas GEORGOPOULOS      Professeur des universités, NTUA, Grèce

M. Diego GONZÁLEZ-AGUILERA      Professeur des universités, Universidad de Salamanca, Espagne

### EXAMINATEURS

M. Pierre CHARBONNIER      Directeur de recherche, Cerema Est, France

M. Fabio REMONDINO      Directeur de recherche, FBK Trento, Italie



Alexander, king of Macedon, began to study geometry; unhappy man,  
because he would thereby learn how puny was that earth of which he had  
seized but a fraction!

---

Seneca the Younger (4 BCE-65 CE), *Epistulae Morales ad Lucilium*



*Doctoral School of Mathematics, Information and Engineering Sciences*

ICube Laboratory, TRIO Group UMR 7357 INSA Strasbourg

# THESIS

presented by

**Arnadi Dhestaratri MURTIYOSO**

defended in public on the 9<sup>th</sup> of July 2020

in order to obtain the title of **Doctor of the University of Strasbourg**

Discipline: Engineering Sciences

Speciality: Surveying and Geomatics

---

## Geospatial Recording and Point Cloud Classification of Heritage Buildings

---

### THESIS SUPERVISOR

Pierre GRUSSENMEYER

Full Professor, INSA Strasbourg, France

### EXTERNAL REVIEWERS

Andreas GEORGOPOULOS

Full Professor, NTUA, Greece

Diego GONZÁLEZ-AGUILERA

Full Professor, Universidad de Salamanca, Spain

### EXAMINATORS

Pierre CHARBONNIER

Research Director, Cerema Est, France

Fabio REMONDINO

Research Director, FBK Trento, Italy



---

*This PhD thesis was financed by a grant from the Indonesian Endowment Fund for Education (LPDP), Republic of Indonesia. All work and publications related to this dissertation was conducted within the period of April 2017 to June 2020 in the TRIO group of the ICube Laboratory at INSA Strasbourg, France.*

---





# Acknowledgements

*"Why do you want to use photogrammetry for heritage? Aren't there other higher-priority applications?"*

Thus spoke one of my old professors several years ago. I must admit that the question bugged me for some time; after all, I have spent the last decade of my life studying photogrammetry for heritage recording. I even travelled 11,000 km to learn about this fascinating mix of technology and my old passion: archaeology. Well, dear reader, I am pleased to tell you that you shall, hopefully, find the answer to the above mentioned question in this 158-page manuscript. And maybe some things more.

But I did not arrive at this point where I am writing the acknowledgements part alone. I certainly did not. As the great Isaac NEWTON aptly put it, "If I have seen further it is by standing on the shoulders of Giants". And one of those giants is undoubtedly my mentor and friend Pierre GRUSSENMEYER, who has believed in me since as long as I can remember. Thank you Pierre for all the lessons and the support you have given me thus far. Another of those giants is none other than the great Fabio REMONDINO. Fabio, if you had not come to that small workshop in Bandung back in 2011, I would not be here right now. You introduced me to Pierre and I am eternally grateful for that. And of course if it wasn't for Deni SUWARDHI, I would not have even met Fabio. Pak Deni was also the one who introduced me to the marvellous world of 3D heritage documentation in the first place, back when I was looking for a mandatory internship in 2010.

To my good friend Niclas BÖRLIN, it was a blast knowing you. Do take me on one of your flights one of these days. I should also thank Pierre CHARBONNIER who is also acting as my jury and was a member of my thesis follow-up committee. Your ideas and insight on the relation of my work to machine learning helped me to determine the path I took during the thesis. The other members of my jury, Andreas GEORGOPOULOS and Diego GONZALEZ-AGUILERA have also been supportive during the difficult moments leading to the defence.

To the members of the PAGE-TRIO laboratory of INSA Strasbourg, past and present. Hélène, you came up with many ideas in my thesis that you should have been my supervisor. My "big brother" Emmanuel MOISAN, I can always depend on you for spiritual support. My fellow PhD warriors: Elise, *Ich vermisse unsere Diskussionen zu vielen Themen über ein Stück Martabak!* Elena, always compassionate, with whom I have shared my struggles in professional and personal matters. Xiucheng, a very kind and intelligent person who is always open for discussion. And my "little brother" Rami, always positive and funny. You helped me so much with my poor math. *Ganbatte, Rami-kun!* To the other members of the lab: Tania, the kind of teacher that I always respected and strive to become. Mathieu, always eager whenever he heard that I'm using GIS in my work. Emmanuel ALBY and Christophe, thanks for the enlightening coffee break discussions. And of course, Samuel! *AHA! Tu penses que je t'avais oublié ?* Immense, immense thanks for all your help, support, and the occasional life lesson during my time in the lab.

---

Discussions with other people I met during the PhD thesis are also vastly important in shaping my research. Francesca MATRONE, thank you for your great ideas which has led to some exciting collaborations. Susana DEL POZO, always a pleasure to meet and talk to you! My other brother from the same continent, Yi-Chou LU, also gave me interesting ideas to explore. And although I cannot name everyone here, the following people with whom I have co-authored some papers were also crucial in moulding my research to its current form: Fabio MENNA, Erica NOCERINO, Budi Heri NUGROHO, and Yustisi LUMBAN-GAOL. A special shout out to Pieter-Jan "PJ" DE VOS for reminding me of the importance of the "social" aspect of heritage documentation through our biennial discussions in CIPA congresses.

To PPI Alsace, past and present, my family away from home. Thank you for all the joy, the arguments, the tension, the gossips, the food, and everything you guys have shared with me this past seven years. Thank you for keeping the organisation running, and please keep on doing it. I will be watching.

For posterity, I guess it is also important to mention that I wrote my thesis during the [2020 Covid-19 pandemic](#) lock-down in France. This surreal moment is a sad but historic point in human history, unprecedented in the entirety of my very short existence on Earth. But some people have been working hard to keep my spirits up, and I must thank them for this. First thanks go to my fellow Strasbourg "castaways" Tia, Gavriila, Jessie, "Mas" Echa. Our Whatsapp discussions were epic (and essential to keep us sane). Special thanks for Thea; thank you for all the *receh* posts and the discussions during the quarantine. They really highlighted my dull days. Thanks also to Filsa for keeping me up-to-date with the situation in Indonesia (and for the occasional mandatory gossiping, of course). Andru "everyone's *Bebek*" for never failing to listen to me even in my most challenging times.

Special thanks to the CIPA EP team: Stathis, Joe, and Rebecca for giving me the excellent opportunity to practice my defence beforehand. Thanks also to Emmanuel CLEDAT and Thibaud for their valuable inputs. A big thanks goes to my proofreading team: Dimas, Ilo, Jessie, Thea, and Peter for English. Thanks to H el ene for the excellent proofreading she did for my poor French.

For the data, I have to acknowledge the following persons for their cooperation: His Majesty the Sultan Sepuh XIV PRA Arief Natadiningrat and Iwan PURNAMA for the Keraton Kasepuhan dataset; Father J erome HESS, who is always friendly and welcoming, for the St-Pierre dataset. I'm proud to be one of the few (or the only one) to be able to point at the crucifix and say "I know the coordinates for Jesus' hand" inside this beautiful church. I thank again Francesca MATRONE for the Valentino dataset and Fabio REMONDINO for the Paestum dataset. For the St-Paul data, I must thank my good friend Julien VANDERMEERSCHEN and Tristan FREVILLE of the company DroneAlsace. Let's grab another beer sometime, on me.

Last but not least, infinite thanks to my parents to whom I dedicate this manuscript. My mother Indrajani NATAWARDAJA, the kindest person on God's good Earth. She is always by my side in all my highs and lows, and I cannot thank her enough for being there. My father, Sutrisno MURTIYOSO. The wisest man in the world, and my biggest role model. Thank you for encouraging me to go to Europe despite the odds. Thank you for being there at the lowest moments of my life with your advices and life lessons. Of truth we know nothing, for truth is in a well.

# Contents

Contents . . . . .	xi
List of Figures . . . . .	xv
List of Tables. . . . .	xix
<b>INTRODUCTION . . . . .</b>	<b>1</b>
<b>I GEOSPATIAL RECORDING AND QUALITY CONTROL . . . . .</b>	<b>7</b>
<b>Chapter 1 Geospatial Recording for Heritage . . . . .</b>	<b>9</b>
1.1 Rationale. . . . .	9
1.2 Theoretical notions . . . . .	11
1.2.1 Photogrammetry . . . . .	11
1.2.2 Laser scanning . . . . .	15
1.2.3 Drones . . . . .	18
1.3 Multi-scalar and multi-sensor workflow . . . . .	19
1.4 Complementarity and data integration . . . . .	21
1.5 Challenges and constraints . . . . .	22
1.6 Derived products . . . . .	23
1.7 Available datasets . . . . .	25
1.8 Summary. . . . .	27
<b>Chapter 2 Quality Control in Photogrammetry. . . . .</b>	<b>29</b>
2.1 Importance of quality control in photogrammetry . . . . .	30
2.2 Related work . . . . .	31
2.2.1 Bundle adjustment . . . . .	31
2.2.2 Available software solutions . . . . .	33
2.3 Design of the experiments and results . . . . .	36
2.3.1 St-Pierre dataset . . . . .	36
2.3.2 St-Paul dataset . . . . .	41
2.4 Summary. . . . .	48

<b>Chapter 3</b>	<b>Integration of Photogrammetry and Laser Scanning . . . . .</b>	<b>51</b>
	3.1 Background and rationale. . . . .	52
	3.2 State-of-the-art . . . . .	53
	3.3 Proposed integration method . . . . .	54
	3.4 Experiments and results . . . . .	56
	3.4.1 Independent georeferencing . . . . .	56
	3.4.2 Free-net registration and subsequent georeferencing . . . . .	57
	3.4.3 Check point validation . . . . .	58
	3.5 Attempts on automation . . . . .	59
	3.6 Summary. . . . .	61
<b>II POINT CLOUD PROCESSING ALGORITHMS . . . . .</b>		<b>63</b>
<b>Chapter 4</b>	<b>The M_HERACLES Toolbox. . . . .</b>	<b>65</b>
	4.1 Principal motives of the development. . . . .	66
	4.2 General state-of-the-art. . . . .	67
	4.2.1 Approaches to point cloud classification . . . . .	68
	4.2.2 Point cloud processing techniques. . . . .	69
	4.3 Supporting datasets . . . . .	74
	4.4 Summary. . . . .	76
<b>Chapter 5</b>	<b>From Heritage Complex to Heritage Buildings . . . . .</b>	<b>79</b>
	5.1 Rationale. . . . .	80
	5.2 Similar work . . . . .	80
	5.3 GIS-aided segmentation and annotation. . . . .	81
	5.4 Experimental results . . . . .	83
	5.5 Comparison with a commercial solution. . . . .	89
	5.6 Summary. . . . .	91
<b>Chapter 6</b>	<b>From Heritage Buildings to Architectural Elements . . . . .</b>	<b>93</b>
	6.1 Motivations of the study . . . . .	94
	6.2 Detection of structural supports . . . . .	94
	6.3 Detection of beams in structural frameworks . . . . .	97
	6.4 Experimental results . . . . .	100
	6.4.1 Structural supports . . . . .	100
	6.4.2 Beams . . . . .	108
	6.5 Summary. . . . .	109

**CONCLUSIONS . . . . . 111**

**Scientific Contributions . . . . . 117**

**Bibliography . . . . . 123**

**Appendices . . . . . 143**



# List of Figures

1	The principal steps in the 3D recording and modelling process of heritage buildings, as well as their respective correspondences with the parts described in this thesis. Notice that the HBIM part is in dashed lines, as it is not be the main emphasis of this thesis. . . . .	4
1.1	General photogrammetric workflow as adapted from <a href="#">LUHMANN ET AL. [2014]</a>	12
1.2	General classification of image matching methods, adapted from <a href="#">REMONDINO [2014]</a> and <a href="#">SZELISKI [2010]</a> . . . . .	14
1.3	Examples of TLS available in the market and their characteristics as of February 2020. . . . .	17
1.4	The multi-scalar and multi-sensor approach illustrated. Here the Kasepuhan dataset was used for illustration purposes. . . . .	19
1.5	Overall workflow for the integration of photogrammetry and laser scanning.	21
1.6	Example of a section of a panoramic image captured by the TLS. The left figure shows the RGB image, with strong overexposure due to sunray intensity. The right figure shows the intensity image, which was preferred in the tie point identification process. . . . .	23
1.7	A photo collage of the possible derivative products from 3D heritage recording. The examples shown here are taken from the various tests performed during the thesis. . . . .	24
1.8	The three main datasets acquired during the period of the thesis. Note that the St-Paul data are incomplete; hence its sole use for photogrammetric quality control in Chapter 2. . . . .	26
2.1	Comparison between the commands in Aperio and menu options in PhotoScan and their roles in the photogrammetric workflow. For Aperio/MicMac, darker green denotes the high level automated commands, while lighter green denotes lower level XML-based commands. . . . .	35
2.2	(a) One image of the main façade in the St-Pierre dataset. (b) An orthophoto of the façade. Red triangles denote GCPs. Green triangles denote CPs. . . . .	36
2.3	Results of the bundle adjustment showing the orientation of the photos in PhotoScan (a) and DBAT (b). . . . .	37
2.4	The radial distortion curves corresponding to the estimated K1-K3 parameters for S1 (a), S2 (b), S3 (c), and S4 (d). "PS" denotes PhotoScan results.	39
2.5	Histogram of the exterior orientation standard deviations for S4. . . . .	41
2.6	Distribution of the ground control points and the check points on the eastern façade of St-Paul's church. Green triangles denote GCPs while orange triangles denote CPs. . . . .	43



2.7	Flow chart of the conducted experiments involving PhotoScan, DBAT, and Apero. Two experiments were performed in order to compare PhotoScan to DBAT and Apero respectively, followed by a project quality control based on the bundle adjustment metrics reported by open source solutions. . . .	44
2.8	Results of the external orientation process shows the positions and attitudes of each camera station, together with the sparse point clouds generated from the respective feature matching process. PhotoScan and DBAT processed the same tie point measurements. Apero generated its own set of tie point measurements from the same input images as PhotoScan. . . .	44
2.9	Project statistics generated by DBAT related to the quality of the images and their exterior orientation parameters. Large point residuals or high standard deviations of the exterior parameters provide clues to problematic areas. . . . .	47
2.10	Project statistics generated by Apero related to the quality of tie points used in the bundle adjustment process. . . . .	47
3.1	The two point clouds used in the experiment: (a) exterior point cloud generated by UAV and terrestrial photogrammetry, and (b) interior point cloud generated by TLS. . . . .	53
3.2	The two approaches to point cloud registration and georeferencing used in this experiment. . . . .	54
3.3	Horizontal profiles of the exterior (green) and interior (red) data showing the tie point distribution used for the free-net registration. . . . .	55
3.4	The two point clouds registered and georeferenced using the independent method. The blue-coloured part represents the exterior while the green-coloured one represents the interior. . . . .	57
3.5	Distribution of the measured checkpoints superimposed on the nadiral orthophoto of the church. The radii of the circles around the points represent the 3D residuals of each point, magnified by a factor of 500 for presentation purposes. . . . .	59
3.6	The steps in the keypoint detection and matching: (a) image of an example window opening. Here is shown the exterior side; (b) segmented point cloud of this window; (c) window pane extracted using a RANSAC algorithm; (d) detected SHOT keypoints. Red denotes exterior keypoints and green interior keypoints. . . . .	60
4.1	The overall pipeline for 3D reconstruction, from a point cloud up to HBIM-compatible 3D models. M_HERACLES development focuses on the manual bottlenecks of the pipeline (red inverted trapeziums) up to before the 3D modelling process (long-dashed elements), although a preliminary result of automatic 3D modelling of beam structures will also be briefly presented in Chapter 6. . . . .	67
4.2	Illustration of the Hough transform concept, with (a) representing arbitrary 2D points and (b) the corresponding curves in the parameter space. Adapted from <a href="#">VOSSELMAN AND MAAS [2010]</a> . . . . .	71
4.3	Examples of the use of RANSAC in detecting 3D geometric objects: (a) detection of a cylinder on a pillar, (b) detection of TLS spheres, and (c) detection of the floor component on a point cloud. . . . .	72

4.4	Different approaches to the 3D region growing problem. <b>M_HERACLES</b> implemented the voxel based with smoothness constraint region growing. . . .	74
4.5	The three supporting datasets used in tests for <b>M_HERACLES</b> in its different scale steps and classes. . . . .	75
4.6	A general flowchart of the <b>M_HERACLES</b> workflow. The first step consists of segmentation from the scale level of a complex/neighbourhood to that of a building, while the second step involves segmentation from a building's scale level to that of architectural elements. Violet rectangles denote the use of third party libraries, while red rectangles are the main functions developed for the tasks. . . . .	76
5.1	Illustration of the cookie-cutter segmentation. Notice that in the result all points of all altitude values within the confines of the GIS vector are included in the cluster. A region growing method was subsequently employed to acquire the desired object. . . . .	82
5.2	The unorganised and unclassified point cloud of Kasepuhan (left) and St-Pierre (right) used as inputs for the algorithm. . . . .	83
5.3	The GIS shapefile data used to help the segmentation process. To the left, three shapefiles were available for the Kasepuhan dataset while to the right, only one shapefile entity was used for the St-Pierre dataset. . . . .	83
5.4	Results of the automatic segmentation and annotation for the Kasepuhan dataset: the BUILDINGS class (a), the WALLS class(b) and the GATES class (c). . . . .	84
5.5	Results of the automatic segmentation and annotation for the St-Pierre dataset. Here only six of the most important buildings are shown, of which <b>BatimentsPublic3</b> represents the St-Pierre church which will be further processed in Chapter 6. Note also that only the aerial LIDAR data are shown as opposed to the denser and more detailed drone and TLS data in Figure 5.4. . . . .	85
5.6	A profile of the Kasepuhan dataset to illustrate the segmented classes and the ordering of the use of the <b>shapeseg</b> function, starting from lower objects to higher ones. . . . .	86
5.7	Visual comparison of the point cloud segmentation process: (a) displays the manual classification used as reference, (b) the results of the Metashape automatic classification, and (c) the results from <b>M_HERACLES</b> . The colour blue denotes the 'buildings' class, orange the 'walls' class, and green the 'trees' class. . . . .	89
5.8	Histogram representation of the classification performance for each class in the Kasepuhan dataset (left) and the median value of the principal quality parameters (right). . . . .	90
5.9	Histogram representation of the quality parameters for the St-Pierre dataset comparing Metashape against <b>M_HERACLES</b> . . . . .	91
5.10	Case of mis-classification in the results of the Metashape point cloud classification on the St-Pierre dataset. Here are shown some examples where the problem is most observable: (a) <b>STPIERRE</b> , (b) <b>PLSJUSTICE</b> and (c) <b>DIRIMPOTS</b> sub-clouds. The colour blue denotes the 'buildings' class and green the 'high vegetation' class. Note also the presence of the crane in (b). . . . .	91

6.1	A flowchart showing the main steps in the <code>beamdetect.m</code> function of <code>M_HERACLES</code> , showing the intermediate and final results. . . . .	98
6.2	Automatically segmented building attic and body for the Kasepuhan Royal Pavilion (left) and Musicians' Pavilion (right) datasets; the attic is shown in green and the body in purple. . . . .	101
6.3	Visual results of the segmentation and classification process for the 5 tested cases. For the segmentation part each colour denotes individual detected objects, while for the classification red denotes the "column" class and blue the "non-column" class. . . . .	102
6.4	Illustration of some problems encountered during the segmentation process with sample pillar result clusters from: (a) Kasepuhan Royal Pavilion, (b) St-Pierre, and (c) Paestum. Blue colour denotes true positive points, while grey ones are false positive points. . . . .	103
6.5	Histogram representation of the quality parameters for the Valentino dataset comparing the results of <a href="#">MALINVERNI ET AL. [2019]</a> against <code>M_HERACLES</code> for the column class. . . . .	108

# List of Tables

2.1	Four self-calibration scenarios were tested. In the first three scenarios, the EXIF value for the focal length was used as initial values. A fourth scenario employs the precalibrated initial values. . . . .	37
2.2	The estimated parameters and standard deviations for the four tested scenarios. "PS" denotes PhotoScan while the "PreCal" column presented the precalibrated values generated by laboratory calibration. . . . .	38
2.3	Results for the different scenarios, showing the $\sigma_0$ , reprojection error RMS, GCP error RMS, and CP error RMS. . . . .	40
2.4	Correlations in the processed projects in DBAT. For the automatic points, the value denotes the percentage of automatic tie points with high correlation values. . . . .	41
2.5	GCP and CP 3D residuals for PhotoScan, DBAT, and Apero. . . . .	46
3.1	RMS of the tie point residuals used for the registration of the exterior and interior point clouds using the free-net approach. . . . .	58
3.2	RMS of the control point residuals used in the georeferencing process using the free-net approach. . . . .	58
3.3	RMS of the checkpoint residuals for all tested methods. . . . .	60
5.1	Table of the quantitative analysis on the results of step 1 for Kasepuhan. In this table, three classes were taken into account (buildings, gates, and walls) with a total of 13 objects. %P is precision, %R is recall, and %F1 is the normalised F1 score. . . . .	87
5.2	Table of the quantitative analysis on the results of step 1 for St-Pierre. In this table, only one class was taken into account (public buildings) with a total of 6 out of 17 objects used in the statistical analysis. %P is precision, %R is recall, and %F1 is the normalised F1 score. . . . .	87
5.3	Comparative table showing the quantitative results of the classification for Kasepuhan using Metashape and M_HERACLES. . . . .	90
6.1	General parameters and overview of the segmentation and classification results. . . . .	101
6.2	Table of the quantitative analysis on the results of step 2 for the detection and classification of columns in the Kasepuhan Royal Pavilion dataset. %P is precision, %R is recall, and %F1 is the normalised F1 score. . . . .	105
6.3	Table of the quantitative analysis on the results of step 2 for the detection and classification of columns in the Kasepuhan Musicians' Pavilion dataset. %P is precision, %R is recall, and %F1 is the normalised F1 score. . . . .	106

6.4 Table of the quantitative analysis on the results of step 2 for the detection and classification of columns in the St-Pierre dataset. %P is precision, %R is recall, and %F1 is the normalised F1 score. . . . . 107

6.5 Table of the quantitative analysis on the results of step 2 for the detection and classification of columns in the Valentino dataset. Note that only the detected columns were taken into account here. %P is precision, %R is recall, and %F1 is the normalised F1 score. . . . . 107

6.6 Table of the quantitative analysis on the results of step 2 for the detection and classification of beams. %P is precision, %R is recall, and %F1 is the normalised F1 score. . . . . 109

# Abstract

The documentation of built heritage has seen a significant development these past few decades due to advancements in new 3D sensors and 3D recording techniques. 3D data serve as reliable and tangible archive for historical sites and monuments. Furthermore, improvements in processing algorithms and hardware have led to the democratisation of 3D recording techniques. The term "reality-based 3D modelling" was thus introduced. Nowadays, two mutually complementary 3D recording techniques dominate the 3D recording domain: the passive image-based method and the active range-based method. Photogrammetry has been the main image-based reconstruction method since its early conception over 150 years ago. The field of photogrammetry has also seen much development with aid from the computer vision domain, such as the Structure from Motion (SfM) and dense matching algorithms. Photogrammetry has since become a winning solution, often possible to employ in a low-cost setting, for the 3D recording of heritage sites. Developments in optical sensors and also the rise of drones have increased further its popularity amongst the heritage community. On the other hand, the LIDAR or laser scanning presents the main method for the range-based active technique. Since 3D data have such importance in the field of heritage documentation, quality control is paramount and must be performed before any point cloud processing is even planned to be conducted. The thesis is therefore divided into two parts. The first part concerns mainly the data acquisition and quality control of the point cloud data using the two techniques most commonly used, i.e. photogrammetry and laser scanning. A particular emphasis is also put on the integration of photogrammetry and laser scanning within the context of a multi-scalar documentation of a heritage site. The second part addresses the processing of the resulting point cloud, particularly its segmentation and classification. The multi-scalar approach proposed in this thesis is an important point to note, as in many cases a historical building of interest is located in a historical neighbourhood. The developed algorithm thus enabled first the automatic segmentation and classification of point clouds from the scale step of a heritage complex into that of individual heritage buildings of interest. The resulting point cloud clusters were thereafter segmented and classified further into architectural elements in order to extract objects automatically. The automatic detection and subsequent segmentation of two classes of architectural elements will be presented, namely pillars and beams. In this thesis, we show that the developed algorithms are practical, fast, and in many cases sufficient. By combining the quality control developed in the first part with these functions, the thesis attempts to address the 3D workflow of heritage sites in a holistic manner, from the 3D data acquisition up to the resulting point clouds' segmentation and classification into individual entities in various scale steps.

**Keywords:** automation, photogrammetry, point cloud, 3D modelling, classification, heritage

# Résumé

La documentation du patrimoine bâti a beaucoup évolué ces dernières années grâce au développement de nouveaux capteurs 3D et de nouvelles techniques de relevé 3D. Les données 3D contribuent à la création d'archives fiables et tangibles des sites et des monuments historiques. Les progrès au niveau des algorithmes et des moyens de calculs ont permis de démocratiser la mise en œuvre des relevés 3D. L'usage du terme « modélisation 3D photoréaliste » s'est généralisé. Aujourd'hui, ce type de modélisation repose sur deux méthodes complémentaires: la méthode passive basée sur le traitement d'images et la méthode active basée sur la mesure directe de distances par balayage laser. La photogrammétrie et ses 150 ans d'existence représente la technique la plus courante dans l'approche basée sur l'utilisation d'images. La photogrammétrie s'est enrichie de techniques issues du domaine de la vision par ordinateur comme la structure par le mouvement (*Structure from Motion* ou SfM) et des algorithmes d'appariement pour le calcul de nuages de points denses. La photogrammétrie est ainsi devenue une solution fiable et à faible coût pour le relevé 3D du patrimoine. Les développements des capteurs photographiques et la généralisation de l'utilisation des drones ont largement contribué à la popularité de la photogrammétrie. Dans l'approche basée sur la mesure de distances, la technologie LiDAR s'est beaucoup développée. Vu l'importance des données 3D dans la documentation du patrimoine bâti, le contrôle de qualité est un aspect primordial qui devrait être abordé avant d'entreprendre le traitement du nuage de points. La thèse est ainsi divisée en deux parties. La première partie concerne principalement l'acquisition et le contrôle de qualité des données. Un point important sera l'intégration de la photogrammétrie et de la lasergrammétrie dans le contexte de la documentation d'un site historique à différentes échelles. La deuxième partie de la thèse va aborder le traitement de nuages de points, plus particulièrement la segmentation et la classification de nuages de points. L'aspect multi-échelle de notre approche est importante car dans beaucoup de cas, un bâtiment remarquable se situe dans un quartier historique qui nécessite une segmentation multi-échelle. L'algorithme développé pendant la thèse permettra la segmentation et la classification automatique du nuage de points d'un quartier historique en autant de nuages de points distincts que de bâtiments. Ces groupes de nuages de points sont ensuite segmentés et classifiés en éléments architecturaux pour extraire automatiquement des objets tels que des piliers et les poutres d'une charpente. Nous avons montré que les algorithmes développés sont pratiques, rapides et dans beaucoup de cas ne nécessitent pas l'utilisation d'une approche plus complexe. En combinant le contrôle de qualité développé dans la première partie avec ces fonctions, nous avons considéré l'ensemble du processus allant de l'acquisition de données 3D jusqu'à la segmentation et la classification en entités à plusieurs échelles.

**Mots-clès :** automatisation, photogrammétrie, nuage de points, modélisation 3D, classification, patrimoine

# Ringkasan

Dokumentasi cagar budaya telah berkembang sangat pesat selama beberapa dekade ini dengan berbagai kemajuan di bidang sensor dan metode pemetaan trimatra (3D). Selain itu, kemajuan dalam bidang perangkat lunak maupun perangkat keras juga telah membuat teknologi pemetaan 3D semakin lazim digunakan. Perkembangan inilah yang kemudian memunculkan istilah "pemodelan 3D berbasis kenyataan". Saat ini terdapat dua teknik yang sering digunakan dalam pemetaan 3D. Kedua teknik ini bersifat komplementer, dan terdiri dari teknik pengukuran pasif berbasis citra dan teknik pengukuran aktif berbasis pengukuran jarak. Teknik untuk menghasilkan data 3D berbasis citra terutama direpresentasikan oleh ilmu fotogrametri sejak 150 tahun yang lalu. Bidang ilmu fotogrametri juga telah berkembang pesat dengan bantuan ilmu komputer, terutama penglihatan komputer melalui algoritma *Structure from Motion* (SfM) dan pencocokan citra untuk menghasilkan awan titik rapat. Dengan demikian fotogrametri sering kali menjadi solusi utama dan relatif berbiaya rendah untuk dokumentasi cagar budaya. Perkembangan di bidang optik dan wahana nir-awak semakin meningkatkan popularitas fotogrametri di kalangan pemerhati cagar budaya. Di sisi lain, teknik berbasis pengukuran jarak sering kali menggunakan teknologi LIDAR atau pemindaian laser dalam melaksanakan pemetaan 3D. Karena data 3D bersifat sangat penting dalam dokumentasi cagar budaya, kontrol kualitas sangat penting untuk dilakukan sebelum awan titik yang dihasilkan dapat diolah lebih lanjut. Untuk itu, disertasi ini dibagi menjadi dua bagian besar. Bagian pertama akan membahas pengambilan data dan kontrol kualitas untuk awan titik yang diperoleh. Selain itu, integrasi antara data fotogrametri dan pemindai laser dalam konteks pemetaan multi-skalar menjadi salah satu tema penting dalam bagian ini. Bagian kedua akan membahas pengolahan awan titik, terutama segmentasi dan klasifikasi. Pendekatan multi-skalar yang diajukan dalam disertasi ini penting karena dalam banyak contoh sebuah bangunan cagar budaya sering kali terletak di dalam sebuah kompleks bersejarah. Untuk itu segmentasi dan klasifikasi awan titik juga dilakukan menggunakan pendekatan multi-skalar. Algoritma yang dikembangkan dalam disertasi ini pertama-tama melakukan segmentasi dan klasifikasi dari skala sebuah kompleks cagar budaya dan menghasilkan sejumlah awan titik individual dari bangunan-bangunan cagar budaya. Hasil dari segmentasi ini kemudian digunakan sebagai data masukan untuk proses segmentasi berikutnya untuk mendeteksi secara otomatis unsur-unsur detail arsitektural seperti pilar dan kisi-kisi bangunan. Dalam disertasi ini, hasil menunjukkan bahwa algoritma yang dikembangkan bersifat praktis, cepat, dan dalam banyak kasus cukup tanpa perlu pengolahan lain yang lebih kompleks. Dengan menggabungkan kontrol kualitas pada bagian pertama dan fungsi-fungsi yang dikembangkan pada bagian kedua ini, disertasi ini telah mencakup keseluruhan alur kerja 3D untuk dokumentasi cagar budaya; dari pengambilan data 3D sampai dengan segmentasi dan klasifikasi awan titik menjadi entitas tunggal dalam berbagai tingkatan skala.

**Kata kunci** : otomatisasi, fotogrametri, awan titik, pemodelan 3D, klasifikasi, cagar budaya





# Acronyms

<b>ALS</b>	Aerial Laser Scanner
<b>BIM</b>	Building Information Model
<b>CAD</b>	Computer Aided Design
<b>CP</b>	Check Point
<b>CSF</b>	Cloth Simulation Filtering
<b>DBAT</b>	Damped Bundle Adjustment Toolbox
<b>DEM</b>	Digital Elevation Model
<b>GCP</b>	Ground Control Point
<b>GIS</b>	Geographical Information System
<b>GNSS</b>	Global Navigation Satellite System
<b>HBIM</b>	Heritage Building Information Model
<b>ICP</b>	Iterative Closest Point
<b>LIDAR</b>	Light Detecting And Ranging
<b>M_HERACLES</b>	HERitage by point CLOUD procESSing for Matlab
<b>PCA</b>	Principal Component Analysis
<b>PCL</b>	Point Cloud Library
<b>RANSAC</b>	RANdom SAmples Consensus
<b>RMS</b>	Root Mean Square
<b>TLS</b>	Terrestrial Laser Scanner
<b>UAV</b>	Unmanned Aerial Vehicle



# Introduction

Documentation of heritage objects by means of surveying techniques has a long history. Indeed, surveying has been an integral part of any conservation effort as well as archaeological missions since the early days of heritage conservation [BANNING, 2002]. The need for geospatial data was, and remains important in order to present a real and tangible archive. While 3D techniques have been used since at least several decades for this purpose, they have seen a very important development since the beginning of the third millennium. This is due to significant improvements in the quality of 3D recording sensors, including the invention of the laser scanning or LIDAR (LIght Detection And Ranging) technique [BRYAN ET AL., 2004]. Fast and accurate heritage recording became feasible, although it remained an expensive endeavour. However, during the last decade, further improvements in both hardware and software have rendered the 3D nature of heritage documentation more and more ubiquitous. The term "reality-based 3D modelling" was introduced, which nowadays depends on mainly two methods: passive or image-based and active or range-based [REMONDINO AND RIZZI, 2010].

Photogrammetry represents the most commonly used technique in the image-based approach. As a branch of science which has a long history in 3D data generation since the advent of aerial photography during the early 20<sup>th</sup> century [GRUSSENMEYER ET AL., 2002], photogrammetry has seen massive improvements in terms of computation capability as well as the results offered. The traditionally surveying-oriented photogrammetric process was augmented by various techniques from the computer vision domain, such as Structure from Motion (SfM) and dense matching, to create a versatile and relatively low-cost solution for 3D heritage recording [REMONDINO AND RIZZI, 2010]. Of course, improvements in lens and sensor capabilities as well as the recent democratisation of drones or UAVs (Unmanned Aerial Vehicles) have also improved photogrammetry's popularity even further amongst the heritage community [MURTIYOSO ET AL., 2016].

As far as the active range-based approach is concerned, the LIDAR technology (including both the Terrestrial Laser Scanning or TLS and the Aerial Laser Scanning or ALS) has also developed tremendously. Using as comparison parameter the scan rate of Time of Flight (ToF) devices produced for example by Trimble, the point per second rate has improved exponentially from 5,000 points/second (Trimble GX) in 2005 [HANKE ET AL., 2006] to 25,000 points/second (Trimble SX10) in 2017 [LACHAT ET AL., 2017] and up to 100,000 points/second (Trimble X7) in 2020<sup>1</sup>. This is also supported by significant improvements in the software part, where the workflow automation is taking a far more important part aided by ever more powerful computing cores available for the user [LACHAT ET AL., 2018; HILLEMANN ET AL., 2019].

Laser scanning technology is the most commonly used technique according to MACHER ET AL. [2015] as it is suitable for both indoor and outdoor applications. However, pho-

---

<sup>1</sup><https://geospatialx7.trimble.com>, retrieved 22 January 2020.

togrammetry both by traditional close-range terrestrial cameras and by UAV or drones are also widely used to complement laser data [CHIABRANDO ET AL., 2016; MURPHY ET AL., 2009]. One of the primary objectives of this thesis is therefore the integration of geospatial data acquired by different sources, in preparation for further processing. Furthermore, the data coming from various sources can be quite heterogeneous in terms of their respective qualities. It is therefore very important to properly register the results while keeping in mind the geometric precision [LACHAT ET AL., 2016]. This precision aspect is an important factor in every spatial application, more so in detail-rich heritage buildings. Quality control is thus an important aspect of the workflow. With this fact in mind, the 3D data acquisition, the quality control, as well as the integration of multi-sensor data constitute the bulk of the first part of this dissertation. In this regard, we present a multi-sensor and multi-scalar 3D recording approach for heritage sites.

The most common product of the 3D recording process is a 3D point cloud, either via direct laser scanning or the deployment of dense matching algorithms on an oriented photogrammetric network. The point cloud stores geometric information (i.e. Cartesian coordinates) which forms a 3D representation of the thence scanned object [BARSANTI ET AL., 2014]. Several other information can also be stored within the point cloud, commonly other geometric features such as point normals, curvatures, linearity, and planarity (relative to a local plane) [WEINMANN ET AL., 2015] as well as RGB colour or scan intensity in the case of a TLS point cloud. However, this information remains singular for each point within the point cloud. In order to be able to perform more meaningful operations on the point cloud, segmentation must be performed and followed by semantic labelling; thus virtually disassembling the raw point cloud into smaller, classified clusters [MURPHY ET AL., 2013]. The segmented clusters of point clouds can thenceforth be treated as classified point cloud, from whence various analyses, 3D modelling, and model predictions could be performed. Indeed, a special subset of Building Information Model (BIM) is dedicated for heritage buildings, dubbed the Heritage Building Information Model (HBIM)[MURPHY ET AL., 2009] which enables such operations to be conducted.

BIM is used by the architecture and civil engineering (ACE) community for the 3D representation of structures. Various multi-disciplinary data and information can be integrated within BIM and shared between the different stakeholders [NIETO ET AL., 2016; QUATTRINI ET AL., 2015]. HBIM has the general advantages of BIM, such as the possibility to perform prototyping, visualisation, collaboration, energy simulation, comparison of different design options, solar study, and energy demand prediction. However, HBIM is specifically tailored for heritage buildings, thus providing an automated documentation in the form of engineering drawings for precise conservation. It also provides a review of the building's exterior and interior. All these advantages lead to the ease and optimisation of the management, survey renovations and eventual changes to the structure, rendering these processes much more effective and efficient [KHODEIR ET AL., 2016; NIETO ET AL., 2016]. HBIM can automatically create cut sections, details and schedules in addition to orthographic projections and 3D models (wire frame or textured) for both the analysis and conservation of historic objects, structures and environments [MURPHY ET AL., 2013]. Some other uses of HBIM include fire prevention of the said heritage site [BIAGINI, 2016], structural simulation and conservation analysis [DORE ET AL., 2015], and heritage tourism as well as the raising of public awareness through augmented reality applications [BARAZZETTI ET AL., 2015]. Being a collaborative platform, HBIM enables architects, archaeologists, art historians, restorers, and photographers and other experts to access the same integrated pool of information [MEGAHED, 2015].

Nevertheless, as has been previously mentioned the point cloud needs to be segmented and classified before the 3D modelling for HBIM even begins. This segmentation and classification step remains largely a manual process; an operator manually segments and labels each point cloud cluster as the intended entity. This process is similar and even analogous to traditional digitising in the 2D case of aerial or satellite images into 2D vectors with attributes in GIS (Geographical Information System). The manual segmentation and labelling is further complicated in the case of heritage objects due to its inherent complexity in terms of architectural style, materials, age of the structure, infinitely diverse decorations, etc. It is therefore typically the most time consuming part in the 3D pipeline [YANG ET AL., 2016] and therefore presents a bottleneck in the general workflow.

The automation of the 3D modelling process has been a very interesting subject of research in recent years. The modelling of planar surfaces or façades has been studied in many works (see e.g. DORE AND MURPHY [2013] and PU AND VOSSELMAN [2006]). These approaches often employ surface-normal similarity and coplanarity in patches of vectorial surfaces. Other uses robust algorithms such as the Hough transform or RANSAC to detect the surface. Once a planar region is detected, the parameters of the plane can be estimated using total least squares fitting or robust methods that are less affected by points not belonging to the plane [TANG ET AL., 2010]. Other methods for point cloud segmentation includes region growing algorithms [BASSIER ET AL., 2017a; VO ET AL., 2015], or approaches based on machine learning [GRILLI ET AL., 2018]. In regards to indoor modelling, the methods are mainly based on geometric primitive segmentation. Some approaches are based on space segmentation before applying a more detailed segmentation [MACHER ET AL., 2015]. The segmentation of planes is then performed using robust estimators such as MLESAC, which uses the same sampling principle as RANSAC but the retained solution maximizes the likelihood rather than just the number of inliers [MACHER ET AL., 2016].

Some challenges in the 3D pipeline of heritage buildings include problems related to the complexity of certain heritage buildings. In this case, the Level of Detail (LoD) requirements of heritage objects are often higher [DORE ET AL., 2015; LACHAT ET AL., 2016; YANG ET AL., 2016]. One of the main topics is the standardisation offered by the current BIM technology used to manage simple buildings and constructions [ORENI ET AL., 2014], limited by the irrelevance of object libraries and the inability of 3D scans to determine structures in buildings of dissimilar age and construction. Indeed, several researches [ELIZABETH AND PRIZEMAN, 2015; ORENI ET AL., 2013] focused directly on enhancing the existing libraries of historical parametric objects for HBIM.

Figure 1 shows an illustration of the different aspects of this thesis. Geospatial data are the base of all operations in this research. Quality control is therefore an important aspect which must be addressed for all datasets to be used. The first part of the thesis concerns the quality control of these data, as well as their integration. The second part of the thesis concerns the segmentation and classification of the resulting point cloud. The focus of this part of the thesis will be on the automation part, in order to reduce as much as possible human intervention in the 3D classification phase. HBIM is the "crown", so to speak, of the whole 3D pipeline as presented in Figure 1. However, this research will not delve into the automation of HBIM models, even though it will be mentioned in several places as it is very much pertinent to the 3D pipeline. Thus, the thesis has two main objectives: firstly the quality control and integration of geospatial data and secondly the automation of point cloud classification; both in the context of heritage buildings. Both goals form the globality of the 3D recording pipeline as showcased by Figure 1.

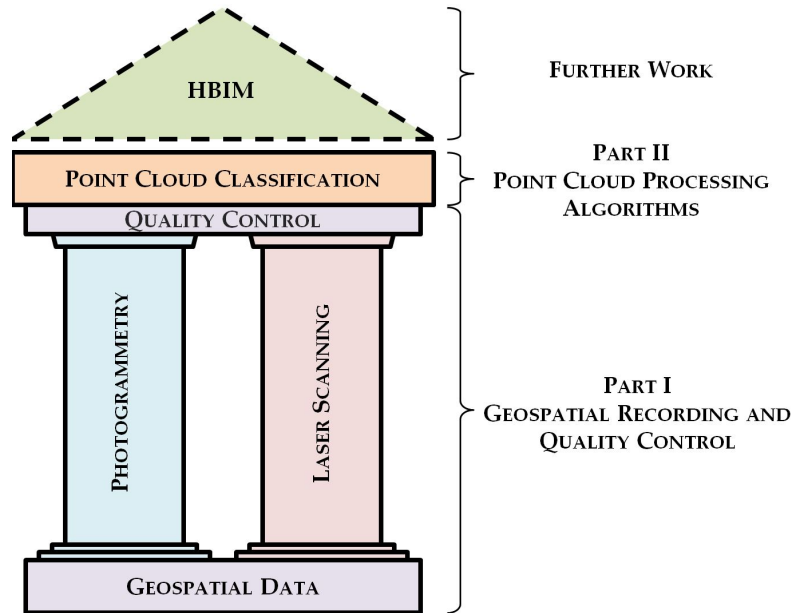


Figure 1: The principal steps in the 3D recording and modelling process of heritage buildings, as well as their respective correspondences with the parts described in this thesis. Notice that the HBIM part is in dashed lines, as it is not be the main emphasis of this thesis.

In the first part, we attempt to answer the question of *how to generate point clouds for heritage buildings, all the while keeping its geometric quality*. To this end, **Chapter 1** will first present a general state-of-the-art in heritage documentation while also presenting our proposed pipeline in 3D data acquisition. **Chapter 2** will discuss the quality control for photogrammetry which is emphasised in this study since photogrammetric workflow presents more steps and therefore more possibilities for error. For the geometric control of photogrammetric projects in particular, the Matlab<sup>©</sup>-based Damped Bundle Adjustment Toolbox (DBAT) [BÖRLIN AND GRUSSENMEYER, 2014, 2013a] developed together with the University of Umeå (Sweden) was used. This toolbox provides a powerful and highly customisable tool for the evaluation of the acquired photogrammetric data. Another open source algorithm which has been tested for these purposes as a comparison is the photogrammetric triangulation module Aperio [RUPNIK ET AL., 2017]. In **Chapter 3**, quality control on laser scanning data as well as the integration between photogrammetry and laser scanning will be discussed. The open source software CloudCompare<sup>2</sup> and the scientific software 3DVEM [MUNUMER AND LERMA, 2015] were employed in this regard. This chapter will also discuss a preliminary attempt in automating the integration process; a topic which nevertheless remains beyond the scope of this thesis but proved to be an interesting starting point for further research.

The second part focused on the research question of *what do we do next once a high quality heritage building point cloud is available to us*. To try to answer this question, the point cloud processing toolbox M\_HERACLES will be introduced in **Chapter 4**. This toolbox was created during the thesis to host the point cloud processing functions as developed in the Matlab<sup>©</sup> language. In relation to the developed multi-scalar approach described in Chapter 1, the following chapters will detail the algorithms developed to perform the segmentation and classification for each scale level of the point cloud. In

<sup>2</sup><https://www.danielgm.net/cc/>, retrieved 11 February 2020

**Chapter 5**, we address the problem of small scale point cloud segmentation. In our case, it consists of the automatic segmentation and semantic annotation of the point cloud of heritage complexes (i.e. the scale of a neighbourhood) towards clusters of point clouds corresponding to individual buildings of interest. **Chapter 6** goes further in segmenting and classifying these buildings into architectural elements, namely pillars and beams. The whole Part II will therefore present a complete multi-scalar automation attempt from small to large scale point clouds in one single M\_HERACLES environment.

As interesting as the topic of heritage point cloud is, the scope of the thesis must remain restricted in order to focus on particular topics. In this regard, the thesis will not address the automatic creation of HBIM models (which was admittedly the first intention, as a logical continuation of previous research by MACHER ET AL. [2017] and YANG ET AL. [2018]). However, some preliminary attempts at the automatic generation of 3D primitives will also be presented in Chapter 6. The conclusion part will therefore try to summarise the work and the results of the thesis, all the while presenting recommendations and suggestions for further work in this particularly interesting research topic.

We see that the thesis essentially boils down to two research questions that we attempt to answer with experiments and developments, while still anchoring on previous work by various researchers. The two parts of this thesis will respectively try to describe in detail the findings and problems we encountered during the duration of the research in our attempt to answer these two general questions.





---

PART I

GEOSPATIAL RECORDING AND QUALITY  
CONTROL

---



# Chapter 1

## Geospatial Recording for Heritage

### Contents

---

<b>1.1 Rationale</b> . . . . .	<b>9</b>
<b>1.2 Theoretical notions</b> . . . . .	<b>11</b>
1.2.1 Photogrammetry . . . . .	11
1.2.2 Laser scanning . . . . .	15
1.2.3 Drones . . . . .	18
<b>1.3 Multi-scalar and multi-sensor workflow</b> . . . . .	<b>19</b>
<b>1.4 Complementarity and data integration</b> . . . . .	<b>21</b>
<b>1.5 Challenges and constraints</b> . . . . .	<b>22</b>
<b>1.6 Derived products</b> . . . . .	<b>23</b>
<b>1.7 Available datasets</b> . . . . .	<b>25</b>
<b>1.8 Summary</b> . . . . .	<b>27</b>

---

### 1.1 Rationale

Heritage documentation is an important aspect in any conservation effort. Indeed, past history has shown the importance of archiving and documentation for the physical state of historical monuments and objects. This is becoming more important in the presence of threats, both natural [ACHILLE ET AL., 2015; BAIOCCHI ET AL., 2013] and anthropological [GRUSSENMEYER AND AL KHALIL, 2017; MENOU, 2019]. The reconstruction of damaged historical sites was, and to some degree still is, the object of contention since the beginnings of archaeological excavations and cultural heritage preservation. One needs only to see some historical examples, ranging from the 19<sup>th</sup> century controversies surrounding VIOLLET-LE-DUC’s work on French cultural heritage up to the debate on the reconstruction of the Ramses II colossus at Luxor happening as recently as 2019. Nevertheless, several attempts at regulating this problem has been made. The Venice Charter of 1964 crystallised the significance of holistic understanding of the object in question before any reconstruction effort is even attempted. Article 16 of the Charter even suggested:

*"In all works of preservation, restoration or excavation, there should always be precise documentation in the form of analytical and critical reports, illustrated with drawings and photographs..."*

The wording "precise" can therefore be safely interpreted in the sense of geospatial data. Indeed, surveying techniques have long been a staple of archaeological missions. Starting with the creation of 2D maps, surveying and thenceafter geospatial techniques evolved to support preservation efforts in many other fields. The invention of the GIS was also welcomed in the heritage documentation community [FABBRI ET AL., 2012], a precursor to modern HBIM. The entry of photogrammetry and laser scanning into the heritage documentation field also generated a small revolution in terms of exponentially decreasing the effort required to record heritage objects. These techniques provide the base for the current trend towards the integration of geospatial data not only for the recording, but also the management of heritage sites.

Apart from the traditional 2D drawings and photographs, digital 3D documentation of historical sites nowadays presents a useful tool in the analysis and interpretation of historic buildings, as well as its eventual reconstruction in the event of damage. Indeed, the addition of the third dimension presents a different perspective to the record which a 2D representation simply is not enough to showcase. That being said, a preference for 2D documentation still persists in some cases [ACHILLE ET AL., 2012; SEMLER ET AL., 2019], especially in historical sites with a long history of 2D archiving. However, this is slowly giving way to three-dimensional forms as 2D representations can be generated from a 3D archive easily [FASSI ET AL., 2017]; the inverse is not true. Furthermore, 3D representation opens the possibility to various useful products such as 3D printing (and also prototyping), augmented reality (AR) as well as virtual reality (VR), and more importantly HBIMs.

Digital documentation uses various sensors in capturing the reality. This may be done either by image-based or range-based techniques [REMONDINO ET AL., 2012]. Each technique has its own advantages and disadvantages; as such it is not rare to see the combination of both techniques in a thorough documentation project [GRENZDÖRFFER ET AL., 2015; MURTIYOSO ET AL., 2018b]. Indeed, the integration of heterogeneous data is an important research issue [LACHAT ET AL., 2018]. The Terrestrial Laser Scanner (TLS) is a range-based system that enables the capture of many points in a short period of time [GRUSSENMEYER ET AL., 2012a], and has seen a lot of use in the field of heritage documentation [LACHAT ET AL., 2017; LERMA ET AL., 2010]. As regards to image-based techniques, photogrammetry has in recent times seen important leaps, and even more so with the advent of drones [MURTIYOSO ET AL., 2016; REMONDINO ET AL., 2013]. Both range-based and image-based techniques generate dense point clouds from which derivative geospatial products may be generated.

In the context of the documentation of heritage complexes, the extent of the site makes it logical to use a multi-scalar approach. Larger areas do not need a fine resolution data, as opposed to smaller buildings or even architectural elements. In this approach, the site is digitised in several scale steps, according to the required resolution for each level [FIORILLO ET AL., 2013]. The use of the multi-scalar approach also means that more than one sensor could be employed in order to cover each scale step. Furthermore, for the case of urban areas, many constraints such as the geography and urban density means that the use of one single sensor may not be sufficient. It is in this regard that multi-sensor and multi-scalar documentation became a logical solution to the problem of documenting historical complexes.

## 1.2 Theoretical notions

Heritage documentation is often characterised by the myriad of 3D sensing techniques available in the market today, as well as difficulties in maintaining a balance between cost, effort, and precision in choosing them. As technical developments advance, many which were cost prohibitive in the past have become more and more low cost (e.g. transition from analogue metric cameras to digital non-metric cameras in photogrammetry). That being said, the laser scanning technology has more or less become the standard solution in heritage documentation [CHIABRANDO ET AL., 2017; HERBIG ET AL., 2019; LACHAT ET AL., 2017]. This technology has been around since the 1980s and was a revolutionary technology in the domain of 3D mapping. Contrary to traditional total stations, the TLS technology enables the recording of the environment in a fast and relatively accurate manner while being fairly easy to use [MURTIYOSO ET AL., 2018b].

Photogrammetry on the other hand, is an older technology dating back to the analogue era. However, in the past two decades it has seen resurgence as image processing algorithms advanced significantly and helped by the availability of more powerful computing resources. As it requires only 2D images and dedicated processing software, photogrammetry has often been seen as a low-cost alternative to laser scanning [BARSANTI ET AL., 2014; CHIABRANDO ET AL., 2015; EVGENIKOU AND GEORGOPOULOS, 2015], even though in many cases both can be complementary; indeed the study of sensor integration is one of the more studied subjects recently [MAGDA RAMOS AND REMONDINO, 2015; MUNUMER AND LERMA, 2015; MURTIYOSO ET AL., 2018b]. Other alternative low-cost sensors also exist and have been tested for heritage objects, e.g. RGB-D cameras [LACHAT ET AL., 2015], fish-eye [PERFETTI ET AL., 2018] or panoramic 360°cameras [BARAZZETTI ET AL., 2019], and even low-end smartphones [NOCERINO ET AL., 2017b; KIM ET AL., 2013].

Architectural documentation has been performed using these methods in many research [GRUSSENMEYER ET AL., 2002]. It is often used in extracting orthophotos [CHIABRANDO ET AL., 2015], vector models of façades [CEFALU ET AL., 2013], 3D models [VERHOEVEN, 2016], as well as inputs for HBIM [QUATTRINI ET AL., 2015]. Both photogrammetry and laser scanning have been employed on the documentation of various types of vernacular buildings [AKBAYLAR AND HAMAMCIOGLU-TURAN, 2007; BROWN ET AL., 2009].

### 1.2.1 Photogrammetry

Photogrammetry has seen a lot of improvements and renewed interest in the last decade, partly due to significant developments in terms of sensor manufacturing [MURTIYOSO ET AL., 2017d]. This is also helped by breakthroughs in the field of dense matching [REMONDINO ET AL., 2014] as well as the democratisation of UAVs or drones [NEX AND REMONDINO, 2014]. As regards to heritage documentation, photogrammetry is often employed in its close-range configuration. Using this technique, flexible ground sampling distance (GSD) can be designed according to requirements. It therefore enables the capturing of intricate details. Added with the possibility to recover photo-realist texture from the images, this gives photogrammetry an advantage over other methods. UAVs give another edge to photogrammetry by adding aerial points of view. However, photogrammetric data processing requires extensive time, experience, and resources.

Historically, photogrammetry requires a well-calibrated and even metric sensors. Today, however, faster computing capabilities and algorithms have widened the scope of sensors that may be processed according to photogrammetrical processes. The DSLR

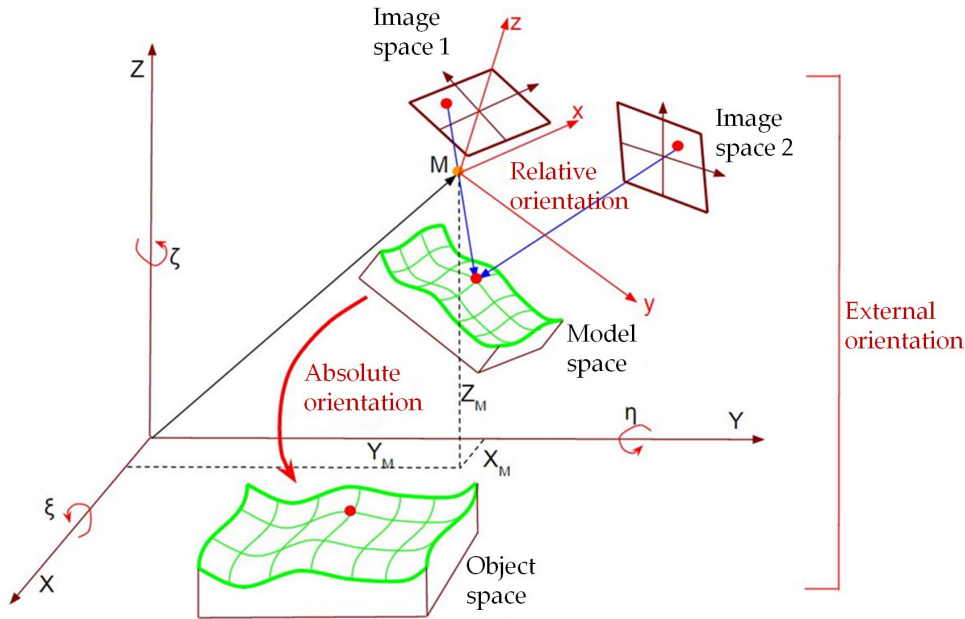


Figure 1.1: General photogrammetric workflow as adapted from LUHMANN ET AL. [2014]

camera is typically used in a standard heritage documentation project due to its relative stability and possibility to choose acquisition parameters [REMONDINO ET AL., 2012; BEDFORD, 2017]. However, other lower-end sensors also have seen a rise in photogrammetric quality and thus usability in recording heritage objects, e.g. smart-phone cameras, or spherical panoramic cameras.

The general photogrammetric pipeline is already well established, and is summarised in Figure 1.1. The traditional pipeline generally begins with interior orientation and camera calibration, followed by relative and absolute orientations. The block bundle adjustment method enables the computation of the point projection from the image space to the object space in one single operation (more details will be described in Chapter 2), otherwise known as the resolution of the external orientation problem. The external orientation concerns the determination of the relation between the image space and the object space. This relation is defined by the position and orientation of the camera in the object space coordinate system [SCHENK, 2005]. Mathematically speaking this is described by six parameters, including three rotations and three translations.

