



Alma Mater Studiorum - Università di Bologna
Dottorato di Ricerca in Disegno dell' Architettura
Scuola di Dottorato in Ingegneria Civile ed Architettura
XXVI Ciclo di Dottorato

A Semantic-Based Framework for Digital Survey of Architectural Heritage

Presentata da: dott. Zheng Sun

Coordinatore Dottorato: prof. Annalisa Trentin

Relatore: prof. Fabrizio Ivan Apollonio

Correlatore: prof. Marco Gaiani

Settore Concorsuale: 08/E1 - Disegno

Settore Scientifico Disciplinare di afferenza: ICAR/17

Esame finale anno 2014

Acknowledgement

I would like to express my thanks to Fabrizio Ivan Apollonio and Marco Gainai, my professors who guided my research and reply my questions with great patience. The research would not develop smoothly without the technical support of Giovanni Bacci and Andrea Ballabeni. Massimo Ballabeni offered me a lot of help in photographing. I received kind suggestions from Maristella Casciato, Francesco Ceccarelli and Francesco Saverio Fera.

I also would like to thank to my family and friends for their support in my life.

Index

Abstract.....	III.
Part 1. Knowledge-based Modeling in BIM platform.....	1.
Chapter 1. Treatise: System of knowledge.....	4.
1.1 Literal origin.....	4.
1.2 Canon of orders.....	5.
1.3 Formal composition.....	7.
Chapter 2. Shape grammar.....	9.
2.1 Parametric primitives.....	9.
2.2 Semantic segmentation.....	10.
2.3 Sweep operation.....	13.
2.4 Bottom-Top Process.....	14.
Chapter 3. BIM-based modeling and management.....	17.
3.1 Object-based parametric modeling.....	18.
3.2 Workflow of object modeling.....	19.
Chapter 4. Case study: Palazzo Barbaran da Porto.....	25.
4.1 Semantic organization.....	25.
4.2 Data enrichment and filter.....	27.
4.3 Interoperability and related database.....	28.
Chapter 5. Summary.....	30.
Reference.....	31.
Part 2. Image-based modeling.....	33.
Chapter 1. Introduction.....	34.
1.1 Range-based modeling and image-based modeling.....	34.
1.2 The division of photogrammetry and computer vision.....	34.
Chapter 2. Evaluation of SfM packages.....	36.
2.1 State of art.....	36.
2.2 Overview of SfM packages.....	36.
2.3 Consideration of SfM features.....	37.
2.3.1 User manipulation.....	38.
2.3.2 Metric accuracy.....	38.
2.3.3 Application to complex objects.....	38.
2.3.4 Efficiency.....	39.
2.4 Experiment.....	39.
Chapter 3. Workflow of image-based modeling.....	44.
3.1 Camera calibration.....	44.
3.1.1 Targetless camera calibration.....	44.
3.2 Image acquisition.....	44.
3.2.1 On- siteCamera settings.....	44.
3.2.2 Camera configuration.....	45.
3.2.2.1 Parallel-axis network.....	45.
3.2.2.2 Convergent network.....	45.
3.2.2.3 Divergent network.....	45.
3.2.3 Instance of failure.....	47.
3.2.3.1 Image blur.....	47.

3.2.3.2 Low resolution.....	47.
3.3 Image processing.....	48.
3.3.1 Color calibration.....	48.
3.3.2 Noise reduction.....	50.
3.3.3 Wallis filter.....	50.
3.4 Feature correspondence.....	51.
3.2.1 Feature detection.....	51.
3.2.2 Feature matching.....	52.
3.2.3 Incremental SfM.....	53.
3.5 Dense image matching.....	54.
3.5.1 State of art.....	54.
3.5.2 Experiment.....	55.
3.6 Further processing.....	57.
3.6.1 Surface reconstruction.....	57.
3.6.2 Texture mapping.....	59.
3.6.3 Segmentation.....	60.
Chapter 4. Case study.....	63.
4.1 Case study 1: Facade of Palazzo Albergati.....	63.
4.2 Case study 2: Portico.....	70.
Reference.....	75.
Part 3. Integrating ideal model with as-built model.....	81.
Chapter 1. Semantic organization of BIM.....	82.
1.1 Three types of family.....	82.
1.2 The logic of host.....	83.
Chapter 2. Towards As-built modeling.....	84.
2.1 State of art.....	84.
2.2 A semantic-based integrated modeling approach.....	84.
Chapter 3. Case study.....	87.
3.1 Case study 1: The structural framing of entrance hall.....	87.
3.2 Case study 2: The vault of portico.....	89.
3.3 Case study 3: Facade of Palazzo Barbaran.....	91.
Reference.....	95.
Summary: Strength, limitation and suture development.....	98.

Abstract

The research aims at developing a framework for semantic-based digital survey of architectural heritage. Rooted in knowledge-based modeling which extracts mathematical constraints of geometry from architectural treatises such as *The Four Books on Architecture*, as-built information of architecture obtained from image-based modeling is integrated with the ideal model in BIM platform. The knowledge-based modeling transforms the geometry and parametric relation of architectural components from 2D printings to 3D digital models, and create large amount variations based on the shape grammar in real time thanks to parametric modeling. It also provides prior knowledge for semantically segmenting unorganized survey data. The emergence of SfM (Structure from Motion) provides access to reconstruct large complex architectural scenes with high flexibility, low cost and full automation, but low reliability of metric accuracy. We solve this problem by combing photogrammetric approaches which consists of camera configuration, image enhancement, and bundle adjustment, etc. Experiments show the accuracy of image-based modeling following our workflow is comparable to that from range-based modeling. We also demonstrate positive results of our optimized approach in digital reconstruction of portico where low-texture-vault and dramatical transition of illumination bring huge difficulties in the workflow without optimization. Once the as-built model is obtained, it is integrated with the ideal model in BIM platform which allows multiple data enrichment such as material, structure, construction phase, and design options. In spite of its promising prospect in AEC industry, BIM is developed with limited consideration of reverse-engineering from survey data. Besides representing the architectural heritage in parallel ways (ideal model and as-built model) and comparing their difference, we concern how to create as-built model in BIM software which is still an open area to be addressed. The research is supposed to be fundamental for research of architectural history, documentation and conservation of architectural heritage, and renovation of existing buildings.

Key words

Knowledge-based modeling, Image-based modeling, BIM, as-built model

Part One. Knowledge-based modeling in BIM platforms

Knowledge-based modeling is founded on the treatises of architecture, and transforms the semantic composition and their ratio to 3D model with mathematical constraints. Stemmed from the ideal Antiquity, the treatises served as a knowledge system that set up the grammar of classical language and guided the rules of architectural practice. Treatises such as Andrea Palladio's *I Quattro Libri dell'Architettura* (The The Four Books on Architecture) exerted profound influence in Europe since its publication in 1570 and later in North America. Palladio's treatment on villas' plan was studied for its similarity with parametric shape grammar (Stiny et al., 1978). *The Four Books on Architecture* is regarded as a system of knowledge from which rules of formal composition are employed to generate vast amounts of variations. Another analogy between the treatises and 3D model is both of them are actually medias of representation (Mitchell et al., 1995) with different dimensions (2D and 3D), graphic modes (Wireframe, hidden line and shaded) (figure 1.1) and projection methods (perspective and orthogonal).

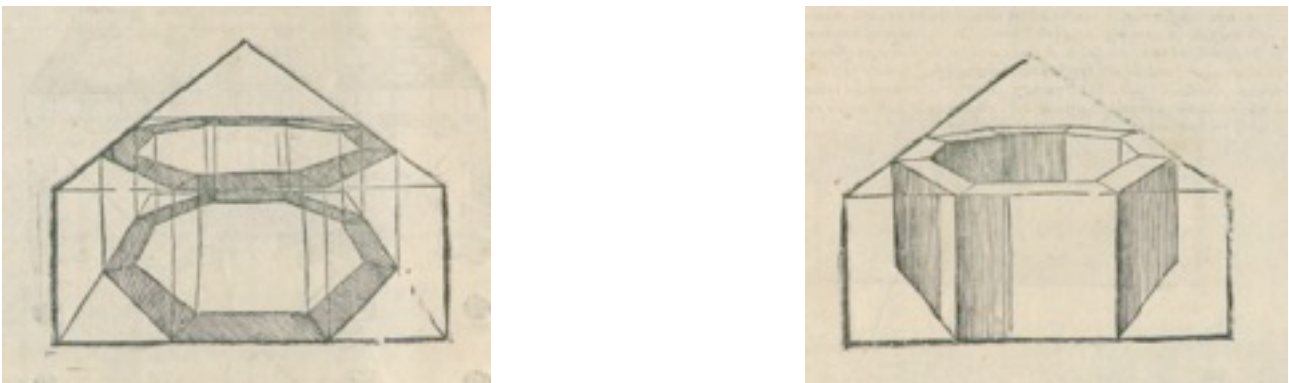


Figure 1.1 Perspective of a 8-polygon ring in wireframe way and shaded way after Serlio's *Tuttel'opere d'architettura er prospettiva di Sebastiano Serlio Bolognese*, Venice, 1619

The strength and limitation of knowledge-based modeling should be discussed with reference to the purpose that it is supposed to satisfy (El-Hakim, S. and Beraldin, J.A., 2002). In the light of geometric authenticity, image-based modeling (Remondino et al, 2006) and range-based modeling (Bernardini et al, 2002) are generally employed, since knowledge-based modeling represents an ideal state, other than actual state with real-life information, such as irregularities, missing elements and alterations (Gainai et al., 1999). Knowledge-based modeling has advantages, however, if following aspects are taken into account (chart 1.1):

1. Digital camera and laser scanner record only the surface of the object. Construction details behind the surface can be detected from historical data (Murphy et al., 2007). Similarly, historical data can be complementary when objects are out of the coverage of digital survey.
2. The data collected from the survey should be extracted into different levels to ease the further retrieval from the data repository (Manferdini et al., 2008). Thus, the semantic classification and hierarchical organization according to architectural drawing conventions related to the different historic periods is necessary (De Luca et al., 2005).
3. The treatises are by no means only collection of ancient surveys, but also consist of composition rules of classical architecture as well its practice. The algorithm developed from the study of shape grammar can be used to generate similar geometrical results automatically (Stiny et al., 2007).

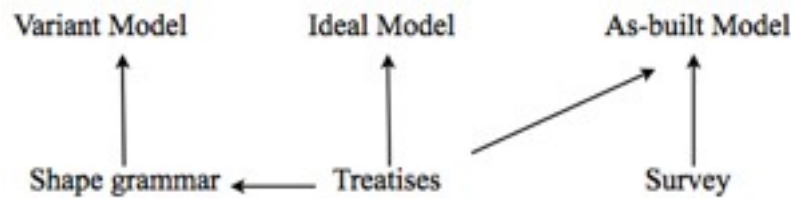


Chart 1.1 Translation from knowledge-based approach to other approaches

In the domain of 3D documentation of architectural heritage today, various modeling platforms are employed in knowledge-based modeling, such as Geometrical Descriptive Language (GDL) of ArchiCAD (McGoverna et al., 2011) and C++ in Maya (De Luca et al., 2007). The common methodology, however, is much similar: 1). Modeling parametric components, such as columns, windows, pediments, etc., according to classical architecture treatises. 2). Collecting geometric data of buildings' actual state with laser scanner or photogrammetry. 3). Mapping predefined components onto points cloud and adjust automatically.

In the framework of this workflow, the paper presents a method based on Revit 2014. It retains parametric features without any embedded language, and components are inherently semantic which ease the segmentation process. The BIM platform supports additional attributes assignment to the components. The simulation and interoperability are manipulated in object level instead of model level. Revit API (Application Programming Interface) provides access to customize and optimize the modeling process and the management of database. Multiple information are supposed to translate among different interfaces to better conserve and share the 3D model.

This chapter represents the mechanism of knowledge-based modeling in BIM platforms. The classical architecture treatises are studied in the first part. The scope is not historical, but stem from the practical aim of grammar development. In the second part, the grammar shown in part one will be developed to the logic and language that is readable to computer. The third part show the common workflow of treatise-based components modeling in BIM platforms. A case study of Palazzo Porto da Barbaran, the work of Palladio, is provided in the fourth part. In the end, summary, problems and future work are to be addressed.

Chapter 1 Treatises: System of knowledge

Classical architecture has always, even in ancient times, depended on precedents and therefore on written treatises. Vitruvius himself indebted to ancient authors, and the later architects and theorists such as Alberti, Palladio and Claude Perrault to a great extent depended on Vitruvius. Before the advent of multi-medias, architectural treatises are one of the most important approaches for knowledge broadcasting.

It is difficult to specify the criteria of treatise in terms of content, number of illustrations and motivations of authors. Vitruvius was probably not a notable architect in his time, and he spent much content depicting military facilities in his treatise. Alberti, who drew inspiration from Vitruvius's text, provided few illustrations for his text. In Vignola's patterns books about orders, however, few texts were written by the author. Vitruvius and Alberti focused on ascetic principles such as ratio and rhythm, while Palladio provided a retrospective exhibition of his own design in his well-known treatise. 18th century French theorists were enthusiastic to compare the tiny variations of orders with regard to precedent treatises and ancient ruins.

The evolution and spread of treatise in time and space, the difference of versions as well as the author's different motivations to write and to draw are out of the scope of the paper. The treatises of classical architecture are regarded as a body of knowledge from which useful data could be retrieved and stored in BIM platforms. It is still necessary, however, to have a basic knowledge of the most influential treatises to understand the grammar of classical architecture.

1.1 Literal origin

Vitruvius's *De architectura* is the only surviving written work about architecture from antiquity. Written in 1st century BC, it was dedicated to the Emperor Augustus. Although earlier Greek treatises had been cited by Vitruvius, none of them were survived till Middle Ages. Since the manuscript of *De architectura* was rediscovered in 15th century, it became the major inspiration for Renaissance architecture. Until Serlio's treatise appeared, it was the various versions of Vitruvius's work that dominated the discourse of architecture. Prominent architects such as Palladio and Claude Perrault were involved in the edition and translation of Vitruvius's work. Palladio wrote in the Foreword of his *The Four Books on Architecture* that: "Since I always held the opinion that the ancient Romans, as in many other things, had also greatly surpassed all those who came after them in building well, I elected as my master and guide Vitruvius, who is the only ancient writer on this art (Palladio, 1965)." Due to the lost of the original illustrations, however, Renaissance authors made illustrations partly with reference to Vitruvius's text and partly from surveys of Roman Ruins. This might be one of the origins of numerous Parallels of orders in 18th century France. *De architectura* consisted of 10 parts in a wide range of topics: Book 1: General rules of town planning, architecture and civil engineering; Qualifications of architect; Book 2: Building material; Book 3 and Book 4: Temples and orders; Book 5: Civil buildings; Book 6: Domestic buildings; Book 7: Pavement and decorative plasterwork; Book 8: Water supplies and aqueducts; Book 9: Sciences influencing architecture - geometry, mensuration, astronomy, sundial; Book 10: Use and

construction of machines - Roman siege engines, water mills, drainage machines, Roman technology, hoisting, pneumatics.

The next notable work about architecture after Vitruvius was Leon Battista Alberti's *De re aedificatoria* (Ten Books on Architecture). It was also the first treatise in the age Humanism. Alberti apparently drew inspiration from Vitruvius, for he divided his work exactly the same number as Vitruvius did and omitted illustrations. Leon Battista Alberti (1404-1472) was the most celebrated and influential classical scholar of his age. Born in Genoa, he was educated in Padua and Bologna and admitted to the priesthood. His interests and written works ranged from architecture to painting, sculpture, Italian grammar, cryptology, geography and even household administration. He was also a practicing architect, painter and playwright, but his great strength was in his scholarship and writing.

1.2 Canon of orders

From Serlio's notable plate of five orders to the Parallel of orders in 18th century French architectural discourse, the compositional rules of orders evolved to not only a canon of architecture, but a issue related to ideal antiquity, debate between Greek language and Latin language and national patterns.

The connection of Entablature and Column is a matter of load transfer. No matter how the geometry varies, the capitals of Doric, Ionic and Corinthian are all transition of beam (rectangular) and column (circle). Vitruvius's text and Roman ruins gave the orders ambiguous meanings, to which metaphors and attributes were assigned. The five species of order have different social meanings that correspond to human figures and building function. Besides the well know example of gender attributes of Doric, Ionic and Corinthian, they were thought to be fit to house of craft man, intellectuals and nobles. Francois Blondel, 17th century French architect, contributed his interpretations to the order: Tuscan=gigantic; Doric=Herculean; Ionic=womanly; Composite=heroic; Corinthian=virginal (Kruft H.W., 1994). Serlio claimed that when the orders were used on the façade of church, the character of the saints or virgins to whom they were dedicated should be taken into account (Summerson J., 1997).

Vitruvius depicted four orders- Tuscan, Doric, Ionic and Corinthian- in his book. Alberti added the fifth order-Composite-a combination of Ionic and Corinthian, and ascribed columns to the category of ornament instead of structures, which was one of the central problems of Renaissance architecture (Wittkover R., 1949). It was Sebastiano Serlio (1475-1553/1555) who put the five orders side by side on his well known plate (figure 1.2) that spread all over the civilized countries. The way Serlio arranged the orders was analogous with the verbs in the grammar of Latin language. Serlio founded the tradition of showing orders in a parallel way. Lots of famous architects like Vignola, Scamozzi, Perrault made similar plate in their works. Another great innovation of Serlio was his emphasis on practical drawings instead of abstract theoretical discussion among intellectuals as Alberti did. Serlio's orders were elaborated and enriched in the preceding centuries in France, German and England.

Vignola's *Regola delli Cinque Ordini d'Architettura* (The Five Orders of Architecture) was one of the most widespread pattern books with more than 250 editions. It might be improper to call it

treatise, for the illustrations much outweighed the texts in the book. This was a different manner from its predecessors. Vignola continued the canon of five orders after Serlio, and popularized it with a set of rules that can be understood by average mind. He took measurements from buildings of Antiquity, thus he set the rules of Doric from the Theatre of Marcellus in Rome. Only for the Tuscan did he use Vitruvius's version, as he knew no instance of Antiquity in Rome (Kruft et al., 1994). Vignola introduced module as the unit of measurement to an universal method of calculation that can be easily applied. His ratios of module to column height for various species were as follow: 14 (Tuscan), 16 (Doric), 18 (Ionic), 20 (Corinthian and Composite). Each species of order were shown in five ways of application: 1) the colonnade, 2) arcade, 3) arcade with pedestal, 4) individual pedestal and base forms, and 5) individual capitals and Entablature forms.

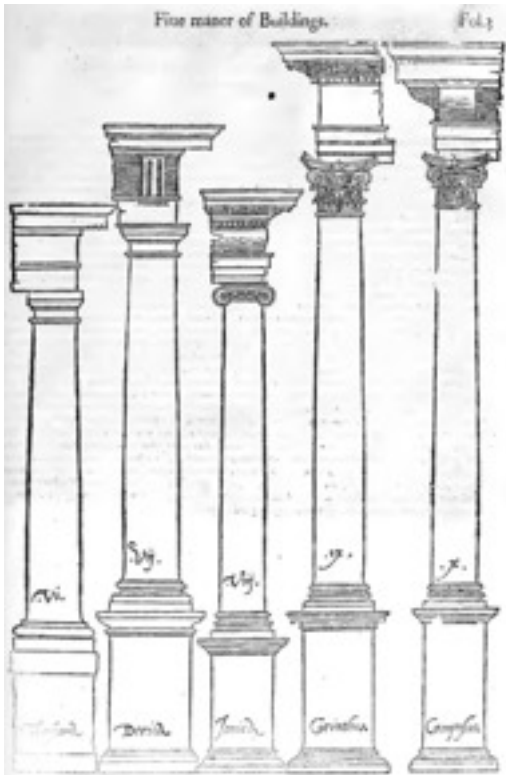


Figure 1.2 Parallel of the five species of orders after Serlio's *Tuttel'opere d'architettura er prospettiva di Sebastiano Serlio Bolognese*, Venice, 1619

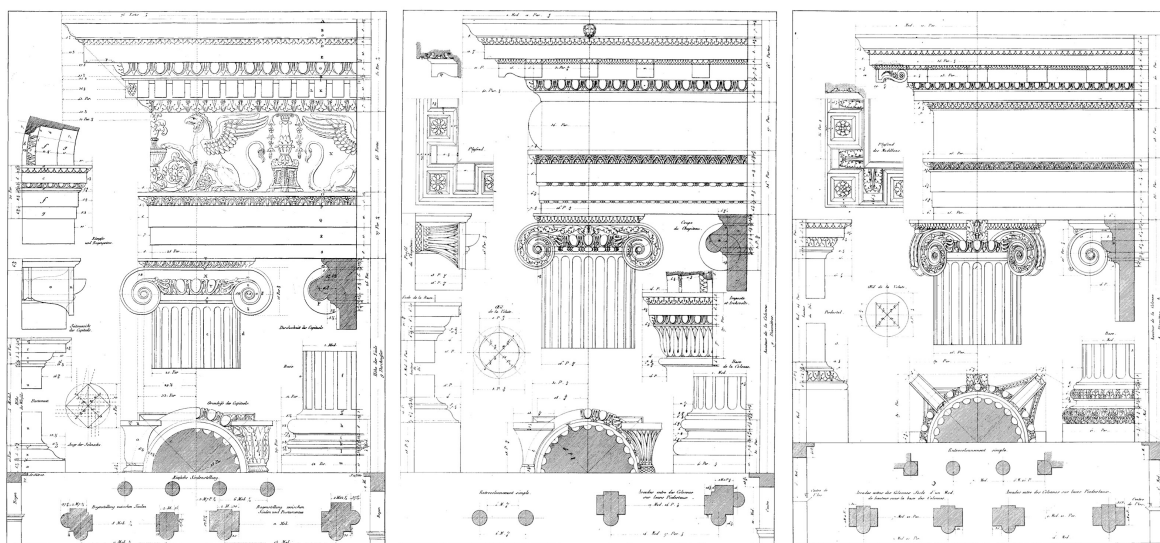


Figure 1.3 Parallel of Ionic order of (from left to right): Vignola, Palladio, Scamozzi. Collected in Rattner D., *Parallel of the classical orders of architecture*, Acanthus, New York, 1998

18th century France witnessed the bloom of “parallels of orders”. Authors of treatises at that time were able to take measurement of ancient ruins by themselves not only in Rome, but also in Greece and Turkey. It enabled them to doubt the tradition that formed after Vitruvius. Chambray (1606-1676) in his *The parallel of the ancient architecture with the modern: in a collection of the principles authors who have written on the five orders* compared in pairs the proportions of orders in previous treatises with regard to antique ruins, and concluded that the three Greek orders- Doric, Ionic and Corinthian contained the principles valuable for architecture, while the two Latin orders- were not dealt with Vitruvius’s rules. The canon of order reached a climax through the systematic research of orders. It was not only the difference among Renaissance authors were analyzed (figure 1.3), but the orders from the ruins of Greek temples (figure 1.4) were compared.

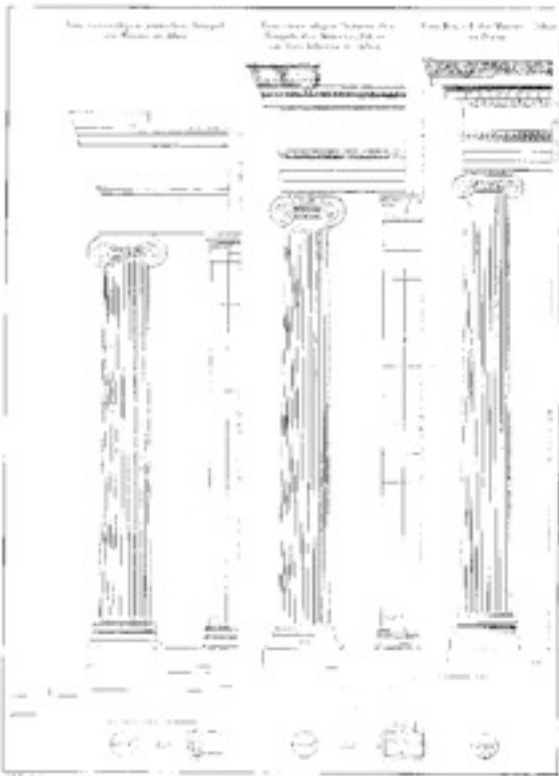


Figure 1.4 Parallel of Ionic orders based on survey of Greek temples (from left to right): Temple on the Ilissus, Erechtheum on the Athenian Acropolis, Temple of Athena at Preine. Collected in Rattner D., *Parallel of the classical orders of architecture*, Acanthus, New York, 1998

The parallels of orders not only supply abundant historical data for modeling, but also touched the essence of parametric modeling in the light of data collection and presentation. Parametric modeling builds a flexible framework instead of fixed geometry. A database laid on the foundation of the parallels of orders will make the data retrieval process easy and rapid.

1.3 Formal composition

Palladio is one of the most influential architects in the history of architecture. Although he designed most of his works in Veneto, Italy, the Palladian architecture was spread over the North America

until 18th century. Born in Padua in 1508, he was not trained as a painter or sculptor in his youth like Bramante, Raphael and Michelangelo, nor a well-rounded humanist like Alberti, but a stonemason. Besides written work, Palladio's achievement as architecture is more notable than Serlio, Alberti and Vitruvius.

The widespread Palladian architecture is to much extent ascribed to his well-known treatise. *The Four Books on Architecture*. In the first book, Palladio described the modular formalization of the orders and other architectural components, basic principles of architecture and the choice of materials. The drawings of Palladio's own design for palaces, villas, public buildings and bridges were exhibited in the second and third book. The last book consisted of the surveys of surviving Roman ruins which he had studied most closely. Palladio provided the majority of images with measurements which are mathematically related base on modules. Therefore, it is feasible to employ the plans, elevations and details in the book to build accurate 3D model of Palladio's architecture.

Furthermore, *The Four Books on Architecture* is structured as a system of knowledge (Baldissini, 2008), offering a concise and clear grammar to its readers. Written in a different way from the precedent architectural treatises, Palladio did not limit himself in the theoretical framework of Vitruvius as Albert, nor directed most attention on illustrations with few texts as Vignola. In the way that the texts and illustrations are balanced, he was similar with Serlio, but inscribe dimensions directly on the plates. Unlike Serlio who designed few buildings in his life, Palladio designed numerous churches, palazzo, especially more than thirty villas, most of which were collected in the *The Four Books on Architecture*. Palladio not only set the rules of architectural composition, but also provided lots of instances to which the rules were applied.

The Four Books on Architecture presents two important properties for computer applications: they show the Palladian buildings theoretical foundation as complete system, and they demonstrate a rigid formalism that well fits computer features and techniques (Wittkover, 1949). For these reasons many methods and techniques were developed during the last 30 years, mainly various generative techniques, such as Palladian villa production with 3D printer (Sass, 2007) and generative algorithm applied to Villa Rotonda (Park, 2006).

Jean-Nicolas-Louis Durand(1760-1834) founded another possibility of formal composition in his *Precis des lecons d'architecture (Precis of the lectures on architecture)*. Published in 1802, the book proved the astonishing potential from grid systems with regard to geometric volumes instead of space. Durand employed 2D illustrations including plans, elevations and sections with uniform scale instead of perspectives to represent the buildings. The way Durand presented his lecture maybe related to the fact that his students in Ecole Polytechnique were engineers rather than architects.

Chapter 2 Shape Grammar

Euclid's elements are the primitives to understand, segment and represent in classical architecture, as today's nonlinear architecture depends on fractal geometry. Le Corbusier wrote in his *Towards a new architecture*: "Egyptian, Greek or Roman architecture is an architecture of prism, cubes and cylinders, pyramids or sphere: the Pyramids, the Temple of Luxor, the Parthenon, the Coliseum, Hadrian's villa." (figure 1.5) The complicated 3D objects are composed by primitives. The primitives are resulted by a set of formal composition and geometric operation that could be shapes with fewer dimensions, such as surfaces, lines and points.

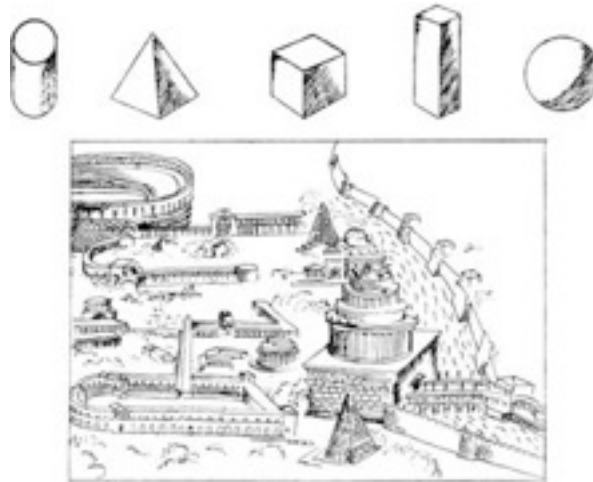


Figure 1.5 Composition of volumes after Le Corbusier, *Towards a new architecture*

2.1 Parametric primitives

We can treat the geometry of classical architecture as formal composition of primitives with certain rules. The primitives include solids and voids. The solids refer to structures and ornaments such as columns and cornices. The voids include space such as rooms, corridors and courts. Classical architecture has its unique geometric primitives and rule, just as vocabularies and grammar in a language. Classical architecture's vocabulary is composed by Doric order, round arch and triangular pediment, while the Gothic pointed arch and flying buttress.

We continue this partition by decomposing the 3D shape into a set of 2D shapes and 1D shape. For example, the shaft of an order is an approximate cylinder formed by extruding a circle along a path. In this way we get two primitives: a circle and a line.

Cylinder (Circle, Height)

If we go further, we can decompose most elements into atoms that cannot be further divided. For each atom, dimensions could be replaced by variables. We can set them as parameters with adjusting values.

Circle [Center (x,y,z), Radius]

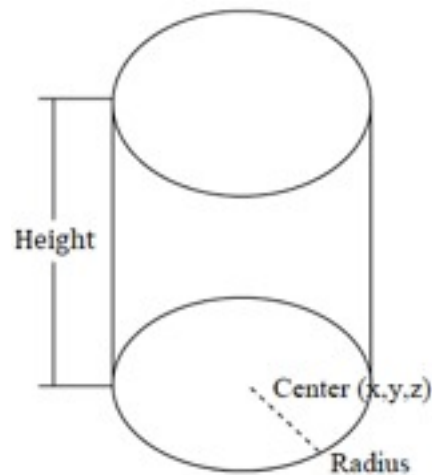


Chart 2.2 Essential parameters to define a cylinder

In this case, we could control the cylinder's form by setting constraints to a parameter or to the ratio of two parameters. If we lock the 'z' value of the center point, for example, it would remain in xy plane. If we lock all of the three values of 'x, y, z', we actually pin the point. The length of radius and height could be constants that always remain their values. If necessary, we can also constraint the ratio of height and radius by setting a formula.

$$\text{Height}=\text{Radius}*4$$

Different species of orders have different ratio is this formula.

$$\text{Height}=\text{Radius}*n$$

Ionic orders usually have larger 'n' than Doric orders, but smaller than Corinthian, for Corinthian orders are the most slender one in the order parallels. Even in the same species the ratio could be different. Vignola assigned 14 to 'n' in his treatise, while the number in Palladio's Doric could be 14 or 15.

