

Extracting regularities from point clouds of indoor scenes

YUNMENG ZHU

February, 2017

SUPERVISORS:

Dr.ing, M.S, Peter

Dr.ir, S.J, Oude Elberink

Advisor: S., Nikoohemat MSc



Extracting regularities from point clouds of indoor scenes

YUNMENG ZHU

Enschede, The Netherlands, February, 2017

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: [Geoinformatics]

SUPERVISORS:

Dr.ing, M.S, Peter

Dr.ir, S.J, Oude Elberink

Advisor: S., Nikoohemat MSc

THESIS ASSESSMENT BOARD:

Prof.dr.ir, M.G, Vosselman (Chair)

Dr.ing, S., Zlatanova (External Examiner, Delft University of Technology)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

Abstract

Nowadays, an increasing number of people are interested in the build-up indoor environment as it is an important place to optimize our quality of life in social, entertainment and economic activities. Furthermore, Indoor Mobile Laser Scanning (IMLS) makes indoor data acquisition and 3D model generation come true. However, the data collected by IMLS is a large set of dense point clouds, i.e. a list of point coordinates that lacks structural information in its raw stage. The data may have data gaps, or outliers which hinder the way of understanding structure and reconstruction of buildings.

One way to solve this problem is to extract regularities in the man-made indoor environment. This can help to understand and reconstruct a fundamental frame of the indoor scenes. This research was designed to detect two main characteristics that exist in an indoor environment. One is symmetry and the other is repetitive patterns. The polygon symmetry detection has 4 main steps.

1. Segmentation (surface growing) for raw point clouds;
2. Extract segments' boundary polygons and deal with intersection;
3. Symmetry centroid selection and symmetry line generation;
4. Over flip polygons to detect symmetry and identify the symmetric parts.

This process has been tested on both 2D floor plan and 3D laser point clouds. The surface growing segmentation method was applied to extract the planar surfaces from unstructured laser point cloud data. The boundary was extracted by 2D alpha shape and the boundary points' x, y coordinates were kept. To detect symmetry, a draft symmetry centroid was selected first and then refined by shifting around neighbouring regions. The remaining symmetric polygons are visualized in different colours for better recognition.

Repetitive structures in the indoor environment has identified in this research include: room width and length, wall thickness, room concave corners, etc. Those repetitive structures can be combined and applied to detect the similar rooms. Two approaches introduced to identify the similar rooms.

1. Identification of the similar rooms by weight. Assigning a stochastic weight for each feature and select best results depends on different weights combination. It produces several good results.
2. Hierarchical identification of the similar rooms. Rooms are grouped by the hierarchical features. The accuracy of this method can reach above 95% in some cases.

These regularities could probably be applied in data compression, registration of point clouds, or deriving a CAD or BIM model.

Keywords:

Indoor Mobile Laser Scanning (IMLS), regularity, symmetry, repetitive patterns, similar rooms, point clouds, 3D model, floor plans

Acknowledgement

I would like to take this opportunity to thank all the people who contributed to the work described in the thesis.

First I offer my sincerest gratitude to my first supervisor, Dr Michael, for giving me his very valuable knowledge and experience, engaging me new ideas, and supporting my thesis writing during the half year. I would like to thank my second supervisor Sander for his interest in my work, and his patience to explain complex ideas into simple terms. The result described in this thesis was accomplished with the help and support of my advisor, Shayan. I greatly benefited from his keen scientific insight and critical thinking.

I am grateful for University of Stanford who provides the high-quality data resources. I cannot complete my thesis without the datasets.

Thanks to all the ITC staff and fellow students, especially Wan and Simon for build up the running team. The study experience and many activities in ITC are unforgettable.

It is with immense gratitude that I acknowledge my friends Qiuqiu, Yeye and Junjun who brought the company and endless happiness. Thanks to Kaaviya, Arman and Raga for supporting and helping in both life and study. Thanks to my friends Tingzi, Yanyan, Jianghua, Yizhen.

To my beloved family, I would like to express my greatest appreciation to my father Xianwen, mother Chengxi, and cute brother Qichao.