One of the classic equations in photogrammetry is the so-called collinearity equations [GRUSSENMEYER AND AL KHALIL, 2002; LUHMANN ET AL., 2014]:

$$\begin{aligned} x' &= x_0 - f \cdot \frac{r_{11}(X - X_s) + r_{12}(Y - Y_s) + r_{13}(Z - Z_s)}{r_{31}(X - X_s) + r_{32}(Y - Y_s) + r_{33}(Z - Z_s)} \\ y' &= y_0 - f \cdot \frac{r_{21}(X - X_s) + r_{22}(Y - Y_s) + r_{23}(Z - Z_s)}{r_{31}(X - X_s) + r_{32}(Y - Y_s) + r_{33}(Z - Z_s)} \end{aligned} \quad (1.1)$$

Where :

$f$  : focal length

$x_0, y_0$  : coordinates of the principal point

$x', y'$  : coordinates in the image space

$r$  : rotation matrix

$X, Y, Z$  : coordinates in the object space

$X_s, Y_s, Z_s$  : coordinates of camera projection centre

Equation 1.1 plays a very important role in the bundle adjustment process as they describe the basic relation between the image and object space. It may also be extended to include the image distortion parameters, thus enabling self-calibration to be computed at the same time as the block bundle adjustment [MOE ET AL., 2010]:

$$\begin{aligned} x' &= x_0 - f \cdot \frac{r_{11}(X - X_s) + r_{12}(Y - Y_s) + r_{13}(Z - Z_s)}{r_{31}(X - X_s) + r_{32}(Y - Y_s) + r_{33}(Z - Z_s)} + \delta x_r + \delta x_t \\ y' &= y_0 - f \cdot \frac{r_{21}(X - X_s) + r_{22}(Y - Y_s) + r_{23}(Z - Z_s)}{r_{31}(X - X_s) + r_{32}(Y - Y_s) + r_{33}(Z - Z_s)} + \delta y_r + \delta y_t \end{aligned} \quad (1.2)$$

With:

$$\begin{aligned} \delta x_r &= \bar{x}(K_1 r^2 + K_2 r^4 + K_3 r^6) \\ \delta y_r &= \bar{y}(K_1 r^2 + K_2 r^4 + K_3 r^6) \end{aligned} \quad (1.3)$$

$$\begin{aligned} \delta x_t &= P_1(2\bar{x}^2 + r^2) + 2P_2\bar{x}\bar{y} \\ \delta y_t &= 2P_1\bar{x}\bar{y} + P_2(2\bar{y}^2 + r^2) \end{aligned} \quad (1.4)$$

The variables  $K_1$ ,  $K_2$  and  $K_3$  denote the radial distortion parameters, while  $P_1$  et  $P_2$  denote tangential distortion parameters. These variables are crucial in the calibration of the camera, and their influence on the precision of the final photogrammetric result will be investigated further in Chapter 2.

Meanwhile, parallel to developments in traditional photogrammetry, the computer vision domain also developed solutions independently to the image orientation problem. The advent of SfM heralded a new revolution in photogrammetry, as it automatised many of the previously labour-intensive processes such as tie point marking. Precedents in the photogrammetric domain to these types of automations also exist [GRUEN, 1985], but the computer vision domain has undeniably contributed massively in the development of 3D data generation algorithms. Dense matching was another development which revolutionised photogrammetry, as it enabled the image-based generation of dense point clouds where previously sparse points were only possible. This brought photogrammetry back from the "shadows" of laser scanning and presents an alternative for dense point cloud generation.

Most 3D reconstruction software which is based on images has their own algorithm for the generation of a dense point cloud. REMONDINO [2014] has tried to classify the different existing approaches to dense matching. The most basic classification is between the matching of features (i.e. comparison of descriptors) and the matching of grayscale value within a set search window. Once the correspondence is done, a simple mathematical calculation is performed in order to determine the coordinates of the object on the object space. The matching of features is otherwise called feature-based matching (FBM) while the other classification is called area-based matching (ABM). A simplified schematic describing this classification can be seen in Figure 1.2.



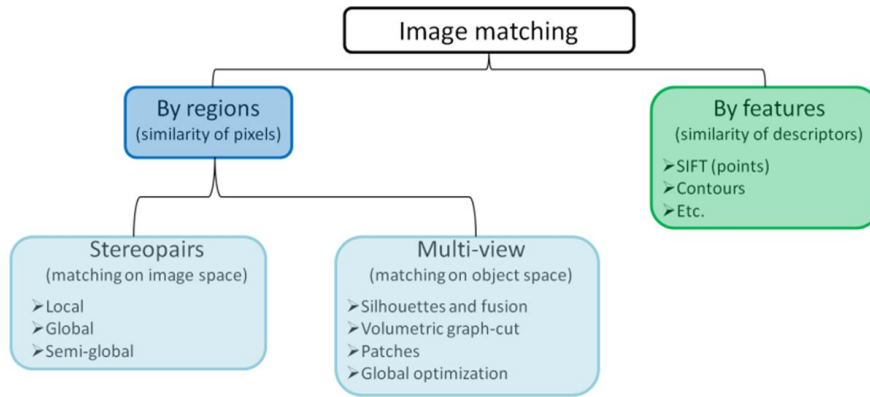


Figure 1.2: General classification of image matching methods, adapted from REMONDINO [2014] and SZELISKI [2010].

Algorithms which follow the ABM classification are very powerful, with the possibility to reach a matching precision of up to  $1/50$  pixels. However, ABM requires objects with textures as well as approximate values. On the other hand, FBM is less dependent on textures, even though the resulting point cloud is often not dense enough. Examples of FBM algorithms include SIFT [LOWE, 2004], SURF [BAY ET AL., 2006], Harris detectors [HARRIS AND STEPHENS, 1988], etc.

Nowadays, FBM is frequently used to generate in the first place a sparse point cloud [SZELISKI, 2010]. This sparse point cloud can be then used as approximate values for the dense matching stage which employs ABM. As regards to ABM, two sub-categories exist which depend on the space where the matching is performed. The first one uses epipolar lines and therefore performs the matching in the image space. SZELISKI [2010] divided this approach into local and global methods. In the local method (based on a search window), the disparity or parallax for each pixel is calculated. The use of a window implies a smoothing on object borders [HIRSCHMÜLLER, 2011]. The global method on the other hand, calculates an energy minimisation on the entirety of the image with an explicit smoothing. Another method called the Semi-Global Matching (SGM) was proposed by HIRSCHMÜLLER [2005], in which an energy minimisation on the length of the epipolar line is calculated, as well as other directions around the pixel.

The other sub-category of ABM performs the matching in the object space. In this approach, a sparse point cloud is necessary as approximate values in the matching process which may be obtained from an FBM matching. When such point cloud is available, the algorithm classifies the images according to their positions relative to the object to be reconstructed. Afterwards the dense matching is performed on a patch around a 3D point of the sparse point cloud seen by a certain group of images. This patch is then enlarged iteratively towards the neighbouring pixels in the image space. A visibility constraint is added to filter the result of this matching [FURUKAWA AND PONCE, 2009].

Photogrammetry for heritage documentation is often employed in its close range configuration, i.e. with images taken from a short distance. This may be done using terrestrial or aerial platforms. The use of DSLR cameras in terrestrial photogrammetry has been much studied, and the general consensus is that as far as non-metric cameras are concerned, DSLR cameras provide the best results [BEDFORD, 2017] and have even been used for metrological and referential purposes [MENNA ET AL., 2018; BÖRLIN ET AL., 2019b]. One disadvantage of the terrestrial technique is that this method has a limited point of

view when dealing with tall structures. This problem is historically solved using various techniques such as the use of cranes [ACHILLE ET AL., 2015; FANGI, 2019] or unpowered platforms such as kites and balloons [VERHOEVEN ET AL., 2012], but nowadays the use of drones is more commonplace [GRENZDÖRFFER ET AL., 2015; HANAN ET AL., 2015; MURTIYOSO AND GRUSSENMEYER, 2017]. When working with drones for close range photogrammetry, a compromise between payload and image quality is often the issue. Higher quality cameras such as a DSLR requires a bigger and heavier (and thus more expensive) drone platform, while lighter and smaller drones can only support lower-quality cameras. In these cases the UAV presents a natural advantage over other terrestrial techniques in its capability to capture aerial images, thus covering angles which would otherwise be impossible to cover from the ground.

While most of the photographic sensors discussed up to this point refer to classical pin-hole projection camera, there exist other types of lenses with their own advantages and disadvantages. The fish-eye lens has the large disadvantage of having a high level of image distortion; however it provides a much larger field of view. In this way, it may be helpful to survey narrow spaces [PERFETTI ET AL., 2018]. Another, extended version, of this fish-eye photogrammetry is the so-called spherical photogrammetry. This technique usually involves the use of equirectangular panoramic 360° images [BARAZZETTI ET AL., 2018; PRAMULYO ET AL., 2017; RAMOS AND PRIETO, 2016]. The main interest point in the use of such sensors is its low-cost nature and ease of data acquisition, in expense of image quality. However, the quality of these sensors and therefore the attainable precision has increased in a significant manner in these last few years [BARAZZETTI ET AL., 2018; MURTIYOSO ET AL., 2019].

## 1.2.2 Laser scanning

Laser scanning has been around since the 1980s and was a revolutionary technology in the domain of 3D mapping. Contrary to traditional total stations, the TLS technology enables the recording of the environment in a fast and relatively accurate manner while being fairly easy to use. This abundance of data, however, may also be a disadvantage as it generates large files of point clouds. This is especially true when the object in question is a complex building, as is often the case in heritage documentation [BARSANTI ET AL., 2014]. Occlusions may also occur, which renders the final point cloud incomplete [LACHAT ET AL., 2016]. The TLS also has limitations when it comes to the point cloud colours and textures, even with the addition of cameras attached to the device [HASSANI, 2015].

The laser scanning technology employs laser light and a type of rotating mirror or prism to measure distances in a regular pattern [GRUSSENMEYER ET AL., 2016], thus enabling a fast and high density measurement of an object. Due to its use of the laser light, it is categorised as an active sensor (contrary to photogrammetry), in that it emits waves and receives their echo. One of its main advantages is therefore the possibility to work in almost any condition, including lack of lighting [RESHETYUK, 2009b]. Furthermore, the use of direct measurements mean that in ideal calibrated conditions the TLS does not need to be scaled like a photogrammetric model. Other additional information can also be recorded by the scanner, for example the intensity value of each point measurement which is useful in applications such as object material identification. When equipped with an RGB camera, the scanner can also register colour values, thus giving a coloured point cloud as the result.

This method can be classified according to several criterions. In terms of the use, it

is usual to classify the technology into aerial laser scanning (ALS) and terrestrial laser scanning (TLS) [VOSSELMAN AND MAAS, 2010]. RESHETYUK [2009b] mentioned that the terminology 'LIDAR' (Light Detection and Ranging) is commonly used, although more often it refers to ALS. The somewhat rarer term 'LADAR' (Laser Detection and Ranging) may also be encountered in some literature. STAIGER [2003] classified the laser scanning technology according to its field of view:

- *Panoramic scanner*, with a field of view that encompasses the entire hemisphere.
- *Camera scanner*, with a limited field of view both horizontally and vertically.
- *Hybrid scanner*, with one rotation axis without restrictions and the other limited.

Another classification is based on the method used to compute the measured distances [GRUSSENMEYER ET AL., 2016; HÉNO AND CHANDELIER, 2014]:

- *Pulse-based scanner*: The basic concept of the pulse based scanner involves the measurement of the time required for the laser pulse to be emitted and returned to the sensor. From this value the distance can therefore be derived, analogous to the process of absolute positioning in GNSS. As with the GNSS technology, the quality of these distance values depend strongly on the quality of the clock embedded within the scanner. This type of scanner is thus better suited for long ranged applications, in which the required precision is lower but the necessity for longer distance is more important (e.g. applications in the mining industry).
- *Phase-based laser scanner*: The phase-based scanner sends a high intensity amplitude-modulated waves and compares the difference of the phase between the emitted and the returned signals analogous to relative positioning in GNSS. This method enables a much higher precision compared to pulse-based methods, but cannot be used for applications requiring very long distances. The reason being that if the ranging distance is farther than the modulating wave length, the risk of phase ambiguity becomes higher. It is thus best suited for middle-range applications (e.g. heritage documentation at a smaller scale). It should also be noted that both the pulse and phase-based types are often referred as time-of-flight scanners.
- *Triangulation-based laser scanner*: This type of scanner relies on solving simple mathematical triangles formed between a single laser emitter, the object, and an optical receiver system. Since using this method the computed distances become less precise as the distance from the scanner to the object increases, it's application is limited to short-range cases. However, in these short-range scenarios, the triangulation-based laser scanner can attain a very high precision.

A quick comparison of some TLS available in the market as of 2020 is presented in Figure 1.3. Meanwhile, registration between TLS stations is equally important in presenting a geometrically correct point cloud. Registration is the term denoting the process of bringing each scan station into the same system, while the term georeferencing is used when this common system constitutes a geodetic coordinate system [RESHETYUK, 2009b]. Georeferencing may be performed either directly or indirectly. Direct georeferencing involves the direct measurement of the coordinates of the centre of the scan stations; this

	Trimble SX10	Z+F Imager 5016	Leica BLK360	Stonex F6	FARO Edge ScanArm HD
	Robotic laser scanner (integration with total station)	Middle-ranged TLS Real-time automatic registration	Short-ranged TLS Weights 1 kg	Hand-held scanner	Triangulation-based scanner for small objects
Type	Pulse-based	Phase-based	Phase-based	Triangulation-based	Triangulation-based
Range	0.9 - 600 m	0.3-365 m	0.6-60m	0.5-4.5 m	<1.2 m
Use	buildings, terrestrial surveys, etc.	buildings, monuments	interiors	statues, carvings, artefacts	artefacts, carvings, inscriptions

Figure 1.3: Examples of TLS available in the market and their characteristics as of February 2020.

may be achieved by placing the scanner on control points which were previously measured using classical surveying means. In this way, the scans are automatically placed in a common geodetic system and therefore theoretically requires no overlap between them.

The second approach is the indirect georeferencing of the scans. In this approach, the scans are registered in a relative system (analogous to photogrammetric relative orientation) before a 3D conformal transformation is applied to place it in a geodetic system (thus analogous to the process of absolute orientation in photogrammetry). In order to perform the registration between the scans, a certain overlap area between the them is required. The registration is often based on a preliminary coarse 3D conformal transformation followed by a fine closest neighbour-based registration [LACHAT ET AL., 2018; FABADO ET AL., 2013]. The coarse transformation can be performed based on common targets which are often artificial. The most common form of artificial targets are planar and spherical [GRUSSENMEYER ET AL., 2016]. 3D conformal (rigid-body) transformation is essentially defined by the following mathematical relation:

$$\begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} = \begin{pmatrix} t_X \\ t_Y \\ t_Z \end{pmatrix} + m \cdot \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \cdot \begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} \quad (1.5)$$

Where:

$X_1, Y_1, Z_1$  : coordinates in the first (or reference) system

$X_2, Y_2, Z_2$  : coordinates in the second (or transformed) system

$t_X, t_Y, t_Z$  : 3D translations

$r$  : rotation matrix containing rotations on the three axes

$m$  : scaling factor

Note that in most cases, the scaling factor may be suppressed since an ideal TLS set-up measures the distances directly without the need for a scaling factor (as opposed to the photogrammetric model). If, however, there is reason to believe that the scaling factor is necessary, it may be computed [VOSSELMAN AND MAAS, 2010].

A finer transformation based on closest point between the individual point clouds can thereafter applied, for example using the Iterative Closest Point (ICP) algorithm [BESL AND MCKAY, 1992]. This iterative process requires good prior alignment between the point cloud, thus the two-step process as mentioned in the previous paragraph. Other types of iterative fine registration algorithms exist in the literature, but within the scope of this thesis they will not be addressed further. Interested readers may refer to VOSSELMAN AND MAAS [2010] for a dedicated section on this issue.

### 1.2.3 Drones

Drones, also known as Unmanned Aerial Vehicle (UAV), is becoming more and more ubiquitous in heritage documentation due to its obvious advantage of an aerial point of view. As the name suggests, an Unmanned Aerial Vehicle comprises a flying platform which is controlled from a ground station using a communication data link [COLOMINA AND MOLINA, 2014]. Although it had its beginnings with military applications, its use in the geomatics domain has continued to increase. Furthermore, the integration of imaging sensors, and recently positioning apparatus such as GNSS receivers, has increased its potential for close-range applications, as it complements classical terrestrial close-range data acquisition [NEX AND REMONDINO, 2014].

Being a versatile system, the UAV has seen many uses in various domains; from photography and video recording to more metric applications. In its role as an aid to photogrammetric work, the UAV has been applied in various fields, such as disaster management [ACHILLE ET AL., 2015; BAIOCCHI ET AL., 2013], 3D building reconstruction [ROCA ET AL., 2013], surveying/mapping [CRAMER, 2013], environmental conservation [BURNS AND DELPARTE, 2017], and heritage documentation [CHIABRANDO ET AL., 2015; MURTIYOSO ET AL., 2016]. While drones are traditionally equipped with imaging sensors (thus the logical use for photogrammetry), recent miniaturisation in aerial LIDAR technology has also enabled its use on UAV platforms.

Several types of UAVs exist [COLOMINA AND MOLINA, 2014; NEX AND REMONDINO, 2014]. The main types available in the surveying and mapping industry usually involves either a fixed-wing or a multi-rotor platform [REMONDINO ET AL., 2011]. A fixed-wing UAV typically has the advantage of a larger mapping zone and lighter weight. However, it is more limited in payload and may be more sensitive to wind conditions [MURTIYOSO ET AL., 2016]. Exceptions to this case include larger sized fixed-wing UAVs, which are often fuelled by combustion engines rather than electric batteries. The multi-rotor type provides a more robust frame with the possibility to take larger payload, but is limited in coverage. The fixed-wing type is therefore suited for mapping large areas following classical aerial photogrammetry, while the multi-rotor one is more suited for close-range applications. Some problems encountered in close-range photogrammetry are specific to this technique, since it involves a more complex geometric configuration of camera stations

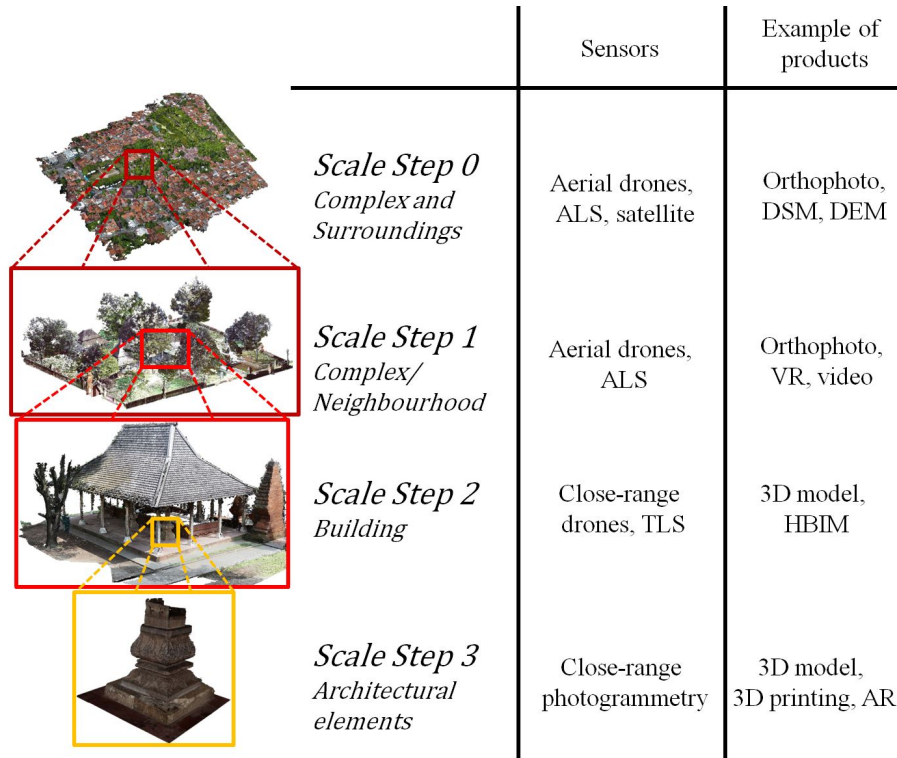


Figure 1.4: The multi-scalar and multi-sensor approach illustrated. Here the Kasepuhan dataset was used for illustration purposes.

when compared to aerial photogrammetry [MURTIYOSO ET AL., 2017a]. For example, close-range photogrammetry can involve an unordered set of camera positions and overlaps with varying GSD values, whereas in aerial photogrammetry these parameters are generally well defined.

### 1.3 Multi-scalar and multi-sensor workflow

Due to the variety of sensors that are today available in the generation of geospatial data, a multi-sensor and multi-scalar approach in data acquisition is a sensible development. This is more so when addressing a historical building within its geographical context; indeed, in many cases the buildings of interest are located within a historical neighbourhood or complex, of which the recording is also interesting for many stakeholders. The immediate need for such type of recording typically concerns a technical issue of project planning. This ideally requires a bird's eye's view of the edifice and its surroundings. This type of small scale recording does not obviously require the same level of resolution as the recording of larger-scale objects such as column capitals; otherwise the size of the dataset would have been immense. In order to address this issue, in some of the main datasets used in this thesis a multi-scalar approach is taken in which several scale levels (or "steps", see also Figure 1.4) were established.

Thus, we propose a multi-sensor workflow in which photogrammetry was mainly used to capture close-range objects with intricate details which requires higher resolution while laser scanning was used to capture the buildings in general. Both aerial photogrammetry and ALS may be employed to get the aerial view of the complex. Close-range drone photogrammetry was used in this thesis, acting as a complement to TLS by providing

an aerial point of view with similar point cloud resolution. This in turn acted as a bridge between the resolution generated by aerial views (ALS or aerial photogrammetry) and that by close range photogrammetry. In terms of the sensors used, the main types of sensors employed in this thesis include the following:

- *Aerial photogrammetry*: Aerial photogrammetry was performed in order to obtain a global view of the sites and their surroundings. We define aerial photogrammetry as a flight mission, often performed using a drone, in which the images were taken from a certain flying height and generally providing a GSD value of between 1 to 5 cm.
- *Aerial LIDAR*: In one instance, notably that of the St-Pierre dataset used in Chapter 5, an aerial LIDAR (ALS) data was used. This is due to two factors: the first being the availability of such data thanks to the open data policy of the Strasbourg city municipality, and the other being the very restrictive legal requirements for a similar drone mission.
- *Close-range photogrammetry*: Close-range photogrammetry was employed in several cases pertaining to the recording of larger-scale objects. We made the distinction between terrestrial and drone close-range photogrammetry. Both are characterised by a close object-to-camera distance; however the terrestrial acquisition was done mostly with a DSLR camera while the drone acquisition used the integrated drone camera.
- *Terrestrial laser scanning*: The TLS method is used in many of the datasets, as they are reliable and easy to acquire. The TLS was used to acquire point cloud data on the scale of buildings, both the exterior and the interior.
- *Topographical surveying*: In parallel with the 3D data acquisition, topographical surveys were conducted to measure several 3D coordinates of some control points. These control points were then integrated into the other projects (photogrammetry and laser scanning) in order to georeference them to the same system. A combination of polygonal traverse network and GNSS measurements were used in this thesis.

The multi-scale aspect of the project refers to the different scale steps that were acquired (see Figure 1.4). Step 0 corresponds to a very general view of the heritage site and its surroundings. This is often useful for technical planning and reconnaissance purposes, for example to plan ahead the placement of surveying and TLS stations. This scale step is not compulsory and may well be replaced by consultation of available satellite images, although depending on the geographical location satellite imagery may not be sufficient. Step 1 involves the point cloud of the heritage complex or neighbourhood and may be acquired using either aerial images or ALS; this scale step is very important in order to understand the context of the historical building in questions.

We then enter into Step 2, in which the building(s) itself is concerned. This scale step naturally requires a higher level of point cloud resolution, hence the use of close-range drones and/or TLS. Finally, scale Step 3 denotes the largest scale, i.e. the most detailed data. This scale step does not need to involve the whole building, and is more often recorded for some remarkable parts or architectural elements such as detailed carvings, column capitals, plinths, etc. Naturally, the average phase-based TLS is not fine enough to obtain these data. Even though it is possible to acquire such resolution using, for

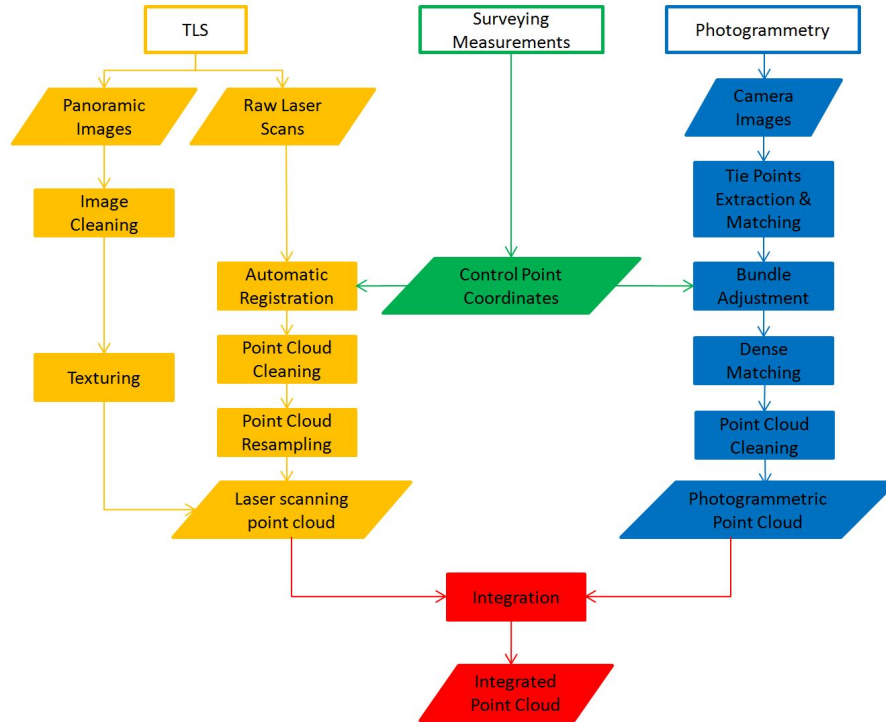


Figure 1.5: Overall workflow for the integration of photogrammetry and laser scanning.

example, triangulation-based TLS, we opted for the simpler (and low-cost) solution of close-range photogrammetry. That being said, this scale step was not the main focus of the thesis, as in many of the developed quality control and point cloud processing experiments only the Step 1 and 2 point clouds were required.

## 1.4 Complementarity and data integration

Due to the advantages and disadvantages of the available recording methods, each can be complementary and their integration is often performed (see for example [MUNUMER AND LERMA \[2015\]](#) and [GRUSSENMEYER ET AL. \[2008\]](#)). Various integration workflows have been proposed, which mainly depends on the particular case. Early attempts include the possibility to integrate the 3D data by means of a Geographic Information System (GIS) [[VON SCHWERIN ET AL., 2013](#)]. Another approach performs independent georeferencing on each dataset in the same coordinate system, with direct integration at the end of each georeferencing [[MURTIYOSO ET AL., 2017c](#); [GRUSSENMEYER ET AL., 2010](#)]. Data integration may be performed on three different levels, namely the raw data level, the feature level, and the decision level [[FARELLA ET AL., 2019](#)]. The raw data level integration is naturally preferred for geometric reasons, as this type means that the data were integrated at the sensor level therefore reducing the possibility of error propagation. However, this is not always a practical solution since it requires a specific hardware set-up. The feature level integration involves extracted features whereas likewise the decision level integration involves decisions of the sensors, such as the use of statistical or fuzzy logic methods [[BASTONERO ET AL., 2014](#)].

The data integration method used in most of the missions during this thesis was described in [MURTIYOSO ET AL. \[2017c\]](#) and can be summarised by the workflow presented



in Figure 1.5. Starting with raw laser scan and photogrammetric data, the final product is a hybrid 3D meshed and textured model of the objects, even though for our purposes the integrated point cloud was often enough as input for further processing down the 3D pipeline. This is the result of the integration of the overall, global, and less detailed 3D model obtained from the laser scanner with the much more detailed photogrammetric 3D models of the points of interest.

The method of 3D data integration was performed as follows: instead of a block integration of all the data which would have taken an immense amount of time and resources, integration was performed by means of independent georeferencing for each dataset. This means that photogrammetric projects were georeferenced using the absolute orientation method, while laser scanning projects used the measured coordinates of some of the artificial spheres to perform a rigid-body transformation. All control coordinates were either measured via total station and GNSS survey or retrieved from the georeferenced TLS point cloud. It should be noted that in cases where control points were obtained from the TLS point cloud, a big downside would be the propagation of error from the TLS georeferencing process. However, this method saved a lot of acquisition time in the field. The coordinates of these points were all linked to the national mapping system of the respective mission sites. The fact that all data are georeferenced to the same absolute system means that future missions can be superposed easily. Another integration method dubbed the "free-network" approach was also tested as an alternative. This comparison will be detailed in Chapter 3.

## 1.5 Challenges and constraints

Various challenges specific to the conditions of the site may be encountered during the recording mission. Several of the commonly found problems will be described in this section. The first and evident problem arises from the architectural styles of the historical buildings. Different types of architecture requires different strategies in the data acquisition. The recording of a European church is quite straightforward as it presents a typical case of indoor scanning. However, tropical architecture, with its lack of walls and prevalence of open pavilions may present different forms and thus requires different strategies. This is not to say that European architecture does not deviate from the habitual scanning process of modern buildings; indeed the presence of niches, statues, or pillars often present serious occlusion problems. Photogrammetry can sometimes be used to help connect the different scans when required.

The climate of the site also poses certain challenges with regards to the acquisition mission. High temperatures and especially high sun exposure may pose problems for some surveying tools which rely on infrared and/or lasers. In this regard, the use of the spatial intersection method for the computation of control points may be useful, since it is based only on angular measurements which are less influenced by the temperature. Solar intensity was also an important influencing factor, as it renders some passive sensors problematic due to overexposure. Problems with overexposure in photos may be rectified in manual mode; however the panoramic images taken by the TLS were more difficult to compensate. On some processing steps requiring point selection, intensity rather than RGB images were therefore used (see illustration in Figure 1.6).

Dense vegetation may also be a challenge, particularly for drone acquisition. Ground Control Points (GCPs) are ideally placed around the site following standard photogrammetric convention, but this could be hindered by the amount of tree canopy. Premark



Figure 1.6: Example of a section of a panoramic image captured by the TLS. The left figure shows the RGB image, with strong overexposure due to sunray intensity. The right figure shows the intensity image, which was preferred in the tie point identification process.

placement for GCPs could also be very limited to small open areas, which do not necessarily correspond to the ideal photogrammetric ground control network. It also posed problem for GNSS measurements, requiring thus a longer measurement time in order to get to the required precision. In regards to TLS and close range terrestrial photogrammetry, the dense vegetation also generates noises which must then be cleaned from the resulting point cloud.

Another unavoidable problem is linked to human presence, as many of the heritage sites scanned were also open for tourists. In this regard, particular care on the handling of surveying tools and TLS spheres were therefore of the utmost importance. Members of the team had to stand by on several spots with artificial TLS spheres in order to avoid them being moved by passers-by. Several control points may also be required to be measured on immobile detail points (e.g. roof edges, brick intersections, etc.) to mitigate problems which may arise from moved targets. This adds complexity during the acquisition and noises in the resulting point cloud. While authorisation for interior scanning was in most cases unproblematic, a complete permit to vacate the premises was almost impossible to attain in all case studies. Furthermore, regulations require a special attention as they may differ for each country (see for example the dedicated section on this issue in [MURTIYOSO AND GRUSSENMEYER \[2017\]](#)).

## 1.6 Derived products

As has been previously established, the main result of the 3D recording operation was the 3D point cloud. Following the proposed workflow as described in section 1.3, a final point cloud was acquired for a particular dataset from the combination of several techniques. While this study focuses on the quality control on this process as well as further point cloud processing, several derivative products can also be obtained from this combined point cloud. These products are typically of interest to stakeholders of the project, as they help in the dissemination of the information regarding the historical site and in some cases helped the management of the said place. Some of the derivative products generated from the tests conducted during this thesis is as follows (see also Figure 1.7):

- *Orthophoto*: The orthophoto of the object, be it aerial or frontal, can be created from photogrammetry. In the case of aerial orthophotos, they are useful as a preliminary overview of the heritage complex and may thus serve as a planning tool for further missions. Naturally the quality of orthophotos depend on the processing method; a true orthophoto may be obtained if required but it would require further processing



Figure 1.7: A photo collage of the possible derivative products from 3D heritage recording. The examples shown here are taken from the various tests performed during the thesis.

in a dedicated photogrammetric software. A frontal orthophoto (i.e. for a façade) is also a useful tool for architects and conservators from which they may derive CAD drawings.

- *Photo-realistic 3D model*: Starting from the registered and combined TLS and photogrammetry point cloud, a 3D mesh may be generated (usually using the Poisson method). Photogrammetric textures can then be applied to the 3D mesh to create a photo-realistic 3D model. For the TLS point cloud, texturing is important since textures from the inherent panoramic images is not good enough, and is often hindered by occlusions from the station's point of view. This 3D model may thereafter be printed (see below).
- *HBIM model*: In terms of site management, the 3D mesh is not sufficient as it could not store semantic information. One way to address this problem is the creation of an HBIM which permits the storing of semantic data within the 3D model. The HBIM model can then be annotated with semantic information later on.
- *Virtual visits and virtual reality (VR)*: In order to create an immersive medium for visualising the results, a virtual visit environment may also be developed. Both the point cloud and photorealistic 3D meshed models can be used in this regard. A combination of texturing methods may be opted, with the Step 2 objects (building exterior and/or interior) displayed in photorealistic textures, while the less interesting parts of Step 1 (e.g. lamp post, ground, trees) only used interpolated colour originating from the point cloud colour. This is done in order to reduce processing time and focus more on the objects of interest. These derivative products are very useful in helping the dissemination of information about the site to the public. The immersive environment enables people to visit the site remotely and appreciate more the historical site. It is also an interesting tool for the architects in order to examine various details of the area without having to go to the site.
- *Augmented reality*: During the period of the thesis, we also experimented with augmented reality solutions, mainly using the `AR.js` library in JavaScript<sup>1</sup>. The developed AR solution is lightweight and has proven to be very useful for education and knowledge dissemination purposes.
- *3D printing*: Lastly, the created 3D models can be printed in 3D, thus creating mock-up architectural models. This application of 3D data severs the illusion of virtuality of the point cloud, and renders it in a physical form. Again, this type of derivative product is very useful and interesting for conservators, especially when dealing with damaged heritage object.

## 1.7 Available datasets

After describing the workflow we proposed for the 3D recording of geospatial data for heritage buildings, this section will list the available main datasets to be used in further experiments regarding quality control and point cloud processing. Three datasets were acquired during the duration of the thesis using the aforementioned acquisition workflow

---

<sup>1</sup><https://github.com/jeromeetienne/AR.js>, retrieved 14 February 2020.

## MAIN DATASETS

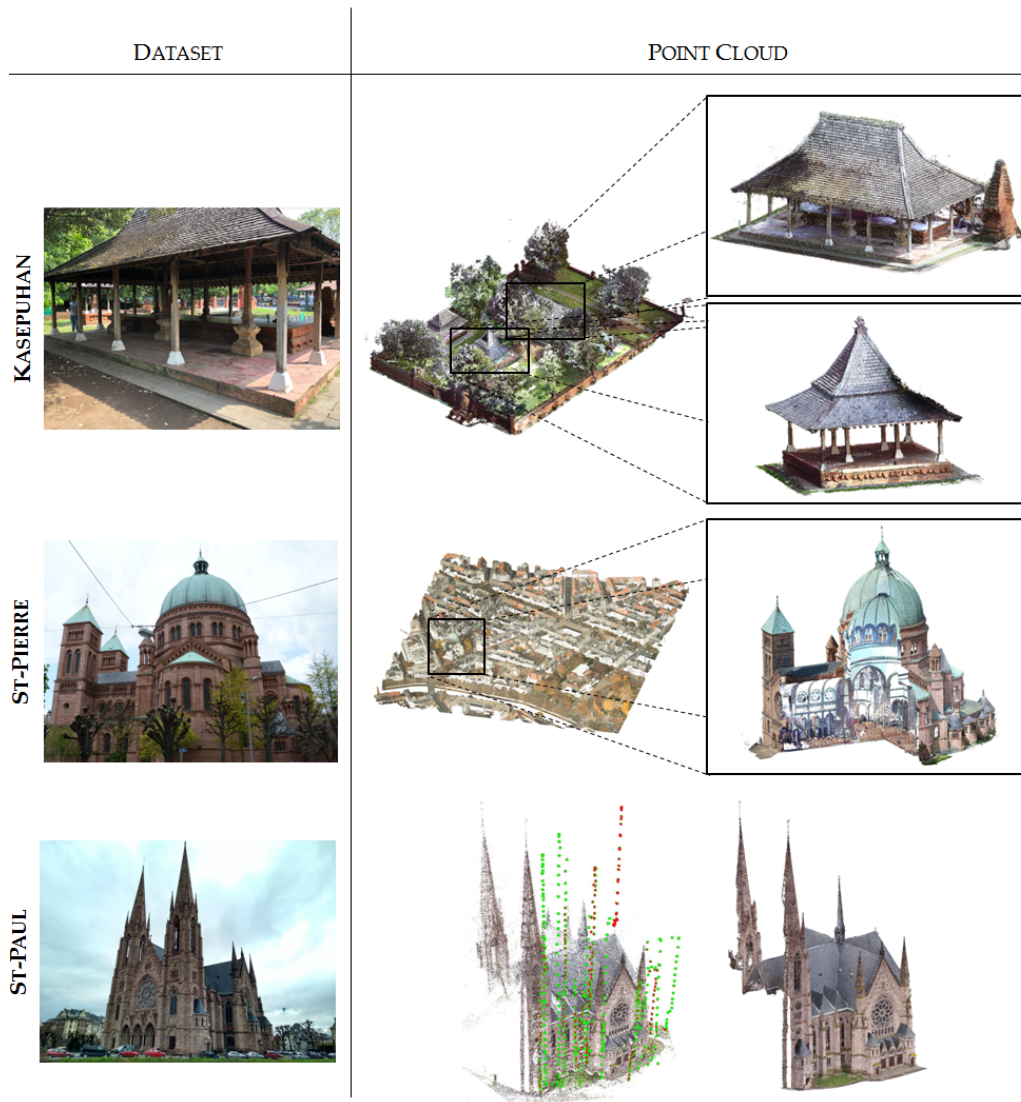


Figure 1.8: The three main datasets acquired during the period of the thesis. Note that the St-Paul data are incomplete; hence its sole use for photogrammetric quality control in Chapter 2.

(Figure 1.8). These three datasets were used during the experiments, while some additional datasets obtained from other parties were also added for the point cloud processing part of the thesis (Part II). These supporting datasets will be described in Chapter 4. The main datasets used in this research are as follows:

- *Kasepuhan Royal Palace, Cirebon, Indonesia ("Kasepuhan")*: This historic area dated to the 13<sup>th</sup> century and includes several historical buildings within its 1,200 m<sup>2</sup> brick-walled perimeters. A particular area of the dataset called Siti Inggil is of particular interest to the conservators as they represent the earliest architectural style in the palace compound. In this paper, the Siti Inggil area is used as a focal point, with two of its pavilions (the Royal Pavilion and the Musicians' Pavilion) used as a case study for the more detailed scale level point cloud processing. Heavy vegetation was present within Siti Inggil, often overlapping with the buildings. The site was digitised in May 2018 using a combination of TLS and photogrammetry (both terrestrial and drone) and georeferenced to the national projection system.

- *St-Pierre-le-Jeune Catholic Church, Strasbourg, France ("St-Pierre")*: The St-Pierre-le-Jeune Catholic Church was built between 1889 and 1893 in Strasbourg during the German era. The church is located in a UNESCO-listed district, the Neustadt, which comprises some other historical buildings of interest such as the Palais du Rhin, formerly the Imperial palace during the German Reichsland era between 1871 and 1918. It is an example of neo-Romanesque architecture crowned by a 50 m high and 19 m wide dome. The neighbourhood around the church was used as a case study in the research along with its interior. The church's surroundings was scanned by aerial LIDAR in 2016 by the city's geomatics service; the point cloud data have since been published as open data<sup>2</sup>. The exterior of the church was also recorded using drones in May 2016 to get a larger-scale and thus more detailed data, while the interior was scanned using a TLS in April 2017.
- *St-Paul Church, Strasbourg, France ("St-Paul")*: This church is an example of neo-Gothic architecture, also built during the German Reichsland era of Strasbourg between 1892 and 1897. Similar to the St-Pierre, this church is located in the Neustadt district, and is thus a part of the UNESCO World Heritage since 2017. The church was recorded solely using close-range drone photogrammetry as a pilot project during the early period of the thesis, and was therefore mostly used in Chapter 2 to perform quality control on the photogrammetric project. Unfortunately no 3D recording of the church's interior was performed.

## 1.8 Summary

This chapter described the rationale behind the use of geospatial data in heritage recording, as well as stressing on its importance in the face of threats. The following points have been discussed in this chapter:

- **3D recording techniques**: A short overview on the theoretical base behind the two most used techniques in 3D heritage recording, photogrammetry and laser scanning, was thereafter presented. These theoretical notions serve as a starting point in developing the 3D pipeline as described afterwards while arguing for a multi-sensor and multi-scalar approach.
- **Need for a multi-scalar and multi-sensor approach**: In the approach proposed by this chapter, the varying components of 3D data were linked together using measured 3D coordinates to create a georeferenced hybrid 3D point cloud. The multi-scale aspect divides the project into four scale steps, going from Step 0 (the heritage complex and its surroundings), Step 1 (the heritage complex *per se* or the neighbourhood where the historical building of interest is located), Step 2 (the historical building of interest), up to Step 3 (architectural and decorative elements). This division in multiple scale levels is important in order to systematise the 3D recording process, in which several different resolution requirements must be present. On the other hand, the multi-scalar aspect enabled the thorough recording of any historical building within the context of its geographical and historical environments. In this way, the proposed workflow presented a way to manage the 3D data for use in further applications.

---

<sup>2</sup>[https://data.strasbourg.eu/explore/dataset/odata3d\\_lidar](https://data.strasbourg.eu/explore/dataset/odata3d_lidar), retrieved 24 January 2020

- **Adaptation in face of challenges and issues:** Furthermore, some challenges during the acquisition have been discussed in this chapter. Some adaptive measures to answer these challenges were also briefly mentioned and proposed. Indeed, the present challenges range not only from technical constraints but also well into other domains such as legal and social issues. A holistic project management is therefore key in identifying these constraints beforehand and proposing feasible solutions before the field work even began.
- **Usability of derived 3D products:** A special section was dedicated to list the products that may be derived from the resulting 3D point cloud. Indeed, these products does not represent the core subject of the thesis itself, but nevertheless offers a panoramic view on the possibilities of the use of 3D data in the field of heritage. This goes beyond technical and geometric issues and well into the field of knowledge transfer and communication, which we may argue is no less important than matters pertaining to precision and accuracy. Indeed, the whole idea of the use of geospatial data for heritage recording is to bridge the gap between the social science (i.e. archaeology, sociology, anthropology, communication science, etc.) and the technologies (i.e. photogrammetry, laser scanning, drones, etc.). Using the products described in section 1.7, architects and conservators may boost their conservation efforts, while policy makers may use them to decide on necessary measures to preserve any particular site in question. The 3D data may also be used for other, more sophisticated purposes. For example, the creation of a HBIM will be very useful in the management of heritage sites. Virtual reality solutions also enable a democratisation of 3D technology and the diffusion of historical information to the public. Finally, with the existing 3D documentation acting as digital archives, a physical reconstruction in the case of damage can be performed via, for example, the 3D printing technology.

Having established the acquisition workflow, the next chapters in this Part I will address the geometric aspect of these 3D data. This will include the photogrammetric quality control (Chapter 2) and 3D data integration (Chapter 3).

# Chapter 2

## Quality Control in Photogrammetry

### Contents

---

<b>2.1 Importance of quality control in photogrammetry</b>	<b>30</b>
<b>2.2 Related work</b>	<b>31</b>
2.2.1 Bundle adjustment	31
2.2.2 Available software solutions	33
<b>2.3 Design of the experiments and results</b>	<b>36</b>
2.3.1 St-Pierre dataset	36
2.3.2 St-Paul dataset	41
<b>2.4 Summary</b>	<b>48</b>

---

This chapter will describe the experiments conducted to try to scrutinise the photogrammetric results obtained from the acquisition missions. Indeed, the use of "black-box" commercial software is not only practical, but had also shown good results in preliminary testing. However, one of the biggest disadvantages of such software is the opaqueness of their algorithm. Even though this is understandable from the economic point of view, for research purposes this greatly hindered a proper analysis of the project. In this chapter, quality control will be performed on photogrammetric projects. Photogrammetry is emphasised and given one dedicated chapter because it involves more processing steps than laser scanning, thus increasing the possibility of error and its propagation through the stages. Laser scanning quality control will nevertheless be explained in Chapter 3, although the analysis will be simultaneous with that of the data integration process. In this chapter, results from our experiments on the results by the software PhotoScan will be discussed. PhotoScan, a photogrammetric software developed by the company Agisoft, is used extensively in the heritage domain [BURNS AND DELPARTE, 2017; BEDFORD, 2017]. The reader should however note that since 2018 PhotoScan has been rebranded as Metashape. In colloquial use, the names PhotoScan and Metashape could therefore be interchangeable. The study presented in this chapter was mostly conducted before this official name change, thus the name "PhotoScan" will be prevalent in this chapter. Moreover, the majority of the content presented here has been published in the papers MURTIYOSO ET AL. [2017a] and MURTIYOSO ET AL. [2018a].



## 2.1 Importance of quality control in photogrammetry

The advent of UAV technology and developments in the domain of computer vision has largely facilitated the traditional photogrammetric workflow [MURTIYOSO AND GRUSSENMEYER, 2017]. Several traditionally manual photogrammetric tasks such as tie point marking and orthophoto production have been automated to a degree where a simple button may be sufficient to perform these tasks [REMONDINO, 2011]. This is a very important advantage for users, more so for those who are not in the photogrammetry community. Indeed, these automations have given an impetus to the democratisation of photogrammetry, where almost anyone can take digital pictures and turn them into 3D models. This has also influenced the mapping industry, since it provides potential users with fast, easy, and fairly low-cost solution to perform mapping activities [CHIABRANDO ET AL., 2015]. However, this advantage is also a double-edged sword precisely due to its simplicity. The black-box nature of some of the SfM based commercial software has hindered a proper statistical analysis of its results [JAUD ET AL., 2016].

Image processing algorithms have also seen a significant improvement in the past two decades. Development of computer vision-derived algorithms has largely facilitated the photogrammetric workflow, enabling the automation of much of the previously manual work. This includes the development of image feature matching [LOWE, 2004] that enables the automatic extraction and matching of tie points. Similarly, improvements in dense matching algorithms [HIRSCHMÜLLER, 2005; FURUKAWA AND PONCE, 2009] have rendered the photogrammetric process very powerful, enabling it to produce dense point cloud up to one point per pixel [ACHILLE ET AL., 2015]. In photogrammetry, UAVs have provided the means to capture images at close ranges but from aerial points of view. This has enabled the 3D mapping of difficult or inaccessible parts [MURTIYOSO ET AL., 2016].

One main and important aspect of the photogrammetric workflow is the external orientation or camera pose estimation step, in which the positions and rotational attitudes of each of the camera stations are determined. This step influences the quality of the rest of the workflow, such as the dense image matching. Exterior orientation is often resolved using a bundle adjustment computation, with initial values calculated from other methods such as the Direct Linear Transformation (DLT), consecutive relative orientation, spatial resection, etc. [LUHMANN ET AL., 2014]. The main aim of this chapter is to demonstrate how open source bundle adjustment solutions can be used to provide valuable diagnostics for the output of a commercial software. The commercial software evaluated was Agisoft PhotoScan, which is a popular software for 3D mapping [BURNS AND DELPARTE, 2017]. The open source solutions used to evaluate the PhotoScan results were DBAT (Damped Bundle Adjustment Toolbox) [BÖRLIN AND GRUSSENMEYER, 2016] and Apero [PIERROT-DESEILLIGNY AND CLERY, 2012].

While Apero is a full photogrammetric module representing the whole photogrammetric process (usually coupled with MicMac for its dense matching step), DBAT has the capability to take PhotoScan projects as input. This effectively means that DBAT can be used to reprocess PhotoScan's results, while providing more flexibility and detailed metrics than PhotoScan (e.g. posterior standard deviations, sigma naught, correlations between parameters, etc.) [MURTIYOSO ET AL., 2017a]. These statistics can be used, for example, to determine the quality of images, presence of outlier observations in the bundle adjustment process, or quality of the computed external and internal parameters. Indeed, compared to classical aerial photography, images provided by terrestrial and UAV close range acquisitions present a particular problem absent in traditional aerial photography,

in that the image and the control point configuration is often irregular. It is therefore in the interest of some users to understand the results in a more detailed manner.

## 2.2 Related work

Close range photogrammetry has often been used to acquire 3D data (e.g. shape, position, and size) from images [GRUSSENMEYER AND AL KHALIL, 2002]. The rise in the use of UAVs and rapid developments in imaging technology and image processing have increased the use of close range photogrammetry for mapping purposes. The previously described developments in imaging sensor technology has also greatly made this method an alternative or complement to terrestrial laser scanners [GRENZDÖRFFER ET AL., 2015; GRUSSENMEYER ET AL., 2010; REMONDINO, 2011]. Furthermore, the use of low-cost sensors such as smartphone images has also increased [NOCERINO ET AL., 2017a]. This thus relatively low-cost solution [BARSANTI ET AL., 2014] for mapping and reality-based 3D modelling is often complemented by commercial, easy-to-use photogrammetric and/or SfM software packages. Although some open source software alternatives exist [RUPNIK ET AL., 2017; GONZÁLEZ-AGUILERA ET AL., 2016], commercial software such as Eos System's Photomodeler, Pix4D, and Agisoft PhotoScan remain very popular, especially outside the photogrammetric community due to their simplicity in creating fairly accurate results [GRUSSENMEYER AND AL KHALIL, 2002; REMONDINO, 2014; BURNS AND DELPARTE, 2017]. Commercial solutions typically hide the algorithms and show a simplified interface to the user in order to make it easy to generate the desired result. This is an advantage for many users, especially those who are not used to the classical photogrammetric workflow. At the same time it complicates a transparent and independent check of the result of each stage of the workflow.

Comparisons of bundle adjustment computation between PhotoScan and other open source solutions have been the object of the study of several other research. Open source solutions are often used to perform the comparison, as they provide a more open algorithm which the users may parametrise to match their need. Apero in particular was used for such comparisons in previous studies [JAUD ET AL., 2016; OUÉDRAOGO ET AL., 2014; REMONDINO ET AL., 2012]. Meanwhile JAMES ET AL. [2017b] focused on optimising PhotoScan processing, while in JAMES ET AL. [2017a] the authors performed PhotoScan assessments in terms of computed covariance of both internal and external parameters, which is similar to DBAT's functionality.

### 2.2.1 Bundle adjustment

A crucial step in the photogrammetric workflow is the determination of the positions and attitudes of the camera stations in 3D space. The associated process involves the calculation of the exterior orientation parameters [SCHENK, 2005; GRUSSENMEYER AND AL KHALIL, 2002], called extrinsic parameters in the computer vision domain [GRANSHAW, 2016]. This analytical procedure is sometimes referred to as "aerotriangulation" or "phototriangulation", although traditionally aerotriangulation was the term applied to the densification of ground controls ("bridging") [WOLF ET AL., 2014].

Several approaches exist for performing the exterior orientation, e.g. image resection and relative orientation [GRUSSENMEYER AND AL KHALIL, 2002]. The bundle adjustment technique enables the simultaneous solving of the exterior orientation problem using image coordinates and the collinearity conditions (Equation 1.1). Mathematically speaking, a

bundle adjustment is a non-linear least-squares optimisation problem on the simultaneous estimation of the 3D point coordinates of the points on the image and the external camera parameters, potentially including the internal parameters [REMONDINO ET AL., 2012] as shown in Equation 1.2. In the latter case, the bundle adjustment process is sometimes referred to as "self-calibration" or "auto-calibration". The bundle adjustment process may involve control points as well as embedded GNSS/IMU data. Any observation can furthermore be weighted according to their precision [GRANSHAW, 2016]. This enables a rigorous solution to the exterior orientation problem.

In general, two approaches to perform bundle adjustment may be followed [REMONDINO ET AL., 2012; GRANSHAW, 1980]:

- *Free-network bundle adjustment*: The free-network approach involves a calculation of the exterior parameters in an arbitrary coordinate system, followed by a 3D similarity transformation to align the network to the coordinate system of the control point ("the real world system"). In the classical aerial photogrammetry, this approach echoes the relative orientation (free network orientation) and the absolute orientation (similarity transformation) steps.
- *Block bundle adjustment*: The block bundle approach involves a simultaneous least-squares estimation of the 3D point coordinates, the external camera parameters and optionally the internal camera parameters, in the coordinate system of the control points. This is done by introducing at least three control points and integrating them within the computation matrix. Appropriate weights can be applied to these observations.

The free-network approach is less rigorous than the block bundle approach, and deformations on the model might occur due to the lack of external constraint and/or imperfect calibration of the internal camera parameters [REMONDINO ET AL., 2012; BEDFORD, 2017]. However, the free-network method is faster since it involves fewer observations. It is therefore often performed as an initial step using a minimal amount of control points. The resulting arbitrary orientation may then be used to guide other control point observations. At the end of this process, fine-tuning using the block bundle adjustment approach is performed. The `Apero` command `Tapas` and the basic workflow of `PhotoScan` use the free-network approach. The 3D similarity transformation in `Apero` and `PhotoScan` is managed by the `GCPBascule` command and the 3D markers, respectively. On the contrary, in the block bundle approach the control points are taken into account directly in the computation. A model deformation is therefore less likely to happen and the network is more stable. The `Campari` command in `Apero` as well as `DBAT` by default employs the block bundle adjustment approach.

The classical least-squares adjustment performed during the bundle adjustment process corresponds to the undamped Gauss-Newton optimisation method [BÖRLIN AND GRUSSENMEYER, 2013a]. In this method, the unknown parameter vector  $X$  (containing the exterior orientation parameters as well as object point coordinates) is iteratively updated as  $X + \delta_X$ , with  $\delta_X$  computed by the minimisation of the following normal equation:

$$(J^T P J) \delta_X = (J^T P X^0) \quad (2.1)$$

Where :

$J$  : Jacobian matrix consisting of first order derivatives of Equation 1.1

$P$  : weight matrix

$\delta_X$  : correction to the unknown parameter matrix  $X$

$X^0$  : approximate initial values to  $X$

However, several factors such as the absence of appropriate initial values or the low intersection angles between the images may cause the iterative solution to not reach convergence [BÖRLIN AND GRUSSENMEYER, 2013b]. Several damping methods may be applied in order to optimise the results [BÖRLIN AND GRUSSENMEYER, 2013a], such as the implementation of the Levenberg-Marquardt algorithm [RUPNIK ET AL., 2017]:

$$[J^T P J + \lambda \text{diag}(J^T P J)] \delta_X = (J^T P X^0) \quad (2.2)$$

In Equation 2.2, the variable  $\lambda$  is the so-called damping factor. This variable enables a dynamic iteration which will adjust its steps accordingly depending on the presented initial values. Using this method, a faster computation can be achieved during the bundle adjustment process, although this does not mean that the system is immune against local minima. The presence of good initial values therefore remains important.

## 2.2.2 Available software solutions

In parallel with improvements in imaging sensors, a significant development in image processing algorithms has led to the inception of various photogrammetric software programmes. In general modern SfM/photogrammetry-based solutions are available of both commercial and open source nature. In terms of commercial software, Agisoft PhotoScan is one of the most commonly used [BURNS AND DELPARTE, 2017; CHIABRANDO ET AL., 2015; VERHOEVEN, 2011]. Other popular solutions include Pix4D, Photomodeler, RealityCapture, and 3DF Zephyr [BEDFORD, 2017; MURTIYOSO ET AL., 2017d]. The openMVG page<sup>1</sup> provides a comprehensive list of the available open source 3D reconstruction libraries.

Some notable examples of open source programmes offering a complete photogrammetric workflow include IGN’s Apero-MicMac package [RUPNIK ET AL., 2017], the ISPRS Scientific Initiative project GRAPHOS [GONZÁLEZ-AGUILERA ET AL., 2016], or VisualSFM. Partial algorithms and software programmes performing a specific part of the general photogrammetric workflow also exists, for example DBAT [BÖRLIN AND GRUSSENMEYER, 2016] that calculates the bundle adjustment step and SURE [WENZEL ET AL., 2013] that generates 3D dense point cloud from pre-oriented images. The programmes listed above perform the computations locally. Other software that perform the computations in the cloud include Autodesk’s Recap, KU Leuven’s Arc3D [BEDFORD, 2017], and the Replicate project [NOCERINO ET AL., 2017a].

In this chapter, a critical analysis of the bundle adjustment results of the commercial software PhotoScan will be discussed. This involves the reprocessing of the PhotoScan project using DBAT, in order to derive detailed metrics. These metrics will then be used to verify the results given by PhotoScan. In addition, an independent bundle adjustment processing using Apero was also performed for the St-Paul dataset. Metrics were generated and used to assess the results and may eventually be used to detect problems in the photogrammetric project.

<sup>1</sup><https://github.com/openMVG>, accessed 17 February 2020

## Agisoft PhotoScan

PhotoScan (or Metashape) is a stand-alone software developed by the company Agisoft LLC. It performs 3D reconstruction of objects from images and employs the whole photogrammetric workflow. PhotoScan has a user-friendly interface with several simplified functionalities and parameters which nevertheless manages to deliver a fairly accurate result. Few things are known about the algorithms employed by PhotoScan. The tie point extraction and detection may employ an improved version of SIFT [CHIABRANDO ET AL., 2015]. Furthermore, the dense matching method used may be a variant of the Semi-Global Matching (SGM) algorithm [REMONDINO ET AL., 2014; HIRSCHMÜLLER, 2005]. In terms of bundle adjustment, little has been published about the approach used by PhotoScan.

As a commercial software, PhotoScan focuses on the results rather than detailed control of the processing parameters. This is advantageous for many users as it simplifies the workflow and renders the 3D reconstruction of objects easier. However, the lack of control may be a drawback in metric applications, as often encountered by photogrammetrists [MURTIYOSO ET AL., 2017a]. In order to render the interface more user-friendly, many processing parameters in PhotoScan are preset with default values. Also, fewer metrics related to the bundle adjustment and dense matching results are given (correlation levels, exterior orientation standard deviation values, etc.). This makes it more difficult to validate the results and detect any existing problems in the project in the case where the resulting precision is unsatisfactory.

However, the authors of PhotoScan seem to have taken notice of this inflexibility for advanced users. As of 17 February 2020, Metashape (formerly PhotoScan) reports standard deviations and correlation coefficients between the interior orientation parameters as well as tie point covariances in its report file. PhotoScan version 1.3.4 was used for the experiments described in this chapter.