Similarly, the capital and base of the Doric order, the other species of orders, and components such as window can all be defined with approximate primitives. By variables replacement, we are able to describe them flexibly. Further refinement could be developed according to details from the treatises. The enormous data collected in the treatises especially "Parallels of orders" provided us sufficient information in this aspect.

2.2 Semantic segmentation

As we know, an order consists three part (Tzonis et al., 1987) 1) the Entablature, a horizontal member above the column; 2) the column, a long vertical cylindrical member; 3) and the crepidoma or stylobate, a stepped platform on which the column rests, or a pedestal, a prismatic member under the column. (figure 1.6)

Order (Entablature, Column, Crepidoma)

Each of these members could be further divided into three parts. The Entablature is articulated into three members: 1) the cornice, the uppermost member projecting in the form of a continuous eave; 2) the frieze, a band made of blocks resting on the architrave below; 3) and the architrave, also made up of blocks, which span the distance between two columns and rest on the capitals. The column consists of 1) the capital, the topmost member; 2) the shaft, the middle part; 3) and the base, the lower part. The base is usually absent in Doric order. When there is a pedestal between the column and stylobate, it too is divided into 1) the cornice, projecting in the form of a continuous eave; 2) the dado, a block of varying height; 3) and the pedestal base.

Entablature (Cornice, Frieze, architrave)

Column (Capital, Shaft, Base)

Pedestal (Cornice, Dado, Base)

The tripartition continues to be applied if we divide these members until we get atoms that cannot be divided any more. These atoms are the basic words that form the vocabulary of classical architecture. We can find the grammars in different treatises to compose the sentences with these words.

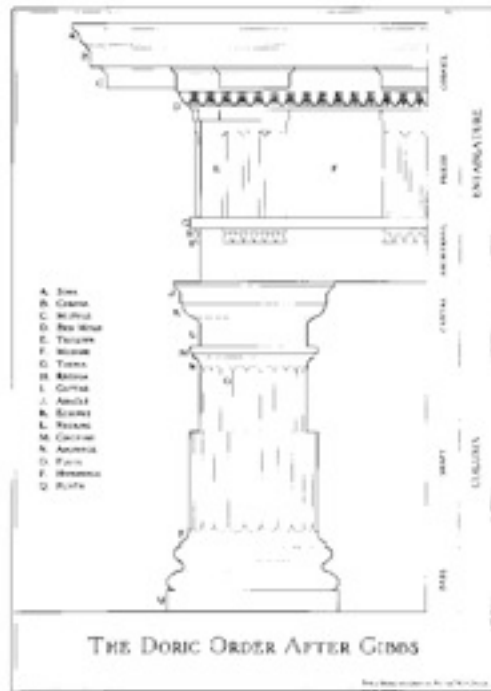
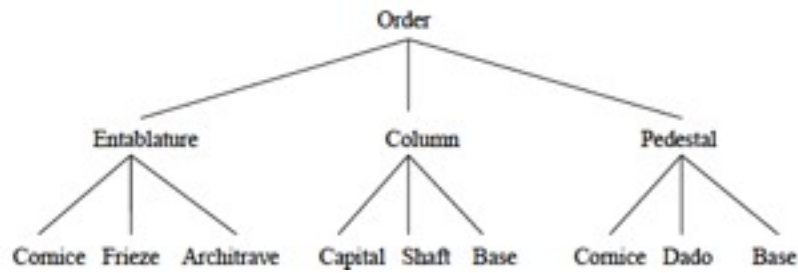


Figure 1.6 The Doric Order after James Gibbs, *Rules for Drawing the Several Parts of Architecture* (London, 1732). Collected in Rattner D., *Parallel of the classical orders of architecture*, Acanthus, New York, 1998

(Rattner, 1998) developed a classification of the moldings showing a systematic way the role they play in the design of classical buildings. The author defines the moldings as the smallest physical units – the atoms – of the classical architecture and provides a way to understand the shape of the architectural elements according to their combination. This classification is based on a variety of 14 moldings (figure 1.7) and uses several criteria of regrouping variety of moldings according to their architectural function in the composition of the building (De Luca et al., 2007).

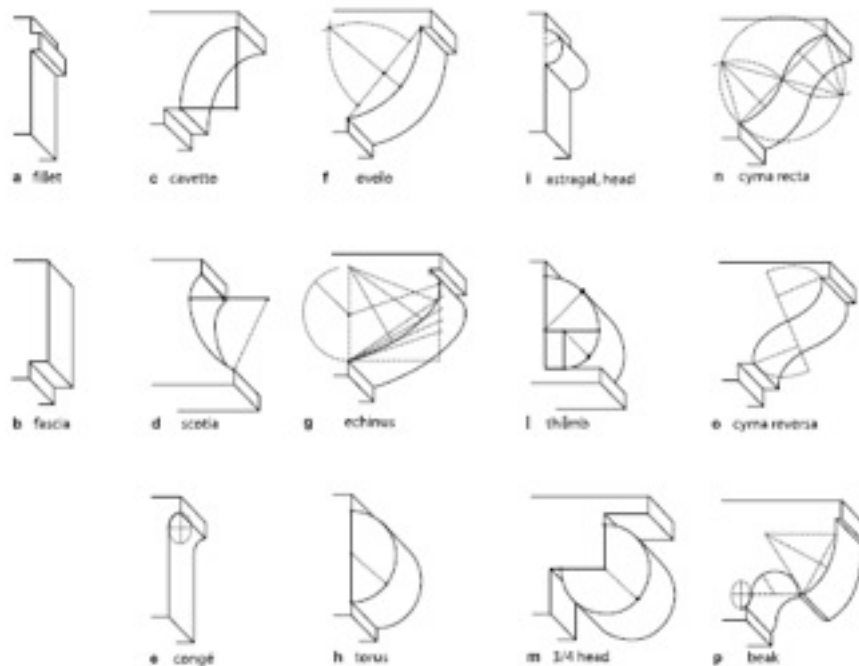


Figure 1.7 Classification of moldings of the classical language suggested by (Rattner, 1998)

We have already set the parametric ratios of the general shapes of the components. Semantic segmentation process will finally lead to the atom units. Recent study analyzed their shapes and parametric rules are applied to them. As a result, we are able to describe the 2D profile of a component from the very bottom of the design process with vast amounts possibilities in each nodes. Next part will focus on how to generate 3D objects from the 2D profile.

2.3 Sweep operation

We can regard a straight line as the path swept out by a translated point. In the same way, a rectangular surface can be swept out by a straight line, and a solid swept out by a surface. (Mitchell, 1995). Common transformation operations of geometry such as extrusion, revolve, blend and even loft can all be regarded as the variations of sweep with specific constraints or conditions. For example, a cylinder can be modeled either by revolving around its central axis or sweeping along the edge of its bottom or top surface. A cone can be constructed by sweeping a diminishing circle along straight line.

Most classical architecture elements can be modeled by sweep operation. It is not an coincidence, but closely analogous to the craftman's technique of using a template to make a profile on a block of stone and then cut off the extra parts (Mitchell, 1995). Components of classical architecture such as cornice, frieze, balusters are all applied to this rule. Even components of Gothic architecture and Chinese architecture have the similar features.

If we reverse the process, a classical architectural element can be decomposed into a set of 2D and 1D shapes (profile and path). In Palazzo Barbaran da Porto, the triangular cornice and curve cornice above the windows have the same profile but different path and the cornices on different levels have different profiles but the same path (figure 1.8).

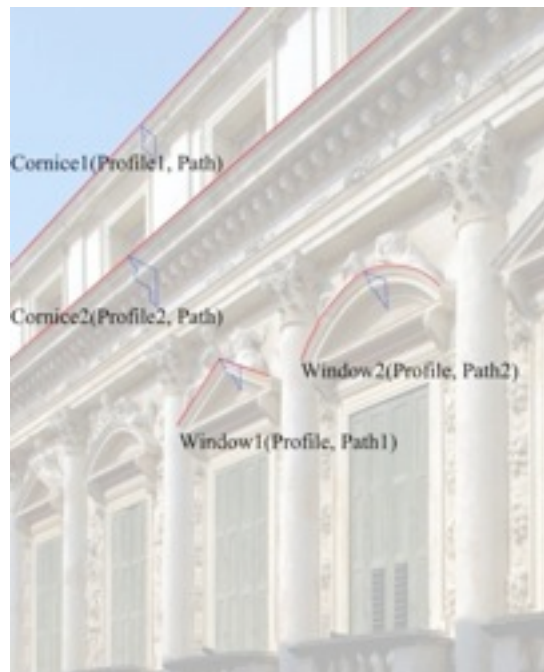
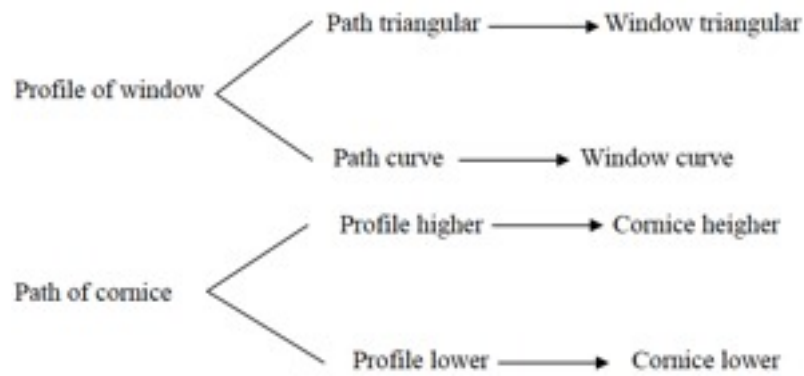


Figure 1.8 Sweep to generate the cornice and the window in Palazzo Barbaran da Porto, designed by Palladio

Sweep operation make transformation of 2D closed loops on construction planes. The profile composed by atoms units are enlarged with multiplied possibilities of geometric composition.

2.4 Bottom-Up process

In this part, we build our analysis much on the previous parts from 2.1 to 2.3. We have explained the geometry of classical architecture is actually a formal composition of Eculid's elements. We can depict their geometry with simple dimensions like points and lines. The dimensions of these primitive could be assigned with parameters. Then we show how a Doric order is decomposed into atom units and the fourteen basic classifications of the units. It is the regroup of these units that form the profiles of Order, Cornice and Entablature of classical architecture. Once we are able to depict the profile of geometry, it is easy to generate the 3D shape by sweeping along the correct paths. In this way, we are able to compose an element such as a Doric order from the very bottom of

the design process. If we zoom out from a single element to group of elements, the basic unit of repetition on façade of Palladio’s Palazzo for example, we may find the rules still work.

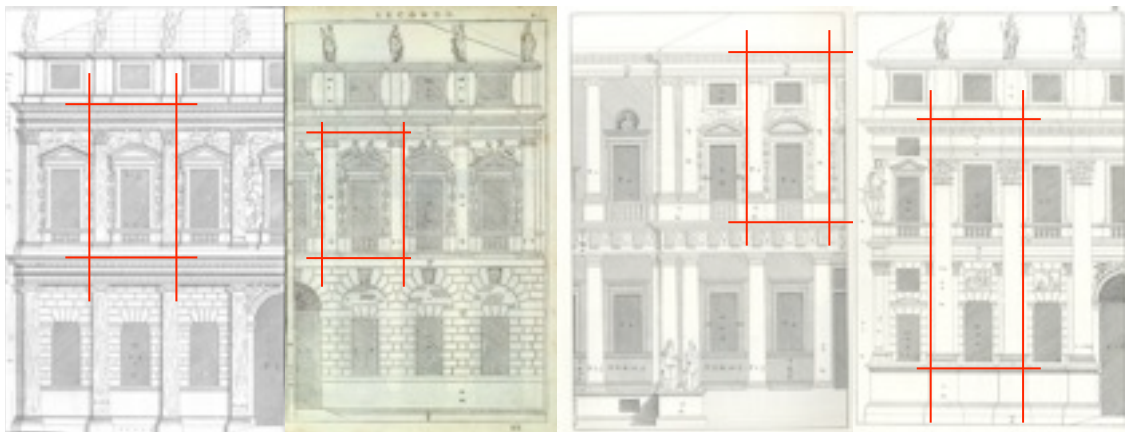


Figure 1.9. Several drawings of façade of Palazzo after Palladio’s *The Four Books on Architecture* (Venezia 1570, I, 18). Red lines show the basic unit of repetition on façade. From left to right: Palazzo Barbaran , Palazzo Porto, Palazzo Chiericati, Palazzo Valmarana

A basic unit that compose a façade on Palladio’s façade (figure 1.9) is usually articulated with four rules: 1) A pair of pilasters whose capital could be either Ionic or Corinthian, shaft cube or cylinder, height dominating one floor or two floors (chart 1.3); 2) A window with top molding’s cornice either triangular or curve (If the pilasters dominate two floors in rule one, then a rectangular window is added either above or below the window on the noble plan); 3) A balcony with molding and an array of balusters, and 4) Two segments of Entablatures on the bottom and top of the façade unit.

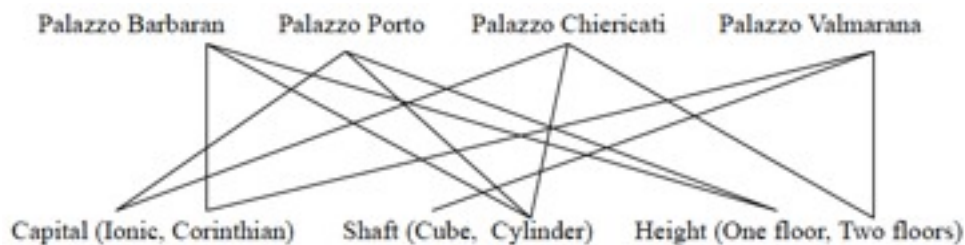


Chart 1.3 Options of rule 1 that applied to the four Palazzo

In this case, we set only variables in the first and the second rules, while keep the third rule and fourth rule constant. The second rule is articulated with a “If, then” sentence. The result of the first rule determines if the part after “then” will be executed. The application of the first rule and the second rule lead to the generation of façade’s unit as shown in figure 1.9.

If we keep in mind that capital of Ionic order and Corinthian order are composed by members as well as sub-members, we realize the Bottom-Up process from atom units to the façade. We can apply the process of order to the other members and their sub-members until the atom units. In this way, we create all the other possibilities from the same compositional rule. In most cases, classical architecture dose not employ new geometries, but different combinations and regroupes of elements based on certain rules. It is in analogy with language that new words are rarely invented but new sentences yielding the grammar are always regrouped.

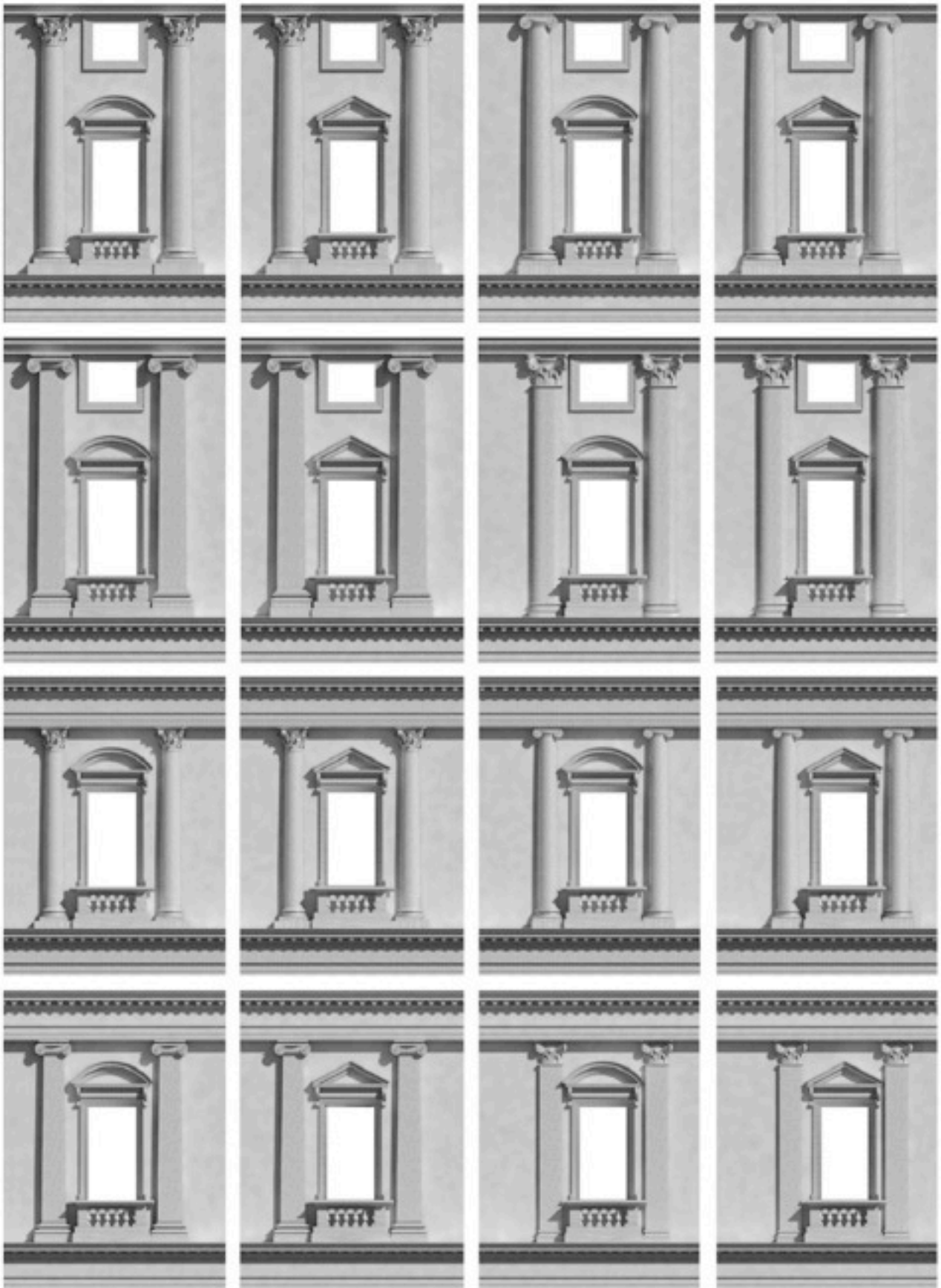


Figure 1.9 Re-combination of façade elements based on four variables including two options.: Capital (Corinthian, Ionic), Shaft (Cube, Cylinder), Height (One floor, Two floors), Window (Triangular, Circle)

Chapter 3 BIM-based modeling and management

BIM (Building Information Modeling) is not only a new modeling tool that produces associative engineering drawings, but also a dataset where multiple information is collected, stored and shared. It has the potential to act as a knowledge system as the rule architectural treatise played in pre-modern society.

The transition from CAD to BIM in AEC (Architecture, Engineering and Construction) industry is already underway (Eastman, 2007), while the documentation and management of architectural heritage has barely benefited from this technological renovation. The application of BIM in new projects reminds us that BIM is not only a more powerful modeling tool, but also provides inherent semantic organization of architectural components with parametric relations. In addition, BIM can serve as a centralized dataset supporting comprehensive data input for life-cycle management and various analytical purposes.

In recent years, as the field of architectural heritage has taken great advantages from the improvement of data capturing technologies (Remondino, 2011), the need to establish centralized repositories creating further simulations is on the schedule (Koller, 2009). The application of BIM is opening considerable possibilities, while a number of issues such as the techniques, platforms and standards are to be solved.

BIM-based modeling and documentation of existing buildings is still in its infancy. Researchers take advantage of different features of BIM for various purposes. Modeling parametric components (Murphy, 2011) with BIM tools is an effective approach which eases the time consuming modeling process in traditional CAD. Based on architectural pattern books, parametric objects are scripted using Geometric Descriptive Language (GDL) in ArchiCAD, and then mapped onto points cloud to automatically produce engineering drawings. Autodesk Research explored the potential of BIM in life-cycle management data (Attar, 2010; Fai, 2011). BIM model can act as a repository of data that not only describes both geometrical and semantic aspects of buildings, but also provides access to various analytical purposes, such as energy saving and fire evacuation. Besides explicit semantic description, BIM also facilitates the study of theoretical and historical knowledge (Pauwels, 2008). Pauwels proposed a schematic spatial model allowing abstract information, such as reference to historical plans or other documentation, to be added during different stages of research (Eastman, 2007). Among the current researches, several challenges are to be addressed.

The first challenge concerns the development of a series tools and techniques of modeling classical architecture. In this paper, we explored a grammar-based modeling approach through building semantic and parametric connections between Palladio's rules and his work. Also, a typology of geometrical shapes of classical architecture (De Luca, 2007) and points cloud mapping (Murphy, 2011) are combined to adjust to particular situations.

The second one concerns the platforms of BIM. Among the dominating BIM platforms, ArchiCAD and Revit have been employed in the researches of architectural heritage currently, and there will be much work to explore the suitable platform and related techniques.

The last one relates to the standard of data exchange among different platforms of BIM, and BIM platforms with other simulation platforms. Existing since early CAD, interoperability became more complicated with the advent of BIM, because not only geometrical shape, but also parametric

relations are supposed to be shared. Information Foundation Class (IFC) is expected to be a promising standard, such as Extensible Markup Language (XML), the language used to send information over the Web (Eastman, 2007). Some of the new exchange formats are XML-based, such as City Geography Markup Language (CityGML) and Green Building XML (gbXML), which provide data exchange to GIS and energy analysis respectively.

3.1 Objects-based parametric modeling

The current generation of BIM architectural design tools, including Autodesk Revit, Bentley Architecture all grew out of the object-based parametric modeling capabilities (Eastman, 2007). BIM tools define, store and exchange data in object level instead of model level. It also creates the possibility to define the inner ratio of the architectural elements by parameters and formulas.

In traditional CAD platform, designers use given tools to generate geometry, and then assign semantic meaning to them. A column and a rail of stair may have the same geometry, but are just defined as a column and a rail. In BIM platforms, however, designers first make choice from the given template which has inherent semantic meanings. A template of column and that of rail are different in terms of structure, material and attachment attributes such as the relation with ceiling, beam and stair. The templates determine the potential behaviors of the components imitating the real construction. A column drawn in traditional CAD may lose its identity when shared in another platform due to the set of layers, but a column modeled in Revit remains its attributes of material and structure when imported to Archicad via IFC.

With regard to classical architecture's coherent feature of semantic organization and ratio, the object-based parametric modeling may greatly ease the manual modeling process in traditional CAD platforms. It is easy to transfer a Doric order from the ratio of Vignola to that of Palladio and make variations. In addition, a type could have several instances. Instances share some parameters in common which are named "type parameter", and have their own parameters which are named "instance parameter." Imagine two Doric orders have the same semantic composition and inner ratio (type parameters), but are placed at different levels and are attached to different objects (instance parameters). The relation of type parameters and instance parameters are closely analogous to the ideal state and actual state of an object. For example, the erosion degree of a set of capitals that locate in the same portico is instance parameters (figure 1.10). When point cloud are captured, more precise mesh model could be used to calibrate the model based on treatises.



Figure 1.10 A pair of Ionic Order with different as-built state

3.2 Workflow of parametric object modeling

BIM enables object-based parametric modeling, and inherently incorporate semantic data pertaining to structural, material and operational information (Eastman, 2007). We can model the parametric architectural objects according to the treatises in the BIM platforms, and assign additional information to these models. Instead of drawing a concrete shape, we can replace dimensions by variables, and set formulas among them. In this way, we create Vignola's Doric order or Palladio's Doric order. In each case, the dimension adjustments are automatic. One modification will reflect to the others in real time. The work is actually a process of information translation between two systems of knowledge.

Palladio described the semantic structure of orders and their ratios by texts and illustrations (figure 1.11). In the case of Doric orders, Palladio wrote:“. . . The capital ought to be in height half the diameter of the column, and is to be divided into three parts. The upper part is given to the *abaco* and *cimacio*. The *cimacio* is two of the five parts thereof, which must be divided into three parts; with the one the *listello* is made, and with the other two the goal. The second principal part is divided into three equal parts, one to be given to the *anelli* or *annulets*, or *gradetti*, which three are equal. The other two remain for the *ovolo*, which projects two thirds of its height. The third part is for the Collarino . . . ” (Palladio, 1965)

In order to create a Doric column according to Palladio's treatise, instead of designing an instance, we first define its family or class which contains a set of predefined behavior. Various families have different predefined behavior with reference to real construction. Column family can be attached to Entablature, ceiling and other structural frame, while wall family's behavior involves the relation with the opening of door and window. In Revit Architecture, a template file is supposed to be selected from the lists of default templates.

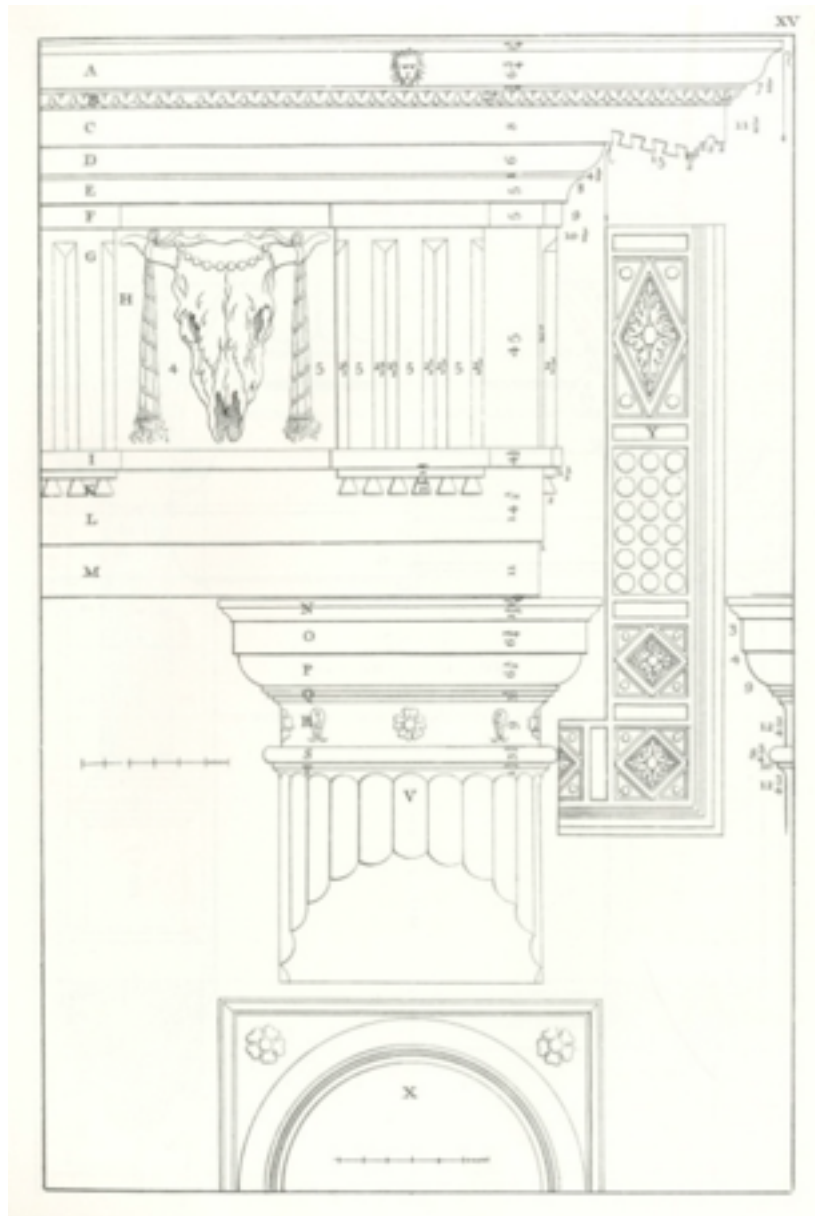
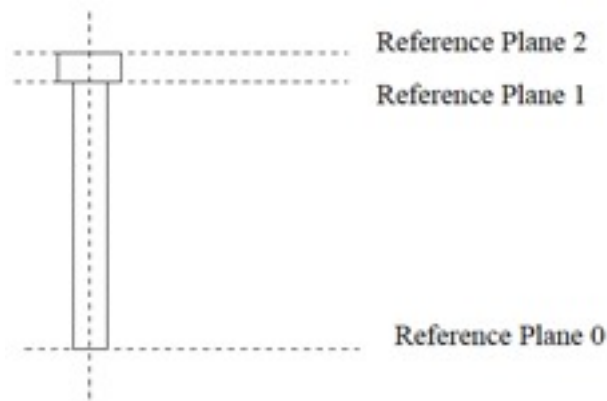
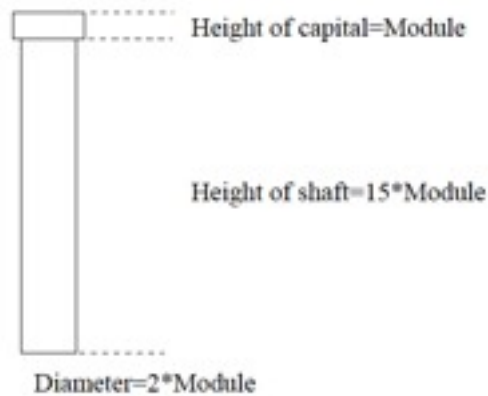


Figure 1.11. Description of Doric' capital with measurement in *The Four Books on Architecture* (Venezia 1570, I, 18)

We should set the framework for preceding operation taking into account of the Doric column's composition. As we already know in the previous parts, a Doric column consists of capital and shaft. The base is absent after Palladio's Doric column. Then we create a set of reference planes to which the edges of two volumes (capital and shaft) are attached. The distance is based on the general shapes of the column.



Palladio set the relations among all the members as well as their sub-members by unit of module. For Doric order, the height of the capital equals a module which is half the diameter of the lower part of the shaft, and the height of the shaft is fifteen modules. We assign parameters to these dimensions by giving them labels, and use formulas to make potential associative connections among them. If a value is assigned to “Height of capital”, for example, “Height of shaft” and “Diameter” will adjust to the change automatically without changing their ratios.



