

Transfert et repérage d'annotations sémantiques pour le patrimoine architectural. Un système méthodologique combinant Intelligence Artificielle, H-BIM et plateformes d'annotations collaboratives reality-based.

Valeria Croce

▶ **To cite this version:**

Valeria Croce. Transfert et repérage d'annotations sémantiques pour le patrimoine architectural. Un système méthodologique combinant Intelligence Artificielle, H-BIM et plateformes d'annotations collaboratives reality-based.. Eco-conception. HESAM Université; Università degli Studi di Firenze, 2022. Français. ⟨NNT : 2022HESAE037⟩. ⟨tel-03708142⟩

HAL Id: tel-03708142

https://pastel.hal.science/tel-03708142v1

Submitted on 29 Jun 2022

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.



HAL Authorization

HESAM UNIVERSITÉ

Arts et Métiers Sciences et Technologies

ÉCOLE DOCTORALE SCIENCES DES MÉTIERS DE L'INGÉNIEUR
[Laboratoire ASTRO–Université de Pise, DICEA–Université de Florence, UMR MAP-CNRS Marseille, LISPEN– Campus d'Aix-en-Provence]

THÈSE

présentée par : **Valeria CROCE**
soutenue le : **21 juin 2022**

pour obtenir le grade de : **Docteur d'HESAM Université**

préparée à : **École Nationale Supérieure d'Arts et Métiers**
Spécialité : **Conception**

Semantic annotation transfer and retrieval for architectural heritage.
A methodological system combining Artificial Intelligence, H-BIM and collaborative reality-based annotation platforms

THÈSE dirigée par :

[Mme. CAROTI Gabriella]
[M. PIEMONTE Andrea]
[M. BEVILACQUA Marco Giorgio]
[M. DE LUCA Livio]
[M. VERON Philippe]

et co-encadrée par :
[M. JACQUOT Kévin]

Jury

Mme Roberta SPALLONE, Professeur, DAD, Politecnico di Torino	Présidente
M. Andreas GEORGOPOULOS, Professeur, NTU Athènes	Rapporteur
M. Manuel A. RODENAS LOPEZ, Professeur, UPCT Cartagena	Rapporteur
M. Philippe VERON, Professeur, EA 7515 LISPEN, Arts et Métiers	Examinateur
M. Livio DE LUCA, HDR CNRS, UMR 3495 MAP	Examinateur
M.me Gabriella CAROTI, Professeur, ASTRO DICI UNIPI	Examinatrice
M. Andrea PIEMONTE, Professeur, ASTRO DICI UNIPI	Examinateur
M. Marco Giorgio BEVILACQUA, Professeur, DESTeC UNIPI	Examinateur

THÈSE

A mia nonna, Teresa.
To my grandmother, Teresa.

Abstract

This work proposes an integrated methodological approach for the transfer, the retrieval and the exchange of semantic annotations of 2D/3D digital heritage models, by exploiting Artificial Intelligence techniques, H-BIM environments and collaborative reality-based annotation platforms as Aïoli (aioli.cloud). The proposed methodology is validated on relevant case studies of the French and Italian heritage, such as the Notre-Dame Cathedral in Paris and the Pisa Charterhouse.

In the cultural and architectural heritage domain, the plurality of available representation methods is often the source of data dispersion and entangles the collection of the wide variety of heterogeneous material related to the study of a heritage object.

Indeed, the increasing diffusion of digital models, even acquired by laser scanning or photogrammetric survey, raises the issue of *digital continuity,* intended as the need to ensure data collection, traceability, maintenance, interpretability and availability regardless the type of digital representation chosen.

In this context, the semantic annotation mechanism, understood as the association of knowledge-related (semantic) information to purely metric digital data, is fundamental to support the correct interpretation and sharing of digital heritage information, the latter including both 2D (images, ortho-photos, drawings) and 3D media (point clouds, meshes, parametric models etc.).

Considering Heritage-Building Information Modeling (H-BIM) systems and reality-based collaborative annotation platforms such as Aïoli as starting point towards the archival and sharing of semantic information, the objective of this work is to propose a methodological approach

enabling the transfer, the retrieval and the exchange of semantic annotations over 2D/3D digital heritage models.

A suitable method for the sharing, update and transfer of information, valid regardless of the type of digital model chosen for the representation, would indeed answer the need to archive, share and update data relating to a certain heritage artefact. To this task, and relying on original survey and analytical data, the proposed methodological approach develops according to the three phases of:

i. Semi-automatic semantic segmentation (classification) of surveying data through Artificial Intelligence.

ii. 2D/3D annotation transfer.

iii. H-BIM reconstruction, semantic structuring and insertion of localized information.

In detail, the semi-automatic semantic enrichment of digital data is investigated by application of Artificial Intelligence algorithms, enabling the interpretation and classification of raw data (e.g., point clouds, images, mesh) obtained from 3D surveying, according to the recognition of architectural components, the detection of degradation patterns, the material mapping, and so forth.

Then, the information obtained is transferred and propagated to multiple representation systems, from 2D to 3D and vice versa, also through the use of the Aïoli collaborative annotation platform.

At a final stage, the AI-based classification information is also exploited in view of the construction of H-BIM models starting from annotated survey data (the so-called scan-to-BIM process). The digital model so obtained results in a semantically structured representation, where the insertion of localized annotations can further be studied for restoration, conservation and dissemination purposes.

Since the thesis and the experimental works are developed within the framework of an international agreement for joint research doctoral thesis (*co-tutelle*), involving Italian and French research institutions, the different phases of the proposed approach are tested and assessed with reference to representative case studies of both the Italian and French architectural heritage: among them, besides a number of significant Churches and Museums of the Italian territory, it is worth noting the Notre-Dame Cathedral in Paris and the Pisa Charterhouse. Each time, the results are evaluated by considering the case-specific representation and restitution needs and requirements.

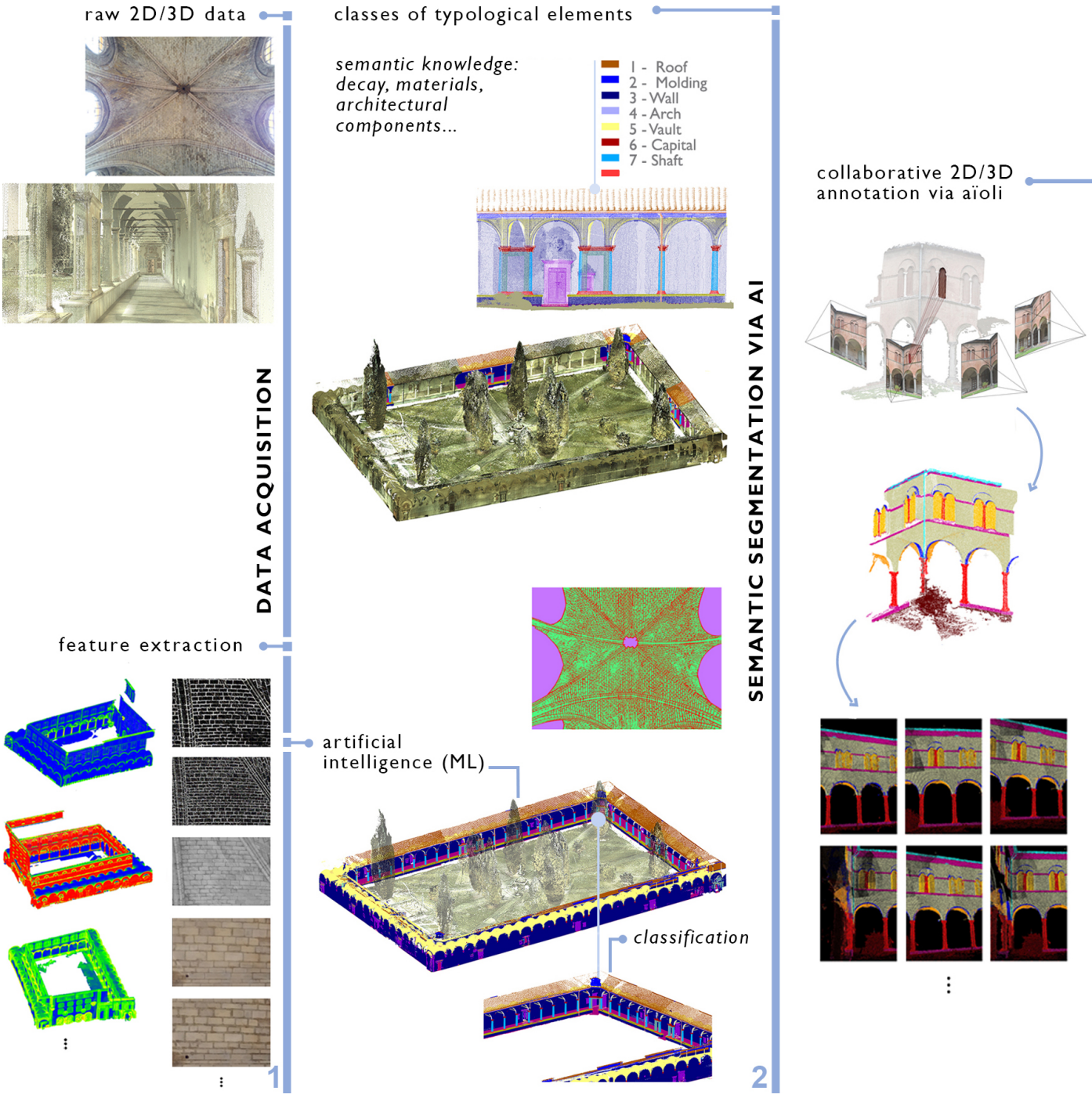
Through the proposed approach, a unified framework towards the exploitation and realization of semantically rich digital models is obtained,

to be applied in fields of study such as archaeology, architecture, civil engineering and art history, for the study of monuments, historical sites, buildings, sculptures, paintings, archaeological finds and works of art in general. The approach will be made available to restorers, engineers, architects, archaeologists, historians, and other experts who continuously deal with the issues of fusion, processing, and digital connection of cultural heritage data.

Keywords

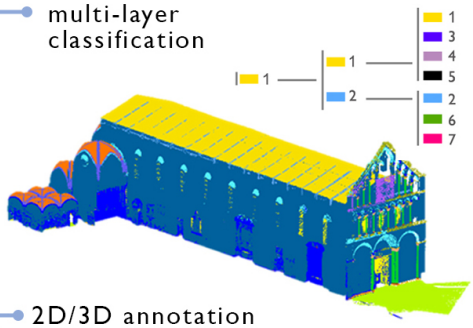
Semantic annotation; digital heritage; BIM; H-BIM; 3D survey; Artificial Intelligence; Machine Learning; semantic segmentation; Scan-to-BIM; point clouds.

Graphical abstract

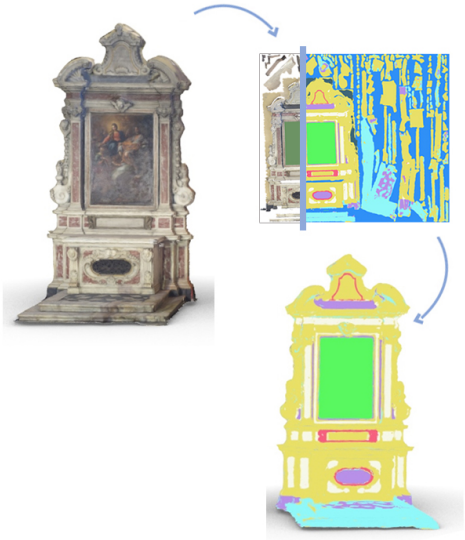


2D/3D INFORMATION PROPAGATION

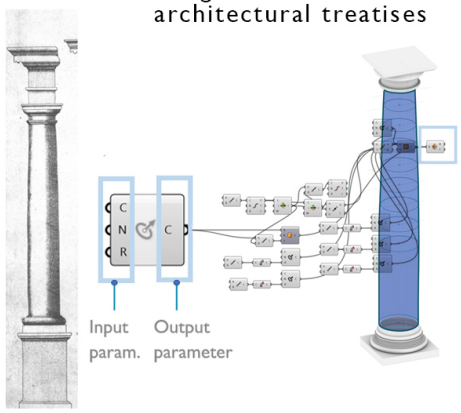
multi-layer classification



2D/3D annotation transfer



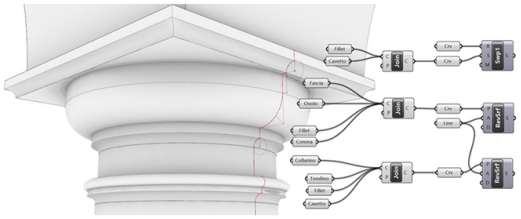
3D geometries: VPL and architectural treatises



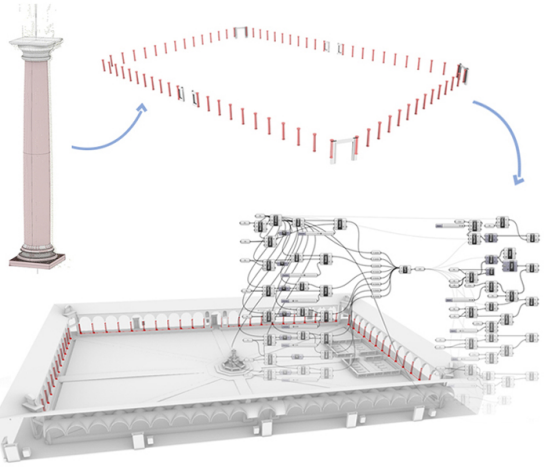
3

H-BIM RECONSTRUCTION, SEMANTIC STRUCTURING, LOCALIZED INFORMATION

template geometries



propagation of semantic information



Insertion of localized information

archival, amendment, updating of semantic annotations



4

Sommario

Trasferimento e recupero di annotazioni semantiche per il patrimonio architettonico.

Definizione di un approccio metodologico basato su Intelligenza Artificiale, H-BIM e piattaforme di annotazione collaborativa.

Nella presente tesi, si propone un approccio metodologico integrato per il trasferimento, il recupero e la condivisione di annotazioni semantiche a partire da modelli digitali 2D e/o 3D per i beni culturali, che sfrutti insieme tecniche di Intelligenza Artificiale, ambienti H-BIM e piattaforme di annotazione collaborativa e *reality-based* come Aïoli (aioli.cloud). La metodologia proposta è validata su casi di studio significativi del patrimonio architettonico francese e italiano, come la Cattedrale di Notre-Dame di Parigi e la Certosa Monumentale di Pisa.

Nell'ambito della documentazione digitale del patrimonio culturale e architettonico esistente, la pluralità di metodi di rappresentazione a oggi disponibili è spesso fonte di dispersione di dati e ostacola notevolmente la raccolta e l'archiviazione della grande varietà di materiale eterogeneo legato allo studio di un bene.

La crescente diffusione di modelli digitali – anche acquisiti a partire dal rilievo laser scanner o fotogrammetrico – pone quindi il problema della *continuità digitale*, intesa come la necessità di garantire la raccolta, la tracciabilità, la conservazione, l'interpretabilità e la disponibilità dei dati indipendentemente dal tipo di rappresentazione digitale che si scelga di utilizzare.

In questo contesto, il meccanismo di annotazione semantica, inteso come l'associazione di informazioni legate alla conoscenza (*semantiche*) al dato digitale puramente metrico, è fondamentale per assicurare la corretta interpretazione e condivisione delle informazioni sul patrimonio culturale digitale, sia in formato 2D (che si tratti, ad esempio, di immagini, ortofoto, disegni) che in formato 3D (come nuvole di punti, mesh, modelli parametrici).

Considerando i sistemi di tipo Heritage-Building Information Modeling (H-BIM) e le piattaforme di annotazione collaborativa di modelli *reality-based* quali Aïoli come un punto di partenza essenziale per l'archiviazione e la condivisione di informazioni semantiche, obiettivo di questo lavoro è la definizione di un approccio metodologico che consenta il trasferimento, il recupero e lo scambio di annotazioni semantiche su modelli digitali 2D/3D dei beni culturali.

La costruzione di un metodo adeguato di condivisione, aggiornamento e trasferimento di informazioni, da applicare indipendentemente dal tipo di modello virtuale scelto per la rappresentazione, risponderebbe infatti all'esigenza di archiviare, condividere e aggiornare costantemente i dati digitali relativi alla documentazione di un certo edificio o di un monumento oggetto di studio.

Per raggiungere questo obiettivo, e a partire da dati di rilievo e analisi, l'approccio metodologico proposto si sviluppa secondo le tre rispettive fasi di:

i. Segmentazione semantica (classificazione) semi-automatica dei dati digitali di rilievo attraverso algoritmi di Intelligenza Artificiale (IA).

ii. Trasferimento delle annotazioni 2D/3D.

iii. Ricostruzione H-BIM, strutturazione semantica e inserimento di informazioni localizzate.

In dettaglio, la prima fase prevede l'applicazione di algoritmi di IA per consentire l'arricchimento semantico dei dati digitali. Con un approccio semi-automatico, vengono classificati e interpretati i dati grezzi ottenuti dal rilievo (nuvole di punti, immagini, mesh) per riconoscere, ad esempio, componenti architettoniche e tipologiche ricorrenti di un edificio, o ancora per restituire lo stato di degrado delle superfici, per mappare i materiali, e così via.

Quindi, le informazioni ottenute vengono trasferite e propagate a più sistemi di rappresentazione, dai supporti 2D ai modelli 3D e viceversa, anche attraverso l'uso della piattaforma di annotazione collaborativa Aïoli.

I dati 2D e 3D, classificati semi-automaticamente tramite IA, sono poi ulteriormente sfruttati in vista della ricostruzione di modelli BIM a partire dal dato di rilievo semanticamente annotato (cosiddetto processo *Scan-to-BIM*).

Il modello digitale ottenuto a conclusione di questo processo risulta una rappresentazione semanticamente strutturata, che può essere poi ulteriormente arricchita tramite inserimento di annotazioni localizzate

(relative, ad esempio, alla mappatura materica o del degrado su porzioni ridotte dal dato), ad esempio per scopi di restauro e conservazione.

I contenuti della presente tesi, e quindi il relativo contesto sperimentale, sono sviluppati nell'ambito di un accordo internazionale di co-tutela di tesi di dottorato, che coinvolge istituti di ricerca italiani e francesi. Di conseguenza, l'approccio metodologico proposto è valutato con riferimento a casi studio rappresentativi del patrimonio architettonico sia italiano che francese: tra questi, oltre ad alcune chiese e musei significativi del territorio italiano, si segnalano in particolare la Cattedrale di Notre-Dame di Parigi e la Certosa di Pisa.

Di volta in volta, i risultati dell'applicazione della metodologia proposta vengono valutati considerando, caso per caso, le specifiche esigenze di descrizione, di analisi e di restituzione.

Attraverso l'approccio così costituito, si delinea un quadro unificato di realizzazione e di fruizione di modelli digitali semanticamente ricchi, da poter applicare in campi di studio quali l'archeologia, l'architettura, l'ingegneria civile e la storia dell'arte, per lo studio di monumenti, siti storici, edifici, sculture, dipinti, reperti archeologici e opere d'arte in generale. L'approccio sarà messo a disposizione di restauratori, ingegneri, architetti, archeologi, storici e altri esperti che affrontano continuamente gli aspetti legati alla fusione, all'elaborazione e alla connessione digitale dei dati relativi al nostro patrimonio culturale.

Parole chiave

Annotazione semantica; patrimonio culturale digitale; BIM; H-BIM; rilievo 3D; Intelligenza Artificiale; apprendimento automatico; segmentazione semantica; Scan-to-BIM; nuvole di punti

Schema di sintesi

Lo schema di sintesi, con relativa traduzione in lingua italiana delle parti di testo, è disponibile in Appendice 1 (Appendix 1).


UNIVERSITÀ DI PISA

UNIVERSITÀ
DEGLI STUDI
FIRENZE

Arts Sciences et Technologies
et Métiers
lispen

culture
mar
cnrs


UNIVERSITÉ
FRANCO
ITALIENNE
UNIVERSITÀ
ITALO
FRANCESE
HESAM
UNIVERSITÉ

Résumé

Transfert et repérage d'annotations sémantiques pour le patrimoine architectural.

Un système méthodologique combinant Intelligence Artificielle, H-BIM et plateformes d'annotation collaborative reality-based.

Cette thèse de doctorat propose une approche méthodologique intégrée pour le transfert, la récupération et le partage d'annotations sémantiques, à partir de modèles numériques 2D et/ou 3D du patrimoine culturel, et en combinant les techniques d'Intelligence Artificielle, les environnements H-BIM et les plateformes d'annotation collaboratives et basées sur la réalité (*reality-based*) telles que Aïoli (aioli.cloud). La méthodologie proposée est validée sur des études de cas significatives du patrimoine architectural français et italien, comme la cathédrale Notre-Dame de Paris et la Chartreuse de Pise.

Dans le cadre de la documentation numérique du patrimoine culturel et architectural existant, la pluralité des modes de représentation aujourd'hui disponibles est souvent source de dispersion des données et constitue un frein considérable à la collecte et à l'archivage de la grande variété de matériaux hétérogènes liés à l'étude d'un artefact patrimonial.

La diffusion croissante des modèles numériques – même issus du balayage laser ou de relevés photogrammétriques – pose donc le problème de la *continuité numérique*, entendue comme la nécessité de garantir la collecte, la traçabilité, la conservation, l'interprétabilité et la disponibilité des données indépendamment du type de représentation numérique que l'on choisit d'utiliser.

Dans ce contexte, le mécanisme d'annotation sémantique, entendu comme l'association d'informations (*sémantiques*) liées à la connaissance à des données numériques purement métriques, est indispensable pour permettre l'interprétation et le partage corrects des informations numériques relatives au patrimoine culturel, tant en

format 2D (par exemple, images, ortho-photos, dessins) qu'en format 3D (par exemple, nuages de points, maillages, modèles paramétriques).

Considérant les systèmes de Modélisation des Informations de la Construction (H-BIM avec l'acronyme anglais) et les plateformes collaboratives d'annotation de modèles basés sur la réalité, telles que Aïoli, comme un point de départ essentiel pour le stockage et le partage d'informations sémantiques, ce travail a pour objectif de définir une approche méthodologique qui autorise le transfert, la récupération et l'échange d'annotations sémantiques sur des modèles numériques 2D/3D du patrimoine culturel.

L'élaboration d'une méthode appropriée de partage, de mise à jour et de transfert de l'information, à appliquer indépendamment du type de modèle virtuel choisi pour la représentation, répondrait en effet aux besoins d'archivage, de partage et de mise à jour des données numériques relatives à la documentation d'un bâtiment ou d'un monument étudié.

Pour atteindre cet objectif, et à partir des données de relevé et d'analyse, l'approche méthodologique proposée repose sur les trois phases suivantes :

i. Segmentation (classification) sémantique semi-automatique de données de relevé numérique par le biais d'algorithmes d'intelligence artificielle (IA).

ii. Transfert d'annotations 2D/3D.

iii. Reconstruction H-BIM, structuration sémantique et insertion d'informations localisées.

Dans le détail, la première phase consiste à appliquer des algorithmes d'IA pour permettre l'enrichissement sémantique des données numériques. Par une approche semi-automatique, les données brutes issues du relevé (nuages de points, images, maillages) sont classées et interprétées afin de reconnaître, par exemple, les composantes architecturales et typologiques récurrentes d'un bâtiment, ou bien afin de restituer l'état de dégradation des surfaces, de cartographier les matériaux, etc.

Ensuite, les informations obtenues sont transférées et propagées vers plusieurs systèmes de représentation, des médias 2D aux modèles 3D et vice versa, également par l'utilisation de la plateforme d'annotation collaborative Aïoli.

Les données 2D et 3D, classées de manière semi-automatique par l'IA, sont alors exploitées en vue de la reconstruction de modèles BIM à partir des données de relevé qui ont été annotées sémantiquement

(processus dit *Scan-to-BIM*).

Le modèle numérique obtenu au terme de ce processus est une représentation sémantiquement structurée, qui peut ensuite être enrichie par l'insertion d'annotations localisées (relatives, par exemple, à la cartographie des matériaux ou à la dégradation sur de petites portions des données), par exemple à des fins de restauration et de conservation.

Le contenu de cette thèse, ainsi que son contexte expérimental, sont développés dans le cadre d'un accord international de *co-tutelle* de thèse de doctorat impliquant des institutions de recherche italiennes et françaises. Par conséquent, l'approche méthodologique proposée est évaluée en se référant à des études de cas représentatives du patrimoine architectural italien et français: parmi celles-ci, outre certaines églises et musées significatifs du territoire italien, on peut notamment mentionner la Cathédrale Notre-Dame de Paris et la Chartreuse de Pise.

A chaque fois, les résultats de l'application de la méthodologie proposée sont évalués en considérant, au cas par cas, les besoins spécifiques de description, d'analyse et de restitution.

Grâce à l'approche ainsi constituée, on esquisse un cadre unifié pour la réalisation et l'utilisation de modèles numériques sémantiquement riches, qui peut être appliqué dans des domaines d'étude tels que l'archéologie, l'architecture, le génie civil et l'histoire de l'art, pour l'étude des monuments, des sites historiques, des bâtiments, des sculptures, des peintures, des découvertes archéologiques et des œuvres d'art en général. L'approche sera mise à la disposition des restaurateurs, ingénieurs, architectes, archéologues, historiens et autres experts qui sont continuellement confrontés aux aspects de la fusion, du traitement et de la connexion numérique des données relatives à notre patrimoine culturel.

Mots clés

Annotation sémantique; patrimoine culturel numérique; BIM; H-BIM; levé 3D; intelligence artificielle; apprentissage automatique; segmentation sémantique; Scan-to-BIM; nuages de points.

Résumé graphique

Le résumé graphique, avec la traduction du texte en français, est disponible en Annexe 2 (Appendix 2).

Index

1. Introduction I

1.1. General context I

1.2. Objective IV

1.3. Methodology V

1.3.1. 2D/3D data acquisition. V

1.3.2. Data processing and interpretation: Semantic segmentation via Artificial Intelligence V

1.3.3. 2D/3D Information propagation VI

1.3.4. H-BIM reconstruction, semantic structuring and transfer of localized information VI

1.4. Experimental context and case studies..... VIII

1.5. Thesis structure X

Part I

2. The role of heritage digitization 3

2.1. Monuments and Cultural Heritage. Definitions 5

2.2. Digital documentation as a requirement for heritage preservation 7

2.3. Architectural survey as a means of knowledge..... 10

2.3.1. Traditional / Digital survey 11

2.3.2. In-situ and ex-situ survey 13

2.4. The digital continuity concept: visualizing and ordering the complexity of heritage artefacts 15

2.5. Heritage- or Historic- Building Information Modeling 17

3. Acquisition, processing and semantic enrichment of digital heritage objects.....21

3.1. 3D surveying techniques..... 26

3.1.1. Laser scanning..... 28

3.1.2. Photogrammetry..... 32

3.1.3. Comparative analysis of existing 3D surveying techniques 41

3.2. Processing of 2D/3D information..... 44

3.2.1. 2D visual and metric descriptors..... 44

3.2.2. 3D visual and metric descriptors..... 48

3.2.3. Comparative analysis of 2D/3D processing methods and research challenges 59

3.3. Digital information systems for the semantic enrichment of heritage models..... 60

3.3.1. Classification based on part/whole relationship 61

3.3.2. Classification based on the degree of automation..... 62

3.3.3. Classification based on type of annotated media 63

3.3.4. The reality-based annotation platform Aïoli 70

3.3.5. Heritage-Building Information Modeling 74

3.3.6. Semantic segmentation 88

3.3.7. Critical analysis of existing 2D/3D annotation platforms 98

3.4. Conclusions and research challenges 101

3.4.1. More automated semantic annotation procedures 102

3.4.2. Annotation transfer between different representation systems 102

3.4.3. Construction of a formalized knowledge base..... 103

Part II

4. The rationale of the method 107

4.1. Semantic segmentation via Artificial Intelligence 112

4.2. 2D/3D Information propagation 113

4.3. H-BIM reconstruction, semantic structuring and transfer of localized information..... 114

4.4. Experimental context and choice of emblematic case studies 115

5. Semantic segmentation via Artificial Intelligence ... 119

5.1. Semantic description of the building's morphology 120

5.2. Machine Learning approaches for 2D/3D semantic segmentation. Methods 123

5.2.1. Geometry-based and texture-based approaches 125

5.2.2. Feature extraction..... 127

5.2.3. Feature selection..... 131

5.2.4. Manual annotation of a training set..... 132

5.2.5. Random Forest classifier and evaluation of the trained model 133

5.3. Results: annotated 2D/3D data..... 136

5.3.1. Geometry-based classification results 136

5.3.2. Texture-based classification results 152

6. 2D/3D Information propagation 167

6.1. Collaborative 2D/3D annotation transfer through Aïoli..... 169

6.1.1. Spatialization of 2D/3D data within Aïoli..... 172

6.1.2. Implementation of semi-automatic annotation methods via Aïoli..... 172

6.1.3. Image-based annotation on Aïoli and construction of a 2D/3D training set..... 176

6.1.4. IA-based point cloud classification and back interpolation 180

6.1.5. 3D-2D back projection and annotation of the whole image set 182

6.2. Additional 2D/3D semi-automatic annotation transfer possibilities 184

6.2.1. UV map – mesh models 184

6.2.2. From classified image sets to 3D point clouds 186

6.2.3. Spatialization of archival images onto the 3D model 190

6.3. Multi-level and multi-resolution point cloud semantic segmentation 194

6.4. Conclusions 201

7. From the Semantic point cloud to Heritage-Building Information Modeling 203

7.1. Methods 205

7.1.1. A semi-automatic Scan-to-BIM workflow for the reconstruction of template models 205

7.1.2. Reference to architectural treatises 207

7.1.3. Generative modeling and Visual Programming Language 208

7.2. Results: Visualization and import of the annotated classes 210

7.3. Results: Libraries of template geometries 212

7.4. Results: Information propagation 218

7.5. Import into BIM-based software for Rhino.Inside Revit 221

7.6. Hints to rules of association between template model – real model 222

7.7. Assessment of the proposed approach 227

8. Conclusions and future perspectives 231

8.1. Assessment of the proposed method and main contributions 231

8.1.1. Semi-automatic segmentation (classification) of surveying data through Artificial Intelligence 233

8.1.2. 2D/3D annotation transfer 234

8.1.3. H-BIM reconstruction, semantic structuring and insertion of localized informatione 235

8.2. Research perspectives 236

8.2.1. Definition of unique development environment and interface 236

8.2.2. Use of hybrid 2D/3D descriptors to improve the manual annotation phase within Aïoli 237

8.2.3. Improvement of existing benchmarks for the construction of Deep Neural Networks 239

8.2.4. Levels of detail within H-BIM environments and integration of localized information 239

8.2.5. Extension to other case studies 239

9. Bibliography 241

10. Sitography 256

11. List of abbreviations 259

12. Appendix 261

Appendix 1 – Schema di sintesi 262

Appendix 2 – Résumé graphique 264

Appendix 3 – Aïoli class indexing script 266

Appendix 4 – Point cloud file formats 269

Appendix 5 – List of publications 270

Appendix 6 – Digital&BIM Italia. Competition panels. 274

Appendix 7 – Résumé en langue française 280

Acknowledgments 300

Chapter 1

Introduction

1.1. General context

The ever-growing development of digital acquisition, graphic restitution and modeling systems has radically changed the cultural heritage sector, proposing a new approach to conservators, architects, engineers, restorers and other experts towards the analysis, preservation, documentation and dissemination of heritage assets.

The solutions proposed over time for the visualization and fruition of digital data are widely proven. Tools and methods enabling bi-dimensional (2D, e.g.: images, ortho-photos etc.) or three-dimensional (3D, e.g.: point clouds, mesh models etc.) virtual representations are available to a plurality of actors, even coming from different and multidisciplinary studies of the cultural heritage sector.

At the same time, the prompt development of acquisition techniques has enabled a high amount of digital data to be available in a relatively short time; 3D surveying from laser scanning or photogrammetry is an effective means to support the graphic restitution of architectural artifacts by detection of their shapes and materials, dimensions and proportions, appearance and contours. As such, a virtual replica of existing heritage objects and sites can be obtained with extreme precision. This replica could be intended as a *digital twin*, serving as a *digital counterpart* of the physical object, in which the information that is embedded within this latter can be digitally stored (Grieves, 2014).

However, although great progress is today accomplished in the field

of acquisition and rendering systems, digital 2D/3D data are still far from being understood as privileged means for the management of the information on studies, analyses, and for the retrieval of all the other documentation that revolves around an architectural heritage object.

To date, digital models are not integrated and collaborative digital tools, as they do not support proper archival, access and reasoned usage of both *qualitative* (metrically correct) and *quantitative* (referring to the knowledge of the asset) information, related to a studied object. Hence, the pure geometric representation given by the digital model is not in itself sufficient to meet the concrete needs of observation, processing and interpretation of heterogeneous data and documentary sources produced in the context of heritage studies.

This is all the truer as if considering the multiplicity of products, spreadsheets, 2D drawings and 3D models used, often the cause of dispersion of data and, consequently, of knowledge. The heterogeneous nature of the information produced, resulting from the involvement of professionals from different study domains, motivates the need for the development of novel digital methods and tools for the storage, organization and arrangement of the documentation related to a heritage object, as well as the implementation of common decision support mechanisms.

Digital information systems aimed at connecting the purely geometric-visual representation to *semantic* data, i.e., linked to the analysis of the asset, are therefore increasingly sought after, and open the way to new multi-modal and multi-user approaches to investigate, enhance, preserve cultural heritage.

In this context, on the one hand, Building Information Modeling (BIM) processes are emerging for the construction of shared digital representations supporting the information management within the design, construction and operation processes, so to form a reliable basis for decisions (ISO 19650:2019). The application of BIM systems to the heritage domain is rather recent: the concepts of BIM, H-BIM, scan-to-BIM, which are used today with disparate meanings and sometimes controversial interpretations, are appropriately introduced in Section 2.5 in the definition that it is intended to give them in the present work.

On the other hand, modern Artificial Intelligence (AI) techniques, if suitably transferred to the architectural representation domain, could improve the interpretation and classification of the multiple digital data available. Their implementation would enable the access, exchange and continuous update of knowledge-related architectural heritage information, that may concern e.g., the alterations of a building, the documentation of its changes over time, the restoration and conservation interventions, the state of degradation of its surfaces, the mapping of materials and so forth.

In this perspective, a more concrete interpretation of what is represented can be envisaged: a semantic reading of surveying data would enable a streamlined process of collection, elaboration, visualization of information directly within the 2D or 3D digital representations. The introduction of semantically enriched models hence offers a privileged tool for the description, analysis and understanding of the entire cognitive universe that hinges on the study of heritage objects and sites.

In current Heritage-Building Information Modeling (H-BIM) or reality-based annotation platforms, the digital representations are associated with meaningful information by inserting suitable tags, usually called annotations, directly over the digital model. However, existing solutions only allow annotations over a single type of 2D or 3D representation used, a point cloud, a mesh, a parametric model, a technical 2D drawing, an ortho-photo, an archival image, a photo. Consequently, when the digital media is modified, further usage, exploitation and updating of crucial information concerning the considered heritage object is prevented, unless the information previously acquired is suitably re-inserted in the new representation.

This is the bottleneck of existing semantic information systems for digital heritage: they remain anchored to the specific type of representation they support, preventing the shift, structuring, and further inter-linking of the information produced by the various professionals involved.

A suitable method for the sharing, update and transfer of information, valid regardless of the type of digital model chosen for the representation, would indeed answer the need to archive, exchange and update data relating to a certain cultural asset.

The present research stems from the necessity to ensure a more objective and shared approach for data storage and organization on digital objects, so to allow to fully understand, structure, access, archive, update and correlate, starting from initial survey data, the wealth of information produced in the context of heritage studies. The work thus presents a unified framework towards the exploitation and fruition of semantically rich heritage models, that can be beneficial to the community of restorers, engineers, architects, conservators, and all other stakeholders involved, who are continuously facing the problem of merging, documenting, and connecting heritage data, as well as linking them for restoration, management, and conservation policies.

1.2. Objective

The objective of this thesis work is to define an integrated methodological approach for the transfer, the retrieval and the exchange of semantic annotations of 2D/3D digital heritage models, by exploiting Artificial Intelligence techniques and the construction of H-BIM environments. Focusing on the semantic annotation mechanism for the enrichment of digital replicas, purpose of the work is to ensure an *informative continuum* in the interrelated phases of heritage data acquisition, observation, processing, representation.

The following open research questions are addressed:

- Identifying a connection between annotations made on different digital supports (a point cloud, a mesh, a parametric model, a drawing, an ortho-photo, a digital photo etc.);
- Classifying 2D and 3D heritage data in a more automated and streamlined way;
- Propagating information between 2D digital documents and 3D models, or even between texts, images, ortho-photos, and 3D models;
- Making a semantic annotation become *interoperable*, in such a way that information can be read and accessed regardless the type of digital support in which it is originally drawn up,

Even foreseeing different levels of information reading, editing, updating so that the different experts of the Cultural Heritage domain can access semantic data associated with a digital model, even based on their reciprocal needs, positions and requirements. In tackling these open issues, the goal of this work is the definition of a common system of input and retrieval of information within semantic digital models, to be:

- Independent of the program used for data spatialization and insertion;
- Insensitive to variations in the real geometry of the studied object, both in terms of the level of precision of the representation and/or type of model used.

As such, the semantic information could be located and stored virtually over the heritage object, regardless the type of digital representation chosen and its modifications over time. This implies that raw or organized data, concerning any kind of information e.g., history, images, textures, degradations, crack patterns, interventions and so on, could be interconnected directly over the digital representation.

1.3. Methodology

The proposed methodological approach relies on the identification of relevant and recurring elements of the heritage object, that are annotated, retrieved and semi-automatically propagated on and between different types of digital models, either 2D or 3D. As such, the information is associated and linked each time to representative elements identified over the digital model, thus ensuring the transfer and retrieval of significant data. The semantic information is indeed located and stored virtually over the heritage object, regardless the type of digital representation chosen and its modifications over time.

The formulation of this methodological approach requires the ability, on the one hand, to manage the digital acquisition, restitution and interpretation of existing geometric, visual, descriptive forms of an architectural asset; on the other hand, to define an appropriate digital system for the storage, transfer and exchange of the variety of information produced. The definition of the method starts from 2D/3D data acquisition and is broken down into three respective phases: *i)* semantic segmentation via Artificial Intelligence; *ii)* 2D/3D information propagation; *iii)* H-BIM reconstruction, semantic structuring and transfer of localized information.

1.3.1. 2D/3D data acquisition

Current systems for the surveying, restitution and modeling of the shapes of a building, aimed at the construction of a *digital twin*, are taken into account. In this phase, the integration of existing methods and techniques for 3D data acquisition – among which laser scanning and photogrammetry – is assessed, considering their applications in the general context of heritage studies.

1.3.2. Data processing and interpretation: Semantic segmentation via Artificial Intelligence

The raw data derived from 3D surveying are deconstructed so to interpret the nature of what is represented (e.g., in terms of identification of typological and recurring elements, of materials or decay mapping etc.). The 3D models acquired from the observation of reality are thus appropriately simplified, by identifying, isolating and classifying each time the different elements that compose them, according to the case-specific needs. This process, known as *semantic segmentation*, takes place via a more automated mechanism that is based on the application

of Artificial Intelligence (AI) techniques to raw survey data (point clouds, meshes, photos) and that allows to suitably identify relevant and recurring architectural components.

1.3.3. 2D/3D Information propagation

In this phase, the semantic information, either obtained manually or semi-automatically (through AI), is propagated from the source digital object to other existing representation types. Different possibilities for semantic annotation transfer are explored, enabling the shift between different digital media, e.g., from a point cloud to an image, from a mesh to an ortho-photo, and so forth.

Collaborative annotation transfer methods are also proposed, leveraging hybrid 2D/3D annotation systems, in the form of an implementation of the Aïoli platform. *Multi-level* and *multi-resolution* workflows for the propagation of annotations are later illustrated for the annotation of point clouds at different levels of detail.

1.3.4. H-BIM reconstruction, semantic structuring and transfer of localized information

In a latter phase, the semi-automatic annotation of architectural components on 3D surveying data is taken into account to foster Scan-to-BIM reconstruction processes. To these purposes, reference *template* geometries are reconstructed in digital modeling environments starting from semantically segmented data. Rules and canons derived from architectural treatises of historic architecture are appropriately translated to formalize and guide the 3D conceptual reconstruction of the buildings' components. Then, again through a semantic annotation process, the information on the reconstructed geometries is propagated to all other elements that present similar characteristics, for the generation of an overall reference model. This approach for the reconstruction of a semantically rich 3D model is in line with the logic of H-BIM environments and streamlines the so-called Scan-to-BIM processes.

At a final stage, a possible research development is identified to enable the insertion of localized information relating, e.g., to portions of the constructed *template* geometries and describing, for instance, degradation patterns, material mappings and conservation states. The operations therein described allow to associate the *template* constructed model to any kind of real representation of the object, always preserving information retrieval and export, regardless of the type of digital media being chosen.

By such a procedure, the available informative documents and spreadsheets concerning the analysis and documentation of a heritage asset can be structured and organized within the reconstructed *template* model, through the mechanism of annotation.

1.4. Experimental context and case studies

The present work focuses on the elaboration and interpretation of survey data for the reconstruction and sharing of semantically rich 3D models. The thesis and the experimental works are developed within the framework of an international agreement for joint research doctoral thesis (*co-tutelle*), involving the following institutions:

- The Laboratory on Scientific and Topographic Applications for Operative Surveying, from the Department of Civil and Industrial Engineering of the University of Pisa (ASTRO, DICI, UNIPI);
- The Department of Energy, System, Territory and Construction Engineering, from the University of Pisa (DESTeC, UNIPI);
- The Joint research unit Models and Simulations for Architecture and Heritage (Modèles et Simulations pour l'Architecture et le Patrimoine) under the authority of the National Centre for Scientific Research and the French Ministry of Culture (UMR 3495 CNRS/MC MAP);
- The Department of Civil and Environmental Engineering from the University of Florence (DICEA, UNIFI), in the framework of the International Doctorate in Civil and Environmental Engineering (Cycle XXXIV, curriculum Construction design, verification and control);
- The Physical and Digital Systems Engineering Laboratory from the Ecole Nationale Supérieure d'Arts et Métiers Paris-Tech in Aix-en-Provence (LISPEN, unit n. 7515, ENSAM Aix-en-Provence), in the framework of the doctoral school in Design Engineering (Ecole doctorale Sciences et Métiers de l'Ingénieur, ED SMI 432).

The work is funded by the Tuscany Region (Borse Pegaso POR FSE TOSCANA 2014/2020). It received the Vinci 2019 award (Chapter II – Mobility fundings for Ph.D. thesis in co-tutelle) from the Université Franco-Italienne, a supranational institution promoting research collaborations between Italy and France.

In December 2021, the Ph.D. thesis was awarded the first prize of the BIM&DIGITAL Awards 21 for Category 8 – Research, organized by the associations CLUST-ER BUILD and DIGITAL&BIM Italia (see Appendix 6 – Digital&BIM Italia. Competition panels).

In May 2022, the work was then recognized the first prix for Ph.D. thesis 'Il restauro nell'era dell'H-BIM' (en.: *restoration in the era of H-BIM*), organized by Rotary Club di Roma with the patronage of Facoltà di

Architettura e Facoltà di Ingegneria Civile e Industriale di Sapienza Università di Roma and in collaboration with Dipartimento di Storia, Disegno e Restauro dell'Architettura Scuola di specializzazione in Beni Architettonici e del Paesaggio Fondazione Almagià – ACER.

Concerning the experimental context, the different phases of the proposed methodological approach are tested and assessed with reference to representative case studies of both the Italian and French historical buildings and monumental complexes: as part of the artistic and cultural heritage of the two countries involved in the co-tutelle agreement, these sites require the provision of the whole geometric, graphic and informative apparatus, at multiple scales of reading and within digital information systems able to express and convey their complexity by annotation of 2D/3D representations. Each time, the results are evaluated by considering the case-specific representation and restitution needs and requirements.

Among the considered case studies, besides a number of significant Churches and Museums of the Italian territory, it is worth noting the Notre-Dame Cathedral in Paris and the Pisa Charterhouse.

For the latter, the presented work is developed as part of the project Studi conoscitivi e ricerche per la conservazione e la valorizzazione del Complesso della Certosa di Calci e dei suoi Poli Museali (en: Cognitive studies and research for the conservation and enhancement of the Calci Charterhouse complex and its museum complexes), promoted by the University of Pisa and coordinated by Professor Marco Giorgio Bevilacqua.

The broad range of case studies collected and analyzed during the three years of the Ph.D. program is considered for the experimental context of this thesis, as well as for the validation and further verification of the proposed methodological approach: churches, monuments, museums widespread over the Italian and French territories surveyed with application and/or integration of different techniques e.g., laser scanning, UAV- or terrestrial- photogrammetry, alongside traditional instrumentations. On the considered case studies:

- Semi-automatic procedures for the semantic segmentation of 3D data, such as those derived from the most recent applications of Artificial Intelligence (Machine Learning and Deep Learning) are applied to classify and propagate information starting from surveying data;
- Current H-BIM platforms constitute the informative basis for the insertion and storage of geometric and semantic information;
- Procedural modeling tools and visual programming

environments are used for the definition and subsequent subdivision of conceptual source 3D models.

1.5. Thesis structure

The thesis is structured as follows: Part I deals with the state of the art on modern digital documentation systems for cultural heritage, with an overview of current methods of acquisition, processing and modeling of 2D or 3D digital data. Contributions, limits and perspectives of use of such systems are recognized, operating comparative analyses and identifying the main research challenges that led to the definition of the proposed methodology.

In detail, Chapter 2 introduces several relevant concepts related to heritage digitization and fruition that will be extensively used throughout this dissertation: Cultural heritage and monuments, digital continuity, 3D surveying, semantic enrichment and H-BIM. In Chapter 3, current practices of digital heritage data acquisition, processing and semantic enrichment are extensively reviewed and discussed, with a thorough analysis on visual and metric descriptors that can be extracted from raw data and on more or less automatic existing semantic annotation systems as well.

In Part II, the proposed methodological approach is presented and its experimental developments and applications to significant case studies are illustrated. The proposed workflow allows retrieval, use and updating of semantic knowledge concerning a heritage artifact via consolidation/ visualization of its digital replica and annotation transfer. An in-depth regard is laid on the application of AI-based classification methods and on the automation of Scan-to-BIM processes.

In detail, Chapter 4 presents the rationale and related overall flowchart of the proposed approach, while Chapters 5, 6 and 7 are devoted to the illustration of the different methods, and related results, obtained for the three respective phases of the proposed workflow. Concluding remarks and future developments of this thesis are finally outlined in Chapter 8.

Part I

Chapter 2

The role of heritage digitization

Documentation is a fundamental and necessary prelude to any effort towards conservation, recovery and preservation of the historical and artistic heritage, as it entails a permanent record of the state and characteristics of a monument or site. This persistence of heritage information is all the more important in the historical period under way, threatened by on-going attrition and wars, health crisis, climate change effects, natural or man-made disasters and negligence in general (Chiabrando et al., 2018; Martínez-Espejo Zaragoza et al., 2017; Achille et al., 2015), which expose monuments and sites to continuous change, evolution, loss, and even total destruction (Remondino and Rizzi, 2010).

Article 3 of the Venice Charter for the Conservation and Restoration of Monuments and Sites states:

'The intention in conserving and restoring monuments is to safeguard them no less as works of art than as historical evidence.'

(ICOMOS, 1964)

In an effort to preserve the memory of these places over time, digital systems now constitute a fundamental resource. In this regard, problems related to the continuous accessibility, retrieval and traceability of the digital material arise. Some fundamental concepts and definitions concerning the digital fruition and exploitation of heritage data are addressed in this chapter, as they will be extensively used and mentioned throughout the dissertation: Cultural Heritage and monuments (Sections 2.1 - 2.2), digital surveying (Section 2.3), digital continuity (Section 2.4), semantic enrichment and Heritage-Building Information Modeling (Section 2.5).

In Sections 2.2 - 2.3, it is stressed how the practices of acquisition, measurement, surveying and restitution have evolved over time, leading to the development of a broad variety of systems and modes of representation enabling to permanently record in the three dimensions the condition and morphology of a building or site. The deep connection that subsists between metric survey, graphic representation and semantics is further illustrated, highlighting how, over time, the most complex work of heritage analysis and interpretation has been transferred from the in-situ acquisition phase to the ex-situ elaboration process.

2.1. Monuments and Cultural Heritage. Definitions

The Venice Charter, the International Charter for the Conservation and Restoration of Monuments and Sites, drafted in 1964 after the 2nd International Congress of Architects and Technicians of Historic Monuments, is a text of historical importance, considered a reference for the definition of the term monument and for the formulation of the international code for restoration and documentation of monuments and sites. According to Article 1 of the Charter:

'The concept of a historic monument embraces not only the single architectural work but also the urban or rural setting in which is found the evidence of a particular civilization, a significant development or a historic event. This applies not only to great works of art but also to more modest works of the past which have acquired significance with the passing of time.'

(ICOMOS, 1964)

This notion of historical monument as evidence and everlasting remembrance of the culture and traditions of a place has a long history. Looking back to the past, Cicero, in his oration *de signis* (4th oration of the cycle *In Verrem*), already sustained the importance of works of art:

'Omnia monumenta dicuntur quae faciunt alicuius rei recordationem' [1]

(Eng.: Monuments are said to be all those that evoke remembrance of something)

Emphasizing their value as historical legacies and heritages to be transmitted to future generations.

The conception of historical monuments as ***living witnesses of their age-old traditions*** (as they are described in the preamble to the Venice Charter) has led, over time, to mature the definition, even broader, of *Cultural Heritage*, thanks to the efforts of the United Nations Educational, Scientific and Cultural Organization Convention (UNESCO Convention), held in Paris in year 1972.

In this framework, the cultural heritage of each nation has been distinguished according to respective three categories of: *monuments, groups of buildings and sites*. Specifically, heritage monuments are defined in Article 1 as:

'Architectural works, works of monumental sculpture and painting, elements or structures of an archaeological nature, inscriptions, cave dwellings and combinations of features, which are of outstanding universal value from the point of view of history, art or science.'

(UNESCO, 1972)

[1] J. C. Orellius, M.Tullii Ciceronis opera Bd. V.2, Zurich 1833. pp. 391/3 f

With regard to the definition given in this context, which in a certain sense confines Cultural Heritage to noteworthy entities of artistic, historical or scientific value, we prefer to refer to the definition provided by the Venice Charter, which also includes those *'modest works of the past which have acquired significance with the passing of time'*. This definition indeed places greater emphasis on the character of transmission, preservation and heritage of past generations, of cultural and historic memory embodied by each monument, group of buildings or sites.

This perception of Cultural Heritage as a witness of the past, after all, is the same that has led to revise the definition given by the UNESCO Commission, including, in 2003, the intangible heritage of oral traditions, performing arts, local knowledge, and traditional skills alongside the physical artifacts that are part of the tangible Cultural Heritage (UNESCO, 2003a).

'The practices, representations, expressions, knowledge, skills – as well as the instruments, objects, artefacts and cultural spaces associated therewith – that communities, groups and, in some cases, individuals recognize as part of their Cultural Heritage.' (UNESCO, 2003a)

In general, the intent to provide an unambiguous definition of such terms demonstrated the development of interest, by supranational institutions and European states, in the issues related to protection and preservation of Cultural Heritage. Even at the national level, the efforts towards these themes have been translated into the will to build a solid legislative apparatus to regulate historic monuments and sites.

At the Italian level, cultural assets are disciplined by the *D. Lgs. 42/2004, Codice dei Beni Culturali e del Paesaggio*, which establishes rules and functions for their protection and enhancement.

At the French level, the *Code du Patrimoine*, promulgated in 2004, brings together the existing provisions and texts of French law concerning heritage and certain cultural services.

2.2. Digital documentation as a requirement for heritage preservation

The concept of documentation of cultural heritage as an essential prerogative of any preservation effort has gained ground especially in the face of the need to safeguard cultural and artistic assets from the disappearance or alteration due to man-made or natural threats. One thinks, for example, of the debate that arose around the 19th century on the restoration of the French heritage by Eugène Viollet-Le-Duc, or again, looking at our more recent history, of the question raised around the reconstruction of most of the roofing system of the Notre-Dame Cathedral in Paris after the fire of April 2019.

For heritage objects, documentation represents a necessary tool for acquiring and disseminating knowledge, for promoting public interest and involvement, and for ensuring the protection and preservation of works overtime. In response to events that have endangered or destroyed the entirety of historic buildings, a greater critical awareness of the heritage documentation process has maturated. It is the same awareness that moved, at the beginning of the 20th century, the need to transform theoretical studies on the restoration and documentation of the built heritage into a body of more defined, legitimately and universally valid norms and principles. On this line, various Charters of Restoration were drafted, among which it is worth noting the Athens Charter (1931), written on the International Conference of Architects, and the Italian Charter of Restoration (1932), the first official guideline of the Italian State on the subject of restoration, issued by the Superior Council for Antiquities and Fine Arts. After the Second World War, with its cycle of physical destruction of the architectural heritage, the interest in the definition of universal practices for the restoration of monuments increased.

This renovated concern led to the drawing up, in 1964, of the already mentioned Venice Charter, which today represents a universal reference text for the protection and safeguard of historical monuments. The final article of this text, Article 16, states the importance of the documentation of heritage assets:

'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. Every stage of the work of clearing, consolidation, rearrangement and integration, as well as technical and formal features identified during the course of the work, should be included. This record should be placed in the archives of a public institution and made available to research workers. It is recommended that the report should be published.'

(ICOMOS, 1964. Article 16).

At the end of the 1960s, therefore, the document affirmed the importance of bringing together and sharing, in a systematic way, the various documents (photos, drawings, analytical results) produced in the course of multidisciplinary studies carried out on a historical monument, in order to avoid its destruction and ensure the conservation of the cultural, historical and artistic heritage it entails.

In the last century, and with the rise of new information and communication technologies, the issue of reorganizing the different information shifted to the digital sphere: it was in this sense that, in 2003, the UNESCO Commission promoted the Charter on the Preservation of Digital Heritage. The document focused on the protection of digital assets, the latter understood as:

'Cultural, educational, scientific and administrative resources, as well as technical, legal, medical and other kinds of information created digitally, or converted into digital form from existing analogue resources.'

(UNESCO, 2003b. Article 1)

Emphasis is also placed on how this heritage is at risk of *being lost to posterity* (Article 3), as is the case of any other heritage asset. With the same concern, the recommendation of the European Commission, published in 2011 and concerning the digitization and online accessibility of cultural material and digital preservation, endorsed the digital library *Europeana*, providing direct access to over 19 million digitized objects, so including as a key issue of the Digital Agenda:

'The preservation of print documents (books, journals and newspapers), photographs, museum objects, archival documents, sound and audiovisual material, digital monuments and archaeological sites.'

(EU Commission, 2011)

Along the same lines, in 2014, the Council of the European Union encouraged the member states to foster:

'Funding, development and dissemination of digital cultural […] and promote public access to those digital heritage resources and services, including via Europeana'

and placed the definition of digital heritage alongside those of tangible and intangible heritage.

The report published in 2019 by the European Commission on the advancements since this 2011 Recommendation, moves beyond the aspects related to the preservation of museum and archival material, or audiovisual documents, and rather focuses mainly on the importance of the preservation -even digital, of monuments and archaeological sites, as particularly vulnerable to threats. This is a renewed awareness accentuated by the event of the fire of the Notre-Dame Cathedral, occurred that same year, which focused attention again on the threats to

which this heritage is subjected and to the importance of implementing funding programs for the digitization of immovable cultural heritage.

On an international level, the International Council of Monuments and Sites (ICOMOS), the UNESCO and other supranational organizations such as the International Center for the Study of the Preservation and Restoration of Cultural Property (ICCROM) have pursued their activities for the protection and conservation of the world's monuments and sites, issuing over time charters, guidelines and declarations that emphasize the importance of documenting the built heritage, starting from the principles expressed in the Venice Charter and even considering the preservation of digital objects.

Many international projects under way, endorsed by the UNESCO Culture Sector, are aimed at promoting the preservation and protection of existing heritage. Through the International Fund for the Promotion of Culture (IFPC), the Joint Programming Initiative on Cultural Heritage and Global Change (JPI CH), the Intangible Cultural Heritage Fund, the World Heritage Fund, and the Heritage Emergency Fund, e.g., the EU supports planning of an innovative, multidisciplinary nature, encouraging the sharing and appreciation of the cultural heritage of member countries, raising awareness of common history and values. Programmatic lines of most of these projects are, nowadays, the digitization of Cultural Heritage and the use of virtual environments where to store, document, access and share heritage-related data. These themes, intimately linked to the so-called digital continuity concept, are further analyzed in Section 2.4.

2.3. Architectural survey as a means of knowledge

As underlined in Section 2.1, the notion of Cultural Heritage includes both tangible (e.g., buildings, monuments, landscapes, books, works of art, and artifacts) and intangible assets (e.g., oral traditions, language, and knowledge) that contain a value to be transmitted, as the legacy of a culture or a place. In the case of tangible heritage, and in particular of architectural buildings, an essential requirement for the transmission of knowledge is the analysis of the state of the art and of the actual or past conditions. This is intended, e.g., in terms of restitution of the shape, the geometry, the appearance of a heritage asset:

'Architectural objects have a shape (spatial extension), they can be described by structures (entity collections and part-whole relation), they have quality attributes (colors, textures, terms, etc.) and they always have an interaction with time; finally various kinds of sources can describe their aspects' (De Luca, 2011).

At the base of the analyses aimed at the scientific documentation of the current state of a building or a historical site lies the concept of surveying. The survey is intended as a complex operation that includes both the measurement of a building (or of one of its parts) and its graphic representation: it is, therefore, at the same time, analysis, selection, synthesis, graphitization and codification of the built reality.

According to Saint-Aubin (1992), the architectural surveying process can be explained as a declination of three different terms: description (*décrire*), visualization (*figurer*) and representation (*représenter*).

The description is intended as the procedure of choice and selection of significant information, the abstraction process through which an architectural object is segmented, and its constituent components are identified.

The visualization process condenses the building into a graphic drawing composed of an ordered assembly of simple or simplified volumes, polygons and then lines delimiting it; a hierarchization eliminates the elements judged to be accessory and selects relevant geometric and visual descriptors. This also implies the translation of codes and symbols in a codified language and shape grammar.

The third process then involves the choices on the level of detail, on the type and mode of representation, on the density of information, for the elaboration of a complete, intelligible graphic drawing.

The term surveying thus indicates simultaneously the operation of restitution of the architectural work through the description of its

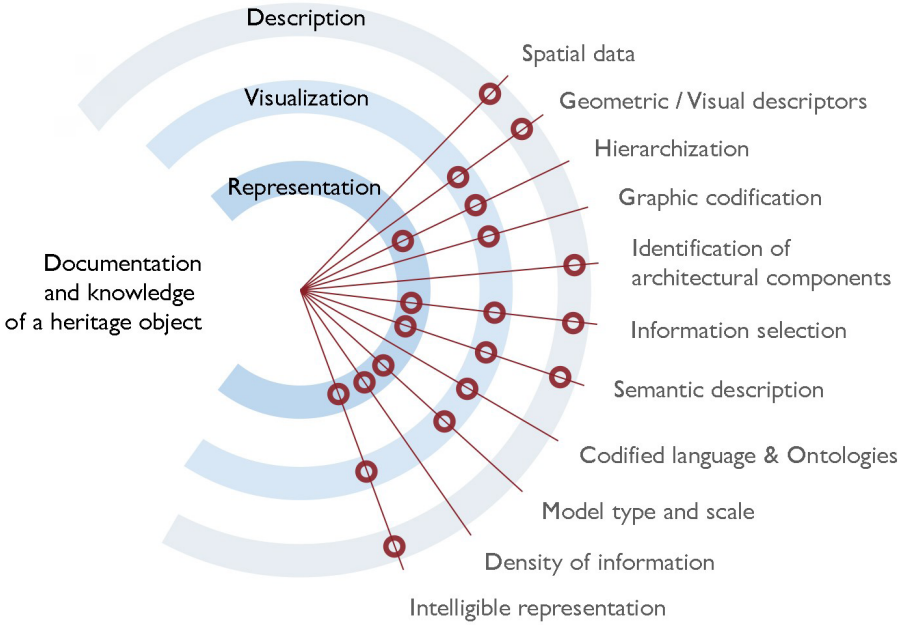


Figure 2-1.

Knowledge and documentation of heritage artifact as processes combing description, representation and visualization.

shapes, attributes, interactions with time and aspects, and the processes of analysis and graphic decomposition that leads one to classify and organize, through description and visualization, the different available information on the built object (Figure 2-1).

As later expressed by (Docci and Chiavoni, 2017), the different layers of graphic expression thus correspond to several and subsequent levels of comprehension of the heritage artifact:

'Used in a structuralist key, drawing becomes a real critical means, with which it is possible to analyze the work through its segmentation and the representation of its constituent elements. [...] The graphic medium is used to create models of knowledge on the architectural object or urban reality, in order to be able to understand and analyze them, just as a scientist would do, artificially reproducing a certain phenomenon in his laboratory.'

(Docci and Chiavoni, 2017)

2.3.1. Traditional / Digital survey

The traditional approach to the surveying of architectural buildings involves the use of direct measuring instruments (direct survey), to restore the dimensions of an architectural organism starting from the in-situ (on-site) measurements of the object. The restitution of the visual appearance, in this way, is performed by empirical decomposition and subsequent graphical transcription, through drawing, of the elements that make up the architectural organism, as in the example of Figure 2-2. The survey is carried out using measuring instruments that are stretched along the surfaces of the objects to be reproduced or in their vicinity, to measure lengths, to set alignments and to note on a graphic

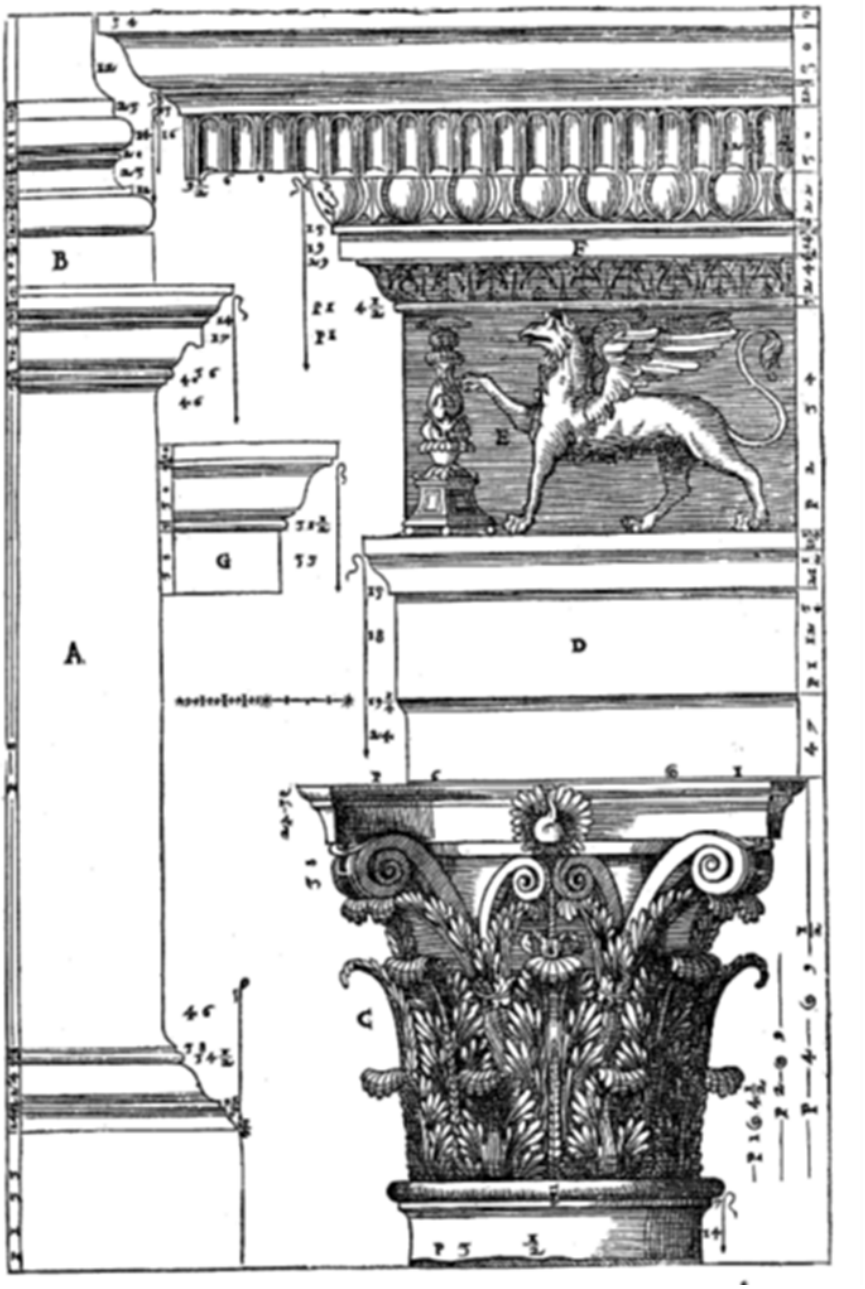


Figure 2-2.

Tempio di Antonino e Faustina in Rome, architectural detail by Andrea Palladio. The drawing is supported by quotation marks and annotations (Tav. XXI, Palladio).

support the extents of significant elements.

The developments witnessed in recent years in the field of design, computer science and metrology, have made it possible to move from direct measurements of reality to the indirect surveying, based on the use of active or passive sensors (Pierrot-Deseilligny et al., 2011) that allow to reconstruct a digital model of reality. Laser scanning and photogrammetry are among the most widely used techniques.

The transition from direct to indirect surveying was also marked, at a lexical level, by the gradual replacement of the term *topography*, which indicated the traditional science of surveying, with the term *geomatics*,

which places more emphasis on the close link between the disciplines of representation and computer sciences (Bevilacqua et al., 2018).

2.3.2. In-situ and ex-situ survey

Technologies as laser scanning and photogrammetry, if compared to traditional surveying devices, offer a higher level of automation in the data acquisition phase, as they allow to capture a massive amount of 2D and 3D data, in a short time and with less invasive methods that do not require direct contact with the object of study (Bevilacqua et al., 2018; Remondino and Rizzi, 2010). Moreover, the use of the third dimension in the restitution of the digital replica has the remarkable advantage of better describing and representing the morphological complexity of each architectural artifact, e.g., in terms of geometric and visual attributes. The 3D virtual representation, compared to the bi-dimensional restitution, allows in fact to depict an artifact in its entirety, capturing a complete situation at a specific time and enabling general views over the object. Therein, any detected information can be easily viewed and retrieved.

Despite the benefit of such techniques in the data acquisition phase, however, significant problems arise in handling and processing the massive amount of 2D and 3D data derived from reality-based surveying. This constraint effectively shifts the focus from the survey acquisition phase to the data processing one, thus significantly changing the approach of scholars towards the study of a heritage object. Abergel

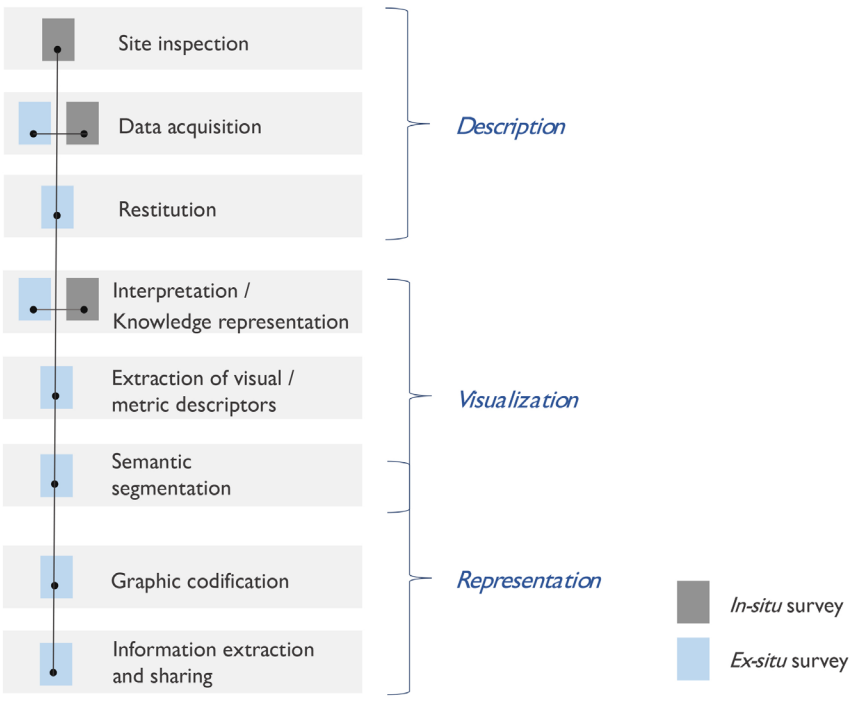
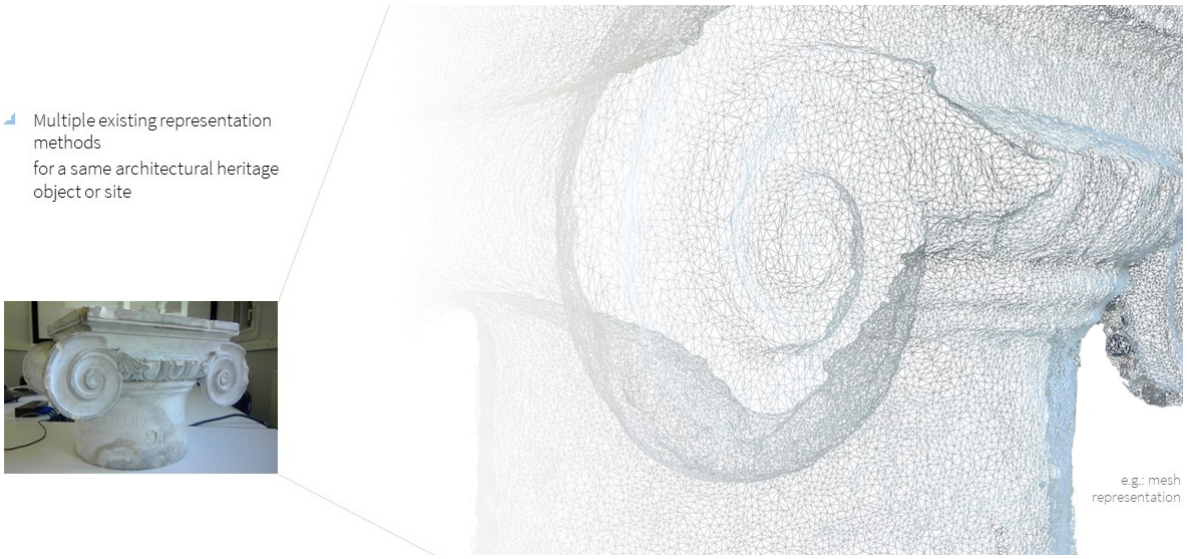


Figure 2-3.

Digital survey: in situ and ex-situ activities.

Adapted from (Abergel, 2020).

(2020) call it a paradigm shift that displaces the most important moment of observation and interpretation of data: from the *in-situ* inspection and survey required by traditional methods to an *ex-situ* survey. The analytical step, involving information selection and segmentation, thus moves from the real to the digital environment (Figure 2-3).



2.4. The digital continuity concept: visualizing and ordering complexity of heritage artefacts

The transition from traditional to digital survey instrumentation has greatly streamlined the collection of cultural heritage data. At the same time, however, it has resulted in a massive availability of digital heritage models, with a consequent multiplicity of representation types and/or surveying outputs.

This huge amount of digital heritage data often causes data dispersion, alteration or loss, so complicating the production of consolidated knowledge. Despite the acceleration of the acquisition phase and the automation of data capturing devices, indeed, methods and tools enabling a streamlined workflow of information handling and structuring, as well as of rational management and semantic enrichment of digital models, is still lacking.

In order to make the scientific community aware of the risk of dispersion of such information, the concept of *digital continuity,* originally proposed in the industrial context to guide the product life cycle management (MacLean and Davis, 1999), can be transposed to the domain of cultural heritage, stressing the need to keep digital heritage data usable, complete and available over time. The UNESCO Commission (2003) affirmed that:

'Continuity of the digital heritage is fundamental. To preserve digital heritage, measures will need to be taken throughout the digital information life cycle, from creation to access. Lovng-term preservation of digital heritage begins with the design of reliable systems and procedures

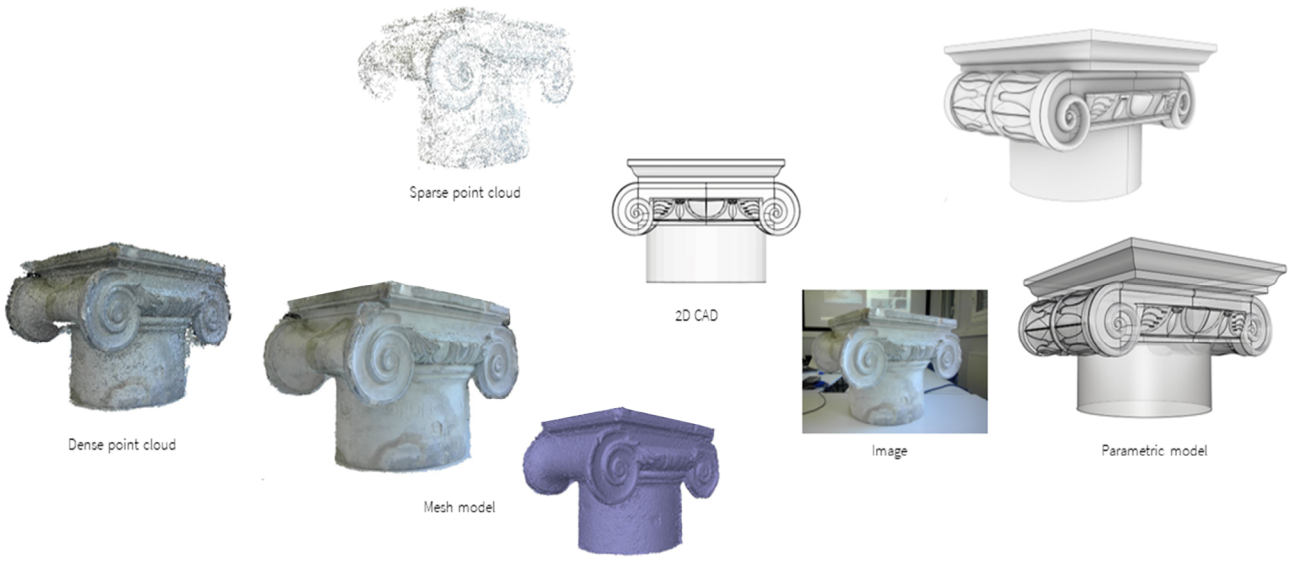


Figure 2-4.

Several types of existing digital representation systems.

which will produce authentic and stable digital objects'.
(UNESCO 2003b, Article 5)

The quality and exchange of digital cultural heritage information is crucial to ensure multidisciplinary knowledge sharing. Digital information that is scarce, missing, or simply not well interlinked and coordinated can be decisive and result in the loss of the wealth of knowledge, historical studies, and temporal states concerning a certain cultural resource.

The same principles were further emphasized in the *London Charter for the Computer-based visualization of Cultural Heritage* (2006) and in the subsequent Seville Principles (2011); these documents, besides demonstrating the potential of computer-based heritage visualization, focused on its many weaknesses and inconsistencies, stressing the need to ensure completeness, availability and permanent access to digital heritage data, both in terms of common vocabulary and representation methods.

Since then, several international projects have raised the awareness on the importance of digital continuity, particularly concentrating on the first aspect, relating to the definition of a common lexicon: the International Organization for Standardization (ISO), a worldwide federation of national standards bodies, proposed a reference ontology for the interchange of cultural heritage information, the BS ISO 21127:2014. Further on, the European Standard EN 15898:2019, approved by the European Committee for Standardization (CEN) in 2019, provided the definition of the main general terms to be used in the cultural heritage conservation context.

In the same year, 27 European states signed up the *Declaration of cooperation on advancing the digitization of Cultural Heritage,* a common initiative to highlight the potential of 3D digitization processes in the preservation and enhancement of Cultural Heritage. Several projects on digital Cultural Heritage have been funded by the EU since then, of which it is worth mentioning INCEPTION (Inclusive Cultural Heritage in Europe through 3D semantic modeling), GRAVITATE (Geometric reconstruction and novel semantic reunification of cultural heritage objects), and Time Machine (Mapping 2000 years of European History).

Despite these international efforts, little has been made on the second aspect, which concerns the definition of common representation methods. The multitude of representation methods and systems that exist today (to give an example, Figure 2-4) and the lack of unambiguous systems for entering, managing and structuring annotations within 3D models still complicate the creation of complex information structures and thus the achievement of the sought-after digital continuity.

2.5. Heritage- or Historic- Building Information Modeling

The intention to associate the purely geometric representation of a digital model with textual, thematic and descriptive information has led to the emergence of Building Information Modeling (BIM) techniques. Born in the Architecture, Engineering and Construction field (AEC) and originally applied to the construction of new buildings, these techniques underwent their most important developments around 2002-2003 (Abdal Noor and Yi, 2018). The logic behind them is the creation of information systems enabling to manage, in an integrated way, the different knowledge-related data (drawings, analytical spreadsheets, images, documents) concerning a certain building and in the different phases of its lifecycle. The BIM technique has been later extended to the existing heritage, starting from the studies by Murphy (Murphy et al., 2009) and his student Dore (Dore and Murphy, 2013) at the Dublin Institute of Technology.

Despite the increasing diffusion of BIM procedures, also in application to historic buildings (Figure 2-5), there is today a great confusion about the different terms and definitions, and it is therefore necessary to clarify many of them before moving on. First of all, in general the term BIM is used to define both a model and a process; the letter M in the acronym, indeed, may be referred either to the terms Model, Modeling,

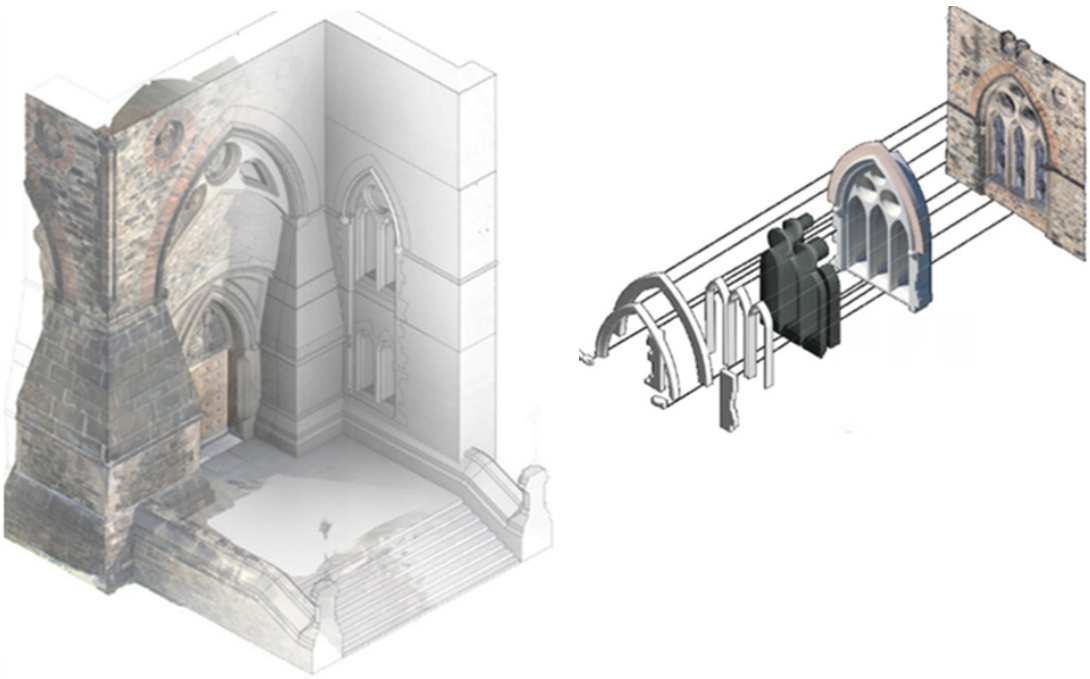


Figure 2-5.

H-BIM reconstruction based on 3D survey data (Murphy, 2009).

Management (Bruno et al., 2019). In detail, one can consider:

- BIM as a process. In the first meaning, BIM is intended as the set of actions that are carried out on the digital replica during the whole life cycle of a building, in order to collect information produced over time (e.g., related to study, analysis, construction, maintenance,…) and condense them within the 3D representation (Pocobelli et al., 2018a; Hichri et al., 2013). As such, the geometric conceptual model could be connected to descriptive multi-source information, graphic attributes and parametric rules.
- BIM as a model. This second meaning relates more to the model, intended as the mere virtual replica of an existing artefact, considered in its physical and functional characteristics (Arayici, 2008). The real object appears decomposed in its architectural and semantic elements, that are virtually reconstructed (Bruno et al., 2019; López et al., 2018; Murphy et al., 2013). The reconstruction is based on smart objects, i.e. objects that are defined by numerical parameters, such as dimensions or materials, and that are embedded with other kinds of information, such as building materials and properties (Pocobelli et al., 2018a; Baik, 2017).

The reconstruction process should entail the three phases of (Macher et al., 2017; Tang et al., 2010):

- Geometric modeling of the architectural components (e.g., the question 'what is the shape of this wall?')
- Object recognition and assignment of a building category (e.g., the statement 'This object is a brick wall.')
- Object-relationship modeling (e.g., the statement 'Wall 1 is connected to Wall 2 at this location').

When using the term BIM in this work it is generally referred to the collaborative process for the production and management of structured data (Bruno, 2017) rather than to the model per se. Indeed, this definition puts more emphasis on the perspective of ensuring a shared data access system, a common syntax and information organization. Moreover, with regard to the application to historical heritage, the acronym H-BIM encompasses at the same time the concepts of Historic-Building Information Modeling, Heritage-Building Information Modeling and BIM for heritage, which have been adopted almost interchangeably over time (Historic England, 2017). The term H-BIM is thus used generally to define the set-up of specific procedures aiming to create BIM-based databases for historical objects, buildings or sites. Closely related to this term, the expression Scan-to-BIM (Rocha et al., 2020; Bosché et al., 2015) is used to describe the construction of H-BIM

systems starting from 3D reality-based data; that is, the construction of parametric models done by interpreting and modelling the data acquired from surveying, such as point clouds captured via laser scanning or photogrammetry.

Chapter 3

Acquisition, processing and semantic enrichment of digital heritage objects

The emergence of the architectural surveying practice, as we conceive it today, is intimately linked to the studies and theoretical principles introduced during the Italian Renaissance. According to Saint-Aubin (1992), the fundamentals of surveying were theorized starting from the 15th century, a period in which two different paths of data collection were established, apparently conflicting but actually complementary, which have since then been applied to the study of existing architectural works: the first one, oriented to preserve the vision of the architectural object; the other one, oriented to preserve its dimensional and metric properties.

The first method relies on the projection of a real object on a perspective plane, and is more related to the invention of perspective, while the second one relies on the collection and restitution of orientations and distances. The two modalities both start, as if from a postulate, from the *visual ray* theory: in the first case, the projecting rays starting from the object intersect the perspective plane and define the elements that make up the scene on a two-dimensional support, as is the case today in photographic sensors. In the second, more abstract case, the object is recorded from a known position (station) and angles and distances are measured with respect to that position (De Luca, 2006).

The first modality, based on the perspective plane, evolved from the introduction of perspective by Filippo Brunelleschi and the subsequent theories outlined by Leon Battista Alberti in the treatises *De Pictura*, through Albrecht Dürer's *perspectographs* to the camera obscura, and on to analogue/digital photography and then photogrammetry.

To demonstrate the effectiveness of this type of representation,

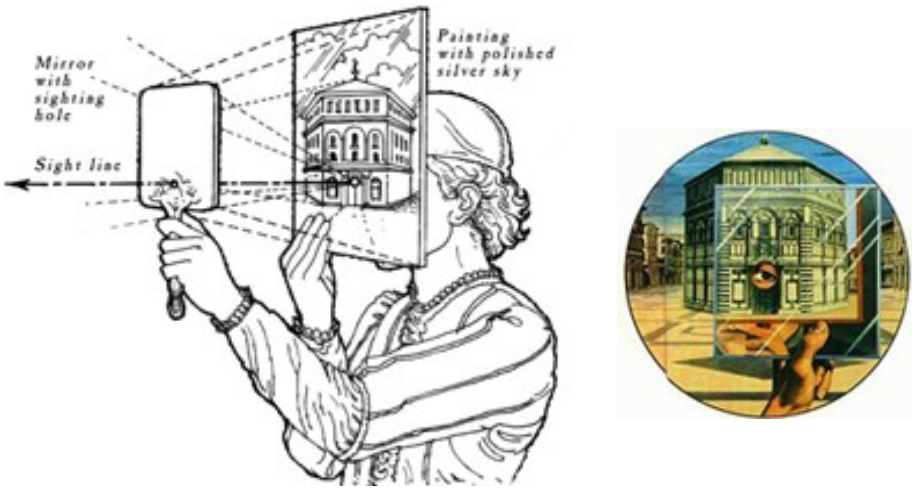


Figure 3-1.

Scheme of the Brunelleschi's method of perspective (Argan, Wittkower, 1990).

Brunelleschi created in 1415 the perspective panel named *tavoletta* for the Baptistery of Florence. He painted the Baptistery with its marble inlays on a square panel and, in order to demonstrate the verisimilitude of the painted image with the real one, he made a flared hole on the back of the panel, corresponding to the vanishing point of the scene.

The observer, by standing at the exact point where the image and the real object are superimposed (located inside the central door of the Cathedral) and looking into the hole with one eye, could perceive the real image of the scene. Then, with the help of a mirror supported by the observer's other hand and adjusted to a suitable distance, he could see the painted image reflected in the mirror and appreciate the perfect coincidence of the painted image with the real one. To make the demonstration of perspective even more effective, Brunelleschi covered the area of the panel corresponding to the sky with burnished silver so that the real sky would be reflected in it with its clouds, increasing the effect of depth in the third dimension of the illustrated architectural elements and of real/virtual hybridization (Figure 3-1).

Further developments of this theory are the perspective machines described by Dürer (1525), drawing machines that for the first time mechanically translated all the parameters of perspective construction (Figure 3-2, 3-3): the eye being a nail, the visual ray being a wire, and the picture being a plane described by the intersection of two wires within a frame (Camerota, 2002).

Brunelleschi's experiment, theorized by Leon Battista Alberti in his treatise *De Pictura*, as well as the subsequent experiments at the intersection of geometry and optics, marked the starting point for the birth of modern photography and photogrammetry, and in fact constituted the theoretical principles on which the image-based reconstruction techniques and tools still leverage today (Saint-Aubin, 1992).

On the other hand, the second mode, which is based on the definition

of a station point and the acquisition of angles and distances between this point and the real object, could be said to provide the principle on which measuring instruments such as the theodolite, the total station and, later on, the 3D laser scanner have been created.

In this context, the experiments on the restitution and documentation of ancient and classical architecture during the Renaissance period were significant: they brought to light the need to study proportions and relationships and to correctly restore the measurements of historical monuments on bi-dimensional supports, in order to enrich with metric



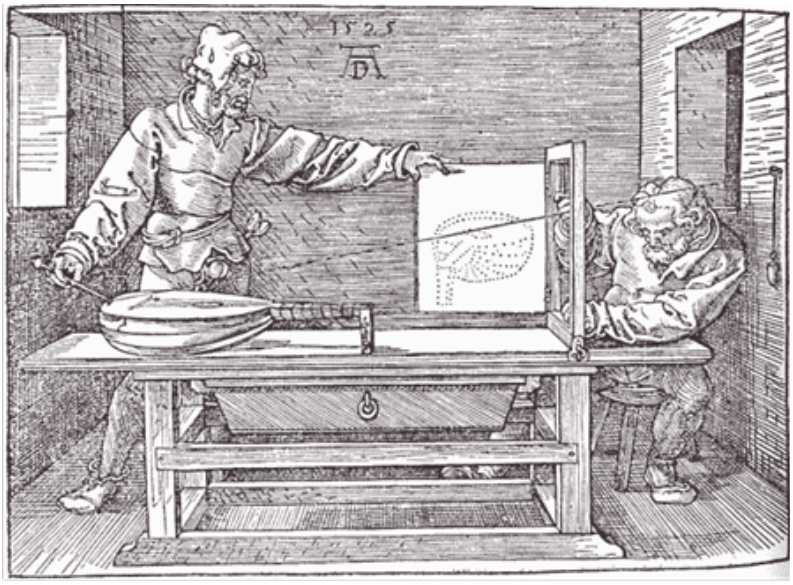


Figure 3-2, 3-3.

Albrecht Dürer's drawing devices: the *perspectographs* (Dürer, 1525)

rigor the figurative and typological repertoire indispensable for a design in the manner of the ancients (Docci and Chiavoni, 2017).

Fundamental is the research by the major painters, architects and engineers of the 15th-16th centuries into methods of measurement based on the recognition and restitution of points of interest, calculated by distance from a fixed point and reported on graphic instruments as drawing sheets and maps. (Saint-Aubin, 1992) gives, in this regard, the example of Leon Battista Alberti's cartographic map of the ancient Rome (Figure 3-4): the topography of the town is reconstructed by means of a system of coordinates allowing to graphically reconstruct the shape of the city, by detecting points of interest identified by polar coordinates and by knowing the position of the surveyor (Zerlenga, 2020). This procedure recalls the operating principles of currently used 3D laser scanners.

The spread of representation techniques based on geometry and optics and the implementation of surveying instruments led indeed to the emergence of the practice of architectural surveying for engraving and printing classical and military architecture, which became widespread from the 17th century onwards, with the principles of cavalier, military and parallel projection; at the end of the 18th century, the descriptive geometry was finally systematically introduced and codified thanks to the work by Gaspard Monge (Docci and Chiavoni, 2017).

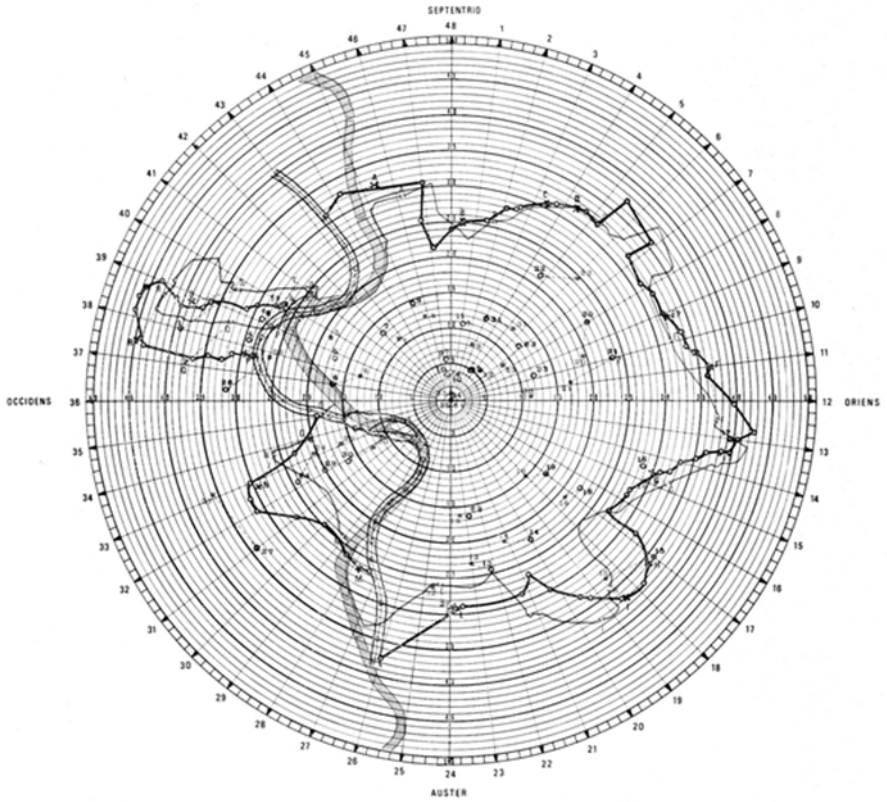


Figure 3-4.

Reconstruction of the urban walls of ancient Rome according to the method described by L. B. Alberti (Vegio, 2019).

However, while instruments and tools for surveying and representation have evolved over time since the revolution introduced in the Renaissance period, the theoretical principles behind the two methods described have remained fundamentally unchanged up to the present day. The developments of the last century in particular in the fields of informatics and computer vision have rather concerned the acquisition of greater accuracy and automation of surveying instruments and have enabled the digital storage of data.

Indeed, in Section 3.1 we illustrate laser scanning and photogrammetry as main techniques for 3D reality-based data collection, showing that the measurement of angles and distances still guides the bases of modern laser scanning principles (Paragraph 3.1.1), and the intersection between the visual rays and the perspective plane is still the core theory behind the construction of photo-cameras and behind the formulation of photogrammetric restitution methods (Paragraph 3.1.2).

The different outputs of existing 3D surveying techniques are then commented on in Section 3.2; existing platforms for the annotation-based semantic enrichment are illustrated in Section 3.3, highlighting, limitations, potentials and perspectives of existing tools enabling the graphical visualization and the semantic annotation of 2D/3D heritage objects are discussed.

3.1. 3D surveying techniques

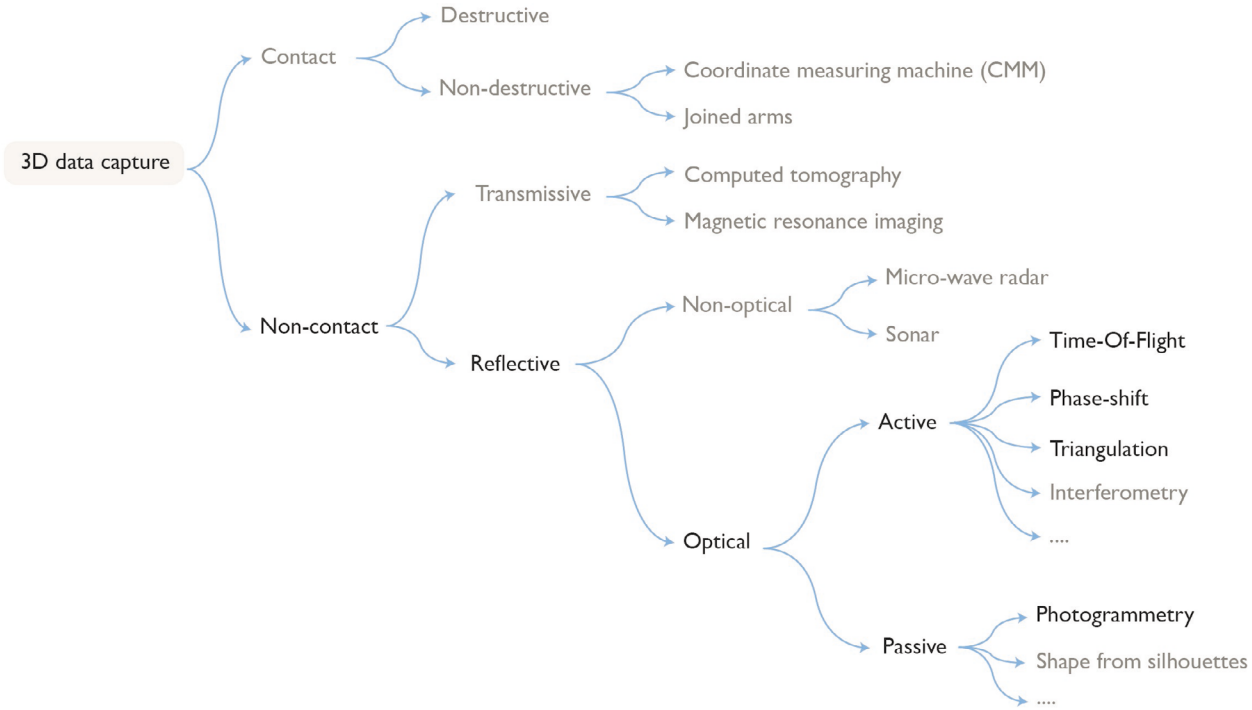


Figure 3 5.

Taxonomy of 3D data capture systems. Bold characters indicate the acquisition techniques discussed herein.

Adapted from (Luhmann et al., 2016; (Caroti et al., 2015; Grussenmeyer et al., 2008).

Remondino and Rizzi (2010) defined reality-based techniques as those methodologies that:

'[…] employ hardware or software to survey the reality as it is, documenting the actual or as-built situation of a site and reconstructing it from real data'.

Such methodologies can be distinguished first based on the direct or indirect contact with the study object, hence considering whether or not they fulfil the requirement that recording devices will not potentially damage the artefact (3D icons, 2014). Among these, laser scanner and photogrammetry, which are the most widely applied technologies in the case of cultural heritage survey, rely on non-invasive sensors since they do not require direct contact with the heritage artifact.

For this case, whether the radiation is only received or even emitted by the instrumentation, non-contact sensors are distinguished in passive or active sensors, respectively. Several hints to this type of categorization are found in literature (Luhmann et al., 2016; Caroti et al., 2015; Grussenmeyer et al., 2008) and allow to subdivide restitution techniques according to the scheme provided in (Figure 3-5).

Active sensors. Active sensors use their own light, or radiation, to produce the scan: techniques relying on such systems are capable of transmitting a signal to the object to be surveyed, in the form of electromagnetic

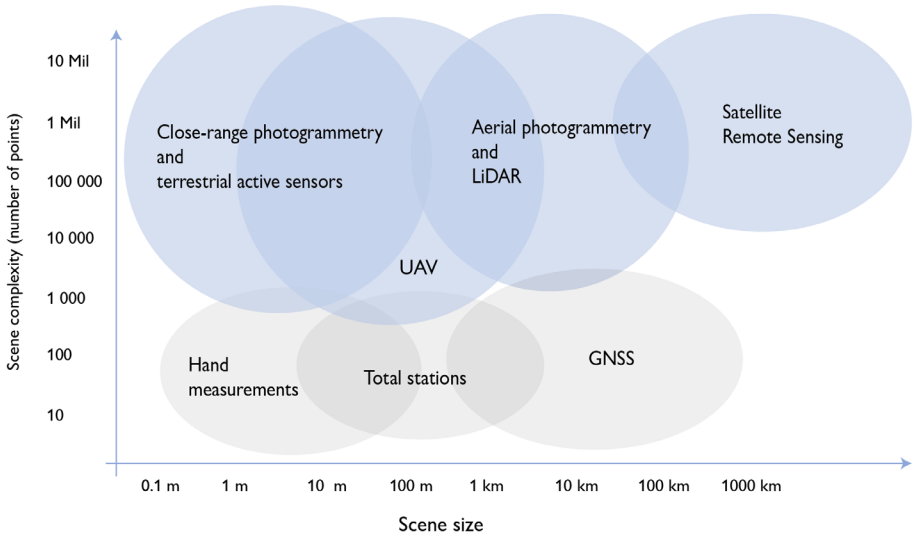


Figure 3-6.

Choice on the type of 3D data capture technique based on object/scene dimensions and complexity of the reconstructed digital model.

Adapted from (Georgopoulos and Stvathopoulou, 2016; Remondino and Campana, 2014).

radiation in either the visible or invisible spectrum (Weinmann, 2016a). A scanner device emitting the signal directly provides, from respective observations of an existing scene, the range data that is necessary to record metric information of the environment (Blais, 2004). Acquisition devices such as pulsed, phase-shift, triangulation-based laser scanners exploit such range-based techniques, as they directly capture distances, thus 3D information (Remondino and Rizzi, 2010). The metric data is acquired directly by on-site surveying (Vosselman and Maas, 2010).

Passive sensors. Unlike active sensors that emit a proper specific signal, passive sensors capture the natural environmental radiation coming from other light sources: they exclusively exploit visible light, or radiation, to capture data, and are said to be passive because they require an external source. Image-based modeling techniques are derived from the use of such sensors: photography is used in this case as a support medium to capture the real conditions of an object, and only then are its spatial coordinates extracted through photogrammetric processing. Image data, indeed, require to formulate mathematical expressions to turn the 2D image measurements into 3D information (Remondino and Rizzi, 2010). Among passive surveying techniques, photogrammetry is doubtlessly the most widely employed to date.

The choice on the use of the most suitable technique may depend on the size of the object to be surveyed and on the required accuracy: Luhmann et al. (2016) and Remondino and Campana (2014) discussed the dependency between selection of measurement techniques and complexity of the reconstructed digital model. Figure 3-6 illustrates this relationship for different existing measurement techniques. In the architectural heritage domain, considering the size of the surveyed objects and the degree of accuracy or visual and metric description required, the preferred techniques for 3D acquisition are laser scanning

(active technique) and photogrammetry (passive technique). As further explained in Section 3.1.3, however, one technique does not exclude the other, rather in many cases an integration between these two, alongside traditional hand measurements and total stations, is necessary to achieve a complete description of geometries and shapes of an architectural artifact.

3.1.1. Laser scanning

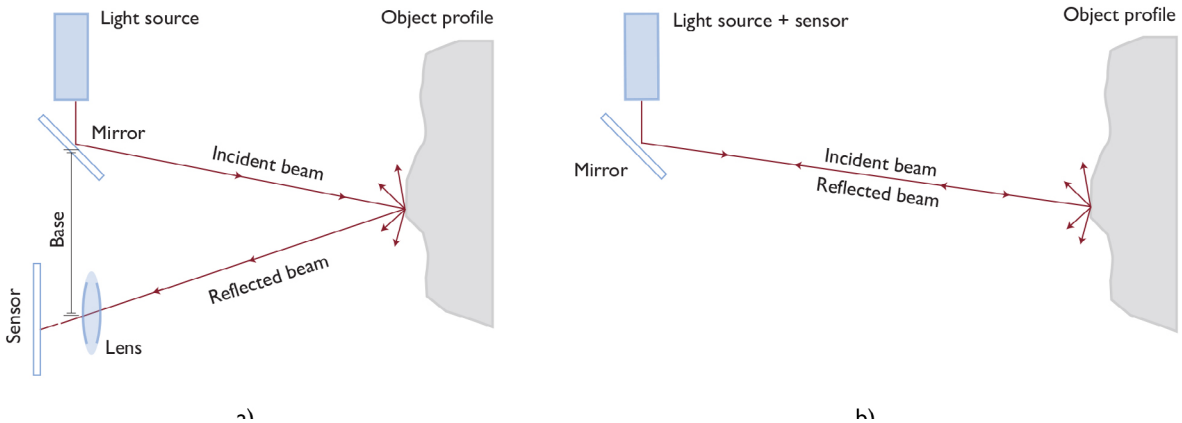


Figure 3-7.

Laser scanner operating principles: triangulation (a) and range scanners (b).

Adapted from (Guidi and Remondino, 2012).

Laser scanning techniques record 3D coordinates by analysis of the reflected light of a laser beam projected onto the objects' surfaces. The surveying instruments include a unit for the emission of the laser beam -with movable mirrors to steer the laser beam- and optical sensors to collect the diffused return signal. The laser scanner provides the measurement of two angles (zenith and azimuth) and a range, as the position of points hit by the laser beam is determined via spherical coordinates that are later converted to cartesian coordinates (Bevilacqua et al., 2018).

According to the working principles used to return the points coordinates, laser scanners can be classified respectively in range scanners (phase-shift and Time-of-Flight, TOF) and triangulation scanners (Figure 3-7). A comparison between the different methods in terms of range of applications and grade of accuracy has been presented in (Guidi and Remondino, 2012; Boehler et al., 2003).

- Triangulation-based laser scanners.

These systems rely on spatially distinct emitters and receivers, which form the base of a triangle with the third vertex on the survey object. The length of one side of the triangle and the distance between the camera

and the laser emitter are known. From the identified configuration parameters of the device, the shape and size of the triangle are fully determined and give the location of the scanned point. Triangulation-based laser scanners are generally adopted when reconstituting the shape of small and medium-sized objects (to be returned at a scale of 1:1–1:10), e.g., statues, capitals, moldings.

The measurement range is in fact very low, as the computed distances become less precise at the increasing distance from the object; however, for short-range applications, such techniques attain a high level of accuracy.

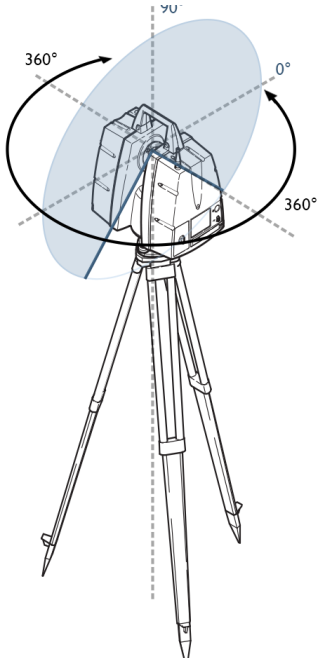


Figure 3-8.
Field of view of the laser scanner.

- Phase-shift laser scanners.

The range is calculated by measuring the phase difference of the modulated signal between emission and reception. If compared to TOF laser scanners, these instruments enable a much higher precision, but they cannot be used to cover long distances between laser scanner and object to be surveyed; that is why they are generally applied in middle-range applications requiring heritage documentation at a smaller scale.

In all cases, the output of the laser scanning survey is influenced by the characteristics of the instrument itself (calibration, method used,…), of the surveyed object (reflection, color, light diffusion and absorption), of the working environment and of the backscattered echo-signal (Remondino and Rizzi, 2010; Beraldin et al., 2007).

For architectural heritage applications requiring the restitution of buildings or monuments, TOF laser scanners are preferred. When performing the acquisition, the instrument is placed at specific locations, i.e., scan stations. For each scan station, the instrument produces a survey of the objects that are in direct view and within the distance of the laser rangefinder, considering a 360° angle of view in the horizontal plane and a slightly smaller angle in the vertical plane (Figure 3-8). A sequence of laser beams is arranged in a dense grid, returning the 3-D coordinates of points that are hit by the laser beams at a given scan position; the result is a 3D dense point cloud (Figure 3-9).

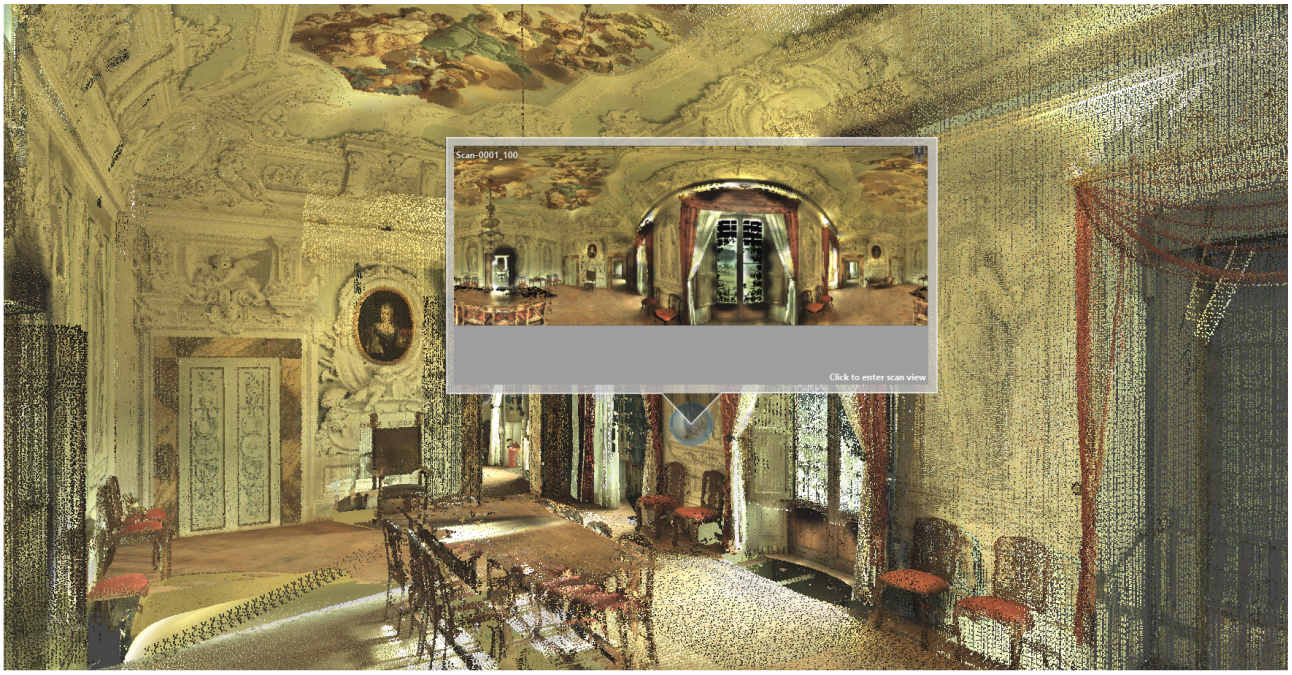


Figure 3-9 (*above*).

Panoramic image acquired via a laser scanner, e.g.: Grand-Ducal apartments of Pisa Charterhouse.