## DBAT

The Damped Bundle Adjustment Toolbox (DBAT) is a set of functions developed in the Matlab<sup>©</sup> language for the purpose of calculating bundle adjustment solutions [BÖRLIN AND GRUSSENMEYER, 2013a]. DBAT was originally developed to test different damping methods to the bundle adjustment process. This has been applied, for example, to help camera self-calibration using the extended collinearity equations to converge by only using the EXIF values of the images [BÖRLIN AND GRUSSENMEYER, 2014, 2013b]. Furthermore, DBAT has been tested to reprocess real world datasets, such as large-format aerial images [BÖRLIN AND GRUSSENMEYER, 2016; LUMBAN-GAOL ET AL., 2018] and close range photogrammetry projects, both terrestrial and UAV-based [DALL'ASTA ET AL., 2015; MURTIYOSO ET AL., 2017a]. Its versatility was also used by MENNA ET AL. [2018] to process underwater photogrammetry and by ABATE AND MURTIYOSO [2019] to compute unordered image sets of kite photogrammetry. The possibility to enter into the bundle adjustment details were further reinforced by its modular nature [BÖRLIN ET AL., 2019b], thus making it a very flexible environment for photogrammetric computations.

DBAT provides comprehensive statistics, such as posterior standard deviations of exterior and interior parameters, intersecting angles between images, sigma naught, correlations between the computed parameters, etc. The reported statistics can be used to validate the results, and if necessary to detect errors and redress the project in order to increase its quality [MURTIYOSO ET AL., 2017a]. Originally DBAT was designed to process export files from Photomodeler, but can today load and process PhotoScan (and there-

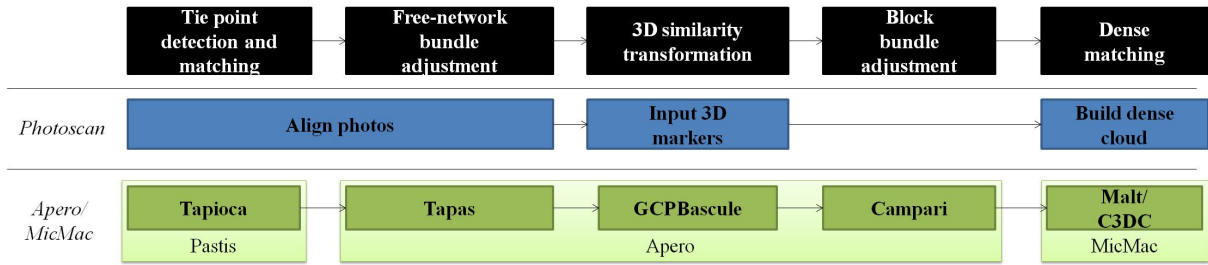


Figure 2.1: Comparison between the commands in Apero and menu options in PhotoScan and their roles in the photogrammetric workflow. For Apero/MicMac, darker green denotes the high level automated commands, while lighter green denotes lower level XML-based commands.

after Metashape) projects as well. The latest development as described in BÖRLIN ET AL. [2019a] indicates that a generalised XML file should theoretically enable input from any type of photogrammetric software; however we have yet to test this new capability.

In this regard, DBAT can be used to reprocess PhotoScan projects and derive more detailed metrics from them. However, DBAT reproduces PhotoScan projects in a broad manner, while generating some bundle adjustment diagnostics. Some inconsistencies are therefore to be expected. For example, PhotoScan seems to relatively scale the tie point image observations by a factor related to the feature scale, which is not reproduced in DBAT. In the discussion of this chapter, the DBAT version used was 0.7.0.1, although the latest version as of 17 February 2020 is the version 0.9.1.3. DBAT is available from GitHub<sup>2</sup>.

## Apero

Apero is part of an open source 3D reconstruction module developed by the French national mapping agency (IGN). Apero performs bundle adjustment with the Levenberg-Marquardt damping method [RUPNIK ET AL., 2017] from the tie points generated and matched by the Pastis module. Pastis is an interface to SIFT++ [PIERROT-DESEILLIGNY AND CLERY, 2012; CHIABRANDO ET AL., 2015] which itself is an improved version of the SIFT feature detection algorithm [LOWE, 2004]. Pastis and Apero are usually coupled with the dense matching module MicMac [PIERROT-DESEILLIGNY AND PAPANODITIS, 2006].

Together, the Pastis-Apero-MicMac family of functions enables the user to perform the complete photogrammetric workflow up to the generation of mesh models and orthophotos. A comparison of the functionalities of PhotoScan and Apero and the corresponding photogrammetric tasks can be seen in Figure 2.1.

The functionality of Pastis-Apero-MicMac suite can be accessed by performing commands on two levels [RUPNIK ET AL., 2017]. The lower level involves the manipulation of XML-files, and the higher level consists of automated commands that may be invoked in a command window environment. In the higher-level command, Tapioca acts as an interface to Pastis, while Apero has several separate commands. Tapas performs a bundle adjustment in the free network mode, which can be followed by 3D similarity transformation by using the command GCPBascule. The resulting initial orientation can be used to guide the user to measure other control points, before performing a simultaneous bundle block adjustment using the command Campari. For dense matching, Malt and C3DC act as interfaces to MicMac.

<sup>2</sup><https://github.com/niclasborlin/dbat>, accessed 17 February 2020

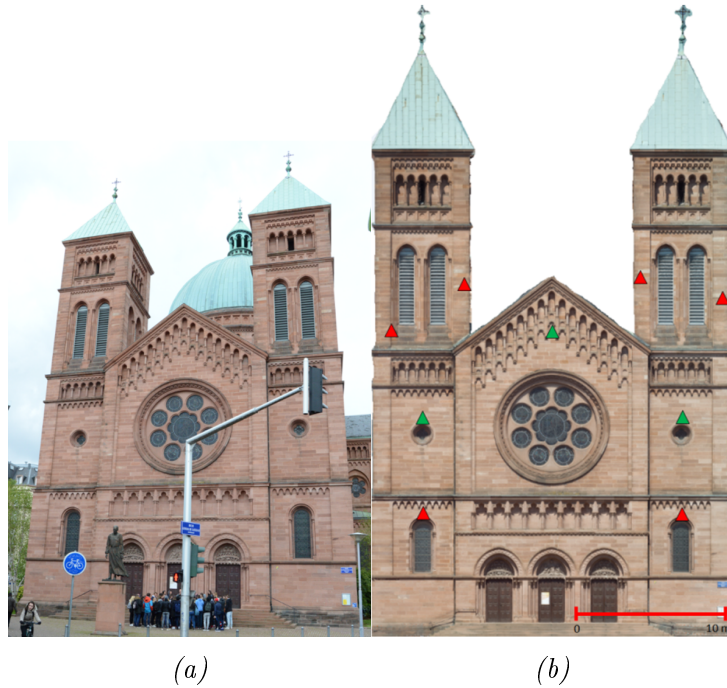


Figure 2.2: (a) One image of the main façade in the St-Pierre dataset. (b) An orthophoto of the façade. Red triangles denote GCPs. Green triangles denote CPs.

In this chapter, only the Aperio part will be used due to the focus on the bundle adjustment step. The tie points were however generated by Pastis (using the `Tapioca` command).

## 2.3 Design of the experiments and results

### 2.3.1 St-Pierre dataset

In this section, DBAT was used to reprocess a UAV project. The results of the self-calibration and control point residuals will then be scrutinised. Potential problems in the dataset were inferred from the available metrics. Note that Aperio is not used in this section, and will rather be discussed in section 2.3.2.

The UAV dataset of the main façade of the St-Pierre-le-Jeune church (Figure 2.2) which has previously been modelled using several software solutions [MURTIYOSO AND GRUSSENMEYER, 2017] was used as a basis for the reprocessing using DBAT. The St-Pierre UAV dataset consisted of 239 images each with a 38 MP resolution. Among these images, 67 were taken from a perpendicular point of view while the rest were oblique images taken with the sensor oriented upwards, downwards, to the left, and to the right. This configuration was used in order to take into account the geometric requirements of a convergent photogrammetric block, as well as to cover difficult parts of the object during the dense matching step (an approach described in MURTIYOSO ET AL. [2017b]). The UAV used for this purpose was the Sensefly Albris, which has the capability to maintain an approximate distance to the object. This enabled the dataset to have a roughly constant camera-to-object distance and therefore constant theoretical GSD. In this case, the theoretical GSD is 1.4 mm for a distance of 8 m.

A total of 9 Ground Control Points (GCPs) were measured on the façade, with an

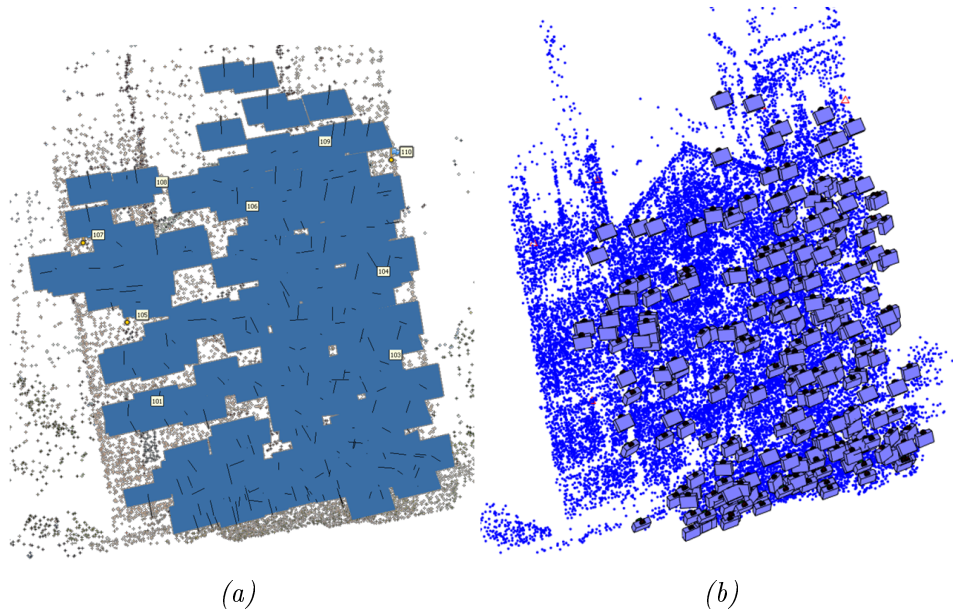


Figure 2.3: Results of the bundle adjustment showing the orientation of the photos in PhotoScan (a) and DBAT (b).

Scenario	Description
S1	Self-calibration with $K_1$ , $K_2$ , $K_3$ , $P_1$ , and $P_2$ using EXIF initial values
S2	Self-calibration with $K_1$ , $K_2$ , and $K_3$ , using EXIF initial values
S3	Self-calibration with $K_1$ , $K_2$ , $P_1$ , and $P_2$ using EXIF initial values
S4	Self-calibration with $K_1$ , $K_2$ , $K_3$ , $P_1$ , and $P_2$ using precalibrated initial values

Table 2.1: Four self-calibration scenarios were tested. In the first three scenarios, the EXIF value for the focal length was used as initial values. A fourth scenario employs the precalibrated initial values.

overall precision of 5 mm. From these 9 GCPs, 3 were selected as Check Points (CPs). The choice of GCPs and CPs follows the convention usually seen in classical aerial photogrammetry [KRAUS AND WALDHÄUSL, 1998]. A polygon network was established around the building which was attached to the French national coordinate system. The photogrammetric control points were thereafter measured from these polygon points. In addition to this field acquisition, the Albris sensor was calibrated beforehand using a set of coded targets that was put in a dedicated room.

The coded-targets were measured using a total station in order to give a rigorous set-up for the calibration. The sensor was then calibrated in PhotoScan. The precalibrated values were used in one of the scenarios tested in this section, while tests using approximate values derived from the images' EXIF file were also performed. An illustration of the result of the bundle adjustment in the two algorithms tested is shown by Figure 2.3. In general, all algorithms managed to reach convergence in their computation and orient all images in all of the proposed scenarios.

Several test scenarios were performed in this research in order to test DBAT's ability to reprocess photogrammetric projects in different conditions (Table 2.1). The main difference between the scenarios lies in the self-calibration parameter configuration. All



	S1				S2				
	PS	$\sigma$ (mm)	DBAT	$\sigma$ (mm)	PS	$\sigma$ (mm)	DBAT	$\sigma$ (mm)	
<b>f (mm)</b>	7.927	N/A	7.921	0.0001	7.927	N/A	7.921	0.0001	
<b>x<sub>0</sub> (mm)</b>	5.057	N/A	5.057	0.0002	5.053	N/A	5.053	0.0001	
<b>y<sub>0</sub> (mm)</b>	3.798	N/A	3.798	0.0002	3.801	N/A	3.801	0.0001	
<b>K<sub>1</sub></b>	3.96E-03	N/A	3.88E-03	2.09E-06	3.96E-03	N/A	3.88E-03	2.09E-06	
<b>K<sub>2</sub></b>	-1.70E-04	N/A	-1.63E-04	1.35E-07	-1.70E-04	N/A	-1.63E-04	1.35E-07	
<b>K<sub>3</sub></b>	2.14E-06	N/A	2.04E-06	2.60E-09	2.14E-06	N/A	2.04E-06	2.61E-09	
<b>P<sub>1</sub></b>	-1.88E-05	N/A	-2.00E-05	1.03E-06	N/A				
<b>P<sub>2</sub></b>	1.62E-05	N/A	-1.74E-05	1.02E-06	N/A				

	S3				S4				PreCal
	PS	$\sigma$ (mm)	DBAT	$\sigma$ (mm)	PS	$\sigma$ (mm)	DBAT	$\sigma$ (mm)	
<b>f (mm)</b>	7.973	N/A	7.966	0.0002	7.927	N/A	7.921	0.0001	7.970
<b>x<sub>0</sub> (mm)</b>	5.053	N/A	5.053	0.0004	5.057	N/A	5.057	0.0002	5.052
<b>y<sub>0</sub> (mm)</b>	3.797	N/A	3.797	0.0004	3.798	N/A	3.798	0.0002	3.787
<b>K<sub>1</sub></b>	2.02E-03	N/A	2.34E-03	1.45E-06	3.96E-03	N/A	3.88E-03	2.09E-06	3.61E-03
<b>K<sub>2</sub></b>	-4.12E-05	N/A	-5.77E-05	4.14E-08	-1.70E-04	N/A	-1.63E-04	1.35E-07	-1.51E-04
<b>K<sub>3</sub></b>	N/A				2.14E-06	N/A	2.04E-06	2.60E-09	1.88E-06
<b>P<sub>1</sub></b>	-5.64E-06	N/A	-5.54E-06	1.55E-06	-1.87E-05	N/A	-2.00E-05	1.03E-06	-2.98E-05
<b>P<sub>2</sub></b>	2.00E-05	N/A	-2.09E-05	1.66E-06	1.56E-05	N/A	-1.74E-05	1.02E-06	-3.93E-05

Table 2.2: The estimated parameters and standard deviations for the four tested scenarios. "PS" denotes PhotoScan while the "PreCal" column presented the precalibrated values generated by laboratory calibration.

scenarios were recreated in DBAT and the results were compared. The quality criteria of interest were chosen to be the RMS values of the GCP errors, and the RMS values of the CP errors. The GCP RMS may be seen as a measure of internal bundle adjustment precision of the respective algorithms, while the CP RMS may give an idea on the accuracy of the solution compared to ground truth data. The precision of the GCP measurements was taken into account during the bundle adjustment in PhotoScan and DBAT as weighting factors. In addition, the *a priori* marking precision for both manual and automatic object points (OPs) in both datasets were fixed at 1 pixel. The choice of this value was done in order to facilitate the comparison between PhotoScan and DBAT.

### Self-calibration

Detailed results of the self-calibration for the St-Pierre dataset can be seen in Table 2.2. As a comparison, a column containing the precalibrated values was also added to Table 2.2 (dubbed the "PreCal" column). In general DBAT had successfully reprocessed the PhotoScan projects in terms of camera calibration values as can be seen in this table. For the focal length, DBAT managed to calculate values with an average difference against PhotoScan of  $6.25 \mu\text{m}$ . As for the principal point offset, DBAT's results were virtually the same as PhotoScan's, within 3 significant numbers. These results show that DBAT managed to closely follow PhotoScan's computations.

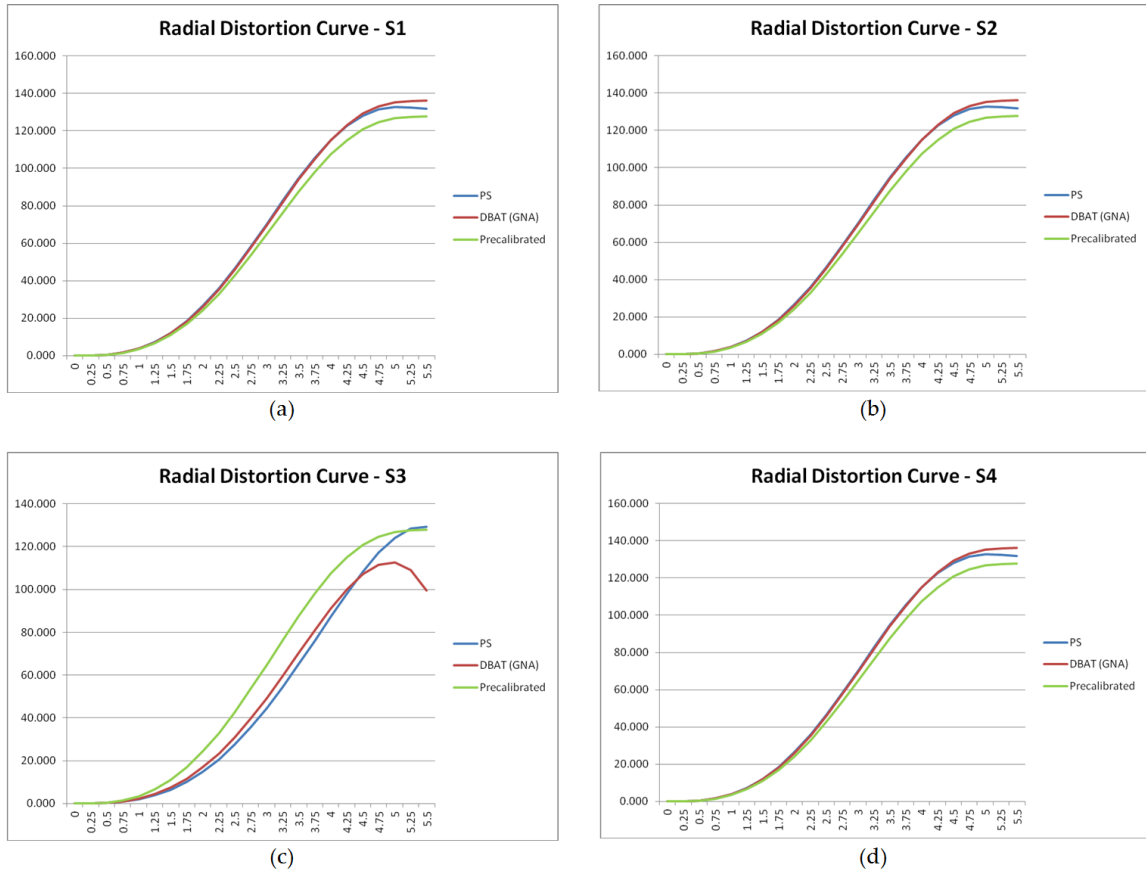


Figure 2.4: The radial distortion curves corresponding to the estimated  $K1$ - $K3$  parameters for  $S1$  (a),  $S2$  (b),  $S3$  (c), and  $S4$  (d). "PS" denotes PhotoScan results.

Differences in terms of the distortion parameters are more difficult to ascertain. To this end, the radial distortion curves were plotted in Figure 2.4. In the cases of  $S1$ ,  $S2$ , and  $S4$ , DBAT managed to generate a similar distortion profile as that of PhotoScan, with small differences beginning at the radial distance of 4.25 mm relative to the projective centre. These minor differences may come from slight errors due to the conversion from PhotoScan to DBAT distortion coefficient format. The DBAT format follows the PhotoModeler convention in presenting distortion parameters as polynomial coefficients scaled by the focal length, while PhotoScan calculated the normalised value of these parameters. Differences with the precalibrated values are to be expected since the conditions during the calibration are not exactly the same as the conditions during the real acquisition. Furthermore, some differences may be expected because PhotoScan most probably performs a free network adjustment followed by a conformal 3D transformation, whereas DBAT includes GCPs directly in its bundle adjustment computation.

It is also interesting to note that both PhotoScan and DBAT arrived at the same calibration values in  $S1$  (with EXIF initial values) and  $S4$  (with precalibrated values as initial values). However, it should be noted that the case of the St-Pierre dataset presents a particular case where oblique photos were also included in the bundle adjustment process; this increases the strength of the acquisition network geometry.

$S3$  presented an interesting observation on its distortion curve. By not calculating  $K_3$  in the self-calibration process, DBAT and PhotoScan's curve diverge almost immediately from the 1 mm radial distance mark. Furthermore, the  $\sigma_0$  value of  $S3$  in DBAT was 2.070 which presented an anomaly compared to the other cases (see also Table 2.2).

	Software	$\sigma_0$	Rep. error RMS (pix)	GCP RMS (mm)	CP RMS (mm)
<b>S1</b>	PhotoScan	N/A	1.490	5.8	7.7
	DBAT	1.142	1.490	5.8	7.5
<b>S2</b>	PhotoScan	N/A	1.492	5.3	7.8
	DBAT	1.144	1.492	5.1	8.1
<b>S3</b>	PhotoScan	N/A	2.700	7.6	8.7
	DBAT	2.070	2.700	7.3	8.5
<b>S4</b>	PhotoScan	N/A	1.490	5.8	7.7
	DBAT	1.142	1.490	5.8	7.5

Table 2.3: Results for the different scenarios, showing the  $\sigma_0$ , reprojection error RMS, GCP error RMS, and CP error RMS.

PhotoScan also gave a reprojection error of 2.700 pixels. Even though the fact that  $K_3$  is not calculated suppressed the correlation (see also the influence of correlations in Table 2.4) between the estimated calibration parameters, this may indicate that for this particular sensor  $K_3$  is nevertheless an important factor.

### GCP and CP verification

Comparison of the GCP and CP RMS for the different scenarios tested in this experiment can be seen in Table 2.3. It should be noted that in this experiment, in order to compare both algorithms, the GCPs were weighted using their precision of 5 mm, while all markings whether automatic or manual were weighted using a uniform marking precision of 1 pixel. Results of the bundle adjustment show that a maximum difference of 0.3 mm for the GCP RMS between DBAT and PhotoScan were observed. The maximum difference of CP RMS was also 0.3 mm, for a theoretical GSD of 1.4 mm.

Several factors may contribute to the final RMS result. The GCPs were distributed evenly on the façade; however the lack of depth variation between the GCPs may contribute to the final RMS. Furthermore, the noise present on the images also generates another source of error. However, the main objective of the experiment is to compare the performance of PhotoScan and DBAT. In this regard, DBAT has managed to reprocess PhotoScan projects under approximately the same conditions and weighting, although a slight difference is always to be expected when dealing with a black-box solution. It may then be used as a tool to verify PhotoScan's results on which a quality check may be performed.

### Quality control

One advantage of DBAT lies in the metrics that it provides the user at the end of the bundle adjustment process. In terms of correlation values, Table 2.4 showed the high correlation values between the different calibration parameters as well as the number of automatic tie points with high correlation values in all the scenarios tested. In the case where  $K_3$  is calculated, the results show a strong correlation between the radial distortion coefficients. The standard deviation values given by DBAT for the calibration parameters

	Calibration parameters	Automatic tie points	
		more than 95%	more than 99%
S1	K1-K2:96.8%; K2-K3:98.6%	26.00%	1.40%
S2	K1-K2:96.8%; K2-K3:98.6%	26.01%	1.40%
S3	-	25.90%	1.38%
S4	K1-K2:96.8%; K2-K3:98.6%	26.00%	1.39%

Table 2.4: Correlations in the processed projects in DBAT. For the automatic points, the value denotes the percentage of automatic tie points with high correlation values.

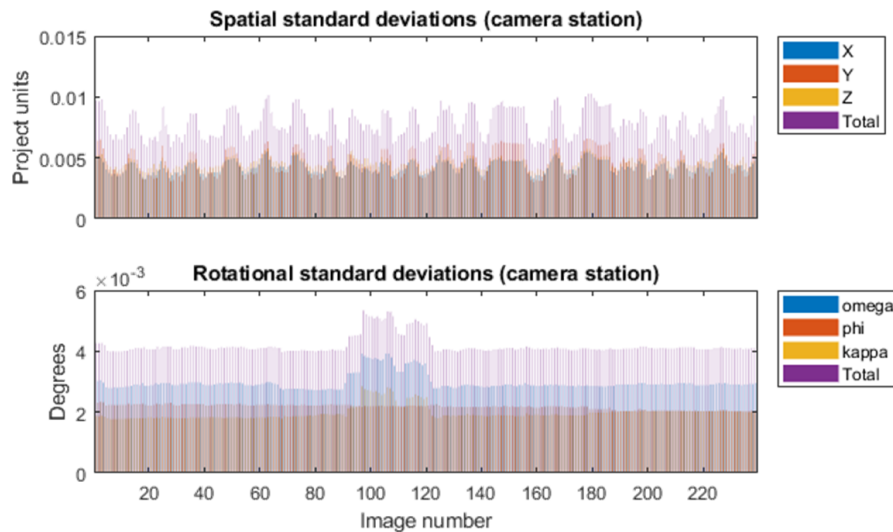


Figure 2.5: Histogram of the exterior orientation standard deviations for S4.

are also useful to assess the quality of the self-calibration process. As regards to the automatic tie points, the St-Pierre dataset shows that over a quarter of its tie points have a strong correlation of more than 95%. Based on this information, a filtering of the automatic tie points for the St-Pierre dataset could be performed in order to increase the quality of the bundle adjustment.

Indeed, by performing this filtering in DBAT, the high correlations disappeared. Another example of metrics which can be derived from DBAT includes the standard deviations for the external orientation parameters. This may be useful in some cases to help users in sorting images which may worsen the results of the bundle adjustment. These metrics are shown in histogram form in Figure 2.5. A slight increase in rotational standard deviation can be seen for the images numbered around 100 and 120. This indicates that the orientation of these images may be subject to some errors; this could serve as a clue to reassess these images and in the worst case suppress them from the project altogether.

### 2.3.2 St-Paul dataset

A second case study on the quality control of photogrammetric projects was conducted on a UAV flight mission over the historic 19th century St-Paul church located in the city of Strasbourg, France (see also the historical information on section 1.7). In terms of building material, the edifice was built using the typical Alsatian red sandstone (the

same main material used for St-Pierre and Strasbourg Cathedral). This provides enough textures for the feature matching step and eventually the dense matching process. The eastern façade of the church was taken for the experiments regarding the quality of its bundle adjustment, by means of comparing the results generated by PhotoScan, DBAT, and Apero.

In this case study, the UAV DJI Phantom 4 Professional was used to acquire the data. The Phantom 4 Professional was released in November 2016, and includes a 20 megapixels on-board still sensor. It is equipped with a 3-axis stabilisation gimbal, and may fly for roughly 30 minutes for each flight. The sensor is a CMOS with a 3  $\mu\text{m}$  pixel size, with an 8.8 mm focal length (24 mm in 35 mm format equivalent). It is a multi-rotor type UAV with four rotors. The Phantom 4 sensor is integrated with the system; the image acquisitions were therefore performed using the on-board sensor, equipped with a standard frame camera lens i.e. not a fish-eye lens. The UAV employs a global shutter as opposed to a rolling shutter.

Before the image acquisition was performed, a topographical survey was conducted around the object of interest using a Trimble S8 robotic total station. Traverse points were measured around the church, from which detail points on the façade were determined. Some of the detail points were used as GCPs, while others were used as CPs. The whole traverse network was attached to the French national cartographic projection system using a GNSS receiver and the RTK (Real Time Kinematic) method. Two traverse points were fixed in planimetric coordinates, while the altitude of one of the two points was fixed. The altitude values of the points were measured by GNSS, as no levelling benchmark was found near the building. Each GCP and CP was measured twice from two stations, in order to enable a spatial intersection computation on their coordinates.

At the end of the topographical survey, all measurements were calculated in a least-squares block adjustment using the software Covadis. This was done in order to determine the standard deviation values for each control point and enable the use of weighting in the bundle adjustment process. The obtained standard deviation values were of the order of 5 mm. In average, the planimetric precision of the control points resulting from this process was 4 mm, while the average altimetric precision was 2 mm. A total of 29 marked points (see Figure 2.6) were thus measured manually. Sixteen points were used as GCP. The remaining thirteen were used as CP. The distribution of GCP followed the classical photogrammetric convention that control points should be placed at the object's perimeters in order to ensure a uniform 3D transformation of the model (again, see [KRAUS AND WALDHÄUSL \[1998\]](#)).

The UAV was flown using a combination of perpendicular and oblique photos in order to cover difficult parts of the building, a method similar to the one used to acquire the main façade of St-Pierre as described in section 2.3.1. The flight strips were performed systematically in a vertical fashion; the oblique photos followed the same flight strip design. A rough flight plan was designed beforehand to ensure enough overlap and sidelap between the images and strips, respectively, and given to the pilot. The final flight configuration resembled the designed one, although some modifications needed to be performed on the field. Furthermore, additional flights were carried out to capture more complex parts of the edifice, such as the central spire. It should be noted that UAV projects are strictly regulated by the French government. However, close range photogrammetry projects such as the one performed in this research fits one of the scenarios set by the regulations, with an 8 kg limit on the UAV weight and a 100 m limit of operation between the pilot and the UAV [[DGAC, 2015](#)].

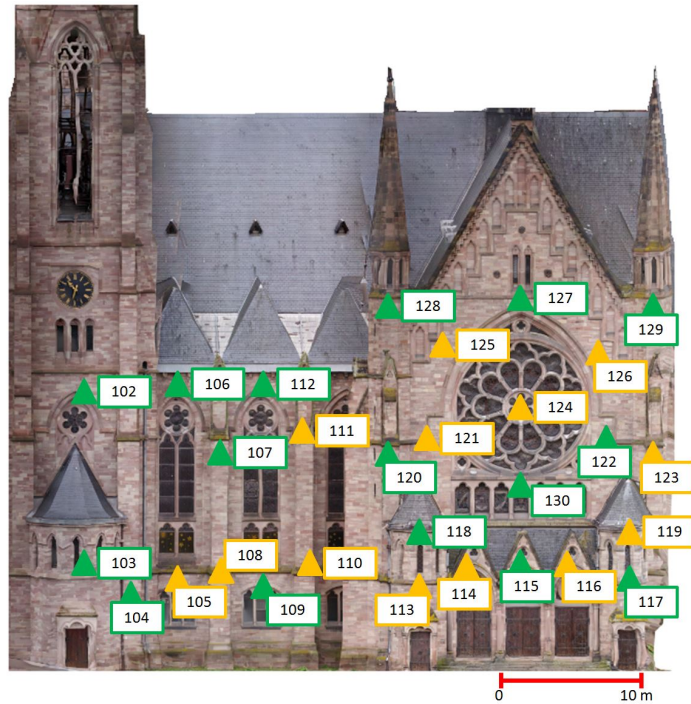


Figure 2.6: Distribution of the ground control points and the check points on the eastern façade of St-Paul's church. Green triangles denote GCPs while orange triangles denote CPs.

The project was also carefully planned to homogenise the lighting conditions and the resulting image texture as much as possible. Thus, flights were performed during the morning or the afternoon to avoid hard sunlight around noon. Cloudy days were preferred for flights over sunny days. In terms of geometric planning, the UAV was flown at an approximate camera-to-object distance of between 5 to 10 meters. This corresponds to a theoretical average GSD of 2 mm.

The images were then processed using the standard PhotoScan pipeline, from the tie point generation, bundle adjustment and self-calibration, up to the creation of 3D models. In this section, we focus on the bundle adjustment results for the 485 images of the eastern façade of the church. The experiments were designed to assess and compare the bundle adjustment results from PhotoScan with DBAT and Apero. DBAT was used to reprocess the bundle adjustment computations of the PhotoScan project. Thus, DBAT computations used the same 2D point measurements and 3D GCP control measurements as PhotoScan to reprocess the project and provide detailed bundle adjustment diagnostics. In contrast, the Apero comparison was performed using only the same images as input. Thus, the Apero computations were based on measurements of different tie points than those generated by PhotoScan.

A graphical representation of the conducted experiments is shown by the Figure 2.7. GCP and CP residuals were compared between PhotoScan and DBAT and then between PhotoScan and Apero. This method of bundle adjustment assessment is often used in the literature [GERKE AND PRZYBILLA, 2016; KIM ET AL., 2013; AI ET AL., 2015]. After the comparisons were performed and the results validated, metrics from the open source algorithms were then used to perform a quality assessment of the project. The objective of the quality assessment was to investigate whether improvements and error detection can be performed on PhotoScan projects by the open source solutions, to potentially enable a more precise photogrammetric end product.

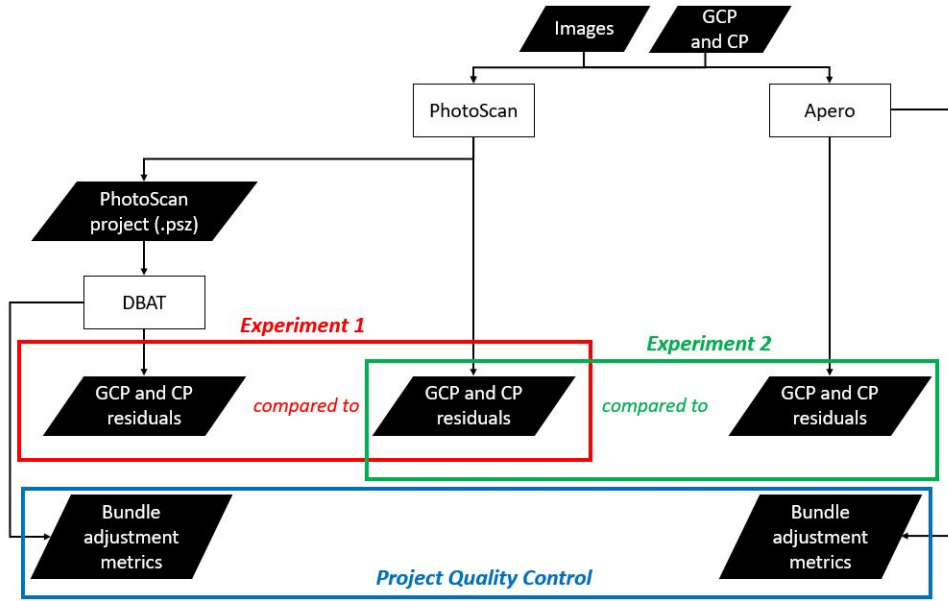


Figure 2.7: Flow chart of the conducted experiments involving PhotoScan, DBAT, and Apero. Two experiments were performed in order to compare PhotoScan to DBAT and Apero respectively, followed by a project quality control based on the bundle adjustment metrics reported by open source solutions.

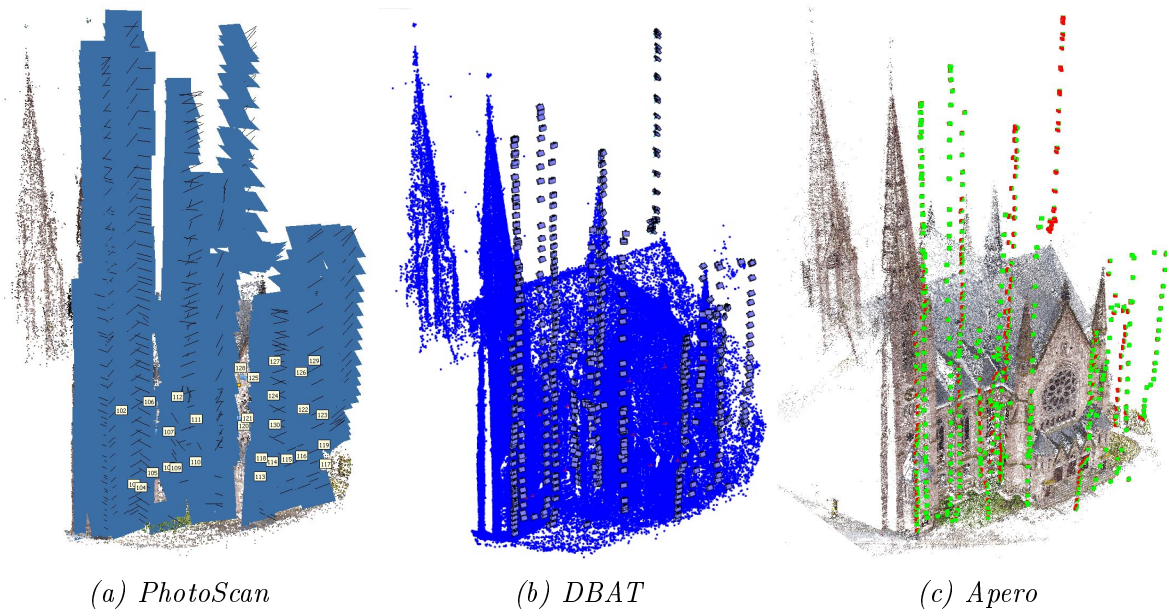


Figure 2.8: Results of the external orientation process shows the positions and attitudes of each camera station, together with the sparse point clouds generated from the respective feature matching process. PhotoScan and DBAT processed the same tie point measurements. Apero generated its own set of tie point measurements from the same input images as PhotoScan.

The GCP standard deviation values obtained from the topographic survey were used as weighting factors. Furthermore, the manual marking error was set to 0.3 pixels, and the automatic tie point error was set to 1 pixel. The weightings of the observations in the bundle adjustment was therefore dictated by the *a priori* values. Identical weighting settings were used for all three algorithms (PhotoScan, DBAT, and Apero) in order to

compare their results under similar computing conditions. All three algorithms succeeded in computing the orientation parameters of all images in the project. Results from PhotoScan were reprocessed in DBAT using the Gauss-Newton-Armijo method, while Apero performed the bundle adjustment based on tie points provided by its own feature matching function, Pastis. A visual representation of the orientation results of the three algorithms can be seen in Figure 2.8.

### Reprocessing of PhotoScan using DBAT

In the first experiment, the PhotoScan project was reprocessed using DBAT. The main objective of this experiment is to investigate whether DBAT can recreate the PhotoScan results. Diagnostics from DBAT can then be used to assess the quality of the project. The evaluation was based on the GCP and CP residual RMS. The GCP and CP residuals are shown in Table 2.5.

Overall, the DBAT residual RMS values for the GCP and CP have differences of 0.0-0.5 mm from those of PhotoScan. These slight differences are virtually negligible, since they correspond to 0.2 pixels when the GSD is taken into account. As the manual marking error was assumed to be 0.3 pixels, an error of this magnitude falls largely within the marking tolerance (hypothesised to be  $1.96\sigma$  for a 95% confidence level). The largest disparity between 3D residuals was for points 123 and 126, with a difference of 3.5 mm between DBAT and PhotoScan. However, the results suggest that DBAT is able to reproduce PhotoScan's results to a reasonable degree and within the expected tolerance, thus validating the results given by PhotoScan.

### Independent Check using Apero

In the second experiment, a similar comparison was performed between PhotoScan and Apero. The number of tie points used by PhotoScan and Apero was on average 4,000 points per image and 15,000 points per image, respectively. The GCP and CP residuals are also shown in Table 2.5. The Apero GCP residuals were higher than that of PhotoScan, amounting to 8.0 mm. This is higher than the theoretical GSD of 2 mm, which roughly represents the lowest geometric value attainable by an observation on an image. However, it is still within the tolerance when considering the 5 mm GCP precision. The Apero CP residual RMS was similar to the ones generated by PhotoScan/DBAT results. The observed difference in residual size may be due to the different processes used to detect and filter tie points. Especially, a less aggressive tie point filtering strategy in Apero might explain the elevated residuals. An alternate hypothesis is that PhotoScan employs a better blunder detection algorithm for the automatic tie points. Nevertheless, the Apero results show that millimetric result can be obtained from this dataset.

### Quality control

Conclusions from the two experiments can thus be used to validate PhotoScan's results. The first experiment (PhotoScan *vs.* DBAT) showed that using the same initial values and under similar weighting conditions, PhotoScan's results can be recreated by DBAT. Meanwhile the second experiment (PhotoScan *vs.* Apero) showed that a similar order of precision can be achieved independently using Apero. In terms of computing time, DBAT took 1 hour and 13 minutes and Apero took 54 minutes to finish the bundle adjustment process, using a computer with a 24-core Intel<sup>(R)</sup> Xeon<sup>(R)</sup> 2.4 GHz processor and 50 GB



GCP 3D residuals (mm)				CP 3D residuals (mm)			
Point	PhotoScan	DBAT	Apero	Point	PhotoScan	DBAT	Apero
102	1.1	2.8	1.5	105	7.3	8.3	7.2
103	4.0	5.2	7.6	108	3.4	2.4	5.2
104	3.5	3.2	5.1	110	9.1	9.2	7.8
106	8.7	8.8	17.3	111	2.1	2.2	1.4
107	4.6	4.6	7.0	113	2.7	3.7	4.3
109	4.1	4.4	5.6	114	6.1	6.6	5.4
112	6.3	4.9	7.6	116	3.2	4.1	9.1
115	2.5	3.3	1.4	119	4.3	3.7	7.7
117	5.5	6.8	12.2	121	3.5	3.6	7.1
118	2.9	3.2	4.5	123	7.0	3.5	7.3
120	5.6	5.8	9.3	124	6.7	5.4	3.7
122	7.1	6.4	8.2	125	11.3	10.8	5.7
127	6.5	5.7	6.0	126	7.3	3.7	7.7
128	4.5	4.5	5.7	<b>RMS</b>	<b>6.3</b>	<b>5.8</b>	<b>6.4</b>
129	5.9	5.1	5.5				
130	6.4	6.5	8.6				
<b>RMS</b>	<b>5.3</b>	<b>5.3</b>	<b>8.0</b>				

Table 2.5: GCP and CP 3D residuals for PhotoScan, DBAT, and Aperio.

of RAM. It should be noted that much of this computing time is used to calculate the covariance matrix for the observations, which means that projects with more observations would typically require longer computing time. In total, 485 images, 16 manual GCPs, and 192,814 automatic tie points were processed. However, the latest version of DBAT has implemented a significant improvement in the computation of the covariance matrix; preliminary unpublished results show an acceleration of a factor of  $10^3$ . Furthermore, the bundle adjustment analysis provided by PhotoScan is limited to the GCP and CP residuals as well as the image residuals of the observations and internal parameter correlations and standard deviations. However, both DBAT and Aperio generate other metrics related to the quality of the result of their respective bundle adjustment processes.

The metrics presented by DBAT include the image coverage, the number of automatic tie points for each image, the overall tie point residuals for each image, as well as the standard deviations for the positions and rotational attitudes of each camera station. The DBAT statistics are presented graphically in Figure 2.9.

<sup>3</sup>These results will be reported in a future paper: BÖRLIN, N.; MURTIYOSO, A.; GRUSSENMEYER, P. 2020. "Efficient computation of posterior covariance in bundle adjustment in DBAT for projects with large number of object points". Proceedings of the XXIV ISPRS Congress, 14-20 June 2020, Nice, France. *In press*.

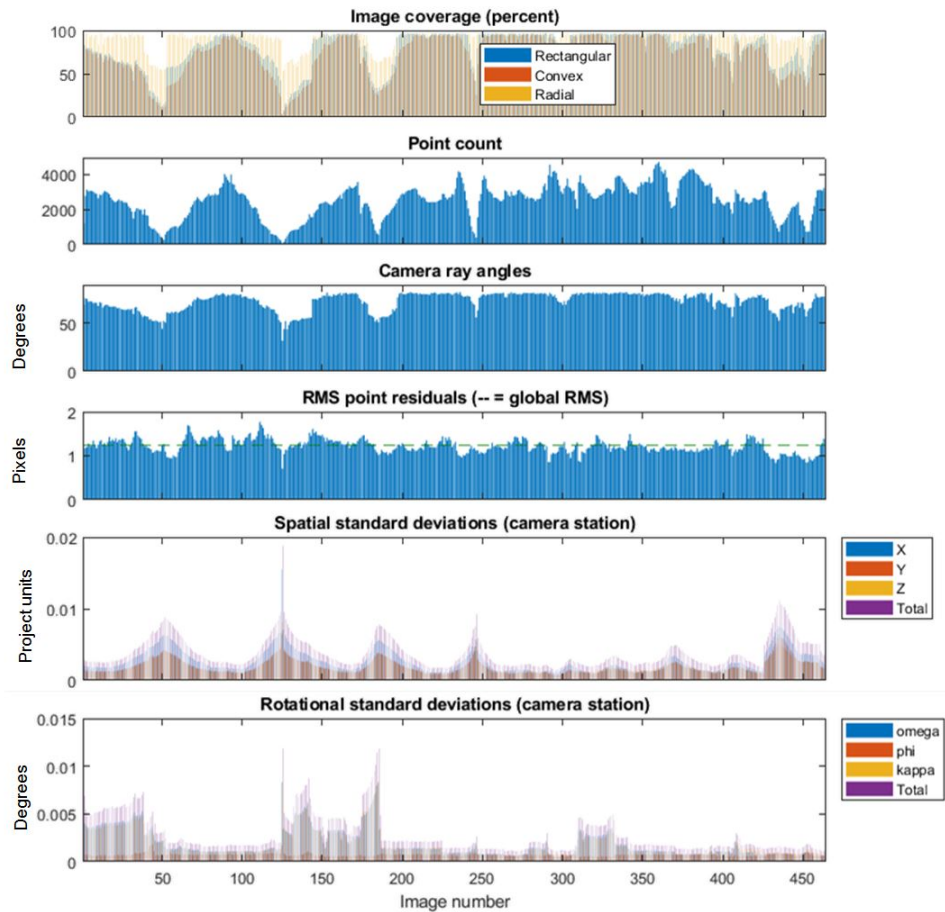


Figure 2.9: Project statistics generated by DBAT related to the quality of the images and their exterior orientation parameters. Large point residuals or high standard deviations of the exterior parameters provide clues to problematic areas.

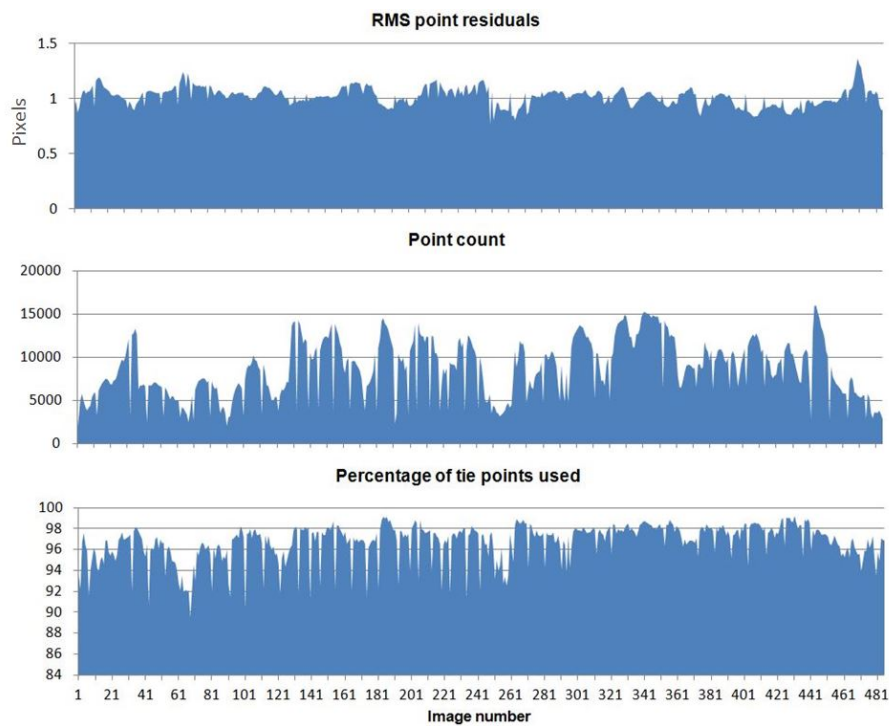


Figure 2.10: Project statistics generated by Aperio related to the quality of tie points used in the bundle adjustment process.

The statistics generated by Apero include the overall tie point residuals, the number of automatic tie points, and the percentage of points deemed useful for the bundle adjustment. A graphical representation of the Apero statistics can be seen in Figure 2.10. An analysis of Figure 2.9 reveals some interesting results. For example, some spatial standard deviation values of the camera stations reached 2 cm. This correlates to low point count for those particular images, sometimes also with a decrease in camera ray angles. In some cases, the rotational standard deviation peaked to more than 0.01 degrees, something that may be due to a lower number of tie points although the connection is less clear. Nevertheless, the correlation between the number of tie points in an image and the precision of its exterior orientation parameters is expected, as fewer observations will generate lower precision due to a lower redundancy. In addition, the image coverage correlates with the number of tie points in the images. Overall, the statistics provide clues as to what parts of the model would benefit from adding more measurements and/or images.

The metrics generated by Apero provide similar information, although with less obvious conclusions of where to improve the model. The Apero plots show that the number of automatic points is much higher than those generated by PhotoScan. It should be noted that in both experiments, no post-bundle adjustment filtering of tie points were carried out. The number of detected points can also be used as an indication of image quality, since images with worse quality, e.g. blurred or not focused, have fewer features available for detection during the image matching step. Indeed, the number of tie points detected by Apero seems to be correlated to the percentage of tie points used in the bundle adjustment process. Figure 2.10 also shows that in general Apero takes more than 90% of the detected tie points for the bundle adjustment, with a small part of the images where less than 90% were used. This can also be an indication on the quality of the images. Finally, in terms of the RMS point residuals, both DBAT and Apero gave an average value of around 1 pixel, which corresponds to the assumed precision of the automatically detected tie points. In terms of posterior manual marking precision, the mean value of 0.28 pixels were obtained. This corresponds also with the expected prior value of 0.3 pixels set in the start of the bundle adjustment process. Even though the exact algorithm used by PhotoScan remains difficult to ascertain, these approaches may help users to better understand the results they receive and eventually improve them. Indeed, with the lack of more metrics, some problems in the dataset cannot be detected, as has been shown in this section.

## 2.4 Summary

This chapter aims to demonstrate how open source solutions can be used to provide valuable survey diagnostics for bundle adjustment results generated by a commercial software programme, in this case PhotoScan. The rationale behind the project quality control using bundle adjustment metrics lies behind the fact that for photogrammetry, this phase serves as the base for further results. Indeed, a bad bundle adjustment result means errors exist in the computed image orientation. This in turn would affect the quality of the dense matching results. The following two points were highlighted in the chapter:

- **Use of open source solutions to recreate bundle adjustment results:** The experiments in this chapter were conducted mainly using a close-range UAV dataset

on two examples of built heritage buildings. The relatively complex network configuration of a close-range UAV dataset compared to classical aerial photogrammetry, the inconsistent nature of its overlaps and intersecting angles, as well as the varying average GSD render it more important but difficult to detect weak parts of the network. In these examples, PhotoScan generated bundle adjustment results that are of the same millimetric order as the theoretical GSD. This was validated by both DBAT and Apero, which produced similar precision as estimated by the GCP and CP residuals. The results were achieved under the same weighting conditions. A downside of PhotoScan is the relative lack of detailed metrics for its computed orientation values. This lack may mask potential problems in the dataset. The DBAT results suggest that some improvements on the project can be done, particularly by adding more overlapping photos for some zones with lower exterior orientation precision. The Apero results provide some indication of the image quality that might be used for similar reasons.

- **Importance of project quality assessment:** Overall, this chapter showed that an objective assessment of photogrammetric results is important. Errors may be hidden within the project, which limit the potential of the photogrammetric data. If the generated precision is within the project requirement, this may not be an issue. However, when problems occur within a dataset, bundle adjustment metrics, such as those provided by DBAT and, to a lesser degree, Apero, can be very helpful to detect the error. Open source methods may therefore be used to check the validity of the projects and to analyse the results in more detail. However, with the ever increasing use of commercial photogrammetry/SfM software for mapping purposes, we have seen many improvements in this regard. If the recent addition of bundle adjustment metrics in Metashape is anything to go by, open source solutions might lead the commercial software authors in the right direction. This will be a very welcome development for photogrammetrists and lay users alike.



# Chapter 3

## Integration of Photogrammetry and Laser Scanning

### Contents

---

<b>3.1 Background and rationale . . . . .</b>	<b>52</b>
<b>3.2 State-of-the-art . . . . .</b>	<b>53</b>
<b>3.3 Proposed integration method . . . . .</b>	<b>54</b>
<b>3.4 Experiments and results . . . . .</b>	<b>56</b>
3.4.1 Independent georeferencing . . . . .	56
3.4.2 Free-net registration and subsequent georeferencing . . .	57
3.4.3 Check point validation . . . . .	58
<b>3.5 Attempts on automation . . . . .</b>	<b>59</b>
<b>3.6 Summary . . . . .</b>	<b>61</b>

---

In this chapter, the quality control will be performed on the integration results between photogrammetric and laser scanning data. Due to the heterogeneous quality of each method, proper means for integration are required in order to generate a geometrically acceptable result. The integration which will be discussed in this chapter involves the exterior point cloud of a building, acquired by photogrammetry, and its interior counterpart generated by TLS. The case is quite particular as each recording mission was conducted separately in two different time epochs; therefore not much overlap was available between the two point clouds. Two integration methods will be tested in this chapter. The first one involves an independent georeferencing of each point cloud separately via a separate set of control points which were measured in the same coordinate system. The second approach attempts to register one point cloud into another, before georeferencing them together to reach the final real-world coordinate system. A comparison of both methods will be conducted using open source means, in order to deduce the advantages and disadvantages of each method. Furthermore, an attempt at automating this indoor-outdoor integration process will also be presented in this chapter, based on the PCL<sup>1</sup> library.

---

<sup>1</sup><http://pointclouds.org/>, accessed 19 February 2020

The contents of this chapter were previously published in [MURTIYOSO AND GRUSSENMEYER \[2018\]](#) and presented as an oral presentation at the 2018 ISPRS Technical Commission II Symposium in Riva del Garda, Italy.

### 3.1 Background and rationale

In the field of 3D heritage documentation, point cloud registration is a relatively common issue. With the rising needs for HBIMs, this issue has become more important as it determines the quality of the data to be used for 3D modelling. Furthermore, in the context of historical buildings, it is often interesting to document both the exterior façades as well as the interior.

As has been mentioned before, this chapter will discuss two proposed approaches for the registration and georeferencing of building exterior and interior point clouds coming from different sensors, namely the independent georeferencing method and the free-network registration and georeferencing. Building openings (mainly windows) were used to establish common points between the systems. These two methods will be compared in terms of geometrical quality, while technical problems in performing them will also be discussed.

Point clouds are a typical product of 3D recording processes. Point cloud generation in this regard is often limited to either the exterior of the object in question or its interior. In the case of heritage buildings, an exterior model is useful for façade analysis and general visualisation [[FRITSCH ET AL., 2013](#)]. On the other hand, interior point clouds had also seen an increase in demand, partly due to the advent of AR/VR technologies, as well as the increasing use of BIM [[QUATTRINI ET AL., 2015](#)].

The recording method for the exterior and interior parts of the building may also differ, depending on the nature of the object. For example, in order to generate exterior point clouds for tall structures, a UAV photogrammetric mission may be more appropriate than a TLS one. Conversely, UAV deployment in interior situations may be complicated, in which case TLS or close-range terrestrial photogrammetry may be better adapted.

In light of the necessity to create a complete 3D model of a heritage building, it is interesting to combine heterogeneous point clouds generated by these different sensors of the building's exterior and interior [[FASSI ET AL., 2011](#)]. Also in the interest of keeping the geometric quality, it is also important to devise a method of quality control for the resulting registered point cloud. This is more so when the resolution of the point clouds are heterogeneous, as in the case where the point cloud sources are different [[GRUSSENMEYER ET AL., 2012b](#)].

One main problem that arises from the registration of exterior and interior data is the lack of overlap between the two point clouds, more so when both datasets were acquired separately and using different types of sensors. While the use of a single coordinate system as explained in section 1.4 is quite easy to implement, one of the objectives of this chapter is to confront this method with another which involves prior registration between the two point clouds. A comparison of two registration approaches in order to combine the exterior and interior point clouds of a heritage building will thus be presented. Assessment of the geometric quality of each approach will be performed using checkpoint residuals. These checkpoints were measured independently from the main point cloud acquisition mission by means of terrestrial surveying methods. In addition, an attempt on the automation of some parts of the exterior-interior registration workflow will also be presented. This

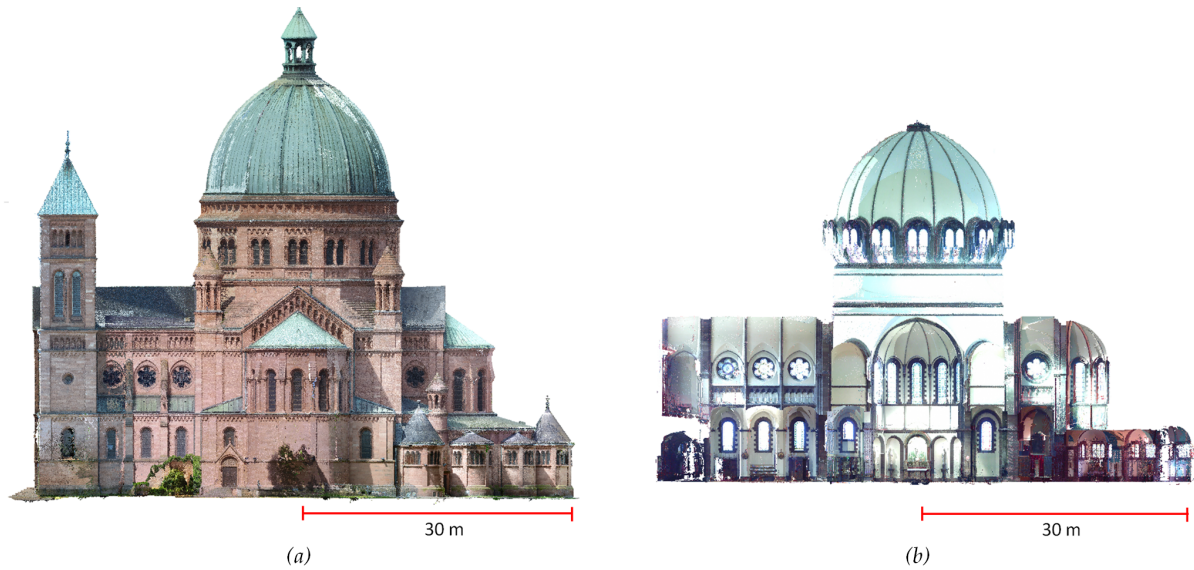


Figure 3.1: The two point clouds used in the experiment: (a) exterior point cloud generated by UAV and terrestrial photogrammetry, and (b) interior point cloud generated by TLS.

is performed essentially using automatic 3D feature detection and matching, followed by 3D transformation.