We can refine the shape of the shaft, especially the order by decomposing them into the atoms. Parameters and formulas are set to them in the same way as the previous procedures. When the 2D profile is drawn, we have two options to generate the 3D shape. The first one is revolve and extrusion. We first generate the 3D shape of every single atoms and then combine them together. It is more time-consuming and tedious, but the potential segmentation process, if there is, of the order would be much easy. The second option is sweep which needs a profile and a path. Since we already have the profile, we just have to select or draw the path. The strength and limitation are converse to those of revolve and extrusion.

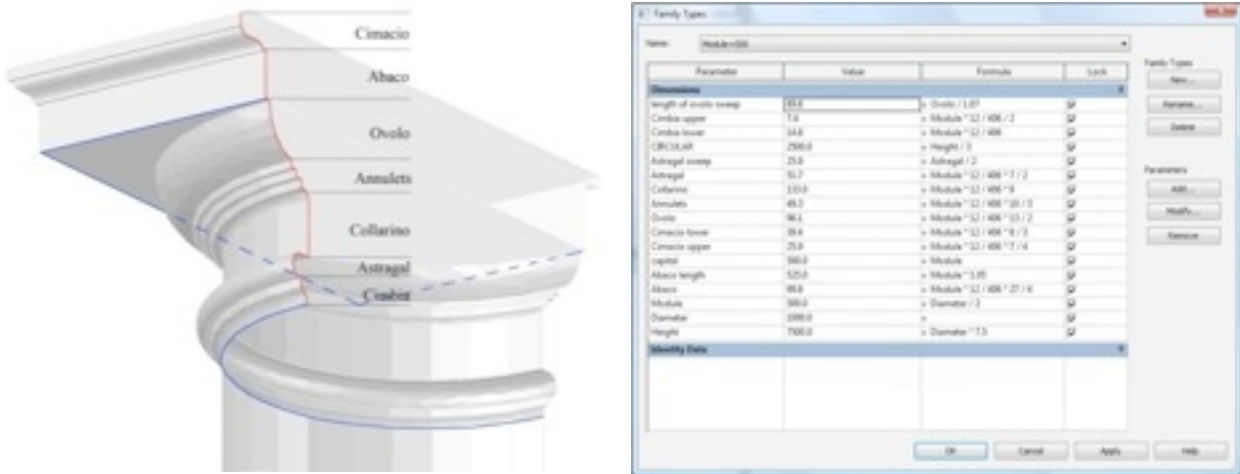


Figure 1.12 Modular relations between different parts of Doric' capital in Revit according to Palladio's rules

Once the modeling process is completed, we can test the behavior of family by modifying parameters. We get a set of Doric columns (family) with module's value vary (types) (figure 1.12). The workflow of modeling a Doric column in Revit Architecture also applies to the most elements of classical architecture, such as doors, windows and balusters (chart 1.4). In some cases, a Boolean operation is necessary to cut the extra part with void when the operations of solids are invalid to model some shapes. In the triangular cornice, for example, the path of sweep is not vertical to the surface. Sweep operation either generates extra or void parts. An extrusion of void would easily solve the problem.

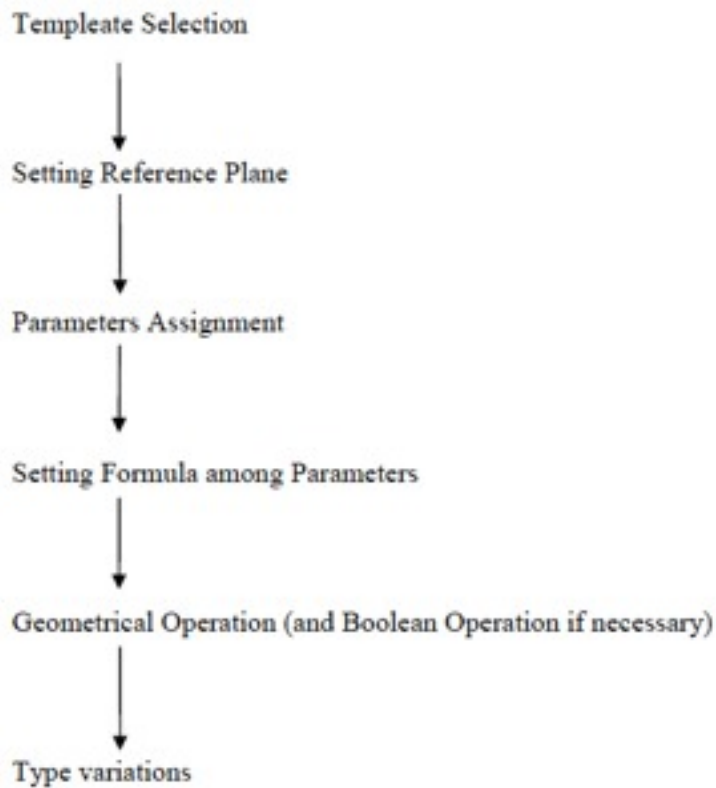
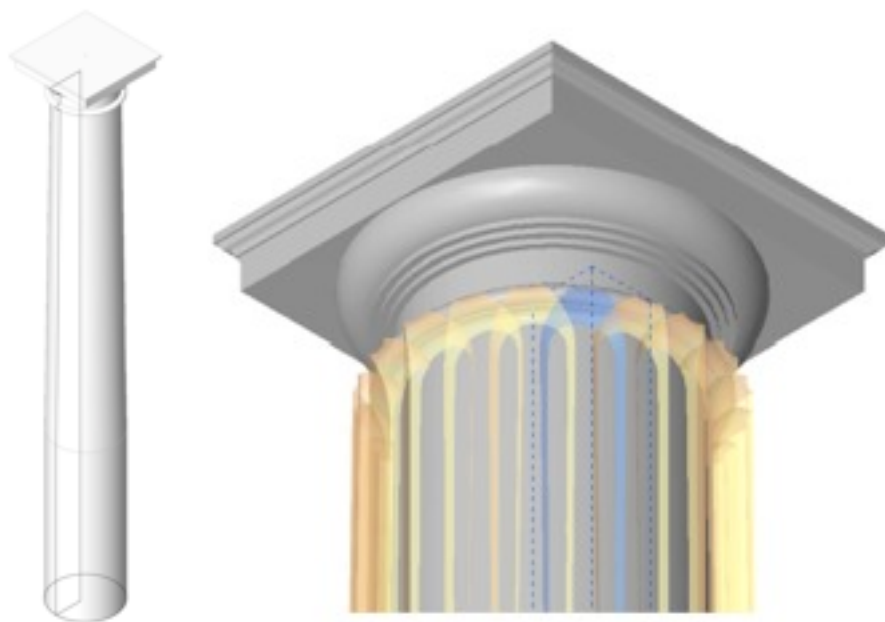


Chart 1.4 Workflow of parametric modeling of components in Revit Architecture

Extracting the semantic organization and modular dimensions from treatises of Palladio, Scamozzi and Vignola, we modeled the parametric Doric orders in Revit. The workflow consists of 1) creating a set of reference planes among whose distances are labeled in the name of semantic nodes; 2) setting up the constraints of parameters by formulas; 3) creating geometry of the semantic nodes and aligning them to the corresponding reference planes; 4) loading the profile to the column family and sweeping it along the bottom edge of the order (figure 1.13). 5) creating another profile and loading it into the column family to generate the fluted order via Boolean operation (figure 1.14).



Left: figure 1.13 Loading profile into sweep operation to generate Doric order
 Right: figure 1.14 Editing profile of flute and Boolean operation

When the modeling process finishes and is proved to be well-performed, we are able to generate a set of types in a wide range of dimensions by assigning new values to the “module” parameter. The ideal models strictly based on treatises enable automatic generation of knowledge-based components and serves as central platforms to which data captured from surveys could be linked. The parallel of Doric orders in BIM platform provides a set of views (figure 1.15, figure 1.16) similar to the parallel of orders in the treatises, and also creates schedules of the data in a parallel way.

The advantage of using profiles in modeling is that it is a reusable family file which could be edited and loaded into other families. The Doric order modeled according to Palladio’s treatise could become that of Vignola or Scamozzi in real time by profile substitution in the option list. In addition, a 3D form is translated into 2D shape via the sweep operation. Common transformation operations of geometry such as extrusion, revolve, blend and even loft can all be regarded as variations of sweep with specific constraints or conditions. For example, a cylinder can be modeled either by revolving around its central axis or sweeping along the edge of its bottom or top surface. A cone can be constructed by sweeping a diminishing circle along a straight line. Most classical

architecture elements can be modeled by sweep operation. It is not a coincidence but closely analogous to the craftsman's technique of using a template to make a profile on a block of stone and then cut off the extra parts (Mitchell W.J., 1994).

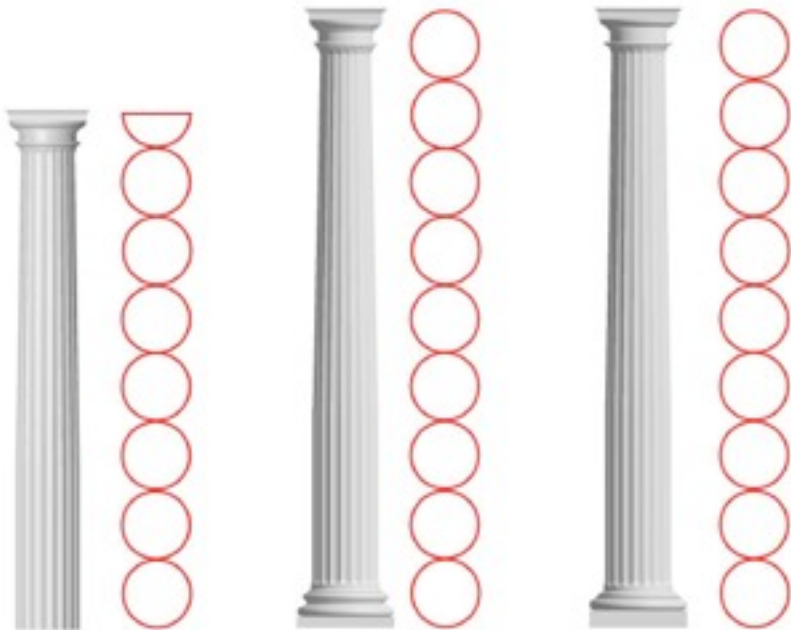


Fig. 1.15: Parallels of Doric orders modeled in Revit Architecture after Palladio, Scamozzi and Vignola (from left to right)

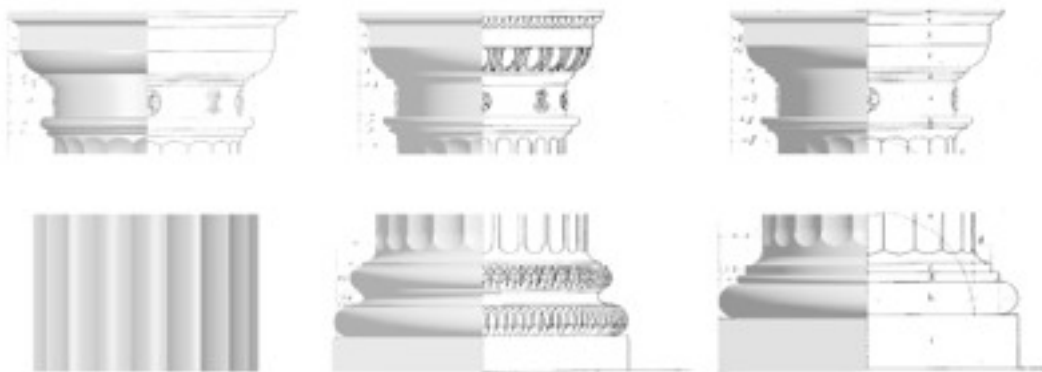


Fig. 1.16: Parallel of Doric orders' detail modeled in Revit and from treatises of (from left to right) Palladio, Scamozzi and Vignola (Rattner, 1998. Parallel of the classical orders of architecture, Acanthus, New York)

Chapter 4 Case study: Palazzo Barbaran da Porto

Palazzo Barbaran da Porto, designed by Andrea Palladio, is one of the most important architecture of the Renaissance in the Western Culture and built in the inner town of Vicenza, Italy. Palazzo Barbaran da Porto was designed in 1569 and built between 1570 and 1575 by Andrea Palladio. It is the only great city palace that Palladio succeeded in fully executing during his life. Commissioned by Montano Barbarano, a Vicentine noble man, Palladio was asked to coordinate the new design and the existing houses belonging to the family of the client. Palladio contributed a set of masterful proposals with a central atrium and symmetric façade with 7 bays. Following the completion of the drawing, Barbarano acquired an adjacent house in the west, which eventually resulted in a façade with 9 bays.

Palazzo Barbaran da Porto has been chosen as a case study due to its peculiarity, related to its design process and historic development (during Palladio's life time) and to its architectural composition, where we can find out Palladio's buildings theoretical principles, methodological process and design rules exemplarily apply.

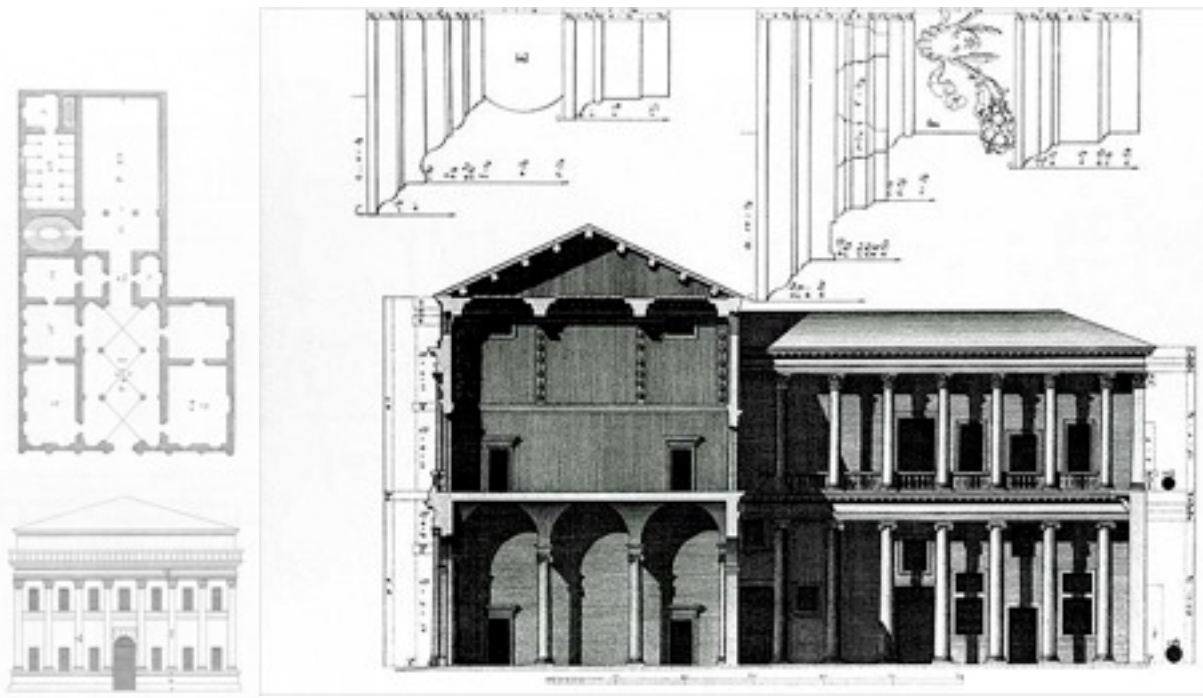
In the end, a set of engineering drawings including plans, elevations, sections and renderings are produced (figure 1.17).



Figure 1.17 Rendering of Palazzo Barbaran da Porto

4.1 Semantic organization

The construction of Palazzo Barbaran da Porto was completed after the death of Palladio. For this reason, the drawing (figure 1.18; figure 1.19) collected in *The Four Books on Architecture* is different from the actual state of the building. Thanks to Ottavio Bertotti Scamozz, Palladio's interpreter in 17th century, we are able to acquire the essential data from his survey. Besides plans, elevations and sections, Scamozz drew much attention on the details of Entablature. He supplied the profiles of Palladio's buildings with dimensions in his survey.



Left: Figure 1.18. Palladio's design for Palazzo Barbaran da Porto collected in the *The Four Books on Architecture* (Venezia 1570, II, 22)

Right: Figure 1.19. Ottavio Bertotti Scamozzi's survey for Palazzo Barbaran da Porto collected in *Le fabbriche e i disegni di Andrea Palladio raccolti ed illustrati*

As the modeling process illustrated in part 2.3.2, we created a set of basic architectural components strictly based on Palladian grammar. The semantic organization starts from choosing the predefined families from the library. Variations and adjustments are made according to various practical situations. Even though only Ionic columns and Corinthian columns are employed by Palladio in Palazzo Barbaran da Porto, 10 variations exist (chart 1.5). Difference of size can be automatically adjusted by modifying parameters, while new geometry of capitals has to be manually edited. It also results a new family of column which enriches the family library. Then, instances of family are mapped onto a points cloud or image based survey (Murphy, 2011). In the same way, windows, doors, railings and other categories of components of the building are adjusted, created and distributed. In addition, Revit Architecture enables users to create subcategories with unique parameters when it is essential to distinguish different parts in a family. A visual connection between the semantic organization and the database of the project can be founded via Revit DB Link. The database (Microsoft Access, Microsoft Excel, or ODBC database) displays the metadata of the families, instances and other information of the project.

Category	Family	Type	Number of instance	Location
Column	Corinthian	Diameter 620mm	13	1st Floor_facade
		Diameter 648mm	9	1st Floor_court
		Diameter 648mm_semi	1	1st Floor_court
	Ionic_1	Diameter 728mm	13	Ground Floor_facade
		Diameter 758mm	8	Ground Floor_court
		Diameter 758mm_semi	1	Ground Floor_court
	Ionic_2	Diameter 758mm	1	Ground Floor_court
	Ionic_3	Diameter 640mm	4	Ground Floor_atrium
		Diameter 640mm_semi	8	Ground Floor_atrium
Diameter 640mm_quart		4	Ground Floor_atrium	

Chart 1.5 BIM metadata of the columns of Palazzo Barbaran da Porto

4. 2 Data enrichment and filter

Documentations of architectural heritage are described in different levels according to analysis need (De Luca, 2011). As a centralized repository of data, BIM enables comprehensive data input besides semantic and geometrical information, and displays the data needed by users via predefined filters. In this project, we add structural, material and other optional data based to the semantic organized components, and set relevant filters for data management (chart 1.6).

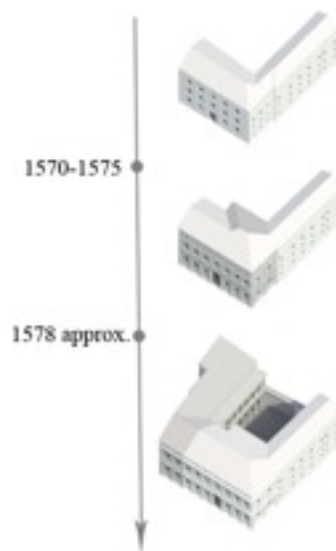


Figure 1.20. The evolution of morphology of Palazzo Barbaran da Porto in 3 construction phases

One of the most notable features is the time line which is unique and important for the documentation and management of architectural heritage. Revit Architecture provides every family with temporal attributes by setting the parameters of construction phase and demolish phase. Palazzo Porto da Barbaran underwent three construction phases generally (Figure 1.20) between some years before 1570 and up to after 1575. The time line feature presents a morphology evolution

of architectural heritage to the public, and fosters a sustainable frame to which properties with temporal dimensions can be attached.

Model components could be chosen to display or hide in the light of instances (ID), category and parameter. Besides default filters, Revit API (Application programming interface) enables users to customize the filters by programming language such as VB.net or C#. Rooms with certain enclosure conditions and reaching specific dimensions can be filtered for analytical need. The approach is efficient in large architectural models and finding similarities between two models.

Format	Description
.RFA	Families of the project
.RFT	Central model containing all BIM data of the project
.PDF	Containing engineering drawings
.IFC	Industry Foundation Class export of Revit file
.gbXML	File for energy analysis
.MBD	Database in Microsoft Access
.KML	3D Model containing semantic structure interactive with Google Earth

Chart 1.6. File structure of Revit model of Palazzo Barbarand da Porto

4.3 Interoperability and related data-base

Interoperability is the ability to exchange data between applications (Eastman, 2007). A considerable functional extension of BIM performs not in BIM models, but via interoperability with other applications and platforms. The simulation tools are the most common application of BIM extension in new projects and various workflows exist, such as energy performance, crowd simulation, environment analysis and fire evacuation, while the application of BIM extension for architectural heritage is a new area of research. Interoperability has traditionally relied on file-based exchange formats limited to geometry, while BIM repositories are distinguished by providing object-based management capabilities, representing attributes, properties and behaviors of geometry (Eastman, 2007).

Palladio 3D geodatabase is a digital 3D GIS web-based system founded on Google Earth, built to offer a deep insight into Palladio's opera for restitution of its architectural surveys, reconstruction for historical reasons, analysis for structural and other options, and assembly and presentation of all certified Palladian documentation (Apollonio, 2010). The 3D modeling workflow is divided into 3 mutually connected phases (modeling, segmentation and visualization). The resulting file is in KMZ (zipped KML) format which includes KML (Keyhole Markup Language) and a series of referenced file (DAE, TIF or PNG). In the workflow, Revit Architecture is used in the modeling phase, while segmentation and visualization are carried out in Maya, which is more powerful in images rendering, but cannot read BIM's inherent semantic structure and additional data. In this way, the interoperability between the 2 dataset repositories, BIM and web-based GIS, is isolated.

A BIM model containing information of different levels, and execute in the full cycle management is expected to smooth the current workflow and facilitates its automation. Revit Architecture

enables BIM model to interoperate with Google Earth via Revit Globe Link, a plug-in of Revit Architecture. Through importing and exporting KML and KMZ files, Revit Globe Link acquires the site information needed in Revit modeling from Google Earth, and publishes Revit building information model into Google Earth. Both semantic and visual attributes remain in Google Earth as they are set in Revit Architecture. Though the advantages, problems and further works exist in 3 aspects: (1) Revit Architecture is not capable of producing high quality rendering; (2) the 3D model is too large to be easily handled in Google Earth; (3) how to transform the metadata of the architectural objects within BIM platform into the description in interface of Palladio 3D geodatabase coded by HTML.

Chapter 5 Summary

In the chapter, we presented a knowledge-based approach for the documentation and management of architectural heritage in BIM platform. The treatises are supposed to be a theoretical frame east the process of segmentation and semantic organization. The enormous historical data that stored in the treatises could be transferred into BIM dataset through parametric modeling and further data enrichment. Architectural components are first modeled according to corresponding architectural grammar and stored as families in the library. Automatic modification and manual variations are made to adjust particular situations with the reference of laser scanner, image-based survey and archives of drawings. BIM's semantic structure enables multiple data enrichment and filters of various analysis need. Finally, simulations and inter-operation with other analytical platforms are executed. In this way, the centralized BIM model serves as a sustainable dataset comprising comprehensive information, and provides access to life-cycle management and potential simulations of various disciplines. The future work in the domain BIM platforms will address 1)the exploration of automatic generation of complicated architectural shapes in BIM. 2) Customize the application of BIM in heritage management via Revit API. Revit API enables users to add external command and external application by scripting in Microsoft Visual Studio, and 3) Establishment of web-based data repository interactive with GIS. In the whole survey process, image-based survey and range-based survey will be employed to collect real-life data of the buildings. The method of integrated techniques will be addressed in the next chapters.

Reference:

1. Ackerman J.S., 1966. Palladio. Penguin.
2. Alberti L.B., 1550. L'architettura, Firenze
3. Payne A.A., 1999. The Architectural Treatise in the Italian Renaissance: Architectural Invention, Ornament, and Literary Culture, Cambridge University Press
4. Apollonio, F. I., Corsi, C., Gaiani, M., & Baldissini, S. (2010). An integrated 3D geodatabase for Palladio's work. *International Journal of Architectural Computing*, 8(2), 111-133.
5. Attar, R., Prabhu, V., Glueck, M., & Khan, A. (2010, April). 210 King Street: a dataset for integrated performance assessment. In *Proceedings of the 2010 Spring Simulation Multiconference* (p. 177). Society for Computer Simulation International.
6. Baldissini S., Beltramini G., Gaiani M., 2008. Rich internet application as paradigm to manage and visualize cultural heritage references and contents: the case study of the first electronic edition of Andrea Palladio's The Four Books on Architecture. DMACH conference proceedings, CSAAR Press, Amman, Jordan, pp. 255-270.
7. Biermann V., Thoenes C., etc., 2003. Architectural Theory, From the Renaissance to the Present, Taschen.
8. Blondel J.F., 1771-1777. Cours d'architecture ou Traite de la Decoration, Distribution & Construction des Batiments, Paris ,
9. Chambray R.F., 1650. Parallele de l'architecture antique et de la modern: avec un recueil des dix principaux auteurs qui ont ecrit de cinq orders, Paris
10. Chevrier, C., Perrin, J.P., 2009. Generation of architectural parametric components. CAAD Future conference, June 17-19, Montreal, Canada, Presses de l'Université de Montréal, pp. 105-118.
11. De Luca L., Bussayarat C., Stefani C., Véron P., Florenzano M., 2011. A semantic-based platform for the digital analysis of architectural heritage. *Computers & Graphics*, vol. 35, n. 2, pp. 227-241.
12. De Luca L., Véron P., Florenzano M., 2005. Semantic-based modelling and representation of patrimony buildings. Semantic-based modelling and representation of patrimony buildings, SVE Workshop towards Semantic Virtual Environments
13. De Luca L., Véron P., Florenzano M., 1802-1805. A Generic Formalism for the Semantic Modeling and Representation of Architectural Elements. *The Visual Computer*, vol. 23, n. 3, 2007, pp. 181-205.
14. Durand J.N.L., *Precis des lecons d'architecture*, Paris.
15. Durand J.N.L., 1800. *Recueil et Parallele des edifies de tout genre anciens et modernes*, Paris.
16. Eastman C., Teicholz P., Sacks R., Liston K., 2007. *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors.* John Wiley & Sons, New Jersey
17. El-Hakim, S. and Beraldin, J.A., 2002. Detailed 3D Reconstruction of Monuments Using Multiple Techniques, *Proceedings of the International Workshop for Cultural Heritage Recording*, Corfu, Greece, pp. 58-64.
18. Fai S., Graham K., Duckworth T., Wood N., Attar R., 2011. Building information modeling and heritage documentation. CIPA 2011 Conference Proceedings, XXIIIrd International CIPA Symposium, Prague.
19. Hart V., Hicks P., 1998. *Paper Palaces: The Rise of The Renaissance Architectural Treatises*, Yale University Press
20. Gaiani M., 1999. Translating the architecture of the world into virtual reality and vice-versa: 7 years of experimentation with "conservation and representation" at OFF. In: *Proceedings of*

Heritage Applications of 3D Digital Imaging. NRC Ottawa

21. Gibbs J., 1728. *A Book of Architecture, Containing Designs of Buildings and Ornament*, London.
22. Gibbs J., 1732. *Rules for Drawing the Several Parts of Architecture*, London
23. Koller D., Frischer B., Humphreys G., 2009. Research Challenges for Digital Archives of 3D Cultural Heritage Models. *Journal on computing and cultural heritage*, vol. 2, n. 3 pp. 1-17
24. Krufft H.W., 1994. *Architectural Theory: From Vitruvius to the Present*, Princeton Architectural Press
25. Manfredini, A.M., Remondino, F., Baldissini, S., Gaiani, M., Benedetti, B., 2008. 3D Modeling and Semantic Classification of Archaeological Finds for Management and Visualization in 3D Archaeological Databases, *Proc. 14th VSMM*, pp. 221–228.
26. Mitchell W.J., 1994. *The Logic of Architecture. Design, Computation, and Cognition*. MIT Press, Cambridge, Massachusetts
27. Mitchell W. J., McCullough M., 1995. *Digital Design Media*, John Wiley & Sons, Inc..
28. Murphy M., McGovern E., Pavia S., 2011. Historical Building Information Modelling-Adding Intelligence to Laser and Image based surveys, 4th International Workshop 3D-ARCH 2011, pp. 7.
29. Murphy M., McGovern E., Pavia S., 2007. Parametric Vector Modelling of Laser and Image Surveys of 17th Century Classical Architecture in Dublin, The 8th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST.
30. Murphy M., McGovern E., Pavia S., 2009. Historic building information modelling (HBIM), *Structural Survey*, Vol. 27 Iss: 4 pp. 311 - 327
31. Palladio A., 1965. *The Four Books on Architecture* translation by Isaac Ware 1738, Dover Publications, New York.
32. Pauwels P., Verstaeten R., De Meyer R., Van Campenhout J., 2008. Architectural information modelling for virtual heritage application, digital heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia, pp. 18-23.
33. Rattner D., 1998. *Parallel of the classical orders of architecture*, Acanthus, New York
34. Revit Architecture 2013 Users' Guide, 2013, Autodesk.
35. Remondino, F., El-Hakim, S., 2006. Image-based 3D modelling: a review. *Photogrammetric Record*, 21(115), pp.269-291.
36. Remondino F., 2011. Heritage Recording and 3D modelling with photogrammetry and 3D scanning. *Remote Sensing*, vol. 3, n. 6, pp. 1104-1138.
37. Park H.J., 2007. Parametric variations of Palladio's Villa Rotonda. *International Journal of Architectural Computing*, vol. 06, n. 02, pp. 145-169.
38. Scamozzi O.B., 1796. *Le fabbriche e i disegni di Andrea Palladio raccolti ed illustrati*, Venice.
39. Sass L., 2007. A Palladian construction grammar—design reasoning with shape grammars and rapid prototyping. *Environment and Planning B: Planning and Design*, vol. 34, n. 1, pp. 87-106.
40. Scamozzi V., 1615, *Idea dell'architettura universal*, Venice.
41. Serlio S., 1619. *Tutte l'opere d'architettura et prospettiva di Sebastiano Serlio Bolognese*, Venice, First published in 1537(book 1), 1540(book 2), 1545(book 3), 1547(book 4), and 1575 (book 5)
42. Stiny, G., Mitchell, W. J., 1978. The palladian grammar, *Environment and Planning B* 5, 5–18.
43. Summerson J., 1977. *The classical language of architecture*, The MIT Press.
44. Tzonis A., Lefavre L., 1987. *Classical Architecture: The Poetics of Orders*, The MIT Press.
45. Vignola G.B., 1607. *Regola delli cinque Ordini d'Architettura*, Rome.
46. Wittkover R., *Architectural Principles in the Age of Humanism*, London, 1949.