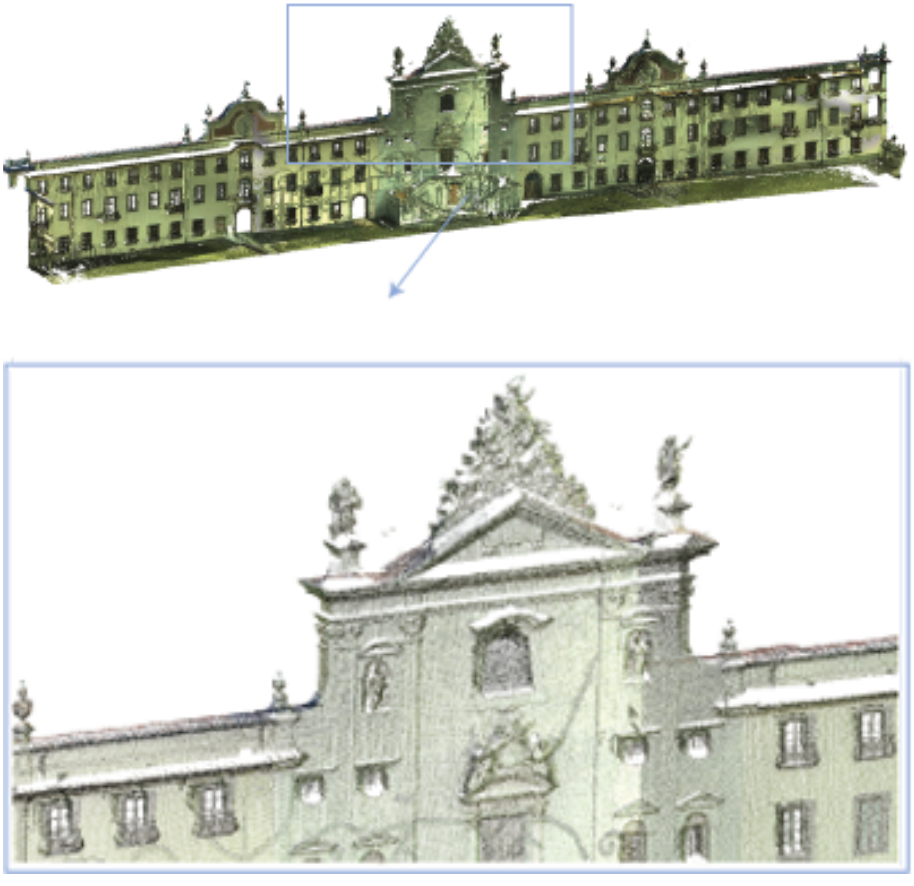


Figure 3-10 (*to the left*).

Dense point cloud obtained via laser scanning. E.g.: Pisa Charterhouse.

In general, the point coordinates can be returned in association with chromatic data in black and white that describe the intensity of the returned laser pulse, that is a function of the power of the received backscattered echo signal (Tan and Cheng, 2020; Xu et al., 2017; Kashani et al., 2015). This information can vary on the basis of the laser scanner system, of the weather and light conditions and of the physical characteristics of the reflective surface (Bolkas, 2019). Moreover, many instruments are generally equipped with a camera that acquires images so as to detect for each point a colorimetric value, in the R, G, B (red, green and blue) space, thus resulting in a colorized point cloud (Figure 3-10). However, as highlighted by (Bevilacqua et al., 2018; Caroti et al., 2015), such cameras do not provide high-quality photography if compared to current image-based acquisition techniques.

3.1.2. Photogrammetry

The etymology of the term photogrammetry is related to the composition of the Greek words photo (φῶς-φωτός, meaning "light"), graphia (γραφία, drawn) and metron (μέτρον, measuring), so the term literally signifies '*a metric representation with light*'. The expression indeed stands for the all analytical, graphic and optical-mechanical procedures through which, starting from a suitable number of photographs of an object, appropriately taken, it is possible to reconstruct the object (or certain projections of it) in digital form. The American Society of Photogrammetry and Remote Sensing provides the following definition of the term:

'The art, science and technology of obtaining reliable information about physical objects and the environment through the process of recording, measuring and interpreting photographic images, or patterns of electromagnetic radiation and other phenomena'

(American Society of Photogrammetry, 1980)

The reconstruction of an object starting from images requires analyzing the optical process by which an image is created. The study of the mathematical relationships and projections between 2D images and 3D scene and vice-versa allow to derive the configuration and position of the cameras at the time of acquisition. The central projection imaging thus constitutes the basic mathematical model based on which it is possible to derive (Kraus, 1993):

- Numbers, i.e., coordinates of object points in a 3D coordinate system;
- Drawings, e.g., maps with planimetric details, contour lines or graphic representations of objects;
- Images: digital images, rectified photographs (ortho-photos).

The development of the branches of photogrammetry has been influenced by the advancements and innovations of the related technologies employed: photography, informatics and electronics. Over time, the discipline has undergone the phases of plane table photogrammetry, analog photogrammetry, analytical photogrammetry, to finally enter the phase of digital photogrammetry (Bezoari and Selvini, 1999).

Since the 90s, the shift to digital and no longer analog images, together with the advances of computer vision, revolutionized and multiplied the use and applications of photogrammetry, so broadening the technique to the heritage domain. Photogrammetry thus stems from the principles of image capturing, and currently leverages digital photographs for the reconstruction process.

3.1.2.1. The digital image

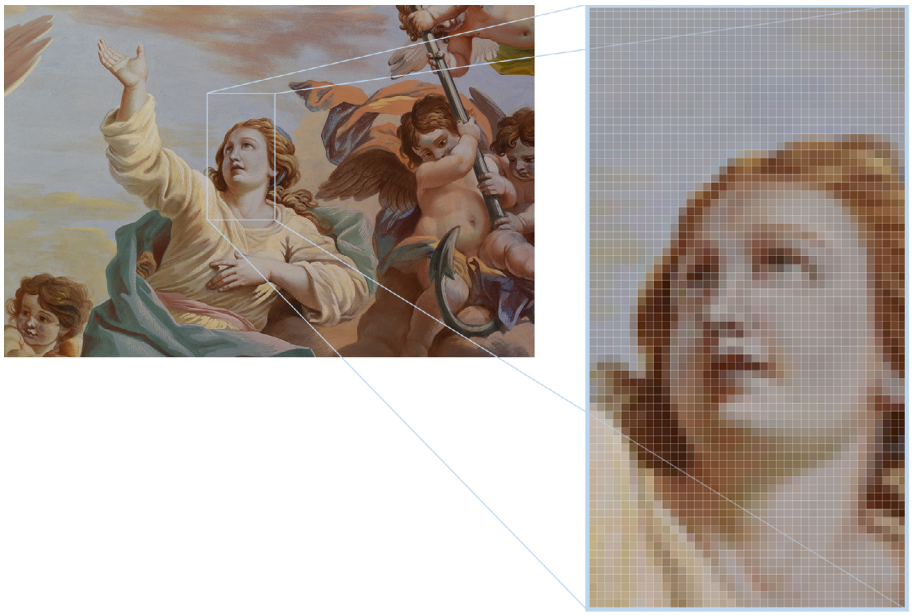


Figure 3-11.

The digital image as a set of digital values, called pixels.

E.g.: Grand-Ducal apartments, Pisa Charterhouse.

In digital images, a camera uses a sensor to record the incoming pattern of light. While analogic cameras captured light onto photographic plate or photographic film, digital cameras use an electronic image sensor: when hit by the incoming light, the sensor returns the received electromagnetic radiation as an electric signal. The light detector is typically a charge-coupled device (CCD) or an active pixel sensor (CMOS), the latter commonly found in modern reflex cameras.

The intensity of light is captured by photosensitive cells that are paired with color filters, which alternately let the red, blue and green color components pass through: each component of the sensor, which constitutes the single pixel (Figure 3-11, Figure 3-12), records the light intensity and converts the light into an image.

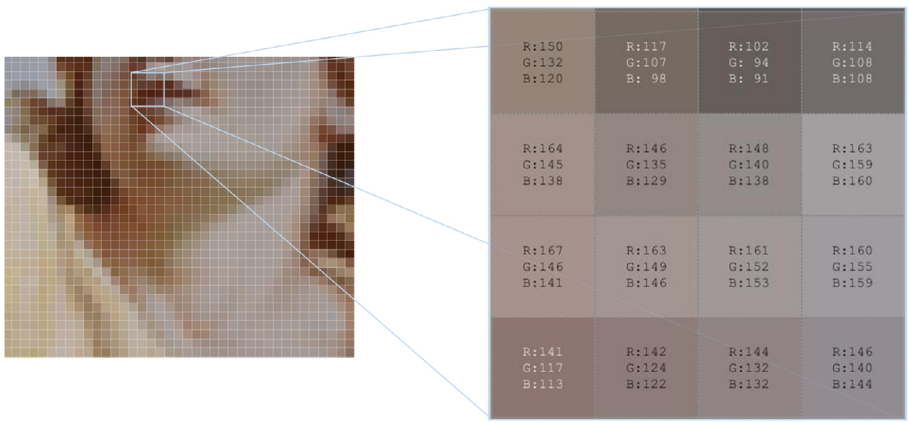


Figure 3-12.

Each pixel pertains color information, e.g. in the form of RGB (red, green, blue) values.

E.g.: Grand-Ducal apartments, Pisa Charterhouse.

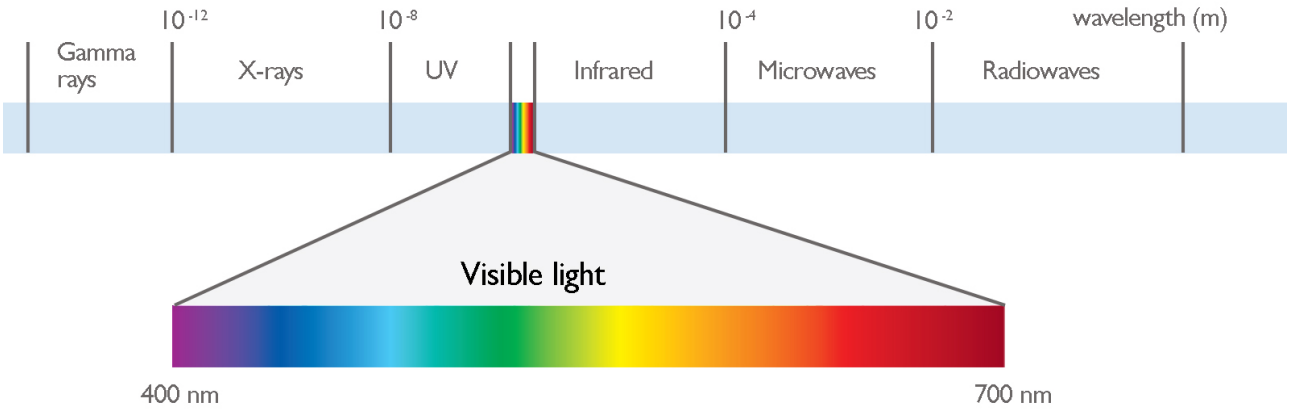


Figure 3-13. ELectromagnetic spectrum.

Depending on the filters used, the sensors can capture specific portions of the electromagnetic spectrum (Figure 3-14) that can also go beyond the visible light; hence, the techniques of multi-spectral imaging. The latter span the wavelength range, exploiting also rays that cannot be seen by the naked eye and that belong to the category of ultraviolet (UV, 200-400 nm) and infrared (IR 760-1700 nm) rays (Marengo et al., 2011).

Thermal infrared information is often considered in the analysis of historical buildings, to reveal further details on state of conservation, padding, older layers, back structure of frescoes.

Near IR is applied to reveal *sinopie* (pentimenti) and preparatory drawings of paintings, while UV rays are used for the detection of varnishes and over-paintings (Vallet et al., 2012; Remondino and Rizzi, 2010; Odile Guillon, ©CICRP, 2010. Figure 3-13, Figure 3-14).

Figure 3-14.

Observation, in semi grazing light (*a*), of the response emitted in the IR range (*b*) and of the X-ray fluorescence under UV radiation (*c*).

e.g.: Frescoes of the Charterhouse of Villeneuve-les-Avignon (Odile Guillon, © CICRP, 2010).



a) b) c)

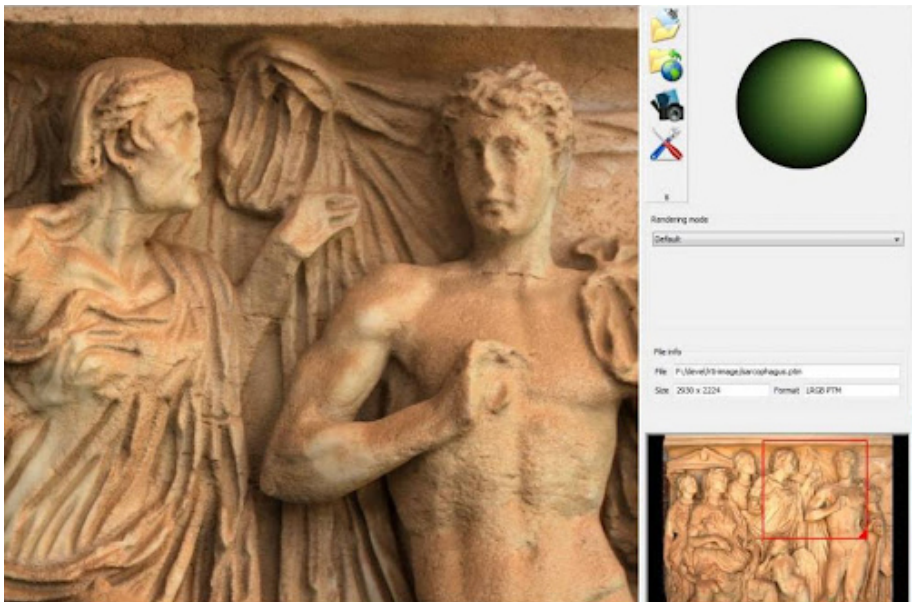


Figure 3-15.

RTI web viewer. The control panel to the right allows to edit the lighting conditions, to improve the perception of the details (Palma et al., 2010).

It is also worth noting, in this section, the Reflectance Transformation Imaging (RTI) as an effective computational photograph technique developed in the domain of digital imagery.

A stack of digital photographs taken from the same location is composed, by varying each time the light source, with known beam direction. This produces a dynamic digital image (Figure 3-15), in which the surface of the detected object can be visualized under different lighting conditions, by appropriately controlling the direction of the illumination and thus allowing interactive relighting for vision, in dedicated web-based software (Palma et al., 2010).

3.1.2.2. The camera model

The photographs produced by the cameras can be considered, with sufficient approximation, central perspectives of the photographed object: the image of a point P is formed by a perspective projection of center C, the optical center of the camera, on the image plane (Figure 3-16, Figure 3-17).

The camera lens, which in fact consists of a group of lenses, is approximated as a single thin converging lens. The center of the diaphragm is also considered to be the optical center of the lens (Kraus, 1993). Thus, the optical laws stating that "any incident ray passing through the optical

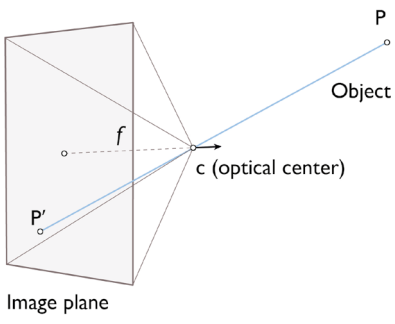
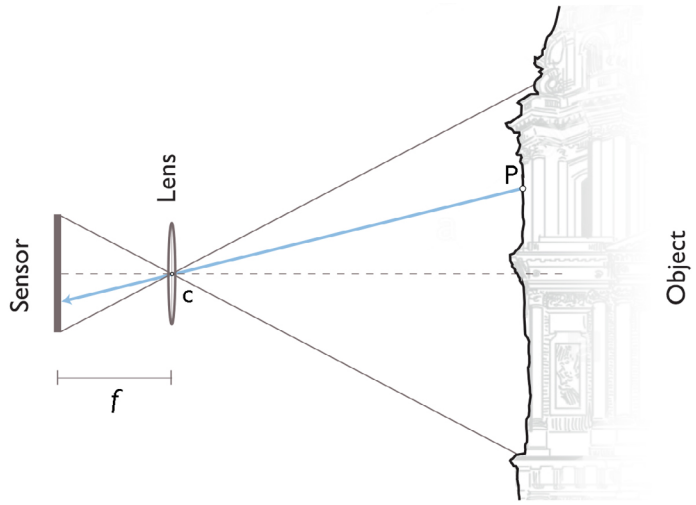


Figure 3-16.

Image as a perspective projection.

Figure 3-17.

Approximation of the camera model.



center of a lens is not deflected by the lens", allow us to schematize the object-camera system with the *pinhole camera model.*

The pinhole aperture of the camera, through which all projection lines must pass, is assumed to be infinitely small, a point. This point in 3D space, located at the intersection of the image rays, represents the perspective center.

From the geometric model of the pinhole camera and considering possible real aberrations from this ideal model, it is possible to derive, for each point of the scene (object coordinates in the 3D space - x, y, z), in which position (coordinates in a sensor plane, in two dimensions - u, v) this point will be projected on the image plane (De Luca, 2011).

To achieve this, the internal and external orientation parameters of the camera are inferred:

- Interior orientation parameters.

The position of the perspective center with respect to the camera reference system (image coordinate system) and the departures from the ideal central projection (image distortion) are the interior orientation parameters, as they describe the internal geometric model of the camera as it existed at the time of data capture (Luhmann et al., 2016). In detail, the position of the perspective center is determined by defining the distance f between image plane and perspective center and the image coordinates of the principal point. Distortion, on the other hand, corresponds

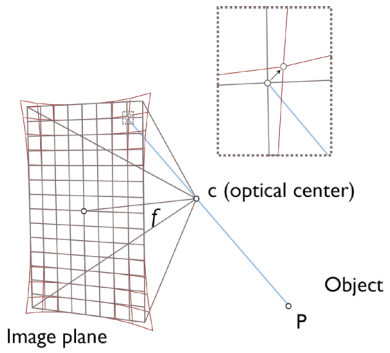


Figure 3-18.

Geometrical model of a camera with radial distortion. Adapted from (De Luca, 2011).

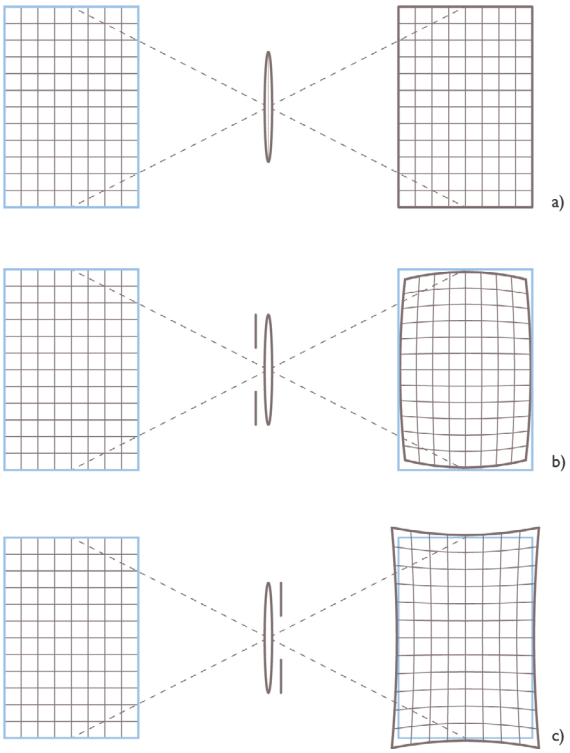


Figure 3-19.

a) absence of distortion;
b) barrel distortion;
c) pincushion distortion.

to systematic representation errors introduced by the shape and imperfections of the lens (Figure 3-18): it can be radial, tangential (Figure 3-19), induced by errors due to the sensor etc. The interior orientation, since including parameters related to internal geometry of the camera, is determined by calibration.

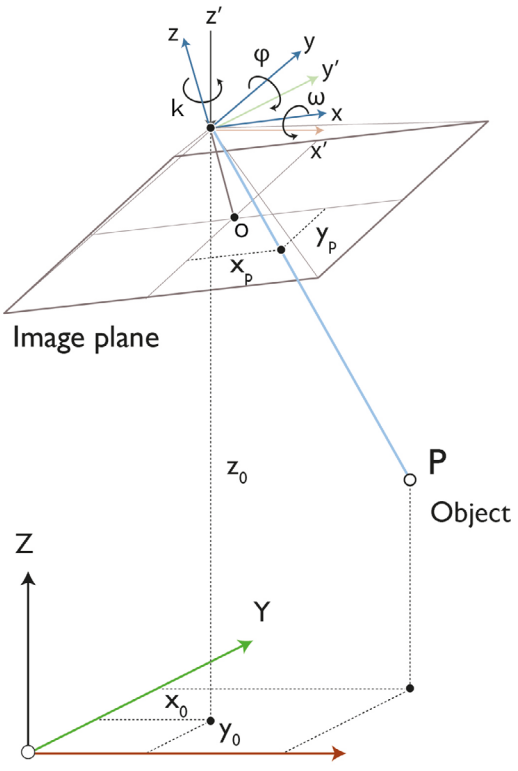


Figure 3-20.

Exterior orientation parameters.

- Exterior orientation parameters.

They describe the spatial position and orientation of the camera with respect to the global coordinate system, as it existed at the time of data capture. These parameters (Figure 3-20) are determined by three translations (coordinates with respect to the X, Y, Z axes) and three rotations (angles around the three axes: ω, ϕ, k). When the calibration and orientation parameters have been determined, the 3D restitution of the scene can be obtained through further measurements of correspondences (homologous points) between the images.

3.1.2.3. Image-based reconstruction methods

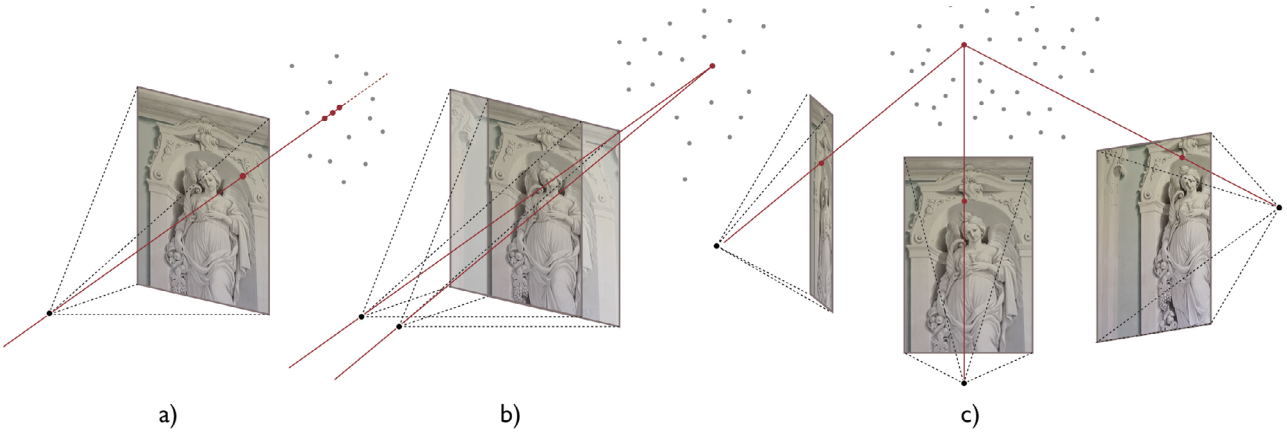


Figure 3-21. Mono-scopic (*a*), stereo-scopic (*b*) and multi-view stereo (*c*) vision.

The 3D coordinates of a point are derived, from photographs, considering that a point on the object, the projection center and the projected point in the image are collinear, i.e., lie on a straight line (*collinearity condition*).

A point on a single image only defines a ray to the object point, thus there is no information on the depth of the scene (Figure 3-21a). To find the position of a point in space, at least two images are needed (*stereopair*), as defined by the principles of stereoscopic vision (Figure 3-21b). From the intersection of at least two corresponding (homologous), spatially separated image rays, an object point can be located in three dimensions (Luhmann, 2013). When dealing with a higher number of images, appropriately taken and with sufficient overlapping, the object can be described from multiple points of view and in a more complete way (*multi-image photogrammetry,* Figure 3-21c). The schemes of restitution from a single image, a stereopair or a stack of images are



Figure 3-22.

Dense point cloud and visualization of camera positions.

e.g.: Saint-Trophime Cathedral, Arles

provided in Figure 3-21.

Starting from these considerations, the reconstruction of 3D structures from two-dimensional image sequences involves at first the search for key points between the images, then their matching to detect the so-called *tie points*. The camera position is estimated for each image and the camera orientation parameters are calculated.

The image alignment results in the creation of a sparse point cloud, together with the visualization of camera positions in the 3D space. Then, for each image, a dense matching reconstruction method generates the depth maps (providing a depth value for each pixel or group of pixels of the images), and these are later used to generate dense point clouds (Figure 3-22).

The creation of correspondences (tie points) between images and the feature extraction and matching, as key steps of the 3D reconstruction starting from images, have been strengthened and automated thanks to the combination of photogrammetry with the principles of Structure-from-Motion (SfM) and Multi-View Stereo (MVS), derived from computer vision (Luhmann, 2013; Pierrot-Deseilligny et al., 2011; Kasser and Egels, 2001). For this reason, nowadays, several software suites enable semi-automatic dense image matching procedures; among these, it is worth mentioning Agisoft Metashape (previously Agisoft Photoscan, agisoft.com), and the open-source platform MICMAC, developed by the French National Geographic Institute IGN (Pierrot Deseilligny and Clery, 2012; Pierrot-Deseilligny et al., 2011). Note, however, that in the case of photogrammetric imaging techniques, the 3D reconstruction is provided in an arbitrary reference system. By considering the insertion of a proper set of known coordinates (Figure 3-23) in a defined reference system, called Control Points (Croce et al., 2019; Croce and Martínez-Espejo Zaragoza, 2018; Martínez-Espejo Zaragoza, 2017; Rizaldy and

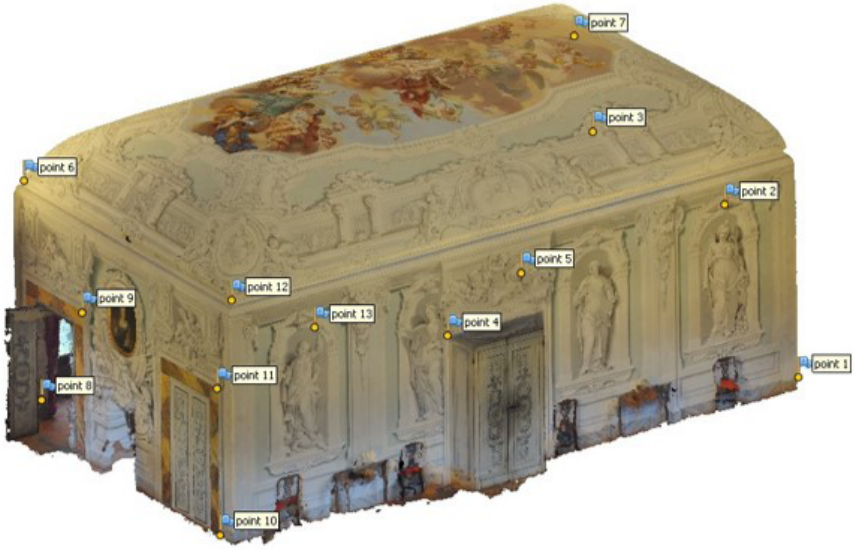


Figure 3-23.
Insertion of Control Points
e.g.: Vaults of the Grand-Ducal apartments, Pisa Charterhouse

Firdaus, 2012), the geo-referencing can be achieved.

- Incremental photogrammetry.

It is worth mentioning, in this section, that current applications of photogrammetry are going towards the achievement of incremental photogrammetry, intended as the spatialization of multiple new images starting from an initial photogrammetric dataset, in an incremental workflow. Progress against this data processing chain would allow consistency and correspondence between images and 3D model to be preserved, even as the original image dataset is enriched with new information.

Towards this effort, Pamart et al. (2020, 2019) have proposed a versatile and robust solution for the incremental registration of image sets, in the form of a Totally Automated Co-registration and Orientations (TACO) work-flow, that is aimed at enriching the 3D scene by adding oriented pictures – coming from a wide range of acquisition modalities –, around an original reference point cloud.

Figure 3-24 shows an example of application of TACO for the spatialization of two historical and archival images of the Cathedral of Arles. This means that images that have been acquired at different times or under different light and weather conditions could be inserted and linked to a source photogrammetric point cloud.

In this context, complementary and alternative tools are also considering the combination of photogrammetry with Neural radiance fields (NeRF), i.e. algorithms derived from Artificial Intelligence (AI) that synthesize novel views of complex scenes by optimizing an underlying

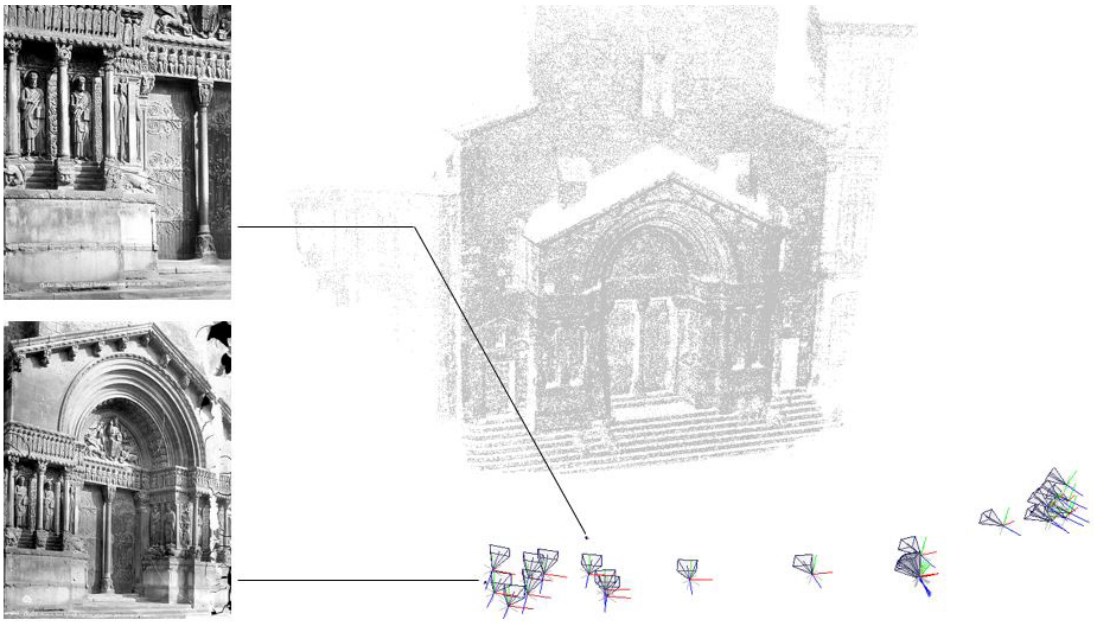


Figure 3-24.

Spatialization of two historical images over an original reference point cloud.

e.g.: Saint-Trophime Cathedral, Arles

continuous volumetric scene function using a sparse set of input views (Mildenhall et al., 2020). These methods use sparse images to create an underlying representation that can be explored and rendered from novel viewpoints with varying lighting condition (Tancik et al., 2021). However, the applications of this technology to cultural heritage have been little explored to date, also due to the difficulty of integrating it with existing SfM and MVS tools (Condorelli et al., 2021).

3.1.3. Comparative analysis of existing 3D surveying techniques

The review of 3D acquisition systems shows how digital survey solutions can be diversified according to several factors, such as the characteristics and complexity of the object to be surveyed (in terms of size, color, shape, chemical-physical and mechanical properties), the desired degree of detail and the surveying conditions. The results vary consequently, in terms of graphic / colorimetric restitution. In the case of active techniques such as laser scanning, the result of the acquisition depends in particular on the instrument-object distance, on the dimensions and on the material characteristics (e.g., surface roughness and reflectance) of the artifact to be detected. The 3D information is obtained directly, and it is at a real scale.

In the case of photogrammetric computation, the survey is strongly influenced not only by the geometry, but also by the brightness and colorimetry characteristics of the object and the scene. There is a

strong dependence on texture and for this reason, by properly planning the image acquisition phase (for example, in terms of overlapping between subsequent images and choice of distance and position of data capture), it is possible to obtain a better result in terms of color, hence of visual information. However, the photogrammetric acquisition may be challenging when dealing with reflective or transparent materials, too repetitive details or elements with uniform colorimetry (Abergel, 2020). The achievable accuracy thus depends on the pattern of the object to be detected, on the light conditions of the scene and also on the quality of the image in terms of brightness, sharpness, contrast, saturation adjustments set by the operator at the moment of data capture. It is also affected by the definition of the Ground Sampling Distance (GSD), which represents the size of an individual pixel relative to the object scale, and, in case of geo-referenced model, by the choice and position of control points (Martínez-Espejo Zaragoza, 2017).

Another term of comparison concerns the cost of instrumentation, which is significantly lower in the case of photogrammetry, where it is the photographic camera itself, with a fixed focal length lens, to constitute the survey instrument. This contributes to a certain extent to the *democratization* of the photogrammetric process: as highlighted by Bevilacqua et al. (2018), given the modest cost of digital camera, even an inexperienced operator using a redundant set of images taken from different points of view could obtain, in a short time, a realistic 3D model of a heritage artifact. This may lead to the absence of metric consistency and to inappropriate results.

3.1.3.1. Multi-modal acquisitions.

As pointed out by Remondino and Rizzi (2010), for 3D heritage documentation there is no single successful technique, but rather it is necessary to integrate different surveying methods in order to exploit the potentialities and advantages of the single one, compensate relative weaknesses, derive different levels of detail and description of the scene, achieve more accurate surveying results. Since the correct balance between accuracy, portability, cost, time of acquisition and flexibility is not often found in a single technique-especially for the case of complex heritage monuments, most surveying projects require the combination of multiple sensors and techniques (Remondino, 2011).

For this reason, research work in recent years has focused on the issue of multi-modal 3D data fusion, especially in terms of laser scanning/ photogrammetry hybridization (Caroti and Piemonte, 2020; Bevilacqua et al., 2018): in today's established integration techniques, the laser scanner survey constitutes the geometric reference, while dense-image matching techniques are applied to improve the color information and

the visual quality of the scenes (Grussenmeyer et al., 2012, 2008).

In fact, the laser scanner enables homogeneous accuracy in detecting 3D coordinates of the scene, and the survey is not influenced by lighting conditions or by the presence of elements with uniform colorimetry, or with too repetitive patterns (Bevilacqua et al., 2018). If the laser scanner survey allows to obtain a geometrically precise and scaled datum, on the other hand the color information obtained is of poor quality. Conversely, photogrammetry returns high quality radiometric and colorimetric data from the photographic campaign, but dense-image matching algorithms may not return a correct 3D model, especially in the presence of inadequate patterns.

Integrating the different techniques appears in any case the best solution to exploit the potential and advantages of the individual instruments, allowing to obtain a better result in terms of precision of the single data, overall accuracy and optimization of the entire process. After the integration, the correspondence between laser scanner and photogrammetric surveys can be verified through cloud-to-cloud or cloud-to-mesh comparison algorithms.

For the given reasons, most case studies presented in this thesis are a result of the integration between UAV-based photogrammetry, terrestrial photogrammetry, and/or terrestrial laser scanning techniques (Croce et al., 2019). These techniques are flanked by traditional topographical surveying techniques (total station, global navigation satellite systems - GNSS) in order to derive each time relevant information on the heritage monument or site studied at different scales.

3.2. Processing of 2D/3D information

When dealing with digital heritage models, a common problem is data interpretation, intended as the process that assigns a meaning to the information analyzed and displayed, determining its significations and implications. Indeed, a correspondence between the information that is extracted from the raw digital data and the interpretation that the human eye can infer from it is to be established.

Regardless of the type of technology used, the outputs of digital survey thus require a processing phase, or at most, the manipulation with specific tools allowing to extract pieces of information about the visual and/or metric content of digital 2D/3D data. For these reasons, the research has been expanding with regard to the definition of appropriate processing sets, as specific structures of 2D and 3D digital data significantly characterizing digital heritage contents. Through a processing phase, the digital information is associated with a certain semantic meaning (Carboni and Luca, 2019; De Luca, 2011). The choice on the type of information to be extracted and on the type of processing to be performed depends on the demands and purposes of restitution and analysis of an asset.

In the following section, current data processing methods applied to 2D imagery and 3D point clouds are analyzed and discussed; in detail, we analyze the structures and methods allowing to add characteristic features to raw data, as a premise for the studies on semantic segmentation through AI.

3.2.1. 2D visual and metric descriptors

Nowadays, several systems allow the restitution of visual attributes starting from 2D images, and there exist various software enabling manipulation and processing of input data, for the extraction of suitable features. As for digital images, editing operations may include, e.g., brightness adjustments, tonal values, cropping, element selections, application of filters or effects. The adjustments are applied to individual pixels of the image to fix, e.g., an overexposure, correct the color, and improve the brightness (Hassaballah et al., 2016). Besides, color rendering methods that include the extraction of the single R, G, B channels of an image for a subsequent mapping in false colors (*image in false colors*, Aveta et al., 2017) or even the enhancement of the color separation to improve visual interpretation and make feature discrimination easier (e.g.: *decorrelation stretching*, Alley, 1999).

The adjustments layers can concern, alternatively, the whole image

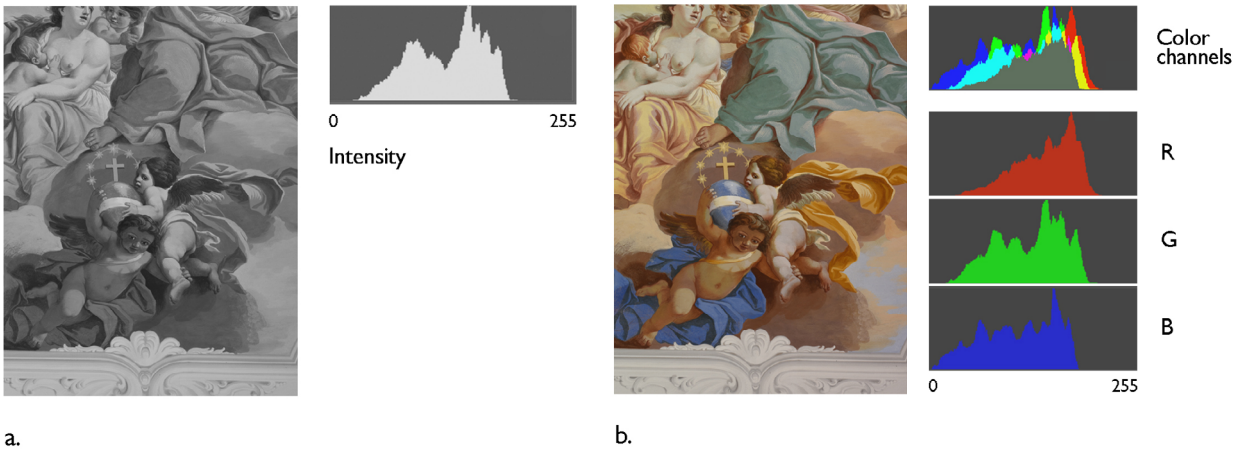


Figure 3-25.

Histograms of a BW (*a*) and color (*b*) image. e.g.: Vaults of the Grand-Ducal apartments, Pisa Charterhouse

in its entirety (*global features*) or patches, i.e., small groups of pixels (*local features*) (Kabbai et al., 2019). They alter each time the histogram of the image (3.2.1.1). Besides the traditional adjustments of hue, saturation and brightness (3.2.1.2), we hereby consider other types of filters, classified according to three categories: edge detectors (3.2.1.3), texture filters (3.2.1.4) and noise-reduction filters (3.2.1.5). They will be used later for the purposes of semantic segmentation.

3.2.1.1. Histogram and image moment.

A descriptor that can be extracted globally is the image histogram, a graph displaying the number of pixels contained in an image as a function of the intensity. In a grayscale image, for the different possible intensities (0 to 255 values), the histogram reports the distribution of pixels amongst the grayscale values; for the case of color images, an individual histogram can be derived to display the variations of red, green and blue channels, respectively (Figure 3-25).

Image moments can also be considered as image descriptors (Messaoudi, 2018; Manuel, 2006); these are scalar quantities used to characterize a function and to capture its significant features. They can be used to derive simple image properties such as areas (or total intensity), barycenter and/or orientations (Flusser, 2002; Teague M.R., 1980; Hu M.K., 1962).

3.2.1.2. Feature extraction: image adjustments.

Features can be extracted by direct manipulation of the images: the tonal range and color balance of an image can be adjusted by controlling, e.g., intensity levels of image shadows, mid-tones, shadows and brightness. The operator either guides the process, by regulating the adjustment of the image, or relies on automated editing tools; in this

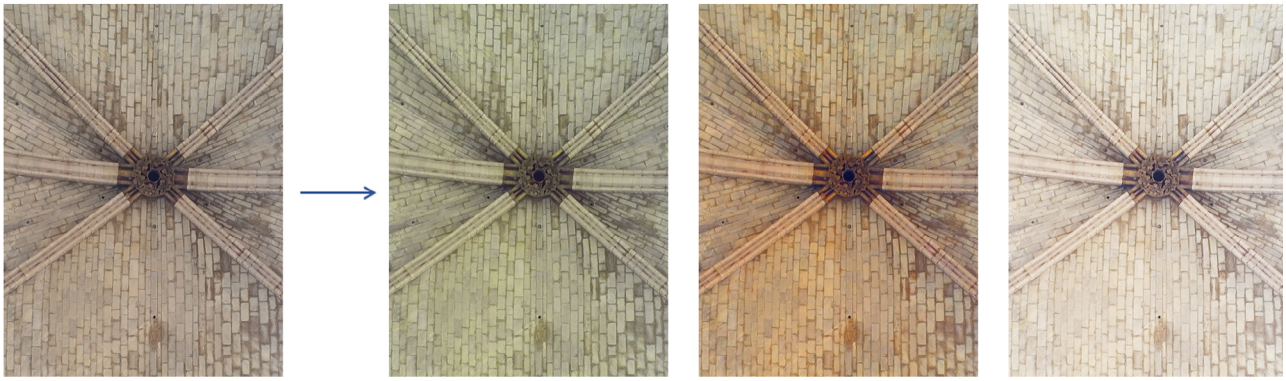


Figure 3-26.

Example of image adjustments from input data (*a*): hue (*b*), saturation (*c*) and brightness (*d*).

e.g.: Vaults of the Notre-Dame Cathedral in Paris.

way, adjustments (in terms of level, contrast, gamma, hue, saturation, and brightness modification), filters and effects can be applied to the image, and regions of interest can be selected to distinguish certain elements or pixels or patterns. Visual guides and tools are provided in common image processing software, e.g., Adobe Photoshop, Adobe Lightroom (adobe.com) and ImageJ (imagej.nih.gov), to achieve these tasks (Figure 3-26).

For instance, when using a set of image acquisitions to provide color information of an elevation drawing or an ortho-photo, light conditions may vary between one image and another; in this case, color and intensity values should be equalized in the post-processing phase, generally based on the elaboration of RAW format files. The operator can intervene by selecting and appropriately correcting an image to make it match-in radiometric terms- to the others. This practice is widely applied in restoration and recovery domains to improve color and graphic rendering of 2D drawings. However, the manual adjustments of heritage images can lead to different solutions depending on the operator, or, for the same operator, depending on the different adjustment level or layer chosen. If filters are applied to a portion of the image, the result may even vary depending on the hand-made selection. This can lead to lack of or poor reproducibility, high user bias, high intra- and inter-user variability, which are incompatible with the search for objectivity needed in the study of cultural heritage.

3.2.1.3. Feature extraction: edge detectors.

Edge detection filters aim at identifying edges and boundary objects of an image, by the recognition of sharpness or brightness changes; discontinuities produced by sharp edges result in changes in the intensities of the pixels that delineate the object borders (Shrivakshan, 2012). Laplacian and Sobel filters, difference of Gaussians, Hessian

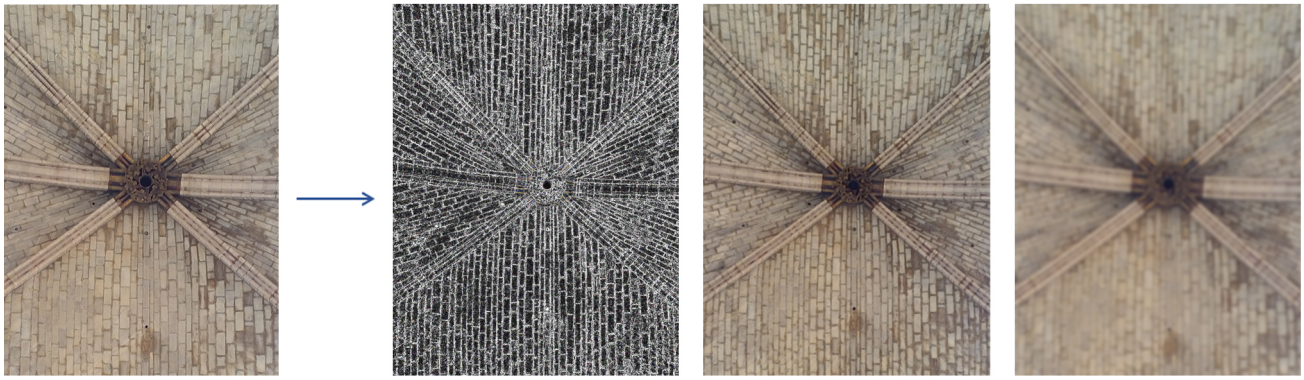


Figure 3-27.

Image adjustments from input data (*a*): edge detectors (Sobel filter, *b*), texture (Minimum, *c*) and noise reduction filters (Gaussian blur, *d*).

e.g.: Vaults of the Notre-Dame Cathedral in Paris.

matrix eigenvalues and Gabor filters belong to this class and allow enhancing, e.g., the distinction of joints and ashlars (Figure 3-27b).

3.2.1.4. Feature extraction: texture filters.

These filters extract texture information by appropriately averaging the color of pixels in a given pixel neighborhood (including filters such as minimum, maximum, median, variance, entropy, structure tensor, etc.). As such, they provide information about the local variability of more or less smooth depending on the greater or lesser variation of the pixels range of values in the chosen neighborhood (Heckbert, 1986) (Figure 3-27c).

3.2.1.5. Feature extraction: noise reduction filters.

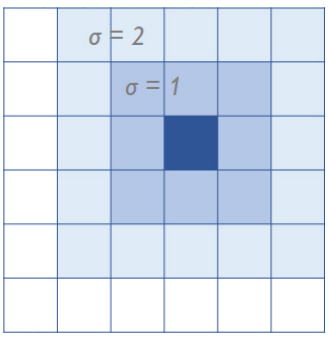


Figure 3-28.

Pixel neighborhood σ.

Filtering algorithms alter the image noise by blurring and/or by replacing the intensity of each pixel with a weighted average of intensity values from nearby pixels (Figure 3-7d). Gaussian blur, bilateral filters, anisotropic diffusion, Kawahara and Lipschitz are filters of this type (Van De Ville et al., 2003).

Such filters are generally computed in a chosen pixel neighborhood; in other words, the pixels within a radius of σ pixels from the target pixel are concerned by the adjustment (Figure 3-28).

Each time a filter is applied to the image, the image histogram changes, so allowing certain features to be emphasized rather than others. Such



Figure 3-29.
Point cloud visualization as an ensemble of unstructured data.

descriptors can therefore reveal quantitative information and visually accentuate certain features; however, it has to be noted that they do not provide information about the content or about the objects represented in the image (at most, only the objects' edges can be recognized).

3.2.2. 3D visual and metric descriptors

The raw data acquired from surveying are expressed as a set of points with 3D coordinates in a given reference system; at most, the points can be associated with color information (photogrammetry or laser scanners equipped with cameras) or intensity information (laser scanners), as highlighted in Section 3.1 (Figure 3-29).

As unstructured data, point clouds do not contain semantic information. Moreover, the large file sizes complicate their use and manipulation. For these reasons, the surveying outputs cannot be exploited directly, but rather must undergo a series of suitable processing methods, aimed at accentuating the visual and/or metric properties of the 3D data. The elaboration strategies presented in this section depend on the purposes of the investigation carried out on the architectural asset and can be adapted according to the objectives of the graphic restitution.

3.2.2.1. Point cloud registration and geo-referencing

When dealing with complex or massive historical buildings in size, the 3D scans of the different surveyed environments are to be linked together, as the survey object may not be visible from a single scan station. Several scans are thus to be considered: each scan is first generated in its own reference system. For laser scanners, this is a reference system proper to the instrument and the given scan station; for photogrammetry, the reference system is again arbitrary, and in addition the model is not full-scale. In any case, the mutual orientation between the scans is unknown

(Bevilacqua et al., 2018).

It is therefore required, at first, to align in a single coordinate system the set of point clouds acquired (Figure 3-30). This operation goes under the name of *point cloud registration* (Guidi and Remondino, 2012) and it is aimed at estimating the transformation matrix between point clouds captured from different scan positions, in order to represent the entire object with respect to a certain reference system (Shanoer and Abed, 2018).

The point cloud registration can be performed by introducing natural or artificial targets in the scene that may serve either as common points visible from adjacent scans or as points of known coordinates in the same reference systems to which the individual scans can be linked (Bevilacqua et al., 2018).

Many commercial TLS solutions are equipped with spherical or planar scan targets, which, if placed in suitable chosen positions, allow the automatic recognition of targets during registration, through best-fit algorithms.

When expressing the 3D point coordinates with respect to a known global reference system, e.g., a topographic system or a GNSS network, and calculating the roto-translation matrix from the original reference system to the global one, the related operation is called *geo-referencing* (Otepka et al., 2013). The performance of the process is influenced by the accuracy with which the coordinates of the GCPs were determined. In yet other cases, the registration operation is performed by *Iterative Closest Point* (ICP) algorithms, which link the different scans by exploiting natural 3D features in the scene as reference and on the basis of a considerable level of overlapping (Guidi and Remondino, 2012). In the latter case, the accuracy depends on the standard deviation of the ICP process.

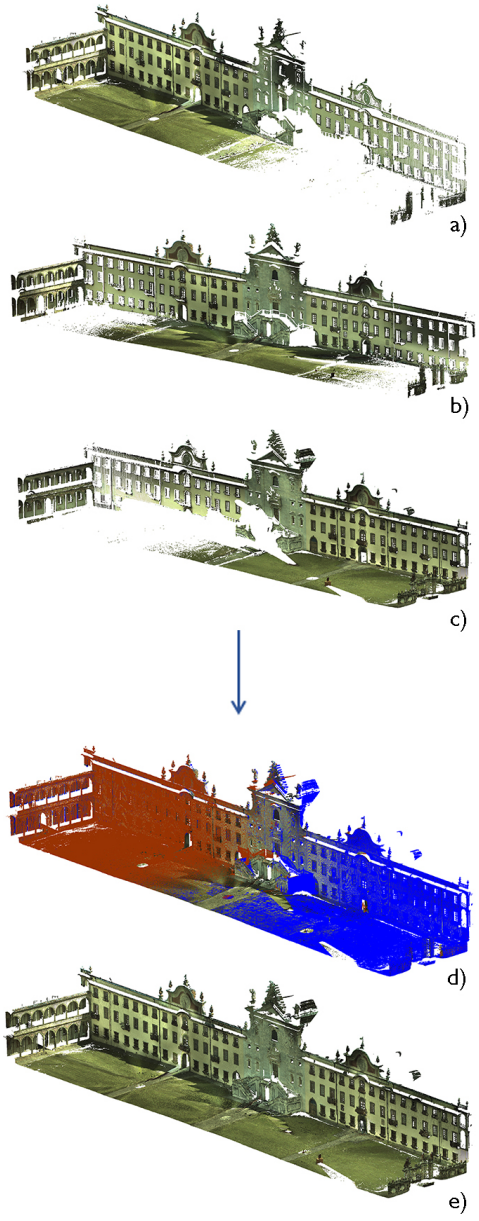


Figure 3-30.

Registration of point clouds acquired from three single scan stations (*a, b, c*) in a same reference system (*d, e*).

e.g.: Main façade, Pisa Charterhouse

3.2.2.2. Filtering, noise reduction, decimation.

In order to produce a consistent data set, processing may require:

- Reducing the point cloud noise, intended as the irregular deviation of the points from the detected objects. Noise can be due to instrumental features or interferences (excessive divergence of the laser beam from the surface, instrument vibrations), positioning factors (excessive proximity of the detected object) and environmental factors (interference due to the presence of obstacles or elements with high reflectivity, non-optimal light and humidity conditions) (Macher, 2017).
- Reducing data size. A resampling phase (*decimation*) enables easiest handling of digital 3D information and speeds up further processing and visualization steps, to the detriment of a loss in resolution and detail; the format type is also crucial, as well as the choice to divide the data into portions of smaller size (Landes, 2020).

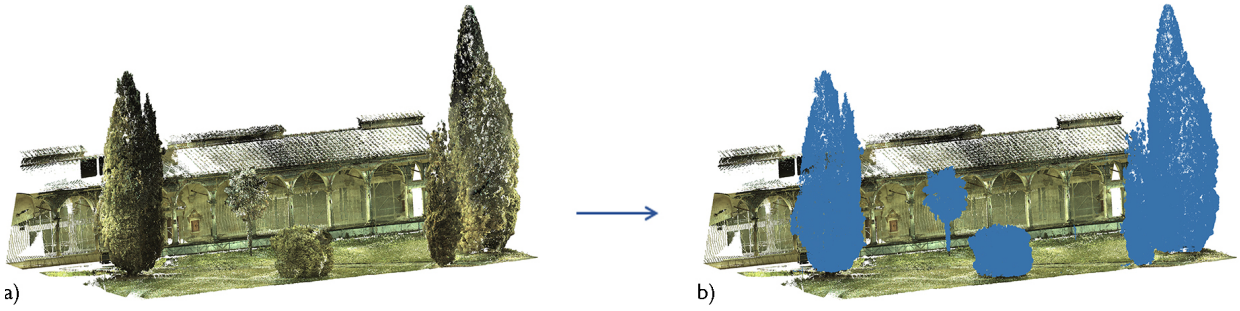
For what concerns *noise reduction and filtering*, tools are now available to correct and/or remove false or unwanted points. On the one hand, the manual deletion of obstructive elements (e.g., trees, street-lamps, cars for outdoor surveys; furniture, tables, chairs for indoor surveys) may take place by delimitation and subsequent removal, by the user, of selected regions (Figure 3-31); on the other hand, automatic tools are currently available for point cloud filtering. The latter rely on the exclusion of groups of points (*outliers*) that occupy wrong positions in the 3D space or even located too far from the rest of the point cloud.

As an example, CloudCompare, i.e., a common point cloud processing software (cloudcompare.org), includes the *Noise reduction filter* and the *Statistical Outlier removal* (Han et al., 2017). A threshold distance is established to determine whether a point should be considered an *outlier* (and thus remove it) or not. In the case of the noise reduction filter, first the average distance of each point from its neighbors is calculated; then, the points that are farther than the average distance plus a certain number of times the standard deviation are deleted.

Figure 3-31.

Example of obstacle removal by selection of trees and low vegetation (*b*) from an original point cloud (*a*).

E.g.: Main cloister, Pisa Charterhouse.



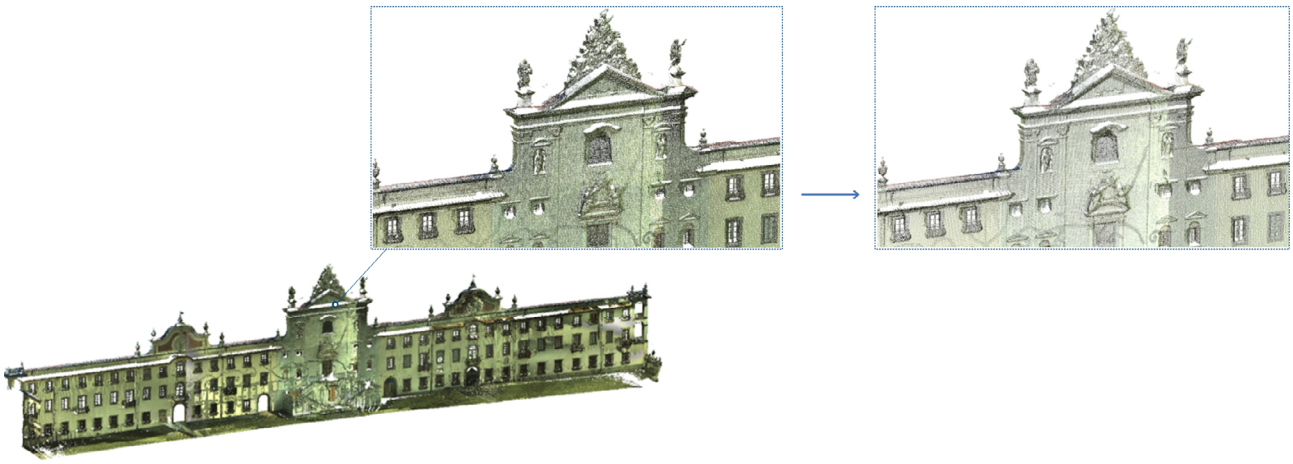


Figure 3-32.
Point cloud decimation.
E.g.: Main façade, Pisa Charterhouse.

As for *sub-sampling operations,* they are aimed at causing the number of points in the cloud to decrease. Again, the decimation can be performed automatically by setting a minimal distance to be preserved between two points. In a homogeneous point reduction process, no point in the output cloud will be closer to another point than the specified value. The larger this value, the fewer points are preserved (Figure 3-32). However, the sampling distance can also be modulated with the scalar values, so that more points can be sampled, e.g., in curvy areas and less points in planar ones, to preserve the level of detail of some curvy elements of the surveyed object.

Although decimation significantly reduces the file size, the number of points (hence, the density) of the resulting cloud should not result too low, so to avoid significant loss of detail. The best compromise between number of points and level of precision and object description is indeed to be established.

3.2.2.3. 3D feature extraction: association of visual and metric descriptors

Metric and visual descriptors can even be associated to 3D data (Section 3.1).

In this case, the concept of *scalar field* is essential and needs to be introduced: a scalar field represents a set of values (one per point), that can result from the application of filters (smooth, gradient,...), of mathematic and statistical operations to each point of the point cloud. The result of these operations can be displayed visually by custom colors and color maps (cloudcompare.org).

It follows that a point cloud can be represented as a row-by-column matrix: the rows correspond to the list of 3D points, while the columns display, apart from X, Y, Z coordinates and color data, the different

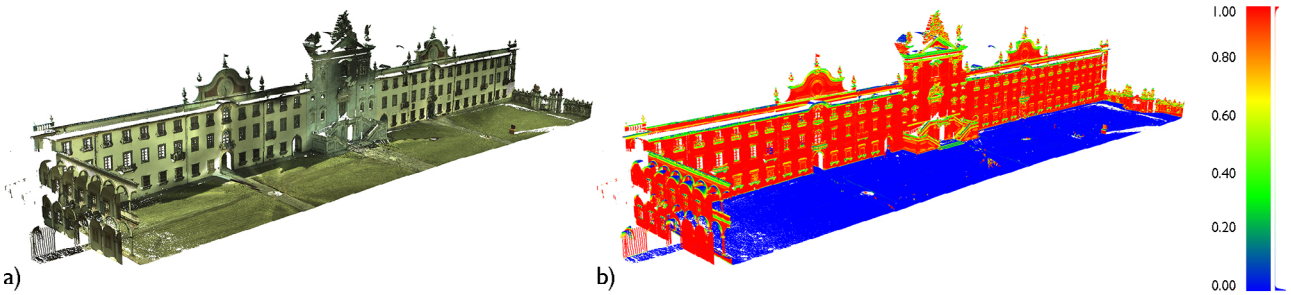


Figure 3-33.

Feature extraction (*b*) from a raw point cloud (*a*).

e.g.: Verticality feature for the main façade of Pisa Charterhouse. with the associated color map.

scalar fields associated with the point cloud (Table 3-1). The number of columns will thus correspond to the number of information that is conveyed by each point. As further explained (Chapter-5), geometric, radiometric, intensity and curvature features can be extracted as metric and visual attributes of the 3D points (Figure 3-33, Figure 3-34), so to assist object recognition and semantic segmentation phases.

Each feature can be calculated by choosing a proper spherical neighborhood of each 3D point, and displayed as a column of the point clouds' matrix and represented graphically through color scales, to show the variation of feature values for the whole collection of 3D points.

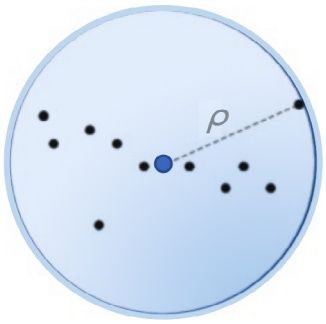


Figure 3-34.

Spherical neighborhood of a 3D point, with radius ρ.

3.2.2.4. Metric information and extraction of 2D views

Bi-dimensional views can be extracted starting from 3D point clouds and models, by appropriately choosing a projection system. This aspect is crucial as deformations due to the projection system used should be limited as much as possible for a correct restitution of 2D views, even according to the method that best suits the given task (Abergel, 2020).

Table 3-1.

Matrix structure of a 3D point cloud.

Point	X	Y	Z	R	G	B	Scalar field
1	X_1	Y_1	Z_1	R_1	G_1	B_1	f_1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	X_n	Y_n	Z_n	R_n	G_n	B_n	f_n



Figure 3-35.

Visualization of a 3D model in perspective (*a*) and orthographic (*b*) projection.

e.g.: Altar of Santa Vittoria, Chiesa del Carmine in Pisa.

Most common types of projection are perspective, parallel projections (Figure 3-35) and panoramic projections (Figure 3-36; De Luca, 2009):

- Perspective projection. It is a linear projection and can be distinguished into one-point, two-point or three-point perspective based on the number of vanishing points. The 3D objects are projected on an image plane, so that closer objects closer appear larger than those further away. The actual visual perception is approximated by creating, starting from the 3D model, a virtual photograph with orientation and angle of view chosen by the operator. Parallel projection can be seen as a perspective projection where the vanishing point is at infinity, as the projection lines are parallel to each other. The orthographic projection is then a form of parallel projection in which the direction of projection is perpendicular to the projection plane. Parallel projections can be used as a measuring support, and it is on the principles of orthographic projection that elevations and ortho-photos can be generated (Figure 3-35).

- Panoramic projections allow creating images with extremely wide fields of view (Conrat, 2010). Cylindrical, spherical, or cubic projections are distinguished depending on whether the scene is projected, from the optical center, onto a cylindrical surface, a spherical surface, or the six faces of a cube, respectively (Figure 3-36). Such projections are achieved either by on-site acquisitions with specific panoramic cameras or by specific software enabling the assemblage of multiple views of a same object (Manuel, 2016).

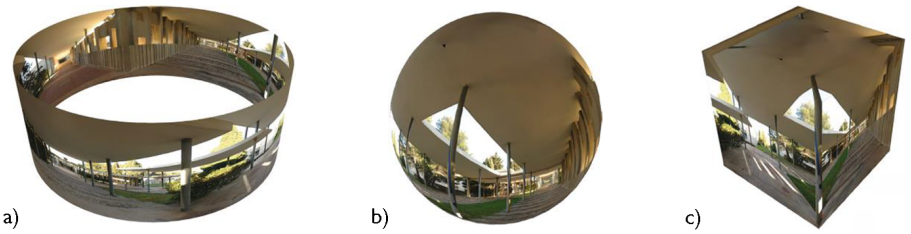


Figure 3-36.

Panoramic projections: cylindrical (*a*), spherical (*b*) and cubic (*c*) projections. (De Luca, 2009).

Considering the elaboration 2D views starting from the application of these projection systems to the result of 3D outputs, the most used tool is undoubtedly the ortho-photo. This media enables the extraction of dimensional-metrically correct- information of the studied object, as optical distortions are corrected and apparent changes in the elements' positions caused by perspective are avoided. By projecting the scene onto a chosen reference plane, the orthographic projection combines precise metric information with detailed visual information. The measurement of distances and areas can thus be obtained from ortho-rectified images in scale. In the architectural heritage context, ortho-photos offer an effective tool to restitute the decorative and material apparatus (Caroti and Piemonte, 2020; Perfetti et al., 2019) as well as to describe of the degree of conservation of wall or frescoed surfaces (Aicardi et al., 2018).

Besides ortho-photos, all types of geometric views used in architecture can be extracted. Plans, sections and elevations can be returned based on point clouds processing, to render the layouts, shapes and dimensions of the object under study according to parallel projections. To this end, some tools implemented within modern point cloud processing software (e.g.: CloudCompare and Autodesk ReCap) allow to interactively define a clipping box around a given cloud, so as to adjust the point cloud extents and orientation in order to segment it and consequently visualize or hide certain portions of the input data. This tool is particularly useful for the elaboration of sections views, or even for segmentation of certain portions of the cloud (Figure 3-37a).

It is worth recalling, however, that 2D views are not directly deducible from the raw acquisition but rather result from the human interpretation of the 3D data, as output of long elaboration processes.

Finally, metric information can be derived, based on the 3D survey, from the recognition of relevant points. The extraction of metric measurements can then be performed from the selection, in the unstructured set of 3D points, of meaningful points with respect to a certain exigence, deriving information such as: coordinates extraction (Figure 3-37b), information on scalar fields, measurement of angles, distances (Figure 3-37c), or level differences (Abergel, 2020). In all cases, the extraction of metric information from a point cloud requires prior selection of points of interest and extraction of their coordinates, and the input data must be scaled or geo-referenced.

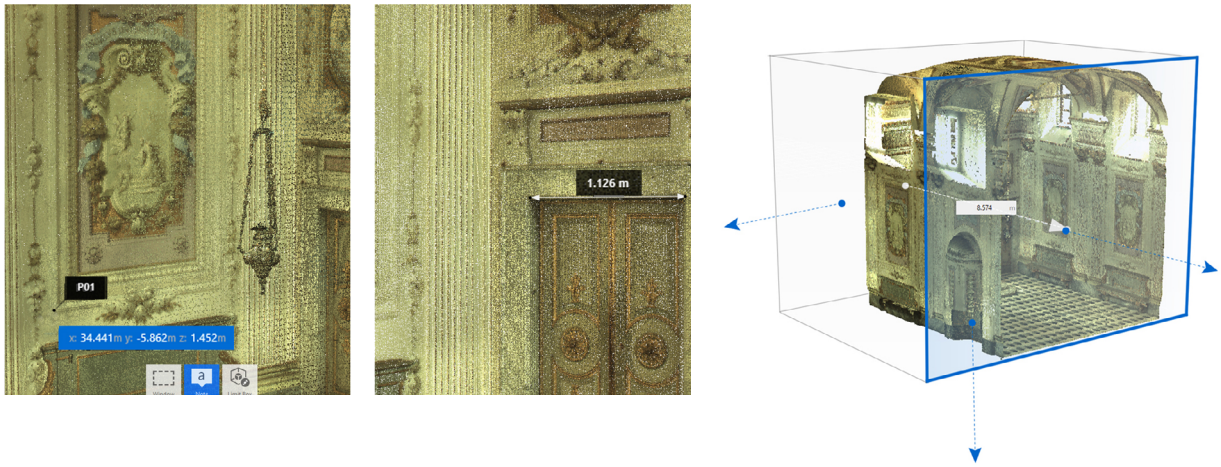


Figure 3-37.

From left to right: Limit box tools allowing: extraction of 3D coordinates *(i)* and distances *(ii)* from a point cloud and the creation of section plans *(iii)*.

e.g.: Cappella del Rosario, Pisa Charterhouse.

3.2.2.5. Mesh generation.

The raw point cloud obtained from surveying can be furthered process to obtain a polygonal mesh model (Figure 3-38). In the case of structured point clouds, meshing from a set of discrete points is a straightforward step; however, in the case of cultural heritage applications, since dealing with rather unstructured point clouds obtained from reality-based surveying, the meshing process is not immediate (Remondino, 2011). The Delaunay triangulation method (Chew, 1993), or the screened Poisson surface reconstruction (Kazhdan, Hoppe, 2013) are generally used algorithms to create surfaces from oriented point sets. The resulting mesh has vertices, edges, and faces, and represents the set of scanned objects. In the case of heritage point clouds, however, such objects are not distinct one another, and the meshing procedure may require a preliminary segmentation step.

Several processing steps can be envisaged later on the mesh, involving e.g., hole filling, removal of unnecessary polygons or collapse of nodes that cause inconsistencies. One can also change from a triangular mesh to a quadrilateral mesh (Tarini et al., 2010), a hexagonal mesh (Ray et al., 2017), or other types of polygons (Owen, 2000).

As in the case of point cloud subsampling, even the mesh can be reduced in file size, for better management in a virtual environment, by a decimation process (also called *simplification*), which aims to reduce the number of mesh faces while keeping the overall shape, volume and boundaries preserved as much as possible: for structural analysis, dissemination or web content purposes, for example, these simplified mesh models are used (Bevilacqua et al., 2018).

Existing web-based platforms, such as 3DHop (3dhop.net) and Sketchfab (sketchfab.com) leverage mesh simplification approaches to offer on-line and real-time visualization of 3D heritage objects and sites.



Figure 3-38.

Sparse point cloud (*a*); dense point cloud (*b*); mesh (*c*); textured model (*d*).

e.g.: Altar of Santa Vittoria, Chiesa del Carmine in Pisa.

3.2.2.6. Texture: visual enrichment and shaders.

After the mesh geometry has been reconstructed, the color information can be enhanced by generation of a texture (Figure 3-38d). In photogrammetry, the images used-and suitably oriented- are exploited to create and project color information over the mesh.

The projection of color information to the mesh surface is achieved by the so-called *UV mapping* process (Figure 3-39), that assigns a texture coordinate to each vertex of the mesh: in this way, the texture can be transferred to the model's surface, by projecting the coordinates from the 2D space of the texture (u and v axes) to the 3D space of the model (x, y and z axes). Texture maps thus represent a 2D projection of the color of the studied object, and as such provide significant information on its visual radiometric attributes; they can guide, for example, the recognition of alterations on the surface or ease the identification of fractures or discontinuity zones. This allows to visually appreciate the objects' features, but also to benefit from the usual image processing methods to manipulate the 3D information (Abergel, 2020). Indeed, the UV map can be extracted to further process the texture on a 2D canvas, and to subsequently project the information onto the 3D model.

As demonstrated later on (Sections 5.3.2 and 6.2), this aspect proves extremely useful for the virtual visualization and 2D/3D transfer of segmentation information.

Finally, other texture maps (Figure 3-40) can be extracted and associated

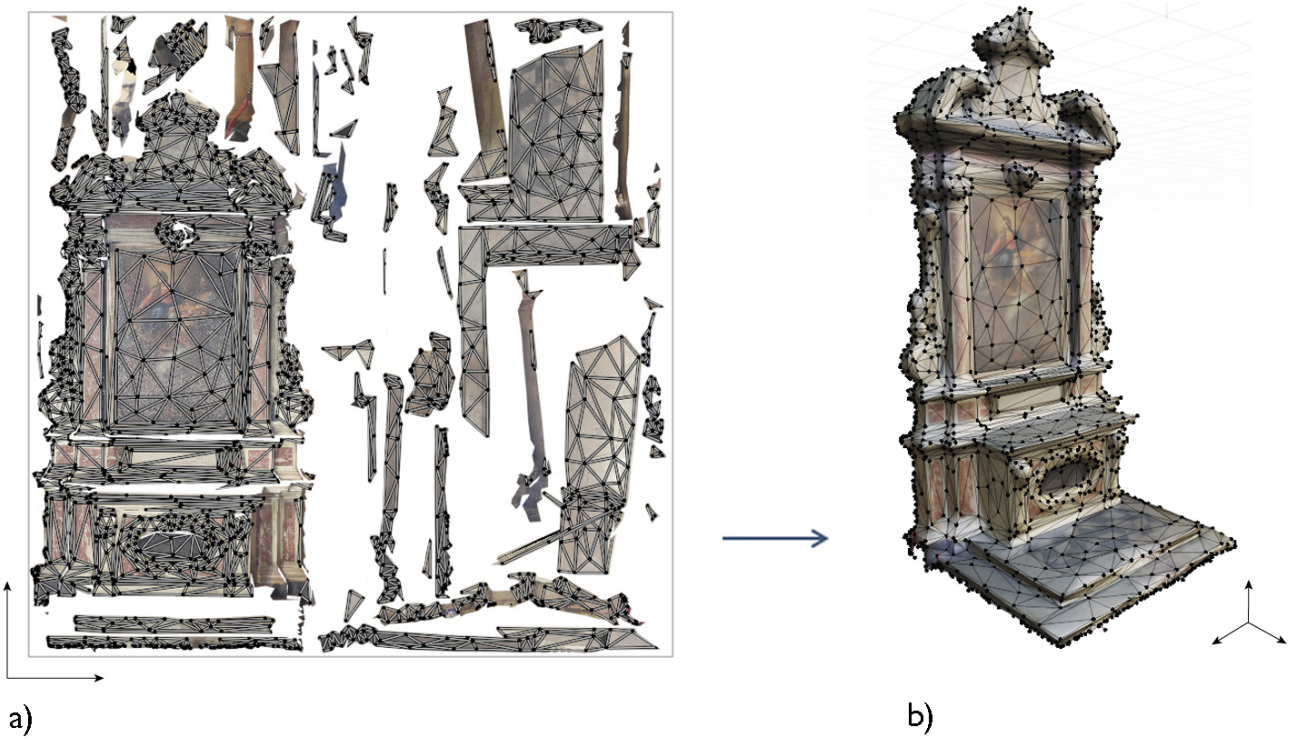


Figure 3-39.

Projection of a UV map (*a*) onto the 3D model (*b*).

e.g.: Altar of Santa Vittoria, Chiesa del Carmine in Pisa.

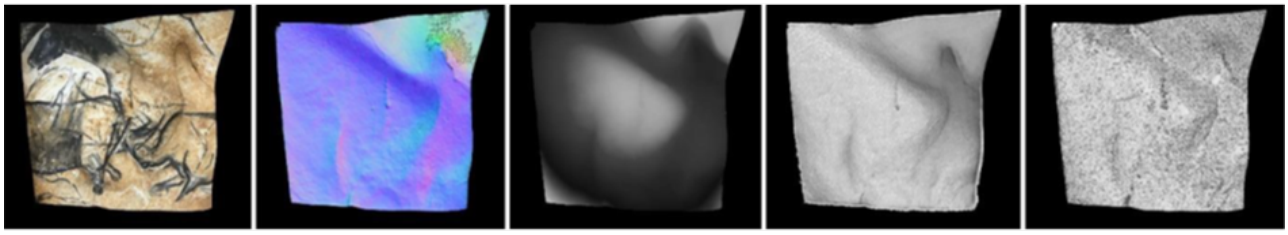
to the main one, to create complementary features that simulate the physical properties of an object: among them, e.g., the normal map of a 3D object facilitates the distinction between flat and curved surfaces; the ambient occlusion can characterize the object's illumination and lighting; the depth map can describe the distance of the represented objects from a certain viewpoint; the roughness map can help detecting surface irregularities that may cause light diffusion.

3.2.2.7. Disparity maps and temporal tracking

Figure 3-40.

Examples of texture maps characterizing geometric or visual attributes of the study object. *From left to right:* colors, normals, depth, ambient occlusion, roughness. (Abergel, 2020)

Reality-based surveying provides an essential tool for recording the state of conservation of a building or historical site at a given time. It follows that it is possible to compare the condition of this heritage object in successive temporal phases. Let us consider, e.g., two point clouds relating to the same heritage object, acquired at different times and placed in the same reference system. In this case, discontinuity



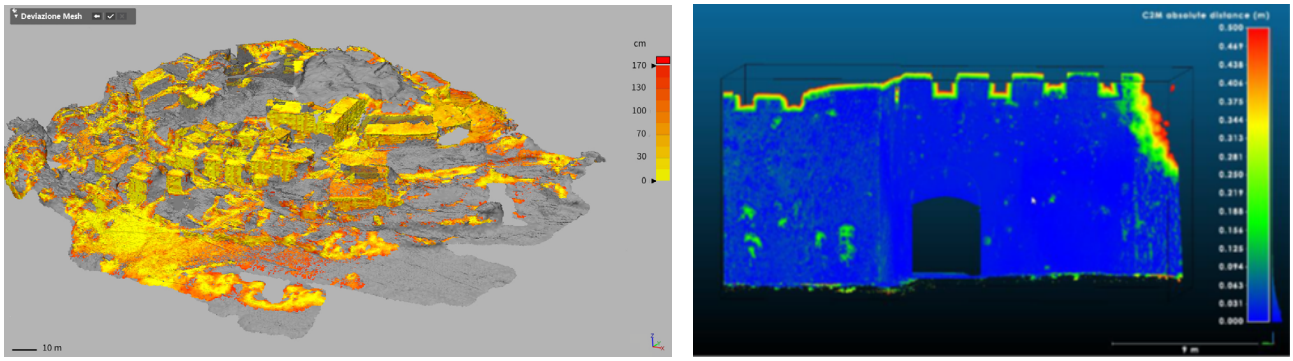


Figure 3-41.

Disparity maps. Comparison between the pre- and post-earthquake assets of Castelluccio di Norcia (Croce and Martinez Espejo Zaragoza, 2018; *to the left)*; Comparison between 2010 and 2019 models of Pisan urban walls (Martinez Espejo Zaragoza et al., 2021; *to the left*).

maps can be generated to highlight the evolutions and changes that the object has undergone from time to time, as well as to identify damage maps of successive states (Figure 3-41, Croce and Martínez-Espejo Zaragoza, 2018).

The distance between the two models can in fact be calculated and the result can be graphically displayed as a scalar field, through the exploitation of point-cloud processing software such as CloudCompare. For this method to be valid, however, the point clouds must have comparable densities (Martinez Espejo Zaragoza et al., 2021).

Similarly, the comparison between a point cloud and a mesh of the same object in two temporal states can be studied.

3.2.3. Comparative analysis of 2D/3D processing methods and research challenges

In this section, processing methods for unstructured 2D/3D data obtained from in-situ acquisitions have been shown, illustrating how the enhancement of geometric and visual attributes stored in digital replicas can provide valuable information for the study and analysis of architectural assets.

Point clouds obtained from laser scanning or photogrammetry can be exploited for subsequent meshing, texturing and optimization processes, or for the elaboration of 2D views (ortho-rectified images, sections, elevations). In the same way, filtering, decimation, obstacles removal, meshing and texturing operations can be performed more or less automatically. As for two-dimensional formats, also in the case of 3D data it is possible to extract appropriate features (geometric, radiometric, intensity, curvature...) to be used for classification applications. The list of features that can be extracted from images and 3D data provided is by no means exhaustive; however, we are hereby interested in concepts that will be taken up later in relation to the semi-automatic processing of digital 2D/3D material, fundamental for the segmentation phase relying on AI. The extraction of metric and visual descriptors is indeed preliminary to the classification phase and to train predictive models for the semantic segmentation task (Chapter 5).

It has to be noted, however, that the amount of information that can be recorded by heritage surveying is vast: as discussed in the present section, methods for digitally synthesizing and manipulating heritage content are today countless and widely proven. However, the plurality of achievable products and representation types is the source of data traceability and interoperability problems and may thus lead to dispersion of knowledge. This forbids bringing together and connecting the huge amount of material produced on a given architectural object and in the framework of multidisciplinary heritage studies.

A semantic reading of the various 2D/3D data produced is thus needed, to ensure correct data archival and retrieval; to this purpose, beyond the plain and consolidated techniques of sharing and visualization for digital heritage, the challenges of current research aim at the generation of digital information systems enabling the connection of the purely geometric representation of buildings and sites with meaningful knowledge. In the next chapter, we will stress how it is precisely in the semantic interpretation of heritage data that a raw surveying output can be distinguished from a digital model rich in information:

'A digital model contains both geometric and semantic information [while] a point cloud represents a building geometry but does not contain semantic information.'

(Macher, 2017)

3.3. Digital information systems for the semantic enrichment of heritage models

The attribution of additional and well-structured information to a digital model is achieved by the so-called semantic annotation. This mechanism enables the link of the purely geometrical and visual attributes derived from 2D and 3D data elaboration with valuable external and complementary information, relating, e.g., to results of documentation, preservation, conservation and restoration activities.

Added contextual information associated to a graphic element (or parts of it) of the digital model could be related, e.g., to analysis and conservation terms (e.g., materials, cracks patterns, intervention and recovery measures).

In the annotation process, Ponchio et al. (2020) identify the twofold process of: a) selection of a location/region over the 3D model and b) generation of the explicit link of that spatial element with structured, semi-structured or unstructured data. When associated to structured and formalized knowledge, the annotation is *semantic* (i.e., allows the semantic enrichment of the digital replica of the object). As such, heterogeneous information produced during research and analysis on an architectural artifact can be sorted within an intelligible digital representation (De Luca, 2011).

However, the hierarchical organization of this added information varies according to the levels of semantic enrichment that are envisaged. The distinction between classes of annotations formerly identified by (Andrews et al., 2012) in the semantic web context can appropriately be adapted to the case of digital heritage models:

- Tags: keywords or plain terms linked to a given resource;
- Attributes: a characteristic of the resource, defined by a name and an associated value;
- Relations: interlinking of provided resources;
- Ontologies: the explicit specification of a shared conceptualization.

Tags represent the most trivial form of annotation, being small and clickable notes that can be put in specific positions of the model. Many on-line repositories of 3D models, e.g. Sketchfab (sketchfab.com/), 3DHop (vcg.isti.cnr.it/3d hop/) and Potree (potree.org/), are equipped with this type of annotation tools, to create simple linkable digital content.

Conversely, ontologies represent the most complete and comprehensive

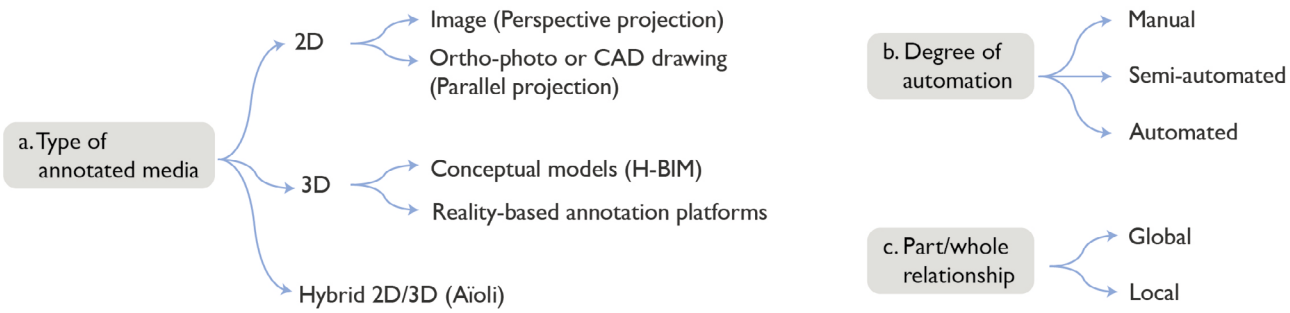


Figure 3-42.

Classification of semantic annotations over digital heritage models as proposed in Croce et al. (2020).

form of annotation, as they contain information structured according to a formal naming and a shared classification; however, since requiring the creation of shared conceptual models and vocabularies, they are the most challenging ones to implement.

Tags, attributes, relations and ontologies allow to spatialize the information, i.e., to directly link it to a 2D or 3D representation of the digital model, on different levels of data organization and management. In (Croce et al., 2020), we concisely reviewed existing methods and platforms for the semantic annotation of heritage models starting from survey data, proposing a distinction based on: type of annotated media, degree of automation and part/whole relationship (Figure 3-42).

The critical discussion of the several solutions proposed over time and the future research challenges that were addressed in the paper are presented in this section. In detail, most consolidated methods, as the traditional mapping of 2D media, are compared with more recent approaches making the most of 3D representations; for the latter, an in-depth regard is directed towards the presentation of existing H-BIM solutions and to the collaborative reality-based annotation platform Aïoli (aioli.cloud), developed by the MAP laboratory of the CNRS.

3.3.1. Classification based on part/whole relationship

Relying on the part/whole relationship, an annotation can be distinguished in:

- Global. In this case, the annotation concerns the heritage object in its entirety and describes its content. Although global annotations are more trivial, as they concern 2D or 3D objects as a whole, they can be applied to simple scenes and are therefore not adapted to complex heritage artifacts.
- Local. In this case, the description relates to restricted or localized portions of the object. The annotation is assigned to specific points, polylines or polygons identified over the 2D or 3D data (Figure 3-43); although being more precise,

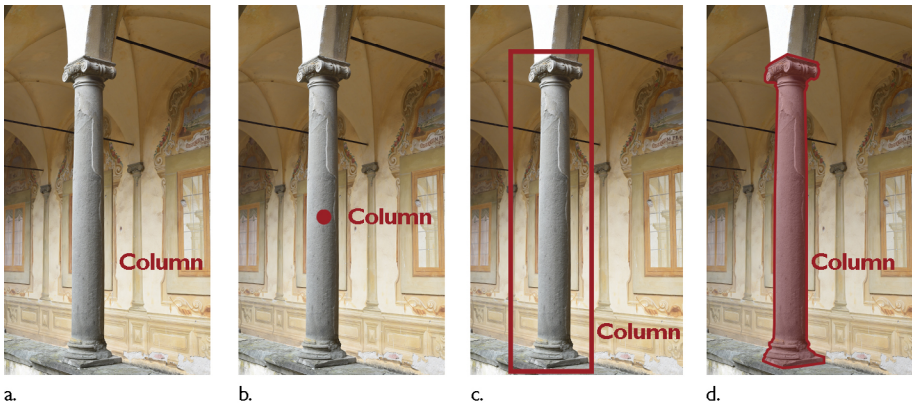


Figure 3-43.

Types of 2D image annotations: global (*a*), point-based (*b*), rectangular (bounding box, *c*), region (group of pixels, *d*). Adapted from (Abergel, 2020).

e.g.: Annotation of a column in the Grand-Ducal cloister, Pisa Charterhouse.

this mechanism may result time-consuming, and this can be a bottleneck when dealing with large amounts of data, as is the case for digital heritage. However, the annotation of limited portions of an object is a key requirement for information systems focused on the restoration and conservation of architectural assets, e.g., for the description of areas affected by localized damage or degradation phenomena.

3.3.2. Classification based on the degree of automation

Considering the degree of automation, annotations are classified as manual, semi-automated, automated.

The classification broadly proposed in the literature refers to the case of 2D media (Manuel et al., 2016), but it can readily be extended to 3D data as well:

- Manual annotations. Annotations of this type are traced manually by the user or operator, who defines their contours or hatches by free-form or polygonal selection tools. Although foreseeing a greater insight and control, as being driven in every step, manual methods appear time-consuming and strongly depend on the visual judgement and on the user's knowledge (Ponchio et al., 2020). Also, they may lead to too arbitrary and subjective choices.
- Semi-automated annotations. In this case, the selection of a region of interest is more automated; nevertheless, the end user performs manual operations over limited portions of the data to guide and control the annotation mechanism. Automated techniques will later provide the remaining annotations.
- Automated annotations. These methods include the automatic selection of regions of interest and their description, according

to the subsequent phases of segmentation, feature extraction, feature selection and classification (Pandya et al., 2014). Such annotations can only be created from a consistent number of data samples, so to suitably train computers to recognize information layers and descriptive attributes.

According to the degree of automation, (Ponchio et al., 2020) propose a similar classification identifying the classes of semantic segmentation and user-driven characterization, respectively. In the first case, the support is partitioned in semi-automated or automated way and the approach is focused more on how to produce the structural decomposition into components rather than the task of linking metadata or some additional knowledge-related information. Besides, user-driven characterization is more guided by the user insights or inputs and the annotation is associated to a tag with semantic data. However, such a distinction appears more labile as there exist many methods combining both semantic and user-dependent annotations and rather identified at the crossroads between manual and more automated methods.

In architectural heritage applications, the complexity of recognition and interpretation of digital data derived from the survey (e.g., in terms of identification of architectural typologies or detection of degradation or alteration patterns) necessarily implies the use of semi-automated approaches, foreseeing the intervention and control of an expert operator in many key parts of the documentation process.

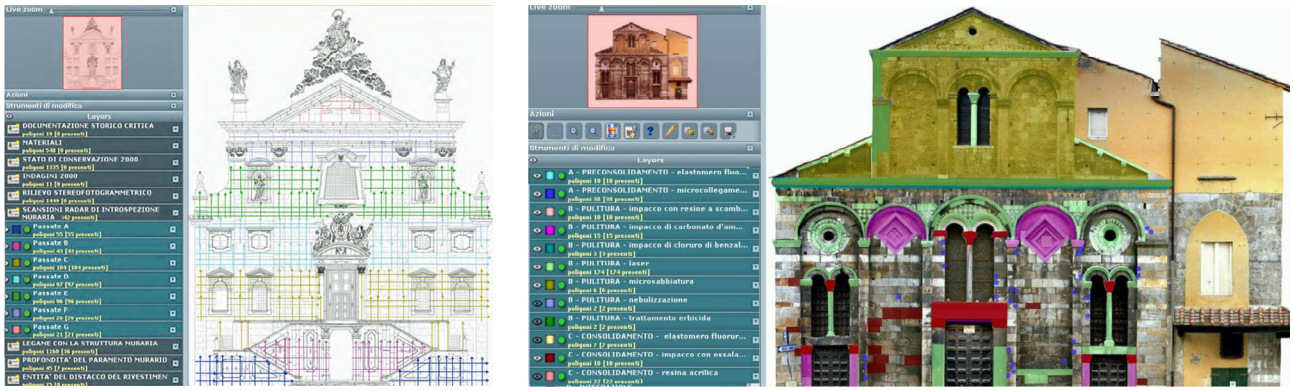
3.3.3. Classification based on type of annotated media

In Cultural Heritage, the annotation implies a thematic mapping process to link the graphic representation of an asset with knowledge-related information. Depending on the type of support used to execute such a thematic mapping, 2D approaches, 3D approaches and hybrid approaches can be classified.

3.3.3.1. 2D approaches.

Originally, the annotation procedure has been developed over 2D supports, for the identification of regions of interest over images (perspective projections), ortho-photos or CAD drawings (orthographic projections). For instance, the conventional approach to map degradation or alterations consisted in applying a certain pattern over a selected region, while clarifying the meaning in a legend.

This process has been transferred from formerly printed documents to the digital environment thanks to the introduction of the first 2D digital annotation platforms. The latter relied on the overlap of information



a. b.

Figure 3-44.

Interface of the SICaR web-based documentation system for the annotation of ortho-images. For further documentation on the SICaR interface, the reader can refer to the web page http://sicar.beniculturali.it

e.g.: Portion of the façade of Pisa Charterhouse and elevation of the Church of San Pietro in Vinculis in Pisa (Fabiani et al., 2016; sicar.beniculturali.it)

over a same bi-dimensional spreadsheet and somewhat transferred the principles of multi-layer and multi-temporal information management typical of Geographic Information Systems (GIS) to the cultural heritage sector. When dealing with raster data, as is the case for images and orthophotos, the annotations concerned single pixels, whereas in the case of CAD drawings, the annotations were vector-type files (Ponchio et al., 2020).

Among such platforms, it is worth mentioning SICaR, a web platform for the documentation of site activities promoted by the Italian Ministry of Culture since 2003 (sicar.beniculturali.it) and presented in (Fabiani et al., 2016).

SICaR was created with the aim of conveying the technical information produced in the framework of restoration works around a heritage object (concerning, e.g., methods, materials and tools used in the intervention, chemical, physical and petrographic analyses), through the graphic support of ortho-photos. A portion of the façade of Pisa Charterhouse was documented via the SICaR as well (Figure 3-44).

However, the reference to 2D media alone implied the need to clarify the meaning of annotations in a legend, and therefore the choice of a graphic code for the association of information with metric data. The communication of the semantic information is therefore not straightforward; in addition, the production and annotation of several diversified graphic drawings is required. The multilayer management is therefore possible on the single orthophoto, but not yet on the whole heritage element.

Although being conceptually referred to a same physical entity -an architectural element or parts of it, from the spatial representation point of view the represented 2D data have no direct connection one another. The problem is described in literature as a *spatial referencing gap* (Messaoudi et al., 2018). This can only be solved by moving to a 3D representation (Croce et al., 2020).

3.3.3.2. 3D approaches.

Besides 2D media, 3D models can also be semantically enriched by the appropriate insertion of annotations. Global and local annotations are possible even in this case; however, the annotations concern no longer the projection of the element, but its real extension and representation in the three dimensions. However, the selection of portions or regions of interest within 3D models is not trivial, and this raises the problems of effectiveness and precision of different available methods and tools for the identification of the critical points of the annotation (Abergel, 2020).

Over time, the availability of different 3D representation methods has also led to the emergence of annotation systems based on different supports. In detail, 3D annotation systems apply both to (Figure 3-45):

- Reality-based models, e.g., point clouds and meshes directly derived from 3D surveying procedures.
- Conceptually modelled mockups, as the ones reconstructed by leveraging existing H-BIM platforms.

For the class of reality-based models, several web platforms - as the already mentioned

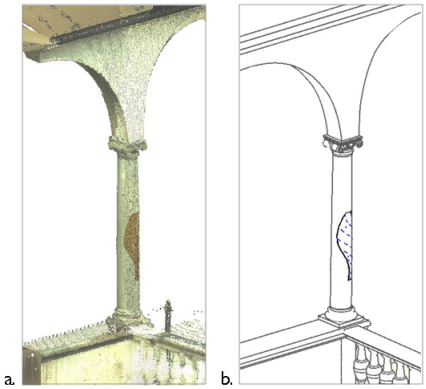


Figure 3-45.

Examples of annotation on a reality-based model (point cloud, *a*) and on a conceptually modelled mockup (H-BIM, *b*).

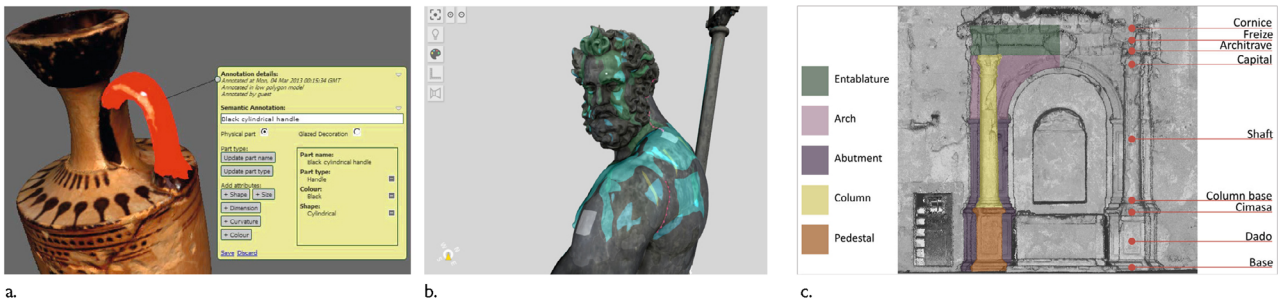
e.g.: Annotation of degradation phenomena over a column of the Grand-Ducal cloister, Pisa Charterhouse

Figure 3-46.

Simple tag annotation of a reality-based model in Sketchfab (sketchfab.com).

e.g.: Cappella del Santo Rosario, Pisa Charterhouse.





a. b. c.

Figure 3-47.

3D annotation systems: interface of 3DSA (Yu et al., 2013; *a*), Neptune IS (Apollonio et al., 2018; *b*) and CulTo (Garozzo et al., 2017; *c*).

Sketchfab, Potree and 3DHop or the applications described in (Soler et al., 2013) and (Attene et al., 2009)- provide user-friendly tools for the selection of regions and subsequent insertion of annotations. In this case, the annotations consist in simple tags (Figure 3-46). The high level of user-friendliness guarantees the use of these platforms even by less-experienced operators; nevertheless, this undeniable advantage is also the limit of such existing tools.

In the Cultural Heritage domain, the realization of inhouse platforms for the creation of more semantic-aware 3D representations has thus become commonplace (Havemann et al, 2009; Manferdini and Remondino, 2010; Serna et al., 2012; Yu et al, 2013). The following list provides a summary of relevant examples of this kind (Croce et al., 2020).

In 2013, Yu et al. presented the 3D Semantic Annotation System (3DSA, Figure 3-47a), a web-based application enabling the creation of points or polygon shapes onto mesh models and supporting the link between the performed selection and a knowledge base. Object meta data were connected to domain-specific ontologies and crowd-sourced annotations. However, the system worked only with single resolution models, and with regions defined by best overlapping.

More specifically in the architectural heritage domain, Apollonio et al. (2018) illustrate Neptune IS (Figure 3-47b, Figure 3-48) an information system allowing to document the restoration activities for the Neptune's Fountain in Bologna and built on the web-viewer 3DHop. The system is 3D-centered, as the reality-based model of the heritage object represents the pivotal element where to store and organize new and previous restoration data. The mesh model of the fountain is segmented according to a hierarchical structure, the latter defined by the restorers to organize the different phases of their work. The user can perform points, polylines and polygonal-type annotations on a chosen area of this segmented model. However, the annotation is in this case limited to the single element that is selected within the hierarchical subdivision (Figure 3-48).

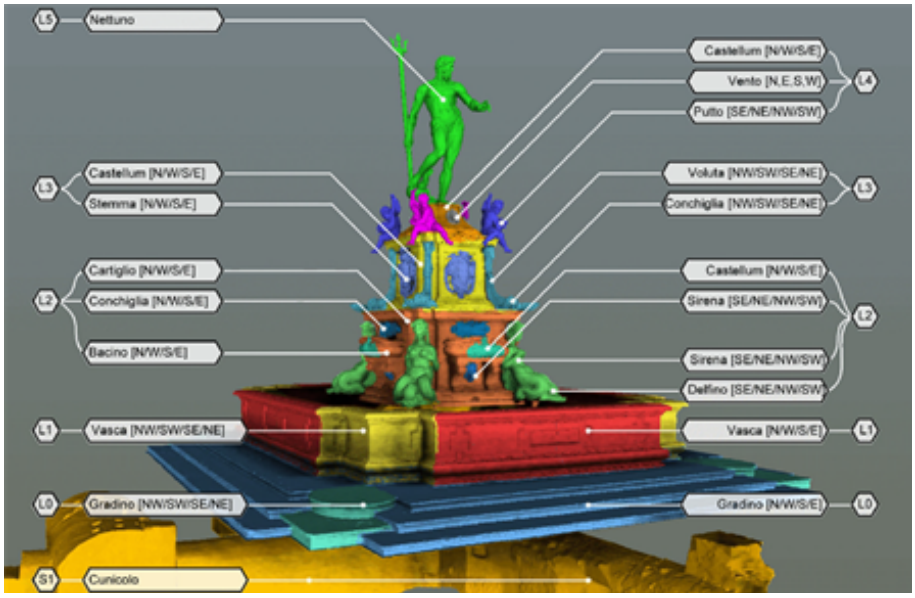


Figure 3-48.

Hierarchical segmentation in levels of Neptune's Fountain to guide the subsequent annotation over reduced portions (Apollonio et al., 2018).

When developing the *Cultural Heritage Tool* based on ontology (CulTO), Garozzo et al. (2017) aimed at supporting heritage experts in the documentation of historical religious buildings (Figure 3 47.). The tool is connected with ontology-driven language, relying on the CIDOC-CRM Conceptual Reference Model (cidoc-crm.org/): the 3D building represents a *Physical Object* that can be segmented into specific sub-structures, and a specific vocabulary is provided for the heritage of churches. A visual ontology supports the annotation phase; however, annotations can only be performed in bi-dimensional images and the connection with the 3D model only remains at a logical (non-informatic or machine-readable) level.

A similar approach, consisting of semantic and thesauri-enhanced searches, was proposed in the Arches platform (*archesproject.org*), an annotation platform designed to link graph structures and spatial data of the historical heritage.

Besides reality-based annotation platforms, another kind of annotation procedure can be identified, as closely related to the spread of Building Information Modeling (BIM) techniques and their application to heritage structures (Bruno et al., 2018; López et al., 2018). It relies on conceptual representations of existing buildings, that are reconstructed, starting from survey data, by making use of reverse engineering processes (Brumana et al., 2020).

Such representations are enriched with non-geometric information; indeed, organized data management is a prerogative of such BIM-based systems. As a key aspect of this work, the semantic enrichment via H-BIM platforms is further addressed in Section 3.3.6.

In any case, it can already be noted that reality-based and H-BIM approaches appear limited to static and non-evolutionary

representations of a heritage object; besides, they are based each time on a single type of representation. To give an example, Apollonio et al. (2018) already recognized a limit of their platform in the impossibility of annotating different temporal states:

'Some of the elements of the fountain changed in shape or color during the restoration project or were removed after the 3D acquisition campaign. In this case, the solution would be […] to provide a "temporal" visualization (interactive switch among different time-tagged models).'

(Apollonio et al., 2018)

Or again, these models are always restricted to the single mesh, point cloud or conceptual model, built at a certain time and with a certain level of detail. As soon as the representation changes (in terms of representation types, resolution, temporal evolution etc.), the annotated data may be lost. This results in the complexity of transferring the annotation-hence, the semantic information- from one 3D model to another. A first attempt to solve this problem was addressed by (Scalas et al., 2017), who developed a method for the annotation transfer between meshes at different resolution. Besides distortion, degeneracy and wrong projection errors already identified by the authors, a limit of their study is that, apart from mesh models, it does not take into account any other existing representation types.

3.3.3.3. Hybrid 2D/3D annotation approaches.

In the case of hybrid approaches, one medium, be it 2D or 3D, is not preferred to the detriment of the other. Rather, the direction is that of a combined use of the data, so as not to lose significant information or repeat semantic annotation operations in the shift from one digital media to another. In the realization of information systems based on

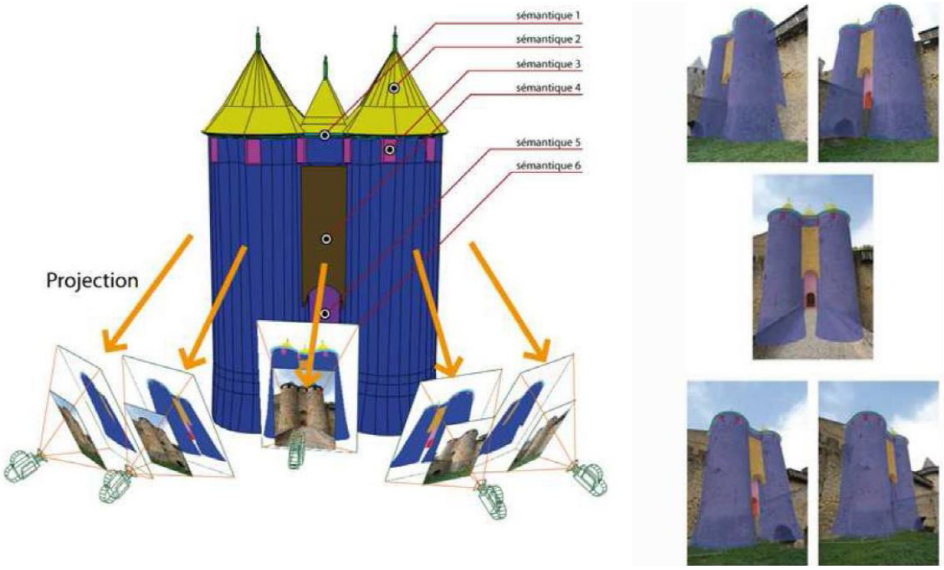


Figure 3-49.

Projection of the annotations from the 3D model to the images (Busayarat et al., 2010).

hybrid approaches of this kind, it is fundamental to define the relative position that the 2D images or drawings occupy with respect to 3D models (point clouds, meshes). In other words, the spatialization of the data is fundamental for the reconstruction of the 2D/3D projective relations that allow the transfer of annotations between the different available media. Indeed, hybrid approaches start from the assumption that it is much easier to select the annotation on a 2D media rather than on the 3D model. They thus allow the performance of an annotation over images or ortho-graphic drawings and then exploit projective relationships to transfer the annotation from 2D to 3D space (Croce et al., 2020).

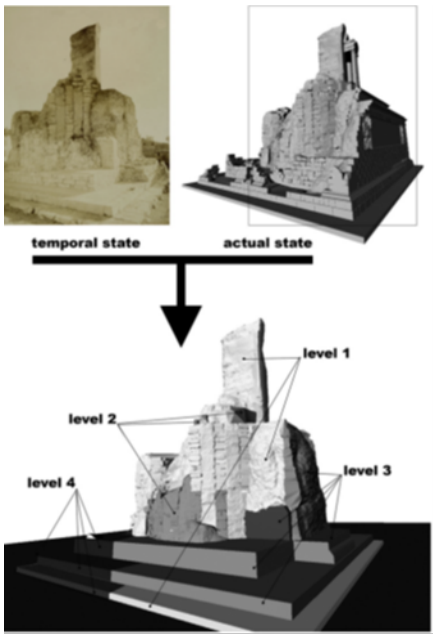


Figure 3-50.

Visual coding of uncertainty applied to the representation of a temporal state (De Luca et al., 2011).

The first experiments in defining hybrid approaches were carried out as part of the NUBES web-based platform development project (Figures 3-49, 3-51): after reconstructing a digital model from survey data, Busayarat (2010) broke down the 3D representation of the building, so to isolate characteristic elements describing its morphology according to architectural vocabularies (capital, base, window, columns, doors, etc.). Then, by establishing a

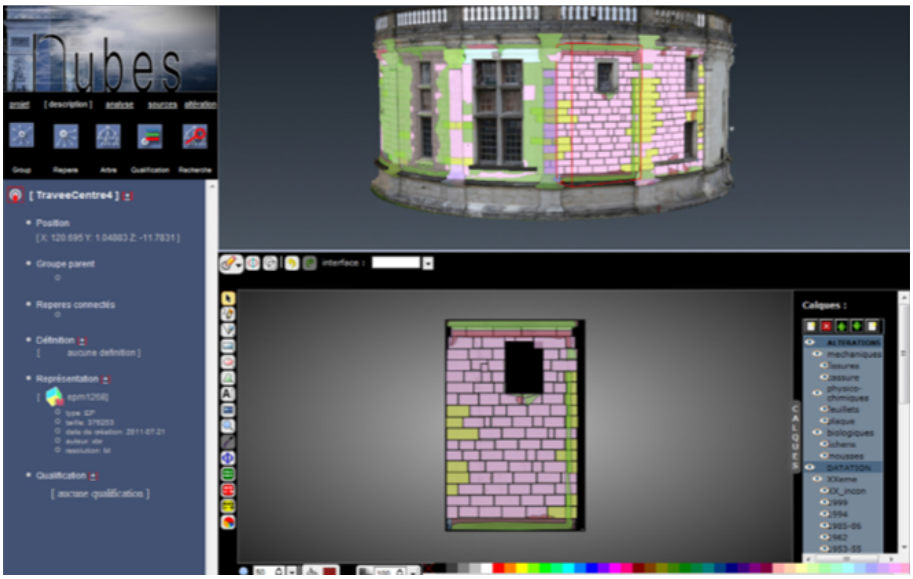


Figure 3-51.

The NUBES web-based platform interface. Transfer of information from 2D mapping to the 3D model (Stefani et al., 2014).

correlation between the model and the 2D images and by reconstructing the projective relations, information was propagated from the 3D model to the set of related iconographic resources. The approach thus moved from 3D to 2D.

De Luca et al. (2011) further developed the platform introducing the temporal referencing of iconographic sources and 3D models, even considering each time a level of reliability of the information and starting from the comparison of 3D models of successive time states. The reference to a historical graph notation enabled the spatio-temporal analysis, as well as the graphical visualization of changes in a building over time (Figure 3-50).

When illustrating the case study of the Chambord Castle, Stefani et al. (2014) demonstrated the use of the NUBES platform to document and monitor the conservation state of a building: the informative system relied on UV mapping procedures to appropriately transfer 2D/3D information related, e.g., to stone degradation, dating and material. UV maps were generated from the 3D model, by a suitable planar, cubic or spherical projection, and annotated with decay or material information. Then, by a proper re-projection matrix, annotations were projected back to the 3D model (Figure 3-51).

Although enabling the link of spatial, visual and temporal attributes, this system was limited to the use of pre-segmented mesh models. Moreover, it required the user to manually reconstruct the position of the images in relation to the 3D model.

Starting from the work by Manuel et al. (2016), a more innovative system was introduced, leveraging a set of images calibrated and spatially oriented on a 3D point cloud.

The method proposed for the point cloud indexation gave birth to the collaborative reality-based annotation platform Aïoli (aioli.cloud/) developed by the MAP laboratory of the French National Center for Scientific Research (CNRS).

Being the fundamental reference for many essential applications of this thesis, the Aïoli platform is provided an in-depth look in the following section.

3.3.4. The reality-based annotation platform Aïoli

The operating principle starts from the consideration that, for an end user, it is more convenient and immediate to perform annotations on images than directly on the 3D (Figure 3-52, Figure 3-53).

By retrieving the projective relations needed to create a photogrammetric model from an initial set of images, the annotation is automatically transferred from a single photo to the 3D and, in re-projection, to all

other images where the same annotation is visible. This is done by storing the correspondences that link pixels of the images to the coordinates of the points in the cloud.

