POLITECNICO DI MILANO

Ph.D Course in
Environmental and Infrastructure Engineering
XXVI Cycle (2011-2013)

Automatic segmentation, classification and extraction of repeated patterns for building façades modelling

Tutor:	Prof. Luigi Mussio (Politecnico di Milano, DICA)
Supervisor:	Prof. Raffaella Brumana (Politecnico di Milano, ABC)
Co-supervisor:	Prof. Marco Scaioni (Tongji University)
Coordinator:	Prof. Alberto Guadagnini (Politecnico di Milano, DICA)

Ph.D Thesis by
Mattia Previtali

Milano, March 2014

Abstract

Availability of realistic 3D building models is important in many applications. In urban planning the chance to explore 3D virtual reality world is much more efficient than the analysis of 2D maps. For public security, accurate 3D building models are indispensable to plan intervention strategies during emergencies. Also, virtual tourism may largely benefit from highly realistic city models.

The rapid development of terrestrial laser scanning (TLS) devices allowed the acquisition of point clouds of urban environments in a relatively reduced time. With the considerable high point density and the explicit 3D coordinates of such point clouds, it is possible to recover both large structures and fine details of buildings. In addition, the increase of automation in the acquisition and registration stage of laser scans, in conjunction with a reduction of the cost of instruments, extended the chance to use terrestrial laser scanners to a larger number of operators. This resulted in a growing attention in generating as-built building models on the basis of TLS point clouds. In many cases such models are generated by manual modelling of each building element. However, this is undoubtedly a rather time consuming and expensive procedure that limits the widespread dissemination of building models. Automation in the reconstruction is essential to speed up the processing, and thus the time needed to deliver the final model can be shortened in comparison with manual techniques. Indeed, only if a high degree of automation is maintained throughout the whole pipeline, from data acquisition to modelling, the economic sustainability of the building model generation can be guaranteed. However, lack of automated approaches to understand the building structures captured in raw data is still underlined by different operators in the Architecture, Engineering, and Construction (AEC) domain.

This thesis introduces a new procedure aimed at the automated production of building façades models. This method is principally designed to generate as-built models of urban construction for thermal retrofitting. Indeed, energy efficient retrofitting of existing buildings, mainly the ones built in the period 1950 - 1975, is a key aspect for reaching the proposed energy consumption reduction targets fixed by national and international authorities. In this field, highly detailed as-built models of buildings are

needed on one side for the thermal assessment, and on the other for producing executive drawings. The developed approach can be considered as a multi-step process in which the building model is iteratively estimated and refined starting from the raw point cloud.

The presented methodology first accomplishes the segmentation of the point cloud of a building façade into its planar elements. Then, starting from the identified planar clusters, façade breaklines are automatically extracted to be used later to generate a 3D vector model. During this step some priors on the urban structures like the prevalence of straight lines and orthogonal intersections are exploited to set additional constraints. Despite considerable effort, data obtained with range scanners may suffers from occlusions. However, building façades exhibit a high degree of self-similarity and redundancy. For this reason an algorithm was developed for the reconstruction of incomplete models with the help of high-level architectural objects and the identification of repeated patterns in urban façades. The final product is a semantically enriched 3D model of the building façade that can be integrated into a Building Information Model (BIM).

To demonstrate the reliability, precision and robustness of the method, several tests on different kinds of datasets are illustrated and discussed. The tests presented here revealed that while reconstruction efficiency is improved by the developed approach, the geometric accuracy of derived models is also comparable to the one achievable by the standard modelling process.

The final building models have several applications. In the last part of this dissertation, the integration of building models with thermal images is addressed in detail. In recent years this task has become quite popular, but actual approaches may not be able to provide accurate and rigorous results. A solution is proposed based on mapping of thermal data on the vector building model. The alignment of both data is obtained with a combined photogrammetric bundle adjustment including both thermal and RGB images.

Finally, some of the algorithms developed for façade modelling are extended, and partially modified, to cope with other applications. In particular, the developed segmentation strategy is tested for scan registration of urban scenes which present the prevalence of some few basic geometric shapes. In such cases, the identification of the same planar features between scans is exploited to determine their registration parameters. A final extension of the developed algorithms concerns modelling of indoor environments. Indeed, automatic reconstruction of buildings' exteriors share many problems with the issues associated to indoor reconstruction, but in the last case the scene may feature a higher degree of clutter and occlusion. For this reason a tailored solution is presented and validated.

Contents

Abstract i

Contents iii

Introduction 1

Motivation of the research 4

Overview of the dissertation 5

Structure of the dissertation 7

PART A 9

Chapter 1
Façade segmentation 10

1.1. Introduction and definitions 11

1.2. State of the art 13

1.2.1. Feature clustering 13

1.2.2. Region-growing 14

1.2.3. Model fitting 16

1.2.4. Hybrid techniques 19

1.3. Critical analysis of existing techniques 19

1.4. Façade segmentation 22

1.4.1. Point cloud pre-processing 24

1.4.2. RANSAC model fitting 25

1.4.2. Under- and over- segmentation removal 29

Chapter 2
Geometric reconstruction 34

2.1. State of the art 35

2.2. Automated façade modelling 39

2.2.1. Extraction of geometric edges 40

2.2.2. Edge smoothing 44

2.2.3. Reconstruction of façade topology and break-line extraction 48

2.2.4. Geometric model export in CAD environment 49

Chapter 3
Semantic model enrichment 52

3.1. State of the art 53

3.2. The hierarchical classification approach 55

3.2.1. Façade knowledge 55

3.2.2. Façade classification rules 58

3.2.3. Semantic object classification 61

3.2.4. Enriched model and export file format 63

Chapter 4
Façade regularity detection 71

4.1. State of the art 73

4.2. Repeated patterns detection 76

4.2.1. Element clustering 78

4.2.2. Estimation of façade repeated patterns 81

PART B 89

Chapter 5
Experimental tests on real façades 90

5.1. Performance evaluation framework 90

5.1.1. Façade segmentation 91

5.1.2. Geometric reconstruction 93

5.1.3. Object classification 93

5.1.4. Regularity detection 93

5.2. Experimental tests 94

5.2.1. ‘D’Oggiono building’ 94

5.2.2. ‘Courtyard no.1 and no.2’ 101

5.2.3. ‘Building no. 34’ 107

5.2.4. ‘Nave building’ 111

5.2.5. Conclusions 116

Chapter 6
Other applications 118

6.1. Façade model integration with IRT images 119

6.1.1. IRT image integration overview 120

6.1.2. Thermal camera intrinsic calibration 121

6.1.3. IRT image orientation 123
6.1.4. Applications 124
6.2. Scan registration using planar features 129
6.2.1. Plane parameterization and scan registration 132
6.2.2. Applications 135
6.3. Indoor reconstruction 139
6.3.1. State of the art 139
6.3.2. Indoor reconstruction method 140
6.3.3. Room surface detection 142
6.3.4. Reconstruction of openings under occlusions and clutter 146
6.3.5. Applications 148
Conclusions and further developments 152
Future investigations 155
Bibliography 157
Acknowledgements 172

Introduction

In many countries all over the word the reduction of energy consumption and the increase of efficiency in exploiting natural energy sources are playing an important role in national and international policies (USDE 2001, CEC 2002, EPA 2002, USDE 2003, CEC 2007). For example, the European Union (EU) fixed ambitious climate and energy targets for 2020 (CEC 2008, Böhringer et al. 2009), and also other countries are following the same road map (Giddens 2009). Indeed, climate change mitigation and sustainable practices are currently at the top of political and technical agendas towards the goal of low carbon cities all over the word. European targets, known as the '20-20-20', are:

- 20% reduction in EU greenhouse gas emissions with respect to 1990 levels;
- raising the share of energy produced from renewable resources to 20% in the countries of the Union; and
- 20% improvement in the EU's energy efficiency.

To meet these goals a significant role is played by *energy efficient buildings*. A building can be defined as 'energy efficient' if it is designed to provide a significant reduction of the energy needed for heating and cooling. In particular, energetic qualification of buildings based on energy consumption (as described in directives such as 2010/31/CE for the EU) becomes a standard to evaluate the thermal efficiency.

Even if a great attention is paid to the topic of 'Zero Energy Buildings' (ZEBs), also the retrofitting of existing ones plays a fundamental role to reach the EU's target. For this reason the development of methodologies for a fast evaluation of thermal efficiency of such buildings and technologies for their retrofit are growing in their interest. In particular, buildings dated between 1950 and 1975 were constructed in an era when little or no consciousness was on taking care of energy efficiency performance (Fig. 1). They are furthermore those that, according to recent reports from the Architect Council of Europe, are requiring envelope refurbishment to guarantee their future exploitation as residential or office buildings (ACE, 2012). Unlike historical buildings for which the refurbishment has to respect precise

restriction and must be controlled by specific authorities, regular civil buildings are subjected to a limited number of constraints (for example, the original appearance of the façade has in general to be preserved). Furthermore civil buildings constructed during 1950 - 1975 fit a similar architectural style which is based on a combination of simple geometrical shapes. In particular, these façade share can be generally modelled as planar objects or at least can be approximately divided into piecewise planar parts.

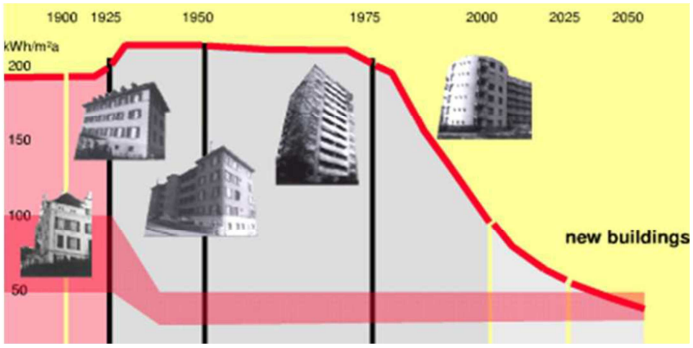


Fig. 1 Energy consumption (expressed in kWh/m^{2}a) in buildings since 1900. Buildings built In 1950 – 1975 show the highest energy consumption.

This research work was partially supported by the European research project EASEE (Envelope Approach to improve Sustainability and Energy efficiency in Existing multi-storey multi-owner residential buildings). This project was developed within 7th Framework Programme for Research, to tackle the initiative Energy efficiency in Buildings Public Private Partnership (EeB PPP). In particular, the EASEE project aims at developing a new holistic approach to energy efficient envelope retrofitting of multi-storey and multi-owner buildings through a combination of modular pre-fabricated components, novel insulation approaches and scaffolding-free installation. The aim of this research inside the EASEE project is to develop innovative methodologies for building survey, modelling and assessment (Fig. 2).

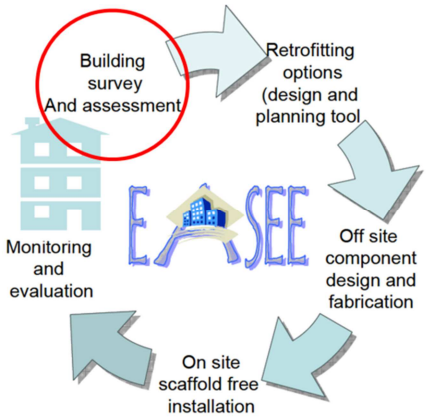


Fig. 2 EASEE approach to envelop retrofitting. Highlighted in red the contribution of the presented research in the project.

Unfortunately, up-to date and complete digital models are not always available, especially for constructions during 1950 - 1975. Executive drawing may have been made on analogue supports. Moreover such products are mainly 2D, and they do not highlight the 3D structure of a façade which is quire relevant for thermal analysis. Also when digital models are available, the model accuracy in the design phase is not ensured throughout the entire construction process. Therefore, a new digital model may be used to check the original one. Second, building models must incorporate all updates due to modification of the structure during its lifetime.

The availability of as-built models could provide a solution to these issues. Three-dimensional information related to architectural data is often complex, and thus it needs an acquisition method that may extract fine details in order to describe every element of the building envelop in the best way. In particular, nowadays an increasing interest is paid to the generation of detailed as-built building models from *terrestrial laser scanning data* (TLS), not only in cultural heritage preservation applications, where surfaces are complex and irregular, but also for large and medium size civil structures. This is mainly due to the fact that automation in acquisition and registration of scans, in conjunction with a reduction of the cost of instruments, allowed to a larger number of operators the chance to use TLS.

However, raw point clouds derived from a TLS survey are generally not directly ready for modelling applications. This is due to the nature of point clouds, which is not associated with any topological relations and presents a very low level of abstraction for exhaustive analysis (for example the measurement of areas, rapid copy and paste operations, etc.). In addition, further problems arise due to the huge size of data to be managed. The resolution of point clouds usually drops in correspondence of edges, corners and other linear features which are the ones playing the most important role in modelling. In order to overcome those problems, point clouds of buildings are generally transformed into vector format by adopting standard reverse engineering approaches. However, this task requires a largely time-consuming manual work that should be carried out by skilled operators. For this reason an increasing interest is paid to automatic modelling of buildings used in conjunction with or as an alternative to computer-aided-design (CAD) techniques. Indeed, manual editing can be avoided in the case a detailed modelling is not required, or may take place only after a basic model has been automatically generated. Automatic processing of point clouds is also important for the timely extraction of useful information and for reducing the cost with respect to a fully manual approach. Although much research work has been done, practical applications of automatic modelling of full façades is still far to cast into practice.

Motivation of the research

Point clouds acquired from real-world façades commonly feature several properties that pose challenges for any algorithm that further processes this raw data to extract information. One major issue is the lack of any structure or topological information in generic point clouds. This characteristic does not depend on the acquisition process but is rather a natural property of such data sets. As a consequence, identification of the true shape of the underlying surface may be a complex task even in densely sampled regions. In the case of real-world scans this difficulty is further intensified by the inevitable presence of noise and outliers (e.g., due to moving objects in the scene). Moreover, acquisition may be incomplete in the sense that large parts of the geometry remain hidden to the scanning device due to occlusions and restrictions on scanner placement. Last but not least, the size of the generated point clouds can easily be in several million points. On the one hand, the large size of point clouds does not necessarily reflect corresponding information content. This is because the scanning device does not adapt the sampling rate to the acquired local geometry and therefore even flat areas (e.g., house walls) will be sampled at the same resolution of fine details. It thus seems necessary and reasonable to investigate data representations that incorporate redundant information on a higher level. In this context, an observation fundamental to this work is that in scenes where man-made objects predominate, like in the cases addressed in the previous section, large parts of the acquired geometry can usually be efficiently represented by a set of simple parametric primitives. Once detected, these primitives achieve the desired effect with respect to the unnecessary redundancy in the point-cloud, because each of them resumes its associated set of corresponding points.

The primary goal of this dissertation is to address the above mentioned challenges posed by point clouds of building façades. A first step towards this end is developing an efficient method for detection of façade planar objects. While fitted shape primitives have been previously used for reverse engineering and in computer vision application, this research strives to further extend these methods to better exploit the primitive representation for building façade modelling.

The detection of high-level structures should achieve computational efficiency even on large point clouds due to their concise representation which can lead to the production of the façade vector model. At the same time they can offer improved quality due to the geometric and semantic cues provided by the primitives in the specific field of façade objects. In particular, the size and shape of a detected primitive along with the spatial relationship with the others can also serve as a rough classification of the façade parts which is helpful for recognition tasks. On the other hand, building façades often exhibit a regular arrangement consisting of repeated patterns and self-similarities. The detection and the parameterization of these regular

patterns can be used to infer missing geometry due to incomplete survey and to reduce holes induced by occlusions.

Overview of the dissertation

In this dissertation a novel methodology for fully automatic façade modelling is presented. In particular, the procedure follows a scheme that goes from scan acquisition up to point-cloud segmentation, automatic breaklines extraction, object classification, façade regularity detection and generation of the final 3D Building Information Model (BIM) of the façade finalized for thermal retrofitting (Fig. 3). Each step of the procedure presents some advances with respect to the state-of-art.

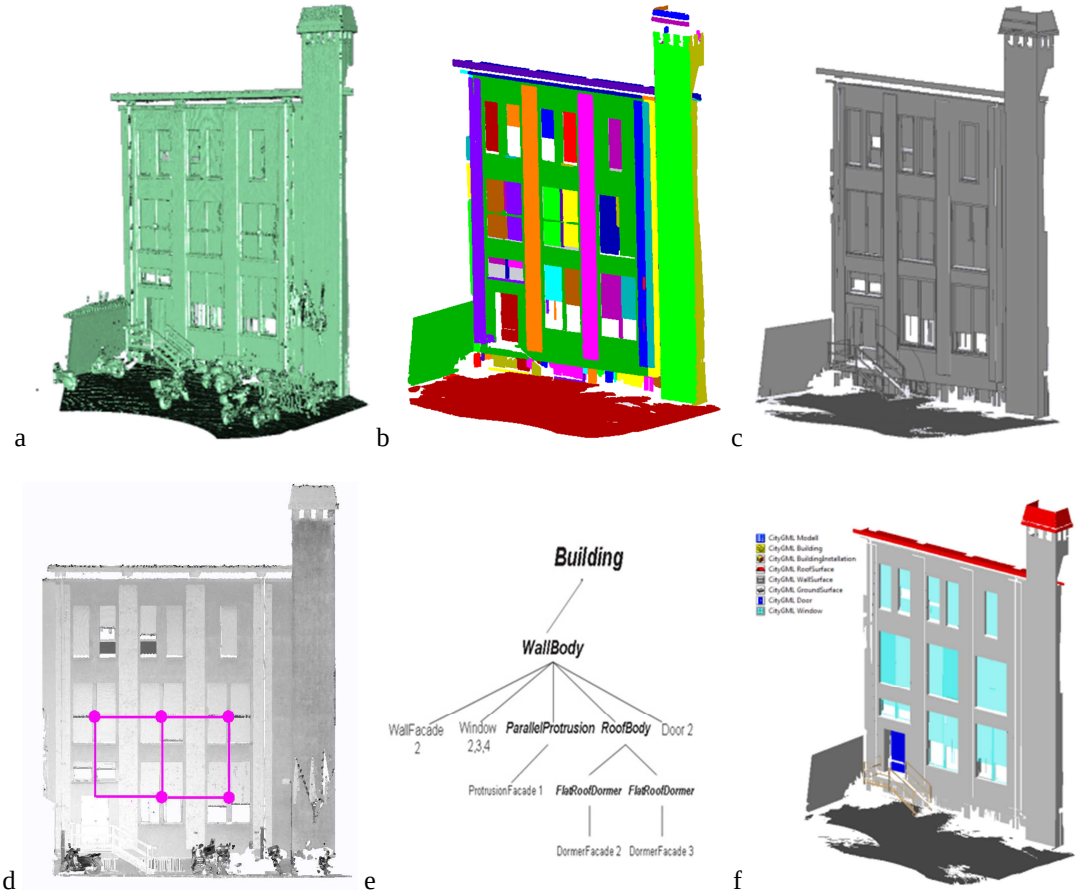


Fig. 3. Building façade at different processing steps: (a) row point cloud; (b) detected planar elements; (c) output vector model; (d) repeated pattern detected; (e) façade object classification; and (f) final CityGML model with semantics.

This dissertation addresses the application to modern building dated between 1950 and 1975 which are the ones for which thermal retrofitting is more demanding. For

residential buildings built in this period elements constituting the façade can be largely modelled as planar objects or at least can be approximately divided into piecewise planar parts. The developed modelling methodology can be applied to unstructured point cloud of tens of millions points. This means that each point is parameterized by its spatial coordinates and may also feature some related attributes (e.g., intensity, colour, normal vector), but does not share any topological relationships with other points in the nearby. The input point cloud can be generated by a single or multiple laser scan station(s). Indeed, after scan registration/geo-referencing, scans are merged together without needing any reorganization into a specific data structure. The overall procedure is presented in Fig. 4.

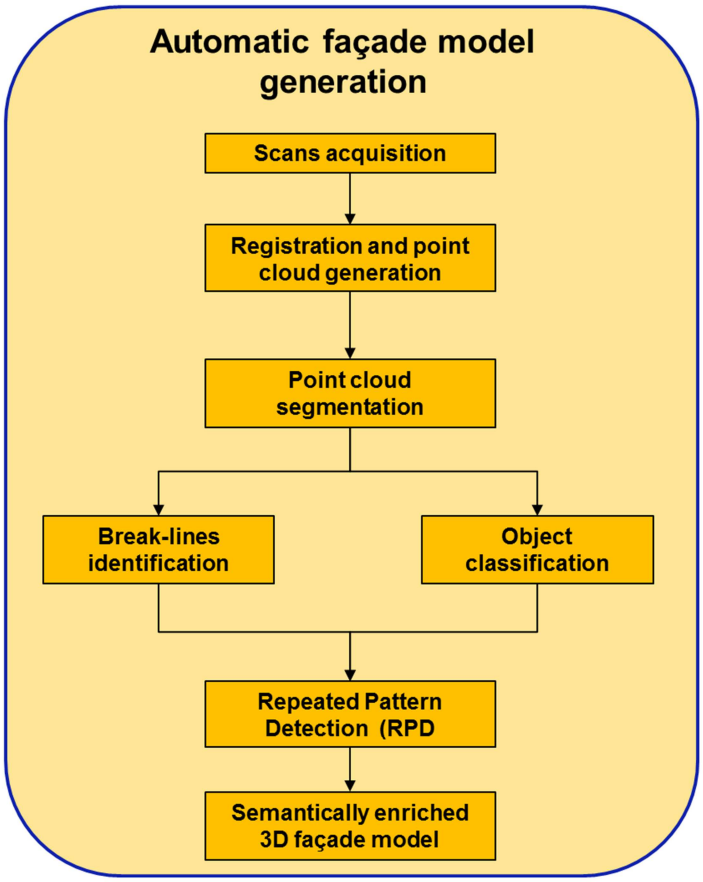


Fig. 4. The flowchart of the developed automatic façade model generation method.

Once all scans are acquired and registered together to output a non-structured point cloud, the main elements constituting the façade are firstly identified by means of a segmentation process based on a modified RANSAC implementation. In particular, the standard RANSAC approach (Boluaassal et al. 2007) for point cloud segmentation is modified by considering topology into the process. This topological information

between scan points and detected clusters are added in order to minimize problems connected to under- and over-segmentation, respectively (Awwad et al. 2009). Once planar clusters constituting the façade are detected, façade breaklines are automatically derived. During this phase some constraints related to façade geometry, like the prevalence of straight lines and orthogonal intersections, are enforced to obtain a regularization effect. At the same time, a further classification is performed, on the basis of some priors on the façade structure organized in a classification tree. In this way detected objects are classified into façade elements (e.g., walls, windows, etc.). Façades of buildings addressed in this work present a high regular structure characterized by the repetition of some elements forming a regular lattice structure. Identification of these regularities can be really useful to partially overcome problems connected to the lack of data and occlusions. For this reason a new automatic algorithm is developed for façade repeated pattern detection (RPD). Finally, detected breaklines along with façade classification results are merged together to obtain a semantically enriched 3D model of the entire façade.

Structure of the dissertation

This dissertation is divided in two parts. Part A gives a description the adopted approach. Part B reports some tests and presents possible integrations and extensions of the algorithms presented in Part A.

In particular, Part A is divided in four Chapters discussing the main steps of the methodology for façade modelling.

In Chapter 1 the developed segmentation strategy for the building façade is presented. Firstly, a review of the state of the art is reported. Particular attention is kept to the limits of existing techniques and to the expected developments. Then, a new segmentation strategy is presented to overcome some important drawbacks of previous approaches.

In Chapter 2 is described the extraction process of edge from the point cloud along with their enhancement and smoothing. The goal is to generate a concise model of the façade by detecting breaklines and a vector model of the building starting from them. The content of the first two Chapters has been also published in Previtali et al. (2013a).

Chapter 3 focuses on the semantic classification of extracted objects. In particular, a series of attributes and spatial relationship are defined for each façade segments which allow their functional classification into façade objects (e.g., wall elements, doors, windows, etc.). For the analysed building type, a series of recognition rules are derived from these features and organized into a hierarchical classification tree.

Finally, possible outputs of the processed data are presented. The content of this Chapter has been also published in Previtali et al. (2013d).

Chapter 4 discusses the reconstruction of incomplete models with the help of high-level architectural objects and identification of repeated patterns in urban façades. Completion of point clouds is often necessary because, despite considerable attention has been put during data acquisition, 3D views obtained with range scanners usually suffer from occlusion. On the other hand, building façades exhibit a high degree of self-similarity and redundancy. For this reason the method presented in this chapter is based on the detection of repeated patterns in the façade and the completion of missing parts by means of high-level architectural features (e.g., windows, doors, etc.). The content of this Chapter has been presented in Previtali et al. (2013c).

Part B is organized in two chapters presenting some results of the developed façade modelling approach (Chap. 5) and discussing its possible extensions (Chap. 6).

Chapter 5 examines the efficiency and accuracy of the developed reconstruction method presented in Part A. Five datasets of typical façades are presented and outcome discussed. Quality evaluation has been done for each step of the developed procedure to evaluate overall accuracy and precision. In particular the obtained façade models are evaluated against manually derived models considering the metrical quality and the completeness of the results.

In Chapter 6 some extensions of the developed automatic façade modelling approach are introduced. In Section 6.2 is presented the integration between the automatically generated building models with other data sources, like RGB and thermal images, to obtain a more complete building representation (Previtali et al. 2013b). Some of the developed algorithms can be used for applications other than façade modelling. In particular, the developed segmentation strategy can be used for scan registration and modelling of indoor environments. Indeed, the segmentation algorithm can be used for registering scans using detected planar features (Sect. 6.2). In addition, point clouds are often used not only for generation of detailed models of building façades but also for indoor environment modelling. Typically, many of the algorithms useful for the reconstruction of exterior building models can be adapted for indoor reconstruction, as well. However, due the large amount of clutter and occlusion in indoor application a tailored solution is needed (Sect. 6.3).

PART A

Chapter 1
Façade segmentation

This chapter presents a novel segmentation method that prepares input data for subsequent processing techniques proposed in this work. Indeed, the segmentation process is the first step for turning out a 3D unstructured point cloud into a building model. Terrestrial laser scanning (TLS) data, as those from other tools for 3D point cloud gathering, are not immediately compatible with numerical models. This is due to the fact that that no surfaces or edges are directly provided. To this end, the segmentation process can be used as a first recognition stage for further modelling as described in the following chapters.

Since the developed shape detection strategy is primarily intended for processing real-world scanned point clouds, it needs to specifically address the challenges of this type of data. In particular, the two most relevant issues arising in this setting are: the unstructured nature of point cloud data, and the corruption induced by noise and outliers.

Instruments for 3D scanning and imaging generally do not provide structured point clouds. Also in the case single scans can be handled as structured data (i.e., data with topological relations like proximity), see also Kang et al. (2013a), when multiple scans are performed and merged together the existing topological relations between points are lost. This typically happens with multi-station TLS surveys. In each scan, points are detected in correspondence of nodes of a regular angular grid established on a sphere around the instrument centre. However, when multiple scans are co-registered, this data structure is lost.

Presence of noise and outliers in the point cloud suggest using robust techniques for segmentation and detection of primitives. A very popular and versatile robust algorithm for segmentation is based on the RANSAC scheme proposed by Fischler and Bolles (1981), which, beside its simplicity, is very general and effective. Compared to other robust algorithms it has only a relatively small set of parameters which in addition have intuitive interpretations. Last but not least, RANSAC requires only very little memory in addition to the point cloud itself, which is important when working with large data sets. However, in its original formulation, the RANSAC

scheme is just an algorithm measuring the consensus to a certain assumed model. Consequently, the extracted primitives may not correspond to real architectonical object determining the so called 'bad-segmentation' problem. Thus, this Chapter presents a modified RANSAC algorithm for shape detection in point-clouds trying to minimize spurious results.

As previously anticipated this dissertation is primarily focused at obtaining as-built façade models for thermal retrofitting purposes. In particular, it addresses to modern building dated between 1950 and 1975 which are the ones for which thermal retrofitting is more demanding. Residential buildings of this period feature a similar architectural style. In particular, elements constituting the façade are generally planar objects or at least can be approximately divided into piecewise planar parts. This aspect strongly influences the segmentation. Indeed, the developed approach is aimed at identifying planar elements constituting the building façade while the detection of other shapes is not relevant for buildings' styles analysed in this research.

1.1. Introduction and definitions

Segments are geometrically continuous elements of object surface that share some similarities (Tóvári 2006). In Rabbani et al. (2006) 'segmentation' is defined as the process of labelling each point in a point clouds so that the points belonging to the same surface or region are given the same label. In this process, points having similar features in a continuous region are grouped to create a 'segment'. Formally segmentation can be expressed as:

$$\Theta P = \{\theta p \mid \forall\ p \in P\} \tag{1.1}$$

where Θ is the segmentation operator in point clouds P and θp is the label assignment for a single point p in P. The results of the segmentation operation are n segments (s) having following properties:

- $S = \{s \mid s \subset P\}$ 1.2
- $\Theta P \Rightarrow S$ 1.3
- $\bigcup s_i = P \ \ where \ \ |s_i| > 0$ 1.4
- $s_i \cap s_j = \emptyset \ \ where \ \ i \neq j$ 1.5

This means that each segment s is a closed subset of point clouds P (1.2) while the segmentation operator Θ determines the character of each segment (1.3). In particular,

every point in the point cloud belongs to a segment only and two different segments do not share common points between them (1.5). Each segment is characterized by its homogeneity measured according to certain features. These features generally represent geometric properties, reflectance intensity of laser pulse and spectral properties of the points. The homogeneity criterion used to perform segmentation is determined by the aim of segmentation itself. For example, planar surfaces are generally sought for building modelling purpose in which segment represents the part of building object such as roof or wall. Typically, in laser scanning data geometrical properties, such as mathematical surface, surface normal, gradients and curvature (Crosilla et al. 2009) in the neighbourhood of a point, are mainly used for segmentation purposes. The reflectance strength of laser pulses, i.e. intensity data, as well as spectral properties of points are rarely used in the segmentation process because of their noisy character (Tóvári 2006).

Property 1.4 is not always valid for unstructured point clouds as there may be some points not assigned to any segments. Other limits exist when focusing on massive unstructured point clouds. These problems are termed in literature as 'bad-segmentation' (Stewart 1997) and may be categorized as:

- under-segmentation, in the case several features are segmented into one;
- over-segmentation, when one feature is segmented into several ones; and
- no segmentation: feature is not segmented or wrongly segmented.

Main reason for bad segmentation is that real point cloud are naturally noise affected, due to the accuracy of the instrument which usually varies from a few centimetres (for example, in Mobile Laser Scanning data sets) up to a few millimetres. Unfortunately, all the segmentation algorithms suffer from noisy data because make more difficult the prediction of homogenous areas or hinder the plane and curve estimations. As a consequence, if segmentation parameters are sensitive enough, most of the proposed methods will lead to high rates of point cloud over-segmentation; in alternative, using insensitive segmentation to face towards noisy data, an under-segmented point cloud will be obtained. In particular, according to Sapkota (2008) the presence of spurious planes and bad segmentation results, i.e., results not useful for users, is typical to any actual segmentation procedure presented in literature. This is due to the fact that in many cases extracted planes and object from the point cloud do not correspond to real architectonical objects.

The different segmentation strategies differentiate in the way how to deal with 'bad segmentation' problems.

1.2. State of the art

Point cloud segmentation is one of the main research areas in the laser scanning field. Indeed, it can be designed to introduce some level of organisation to the data before extraction of useful information (Filin and Pfeifer 2006), or as a precursor to object recognition and model fitting (Rabbani 2006). Finally, can be also employed as a pre-processing step before the application of filtering algorithms (Tóvári and Pfeifer 2005). In the past decades, various algorithms have been designed to extract planar surfaces from point clouds, mainly using geometric criteria to detect homogeneous regions in data. However, although the large number of works in literature focusing on segmentation, this problem is far from being solved even for planar features. The different segmentation methods proposed can be categorized into four groups (Vosselman and Maas 2010):

- Feature clustering;
- Surface growing;
- Model fitting; and
- Hybrid techniques.

In the next subsections the basic principles of the above mentioned methods are explained and some of the relevant algorithms proposed in literature are discussed.

1.2.1. Feature clustering

The method based on feature clustering offers a general and flexible way to identify homogeneous segments in data, without being restricted to one specific pattern. *Feature clustering* methods can be seen as a combination of two processes: first, patterns in the data are identified on the basis of some attributes, and then they are clustered based on them.

Firstly, representative measures (‘features’) are defined for each point based on some geometrical and radiometric characteristics. These features generally include position of each point, locally estimated surface normal, residuals of best fitting surface, and reflectance of laser scanning points. All these items of information are used to generate an nD feature. Thereafter, clusters are identified in an nD feature space. The points belonging to each cluster in the feature space are labelled as unique segment in the object space.

Even if this method is quite general, identification of proper features may be complex and segmentation results are highly dependent upon this selection as well as the methods used for partitioning the feature space. Since in many cases the features of individual points are described using points in local neighbourhood, this segmentation technique is also sensitive to noise in the dataset and is influenced by the definition of neighbourhood. Therefore, an additional robust method is needed to eliminate the

noisy data and outliers. In addition, clustering multidimensional features can be computationally infeasible for large point clouds.

Filin (2002) presented a clustering algorithm using an unsupervised classification technique for extracting homogeneous segments in Airborne Laser Scanning (ALS) data. A 7D feature vector is defined for each point. Features consist in point coordinates, surface parameters of a plane fitted to the point neighbourhood and the relative height difference between the point and its neighbours. Then, instead of creating a 7D feature space, the author separates positional information to create a 4D feature space. After defining surface classes, points are grouped in object space by using spatial proximity measures. The goal is to find clusters that are spatially meaningful and at the same time to avoid over-segmentation.

Hofmann et al. (2003) defined a feature clustering segmentation method for TIN structures, which is derived from ALS point clouds. For each triangle of the data a 2D (slope and orientation) or a 3D (slope, orientation and O-distance) feature is defined. According to the authors the O-distance is calculated from the origin O to the plane containing the triangle.

An unsupervised clustering approach based on fuzzy methods is presented in Biosca and Lerma (2008). In this case both the Fuzzy C-Means (FCM) algorithm and the Possibilistic C-Means (PCM) mode-seeking algorithm are used in combination with a similarity-driven cluster merging method.

1.2.2. Region-growing

Region-growing methods relay on the main assumption that neighbouring points in the cloud presents similar characteristics. For this reason region-growing methods are based on the aggregation of homogenous points, basing on certain similarity criteria, starting from a point which is called *'seed'*. Therefore, it can be seen as a combination of two steps: (i) identification and (ii) growing of the 'seed' surface.

A 'seed' surface consists of a group of neighbour points that fits well into a defined geometric shape (e.g., a plane). For the 'seed' surface selection, a group of adjacent points are identified and tested whether they fit well the defined geometric shape or not. If a shape is found to fit within some predefined thresholds, it is accepted as seed surface; otherwise another point is tested. Once the 'seed' surface is selected, every point in it is examined to find the neighbour points that may fit to the defined shape. This operation is basically intended to grow the surface towards its neighbourhood. Points are added in the growing surface if they meet the predefined criteria. After adding a point, the equation of the fitted shape is updated. The acceptance decision for a point can be based on one or more of the following criteria:

a. *proximity of points*: only points that are within a certain distance from the current 'seed' surface are added to it. This proximity can be evaluated by checking the distance between two points on an edge of a Delaunay triangulation or, otherwise, considering the Euclidean or the Manhattan distance, i.e., distance between two points measured along axes at right angles (Niedermeier and Sanders 1996). This condition can, however, be too strict if some outliers are present (Vosselmann et al. 2004);
b. *global planarity* (in the case the fitting model is a plane): a plane equation is determined by fitting a plane through all points within a given radius around the selected 'seed'. A candidate point is accepted in the segment if the orthogonal distance from the plane evaluated by considering all the points already in the segment is within a defined threshold;
c. *surface smoothness*: local surface normal for each point in the point clouds is estimated and a candidate point is accepted if the angle between the local surface normal of the point and the normal of the growing surface is below a threshold;
d. *height difference*: the height difference between a point and its neighbours is computed as the distance from the point to the best fitting plane to its neighbours along the normal direction; and
e. *principal curvatures through Principal Component Analysis (PCA)*: the geometrical properties of an object in a point are intrinsically described by means of principal curvatures from PCA. The curvature ray is maximum in correspondence of planar areas while a smaller radius represents the proximity to an edge.

However, there is no universal criterion which has ubiquitous validity (Biosca and Lerma 2008).

Several variations in surface growing techniques of segmentation are suggested in the literature.

Vosselman et al. (2004) proposed an approach to automatically extract planar surfaces from TLS point clouds. In this approach, several parameters need to be specified for the planar surface-growing algorithm, such as the number of 'seeds', the surface-growing radius and the maximum distance between surfaces. Using different values for these parameters, it is easy to obtain significant changes in the segmentation results.

Tóvári and Pfeifer (2005) presented a segmentation method for ALS data. First, the normal vector at each point is estimated by using k-nearest neighbours (Samet 2006). Then a point is selected randomly and the adjacent points are examined for certain criteria. If the criterion meets, the adjusting plane is estimated using those 'seed'

points. During growing, the neighbouring points are added to the segment if they meet criteria (a), (b) and (c) in the above list. For plane adjustment, eigenvector/eigenvalue approach using the second moments of point coordinates are used.

Rabbani et al. (2006) presented a method to segment unstructured 3D point clouds of industrial scene based on smoothness constraints. Also in this case in a first step, normal for each point is estimated by using k-nearest neighbours and Least Squares plane fitting. Then residuals of plane fitting are used to approximate the local surface curvature. These residuals are sorted and used to select 'seed' points. In particular, points with minimum residual are considered suitable 'seed' points for detecting planar segments. The growing phase is carried out by using previously estimated point normal and their residuals. In particular, criteria (a) and (c) are implemented.

Another typical variation of region growing algorithm for ALS data is presented in Gorte (2002), where triangles are used as basic surface units. The merging of triangular meshes is carried out by comparing the plane equations of neighbouring triangles.

1.2.3. Model fitting

Methods in this category are based on the identification of geometrical primitives in the point cloud. Indeed, many man-made objects can be decomposed into geometric primitives like planes, cylinders, cones, tori, spheres, etc. This approach tries to fit primitive shapes in point cloud data. Thus, those points conforming to the sought primitive are labelled as one segment. However, outliers caused by noise, registration errors or miscalibrations are frequently encountered in laser scanning point clouds. For this reason a robust parameter estimation methods is needed to extract geometrical shapes in the presence of outliers. The two most important algorithms in the line of robust fitting are Hough transform (Hough 1962) and RANdom SAmple Consensus (RANSAC) introduced by Fischler and Bolles (1981).

The Hough transform maps, for a given type of parameterized primitive, every point in the data to manifold in the parameter space. The manifold describes all possible variants of the primitive that contain the original point. In practice each point casts votes for many cells in a discretized parameter space. Shapes are extracted by selecting those parameter vectors that have received a significant amount of votes. In the case of detection of planar objects, a general plane can be represented as:

$$Z = s_x X + s_y Y + d \qquad 1.6$$

where s_x and s_y represent the slope of the plane along X- and Y-axis, respectively, and d denotes the intersect of the plane on the Z-axis. This parameterization for a plane in 3D space does not describe vertical planes. To overcome this limitation, the *normal form* of the plane equation can be used:

$$\cos\theta \cdot \cos\phi \cdot X + sen\,\theta \cdot \cos\phi \cdot Y + sen\,\phi \cdot Z = d \qquad 1.7$$

where $\theta \in [0, 2\pi]$ and $\Phi \in [-\pi / 2, \pi / 2]$ denote two angles of the plane normal and d represents the distance from the origin to the plane. The major application area of the Hough transform remains the 2D domain where the number of parameters typically is quite small. However, also 3D applications are reported in the literature.

Maas and Vosselman (1999) adopted 3D Hough transform for detecting roof planes in 3D point clouds. The Hough space is described by two slope parameters and a distance (Eq. 1.6). This description of clustering space is only suitable for ALS data as this form of parameterization cannot describe vertical planes which are common in TLS data.

Vosselman et al. (2004) suggested a variation in Hough transform using surface normal to speed up the process of planar surface detection with increased reliability. The normal vector and the position of a point is adopted to define a plane whose parameters can be directly mapped to a single point in the parameter space. This solution avoids the process of computing the intersection of the plane with the corresponding bin. Only the increment of counter of each single bin is enough. The authors also proposed a two-step procedure for the Hough-based detection of cylinders. Similarly Rabbani and Van den Heuvel (2005) decomposed the sphere or cylinder detection problem into two sub-problems of a low complexity: the detection of the cylinder axis direction (2 parameters) and the direction of a circle in a plane (3 parameters).

The RANSAC paradigm is used to extracts shapes by randomly drawing minimal sets from the point data and constructing corresponding shape primitives. In particular, a minimal set is the smallest number of points required to uniquely compute a given type of geometric primitive (for example, 3 non collinear points to estimate a plane in space). The resulting candidate shape is tested against all points in the data to determine how many of the points are well approximated by the primitive (called the ‘score’ of the shape). After a given number of trials, the shape which approximates the most points (points well approximated by the primitive are also referred to as

'inliers') is reported. All inliers are then extracted from the dataset and used to estimate the detected geometric primitive.

Bretar and Roux (2005) proposed an algorithm for the detection of roof facets of buildings based on normal-driven RANSAC (ND-RANSAC). For this purpose, they first calculate the normal vectors for each point and then randomly select sets of three points having the same orientation of the normal vectors. The number of random draws is managed automatically by statistical analysis of the distribution of normal vectors using the Gaussian sphere of the scene.

The extension of RANSAC algorithm for roof plane detection is proposed in Tarsha-Kurdi et al. (2007). The authors used the number of trials as an input rather than probabilistic calculation. They suggested calculating it by using the point density and the foreseeable size of urban roof plane. Another adaptation over the standard RANSAC technique is that they use criteria based on standard deviation of distance from the plane to select a reduced point-set instead of the original one to evaluate the candidate feature.

RANSAC based algorithm for the detection of several geometrical shapes such as planes, spheres, cylinders, cones and tori is presented in Schnabel et al. (2007). In this method, they use localized sampling strategy using octree data structure (Samet 2006) for the random selection of the minimal subset of points. While evaluating the score of the candidate feature, several parameters (e.g., number of points within the tolerance distance of the feature, minimum deviation of the surface normal, etc.) are taken into account.

An application of RANSAC to massive unstructured 3D point clouds is reported in Boulaassal et al. (2007). In this case, large point clouds of building façades are analysed and segmented showing that a sequential application of RANSAC allows the automatic segmentation of planar surfaces from 3D point clouds acquired by TLS. However, some problems are reported in the case of adjacent planes and in the plane connectivity. In addition, the method has been proved only on small point clouds, limiting in this way computational problems.

To cope with the relative computational inefficiency of RANSAC in Kang et al. (2013b) a conditional sampling method based on the Bayesian sampling consensus (BaySAC) is proposed. This method selects the minimum number of data required with the highest inlier probabilities as a hypothesis set. Thus, reducing the number of iterations needed to find a good model.

1.2.4. Hybrid techniques

Some authors tried also to combine different methods to better exploit the potential of each of them. In general, region growing is combined with other plane detection methods as it takes into account the spatial proximity of the points in a more natural way.

Roggero (2002) combined hierarchical region growing and principal component analysis (PCA) to segment ALS data. PCA is used to define the aggregation criteria and to describe the geometrical properties of the surfaces. Two algorithms differing in PCA and in aggregation criteria are proposed. One of the algorithms is based on descriptor mapping. First, one or more properties like static moment or curvature or pulse intensity are computed and mapped to each point. Then the region growing is performed with reference to the property map. The second algorithm does not perform descriptor mapping and uses PCA in region growing phase to speed up the method.

Elberink and Vosselman (2006) used Hough transform for ‘seed’ surface selection combined with a surface growing approach. For some arbitrary point, k-nearest neighbour points are selected and Hough transform is applied to these points only. If the minimum number of points is identified to be in a plane by Hough transform, Least Squares are used to fit the parameters of the plane and the points are taken as ‘seed’ surface. The acceptable ‘seed’ surface is used here instead of the optimal ‘seed’ surface (having maximum number of points with minimum residuals) at the cost of computation speed. In the growing phase, the orthogonal distance of the adjacent points to the growing plane is analysed and the points are added to the surface if the distance is below a threshold.

1.3. Critical analysis of existing techniques

Summary of the state of the art of segmentation methods presented in the previous section is provided in Tab. 1.1. Based on the review of different segmentation methods, following conclusion can be drawn for each category.

The results of the segmentation process based on clustering of features are dictated by the choice and the quality of the representative features of each point. In addition, clustering multidimensional features for large data volume is computationally very expensive while dealing with large data volume is quite natural in laser scanning point clouds processing.

Region-growing based methods always consider spatial proximity of points. However, the quality of results from region-growing-based segmentation depends on the methodology used for ‘seed’ surface selection and the criteria applied for growing.

Segmentation method		**Work**	**Description**
Clustering		Filin (2002)	7D feature space is defined to cluster points.
		Hofmann et al. (2003)	2D and 3D feature space is defined for TIN structures
		Biosca and Lerma (2008)	FCM and PCM in combination with a similarity-driven clustering
Region growing		Gorte (2002)	Grouping of TIN, more suitable for ALS data.
		Vosselman et al. (2004)	Several parameters (number of seeds, the surface-growing radius, etc.) used to grow.
		Rabbani et al. (2006)	Local surface normal similarity as criteria to grow.
		Tóvári and Pfeifer (2005)	Surface normals, spatial proximity and distance to plane as criteria to grow.
Model fitting	Hough transform	Maas and Vosselman (1999)	Utilises slope form of plane equation to create 3D Hough space for roof plane detection.
		Vosselman et al. (2004)	Surface normal is introduced in Hough space.
	RANSAC	Bretar and Roux (2005)	Utilizes local surface normal for each point, uses RANSAC selecting three points with similar orientation of normals.
		Schnabel et al. (2007)	Localized sampling strategy using octree data structure. Different parameters are used in candidate evaluation.
		Tarsha-Kurdi et al. (2007)	Uses standard deviation of distance to plane and minimum number of points as candidate evaluation criteria.
		Boulaassal et al. (2007)	Uses RANSAC for façade segmentation. Problems are reported in the case of adjacent planes and in the plane connectivity.
		Kang et al (2013b)	Conditional sampling method based on the Bayesian sampling consensus
Hybrid		Roggero (2002)	Combines PCA and region growing
		Elberink and Vosselman (2006)	Hough transform for seed plane selection and then region growing based on the distance to plane criteria.

Tab. 1.1. Summary of the state of the art point-cloud segmentation methods.

The selection of the 'seed' points is of primary importance, because the final segmentation results depend on that: in fact, some problems occur when 'seed' points are extracted close to breaklines or outliers. As a consequence, none of these algorithms can be considered a robust solution. It can be noticed that several parameters need to be specified for the surface growing algorithms whose selection is not always an easy task. Indeed, by using slightly different values for these parameters, it is easy to obtain various bad segmentation problems (over-, under- and/or no segmentation). With the same number of 'seeds', larger surface-growing radius or larger maximum distance from surface may lead to under-segmentation. On the other hand, smaller surface-growing radius or smaller distance from surface may lead to over-segmentation. Pu and Vosselmann (2008) stated that over-segmentation is preferable to under-segmentation, because over-segmented parts may have some similar properties that can be exploited to merge these segments together at a later stage.

The segmentation based on surface fitting using Hough transform or RANSAC is effective in presence of noise and outliers. However, the straightforward implementation of both techniques is computationally inefficient. In addition, points classified as belonging to the same detected shape by these techniques may not necessarily belong to the same object surface. To separate points belonging to different object surfaces specific strategies should be employed. While both Hough transform and RANSAC strategies were used for processing point clouds, an important comparison between them was made in Tarsha-Kurdi et al. (2007) in terms of processing time and sensitivity to cloud characteristics using ALS data. The authors show that RANSAC is more efficient than the Hough transform algorithm, since the difference in processing time is negligible even when data size is very large and Hough-transform is also very sensitive to segmentation parameters values.

The hybrid techniques have some desirable strength as they exploit the benefits of more methods. Combination of model fitting on local region with expansion towards adjacent points using region growing is one of the most efficient segmentation strategies.

As a general consideration, segmentation techniques analysed in the former section had resulted in showing good results in different application domains. Nevertheless, none of them has achieved results to be considered the optimal segmentation method at least for a specific category of applications: the final results needs to be improved by manual editing or post-processing. Furthermore, algorithms lacks of flexibility: a method that works well in a specific case can hardly be used for other applications. Many methods have a large number of parameters, whose meaning and effects on final segmentation are not always clear (Rabbani et al. 2006) and make the reliability

of the results very variable according to the specific conditions. Some methods use fuzzy logic to interactively estimate the best set of parameters and their values, even if these approaches require a training phase that it is not always possible.

To partially overcome the previous listed limitations of state of the art techniques a new automatic approach is presented for the segmentation of planar surfaces based on the combination of RANSAC algorithm and region-growing approaches. The aim of this strategy is to derive 'meaningful' segments from building-façade point clouds. This means that extracted segments would correspond to objects of interest (e.g., roofs, walls, doors, etc.) or parts thereof instead of being simply those which best fit some mathematical models. In particular the main improvements of the presented methodologies are:

- creating a new approach that can deal directly with 3D point clouds from TLS;
- segmenting and extracting different planes with best fit to reality for complex objects starting from massive unstructured 3D point cloud in the presence of noisy data;
- reducing over-segmentation problems by introducing a new score function based both on signed distances and normal directions, and defining topology information (i.e., definition of neighbourhood between points) in the point clouds; and
- reducing under-segmentation by performing a clustering of extracted planes based on topology properties for surfaces.

1.4. Façade segmentation

Given a point cloud $P = \{p_1,\ldots, p_N\}$ with associated normals $N = \{n_1,\ldots,n_N\}$ the output of the implemented algorithm is a set of planes $\Psi = \{\psi_1,\ldots, \psi_N\}$ with corresponding disjoint sets of points $P_\Psi = \{P_{\psi 1} \subset P,\ldots, P_{\psi N} \subset P\}$ and a set of remaining points $R = P \setminus \cup_\psi P_\psi$.

The overall structure of the method is outlined in Fig. 1.1. In each iteration of the algorithm, the primitive with maximal score is searched using the RANSAC paradigm. The score function has two free parameters: (i) ε specifies the maximum distance of a compatible point while (ii) α restricts the deviation of a points' normal from the one of the plane to judge.

Plane candidates are generated by randomly sampling minimal subsets of P. After a new candidate has been generated, the one with the highest score is computed. The best candidate is accepted if, given the number of inliers $|m|$ of the candidate and the number of drawn candidates $|C|$, the probability $P(|m|, |C|)$ that no better candidate was overlooked during sampling is high enough.

After having determined the maximum consensus plane a check is performed to verify if the set of points $P_{\Psi j}$ belong to a single façade object or more (Forlani et al. 2006). This check is performed by setting up a bitmap and finding all connected components in it. In this phase some spurious segments might be found. Indeed, each segment whose area is far lower than the mean value is rejected. When a candidate is accepted, the corresponding points $P_{inliers}$ are removed from P and the candidates C_m generated with points in $P_{inliers}$ are added to Ψ. The algorithm terminates as soon as $P(\tau,|C|)$ for a user defined minimal shape size τ is large enough.

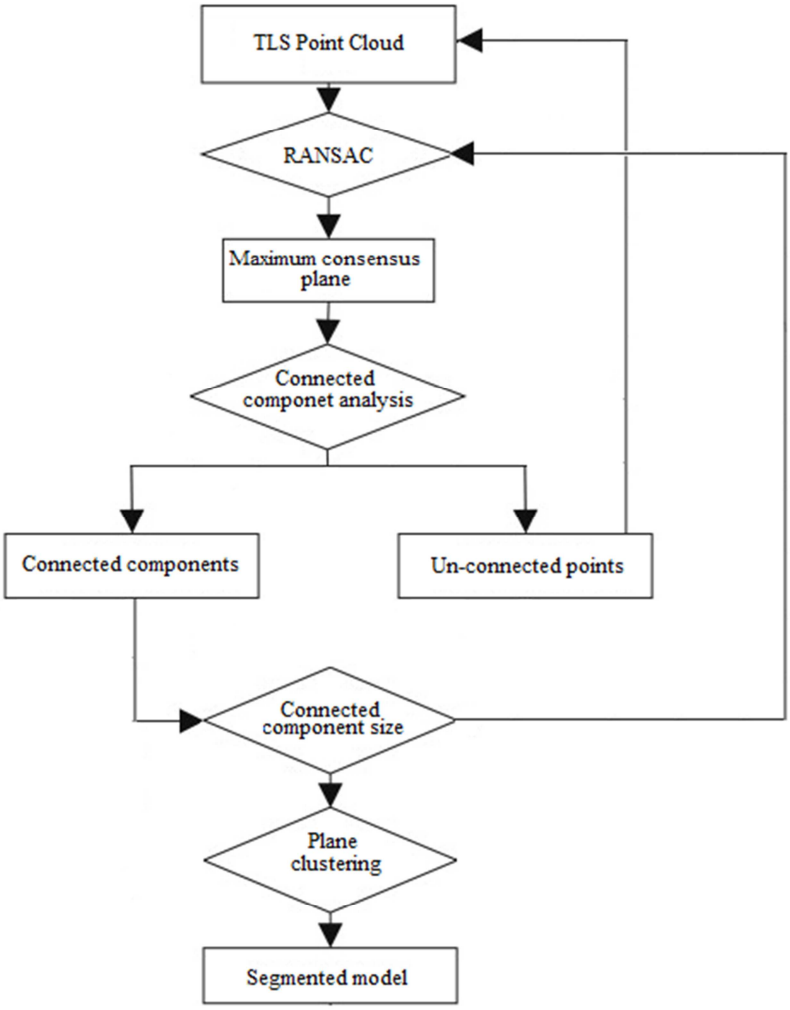


Fig. 1.1. Workflow of the developed façade segmentation process.

After having determined all planar elements a clustering of extracted object is performed to minimize over-segmentation problems. Clustering is performed by means of mean shift clustering in the plane domain on the base of: (i) plane normals, (ii) distance between extracted segments and (iii) segment intersection.

1.4.1. Point cloud pre-processing

The term 'point cloud' usually refers to an unordered collection of 3D spatial locations that can additionally be equipped with a set of attributes at the respective position in space:

$$P = \left\{ \left(p_1, a_1^1, ..., a_1^k \right), \left(p_2, a_2^1, ..., a_2^k \right), ..., \left(p_N, a_N^1, ..., a_N^k \right) \right\} \quad 1.8$$

P denote a point cloud with coordinates $p_i \in R^3$ and attributes a_i^j. Typical attributes are the local normal $n_i \in S^2$, point colours $c_i \in [0, 255]^3$, and laser intensity $I_i \in [0, 255]$.

The most popular acquisition techniques of point clouds, laser range scanning, structured light scanning, shape from shading and multi-view photogrammetry in general all produce unstructured and irregular 3D point clouds that are sampled from the acquired surface geometry. This means that no information of proximity (topology) is assigned to any point.

In the segmentation stage local normal is used in the score function definition. However, while some attributes like colours or intensity can be directly acquired from scan scanning other like normal vector information at the point locations need pre-processing.