The case study used for this chapter’s experiments is the St-Pierre dataset’s exterior and interior point clouds. The exterior point cloud was obtained using a combination of UAV and close-range terrestrial photogrammetry (Figure 3.1a); while the interior point cloud was acquired using a TLS (Figure 3.1(b)). Unfortunately, we didn’t manage to obtain the interior point cloud for the St-Paul dataset, while in the case of Kasepuhan no real interior is present as the buildings consist of open-air pavilions.

## 3.2 State-of-the-art

The registration of separate point clouds into the same system has been much addressed. Basic principles of 3D registration involves similar methods to the ones used, for example, in photogrammetric absolute orientation [WOLF ET AL., 2014]. The 3D registration of point clouds typically commences with a coarse transformation, often computed using the classical 7-parameters 3D similarity transformation described previously in Equation 1.5. Normally this computation is based on the least-squares method, which enables a block adjustment on the whole system and is based on common points between the two point clouds. Several algorithms for solving the transformation computation can be found, e.g. Singular Value Decomposition (SVD), Principal Component Analysis (PCA), etc. [BELLEKENS ET AL., 2014]. The definition of these common tie points can be addressed using several methods, including 3D feature detection, automatic detection of artificial targets (e.g. spheres and coded targets), or manual point measurement. A further refinement of the resulting transformation can then be performed using algorithms such as the ICP method [BESL AND MCKAY, 1992]. Aside from the relative registration, the point clouds also need to be georeferenced to a common real-world system, so that accurate measurements may be performed on them. Several georeferencing approaches exist in the literature [LACHAT ET AL., 2017; BÖHM, 2005; RESHETYUK, 2009a], which mainly involve



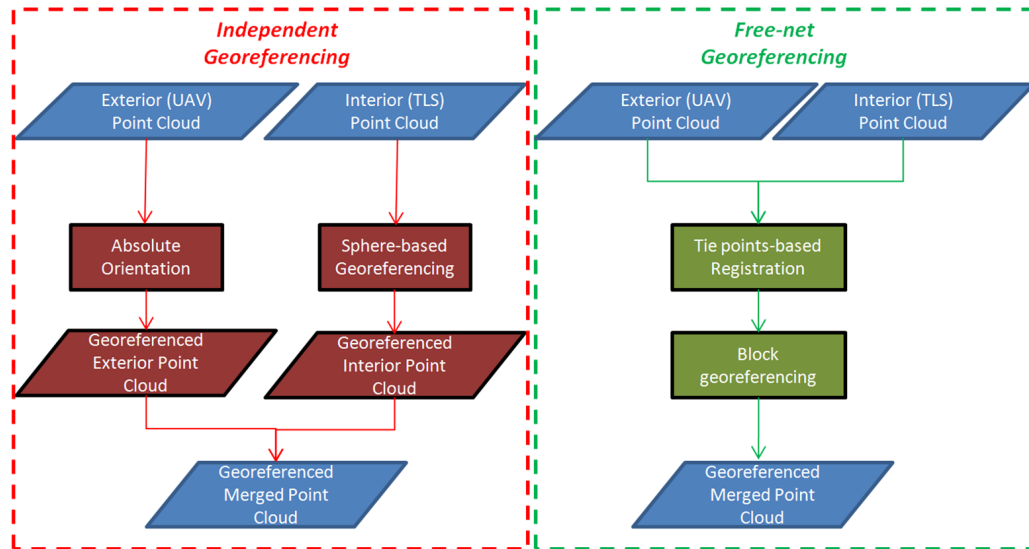


Figure 3.2: The two approaches to point cloud registration and georeferencing used in this experiment.

either an independent georeferencing for each point cloud ("direct" or sensor-driven) or free-network ("indirect" or data-driven) georeferencing.

An independent georeferencing involves separate transformations for each point cloud, which puts the point clouds directly in the same system. On the contrary, free-network (henceforth "free-net") georeferencing involves a relative registration of the point clouds, and followed by another 3D similarity transformation towards the cartographic projection system. Independent georeferencing of individual point clouds may provide a faster result, given the fact that they come from different sources, which may already involve a georeferencing process (e.g. absolute orientation in photogrammetry). An ICP process can be performed at this stage to combine both point clouds; however in the case of the combination of exterior and interior data this is not ideal since only small overlap exists between the two data. As such, a block transformation computation between the point clouds is proposed in the free-net approach, in order to link the point clouds through a block adjustment process.

The identification of tie points between the varying point clouds is an important task in the registration workflow. One method to do this is to identify the tie points manually [LACHAT ET AL., 2016; MUNUMER AND LERMA, 2015]. Various experiments on the automation of this task can be found in the literature. Some approaches transform the 3D point cloud into 2D depth maps [WEINMANN, 2016] and perform image matching on the two resulting images to find their correspondences [FORKUO AND KING, 2004]. Another approach involves the detection of 3D keypoints and the computation of feature descriptors on both point clouds [HÄNSCH ET AL., 2014; HOLZ ET AL., 2015; HILLEMANN ET AL., 2019].

### 3.3 Proposed integration method

In this study, two registration and georeferencing approaches will be considered (see Figure 3.2). The first corresponds to the independent georeferencing method, in which both point clouds (exterior and interior) were georeferenced separately. The exterior photogrammet-

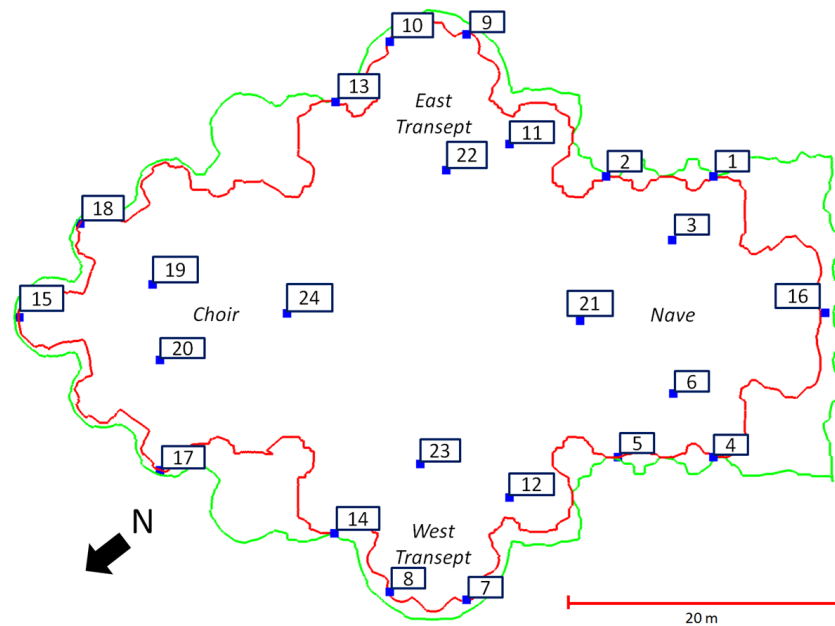


Figure 3.3: Horizontal profiles of the exterior (green) and interior (red) data showing the tie point distribution used for the free-net registration.

ric point cloud was georeferenced by means of absolute orientation, while the interior TLS point cloud was georeferenced using the coordinates of the 3D spheres measured during the acquisition. Both georeferencing were performed on the same system, thereby giving a merged result directly in the absolute system. The second free-net approach registered both ungeoreferenced point clouds in an arbitrary coordinate system, before applying 3D similarity transformation to attach it to the georeferenced system via control points. The registration is based on a set of tie points which are evenly distributed throughout the building. These tie points concern mainly openings such as windows. The tie points were identified in the first place by manual measurements, with results to be discussed in section 3.4. Afterwards, an attempt to automate the tie point detection and matching will be presented in section 3.5. In both cases, an ICP fine registration procedure is not feasible due to the minimum overlapping zones between the two datasets.

In the free-net approach, a total of 24 tie points were identified on the church (Figure 3.3). The windows which represent overlapping spaces between the exterior and interior were used as point candidates, although physically they are not the same points due to the thickness of the window panes. In this regard, by assuming a window pane thickness of around 2 cm, an error of this order should be expected on the final results. Furthermore, the point clouds used in the experiment were subsampled to 1 cm in order to give a more manageable dataset while keeping enough details to identify as tie point candidates.

In order to assess the quality of the registration and georeferencing of both approaches, a set of independent check points were measured on the exterior and interior of the building using a total station. The spatial intersection method was chosen to measure these points, in order to generate high-precision coordinates with standard deviation values available. Measurements were based on a traverse network established in and around the church. Each check point was measured from at least two stations. Similar to the control points on the St-Paul dataset, the Trimble S8 robotic total station was used to perform the measurements. Coordinates of these points on the processed point cloud issuing from both approaches were then compared to the ones measured using the total station. The

comparison of their residuals was then used as a criterion for quality assessment.

In terms of software used in the experiments, the open source software CloudCompare (version 2.9.1) and 3DVEM were used. CloudCompare offers a 3D transformation module in which a standard similarity transformation is performed. The software gives the users possibility to choose between a 6-parameters (no scaling) or 7-parameters (with scaling) transformation. 3DVEM is a scientific software programme developed by the University of Valencia (Spain) which performs point cloud registration and georeferencing with the possibility to use the standard least-squares based method or robust estimators (namely the modified Danish and minimum sum estimators) [FABADO ET AL., 2013]. One of its advantages is access towards supplementary statistics (e.g. standard deviations) which is often useful as means of project quality control. 3DVEM does not, however, support registrations with scaling factors. The automation experiment uses the Point Cloud Library (PCL) and its functions which were written in the C++ language [RUSU AND COUSINS, 2011].

## 3.4 Experiments and results

In this section, the two proposed integration methods hitherto described in this chapter and illustrated by Figure 3.2 will be described. The technical step-by-step process will first be explained to achieve both results. The first method in section 3.4.1 is more or less a single step process albeit performed separately for each point cloud. The second method of section 3.4.2 included two steps reminiscent of relative and absolute orientation in photogrammetry and indirect georeferencing in TLS processing. A check-point based method to assess the result quantitatively will thereafter be presented, with critical discussions on the potential sources of error.

### 3.4.1 Independent georeferencing

In the independent georeferencing approach, each dataset was georeferenced separately using methods employed in their respective workflow. The exterior point cloud was generated entirely from images taken using UAVs and cameras; thus the georeferencing follows the absolute orientation method normally seen in photogrammetry. A total of 25 ground control points (GCPs) were used to this end, yielding an overall RMS value of 1.8 cm on the GCP residuals.

The interior of the church was scanned entirely using the FARO Focus X330 terrestrial laser scanner (TLS). The registration between the stations was performed using the FARO Scene software, with the aid of automatically detected spheres and targets. The coordinates of these spheres were measured using a total station, enabling the georeferencing of the interior point cloud. A total of 31 spheres and targets were used in the georeferencing process, yielding an overall final precision of 1.0 cm.

The control points used in both datasets were measured from traverse network points. This network was measured in the French national projection system, with several points fixed using GNSS-measured coordinates. Since no levelling benchmark was to be found near the site, the GNSS altitudes were used in the coordinate computation. This common projection system used by the control points in both datasets means that at the end of their respective georeferencing process the interior and exterior point clouds were directly merged in one coordinate system. The visual results of this approach were illustrated in Figure 3.4.

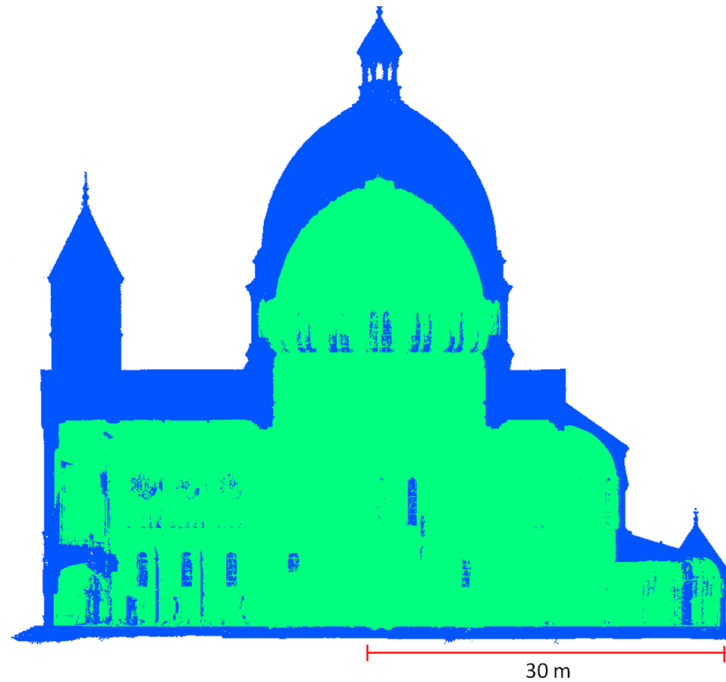


Figure 3.4: The two point clouds registered and georeferenced using the independent method. The blue-coloured part represents the exterior while the green-coloured one represents the interior.

### 3.4.2 Free-net registration and subsequent georeferencing

The free-net approach, on the other hand, assumed that both datasets have been neither registered on each other nor georeferenced to the reference system. In practice, in the interior TLS data this means that the point cloud was simply not georeferenced. In the photogrammetric exterior point cloud, the absence of absolute orientation implies that the relative model is not scaled either. Since 3DVEM does not support scaling during the registration step, absolute orientation was nonetheless performed on the photogrammetric data. However, noises on the rotational and translational parts were introduced on purpose to generate a simulated data for the purposes of the experiment.

The registration was based on the 24 tie points manually measured on both point clouds. The process was performed on the software CloudCompare as well as 3DVEM. In 3DVEM, a standard least-squares (LS) based transformation as well as the minimum sum (MS) robust estimator [SHARON ET AL., 2009] were tested. The results in terms of tie point residuals are showed by Table 3.1.

All proposed algorithms gave similar results of the order of 3.0 cm in RMS. Assuming a point marking precision of 1.0 cm, this value is only slightly outside the *a priori* tolerance of  $2\sigma$  (for a level of confidence of 95%) but falls within the  $3\sigma$  (99.8% level of confidence) range. It is also worth noting here that the use of robust estimators in this case does not show a significant difference from the use of standard least squares solution.

Once the registration was performed, the georeferencing step followed which used 15 photogrammetric GCPs on the exterior and 6 measured 3D spheres in the interior. This georeferencing was performed in block, using all 21 control points at the same time. Again, CloudCompare and 3DVEM were used to this end. Both implement a standard rigid-body transformation to perform this step. Results to this georeferencing process can be consulted in Table 3.2.

Free-net registration		
Software		Residuals RMS (cm)
CloudCompare		2.8
3DVEM	LS	2.9
	MS	2.9

Table 3.1: RMS of the tie point residuals used for the registration of the exterior and interior point clouds using the free-net approach.

Free-net georeferencing		
Software		Residuals RMS (cm)
CloudCompare		3.1
3DVEM		3.1

Table 3.2: RMS of the control point residuals used in the georeferencing process using the free-net approach.

The RMS of the control point residuals in both solutions shows similar results. The values obtained are also of the same order with the registration RMS, further showing the georeferencing precision of the process and the absence of systematic error.

### 3.4.3 Check point validation

In order to validate the accuracy of the results from the independent and free-net approaches, checkpoints were compared between those measured separately using a total station and those measured on the resulting merged point clouds. While the georeferencing RMS in each method shows their respective precisions, this checkpoint analysis will enable us to determine the accuracy of each result. The average standard deviation of these check point coordinates is 0.8 cm.

The results of the checkpoint analysis in terms of RMS values can be seen in Table 3.3. The 3D RMS values of the free-net approach all show consistencies with their respective registration and georeferencing RMS values. When compared against each other, the free-net RMS values have slight differences, but these are of the order of 1 to 3 mm and may well reflect the influence of random error during the checkpoint measurement on the point clouds. The independent approach shows a slightly higher value compared to the precision of the respective georeferencing process of the exterior and interior point clouds. In order to detect any irregularities, a comparison was performed for each individual checkpoint.

Upon closer look, the independent method showed signs of systematic error towards the front of the building (see Figure 3.5). This is particularly true for the exterior data, and may have originated from error propagated from the absolute orientation process of the photogrammetric data. It should be noted that the photogrammetric point cloud was generated in smaller parts before combined in one merged point cloud of the exterior with the aid of GCPs. Furthermore, the front façade of the church presented a case where the imaging sensor is different from the rest of the building. This may have also generated small errors, which results in the systematic trend of the checkpoint residuals.

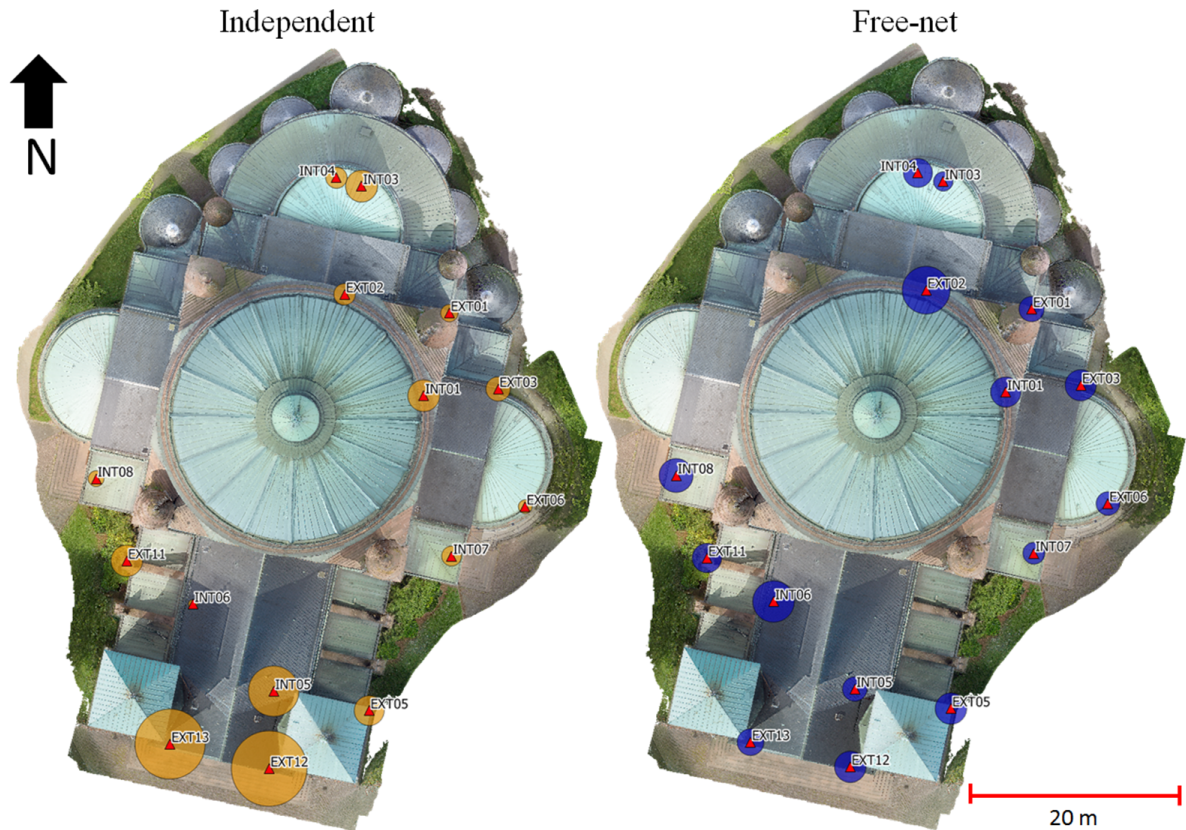


Figure 3.5: Distribution of the measured checkpoints superimposed on the nadir orthophoto of the church. The radii of the circles around the points represent the 3D residuals of each point, magnified by a factor of 500 for presentation purposes.

The free-net checkpoint residuals, as can also be seen in Figure 3.5, showed a more even distribution of error throughout the whole site. This is to be expected, due to the fact that the georeferencing was performed as a unique block to the reference system. The systematic errors present in either of the datasets were therefore evenly adjusted in this case.

### 3.5 Attempts on automation

In the previous section 3.4.2, the free-net approach has been demonstrated to be useful in cases where independent georeferencing may induce systematic errors. However, in terms of processing time, the independent approach requires less time in the context of an exterior-interior data acquisition. This is because in many projects, the presence of control points are part of the accepted workflow and will in any case be used to georeference the data. The free-net approach, on the other hand, requires additional work in terms of the identification of tie points for the purpose of the registration. This can be a tedious process, and a way to automate this step can greatly reduce the overall processing time.

In this regard, this section will describe some experiments on the automation of the tie point identification. In these experiments, the PCL library of functions was used. This section will mainly describe the detection of 3D keypoints, the computation of the feature detectors, and the matching of these 3D tie points using some of the functions already implemented in PCL version 1.8.0.

Methods		RMS			
		X (cm)	Y (cm)	Z (cm)	3D (cm)
Independent		1.7	1.5	1.8	2.9
Free-net	CloudCompare	1.4	1.2	1.6	2.4
	3DVEM - LS	1.2	1.3	1.8	2.5
	3DVEM - MS	1.5	1.3	1.8	2.7

Table 3.3: RMS of the checkpoint residuals for all tested methods.



Figure 3.6: The steps in the keypoint detection and matching: (a) image of an example window opening. Here is shown the exterior side; (b) segmented point cloud of this window; (c) window pane extracted using a RANSAC algorithm; (d) detected SHOT keypoints. Red denotes exterior keypoints and green interior keypoints.

The first step to the proposed approach involves the pre-segmentation of the exterior and interior point clouds into smaller ones centred on several openings (i.e. windows). The automatic detection of openings has been much discussed in the literature (see for example BOULAASSAL ET AL. [2007] and ROCA ET AL. [2013]); however this will not be discussed in this section. For the purposes of our tests, the segmentation was performed manually, using the same distribution of tie points as the one used in section 3.4.2. Figure 3.6 shows some of the steps taken in this experiment.

Direct keypoint detection and matching to this segmented point cloud proved to be insufficient, as the algorithm takes the exterior and interior parts to be the same object. A RANSAC-based plane fitting was therefore employed in order to extract the window panes, the part which involves the overlapping zone between the two data (Figure 3.6b). Working with these data, 3D keypoints were detected on the window glasses. To this end, the 3DSIFT algorithm was used. The 3DSIFT is an implementation of the SIFT method

[LOWE, 2004] for 3D data. Results of the keypoint detection can be seen in Figure 3.6c.

In order to find correspondence matches between the detected keypoints, descriptors need to be computed for each keypoint on each data. In this experiment, the SHOT (Signature of Histograms of Orientations) [TOMBARI ET AL., 2010] and FPFH (Fast Point Feature Histograms) [RUSU ET AL., 2009] descriptors were tested. SHOT gave better results in this particular case. After the computation of the descriptors, a reciprocal correspondence was computed on both keypoint sets to find matches. Finally, a filtering was performed on the detected correspondences to reject bad matches. This filtering step also encountered problems, mainly when faced against noisy point cloud. As the object in question constituted a reflective material, noise is a very important issue, both for photogrammetry and TLS. This becomes worse with the presence of vegetations on the exterior and bad scanning angles in the interior.

The approach proposed in this section managed to generate a total of 35 tie points from 24 pre-segmented openings. Upon performing the point cloud registration using this set of automatic tie points, an RMS on the computed 3D residuals gave a value of 4.4 cm. This is admittedly still not as good as the results obtained from manual measurements; however this early development may prove to be a very useful tool in reducing the processing time of tie point identification.

## 3.6 Summary

This chapter discussed two important points:

- **Pros and cons of the tested integration methods:** Analysis of the two tested approaches showed that while it is possible to obtain similar final results from both methods, there are some advantages as well as disadvantages to each approach. The independent approach provided a faster solution. Indeed, in projects where control points are measured as part of the established workflow, independent georeferencing is unavoidable. However, the quality of the end result will depend strongly on the separate georeferencing processes implemented in the interior and exterior point clouds. The free-net approach, on the other hand, enables the combining of both data in the same adjustment system. This compensated any eventual systematic error, as has been demonstrated with the experiment. However, the downside of this approach is the necessity to identify tie points on both sides of the building, which may take a lot of time when performed manually.
- **Tie point automation to support the free-net method:** A simple chain of automation implemented using PCL was also proposed, which enabled the partial automation of this tie point identification process<sup>2</sup>. However, as this is not part of the main scope of the thesis, further development had been annulled in order to focus more on the point cloud segmentation and classification part of the thesis. Nevertheless, the developed functions served as a base for other work in our research group, especially the work by ASSI ET AL. [2019]. It was also investigated further by two master projects for different purposes: PEGOSSOFF [2018] investigated the possibility to use the same approach to automatically register heterogeneous point

---

<sup>2</sup>The code written for this purpose is open source and can be consulted in the following link: <https://github.com/murtiad/HERACLES>, accessed on 20 February 2020.



clouds, while [AMZIL \[2020\]](#) presented a similar work related to the automatic registration of indoor and outdoor point cloud. The reader is therefore advised to consult these papers and reports in order to know our group's latest update on this issue.

The experiments aimed to test two registration and georeferencing methods for exterior and interior point clouds, in the case of heritage building recording. This is done primarily as a preliminary quality assessment for the point cloud generated by different sensors, before further use in other purposes such as point cloud segmentation and classification. Indeed, geometric accuracy quality in 3D models becomes more and more important in order to faithfully represent the reality. This has led to the necessity to properly validate point cloud data obtained from reality-based techniques.

This chapter concludes the first part of the dissertation concerning the geospatial data recording and quality control for heritage buildings. The next chapters will address the point cloud processing part of the thesis; that is, algorithms and functions developed to process the point clouds further. Indeed, the point cloud as the primary result of the recording process presents hitherto only the geometric aspect of the object representation. The next part will describe our approach to process the point cloud further by proposing methods to automate the semantic annotation process, therefore rendering them more tangible and useful for more advanced applications such as 3D GIS or HBIM.

---

PART II

POINT CLOUD PROCESSING  
ALGORITHMS

---



# Chapter 4

## The M\_HERACLES Toolbox

### Contents

---

<b>4.1 Principal motives of the development . . . . .</b>	<b>66</b>
<b>4.2 General state-of-the-art . . . . .</b>	<b>67</b>
4.2.1 Approaches to point cloud classification . . . . .	68
4.2.2 Point cloud processing techniques . . . . .	69
<b>4.3 Supporting datasets . . . . .</b>	<b>74</b>
<b>4.4 Summary . . . . .</b>	<b>76</b>

---

When addressing the 3D documentation of heritage, in many cases the historical edifice of interest is located within a larger heritage complex or site. A thorough documentation may therefore incorporate this larger area, often at the scale of a neighbourhood, into the mission. In this regard a multi-scalar and multi-sensor approach is unavoidable; each sensor is usually adapted only for one object scale, e.g. close range photogrammetry for statues or phase-based TLS for building interiors (cf. Figure 1.4). Indeed, smaller scale objects or larger areas do not need the same fine resolution as an artefact or architectural detail. To address this issue, we propose not only a thorough multi-scalar recording of heritage complexes, but also a progressive point cloud processing. The multi-scalar 3D data acquisition pipeline has been adequately explained in [MURTIYOSO ET AL. \[2018b\]](#) and described in section 1.3. This chapter will address the multi-scalar point processing part, starting from the scale of a neighbourhood up to that of architectural elements. This was done by means of a modular toolbox called M\_HERACLES. In this modular approach, inspired by [BÖRLIN ET AL. \[2019b\]](#), independent functions were developed for addressing particular objects (e.g. columns, wooden beams, etc.) while still following a systematic multi-scalar progression. This approach will also help the 3D modelling pipeline down the line, by attributing semantic information on an object’s class depending on the function employed. The individual algorithms developed to segment and classify the heritage point cloud will be described in Chapters 5 and 6. Meanwhile this chapter will discuss about the general background to the development of M\_HERACLES, some theoretical notions, as well as description on some additional datasets to the main ones already mentioned in section 1.7.

## 4.1 Principal motives of the development

The 3D documentation of heritage complexes or quarters often requires more than one scale due to its extended area. While the documentation of individual buildings requires a technique with finer resolution, that of the complex itself may not need the same degree of detail. This has led to the use of a multi-scale approach in such situations, which in itself implies the integration of multi-sensor techniques. The challenges and constraints of the multi-sensor approach are further added when working in urban areas, as some sensors may be suitable only for certain conditions. In the context of the documentation of heritage complexes, the extent of the site makes it logical to use a multi-scalar approach. The site can be digitised in several scale steps, according to the required resolution for each level [FIORILLO ET AL., 2013]. The use of the multi-scalar approach also means that more than one sensor could be employed in order to cover each scale step. Furthermore, for the case of urban areas, many constraints such as the geography and urban density mean that the use of one single sensor may not be sufficient. It is in this regard that multi-sensor and multi-scalar documentation became a logical solution to the problem of documenting historical complexes [MURTIYOSO ET AL., 2018b].

Furthermore, the geometric recording alone is in many cases not enough. While many useful derivatives can already be generated from a geometric point cloud (sometimes with the RGB information attached) as has been explained in section 1.6, a more advanced application such as 3D GIS and HBIM requires the user to classify the points into cluster entities. This is similar to the relation between CAD and GIS, in which the latter is annotated by non-geometric properties and is often labelled according to pre-established classes or entities. In GIS, the presence of semantic information enables further modelling, analysis, and even future predictions according to the existing data. Indeed, semantic information imbues tangible meaning into the otherwise uniquely numerical point cloud (i.e. a set of points with Cartesian coordinates). Analogous to this 2D relation between CAD and GIS, the same can also be said about point cloud semantic segmentation in which the 3D point cloud is segmented according to their respective classes and labelled.

The task of point cloud segmentation and subsequent classification is thus an important part of the 3D reconstruction process, as exemplified by Figure 4.1. However, several bottlenecks exist in this pipeline (shown as red inversed trapeziums in Figure 4.1) which are mainly caused by the requirement for manual intervention. This includes the segmentation of a raw point cloud into meaningful clusters and the labelling (semantisation) of these clusters into appropriate classes. The developments in this study focus on these two particular bottlenecks in the 3D pipeline, which present a more complicated case due to the complexity of heritage buildings. The 3D modelling of these clusters into geometrical primitives or solid models as required by HBIM and/or 3D GIS also presents a manual bottleneck; however this thesis will only briefly address this topic.

In our proposed workflow, the multi-scalar approach is used to progressively segment the point cloud of a heritage complex; first into building units and then further into architectural elements (more precisely wooden frames and structural supports). In addition to progressive segmentation, the developed method will also try to classify the results as automatically as possible in order to add the semantic dimension to the data, which is vital in BIM and 3D GIS environments. In this regard, a toolbox was created in the Matlab<sup>©</sup> environment to host all the codes and functions written for the research under one project: HERitAge by point CLOUD procESSing for Matlab<sup>©</sup> (*M\_HERACLES*). The aim of *M\_HERACLES* is to develop simple algorithms to help in the automation effort of point

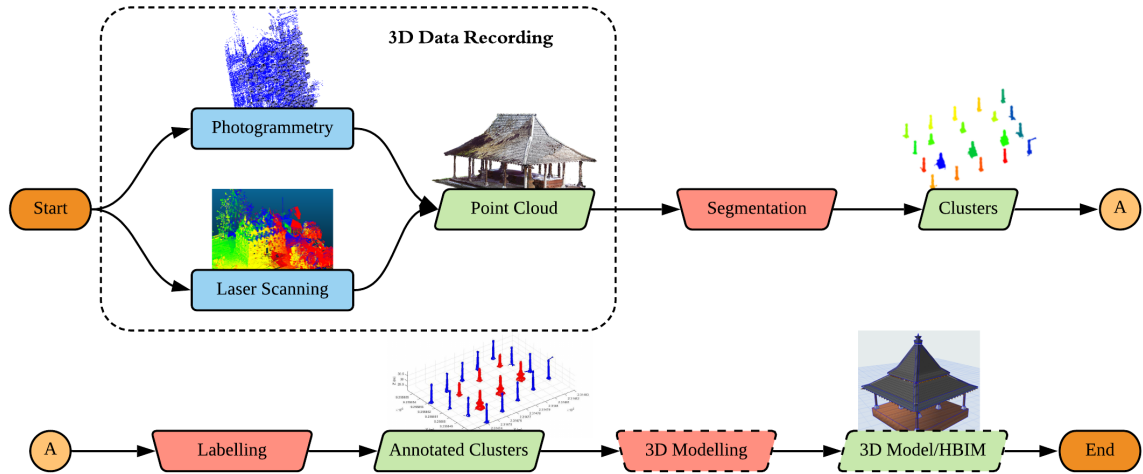


Figure 4.1: The overall pipeline for 3D reconstruction, from a point cloud up to HBIM-compatible 3D models. *M\_HERACLES* development focuses on the manual bottlenecks of the pipeline (red inverted trapeziums) up to before the 3D modelling process (long-dashed elements), although a preliminary result of automatic 3D modelling of beam structures will also be briefly presented in Chapter 6.

cloud processing in the context of cultural heritage. This includes among others segmentation, semantic annotation, and eventually 3D primitive generation. The toolbox is open source and available online via GitHub<sup>1</sup>. *M\_HERACLES* is mostly developed in Matlab<sup>®</sup> R2018a using its Computer Vision Toolbox and several other third party libraries.

## 4.2 General state-of-the-art

The documentation of heritage objects has been addressed in a lot of literature. As has been established beforehand, the documentation process takes more and more the form of 3D recording. Nowadays, the use of image-based (e.g. photogrammetry) and range-based techniques is very common [REMONDINO AND RIZZI, 2010; HASSANI, 2015] and may even be complementary to each other. Several useful guidelines also exist to advise stakeholders who do not have a surveying background on good practices in the subject [BARSANTI ET AL., 2014; BEDFORD, 2017]. Numerous examples exist in the literature on the use of 3D techniques for heritage documentation, e.g. the work of FANGI [2019], FIORILLO ET AL. [2013] and HERBIG ET AL. [2019] to cite a few. Another trend which has surfaced as a logical consequence of the availability of multiple sensors is data integration, both in the sensor level and the point cloud level [GRENZDÖRFFER ET AL., 2015; MURTIYOSO ET AL., 2017c; FARELLA ET AL., 2019; MUNUMER AND LERMA, 2015].

The automation of the 3D pipeline is a recurring topic in the literature. Notably this work is inspired by similar work in our group by MACHER ET AL. [2017] and YANG ET AL. [2018]. While MACHER ET AL. [2017] attempted to automate the scan-to-BIM process for modern buildings, YANG ET AL. [2018] focused more on the mesh-to-BIM process for heritage buildings. The first research saw limited results when applied to heritage buildings, while the latter encountered the problem of complex architecture when addressing heritage buildings. MAALEK ET AL. [2019] presented a research in an industrial

<sup>1</sup>[https://github.com/murtiad/M\\_HERACLES](https://github.com/murtiad/M_HERACLES), accessed 21 February 2020

architectural setting, similar to MACHER ET AL. [2017]. Some of the ideas presented by MAALEK ET AL. [2019] and LUO AND WANG [2008] served as an inspiration to parts of Chapter 6. Meanwhile, BASSIER ET AL. [2017b], MALINVERNI ET AL. [2019], and GRILLI AND REMONDINO [2019] attempted to use machine and deep learning to solve the automation issue. In this thesis the algorithmic approach similar to the one proposed by MAALEK ET AL. [2019] is preferred, in part in order to help with training data generation for future potential machine learning solutions.

Apart from the algorithms developed by MACHER ET AL. [2017] and YANG ET AL. [2018] some other libraries already exist for similar purposes, such as the PolyFit library [NAN AND WONKA, 2017]. Furthermore, the Point Cloud Library (PCL) [RUSU AND COUSINS, 2011] which was already mentioned in Chapter 4 also contains functions that could be used to automate the 3D modelling process. However, this is not the main purpose of PCL.

### 4.2.1 Approaches to point cloud classification

Several approaches to point cloud processing within the context of classification exist in the literature. A very general division of point cloud segmentation and classification is given in NGUYEN AND LE [2013], in which the existing algorithms are divided into either the use of geometric axioms and mathematical functions, or the use of machine learning techniques. This division is concurrent with ideas presented by MAALEK ET AL. [2019], in which the former is mentioned as the use of geometrical, spatial and contextual constraints. BASSIER ET AL. [2017b] mentioned a distinction between heuristic and machine learning techniques. Another attempt to classify the existing approaches was proposed by GRILLI ET AL. [2017], in which the authors added region growing algorithms [BASSIER ET AL., 2017a; VO ET AL., 2015], edge-based segmentation [BOULAASSAL ET AL., 2007], and model fitting [SANCHEZ AND ZAKHOR, 2012] as other possible segmentation approaches, while point cloud classification is divided into supervised (data-training), unsupervised, or interactive manner.

It is worth noting that most of the examples seen in the literature address a particular scale level for the point cloud. For example, RIZALDY ET AL. [2018] focused on small scale point clouds or larger areas mainly done to support surveying purposes, while BASSIER ET AL. [2017b] performed the point cloud processing at the larger scale of a building. Many algorithms were also developed with modern objects in mind [MACHER ET AL., 2017; RIVEIRO ET AL., 2016] even though forays into the heritage domain is becoming more numerous in recent years [GRILLI AND REMONDINO, 2019; MALINVERNI ET AL., 2019]. The goal of this research is to develop a toolbox which enables the processing of multi-scalar heritage point cloud, from the scale of a neighbourhood (heritage complexes) up to that of architectural elements. This is encouraged by the increasing trend towards multi-sensor and multi-scalar recording missions for heritage sites [MURTIYOSO ET AL., 2018b; FARELLA ET AL., 2019].

### Machine learning and deep learning approaches

In general, machine learning and its subset deep learning solutions have seen a surge in popularity these recent years since the advent of big data [LIU ET AL., 2019]. Machine learning approaches are robust against noise and occlusions, and are generally reliable. Its main disadvantage, however, is the necessity of a large amount of training data and the computing power needed to train the algorithm. The usual method to create training data

is to segment and classify point clouds manually [ANTONOPOULOS AND ANTONOPOULOU, 2017], although synthetic training data can also be generated in some cases [ROS ET AL., 2016]. It also remains a largely black-box solution and therefore leaves very little room for user intervention [BASSIER ET AL., 2017b].

Various types of machine learning and deep learning techniques are available, as described in LIU ET AL. [2019]. In GRILLI ET AL. [2019], a comparison on several machine learning and deep learning techniques were performed. MALINVERNI ET AL. [2019] described a deep learning approach to classify outdoor point clouds in the case of heritage sites, while WANG ET AL. [2015b] proposed the use of a multi-scalar approach for classifying multi-resolution TLS data. As deep learning is a well established technique in the realm of 2D image recognition, one way to perform point cloud classification is to apply deep learning on 2D images created from point cloud colour (orthophotos, UV textures, etc.) [GRILLI AND REMONDINO, 2019]. The technique is also often used to perform the segmentation and classification of point cloud generated by aerial platforms (aerial photogrammetry, LIDAR) as it enables the reduction of the (usually more complex) 3D point cloud into a 2.5D problem [RIZALDY ET AL., 2018].

While the appeal of machine learning is strong for performing point cloud processing in the case of complex geometries as encountered in heritage objects, the main bottleneck remains the generation of the training dataset [MAALEK ET AL., 2019]. In this study, an algorithmic approach is considered in order to provide a fast result which may eventually be used to help generate training data for future machine learning techniques. Indeed, manual labelling of heritage objects also present a particular difficulty since objects in the same class may have many variations.

### Algorithmic approach

The algorithmic approach employs geometric rules and mathematical functions to perform point cloud segmentation [MACHER ET AL., 2015]. This approach is often heuristic in nature, but maybe enough for certain purposes as they are fast and simple to implement [BASSIER ET AL., 2017b]. Algorithmic segmentation uses mathematical rules and functions as constraints during the segmentation (and possibly also classification). These rules may range from simple rules (e.g. "floors are flat and located below each storey" or "pillars are cylinders") [MAALEK ET AL., 2019] to the implementation of ontological relations [POUX ET AL., 2018; LU ET AL., 2018; DRAP ET AL., 2017].

The rules and constraints in this type of method are often determined differently according to the encountered case. MACHER ET AL. [2017] employed a type of multi-scalar approach by sub-dividing the point cloud into floors, rooms, and thence walls. RIVEIRO ET AL. [2016] similarly used geometric constraints to segment the walls of bridges. In POUX ET AL. [2018], relational ontology was used as constraints in determining the classes of point clouds segmented by connected component segmentation.

## 4.2.2 Point cloud processing techniques

In this section, several well-established algorithms which were used and were crucial during the development of M\_HERACLES will be briefly described. Furthermore, the specific use of each algorithm in M\_HERACLES will be mentioned at the end of each description. The description provided here will concern some basic concepts in point cloud processing and will not dive into details. Interested readers are strongly encouraged to consult the references for an in-depth explanation for each algorithm.



## Hough transform

The Hough transform is a type of robust optimisation method for the classical data fitting problem. It was first introduced in a patent by HOUGH [1962] and named after the author. However, the modern method as we know it today was developed by DUDA AND HART [1972] and then further popularised for use in computer graphics by BALLARD [1981]. This method is most often used to detect lines in a set of 2D points, but may also be expanded into the 3D space to detect surfaces and other forms such as cylinders [VOSSELMAN AND MAAS, 2010]. In this thesis, the classical 2D version of the algorithm was used to some extent in section 6.3 to detect lines, while 3D data fitting was mainly performed using the RANSAC paradigm.

In order to use this method to detect lines from a set of points, the main idea is based on the fact that any line passing on a Euclidean plane will have a constant bearing (i.e. angle as measured from the Y-axis). Indeed, a 2D line may be defined as such:

$$d = X \cos \alpha + Y \sin \alpha \quad (4.1)$$

in which the angle  $\alpha$  stands for the angle between the line and the Y-axis, and  $d$  denotes the distance from the line to the system origin. Equation 4.1 describes an infinite amount of points passing through the line. When we inversed the problem by placing a set of points with known Cartesian coordinates as the input, each point will thus generate a unique sinusoidal curve in the parameter space. In Figure 4.2, an example is given to help explain these proceedings. A line is to be determined from the set of points in Figure 4.2a, where some noises exist. Each point on Figure 4.2a is transformed, so to speak, into the parameter space of  $\alpha$  and  $d$  in Figure 4.2b. For the red point located at the coordinates  $(-2, -1)$ , the generated curve is shown by the red curve. This curve represents all the possible lines that may intersect on the red point with varying values of  $\alpha$  and  $d$ . The same curves were traced for all the other points (green and blue), and we can see that the curves intersect at several points. By determining the point where most curves intersect (denoted as the point A in Figure 4.2b), we may determine the most probable line to pass by the set of points in Figure 4.2a all the while eliminating the noises.

Note that in Figure 4.2, one blue point is located approximately on the line's trajectory, but located far from the other inliers. The Hough transform method cannot detect this information; this means that even noises that are located within the same line may be considered as part of the detected line. A possible solution to this problem is a subsequent clustering by e.g. the region growing algorithm.

In order to detect the intersection point with the most curves in the parameter space (also called the Hough space) using computers, Figure 4.2b must be discretised [VOSSELMAN AND MAAS, 2010]. This means that the choice of the discretisation bin is very important, as the curves may not intersect *per se*, but are only very close to each other. A bin distance that is too large may consider noises as part of the detected line, while a bin distance which is too small may slow down the computation.

The Hough transform was used within M\_HERACLES during the automatic beam detection (section 6.3). It was mainly used in its 2D form, since one of the techniques used in section 6.3 is projection of 3D space into 2D space. It was thus utilised to determine the beam axes in a building roof framework setting.

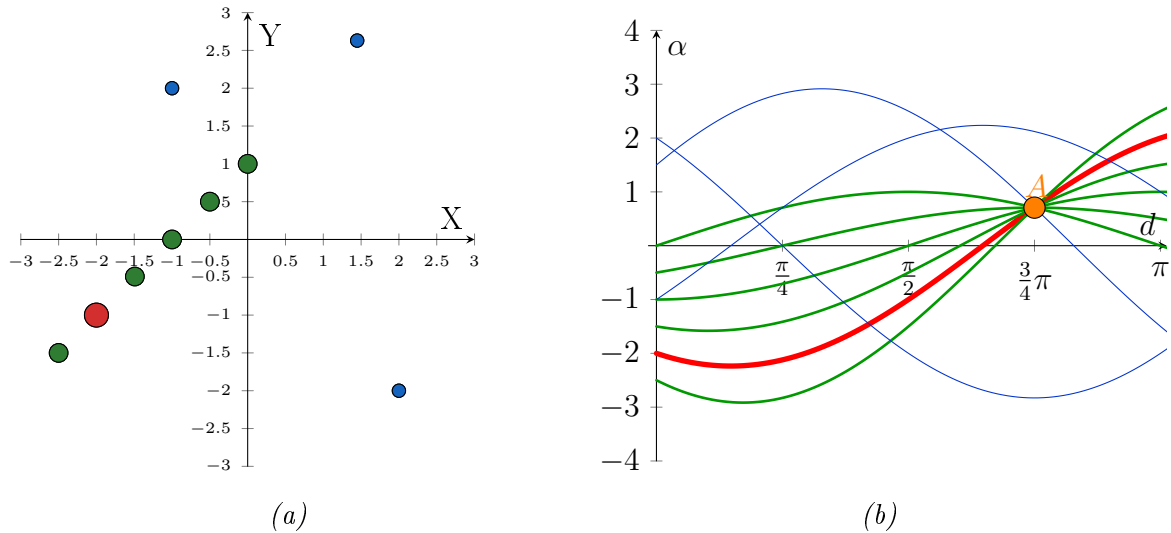


Figure 4.2: Illustration of the Hough transform concept, with (a) representing arbitrary 2D points and (b) the corresponding curves in the parameter space. Adapted from [VOSSELMAN AND MAAS \[2010\]](#).

## Random sample consensus

The random sample consensus or RANSAC algorithm also presents a robust approach to the data fitting question. First developed by [FISCHLER AND BOLLES \[1980\]](#), the RANSAC paradigm is based on a voting system in which a score is attributed to each model fit candidate. RANSAC is versatile and based on a simple premise [[VOSSELMAN AND MAAS, 2010](#); [BOULAASSAL ET AL., 2007](#)] and is therefore adaptable to different kinds of mathematical model. In the 2D space, a simple application of RANSAC is for robust linear regression. In 3D, the chosen model to be fitted may be more diverse, e.g. planes, spheres, and cylinders (cf. Figure 4.3).

RANSAC starts by taking a minimal set of points as required for the definition of a particular geometric model. For example, 2 points are required to define a line and 3 points are required to define a plane. This minimal set of points are chosen randomly, and then used to compute the geometric model required. Afterwards, all the other points in the data are tested against this acquired model. A certain distance from the model is introduced as a tolerance, with points located within this tolerance labelled as inliers and the others as outliers. The number of inliers linked to a particular minimal set and its geometric model serves thus as its consensus set or score. The process is then repeated a certain number of times, with better scoring sets replacing previous ones as the best fit model.

While the basic concept of RANSAC is quite simple, problems may arise when handling large datasets. This is because the algorithm is supposed to test all the possibilities in the dataset to create the geometric model from the minimal set. Thus with increase in point number, the available possibilities also increase exponentially. To address this problem, RANSAC offers a probabilistic approach in which the number of iterations is determined prior to the computation itself. The following Equation 4.2 is generally used to determine this iteration number [[VOSSELMAN AND MAAS, 2010](#)]:

$$t = \frac{\log(1 - z)}{\log(1 - (1 - e)^n)} \quad (4.2)$$

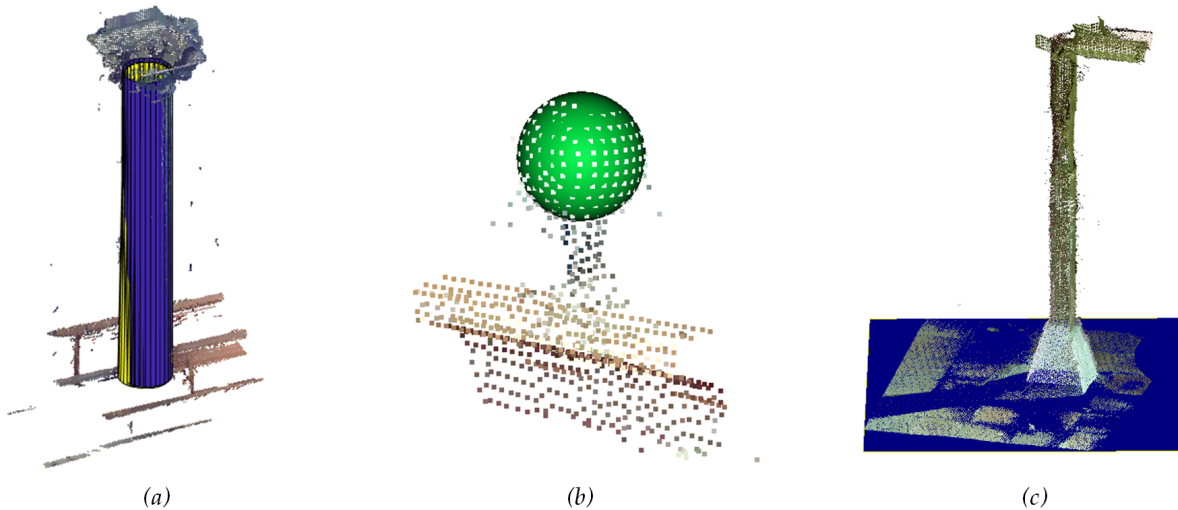


Figure 4.3: Examples of the use of RANSAC in detecting 3D geometric objects: (a) detection of a cylinder on a pillar, (b) detection of TLS spheres, and (c) detection of the floor component on a point cloud.

Where:

$t$  : number of iterations

$z$  : success probability

$e$  : probability that a point is an outlier

$n$  : number of points required for the minimal set

In this way, the number of iterations can be limited and the computation time greatly optimised. However, another problem related to RANSAC is the choice of the tolerance between the model and the inliers. A tolerance threshold which is too small may be too restrictive and thus does not include correct inliers. On the other hand, a large tolerance threshold may include noises in the solution. Some modifications to the classical RANSAC paradigm can be found to address the problem of a large tolerance, notably in [TORR AND ZISSERMAN \[1998\]](#) who proposed two modifications: M-SAC (M-estimator Sample Consensus) and MLESAC (Maximum Likelihood Estimate Sample Consensus) which maximises the likelihood that the data were part of the computed geometric model (i.e. addition of another parameter of scoring other than the number of inliers) [[MACHER ET AL., 2016](#)]. The M-SAC algorithm is employed by the model fitting functions of Matlab<sup>®</sup> Computer Vision Toolbox.

In M\_HERACLES, the RANSAC paradigm is used in many instances. The most frequent use is to detect planar surfaces (e.g. floors) and segment and/or suppress them accordingly. Basic cylinder fitting was used to generate a crude 3D primitive for pillars in section 6.2, while the implementation of RANSAC for 3D cuboids was used to generate beams in section 6.3.

### 3D region growing

Region growing was originally a method developed for 2D image segmentation. The algorithm takes a "seed" pixel and clusters nearby pixels into the same object when

they follow a set criterion (e.g. similarity of pixel colour or grey scale value). This is a useful approach to image segmentation which enables an efficient data clustering since theoretically it can group similar objects and may be implemented with more than one criterion.

The concept of region growing may be expanded into the 3D point cloud. The seed in this case involves a point in the said point cloud. Moreover, in the 3D space a host of other criteria also becomes possible to be enforced. Instead of relying solely on RGB values of each point, the region growing criterion can be based on the similarity of geometrical features such as point normals, curvature, planarity, etc [WEINMANN, 2016]. In this case, these geometrical features are often computed locally; for example the notion of "point normal" indicates the directional vector of the point towards the normal of the local surface constituting of the point itself and a determined number of its neighbours. Indeed, in the case of 3D point cloud, RGB values may not be a feasible criterion since depending on the point cloud source they may not be enough to properly represent the geometry of the object and thus its understanding.

Several types of region growing can be found and classed according to the similarity criterion it applied as well as the way it manages the 3D data (cf. Figure 4.4):

- *Euclidean distance*: this approach is quite straightforward as it takes the 3D distance between each point as the criterion for the region growing. After a random seed point is chosen, it checks for its nearest neighbours and thereafter includes those located within the radius of the threshold in the same region. When no other neighbours can be found, the algorithm reiterates by taking a new seed point from the remaining unsegmented point cloud. While straightforward to implement, this method may suffer from very long processing time especially when dealing with large point clouds. This approach is implemented in the Matlab<sup>©</sup> 2018a Computer Vision Toolbox which was used for the development of M\_HERACLES.
- *Voxel adjacency*: this type of region growing tries to address the problem of long processing time in distance-based region growing by subdividing the point cloud into voxels. The manner with which the voxel subdivision is carried out is also varied, e.g. kd-tree or octree. After subdivision, the seed takes the form of one voxel; neighbouring voxels are subsequently categorised as the same region. Finally, a winner-takes-all approach is used to include all the points in the voxels of the detected region in the new cluster. The subdivision of the point cloud into voxels means that the method is much faster than its greedy counterpart, but the same reason may cause the identified regions to have jagged edges. This approach was implemented in MACHER ET AL. [2017] as well as the label connected components module in CloudCompare.
- *Smoothness constraint*: when using adjacency-based constraints, whether greedy distance-based or voxel-based, the algorithm only takes into account the affinity of the seeds as regards to their neighbours. While this is good for data clustering, it does not permit the distinction between different faces with different normal directions; something which is useful when dealing with indoor scenes. The region growing algorithm may therefore use instead the normal and curvature values as its criteria. In this approach, for each seed the angles between its normal vector and that of its neighbours are computed. Points are regarded as belonging to the same region when the normal angles as well as the difference of the curvature are below

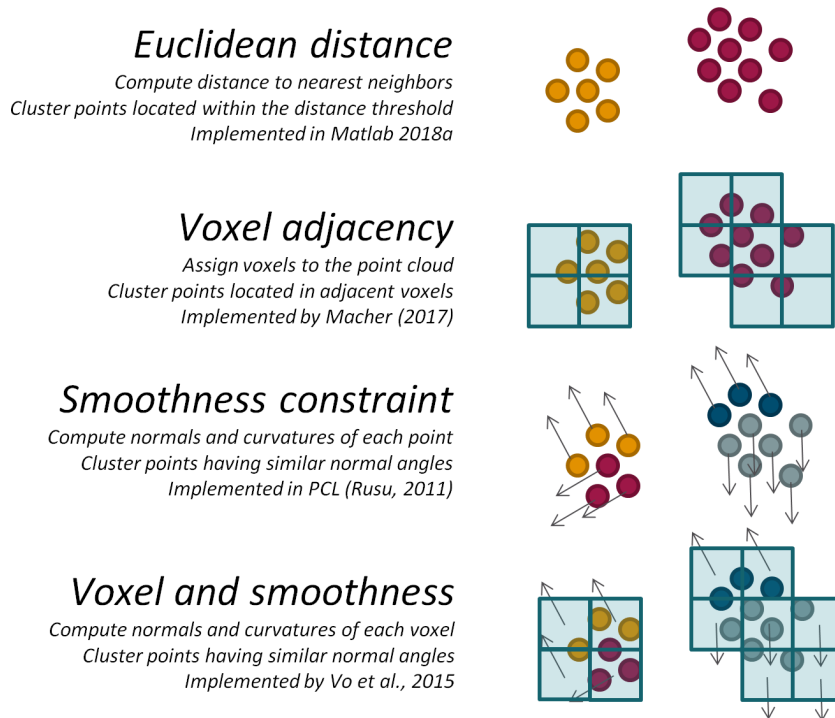


Figure 4.4: Different approaches to the 3D region growing problem. *M\_HERACLES* implemented the voxel based with smoothness constraint region growing.

the set threshold. This approach was implemented by [RUSU AND COUSINS \[2011\]](#) for the PCL library.

- *Voxel and smoothness*: similar to the problem with greedy distance-based region growing, greedy smoothness-constrained region growing may take a long processing time when dealing with large point clouds. To accelerate this process, the voxel subdivision can also be applied. In this approach, an average or median normal and curvature are computed for each voxel from the values of its constituting points. These values are then used to perform smoothness-constrained region growing. Finally, a winner-takes-all approach is also applied to expand the regions from mere voxels into points. The disadvantage of the approach is the fact that the data are subdivided into cubes of voxels; therefore rendering region edges as jagged. This approach was implemented in [VO ET AL. \[2015\]](#).

Within *M\_HERACLES*, the region growing method is also widely used. The Euclidean distance-based method as implemented in Matlab<sup>©</sup> was used in both Chapters 5 and 6 to isolate point cloud "islands". This is particularly useful when a segmented point cloud constitutes an unwanted noise, but also in order to segment free-standing objects individually. The smoothness-constrained method, both the greedy and voxel-based versions, are implemented in *M\_HERACLES* as functions. The smoothness and voxel based method is particularly used in section 6.3 to detect individual beam facets.