The operator chooses the image that better displays the region to be annotated; then, by use of drawing tools, the selection of this region of interest is performed.

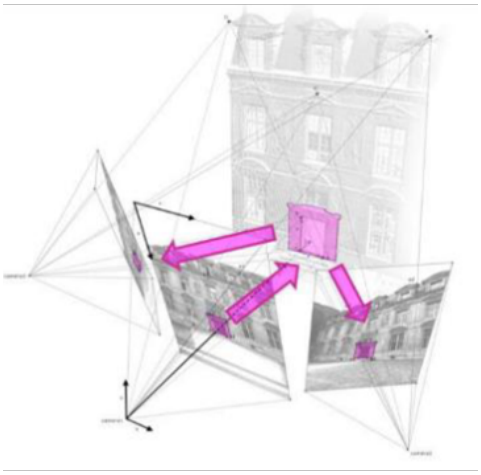


Figure 3-52.

Schema of the approach proposed by Manuel (2016) for the 2D/3D propagation of annotations (Manuel, 2016).

By the 2D/3D correspondences, the selection is transferred to the 3D model, hence projected back to all other images where it appears. In doing so, the semantic annotation is not only transferred from the initial image to the 3D representation, but it is also propagated to all other photographs where it is shown.

The continuous projective relationship created enables moving from 2D to 3D and back from 3D to 2D.

The platform's software development environment includes (Abergel et al., 2021):

- A database, where application data (users accounts, projects, annotations etc.) are stored;
- A processing server hosting the TACO engine (Pamart et al., 2019, already mentioned in Paragraph 3.1.2.3) to calculate relative poses of 2D images within a photogrammetric project;
- A web server, managing project data and communicating with the processing server;
- A thin client, integrating web 2D and 3D viewers built on the



Figure 3-53.

Aïoli operating principle: the annotation is transferred from a input image to the 3D and, in reprojection, to all other images where the same annotation is visible (Abergel, 2020).

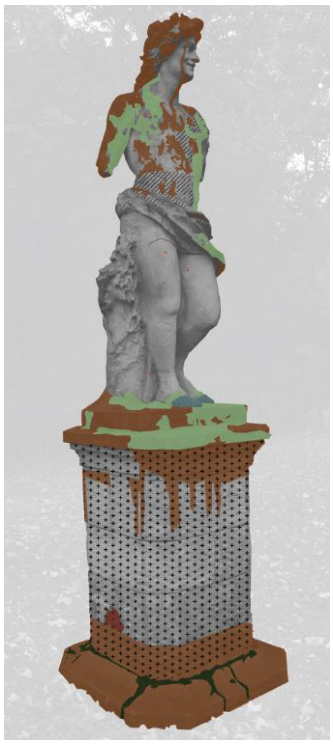


Figure 3-54.

Alterations of the statue of the Magalone garden in Marseille mapped onto the 3D model. (Roussel et al., 2019).

JavaScript 3D libraries Three.JS and Potree, enabling the display of image sets or point clouds as well as the relative positions of images with respect to the 3D model.

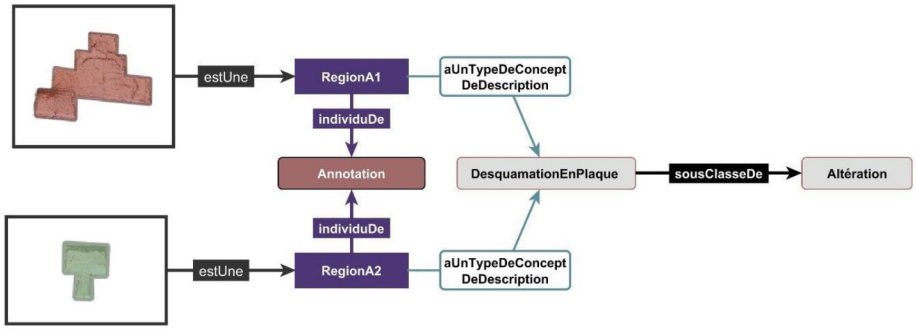
For further details on the Aïoli computer base, as well as on the software implementation by use of the TACO algorithm, the reader can refer to (Abergel et al., 2021; Pamart et al., 2020). In any case, it must be noted that documentation through this platform is intended to be collaborative, as annotations are made and further shared via cloud computing services. The end user can structure annotations in layers or groups of layers and associate them with geometric (automatically calculated) and semantic (user-defined) descriptors.

Roussel et al. (2019) provided a demonstration of the use of Aïoli for the diagnosis of sculptural elements, documenting the evolution of degradation patterns of the Autumn Statue of the Magalone garden in Marseille (Figure 3-54).

In a further extension of the approach, Messaoudi et al. (2018) proved how annotations in Aïoli can be linked and supported by a domain ontology model (Figure 3-55). Referring to the case of masonry structures, the study illustrated the correspondence between the annotation layers and the thematic and lexical classes defined by cultural heritage ontologies. From this link, spatial relationships were studied, and data were cross-referenced to analyze occurrences (e.g., the relation between degradation phenomena, materials, etc.) as well as to characterize (by spatial and morphological attributes) the domain

Figure 3-55.

Part of the ontological model developed by Messaoudi (2018): an annotation is linked to thematical classes defined by existing ontologies. (Messaoudi, 2018)



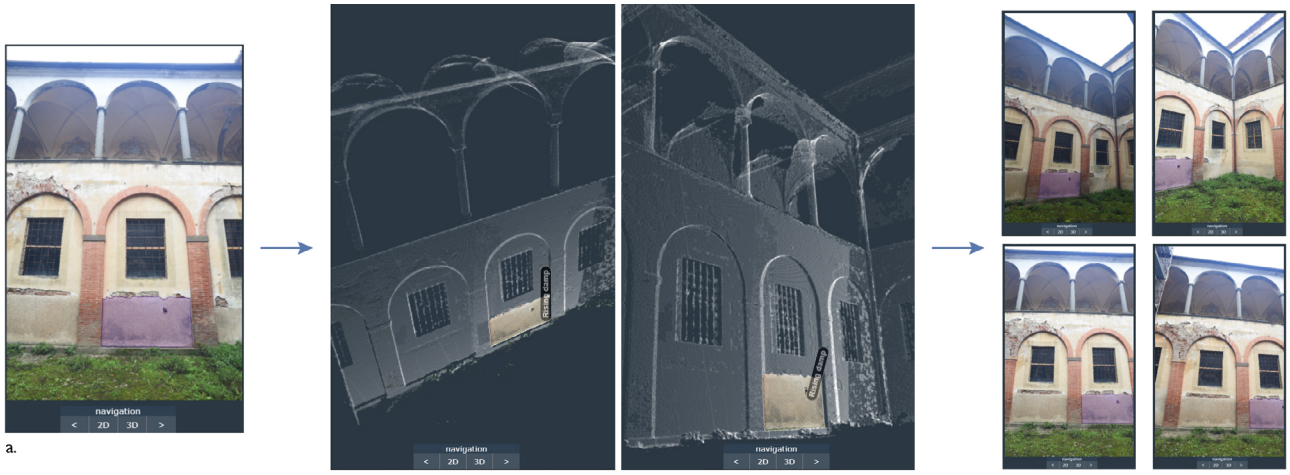


Figure 3-56.

Processing steps via Aïoli: (a) selection of the annotation on the first image, (b) detection of 3D points corresponding to the selection and (c) propagation of the annotation on the other images of the set.

e.g.: Grand-Ducal cloister, Pisa Charterhouse

ontology concepts. With the same objective, Roussel et al. (2019) envisaged an integration of the platform with reference conceptual models to improve the documentation processes of sculptural heritage. Later on, Abergel et al. (2021) implemented Aïoli with an augmented reality environment, allowing the on-site interaction of the digital annotated model with the real object. The work relied in this case on the alignment between real and virtual environments to display real-time annotations on screen when on site. In an in-depth study aimed at the 3D annotation of rock art, the authors further explored the addition of decorrelation stretching and edge detection filters (Paragraph 3.2.1) to ease the on-site observation of imperceptible phenomena.

The promising advances introduced in recent years through the development of this platform point in the direction of 2D and 3D information transfer, although the core of the annotation process still relies on photogrammetric models. The point cloud, at the center of the digital information system (Figure 3-56, Figure 3-57), is used to federate knowledge around the object, thus limiting the effect of dispersion usually found in the heritage field.

However, some methodological limitations of the platform may still be identified: the 3D representation exclusively relies on point-based models, and therefore there is no provision for linking the annotations with mesh or even parametric mockups, as those that are constructed in (H)BIM environment.

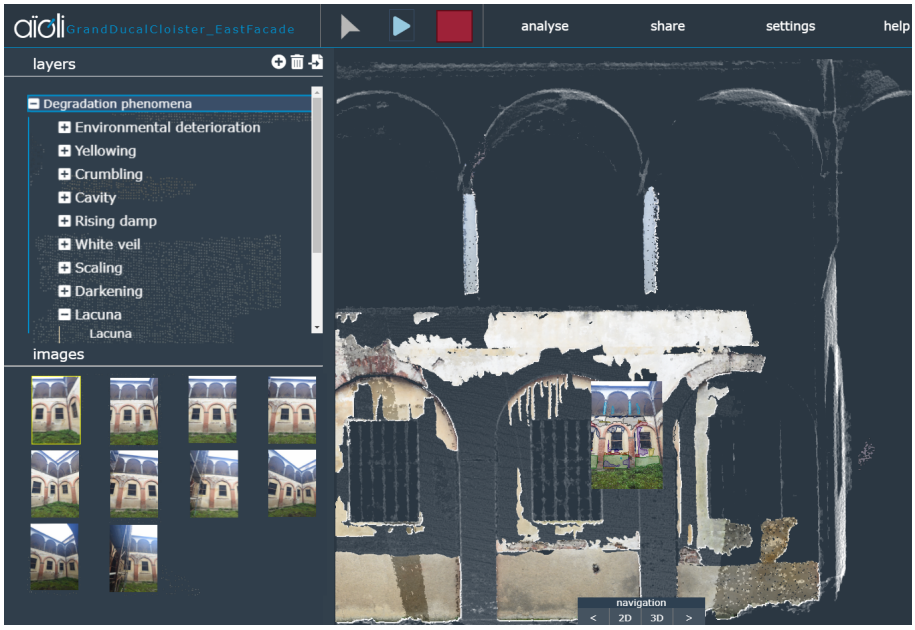
Moreover, the annotation over a single image may be challenging when the element to be annotated is not extended to the whole image or even in the case of complex architectural forms (e.g., if whole vaults or columns surfaces are to be annotated).

Finally, it should be noted that the indexing method proposed in Aïoli is

Figure 3-57.

Interface of the Aïoli platform: projection of the annotations in 2D/3D.

e.g.: Grand-Ducal cloister, Pisa Charterhouse



so far only developed in the 2D-> 3D direction, but this does not exclude the transfer of information in the opposite direction 3D-> 2D, intended as the possibility to perform manual annotations over the point cloud and no longer on images.

3.3.5. Heritage-Building Information Modeling

The BIM technology was initially conceived for new buildings. Its origin is traced back to the studies conducted, since 1974, by Professor Charles 'Chuck' Eastman and his research group at the Carnegie-Mellon University in Pittsburgh (USA). They initially introduced the theory of *Building Description Systems and Building Product Modeling*, while pioneering those tool developments that later evolved into the *Building Information Modeling* concept (Eastman, 1974; Zhang et al., 2013).

Since these early studies, applications of BIM in the AEC domain have expanded, and the widespread use of such methodology has led the National Institute of Building Sciences to the provision of a more comprehensive definition:

'A building Information Model (BIM) is a digital representation of physical and functional characteristics of a facility. As such it serves as a shared knowledge resource for information about a facility forming a reliable basis for decisions during its life-cycle from inception onward.'

(NIBMS, 2007)

BIM techniques have indeed been originally developed to organize, around the 3D geometric representation of a new building, the complex wealth of knowledge related to textual, thematic and descriptive data.

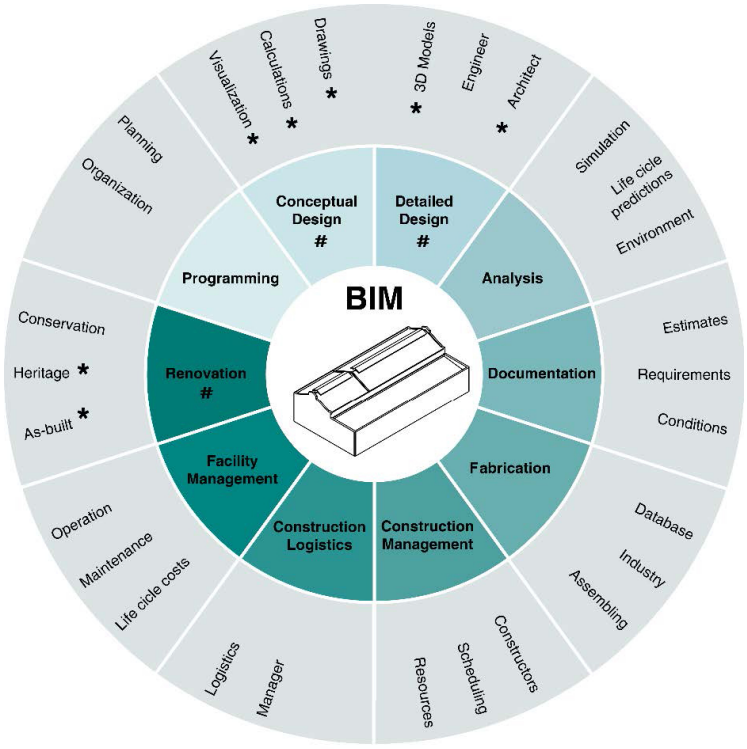


Figure 3-58.

BIM as a process. Diagram of a Building Information Modeling working process (Rocha et al., 2020).

In the same representation environment, geometric and semantic features are brought together, and the information is intended to be stored, analyzed and accessed in a coordinated way (Bruno et al., 2018).

The recourse to BIM modeling can thus be crucial for decision making on buildings, as pointed out in the definition of BIM provided by the International Standard ISO 19650 (2019):

'The use of a shared digital representation of a built asset to facilitate design, construction and operation processes to form a reliable basis for decisions.'

ISO 19650:2019)

As already outlined in Section 2.5, BIM can be referred both to the general process of information update, retrieval and amendment (BIM as a process) than to the specific geometric representation where such information is added (BIM as a model).

The first concept is the broadest one, as BIM is rather understood as a process allowing to create, manage, share and derive the information related to the object of study, by referring to a same 3D digital environment (Figure 3-58). In this sense, the BIM definition includes at the same time, as key elements, a 3D object-oriented model, a relational database containing the information on the asset and the data exchange flow between the actors involved (Eastman et al., 2008).

Considering the latter aspect, the communication between the different professionals working on an asset may take place directly on the digital mockup, and the information flow may cover various moments of the building's life cycle (Figure 3-58). As for the case of new constructions, reference is generally made respectively to the phases of pre-planning, design, construction and maintenance (Hong et al., 2015; Volk et al., 2014), interconnected onto the collaborative digital model (López et al., 2018).

According to Arayici (2008), behind the connection between geometry, integrated information, semantic richness and lifecycle support, a relevant attribute of BIM is identified in the presence of *'comprehensible and extensible object properties that expand the meaning of each represented object'*. The building is intended to be structured according to its typological or constructive components, e.g., walls, doors, windows etc., being connected one another by specific relationships and each one being defined by a suitable set of parameters (*parametric modeling*). In addition to containing significant geometric information, such elements are enriched with analytical data describing different physical, material and constructive properties (Pocobelli et al., 2018a; Volk et al., 2014).

All the available information is thus connected to each 3D component of the model, thanks to the use of such parameters and properties made available in specific object libraries (Pocobelli et al., 2018a; López et al., 2018; Logothetis et al., 2015).

In bringing together the geometric, physical and functional characteristics of the building organism around a consistent digital representation, reference will rather be made to the idea of BIM as a model: the digital replica is obtained by reconstructing real objects (e.g., walls, doors, windows etc.), connected to each other by specific relationships and each of them defined by a wide series of properties. These are parametric objects (Volk et al., 2014), i.e., model components that can be modified or updated when needed (Pocobelli et al., 2018a; Fai and Sydor, 2013; Murphy et al., 2013; Oreni et al., 2013) by definition of suitable parameters and descriptors, enriching the digital model with analytical information related, e.g., to physical, material, constructive attributes (HISTORIC England, 2017; Figure 3-59).

Each element is thus connected to descriptive multi-source information (Eastman, 2008). The perspective is therefore that of an *object-oriented modeling*, which places the digital mockup at the center of the information process (De Luca et al., 2006).

The BIM key features identified by (Arayici, 2008) suitably apply to the case of new structures. Indeed, current BIM software as Autodesk Revit, Archicad (Graphisoft), Tekla Structures and Bentley system (López et al., 2017), were originally conceived to describe the lifecycle of a new building from conception to demolition.

Information protocols were established to ease the construction of *object-oriented models* within such kind of platforms. The open global standard Industry Foundation Classes (IFC), promoted and distributed by buildingSMART, was built to achieve this task (EN ISO 16739:2013; buildingSMART International): each element of a new building has a standardized and open data structure, so to guarantee information exchange between different BIM authoring software applications for new constructions (Pocobelli et al., 2018a; Rodríguez-Moreno et al., 2018).

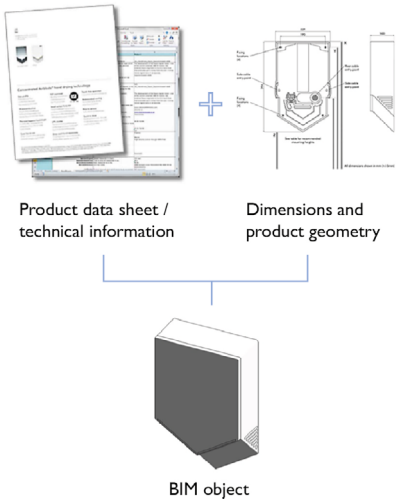


Figure 3-59.

BIM objects as containing both geometry (2D or 3D) and associated information. Adapted from (HISTORIC ENGLAND, 2017).

The concepts of Level of Detail (LoD)-related to the degree of certainty of graphical elements of the model-, and of Level of Information (LoI) -related to non-graphical information-, or even the notion of BIM maturity levels (BIM Task Group, 2011) have been introduced for new constructions, to adapt representation styles and uniform the levels of accuracy of the digital model (Rocha et al., 2020).

As for new constructions, the BIM modeling starts from the very first preplanning, design and construction phases (Hong et al., 2015) and includes part of the project delivery.

For the case of existing buildings, however, the conditions for the construction of a BIM-based representation markedly differ: first of all, a model of the initial state of construction of the building is generally not-available (Volk et al., 2014), thus information must be derived through data deduction and interpretation processes.

In the following sections, it is stressed how the application of BIM to heritage objects raises several issues with respect to the streamlined and consolidated workflow that applies for new constructions.

In detail, the difficulties related to the set-up of BIM models starting from 3D surveying data are introduced (Paragraph 3.3.5.1); then, traditional and more automated approaches proposed over time for the solution of the so-called Scan-to-BIM problem are outlined and compared (Paragraphs 3.3.5.2, 3.3.5.3). Finally, existing methods for annotating Heritage Building Information Models (H-BIM) are discussed (Paragraph 3.3.5.4).

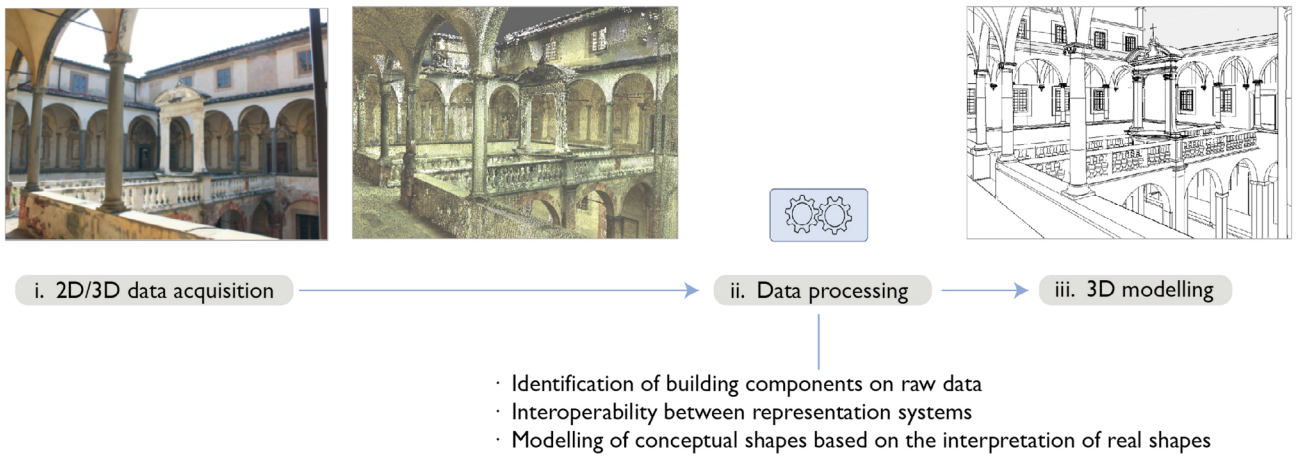


Figure 3-60.
Steps of the Scan-to-BIM workflow as identified by Hichri et al. (2013).
e.g.: Grand-Ducal cloister, Pisa Charterhouse.

3.3.5.1. BIM for historic or heritage buildings

As for built heritage, the production of BIM models implies the reference to existing, reality-based data, as derived from surveying. The earliest contributions to the definition of tools and approaches to the documentation of the built heritage by means of HBIM applications are owed to Maurice Murphy (Murphy et al., 2009), and Conor Dore (Dore and Murphy, 2013), from the Dublin Institute of Technology. They first theorized a system enabling to:

> *'[…] apply the consolidated BIM knowledge acquired for new constructions to the case of existing heritage objects and sites, to help the documentation and management activities in view of conservation work, retrofitting, renovations and buildings analysis.'*
>
> *(Dore and Murphy, 2013)*

Since then, the scientific literature on HBIM has expanded, aiming at illustrating how geometrical data can be linked to: architectural grammar and styles (Miceli et al., 2020; Oreni et al., 2013; Arayici, 2008); material characterization (Fai and Sydor, 2013); degradation patterns (Bacci et al., 2019); façade interventions and historical layers (Angulo-Fornos and Castellano-Román, 2020); structural damage and FEM analysis (Brumana et al., 2020; Pepe et al., 2020; Barazzetti et al., 2015); data collection and simulation of environmental parameters (Pocobelli et al., 2018b); archival photographs (Martinez Espejo Zaragoza et al., 2021) or text documents (Brumana et al., 2020; Bruno et al., 2019). Simulations on the buildings' temporal states and evolutions can also be included in BIM platforms (Brumana et al., 2013).

To further explore HBIM applicative fields, the reader can refer to the wide literature reviews that exist on the subject (López et al., 2018; Pocobelli et al., 2018a; Logothetis et al., 2015; Pătrăucean et al., 2015; Volk et al., 2014; Hichri et al., 2013).

In detail, the work by (Hichri et al., 2013) emphasized that H-BIM techniques require the transition from the real -existing- condition of the object to the modeling environment. This shift from an 'as-built' condition (registration of a building after construction) to the 'as-is' representation (registration of the current condition of a building) necessarily implies the reference to surveying data and the application of reverse engineering techniques.

If the HBIM model is reconstructed based on a 3D scan, formerly acquired via TLS or photogrammetry, the process is called Scan-to-BIM. As pointed out by (Dore and Murphy, 2013), in this case:

'Interactive objects representing architectural elements are constructed from historic data. These elements (including detail behind the scan surface) are accurately mapped onto a point cloud or image-based survey'

(Dore and Murphy, 2013)

The system is articulated in three main steps: i) data acquisition, ii) processing of survey data and iii) 3D modelling (Figure 3-60).

The processing phase, since involving interpretive steps and deductive reasoning, represents the bottom line of current research: indeed, the H-BIM modeling from 3D surveying data requires to properly recognize and characterize the elements composing the real object, as well as segmenting them and assigning them to object categories (Macher et al., 2017; Hichri et al., 2013).

In other words, architectural elements are not only to be represented in their geometry, but also with their corresponding semantic attributes (Rocha et al., 2020). Moreover, the relationship between components must be stressed (Tang et al., 2010).

The main limits of the Scan-to-BIM processing workflow are thus widely identified in the following:

- Difficulties in representing the variety of complex irregular shapes and architectural details characterizing heritage objects (Yang et al., 2018; Bruno et al., 2018; Dore and Murphy, 2013), which necessarily entail classification, hierarchical organization and simplification assumptions (Angulo-Fornos and Castellano-Román, 2020; Macher et al., 2017);
- Measurement uncertainties (Macher et al., 2017), as the original data may include occlusions (Previtali et al., 2014) or non-visible objects (Pătrăucean et al., 2015);
- Absence of pre-defined resourceful libraries of parametric objects (López et al., 2018) and lack of existing standard if compared to new constructions (Rocha et al., 2020; Dore and Murphy, 2013);
- High conversion effort (Dore and Murphy, 2013), since most

Figure 3-61.

Instantiation of a capital by direct reconstruction over the raw point cloud.

e.g.: Grand-Ducal cloister, Pisa Charterhouse (Croce et al., 2019; Grassi, 2019)

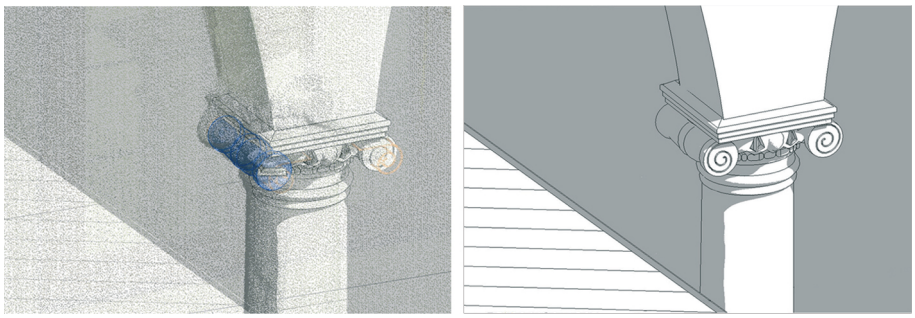
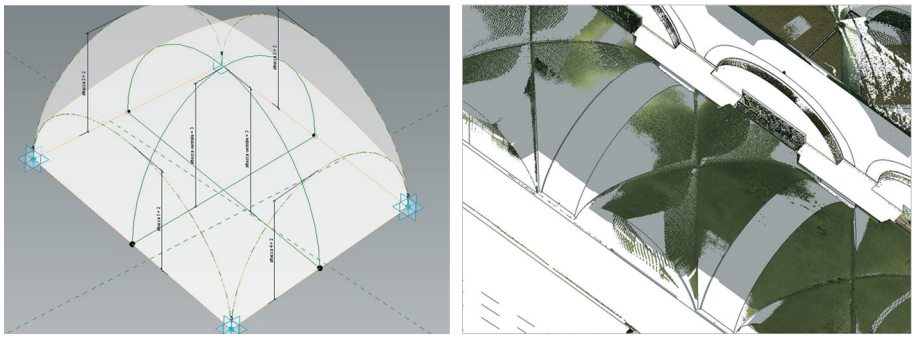


Figure 3-62.

Instantiation of vaults by manual reconstruction methods.

e.g.: Grand-Ducal cloister, Pisa Charterhouse (Croce et al., 2019; Grassi, 2019)



BIM software for new buildings offer tools for the construction of regular and standardized objects and the free-form geometry modeling functions available within are limited (Bruno and Roncella, 2019; Yang et al., 2018; Tommasi et al., 2016; Barazzetti et al., 2015).

As a direct consequence of such limitations, (Bruno et al., 2019) further mention the lack of clear, shared regulatory references and guidelines for H-BIM artefacts and the lack of knowledge and training in H-BIM technology of most professionals.

To overcome such constraints, semi-automated methods have been presented alongside more consolidated Scan-to-BIM workflows.

3.3.5.2. Manual Scan-to-BIM methods.

Most traditional approaches to the Scan-to-BIM are manual, i.e., they involve the reconstruction of components starting from a point cloud, by means of visual recognition and subsequent manual tracing of geometries (Figure 3-61, Figure 3-62).

Formerly introduced by (Dore and Murphy, 2017; Murphy et al., 2013, 2009), manual methods appear today widely consolidated, as proven by the extensive literature reviews presented by (Logothetis et al., 2015; Volk et al., 2014; Tang et al., 2010), and more recently by (López et al., 2018; Pocobelli et al., 2018a). However, such approaches have

resulted in overall time-consuming and laborious processes (Macher et al., 2017; Volk et al., 2014), in which the operator is asked to manually identify, isolate and then reconstruct each class of building elements.

This entails a considerable amount of time and resources, besides implying the risk of making too subjective choices (Xiong et al., 2013).

For these reasons, beyond the more traditional approaches, the research has dwelt on the development of more automated approaches, so to reduce the human involvement and speed up the Scan-to-BIM workflow. Therefore, automation of the process currently appears as a very active research area.

3.3.5.3. Semi-automated Scan-to-BIM methods.

The fundamental problems in the definition of semi-automated methods are the automatic recognition and the labeling of data points or segments with a named object or object class (e.g., windows, columns, walls, roofs etc.), starting from the raw point cloud data (Dore and Murphy, 2017; Jung et al., 2014; Xiong et al., 2013; Zhang et al., 2013; Tang et al., 2010). Existing methods can be distinguished according to the solution identified over time to these issues.

- Primitive fitting methods.

They operate on the recognition of simple geometric elements, such as planes, cylinders (Figure 3-63), cones and spheres (Schnabel, 2009). The robust estimation of primitive parameters takes place via primitive fitting algorithms, among which it is worth recalling the Random Sample and Consensus (RANSAC; Fischler and Bolles, 1981) and the Hough Transform (Hough, 1962); a comparative study of these two algorithms was previously proposed by Tarsha-Kurdi et al. (2007).

Primitive fitting methods are to date limited to indoor applications (Thomson and Boehm, 2015; Jung et al., 2014; Previtali et al., 2014; Zhang et al., 2013), for the recognition and subsequent reconstruction of mainly planar elements, as floors and walls, or at most for the detection of openings (Macher et al., 2017; Hong



Figure 3-63.

Reconstruction of a column shaft by primitive fitting of a cylinder shape.

e.g.: Grand-Ducal cloister, Pisa Charterhouse.

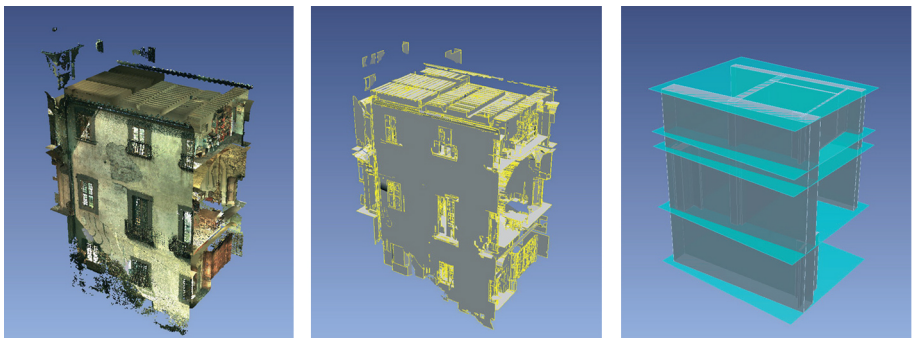


Figure 3-64.

Primitive fitting of perimetral walls and slabs by EdgeWise Building. Raw point cloud (*a*); fit planes (*b*); BIM objects (*c*).

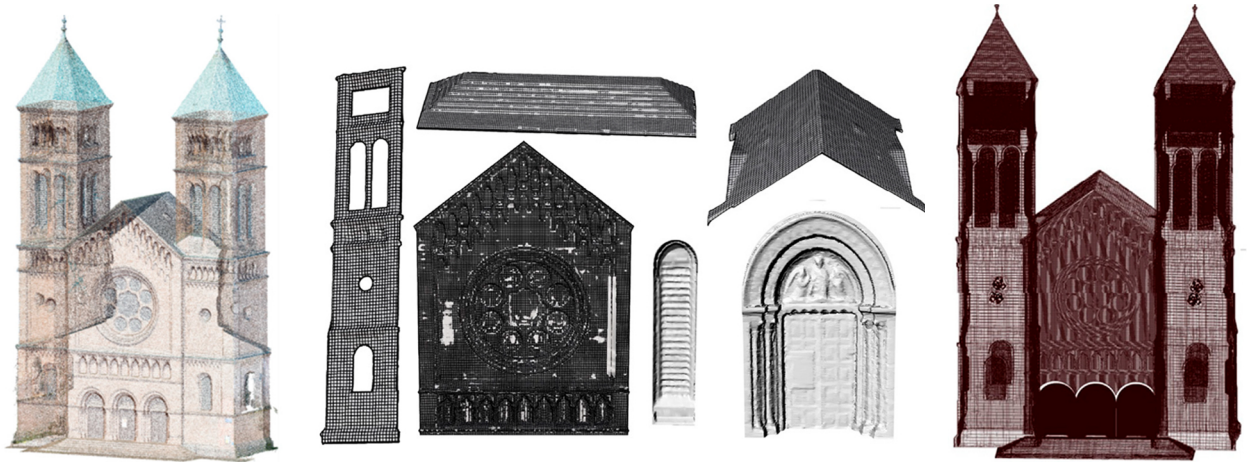
e.g.: Grand-Ducal cloister, Pisa Charterhouse (Grassi, 2019).

et al., 2015; Jung et al., 2014). A conversion of the fitted primitives into BIM-readable objects is in any case needed: Macher et al. (2017), for instance, exploited the open-source software FreeCAD in order to convert building entities from imported data into the IFC format, while Thomson and Boehm (2015) leveraged the eXtensible Building Information Modelling (xBIM) toolkit to turn segmented walls, floors and ceilings into IFC compliant files.

Many commercial solutions proposed by software companies offer tools for the semi-automatic recognition of walls, slabs and pipes (López et al., 2018; Bosché et al., 2015; Wang et al., 2015; Thomson and Boehm, 2015); among them, it is worth mentioning EdgeWise Building (Figure 3-64), a stand-alone software complement for Autodesk Revit promoted by ClearEdge3D (clearedge3d.com); Scan-to-BIM by IMAGINiT Technologies, developed as a plug-in for Autodesk Revit (imaginit.com); Kubit PointSense Buildings and Pointfuse from Arithmetica (pointfuse.com). However, these platforms are limited to the extraction of simple planar features and elements with standardized dimensions: the more complex the model, the less successful these approaches are, and this hampers the application to heritage objects and sites. In presence of multiple planes oriented differently, or of occlusion and clutter errors in the original point cloud, the method may lead to incorrectly segmented and misclassified objects, improperly fitted geometries as well as incorrect assumptions for missing data (Dore and Murphy, 2017).

- Mesh-reconstruction methods.

In this case, models appear as hybrid objects, integrating conceptual shapes and meshes (Macher et al., 2017). The 3D survey is manually segmented, but the elements that are more complex or that require a higher level of detail are imported separately in BIM environment as mesh models, reconstructed from portions of the original point cloud (Figure 3-65). Andriasyan et al. (2020), Pepe et al. (2020), Santagati et al. (2018), Yang et al. (2019, 2018), Rodríguez-Moreno et al. (2016, Figure 3-66) and Barazzetti et al. (2015) proposed several methodological approaches for the conversion of 3D textured meshes derived from surveying into BIM objects; however, the mesh



a. b. c.

Figure 3-65.

Reconstruction method by Yang et al. (2019). Raw point cloud (a); manually segmented meshes (b); final mesh-to-BIM model (c)

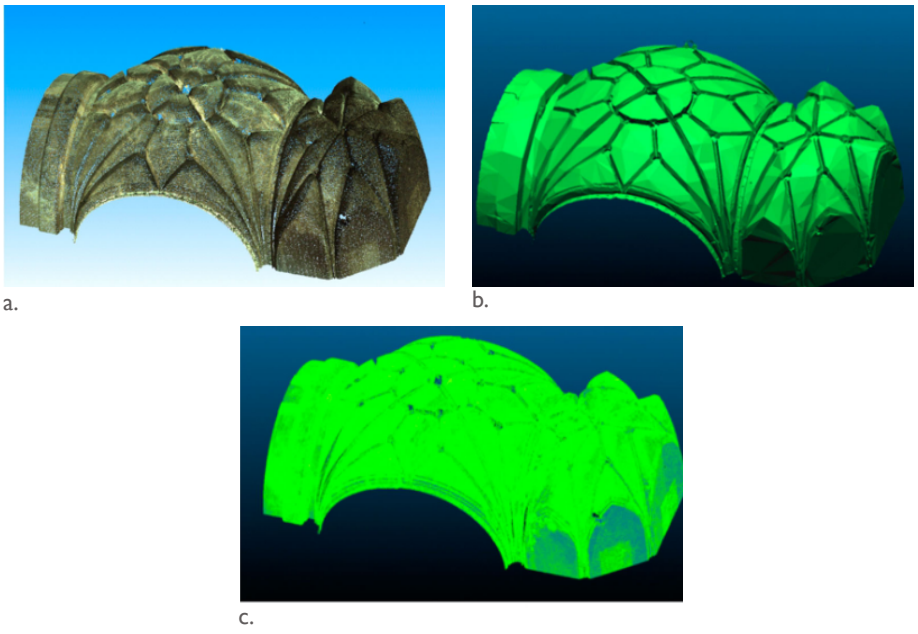
e.g.: St-Pierre-le-Jeune Church, Strasbourg, France (Yang, 2019).

manipulation and geometric modification are limited since the models cannot be edited and controlled by the insertion of parameters (non-*parametric* modeling).

- Reconstruction by shape grammar and object libraries.

Even more automated approaches that have been presented to describe the complexity of existing heritage relied on the construction of suitable 3D libraries of building components (families), to model recurring architectural elements.

De Luca et al. (2007) studied the extraction and formalization of architectural knowledge based on the analysis of architectural treatises, for the generation of template shape libraries of classical architecture (Figure 3-67, Figure 3-68). Similarly, Murphy et al. (2013) modelled



a. b. c.

Figure 3-66.

Transept and chevet modelled with mesh-based reconstruction methods (*a*) based on the original point cloud (*b*) and related cloud-to-mesh distance computation (*c*).

e.g.: Saint Jeromés Church in Baza, Spain (Rodríguez-Moreno et al., 2016).

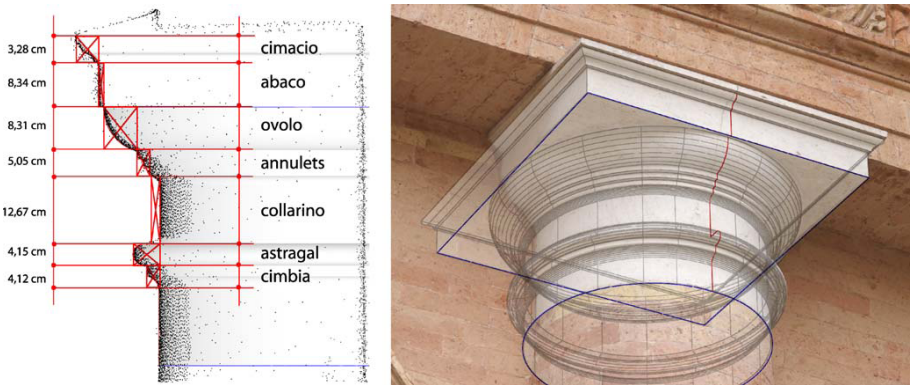


Figure 3-67.

Instantiation of a capital on a point cloud *(to the left)* by reference to prior knowledge on architectural treatises *(to the right).* (De Luca, 2007)

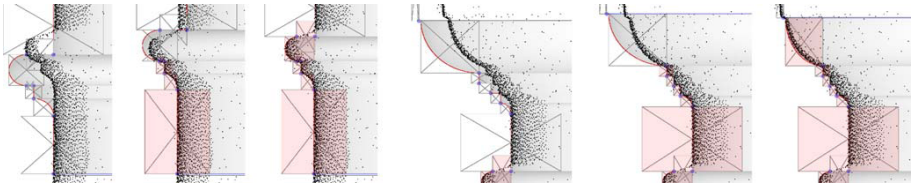


Figure 3-68.

Under-constraint deformation of geometrical atoms for the instantiation of the template geometry. (De Luca, 2007)

interactive parametric objects relying on manuscripts ranging from Vitruvius to Palladio to the architectural pattern books of the 18th century, to derive parametric and shape rules. The objects are dynamic, i.e., they vary based on the control of specific parameters, and their assembling enables the construction of a parametric building façade, in such a way that the information derived from architectural pattern books can be compared to the real dimensions of the elements acquired from surveying.

The work by Quattrini et al. (2015) was devoted to the construction of ontological models and shared libraries for Romanesque buildings, while Oreni et al. (2013) implemented a collection of wooden beam floors, brick vaults and wooden vaults to describe the structural covering system of many historical buildings of northern Italy: the hierarchical aggregation of the objects composing a parametric family, as well as the information on material associated to each, were considered for this study. Analogously, (Baik, 2017) described the Islamic architecture of many relevant heritage buildings in the city of Jeddah, Saudi Arabia, and (Fai and Rafeiro, 2014) documented the openings of the West Block of Canada's Parliament Hill, at three different levels of detail, for the construction of a library of Gothic arches.

These studies all pointed out, to a wider or lesser extent, that the formalization of specific libraries allows the management of information over a single and shared representation of the object; the reference to a conceptual geometry preserves the data independently of a specific modeling language (De Luca et al., 2007).

Being based on the codification of architectural languages and models, these systems allow describing the morphological and formal complexities of historic buildings, by somewhat re-proposing the principles of 2D/3D decomposition of existing buildings that were typical of architectural treatises. The reference, in fact, is always to historic architectural data.

- Reconstruction by generative modeling.

In order to fasten the reconstruction of the identified building components, most recent approaches on shape grammar focused on the use of visual programming environments: the latter consist in the creation of 3D content on a sequence of graphic generation instructions, rules or algorithms that can be repeated with varying characteristics (Kelly, 2006). In a visual programming approach, the components of the 3D model are generated via the manipulation of a visual *node* (algorithm), and the connection of multiple nodes (*wires*) allows to link together the different 3D shapes (Figure 3-69). Visual programming thus allows designers to manipulate geometries by programming interactively, via a graphical coding language, the set of modeling procedures, primitive adjustments and duplication operations performed in 3D space (Yang et al., 2019; Tommasi et al., 2016; De Luca et al., 2007).

Available options for such kind of reconstruction procedures are Grasshopper, a visual programming interface for Rhino3D, and Autodesk Dynamo, a plug-in for Autodesk Revit. These platforms offer a high flexibility for parameter variation and scalable geometric representation, that fit with the logic of BIM based modeling. Besides few attempts (Andriasyan et al., 2020; Pepe et al., 2020; Yang et al., 2019; Capone and Lanzara, 2019; Tommasi et al., 2016), however, the approaches based on visual programming have little been explored for the application of BIM-based reasoning on existing monuments and sites.

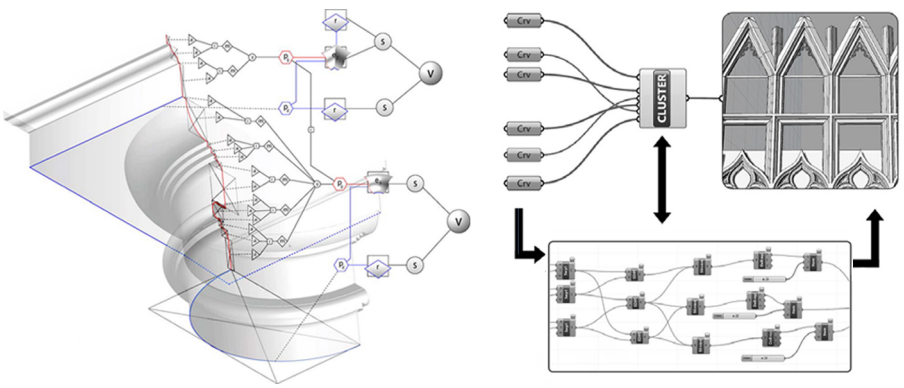


Figure 3-69.

Generative modeling approaches applied to the reconstruction of a capital (De Luca, 2007) and a *falconatura* (Tommasi et al., 2016).



Figure 3-70.

Annotation of localized degradation phenomena on BIM elements. The colors are associated to a legend of decay patterns (Chiabrando et al., 2017).

3.3.5.4. Annotation levels within H-BIM platforms.

A key issue in the application of BIM to existing heritage concerns the type of annotations made available on the 3D model. In H-BIM approaches, whatever conceived, semantic annotations are trivial when related to a single BIM component considered in its entirety.

Simple annotations, as those of the tag or attribute type (Paragraph 3.3.1) can be performed on whole objects (e.g., a wall, a slab, a column). Supplementary sources, e.g., images, CAD drawings or analytical spreadsheets can be associated to these elements, following the logic of the BIM-based documentation as derived from the new constructions industry. However, the annotation process becomes complex when localized information needs to be inserted; this issue occurs, e.g., when mapping degradation, decay conditions or materials, as this may require the insertion of annotations not concerning the parametric element as a whole, but rather parts of it.

In this second case, the annotations are performed by suitably drawing flat or curved shapes over the surfaces of each parametric element. In Autodesk Revit, this is done by referring to the mechanism of the so-called adaptive components (Figure 3-70), i.e., elements that, since not being directly available from the object library, are created from scratch by the user to suit the annotation selection needs (Chiabrando et al., 2017; Lo Turco et al., 2017). The work by Malinverni et al. (2019) relied on the mechanism of adaptive components to derive a thematic map of degradation patterns, with related legend. Each annotation was categorized with a proper ID code assignment, to ease the accomplishment of specific queries within the model. Simeone et al. (2019) applied this procedure to detect transformations and reused elements within the northern wall of the Oratory of San Saba, in Rome.

(Pocobelli et al., 2018b) performed annotations in H-BIM environment so to construct a weathering forecasting model, showing the spatial distribution of moisture along surface walls. Bacci et al. (2019), with a similar procedure, investigated the logical structuring of annotations, by

associating the adaptive components to multi-category tags, related to degradation level, material and/or conservation interventions. Noardo (2018) further studied the connection between annotations of heritage objects made in BIM environments and existing data models used in cartography, as the CityGML.

The introduction of suitable parametric adaptive components extends the annotation capability to localized elements. This feature is an essential requirement for information systems focused on the restoration and conservation of cultural heritage, where also localized areas need to be annotated, for example, when affected by localized damage or degradation phenomena (Croce et al., 2020).

3.3.6. Semantic segmentation

The state-of-the-art on existing 3D reconstruction methods and Scan-to-BIM practices, proposed in Paragraph 3.3.5, demonstrated that data segmentation and classification constitute a fundamental step. In detail, they enable:

- a decomposition of the survey data, whether 2D or 3D, into subsets of elements (pixels or points) sharing the same features, whether geometric, radiometric, spectral, semantic or temporal, extracted from 2D or 3D descriptors and according to predefined criteria (*segmentation*);

- the assignment of a label to each subset (*classification or semantic segmentation*).

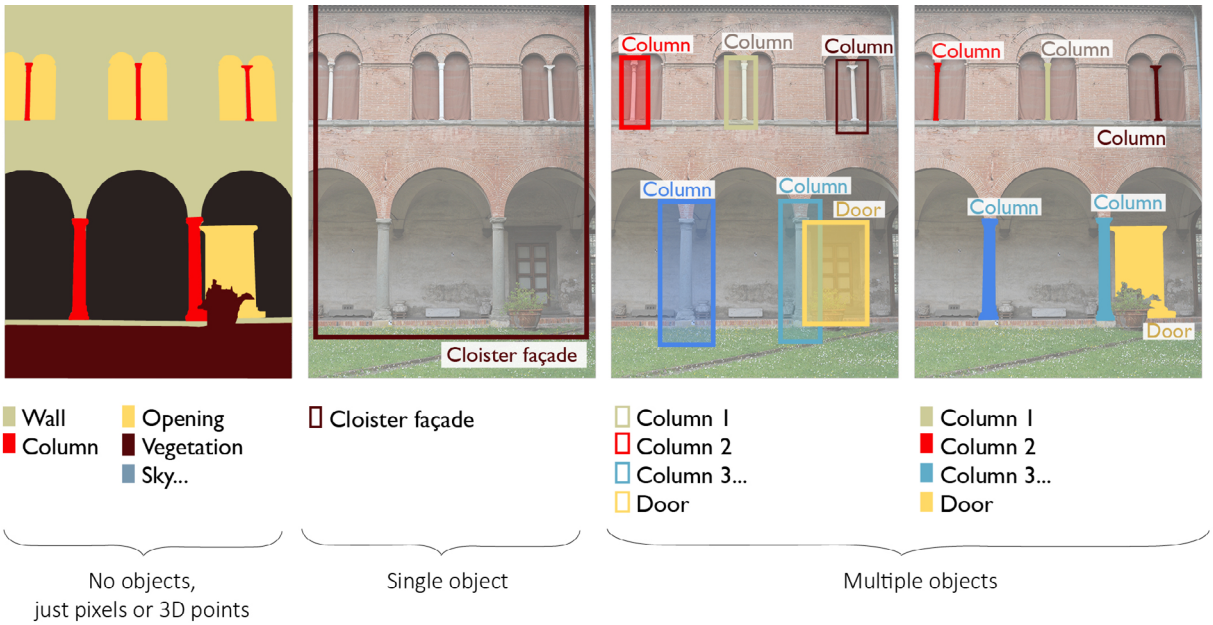
In comparison with other common computer vision tasks as object recognition, instance localization and segmentation, the semantic segmentation process classifies pixels or points as belonging to a certain label and performs this operation for multiple objects of the 2D image or 3D scene (Figure 3-71). If the decomposition is performed by referring to prior knowledge on the studied 2D/3D scenes, the operation is referred to as a semantic segmentation.

Figure 3-71.

Semantic segmentation compared to other computer vision tasks. Adapted from *nanonets.com/blog/semantic-image-segmentation-2020/*

e.g.: Cloister of the National Museum of San Matteo, Pisa

First conceived for image processing and only later applied to 3D point cloud data interpretation, the technique has been developed over time outside the Cultural Heritage domain, e.g., in autonomous driving medical imaging, virtual try-on, visual image search, video surveillance systems (Kar et al., 2021; Xie et al., 2020). The applications to the various fields were however led by the common goal of simplifying complex scenes by grouping objects/points/pixels with same characteristics, so to facilitate and automate further processing (Landes, 2021).



3.3.6.1. Overview of existing segmentation and classification methods.

Numerous segmentation and classification approaches have been proposed over time, extended from the segmentation of 2D images to that of point clouds: on the latter, this is demonstrated by the comprehensive literature reviews recently proposed by Landes (2021), Grilli et al. (2017) and Poux et al. (2016) who enriched the state of the art by (Weinmann, 2016; Vosselman et al., 2003).

According to the distinction proposed by (Grilli et al., 2017), the different approaches proposed over time to the segmentation problem can be grouped as follows, both for 2D and 3D data: region-growing, edge-based, primitive fitting, hybrid methods, and AI-based applications. AI-based methods are thoroughly investigated in Paragraph 3.3.6.2.

- Region-growing.

Region-based methods rely on a two-step procedure: they start by the identification of initial points (*seed points*), and then grow the selection based on similarity and proximity criteria, as surface orientation, curvature constraints etc. (Klasing et al., 2009; Belton and Lichti, 2006).

Initially introduced by Besl et al. (1988) and relying solely on geometric features, the method was refined by Vosselman et al. (2003), who added the color properties as further similarity criteria. Gonizzi Barsanti et al. (2017) tested region-growing methods for point cloud segmentation on several heritage objects (Figure 3-72).

Although resulting more robust to noise if compared to edge-based methods, these approaches are highly influenced by the initial location of seed points, as improper selections may lead to under- or over-segmentation results, as well as by the inaccurate calculations of normals and curvature values near region boundaries (Grilli et al., 2019) that may make the clustering process fail (Gonizzi Barsanti et al., 2017).

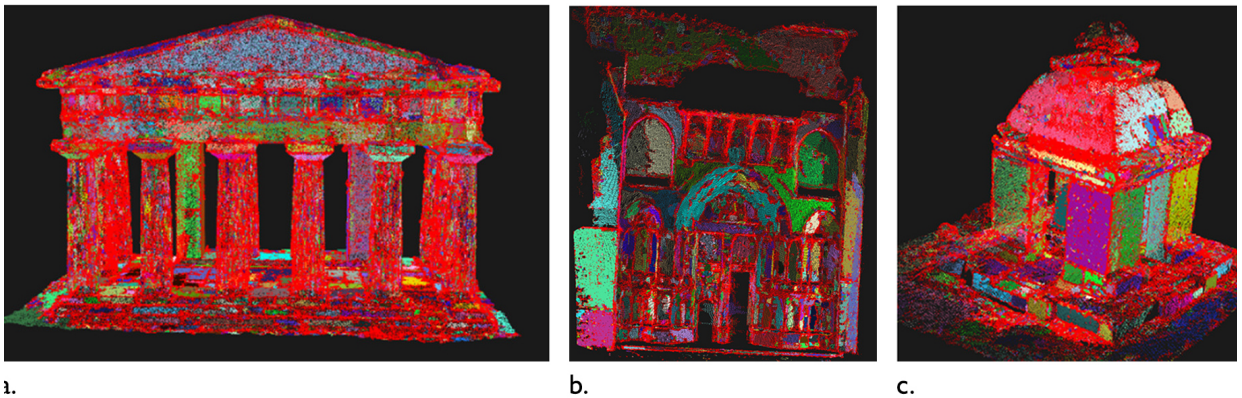


Figure 3-72.

Segmentation via region-growing for heritage point clouds.

e.g.: the Temple of Neptune in Paestum (a); the S. Emilion portal (b) and the Prasat in Angkor (c).

(Gonizzi Barsanti et al., 2017)



Figure 3-73.

Hybrid segmentation methods: extraction of points lying on silhouettes (*a*); detection of points with high variation in normals as edges (*b*); grouping of smooth surfaces following region-growing (*c*).

(Che and Olsen, 2017).

- Edge-based methods.

These methods operate through edge detection filters, as the ones described in Paragraph 3.2.1.3, in which the abrupt change of uniformity criteria enables highlighting the edges of an element (Al-Amri et al., 2010; Arbelaez et al., 2010; Kaur et al., 2012). As for point clouds, features as the gradients, the normals or the principal curvatures are used to first detect the edges that outline the borders (i) and to later group them inside the boundaries (ii) to convey the different segments (Rabbani et al., 2006 ; Ni et al, 2016; Wang and Shan, 2009). However, such methods are very sensitive to noise or irregular density, situations that commonly occur in point clouds.

- Primitive fitting methods.

Primitive fitting methods rely on the decomposition of the shape into simple geometric primitives, as planes, cylinders, spheres. They have been studied also in view of the construction of BIM-based representation for heritage objects: the methods, as well as their applications in current Scan-to-BIM approaches, have been formerly described in Paragraph 3.3.5.3, to which the reader can refer.

- Hybrid approaches.

They are based on combining two or more of the previously described approaches to exploit the potential and overcome the limitations of the single strategies taken individually. Typically, a combination of region-based and edge-detection methods is considered, as illustrated in the example of Figure 3-73 (Che and Olsen, 2017; Vieira and Shimada, 2005; Lavoué et al., 2005).

3.3.6.2. AI-based methods.

Such segmentation systems exploit AI, and in detail Machine Learning (ML) algorithms, to create aggregations of basic elements. Analyses and predictions via ML are performed relying on empirical training data, which are considered as examples that guide the segmentation process. Deep Learning (DL) systems, which allow learning through artificial neural networks, are also part of these techniques (Figure 3-74).

Landes (2021) refers to ML-methods as aggregation-based approaches (fr: *approches basées sur l'agrégation*). They can be performed either via supervised or unsupervised learning techniques (Jatana, 2020), the difference being that the former uses labeled data to make the prediction, while the latter does not (Figure 3-75).

In supervised learning, in fact, each training example is described by different inputs (classes, *features*) along with the related output. For the classification task, the output is given in the form of a class or category and the supervised learning algorithm trains a predictive model to recognize the class associated with the new inputs.

In contrast, unsupervised learning algorithms work on unlabeled datasets: they uncover the inherent structure and hidden patterns from the given data in order to extract useful representations (*features*) based on specific criteria (Fiorucci et al., 2020). Unsupervised ML is mainly based on *clustering* (grouping unlabeled data based on similarities or differences) or *association method*s (identifying relationships in a given set of input data).

The choice of appropriate features is critical in defining these problems, as properly optimized feature extraction improves the performance of the predictive models, both in terms of training time and of accuracy of the resulting prediction.

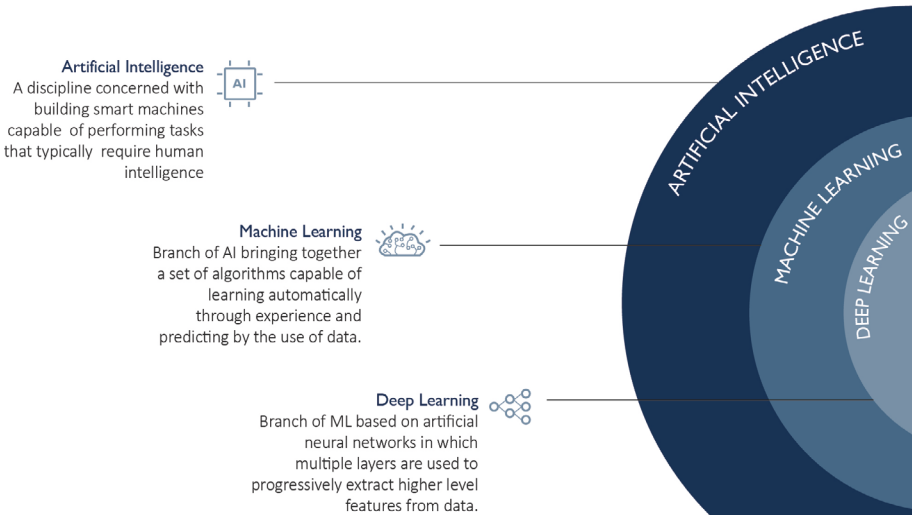


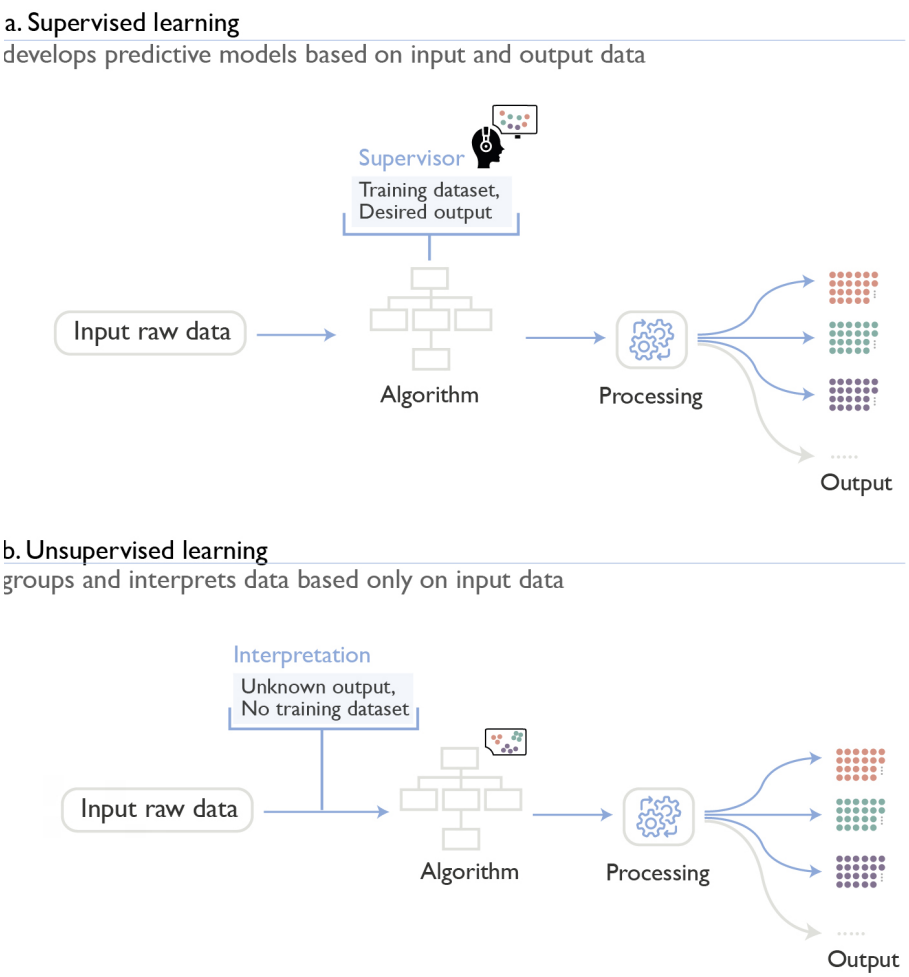
Figure 3-74.

Relationship between AI, ML and DL. Adapted from *intesa.it/intelligenza-artificiale-machine-learning-e-deep-learning*

Figure 3-75.

Difference between supervised (*a*) and unsupervised (*b*) ML approaches.

Adapted from Jatana, 2020.



The appropriate selection of most robust and relevant features addresses the problem of heterogeneity, which specifically characterizes point cloud data (Grilli et al., 2017; Weinmann et al., 2015).

Most robust supervised ML algorithms are Support Vector Machines, Decision Trees, Random Forest, and Artificial Neural networks.

- Support Vector Machines (SVMs).

SVMs are binary classifiers whose objective is to find the hyperplane that distinctly classifies data points of one class from those of another class. The separation of the two classes can be performed with identification of many possible planes, but the goal of such algorithms is to identify the plane having the *maximum margin*, i.e., that maximizes the width of the slab parallel to the hyperplane that has no interior data points (Figure 3-76a). *Support Vectors* are a subset of the training observations that are closest to the separating hyperplane, demarcating the boundary of the slab.

SVMs have been extensively applied to the classification of urban environments as well as in the geo-spatial field (Kavzoglu and Colkesen,

2009); their use in these domains even allowed the classification of point clouds (Bogdan and Cousins, 2011; Caputo et al., 2015) for the distinction, e.g., of vegetation, road surfaces, land cover and construction sites (Grilli, 2020). However, these algorithms require a long training time in the case of large datasets, and they do not perform well when the target classes are overlapping. They are thus suitable only when dealing with small or medium datasets.

- Decision trees and Random Forest (RF).

Decision trees (Figure 3-76b) are tree structures used to split and recursively partition a complex problem into disjoint sub-sets, based on binary decision rules corresponding to certain cut-off values in the features (Fiorucci et al., 2020). Each branch of the tree can be seen as a yes/no question, and the series of decision nodes leads to a certain leaf node, corresponding to the prediction.

Single decision trees, however, may create over-complex models that do not generalize well across the different data: this problem is known as overfitting, and it occurs when the algorithm captures noise in the data. Moreover, small variations in the data might result in a completely different tree being generated (*variance*).

These issues can be tackled by the construction of an ensemble of de-correlated decision trees operating together as a forest: the Random Forest model originally proposed by Breiman (2001) relies on the idea of averaging many simple noisy un-correlated decision trees. Each decision tree of the forest spits out a class prediction, and the class with most votes becomes the model's prediction.

To ensure that each individual tree is non-correlated with the behavior of any other tree, the RF uses two fundamental methods: the Bootstrap Aggregation (Bagging, i.e., data are randomly sub-sampled with replacement) and the feature randomness (i.e., node splitting is based on a random subset of features).

Two important components must be set for the learning of the tree structure: i) the number of trees in the forest (even called *estimators* or *learners*); ii) the number of variables to be selected and tested for the best split when growing the trees (Belgiu and Drăguţ, 2016).

RFs have been widely applied in remote sensing, to classify point clouds of urban scenes, as demonstrated by the wide literature review already proposed by (Belgiu and Drăguţ, 2016).

Unsupervised learning methods are introduced in the following.

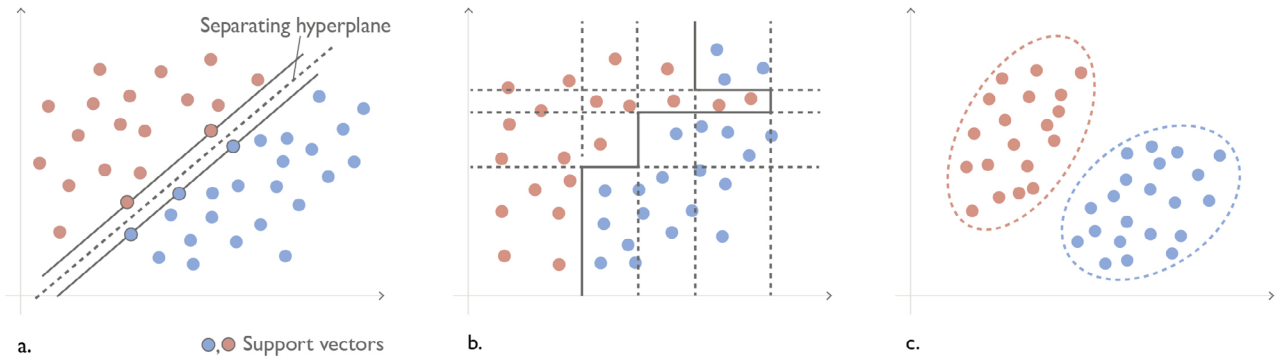
- K-means clustering.

This unsupervised learning method (Figure 3-76c) relies on a clustering algorithm that partitions the given data into a K number of clusters (groups) using similarity features. As such, data points falling in a same cluster will result more similar to each other than those falling into other clusters.

The algorithm randomly assigns, at first, the data points to any of the K clusters. Then, the centroid of each cluster is calculated and the distance of data points from the centroid of each cluster is derived. The data points are later assigned to the nearest cluster. This operation is repeated many times and the centroid of the new cluster is calculated until data points do not change clusters, or until a given number of iterations is reached.

The algorithm initially proposed by MacQueen et al. (1967) has later been tailored to point cloud classification (Grilli et al., 2017; Kuçak et al., 2017). Although being simple to implement and suitable for large datasets, K-means algorithms do not perform well in case of clustering data with varying sizes and density. The result can also be biased by the presence of outliers: prior statistical outlier removal or noise reduction filters can thus be considered before clustering (Chehata et al., 2009).

Figure 3-76.
Support Vector Machine (*a*), Decision tree (*b*) and K-means clustering (*c*) working principles.



- Artificial neural networks (ANNs)

The Artificial Neural Networks (ANNs) are mathematical models that imitate the functionalities of human brain (Bre and Gimenez, 2017). Inspired by the brain recognition process, they consist of an input layer i, one or more hidden layers (h_1, h_2, ..., h_n), and an output layer o: highly connected networks of neurons relate the input to the desired outputs (Figure 3-77).

The information input is elaborated through the construction of hierarchical levels of information abstraction (layers), simulating the way organic neurons work. Neurons belonging to a same layer recognize

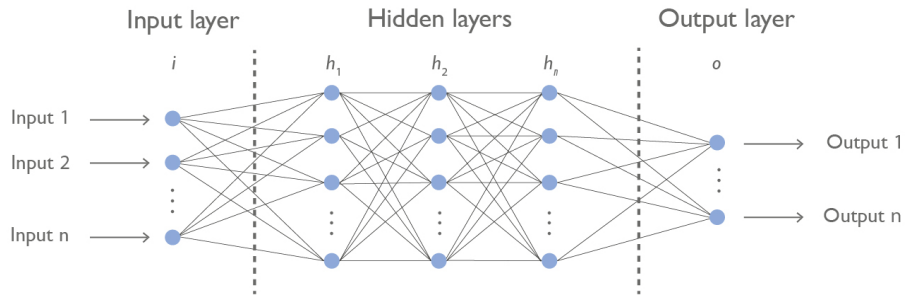


Figure 3-77.

Artificial neural network architecture (input layer i, hidden layers h_1, h_2, ..., h_n; output layer o). Adapted from Dastres and Mohsen, 2021.

similar features: for instance, low level layers are generally able to recognize lines and edges of an image, as simple features, whereas high level layers learn features such as more complex textures and patterns.

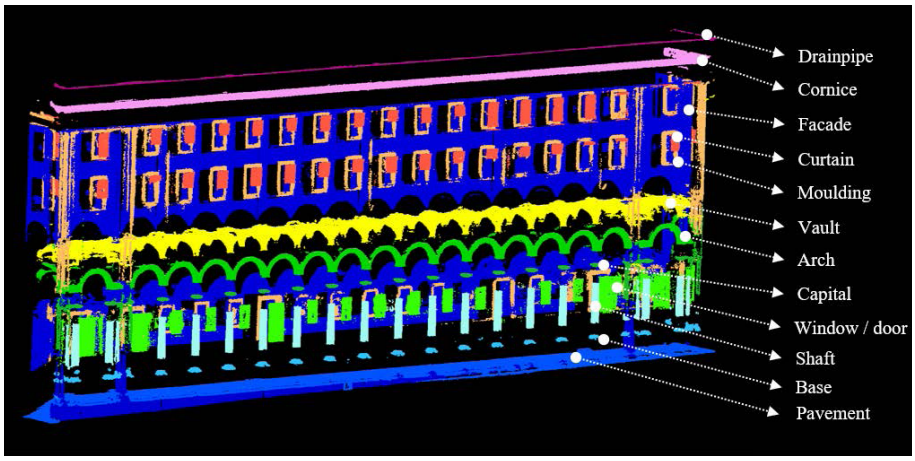
In the case of multiple hidden layers, ANNs are referred to as deep neural networks: the term 'deep', indeed, refers to the depth of the built mathematical model. Deep learning (DL) neural networks have constituted a relevant advance in AI techniques, and they can be considered the evolution of ML approaches, as the goal of DL is to perform end-to-end learning, i.e., the data themselves are the inputs and both the features and the classification are learned as part of the training process.

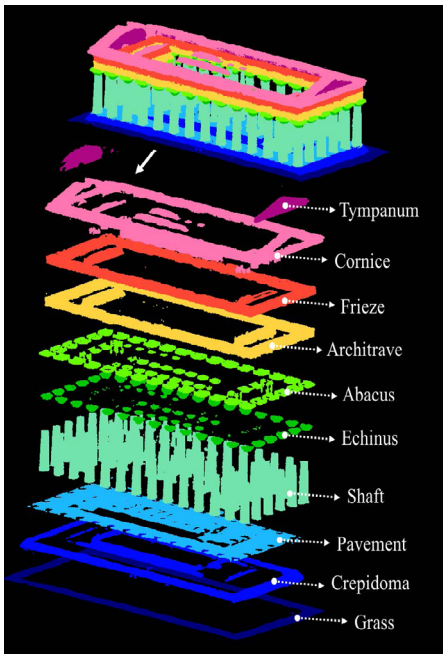
Extensively studied for image processing (Le Cun et al., 2015), deep learning techniques have recently been tested for 3D data classification, as demonstrated by the literature review proposed by Griffiths and Boehm (2019). The advantage of such methods is the possibility to constantly update the model with new training information, besides the fact that representations required for detection or classification are automatically derived. However, a relevant amount of data is required to train DL models (Schmidhuber, 2015; Goodfellow, 2016).

3.3.6.3. Semantic segmentation of heritage objects and sites.

In the Cultural heritage domain, ML and DL are emerging to assist digital data interpretation, logical organization, and semantic enrichment of a given asset being studied (Fiorucci et al., 2020) e.g. in terms of recognition of architectural elements (Pierdicca et al., 2020), re-assembly of dismantled fragments (Paumard et al., 2020), detection of occluded or damaged wall regions (Ibrahim et al., 2020), mapping of the spatial distribution of historical phenomena and ancient trade networks (Varinlioglu and Balaban, 2021).

In the case of surveying data, attention is directed towards the automatic recognition and annotation of relevant characteristics of a building or site: from the former experiments geared towards the semantic segmentation of heritage 2D images (Ibrahim et al., 2020; Korc and Forstner, 2009; Manfredi et al., 2013), the investigations moved on





Figures 3-78, 3-79.

Point cloud classification results by Grilli et al., 2019.

e.g.: Temple of Neptune in Paestum (*above*); Building with porticoes in Bologna (*to the left*).

to studying more automatic annotations directly on 3D media, i.e. point clouds (Grilli et al., 2018) and/or textured polygonal meshes (Grilli and Remondino, 2019).

The use of ML algorithms for the classification of heritage point clouds remains however limited to few applications, that either explore the classification of architectural objects into their typological components (*geometry-based approaches*) (Matrone et al., 2020; Morbidoni et al., 2020; Grilli et al., 2019; Malinverni et al., 2019) or look for the detection of degradation patterns or material mapping (*texture-based approaches*) (Bassier et al., 2020; Grilli and Remondino, 2019; el Kadi, 2014). In any case, supervised ML methods, i.e., relying on annotated training data as input, appear more suitable to describe the complexity and morphology of historical buildings.

Grilli et al. (2019) transferred the classification techniques that exploit geometric covariance features to the realm of cultural heritage (Figure 3-78, Figure 3-79), previously restricted solely to the geospatial sphere for the semantic segmentation of urban scenes (Özdemir et al., 2019;



Figure 3-80.

Examples of labelled point clouds from the ArCH dataset (Matrone et al., 2020).

Weinmann et al., 2015). They investigated the effectiveness of geometric features in training a RF classifier and explored the correlation between such features and the dimensions of main architectural elements.

Similarly, Bassier et al. (2017) used SVMs to perform multiclass learning for the segmentation of floors, ceilings, roofs, beams and walls in a group of existing buildings, including regular houses, castles and churches. However, the use of SVMs proved not to be feasible for large datasets, since requiring a too long training time.

Foreseeing the construction of training datasets for Machine Learning applications on heritage objects, Murtiyoso and Grussenmeyer (2019) presented a series of functions and algorithms for point cloud segmentation, in the form of a toolbox named M_HERACLES (github.com/murtiad/M_HERACLES).

Besides ML techniques, the application of DL classification methods has also been investigated starting from the introduction of PointNet (github.com/charlesq34/pointnet; Qi et al., 2017) and its later extension PointNet++ (Qi et al., 2019), deep neural networks capable of calculating three-dimensional geometric data of simple shapes. The application of PointNet++ to heritage point clouds has been previously explored in Malinverni et al. (2019). Pierdicca et al. (2020) tested PointNet and PointNet++ with an improved dynamic graph convolutional neural network and assessed both methods on 11 labeled point clouds

representing indoor and outdoor scenes of heritage architectures of different historical periods.

Several researches are also emerging in the heritage domain to exploit AI in noise-masking of heritage scenes: starting from photogrammetric models, Murtiyoso and Grussenmeyer (2022), benefited from the deep learning-based semantic segmentation of images and projected 2D semantic labels into the 3D space, to exclude noise and unwanted object classes from the 3D scene. A similar approach was presented by Pellis et al. (2022).

When comparing the performance of the ML and DL classifiers for large scale 3D heritage semantic segmentation, however, Matrone et al. (2020a) state that there is no winning solution, since the segmentation results are alternatively better depending on the case to be studied.

Nevertheless, DL systems need large amounts of annotated datasets to classify new observations, and the current unavailability of this kind of information in the heritage domain is a limitation to their application in this field. The ArCH dataset, presented by Matrone et al. (2020b), was a first attempt to reorder and archive multiple sets of annotated point clouds to provide large training datasets of heritage objects (Figure 3-80).

The above-mentioned studies all indicated, to a greater or lesser extent, how the use of ML and DL techniques in 3D heritage data can facilitate the scan-to-BIM reconstruction by semantically separating elements in point clouds as a prior step to the import within BIM environments. However, to the author's knowledge, no research has to date dwelt on the difficulties of switching from the annotated point cloud to a H-BIM model. This still-uncertain and unclear transition, deserving more in-depth analysis worth the exploration of the operational challenges of a scan-to-BIM via ML model (Croce et al., 2021a,b), constitutes a research line of this Ph.D. work.

3.3.7. Critical analysis of existing 2D/3D annotation platforms

Existing 2D or 3D annotation solutions provide the cornerstone for the construction of semantically rich digital representations, as they offer new perspectives of data analysis and correlation that go beyond the pure visual or geometric descriptors extracted from raw surveying (Figure 3-81).

Semantic enrichment is a key aspect in preventing the dispersion of knowledge, as it enables bringing together the outcomes of multidisciplinary heritage studies and linking in a same digital environment both quantitative information (obtained from acquisition

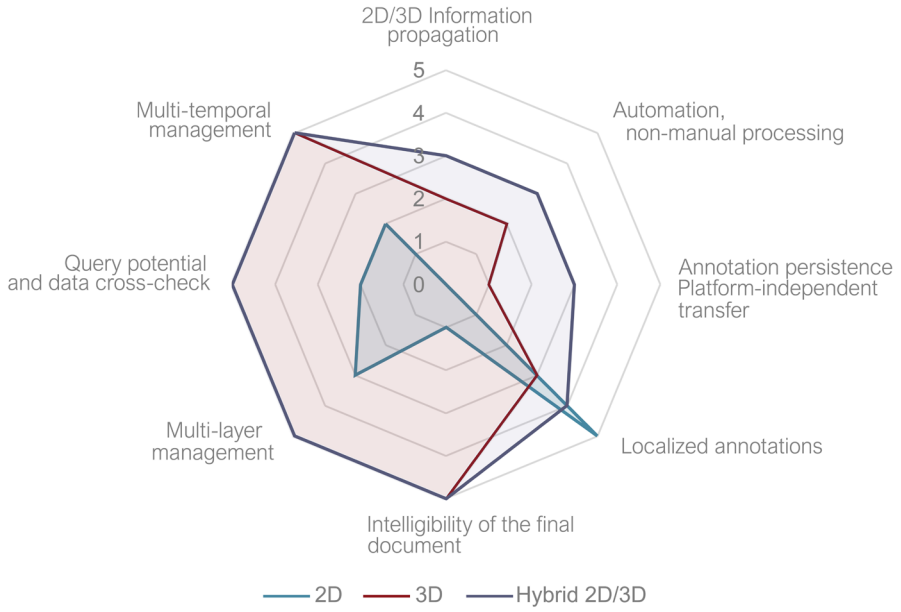


Figure 3-81.

Graphical comparison between 2D, 3D and hybrid 2D/3D annotation approaches.

and geometric survey processing) and qualitative aspects (relating to the semantic data interpretation).

Existing systems for 2D, 3D or hybrid annotation have been presented and classified in this chapter, respectively according to the criteria of: part/whole relationship (Paragraph 3.3.1), degree of automation (Paragraph 3.3.2) and type of annotated media (Paragraph 3.3.3).

The described experiments show that semantic and morphological dimensions can be put together within a knowledge model, so to access new levels of analysis and information management, particularly in collaborative work settings. Both 2D and 3D digital supports can be semantically enriched through an annotation process. However, the shift to a 3D media allows a complete and clear representation of the heritage object. To this latter regard, different solutions for 3D annotations have been analyzed, with an in-depth regard on the reality-based annotation platform Aïoli (Paragraph 3.3.4) and on H-BIM-based annotation systems (Paragraph 3.3.5).

Figure 3-81 reports a graphical comparison between 2D, 3D and hybrid 2D/3D annotation systems reviewed. In detail, 2D annotation systems were the earliest ones to be developed, and they first enabled the management of multiple layers of description (*multi-layer management*), even referred to multiple temporal states (*multi-temporal management*), over a single, bi-dimensional, digital representation, whether it be an ortho-photo, an image a CAD drawing.

Within such kind of systems, localized annotations (e.g., related to the description of degradation phenomena, of material mapping, of decorated portions, of crack patterns etc.) are possible as long as they

refer to the single 2D spreadsheet chosen (*low query potential and low data cross-check*).

3D annotation systems partially solve this issue, as annotations can be linked to a same semantic 3D model, virtually tying up annotations to its parts and not requiring the production of multiple annotated documents. BIM-based representations for cultural heritage leverage this type of annotations. Data overlapping is permitted, and information do not require to be duplicated, as is the case of traditional cartographic mapping (*higher level of multi-temporal and multi-layer management, higher intelligibility of the final model*). Cross-check of data and information retrieval through queries are thus enabled to explore correlation elements (e.g., between materials and state of degradation, or between state of degradation, crack pattern and outcomes of past interventions, also in view of the assessment of their evolution).

These advantages are also found in hybrid 2D/3D annotation platforms as Aïoli, but the latter even present the additional benefit of enabling a first type of annotation transfer between image and point cloud (*2D/3D information propagation*). A *higher level of automation* is thus granted with respect to other annotation approaches studied.

However, the collaborative management of these platforms is still constrained by the low level of interoperability and information transfer both 3D and hybrid approaches have. In other words, an annotation inserted at the level of the single platform or the single model (either a point cloud, a mesh, a parametric model, an image), might not be imported and preserved when switching to a different type of representation or from one platform to another. In Scalas et al. (2020), this issue is related to the so-called *annotation persistence*: the lack of persistence of information when passing from one model to another can again be the cause of data leakage and thus may lead to dispersion of knowledge.

Sounder support of interoperability scenarios remains, and thus requires further studies (Croce et al., 2020).

Nevertheless, it might be noted that the availability of more automatic approaches to 2D/3D data annotation, based on AI, paves the way for more insightful systems of information transfer and propagation. Indeed, the application of supervised ML methods to heritage segmentation and classification might improve 3D data interpretation and automate the elaboration of surveying outputs. Existing 2D/3D semantic segmentation methods could be combined with the data structuring and organization logic typical of H-BIM systems or reality-based annotation platforms, to create better and more effective matches between survey data and reconstructed digital models rich in information content; this also in view of the automation of Scan-to-BIM processes.

3.4. Conclusions and research challenges

In Chapter 3, existing methods for the acquisition, processing and semantic enrichment of digital heritage data are described.

Surveying techniques can today be considered consolidated, lessening the time for on-site acquisitions and enabling the digital restitution of objects or sites, via two- and three-dimensional visualizations. Available surveying techniques can be diversified according to the complexity, characteristics and conditions of the objects to be surveyed. Among these, laser scanning and photogrammetry are widely used and integrated for the digital documentation of heritage artifacts (Section 3.1).

However, surveying implies the manipulation of heterogeneous data of various nature, origin and level of detail. This opens many possibilities to the research communities in terms of information retrieval and analysis, as numerous 2D and 3D descriptors can be extracted from raw survey data to enhance geometric and/or visual attributes (Section 3.2). Yet, this richness is also a source of vulnerability, as data often remain under-exploited due to the lack of effective tools to allow complete manipulation, handling and connection. In addition, processing the raw survey data requires complex interpretation procedures.

Data segmentation and classification are crucial aspects of the data processing phase, as they enable the subdivision of complex 2D or 3D data according to the logic of the represented components (*segmentation*) and the association of each group of elements to a semantic label (*classification*).

The semantic annotation mechanism is fundamental for the definition of such segmentation and classification strategies, to enable the association of geometric and visual attributes with semantic information. In Section 3.3, existing systems for 2D and 3D data annotation are presented, with particular emphasis on the existing collaborative reality-based annotation platform Aïoli and on modern H-BIM systems. These two methodologies allow at different levels: multilayer management, multitemporal management, knowledge-based management and data interpolation, but several open issues need deeper investigations.

First of all, the annotation of heritage models is today mostly manual, thus resulting a time-consuming process also prone to too subjective choices. In addition, the insertion of annotations, and additional information on 2D/3D models, intrinsically depends on the type of representation used, either a point cloud, a mesh, a parametric model, a 2D drawing, a photo etc. Consequently, when the digital

model is modified, further usage, exploitation and updating of crucial information concerning the considered heritage object is prevented, unless the information previously acquired is suitably *re-inserted* in the new model. Interoperability problems thus persist in the transition between different types of representation, i.e., in the exchange of information within models of different nature or created at different times (Paragraph 3.3.7).

In order to solve this *spatial and temporal reference gap*, the creation of a semantic annotation transfer system allowing dialogue between different types of models, regardless of the refinement and precision of the representation, is to be envisaged. A new framework for the archival and access of semantically annotated data on digital models would make it possible to spatialize, visualize, retrieve, update the different information produced on the object of study, within a virtual representation. The data to be associated to the model could be related, for instance, to mapping of degradation, materials, intervention techniques and measures, different temporal states or evolution phenomena, on both 2D and 3D media.

As a consequence of these considerations, the present work considers the improvement of the following relevant topics: more automated semantic annotation procedures (Paragraph 3.4.1); annotation transfer between different representation systems (Paragraph 3.4.2); construction of a formalized knowledge base (Paragraph 3.4.3).

3.4.1. More automated semantic annotation procedures

The semantic annotation procedure, also in view of the *annotation persistence,* could be greatly expedited with the introduction of automatic or semi-automatic ML techniques. Their application would indeed lead to a significant reduction of the human involvement in the brute annotation phase, lessening the space of arbitrary or subjective choices.

This results particularly promising for detection and labeling, over 3D models, recurrent and repeated typological elements, such as walls and slabs, columns, capitals, vaults, mullioned windows, rose windows or even more complex elements, whatever the scope and the extension of the annotation.

3.4.2. Annotation transfer between different representation systems

As shown above, in Paragraph 3.3.3, semantic annotations can be performed over different media; for this reason, the output of a same

semantic annotations may vary based on the model resolution and on the type of representation (e.g., different number of triangles or polygonal surfaces of a mesh, different point clouds density).

A research challenge thus concerns the development of a system of input and retrieval of information, to be:

- Independent of the type of representation (software used, accuracy and scale of the 2D or 3D representation and so on);
- Insensitive to changes in geometry due to refinements of the model, e.g., availability of additional data, combination of different representation techniques (point clouds, ortho-photos, mesh, parametric models).

This task is partly raised in the study by Scalas et al. (2020), that developed a first pipeline restricted to the annotation transfer between mesh models at different resolution. The achievement of the *annotation persistence*, leveraging annotation transfer between multiple representations, would ensure continuous update and refinement of semantic annotations over time, regardless of the type of model used (Croce et al., 2020).

3.4.3. Construction of a formalized knowledge base

Heterogeneous information should be organized within the model by setting up a unique data entry process, aiming to form a common basis to be recognized and shared by the different experts working on the heritage asset. By formalizing the knowledge related to a specific domain, different data can be collected and structured based on a machine-readable rationale and on controlled specific glossary and vocabulary. It is a matter of semantically enriching the representation through the use of digital information systems, aiming in particular at connecting the threefold facets of digital heritage models, concerning visual, metric and semantic features (Croce et al., 2020).

Considering these aspects, the approach proposed hereafter aims at streamlining the annotation procedure as well as fostering the exchange of research results on architectural heritage assets, so to better ensure the archival and retrieval of digital heritage information, for documentation and dissemination purposes.

Part II

a. SEMANTIC SEGMENTATION VIA ARTIFICIAL INTELLIGENCE



- Identification of relevant and recurring elements
- Segmentation of unstructured 2D/3D data
- Interpretation of raw surveying information via ML

Chapter 4

The rationale of the method

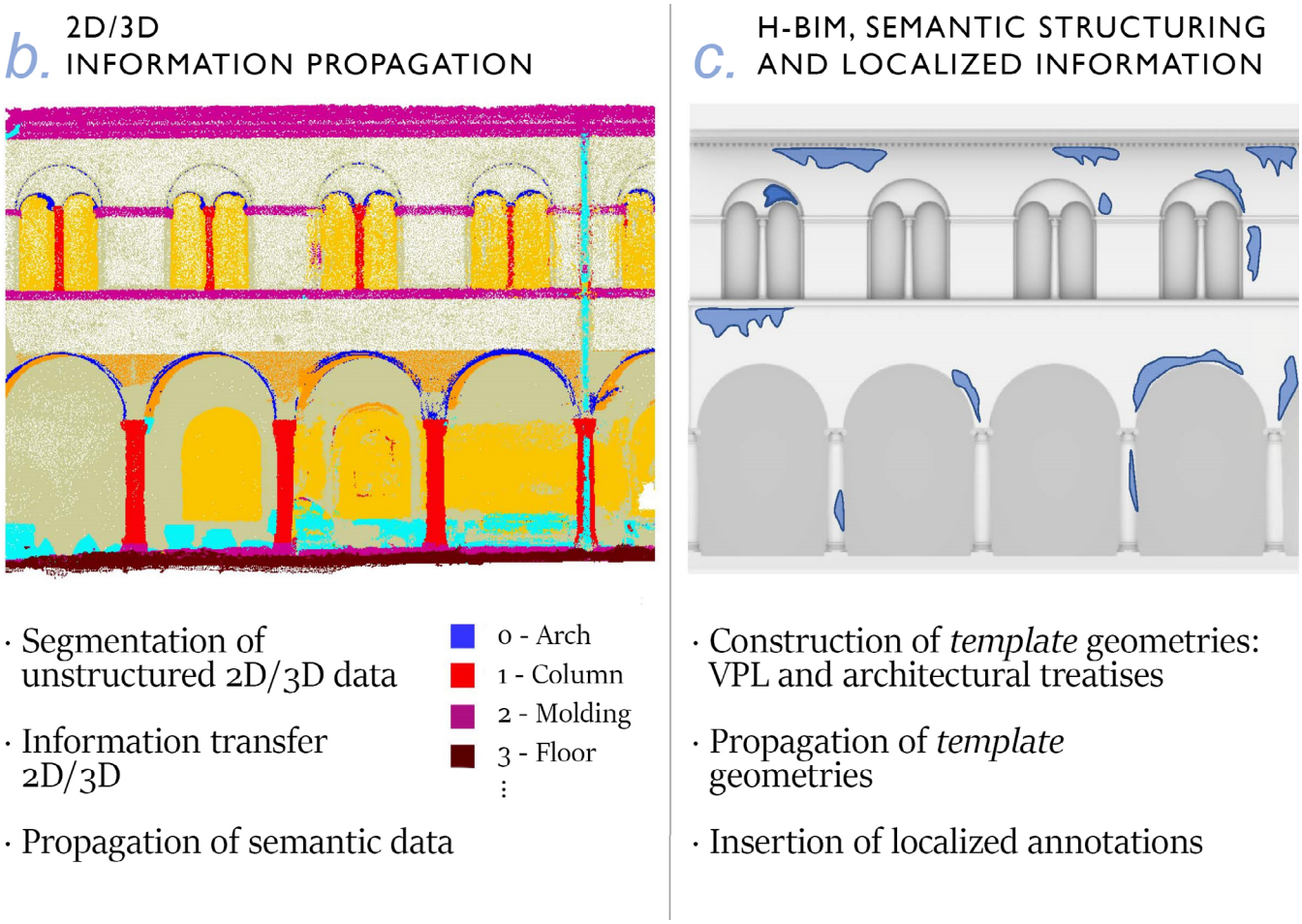


Figure 4-1.

Phases of the proposed methodological approach.

In light of the state-of-the-art on existing systems for the acquisition and processing of survey data proposed in Part I, and considering the related research challenges that emerged in Section 3.4, aim of this thesis is the definition of an integrated methodological approach enabling the structuring, transfer and propagation of semantic annotations between different types of digital heritage models, by exploiting more automated annotation procedures and toward the construction of collaborative heritage information systems.

In tackling this issue, we refer to the definition of a common reference model to which information is each time associated, this model being established a priori and based on the segmentation and classification of heritage objects into typological elements and components.

For the definition of such methodological approach, in detail:

- More automated semantic annotations strategies, relying on AI, are studied to ease the interpretation, segmentation and classification of surveying data, even in view of the automation of Scan-to-BIM practices;
- The Aïoli platform and H-BIM systems are used as the informative basis, since they provide collaborative environments where visual, metric and semantic data can be archived, retrieved and amended directly on the digital representation.

As such, the ambition is to enable more effective, semi-automated interpretation and transfer of semantic data, through the use of annotation systems.

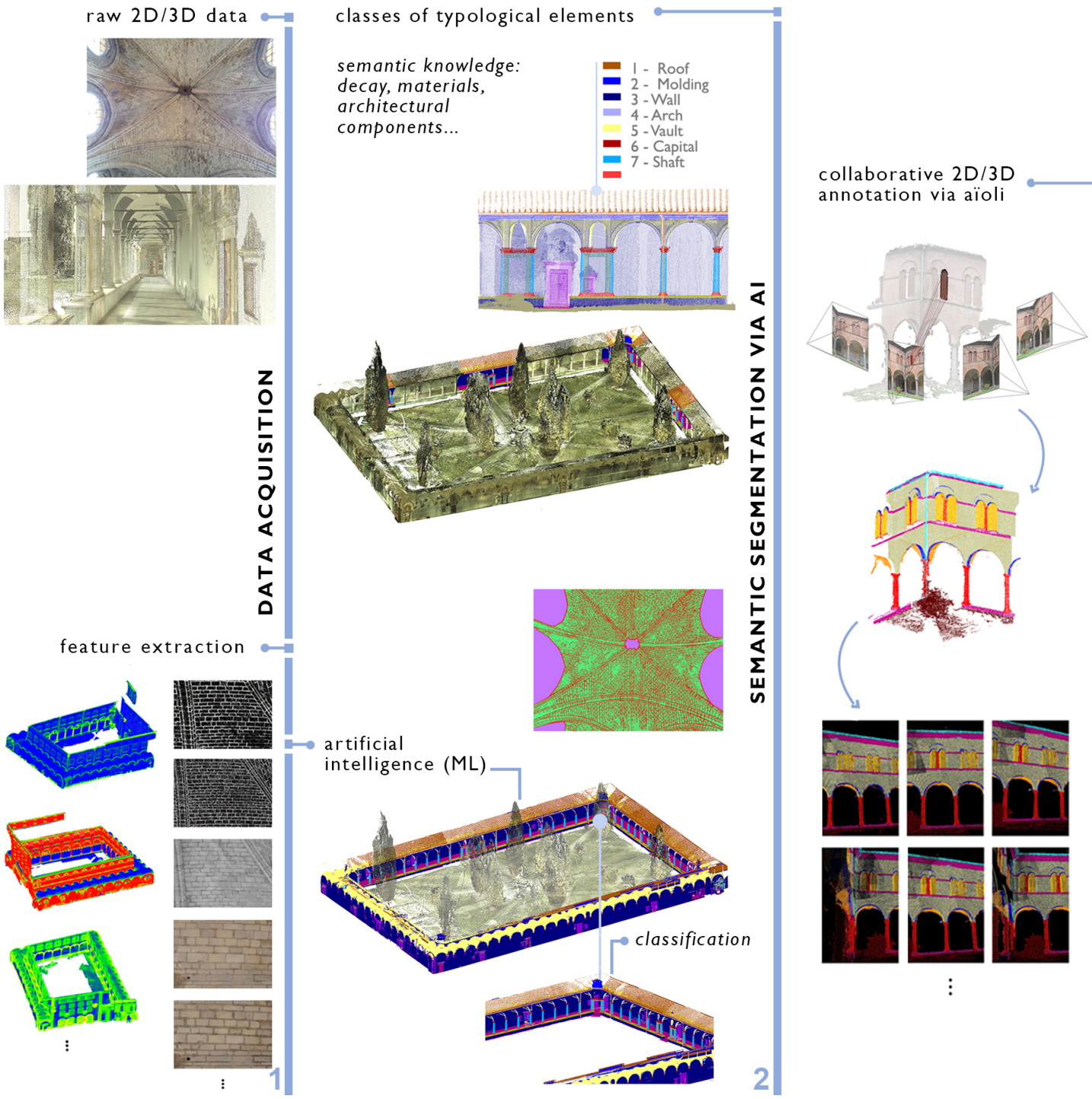
The starting point for the definition of the methodological approach is data acquisition: the architectural survey is intended as the very first phase, where the description and representation of an existing architectural object take place by detection of its 2D/3D characteristics of shape, geometry, size and/or color. Laser scanning and photogrammetry techniques are used as privileged tools to acquire 3D data with a high spatial density.

The raw data acquired from surveying is thus considered, in this work, as input information. On this basis, the proposed methodological approach is structured as follows (Figure 4-1, Figure 4-2):

- Semantic segmentation via AI (Section 4.1);
- 2D/3D information propagation and H-BIM reconstruction (Section 4.2);
- Semantic structuring and transfer of localized information (Section 4.3).

Chapters 5-7 provide a detailed description of the assessed methods and related results obtained, in the respective phases. The different case

studies considered to validate the proposed methodological approach are provided in Section 4.4.


raw 2D/3D data
classes of typological elements
semantic knowledge: decay, materials, architectural components...
1 - Roof
2 - Molding
3 - Wall
4 - Arch
5 - Vault
6 - Capital
7 - Shaft
collaborative 2D/3D annotation via aïoli
DATA ACQUISITION
SEMANTIC SEGMENTATION VIA AI
feature extraction
artificial intelligence (ML)
classification
1
2

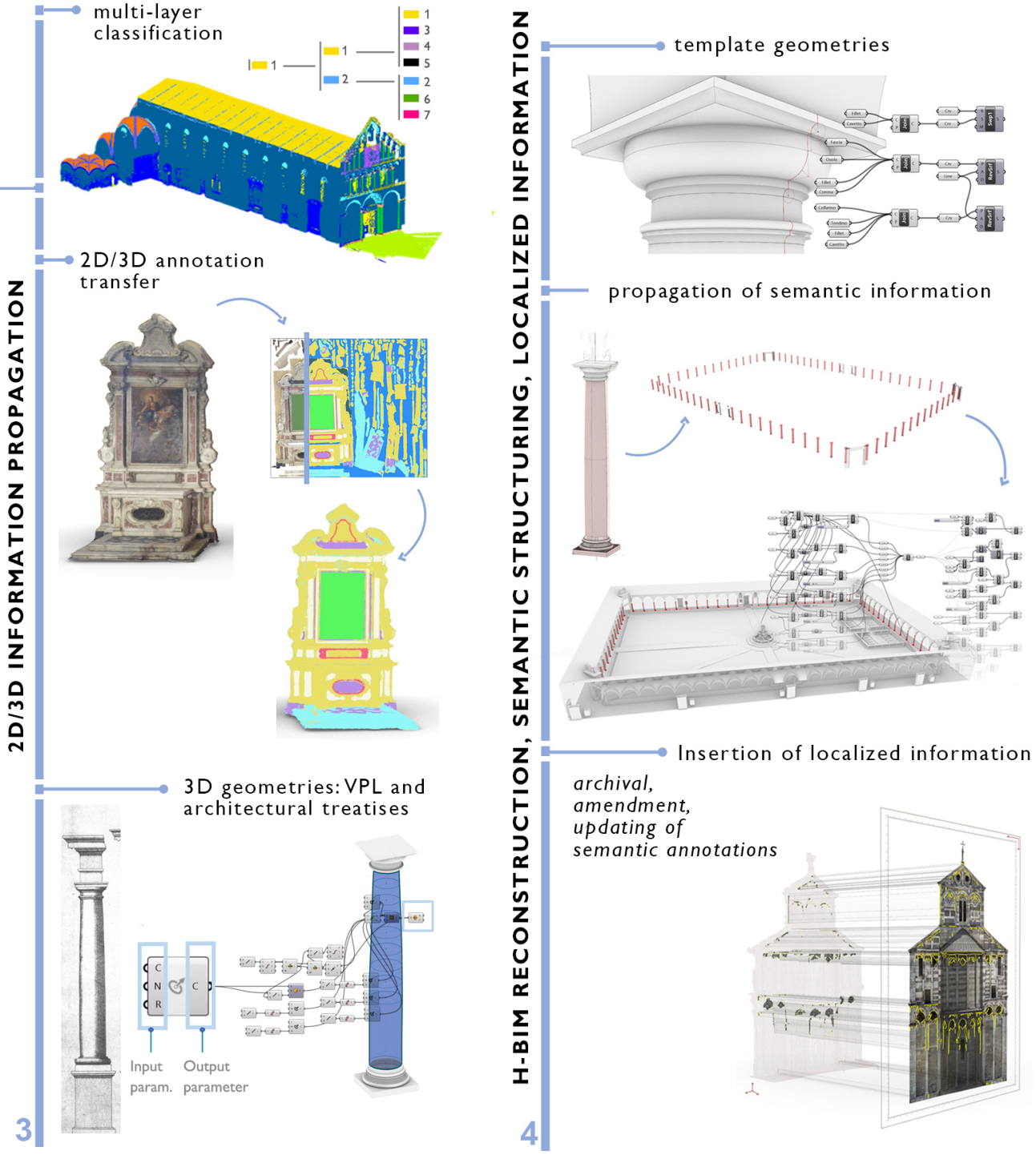


Figure 4-2.

Outline of the proposed methodological approach.

4.1. Semantic segmentation via Artificial Intelligence

In this phase, semi-automated systems are exploited for 2D and 3D data interpretation, to better characterize the surveyed architectural scene, improving the knowledge of the heritage object by description and recognition of its forms, materials, state of conservation.

Techniques derived from AI are considered for the construction of appropriate tools supporting the processing, management and semantic enrichment of digital heritage objects. Starting from the extraction of appropriate *features*, and by providing a relevant set of *training data*, an artificial intelligence is trained to classify, within digital models: the architectural elements that make up a historic building; the thematic mapping of degraded surfaces; damage maps and specification of intervention types; virtual reconstruction of lost elements and organization of archival documentation; definition of materials and their assembly, construction techniques, etc.

The semantic classification phase via ML thus allows the acquisition of 2D or 3D data that are semantically annotated, hence more easily interpretable. The methodology is studied both with reference to 2D data (photos, ortho-photos, drawings, archival images) than to 3D models (meshes, point clouds) of the architectural heritage and leverages supervised ML algorithms. In detail, the Random Forest is chosen as the reference algorithm following the successful tests by Grilli (2020), that demonstrated that RF outperforms other ML approaches for the classification of heritage objects. Due to the lack of a sufficient amount of training data and to the variety of historical buildings, DL approaches are discarded.

4.2. 2D/3D Information propagation

In this second phase, we focus on the development of new modalities of transfer and propagation of semantic annotations, from 2D to 3D data and vice versa.

At first, an implementation of the Aïoli platform with semi-automatic annotation methods is presented: combining the hybrid 2D/3D annotation system of the platform with semi-automatic (AI-driven) annotations, we improve the possibilities of propagation of semantic data to date available in Aïoli. Towards novel methods of reality-based documentation of heritage objects, a collaborative and web-based approach is developed, relying on AI for the semi-automatic 2D/3D annotation transfer.

Later on, several workflows are explored to ensure the propagation of information between different types of representation, e.g., meshes, point clouds, images, ortho-photos and so forth. The aim is to provide a possible future user with a range of feasible methods to be chosen as most suitable for data transfer, according to the case-specific needs and requirements for semantic data enrichment (shifting, e.g., from point clouds and meshes to images, UV maps, ortho-photos and vice versa).

Multi-layer and *multi-resolution* workflows are later offered to enable the transfer of semantic data at different levels of detail.

4.3. H-BIM reconstruction, semantic structuring and transfer of localized information

In case of classification of a heritage model into its recurring architectural components, the semantic segmentation is considered for the further construction of a H-BIM model.

We leverage in this phase classified point clouds to reconstruct, based on annotated 3D data, the *conceptual* geometries of each class of architectural elements detected (e.g., column, vault, molding etc.).

Referring to architectural treatises and exploiting generative modeling environments, a series of rules and graphic conventions are identified to reconstruct *template* reference geometries, and to later propagate the latter to all elements belonging to a same typological class.

The operation, if performed on the whole model, is essential in view of the realization of a complete information system of the H-BIM type, in which the connection with the original point cloud is preserved and where knowledge-related data can be stored.

Indeed, at a final stage, we hint at further possibilities of semantic enrichment of the model, by suggesting ways to combine raw, coarse and even localized information over the reconstructed digital object.

Some inputs are provided concerning ways of association of descriptive information on each reconstructed *template* geometry.

As such, the conceptual meta-model built might be linked to all other existing representation models (mesh, point clouds, photos, ortho-photos, 2D graphical drawing etc…), even enabling, in a H-BIM perspective, the association of the information to localized areas, so to note, e.g.: historical, architectural or artistic data, the presence of frescoes and decorative parts, degradation phenomena, crack patterns, repair and restoration interventions and so on.

4.4. Experimental context and choice of emblematic case studies

In order to develop the proposed methodological approach, a plurality of needs is considered:

- Tools for manipulating, filtering, exporting and extracting geometric and visual descriptors from point clouds and meshes;
- Tools for extracting geometric and visual descriptors from images while visualizing the related result;
- Functions to semantically annotate portions of 2D/3D data, for the definition of training sets of classified digital models;
- Supervised ML classifiers training functions, for the semi-automatic annotation of heritage models;
- Tools for the 2D/3D transfer and propagation of annotations;
- Tools for the geometric reconstruction of template geometries, allowing control and customization of shapes and 2D/3D graphics.

The experimental context is thus conceived in order to satisfy the variety of processes envisaged. In order to experiment and adapt existing tools to the new needs of description and documentation, the manipulation, integration and application of existing and already available software is taken into account. Beyond the development of a single user interface and beyond the construction of a unique environment where to manage all steps of the methodological approach, it is thus preferred to consider the application of existing licensed or open-source software so to investigate the best attainable workflow for the insertion, transfer and semi-automatic archival of semantic annotations on digital heritage models.

The experimental context is therefore aimed at giving a practical and direct answer, through the manipulation and adaptation of already available software, to the needs for documentation and information sharing via collaborative platforms. The designed workflow, in detail, is aimed at demonstrating the feasibility of directing existing software tools towards the improvement of Aïoli and H-BIM systems, and towards the management of collaborative annotation platforms in general, based on initial on-site acquisitions.

It is thus chosen to combine several software, each of them preferred -or adapted- every time according to the specific needs: among these, MATLAB and its Statistics and Machine Learning Toolbox (mathworks.com), ImageJ (danielgm.net) for the AI-based classification; CloudCompare

a.



1. Chiostro Granducale, Certosa di Calci



2. Chiostro Grande, Certosa di Calci



3. Chiostro del Capitolo, Certosa di Calci



4. Interior façade, Certosa di Calci



5. Notre-Dame Cathedral, Paris



6. Museo Nazionale di San Matteo, Pisa



7. Chiesa del Carmine, Pisa



8. Pisa Cathedral, Southern transept

...

(cloudcompare.org) for point cloud processing and feature extraction; Rhino and its visual programming interface Grasshopper (rhino3d.com) for the transition to H-BIM models; Autodesk Revit (autodesk.it) and the plug-in Rhino.Inside.Revit (rhino3d.com/inside/revit/1.0/) for the creation of template geometries as well as for the management of BIM environments.

Finally, as mentioned above, the geometric and visual aspects of digital heritage data are combined with the semantic ones through the use of collaborative annotation platforms as Aïoli and the H-BIM environments.

For the study of such systems, the present research project builds on the work conducted, in recent years, respectively by the ASTRO laboratory of the University of Pisa, the LISPEN laboratory of the ENSAM of Aix-en-Provence and the research unit UMR MAP 3495 of the CNRS of Marseille. The topic of this Ph.D. thesis thus draws on most recent studies carried out by the Italian and French research institutions involved.

On the one hand, the experience gained by the ASTRO laboratory on Scan-to-BIM processes and on the creation of information models from the 3D surveying of the built environment; on the other hand, the experience gained by the MAP laboratory of Marseille on the topics of digitization and virtual reconstruction of the built heritage and on the development of the hybrid 2D/3D annotation platform Aïoli for the documentation of heritage artefacts.

The cooperation between Italian and French institutions has thus led to a mutual enrichment in terms of research activities and dissemination of scientific work in the supranational context, the latter recognized and supported by the Université Franco-Italienne, in the context of the Vinci2019 prize for Ph.D. thesis in co-tutelle agreement (Chapter

b.

	Ground-based photogrammetry	UAV-based photogrammetry	Laser scanning
1. Chiostro Granducale, Certosa di Calci	x	x	x
2. Chiostro Grande, Certosa di Calci		x	x
3. Chiostro del Capitolo, Certosa di Calci	x	x	x
4. Interior façade, Certosa di Calci	x	x	x
5. Notre Dame Cathedral, Paris	x	x	x
6. Museo Nazionale di San Matteo, Pisa	x		
7. Chiesa del Carmine, Pisa	x		x
8. Pisa Cathedral, Pisa	x	x	x
⋮			

Figure 4-3.

Many relevant case studies considered of the Italian and French heritage (*a*) and related surveying techniques applied (*b*).

II-Mobility fundings for Ph.D. thesis in co-tutelle).

Again, in the wake of the international agreement, the case studies on which the approach is tested are related to architectural objects making part of the Italian and French cultural heritage. The digital models considered indeed refer to historical buildings, surveyed by laser scanner, ground-based and drone-based photogrammetry, as well as by integration of these techniques, alongside traditional surveying instruments: among the case studies examined, the most representative ones are the Pisa Charterhouse and the Notre-Dame Cathedral in Paris.

Besides, several churches, museums, historical buildings on the Italian and French national territory are considered. An overview of these case studies is provided in Figure 4-3.