Several approaches are presented in the literature for estimating local normals in a point cloud (Hoppe et al. 1992, Pauly et al. 2005, Jenke et al. 2008). All of them rely on a similar framework: for each point a plan is locally fitted considering a certain number of neighbour points. In the developed pre-processing step the local normal is estimated using the method proposed in Jenke et al. (2008). This method iteratively increases the size of the neighbourhood until the estimated normal and the eigenvalues of the weighted covariance matrix C (N_k *(p)*) stabilize:

$$C(N_k(p)) = \sum_{j \in N_k(p)} (p_j - \mu_k)(p_j - \mu_k)^T \phi\left(\| p_j - p \| / h_k \right) \quad 1.9$$

where μ_k is the weighted average of all points is N_k *(p)*, h_k is the radius of the smallest sphere containing N_k *(p)* cantered at p and ϕ is a positive monotonously decreasing weighting function. With respect to other approaches, this has the advantage that potentially less neighbourhood sizes have to be considered if stability is detected early

on. Even if this approach is heuristic since the algorithms in this work do not require correct normal close to sharp features and can easily deal with some smoothing in these areas, the method of Jenke et al. (2008) was found to work quite well in practice.

1.4.2. RANSAC model fitting

As mentioned above the façade types considered in this research are mainly characterized by planar objects. For this reason the segmentation step consider only identification and extraction of planar segments from the point cloud. Every 3D point p_i fixes only one parameter of the shape so that for a plane, $\{p_1, p_2, p_3\}$ constitutes a minimal set. To confirm the plausibility of the generated plane, the minimal set is accepted only if the three point normal n_1, n_2, n_3 deviations are less than the predefined angle α. Once the subset is accepted and the plane passing through these tree points is estimated, the score function F is used for measuring the quality of a given plane candidate. The score function F takes into account the following aspects:

- the number of points that fall within the ε-band around the plane; and
- to ensure that the points inside the band roughly follow the direction of given plane, only those points inside the band whose normal do not deviate from the normal of the plane more than defined angle α are considered as inliers for the guessed plane.

More formally, given a candidate shape C whose fidelity is to be evaluated, F is defined as follows:

$$F(C) = \left| P_\psi \right| \tag{1.10}$$

i.e., $F(C)$ counts the number of points in P_ψ. Being P_ψ defined as (Fig. 1.2):

$$P_\psi = \{p \mid p \in P \wedge |d(\psi,p)| < \varepsilon \wedge arcos\,[n(p)\cdot n(\psi,p)] < \alpha\} \tag{1.11}$$

where $d(\psi, p)$ is the signed distance of point p to the plane Ψ, $n(p)$ is the normal in p and $n(\psi, p)$ is the normal of Ψ in p's projection on ψ.

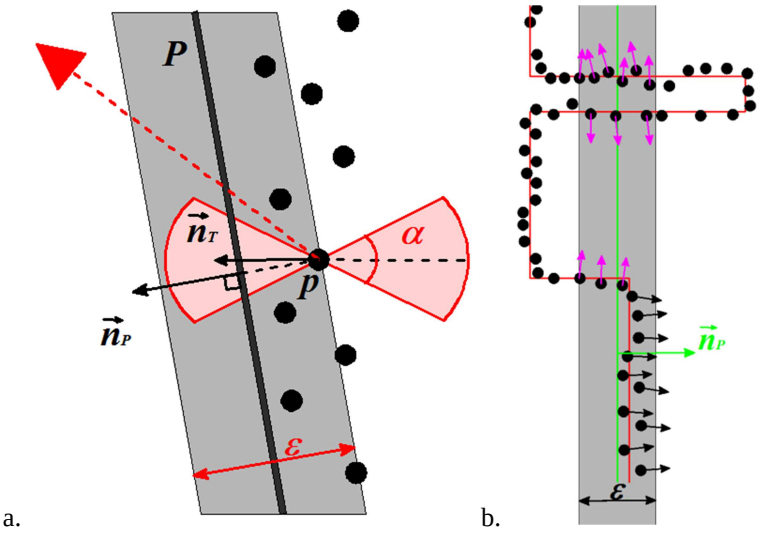


Fig. 1.2. RANSAC score function: (a) selection criteria of inlier point *p* for the plane *P*; and (b) effect of the defined score function for inlier detection: points marked in purple are excluded from inlier detection because their normal *n* deviate from the guessed plane normal n_P.

In particular, the signed distance function for a plane is given by:

$$d(x) = < n,x - p > = < n,x > - < n,p > \qquad 1.12$$

where *n*, |*n*|=1 is the normal to the plane and *p* is an arbitrary point in the plane. The intuitive threshold value ε for the Euclidean distance between a point and an estimated plane can be easily found by the user according to the instrumental noise and minimum point density for the acquired point clouds.

As specified in the definition of the score function F, in contrast to other RANSAC approaches (Boluaassal et al. 2007), in the implemented segmentation strategy not only the signed distance is evaluated but also the compatibility of the local point normal with the estimated plane normal.

While comparison of local surface normal is not necessary for planar detection, where distance to the plane is sufficient criteria, by employing this further check there is no way that a maximum consensus plane is obtained from spurious surfaces containing points having different normal directions (Fig. 1.3). This reduces some of the bad-segmentation results reported in Awwad et al. (2009). It must be noted that the plane direction, which is obtained by RANSAC, is based on the sequence of choosing the three points; therefore, the check in Eq. (1.11) also considers the opposite direction for the plane normal.

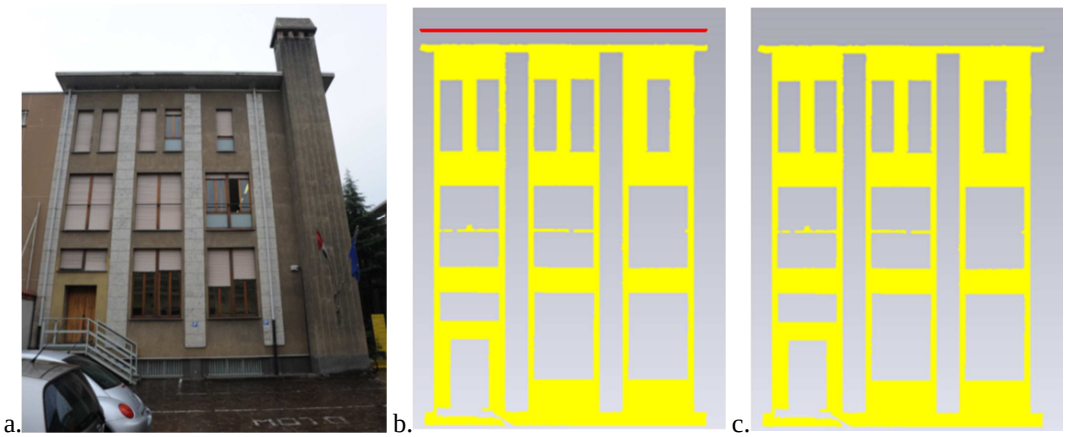


Fig. 1.3. Example of bad segmentation removal with the defined score function for 'D'Oggiono building' dataset (a). Results with simple signed distance criterion: points highlighted in red (belonging to the roof) are erroneously clustered with the main façade wall (b). Segmentation results with point vector consistency check (b)

However, this two-step approach increases the computational complexity. To face this problem, the efficiency of RANSAC is improved by implementing an adaptive approach to determine the number of iterations necessary to extract, with a probability *p*, the shape that achieve the highest possible score (Hartley and Ziesserman 2004). This problem turns out in evaluating the number of candidates that have to be considered to guarantee, that the minimal set is drawn to define this shape, given a predefined probability *p*. This solution is quite popular in RANSAC applications when exploiting all possible combinations of points to form the minimal set is computationally infeasible.

In particular, the minimum number of trials (*T*) to extract with probability *p*, the plane that achieve the highest score is:

$$T \geq \frac{\ln(1-p)}{\ln\left(1-(1-\varepsilon)^n\right)} \qquad 1.13$$

where ε is the outlier percentage and *n* is the number of parameters to be estimated (in the case of a plane $n = 3$). Being the percentage of outliers unknown at the first iteration the outlier percentage is fixed near 100%. When a new maximum consensus plane is detected the outlier percentage is calculated according to the current number of inliers and the number of required iterations is updated (Tab. 1.2). In the case the requested number of iterations has been reached the process stops, otherwise the loop iterates.

n = 3	Outlier percentage						
	5 %	30 %	50 %	70 %	80 %	85 %	90 %
p_t =99 %	2	11	34	168	573	1362	4602
p_t =95 %	2	7	22	109	373	886	2994

Tab. 1.2. Minimum number of iteration for plane detection wit p = 95% and p = 99%.

When a candidate shape C is accepted as the maximum consensus in the point cloud the corresponding points $P_{inliers}$ are removed from P and the process iterates until new shapes are detected having a size larger than a user defined threshold, defined as a percentage of the original point clouds (e.g., 3% of the original point cloud).

Algorithm 1 Extract planes in the point cloud P

```
Ψ ← Ø{extracted shape}
repeat
    C ← Ø {candidate shape}
    nBest ← 0 {number of points in the best candidate}
    i ← 0 {number of iterations}
    T ← ∞ {number of iteration for having P (|m|, |C|)>p_t}
    repeat
        P3 = SelectRandomSubset(P) {select minimum subset of points i.e. 3}
        ψ = ComputePlane(P3) {compute plane from subset}
        nInliers = ComputeInliers(ψ, P) {compute no. of consistent points based on a F}
            if nInliers>nBest
                nBest = nInliers
                C ← ψ
                T=NumberOfIterations(nInliers, P)
            end if
    until i< T
    return C
    P ← P \ P_inliers {remove inliers points}
    Ψ ← Ψ ∪ C{add extracted shape}
until P(τ, |C|)>p_t
return Ψ
```

When a shape having a number of inliers lower than the threshold is detected as the one having the maximum consensus the segmentation stops. The workflow of the developed RANSAC algorithm for segmentation is reported in Algorithm 1.

1.4.2. Under- and over- segmentation removal

A new approach based on topology information was developed to cope with both under- and over- segmentation problems.

Under-segmentation comes up when several features in the point cloud (ψ_1,ψ_2, ..., ψ_n) are erroneously detected during the segmentation process an assigned to the same segment (ψ_0). This problem mainly arises when encountering special cases such as two or more planar surfaces having the same normal and being at the same level. This situation typically occurs for windows (Fig. 1.4).

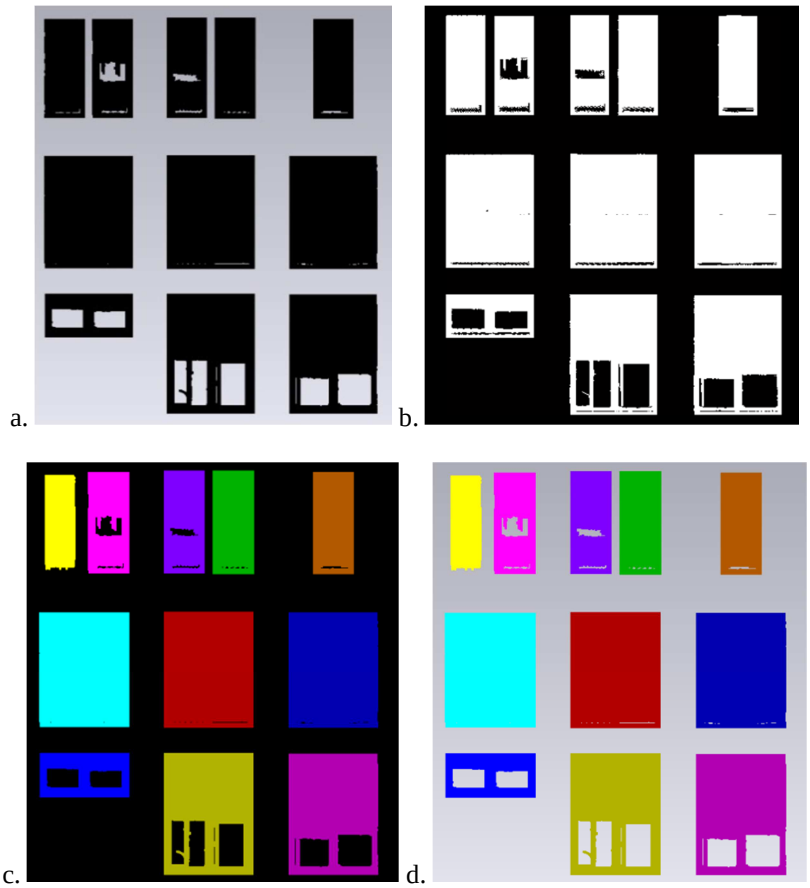


Fig. 1.4. Example of under-segmentation removal for a set of windows: (a) all windows belonging to the same plane are clustered in the same segment; (b) an occupancy map is generated with pixel size β=1cm; (c) connected components in the bitmap are detected; and (d) final segmentation results.

Indeed, all windows are generally on the same plane and are erroneously classified as one single object. This is mainly due to the fact that RANSAC does not explicitly held connectivity. Indeed, all points which are classified in one group should form a connected component in the object space while features representing different objects should have spacing between them due to features which are segmented into other groups. On the other hand, in unstructured point clouds the connectivity requirement,

which is straightforward in image processing (since there, connectivity is defined in terms of the 4- or 8- neighbourhood on the pixel grid, see Dorninger and Nothegger 2007), needs a specific definition of neighbourhood (topology). This could be done, for example, by considering two points as neighbours if they are connected by an edge in the Delaunay triangulation; however, obtaining the Delaunay triangulation is computationally expensive (in particular, for large point clouds).

Hence, this research uses a criterion based on the definition of a bitmap located in the parameter domain of the shape (Fig. 1.5).

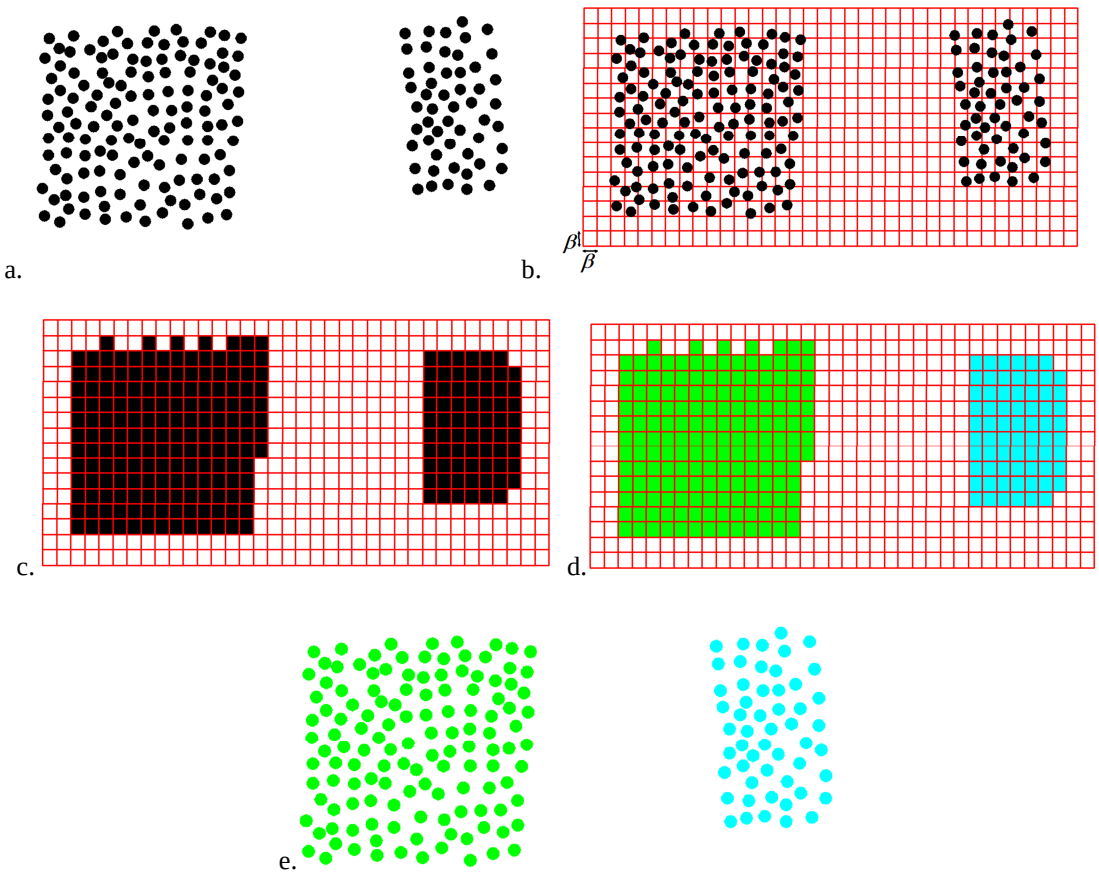


Fig. 1.5. Under-segmentation removal: (a) RANSAC detected cluster; (b-c) generation of the raster bitmap; (d) detection of connected components in the bitmap; and (e) correctly detected segments.

In particular, the bitmap is obtained by projecting each point along the plane normal direction. A cell in the bitmap is defined as occupied (and its value set to 1) if at least one point is projected into it. Otherwise it remains equal to 0. Ideally, the size β of the cells in the bitmap should correspond to the distance between neighbouring points in the data, i.e. the sampling resolution. However, in the case of point clouds acquired by TLS, data are irregularly sampled. For this reason the cell size β is chosen as the mean sampling resolution in the point cloud. Larger cells may sample in a too coarse way

the point cloud and do not overcome the problem. Indeed, β should be lower than the distance between objects lying in adjacent planes. On the other hand. a too small cell size may fit the limit of the scan ground sampling distance (GSD). A reasonable value for β can be easily set up by the used. In reality, segmentation results are not influenced so much, giving a reasonable cell size, from this parameter. Once the bitmap is setup, cells representing a connected component can be easily found. Then all points whose projection belongs to the same connected component are grouped into the same segment. In this phase some spurious segments might be found. Indeed, each segment whose area is far lower than the mean value is rejected.

Over-segmentation occurs when one feature in the point cloud (ψ_0) is segmented into several ones (ψ_1,$\psi_{2, ..., }$ ψ_n). It is generally associated with noise or irregularities in data. Indeed, many façades presents several irregularities like out of plumbs, variation of shapes and the like, that are not evaluated in the RANSAC inliers estimation. For example in the case of a façade presenting an out of plumb, the segmentation may result in subdividing a single façade wall into several objects. Obviously over-segmentation problems can be prevented increasing the RANSAC tolerances. However, this may lead to under-segmented parts which may be difficult to split again. For this reason, here restrictive RANSAC tolerances are selected causing a moderate over-segmentation.

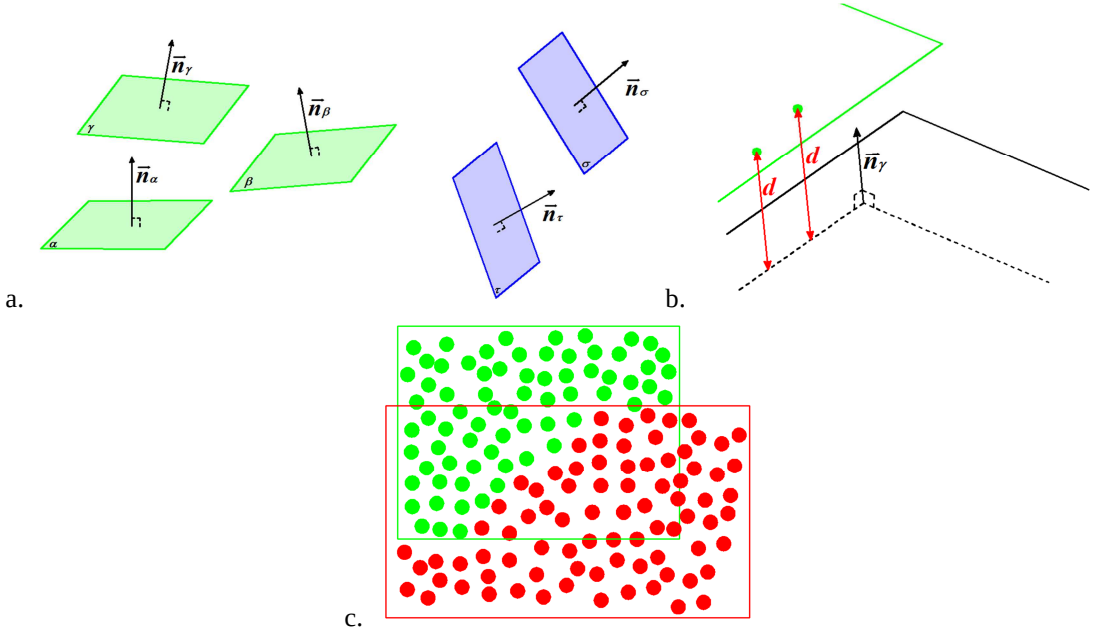


Fig. 1.6. Parameters evaluated for over-segmentation removal: (a) similarity of plane normal vector; (b) small perpendicular distance between points classified in different segments; and (c) an intersection zone between clusters.

This over-segmentation can be reduced by introducing topology properties for point clouds and planes. Indeed, group of points which belong to the same surface should share the following parameters (Fig. 1.6):

- similarity of plane normal vector;
- small perpendicular distance between points classified in different segments; and
- an intersection zone between them.

The methodology to overcome over-segmentation is based on two steps.

First, the whole group of detected segments are clustered by using the mean shift clustering algorithm (Comaniciu and Meer 2002). This exploits the normal vectors using as bandwidth the user-defined tolerance angle α. The main advantage of mean shift algorithm is that it is a non-parametric clustering technique which does not require prior knowledge of the number of clusters, and does not constrain the shape of the clusters. In particular, given n data points $x_i \in R^4$, the multivariate kernel density estimate using a radially symmetric kernel *K(x)*, is given by:

$$f_K = \frac{1}{nh^d}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{1.14}$$

where *h* (termed as the *bandwidth parameter*) defines the radius of kernel. The radially symmetric kernel is defined as:

$$K(x) = c_k k\left(\|x\|^2\right) \tag{1.15}$$

where c_k represents a normalization constant which assures *K(x)* integrates to 1. For a Gaussian kernel, as used in our case:

$$K(x) = \frac{1}{2\pi} e^{-c\|x\|^2} \tag{1.16}$$

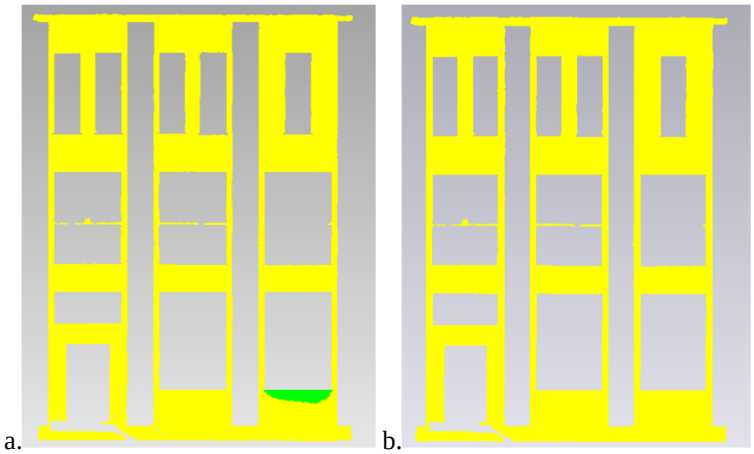
Taking the gradient of the density estimator Eq. (1.14) and some further algebraic manipulation yields:

$$\nabla f(x) = \frac{2c_{k,d}}{nh^{d+2}}\left[\sum_{i=1}^{n} g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)\right]\left[\frac{\sum_{i=1}^{n} x_i g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)} - x\right] \tag{1.17}$$

where $g(x) = -k'(x)$ denotes the derivative of the selected kernel profile. The first term is proportional to the density estimate at x (computed with the kernel $G=c_g g(||x||^2)$). The second term, is the *mean shift vector* (m), that points toward the direction of maximum increase in density and is proportional to the density gradient estimate at point x obtained with the kernel K. The mean shift procedure for a given point x_i is as follows:

1. compute the mean shift vector $m(x_i^t)$;
2. translate density estimation window: $x_i^{t+1} = x_i^t + m(x_i^t)$; and
3. iterate steps 1 and 2 until convergence (i.e., $\nabla f(x_i) = 0$).

During the clustering phase, since the plane normal direction is based on the sequence for choosing the points defining the plane, the check considers also the opposite normal direction.



a. b.

Fig. 1.7. Example of over-segmentation removal: (a) a small out of plumb (1 cm) is detected in the façade (green cluster); and (b) clusters are merged after clustering all planes.

Once extracted planes are clustered according to their normal, the perpendicular distance between points classified as representing different objects is evaluated, into each family of planes as detected by clustering of normals. If the mean value of these distances is lower than the user defined RANSAC threshold and the convex hulls of the point clusters intersects, they are recognized as a single object and merged together (Fig. 1.7).

Chapter 2
Geometric reconstruction

In Chapter 1 has been described the segmentation of the building façade into planar features. This Chapter presents the developed automatic procedure for geometric reconstruction and vectorization of the building façade starting from the features detected during the segmentation stage. The generation of a digital vector model of the façade also requires a precise description of edges and breaklines, which are not achieved during segmentation of planar features. For this reason, a procedure to detect façade breaklines along with their post-processing and smoothing is described.

Point clouds obtained from TLS can be directly used to reconstruct façade models by means of simple triangulation of the acquired data. However, these triangulated irregular networks (TINs) models are generally computed on the mere basis of an irregular distributed point cloud. Therefore, they do only implicitly store break-line information. The quality of the break-line description within these models depends on the original point sampling interval and on the size of triangles. This representation can be adequate for '*free form objects*', i.e., object whose form cannot be parameterized by means of simple geometric shapes (like some decorations, bas-reliefs, sculptures, etc.), where discontinuities are smooth. However, in the case of objects that are made up of several 3D regular solid shapes (like mechanical parts, building façades, etc.), transitions between different elements is sharp and explicit description of these discontinuities is essential for generating high quality models. In addition, TIN models obtained by conventional triangulation methods generally results in a tremendous number of triangles. This is because TLS systems acquire data at a given sampling rate and therefore even flat areas are sampled with high point density. The large size of these models may cause problems in handling such huge data sets.

For these reasons, when dealing with building façades their modelling is performed by individuating in the point cloud the basic shapes constituting the façade and fitting polygons to them. At the end a 3D vector description of the façade is obtained. Indeed, a vector model exploiting the primitive shapes may achieve the desired effect of generating a lighter model compared with the original point cloud or TIN model. In

addition vector models give the possibility to explicitly describe 3D breaklines. Indeed, a 'break-line' can be defined as the intersection of two surfaces, each surface interpolating the points on either side of the break-line itself.

Until today, façade vector models have been usually generated with a manual approach. In standard reverse engineering, generally, the point cloud is first analysed by identifying the basic shapes constituting the façade. Then breaklines are digitized by the measurement of relevant points along edges. The selection of the important lines and the way the line are discretised is up to the individual human interpretation. Then, starting from the detected edges the geometry of a façade element is recovered. However, this requires a largely time-consuming manual work performed by skilled operators. To speed up the façade modelling, in the field of thermal retrofitting of existing buildings, an automatic modelling approach is presented in this Chapter. Obtained models can then be used in conjunction with or as an alternative to Computer-Aided-Design (CAD).

In the developed methodology automatic modelling is obtained thought a multi-step approach. First, the original point cloud is segmented to identify all planar clusters constituting the building façade as described in Chapter 1. Then, considering each detected object, the segment edges are identified giving a first rough approximation of façade breaklines. The extracted edges are random noise affected and some blunders may still present and thus cannot be directly used as final product. For this reason a further processing is need to delete gross errors and smooth edges. Here some architectural priors are added. Only after these steps, the collected information can be used for defining the boundaries of each element of the façade. Then, the topological relations between segmented regions are enforced evaluating distance between estimates of their edges. Finally the segments, their boundaries, the intersections lines and the topological relations, available at this point of the process chain, are used to generate the 3D façade model.

2.1. State of the art

A first key element for geometric reconstruction of building façade from point cloud is a prior modelling of breaklines, based on the original unclassified TLS points. The breaklines problems was firstly analysed in the ALS field for Digital Terrain Model (DTM) production (Vosselman and Maas 2010). In this field, an edge in a point cloud is defined by those points where changes in the local surface properties exceed a given threshold. They can be divided in: (i) step discontinuity (a 'jump' occurring in the data), (ii) slope discontinuity and (iii) curvature discontinuity (one of the principal curvature changes locally). The local surface properties mostly used in the edge base methods are surface normals, gradients, principal curvatures, or higher order derivatives. Edge detection methods look for abrupt changes, therefore they are very

sensitive to noise in the range data. From the methodological point of view, the edge detection can be carried out by (at least) thee methods, classified as (i) *analytical direct*, (ii) *analytical indirect* and (iii) *geometrical by decimation* (Beinat et al. 2007). The algorithms involving surface interpolations by any analytical direct function belong to the first class (i). These have the common property to provide one or more local numerical values directly revealing singularities in the laser point cloud (Briese 2004, Alshawabkeh et al. 2006). The analytical indirect methods (ii) start with the suitable estimation of continuous surfaces in order to interpolate in the best way the laser data. Only in a second step, the edges are detected by considering the space intersection of such surfaces or simply analysing to which surface each point has been assigned. Finally, the geometrical decimation (iii) regards the optimization of the mesh, coming out from a triangulation process, and does not involve the coordinate points; nevertheless, while the edges are strongly correlated with the result of vertex decimation, this approach does not succeed in the slope discontinuity detection.

In ALS domain, a first break-line detection approach, starting from an ALS range image leading to smooth vector breaklines, was presented by Brügelmann (2000). Within this method the first and essential step is the extraction of edge pixels in the range image. This is performed with the help of a raster-based method using a hypothesis testing method presented by Förstner (1998). This method of second derivatives and hypothesis testing treats break-line detection in range images on the same principle as finding edges in intensity images. The basic idea is that edge pixels are borders of homogeneous regions. Therefore the following two basic properties for edge pixels should be valid:

- the homogeneity measure on edge pixels should differ significantly from the one determined in homogeneous regions; and
- the homogeneity measure should be locally maximum across the edge.

In order to derive this homogeneity measure for range images the gradient image is calculated. Whereas for intensity images the squared gradient magnitude is often used as homogeneity measure, a multi-channel extension is applied to the gradient image. This leads to a homogeneity measure called quadratic variation which is closely related to the second derivatives based on the two principal curvatures and describing the maximum and minimum normal curvature at every pixel (Émery and Meyer 1989). The results of this process are pixels marked with the attribute ‘edge pixel’. To generate the 2D position of the break-line within the broad regions of edge pixels a thinning operation is applied after a non-maximal-suppression, taking into account the direction of the maximum curvature. Then, a further raster to vector conversion allows generating 2D vector breaklines. Within this step some smoothing using a 2D cubic polynomial spline method is performed in order to perform a certain

generalisation and elimination of 'zig-zag' effects caused by the original raster data structure. Next to this interesting approach, a lot of further algorithms within the area of edge-based segmentation techniques based on raster ALS surface models were developed. All have in common to use image processing techniques in order to extract 2D break-line pixels, which have to be refined in further steps.

In TLS domain Boulaassal et al. (2009) presented a contour extraction algorithm for building façades. Once having performed façade segmentation and detected planar clusters in a façade the extraction of their contour is carried out. The main idea exploited in this algorithm is based on the hypothesis stipulating that contour points belong to the long sides of Delaunay triangles for detected clusters. This algorithm proved to be able to detect contour points. However, due to noise in the dataset and the random nature of points acquired by TLS systems, the derived contours present a very irregular and jagged shape.

Becker and Haala (2007) presented a procedure for breaklines extraction from point cloud of building façades combining two different phases. In a first step a cell decomposition of the façade is performed by identifying contour points using a rasterization of the façade similar to Brügelmann (2000). Then, façade edges are refined by means of an edge matching procedure combining photos and TLS data.

A procedure for breaklines detection combining multi-image matching and TLS data is also presented in Nex and Rinaudo (2009). In this case, edges are extracted from a set of oriented images and their 3D position is retrieved trough a multi-image matching algorithm. Then, after edge denoising and smoothening, matched edges are checked within the TLS data.

Concerning façade modelling, a few automatic reconstruction methods have been proposed, while some commercial packages such as Cyclone®, Phidias® and CC-modeler® provide only semi-automated reconstruction functionalities. A common aspect of all these methods is that while such systems works well with the data, they usually require the user to set up several control parameters. This kind of parameterization is very common in fully automatics methods but it turns out to be also an under-estimated obstacle, since the search for proper parameters can be very time consuming.

Automatic façade modelling is generally performed with a bottom-up approach, i.e. they directly extract features such as points or edges from multi-source data and then try to aggregate them into 3D models.

In earlier works, Stamos and Allen (2000) developed a system for reconstruction of buildings from range scans combined with sets of unordered photographs. This method is based on fitting planar polygons into pre-clustered point clouds.

Bauer et al. (2003) also proposed an approach for the detection and partition of planar structures in dense 3D point clouds of façades, like polygonal models, with a considerably lower complexity than the original data.

Frueh et al. (2005) presented an early attempt which integrated TLS data and digital image to generate façade meshes. Laser point clouds are first used to generate depth images, and foreground (occlusion objects) and background (building façades) are distinguished by histogram analysis. TIN mesh models are generated for background points and textured with selected photos. The texture holes caused by occlusions are filled with texture from similar areas. This approach achieves very high automation and realistic results, but no simple geometric shapes (such as polygon) are reconstructed. In addition, the generated huge amount of triangles may result in slow visualization.

Pu and Vosselman (2006) presented an automatic approach to extract building façade features from a terrestrial point cloud. The method first defines several important building features. Then the point cloud is segmented into planar segments. Finally each segment is compared with building features. However, the feature extraction method does not work for windows.

Becker and Haala (2007) reconstructed polyhedron models by integrating TLS data and digital photos. To achieve registration between different datasets, intensity images generated from laser point cloud are used for bridging. This turns out the problem into image to image registration, which is solved by the SIFT algorithm (Lowe 1999). Assuming that no laser points are available from windows, laser points on window edges are extracted. Edges extracted by Sobel filter from digital photos are then used to refine the window edges. The resulting model contains windows frames and crossbar. Problems may arise when the laser point density is too low to reliably match digital images. Moreover, the shape of detectable windows is limited to rectangles.

Recently, Venagas et al. (2012) proposed an approach for the reconstruction of buildings from 3D point clouds with the assumption of Manhattan World Building geometry (or ‘Legoland’ as used in Förstner 2010). This system detects and classifies features in the data and organizes them into a connected set of clusters from which volumetric model description is extracted. However, derived models present a very low geometric level-of-detail.

2.2. Automated façade modelling

In this section, a detailed description of the implemented automated façade modelling algorithm is given (Fig. 2.1).

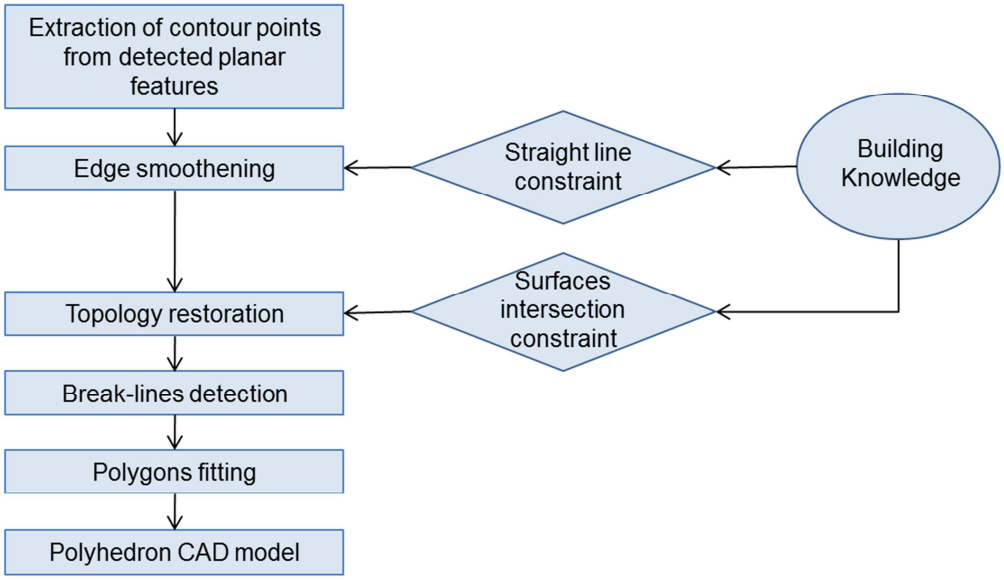


Fig. 2.1. Workflow of the developed automated façade modelling process.

In a first step, once planar clusters constituting the façade are detected by the segmentation approach presented in Chapter 1, the extraction of their contours is carried out. To achieve this task the algorithm described in Section 2.2.1 has been developed.

However, contours extracted in this way are affected by random noise and some blunders might still be found. Blunders are firstly removed from the extracted edges by evaluating their size. In particular, points forming too small contours are evaluated as small holes in the point cloud and filled. Due to the presence of noise and the acquisition scheme of laser scanners, edges and corners may present an irregular and jagged shape. Indeed, in these locations the spatial resolution of the point cloud is not enough to give a complete description of the geometry. On the other hand for the considered building types, straight lines prevail. For this reason, the noisy edges are split into basic elements (linear and curved elements) and each of these is smoothed and eased trough automatic transformation into lines and B-splines. Finally, the basic elements are recollected in a unique smoothed edge.

Due to the RANSAC tolerances in the segmentation steps, topological relations between surfaces are lost. In order to restore correct topology in data segment, adjacency is evaluated. In this way, starting from the façade segments, their boundaries and topological relations, the intersection lines, i.e. breaklines, can be

determined. Finally, the geometric reconstruction of the façade, represented by means of polygon elements, is exported in CAD standard format in order to give a good preliminary data for the graphic visualization of the survey and for an evaluation of the achieved results.

2.2.1. Extraction of geometric edges

The first step of the presented approach for automatic façade modelling is the extraction of contour points for each detected planar segment. Before further processing, a preliminary step is performed by defining a new coordinate system for each planar cluster. For this purpose, a Principal Component Analysis (PCA) is calculated based on the points of the planar segment. The coefficients of the first two principal components define vectors that form an orthogonal basis for the plane. The third one is orthogonal to the first two, and its coefficients define the normal vector of the plane. The original coordinates ($X_{or.}$, $Y_{or.}$, $Z_{or.}$) are then transformed in this new local space aligned along the principal component directions (X_{new}, Y_{new}, Z_{new}). In particular, the component Z_{new} which is approximately directed along the normal direction to the planar cluster may be considered as negligible.

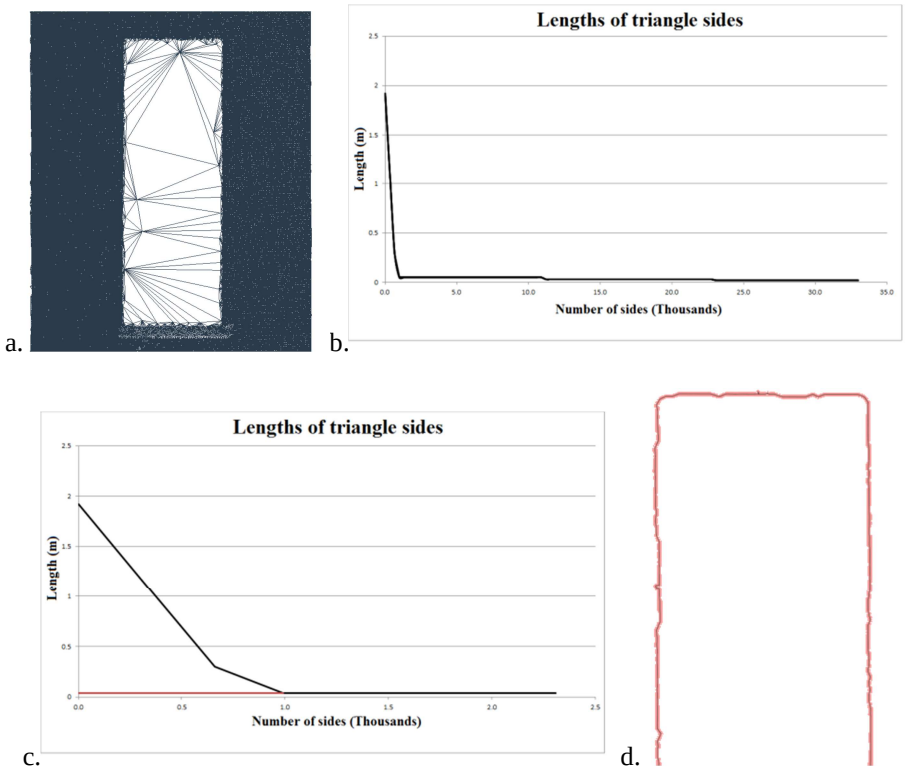


Fig. 2.2. Delaunay triangulation of a point cloud portion of a building façade (a); curve of lengths of triangle sides (b) with a zoom (c) on the curvature changing point; and extracted contours (d).

At the end, when contour points are extracted, they will be transferred again in the original reference system. The presented algorithm for contour point detection relies

on the observation (Boulaassal et al. 2007) that points belonging to a cluster contour can be individuated after Delaunay triangulation of planar segments. In particular, contour points are defined as extremities of long sides of Delaunay triangles. For this reason, according to the implementation in Boulaassal et al. (2007) the lengths of all triangle sides are ordered in an ascending way. Two classes of side length can be distinguished (Fig. 2.2).

The first one contains short sides that are located on the horizontal part of the curve. The second one contains long sides that are represented by the vertical branch. This is given by the fact that the number of points belonging to the contour is negligible compared to the total number of points. The threshold separating these two classes is typically determined around the point *P* where the curvature is changing. In particular the numerical first derivative for each point location in the graph is calculated. The threshold value is then fixed in correspondence of a sudden jump of the derivative toward almost infinite values.

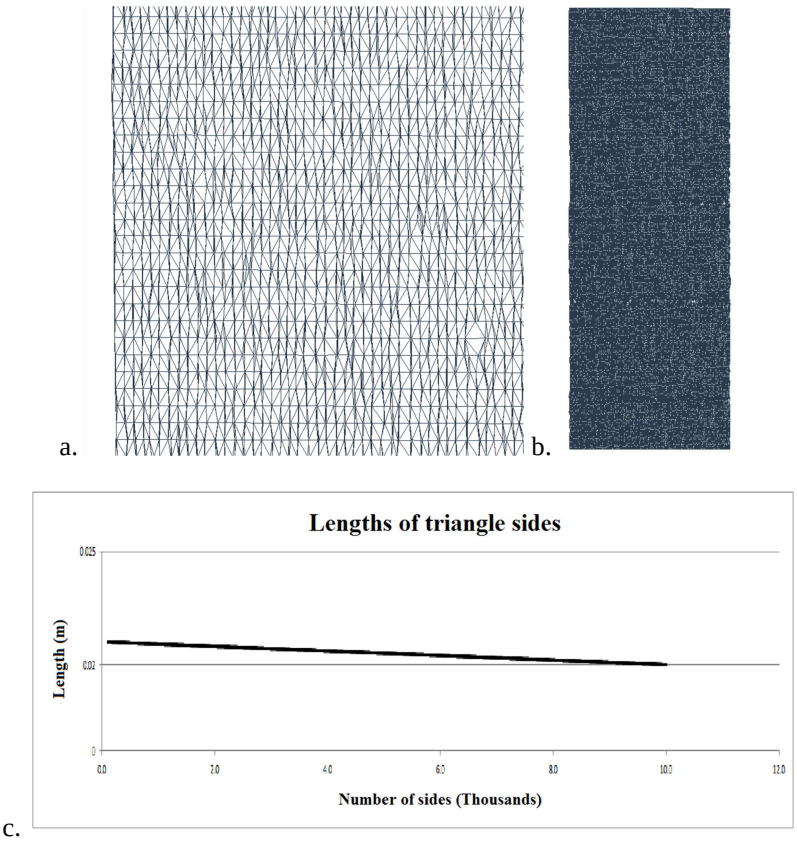


Fig. 2.3. Limitations of the Boulaassal et al. (2007) contour extraction algorithm. Triangle sides at cluster external contours generally do not exhibit a significant increase with respect to the other ones (a); and for compact elements (b) the identification of a point where curvature changes is not evident (c).

Even if this method proved to work quite well in the extraction of contours in previous applications (Alshawa et al. 2009, Boulaassal et al. 2007, Martinez et al. 2012), some limitations were found with the building addressed in this research:

- even if it is true that long sides of Delaunay triangulation generally determines the presence of a contour, this is mainly true for the identification of holes in a cluster, e.g., windows contours inside main façade segment. However, triangle sides at cluster external contours generally do not exhibit a significant increase with respect to the other ones, e.g., external wall contour sides generally have the same size of internal triangles' sides (Fig. 2.3a);
- due to the observation made at the previous point the selection of a proper threshold between long and short sides is not a simple task. Indeed, while the point *P*, representing the curvature change in the graph, can be properly used to identify the presence of points forming a contour hole in the segment, it is not appropriate for identifying external contours. On the other hand the selection of a lower threshold may generate spurious results;
- for some specific segments, e.g. compact elements, the identification of a point where curvature changes is not evident (Fig. 2.3b-c); and
- contours obtained in this way are quite irregular and typically show a characteristic jagged shape (Fig. 2.2d).

For these reasons a new approach for automated contour identification was developed (Fig. 2.4). Once the new coordinates are calculated (X_{new} , Y_{new}) a bi-dimensional Delaunay triangulation is performed. By identifying the threshold between long and short side triangles, all the ones having a side exceeding the threshold are removed. In this way, all Delaunay triangles that have been wrongly generated in correspondence of holes are removed while external contour sides are generally kept.

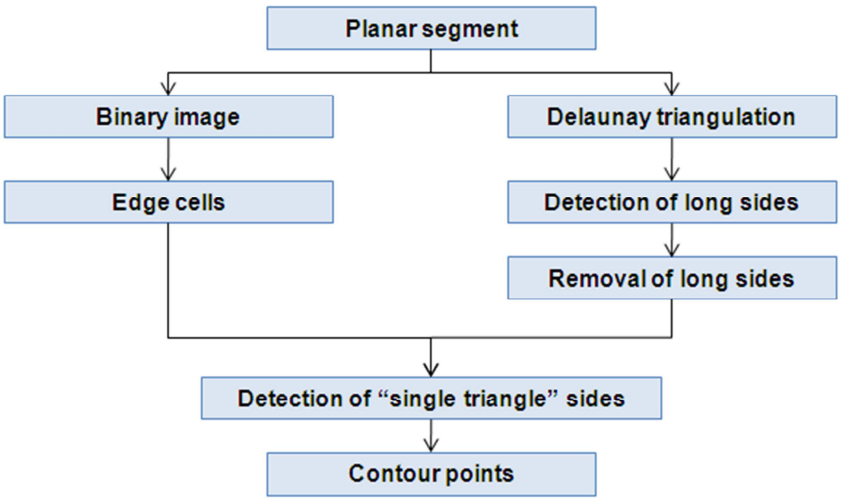


Fig. 2.4. Workflow of the developed contour extraction process.

By analysing the obtained triangulated surfaces, contour points are then derived by identifying triangle sides that belong to a single Delaunay triangle. Indeed, points in the inner part of the planar clusters belong to triangle sides shared by two triangles. On the other hand contour points generate triangle sides belonging to a single Delaunay triangle (Fig. 2.5).

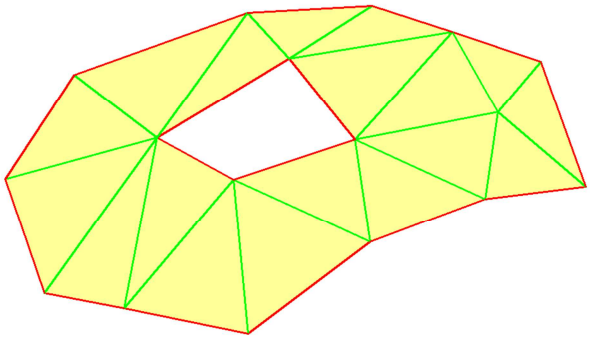
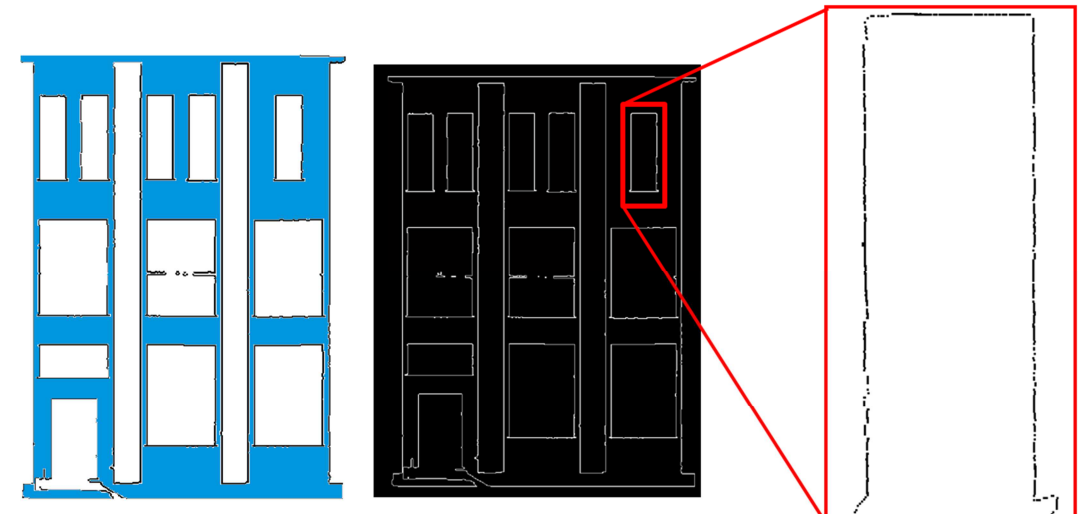


Fig. 2.5. Detection of contour points principle: contours lines (red) belong to one triangle only while others belong to a couple of triangles.

The search of these sides can result in a significant computational effort. For this reason, a simple way is implemented to quickly identify and ignore the ones that are not close to the boundary. Firstly the plane is discretized into cells of size β x β. Then a binary image for the cluster is generated. In this image white cells represent elements where TLS data are available, while black cells are grid elements with no data.



a. b.

Fig. 2.6. Detection of contour points: (a) triangulated segment model; (b) contours raster map and detected contour points.

Points lying inside cells whose 8-neighbors all contain data points are pruned out. In this way a set of edge cells is obtained and the search of contour points is performed only in these cells (Fig. 2.6).

The optimal cell size β, as previously discussed, is the mean sampling resolution in the point cloud. Of course, accuracy of the extracted edge cells is influenced by the cell size. However, a coarse rasterization does not affect the accuracy of the extracted contour. Indeed, contour points are computed considering the original data while the generation of a raster map is only used to speed up the identification of contour point by eliminating a series of non-contour sides.

2.2.2. Edge smoothing

As previously observed contour points detected with the presented approach define quite irregular and jagged contours showing a characteristic 'saw-tooth' shape. In fact, these edges have an irregular and waved shape due to the noise and random measurement errors. However, this is in contrast with the characteristic façade geometry where straight lines are predominant. For this reason, the achieved edges cannot be directly used for façade modelling and smoothing is needed to define a regular shape of the object. This process has to consider the different typologies of edges that compose the object. In particular, edges can be usually split into different basic entities (linear or curved parts). Then, each of these entities can be simplified in lines and curves. The line and the curve equations must be fitted in the best possible way using the dominant points information, while the whole edge can be finally reconstructed by linking these entities together. In this paragraph the automated reconstruction of edges is presented.

One of the most common edge simplification methods is the Douglas-Peucker algorithm (Douglas and Peucker 1973). It is very simple to implement and it works for every edge dimension, once it only relies on the distance between points and lines. Its basic rule is that the approximation must contain a subset of the original data points and all the original data points must lie within a certain predefined distance to the approximation (Wu-Shin and Gonzales-Marquez 2003). The Douglas-Peucker algorithm has a hierarchical structure starting with a crude initial guess, namely the single edge e joining the first and last vertices of the polyline. Then, the remaining vertices are tested for closeness to that edge. If there are vertices further than a specified tolerance ε away from the edge, then the vertex farthest from it is added to the previously simplified polyline. This creates a new approximation for the original polyline. Using recursion, this process continues for each edge until all vertices of the original polyline are within ε. This algorithm has $O(mn)$ worst case time and $O(n \log n)$ expected time, where n is the number of input vertices and m is the number of the segments of the simplified polyline.

Other simplification algorithms extracts the points that lie on longer line segments by means of a sequential approach, looking for positions where the slopes of two consecutive edges with similar slopes may be grouped into one line segment (Sampath and Shan 2007). In alternative, the 'sleeve fitting' algorithm can simplify into a line the points that are inside a sector bound defined by an angle (Zhao and Saalfeld 1997).

All these simplification algorithms have been successfully used in aerial applications or for map digitalization. They essentially reduce the number of points per edge, smooth data into a polyline composed by linear elements (i.e., Douglas-Peucker and the Shao-Gonzales methods do that), or they are able to divide an edge in linear elements (like 'sleeve fitting' method) and regularize them in order to extract simple building roofs.

Unfortunately this kind of solution is not sufficient in the edge reconstruction of the architectonic applications as they have usually generic shapes, different length and complexity. Anyway, in the great part of residential building façade in the period 1950 - 1975, long edges that describe the main geometry of the building, can be eased in a sequence of lines and curves linked together. In these conditions, the linear and curved features have to be recognized and then a specific simplification algorithm has to be applied. In order to do that, a specifically designed smoothing algorithm has been developed (Fig. 2.7).