Part 2 Image-based modeling

Chapter 1

1.1 Range-based modeling and image-based modeling

Range-based modeling (Bernardini et al., 2002) has been commonly used in architectural heritage. It captures 3D geometric information of objects based on active sensor and generate accurate results. The workflow has been quite straight forward since its wide application in cultural heritage for 20 years: scan the objects from multiple stations, register the points cloud from different stations and merge them in one model. In spite of the high accuracy level, range-based modeling is limited in flexibility, portability and high cost. Image-based modeling (Remondino et al., 2006) uses 2D images to recover 3D information through mathematical models. In contrast to range-based modeling, it is more economical, portable, flexible and gives rise to photorealistic models. Range-based modeling is an ideal approach in ruins of architectural heritage such as Pompei, where negative effects of traffic could be minimized. In large-scale civil buildings such as Bologna's 42km portico, however, the traffic and pedestrians in city center are difficult to avoid given each scan lasts more than 20 minutes. In addition, image-based modeling could reconstruct non-existing objects from the imagery archives. The two approaches are usually integrated in survey of large complex architectural heritages (El-Hakim et al., 2004; Remondino 2011). Comparisons of range-based modeling and image-based modeling are presented in (Boehler et al., 2004; Baltsavias 1999). Image-based modeling is gaining attentions in recent years thanks to the robust automatic image feature correspondence techniques developed in computer vision fields. The algorithms ease the time-consuming manual operation and yield model with high completeness, but limited metric accuracy. How to utilize this technique to achieve high accuracy for metric survey is the focus of this chapter.

1.2 The division of photogrammetry and computer vision

For architects who utilize the fruits of image-based modeling, the division of photogrammetry and computer vision could barely arises any particular interests. It is not unwise to bear in mind, however, that the two communities develop the shared techniques out of different aims, and hence lead to distinctions in terms of accuracy, completeness and efficiency (Förstner 2002).

Communities of photogrammetry community and communities of computer vision contributed to image-based modeling out of different aims: the central theme of photogrammetry is accuracy, while computer vision emphasized more on automation. Photogrammetry started in the middle of the 19th century when the first focal-plane cameras appeared (Havlena, 2012), concentrates on the accuracy of image orientation and camera calibration out of the purpose of measure and mapping (Brenner et al., 1998; Haala et al., 1997; Vosselman et al., 2001), and therefore requires rigid image acquisition. Computer vision community, however, focuses more on the automation aspect of the process for better and easier visualization (Remondino et al., 2006). A well-organised image capture is necessary in traditional photogrammetry. The viewpoints are supposed to be highly overlapped in adjacent pair and each one has similar depth of view. It is not advisable to take one photo with the whole target in the frame, but the other focusing one some detail. Meanwhile, the camera parameters should be fixed during the capture process to avoid the chaos of camera calibration. Techniques coming from computer vision community such as SfM (Structure from Motion), due to its SIFT algorithm (Lowe, 2004), is able to reconstruct 3D scenes from images that vary in camera, illumination and season. Furthermore, SfM (Structure from Motion) have been applied to large scale reconstruction such as Piazza San Marco and Rome from unorganized photo collections (Hartley et al., 2000; Snavely et al., 2006; 2008; 2010; Agarwal et al., 2009; Wu, 2011). From an architectural point of view the progress in computer vision is promising, since it eased the

cumbersome process and provide architects with results. Thanks to this technique, people from architectural community (design, restoration, history) are able to reconstruct 3D scenes from photo collections with a basic skill of computer and a decent device. In spite of the great potential, SfM still has the following drawbacks: Level of accuracy and 'black box'. SfM employs self-calibration to extract camera parameters. The accuracy level is not as high as that of laser scanner, and different solutions may vary given the same image input (Brutto, 2012). Some web-based SfM packages (Arc3D, 123D Catch, Recap) provide users with few parameters to handle which lead to the unpredictable result. Consequently, an integrated approach that combines accuracy with automation and complex architectural scene is within the scope of the paper.

Chapter 2. Evaluation of SfM packages

2.1 State of art

The field of computer vision witnessed the boom of SfM (Structure from Motion) in the last 20 years (Fitzgibbon et al., 1998; Nister et al., 2004; Vergauwen et al., 2006; Snavely et al., 2008; Agarwal et al., 2009; Wu, 2011). The evaluation of SfM has become an issue that arose field-across interests. Parallel comparison of different SfM packages have been reported in terms of visual completeness and metric accuracy. Computer vision community, the field where SfM grew from, employed benchmark datasets ranging from small hand crafts to architectural reconstruction. In the latter cases, the reported geometric accuracy is at meter-level (Crandall et al., 2011; Wu, 2013). The metric accuracy of SfM, however, is the main concern of the architectural heritage and archaeology. The prevailing method compares the geometric deviation between 3D model obtained by laser scanner and that from SfM in terms of deviation of 3D shape or 2D cross section (Brutto et al., 2012; Manferdini et al., 2013; Santagati et al., 2013; Donets et al., 2011; Dante, ???). Besides, (Dellepiane et al., 2011) evaluated the repeatability of SfM tools from different photo sets in the same object. (Remondino et al., 2012) presented a comprehensive review in by employing ground truth of GCP, objects with known geometric shape, and internal parameter of calibrated camera. Besides parallel comparison of SfM packages, factors such as image resolution (Nguyen el al., 2012) and camera options (Rasztovits et al., 2013) have also been reported. The variables that effect the evaluation of SfM accuracy include the internal error of ToF (Time of Flight) laser scanner, scaling factor, and alignment algorithms. Problems to be addressed are: 1). The reconstruction sequence of input image makes a difference in the 3D model; 2). The images photographed in a fixed focal length are considered different in SfM.

2.2 Overview of SfM packages

We generally classify the current SfM packages into three categories: web-based solutions, open source programmes and commercial softwares (chart 2.1):

Visual SfM

Visual SfM (Wu, 2011), developed by University of Washington and Google, is an open source programme with implementation of siftGPU(Wu, 2007) and multicore bundle adjustment (Wu et al., 2011). When the feature detection and feature matching are completed, it first generates sparse points with recovered camera poses, and then provides PMVS2 (Furukawa et al., 2010) or CPMVS (Jancosek, 2001) for dense model. The model could be adjusted to real-coordinates via GCP (Ground Control points) or GPS data embeded in photographs. A distinctive feature of Visual SfM is that it provides user interface to show incremental SfM process.

Bundler

Bundler (Snavely et al., 2006; 2008) developd by University of Washington and Microsoft is created for handle huge amout unorganized public image collections, such as those available on Google and Flickr. A well known example is reconstructing famous scense of Rome in less than 24 hours (Agarwal et al., 2009). Bundler yields a set of sparse points with recovered camera parameters, from which dense model could be generated by PMVS2.

Photosynth

Released by Microsoft Live Labs, photosynth (Snavely et al., 2008) is largely based on (Snavely et al., 2006). It is the web-service version of Bundler developed for amateurs of computer vision. In constrast to other SfM systems, Photosynth focuses more on navigation and visual transition among

points cloud and photos. The default package provides the workflow from photos to 3D navigation of sparse 3D model during which no further operation is available. PMVS2 is needed for dense reconstruction.

Arc3D

Arc3D (Tingdahl et al., 2011; Vergauwen et al., 2006), developed by Luwen university, is a fully automatic 3D reconstruction solution for general public. It employs depth map to create textured 3D mesh, and provide to a few options such as image subsampling, image masking, and depth maps for mesh generation. Arc3D does not enforce ordered images and constant camera parameters (Tingdahl D., 2011), but they are still encouraged.

123D Catch

Evolved from PhotoFly, 123D Catch is a web-based solution released by Autodesk. In contrast to other web-based packages such as Arc3D and Photosynth, 123D Catch provides an interface for visualization and modification. Users are allowed to manually stitch the unoriented images and scaling the model according to a known distance. The maximum input images amount hinders its appliance in large-scale reconstruction.

Recap photo

Recap photo is a newly-released Autodesk product developed for reverse engineering. It allows to select photos for modeling and texture mapping respectively. Similarly with 123D Catch, Recap Photo is highly automatic that the workflow is from input images to textured mesh during which no user manipulations are available.

Apero

Apero (Deseilligny et al., 2011), usually coupled with MicMac (Deseilligny et al., 2006) is an open source programme employing SIFT++ feature extractor (Vedaldi et al., 2010) and allows multiple camera models (Brown's, fish eyes, etc.) selection.

Agisoft PhotoScan

Agisoft PhotoScan is a low-cost commercial software. It provides options in each step of the workflow, but the algorithms are unknown due to commercial reasons, such as image enhancing algorithms which facilitates feature extraction. There are two remarkable features that useful to large datasets: 'chunk' and realignment of subset images. Performing global image orientation requires high computation resource and long processing time, and sometimes is impossible to be completed. A 'chunk' is a splitted sub dataset where image orientation, geometry modeling and texture mapping are performed separately. Then resulting 3D model could be combined together. Most SfM systems are incapable to handle the failure of image orientation, while PhotoScan allows to realign a specified subset of images while keeping the successful ones fixed.

2.3 Consideration of SfM features

Although the general workflow of SfM packages are similar, they are developed with different emphasis such as automation/customization, or accuracy/efficiency. For example, web-based solutions are employed in archaeology where automation and efficiency are of prime importance, while the expected accuracy may not meet the needs for architectural survey; In the light of result rendering, a textured mesh is acceptable in most cases, but could be a limitation when mesh and texture mapping are supposed to be generated from different photos. Taking into account of its

applicane in architectural heritage, we conclude several important features of SfM packages: user manipulation, metric accuracy, appliance to complex objects, and efficiency.

2.3.1 User manipulation

SfM systems are generally considered as ‘black box’ for their one-click operation and opaque mechanism (Nguyen et al., 2012; Remondino et al., 2012). It is true for web-based solutions and commercial softwares, but may not apply to open source programmes. Users of web-based solutions are not allowed to control the process but simply upload images and download the results. If the reconstruction succeeds, the models are rendered with texture and post-processing such as ‘fill how’ which optimizes visual completeness with the cost of accuracy. If the failure happens, probably due to the weakness of camera network, users are notified without understanding how to improve it. Commercial softwares provide more operations such as camera calibration parameters, clusters of reconstruction and level of detail, but the unknown algorithms hinders further improvement once the image acquisition phase is over. In contrast, open source programmes enable threshold tuning in steps including feature detection, bundle adjustment, matching accuracy and dense reconstruction. Once understanding the threshold, user could customize the workflow depending on his/her actual need. Besides, as each step is independent with the others, external codes could be in certain step to optimized the workflow. In the workflow of Visual SfM, for example, we could substitute siftGPU, the default sift, to ASIFT (Morel et al., 2009) which performs better when the baseline to depth ratio of input images is high; When the objects have repetitive patterns, we could use SURE instead of PMVS2 in dense reconstruction, as the latter one tends to define and discard redundant images.

2.3.2 Metric accuracy

Geometric accuracy is the top consideration of SfM when applied to metric survey of architectural heritage. The nature of objects, property of dataset (camera configuration, image amount, image resolution) and SfM algorithms (feature detection, feature matching, and dense reconstruction) all affect the accuracy level. Different dataset of the same object lead to different results in the same SfM system, ; Given the same dataset, various SfM packages give rise to 3D models with subtle deviation to obvious distortion or deformation. In order to achieve high accuracy, the camera configuration is supposed to well-designed for web-based solutions and commercial softwares, while open source programmes provide access to additional improvement such as threshold of image contrast in feature detection, error tolerance in feature matching, iterative value in outlier removal, and minimal cameras for model generation.

2.3.3 Application to complex objects

Although SfM packages demonstrate slight difference when applied to small/medium scale reconstruction, such as work of art on desktop or archaeological sites, they make significant distinction in architectural reconstruction as the typology varies. In the former case, the view points are usually organized without any interference, such as traffic, pedestrian, plants and occlusion, all of which are commonly encountered in photographic survey of architecture. More importantly, the camera configuration is strongly effected by the typology of architecture some of which are incapable to be reconstructed or problematic for some SfM systems. Arc3D performs well when the cameras viewpoints compass the object, but tends to fail when the cameras are surrounded by the objects, such as a room or a piazza. In the latter case, only a piece of 2D surface could be reconstructed. We could conclude that the expected camera configuration for Arc3D is of convergent-axis or parallel-axis, but not divergent-axis. 123D Catch is not applicable to complex

objects for the maximum input images is 70, while a five-bay portico, for example, is supposed to reach a satisfactory completeness with around 220 images.

2.3.4 Efficiency

It is not neglectable to keep the processing of SfM in a reasonable time without taking up too much computation resource especially in large-scale reconstruction. Long time of processing and high demand for computer device decrease the utilization of SfM. The factors that influence the processing time include image properties (amount, resolution) and algorithms of SfM (feature detection, feature matching and dense reconstruction). It is possible to speed up the process by tweaking some thresholds with loss of accuracy and completeness, among which the most common approach is image subsampling. Implemented by most SfM systems even web-based solutions, it decreases the processing time in feature detection, feature matching and dense reconstruction. Feature matching is the most time consuming step in large dataset. The global pairwise matching complexity is $O(n^2)$ given the image amount is n . (Wu, 2013) proposes a method that match only the subset of features of a image pair, and skip or continue the matching according to threshold of matching results. Our test show that the feature matching of more than three thousand images that could have lasted several days was reduced to couple of minutes with a slight loss of image orientation number compared to that of full pairwise matching. It could be useful when accuracy is not considered as the most crucial factor.

	Type	Interface	Camera models	User manipulation	Output
Arc3D	web-based	GUI (meshlab)	no	image subsampling	textured mesh
123D Catch	web-based	GUI	no	level of detail manual realignment	textured mesh
Recap Photo	web-based	web page	no	texture mapping	textured mesh
Photosynth	web-based	web page	no	no	sparse points camera parameters
Agisoft	commercial	GUI	multiple	full cycle	sparse points camera parameters dense point texture mesh
Apero	open source	GUI	multiple	full cycle	sparse points camera parameters
Bundler	open source	command line	no	no	sparse points camera parameters
Visual SFM	open source	GUI	fixed	full cycle	sparse points camera parameters

Chart 2.1 A feature overview of SfM packages

2.4 Experiment

The images of the two dataset (chart 2.2) are captured with SLR camera (Nikon D3100) in fixed focal length (18mm). NEF format with highest resolution (4608*3072) is employed for it is more accurate than JPEG and allows further processing such as color calibration and noise reduction,

both of which are applied to the two datasets to improve the image orientation. The configuration of camera network is based on the rule of three-each surface is covered by at least three cameras, and consideration of accuracy-convergent camera network with camera rotation. The reference model is obtained by Leica C10 laser scanner with 20mm resolution.



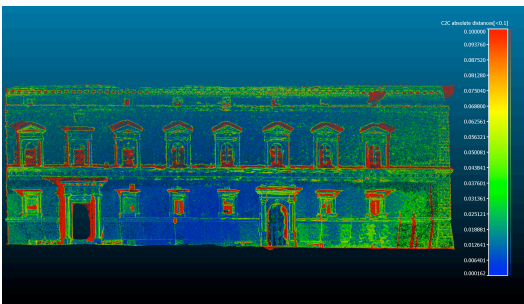
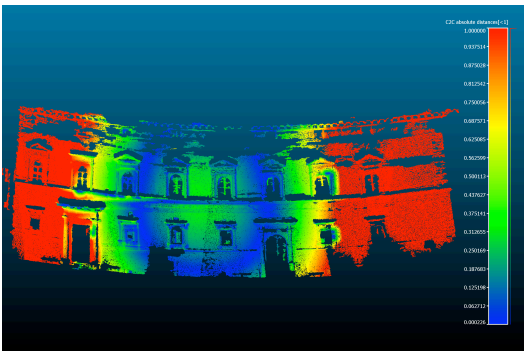

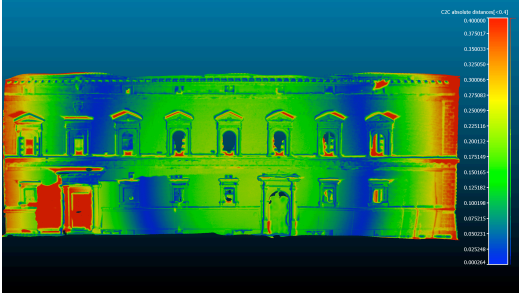

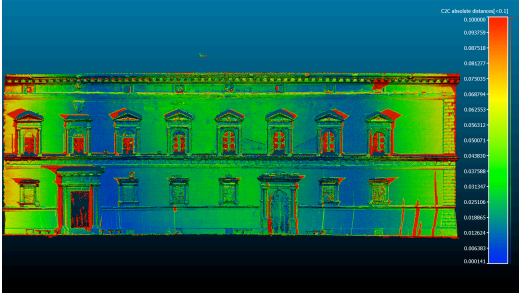


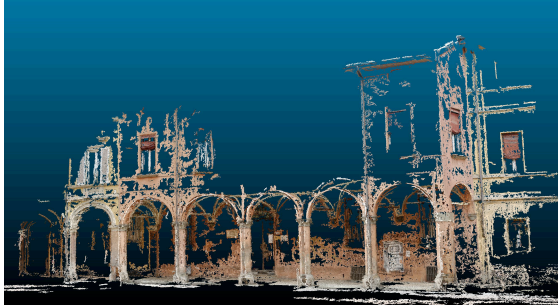
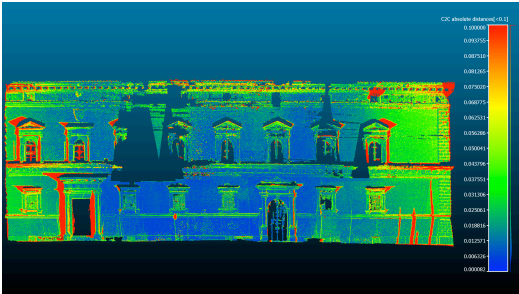

Dataset	Image amount	Resolution (pixel)	Camera	Focal length (mm)	Image filter	3D visualization
Facade	59	4608*3072	Nikon D3100	18	noise reduction color calibration	
Portico	212	4608*3072	Nikon D3100	18	noise reduction color calibration	

Chart 2.2 Description of two datasets for SfM evaluation

Once the SfM models are generated, they are roughly aligned to the laser scanned model, and cleaned with the same mask to remove variables in approximate deviation comparison-mean deviation and standard deviation (chart 2.3). An iterative refinement of alignment is applied to further improve the reliability of alignment. In the end the geometric deviation between SfM models and reference is shown in red-blue map.

SfM	Facade-accuracy	Portico-completeness
123D Catch		Exceed maximum input image (70)
Arc3D		

SfM	Facade-accuracy	Portico-completeness
Recap		
Agisoft		
Photosynth +PMVS2		
Visual SfM +PMVS2		

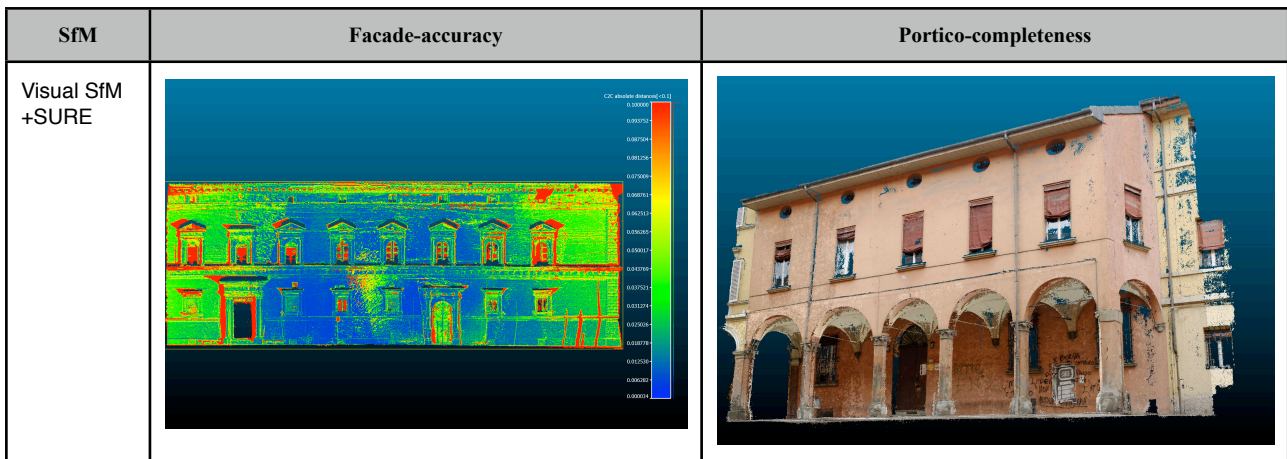


Chart 2.3. A visual comparison of SfM packages in two datasets. Dense reconstruction packages (PMVS2, SURE) are employed to complete the 3D models.

SfM	facade			portico	
	time (min.)	Mean dist. (mm)	Standard dist. (mm)	time (min.)	oriented images
123D Catch	40	24	104	-	-
Arc3D	107	926	4400	180	19/212
Recap Photo	139	142	216	665	-
Agisoft	103	23	109	523	212/212
Photosynth+PMVS2	133	44	182	723	194/212
Visual SfM+PMVS2	90	13	80	610	188/212
Visual SfM+SURE	65	15	86	210	188/212

Chart 2.4. A quality comparison of SfM packages in two datasets. 123D Catch is failed to generate model in 'portico' dataset for more than 70 images is not allowed. The camera information is unknown in Recap Photo

The tested SfM packages perform distinctively in the two types of architecture (chart 2.4). In the 'facade' dataset, distortion is observed in most SfM packages, among which the Arc3D model has the most apparent distortion with 3-4m deviation (figure 2.1) in two sides probably due to the different depth of image series. Visual SfM has the most slight discrepancy with mean deviation of 13mm. Photosynth+PMVS2 and Visual SfM + PMVS2 yields model with holes. The incompleteness was caused by redundancy removal of PMVS2. Misalignment is observed in the Recap Photo model. The highly repetitive structure of classical architecture leads to outliers which could be removed by tweaking the RANSAC thresholds in open source softwares, but problematic web-based solution. More distinctions are presented in the 'portico' dataset. Arc3D yields only a piece of distorted model because of the complex camera configuration. Photosynth orients the second most images following Agisoft PhotoScan, but generates model with low completeness and long processing time which are not suggested for even low-accuracy focused survey. Visual SfM yield different results when coupled with PMVS2 and SURE in dense reconstruction. SURE is more robust and efficient than PMVS2. Recap Photo performs astonishingly well, but the maximum input image amount (300) might impedes its appliance large datasets. In general, Agisoft PhotoScan covers the whole workflow rendering texture model and is robust in image orientation even when low-texture surfaces exist, but the unknown algorithms hinders further improvement when image acquisition is completed; Visual SfM+SURE is suggested for metric survey for its accuracy, high-

completeness, reasonable efficiency, and user manipulation in each step; Recap Photo is an option when automation is emphasized.



Figure 2.1. Modeling errors generated by web-based SfM packages. Left: Incorrect wall position in model from Recap Photo; Right: Obvious distortion in model from Arc3D

Chapter 3. Workflow of image-based modeling

Image-based modeling is the process that produces 3D models from collections of correlated images via digital cameras. The general workflow consists of camera calibration, image acquisition, image filters, feature correspondence, and dense reconstruction. The on-site work begins with image acquisition, but a pre-planning that depends on the expected level of accuracy on one hand, and analysis of objects and context on the other is always necessary. Although the main principles of photographic acquisition are constant, strategies may vary from case to case. Each procedure of the workflow will be discussed in the following contents.

3.1 Camera calibration

Calibrated camera plays a fundamental role in image-based modeling. Both intrinsic parameters and extrinsic parameters of camera are supposed to be estimated during the camera calibration. From a photogrammetric point of view, a camera is considered calibrated if the principle distance, principle point offset, and lens distortion parameters are known (Remondino et al., 2006). While computer vision community employs self-calibration algorithms (Pollefev et al., 1999) which initially supposes the camera parameters are unknown, and only extract focal length and doubtful distortion parameter at the end of camera recovery. An instance is that the coded calibration of Iwitness yields three parameters of distortion (k_1, k_2, k_3), while Visual SfM allows only input parameter.

3.1.1 Targetless calibration

The current state employs coded targets to accurately determine image correspondence, while objects with distinctive features could also be the targets of calibration if certain image acquisition guidelines are obeyed (Barazzetti, et al., 2010). By calibrating cameras from the on-site images, targetless calibration could speed up the workflow while maintaining the accuracy in a similar level. Targetless calibration requires high favorable camera configuration as follows:

- Highly convergent network - larger convergent angle leads to larger base-to-depth ratio
- Orthogonal camera rotation - break the projective coupon between IO and EO parameter
- Various distance - maximum possibility of image scale variation
- Eight or more images - ensure sufficient feature correspondence

3.2 Image acquisition

3.2.1 On-site camera setting

Focal length

The length of focal lens determines the expected accuracy and the amount of images needed to cover the object. Given the distance and image resolution fixed, the longer the focal length, the higher the accuracy of the reconstruction: f/z (f : focal length; z : distance between camera to object) Wider lenses (18 to 30mm) provides larger field of view and coherently less images, while longer lenses capture more details but need more images, and potentially higher possibilities of error in image undistortion The selection of lens depends on the situation of field and accuracy expectation. Fixed focal length is always a prerequisite of photogrammetry, but optional for SfM packages, some of which encourages constant focal length such as Recap Photo and Arc3D, while some allows multiple lenses such as Visual SfM, Bundler, Agisoft PhotoScan.

Focus method

Most SLR provides both manual focus and auto focus for focalization. Although auto focus provides a wide range of modes (single-point, dynamic-area auto-area, 3D tracking) covering various situations, manual focus is still suggested to avoid subtle change of focus length due to the change of focus points and the distance in-between.

Aperture

Aperture controls the amount of light that comes into the sensor, and the depth of field (The point at which begins to blur). Small aperture ensures the entire object is in good focus and sharpness. For this reason, small aperture (large F-stop e.g. >4.8) is suggested.

Shutter speed

Shutter speed is automatically determined when the camera is in aperture priority mode. But it is better to bear in mind that at least speed of 1/200 second is required for handheld photography to avoid image blur caused by hand shaking.

ISO

ISO is the sensitivity of sensor to light. Higher ISO is used in dark space to get faster shutter speed, but it also leads to more noise in the image. Therefore, lowest ISO (usually 100) is suggested.

Image format

Raw/NEF is the suggested format for it is not compressed, allows modification of white balance, exposure without quality decrease, and lead to more accurate modeling results (Stamatopoulos et al., 2012).

Image resolution

Higher image resolution would benefit feature extraction, and potentially lead to more oriented images (figure 2.2). Correspondingly more computation resource and time is needed with larger image resolution.



Figure 2.2 Dense-reconstructed points generated by same photo set but compressed with different resolution. The left model was reconstructed from photos with 1024 pixel in larger dimension, the right one 2048 pixel. It showed that larger resolution leads to higher density of points cloud.

3.2.2 Camera configuration

The design of camera network draws a significant influence on the result of image-based modeling in terms of completeness and accuracy. A well-designed camera network for architectural heritage should achieve a balance between completeness and accuracy, as well as reasonable efficiency which is favorable in large scale reconstruction. Full automation has been realized in the image-based modeling workflow, but design of camera network is still a challenging task to much extent based on experience of operators. Automatic approach has been proposed (Alsadik et al., 2013), but it requires a preliminary reconstruction with low-resolution video frames and only applies to small objects with simple shapes.

From a traditional photogrammetric point of view, the accuracy of network increases with the increase of base to depth ratio (B:D), which explicitly lead to a convergent camera network (Wenzel et al., 2013).

Terrestrial camera network is strongly effected by the character of the surveyed architecture (figure 2.3). The camera view points are supposed to spread out along the expansion of architectural layout, and yield the field of view of focal length in the limited vacancy of architecture or narrow space with adjacents objects. Pedestrians, cars, and plants are potential occlusions and obstacles to ideal viewpoints. Image color balance alters subtly due to the interaction between illumination and form as well as diffuse of building material. Another neglectable issue is to define the minimum camera network which is critical in city-scale reconstruction.

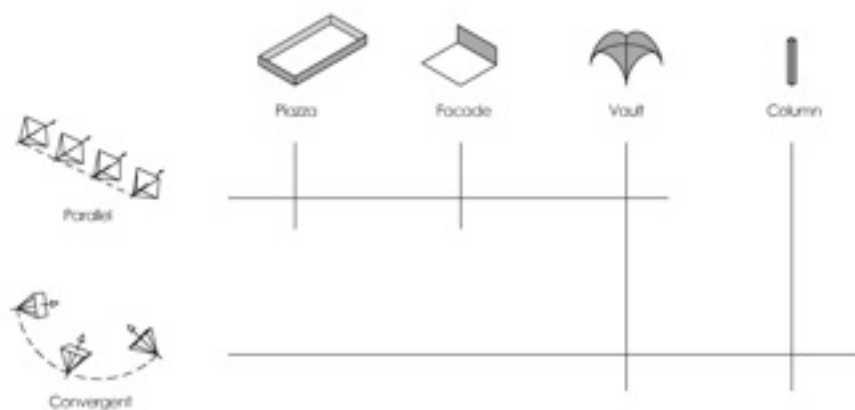


Figure 2.3 Application of camera network in various types of architecture or component

3.2.2.1 Parallel-axis camera network

Parallel-axis camera network eliminates the influence of image tilting in convergent network and guarantees high similarities in adjacent images which is beneficial to feature matching and potentially gives rise to more 3D points. For this reason, parallel-axis camera network is suggested by tutorials of SfM softwares such as (Photosynth Photography Guide, 2008; Agisoft PhotoScan Image Capture Tips, 2010; Recap Photo Getting Started Guide, 2013) in capturing planar surface. Another advantage of parallel-axis-images is the average image weight in texture mapping. The drawback of parallel-axis camera network is it provides poor depth information for recovery of camera poses and distortion parameters. Besides, planar surface is rare in architectural heritage, even buildings facade contains surfaces that such as windows, balcony which could be missed. Therefore, parallel-axis could be applied to 3D visualization with low accuracy expectation to speed

up the workflow, but not metric survey. The step size (base line) of parallel-axis should be determined with respect to the depth, focal length, image resolution and expected precision (Wenzel et al., 2013). Complex shape and level of distinction of texture also essential considerations.

3.2.2.2 Convergent camera network

Convergent camera network leads to small base to depth ratio which is favorable to geometric accuracy. It is one of the necessities of image acquisition for camera calibration. Convergent camera network undoubtedly apply to reconstruction of statue, column or villa where the objects are encircled by the cameras in any case. Given a planar surface such as building facade for example, however, dividing the surface into several pieces and setting convergent network for each piece is not reasonable. It not only demands much more images than parallel-axis network, but also increase the ratio of divergent cameras which may reduce the level of accuracy. Consequently initializing the SfM process from a convergent part and resuming the rest in parallel-axis network is suggested (chapter 4.1).