Many settings are related to the monumental complex of the Pisa Charterhouse (ita: *Certosa di Pisa*), a monastic complex whose original structure dates back to the thirteenth century and which today houses the Museum of the Charterhouse and the Natural History Museum of the University of Pisa. The current layout of the monastery results from the ceaseless expansion and renovation that took place since its very foundation and many spaces of the Carthusian life -where hermitical and *coenobitical* life were harmoniously fused together-, can still be found today: the *Case basse* (low houses), the *Corte d'onore* (main courtyard), the Church and Cenoby define a progressive ascent towards the *desertum* at the heart of the complex, the Cloister of the Fathers or main cloister (Croce et al., 2021; Bandinelli et al., 2016).

Since year 2013, the Pisa Charterhouse has been subject to an intense work of surveying, analysis, documentation, promoted by the University of Pisa and coordinated by Professor Marco Giorgio Bevilacqua,

under the project *Studi conoscitivi e ricerche per la conservazione e la valorizzazione del Complesso della Certosa di Calci e dei suoi Poli Museali* (en: Cognitive studies and research for the conservation and enhancement of the Calci Charterhouse complex and its museum complexes).

The architectural survey of the complex was conducted by the ASTRO laboratory, with application of integrated laser scanner and photogrammetric methodologies.

Another significant case study is the Notre Dame Cathedral in Paris, the gothic church that is a symbol of the French capital. After the fire that struck the cathedral in April 2019, the CNRS has been designated by the French Ministry of Culture to guide the organization of a scientific project in support of the restoration activities on the Cathedral. Among the nine thematic working groups set up, the one on Digital data (fr.: *données numériques*), coordinated by Livio De Luca, has the ambition to build a *digital ecosystem*. The latter should be put into practice in order to assist the scientific study and the restoration of the cathedral by progressively integrating the data, information and knowledge coming from several actors, thanks to: the collection and integration of existing data, the production of new ones, their sharing and permanent archiving, their structuring and semantic enrichment (Notre-Dame Science, 2021).

For the different case studies considered on the French and/or Italian territory, the proposed methodological approach has been tested and adapted each time case by case. In detail, it has to be noted that we do not propose a single end-to-end workflow that applies to each case study, but rather the general approach defined and divided into the three respective phases -i.e., semantic segmentation via AI, 2D/3D information propagation, H-BIM reconstruction, semantic structuring and transfer of localized information-, can be adapted, integrated and validated each time, according to the case-specific needs for description and documentation.

Chapter 5

Semantic segmentation via Artificial Intelligence

This chapter presents methods and results obtained for the first phase of the studied workflow, concerning the use of AI to semi-automatically classify digital heritage objects.

As detailed in Section 4.1, the reading and interpretation of raw surveying data of a heritage object, while being strictly natural for the human brain, is highly challenging for computers. The classification based on semantics, even before the 3D modeling and reconstruction phase, is therefore essential to assist the interpretation and dissemination of surveying data, and to characterize digital information, e.g., in terms of material mapping, degradation mapping or identification of the architectural components of a building.

As regards this latter category, we are primarily interested in those buildings of historical heritage in which recurring typological and architectural elements can be recognized (Section 5.1); as such, the semantic description can be guided by the study of the building's morphology.

In Section 5.2, AI-based methods for the classification of 2D/3D digital heritage data are proposed, either based on the geometry of the represented elements (*geometry-based*) or on the texture characterization (*texture-based*).

The results on annotated 3D data of the considered case studies, as well as the evaluation of the proposed approach in terms of reliability of the predictive models built, are presented in Section 5.3.

5.1. Semantic description of the building's morphology

The operation of semantic description based on the subdivision of the building into its architectural components has a long history.

The classification process is in fact the very one that led Vitruvius and other major essayists to investigate, and convey to posterity, the proportions of architectural orders, thus codifying languages and models of classical architecture (Croce et al., 2021). In this logic, the current work stems from the consideration that some recurring architectural typologies and shapes are still recognizable in the morphological and constructive complexity of historical constructions. This allows for effective differentiation and, as a result, representation by reconstruction of a *shape grammar*.

As for treatises of classical architecture, a common language or vocabulary is to be established in such a way that all elements making up a historic building are defined. This vocabulary should consider all intrinsic features of each architectural element, and, at the same time, it should offer a unique basis for a common language.

In the digital modeling domain, as already stressed in Paragraph 3.3.5.3, this need has been formerly translated into the creation of ontologies and shared libraries of architectural components to appropriately represent the parts of a building (Quattrini et al., 2015; Oreni et al., 2013; Murphy et al., 2013; De Luca et al., 2007), again based on semantic rules and classifications derived from treatises. De Luca et al. (2007) emphasized in particular that the shape grammar is in any case guided by the objective of analysis, thus the description may change, e.g., depending on the physical or material subdivision, on the distinction of architectural components or on the material mapping (Figure 5-1). This ambiguity might be solved by considering: the use of a vocabulary to isolate elements from the building, the identification of a logic to regroup the elements into a coherent unit, the appropriate choice of the representation type (Figure 5-2).

Starting from the geometrical analysis of the elements that compose a building and by considerations on the architectural treatises, a semantic layer can thus be added and connected with the description of the morphology of an existing construction. This enables the improvement of the geometric and semantic interpretation of a heritage model, also with a view to reconstruction in HBIM environments or sharing via the Aïoli platform. A correspondence between the semantically classified point cloud and the digital model on Aïoli or in BIM can in this way be created, thanks to

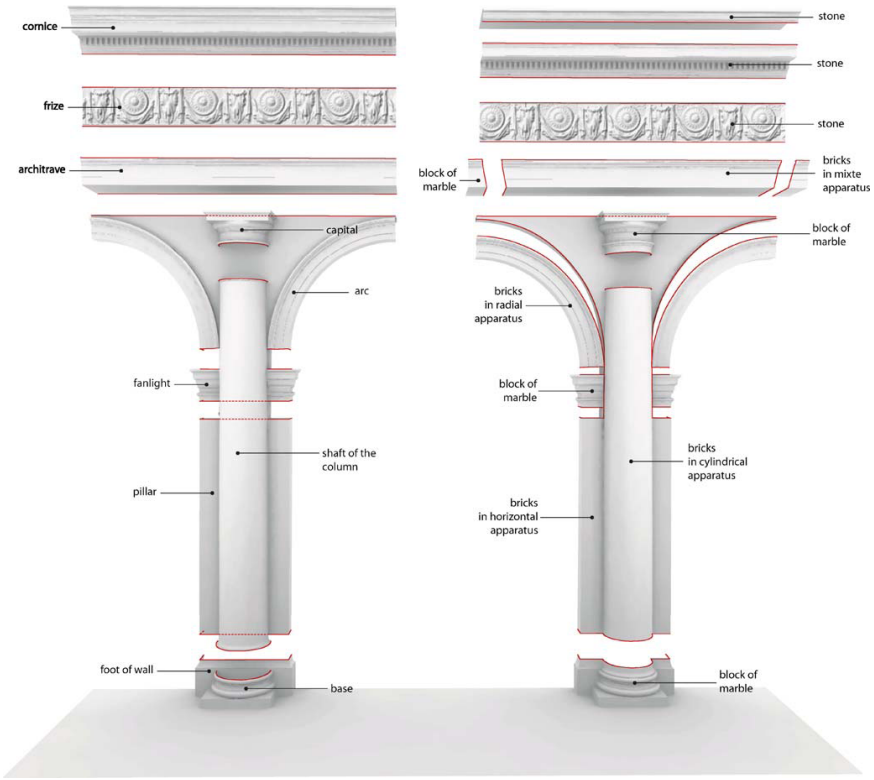


Figure 5-1.

Morphological segmentation of the object according to two different analyses: the vocabulary of the architect (*to the right*) and the construction techniques (*to the left*).

(De Luca et al., 2007)

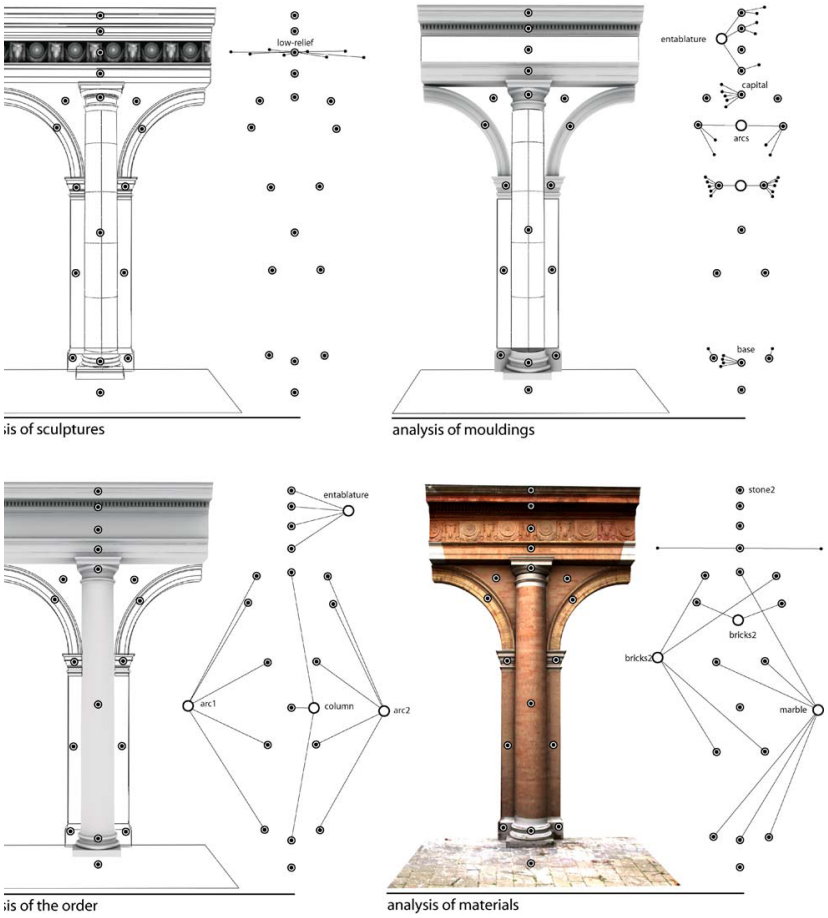


Figure 5-2.

Four "points of view" on the studied object elaborated starting from the same morphological cutting.

(De Luca et al., 2007)

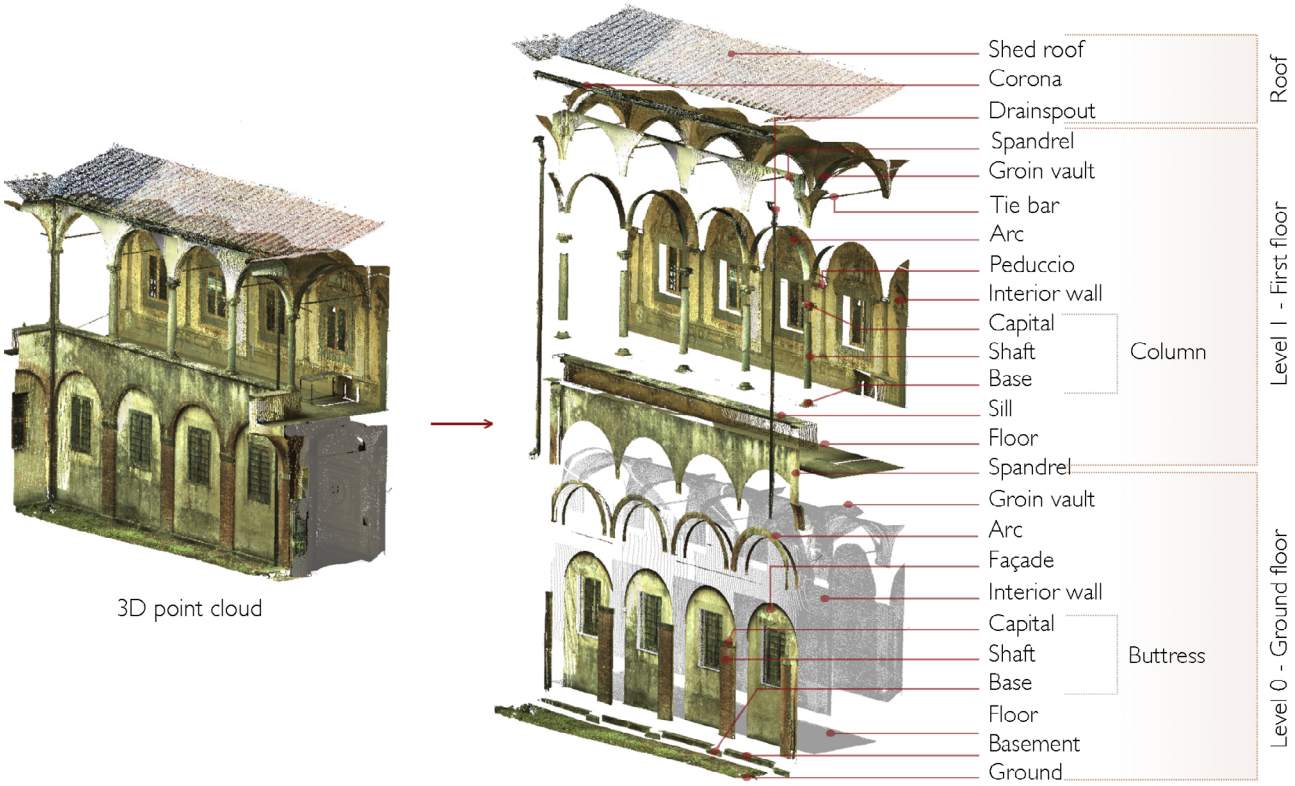


Figure 5-3.

Decomposition of a raw point cloud based on the identification of architectural components.

e.g.: Grand-Ducal cloister, Pisa Charterhouse

the decomposition of the architectural object.

By analyzing the codes of architectural representation, in fact, one can on one hand identify relevant information to be extracted from survey, while on the other hand choose the process of construction better suited to guarantee the geometrical restitution of elements (De Luca et al., 2007), even depending its expertise and scope (Cera, 2018).

The semantic classification of the architectural elements is therefore to be considered before the modeling phase, so as to define a shape grammar approach guiding the deconstruction of the raw surveying data. This decomposition of the digital object (Figure 5-3) follows the logic of BIM reconstruction and generative design, whereby the architectural components of the building are first identified and then reconstructed each, following a sequence of algorithms defining proportional and constructive rules.

The transcription of the rules handed down by historical treatises in the H-BIM language and through generative modeling, starting from classified digital heritage data, becomes a natural consequence of these considerations.

5.2. Machine Learning approaches for 2D/3D semantic segmentation. Methods

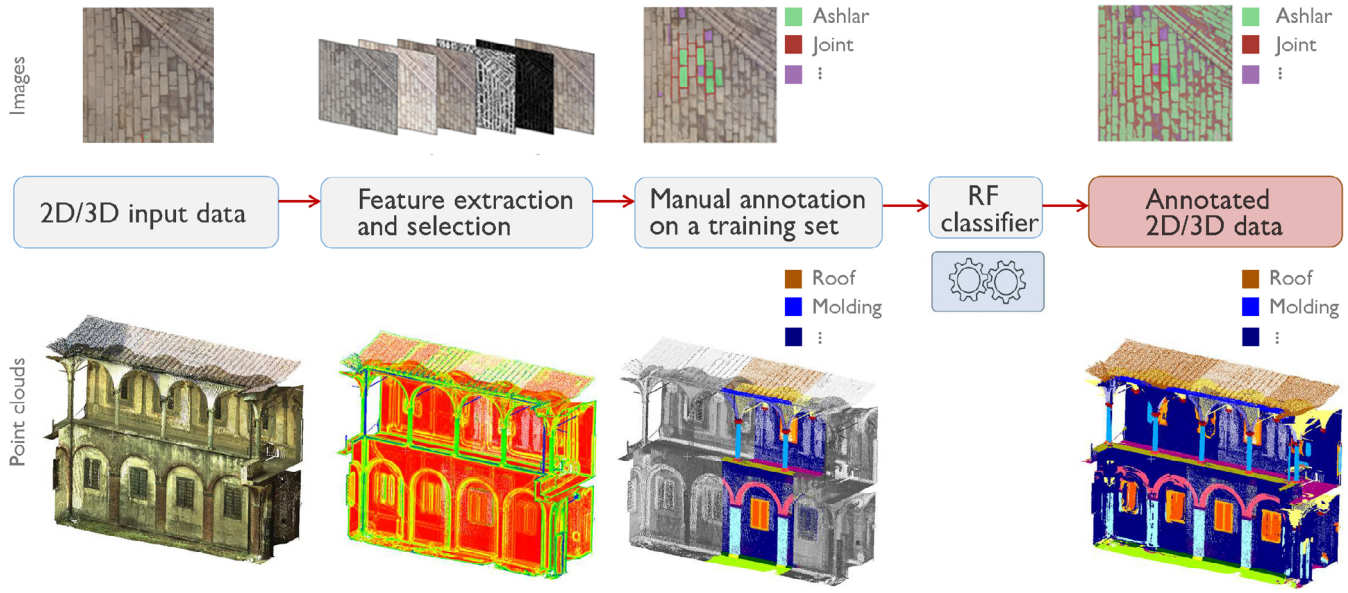


Figure 5-4.

Overview of the ML approach for 2D/3D semantic segmentation.

e.g.: vaults of the Notre-Dame Cathedral in Paris (*above*); Grand-Ducal cloister of the Pisa Charterhouse (*below*).

The proposed methodological approach (Figure 5-4) leverages AI for the semantic segmentation of 2D (images, UV maps, ortho-photos) and 3D (point clouds, meshes) heritage data. Starting from the extraction of appropriate features, and providing a suitable set of training data, an artificial intelligence is trained to recognize and classify digital heritage, e.g., in terms of thematic mapping of degraded surfaces, framing of damages, definition of materials and their assembly, classification of the different architectural elements that make up a historic building, etc. The methodology is studied with reference to both 2D/3D digital representations of the historical architectural heritage and it is based on a supervised ML algorithm, the Random Forest.

The semantic segmentation procedure, relying on AI, involves the following steps:

i. Neighborhood selection and feature extraction;

ii. Feature selection;

iii. Manual annotation on a reduced portion of the dataset (*training set*), to identify classes of elements;

iv. RF classifier, and consequent accuracy evaluation;

v. Annotated 3D point cloud.

The procedure is in line with the workflow shown by Hackel et al., (2018, 2016a) and Weinmann (2016) for 3D scene interpretation, and by Grilli et al. (2019a) for its former applications to the heritage domain (Croce et al., 2021).

At first, a set of features is extracted in a chosen local neighborhood of each 3D point or image pixel (*i*). The choice of appropriate features - and consequently, the choice of the local neighborhood in which they are computed, is crucial for the ML model to be more amenable to the classification task. The predictive model, in fact, is built to make predictions by recognizing the features that distinguish one class of elements from another. As the initial set of features may appear redundant or too large to be managed, the features are iteratively selected (*ii*).

Classes of recurring architectural elements are identified and labeled on the training set. A multi-scale classification using a RF enables one to iteratively select the most relevant features and perform the classification process by considering a subset of features each time. Steps (*ii*) and (*iii*) are thus strictly interrelated.

Once built thanks to the training set of annotated data, the predictive model is applied to the remaining part of the point cloud or image (non-manually annotated), so to semantically label the classes of typological elements in the rest of the dataset. The accuracy of the classifier is finally assessed (*iv*).

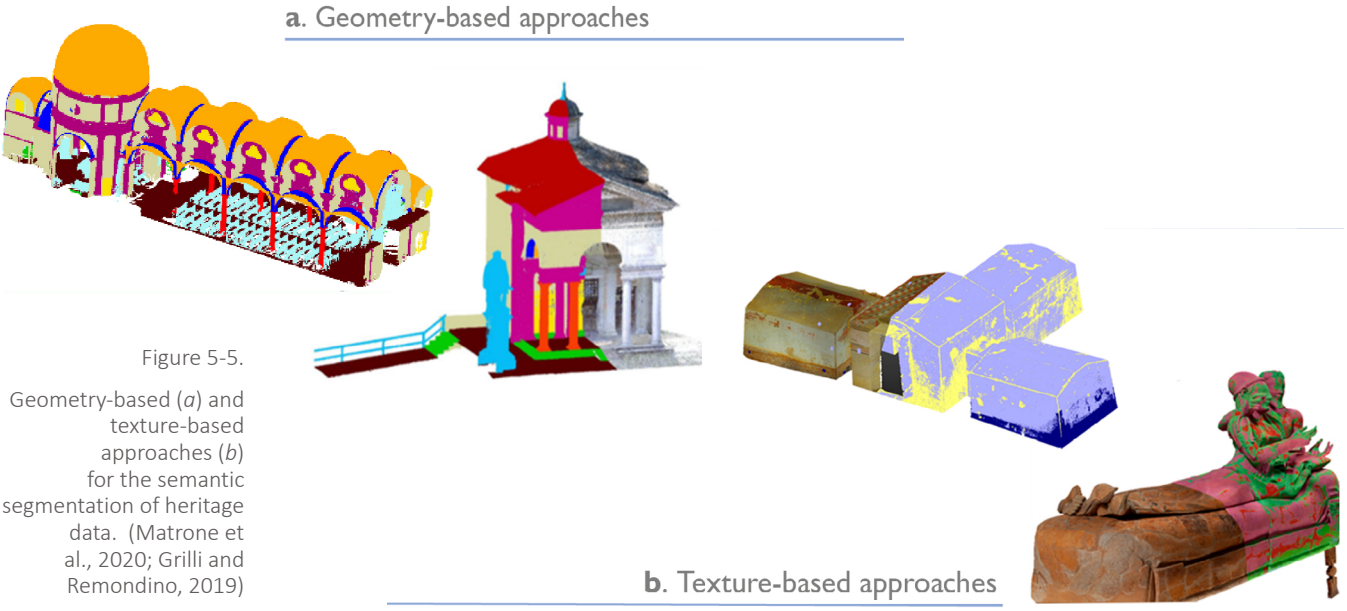


Figure 5-5.

Geometry-based (*a*) and texture-based approaches (*b*) for the semantic segmentation of heritage data. (Matrone et al., 2020; Grilli and Remondino, 2019)

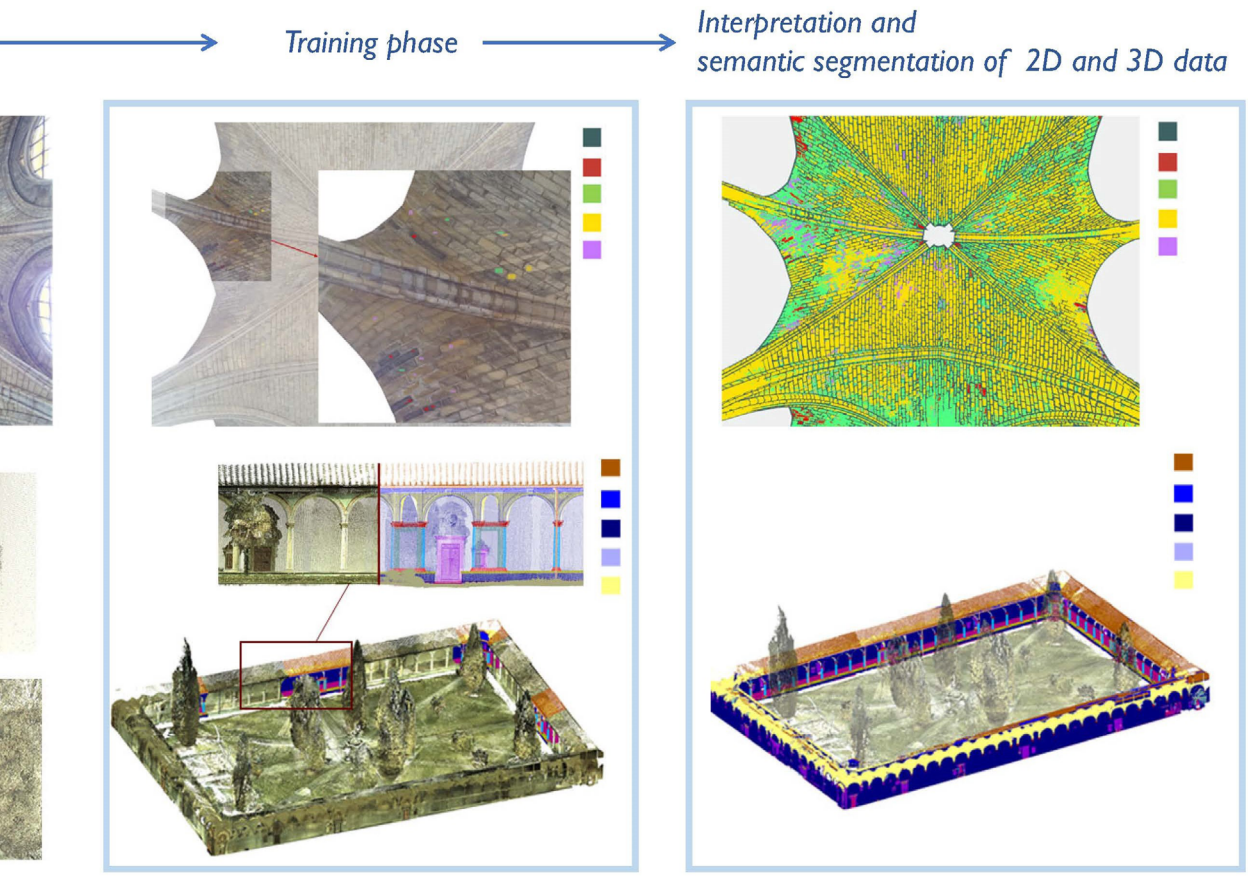


Figure 5-6.

A reduced portion of 2D and 3D input data is manually labeled, and the ML-based segmentation enables the classification of the whole dataset.

e.g.: Vaults of the Notre-Dame Cathedral, Paris (*above*); main cloister of the Pisa Charterhouse (*below*).

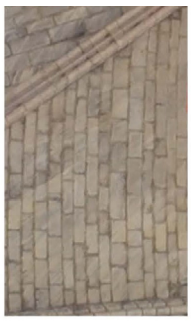
At the end of the process, a semantically annotated dataset is obtained (*v*).

5.2.1. Geometry-based and texture-based approaches

The classification process may follow different approaches (Figure 5-5, respectively based on the use of radiometric data of the model's textures (*texture-based approach*) or on the analysis of the geometric properties of point clouds (*geometry-based approach*) (Grilli and Remondino, 2019). Both approaches comprise the five steps of the procedure described in Section 5.2.

In the case of texture-based approaches, segmentation is mainly driven by the color data that can be extracted from a textured mesh model (pixel-wise classification). For instance, if maps of degradation, surface alterations and materials are to be derived from the digital data, the purely metric model itself may reveal not to be sufficient to describe proper colorimetric characteristics of the real object, hence texture-based approaches are preferable. On the other hand, in segmenting parts of the model with uniform staining (e.g., in the case of objects with similar materials or even with the same material), the color data

alone does not allow proper distinction of data. In this case, the geometry-based approach may be preferred, e.g., in view of classifying the architectural components of a building, as in this case the semantic segmentation mainly follows the geometric properties of the digital model (3D point classification).



• **2D** →

In this work, both methods are considered and the choice between one or another is guided each time by the case-specific description and representation needs (Figure 5-6): in general, for instance, geometry-based approaches are better suited to the distinction of the architectural components of a building and can be applied to both photogrammetry and laser scanner models, while texture-based approaches appear more suitable in the case of photogrammetric models, for material and degradation mapping purposes. For the latter, we will see (Section 5.3.2) that the classification operation can be performed in 2D, on the UV map extracted from the mesh model, and then projected onto the 3D model: this allows to work in a 2D-3D reprojection space where it is possible to exchange annotation information between image and mesh and vice versa.

• **3D** →



In this work, both geometry-based and texture-based approaches are considered and the choice between one or another is guided each time by the case-specific description and representation needs (Figure 5-6): in general, for instance, geometry-based approaches are better suited to the distinction of the architectural components of a building and can be applied to both photogrammetry and laser scanner models, while texture-based approaches appear more suitable in the case of photogrammetric models, for material and degradation mapping purposes. For the latter, we will see (Section 5.3.2) that the classification operation can be performed in 2D, on the UV map extracted from the mesh model, and then projected onto the 3D model: this allows to work in a 2D-3D re-projection space where it is possible to exchange annotation information between image and mesh and vice versa.

In the case of geometry-based segmentation, the feature extraction is performed with CloudCompare (danielgm.net), while the ML classifier is trained in Matlab, starting from the Statistics and Machine Learning Toolbox (en.mathworks.com).

In case of texture-based approaches, the classification is run via the



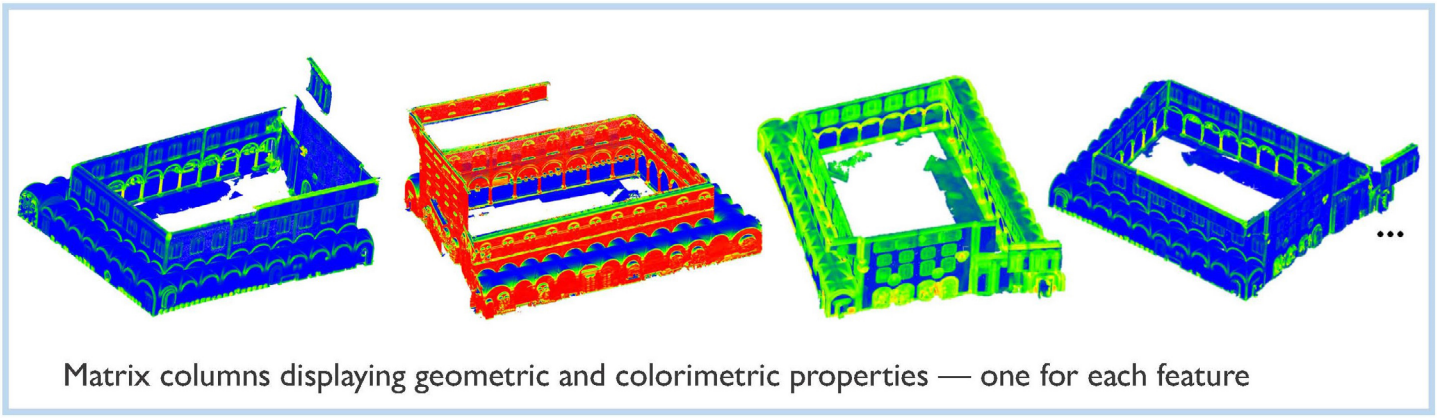


Figure 5-7.

Samples of feature extraction on 2D (*above*) and 3D data (*below*).

e.g.: Vaults of the Notre-Dame Cathedral, Paris (above); cloister of the National Museum of San Matteo in Pisa (below).

Fiji plug-in of ImageJ called *Trainable Weka Segmentation* (imagej.net/plugins/tws/), i.e., an image-processing tool combining a collection of ML algorithms with a set of selected image features to produce pixel-wise segmentations.

5.2.2. Feature extraction

The feature extraction, and subsequent selection, are fundamental steps in the classification process as they enable the association of initial raw data with a given set of accessory information. This information will help distinguishing between classes of elements, thus assisting learning and generalization steps.

Feature extraction is based on the calculation of visual and geometric descriptors (Figure 5-7): the latter can either be 2D descriptors, as the ones illustrated in Section 3.2.1, or 3D descriptors, as the ones described in Section 3.2.2.

In the case of 2D information, the features are displayed as a stack of

additional images-one per feature-, connected to the input image.

For the case of 3D data, each feature constitutes an additional column of the point clouds' matrix, and the variation of the feature value in a given dataset can be graphically visualized by the use of color scales.

5.2.2.1. 2D features.

In 2D, the segmentation problem is in the form of a pixel classification problem as each pixel of the image can be classified as belonging to a specific segment or class.

The features extracted from an input image can be (Arganda-Carreras et al., 2014):

- *Image adjustments*, such as hue, brightness and saturation (Paragraph 3.2.1.2);
- *Edge detectors*, aiming at indicating boundaries of objects in an image (Paragraph 3.2.1.3);
- *Texture filters*, including filters such as minimum, maximum, median, variance, entropy, structure tensor, etc. (Paragraph 3.2.1.4);
- *Noise reduction filters*, such as Gaussian blur, bilateral filter, Anisotropic diffusion, Kuwahara and Lipschitz (Paragraph 3.2.1.5).

The features are computed each time for given values of the pixel neighborhood.

5.2.2.2. 3D features.

In case of 3D data, the classification is performed by considering as distinctive features:

- *Covariance features*, enclosing the mutual geometric relations between each 3D point of the cloud and the other points that are its nearest neighbors;
- *Changes of curvature*, calculated based on (Zhang, 2008);
- *Radiometric features*, expressing the color value associated to each 3D point, when available;
- *The Z coordinate*, as height feature.

In detail, covariance features (Table 5-1) are geometric features derived from the eigenvalues λ_1, λ_2, λ_3 of the covariance matrix, computed within a local neighborhood of the 3D point. The covariance matrix, also referred to in literature as 3D structure tensor (Rodriguez-Gonzalvez

and Rodriguez-Martin, 2019; Hackel et al., 2016b), is given by:

$$[Cov] = \begin{bmatrix} Var[A_1] & Cov[A_1, A_2] & Cov[A_1, A_3] \\ Cov[A_1, A_2] & Var[A_2] & Cov[A_2, A_3] \\ Cov[A_1, A_3] & Cov[A_2, A_3] & Var[A_3] \end{bmatrix}, \quad (5\text{-}1)$$

Let A be a point of the Euclidean space, $\mathcal{N}(A;\varrho)$ be a spherical local neighborhood of point A, whose radius is ϱ, and $A_k=(A_{1k},A_{2k},A_{3k})$ a point belonging to $\mathcal{N}(A;\varrho)$.

Referring to $\mathcal{N}(A;\varrho)$, the elements of the covariance matrix are given by:

$$Var[A_i] = \frac{1}{N} \sum_{A_k \in \mathcal{N}(A;\varrho)} (A_{i_k} - \bar{A}_i)^2 \quad (5\text{-}2)$$

$$Cov[A_i, A_j] = \frac{1}{N} \sum_{A_k \in \mathcal{N}(A;\varrho)} (A_{ik} - \bar{A}_i) \cdot (A_{j_k} - \bar{A}_j) \quad (5\text{-}3)$$

where the sum of $A_k \in \mathcal{N}(A;\varrho)$ is extended to the N points A_k of the cloud belonging to $\mathcal{N}(A;\varrho)$.

A_{ik} is the i-th component of A_k (in x, y and z dimensions), and $\bar{A}_i$ is the mean of all observations in the i dimension.

The covariance matrix (5-1) is a symmetric matrix. The non-negative eigenvalues of the matrix, ordered according to $\lambda_1 \geq \lambda_2 \geq \lambda_3$, correspond to an orthogonal system of eigenvectors $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$ and are leveraged to calculate the nine 3D shape features displayed in Table 5-1, for each 3D point.

Feature name	*Expression*	
Linearity	$L_\lambda(\varrho) = \frac{\lambda_1 - \lambda_2}{\lambda_1}$	(5-4)
Planarity	$P_\lambda(\varrho) = \frac{\lambda_2 - \lambda_3}{\lambda_1}$	(5-5)
Sphericity	$S_\lambda(\varrho) = \frac{\lambda_3}{\lambda_1}$	(5-6)
Omnivariance	$O_\lambda(\varrho) = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3}$	(5-7)
Eigenentropy	$E_\lambda(\varrho) = -\sum_{i=1}^{3} \lambda_i \ln(\lambda_i)$	(5-8)
Surface variation	$SV_\lambda(\varrho) = \frac{\lambda_3}{\sum_{i=1}^{3} \lambda_i}$	(5-9)
Sum of eigenvalues	$\Sigma_\lambda(\varrho) = \sum_{i=1}^{3} \lambda_i$	(5-10)
Anisotropy	$A_\lambda(\varrho) = \frac{\lambda_1 - \lambda_3}{\lambda_1}$	(5-11)
Verticality	$V_\lambda(\varrho) = 1 - \lvert([0\ 0\ 1], \mathbf{e_3})\rvert$	(5-12)

Table 5-1. Geometric features considered (Hackel et al., 2016).

More specifically, linearity (5-4), planarity (5-5) and sphericity (5-6) are dimensionality parameters, as they express the probability of each 3D point to be labeled as 1D (linear), 2D (planar) or 3D (volume) structures respectively (Weinmann et al., 2015, 2013). Omnivariance (5-7) and eigenentropy (5-8) are a measure of the order/disorder of points within the chosen local 3D neighborhood (Weinmann, 2016).

Other descriptors as surface variation (5-9), sum of eigenvalues (5-10), anisotropy (5-11) and verticality (5-12) can be calculated – the last one depending on the eigenvector $\mathbf{e_3}$ of the covariance matrix.

Features are computed each time at a given value of the radius of the local spherical neighborhood of a 3D point, ϱ. For the same dataset, the same feature can vary significantly as the value of ϱ varies (Figure 5-8).

Also, note here that two point clouds representing a same object but acquired with different point densities may yield different feature values, again due to the dependence on the value of ϱ.

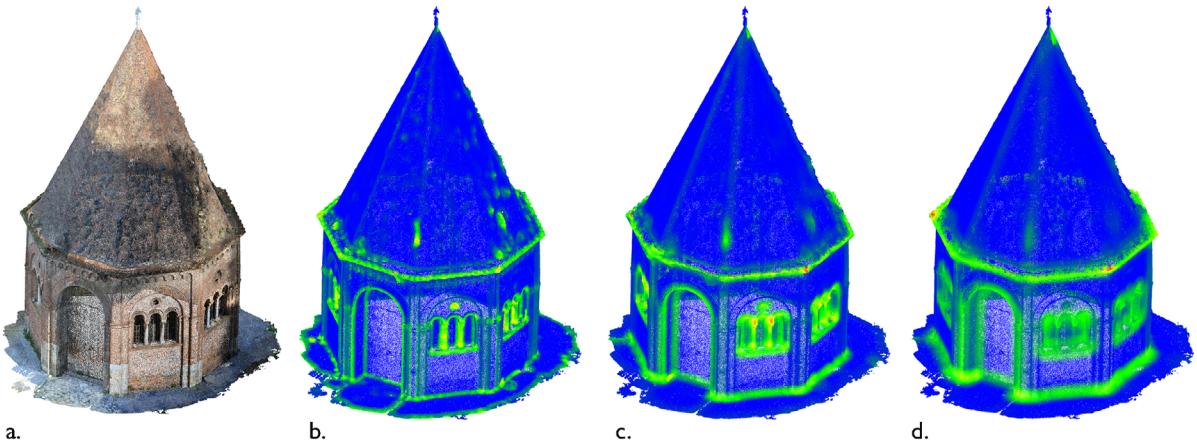


Figure 5-8.

Cappella Sant'Agata in Pisa (*a*): Surface Variation feature computed at different values of the local neighborhood radius: 0,2 m (*b*); 0,4 m (*c*); 0,6 m (*d*).

5.2.3. Feature selection

The choice of features used for training is fundamental in determining the heterogeneity between classes, therefore the performance of a classifier.

For the case of 3D data, a multi-scale approach is adopted first, as features are computed with varying values of the local neighborhood radius. Then, in order to reduce the training time, and also in order to improve the classifier performance, a feature selection phase is considered, so to remove redundant and less relevant features and consequently train the RF with a reduced set of features.

The choice of the RF is indeed motivated by the fact that this algorithm has an indicator for feature selection: when training a tree, it is possible to determine how much each feature decreases the impurity (the more a feature decreases the impurity, the more important the feature), and consequently average the impurity decrease across all trees to estimate the final importance of each feature (*towardsdatascience.com*).

At first, different sizes of the local neighborhood radius are established. The features are extracted considering these varying values of ϱ and the RF is originally trained with the whole set of features. The choice on the values of ϱ by which the features should be computed is driven each time by considerations made on the recurrent dimensions of the elements found in the dataset, relating, e.g., to the thickness of certain architectural moldings and to the radii and diameters of the columns. Moreover, as demonstrated by (Kaiser et al., 2019), we consider that lower values of the neighborhood radius provide better performance in describing finer details, while higher values apply a sort of smoothing filter.

At the end of the first training phase (with all features), a relevance score can be extracted, showing the relative importance of each

feature to the classification task. Starting from the importance ranking extracted, redundant and less relevant features are removed, and the RF is trained with a reduced subset of features. The classification results for the different features combinations are finally compared.

5.2.4. Manual annotation of a training set

In the case of digital heritage data, the specificity and complexity of the architectural heritage under study make it necessary to train the predictive model considering labeled data as input. The manual annotation of a certain portion of the data is indeed demanded in supervised ML approaches.

As previously shown (Section 3.3), the manual annotation is intended as the twofold operation of manually selecting a portion of the data and associating a semantic label to that information. The manual annotation operation is performed in our case on a small but significant portion of the data, to identify the classes into which the dataset is to be subdivided. For a proper classification, the chosen portion of the data should be representative of all classes under investigation.

In the case of image classification, the manual annotation takes place via the selection of regions or segments that comprise sets of pixels sharing certain visual characteristics, and via the subsequent assignment to a specific label. A subset of input pixels that has been labeled is represented in the feature space and then used as a training set for the classifier.

In the case of point clouds, instead, the manual annotation is performed through CloudCompare. The point cloud is segmented according to the classes to be identified, and a class index is associated to each point of the segmented part.

In this way, the part of the point cloud that is manually segmented will be organized with a structure analogous to that of Figure 5-9: the n points of the cloud will be associated with coordinates, color values (where available), m geometric features (calculated in a local neighborhood radius ϱ), and a final column corresponding to the class index. This last value will be known for the manually annotated portion of the dataset,

Figure 5-9.
Matrix structure of a point cloud.

Point	X	Y	Z	R	G	B	Intensity	Geometric features $f(r)$	Label (response value)
1	X_1	Y_1	Z_1	R_1	G_1	B_1	i_1	$f_{11}(\varrho) \dots f_{1m}(\varrho)$	l_1
2	X_2	Y_2	Z_2	R_2	G_2	B_2	i_2	$f_{21}(\varrho) \dots f_{2m}(\varrho)$	l_2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮ ⋱ ⋮	⋮
n	X_n	Y_n	Z_n	R_n	G_n	B_n	i_n	$f_{n1}(\varrho) \dots f_{nm}(\varrho)$	l_l

while it will be deduced for the rest of the dataset to be automatically classified.

5.2.5. Random Forest classifier and evaluation of the trained model

The classifier can be trained using the selected features, together with the manually labeled set, as input data. For the ML framework, the choice to train a RF classifier (Figure 5-10) is in continuity with the first applications of such supervised learning algorithms to the heritage domain (Grilli et al., 2019b).

The RF is an ensemble method: it creates multiple decorrelated decision trees, resulting in a *forest of trees*, and then combines the output class of each individual learner, selecting the best solution by means of voting (Paragraph 3.3.6.2). The parameters to be chosen in training a RF are the number of trees $\mathcal{N}_{trees}$ and the number of variables to be selected and tested for the best split when growing the trees (Ni et al., 2017). This algorithm was chosen because:

- Its combination of learning models increases classification accuracy, averaging noisy and unbiased models.
- It requires a smaller amount of annotated data for learning, compared to other ML algorithms.
- It presents estimates for features' importance: the less

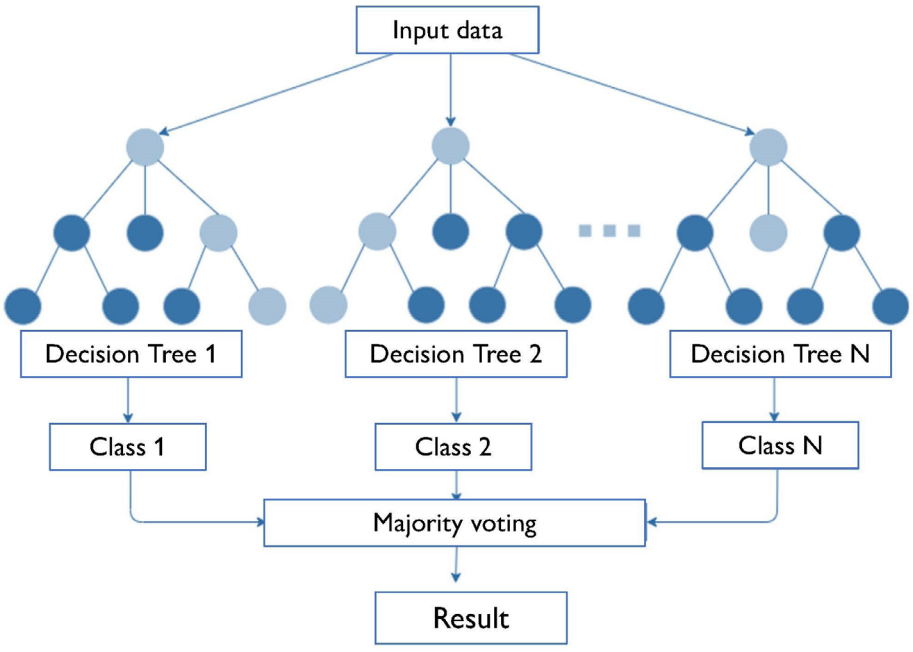


Figure 5-10.
Workflow of the RF.

significant variables can be removed so that the model is trained on a subset of features, greatly lessening the time for learning and increasing predictive accuracy (Croce et al., 2021).

For the assessment of the classifier constructed, a small portion of unseen data is used as *evaluation dataset*, where to test the correspondence between the true-manually annotated- labels and the predicted ones.

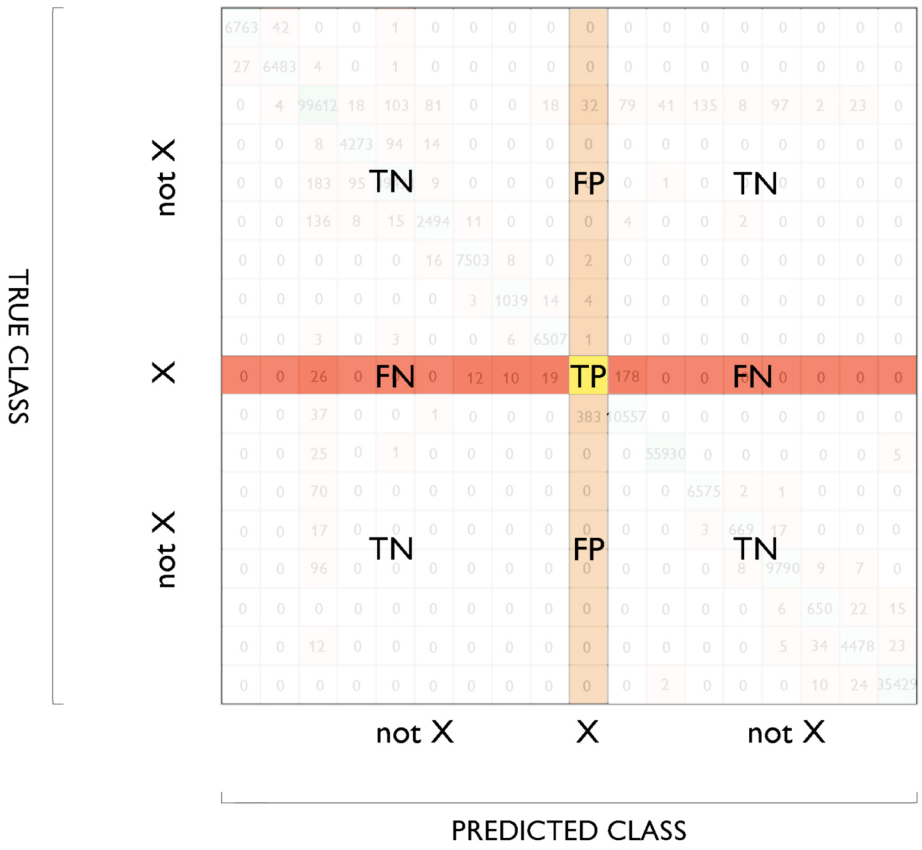
The correspondence is sorted in the form of a summary, the so-called *confusion matrix*, which provides a measure of the number of correct and incorrect predictions, class by class.

The rows of the confusion matrix display the true – manually annotated – values, while the columns display the predicted ones (Figure 5-11), for each class. As a consequence, the on-diagonal elements stand for the True Positive (TP) values (correctly classified instances of the dataset), while the off-diagonal elements stand for misclassified values: True Negatives (TN), False Positives (FP) and False Negatives (FN) values.

Figure 5-11.

Structure of a confusion matrix.

Based on the numbers or percentages of TP, TN, FP and FN, many performance measures can be derived from the confusion matrix.



Precision, *Recall*, *Overall accuracy,* and *F-measure* are commonly used parameters in the assessment of ML models. They are expressed as follows:

$$Precision = \frac{TP}{TP + FP} \tag{5-13}$$

$$Recall = \frac{TP}{TP + FN} \tag{5-14}$$

$$Overall\ accuracy = \frac{TP\ +\ TN}{TP + TN + FP + FN} \tag{5-15}$$

$$F - measure = 2 \cdot \frac{Recall\ \cdot Precision}{Recall + Precision} \tag{5-16}$$

Among these, *Precision* is the ratio of *TP* values and the total number of positives predicted by a model, while *Recall* is the ratio of correctly predicted observations to all observations in the actual class.

Overall accuracy represents the number of correctly classified data over the total number of data instances. Finally, *F-measure,* also known as *F1-score*, represents the weighted average of *Precision* and *Recall*, and thus takes into account both *FP* and *FN* values.

5.3. Results: annotated 2D/3D data

5.3.1. Geometry-based classification results

For the case of geometry-based approaches, the input information is represented by the 3D point cloud, as the classification is mainly driven by the recognition of geometric properties of the digital model and by the identification of relevant architectural elements (Section 5.2.1).

The approach to the semantic segmentation of heritage data leverages the RF as a supervised learning algorithm (Paragraph 3.3.6.2). According to the five steps of the learning process illustrated by (Weinmann, 2016) and starting from an initial point cloud obtained by laser scanner or photogrammetry, the process is articulated in the following phases: neighborhood selection, feature extraction, feature selection, manual annotation and classification.

The case studies analyzed in this part for the geometry-based classification refer to the architectural typology of the cloister, where recurring typological elements (e.g., columns, vaults, arches and so on) are reunited in a courtyard space closed along the perimeter by open vaulted galleries. Three cloisters within the territorial scope of the Province of Pisa (Tuscany, Italy) are considered (Figure 5-12):

- The Grand cloister or main cloister of the Pisa Charterhouse.

The cloister was built starting from year 1375, and it undertook important renovations during the 17th century. The construction develops on a single level, with cross-vaulted spans that open towards the internal garden and a system of marble columns. The doors along the perimeter of the cloister mark the access to the ancient cells of the

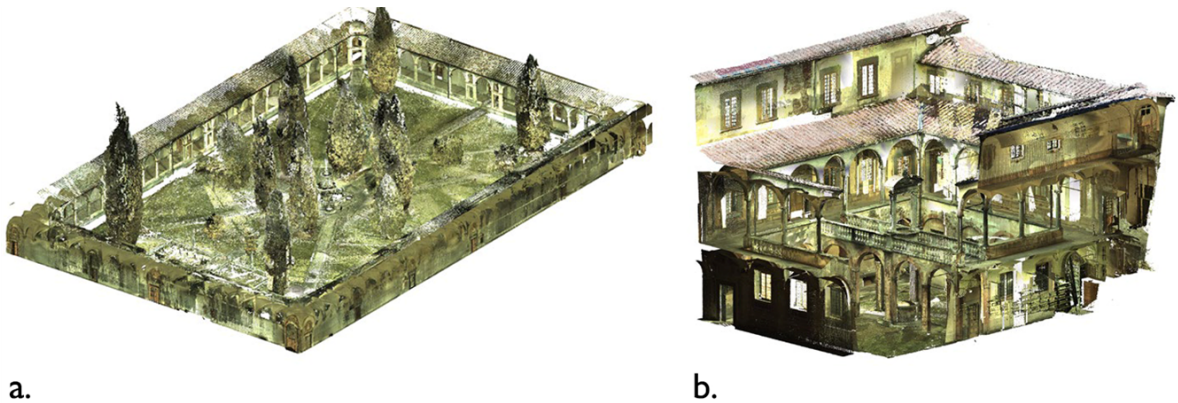


Figure 5-12.

Point clouds of the cases studied.

e.g.: The Grand cloister (*a*, 10 M points) and Grand-Ducal cloister (*b*, 6M points) of the Pisa Charterhouse and the cloister of the National Museum of San Matteo (*c*, 12 M points).

Carthusian Fathers. The cloister has a surface of about 70x45 m and the considered point cloud is the result of a Leica ScanStation C10 laser scanner survey (about 10 M points).

- The Grand-Ducal cloister of the Pisa Charterhouse.

The original structure dates back to the 14th century, but several transformations performed during the 17th century have granted it its current layout, marked by a series of vaulted galleries facing the internal courtyard and a central cistern. In plan, the internal perimeter of the cloister has a rectangular shape of almost 12x14 m. The survey is performed by laser scanner and later integrated with drone-based photogrammetry to restitute the roofing elements; the point cloud of this dataset is made up of 6 M points.

- The cloister of the medieval convent of San Matteo in Pisa.

The convent of San Matteo was founded in the 11th century and is today a National Museum. The cloister underwent major changes during the 16th century with the construction of the *portico*: a granite-columned loggia with Gothic windows and a cross-vaulted *ambulacrum* currently close the central space, which covers around 20x35 m. The survey was carried out by terrestrial photogrammetry and the resulting point cloud consists of about 12 M points.

The considered point clouds have different densities, but they are all set to a minimum space between points of 0,01 m.

For the three case studies, the semantic segmentation approach relying on the RF algorithm is applied on a case-by-case basis: a suitable set of features is extracted and selected from the input point cloud. Then, a sufficiently exhaustive portion of the model, describing all the classes that are present, is chosen as the training set, i.e., a dataset of examples used for the learning process. With a view to inserting this study in a wider context and to share the results more effectively, the identification of the classes of elements over these point clouds follows the above-mentioned state-of-the-art 3D ArCH benchmark (Matrone et al., 2020), that distinguishes ten classes: arch, column, molding, floor, door/ window, wall, stair, vault, roof, other (Croce et al., 2021).



c.

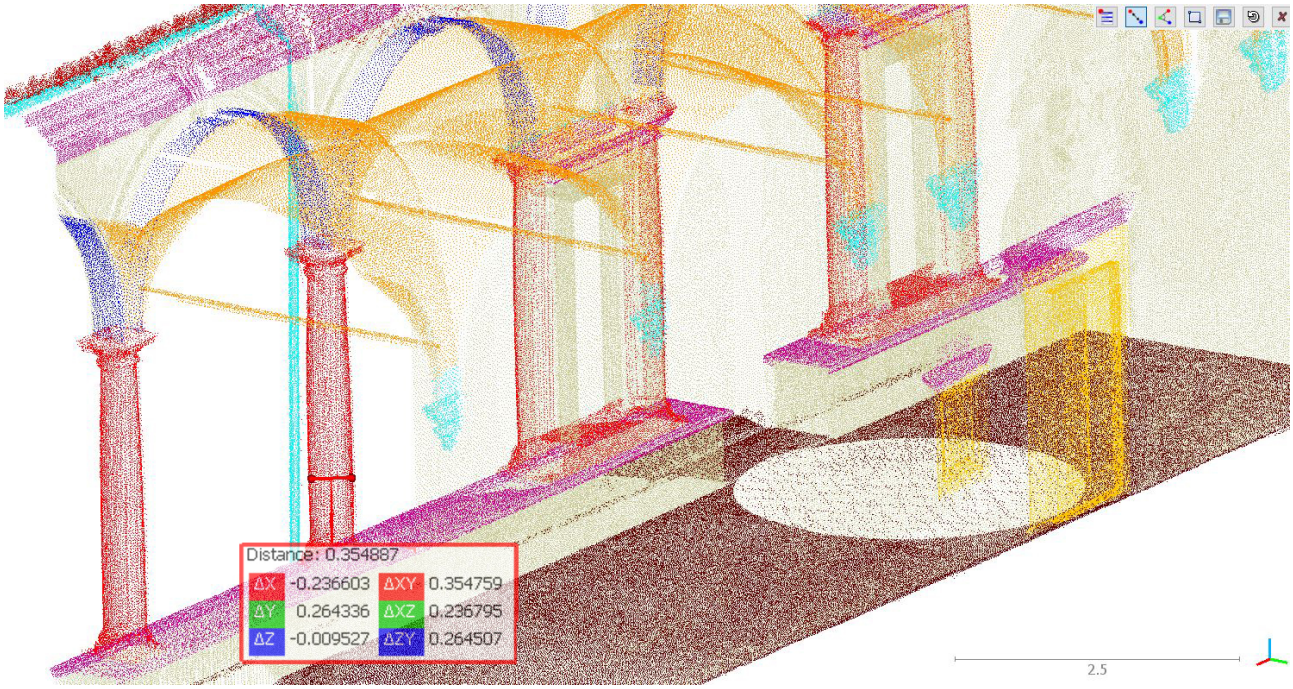


Figure 5-13.

The choice of the local neighborhood radius is influenced by evaluations on recurrent dimensions of the elements of the dataset.

e.g.: Grand cloister, Pisa Charterhouse.

5.3.1.1. Neighbourhood selection and feature extraction.

At an initial stage, geometric features, i.e., covariance features and changes of curvature (Paragraph 5.2.2.2) are extracted with varying local neighborhood radii (*multi-scale* approach), and they are later selected via a predictor importance estimate process.

Firstly, the choice of the local neighborhood in which features vary is guided by considerations on the recurring size of some elements that make up the dataset. The distance between two points of the point cloud is calculated to provide a first estimate, e.g., of the diameter of the columns (Figure 5-13), or of the thickness of certain architectural moldings, as well as other repetitive dimensions of the elements, so to extract geometric features considering selected ranges of the local neighborhood (e.g., $\varrho = 0.2\ m; \varrho = 0.4\ m, \varrho = 0.6\ m$).

Within the chosen ranges of ϱ, the covariance features are extracted from the covariance matrix (5-1) following Equations (5-4) to (5-12), while the Normal Change Rate is extracted as a curvature measure describing for each point the speed of the orientation change (Zhang, 2008).

Radiometric features, derived from the decomposition of the color space into single R, G, and B scalar fields, and a height feature (the Z coordinate), are then considered as additional characteristics for each dataset (Figure 5-14, Figure 5-15).

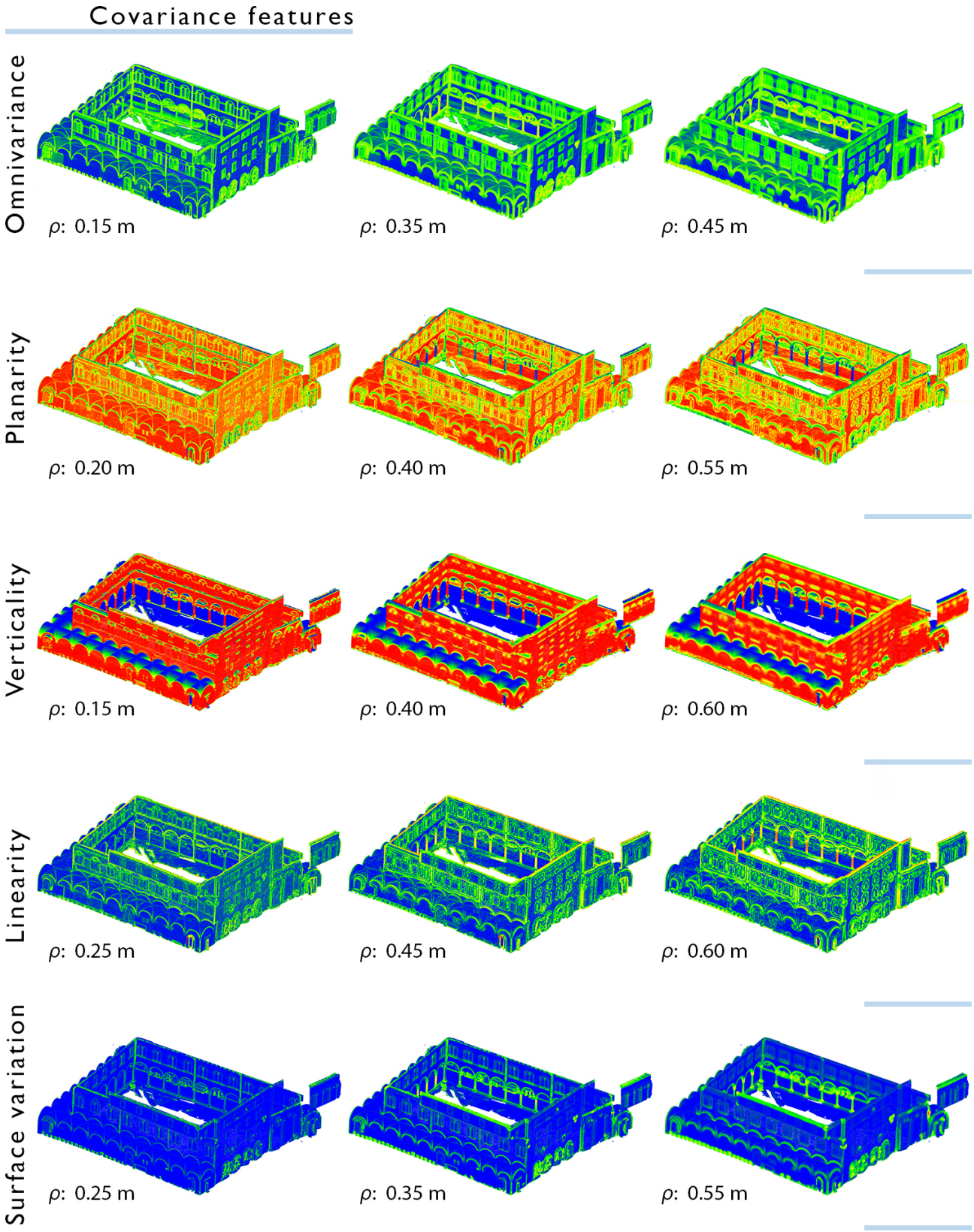
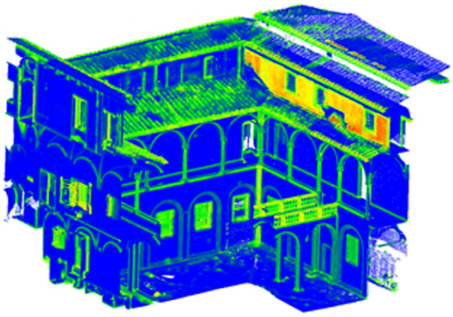


Figure 5-14.

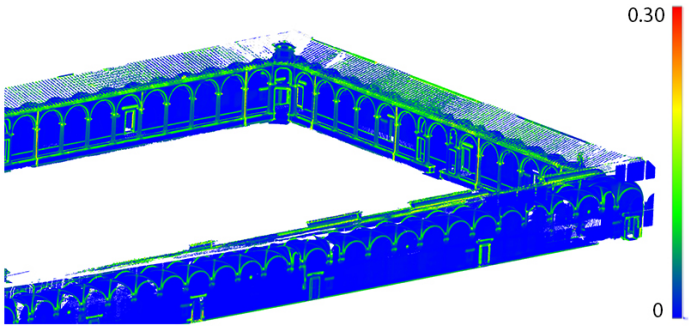
Examples of covariance features computed for different values of the local neighborhood radius.

e.g.: Cloister of the National Museum of San Matteo.

Changes of curvature

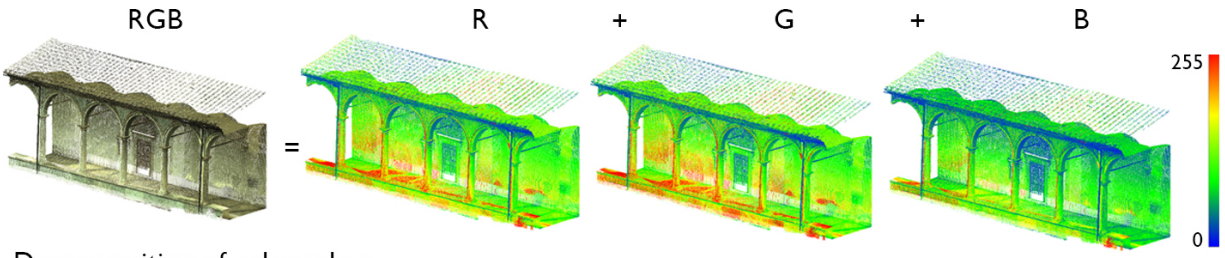


Normal change rate (ρ : 0.10 m)
e.g.: Grand-Ducal cloister,
Pisa Charterhouse



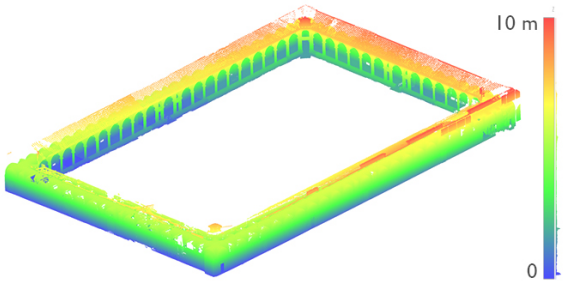
Normal change rate (ρ : 0.20 m)
e.g.: Main cloister,
Pisa Charterhouse

Radiometric features

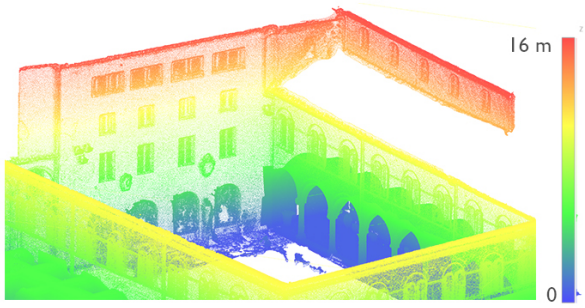


Decomposition of color values
into single scalar fields
e.g.: Main cloister, Pisa Charterhouse

Height feature



Z coordinate
e.g.: Main cloister, Pisa Charterhouse



e.g.: Cloister Museum of San Matteo, Pisa

Figure 5-15.

Examples of changes of curvature, radiometric features and height features computed for the different case studies.

5.3.1.2. Feature selection and manual annotation of the training set.

identification of classes over a portion of data allows training a RF classifier, so to evaluate first most important features and remove less relevant or redundant ones, thus improving both the classifier performance and the training time.

The so-called training set is specified each time for each case study, and it constitutes almost the 20% of the total number of points of the dataset (Figure 5-16). This subset of the overall point clouds is manually annotated: the point cloud is segmented according to the distinction into classes defined by (Matrone et al., 2020); then, a label is associated to each segment, by a class index varying from 0 to 10 and a related color. The training set is chosen in a representative portion of the dataset, where all classes to be annotated are visible.

Figure 5-17 shows the training sets identified for the three case studies, respectively. The distinction into classes is suggested graphically by a color legend.

In two cases, the number of classes is reduced from 10 to 9: for the Grand cloister, in fact, there are no objects belonging to the class '6 – Stair', while for the San Matteo point cloud the class '8 – Roof' is missing, as the photogrammetric survey is only performed at the ground level and does not allow the description of the roofing system.

For all the case studies, class '9 – Other' is inserted to group together all the elements not belonging to the previous classes (0 to 8), to include, e.g., downspouts or exhibition pieces.

After the manual annotation of the training set, a first classification test is run, via the RF algorithm, considering the multi-scale feature extraction (i.e., the whole set of features extracted at different local neighborhood radii). At the end of the learning process, an importance ranking can be displayed, showing the relevance score of each feature.

An example of a multi-scale predictor importance estimate is provided in Figure 5-18 for the case study of the main cloister of Pisa Charterhouse.

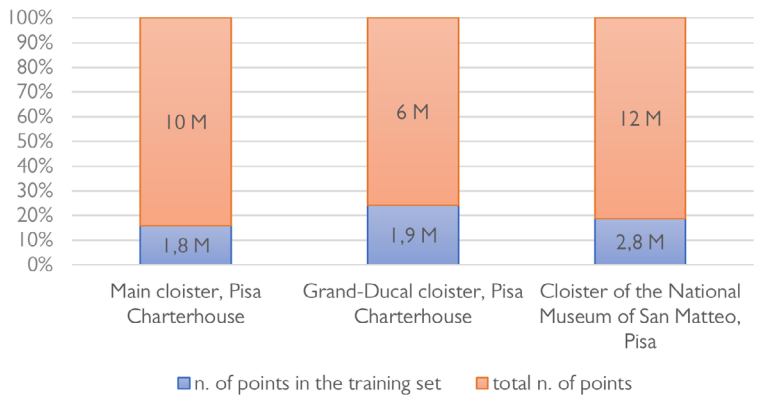


Figure 5-16.

Percentage of training set used for learning for the three case studies.

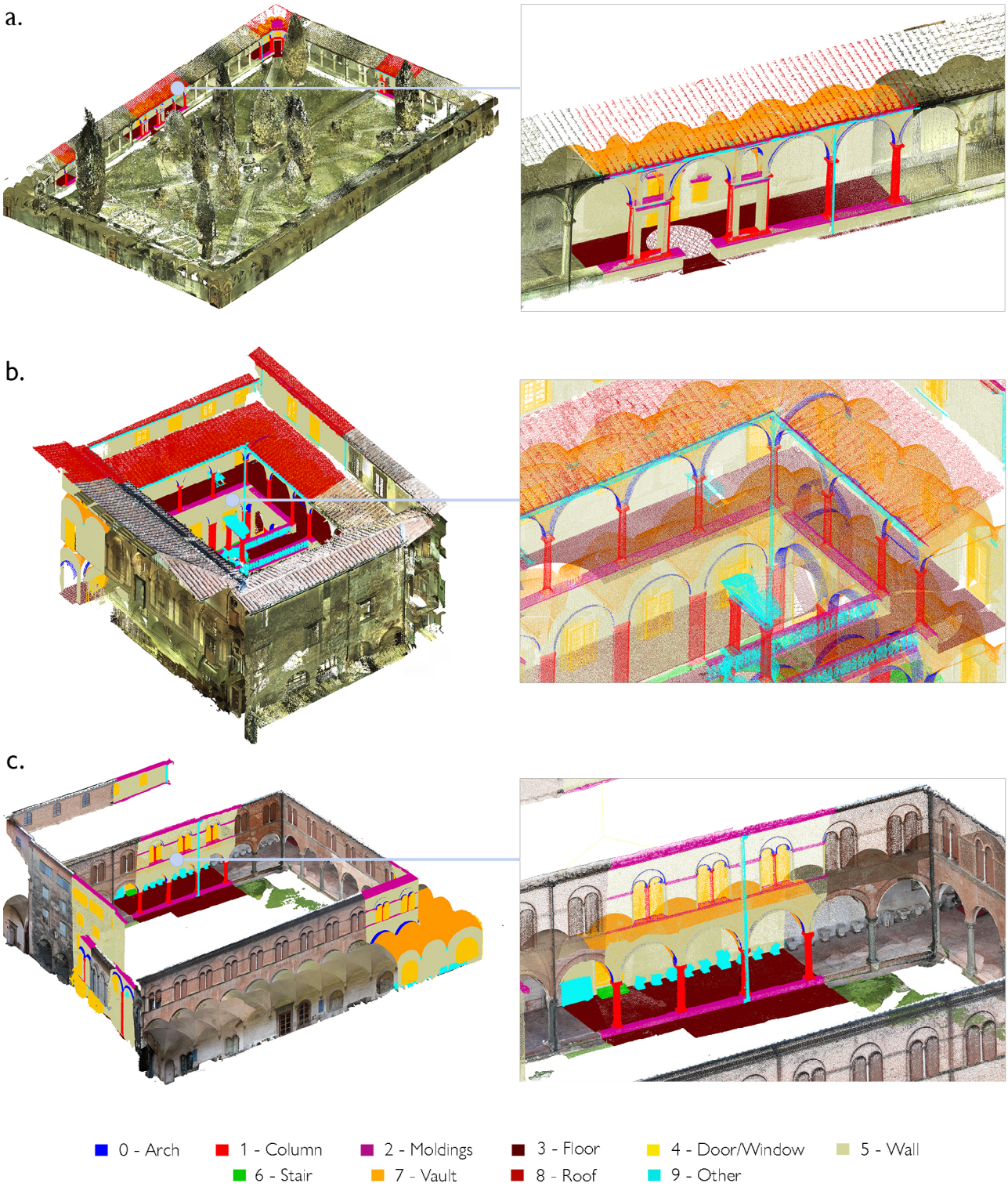


Figure 5-17.

Training set for the three case studies.

e.g.: the Grand cloister (*a*) and Grand-Ducal cloister (*b*) of the Pisa Charterhouse; the cloister of the National Museum of San Matteo (*c*).

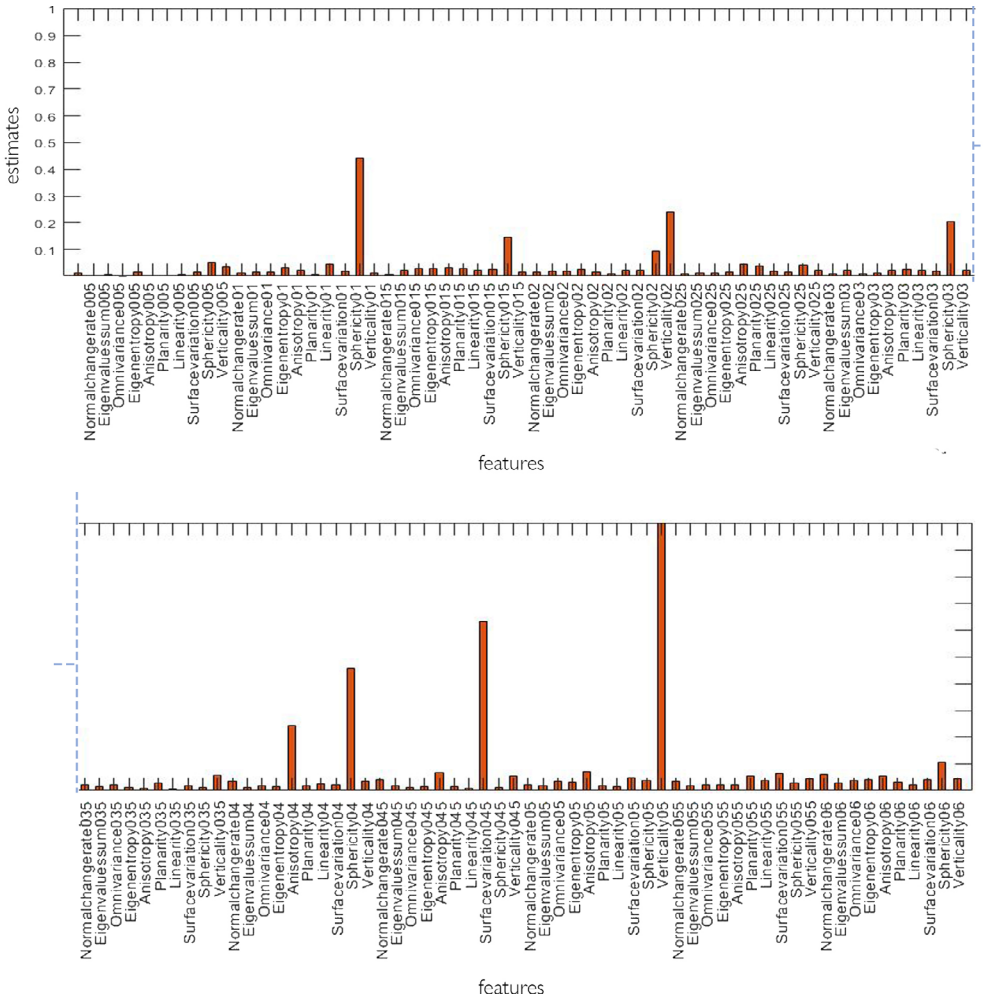


Figure 5-18.

Example of predictor importance estimate run for the main cloister of Pisa Charterhouse.

The number in parentheses at the end of each feature name stands for the related value of the local neighborhood radius in which the feature is computed.

Starting from the importance ranking extracted, redundant and less relevant features can be iteratively removed, and the RF trained with a reduced subset of features. This step allows insight to the data, showing which features are more relevant to the classification task, so to reduce data dimensionality to a set of about 10-15 features, and to allow the selection of a subset of predictors that suitably describe the identified classes. The feature ranking process (Figure 5-18) is run each time for the three different case studies.

At the end of the feature selection process, most important features extracted result in:

- Sphericity, surface variation, anisotropy, verticality, normal change rate for the main cloister dataset;
- Linearity, planarity, sphericity, omnivariance, surface variation,

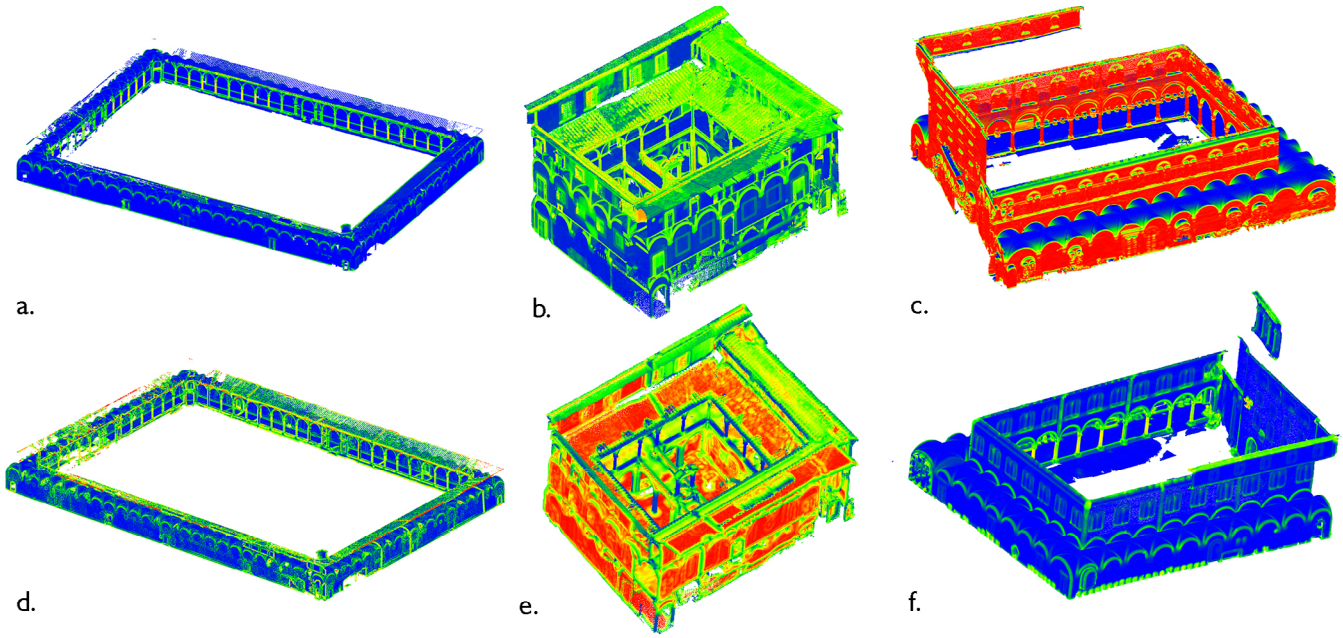


Figure 5-19.

Many geometric features selected: normal change rate (ρ= *0.4, a*) and sphericity (ρ= *0.4, d*) for the Grand cloister; omnivariance (ρ= *0.2, b*) and planarity (ρ= *0.6, e*) for the Grand-Ducal cloister; verticality (ρ= *0.2, c*) and surface variation (ρ= *0.4, f*) for the National Museum of S. Matteo.

anisotropy, verticality and normal change rate for the Grand-Ducal cloister dataset;

- Sphericity, omnivariance, surface variation, anisotropy, verticality, normal change rate for the San Matteo dataset.

By comparing the results of the three different datasets in terms of feature selection, and by visually considering the variation of features along the datasets (Figure 5-19, Figure 5-20), the following consideration can be drawn up:

- Surface variation, sphericity, anisotropy and verticality always appear among most relevant features. In particular, verticality is more apt to the distinction between elements of the dataset that are mostly horizontal (floors, ground) or vertical (columns, walls). Omnivariance and sphericity enable the recognition of architectural moldings, arches, vaults and columns; anisotropy and planarity further support the classification of columns and the distinction of windows and doors from the wall. Moreover, they are valuable for depicting finer elements with horizontal development, e.g., the windowsill and the under-roof moldings.
- Normal change rate, as curvature feature, is appropriate to depict columns and arches, and to identify elements belonging to 'class 9 – other', as drain spouts.
- Sum of eigenvalues and eigenentropy are discarded as least relevant features.

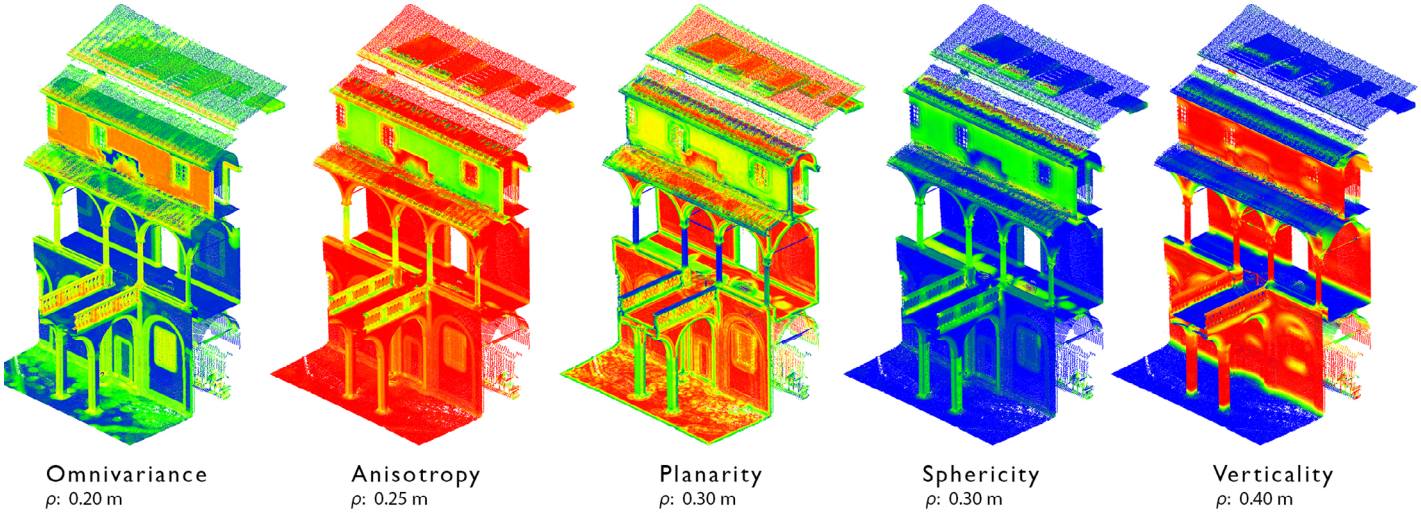


Figure 5-20.

Many relevant features identified on a portion of the Grand-Ducal cloister dataset, Pisa Charterhouse.

5.3.1.3. Classification results.

The selected features, associated to the manually annotated training set, allow to train the RF classifier, in order to extend the classification to the whole point cloud. This procedure is followed for each one of the cases under investigation and leads to the semantic segmentation results illustrated in Figure 5-21. Once classified, the point cloud datasets can be differentiated according to the distinction of the architectural elements that compose them: to give an example, Figure 5-22 provides an exploded view of the several classes of recurring architectural elements, recognized on a portion of the main cloister of the Pisa Charterhouse.

In order to evaluate the classifier performance, a validation set is considered, consisting of almost the 25% of annotated portion of the data, not previously used for training. On such validation set, the correspondence between the true -manually annotated- labels and the predicted ones is tested by referring to the values of the confusion matrix and to the performance measures: ***Precision*** (5-13), ***Recall*** (5-14), ***Overall accuracy*** (5-15) and ***F-measure*** (5-16) (Paragraph 5.2.5).

The average values of the performance measures obtained for the three case studies are summarized in Table 5-2.

Table 5-2.

Average performance measures obtained for the classification of the three datasets.

	Grand cloister, Pisa Charterhouse	*Grand-Ducal cloister, Pisa Charterhouse*	*Cloister Museum of San Matteo, Pisa*
n. of classes	9	10	9
Avg. Recall	93,49%	83,03%	84,37%
Avg. Precision	95,56 %	82,07 %	89,71 %
Avg. Accuracy	99,30 %	98,04 %	98,73 %
Avg. F1-score	94,44 %	81,53 %	85,98 %

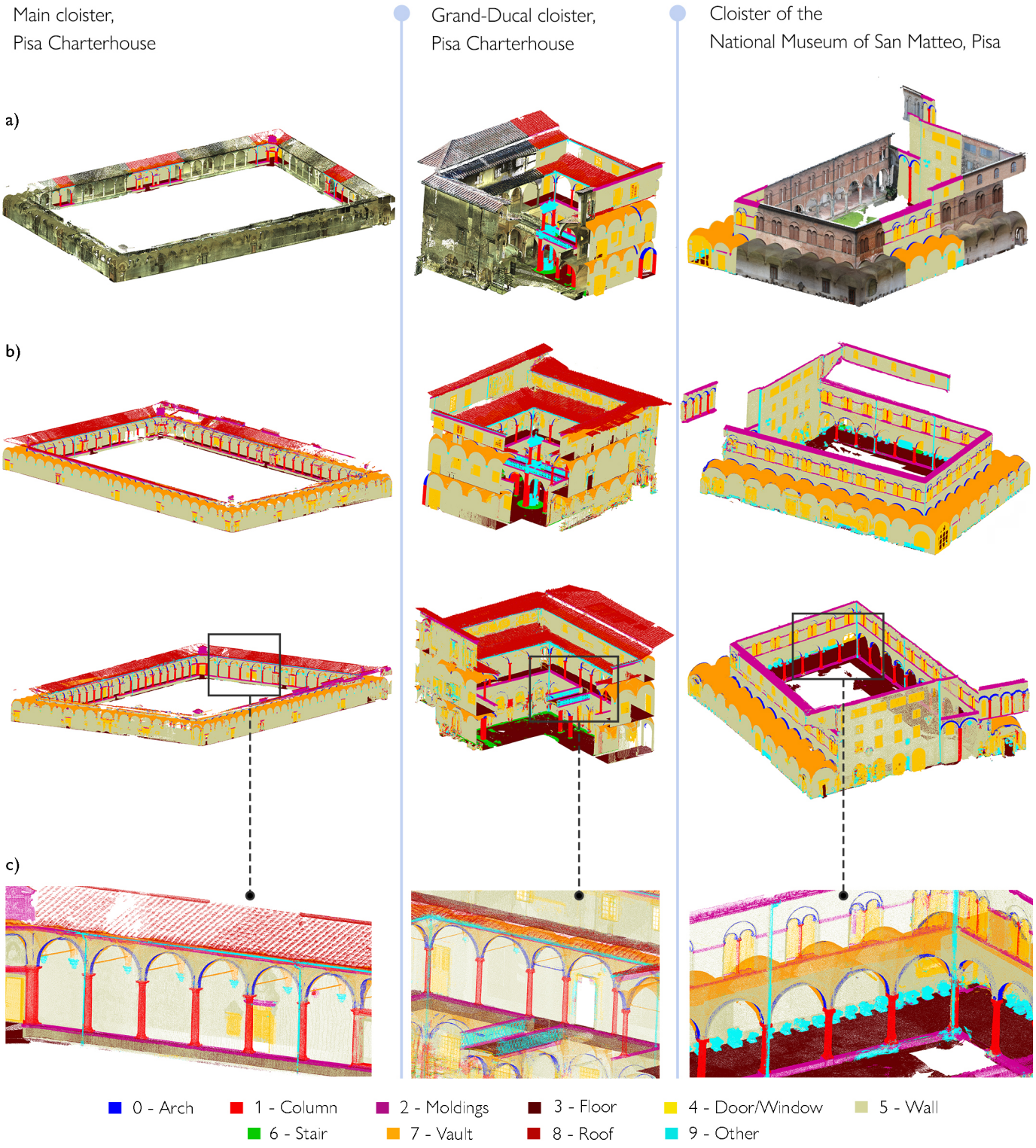


Figure 5-21.

Training set (*a*) and segmentation results (*b, c*) for the three case studies.

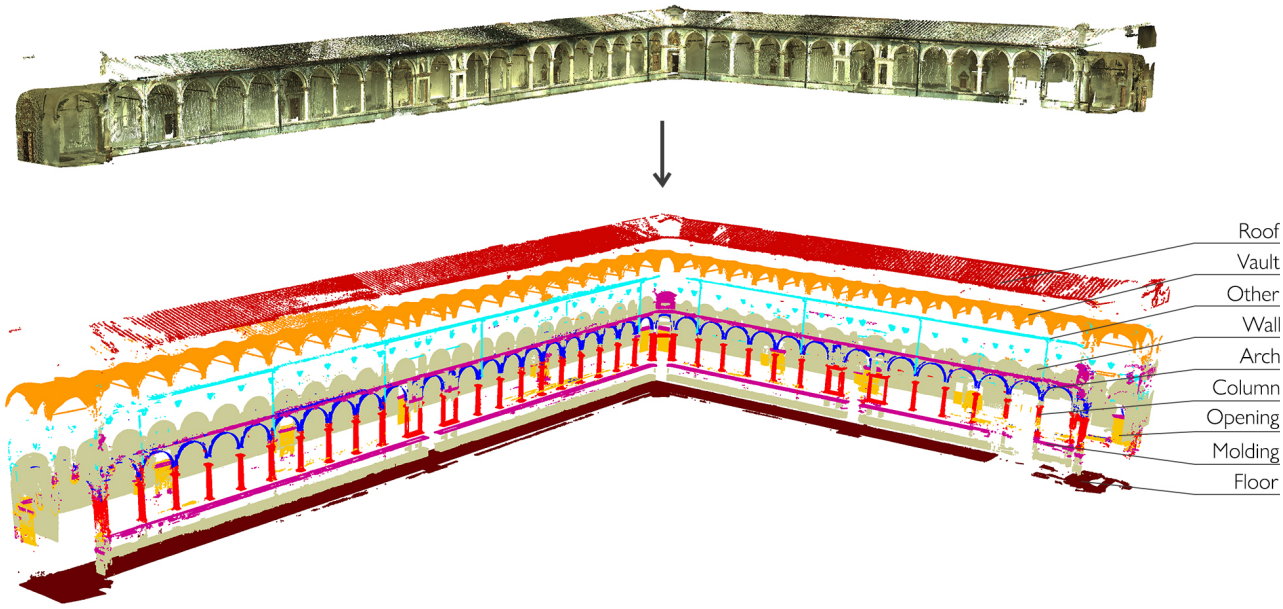


Figure 5-22.

Semantic segmentation: exploded view of the classified components.

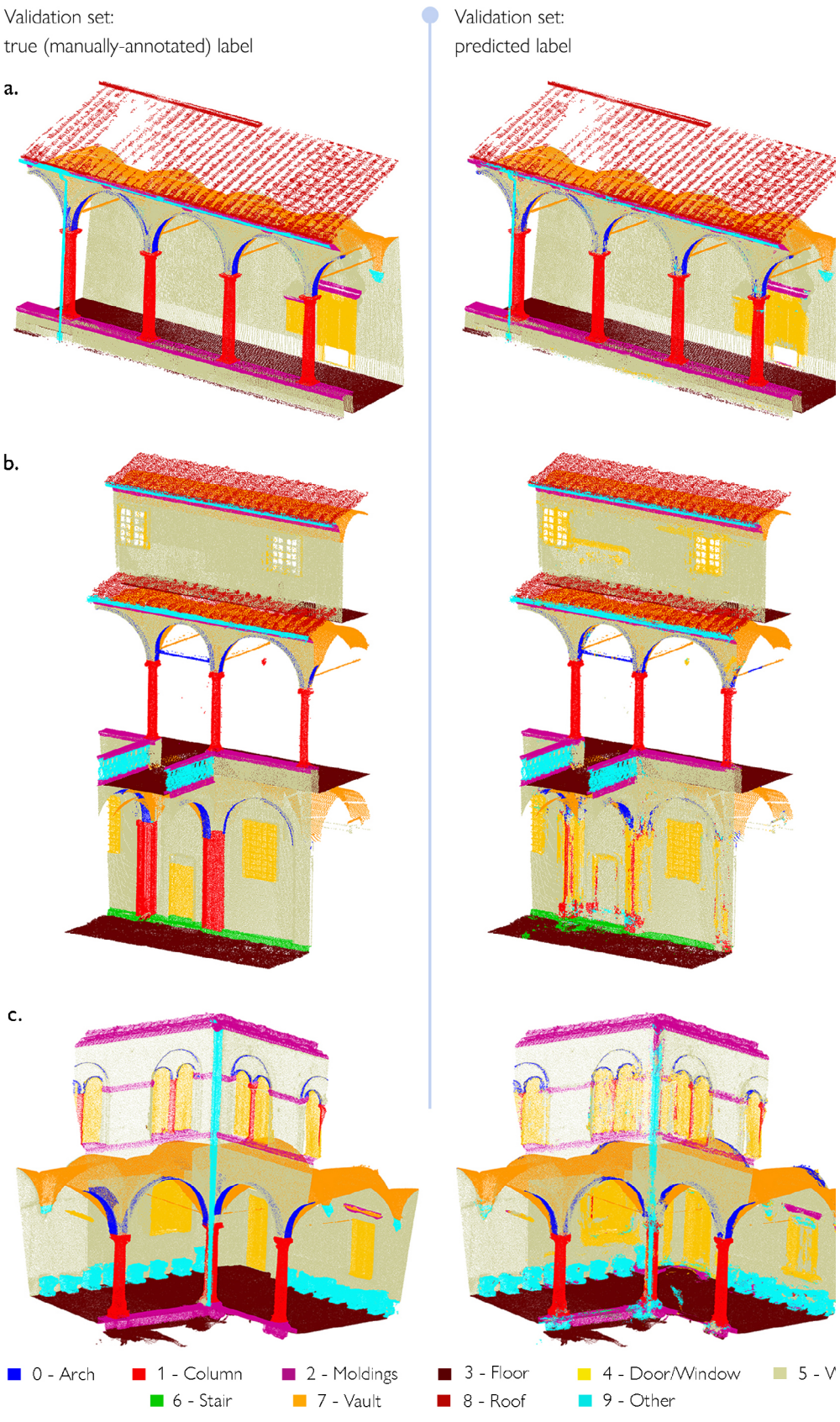
e.g.: Grand cloister, Pisa Charterhouse

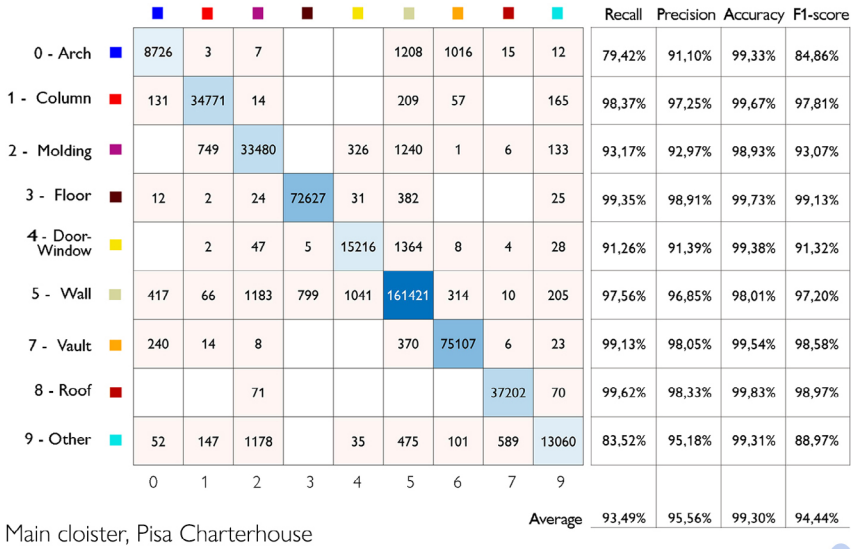
Figure 5-23 displays the visual comparison between true (manually annotated) and predicted labels (b), while the confusion matrix and the performance values, broken down class by class, are provided in Figure 5-24.

Figure 5-23.

Validation sets: comparison between true and predicted labels.

e.g.: the Grand cloister (*a*) and Grand-Ducal cloister (*b*) of the Pisa Charterhouse; the cloister of the National Museum of San Matteo (*c*).





	0	1	2	3	4	5	6	7	9	Recall	Precision	Accuracy	F1-score
0 - Arch	8726	3	7			1208	1016	15	12	79,42%	91,10%	99,33%	84,86%
1 - Column	131	34771	14			209	57		165	98,37%	97,25%	99,67%	97,81%
2 - Molding		749	33480		326	1240	1	6	133	93,17%	92,97%	98,93%	93,07%
3 - Floor	12	2	24	72627	31	382			25	99,35%	98,91%	99,73%	99,13%
4 - Door-Window		2	47	5	15216	1364	8	4	28	91,26%	91,39%	99,38%	91,32%
5 - Wall	417	66	1183	799	1041	161421	314	10	205	97,56%	96,85%	98,01%	97,20%
7 - Vault	240	14	8			370	75107	6	23	99,13%	98,05%	99,54%	98,58%
8 - Roof			71					37202	70	99,62%	98,33%	99,83%	98,97%
9 - Other	52	147	1178		35	475	101	589	13060	83,52%	95,18%	99,31%	88,97%
Average										93,49%	95,56%	99,30%	94,44%

Main cloister, Pisa Charterhouse

	0	1	2	3	4	5	6	7	8	9	Recall	Precision	Accuracy	F1-score
0 - Arch	5165	57	4	4	193	1092	5	622	1	46	71,85%	77,11%	99,08%	74,39%
Column	26	10915	185		5323	2362	35	6		300	56,99%	90,48%	97,57%	69,93%
Molding	90	13	17874	106	7	653	1	75		293	93,52%	92,37%	99,30%	92,94%
3 - Floor	16	22	30	77893	40	338	2267	170	100	54	95,75%	99,43%	98,95%	97,56%
4 - Door-Window	9	277			10639	3189		3		143	74,61%	41,16%	95,13%	53,05%
5 - Wall	334	471	521	185	9375	97392	77	287	25	606	89,13%	89,70%	94,03%	89,41%
6 - Stair		259		102	1	954	2434	5		31	64,29%	47,82%	98,96%	54,84%
7 - Vault	1022	20	10	44	228	1403	12	88625	126	137	95,98%	98,68%	98,73%	97,31%
8 - Roof			211	1	1	143		11	18239	858	93,71%	97,39%	99,56%	95,51%
- Other	3	1	514	18	22	126	1	13	237	15957	94,46%	86,59%	99,12%	90,35%
Average											83,03%	82,07%	98,04%	81,53%

Grand-Ducal cloister, Pisa Charterhouse

	0	1	2	3	4	5	6	7	9	Recall	Precision	Accuracy	F1-score
0 - Arch	18532	31	247		1056	1389		652	10	84,56%	85,78%	99,12%	85,17%
1 - Column	68	40240	1010		643	524		11	145	94,37%	98,40%	99,59%	96,34%
2 - Molding	19	190	59693	1525	376	1358	37	28	393	93,83%	88,52%	98,42%	91,10%
3 - Floor			2445	169134	50	138	9		649	98,09%	97,45%	98,95%	97,77%
4 - Door-Window	111	158	472		19541	3269			129	82,52%	69,28%	98,26%	75,33%
5 - Wall	523	199	488	424	2909	185032	4	338	1340	96,75%	93,89%	97,52%	95,29%
6 - Stairs	4		340	388	276	449	2363	1	1570	43,83%	89,88%	99,55%	58,93%
7 - Vault	2346	13	1		14	2196		176850	11	97,48%	99,42%	99,24%	98,44%
9 - Other		64	2741	2086	3339	2724	216		23657	67,93%	84,78%	97,91%	75,42%
Average										84,37%	89,71%	98,73%	85,98%

Cloister of the National Museum of San Matteo, Pisa

Figure 5-24.

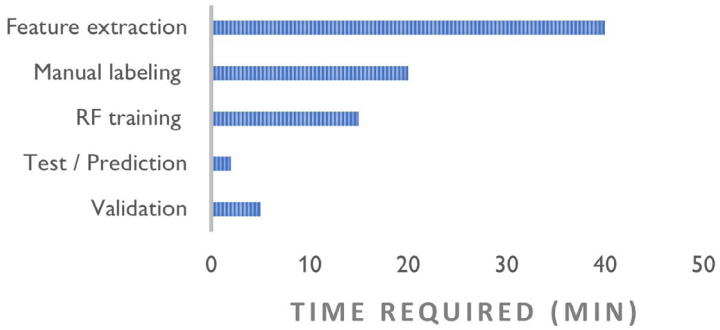
Confusion matrix and performance scores for the three case studies.

5.3.1.4. *Assessment of geometry-based classification results.*

From the evaluation of the performance scores of the three datasets considered, the geometry-based data classification returns an average accuracy of 98,73 % and a F-score of 87,13%. The models accurately predict the semantic segmentation results, and the data obtained are deemed satisfactory, particularly in view of the use of the semantically annotated point cloud for the subsequent reconstruction and modeling phase in BIM. The ML-based semi-automatic segmentation process significantly reduces the annotation time: Table 5-3 summarizes elaboration times for the different phases of the workflow.

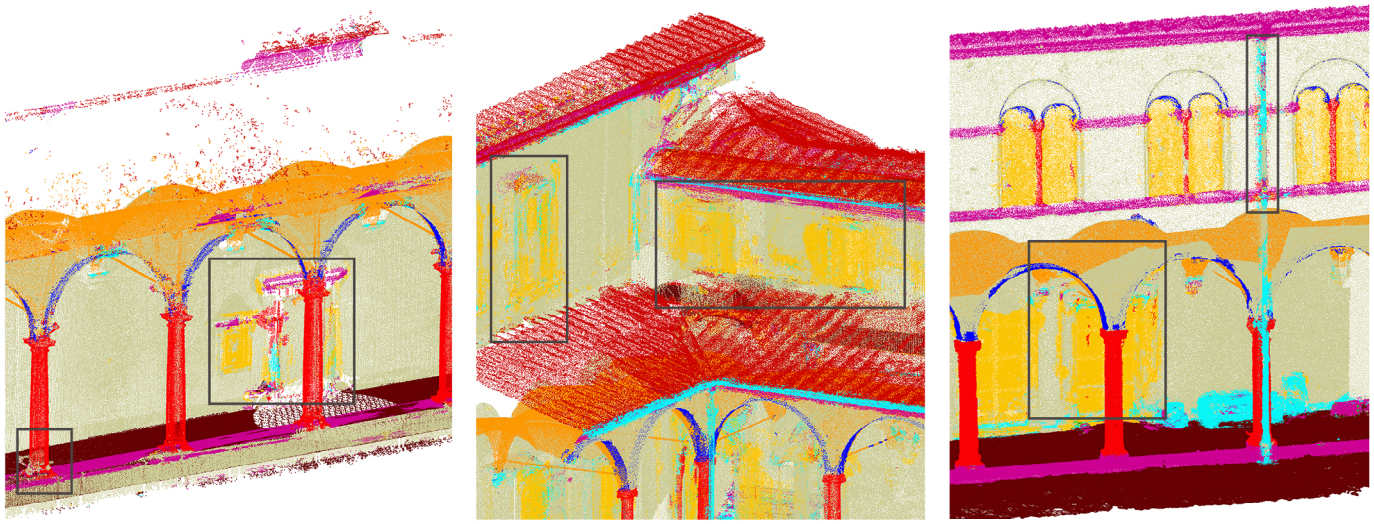
Table 5-3.

Average time required for the different phases of the classification process.



However, some considerations need to be raised on the validation of this methodology:

- By observing the confusion matrices (Figure 5-24) and by visually verifying the data with the segmentation results, misclassification errors appear at the boundary regions, i.e., over those areas that mark the limit between one class and another. In detail, since the features are computed in a given local neighborhood, the feature extraction may be misleading for those 3D points that are located in border regions between one class and another (Figure 5-25). These errors increase at increasing radius of the spherical neighborhood. This situation has been mitigated, on the one hand, by adding radiometric features discriminating based on color information, and on the other hand, by choosing low values of ρ, below 0.6 m.
- Regions with similar development (e.g., planar or cylindrical) can be mis-classified as belonging to the same class. Considering the off-diagonal elements of the confusion matrix, it is noted that 'class 5 – wall' and 'class 4 – door and window', are often mistaken amongst themselves, as they are characterized by mostly planar behavior. For regions of this type, in fact, the geometric features appear with similar values, and this may lead to incorrect classifications.



a. b. c.

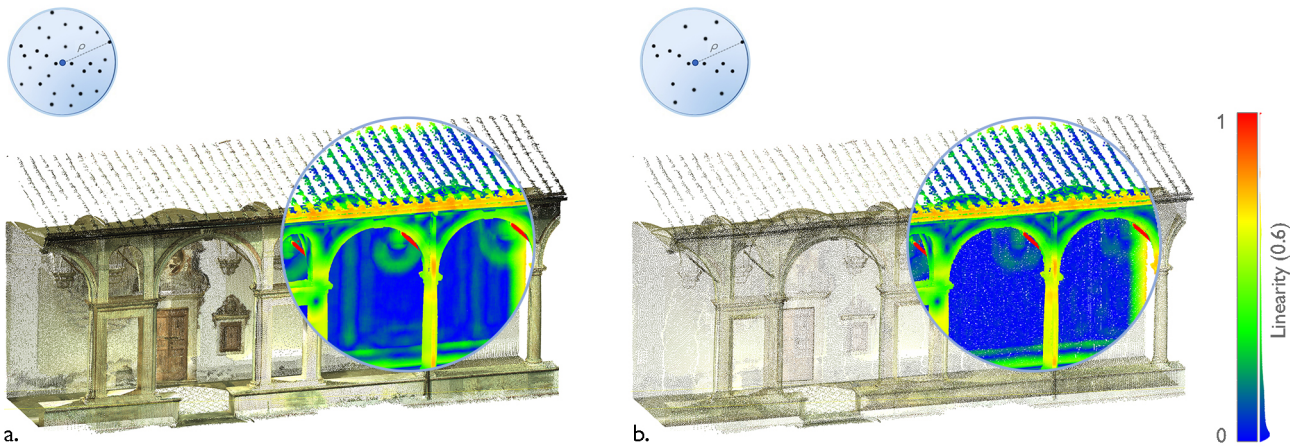
Figure 5-25 *(above)*.

Misclassified elements in boundary regions.

e.g.: the Grand cloister (*a*) and Grand-Ducal cloister (*b*) of the Pisa Charterhouse; the cloister of the National Museum of San Matteo (*c*).

- Since the covariance and curvature features are computed in a given local neighborhood, the density of the point cloud influences the classification results. Indeed, considering two point clouds for the same case study but with different point densities, the feature selection may sort two distinct results. As for the example of Figure 5-26, the values of a same geometric feature, extracted in a same local neighborhood, vary if the two point clouds have different point densities. In order to align feature selection for different surveys of the same dataset, it could thus envisaged to restore the point clouds to the same density, by means of a subsampling operation.

In addition to these observations, it should be pointed out that the number of semantic classes chosen can strongly influence the quality of the classification results, since, for example, as the number of classes



a. b.

Figure 5-26 *(below)*.

Example of a same feature (linearity, $\varrho = 0.6\ m$) computed on two point clouds with higher (*a*) and lower (*b*) densities.

e.g.: Main cloister, Pisa Charterhouse

increases, a greater similarity between them can be found; likewise, the result can be influenced by the amount of training data selected for the classification task.

Finally, it should be noted that the case studies presented are all related to the medieval cloister typology, which is characterized by recurring typological elements. Future work may thus consider:

- Extending the methodology to other datasets, related to different architectural types and/or built in different periods. In any case, the classification should be driven by the recognition of typological architectural elements;
- Using the same training set to generalize the classification of several datasets related to a same architectural typology. In this case, the point clouds should have comparable densities, and should identify objects with very similar characteristics, so enabling the identification of a unique set of classes.

In any case, the results obtained through geometry-based approaches allow distinguishing different datasets based on the recognition of recurring architectural elements.

The separation of semantic parts is the fundamental prelude for the automation of the scan-to-BIM process and for a better correlation between point cloud and parametric model.

5.3.2. Texture-based classification results

Texture-based classification approaches rely on colorimetric information, and they are demanded in cases where geometric data alone are not sufficient to satisfy the tasks of description and semantic representation, e.g., if maps of degradation, surface alterations and materials are to be derived from the digital data.

Since being based on the elaboration of textures, the approaches apply either to the classification of images, ortho-photos or UV maps, the latter

Figure 5-27.

Texture-based approaches can be applied to images (*a*), ortho-photos (*b*) or UV maps (*c*).

e.g.: Cappella Sant'Agata in Pisa



a. b. c.

extracted from high resolution photogrammetric models (Figure 5-27). Each case study presented in this section relates to the classification of one such digital data type, respectively. In detail:

- For the classification of ortho-images, the case study refers to a medieval Chapel in Pisa, the *Cappella di Sant'Agata*.
- For the classification of photos, reference is made to a set of images of the vaults of the Notre-Dame Cathedral in Paris.
- For the UV map classification, the photogrammetric models of the altars of the *Chiesa del Carmine,* a Pisan church, are studied.