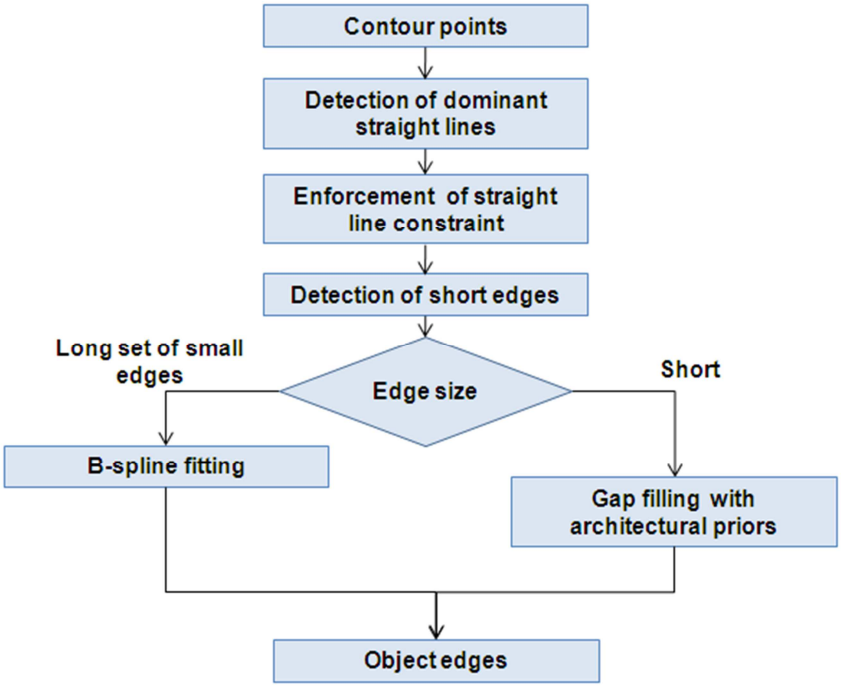
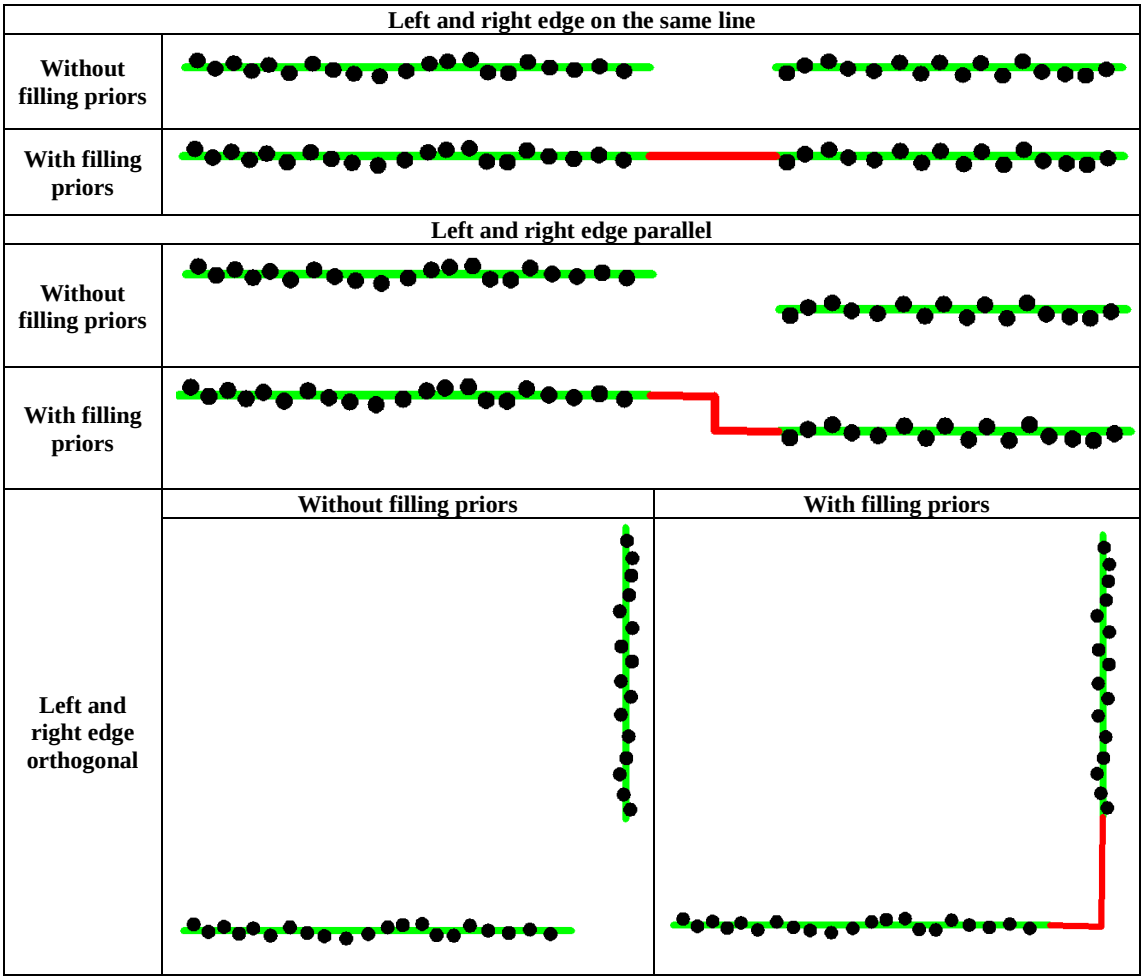


Fig. 2.7. Workflow of the developed edge smoothing and regularization process.

As previously noticed, an important aspect of the façades targeted in this research is given by the fact that the straight lines are predominant in façade elements. In addition, these straight lines generally intersect in an orthogonal way. For this reason, once the contour points are determined, the dominant edge directions are identified by using a sequential RANSAC implementation aimed at extracting linear features similar to the one presented in Chapter 1 for the detection of planes. Once dominant lines are detected, inliers points are removed and substituted with a straight line. The remaining contour edges are then evaluated. Indeed, sometimes, small occlusions on a façade or segmentation errors cause irregular edges on the generated outline. These irregular edges should be removed by observing that they form short segments, which result in a gap on the outline. If the left long edge (to the gap) and right long edge belong to the same line, the gap is just filled by connecting a line segment. If the two edges are parallel, a line segment which is perpendicular to both is generated, and the two edges are extended to reach the perpendicular segment. Finally, in the case the two initial edges are orthogonal they are just extended or shortened until they intersect at a point to fill the gap (Tab. 2.1).



Tab. 2.1. Filling of boundaries for different arrangements of edges.

In some cases instead of a single small edge a long set of small edges can be found. This is generally due to the presence of a curved feature. In the implemented procedure instead of detecting the curved shape first and then apply the corresponding fitting algorithm, a B-spline curve representation is adopted as uniform mathematical model. A B-spline curve of degree p is defined by n+1 control points $P_0,\ldots,P_n$ and a knot vector of m+1 knots:

$$U = \{u_0, u_1, ..., u_m\} \tag{2.1}$$

where U is a non-decreasing sequence with $u_i \in [0,1]$ and n, m and p must satisfy:

$$p \equiv m - n - 1 \tag{2.2}$$

The B-spline parametric curve function is of the form:

$$C_p(u) = \sum_{i=0}^{n} P_i N_{i,p}(u) \tag{2.3}$$

$N_{i,p}\,(u)$ is the basis function of the B-splines, defined by:

$$N_{i,0}(u) = \begin{cases} 1 & if \quad u_i \le u \le u_{i+1} \\ 0 & otherwise \end{cases}$$
$$N_{i,p}(u) = \frac{u - u_i}{u_{i+p} - u_i} N_{i,p-1}(u) + \frac{u_{i+p+1} - u}{u_{i+p+1} - u_{i+1}} N_{i+1,p-1}(u) \tag{2.4}$$

The objective function for a Least Squares curve fitting with a spline function of degree *k* is:

$$U = \sum_x \left\{ W(x) \left[y(x) - \sum_i P_i N_{i,p}(x) \right] \right\}^2 \tag{2.5}$$

where $W(x)$ is a weight and $y(x)$ is the input value at x. The coefficients P_i are the parameters to be determined. The knot values may be fixed or they too may be treated as parameters. In the developed implementation weights $W(x)$ are fixed equal to 1

because no a priori information about noise in the data are available and knots are treated as parameters. For determining the number of knots to use and where they should be placed, the 'curvature-based' square distance minimization proposed by Wang et al. (2006) was used.

Once having individuated dominant lines and possibly curved elements are fitted with B-spline the whole edge is reconstructed by linking these entities together. All computations are performed in the local principal component space. Finally, reconstructed edges are transposed in the original datum (Fig. 2.8).

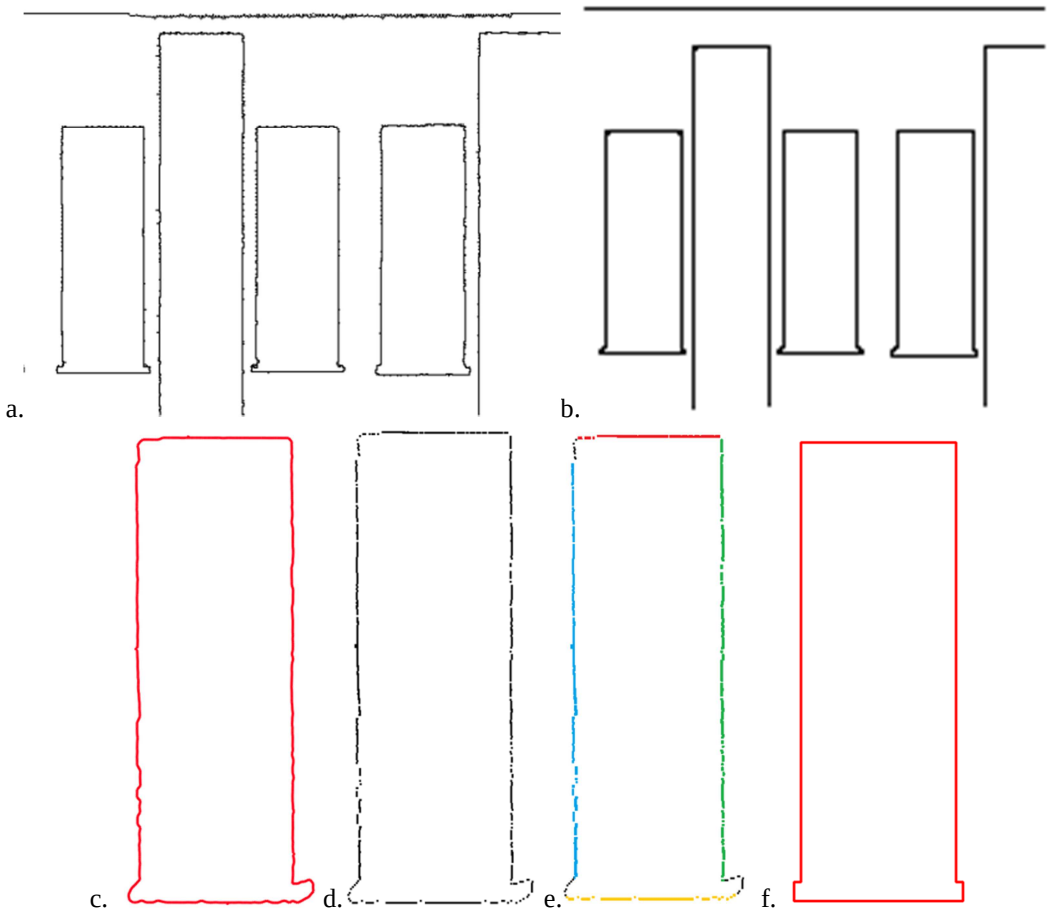


Fig. 2.8. Boundary edges before (a-c) and after (b-f) the straight line constraint application. Example of dominant line detection for a window (d): dominant lines are represented with different colours (e).

2.2.3. Reconstruction of façade topology and break-line extraction

Because of RANSAC is used and a threshold tolerance was fixed in the segmentation phase, some topological relations like intersection between surfaces are lost. Indeed, in the case of intersecting surfaces the points lying on the intersection (that belong to both surfaces) are assigned only to one cluster (generally to the larger one). Due to this shortcoming, segments adjacency is lost. In addition, detected edges for adjacent

surfaces do not intersect. These parts are mostly located in the intersection zones between features, for example between sidewalls and walls, between sidewalls and doors of windows. On the other hand these intersections between features exist in reality. To restore them a surface intersection constraint is added into modelling pipeline.

The first step is the detection of adjacent segments. Two segments are considered adjacent if at least one pair of line segments from their rough 3D boundaries is close to one another. In particular, segments are considered adjacent if the distance between them is lower than the RANSAC threshold used for plane detection. Once two adjacent segments are detected, surface intersection is enforced. In particular, feature's outline edges which are close to the other feature are replaced, in both adjacent segments, with the intersection line of the planes that the two features belong to (Fig 2.9).

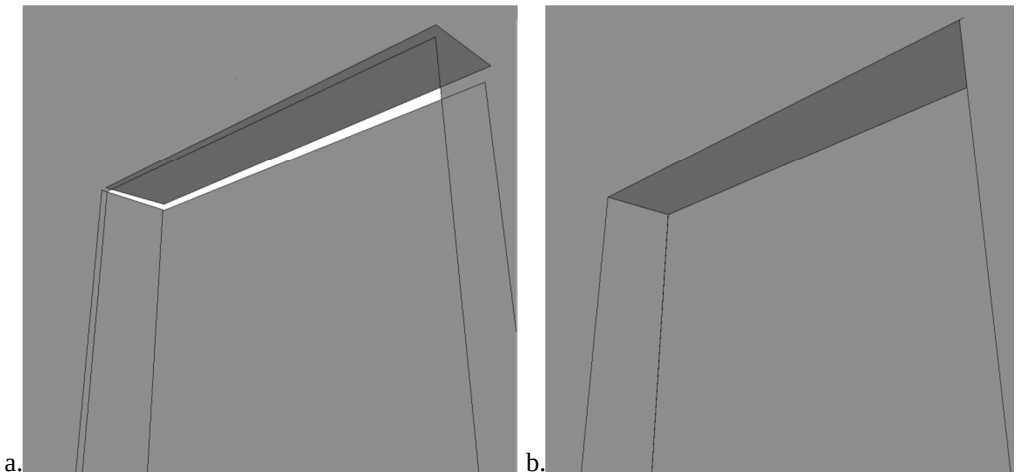


Fig. 2.9. Plane connection before (a) and after (b) the surface intersection constraint.

Surface intersection constraint, not only re-establish the topology between objects but also increase the accuracy of detected breaklines. Indeed, by means of surface intersection constraints, breaklines are calculated as the intersection of planes which are estimated from a large set of points. While edge calculation from contour points may be more affected by noise in the point cloud. In addition, as previously discussed, the accuracy of laser scanning measurements in correspondence of edges is generally lower than the one on smooth surfaces (see also Soudarissanane et al. 2011).

2.2.4. Geometric model export in CAD environment

At this point of the processing chain the segments, their boundaries and intersections are available. Thus, the 3D building model can be reconstructed now.

The geometry of a reconstructed model can be described in there different ways: (i) *spatial enumeration*, (ii) *boundary representation*, and (iii) *constructive solid geometry (CSG).* In *spatial enumeration* models (Frueh et al. 2005) space is described as a regular array of cells (usually cubes). Each cell is termed as *voxel* and a 3D object is represented as a list of occupied voxels. This king of representation generally requires high memory consumption. In addition, model resolution is limited to size and shape of voxels. In the *boundary representation* a closed 2D surface define a 3D object. Object boundaries can be defined in two ways: (i) in a primitive-based way where a collection of primitives form the boundary (e.g., a set of polygons) or (ii) in a free form-based approach where boundaries are defined by means of splines, parametric surfaces or implicit forms. In the boundary representation models (Pu and Vosselman 2009, Becker 2009, Tian et al. 2009) the outlines of planar features are extracted as polyhedron, so the model sizes are much smaller than the spatial enumerations models. The data reduction of boundary representation also leads to much faster visualization. With *constructive solid geometry (CSG),* a building model is composed by some fixed primitives arranged by means of Boolean operators (union, difference, and intersection). The CSG models are always watertight, and require even fewer parameters than boundary representation. However, it is hard to represent complex objects with CSG when they can be hardly decomposed into simple shapes. Besides, the heavy conversion of a CSG model to the boundary representation is always necessary for visualization purposes. In addition, representation is not unique. Only a few methods (Haala and Brenner 1999, Suveg and Vosselman 2004) are based on CSG building models.

Interoperability is an important need in order to make the derived model relay available to operators (Brumana et al. 2013). Indeed, the presented workflow has been implemented in *Matworks Matlab*® code. This language allowed all the workflow steps to be accomplished, but it did not succeed in production of suitable drawing for the end-users (e.g., architects and civil engineers). For this reason, the obtained faced model is exported by using the *Drawing eXchange Format* (DXF) format (McHenry and Bajcsy 2008). This format enables conversion of data into a format compatible to a great number of design and CAD software packages.

The data format of a DXF is called a ‘tagged data’ format, which means that each data element in the file is preceded by an integer number that is called ‘group code’. A group code value indicates what type of data element follows. In the presented approach, geometric reconstruction can be seen as a process of polygon fitting. For this reason, the boundary representation is the closest to the developed method. In addition, several standards, as CityGML (Gröger and Plümer 2012.), rely on boundary representation for geometry definition. In particular, in the DXF file the building boundaries are represented by a set of polygons defined as region objects.

In some cases also a polygon mesh representation can be useful. For this reason the obtained model is also exported into Polygon File Format (PLY) file format. In this case a constrained Delaunay triangulation, where façade breaklines represent constraints, is used. The derived model presents the advantage of maintaining the detected breaklines. In addition the size of the final model is significantly reduced with respect the triangularization of the original point cloud.

Chapter 3
Semantic model enrichment

This chapter presents an automated procedure for labelling several important building features (walls, doors, roof, protrusions, etc.) starting from the planar segments detected by the point-cloud segmentation strategy presented in Chapter 1. Extracted semantic information are then associated with their geometric properties (e.g., size and shape) and spatial relationships. These items are then exploited to produce a compact, semantically enriched 3D model that contains the geometry and identity information that substantially are required to feed a Building Information Model (BIM).

In the last years, in the Architecture, Engineering, and Construction (AEC) domain, the importance of semantically rich 3D models of buildings has been continuously growing. In particular, their importance is not limited only to design and construction phases, but also extended to throughout the entire building serviceability and the facility management phase. Such models are generally known as BIMs. According to the United States National Building Information Modelling Standard (NIBS, 2007) a 'Building Information Modelling (BIM) is a digital representation of physical and functional characteristics of a facility. A BIM is a shared knowledge resource for information about a facility forming a reliable basis for decisions during its life-cycle; defined as existing from earliest conception to demolition.' As can be clearly seen from this definition (in contrast to traditional building design approaches) a BIM is much more than a simple geometric model. It manages not only graphics, but also information that allows the automatic generation of drawings and reports, design analysis, schedule simulation, facilities management, and more. Last but not least it enables the building team to make better-informed decisions. In particular, the BIM logic is based not only on a 3D building geometry but also on semantic and descriptive information.

In particular, the process of converting point-cloud data into a BIM is known as 'scan-to-BIM' process. Geometric surfaces or volumetric primitives are fitted to the 3D point cloud to model walls, floors, ceilings, columns, beams, and other architectural elements of interest. The modelled primitives are annotated with their identity labels (e.g., 'wall') and metadata, such as the surface material (e.g., 'concrete'). Spatial and

functional relationships between nearby structures and spaces are also established. Currently, the 'scan-to-BIM' process is primarily a labour-intensive manual operation (Anil et al. 2011). Researchers in the AEC domain recognized the need for automation tools to speed up such processing (Brilakis et al. 2010, Tang et al 2010). However, despite the progress in sensor development, autonomous recognition of the building structures remains the most challenging task among the whole reconstruction process.

3.1. State of the art

A core task of as-built BIM construction is *object recognition*. It can be defined as the process of labelling a set of data points or geometric primitives extracted from the data with a named object or object class (Xiong et al. 2013). Whereas the modelling task would find a set of points to be a vertical plane, the recognition task would label that plane as being a wall. Object recognition algorithms may be divided into two categories: recognizing object instances of an exact shape (e.g., instances of a specific type of beam), or recognizing classes of objects, where the shape may vary among instances from the class (e.g., windows that may vary in height, width, etc.). In the case of façade objects because of the great variety of building styles, the classification problem has to be formulated in the term of recognizing classes of objects.

The most common approach uses 'global-shape' descriptors, which are less discriminative than 'semi-local' descriptors used for the recognition of simpler object instances. In those cases, the recognition process can be divided in two main steps. First, in an off-line process, global shape descriptors are computed for each object to be detected. These descriptors are stored in a database that is designed to facilitate rapid lookup of descriptors based on the similarity to a query descriptor. Second, at runtime, the recognition is performed in a scene where instances of the target object are to be detected, or possibly with a pre-segmented data instance to be recognized. Shape descriptors are computed at locations in the query scene, either randomly or at salient points, and the most similar descriptors in the model database are retrieved. In particular, the query object is matched against examples from the entire class. The descriptor similarity measure is designed so that similar shaped objects will result in similar descriptors. As a result, the closest matching descriptors from the database will give an indication of the object class. A complete review presenting a large number of global descriptor methods can be found in Shilane et al. (2004). One disadvantage of the global descriptor approach is that it is unable to handle dataset suffering from occlusions or clutters, both of which are commons in laser-scanning data. In particular, the back sides and bottoms of most objects are not visualized. Another limiting aspect is the long out-of-core training phase for the descriptors to cover various instances in a class. Due to these limits of existing methods, the recognition of BIM-specific components such as walls, windows, and doors, is still in

its early stage. Methods in this category typically perform an initial shape-based segmentation of the scene into planar regions, for example, and then use features derived from the segments to recognize objects. This approach is exemplified by Rusu et al. (2008) who use heuristics to detect walls, floors, ceilings, and cabinets in a kitchen environment.

Another opportunity is associated with introducing *context information* into the recognition phase. In particular, some researchers proposed leveraging the spatial relationships between objects or geometric primitives to reduce the ambiguity of recognition results. Such approaches generate semantic labels of geometric primitives, and test the validities of these labels with a spatial relationship knowledge base. Usually, such a knowledge model is represented by a semantic net (Nüchter and Hertzberg 2008). For example, it may specify the relationships between entities such as 'floors are orthogonal to walls and doors and parallel to ceilings'. During the recognition process, if a surface is recognized as 'floor,' then the algorithm will identify that the valid semantic labels of a surface orthogonal to it can only be 'wall' or 'door,' but not 'ceiling,' thereby reducing the search space (Cantzler 2003). Such validity checking approaches provide ways to integrate domain knowledge into the object recognition process making the recognition more robust. An automatic extraction and classification algorithm of building features is presented in Pu and Vosselman (2009). In this case, objects derived from segmentation are classified according to some parameters like their position and extent. Similarly, Luo and Sohn (2010) presented an approach based on classification of façade elements according to a different set of parameters (e.g., direction, area, depth, shape index, etc.).

Finally, several works in literature focus on exploiting the *regularity* of simple façade elements, especially windows, to derive semantic information from façade point cloud. In particular, some specific *grammar rules*, either derived in a top-bottom or in a bottom-up manner, are used for the description of such repetitive patterns. The work of Ripperda (2008) aimed at interpreting building façades with a description grammar. Façade elements (such as 'walls,' 'doors,' and 'windows') and abstract elements (such as repetition, symmetry and array) are written as 'terminals,' and then building façades are described by hierarchical composition of these 'terminals.' The reconstruction of a building façade can be seen as a stochastic process of interpreting sensor data with the grammar. The probability distribution of the terminals is searched under supervision of the reversible jump Markov Chain Monte Carlo (rjMCMC) method (Green 1995). Regular shaped buildings with flat façades can be well represented with the defined grammar. However, the stochastic analysis for complex façades might not lead to meaningful results. Becker (2009) first reconstructed polyhedron models from TLS data and images, and then used the clues extracted from the models to synthesize areas without laser data. After the reconstruction step, the

structural and hierarchical relations between façade objects are described in the form of shape grammar similar to the one in Ripperda (2008). The discovered patterns can be propagated to the upper regions of the same façades to compensate the low point density, or even propagated to other buildings to create similar styled building models.

The above mentioned procedures generally work well on simple façade elements recognition, like walls and windows. However, a more detailed set of façade elements need to be represented in building reconstruction at a level-of-detail adequate for thermal building analysis. In addition, some of the previous reported methods are based on regularity assumptions for building façades that may fail in the case of clutter and large occlusions. Finally, detected semantic objects are generally not combined with geometric modelling in a unique framework. For these reasons a new classification strategy has been developed which is aimed at recognizing a more detailed set of façade elements from TLS data, such as sidewalls, roofs, and doors to allow modelling at a higher level-of-detail. The developed approach does not make use of strong assumptions about façade regularity but exploits a set of generic architectural priors for buildings built in 1950 - 1975 for the generation of a semantic rich façade model.

3.2. The hierarchical classification approach

In this section the developed approach for façade object classification is presented. The process is based on some general knowledge derived from the nature of building structures. In particular, for the building type considered in this research, façades have a dominant planar structure, characterized by a flat dominant surface and with other façade's components having off-plane depth variations with respect to this plane, either positive (extrusions) and negative (intrusions). This assumption is also generally valid for a great variety of modern building architectural styles. These priors are then translated into logic-level terms to derive the rules describing relations between concepts which are organized into a hierarchical classification tree.

3.2.1. Façade knowledge

As the products of human construction, a building is constituted by various components. These components, or 'features,' serve as important context 'nodes' during both construction and recognition process. A building can be decomposed either according to component functionality, such as: *wall*, *roof*, *door*, *window*, *balcony*, *dormer*, *awning*, *chimney*, *water pipe* and *decoration*, or according to the spatial partitions, such as: *floor*, *floor 1,* and so on. In this dissertation façade object classification is performed considering the former principle. Indeed, objects with different functions usually appear in different geometries, which are easily distinguishable from TLS data. On the other hand, the possibility to define the

different functionality of each façade object is one of the main aspects of a semantic rich model.

It is not necessary and somewhat redundant to assemble all functional components into the knowledge base classification of building features. Thus, the selection of features follows three considerations:

1. the features which are within the interests of the reconstruction are only considered. For example, awnings and water pipes are not included in the classification phase because they are not often desired in the reconstructed model;
2. the features which can be hardly captured in the sensor data are not considered. For example, the chimneys are seldom scanned by using TLS; and
3. the features can be defined in the level of primitive or plane. At this moment, only planar features are considered because buildings built in 1950 - 1975 presents façades that are mainly composed by planar features, and planes can be easily extracted from laser point clouds with existing methods.

In particular, only the following feature classes (or semantic type in other words) are defined (Fig. 3.1):

Ground: although ground is not a building feature, it is included in the feature type list because generally it is acquired by TLS and it is useful for recognizing other façade elements.

Façade wall: is the remaining part after removing all protrusions (e.g., roof, railings, etc.) and intrusions (windows, doors, and sidewalls). Since protrusion locates at the top/bottom of the façade (e.g., roof), or the inside of the wall (e.g., railings) and intrusions locates at the inside of the wall, the wall appear continuous and keep coplanar in the area of a façade. Therefore, wall element covers nearly the whole façade and is the largest element.

Sidewall attaching window/door: is the side face of the wall and perpendicular to the frontal façade wall. Sidewall is the connection between the main façade plane and windows/doors elements. It is different from other intrusions like doors and windows, which always appear parallel to the frontal face of the wall.

Roof: this element is generally located at the top of the façade. In addition, it is generally a horizontal or mildly sloped element protruding out from the façade to protect the façade from rain.

Fig. 3.1. Example of a façade point cloud and the defined feature classes (highlighted in purple).

Window: is an intrusion element which is parallel to the frontal façade wall. A window can be found at each floor and generally it is inside a ‘O shaped’ hole in the façade wall.

Door: is an intrusion element which is parallel to the frontal façade wall. A door can be found only at the first floor and generally it is inside an ‘inverted U shaped’ hole in the façade wall.

Wall attachment: is an element protruding out of the façade wall. Possible wall attachments include window sill, stairs, railing and so on.

3.2.2. Façade classification rules

In a façade, each feature has some peculiar characteristics which may be described as *recognition rules* to be derived either from the statistical analysis of a training set or based on the knowledge of some priors about the façade and urban scene. The former approach has the main disadvantage of an out-of-core training phase based on a wide number of manually classified cases. The latter is applicable only for the building types featuring the defined priors. The classification presented in this dissertation belongs to this second group and are designed for buildings in the period 1950 - 1975. On the other hand, to increase the application domain, the defined rules are quite general and they can be used for a wide variety of building types. Only for a limited number of architectural styles the defined priors may not be adequate.

Each building feature has a number of attributes, which are similar within the same feature type. Strong clues can be extracted from the combination of feature attributes to suggest particular feature types. Some common attributes are considered below:

Size: the size is probably the most distinguishable attribute. It is most likely that walls occupy the largest space on a building, and usually windows are not longer than the height of a floor. The term size can refer to a feature's length, width, height, area, or volume.

Position: some features can be expected at certain relative positions inside a scene. For example, ground is usually the lowest part, and roofs are seldom in a higher position with respect to other structures.

Orientation: features' orientations are also predictable. For example, ground surfaces are almost horizontal; walls are usually vertical as they are the supporting body of buildings; roofs are never vertical.

Besides a feature's own attributes, some spatial relations between features can be predicted. These spatial relations are:

Intersection: intersection is a basic spatial relation indicating the neighbouring relationship of two features. For example, neighbouring walls often intersect perpendicularly; dormers must intersect a roof plane.

Angle: angles between building features are often parallel or orthogonal. For example, windows on a wall usually are parallel with respect to wall; sidewalls on a window often are perpendicular to the main wall.

Inside: if a feature is part of another feature, it is usually inside the boundary of the other feature. For example, windows of a wall must be inside the wall's boundary.

To a certain direction of: the locations of some features can be expected to a certain direction of other features. For example, a roof can be expected above a wall; sidewalls should be perpendicular to the wall façade, etc.

All these information about feature attributes can be formulated in terms of first-order logic (Russell and Norvig 2003) to define the initial rules of the knowledge base classification. Firstly, the statement of a feature type is represented with a binary function:

$$IsType(arg1,arg2) = TRUE$$

where *arg1* is a geometry feature while *arg2* is a constant indicating *arg1*'s feature type. For example, the assertion $IsType(F_1;Wall) = True$ states that the type of F_1 is *Wall*.

Feature attributes and spatial relations are also represented with functions, which take the index of a particular feature as their sole argument. Sometimes, a feature's own attributes does not give a useful clue if the relative situation is unknown. Therefore, some relative attributes and their enquiry functions are introduced. Determination of relative information requires sorting of particular attributes throughout a group of features. For example, a feature is considered Low if its position is ranked as the lowest among the other feature types. The attribute functions are:

- position functions (*IsLow*),
- size functions (*IsLarge*),
- orientation functions (*IsVertical, IsHorizontal*).

The spatial relation functions are associated with a reference feature. They are:

- 'to a certain direction of' functions (*IsOver, IsOut*),
- angle functions (*IsPerpendicular, IsParallel*),
- inside functions (*IsInside*).

In particular, the rules for the previously defined façade features are here reported:

1. A 'ground' feature (*f*) is a low, large, and horizontal plane:

$$\forall f \quad IsType(f, Ground)$$

$$\Rightarrow IsLow(f) \quad \wedge \quad IsHorizontal(f)$$

2. A 'façade wall' feature is a large and vertical plane:

$$\forall f \quad IsType(f, WallFacade)$$

$$\Rightarrow IsLarge(f) \quad \wedge \quad IsVertical(f)$$

3. A 'roof' feature (f_1) is a large mild slope element at the top of a wall and protruding out of the façade:

$$\forall f_1 \quad IsType(f_1, Roof)$$

$$\Rightarrow \exists f_2 \quad \neg IsVertical(f_1) \quad \wedge \quad IsOver(f_1, f_2)) \quad \wedge \quad IsOut(f_1, f_2) \quad \wedge \quad IsType(f_2, WallFacade)$$

4. A 'wall attachment' (f_1) is an element protruding out of the wall façade:

$$IsType(f_1, WallAttachment)$$

$$\Rightarrow \exists f_2 \quad IsOut(f_1, f_2) \quad \wedge \quad IsType(f_2, WallFacade)$$

5. A 'side wall' (f_1) of a window/door is an intrusion of the wall façade perpendicular to façade walls:

$$\forall f_1 \quad IsType(f_1, SideWall)$$

$$\Rightarrow \exists f_2 \quad \neg IsOut(f_1, f_2) \quad \wedge \quad IsPerpendicular(f_1, f_2)) \quad \wedge \quad IsType(f_2, WallFacade)$$

6. A 'window' (f_1) is an intrusion of the wall façade perpendicular to the wall and is inside a 'O-shape' of the façade wall:

$$\forall f_1 \quad IsType(f_1, Window)$$

$$\Rightarrow \exists f_2 \quad \neg IsOut(f_1, f_2) \quad \wedge \quad IsParallel(f_1, f_2)) \quad \wedge \quad IsInsideO(f_1, f_2) \quad \wedge \quad IsType(f_2, WallFacade)$$

7. A ‘door’ (f_1) is an intrusion of the wall façade perpendicular to the wall and is inside a ‘U shape’ of the façade wall:

$$\forall f_1 \quad IsType(f_1, Window)$$

$$\Rightarrow \exists f_2 \quad \neg IsOut(f_1, f_2) \quad \wedge \quad IsParallel(f_1, f_2)) \quad \wedge \quad IsInsideU(f_1, f_2) \quad \wedge \quad IsType(f_2, WallFacade)$$

3.2.3. Semantic object classification

In contrast to other supervised classification strategies, the developed method obtains the classification criteria not from statistical analysis of training sets but from semantic interpretation of the façade based on the feature types, attributes and rules described in the previous section. All these information are stored into a *hierarchical classification tree*. Rules presented in the previous section rely on some information associated to object orientation (like verticality and horizontality) and object position (e.g., a roof is always over a wall, ground is at the lowest level, etc.). In order to manage these relationships, a proper reference system for the point cloud is assumed. In particular, a point cloud needs a pre-processing step where it is aligned into a local cartesian coordinate system having the Z-axis aligned to the ground up-vector. This can be simply obtained by levelling the TLS during data acquisition. In addition, actual laser scanners are generally equipped with an electronic inclinometer allowing the correction of verticality.

With reference to Fig. 3.2 the classification process can be divided into two main steps. First, simple and easy detectable objects are classified (ground, wall and roof) by evaluating both area and position of all the detected objects. The ground is detected at first by looking for the lowest object between horizontal (or pseudo-horizontal) features.

Then façade walls are extracted since they are perpendicular to the ground and with the largest area with respect to all remaining vertical objects. Indeed, the wall area in a façade is generally much larger than the one covered by other vertical objects, like windows or doors. In addition, main façade planes are generally less than sidewalls, windows and doors, so they can be considered as outliers and can be easily detected by robust statistics. According to probability theory and mathematical statistic, more than 95.4 percent of the sample should fall into $\pm 2\sigma$ interval. Therefore, objects whose size is beyond $+ 2\sigma$ are classified as wall.

Finally, non-vertical objects over walls and protruding out from the façade are classified as roof.

In the second step all non-classified objects are analysed. Their position with respect to the main façade wall is evaluated. Objects in front of the façade are considered as extrusion objects and then classified in a general way as wall attachments. Objects like stars, railing and balconies can be classified into this category.

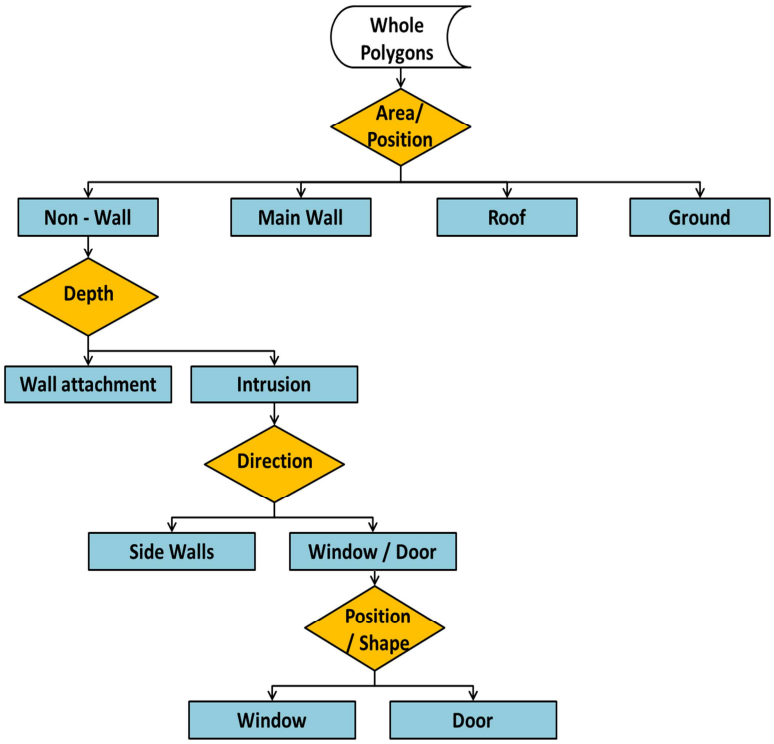


Fig. 3.2. Hierarchical classification tree, orange diamonds are conditions while blue rectangles represent façade elements.

Objects beyond the plane of the façade are first generically assigned as intrusions. Then they are split into sidewalls and windows/doors according to their orientation with respect to the main façade. Indeed, sidewalls can be easily recognized because, in contrast to other intrusions, their orientation is perpendicular to the façade walls. Intrusions parallel to the façade are generically classified as windows/doors. In order to distinguish between doors and windows, it is verified their position and the shape of the gap in the main façade plan. In particular, doors are searched only at the bottom floor and in correspondence of a characteristic gap of the main façade plane having an 'inverted U shape.' Other intrusions parallel to the façade plane and in correspondence of 'O-shape' gaps in the main wall are classified as windows.

In the presented approach, classification clauses consider some quite strict spatial relations (e.g., vertical, horizontal, perpendicular, and parallel). However, due to

imperfections during the construction and noise in the data, the defined rules generally do not strictly hold. For example, it is very unlikely that a sidewall is perfectly perpendicular to the façade wall. For this reason all the previously defined rules concerning spatial relations are relaxed of a predefined angle α, which was fixed in the experimental application described in Chapter 5 equal to 5°.

3.2.4. Enriched model and export file format

Once the geometrical model of the building is generated and objects are classified in façade features according to their functionality a semantically enriched model of the façade can be generated. Due to the presence of occlusions or lacks in the point cloud, some part of the model might be pending. In the next Chapter 4 a procedure for recovering missing data on the basis of regular pattern detection will be addressed.

Even if several geometric file formats have been developed in both Computer Graphic and CAD fields, the number of those allowing the semantic definition is relatively small. The two most prominent standards are Industry Foundation Classes (IFC) and City Geography Markup Language (CityGML). Even if the two data structures share some similarities, several works in the literature (Benner et al. 2005, Isikdag and Zlatanova 2009, Nagel et al. 2009) showed differences between the two standards both in the description of geometry and in semantic object definition. To allow a higher interoperability, once the building geometry has been defined (Sect. 2.2) and each façade objects has been classified (Sect. 3.2.3), the 3D model of the building façade is generated both in CityGML and IFC standards.

3.2.4.1. CityGML

CityGML (Gröger and Plümer 2012) is the international standard of the Open Geospatial Consortium (OGC) for the representation and exchange of 3D city models, see also OGC (2014). It defines the 3D geometry, topology, semantics and appearance of the most relevant topological objects in urban or regional context. The principal focus is on the semantic definition of all objects (features) that are relevant for applications of 3D city models (e.g., building and their parts: walls, dormers, doors, windows, etc.). Furthermore, the relations between those features (e.g., the relation of a door to the wall it contains) are represented in explicit way. For the representation of geometry (and topology), CityGML uses a standardized model provided by the Geography Markup Language (GML) and the eXtended Markup Language (XML). CityGML is not just restricted to modelling buildings. It extends to tunnels, bridges, transportation infrastructures, water body and vegetation. However, the building model is the most important component of CityGML as it enables the representation of buildings and their component part with regard to geometry (both outdoor and indoor) as well as to semantics (feature types and properties).

The geometrical feature representation is based on the geometrical model provided by GML 3.1.1 which is an implementation of ISO 19107 ‘Spatial Schema.’ CityGML is restricted to planar polygons: all coordinates of the outer boundary and of the optional interior boundaries must be located in the same plane. Non-linear structures can be approximated by planar surfaces. For each position of the geometrical representation, absolute 3D coordinates must be given explicitly. Features are represented geometrically by the well-known boundary representation model. Surfaces must be mutually non-overlapping and non-penetrating. Building and their parts have common attributes (like a class, creation and destruction date, the owner of a building, measured height, roof type, the number of stories above and below ground).

The features in CityGML can be represented in five discrete Levels-of-Detail (LoD). In particular, the LoD concept of CityGML is not restricted to geometrical aspects but covers also semantic ones: with increasing LoD, the semantic richness also increases. The LoD concept is characterized by the following properties:

- data integration and interoperability is facilitated, since features that are represented in the same LoD can be integrated more easily than features of different LoDs;
- each LoDs reflects specific application requirements and hence is suitable for a certain class of applications;
- the LoDs may depend on the reconstruction method; and
- the same feature can be simultaneously represented in different LoDs. This facilitates analysis and visualization tasks, since tools can select dynamically the most appropriate LoD for the task.

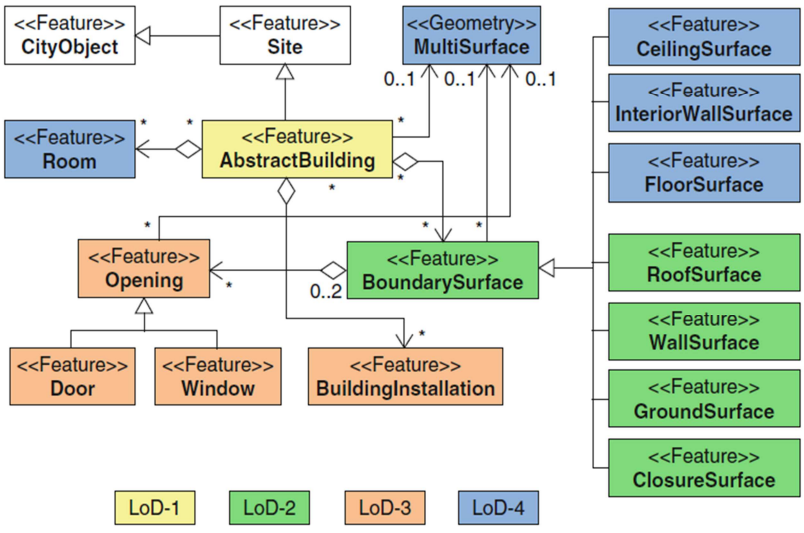


Fig. 3.3. CityGML file representation of a building at different LODs showed as UML instance diagram.

A building can be represented in multiple LoDs simultaneously and it is formally described by UML diagrams (Fig. 3.3).

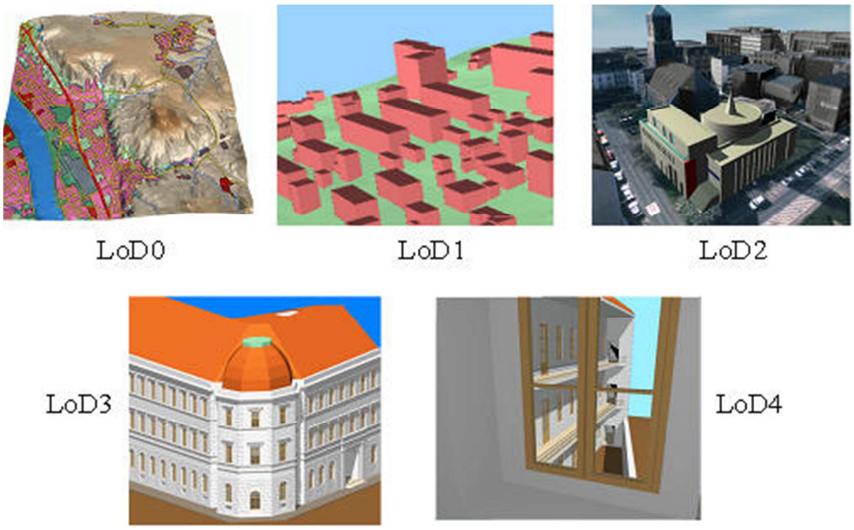


Fig. 3.4. The five levels of detail (LoD) defined by CityGML [from Kolbe et al. 2005].

In particular, the geometry representation becomes more detailed with increasing LoD. Different LoDs for building models are described from 0 to 4 (see Fig. 3.4 for reference).

A building in **LoD0** can be either represented by horizontal 2.5D polygons with roof level height or with footprint level height. The semantics is modelled by *Building* instance with corresponding attribute values.

In **LoD1** a building is represented as block model that is either represented as a solid or as multi surface. A *Building* can be partitioned into different *BuildingParts*. Each part has its own roof type and representation as solid. The roof type refers to the shape of the building in reality, not to the representation in LoD1. Hence, the value may be 'gabled roof' while the representation in LoD1 always has a flat, horizontal roof.

LoD2 adds generalized roof structures to LoD1. In addition, boundary surfaces of a building can be represented as thematic features. Vertical walls surfaces are represented as *WallSurface*, surfaces that cover the building from above as *RoofSurfaces* and horizontal surfaces that delimit the bottom of the building from the ground are represented as *GroundSurfaces*.

If LoD2 is extended by openings (windows, doors), detailed roof structures (dormers, chimneys, roof overhanging) and detailed façade structures, **LoD3** can be achieved. These objects can be represented as features with their own attributes and surface

geometry. Windows and doors are assigned to the corresponding wall/roof surface and each feature has a surface geometry representing its spatial properties. High-resolution textures can be mapped onto these structures.

In **LoD4**, interior structures of buildings are considered as well. The geometry of each room is represented and interior structures for 3D objects (e.g. rooms, interior, stairs, furniture, etc.) are added.

CityGML allows its extension for applications requiring specific feature types, attribute and relations. For that purpose, CityGML provides a mechanism called Application Domain Extension (ADE). An ADE is specified by an application schema of CityGML in different XML name spaces (in the same fashion as CityGML is an application schema of GML). It defines new feature types (with new attributes, geometries and associations), with may be subtypes of existing types. Furthermore, new attributes, geometries, and associations can be added to the existing types by using a hooking mechanism. One CityGML dataset may use multiple ADEs simultaneously, facilitating the multi-functional use of 3D models. Furthermore, the ADE mechanism can be applied iteratively, generating an ADE of an ADE, and using object-oriented concepts in order to define application schemas at different LoDs. An alternative mechanism to extend City GML is the use of generic grammar attributes to add additional features and attributes on demand.

3.2.4.2. IFC

The representation of buildings and their structures is also the objective of Building *Information Models (BIM)*, of *Computer-aided architectural design (CAAD)* and *Architecture, Engineering, Construction (ACE)* models. For data exchange in the BIM world, there is one important ISO standard called IFC (Industrial Foundation Classes) in its version 2x. Currently version is 2x Edition 3. IFC is an object oriented format developed by the International Alliance for Interoperability (IAI). The goal of IFC is to specify a common language for building industry technology aimed at improving communication, productivity, delivery time, cost, and quality throughout the design, construction and maintenance life cycle of buildings (Hallberg and Tarandi 2009).

IFC defines an EXPRESS based entity-relationship model consisting of several hundred entities organized into an object-based inheritance hierarchy.

In particular, IFC divides all entities into rooted and non-rooted entities. Rooted entities derive from *IfcRoot* and have a concept of identity (having a globally unique identifier - GUID), along with attributes for name, description, and revision control. Non-rooted entities do not have identity and instances only exist if referenced from a rooted instance directly or indirectly. In the IFC file format each *IfcRoot* specification (called “class”) is used to describe a range of things that have common characteristics.

In particular, there are three kinds of IFC classes: object classes, relation classes and property sets.

The object classes consist of a triplet (GUID, OS, FU), where GUID defines the general identifier of the current IFC object, OS defines the ownership features of this object and FU are the functional units. These functional units define the context of use of the classes (i.e., the geometrical representation, its localization, its composition, etc.). Fig. 3.5 shows 12 types of building elements that can represent a building structure in IFC standard.

The relation classes represent the various relations between the object classes and their functional units. There are five fundamental relationship types: composition, assignment, connectivity, association, and definition. Indeed, IFC format is made of objects and connections between them. Object attributes describe the 'semantic' of the object while connections between objects are represented by 'relation elements'. For example, building elements and opening elements are subtypes of structural element. Each building element has zero or more opening elements, i.e., a wall without any door or window has zero openings, whereas each opening element (like door, window) is attached to only one building element.

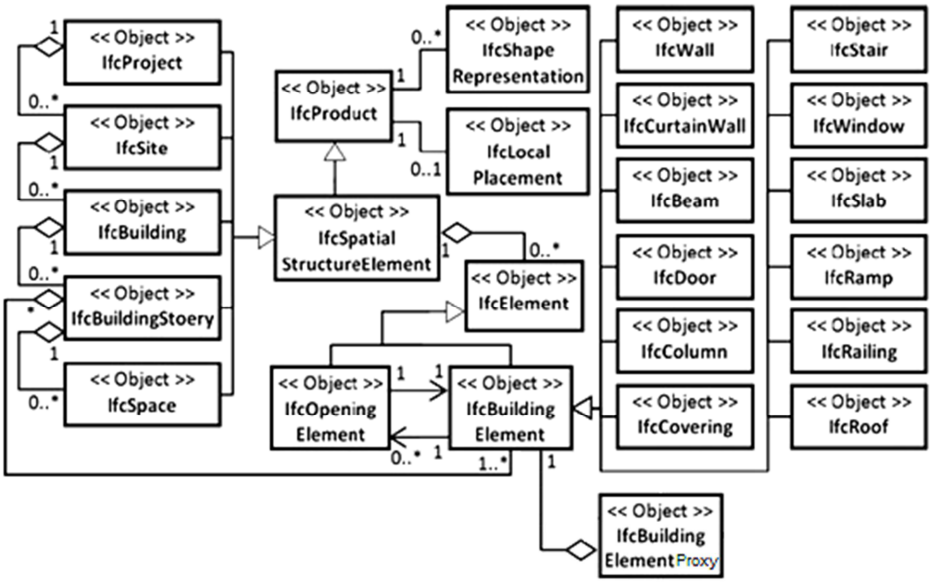


Fig. 3.5. Simplified IFC file structure.

The property classes capture dynamically extensible property sets. A property set contains one or more properties which may be a single value (e.g. string, number, unit measurement), a bounded value (having minimum and maximum), an enumeration, a list of values, a table of values, or a data structure. While IFC defines several hundred property sets for specific types, custom property sets may be defined by application vendors or end users.

3.2.4.3. CityGML and IFC comparison

Common to CityGML and IFC is a detailed semantic and geometrical model of building and its interior structures. In the latest version of IFC, coordinates for position and height values in the global system WGS 84 can be specified, in addition to the local coordinate reference systems typically used in CAD. However, there are significant differences between both models:

- The definition of semantic objects differs in both models. IFC focuses on the construction and design of buildings and provides construction elements like slab, beam, or wall. Such objects typically are in the boundary of multiple rooms and simultaneously are part of the outer boundary of a building. In contrast, the definition of CityGML describes how buildings are observed or used. Hence, objects such as rooms or wall/ceilings of a single room are defined. In addition, IFC holds more detailed information about building objects than CityGML.

- CityGML uses Boundary Representation for the definition of spatial properties, since the focus is on how buildings are used and observed. According to the constructive nature of CAAD, in IFC additionally Constructive Solid Geometries (CSG) and sweep geometries are applied.

- The objects in IFC are represented in one LoD only. A multi-resolution representation as in CityGML is not available.

The relationships between CityGML and IFC have been discussed intensively in the last years. Particularly, the multiplicities of relationship between the two are analysed. Indeed, 1:1 relations occur rarely, whereas 1:n (one CityGML feature corresponds to n IFC features) and n:1 relations prevail. For this reason no straightforward translation rules between IFC and GityGML standards are available.

3.2.4.4. CityGML and IFC output

The object features derived from the previous classification step (Sect. 3.2.3) can be used in a straightforward way to generate a CityGML model at LOD3. Indeed, the features defined in CityGML meta-language as *RoofSurface*, *WallSurface*, *GroundSurface*, *Window*, *Door*, *BuildingInstallation*, *GroundSurface* have a clear

correspondence with previously classified objects (Fig. 3.6). Each feature is represented by a surface geometry (Sect. 2.2) describing its spatial properties. Focusing on the façade geometry, a solid representation of the building is not always available (e.g., roofs are missing in many cases). For this reason a multi-surface representation, that do not completely seal the building, is given.

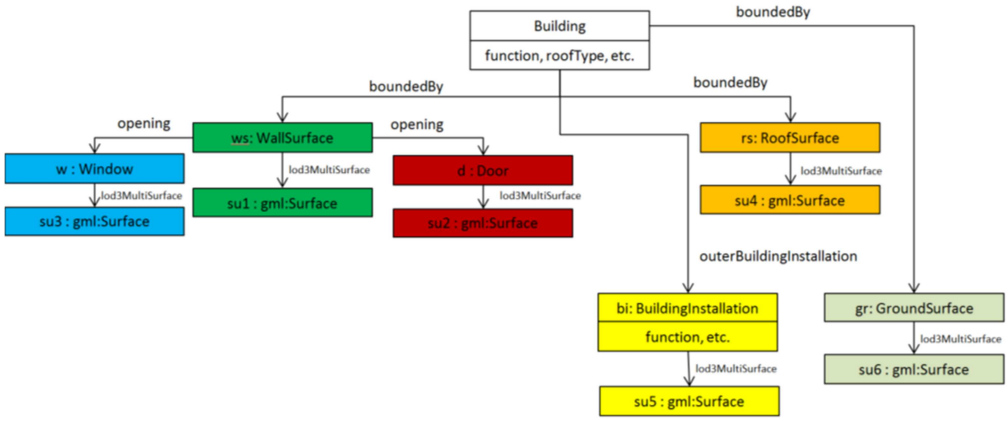


Fig. 3.6. Illustration of a LoD3 building represented as CityGML feature structure as UML instance diagram.

As previously mentioned CityGML standard has a high flexibility. This allows high interoperability with several CAD environments as well as thermal evaluation performance software packages. An important aspect is given by the fact that additional information can be added to the model, e.g., the *Location* which indicates the global position of the building and its orientation. This can be used to evaluate façade exposition and sunlight or other descriptive data which are of major interest for energy efficiency evaluation, like the *insulation value* (U-value) for windows, walls, floors and roofs (see Africani et al. 2013) can be added as generic attribute to the building features.

Concerning IFC, compared to the high semantic differentiation the previously defined classification (Sect 3.2.3) appears too coarse for this standard. Indeed objects that were previously classified in a general way as wall-attachments, in a IFC files can be classified more specifically, according to their functionality, as different features (e.g. *IfcStair*, *IfcRailing*, etc.). This is the multiplicities of relationship underlined by different authors for direct conversion between CityGML and IFC. However, in many cases and for a variety of applications it is not necessary and somewhat redundant to assemble all functional components of a building into the associated IFC class. For this reason, a simplified modelling, both in the geometrical representation and in the semantic definition, can be used. In simplified modelling, objects belonging to a different IFC classes can be grouped in the same feature. Considering thermal analysis and retrofitting of buildings the most important semantic classes concern

wall, window, door and roof elements. Indeed these elements are the ones playing the most important role in energy dispersion and the first elements to be retrofitted/updated for increasing energy efficiency. Lower interest is paid to elements like railing and stairs. Then, considering a simplified semantic modelling object classified as Wall Attachments within the hierarchical classification step can be represented in the IFC standard as *IfcBuildingElementProxy*, i.e. a definition that provides the same functionality as an *IfcBuildingElement*, but without having a defined meaning of the special type of building element it represent. In a similar way, the ground can be represented in IFC as a particular kind of *IfcSlab* that can be defined for example as *BaseSlab*. On the other hand wall, window, door and roof object has a direct correspondence with IFC classes. Spatial properties of each class are represented by means of surface elements in a Boundary Representation.

Chapter 4
Façade regularity detection

This chapter addresses the reconstruction of incomplete models with the help of high-level architectural objects and the identification of repeated patterns in urban façades. Indeed, building façades generally exhibit a high degree of self-similarity and redundancy. The presented approach explicitly makes use of these characteristics of the urban scenes to enable plausible recovery of missing geometry.

Completion of point clouds is often necessary because, even though a considerable effort is put on data acquisition planning, data obtained with range scanners usually suffers from occluded or defective portions of objects that either could not be perceived during acquisition or might have adverse material properties that hinder the scanning device (for example, marble, see Godin et al. 2001). Nonetheless, a complete surface representation without holes is usually required for further processing or rendering. Therefore reconstruction algorithms must not only be able to recover the surface parts that have been captured, but should also synthesize plausible geometry in missing areas. However, this is a challenging problem. This is mainly due to the wide collection of architectural elements and styles that could be combined in façades, with large changes between different countries, to variations in sampling density, and to noise and outliers which might be present in the point cloud. For these reasons approaches based on general smoothness assumptions or relying on a database of example cases from which a completing surface can be retrieved, may be suitable to complete small holes but generally fail in the case of large missing parts.

The method presented in this Chapter is based on the completion of missing parts by means of high-level architectural features (e.g., windows, doors, etc.) detected in the building façade by means of the methodologies described in the previous chapters. Indeed, façades of the building type target of this research often exhibits a regular arrangement consisting of repeated patterns and self-similarities (Fig. 4.1). The presence and the *regular patterns detection (RPD)* can be used to reduce holes induced by occlusions and enhance automatic façade modelling. The key observation is that the same geometry is scanned multiple times over recurrences of the repeated elements. The non-local multitude of geometry provides opportunities to complete missing parts using information from other regions. Exploiting these repetitions is

very powerful, as it can reveal fully or partially occluded elements which might otherwise escape detection. In particular, the detected façade features can be extended into the empty regions and serve as guidance for hole-filling.