3.2.2.3 Divergent camera network

Divergent camera network is not a necessity but a by-product in complex architectural scene, such as portico whose column is supposed to be encircled by convergent cameras but inevitably divergent ones from adjacent columns. There is no evidence that the existence of divergent cameras decrease the accuracy level in a camera network where parallel-axis cameras and convergent cameras also present, but it would definitely fail to reconstruct 3D model from images photographed from a fixed stand point simply rotating towards different orientations. The reason of the failure is the low ratio between base-line and depth instead of simply stationary viewpoint. Failure would still happen when base-line exist but relatively small compared to the depth, for instance photographing bird view images from a 70-meter-high tower with an approximate square plan (5m*5m). The maximum base-line is 5 meter (no diagonal views are available as obstacle exists in the center of the plan), while the minimal depth is 70 meter. The base to depth ratio is too low to provide information for depth determination.

3.2.3 Instance of failure

Besides the inherent drawback of camera configuration, there are factors one should try to avoid in image acquisition. A negative example is reconstructing a 300-meter-long portico with a hand-held camera in video mode in order to speed up the workflow. The idea is to walk along the portico while keeping the cameras oriented to certain semantic parts - for example facade, vault and wall - of portico. Supposing six rounds are needed to cover all surfaces and each round takes 4 minutes, the time of image capture is less than half an hour which is much more efficient than photographing. Unfortunately the experiments failed due to following reasons:

3.2.3.1 Image blur

Image blur caused by shake of hand-held camera should be avoided, as it leads to incorrect feature detection. Currently anti-blur algorithms are not reliable to recover the damage precisely. In order to avoid the image blur, tripod is always suggested to use if possible. Reducing exposure time also reduces the chance of blur, but may cause insufficient brightness in frame.

3.2.3.2 Low resolution

Although SfM techniques are developed to utilize low-resolution images and medium-resolution images with incomplete EXIF data downloaded from Internet, problems occur when the texture of the object is less distinctive such as white-painted wall or vault built by plaster. Currently SLR

cameras allows maximum video output as 1080 pixel, while the same parameter of photo is larger than 5000 pixel. It is suggested to employ high resolution images for low-texture surfaces or enlarge the ratio of object in the frame by zooming in or moving closer.

3.3 Image processing

In spite the correct camera setting and photographing skills, the direct output of on-site image acquisition is not qualified for initializing the modeling process for the deviation of white balance and incorrect exposure. Our test show the improvement Sharpness, A set of image filters are employed to calibrate the deviation caused by on-site illumination and situation, and to enhance the image features to ease the feature correspondence.

3.3.1 Color calibration

The same object may have tremendous color deviation with different camera set or illumination conditions. Color calibration is hence necessary to discard such effects and lead to a uniform texture. In addition, our test show that color calibration also benefit to feature correspondence and 3D sparse reconstruction. The workflow of color calibration is:

1. Image capture plan

Photos are divided into groups with respect to semantic components and shooting angles. A set of photos that share the same color profile should has similar illumination. For instance, the images perpendicular towards building facade should employ different color profile with those photographed with 45 degree.

2. Modify camera set (Raw, Adobe RGB)

Raw format is more accurate than jpg format (Stamatopoulos, et al., 2012) in the light of color space and hence add less noise to the resulting points cloud. Besides, it allows much possibilities for further modification such as exposure and white balance. Adobe RGB has larger range of color space than sRGB, so we use Adobe RGB in image acquisition and sRGB as input for SfM.

3. Image capture

A photo with ColorChecker in frame is photographed in each group (figure 2.5). In each group, photos are shot keeping camera set and shooting angle fixed to get as similar illumination as possible. In addition, the ColorChecker should has even illumination, and avoid brighter object than the white square of ColorChecker in the frame.



Figure 2.4 Image containing ColorChecker in the frame

4. White balance adjustment

Photos too light or too dark would reduce the detected features. Hence we find the suitable exposure and apply it to the other photos in the group by adjusting white balance value of the ColorChecker (figure 2.5). In the end photos are stored as TIF format to avoid data loss caused by compression in JPG format.

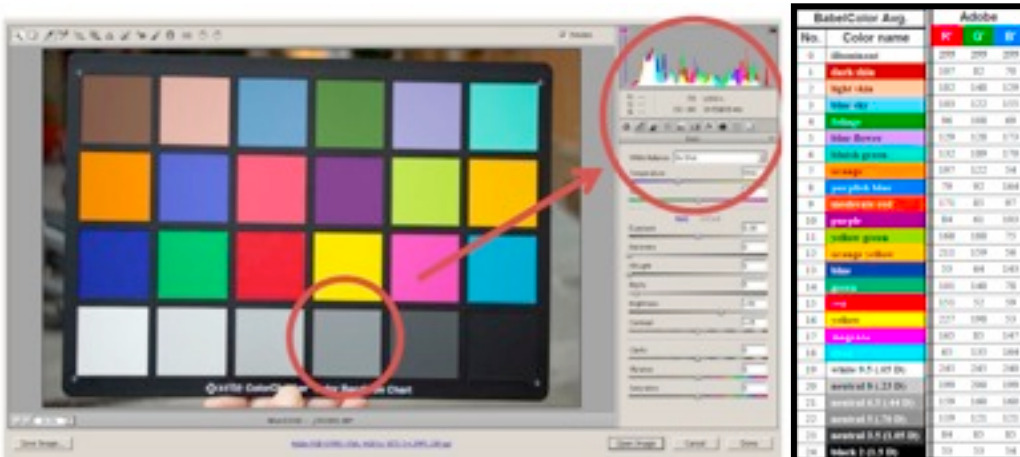


Figure 2.5 Adjusting the white balance and exposure according to the value of ColorChecker

5. Color profile extraction and assignment

We extract the color profile of the photo with ColorChecker via ProfileMaker, and assign the same profile to the other photos in the group.

6. Result evaluation

We use Imatest to evaluate the photo with ColorChecker that has been profile assigned (figure 2.6). The crucial parameters include error and exposure. If the error is a bit higher than expected value, we could reassign the profile to the photo with ColorChecker. If the value is far beyond the acceptable range, we may have to restart from the very beginning.

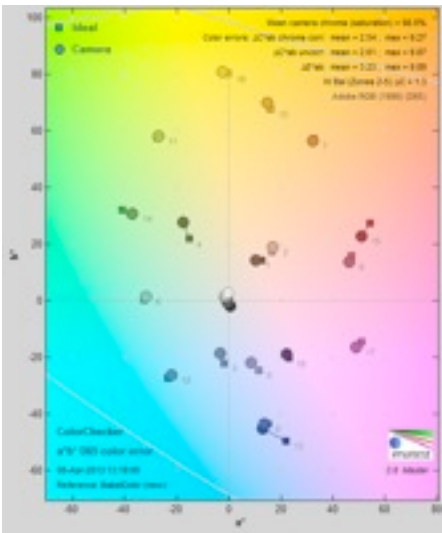


Figure 2.6. Color calibration report which provides error value and show the deviation in color space

7. Image output

If the evaluation is satisfactory, we convert the color space to sRGB and export the photos in JPG format, because by default OpenGL employs sRGB for 3D visualization.

3.3.2 Noise reduction

Image noise is a by-product of image acquisition when the digital sensor attempts to record tiny amounts of light. Generally higher ISO leads to more noise in the frame. The noise could interference the feature detector which extracts in a pixel level. We employ Noiseware (<http://www.imagenomic.com/about.aspx>) to reduce the noise in frame (figure 2.7). The experiments show improvement in terms of completeness and noisy points removal in the resulting model.



Figure 2.7 Original image (left) and noise-reduced-image (right).

3.3.3 Wallis filter

The Wallis filter is a digital processing function that enhances the contrast levels of an image by retaining the edge details and removing low-frequency information. Studies have found that interest operators typically find more suitable points on imagery that has been pre-processed with the Wallis filter (Remondino, 2006). We apply wallis filters to image for flatten the different exposure and acquire more features on the low-texture areas (figure 2.8).

The algorithm is adaptive and adjusts pixel brightness values in local areas only, as opposed to a global contrast filter which applies the same level of contrast throughout an entire image. The resulting image contains greater detail in both low and high level contrast regions concurrently, ensuring that good local enhancement is achieved throughout the entire image. The Wallis filter requires the user to input a target mean and standard deviation, and adjusts local areas to match the target values accordingly. A weight value is used to control to what extent the target values are used and how much of the original image is kept.



Figure 2.8 Comparison of original images and wallis-filtered images. Wallis-filter flattens the different exposure and reveal more details on the low-texture areas. Left: original image got correct exposure in the doorway while overexposed in the court yard; Middle left: the same image with wallis-filter; Middle right: original image got correct exposure in the court yard while underexposed in the court yard; Right: the same image with wallis-filter.

3.4 Feature correspondence

3.4.1 Feature detection

The image orientation of photogrammetry and computer vision relies on feature extraction of input images. The correspondence stage first finds interest points by feature detectors and feature descriptors, and then match the interest points between pairs of images (figure 2.9). Lots of algorithms were developed to robust transformation of illumination, geometric transition (Remondino, 2006).



Figure 2.9 Extracted features shown in vectors representing the location and intensity

(Lowe, 2004)'s SIFT (Scale Invariant Feature Transformation) algorithm transforms the features into a collection of vectors that are invariant to image transformation and partially invariant to change of illumination and viewpoints. The result of feature detection is relevant to a set of factors ranging from image quality (dimension, sharpness, contrast, exposure) to object and ambient (texture and reflection). One of the advantages of SIFT is that it detects features at multiple scales, for instance images with the same resolution but different depths, which is very common in image acquisition.

ASIFT (Morel et al., 2009) aims to correct the SIFT problem of view angle is too different: if the object is under view has similar illumination conditions, has a rather flat surface, and is not a mirror, then ASIFT retrieves the object even under extreme changes of angle. In technical terms, ASIFT is more affine invariant than SIFT.

LADASHash (Strecha et al., 2012) is SIFT-like local feature binary descriptors that map the descriptor vectors into the Hamming space, in which the Hamming metric is used to compare the resulting representations. To better take advantage of training data composed of interest point descriptors corresponding to multiple 3D points seen under different views, it has introduced a global optimization scheme. The approach first aligns the SIFT descriptors according to the problem-specific covariance structure. In the resulting vector space, all SIFT descriptors have diagonal covariance. The advantage of LADASHash is that it deduces the size of the descriptors by representing them as short binary strings and learns descriptor invariance from examples. This approach is very fast and can be used for many other applications for which similar training data is available.

SGLOH (Mikolajczyk et al., 2005) differs from SIFT in the respects of being computed over a log-polar grid as opposed to a rectangular grid, using a larger number of 16 bins for quantizing the gradient directions as opposed to 8 bins as used in the regular SIFT descriptor, and using principal component analysis to reduce the dimensionality of the image descriptor. From their experimental results, the authors argued that the GLOH descriptor leads to better performance for point matching on structured scenes whereas the SIFT descriptor performed better on textured scenes.

3.4.2 Feature matching

Once feature detection is completed, matching is processed among pairs of image (figure 2.10). In contrast to feature detection which plays a fundamental role in accuracy of resulting model, feature matching to much extent determines the processing time of the pipeline. Given the image amount n , the time needed for global matching is $O(n^2)$. Instead of global matching, it is possible to match only the sequential images within a certain amount which is useful for video frames, or geometrically related images according to GPS data. (Wu , 2013) proposed an approach to match only a subset of image features, and then decide to skip or resume the full matching. Similarly, (Snavely et al., 2008) presented to compute a small skeletal subset of images, and adds the remaining images using pose estimation. Thresholds regarding to feature matching include error tollerante with the unit of pixel. Smaller error tollerante leads to more accurate inliers matching, while tweaking the parameters allows more matches and potentially more sparse 3D points.

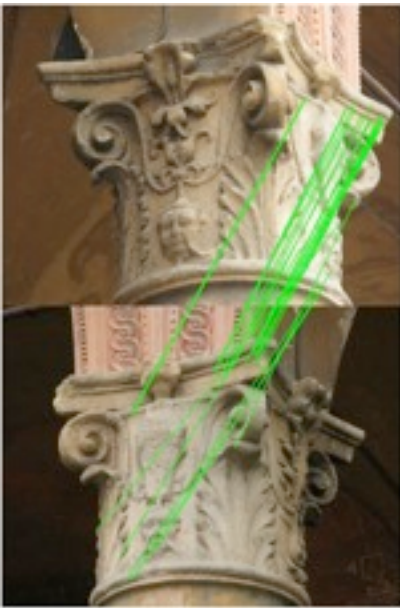


Figure 2.10 Feature correspondence between a pair of images.

3.4.3 Incremental SfM

The SfM process by default initializes from an image pair that has large number of matches and large baseline. A 3D point could be reconstructed from its two projections by computing the intersection of the two space corresponding to it. New images are continuously triangulated with existing images during which outliers are removed and bundle adjustment is applied, until no more image yields the threshold of minila inlier matches. In the end of the SfM process, a set of sparse points with recovered camera poses are output (figure 2.12). The imagery sequence of the SfM process makes different results from the same photo set. For this reason, it would contribute to the accuracy by initializing the SfM process from a image pair with large base-to-depth ratio and abundant distinctions. The outliers - incorrectly matched features - are supposed to be removed in incremental SfM process by RANSAC. The random seed of the algorithm also leads to inconstant 3D points.



Figure 2.11 Output of SfM process: sparse points and recovered camera poses

3.5 Dense image matching

In the dense image matching stage, images that are oriented in the previous steps are further matched to yield a complete 3D models. As both the image orientation and camera parameters estimation have been completed, the difference of various dense image matching algorithms mainly exist in the completeness rather than accuracy. Besides the nature of the algorithms, the repetitive pattern of architecture, image acquisition angle with the surface of object, and the material of the surface all exert influence on the result of the dense image matching.

3.5.1 State of art

(Strecha et al., 2004) proposed a PED-based formulation which is applicable to a small set wide-baseline images with high resolution. (Furuawa et al., 2010) proposed a patch-based multi view stereo matching approach. The initialization does not rely on any prior estimation of the surface, but follows a multi-step approach. A patch is a local tangent rectangle approximation of a surface that geometrically determined by its center, normal vector oriented towards the cameras and a reference image in which the patch is visible. The surface growing is decomposed into clusters to save memory demand. Micmac (Dellepiane et al., 2011) implements a coarse to fine modification of the maximum flow matching algorithm proposed by (Roy et al., 1998). A patch in the master image is first identified from the potential 3D points and then projected onto adjacent images for global matching. The procedure was originally developed to deal with high-resolution satellite images, but now it can also handle complex terrestrial datasets. Micmac enforces users to select a set of master images for the correlation procedure. SURE (Rothermel et al., 2012) uses a semi-global matching approach (Hirschmuller, 2008) to compare depth map for only adjacent image pairs. The memory demand and processing time is relatively low. (Jancosek et al., 2011) presented an approach to achieve completeness in low-texture areas. It uses a visual-hull to reconstruct the weakly-supported surface without dense points. Besides the open source solutions, the dense image matching algorithms of commercial softwares such as Agisoft PhotoScan is unknown for commercial reasons.

(Seitz et al., 2006) categorized the algorithms according to six properties: the scene representation, photo consistency measure, visibility model, shape prior, reconstruction algorithm and initialization requirement. Based on the datasets of small-size hand crafts, sub-millimeter translational offsets are

reported among the tested algorithms (Kolmogorov et al., 2002; Vomitassi et al., 2005; Pons et al., 2005; Goedele et al., 2006; Furukawa et al., 2006). (Strecha et al., 2008) compared dense image matching algorithms in outdoor 3D reconstruction, and revealed that (Furukawa et al., 2007) performed better in accuracy and completeness than the other algorithms (Strecha et al., 2004; 2006). (Recondito et al., 2013) presents a critical review and analysis on SURE, Micmac, PMVS and Agisoft with datasets vary in camera configuration (convergent and parallel-axis) and image scale. The tested algorithms have advantages and limitations, and highly depend on the input parameters.

3.5.2 Experiment

We apply PMVS, SURE and CMPMVS to the dataset of portico where the plaster vault and repetitive facade pattern cause potential challenges to the completeness of the model (chart 2.5). 212 images are photographed to ensure sufficient converge of the surfaces. Images are color-calibrated and processed with noise filter which are proved to enhance the feature correspondence. (Vedaldi et al., 2010)'s vlfeat (peak=6) is utilized to extract features, and the sparse model is built in Visual SfM (Wu, 2011). The output of Visual SfM is input of the tested dense image matching approaches.

Dataset	Image amount	Resolution (pixel)	Camera	Focal length (mm)	Image filter	SfM and sift
Portico	212	4608*3072	Nikon D3100	18	noise reduction color calibration	Visual SfM VLfeat (peak=6)

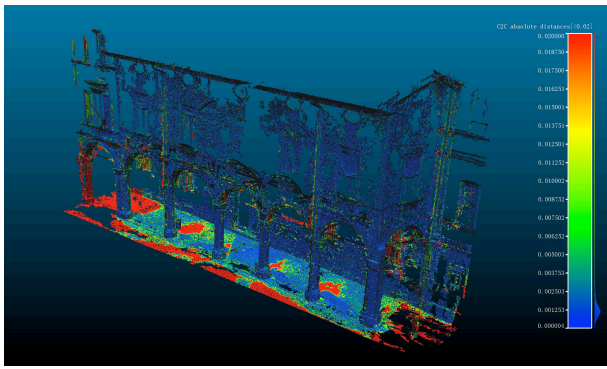
Chart 2.5. Description of dataset properties.



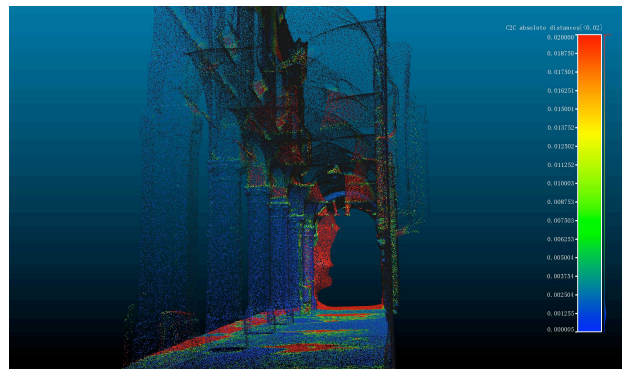
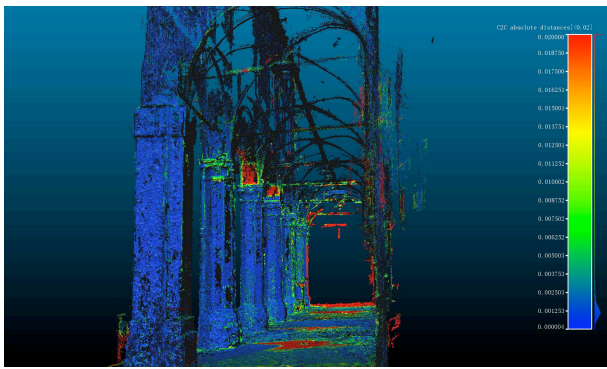
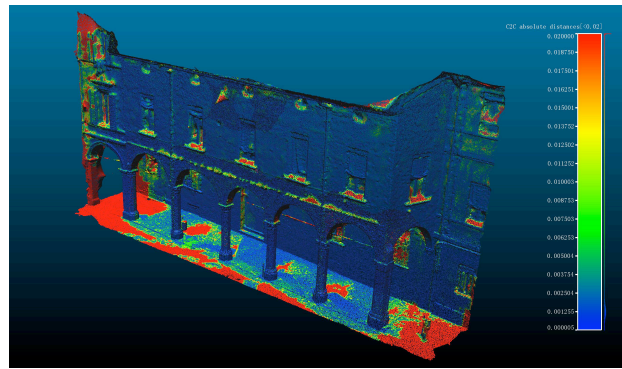
Figure 2.12. A parallel visualization of dense image matching results. From left to right: SURE, PMVS, CMPMVS. The deformation of vault in the CMPMVS model is caused by visual-hull reconstruction in weak surfaces with sparse input points.

The model from SURE demonstrates the best completeness, while small empty areas are left on the pavement and on top of the columns, both of which are reconstructed in the model of PMVS and CMPMVS (figure 2.12). The holes on pavement are caused by the sharp image acquisition angle between camera and the object's surface, while the one on the column are due to insufficient image coverage. SURE tends to require more image with small baselines for modeling isolated small surfaces given the same conditions than PMVS and CMPMVS. The PMVS model is poorly reconstructed with huge blank on the facade and vault. PMVS discards redundant images to speed up the process, but the removal tends to be incorrect in the case of architectural facade with repetitive pattern. CMPMVS output textured mesh instead of points cloud. The points cloud are sampled from the surface of the mesh for comparison. Deformation is observed on the vaults where the input points are sparse (chart 2.6).

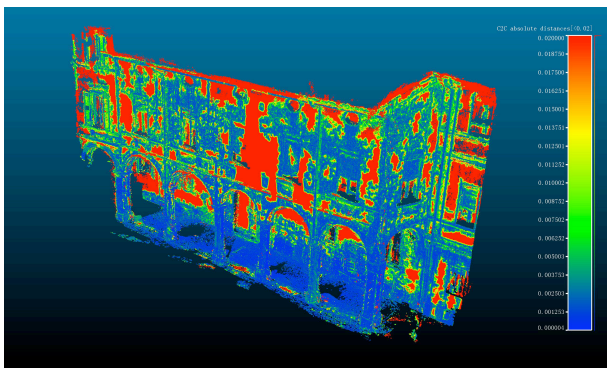
Reference: SURE. Visualized: PMVS



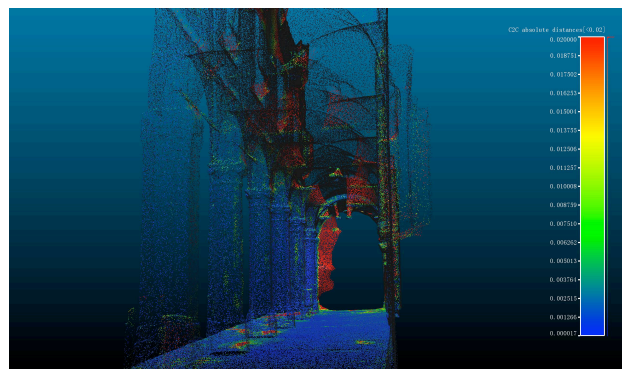
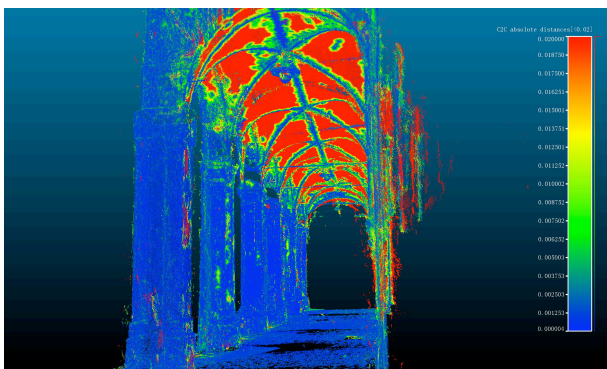
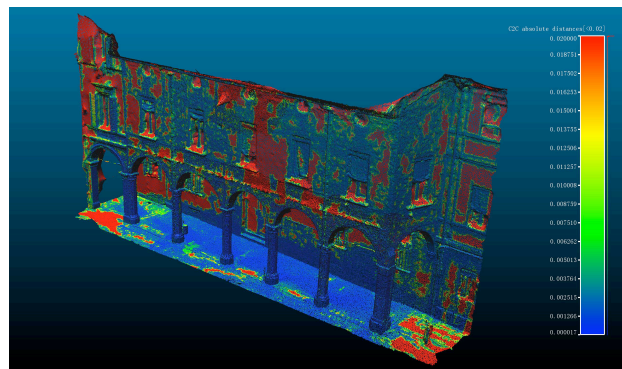
Reference: SURE. Visualized: CMPMVS



Reference: PMVS. Visualized: SURE



Reference: PMVS. Visualized: CMPMVS



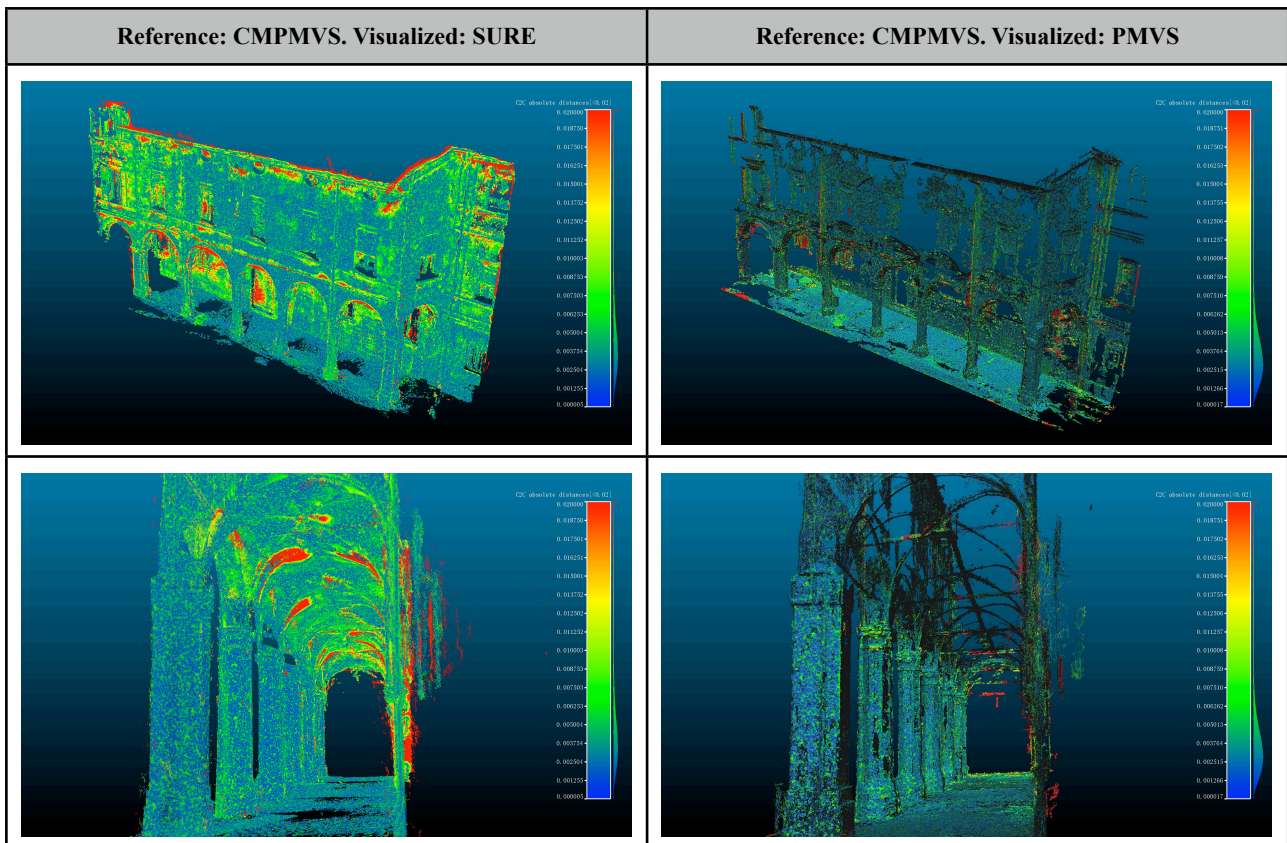


Chart 2.6. Completeness comparisons of SURE, PMVS and CMPMVS with respect to each other. The deviation is shown in red-blue map. Red areas represents the surfaces that miss in visualized model but exist in the reference model.

3.6 Further processing

3.6.1 Surface reconstruction

The output of range-based modeling and most SfM approaches output points could in the end. Points cloud are a set of points each defined by x, y, z coordinates, and often retains RGB color value and normal vector. While points cloud can be visualized, analyzed and measured, they are usually converted to 3D surfaces for further processing to cater various purposes, such as curvature analysis, extraction of profile, texture mapping and segmentation. The challenge of surface reconstruction relate to points sampling, noise cleaning and uncovered areas.



Figure 2.13. Points cloud cleaning. Left: Original points cloud; Middle: cleaning by vertex RGB color filter. Noisy white points on the upper edge of the facade ' $R > 245$ ' are selected and removed. Right: points cloud after cleaning

Before surface reconstruction, points cloud have to be processed for better handling and noise removal. Subsampling is necessary in order to maintain the points cloud with in a reasonable size. Points are subsampled in terms of metric distance or expected amount. Noise removal is operated by manual operation or filters according to the vertex attributes such as RGB color (figure 2.13). For dense image matching algorithms that produce separate models from corresponding image depth map, an alignment could be applied instead of simply merging the points to decrease the errors of image orientation (figure 2.14).

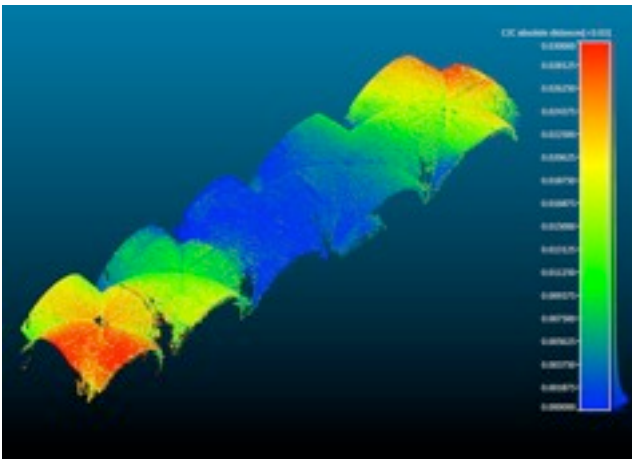


Figure 2.14. Points deviation between the direct output of dense image matching and the aligned points cloud. A tendency of deviation is observed from the referenced model to two sides.

Various algorithms of surface reconstruction exist in the field, such as Delaunay (Bassinet et al., 1984; Kolluri et al., 2004), Alpha shapes (Edelsbrunner et al., 1994; Bajaj et al., 1995; Bernardini et al., 1994), and Moroni diagrams (Amenta et al., 1998; 2001). We used screened poisson surface reconstruction (Kazhdan et al., 2006) which provides a set of parameters according to the ratio of noise in the points cloud. High depth gives rise to more sharpness in the edges such as relief or cornice but more noise on the flat surfaces such walls, and vice versa. A value of 10 is acceptable in most cases (figure 2.15). Once the mesh is built, operations such as noise removal, smoothing (Herrmann, 1976), and decimation could be applied.