TABLE OF CONTENTS

List of figures	6
List of tables	7
1. Introduction	8
1.1. Background and motivation	8
1.2. Research objective and question.....	9
1.3. Research approach and innovation	9
1.4. Thesis structure.....	10
2. Literature Review.....	11
2.1. Overview.....	11
2.2. Indoor laser scanning.....	11
2.3. Segmentation.....	13
2.4. Symmetry property.....	13
2.5. Repetitive patterns.....	14
2.6. Summary	14
3. Symmetry detection.....	15
3.1. Description of symmetry	15
3.2. Framework.....	16
3.3. Symmetry detection.....	17
4. Repetitive patterns.....	30
4.1. Repetitiv patterns description.....	30
4.2. Edges length	31
4.3. Wall thickness.....	32
4.4. Floor center and polygon centroid.....	33
4.5. Distance from center	34
4.6. Polygon area, Bounding box area, and concave corner area	34
4.7. Arbitrary orientation of point clouds.....	35
4.8. Bbox Length and Width.....	36
4.9. Identification of similar polygons.....	36
5. Result and Discussion.....	40
5.1. Study area and datasets	40
5.2. Implementation.....	41
5.3. Symmetry result and discussion	41
5.4. Repetitive patterns result and discussion	48
6. Conclusion and Recommendations.....	50
6.1. Conclusions	50
6.2. Answer to research questions.....	50
6.3. Recommendation	51
List of references	52
Appendix	55

LIST OF FIGURES

Figure 1-1 Regular patterns.....	8	
Figure 2-1 Indoor laser scanner	12	
Figure 3-1 Axial symmetry.....	15	
Figure 3-2 Framework of symmetry detection	17	
Figure 3-3 Difference of floor boundary.....	18	
Figure 3-4 Boundary symmetry detection results with different angles	20	
Figure 3-5 Intersection area and difference area for the line of symmetry with angles from -89 to 90 degrees	20	
Figure 3-6 Buffer distance (1 meter), buffer region (light green circle region), draft centroid (blue triangle), 29 candidate centroids (black), one optimal centroid (red point)	21	
Figure 3-7 Overview of polygons per floor (left) and close look of boundaries intersect (right)	21	
Figure 3-8 Digitized 2D floor map.....	23	
Figure 3-9 Inner building symmetry.....	24	
Figure 3-10 Identification of room polygon symmetry by area	25	
Figure 3-11 Identification of segment polygon symmetry by area.....	25	
Figure 3-12 Identification of room polygon symmetry by number	26	
Figure 3-13 Identification of segment polygon symmetry by number	27	
Figure 3-14 Segmentation and projection of 3D point clouds into 2D plane.....	28	
Figure 3-15 Identification of edge symmetry.....	29	
Figure 4-1 Framework of repetitive pattern detection	31	
Figure 4-2 Edges length	32	
Figure 4-3 Edges length	32	
Figure 4-4 Repetitive distance patterns.....	33	
Figure 4-5 Floor center and polygon centroid, center centroid connection, and distance from center	34	
Figure 4-6 Polygon area (white bar), bounding box area is a bit bigger (blue) than polygon area.....	34	
Figure 4-7 Minimum area bounding box overview of floor, detail polygon, and orientation angles	35	
Figure 4-8 Round value of width and length of bounding box	36	
Figure 4-9 Ground truth data.....	37	
Figure 4-10 Good result of identification of similar polygons by weight in table 4-2	38	
Figure 4-11 Hierarchical identification of similar polygons	39	
Figure 5-1 Original 2D floor map and its vector data.....	40	
Figure 5-2 3D laser point clouds datasets	41	
Figure 5-3 Building symmetry detection.....	42	
Figure 5-4 Interior symmetry with the line of symmetry in 90 degrees	42	
Figure 5-5 Interior symmetry with the line of symmetry in 0 degrees.....	42	
Figure 5-6 Identification of room polygon symmetry for dataset 5.....	43	
Figure 5-7 Identification of segment polygon symmetry for dataset 5.....	43	
Figure 5-8 Identification of room polygon symmetry for dataset 5.....	44	
Figure 5-9 Identification of segment polygon symmetry for dataset 5.....	44	
Figure 5-10 Original edges test datasets	45	
Figure 5-11 Result of edge symmetry detection for dataset 1	46	
Figure 5-12 Result of edge symmetry detection for dataset 4.....	46	
Figure 5-13 Result of edge symmetry detection for dataset 5.....	47	
Figure 5-14 Ground truth of dataset 5	Figure 5-15 Identification of similar room by weight.....	48
Figure 5-16 Hierarchical identification of similar polygons for dataset 4	49	