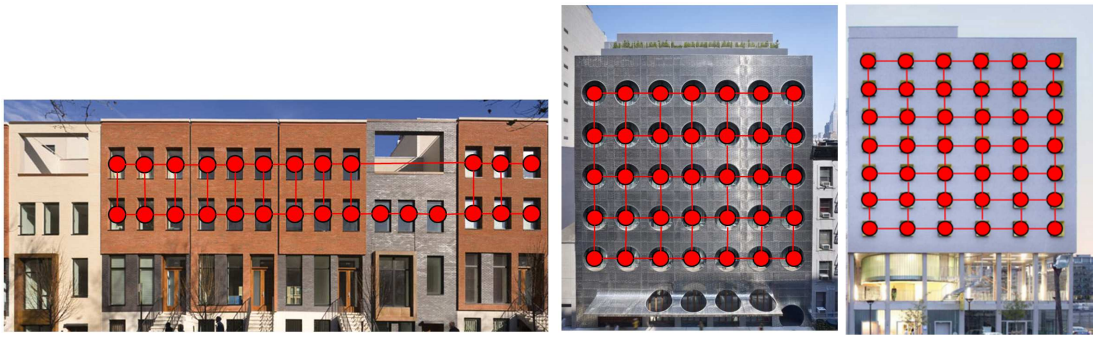


Fig. 4.1. Examples of buildings presenting repeated patterns.

Regularity and self-symmetry in urban buildings is widely demonstrated across countries and cultures. Such large scale repetitions arise from manufacturing ease, build-ability, functional requirements and aesthetic. For thermal retrofitting purposes the detection of these repetitions, e.g. windows, is of primary importance because they are the first element to be updated to improve thermal efficiency. While in recent years many techniques have been developed to detect repeated parts in models (Debevec et al. 1996, Hays et al. 2006, Mitra et al. 2006, Korah and Rasmussen 2007, Pauly et al. 2008, Musialski et al. 2009), most of these works do not investigate how to optimize the use of strong regularity in urban buildings. Moreover, most of the techniques are applied in image space by analysing 2D images sampled over a regular domain. Only few attempts have been made towards detection of regularity directly on 3D geometry.

The challenge lies in the automatically determination of which elements repeat in the façade and the regular pattern they form. The complexity of the problem is increased when missing data due to occlusions and the variability in data resolution have to be considered. In the presented approach, instead of making strong prior assumption about the models and blindly recreating the geometry using predefined procedural rules, a bottom-up method is used attempting to extract maximum information from the point cloud. Indeed, most buildings are designed and generated in a procedural and modular fashion. In particular, building façades that are targets of this research can be represented as the repetition of some *basic-geometries* into *multiple 2D periodic structures*. However, instead of learning parameters from a codebook of rules, repeated elements and lattice structures are learned directly from empirical data.

4.1. State of the art

Given the large volume of work on urban modelling, we refer the reader to the recent survey by Musialski et al. (2012) for a comprehensive review. Here only previous works addressing repetition detection of urban scenes are recalled.

Detection of symmetries and repeated elements such as windows and balconies in urban scenes is a problem that has received significant attention in the field of image analysis while detection in point clouds received significant attention lately.

Image-based façade analysis methods can be based either on the processing of single images (Müller et al. 2007, Musialki et al. 2010) or multi-view images jointly with the extracted point clouds (Xiao et al. 2009).

Concerning single image analysis, Schaffalitzsky and Ziesserman (1999) used edge detection to find interesting elements and successively recognized patterns by using a grouping strategy for translational grids based on maximum likelihood estimation. Korah and Rasmussen (2007) addressed the problem of automatically detecting 2D grid structures such as windows on building façades from images.

In Müller et al. (2007) and Lee and Nevatia (2010) the unique characteristics of façade structures such as regularity and orthogonality are exploited in a statistical model to detect translational symmetries and repetitive window structures. However, those techniques are largely based on the strong assumption that a façade is governed by a single hidden global rectilinear grid and global correlation between repeated elements in the scene. A more general approach is presented in Jahangiri and Petrou (2009), where repetitive elements, and especially windows, are detected as blobs in a colour image but the underlying structure is not analysed.

Wu et al. (2010) presented a feature-based method that extracts repetition and symmetry patterns from a rectified image. They made simplifying assumptions such as constant repetition height and no gaps between floors. While in Park et al. (2011) a feature extraction method is used to efficiently detect multiple façade regularities, but this algorithm implicitly relies on structured data and then it is not suitable for point clouds.

Finally, shape grammar methods are presented in Ripperda and Brenner (2009), where a limited manual interaction is required and in Teboul et al. (2011) where an out-of-core training phase is performed. The grammar yields segmentation of façades in semantic parts (e.g., walls, windows, balconies, etc.). However, in both cases repetitive structures are not detected.

Works on reconstruction of urban scenes have also been based on collections of photos or multi-view stereo relying on photogrammetric reconstruction and image-based modelling techniques.

Debevec et al. (1996) proposed an iterative image-based modelling method that exploited characteristics of architectural objects coupling an image-based stereo algorithm with manually specified 3D model constraints. More recently, Sinha et al. (2008) presented an interactive modelling system using unoriented sets of photographs, leveraging the piecewise-planarity of architectural models. Xiao et al. (2009) efficiently modelled façades from images by decomposing them into rectilinear elementary patches. Later they extended the semantic segmentation and analysis to more general scenes, in order to produce visually compelling results by imposing strong priors on urban regularity.

Fewer approaches exist in the case regularity identification is performed in large point clouds. State of art works using structure repetition in urban façades can be classified in interactive and automatic procedures.

While laser scans are in general dense and relatively regular, thus perfectly suited for architectural reconstruction the acquisition process may result in corrupted and incomplete data. In order to overcome such problems, several methods propose to process the data with user control. Interactive tools rely on similar frameworks (Zheng et al. 2010, Nan et al. 2010). The user defines first some basic façade objects which are then snapped to similar elements in the point cloud.

Boehm (2008) published a method for completion of TLS point clouds, which is done by iteratively utilizing the repetitive information typically present in urban buildings. Another approach aiming at a similar goal was introduced by Zheng et al. (2010). It is also an interactive method for consolidation which completes holes in scans of building façades. This method exploits repetitions to consolidate the imperfect data, denoise it, and complete the missing parts. Another interactive tool for assembling architectural models was introduced by Nan et al. (2010). In this system, the user defines simple building blocks (Smart-Boxes), which snap to common architectural structures like windows and balconies. They are assembled through a discrete optimization process that balances between fitting the point-cloud data and their mutual similarity. In combination with user interaction, the system can reconstruct complex buildings and façades.

Discovering regular structures in an automated way is a challenging task since there is no a priori knowledge of size, shape, or location of elements describing the pattern. In addition, façade elements can be incomplete or corrupted by noise. For this reason

automatic solutions generally rely on some architectural assumptions about the building façade.

An automatic data-driven façade reconstruction by cell decomposition is introduced in Becker and Haala (2009), which requires a coarse 3D building model as input. They also proposed to automatically derive a façade-grammar from data in a bottom-up way. However, this algorithm focused on façades containing windows only and this could not handle more complex façades.

In Stamos and Allen (2002) window-like rectangular features are extracted by using 3D edge detection on high-resolution 3D data but regularity is not enforced. In Friedman and Stamos (2011) and Mesolongitis and Stamos (2012) methods for detecting regularities in building façades are presented. However, they are mainly addressed to detect only one feature type (e.g. windows) and require scan points being organized in 2D vertical scan-lines which may be a limiting aspect if multiple scans are fused together.

In Pauly et al. (2008) a general regularity detection method for 3D models is presented. This general approach can be used for extracting a single façade pattern. However, in this case similarities in the model are detected by considering a local similarity measure of the point cloud curvature which is more prone to output outliers. In Triebel et al.(2006) a Markov Network approach that requires training is used to label points as windows. In Shen et al. (2011) façades are adaptively partitioned by horizontal and vertical planes based on the boundary features of planar regions. However, wrong horizontal or vertical splitting may result in wrong façade structure identification. In addition, this method can be seriously affected by the variation in the resolution and the window appearance, which may co-exist in a single scan.

These automated methods generally rely on the assumption that the façade can be split into building blocks by a single rectilinear grid. Even if there is a certain number of façades that satisfy this assumption, in many cases façades presents a more complex structure.

Structure discovery is also addressed in the field of Computer-Aided-Design (CAD). Shikhare et al. (2001) proposed a compression scheme that exploits geometric patterns in CAD models. This method is most effective for procedural designed model where the repetitive elements appear as separate connected components. Li et al. (2010) introduced regular features trees that provide a concise description of symmetry features in order to capture important aspects of the aim of the geometric design. This method is specifically designed for shapes that are bounded by planar spherical, cylindrical, conical and toroidal surfaces. However, those techniques can be used only for simple geometry of these typical CAD models.

In the next sections the developed procedure for *regular patterns detection (RPD)* is presented. In contrast to other works (Müller et al. 2007, Xiao et al. 2009, Musialski et al. 2009) that rely on the assumption that a building façade can be split into building blocks by a single rectilinear grid, this work presents a more flexible strategy aimed at detecting concatenated and\or interlaced grids of elements. Indeed, even if there are a certain number of façades satisfying the rectangular lattice assumption, in many cases façades presents more complex structure. In Fig. 4.2 some synthetic façades are represented demonstrating some cases consist of multiple periodic regions. In addition, in contrast to other developed techniques regular patterns are not only detected but these repetitions are also exploited to reveal fully or partially occluded elements and complete missing areas with high level structures.

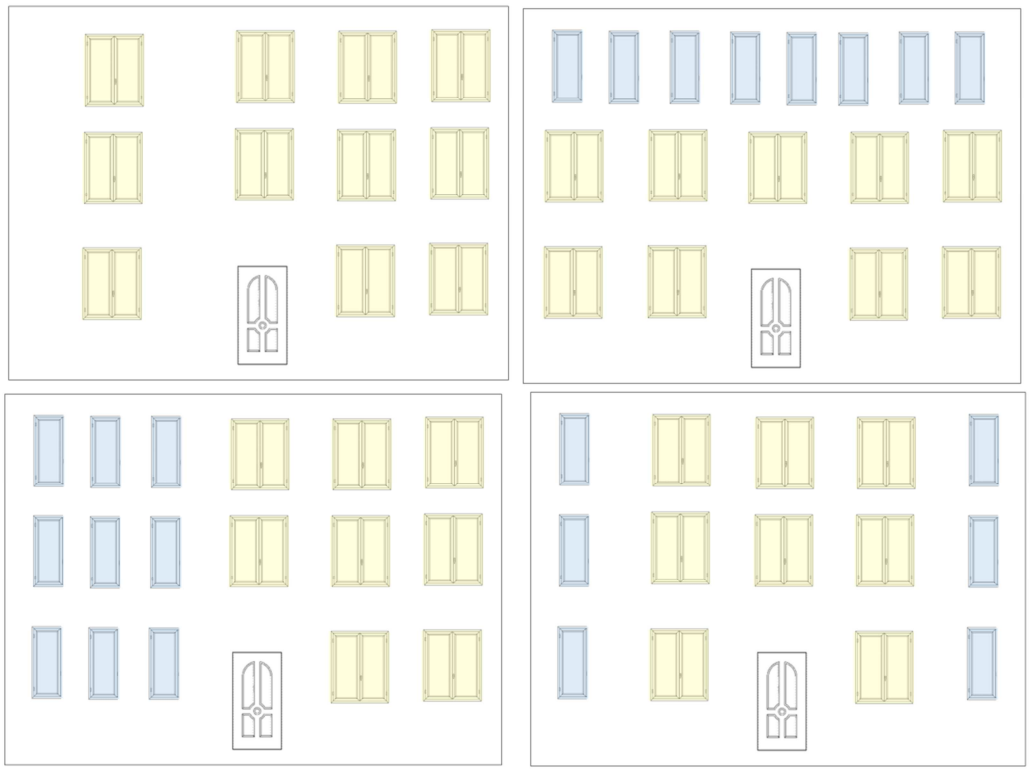


Fig. 4.2. Synthetic examples demonstrating some cases consisting of multiple periodic regions.

4.2. Repeated patterns detection

As summarized in Fig. 4.3, the developed methodology for RPD consists of two phases: *element grouping* and *structure regularity estimation.* The presented approach can be carried out to generate a hierarchical representation of the façade as a series of *basic-geometries* repeated into *multiple 2D periodic structures* (i.e., a *lattice*). After an initial estimate of the repeated element locations, the refinement and the recovery of missed locations undetected in the first step are performed by a voting scheme in which each location votes for lattices that are considered good fit for the object

distribution and represent the local periodicity in the vertical and horizontal directions.

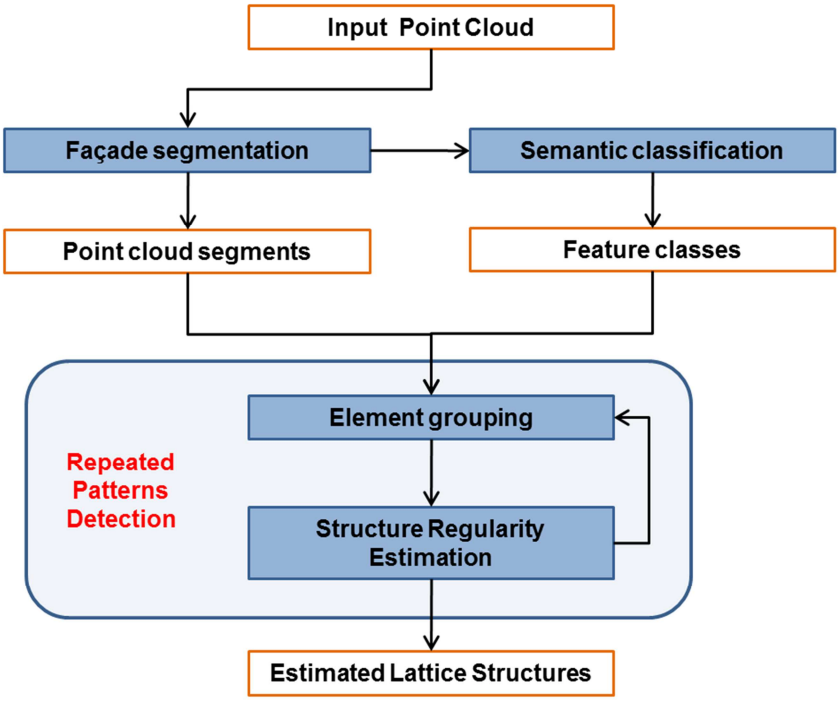


Fig. 4.3. Overview of the developed Repeated Patterns Detection (RPD) approach.

The developed procedure for RPD starts from the point cloud segments derived from façade segmentation (Chap. 1) and the associated feature classes obtained by semantic classification (Chap. 3). Façade segments are used to detect groups of similarity elements in the data during *element grouping* by means of a similarity measure. During this phase, similarity is evaluated between pairs of patches. In particular, similarities between objects into the same category are searched for. This means that similarities are not verified between objects belonging to different façade classes, (e.g. no similarity is sought between a door and a window). This reduces the computational time and prevents wrong regularity estimation. The *element grouping* step not only allows the detection of similar façade patches but also allows an initial estimation of the similarity transformation between those patches. The grouping is achieved by a geometric registration followed by a iterative bottom-up clustering of façade object pairs.

In the *structure regularity estimation* phase the parameters of the generative grid models of repeated patterns are estimated. This estimation is inspired by Generalized Hough Transform (Ballard 1981) and a lattice voting scheme (Pauly et al. 2008). In this phase, grouping information derived in the previous step is used to perform a global optimization towards alignment of repetitive façade elements by using Least

Squares. The final output is a series of lattices representing the local periodicity in the vertical and horizontal direction of each detected cluster of similar objects. Since to some extent the result of the regularity estimation is dependent on the element grouping, these steps are iteratively repeated until convergence

4.2.1. Element clustering

Once the identified planar clusters are subdivided into façade classes, the goal of element grouping is to gather similar objects together so as to identify repetitive elements. This step also provides the information needed to estimate structure regularity as described in Subsection 4.2.2. The grouping is achieved by a two-step clustering algorithm.

It can be assumed that the façade has a dominant planar structure, characterized by a flat dominant surface and with other façade's components having off-plane depth variations with respect to this plane, either positive (outwards) and negative (inwards). This assumption is generally valid for a large variety of modern building styles, in particular holds quite well for those built-up in the period 1950 - 1975.

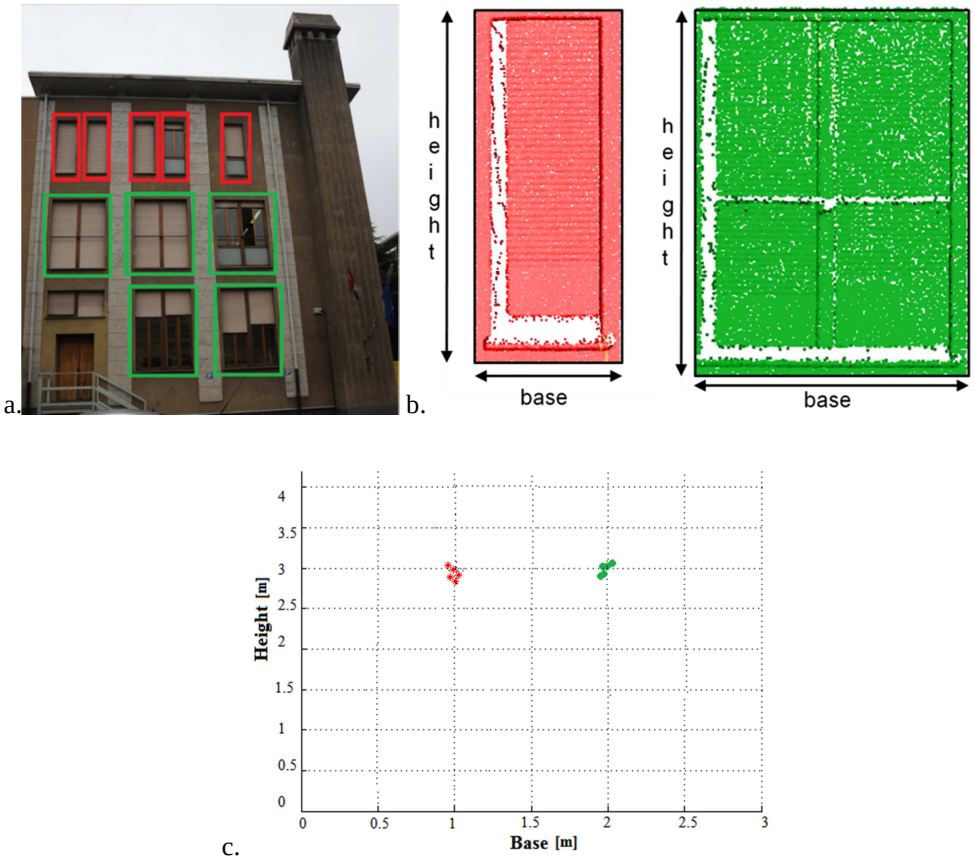


Fig. 4.4. Example of a façade with two different window types (red and green) (a-b); and (c) clustering results using base and height of the bounding box of detected objects.

First, elements classified in the same façade feature class (e.g., window, door, etc.) are clustered according to their shape (base and height of the bounding box). Indeed, façade objects belonging to the same feature may present significant geometric differences and follow different repetitive patterns. For example in Fig. 4.4, a façade with two different types of windows is shown. This first step is performed to have a rough clustering of similar objects and to reduce the number of misclassifications. Indeed, it is unlikely that objects having a significant difference in the bounding box shape present a high similarity between them.

Once two objects S_i and S_j belonging to the same façade feature are clustered together, they are aligned to measure the similarity between them. This task is performed by computing a rigid-body transformation T_{ij} (Fig. 4.5a) using a standard *Iterative Closest Point* (ICP) implementation (Besl and McKay, 1992). In particular, the transformation T_{ij} is restricted to be a translational one, along both directions of the façade plane, being the off plane shift negligible. Indeed, in this research only repeated structures that can be obtained by translating a base object are looked for. Similarity transformations involving scaling and rotations are not addressed.

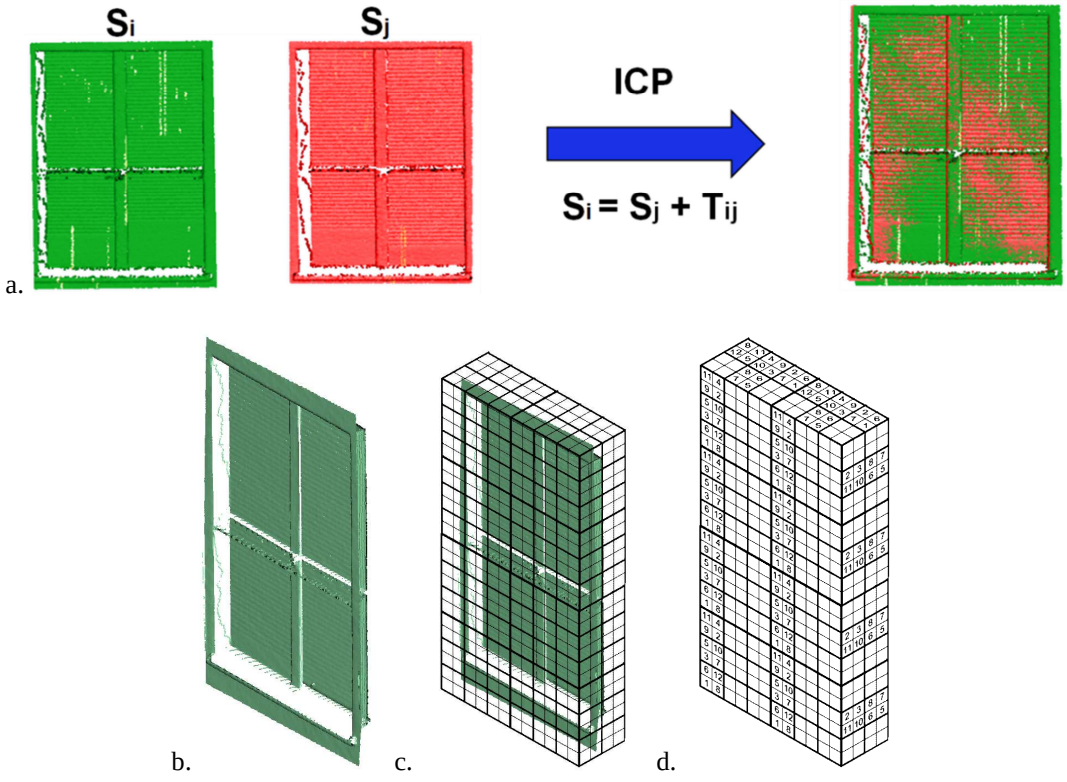


Fig. 4.5. Example of a calculation of SM. Two objects, S_i and S_j, are aligned with ICP (a). Volumetric. representation of a window: the original point cloud for object S_i (b), the voxel splitting (c) and the final tensor representation v_i (d).

Once S_i and S_j are aligned the *similarity measure* (SM) is evaluated. To tolerate poor quality input data the space of the overlapping region of the aligned slices is quantized and the similarity between S_i and S_j is calculated in this quantized space. Specifically, the aligned slices S_i and S_j are embedded into a volumetric grid whose size is determined by the bounding box of the overlapping region B between the two slices (Fig. 4.5b-c).

The grid resolution is fixed a little bit larger than the mean sampling distance in the point cloud. For each resulting voxel two functions v_i and v_j are defined to indicate the number of points contained in the voxel from S_i and S_j, respectively. At the end, the original point clouds are quantized into a tensor representation (Fig. 4.5d). The *SM* between slices S_i and S_j is defined as:

$$SM = \frac{\left[v_i - \overline{v_i}\right]^T \left[v_j - \overline{v_j}\right]}{\sqrt{\left\|v_i - \overline{v_i}\right\|^2} \sqrt{\left\|v_j - \overline{v_j}\right\|^2}} \tag{4.1}$$

where

$$v_i = \begin{bmatrix} v_i(1) \\ v_i(2) \\ \dots \\ \dots \\ v_i(K) \end{bmatrix}; \quad v_j = \begin{bmatrix} v_j(1) \\ v_j(2) \\ \dots \\ \dots \\ v_j(K) \end{bmatrix}; \quad \begin{aligned} \overline{v}_i &= \frac{1}{K}\sum_{k=0}^{K} v_i(k) \\ \overline{v}_j &= \frac{1}{K}\sum_{k=0}^{K} v_j(k) \end{aligned} \tag{4.2}$$

The adopted *SM* is a generalization of the *normalized linear correlation* coefficient (ρ) and in a similar way it may range from -1 (full inverse correlation) to +1 (full direct correlation). For this reason *SM* values close to +1 indicate high similarity between S_i and S_j, while in the case *SM* is close to zero or negative, they are assumed to be different each other. *SM* also supports partial matching of two slices, since the similarity is defined on the overlapping region of their aligned versions.

Once the similarity is measured for each pair of slices, the ones having the maximum similarity are automatically clustered by using a bottom-up method as far as no more clusters (C_1,C_2,…, C_n) can be created. The clustering process is stopped until *SM* is lower than a user-defined threshold (SM = 0.7 has been used in the experiments). In this way, elements with low similarity with respect to the others in the cluster are discarded, improving the robustness of the method.

4.2.2. Estimation of façade repeated patterns

The next step estimates the structure regularity for each group of previously detected similar elements (C_1,C_2,…, C_n). In this step, the elements of each set (C_i) are iteratively aligned and the new ones, based on the local periodicity of the lattice histogram (H), are assumed. The inputs of this step are the detected sets of similar façade elements (C_1,C_2,…, C_n) and their locations (L_1, L_2, …, L_n), while the output is a set of tuples ($S_{0,i}$, $G_{MN,i}$) where $S_{0,i}$ is the basic repeated element in the structure and $G_{MN,i}$ is a transformation group, having a lattice structure consisting of M rows and N columns, acting on $S_{0,i}$. The developed iterative procedure is shown in Fig. 4.6a for a single set of façade elements while details of a single step of the procedure are summarized in Fig. 4.6b.

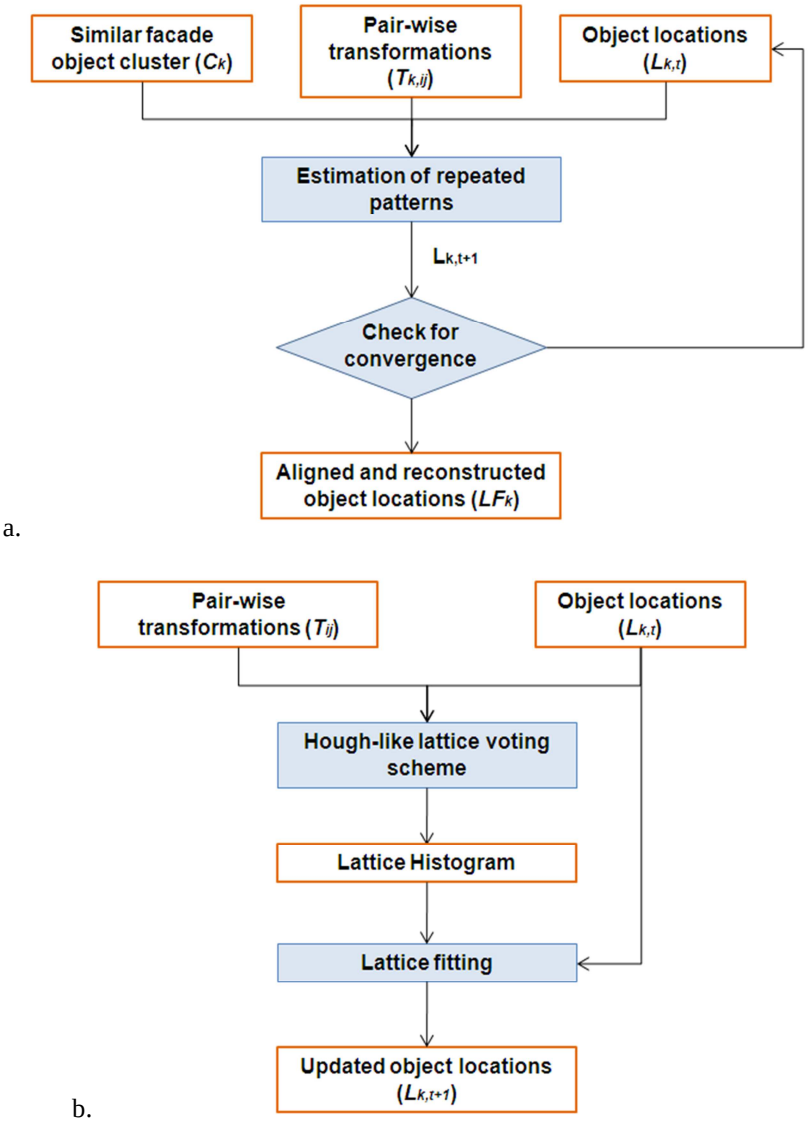


Fig. 4.6. Overview of the iterative repeated patterns estimation (a); and (b) workflow of a single step of repeated patterns estimation.

In other words, it is a set of Q x R 2D points arranged in a regular grid with variable spacing between rows and columns and the element l_{11} is the upper left element. Each of the $l_{ij} \in L$ can be considered as a '*seed*' or '*node*' of the lattice. In the cases where Q=0 or R=0, the lattice becomes one-dimensional.

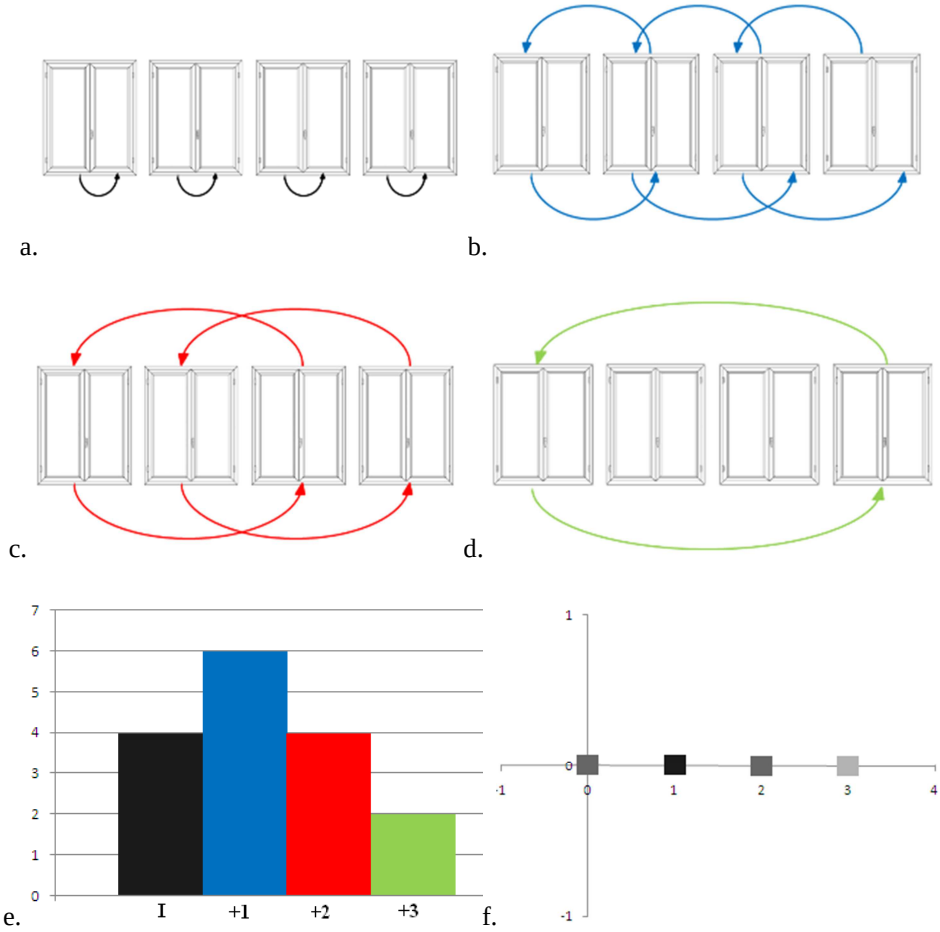


Fig. 4.7. Example of a simple 1D regular structure with all possible pairwise transformations (a-d) which form a characteristic cumulative pattern in the transformation histogram H (e-f).

The first step for the generation of the final transformation group G_{MN} is the construction of a lattice histogram (H) through a *lattice voting* procedure based on a Generalized Hough Transform voting scheme. The basic idea behind lattice voting is to consider similar façade object and to evaluate local lattices that are a good fit for data. In this way, it is possible to have a representation of the local periodicity of the region in which façade elements ($c_{i,1}$,$c_{i,2}$,…, $c_{i,k}$) belongs to. In addition, the aggregated information contained in the lattice can help in estimating updated object

locations c'_k. In particular, a *Hough-like lattice voting scheme* is used. In this scheme once a couple of similar object c_i, c_j is found the associated pair-wise transformation T_{ij} votes in a *Lattice Voting Space (Ω)*, see Fig. 4.7. The resulting histogram (H) has picks in locations that will be elements for a new iteration C'_i. Indeed, C'_i contains not only elements also contained in C_i but it may also include new points if there are peaks in locations where no element C_i in were found (Fig. 4.8). In this way, if the assumption of underlying multiple regular structures holds occluded or yet undetected features may be revealed.

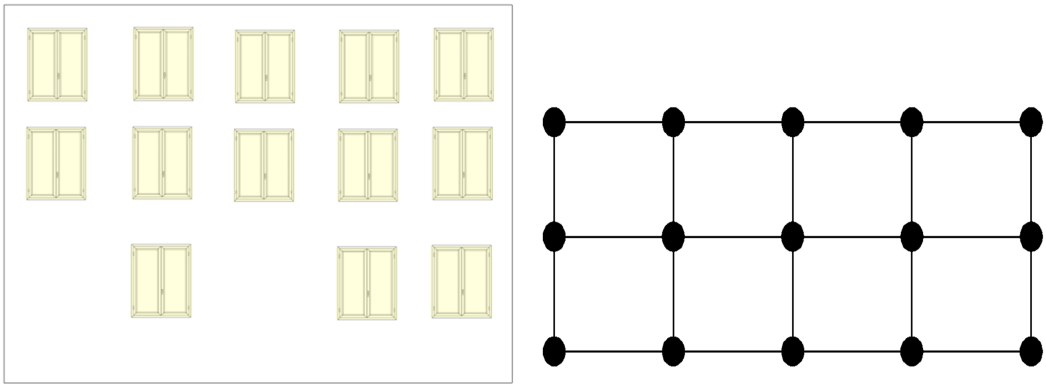


Fig. 4.8. Illustration of lattice voting: notice the reconstruction of the lower left and lower middle center.

Once the lattice histogram is set up a lattice structure is fitted to it. However, the set of pair-wise transformations T_{ij} may present different values due to noise in the model, local variations of sample position, and non-perfect alignment within the ICP registration phase. All these contributions lead to some inaccuracies in the transformation estimation. This dispersion of the transformations T_{ij} reflects in the Lattice Voting Space (Fig 4.9). In addition, some transformations may miss due to holes in the input data. For these reasons, the reliably detection of regular structures calls for a grid fitting approach that is robust against noise and holes. To this end, a global optimization method based on the work of Pauly et al. (2008) has been applied. However, while in that work the estimation procedure was aimed at detecting only rectangular lattices that were represented as the composition of two base vectors (one for row and the other for column description), this methodology takes into account a more general configuration considering also the possibility that column and row spacing may vary within the lattice.

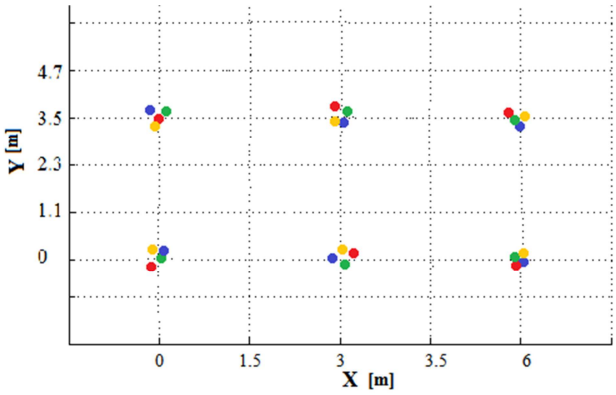


Fig. 4.9. Distribution of pairwise translation vectors for similar façade object pairs.

The unknown grid position for a lattice structure of M rows and N columns are represented by the row coordinates $Xg_{1,\,2,\ldots,\,i,\,\ldots,\,M}$ and column coordinates $Yg_{1,\,2,\ldots,\,i,\,\ldots,\,N}$ (Fig. 4.10) The input data are the set of pairwise transformations (Fig. 4.9) represented by a vector T_k (X_{Tk}, Y_{Tk}) and the location of detected features c_k. To find the unknown grid positions g_{ij}, an optimization scheme combining four *energy terms* is applied.

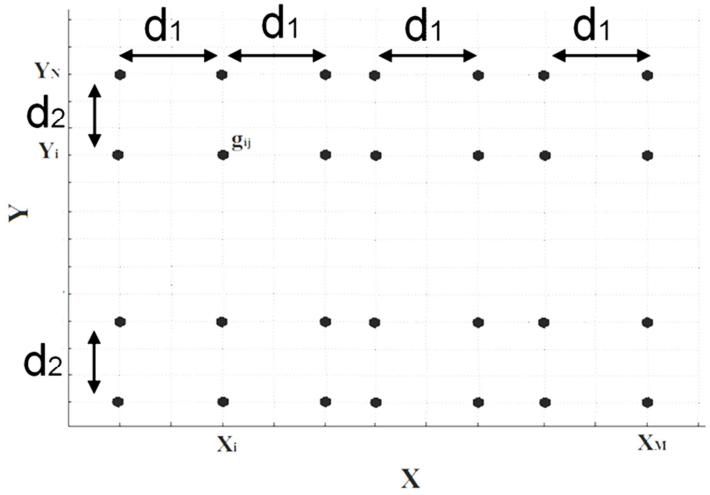


Fig. 4.10. Lattice estimation: unknown grid positions.

The first term takes into account the distance between the grid location g_{ij} to the closest feature location c_k (Fig. 4.11a):

$$E_{C\_1} = \sum_i \sum_j \alpha_{ij}^2 \left[\left(Xg_i - Xc_k \right)^2 + \left(Yg_j - Yc_k \right)^2 \right] \qquad 4.4$$

Similarly the second energy term takes into account the distance between the transformation T_k (X_{Tk}, Y_{Tk}) and the closest grid location $x(k)$ (Fig. 4.11b):

$$E_{C\_2} = \sum_{k=1}^{|T|} \beta_k^2 \left\| T_k - x(k) \right\|^2 \qquad 4.5$$

The continuous variable α_{ij} and β_k are weights measuring how reliably a grid location is mapped to a cluster centre and vice-versa. They are included as additional unknowns in the optimization process accounting for holes and outliers in the distribution of the feature locations. Indeed, values of α_{ij} and β_k close to zero indicate a hole or an outlier, respectively, while values close to 1 represent a reliable matching between transformation cluster and grid location.

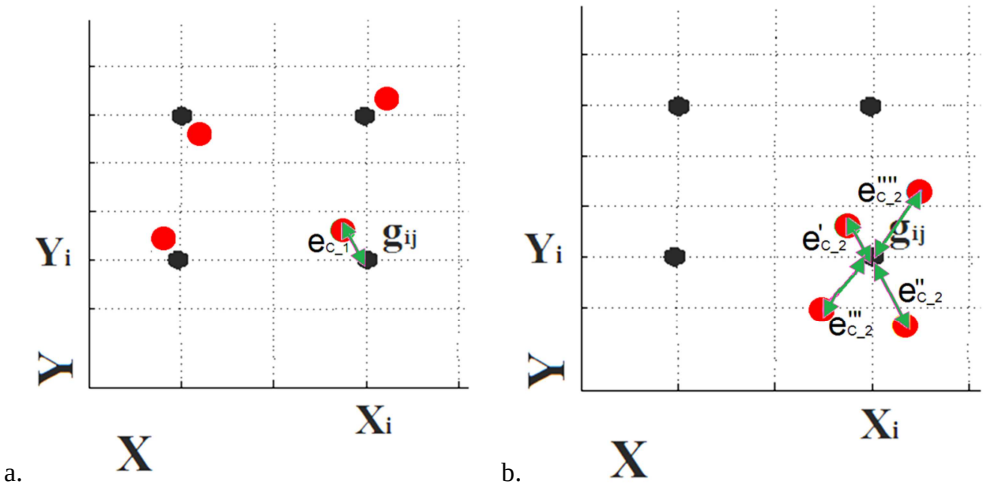


Fig. 4.11. Lattice estimation: example of contribution e_{c_1} to the first energy term (a) and e_{c_2} to the second one (b).

The last two energy terms are aimed at maximizing the number of valid correspondences between grid location and cluster centres:

$$E_\alpha = \sum_i \sum_j (1-\alpha_{ij}^2)^2 \quad and \quad E_\beta = \sum_k (1-\beta_{ij}^2)^2 \qquad 4.6$$

The combination of the four energy terms gives the total objective function to be minimized:

$$E = \gamma\left(E_{C\_1} + E_{C\_2}\right) + \left(1-\gamma\right)\left(E_{\alpha} + E_{\beta}\right) \qquad 4.7$$

where the coefficient γ balances the two energy terms. In the tests carried out, $\gamma = 0.8$ has been adopted.

In order to find repeated similarity in the lattice structure, the spacing between consecutive columns and rows are calculated and clustered. In the case some grids present a similar spacing, in the minimization process these additional constraints will be enforced. For example, in the case the spacing between columns Yc_j-Yc_{j+1} and Yc_{j+1}-Yc_{j+2} are clustered together, the following constraint equation is added to the minimization:

$$\left(Yc_{i+1} - Yc_i\right) - \left(Yc_{i+2} - Yc_{i+1}\right) = 0 \qquad 4.8$$

To minimize the above objective function an iterative Gauss-Newton Least Squares technique (Triggs et al. 2000) is applied. As can be seen, the minimization process is non-linear and a set of initial estimates of the size of the lattice in terms of rows and columns is necessary as well as the initial values for the grid locations g_{ij} and the weights α_{ij} and β_k.

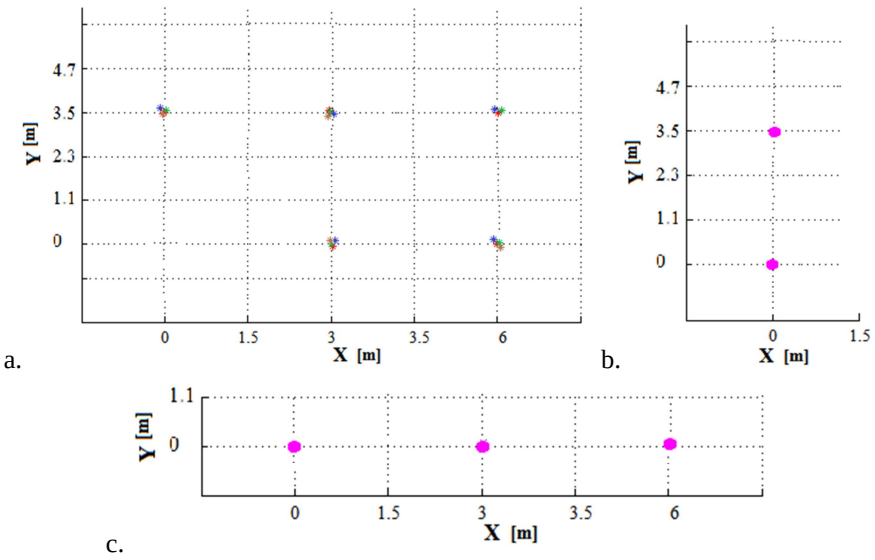


Fig. 4.12. Determination of first iteration approximate values. The pairwise transformations initially estimated (a) are then clustered along the two dominant directions of the lattice (b-c) giving the initial estimates of the lattice nodes.

In the first iteration, approximate values are determined by clustering in H the pair-wise transformations T_{ij} along the two dominant directions of the lattice (Fig. 4.12). Detected cluster centres are used as approximations of the lattice row ($Xg_{1,\,2,\ldots,M}$) and column ($Yg_{1,\,2,\ldots,N}$) coordinates. The correspondent weights α_{ij} and β_k are initialized to one, since no a priori knowledge on holes and outliers is assumed.

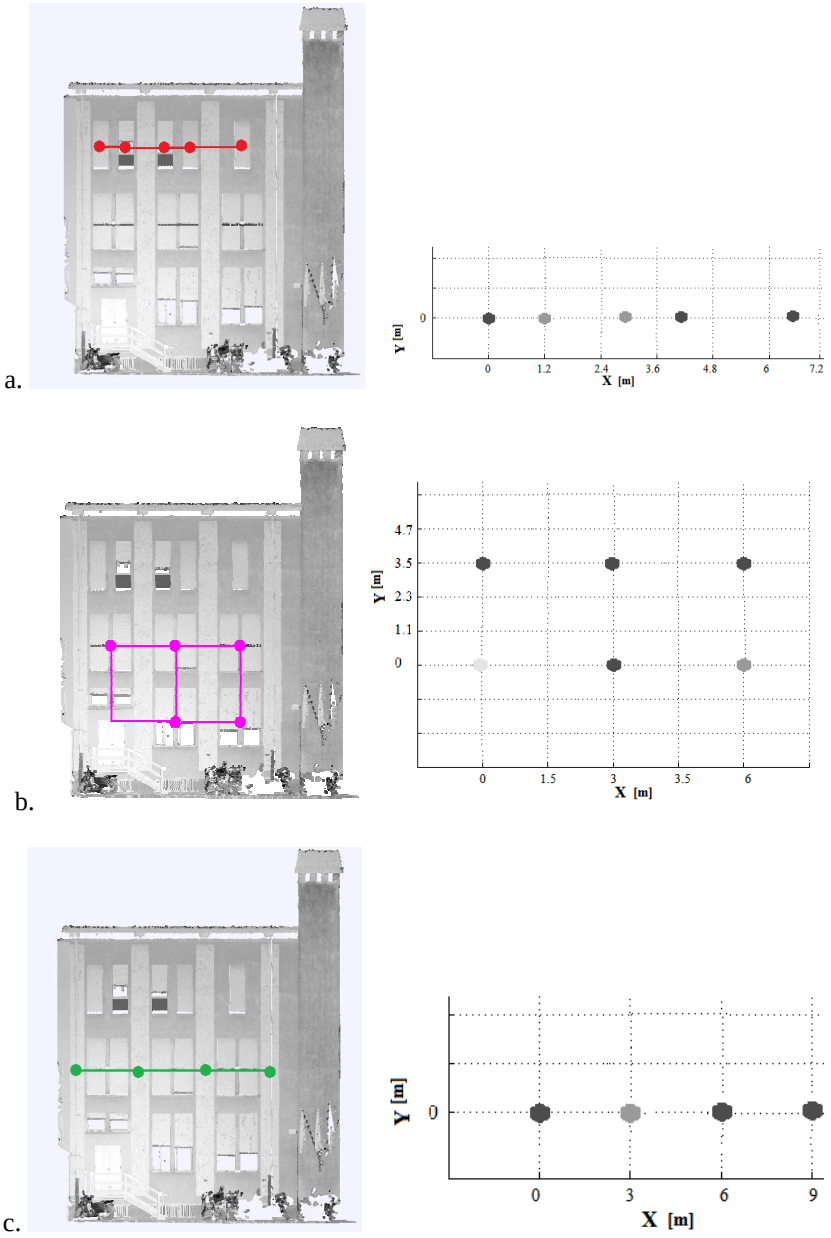


Fig. 4.13. Regularity structure superimposed to the point cloud (left) and lattice estimation (right). Lattices centres are coloured according to the associated weight (white for 0 and black for 1).

Once the lattice is estimated, the algorithm operates iteratively by performing a new element grouping in correspondence of each node of the lattice. In particular, in correspondence of nodes which are labelled as missing (holes) during the lattice fitting a feature is searched. In the case SM confirms the assumption this new feature is included in the similar feature cluster C_i and a new iteration takes place. The process halts when no changes in the lattice nodes are observed between two successive iterations. In particular, the presented algorithm can efficiently detect repeated features with non-rectangular lattice patterns (Fig. 4.13a), where the spacing between rows and columns may vary, and may also deal with missing elements (Fig. 4.13b). In the second case it is possible to observe that the weight to the bottom-left element is close to zero.

PART B

Chapter 5
Experimental tests on real façades

In Part A of this dissertation a detailed description of the façade modelling procedure has been given. This chapter examines the efficiency and accuracy of reconstruction methods. Several examples of common façades from buildings constructed during 1950 - 1975 are presented. The whole procedure from segmentation of raw point cloud to repeated pattern detection is applied to every case study. Evaluation of results is made on the basis of common metrics whose description is given in Section 5.1.

All the above-mentioned steps for façade modelling are fully implemented in *Matworks Matlab*® environment. Exception has been made with the visualization of final digital models. For this purpose, the obtained faced models are exported in different file formats: (i) DXF and PLY format for vector model of the façade and (ii) CityGML and IFC file for the semantic enriched model.

5.1. Performance evaluation framework

Overall accuracy and precision are used to evaluate the performances of the presented façade modelling procedure. In particular, quality evaluations are performed for each step of the procedure by following the same scheme for each case study:

- *façade segmentation*: the point cloud of the building façade has been manually classified into planar clusters and used as ground truth façade;
- *geometric reconstruction*: the presented datasets were manually vectorized, starting from the point cloud, and façade breaklines were identified. Those breaklines are used for evaluating the accuracy of automatically detected ones;
- *object classification*: points of each type of façade element have been manually selected from the original point clouds and have been compared with object classification results; and
- *façade regularity detection*: presence of repeated pattern in the façade are manually selected and compared with the ones detected in an automated way.

5.1.1. Façade segmentation

The performance measurement of a segmentation algorithm is conducted on the basis of the analysis of resulted segments. This task requires a set of benchmarking data against which the resulted segments can be compared.

Hoover et al (1996) provided methodology to evaluate the result of segmentation, originally designed for range images. A similar framework was adopted by Geibel and Stilla (2000) for the comparison of different procedures in segmentation of ALS and by Nyaruhuma (2007) for the performance evaluation of different algorithms for detecting roof faces in 3D point clouds.

The underlying principle is based on the comparison of the resulted segments with the corresponding reference segments and, and on the evaluation of how many common points belong to both datasets. Once the correspondences are established, correct segments are found. A segment is correctly detected if the majority of points in the reference segment are also labelled as a single segment in the result. Correct segmentation can be expressed as:

$$M_r = \frac{N_{ro}}{N_r} \tag{5.1}$$

where M_r is the ratio of intersection points (N_{ro}) to the total number of points in reference segment (N_r). A pair of segments in the benchmarking data and in segmentation results are classified as an instance of correctly detected segment if the correspondence percentage is greater than certain tolerance value (T):

$$M_r > T \tag{5.2}$$

In the cases where one segment in reference data (S_r) does not correspond to any segmented element, three instances may occur:

- **Over-segmentation:** this is the case where one segment in reference data (S_r) is represented by more segments ($S_{o1,}$ S_{o2},…, S_{on}) in the output of the algorithm. In this case, the total number of intersection points is the sum of intersection points in many S_o segments:

$$M_{rove} = \frac{N_{ro1} + N_{ro2} + ... + N_{ron}}{N_r} \tag{5.3}$$

where $N_{o1,}$ N_{o2},…, N_{on} corresponds to the total number of points in segments $S_{o1,}$ S_{o2},…, S_{on} respectively. If a reference segment is classified as over-segmented the number of detected segments ($S_{o1,}$ S_{o2},…, S_{on}) constituting it can be determined.

- **Under-segmentation:** if one segment (S_o) in the output of the segmentation process intersects with more than one reference segments ($S_{r1,}$ S_{r2},…, S_{rn}), this result in under-segmentation. This can be considered as the case of insufficient separation of multiple planar surfaces. Likewise in over-segmentation, the total number of intersection points is the sum of the intersection points in many S_r segments:

$$M_{ou} = \frac{N_{r1o} + N_{r2o} + ... + N_{rno}}{N_o} \quad 5.4$$

The values of $N_{r1,}$ N_{r2},…, N_{rn} are the total number of points in segments $S_{r1,}$ S_{r2},…, S_{rn} respectively. If a detected segment is classified as under-segmented, the number of real segments ($S_{r1,}$ S_{r2},…, S_{rn}) erroneously clustered together can be found; and

- **Missed segment:** a reference segment (S_r) is classified as missed segment if it does not have any correspondence with the obtained segments.

In addition, a segment (S_o) in the output of the segmentation process is defined as 'noisy segment' if the segment cannot be classified in any of the previous categories.

Starting from the segment classification, also the following parameters can be computed:

- **Commission error (I Type):** the probability of erroneously detecting a plane; it is evaluated as the ratio between the number of wrong segments and the total number of real planes; and
- **Omission error (II Type)**: the probability that of a real plane is undetected; it is evaluated as the ratio between undetected planes and the total number of real planes.

The performance of the developed algorithm has been assessed on the basis the framework described above. Results will be provided per each of the case studies described in Section 5.2.

5.1.2. Geometric reconstruction

The results achieved by the proposed approach need to be checked in order to quantify the geometrical accuracy of the obtained vector models. In order to do that a manual reconstruction of the same façade was performed (Nex and Rinaudo 2009) which is considered in the literature the most precise method to vectorize a point cloud. For this reason, an experienced operator performed the vectorization of the building model staring from the point cloud. Manual models are then compared with the automatically generated ones. In particular, the detected breaklines are compared with the manually benchmarking model. Firstly, the accuracy of the reconstructed breaklines was derived by comparing the ground truth position of each line with the position estimated by the automatic algorithm. In particular, for each edge the absolute modelling error is defined as the absolute magnitude of the difference between the ground truth and the model position. An edge is considered as correctly detected if the distance between the manually generated edge and the closest automatically generated edge is lower than a predefined threshold T. Then the reliability of break-line evaluated by comparing the number of commission and omission errors in a similar way to that one described in Subsection 5.1.1.

5.1.3. Object classification

Overall classification accuracy, precision and recall are used to evaluate the classification performance.

At first, actual points of façade point clouds are manually classified into façade elements to derive a reference dataset. Then the actual classification results and those predicted from the automatic classification method are evaluated by means of visual interpretation and a confusion matrix is created, where each row represents the instances in a predicted class, and each column represents the instances in an actual class. From the confusion matrix, the overall classification accuracy can be retrieved, as the sum of correctly classified segments divided by the total number. In a similar way, omission and commission errors can be defined. The commission error reflects the probability that, given an object from a certain predicted class, it does not belong to the same class in the reference data. And the omission error measures the probability that, given an object of the reference data, it has not been correctly classified in the prediction class. In short, commission is a measure of the exactness, whereas omission is a measure of the recall. The lower these two values, the more excellent the classification performs.

5.1.4. Regularity detection

In order to evaluate the performance of the regularity detection algorithm, a manual identification of repeated object is performed and compared with the detected lattices. In particular, the fractions of commission errors (wrong detections) and omission

errors (missing detections) in the detection of lattice nodes have been evaluated with respect to manually detected nodes, which have been assumed as benchmarking values.

5.2. Experimental tests

This section presents the application of the developed modelling procedure to some real building façades. In particular, these test are focused to assess the reliability of the approach under different operational conditions and façades typologies. All experiments have been performed on buildings in different campus of the Politecnico di Milano University. In total, five data sets have been used: (i) 'D'Oggiono building,' (ii) 'Courtyard no.1,' (iii) 'Courtyard no.2,' (iv) 'Building no. 34' and (v) 'Nave building'.