As for the case of geometry-based approaches, the RF classification algorithm is applied on a case-by-case basis: a suitable set of features is extracted from the input point cloud, then the classes that are to be identified are manually annotated over reduced portions of the image to train the RF classifier. Once the semantic segmentation result is obtained, the classifier performance can be evaluated by comparing the predicted and real classes on a manually annotated portion of the model, i.e., the validation set. In this case, the classified elements are pixels (as opposed to the 3D points of the geometry-based case) and the classification results are graphically displayed on the image.

5.3.2.1. Classification of ortho-photos.

The classification results on ortho-photos are presented with reference to the case study of the chapel *Cappella di Sant'Agata*.

The chapel, built on an octagonal plan, is located behind the apse of the church of *San Paolo a Ripa d'Arno*, in Pisa. The construction was ordered by the *Vallombrosian* monks of the church in the second half of the 11th century; initially connected to the church by medieval buildings, the chapel is now surrounded by public greenery. The building is closed by perimetral masonry walls articulated by pillars, three-mullioned windows, under-roof arches and it concludes upwards with a spire.

In the 1960s, the restoration works of the interiors, entailing the removal of the outermost layer of plaster that covered the original walls, led to the discovery of frescoes dating back to the end of the 13th century – beginning of the 14th century and attributed to an anonymous Pisan master. In the years intervened between the restoration works and today, a progressive loss of the discovered frescoes has occurred, even caused by the exposure of the chapel to the atmospheric agents.

For this case study, the survey of the chapel is exploited to describe the current condition of the building's internal space, revealing in particular the parts of the brickwork that still show remains of wall painting and traceable back to the original decorations.



Figure 5-28 *(above)*.

Section of the chapel (*a*) and manually-annotated training set with related legend (*b*).

e.g.: Cappella Sant'Agata in Pisa

The classification is conducted considering the survey carried out by the ASTRO laboratory of the University of Pisa in year 2016, involving the use of total station, laser scanner and terrestrial photogrammetry alongside traditional survey instruments for the restitution of the interior and exterior walls of the Chapel (Tenchini, 2016). In detail, the annotation is performed over a section of the chapel, derived from the elaboration of the ortho-mosaic of the photogrammetric model (Figure 5-28a). In the training phase, four annotation classes are manually identified on the orthophoto, associated to relative colors and labels (Figure 5-28b): brick masonry still presenting traces of wall painting (class 0), chromatic alteration of brick masonry (1), plaster (2) and marble (3). The white background of the image is excluded from the classification, and the training time takes almost 10 minutes.

The manually annotated portion of data is then associated with the

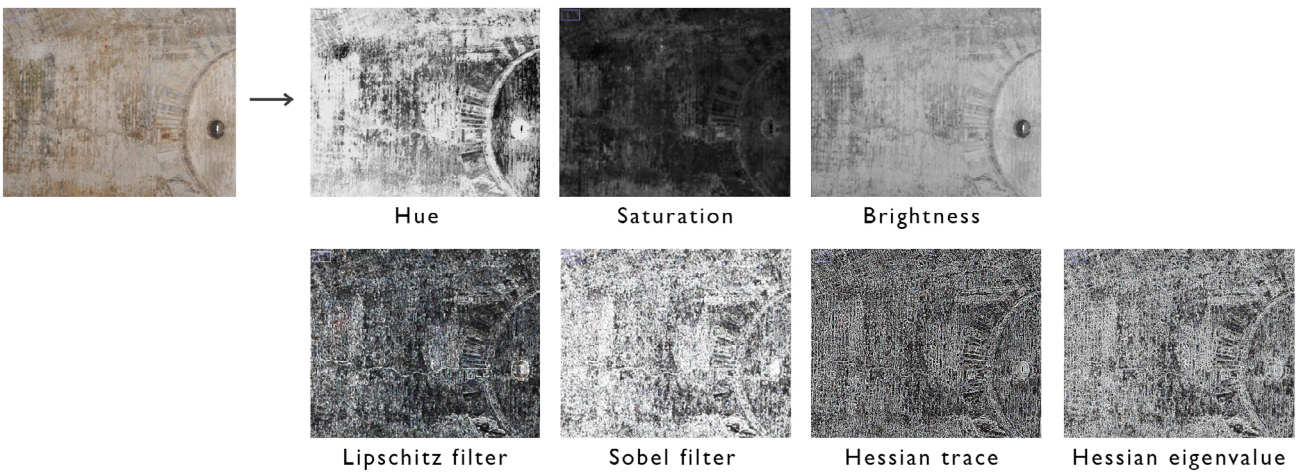


Figure 5-29 *(below)*.

Image adjustments, edge-detection and noise-reduction filters extracted from the input image.

e.g.: Cappella Sant'Agata in Pisa

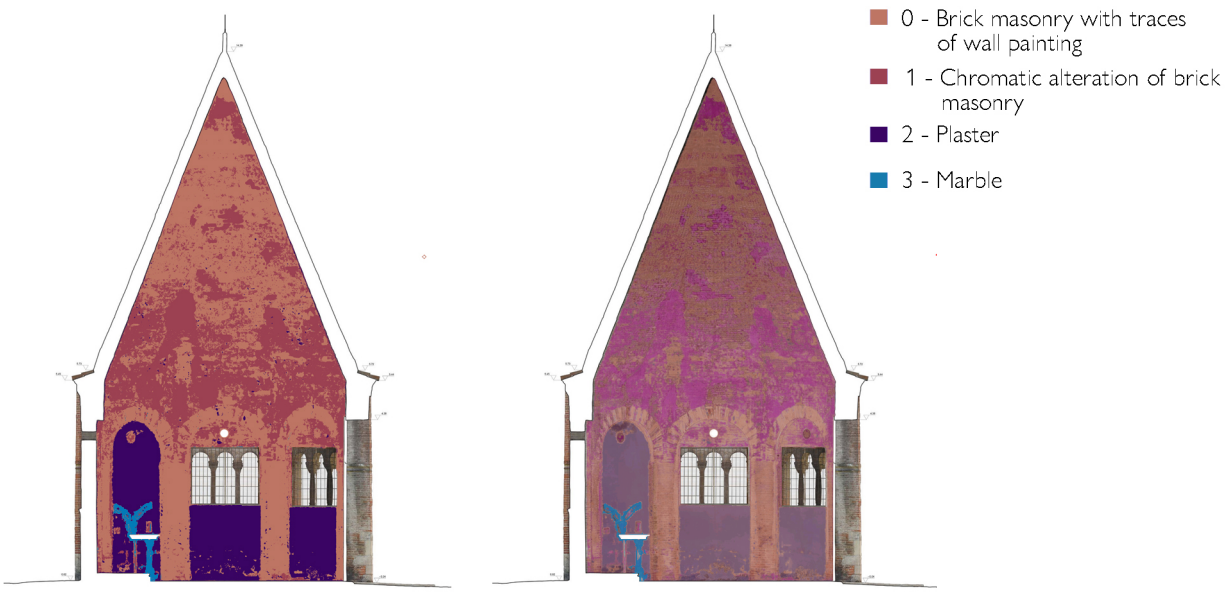


Figure 5-30.

Semantic segmentation results for a section of the Cappella di Sant'Agata in Pisa, in opacity (*a*) and in transparency (*b*).

e.g.: Cappella Sant'Agata in Pisa

extraction of pertinent features (Figure 5-29, Figure 5-30): image adjustments, edge-detection and noise-reduction filters as well as texture filters.

The results of the automatic annotation by RF training are provided in Figure 5-31, also showing the results in transparency over the initial image. In particular, the distinction between parts of brick altered by the presence of plaster (chromatic alteration) and parts of brick where traces of the original frescoes are still recognizable is highlighted graphically by color maps.

In order to validate the classification results, the automatic annotation is compared to the manual annotation on a portion of the data (validation set of Figure 5-32 a-c). Figure 5-32d further displays the graphical comparison between true and predicted labels, resulting from the

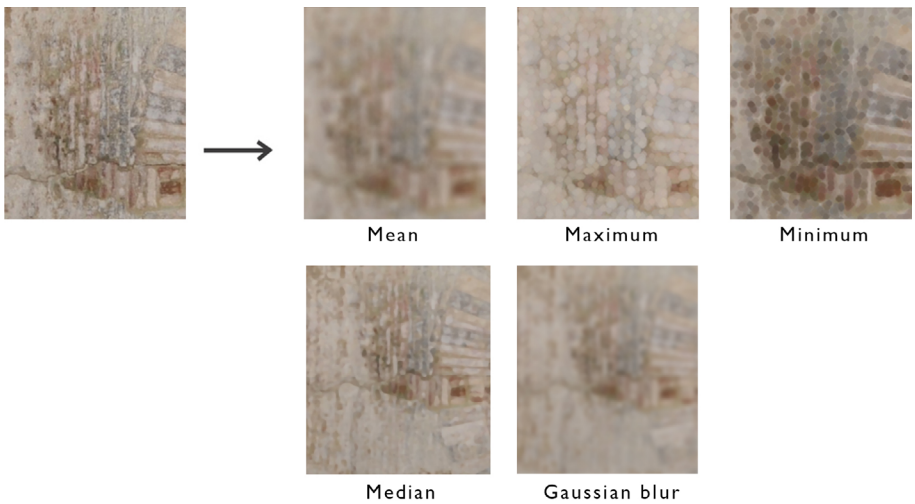
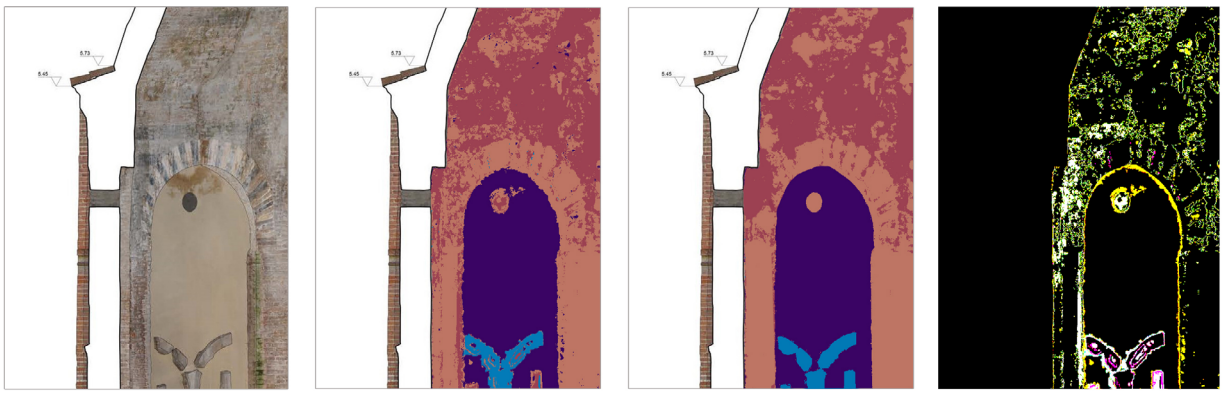


Figure 5-31.

Texture filters extracted from the input image.

e.g.: Cappella Sant'Agata in Pisa



a. b. c. d.

Figure 5-32.

Validation set (*a*). Comparison between predicted (*b*) and true (manually-annotated, *c*) labels, and false colors map highlighting the differences between one class and another (*d*).

e.g.: Cappella Sant'Agata in Pisa

juxtaposition of the two images: the false colors denote the difference between true and predicted classes.

Again, from the comparison between true and predicted data, the confusion matrix and related performance measures can be extracted. The values of the confusion matrix are provided in terms of number of correctly/incorrectly classified pixels (Figure 5-33).

The results provide optimal values of accuracy and precision, on the order of 90%. To these results it should be added that, compared to the manual classification performed by the operator, the supervised automatic classification allows to describe more details and differences in the material composition of the surfaces.

Moreover, this classification requires a remarkably low amount of time, especially if compared to the manual annotation time of the whole ortho-photo, which would require, according to the estimates, from 6 to 9 hours. The control, by the operator, of the annotated data, and the eventual refinement, when necessary, of some parts not correctly classified, could be further considered in order to optimize results and classification times, always in the perspective of a supervised approach.

Successive segmentations are considered over the image dataset: at first, joints and ashlars classes are identified on the vaults' intrados (the

Figure 5-33.

Confusion matrix and related performance measures extracted based on the validation set.

e.g.: Cappella Sant'Agata in Pisa

	0	1	2	3	Recall	Precision	Accuracy	F1-score
0 - Brick masonry with traces of wall painting	76161	7413	559	25	90,50%	89,04%	93,01%	89,76%
1 - Chromatic alteration of brick masonry	6381	73752	16	3	92,02%	90,17%	94,19%	91,08%
2 - Plaster	2652	199	74555	1020	95,06%	99,03%	98,15%	97,01%
3 - Marble	346	428	157	4905	84,05%	82,40%	99,20%	83,21%
Average					90,41%	90,16%	96,14%	90,27%

5.3.2.2. Classification of photos.

For the classification of 2D images, the example of the vaults of the Notre Dame-Cathedral in Paris is provided. In this case, the classification test aims at semantically annotating, in a semi-automatic way, the intrados of the vaults of the choir and the nave, distinguishing the ashlars by type and color.

For the study of the remaining part of the vaults of Notre Dame Cathedral following the fire of April 2019, the semantic segmentation operation, in this case, is even carried out in view of the combination between manual and semi-automatic annotations within the Aïoli platform and in view of the consequent transfer of all annotations (manual or semi-automatic ones) from 2D to 3D (e.g. in Figure 5-34).

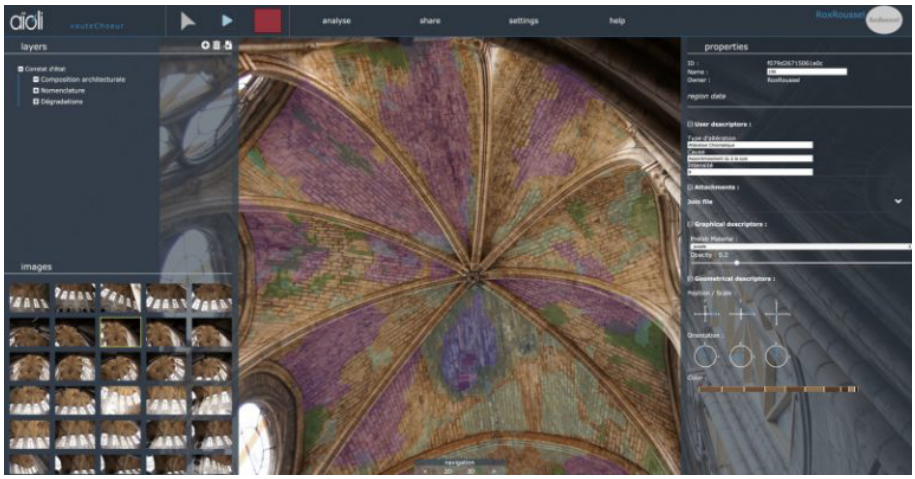


Figure 5-34.

2D/3D manual annotation of the alterations detected on the choir vaults on the Aïoli platform

(*source:* UMR MAP CNRS/MC)

e.g.: Notre-Dame Cathedral in Paris

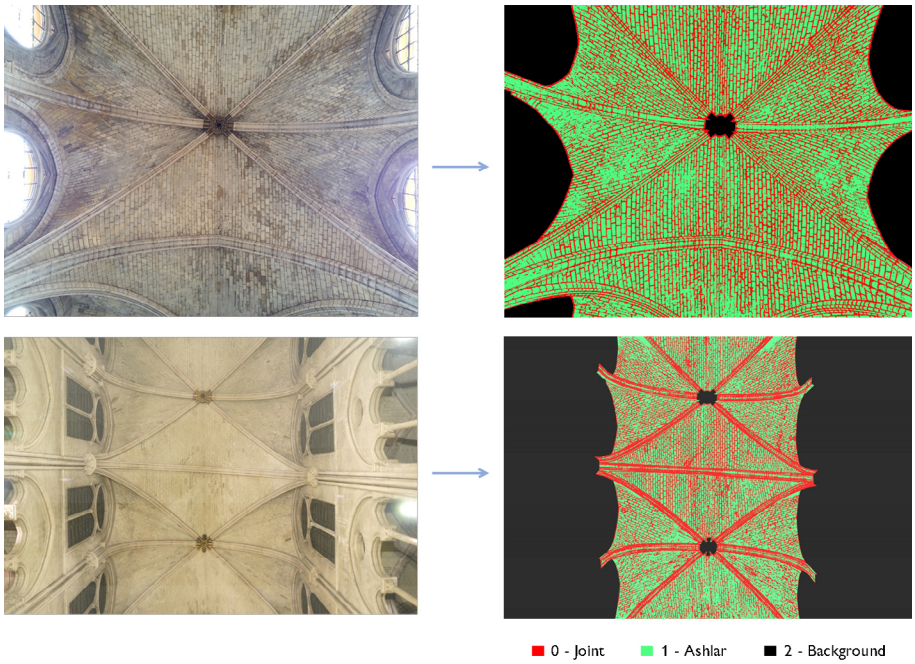


Figure 5-35.

Segmentation results for the image sets of the intrados of the vaults (first segmentation phase).

e.g.: Notre-Dame Cathedral in Paris

rest of the images-background- is excluded from the annotation). the extracted features and the training data are illustrated in the diagrams of Figure 5-36 and Figure 5-37, respectively.

The results for the first segmentation phase are shown in Figure 5-38.

Figure 5-36.

Image adjustments (1st row), edge detectors and noise reduction filters (2nd row) extracted from an image of the vaults.

e.g.: Notre-Dame Cathedral in Paris

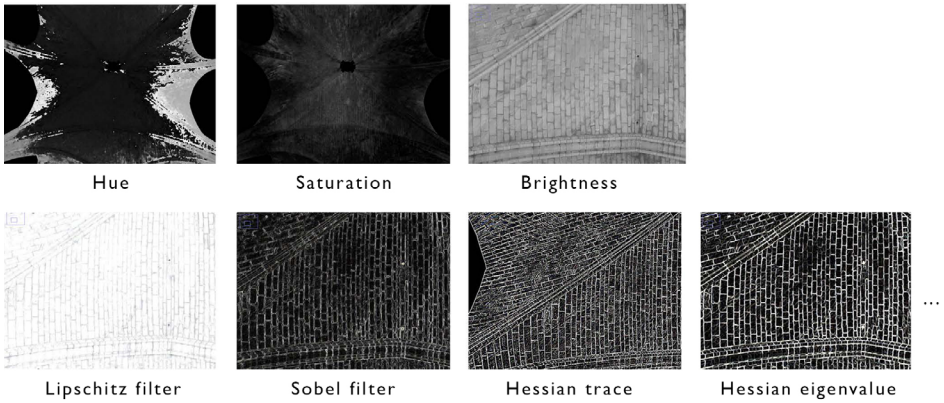


Figure 5-37.

Texture filters extracted from an image of the vaults.

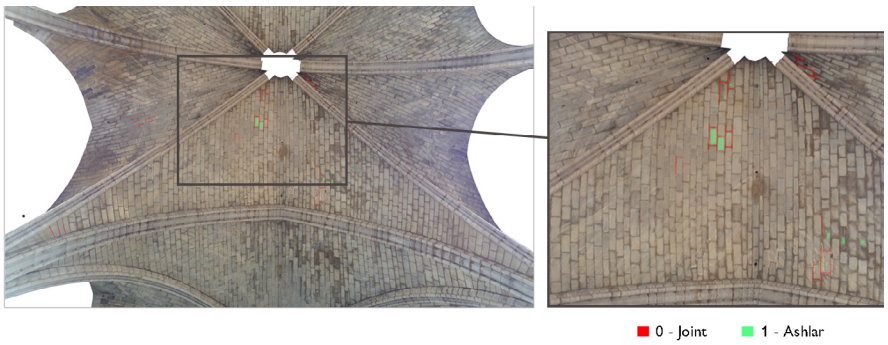
e.g.: Notre-Dame Cathedral in Paris



Figure 5-38.

Training set for the first segmentation phase.

e.g.: Notre-Dame Cathedral in Paris



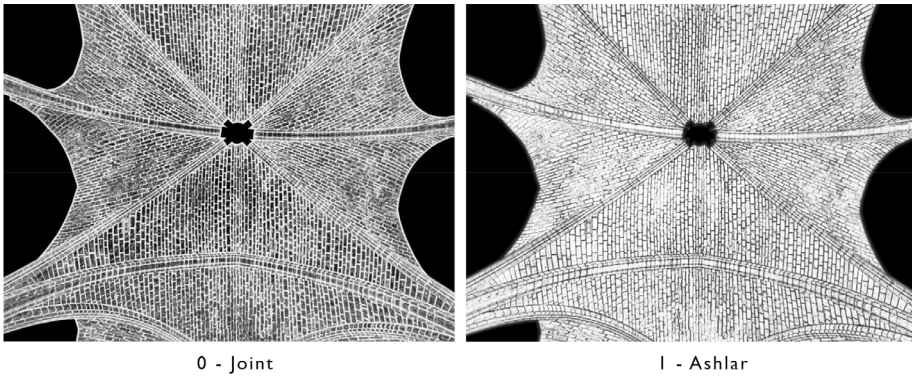


Figure 5-39.

Probability maps: the whiter the pixel is, the higher the probability that the pixel belongs to the given class, respectively class 0 – Joint, and class 1 – Ashlar.

e.g.: Notre-Dame Cathedral in Paris



Figure 5-40.

Selection of the class 'ashlar' to be processed for further segmentation.

e.g.: Notre-Dame Cathedral in Paris

For each class of elements, the probability maps of Figure 5-39 show the probability that a given pixel belongs to the chosen class: the whiter the pixel is, the higher the probability that the pixel belongs to the given class. From the observation of these maps, it is noted that most uncertainties in classification are found in the external areas of the images, i.e., at the impost of the vaults. In such areas, in fact, the proximity of stained glass windows of the cathedral causes light alterations that affect the classification results.

Figure 5-41.

Training set for the second segmentation phase.

e.g.: Notre-Dame Cathedral in Paris

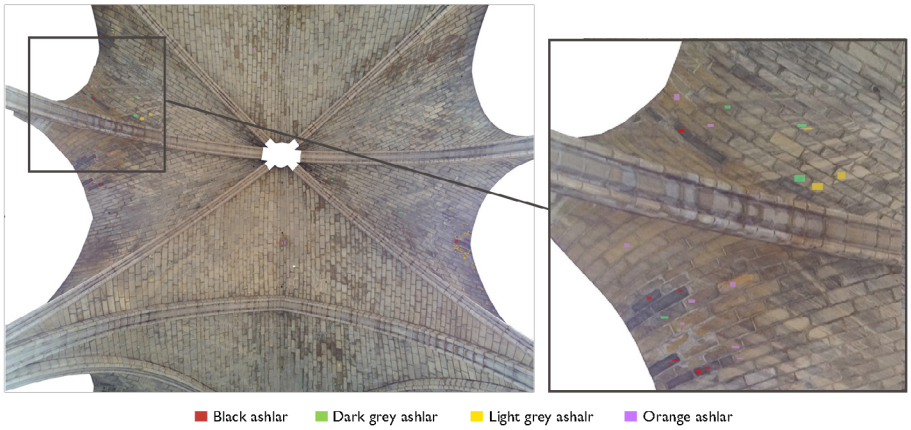


Figure 5-42.

Segmentation results for the image sets of the intrados of the vaults (second segmentation phase).

e.g.: Notre-Dame Cathedral in Paris



Once joints are distinguished from ashlars, and the image pixels corresponding to the ashlar class (class 1) isolated (Figure 5-40), the second segmentation phase is performed: in this case, the ashlars are distinguished according to their color, respectively detecting the classes of black, dark grey, light grey or orange ashlars.

Starting from the annotation of the classes on a reduced portion (Figure 5-41) and, for this second phase, from the extraction of colorimetric features alone, the result of Figure 5-42 is obtained.

The individual classes of ashlars obtained as a result of the segmentation process can be isolated and distinguished based on color values, also showing the related pixel variation in the RGB space (Figure 5-43).

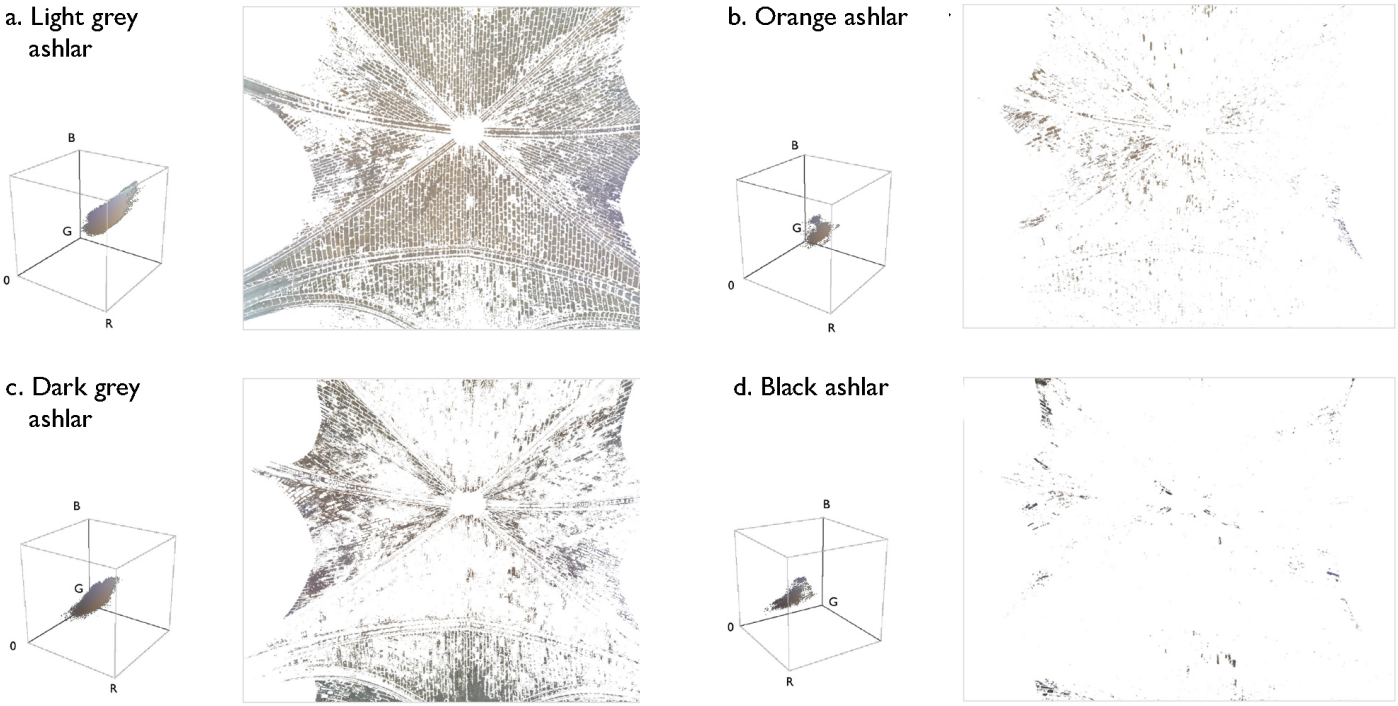


Figure 5-43 *(above)*.

Distinction of ashlars based on color values, with related pixel variation in the R, G, B space.

It is worth highlighting, in this section, that the possibility of distinguishing the ashlars from the recourses and of selecting them by color classes opens the way to further data analysis and comparison tests. For instance, starting from the number of existing pixels for each class, a certain series of shape descriptors can suitably be extracted for the single ashlars identified: to give an example, the regions of interest concerning a single ashlar can be assigned a name and a number, and analytical values can be derived considering the number of pixels constituting the single ashlar. The latter can relate, e.g., to perimeter, length and width of a bounding rectangle (smallest rectangle enclosing the selection), average width and height of each region, area of the selection (in square pixels or calibrated square units) and so on.

Figure 5-44 provides an example for this kind of descriptor extraction process, derived thanks to the Analyze Particles tool via ImageJ (imagej.nih.gov/ij/docs/guide/146-30.html).

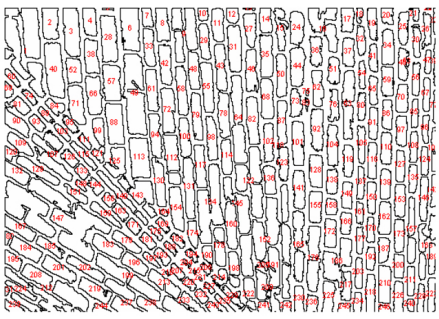
Figure 5-44 *(below)*.

Extraction of information on pixels: classified ashlars (a); regions of interest numbered (b) and extracted properties (c).

e.g.: Notre-Dame Cathedral in Paris



a.



b.

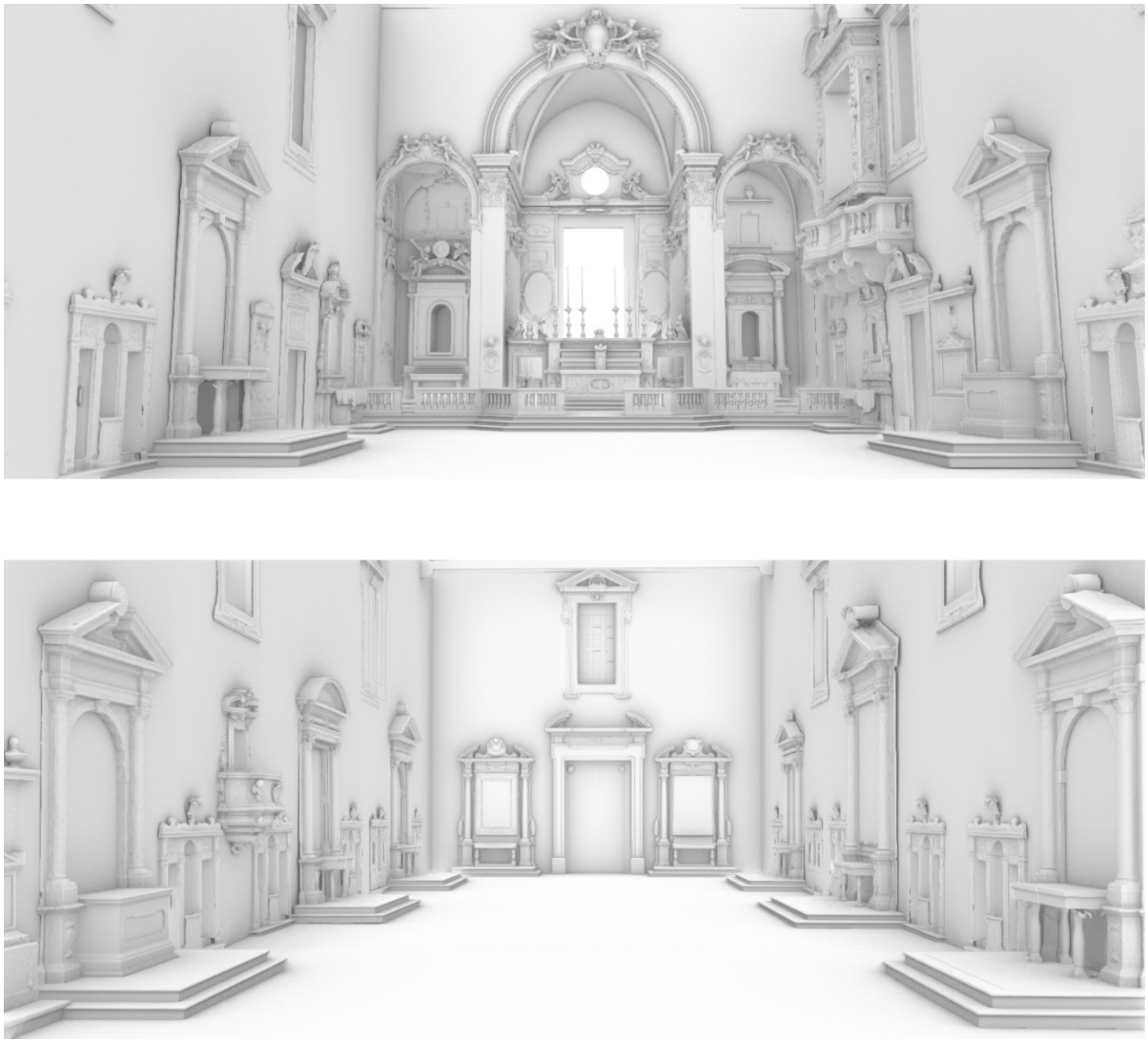
	Area	Mean	Perim.	BX	BY	Width	Height
1	2975	141.576	562.215	0	0	50	109
2	695	142.053	118.184	45	0	21	38
3	1061	148.574	158.326	71	0	21	60
4	507	151.049	98.527	95	0	18	34
5	29	151.793	31.071	119	0	14	3
6	1084	150.171	166.912	141	0	24	53
7	271	144.572	70.627	167	0	15	23
8	602	139.159	118.426	184	0	18	41
9	1358	144.440	314.066	203	0	33	66
10	174	147.713	53.213	236	0	13	16
11	545	142.158	105.941	252	0	16	41
12	250	134.644	60.385	271	0	15	19
13	25	141.880	33.314	291	0	15	3
14	418	147.278	88.042	311	0	16	30

c.

Figure 5-45.

3D models of the interior, completed with mesh models of the altars.

e.g.: Chiesa del Carmine in Pisa.



The shape descriptors are provided here in pixel values, but it is envisaged that a proper spatialization of the image with respect to the 3D model would allow to retrieve 2D/3D information at scale. Moreover, in the case of photogrammetric models, the retrieval of correspondences between annotated image and 3D model would allow to transfer classified regions of interest to the point cloud, and further propagate 2D/3D information.

5.3.2.3. Classification of UV maps.

As for the case of orthophotos and images, UV maps (Paragraph 3.2.2.6) can also be considered for the classification task, if suitably extracted from 3D photogrammetric models.

To conclude the experimentation results of this chapter, the case study

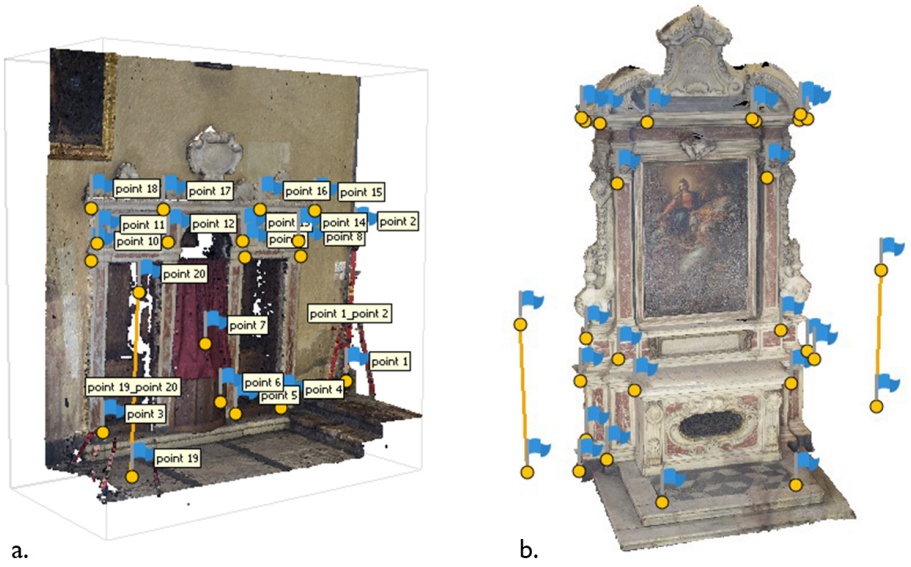


Figure 5-46.

Photogrammetric models of a confessional (a) and of the altar of Santa Vittoria (b).

e.g.: Chiesa del Carmine in Pisa.

of the *Chiesa del Carmine* in Pisa is herein presented.

Built by the Carmelites at the end of the 14th century, the church is composed of a single nave covered with a gable roof. The enrichment of the internal decorative face dates back to the 17th century: along the side and back walls of the church, sumptuous baroque altars stand, placed on raised floors and closed by two columns, mounted on a base, which flank the altar table and conclude at the top with a variously shaped crowning.

The complexity of the liturgical decoration of the church (Figure 5-45) required a detailed survey of the altars and confessionals, which allowed to work on textured mesh models at high resolution and then on the elaboration of UV maps for the semantic mapping of materials (Taddeucci, 2021). In particular, the existing laser scanner survey of the Church is integrated with the ground-based photogrammetric survey (Figure 5-46), to return the colorimetric characteristics of altars and liturgical furnishings.

From the textured mesh of each altar and confessional, and through the control of the seam lines, the UV maps are exported (according to the procedure previously described in Paragraph 3.2.2.6) and later annotated in 2D for the semantic segmentation process (Figure 5-47a).

The classes identified, in this case, relate to the semantic mapping of materials. The paintings of the altar (the canvas material) are manually excluded from the classification, since trivial to segment over the UV map. Then, the different colors of the marble slabs (white, pink, yellow marble) are distinguished on a reduced training portion (Figure 5-47b). Figure 5-47c shows the segmentation results for the altar of Santa Vittoria, annotated based on the type of material.

Figure 5-47.

Altar of Santa Vittoria, UV map (*a*), training set (*b*) and segmentation results with related legend (*c*).

e.g.: Chiesa del Carmine in Pisa

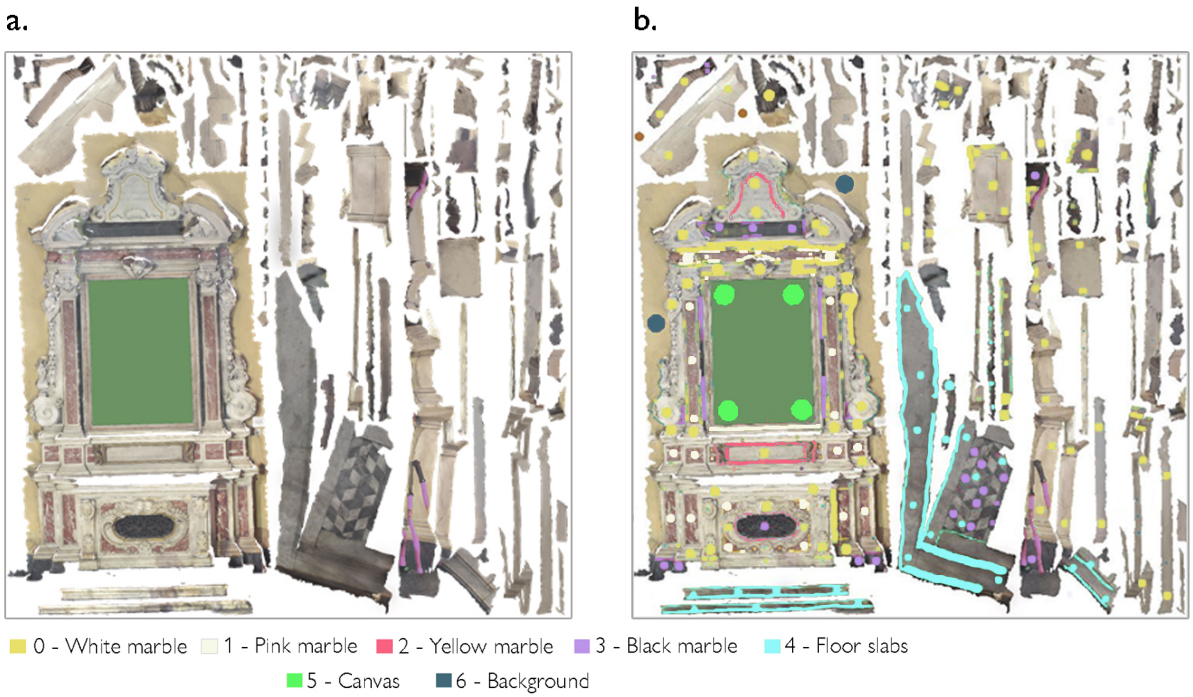
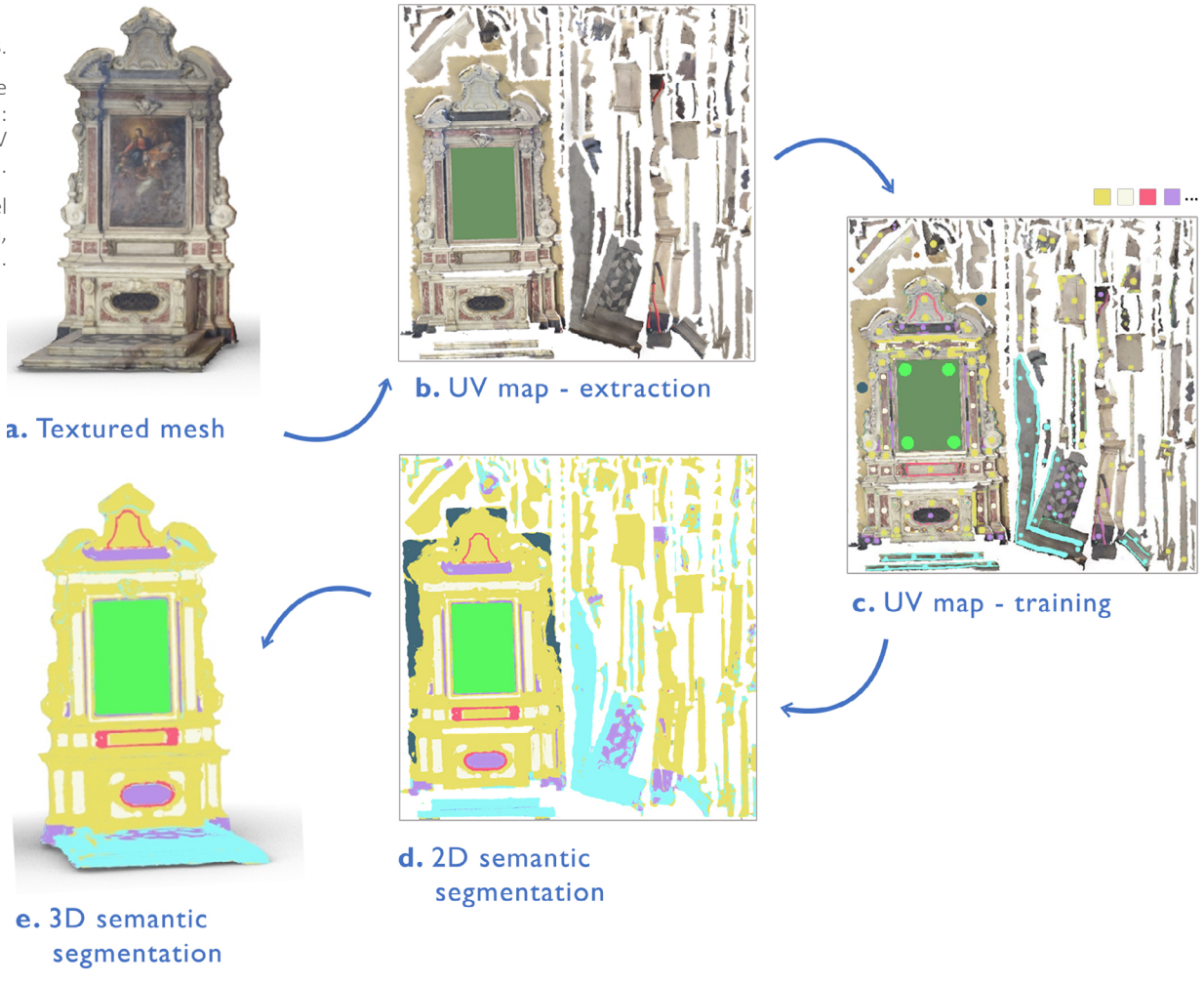


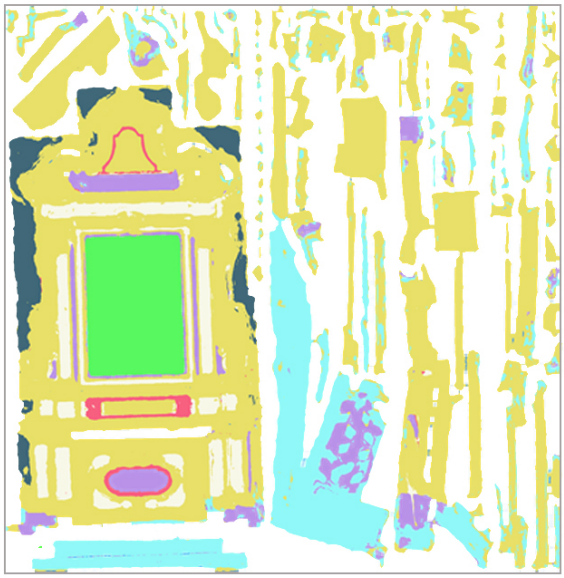
Figure 5-48.

Steps of the classification process: from the classified UV map to the 3D model.

e.g.: Chiesa del Carmine in Pisa, Altar of Santa Vittoria.



c.



At a final stage (Figure 5-48), the supervised classification obtained on the UV map can be projected back in 3D by exploiting the projective relationships between images and model, so to obtain an overall mesh model in which different colors correspond to different degrees of degradation or different types of material (Croce et al., 2022).

Class 6- Background, related to the plaster of the side wall of the church, could even be used to support the appropriate selection of the mesh faces related to the altar, suitably excluding the faces related to the back wall.

5.3.2.4. Assessment of texture-based classification results.

Together with geometry-based approaches, texture-based classification is an essential tool to support the semi-automatic enrichment and semantization of digital heritage data, facilitating digital data interpretation and sharing.

Supervised annotations on textures can be fundamental as a support to restoration operations, which require semantic mapping and comparison between successive conditions of degradation of surfaces and/or of material states. Once again, the annotation time is relatively short compared to traditional manual methods.

However, as a supervised approach, the operator's choice of the classes and training set used for learning is fundamental in determining the success of the classification process. Since solely relying on texture, in fact, the identification of classes can be tricky in cases where two classes have similar color characteristics or patterns. Analogously, the presence of differently lit surfaces or transparent elements may result in misclassifications. To overcome this problem, it could be envisaged to preliminary and rapidly uniform and adjust color values in those portions of the image that are expected to be misclassified.

In any case, it should be noted that, through the photogrammetric process, several images together, if properly acquired, can be used to build a 3D model. As a consequence, if the classification is performed over an image or a set of images, there is a way to transfer information (in the form of semi-automated or manual annotations) from the 2D to

the 3D space. As such, the application of texture-based classification approaches to imagery, orthophotos and UV maps can be directly linked to the use of imagery in the photogrammetric process. It can thus be expected to transfer semantic information from 2D to 3D and vice versa:

- By retrieving the projective relationships between textured mesh and UV map.
- When considering an image set used for photogrammetry, by propagating annotations from image to image and from image to 3D model (and vice versa).
- On collaborative reality-based platforms such as Aïoli, by combining manual and semi-automatic annotations, always with a view to 2D/3D information branching.

These considerations open the way to the topics that will be presented in Chapter 6, the latter specifically addressing the propagation of semi-automated semantic annotations between different 2D/3D media.

Chapter 6

2D/3D Information propagation

The previous sections of this work underlined how the use of semantic annotations, more or less automatic, enables the enrichment of 2D and 3D digital heritage representations, by addition of a semantic attribute to the raw data obtained from the observation and surveying of the built environment. In Chapter 5, in detail, more automatic annotation methods were investigated, relying on the application of AI and supervised ML algorithms.

In the present chapter, the semantically annotated data (obtained either manually or more automatically) is exploited not only as the informative resource of a single digital representation *per se*, but rather as a starting point to enrich multiple types of representations of the same architectural object. In other words, the second stage of the proposed methodological approach is focused on preserving semantic annotation information in the transition:

- to a different type of representation (point cloud, mesh, image, ortho-photo...);
- to successive levels of digital resolution and/or detail.

With the aim of ensuring the *digital continuity* of heritage data, as intended in Section 2.4, we wish to guarantee that the annotation information, whether automatic or manual (hence, regardless of how it was obtained), is preserved in the transition, e.g., from images to 3D model, or from more detailed 3D model to less detailed 3D model and so on.

In this regard, the information transfer and propagation is investigated in this chapter with a multi-source and multi-scale approach, exploring

in detail:

- The automatic annotation transfer from 2D to 3D and vice versa, by retrieval of the projective relationships linking images and models (Sections 6.1, 6.2).
- The multi-level and multi-resolution semantic segmentation for the transfer of information within 3D models at different levels of detail (Section 6.3).

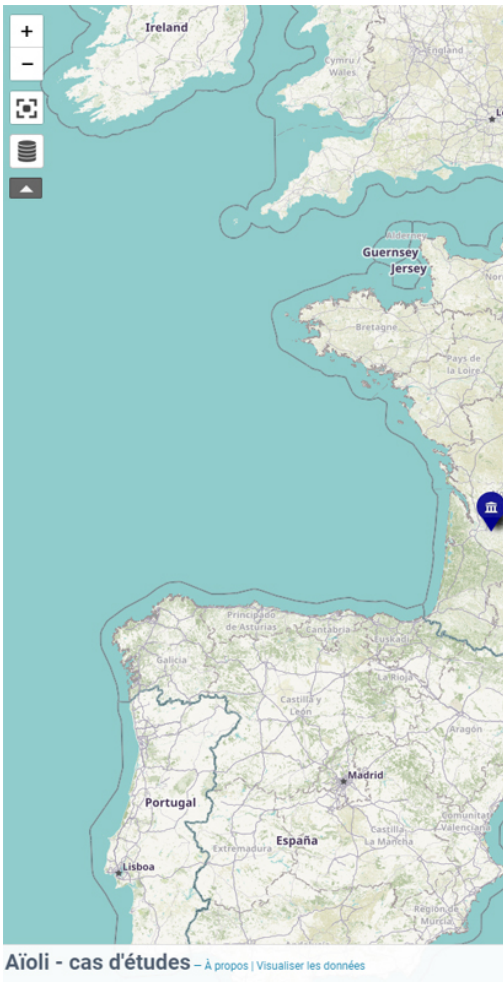
As for the first point, Section 6.1 features some further possibilities for the semantic enrichment of digital representations in hybrid 2D/3D annotation mode that take advantage of the most recent developments of the Aïoli collaborative reality-based annotation platform, developed by the MAP CNRS/MC research unit.

As previously introduced in Paragraph 3.3.3.3, the platform relies on the propagation of information from 2D to 3D and vice versa, thanks to the retrieval of projective relations that guide the construction of photogrammetric models. This underlying principle of Aïoli is exploited herein to transfer semi-automatically obtained annotations through the reprojection space between images-point cloud-images.

After a concise overview of Aïoli's operating rules, an implementation of the platform is presented, with reference to the insertion of semi-automatic annotation processes. The collaborative annotation tools, enabling 2D/3D transfer of semantic data, further confirm the platform as a fundamental support tool in the management and multi-source sharing of different 3D models and images, related to heritage artifacts all over the world.

The transfer of classification information is further addressed in Section 6.2, either considering as a starting point the image-based segmentation (annotation transfer 2D-> 3D) or the point cloud- or mesh-based segmentation (annotation transfer 3D -> 2D).

Finally, a multi-level and multi-resolution approach is presented for the transfer of semantic data between point clouds at different levels of detail: applied to the significant case studies of the Notre-Dame Cathedral in Paris and of the Church of Santa Caterina in Pisa, this approach allows to segment point clouds based on the identification of successive levels of detail and resolution, hence of interpretation, of the 3D raw data (Section 6.3).



6.1. Collaborative 2D/3D annotation transfer through Aïoli

6.1.1. Spatialization of 2D/3D data within Aïoli

Aïoli is a collaborative annotation platform designed with cloud-computing services and dedicated to the semantic annotation of heritage objects, including buildings, monuments, archaeological sites and so on. The annotation is said to be collaborative as a single project can be shared via web with multiple actors (conservators, architects, engineers, restorers...) involved in the documentation and analysis of the heritage artefacts.

As anticipated in Section 3.3.3, the platform can be accessed on-line

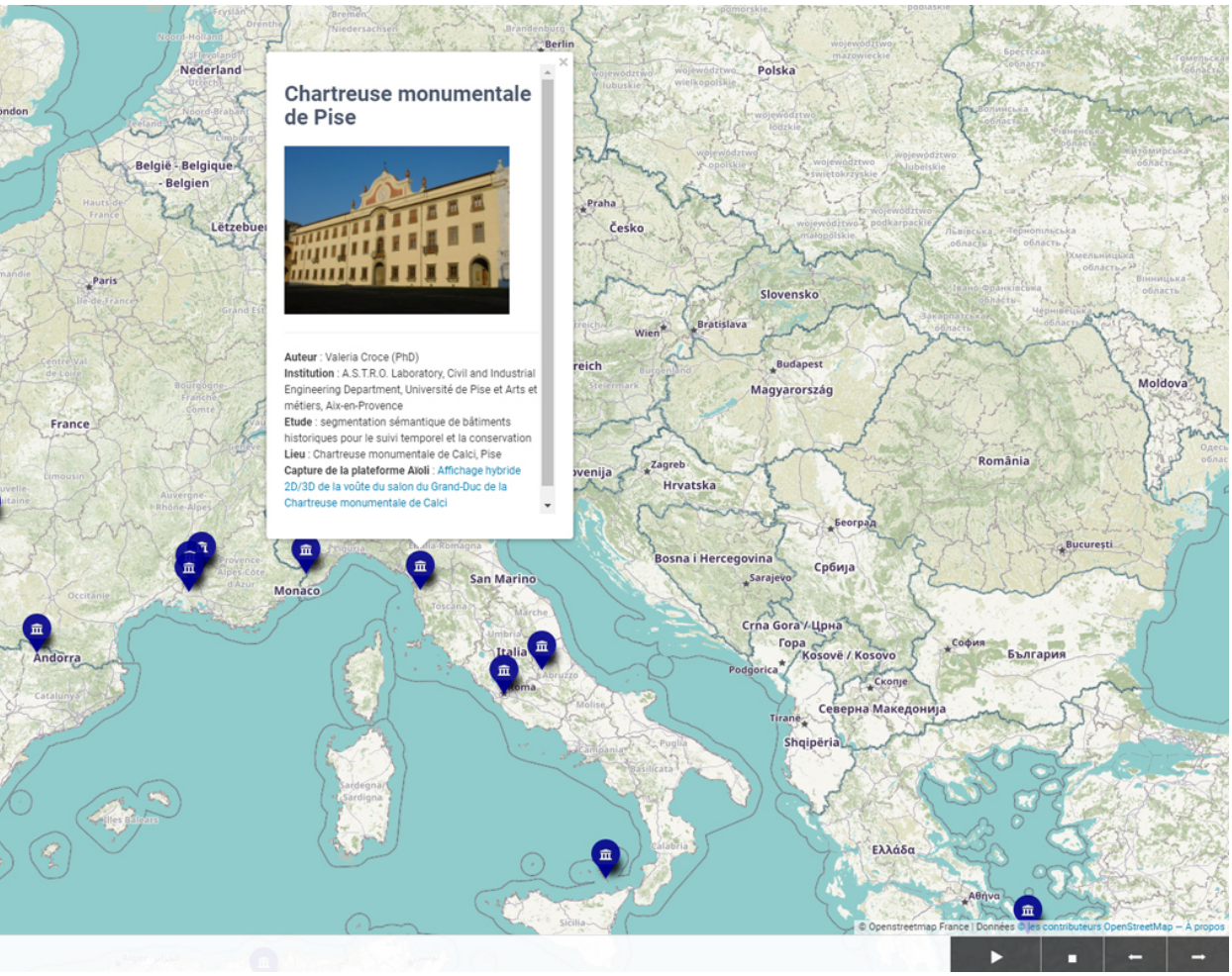


Figure 6-1.

Interactive map featuring Aïoli projects for different case studies of the European heritage (*source*: aioli-cloud; framacarte.org/fr/map/aioli-cas-detudes_69435)

e.g.: Visualization of the Pisa Charterhouse project

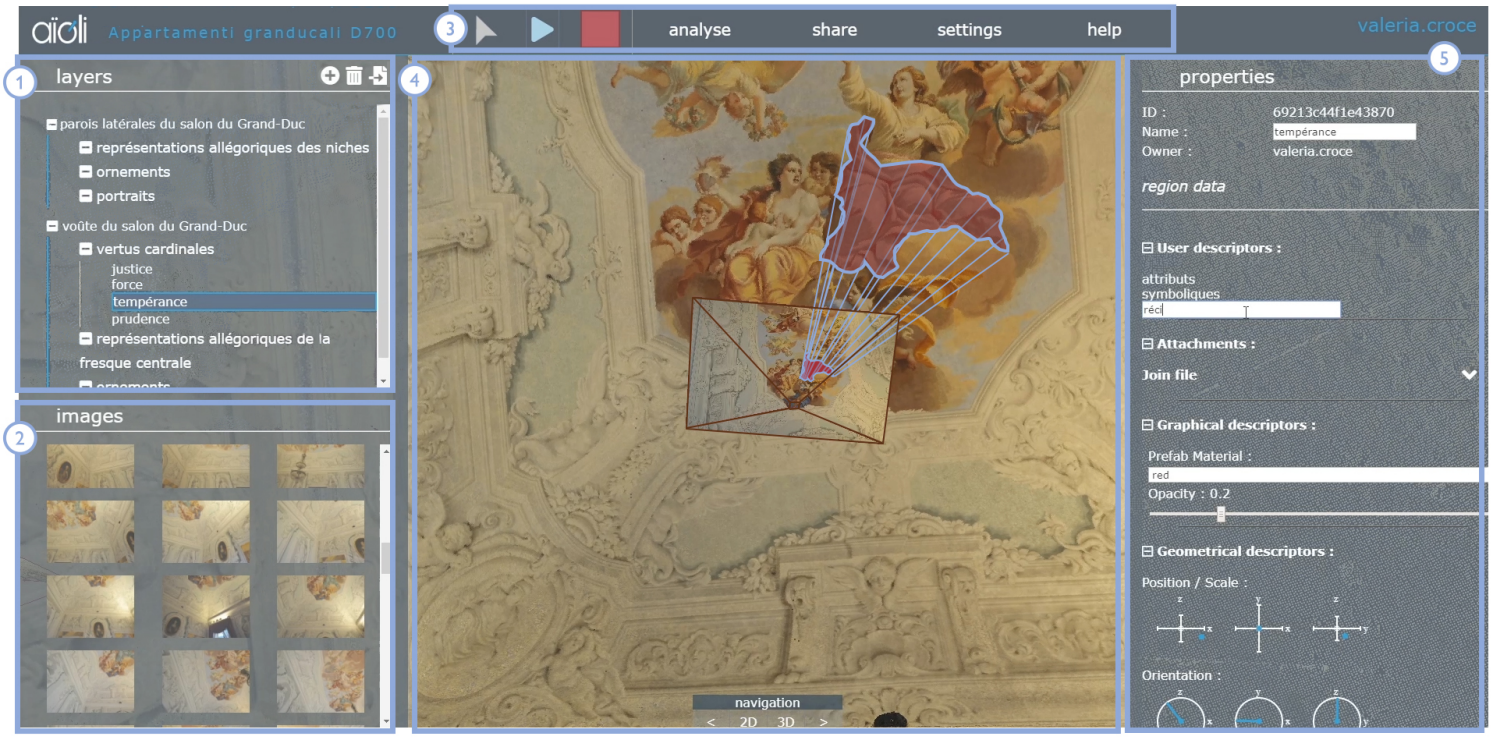


Figure 6-2 *(above)*.

Aïoli interface: annotation layers (1); image sets (2); annotation tools, sharing and visualization options (3); 2D/3D viewer for the spatialization of annotations (4); annotation descriptors (5).

e.g.: Vaults of the Grand-Ducal apartments, Pisa Charterhouse

by the end user, who uploads a series of photographs, appropriately acquired, representing a studied object from different points of view.

The uploaded photographs are processed to generate a dense point cloud, around which the original images are spatialized (Figure 6-2).

By means of a series of annotation tools, the user can select each time a relevant image of the dataset and make the annotation.

Each annotation performed is uniquely determined by an identification code (ID), automatically generated. The annotation can even belong to different layers of information (corresponding to different levels of observation and representation), and it can be associated with textual or numerical descriptors specified by the user.



At the end of the selection process, the annotation is propagated onto the 3D model, and it is later projected back onto all the remaining images where the same region of interest is visible. As such, the annotation made appears on the initial image, but also on all other 2D (images) and 3D sources (point cloud) in which it is present.

These operations are made available

within the Aïoli platform through the manipulation of annotation tools and utilities, provided with a user-friendly interface (Figure 2-1, Figure 6-2): besides the annotation layers menu (1) and the image set pane (2), a command bar (3) is conceived with tools and instruments supporting annotation selection and sharing, editing and visualization options. The 2D/3D viewer for the spatialization of annotations (4) is put at the center of the web-page, while, on the right, a properties menu enables browsing the ID associated to each annotation, as well as inserting and controlling user descriptors (5).

Annotation layers in 2D/3D can thus be associated with semantic descriptors, which qualify them as a whole, and each region inherits the descriptions of the parent layer. At a final stage, the annotated 2D/3D data can be exported and shared externally by means of an export component included in the settings toolbar.

Aïoli builds on photogrammetry, image processing, and indexing procedures to run. Starting form an initial image dataset, the image orientation and the dense point cloud generation are realized thanks to the principles of photogrammetry and from the computational algorithms of the open source photogrammetric software MicMac (Rupnik et al., 2017). Then, an indexing procedure is considered to associate the selected image annotation in pixel coordinates (i, j) with the corresponding points in 3D coordinates (x, y, z). Once the coordinates of the points concerned by the annotation are found, the inverse relationship 3D-> 2D is then used to back-project the annotation from the point cloud to the remaining images (Figure 6-3).

It has also to be noted that, when a point cloud is built on a set of images thanks to the photogrammetric process, Aïoli stores the 3D point cloud file with the basic metric-visual information of the 3D points (X, Y, Z coordinates, R, G, B and normals, as specified in Section 3.2), plus additional characteristics shaping the object's surface, e.g., curvature, roughness, ambient occlusion and depth maps, extracted

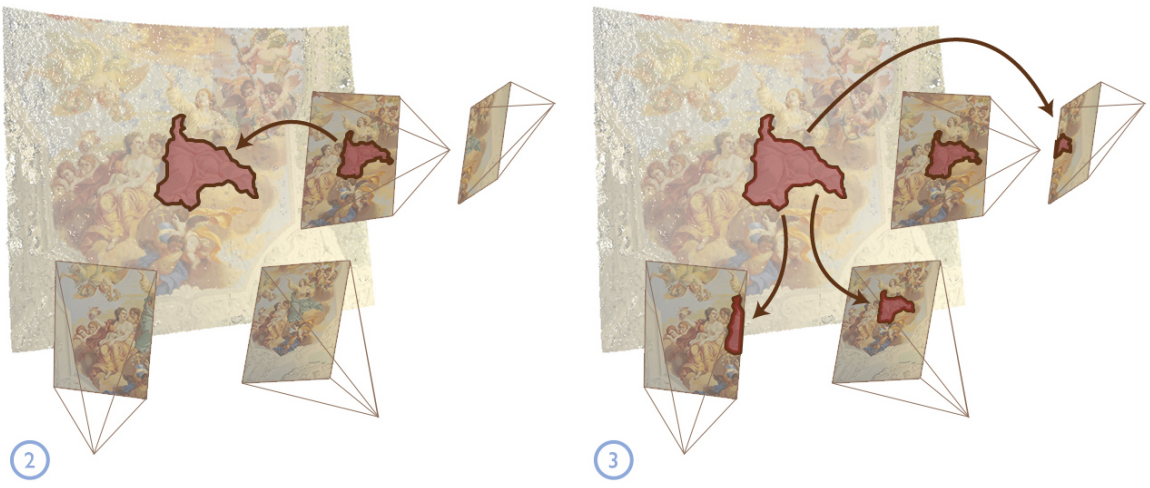


Figure 6-3 *(below)*.

Selection of a region over the first image (1); identification of the coordinates of the 3D points corresponding to the selection (2); back projection on other images. Adapted from (Manuel, 2016).

e.g.: Vaults of the Grand-Ducal apartments, Pisa Charterhouse

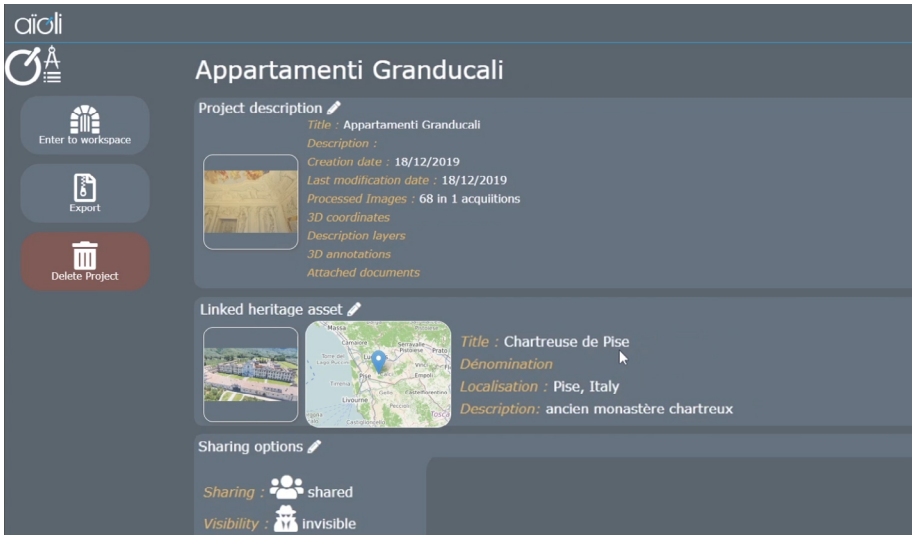


Figure 6-4.

Collaborative management: example of project sharing options within the Aïoli platform

e.g.: Vaults of the Grand-Ducal apartments, Charterhouse project).

from the raw data through a series of dedicated algorithms (Manuel, 2016). Considering the projective relationship image-point cloud, all this information can also be projected, by 2D/3D transfer, from the 3D model to the images, so to morphologically characterize the whole image dataset and express the various information as additional layers of the master photographs (Abergel, 2021). Complementary information can thus be associated with both 2D and 3D data when indexing an image.

The produced point cloud, images and annotations can finally be shared on a web-base with other actors involved (Figure 6-4).

A demonstrative video on the use of Aïoli is available at the link aioli.cloud/#GALLERY?the-video=2, and concerns the annotation of the frescoed vaults of the Grand-Ducal apartments of the Pisa Charterhouse. The video was realized to illustrate the platform in the context of the International Conference *Xèmes Rencontres Internationales Monaco et la Méditerranée,* held in Monaco in March 2020 (*Cas d'étude-Chartreuse Monumentale de Calci. Annotation du projet : Appartement de Grand-Duc de la Chartreuse Monumentale de Calci*).

6.1.2. Implementation of semi-automatic annotation methods via Aïoli

Considering on one hand the 2D/3D annotation tools offered by Aïoli (Figure 6 5), and on the other hand the ML-based classification process illustrated in Chapter 5, a question can be raised on the integration of semi-automatic annotation methods exploiting AI within the platform, so to fasten and better serve the collaborative annotation process.

In this context, the present work combines ML-based annotation methods with Aïoli's potentials in terms of annotation transfer between

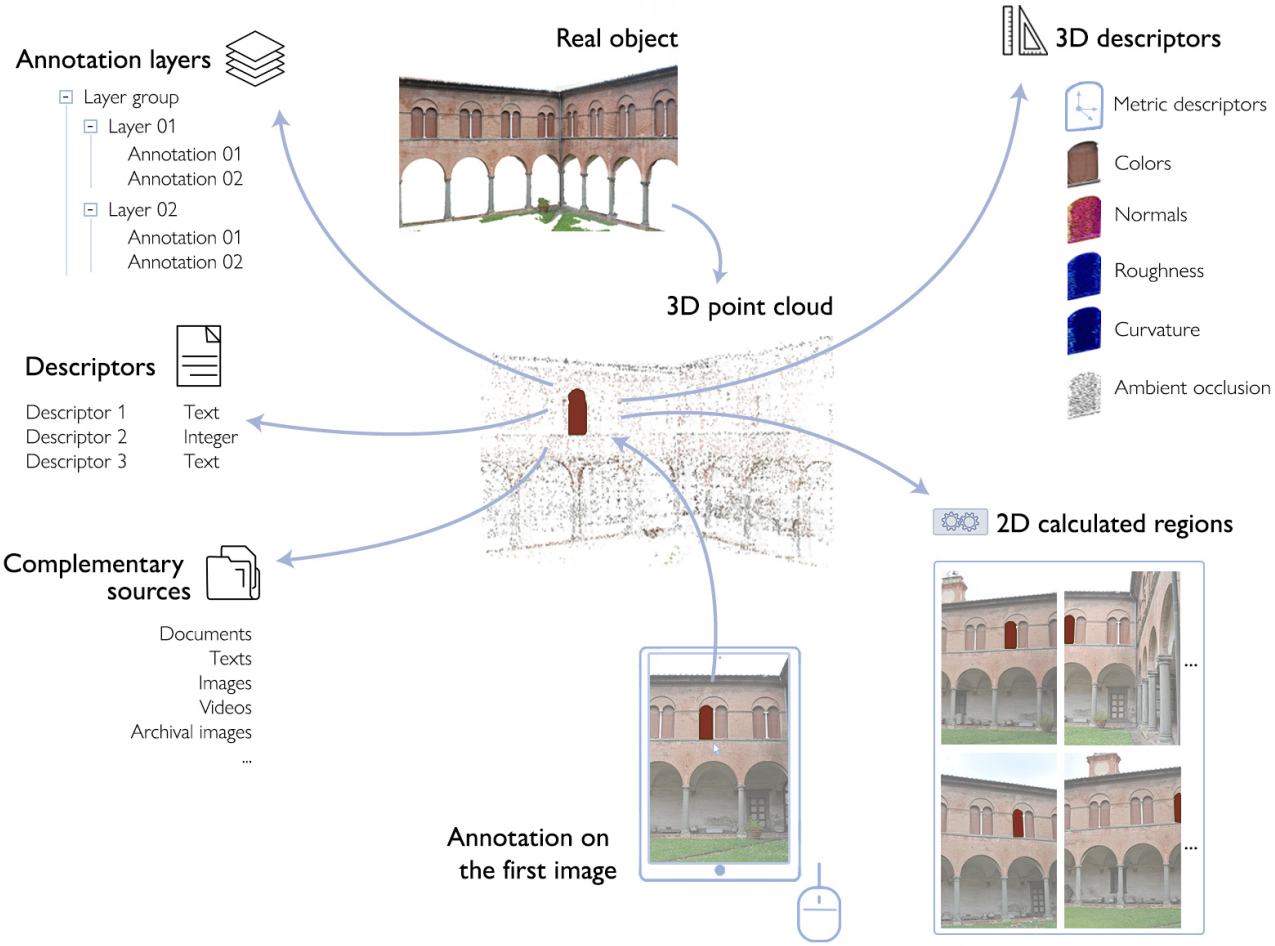


Figure 6-5.

2D/3D propagation and information insertion within the Aioli platform. Adapted from (Abergel, 2021; Manuel, 2016).

e.g.: National Museum of San Matteo, Pisa.

images-point clouds and between point clouds- images.

The basic principle guiding the following experimentations is to leverage the Aïoli platform for the construction of a 2D/3D training set, starting from the annotation of images and transferring onto the 3D point cloud the semantic information (the latter related, e.g., to architectural components, materials, degradation patterns and so on). The training set so identified in 3D is exploited to classify the whole point cloud, via the RF algorithm; at a final stage, the annotations propagated on the 3D point cloud are exploited to automatically annotate the images, by back projection of the semantic segmentation information.

The semi-automatic 2D/3D annotation approach developed is summarized in Figure 6-6 and Figure 6-7 and it is presented here with reference to the annotation of recurrent architectural components over the photogrammetric model of the cloister of San Matteo.

Firstly, using the Aïoli platform interface, the images are annotated by identifying, on a reduced portion of the image set, the classes of architectural components in which the dataset is to be subdivided.

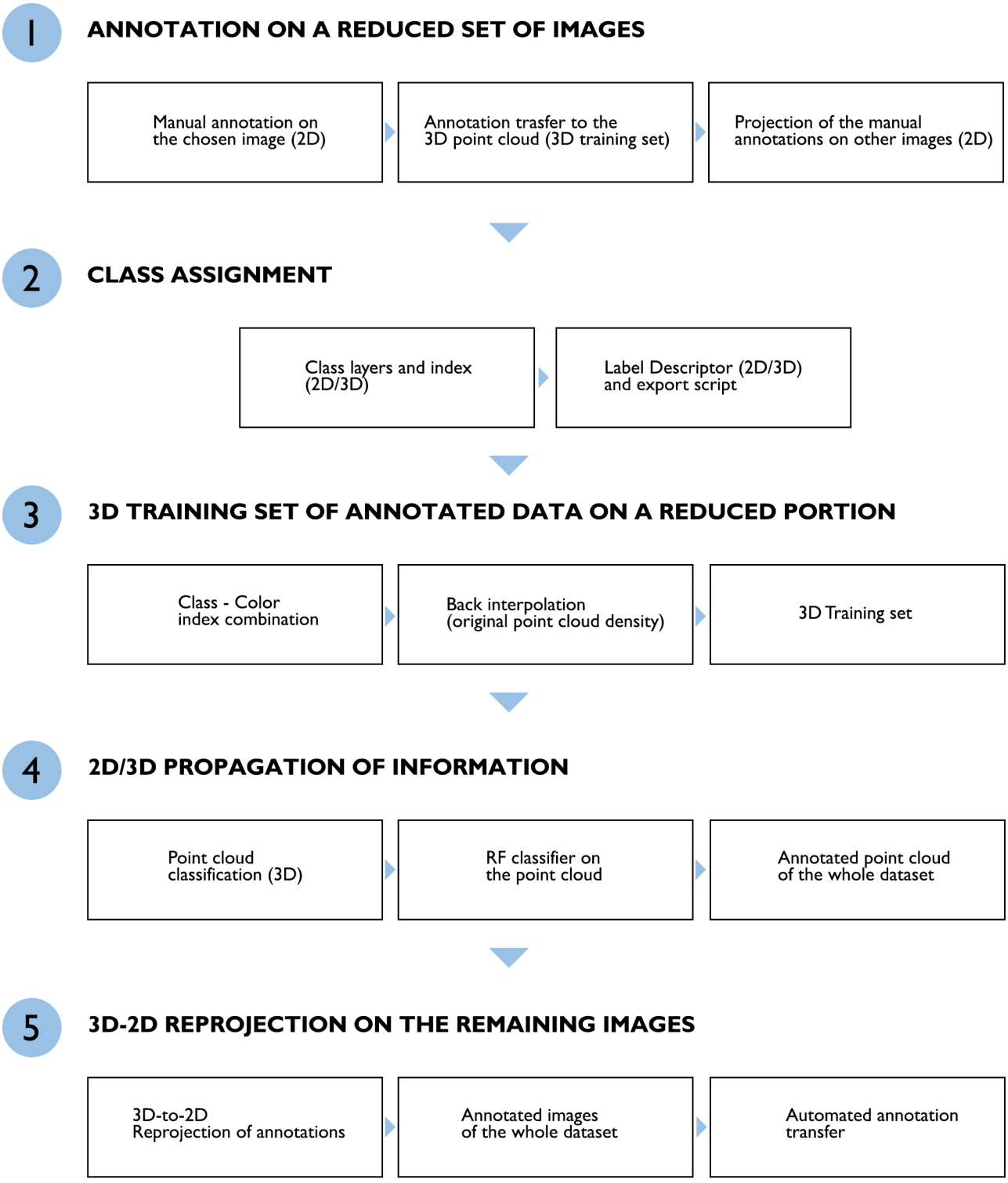


Figure 6-6.

2D/3D transfer of semantic information within the Aïoli collaborative reality-based annotation platform: overall workflow.

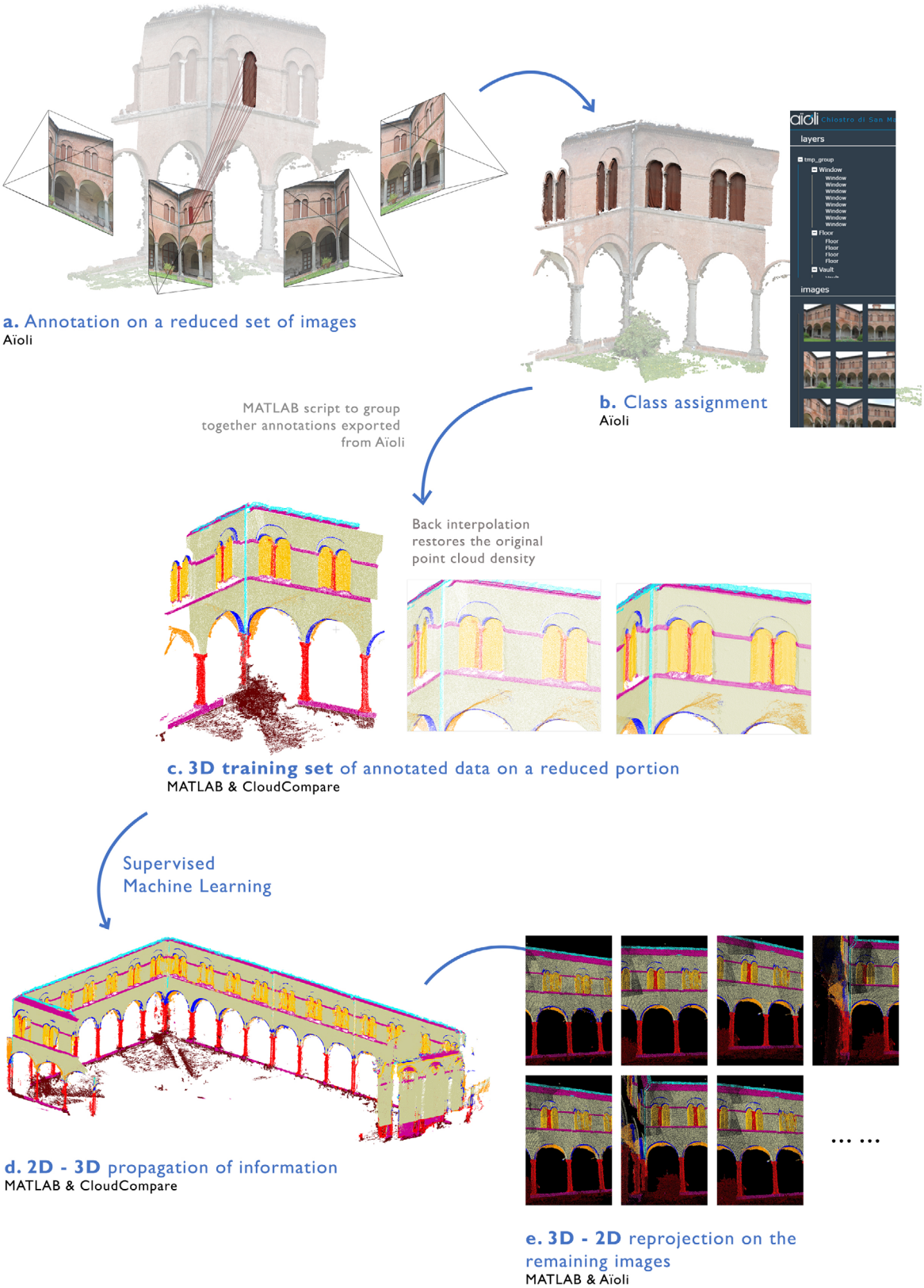


Figure 6-7.

Proposed approach for the 2D/3D transfer of semantic information within the Aïoli collaborative reality-based annotation platform.

e.g.: National Museum of San Matteo, Pisa.

Then, through the manipulation of textual and numerical descriptors, a class and a class index are assigned to the annotations performed, in the form of user-defined attributes.

Later on, the annotations so realized on the images are projected on the points of the 3D model, and the training set thus constructed is visualized (*projected*) on the point cloud. This 3D data is used to classify the entire point cloud, reiterating the segmentation approach based on supervised ML and previously described in Section 5.2.

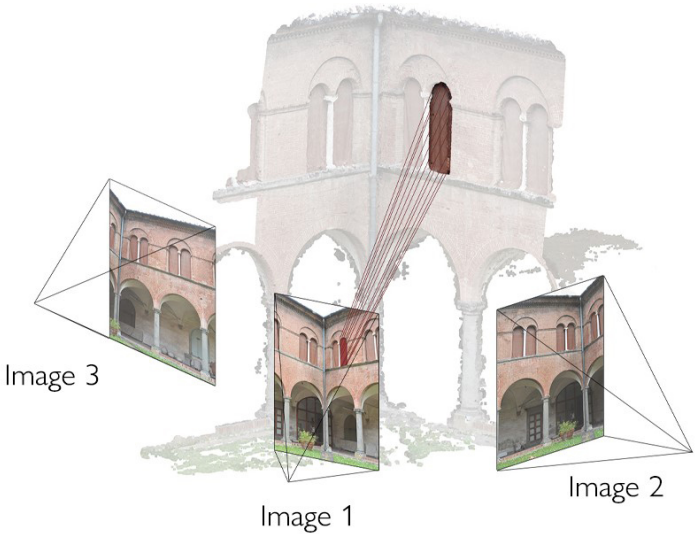
Finally, the annotations on the semantically annotated point cloud are exploited to annotate, via back projection, the remaining set of images.

6.1.3. Image-based annotation on Aïoli and construction of a 2D/3D training set

The San Matteo dataset considered for the 2D/3D semi-automatic segmentation test comprises a set of 96 images. The 3D point cloud is originally built by photogrammetric processing via the Agisoft Metashape software, and it is put in real scale thanks to the integration with a total station surveying.

A *Metashape-to-Aïoli import* script, developed in Python and previously released by the MAP research unit (Pamart et al., 2020), is exploited to properly import point cloud and related cameras in Aïoli.

Upon insertion of the point cloud, Aïoli stores the 3D coordinates corresponding to each pixel as a TIFF file, called XYZ file (Manuel et al., 2013a). In the TIFF file, color information is collected in form of a three-layer matrix where each column corresponds to a primary color (R, G, B). Instead of storing color information of each pixel, the matrix structure is exploited to store the calculated 3D coordinates corresponding to pixels: in other words, X coordinates are stored on the red layer, Y coordinates on the green layer and Z coordinates on the blue layer: by this process, the size of the XYZ



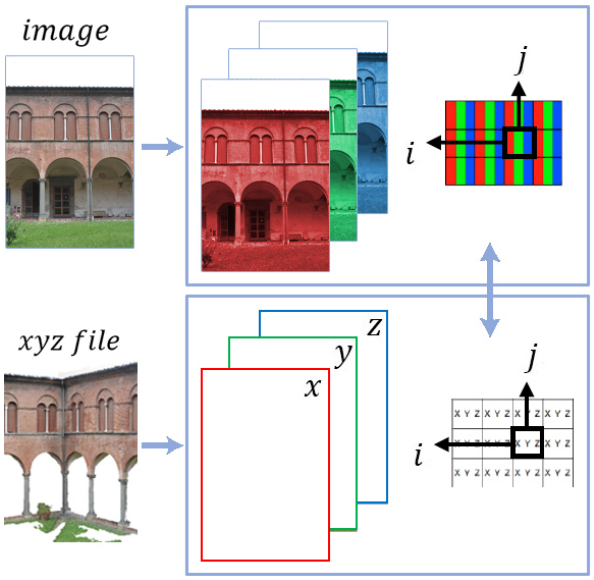


Figure 6-8.

Structure of TIFF and XYZ files. Adapted from (Manuel, 2013b).

e.g.: National Museum of San Matteo, Pisa.

file and the size of the TIFF image become comparable.

In order to identify the 3D coordinates of a pixel located at the *i*-th row and the *j*-th column of the image, the value of each layer of the *XYZ* file of the image is read at the same position (Figure 6-8, Figure 6-9). By this

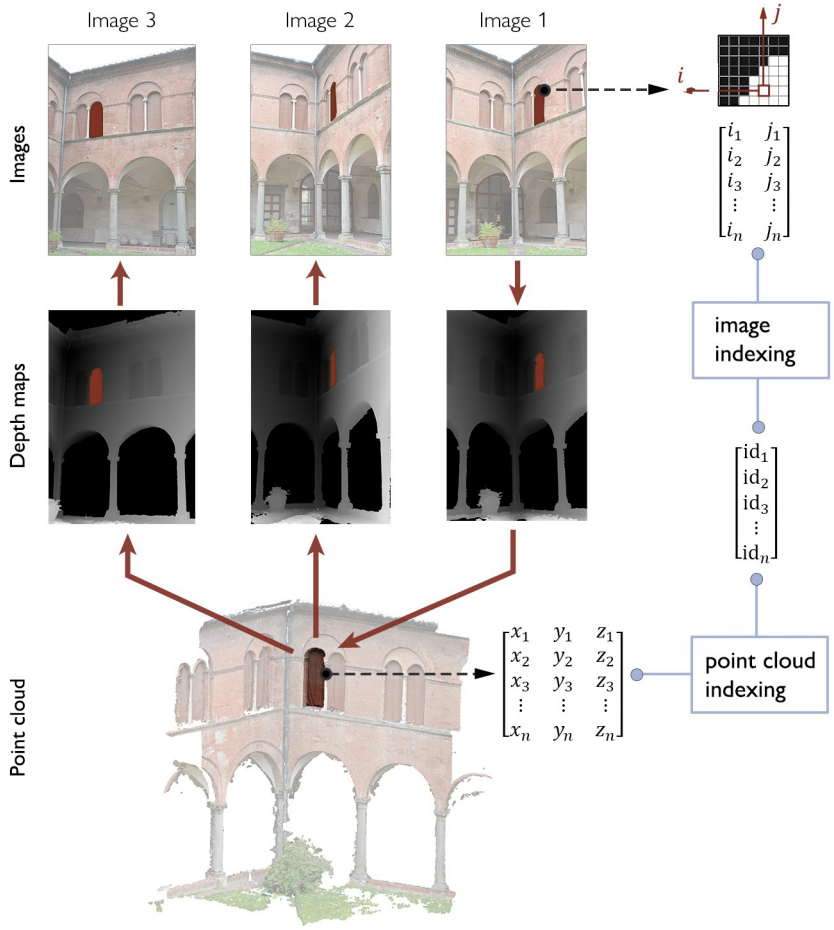


Figure 6-9.

2D/3D annotation transfer scheme: the operating principle of the Aïoli platform. Adapted from (Manuel, 2016).

e.g.: National Museum of San Matteo, Pisa.

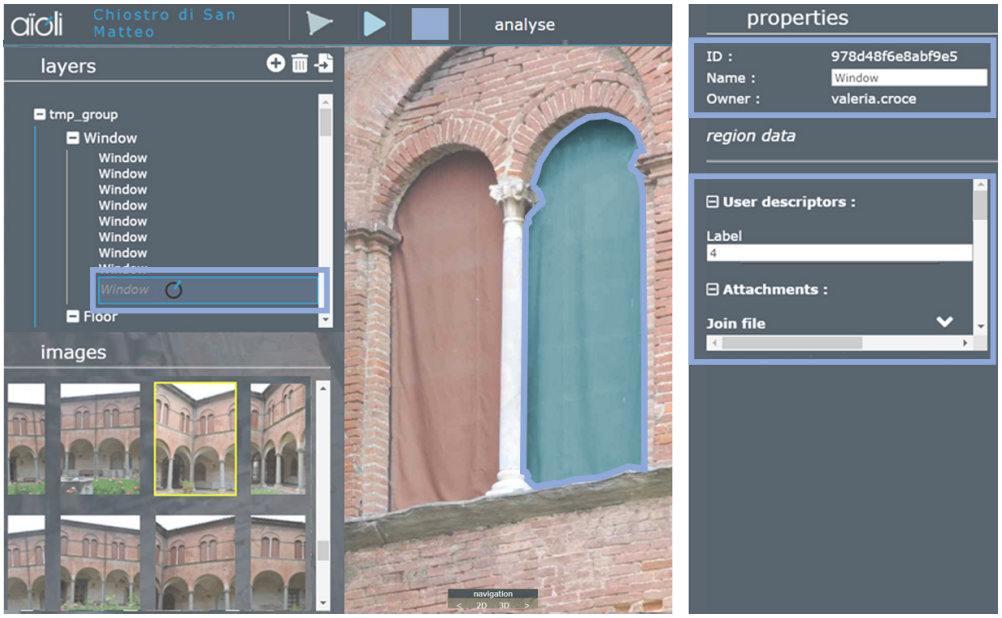


Figure 6-10.

An image is annotated with related class index as user-defined descriptor

e.g.: class 4- Window, National Museum of San Matteo, Pisa.

correspondence, the 3D coordinates associated to a pixel of the image can be extracted from the *XYZ* file (Manuel et al., 2013b).

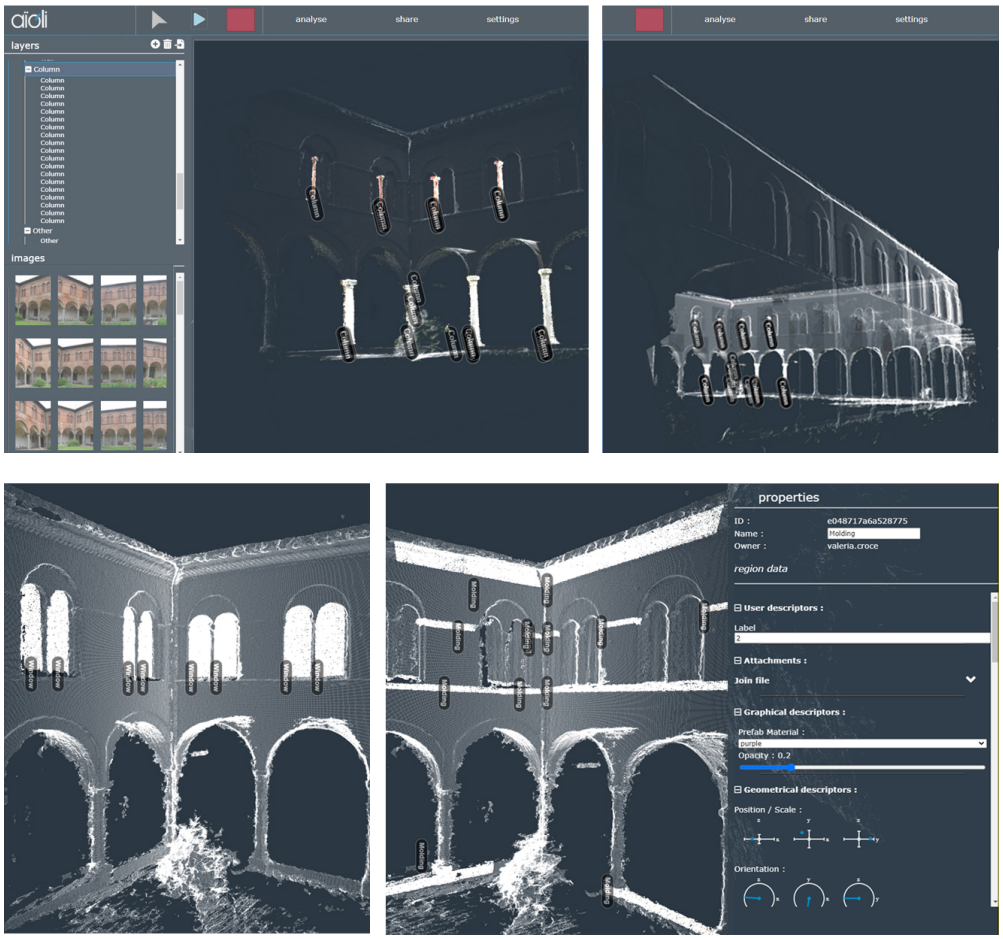


Figure 6-11.

Training set: Classes of Columns, Windows and Moldings, with related label descriptors, visualized on the 3D point cloud.

e.g.: National Museum of San Matteo, Pisa.

Once the project file on Aïoli is properly imported and structured according to this image-point cloud correspondence, the training set is constructed by annotating (in 2D) some relevant images (e.g.,Figure 6-10), in order to identify a sample data portion on which the ML classifier can be trained. For the San Matteo dataset, the classes of recurring architectural elements are sorted based on the ArCH dataset (Matrone et al., 2020), as in previous geometry-based segmentation tests (Section 5.3).

Thanks to the correspondence created within Aïoli between pixels and 3D points, the manually-annotated training set is propagated from the images to the point cloud. In detail, to properly characterize the 2D/3D regions of interest so constructed:

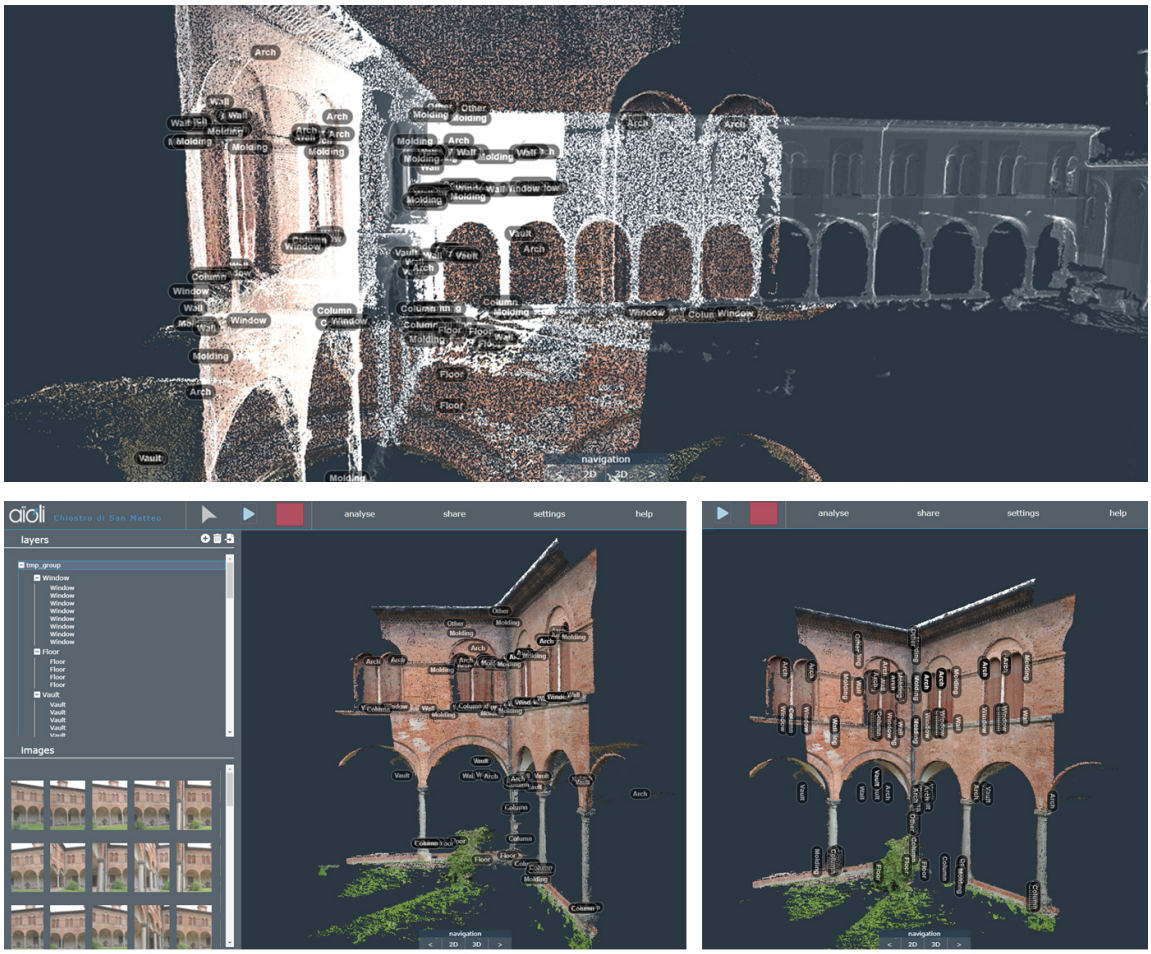
- Annotations are grouped according to the relative class they belong to, through the project layer control pane (Figure 6-11);
- A dedicated numerical user descriptor named 'Label' is created to feature for each annotation a class index (the value 4 in the example of Figure 6-10).

At the end of this first annotation process (Figure 6-12), a hybrid 2D/3D

Figure 6-12.

The annotations are transferred from images to point cloud to identify a training set in the 3D space.

e.g.: National Museum of San Matteo, Pisa.



training set is obtained, and it can be used to train a ML classifier, by applying a supervised classification approach.

6.1.4. IA-based point cloud classification and back interpolation

At a later stage, the training set of annotated 2D/3D data is exported from the platform: for each annotation, Aïoli stores in file folders, respectively, the contour lines of the 2D annotations traced, the image masks and the calculated points in the 3D space. The latter are defined both by coordinates and RGB values.

According to the structure of the export file, each annotation is included in the file folder corresponding to the parent layer group, i.e., the class of recurring typological elements, and it is identified by a unique alphanumeric code.

For the supervised classification, a scripting algorithm written and developed by the author in MATLAB (See Appendix 3 Aïoli class indexing script) enables to appropriately organize and unpack the exported data: in detail, the algorithm reorders the export folders and groups together in a single .ply file (Appendix 4 – Point cloud file formatas) the annotations that relate to a same class. The Label descriptor, identified by the class index, is later associated to each 3D point, together with X, Y, Z coordinates and RGB values. By doing so, the resulting .ply file created is a point cloud with an internal matrix structure that is similar to the one previously illustrated in Figure 5-9 (Section 5.2.4).

The portion of the annotated point cloud thus obtained is shown in Figure 6-13 and the identified classes can be displayed in 3D with related color legend. However, a problem of spatial density is observed on this initial training set: indeed, with the image-based annotation procedure in Aïoli, the semantic information is only transferred to a limited number of points corresponding to the selected annotation region, and

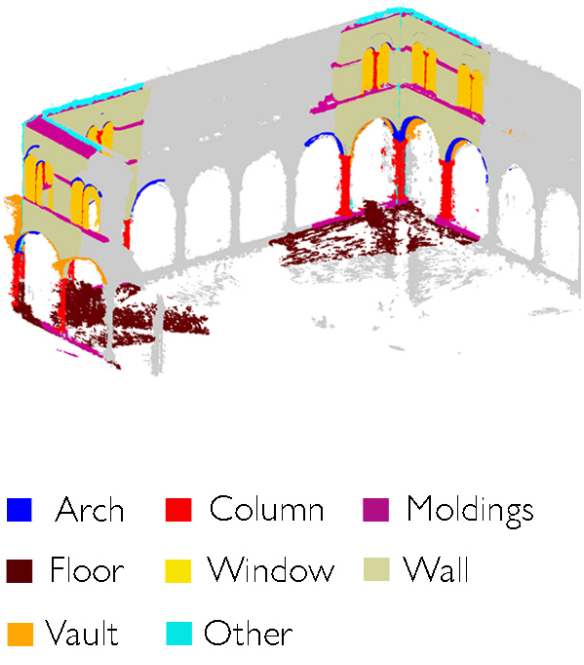


Figure 6-13.
Training set visualized on the 3D point cloud with color classes and related legend.
e.g.: National Museum of San Matteo, Pisa.

this necessarily implies a referencing problem between the density of the initial point cloud and that of the point cloud after annotation. This issue may also influence the extraction of geometric features, as the latter depend on the density of the point cloud in a defined local neighborhood (as previously observed in Paragraph 5.3.1.4).

To overcome this problem, a back interpolation of the label color is performed between the initial dataset and the annotated one, so restoring the initial point cloud density while preserving the classification information obtained (Figure 6-14). The back interpolation process is realized in CloudCompare (cloudcompare.org/doc/wiki/index.phptitle=Rasterize#Interpolating_empty_cells).

Considering this dense point cloud, the semantic segmentation process is put into practice, following the procedures of neighborhood selection, feature selection and extraction and RF training previously described in Section 5.2.

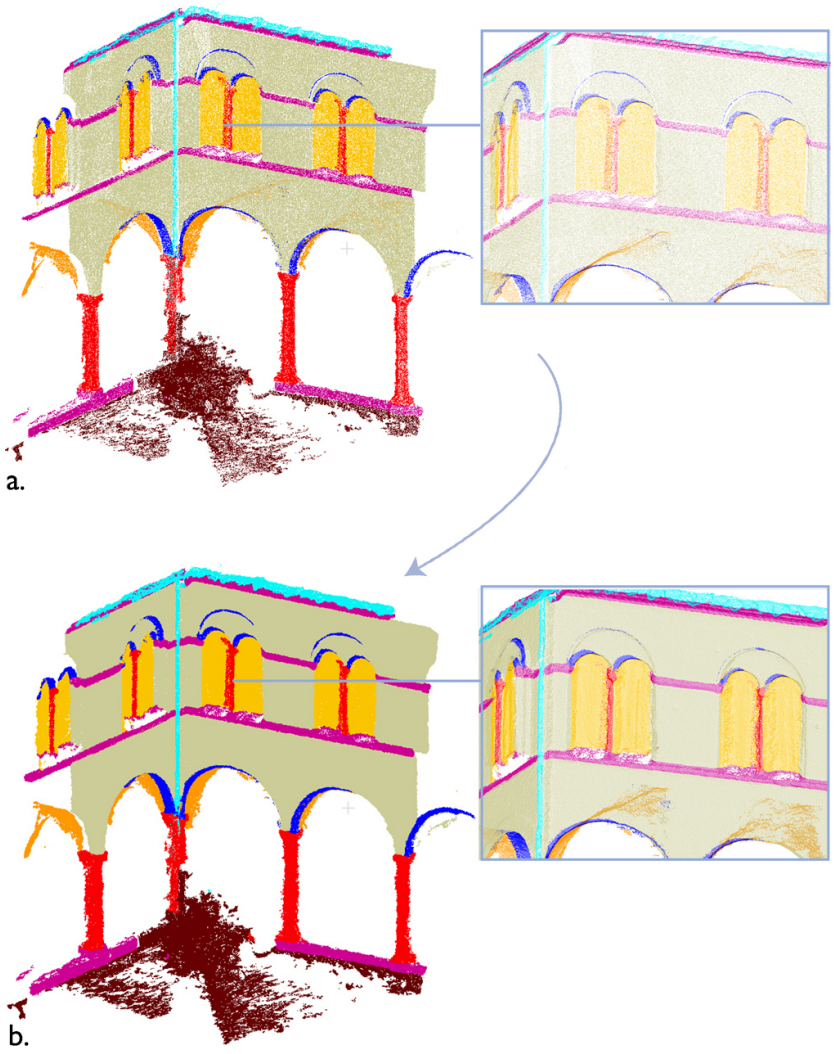


Figure 6-14.

Back interpolation of color values from lower (*a*) to higher (*b*) point cloud density.

e.g.: National Museum of San Matteo, Pisa.



Figure 6-15.

Training set (*above*) and segmentation results (*below*) on the point cloud.

e.g.: National Museum of San Matteo, Pisa.

6.1.5. 3D-2D back projection and annotation of the whole image set

The outcome of the semantic segmentation process by supervised ML can be visualized in Figure 6-15 over the point cloud.

At a final stage, the results obtained in terms of 3D annotation are leveraged to test the propagation of information from 3D to photos, so as to transfer the semantic datum to all other images of the initial set. Again, recovering the projective relations between images and the 3D model, the classification is propagated, in a semi-automatic way, to the remaining part of the image dataset, via back information transfer 2D -> 3D (Figure 6-16).

As the 3D points belonging to each annotated class are determined, in fact, the triplets of coordinate can be leveraged to retrieve the relative position (i, j) of each annotation in the XYZ file. In so doing, annotations will result simultaneously defined by areas on images and by a set of 3D coordinates in space, in a hybrid 2D/3D mode.

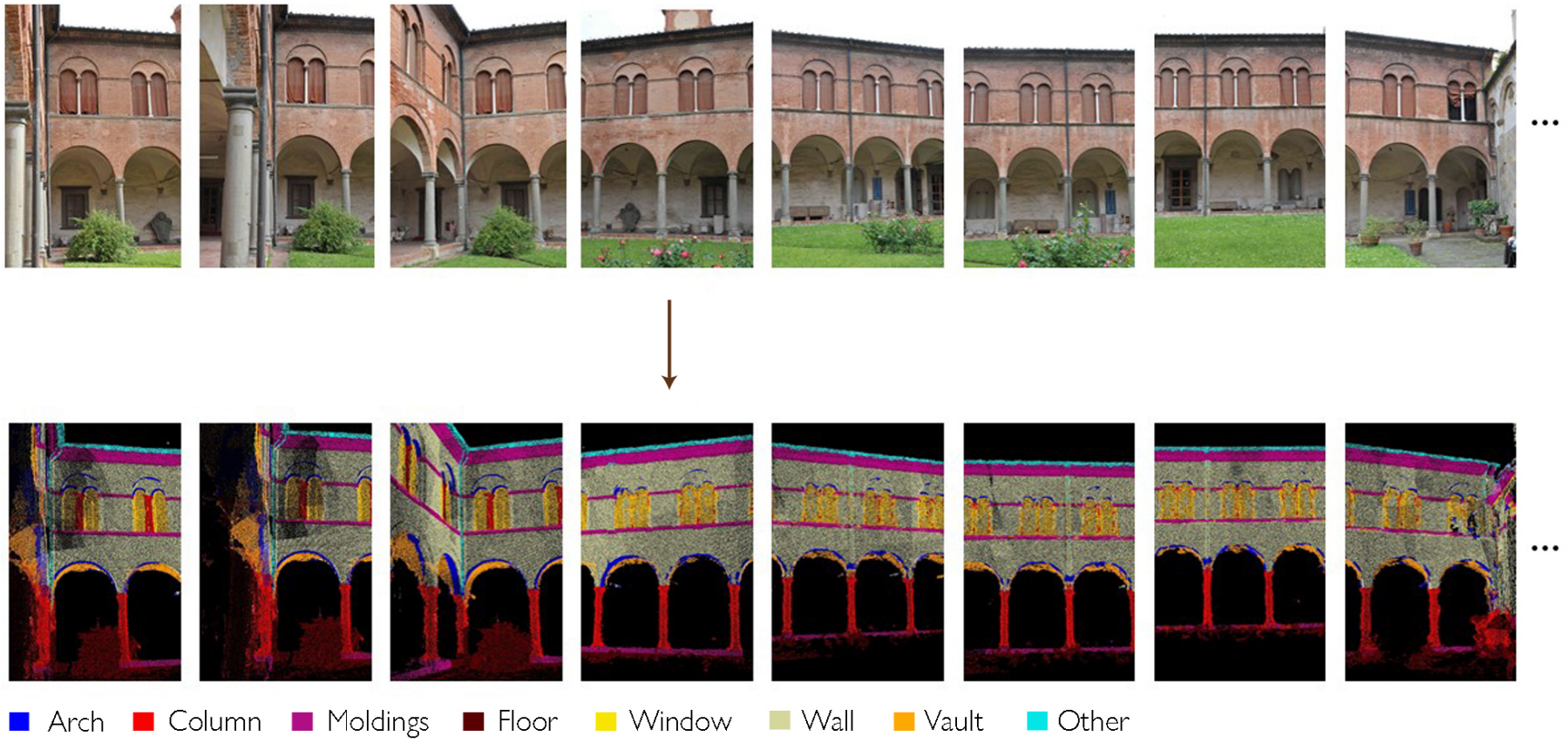


Figure 6-16.

Segmentation results on the images after back projection.

e.g.: National Museum of San Matteo, Pisa.

6.2. Additional 2D/3D semi-automatic annotation transfer possibilities

In the present Section, further possibilities of semantic enrichment are investigated, relying on the transfer of annotations between digital data of different types.

6.2.1. UV map – mesh models

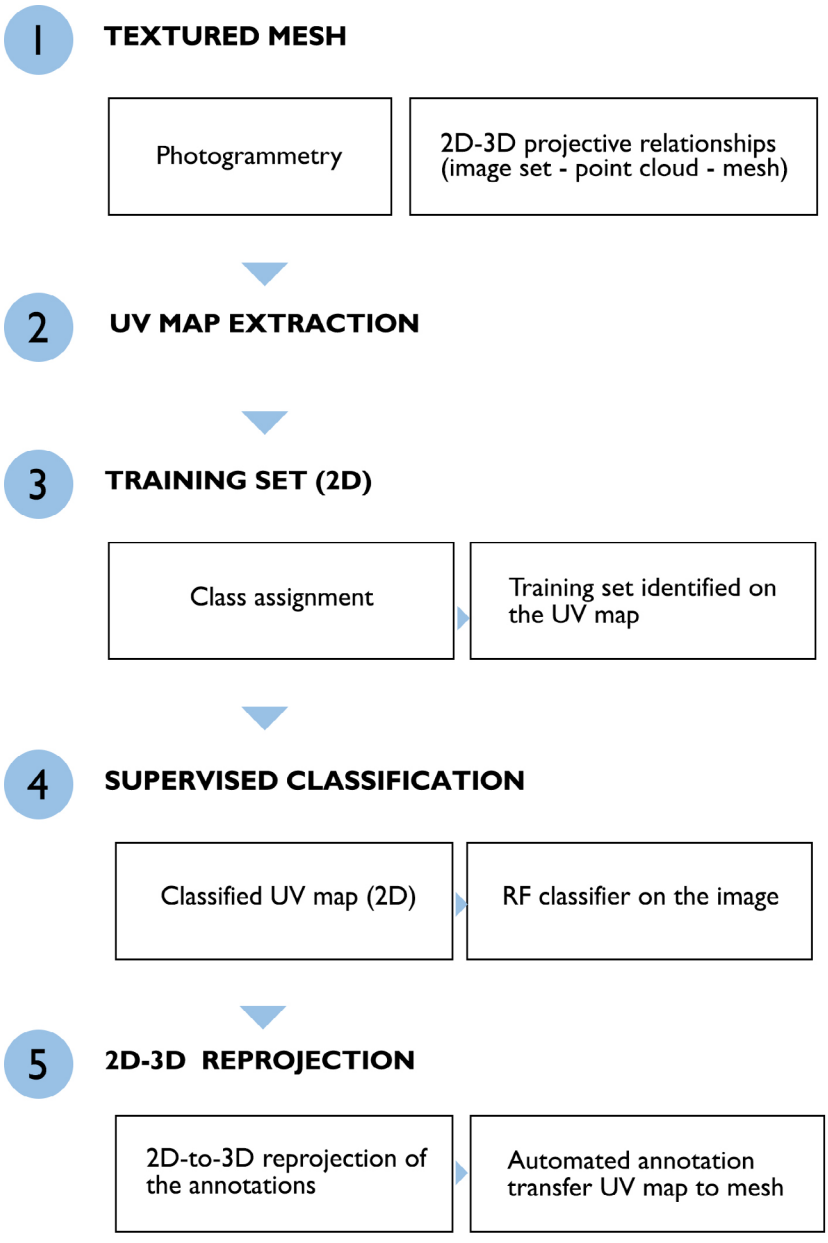


Figure 6-17.