LIST OF TABLES

Table 3-1 Parameter of room polygon symmetry detection by area.....	25
Table 3-2 Parameter of segment polygon symmetry detection by area.....	25
Table 3-3 Parameter of room polygon symmetry detection by number.....	26
Table 3-4 Parameter of segment polygon symmetry detection by number.....	27
Table 3-5 Parameter of edge symmetry detection.....	29
Table 4-1 Repetitive patterns.....	30
Table 4-2 4 different weight combination.....	38
Table 5-1 Disjoint space statistics per building area.....	41
Table 5-2 Building element statistics per building area.....	41
Table 5-3 Parameters of room polygon symmetry detection by area for dataset 5.....	43
Table 5-4 Parameters of room polygon symmetry detection by area for dataset 5.....	43
Table 5-5 Parameters of segment polygon symmetry detection by room for dataset 5.....	44
Table 5-6 Parameters of segment polygon symmetry detection by number for dataset 5.....	44
Table 5-7 Parameters of edge symmetry detection for dataset 1.....	46
Table 5-8 Parameters of edge symmetry detection for dataset 4.....	46
Table 5-9 Parameters of edge symmetry detection for dataset 5.....	47
Table 5-10 Symmetry detection parameters.....	47
Table 5-11 Different weight combination.....	48
Table 5-12 Advantage and disadvantage of two method of identification of similar polygons.....	49

1. INTRODUCTION

1.1. Background and motivation

Nowadays, an increasing number of people are interested in the indoor scenes as it is an important place to optimize our quality of life in social, entertainment and economic activities. Furthermore, with the development of computer vision and graphics, photonics, solid-state electronics, laser scanning has become a significant technique for acquiring 3D geo-information. And the possibility of efficient processing for dense point clouds makes 3D indoor modelling reality (Vosselman & Maas, 2010).

However, the data collected by laser scanner is a large set of dense point clouds with a list of point coordinates (Budroni, 2013). There is no structural information and semantic information, moreover, because of the complex indoor furnishing and decoration, the collected data has a high variable point density and anisotropy. Collected point clouds may include data gaps, as some part of them are hidden among large objects or sheltered by clutters. It is difficult to remove clutter objects before data collection. Outliers occur as the scanner scans into the outside of the building through windows or opening doors or corrupted by noise. These factors cannot be avoided and they are occurring often in the indoor scene which will result in data gaps. Further challenging arises after the collection of those point clouds, that is registration errors as laser scanners scanning indoor scenes in different positions. Those factors will make the indoor structure not clear.

To solve those problems, extracting regular patterns in the indoor environment may provide help. Regular patterns are ubiquitous in the indoor construction. Many objects are characterized by the presence of such patterns. Like many repetitive regularities present in nature, for example, patterns of zebras, scales of fish, patterns of leaves (“Principle of Repetition ; Pattern | Visual Communication Design,” 2013). Likewise, many repetitive regularities also abound in manmade indoor scenes, such as repetitive regular fenestration (see Figure 1-1 left). It has same rectangle window size which repetitively presents in the whole facade. All the windows are organized well which is parallel in each store and perpendicular to the horizontal surface (Furukawa, Curless, Seitz, & Szeliski, 2009). From figure 1-1 right, a symmetric room distribution pattern can be seen from left to right side. Reading rooms in the left side and right side of the building have the same size, also for all the WC. Some structural elements (e.g. doors, rooms) have repetitive shapes and sizes. These regular geometries are often simple and repeated many times which constitute the indoor scene frame.