5.2.1. 'D'Oggiono building'

This section provides the experiment details of 'D'Oggiono building'. A portion of the Lecco campus headquarter was scanned by using a TLS FARO-FOCUS 3D in 3 scans (Fig. 5.1). Some technical specifications of the laser scanner adopted are reported in Tab. 5.1.

Faro Focus 3D technical specifications		**Riegl LMS- 420i technical specifications**	
Range measurement mode	Phase-shift	**Range measurement mode**	Time-of-flight
Operational range	0.6 – 150 m	**Operational range**	2.0 – 1000 m
Angular resolution	0.009°	**Nominal accuracy**	± 10 mm
Measurement speed	120.000 – 976.000 points/sec	**Nominal precision**	± 4 mm c
Precision at 10 m	0.6 mm (90% reflectivity) 1.2 mm (10% reflectivity)	**Acquisition rate**	3000 – 9000 pts/sec
Precision at 25 m	0.95 mm (90% reflectivity) 2.2 mm (10% reflectivity)	**Vertical field of view**	80 °

Tab. 5.1. Technical specifications of the adopted TLS instrument.

Multiple targets have been set up over the area, so that they could be detected by SCENE 3D® (the post-processing software of FARO scanners) as control points for registration. The final registration accuracy, evaluated in terms of sigma naught, has resulted as 3.1 mm. The laser data acquisition has taken 45 minutes in total. Actually, each single scan only has taken 8 minutes while the remaining time has been spent on placing targets and for the instrument setup. The average point density is about 20 thousand points per square meter on the walls. In total, 2 building façades were selected for the experiment (Fig. 5.1) and some results are shown in the previous

chapters. This dataset includes façade elements such as walls, doors, windows, roofs and stairs. Some occlusions are present in the lower part of the façade due to some cars and bikes that could not be removed during scanning.

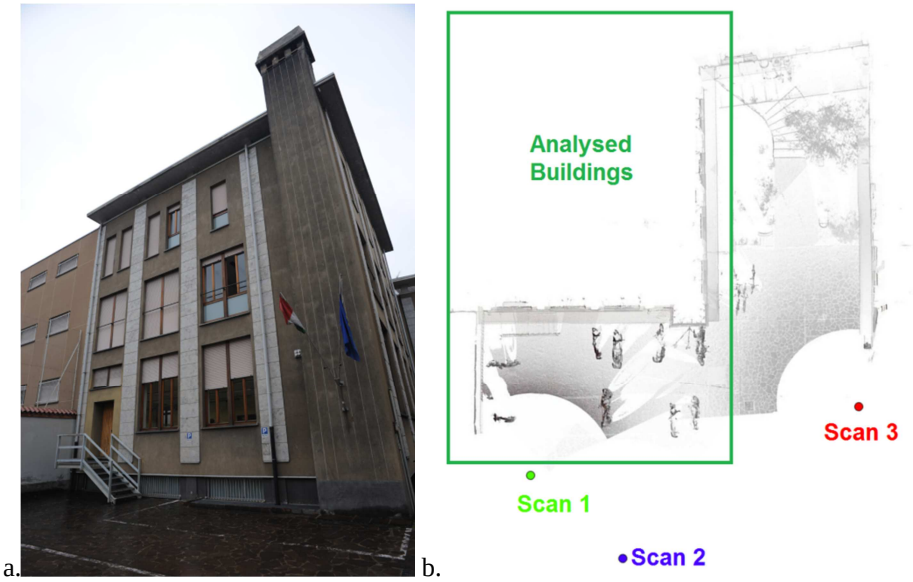


Fig. 5.1. 'D'Oggiono building' dataset: (a) a picture of the analysed building; and (b) the scan acquisition scheme.

First step of the developed methodology is façade segmentation. The parameters presented in Tab. 5.2 were used in the processing.

	'D'Oggiono Building' 'Courtyard no.1' 'Courtyard no.2' 'Building no. 34'	**'Nave Building'**
RANSAC plane threshold ε	1 cm	5 cm
RANSAC normal threshold α	20 °	20 °
Bitmap cell size β	1 cm	2 cm
RANSAC dominant line threshold ε	0.7 cm	1 cm

Tab. 5.2. Parameters used for façade modelling for the analysed datasets.

These parameters have been chosen taking into consideration the instrument accuracy and the scan registration statistics. Indeed, a too restrictive selection of the RANSAC plane threshold ε fit the limitation given by registration accuracy. At the same way the cell size β was selected considering the average point density. Indeed, a square meter in the bitmap is represented by 10,000 cells while the point density is about 20,000 points. On the other hand, these parameters are quite restrictive with respect to the ones presented in literature.

The segmentation results are presented in Fig. 5.2. A visual observation of this result shows that the algorithm is able to detect all planar surfaces present in the data. Furthermore, larger surfaces (e.g., walls and roofs) as well as smaller surfaces (e.g., some parts of windows frames) are extracted.

a. b.

Fig. 5.2. 'D'Oggiono building' segmentation results: (a) original point cloud; and (b) segmentation results, each detected segment is represented with a different colour.

It can be observed that there are many small segments near the ground, which are mainly due to the presence of other objects like cars. A more detailed data analysis can be obtained by evaluating the parameters described in Subsection 5.1.1. Results are summarized in Tab. 5.3 and Tabs. 5.4 - 5.5. In particular, Tab. 5.4 reports, for different tolerances (T), the number of reference segments that are either correctly detected or are instances of bed segmentation. Tab. 5.5 reports instead results for the predicted planes, indicating the correctly detected segments and the fraction of segments contributing to over-, under-segmentation and noisy elements.

	'D'Oggiono building'	**'Courtyard no.1'**	**'Courtyard no.2'**	**'Building no. 34'**	**'Nave Building'**
Automatically detected Planes	127	119	119	21	650
Manually labelled Planes	120	112	108	22	605

Tab. 5.3. Segmentation results for for the analysed datasets.

Tolerance (%)	**'D'Oggiono building' results for reference planes**					
	No. of correct segmentation	**No. of over-segmentation**	**No. of under-segmentation**	**No. of missed segment**	**Commission error (%)**	**Omission error (%)**
0.8	120	0	0	0	5.5	0.0
0.85	119	1	0	0	6.3	0.8
0.9	117	3	0	0	7.9	2.5
0.95	112	7	0	0	11.8	6.7

Tab. 5.4. 'D'Oggiono building' results for reference planes.

Tolerance (%)	'D'Oggiono building' results for detected planes			
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of noise segment
0.8	120	0	0	7
0.85	119	2	0	6
0.9	117	6	0	4
0.95	112	15	0	0

Tab. 5.5. 'D'Oggiono building' results for detected planes.

The result reveals that the numbers of correctly detected segments decreases as the tolerance increases as expected. However, the numbers of correctly detected segments with different tolerance classes are quite similar up to T = 0.9. This means a high reliability of the segmentation. In addition, the algorithm is able to detect significant numbers of correct segments even at the highest tolerance of T = 0.95 showing satisfactory results in terms of both omission and commission errors. The robustness of the method is also proved by the fewer instances of missing and noisy segments. This means that there is a strict matching between reference and predicted results. Some instances of over-segmentation are mainly due to non-conformance of some wall segments to planar surface assumption. No instances of under-segmentation in the dataset are observed. The effectiveness of the method can be observed by comparing results obtained with a 'naive segmentation' approach based on a simple sequential RANSAC plane estimation (Tab. 5.6). In this case, it is possible to observe that the numbers of correctly detected segments is comparable with over and under segmented instances and a high number of missing and noisy segments is observable. In addition, their number rapidly grows when the tolerance is restricted.

Tolerance (%)	'D'Oggiono building'						
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of missed segment	No. of noise segment	Commission error (%)	Omission error (%)
0.5	50	34	32	9	34	66.7	60
0.6	48	34	32	11	36	68.0	61.6
0.7	40	30	31	24	49	73.3	68
0.8	37	24	28	36	61	75.3	70.4
0.9	35	20	22	48	73	76.7	72
0.95	10	12	10	93	118	93.3	92

Tab. 5.6. 'D'Oggiono building' results for a 'naive segmentation' sequential RANSAC implementation.

Starting from the derived segments the contour points and extracted and then the edge smoothing is performed. The parameters used in this phase are summarized in Tab. 5.2. The obtained building model in CAD format is presented in Fig. 5.3.

a. b.

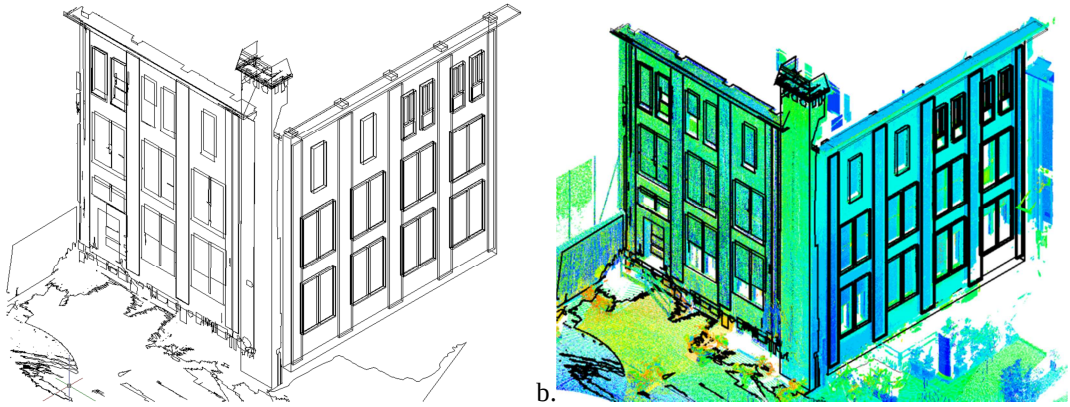


Fig. 5.3. Geometric reconstruction of 'D'Oggiono building' dataset: (a) final 3D digital model of the façade; and (b) façade model with overlaid point cloud.

The comparison of this model with the one obtained by manual modelling is summarized in Tab. 5.7 and Fig. 5.4.

	'D'Oggiono building'	**'Courtyard no.1'**	**'Courtyard no.2'**	**'Building no. 34'**	**'Nave Building'**
Automatically extracted Breaklines	505	524	457	106	650
Manually derived Breaklines	530	536	478	110	605

Tab. 5.7. Breaklines results for the analysed datasets.

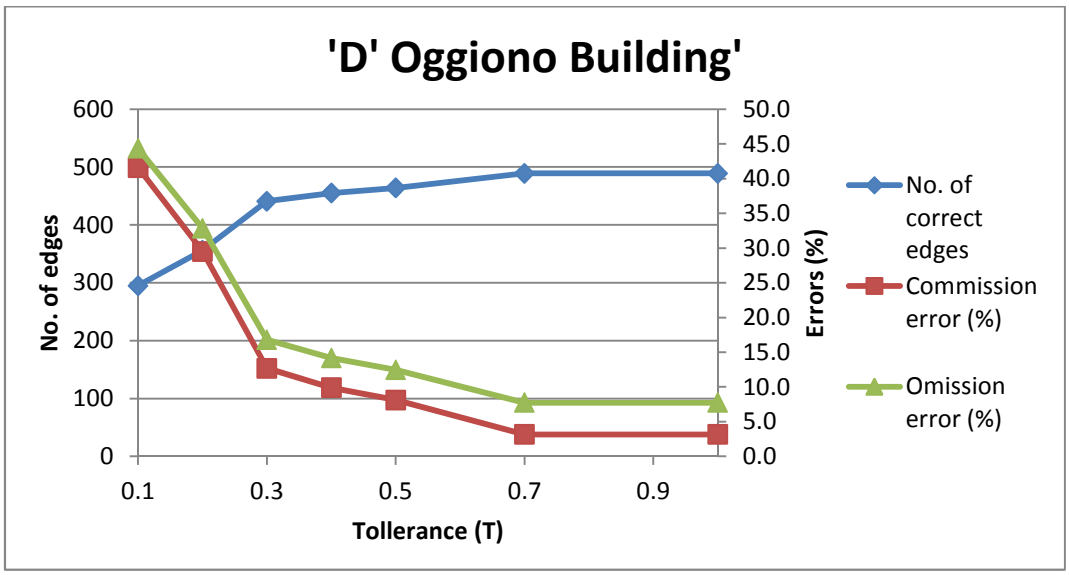


Fig. 5.4. Geometric reconstruction analysis for 'D'Oggiono building' dataset.

As expected, the numbers of correctly detected edges decreases as the tolerance increases. However, this decrease presents a significant discontinuity in

correspondence of T = 3.0 mm. This means that the accuracy of the detected edge is in this order of magnitude. However, an important element needs to be observed. Manual modelling of a point cloud is indeed influenced by human interpretation and the definition of breaklines with accuracy higher than 2.0 – 3.0 mm is almost impossible also for a skilled operator.

In Figure 5.5 the detected repeated patterns are presented. In particular, it is possible to observe that the presented algorithm can efficiently detect repeated features with non-rectangular lattice patterns, i.e. variable spacing between columns (Fig. 5.5d). In addition, the method may also deal with missing elements (Fig. 5.5c). In this second case it is possible to observe that the weight to the bottom-left element is close to zero meaning that the element is effectively missing.

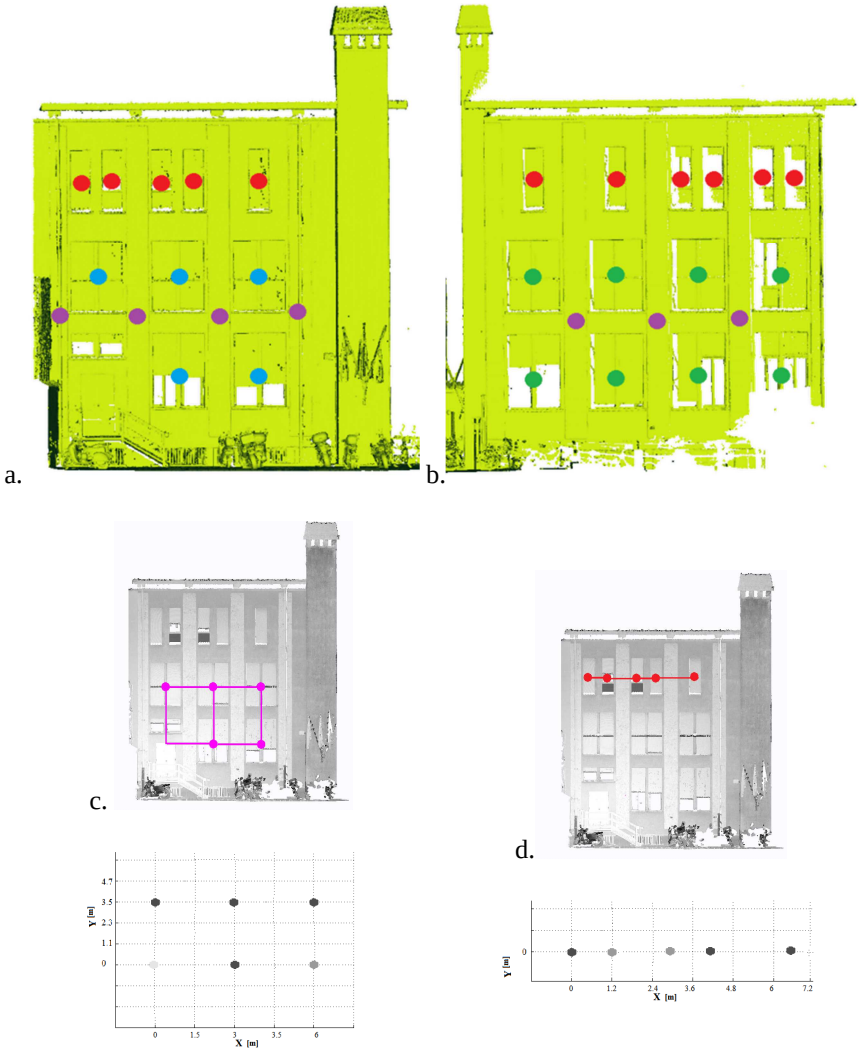


Fig. 5.5. Results of RPD algorithm for 'D'Oggiono building' dataset: the repeated pattern for the two analysed façades superimposed to the point cloud (a-b); and two examples of detected lattices (c-d). Lattices centres are coloured according to the associated weight (white for 0 and black for 1).

Classification results lead to the generation of the semantic rich model shown in Fig. 5.6.

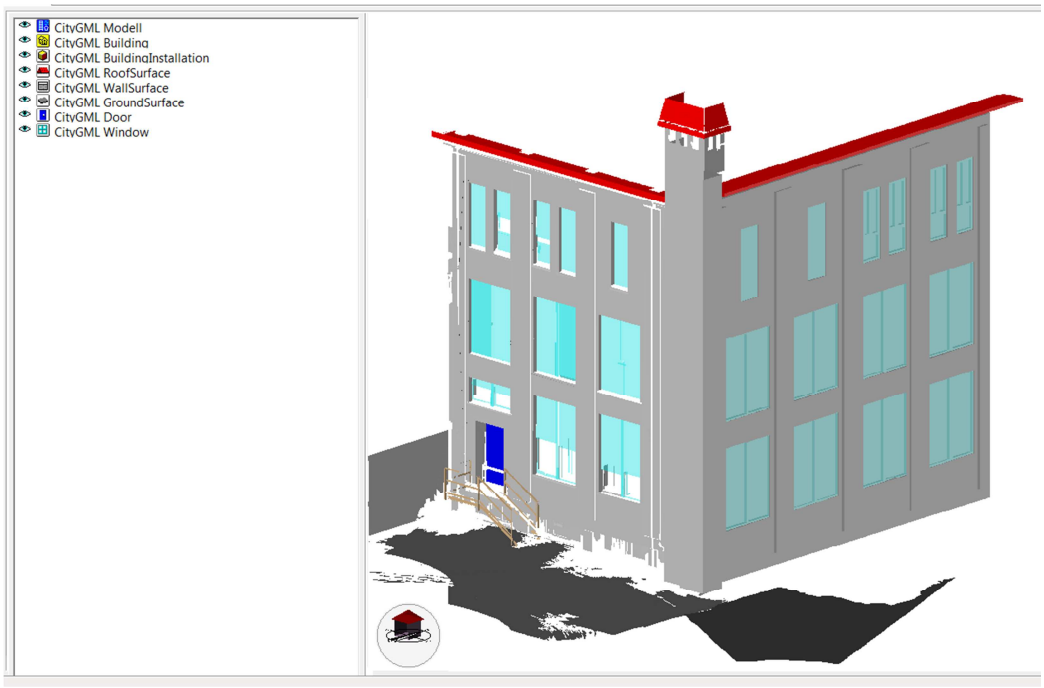


Fig. 5.8. Semantic reach model of the 'D'Oggiono building' in CityGML format. The model is visualized using the FZKViewer developed by the Karlsruhe Institut fur Technologie.

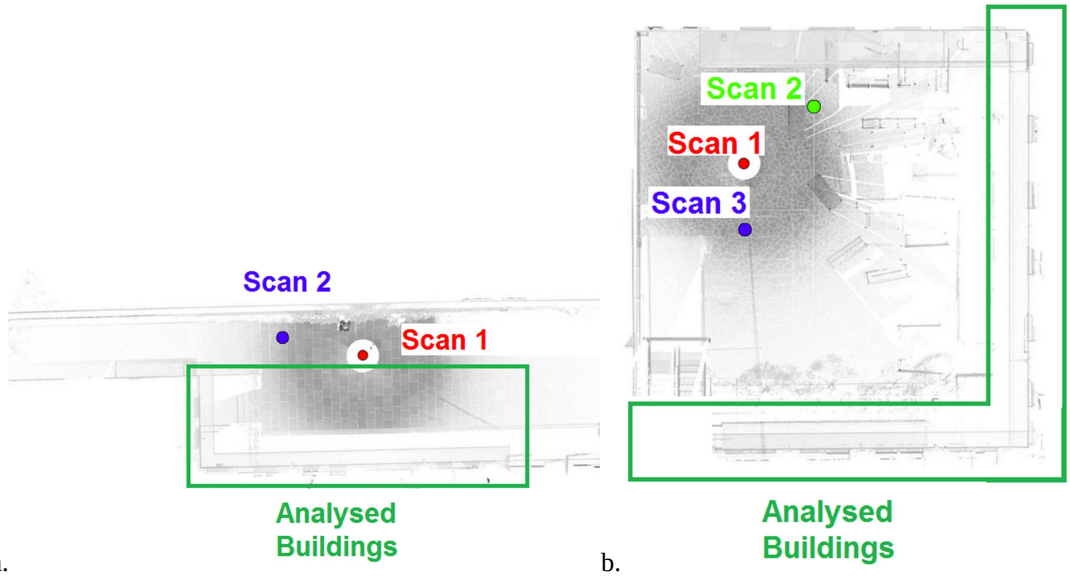
'D'Oggiono building'		Reference								
		No. of wall	**No. of roof**	**No. of sidewall**	**No. of window**	**No. of door**	**No. of ground**	**No. of attachment**	**Total**	**Commission error [%]**
Classified	**No. of wall**	20	0	0	0	0	0	1	21	4.8
	No. of roof	1	6	0	0	0	0	0	7	16.7
	No. of sidewall	0	0	40	0	0	0	0	40	0.0
	No. of window	0	0	0	30	0	0	0	30	0.0
	No. of door	0	0	0	0	1	0	0	1	0.0
	No. of ground	0	0	0	0	0	1	0	1	0.0
	No. of attachment	0	0	0	0	0	0	20	20	0.0
	Total	21	6	40	30	1	1	21	120	
	Omission error [%]	4.8	0.0	0.0	0.0	0.0	0.0	4.8		

Tab. 5.8. Confusion matrix for 'D'Oggiono building' dataset.

The confusion matrix is shown in Tab. 5.8. The overall classification accuracy, which is calculated as the sum of the correct classified objects divided by the total number is 98.3%. In addition a low confusion can be observed in Tab. 5.8. Commission errors are zero for five classes meaning that most of detected façade objects correctly fall down their actual classes. Also commission errors are quite low, exception made for roof objects (even if lower than 20%). Indeed, a vertical roof element is erroneously classified as wall object. This is given by the fact that rules for roof elements assume a roof cannot be vertical.

5.2.2. 'Courtyard no.1 and no.2'

A second test was performed on two different courtyards which were acquired with a TLS FARO-FOCUS 3D. The scanned scenes present a quite high number of clutters generating occlusions on the building façades and resulting in large missing parts. Furthermore, façades are constituted by different façade objects, e.g. walls, roofs, parts of windows frames. The average point density of the 'Courtyard' datasets is about 25 thousand points per square meter on the walls. Four building façades were modelled. In Fig. 5.9 the acquisition schemes for the two datasets are presented. In particular, the two datasets were processed in an independent way by using parameters summarized in Tab. 5.2.



a. b.

Fig. 5.9. The scan acquisition schemes for 'Courtyard no.1' (a) and 'Courtyard no.2' (b) datasets.

A visual representation of segmentation results for both datasets is shown in Fig. 5.10, while segmentation evaluation results are summarized in Tabs. 5.3 – 5.9 – 5.10 – 5.11 – 5.12.

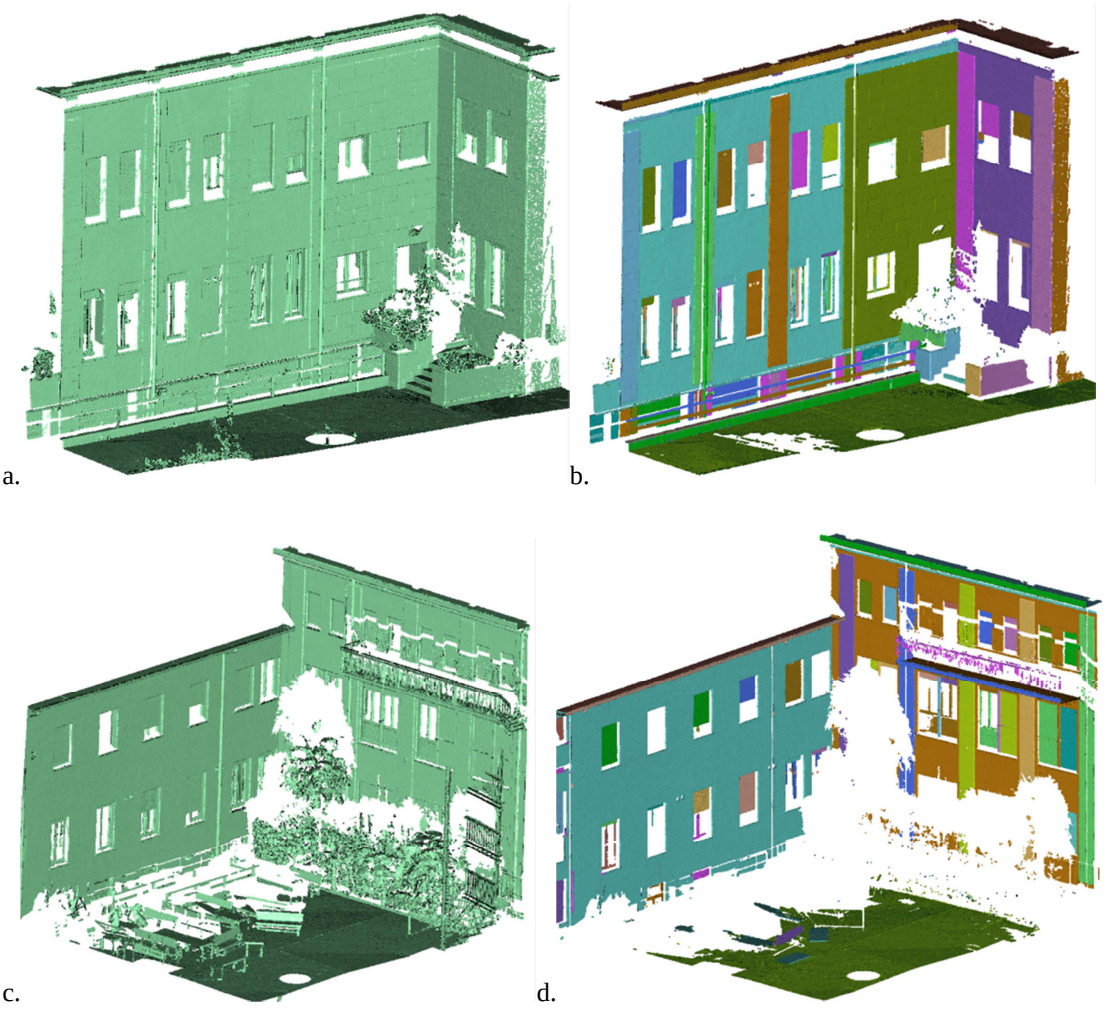


Fig. 5.10. 'Courtyard no.1'and 'Courtyard no.2' (top and bottom row respectively) segmentation results: (a-c) original point clouds; and (b-d) segmentation results, each detected segment is represented with a different colour.

The result are similar for both datasets. In particular, robustness of the method is confirmed by the high number of correct segments wile few instances of missing and noisy segments are observed.

Tolerance (%)	'Courtyard no.1' results for reference planes					
	No. of correct correct segmentation	No. of correct over-segmentation	No. of correct under-segmentation	No. of correct missed segment	Commission error (%)	Omission error (%)
0.8	110	0	0	2	7.6	1.8
0.85	108	2	0	2	9.2	3.6
0.9	107	3	0	2	10.1	4.5
0.95	102	8	0	2	14.3	8.9

Tab. 5.9. 'Courtyard no.1' results for reference planes.

Tolerance (%)	'Courtyard no. 2' results for reference planes					
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of missed segment	Commission error (%)	Omission error (%)
0.8	106	0	0	0	10.9	1.9
0.85	106	0	0	0	10.9	1.9
0.9	103	3	0	0	13.4	4.6
0.95	98	8	0	0	17.6	9.3

Tab. 5.10. 'Courtyard no.2' results for reference planes.

Tolerance (%)	'Courtyard no. 1' results for detected planes			
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of noise segment
0.8	110	0	0	9
0.85	108	4	0	7
0.9	107	7	0	5
0.95	102	17	0	0

Tab. 5.11. 'Courtyard no.1' results for detected planes.

Tolerance (%)	'Courtyard no.2' results for detected planes			
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of noise segment
0.8	110	0	0	9
0.85	108	4	0	7
0.9	107	7	0	5
0.95	102	17	0	0

Tab. 5.12. 'Courtyard no.2' results for detected planes.

The comparison of derived building models with the ones obtained by manual modelling are summarized in Tab. 5.2 and Figs. 5.11 – 5.12.

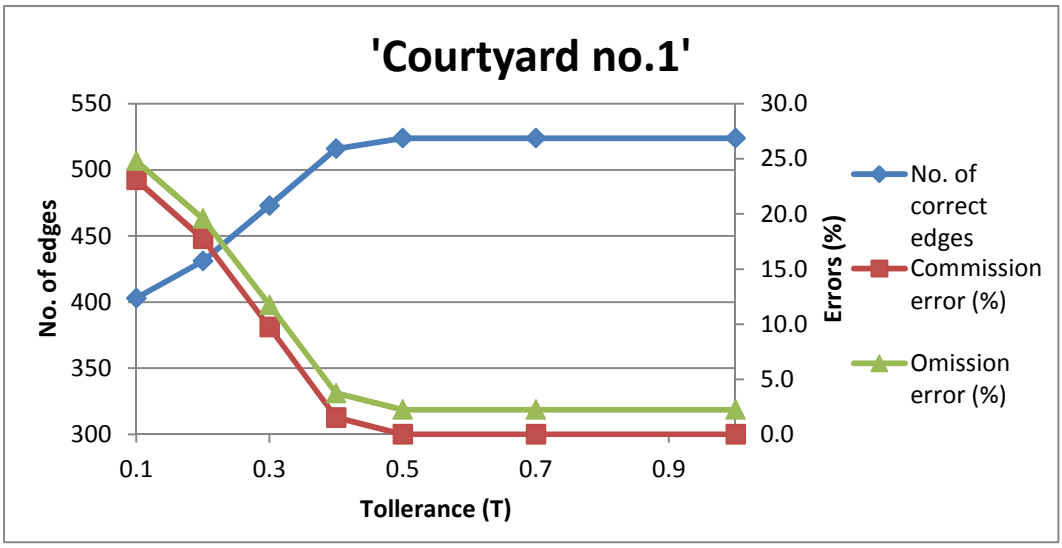


Fig. 5.11. Geometric reconstruction analysis for 'Courtyard no.1' dataset.

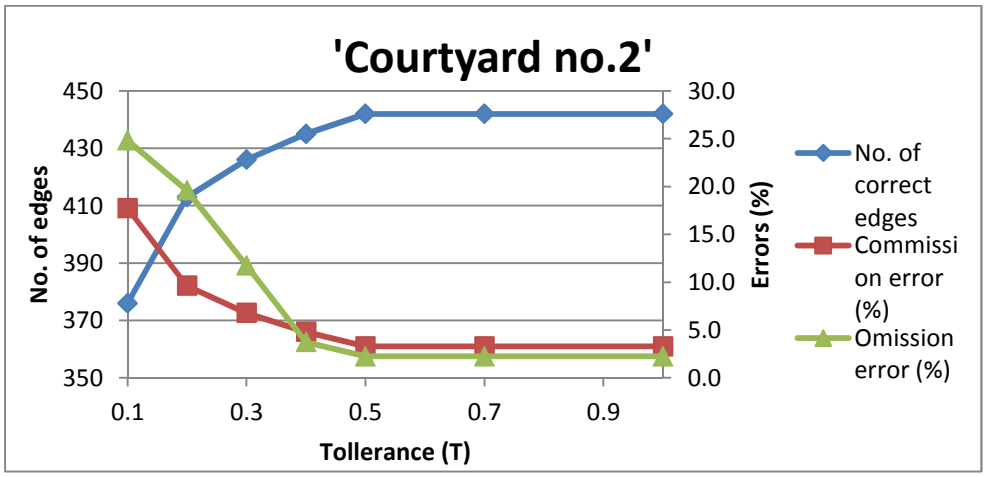


Fig. 5.12. Geometric reconstruction analysis for 'Courtyard no.2' dataset.

The obtained building models in CAD format is presented in Fig. 5.13.

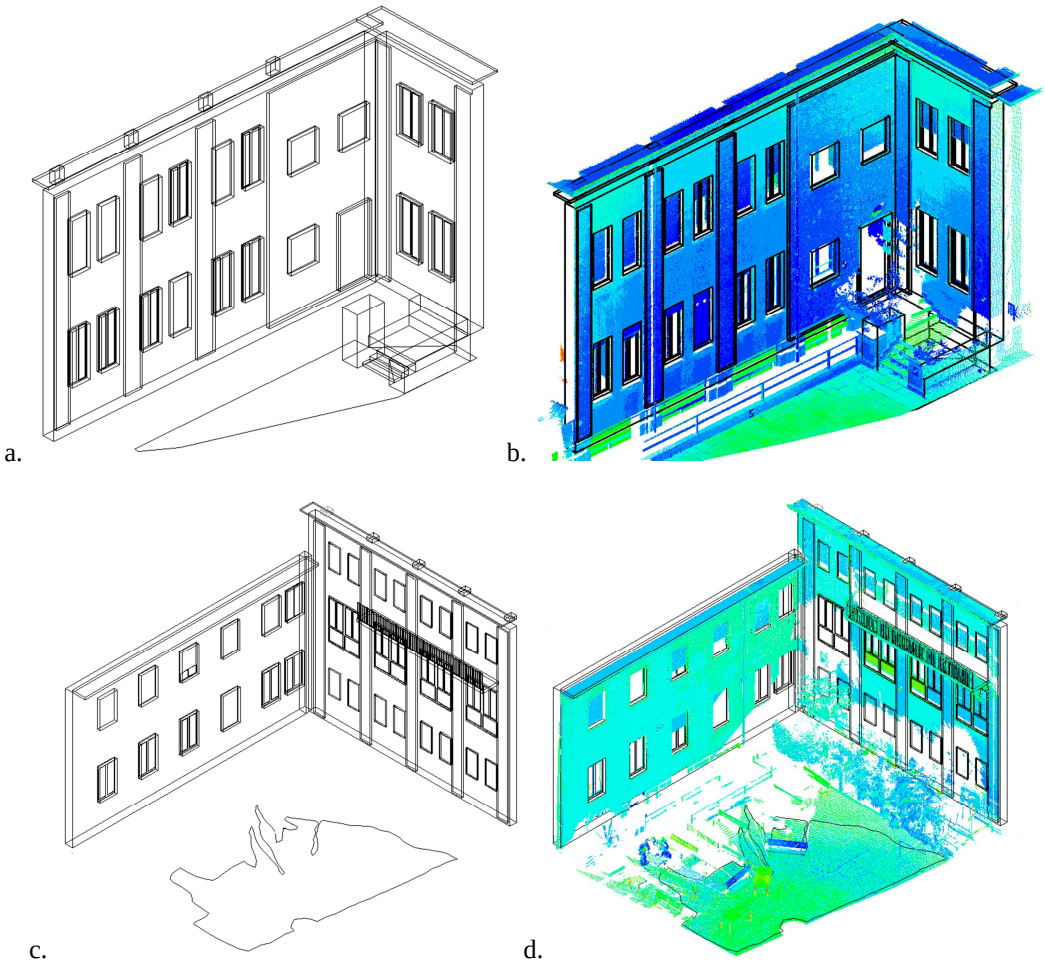


Fig. 5.13. Geometric reconstruction of 'Courtyard no.1and no.2' (top and bottom row respectively): (a-c) final 3D digital models of the façades; and (b-d) façade models with overlaid point cloud.

Also for these two datasets a good matching between extracted and manually identified edges is observed up to a 3.0 mm tolerance. In addition also at lower tolerances a high number of corresponding segments has been found.

In Fig. 5.14 the detected repeated for both datasets are shown. In particular, it is possible to observe that the developed algorithm can efficiently detect missing objects (like windows) also in the case of severe occlusions. Detected regularities can be used to complete the building model.

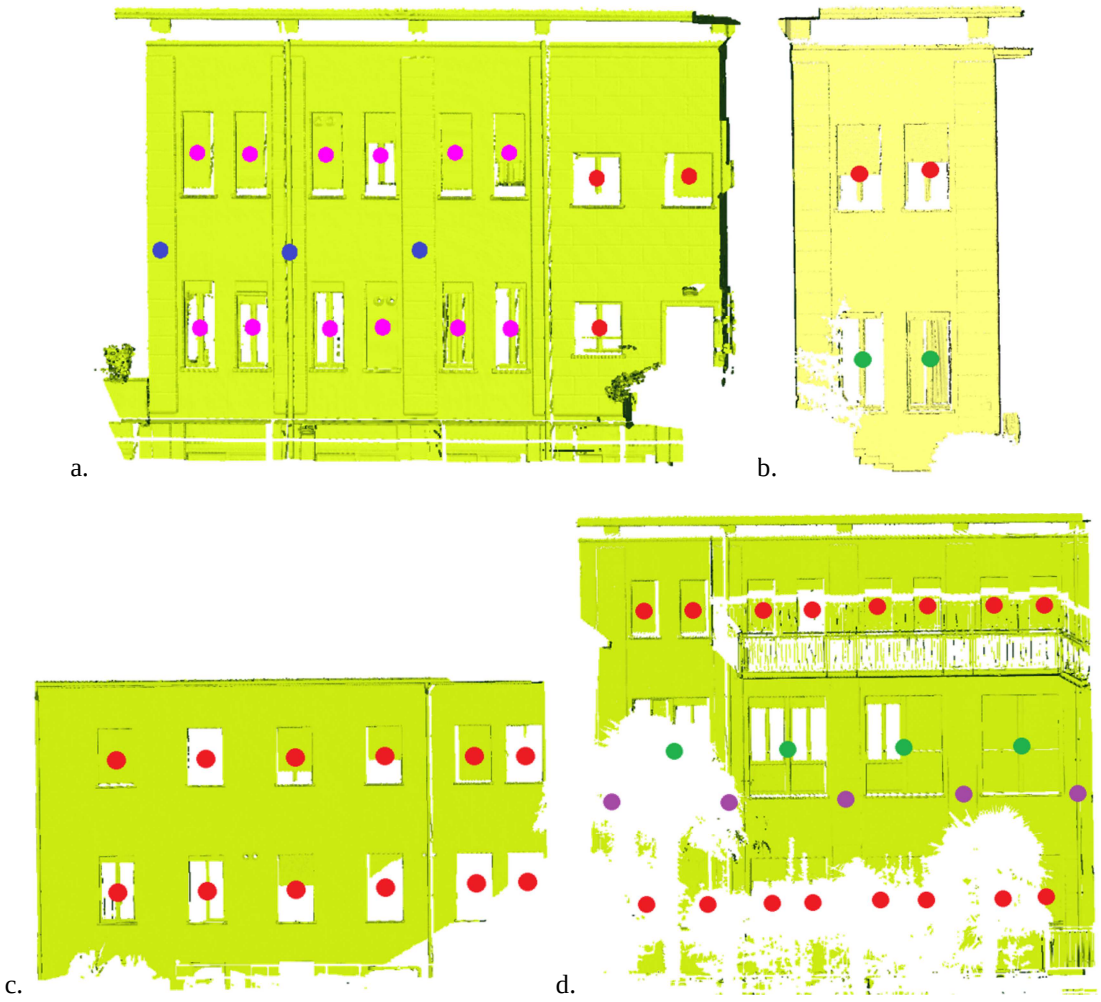


Fig. 5.14. Results of RSD algorithm for 'Courtyard no.1' (a-b) and 'Courtyard no.2' (c-d). Each repeated pattern is shown with a different colour.

Classification results are presented in Tabs. 5.13 – 5.14, where confusion matrices are reported. The overall classification accuracy is 99.1% and 98.1% for dataset 1 and 2, respectively. Also in these datasets the highest confusion exists between roof and walls because vertical roof elements have been erroneously classified as walls.

Classification results lead to the generation of the semantic rich models shown in Fig. 5.15 and Fig. 5.16.

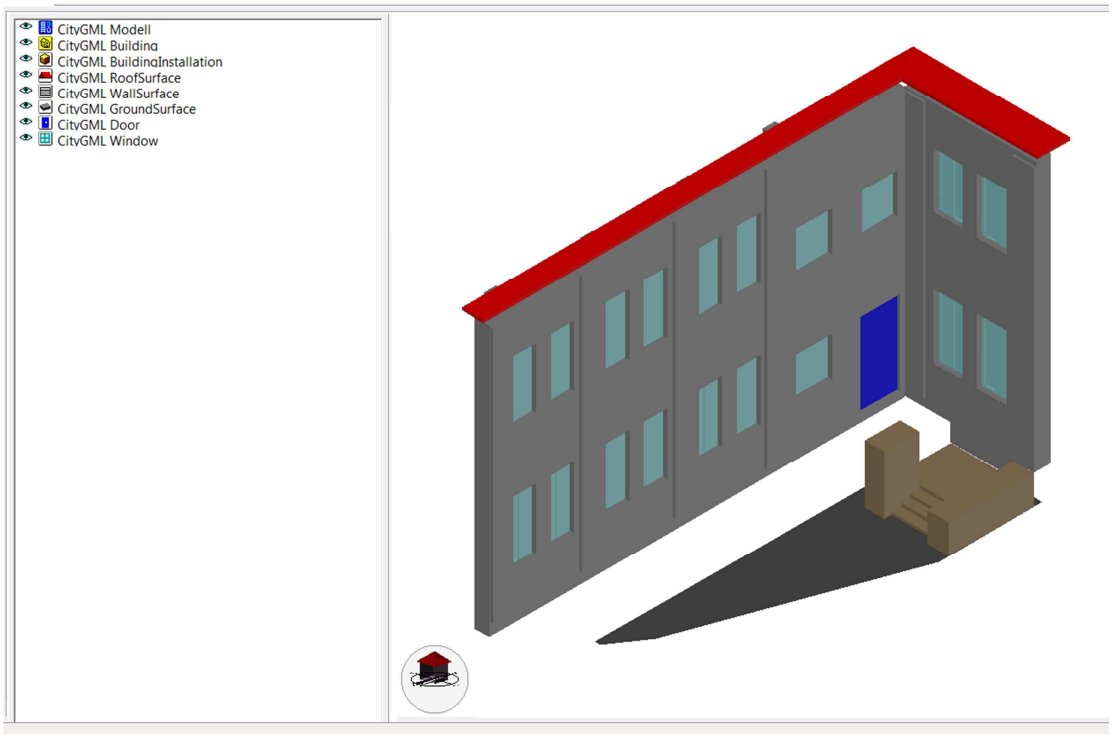


Fig. 5.15. Semantic reach model of the 'Courtyard no.1' in CityGML format. The model is visualized using the FZKViewer developed by the Karlsruhe Institut fur Technologie.

'Courtyard no.1'		Reference								
		No. of wall	**No. of roof**	**No. of sidewall**	**No. of window**	**No. of door**	**No. of ground**	**No. of attachment**	**Total**	**Commission error [%]**
Classified	**No. of wall**	15	0	0	0	0	0	0	15	0.0
	No. of roof	1	6	0	0	0	0	0	7	16.7
	No. of sidewall	0	0	52	0	0	0	0	52	0.0
	No. of window	0	0	0	19	0	0	0	19	0.0
	No. of door	0	0	0	0	1	0	0	1	0.0
	No. of ground	0	0	0	0	0	2	0	2	0.0
	No. of attachment	0	0	0	0	0	0	15	15	0.0
	Total	16	6	52	19	1	2	15	111	
	Omission error [%]	6.3	0.0	0.0	0.0	0.0	0.0	0.0		

Tab. 5.13. Confusion matrix for 'Courtyard no.1' dataset.

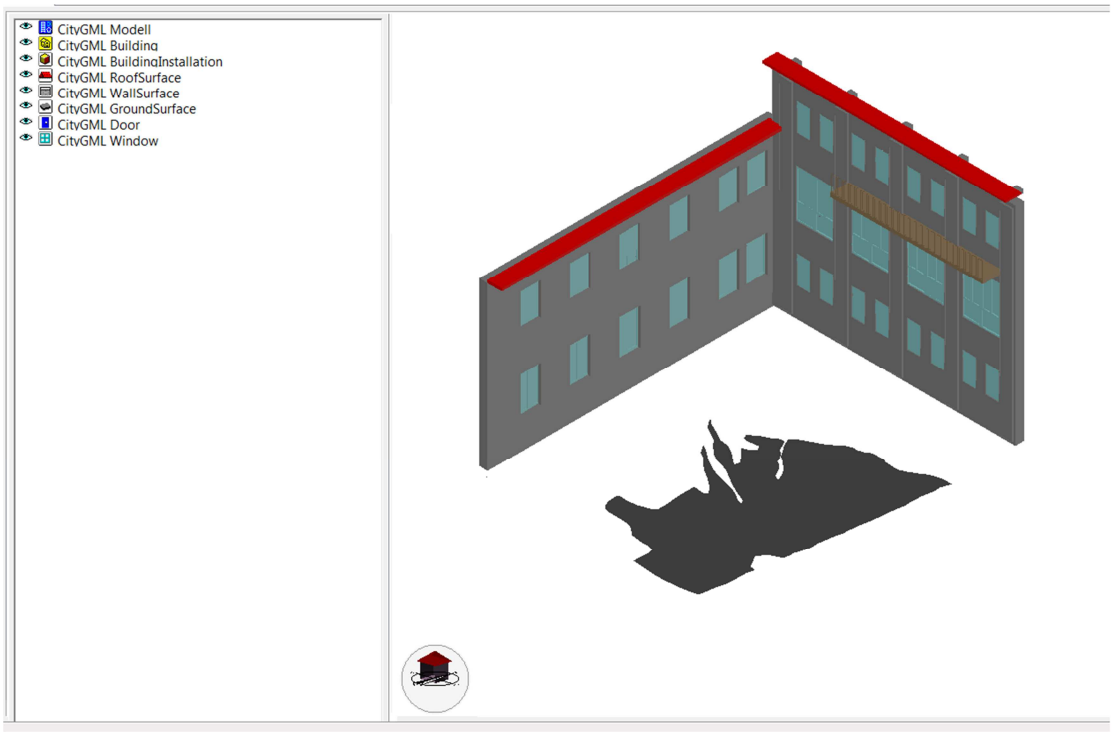


Fig. 5.16. Semantic reach model of the 'Courtyard no.2' n CityGML format. The model is visualized using the FZKViewer developed by the Karlsruhe Institut fur Technologie.

'Courtyard no.2'		Reference								
		No. of wall	**No. of roof**	**No. of sidewall**	**No. of window**	**No. of door**	**No. of ground**	**No. of attachment**	**Total**	**Commission error [%]**
Classified	**No. of wall**	18	0	0	0	0	0	0	18	0.0
	No. of roof	2	8	0	0	0	0	0	10	25.0
	No. of sidewall	0	0	45	0	0	0	0	45	0.0
	No. of window	0	0	0	25	0	0	0	25	0.0
	No. of door	0	0	0	0	0	0	0	0	0.0
	No. of ground	0	0	0	0	0	4	0	4	0.0
	No. of attachment	0	0	0	0	0	0	5	5	0.0
	Total	20	8	45	25	0	4	5	107	107
	Omission error [%]	10.0	0.0	0.0	0.0	0.0	0.0	0.0		

Tab. 5.14. Confusion matrix for 'Courtyard no.2' dataset.

5.2.3. 'Building no. 34'

A further test was carried out on a façade of the seven storey building named 'Building no. 34' located in the Leonardo Campus of Politecnico di Milano. The chosen test façade presents existing precast panels at the two top floors, while the

remaining part features a mortar finishing. An important issue related to the survey is the presences of a large tree just in front of the façade, resulting in large occlusions.

Scans were registered with 8 checkerboard targets measured with a theodolite Leica TS30. The same points were also used for the registration of a block of RGB images, along with some additional natural points (e.g., window and door corners) that have been used for registration of a set of Thermal Infrared (TIR) images as presented in the following Section 6.1. The network scheme is presented in Fig. 5.17. After the adjustment of the geodetic network, the estimated accuracy in checkerboard target measurements resulted in ± 2.0 mm.

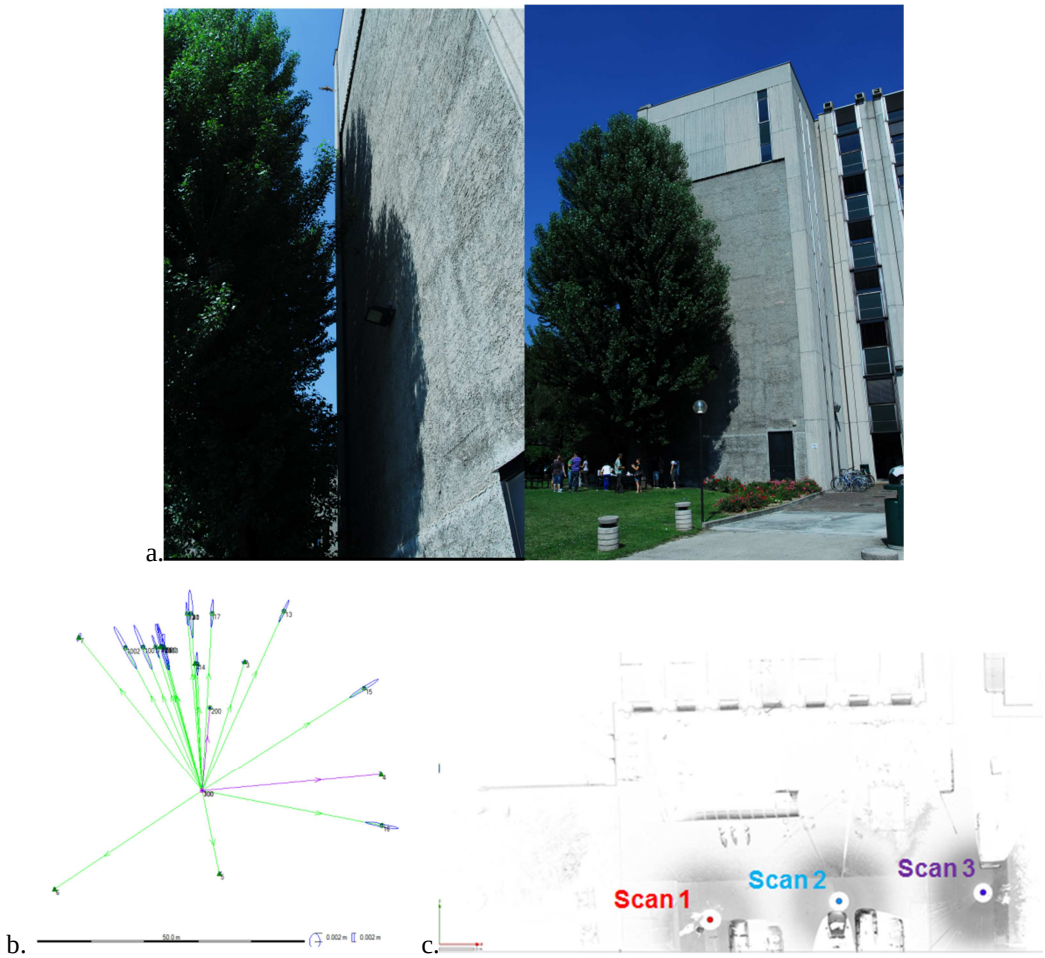
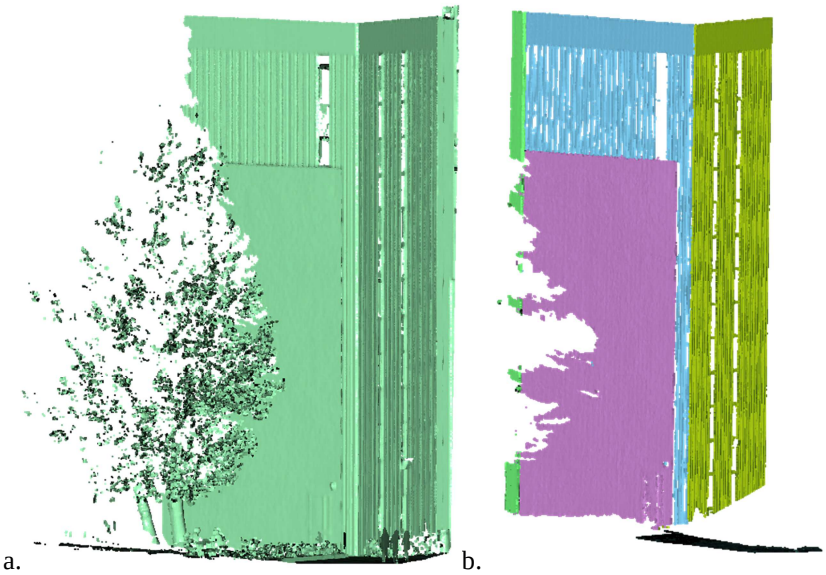


Fig. 5.17. 'Building no. 34' dataset: (a) some pictures of the analysed building; and the schemes of the geodetic network (b) and the scan acquisitions (c).

The laser scanning survey was carried out by using a TLS FARO-FOCUS 3D and consisted in 3 scans (Fig. 17c) acquired from different standpoints in order to survey the entire western façades of the building. The ground sampling distance (GSD) ranges from 1.5 mm in the lower part of the façade up to 4 mm in the upper part. As mentioned, scan referencing was performed by using as GCPs the checkerboard targets complemented by 5 spherical targets used to strengthen the precision of scan

referencing. The mean precision after georeferencing, evaluated from the residuals on GCP, was about ± 3 mm.

Façade segmentation results are summarized in Fig. 5.18 and Tabs. 5.2 – 5.15 – 5.16.



a. b.

Fig. 5.18. 'Building no. 34' segmentation results:.(a) original point cloud; and (b) segmentation results, each detected segment is represented with a different colour.

Compared to the previously cases this is a simpler situation where planes are quite wide and separation between them is clear. For this reason fewer errors have been found.

Tolerance (%)	'Building no. 34' results for reference planes					
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of missed segment	Commission error (%)	Omission error (%)
0.8	21	0	0	1	0.0	4.5
0.85	21	0	0	1	0.0	4.5
0.9	21	0	0	1	0.0	4.5
0.95	21	0	0	1	0.0	4.5

Tab. 5.15. 'Building no. 34' results for reference planes.

Tolerance (%)	'Building no. 34' results for detected planes			
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of noise segment
0.8	21	0	0	0
0.85	21	0	0	0
0.9	21	0	0	0
0.95	21	0	0	0

Tab. 5.16. 'Building no. 34' results for detected planes.

As previously observed the large tree in front of the façade resulted in a large occlusion. Unfortunately, in this case also the developed completion strategy could not be exploited because no repetition could be observed in this area. For this reason, a manual editing was needed. Results of geometric modelling and classification steps are presented in Tabs. 5.3 – 5.17 and Figs. 5.19 – 5.20.