From classified UV maps to mesh models. Overall workflow.

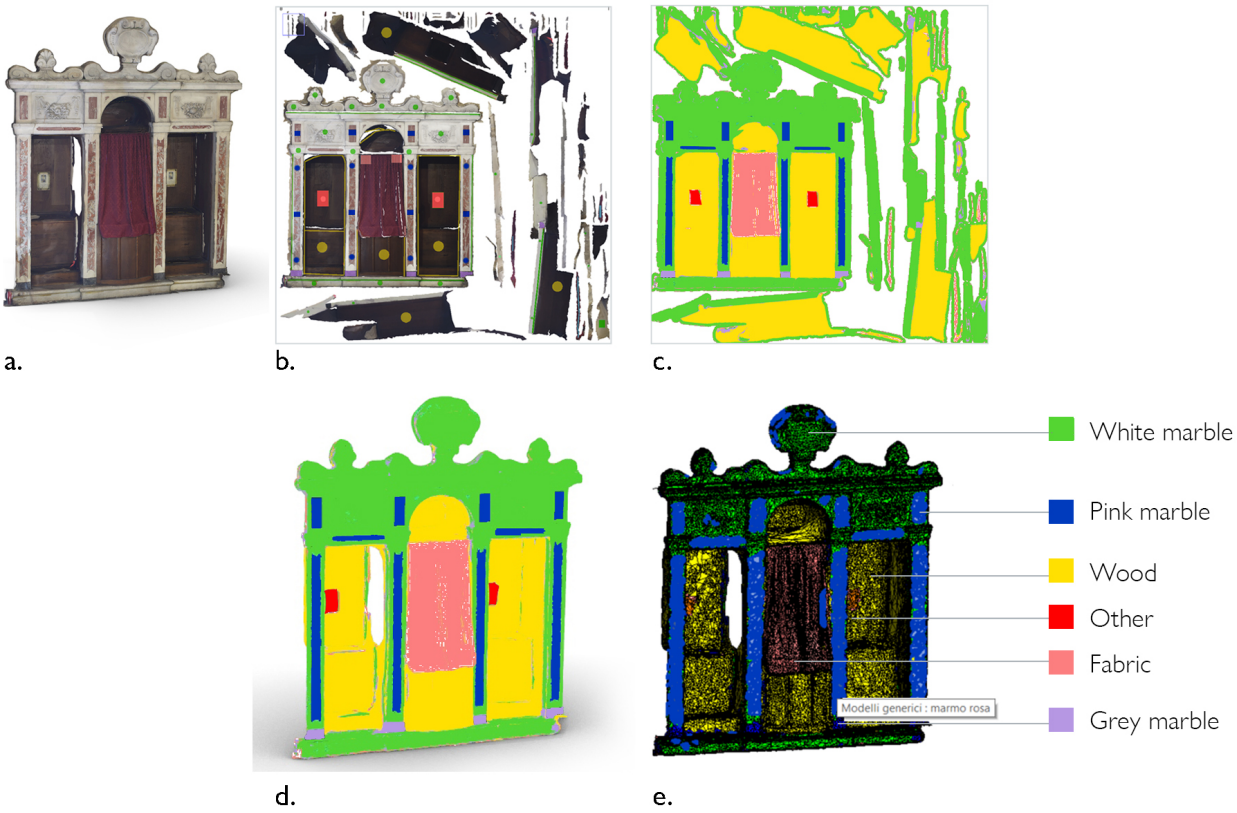


Figure 6-18.

UV map-mesh model transfer: photogrammetric model (*a*), training set (*b*), segmented UV map (*c*), segmented textured model via back projection (*d*), selection of the single mesh faces based on color values (*e*).

e.g.: Confessional, Chiesa del Carmine

As previously outlined in Paragraph 5.3.2.3, the UV map extracted from a mesh model can be used for 2D classification and subsequent projection of the semantic information onto the 3D mesh (Figure 6-17).

In the case of the dataset of the *Chiesa del Carmine*, this approach reveals fundamental in characterizing high-resolution photogrammetric models of the altars and confessionals, e.g., in terms of assessment of materials and degradation levels.

In this regard, it should be remarked that the transfer of semantic information, as well as the appropriate retrieval of the projective relation between UV map and mesh model, could also support the automatic segmentation – hence, the selection – of the single mesh faces according to the different colors found on the texture (Figure 6-18). In so doing, different portions of the mesh could be detected and segmented, each corresponding to one of the identified classes of materials or decay.

This operation, as will be shown in Section 7.5, is essential to guarantee the autonomous management of the parts (the identified classes) within BIM-based software as Autodesk Revit, to assist the construction of building information models.

6.2.2. From classified image sets to 3D point clouds

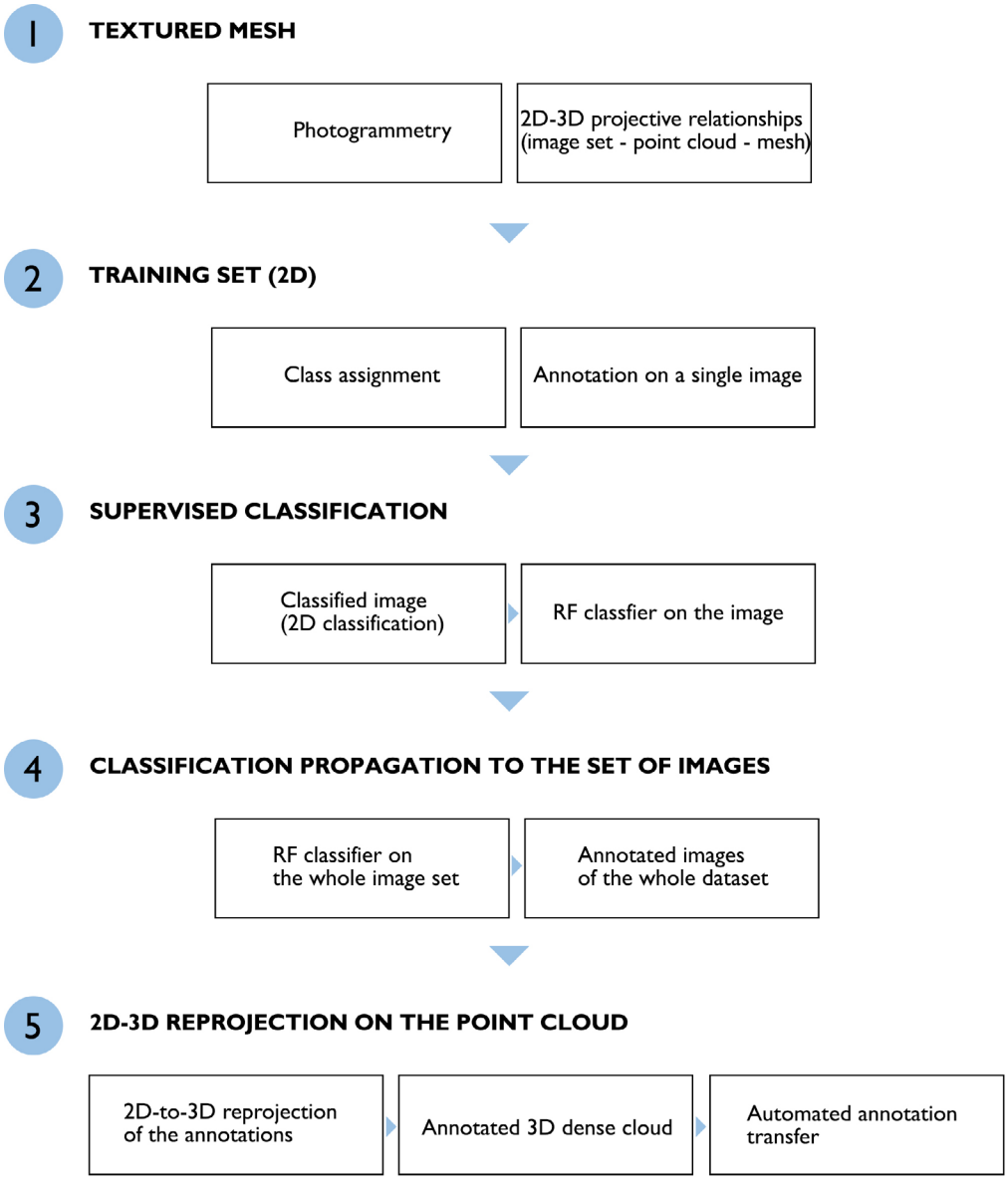


Figure 6-19. From classified image sets to 3D point cloud – Overall workflow.

In the case of photogrammetric models, the projective relationships between images and point clouds can be retrieved, with a similar approach to the one shown for Aïoli in Section 6.1, so as to transfer semantic data from 2D to 3D and vice-versa.

As for the semi-automatic classification approach implemented within Aïoli, the annotation of a single image is leveraged for the construction of a 3D training dataset, and the latter, in turn, allows the information to be propagated to the rest of the images.

However, it is worth noting the classification operation can also be performed with an inverse approach, i.e., by annotating the whole

image set by supervised ML first, and later transferring the information to the 3D model (Figure 6-19). In this case, the approach is more texture-oriented, since the automatic classification is formerly carried out on the images rather than on the point cloud, and it follows the steps of Figure 6-20. The example provided refers to the identification of material patterns and the detection of the different layers of construction of a track of the medieval urban walls of Pisa, starting from a set of images acquired on site.

Initially, the RF is trained on a single image; then, the classifier is applied to the rest of the dataset (previously unseen images), in batch processing (i.e., without further training). In this way, a semi-automatically annotated image set is first obtained; then, the classification information is transferred from the whole set of images to the 3D point cloud by back projection.

For this latter stage, in detail, the relationship between 3D coordinates and images is preserved, but the original images are replaced with the annotated ones in order to change the color information of the dense cloud (the so-called *colorized dense cloud process*), as well as in order to generate the texture of the mesh model. In other terms, the 3D coordinates of the points remain unchanged, while the color of the 3D model is modified according to the classification result.

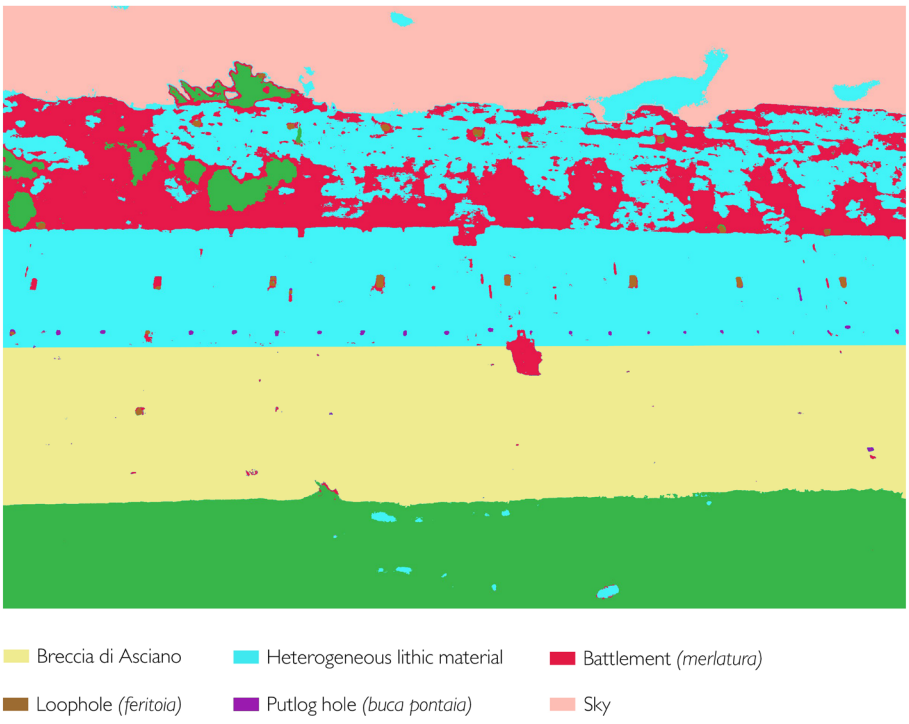
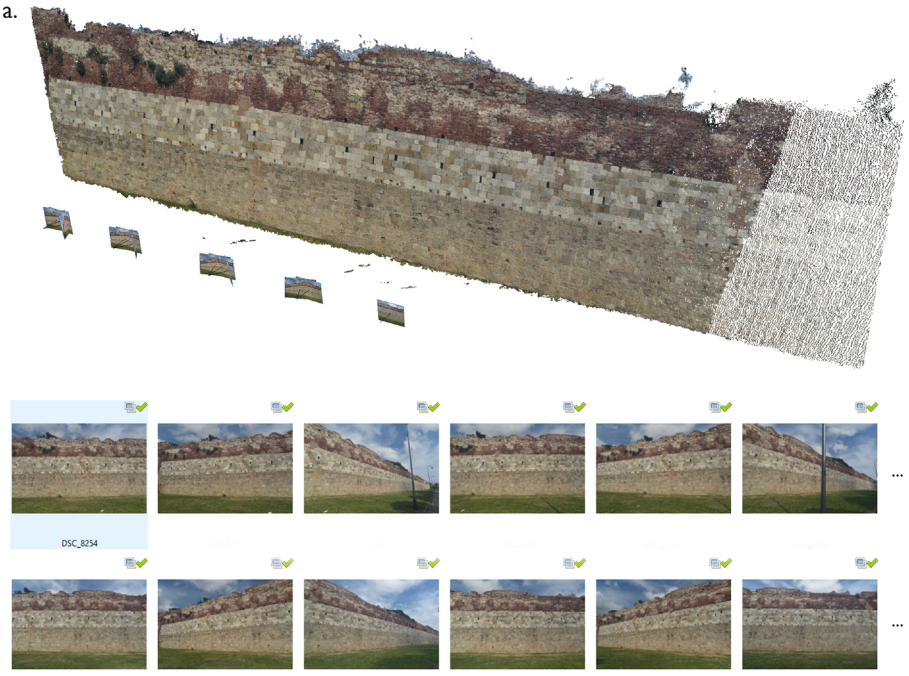


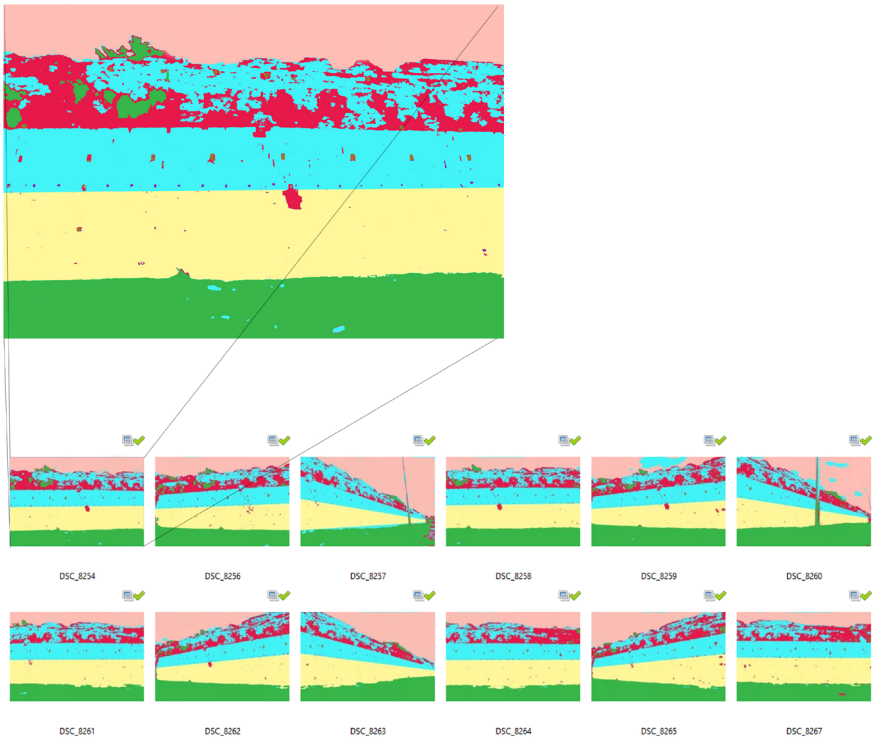
Figure 6-20.

Image sets-point cloud segmentation.
e.g.: Truck of the urban walls of Pisa in *Tersanaia*

a.


DSC_8254

b.


DSC_8254
DSC_8256
DSC_8257
DSC_8258
DSC_8259
DSC_8260
DSC_8261
DSC_8262
DSC_8263
DSC_8264
DSC_8265
DSC_8267

C.

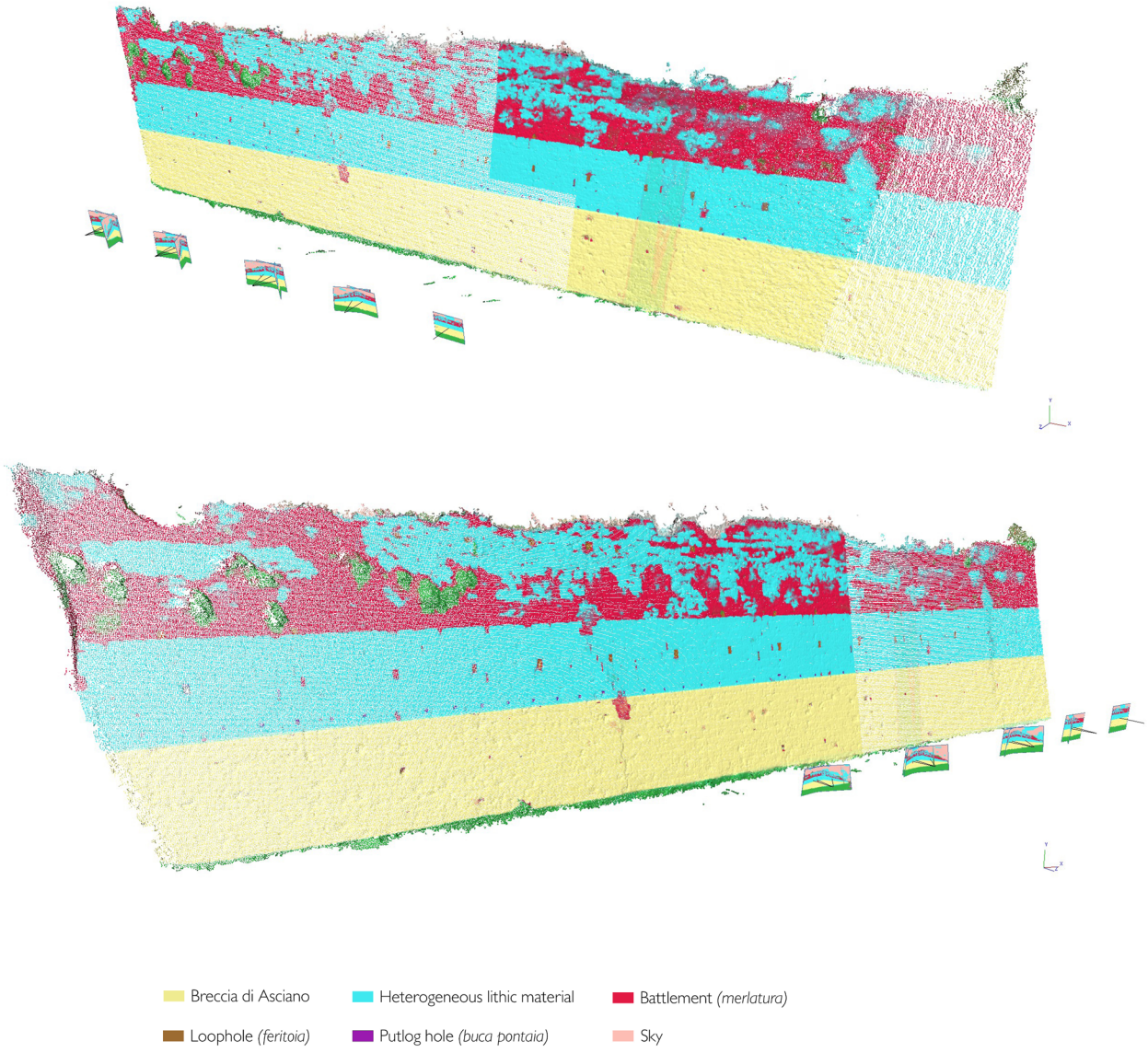


Figure 6-20.

Image sets-point cloud segmentation: 2D photogrammetric model *(a)*; 2D segmentation *(b)*; 3D segmentation *(c)*.

e.g.: Truck of the urban walls of Pisa in *Tersanaia*

6.2.3. Spatialization of archival images onto the 3D model

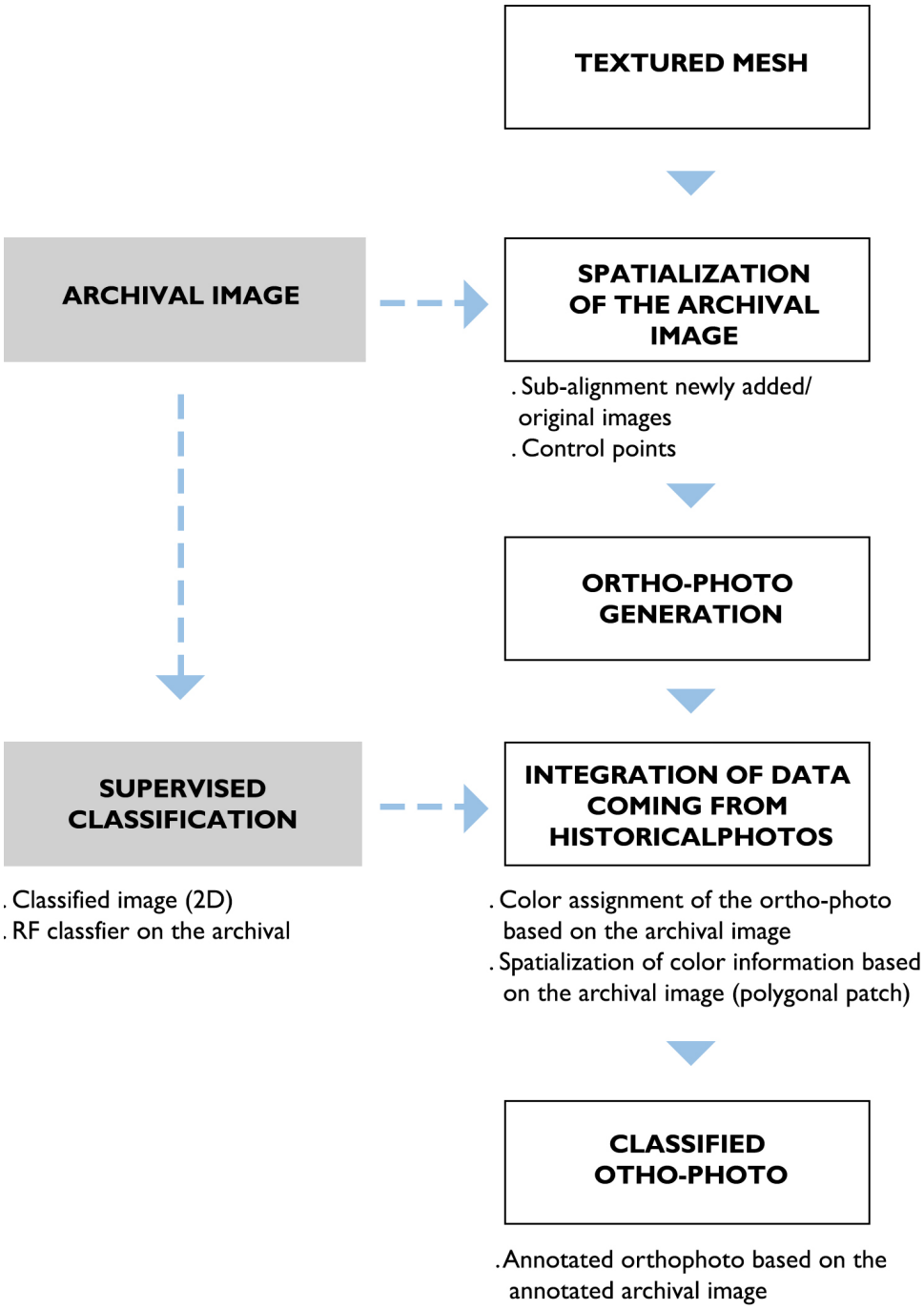


Figure 6-21.

Spatialization of archival image onto the 3D model – Overall workflow.

Another major feature of 2D/3D annotation transfer concerns the spatialization of new images on an existing photogrammetric model. Previous studies (Martinez Espejo Zaragoza et al., 2021; Bevilacqua et al., 2017) underlined the possibility to integrate the surveying models with data coming from historic and archival photographs, as well as with

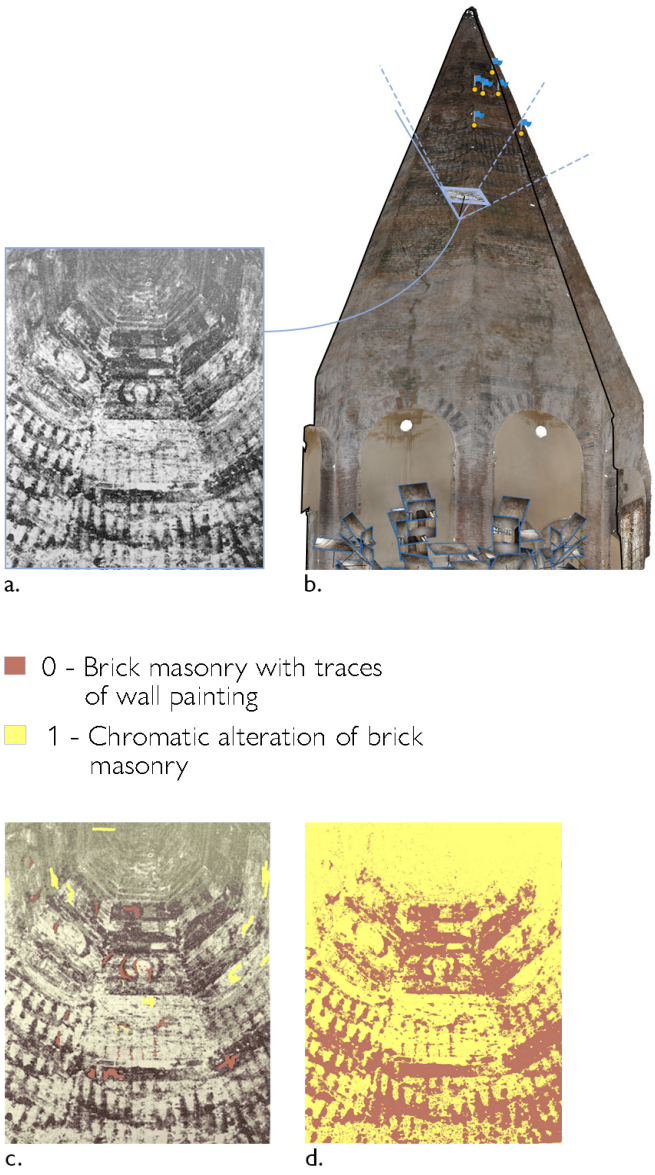


Figure 6-22.

Spatialization of archival images. The input archival image (*a*) is suitably inserted within the existing-actual-photogrammetric model (*b*). The archival image is then trained (*c*) and classified (*d*).

images of the same object acquired in different times and/or lighting conditions.

In this case, the extra images should be sub-aligned to the set of already aligned images, and the orientation parameters and shooting geometry of the cameras be derived by appropriate insertion of control points, both identified in the newly added image as well as in the original ones (Figure 6-21).

If the spatialization is feasible, the semantic information can be propagated from the newly added image to the rest of the image set, as well as to the 3D model.

For the example of the *Cappella di Sant'Agata* chapel in Pisa, already mentioned in Paragraph 5.3.2.1, the spatialization of archival images (Figure 6-22) provides a better comprehension on the remaining of the



Figure 6-23.

Spatialization of an archival image (Burresi and Caleca, 2003) over an existing photogrammetric model.

e.g.: Cappella di Sant'Agata, Pisa

wall paintings. In detail, the insertion of additional images (Burresi and Caleca, 2003), dating back to the 1960s (Figure 6-23), allows to characterize the original frescoes found on the inside cusp, as well as to compare them with the current situation of the walls, affected by advanced degradation patterns (Bevilacqua et al., 2017).

By evaluating the actual and previous states of the wall and by generating an ortho-photo of the 3D model in the two respective conditions, a classification approach is applied to compare the mural traces of the 1960s and of today, respectively.

The results are provided in Figure 6-24, and the segmentation of the archival image reveals the lost wall traces over the ortho-photo (Figure 6-24 c).

The integration between archival imagery, 3D surveying, and AI classification approaches allows to better describe and characterize the state of conservations of the building's interiors over time. The reality-based model of the Chapel is considered to compare the actual condition of the interior frescoes with the situation that existed at the time the frescoes were discovered, in the late 1960s. The temporal evolution of the frescoes is depicted by providing the evidence of previous wall traces, and by leveraging a semantic segmentation process.

Considering the importance of the integration of such techniques, future work on this topic goes in the direction of a deeper fusion between 3D models, images, temporal states and classification techniques. In detail, the author is currently working on the construction of a dedicated application in which the reality-based model derived from surveying and the classification information derived from the study of archival documentation can be directly accessed, integrated and visualized. In such immersive experience, the integration and dissemination of reality-based models and classification results may be achieved by considering advanced real-time visualization tools as Unreal Engine (Figure 6-25).

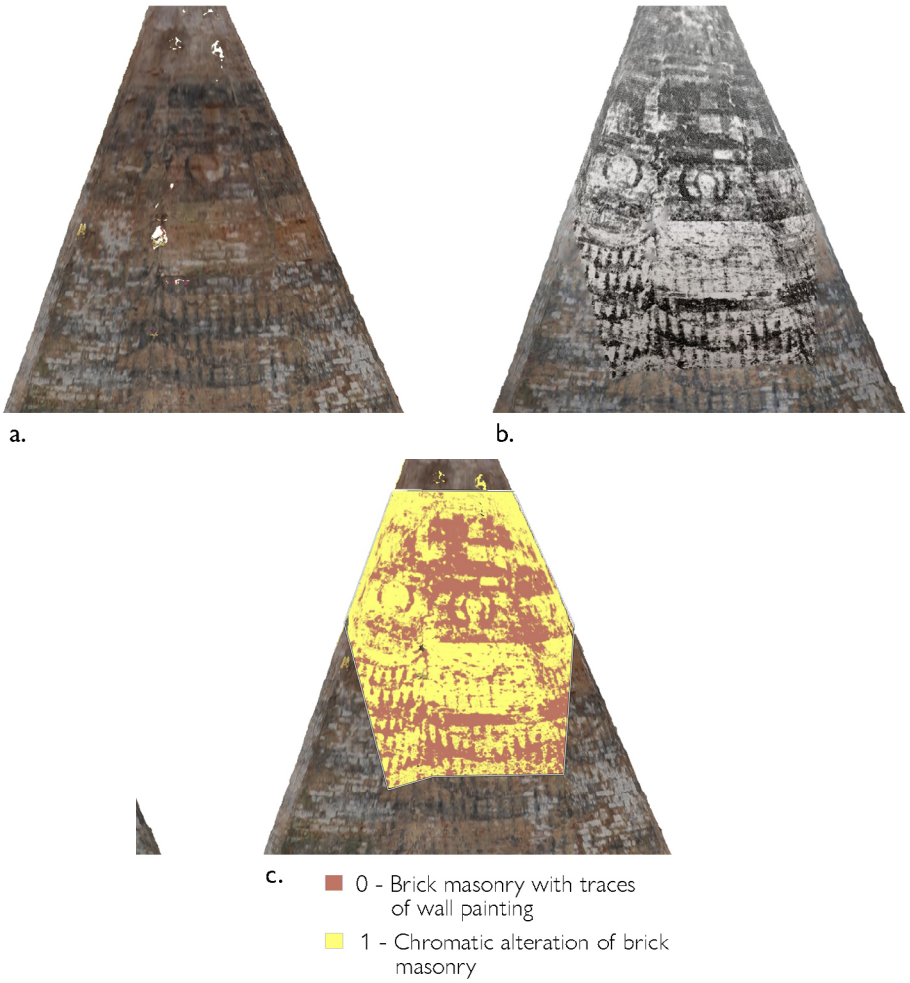


Figure 6-24 *(to the left).*

Portion of the ortho-photo of the interior walls of the chapel: actual condition (a); spatialization of the archival image (b); evidence of previous wall traces highlighted by segmentation and projection of the archival image over the ortho-mosaic (c).

e.g.: Cappella di Sant'Agata, Pisa



Figure 6-25 *(below).*

A digital rendering of the reality-based model of the Chapel via Unreal Engine, toward the integration and dissemination of the classification results within Virtual Reality applications.

6.3. Multi-level and multi resolution point cloud semantic segmentation

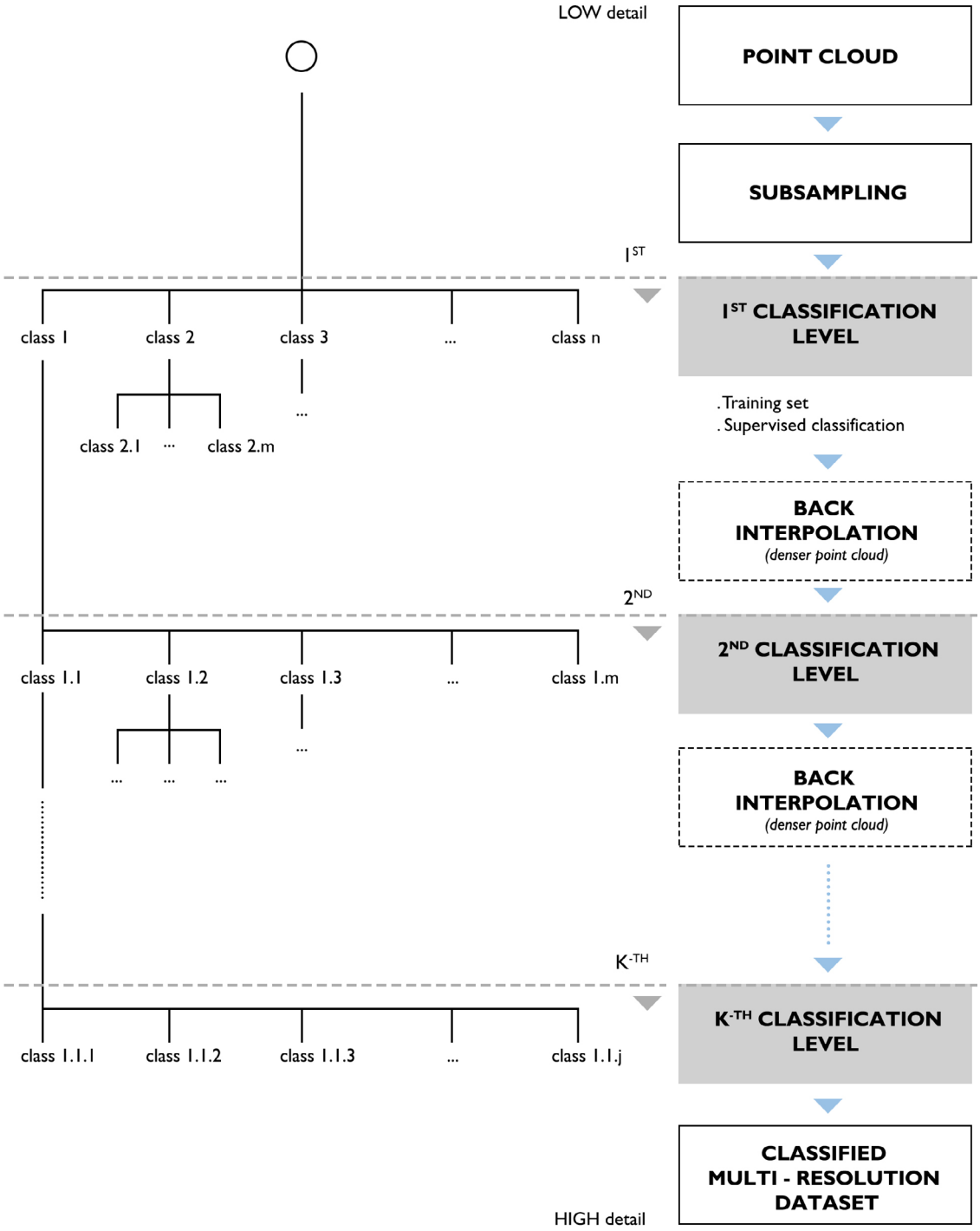


Figure 6-26.

Multi-level point cloud semantic segmentation scheme.

With a multi-layer and multi-resolution approach, subsequent levels of detail and representation can be considered on point clouds, to suitably transfer semantic information.

As observed in Paragraph 5.3.1.4, in the case of geometry-based classification approaches, the choice of the classes (in which the original dataset should be subdivided) is essential in determining the success of the classification results: indeed, as the number of classes increases, so do the similarities between one class and another, so enhancing the risk of misclassification. Furthermore, it has to be observed that, when training the RF classifier:

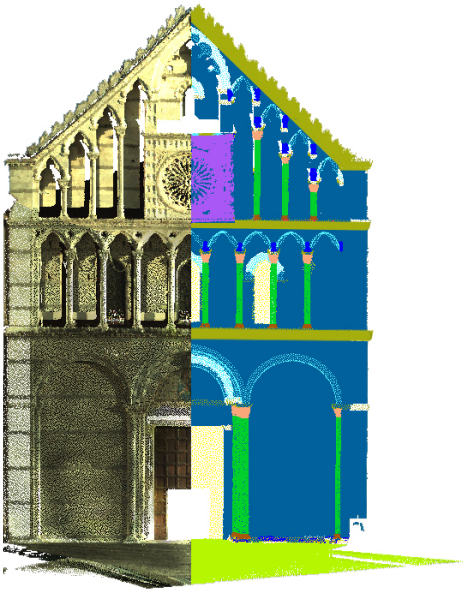


Figure 6-27.

3D laser scanner survey of the façade and related semantic segmentation.

e.g.: Chiesa di Santa Caterina, Pisa

- The time required for feature extraction increases as the density of the point cloud increases, requiring more and more performing computing machines;
- The elevated file size of the point clouds may entangle difficulties in the manual annotation of training datasets.

These issues can become critical in the validation of the proposed methodological approach, demanding significant amounts of time and resources.

Starting from these considerations, emphasis is placed here on the possible definition of multi-level and multi-resolution systems (Figure 6-26), based on the annotation of point clouds with different point densities, even addressed by the work of (Teruggi et al., 2020).

Initially, the original point cloud is subjected to successive subsampling operations (Section 3.2.2.2). In this way, point clouds with different densities are obtained, describing the considered data set in greater or lesser detail.

A first semantic segmentation is thus performed on the less dense point cloud, and the classes are identified with a lower level of detail. Then, a first RF classifier is trained, and a classification result is obtained. The latter is hence transferred to the point cloud with higher density by *back interpolation* operations (Section 6.1.4), and so on up to the most detailed segmentation level.

The following figures show the results of the application of such

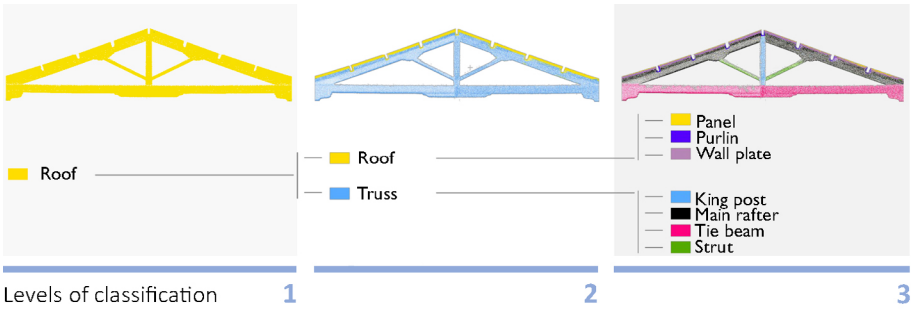
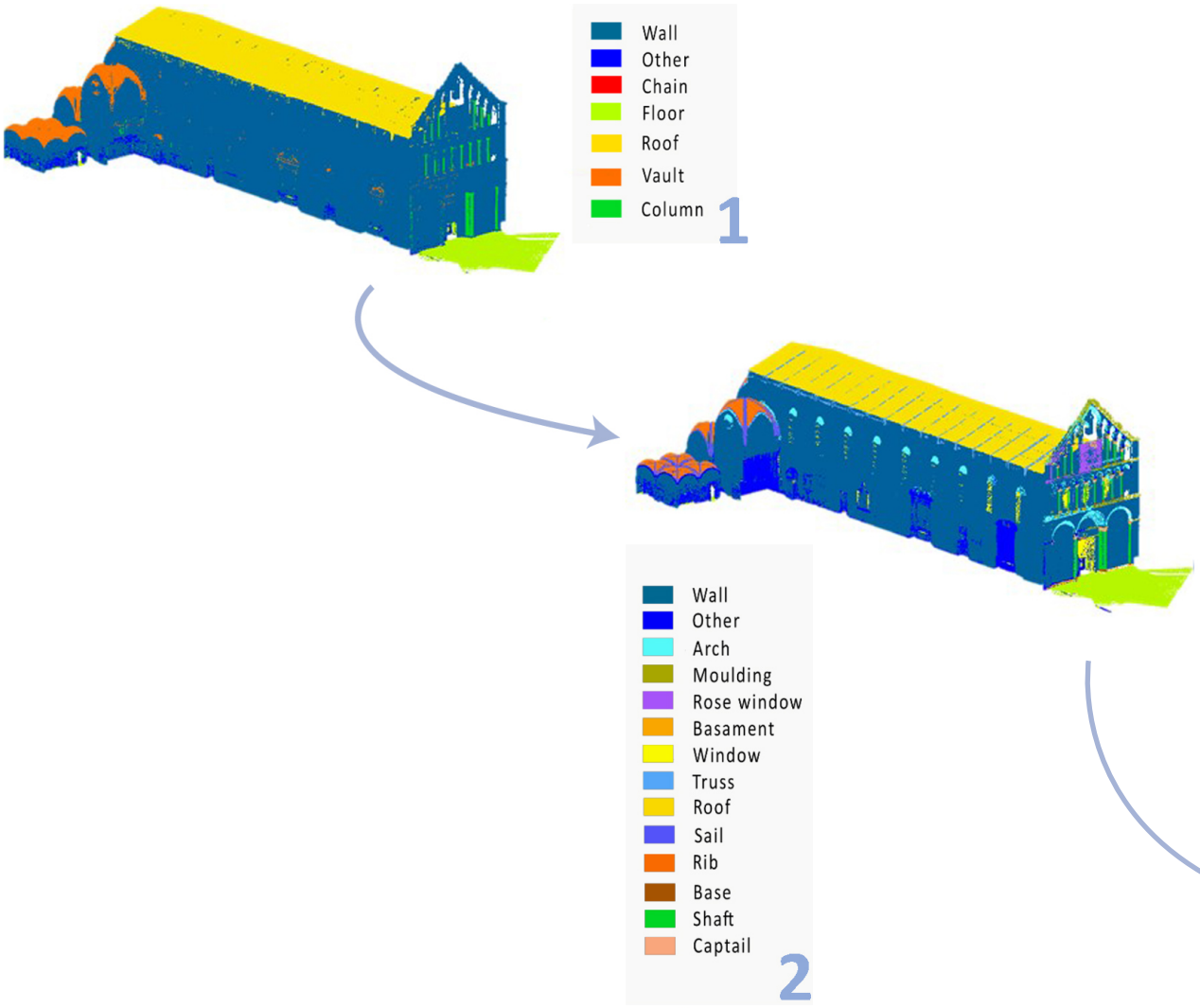


Figure 6-28.

Example of multi-class segmentation approach for the roofing system.

e.g.: Chiesa di Santa Caterina, Pisa

approach to the two case studies of the *Chiesa di Santa Caterina* church in Pisa (Figure 6-27) and of the framing structure (fr: *charpente*) of Notre-Dame Cathedral in Paris, respectively.



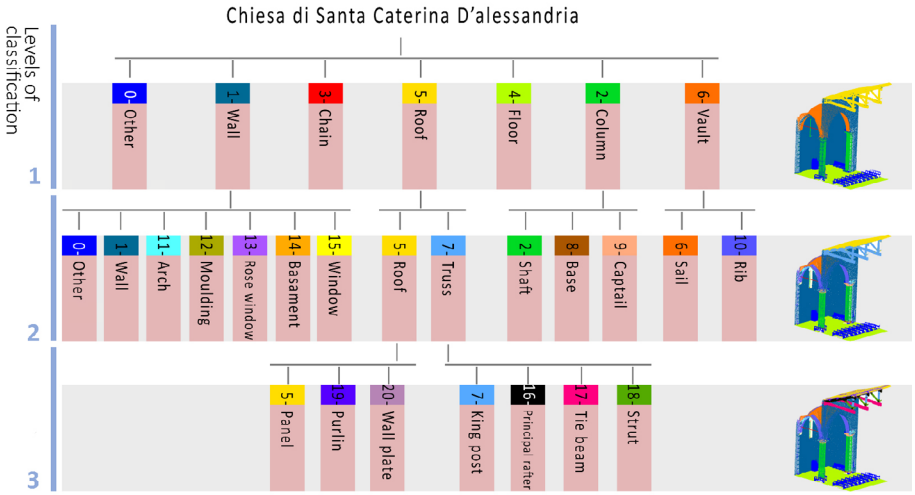


Figure 6-29.

Multi-level classes identified for the point cloud classification purpose

e.g.: Chiesa di Santa Caterina, Pisa

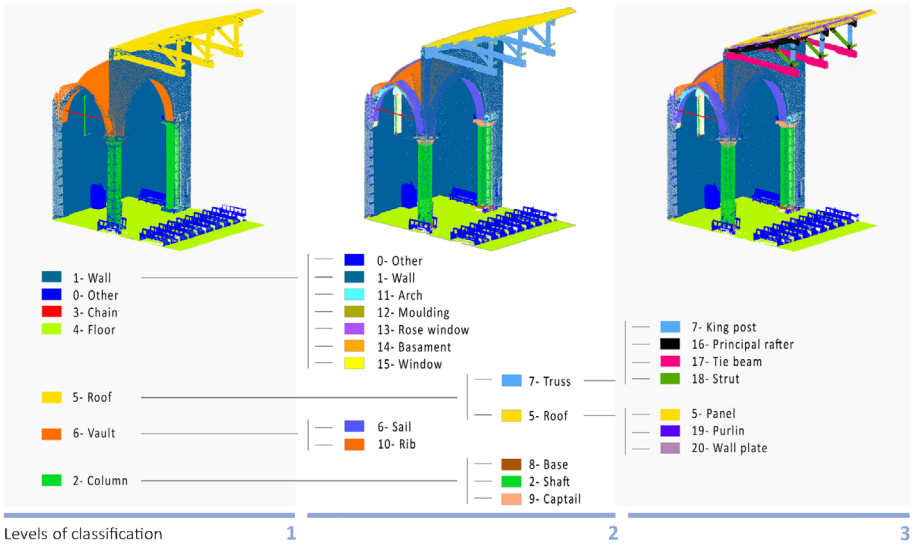


Figure 6-30.

Multi-level training set identified

e.g.: Chiesa di Santa Caterina, Pisa

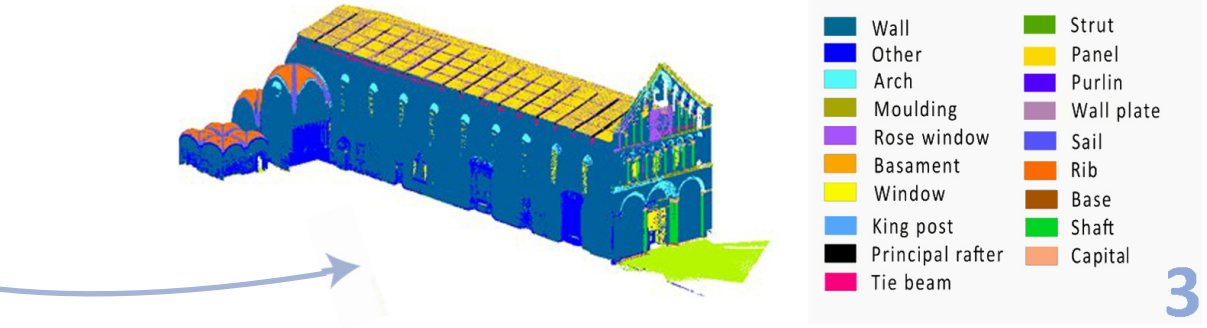


Figure 6-31.

Multi-level classification results

e.g.: Chiesa di Santa Caterina, Pisa

Figure 6-32.

Point cloud of the framing system before the fire (UMR MAP CNRS/MC).

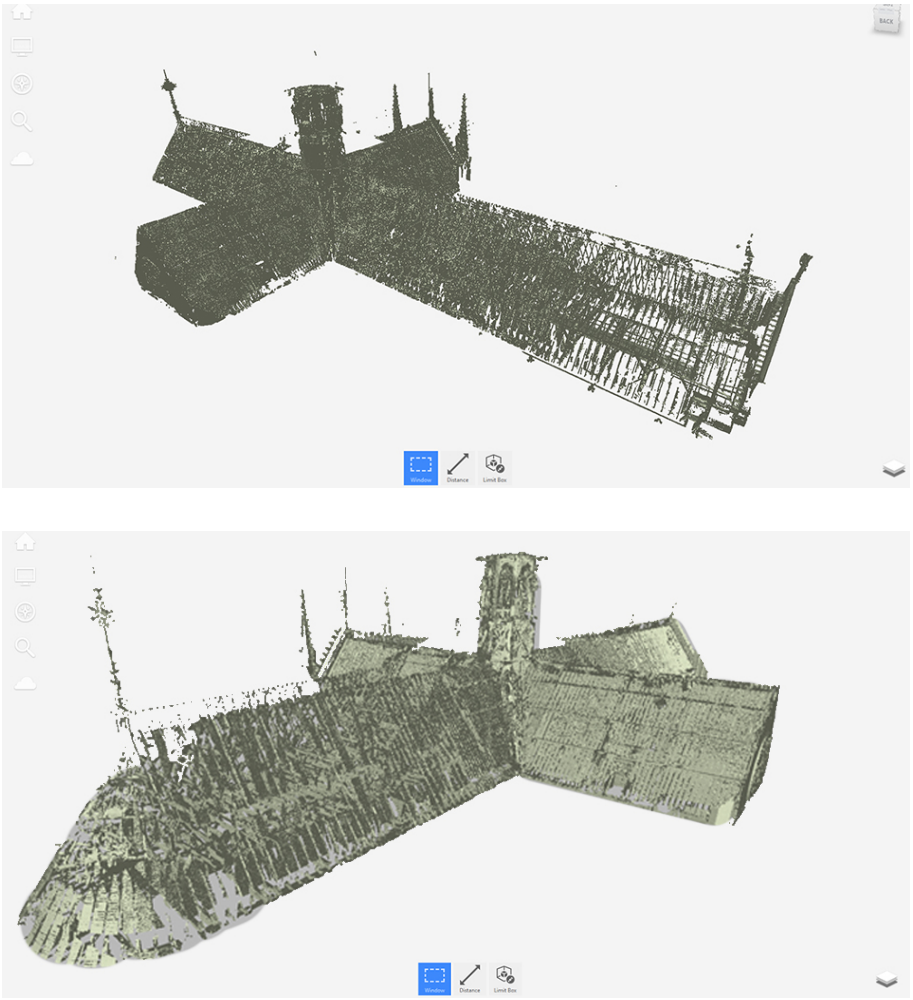
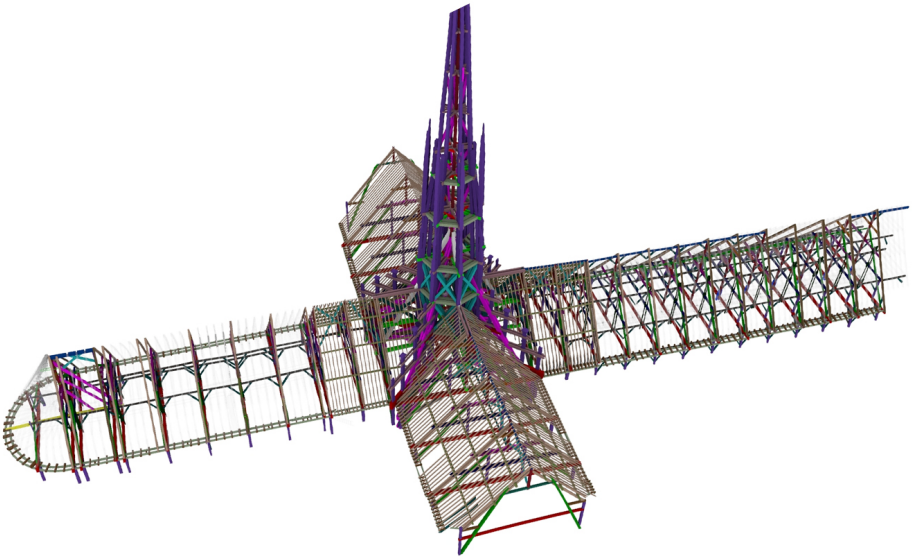


Figure 6-33. (*to the right*)

3D components of the framing system re-modelled in CAD.

Source: Kévin Jacquot, UMR MAP CNRS / MC

e.g.: Notre-Dame Cathedral, Paris



The classification levels identified for the *Chiesa di Santa Caterina* dataset, from 1 to 3, range from least to most detailed. For the example

of the 'Roof' class provided in Figure 6-28, the class is decomposed at the 2nd segmentation level into 'Roof' and 'Truss' classes, and the 'Truss' class is further segmented into its components 'Strut', 'Tie beam', 'Principal rafter' and 'King post' at the 3rd segmentation level.

A classification approach of this type extended to the whole dataset leads to the detection of the classes of Figure 6-29, with the related training set (Figure 6-30). The results of the classification are provided in Figure 6-31.

In the second relevant case of Notre-Dame, the considered dataset relates to a survey (Figure 6-32, Figure 3-33) carried out before the fire that hit the cathedral in April 2019 and is used as a reference for the reconstruction of the semantic 3D model of the roofing structure. In particular, relying on the study of such surveying as well as on the inspections carried out by Fromont and Trentesaux (2015), the MAP research unit reconstructed, via virtual 3D modelling, the different components of the roof and structured them according to a semantic organization (Figure 6-32, Figure 6-33).

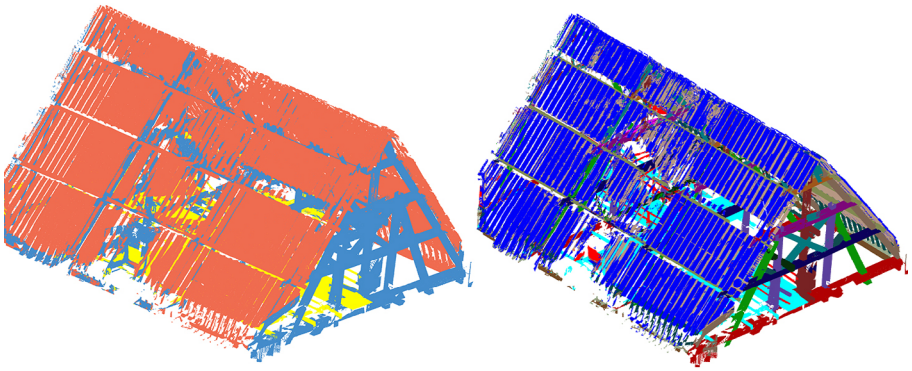


Figure 6-34.

Hierarchical classification approach identified for the segmentation of the roofing structure (fr: *charpente*).

e.g.: Notre-Dame Cathedral, Paris. Classification results for the southern transept

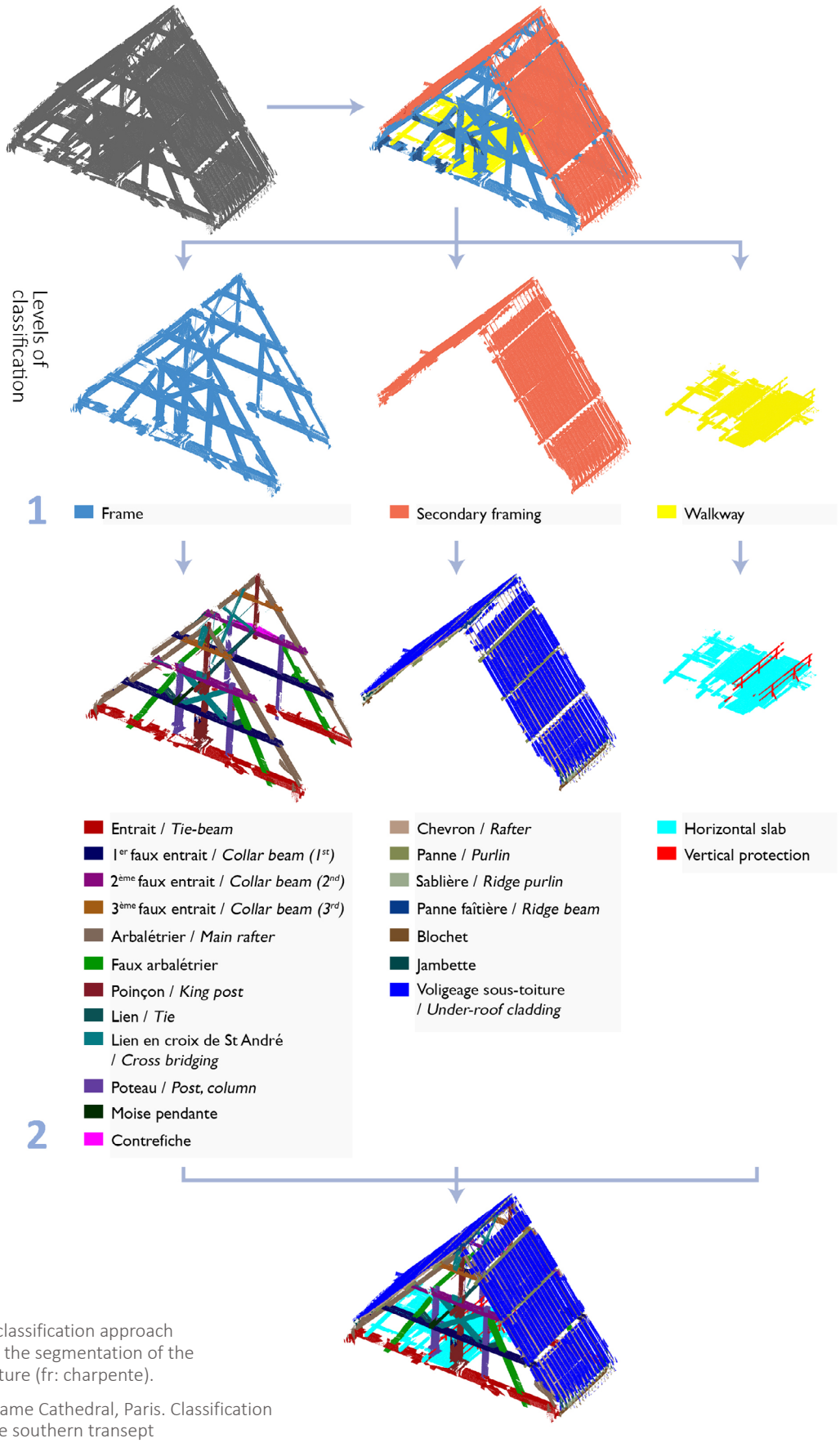


Figure 6-35.

Hierarchical classification approach identified for the segmentation of the roofing structure (fr: charpente).

e.g.: Notre-Dame Cathedral, Paris. Classification results for the southern transept

Once again, however, in the transition to the reconstructed 3D data, the reference to the original point cloud is lost. In order to fill this gap, the multi-level segmentation approach is proposed so to restore the link between the reconstructed model and the initial point cloud, thus detecting for each 3D component the points of the cloud to which it corresponds. Given the complexity of the dataset, a first manual segmentation is considered so to divide the framing structure into choir, northern transept, southern transept, spire and nave.

Then, for each group, the segmentation levels are identified. The training set and the results of the classification are presented in Figure 6-34 and Figure 6-35, respectively, for the example of the south transept.

6.4 Conclusions

In this chapter, further possibilities for semantic enrichment are presented to enable the transfer of information between 3D models (mesh or point clouds) and 2D models (images, UV maps or ortho-photos). The proposed approaches further contribute to a better understanding of complex heritage architectures, as well as to collaborative data sharing and information management.

As regards the implementation of the Aïoli platform (Section 6.1), we demonstrated how, starting from photogrammetric models, the hybrid annotation system can be combined with supervised classification methods based on the RF. Besides allowing the propagation of semantic data between images and point clouds and vice versa, the proposed approach offers novel semi-automatic modalities of data annotation.

In Section 6.2, further approaches for the annotation propagation from point clouds or meshes (3D) to images, UV maps or ortho-photos (2D) are suggested. Once again, the combination of these approaches and ML-driven classification offers promising developments in terms of digital data collection, collaborative work, traceability and interoperability for heritage documentation and analysis purposes. Finally, the hierarchical classification approach proposed in Section 6.3 further allows the construction of point clouds at different densities, that could result beneficial also in view of the generation of H-BIM models, of a same heritage object, at different levels of detail.

The need for more supports and the availability of multiple levels of resolution and detail are no longer to be seen as a constraint, but rather as a potential: starting from the existence of several formats, in fact, the operator may choose, according to his specific requests and needs, which is the easiest and fastest support to annotate, so to transfer,

afterwards, the information from 2D to 3D or vice versa.

However, it has to be noted that the cases presented in this chapter mostly rely on the use of digital models built from photogrammetric surveys, in which the relationship between image / UV map / ortho-photo and 3D model is known a priori as an effect of the construction of the 3D model itself. Indeed, the image orientation and calibration, are relevant phases of the photogrammetric process. The alignment between images and point cloud is thus totally consistent (Manuel et al., 2016).

When dealing, e.g., with the spatialization of images over a point cloud or a mesh obtained by laser scanner, the annotation transfer, although always possible, becomes less automatic and direct. In fact, the 2D-3D projective relations are not so trivial and they should be derived, e.g., by the study of common points between image and model, as in the case of the spatialization of archive images (Bevilacqua et al., 2017). Otherwise, the use using panoramic images from a laser scanner surveying may be considered for the annotation transfer purpose.

Other alignment scenarios, more focused on laser scanning surveys, could thus be further considered to test the effectiveness of the presented strategies.

As a final remark to this chapter, it should be noted that, for the supervised classification task, the training set has been performed, in the cases presented, either starting from the annotation of 2D supports or starting from the annotation of 3D media. The hybrid 2D/3D mode, in other words, has not yet been considered for the manual annotation of training data.

As a matter of fact, some annotations may be difficult to perform on images or, vice versa, on 3D models: to give an example, in the case of image-based annotations on Aïoli, the selection of objects with cylindrical development (e.g., columns) is tricky on photos. Conversely, the annotation of a particular form of degradation of a flat wall, for instance, can be challenging on the point cloud.

To overcome this issue, while still considering the possibilities of integration between hybrid annotation modes and classification approaches, a multi-modal 2D/3D annotation system could be developed in order to take full advantage of each annotation medium. In this way, the annotations could be made both in 2D and 3D, according to the needs defined by the operator, so realizing a training set of semantically-annotated hybrid data that could be used for further classification and information propagation.

Chapter 7

From the Semantic point cloud to Heritage-Building Information Modeling

In the present Chapter, a semi-automated procedure is proposed to enable the reconstruction of Heritage-Building Information Models starting from annotated 3D survey data. In cases where the classification process relies on the identification of recurring architectural elements, in fact, the semantically segmented point cloud, obtained as a result of the application of AI, can be considered in view of the reconstruction of HBIM models (Figure 7-1).

As highlighted in Chapter 5, the introduction of AI-based techniques to the heritage domain constitutes a significant advance in the interpretation of raw survey outputs, improving the automatic recognition of materials, decay patterns or architectural components within initial 2D/3D digital survey data. Geometry-based approaches (Section 5.3.1), in particular,

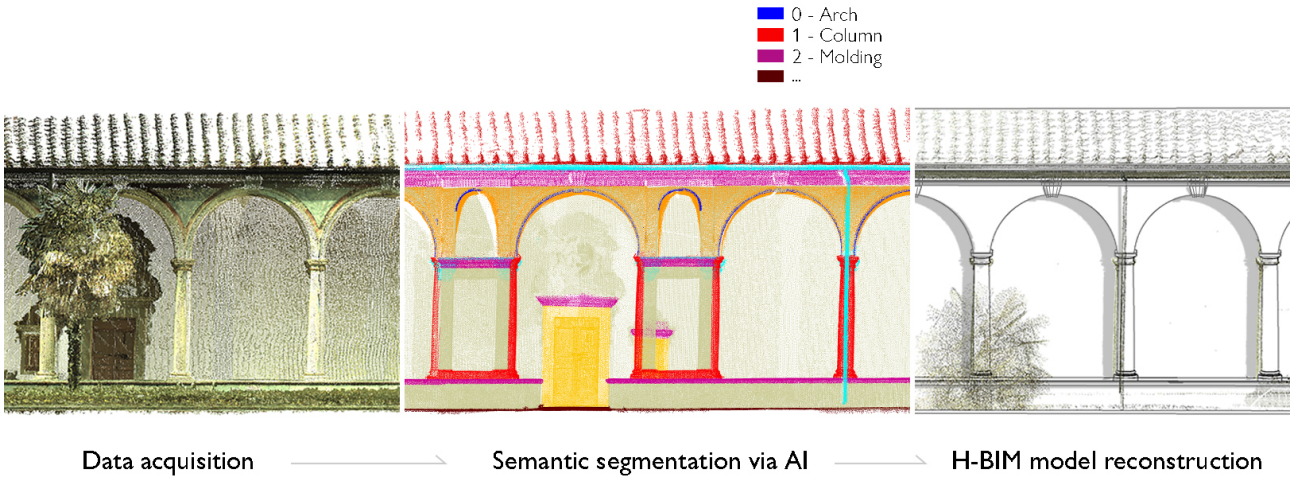


Figure 7-1.

The semantic point cloud with identified typologies of recurring architectural elements is taken as input for the construction of H-BIM representations.

e.g.: Grand cloister, Pisa Charterhouse

support the detection of typological elements as walls, columns, floors, vaults etc.

Considering that such a semantic structuring of heritage data could be beneficial in view of the construction of BIM-based representations, the classified point cloud (resulting from AI-based classification techniques with recognition of the main architectural components) is taken as initial datum. The hierarchical organization of the classes over the annotated point cloud, in fact, reflects the logic of H-BIM environments, where each element is defined by specific names, attributes and properties.

The recourse to architectural treatises, as well as the use of visual programming languages, allow deriving and propagating 3D shapes belonging to a same typological class. These operations become essential in view of the realization of a complete information system of the H-BIM type where to store knowledge-related data.

In detail, the reconstruction process proposed relies on the following steps:

- Import of the annotated point cloud in 3D modeling environment, broken down class by class;
- Reconstruction of a template geometry for each class of architectural elements identified, also referring to architectural treatises;
- Propagation of the template geometry to all elements belonging to a typological class.

The classes (families of parametric elements, e.g., columns, moldings, roofs, walls, vaults) are reconstructed, also based on architectural canons, through visual programming languages (VPL).

The model obtained at the end of the process, after 3D reconstruction of all classes of typological elements, can be used to build H-BIM type representations, i.e., to construct 3D repositories of the architectural heritage, that can be further enriched with information related to conservation and documentation.

The Chapter is structured as follows: in Section 7.1, an in-depth illustration of the proposed approach is provided, while in Sections 7.2, 7.2, 7.3 and 7.4 the results obtained by application of the methodology to some relevant case studies are presented. The assessment of the method and the concluding remarks are finally discussed is Section 7.6.

7.1. Methods

7.1.1. A semi-automatic Scan-to-BIM workflow for the reconstruction of *template* models

The supervised ML procedure, followed for the identification of typological and recurring components on raw 3D data of our architectural heritage, was previously illustrated in Section 5.3.1 and led to the generation of a semantic point cloud: therein, the different architectural elements composing the surveyed object are distinguished and classified in accordance with the classes recognized by (Matrone et al., 2020) for the 3DArCH dataset.

Starting from the semantically annotated data so achieved, a conceptual model can be reconstructed, by treating and processing separately each class of architectural components, so that each one results independent of the other. This way of classifying data appears in line with the rationale of the H-BIM process, whereby the model is generated through smart objects, appropriately distinguished in terms of typology and morphology (e.g., roof, wall, floor, column etc.) and regrouped into *families* of architectural elements. In the overall model, each component retains, indeed, its own identity in relation to the other.

Based on these considerations, for each class detected by AI-based semantic segmentation on the segmented point cloud, a template geometry is reconstructed. This latter constitutes the basis for the reconstruction of H-BIM *families*, each corresponding to a class and regrouping typological elements with similar geometries.

For the reconstruction of *template* shapes, in detail, reference geometries and proportions are derived, where available, from treatises of historical architecture, as detailed in Section 7.1.2.

Ideal shapes are so generated for each class through recognition and parametric reconstruction of elementary parts, profiles and surfaces that compose them. The procedure is broken down class by class and it is accomplished by exploiting the visual programming language, respectively following these steps (Figure 7-2):

- Import of the segmented point cloud and extraction of the single class concerned by the reconstruction process (Section 7.2);
- Structuring of the *template* element to be reconstructed by

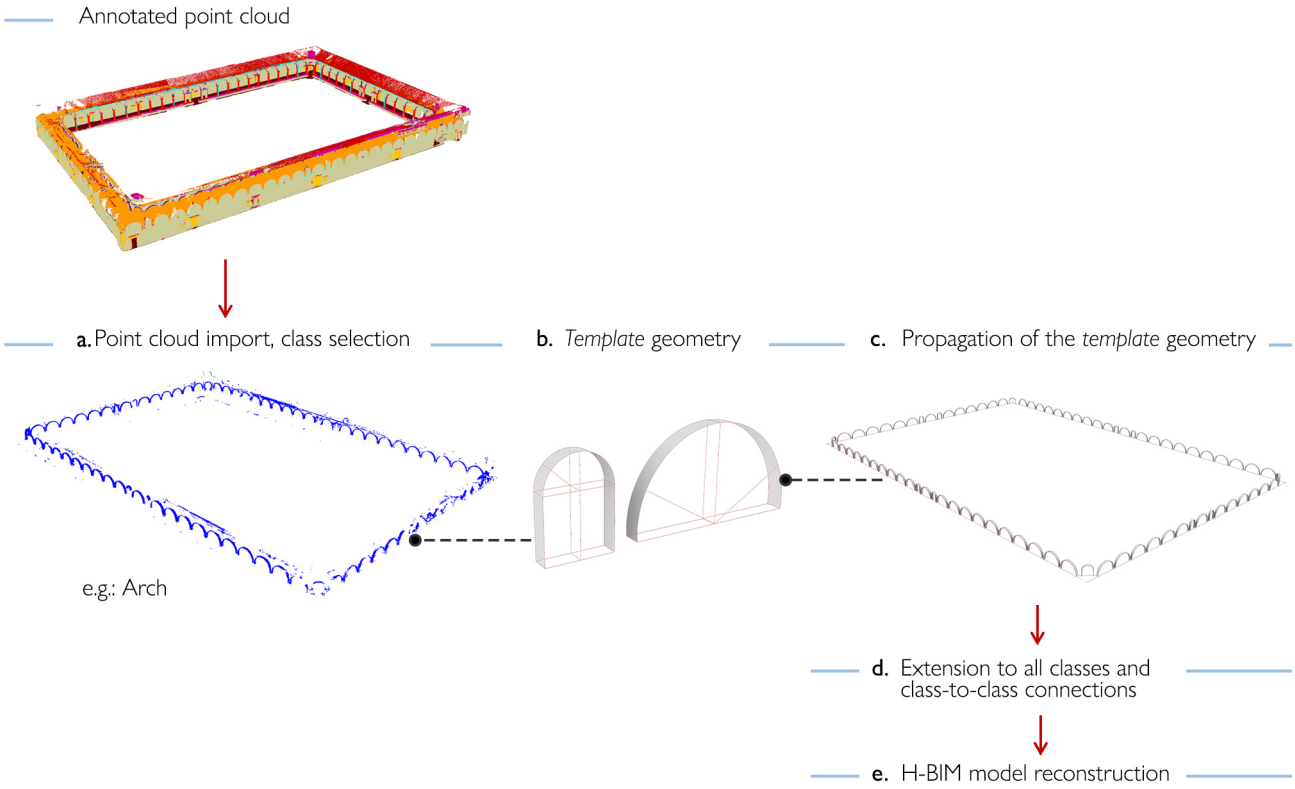


Figure 7-2.

Scan-to-BIM workflow starting from the input semantic point cloud

e.g.: Grand cloister, Pisa Charterhouse. *Template* reconstruction of Class 0 – 'Arch'

definition of base construction plans, constraints, generating primitives, base profiles and ensuing functions of extrusion, loft, sweep etc. (Section 7.3);

- Definition of element replica operations, so to enable the propagation and duplication of the defined geometry to multiple elements presenting same characteristics (Section 7.4).

The mathematical and conceptual representation of each class is managed through generative modeling procedures, relying on the creation of Non-Uniform Rational B-Splines (NURBS).

In detail, we leverage on the graphical algorithm editor Grasshopper, integrated with Rhino, for the generation, real-time modification, and graphical control of the architectural forms. Grasshopper, in particular, enables the visual manipulation of the 3D geometries' reconstruction procedures, by the creation of nodes (*algorithms*) and by their connection via wires. Finally, the plug-ins Volvox (food4rhino.com/en/app/volvox) and *Rhino. Inside. Revit* (rhino3d.com/inside/revit/1.0/) are leveraged to connect the reconstructed 3D model with the point cloud processing software and with the BIM platforms, respectively.

Concerning these latter aspects, the previously-mentioned concepts of VPL and generative modeling (Paragraph 3.3.5.3) are further clarified in Section 7.1.3.

7.1.2. Reference to architectural treatises

The reconstruction of some relevant architectural typologies, such as the parts of the column or the moldings, follows the reference to architectural treatises.

The major treatise writers, starting from the work by Vitruvius and his treatise *De Architectura*, conceived indeed their masterworks so to organize first, in an encyclopedic way, the morphological complexity of historical buildings. They set up a series of proportional rules and canons of geometric construction considered to be the basis of the construction of existing architectures.

In detail, their work has always been focused on:

- The decomposition of the building organism according to the recognition of recurring architectural elements.
- The definition of a vocabulary of forms, enabling the semantic organization, even at a lexical (textual) level, of an architectural object.
- The search for the rules and constructive proportions that inspired the design of a built heritage artefact.

As such, architectural treatises allow an effective differentiation and representation of typological architectural components, while identifying for each element of an historic building its intrinsic *shape grammar.*

The idea of decomposing a built object according to the identification of recurring architectural elements, therefore, has a long history that can be inferred to the work of the major treatise-writers, and can be evoked when dealing with the semantic reconstruction of digital 3D models. The process of semantic segmentation is in fact the very one that led Vitruvius and other essayists to investigate, and pass on, the typical proportions of architectural orders, thus codifying languages and models of classical architecture.

Comparing the theoretical reflections posed by these masterworks with the observation and organization of the real and surveyed data, our aim is to digitally represent each component of the building according to its conceptual form, which hypothetically is the one at the base of the descriptive and representative canons that led to its realization.

The semantic structuring of concepts is so resumed and re-organized around a geometric digital representation, through the construction of libraries of *template* shapes dedicated to buildings surveying.

The purpose of this part of the work is therefore to reproduce the constructive and proportional process that guided the original

conception of a building, in order to derive an almost conceptual representation of the object of study. This *quasi-ideal* representation of the building should be based on the definition of template geometries to which, from time to time, semantic data is attributed and to which information related to knowledge and analysis is associated.

It is therefore intended to manipulate in this way a quasi-conceptual representation of the building, which is no more than the transfer of a semantic segmentation process.

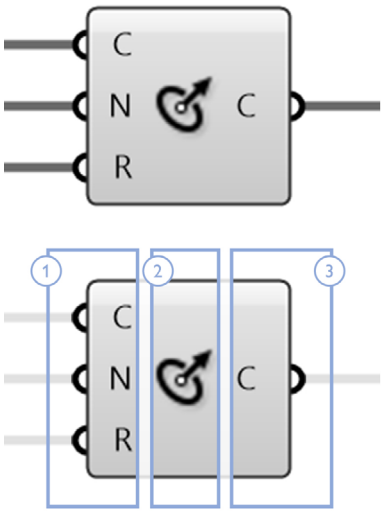
7.1.3. Generative modeling and Visual ProgrammingLanguage

For the reconstruction and propagation of the *template* geometries, VPL and generative design environments are considered. These systems, in fact, allow to conceive, edit and manipulate 3D parametric objects algorithmically and in a logical way, working in the programming environment while graphically observing the constructed or edited geometries onto a 3D viewer.

In the example of generative modeling considered, Grasshopper constitutes the programming environment, and Rhino 3D is the graphical visualization interface.

Figure 7-3.

Structure of a node in VPL. Input port with related options (1); body (2); output port with related options (3).



In the visual programming environment, each algorithm (in our case, whether it be implemented for the construction of a template geometry or for its propagation as well), is defined by a series of *nodes*, that are connected to each other through *wires*. Each node identifies a single operation, and it is composed of three parts (Figure 7-3): an input port with related options (1), a body with related name (2) and an output port with related options as well (3).

The wires link together multiple nodes to define more complex operations (hence, algorithms): data comes into the node through the input port on the left and sorts from the node through the output port on the right, so defining a step step-by-step action that follows a

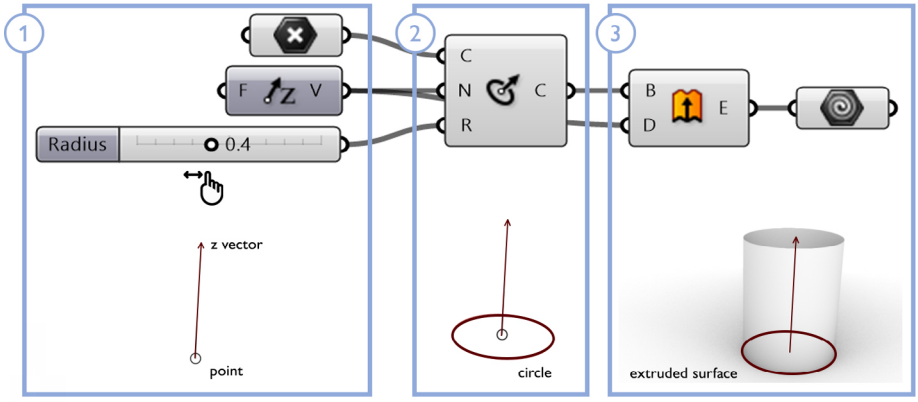


Figure 7-4.

Example of application of VPL for the creation of a cylindrical surface.

Center point, normal vector and radius (1) define the base circle (2), that is later subject to extrusion (3). The radius is set by use of a slider.

The changes are displayed in real-time in Rhino3D viewer (*second row in the image*).

basic flow of input, processing and output. A modification of the visual scripting language in Grasshopper changes in real-time the Rhino3D modeling environment (Figure 7-4).

By connection of multiple nodes, a repeatable set of instructions can be generated, so defining the relationships and the sequences of actions that compose custom algorithms. Finally, as in common programming languages, data can be organized in lists to perform the same action simultaneously over multiple elements.

The principles of generative modeling, based on the definition of a series of operations and rules allowing the generation of a certain graphical element, recall and evoke the architectural canons, and the series of constructive and proportional rules provided in architectural treatises for the study of ancient buildings. This correspondence was previously addressed in the reconstruction processes proposed by Bianconi et al. (2018), Apollonio et al. (2012) and De Luca et al. (2007), that exploited the application of VPL and generative modeling for the digital reconstruction of classical architecture. As stated by Bianconi et al. (2018), in fact, although being conceived for new and contemporary buildings, the generative modeling approach is actually close to the study of Cultural Heritage.

On this premise, we consider here the possibility of integration of reconstruction modalities based on generative modeling and semantically annotated point clouds obtained via AI. A series of visual algorithms is implemented in order to enable the visualization and management of the semantic point cloud, as well as to ensure the suitable creation, adjustment and propagation of *template* geometries.

Specifically, the construction of *template* geometries takes place via the development of custom algorithms based on the geometric modeling of curves and basic profiles and on their subsequent manipulation through functions of extrusion, loft, sweep, revolution etc. For each class identified over the point cloud, the reconstruction is performed by interpreting the idea of design which is upstream of its realization, starting from both the treatises and the 3D annotated data.

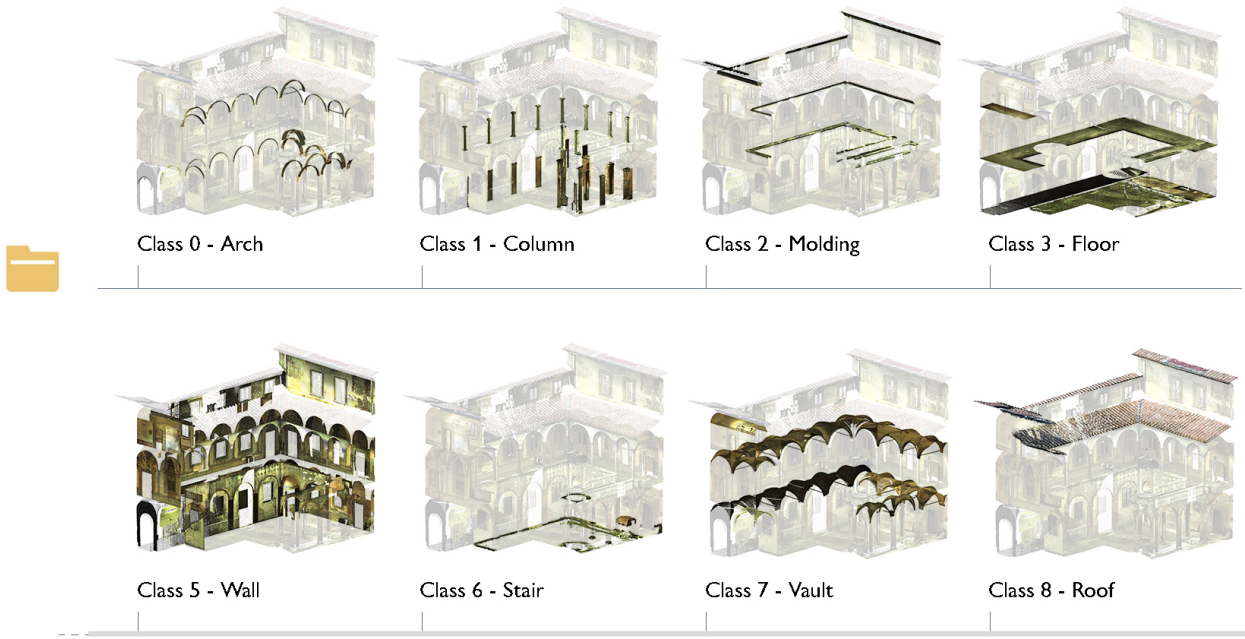
In the case of the propagation operation, instead, reference is made to nodes allowing array, copy, translation operations (e.g., the *'Move' node*), while a list of items can be considered to perform translation and replication operations involving multiple elements concurrently.

Finally, note that Rhino and Grasshopper are often referred to in the following paragraphs for the implementation of these algorithms. However, representation by visual programming has an incredible inherent advantage: whatever the real model of the building, and beyond the modeling software used, the method can be rewritten in various environments, as the defined construction rules allow semantic attributes to be reproduced elsewhere, even regardless of the real or survey data of the building.

7.2. Results: Visualization and import of the annotated classes

Figure 7-5.
Isolated classes of architectural components from the initial point cloud
e.g.: Grand-Ducal cloister, Pisa Charterhouse

The results are presented in this Section with reference to the annotated point clouds of the Grand-Ducal cloister and the main cloister of Pisa Charterhouse and of the cloister of the National Museum in San Matteo, Pisa. The semantically segmented point clouds for these datasets are obtained after application of AI-based geometric



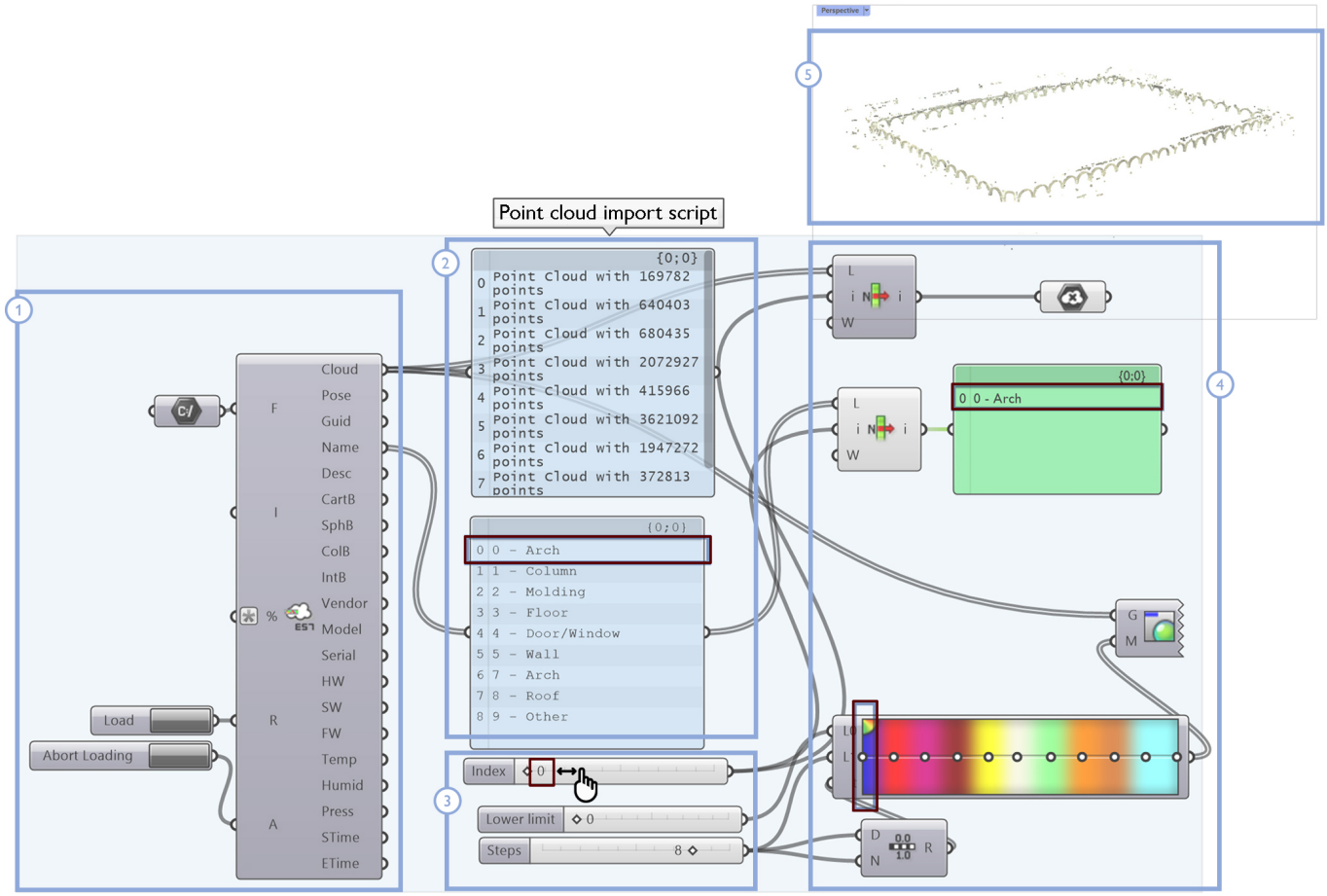


Figure 7-6 *(above)*.

Point cloud import script. The meaning of each overlay rectangle is specified in the text.

e.g.: class 0 – 'Arch',

Grand cloister, Pisa Charterhouse



Class 4 - Door / Window



Class 9 - Other

classification approaches (Section 5.3) and they relate to the architectural typology of the medieval cloister.

A first problem in the reconstruction process is identified in the need to preserve the classified 3D data even when importing the point cloud within the modeling environment. In other words, the result of the semantic segmentation procedure, i.e., the annotated point cloud, provides a 3D datum that is already segmented, and this means that each class of typological elements detected therein might be isolated and shown singularly (as in the example of Figure 7-5). The information related to the class might thus be made available and visible even when shifting to the 3D modeling environment, so as to start the generation of template geometries, class by class, on the basis of the annotated point cloud.

To this task, a point cloud import algorithm is initially developed and built up in Grasshopper (Figure 7-6), leveraging some nodes available in the Volvox plugin, that enable the link between point clouds and modeling

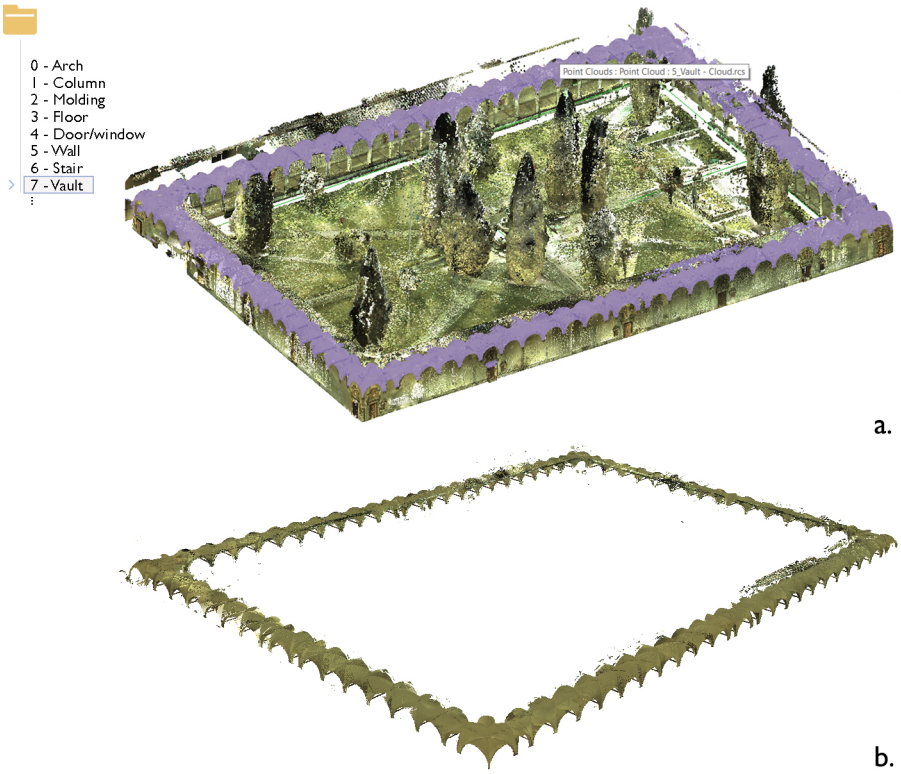


Figure 7-7.

Segmented point cloud import. Selection (*a*) and isolation (*b*) of the 'Vault' class.

e.g.: Grand cloister, Pisa Charterhouse

environments. In detail, the developed algorithm reads and sorts the point cloud file (1), while recognizing indices and colors that are associated to each class of typological elements. The original point cloud is segmented into multiple point clouds corresponding each to a single class (2); then, the management of a dedicated slider allows the user to directly select the index of a desired class (3). The colors, names and points associated to the selected class are this way browsed (4), and the corresponding point cloud is displayed in the graphical interface (5).The advantage of such a procedure lies in the user capability to toggle on or off, according to its requirements or needs, the visualization of a given class: all of the objects that are not selected are in fact immediately hidden from the original point cloud and only 3D points associated to the selected class remain visible (Figure 7-7). This direct manipulation of the classes of architectural components over the original point cloud simplifies the modeling phase even at its first stage, as the 3D reconstruction work can be carried out on an already segmented (reduced) dataset.

7.3. Results: Libraries of *template* geometries

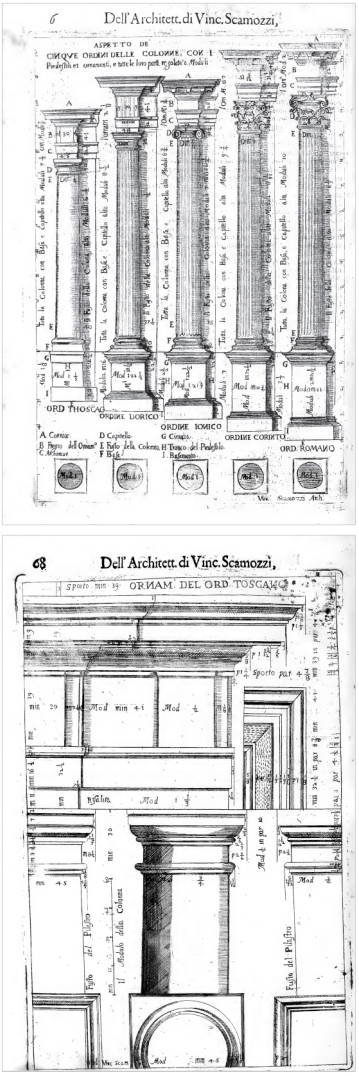


Figure 7-8.

Studies on the architectural orders (*a*); entablature of the Tuscan order (*b*) in the treatise by Vincenzo Scamozzi.

(Parte Seconda, Libro Sesto, Cap XVII)

Once the semantically segmented point cloud is properly imported in modeling environment, the *template* geometries are reconstructed for each class in the form of parametric 3D objects, starting from the distinction into class and thus re-proposing the classification into recurring typological architectural elements.

Again, the need to reconstruct geometries with editable and adaptable parameters leads to the choice of visual programming languages, as the latter allow to:

- Guide the reconstruction of conceptual shapes by a series of rules and processing operations (i.e., by connection of nodes and wires in VPL).
- Directly modify the reconstructed shapes, editing their size and properties by control of the different graphic elements that compose them and by direct manipulation of sliders related to their dimensions.

Many relevant classes of architectural typologies – such as the column, the arches and the vaults –, are reconstructed with reference to architectural canons. For the case studies considered, the treatise *The idea of the universal architecture* by Vincenzo Scamozzi (Figure 7-8), is taken into account since being the work that inspired, during the 17th century, the renovation of many settings of the Pisa Charterhouse.

In particular, for the *template* shapes reconstruction process, the formalization approach followed (De Luca et al., 2007) is based on: the interpretation of any knowledge referring to the shape; the identification of the modeling methods required; the identification of the relations between the different elements.

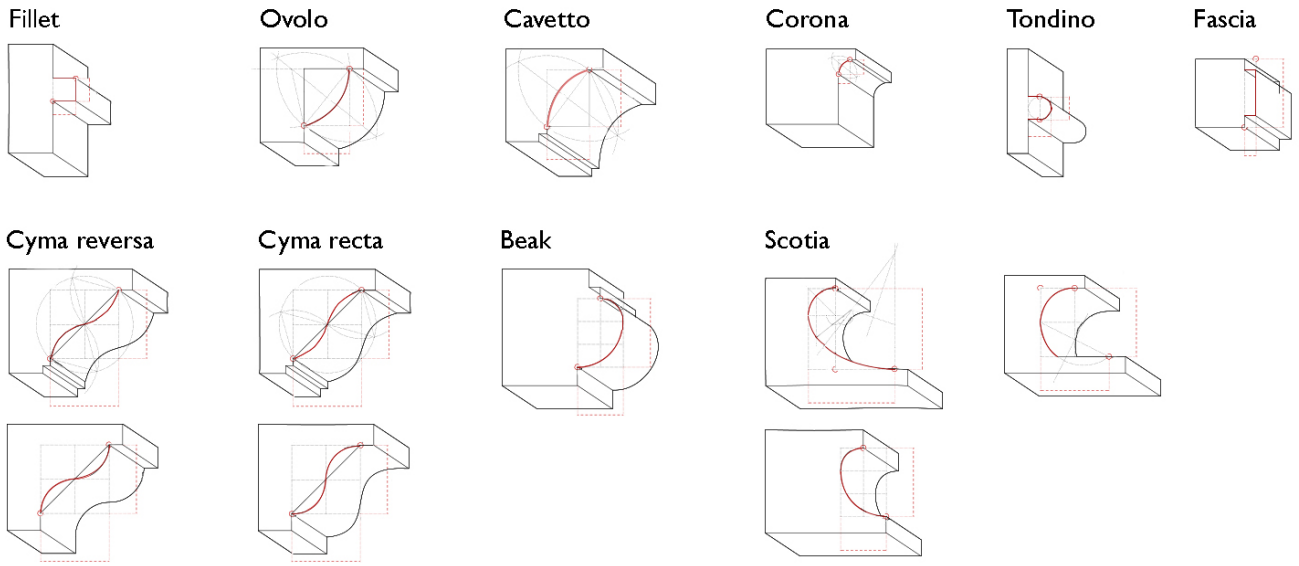


Figure 7-9.

Study on *template* moldings with formalization of the *shape grammar.*

As such, we identify the elementary entities and formalize architectural primitives on which the representation should be based.

Once the 3D points belonging to each class of elements have suitably been isolated on the viewer, in fact:

- Reference construction planes are detected and properly oriented.
- A generating profile is created by VPL and a directing path, where needed, is delineated.
- Functions such as revolution, sweep, extrusion, loft etc. are applied to the identified profile to obtain the desired surface.

The parametric geometry so constructed, representative of each class

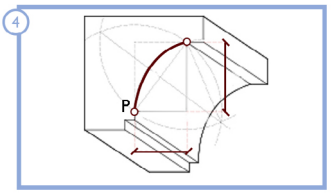
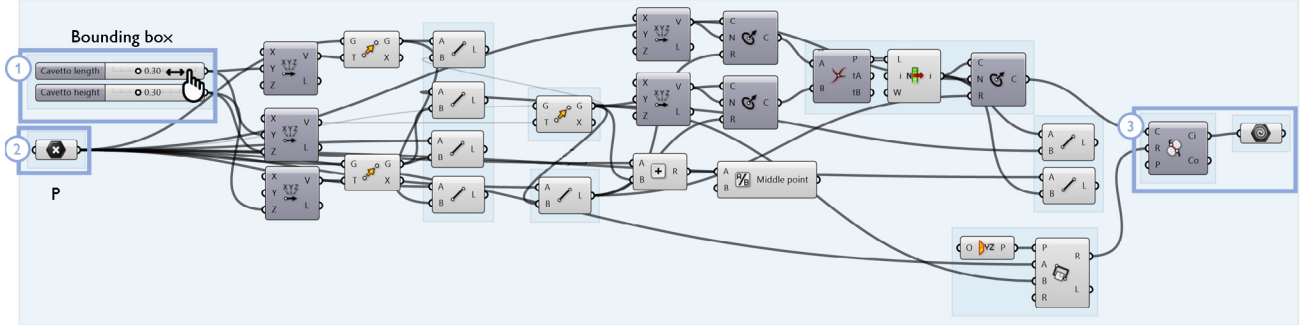


Figure 7-10.

Study on the construction of the molding '*cavetto*'. The molding can be modified by editing its bounding box (1) and its anchoring point (2). The resulting curve is modified immediately on the VPL interface (3) and in the viewer (4).



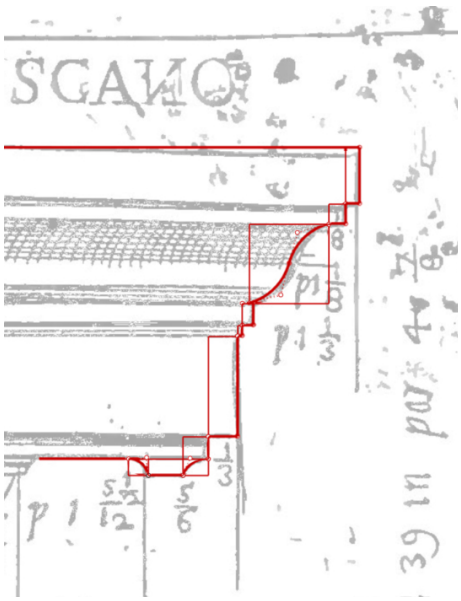


Figure 7-11.

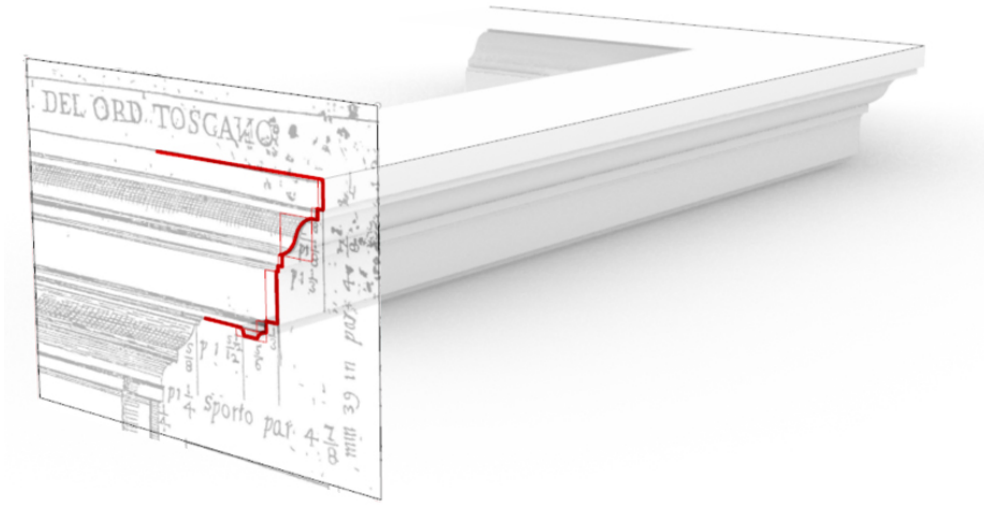
Study on the combination of multiple moldings for the creation of a profile *(above)*.

The generating profile (*a*) is swept along the directing path (*b*) to create a 3D surface of the entablature *(below)*.

of elements, can later be adjusted to the original survey data through the management of sliders and parameters, as well as through the direct visualization of the correlated outputs on the point cloud in Rhino3D graphic environment.

For the formalization of the template geometries, the moldings are considered first because they constitute the smallest and most trivial units of architectural elements corresponding to a semantic description of the building, and also because their combination can give rise to more complex architectural elements.

Many moldings of the treatise by Vincenzo Scamozzi have been analyzed and studied (Figure 7-9), and an example of their generating algorithms' structure is provided in Figure 7-10. It is noted that each molding is represented by a curve (drawn on a selected construction plane) and is included within a bounding box, i.e., a deformable section defining the height and width dimensions of the element. A starting point and an end point provide the anchoring points of any other moldings connected to it.



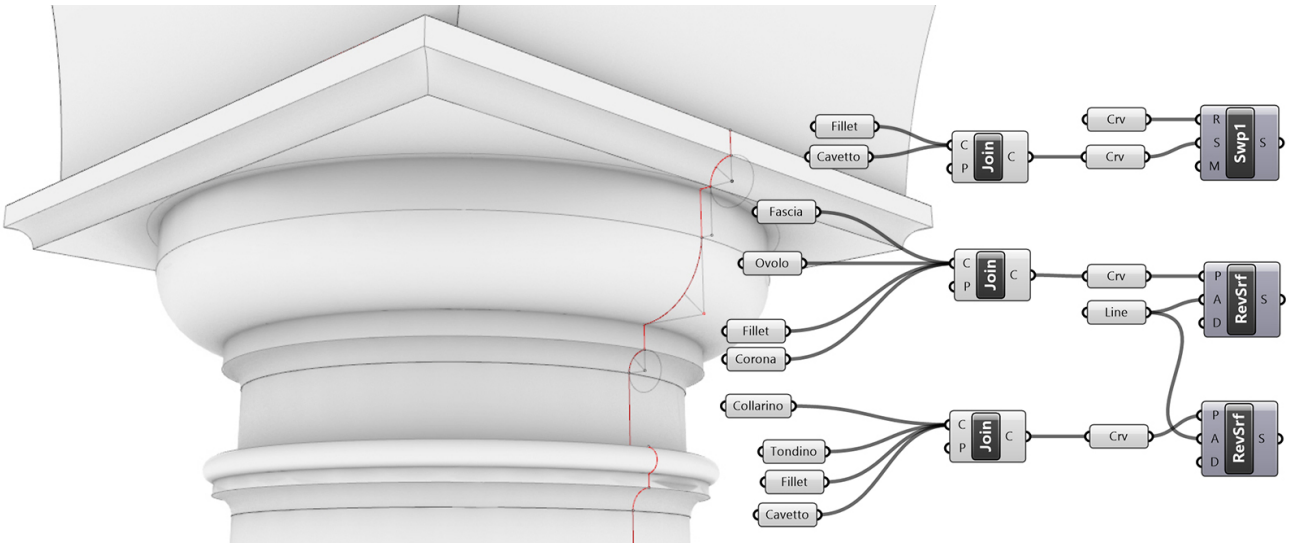


Figure 7-12.

Creation of a capital of the Tuscan order for instantiation on the point cloud of the Grand-Ducal cloister, Pisa Charterhouse.

In particular, the overlay rectangles of Figure 7-10 (1) highlight the sliders that the user can manipulate to modify the various moldings, e.g., in terms of width and height of the bounding boxes. A modification of the sliders in VPL automatically activates the modification of the concerned curve. Considering the insertion points of the moldings as anchoring elements for the construction of subsequent moldings, more complex profiles can later be provided, as in the example of Figure 7-11.

The so constructed generating profiles and directing paths can be subject to application of extrusion, loft, sweep and revolution functions, so determining the *template* shape of a class.

Figure 7-11 provides the example of a capital: the generating profile is given by the combination of several architectural moldings, as theorized in architectural canons, and the use of sweep and revolution operations allow delineating its shape in the 3D space.

Again, Figure 7-12 displays the construction process followed for the column shaft of the Grand cloister dataset: at first, descriptors and geometrical attributes are defined for the construction of this architectural component. Hence, the column shaft is constructed relying on the study of the dimensional relationships between the column base diameter and its height, which allows to establish, at different heights, reference circles. Later on, these selected circles are used to define, via a loft function, the conceptual surface shape of the concerned object, even representing the *enthesis* of the column shaft.

By extending these processes to the whole set of classes, it is possible to defined for each one a *template* geometry, i.e., a series of rules and visual scripting nodes enabling its construction.

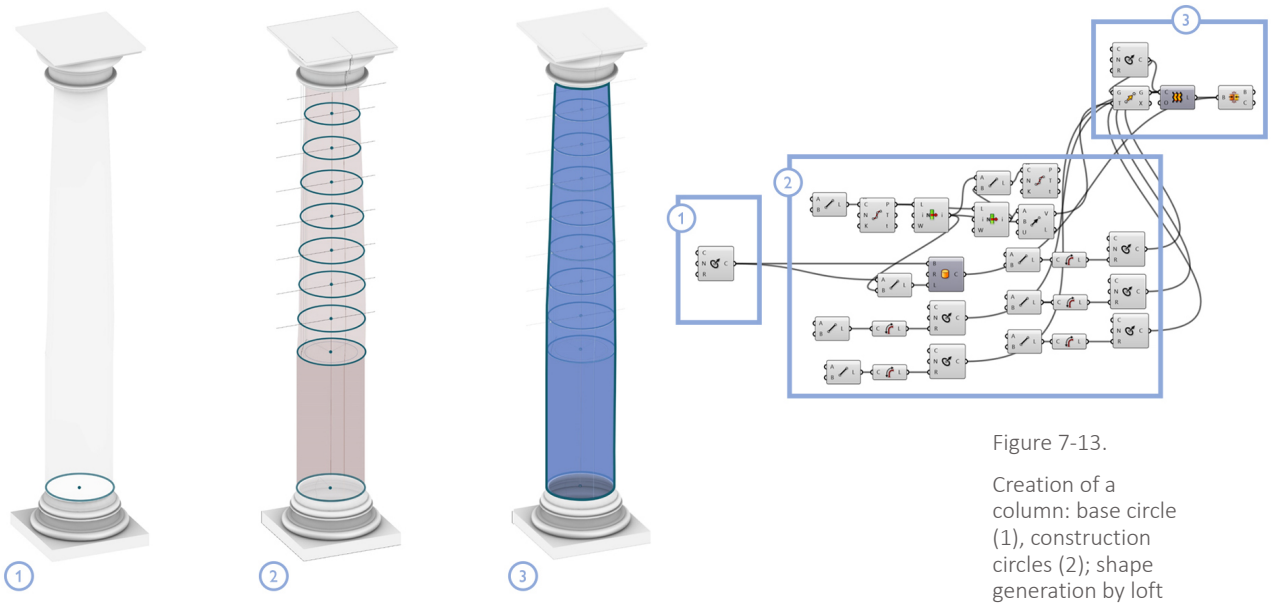


Figure 7-13.