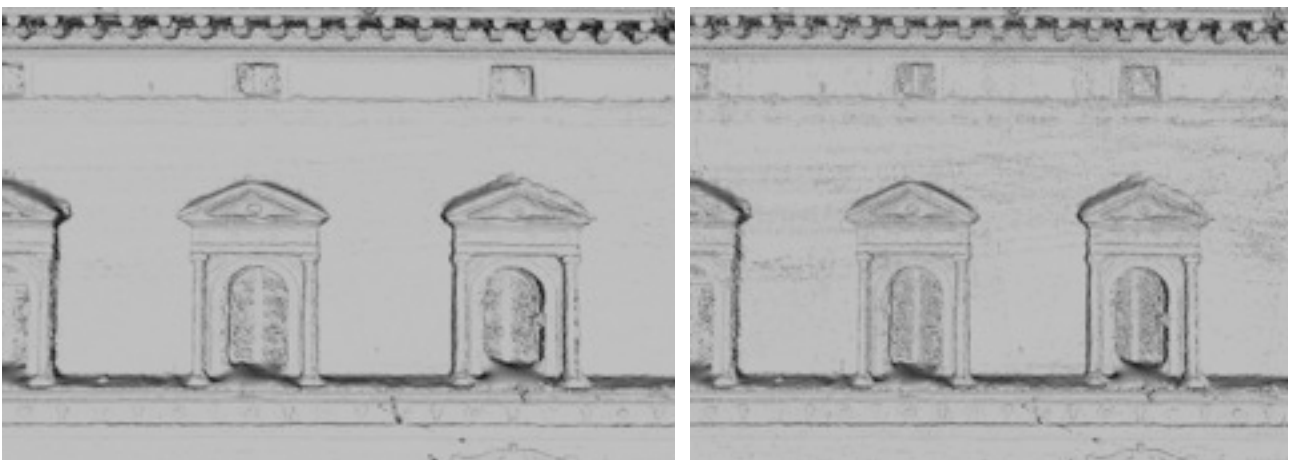


Figure 2.15 Surface reconstruction using Screened poisson (Kazhdan et al., 2006) with depth 11 (left) and depth 12 (right).

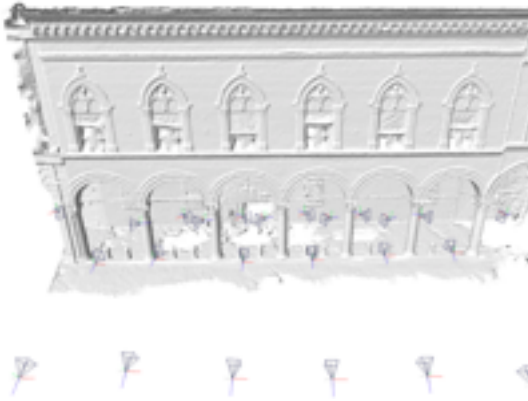


Figure 2.16 A sub set of cameras are employed for texture mapping.

3.6.2 Texture mapping

Although color calibration of all the photos would improve the feature correspondence, it makes no sense to employ all of them for texture mapping. Excessive photos are proved to be vulnerable to computer crash and would bring chaos in color distribution. A small amount of photos that well-covered all the surfaces and well color-calibrated would be qualified for texture mapping (figure 2.16).

Meshlab offers two ways of texture mapping, and they are proved to work in different situations:

1. Project raster colors onto mesh and filling the texture.
2. Parametrization + texturing from registered rasters (Pietroni et al., 2010) (figure 2.17)



Figure 2.17 Mesh and textured mesh with different algorithms. Mesh without texture (left). Mesh textured by perspective-projected cameras (middle). Mesh textured in parametrization (right)

The first filter perspective-project color information from all the active rasters on the mesh and filling the texture. It generates well-distributed texture on plane surfaces with fine detail (figure 2.18), but lead to chaos of texture when objects have different depth of view. Because of the projection of perspective instead of orthogonal, it projects the texture of the object closer to the camera to the object further. In the light of architectural scene, this approach is applied to simple and plane geometry such as facade, but not complicated scene with different depth. In contrasts to the first filter, the second filter is proved to be effective in complex shape and topology. Tailored to minimize isometric distortion, the mesh is parameterized and textured by creating some patches that correspond to projection of portions of surfaces onto the set of registered rasters.



Figure 2.18 Comparison of two methods of texture mapping. Left: Color projection Right: Parametrization. In the left image, a trash can is incorrectly projected onto the wall.


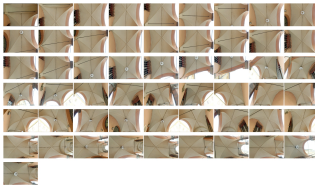


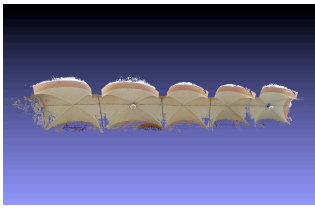
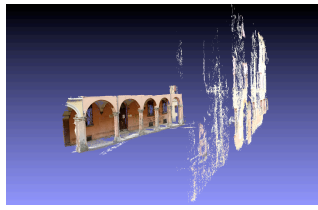

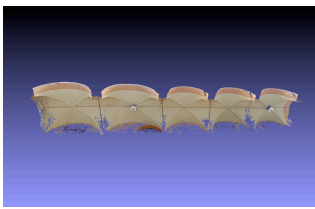

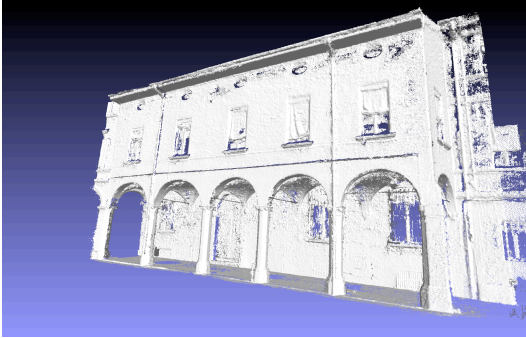
3.6.3 Segmentation

Segmentation is a critical step that transforms the raw survey data to semantic BIM objects. Range-based modeling captures data within a specified range of radius, while image-based modeling yields points from pixels in the frame as long as observed by two cameras. Both of the approaches could give rise to huge amount data without any consideration of the semantic organization of the surveyed architecture. Segmentation is necessary to maintain the data in a reasonable size to ease the operation, distinguish the subtle discrepancies between the components of same hierarchy, and cater the objects-oriented modeling, the mechanism of BIM platforms.

Automatic segmentation techniques have been developed in the filed of computer vision (Atene et al., 2006; Benhabiles et al., 2009; Chen et al., 2009). In the filed of architecture, relevant techniques consists of color similarities and spatial proximities (Zhana et al., 2009), shape detection (Ning et al, 2010), and distance measurement among planar surfaces (Dorninger et al., 2007). Appropriate techniques that apply to cultural heritage barely exist (Whiter et al., 2013). Furthermore, the automatic segmentation have limitations due to geometric complexity and data ambiguities (Barazzetti et al., 2010). Hence, manual intervention is still needed in most cases.

The segmentation of architecture is highly relied on the knowledge of the construction grammar. Prior-knowledge is needed to distinguish a component from the others. Treatises of architecture play a crucial role in understanding the nature of architecture, such as *Four Books on Architecture* on Palladian architecture, or *Yinzaο Fashi*¹ on ancient Chinese wooden architecture. The first part of the paper has been developed based on the treatises of classical architecture.

¹ Yinzaο Fashi (literally ‘Treatise on Architectural Methods or State Building Standards’), published in 1103AC, China, is a technical treatise dedicated to building and construction in Song Dynasty. It exerted a profound influence on wooden architecture in China.

Software	Workflow			
Cloud compare	1.Segmentation	Facade	Vault	Under Portico
				
	Points / Size	33462430 points / 989MB	54428728 points / 1.57G	52980385 points / 1.52G
	2.Subsample (1cm subsampled)			
		Points / Size	4176130 points / 123MB	1017622 points / 30MB
	Meshlab	3.Clean		
Points / Size			3902048 points / 59.5 MB	993701 points / 15.1MB
4.Merge and resample				
Points / Size	5786201 points / 88.2 MB			
Meshlab command line	5.Normal (normal neighbor=80)			
		Points / Size	7100747 points / 162 MB	



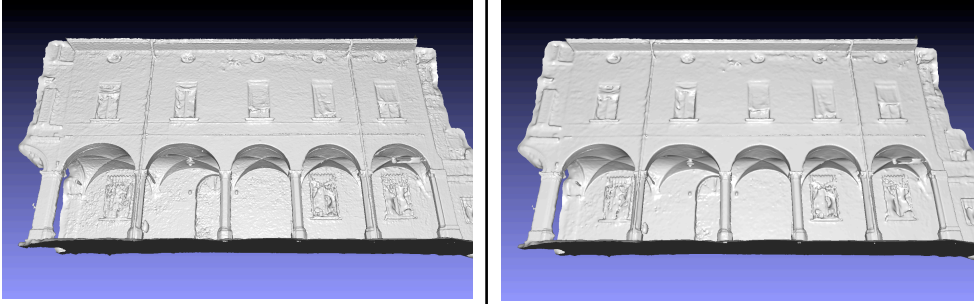

Software	Workflow		
Screened Poisson	6.Poisson (Depth=10)		
Points / Faces / Size		7100747 points / 2370283 faces / 68.6 MB	
	7.Trim (Trim=7)		
Points / Faces / Size		1162039 points / 2321716 faces / 47.4 MB	
	8. Smooth		
	9. Texture mapping		

Chart 2.6. A workflow from points cloud to texture mesh

Chapter 4 Case study

We compare the result of our optimized workflow with that of the original workflow in two case studies. The first case study, facade Palazzo of Albergati ('facade'), is a 55 meter-long facade built by brick and decorated with marble in Entablature and cornice. The second case study (portico) is a five-bay portico most of whose material is plaster. The two buildings were selected for their different architectural typology and material, both of which lead to different strategies in design of camera network, bundle adjustment, image filter, feature descriptor and dense reconstruction. 'Facade' is considered as a 2D planar surface where subtle distortion caused by different approaches reflect. In the 'portico' dataset, the main challenges come from the dramatically altered illumination and low texture surfaces, which caused difficulties in image orientation and model completeness in the original workflow.

4.1 Case study: Facade of Palazzo Albergati

Camera network.

In the original camera network, the axis of cameras was parallel to the facade with 2.1m step size. All the camera poses (figure 2.19) were perpendicular to the facade as suggested by several SfM guides (Photosynth Photography Guide, 2008; Agisoft PhotoScan Image Capture Tips, 2010; Recap Photo Getting Started Guide, 2013). The advantage of this camera network is that it maintains high similarities among adjacent images and consequently lead to high completeness, as well as keeps the image within a reasonable amount. From a photogrammetric point of view, however, the parallel/perpendicular camera network dose not afford good precision because of the low base line to depth ratio (B:D). Large base line to depth ratio, explicitly leading to convergent camera network, is supposed to improve the geometric accuracy (Remondino et al., 2006) and to be necessary for on site camera calibration (Barazzetti et al., 2011), but might cause difficulties in feature matching due to the image angle transition. Furthermore, applying multiple partial convergent camera network to long-linear space is not practical, for it highly increases the input images.



Figure 2.19 Original photo set with 32 images

Our optimized approach integrates on site calibration with the original parallel-axis-camera network to balance accuracy and efficiency. On site camera calibration extracts features from a distinctive object instead of coded target to speed up the workflow while maintains the similar accuracy. A camera network that fulfills the requirement of on site calibration should has:

- 1.Highly convergent network -large baseline/depth ratio
- 2.Orthogonal camera rotation - break the projective coupon between IO and EO parameter
- 3.Eight or more images

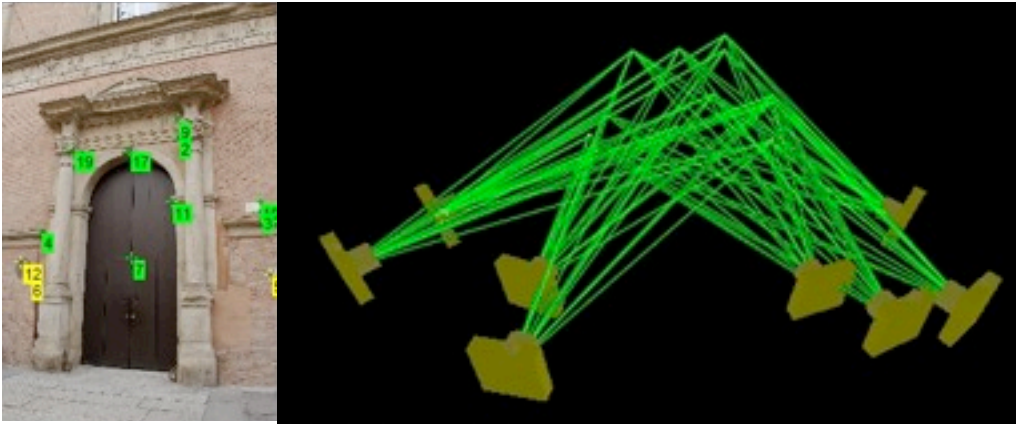


Figure 2.20 Targetless calibration

	Target calibration	On site calibration
focal length	18.6546	18.45mm
xp	-0.0405	0.046mm
yp	-0.0133	-0.084mm
k1	1.9007E- 004	2.076E- 004
k2	4.9026E- 007	6.993E- 007
k3	-2.3789E- 009	-2.288E- 009

Chart 2.7 Comparison of camera calibration via coded targets and targetless calibration

We chose one of the two doors of facade as the target of on site camera calibration for its distinctive features (figure 2.20). The obtained camera parameters are close to that of camera calibration from coded targets (chart 2.7). Besides the photo set for self-calibration (F0, #8) , we photographed two additional photo sets in parallel axis with similar step size (2.1m) but different depth and orientation angle: one portrait-oriented (F1, #31) and the other landscape oriented (F2, #20) (figure 2.21). It is supposed to provide better geometric depth information than single-parallel-axis camera network as original camera network. Evaluated by points cloud obtained by laser scan, the distortion is obvious when the full photo (#59) was employed (chart 2.8). A possible explanation is that F2, different from F1 which was photographed in panorama view, captures only facade of ground floor in the frame. The accuracy of 3D model is high in areas where F1 and F2 overlap, but low in the other areas. For this reason, we maintain only several images of F2 in crucial areas (doors) for global accuracy. The optimized camera network consists of 39 images.



Figure 2.21 Full photo set of 'facade'. Left: F0; Middle: F1; Right: F2.

photo set (# image)		camera network	points deviation
original (32)			
new	F1 (27)		
	F1+F2 (31)		
	F0+F1+F2 (39)		


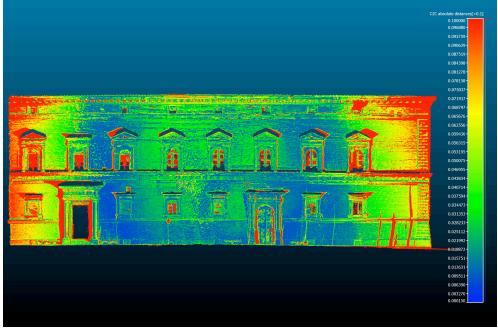
photo set (# image)		camera network	points deviation
	Full (59)		

Chart 2.8 Comparison of various camera network and points deviation from points cloud obtained by laser scan (Leica C10 laser scanner one shot, resolution=20mm). The decrease of gray level represents the reduction of image amount from the full photo set

Bundle adjustment

Most SfM systems are incremental in image orientation (Agarwal et al., 2009) as well as Visual SFM (Wu, 2013). Initializing from a well-matched image pair, 3D model is reconstructed and grows by repeatedly triangulating new cameras until no new cameras yield the minimum requirement of inlier matches. The incremental procedure employs bundle adjustment and outliers filter to reduce projection errors, but the misalignment still accumulates as the model grows. Therefore, initializing from a solid camera network would improve the accuracy. In the default workflow, the bundle adjustment is completed from a randomly selected pair; while in the optimized workflow we first built the sparse model from a solid base - the photo set F0 which is employed for self-calibration, and then oriented the other images by resuming incremental SfM. (figure 2.22). The test show a improvement of accuracy (chart 2.9).



Figure 2.22 Two-steps bundle adjustment. Left: Step 1. convergent network used for self-calibration; Right: Step: 2 Adding parallel-axis images

bundle adjustment	mean devi./ standard devi	precise deviation
one step	0.022 0.134	
two steps	0.015 0.092	

Chart 2.9 Comparison of bundle adjustment order. Range: 6cm

Image filter

In the default workflow, the images were photographed in JPEG format without any further processing. In the optimized workflow, however, we photographed RAW/NEF format for it contain full information enabling further processing, and also yields photogrammetric accuracy improvement over JPEG imagery (Stamatopoulos et al., 2012). We first tuned white balance and exposure based on the ColorChecker which we embed when photographing in every photo set. A photo set is a collection photos that were photographed with similar view angle and illumination condition - for instance, photos that are perpendicular to the vault or photos at 30 degree to the inner face of columns. A noise reduction filter is employed to discard the noise in the frame (figure 2.23). Then the color profile of the ColorChecker image is extracted, assigned in sRGB color space, and evaluated. If the error tolerance is acceptable (figure 2.24), the color profile is assigned to the other images of the same photo set and exported to JPG in sRGB color space.



Figure 2.23 Comparison of original image (left) and noise-filtered image (middle), and color-calibrated image (right)



Figure 2.24 Error report of color deviation shown in RGB distribution

Feature detection

Most SfM systems utilize sift (scale invariant feature transformation) for feature detection. The default sift of Visual SfM, siftGPU (Wu, 2007), is an implementation of SIFT (Lowe, 2004) for GPU. It pursues a balance between efficiency and accuracy by limiting the maximum number of feature. We tried various sift versions and parameters, among which a crucial parameter is peak. It is a image contrast threshold that closely relate to accuracy. A value of 6 appears regularly across the image pairs as the best choice (May et al., 2010). Our test of points deviation testify the improvement of accuracy from default siftGPU to VLfeat with peak of 6 (chart 2.10).

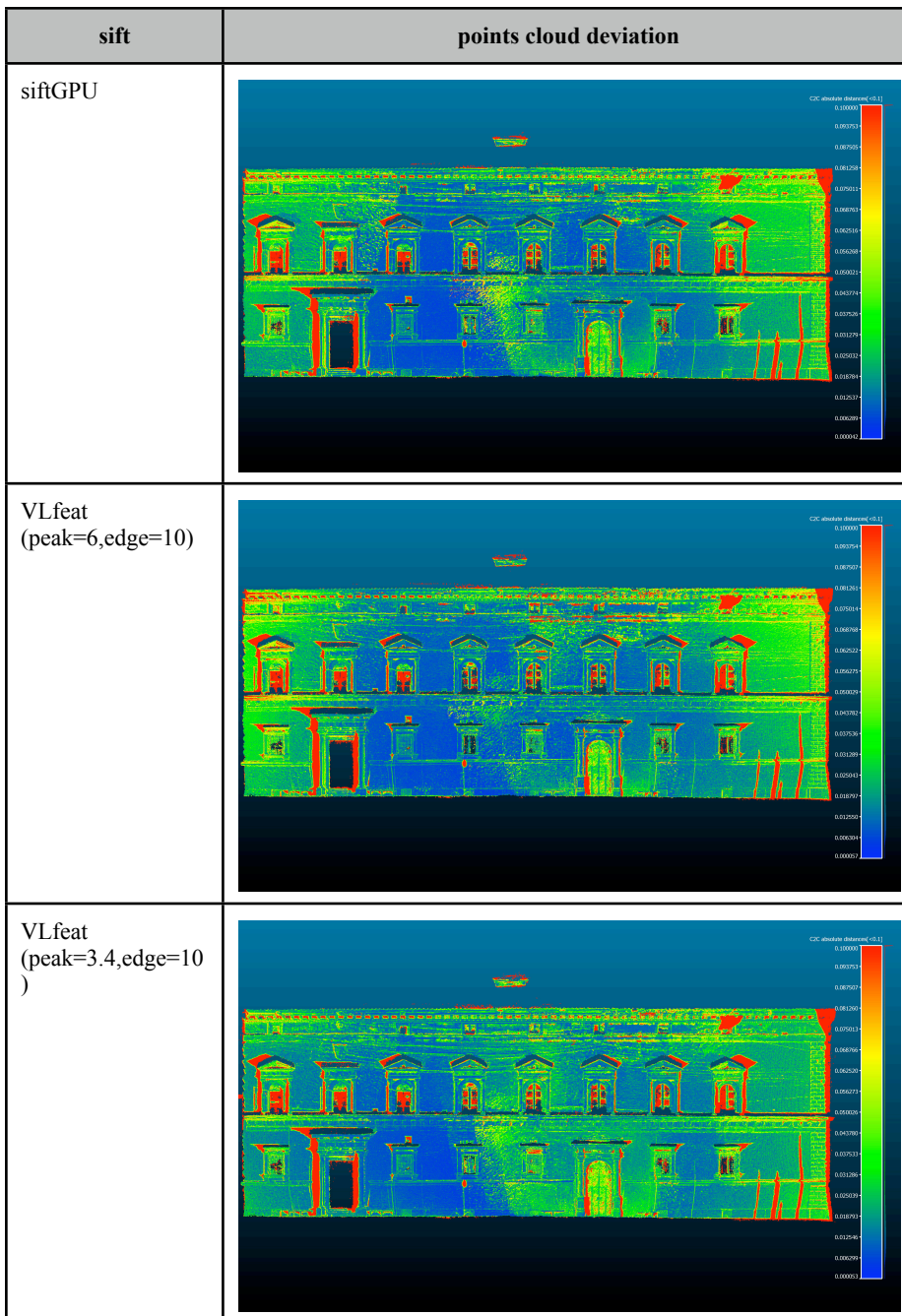


Chart 2.10 Sift comparison in terms of points deviation to laser scanned points cloud

Dense reconstruction

The dense reconstruction tools in original workflow are CMVS and PMVS2 (Furukawa et al., 2010). CMVS first decomposes the oriented images into clusters, and PMVS2 uses a patch-based algorithm to reconstruct dense 3D points for each cluster independently. Although a few parameters are allowed to be tuned before the modeling process, such as patching size, image subsampling ratio, and minimum cameras a point is observed, PMVS2 tends to automatically discard redundant images to avoid computational effort and clean the noise (Furukawa et al., 2010). When the structure of architecture is highly repetitive, such as facade of classical architecture, it is prone to discover redundant images improperly and cause holes on facade (chart 2.12). For this reason, we use SURE (Rothermel et al., 2012) in the optimized workflow. Different from PMVS2, SURE employs an approach of depth map, which is proved to be more robust and avoid the holes caused by redundant image removal. We could use filters such as RGB color selection to clean the noisy points caused by depth map.

Conclusion

In the case of facade of Palazzo Albergati, we prove that the geometric accuracy of SfM is comparable to laser scanner (around 5mm) via optimization of camera network, bundle adjustment, and feature detection (chart 2.11). Besides, we improve the model completeness by appropriate dense reconstruction package. Color calibration in image processing procedure ensures a true color texture mapping without interference of environment illumination.

	camera network	bundle adjustment	image filter	feature detection	dense reconstruction
original	parallel	one step	no	siftGPU	PMVS2
optimized	convergent+ closer parallel + further parallel	two steps	noise reduction +color calibration	VLfeat (peak=6, edge=10)	SURE

Chart 2.11 Workflow comparison of original workflow and optimized workflow

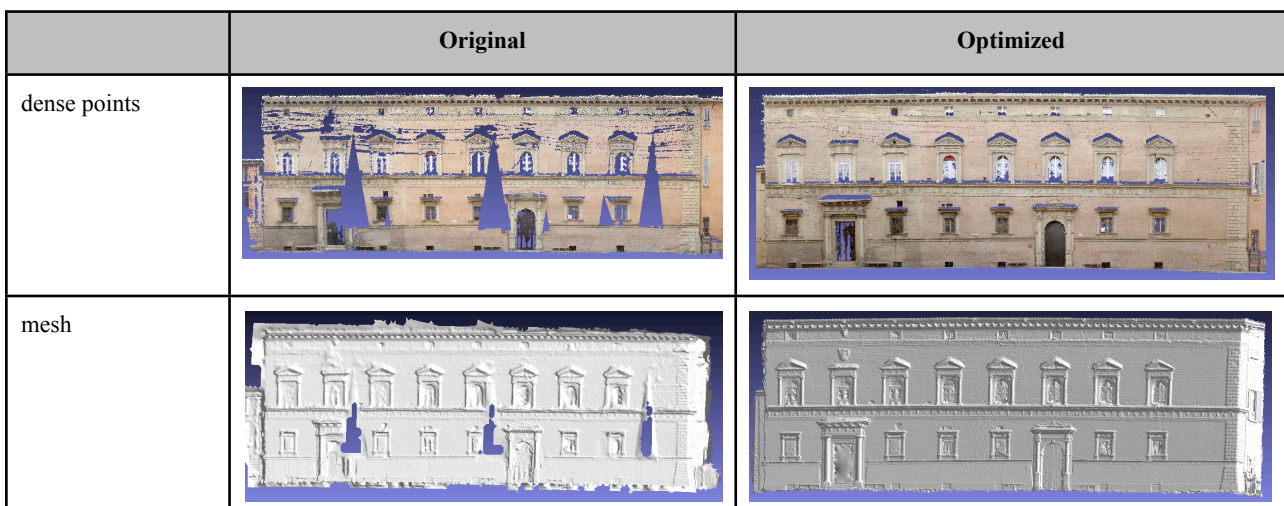


Chart 2.12 Comparison of original workflow and optimized workflow shown in points cloud and mesh. The holes on original model were caused by in redundant image removal in dense reconstruction.

3.6.2 Case study: Portico

Camera network

The camera network of portico is more challenging than that of facade, piazza or an isolated building. The porticos, owing to their various-exposed components, have dramatically altered illumination and limited view angle and view depth, as well as low textured areas (when vault built by plaster). According to our experience, the modeling tends to be failed at vault and inner surface of column even though surfaces and their transition were covered - as in the original camera network. Besides the rule of three, one probably has to take extra photos for such surfaces. Hence, it is necessary to design the camera network with respect to semantic components before survey.

1.1) Facade: The approach is similar with that of 'facade', including one convergent photo set for self-calibration and two parallel-axis photo set with different depth. (figure 2.25)

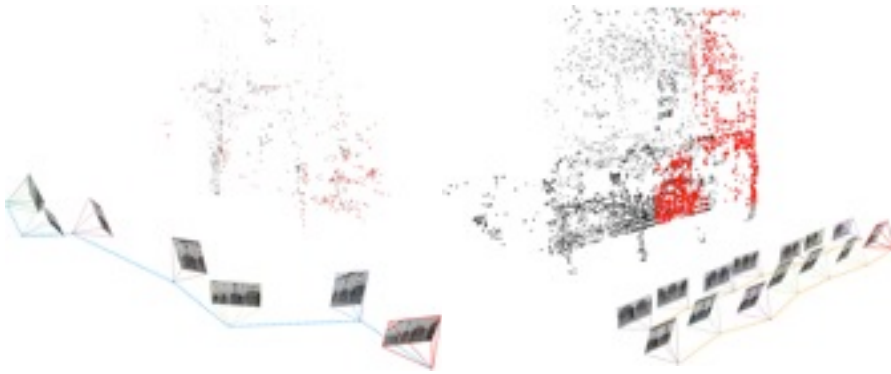


Figure 2.25 Camera network for facade

1.2) Column: Generally there are two types of column: cylinder and cube. We roughly consider them as four surfaces oriented :exterior, interior and two sides. The exterior surface is covered by photo set for facade. The cameras are supposed to revolve around the other three faces of each column with high overlap (figure 2.26). The exterior face is covered by the photo set for facade.



Figure 2.26 Camera network for facade

1.3) Wall: The exterior wall of ground floor inside portico is covered by the photo set for facade.

1.4) Vault: Sky-oriented cameras with small step size (40cm) is necessary when the vault is painted by plaster. In addition, convergent cameras from six corners of each bay are necessary to ensure the accuracy and improve the connection between adjacent vaults (figure 2.27).

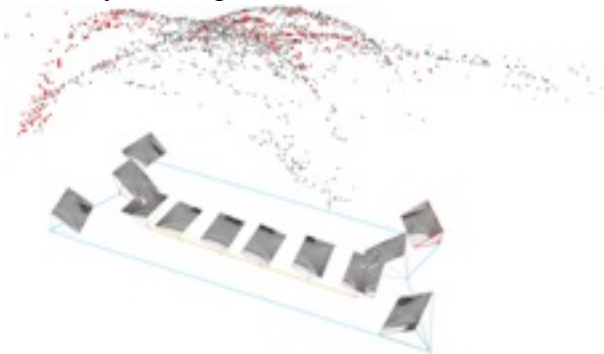


Figure 2.27 Camera network for vault

1.5) Pavement: When hand-held camera is employed, one might have to lift the camera as high as possible to increase the shooting angles. Sharp angles do not ensure model generation in dense reconstruction phase. Convergent network shooting from both exterior and interior based on each bay is suggested. (figure 2.28)

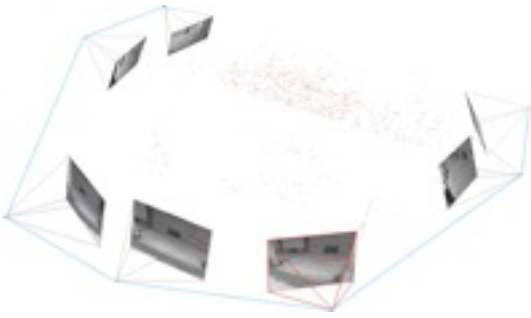


Figure 2.28 Camera network for pavement

Bundle adjustment

Similar with that in ‘facade’ photo set, stepwise bundle adjustment is employed to reduce potential error accumulation. Initializing from a convergent camera network consisting of eight images with camera rotation, cameras outside the portico and cameras inside the portico were oriented sequentially.

Image filter

Besides filters of noise reduction and color calibration which we applied to ‘facade’, we added wallis filter (Wallis, R., 1976) to ‘portico’ photo set to solve the problems of image orientation caused by low-texture areas and transition of illumination. The vault of portico, painted by plaster, is smooth and uniform that only a limited number of features could be extracted and consequently fail to be reconstructed. Possible solutions include zooming in focal lense, or reducing the ratio of low-texture areas in the frames. Furthermore, the illumination varies dramatically between interior and exterior. Appropriate exposure for exterior is probably too dark for interior, and vice versa. Using a wallis filter, the difference of image depth and exposure is flattened, and features of low-texture are extracted in a well distributed way (figure 2.29). Compared to photo set without wallis filter, the test show an obvious improvement in image orientation (chart 2.13).