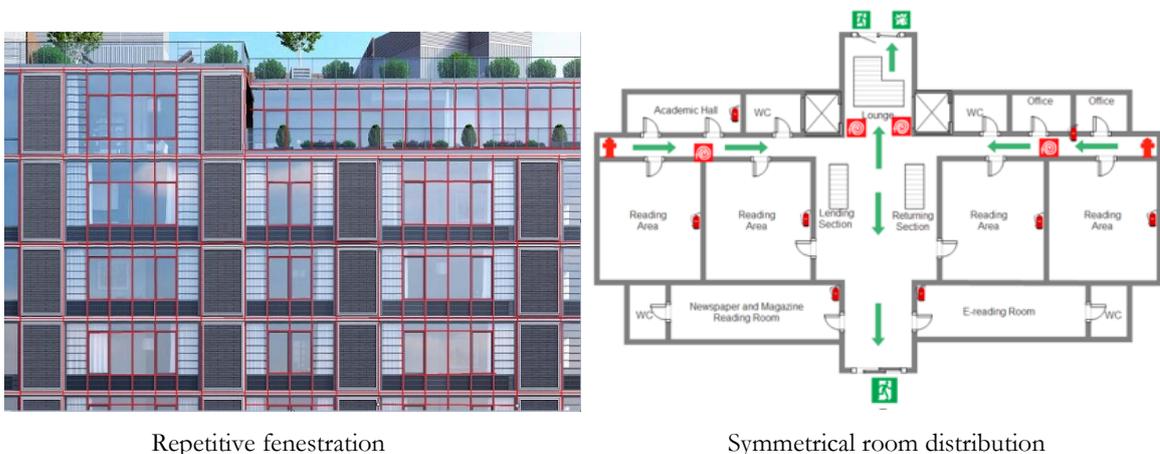


Figure 1-1 Regular patterns

Many regularity detections and descriptions are based on images or outdoor point clouds. Few researchers tried to apply regularities to reconstruct the indoor environments. Kim, Mitra, Yan, and Guibas (2012) use learning and recognition phases to acquire 3D indoor environments with variability and repetition. Mattausch, Panozzo, Mura, Sorkine-Hornung, and Pajarola (2014) segment and classify indoor objects by detecting repeated patterns.

Extracting those repetitive regularity patterns can help to understand and improve the reconstruction of a fundamental frame structure of the indoor scenes. For example, the retrieved regularities used in compression and registration of point clouds, or filling the data gap, or further utilized in deriving a CAD or BIM model. Thus, explore an approach to describe and detect those regularities is very significant.

1.2. Research objective and question

1.2.1. Research objectives

The main objective of this research is description and detection of regularities from the point clouds of the indoor scenes. In order to achieve this goal, some sub-objectives should be addressed which are listed as following:

1. Developing a method to describe indoor environment regularities;
2. Detecting regularities after description;
3. Applying the detected regularities in various user cases;
4. Analysing detected results.

1.2.2. Research questions

The following questions should be addressed to achieve above objectives:

1. How can regularities be described and expressed in a concise and comprehensive way from 3D point clouds of indoor scenes?
2. Because of heavy occlusion and varying point density, how to improve the identification of the potential regularities?
3. How can those regularities be detected?
4. How to evaluate the results?
5. How to apply regularities for improvement of 3D reconstruction?

1.3. Research approach and innovation

Nowadays technology laser scanning, a rapid and flexible method for extracting high-resolution and 3-dimensional data, makes indoor environment modelling grow fast. However, problems such as clutter (e.g. furniture) which cause occlusion in the data and also gaps which make it challenging to reconstruct a 3D model. In order to get overcome these problems, few researchers work on the repetitive pattern and potential regularities existing inside the build-up indoor environment.

This research was conducted in the indoor environment. It concentrates on describing regularities existing in the manmade environment. Finding the similar rooms and detecting whether parts of the floor are symmetric are the two regularities discovered in this research. First, we introduce definition for regularities (symmetry, repetition). Then, methods were developed to detect those regularities from 3D point clouds of indoor scenes with high accuracy and completeness. The approach helps to decide about the identification or potential existence of regularities inside the indoor space. It will be applied to identify whether these regularities are also existing in various indoor environments. Finally, the retrieved regularities could be applied in compression, registration of point clouds, or further utilized in deriving a CAD or BIM model.

The approach of research study depends on several factors, for example, data quality (including point density and precision), down sampling parameter decision, segmentation method, principle component analysis result, filter parameters. Several concepts will be introduced in the literature review.

Many problems are on the way of achieving the objectives. However, the main problem in the research is how can indoor environment regularities be described and expressed from 3D point clouds? To solve the main research problem, this project tested a variety of datasets from both 2D floor plans and 3D laser point clouds.