### 4.3 Supporting datasets

The research described in this paper utilises several datasets which are mainly heritage sites (cf. section 1.7). As has been previously described, the main datasets both involves

## SUPPORTING DATASETS

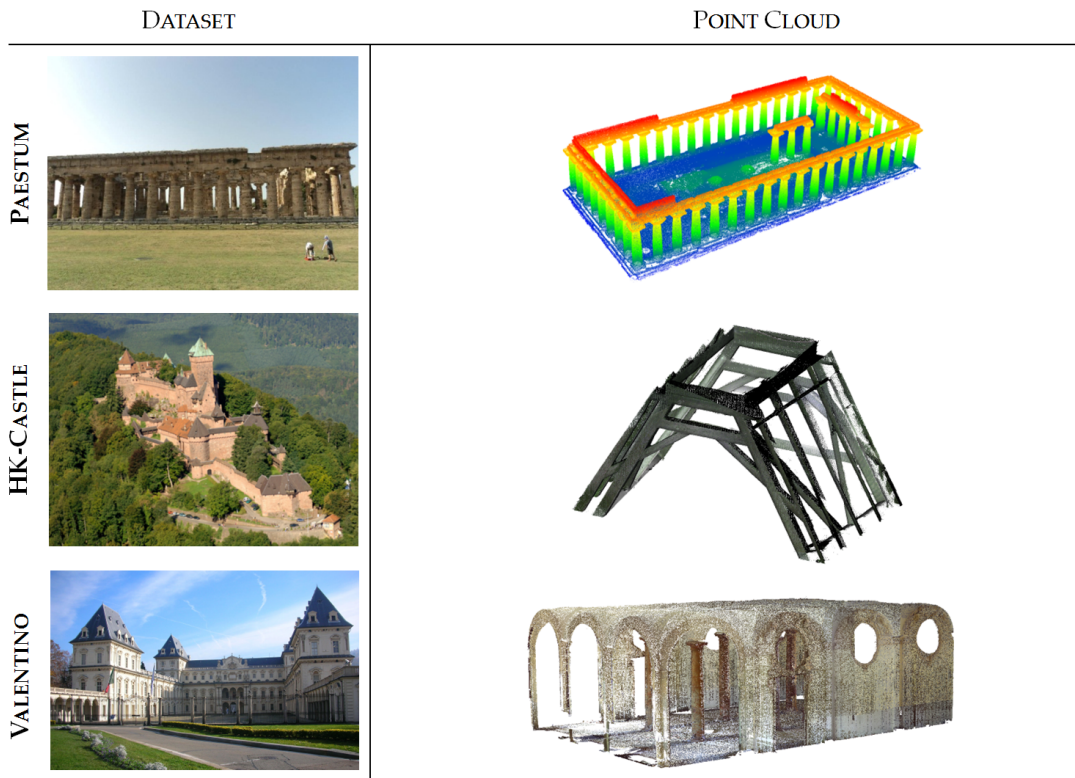


Figure 4.5: The three supporting datasets used in tests for *M\_HERACLES* in its different scale steps and classes.

multi-sensor and multi-scalar data. The multi-sensor aspect is due to the fact that the final point cloud is a result of the combination of several 3D sensors, in most cases photogrammetry (aerial and close range drone as well as close range terrestrial photos) and laser scanning (terrestrial, but also aerial LIDAR in the case of the St-Pierre dataset).

In addition to the main datasets, three supporting datasets were also used to augment the research and serve as an objective experiment on the developed algorithm's performance (Figure 4.5). These datasets are generally specific in nature, and does not possess the multi-sensor and multi-scalar attribute of the two main datasets. They are, however, useful in order to give another perspective and test the capabilities of the algorithm.

The supporting datasets are as follows:

- *Temple of Hera, Paestum, Italy ("Paestum")*: Paestum is a complex of Greco-Roman ruins located in the south of Italy and dating back to the 7<sup>th</sup> century BCE. It was inscribed as UNESCO world heritage in 1998 [FIORILLO ET AL., 2013]. The site was scanned using a combination of drone and TLS in 2013. The Temple of Hera, whose point cloud is used in this research, is one of the three remaining Doric temples in Paestum. The vastly different architectural style of this dataset compared to the mains presents an interesting contrast for the developed algorithm. The point cloud of the Temple of Hera has been graciously shared by the FBK Trento team for our experimental use. The Paestum dataset is used exclusively for the pillar detection part of the research.
- *Haut-Koenigsbourg Castle, Orschwiller, France ("HK-Castle")*: The Haut-Koenigs-

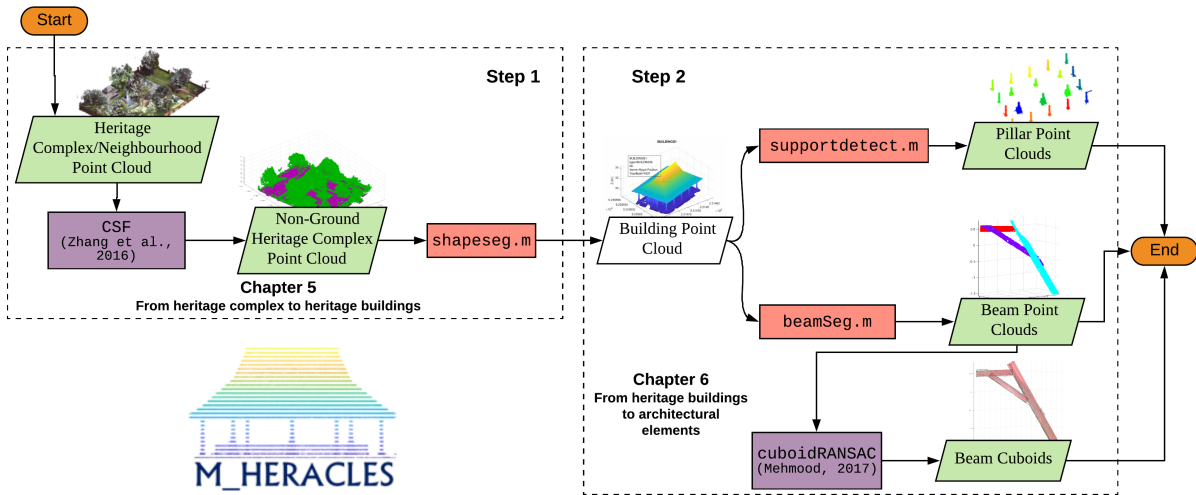


Figure 4.6: A general flowchart of the *M\_HERACLES* workflow. The first step consists of segmentation from the scale level of a complex/neighbourhood to that of a building, while the second step involves segmentation from a building's scale level to that of architectural elements. Violet rectangles denote the use of third party libraries, while red rectangles are the main functions developed for the tasks.

bourg is a medieval castle (dated to at least the 12<sup>th</sup> century) located in the Alsace region of France. Badly ruined during the Thirty Years' War, it underwent a massive, if somewhat controversial, restoration from 1900 to 1908. The resulting reconstruction shows the romantic and nationalistic ideas of the German empire at the time, the sponsors of the restoration. The castle has been listed as a historical monument by the French Ministry of Culture since 1993. In this research, only a part of the timber beam frame structure of the castle scanned using a TLS was used to perform tests on the beam detection algorithm. The beams are mostly oblique and distributed in the 3D space. The beams are of very regular shape and relatively unbroken [YANG ET AL., 2017]. The HK-Castle dataset is used exclusively for the beam detection part of the research.

- *Valentino Castle, Turin, Italy ("Valentino")*: The Castle of Valentino is a 17<sup>th</sup> century edifice located in the city of Turin, Italy. It was used as the royal residence of the House of Savoy and was inscribed into the UNESCO World Heritage list in 1997. Today the building is used by the architecture department of the Polytechnic University of Turin. The particular "Sala delle Colonne" or Room of Columns inside the castle was used in this study. This point cloud has been graciously shared by the Turin Polytechnic team for our experimental use. The Valentino dataset is used exclusively for the pillar detection part of the research.

## 4.4 Summary

This chapter attempts to describe the background to the development of the toolbox *M\_HERACLES*. As the 3D processing workflow (Figure 4.1) became more and more established, several bottlenecks have been identified which may benefit from an automation process. This mainly involves the point cloud segmentation and labelling phase, as well

as the 3D modelling. The potential complexity of heritage buildings has also been established, rendering the 3D pipeline more difficult to automate. In this chapter the following questions were addressed:

- **Automation methods for point cloud classification:** Several approaches to automation were discussed in this chapter, including the machine/deep learning and the algorithmic approach. This thesis was a continuation of previous work on scan-to-BIM automation, and thus puts an emphasis on the algorithmic approach. This was also done with another objective in mind; namely to help the segmentation and labelling process of training data for future machine/deep learning applications.
- **Novel approach to multi-scalar classification:** As can also be seen in section 4.2, many studies focused on a single scale step; i.e. either the larger scale of a building or the smaller scale of an area. *M\_HERACLES* attempts to present a continuous solution for multi-scalar projects, enabling the progressive segmentation of data from the scale of neighbourhoods up to that of architectural elements. This is done within one single toolbox, as the functions were created in modules. The modular nature of the toolbox accounts for its flexibility in use, depending on the encountered case and scale step.
- **How *M\_HERACLES* manages the multi-scalar question:** Figure 4.6 showcases the general pipeline of *M\_HERACLES*. Indeed, within *M\_HERACLES*, the first step (**Step 1**) involves point cloud processing from complexes to buildings aided by 2D GIS data. This will be detailed further in Chapter 5. Further, the second step (**Step 2**) will process the resulting building point cloud from Step 1 as well as the supporting datasets (section 4.3) into architectural elements. At the moment of writing of this thesis, two classes are available: pillars (structural supports) and beams (building frameworks). The algorithms concerning this part of the thesis will be described in Chapter 6.

All datasets were processed using functions available in the *M\_HERACLES* toolbox; two of the three main datasets (Kasepuhan and St-Pierre) were processed firstly on the heritage complex to building scale level (Chapter 5) before processed further on the building to architectural element scale level (Chapter 6). More specifically, Kasepuhan, St-Pierre, and the supporting datasets Paestum and Valentino were tested for pillar detection (section 6.2), while the HK-Castle dataset was used to test the beam detection function (section 6.2) in *M\_HERACLES*. All datasets were processed using an Intel<sup>(R)</sup> Xeon<sup>(R)</sup> E5645 2.4 GHz CPU.





# Chapter 5

## From Heritage Complex to Heritage Buildings

### Contents

---

5.1 Rationale . . . . .	80
5.2 Similar work . . . . .	80
5.3 GIS-aided segmentation and annotation . . . . .	81
5.4 Experimental results . . . . .	83
5.5 Comparison with a commercial solution . . . . .	89
5.6 Summary . . . . .	91

---

In the documentation of heritage buildings, it is often necessary to contextualise the site on its geographical location. This is useful in order to give insight on other parts of the historical building, including its socio-economical and original planning contexts. In order to do so, it is not uncommon to extend the 3D documentation of buildings into that of its surroundings. This neighbourhood-scale data present different levels of complexities and details than the buildings of interest itself, and a straight-forward segmentation approach may not work due to the vast differences of the scenes. In this regard, the **Step 1** of **M\_HERACLES** involves the processing of point clouds from the neighbourhood scale to that of a building. In order to further help the semantic classification of the point cloud, an automatic semantic annotation approach will also be presented in this chapter. In order to do both the geometric segmentation and the semantic annotation, 2D GIS files were used to facilitate the process. The main reasoning behind this is because in many heritage sites, 2D GIS systems are often already in place. We used these 2D data to help in the segmentation of a 3D point cloud, with the added benefit of automatic extraction and annotation of the related semantic information directly to the segmented clusters. In addition to the GIS data, several point cloud processing algorithms were also employed in this phase to take into account several scene complexities such as vertical overstacking of objects. Both the Kasepuhan and St-Pierre main datasets were used in the experiments in this chapter, as the neighbourhood-scale point clouds were available for these two datasets. The readers should note that the majority of this chapter has been previously published in [MURTIYOSO AND GRUSSENMEYER \[2019b\]](#).

## 5.1 Rationale

The segmentation of unorganised point cloud data is a much studied research theme, which stems partly from the needs of the remote sensing community to classify aerial LIDAR data. The developments in photogrammetry and particularly dense matching meant that 3D aerial point cloud can now be obtained for larger scale objects via the use of UAVs [CHIABRANDO ET AL., 2015; MURTIYOSO ET AL., 2016]. For heritage sites in the scale level of complexes with several objects spread out in a small area, documentation is often performed using a combination of aerial and terrestrial techniques [GRENZDÖRFFER ET AL., 2015; MURTIYOSO ET AL., 2018b]. This provides a better level of point cloud resolution than simple aerial data as well as covering more difficult angles. While this gives a winning solution to record completely a heritage site, it also adds to the complexity of the segmentation. This further increases the need for the automation of point cloud segmentation.

Furthermore, while many in the literature focuses on the segmentation and classification of aerial point cloud into certain generic classes (e.g. ground, buildings, vegetation, etc.), the field of heritage documentation often requires various different semantic attributes (e.g. historical information, architectural styles, etc.) to be stored [DRAP ET AL., 2017]. In many occasions, these semantic attributes are often already stored in the form of 2D Geographical Information Systems (GIS). The idea behind this research is therefore to use these pre-existing GIS data in helping the point cloud segmentation process, while in the same time performing the annotation of the attributes of each GIS layer field onto the segmented point cloud. This will facilitate further processes down the workflow pipeline, such as the creation of 3D GIS and HBIM.

## 5.2 Similar work

Within the spatial scale of a heritage complex site, the use of (2D) GIS for site management is also common place, as is indicated in several publications [FABBRI ET AL., 2017; FLETCHER ET AL., 2007; SEKER ET AL., 2010]. GIS enables the annotation of semantic aspects such as historical, social, and cultural information into the geometric data [FLETCHER ET AL., 2007]. This in turn enables GIS to become a powerful tool for heritage site management, where spatial analysis is made possible. One of the widely used format for GIS data is the ESRI shapefile (.shp) format [KASTUARI ET AL., 2016].

The most natural approach when addressing geospatial data of a heritage complex site is the use of aerial data. As regards to the segmentation of aerial-based point cloud, the most basic segmentation involves the extraction of ground and non-ground points. This has been done, amongst others, by creating an approximate surface of the DEM (Digital Elevation Model) [ZHANG ET AL., 2016] or by the use of deep learning technique in order to teach the algorithm to recognise ground characteristics [RIZALDY ET AL., 2018].

To further segment the non-ground points, a region-growing segmentation may be implemented [OMIDALIZARANDI AND SAADATSERESHT, 2013]. Another study performed normal analysis or tensor voting to determine man-made objects in an aerial point cloud [KIM AND MEDIONI, 2011]. These approaches have some limitations, especially when dealing with an object the scale of a heritage complex site with high resolution point cloud from heterogeneous sources. A naïve region-growing method when applied to such datasets would require enormous resources, while a normal analysis may encounter problems with

heterogeneous objects. Another approach [LIU ET AL., 2018] suggested using 2D vectorial GIS data to aid the segmentation, but lacks the semantic annotation part and was performed on a projected 2.5D point cloud. A similar approach was also used in segmenting 2D aerial images [KAISER ET AL., 2017].

The region-growing algorithm may also be used to perform segmentation and classification in larger scale objects [BASSIER ET AL., 2017a; SPINA ET AL., 2011]. Other approaches may also use machine learning in the image space [GRILLI ET AL., 2018] or 3D object space [BASSIER ET AL., 2017b; MALINVERNI ET AL., 2019]. After the classification step is completed, semantic information still needs to be annotated for each segmented part in order to create a semantically rich point cloud [POUX ET AL., 2017].

The objective of this study is to benefit from the often pre-existing 2D GIS data in the heritage documentation domain in aiding the segmentation, classification, and semantic annotation of 3D point cloud simultaneously. The existing shapefiles' geometric data will be used to guide the segmentation, while annotation of the semantic data from the shapefile attribute will be performed concurrently.

### 5.3 GIS-aided segmentation and annotation

The proposed approach used 2D GIS layers in the form of the commonly used ESRI shapefiles. These shapefiles contain 2D overhead vectorial representations of objects within the heritage complex, each with its own semantic attributes stored inside the file's database part. In some cases, these shapefiles may already exist. In other cases where a 2D GIS of the site has not been developed, the shapefiles can be obtained from simple digitisation. This may be performed based on existing topographical maps, online maps, or from orthophotos generated by photogrammetry. The 3D point cloud may be obtained using any kind of tools; including laser scanning, photogrammetry, or a combination of these techniques. The point cloud should be georeferenced to the same system as the GIS, which often times translates into the common national projection coordinate system. This procedure is well integrated in the classical heritage documentation workflow. However, in the case of the impossibility of an absolute georeferencing, a preliminary 3D transformation may also be performed beforehand.

The first step in the workflow involves the extraction of the ground. This approach used the Cloth Simulation Filtering (CSF) method [ZHANG ET AL., 2016] to extract the ground. Each polygonal vector in the shapefile was then used to segment the non-ground point cloud, using a "cookie-cutter" approach. Algorithm 1 displays the pseudocode of the proposed segmentation algorithm used at the aftermath of the ground extraction process, as written in the function `shapeseq.m`. Each shapefile effectively represents one object class. Since both the GIS and the point cloud are already in the same coordinate system, this would effectively create a bounding box in the form of the shapefile vector for the object in the point cloud, from which a segmented portion is extracted. A buffer area threshold was introduced to the 2D vectors in order to provide a tolerance with regards to digitising or georeferencing precision. This ensures that the algorithm will still be able to extract the object even if the digitising was not very precise. We call this type of segmentation the "cookie-cutter" method due to its resemblance to the functionality of said cooking apparatus (see Figure 5.1 for a visual illustration).

Since the algorithm uses this cookie-cutter approach, all points of all altitude values within the bounding box were therefore segmented. This may cause some problems when some objects overlap or are stacked with each other. For example, the existence of a

## ALGORITHM 1 - Semantic segmentation of heritage complexes aided by GIS data

---

```

function [Struct] = shapeseq(PtCloud,shpFile);
for  $i \leftarrow 1$  to numObjects do
    CookieCutter  $\leftarrow$  shpFile.geometry ;           // create a 2.5D cookie cutter
    index  $\leftarrow$  isinterior(CookieCutter,PtCloud);
    if index = 'TRUE' then
        | PtCloudIn  $\leftarrow$  PtCloud(index,:);       // check if any points is inside
    else
        | PtCloudOut  $\leftarrow$  PtCloud(index,:);
    end
    PtCloudIn2  $\leftarrow$  max(pcasegdist(PtCloudIn)); // region growing to refine the result
    Struct(i).PtCloud  $\leftarrow$  PtCloudIn2 ;         // store the segmented point cloud
    for  $l \leftarrow 1$  to nbAttributes do
        | Struct(i).(attributeList(l))  $\leftarrow$  shpFile.(attributeList(l)) ; // get attributes
    end
end
end of function

```

---

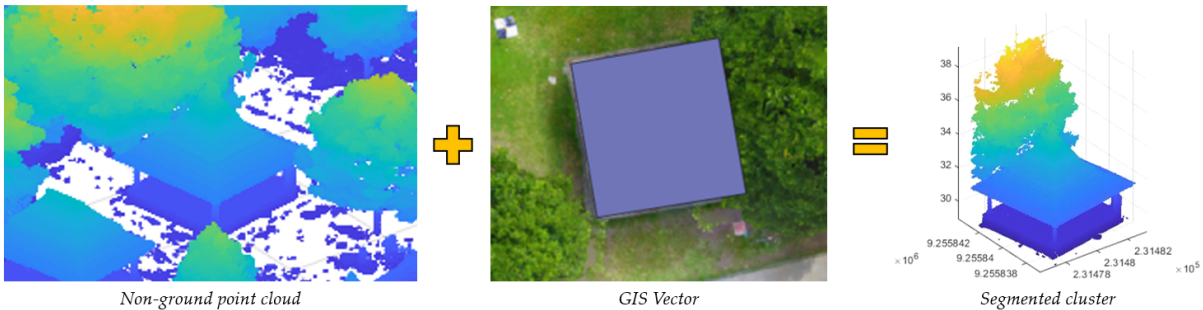


Figure 5.1: Illustration of the cookie-cutter segmentation. Notice that in the result all points of all altitude values within the confines of the GIS vector are included in the cluster. A region growing method was subsequently employed to acquire the desired object.

tree crown above a building would mean that the segmented point cloud also includes a part of the tree crown. This problem was addressed by performing another consequent segmentation algorithm on the previously segmented point cloud.

A distance-based region growing algorithm was implemented in this case to extract the point cloud clusters. A filtering was then performed to exclude any remaining noise, and the largest cluster was taken as the most probable object of interest. The excluded points were, however, not deleted. Rather, they were re-merged with the remaining unclassified point cloud, to be used in the next iteration of the algorithm. In this regard, overlapped or stacked point clouds may be extracted individually without losing parts of each object's data. The next iteration will then use the remaining point cloud to perform the computation, thereby reducing the size of the processed data with each iteration count. Finally, as the 2D vector geometry was taken from the shapefile, the associated attribute fields may be annotated directly to the segmented point clouds. All the attributes of the 3D object as recorded in the 2D GIS are thus stored automatically.

The function will therefore generate a structure for each class which consists of the different objects (records or instances in the shapefile database). These objects themselves are also structures in which the segmented point cloud and attributes are stored. The function also gives as an output the remaining unsegmented data, which can then be used as input for the further segmentation of other shapefile classes by reiterating the function.

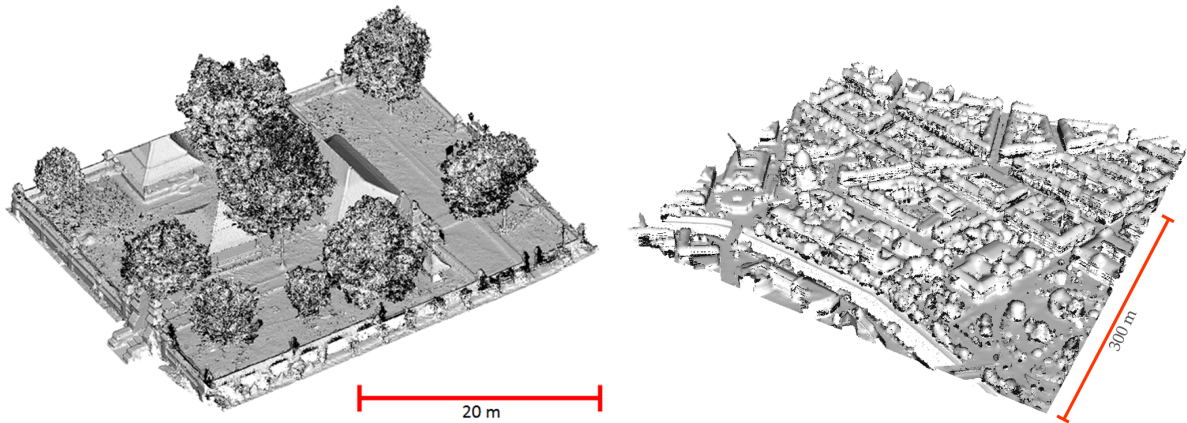


Figure 5.2: The unorganised and unclassified point cloud of Kasepuhan (left) and St-Pierre (right) used as inputs for the algorithm.

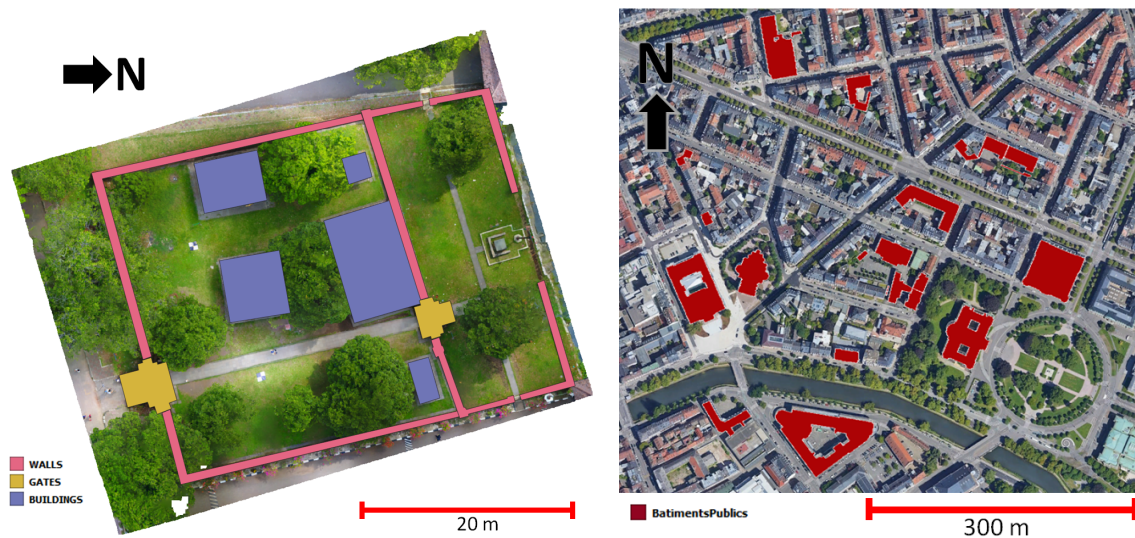
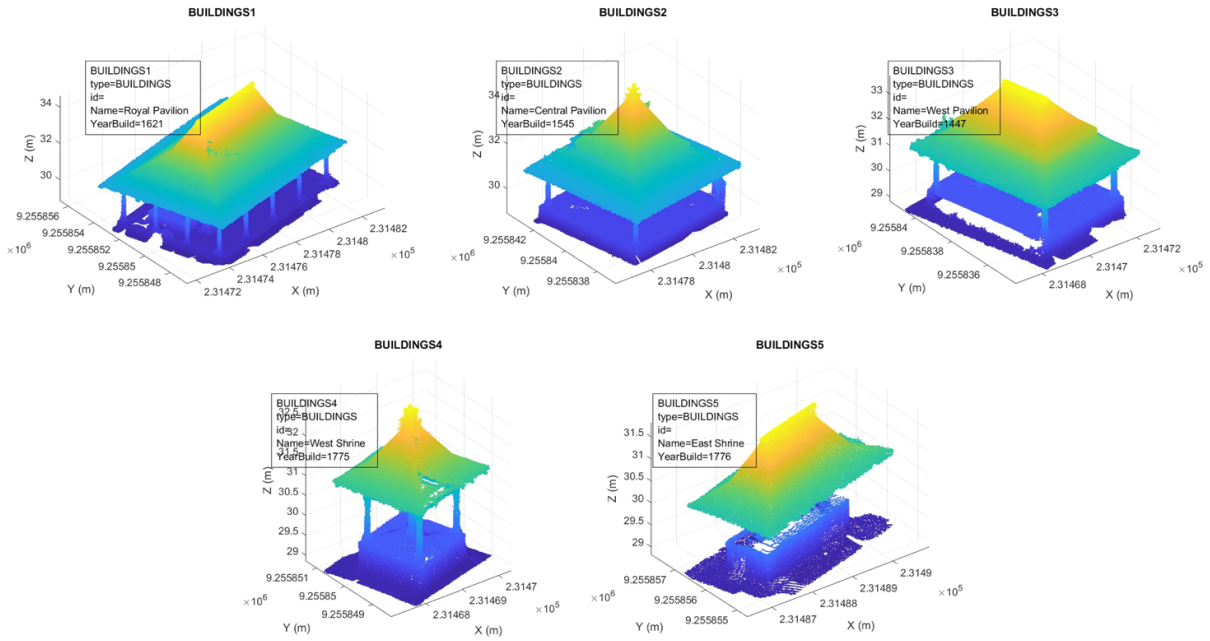


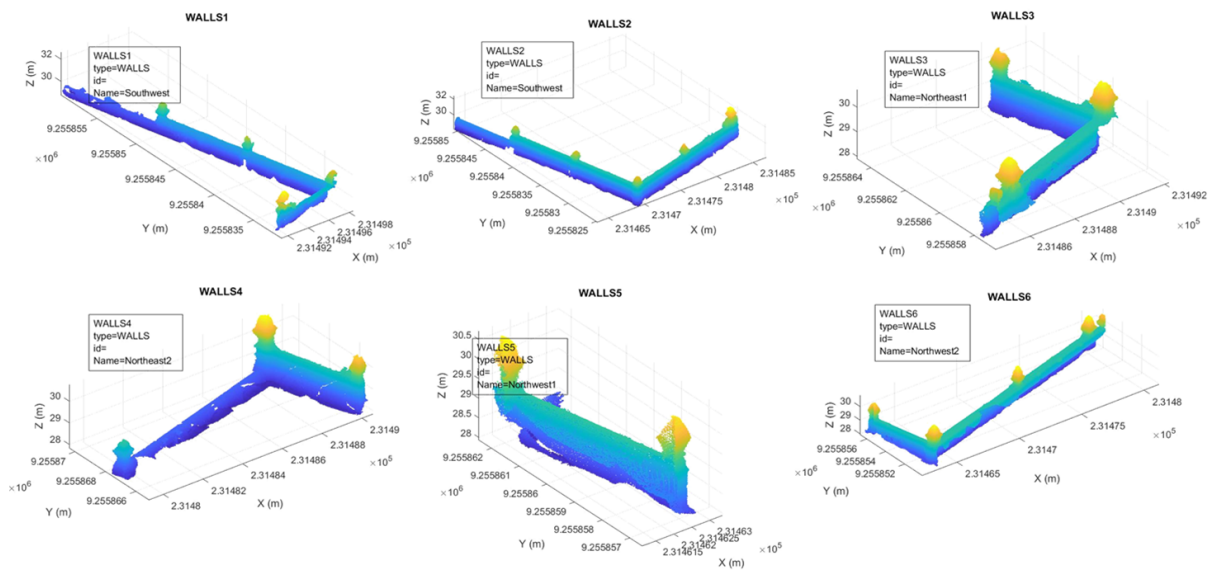
Figure 5.3: The GIS shapefile data used to help the segmentation process. To the left, three shapefiles were available for the Kasepuhan dataset while to the right, only one shapefile entity was used for the St-Pierre dataset.

## 5.4 Experimental results

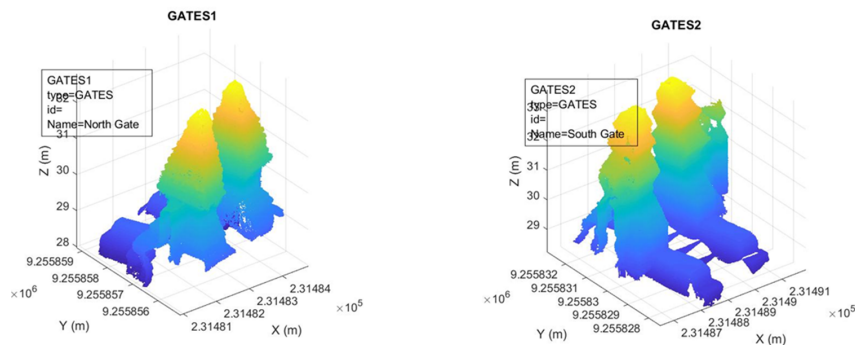
The inputs for the experiment were unorganised point clouds of a heritage complex. For Kasepuhan, this is represented by the point cloud of the Siti Inggil area while for St-Pierre a tile of the aerial LIDAR data over the Neustadt area of Strasbourg was used (Figure 5.2). The GIS shapefile data as shown in Figure 5.3 were used to aid the segmentation process. In the case of the St-Pierre dataset, the shapefile was acquired through the open data framework of the Strasbourg municipal council, the "*Référentiel topographiques simplifié*" (RTS) or simplified topographic reference. The RTS shapefile data consist of several classes, but for the purposes of this study, only the "public building" class will be addressed. For the Kasepuhan dataset, however, no prior shapefile was available for the site. The shapefiles of several object classes were therefore generated via digitisation of the orthophoto of the site, which was also made available during the acquisition mission. The digitisation was made with low precision in order to test the robustness of the function.



(a) BUILDINGS



(b) WALLS



(c) GATES

Figure 5.4: Results of the automatic segmentation and annotation for the Kasepuhan dataset: the BUILDINGS class (a), the WALLS class (b) and the GATES class (c).

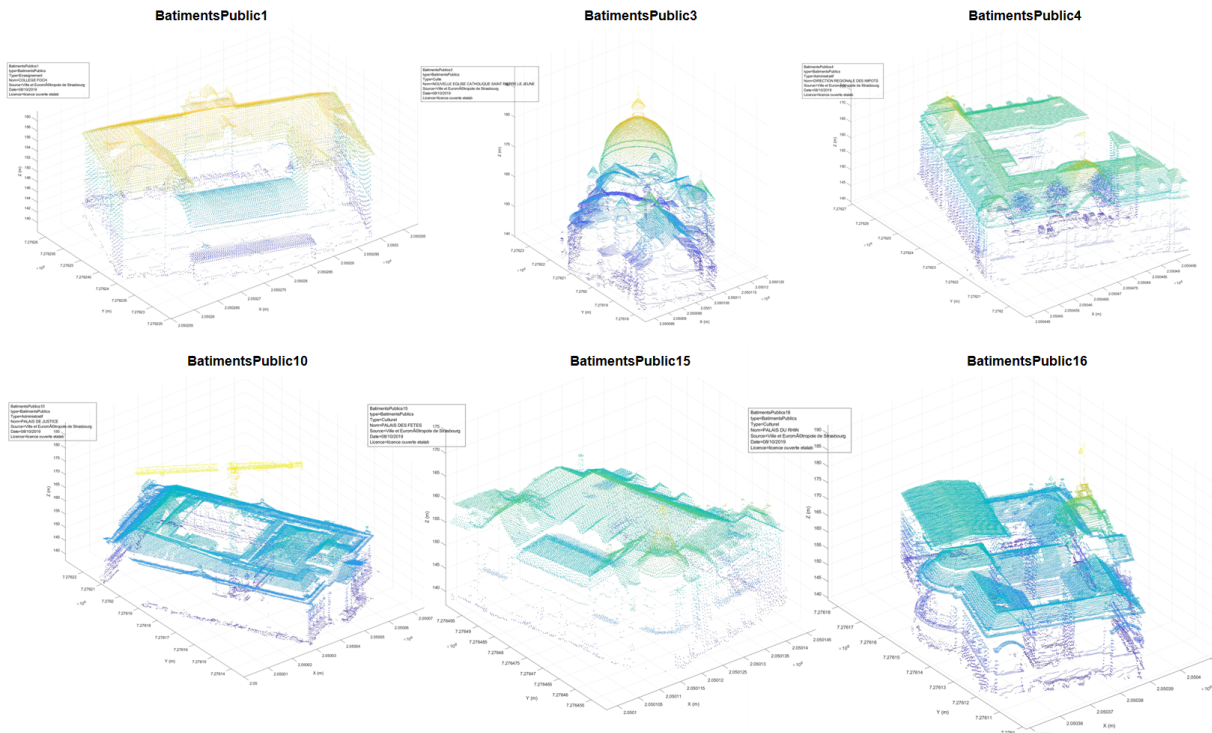


Figure 5.5: Results of the automatic segmentation and annotation for the St-Pierre dataset.

Here only six of the most important buildings are shown, of which *BatimentsPublic3* represents the St-Pierre church which will be further processed in Chapter 6. Note also that only the aerial LIDAR data are shown as opposed to the denser and more detailed drone and TLS data in Figure 5.4.

Figure 5.4 shows the visual results of the segmentation and annotation process for the Kasepuhan. The original Kasepuhan dataset consisted of 10.4 million points, and were segmented into four classes (buildings, walls, gates and the ground) and 13 different annotated objects in about 10 minutes. In the case of the St-Pierre dataset, the algorithm was visually more successful in segmenting the LIDAR point cloud into the public building class and annotating them. The St-Pierre dataset consisted of 5.9 million points and was segmented into 17 objects of one class in about 7 minutes. Here again, the visually higher success of the St-Pierre dataset may be due to the better CSF results, thus giving a cleaner result than Kasepuhan. Figure 5.5 shows the results for six of the most important and thus interesting heritage buildings within the class.

Since Kasepuhan possessed more than one class (i.e. more than one shapefile), the `shapeseq` function was progressively employed for each shapefile while taking into account the order of which shapefile was processed first. Generally speaking, lower objects were segmented first, moving towards higher objects (see Figure 5.6). After the classification of the ground using the CSF algorithm, the low brick walls were the first to be processed. This was then followed by the gates and the buildings. Finally, the trees were segmented (however, note that tree segmentation is not the main objective in this case; hence its absence in the statistics of Table 5.1). This ensures that vertically stacked objects (e.g. walls and building roofs or building roofs and trees) are segmented correctly and thus avoids ambiguity during the region growing step. Since the remaining point cloud from the previous processing is used in the next iteration, the combination of "cookie-cutter" style of segmentation and region growing was able to properly segment stacked parts of



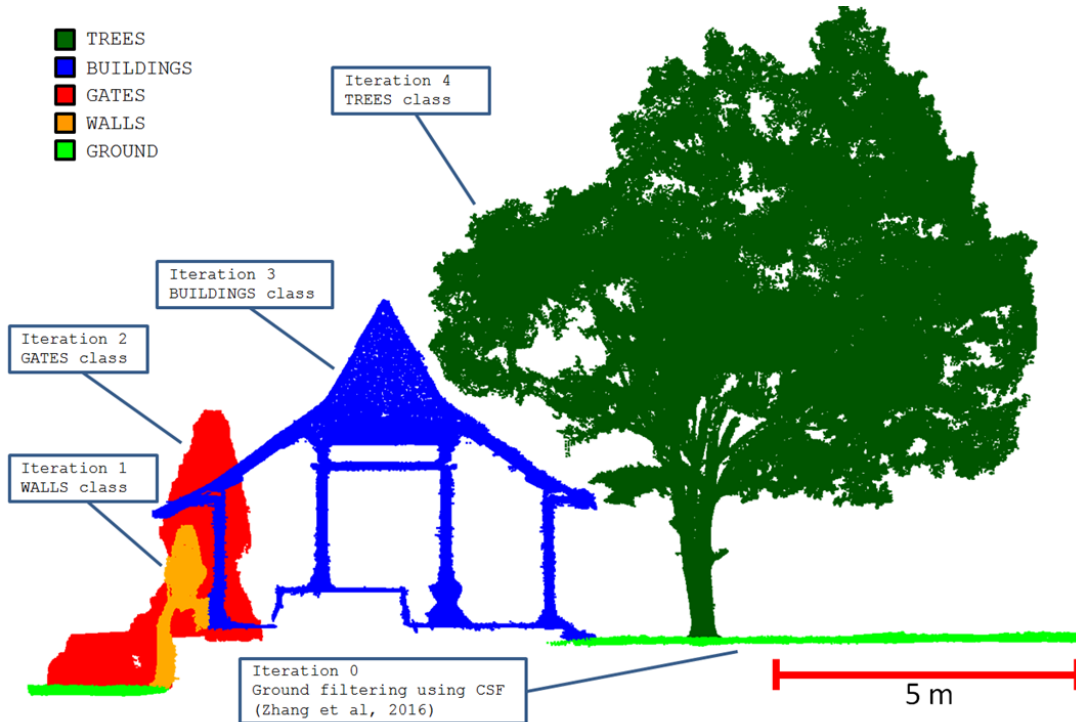


Figure 5.6: A profile of the Kasepuhan dataset to illustrate the segmented classes and the ordering of the use of the *shapeseq* function, starting from lower objects to higher ones.

the input point cloud. This also results in a faster processing because the input point cloud becomes smaller as the programme proceeds.

The results in Figure 5.4 show that the developed approach was successful in segmenting the Siti Inggil area and annotating each segmented portion with the relevant semantic information, even by using roughly digitised shapefiles (centimetric precision). The results for the first use of the algorithm on the walls class is shown in Figure 5.4b. Most of the walls were segmented correctly, except for the object WALLS5 where a large part of it was considered by the algorithm as another object altogether. This may have been due to the low resolution of the TLS on this particular part of the wall, as the critical junction that caused the segmenting failure is found behind a tree. Only two objects are present within the gates class (Figure 5.4c), and the algorithm seems to have been able to perform the segmentation correctly.

Five buildings were included in the buildings class (Figure 5.4a), all of which were segmented properly. BUILDINGS4 and 5 showed that a small portion of ground point cloud is still included in their respective clusters; this is due to the parametrising of the CSF algorithm.

In general, smaller buildings yielded a slightly worse result in the CSF ground classification, as seen here in the case of BUILDINGS4 and 5. In addition, the class ordering is shown to be important to avoid errors in stacked cases. In the case where the buildings were segmented before the lower walls, some parts of the walls were included in BUILDINGS1. Conversely, by segmenting first the lower objects as illustrated in Figure 5.6, the results in both the buildings and walls classes were correct.

Tables 5.1 and 5.2 present quantitative analysis on the obtained results for the two datasets. In Table 5.1, the number of segmented points is also used as a parameter of segmentation quality.

Object	Point Number		Misclassified		True Positive	%Unclassed	%P	%R	%F1
	Manual	Auto	Overclassified	Unclassed					
BUILDINGS1	703 500	680 386	10 592	33 706	669 794	4.79	98.44	95.21	96.80
BUILDINGS2	643 350	633 897	6 630	16 083	627 267	2.50	98.95	97.50	98.22
BUILDINGS3	317 459	300 283	9 873	27 049	290 410	8.52	96.71	91.48	94.02
BUILDINGS4	58 532	60 838	8 296	5 990	52 542	10.23	86.36	89.77	88.03
BUILDINGS5	52 026	58 047	7 415	1 394	50 632	2.68	87.23	97.32	92.00
GATES1	101 196	95 754	4 017	9 459	91 737	9.35	95.80	90.65	93.16
GATES2	151 040	146 133	4 955	9 862	141 178	6.53	96.61	93.47	95.01
WALLS1	216 951	151 520	683	66 114	150 837	30.47	99.55	69.53	81.87
WALLS2	417 768	351 818	3 168	69 118	348 650	16.54	99.10	83.46	90.61
WALLS3	84 516	81 520	5 762	8 758	75 758	10.36	92.93	89.64	91.25
WALLS4	64 877	56 804	4 595	12 668	52 209	19.53	91.91	80.47	85.81
WALLS5	63 014	34 752	1 814	30 076	32 938	47.73	94.78	52.27	67.38
WALLS6	177 399	175 862	13 371	14 908	162 491	8.40	92.40	91.60	91.99
<b>Mean</b>						<b>13.66</b>	<b>94.68</b>	<b>86.34</b>	<b>89.71</b>
<b>Median</b>						<b>6.53</b>	<b>95.80</b>	<b>90.65</b>	<b>91.99</b>

Table 5.1: Table of the quantitative analysis on the results of step 1 for Kasepuhan. In this table, three classes were taken into account (buildings, gates, and walls) with a total of 13 objects. %P is precision, %R is recall, and %F1 is the normalised F1 score.

Object	Point Number		Misclassified		True Positive	%Unclassed	%P	%R	%F1
	Manual	Auto	Overclassified	Unclassed					
COLLFOCH	34 011	32 384	217	1 844	32 167	5.69	99.33	94.58	96.90
STPIERRE	36 858	34 960	757	2 655	34 203	7.59	97.83	92.80	95.25
DIRIMPOTS	52 520	56 586	6 099	2 033	50 487	3.59	89.22	96.13	92.55
PLSJUSTICE	81 074	69 559	637	12 152	68 922	17.47	99.08	85.01	91.51
PLSFETES	37 663	35 823	0	1 840	35 823	5.14	100.00	95.11	97.50
PLSRHIN	84 833	74 738	1 026	11 121	73 712	14.88	98.63	86.89	92.39
<b>Mean</b>						<b>9.06</b>	<b>97.35</b>	<b>91.75</b>	<b>94.35</b>
<b>Median</b>						<b>6.64</b>	<b>98.86</b>	<b>93.69</b>	<b>93.90</b>

Table 5.2: Table of the quantitative analysis on the results of step 1 for St-Pierre. In this table, only one class was taken into account (public buildings) with a total of 6 out of 17 objects used in the statistical analysis. %P is precision, %R is recall, and %F1 is the normalised F1 score.

Four values were used to assess the quality of the segmentation: percentage of unclassified points (i.e. Type II error rate or false negative), precision, recall, and the F1 normalised score. These latter three values are often used in assessing the quality of point cloud classification algorithms and stems from the data science domain, while the percentage of the unclassified points was used as a general parameter to broadly assess how many points within the input point cloud were correctly classified. The values of precision, recall, and F1 score respectively are computed using the following equations:

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (5.1)$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (5.2)$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5.3)$$

Thus the precision represents the classification success rate when compared to the predicted positive, while recall denotes the success rate relative to the real or actual positive. The F1 score is the harmonic mean of precision and recall, and may also be useful to balance both values as they may have vastly different values.

On the Kasepuhan data, overall the unclassified rate from all 13 objects yielded an average value of 13.66% and a median value of 6.53%. Meanwhile, the precision median is also quite high at 95.80% with a lower median recall value of 90.65%, thus giving a median F1 score of 91.99%. While this value is seen to be good enough, the quality for each class differs. The buildings and gates class fared the best, with an average F1 score of 93.81% and 94.09% respectively, although the score for the walls class may be biased since it only consists of two objects. As far as the buildings are concerned, BUILDINGS4 presented the largest error, which is caused by the remaining unfiltered ground around the structure. The walls class presented the worst results with an average F1 score of 84.82%. The WALLS class presents poorer recall value, which may be due to the significant presence of noises. This is particularly true for WALLS1 where the presence of large flower pots rendered the point cloud very noisy. Meanwhile, WALLS5 presented a breakaway wall which was the otherwise intact in the ground truth. This is due to the difference of point cloud resolution which caused its division in two clusters where it should have been one.

Apart from the obvious reasons for outlier or noisy data, the walls class does indeed show a systematic trend in its error. The median value of its unclassified rate is 18.04%, which is quite high compared to similar values from the other classes (between 4%-8%). Several aspects can play a role in this regard. First of all, it is quite noticeable that the result of the CSF ground filtering around the walls is quite noisy. This may be due to (i) the uneven ground around the walls, owing to the fact that many tree roots and tall grass are present; and (ii) the small surface area of the walls compared to the buildings.

A similar albeit less-evident phenomenon can be seen in the buildings class, where as the object's ground surface decreases, more unfiltered ground is present. A better parametrising for the CSF algorithm should therefore be experimented to improve the results.

For the St-Pierre dataset, quantitative values validate the visual conclusion where the algorithm worked better than for Kasepuhan. In the statistical analysis presented in Table 5.2, only six of the most important heritage sites located within the neighbourhood of the St-Pierre church were taken into account. In this reduced sample, the median unclassified percentage amounts to 6.64% (i.e. comparable to that of Kasepuhan), but the median precision attained a value of 98.86% and the median recall that of 93.69%, thus yielding a median F1 score of 93.90%. The best result was obtained for the Palais des Fêtes (PLSFETES) building, with 100% precision (97.5% F1 score). The worst F1 score was obtained by the Palais de Justice (PLSJUSTICE) building. This is due to a chunk of the point cloud of the aforementioned building which was visibly not segmented into the cluster. This unsegmented chunk corresponds to the scaffoldings on the building, erected due to renovations. The LIDAR points of the scaffoldings were too few, that M\_HERACLES considered them as noise. Apart from this outlier, the most frequently encountered error

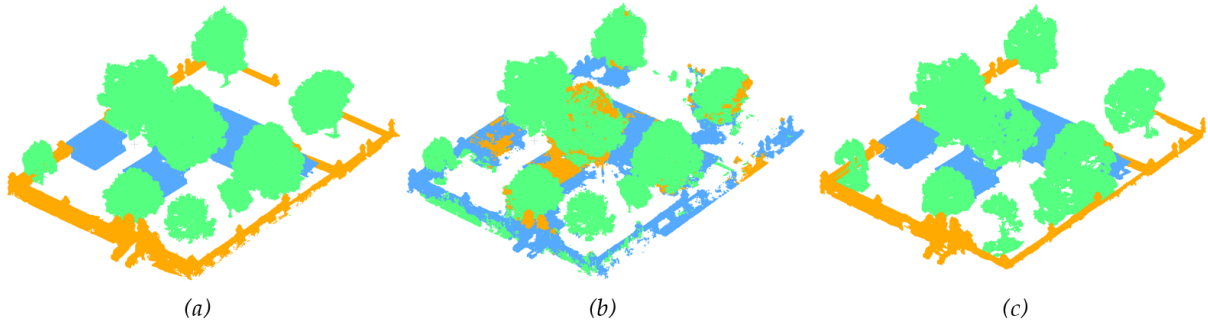


Figure 5.7: Visual comparison of the point cloud segmentation process: (a) displays the manual classification used as reference, (b) the results of the Metashape automatic classification, and (c) the results from *M\_HERACLES*. The colour blue denotes the 'buildings' class, orange the 'walls' class, and green the 'trees' class.

seems to be related to the presence of vegetation or, analogous to the Kasepuhan data, the minor errors due to prior ground extraction.

An interesting point to note in summarising these results is the speed of the processing when compared to manual segmentation and labelling. The algorithm, while having several outliers (especially in the presence of important noise), generated good results. This is particularly true in the case of the St-Pierre dataset, where the urban density and particularly flat terrain generated very good results.

## 5.5 Comparison with a commercial solution

In order to assess the quality of our developed approach, a comparison was performed with the automatic point cloud classification results of the commercial software Agisoft Metashape. While Metashape is chiefly a photogrammetric software known for its use in image-based reconstruction, it was also recently augmented with a function for multi-class point cloud semantic segmentation. According to the official documentation, Metashape employs a machine learning technique to perform this task; indeed, it asks its users to submit training datasets in order to improve the classification quality in the future. For the purposes of our comparison, Metashape version 1.5.3 build 8469 (release date 24 June 2019) was used.

The Metashape automatic classification was performed on both main datasets: Kasepuhan and St-Pierre. Visual results for the Kasepuhan can be seen in Figure 5.7. Three classes were defined, namely the 'buildings', 'walls' and 'trees' classes. In Metashape, this corresponds respectively to 'buildings', 'man-made objects', and 'high vegetation' classes. Figure 5.7 shows that the Metashape automatic classification had difficulties in distinguishing between buildings and walls, with most of the walls classified as buildings. Some parts of the walls were also classified as high vegetation. This may be due to the fact that the Kasepuhan dataset presented a large-scale and thus more complex scene not entirely suitable for the machine learning-trained function. Unfortunately there is no way to verify this hypothesis since Metashape understandably does not divulge their machine learning method in detail. On the contrary, *M\_HERACLES* managed to classify the objects fairly well thanks to the use of shapefiles to guide it. Visually, Figure 5.7c also showed that some parts were nevertheless unclassified, notably the walls at the back of the dataset. This may be due to the low resolution of the point cloud for this part of the

Class	%Precision		%Recall		%F1 Score	
	Metashape	M_HERACLES	Metashape	M_HERACLES	Metashape	M_HERACLES
Buildings	51.44	95.40	73.49	77.14	60.52	85.30
Walls	6.48	96.61	3.17	77.21	4.26	85.83
Trees	92.15	88.23	85.12	74.80	88.50	80.96
<b>Median</b>	<b>51.44</b>	<b>95.40</b>	<b>73.49</b>	<b>77.14</b>	<b>60.52</b>	<b>85.30</b>

Table 5.3: Comparative table showing the quantitative results of the classification for Kasepuhan using Metashape and M\_HERACLES.

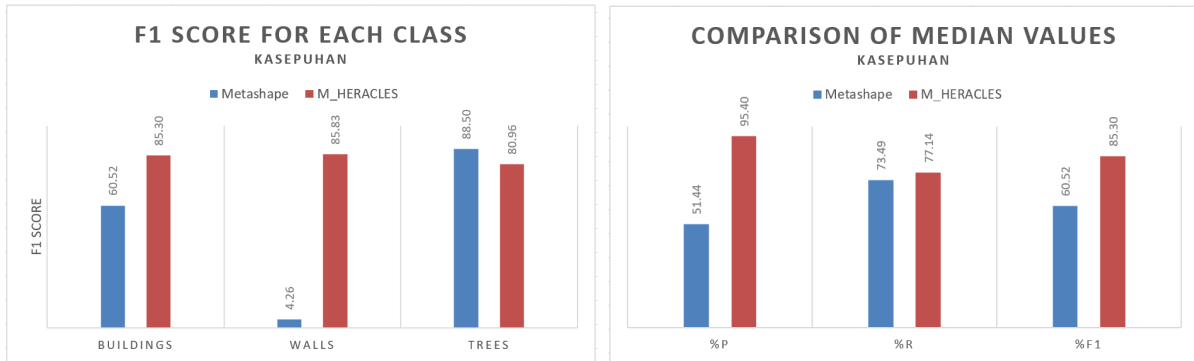


Figure 5.8: Histogram representation of the classification performance for each class in the Kasepuhan dataset (left) and the median value of the principal quality parameters (right).

site (note the same observation on Metashape results).

Table 5.3 displays a quantitative comparison of the two tested algorithms for Kasepuhan, also visualised via histograms in Figure 5.8. M\_HERACLES managed to outperform Metashape in most cases (yielding a slightly lower F1 score in the trees class), especially in the walls class. The median value of the F1 score for M\_HERACLES was 85.30% compared to Metashape's 60.52%. It showed a lower recall value and higher precision, which may be explained by the fact that the use of shapefiles disproportionately increased Type I error.

When implemented on the St-Pierre dataset, M\_HERACLES notably still performed better than Metashape as can be consulted in Figure 5.9. Metashape produced a very high precision rate; however this must be understood with a caveat. Indeed, Metashape performed automatic segmentation on all buildings on the scene, whereas M\_HERACLES only performed one on the "public building" class as dictated by the related shapefile. This distinction between public buildings and other buildings follows the official categorisation as set by the Strasbourg city geomatics service. In this regard the results of Metashape was therefore manually segmented to include only the so-called public buildings, thus yielding a slight bias towards higher precision. However, as far as the recall value is concerned, M\_HERACLES again outperformed Metashape. This is mainly due to the mis-classification rate of Metashape as can be seen in Figure 5.10. For example in Figure 5.10(a), much of the St-Pierre church dome and church towers were misclassified as high vegetation. This played a large role in explaining the low recall value for Metashape. Overall, in terms of F1 score M\_HERACLES also managed to outperform Metashape in this case of highly urban scene as opposed to Kasepuhan's with more closed and isolated complex situation.

As can be seen in this section, the proposed M\_HERACLES algorithm managed to perform the classification of point clouds in the neighbourhood scale fairly well. Comparison with Metashape also showed that our solution presents very promising and interesting

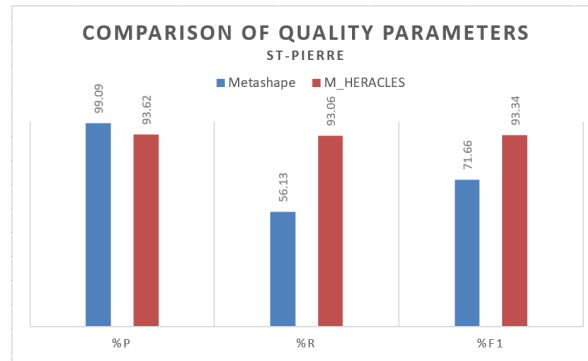


Figure 5.9: Histogram representation of the quality parameters for the St-Pierre dataset comparing Metashape against M\_HERACLES.

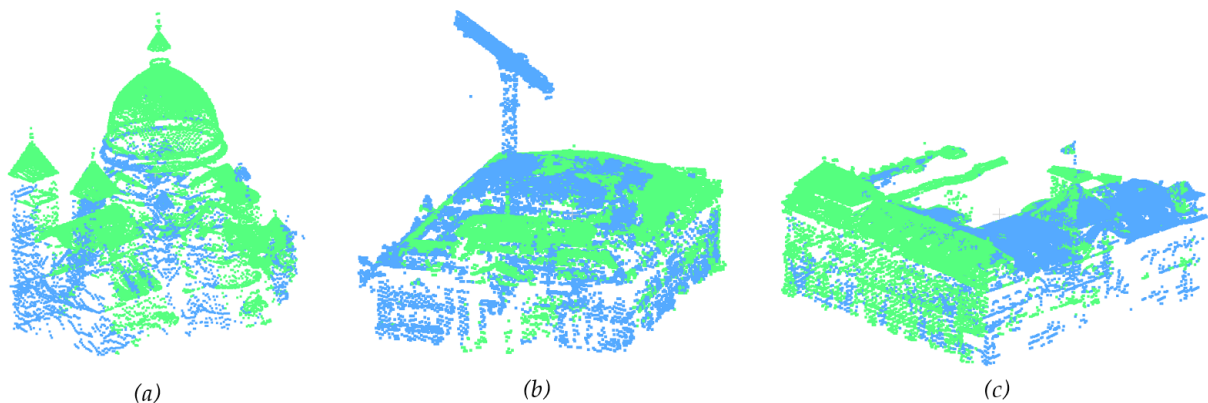


Figure 5.10: Case of mis-classification in the results of the Metashape point cloud classification on the St-Pierre dataset. Here are shown some examples where the problem is most observable:

(a) STPIERRE, (b) PLSJUSTICE and (c) DIRIMPOTS sub-clouds. The colour blue denotes the 'buildings' class and green the 'high vegetation' class. Note also the presence of the crane in (b).

results. Another advantage of M\_HERACLES is the possibility to retrieve individual objects instead of a single cluster comprising all of the instances in the same class. This is useful when working with heritage sites, since in many cases and for various reasons the user may wish to acquire the point cloud of one or more specific buildings. Furthermore the possibility to annotate these individual buildings with semantic information derived from the GIS shapefiles also presented an advantage for the developed algorithms. As far as the processing time is concerned, Metashape clocked a much faster time at around 5 minutes for both Kasepuhan and St-Pierre. This also shows the necessity for further optimisation of M\_HERACLES in terms of the processing time; although the current result is already quite satisfactory considering the results obtained.

## 5.6 Summary

In summary of this chapter, several takeaway messages can be made based on the experiments and analyses:

- **The M\_HERACLES method as a viable solution for point cloud classification:** The approach demonstrated in this chapter has shown to be able to perform well in segmenting the 3D point cloud of heritage complexes. The results also generated

better metrics in both case studies than the classification result of a commercial software. Furthermore, using this method each object was segmented individually but retains a classification according to the input shapefile. Another advantage is that the semantic attribute present in the shapefile database is not lost. This information is annotated automatically to each object cluster, therefore facilitating further work on site management in the form of 3D GIS and/or HBIM.