Figure 2.29 Detected features on photo of vault. Left: colorful image. Right: Wallis-filtered image

	image orientation	3+ pts
color	185/212	22580
wallis	211/212	25938

Chart 2.13 Comparison of sparse reconstruction result between colorful images and wallis-filtered images

Feature detection

Similar with ‘facade’, we substitute default siftGPU to VLfeat (peak=6, edge=10). The test (chart 2.14) show that the feature matching and number of 3+ points are greatly improved, in spite of slightly less oriented images.

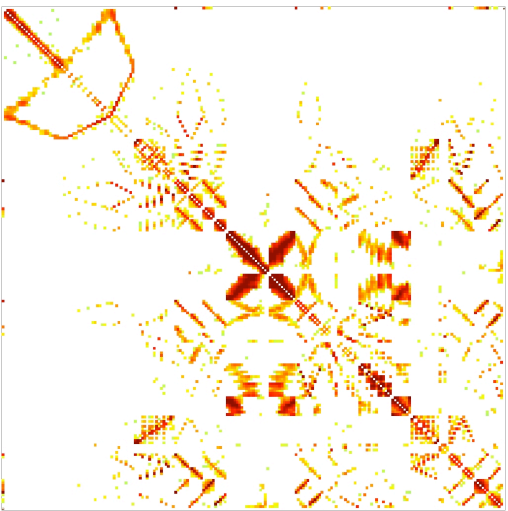
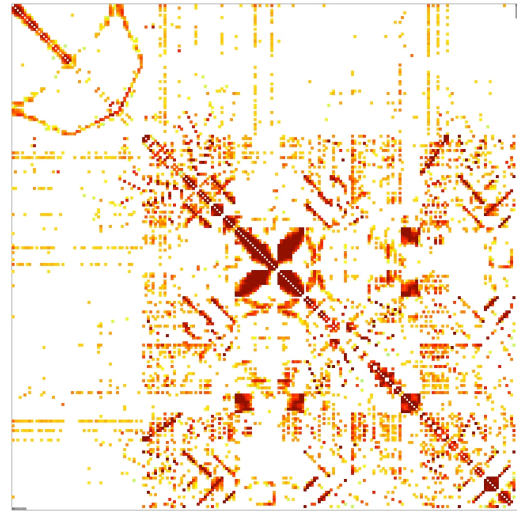
	default siftGPU	VLfeat (peak=6, edge=10)
oriented images	211/212	205/212
3+ pts	25938	57538
PBA	1.451	0.634
matching matrix		

Chart 2.14 Comparison of default siftGPU and VLFeat in sparse reconstruction

Dense reconstruction

We employed SURE instead of PMVS2 for dense reconstruction. Besides high completeness as mentioned in ‘facade’ photo set, SURE generates models via depth map from successfully orientated images. Instead of simply merging all the models, we could apply global alignment before merging them which is supposed to reduce the geometric error caused by misalignment during image orientation. Then we replace wallis images to images that have been color-calibrated, because the latter one tends to generates less noise in dense model.

Conclusion

From an architectural typology of view, portico is a challenging case for image-based modeling, for its complicated camera network and dramatical illumination transition. The optimized workflow in five steps (chart 2.15) demonstrates an obvious improvement compared to original model in terms of completeness (chart 2.16).

Approach	camera network	bundle adjustment	image filter	feature detection	dense reconstruction
original	rule of 3	one step	no	siftGPU	PMVS2
optimized	accuracy +minimal coverage	three steps	noise reduction + color calibration / wallis	VLfeat(peak=6, edge=10)	SURE

Chart 2.15 Comparison of different approaches between original workflow and optimized workflow of ‘portico’ photo set



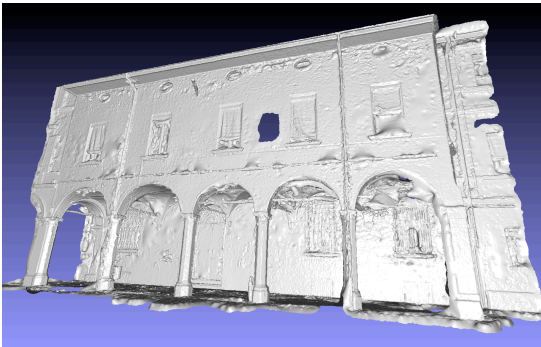
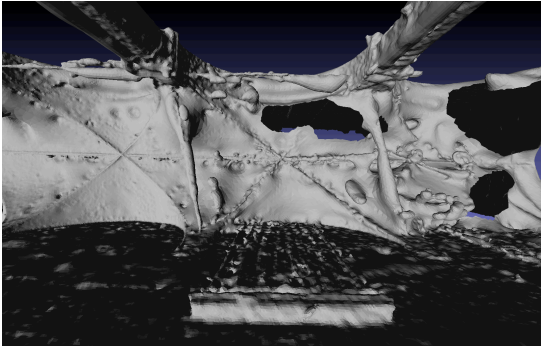
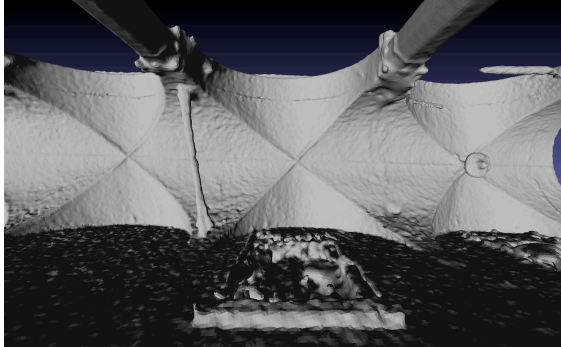
	Original	Optimized
dense points		
mesh		
mesh detail		

Chart 2.16 Comparison of original workflow and optimized workflow shown in points cloud and mesh

Reference:

1. Agarwal, S., Snavely, N., Simon, I., Seitz, S. M., & Szeliski, R. (2009, September). Building rome in a day. In *Computer Vision, 2009 IEEE 12th International Conference on* (pp. 72-79). IEEE.
2. Agisoft PhotoScan Image Capture Tips: Equipments and Shooting Scenarios, 2010. <http://downloads.agisoft.ru/pdf/Image%20Capture%20Tips%20-%20Equipment%20and%20Shooting%20Scenarios.pdf>
3. Alsadik, B., Remondino, F., Menna, F., Gerke, M., & Vosselman, G. (2013). Robust extraction of image correspondences exploiting the image scene geometry and approximate camera orientation. *Int. Arch. of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40, 5.
4. Amenta, N., Bern, M., & Kamvysselis, M. (1998, July). A new Voronoi-based surface reconstruction algorithm. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques* (pp. 415-421). ACM.
5. Amenta, N., Choi, S., & Kolluri, R. K. (2001). The power crust, unions of balls, and the medial axis transform. *Computational Geometry*, 19(2), 127-153.
6. Arc 3D Web Service. (2005). How to transform your images into 3D models.<http://homes.esat.kuleuven.be/~visit3d/webservice/v2/manual2/Webservice%20training.pdf>
7. Bajaj, Chandrajit L., Fausto Bernardini, and Guoliang Xu. "Automatic reconstruction of surfaces and scalar fields from 3D scans." *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. ACM, 1995.
8. Baltsavias, E. P. (1999). A comparison between photogrammetry and laser scanning. *ISPRS Journal of photogrammetry and Remote Sensing*, 54(2), 83-94.
9. Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3), 346-359.
10. Barazzetti, L., Mussio, L., Remondino, F., & Scaioni, M. (2011). Targetless camera calibration. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(5/W16).
11. Barazzetti, L. (2011). Automatic Tie Point Extraction from Markerless Image Blocks in Close-range Photogrammetry
12. Barazzetti, L., Scaioni, M., & Remondino, F. (2010). Orientation and 3D modelling from markerless terrestrial images: combining accuracy with automation. *The Photogrammetric Record*, 25(132), 356-381.
13. Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., & Taubin, G. (1999). The ball-pivoting algorithm for surface reconstruction. *Visualization and Computer Graphics, IEEE Transactions on*, 5(4), 349-359.
14. Boissonnat, Jean-Daniel. "Geometric structures for three-dimensional shape representation." *ACM Transactions on Graphics (TOG)* 3.4 (1984): 266-286.
15. Boehler, W., & Marbs, A. (2004, June). 3D scanning and photogrammetry for heritage recording: a comparison. In *Proceedings of the 12th International Conference on Geoinformatics* (pp. 291-298.)
16. Brutto, M. L., & Meli, P. (2012). Computer Vision Tools for 3D Modelling in Archaeology. *International Journal of Heritage in the Digital Era*, 1, 1-6.
17. Chandler, J., Fryer, J., Accuracy of Autodesk 123D Catch? <http://homepages.lboro.ac.uk/~cvjh/c/otherfiles/accuracy%20of%20123dcatch.htm>
18. Cignoni, P., Corsini, M., Dellepiane, M., Ranzuglia, G., Vergauwen, M., & Van Gool, L. (2008). MeshLab and Arc3D: photo-reconstruction and processing 3D meshes. *status: published*.

19. Crandall, D., Owens, A., Snavely, N., & Huttenlocher, D. (2011, June). Discrete-continuous optimization for large-scale structure from motion. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on* (pp. 3001-3008). IEEE.
20. Dante, A., La Modellazione di una Lapide Sepolcrale Attraverso Tecniche Low-Cost Range ed Image Based. http://3dom.fbk.eu/files/lc3d/Abate_lowcost3d-2012-Trento.pdf
21. Dellepiane, M., Calieri, M., & Dell'Unto, N. (2011). Monitoring archeological excavation using dense stereo matching techniques. tech. report CNR-ISTI, Pisa, Italy.
22. Deseilligny, M. P., & Clery, I. (2011, March). Apero, an open source bundle adjustment software for automatic calibration and orientation of set of images. In *Proceedings of the ISPRS Symposium, 3DARCH11*.
23. Deseilligny, M. P., Paparoditis, N. (2006). A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery. *Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(1/W41).
24. Doneus, M., Verhoeven, G., Fera, M., Briese, C., Kucera, M., & Neubauer, W. (2011). From deposit to point cloud: a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. In *XXIIIrd International CIPA Symposium* (Vol. 6, pp. 81-88).
25. Edelsbrunner, H., & Mücke, E. P. (1994). Three-dimensional alpha shapes. *ACM Transactions on Graphics (TOG)*, 13(1), 43-72.
26. El-Hakim, S. F., Beraldin, J. A., Picard, M., & Godin, G. (2004). Detailed 3D reconstruction of large-scale heritage sites with integrated techniques. *Computer Graphics and Applications, IEEE*, 24(3), 21-29.
27. Förstner, W. (2002). Computer Vision and Photogrammetry—Mutual Questions: Geometry, Statistics and Cognition. *Bildtechnik/Image Science, Swedish Society for Photogrammetry and Remote Sensing*, 151-164.
28. Fortune, S. (1992). Voronoi diagrams and Delaunay triangulations. *Computing in Euclidean geometry, 1*, 193-233.
29. Furukawa, Y., & Ponce, J. (2010). Accurate, dense, and robust multiview stereopsis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(8), 1362-1376.
30. Furukawa, Y., Curless, B., Seitz, S. M., & Szeliski, R. (2010, June). Towards internet-scale multi-view stereo. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on* (pp. 1434-1441). IEEE.
31. Fraser, C. S. (1997). Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(4), 149-159.
32. Fraser, C., Jazayeri, I., and Cronk, S., (2010). A feature-based matching strategy for automated 3D model reconstruction in multi close-range photogrammetry. In: 2010 ASPRS Conference
33. Fraser, C. S. (2013). Automatic Camera Calibration in Close Range Photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 79(4), 381-388.
34. Furukawa, Y., & Ponce, J. (2009). Accurate camera calibration from multi-view stereo and bundle adjustment. *International Journal of Computer Vision*, 84(3), 257-268.
35. Garland, M., & Heckbert, P. S. (1997, August). Surface simplification using quadric error metrics. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (pp. 209-216). ACM Press/Addison-Wesley Publishing Co..
36. Goesele, M., Snavely, N., Curless, B., Hoppe, H., & Seitz, S. M. (2007, October). Multi-view stereo for community photo collections. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on* (pp. 1-8). IEEE.

37. Goesele, M., Curless, B., & Seitz, S. M. (2006). Multi-view stereo revisited. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on* (Vol. 2, pp. 2402-2409). IEEE.
38. Hartley, R. I., & Mundy, J. L. (1993, September). Relationship between photogrammetry and computer vision. In *Optical Engineering and Photonics in Aerospace Sensing* (pp. 92-105). International Society for Optics and Photonics.
39. Herrmann, L. R. (1976). Laplacian-isoparametric grid generation scheme. *Journal of the Engineering Mechanics Division*, 102(5), 749-907.
40. Hirschmuller, H. (2008). Stereo processing by semiglobal matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2), 328-341.
41. Jancosek, M., & Pajdla, T. (2011, June). Multi-view reconstruction preserving weakly-supported surfaces. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on* (pp. 3121-3128). IEEE.
42. Jazayeri, I., Fraser, C.S. (2009) Interest Operators for Feature-Based Matching in Close-Range Photogrammetry. *Photogrammetric Record*, 25(129): 24–41 (March 2010)
43. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.
44. Kazhdan, M., Bolitho, M., & Hoppe, H. (2006, June). Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*.
45. Kolluri, R., Shewchuk, J. R., & O'Brien, J. F. (2004, July). Spectral surface reconstruction from noisy point clouds. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing* (pp. 11-21). ACM.
46. Kolmogorov, V., & Zabih, R. (2002). Multi-camera scene reconstruction via graph cuts. In *Computer Vision—ECCV 2002* (pp. 82-96). Springer Berlin Heidelberg.
47. Manfredini, A. M., Galassi, M. (2013). Assessments for 3D Reconstruction of Cultural Heritage Using Digital Technologies, 3D-ARCH-3D Virtual Reconstruction and Visualization of Complex Architectures, Trento, Italy.
48. Matthews, N. A. (2008). Aerial and close-range photogrammetric technology: Providing resource documentation, interpretation, and preservation. *Technical Note, 428 Bureau of Land Management, Denver, Colorado*.
49. May, C., Turner, M.J. and Morris, T., (2010). Scale Invariant Feature Transform: A Graphical Parameter Analysis. In: *Proceedings of the BMVC 2010 UK postgraduate workshop*. 5.1-5.11
50. Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(10), 1615-1630.
51. Morel, J. M., & Yu, G. (2009). ASIFT: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2(2), 438-469.
52. Nguyen, H. M., Wünsche, B., Delmas, P., & Lutteroth, C. (2012, June). 3D models from the black box: investigating the current state of image-based modeling. In *Proceedings of the 20th International Conference on Computer Graphics, Visualisation and Computer Vision (WSCG 2012), Pilsen, Czech Republic, June*.
53. Olague, G., & Mohr, R. (2002). Optimal camera placement for accurate reconstruction. *Pattern Recognition*, 35(4), 927-944.
54. Photosynth photography guide, 2008. <http://cdn1.ps1.photosynth.net/docs/Photosynth%20Guide%20v8.pdf>
55. Pietroni, N., Tarini, M., & Cignoni, P. (2010). Almost isometric mesh parameterization through abstract domains. *Visualization and Computer Graphics, IEEE Transactions on*, 16(4), 621-635.

56. Pons, J. P., Keriven, R., & Faugeras, O. (2005, June). Modelling dynamic scenes by registering multi-view image sequences. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (Vol. 2, pp. 822-827). IEEE.
57. Pollefeys, M., Koch, R., & Van Gool, L. (1999). Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *International Journal of Computer Vision*, 32(1), 7-25.
58. Rasztovits, S., Dorninger, P., (2013). COMPARISON OF 3D RECONSTRUCTION SERVICES AND TERRESTRIAL LASER SCANNING FOR CULTURAL HERITAGE DOCUMENTATION. IN: XXIV International CIPA Symposium, Strasbourg, France
59. Recap getting started guide, 2013. https://adsk-recap-public.s3.amazonaws.com/Getting_Started_Guide_ReCap_Photo.pdf
60. Remondino, F., Del Pizzo, S., Kersten, T., Troisi, S., 2012: Low-cost and open-source solutions for automated image orientation – A critical overview. Proc. EuroMed 2012 Conference, M. Ioannides et al. (Eds.), LNCS 7616, pp. 40-54. Springer, Heidelberg
61. Remondino, F. (2003, February). From point cloud to surface: the modeling and visualization problem. In *International Workshop on Visualization and Animation of Reality-based 3D Models* (Vol. 34, p. 5).
62. Remondino, F. (2006). *Detectors and descriptors for photogrammetric applications. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36 (3), 49-5
63. Remondino, F., & El-Hakim, S. (2006). Image-based 3D Modelling: A Review. *The Photogrammetric Record*, 21(115), 269-291.
64. Remondino, F., & Fraser, C. (2006). Digital camera calibration methods: considerations and comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(5), 266-272.
65. Remondino, F. (2011). Heritage recording and 3D modeling with photogrammetry and 3D scanning. *Remote Sensing*, 3(6), 1104-1138.
66. Remondino, F., Spera, M. G., Nocerino, E., Menna, F., Nex, F., & Gonizzi-Barsanti, S. Dense image matching: comparisons and analyses.
67. Remondino, F., Spera, M.G., Nocerino, E., Menna, F., Nex, F. and Gonizzi Barsanti, S. (2013). Dense image matching: comparisons and analyses. 2013 Digital Heritage International Congress (DigitalHeritage), Vol. 1, pp. 47-54
68. Remondino, F.; El-Hakim, Sabry; Gruen, A.; Li Zhang, "Turning images into 3-D models," *Signal Processing Magazine, IEEE* , vol.25, n.4, pp. 55-65, July 2008
69. Rothermel, M., Wenzel, K., Fritsch, D., & Haala, N. (2012, December). SURE: Photogrammetric Surface Reconstruction from Imagery. In *Proceedings LC3D Workshop, Berlin*.
70. Roy, S., & Cox, I. J. (1998, January). A maximum-flow formulation of the n-camera stereo correspondence problem. In *Computer Vision, 1998. Sixth International Conference on* (pp. 492-499). IEEE.
71. Saint-Marc, P., Chen, J. S., & Medioni, G. (1989, June). Adaptive smoothing: A general tool for early vision. In *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR'89., IEEE Computer Society Conference on* (pp. 618-624). IEEE.
72. Santagati, C., & Inzerillo, L. (2013). 123D Catch: efficiency, accuracy, constraints and limitations in architectural heritage field. *International Journal of Heritage in the Digital Era*, 2 (2), 263-290.

73. Santagati, C., Inzerillo, L., & Di Paola, F. (2013). Image-based modeling techniques for architectural heritage 3d digitalization: limits and potentialities. *The International Archives of Photogrammetry, Remote Sensing and Spatial and Information Sciences*, 40(5), 550-560.
74. Schenk, T. (2005). Introduction to photogrammetry. *The Ohio State University, Columbus*.
75. Seitz, S. M., Curless, B., Diebel, J., Scharstein, D., & Szeliski, R. (2006, June). A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer vision and pattern recognition, 2006 IEEE Computer Society Conference on* (Vol. 1, pp. 519-528). IEEE.
76. Snavely, N., Seitz, S. M., & Szeliski, R. (2006, July). Photo tourism: exploring photo collections in 3D. In *ACM transactions on graphics (TOG)* (Vol. 25, No. 3, pp. 835-846). ACM.
77. Snavely, N., Seitz, S. M., & Szeliski, R. (2008). Modeling the world from internet photo collections. *International Journal of Computer Vision*, 80(2), 189-210.
78. Snavely, N., Seitz, S. M., & Szeliski, R. (2008, June). Skeletal graphs for efficient structure from motion. In *CVPR* (Vol. 1, p. 2).
79. Stamatopoulos, C., Fraser, C. S., & Cronk, S. (2012) ACCURACY ASPECTS OF UTILIZING RAW IMAGERY IN PHOTOGRAMMETRIC MEASUREMENT. In: XXII Congress of the International Society for Photogrammetry, Remote Sensing
80. Strecha, C., Bronstein, A. M., Bronstein, M. M., & Fua, P. (2012). LDAHash: Improved matching with smaller descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(1), 66-78.
81. Strecha, C., Fransens, R., & Van Gool, L. (2004, June). Wide-baseline stereo from multiple views: a probabilistic account. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on* (Vol. 1, pp. I-552). IEEE.
82. Strecha, C., Fransens, R., & Van Gool, L. (2006). Combined depth and outlier estimation in multi-view stereo. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on* (Vol. 2, pp. 2394-2401). IEEE.
83. Strecha, C., von Hansen, W., Van Gool, L., Fua, P., & Thoennessen, U. (2008, June). On benchmarking camera calibration and multi-view stereo for high resolution imagery. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on* (pp. 1-8). IEEE.
84. Tingdahl, D., & Van Gool, L. (2011). A public system for image based 3D model generation. In *Computer Vision/Computer Graphics Collaboration Techniques* (pp. 262-273). Springer Berlin Heidelberg.
85. Vedaldi, A., & Fulkerson, B. (2010, October). VLFeat: An open and portable library of computer vision algorithms. In *Proceedings of the international conference on Multimedia* (pp. 1469-1472). ACM.
86. Vergauwen, M., & Van Gool, L. (2006). Web-based 3d reconstruction service. *Machine vision and applications*, 17(6), 411-426.
87. Vogiatzis, G., Torr, P. H., & Cipolla, R. (2005, June). Multi-view stereo via volumetric graph-cuts. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (Vol. 2, pp. 391-398). IEEE.
88. Wallis, R. (1976, November). An approach to the space variant restoration and enhancement of images. In *Proc. of symp. on current mathematical problems in image science*, naval postgraduate school, Monterey CA, USA, November.
89. Wenzel, K., Rothermel, M., Fritsch, D., Halaa, N. (2013). Image acquisition and model selection for multi-view stereo. *3D-ARCH-3D Virtual Reconstruction and Visualization of Complex Architectures*, Trento, Italy.
90. Wu, C. (2013) Towards Linear-time Incremental Structure from Motion. In *International conference on 3DV conference*, 2013 IEEE 127-134

91. Wu, C. (2007). "SiftGPU: A GPU implementation of Scale Invariant Feature Transform (SIFT)". <http://cs.unc.edu/~ccwu/siftgpu/>;
92. Wu, C. (2011). "VisualSFM: A Visual Structure from Motion System", <http://ccwu.me/vsfm/>
93. Wu, C., Agarwal, S., Curless, B., & Seitz, S. M. (2011, June). Multicore bundle adjustment. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on* (pp. 3057-3064). IEEE.

Part Three. Integrating ideal model with as-build model

In the last chapter, we enrich the semantic models in BIM platform (Revit) based on the foundation of the previous chapters: knowledge-based model and reality-based model. In the first chapter, knowledge-based modeling, we transform the treatises of classical architecture into parametric relationships of semantic components in BIM. The parametric modeling facilitates great flexibility to 3D modeling by predefined parametric families. The transitions from Doric column after Palladio to that of Vignola, for example, is only a matter of parameters tweaking instead of time-consuming 3D drawing from the very beginning. However, discrepancies exist between the theory and the practice for several reasons. That is why reality-based modeling is required. Reality-based modeling captures as-built information consisting of all discrepancies to the theoretical models. The limitations are less-segmented information and only the superficial features are documented. Hence, a semantic-based model that enriched with multiple information is demanded to meet various needs such as historical research, architectural renovation or virtual reality. In this chapter, we will propose how to segment the as-built models and fill them into the semantic framework to cater different levels of detail, accuracy and flexibility.

Chapter 1. Semantic organization of BIM

1.1 Three types of family

As an object-oriented BIM platform, Revit offers three kinds of family to virtually construct a project. These three kinds of family, system family, loadable family, and in-place family, cover a wide range of levels of intelligence and possibility of customization. The level of accuracy of 3D model, owing to the flexibility of the family.

System families are highly-predefined and embedded in projects. They are highly intelligent families, but hardly to be customized. System families, such as wall, floor and ceiling, have the least recognizable features among different styles or eras. Ignoring the difference of color, a piece of wall in a Palladian villa may look similar with that in a Chinese palace. Most of their distinctions are quantity parameters (thickness, length, height) and quality parameters (material, color), which could be managed by values. The behaviors of system are constrained. It is not possible to draw a vertically inclined wall from system family, for its behaviors are constrained. System families are supposed to be defined numerically rather than geometrically.

Loadable families are most commonly used in Revit model. They are supposed to be created from templates (.rft), or modified from existing families (.rfa) from library components of Revit. They are the only family edited and saved in external environment of a project. Once founded, they could be reused in different projects. Hence loadable families are usually parametric. The inner lines of a window's frame are attached to the opening's edges which are constrained by height and width. The frame would adjust itself to the modification when height and width are changed. In a template file, components' behavior and attributes are set according to semantic category (column or beam) and construction category (architectural column or structural column). Loadable family provides a more flexible modeling environment than system family through operations like extrusion, revolve, and sweep.

In-place family facilitates to customize components to the largest flexibility. Created in a project, the difference of category is bridged, for the form could be assigned to any category. It is in in-place family that components could have 'illogical' behavior, but also intelligence loss. An inclined wall,

for instance, could be built as in-place family by sweeping along a 3D curve. In-place family could not be saved, nor reused in other projects.

Kinds of family	Intelligence	Customization	Call method	Category	Environment
System family	High	Low	Embedded	Wall, ceiling	project
Loadable family	middle	middle	Template	Column, profile	family
In-place family	low	High	In-place edit	All category	project

Chart 3.1 Properties of three types of family

1.2 Logic of host

The logic of host family in Revit represents the hierarchy of a component and its sub-component, and the connected relationship of different components.

A family could be a host when it contains another loadable family, nested family. The geometric relations and behaviors of the host family and nested family could be constrained by associating parameters. For example, a capital (nested family) could be modeled and loaded into a column (host family) with associated parameter such as radius. Once a new value is assigned to the column, the nested family will adjust to the modification automatically. Besides sub-components, profile could also be defined as nested family. When generating a form by sweeping along a path in the host family, the profile could be specified from the nested family instead of in-place drawing. For architectural heritage containing multiple moldings, nested profile is an effective way to transform reality-based information into semantic model in BIM.

A family is allowed to pick a reference plane (A) to be A-hosted in Revit, such as furnitures are floor-hosted and doors are wall-hosted. The hosted relations, owing to reasons such as gravity, function and architectural grammar, are quite common in architecture. Taking cornice which is swept along a path of wall, pediment, or window for example. We can regard the cornice as hosted by the latter components. Once a host is picked, the cornice could be only located attached to it even in further modification. Revit enables users to create families that are line-hosted, face hosted, and less ambiguously ceiling hosted, floor hosted.

Chapter 2. Towards as-built model in BIM

2.1 State of art

When a textured mesh is rendered, image-based modeling is completed while begins the process towards as-built BIM. BIM is considered to be highly potential in architectural heritage documentation for parametric modeling and data enrichment. The parametric modeling is close to the geometric nature of classical architecture which is built on the foundation of modular relation and repetitive pattern. Treatises of architecture is an ideal source of knowledge-based modeling (Apollonio et al., 2012) which outputs ideal model. However, the ideal model is as-built model (Gainai, 2000).

Building Information modeling (BIM) serving as knowledge resource of multiple information is considered to be high potential in documentation of architectural heritage. Researches on as-built BIM model emerge in recent years thanks to the progress in digital survey and BIM. How to link the reverser workflow has arisen wide interests. The workflow consists of data acquisition, segmentation, and as-built modeling (Hichri et al., 2013). The first stage, as illustrated in part two, mainly employs range-based modeling and image-based modeling - two approaches usually integrated - to acquire reality-based model with high accuracy and automation. Segmentation of survey data is a time consuming work. Although robust methods have been developed (Tang et al., 2010; Xiong et al., 2013), but manual intervention is still mandatory in most cases. The largest challenge comes from the last procedure, how to reconstruct as-built model based on the survey data in BIM platforms.

Various practice of BIM based on points cloud have been reported recently. (Giudice et al., 2013) explored the pipeline from points cloud to BIM in thermal power of a building for restoration use. The points cloud are served as as-built information to which the archives are compared. (Bianco et al., 2013) presented how to compromise the points cloud with the rigid Revit model in a parallel way to cater different purposes. (Nicolas et al., 2013) compares the deviation of points cloud with the average flat plane in a room size for renovation purpose. However, the practice above do not address the central issue of as-built BIM in cultural heritage-how to achieve high accuracy.

In spite of its great ability in multiple-data documentation, BIM platforms are developed with barely any attentions to facilitate to existing architecture, not to mention architectural heritage. Although the points cloud could be imported into modeling environment by various plug-in (Imaginit's *ScanToBIM*) for visualization. They could hardly benefits to the accurate reality-based modeling. The points are not recognized nor captured, but defined as frozen scaffolding allowing no modification. (Dore et al., 2014) presented an approach that combines procedural modeling with detailed parametric modeling to build components library, and map the components onto the survey data. The parametric modeling is realized in GDL of ArchiCad. (Garagnani, 2013) presented an approach to transform points cloud into reference points in Revit. The plug in (Green Spider), developed in Revit API, allows users to import the points or trace a spline interpolation among imported points into mass family. High accuracy could be achieved by lofting the spline decimated from the cross section of survey data.

2.2 A semantic-based integrated modeling approach

An integrated modeling approach is proposed in this chapter aiming at balancing intelligence with accuracy, as well as keeping manual modeling within a reasonable degree by application of parametric modeling.

We first categorize the semantic components of classical architecture into categories that are employed by BIM today. Favorably almost all components of find their counterparts in today’s AEC industry, although we have to classify cornice to beam which has structural function instead of merely decoration.

Then we consider the factors that tend to cause negleable discrepancies between primitive models and as-built models. Since it is impossible to strictly build the as-built model in BIM softwares, it is not unwise to define a general rule. The first factor is the organic form of classical architecture: capital of order, moulding and relief. In these cases, the variation not only exist between knowledge-based model and as-built model, but also vary from instance to instance. Another factor comes from components which are difficult to measure and to construct, such as valut. Vault is by no means primitives not only considering the errors of measurement, construction, and restoration, but also because it is a compromise of irregular disposition of building, unflatten site and joints of problematically constructed column and wall. Consequently capital of column, molding, relief and vault are suggested to be generated from the points cloud of survey.

In contrast to elements above, there are also components which relatively close to primitives such as wall, floor and flat ceiling. In general cases, the subtle inclined degree of a piece of wall is of minor importance. Such components could be created as system family to retain the maximum possibility of smart behavior as host family to be picked by window, door or cornice.