Creation of a column: base circle (1), construction circles (2); shape generation by loft function (3).

The above-described algorithms, in fact, together with defining a series of operations allowing the construction of the *template* geometry of each class, further allow controlling and editing the dimensions of this latter via the manipulation of sliders. The reconstructed geometries of other classes, such as the ones relating to the vaults and the arches, are displayed in the following section.

7.4. Results: Information propagation

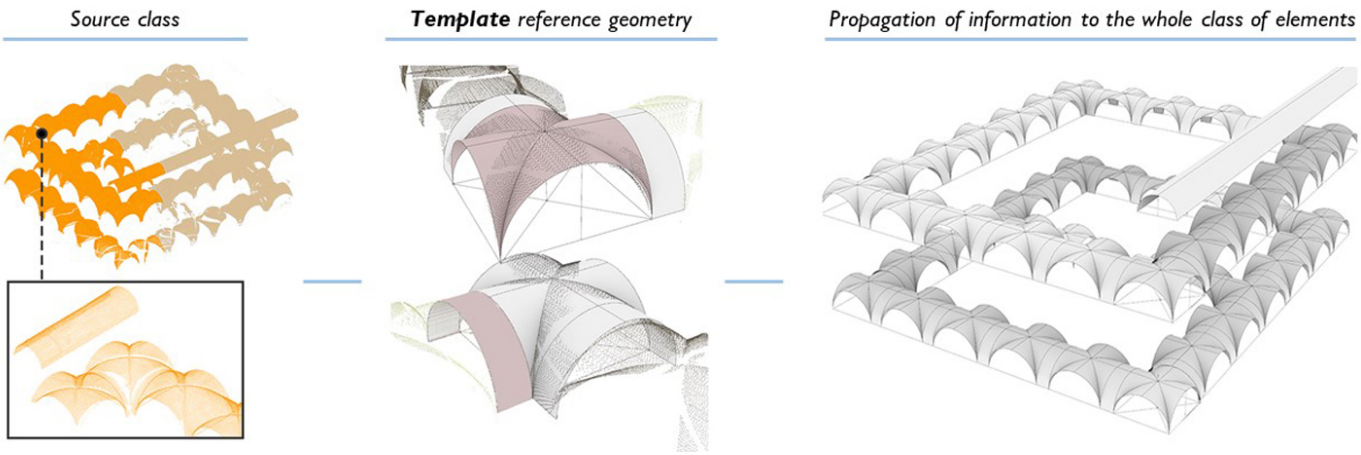


Figure 7-14.
Propagation workflow.
e.g.: Class 7 – Vault, Grand-Ducal cloister, Pisa Charterhouse

Once defined for each class, the *template* geometry is later propagated to other parts of the point cloud, where points have been recognized as belonging to the same architectural typology (Figure 7-14).

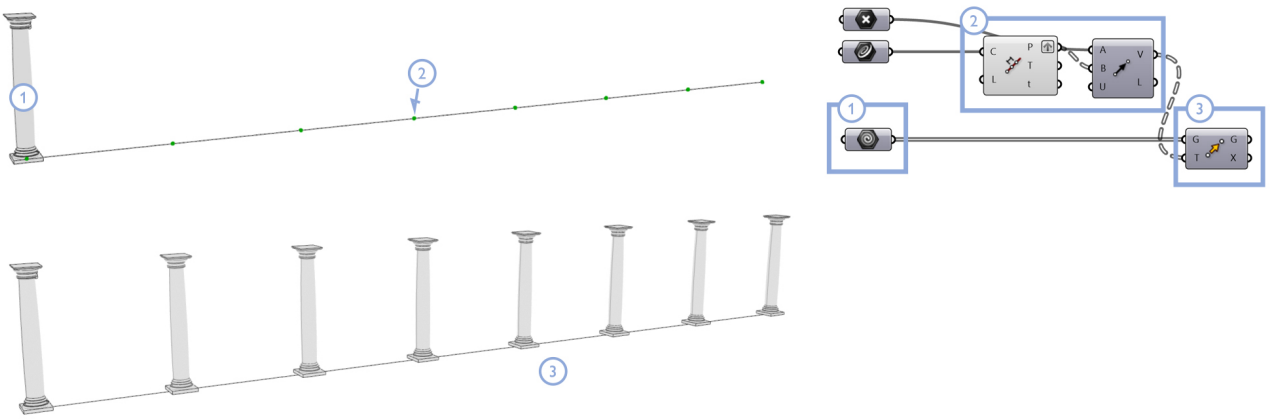
The procedure takes place via duplication and displacement nodes, in particular considering the array, copy and translation operations (e.g., the *'Move' node* in Grasshopper).

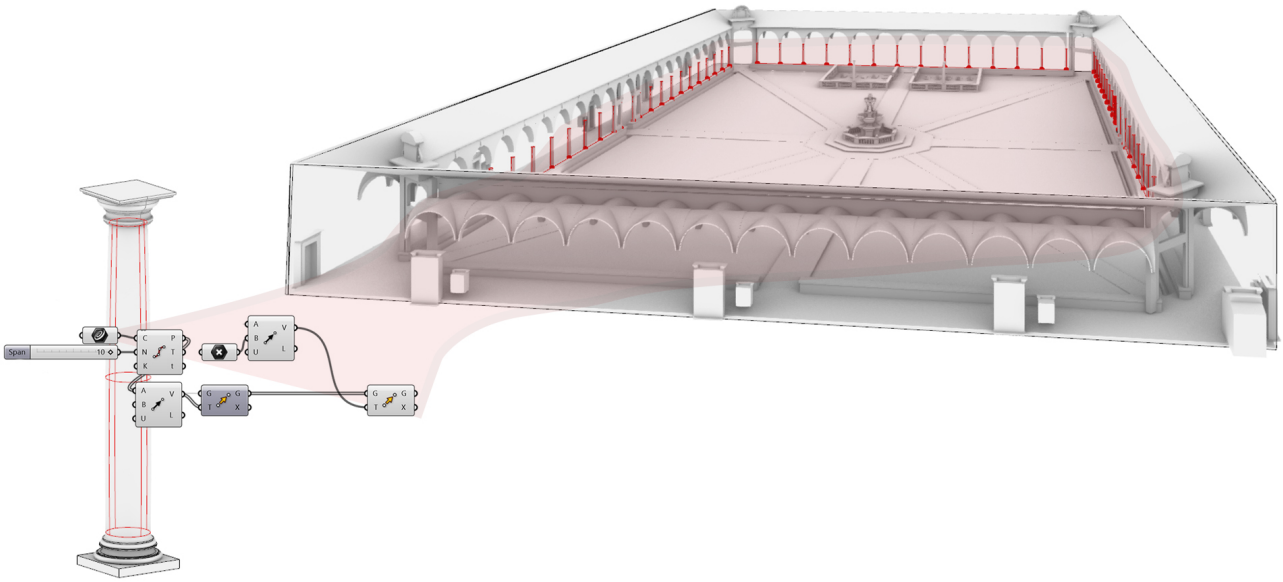
A list of items can be considered to perform translation and replication operations involving multiple elements concurrently (Figure 7-15).

These operations enable the repetition, as many times as necessary, of the 3D geometry for each class of architectural components (Figure 7-16).

Figure 7-17 shows the results obtained in the construction of some

Figure 7-15.
Example of propagation algorithm: template geometry *(1)*; insertion points and Move node *(2)*; propagated geometries *(3)*.
e.g.: Class 1- Column, Grand cloister, Pisa Charterhouse





significant classes extracted from the considered datasets: the ‘column’ class for the Grand cloister, the ‘vault’ class for the Grand-Ducal cloister and the ‘arch’ class for the National Museum of San Matteo.

The creation and propagation of conceptual geometries via generative design rules allows to manipulate each time the repetition – and, eventually, the parameters’ modification – of these reconstructed graphic elements.

Figure 7-16.

Propagation of the Column class.

e.g.: Class 1-Column, Grand cloister, Pisa Charterhouse

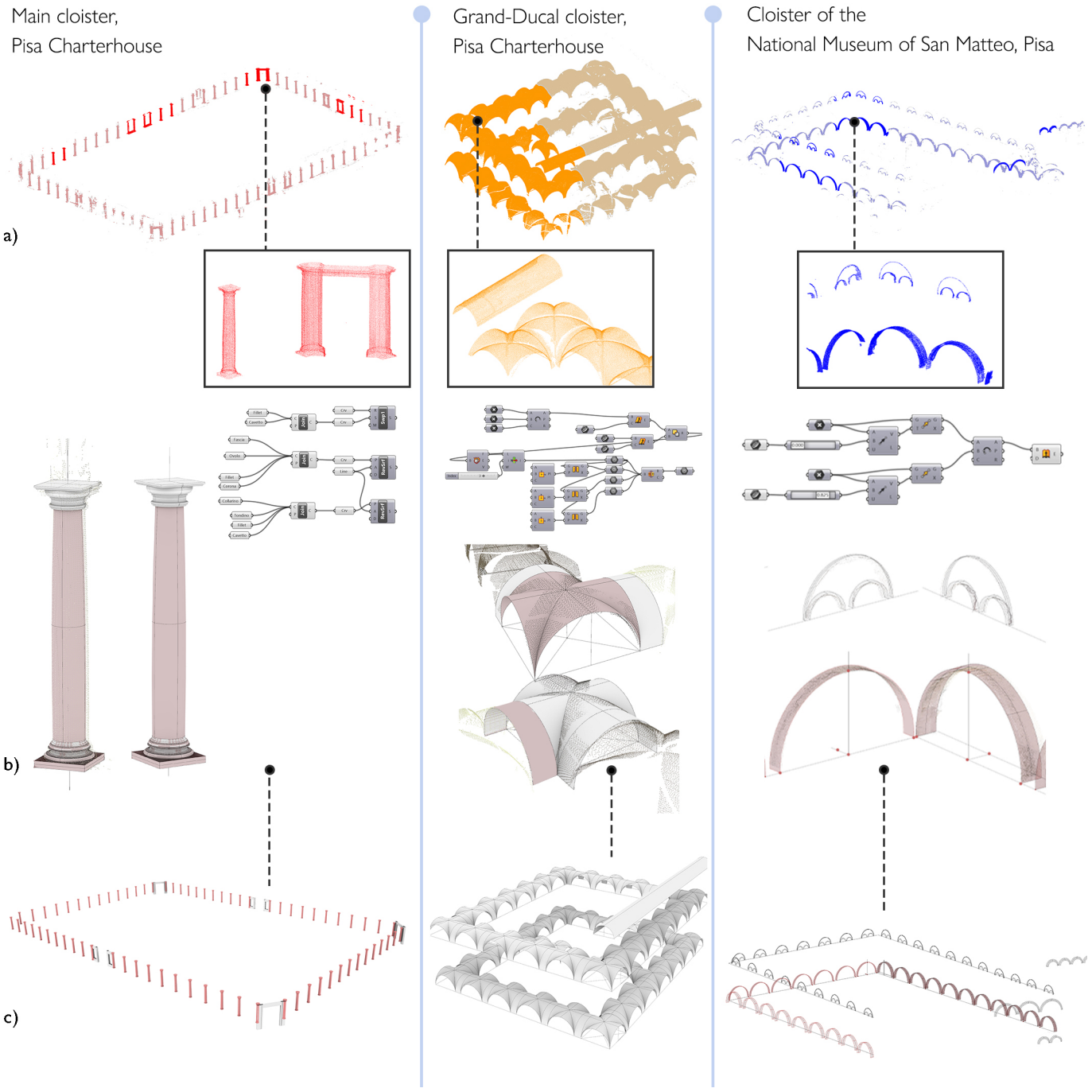


Figure 7-17.

Examples of reconstruction of parametric components: original class (*a*), conceptual reference geometry (*b*) and propagation of the information to the whole class (*c*).

7.5. Import into BIM-based software for Rhino.Inside Revit

At the end of the process, if the propagation of conceptual geometries is extended to the whole set of classes, a final model of the considered object can be obtained (Figure 7-22).

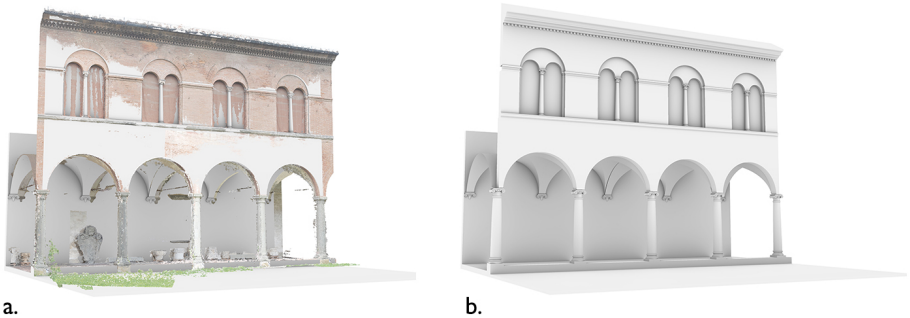


Figure 7-18.

Resulting conceptual model for a portion of the cloister.

e.g.: National Museum of San Matteo.

Each architectural component identified within the individual classes maintains its own identity and semantic description, i.e., it is linked to the semantic decomposition of the building, and it is defined by a template reference geometry (Figure 7-19).

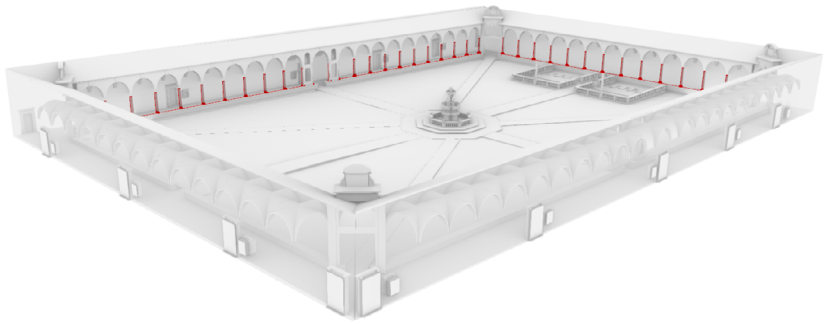


Figure 7-19.

Resulting conceptual model and selection of the 'Column' class.

e.g.: main cloister, Pisa Charterhouse.

This means that each object making part of the model can be reconstructed whatever the software used and be imported as well, e.g., in BIM environments (Figure 7-20), so to be further enriched with non-geometric information (related, e.g., to materials, restoration and consolidation interventions, documentary and analytical sources, preservation state and so forth).

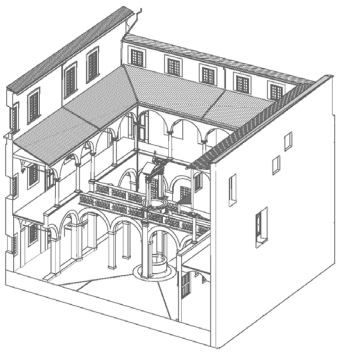


Figure 7-20.

Resulting model imported into Autodesk Revit.

e.g.: Grand-Ducal cloister, Pisa Charterhouse.

In the presented case, for instance, the VPL in Grasshopper could be provided a connection with the BIM software Autodesk Revit, thanks to the **Rhino.**

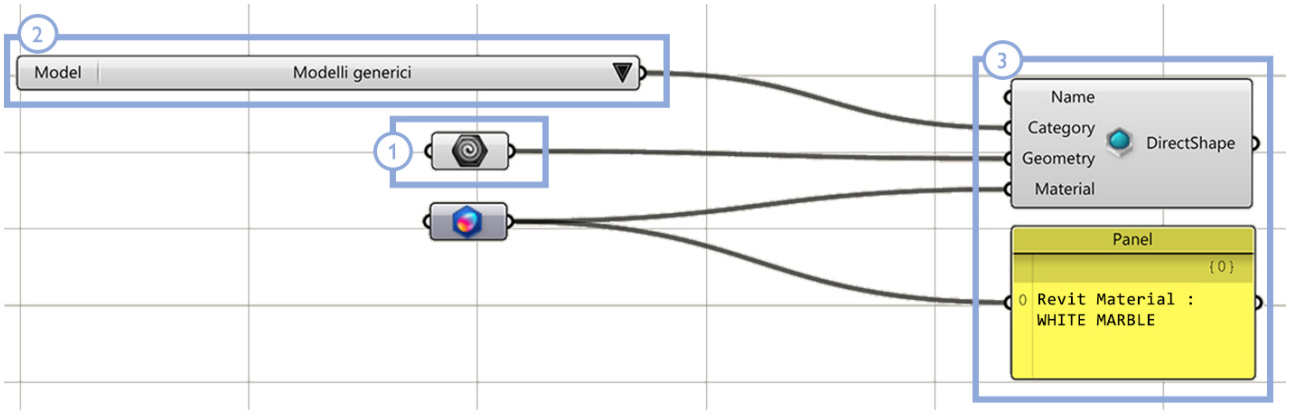


Figure 7-21.

VPL import script via *Rhino. Inside. Revit*: the selected template geometry (1) is imported in Autodesk Revit as a generic model component (2), and it is associated the Revit Material 'White marble' (3).

e.g.: Grand-Ducal cloister, Pisa Charterhouse.

Inside. Revit plug-in. Among others, this is a possible way of importing the classified model into Revit: the plug-in, in fact, enables the selection of the architectural objects belonging to the single classes of elements and their association to a Revit *family*, thus preserving their semantic description level.

7.6. Hints to rules of association between template model–real model

In the previous part of the chapter, we focused on the reconstruction of object classes from a semantically annotated 3D datum. The resulting representation, based on primitive geometries, can be used as the basis for the generation of an information system where to store knowledge-related data: the latter may either be graphically associated to a whole class of elements, or to the single elements, as well as to parts of them.

Although the annotation of classes of objects or of single objects on this template reference model seems trivial, as it follows the logic of H-BIM environments, the annotation process might become challenging when dealing with localized annotations.

Considering this latter case, this paragraph presents several insights on possible future developments of this study towards the insertion of more localized information e.g., relating to portions of a single architectural component, so to note, *inter alia*: historical, architectural or artistic data, the presence of frescoes and decorative parts, degradation phenomena, crack patterns, repair and restoration interventions and so on. As in the example of the ortho-photo of Figure 7-22, we want annotations performed on parts of the single component (e.g., the

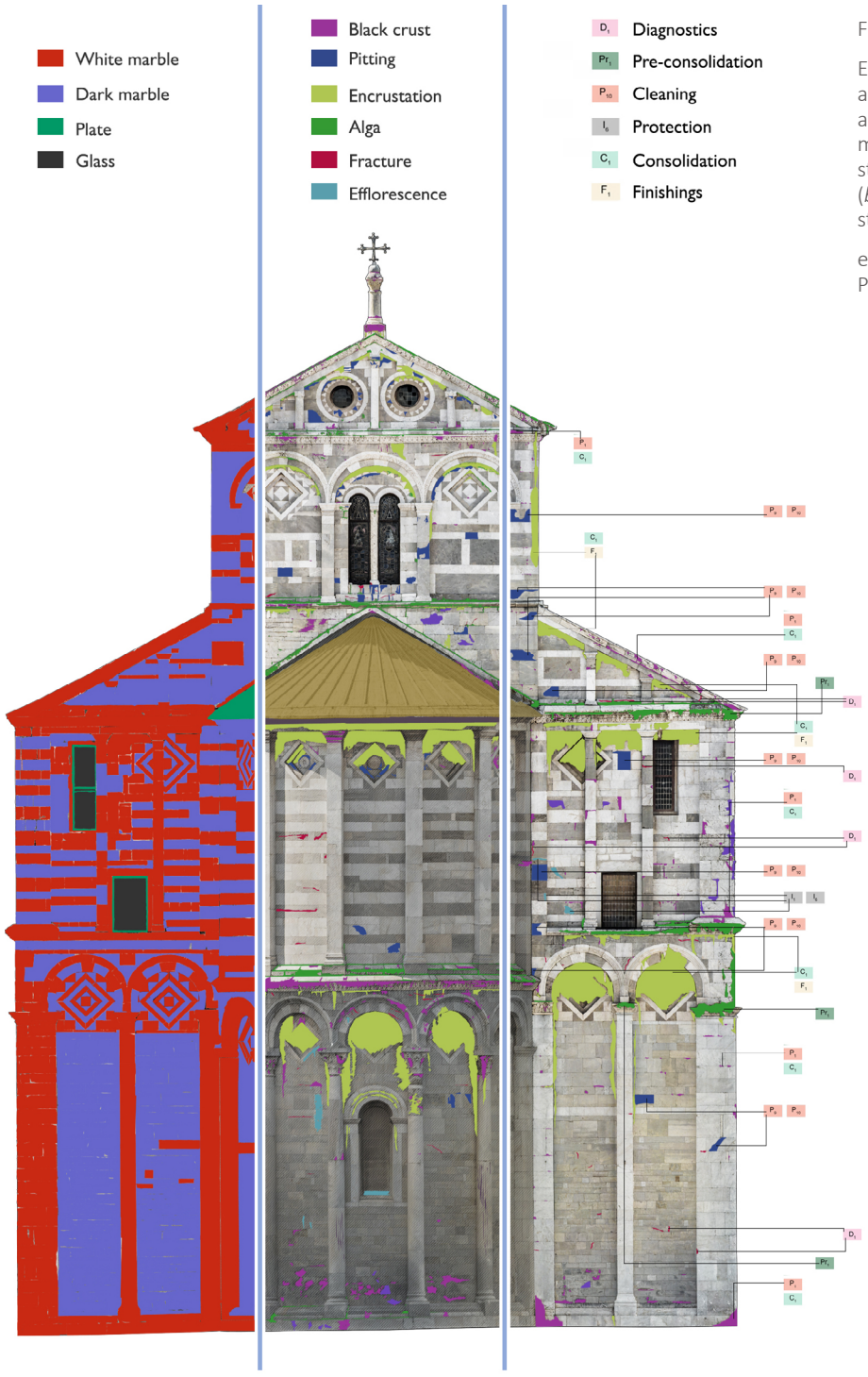


Figure 7-22.

Example of localized annotations over an ortho-photo: material survey (*a*); state of degradation (*b*); intervention strategies (*c*).

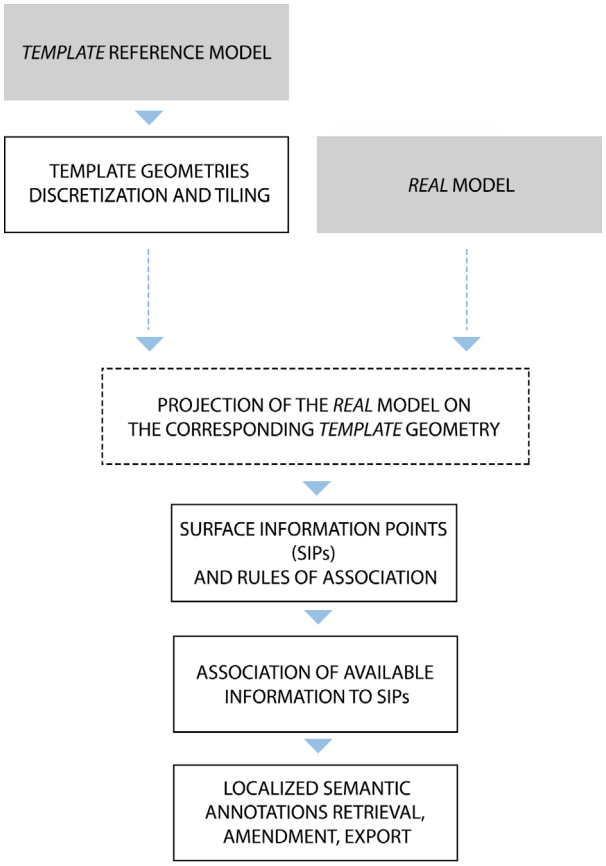
e.g.: North transept, Pisa Cathedral.

column, the wall, the capital and so forth) to be transferred from the 2D to specific portions of the corresponding 3D *template* reference geometry.

In order to refer the semantic information to specific portions of each template geometry, a possible workflow could be the one presented in Figure 7-23.

Figure 7-23.

Hint to a possible workflow enabling association between template geometries-real models.



In this case, the *template* model obtained at the end of Section 7.4, since based on conceptual shapes (mathematical relationships), may be used as a reference when producing the documentation of different types. By referring each information to the *template* reference system of the building, the different analyses produced by the working groups could in fact be spatialized on a common base, regardless the type of support on which the different information are contained.

The process is envisaged as follows: starting from the reconstructed model, an adequate discretization and tiling procedure could be

Figure 7-24.

Example of tiling structure for the portion of a capital with identification of SIPs (points) representative of each tiling surface.





Figure 7-25.

Example of *real-template* model association: *a)* thematic mapping over the orthophoto; *b)* suitable projection of annotations on the *template* representation; *c)* point-based representation and *d)* point-based thematic mapping over the 3D model.

e.g.: North transept, Pisa Cathedral

proposed for each template reference geometry, so to identify significant and representative elements (such as points) to which information is associated each time.

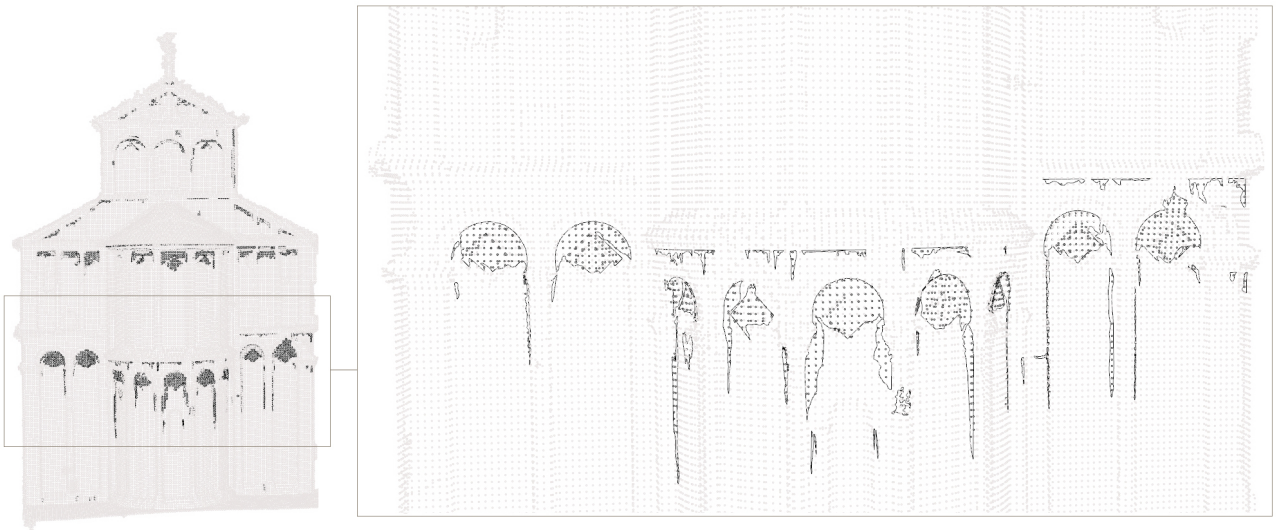
As such, each part of the *template* geometry could be represented by a finite quantity of points, suitably numbered and representative of each face of this tiling structure (Figure 7-24), to which semantic information could be connected each time.

This process would allow the association of new and raw data, whatever

Figure 7-26 *(below)*.

Example of point-based annotation of the 'encrustation' phenomenon projected on the reference model.

e.g.: North transept, Pisa Cathedral



the tiling of the real model is, to a point or to a sample of points over the *template* geometry. As such, the elements of the *real* representation (either a point cloud, an image, a mesh, an ortho-photo and so forth) could always be put in correspondence with a point of the template model (Figure 7-25); this would allow to interrogate the real model so to understand what kind of information is associated to that point (Figure 7-26, e.g., in terms of degradation patterns, conservation states, presence of signs of degradation, decorations).

For this task, association rules are demanded to describe a univocal correspondence between tiles of the *template* model and parts of the real model. This could be achieved, for instance, by means of a suitable projective relationship connecting the real tessellation to the *template* surface.

The information retrieval process, in this way, could be either developed in input or in output: indeed, information could be entered by associating the descriptive data to the representative point on the *template* model – and, through this point, to the points/elements of the *real* model (input). Conversely, any information needed concerning the *real* model could be obtained referring to the representative point of the template model.

This tiling operation, point by point, could also be performed at several levels of detail or even concern the entire element, so enabling the association of information to multiple sets of points or surface elements (Croce et al., 2021b).

7.7. Assessment of the proposed approach

In this chapter, a workflow for the semi-automated reconstruction of 3D models starting from annotated surveying data is proposed, relying on visualization and import of annotated data, and on the construction and propagation of *template* geometries. By this approach, the geometric nature of the building components is reconstructed, the originally designed shape is interpreted and a reference geometry is identified and modelled for each class, following the logic of the parametric BIM families. Upon completion of the process, visual programming techniques allow for propagation and dimensional comparison of repetitive elements.

First of all, the integration of semantic segmentation methods exploiting AI and graphical algorithm editors for 3D modeling introduces relevant developments towards the automation of Scan-to-BIM workflows. A direct link between the point cloud and the reconstructed model is indeed established at the level of the single classes of elements, and this avoids the loss or possible dispersion of information in the transition from 3D survey to parametric model.

The conceptual representation obtained provides an effective support tool in the documentation of each architectural asset: the digital model obtained, since being based on *template* geometries (even generated referring to architectural canons and treatises), becomes valid regardless the changes and modification of the real object, and additional information can thus be inserted, retrieved, edited and updated therein, as a relevant foundation for H-BIM type systems. Indeed, the use of architectural canons and the definition of a series of rules allowing to reconstruct or edit a 3D object makes this system independent of the software used. This also implies that Grasshopper and Rhino3D do not represent the unique environment where such methods could be implemented, but rather other generative modeling software (such as Dynamo, implemented in the Autodesk Revit software) could be leveraged to the same task, if the same construction, modeling and operation processes are recovered and re-proposed therein.

Moreover, as a consequence of the established connection between the reconstructed classes and the semantic point cloud, it is relevant to note that the two representation types could be compared one another, e.g., in terms of relative distances, so as to derive to what extent, for each class of elements, an almost conceptual (digitally reconstructed) model diverges from the *real* data (e.g., from a point cloud acquired by surveying). With this intention, the comparison between the two models (real-ideal) could lead to the construction of disparity maps

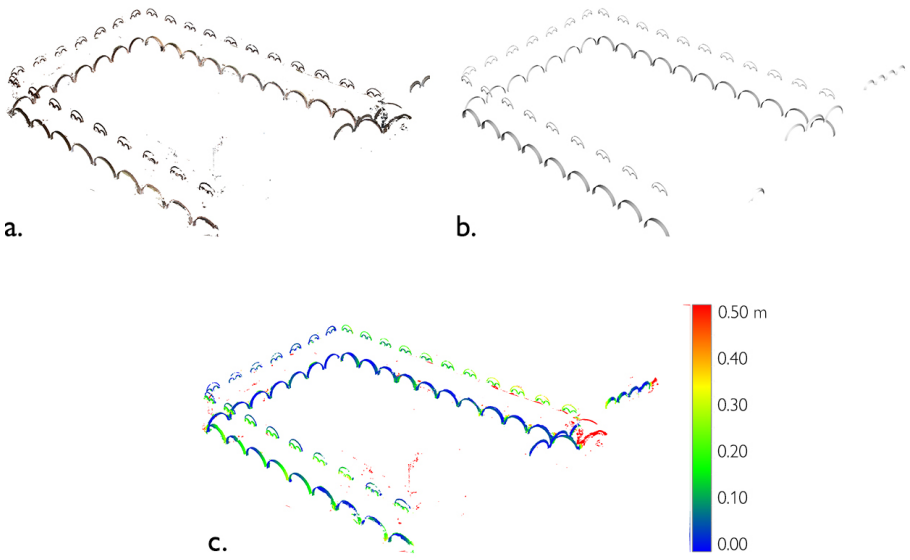


Figure 7-27.

Real model (point cloud, *a*) and template model (*b*) are compared to extract a disparity map (*c*). The signed distances between the two models are expressed by a dedicated color scale.

e.g.: class 0 – 'Arch', Grand-Ducal cloister, Pisa Charterhouse.

showing the variability, in space and time, of *real* (existing) architectural elements with respect to their related *ideal* model (Figure 7-27). For instance, such a study could further allow the refinement of the *template* geometry or could impact the study on the evolution in time of an architectural style as well. A similar approach to the comparison of *real-template* geometries was previously presented in the work by (De Luca and Lo Buglio, 2014) who identify these process as *the hybridization of the complexity (representing instances without interpretation) with semantic concepts (linked to the notion of intelligibility).*

The connection point cloud-3D conceptual model allows to enrich each element belonging to a given class with further analytical or technical data. However, it has to be observed that the transfer of localized information – i.e., related to small portions of the template geometries –, is not yet possible, unless appropriate subdivisions of the reconstructed surfaces are provided to link the information at a higher level of detail.

In this latter regard, the last part of this Chapter introduces future developments of the present work, that aim at improving the insertion of localized annotations on the 3D model, as well as at transferring such type of information from the point cloud to the BIM environment. Some inputs are thus provided concerning ways of association between this information and characteristic elements identified over each reconstructed 3D *template* geometry, so hinting at further possibilities of semantic enrichment.

As such, the conceptual meta-model built might be linked to all other existing representation models (mesh, point clouds, photos, ortho-photos, 2D graphical drawing etc…), even enabling, in a H-BIM

perspective, the association of the information to localized areas.

The 3D reference model so constructed, if suitably tiled, could be valid once and for all and it could be used whenever the information update, recover, validation, or amendment on the real/digital model is required. To accomplish this, it is assumed that descriptive data might be inserted considering a suitable set of representative points of this reference model, to which information is associated each time. So doing, it would indeed be possible to shift to a representation based on Surface Information Points (SIPs) identified over the source model, each one identifying a specific surface tile, to which information related to the thematic mapping of each tile is linked. This topic, since being the subject of future work, was presented in this Chapter as a possible further extension of the approach proposed for annotation transfer.

Chapter 8

Conclusions and future perspectives

8.1. Assessment of the proposed method and main contributions

This thesis work proposed an integrated methodological approach for the transfer, the retrieval and the exchange of semantic annotations of 2D/3D digital heritage models, based on the combination of Artificial Intelligence techniques, H-BIM environments and collaborative reality-based annotation platforms as Aïoli. The approach is designed to support restorers, conservators, historians of art, architects, engineers, archaeologists and other experts, working on existing heritage objects, monuments and sites, involved in documentation, restoration and enhancement process and continuously dealing with the problems of collection, retrieval, fusion and updating of heterogeneous information over digital representations of heritage artifacts.

The analysis of the state-of-the-art conducted in Part I revealed the existence of a plurality of representation methods, both in 2D (images, ortho-photos, drawings, UV maps…) than in 3D (point clouds, textured meshes, parametric models…), for the digital documentation of our cultural and architectural heritage. At the same time, the literature review unveiled the evident lack of streamlined methods and useful tools for the correct archival and collection of the heterogeneous knowledge produced within the study of heritage objects.

The plurality of information and the existence of multiple 2D and/or 3D representation types, in fact, is often the source of data dispersion, and raises the issue of the so called *digital continuity* of data, intended as the need to ensure suitable retrieval, traceability, maintenance and interpretability of the information regardless the type of digital representation chosen.

To bridge this gap -and exploiting the semantic annotation mechanism for the association of knowledge-related (*semantic*) information to purely metric digital data- this thesis proposed a series of methods and tools to support appropriate interpretation, collection and sharing of digital heritage information. The methodological approach illustrated and developed in Part II thus aimed at enabling the transfer, the retrieval and the exchange of semantic annotations over 2D/3D media, to enhance documentation and dissemination practices.

In detail, the methodological aspects have been developed by exploiting and combining raw surveying data, Heritage-Building Information Modeling (H-BIM) systems and reality-based collaborative annotation platforms such as Aïoli. The semi-automated annotation mechanisms proposed are structured according to the three respective phases of:

i. Semi-automatic semantic segmentation (classification) of surveying data through Artificial Intelligence.

ii. 2D/3D annotation transfer.

iii. H-BIM reconstruction, semantic structuring and insertion of localized information.

The approach, although requiring the supervision of skilled users (engineers, architects, conservators or restorers involved in the documentation process), ensures a high level of automation and enables the association of *semantic* information to raw digital representations derived from surveying, e.g., in terms of mapping of degradation phenomena, materials, intervention techniques and measures, different temporal states, on both 2D and 3D media.

The proposed tools consider the application of existing licensed or open-source software so to investigate the best attainable workflow for the insertion, transfer and semi-automatic archival of semantic annotations on digital heritage models. They allow to:

- Classify and interpret raw data obtained from surveying, proposing procedures for a more effective management of digital heritage data.
- Ease the comprehension and knowledge of the object, by

restitution of its visual, geometric and descriptive attributes, and by addition of knowledge-related annotation referring, e.g., to degradation phenomena, historical information, alterations or material mapping and so on;

- Enable the multi-modal connection and interlinking of different digital sources (2D or 3D) and models of a same heritage artifact;
- Guide existing software tools towards the improvement of Aïoli and H-BIM systems, and towards the management of collaborative annotation platforms in general, based on initial on-site acquisitions;
- Build a reference model of a heritage object, that can be leveraged for the construction of H-BIM type informative systems, for further semantic enrichment.

The different tools proposed for the three phases of the approach may be adapted on a case-by-case basis, according to the restitution needs and requirements established each time by the expert working on the heritage object. As previously stated in Section 4.4, we did not intend to propose a single *end-to-end* workflow to be applied every time to each case study, but rather the general approach defined and divided into the three respective phases -i.e., semantic segmentation via AI, 2D/3D information propagation, H-BIM reconstruction, semantic structuring and transfer of localized information-, was proposed to be adapted, integrated and validated each time, according to the case-specific needs for description and documentation.

By considering the results obtained and the experiments conducted on several case studies of the Italian and French heritage, several concluding remarks and related limits may be drawn for the different phases of the approach.

8.1.1. Semi-automatic semantic segmentation (classification) of surveying data through Artificial Intelligence.

In the first phase, Artificial Intelligence algorithms are exploited to ease the reading and interpretation of raw surveying data of a heritage object. Geometry-based and texture-based approaches enable, to different extents, the addition of a semantic label associated, e.g., to the decomposition of a building into its recurring architectural elements, to the material mapping, to the degradation states and so on.

Among them, geometry-based approaches are more tailored to the recognition of recurring architectural elements and components, even as a fundamental and necessary prelude to Scan-to-BIM reconstruction

applications.

On the other hand, texture-based approaches, since being based on color information, may support restoration operations which require semantic mapping and comparison between successive conditions of degradation of surfaces and/or of material states.

By application of AI, and ML in particular, the human involvement is reduced in the brute annotation phase, lessening the space for arbitrary and too subjective choices and significantly reducing the annotation time. A semi-automatic annotation process is so realized by leveraging the suitable manipulation, export and extraction of geometric and visual descriptors (*features*) from raw 2D or 3D data.

However, the user's supervision, in terms of choice of the training set and decision on which classes should partition the digital data, is crucial in determining the success of the classification and labeling process. In addition, high performance computing machines are required in the feature extraction and data-driven algorithm training phase. Moreover, mis-classification may occur in boundary regions, as well as in regions with similar development (geometry-based approaches) or with similar color characteristics or patterns (texture-based approaches).

The generalization of a same ML algorithm to other datasets, related to different architectural types and/or built in different periods, and the creation of larger annotated datasets to train deep neural networks, are possible future developments in this domain.

8.1.2. 2D/3D annotation transfer

The second phase (*ii*) focused on the importance of considering the semantically annotated data (obtained either manually or more automatically via the segmentation process) not only as the informative resource of a single digital representation *per se*, but rather as a starting point to enrich multiple types of representations of the same architectural object.

The different annotation transfer tools proposed focus on preserving semantic annotation information even in the transition:

- to a different type of representation (point cloud, mesh, image, orthophoto...);
- to successive levels of digital resolution and/or detail.

To this task, possible mechanisms for propagating annotations from 2D to 3D digital objects (and vice versa) are developed, to guarantee that the annotation information, whether automatic or manual (hence, regardless of how it was obtained), might be preserved in the transition, e.g., from images to 3D model, or from more detailed 3D model to less

detailed 3D model and so on.

Several methods were featured for the semantic enrichment of digital representations in hybrid 2D/3D annotation mode, even taking advantage of the most recent developments of the Aïoli collaborative reality-based annotation platform, developed by the MAP CNRS/MC research unit. The implementation developed for the Aïoli platform consisted, indeed, in the combination of its hybrid 2D/3D annotation system with supervised classification methods. It was shown, in this regard, how information can be semi-automatically propagated with a collaborative and web-based annotation approach.

The collaborative annotation tools, enabling 2D/3D transfer of semantic data, confirmed the platform as a fundamental support tool in the management and multi-source sharing of different 3D models and images, related to heritage artifacts all over the world.

Then, additional 2D/3D semi-automatic annotation transfer methods were explored, so as to provide a future user with a range of possible methods to be chosen as most suitable for data transfer, according to the case-specific needs and requirements for semantic data enrichment (shifting, e.g., from point clouds and meshes to images, UV maps, ortho-photos and vice versa). The proposed workflows further contribute to a better understanding of complex heritage architectures, as well as to collaborative data sharing and information management of digital replicas of monuments and sites.

However, it has to be noted their application mostly referred, in this work, to photogrammetric surveys. Other alignment or annotation transfer scenarios, for instance involving the use of laser scans, would deserve further investigations.

Moreover, in the presented scenarios, the manual annotation of supervised (*training*) data is always performed by the user on a single support, and hybrid manual annotation modes are not yet considered.

8.1.3. H-BIM reconstruction, semantic structuring and insertion of localized information.

At a later stage, the work demonstrated the importance of leveraging semantically annotated and structured data for the construction of BIM-based representations.

In cases where the classification process relied on the identification of recurring and relevant architectural components, the semantically segmented point clouds were later considered for reconstruction of a reference model made out of template geometries, with the logic of H-BIM information systems (phase *iii*).

A semi-automated procedure was thus proposed for the Scan-to-BIM

reconstruction, i.e., to enable the reconstruction of Heritage-Building Information Models starting from annotated 3D survey data.

By this approach, the geometric nature of the building components was reconstructed, the originally designed shape was interpreted and a reference geometry was identified and modeled for each class, following the logic of the parametric BIM families. Upon completion of the process, visual programming techniques allowed the propagation of *template* elements. The reconstructed H-BIM model preserves the semantic link with the semantically annotated point cloud, at the level of the single classes of detected architectural components, and can be leveraged for further enrichment with non-geometric (analytical, knowledge-related) information. However, it is noted that the annotation process, although being trivial when referred to the single architectural component, becomes tricky in cases of localized annotations.

Based on this observation, the last part of the workflow introduced additional possibilities of semantic structuring and transfer of localized information, relying on the definition of suitable tiling procedures of the *template* model geometries. These latter experiments were aimed at noting, *inter alia,* the presence of frescoes and decorative parts, degradation phenomena, crack patterns, repair and restoration intervention and so on.

8.2. Research perspectives

During the drafting and development of this thesis, several problems and questions came out. The research perspectives, that could be the subject of further investigation together with the already mentioned work on localized annotations (Section 7.6 about 'Hints to rules of association between *template* model – *real* model'), are presented in this last Section.

8.2.1. Definition of unique development environment and interface

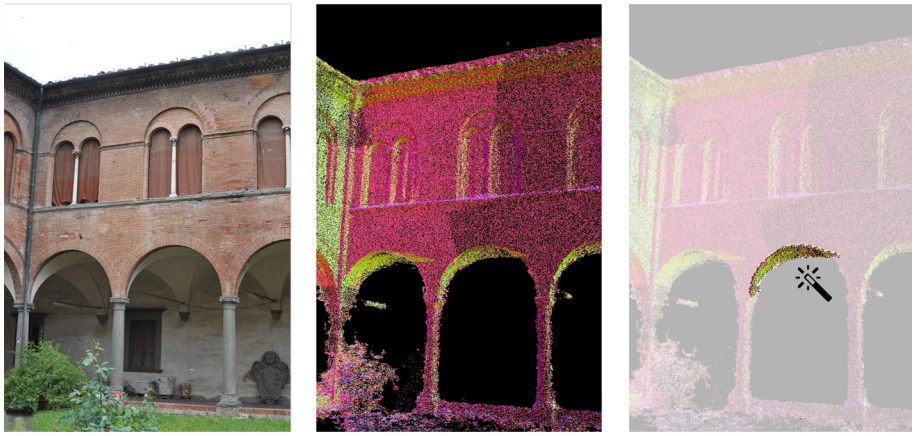
At present, a general algorithm implementing all the steps of the proposed methodology is not yet available, therefore several software programs and tools have been combined, each one specialized to perform a specific task. Indeed, the goal of this work was to test the efficiency of the proposed approach from a methodological point of view, in order to satisfy the variety of processes envisaged. Existing and already available software tools were thus tested, manipulated and adapted to fulfill new needs of description and documentation with a more direct and practical answer to the exigences of insertion, transfer and semi-automatic archival of semantic annotations on digital heritage models. Despite the complication derived by the use of multiple software, the procedure itself has been validated, demonstrating that the process can be generally applied.

The results of the proposed methodology undoubtedly suggest, as a forthcoming development of the research, the creation of a dedicated interface as consistent as possible that minimizes the use of different software, in which the different phases of the proposed methodological approach could be reunited and made available to the community of heritage experts.

8.2.2. Use of hybrid 2D/3D descriptors to improve the manual annotation phase within Aïoli

Even considering the above-mentioned research perspective (Section 9.2.1), the Aïoli platform could be implemented with further tools enabling at the same time hybrid 2D/3D annotation methods on multiple supports, template geometries reconstruction and/or additional semi-automatic annotation possibilities.

In particular, implementations that are planned to be applied on the platform in the very near future involve the following aspects:



a. b. c.

Figure 8-1.

Possible advances on hybrid 2D/3D annotation mode: the selection of a region corresponding to the 'arc' and 'vault' classes on an input image (*a*) is performed over its descriptor (*b*) via a magic wand tool (*c*).

The example refers to the *'Normal'* descriptor, that can be extracted in Aïoli from an input image (Manuel, 2016).

development of tools and methods supporting the user in the 2D/3D manual annotation phase (Paragraph 8.2.2.1); use of the annotated images to train deep neural networks and build libraries of segmented images for the built heritage (Paragraph 8.2.2.2).

8.2.2.1. Tools and methods supporting the user in the 2D/3D manual annotation phase.

Since Aïoli currently allows the manual annotation solely on the images and not in 3D, the development of combined hybrid 2D/3D segmentation approaches is envisaged, so to take full advantage of each annotation medium, in such a way that a possible user can perform the annotation either in 2D (e.g., on the image) or in 3D (e.g., on the point cloud), choosing which support is the easiest one to annotate.

Moreover, it is thought in this regard that geometric or radiometric features (descriptors) associated with data – both 2D and 3D – could be used to ease the annotation of objects in a semi-automatic and supervised way. For example, in the case of images, this could result in the development of a sort of *magic wand* tools, assisting the selection of regions of interest in which a certain descriptor appears with homogeneous tones and colors (Figure 8-1). For a list of image-point cloud descriptors that can currently be extracted from Aïoli, the reader can refer to the work by Manuel (2016).

8.2.2.2. Use of the annotated images via Aïoli to train Deep Neural Networks and build libraries of segmented images for the built heritage.

In this work, we explored the possibility to transfer the semantic segmentation information from the point cloud to the set of existing images, by exploiting the photogrammetric processing and the 2D/3D hybrid propagation function (Section 6.1.5). As a future development of

this work, if a consistent number of annotated images is provided, the latter could be used in view of the construction of libraries of semantically segmented images, to train Deep Neural Networks in recognizing, e.g., architectural components on a set of images of existing buildings.

8.2.3. Improvement of existing benchmarks for the construction of Deep Neural Networks

Still taking into account the developments toward the construction of Deep Neural Networks, the semantically annotated 3D datasets presented in this thesis could be used to implement existing benchmarks of annotated digital heritage objects and sites, as the 3DArCH one (Matrone et al., 2020). Indeed, the construction of a significant volume of annotated 3D data (especially, point clouds) could improve the application of DL strategies to the heritage domain, to date not yet applicable due to the limited amount of available training information.

8.2.4. Levels of detail within H-BIM environments and integration of localized information

Regarding the Scan-to-BIM reconstruction phase, future developments of this research have been already introduced in Chapter 8, in terms of insertion of localized information within the reconstructed *template* geometries.

In addition to this, it has to be noted that successive levels of classification might also be included as part of the Scan-to-BIM reconstruction approach. The multi-level hierarchical segmentation method, which has been proposed here with reference to point clouds, could indeed be exploited for the construction of template models, at different levels of detail and resolution.

Future advances of the present work may thus foresee the application of the Scan-to-BIM reconstruction approach even to the multi-level case, i.e., considering successive levels of semantic classification of the 3D model.

8.2.5. Extension to other case studies

The principles introduced in this work for the insertion, organization and transfer of semantic information have been tested only to the case of buildings whose architectural composition is characterized by recognizable and recurring elements. The identification of typological classes, indeed, is a fundamental premise for the ML-driven

decomposition of the building.

To this concern, several issues may probably arise from the application of this approach to other types of artifacts – e.g., sculptures, archaeological remains, paintings – and more diversified buildings – e.g., referred to different construction types or built in different periods. The extension of the proposed approach to other case studies is finally a possible development of the present thesis.

Chapter 9

Bibliography

Abdal Noor, B., Yi, S., 2018. Review of BIM literature in construction industry and transportation: meta-analysis. *Construction Innovation*, 18, 4, pp. 433-452. doi.org/10.1108/CI-05-2017-0040.

Abergel, V., 2020. Relevé numérique d'art pariétal : définition d'une approche innovante combinant propriétés géométriques, visuelles et sémantiques au sein d'un environnement de réalité mixte. Synthèse d'image et réalité virtuelle (Ph.D. thesis). HESAM Université. Français. NNT: 2020HESAE02.

Abergel V., Jacquot K., De Luca L., Véron P., 2021. Combining on-site and off-site analysis: towards a new paradigm for cultural heritage surveys. *DISEGNARECON*, 2021, 14 (26), doi. org/10.20365/disegnarecon.26.2021.13.

Achille, C., Adami, A., Chiarini, S., Cremonesi, S., Fassi, F., Fregonese, L., 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, 15520–15539. doi.org/10.3390/s150715520.

Aicardi, I., Chiabrando, F., Lingua, A.M., Noardo, F., 2018. Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach. *Journal of Cultural Heritage* 32, 257–266. doi.org/10.1016/j.culher.2017.11.006.

Aitelkadi, K., Tahiri, D., Simonetto, E., Sebari, I., Boulaassal, H., 2014. Automatic Extraction of Facade Details of Heritage Building Using Terrestrial Laser Scanning Data. *J. Archit. Eng. Tech.* 3(133). doi.org/10.4172/2168-9717.1000133.

Al-Amri, S.S. and Kalyankar, N.V., 2010. Image segmentation by using threshold techniques. *Journal of Computing* 2(5), 83-86. ISSN 2151-9617.

Andrews, P., Zaihrayeu, I., Pane, J., 2012. A classification of semantic annotation systems. *Semantic Web* 3, 223–248. doi.org/10.3233/SW-2011-0056.

Andriasyan, M., Moyano, J., Nieto-Julián, J.E., Antón, D., 2020. From Point Cloud Data to Building Information Modelling: An Automatic Parametric Workflow for Heritage. *Remote Sensing* 12, 1094. doi.org/10.3390/rs12071094.

Angulo-Fornos, R., Castellano-Román, M., 2020. HBIM as Support of Preventive Conservation Actions in Heritage Architecture. Experience of the Renaissance Quadrant Façade of the Cathedral of Seville. *Applied Sciences* 10, 2428. doi.org/10.3390/app10072428.

Apollonio, F.I., Basilissi, V., Callieri, M., Dellepiane, M., Gaiani, M., Ponchio, F., Rizzo, F., Rubino, A.R., Scopigno, R., Sobra', G., 2018. A 3D-centered information system for the documentation of a complex restoration intervention. *Journal of Cultural Heritage* 29, 89–99. doi.org/10.1016/j.culher.2017.07.010.

Apollonio, F.I., Gaiani, M., Sun, Z., 2012. BIM-based Modeling and Data Enrichment of Classical Architectural Buildings. *SCIRES-IT* 2. doi.org/10.2423/i22394303v2n2p41.

Arayici, Y., 2008. Towards building information modelling for existing structures. *Structural Survey*, 26, 210–222. doi.org/10.1108/02630800810887108.

Arbelaez, P., Maire, M., Fowlkes, C. and Malik, J., 2010. Contour detection and hierarchical image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 33(5), 898-916.

Arganda-Carreras, I., Kaynig, V., Schindelin, J., Cardona, A., Seung, H.S., 2017. Trainable Weka Segmentation: A Machine Learning Tool for Microscopy Image Segmentation. *Bioinformatics*, 33 (15), 2424–2426. doi.org/10.1093/bioinformatics/btx180.

Attene, M., Robbiano, F., Spagnuolo, M., Falcidieno, B., 2009. Characterization of 3D shape parts for semantic annotation. *Computer-Aided Design,* 41, 756–763. doi.org/10.1016/j.cad.2009.01.003.

Aveta, C., Salvatori, M., Vitelli, G.P., 2017. The complex point cloud for the knowledge of the architectural heritage. Some experiences. *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* 42W4, 235–244. doi.org/10.5194/isprs-archives-XLII-5-W1-235-2017.

Bacci, G., Bertolini, F., Bevilacqua, M.G., Caroti, G., Martínez-Espejo Zaragoza, I., Martino, M., Piemonte, A., 2019. HBIM methodologies for the architectural restoration. The case of the ex-church of San Quirico all'Olivo in Lucca, Tuscany. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2/W11, 121–126. doi.org/10.5194/isprs-archives-XLII-2-W11-121-2019.

Baik, A., 2017. From point cloud to Jeddah Heritage BIM Nasif Historical House – case study. *Digital Applications in Archaeology and Cultural Heritage* 4, 1–18. doi.org/10.1016/j.daach.2017.02.001.

Bandinelli, A., Bevilacqua, M.G., Karwacka, E., 2016. Il disegno di rilievo come strumento di indagine. Il caso della Certosa monumentale di Calci/The architectural survey drawing as an investigation instrument. The case of Calci Charterhouse. In: Bertocci S., Bini, M. (Eds.). Le ragioni del disegno. Pensiero, forma e modello nella gestione della complessità. Atti del 38° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/The reasons of drawing. Thought, shape and model in the complexity management. Proceedings of the 38th International Conference of Representation Disciplines Teachers. Roma: Gangemi Editore, 63-68.

Barazzetti, L., Banfi, F., Brumana, R., Gusmeroli, G., Previtali, M., Schiantarelli, G., 2015. Cloud-to-BIM-to-FEM: Structural simulation with accurate historic BIM from laser scans. *Simulation Modelling Practice and Theory* 57, 71–87. doi.org/10.1016/j.simpat.2015.06.004.

Bassier, M., Vergauwen, M., Van Genechten, B., 2017. Automated classification of heritage buildings for as-built BIM using machine learning techniques. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* IV-2/W2, 25–30. doi.org/10.5194/isprs-annals-IV-2-W2-25-2017.

Bassier, M., Yousefzadeh, M., Vergauwen, M., 2020. Comparison of 2D and 3D wall reconstruction algorithms from point cloud data for as-built BIM. *Journal of Information Technology in Construction ITcon* 25, 173–192. doi.org/10.36680/j.itcon.2020.011.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114, 24–31. doi.org/10.1016/j.isprsjprs.2016.01.011.

Belton, D., Lichti, D.D., 2006. Classification and segmentation of terrestrial laser scanner point clouds using local variance information. ISPRS- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(5), pp.44-49.

Beraldin, J.-A., Blais, F., El-Hakim, S.F., Cournoyer, L., Picard, M., 2007. Traceable 3D imaging metrology: evaluation of 3D digitizing techniques in a dedicated metrology laboratory. *Optical 3D Measurement Techniques* 25, 310-318. doi.org/10.36680/j.itcon.2020.011.

Besl, P. J., Jain, R. C., 1988. Segmentation through variable-order surface fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3 (10), 167-192.

Bevilacqua, M.G., Caroti, G., Piemonte, A., Ruschi, P., Tenchini, L., 2017. 3D survey techniques for the architectural restoration: the case of St. Agata in Pisa. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-5/W1, 441–447. doi.org/10.5194/isprs-archives-XLII-5-W1-441-2017.

Bevilacqua, M.G., Caroti, G., Piemonte, A., Terranova, A.A., 2018. Digital Technology and Mechatronic Systems for the Architectural 3D Metric Survey, in: Ottaviano, E., Pelliccio, A., Gattulli, V. (Eds.), Mechatronics for Cultural Heritage and Civil Engineering, Intelligent Systems, Control and Automation: Science and Engineering. Springer International Publishing, Cham, pp. 161–180. doi.org/10.1007/978-3-319-68646-2_7.

Bezoari, G., Selvini, A., 1999. Gli strumenti per la fotogrammetria. Storia e tecnica. Liguori Ed.

Bianconi, F., Filippucci, M., Magi Meconi, F., 2018. Parametrical Vitruvius. Generative modelling of the architectural orders. *SCIRES-IT* 8, 29–48. doi.org/10.2423/i22394303v8n2p29.

Bogdan, R. R. and Cousins, S., 2011. 3D is here: Point Cloud Library (PCL). 2011 *IEEE International Conference on Robotics and Automation*, 1-4, doi: 10.1109/ICRA.2011.5980567.

Boehler, W., Vicent, M.B., Marbs, A., 2003. Investigating laser scanner accuracy. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIV(5)/W15, 696–701. doi.org/10.1.1.485.5642.

Bolkas, D., 2019. Terrestrial laser scanner intensity correction for the incidence angle effect on surfaces with different colours and sheens. *International Journal of Remote Sensing* 40, 7169–7189. doi.org/10.1080/01431161.2019.1601283.

Bosché, F., Ahmed, M., Turkan, Y., Haas, C.T., Haas, R., 2015. The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction* 49, 201–213. doi.org/10.1016/j.autcon.2014.05.014.

Breiman, L., 2001. Ramdom forests. *Machine Learning* 45, 5–32.

Brumana, R., Georgopoulos, A., Oreni, D., Raimondi, A., Bregianni, A., 2013. HBIM for Documentation, Dissemination and Management of Built Heritage. The Case Study of St. Maria in Scaria d'Intelvi. *International Journal of Heritage in the Digital Era* 2, 433–451. doi.org/10.1260/2047-4970.2.3.433.

Brumana, R., Oreni, D., Barazzetti, L., Cuca, B., Previtali, M., Banfi, F., 2020. Survey and Scan to BIM Model for the Knowledge of Built Heritage and the Management of Conservation Activities, in: Daniotti, B., Gianinetto, M., Della Torre, S. (Eds.), Digital Transformation of the Design, Construction and Management Processes of the Built Environment. Springer International Publishing, Cham, pp. 391–400. doi.org/10.1007/978-3-030-33570-0_35.

Bruno, N., Roncella, R., 2019. HBIM for Conservation: A New Proposal for Information Modeling. *Remote Sensing* 11, 1751. doi.org/10.3390/rs11151751.

Bruno, S., De Fino, M., Fatiguso, F., 2018. Historic Building Information Modelling: performance assessment for diagnosis-aided information modelling and management. *Automation in Construction* 86, 256–276. https://doi.org/10.1016/j.autcon.2017.11.009.

Bruno, S., Musicco, A., Fatiguso, F., Dell'Osso, G.R., 2019. The role of 4D historic Building Information Modelling and management in the analysis of constructive evolution and decay condition within the refurbishment process. *International Journal of Architectural Heritage* 1–17. doi.org/10.1080/15583058.2019.1668494.

Burresi M., Caleca A., 2003. *Affreschi medievali a Pisa*, Pacini Editore (Eds.), Pisa.

Capone, M., Lanzara, E., 2019. Scan-to-BIM vs 3D ideal model. HBIM: Parametric tools to study domes geometry. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-2/W9, 219–226. doi.org/10.5194/isprs-archives-XLII-2-W9-219-2019.

Caputo, M., Denker, K. and Franz, M. O., 2015. (2015) Support Vector Machines for Classification of Geometric Primitives in Point Clouds. In: Boissonnat JD. et al. (Eds.) Curves and Surfaces. Curves and Surfaces 2014. Lecture Notes in Computer Science, 9213, 80-95. Springer. doi.org/10.1007/978-3-319-22804-4_7 S

Carboni, N., 2019. An Ontological Approach to the Description of Visual and Iconographical Representations. *Heritage* 2, 1191–1210. doi.org/10.3390/heritage2020078.

Caroti, G., Martínez-Espejo Zaragoza, I., Piemonte, A., 2015. Range and image based modelling: a way for frescoed vault texturing optimization. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XL-5/W4, 285–290. doi.org/10.5194/isprsarchives-XL-5-W4-285-2015.

Caroti, G., Piemonte, A., 2020. Integration of Laser Scanning and Photogrammetry in Architecture Survey. Open Issue in Geomatics and Attention to Details, in: Parente, C., Troisi, S., Vettore, A. (Eds.), Geomatics: Research, Results and Review, Communications in Computer and Information Science. Springer International Publishing, Cham, pp. 170–185. doi.org/10.1007/978-3-030-62800-0_14.

Ceccarelli, L., 2021. Applicativi di segmentazione semantica nel processo from Scan to BIM: Il caso studio della Chiesa di Santa Caterina d'Alessandria a Pisa (M.Sc. thesis). Corso di Laurea Magistrale in Ingegneria Edile Architettura, Università di Pisa.

Cera, V., 2018. Il rilievo dell'Architettura tra identificazione tipologica e strutturazione semantica. La Certosa di San Lorenzo a Padula nella rappresentazione digitale per il Cultural Heritage. (Ph.D. thesis), fedOA Università degli Studi di Napoli Federico II.

Che, E. and Olsen, M. J., 2017. Fast Edge Detection for Terrestrial Laser Scanning through Normal Variation Analysis. *ISPRS Annals Photogramm. Remote Sens. Spatial Inf. Sci.*, 9, IV-2/W4, 51-57. doi.org/10.5194/isprs-annals-IV-2-W4-51-2017

Chehata, N., Guo, L. and Mallet, C., 2009. Airborne lidar feature selection for urban classification using random forests. *Laserscanning*, Sep. 2009, Paris, France.

Chiabrando, F., Colucci, E., Lingua, A., Matrone, F., Noardo, F., Spanò, A., 2018. A European interoperable database (EID) to increase resilience of Cultural Heritage. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-3/W4, 151–158. doi.org/10.5194/isprs-archives-XLII-3-W4-151-2018.

Chew, L., 1993. Guaranteed-quality mesh generation for curved surfaces. in: Proc. of the 9th annual symposium on Computational geometry SGC '93. San Diego, California, USA, 274280.

Chiabrando, F., Lo Turco, M., Rinaudo, F., 2017. Modeling the decay in a HBIM starting from 3D point clouds. A followed approach for Cultural Heritage knowledge. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2/W5, 605–612. doi.org/10.5194/isprs-archives-XLII-2-W5-605-2017.

Condorelli, F., Rinaudo, F., Salvadore, F., Tagliaventi, S., 2021. A comparison between 3D reconstruction using NeRF neural networks and MVS algorithms on Cultural Heritage images. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLIII-B2-2021, 565–570. doi.org/10.5194/isprs-archives-XLIII-B2-2021-565-2021.

Conrat J.M., Pajusco P., Dunand A., 2010. Utilisation des photos panoramiques pour la compréhension des phénomènes physiques du canal de propagation, *Revue Générale de l'Electricité*, Société de l'Électricité, de l'Électronique et des Technologies de l'Information et de la Communication, 2, 73- 82. hal-00460795

Croce, V., Caroti, G., Piemonte, A., 2019. Assessment of earthquake-induced damage level on buildings: analysis of two different survey methods for a case study. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2/W15, 351–358. doi.org/10.5194/isprs-archives-XLII-2-W15-351-2019.

Croce V., Caroti G., Piemonte A., Bevilacqua M.G.B. 2021. From survey to semantic representation for Cultural Heritage: the 3D modeling of recurring architectural elements. *e-journal ACTA Imeko* 10 (1). doi.org/10.21014/acta_imeko.v10i1.842

Croce, V., Bevilacqua, M.G., Caroti, G., Piemonte, A., 2021. Connecting geometry and semantics via Artificial Intelligene: from 3D classification of heritage data to H-BIM representations. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLIII-B2-2021, 145–152. doi.org/10.5194/isprs-archives-XLIII-B2-2021-145-2021.

Croce, V., Caroti, G., De Luca, L., Jacquot, K., Piemonte, A., Véron, P., 2021a. From the Semantic Point Cloud to Heritage-Building Information Modeling: A Semiautomatic Approach Exploiting Machine Learning. *Remote Sensing* 13, 461. doi.org/10.3390/rs13030461.

Croce, V., Caroti, G., De Luca, L., Piemonte, A., Véron, P., 2020. Semantic annotations on heritage models: 2D/3D approaches and future research challenges. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLIII-B2-2020, 829–836. doi.org/10.5194/isprs-archives-XLIII-B2-2020-829-2020.

Croce, Valeria, Caroti, G., Piemonte, A., 2021b. Propagation of semantic information between orthophoto and 3D replica: a H-BIM system for the north transept of Pisa Cathedral. *Geomatics, Natural Hazards and Risk* 12, 2225–2252. doi.org/10.1080/19475705.2021.1960432.

Croce V., Caroti G., De Luca L., Piemonte A., Véron P., Bevilacqua M.G., 2021. Tra Intelligenza Artificiale e H-BIM per la descrizione semantica dei beni culturali: la Certosa di Pisa/Artificial Intelligence and H-BIM for the semantic description of cultural heritage: the Pisa Charterhouse. in: Arena A., Arena M., Mediati D., Raffa P. (Eds.). Connettere. Un disegno per annodare e tessere. Linguaggi Distanze Tecnologie. Atti del 42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/Connecting. Drawing for weaving relationship. Languages Distances Technologies. Proceedings of the 42nd International Conference of representation Disciplines Teachers. Milano: FrancoAngeli, 608-625.

Croce, V., Martínez-Espejo Zaragoza, I., 2018. UAV-based 3D photogrammetry for post-earthquake studies on seismic damaged cities – A case study: Castelluccio di Norcia. *Proc. of International Conference on Education and Information Systems, Technologies and Applications* EISTA 2018 Orlando (USA).

Dastres, R., Soori, M., 2021. Artificial Neural Network Systems. I*nternational Journal of Imaging and Robotics (IJIR)* 21 (2), 13-25. hal-03349542f

De Luca, L., 2006. Relevé et multi-représentations du patrimoine architectural Définition d'une approche hybride pour la reconstruction 3D d'édifices (Ph.D. thesis). Sciences de l'Homme et Société. Arts et Métiers Paris-Tech. Français. NNT: 2006ENAM0009.

De Luca, L., 2011. Methods, formalisms and tools for the semantic-based surveying and representation of architectural heritage. *Appl Geomat.* doi.org/10.1007/s12518-011-0076-7.

De Luca, L., Busayarat, C., Stefani, C., Véron, P., Florenzano, M., 2011. A semantic-based platform for the digital analysis of architectural heritage. *Computers & Graphics* 35, 227–241. doi.org/10.1016/j.cag.2010.11.009.

De Luca, L., Lo Buglio, D., 2014. Geometry vs Semantics: Open Issues on 3D Reconstruction of Architectural Elements, in: Ioannides, M., Quak, E. (Eds.), 3D Research Challenges in Cultural Heritage, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, 36–49. doi.org/10.1007/978-3-662-44630-0_3.

De Luca, L., Véron, P., Florenzano, M., 2007. A generic formalism for the semantic modeling and representation of architectural elements. *Visual Comput* 23, 181–205. doi.org/10.1007/s00371-006-0092-5.

De Luca, L., Véron, P., Florenzano, M., 2006. Reverse engineering of architectural buildings based on a hybrid modeling approach. *Computers & Graphics* 30, 160–176. doi.org/10.1016/j.cag.2006.01.020.

Docci, M., Chiavoni, E., 2017. Saper leggere l'architettura, Laterza Ed., Bari.

Dore, C., Murphy, M., 2017. Current state of the art Historic Building Information Modeling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII–2, 185–192. doi.org/10.5194/isprs-archives-XLII-2-W5-185-2017.

Dore, C., Murphy, M., 2013. Semi-automatic techniques for as-built BIM façade modeling of historic buildings, in: Proc. of the 2013 Digital Heritage International Congress, IEEE, Marseille, France, 473–480. doi.org/10.1109/DigitalHeritage.2013.6743786.

Eastman, C., Teicholz, P., Sacks, R. & Liston, K., 2008. BIM Handbook – A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers, and Contractors, John Wiley & Sons, Inc.

Eastman, C., 1975, The Use of Computers Instead of Drawings, *AIA Journal* 63(3), 46–50.

Fabiani, F., Grilli, R., Musetti, V., 2016. Verso nuove modalità di gestione e presentazione della documentazione di restauro: SICaR web la piattaforma in rete del Ministero dei Beni e delle Attività Culturali e del Turismo. *Bollettino Ingegneri,* Collegio Ingegneri della Toscana 3–13.

Fai, S., Rafeiro, J., 2014. Establishing an appropriate Level of Detail (LoD) for a Building Information Model (BIM) – West Block, Parliament Hill, Ottawa, Canada. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II–5, 123–130. doi.org/10.5194/isprsannals-II-5-123-2014.

Fai, S., Sydor, M., 2013. Building Information Modelling and the documentation of architectural heritage: Between the 'typical' and the 'specific,' in: Proc. of the 2013 Digital Heritage International Congress, IEEE, Marseille, France. 731–734. doi.org/10.1109/DigitalHeritage.2013.6743828

Fiorucci, M., Khoroshiltseva, M., Pontil, M., Traviglia, A., Del Bue, A., James, S., 2020. Machine Learning for Cultural Heritage: A Survey. *Pattern Recognition Letters* 133, 102–108. doi.org/10.1016/j.patrec.2020.02.017.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 24, 381–395. doi.org/10.1145/358669.358692.

Flusser J., 2002. On the inverse problem of rotation moment invariants, *Pattern Recognition*, 35(12), 3015-3017. doi.org/10.1016/S0031-3203(02)00093-6.

Fromont R., Trentesaux C., 2005. Observation visuelle des charpentes de la nef et du chœur - Relevés complémentaires et bilan sanitaire des charpentes de la Cathédrale Notre-Dame de Paris, France.

Garozzo, R., Murabito, F., Santagati, C., Pino, C., Spampinato, C., 2017. Cul-TO: an ontology-based annotation tool for data curation in Cultural Heritage. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-2/W5, 267–274. doi.org/10.5194/isprs-archives-XLII-2-W5-267-2017.

Georgopoulos, A., Stathopoulou, E.K., 2016. Data acquisition for 3D geometric recording: state of the art and recent innovations, 2017. In: Heritage and Archaeology in the Digital Age. Acquisition, Curation, and Dissemination of Spatial Cultural Heritage Data. Vincent, M.L., López-Menchero Bendicho, V.M., Ioannides, M., Levy, Th.E. (Eds.). ISBN 978-3-319-65370.

Gonizzi Barsanti, S., Guidi, G., De Luca, L., 2017. Segmentation of 3D models for Cultural Heritage structural analysis: some critical issues. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci*. IV-2/W2, 115–122. doi.org/10.5194/isprs-annals-IV-2-W2-115-2017.

Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep Learning. MIT Press.

Grieves, M. W., 2014. Digital Twin: Manufacturing Excellence through Virtual Factory Replication. Digital Twin white paper. LLC 2014

Griffiths, D., Boehm, J., 2019. A Review on Deep Learning Techniques for 3D Sensed Data Classification. *Remote Sensing* 11, 1499. doi.org/10.3390/rs11121499.

Grilli, E., 2020. Automatic classification of architectural and archaeological 3D Data (Ph.D. thesis). Alma Mater Studiorum – Università di Bologna, XXXII cycle.

Grilli, E., Dininno, D., Petrucci, G., Remondino, F., 2018. From 2D to 3D supervised segmentation and classification for Cultural Heritage applications. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII–2, 399–406. doi.org/10.5194/isprs-archives-XLII-2-399-2018.

Grilli, E., Farella, E.M., Torresani, A., Remondino, F., 2019a. Geometric features analysis for the classification of Cultural Heritage point clouds. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-2/W15, 541–548. doi.org/10.5194/isprs-archives-XLII-2-W15-541-2019.

Grilli, E., Menna, F., Remondino, F., 2017. A review of point clouds segmentation and classification algorithms. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-2/W3, 339–344. doi.org/10.5194/isprs-archives-XLII-2-W3-339-2017.

Grilli, E., Özdemir, E., Remondino, F., 2019b. Application of machine and deep learning strategies for the classification of heritage point clouds. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-4/W18, 447–454. doi.org/10.5194/isprs-archives-XLII-4-W18-447-2019.

Grilli, E., Remondino, F., 2019. Classification of 3D Digital Heritage. *Remote Sensing* 11, 847. doi.org/10.3390/rs11070847.

Grussenmeyer, P., Alby, E., Landes, T., Koehl, M., Guillemin, S., Hullo, J.F., Assali, P., Smigiel, E., 2012. Recording approach of heritage sites from high resolution photogrammetry and terrestrial laser scanning. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XXXIX-B5, 553–558. doi.org/10.5194/isprsarchives-XXXIX-B5-553-2012.

Grussenmeyer, P., Landes, T., Voegtle, T., Ringle, K., 2008. Comparison methods of terrestrial laser scanning, photogrammetry and tacheometry data for recording of Cultural Heritage buildings. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XXXVI, 213–318.

Guidi, G., Remondino, F., 2012. 3D Modelling from real data, in: Alexandru, C. (Ed.), Modeling and Simulation in Engineering. InTech. doi.org/10.5772/30323.

Hackel, T., Wegner, J.D., Savinov, N., Ladicky, L., Schindler, K., Pollefeys, M., 2018. Large-Scale Supervised Learning For 3D Point Cloud Labeling: Semantic3d.Net. *Photogramm. Eng. Remote Sensing* 84, 297–308. doi.org/10.14358/PERS.84.5.297.

Hackel, T., Wegner, J.D., Schindler, K., 2016a. Fast semantic segmentation of 3D point clouds with strongly varying density, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.,* 177-184. doi.org/10.5194/lSPRS-ANNALS-III-3-177-2016.

Hackel, T., Wegner, J.D., Schindler, K., 2016b. Contour Detection in Unstructured 3D Point Clouds, in: Proc. of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, NV, USA, 1610–1618. doi.org/10.1109/CVPR.2016.178.

Han, X.-F., Jin, J.S., Wang, M.-J., Jiang, W., Gao, L., Xiao, L., 2017. A review of algorithms for filtering the 3D point cloud. *Signal Processing: Image Communication* 57, 103–112. doi.org/10.1016/j.image.2017.05.009.

Hassaballah, M., Abdelmgeid, A.A., Alshazly, H.A., 2016. Image features detection, description and matching, in: Awad, A.I., Hassaballah, Mahmoud (Eds.), Image Feature Detectors and Descriptors, Studies in Computational Intelligence. Springer International Publishing, Cham, pp. 11–45. https://doi.org/10.1007/978-3-319-28854-3_2.

Havemann, S., Settgast, V., Berndt, R., Eide, Ø., Fellner, D., 2009. The arrigo showcase reloaded—towards a sustainable link between 3D and semantics. *Journal on Computing and Cultural Heritage* 2 (1), 113. doi.org/10.1145/1551676.1551680.

Heckbert, P.S., 1986. Survey of texture mapping. *IEEE Computer Graphics and Applications* 6, 56–67. doi.org/10.1109/MCG.1986.276672.

Hichri, N., Stefani, C., De Luca, L., Véron, P., Hamon, G., 2013. From point cloud to BIM: a survey of existing approaches. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XL-5/W2, 343–348. doi.org/10.5194/isprsarchives-XL-5-W2-343-2013.

Hong, S., Jung, J., Kim, S., Cho, H., Lee, J., Heo, J., 2015. Semi-automated approach to indoor mapping for 3D as-built building information modeling. *Computers, Environment and Urban*

Systems 51, 34–46. doi.org/10.1016/j.compenvurbsys.2015.01.005.

Hough, P V.C., 1962. Tue, Method and Means for Recognizing Complex Patterns. USA.

Hu M.K., 1962. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory* 8 (2), 179-187. doi.org/10.1109/TIT.1962.1057692.

Ibrahim, Y., Nagy, B., Benedek, C., 2020. Deep Learning-Based Masonry Wall Image Analysis. *Remote Sensing* 12, 3918. doi.org/10.3390/rs12233918.

Jacquot, K., 2014. Numérisation et restitution virtuelle des maquettes de la collection de Louis XIV, le cas des fortifications bastionnées (Ph.D. thesis). Architecture, aménagement de l'espace. Université de Lorraine. Français. NNT: 2014LORR0010.

Jung, J., Hong, Sungchul, Jeong, S., Kim, S., Cho, H., Hong, Seunghwan, Heo, J., 2014. Productive modeling for development of as-built BIM of existing indoor structures. *Automation in Construction* 42, 68–77. https://doi.org/10.1016/j.autcon.2014.02.021.

Kar, M.K., Nath, M.K., Neog, D.R., 2021. A review on progress in semantic image segmentation and its application to medical images. *SN Computer Science* 2, 397. doi.org/10.1007/s42979-021-00784-5.

Kashani, A., Olsen, M., Parrish, C., Wilson, N., 2015. A Review of LIDAR Radiometric Processing: From Ad Hoc Intensity Correction to Rigorous Radiometric Calibration. Sensors 15, 28099–28128. https://doi.org/10.3390/s151128099.

Kasser, M., Egels, Y., 2001. Digital photogrammetry. Taylor & Francis, London - New York. doi.org/10.1201/9780203305959.

Kavzoglu, T. and Colkesen, I., 2009. A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation* 11(5), 352-359. doi.org/10.1016/j.jag.2009.06.002.

Kazhdan, M., Hoppe, H., 2013. Screened poisson surface reconstruction. *ACM Transactions on Graphics* 32(3), 113. doi.org/10.1145/2487228.2487237.

Kaur, J., Agrawal, S., Vig, R., 2012. A comparative analysis of thresholding and edge detection segmentation techniques. *International Journal of Computer Applications* 39(2), 29-34.

Kelly, G., 2006. A survey of procedural techniques for city generation. *The ITB Journal*, 7. doi.org/10.21427/D76M9P.

Klasing, K., Althoff, D., Wollherr, D., Buss, M., 2009. Comparison of surface normal estimation methods for range sensing applications. in: Proc. of the 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan. 3206-3211. doi: 10.1109/ROBOT.2009.5152493.

Korc, F., Forstner, W., 2009. eTRIMS Image Database for Interpreting Images of Man-Made Scenes 12.

Kuçak, R.A., Özdemir, E., Erol, S., 2017. The segmentation of point clouds with K-means and ANN (artificial neural networks). *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-1/W1, 595–598. doi.org/10.5194/isprs-archives-XLII-1-W1-595-2017.

Lavoué, G., Dupont, F. and Baskurt, A., 2005. A new CAD mesh segmentation method, based on curvature tensor analysis. *Computer-Aided Design* 37(10), 975-987.

LeCun, Y., Bengio, Y., Hinton, G, 2015. Deep learning. *Nature* 521(7553), 436–444.

Lo Turco, M., Mattone, M., Rinaudo, F., 2017. Metric survey and BIM technologies to record decay conditions. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-5/W1, 261–268. doi.org/10.5194/isprs-archives-XLII-5-W1-261-2017.

Logothetis, S., Delinasiou, A., Stylianidis, E., 2015. Building Information Modelling for Cultural Heritage: A review. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-5/W3, 177–183. doi.org/10.5194/isprsannals-II-5-W3-177-2015.

López, F., Lerones, P., Llamas, J., Gómez-García-Bermejo, J., Zalama, E., 2018. A Review of Heritage Building Information Modeling (H-BIM). *MTI* 2, 21. doi.org/10.3390/mti2020021.

López, F.J., Lerones, P.M., Llamas, J., Gómez-García-Bermejo, J., Zalama, E., 2017. A framework for using point cloud data of heritage buildings toward geometry modeling in a BIM context: a case study on Santa Maria La Real De Mave Church. *International Journal of Architectural Heritage* 1–22. doi.org/10.1080/15583058.2017.1325541.

Macher, H., Landes, T., Grussenmeyer, P., 2017. From Point Clouds to Building Information Models: 3D Semi-Automatic Reconstruction of Indoors of Existing Buildings. *Applied Sciences* 7, 1030. doi.org/10.3390/app7101030.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. in: Proc. of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, California, USA. 281-297.

Malinverni, E.S., Pierdicca, R., Paolanti, M., Martini, M., Morbidoni, C., Matrone, F., Lingua, A., 2019. Deep learning for semantic segmentation of 3D point cloud. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2/W15, 735–742. doi.org/10.5194/isprs-archives-XLII-2-W15-735-2019.

Manferdini, A., Remondino, F., 2010. Reality-Based 3D Modeling, Segmentation and Web-Based Visualization. in: Ioannides, M., Fellner, D., Georgopoulos, A., Hadjimitsis, D. (Eds.), Digital Heritage. Berlin, Heidelberg: Springer, 110124. ISBN 978-3-642-16872-7.

Manfredi, M., Grana, C., Cucchiara, R., 2013. Automatic Single-Image People Segmentation and Removal for Cultural Heritage Imaging, in: Petrosino, A., Maddalena, L., Pala, P. (Eds.), New Trends in Image Analysis and Processing – ICIAP 2013. Springer Berlin Heidelberg, Berlin, Heidelberg, 188–197.

Manuel, A., 2016. Annotation sémantique 2D/3D d'images spatialisées pour la documentation et l'analyse d'objets patrimoniaux (Ph.D. thesis). Génie des procédés. Ecole nationale supérieure d'arts et métiers- ENSAM. Français. NNT: 2016ENAM0009.

Manuel, A., Gattet, E., De Luca, L., Véron, P., 2013a. An approach for precise 2D/3D semantic annotation of spatially-oriented images for in situ visualization applications, in: Proc. of the 2013 Digital Heritage International Congress, IEEE, Marseille, France, 289–296. doi.org/10.1109/DigitalHeritage.2013.6743752.

Manuel, A., Stefani, C., De Luca, L., Véron, P., 2013b. Distribution of semantic annotation towards a set of spatially oriented photographs. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XL-5/W1, 175–180. doi.org/10.5194/isprsarchives-XL-5-W1-175-2013.

Manuel, A., Véron, P., De Luca, L., 2016. 2D/3D Semantic Annotation of Spatialized Images for the Documentation and Analysis of Cultural Heritage, in: Proc. of the 14th Eurographics Workshop on Graphics and Cultural Heritage (GCH '16). Eurographics Association, Goslar, DEU, 101–104. doi.org/10.2312/GCH.20161391.

Marengo, E., Manfredi, M., Zerbinati, O., Robotti, E., Mazzucco, E., Gosetti, F., Bearman, G., France, F., Shor, P., 2011. Development of a technique based on multi-spectral imaging for monitoring the conservation of cultural heritage objects. *Analytica Chimica Acta* 706, 229–237. doi.org/10.1016/j.aca.2011.08.045.

Martinez Espejo Zaragoza, I., Caroti, G., Piemonte, A., 2021. The use of image and laser scanner survey archives for cultural heritage 3D modelling and change analysis. *ACTA IMEKO* 10, 114. doi.org/10.21014/acta_imeko.v10i1.847.

Martínez-Espejo Zaragoza, I., Caroti, G., Piemonte, A., Riedel, B., Tengen, D., Niemeier, W., 2017. Structure from motion (SfM) processing of UAV images and combination with terrestrial laser scanning, applied for a 3D-documentation in a hazardous situation. *Geomatics, Natural Hazards and Risk* 8, 1492–1504. doi.org/10.1080/19475705.2017.1345796.

Matrone, F., Grilli, E., Martini, M., Paolanti, M., Pierdicca, R., Remondino, F., 2020. Comparing Machine and Deep Learning Methods for Large 3D Heritage Semantic Segmentation. *IJGI* 9,

535. doi.org/10.3390/ijgi9090535.

Matrone, F., Lingua, A., Pierdicca, R., Malinverni, E.S., Paolanti, M., Grilli, E., Remondino, F., Murtiyoso, A., Landes, T., 2020. A benchmark for large-scale heritage point cloud semantic segmentation. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLIII-B2-2020, 1419–1426. doi.org/10.5194/isprs-archives-XLIII-B2-2020-1419-2020.

Messaoudi, T., 2017. Proposition d'une ontologie de domaine dédiée à l'annotation d'images spatialisées pour le suivi de la conservation du patrimoine culturel bâti (Ph.D. thesis). Ecole nationale supérieure d'arts et métiers- ENSAM. Français. NNT: 2017ENAM0021

Miceli, A., Morandotti, M., Parrinello, S., 2020. 3D survey and semantic analysis for the documentation of built heritage. The case study of Palazzo Centrale of Pavia University. *VITRUVIO* 5, 65. doi.org/10.4995/vitruvio-ijats.2020.13634.

Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R., 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, *CoRR*, abs/2003.08934. arXiv:2003.08934.

Morbidoni, C., Pierdicca, R., Paolanti, M., Quattrini, R., Mammoli, R., 2020. Learning from synthetic point cloud data for historical buildings semantic segmentation. *J. Comput. Cult. Herit.* 13, 1–16. doi.org/10.1145/3409262.

Murphy, M., McGovern, E., 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. doi.org/10.1016/j.isprsjprs.2012.11.006.

Murphy, M., McGovern, E., Pavia, S., 2009. Historic building information modelling (HBIM). *Structural Survey* 27, 311–327. doi.org/10.1108/02630800910985108.

Murtiyoso, A., Grussenmeyer, P., 2022. Automatic point cloud noise masking in close range photogrammetry for buildings using AI-based semantic labelling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVI-2/W1-2022, 389–393, doi.org/10.5194/isprs-archives-XLI-2-W1-2022-389-2022, 2022.

Murtiyoso, A., Grussenmeyer, P., 2019. Point cloud segmentation and semantic annotation aided by GIS data for Heritage complexes. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2/W9, 523–528. doi.org/10.5194/isprs-archives-XLII-2-W9-523-2019.

Ni, H., Lin, X., Ning, X. and Zhang, J., 2016. Edge Detection and Feature Line Tracing in 3D-Point Clouds by Analyzing Geometric Properties of Neighborhoods. *Remote Sensing*, 8(9):710. doi.org/10.3390/rs8090710.

Ni, H., Lin, X., Zhang, J., 2017. Classification of ALS Point Cloud with Improved Point Cloud Segmentation and Random Forests. *Remote Sensing* 9, 288. doi.org/10.3390/rs9030288.

Noardo, F., 2018. Architectural heritage semantic 3D documentation in multi-scale standard maps. *Journal of Cultural heritage*, 32, 156-165. doi.org/10.1016/j.culher.2018.02.009.

Oreni, D., Brumana, R., Georgopoulos, A., Cuca, B., 2013. HBIM for conservation and built heritage: towards a library of vaults and wooden bean floors. I*SPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-5/W1, 215–221. doi.org/10.5194/isprsannals-II-5-W1-215-2013.

Otepka, J., Ghuffar, S., Waldhauser, C., Hochreiter, R., Pfeifer, N., 2013. Georeferenced point clouds: a survey of features and point cloud management. *ISPRS International Journal of Geo-Information* 2, 1038–1065. doi.org/10.3390/ijgi2041038.

Owen, S., 2000. A Survey of Unstructured Mesh Generation Technology. *International Meshing Roundtable* 3, 239- 267.

Özdemir, E., Remondino, F., Golkar, A., 2019. Aerial point cloud classification with deep learning and machine learning algorithms. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-4/W18, 843–849. doi.org/10.5194/isprs-archives-XLII-4-W18-843-2019.

Palma, G., Corsini, M., Cignoni, P., Scopigno, R., Mudge, M., 2010. Dynamic shading enhancement for reflectance transformation imaging. *J. Comput. Cult. Herit.* 3, 1–20. doi.org/10.1145/1841317.1841321.

Pamart, A., Morlet, F., De Luca, L., 2019. A fully automated incremental photogrammetric processing dedicated for collaborative remote-computing workflow. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2/W9, 565–571. doi.org/10.5194/isprs-archives-XLII-2-W9-565-2019.

Pamart, A., Morlet, F., De Luca, L., Véron, P., 2020. A robust and versatile pipeline for automatic photogrammetric-based registration of multimodal Cultural Heritage documentation. *Remote Sensing* 12, 2051. doi.org/10.3390/rs12122051.

Pandya D., Shah B., 2014. Comparative Study on Automatic Image Annotation. *International Journal of Emerging Technology and Advanced Engineering*, 4(3). ISSN 2250-2459.

Pătrăucean, V., Armeni, I., Nahangi, M., Yeung, J., Brilakis, I., Haas, C., 2015. State of research in automatic as-built modelling. *Advanced Engineering Informatics* 29, 162–171. doi.org/10.1016/j.aei.2015.01.001.

Paumard, M.-M., Picard, D., Tabia, H., 2020. Deepzzle: Solving Visual Jigsaw Puzzles With Deep Learning and Shortest Path Optimization. *IEEE Trans. on Image Process*. 29, 3569–3581. doi.org/10.1109/TIP.2019.2963378.

Pellis,, E., Murtiyoso, A., Masiero, A., Tucci, G., Betti, M., Grussenmeyer, P., 2022. An image-based deep learning workflow for 3D heritage point cloud semantic segmentation, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVI-2/W1-2022, 429–434. doi.org/10.5194/isprs-archives-XLVI-2-W1-2022-429-2022, 2022.

Pepe, M., Costantino, D., Restuccia Garofalo, A., 2020. An efficient pipeline to obtain 3D model for HBIM and structural analysis purposes from 3D point clouds. *Applied Sciences* 10, 1235. doi.org/10.3390/app10041235.

Perfetti, L., Fassi, F., Gulsan, H., 2019. Generation of gigapixel orthophoto for the maintenance of complex buildings. Challenges and lesson learnt. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-2/W9, 605–614. doi.org/10.5194/isprs-archives-XLII-2-W9-605-2019.

Pierdicca, R., Paolanti, M., Matrone, F., Martini, M., Morbidoni, C., Malinverni, E.S., Frontoni, E., Lingua, A.M., 2020. Point Cloud Semantic Segmentation Using a Deep Learning Framework for Cultural Heritage. *Remote Sensing* 12, 1005. doi.org/10.3390/rs12061005.

Pierrot Deseilligny, M., Clery, I., 2012. APERO, an open source bundle adjustment software for automatic calibration and orientation of set of images. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XXXVIII-5/W16, 269–276. doi.org/10.5194/isprsarchives-XXXVIII-5-W16-269-2011.

Pierrot-Deseilligny, M., De Luca, L., Remondino, F., 2011. Automated Image-Based Procedures for Accurate Artifacts 3D Modeling and Orthoimage Generation. *Geoinformatics FCE CTU* 6, 291–299. doi.org/10.14311/gi.6.36.

Pocobelli, D.P., Boehm, J., Bryan, P., Still, J., Grau-Bové, J., 2018a. BIM for heritage science: a review. *Heritage Science* 6, 30. doi.org/10.1186/s40494-018-0191-4.

Pocobelli, D.P., Boehm, J., Bryan, P., Still, J., Grau-Bové, J., 2018b. Building Information Modeling for monitoring and simulation data in heritage buildings. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII–2, 909–916. doi.org/10.5194/isprs-archives-XLII-2-909-2018.

Ponchio, F., Callieri, M., Dellepiane, M., Scopigno, R., 2020. Effective Annotations Over 3D Models. *Computer Graphics Forum* 39, 89–105. doi.org/10.1111/cgf.13664.

Poux, F., Hallot, P., Neuville, R., Billen, R., 2016. Smart point cloud: definition and remaining challenges. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* IV-2/W1, 119–127. doi.org/10.5194/isprs-annals-IV-2-W1-119-2016.

Previtali, M., Barazzetti, L., Brumana, R., Scaioni, M., 2014. Towards automatic indoor reconstruction of cluttered building rooms from point clouds. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci*. II–5, 281–288. doi.org/10.5194/isprsannals-II-5-281-2014.

Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2017. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *CoRR*, abs/1612.00593. arXiv:1612.00593.

Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *CoRR*, abs/1706.02413. arxiv.org/abs/1706.02413.

Quattrini, R., Malinverni, E.S., Clini, P., Nespeca, R., Orlietti, E., 2015. From TLS to HBIM. High quality semantically-aware 3D modeling of complex architecture. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XL-5/W4, 367–374. doi.org/10.5194/isprsarchives-XL-5-W4-367-2015.

Rabbani T., Van Den Heuvel F., Vosselmann G., 2006. Segmentation of point clouds using smoothness constraint. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* 36(5), 248-253.

Rupnik E., Daakir M., Pierrot-Deseilligny M., 2017. MicMac – a free, open-source solution for photogrammetry. In : Open Geospatial Data, Software and Standards. Vol. 2, n° 1. 10.1186/s40965-017-0027-2.+

Remondino, F., 2011. Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sensing* 3, 1104–1138. doi.org/10.3390/rs3061104.

Remondino, F., Campana, S., 2014. 3D recording and modelling in archaeology and Cultural Heritage. Theory and best practices. BAR International Series ISBN-1407312308.

Remondino, F., Rizzi, A., 2010. Reality-based 3D documentation of natural and cultural heritage sites—techniques, problems, and examples. *Applied Geomatics* 2, 85–100. doi.org/10.1007/s12518-010-0025-x.

Rizaldy, A., Firdaus, W., 2012. Direct georeferencing: a new standard in photogrammetry for high accuracy mapping. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XXXIX-B1, 5–9. doi.org/10.5194/isprsarchives-XXXIX-B1-5-2012.

Rocha, Mateus, Fernández, Ferreira, 2020. A Scan-to-BIM Methodology Applied to Heritage Buildings. *Heritage* 3, 47–67. doi.org/10.3390/heritage3010004.

Rodriguez-Gonzalvez, P., Rodriguez-Martin, M., 2019. Weld Bead Detection Based on 3D Geometric Features and Machine Learning Approaches. *IEEE Access* 7, 14714–14727. doi.org/10.1109/ACCESS.2019.2891367.

Rodríguez-Moreno, C., Reinoso-Gordo, J.F., Rivas-López, E., Gómez-Blanco, A., Ariza-López, F.J., Ariza-López, I., 2018. From point cloud to BIM: an integrated workflow for documentation, research and modelling of architectural heritage. *Survey Review* 50, 212–231. doi.org/10.1080/00396265.2016.1259719.

Roussel, R., Bagnéris, M., De Luca, L., Bomblet, P., 2019. A digital diagnosis for the “autumn” statue (Marseille, France): photogrammetry, digital cartography and construction of a thesaurus. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-2/W15, 1039–1046. doi.org/10.5194/isprs-archives-XLII-2-W15-1039-2019.

Santagati, C., Lo Turco, M., Garozzo, R., 2018. Reverse information modeling for historic artefacts: towards the definition of a level of accuracy for ruined heritage. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII–2, 1007–1014. doi.org/10.5194/isprs-archives-XLII-2-1007-2018.

Scalas, A., Mortara, M., Spagnuolo, M., 2020. A pipeline for the preparation of artefacts that provides annotations persistence. *Journal of Cultural Heritage* 41, 113–124. doi.org/10.1016/j.culher.2019.06.006.

Scalas, A., Mortara, M., Spagnuolo, M., 2017. 3D annotation transfer. in: Schreck, T., Weyrich, T., Sablatnig, R., Stular, B. (Eds.), Proc. of the Eurographics Workshop on Graphics and Cultural Heritage. The Eurographics Association, 1-5. doi.org/10.2312/gch.20171295.

Scamozzi, V., 1615. L'Idea Dell'architettura Universale; Venezia, 1615.

Schmidhuber, J., 2015. Deep Learning in Neural Networks: An Overview. *Neural Networks* 61, 85-117.

Schnabel, R. Efficient Point-Cloud Processing with Primitive Shapes (Ph.D. thesis), University of Bonn, Bonn, Germany, 2009.

Serna, S., Schmedt, H., Ritz, M., Stork, A., 2012. Interactive Semantic Enrichment of 3D Cultural Heritage Collections. in: Arnold, D., Kaminski, J., Niccolucci, F., Stork, A. (Eds.), VAST: International Symposium on Virtual Reality. The Eurographics Association, 33–40. doi.org/10.2312/VAST/VAST12/033-040.

Shanoer, M.M., Abed, F.M., 2018. Evaluate 3D laser point clouds registration for cultural heritage documentation. *The Egyptian Journal of Remote Sensing and Space Science* 21, 295–304. doi.org/10.1016/j.ejrs.2017.11.007.

Shrivakshan, G.T., 2012. A comparison of various edge detection techniques used in image processing.

Simeone, D., Cursi, S., Acierno, M., 2019. BIM semantic-enrichment for built heritage representation. *Automation in Construction* 97, 122–137. doi.org/10.1016/j.autcon.2018.11.004.

Soler, F., Torres, J.C., León, A.J., Luzón, M.V., 2013. Design of cultural heritage information systems based on information layers. *J. Comput. Cult. Herit.* 6, 1–17. doi.org/10.1145/2532630.2532631.

Stefani, C., Brunetaud, X., Janvier-Badosa, S., Beck, K., De Luca, L., Al-Mukhtar, M., 2014. Developing a toolkit for mapping and displaying stone alteration on a web-based documentation platform. *Journal of Cultural Heritage* 15, 1–9. doi.org/10.1016/j.culher.2013.01.011.

Taddeucci, S., 2021. Metodologie integrate H-BIM per la conservazione e valorizzazione del patrimonio storico. Il caso della Chiesa di Santa Maria del Carmine a Pisa (M.Sc. thesis). Corso di Laurea Magistrale in Ingegneria Edile Architettura, Università di Pisa.

Tan, K., Cheng, X., 2020. Distance Effect Correction on TLS Intensity Data Using Naturally Homogeneous Targets. *IEEE Geosci. Remote Sensing Lett.* 17, 499–503. doi.org/10.1109/LGRS.2019.2922226.

Tancik, M., Mildenhall, B., Wang, T., Schmidt, D., Srinivasan, P.P., Barron, J.T., Ng, R., 2021. Learned Initializations for Optimizing Coordinate-Based Neural Representations. arXiv:2012.02189.

Tang, P., Huber, D., Akinci, B., Lipman, R., 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, 829–843. doi.org/10.1016/j.autcon.2010.06.007.

Tarini, M., Pietroni, N., Cignoni, P., Panozzo, D., et Puppo, E., 2010. Practical quad mesh simplification. *Computer Graphics Forum* 29 (2), 407418. doi.org/10.1111/j.1467-8659.2009.01610.x.

Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P., 2007. Hough-transform and extended RANSAC algorithms for automatic detection of 3D building roof planes from Lidar data. in: Proc. of the ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, Espoo, Finland. 407-412.

Teague M.R., 1980. Image Analysis vie the General Theory of Moments. *JOSA*, 70(8), 920-930.

Tenchini, L., 2016. La Cappella di Sant'Agata a Pisa. Analisi storica e rilievo tridimensionale finalizzati al restauro della cappella e alla riqualificazione dell'area circostante (M.Sc. thesis). Corso di Laurea Magistrale in Ingegneria Edile-Architettura, Università di Pisa.

Teruggi, S., Grilli, E., Russo, M., Fassi, F., Remondino, F., 2020. A Hierarchical Machine Learning Approach for Multi-Level and Multi-Resolution 3D Point Cloud Classification. *Remote Sensing* 12, 2598. doi.org/10.3390/rs12162598.

Thomson, C., Boehm, J., 2015. Automatic Geometry Generation from Point Clouds for BIM. *Remote Sensing* 7, 11753–11775. doi.org/10.3390/rs70911753.

Tommasi, C., Achille, C., Fassi, F., 2016. From point cloud to BIM: a modelling challenge in the Cultural Heritage field. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLI-B5, 429–436. doi.org/10.5194/isprsarchives-XLI-B5-429-2016.

Vallet, J.-M., De Luca, L., Feillou, M., 2012. Une nouvelle approche spatio-temporelle et analytique pour la conservation des peintures murales sur le long terme. *In Situ*, 19 doi. org/10.4000/insitu.9829 insitu. https://doi.org/10.4000/insitu.9829.

Van De Ville, D., Nachtegael, M., Van der Weken, D., Kerre, E.E., Philips, W., Lemahieu, I., 2003. Noise reduction by fuzzy image filtering. *IEEE Transactions on Fuzzy Systems* 11, 429–436. doi. org/10.1109/TFUZZ.2003.814830.

Varinlioglu, G., & Balaban, Ö., 2021. Artificial intelligence in architectural heritage research. *The Routledge Companion to Artificial Intelligence in Architecture*, 207. doi.org/10.4324/9780367824259-13

Vieira, M., & Shimada, K., 2005. Surface mesh segmentation and smooth surface extraction through region growing. *Computer aided geometric design* 22 (8), 771-792.

Volk, R., Stengel, J., Schultmann, F., 2014. Building Information Modeling (BIM) for existing buildings — Literature review and future needs. *Automation in Construction* 38, 109–127. doi. org/10.1016/j.autcon.2013.10.023.

Vosselman, G., Gorte, B., Sithole, G., B, T., 2003. Recognising structure in laser scanner point clouds. in: Proc. of the ISPRS working group VIII/2: laser scanning for forest and landscape assessment, Freiburg, Germany, October 3-6, 2004, 33-38.

Wang, C., Cho, Y.K., Kim, C., 2015. Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. *Automation in Construction* 56, 1-13. doi. org/10.1016/j.autcon.2015.04.001.

Wang, J., Shan, J., 2009. Segmentation of LiDAR point clouds for building extraction. In: Proc. of the American Society for Photogramm. Remote Sens. Annual Conference, Baltimore, MD, 9-13.

Weinmann, M., 2016. Reconstruction and Analysis of 3D Scenes. Springer International Publishing: Cham, Switzerland. ISBN 978-3-319-29244-1.

Weinmann, M., Jutzi, B., Hinz, S., 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. doi.org/10.1016/j.isprsjprs.2015.01.016.

Weinmann, M., Jutzi, B., Mallet, C., 2013. Feature relevance assessment for the semantic interpretation of 3D point cloud data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-5/W2, 313–318. doi.org/10.5194/isprsannals-II-5-W2-313-2013.

Xie, Y., Tian, J., Zhu, X.X., 2020. Linking points with labels in 3D: A review of point cloud semantic segmentation. *IEEE Geosci. Remote Sens.* 8, 38–59. doi.org/10.1109/MGRS.2019.2937630.

Xiong, X., Adan, A., Akinci, B., Huber, D., 2013. Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in Construction* 31, 325–337. doi. org/10.1016/j.autcon.2012.10.006.

Xu, T., Xu, L., Yang, B., Li, X., Yao, J., 2017. Terrestrial Laser Scanning Intensity Correction by Piecewise Fitting and Overlap-Driven Adjustment. *Remote Sensing* 2017, 9, 1090. doi. org/10.3390/rs9111090.

Yang, X., Koehl, M., Grussenmeyer, P., 2018. Mesh-To-BIM: from segmented mesh elements to BIM model with limited parameters. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII–2, 1213–1218. doi.org/10.5194/isprs-archives-XLII-2-1213-2018.

Yang, X., Lu, Y.-C., Murtiyoso, A., Koehl, M., Grussenmeyer, P., 2019. HBIM Modeling from the Surface Mesh and Its Extended Capability of Knowledge Representation. *IJGI* 8, 301. doi.org/10.3390/ijgi8070301.

Yu, C.-H., Groza, T., Hunter, J., 2013. Reasoning on Crowd-Sourced Semantic Annotations to Facilitate Cataloguing of 3D Artefacts in the Cultural Heritage Domain, in: Salinesi, C., Norrie, M.C., Pastor, Ó. (Eds.), Advanced Information Systems Engineering, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 228–243. doi.org/10.1007/978-3-642-41338-4_15.

Zerlenga, O., 2020. Ludi Matematici by Leon Battista Alberti. *Diségno* 233–239. doi.org/10.26375/disegno.7.2020.22.

Zhang, S., Teizer, J., Lee, J.-K., Eastman, C.M., Venugopal, M., 2013. Building Information Modeling (BIM) and Safety: Automatic Safety Checking of Construction Models and Schedules. *Automation in Construction* 29, 183–195. doi.org/10.1016/j.autcon.2012.05.006.

Chapter 10

Sitography

Zhang, X., 2008. Curvature estimation of 3D point cloud surfaces through the fitting of normal section curvatures. Proc. of AsiaGraph 2008, 72-79.

Agisoft Metashape [on-line]. [Accessed on July 11, 2021]. Web-site: agisoft.com/

Arches - An open-source data management platform for the heritage field [on-line]. [Accessed on March, 26, 2022]. Web-site: www.archesproject.org/

Autodesk Revit [on-line]. [Accessed on July 11, 2021]. Web-site: www.autodesk.it/products/revit/overview

BIM Task Group, 2011. BIM Industry Working Group. A report for the Government Construction Client Group—March 2011. Communications. 2011. [on-line]. [Accessed on October 29, 2021]. Web-site: bimtaskgroup.org/wp-conte nt/uploa ds/2012/03/BIS-BIM-strat egy-Report.pdf

BuildingSMART International. IFC Overview Summary. 2012. [on-line]. [Accessed on October 29, 2021]. Web-site: buildingsmarttech.org/specifications/ifc-overview/ifc-overview-summary

Bundler: Structure from Motion (SfM) for Unordered Image Collections [on-line]. [Accessed on July 11, 2021]. Web-site: cs.cornell.edu/~snavely/bundler/

D. Lgs. 42/2004, Codice dei beni culturali e del paesaggio, pubblicato nella Gazzetta Ufficiale n. 45 del 24 febbraio 2004- Supplemento Ordinario n. 28

Cidoc-CRM Conceptual Reference Model [on-line]. [Accessed on October 18, 2021]. Web-site: www.cidoc-crm.org/

CloudCompare. [on-line]. [Accessed on October 2, 2021]. Web-site: danielgm.net/cc/

ClearEdge3D. [on-line]. [Accessed on October 10, 2021]. Web-site: clearedge3D.com

Code du Patrimoine, ordonnance numéro 2004-178 du 20 février 2004 [on-line]. [Accessed on June 6, 2021]. Web-site: legifrance.gouv.fr/codes/id/LEGITEXT000006074236/

COTAC BIM4C Integrating HBIM Framework Report Part 1 [on-line]. [Accessed on July 5, 2021]. Web-site: cotac.global

EN ISO 19650:2019. British Standards Institution BSI EN ISO 19650: Organization and digitization of information about buildings and civil engineering works, including building information modelling- Information management using building information modelling, London: BSI

EN ISO 16739:2013. Industry Foundation Classes (IFC) for Data Sharing in the Construction and Facility Management Industries. International Standard. 2013. [on-line]. [Accessed on October 29, 2021]. Web-site: iso.org/standard/51622.html

EU Commission, 2019. European Commission report on Cultural Heritage: Digitization, Online Accessibility and Digital Preservation. Progress report on the implementation of Commission Recommendation 2011/711/EU, 12 June 2019. [on-line]. [Accessed on September 27, 2021]. Web-site: digital-strategy.ec.europa.eu

EU Commission, 2014. Council conclusions of 21 May 2014 on cultural heritage as a strategic resource for a sustainable Europe, 2014 OJ C 183. [Accessed on September 27, 2021]. Web-site: eur-lex.europa.eu

EU Commission, 2011. Commission Recommendation of 27 October 2011 on the digitization and online accessibility of cultural material and digital preservation, 2011/711/EU OJ L 283 [on-line]. [Accessed on September 27, 2021]. Web-site: eur-lex.europa.eu

GRAVITATE, Geometric reconstruction and novel semantic reunification of cultural heritage objects [on-line]. [Accessed on July 5, 2021]. Web-site: cordis.europa.eu/project/id/665155/it

HISTORIC ENGLAND, 2017. Historic England 2017 BIM for Heritage: Development of Historic Building Information Model. Swindon. Historic England [on-line]. [Accessed on October 10, 2021]. Web-site: historicengland.org.uk

IMAGEJ [on-line]. [Accessed on July 15, 2021]. Web-site: imaginit.com: /imagej.nih.gov/ij/download.html

IMAGINiT – Autodesk platinum partner. [on-line]. [Accessed on October 10, 2021]. Web-site: imaginit.com

INCEPTION, Inclusive Cultural Heritage in Europe through 3D semantic modelling [on-line]. [Accessed on July 5, 2021]. Web-site: inception-project.eu/en

INTESA.IT Intelligenza Artificiale, Machine Learning e Deep Learning. [on-line]. [Accessed on November 31, 2021]. Web-site: intesa.it/intelligenza-artificiale-machine-learning-e-deep-learning

Jatana, 2020. Dive into Supervised and Unsupervised Machine Learning [on-line]. [Accessed on November 30, 2021]. Web-site: medium.com/analytics-vidhya/dive-into-supervised-machine-learning-90eaeeb8b8bd

JPI CH. Joint Programming Initiative on Cultural Heritage and Global Change. [Accessed on January 21, 2022]. Web-site: heritageresearch-hub.eu/homepage/joint-programming-initiative-on-cultural-heritage-homepage/

LONDON CHARTER, 2006. London Charter for the computer-based visualization of cultural heritage [on-line]. [Accessed on September 27, 2021]. Web-site: 3dicons-project.eu/guidelines-and-case-studies

M-HERACLES. Heritage by Point cloud processing [on-line]. [Accessed on July 1, 2021]. Web-site: github.com/murtiad/HERACLES

MATLAB, Statistics and Machine learning Toolbox [on-line]. [Accessed on July 11, 2021]. Web-site: it.mathworks.com/products/statistics.html

MacLean, Margaret; Davis, Ben H (eds) (1999). Time & Bits: Managing Digital Continuity. Getty Publications.