- **Limitations due to the algorithmic approach taken:** Some shortcomings are still present, however. The fine tuning of segmentation parameters is important in order to extract the correct objects. This is particularly true for the region growing segmentation parameters for two different objects which superposed closely. The ordering of the class segmentation is also important to avoid clustering ambiguity as can be seen in Figure 5.6. The correct iteration ordering, in this case from lower to higher object heights, was essential in segmenting the data correctly when more than one class is concerned. When the ordering is wrong, the region growing segment of the algorithm encounters ambiguities in determining which segmented cluster is part of the desired class. This remains as one of the aspects that require manual human interpretation and intervention.
- **Importance of the quality of the ground extraction process:** The ground extraction at the beginning of the algorithm is also an important factor that influences the final product. Indeed, when applying the algorithm to segment the trees class, some residual ground point cloud was still present in the end. Irregularities of the ground near the trees (presence of grasses, roots, etc.) for example may appear as the cause of the problem. Although tree extraction is not the main objective of the research, this particular point still merits a further improvement of the algorithm as vegetations can sometimes also be important, particularly for environmental analysis. Furthermore, some objects which didn't have enough points were unpurposely filtered. That being said, the attained median F1 scores of 91.99 % for Kasepuhan and 93.90% for St-Pierre are very encouraging.
- **Possible extensions of the presented idea:** Several other ideas may also be added to develop the algorithm further. For example, in the absence of shapefiles the use of CAD files may be interesting. This is also interesting for the case of the next scale step (buildings to architectural elements). Indeed, in many heritage buildings a CAD file may already exist which may serve as the same basis as the shapefile in this chapter. While this thesis does not address this path in more detail, a study was conducted in our group by [SEMLER ET AL. \[2019\]](#). Interested readers are encouraged to consult this publication.

Having performed the Step 1 of Figure 4.6, the following Chapter 6 will address the next scale step. The results from this Chapter may thereafter be used as input for the algorithms of Step 2, which will be described in the next chapter.

# Chapter 6

## From Heritage Buildings to Architectural Elements

### Contents

---

<b>6.1 Motivations of the study . . . . .</b>	<b>94</b>
<b>6.2 Detection of structural supports . . . . .</b>	<b>94</b>
<b>6.3 Detection of beams in structural frameworks . . . . .</b>	<b>97</b>
<b>6.4 Experimental results . . . . .</b>	<b>100</b>
6.4.1 Structural supports . . . . .	100
6.4.2 Beams . . . . .	108
<b>6.5 Summary . . . . .</b>	<b>109</b>

---

In this chapter, the **Step 2** of `M_HERACLES` as illustrated in Figure 4.6 will be described. This step involves the segmentation and classification of heritage building point clouds into architectural elements. It presents the next logical continuation of the previous chapter, and is one of the most important phases in 3D modelling for HBIM purposes. The classification of heritage building point clouds into clusters of smaller elements also imbues semantic information into the geometric data, thus enabling further operations and analysis to be conducted. The automation for two particular classes of architectural elements have been developed and will be detailed in this chapter, namely that of pillars and beams. Pillars or structural supports present a very interesting class from the heritage documentation perspective as they often show not only artistic taste but also the engineering know-hows of the original builders. It is therefore unsurprising to find that pillars and their decorations are one of the most diverse element in heritage buildings. Their styles and forms vary not only between different architectural types but also between individual instances. Meanwhile, the building framework show the craftsmanship of heritage buildings, crystallised in the form of beams. In `M_HERACLES`, the detection and classification of pillars from building point clouds were performed using a 2.5D approach as extended from the techniques used in Chapter 5. However the detection of beams requires a 3D approach due to the manner in which building frameworks use space. This presents two different but still related approach for the two classes available at the moment in `M_HERACLES` in keeping with the spirit of modularity.



Readers should be aware that the majority of the research that will be presented in this chapter has been previously published in [MURTIYOSO AND GRUSSENMEYER \[2019a\]](#). The paper was presented orally at the 27<sup>th</sup> CIPA Symposium in Ávila, Spain.

## 6.1 Motivations of the study

There are two often distinct parts in the pre-processing of point clouds within the 3D modelling workflow. The raw point cloud is usually segmented into smaller clusters representing certain elements of the object, and then classified into object classes (e.g. pillar, arch, floor, etc.). Unfortunately, the 3D modelling process still retains a large part which requires manual intervention [[MACHER ET AL., 2017](#)]. Attempts to automate any part of this process will greatly save both time and resources in the overall workflow.

Segmentation is an important part of the 3D modelling process, from which we may derive semantic-rich models such as 3D GIS or BIM (Building Information Model) [[CAMPANARO ET AL., 2016](#); [WANG ET AL., 2015a](#)]. Depending on the level of complexity, the process of manual segmentation may take a lot of time. Furthermore, point cloud automatic classification has also become more and more important in this regard. Classification will confer classes into the point cloud clusters resulting from the segmentation step. It is therefore an important process to be performed on the point cloud in order to give tangible information on the point cloud.

Another motivation to the development of these series of functions is the increasing popularity of machine learning techniques. Machine learning, and more precisely deep learning techniques, has seen a surge in overall interest in this age of big data. The possibility to use a large quantity of data to train the computer to do the semantic annotation automatically is indeed a very interesting concept, as it provides a robust segmentation result with a fairly quick processing time. However, the main bottleneck problem in implementing deep learning techniques is again mainly related to the availability of labelled datasets [[MAALEK ET AL., 2019](#)] which heretofore is still performed mostly manually. In the case of heritage point cloud, this problem is exacerbated by the diversity of classes and architectural features, as well as the general lack of labelled datasets, hence the prevalence of manual annotation to generate training data [[MALINVERNI ET AL., 2019](#)]. In this regard, the algorithms developed in this study may serve as a base to automate this otherwise tedious task of manual data labelling.

This chapter will describe an approach to automatically segment a historical building point cloud into architectural elements. In doing this, several geometry-based constraints as well as point cloud processing algorithms were used in order to distinguish the different building element units e.g. roofs, structural support, floors, framework, etc.

## 6.2 Detection of structural supports

Pillars or structural supports in a historical setting is often an interesting architectural element, since they showcase both the engineering know-how and the architectural taste of the builders. It is with this reason in mind that the first function was developed to segment structural supports automatically. Additionally, simple geometric rules were implemented in order to be able to identify a column from other types of structural supports. Indeed, this kind of development has been addressed before in the scope of simple pillars, often in an industrial setting [[LUO AND WANG, 2008](#); [RIVEIRO ET AL., 2016](#)].

## ALGORITHM 2 - Attic detection

---

```

function [Struct] = attigsegment(PtCloud);
Slices = slices(ptCloud) ;                               // create vertical slices
for  $i \leftarrow 1$  to numSlices do
  Area  $\leftarrow$  convhull(thisSlice) ;                    // compute the 2D area of the slice's convex hull
  AreaList(i)  $\leftarrow$  Area;
  for  $i \leftarrow 1$  to size(AreaList)-1 do
    delta=abs(diff(AreaList(i));AreaList(i+1));
    if delta>tolerance then
      | Zlimit=ThisSlice(z) ;                            // determine limit between attic and body
    end
  end
  Struct.Body  $\leftarrow$  find(ptCloud(z)<=Zlimit) ;        // lower than limit is body
  Struct.Attic  $\leftarrow$  find(ptCloud(z)>Zlimit) ;        // higher than limit is attic
end
end of function

```

---

## ALGORITHM 3 - Pillar segmentation and classification

---

```

function [Struct,SupportType] = supportdetect(Body);
Slices = slices(Body);
SliceUsed  $\leftarrow$  size(slices) ;                          // create vertical slices and take the middle one
Clusters = psegdist(SliceUsed) ;                         // region growing to isolate the clusters
for  $i \leftarrow 1$  to numClusters do
  Area = convhull (ThisSlice) ;
  circularity(i)  $\leftarrow$  (chull_perim2)/(4*pi*Area) ;    // compute the circularity
  if circularity<threshold then
    | SupportType(i)  $\leftarrow$  "column";                    // if circular, it's a pillar
  else
    | SupportType(i)  $\leftarrow$  "not column";
  end
  CookieCutter=polybuffer(chull) ;                       // create a 2.5D cookie-cutter
  index  $\leftarrow$  isinterior(CookieCutter,Body);
  if index = 'TRUE' then
    | PtCloudIn  $\leftarrow$  PtCloud(index,:) ;                // check if any points is inside
  else
    | PtCloudOut  $\leftarrow$  PtCloud(index,:);
  end
  [PtCloudIn2,PtCloudOut2]=pcfitplace(PtCloudIn) ;       // remove floor/ceiling
  PtCloudIn3  $\leftarrow$  max(psegdist(PtCloudIn)) ; // region growing to refine the result
  Struct(i).ptCloud  $\leftarrow$  PtCloudIn3 ;                 // store the result in a structure
end
end of function

```

---

In the field of heritage, pillars or supports can be very variable depending on the architectural style and geographical situation; hence making this operation more difficult. Some authors solved this problem by creating a dedicated library of parametric objects [MURPHY ET AL., 2013], while the most common solution remains a manual segmentation [ANTONOPOULOS AND ANTONOPOULOU, 2017].

In this study, geometric characteristics (also called hard-coded knowledge, as exemplified in MAALEK ET AL. [2019]) were used to help identify the class of the segmented point cloud cluster. In particular, the circular cross-section characteristic of most columns will be used as the main rule in determining if a segmented point cloud is a column or not. This approach has been used in several other studies, for example MACHER ET AL. [2017]

for the creation of as-built BIM elements or RIVEIRO ET AL. [2016] for engineering purposes. LUO AND WANG [2008] also developed a similar approach to the one presented in this paper, albeit implemented for modern columns and without semantic classification.

The proposed method employs several geometrical characteristics of typical historical buildings in performing the segmentation and classification process. The algorithm used as a starting point the point cloud of a heritage building. The first part of the developed approach was the identification and segmentation of the building's body and attic (Algorithm 2). The attic in this case is defined as the space between the roofs and ceilings of the uppermost storey. This vertical segmentation is meant to facilitate further segmentation process and point cloud management. To this end, horizontal profiles of the object were extracted and their geometric properties were used to identify the attic from its main body. The area surface of the profiles was assessed, and a significant reduction of area surface was interpreted as the limit between the attic and the main body. In this manner, the algorithm was able to quickly and quite reliably determine these two parts of the building. In the absence of a tilted roof (such as the case with some modern buildings), the algorithm will simply determine that the building has no attic.

Further segmentation was performed to detect architectural elements from the building's body. Algorithm 3 shows a pseudocode of the developed approach to specifically detect structural supports. The `supportdetect.m` function consists of two parts. The first part concerns the detection of the structural supports and successive segmentation into potential point cloud clusters. In this case, a 2D approach to a 3D problem was used to help with the process; a method similar to the one described in MACHER ET AL. [2017]. Consequently, a cross-section of the building's body (result of the previous `attigsegment.m` function described in Algorithm 2) was extracted. From the cross-section, various "islands" represent different vertical elements of the building. In order to segment these elements into individual clusters, a region growing segmentation based on Euclidean distance was performed. A preliminary filtering and immediate classification was then performed to distinguish between potential structural supports, walls, and point cloud noise. The filtering was done using the convex hull area criterion.

From the list of structural support clusters generated from this process, other geometrical rules were then used to determine if a structural support is a column or a not. While there is no single agreed definition as to the definition of a column, this study defines a column as a vertical support which mostly possesses a circular cross-section. On the other hand, a non-column is defined as any other structural support having a non-circular cross-section. This definition corresponds to the one taken from the UK-based Designing Buildings website<sup>1</sup>.

In order to distinguish between a circular and non-circular cross-section, again the convex hull is computed for each support candidate's cross-section. For each structural support candidate, the circularity parameter is computed from the convex hull parameters. This value follows the following Equation 6.1, slightly modified from TAKASHIMIZU AND IYOSHI [2016]:

$$circularity = \frac{Perimeter^2}{4\pi \cdot Area} \quad (6.1)$$

In this set-up, a circularity value of a perfect circle is 1, while as the value increases the form of the object departs from a circular form. While the circularity parameter is very easy to compute, it should be noted that it is not robust and is therefore prone to errors

<sup>1</sup>[https://www.designingbuildings.co.uk/wiki/Types\\_of\\_column](https://www.designingbuildings.co.uk/wiki/Types_of_column), retrieved 27 January 2020

due to noises which may distort the form of the cross-section's convex hull. The success of this approach depends therefore on the quality of the point cloud data acquisition process (acquisition noise, errors during registration, etc.).

In the final part of the code shown by Algorithm 3, the segmentation and classification was extended back into the 3D space. Note that until now, only the building's cross-section's clusters of islands were segmented and classified. In order to do so, a similar approach to Chapter 5 was used. In this approach, the convex hull of each support's cross-section is used as a "cookie-cutter" to obtain the 3D point cloud of all elevations corresponding to each island cluster. A buffering threshold was applied to the convex hull in order to give a tolerance value to the process. A RANSAC plane fitting was then subsequently applied to remove the horizontally planar parts (floors and/or ceilings) of the segmented result. Finally, a last Euclidean distance-based region growing segmentation was performed in order to delete remaining noises and therefore refine the segmentation. In this way, a form of automatic classification of the segmented point cloud clusters was conducted. The output of the general workflow consists of clusters of point clouds - segmented and classified into the attic and the main building body- which were further classified into columns and non-columns.

### 6.3 Detection of beams in structural frameworks

This section describes an early work on the automatic detection of beams in building frameworks. The rationale of this research path is the importance of building frames in the context of historical buildings, as they encapsulate the core of the construction knowledge and know-how of the builders [DA COSTA SALAVESSA, 2012]. The recent burning of Notre-Dame de Paris cathedral in April 2019 also emphasised importance in the documentation of the timber framework of other similar structures [MENOUE, 2019].

The automatic parametric modelling of wooden beams has been addressed in another research conducted by our group, as presented in YANG ET AL. [2017]. However, in that research, the authors relied on total station measurements to automatically create parametric models of wooden beams. The idea within this particular part of M\_HERACLES is to benefit from the availability of point cloud data, which is much faster to acquire and provides more detailed data than traditional total station measurements. Although one might even argue for over-abundance of data in point clouds, the ease of acquisition of points clouds compared to traditional surveying is undeniable. Another similar work of automatic parametric modelling of wooden beams was presented in PÖCHTRAGER ET AL. [2017] and PÖCHTRAGER ET AL. [2018]; indeed the algorithm described in this section took some inspiration from their approach.

Contrary to the developments in Chapter 5 and section 6.2 where a 2.5D approach was taken, wooden beams present a true 3D environment where a 2.5D approach was insufficient to solve. The algorithm described in this section therefore takes a departure from the previous lines of reasoning by considering the problem as a 3D one, while still taking notes from the previous algorithms. The idea behind the developed function is to first take as input the point cloud of a building framework (irrespective of the material) and decompose them first into distinct facets or segments. Afterwards several geometric constraints were applied to extract the point cloud of individual beams from those of the facets. The function was created to reach this point of the segmentation process; however, an optional third party library also enables the creation of parametric best-fit cuboids from the segmented beams. The overall workflow is described in Figure 6.1.

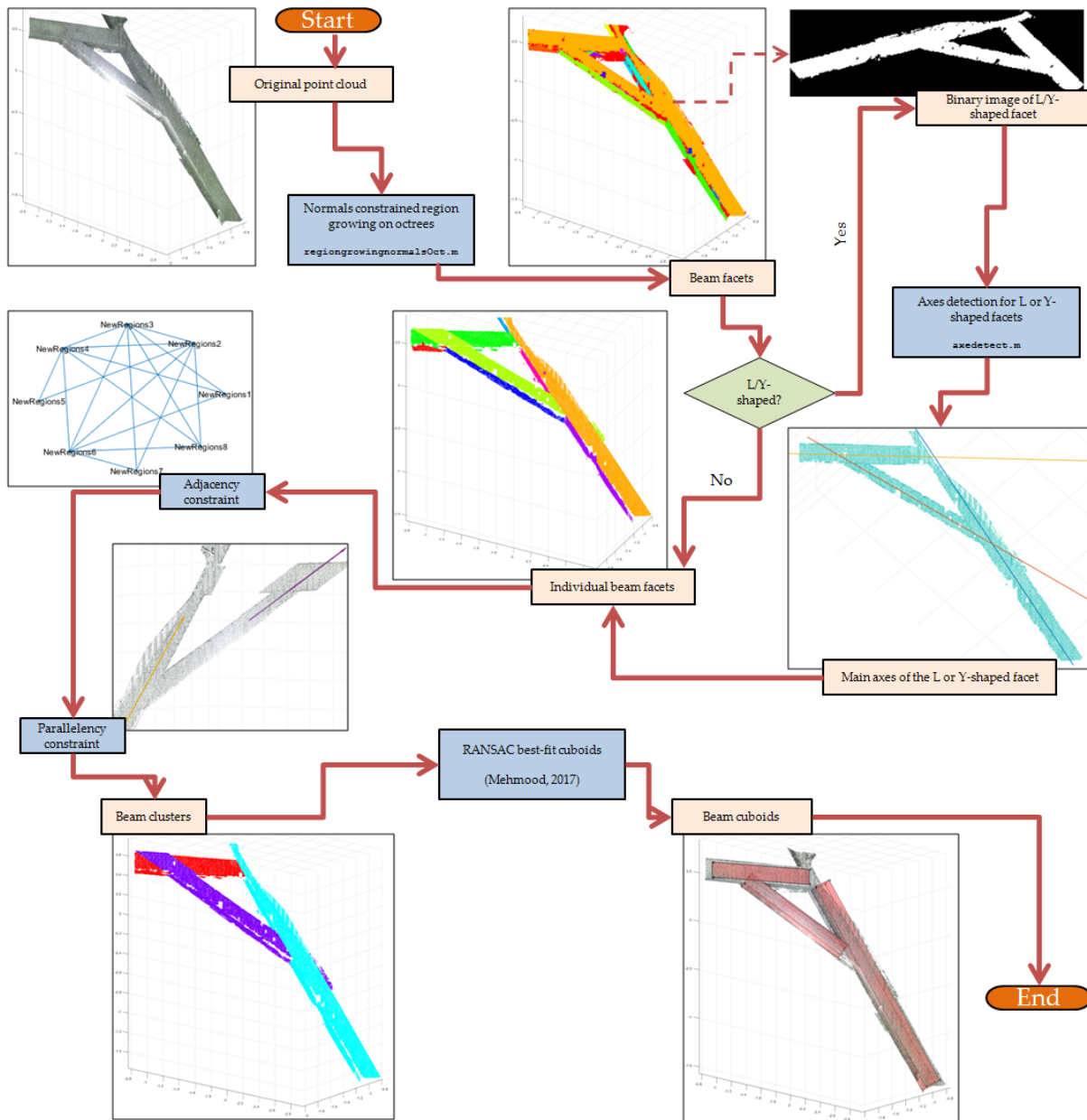


Figure 6.1: A flowchart showing the main steps in the `beamdetect.m` function of `M_HERACLES`, showing the intermediate and final results.

The facet detection was performed using the region growing method. The theoretical primer of the region growing method is well known; and in this case, we used the same approach as the Point Cloud Library (PCL) [RUSU AND COUSINS \[2011\]](#) implemented in Matlab<sup>®</sup>. This implementation employs point cloud normals and curvatures as constraints, as opposed to the function `pcsegdist` in Matlab<sup>®</sup> which uses Euclidean distance as the principal constraint.

The use of normals and curvature as constraints is important in order to distinguish the different facets; indeed, the use of either Euclidean distance proximity or voxel adjacency would not be able to distinguish neighbouring facets from each other. However, the use of greedy region growing algorithm, implemented to all points, takes too much resource and computing time and is therefore impractical. A solution to this problem is a slight tweak in the algorithm to perform the region growing on voxel bins instead of the points

themselves VO ET AL. [2015]. A similar implementation of this algorithm was the fast-marching approach described in DEWEZ ET AL. [2016]. From our observations, this voxel-based region growing has shown to increase the computing time up to a factor of 10 in empiric tests.

The post-segmented point cloud created clusters of facets. However, in the case of branching (L or Y-shaped) facets, additional segmentation was necessary in order to segment the faces properly for each beam. In our approach, the facet cluster was projected into a 2D binary image via PCA (Principal Component Analysis) transformation using the first and second components as the main X and Y axes; an approach similar to MACHER ET AL. [2017] and PÖCHTRAGER ET AL. [2017]. Afterwards, a Hough Transform analysis was performed on the binary image in order to detect the edges. The computed edges for each beam facet were thereafter averaged to obtain the centre axis for each beam facet. When the axis is detected, the L or Y-shaped facet was segmented into individual elongated or I-shaped clusters as is necessary for the next step of the approach.

Once the individual facets of individual beams were detected, two geometrical constraints were applied to group the facets into clusters of beams. The two constraints were similar to the ones used in PÖCHTRAGER ET AL. [2017], although in this algorithm only two out of the three mentioned in that paper were used. This reduction in geometric constraints was done in order to prevent over-constraining the problem. The two constraints applied in the algorithm are as follows:

- *Adjacency constraint*: the neighbourhood or adjacency constraint was enforced to limit candidate facets of each beam to only facet clusters which are located adjacent to the current reference facet. In PÖCHTRAGER ET AL. [2017], this constraint was defined by the distance between the facet centroids. The main problem in this approach is of course when dealing with short facets located next to elongated ones, where the geometric constraint of centroid neighbourhood is not enough. In M\_HERACLES, we modified this approach by implementing voxel-based region growing on the facets, this time around by enforcing a distance threshold between adjacent voxels from different facets as the constraint. In this way, adjacency is defined by whether any edge of the facet cluster is near to another one instead of only on the facet centroid's neighbourhood.
- *Parallelism constraint*: once the adjacency between the different facets is defined (established via an adjacency matrix), the search for candidate beam facets is reduced to the defined neighbours. Between neighbours, another geometric constraint on the parallelism of clusters was enforced. The underlying premise of this constraint is that the main axis of facets belonging to the same beam would normally be approximately parallel to each other. Firstly, the major principal axis of the facet clusters was computed using PCA. Two facets are considered parallel if their first PCA components satisfy equation 6.2:

$$\overrightarrow{OA_1} - \overrightarrow{OA_2} \approx 0 \quad (6.2)$$

where  $\overrightarrow{OA_1}$  is the first PCA component of the first (or reference) facet cluster and  $\overrightarrow{OA_2}$  the analogous vector for the second (or tested) facet cluster. A certain tolerance threshold was naturally required to be imposed, as perfect parallelism is almost mathematically impossible due to the classical quantization problem. Since the first

adjacency constraint already limited the candidate facet clusters for a beam, this second geometric constraint was deemed enough to detect the beam.

PÖCHTRAGER ET AL. [2017] proposed a third constraint: the angle formed between the normals should be either a right or straight angle:

$$\cos^{-1}(\vec{n}_1 \cdot \vec{n}_2) = \left[ 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2} \right] \quad (6.3)$$

with  $\vec{n}_1$  the normal vector for the first or reference facet and  $\vec{n}_2$  the comparable vector for the second or tested facet. However, M\_HERACLES chose not to implement this third constraint, as it may over-constrain the detection problem. Indeed, using the modified adjacency constraint has greatly improved the process by reducing the potential facets to true neighbours (instead of centroid neighbours) that the simpler parallelism check was deemed enough to correctly detect beam candidates from the pool of available facets.

The output of the function is therefore a structure containing the point cloud of the individual beams. A further processing may then be performed by generating best-fitted cuboids into these beam point clouds; however M\_HERACLES is not yet equipped with a cuboid-fitting function. For test purposes we used the RANSAC cuboid-fitting function developed by Usama Mehmood<sup>2</sup>.

## 6.4 Experimental results

This section will present the results from the tests conducted on the developed functions described in sections 6.2 and 6.3. This chapter involves the most number of tested datasets, as within this scale level more data are available. Two of the three main datasets were both used for tests, only the St-Paul dataset was not used since we didn't manage to acquire the interior point cloud. All the supporting datasets were added specifically to test the algorithms of this chapter, including the Paestum and Valentino for the detection of structural supports and the HK-Castle for the detection of beams.

### 6.4.1 Structural supports

The first function as described in Algorithm 2 was applied to the Kasepuhan dataset in order to separate the building's body from its attic, while Algorithm 3 was applied to all four datasets (Kasepuhan, St-Pierre, Paestum, and Valentino) to detect, segment, and classify their respective structural supports. Furthermore, two pavilions in Kasepuhan were used for the tests, namely the Royal Pavilion and the Musicians' Pavilion (BUILDINGS1 and BUILDINGS2 respectively in Figure 5.4a), therefore yielding a total of five case studies.

The first result concerning the attic segmentation algorithm for the two Kasepuhan buildings can be seen in Figure 6.2. The algorithm detects an abrupt change in overall cross-section convex hull area and determines automatically the upper part as the attic and the lower part as the building's body. In this regard, the programme managed to detect the attic automatically and quickly (less than 5 seconds for both objects). This part of the algorithm is aimed as a sort of pre-processing for the point clouds of buildings which possess an attic.

---

<sup>2</sup><https://mathworks.com/matlabcentral/fileexchange/65168-cuboid-fit-ransac>, retrieved on 4 March 2020.

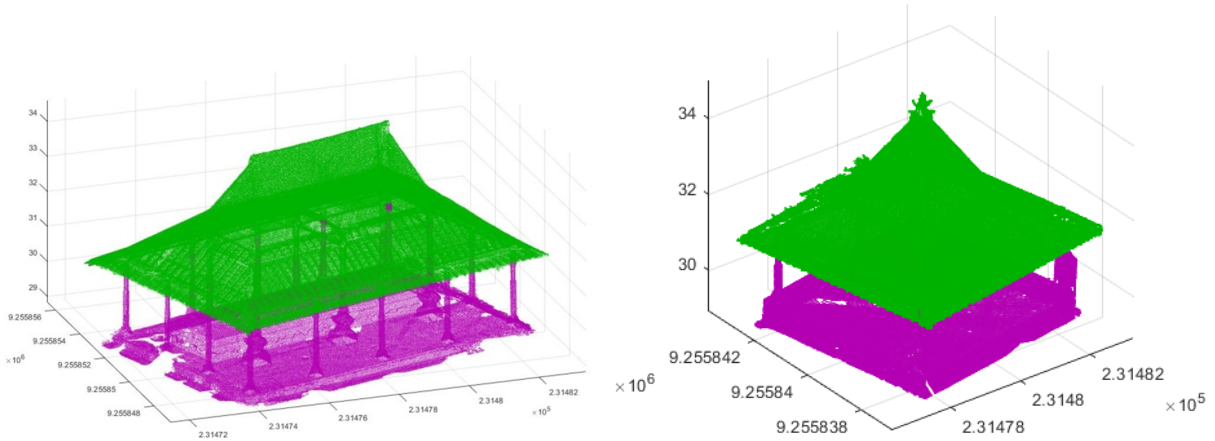


Figure 6.2: Automatically segmented building attic and body for the Kasepuhan Royal Pavilion (left) and Musicians' Pavilion (right) datasets; the attic is shown in green and the body in purple.

	Kasepuhan		St-Pierre	Paestum	Valentino
	Royal	Musicians'			
<b>Object type</b>	Pavilion		Church choir	Greek temple ruins	Room/Hall
<b>Point count</b>	155 115	127 223	1 852 162	1 181 025	3 513 341
<b>Proc.time</b>	38.28 secs	25.1 secs	83.56 secs	341.16 secs	307.8 secs
<b>Segmentation</b>					
<b>Structures</b>	20	8	8	58	20
<b>Ground truth</b>	20	8	8	58	19
<b>Notes</b>	N/A	N/A	4 twin columns	N/A	6 free standing, 13 engaged
<b>Classification</b>					
<b>Columns</b>	6	0	8	56	6
<b>Non-columns</b>	14	8	0	2	14
<b>Ground truth</b>	6 columns, 14 rectangular posts	8 rectangular posts	8 columns	56 columns, 2 antae	19 columns, of which 13 are engaged

Table 6.1: General parameters and overview of the segmentation and classification results.

This preliminary processing will enable a fully automatic workflow which begins with the point cloud of the entire building as input, therefore minimising as much as possible any human intervention during the process. In this way, this operation serves as a precursor to the structural support detection part of the developed algorithm.

In the developed workflow, this step is followed immediately by the structural support detection as expressed in the pseudocode of Algorithm 3. For the Kasepuhan dataset, the building body parts which was previously segmented were used as input, while for the other datasets the original point clouds were either directly used as inputs due to the inexistence of attics (e.g. Paestum and Valentino) or manually segmented for this particular purpose (St-Pierre). As has been previously described in section 6.2, there are two main parts of the algorithm which are performed simultaneously; namely the segmentation and the classification. The results can be consulted in Table 6.1 and visually displayed in Figure 6.3.



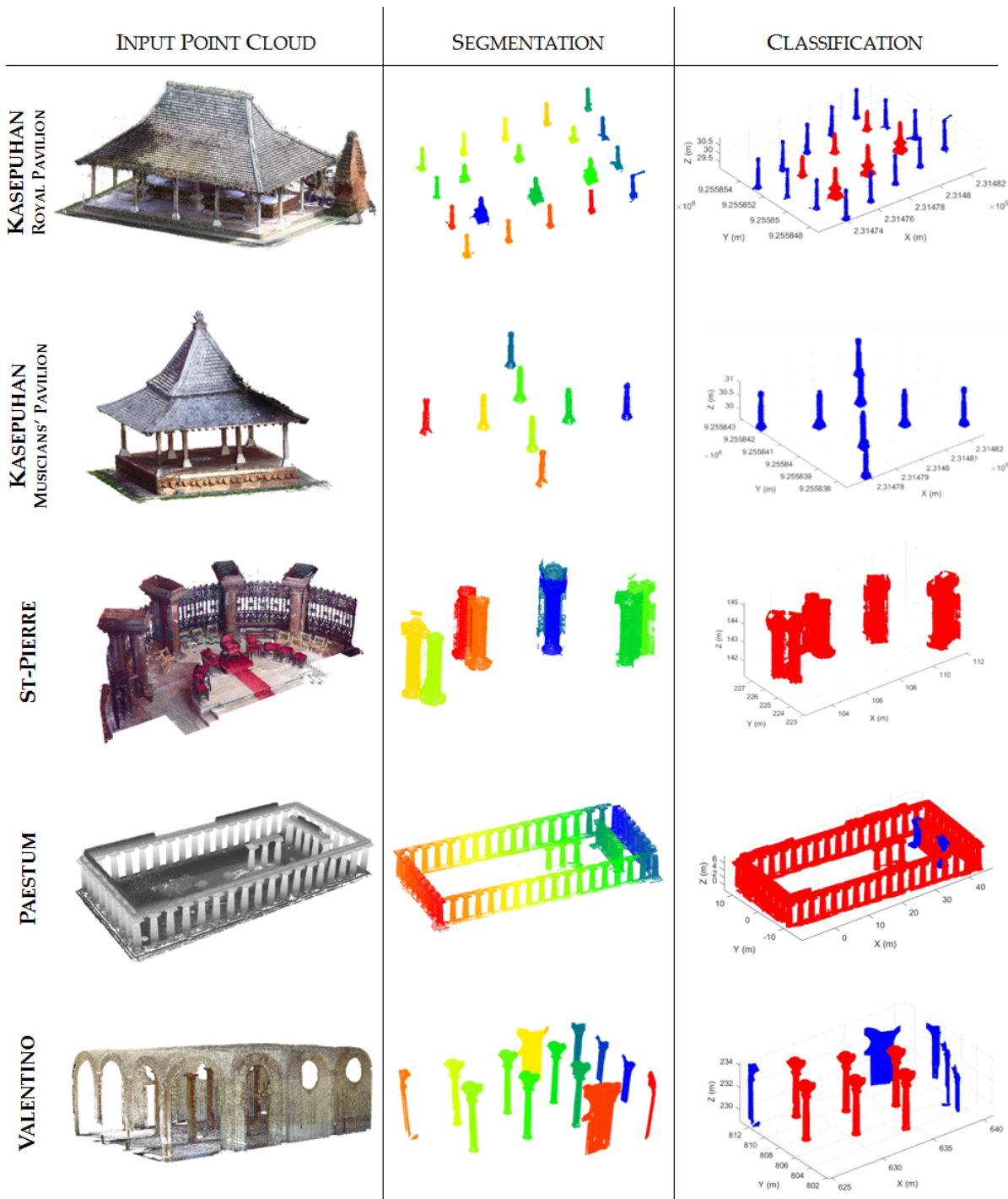


Figure 6.3: Visual results of the segmentation and classification process for the 5 tested cases. For the segmentation part each colour denotes individual detected objects, while for the classification red denotes the "column" class and blue the "non-column" class.

The Kasepuhan datasets are the smallest in size, with a little over 155K points for the Royal Pavilion and 127K points for the Musicians' Pavilion (note that in these cases only the body point cloud post-segmentation by Algorithm 2 is used as inputs). The St-Pierre, Paestum and Valentino datasets presented much larger point clouds with over 1.8M, 1.1M and 3.5M points respectively. It is also interesting to note that the tested datasets possess different styles of architecture; the Kasepuhan dataset consists mainly of open pavilions with many free-standing columns, while the Valentino presents an example of an interior

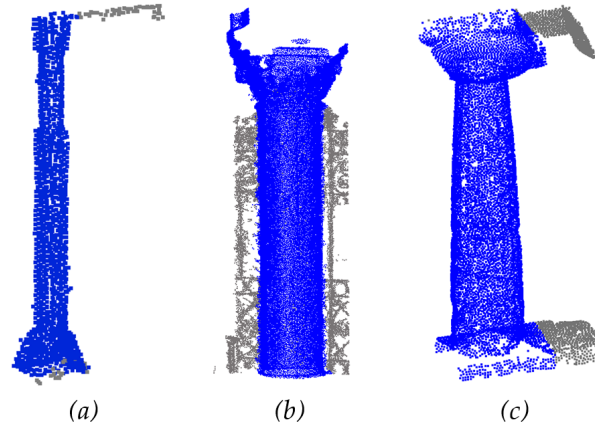


Figure 6.4: Illustration of some problems encountered during the segmentation process with sample pillar result clusters from: (a) Kasepuhan Royal Pavilion, (b) St-Pierre, and (c) Paestum. Blue colour denotes true positive points, while grey ones are false positive points.

point cloud case. Paestum also constituted a colonnade of free standing columns similar to Kasepuhan, but with a different layout and style. Meanwhile the St-Pierre church choir was chosen especially due to the presence of twin pillars.

As can be seen in Table 6.1, in terms of the segmentation process the algorithm managed to detect 20 supports for the Royal Pavilion and 8 for the Musicians' Pavilion. This corresponds exactly with the ground truth data. In the case of the St-Pierre, interestingly it managed to correctly detect the 8 pillars individually despite the twin pillar nature. Likewise, the correct number of pillars were detected for Paestum (58 out of 58). For the Valentino, 20 structures were detected in lieu of the actual 19 structures as can be found in the ground truth.

While the detection of the supports in the two Kasepuhan objects and Paestum are quite straightforward due to the fact that in both datasets the supports are fairly apart from each other (i.e. free standing pillars), the cases of St-Pierre and Valentino are more complex. In the St-Pierre choir data, the eight pillars are actually four pairs of twin pillars, with each pair consisting of two columns conjoined at the plinth and capital levels. Furthermore, the posterior columns of each pair are attached to an iron fence which links the four pairs and forms a barrier between the choir and the ambulatory located behind it. Difficulties arose when applying the algorithm by default, because the function arbitrarily takes the middle altitude cross-section of the point cloud to perform the detection part. In this regard, the iron fence hindered a proper detection of the posterior columns as stand-alone supports. A tweak was necessary to be applied to the algorithm in order to properly detect each support, namely by setting the cross-section profile to be used in the detection part to the one just beneath the capitals where the iron fence ends. The buffering of the convex hull cookie-cutter polygon also needed to be adjusted as to take into account the short space between the two columns in each pair. The Valentino data also presented a particular challenge since 13 out of the 19 pillars present in the dataset are in the form of engaged pillars, i.e. semi-pillars or columns which are part of the wall. As can be seen from the results, the algorithm had difficulties in segmenting these kinds of structural supports, contrary to cases with only free-standing pillars.

Figure 6.4 shows a sample of some of the results detected by the algorithm while showcasing some of the problems encountered. For the Kasepuhan Royal Pavilion (Figure 6.4a), some erroneous points were segmented together with a cluster, effectively presenting

a case of overclassification or false positive points. In this case, the overclassified points are those belonging to a sign post which was attached to the structure. This error typically manifests due to the use of the cookie-cutter and Euclidean distance-based region growing approaches. Although the RANSAC-based plane-fitting filter managed to exclude the floor part of the segmented cluster, the use of distance-based region growing did not manage to exclude the one of the sign post. This is because the sign post is attached to the pier, effectively telling the algorithm that these points belong to the same cluster.

A similar problem can be observed with the St-Pierre dataset. As has been previously explained, the posterior columns are attached to an iron fence. Here the same problem with the combination of the cookie-cutter and distance-based region growing manifested itself. Indeed, the same reasoning can be followed to explain why a small part of the iron fence on each side of the support was included in the cluster (Figure 6.4b). A similar, albeit more curious, problem can be seen in the Paestum dataset. Here the same argument regarding the disadvantages of the cookie-cutter and distance-based region growing segmentation may explain the existence of the false positive points (Figure 6.4c). However, the Paestum dataset displays a systematic tendency to this, in which the same case happens not only to one support, but indeed many. This may be explained by the iterative nature of the algorithm. A possible solution would be to refine the algorithm's parameter, for example by fine tuning the buffering radius of the support's cross section convex hull.

As has been previously mentioned, a preliminary segmentation was performed to divide the two Kasepuhan datasets into the building body and attic, with the body used as input for the function `supportdetect.m` as described in Algorithm 3. In these two datasets, some points mainly at the top of the pillars remained unclassified. This is due to the fact that in the pre-segmentation of the building body and attic, the algorithm considered the change in the surface of the cross-sections of the building to determine the two parts. These cross-section surfaces being calculated from the surface of the bounding box, only the exterior point cloud was considered leaving the interior out. This is reflected numerically in Tables 6.2 and 6.3, where the recall value for this dataset is visibly low despite a very high precision. The inverse is seen in the Valentino dataset, where the non-planar ceiling created a case of oversegmentation. Indeed, the statistics shown in Table 6.5 shows a high rate of recall but lower precision.

Statistically speaking, within Tables 6.2, 6.3, 6.4 and 6.5 the overclassified column describes the number of points considered as false positives, while the unclassified column denotes the true negative points. False negative points are not shown since the values are negligible due to the cookie-cutter approach of taking all points of all elevations of a particular polygon shape. Similar to the analysis conducted in Chapter 5, four statistical values were presented to assess the quality of the algorithm, namely percentage of unclassified points, precision, recall, and F1 score. In terms of the unclassified percentage, the Musician's Pavilion showed a higher rate (median of 34.36%) which is most probably caused by the same reasons as the one established before regarding errors during the pre-segmentation between the body and the attic. The median precision of Kasepuhan is 99.60% for the Royal Pavilion and 100% for Musicians' Pavilion, which is very satisfactory. However, as has been previously mentioned their recall value is lower at 77.27% and 65.64% respectively. This loss in recall value also seems to be systematic, again validating the points as argued in the previous paragraph. The overall median F1 score for the Kasepuhan dataset was 86.42% for the Royal and 79.23% for the Musicians' Pavilions.

The statistics for the St-Pierre dataset displayed a similar trend to that of the two

Object	Point Number		Misclassified		True positive	% Unclassed	%P	%R	%F1
	Manual	Auto	Overclassified	Unclassed					
COLUMN1	4 438	3 325	1	1 114	3 324	25.10	99.97	74.90	85.64
COLUMN2	2 222	1 341	9	890	1 332	40.05	99.33	59.95	74.77
COLUMN3	4 431	3 360	30	1 101	3 330	24.85	99.11	75.15	85.48
COLUMN4	2 318	1 286	0	1 032	1 286	44.52	100.00	55.48	71.37
COLUMN5	3 973	3 059	27	941	3 032	23.68	99.12	76.32	86.23
COLUMN6	1 728	1 023	1	706	1 022	40.86	99.90	59.14	74.30
OTHER1	1 837	1 475	24	386	1 451	21.01	98.37	78.99	87.62
OTHER2	1 809	1 265	7	551	1 258	30.46	99.45	69.54	81.85
OTHER3	1 891	1 415	1	477	1 414	25.22	99.93	74.78	85.54
OTHER4	1 757	1 413	2	346	1 411	19.69	99.86	80.31	89.02
OTHER5	1 754	1 532	98	320	1 434	18.24	93.60	81.76	87.28
OTHER6	1 896	1 631	129	394	1 502	20.78	92.09	79.22	85.17
OTHER7	1 644	1 425	52	271	1 373	16.48	96.35	83.52	89.48
OTHER8	1 421	1 131	26	316	1 105	22.24	97.70	77.76	86.60
OTHER9	1 615	1 243	3	375	1 240	23.22	99.76	76.78	86.77
OTHER10	1 720	1 282	3	441	1 279	25.64	99.77	74.36	85.21
OTHER11	1 609	1 291	3	321	1 288	19.95	99.77	80.05	88.83
OTHER12	1 763	1 418	21	366	1 397	20.76	98.52	79.24	87.83
OTHER13	1 273	1 006	1	268	1 005	21.05	99.90	78.95	88.20
OTHER14	1 605	1 309	1	297	1 308	18.50	99.92	81.50	89.77
<b>Mean</b>						25.12	98.62	74.88	84.85
<b>Median</b>						22.73	99.60	77.27	86.42

Table 6.2: Table of the quantitative analysis on the results of step 2 for the detection and classification of columns in the Kasepuhan Royal Pavilion dataset. %P is precision, %R is recall, and %F1 is the normalised F1 score.

Kasepuhan datasets in that it registered higher precision than recall rate. With a median precision of 81.20% and recall of 71.39%, the results for this dataset is nevertheless quite promising. It should be well noted that the St-Pierre choir dataset is quite complex due to the existence of the twin pillars and the presence of many noises (folded chairs were placed against the twin pillars in addition to the presence of the iron fence on the posterior pillar). Indeed, manual segmentation and labelling (to generate reference data) took quite some time to perform the task due to these conditions. Granted, the automatic results still had remaining noises and must be cleaned further manually. However, with a fast processing time (a little under one and a half minute) this solution may prove to be very useful in performing the segmentation task, or at least provide a first approximate result. St-Pierre displays the highest percentage of unclassified points, amounting to 28.61% in median. This is easily explained by ambiguities due to the existence of the iron fence between the choir and the ambulatory, while the twin nature of the columns also generated errors. Furthermore, the input point cloud was not pre-processed or cleaned beforehand.

For the Valentino dataset, the unclassified rate stands at a median value of 11.84%. The precision level is low for the Valentino, at 66.68% which suggested an overclassification.

Object	Point Number		Misclassified		True positive	%Unclassed	%P	%R	%F1
	Manual	Auto	Overclassified	Unclassed					
K01	2 963	2 106	2 106	0	857	28.92	100.00	71.08	83.09
K02	2 543	1 819	1 815	4	728	28.63	99.78	71.37	83.22
K03	2 577	1 787	1 783	4	794	30.81	99.78	69.19	81.71
K04	2 379	1 618	1 618	0	761	31.99	100.00	68.01	80.96
K05	3 698	2 340	2 340	0	1 358	36.72	100.00	63.28	77.51
K06	3 440	2 158	2 158	0	1 282	37.27	100.00	62.73	77.10
K07	3 646	2 282	2 282	0	1 364	37.41	100.00	62.59	76.99
K08	3 361	2 117	2 117	0	1 244	37.01	100.00	62.99	77.29
<b>Mean</b>						33.60	99.94	66.40	79.73
<b>Median</b>						34.36	100.00	65.64	79.23

Table 6.3: Table of the quantitative analysis on the results of step 2 for the detection and classification of columns in the Kasepuhan Musicians' Pavilion dataset. %P is precision, %R is recall, and %F1 is the normalised F1 score.

As mentioned earlier, this is mainly due to the ceilings of the dataset that represent arcs instead of planar surfaces as had been hard-coded in the algorithm. A further improvement of the algorithm may incorporate this possibility into account, as this type of ceiling can be found in many heritage datasets. However, the recall value is quite high with a median value of 88.16%, thus yielding an F1 score of 75.92%. The high recall value is also reflected to some degree by the lowest unclassified value amongst the tested datasets (median of 11.84%). This means that the algorithm does nevertheless give promising results. Indeed, in some applications where high precision is not necessarily required (e.g. training data generation for deep learning techniques) these results may be sufficient.

No detailed table was made for the Paestum dataset, as the amount of structures (58) was too impractical to put into a table. However, the dataset did yield an average precision value of 90.67%, with recall value at 87.33% (thus an F1 score of 88.97%). These values are very good and is thus encouraging; however as has been noted before, systematic error is present in the results (Figure 6.4c). Paestum also yielded a 12.67% overall unclassified percentage, which is more or less comparable to the results of Valentino.

For the classification of columns *vs.* non-columns, the algorithm utilised the circularity value of the cross-section convex hull to determine if a cluster is attributed the column class; columns being characterised by a more circular form denoted by a circularity value of around 1. An empiric value of 1.12, computed from the average circularity of ground truth columns, was used as the threshold between the "column" and "non-column" classes. Both Table 6.1 and Figure 6.3 also displays the results of the classification phase.

For the Royal Pavilion dataset, the algorithm managed to detect the 6 columns located at the inner part of the pavilion, of which three are located on an elevated dais. The surrounding 14 wooden piers were also correctly identified as non-columns. The algorithm took 38.28 seconds to generate this result. The algorithm also managed to correctly classify all the pillars of the Musicians' Pavilion as non-columns. Indeed, under the definition of columns as set in this study and contrary to classical columns, the eight structures in the Musicians' Pavilion dataset cannot be classified as columns as they are in fact rectangular shaped posts (otherwise known architecturally as "piers"). The processing time for this dataset was 25.10 seconds.

Object	Point Number		Misclassified		True positive	%Unclassed	%P	%R	%F1
	Manual	Auto	Overclassified	Unclassed					
S01	72 587	54 995	47 709	7 286	24 878	34.27	86.75	65.73	74.79
S02	66 298	64 952	52 922	12 030	13 376	20.18	81.48	79.82	80.64
S03	74 430	55 979	50 435	5 544	23 995	32.24	90.10	67.76	77.35
S04	71 667	59 277	43 647	15 630	28 020	39.10	73.63	60.90	66.67
S05	64 893	54 969	54 343	626	10 550	16.26	98.86	83.74	90.68
S06	66 678	61 804	50 018	11 786	16 660	24.99	80.93	75.01	77.86
S07	67 316	75 062	51 996	23 066	15 320	22.76	69.27	77.24	73.04
S08	60 165	49 212	35 814	13 398	24 351	40.47	72.77	59.53	65.49
<b>Mean</b>						28.78	81.72	71.22	75.81
<b>Median</b>						28.61	81.20	71.39	76.07

Table 6.4: Table of the quantitative analysis on the results of step 2 for the detection and classification of columns in the St-Pierre dataset. %P is precision, %R is recall, and %F1 is the normalised F1 score.

Object	Point Number		Misclassified		True positive	%Unclassed	%P	%R	%F1
	Manual	Auto	Overclassified	Unclassed					
V01	35 370	46 666	15 594	4 298	31 072	12.15	66.58	87.85	75.75
V02	35 845	47 358	15 744	4 231	31 614	11.80	66.76	88.20	75.99
V03	39 169	51 853	17 333	4 649	34 520	11.87	66.57	88.13	75.85
V04	40 155	51 923	17 010	5 242	34 913	13.05	67.24	86.95	75.83
V05	38 288	52 623	17 575	3 240	35 048	8.46	66.60	91.54	77.10
V06	39 689	53 016	17 406	4 079	35 610	10.28	67.17	89.72	76.82
<b>Mean</b>						11.27	66.82	88.73	76.23
<b>Median</b>						11.84	66.68	88.16	75.92

Table 6.5: Table of the quantitative analysis on the results of step 2 for the detection and classification of columns in the Valentino dataset. Note that only the detected columns were taken into account here. %P is precision, %R is recall, and %F1 is the normalised F1 score.

In the case of the Paestum dataset, the ground truth data gave a total of 58 pillars comprised of 56 circular columns and 2 antae. The antae are pillars or posts, most often rectangular, often located on either side of the entrance to the inner part of a Greek temple. Within the predefined context identified at section 3, these two antae were therefore considered as non-columns. The algorithm managed to detect the same number of support types, and correctly determined which support belonged to which class. The whole processing of the Paestum data, comprising the detection, segmentation, and classification, took a total of 341.16 seconds.

The St-Pierre data also showed promising results, as the programme managed to identify the eight supports as columns, their twin nature notwithstanding. In this case, the processing time amounts to 83.56 seconds. For the Valentino, it also managed to correctly identify that the free-standing pillars are columns, while the rest of the detected structures were classified as non-columns. The segmentation and classification of Valentino took 307.8 seconds.

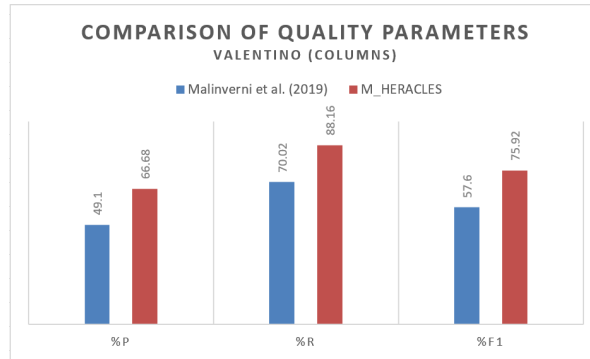


Figure 6.5: Histogram representation of the quality parameters for the Valentino dataset comparing the results of MALINVERNI ET AL. [2019] against M\_HERACLES for the column class.

The processing time of the datasets showed that they may be, at least in part, linked to the number of points inside the input data. However it is more probable that the bulk of the processing time is linked to the number of the detected elements. For example, for the 127K points Musicians’ Pavilion data, the algorithm managed to detect, segment, and classify the objects in 25.1 seconds. This was done in 83.56 seconds for the 1.8M points St-Pierre dataset, also with 8 detected structures. Conversely, the Valentino dataset which consists of almost 10 times more points than Kasepuhan were processed in a little over 5 minutes in order to detect 20 structures. However the Paestum data took more time to process compared to the denser Valentino data due to the amount of supports that had to be detected (almost three times that of Valentino).

The segmentation is therefore the part that takes more time depending on the point cloud density and the number of identified structural supports. However, the overall processing time is still faster by at least a factor of 2 when roughly compared to the time it takes to perform the same task manually, without taking into account the time required to identify and classify each cluster into the appropriate classes. Although the algorithm managed to perform the classification task well enough, it should be noted that fine tuning is still required in specific cases.

A quick comparison was also performed between our results and the results presented in MALINVERNI ET AL. [2019] which also used the Valentino in their experiments with PointNet++ DL approach. As has been mentioned previously, M\_HERACLES managed to yield a median precision value of 66.68%, recall value of 88.16%, and F1 score of 75.92%. In MALINVERNI ET AL. [2019], Valentino was used as the test dataset after the authors’ DL algorithm was trained using another dataset and was classified into four classes, including columns. For the columns class, the authors cited a value of 49.10% in precision, 70.02% in recall, and 57.60% in F1 score (Figure 6.5). Although our algorithm managed to provide better results than the compared study, several remarks should nevertheless be taken into account. Firstly, in our study only free-standing pillars were accounted for, whereas MALINVERNI ET AL. [2019] also included engaged columns. Indeed, M\_HERACLES did not manage to correctly detect the engaged columns. Secondly, the DL approach used in the other study has the potential to generate better result with more training data.

### 6.4.2 Beams

This part of the algorithm has not been tested as thoroughly as the pillar detection part; however a preliminary result (as shown in Table 6.6) conducted on a subset of the HK-

Object	Point Number		Misclassified		True Positive	%Unclassed	%P	%R	%F1
	Manual	Auto	Overclassified	Unclassed					
Beam1	15 036	10 960	608	4 684	10 352	42.74	94.45	68.85	79.64
Beam2	57 986	43 826	0	14 160	43 826	32.31	100.00	75.58	86.09
Beam3	28 789	26 141	2 355	5 003	23 786	19.14	90.99	82.62	86.60
					<b>Mean</b>	<b>31.40</b>	<b>95.15</b>	<b>75.68</b>	<b>84.11</b>
					<b>Median</b>	<b>32.31</b>	<b>94.45</b>	<b>75.58</b>	<b>86.09</b>

Table 6.6: Table of the quantitative analysis on the results of step 2 for the detection and classification of beams. %P is precision, %R is recall, and %F1 is the normalised F1 score.

Castle dataset showed that the algorithm managed to correctly identify the individual beams. The small dataset consists of 100k points and was processed in 3 minutes 42 seconds. The algorithm gave very good results in terms of precision (median value of 94.45%), but quite low values of recall (median value of 75.58%). The low recall value can be explained by the fact that the algorithm also performs noise reduction, in which detected regions having less than a set threshold number of points are eliminated. The resulting cluster is therefore cleaner than the manual segmentation, but this means a sharp decrease in recall value. The precision value is however very satisfactory. Furthermore, the algorithm correctly deduced the number of beams that are present in the input point cloud. However, the algorithm still suffers in terms of processing time. More than half of the processing time was taken by the curvature computation at the beginning of the function; this requires therefore more investigation and optimisation. Moreover, these preliminary results concern only a small dataset. More investigations must be conducted to assess the quality of the algorithm, namely by processing a larger dataset.

## 6.5 Summary

This chapter presented a series of functions dedicated to point cloud processing in the context of heritage objects and specifically for the scale step of a heritage building. The main driving cause of the development of these algorithms is to address the increasingly multi-sensor and multi-scalar nature of heritage documentation. The presented pillar and beam detection functions enable the user to automate one of the main bottlenecks of 3D modelling; namely the segmentation and classification of architectural elements. Automation in this regard may therefore reduce human intervention and thus human error. Results for the presented functions look promising. Experimental results show the following important points:

- **Promising results for the pillar detection task, with some caveats:** In section 6.2, the tests on five datasets showed that the algorithm is useful in performing fast segmentation and classification for structural supports. However, as has been shown in the results, while the algorithm is fast and easy to use it remains prone to noise and deviations from the hard-coded geometrical rules. This is evidenced by the stark contrast between the five datasets, among which Kasepuhan presented cases with higher precision but lower recall (suggesting underclassification), the Valentino showing higher recall but lower precision (suggesting overclassification) and the St-Pierre and Paestum presenting somewhat of a mix between these two cases. These



differences were caused by deviations from the general rule; both of the Kasepuhan interior ceilings does not correspond to the altitude of the roofs on the exterior, while Valentino's non-planar ceiling caused the error. The St-Pierre case showed that the proneness of the algorithm against noises, while Paestum warned on the effects of systematic error. However, the results remain promising and some lessons can be learned from this experiment, to be the subject of further improvement. Furthermore, the fast nature of the segmentation and classification process is in contrast to training-intensive machine learning/deep learning and resource-intensive manual segmentation and labelling. It may therefore be used to complement machine learning algorithms, especially in the generation of training data.

- **Early results for the beam detection task with more tests required:** The beam detection algorithm described in section 6.3 is still in its early stages, and more tests must be conducted in order to better assess its efficacy. A test with a small dataset yielded a very satisfactory median precision value of 94.45%, but with low median recall value of 75.58% which is mostly due to the noise-filtering function applied in the algorithm. Although these preliminary results are promising, processing time remains an important issue. Most of the processing time was used for the computation of the normals and especially curvatures. More research should be conducted to optimise this part of the algorithm better.
- **Potential to act as a complement to ML/DL methods:** One interesting idea which has been planned to be tested is to use the results from the algorithms described in this chapter, whether the detection of structural supports or beams, to help generate training data for machine learning and deep learning techniques. As has been previously established, one of the bottlenecks in these approaches is the creation and labelling of training data which is performed manually. The algorithms proposed in this chapter may help to automate (or at least provide an "approximate value") this training data generation process, thus rendering the overall 3D processing pipeline more automatic.

# Conclusions

Suppose a man can convince me of error and bring home to me that I am mistaken in thought or act; I shall be glad to alter, for the truth is what I pursue, and no one was ever injured by the truth, whereas he is injured who continues in his own self-deception and ignorance.

---

Marcus Aurelius (121-180 CE), *Meditations*, *Book VI*

Throughout this thesis, the overarching topic has been the attempt to automate as much as possible the 3D reconstruction process. This went from the acquisition of geospatial data from techniques such as photogrammetry and laser scanning, their integration and geometric quality control, up to the post-processing of the point cloud as the main result. All of these were also done in the context of heritage documentation, to better serve the heritage community in archiving and presenting tools in order to understand them better. In this sense, the thesis has revolved around two main research questions which were thereafter elaborated in the two parts of this manuscript. In respect to the point cloud as the main product of the whole 3D pipeline, the first problem that we tried to address in Part I involves the manners pertaining to its generation (**Chapter 1**). This includes the basics of 3D data acquisition, as well as a proposed multi-scalar and multi-sensor workflow. This opening chapter also introduced the main datasets that were used for the subsequent experiments. Furthermore, special emphasis was made on the quality control especially for the photogrammetric results (**Chapter 2**). Photogrammetric quality control was accentuated because the photogrammetric workflow was deemed to be more prone to errors due to the many involved steps relative to laser scanning. This is exacerbated by the fact that most of the commonly used photogrammetric software involve a black-box approach. Of course, this does not necessarily mean that laser scanning data do not require quality control. To a certain degree this was performed in **Chapter 3**, although it was conducted in the context of data integration with photogrammetry. Indeed, data integration presented the other emphasis of Part I, as the proposed multi-scalar approach is logically followed by a multi-sensor nature and thus the imperative to perform good data integration.