Column, window and door are usually the most distinctive elements of a building. The features of such elements are to much extent profile-based. The geometry of the frame of window and door is quite straightforward. The variation mainly relies on the profile which is swept out along the frame. The profiles could first be extracted from the points cloud of survey, and then nested to the loadable families of such element. On the one hand the host family retains parametric relationship, on the other the profiles provide as-built information.

Components		System family	Loadable family	In-place family	Survey-based
Wall		•			
Floor		•			
Ceiling	Flat	•			
	Vault			•	•
Column	Shaft		•		
	Capital		•		•
Beam/ Cornice	Path				
	Profile		•		•
Window	Path				

Components		System family	Loadable family	In-place family	Survey-based
	Profile		•		•
Door	Path				
	Profile		•		•

Chart 3.2 Category of semantic components with respect to family types and host logic

Chapter 3. Case study

3.3.1 Case study 1. The structural framing of entrance hall

Villa Badoer, located in Rovigo, Italy, was designed by Andrea Palladio in 1556 and built between 1557 and 1563. Being one of the remains of Palladian villa, it is part of UNESCO World Heritage Site ‘City of Vicenza and the Palladian Villas of the Veneto’. The digital model of the villa was reconstructed using Visual SfM plus PMVS from 326 photos photographed with Nikon D3100. The SfM techniques demonstrate its advantages in reconstructing complex objects. It proves that both interior and exterior of a building could be reconstructed in one model as long as the transition of position is well-covered by images (figure 3.1). As the difficulties mainly come from the transition of illumination when linking interior with exterior, it would be much straightforward to expand the model to the rest rooms on the ground floor. This feature of SfM makes it more efficient than range-based modeling which scans the building from multiple stations (at least one station per room) and registers all the scanned data into one model.



Figure 3.1. Camera positions and dense points of Villa Badoer obtained by SfM techniques. The entrance hall and the exterior walls of the villa are reconstructed in the same model without further registration.

The points cloud of the entrance hall gives access to as-built visualization but contains no semantic information. Consequently a BIM model is created based on semantic components with reference to the survey data (figure 3.2). The ideal model is originally created by considering the beams under the ceiling as regularly-distributed and have the same dimension. Mapping the beams onto the points cloud, however, it turns to be the beam system is irregular in the way the beams are distributed and the dimension of each beam. Generally two types of beam exist with two dimensions of 235mm*286mm and 184mm*286mm (chart 3.3). The as-built BIM model of the beam system is created based on these two types of beam and located according to the survey data. The structural deformation of the beam system could be visualized by registering the survey data with the BIM model and computing their distance (figure 3.3).

A	B	C	D	E	F
Category	Family	Family and Type	Material:Area	Material:Volume	Type
Structural Frame	Timber	Timber:235*286	6.46m ²	0.42m ³	235*286
Structural Frame	Timber	Timber:235*286	6.46m ²	0.42m ³	235*286
Structural Frame	Timber	Timber:235*286	6.46m ²	0.42m ³	235*286

A	B	C	D	E	F
Structural Frame	Timber	Timber:235*286	6.46m ²	0.42m ³	235*286
Structural Frame	Timber	Timber:235*286	6.46m ²	0.42m ³	235*286
Structural Frame	Timber	Timber:235*286	6.46m ²	0.42m ³	235*286
Structural Frame	Timber	Timber:235*286	6.46m ²	0.42m ³	235*286
Structural Frame	Timber	Timber:235*286	6.46m ²	0.42m ³	235*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286
Structural Frame	Timber	Timber:184*286	5.83m ²	0.33m ³	184*286

Chart 3.3. Material list of structural frame



Figure 3.2. Ideal model and as-built model of entrance hall integrated in BIM platform

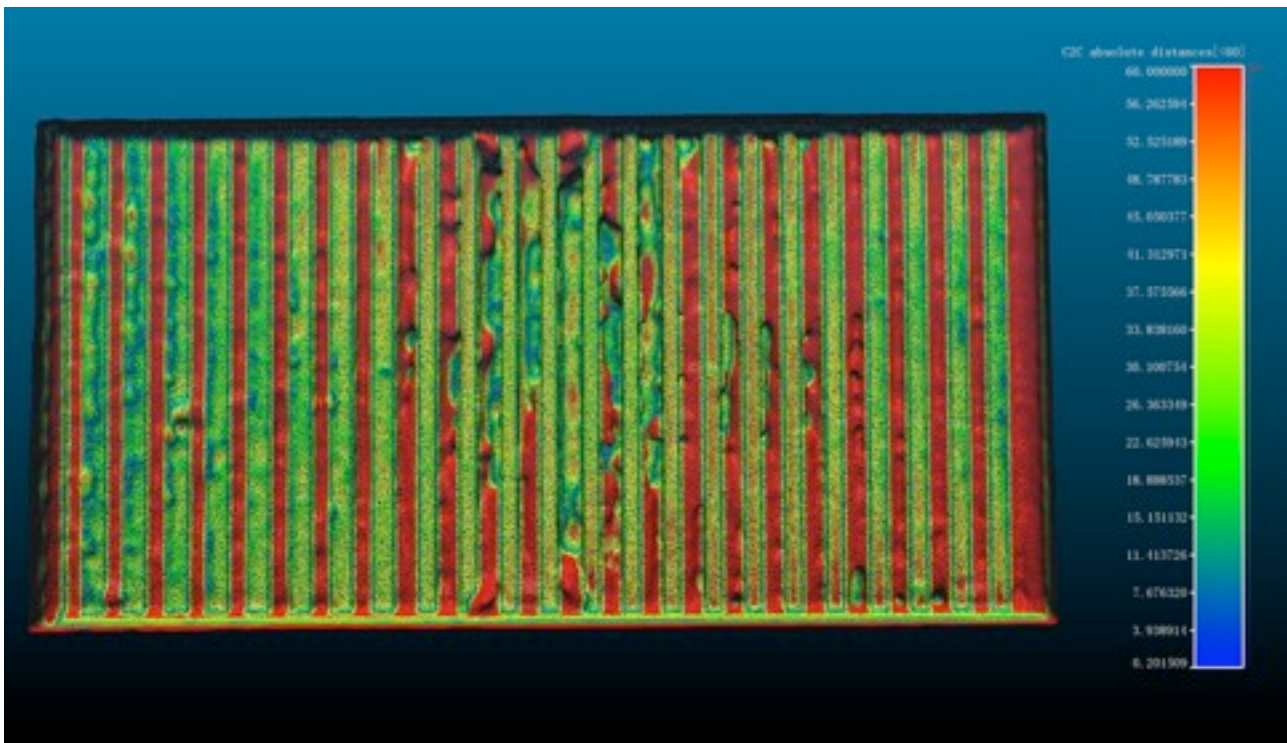


Figure 3.3. The distance computation between ideal model and as-built model. (Unit:m). Change of color distribution that indicates the structural deformation of beam is observed along the shorter axis.

Case study 2. The vault of portico

Portico is a typical architectural typology in Bologna, which is known for its 42 km-long portico. The image-based modeling of Portici Bolognese was launched to cater its application of World Heritage of UNESCO. The vaults of portico are the most difficult parts for documentation and modeling, not only for the low-texture surfaces, but also for the irregular forms. Range-based modeling could yield accurate results, but it is not practical considering the 42 km-long portico and traffic in the city center as well as pedestrians. The SfM approaches was selected for its high flexibility and low-cost, and it proved to be satisfactory in terms of modeling completeness and efficiency.

The five-bay portico located at Via Saragozza was digitally reconstructed with 212 images, a considerable amount of which were oriented towards the vault (figure 3.4). Dense points are obtained by running SURE (figure 3.5). The vaults are sliced into 10 cross sections each of which is decimated with a distance of 20cm between adjacent two points (figure 3.6). The decimated points are imported into Revit (figure 3.7) and constituted splines (figure 3.8). Finally the vault is built by lofting the splines (figure 3.9). Once the vault gets its form, the mass family is imported into Revit project environment and is defined as ceiling through in-place modeling.

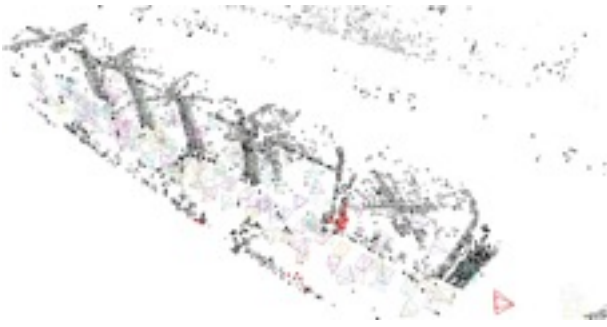


Figure 3.4. Sparse points with camera poses recovered by Visual SfM



Figure 3.5. Dense points acquired in SURE from the output of visual SfM

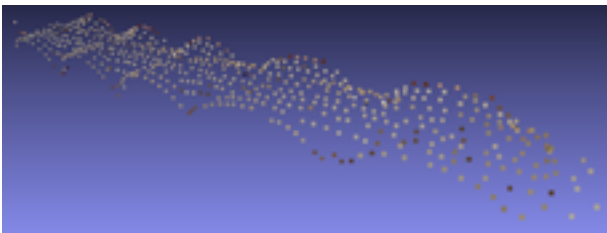


Figure 3.6 Points decimated with a distance of 20cm in adjacent two points

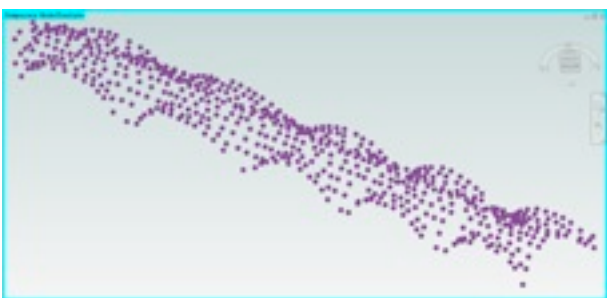


Figure 3.7. Points imported into Revit and recognized as reference points

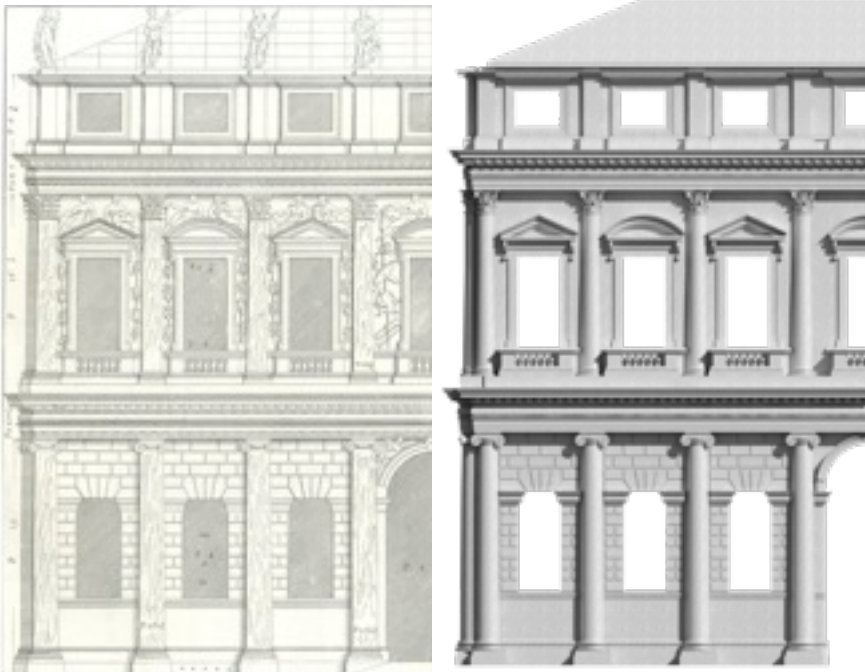


Figure 3.8. Curves generated by interpolating the reference points



Figure 3.9. Vault generated by lofting the sliced curves. It is classified as ceiling by IFC (IfcCovering)

3.3.3 Case study 3. Facade of Palazzo Barbaran



Left: Figure 3.10 Drawing of Palazzo Barbaran from *The Four Books on Architecture*.
 Right: Figure 3.11 Ideal model based on the left drawing built in Revit 2014.

Palazzo Barbaran is one of the remains of Palladio's Palazzo that collected in the third part of his treatise-*The Four Books on Architecture* (figure 3.10). Although the building was expanded with two extra bays on facade after Palladio's death, the existing part was completed under his full supervision. We reconstruct 3D model of this facade in two approaches - knowledge-based

modeling and reality-based modeling - with the aim to find the discrepancies between the ideal model and the as-built model and integrated both of them into BIM database.

The knowledge-based approach (figure 3.11) yields ideal model as it is strictly based on Palladio's grammar presented in his book. Palladio set up universal rules of architectural components such orders, and explained how the rules could be applied to practice with flexibility. As in the case of facade of Palazzo Barbaran, the columns, windows and moulding are locally created in Revit's Family environment based on Palladio's universal rules, and then globally composed according to the dimensions engraved on the drawing of the facade. Besides geometric information, thanks to the

In the reality-based approach, the points cloud of as-built model is reconstructed from 172 images by SfM techniques (figure 3.12) (Wu, 2011; Furukawa et al., 2010). The points cloud is further processed to mesh surface by Screened Poisson (Kazhdan et al., 2006), and scaled into right scale according to a known distance.



Figure 3.12. As-built model of facade of Palazzo Barbaran. The points cloud is yielded from 172 images by SfM techniques.

In order to compare the discrepancies between the ideal model and the as-built model, points cloud are extracted from the polygonal surfaces of the ideal model, and registered to the points cloud from survey with further refinement to ensure the accuracy. For each point in the ideal model, the closest point is found from the as-built model and the distance in-between is shown by red-blue distribution shown in figure . The large red areas are caused by the lack of survey data in corresponding parts due to the limited image capture angles from street level. However, red areas are observed from the Entablature of Ionic Order on the ground floor and Corinthian Order on the first floor where are covered in the points cloud of survey.

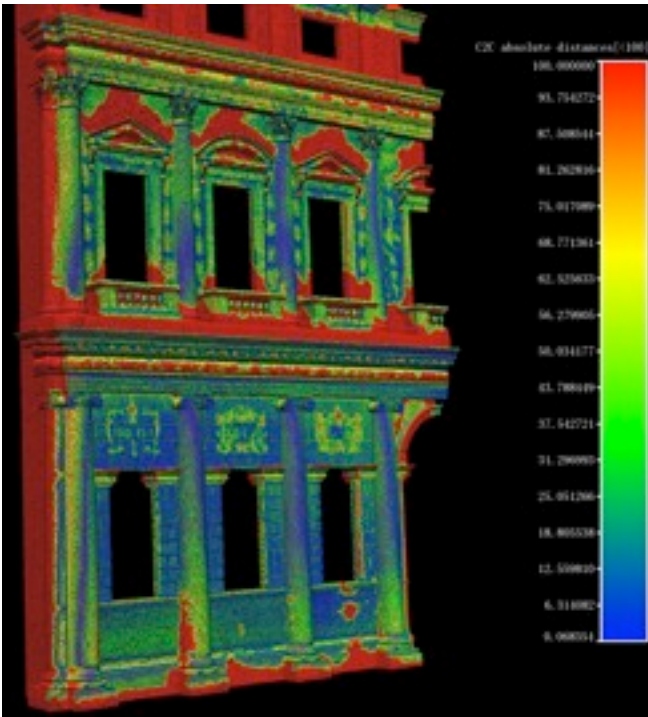
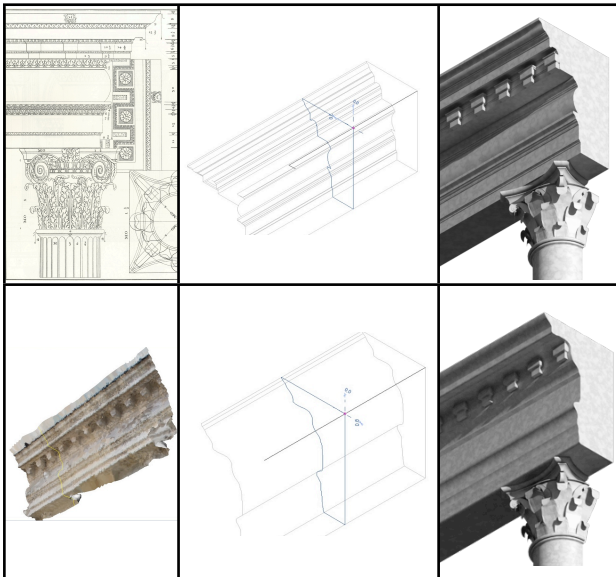


Figure 3.13. The difference of ideal model and as-built model of facade of Palazzo Barbaran revealed in red-blue map. The difference increases from blue to red. (Unit: mm). The red distribution is caused by two possibilities: lack of corresponding points from survey; discrepancies between what Palladio's drawing and real construction.

The Entablature of the two orders in ideal models are created identically to Palladio's universal rules (semantic composition of orders) and their application in practice (drawing of the facade). However, the as-built models are distinctive to the ideal models in the way moldings are composed. The curves in Entablature of both Ionic Order and Corinthian Orders are not presented in the construction.



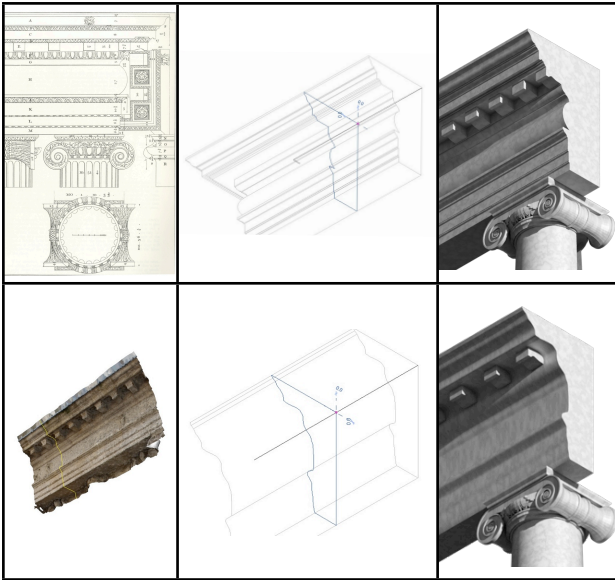


Chart 3.4

Top: Entablature of Corinthian Order based on knowledge and reality.

Bottom: Entablature of Ionic Order based on knowledge and reality.

In BIM platform, both the ideal models and the as-built models are retained in family environment (beam model) and project environment (facade model). The models of Entablature are created in beam family by nesting two different profiles - one from treatise and one from survey- respectively. Both the ideal model and the as-built model of certain component could exist in the same position (space) and construction phase (time) of global model by defining them as different design options (chart 3.4). For instance, the Entablature of the facade is defined as design options that contains two possibilities: the ideal model and the as-built model. When one solution is activated, the other one would be filtered. The approach could be applied to the components where discrepancies exist such as the Entablature.

Reference

1. Arayici, Y. (2008). Towards building information modelling for existing structures. *Structural Survey*, 26(3), 210-222.
2. Arayici, Y. (2007). An approach for real world data modelling with the 3D terrestrial laser scanner for built environment. *Automation in Construction*, 16(6), 816-829.
3. Attene, M., Katz, S., Mortara, M., Patané, G., Spagnuolo, M., & Tal, A. (2006, June). Mesh segmentation-a comparative study. In *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on* (pp. 7-7). IEEE.
4. Apollonio, F.I., Gaiani, M., & Zheng, S. (2012). Bim-based Modeling and Data Enrichment of Classical Architectural Buildings. *SCientific RESearch and Information Technology*, 2(2), 41-62.
5. Barazzetti, L., Remondino, F., & Scaioni, M. (2010). Automation in 3D reconstruction: results on different kinds of close-range blocks. In *ISPRS Commission V Symposium Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Newcastle upon Tyne, UK* (Vol. 38, No. 5).
6. Benhabiles, H., Vandeborre, J-P., Lavoué, G., Daoudi, M. (2009). A framework for the objective evaluation of segmentation algorithms using a ground-truth of human segmented 3D-models, *Proc. IEEE International Conference on Shape Modeling and Applications 2009*, IEEE Computer Society, Washington DC, USA, 8 pages.
7. Bianco, I., Del Giudice, M., and Zerbinatti, M. (2013). A DATABASE FOR THE ARCHITECTURAL HERITAGE RECOVERY BETWEEN ITALY AND SWITZERLAND. *XXIV International CIPA Symposium, Strasbourg, France*
8. Chen, X., Golovinskiy, A., Funkhouser, T. (2009). A Benchmark for 3D Mesh Segmentation, *Proc. ACM Transactions on Graphics*, 28 (3), 12 pages
9. Cheng, X. J., & Jin, W. (2006, October). Study on reverse engineering of historical architecture based on 3D laser scanner. In *Journal of Physics: Conference Series* (Vol. 48, No. 1, p. 843). IOP Publishing.
10. Del Giudice, M. and Osello, A.: BIM FOR CULTURAL HERITAGE, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-5/W2, 225-229
11. Dore, C., Murphy M. (2013)Semi-Automatic Techniques for As-Built BIM Façade Modeling of Historic Buildings. In: *Proceedings of the 2013 Digital Heritage International Congress*. Marseille, France.
12. Dorninger, P., and C. Nothegger. (2007). 3D segmentation of unstructured point clouds for building modelling. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol 35, pp 191–196.
13. Gaiani, M., Balzani, M., & Uccelli, F. (2000, September). Reshaping the Coliseum in Rome: an integrated data capture and modeling method at heritage sites. In *Computer Graphics Forum* (Vol. 19, No. 3, pp. 369-378). Blackwell Publishers Ltd.
14. Gaiani, M. (1999). Translating the Architecture of the Real into the Virtual. *Proceedings of the Heritage Applications of 3D Digital Imaging*, 1-16.
15. Garagnani, S. (2013). Building Information Modeling and real world: A methodological approach to accurate semantic documentation for the built environment. In: *Proceedings of the 2013 Digital Heritage International Congress*. Marseille, France.
16. Garagnani, S., Manfredini, A.M. (2013). PARAMETRIC ACCURACY: BUILDING INFORMATION MODELING PROCESS APPLIED TO THE CULTURAL HERITAGE PRESERVATION. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Trento, Italy, Volume XL-5/W1*

17. Giudice, M. D., Osello, A. (2013). BIM for Cultural Heritage. In: XXIV International CIPA Symposium, Strasbourg, France
18. Hichri, N., Stefani, C., De Luca, L., Veron, P. (2013). Review of the <As-Built BIM> Approaches. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-5/W1, Italy, Trento
19. Manferdini, A.M., Remondino, F., Baldissini, S., Gaiani, M., Benedetti, B., 3D Modeling and Semantic Classification of Archaeological Finds for Management and Visualization in 3D Archaeological Databases, Proc. 14th VSMM, pp. 221–228, 2008
20. Manferdini, A. M., & Remondino, F. (2010). Reality-based 3D modeling, segmentation and web-based visualization. In Digital Heritage (pp. 110-124). Springer Berlin Heidelberg.
21. Manferdini, A. M., & Remondino, F. (2012). A review of reality-based 3D model generation, segmentation and web-based visualization methods. International Journal of Heritage in the Digital Era, 1(1), 103-124.
22. Murphy, M., McGovern, E., & Pavia, S. (2009). Historic building information modelling (HBIM). Structural Survey, 27(4), 311-327. ITcon Vol. 19, pg. 20-46
23. Murphy, M., McGovern, E., & Pavia, S. (2011, May). Historic Building Information Modelling- Adding Intelligence to Laser and Image Based Surveys. In *Proceedings of the 4th ISPRS International Workshop* (Vol. 38, p. 5).
24. Nicolas, G., Landrieu, J., Nugraha, Y., and Pèrè, C. (2013) HYBRID REPRESENTATION OF DIGITAL MOCKUP FOR HERITAGE BUILDINGS MANAGEMENT. In: XXIV International CIPA Symposium, Strasbourg, France
25. Ning, X., Zhang, X., Wang, Y., and Jaeger, M. (2010). Segmentation of architecture shape information from 3D point cloud.
26. Oreni, D., Brumana, R., Georgopoulos, A., and Cuca, B. (2013). HBIM FOR CONSERVATION AND MANAGEMENT OF BUILT HERITAGE: TOWARDS A LIBRARY OF VAULTS AND WOODEN BEAN FLOORS, XXIV International CIPA Symposium, Strasbourg, France
27. Overend, M. A Tool that Combines Building Information Modeling and Knowledge Based Engineering to Assess Façade Manufacturability.
28. Pomaska, G. (2009, October). Utilization of photosynth point clouds for 3D object reconstruction. In Proceedings of the 22nd CIPA symposium, Kyoto, Japan.
29. Remondino, F., & Rizzi, A. (2010). Reality-based 3D documentation of natural and cultural heritage sites—techniques, problems, and examples. Applied Geomatics, 2(3), 85-100.
30. Remondino, F. (2011). Heritage recording and 3D modeling with photogrammetry and 3D scanning. Remote Sensing, 3(6), 1104-1138.
31. Rothermel, M., Wenzel, K., Fritsch, D., & Haala, N. (2012, December). SURE: Photogrammetric Surface Reconstruction from Imagery. In Proceedings LC3D Workshop, Berlin.
32. Scan to BIM software, <http://imaginit.com/software-solutions/building-architecture/scan-to-bim>
33. Tang P., Huber D., Akinici B., Lipmand R. and Lytle A. (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques', automation in construction, vol. 19, no. 7, pp. 829-43.
34. Trevisan, C., (2002). Proporzioni e vera forma di particolari architettonici rilevati con scanner 3D: caratteristiche di un software specifico. Disegnare idee immagini : rivista semestrale del Dipartimento di rappresentazione e rilievo , pp.44-49, n.24
35. Vergauwen, M., & Van Gool, L. (2006). Web-based 3d reconstruction service. Machine vision and applications, 17(6), 411-426.
36. Xiong X., Adan A., Akinici B., Huber D. (2013). Automatic creation of semantically rich 3D building models from laser scanner data, Automation in Construction, Vol. 31, pp 325-337.

37. Zhana, Q., Liangb, Y., and Xiaoa, Y. (2009). Color-based segmentation of point clouds . IAPRS Vol XXXVIII, Part 3/W8.

Summary: Strength, limitation and future development

The paper focuses on how to virtually reconstruct architectural heritage based on emerging digital techniques and medias. Two modeling approaches are discussed in the research. The strength, limitation and future work of knowledge-based modeling and image-based modeling will be presented in the following content.

In the first part (knowledge-based modeling), 3D models are created according to architectural pattern books. In contrast to CAD-based modeling approach, the BIM-based modeling builds also the parametric relations extracted from shape grammar. It allows modification of dimension and adjustment in real time and maintains potential for further data enrichment. The object-oriented modeling approach of BIM is inherently close to the nature of classical architecture and provide prior knowledge for segmenting survey data.

Given the advantage of associative relation and multi data documentation, the drawbacks of BIM softwares are that organic forms and irregular disposition are what it is supposed to cater, while both of them are common cases of classical architecture. Developed for production-based modeling, BIM provides limited flexibility in handling hand-made shape. Revit provides only four generative operation in loadable family (extrusion, sweep revolve, swept blend). It is sufficient for modeling geometry composed by combination of moulding such as Entablature, but is much time-consuming to create capital of Corinthian order. ArchiCAD supplies GDL code writing to achieve flexibility besides semantic components which are loaded from its library, but the coding is a challenging work for architects. In addition, BIM softwares enable few possibilities for irregular disposition of pre-defined semantic components. For example, inclined wall is not permitted in Revit by default unless created as in-place family with the loss of intelligent behavior and semantic metadata. It is possible to create inclined wall with user-defined angles in ArchiCAD while no more further irregular transformation is allowed. It is difficult to contains high level of accuracy from as-built state and maintain semantic behavior at the same time.

Future development of BIM software in modeling historical architecture may come from the development of components library and user customization via API. Both Revit and ArchiCAD provide library of components by default, but only a small portion of the library are historical components. On the other hand, it is possible to download BIM models in certain searching engines, but few of them contain parametric relations. Therefore, it is necessary to complete the models library of architectural heritage from orders to other semantic categories and from classical architecture to Gothic architecture or Chinese architecture. The tolerance of irregular disposition of geometry is less technically challenging from a developer's point of view, but no evidence show update in such field will be released in next versions Revit and ArchiCAD. Therefore, user-involved customization should be a possible development such as Revit API. A large series of plug-in exist to solve problems in design such as visibility or user interface, so it is reasonable to expect such progress in the filed of architectural heritage.

Image-based modeling generated reality-based modeling from digital images. It is more flexible, economical and photorealistic than range-based modeling which has been widely used in the field of architecture. Our workflow integrates the full automation of SfM developed by computer vision to metric accuracy concerned by photogrammetry. The experiments show our optimized workflow leads to results comparable to that from laser scanner. Taking into consideration of its wide application of architectural typology from facade to portico, ability to merging interior and exterior and large-scale reconstruction even at city level, the image-based modeling is prospecting in survey and modeling of architectural heritage.

One of the limitations of image-based modeling is that it is dependent on the illumination condition. Highlights and shadows casted on surfaces cause chaos in feature correspondence and color authenticity. Unlike range-based modeling which is able to work during night, a photographer may not work in majority time of a day. Furthermore, the on-site image acquisition is still a tedious work in large scenes which requires considerable skill and experience, although full automation has been realized from image input to textured mesh model. In our experiment, a five-bay portico requires about 200 images which cost at least 1 hour by a skilled operator. The consumed time can not be neglected when Bologna's 42km portico is the target of digital reconstruction. Video cameras that are fixed on vehicles are expected to speed up the image acquisition, but the resolution of exported frames is not sufficient to extract features from low-textured surfaces such as plasterwork. Semantically segmenting the point cloud or mesh is one of the most challenging works. In spite of the algorithms embedded in commercial software such as RapidForm, manual intervention is still mandatory in most cases. Future developments are supposed to be originated from the issues above.

Integrating survey data with parametric models in BIM is an open topic that needs to be addressed. Revit by default allows point clouds to be imported and adds visualization options of point clouds since the last version, but the point cloud is not regarded as points created in Revit but just frozen mass. Meshes in .txt and .xyz format could be imported into ArchiCAD for generating terrain, but other semantic categories can not utilize the survey data. (Garagnani, 2013) developed a plug-in via Revit API that allows forms to be generated from imported point clouds in the mass family of Revit. Although it is constrained to sub-components modeling due to the computing resources of point clouds in Revit, it is an approach towards as-built BIM models. It is not reasonable to expect the full reverse-engineering chain to be completed in a software that aims at design and construction, so the interoperability among different softwares and plug-in development may be the tendency in the future.