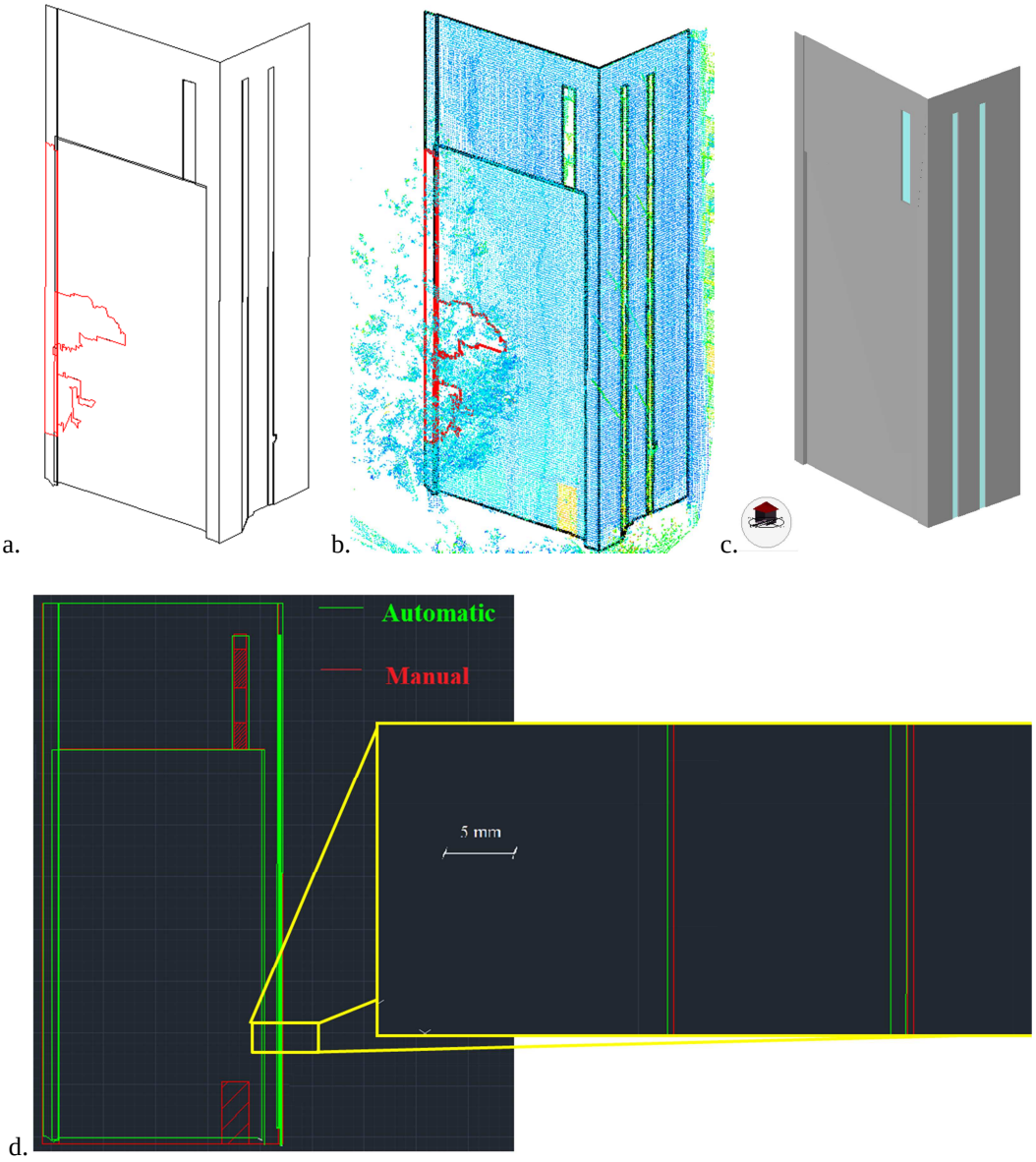


Fig. 5.19. 'Building no. 34' processing results: (a) final 3D digital models of the façade (automatic model, in black, and manual editing, in red); (b) façade models with overlaid point cloud; (c) CityGML model; and (d) visual comparison between automatic and manual modelling results.

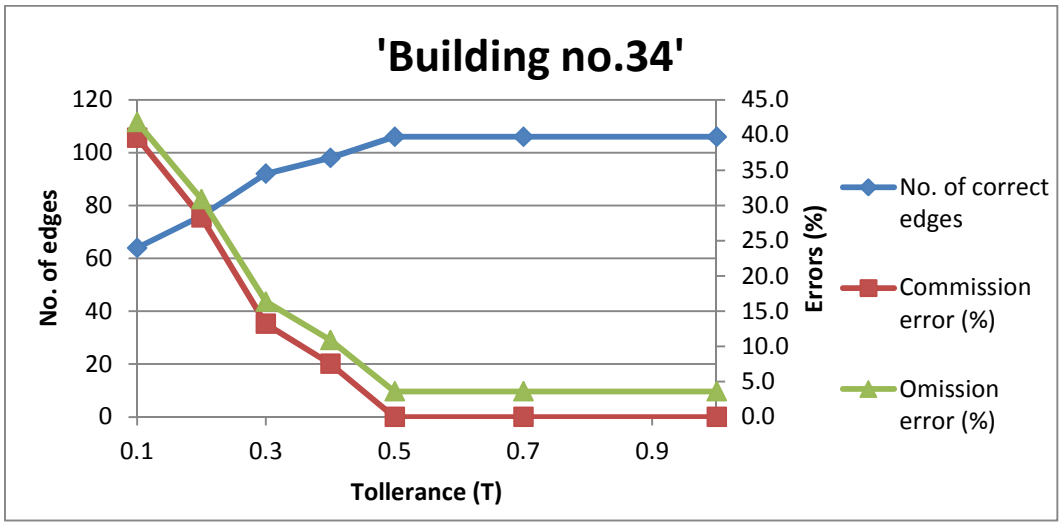


Fig. 5.20. Geometric reconstruction analysis for 'Building no. 34' dataset.

'Building no. 34'		Reference								
		No. of wall	**No. of roof**	**No. of sidewall**	**No. of window**	**No. of door**	**No. of ground**	**No. of attachment**	**Total**	**Commission error [%]**
Classified	**No. of wall**	9	0	0	0	1	0	0	10	10.0
	No. of roof	0	0	0	0	0	0	0	0	0
	No. of sidewall	0	0	8	0	0	0	0	8	0.0
	No. of window	0	0	0	3	0	0	0	3	0.0
	No. of door	0	0	0	0	0	0	0	0	0
	No. of ground	0	0	0	0	0	1	0	1	0.0
	No. of attachment	0	0	0	0	0	0	0	0	0
	Total	9	0	8	3	1	1	0	21	
	Omission error [%]	0.0	0	0.0	0.0	100.0	0.0	0		

Tab. 5.17. Confusion matrix for 'Building no. 34' dataset.

In this case, 100% omission error is reported in the case of door element. This is motivated by the fact that the only door in the scene is coplanar with the façade wall while classification rules assume that such element should be a façade intrusion.

5.2.4. 'Nave building'

The last dataset presented here is the so called 'Nave' office building which was designed by Giò Ponti inside the Leonardo Campus of Politecnico di Milano. The southern façade of the building was scanned by using a long –range scanner RIEGL – LMS 420i. Some technical specifications of the laser scanner adopted are reported in Tab. 5.1.

Fig. 5.21. 'Nave building' dataset: a pictures of the analysed building, due to the presence of a large tree in front of the façade a large occluded area has resulted in the final point cloud .

Four scans have been gathered to cover in a uniform way the entire façade. Registration has been carried out in automatic way by using 24 retro-reflective targets that have been set up over the area. The reference system of the first scan has been kept fixed. Registration adjustment has provided a final sigma naught of 3.5 mm. In this case the data acquisition has taken roughly 5 hours. The average point density of the dataset is about 5 thousand points per square meter on the walls. Due to the presence of a large tree in front of the façade, a large occluded area has resulted in the final point cloud (see Fig. 5.21 and Fig. 5.22).

Due to the large level of noise in the dataset, connected to the nature of the adopted instrument, the RANSAC tolerances for façade segmentation have been changed with respect to the previous tests. In the same way, cell size of the bitmap was increased up to 2 cm to take into account the lower data density. In Tab. 5.2 a summary of the adopted parameters for the modelling process is presented.

A visual representation of segmentation results is given in Fig. 5.22. A more detailed analysis of the data is reported in Tab. 5.3 and Tabs. 5.18 – 5.19.

Tolerance (%)	'Nave building' results for reference planes					
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of missed segment	Commission error (%)	Omission error (%)
0.8	604	0	0	1	0.2	7.1
0.85	599	5	0	1	1.0	7.8
0.9	593	11	0	1	2.0	8.8
0.95	587	17	0	1	3.0	9.7

Tab. 5.18. 'Nave building' results for reference planes.

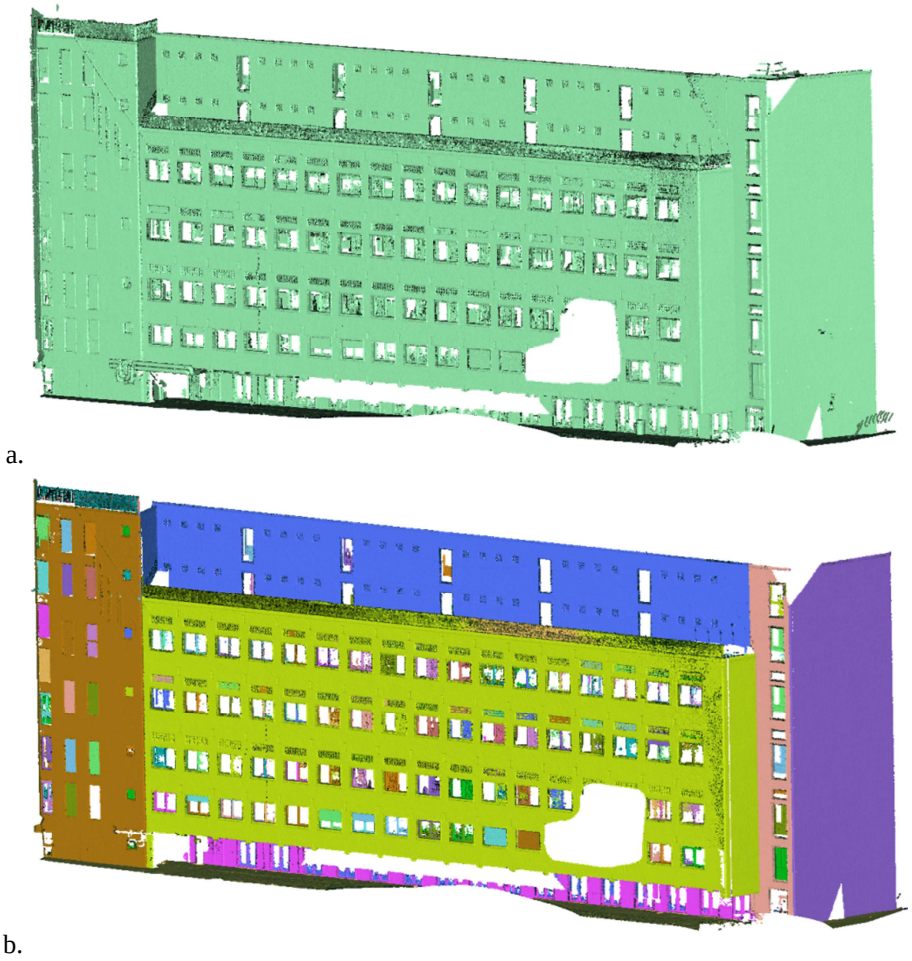


Fig. 5.22. 'Nave building' segmentation results: (a) original point cloud; and (b) segmentation results; each detected segment is represented with a different colour.

Tolerance (%)	'Nave building' results for detected planes			
	No. of correct segmentation	No. of over-segmentation	No. of under-segmentation	No. of noise segment
0.8	604	0	0	46
0.85	599	12	0	39
0.9	593	32	0	25
0.95	587	47	0	16

Tab. 5.19. 'Nave building' results for detected planes.

Also in this case, in the tolerance range 0.8 ÷ 0.9 the number of correctly detected segments is almost constant. Unlike, the previous cases a quite significant amount of over-segmentation and missed instances can be noticed. These results may be due to the higher noise present in the dataset which has influenced the estimation of the local normal. In particular, over-segmentation results are connected to some small segments at the edge of larger regions or inside some regions where the noise and the lower data density made quite difficult the proper estimation of the local normal vector.

The obtained building model in CAD format is presented in Fig. 5.23, while comparison results between manual and automatic models are presented in Tab. 5.3 and Fig. 5.24. Also for the geometric evaluation the tolerances need to be relaxed to take into account for the lower point density.

The number of correctly detected edges at different tolerances has followed the same distribution as in previous experiments. Obviously, due to the higher noise and the lower density, a good accordance between the manual and automatic results can be observed only up to ±1.0 cm tolerance.

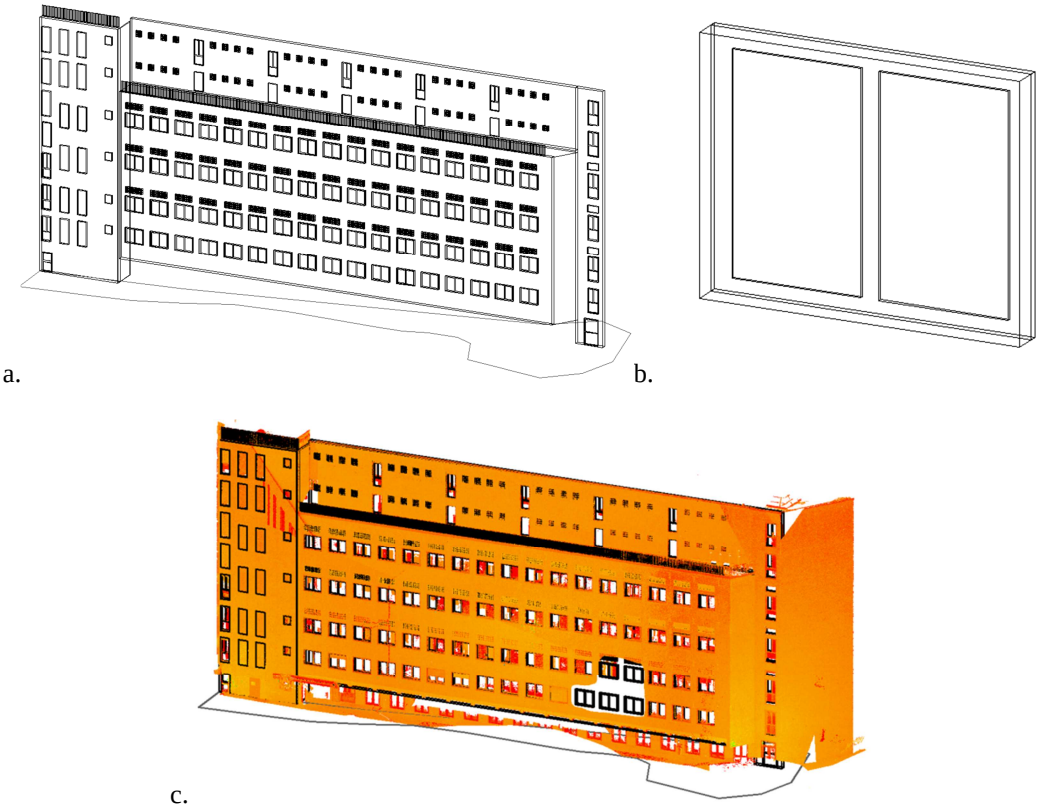


Fig. 5.23. Geometric reconstruction of 'Nave building' dataset: (a) final 3D digital model of the façade with a detail view of a window (b); and (c) façade model with overlaid point cloud.

As previously anticipated, a large missing area due to an occlusion is present in the dataset. However, the developed procedure for RPD has given the possibility of recovering the missing façade objects (Fig. 5.25a). Although this façade shows a quite complex repetition scheme, the developed procedure has succeeded in detecting all the repeated patterns.

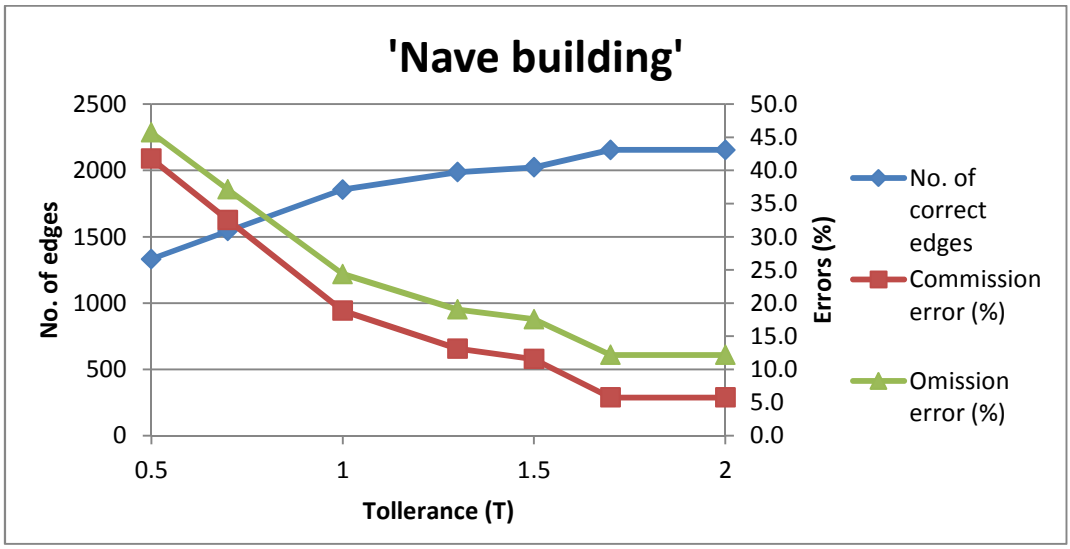


Fig. 5.24. Geometric reconstruction analysis for 'Nave building' dataset.

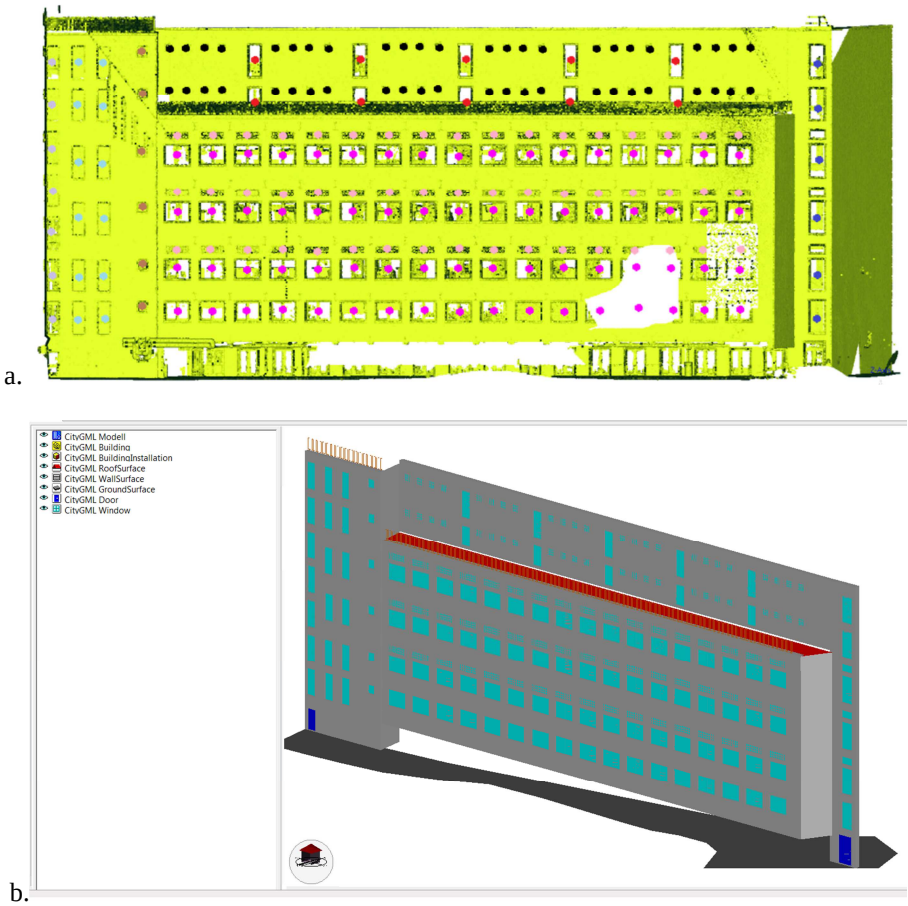


Fig. 5.25. Results of RSD algorithm for 'Nave building', each repeated pattern is shown with a different colour (a), and (b) semantic reach model of the 'Nave building' in CityGML format.

The confusion matrix for this dataset is presented in Tab. 5.20. The overall classification accuracy is 99.5%. A misclassification of a door element is observable. This is given by the fact the door is in the same geometric plane of the main façade wall, while classification rules assume a door element as an intrusion. The final CityGML model is presented in Fig. 5.25b.

'Nave building'		Reference								
		No. of wall	**No. of roof**	**No. of sidewall**	**No. of window**	**No. of door**	**No. of ground**	**No. of attachment**	**Total**	**Commission error [%]**
Classified	**No. of wall**	7	0	0	0	0	0	2	9	22.2
	No. of roof	0	4	0	0	0	0	0	4	0.0
	No. of sidewall	0	0	388	0	0	0	0	388	0.0
	No. of window	0	0	0	188	0	0	0	188	0.0
	No. of door	1	0	0	0	1	0	0	2	50.0
	No. of ground	0	0	0	0	0	2	0	2	0.0
	No. of attachment	0	0	0	0	0	0	12	12	0.0
	Total	8	4	388	188	1	2	14	605	
	Omission error [%]	12.5	0.0	0.0	0.0	0.0	0.0	14.3		

Tab. 5.20. Confusion matrix for 'Nave building' dataset.

5.2.5. Conclusions

In this section different tests on some reals façades are presented to evaluate performance of the developed modelling procedure.

The developed segmentation strategy proved to be quite robust. Indeed, statistics revealed a low number of commission and omission errors proving the method is able to detect in a correct way planes also in the case of restrictive tolerances. The robustness of the method is also proved by the fewer instances of missing and noisy segments, proving that there is a strict matching of points between reference and the obtained segmentation results. Some instances of over segmentation are mainly due to non-conformance of some features segments to the planar assumption. Also noise in the local normal evaluation may generate over segmentation problems. On the other hand, in Chapter 1 it has been already observed that over-segmentation may be recovered at a late stage by grouping together several segments. The developed segmentation strategy was also successful in detecting both larger surfaces (e.g., walls, roofs) as well as smaller features (e.g., parts of windows frames).

The geometric accuracy of the reconstructed models is also evaluated. Obviously, different results have been found according to the point density of each dataset. However, the obtained results can be considered as similar to the ones obtainable with

a traditional manual modelling procedure. Some problems still exist in the case of high fraction of missing data. For those cases, a first automatic modelling followed by a manual editing to fix the problems seems the most adequate procedure to save time and minimize the human effort.

The algorithm for façade regularity detection proved to be able to detect complex and interlaced grids of repeated elements. Detection and exploitation of repeated patterns was effective for completion of missing parts in the case of high repetitive structures.

The experiment proves that the developed object classification strategy is effective. Indeed, overall classification accuracy, precision and recall are always satisfactory. Some problems still exists for correct classification of some features. In particular, some confusion situation may be found among doors, walls, and roofs. This is given by the fact that some of the defined rules does not cover a complete set of clauses to ensure correct classification. A higher specialization of these rules can be used to partially overcome this limitation.

Chapter 6
Other applications

This chapter presents some extensions of the developed automatic façade modelling approach described in the first part of this dissertation. Indeed, the obtained building models can be integrated with other data sources to obtain a more complete representation.

For thermal retrofitting of existing buildings, evaluation of the thermal efficiency is of primary importance to properly design the intervention. Infrared Thermography (IRT) has been proved to be a valuable diagnostic tool for characterization of buildings' thermal behaviour and detecting thermal bridges and heat losses from the envelop (Maldague 2001). However, the use of IRT data for metric purposes presents several limitations due to the reduced resolution of thermal sensors, the large distortion introduced by the thermographic lens systems, and the impossibility of making precise geometric measurements directly on the images. To partially overcame these limitations, in Section 6.1 a procedure based on the combination of thermal information derived from IRT and geometric information of building structure is presented.

On the other hand some of the algorithms developed for façade modelling can be extended, and partially modified, for other applications.

In particular, the developed segmentation strategy can be used for scan registration. Point cloud acquisition by using laser scanners provides an efficient way for 3D as-built modelling of urban environments. In the case of large structures, several scan-points are needed to cover the entire scene and this result in a registration problem (Vosselman and Maas 2010). Several solutions based on artificial targets are nowadays available in commercial software packages. However, artificial targets may not be placed in all situations and a different registration strategy is required in those cases. Identification of the same geometric features among a series of scans can be used to work out the rigid-body transformation useful for the registration of each scan into the global reference system of the final point cloud. This may be an interesting solution for urban and architectural scenes which presents the prevalence of some few

basic geometric shapes. For this reason, a method based on the extraction of planar features for scan registration is presented in Section 6.2.

Laser scanners are often used not only for generation of detailed models of building façades but also for indoor environments (Budroni and Boehm 2005). Automatic reconstruction of buildings' exteriors share many properties (and problems) with the issues associated with indoor modelling. Typically, many of the algorithms useful for the reconstruction of exterior building models can be adapted for indoor reconstruction, as well. However, exterior façade modelling methods operate under the assumption that the surface being modelled is relatively free from obstructions. Even if this may be a reasonable assumption for outdoor scanning, in indoor environments objects like furniture and wall-hangings frequently may occlude the wall surfaces, making the modelling problem more challenging. For this reason, the approach developed for façade modelling was extended to indoor environments (Sect. 6.3) to work with significant amounts of clutter and occlusion.

6.1. Façade model integration with IRT images

In the domain of conservation and maintenance of existing buildings IRT has proved to be an adequate and efficient technique (Maldague 2001, Martín-Ocaña et al. 2004, Ribarić et al. 2009). However, both surface temperature and geometry are needed for a reliable evaluation of thermal efficiency, where spatial relationships are important to localize thermal defects and quantify affected surfaces. For this reason a procedure is developed to combine the geometric content of automatically derived building models and the temperature information derived from IRT into a single framework. These result in a thermography-textured 3D digital model of a building. This model can be interactively browsed, opening in this way new possibilities for the investigators. In addition, starting from the textured models, also raster products can be obtained like thermographic-mosaics, orthophotos, and rectified images.

The key factor for a fruitful integration is the co-registration of the thermal images and the geometric 3D model of the building. According to the structure of the surface and to the image acquisition procedure, the problem can be coped with in different ways. A simple homographic transformation can be correctly used only when the 3D model of the building façade is flat (González-Jorge et al. 2012). Homography estimation requires the identification of at least four corresponding control points (CPs) on both image and object surface. If the façade surface has a more complex 3D shape, homographic model does not hold any more (exception is made for almost flat objects with only small off-plane parts). A more comprehensive approach for image registration is based on collinearity equations which are normally used in photogrammetry to describe the perspective transformation process behind image formation (Luhmann et al. 2006). In this case there are two opportunities both based

on 3D perspective transformations: (i) single image resection or (ii) bundle adjustment of a block of images.

The former technique (Abdel Aziz and Karara 1971) is the most popular in most commercial software packages. In this case the registration of each image can be directly performed by using collinearity equations and by knowing coordinates of at least 3 GCPs (e.g., by using a theodolite or directly from the point cloud derived from TLS). Coordinates of GCPs have to be measured on the images as well to obtain an estimate of camera parameters (position and attitude). However, images are processed independently increasing the number of points to be measured and determining some problems in overlapping areas between consecutive images. Furthermore, the achievable accuracy of the orientation with texture-less images can be questionable.

In photogrammetry, a bundle adjustment approach is used to partially overcome these problems (Luhmann et al. 2006). Several images are registered in a common reference system through the solution of a linearized system of collinearity equations. The unknowns of the system are the six exterior orientation (EO) parameters of the images, while the intrinsic calibration parameters are usually considered as fixed after their estimate with a preliminary calibration project. Additional GCPs are used to control the solution and setup the reference system. Bundle adjustment has also the advantage of exploiting common points between images, reducing so that the total number of points to be measured. However, as thermographic cameras have intrinsic parameters similar to a telephoto lens (narrow field-of-view and long focal lens) it is rather difficult to obtain a block of thermal images suitable for a stable adjustment. Indeed, because of the limited field-of-view thermal image blocks generally present a low ratio between image baselines and camera-object distance. For this reason a simple bundle adjustment of thermal images could provide unreliable results. Here a different approach is used to avoid instability problems and to increase precision of EO estimation. The developed methodology makes use of thermal and RGB images acquired independently, e.g., even in different days. However, the combined orientation of both datasets and the larger resolution and format of RGB images help compute the EO of thermal images.

6.1.1. IRT image integration overview

As previously anticipated the integration between thermal data and geometric building model is obtained by mapping thermal images on the 3D semantically enriched model of the building derived from the procedure described in Part A. Main steps of processing IRT data are reported in the following subsections, while more details can be found in (Previtali et al. 2013b). As shown in the workflow in Fig. 6.1, the procedure can be divided into two main parts: thermographic image processing and automatic façade model generation.

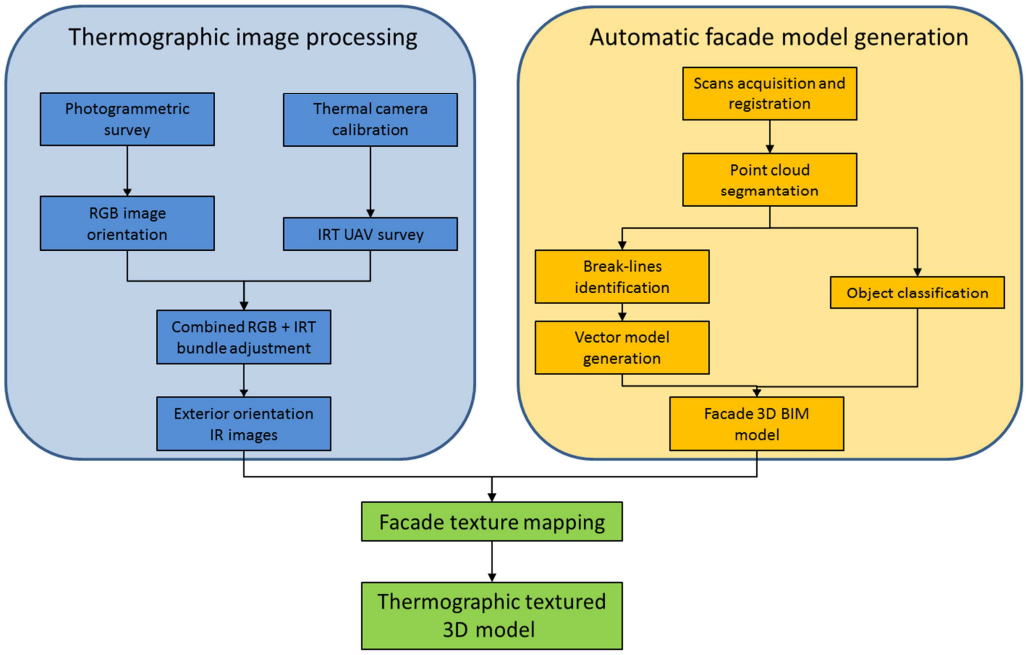


Fig. 6.1. Workflow of the proposed methodology for the generation of a thermographic textured 3D façade model.

Once both parallel processing of image and point cloud processing are completed, their integration is accomplished by texturing the building model with thermal images. The procedure adopted for texture mapping is described in (Previtali et al. 2012).

6.1.2. Thermal camera intrinsic calibration

In order to use a camera for photogrammetric purposes, its calibration should be carried out to determine the interior orientation parameters and to compensate for the effects of lens geometric distortion. In the applications considered in this study, neglecting the correction of distortions may lead to a significant worsening of the final quality of the textured model.

In the case of IRT sensors, the *pinhole camera* model can be assumed and calibration applied by using standard photogrammetric methods. However, IRT cameras are not designed for metric purposes and their calibration may not be an easy task for a series of reasons (see also Luhmann et al. 2013):

- geometric lens distortion could be quite large, especially at the borders of the images;
- because of the shorter wavelength of IR spectrum with respect to the visible one, the diffraction disk diameter is much larger, resulting in a larger pixel size in the thermal sensors;

- in terrestrial applications the IR image resolution is generally in the order of 640x480 pixels or lower (except in some highly expensive sensors);
- auto-focusing systems cannot be usually turned off causing instability in the interior orientation parameter estimation; and
- the limited field of view makes more difficult to carry out a calibration project.

All these aspects should to be taken into account for planning the calibration project, in order to fix a proper set of 'best practice rules' to be generally adopted.

Cameras used in this research for IRT surveys were calibrated by using the Brown's model (Brown 1971), which is based on 8 parameters (principal distance, principal point coordinates, 3 coefficients for radial distortion compensation, and 2 parameters for decentring distortion). These parameters can be estimated by using a proper calibration target set (Fig. 6.2), which must be imaged from different positions. The solution here adopted is based on a set of 40 iron nails fixed in a wooden structure. When exposed to sunlight, nails warm up faster than wooden background and become clearly visible in IR images (Gianinetto et al. 2005).

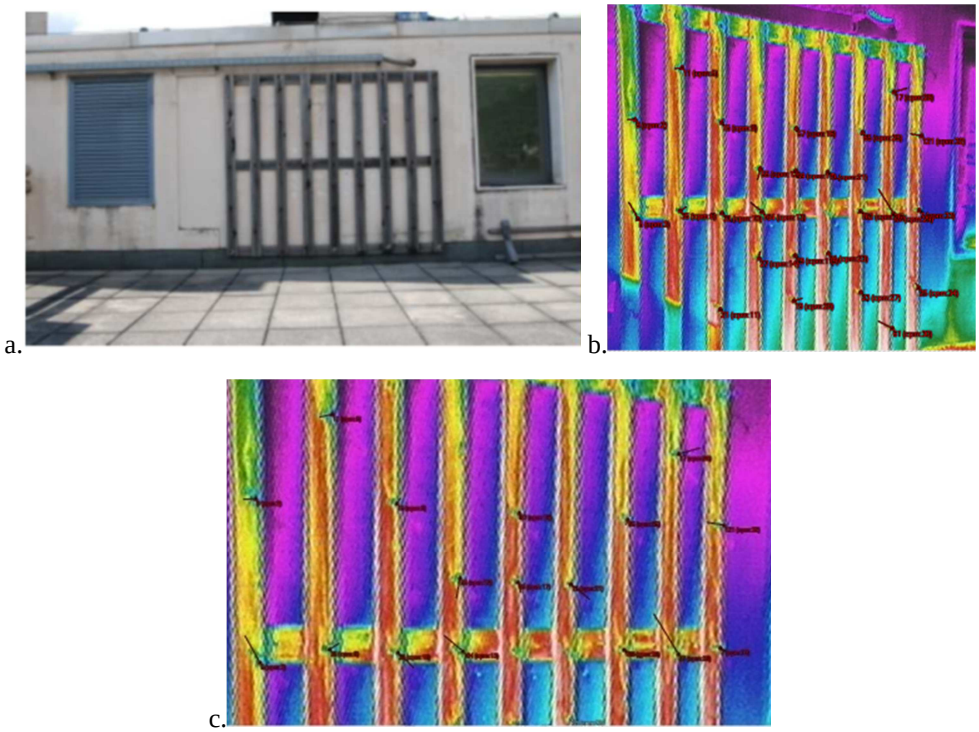


Fig. 6.2. The wooden panel with the iron nails used for IR camera calibration as it is depicted in an RGB image (a) and in IR image (b-c), respectively.

Each of these nails was also measured with a first order theodolite Leica TS30. The 3D coordinates of all nails and the corresponding image coordinates, manually measured on all images, are included in a bundle adjustment, whose solution includes

also the Brown's model calibration parameters. The average precision of this procedure can be evaluated through the estimated statistical quality parameters computed along with the bundle adjustment. In the case under consideration, the computed set of parameters allowed to transform points from the images to the real word with an average residual error of ±2 mm. Calibration results for camera FLIR-Tau 640 adopted in the test described in Paragraph 6.1.4.2 are presented in Tab. 6.1.

Parameter	Calibration result	σ
Focal length c	19.0851 mm	5.24×10^{-3}
Principal point x_p	0.0616 mm	7.10×10^{-3}
Principal point y_p	0.1216 mm	6.24×10^{-3}
K_1	-1.16 x 10-3	2.61×10^{-5}
K_2	6.43 x 10-6	2.20×10^{-5}
K_3	0	0
P_1	0	0
P_2	0	0

Tab. 6.1. Calibration results for thermal camera FLIR-Tau 640.

6.1.3. IRT image orientation

The solution presented here is based on a global photogrammetric bundle adjustment combining both IR and RGB images, which tries to overcome some drawbacks connected to standard space resection and to the bundle adjustment of IR images only (Previtali et al. 2013).

The procedure starts with the acquisition of an adequate set of RGB images with a calibrated camera, meaning that the image block should satisfy the standard requirements of a close-range survey in terms of image overlap, baseline between consecutive images, image resolution (Fraser 1984). For instance, in the case of a planar-like façade, a simple strip of images with an overlap of about 80% can be a good choice which allows one to find Tie Points (TPs) on 3-4 images. In building surveys some factors like occlusions caused by surroundings buildings and/or trees, logistics limitation and the like, might influence the block design. Consequently, a trade-off between concurring requirements has to be defined.

Then, RGB images are oriented within a standard photogrammetric bundle adjustment, which is based on a set of TPs measured on the images, and some GCPs that are used to register the project in the reference system of the building model. An important consideration deserves to be mentioned: TPs individuated in this first step will be used for the registration of IR images. For this reason TPs should be preferably measured in correspondence of elements that are clearly visible in both RGB and IR images (e.g., window and door corners). This strategy may turn out in

increasing the processing time of the bundle adjustment of the RGB images, if compared to a standard photogrammetric project where the integration of IR images is not needed. On the other hand, the larger number of TPs will help the registration of the IR data.

After the registration of all RGB images IR images can be added to the block by measuring some TPs between RGB-to-thermal and thermal-to-thermal points. A final combined bundle adjustment including all images is finally carried out to obtain the EO parameters of all images simultaneously. This allows the creation of a more robust image block made up of both images, where RGB images strengthen the bundle adjustment solution and allow the estimation of reliable EO also for thermal images.

6.1.4. Applications

In this subsection two application of the developed integration procedure are reported.

6.1.4.1. 'Nave building'

A first application example of the developed orientation technique is the 'Nave building' of Politecnico di Milano.

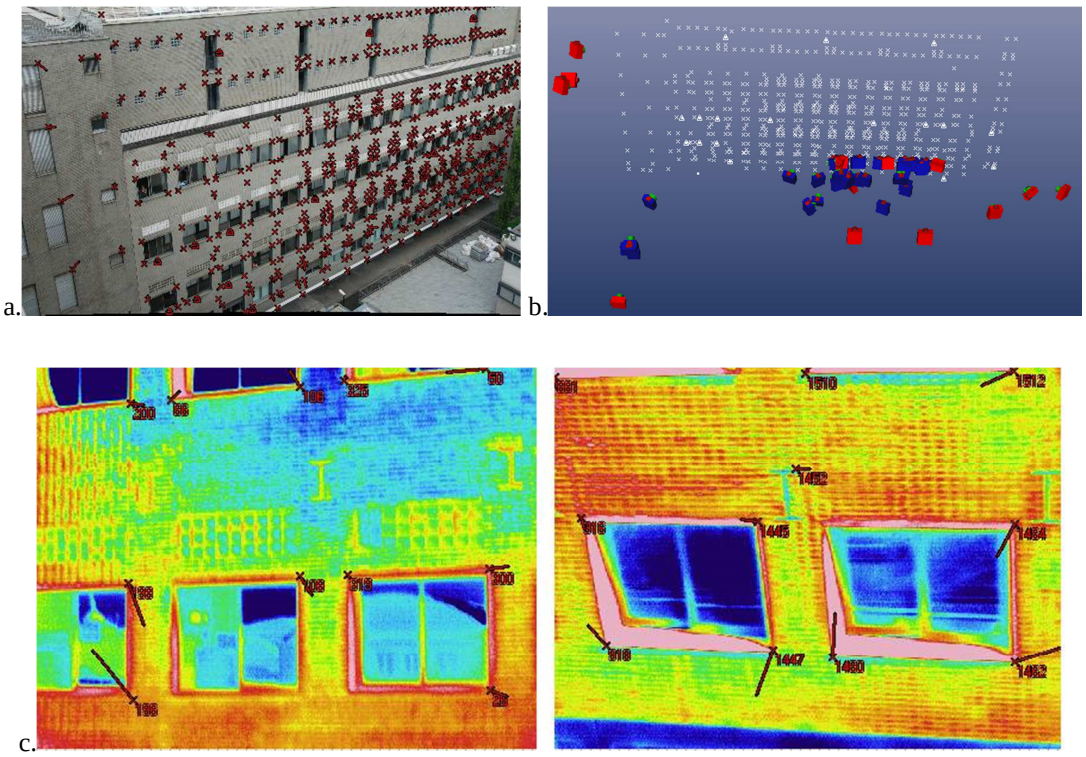
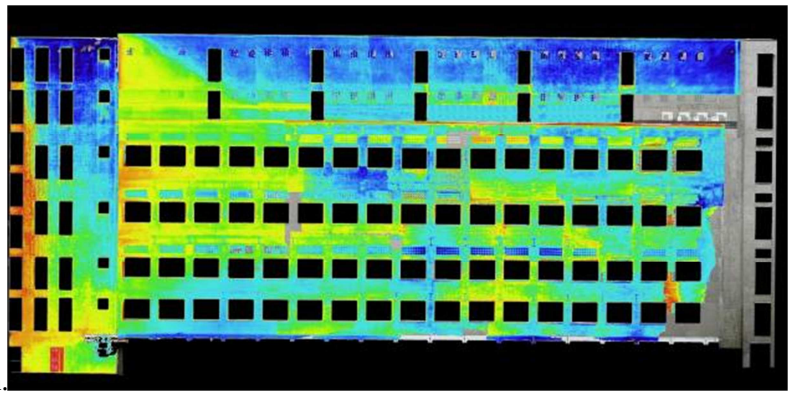


Fig. 6.3. In the upper row, (a) an RGB image with the measured TPs (red crosses); and a 3D view showing camera poses of both RGB (red) and thermal (blue) datasets (b). In the lower row (c), some IR images used for texturing the 3D model with measured TPs; red lines are a graphical representation of the residuals after registration through the bundle adjustment (a magnification factor 50 is applied).

Firstly, 17 RGB images acquired with a Nikon D80 equipped with a 20 mm Sigma lens were registered within a bundle adjustment. In this project 16 retro-reflective targets, measured with a theodolite Leica TS30 and also employed for the registration of laser scans were used as GCPs during bundle adjustment. As can be seen in Fig. 6.3b the irregular distribution of both RGB (red cameras) and IR images (blue cameras) depends on the presence of several occlusions due to obstacles like other buildings and trees. Then, 65 IR images were included in the bundle adjustment by using more than 600 TPs previously measured in the RGB images. Thermal images were acquired with two different IRT cameras (AVIO TVS700 and NEC TH9260) and were included in a unique adjustment with different sets of intrinsic calibration parameters. The high number of TPs used is connected to the narrow field-of-view of both thermal cameras that limits the number of elements clearly measurable in each image (Fig. 6.3c).

a.



b.

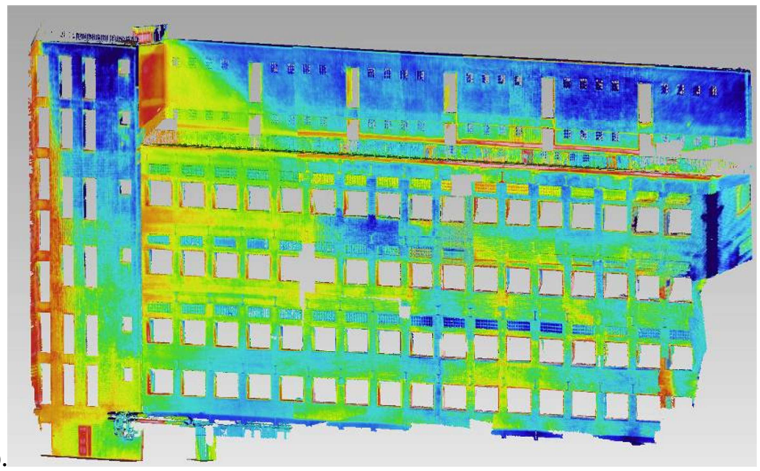


Fig. 6.4. IR orthoimage (a) of the 'Nave building' of Politecnico di Milano and the 3D model textured by IR thermal images (b).

Statistics on the combined bundle adjustment show a final sigma nought of about 1.2 pixels. This result can be considered as fully acceptable due to the low geometric resolution of IR images. In fact, the ground sampling distance (GSD) of thermal images was about 1 cm while the one of RGB images was 1 mm, meaning one order

of magnitude difference. In Previtali et al. 2013 a quantitative comparison of different registration techniques for IRT, RGB and building model are reported and discussed. This analysis involved the use of a set of check point, with respect to which residuals after image orientation have been worked out. In particular, discrepancies for homography and space resection method revealed being much larger than those obtained from the combined bundle adjustment. Final orthophoto results are presented in Fig. 6.4.

6.1.4.2. 'Building no. 34'

A second case study is the 'Building no. 34'. In this case the thermal survey was carried out by using an unmanned aerial vehicle (UAV) platform. Indeed, UAVs may be really useful for thermographic analysis. The use a UAV platform allows exploring areas inaccessible from the ground like roofs. In addition the possibility to reduce the camera-object distance allows one to enhance the ground sampling distance (GSD). This effect is particularly evident in the case of tall buildings. In these cases, images acquired from the ground may present a GSD of several centimetres at top floors, while the use of a UAV gives the chance to obtain a uniform GSD all over the building. The thermographic survey was carried out by using UAV platform (Fig. 6.5a-c) AscTec Falcon 8 (70 cm x 60 cm, weight 2 kg), equipped with 8 motors and able to fly up to 20 minutes with a single battery.

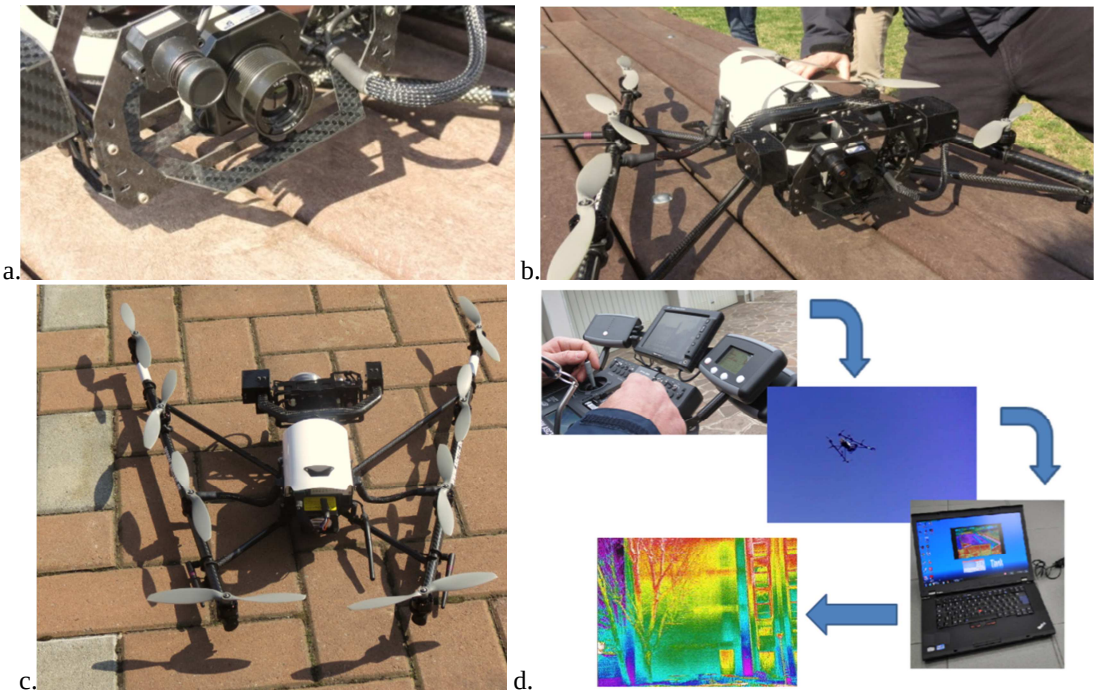


Fig. 6.5. The ASCTEC Falcon 8 equipped with a FLIR Tau 640 thermal camera (a-c); and (d) the thermal image acquisition procedure.

The system presents an actively stabilized camera systems that allows mounting different payloads. In particular, in experiments 5 RGB images were acquired with a Sony NEX-5N camera and 11 thermal images were recorded with FLIR Tau 640 camera equipped with a 19 mm lens system. The electronic equipment of the Falcon 8 includes a GPS antenna and a system of accelerometers determining the system roll, pitch and yaw. The communication system allows the ground station to receive telemetry data and video signals from the on-board sensors. During the survey the UAV can be remotely controlled by a human operator while thermal images were acquired by using a laptop to record the video signal from the thermal camera (Fig. 6.5d).

The survey was completed by acquiring 18 RGB images from the ground. Two thermographic campaigns were performed, the first one in winter (March 16th, 2013), the second in summer (July 9th, 2013). In both cases thermal images were acquired with the UAV following a vertical strip.

As described in Sect. 6.1.3, the orientation of these images was performed in two steps. First, the RGB images were registered within a bundle adjustment. In this project 10 natural points (e.g., window and door corners) measured with a Leica TS30 were used as GCPs during the bundle adjustment. Then, starting from 30 TPs measured in the RGB images, 11 IR images were included in the bundle adjustment. Statistics of the combined bundle adjustment show a final RMS of about 0.9 pixels (Fig. 6.6).

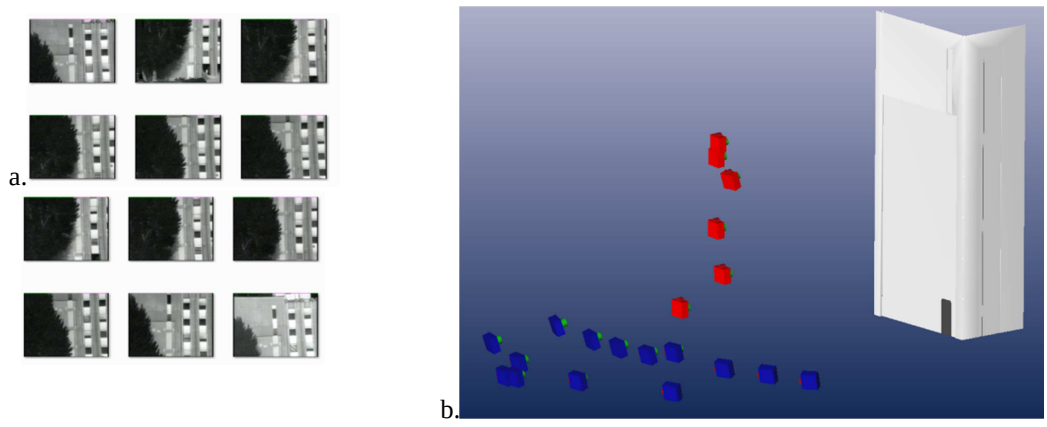


Fig. 6.6. Thermal image processing: (a) the acquired thermal images; and (b) camera poses: red cameras are IR images, whereas blue cameras represent RGB images.

After the registration of IR images in the same reference system of laser scanning point cloud, data were mapped and mosaicked on the digital model of the façade, and then the final thermal orthoimages were derived by simply projecting the data on a plan parallel to the façade. Orthophotos shows the presence of some thermal

anomalies on the façade. In particular, the presence of some closed doors and other objects that are not visible in standard RGB pictures (Fig. 6.7) have been evidenced.

IR March 2013 IR July 2013 RGB March 2013

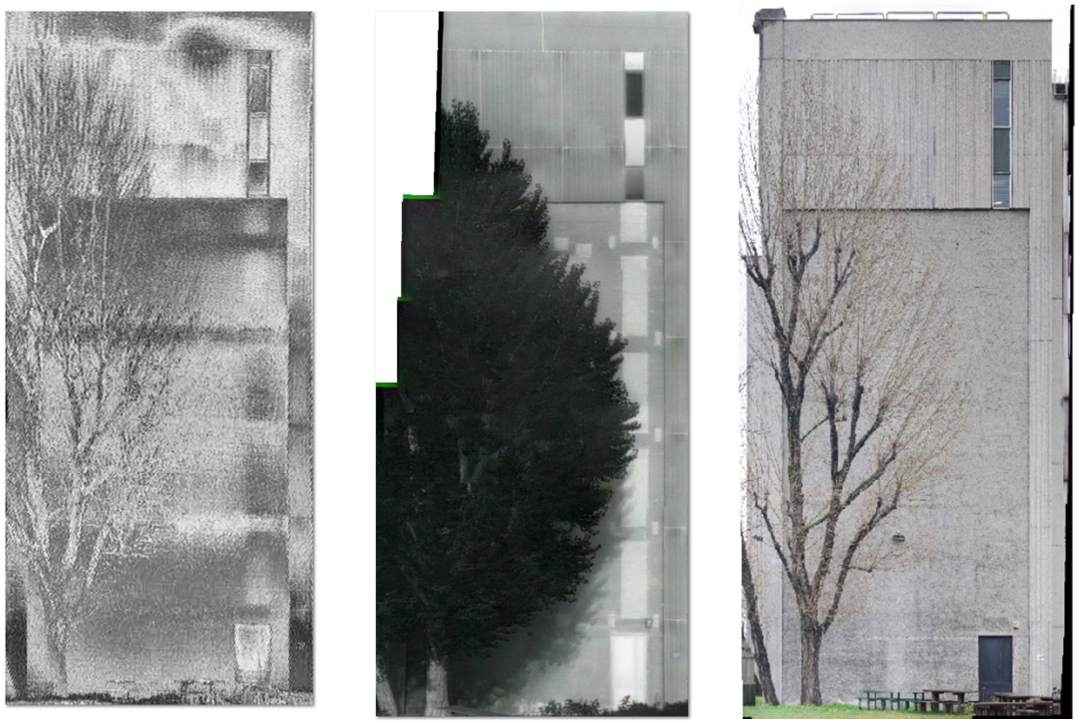


Fig. 6.7. Façade thermal orthophoto: March 2013 (a), July 2013 (b) and RGB (c).

6.2. Scan registration using planar features

Point cloud registration plays an essential role in static terrestrial laser scanning data processing. In this paragraph, a methodology based on planar features is used to combine multiple scans.

Even if different approaches were developed for scan registration (the reader is referred to Barnea and Filin (2010) and Vosselman and Maas (2010) for a more comprehensive overview on this topic) *target-based* approaches are the most commonly employed. Indeed, given a sufficient number of 3D point correspondences (at least 3) parameters of a rigid-body transformation between the instrumental system, termed *Intrinsic Reference System* (IRS), and a *Ground Reference System* (GRS) can be achieved (see Eq. 6.1 and Fig. 6.8):

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = R \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + T \tag{6.1}$$

where (x_1, y_1, z_1) and (x_2, y_2, z_2) are the reference system to register, IRS and GRS respectively, R is a 3×3 rotation matrix and T is a 3×1 translation vector.

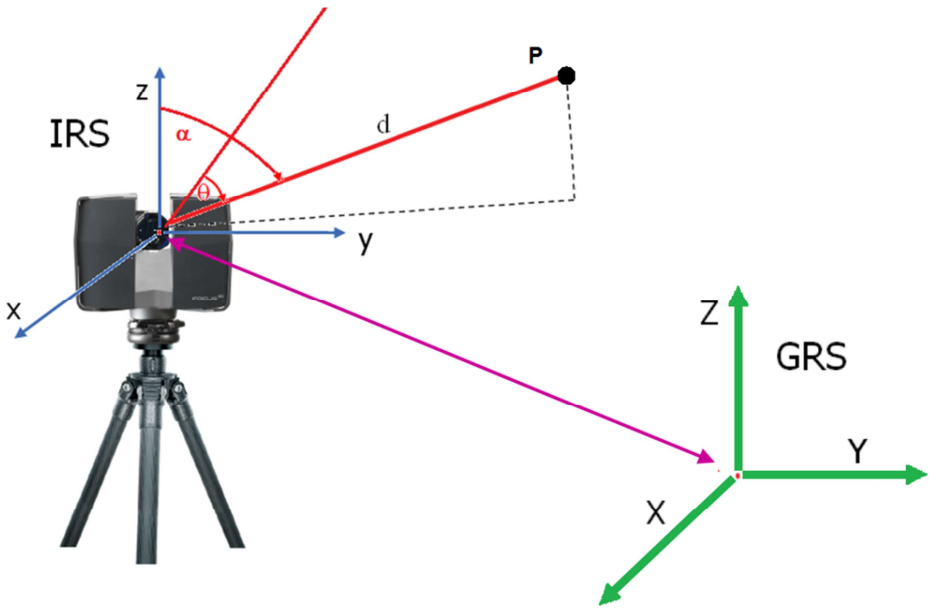


Fig. 6.8. The general scheme of laser scanning measurements and registration procedure.