Part II of the thesis shifted the discussion to address the second research question of the study; namely what do we do with the point cloud generated at the aftermath of Part I. We notice here that the point cloud is central to the research, and this second part therefore focuses on the point cloud post-processing. **Chapter 4** introduced the solution we proposed to answer this question: the `M_HERACLES` toolbox. The chapter also discussed several theoretical notions that were used within the toolbox, as well as a quick presentation of the supporting datasets. In line with the multi-scalar approach to the data acquisition, `M_HERACLES` was designed to address the problem of multiple scale point clouds by deconstructing them into two scale steps. The first scale step involves

the segmentation and semantic annotation of point clouds at the neighbourhood or complex scale into individual clusters of heritage buildings (**Chapter 5**). Thus segmented, the heritage building point clouds can then be used for the next step: the detection and classification of architectural elements from buildings (**Chapter 6**). This chapter discussed the two classes that have thus far been implemented in `M_HERACLES`, namely the pillars and beams. `M_HERACLES` was therefore designed as a modular toolbox to answer the multi-scalar nature of the input data.

## Key findings

As has been previously established, the crux of this thesis essentially crystallises into two research questions. We attempted to answer these questions through the development of methods, algorithms, as well as experimental tests conducted on the available datasets. Some of the key findings from these observations and deductions will be elaborated as follows.

### How do we generate high quality point clouds for heritage buildings?

Chapter 1 established the *raison d'être* of the field of 3D heritage documentation. The necessity of archiving, as well as the possibility to diverge into various interactive 3D applications form the two main driving forces behind the necessity for heritage 3D recording. The chapter also presented a panorama on the myriad of available 3D techniques which may be used to perform the task of heritage building 3D recording. The two most commonly used techniques are the image-based photogrammetry and the range-based laser scanning, while the ever increasing popularity of drones was also noted. A multi-scalar and multi-sensor approach to 3D heritage recording was proposed, which is a logical output of the requirement to thoroughly document a heritage building within its geographical, and thus socio-historical, contexts. In consequence, data integration as well as quality control also became crucial parts of the overall 3D pipeline.

In order to generate high quality point clouds, it is necessary to perform quality control. This is even more so in the case of photogrammetric operations, where the prevalence of black-box solutions often hide potential problems and hinder the full potential of the project. In Chapter 2, photogrammetric quality control was performed using `DBAT`. The various experiments showed that some statistical values can be very useful in determining the quality of the project and detect potential problems. Some key findings include the behaviour of the self-calibration in the image orientation process. While small differences exist between `DBAT` and the tested software (`PhotoScan`), it is rather safe to assume that `PhotoScan` calibration yielded virtually identical results to those of the open alternative. Another interesting finding was the presence of some problems with exterior orientation standard deviation values in the project detected by `DBAT` which would otherwise have gone undetected in `PhotoScan`. This is interesting because by observing these statistics, we may identify problems in the project and eventually improve them by addressing these problems directly. However, we also acknowledge that the field of photogrammetry and SfM is evolving swiftly. The demand for quality control has also interestingly increased quite dramatically in the past few years. This can be observed in the evolution of `PhotoScan` (or `Metashape`), which incorporates more and more quality metrics in their software. This is indeed a very welcome development for the photogrammetric community.

As has been previously established in the first chapter, nowadays a complete multi-scalar point cloud consists also of multi-sensor data. In Chapter 3, this integration was investigated through the experiments on the two georeferencing approaches; namely the independent and free-network approaches. The experiments have shown that independent georeferencing is faster and easier to implement, especially when working in an established heritage recording standard of procedure. This is because in both photogrammetry and laser scanning, georeferencing (or absolute orientation) is already in the workflow. The independent georeferencing method as described in this thesis is therefore straightforward. However, our experiments also showed that the risk of systematic error is higher with this approach. The free-network method, while slower due the necessity of more steps requiring manual intervention, reduces this risk of systematic error as they are minimised in the several optimisation operations involved in the workflow (e.g. error sum minimisation during the 3D rigid body transformation and ICP). However, we have also established that automation is a crucial aspect of the thesis, and the free-network approach with its manual steps strays from this main idea. To remedy this, we have experimented with automatic 3D tie point extraction and matching. Unfortunately, further development into this research area is beyond the scope of the thesis. Only preliminary results were therefore presented in this manuscript. However, the results were promising and other studies based on our preliminary development have also shown encouraging results.

Judging from these points, we may deduce that point cloud generation, including that in the case of heritage recording, is nowadays a straightforward process. Using both photogrammetry and laser scanning, the basic operational knowledge is quite easy even for the uninitiated due to the various user-friendly developments of these past years. However, a high-quality point cloud nevertheless requires a more advanced understanding as to the inner workings of both techniques. Part I of the manuscript dealt with these details which are crucial in order to generate high-quality point clouds, and more so in the case of complex heritage buildings and sites.

### **What do we do next with our high quality heritage building point cloud?**

The point cloud remained a geometric entity with few semantic information. Therefore, the next logical step is to try to imbue the point cloud with attributes and classes. This will enable the point cloud to be used in other applications such as 3D GIS and HBIM while not limited to its geometric nature. In Chapter 4, we propose a toolbox called `M_HERACLES` to address this issue. `M_HERACLES` was designed following the multi-scalar approach elaborated in Part I of the thesis. The idea is to use point cloud processing algorithms (the so called algorithmic or hard-coded knowledge approach) to perform point cloud classification and semantic annotation. This is done progressively, first from the scale step of a heritage complex or neighbourhood to that of a building, and then subsequently from the scale step of a heritage building to that of architectural elements. This algorithmic approach employs various techniques which were also described in the chapter. With the advent of machine and deep learning, `M_HERACLES` was also designed to be able to support this development by aiding the generation of training data.

The first step when dealing with multi-scalar point cloud is to segment the point cloud of the area into clusters of heritage buildings of interest. Chapter 5 described how `M_HERACLES` performed this automatically. By using pre-existing 2D GIS data, the algorithm managed to vastly reduce the time required for manual segmentation. However, segmentation is not the only aspect that this chapter addressed. Indeed, by using the pre-existing GIS files, we took one little step further by annotating the resulting segmented

clusters automatically with the classes and attributes of the GIS. This is effectively the semantic classification aspect of the point cloud segmentation process, and further increases the effectiveness of the developed algorithm. Through qualitative analysis, a very satisfactory average F1 score of 92.95% was achieved from the experiments. This algorithm in itself is already very useful when dealing with aerial point clouds and not necessarily for heritage documentation purposes. Indeed, the results show that the algorithm is suitable for urban settings and may therefore also be used for other more general purpose applications such as city mapping and urban 3D modelling.

Once the point cloud of the buildings are acquired, the next logical step would be to address the larger scale segmentation of heritage building point clouds into architectural elements. In the context of heritage management such as HBIM this scale step is crucial as it describes the detailed components of the building. Chapter 6 presented our developed approach within the `M_HERACLES` toolbox. Two architectural classes were presented, namely pillars and beams. The pillar detection and segmentation was performed using a quasi 2.5D approach, similar to the one employed in the previous chapter. The main hard-coded knowledge in this function is the fact that columns possess circular cross-sections. The developed algorithm was tested on a total of five case studies with an average F1 score of 81.32%. The algorithm fared better in cases where free-standing pillars are present, and fared less in specific cases such as the presence of twin or engaged pillars (i.e. deviation from the hard-coded knowledge). This is an important lesson as it showed the limitations of the algorithmic approach, but also its advantage as a fast tool to generate classified datasets. Even if the results are imperfect, they are very promising as inputs for deep learning techniques or in some cases even for direct use in further 3D modelling. The beam detection function was the last to be developed during the thesis, and incorporated several efficient geometrical rules. However it has only been tested on one dataset, even though the results were also promising with an average F1 score of 86.09%. Nevertheless, in hindsight the developed beam detection function may have been more useful for direct generation of geometric primitives (i.e. cuboids for beams) instead of simple segmentation and classification.

What we can do with the generated high quality point cloud is indeed very important. Part II of this thesis provided the developments and tests that we performed to try to answer this question. Segmentation and classification of point clouds is very important in order to give a tangible sense to it, arguably more so in the case of heritage buildings. Although other modern approaches such as deep learning exist, they still have many shortcomings in the heritage domain. `M_HERACLES` was deliberately not designed to compete against deep learning, but rather to complement it to strive towards the common goal of fully automatic heritage point cloud classification.

## Room for improvements

While the results shown thus far were promising, potential improvements can still be identified. The basic premise of Chapter 2 involves the quality control of the photogrammetric project by means of open source solutions. `DBAT` was chosen as the primary tool to perform this. However at the moment it is not yet integrated with `M_HERACLES` to offer a fully automatic chain from data acquisition up to point cloud classification. However, since the experiments of Chapter 2 some other improvements have been added to `DBAT`; for example modularity of bundle adjustment components to enable a more flexible computation depending on the case [BÖRLIN ET AL., 2019b]. As far as photogrammetric

projects are concerned, automation is at the moment only possible partially since DBAT still requires third party photogrammetric software project as input.

As has been described previously, the results from Chapter 3 showed that the free-network georeferencing displayed a tendency towards less significant systematic error. The main issue with this approach is the largely manual tie point selection process. The developed automation algorithm was tested and the results presented in this thesis is still far from satisfactory. However, other research based on our initial algorithm has shown promising results [ASSI ET AL., 2019]. These results may be improved further by testing other types of 3D point detectors; this is however beyond the scope of this thesis.

The bulk of the programme and algorithm development of the thesis was done in Part II. In Chapter 5, the small scale segmentation of heritage neighbourhoods to heritage buildings did show interesting results. However some crucial problems were also identified. For example, the segmentation order of the classes is still very important in order to avoid vertical stacking. The developed approach also only employed one ground filtering method (the CSF by ZHANG ET AL. [2016]), and a study and comparison with other available ground filtering methods may be beneficial. Manual parameterisation of the function also remains one of the main issues; these parameters were determined empirically and fine tuning is still very much possible at this point. Another idea derived from the same reasoning of Chapter 5 was to use CAD files to perform a similar segmentation and semantic annotation for architectural elements. We explored this idea further in the paper by SEMLER ET AL. [2019].

The algorithmic approach used in M\_HERACLES was again used in Chapter 6 to detect pillars and beams. As far as the pillar detection is concerned, the algorithm worked very well with free-standing pillars but needed specific tweaks when encountering unique cases (e.g. twin and engaged pillars). As has been hypothesised, the approach is also prone to noises and the unexpected presence of unwanted objects on the point cloud. In the majority of cases involving free-standing pillars this is not an issue; however, as the experiments have shown, in some cases manual intervention is still required to clean the results. A pre-processing may be envisaged in this regard to clean the input point cloud prior to use as input for the algorithm. The same problem with proneness to noise can be encountered in the beam detection algorithm. Furthermore, processing time especially for the beam detection part still requires optimisation. Indeed, the most time-consuming part of the process was the computation of curvature values. This may be optimised further by using other methods for computation, for example voxel sampling. The generation of geometric primitives from the segmented point cloud classes is still in the early stages of development. In this thesis, we tested a RANSAC-based method of cuboid fitting to produce 3D models for the beams. This obviously may benefit from further research in the topic of automatic 3D modelling, for example by looking into developments in the scan-to-BIM domain.

## Concluding remarks

This thesis had attempted to answer the need for an efficient method for heritage building recording in 3D. 3D documentation is deemed important in this digital information age as they may serve not only as a digital archive, but also as a base for further analysis and management. The thesis also acknowledged that the massive and significant developments in 3D sensor and processing technologies over these past few decades have reached the point where anybody can perform 3D documentation even using low-cost sensors. This in

turn gave rise to black-box solutions; thus the imperative for quality control especially in the most vulnerable of the 3D pipeline stages (e.g. photogrammetric bundle adjustment assessment, integration between photogrammetry and laser scanning). While black-box solutions are useful in the sense that they enable the democratisation of 3D technologies, we nevertheless insist on the rigorous assessment of the products' quality. The proposed multi-scalar and multi-sensor workflow is also a logical outcome of the availability of various 3D recording sensors.

Furthermore, the semantisation of point cloud has always been a major issue addressed by researchers when confronted with point cloud data. We argue that semantic annotation and point cloud classification are even more important in the case of heritage buildings in order to take into account the diverse complexity of historical architectural styles. The thesis had started with the algorithmic approach in mind, but recent trends suggested that deep learning is becoming more and more feasible as a robust solution to the semantisation problem. However we also identified the shortcomings of deep learning approaches, namely the necessity for labelled training datasets. We also observed that this is in part due to the lack of open data structures for heritage buildings. We argue that conventional algorithmic approach may be used to address this issue and thus complement rather than compete against deep learning techniques. Nevertheless, another downfall that we managed to identify from deep learning approaches is its somewhat black-box nature in performing its task. This is also another argument to put forward for the development of an algorithmic-based toolbox in order to provide a more open alternative.

Various topics had been discussed in this thesis within the topic of the use of geospatial data for the 3D recording and point cloud classification of heritage buildings. Indeed, investigation in this domain quickly became a multi-disciplinary vast wilderness in which geomatics meets computer and social sciences. This multi-disciplinarity aspect is nonetheless the right way forward; indeed, we argue that as the digital age gave way to big data, no single domain may be able to fully describe the problem of 3D heritage recording. A multi-disciplinarity approach is therefore not only logical, but also necessary. Furthermore, we believe that in the future more and more training data will be available to support machine learning approaches; for this purpose we support the creation of open data portals for labelled heritage building point clouds. Through this thesis, we also hope to be able to accelerate this process towards full automation of the 3D recording workflow of heritage buildings while keeping the standards of geometric quality high.

# Scientific Contributions

## PUBLICATIONS

During the period of the PhD thesis from April 2017 to June 2020, a total of 30 scientific contributions were published or currently in press.

### *Publications in peer-reviewed international journals*

- **MURTIYOSO, A.** ; MATRONE, F. ; MARTINI, M. ; LINGUA, A. ; GRUSSENMEYER, P. ; PIERDICCA, R. 2020. "Automatic Training Data Generation for Deep Learning-Based Semantic Segmentation for Heritage Structures". *ISPRS International Journal of Geo-Information*. In preparation.
- **MURTIYOSO, A.** ; GRUSSENMEYER, P. 2020. "Virtual Disassembling of Historical Edifices: Experiments and Assessments of an Automatic Approach for Classifying Multi-Scalar Point Clouds into Architectural Elements". *Sensors*. Volume 20, Issue 8, 2161. <https://doi.org/10.3390/s20082161>.
- **MURTIYOSO, A.**; GRUSSENMEYER, P.; SUWARDHI, D.; AWALLUDIN, R. 2018. "Multi-Scale and Multi-Sensor 3D Documentation of Heritage Complexes in Urban Areas". *ISPRS International Journal of Geo-Information*, 7, 483. <https://doi.org/10.3390/ijgi7120483>.
- **MURTIYOSO, A.**; GRUSSENMEYER, P.; BÖRLIN, N.; VANDERMEERSCHEN, J.; FREVILLE, T. 2018. "Open Source and Independent Methods for Bundle Adjustment Assessment in Close-Range UAV Photogrammetry". *Drones*, Volume 2, Issue 1, 3. <https://doi.org/10.3390/drones2010003>.
- **MURTIYOSO, A.**; GRUSSENMEYER, P. 2017. "Documentation of heritage buildings using close range UAV images: dense matching issues, comparison, and case studies". *The Photogrammetric Record*, Volume 32, Issue 159. pp. 206-229. <https://doi.org/10.1111/phor.12197>.

### *Publications in peer-reviewed conference proceedings*

- **MURTIYOSO, A.**; GRUSSENMEYER, P.; SUWARDHI, D. 2019. "Technical Considerations in Low-Cost Heritage Documentation". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2/W17. pp. 225-232. <https://doi.org/10.5194/isprs-archives-XLII-2-W17-225-2019>.



- **MURTIYOSO**, A.; GRUSSENMEYER, P. 2019. "Automatic heritage building point cloud segmentation and classification using geometrical rules". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2/W15. pp. 821-827. <https://doi.org/10.5194/isprs-archives-XLII-2-W15-821-2019>.
- **MURTIYOSO**, A.; GRUSSENMEYER, P.; SUWARDHI, D.; FADILAH, W.A.; PERMANA, H.A.; WICAKSONO, D. 2019. "Multi-sensor 3D recording pipeline for the documentation of Javanese temples". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2/W15. pp. 829-834. <https://doi.org/10.5194/isprs-archives-XLII-2-W15-829-2019>.
- **MURTIYOSO**, A.; GRUSSENMEYER, P. 2019. "Point cloud segmentation and semantic annotation aided by GIS data for heritage complexes". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2/W9. pp. 523-528. <https://doi.org/10.5194/isprs-archives-XLII-2-W9-523-2019>.
- **MURTIYOSO**, A.; GRUSSENMEYER, P.; SUWARDHI, D.; SUMANTRI, D.; WAHYU, S.; PURNAMA, I.; MURTIYOSO, S.; ANDRYANA, G.; AWALLUDIN, R. 2018. "Digital Documentation Workflow and Challenges for Tropical Vernacular Architecture in the Case of the Kasepuhan Palace in Cirebon, Indonesia". 2018 3rd Digital Heritage International Congress (Digital Heritage) held jointly with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018). <https://doi.org/10.1109/DigitalHeritage.2018.8810090>.
- **MURTIYOSO**, A.; GRUSSENMEYER, P. 2018. "Comparison and Assessment of 3D Registration and Georeferencing Approaches for Exterior and Interior Point Clouds in the Case of Heritage Building Recording". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2. pp. 745-751. <https://doi.org/10.5194/isprs-archives-XLII-2-745-2018>.
- **MURTIYOSO**, A.; GRUSSENMEYER, P.; BÖRLIN, N. 2017. "Reprocessing Close Range Terrestrial and UAV Photogrammetric Projects with the DBAT Toolbox for Independent Verification and Quality Control". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2/W8. pp. 171-177. <https://doi.org/10.5194/isprs-archives-XLII-2-W8-171-2017>.
- **MURTIYOSO**, A.; KOEHL, M.; GRUSSENMEYER, P.; FREVILLE, T. 2017. "Acquisition and processing protocols for UAV images: 3D modelling of historical buildings using photogrammetry". *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume IV-8/W2. pp. 163-170. <https://doi.org/10.5194/isprs-annals-IV-2-W2-163-2017>.
- **MURTIYOSO**, A.; GRUSSENMEYER, P.; GUILLEMIN, S.; PRILAUD, G. 2017. "Centenary of the Battle of Vimy (France, 2017): Preserving the memory of the Great War through 3D recording of the Maison Blanche souterraine". *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume IV-8/W2. pp. 171-177. <https://doi.org/10.5194/isprs-annals-IV-2-W2-171-2017>.

**Book chapters**

- **MURTIYOSO, A.**; SUWARDHI, D. 2017. "Teknik Pencocokan Citra dalam Fotogrametri untuk Dokumentasi Cagar Budaya: Tinjauan Pustaka dan Contoh Kasus". Bunga Rampai ForMIND 2017. Wikantika, K., Ariadji, F.N., Prastiwi, P.A., Eds. Bandung: Penerbit ITB. pp. 242-260. ISBN:978-602-5417-37-5.

**Co-authorship of scientific papers (journals and proceedings)**

- BÖRLIN, N.; **MURTIYOSO, A.**; GRUSSENMEYER, P. 2020. "Efficient computation of posterior covariance in bundle adjustment in DBAT for projects with large number of object points". Proceedings of the XXIV ISPRS Congress, 14-20 June 2020, Nice, France. *In press*.
- MATRONE, F.; LINGUA, A.; PIERDICCA, R.; MALINVERNI, E.; PAOLANTI, M.; GRILLI, E.; REMONDINO, F.; LANDES, T.; **MURTIYOSO, A.** 2020. "A Benchmark for Large-Scale Heritage Point Cloud Classification". Proceedings of the XXIV ISPRS Congress, 14-20 June 2020, Nice, France. *In press*.
- GASICA, T.; BIORESITA, F.; **MURTIYOSO, A.** 2020. "Identification of temporary surface water using Sentinel-1 SAR data, case study: Sentani flash flooding, Indonesia". Proceedings of the XXIV ISPRS Congress, 14-20 June 2020, Nice, France. *In press*.
- ABATE, D.; **MURTIYOSO, A.** 2019. "Bundle Adjustment Accuracy Assessment of Unordered Aerial Dataset Collected through Kite Platform". The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W17. pp. 1-8. <https://doi.org/10.5194/isprs-archives-XLII-2-W17-1-2019>.
- BÖRLIN, N.; **MURTIYOSO, A.**; GRUSSENMEYER, P. 2019. "Implementing Functional Modularity for Processing of General Photogrammetric Data with the Damped Bundle Adjustment Toolbox (DBAT)". The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W17. pp. 69-75. <https://doi.org/10.5194/isprs-archives-XLII-2-W17-69-2019>.
- SUWARDHI, D.; TRISYANTI, S.; KAMAL, L.; PERMANA, H.; **MURTIYOSO, A.**; FAUZAN, K. 2019. "Polyfit Assisted Monoscopic Multi-Image Measurement Systems". The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W17. pp. 347-354. <https://doi.org/10.5194/isprs-archives-XLII-2-W17-347-2019>.
- TRISYANTI, S.; SUWARDHI, D.; **MURTIYOSO, A.**; GRUSSENMEYER, P. 2019. "Low Cost Web-Application for Management of 3D Digital Building and Complex Based on BIM and GIS". The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W17. pp. 371-375. <https://doi.org/10.5194/isprs-archives-XLII-2-W17-371-2019>.
- YANG, X.; LU, Y.; **MURTIYOSO, A.**; KOEHL, M.; GRUSSENMEYER, P. 2019. "HBIM Modeling from the Surface Mesh and Its Extended Capability of Knowledge Representation". *ISPRS International Journal of Geo-Information* 8(7):301. <https://doi.org/10.3390/ijgi8070301>.

- BÖRLIN, N.; MURTIYOSO, A. ; GRUSSENMEYER, P.; MENNA, F.; NOCERINO, E. 2019. "Flexible Photogrammetric Computations using Modular Bundle Adjustment". *Photogrammetric Engineering & Remote Sensing* Vol. 85, No. 5, May 2019, pp. 25-32. <https://doi.org/10.14358/PERS.85.5.361>. **2020 ASPRS Talbert Abrams Award First Honorable Mention.**
- SEMLER, Q.; SUWARDHI, D.; ALBY, E.; MURTIYOSO, A.; MACHER, H. 2019. "Registration of 2D plans on a 3D point cloud as support for the modeling of complex architectures". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2/W15. pp. 1083-1087. <https://doi.org/10.5194/isprs-archives-XLII-2-W15-1083-2019>.
- ASSI, R.; LANDES, T.; MURTIYOSO, A.; GRUSSENMEYER, P. 2019. "Assessment of descriptors for the registration of openings in indoor and outdoor heritage point clouds". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2/W15. pp. 133-138. <https://doi.org/10.5194/isprs-archives-XLII-2-W15-133-2019>.
- BÖRLIN, N.; MURTIYOSO, A. ; GRUSSENMEYER, P.; MENNA, F.; NOCERINO, E. 2018. "Modular Bundle Adjustment for Photogrammetric Computations". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2. pp. 133-140. <https://doi.org/10.5194/isprs-archives-XLII-2-133-2018>.
- LUMBAN-GAOL, Y.; MURTIYOSO, A.; NUGROHO, B. H. 2018. "Investigations on the Bundle Adjustment Results from SFM-based Software for Mapping Purposes". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2. pp. 699-705. <https://doi.org/10.5194/isprs-archives-XLII-2-623-2018>.
- MENNA, F.; NOCERINO, E.; DRAP, P.; REMONDINO, F.; MURTIYOSO A. ; GRUSSENMEYER, P.; BÖRLIN, N. 2018. "Improving Underwater Accuracy by Empirical Weighting of Image Observations". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-2. pp. 623-628. <https://doi.org/10.5194/isprs-archives-XLII-2-699-2018>.

## ACADEMIC SUPERVISING

### *Research and Technological Project (Master student)*

- PIRANDA, G. (INSA Strasbourg, France) Utilisation de caméras sphériques en photogrammétrie pour la génération de nuages de points denses. Defended in January 2019.
- PEGOSSOFF, P. (INSA Strasbourg, France) Comparaison de méthodes de recalage de données lasergrammétriques : utilisation combinée d'un scanner laser terrestre et d'un scanner à main. Co-supervision with Elise LACHAT (ICUBE-TRIO). Defended in January 2018.

### ***Master's thesis (Master student)***

- THIERRY, J. (INSA Strasbourg, France) Détection des fenêtres dans le nuage de points d'un mur de façade et positionnement semi-automatique dans un logiciel BIM. Thesis reviewer. Defended in September 2019.

### ***Bachelor's thesis (Bachelor student)***

- WICAKSONO, D.E. (ITB, Indonesia) Georegistrasi dan integrasi awan titik dari pemindai laser terestris dan fotogrametri untuk berbagai kedetilan objek. Studi kasus: Candi Sari. Co-supervision with Deni SUWARDHI (KKINSIG-ITB). Defended in July 2019.
- PERMANA, H.A. (ITB, Indonesia) Rekonstruksi permukaan polygon 3D dari awan titik menggunakan perangkat lunak PolyFit. Co-supervision with Deni SUWARDHI (KKINSIG-ITB). Defended in July 2019.
- NAFI'AH, G.R. (ITB, Indonesia) Mono dan stereoploting fotogrametri menggunakan foto udara miring. Co-supervision with Deni SUWARDHI (KKINSIG-ITB). Defended in June 2018.

## **TEACHING RESPONSIBILITIES**

- **5<sup>th</sup> year engineering students:** lectures and practical work in bundle adjustment with and without damping, practical work in drone photogrammetry.
- **4<sup>th</sup> year engineering students:** lectures and practical work in computer vision.
- **3<sup>rd</sup> and 2<sup>nd</sup> year engineering students:** practical work in stereo vision and photography.

## **OTHER SCIENTIFIC ACTIVITIES**

### ***Journal Reviewer***

- Aerospace (MDPI): <https://www.mdpi.com/journal/aerospace>
- Applied Sciences (MDPI): <https://www.mdpi.com/journal/applsci>
- Drones (MDPI): <https://www.mdpi.com/journal/drones>
- Environmental Earth Sciences (Springer): <https://link.springer.com/journal/12665>
- Heritage (MDPI): <https://www.mdpi.com/journal/heritage>
- International Journal of Geo-Information (MDPI): <https://www.mdpi.com/journal/ijgi>
- ISPRS Journal of Photogrammetry and Remote Sensing (Elsevier): <https://www.sciencedirect.com/journal/isprs-journal-of-photogrammetry-and-remote-sensing>

- Measurement (Elsevier): <https://www.sciencedirect.com/journal/measurement>
- Remote Sensing (MDPI): <https://www.mdpi.com/journal/remotesensing>
- The Photogrammetric Record (Wiley): <https://onlinelibrary.wiley.com/journal/14779730>

### *Scientific Committee for Conferences*

- XXIV ISPRS Congress, 14-20 June 2020 (**postponed to 2021 due to Covid-19 pandemic**), Nice, France: <http://www.isprs2020-nice.com/>
- 6<sup>th</sup> International Workshop LowCost 3D - Sensors, Algorithms, Applications, 2-3 December 2019, Strasbourg, France: <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W17/>
- 27<sup>th</sup> CIPA International Symposium, 1-5 September 2019, Avila, Spain: <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W15/>

### *Organisation of Scientific Events*

- 6<sup>th</sup> International Workshop LowCost 3D - Sensors, Algorithms, Applications, 2-3 December 2019, Strasbourg, France
- Nusantara 2019 Digital Heritage Tropical School, 3-9 August 2019, Sewu Temple Complex, Indonesia

# Bibliography

- ABATE, D. AND MURTIYOSO, A. (2019). Bundle adjustment accuracy assessment of un-ordered aerial dataset collected through Kite platform. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W17, pages 1–8. *Cited on page 34*
- ACHILLE, C., ADAMI, A., CHIARINI, S., CREMONESI, S., FASSI, F., FREGONESE, L., AND TAFFURELLI, L. (2015). UAV-based photogrammetry and integrated technologies for architectural applications - methodological strategies for the after-quake survey of vertical structures in Mantua (Italy). *Sensors*, 15(7):15520–15539. *4 citations on pages 9, 15, 18, and 30*
- ACHILLE, C., FASSI, F., AND FREGONESE, L. (2012). 4 Years history: From 2D to BIM for CH - The main spire on Milan Cathedral. In *Proceedings of the 2012 18th International Conference on Virtual Systems and Multimedia, VSMM 2012: Virtual Systems in the Information Society*, pages 377–382. *Cited on page 10*
- AI, M., HU, Q., LI, J., WANG, M., YUAN, H., AND WANG, S. (2015). A robust photogrammetric processing method of low-altitude UAV images. *Remote Sensing*, 7(3):2302–2333. *Cited on page 43*
- AKBAYLAR, I. AND HAMAMCIOGLU-TURAN, M. (2007). Documentation of a Vernacular House With Close-Range Digital Photogrammetry. In *XXI International CIPA Symposium, Athens, Greece*. *Cited on page 11*
- AMZIL, S. (2020). Détecteurs et descripteurs de points d'intérêt 3D. Technical report, INSA Strasbourg. *2 citations on pages 62 and 150*
- ANTONOPOULOS, A. AND ANTONOPOULOU, S. (2017). 3D survey and BIM-ready modelling of a Greek Orthodox Church in Athens. In *IMEKO International Conference on Metrology for Archaeology and Cultural Heritage, Lecce, Italy, October 23-25, 2017 3D*, Lecce, Italy. *2 citations on pages 69 and 95*
- ASSI, R., LANDES, T., MURTIYOSO, A., AND GRUSSENMEYER, P. (2019). Assessment of a Keypoints Detector for the Registration of Indoor and Outdoor Heritage Point Clouds. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W15, pages 133–138. *3 citations on pages 61, 115, and 150*
- BAIOCCHI, V., DOMINICI, D., AND MORMILE, M. (2013). UAV application in post-seismic environment. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-1/W2, pages 21–25. *2 citations on pages 9 and 18*

- BALLARD, D. (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122. *Cited on page 70*
- BANNING, E. (2002). *Archaeological Survey*. Springer Science & Business Media. *Cited on page 1*
- BARAZZETTI, L., BANFI, F., BRUMANA, R., ORENI, D., PREVITALI, M., AND RONCORONI, F. (2015). HBIM and augmented information: towards a wider user community of image and range-based reconstructions. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W7, pages 35–42. *Cited on page 2*
- BARAZZETTI, L., PREVITALI, M., AND RONCORONI, F. (2018). Can we use low-cost 360 degree cameras to create accurate 3D models? In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2, pages 69–75. *Cited on page 15*
- BARAZZETTI, L., PREVITALI, M., RONCORONI, F., AND VALENTE, R. (2019). Connecting inside and outside through 360° imagery for close-range photogrammetry. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W9, pages 87–92. *Cited on page 11*
- BARSANTI, S. G., REMONDINO, F., FENÁNDEZ-PALACIOS, B. J., AND VISINTINI, D. (2014). Critical factors and guidelines for 3D surveying and modelling in Cultural Heritage. *International Journal of Heritage in the Digital Era*, 3(1):141–158. *5 citations on pages 2, 11, 15, 31, and 67*
- BASSIER, M., BONDUEL, M., GENECHTEN, B. V., AND VERGAUWEN, M. (2017a). Octree-Based Region Growing and Conditional Random Fields. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W8, pages 25–30. *3 citations on pages 3, 68, and 81*
- BASSIER, M., VERGAUWEN, M., AND VAN GENECHTEN, B. (2017b). Automated Classification of Heritage Buildings for As-Built BIM using Machine Learning Techniques. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume IV-2/W2, pages 25–30. *3 citations on pages 68, 69, and 81*
- BASTONERO, P., DONADIO, E., CHIABRANDO, F., AND SPANÒ, A. (2014). Fusion of 3D models derived from TLS and image-based techniques for CH enhanced documentation. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume II-5, pages 73–80. *Cited on page 21*
- BAY, H., TUYTELAARS, T., AND VAN GOOL, L. (2006). SURF: Speeded up robust features. *Lecture Notes in Computer Science*, 3951 LNCS:404–417. *Cited on page 14*
- BEDFORD, J. (2017). *Photogrammetric Applications for Cultural Heritage*. Historic England, Swindon. *6 citations on pages 12, 14, 29, 32, 33, and 67*
- BELLEKENS, B., SPRUYT, V., BERKVEN, R., AND WEYN, M. (2014). A Survey of Rigid 3D Pointcloud Registration Algorithms. In *AMBIENT 2014, The Fourth International Conference on Ambient Computing, Applications, Services and Technologies*, pages 8–13. *Cited on page 53*

- 
- BESL, P. AND MCKAY, N. (1992). A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256. *2 citations on pages 18 and 53*
- BIAGINI, C. (2016). HBIM and fire prevention in historical building heritage management. In *33rd International Symposium on Automation and Robotics in Construction*. *Cited on page 2*
- BÖHM, J. (2005). Terrestrial Laser Scanning - A Supplementary Approach for 3D Documentation and Animation. In *Photogrammetric Week*, pages 263–271. *Cited on page 53*
- BÖRLIN, N. AND GRUSSENMEYER, P. (2013a). Bundle adjustment with and without damping. *The Photogrammetric Record*, 28(144):396–415. *4 citations on pages 4, 32, 33, and 34*
- BÖRLIN, N. AND GRUSSENMEYER, P. (2013b). Experiments with Metadata-derived Initial Values and Linesearch Bundle Adjustment in Architectural Photogrammetry. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume II-5/W1, pages 43–48. *2 citations on pages 33 and 34*
- BÖRLIN, N. AND GRUSSENMEYER, P. (2014). Camera Calibration using the Damped Bundle Adjustment Toolbox. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume II-5, pages 89–96. *2 citations on pages 4 and 34*
- BÖRLIN, N. AND GRUSSENMEYER, P. (2016). External Verification of the Bundle Adjustment in Photogrammetric Software Using the Damped Bundle Adjustment Toolbox. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLI-B5, pages 7–14. *3 citations on pages 30, 33, and 34*
- BÖRLIN, N., MURTIYOSO, A., AND GRUSSENMEYER, P. (2019a). Implementing functional modularity for processing of general photogrammetric data with the damped bundle adjustment toolbox (DBAT). In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W17, pages 69–75. *Cited on page 35*
- BÖRLIN, N., MURTIYOSO, A., GRUSSENMEYER, P., MENNA, F., AND NOCERINO, E. (2019b). Flexible Photogrammetric Computations Using Modular Bundle Adjustment: The Chain Rule and the Collinearity Equations. *Photogrammetric Engineering & Remote Sensing*, 85(5):361–368. *4 citations on pages 14, 34, 65, and 114*
- BOULAASSAL, H., LANDES, T., GRUSSENMEYER, P., AND KURDI, F. (2007). Automatic segmentation of building facades using terrestrial laser data. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVI, 3/W, pages 65–70. *3 citations on pages 60, 68, and 71*
- BROWN, N., LAING, R., AND SCOTT, J. (2009). The doocots of Aberdeenshire: An application of 3D scanning technology in the built heritage. *Journal of Building Appraisal*, 4(4):247–254. *Cited on page 11*
-



- BRYAN, P., BARBER, D., AND MILLS, J. (2004). Towards a Standard Specification for Terrestrial Laser Scanning in Cultural Heritage ? One Year on. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.  
Cited on page 1
- BURNS, J. H. R. AND DELPARTE, D. (2017). Comparison of Commercial Structure-From-Motion Photogrammetry Software Used for Underwater Three-Dimensional Modeling of Coral Reef Environments. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W3, pages 127–131.  
5 citations on pages 18, 29, 30, 31, and 33
- CAMPANARO, D. M., LANDESCI, G., DELL'UNTO, N., AND LEANDER TOUATI, A. M. (2016). 3D GIS for cultural heritage restoration: A 'white box' workflow. *Journal of Cultural Heritage*, 18:321–332.  
Cited on page 94
- CEFALU, A., ABDEL-WAHAB, M., PETER, M., WENZEL, K., AND FRITSCH, D. (2013). Image based 3D Reconstruction in Cultural Heritage Preservation. In *Proceedings of the 10th International Conference on Informatics in Control, Automation and Robotics*, pages 201–205, Reykjavík.  
Cited on page 11
- CHIABRANDO, F., DONADIO, E., AND RINAUDO, F. (2015). SfM for orthophoto generation: a winning approach for cultural heritage knowledge. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W7, pages 91–98.  
7 citations on pages 11, 18, 30, 33, 34, 35, and 80
- CHIABRANDO, F., LO TURCO, M., AND SANTAGATI, C. (2017). Digital invasions: From point clouds to Historical Building Object Modeling (H-BOM) of a UNESCO WHL site. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W3, pages 171–178.  
Cited on page 11
- CHIABRANDO, F., SAMMARTANO, G., AND SPANÒ, A. (2016). Historical Buildings Models and Their Handling Via 3D Survey: From Points Clouds To User-Oriented HBIM. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLI-B5, pages 633–640.  
Cited on page 2
- COLOMINA, I. AND MOLINA, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92:79–97.  
Cited on page 18
- CRAMER, M. (2013). The UAV @ LGL BW project - a NMCA case study. In *Photogrammetric Week*, pages 165–179.  
Cited on page 18
- DA COSTA SALAVESSA, M. E. (2012). Historical Timber-Framed Buildings: Typology and Knowledge. *Journal of Civil Engineering and Architecture*, 6(2):151–166.  
Cited on page 97
- DALL'ASTA, E., THOENI, K., SANTISE, M., FORLANI, G., GIACOMINI, A., AND RONCELLA, R. (2015). Network design and quality checks in automatic orientation of close-range photogrammetric blocks. *Sensors*, 15(4):7985–8008.  
Cited on page 34

- DEWEZ, T. J., GIRARDEAU-MONTAUT, D., ALLANIC, C., AND ROHMER, J. (2016). Facets : A cloudcompare plugin to extract geological planes from unstructured 3d point clouds. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLI-B5, pages 799–804. *Cited on page 99*
- DGAC (2015). Arrêté du 17 décembre 2015 relatif à l'utilisation de l'espace aérien par les aéronefs qui circulent sans personne à bord. *Cited on page 42*
- DORE, C. AND MURPHY, M. (2013). Semi-Automatic Modelling of Building Façades With Shape Grammars Using Historic Building Information Modelling. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W1, pages 57–64. *Cited on page 3*
- DORE, C., MURPHY, M., MCCARTHY, S., BRECHIN, F., CASIDY, C., AND DIRIX, E. (2015). Structural simulations and conservation analysis-historic building information model (HBIM). In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5-W4, pages 351–357. *2 citations on pages 2 and 3*
- DRAP, P., PAPINI, O., PRUNO, E., NUCCIOTTI, M., AND VANNINI, G. (2017). Ontology-based photogrammetry survey for medieval archaeology: Toward a 3D geographic information system (GIS). *Geosciences*, 7(4):1–33. *2 citations on pages 69 and 80*
- DUDA, R. O. AND HART, P. E. (1972). Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Communications of the ACM*, 15(1):11–15. *Cited on page 70*
- ELIZABETH, O. AND PRIZEMAN, C. (2015). HBIM and matching techniques : considerations for late nineteenth- and early twentieth-century buildings. *Journal of Architectural Conservation*, 21(September):145–159. *Cited on page 3*
- EVGENIKOU, V. AND GEORGOPOULOS, A. (2015). Investigating 3D reconstruction methods for small artifacts. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W4, pages 101–108. *Cited on page 11*
- FABADO, S., SEGUÍ, A. E., CABRELLES, M., NAVARRO, S., GARCÍA-DE-SAN-MIGUEL, D., AND LERMA, J. L. (2013). 3DVEM Software Modules for Efficient Management of Point Clouds and Photorealistic 3D Models. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W2, pages 255–260. *2 citations on pages 17 and 56*
- FABBRI, K., ZUPPIROLI, M., AND AMBROGIO, K. (2012). Heritage buildings and energy performance: Mapping with GIS tools. *Energy and Buildings*, 48:137–145. *Cited on page 10*
- FABBRI, S., SAURO, F., SANTAGATA, T., ROSSI, G., AND DE WAELE, J. (2017). High-resolution 3-D mapping using terrestrial laser scanning as a tool for geomorphological and speleogenetical studies in caves: An example from the Lessini mountains (North Italy). *Geomorphology*, 280:16–29. *Cited on page 80*

- FANGI, G. (2019). Aleppo - Before and after. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII, pages 333–338. *2 citations on pages 15 and 67*
- FARELLA, E. M., TORRESANI, A., AND REMONDINO, F. (2019). Quality Features for the Integration of Terrestrial and UAV Images. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W9, pages 339–346. *3 citations on pages 21, 67, and 68*
- FASSI, F., ACHILLE, C., AND FREGONESE, L. (2011). Surveying and modelling the main spire of Milan Cathedral using multiple data sources. *The Photogrammetric Record*, 26(136):462–487. *Cited on page 52*
- FASSI, F., FREGONESE, L., ADAMI, A., AND RECHICHI, F. (2017). BIM System for the Conservation and Preservation of the Mosaics of San Marco in Venice. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W5, pages 229–236. *Cited on page 10*
- FIORILLO, F., JIMÉNEZ FERNÁNDEZ-PALACIOS, B., REMONDINO, F., AND BARBA, S. (2013). 3d Surveying and modelling of the Archaeological Area of Paestum, Italy. *Virtual Archaeology Review*, 4(8):55–60. *4 citations on pages 10, 66, 67, and 75*
- FISCHLER, M. AND BOLLES, R. (1980). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Technical Report 213, SRI International. *Cited on page 71*
- FLETCHER, R., JOHNSON, I., BRUCE, E., AND KHUN-NEAY, K. (2007). Living with heritage: Site monitoring and heritage values in Greater Angkor and the Angkor World Heritage Site, Cambodia. *World Archaeology*, 39(3):385–405. *Cited on page 80*
- FORKUO, E. AND KING, B. (2004). Automatic fusion of photogrammetric imagery and laser scanner point clouds. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. *Cited on page 54*
- FRITSCH, D., BECKER, S., AND ROTHERMEL, M. (2013). Modeling Façade Structures Using Point Clouds From Dense Image Matching. *Proceedings of the Intl. Conf. on Advances in Civil, Structural and Mechanical Engineering*, pages 57–64. *Cited on page 52*
- FURUKAWA, Y. AND PONCE, J. (2009). Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1362–1376. *2 citations on pages 14 and 30*
- GERKE, M. AND PRZYBILLA, H.-J. (2016). Accuracy Analysis of Photogrammetric UAV Image Blocks: Influence of Onboard RTK-GNSS and Cross Flight Patterns. *Photogrammetrie - Fernerkundung - Geoinformation*, 2016(1):17–30. *Cited on page 43*
- GONZÁLEZ-AGUILERA, D., LÓPEZ-FERNÁNDEZ, L., RODRIGUEZ-GONZÁLEZ, P., GUERRERO, D., HERNÁNDEZ-LOPEZ, D., REMONDINO, F., MENNA, F., NOCERINO, E., TOSCHI, I., BALLABENI, A., AND GAIANI, M. (2016). Development of an all-purpose free photogrammetric tool. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 41, pages 31–38. *2 citations on pages 31 and 33*

- GRANSHAW, S. I. (1980). Bundle Adjustment Methods in Engineering Photogrammetry. *The Photogrammetric Record*, 10(56):181–207. *Cited on page 32*
- GRANSHAW, S. I. (2016). Photogrammetric Terminology: Third Edition. *The Photogrammetric Record*, 31(154):210–252. *2 citations on pages 31 and 32*
- GRENZDÖRFFER, G. J., NAUMANN, M., NIEMEYER, F., AND FRANK, A. (2015). Symbiosis of UAS Photogrammetry and TLS for Surveying and 3D Modeling of Cultural Heritage Monuments - a Case Study About the Cathedral of St. Nicholas in the City of Greifswald. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-1/W4, pages 91–96. *5 citations on pages 10, 15, 31, 67, and 80*
- GRILLI, E., DININNO, D., PETRUCCI, G., AND REMONDINO, F. (2018). From 2D to 3D supervised segmentation and classification for cultural heritage applications. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2, pages 399–406. *2 citations on pages 3 and 81*
- GRILLI, E., MENNA, F., AND REMONDINO, F. (2017). A Review of Point Clouds Segmentation and Classification Algorithms. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W3, pages 339–344. *Cited on page 68*
- GRILLI, E., ÖZDEMİR, E., AND REMONDINO, F. (2019). Application of Machine and Deep Learning Strategies for the Classification of Heritage Point Clouds. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-4/W18, pages 447–454. *Cited on page 69*
- GRILLI, E. AND REMONDINO, F. (2019). Classification of 3D digital heritage. *Remote Sensing*, 11(7):1–23. *2 citations on pages 68 and 69*
- GRUEN, A. (1985). Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*, 14(3):175–187. *Cited on page 13*
- GRUSSENMEYER, P. AND AL KHALIL, O. (2002). Solutions for exterior orientation in photogrammetry: a review. *The Photogrammetric Record*, 17(May):615–634. *2 citations on pages 12 and 31*
- GRUSSENMEYER, P. AND AL KHALIL, O. (2017). From metric image archives to point cloud reconstruction: case study of the Great Mosque of Aleppo in Syria. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W5, pages 295–301. *Cited on page 9*
- GRUSSENMEYER, P., ALBY, E., LANDES, T., KOEHL, M., GUILLEMIN, S., HULLO, J. F., ASSALI, P., AND SMIGIEL, E. (2012a). Recording approach of heritage sites based on merging point clouds from high resolution photogrammetry and Terrestrial Laser Scanning. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 39, pages 553–558. *Cited on page 10*

- GRUSSENMEYER, P., BURENS, A., MOISAN, E., GUILLEMIN, S., CAROZZA, L., BOURRILLON, R., AND PETROGNANI, S. (2012b). 3D multi-scale scanning of the archaeological cave les Fraux in Dordogne (France). *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7616 LNCS:388–395. *Cited on page 52*
- GRUSSENMEYER, P., HANKE, K., AND STREILEIN, A. (2002). Architectural Photogrammetry. In Kasser, M. and Egels, Y., editors, *Digital Photogrammetry*, pages 300–339. Taylor & Francis. *2 citations on pages 1 and 11*
- GRUSSENMEYER, P., LANDES, T., ALBY, E., AND CAROZZA, L. (2010). High Resolution 3D Recording and Modelling of the Bronze Age Cave "Les Fraux" in Perigord (France). In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVIII, pages 262–267. *2 citations on pages 21 and 31*
- GRUSSENMEYER, P., LANDES, T., DONEUS, M., AND LERMA, J. L. (2016). Basics of Range-based Modelling Techniques in Cultural Heritage 3D Recording. In Stylianidis, E. and Remondino, F., editors, *3D Recording, Documentation and Management of Cultural Heritage*, pages 305–368. Whittles Publishing. *3 citations on pages 15, 16, and 17*
- GRUSSENMEYER, P., LANDES, T., VOEGTLE, T., AND RINGLE, K. (2008). Comparison methods of terrestrial laser scanning, photogrammetry and tachometry data for recording of cultural heritage buildings. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Vol. XXXVI, pages 213–218. *Cited on page 21*
- HANAN, H., SUWARDHI, D., NURHASANAH, T., AND BUKIT, E. S. (2015). Batak Toba Cultural Heritage and Close-range Photogrammetry. *Procedia - Social and Behavioral Sciences*, 184(August 2014):187–195. *Cited on page 15*
- HANKE, K., GRUSSENMEYER, P., GRIMM-PITZINGER, A., AND WEINOLD, T. (2006). First Experiences with the Trimble GX Scanner. In *ISPRS Comm. V Symposium*, pages 1–6. *Cited on page 1*
- HÄNSCH, R., WEBER, T., AND HELLWICH, O. (2014). Comparison of 3D interest point detectors and descriptors for point cloud fusion. In *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume II-3, pages 57–64. *Cited on page 54*
- HARRIS, C. AND STEPHENS, M. (1988). A Combined Corner and Edge Detector. *Proceedings of the Alvey Vision Conference 1988*, pages 147–151. *Cited on page 14*
- HASSANI, F. (2015). Documentation of cultural heritage techniques, potentials and constraints. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W7, pages 207–214. *2 citations on pages 15 and 67*
- HÉNO, R. AND CHANDELIER, L. (2014). *Numérisation 3D de bâtiments: cas des édifices remarquables*. Hermes Science Publishing, London. *Cited on page 16*

- HERBIG, U., STAMPFER, L., GRANDITS, D., MAYER, I., PÖCHTRAGER, M., AND SETYASTUTI, A. (2019). Developing a Monitoring Workflow for the Temples of Java. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W15, pages 555–562. 2 citations on pages 11 and 67
- HILLEMANN, M., WEINMANN, M., MUELLER, M. S., AND JUTZI, B. (2019). Automatic extrinsic self-calibration of mobile mapping systems based on geometric 3D features. *Remote Sensing*, 11(16):1–29. 2 citations on pages 1 and 54
- HIRSCHMÜLLER, H. (2005). Accurate and efficient stereo processing by semi-global matching and mutual information. *IEEE International Conference on Computer Vision and Pattern Recognition*, 2(2):807–814. 3 citations on pages 14, 30, and 34
- HIRSCHMÜLLER, H. (2011). Semi-Global Matching Motivation, Developments and Applications. In *Photogrammetric Week*, pages 173–184. Cited on page 14
- HOLZ, D., ICHIM, A. E., TOMBARI, F., RUSU, R. B., AND BEHNKE, S. (2015). Registration with the Point Cloud Library: A Modular Framework for Aligning in 3-D. *IEEE Robotics & Automation Magazine*, pages 110–124. Cited on page 54
- HOUGH, P. (1962). A method and means for recognition complex patterns. US Patent No. US3069654A. Cited on page 70
- JAMES, M. R., ROBSON, S., D’OLEIRE-OLTMANN, S., AND NIETHAMMER, U. (2017a). Optimising UAV topographic surveys processed with structure-from-motion: Ground control quality, quantity and bundle adjustment. *Geomorphology*, 280:51–66. Cited on page 31
- JAMES, M. R., ROBSON, S., AND SMITH, M. W. (2017b). 3-D uncertainty-based topographic change detection with structure-from-motion photogrammetry: precision maps for ground control and directly georeferenced surveys. *Earth Surface Processes and Landforms*, 42(12):1769–1788. Cited on page 31
- JAUD, M., PASSOT, S., LE BIVIC, R., DELACOURT, C., GRANDJEAN, P., AND LE DANTEC, N. (2016). Assessing the accuracy of high resolution digital surface models computed by PhotoScan® and MicMac® in sub-optimal survey conditions. *Remote Sensing*, 8(6):1–18. 2 citations on pages 30 and 31
- KAISER, P., WEGNER, J. D., LUCCHI, A., JAGGI, M., HOFMANN, T., AND SCHINDLER, K. (2017). Learning Aerial Image Segmentation from Online Maps. *IEEE Transactions on Geoscience and Remote Sensing*, 55(11):6054 – 6068. Cited on page 81
- KASTUARI, A., SUWARDHI, D., HANAN, H., AND WIKANTIKA, K. (2016). State of the Art of the Landscape Architecture Spatial Data Model From a Geospatial Perspective. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume IV-2/W1, pages 63–71. Cited on page 80
- KHODEIR, L. M., ALY, D., AND TAREK, S. (2016). Integrating HBIM (Heritage Building Information Modeling) Tools in the Application of Sustainable Retrofitting of Heritage Buildings in Egypt. *Procedia Environmental Sciences*, 34:258–270. Cited on page 2
- KIM, E. AND MEDIONI, G. (2011). Urban scene understanding from aerial and ground LIDAR data. *Machine Vision and Applications*, 22(4):691–703. Cited on page 80

- KIM, J., LEE, S., AHN, H., SEO, D., PARK, S., AND CHOI, C. (2013). Feasibility of employing a smartphone as the payload in a photogrammetric UAV system. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79:1–18. 2 citations on pages 11 and 43
- KRAUS, K. AND WALDHÄUSL, P. (1998). *Manuel de photogrammétrie*. Hermes, Paris. 2 citations on pages 37 and 42
- LACHAT, E., LANDES, T., AND GRUSSENMEYER, P. (2016). Combination of TLS Point Clouds and 3D Data From Kinect V2 Sensor To Complete Indoor Models. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLI-B5, pages 12–19. 4 citations on pages 2, 3, 15, and 54
- LACHAT, E., LANDES, T., AND GRUSSENMEYER, P. (2017). First Experiences with the Trimble SX10 Scanning Total Station for Building Facade Survey. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W3, pages 405–412, Nafplio, Greece. 4 citations on pages 1, 10, 11, and 53
- LACHAT, E., LANDES, T., AND GRUSSENMEYER, P. (2018). Comparison of Point Cloud Registration Algorithms for Better Result Assessment - Towards an Open-Source Solution. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2, pages 551–558. 3 citations on pages 1, 10, and 17
- LACHAT, E., MACHER, H., LANDES, T., AND GRUSSENMEYER, P. (2015). Assessment and calibration of a RGB-D camera (Kinect v2 Sensor) towards a potential use for close-range 3D modeling. *Remote Sensing*, 7(10):13070–13097. Cited on page 11
- LERMA, J. L., NAVARRO, S., CABRELLES, M., AND VILLAVERDE, V. (2010). Terrestrial laser scanning and close range photogrammetry for 3D archaeological documentation: the Upper Palaeolithic Cave of Parpallo as a case study. *Journal of Archaeological Science*, 37(3):499–507. Cited on page 10
- LIU, C.-J., KRYLOV, V., AND DAHYOT, R. (2018). 3D point cloud segmentation using GIS. In *20th Irish Machine Vision and Image Processing Conference*, pages 41–48. Cited on page 81
- LIU, W., SUN, J., LI, W., HU, T., AND WANG, P. (2019). Deep learning on point clouds and its application: A survey. *Sensors*, 19(19):1–22. 2 citations on pages 68 and 69
- LOWE, D. G. (2004). Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110. 4 citations on pages 14, 30, 35, and 61
- LU, Y. C., SHIH, T. Y., AND YEN, Y. N. (2018). Research on Historic BIM of Built Heritage in Taiwan -a Case Study of Huangxi Academy. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2, pages 615–622. Cited on page 69
- LUHMANN, T., ROBSON, S., KYLE, S., AND BOEHM, J. (2014). *Close-Range Photogrammetry and 3D Imaging*. De Gruyter, 2nd edition. 3 citations on pages xv, 12, and 30
- LUMBAN-GAOL, Y., MURTIYOSO, A., AND NUGROHO, B. (2018). Investigations on the bundle adjustment results from sfm-based software for mapping purposes. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2. Cited on page 34

- LUO, D. AND WANG, Y. (2008). Rapid extracting pillars by slicing point clouds. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVII, pages 215–218. *3 citations on pages 68, 94, and 96*
- MAALEK, R., LICHTI, D. D., AND RUWANPURA, J. Y. (2019). Automatic recognition of common structural elements from point clouds for automated progress monitoring and dimensional quality control in reinforced concrete construction. *Remote Sensing*, 11(9). *5 citations on pages 67, 68, 69, 94, and 95*
- MACHER, H., LANDES, T., AND GRUSSENMEYER, P. (2015). Point clouds segmentation as base for as-built BIM creation. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume II-5/W3, pages 191–197. *3 citations on pages 1, 3, and 69*
- MACHER, H., LANDES, T., AND GRUSSENMEYER, P. (2016). Validation of Point Clouds Segmentation Algorithms through their Application to Several Case Studies for Indoor Building Modelling. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLI, pages 12–19. *2 citations on pages 3 and 72*
- MACHER, H., LANDES, T., AND GRUSSENMEYER, P. (2017). From Point Clouds to Building Information Models: 3D Semi-Automatic Reconstruction of Indoors of Existing Buildings. *Applied Sciences*, 7(1030):1–30. *9 citations on pages 5, 67, 68, 69, 73, 94, 95, 96, and 99*
- MAGDA RAMOS, M. AND REMONDINO, F. (2015). Data fusion in cultural heritage - A review. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W7, pages 359–363. *Cited on page 11*
- MALINVERNI, E. S., PIERDICCA, R., PAOLANTI, M., MARTINI, M., MORBIDONI, C., MATRONE, F., AND LINGUA, A. (2019). Deep learning for semantic segmentation of point cloud. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W15, pages 735–742. *8 citations on pages xviii, 68, 69, 81, 94, 108, 153, and 154*
- MEGAHED, N. A. (2015). Towards a Theoretical Framework for HBIM Approach in Historic Presevation and Management. *International Journal of Architectural Research*, 9(3):130–147. *Cited on page 2*
- MENNA, F., NOCERINO, E., DRAP, P., REMONDINO, F., MURTIYOSO, A., GRUSSENMEYER, P., AND BÖRLIN, N. (2018). Improving underwater accuracy by empirical weighting of image observations. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2. *2 citations on pages 14 and 34*
- MENOU, J.-C. (2019). Requiem pour la charpente de Notre-Dame de Paris. *Commentaire*, 166(2):395–397. *2 citations on pages 9 and 97*
- MOE, D., SAMPATH, A., CHRISTOPHERSON, J., AND BENSON, M. (2010). Self calibration of small and medium format digital cameras. In Wagner, W. and Székely, B., editors, *ISPRS TC VII Symposium 100 Years ISPRS*, volume XXXVIII-7B, pages 395–400. *Cited on page 13*