1.4. Thesis structure

The thesis consists of 6 chapters. This Chapter is an introduction of motivation, problems, research objectives, research approach and innovation. Chapter 2 reviewed several related works. Methodology designed in the thesis will be explained in Chapter 3 and Chapter 4. Chapter 3 discussed the approach of detecting symmetry and Chapter 4 is about the approach of detecting repetitive patterns. Chapter 5 showed the performance of the different methods implemented in multiple datasets and discussing the results. The final Chapter summarized the findings from the experiments, answered the research questions, and recommended the future work.

2. LITERATURE REVIEW

2.1. Overview

Indoor scene reconstruction has become an active research topic. The detection and description of regularities is a significant mechanism in how we recognize and understand the indoor environment. Some researchers have contributed several methods on how to reconstruct part of indoor scene through indoor regularities from 3D point clouds of indoor scenes, for example, Okorn, Xiong, Akinci, and Huber (2010) use a histogram of height data to differentiate floor and ceiling as floors and ceilings are in different heights. Using histogram to differentiate features is helpful which was applied for grouping the similar polygons in the repetitive pattern detection which will be explained in fourth chapter. Oesau, Lafarge, and Alliez (2014) presented a method to reconstruct structures which have Manhattan characteristics. Manhattan world is a way of describing an environment which composed of planar surfaces (Cruz, Kim, & Manduchi, 2014). Each of the surface is oriented along one of three vectors which are canonical mutually orthogonal, for example, the floor, the wall and the ceiling. However, those methods always have the limitation of horizontal and vertical directions. Ochmann, Vock, Wessel, Tamke, and Klein (2014) tried to use visibility estimation to do probabilistic point-to room labelling and intersection of the segmented point to detect doors.

Other researchers would like to reconstruct the whole building structure based on RGBD images generated point clouds (Ikehata, Yang, & Furukawa, 2015). Ochmann, Vock, Wessel, and Klein (2016) present an automatic method for reconstruction of parametric 3D building models from indoor point clouds which enable editing operations and measurements. Becker, Peter, and Fritsch (2015) use Grammar-supported approach to modelling rooms. Xiao and Furukawa (2014) use “inverse constructive solid geometry (CSG)” to reconstruct Metropolitan Museum of Art in New York city. That method is used for a large-scale system. In this research, we detected the symmetry and repetitive patterns per floor of the building.

2.2. Indoor laser scanning

2.2.1. indoor laser scanner

With the development of Building Information Modelling (BIM), the process time of capturing detailed and accurate building data need to be improved. Two main types of indoor laser point clouds collection are listed (Thomson, Apostolopoulos, Backes, & Boehm, 2013).

- The terrestrial laser scanner, see figure 2-1 top left. This is a traditional surveying method. Researchers first mark several tie points manually according to the accuracy requirement and indoor structural complexity. Then they settle the laser scanner on the top of a tripod on several pre-determined stations. The selected tie points are used to provide reference frame information for the data registration purpose. This procedure is supposed to provide high accuracy point clouds. However, it will cost a lot of time spending on tie points mark and tripod settle down. Furthermore, the manual work of selecting tie points, designing the optimal network and settling down the laser scanners need a professional background knowledge.
- Mobile laser scanning (MLS). There are three forms of mobile laser scanning: trolley based and hand held based, and backpack based.
 - ZEB1 (figure 2-1 top right) is a hand-held laser mapping system. Unlike terrestrial laser scanner, it doesn't need prior knowledge of surveying methods (“ZEB1 - 3D Laser Mapping,” 2016).

ZEB1 has the advantages of self-contained, lightweight, independent of GPS and hand-carried by an operator. This makes ZEB1 easy to use in a multiple level environments like stairways or mines.

- The typical trolley based instrument equipped with several laser scanners and cameras. For example, M3 Trolley (Figure 2-1 bottom left), which produced by NavVis company. When M3 Trolley starts to collect data, all the scanners start to work by rotating which covering 270 degrees and 6 cameras offering a fully calibrated 360 degrees' view. The trajectory is guided by simultaneous localization and mapping (SLAM) instead of GPS. Because GPS is not working in an indoor environment. SLAM is a process by which a mobile robot can build a map of an environment and at the same time use this map to deduce its location (Durrant-Whyte & Bailey, 2006). Such system can achieve the accuracy of 5mm at a speed of 5000 m² per hour (“M3 Trolley - NavVis,” 2016).
- The backpack (Figure 2-1 bottom right) is a wearable reality capturing sensor platform used in both indoor and outdoor environment (“Leica Pegasus: Backpack | C.R.Kennedy Survey Division,” 2016). The Leica Pegasus: Backpack consists of five cameras covering 360 degrees and two LIDAR profilers. It synchronises point cloud and imagery data to assure the completeness of the document. The Simultaneous Localisation and Mapping (SLAM) and a high precision IMU ensure the accurate positions.