In target based registration approaches targets are used as corresponding elements between scans. They are made of special materials and/or with a particular shape and can be automatically detected and matched in each laser scan. However, artificial elements have to be added to the scene, which is not always possible, and have to be

placed in positions which are stable in time to allow a good scan registration. This may be a difficult task in the case the same object needs to be scanned at different epochs. In addition, a good distribution of targets in the analysed scene is not always achievable because of practical reasons (e.g., when some part of the object cannot be reached). The use of geometric features can be a partial solution to these problems. Indeed, no artificial targets need to be placed and the geometric variation of feature object between different epochs is generally negligible.

In many applications also *surface matching* techniques are used for scan registration. Among all these methods, the Iterative Closest Point (ICP) algorithm (Besl and McKay 1992) is the most exploited. ICP for point cloud registration works without any pre-knowledge about the point-to-point correspondences. It establishes point-to-point correspondence iteratively based on the minimum Euclidean distance. If the initial values are good enough, this procedure usually converges. ICP is a pair-wise registration procedure, and cannot reliably handle simultaneous registration of multiple scans. This results in the propagation of registration errors as more scans are acquired and added to a project. In addition, ICP requires good initial approximations to solve for the relative orientation unknown parameters. For this reason some manual measurements of a few correspondences are generally needed. Different approaches have been proposed to make ICP more robust by filtering the correspondences for effective handling of occlusions (Zhang 1994, Guehring 2001). A comparison of different ICP variations is given by Rusinkiewicz and Levoy (2001). Compared with standard surface registration approaches (like ICP) registration with features allows for global registration (Scaioni and Forlani 2003). Indeed, standard ICP registration approaches can only handle a pair of scans at a time. This leads to accumulation of errors and sub-optimal use of available information. Conversely by using global registration leads to optimal use of available data. In addition, as noted by Bennamoun and Mamic (2002) ICP just produces registration without giving any information about reliability and confidence of the estimated registration parameters. Conversely the presented approach is a direct application of Least Squares fitting and gives a full covariance matrix of the fitted parameters.

This section presents a registration method based on the identification of planar features in the acquired point clouds. Unlike the procedure adopted in Part A, rather than registering scans first and then modelling, first the unregistered scans are analysed and planar objects detected. Then the corresponding models are used to co-register scans.

The possibility to used geometric features for scan registration solution is a well know topic in the literature. Similar approaches are presented in Dold and Brenner (2006), Rabbani et al. (2007), Wang and Brenner (2008) and Van Gooret al. (2011). For this

reason the approach discussed in this section is not revolutionary with respect to methods already published, but it attempts to increase the automation degree and the robustness, as well as to get rid of manual measurements, by using the segmentation algorithm described in Chapter 1 as the basis for scan registration.

Man-made and urban environments consist mainly of objects that can be modelled by using a set of a few well defined objects characterized by a planar geometric model. Determining the transformation parameters of different scans can therefore be based on aligning planar features often present in these environments (Dijkman and van den Heuvel 2000).

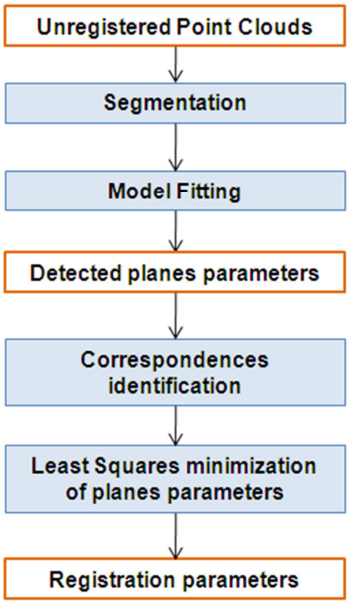


Fig. 6.9. Workflow of the developed methods for scan registration with planar features.

The developed method, presented in Fig. 6.9, works toward the solution in the following four steps:

1. points are labelled as belonging to a certain planar object. This is done for several objects in different unregistered scans. This step can be automated by using the segmentation algorithm presented in Chapter 1;
2. a Least Squares-based fitting algorithm calculates the object parameters for every object in each scan;
3. correspondences between planes belonging to different scans are established. Although the correspondence specification in the actual implementation is done manually this process can be automated using either exhaustive search or RANSAC based random search through the parameter space. In particular, in the case of planes, a minimum of three correspondences are required for registration; and

4. the final transformation parameters are calculated by Least Squares.

The first steps of the developed approach are already discussed in the previous chapters or are simple tasks. Only the Least Squares minimization of planes parameters is presented in the next section.

6.2.1. Plane parameterization and scan registration

In urban environments many objects consist of one or more planar faces that can be used for registration. Describing a plane by the normal vector $\boldsymbol{n} = [n_x\ n_y\ n_z]^T$ and the perpendicular distance from the origin ρ (Fig. 6.10) provide a singularity free representation for infinite planes (Van den Heuvel 1999). This representation is also known as Hesse form of the plane and is more suitable for the Least Squares solution than other parameterizations. As a plane has only three degrees of freedom, a constraint on the length of normal vector $\boldsymbol{n}$ is introduced:

$$\|n\| = \sqrt{n_x^2 + n_y^2 + n_z^2} = 1 \qquad 6.2$$

The estimation of registration parameters is obtained by Least Squares minimization of the sum of squared differences. Mathematically it can be expressed as follows:

$$\min_{\{R\},\{T\}} \sum_{i=1}^{C} \left[\Psi_{R,T}\left(n_{i,1}, \rho_{i,1}\right) - \Psi_{R,T}\left(n_{i,2}, \rho_{i,2}\right)\right]^2 \qquad 6.3$$

where there are C correspondences between planes belonging to different scans. Each correspondence is established between two objects *1* and *2*. $\Psi_{R,T}$ is an operator that applies the transformation, defined by R and T, to the plane parameters $\boldsymbol{n_{i,k}}$ and $\rho_{i,k}$ (h = 1, 2). By using this strategy it is also possible to simultaneously deal with N number of scans.

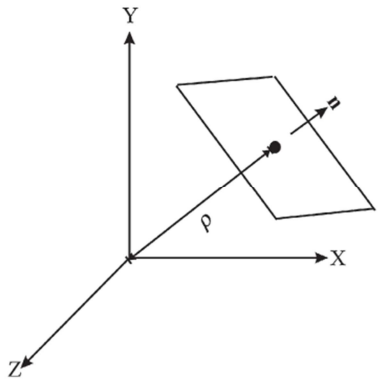


Fig. 6.10. Parameters of the plane, normal **n** and distance from the origin ρ.

In particular each plane provides three equations for the difference in normal vector $\boldsymbol{n}_\delta$ and one equation for the difference in the distance from the origin ρ_δ which is given by:

$$n_\delta = n_1 - \left(R \cdot n_2\right) \tag{6.4}$$

$$\rho_\delta = \rho_2 - \rho_1 + \left(R \cdot n_2\right) \cdot T \tag{6.5}$$

From Eq. (6.4) it is possible to observe that the normal is only affected by the rotation, whereas Eq. (6.5) shows that the change in ρ is a function of both translation and rotation of the scan. To solve the non-linear Least Squares problem in Eq. (6.3) the Gauss-Newton method is used. In particular, by parameterizing R with Euler angles Ω, Φ and K, Eqs. (6.4 and 6.5) can be linearized as follows:

$$A \cdot \Delta x = \delta + V \tag{6.6}$$

where:

$$A = \begin{bmatrix} \frac{\partial n_{\delta,x}}{\partial \Omega} & \frac{\partial \rho_{\delta,x}}{\partial \Phi} & \frac{\partial \rho_{\delta,x}}{\partial \mathrm{K}} & 0 & 0 & 0 \\ \frac{\partial n_{\delta,y}}{\partial \Omega} & \frac{\partial \rho_{\delta,y}}{\partial \Phi} & \frac{\partial \rho_{\delta,y}}{\partial \mathrm{K}} & 0 & 0 & 0 \\ \frac{\partial n_{\delta,z}}{\partial \Omega} & \frac{\partial \rho_{\delta,z}}{\partial \Phi} & \frac{\partial \rho_{\delta,z}}{\partial \mathrm{K}} & 0 & 0 & 0 \\ \frac{\partial \rho_\delta}{\partial \Omega} & \frac{\partial \rho_\delta}{\partial \Phi} & \frac{\partial \rho_\delta}{\partial \mathrm{K}} & \frac{\partial \rho_\delta}{\partial T_x} & \frac{\partial \rho_\delta}{\partial T_y} & \frac{\partial \rho_\delta}{\partial T_z} \end{bmatrix} \tag{6.7}$$

represent the design matrix containing the partial derivatives of Eqs. (6.4 and 6.5) with respect to the registration parameters evaluated at the initial approximations; Δx is the vector of unknowns corrections to the approximate values of registration parameters:

$$\Delta x = \begin{bmatrix} d\Omega \\ d\Phi \\ d\mathrm{K} \\ dT_x \\ dT_y \\ dT_z \end{bmatrix} \tag{6.8}$$

vector δ contains measured minus computed plane parameters

$$\delta = \begin{bmatrix} n_{1,x} - \left(R \cdot n_{2,x}\right) \\ n_{1,y} - \left(R \cdot n_{2,y}\right) \\ n_{1,z} - \left(R \cdot n_{2,z}\right) \\ \rho_2 - \rho_1 + \left(R \cdot n_2\right) \cdot T \end{bmatrix} \tag{6.9}$$

and V contains residuals.

As can be noticed from previous equations partial derivatives of Eq. (6.4) and (6.5) with respect to parameters of scan rotation R and translation T are needed. In particular, by parameterizing R with Euler angles Ω, Φ and K, the partial derivatives with respect to rotation are as follows:

$$\frac{\partial n_\delta}{\partial \Omega} = -\frac{\partial R}{\partial \Omega} n_2 ; \; \frac{\partial n_\delta}{\partial \Phi} = -\frac{\partial R}{\partial \Phi} n_2 ; \; \frac{\partial n_\delta}{\partial \mathrm{K}} = -\frac{\partial R}{\partial \mathrm{K}} n_2 \tag{6.10}$$

$$\frac{\partial \rho_\delta}{\partial \Omega} = \left(\frac{\partial R}{\partial \Omega} n_2\right) \cdot T ; \; \frac{\partial \rho_\delta}{\partial \Phi} = \left(\frac{\partial R}{\partial \Phi} n_2\right) \cdot T ; \; \frac{\partial \rho_\delta}{\partial \mathrm{K}} = \left(\frac{\partial R}{\partial \mathrm{K}} n_2\right) \cdot T \tag{6.11}$$

For translation T:

$$\frac{\partial n_\delta}{\partial T_i} = 0 \tag{6.12}$$

$$\frac{\partial \rho_\delta}{\partial T_i} = \left(R \cdot n_2\right) \cdot \frac{\partial T}{\partial T_i} \tag{6.13}$$

As can be seen from Eqs. (6.4 and 6.5) planes contributes to T only to the direction of the normal vector n. For this reason in order to have a reliable scan registration planes should be evenly distributed on the scene in order to give a proper estimation of T. For this reason a typical proper environment for the presented scan registration strategy using planar features is the case of indoor scans. Indeed, in this case room walls, floors and ceilings give a robust plane configuration for their registration using the presented method.

The proposed cost function presumes equal weighting and uncorrelated object parameters. These assumptions may not be appropriate since significant differences in parameter precision and significant correlations between parameters may exist, depending upon data coverage of the object. These correlations can have negative effects on the convergence of this procedure. This can be taken care of by weighting the equations by the inverse of the covariance matrix obtained during model fitting.

6.2.2. Applications

The procedure for scan registration through the use of planar features was attempted with datasets featuring different characteristics in terms of network geometry and data density. Two examples are illustrated here to present the main advantages and disadvantages of this method in real case studies.

The first example consists in the registration of a single scan pair. These data were acquired with a FARO-FOCUS 3D laser scanner, which is based on phase shift principle for range measurement. The object is a university classroom. The dimensions of the room are 8 m × 4.5 m with a height of 3 m. Each scan (Fig. 6.11) consisted of 28 million points. The scans were segmented and planes recovered using the approach presented in Chapter 1. In particular, the parameters used to segment the two scans are presented in Tab. 6.1. The same parameters are used for both scans to have a uniform accuracy in the estimation of planes.

RANSAC plane threshold ε	1 cm
RANSAC normal threshold α	10 °
Bitmap cell size β	1 cm

Tab. 6.1. Parameters used for "Classroom" and "Office-room" datasets segmentation.

The number of recovered planes in Scan 1 is 15 while in Scan 2 are 16. As previously discussed, the correspondence specification has been accomplished manually although this process can be automated using RANSAC. The coordinate system of Scan 1 was defined as the global coordinate system. Approximate registration values for Scan 2 were obtained by selecting a minimal set of correspondences and solving for the corresponding system of equations. The final sigma nought (σ_0) was 3.0 mm, while standard deviations of registration parameters are presented in Tab. 6.2. The standard deviations of the transformation parameters are obtained by propagating the standard deviations of the point measurements to the object parameters. Some steps of the registration procedure are shown in Fig. 6.11.

Scan	Ω (rad)	Φ (rad)	K(rad)	T_x (m)	T_y (m)	T_z (m)
2	3.105×10^{-5}	2.221×10^{-5}	2.955×10^{-5}	1.504×10^{-4}	2.137×10^{-5}	1.145×10^{-4}

Tab. 6.2. Standard deviations of estimated transformation parameters for 'Classroom' dataset.

To register the same set of point clouds a set of artificial targets were used and compared with the results obtained with the developed registration method. By using target based alignment, a sigma nought of 2.0 mm was obtained. The difference between the translation vectors and the Euler angles Ω, Φ and K obtained from both registration techniques directly provides information about the misalignment (Tab. 6.3). A further check was performed on the coordinates of the artificial targets showing a mean difference of 2.1 mm and a standard deviation of 0.9 mm.

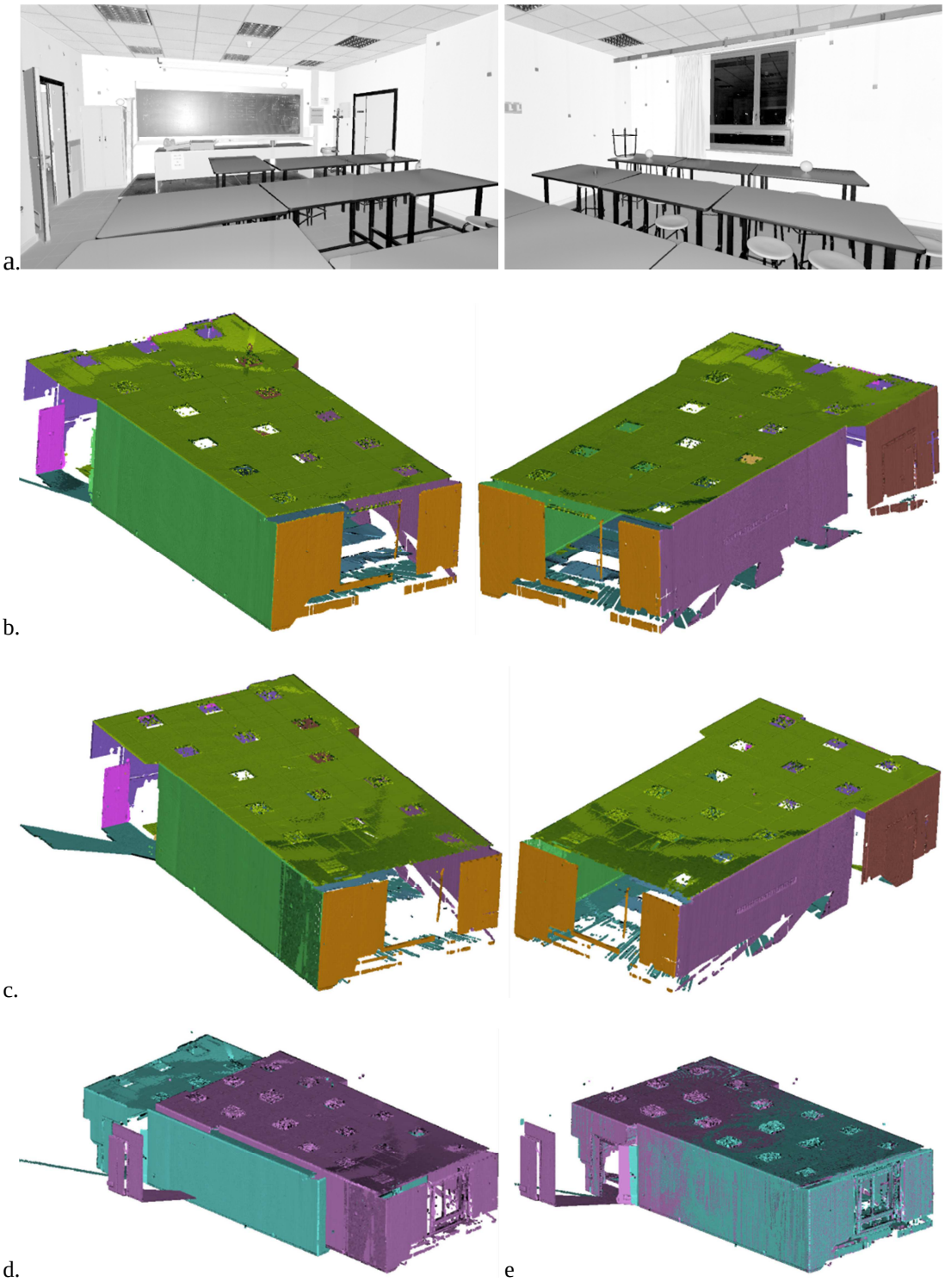


Fig. 6.11. 'Classroom' dataset registration results. An overview of the classroom (a); segmentation results for Scan 1 (b) and Scan 2 (c), corresponding segments between the two scans are represented with the same color; scan alignment before (d) and after the registration (e).

These results are in good accordance with the precision obtainable for target alignment with the laser scanner used in this test.

$\Delta\Omega = 0.0033°$	$\Delta\Phi = 0.0023°$	$\Delta K = 0.0019°$
$\Delta T_x = 0.00057$ m	$\Delta T_y = 0.00032$ m	$\Delta T_z = 0.00051$ m

Tab. 6.3. Comparison between results of proposed and standard target based registration parameters for a pair of scans of 'Classroom'.

The second example consists in the contemporary registration of three scans of an office room (Fig. 6.12). Also in this case the scans are acquired with FARO-FOCUS 3D laser scanner. The parameters used for the segmentation of the scans are the ones presented in Tab. 6.1. The number of recovered objects in Scan 1 – 3 is 20, 18 and 22, respectively. The coordinate system of Scan 1 is defined as the global coordinate system. Once pairwise correspondences between scans are defined, approximate registration values for Scans 2 and 3 were obtained by selecting a minimal set of correspondences and performing an independent pairwise registration to Scan 1. Then the final 3D global adjustment of all scans was performed using pairwise correspondences as described in the previous section. The final sigma nought was 3.2 mm. Tab. 6.4 presents the standard deviations of registration parameters.

Scan	Ω (rad)	Φ (rad)	K(rad)	T_x (m)	T_y (m)	T_z (m)
2	2.988×10^{-5}	3.157×10^{-5}	3.126×10^{-5}	1.998×10^{-4}	1.567×10^{-4}	2.346×10^{-4}
3	9.465×10^{-6}	2.873×10^{-5}	8.793×10^{-6}	2.731×10^{-4}	1.987×10^{-5}	1.312×10^{-4}

Tab. 6.4. Standard deviations of estimated transformation parameters for 'Office room' dataset.

Also in this case a comparison with target based registration was carried (see Tab. 6.5). Differences on artificial targets show a mean difference of 2.4 mm and a standard deviation of 1.0 mm for Scan 2 and mean difference of 3.1 mm and a standard deviation of 0.9 mm for Scan 3.

Scan	ΔΩ (°)	ΔΦ (°)	ΔK(°)	ΔT_x (m)	ΔT_y (m)	ΔT_z (m)
2	0.0025	0.0041	0.0035	0.00050	0.00041	0.00041
3	0.0034	0.0037	0.0022	0.00048	0.00057	0.00048

Tab. 6.5. . Comparison between results of proposed and standard target based registration parameters for 'Office room' dataset.

These results show registration statistics comparable with the ones obtainable by using coded target. This is mainly given by the fact that the indoor scene presents a high number of planar features having an extent covering the entire scene. Also their geometrical distribution was optimal for estimating in a reliable way the registration parameters. Indeed, planes were evenly distributed in all directions. Scenes with a lower number of planar features and a non-regular distribution of planes in space would give significantly worst results. This prevents a large and extensive use of the

presented solution for registration of outdoor scenes. Indeed in this case there are few planes having horizontal direction preventing a reliable estimation of the shift along the vertical.

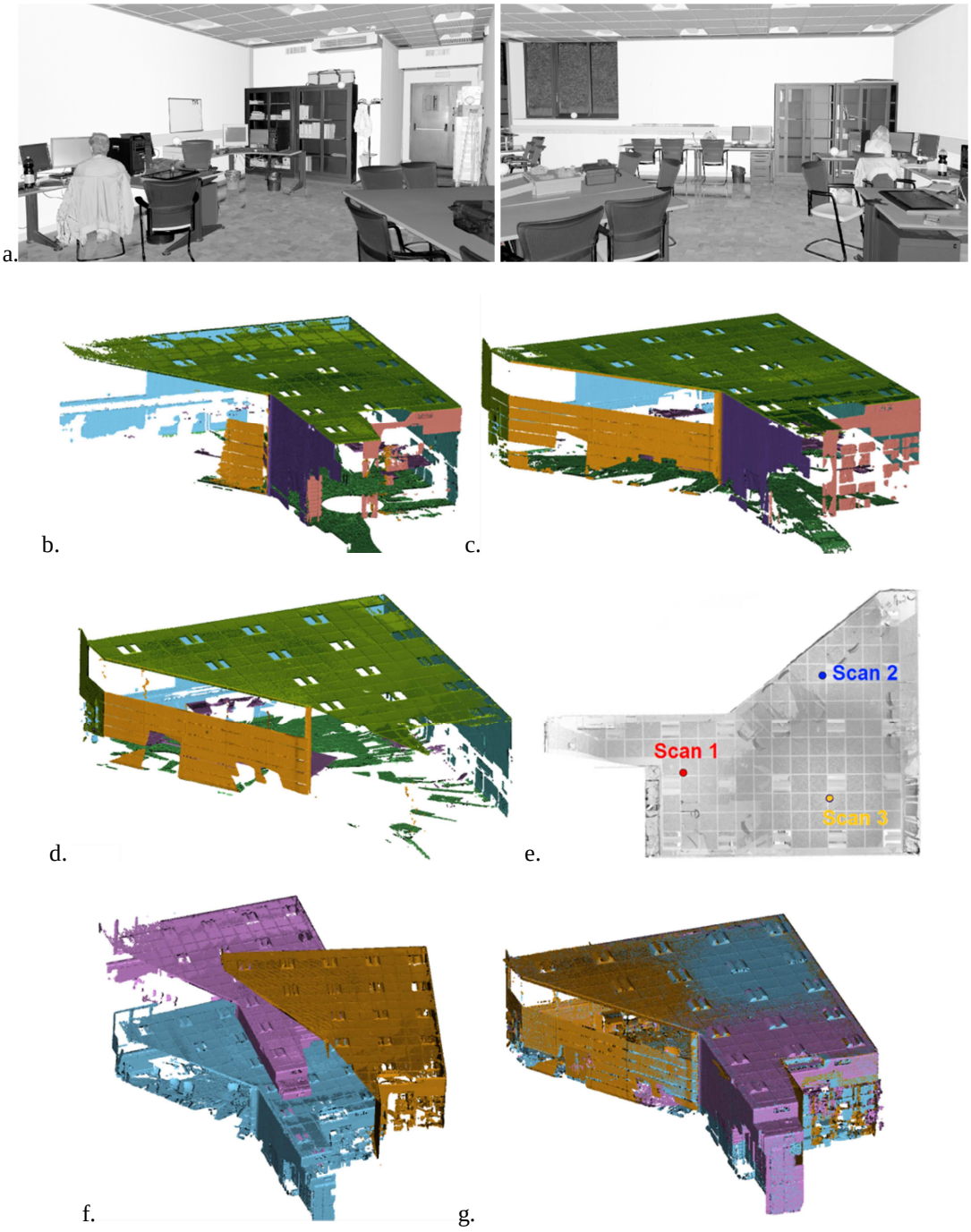


Fig. 6.12. 'Office room' dataset registration results. An overview of the office room (a); segmentation results for Scan 1 (b), Scan 2 (c) and Scan3(c), corresponding segments among the scans are represented with the same color; estimated scan positions (e); and scan alignment before (f) and after the registration (g).

6.3. Indoor reconstruction

Laser scanners are often used not only for generation of detailed models of building façades, but also for indoor environment reconstruction. However, also in this case it is a manual and time-consuming process. The most experiences reported in the literature (see next Subsection 6.3.1) concern the use of static TLS, but the recent improvements in the indoor mobile mapping systems (Biswas and Veloso 2011) are opening new great possibilities in this field. Indeed, the demand for indoor models for different purposes (planned maintenance, preservation and documentation, etc.) has recently increased. Thus a higher degree of automation would better satisfy different applications and speed up the processes.

This section presents an automatic method for modelling predominantly planar indoor environment using laser scanner data. The presented methodology aims at generating reliable models despite the presence of significant amounts of clutter and occlusion, which frequently occur in building indoor. In particular, the developed approach can be used to detect and fill gaps in the input data by detecting occluded regions and model windows or door openings.

6.3.1. State of the art

As previously described in Subsections 1.2 – 2.1 – 3.1 – 4.1 several methods have been proposed for production of building façades models by using laser scanners (Pu and Vosselman 2009, Ripperda and Brenner 2009). These methods operate under the assumption that the surface being modelled is relatively unobstructed. Whether this may be a reasonable assumption for outdoor scanning, in indoor environments objects like piece of furniture and wall-hangings frequently occlude the wall surfaces, making the modelling problem more challenging. Indeed, for façade modelling in the case of occlusions (e.g., due to trees, parked cars, etc.) it is assumed that occluded regions can be observed from different viewpoints (Pu and Vosselman 2009) during the acquisition step. Obviously, such method would not work with significant, unavoidable occlusions as it is the case of indoor environments. An alternative in façade reconstruction is given by model-based approaches (Becker 2009, Koutsourakis et al. 2009). In particular, it is assumed that the occluded region is part of a repeated pattern. A top-down processing is used to predict patterns in façades and to replace missing data with other region that match the pattern, such as one of many identical windows on a wall. However, in the case of interior modelling, repetitive patterns of walls and windows are more unlucky to be identified and ad-hoc algorithms which are robust to clutter are needed.

Currently, model reconstruction and visualization of generic indoor scenarios is still a difficult task (Furukawa et al. 2009). In fact, the reconstruction of interiors is mostly performed using interactive or semi-automatic approaches (Cyclone®, Pointools® . . .).

Many researchers have studied the problem of reconstruction of building interiors using laser scanner data (El-Hakim et al. 1997, Hahnel et al. 2003, Thrun et al. 2004, Budroni and Boehm 2005, Okron et al. 2010, Adan and Huber 2011). For some works the emphasis was given on creating visually realistic models rather than geometrically accurate ones (e.g., El-Hakim et al. 1997). Many authors used TLS data to construct detailed models of walls. Thrun et al. (2004) developed a plane extraction method based on the *'expectation–maximization'* algorithm, while Hahnel et al. (2003) used a *'plane sweep'* approach to find planar regions. However, even if these algorithms work well for extracting planar patches from the laser data, they do not explicitly recognize the identity of components, such as walls, ceilings, and floors. In addition, several works (Hahnel et al. 2003, Thrun et al. 2004, Budroni and Boehm 2005) do not consider the occlusion problem because they focused on modelling of hallways with no furniture or other potentially occluding objects.

'Context-based' building modelling was studied by several other researchers (Cantzler 2003, Nüchter and Hertzberg 2008, Rusu et al. 2008). These approaches rely on hand-coded rules. Recently, Koppula et al. (2011) used a graphical model to represent contextual relationships for recognizing objects in indoor scenes using 3D+colour images (RGBD) from a Microsoft Kinect® sensor (Zhang 2012). However, defined rules are usually brittle and break down in the case of noisy measurements or significant lack in the data.

A very specific area of indoor modelling, which has been the topic of intense efforts for automation, is the reconstruction of industrial environments. These scenes are characterized by repetitive elements such as beams or pipes. Current state of the art in commercial reconstruction tools is manual pre-segmentation in combination with automated fitting (Rabbani et al. 2007).

This section presents an automatic method for modelling indoor environments by using point clouds obtained from static laser scanning. In particular, the proposed methodology addresses the challenges of clutter and occlusion by explicitly reasoning about them throughout the process. First, surfaces representing the room walls, ceiling and floors are detected in a robust way. Then, to understand the nature of occlusions, a ray-tracing algorithm is used to identify regions that are occluded from every viewpoint and to distinguish these regions from openings in the surface (e.g., due to doorways or windows).

6.3.2. Indoor reconstruction method

The developed methodology for automatic indoor reconstruction takes as input a set of registered scans with a known 'up' direction and the location of scanning point in the room. All these prerequisites can be easily obtained in practice. Indeed, scan registration is a well-studied problem, and methods to manual or automatically

register scans are available both in scientific and commercial software packages. Once scans are registered together, scan locations (position and attitude) are known. Also the vertical direction is generally known. This is typically directly provided by the scanner, since it is generally levelled. If the vertical direction is unknown, the orientation can be estimated using statistics on the data. For efficiency, the algorithm operates independently on each room.

The first step in the presented approach (Fig. 6.13) is the detection and estimate of the surfaces to be modelled, i.e., walls, ceiling and floor. However, due to occlusions and clutter some wall may be missing in the dataset. For this reason an automatic procedure is implemented to complete missing elements in a plausible way. To achieve this, the developed algorithm incorporates architectural priors on indoor scenes, notably the prevalence of orthogonal intersections between walls. Once the surfaces describing the room are defined the remaining steps operate on each surface individually. In the second phase, each planar surface is analysed to identify and model the occluded regions and openings by using a ray-tracing algorithm. Openings in the data are detected by using labelling information while contemporarily a further classification is operated between windows and doors. Finally, occluded regions are completed in a realistic way.

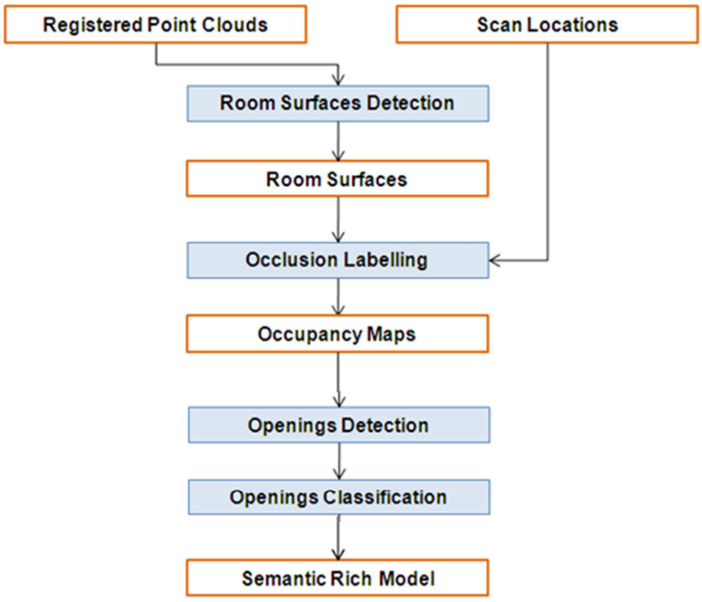


Fig. 6.13. Workflow of the developed methods for indoor modelling.

The output consists of a set of labelled planar patches (wall, floor, and ceiling), adjacency maps indicating which patches are connected to one another, and a set of openings detected within each planar patch. These patches are intersected with one another to form a simple surface-based model of the room. The geometric models of

the room along with semantic information are combined together to derive a semantic rich models of the room in CityGML and/or IFC format.

6.3.3. Room surface detection

The general workflow for room surface detection is presented in Fig. 6.14. The first step in the developed methodology is to detect and estimate the surfaces (walls, roof and ceiling) constituting the room to be modelled. This detection is performed by using the segmentation strategy described in Chapter 1. Once the point cloud segments have been detected a first semantic classification, similar to the one described in Chapter 3, is performed to detect roof and ceiling. Indeed, by analysing the detected segments, the ceiling can be designed as the non-vertical plane having the lower height. Conversely the roof is detected as the non-vertical plane located at the highest level. In this way, the height of the ceiling and roof are determined (Fig. 6.15a).

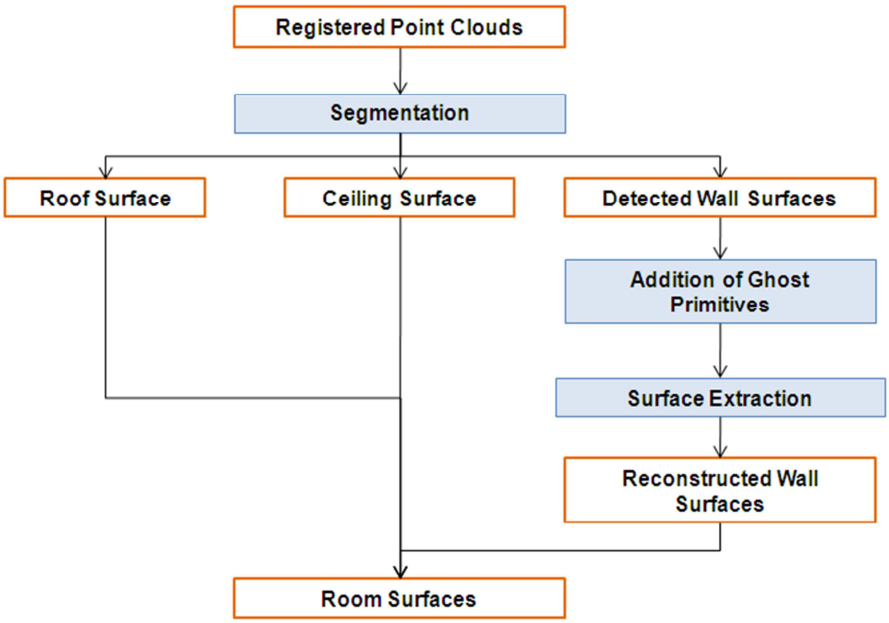


Fig. 6.14. Pipeline for room surfaces detection.

At this stage the floor plan of the room is determined. To this end, the walls constituting the room need to be detected. However, this problem is made more difficult due to clutter and occlusions. Some walls may not be sensed by the laser scanner and may miss in the point cloud. For this reason a proper completion is necessary to reconstruct in a plausible way these missing walls (Fig. 6.15b).

A first rough floor plan of the room can be obtained by projecting the points belonging to the ceiling onto a horizontal plane. Indeed, the acquisition of the ceiling surface, due its location, is generally less influenced by clutter and occlusion than

other surfaces in the room. The horizontal plane is discretized into cells of size β x β and then an occupancy map is generated. In this map, white pixels represent elements where TLS data are available, while black pixels are grid elements with no data. Starting from this binary image it is possible to derive pixels representing the boundary of occupied cells which represent a first rough floor plan.

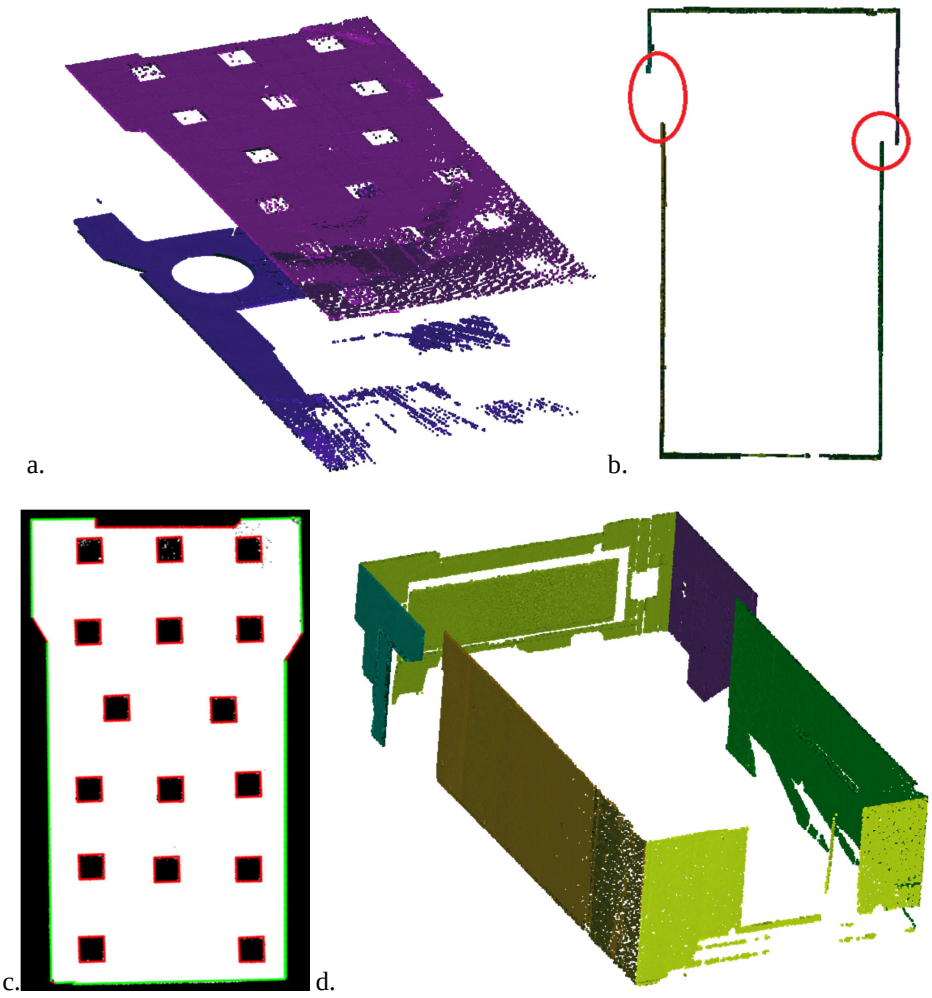


Fig. 6.15. Detection of room surface for 'Classroom' dataset: (a) detected ceiling and floor; (b) some wall portions are missing (red circles) due to occlusions, (c) occupancy map of the ceiling with real walls (green segments) and spurious boundaries (red segments); (d) detected real wall surfaces.

The optimal cell size β, as previously discussed, is a function of the mean sampling resolution in the point cloud. Due to occlusions the obtained plan may contain some spurious boundaries, i.e., the ones not associated to a wall (Fig. 6.15c). To validate the obtained boundaries a check is done with the segmentation results. In particular, only vertical segments falling inside the cells labelled as boundary are considered as real wall surfaces (Fig. 6.15d).

As previously discussed, a single small missing wall may jeopardize the entire reconstruction of the floor plan. In the developed strategy, these gaps are filled by incorporating additional, unseen '*missing walls*' (Chauve at al. 2010). In indoor environment it is possible to observe that walls intersect orthogonally. For this reason missing walls are guessed as orthogonal to detected walls and are added from boundary of detected ones (Fig. 6.16a).

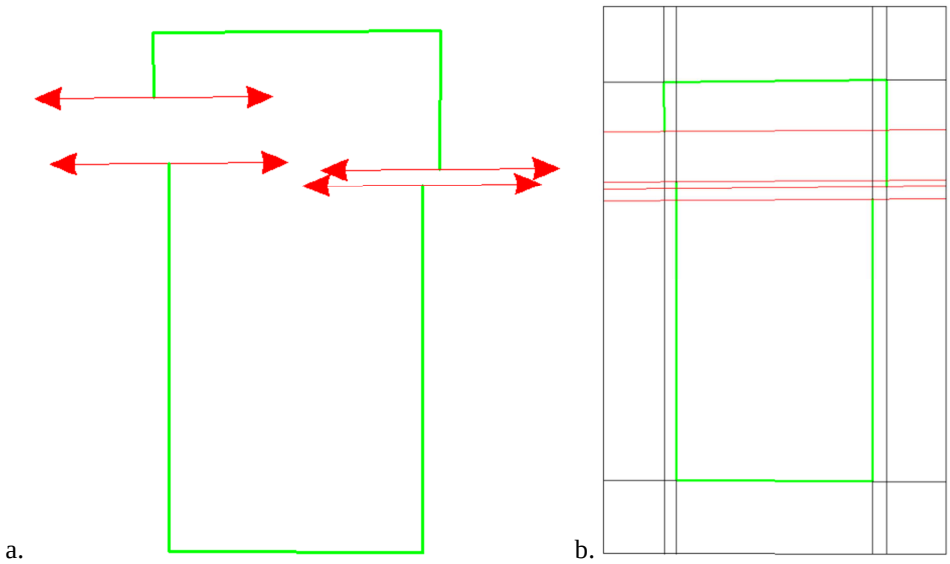


Fig. 6.16. 'Cell complex' construction for 'Classroom' dataset: (a) detected primitives with *'missing walls'*; (b) induced 'cell complex'.

To obtain a continuous floor plan from detected and missing walls, a procedure based on a 'cell complex labelling' is performed. As a first step a 2D arrangement (Edelsbrunner et al. 1986) is set up by using both detected and 'missing' walls. The arrangement generates a partition of the original space domain into convex polygonal cells (Fig. 6.16b). The final partition of the complex do not depends on the order in which primitives are inserted because each primitive cuts the entire complex domain.

The floor plan reconstruction can be formulated as an optimal binary labelling of the 'cell complex'. Each cell is labelled as empty or occupied, and the floor plan can be extracted as the union of all facets separating an occupied cell to an empty one, obtaining this way a watertight and intersection-free boundary. This labelling problem is handled within the framework of minimum s - t cut on the cell-adjacency graph G =(V, E) of the partition: the vertices V are the cells of the polygonal 'cell complex' while the edges E link adjacent cells, i.e., they correspond to the facets of the complex. V is augmented with two additional seeds, a source s and a sink t, with edges from s to each cell and from each cell to t. All edges have non-negative weights w. A s - t cut (S,T) is a partition of V into two disjoint sets S and T such that $s \in S$ and $t \in T$. The cost of an s - t cut is the sum of the weights of the edges from S to T.

Efficient algorithm with low-polynomial complexity exists to find the *s* - *t* with minimal cost, allowing a global minimization of the energy. A graph partitioning (*S*,*T*) corresponds to a binary labelling of cells (Fig. 6.17), where cells in *S* and *T* are respectively empty and occupied, and the cost of the cut to the energy of the associated surface. Weights of edges joining the source or the sink penalize the associated cells, while weights of edges between two cells penalize the associated facets.

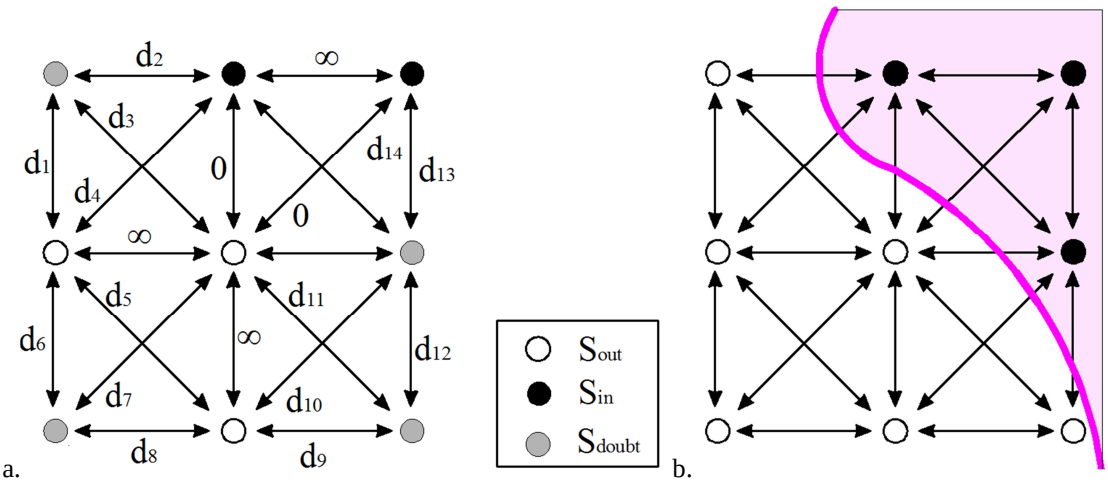


Fig. 6.17. Example of s – t cut: (a) graph construction and cost assignment; (b) final binary labelling with minimum cost.

Starting from the available data some cells can be directly categorized as occupied. In particular, all cells occupied by points belonging to the ceiling can be directly assigned to set *T*. In a similar way cells bordering an occupied cell and separated from it by a detected wall segment are set as empty. For this reason weights of edges joining the sink to cells labelled as occupied are set to infinite and, in a similar way, edges joining the source to empty cells are set to infinite. Then, weights of edges between two occupied cells and between two empty are set to infinite (Fig. 6.18a) and weights of edges connecting an empty and an occupied cell are set to zero. In this way, it is prevented that cells forming the inner part of the room are erroneously labelled as empty or vice-versa. For all other cells an equal unitary weight is associated to edges joining the cells to the source and to the sink because there is no a priori knowledge about the occupancy of a cell. Weights of remaining edges between cells are fixed equal to the length of the edge between the cells. This means that the *s* - *t* cut problems is aimed at minimizing the length of guessed walls segments (Fig. 6.18b). To perform s – t cut, the Kolmogorov's max-flow algorithms is used (Boykov and Kolmogorov 2004). Once having computed the *S*, *T* partitioning, the boundary of the occupied cells of the polygon partition gives the floor plan.

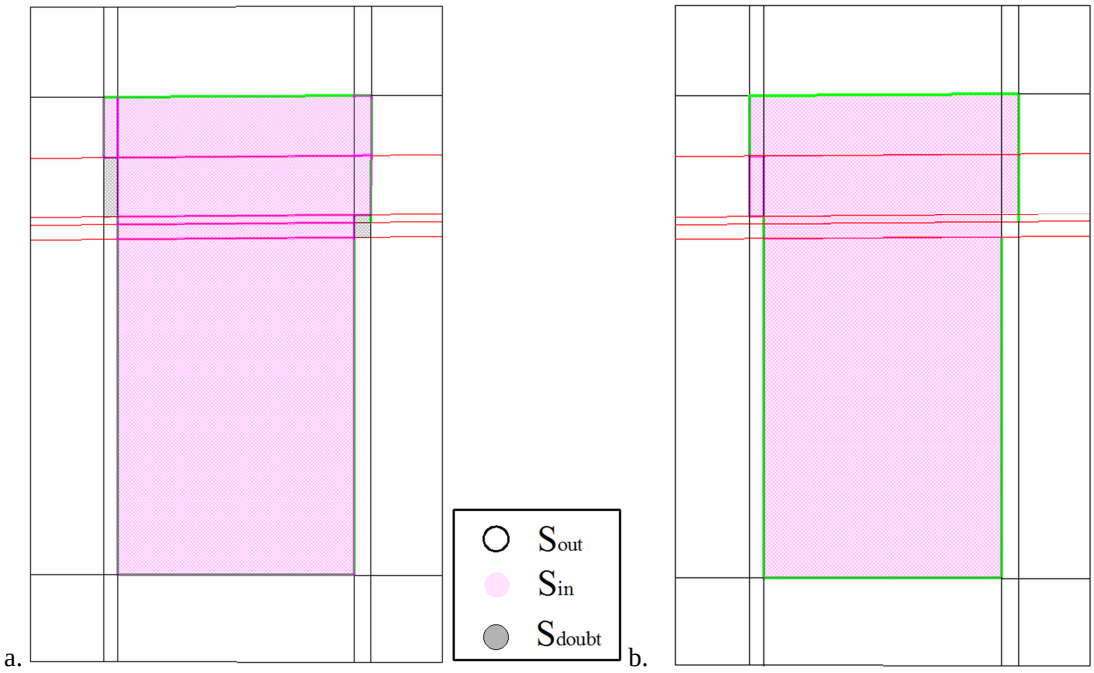


Fig. 6.18. 'Cell complex' labelling for 'Classroom' dataset: initial (a) and final (b) labelling of the 'cell complex'.

Finally, having obtained the floor plan and having previously defined the ceiling and floor heights the surfaces constituting the room can be easily defined.

6.3.4. Reconstruction of openings under occlusions and clutter

Once all surfaces constituting the room are detected the presence of openings is investigated. Detecting the boundaries of openings, such as windows or doors, in a wall is a difficult task. While in façade reconstruction applications windows are detected as holes in the façade point cloud, this does not generally hold for indoor environments. Indeed, also occlusions and clutter produce significant holes in the point cloud which have to be distinguished from real openings. In addition, also an opening may be partially occluded increasing the complexity of the problem. To this end a ray-tracing labelling is performed and an occupancy map is generated (Adan and Huber 2011).

In this step each surface is separately processed. Once having defined previously the wall surface, points representing inliers for the defined plane can be easily detected. The detected plane is then discretized into cells of size β x β and then an occupancy map (denoted as M_0) is generated based on whether inlier points are detected at each pixel location or not. Without additional information, it is not possible to distinguish between a pixel that is truly empty and one that is merely occluded. This problem can be solved by using a ray-tracing labelling to detect occlusion between the sensor and

the surface being modelled. For this reason the scanning locations (position and attitude) should be known.

Let $S = \{S_1, S_2, \ldots, S_N\}$ be the set of scan for the room to be modelled. For each scan position S_K, a labelling L_K is generated by tracing a ray from the scan location to each pixel $P_i(x,y,z)$ labelled as empty in M_0. Starting from cartesian (xyz) coordinates of P_i its spherical coordinates in the Intrinsic Reference System (IRS) of the scanner device can be determined. Indeed, by knowing the scanning location (position and attitude) the cartesian coordinates in the IRS can be obtained. Then from Eq. 6.14 spherical coordinates (range distance d, vertical or nadir angle α, horizontal or azimuth angle θ) can be derived.

$$\begin{bmatrix} x_{IRS} \\ y_{IRS} \\ z_{IRS} \end{bmatrix} = d \cdot \begin{bmatrix} \cos\ \alpha\ \cdot \cos\ \theta \\ \cos\ \alpha\ \cdot \sin\ \theta \\ \sin\ \alpha \end{bmatrix} \qquad 6.14$$

Having defined cell location in spherical coordinates, the 20-nearest neighbours for P_i can be easily defined. As a measuring distance between P_i and other points, θ and α angles are used. In the case the 20-nearest neighbours points have angular distance (θ, α) far larger than the predefined angular scanning resolution, this would mean that no reflected signal come to the laser scanner due to the presence of an opening. In this case the pixel P_i is labelled as empty.

Conversely, if the angular distance is compatible with the predefined scanning resolution, the mean distance of the nearest neighbours is evaluated d_{mean} = mean (d_1, d_2, .., d_{20}). In particular, in the case the mean distance is lower than the P_i distance, this would mean that P_i is occluded by some points in the scan and the cell is consequently labelled. On the other hand if the mean distance is larger than the P_i distance, the cell is labelled as empty. After this ray-tracing labelling for all the scans, K labels for each pixel are obtained (Fig. 6.19). Then all the labels are combined in a final occupancy map L_F adopting the following labelling rule:

$$If \quad L_0(i) = empty \quad and \quad L_j(i) = occluded\ , \forall j = 1,2,\ldots, K \Rightarrow L_F(i) = occluded$$

In other words, a cell is considered occluded if it is occluded from every scan-point (Fig. 6.19).

Having obtained the occupancy map, openings can be easily detected by identifying cells' labels. Then a procedure similar to the one described in Paragraphs 2.2.1 and 2.2.2 is used to determine the shape of openings. A further classification is then performed between openings to distinguish between doors and windows. In particular, are classified as doors those openings intersecting with the ground.

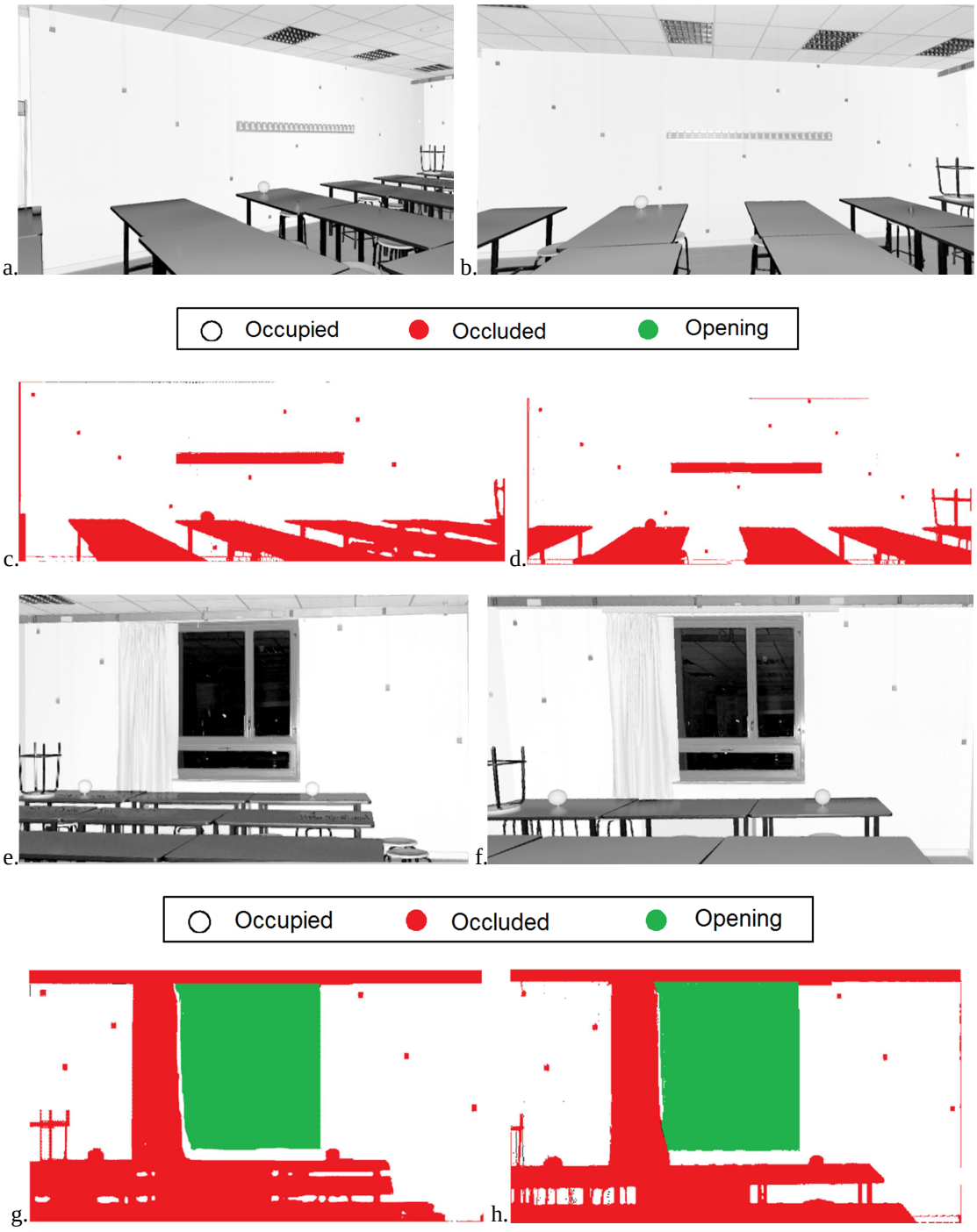


Fig. 6.19. Reconstruction of openings. Reflectance images of Scan 1 (a) and Scan 2 (b) used to model the wall; cell labelling for Scan 1 (c) and Scan 2 (d). Reflectance images of a wall with an opening of Scan 1 (e) and Scan 2 (f); cell labelling for Scan 1 (g) and Scan 2 (h).