- MUNUMER, E. AND LERMA, J. L. (2015). Fusion of 3D data from different image-based and range-based sources for efficient heritage recording. In *2015 Digital Heritage*, volume 304, pages 83–86. *5 citations on pages 4, 11, 21, 54, and 67*
- MURPHY, M., MCGOVERN, E., AND PAVIA, S. (2009). Historic building information modelling (HBIM). *Structural Survey*, 27(4):311–327. *Cited on page 2*
- MURPHY, M., MCGOVERN, E., AND PAVIA, S. (2013). Historic Building Information Modelling - Adding intelligence to laser and image based surveys of European classical architecture. *ISPRS Journal of Photogrammetry and Remote Sensing*, 76:89–102. *2 citations on pages 2 and 95*
- MURTIYOSO, A. AND GRUSSENMEYER, P. (2017). Documentation of heritage buildings using close-range UAV images: dense matching issues, comparison and case studies. *The Photogrammetric Record*, 32(159):206–229. *4 citations on pages 15, 23, 30, and 36*
- MURTIYOSO, A. AND GRUSSENMEYER, P. (2018). Comparison and assessment of 3D registration and georeferencing approaches of point clouds in the case of exterior and interior heritage building recording. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2, pages 745–751. *Cited on page 52*
- MURTIYOSO, A. AND GRUSSENMEYER, P. (2019a). Automatic Heritage Building Point Cloud Segmentation and Classification Using Geometrical Rules. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W15, pages 821–827. *Cited on page 94*
- MURTIYOSO, A. AND GRUSSENMEYER, P. (2019b). Point cloud segmentation and semantic annotation aided by GIS data for heritage complexes. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII, pages 523–528. *Cited on page 79*
- MURTIYOSO, A., GRUSSENMEYER, P., AND BÖRLIN, N. (2017a). Reprocessing Close Range Terrestrial and UAV Photogrammetric Projects with the DBAT Toolbox for Independent Verification and Quality Control. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2 W8, pages 171–177. *4 citations on pages 19, 29, 30, and 34*
- MURTIYOSO, A., GRUSSENMEYER, P., BÖRLIN, N., VANDERMEERSCHEN, J., AND FREVILLE, T. (2018a). Open Source and Independent Methods for Bundle Adjustment Assessment in Close-Range UAV Photogrammetry. *Drones*, 2(1):3. *Cited on page 29*
- MURTIYOSO, A., GRUSSENMEYER, P., AND FREVILLE, T. (2017b). Close Range UAV Accurate Recording and Modeling of St-Pierre-le-Jeune Neo-Romanesque Church in Strasbourg (France). In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W3, pages 519–526. *Cited on page 36*
- MURTIYOSO, A., GRUSSENMEYER, P., GUILLEMIN, S., AND PRILAUX, G. (2017c). Centenary of the Battle of Vimy (France, 1917): Preserving the Memory of the Great War through 3D recording of the Maison Blanche souterraine. In *ISPRS Annals of the*

- Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume IV-2/W2, pages 171–177. *2 citations on pages 21 and 67*
- MURTIYOSO, A., GRUSSENMEYER, P., KOEHL, M., AND FREVILLE, T. (2016). Acquisition and Processing Experiences of Close Range UAV Images for the 3D Modeling of Heritage Buildings. In Ioannides, M., Fink, E., Moropoulou, A., Hagedorn-Saupe, M., Fresa, A., Liestøl, G., Rajcic, V., and Grussenmeyer, P., editors, *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection: 6th International Conference, EuroMed 2016, Nicosia, Cyprus, October 31 – November 5, 2016, Proceedings, Part I*, pages 420–431. Springer International Publishing. *5 citations on pages 1, 10, 18, 30, and 80*
- MURTIYOSO, A., GRUSSENMEYER, P., AND SUWARDHI, D. (2019). Technical Considerations in Low-Cost Heritage Documentation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W17(December):225–232. *Cited on page 15*
- MURTIYOSO, A., GRUSSENMEYER, P., SUWARDHI, D., AND AWALLUDIN, R. (2018b). Multi-Scale and Multi-Sensor 3D Documentation of Heritage Complexes in Urban Areas. *ISPRS International Journal of Geo-Information*, 7(12):483. *6 citations on pages 10, 11, 65, 66, 68, and 80*
- MURTIYOSO, A., KOEHL, M., GRUSSENMEYER, P., AND FREVILLE, T. (2017d). Acquisition and Processing Protocols for UAV Images: 3D Modeling of Historical Buildings using Photogrammetry. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume IV-2/W2, pages 163–170. *2 citations on pages 11 and 33*
- NAN, L. AND WONKA, P. (2017). Polyfit: Polygonal surface reconstruction from point clouds. In *ICCV*. *Cited on page 68*
- NEX, F. AND REMONDINO, F. (2014). UAV: platforms, regulations, data acquisition and processing. In Remondino, F. and Campana, S., editors, *3D Recording and Modelling in Archaeology and Cultural Heritage: Theory and Best Practices*, chapter Photogramm, pages 73–86. Archaeopress, Oxford, England. *2 citations on pages 11 and 18*
- NGUYEN, A. AND LE, B. (2013). 3D Point Cloud Segmentation : A survey. In *2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, pages 225–230. IEEE. *Cited on page 68*
- NIETO, J. E., MOYANO, J. J., RICO DELGADO, F., AND ANTÓN GARCÍA, D. (2016). Management of built heritage via HBIM Project: A case of study of flooring and tiling. *Virtual Archaeology Review*, 7(14):1–12. *Cited on page 2*
- NOCERINO, E., LAGO, F., MORABITO, D., REMONDINO, F., PORZI, L., POIESI, F., ROTA BULO, S., CHIPPENDALE, P., LOCHER, A., HAVLENA, M., VAN GOOL, L., EDER, M., FÖTSCHL, A., HILSMANN, A., KAUSCH, L., AND EISERT, P. (2017a). A smartphone-based 3D pipeline for the creative industry - The replicate eu project. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W3, pages 535–541. *2 citations on pages 31 and 33*

- NOCERINO, E., POIESI, F., LOCHER, A., TEFERA, Y., REMONDINO, F., CHIPPENDALE, P., AND VAN GOOL, L. (2017b). 3D Reconstruction with a Collaborative Approach Based on Smartphones and a Cloud-Based Server. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W8(November):187–194. *Cited on page 11*
- OMIDALIZARANDI, M. AND SAADATSERESHT, M. (2013). Segmentation and classification of point clouds from dense aerial image matching. *The International Journal of Multimedia & Its Applications*, 5(4):33–51. *Cited on page 80*
- ORENI, D., BRUMANA, R., DELLA TORRE, S., BANFI, F., BARAZZETTI, L., AND PREVITALI, M. (2014). Survey turned into HBIM: the restoration and the work involved concerning the Basilica di Collemaggio after the earthquake (L’Aquila). In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume II-5, pages 267–273. *Cited on page 3*
- ORENI, D., BRUMANA, R., GEORGOPOULOS, A., AND CUCA, B. (2013). HBIM for Conservation and Management of Built Heritage: Towards a Library of Vaults and Wooden Beam Floors. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume II-5/W1, pages 215–221. *Cited on page 3*
- OUÉDRAOGO, M. M., DEGRÉ, A., DEBOUCHE, C., AND LISEIN, J. (2014). The evaluation of unmanned aerial system-based photogrammetry and terrestrial laser scanning to generate DEMs of agricultural watersheds. *Geomorphology*, 214:339–355. *Cited on page 31*
- PEGOSSOFF, P. (2018). Comparaison de methods de recalage de données lasergramétriques : utilisation combine d’un scanner laser terrestre et d’un scanner à main. Technical report, INSA Strasbourg. *2 citations on pages 61 and 150*
- PERFETTI, L., POLARI, C., FASSI, F., TROISI, S., BAIOCCHI, V., DEL PIZZO, S., GIANNONE, F., BARAZZETTI, L., PREVITALI, M., AND RONCORONI, F. (2018). Fisheye Photogrammetry to Survey Narrow Spaces in Architecture and a Hypogea Environment. In *Latest Developments in Reality-Based 3D Surveying and Modelling*, pages 3–28. MDPI. *2 citations on pages 11 and 15*
- PIERROT-DESEILLIGNY, M. AND CLERY, I. (2012). Apero, an Open Source Bundle Adjustment Software for Automatic Calibration and Orientation of Set of Images. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVIII, pages 269–276. *2 citations on pages 30 and 35*
- PIERROT-DESEILLIGNY, M. AND PAPANODITIS, N. (2006). A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVI. *Cited on page 35*
- PÖCHTRAGER, M., STYHLER-AYDIN, G., DÖRING-WILLIAMS, M., AND PFEIFER, N. (2017). Automated reconstruction of historic roof structures from point clouds - development and examples. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W2:195–202. *3 citations on pages 97, 99, and 100*

- PÖCHTRAGER, M., STYHLER-AYDIN, G., DÖRING-WILLIAMS, M., AND PFEIFER, N. (2018). Digital reconstruction of historic roof structures: developing a workflow for a highly automated analysis. *Virtual Archaeology Review*, 9(19):21. *Cited on page 97*
- POUX, F., NEUVILLE, R., AND BILLEN, R. (2017). Point cloud classification of tesserae from terrestrial laser data combined with dense image matching for archaeological information extraction. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume IV-2/W2, pages 203–211. *Cited on page 81*
- POUX, F., NEUVILLE, R., NYS, G. A., AND BILLEN, R. (2018). 3D point cloud semantic modelling: Integrated framework for indoor spaces and furniture. *Remote Sensing*, 10(1412):1–35. *Cited on page 69*
- PRAMULYO, H., HARTO, A., MERTOTAROENO, S., AND MURTIYOSO, A. (2017). Towards better 3D model accuracy with spherical photogrammetry. In Ivan, I., Singleton, A., Horák, J., and Inspektor, T., editors, *The Rise of Big Spatial Data - Lecture Notes in Geoinformation and Cartography*, pages 107–120. Springer International Publishing. *Cited on page 15*
- PU, S. AND VOSSELMAN, G. (2006). Automatic extraction of building features from terrestrial laser scanning. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36, pages 25–27. *Cited on page 3*
- QUATTRINI, R., MALINVERNI, E. S., CLINI, P., NESPECA, R., AND ORLIETTI, E. (2015). From TLS to HBIM. High quality semantically-aware 3d modeling of complex architecture. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-5/W4, pages 367–374. *3 citations on pages 2, 11, and 52*
- RAMOS, A. P. AND PRIETO, G. R. (2016). Only image based for the 3D metric survey of gothic structures by using frame cameras and panoramic cameras. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 41, pages 363–370. *Cited on page 15*
- REMONDINO, F. (2011). Heritage recording and 3D modeling with photogrammetry and 3D scanning. *Remote Sensing*, 3(6):1104–1138. *2 citations on pages 30 and 31*
- REMONDINO, F. (2014). Photogrammetry - Basic theory. In Remondino, F. and Campana, S., editors, *3D Recording and Modelling in Archaeology and Cultural Heritage: Theory and Best Practices*, pages 63–72. Archaeopress, Oxford, England. *4 citations on pages xv, 13, 14, and 31*
- REMONDINO, F., BARAZZETTI, L., NEX, F., SCAIONI, M., AND SARAZZI, D. (2011). UAV photogrammetry for mapping and 3D modeling - current status and future perspectives. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVIII, pages 25–31. *Cited on page 18*
- REMONDINO, F., DEL PIZZO, S., KERSTEN, T. P., AND TROISI, S. (2012). Low-cost and open-source solutions for automated image orientation - a critical overview. *Progress in Cultural Heritage Preservation*, 7616 LNCS:40–54. *4 citations on pages 10, 12, 31, and 32*

- REMONDINO, F. AND RIZZI, A. (2010). Reality-based 3D documentation of natural and cultural heritage sites-techniques, problems, and examples. *Applied Geomatics*, 2(3):85–100. *2 citations on pages 1 and 67*
- REMONDINO, F., SPERA, M. G., NOCERINO, E., MENNA, F., AND NEX, F. (2014). State of the art in high density image matching. *The Photogrammetric Record*, 29(146):144–166. *2 citations on pages 11 and 34*
- REMONDINO, F., SPERA, M. G., NOCERINO, E., MENNA, F., NEX, F., AND GONIZZI-BARSANTI, S. (2013). Dense image matching: Comparisons and analyses. In *Proceedings of the Digital Heritage 2013*, volume 1, pages 47–54. *Cited on page 10*
- RESHETYUK, Y. (2009a). *Self-calibration and direct georeferencing in terrestrial laser scanning*. PhD thesis, Royal Institute of Technology (KTH), Stockholm, Sweden. *Cited on page 53*
- RESHETYUK, Y. (2009b). *Terrestrial Laser Scanning: Error sources, self-calibration and direct georeferencing*. VDM Verlag Dr. Müller, Saarbrücken. *2 citations on pages 15 and 16*
- RIVEIRO, B., DEJONG, M. J., AND CONDE, B. (2016). Automated processing of large point clouds for structural health monitoring of masonry arch bridges. *Automation in Construction*, 72:258–268. *4 citations on pages 68, 69, 94, and 96*
- RIZALDY, A., PERSELLO, C., GEVAERT, C. M., AND OUDE ELBERINK, S. J. (2018). Fully Convolutional Networks for Ground Classification from LiDAR Point Clouds. In *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume IV-2, pages 231–238. *3 citations on pages 68, 69, and 80*
- ROCA, D., LAGUELA, S., DIAZ-VILARINO, L., ARMESTO, J., AND ARIAS, P. (2013). Low-cost aerial unit for outdoor inspection of building façades. *Automation in Construction*, 36:128–135. *2 citations on pages 18 and 60*
- ROS, G., SELLART, L., MATERZYNSKA, J., VAZQUEZ, D., AND LOPEZ, A. M. (2016). The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2016-Decem, pages 3234–3243. *Cited on page 69*
- RUPNIK, E., DAAKIR, M., AND PIERROT DESEILLIGNY, M. (2017). MicMac - a free, open-source solution for photogrammetry. *Open Geospatial Data, Software and Standards*, 2(1):14. *4 citations on pages 4, 31, 33, and 35*
- RUSU, R. B., BLODOW, N., AND BEETZ, M. (2009). Fast Point Feature Histograms (FPFH) for 3D registration. In *2009 IEEE International Conference on Robotics and Automation*, pages 3212–3217. *Cited on page 61*
- RUSU, R. B. AND COUSINS, S. (2011). 3D is here: point cloud library. In *IEEE International Conference on Robotics and Automation*, pages 1 – 4, Shanghai. IEEE. *4 citations on pages 56, 68, 74, and 98*

- SANCHEZ, V. AND ZAKHOR, A. (2012). Planar 3D modeling of building interiors from point cloud data. In *Proceedings - International Conference on Image Processing, ICIP*, pages 1777–1780. *Cited on page 68*
- SCHENK, T. (2005). *Introduction to Photogrammetry*. Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University. *2 citations on pages 12 and 31*
- SEKER, D. Z., ALKAN, M., KUTOGLU, H., AKCIN, H., AND KAHYA, Y. (2010). Development of a GIS Based Information and Management System for Cultural Heritage Site ; Case Study of Safranbolu. In *FIG Congress 2010*. *Cited on page 80*
- SEMLER, Q., SUWARDHI, D., ALBY, E., MURTIYOSO, A., AND MACHER, H. (2019). Registration of 2D Drawings on a 3D Point Cloud As a Support for the Modeling of Complex Architectures. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W15, pages 1083–1087. *4 citations on pages 10, 92, 115, and 153*
- SHARON, Y., WRIGHT, J., AND MA, Y. (2009). Minimum sum of distances estimator: Robustness and stability. In *Proceedings of the American Control Conference*, pages 524–530. *Cited on page 57*
- SPINA, S., DEBATTISTA, K., BUGEJA, K., AND CHALMERS, A. (2011). Point Cloud Segmentation for Cultural Heritage Sites. In *VAST11: The 12th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, pages 41–48. *Cited on page 81*
- STAIGER, R. (2003). Terrestrial Laser Scanning Technology, Systems and Applications. In *Proceedings of the 2nd FIG Regional Conference*, pages 1–10. *Cited on page 16*
- SZELISKI, R. (2010). *Computer Vision : Algorithms and Applications*, volume 5. Springer. *2 citations on pages xv and 14*
- TAKASHIMIZU, Y. AND IYOSHI, M. (2016). New parameter of roundness R: circularity corrected by aspect ratio. *Progress in Earth and Planetary Science*, 3(1). *Cited on page 96*
- TANG, P., HUBER, D., AKINCI, B., LIPMAN, R., AND LYTLE, A. (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction*, 19(7):829–843. *Cited on page 3*
- TOMBARI, F., SALTI, S., AND DI STEFANO, L. (2010). Unique signatures of histograms for local surface description. In Daniilidis, K., Maragos, P., and Paragios, N., editors, *ECCV 2010, Part III, LNCS 6313*, pages 356–369, Heraklion, Greece. Springer Berlin Heidelberg. *Cited on page 61*
- TORR, P. AND ZISSERMAN, A. (1998). Robust Computation and Parametrization of Multiple View Relations. In *Proceedings of the 6th International Conference on Computer Vision (ICCV'98)*. *Cited on page 72*

- VERHOEVEN, G., TAELEMAN, D., AND VERMEULEN, F. (2012). Computer Vision-Based Orthophoto Mapping Of Complex Archaeological Sites: The Ancient Quarry Of Pitaranha (Portugal-Spain). *Archaeometry*, 54(6):1114–1129. Cited on page 15
- VERHOEVEN, G. J. (2011). Taking Computer Vision Aloft - Archaeological Three-dimensional Reconstructions from Aerial Photographs with PhotoScan. *Archaeological Prospection*, 18:67–73. Cited on page 33
- VERHOEVEN, G. J. (2016). Mesh Is More-Using All Geometric Dimensions for the Archaeological Analysis and Interpretative Mapping of 3D Surfaces. *Journal of Archaeological Method and Theory*, 35:1–35. Cited on page 11
- VO, A.-V., TRUONG-HONG, L., LAEFER, D. F., AND BERTOLOTTI, M. (2015). Octree-based region growing for point cloud segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104:88–100. 4 citations on pages 3, 68, 74, and 99
- VON SCHWERIN, J., RICHARDS-RISSETTO, H., REMONDINO, F., AGUGIARO, G., AND GIRARDI, G. (2013). The MayaArch3D project: A 3D webGIS for analyzing ancient architecture and landscapes. *Literary and Linguistic Computing*, 28(4):736–753. Cited on page 21
- VOSSELMAN, G. AND MAAS, H. G. (2010). *Airborne and Terrestrial Laser Scanning*. Whittles Publishing, Dunbeath. 5 citations on pages xvi, 16, 18, 70, and 71
- WANG, C., CHO, Y. K., AND KIM, C. (2015a). Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. *Automation in Construction*, 56:1–13. Cited on page 94
- WANG, Z., ZHANG, L., FANG, T., MATHIOPOULOS, P. T., TONG, X., QU, H., XIAO, Z., LI, F., AND CHEN, D. (2015b). A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5):2409–2425. Cited on page 69
- WEINMANN, M. (2016). Preliminaries of 3D Point Cloud Processing. In *Reconstruction and Analysis of 3D Scenes: From Irregularly Distributed 3D Points to Object Classes*, pages 17–38. Springer. 2 citations on pages 54 and 73
- WEINMANN, M., JUTZI, B., HINZ, S., AND MALLET, C. (2015). Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105:286–304. Cited on page 2
- WENZEL, K., ROTHERMEL, M., HAALA, N., AND FRITSCH, D. (2013). SURE : The IfP software for dense image matching. In *Photogrammetric Week*, pages 59–70. Cited on page 33
- WOLF, P., DEWITT, B., AND WILKINSON, B. (2014). *Elements of Photogrammetry with Applications in GIS*. McGraw-Hill Education, 4th edition. 2 citations on pages 31 and 53
- YANG, X., KOEHL, M., AND GRUSSENMEYER, P. (2017). Parametric modelling of as-built beam framed structure in bim environment. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLII-2/W3, pages 651–657. 2 citations on pages 76 and 97

- YANG, X., KOEHL, M., AND GRUSSENMEYER, P. (2018). Automating Parametric Modelling From Reality-Based Data by Revit Api Development. In Remondino, F., Georgopoulos, A., González-Aguilera, D., and Agrafiotis, P., editors, *Latest Developments in Reality-Based 3D Surveying and Modelling*, pages 307–325. MDPI, Basel, Switzerland. *3 citations on pages 5, 67, and 68*
- YANG, X., KOEHL, M., GRUSSENMEYER, P., AND MACHER, H. (2016). Complementarity of Historic Building Information Modelling and Geographic Information Systems. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLI-B5, pages 437–443. *Cited on page 3*
- ZHANG, W., QI, J., WAN, P., WANG, H., XIE, D., WANG, X., AND YAN, G. (2016). An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sensing*, 8(6):1–22. *4 citations on pages 80, 81, 115, and 153*





# Appendices



## Résumé en français

# Relevé 3D et classification de nuages de points du patrimoine bâti

## 1 Introduction

La documentation du patrimoine bâti a beaucoup évolué ces dernières années grâce au développement de nouveaux capteurs 3D et de nouvelles techniques de relevé 3D. Les données 3D contribuent à la création d'archives fiables et tangibles des sites et des monuments historiques. Les progrès des algorithmes et des moyens de calculs ont permis de démocratiser la mise en œuvre des relevés 3D. L'usage du terme « modélisation 3D photoréaliste » s'est généralisé. Aujourd'hui, ce type de modélisation repose sur deux méthodes complémentaires : la méthode passive basée sur le traitement d'images, et la méthode active basée sur la mesure directe de distances par balayage laser. La photogrammétrie, forte de ses 150 ans d'existence, représente la technique la plus courante dans l'approche basée sur l'utilisation d'images. La photogrammétrie s'est enrichie de techniques issues du domaine de la vision par ordinateur comme la structure par le mouvement (*Structure from Motion* ou SfM) et des algorithmes d'appariement pour le calcul de nuages de points denses. La photogrammétrie est ainsi devenue une solution fiable et à faible coût pour le relevé 3D du patrimoine. Les développements des capteurs photographiques et la généralisation de l'utilisation des drones ont largement contribué à la popularité de la photogrammétrie. Dans l'approche basée sur la mesure de distances, la technologie lidar s'est beaucoup développée.

Le relevé 3D génère un nuage de points obtenu, soit par des mesures lasergrammétriques, soit par une solution d'appariement dense implémentée sur un bloc d'images orientées. Le nuage de points conserve des informations géométriques (c'est-à-dire les coordonnées XYZ de chaque point) qui traduisent la géométrie 3D de l'objet scanné. D'autres informations peuvent être stockées dans le nuage de points telles que les directions des normales, les valeurs de courbure, de linéarité ou encore de planéité (relative à une surface locale). Les couleurs RVB et les valeurs d'intensité (dans le cas de la lasergrammétrie) peuvent également être enregistrées dans le fichier du nuage de points. Cependant, ces informations ont une nature singulière, c'est-à-dire qu'elles sont uniques pour chaque point dans le nuage. Afin de pouvoir effectuer des opérations sur la donnée 3D, une information sémantique (appelée également attribut) doit être rajoutée. Ceci nous permet d'interpréter les nuages de points selon leurs caractéristiques tangibles et

non pas seulement géométriques. C'est dans ce contexte que les notions de segmentation et de classification de nuages de points interviennent. Il s'agit du « découpage » du nuage en éléments plus petits, plus gérables et surtout labellisés par des attributs. Un tel nuage de points segmenté et classifié nous permet d'effectuer plusieurs analyses, modélisations, et même des prédictions sur l'objet. Ces informations sont utilisées dans le processus de création de la maquette numérique du bâtiment (« BIM » ou *Building Information Model*). Dans le domaine du patrimoine, cette maquette est appelée maquette numérique du bâtiment historique ou « HBIM » (*Heritage Building Information Model*).

Vu l'importance des données 3D dans la documentation du patrimoine bâti, le contrôle de qualité est un aspect primordial qui devrait être abordé avant d'entreprendre le traitement du nuage de points. La thèse est ainsi divisée en deux parties. La première partie concerne l'acquisition et le contrôle de qualité des données. Un point important sera l'intégration de la photogrammétrie et de la lasergrammétrie dans le contexte de la documentation d'un site historique à différentes échelles. La deuxième partie de la thèse va aborder le traitement de nuages de points, plus particulièrement la segmentation et la classification de nuages de points. L'aspect multi-échelle de notre approche est important car dans beaucoup de cas, un bâtiment remarquable se situe dans un quartier historique qui nécessite une segmentation multi-échelle. L'algorithme développé pendant la thèse permettra la segmentation et la classification automatique du nuage de points d'un quartier historique en autant de nuages de points distincts que de bâtiments. Ces groupes de nuages de points sont ensuite segmentés et classifiés en éléments architecturaux pour extraire automatiquement des objets tels que des piliers et des poutres de charpentes.

## 2 Relevé 3D et contrôle de qualité

La première partie de la thèse abordera tout d'abord les techniques d'acquisitions 3D dans le cadre de la documentation du patrimoine. Nous parlerons ainsi des techniques de base comme la photogrammétrie et la lasergrammétrie, mais aussi de l'approche de relevé multi-échelle développé pendant la thèse. Le contrôle de qualité photogrammétrique est décrit de manière plus détaillée car la photogrammétrie présente plusieurs étapes exigeant des vérifications rigoureuses. L'utilisation des logiciels commerciaux de type « boîte noire » constitue une raison supplémentaire pour approfondir cette étude sur le contrôle de qualité photogrammétrique. Un second sujet qui sera également abordé dans cette partie est la combinaison des données lasergrammétriques avec des données photogrammétriques. En effet, une telle intégration dans le cadre d'une approche multi-capteurs est la conséquence logique de l'approche multi-échelle que nous avons développée.

### 2.1 Relevé 3D du patrimoine

La documentation du patrimoine est un aspect très important dans le travail de la conservation. Historiquement, l'archivage et la documentation physique des monuments remarquables sont souvent mis en évidence. Avec les différents dangers auxquels un monument historique peut être exposé, qu'ils soient naturel ou anthropologique, la documentation devient un élément primordial. En effet, la reconstruction physique d'un site en ruines génère souvent des polémiques en absence d'une bonne documentation ; nous pouvons citer par exemple la controverse liée à la reconstruction du Château du Haut-Koenigsbourg par les autorités allemandes pendant le 19e siècle, ou alors plus récemment le débat sur la restitution du colosse de Ramsès II à Luxor. En fait, il existe des tentatives au niveau

international pour régler ce problème, notamment la charte de Venise de 1964 qui mettait en avant l'importance d'une compréhension holistique d'un objet historique avant même qu'une reconstruction physique soit planifiée. La charte exige également une documentation « précise » dans chaque mission qui concerne des opérations autour d'un objet de patrimoine. Le mot « précise » pourrait être interprété comme l'utilisation de données géospatiales.

Les données géospatiales concernent souvent des coordonnées et des mesures 2D, mais celles-ci ont récemment évolué vers l'espace 3D. Les techniques de la photogrammétrie et la lasergrammétrie sont parmi les méthodes les plus utilisées dans la documentation 3D. La photogrammétrie, du haut de ses 150 ans, représente la technique la plus courante dans l'approche basée sur l'utilisation d'images. La photogrammétrie traditionnelle s'appuyait sur des principes proches de la topographie. Elle s'est enrichie de techniques issues du domaine de la vision par ordinateur pour devenir une solution fiable et à faible coût pour le relevé 3D du patrimoine. Les développements des capteurs photographiques et la généralisation des drones ont largement contribué à la popularité de la photogrammétrie.

Dans l'approche basée sur la mesure de distances, la technologie lidar (dès lors appelée la lasergrammétrie ou balayage laser pour la distinguer du lidar aérien) s'est également beaucoup développée. Une comparaison du constructeur Trimble montre la croissance exponentielle du taux d'acquisition ; en 2005, le Trimble GX pouvait générer 5.000 points par seconde alors qu'en 2017, le Trimble SX10 a atteint 25.000 points par seconde, soit une vitesse cinq fois plus rapide. La toute dernière édition du scanner Trimble, le Trimble X7 peut mesurer jusqu'à 500.000 points par seconde en 2020. Les logiciels de traitements ont également évolué vers l'automatisation de la chaîne de traitements.

Tenant compte de la variété des capteurs disponible aujourd'hui sur le marché, une approche multi-capteurs et multi-échelles pour le relevé 3D peut être envisagée. En plus, le développement d'une telle approche est devenu logique face à la variété des bâtiments se trouvant dans un quartier historique. La possibilité d'avoir les données géospatiales à l'échelle d'un quartier nous permet de mieux comprendre l'objet d'intérêt dans son contexte socio-géographique. Dans le cadre de la mission elle-même, une telle approche peut nous aider dans la planification d'autres opérations de relevé telles que les mesures topographiques.

Dans l'approche multi-échelle développée dans cette thèse, nous identifions quatre niveaux d'échelle (Figure 1.4). Le niveau 0 concerne la plus petite échelle dans laquelle nous trouvons le quartier historique étudié et ses environs. Nous pouvons numériser ce niveau en utilisant des techniques de mesures à petite échelle telles que la photogrammétrie aérienne par drone, des images satellites, ou bien la lasergrammétrie aéroportée. Le quartier historique lui-même constitue le niveau 1 de notre approche. Pour ce niveau d'échelle, des techniques terrestres ne sont pas encore pertinentes ; elles interviendront donc dans le niveau 2 qui représente un bâtiment d'intérêt. Des techniques comme la photogrammétrie rapprochée (terrestre ou aérienne par drone) ou la lasergrammétrie peuvent être utilisées dans ce niveau. Finalement, dans le niveau 3, nous trouvons les éléments architecturaux qui nécessitent la résolution la plus fine de nuage de points. Nous privilégions dans ce cas-là la technique de la photogrammétrie rapprochée.

## **2.2 Contrôle de qualité photogrammétrique**

Pour le contrôle de qualité des données 3D, nous avons décidé de nous intéresser plus particulièrement au contrôle des données photogrammétriques, car la majorité des logi-

ciels commerciaux sont des « boîtes noires ». De plus, la photogrammétrie, contrairement à la lasergrammétrie, nécessite plus d'étapes de traitements et donc plus de contrôles. Pour ce faire, nous nous appuyons sur la boîte à outils DBAT (*Damped Bundle Adjustment Toolbox*) développée dans Matlab (MathWorks) en collaboration avec l'Université d'Umea (Suède). Il s'agit d'un ensemble de fonctions permettant de recalculer l'ajustement des faisceaux des projets photogrammétriques issus de logiciels commerciaux. L'orientation externe ou le calcul des poses des caméras constituent une étape majeure dans un projet photogrammétrique. L'outil DBAT (disponible sur GitHub) fournit des statistiques sur les données et les résultats des calculs, permettant le contrôle qualité de données notamment issues de drones et de photogrammétrie terrestre.

Deux jeux de données ont été utilisés dans cette expérience : des images de drones prises sur les façades de l'église St-Pierre-le-Jeune catholique et l'église St-Paul de Strasbourg. Ces jeux de données sont traités avec le logiciel Agisoft PhotoScan (appelé aujourd'hui « Metashape »). PhotoScan est un logiciel très répandu pour traiter des données de numérisation 3D, y compris dans le domaine du patrimoine. Cependant, étant un logiciel commercial, le gros inconvénient de PhotoScan est sa nature « boîte noire ». Pour certains cas nécessitant de grandes précisions, ceci nous empêche d'effectuer un contrôle robuste sur sa qualité. C'est dans ce cas-là que DBAT intervient en tant que solution libre et ouverte pour le calcul d'ajustement des faisceaux. DBAT a été utilisé pour recalculer les projets PhotoScan des deux jeux de données disponibles. Pour les données de St-Paul (Strasbourg), les images ont été traitées aussi par le logiciel photogrammétrique Apero développé par l'IGN.

Les résultats montrent qu'en utilisant une solution alternative et ouverte comme DBAT, nous avons mis en évidence des problèmes qui existent dans certains jeux de données photogrammétriques. Il ressort de ces expériences qu'il est primordial d'adopter une procédure d'évaluation pour tout projet de photogrammétrie, afin d'être en mesure de détecter les erreurs dans des jeux de données. En effet, la complexité de la configuration des images dans le cas de la photogrammétrie rapprochée, la nature incohérente de leurs taux de recouvrement et la variation de la taille du pixel-objet au sein d'un même projet cachent souvent des problèmes. Il est d'autant plus important d'avoir un moyen pour contrôler les résultats, y compris les résultats intermédiaires, et détecter ainsi les problèmes avant de continuer l'étape suivante de la chaîne de traitements photogrammétriques.

Dans les expérimentations que nous avons effectuées, PhotoScan a généré des résultats de l'ajustement des faisceaux avec une valeur de précision dans le même ordre de grandeur que la taille théorique du pixel-objet. Ceci a été validé par DBAT et Apero qui ont générés des précisions similaires au vu des résidus sur les points d'appui et les points de contrôle (Tableaux 2.3 et 2.5). Ces résultats ont été obtenus en utilisant la même pondération. L'inconvénient de PhotoScan est le manque de valeurs statistiques dans le calcul de l'orientation externe. Ceci pourrait éventuellement cacher des problèmes dans les jeux de données. Les résultats de DBAT montrent que des améliorations peuvent encore être envisagées sur le projet, notamment en ajoutant plus d'images sur les zones ayant une précision faible. Apero montre quelques indices sur la qualité des images que nous pourrions utiliser dans le même but que DBAT.

Ces expérimentations nous montrent qu'une vérification objective des résultats d'un projet photogrammétrique est importante. Des erreurs peuvent se cacher dans le projet, ce qui pourrait affecter sa qualité. Lorsque la précision obtenue correspond au cahier des charges du projet, cela ne pose pas de soucis. Cependant dans le cas où un problème survient, les valeurs statistiques fournies par DBAT et Apero pourraient être utilisées pour

l'identifier. Les solutions libres sont ainsi utiles pour valider les résultats d'un projet et de faire des analyses détaillées. La photogrammétrie d'aujourd'hui avec l'augmentation des moyens de calcul et d'appariement dense génère de plus en plus une demande vers une solution plus métrique, surtout pour des utilisations géospaciales. En effet, la mise à jour récente de PhotoScan (désormais Metashape) qui intègre plus de valeurs statistiques dans son rapport est encourageante pour l'avenir. Nous souhaitons que les logiciels libres puissent ainsi jouer leur rôle de guide pour les logiciels commerciaux pour privilégier des pistes de développement.

### 2.3 Intégration de la photogrammétrie et de la lasergrammétrie

Dans cette section, le contrôle de qualité au niveau de l'intégration des résultats de la photogrammétrie et de la lasergrammétrie sera présenté. Puisque la qualité des données provenant des différents capteurs peut être hétérogène, une bonne méthode d'intégration est nécessaire afin de générer un nuage de points final d'une qualité géométrique acceptable. L'intégration abordée ici concerne le nuage de points de l'extérieur de l'église St-Pierre généré par photogrammétrie et celui de l'intérieur acquis par un scanner laser terrestre. Le jeu de données de St-Pierre présente un cas particulier, car chaque mission a été effectuée séparément à des périodes différentes. Par conséquent, il n'existe pas beaucoup de zones de recouvrement entre les nuages de points de l'extérieur et de l'intérieur. Deux méthodes sont ainsi testées : la première consiste en un géoréférencement indépendant de chaque nuage de points, grâce aux points d'appuis qui sont mesurés séparément, mais qui sont dans le même système de coordonnées. La deuxième méthode consiste à consolider les deux nuages de points en utilisant des points homologues, avant de géoréférencer le nuage consolidé vers un système absolu. La comparaison entre ces deux méthodes est faite en utilisant des solutions libres afin de pouvoir déduire les avantages et les inconvénients de chaque approche. De plus, nous avons essayé d'automatiser la détection des points homologues à l'aide de la bibliothèque de fonctions PCL (*Point Cloud Library*).

L'analyse des résultats montre qu'il est possible d'obtenir un résultat final similaire en utilisant les deux approches. Cependant, en regardant plus en détail nous verrons qu'il existe différents avantages liés à chaque approche. La méthode du géoréférencement indépendant est la plus rapide. Ceci est par ailleurs l'approche la plus logique puisque la chaîne d'acquisition classique pour les données 3D exige déjà un géoréférencement pour chaque méthode individuelle. Par contre la qualité finale du nuage de points consolidé dépend très fortement de la qualité du géoréférencement individuel de l'extérieur et l'intérieur. L'approche « réseau libre » effectue quant à elle un calcul en bloc pour minimiser les erreurs. Comme nous pouvons voir sur le Figure 3.5, la première approche montre une présence de systématismes, alors que les erreurs sur la deuxième approche sont bien mieux réparties. Alors que l'approche réseau libre nous donne un meilleur résultat géométrique, son inconvénient principal est lié à l'identification des points homologues fait jusqu'ici de manière manuelle. Ceci représente la partie la plus chronophage de l'approche réseau libre.

Une chaîne d'automation simplifiée a été développée en utilisant la bibliothèque PCL afin de résoudre ce problème. Les fonctions permettent d'automatiser une partie de la détection de points homologues 3D et leur appariement. Nous réduisons ainsi le temps nécessaire pour identifier les points homologues manuels. Cependant, cette problématique ne constitue pas le point central de la thèse et nous ne poursuivons pas son développement



plus loin. Les fonctions d'automatisation qui ont été développées sont toutefois utilisées comme base pour d'autres travaux de notre équipe de recherche. Le travail d'ASSI ET AL. [2019], un doctorant de notre laboratoire, utilise notamment ces fonctions pour faciliter la consolidation automatique des nuages de points d'intérieur et d'extérieur. Les codes ont été également utilisés dans deux projets de recherche technologique à l'INSA de Strasbourg : PEGOSSOFF [2018] travaillait sur la consolidation des nuages issus de deux différents types de scanner laser (scanner terrestre et scanner à main), tandis que le travail d'AMZIL [2020] fait partie de la suite de la recherche d'ASSI ET AL. [2019].

Les expérimentations menées dans cette partie de la thèse ont comme objectif de tester les deux approches de consolidation et de géoréférencement de nuages de points de l'intérieur et de l'extérieur. Ceci est effectué surtout afin de pouvoir vérifier la qualité du nuage de points final dit « hybride ». Le nuage de points hybride sert en effet pour les données d'entrée des étapes suivantes : la segmentation et la classification. Un tel contrôle est devenu de plus en plus important afin de garantir une représentation fiable de la réalité.

### 3 Algorithmes de traitements de nuages de points

La seconde partie de la thèse concerne le développement d'une chaîne de traitements de nuages de points. Dans cette étude, nous avons essayé d'aborder la problématique particulière du relevé 3D multi-échelle. L'algorithme développé nous permet de segmenter progressivement le nuage de points d'un quartier ; tout d'abord en unités individuelles de bâtiments, puis en éléments architecturaux. Les fonctions sont développées dans Matlab afin de profiter de la continuité avec DBAT, mais aussi avec d'autres recherches menées auparavant dans le laboratoire. L'ensemble des fonctions sont rassemblées dans une boîte à outils Matlab appelée *M\_HERACLES* (*HERitAge by point CLOUD procESSing for Matlab*) qui est à la fois ouverte et libre (et disponible sur GitHub).

#### 3.1 La boîte à outils M\_HERACLES

Dans la plupart des projets de documentation 3D du patrimoine, un bâtiment historique se situe dans un quartier ayant un intérêt particulier au niveau historique. Une documentation complète du site peut être envisagée afin de comprendre l'objet d'intérêt dans le contexte plus large de son quartier. Dans ce cas particulier, une approche multi-capteur et multi-échelle est la réponse logique, car chaque niveau d'échelle a besoin d'une résolution différente. Chaque technique d'acquisition est normalement adaptée pour une échelle donnée, par exemple la photogrammétrie rapprochée pour des statues ou le scanner laser terrestre pour l'intérieur des bâtiments. En fonction de la taille de l'objet, la résolution de son nuage de points correspondant peut être variée. Comme nous l'avons déjà expliqué, dans cette thèse nous proposons une chaîne de traitement multi-échelle y compris dans le traitement de nuages de points. Cette approche vise à segmenter et à classifier le nuage de points de manière progressive ; tout d'abord de l'échelle d'un quartier vers celle d'un bâtiment et puis de l'échelle d'un bâtiment vers celle des éléments architecturaux. Concrètement, les fonctions que nous avons développées en ce sens sont rassemblées dans la boîte à outils *M\_HERACLES*. *M\_HERACLES* a une nature modulaire inspirée de DBAT. Des fonctions indépendantes sont créées dans des buts spécifiques, avec la possibilité de les lancer successivement. *M\_HERACLES* s'intéresse de plus à la question de la classification, c'est-à-dire l'ajout de l'information sémantique aux résultats.

Il existe plusieurs approches permettant d'automatiser le traitement de nuages de

points. Une méthode beaucoup utilisée dans les dernières années s'appuie sur les techniques d'apprentissage profond. Une autre possibilité d'automatisation utilise plutôt des règles géométriques pour identifier et classer les objets, qui correspond à l'approche dite algorithmique. Cette thèse étant la continuité des travaux de notre équipe sur l'utilisation de la méthode algorithmique, **M\_HERACLES** a été conçu pour classer les nuages de points à l'aide de connaissances préliminaires (par exemple, un pilier a la forme d'un cylindre). Néanmoins, le but de cette thèse n'est pas de confronter notre approche algorithmique à celle de l'apprentissage profond, mais plutôt de la supporter. En effet, l'un de nos travaux en cours concerne l'utilisation de **M\_HERACLES** pour générer des données d'entraînement pour un algorithme d'apprentissage profond.

Une revue de la littérature existante nous montre que la plupart des solutions disponibles à ce jour n'abordent qu'un seul niveau d'échelle de la scène. Il existe des algorithmes de classification pour des nuages de points à petite échelle (par exemple l'échelle d'un quartier ou alors d'une ville) et pour de plus grandes échelles (par exemple au niveau de l'intérieur d'un bâtiment), mais ces algorithmes concernent très rarement un scénario multi-échelle. **M\_HERACLES** propose une solution holistique pour un tel cas avec son approche de classification progressive. Tous les calculs se font en utilisant une seule boîte à outils pour faciliter sa mise en œuvre. La nature modulaire de la boîte à outils est conçue pour être flexible en fonction des cas rencontrés et des niveaux d'échelles souhaités.

Figure 4.6 montre la chaîne générale du traitement de nuages de points sous **M\_HERACLES**. La première « étape » concerne la segmentation du nuage de points d'un quartier vers des sous-nuages des objets d'intérêt. Ce processus s'appuie sur des fichiers SIG en 2D, qui permettent d'annoter directement les résultats avec une information sémantique issue de l'attribut du SIG. La seconde « étape » permet la détection des éléments architecturaux en se basant sur le nuage de points généré préalablement. Deux classes d'éléments architecturaux sont actuellement disponibles dans **M\_HERACLES** : les piliers (c'est-à-dire les supports structuraux) et les poutres de charpentes.

Tous les jeux de données sont traités par **M\_HERACLES**. Les jeux de données de Kasepuhan et St-Pierre sont utilisés pour la première étape de segmentation. Les résultats de cette opération sont ensuite utilisés en tant que données d'entrée pour l'étape suivante. En plus, trois autres jeux de données ont également été traités pour illustrer la deuxième étape : Paestum et Valentino sont utilisés pour la détection des piliers, alors qu'un troisième nuage de points de la charpente du Château du Haut-Koenigsbourg est utilisé pour la détection des poutres. Tous les traitements sont effectués en utilisant un seul ordinateur avec un processeur Intel<sup>(R)</sup> Xeon<sup>(R)</sup> E5645 2.4 GHz.

### 3.2 Segmentation d'un quartier historique vers des bâtiments remarquables

La segmentation de nuage de points est un sujet très répandu dans le domaine de la 3D. Ceci est dû au besoin de la communauté de la télédétection de classer des données lidar aériennes. Des développements dans le domaine de la photogrammétrie et notamment l'appariement dense ont révolutionné la recherche sur ce sujet en ouvrant la possibilité d'acquérir des nuages de points très détaillés à plus grande échelle grâce notamment à l'utilisation des drones. Dans le cas de la documentation d'un site historique à l'échelle d'un quartier, cette opération est plus souvent effectuée en utilisant à la fois des techniques aériennes et terrestres. Ceci nous donne un nuage de points complet du site avec la résolution fine de la technique terrestre et enrichi par un point de vue aérien.

Alors que cette combinaison est une solution pertinente pour répondre aux enjeux de la documentation 3D du patrimoine, elle ajoute de la complexité au nuage de points final. Une automatisation est d'autant plus importante afin de nous aider à traiter ces données complexes.

Il existe plusieurs études sur la segmentation et la classification d'un nuage de points aérien vers des classes génériques comme le sol, les bâtiments, la végétation, etc. Cependant, le domaine du patrimoine a souvent besoin d'informations sémantiques (par exemple l'information historique, le style d'architecture, la date de construction, etc.). Dans beaucoup de cas, ces données sémantiques sont déjà stockées sous forme d'un Système d'information géographique (SIG), souvent en 2D. L'idée principale de l'algorithme développé dans cette thèse est de profiter de ces SIG déjà existants pour aider le processus de la segmentation 3D du nuage de points. Puisque le SIG contient aussi les attributs pour chaque objet, nous pouvons ensuite annoter directement les sous-nuages segmentés avec l'information sémantique liée à sa géométrie. La segmentation et l'annotation peuvent de cette manière être utiles dans la création éventuelle d'un SIG 3D ou un HBIM.

Pour la segmentation et la classification de nuages de points des quartiers vers des bâtiments individuels, l'algorithme développé s'appuie sur des fichiers de type SIG, souvent existants pour les quartiers ou les sites historiques, bien qu'ils soient en 2D. L'idée est d'utiliser la géométrie 2D enregistrée dans les fichiers SIG pour guider la segmentation des nuages de points. La géométrie 2D est exploitée en tant que segmenteur 2.5D à l'image d'un emporte-pièce : tous les points du nuage de points se trouvant dans les limites 2D (XY) de l'objet sont considérés comme appartenant à l'entité SIG. Le résultat est ensuite affiné avec l'implémentation de deux algorithmes principaux : la détection (et suppression) du sol et le débruitage du nuage de points par le biais de la méthode du plus proche voisin.

Cette approche a été testée sur deux jeux de données avec des résultats plutôt satisfaisants. Les sous-nuages générés par l'algorithme sont corrects même dans le cas de la présence de végétation dense à Kasepuhan. En utilisant cette approche, chaque objet est segmenté individuellement en gardant la classe attribuée par le fichier shapefile d'entrée. L'autre avantage de notre approche est que les attributs d'une instance de la classe sont directement annotés sur les sous-nuages du résultat. L'annotation se fait automatiquement et individuellement pour chaque sous-nuage afin de pouvoir éventuellement faciliter les travaux de management du site sous forme d'un SIG 3D et/ou un HBIM. La comparaison avec la fonction de classification automatique du logiciel Metashape a également montré des résultats prometteurs ; à savoir que Metashape utilise une approche d'apprentissage profond. Metashape obtient une valeur moyenne de l'indice F1 de 66.09% pour les deux jeux de données testés, contre 89.32% pour M\_HERACLES.

La valeur médiane de l'indice statistique (score) F1 caractérisant la qualité de la segmentation est de 91.99% pour Kasepuhan et 93.90% pour St-Pierre. Ces valeurs sont très prometteuses pour cette tâche. En plus de la segmentation géométrique du nuage de points, l'utilisation des fichiers SIG nous permet d'annoter les nuages de points segmentés avec l'attribut contenant l'information sémantique, ce qui présente un autre atout de notre approche.

Néanmoins, il reste encore quelques défauts dans notre approche. La nature algorithmique se traduit par le besoin d'identifier quelques paramètres principaux. C'est le cas par exemple dans la détermination des seuils de tolérance pour l'algorithme de croissance de région. L'ordre d'itérations est aussi important afin de bien distinguer les classes qui sont superposées verticalement (par exemple la couronne d'un arbre au-dessus du toit

d'un bâtiment). Nous avons observé que l'algorithme doit être appliqué tout d'abord sur les classes ayant une altitude plus basse et puis itéré progressivement vers les classes plus hautes (Figure 5.6). Le choix sur l'ordre d'itérations à prendre reste une intervention manuelle.

L'algorithme d'extraction du sol lancé au début de la segmentation est aussi un élément très important qui influe sur le résultat. L'effet de cette partie de la fonction est plus évident pour la classe de végétation, où la présence des herbes ou des racines empêche une extraction propre du sol. Par conséquent, le sous-nuage segmenté est erroné. Pour le moment seul l'algorithme CSF de ZHANG ET AL. [2016] a été testé. Des tests avec d'autres algorithmes d'extraction du sol peuvent être envisagés pour améliorer les résultats.

D'autres idées ont également été explorées pendant la thèse afin d'améliorer l'algorithme développé. Une extension de notre algorithme qui utilise des fichiers de DAO (Dessin assisté par ordinateur) peut être envisagée surtout pour le prochain niveau d'échelle d'un bâtiment. Un fichier DAO est souvent disponible pour les bâtiments historiques et servira comme un fichier de SIG dans ce cas-là. En effet, cette idée a été mise en œuvre dans une publication de notre équipe de recherche [SEMLER ET AL., 2019].

### 3.3 Segmentation d'un bâtiment remarquable vers des éléments architecturaux

En ce qui concerne l'étape de segmentation des bâtiments vers les éléments architecturaux, nous nous limitons à ce stade de la thèse à deux classes qui ont beaucoup d'intérêt dans le domaine de la documentation du patrimoine : les piliers et les charpentes. Les piliers sont très intéressants, car ils représentent l'architecture alors que les charpentes montrent le savoir-faire des bâtisseurs. Ces éléments sont parmi les plus complexes à segmenter et à classifier, car la forme des différents objets est fonction du style d'architecture. Quelques travaux de recherches existent déjà sur la même thématique, avec une tendance récente vers l'utilisation des algorithmes d'apprentissage automatique et l'application de l'apprentissage profond. Le problème souvent rencontré dans ce genre d'approche est la génération des données d'entraînement, jusque-là largement manuelle. Dans cette étude, nous privilégions l'utilisation des règles géométriques et des contraintes sémantiques pour automatiser la segmentation et la classification de ces éléments architecturaux. L'idée est de ne pas confronter cette approche à celle de l'apprentissage profond. Au contraire, nous présentons un moyen pour faciliter la création des données d'entraînement et ainsi créer une approche complémentaire.

La détection, la segmentation, et la classification des piliers commencent par la découpe verticale du bâtiment. La coupe centrale est sélectionnée, puis un algorithme de croissance de région nous permet d'isoler les « îlots » correspondant à chacun des candidats. Une règle géométrique est appliquée : un « îlot » de nuage de points est labellisé en tant que pilier si la forme de sa coupe est circulaire. Ensuite, un algorithme similaire à celui appliqué sur la segmentation d'un quartier vers un bâtiment est utilisé pour récupérer le nuage de points de l'objet. Le sol est supprimé en appliquant l'algorithme de RANSAC. Pour la détection des piliers, un score F1 moyen de 88.97 % a été obtenu avec un temps de calcul inférieur à 5 minutes pour les trois jeux de données.

Nous avons effectué une comparaison entre les résultats de notre approche et les résultats de l'approche développée par MALINVERNI ET AL. [2019]. Dans leur étude, les auteurs utilisent une approche d'apprentissage profond (algorithme PointNet++) pour effectuer la classification automatique de quelques classes dont les piliers. Une comparaison a pu

être faite seulement sur le jeu de données de Valentino, car les auteurs l'utilisent également dans leur article. Cette étude montre que l'approche de l'apprentissage profond nécessite une classification simultanée de toutes les classes souhaitées confondues ; nous interrogeons seulement le résultat de sa classe « pilier ». Pour Valentino, [M\\_HERACLES](#) a obtenu une valeur F1 de 75.92%. Pour la classe de pilier, l'approche de [MALINVERNI ET AL. \[2019\]](#) obtient une valeur F1 de 57.60% pour Valentino en utilisant un autre jeu de données comme données d'apprentissage.

Alors que notre approche donne un résultat nettement meilleur que celui de l'apprentissage profond, quelques remarques importantes doivent néanmoins être prises en compte. Premièrement, [M\\_HERACLES](#) a été utilisé pour détecter seulement les piliers autoportants en laissant quelques piliers engagés dans la scène de Valentino. Au contraire, l'apprentissage profond tente toujours de détecter tous les objets de la même classe dans la scène. Deuxièmement, l'approche de l'apprentissage profond est par nature améliorable en ajoutant plus de données d'entraînement. À l'avenir, nous pouvons donc nous attendre à une amélioration importante des résultats de l'apprentissage profond.

Pour la détection de poutres individuelles de la charpente, l'approche utilise un algorithme de croissance de région modifié qui permet de détecter les faces des poutres. Une détection des axes principaux des poutres est réalisée grâce à une transformation de Hough et deux règles géométriques liées au voisinage et au parallélisme de chaque face. Les résultats pour ces deux algorithmes sont aussi prometteurs. Pour la détection des poutres, le score F1 obtenu est de 86.09 %.

L'objectif principal du développement de cette partie de la thèse est donc de considérer la nature multi-capteur et multi-échelle de la documentation du patrimoine. Les fonctions pour la détection des piliers et des poutres nous permettent d'automatiser le processus de la segmentation et la classification du nuage de points jusque-là largement manuel. L'automatisation sert à réduire l'intervention humaine et donc l'erreur liée aux opérateurs.

Alors que les résultats sont globalement prometteurs, nous pouvons identifier les limites de l'algorithme développé. L'algorithme développé a du mal à gérer les cas où les caractéristiques de données d'entrée sont différentes des règles géométriques définies dans les fonctions. Par exemple, Valentino donne une valeur de précision plus basse à cause de la présence des piliers engagés (c'est-à-dire des demi-colonnes collées contre un mur). Les résultats de St-Pierre mettent en évidence la sensibilité de l'algorithme en présence du bruit (présence des objets divers dans la scène), alors que Paestum nous montre le risque du systématisme. L'algorithme fonctionne bien pour Kasepuhan certes, mais il s'agit d'un cas avec uniquement des piliers autoportants (voir Figure 6.3). Ceci dit le temps de traitement reste faible et les résultats sont acceptables dans la plupart des cas, surtout si l'objectif principal est de générer une « valeur approchée » pour une classification subséquente par l'apprentissage profond. La rapidité et la simplicité de cette approche algorithmique est même plus avantageuses dans des cas simples.

L'algorithme de détection des poutres est toujours en cours de développement et d'autres tests devront être effectués afin de pouvoir considérer son efficacité. Le test effectué dans cette thèse sur un petit échantillon des données de Haut-Koenigsbourg est prometteur donnant une valeur du score F1 médiane de 86.09%. Néanmoins, le temps de calcul reste élevé et le taux d'exactitude est plutôt faible bien que ceci ne pose pas trop de problèmes pour la création automatique des primitives géométriques. La majorité du temps nécessaire est liée au calcul des directions de normales et des valeurs de courbures; une stratégie doit donc être implémentée afin d'accélérer cette étape.

## 4. Conclusions et perspectives

Nous avons montré que les algorithmes développés sont pratiques, rapides et dans beaucoup de cas ne nécessitent pas l'utilisation d'une approche plus complexe. En combinant le contrôle de qualité développé dans la première partie avec ces fonctions, nous avons considéré l'ensemble du processus allant de l'acquisition de données 3D jusqu'à la segmentation et la classification en entités à plusieurs échelles.

Dans cette thèse, deux questions de recherche ont été formulées. Premièrement, nous nous interrogeons sur la problématique de la génération du nuage de points de haute qualité dans le cadre de la documentation du patrimoine. À partir de nos expériences décrites en détail dans cette thèse, nous pouvons considérer que la chaîne de traitements 3D, quelle que soit la technique utilisée (photogrammétrie, lasergrammétrie, ou autre), est aujourd'hui très directe et relativement facile à apprendre. La création d'un nuage de points 3D n'a jamais été aussi facile. Néanmoins, pour créer un nuage de points de haute qualité, il est indispensable de comprendre le fonctionnement de chaque technique en détail.

La deuxième question à laquelle nous essayons de répondre dans cette thèse concerne l'exploitation du nuage de points de haute qualité. En effet, un nuage de points, même d'une haute qualité, reste un ensemble de points géométriques sans classification. Nous avons montré par le biais de la boîte à outils `M_HERACLES` et les tests effectués que la segmentation et la classification du nuage de points nous permettent de répondre à cette question.

Dans les travaux futurs, il sera intéressant d'intégrer les résultats de nos algorithmes dans d'autres solutions comme l'apprentissage profond, car ceux-ci constituent un apport sur la nature présumée des objets. Au cours de la thèse, nous avons constaté un manque de jeu de données 3D en libre accès dans le domaine du patrimoine. La création d'un portail de partage des données 3D patrimoniales pourra être considérée dans le futur.



Que l'avenir ne te trouble pas ; tu l'aborderas, s'il le faut, en portant dans tout ce qu'il te réserve cette même raison qui t'éclaire sur les choses du moment.

---

Marc Aurèle (121-180 après J.C.), *Pensées pour moi même, Livre VII*



## Résumé

La documentation du patrimoine bâti a beaucoup évolué ces dernières années grâce au développement de nouveaux capteurs 3D et de nouvelles techniques de relevé 3D. Les données 3D contribuent à la création d'archives fiables et tangibles des sites et des monuments historiques. Vu l'importance des données 3D dans la documentation du patrimoine bâti, le contrôle de qualité est un aspect primordial qui devrait être abordé avant d'entreprendre le traitement du nuage de points. La thèse est ainsi divisée en deux parties. La première partie concerne principalement l'acquisition et le contrôle de qualité des données. Un point important sera l'intégration de la photogrammétrie et de la lasergrammétrie dans le contexte de la documentation d'un site historique à différentes échelles. La deuxième partie de la thèse va aborder le traitement de nuages de points, plus particulièrement la segmentation et la classification de nuages de points. L'aspect multi-échelle de notre approche est importante car dans beaucoup de cas, un bâtiment remarquable se situe dans un quartier historique qui nécessite une segmentation multi-échelle. En combinant ces deux parties, nous avons considéré l'ensemble du processus allant de l'acquisition de données 3D jusqu'à la segmentation et la classification en entités à plusieurs échelles.

**Mots-clés :** automatisation, photogrammétrie, nuage de points, modélisation 3D, classification, patrimoine

## Summary

The documentation of built heritage has seen a significant development these past few decades due to advancements in new 3D sensors and 3D recording techniques. 3D data serve as reliable and tangible archive for historical sites and monuments. Since 3D data have such importance in the field of heritage documentation, quality control is paramount and must be performed before any point cloud processing is even planned to be conducted. The thesis is therefore divided into two parts. The first part concerned mainly the data acquisition and quality control of the point cloud data using the two techniques most commonly used, i.e. photogrammetry and laser scanning. A particular emphasis was also put on the integration of photogrammetry and laser scanning within the context of a multi-scalar documentation of a heritage site. The second part will address the processing of the resulting point cloud, particularly its segmentation and classification. The multi-scalar approach proposed in this thesis is an important point to note, as in many cases a historical building of interest is located in a historical neighbourhood; thus the requirement for a multi-scalar segmentation. By combining these two parts, the thesis had attempted to address the 3D workflow of heritage sites in a holistic manner, from the 3D data acquisition up to the resulting point clouds' segmentation and classification into individual entities in various scale steps.

**Keywords:** automation, photogrammetry, point cloud, 3D modelling, classification, heritage