NBIMS, 2007. National Building Information Modeling Standard, 2007- National Institute of Building Sciences. Facilities Information Council on National BIM standards [on-line]. [Accessed on October 21, 2021]. Web-site: 3dicons-project.eu/guidelines-and-case-studies

Notre-Dame Science, 2021. Chantier Scientifique Notre-Dame de Paris. Groupe de travail données numériques. [on-line]. [Accessed on October 21, 2021]. Web-site: notre-dame.science/gt-donnees-numeriques/

Pointfuse. [on-line]. [Accessed on July 12, 2021]. Web-site: pointfuse.com

PointNet. [on-line]. [Accessed on January 21, 2022]. Web-site: github.com/charlesq34/pointnet

Potree- WebGL point cloud viewer for large datasets [on-line]. [Accessed on October 18, 2021]. Web-site: buildinginformationmanagement.files.wordpress.com

Ray, N., Sokolov, D., Reberol, M., Ledoux, F. et Lévy, B., 2017. [on-line] Hexahedral Meshing: Mind the Gap! [Accessed on October 20, 2021]. Web-site: https://hal.inria.fr/hal-01551603.

Rhino3D – Rhinoceros by Robert McNeel & Associates. [on-line]. [Accessed on June 10, 2021]. Web-site: www.rhino3d.com/

Rhino.Inside Revit v. 1.0 [on-line]. [Accessed on June 27, 2021]. www.rhino3d.com/inside/revit/1.0/

SEVILLE PRINCIPLES, 2011. International Principles of Virtual Archaeology, The Seville Principles 2011. [on-line]. [Accessed on September 27, 2021]. Web-site: londoncharter.org

SICaR – Sistema Informativo per i Cantieri di Restauro. [on-line]. [Accessed on January 21, 2022]. Web-site: sicar.beniculturali.it

Sketchfab – 3D models. [on-line]. [Accessed on October 18, 2021]. Web-site: www.sketchfab.com

Time Machine [on-line]. [Accessed on July 5, 2021]. Web-site: www.timemachine.eu/

Trainable Weka Segmentation, ImageJ [on-line]. [Accessed on July 26, 2021]. Web-site: imagej.net/plugins/tws/

UNESCO, 1972. Convention concerning the protection of the World Cultural and Natural Heritage [on-line]. [Accessed on June 6, 2021]. Web-site: whc.unesco.org/en/conventiontext/.

UNESCO, 2003a. Convention for the Safeguarding of the Intangible Cultural Heritage [on-line]. [Accessed on June 6, 2021]. Web-site: ich.unesco.org/en/convention

VOLVOX, Grasshopper plug-in. [Accessed on October 16, 2021]. Web-site: food4rhino.comen/app/ volvox

3D Icons, 2014. 3D Icons, Guidelines & case studies [on-line]. [Accessed on July 11, 2021]. Web-site: 3dicons-project.eu/guidelines-and-case-studies

Notre-Dame Science, 2021. Chantier Scientifique Notre-Dame de Paris. Groupe de travail données numériques. [on-line]. [Accessed on October 21, 2021]. Web-site: notre-dame.science/gt-donnees-numeriques/

Pointfuse. [on-line]. [Accessed on July 12, 2021]. Web-site: pointfuse.com

PointNet. [on-line]. [Accessed on January 21, 2022]. Web-site: github.com/charlesq34/pointnet

Potree- WebGL point cloud viewer for large datasets [on-line]. [Accessed on October 18, 2021]. Web-site: buildinginformationmanagement.files.wordpress.com

Ray, N., Sokolov, D., Reberol, M., Ledoux, F. et Lévy, B., 2017. [on-line] Hexahedral Meshing: Mind the Gap! [Accessed on October 20, 2021]. Web-site: https://hal.inria.fr/hal-01551603.

Rhino3D – Rhinoceros by Robert McNeel & Associates. [on-line]. [Accessed on June 10, 2021]. Web-site: www.rhino3d.com/

Rhino.Inside Revit v. 1.0 [on-line]. [Accessed on June 27, 2021]. www.rhino3d.com/inside/revit/1.0/

SEVILLE PRINCIPLES, 2011. International Principles of Virtual Archaeology, The Seville Principles 2011. [on-line]. [Accessed on September 27, 2021]. Web-site: londoncharter.org

SICaR – Sistema Informativo per i Cantieri di Restauro. [on-line]. [Accessed on January 21, 2022]. Web-site: sicar.beniculturali.it

Sketchfab – 3D models. [on-line]. [Accessed on October 18, 2021]. Web-site: www.sketchfab.com

Time Machine [on-line]. [Accessed on July 5, 2021]. Web-site: www.timemachine.eu/

Trainable Weka Segmentation, ImageJ [on-line]. [Accessed on July 26, 2021]. Web-site: imagej.net/plugins/tws/

UNESCO, 1972. Convention concerning the protection of the World Cultural and Natural Heritage [on-line]. [Accessed on June 6, 2021]. Web-site: whc.unesco.org/en/conventiontext/.

UNESCO, 2003a. Convention for the Safeguarding of the Intangible Cultural Heritage [on-line]. [Accessed on June 6, 2021]. Web-site: ich.unesco.org/en/convention

VOLVOX, Grasshopper plug-in. [Accessed on October 16, 2021]. Web-site: food4rhino.comen/app/ volvox

3D Icons, 2014. 3D Icons, Guidelines & case studies [on-line]. [Accessed on July 11, 2021]. Web-site: 3dicons-project.eu/guidelines-and-case-studies

Chapter 11

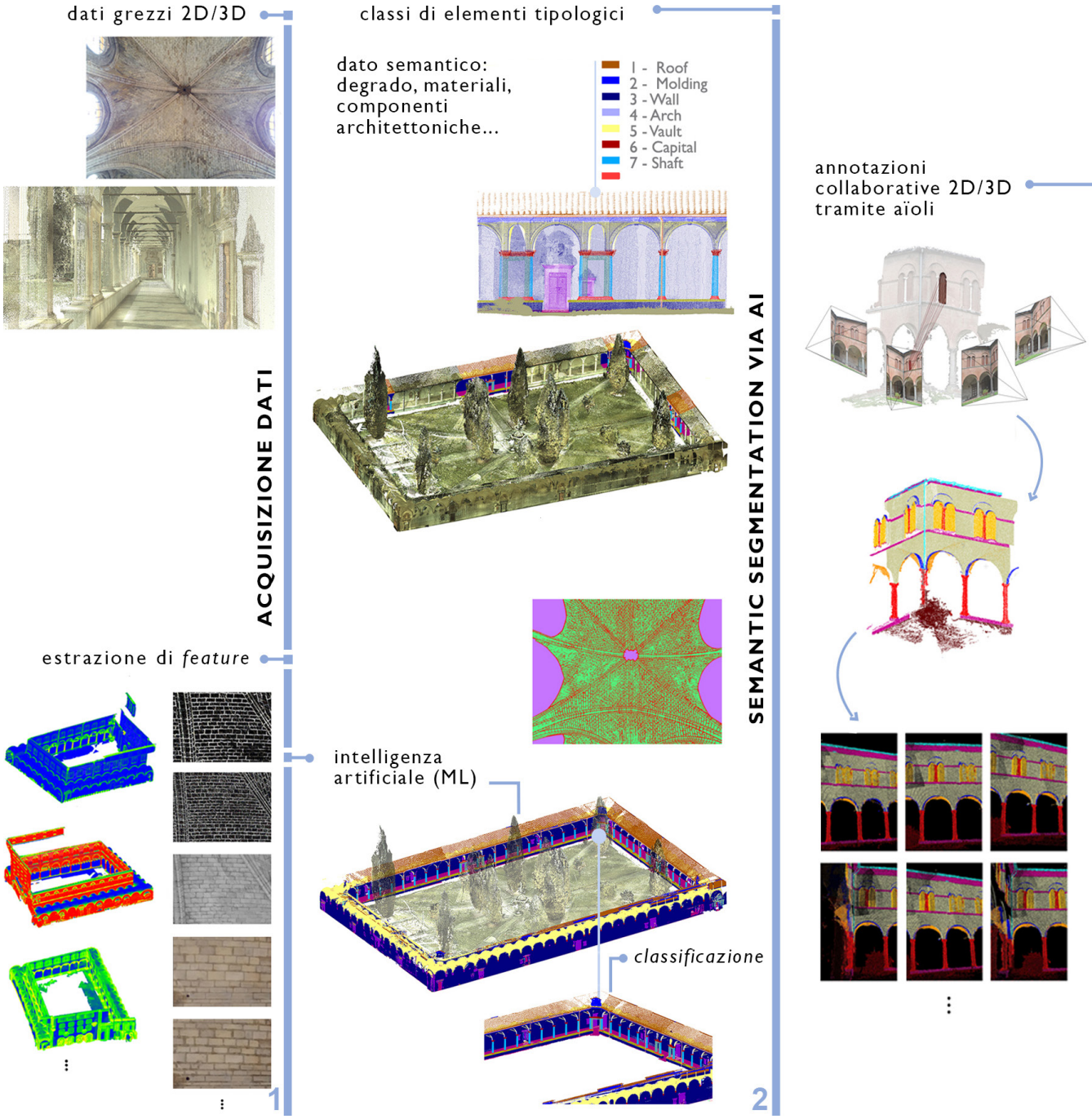
List of abbreviations

AEC	Architecture, Engineering and Construction
AI	Artificial Intelligence
ASTRO	Laboratory on Scientific and Topographic Applications for Operative Surveying (ita: *Applicazioni Scientifiche e Topografiche per il Rilievo Operativo*)
BIM	Building Information Modeling
CAD	Computer Aided Design
CEN	European Committee for Standardization
CMM	Coordinate Measuring Machine
CNRS	National Center for Scientific Research, French Ministry of Culture (fr: *Centre Nationale de la Recherche Scientifique, Ministère de la Culture*)
DICEA	Department of Civil and Environmental Engineering (ita: *Dipartimento di Ingegneria Civile e Ambientale*)
DICI	Department of Civil and Industrial Engineering (ita: *Dipartimento di Ingegneria Civile e Industriale*)
DESTeC	Department of Energy, Systems, Territory and Construction Engineering (ita: *Dipartimento di Ingegneria dell'Energia, dei Sistemi, del Territorio e delle Costruzioni*)
DL	Deep Learning
ENSAM	*Ecole Nationale Supérieure d'Arts et Métiers ParisTech*
FN	False Negatives
FP	False Positives
GIS	Geographic Information Systems
GNSS	Global Navigation Satellite Systems
H-BIM	Heritage-/Historic- Building Information Modeling
ICCROM	International Centre for the Study of the Preservation and Restoration of Cultural Property
ICOMOS	International Council on Monuments and Sites
ID	Identification code
IFC	Industry Foundation Classes

ISO	International Organization for Standardization
JPI CH	Joint Programming Initiative on Cultural Heritage and Global Change
LiDAR	Light Detection and Ranging
LoD	Level of Detail
LoI	Level of Information
LISPEN	Physical and Digital Systems Engineering Laboratory (fr: *Laboratoire d'Ingénierie des Systèmes Physiques et Numériques*)
MAP	Models and Simultations for Architecture and Heritage (fr: *Modèles et simulations pour l'architecture et le patrimoine*)
ML	Machine Learning
MVS	Multi-View Stereo
NeRF	Neural Radiance Fields
NURBS	Non-Uniform Rational B-Splines
RANSAC	Random Sample and Consensus
RGB	Red, Green and Blue
RF	Random Forest
SIP	Surface Information Points
SfM	Structure from Motion
SVMs	Support Vector Machines
TACO	Totally Automated Co-Registration and Orientation
TLS	Terrestrial Laser Scanner
TN	True Negatives
TOF	Time-of-Flight
TP	True Positives
UAV	Unmanned Aerial Vehicle
UMR	Joint Research Unit (fr: *Unité Mixte de Recherche)*
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNIPI	University of Pisa (ita: *Università di Pisa*)
VPL	Visual Programming Language
xBIM	eXtensible Building Information Modeling
2D	Bi-dimensional
3D	Three-dimensional

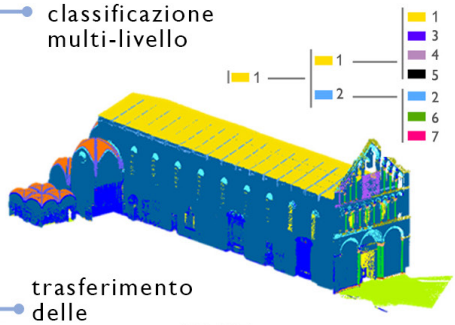
Appendix

Appendix 1 - Schema di sintesi

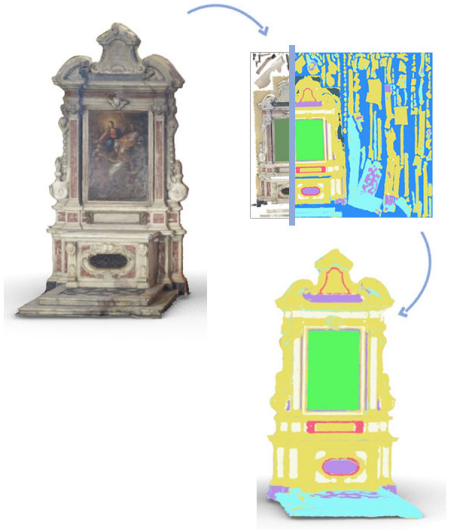


PROPAGAZIONE DELLE INFORMAZIONI 2D/3D

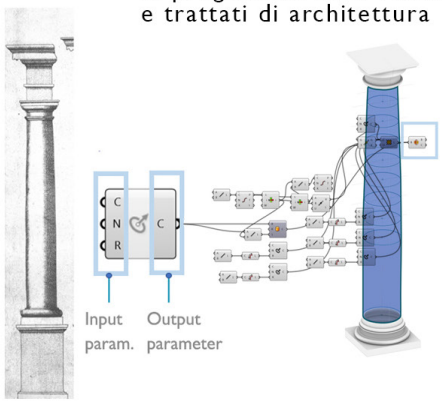
classificazione multi-livello



trasferimento delle annotazioni 2D/3D



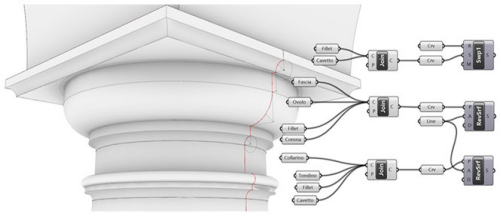
geometrie 3D: linguaggi di programmazione visuale e trattati di architettura



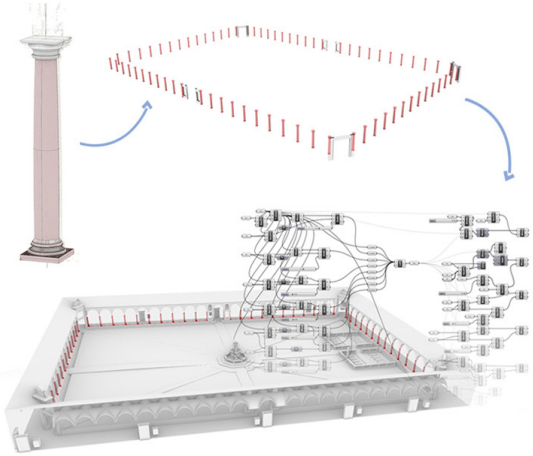
3

H-BIM, STRUTTURAZIONE SEMANTICA, ANNOTAZIONI LOCALIZZATE

geometrie *template*

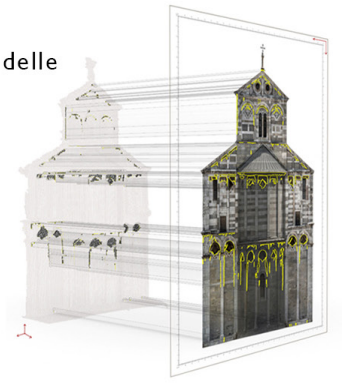


propagazione di informazioni semantiche



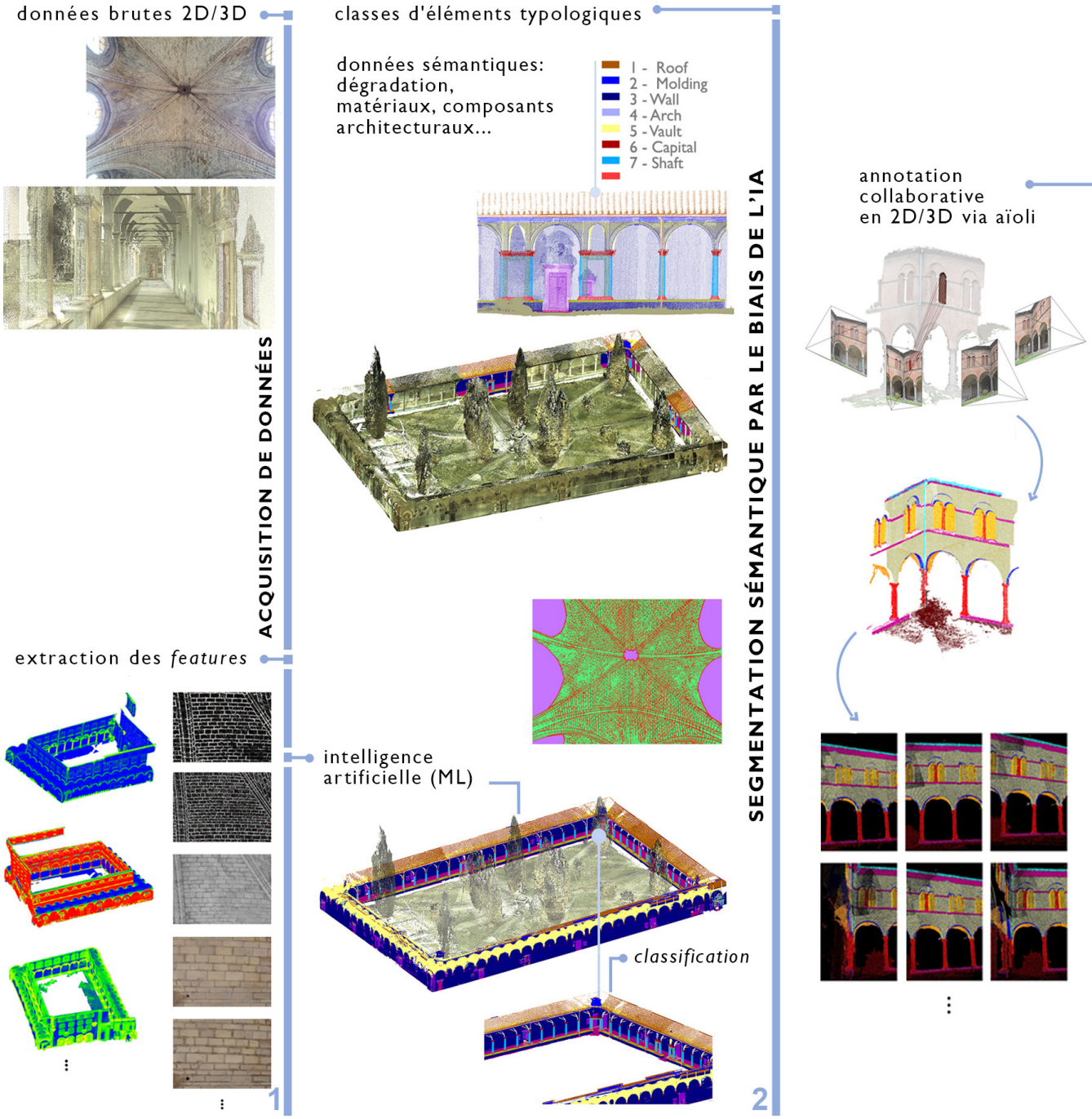
Inserimento di annotazioni localizzate

archiviazione, recupero, aggiornamento delle annotazioni semantiche



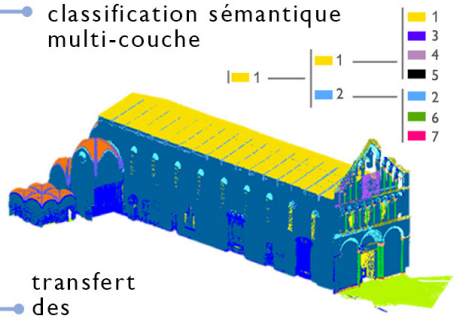
4

Appendix 2 - Résumé graphique

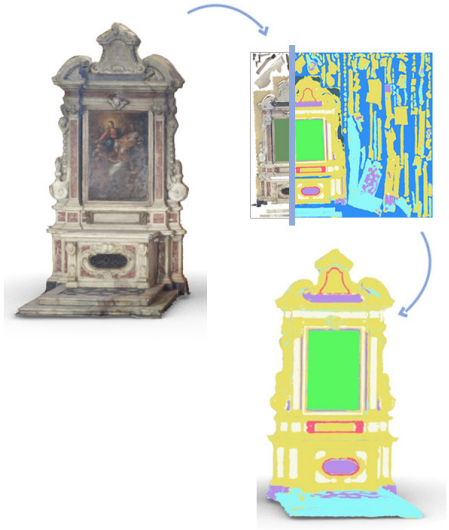


PROPAGATION DE L'INFORMATION 2D/3D

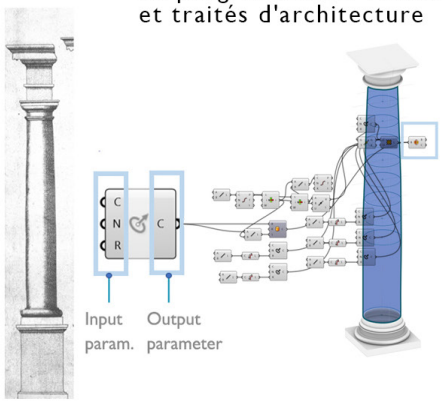
classification sémantique multi-couche



transfert des annotations 2D/3D



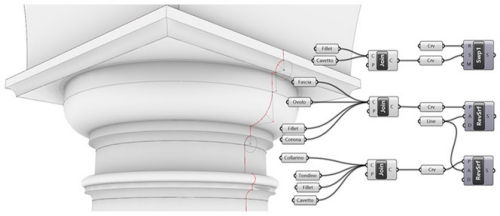
géométries 3D : langages de programmation visuelle et traités d'architecture



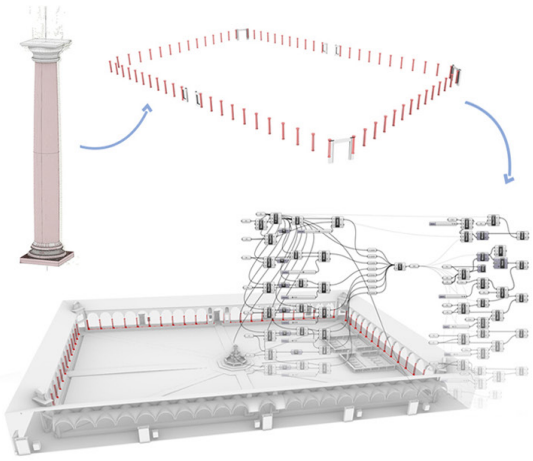
3

H-BIM, STRUCTURATION SÉMANTIQUE, ANNOTATIONS LOCALISÉES

geometries de référence (*template*)



propagation des informations sémantiques



introduction d'annotations localisées

archivage, récupération, mise à jour des annotations sémantiques



4

Appendix 3 – Aïoli class indexing script

Script – Annotations created via the Aioli collaborative reality-based annotation platform are exported and regrouped into classes based on the value of the 'Label' descriptor

```
class = 'C:\...Aioli\AnnotationArchive\3fe0353e361827b7\376c619270091dc8'
```

```
 % Selects directory (folder containing annotations of a single
class)

 class = uigetdir()

 % Assigns class index
 c = 3 % Insert class index

 % Extracts plyfiles as a list of all 'Prop.ply' files present
in the directory
 % of a single class (folders and sub-folders).
 % This operation is broken down class by class.

 plyfiles = dir(fullfile(class, '**/Prop.ply'))
```

plyfiles = 4×1 struct

Fields	name	folder	date	bytes	isdir	datenum
1	'Prop.ply'	'C:\...Aioli\AnnotationArchive\3fe0353e361827b7\376c619270091dc8\8f94c804d7e9dffd'	'20-lug-2021 17:36:18'	1264293	0	7.3836e+05
2	'Prop.ply'	'C:\...Aioli\AnnotationArchive\3fe0353e361827b7\376c619270091dc8\b975d71442f36fc1'	'20-lug-2021 17:36:18'	1581104	0	7.3836e+05
3	'Prop.ply'	'C:\...Aioli\AnnotationArchive\3fe0353e361827b7\376c619270091dc8\c03e39b70b25f5d7'	'20-lug-2021 17:36:18'	779390	0	7.3836e+05
4	'Prop.ply'	'C:\...Aioli\AnnotationArchive\3fe0353e361827b7\376c619270091dc8\c86e20a55df39e43'	'20-lug-2021 17:36:18'	972297	0	7.3836e+05

```
 % Reads all plyfiles in the directory -each corresponding to a
single - annotation and inserts them in a struct file
named'annotatedpc'.

 for k = 1:numel(plyfiles)

     % Loads files in the workspace
     % 'annotatedpc': a struct file containing the annotations
of each single class as a struct file

      annotatedpc(k) = pcread([plyfiles(k).folder '\'
plyfiles(k).name]);
 end
```

annotatedpc = 1×13 object

	1	2	3	4	5	6	7	...
1	1×1 pointCloud	1×1 pointCloud	1×1 pointCloud	1×1 pointCloud	1×1 pointCloud	1×1 pointCloud	1×1 pointCloud	

```
 clear k
 clear plyfiles % clears temporary variables

 % Stores xyz and rgb values in a single table

 xyzrgb=[];
 for nann = 1:numel(annotatedpc) %nann = number of annotations
performed in Aioli
     xyzrgb = vertcat(xyzrgb,[annotatedpc(:,nann).Location
double(annotatedpc(:,nann).Color)])
 end
```

```
xyzrgb = 18343×6 single matrix
   85.7587   97.6904  101.9672  129.0000  128.0000   73.0000
   85.7337   97.7034  101.9573  105.0000   99.0000   78.0000
   85.7575   97.6815  101.9578  121.0000  116.0000   87.0000
      ⋮
xyzrgb = 41216×6 single matrix
   85.7587   97.6904  101.9672  129.0000  128.0000   73.0000
   85.7337   97.7034  101.9573  105.0000   99.0000   78.0000
   85.7575   97.6815  101.9578  121.0000  116.0000   87.0000
      ⋮
```

```
 % Converts array to table and assigns variable names (xyz
coordinates and rgb values)

 xyzrgb = array2table(xyzrgb,...
     'VariableNames', {'X', 'Y', 'Z', 'R', 'G', 'B'})
```

xyzrgb = 95792×6 table

	X	Y	Z	R	G	B
1	85.7587	97.6904	101.9672	129	128	73
2	85.7337	97.7034	101.9573	105	99	78
3	85.7575	97.6815	101.9578	121	116	87
⋮						

```
% Assigns class index to the 'Label' column
xyzrgb{:, 7} = c
xyzrgb.Properties.VariableNames {'Var7'} = 'Label'
```

xyzrgb = 95792×7 table

	X	Y	Z	R	G	B	Label
1	85.7587	97.6904	101.9672	129	128	73	3
2	85.7337	97.7034	101.9573	105	99	78	3
3	85.7575	97.6815	101.9578	121	116	87	3
⋮							

```
% Writes table to local .csv file
writetable(xyzrgb, '3_Floor.csv')
```

Appendix 4 – Point cloud file formats

Type	Extension	Description	Supported features
BIN	.bin	CloudCompare own format	Normals, colors (RGB), scalar fields (>1), labels, viewports, display options, etc.
ASCII	.asc,.txt,.xyz, .csv,.pts	ASCII point cloud file (X,Y,Z,etc.)	Normals, colors (RGB), scalar fields (all)
LAS	.las	ASPRS lidar point clouds	Colors (RGB) and various scalar fields (see LAS 1.4 specifications)
E57	.e57	ASTM E57 file format	Normals, colors (RGB or I), scalar field (intensity)
PTX	.ptx	LEICA point cloud export format	Robust normals can be computed at loading time
FARO	.fls, *.fws	FARO formats	Scalar field (reflection value)
PCD	.pcd	Point Cloud Library format	Colors (RGB), normals, scalar fields (>1)
PLY	.ply	Stanford 3D geometry format (cloud or mesh)	Normals, colors (RGB or I), one or several scalar fields, a single texture
OBJ	.obj	Wavefront mesh	Normals, materials and textures
VTK	.vtk	VTK file format (triangular mesh or cloud only)	Normals, colors (RGB), scalar field(s) (>1)
STL	.stl	STereoLithography file format (mesh)	Normals
FBX	.fbx	Autodesk (Filmbox) File Format	Normals, colors (RGB), materials and textures
DXF	.dxf	Autocad DXF format	Normals, colors (RGB)
SHP	.shp	ESRI Shape file format	Scalar fields (1 per entity)
RASTER	.geotiff, etc.	Common raster formats (GDAL)	Layers (as scalar fields)
PV	.pv	Point cloud + scalar field	Scalar field (1)
PN	.pn	Point cloud + normals	Normals

Source: CloudCompare, File I/O
[on-line]. [Accessed on January 30, 2021]
Web-site: cloudcompare.org/doc/wiki/index.php?title=FILE_I/O

Appendix 5 – List of publications

Refereed articles and book chapters.

- Croce V., Caroti G., Piemonte A., De Luca L., Véron P., 2022.
  H-BIM and Artificial Intelligence. Semi-automatic Scan-to-BIM processes via classified digital heritage data
  Accepted for publication in Sensors, MDPI.

- Croce V., Caroti, G., Piemonte, A. 2021.
  Propagation of semantic information between orthophoto and 3D replica: a H-BIM system for the north transept of Pisa Cathedral.
  In: Geomatics, Natural Hazards and Risks, Vol. 12 (1), pp. 2225-2252.
  doi.org/10.1080/19475705.2021.1960432

- Croce V., Caroti G., Piemonte A., Bevilacqua M.G.B. 2021.
  From survey to semantic representation for Cultural Heritage: the 3D modeling of recurring architectural elements.
  In: e-journal ACTA Imeko, Vol. 10 (1).
  dx.doi.org/10.21014/acta_imeko.v10i1.842

- Croce V., Caroti G., De Luca L., Jacquot K., Piemonte A., Véron P. 2021.
  From the Semantic Point Cloud to Heritage-Building Information Modeling: A Semiautomatic Approach Exploiting Machine Learning.
  In: Remote Sens., 13, 461. doi.org/10.3390/rs13030461

- Croce V., 2019.
  Castelluccio di Norcia. Proposte di recupero per il borgo distrutto dal terremoto.
  In: Centri storici minori – Formazione Abbandono Riappropriazione, Bascherini E. (a cura di), Campano Ed., Pisa, pp. 83-94

- Croce V., Martínez-Espejo Zaragoza I. 2019.
  Survey technologies aimed at preserving the built heritage.
  In: The Matter of Future Heritage, TU Delft Open, CPCL Journal, Bologna, pp. 21-34

Refereed papers.

- Croce V., 2022.
  Classification of the fortified heritage of Pisan walls: Neural Networks and photogrammetry on crowd-sourced images
  Accepted for publication at FORTMED 2023 - Fortifications of the Mediterranean Coast International Conference (23-24-25 March 2023)

- Croce V., 2022.
  The Chapel of Sant'Agata in Pisa. 3D surveying, Artificial Intelligence and archival heritage
  Accepted for publication at the 43rd International Conference of Representation Disciplines Teachers, Unione Italiana del Disegno UID (15-16-17 September 2022)

- Menchelli S., Caroti G., Maltagliati F., Piemonte A., Belcari D., Croce V., Tempesti J.
  Un progetto transdisciplinare multidisciplinare sul relitto "Melania" (Vada, Rosignano M.mo Livorno)
  Accepted for publication at VII National Congress on Underwater Archaeology, Sardegna, Olbia – La Maddalena (11-12-13-14 May 2023)

- Croce V., 2021.
  Towards the automation of Scan-to-BIM processes via annotated digital data. The main cloister of Pisa Charterhouse.
  Presented and accepted for publication at DREAM Silesia, Nature-Architecture-Culture, Panel II – Digital Heritage (22-23 April 2022)

- Croce V., Taddeucci S., Caroti G., Piemonte A., Martino M., Bevilacqua M.G., 2021.
  Semantic mapping of decay and materials through Artificial Intelligence and integrated H-BIM management.
  Presented and accepted for publication at the REAACH-ID conference – Representation for Enhancement and management through Augmented reality and Artificial intelligence: Cultural Heritage and Innovative Design (12-13 October 2021)

- Croce V., Caroti G., De Luca L., Piemonte A., Véron P., Bevilacqua M.G., 2021.
  Tra Intelligenza Artificiale e H-BIM per la descrizione semantica dei beni culturali: la Certosa di Pisa/Artificial Intelligence and H-BIM for the semantic description of cultural heritage: the Pisa Charterhouse.
  In: Arena A., Arena M., Mediati D., Raffa P. (a cura di). Connettere. Un disegno per annodare e tessere. Linguaggi Distanze Tecnologie. Atti del 42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/ Connecting. Drawing for weaving relationship. Languages Distances

Technologies. Proceedings of the 42nd International Conference of Representation Disciplines Teachers. Milano: FrancoAngeli, pp. 608-625.

- Croce V., Caroti G., Piemonte A., Bevilacqua M.G., 2021.
  Connecting geometry and semantics via Artificial Intelligence: from 3D classification of heritage data to H-BIM representations
  In: Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2021, 145–152, https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-145-2021, 2021.

- Croce V., Caroti G., De Luca L., Piemonte A., Véron P. 2020.
  Semantic annotations on heritage models: 2D/3D approaches and future research challenges. ISPRS Int Arch Photogramm Remote Sens Spat Inf Sci., XLIII-B2-2020: 829–836. doi.org/10.5194/isprs-archives-XLIII-B2-2020-829-2020, 2020.

- Croce V., Caroti G., Piemonte A., Bevilacqua M.G. 2019.
  Geomatics for Cultural Heritage conservation: Integrated survey and 3D modeling.
  In: Proceedings of the IMEKO TC4 International Conference on Metrology for Archaeology and Cultural Heritage, MetroArchaeo, Florence, Italy, 4–6 December 2019; pp. 271–276.

- Croce V., Caroti G., Piemonte A. 2019.
  Assessment of earthquake-induced damage level on buildings: analysis of two different survey methods for a case study.
  In: ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, XLII-2/W15, pp. 351-58. doi.org/10.5194/isprs-archives-XLII-2-W15-351-2019, 2019.

- Croce V., Diamantidis D., Sýkora M. 2020.
  Planning post-earthquake surveys: assessment and reconstruction of small historical centers.
  In: Proc. of the IABSE Symposium Synergy of Culture and Civil Engineering, IABSE Symposium Synergy of Culture and Civil Engineering, 7-9 October 2020, Wroclaw, Poland, pp. 332-339.

- Croce V., Diamantidis D., Sýkora M. 2021.
  Seismic damage evaluation and decisions on interventions supported by UAV-based surveys. PROHITECH 2021 4th International Conference on Protection of Historical Constructions, 25-27 October 2021, Athens, Greece.

Awards and recognitions

- **Primo premio per tesi di dottorato 'Il restauro nell'era dell'H-BIM'. Categoria B – Tesi di dottorato e specializzazione (May 2022)**
  First prix for Ph.D. thesis, organized by Rotary Club di Roma with the patronage of Facoltà di Architettura e Facoltà di Ingegneria Civile e Industriale di Sapienza Università di Roma and in collaboration with Dipartimento di Storia, Disegno e Restauro dell'Architettura Scuola di specializzazione in Beni Architettonici e del Paesaggio Fondazione Almagià – ACER

- **BIM&DIGITAL Awards 2021 – Primo premio per tesi di dottorato nella Categoria 8 – Ricerca**
  First prize for Ph.D. thesis in Category 8 – Research

  First prize awarded to the Ph.D. thesis on the topic 'Artificial Intelligence to support BIM restoration works'. BIM&DIGITAL Awards 2021, organized by CLUST-ER BUILD and DIGITAL&BIM (https://www.concorsiarchibo.eu/digitalandbimawards-2021/home)

- **Prix VINCI2019 for Ph.D. thesis in co-tutelle, Université Franco-Italienne**
  Chapter II - Mobility grants for Ph.D. these in co-tutelle, scolarship (Aides à la mobilité pour thèses de doctorat en co-tutelle) Università Italo Francese, Università degli Studi di Torino Area Internazionalizzazione, Direzione Attività Istituzionali, Programmazione, Qualità e Valutazione

- **Bright 2019-La notte europea dei ricercatori, University of Florence. Finalist project at 'La ricerca in 3 minuti'**
  Project selected among the 25 finalists for presentation at the competition 'Research in 3 minutes', organized in the framework of the European night of researchers, University of Florence.

Appendix 6 - Digital & BIM Italia.
Competition panels

INTELLIGENZA ARTIFICIALE AL SERVIZIO DEI CANTIERI BIM DI RESTAURO

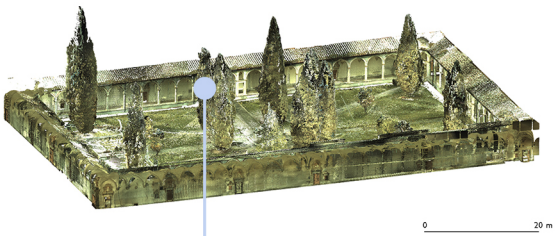
OBI

- Migliorare il proces
modelli BIM per il
(Scan-to-BIM) attrav
automatiche di clas
(nuvole di punti) ch
- l'Intelligenza Artific
il Machine Learning

NUVOLA DI PUNTI 3D

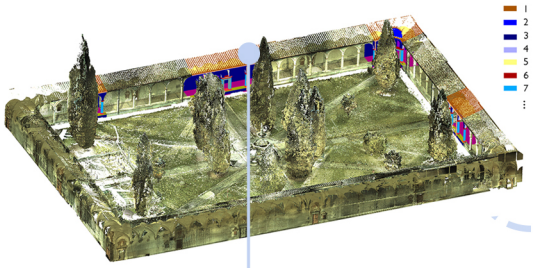
DATI DI TRAINING

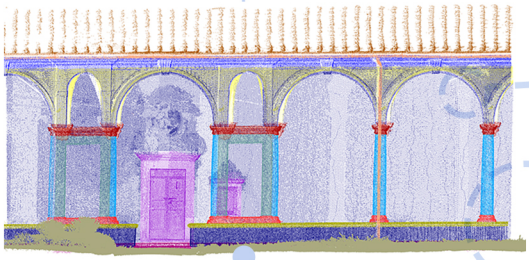
MACHINE











1 - Copertura
2 - Modanatura
3 - Muro
4 - Arco
5 - Volta
6 - Capitello
7 - Fusto
...

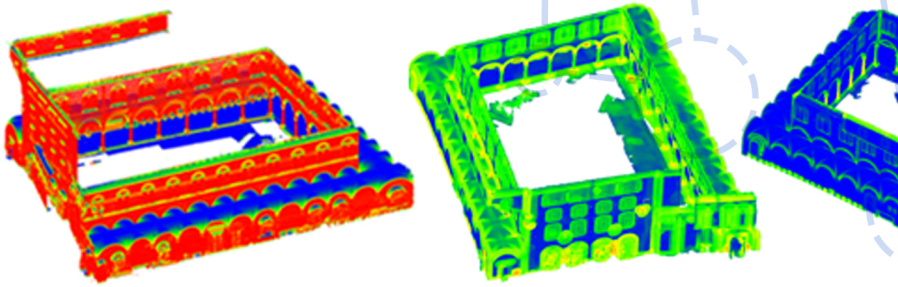
Nuvole di punti 3D, ottenute da laser scanner o fotogrammetria, costituiscono il dato grezzo di partenza. Tali dati, essendo non strutturati, rendono la ricostruzione di modelli 3D una procedura lunga, complessa e perlopiù manuale.

La strutturazione del dato 3D parte da procedure di *training*. Un modello di IA è allenato a riconoscere e classificare determinati elementi sulla nuvola di punti (colonne, modanature, muri, pavimenti), sulla base di un sufficiente numero di dati di esempio.

NUVOLA DI PUNTI 3D



ESTRAZIONE DI FEATURE GEOMETRICHE E COLORIM



ETTIVI

sso di generazione di
costruito
erso procedure più
ssificazione di dati 3D
ne sfruttino
iale e, in particolare,
g (ML).

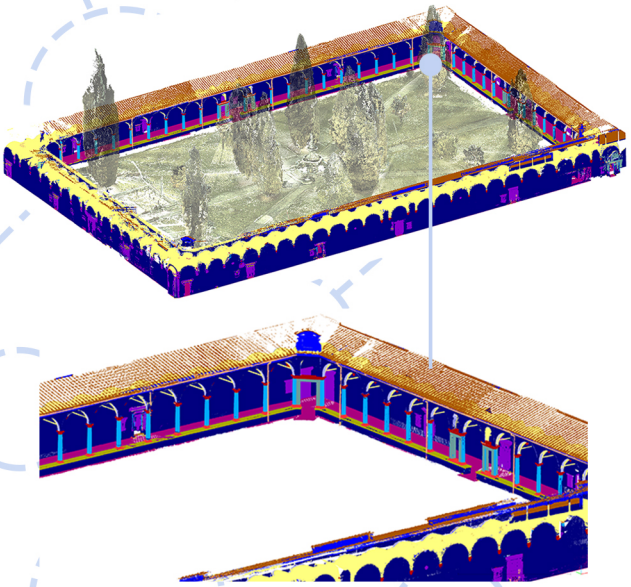
PROBLEMA

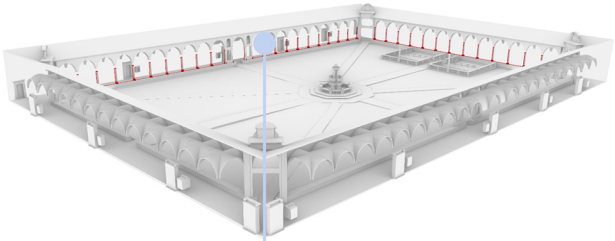
- Dispersione di conoscenze sui beni culturali
- Interpretazione dei dati di rilievo
- Gestione poco automatica e soggettiva della ricostruzione di modelli 3D a partire da nuvole di punti
- Complessità morfologica dei beni architettonici esistenti

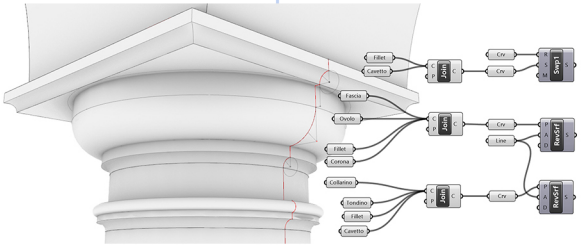
SOLUZIONE

- IA per la classificazione e la strutturazione semantica di dati 3D
- Identificazione e riconoscimento automatico di classi di elementi architettonici sulla nuvola di punti
- Automatizzazione dei processi Scan-to-BIM tramite classificazione automatica ML.

LEARNING → **CLASSIFICAZIONE SEMANTICA** → **RICOSTRUZIONE SCAN-TO-BIM**







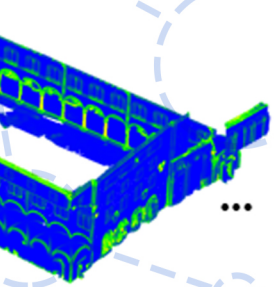
IA

Automatizzazione del processo Scan-to-BIM a partire da dati di rilievo 3D

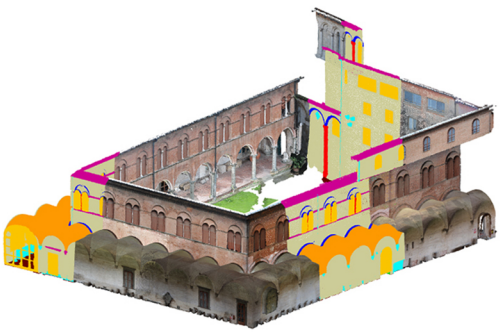
L'allenamento di un modello predittivo basato sull'Intelligenza Artificiale (Machine Learning) consente di classificare l'intero dataset riconoscendo elementi tipologici e architettonici ricorrenti a partire dalla nuvola di punti 3D.

La nuvola di punti ottenuta, semanticamente annotata, è sfruttata per ricostruire un modello BIM dell'edificio. Classi di elementi tipologici identificate sulla nuvola di punti sono modellate in ambiente BIM, con una procedura più automatica e veloce.

METRICHE

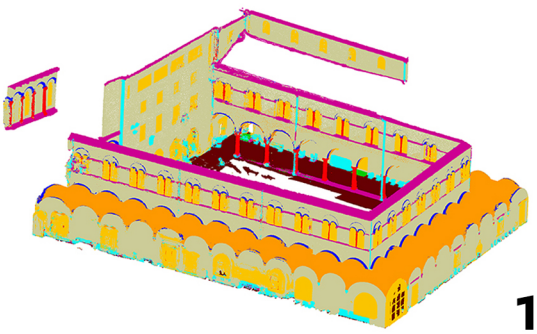


ANNOTAZIONE MANUALE



IA

CLASSIFICAZIONE SEMANTICA

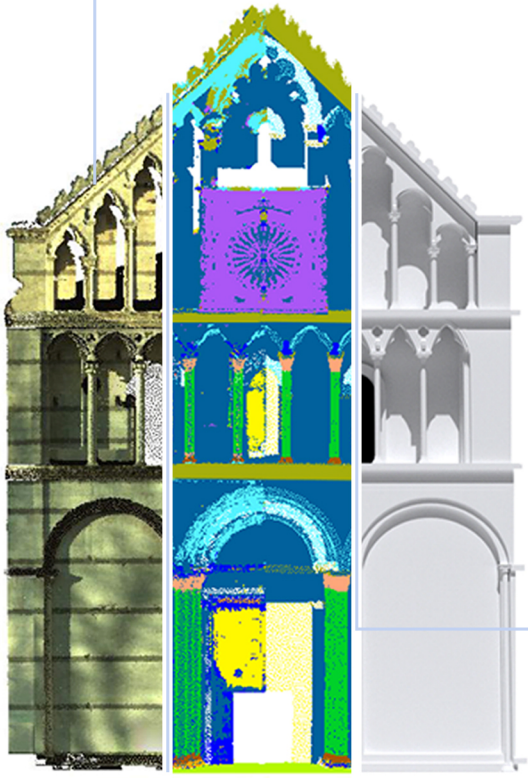


INTELLIGENZA ARTIFICIALE AL SERVIZIO DEI CANTIERI BIM DI RESTAURO

1. RI

- Acquisizione dati c
  punti e mesh grezz
- Rilievo laser scanne
  terrestre o da dron
- Integrazione con te
  tradizionali (stazion
  topografico, sistem

La **Chiesa di Santa Caterina a Pisa**

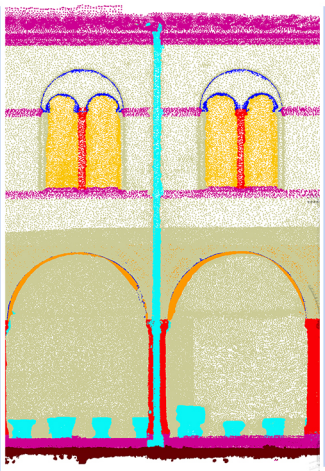


1. 2. 3.

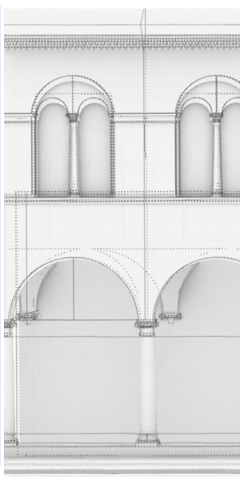
Il Chiostro del **Museo Nazionale di San Matteo a Pisa**



1.



2.



3.

1 - Arco
2 - Colonna
3 - Rosone
...

1 - Arco
2 - Colonna
3 - Modanatura
...

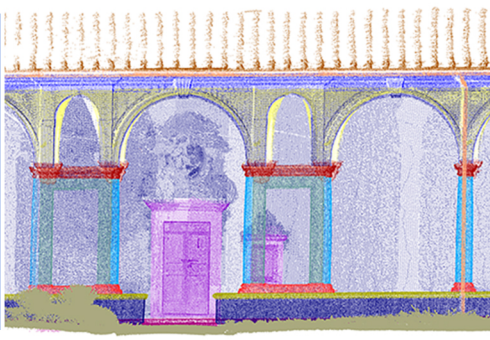
1 - Copertur
2 - Modanatur
3 - Mur
...

CASI STUDIO

La procedura è validata su una pluralità di edifici storici e monumentali del patrimonio architettonico italiano o francese.
Nel dettaglio, la Cattedrale di Notre Dame a Parigi e la Certosa di Calci costituiscono casi studio rappresentativi ed emblematici su cui viene testata la metodologia proposta.



1.



2.



Chiostro Granducale,
Certosa di Calci



Chiostro Grande,
Certosa di Calci



Chiostro del Capitolo,
Certosa di Calci



Facciata del cortile principale
Certosa di Calci

ILIEVO 3D

di rilievo 3D: nuvole di
ze
er e fotogrammetria
e
ecniche di rilievo
e totale, rilievo
i GNSS)

2. SEGMENTAZIONE SEMANTICA via IA

- Scelta di un intorno 3D e estrazione di feature geometriche e colorimetriche
- Annotazione manuale delle classi di elementi tipologici su una porzione
- Allenamento di un algoritmo Random Forest
- Annotazione semantica di classi di elemetni tipologici su tutta la nuvola di punti

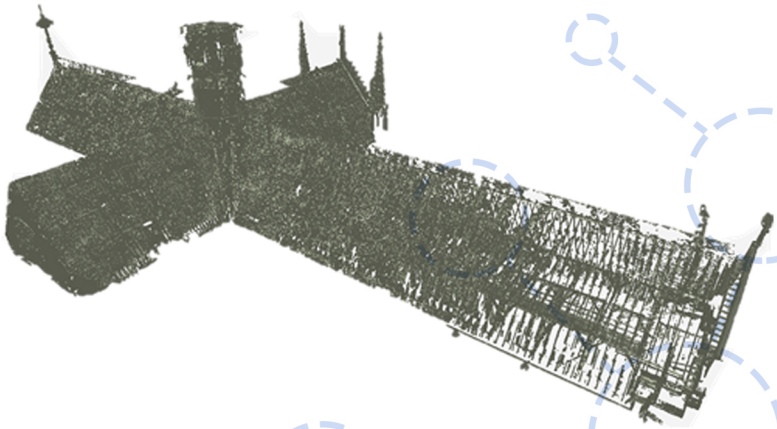
3. MODELLO H-BIM

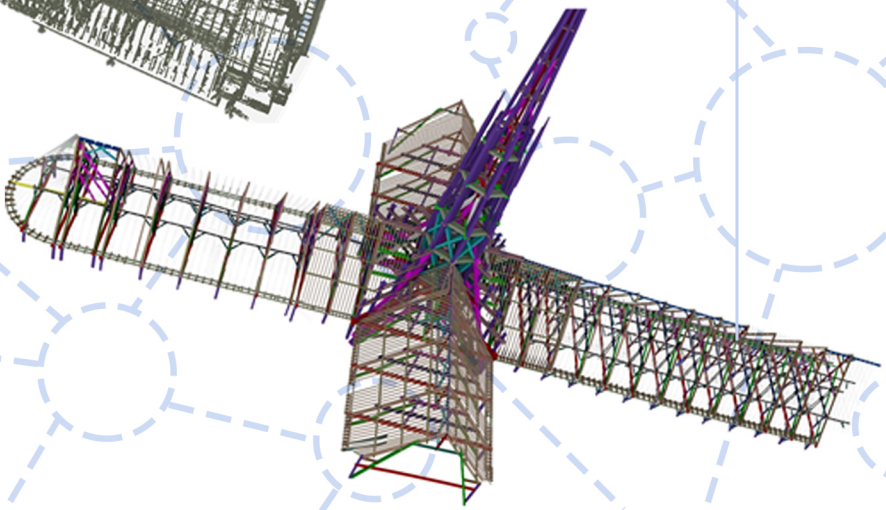
- Importazione delle classi di elementi tipologici in ambiente BIM
- Ricostruzione di geometrie *template* in base a canoni architettonici
- Propagazione delle informazioni tramite visual programming e modello H-BIM
- Automazione processo Scan-to-BIM

Il Chiostro grande della **Certosa Monumentale di Calci**

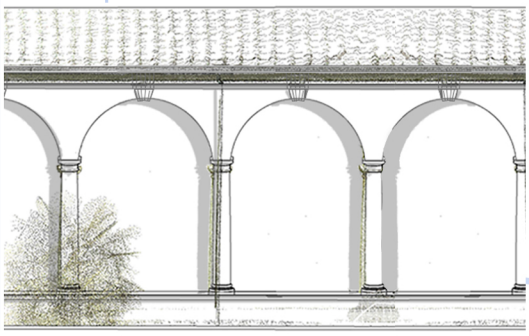
La copertura di **Notre Dame di Parigi**











3.

Come parte del patrimonio artistico e culturale dei due paesi, questi siti storici richiedono la definizione di un apparato grafico e informativo complesso, che consenta, tramite sistemi digitali, la condivisione e il recupero delle informazioni metriche, visuali e semantiche prodotte nelle diverse operazioni di documentazione, recupero e conservazione.



attedrale di Notre Dame,
arigi



Museo Nazionale di San Matteo,
Pisa



Cattedrale di Pisa, transetto sud
Pisa



Chiesa di Santa Caterina
Pisa

...

INTELLIGENZA ARTIFICIALE AL SERVIZIO DEI CANTIERI BIM DI RESTAURO



Classe 1 Copertura | Classe 2 Modanatura | Classe 3 Muro | Classe 4 Arco | Classe 5 Volta | Classe 6 Capitello | Classe 7 Fusto

1. Software per l'elaborazione nuvole di punti.
2. Software per la costruzione del modello H-BIM.

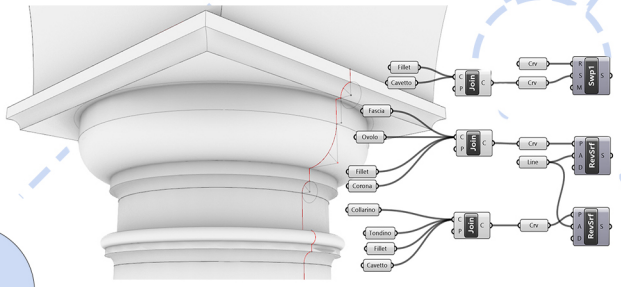
1. MATLAB

2. Rhinoceros + grasshopper | AUTODESK REVIT | Dynamo

- Le classi di elementi architettonici individuate tramite IA sono importate in ambiente BIM, e ricostruite in modo da farle corrispondere con la struttura IFC.
- Ogni classe di elementi è ricostruita a parte, ricorrendo a algoritmi di modellazione generativa (*Grasshopper e Dynamo*).

- Riconoscendo le geometrie ideali dai dati annotati, e a partire dal riferimento a trattati di architettura, viene derivata una rappresentazione template concettuale, che può essere propagata, tramite programmazione visuale, agli elementi della stessa classe tipologica.

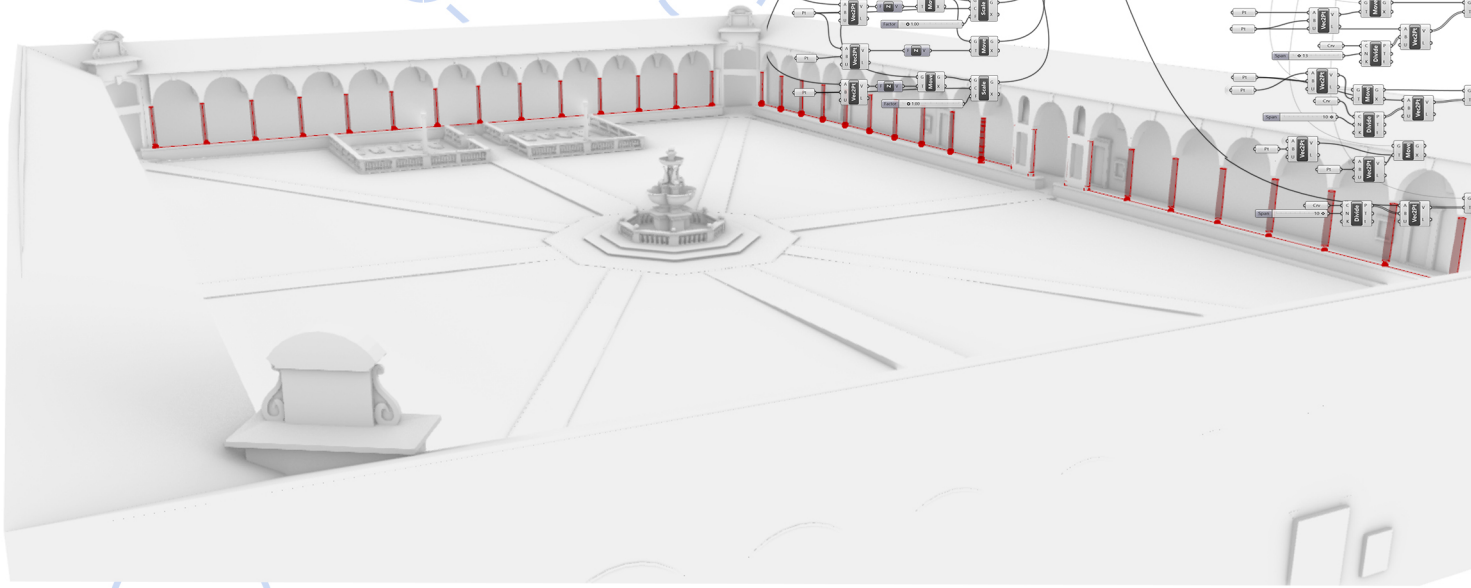
- Questa operazione, eseguita per tutto modello, realizza un sistema informat HBIM completo dove immagazzinare volta i dati relativi alla conoscenza, pr nell'ambito di analisi multidisciplinari bene, volte al recupero, al restauro, a conservazione.



Geometria *template* della classe di elementi

Funzioni di loft, estrusione, sweep...

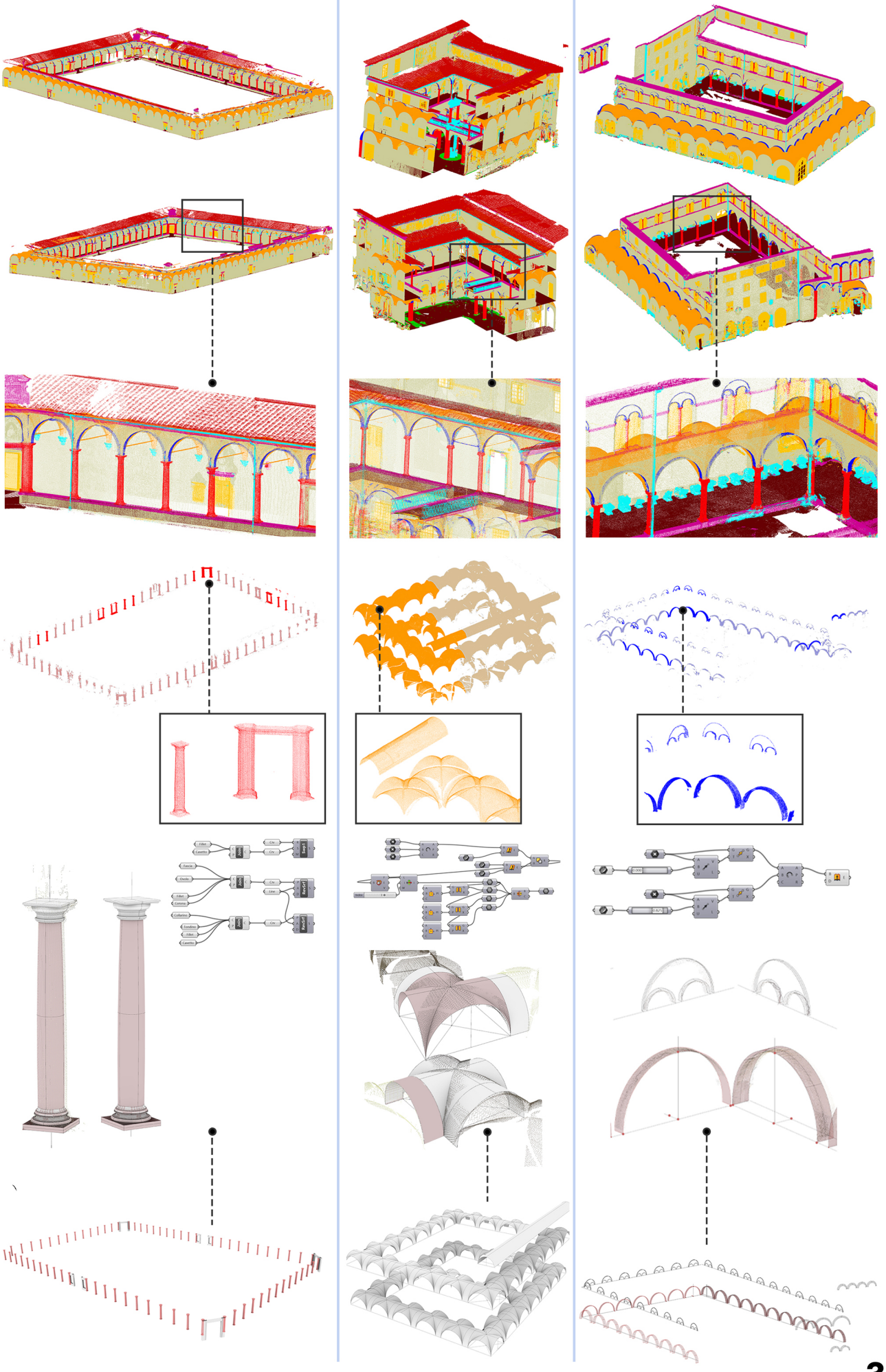
Propagazione della geome *template* all'intero modello



Chiostro grande,
Certosa di Pisa. **Colonna**

Chiostro granducale,
Certosa di Pisa. **Volta**

Chiostro del Museo Nazionale
di S.Matteo, Pisa. **Arco**



NUVOLA DI PUNTI CLASSIFICATA TRAMITE IA

SELEZIONE SINGOLA CLASSE

MODELLO 3D TEMPLATE GEOMETRY

PROPAGAZIONE A TUTTA LA CLASSE

Appendix 7

Résumé en langue française

Transfert et repérage d'annotations sémantiques pour le patrimoine architectural.

Un système méthodologique combinant Intelligence Artificielle, H-BIM et plateformes d'annotation collaborative reality-based.

Cette thèse de doctorat propose une approche méthodologique intégrée pour le transfert, la récupération et le partage d'annotations sémantiques, à partir de modèles numériques 2D et/ou 3D du patrimoine culturel, et en combinant les techniques d'Intelligence Artificielle, les environnements H-BIM et les plateformes d'annotation collaboratives et basées sur la réalité (*reality-based*) telles que Aïoli (aioli.cloud). La méthodologie proposée est validée sur des études de cas significatives du patrimoine architectural français et italien, comme la cathédrale Notre-Dame de Paris et la Chartreuse de Pise.

1. Contexte général

Le développement croissant des systèmes numériques d'acquisition, de restitution graphique et de modélisation a radicalement changé le secteur du patrimoine culturel, en proposant une nouvelle approche de l'analyse, de la préservation, de la documentation et de la diffusion des biens patrimoniaux aux conservateurs, architectes, ingénieurs, restaurateurs et autres experts.

Les solutions proposées au fil du temps pour la visualisation et l'exploitation des données numériques ont amplement démontré leur

efficacité. Des outils et des méthodes permettant des représentations virtuelles bi-dimensionnelles (2D, par exemple : images, ortho-photos, etc.) ou tri-dimensionnelles (3D, par exemple : nuages de points, modèles maillés, etc.), même issus d'études différentes et multidisciplinaires du secteur du patrimoine culturel, sont à la disposition d'une pluralité d'acteurs.

En même temps, le développement rapide des techniques d'acquisition permet de disposer d'une grande quantité de données numériques en un temps relativement court ; le relevé 3D à partir du balayage laser ou de la photogrammétrie est un moyen efficace pour soutenir la restitution graphique par la détection des formes et des matériaux, des dimensions et des proportions, de l'aspect et des contours.

Ainsi, une réplique virtuelle des objets et sites patrimoniaux existants peut être obtenue avec extrême précision. Cette réplique pourrait être conçue comme un jumeau numérique, servant de contrepartie virtuelle à l'objet physique, dans lequel les informations qui sont intégrées dans ce dernier peuvent être stockées numériquement (Grieves, 2014). Cependant, bien que de grands progrès soient aujourd'hui accomplis dans le domaine des systèmes d'acquisition et de relevé, les données numériques 2D/3D sont encore loin d'être entendus comme des moyens privilégiés pour la gestion des informations sur les études, les analyses, et sur toute autre documentation qui gravite autour d'un objet du patrimoine architectural. A ce jour, les modèles numériques ne sont pas des outils de gestion intégrés et collaboratifs, ne permettant pas l'archivage, l'accès et l'utilisation raisonnée des informations qualitatives (métriquement correctes) et quantitatives (se référant à la connaissance du bien), liées à un objet étudié. Ainsi, la représentation géométrique pure donnée par le modèle numérique n'est pas en soi suffisante pour répondre aux besoins concrets d'observation, de traitement et d'interprétation de données et de sources documentaires hétérogènes produites dans le cadre d'études patrimoniales.

Ceci est d'autant plus vrai si l'on considère la multiplicité des produits, documents, dessins 2D et modèles 3D utilisés, multiplicité qui est souvent à l'origine d'une dispersion des données et, par conséquent, des connaissances. La nature hétérogène des données produites, résultant de l'implication de professionnels de différents domaines d'étude, motive le besoin de développer de nouvelles méthodes et de nouveaux outils numériques pour le stockage, l'organisation et la classification de la documentation relative à un objet patrimonial, ainsi que la mise en œuvre de stratégies communes de support à la décision. Les systèmes d'information numériques visant à connecter la représentation purement géométrique et visuelle à des données sémantiques, c'est-à-dire liées à l'analyse de l'objet, sont donc de plus en plus recherchés et ouvrent la voie à de nouvelles approches multimodales et multi-utilisateurs pour étudier et préserver le

patrimoine culturel.

Dans ce contexte, d'une part, des processus de modélisation des informations du bâtiment (BIM) émergent pour la construction de représentations numériques partagées soutenant la gestion de l'information au sein des processus de conception, de construction et d'exploitation, et formant une base fiable pour les décisions (ISO 19650:2019). L'application des systèmes BIM au domaine du patrimoine est plutôt récente : les concepts de BIM, H-BIM, scan-to-BIM sont utilisés aujourd'hui avec des significations disparates pour définir la création de modèles BIM à partir des données issues du relevé architectural.

D'autre part, les techniques modernes d'Intelligence Artificielle (IA), si transférées de manière appropriée au domaine de la représentation architecturale, pourraient améliorer l'interprétation et la classification des multiples données numériques disponibles. Leur mise en œuvre permettrait l'accès, l'échange et la mise à jour continue d'informations sur le patrimoine architectural associées à des connaissances, qui peuvent concerner, par exemple, les transformations d'un bâtiment, la documentation de ses changements dans le temps, les interventions de restauration et de conservation, l'état de dégradation de ses surfaces, la cartographie des matériaux, etc. Dans cette perspective, une interprétation plus concrète de ce qui est représenté peut être envisagée : une lecture sémantique des données topographiques permettrait un processus rationalisé de collecte, d'élaboration, de visualisation des informations directement au sein des représentations numériques 2D ou 3D.

L'introduction de modèles enrichis sémantiquement offre un outil privilégié pour la description, l'analyse et la compréhension de tout l'univers cognitif qui s'articule autour de l'étude des objets et des sites du patrimoine.

Dans les plateformes actuelles d'annotation basées sur la réalité (*reality-based*) ou H-BIM, les représentations numériques sont associées à des informations significatives par l'insertion d'étiquettes appropriées, appelées *annotations*, directement sur le modèle numérique. Cependant, les solutions existantes ne permettent d'annoter qu'un seul type de représentation 2D ou 3D utilisé, soit un nuage de points, un maillage, un modèle paramétrique, un dessin technique 2D, une ortho-photo, une image d'archive, une photo.

Par conséquent, lorsque le support numérique est modifié, l'utilisation, l'exploitation et la mise à jour ultérieures d'informations cruciales concernant l'objet patrimonial considéré sont empêchées, à moins que les informations précédemment acquises ne soient convenablement réinsérées dans la nouvelle représentation.

C'est le goulot d'étranglement des systèmes d'information sémantique existants pour le patrimoine numérique : ils restent ancrés au type spécifique de représentation qu'ils supportent, empêchant

le transfert, la structuration et l'interconnexion ultérieure des informations produites par les différents professionnels impliqués. Une méthode appropriée pour le partage, la mise à jour et le transfert d'informations, valable indépendamment du type de modèle numérique choisi pour la représentation, répondrait en effet au besoin d'archiver, d'échanger et de mettre à jour les données relatives à un certain bien culturel.

La présente recherche découle de la nécessité d'assurer une approche plus objective et partagée pour le stockage et l'organisation des données sur les objets numériques, afin de permettre de comprendre pleinement, de structurer, d'accéder, d'archiver, de mettre à jour et de corréler, à partir de données de relevé initiales, la richesse de l'information produite dans le contexte des études sur le patrimoine.

Ce travail présente donc un cadre unifié pour l'exploitation et la réalisation de modèles patrimoniaux sémantiquement riches, qui peuvent être bénéfiques à la communauté des restaurateurs, des ingénieurs, des architectes, des conservateurs et de tous les autres acteurs concernés, qui sont continuellement confrontés au problème de la fusion, de la documentation et de la connexion des données patrimoniales, ainsi que de leur liaison pour les politiques de restauration, de gestion et de conservation.

2. Objectif

L'objectif de ce travail de thèse est la définition d'une approche méthodologique intégrée pour le transfert, la récupération et l'échange d'annotations sémantiques de modèles numériques 2D/3D du patrimoine, en exploitant les techniques d'Intelligence Artificielle et la construction d'environnements H-BIM. En se concentrant sur le mécanisme d'annotation sémantique pour l'enrichissement des répliques numériques, le but du travail est d'assurer la continuité des informations échangées dans les phases interdépendantes d'acquisition, d'observation, de traitement et de représentation des données du patrimoine. Les problèmes ouverts ci-dessous sont abordés :

- Identification d'un lien entre les annotations effectuées sur différents supports numériques (un nuage de points, un maillage, un modèle paramétrique, un dessin, une ortho-photo, une photo numérique, etc.) ;
- Classification des données patrimoniales 2D et 3D de manière plus automatisée et plus rationnelle ;
- Propagation de l'information entre les documents numériques 2D et les modèles 3D, ou même entre les textes, les images, les orthophotos et les modèles 3D ;

- Permettre l'interopérabilité d'une annotation sémantique, de sorte que les informations puissent être lues et consultées indépendamment du type de support numérique dans lequel elles ont été initialement rédigées ;
- Prévoir différents niveaux de lecture, de traitement et de mise à jour des informations afin que les différents experts du domaine du patrimoine culturel puissent accéder aux données sémantiques associées à un modèle numérique, même en fonction de leurs besoins, fonctions et exigences réciproques.

En abordant ces questions ouvertes, l'objectif de ce travail est la définition d'un système commun d'entrée et de récupération de l'information dans les modèles numériques sémantiques, à être :

- Indépendant du programme utilisé pour la spatialisation et l'insertion des données ;
- Insensible aux variations de la géométrie réelle de l'objet étudié, tant en termes de niveau de précision de la représentation que de type de modèle utilisé.

Ainsi, les informations sémantiques pourraient être localisées et stockées virtuellement sur l'objet patrimonial, indépendamment du type de représentation numérique choisi et de ses modifications dans le temps. Cela implique que les données brutes ou non-organisées, concernant tout type d'information, par exemple l'histoire, les images, les textures, les dégradations, les schémas de fissures, les interventions, etc. peuvent être interconnectées directement sur la représentation numérique.

3. Contexte expérimental et études de cas

La thèse et les travaux expérimentaux sont développés dans le cadre d'un accord international de thèse de doctorat de recherche conjointe (co-tutelle), impliquant les institutions suivantes :

- Le Laboratoire ASTRO - Applications scientifiques et topographiques pour le relevé opérationnel, du Département de Génie Civil et Industriel de l'Université de Pise (ASTRO, DICI, UNIPI) ;
- Le Département d'I ngénierie de l'énergie, des systèmes, des territoires et des constructions de l'Université de Pise (DESTeC, UNIPI) ;
- L'unité mixte de recherche Modèles et Simulations pour l'Architecture et le Patrimoine sous l'autorité du Centre National de la Recherche Scientifique et du Ministère de la Culture français (UMR 3495 CNRS/MC MAP) ;

- Le Département de Génie Civil et Environnemental de l'Université de Florence (DICEA, UNIFI), dans le cadre du Doctorat International en Génie Civil et Environnemental (Cycle XXXIV, curriculum *Conception, vérification et contrôle de la construction*) ;
- Le Laboratoire d'Ingénierie des Systèmes Physiques et Numériques de l'Ecole Nationale Supérieure d'Arts et Métiers Paris-Tech d'Aix-en-Provence (LISPEN, unité n. 7515, ENSAM Aix-en-Provence), et l'Ecole Doctorale Sciences et Métiers de l'Ingénieur ED SMI 432.

Le travail est financé par la Région Toscane (Borse Pegaso POR FSE TOSCANA 2014/2020). Il a reçu le prix Vinci 2019 (Chapitre II - Aides à la mobilité pour thèses de doctorat en co-tutelle) de l'Université Franco-Italienne, une institution supranationale qui encourage les collaborations de recherche entre l'Italie et la France.

En décembre 2021, la thèse de doctorat a reçu le premier prix pour la catégorie 8- Recherche des BIM&DIGITAL Awards 21, prix organisé par les associations CLUST-ER BUILD et DIGITAL&BIM Italia (voir Annexe 6- Digital&BIM Italia. Panels du concours). En mai 2021, la thèse a reçu le Premier prix de thèse "Il restauro nell'era dell'H-BIM", organisée par le Rotary Club de Rome avec le patronage de la Faculté d'Architecture et de la Faculté d'Ingénierie Civile et Industrielle de l'Université Sapienza de Rome et en collaboration avec le Département de l'Histoire, de la Conception et de la Restauration de l'Architecture, l'Ecole de spécialisation en Biens Architecturaux et du Patrimoine Fondation Almagià- ACER.

En ce qui concerne le contexte expérimental, les différentes phases de l'approche méthodologique proposée sont testées et évaluées en référence à des études de cas représentatives des bâtiments historiques et des complexes monumentaux italiens et français : en tant que partie du patrimoine artistique et culturel des deux pays impliqués dans l'accord de co-tutelle, ces sites nécessitent la mise à disposition de tout l'appareil géométrique, graphique et informatif, à des échelles de lecture multiples et au sein de systèmes d'information numériques, capables d'exprimer et de transmettre leur complexité par l'annotation de représentations 2D/3D. Chaque fois, les résultats sont évalués en tenant compte des besoins et des exigences de représentation et de restitution propres à chaque cas d'étude. Parmi les études de cas considérées, outre un certain nombre d'églises et de musées relevants du territoire italien, citons la cathédrale Notre-Dame de Paris et la Chartreuse de Pise.

De nombreux milieux sont liés au complexe monumental de la Chartreuse de Pise (ita : Certosa di Pisa), un complexe monastique dont la structure originale remonte au XIIIe siècle et qui abrite aujourd'hui le Musée de la Chartreuse et le Musée d'histoire naturelle de l'Université

de Pise. La disposition actuelle du monastère résulte de l'expansion et de la rénovation incessantes qui ont eu lieu depuis sa fondation et de nombreux espaces de la vie chartreuse - où la vie hermétique et la vie coénobitique étaient harmonieusement fusionnées -, peuvent encore être trouvés aujourd'hui.

Depuis l'année 2013, la Chartreuse de Pise fait l'objet d'un intense travail de relevé, d'analyse, de documentation, promu par l'Université de Pise et coordonné par le professeur Marco Giorgio Bevilacqua, dans le cadre du projet *Studi conoscitivi e ricerche per la conservazione e la valorizzazione del Complesso della Certosa di Calci e dei suoi Poli Museali* (fr : Études et recherches pour la conservation et la valorisation du complexe de la Chartreuse de Calci et de ses ensembles muséaux).

Le relevé architectural du complexe a été réalisé par le laboratoire ASTRO, avec l'application de méthodologies intégrées de scanner laser et de photogrammétrie.

Une autre étude de cas significative est la cathédrale Notre Dame de Paris, l'église gothique symbole de la capitale française. Après l'incendie de la cathédrale en avril 2019, le CNRS a été désigné par le Ministère de la Culture pour guider l'organisation d'un projet scientifique en soutien aux activités de restauration de la cathédrale. Parmi les neuf groupes de travail thématiques mis en place, celui sur les *données numériques*, coordonné par Livio De Luca, a pour ambition de construire un écosystème numérique. Ce dernier devrait être mis en pratique afin d'aider l'étude scientifique et la restauration de la cathédrale en intégrant progressivement données, informations et connaissances provenant de plusieurs acteurs, grâce à : la collecte et l'intégration des données existantes, la production de nouvelles données, leur partage et leur archivage permanent, leur structuration et leur enrichissement sémantique (Notre-Dame Science, 2021).

Pour les différentes études de cas considérées sur le territoire français et/ou italien, l'approche méthodologique proposée a été testée et adaptée chaque fois au cas par cas, en identifiant trois phases principales de travail: *i)* la segmentation sémantique via l'IA; *ii)* la propagation de l'information 2D/3D; *iii)* la reconstruction H-BIM, la structuration sémantique et le transfert d'annotations localisées. Il faut noter qu'on ne propose pas un flux de travail unique à appliquer à chaque étude de cas, mais plutôt l'approche générale définie et divisée en trois phases peut être adaptée, intégrée et validée à chaque fois, en fonction des besoins spécifiques de description et de documentation.

Un large éventail d'études de cas recueillies et analysées au cours des trois années du programme de doctorat est considéré pour le contexte expérimental de cette thèse, ainsi que pour la validation et la vérification ultérieure de l'approche méthodologique proposée : églises, monuments, musées répandus sur les territoires italien et français étudiés avec l'application et/ou l'intégration de différentes

techniques, par exemple, le balayage laser, la photogrammétrie par drone ou terrestre, en plus des instruments traditionnels. Sur les études de cas considérées :

- Des procédures semi-automatiques de segmentation sémantique des données 3D, telles que celles dérivées des applications les plus récentes de l'intelligence artificielle (Machine Learning et Deep Learning) sont appliquées pour classer et propager l'information à partir des données de relevé ;
- Les plateformes H-BIM constituent la base informative pour l'insertion et le stockage des informations à la fois géométriques et sémantiques ;
- Des outils de modélisation procédurale et des environnements de programmation visuelle sont utilisés pour la définition et la subdivision ultérieure de modèles 3D conceptuels.

4. Méthodologie

L'approche méthodologique proposée repose sur l'identification d'éléments pertinents et récurrents de l'objet patrimonial, qui sont annotés, récupérés et propagés de manière semi-automatique sur et entre différents types de modèles numériques, en 2D ou en 3D. Ainsi, l'information est associée et liée à chaque fois aux éléments représentatifs identifiés sur le modèle numérique, assurant ainsi le transfert et la récupération de données significatives. L'information sémantique est comme ça localisée et stockée virtuellement sur l'objet patrimonial, quel que soit le type de représentation numérique choisi et ses modifications dans le temps.

La formulation de cette approche méthodologique requiert la capacité, d'une part, de gérer l'acquisition numérique, la restitution et l'interprétation des formes géométriques, visuelles, descriptives existantes d'un bien architectural ; d'autre part, de définir un système numérique approprié pour le stockage, le transfert et l'échange de la variété d'informations produites. La définition de la méthode part de l'acquisition des données 2D/3D et se décompose en trois phases respectives :

i) segmentation sémantique via l'Intelligence Artificielle ;

ii) propagation de l'information 2D/3D ;

iii) reconstruction H-BIM, structuration sémantique et transfert d'annotations localisées.

Phase i) - Segmentation sémantique via l'Intelligence Artificielle

Dans cette phase, des systèmes semi-automatiques sont exploités pour l'interprétation des données 2D et 3D, afin de mieux caractériser la scène architecturale étudiée, en améliorant la connaissance de l'objet patrimonial par la description et la reconnaissance de ses formes, matériaux, état de conservation.

Des techniques dérivées de l'IA sont envisagées pour la construction d'outils appropriés soutenant le traitement, la gestion et l'enrichissement sémantique des objets numériques du patrimoine. En partant de l'extraction de caractéristiques appropriées, et en fournissant un ensemble pertinent de données de *training*, une intelligence artificielle est entraînée à classifier, au sein de modèles numériques : les éléments architecturaux qui composent un bâtiment historique ; la cartographie thématique des surfaces dégradées ; les cartes de dommages et la spécification des types d'intervention ; la reconstruction virtuelle des éléments perdus et l'organisation de la documentation archivistique; la définition des matériaux et leur assemblage, les techniques de construction, etc.

La phase de classification sémantique via apprentissage automatique (Machine Learning, ML) permet l'obtention de données 2D ou 3D *annotées sémantiquement,* donc plus facilement interprétables. La méthodologie est étudiée aussi bien en référence à des données 2D (photos, orthophotos, dessins, images d'archives) qu'à des modèles 3D (maillages, nuages de points) du patrimoine architectural et s'appuie sur des algorithmes de ML supervisé. En détail, le *random forest* (RF) est choisi comme algorithme de référence. En raison de l'absence d'une quantité suffisante de données d'entraînement et de la variété des bâtiments historiques, les approches d'apprentissage profond (Deep Learning, DL) sont écartées.

Phase ii) - Propagation de l'information 2D/3D

La deuxième phase de l'approche est consacrée au développement de nouvelles modalités de transfert et de propagation des annotations sémantiques, des données 2D vers 3D et vice versa.

Dans un premier temps, des méthodes d'annotation semi-automatiques sont présentées sous forme d'implémentation de la plateforme Aïoli, une plateforme d'annotation collaborative *reality-based* des modèles numérique du patrimoine qui a été développée par l'UMR MAP du CNRS: en combinant le système d'annotation hybride 2D/3D de la plateforme avec des annotations semi-automatiques (basées sur l'IA), on améliore les possibilités de propagation des données sémantiques à ce jour disponibles dans Aïoli.

Dans le but de développer de nouvelles méthodes de documentation

des objets du patrimoine, une approche collaborative et basée sur le web est développée, en s'appuyant sur l'IA pour le transfert semi-automatique des annotations 2D/3D.

Par la suite, plusieurs flux de travail sont explorés pour assurer la propagation des informations entre différents types de représentation, par exemple, des maillages, des nuages de points, des images, des orthophotos, etc. L'objectif est de fournir à un éventuel futur utilisateur une gamme de méthodes réalisables à choisir comme étant les plus appropriées pour le transfert de données, en fonction des besoins spécifiques du cas et des exigences d'enrichissement sémantique des données (passage, par exemple, de nuages de points et de maillages à des images, des cartes UV, des orthophotos et vice versa).

Des flux de travail multi-couches et multi-résolution sont ensuite proposés pour permettre le transfert de données sémantiques à différents niveaux de détail.

Phase iii) - Reconstruction H-BIM, structuration sémantique et transfert d'annotations localisées

Dans le cas de la classification d'un modèle patrimonial en ses composants architecturaux récurrents, la segmentation sémantique est considérée pour une ultérieure construction d'un modèle H-BIM.

Dans cette phase, les nuages de points classifiés sont exploités pour reconstruire, sur la base de données 3D annotées, les géométries conceptuelles de chaque classe d'éléments architecturaux détectée (par exemple, colonne, voûte, moulure, etc.).

En se référant aux traités d'architecture et en exploitant les environnements de modélisation générative, une série de règles et de conventions graphiques est identifiée pour reconstruire les géométries de référence des modèles *(template geometries)*, et pour propager ensuite ces derniers à tous les éléments appartenant à une même classe typologique.

L'opération, si effectuée sur l'ensemble du modèle, donne lieu à la réalisation d'un système d'information complet de type H-BIM, dans lequel la connexion avec le nuage de points original est préservée et où les données relatives aux connaissances peuvent être stockées.

Dans une dernière étape, on évoque d'autres possibilités d'enrichissement sémantique du modèle, en suggérant des moyens de combiner des informations brutes et même localisées sur l'objet numérique reconstruit. Certains suggestions sont fournies concernant les moyens d'association d'informations descriptives sur chaque géométrie de modèle reconstruite.

En tant que tel, le *méta-modèle conceptuel* construit pourrait être lié à

tous les autres modèles de représentation existants (maillage, nuages de points, photos, ortho-photos, dessin graphique 2D etc...), permettant même, dans une perspective H-BIM, l'association des informations à des zones localisées, afin de noter, par exemple : des données historiques, architecturales ou artistiques, la présence de fresques et de parties décoratives, des phénomènes de dégradation, des modèles de fissures, des interventions de réparation et de restauration et ainsi de suivi.

5. Structure de la thèse

La thèse est structurée comme suit : la première partie (Part I) traite de l'état de l'art sur les systèmes modernes de documentation numérique du patrimoine culturel, avec un aperçu des méthodes actuelles d'acquisition, de traitement et de modélisation des données numériques 2D ou 3D. Les apports, les limites et les perspectives d'utilisation de ces systèmes sont présentés, en opérant des analyses comparatives et en identifiant les principaux défis de recherche qui ont conduit à la définition de la méthodologie proposée.

En détail, le Chapitre 2 introduit plusieurs concepts pertinents liés à la numérisation du patrimoine qui sont largement utilisés tout au long de cette thèse : Patrimoine culturel et monuments, continuité numérique, relevé 3D, enrichissement sémantique et H-BIM. Dans le Chapitre 3, les pratiques actuelles d'acquisition, de traitement et d'enrichissement sémantique des données du patrimoine numérique sont examinées et discutées en détail, avec une analyse approfondie sur les descripteurs visuels et métriques qui peuvent être extraits des données brutes et sur les systèmes d'annotation sémantique existants, qui plus ou moins automatiques.

Dans la deuxième partie, l'approche méthodologique proposée est présentée en illustrant ses développements expérimentaux et applications à des études de cas significatives. Le flux de travail proposé permet la récupération, l'utilisation et la mise à jour des connaissances sémantiques concernant un artefact patrimonial via la consolidation/ visualisation de sa réplique numérique et le transfert des annotations. Un regard approfondi est porté sur l'application de méthodes de classification basées sur l'IA et sur l'automatisation des processus Scan-to-BIM.

En détail, le Chapitre 4 présente la logique et l'organisation général de l'approche proposée, tandis que les Chapitres 5, 6 et 7 sont consacrés à l'illustration des différentes méthodes et des résultats connexes obtenus pour les trois phases respectives du flux de travail proposé.

Les remarques finales et les développements futurs de cette thèse sont enfin exposés dans le Chapitre 8.

6. Conclusions et évaluation de la méthode proposée

Le travail de thèse propose une approche méthodologique intégrée pour le transfert, la récupération et l'échange d'annotations sémantiques de modèles numériques 2D/3D du patrimoine, basée sur la combinaison de techniques d'IA, d'environnements H-BIM et de plateformes d'annotation collaboratives *reality-based* comme Aïoli.

L'approche est conçue pour aider les restaurateurs, les conservateurs, les historiens de l'art, les architectes, les ingénieurs, les archéologues et d'autres experts travaillant sur des objets, des monuments et des sites patrimoniaux existants, qui sont impliqués dans le processus de documentation, de restauration et de mise en valeur et qui sont toujours confrontés aux problèmes de collecte, de récupération, de fusion et de mise à jour d'informations hétérogènes sur des représentations numériques du patrimoine.

L'analyse de l'état de l'art menée dans la première partie de la thèse a révélé l'existence d'une pluralité de méthodes de représentation, tant en 2D (images, orthophotos, dessins, cartes UV...) qu'en 3D (nuages de points, maillages texturés, modèles paramétriques...), pour la documentation numérique de notre patrimoine culturel et architectural. En même temps, recherche bibliographique a révélé le manque évident de méthodes efficaces et d'outils harmonisés pour l'archivage et la collecte des connaissances hétérogènes produites dans le cadre de l'étude des objets patrimoniales.

La pluralité des informations et l'existence de multiples types de représentation 2D et/ou 3D, en fait, est souvent source de la dispersion des données, et soulève la question de la *continuité numérique des données*, c'est-à-dire la nécessité d'assurer une récupération, une traçabilité, une maintenance et une interprétabilité appropriées des informations, quel que soit le type de représentation numérique choisi.

Pour combler cette lacune - et en exploitant le mécanisme d'annotation pour l'association d'informations liées à la connaissance (*sémantiques*) aux données numériques purement métriques - cette thèse propose une série de méthodes et d'outils pour supporter, de façon plus automatique, l'interprétation, la collecte et le partage appropriés des informations du patrimoine numérique. L'approche méthodologique illustrée et développée dans la Partie 2 visait donc à permettre le transfert, la récupération et l'échange d'annotations sémantiques sur des supports 2D/3D, afin d'améliorer les pratiques de documentation et de dissémination du patrimoine culturel architectural.

Dans le détail, les aspects méthodologiques ont été développés en exploitant et en combinant des données brutes issues de relevé, des systèmes de modélisation des informations du type H-BIM et des plateformes d'annotation collaborative basées sur la réalité telles que Aïoli.

L'approche, bien que nécessitant la supervision d'utilisateurs qualifiés (ingénieurs, architectes, conservateurs ou restaurateurs impliqués dans le processus de documentation), assure un haut niveau d'automatisation et permet l'association d'informations sémantiques aux représentations numériques brutes dérivées du relevé, par exemple, en termes de cartographie des phénomènes de dégradation, des matériaux, des techniques et mesures d'intervention, des différents états temporels, sur des supports 2D et 3D.

Les outils proposés envisagent l'application de logiciels sous licence ou open-source existants afin d'étudier le meilleur flux de travail possible pour l'insertion, le transfert et l'archivage semi-automatique d'annotations sémantiques sur des modèles numériques du patrimoine. Ils permettent de :

Classifier et interpréter les données brutes issues de l'arpentage, en proposant des procédures pour une gestion plus efficace des données du patrimoine numérique.

- Faciliter la compréhension et la connaissance de l'objet, par la restitution de ses attributs visuels, géométriques et descriptifs, et par l'ajout d'annotations liées à la connaissance faisant référence, par exemple, à des phénomènes de dégradation, à des informations historiques, à des altérations ou à la cartographie des matériaux, etc ;
- Permettre la connexion et l'interconnexion multimodales de différentes sources numériques (2D ou 3D) et de modèles d'un même objet du patrimoine ;
- Orienter les outils logiciels existants vers l'amélioration des systèmes Aïoli et H-BIM, et vers la gestion des plateformes d'annotation collaborative en général, sur la base des acquisitions initiales sur site ;
- Construire un modèle de référence d'un objet patrimonial, qui peut être utilisé pour la construction de systèmes informatifs de type H-BIM, pour un enrichissement sémantique ultérieur.

Les différents outils proposés pour les trois phases de la démarche pourront être adaptés au cas par cas, en fonction des besoins et exigences de restitution établis à chaque fois par l'expert travaillant sur l'objet patrimonial. l'intention n'était pas de proposer un seul flux de travail de bout en bout à appliquer à chaque fois à chaque étude de cas, mais plutôt l'approche générale définie et divisée en trois phases respectives - c'est-à-dire la segmentation sémantique via l'IA, la propagation des informations 2D/3D, la reconstruction H-BIM, la structuration sémantique et le transfert des informations localisées-, a été proposée pour être adaptée, intégrée et validée à chaque fois, en fonction des besoins spécifiques de description et de documentation.

En considérant les résultats obtenus et les expériences menées sur plusieurs études de cas du patrimoine italien et français, plusieurs remarques conclusives et limites connexes peuvent être tirées pour les différentes phases de l'approche.

Phase i) - Segmentation sémantique via l'Intelligence Artificielle

Dans la première phase, des algorithmes d'Intelligence Artificielle sont exploités pour faciliter la lecture et l'interprétation des données brutes d'un objet patrimonial. Les approches basées sur la géométrie et sur la texture permettent, à plusieurs niveaux, d'ajouter un label sémantique associé, par exemple, à la décomposition d'un bâtiment en éléments architecturaux récurrents, à la cartographie des matériaux, aux états de dégradation, etc.

Parmi elles, les approches basées sur la géométrie sont plus adaptées à la reconnaissance d'éléments architecturaux récurrents et de composants, même en tant que prélude fondamental et nécessaire aux applications de reconstruction Scan-to-BIM.

D'autre part, les approches basées sur la texture, qui reposent sur des informations de couleur, peuvent soutenir les opérations de restauration qui nécessitent une mise en correspondance sémantique et une comparaison entre les conditions successives de dégradation des surfaces et/ou des états des matériaux.

Grâce à l'application de l'IA, et en particulier du ML, les interventions humaines sont réduites dans la phase d'annotation brute, ce qui diminue l'espace pour des choix arbitraires et trop subjectifs et réduit considérablement le temps d'annotation. Un processus d'annotation semi-automatique est ainsi réalisé, en exploitant la manipulation, l'exportation et l'extraction appropriées de descripteurs géométriques et visuels (caractéristiques) à partir de données brutes en 2D ou 3D.

Cependant, la supervision de l'utilisateur, en termes de choix de l'ensemble d'apprentissage et de décision sur les classes à partitionner les données numériques, est cruciale pour déterminer le succès du processus de classification et de *labeling*. En outre, des machines de calcul à haute performance sont nécessaires pour l'extraction des caractéristiques et pour la phase d'apprentissage des algorithmes basés sur les données. Par ailleurs, des erreurs de classification peuvent se produire dans les zones limites, ainsi que dans les zones présentant un développement similaire (approches basées sur la géométrie) ou des caractéristiques ou motifs de couleur semblables (approches basées sur la texture).

La généralisation d'un même algorithme ML à d'autres ensembles de données, liés à différents types d'architecture et/ou construits à différentes périodes, et la création d'ensembles de données annotées plus vastes pour entraîner des réseaux de neurones profonds *(Deep*

Neural Networks), sont des développements futurs possibles dans ce domaine.

Phase ii) - Propagation de l'information 2D/3D

Au cours de la deuxième phase (ii), notre travail s'est concentré sur l'importance de considérer les données annotées sémantiquement (obtenues manuellement ou plus automatiquement via le processus de segmentation) non seulement comme la ressource informative d'une représentation numérique unique en soi, mais plutôt comme un point de départ pour enrichir plusieurs types de représentations du même objet architectural.

Les différents outils de transfert d'annotations proposés se consacrent à la préservation des informations d'annotation sémantique, même lors de la transition:

- vers un type de représentation différent (nuage de points, maillage, image, orthophoto...) ;
- vers des niveaux successifs de résolution numérique et/ou de détail.

Pour cette tâche, des mécanismes possibles pour la propagation des annotations des objets numériques 2D à 3D (et vice versa) sont développés, afin de garantir que l'information d'annotation, qu'elle soit automatique ou manuelle (donc, indépendamment de la façon dont elle a été obtenue), puisse être préservée dans la transition, par exemple, des images au modèle 3D, ou d'un modèle 3D plus détaillé à un modèle 3D moins détaillé et ainsi de suite.

Plusieurs méthodes d'enrichissement sémantique de représentations numériques en mode d'annotation hybride 2D/3D ont été présentées, qui tirent parti des développements les plus récents de la plateforme d'annotation collaborative basée sur la réalité Aïoli, développée par l'unité de recherche MAP CNRS/MC. L'implémentation développée pour la plateforme Aïoli consiste, en effet, en la combinaison de son système d'annotation hybride 2D/3D avec des méthodes de classification supervisée. Il a été montré, à cet égard, comment l'information peut être propagée de manière semi-automatique avec une approche d'annotation collaborative et basée sur le web.

Les outils d'annotation collaborative, permettant le transfert 2D/3D de données sémantiques, ont confirmé la plate-forme comme un outil de soutien fondamental dans la gestion et le partage multi-source de différents modèles et images 3D, liés aux artefacts du patrimoine à travers le monde.

Ensuite, des méthodes supplémentaires de transfert semi-automatique d'annotations 2D/3D ont été explorées et présentées au cas par cas en fonction des besoins et des exigences spécifiques d'enrichissement des

données sémantiques. Les flux de travail proposés contribuent à une meilleure compréhension des architectures patrimoniales complexes, ainsi qu'au partage collaboratif des données et à la gestion de l'information contenue dans les répliques numériques des monuments et des sites.

Cependant, il faut noter que leur application se réfère principalement, dans ce travail, aux relevés photogrammétriques. D'autres scénarios de transfert d'alignement ou d'annotation, impliquant par exemple l'utilisation de balayages laser, mériteraient d'être étudiés de manière plus approfondie.

De plus, dans les scénarios présentés, l'annotation manuelle des données supervisées (formation) est toujours effectuée par l'utilisateur sur un seul support, et les modes d'annotation manuelle hybrides ne sont pas encore considérés.

Phase iii) - Reconstruction H-BIM, structuration sémantique et transfert d'annotations localisées

À un stade ultérieur, l'importance de l'exploitation de données sémantiquement annotées et structurées a été démontrée en vue de la construction de représentations H-BIM .

Dans les cas où le processus de classification s'est appuyé sur l'identification de composants architecturaux récurrents et pertinents, les nuages de points sémantiquement segmentés ont donc été pris en compte pour la reconstruction d'un modèle de référence constitué de géométries *template*, selon la logique des systèmes d'information H-BIM (phase iii).

Une procédure semi-automatique a donc été proposée pour la reconstruction Scan-to-BIM, c'est-à-dire pour permettre la reconstruction de modèles d'information sur le patrimoine bâti à partir de données de levés 3D annotées.

Grâce à cette approche, la nature géométrique des composants du bâtiment a été reconstruite, la forme conçue à l'origine a été interprétée et une géométrie de référence *(template geometry)* a été identifiée et modélisée pour chaque classe, suivant la logique des familles BIM paramétriques. Une fois le processus terminé, des techniques de programmation visuelle ont permis la propagation des éléments de modèle.

Le modèle H-BIM reconstruit préserve le lien sémantique avec le nuage de points sémantiquement annoté, au niveau des classes individuelles de composants architecturaux détectés, et peut être exploité pour un enrichissement supplémentaire avec des informations non géométriques (analytiques, liées aux connaissances). Cependant,

il est à noter que le processus d'annotation, bien que simple lorsqu'il se réfère à un seul composant architectural, devient délicat dans le cas d'annotations localisées.

Sur la base de cette observation, la dernière partie du processus a introduit des possibilités supplémentaires de structuration sémantique et de transfert d'informations localisées, reposant sur la définition de procédures appropriées de *discrétisation des géométries* du modèle.

Ces dernières expériences visaient à noter, entre autres, la présence de fresques et de parties décoratives, les phénomènes de dégradation, les motifs de fissures, les interventions de réparation et de restauration, etc.

7. Perspectives de recherche

Au cours de la rédaction et du développement de cette thèse, plusieurs problèmes et questions ouvertes sont apparus. Les perspectives de recherche, qui pourraient faire l'objet d'une étude plus étendue, sont présentées dans cette dernière Section.

Définition d'un environnement de développement et d'une interface unique

A présent, un algorithme général implémentant toutes les étapes de la méthodologie proposée n'est pas encore disponible, c'est pourquoi plusieurs logiciels et outils ont été combinés, chacun étant spécialisé pour effectuer une tâche spécifique. En effet, le but de ce travail était de tester l'efficacité de l'approche proposée d'un point de vue méthodologique, afin de satisfaire la variété des processus envisagés. Les outils logiciels existants et déjà disponibles ont donc été testés, manipulés et adaptés pour répondre aux nouveaux besoins de description et de documentation avec une réponse plus directe et pratique aux exigences d'insertion, de transfert et d'archivage semi-automatique des annotations sémantiques sur les modèles du patrimoine numérique. Malgré la complication dérivée de l'utilisation de plusieurs logiciels, la procédure elle-même a été validée, démontrant que le processus peut être appliqué de manière générale.

A titre de développement futur de la recherche, les résultats de la méthodologie proposée suggèrent certainement la création d'une interface dédiée qui puisse minimiser l'utilisation de différents logiciels, dans laquelle les différentes phases de l'approche méthodologique proposée pourraient être réunies et mises à la disposition de la

communauté des experts du patrimoine.

Utilisation de descripteurs hybrides 2D/3D pour améliorer la phase d'annotation manuelle dans Aïoli.

La plateforme Aïoli pourrait être implémentée avec d'autres outils permettant à la fois des méthodes d'annotation hybrides 2D/3D sur des supports multiples, la reconstruction de géométries modèles et/ou des possibilités d'annotation semi-automatiques supplémentaires.

En particulier, les implémentations prévues pour être appliquées sur la plateforme dans un futur très proche concernent les aspects suivants:

- développement d'outils et de méthodes assistant l'utilisateur dans la phase d'annotation manuelle 2D/3D.

Etant donné que Aïoli permet actuellement l'annotation manuelle uniquement sur les images et non en 3D, il est envisagé de développer des approches hybrides combinées de segmentation 2D/3D qui tirent pleinement parti de chaque support d'annotation, de telle sorte qu'un utilisateur éventuel puisse effectuer l'annotation soit en 2D (par exemple, sur l'image), soit en 3D (par exemple, sur le nuage de points), en choisissant le support le plus facile à annoter.

Par ailleurs, on estime à cet égard que des caractéristiques géométriques ou radiométriques (*descripteurs*) associées aux données - tant en 2D qu'en 3D - pourraient être utilisées pour faciliter l'annotation d'objets de manière semi-automatique et supervisée. Par exemple, dans le cas des images, cela pourrait se traduire par le développement d'une sorte de baguette magique, aidant à la sélection de régions d'intérêt dans lesquelles un certain descripteur apparaît avec des tons et des couleurs homogènes.

- Utilisation des images annotées pour entraîner des réseaux neuronaux profonds et construire des bibliothèques d'images segmentées pour le patrimoine bâti.

Une partie de ce travail de thèse s'est concentré sur la possibilité de transférer les informations de segmentation sémantique du nuage de points vers l'ensemble des images existantes, en exploitant le traitement photogrammétrique et la fonction de propagation hybride 2D/3D. Dans le cadre d'un développement futur de ce travail, si un nombre suffisant d'images annotées est fourni, ces dernières pourraient être utilisées en vue de la construction de bibliothèques d'images sémantiquement segmentées, pour entraîner les réseaux neuronaux profonds à reconnaître, par exemple, des composants architecturaux sur un ensemble d'images de bâtiments existants.

- Amélioration des repères existants pour la construction de réseaux neuronaux profonds

Toujours en tenant compte des développements en matière de la construction de réseaux neuronaux profonds, les ensembles de données 3D sémantiquement annotées présentés dans cette thèse pourraient être utilisés pour mettre en œuvre des repères existants d'objets et de sites du patrimoine numérique annotés, comme celui de 3DArCH (Matrone et al., 2020). En effet, la construction d'un volume significatif de données 3D annotées (en particulier, des nuages de points) pourrait améliorer l'application des stratégies DL au domaine du patrimoine, à ce jour non encore applicable en raison de la quantité limitée de données disponibles.

Niveaux de détail dans les environnements H-BIM et intégration d'informations localisées

En ce qui concerne la phase de reconstruction Scan-to-BIM, la dernière partie du Chapitre 8 a introduit le problèmes liés à l'insertion d'informations localisées dans les géométries de modèle *template* reconstruites.

En plus de cela, il faut noter que des niveaux successifs de classification pourraient également être inclus dans l'approche de reconstruction Scan-to-BIM. La méthode de segmentation hiérarchique à plusieurs niveaux, qui a été proposée ici en référence aux nuages de points, pourrait en effet être exploitée pour la construction de modèles, à différents niveaux de détail et de résolution.

Les avancées futures du présent travail peuvent donc prévoir l'application de l'approche de reconstruction Scan-to-BIM même au cas multi-niveaux, c'est-à-dire en considérant des niveaux successifs de classification sémantique du modèle 3D.

Extension à d'autres études de cas

Les principes introduits dans ce travail pour l'insertion, l'organisation et le transfert d'informations sémantiques ont été testés uniquement dans le cas de bâtiments dont la composition architecturale est caractérisée par des éléments reconnaissables et récurrents. L'identification des classes typologiques, en effet, est une prémisse fondamentale pour la décomposition du bâtiment par ML.

A cet égard, plusieurs problèmes peuvent probablement surgir de l'application de cette approche à d'autres types d'artefacts - par exemple, des sculptures, des vestiges archéologiques, des peintures - et à des bâtiments plus diversifiés - par exemple, faisant référence à différents types de construction ou édifiés à différentes périodes.

L'extension de l'approche proposée à d'autres études de cas est enfin un développement possible de cette thèse.

Acknowledgements

I gratefully acknowledge Prof. Roberta Spallone, Prof. Andreas Georgopoulos and Prof. Manuel A. Ródenas López, for having accepted to be part of the final committee, for their availability to review this work and for the time they dedicated to it.

I express my deep sense of thanks and gratitude to Profs. Gabriella Caroti and Andrea Piemonte. You guided me all along these three years, supporting me and letting me cope serenely and positively all along my Ph.D..

I thank Gabriella for welcoming me in the ASTRO laboratory of the University of Pisa with an almost maternal affection, for her precious advices, prompt inspirations, constant guidance and teachings. I thank Andrea for his constant availability, for his dedication and keen interest, for guiding me so positively all along the development of this work. I am proud of being your student and continuing to work alongside you. I also thank you for the opportunity you gave me to undertake the *co-tutelle* agreement with the French institutions, once again trusting me and encouraging me.

I am extremely thankful to Prof. Marco Giorgio Bevilacqua, for his constant availability and for making me love architecture and conception since I was an undergraduate student. I am happy to work alongside you and I acknowledge you for your overwhelming attitude to help, for your constant feedbacks and precious guidance.

I owe a deep sense of gratitude to Dr. HDR. Livio De Luca and to Prof. Philippe Véron, for having accepted me as a visiting Ph.D. student and for encouraging me in developing this research work. I sincerely thank Livio for welcoming me in Marseille, for making me feel at home in a foreign country, in a familiar and friendly environment, and for introducing me to the studies and work of the MAP laboratory in the digital heritage domain. I gratefully thank Philippe Véron, for his valuable methodological advice that allowed me to pursue the study, for giving me the opportunity to undertake this project, also involving French institutions.

I would like to thank Kévin Jacquot, for his precious comments and remarks on this work, and for agreeing to co-supervise the thesis.

I acknowledge the institution Université Franco-Italienne, for providing external support for the development of this research work.

I sincerely thank all the fellow lab mates I had the opportunity to meet during these years.

At the ASTRO laboratory, Andrea Bedini, Isabel Martinez, Federico Capriuoli, Benedetta Binetti, Gabriele Pacini, Noemi Leonardi, Jacopo Raffaelli. At the MAP laboratory, Ariane Neroulidis, Adeline Manuel, Anthony Pamart, Antoine Gros, Eloi Gattet, Isabelle Cao, Laura Willot, Laurent Bergerot, Menehould Chataignier, Miorra Rabefandroana, Noémi Renaudin, Pascal Bénistant, Roxane Roussel, Violette Abergel, Jean-Yves Blaise and Iwona Dudek.

I thank you all for your stimulating discussions, for your friendship, for all the good times spent together, and for always being present, supporting and encouraging me. Your presence has been and still is so valuable to me, you made me feel beloved both during the times spent in Italy and in France.

Finally, I cannot forget to thank my family for all the support and affection in these full years, and Simone for his unconditional love.

For all the beautiful moments you have brought me over the years and for those yet to come.