6.3.5. Applications

To evaluate the presented modelling methodology the two datasets presented in Subsection 6.2.2 are used.

The ‘Classroom’ dataset presents significant occlusions and clutter. In addition, a recess of the room is partially occluded by another wall. Once having registered the scan, the first step for indoor modelling is the segmentation of the point cloud. The following parameters were used:

RANSAC plane threshold ε	1 cm
RANSAC normal threshold α	20 °
Bitmap cell size β	1 cm
Bitmap cell size for wall detection	5 cm

Tab. 6.6. Parameters used for indoor modelling.

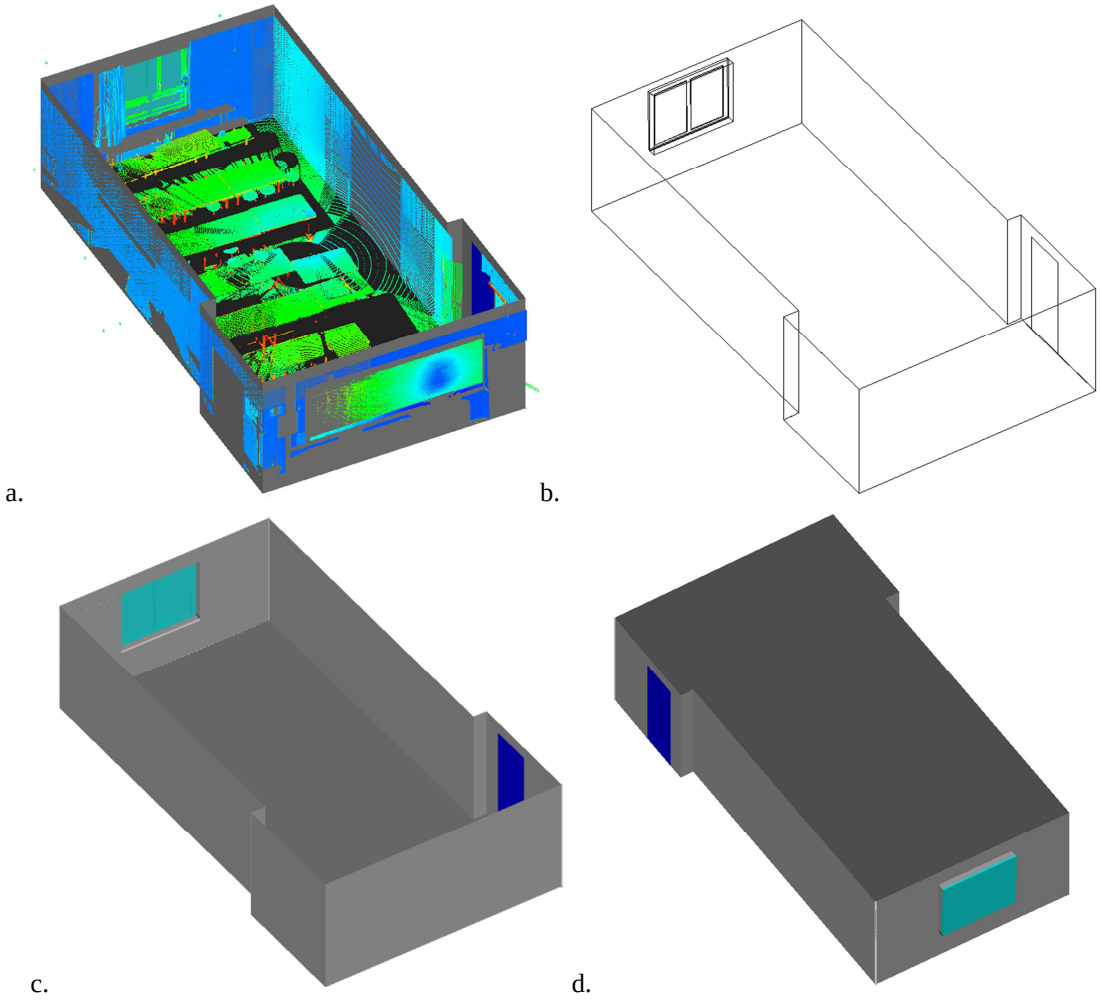


Fig. 6.20. Geometric reconstruction of ‘Classroom’ dataset: (a) room model with overlaid point cloud; and (b-c-d) final 3D room model.

In Fig. 6.15a-b segmentation results are presented and distinction between room elements (floor, ceiling and walls) is presented. Once having defined the detected wall surfaces, ghost primitives are added and the complex arrangement is set up (Fig. 6.16). Finally, occupancy maps are generated for all the detected wall surfaces. The

obtained room model is presented in Fig. 6.20. To evaluate the geometric accuracy of the model a performance analysis similar to the one presented in Section 5.1.2 is carried out (Fig. 6.21).

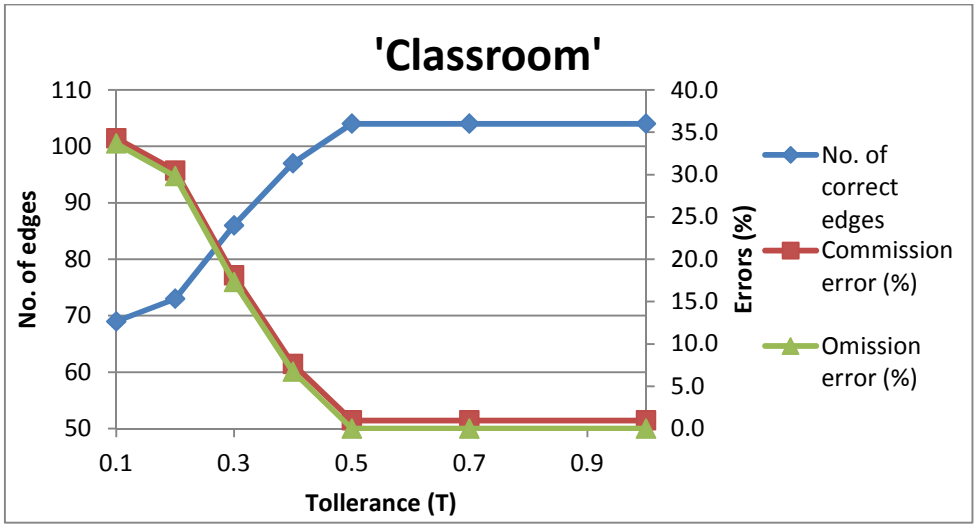


Fig. 6.21. Geometric reconstruction analysis for 'Classroom' dataset.

In particular, it can be noticed that the accuracy of the wall and opening boundaries is similar to the one obtainable with manual modelling of the point cloud. These results are confirmed also for the 'Office room' dataset (Fig . 6.22). In Fig. 6.23 a summary of the main processing step for the 'office-room' dataset are presented.

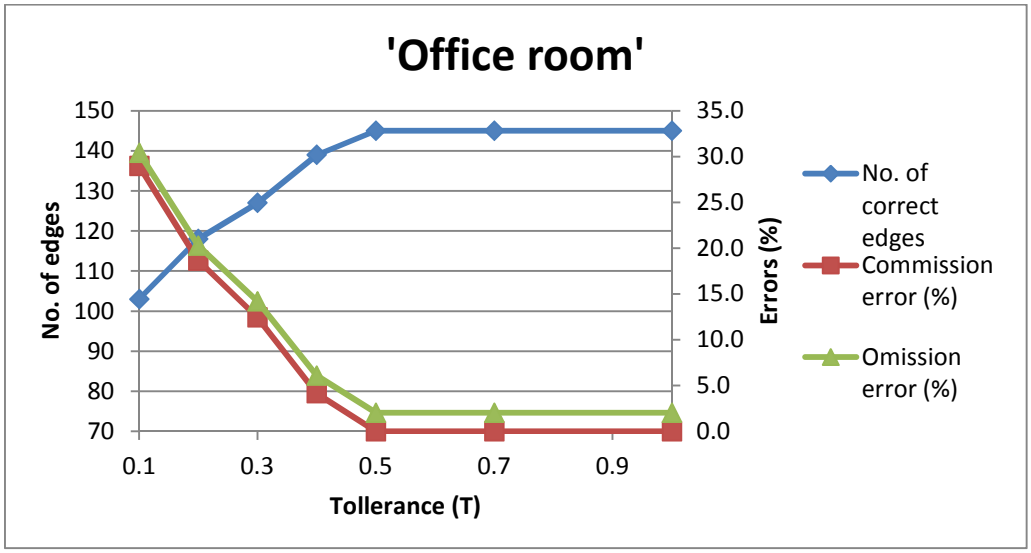


Fig. 6.22. Geometric reconstruction analysis for 'Office room' dataset.

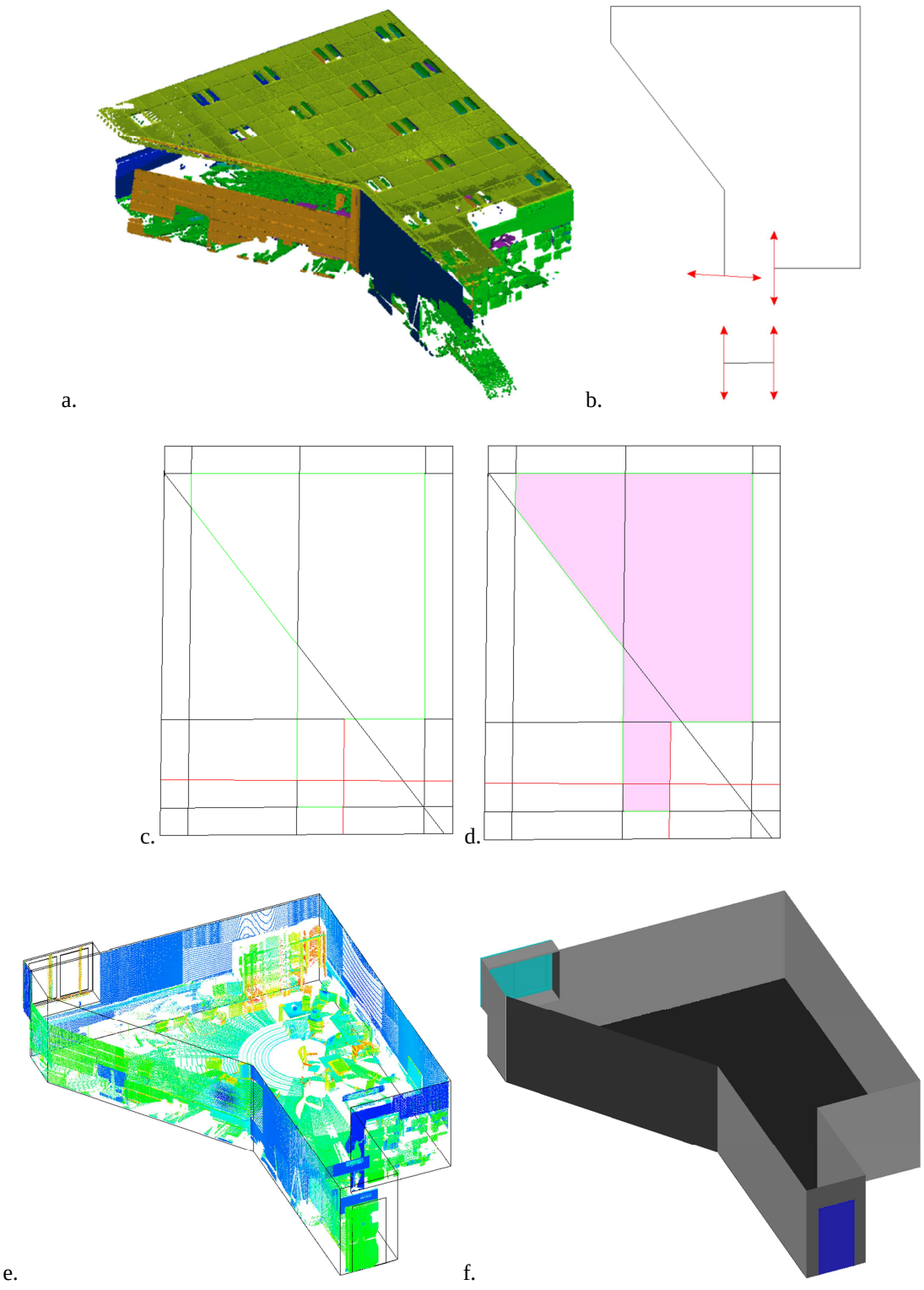


Fig. 6.23. Geometric reconstruction of 'Office room' dataset: (a) segmentation results; (b) detected primitives with '*missing walls*'; (c) induced 'cell complex' and (d) final labelling; (e) room model with overlaid point cloud; and (f) final 3D room model.

Conclusions and further developments

This dissertation presented an automated methodology for building façade reconstruction aimed at supporting the thermal retrofitting of buildings built in the period 1950 - 1975. To achieve this final goal, a number of algorithms were developed to overcome different problems within the general framework. Both existing and new techniques were implemented.

This research is based on the observation that bottlenecks in the terrestrial laser scanning (TLS) field, which is still the primary source of point clouds adopted in façade modelling, has been rapidly changed during the last decade. Indeed, for a long period the major limitations were due to the slow acquisition speed and manual registration problems. Fortunately, the hardware of terrestrial laser systems has been significantly improved in recent years, and the bottleneck of data acquisition has become less severe. Hereby, terrestrial laser scanning (TLS) of large areas is getting much faster and easier. Also the registration problem has been mainly overcome thanks to many consolidated techniques that have been introduced into practice. Consequently the bottleneck moved from the acquisition and registration to the modelling stage. This is partially due to the fact that the extraction of façade models requires the semantic interpretation of a very large variety of possible architectural elements, depending on the building typology, construction time, and style. Whether these tasks may be accomplished with ease (although in a time consuming fashion) by trained human operators, they are a critical issue for machine learning systems. In addition, point clouds may typically exhibit significant missing data due to occlusions increasing the complexity of the problem. However, even if façade modelling is a very active research field up to now, three are no commercial software packages allowing for automated façade modelling starting from point clouds.

This research tries to give contribution to increase automation in the process that goes from the raw point cloud, to the geometrically enriched model of a building and finally to feed a Building Information Model..

The developed approach can be considered as a multi-step process in which the building model is progressively and iteratively estimated and refined.

The segmentation algorithms developed in this work show that detection of planar façade objects can indeed be helpful for efficient processing of large point-clouds. This is because primitive shapes possess several advantageous aspects which can be directly exploited in the modelling pipeline. First of all, because of the simple nature of the primitives, the user-controllable settings for the detection algorithm are very intuitive and are directly related to easily understandable geometric properties such as distance and normal deviation, which moreover are often of immediate relevance to the ultimate application. In addition, a single primitive can resume a large number of points and can serve as a basis for an efficient encoding of the building geometry in a more compact way. It is foreseen that the method will fail in detecting correct building outlines for complex building structures, such as curved walls, curved protrusions, or columns. More complete set of primitives might be included to detect these features. However, for the buildings target of this work almost all relevant façade elements can be modelled with planar shapes. Even if a lot of works was done for reducing bad-segmentation, some problems still exists and some over-segmentation was found. For example, highly non-uniform distributed laser points can be seriously over-segmented at the sparse regions such that there is no sufficient clue for merging. In addition, the optimality of the decomposition is not guaranteed. For comparatively small tolerances, as used throughout this work, the suboptimality of the detection algorithm is mostly irrelevant because there is only a very limited set of possible segmentations since they have to fit the data tightly. Once the allowed tolerances are relaxed though, the number of possible segmentations dramatically increases and the suboptimality of the algorithm clearly shows up. Finally, the definition of proper parameters in the processing has to be defined by the user. The chance to have them defined in an automatic way would give a further important step to the automation of the entire process.

The generation of the vector model of a building is the key task in the overall procedure because the geometric accuracy of the final model is the most important requirement for the digital model to be used for retrofitting drawing generation. The key factor is the implementation of some architectural priors like the prevalence of straight lines and orthogonal intersections in modern buildings, especially in the ones constructed in 1950 - 1975. The exploitation of this further constraint gives an important contribution to overcome problems reported by different works in the literature. In particular, the edge smoothing has provided good results in the regularization into lines and curves of long edges. In addition, the enforcement of surface intersections not only helps define the façade topology, but also increases break-line accuracy. Nevertheless, some problems still remain in correspondence of

curved and complex surfaces. In correspondence of these edges, the implemented algorithm may be significantly affected by noise. In addition further problems may arise when enforcing intersection of surfaces having curved edges. Some problems still exist in the case of high fraction of missing data. For those cases, a first automatic modelling followed by a manual editing to fix the problems seems the most adequate procedure solution to save time and minimize the human effort.

The procedure for extracting semantic features from TLS point cloud is of primary importance for enriching with semantics the building model. This task gives an important contribute to the automation of 'scan-to-BIM' processing. Although the method worked well with the target datasets, the applicability is limited to the building shapes defined in the classification rules. For example, in the current implementation doors are described as 'low features inside a wall façade,' but doors might also appear inside a protrusion/intrusion façade (incomplete knowledge). This limitation can only be solved by improving and extending the defined rules. This means that for different building types, new façade classification rules need to be defined. However, the success of feature extraction not only depends on the classification rules, but also relies on the uniformity and completeness of the TLS data acquisition. Indeed, if a feature is partially occluded, the calculated geometric attributes and topology relations with other features would be unreliable. This may lead in turn to incorrect feature extractions.

In some cases, building features may be partially or completely occluded during scanning. If a feature is completely occluded, it might be recovered later by exploiting high-regularity of building façades. Indeed, building façades target of this work can be represented as the repetition of some basic-geometries into multiple 2D periodic structures (lattices). Identification of these lattices may help fill holes and reducing occlusion problems. However, this technique can be used only in the case of building types presenting high regularity while it is not applicable to façades which cannot be described by a rectilinear mixture model, e.g., façades with completely irregular repetitive patterns.

The examples proposed in Section 5 show that the proposed approach can successfully deal with different data acquired with different sensors.

The application of this methodology to support thermal retrofitting may take advantage from integration of other data, for example RGB and thermal images, into the reconstructed model. An ad-hoc procedure has been developed to this purpose. Detection and matching of geometric features could also be used for registration of laser scans in urban environments. Indeed, urban environments can be mainly decomposed in a set of few basic primitive shapes and planar features are

undoubtedly the more frequent ones in architectural scenes. Although, as things stand now, it is difficult to forecast a massive use of such a method in complex practical projects. For some indoor surveys where no artificial targets can be used the presented approach may give satisfactory results.

Finally, the procedure was adapted to reconstruct indoor environments. Even if some common problems are shared by outdoor and indoor environments, the latter case presents some specific issues. For this reason the previously developed algorithms were adapted to fit these situations.

Future investigations

While this research has been able to demonstrate the merit of a fully automated methodology for façade modelling, there still remain many avenues to pursue for future research. Indeed, quality and completeness of manual models is still superior to the ones obtained from automatic modelling. As previously anticipated this is related to the fact that generation of façade models requires the interpretation of point cloud data from a geometrical, functional and sematic point of view. This is even more complicated by the large variety of possible architectural elements, depending on the building typology, construction time, and style.

For detection of primitives, further improvements in efficiency can be envisioned, especially if additional information is known in advance, for example the largest possible extent of a single primitive or the relative pose of certain ones. This means integrating the segmentation and the feature detection phase in a single step. Also the feature classification should be further developed to achieve more accurate feature recognition and incorporate more feature types. The improvements should be focused on two aspects. First, a more exact model to represent the uncertain situations should be raised. For example, the conditional belief between clauses should be considered. Second, further building types should be investigated and new rules formulated. However, extending the developed strategy to further building styles means that the ability of choosing the right classification rules should be developed. For this reason in the next level of research, the ability to ‘remember’ and ‘learn from’ experiences needs to be integrated into the processing pipeline. In this way the feature recognition should be more accurate when the procedure operates on a similar building.

To achieve more accurate reconstruction, integration of imagery can be exploited. Image-based building reconstruction has been investigated for many years. From multiple 2D images captured from different positions, 3D coordinates of the image features (lines for example) can be calculated. Although acquisition of images is cheap and easy, the problems of image understanding make it still difficult to

automate the reconstruction process on the basis of images only. On the other hand, laser scanners allows for explicit and accurate 3D information extraction. For this reason, the laser data and images are complementary to each other, and efficient integration of the two data types will lead to a more accurate and reliable extraction of three dimensional features. Therefore, an image-based refinement method which uses strong line features extracted from images could be exploited to improve the building façade models generated from TLS point. This refinement not only may fix the models' geometry errors, but also could help overcome inconsistencies between laser and image data like occlusions. Even if different authors worked on that (Vosselman 2004, Brenner 2005, Nex and Rinaudo 2009), a reference solution is not available yet.

The developed idea of repeated pattern detection for model completion can be extended to other building styles than the ones target of this research. Indeed, the highly regular nature of building façades is widely demonstrated all over the world. For example, also historical 20th century buildings present complex decoration but arranged in regular patterns. The detection of these elements and their regularity can be used to speed up the modelling process of historical buildings. The base element can be modelled in an automatic or semi-automatic way, and then its repetition could be automatically recognized reducing this way the time needed for manual modelling. Also the detection of symmetries might also be beneficial to the completion of objects in several respects.

Bibliography

Abdel Aziz, Y.I. and Karara, H.M., 1971. Direct linear transformation from comparator coordinates into object-space coordinates. Close-range Photogrammetry. American Society of Photogrammetry, Falls Church, Virginia. 433 pages: 1-18, (1971).

ACE, 2012. Annual report of the Architects' Council of Europe, 2012.

Adan, A. and Huber, D., 2011. 3D reconstruction of interior wall surfaces under occlusion and clutter. In 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2011 International Conference on (pp. 275-281).

Alshawa, M., Boulaassal, H., Landes, T., Grussenmeyer, P., 2009. Acquisition and automatic extraction of facade elements on large sites from a low cost laser mobile mapping system. In In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences/3DARCH09.

Alshawabkeh, Y., Haala, N., Fritsch, D., 2006. 2D-3D feature extraction and registration of real world scenes. In ISPRS Commission V Symposium Image Engineering and Vision Metrology (pp. 32-37).

Africani, P., Bitelli, G., Lambertinia, A., Minghetti, A., Paselli, E., 2013. Integration of lidar data into a municipal GIS to study solar radiation. In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W1, ISPRS Hannover Workshop 2013, 21 – 24 May 2013, Hannover, Germany, 1-6.

Anil, E., Akinci, B., Huber, D., 2011. Representation requirements of as-is building information models generated from laser scanned point cloud data. In: Proceedings of the International Symposium on Automation and Robotics in Construction (ISARC), (Seoul, Korea).

Awwad, T. M., Zhu, Q., Li, Q., 2009. An Extension of. In Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on (pp. 1-4). IEEE.

Ballard D., 1981, Generalizing the hough transform to detect arbitrary shapes. Pattern Recognition, 13(2):111-122.

Barnea, S. and Filin, S., 2010. Geometry-image-intensity combined features for registration of terrestrial laser scans. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38, Part 3A: 145-149.

Bauer, J., Karner, K., Schindler, K., Klaus, A., Zach, C., 2003. Segmentation of building models from dense 3D point-clouds. In Proc. 27th Workshop of the Austrian Association for Pattern Recognition (pp. 253-258).

Becker, S. and Haala, N., 2007. Refinement of building facades by integrated processing of LIDAR and image data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Science, 36, 7-12.

Becker S. and Haala N., 2009. Grammar supported façade reconstruction from mobile LIDAR mapping. In Proc. CMRT Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci., 229-234.

Becker, S., 2009. Generation and application of rules for quality dependent façade reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing 64 (6), 640 - 653.

Beinat, A., Crosilla, F., Visintini, D., Sepic, F., 2007. Automatic non parametric procedures for terrestrial laser point clouds processing. Proceedings of: PIA07, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (3/W49B).

Bennamoun, M. and Mamic, G.J., 2002. Object Recognition: Fundamentals and Case Studies. Springer-Verlag, New York.

Benner, J., Geiger, A., Leinemann, K., 2005. Flexible generation of semantic 3D building models. In: Groeger, G. et al. (Eds.): Proceedings of the 1st International Workshop on Next Generation 3D City Models, Bonn.

Besl, P.J., McKay, N.D., 1992. A method for registration of 3-d shapes. IEEE Tans. Pattern Anal. Mach. Intell., 14: 239-256.

Biosca, J. M. and Lerma, J. L., 2008. Unsupervised robust planar segmentation of terrestrial laser scanner point clouds based on fuzzy clustering methods. ISPRS Journal of Photogrammetry and Remote Sensing, 63(1): 84–98.

Biswas, J. and Veloso, M., 2011. Depth camera based localization and navigation for indoor mobile robots. Robotics and Automation (ICRA), 2012 IEEE, International Conference, pp. 1697-1702.

Boehm, J., 2008. Facade Detail from Incomplete Range Data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVII (Part B5), 653-658.

Böhringer, C., Rutherford, T. F., Tol, R. S., 2009. The EU 20/20/2020 targets: An overview of the EMF22 assessment. Energy Economics, 31, S268-S273.

Boykov, Y. and Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 26(9), 1124-1137.

Boulaassal, H., Landes, T., Grussenmeyer, P. and Tarsha-Kurdi, F., 2007. Automatic segmentation of building facades using terrestrial laser data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3/W52): 65–70.

Boulaassal, H., Landes, T., Grussenmeyer, P., 2009. Automatic extraction of planar clusters and their contours on building façades recorded by terrestrial laser scanner. International Journal of Architectural Computing, 7(1), 1-20.

Bretar, F. and Roux, M., 2005. Hybrid image segmentation using LiDAR 3D planar primitives. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI, part 3/W19, Enschede, the Netherlands, pp. 232.

Brenner, C., 2005. Building reconstruction from laser scanning images. In: International Journal of Applied Earth Observation and Geoinformation Volume 6, Issues 3-4, March 2005, Pages 187-198 Data Quality in Earth Observation Techniques

Briese, C., 2004. Three-dimensional modelling of breaklines from airborne laser scanner data. In International Archives of Photogrammetry and Remote Sensing, Vol.XXXV, B3, Istanbul, Turky.

Brilakis, I., Lourakis, M., Sacks, R., Savarese, S., Christodoulou, S., Teizer, J., Makhmalbaf, A., 2010. Toward automated generation of parametric BIMs based on hybrid video and laser scanning data, Advanced Engineering Informatics 24 (4) (Nov.2010) 456–465.

Brown, D., 1971. Close-range camera calibration. Photogrammetric Engineering and Remote Sensing, 37(8), pp 855-866.

Brügelmann, R., 2000. Automatic breakline detection from airborne laser range data. In International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, B3, Amsterdam, Netherlands, pp. 109–115.

Brumana, R., Georgopoulos, A., Oreni, D., Raimondi, A., Bregianni, A., 2013. HBIM for Documentation, Dissemination and Management of Built Heritage. The Case Study of St. Maria in Scaria d'Intelvi. International Journal of Heritage in the Digital Era, 2(3), 433-452.

Budroni, A. and Böhm, J., 2009. Toward automatic reconstruction of interiors from laser data. Proceedings of Virtual Reconstruction and Visualization of Complex Architectures (3D-Arch).

Cantzler, H., 2003. Improving Architectural 3D Reconstruction by Constrained Modelling, College of Science and Engineering. School of Informatics, University of Edinburgh, PhD Edinburgh.

CEC, 2002. Commission of the European Communities, Green Paper: Towards a European Strategy for the Security of Energy Supply, Brussels, November 2000.

CEC, 2007. Communication from the Commission to the European Council and the European Parliament: An Energy Policy for Europe. COM 2007, 1 Final, Brussels.

CEC, 2008. Proposal for a decision on the European Parliament and of the Council on the effort of member states to reduce their greenhouse gas emissions to meet the community's greenhouse gas emission reduction commitments up to 2020 COM(2008) 17 final. Commission of the Euopean Communities, Brussels (2008).

Chauve, A. L., Labatut, P., Pons, J. P., 2010. Robust piecewise-planar 3D reconstruction and completion from large-scale unstructured point data. In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on (pp. 1261-1268). IEEE.

Comaniciu, D. and Meer, P., 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Machine Intell., 24:603–619.

Crosilla, F., Visintini, D., Sepic, F., 2009. Reliable automatic classification and segmentation of laser point clouds by statistical analysis of surface curvature values. Applied Geomatics, 1(1-2), 17-30.

Dijkman, S., van den Heuvel, F.A., 2000. Semi-automatic registration of laser scanner data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 33 (Part B 5/1), 215–221 (September).

Debevec, P. E., Taylor, C. J., Malik, J., 1996. Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach. In Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (pp. 11-20). ACM.

Dold, C. and Brenner, C., 2006. Registration of terrestrial laser scanning data using planar patches and image data. IAPRS, XXXVI, 5, 78-83.

Dorninger, P. and Nothegger, C., 2007. 3D segmentation of unstructured point clouds for building modelling. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(3/W49A), 191-196.

Douglas, D. H. and Peucker, T. K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: The International Journal for Geographic Information and Geovisualization, 10(2), 112-122.

Edelsbrunner, H., O'Rourke, J., Seidel, R., 1986. Constructing arrangements of lines and hyperplanes with applications. SIAM Journal on Computing, 15(2), 341-363.

Elberink, S.J. and Vosselman, M.G., 2006. Adding the Third Dimension to a Topographic Database Using Airborne Laser Scanner Data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI, part 3, Bonn, Germany, pp. 92-97.

El-Hakim, S. F., Boulanger, P., Blais, F., Beraldin, J. A., 1997, July. System for indoor 3D mapping and virtual environments. In Optical Science, Engineering and Instrumentation'97 (pp. 21-35). International Society for Optics and Photonics.

Émery, M. and Meyer, P. A., 1989. Stochastic calculus in manifolds. Berlin.

EPA, 2002. Directive 2002/91/CE of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings, 2002.

Filin, S., 2002. Surface clustering from airborne laser scanning data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIV, part 3A/B, Graz, Austria, pp. 119-124.

Filin, S. and Pfeifer, N., 2006. Segmentation of airborne laser scanning data using a slope adaptive neighborhood. ISPRS Journal of Photogrammetry and Remote Sensing, 60(2): 71–80.

Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6): 381–395.

Forlani, G., Nardinocchi, C., Scaioni, M., Zingaretti, P., 2006. Complete classification of raw LIDAR data and 3D reconstruction of buildings. Pattern Analysis and Applications, 8(4), 357-374.

Förstner, W., 1998. Image processing for feature extraction in digital intensity, color and range images. In Proceedings of the International Summer School on Data Analysis and Statistical Foundations of Geomatics, Greece, Springer Lecture Notes on Earth Sciences.

Förstner, W., 2010. Optimal vanishing point detection and rotation estimation of single images from a legoland scene. In Proc. ISPRS Commission III Symp. Photogramm. Comput. Vis. Image Anal, pp. 157-162.

Fraser, C. S., 1984. Network design considerations for non-topographic photogrammetry. Photogrammetric Engineering and Remote Sensing, 50(8), 1115-1126.

Friedman, S., and Stamos, I., 2011. Real Time Detection of Repeated Structures in Point Clouds of Urban Scenes. In 3DIMPVT (pp. 220-227).

Frueh, C., Jain, S., and Zakhor, A., 2005. Data processing algorithms for generating textured 3D building facade meshes from laser scans and camera images. International Journal of Computer Vision, 61(2), 159-184.

Furukawa, Y., Curless, B., Seitz, S. M., & Szeliski, R. 2009. Reconstructing building interiors from images. In Computer Vision, 2009 IEEE 12th International Conference on (pp. 80-87). IEEE.

Geibel, R. and Stilla, U., 2000. Segmentation of laser-altimeter data for building reconstruction: Comparison of different procedures. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIII, part B3, Amsterdam, The Netherlands, pp. 326-334

Giddens, A., 2009. The politics of climate change. Cambridge, UK.

Gianinetto, M., Roncoroni, F. and Scaioni, M, 2005. Calibration of Close-Range Thermal Imagery for Integration into 3D VR Models, In: Proc. of Work. Italy-Canada 2005 "3D Digital Imaging and Modeling: Applications of Heritage, Industry, Medicine and Land", Padova (Italy), 17-18 Maggio.

González-Jorge, H., Lagüela, S., Krelling, P., Armesto J., Martínez-Sánchez, J.P., 2012. Single image rectification of thermal images for geometric studies in façade inspections. Infrared Phys Technol 55(5):421-6.

Green, P. J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika, 82(4), 711-732.
003yopp
x
Gröger, G. and Plümer, L., 2012. CityGML–Interoperable semantic 3D city models, ISPRS Journal of Photogrammetry and Remote Sensing, 71, 12-33.

Grubb, M., Vrolijk, C., Brack, D. Forsyth, T., 1999. The Kyoto Protocol: a guide and assessment (pp. 64-65). London: Royal Institute of International Affairs.

Godin, G., Beraldin, J. A., Rioux, M., Levoy, M., Cournoyer, L., 2001. An assessment of laser range measurement of marble surfaces. In Videometrics and Optical Methods for 3D Shape Measurement, 22-23 January 2001, 174-184.

Gorte, B.G.H., 2002. Segmentation of tin-structured surface models, Joint Conference on Geo-spatial theory, Processing and Applications, Ottawa, Canada, 8 - 12 July 2002.

Guehring, J., 2001. Reliable 3d surface acquisition, registration and validation using statistical error models. Proc.3rd International Conference on 3-D Digital Imaging and Modeling, Quebec, May 28–June 1, pp. 224–231.

Haala, N. and Brenner, C., 1999. Extraction of buildings and trees in urban environments. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 130-137.

Hähnel, D., Burgard, W., Thrun, S., 2003. Learning compact 3D models of indoor and outdoor environments with a mobile robot. Robotics and Autonomous Systems, 44(1), 15-27.

Hartley, R.I. and Zisserman A., 2004. Multiple View Geometry in Computer Vision. Second edition. Cambridge University Press, Cambridge, 672 pages.

Hays, J., Leordeanu, M., Efros, A. A., Liu, Y., 2006. Discovering texture regularity as a higher-order correspondence problem. In Computer Vision–ECCV 2006 (pp. 522-535). Springer Berlin Heidelberg.

Hallberg, D. and Tarandi, V., 2009. On the use of 4D BIM in LMS for Construction Works, Journal of Information Technology in Construction (Itcon).

Hofmann, A. D., Maas, H.-G. and Streilein, A., 2003. Derivation of roof types by cluster analysis in parameter spaces of airborne laser scanner point clouds. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 34(3/W13): 112–117.

Hoover, A., Jean-Baptiste, G., Jiang, X., Flynn, P. J., Bunke, H., Goldgof, D. B., Bowyer, K., Eggert, D. W., Fitzgibbon, A., Fisher, R. B. 1996. An experimental comparison of range image segmentation algorithms. IEEE Trans. Pattern Anal. Mach. Intell. 18, 7, 673– 689.

Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W., 1992. Surface reconstruction from unorganized points (Vol. 26, No. 2, pp. 71-78). ACM.

Hough, P. V. C., 1962. Method and means for recognizing complex patterns. U.S. Patent 3,069,654.

Jahangiri M. and Petrou M., 2009, An attention model for extracting components that merit identification. In ICIP'09, pages 961-964, Piscataway, NJ, USA.

Jenke, P., Krückeberg, B., Straßer, W., 2008. Surface Reconstruction from Fitted Shape Primitives. In VMV (pp. 31-40).

Kang, Z., Zhang, L., Yue, H., Lindenbergh, R., 2013a. Range Image Techniques for Fast Detection and Quantification of Changes in Repeatedly Scanned Buildings. Photogrammetric engineering and remote sensing, 79(8), 695-707.

Kang, Z.; Zhang, L.; Wang, B.; Li, Z.; Jia, F., 2013b. An Optimized BaySAC Algorithm for Efficient Fitting of Primitives in Point Clouds, Geoscience and Remote Sensing Letters, IEEE, vol.PP, no.99, pp.1,5.

Kolbe, T. H., Gröger, G., Plümer, L., 2005. CityGML: Interoperable access to 3D city models. In Geo-information for disaster management (pp. 883-899). Springer Berlin Heidelberg.

Koppula, H. S., Anand, A., Joachims, T., Saxena, A., 2011. Semantic labeling of 3d point clouds for indoor scenes. In Advances in Neural Information Processing Systems (pp. 244-252).

Korah, T. and Rasmussen, C., 2008. Analysis of building textures for reconstructing partially occluded facades. In Computer Vision–ECCV 2008 (pp. 359-372). Springer Berlin Heidelberg.

Koutsourakis, P., Simon, L., Teboul, O., Tziritas, G., Paragios, N., 2009. Single view reconstruction using shape grammars for urban environments. In Computer Vision, 2009 IEEE 12th International Conference on (pp. 1795-1802). IEEE.

Isikdag, U. And Zlatanova, S., 2009. Towards defining a framework for automatic generation of buildings in CityGML using BIM. In: Lee, J., Zlatanova, S. (Eds.): 3D Geo-information Sciences, LNG&C, Springer-Verlag, pp. 79–96.

Lee S.C. and Nevatia R., 2004, Extraction and integration of window in a 3D building model from ground view images. In CVPR, volume 2, pages 113-120.

Li, M., Langbein, F. C. and Martin, R. R. 2010. Detecting design intent in approximate CAD models using symmetry. Computer-Aided Design 42(3), pp. 183-201.

Lowe, D. G., 1999. Object recognition from local scale-invariant features. In Computer vision, 1999. The proceedings of the seventh IEEE international conference on (Vol. 2, pp. 1150-1157). Ieee.

Luhmann, T., Robson, S., Kyle, S., Harley, I., 2006.Close range photogrammetry: Principles, methods and applications. (pp. 1-510). Whittles.

Luhmann, T., Piechel, J., Roelfs, T., 2013. Geometric calibration of thermographic cameras. In Thermal Infrared Remote Sensing (pp. 27-42). Springer Netherlands.

Luo, C and Sohn, G., 2010. A knowledge based hierarchical classification tree for 3d facade modeling using terrestrial laser scanning data. In the 2010 Canadian Geomatics Conference and Symposium of Commission I, ISPRS, Volume XXXVIII, Calgary,Canada, p.195-20 (2010).

Maas, H.-G. and Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2-3): 153-163.

Maldague, X., 2001. Nondestructive Testing Handbook: Infrared and Thermal Testing 3rd ed., ASNT: Columbus, OH, USA.

Martínez, J., Soria-Medina, A., Arias, P., Buffara-Antunes, A. F., 2012. Automatic processing of Terrestrial Laser Scanning data of building façades. Automation in Construction, 22, 298-305.

Martín-Ocaña, S., Cañas- Guerrero, I.,. González Requena, I., 2004. Thermographic survey of two rural buildings in Spain. Energy Build 36(6), 515-23.

McHenry, K. and Bajcsy, P., 2008. An overview of 3d data content, file formats and viewers. National Center for Supercomputing Applications, 1205.

Mesolongitis, A., and Stamos, I., 2012. Detection of windows in point clouds of urban scenes. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on (pp. 17-24). IEEE.

Mitra, N. J., Guibas, L. J., Pauly, M., 2006. Partial and approximate symmetry detection for 3D geometry. ACM Transactions on Graphics (TOG), 25(3), 560-568.

Müller P., Wonka P., Haegler S., Ulmer A. and Van Gool L, 2006, Procedural modeling of buildings. ACM Transactions on Graphics 25, 3, 614-623.

Musialski, P., Wonka, P., Recheis, M., Maierhofer, S., and Purgathofer, W., 2009. Symmetry-Based Façade Repair. In VMV (pp. 3-10).

Musialski P, Recheis M., Maierhofer S., Wonka P. and Purgathofer W, 2010. Tiling of ortho-rectified façade images. In Spring Conference on Computer Graphics.

Musialski P, Wonka P., Aliaga D.G., M. Wimmer, van Gool L. and Purgathofer W., 2012. A survey of urban reconstruction. In EUROGRAPHICS 2012 State of the Art Reports.

Nan L., Sharf A., Zhang H., Cohen- OR D. and Chen B., 2010. SmartBoxes for interactive urban reconstruction. ACM Transactions on Graphics 29, 4 (July 2010), 1.17,19.

Nagel, C., Stadler, A., Kolbe, T., 2009. Conceptual Requirements for the Automatic Reconstruction of Building Information Models from Uninterpreted 3D Models, Academic Track of Geoweb 2009 Conference, Vancouver.

NIBS. 2007. United States National Building Information Modeling Standard Version 1 — part 1 Overview, Principles, and Methodologies.

Niedermeier, R. and Sanders, P., 1996. On the Manhattan Distance Between Points on Space Filling Mesh Indexings. Univ., Fak. für Informatik.

Nex, F. and Rinaudo, F., 2009. New integration approach of Photogrammetric and LIDAR techniques for architectural surveys. Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci, 38, 3.

Nyaruhuma, A.P., 2007. Performance analysis of algorithms for detecting roof faces in airborne laser scanner data. MSc Thesis, ITC, Enschede, 80 pp.

Nüchter, A. and Hertzberg, J., 2008. Towards semantic maps for mobile robots. Robotics and Autonomous Systems, 56(11), 915-926.

OGC, 2014. http://www.opengeospatial.org/

Okorn, B., Xiong, X., Akinci, B., Huber, D., 2010. Toward automated modeling of floor plans. In Proceedings of the Symposium on 3D Data Processing, Visualization and Transmission (Vol. 2).

Park M., Brocklehurst K., Collins T.A. and Liu Y., 2011. Translation-symmetry-based perceptual grouping with applications to urban scenes. In Proceedings of the 10th Asian Conference on Computer Vision – Volume Part III, ACCV'10, pages 329-342, Berlin, Heidelberg.

Pauly, M., Mitra, N. J., Giesen, J., Gross, M. H., and Guibas, L. J., 2005. Example-Based 3D Scan Completion. In Symposium on Geometry Processing (pp. 23-32).

Pauly, M., Mitra, N. J., Wallner, J., Pottmann, H., and Guibas, L. J., 2008. Discovering structural regularity in 3D geometry. In ACM Transactions on Graphics (TOG) (Vol. 27, No. 3, p. 43). ACM.

Previtali, M., Barazzetti, L., Scaioni, M., 2012. An automated and accurate procedure for texture mapping from images. Proc. 18th IEEE Intern. Conference on Virtual Systems and MultiMedia (VSMM), G. Guidi / L. Addison (Eds.), pp. 591-594. Milan, Italy

Previtali, M., Barazzetti, L., Brumana, R., Cuca, B., Oreni, D., Roncoroni, F., Scaioni, M., 2013a. Automatic façade segmentation for thermal retrofit. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 3D-ARCH, 25-26.

Previtali, M., Barazzetti, L., Redaelli, V., Scaioni, M., Rosina, E., 2013b. Rigorous procedure for mapping thermal infrared images on three-dimensional models of building façades.Journal of Applied Remote Sensing, 7(1), 073503-073503.

Previtali, M., Scaioni, M., Barazzetti, L., Brumana, R., Roncoroni, F., 2013. Automated detection of repeated structures in building façades. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-5/W2, 2013, ISPRS Workshop Laser Scanning 2013, 11–13 November 2013, Antalya, Turkey, 6 pages.

Previtali, M., Barazzetti, L., Brumana, R., Cuca, B., Oreni, D., Roncoroni, F., Scaioni, M., 2013d. Automatic façade modelling for energy efficient retrofitting. Applied Geomatics (submitted)

Pu, S. and Vosselman, G., 2006. Automatic extraction of building features from terrestrial laser scanning. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(5), 25-27.

Pu, S. and Vosselman, G., 2009. Knowledge based reconstruction of building models from terrestrial laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing, 64(6), 575-584.

Rabbani, T. and Van den Heuvel, F., 2005. Efficient Hough transform for automatic detection of cylinders in point clouds. ISPRS Proceedings. Workshop Laser scanning. Enschede, the Netherlands, September 12-14, 2005.

Rabbani, T., 2006. Automatic reconstruction of industrial installations using point clouds and images. Ph.D. thesis, Delft University of Technology, Delft, the Netherlands. 154 pages.

Rabbani, T., van den Heuvel, F.A. and Vosselman, M.G., 2006. Segmentation of point clouds using smoothness constraints. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI, part 5, Dresden, Germany, pp. 248-253.

Rabbani, T., Dijkman, S., van den Heuvel, F., Vosselman, G., 2007. An integrated approach for modelling and global registration of point clouds. ISPRS journal of Photogrammetry and Remote Sensing, 61(6), 355-370.

Ribarić, S., Marčetić, D., Vedrina, D.S., 2009. A knowledge-based system for the non-destructive diagnostics of façade isolation using the information fusion of visual and IR images. Expert Syst Appl 36(2), 3812-23.

Ripperda, N., 2008. Grammar Based Facade Reconstruction using RjMCMC. PhotogrammetrieFernerkundungGeoinformation 2008 (2), 83.

Ripperda N. and Brenner C., 2009. Application of a Formal Grammar to Façade Reconstruction in Semiautomatic and Automatic Environments. In 12th AGILE International Conference on Geographic Information Science.

Roggero, M., 2002. Object Segmentation with Region Growing and Principal Component Analysis. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIV, part 3A, Graz, Austria, pp. 289-294.

Rusinkiewicz, S. and Levoy, M., 2001. Efficient variants of the ICP algorithm. Proc. of the 3rd International Conference on 3D Digital Imaging and Modeling, Quebec, May 28–June 1, pp. 145–152.

Russell, S. and Norvig, P., 2003. Artificial Intelligence: A Modern Approach. Pearson Education.

Rusu, R. B., Marton, Z. C., Blodow, N., Dolha, M., Beetz, M., 2008. Towards 3D point cloud based object maps for household environments. Robotics and Autonomous Systems, 56(11), 927-941.

Samet, H., 2006. Foundations of multidimensional and metric data structures. Morgan Kaufmann.

Sampath, A. and Shan, J., 2007. Building Boundary Tracing and Regularization from Airborne Lidar Point Clouds. In: Photogrammetric Engineering & Remote Sensing, Vol. 73, No. 7, July 2007, pp. 805–812. DOI 0099-1112/07/7307

Sapkota, P.P., 2008. Segmentation of coloured point cloud data. M.Sc. thesis, ITC, Enschede, the Netherlands. 67 pages.

Scaioni, M. and Forlani, G., 2003. Independent model triangulation of terrestrial laser scanner data. International archives of photogrammetry remote sensing and spatial information sciences, 34(5/W12), 308-313.

Schaffalitzky F and Zissermann A., 1999, Geometric grouping of repeated elements within images. In Shape, Contour and Grouping in Computer Vision, 165-181.

Schnabel, R., Wahl, R. and Klein, R., 2007. Efficient RANSAC for Point-Cloud Shape Detection. Computer Graphics Forum, 26(2): 214-226.

Shen C.H., Huang S.S., Fu H. and Hu S.M., 2011, Adaptive partitioning of urban facades. SIGGRAPH ASIA, 30(6):184:1– 184:9, 2011.

Shikhane D., Bhakar S. and Mudur S.P., 2001, Compression of large 3D engineering models using automatic discovery of repeating geometric features. In Vision, Modeling, and Visualization, 244-240.

Shilane, P., Min, P., Kazhdan, M., Funkhouser, T., 2004. The princeton shape benchmark. In Shape Modeling Applications, 2004. Proceedings (pp. 167-178). IEEE.

Sinha, S. N., Steedly, D., Szeliski, R., Agrawala, M., and Pollefeys, M., 2008. Interactive 3D architectural modeling from unordered photo collections. In ACM Transactions on Graphics (TOG) (Vol. 27, No. 5, p. 159). ACM.

Soudarissanane, S., Lindenbergh, R., Menenti, M., Teunissen, P., 2011. Scanning geometry: Influencing factor on the quality of terrestrial laser scanning points. ISPRS Journal of Photogrammetry and Remote Sensing, 66(4), 389-399.

Stamos, I. and Allen, P. E., 2000. 3-D model construction using range and image data. In Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on (Vol. 1, pp. 531-536). IEEE.

Stamos I. and Allen P.K, 2002. Geometry and texture recovery of scenes of large scale. Journal of Computer Vision and Image Understanding, 88(2):94–118, 2002.

Stewart, C.V., 1997. Bias in robust estimation caused by discontinuities and multiple structures. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19:818-833

Suveg, I. and Vosselman G., 2004. Reconstruction of 3D building models from aerial images and maps. ISPRS Journal of Photogrammetry and remote sensing 58 (3-4), 202{224.

Tang, P., Huber, D., Akinci, B., Lipman, R., Lytle, A., 2010. Automatic reconstruction of as-built building information models from laser-scanned point clouds: a review of related techniques, Automation in Construction 19 (7) (2010) 829–843.

Tarsha-Kurdi, F., Landes, T. and Grussenmeyer, P., 2007. Hough-Transform and Extended RANSAC Algorithms for Automatic Detection of 3D Building Roof Planes from Lidar Data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI, part 3/W52, Espo, Finland, pp. 407-412.

Teboul O., Kokkinos I., Simon L., Koutsourakis P. And Paragios N., 2011, Shape grammar parsing via reinforced learning. In CVPR, 0:2273-2280.

Tian, Y., Zhu, Q., Gerke, M., Vosselman, G., 2009. Knowledge-based topological reconstruction for building facade surface patches . International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (5).

Thrun, S., Martin, C., Liu, Y., Hahnel, D., Emery-Montemerlo, R., Chakrabarti, D., Burgard, W., 2004. A real-time expectation-maximization algorithm for acquiring multiplanar maps of indoor environments with mobile robots. Robotics and Automation, IEEE Transactions on, 20(3), 433-443

Tóvári, D., 2006. Segmentation Based Classification of Airborne Laser Scanner Data. PhD Thesis, Universität Karlsruhe, Karlsruhe, Germany.

Tóvári, D. and Pfeifer, N., 2005. Segmentation based robust interpolation-a new approach to laser data filtering. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3/W19): 79–84.

Triebel R., Kersting K. and Burgard W., 2006. Robust 3D scan point classification using associative markov networks. In IEEE ICRA, pages 2603 –2608.

Triggs, B., McLauchlan, P. F., Hartley, R. I., Fitzgibbon, A. W., 2000. Bundle adjustment—a modern synthesis. In Vision algorithms: theory and practice (pp. 298-372). Springer Berlin Heidelberg.

USDE, 2001. Energy Information, Administration, Residential Buildings Energy consumption Survey (RECS), U.S. Department of Energy, 2001.

USDE, 2003. Energy Information, Administration, Commercial Buildings Energy consumption Survey (CBECS), U.S. Department of Energy, 2003.

Van den Heuvel, F.A., 1999. A line-photogrammetric mathematical model for the reconstruction of polyhedral objects. In: El-Hakim, S.F. (Ed.), Videometrics VI, Proceedings of International Society for Optical Engineering, vol. 3641, pp. 60–71.

Van Goor, B., Lindenbergh, R., and Soudarissanane, S., 2011. Identifying corresponding segments from repeated scan data. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XXXVIII-5/W12, 295-300, doi:10.5194/isprsarchives-XXXVIII-5-W12-295-2011, 2011.

Vanegas, C. A., Aliaga, D. G., Benes, B., 2012. Automatic extraction of manhattan-world building masses from 3d laser range scans. Visualization and Computer Graphics, IEEE Transactions on, 18(10), 1627-1637.

Vosselman, G., 2004. Fusion of Laser Scanning Dara and Aerial Photographs for Building Reconstruction. In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV, Part 3, Istanbul 2004.

Vosselman, G., Gorte, B. G. H., Sithole, G. and Rabbani, T., 2004. Recognising structure in laser scanner point clouds. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(8/W2): 33–38.

Vosselman, G. and Maas, H. G. (Eds.), 2010. Airborne and terrestrial laser scanning (Vol. 318). Dunbeath, UK: Whittles.

Wang, W., Pottmann, H., Liu, Y., 2006. Fitting B-spline curves to point clouds by curvature-based squared distance minimization. ACM Transactions on Graphics (ToG), 25(2), 214-238.

Wang, Z. and Brenner, C., 2008. Point based registration of terrestrial laser data using intensity and geometry features. In ISPRS Congress ('08), Beijing, China.

Wu C., Frahm J. M. and Pollefeys, 2010, Detecting large repetitive structures with salient boundaries. In ECCV, pages 145-155.

Wu Shin and Gonzales Marquez, 2003. A non-self-intersection Douglas-Peucker Algorithm. Proceedings of the XVI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'03) 1530-1834/03, 2003 IEEE

Xiao J., Fang T. Tan P., Zhao P, Ofek E. and Quan L., 2009. Image-based street-side city modeling. ACM Transactions on Graphics 28, 5, 114:1- 114:12.
Xiong, X., Adan, A., Akinci, B., Huber, D., 2013. Automatic creation of semantically rich 3D building models from laser scanner data. Automation in Construction, 31, 325-337.

Zhao, Z. and Saalfeld, A., 1997. Linear-Time Sleeve-Fitting Polyline Simplification Algorithms. In: Autocarto 13, ACSM/ASPRS'97 Technical Papers, Seattle, Washington, April 1997, vol. 5, pp. 214–223.

Zhang, Z., 1994. Iterative point matching for registration of free-form curves and surfaces. International Journal of Computer Vision 13(2), 119–152.

Zhang, Z., 2012. Microsoft kinect sensor and its effect. Multimedia, IEEE, 19(2), 4-10.

Zheng Q., Sharf A., Wan G., Li Y., Mitra N. J., Cohen-OR D. and Chen B., 2010. Non-local scan consolidation for 3D urban scenes. ACM Transactions on Graphics 29, 4 (July 2010), 1. 11.

Acknowledgements

The period of the PhD study has been a unique experience in my life. I would never have been able to finish my dissertation without the guidance of my supervisors, help from friends, and support from my family. Here, I would like to acknowledge their help.

First of all, I express my gratitude to Prof. Marco Scaioni for giving me the chance to pursue a PhD and for his constant guidance during the PhD study. I thank him for his assistance and supervision of my thesis; for his continual support throughout this work, his scientific advice and knowledge.

I gratefully acknowledge Prof. Raffaella Brumana. All her energetic enthusiasm for my research has never ceased to be a great source of motivation for me. Without her inspiration this work would not have been possible.

Then, I wish to thank my tutor Prof. Luigi Mussio for what he had taught me during the PhD. Furthermore, his rigorous and efficient style has deeply influenced me within and beyond research.

A special thanks to Prof. Alberto Giussani for giving me the chance to participate to some monitoring campaigns of buildings. The possibility to accomplish this work allowed me to investigate a different research field.

I would like to express my gratitude to my colleagues in the IC&T Laboratory for the professional atmosphere they have created. I would like to mention some people in particular. First of all, I would like to express my gratitude to Dr. Luigi Barazzetti for his constructive advices and inspirations in many works and projects as well as his scrupulous contribution to this thesis. In addition, I wish to thank Dr. Fabio Roncoroni for his long-term assistance and patience.

I thank all my friends for all the happy moments we have spent together. Their support and encouragement throughout were crucial. Special tanks to: Sara Galbussera, Pamela Buratti, Luca Brembilla, Mattia Montesano and Federico Benaglia. I hope to keep the party with my friends for the rest of my life.

Last but not least, I wish to thank my parents, whose love and support were always there when I needed.