Terrestrial laser scanner



ZEB1



M3 Trolley



Package: Leica Pegasus

Figure 2-1 Indoor laser scanner

2.3. Segmentation

Segmentation is the footstone on the way of finding regularities in the indoor environment. In the research process, the surface growing segmentation methods will be applied for extracting the primitives for the further process.

2.3.1. Surface growing

To apply the surface growing, select few arbitrary points first. Then select a few neighbouring points to fit a plane. Calculating the square sum of the distance between each point and the fitted plane. If plane fitting results in low residuals, use these points as seed surface. Otherwise, abandon the points. Expanding this planar surface with further neighboring points. The assumption presented in this method is that a part of the dataset where all the points within a certain distance are on the same surface (Vosselman, Gorte, Sithole, & Rabbani, 2004). The criteria of the method are listed as following:

- The determination of seeds in local smooth patches based on:
 - local surface fitting;
 - local smooth normal vector field;
 - detected planes.
- The growing of surfaces based on:
 - proximity (TIN, KNN);
 - surface fit or height continuity;
 - normal vector direction.

2.4. Symmetry property

Symmetry is ubiquitous no matter in the nature environment or in the manmade world. It refers to a sense of harmonious and beautiful proportion and balance (“symmetry,” 2016). Also, symmetrical shape reorganization helps understand and guide the aesthetic and construct of synthetic objects (Simari, Kalogerakis, & Singh, 2006). In this research, geometrical symmetries are discussed a lot.

It is common to see the perfect geometric symmetry patterns in CAD model which are used for editing or compression (Golovinskiy, Podolak, & Funkhouser, 2009). However, most of the structural patterns both in 2D floor plans or 3D point clouds are partial symmetric regions or approximate symmetries, this phenomenon makes the detection of symmetry in computer world more complicated. Even though, many researchers create a lot of methods to find the unperfect symmetric patterns in a variety of fields including computer vision, computational geometry, robotics and visual perception.

Mitra et al. (2006) tried to match simple local shape signatures in pairs first, then use these matches to explore the potential symmetries in the transformation space. Zabrodsky, Peleg, and Avnir (1997) defined that the symmetry distance (SD) of a shape is the minimum mean squared distance by moving points of the original shape in order to get a symmetrical shape. They provided several algorithms for finding different types of symmetries in both 2D and 3D object by calculating the symmetry distance. Mitra, Guibas, and Pauly (2007) presented the symmetrizing deformations which are coupled by spatial domains and transformation configuration spaces to keep the minimal altering of its original shape. Kazhdan, Chazelle, Dobkin, Funkhouser, and Rusinkiewicz (2003) used a reflective symmetry descriptor to record global properties of reflective symmetry in 3D shape. Jiang and Bunke (1991) worked in a way that finding the hypothetical rotational symmetry axes on polyhedron object first, then verifying the assumptions based on a scheme called the generation and the test. The usage of symmetry property is also multi-directions. For example, Golovinskiy et al. (2009) use detection of large-scale symmetries to guide processing of 3D meshes.

In mathematics, symmetry is defined as that an object is variant to any of various transformations, including reflection, rotation or scaling (Petitjean, 2007). This research is going to discuss an axial symmetry as radial symmetry is not very common in the indoor environment.

2.5. Repetitive patterns

Basic elements are arranged in a regular pattern and it is showed in man-made environments with high frequencies. For example, the shape and size of rooms in one building, or the location of windows, and the height of doors. Extracting those repetitive patterns can help in exploiting the occluded information or compressing the dense information. Furthermore, it can be used to complete the building structure (Spinello, Triebel, Vasquez, Arras, & Siegart, 2010).

Many people tried to use different methods to extract the repetitive patterns. Hsu, Li-Chang Liu, and Li (2001) relied on the Fourier transform and autocorrelation to discover the regular texture in the image include intensities, colours and features. A good example is in the Online detection of repeated structures in point clouds of urban scenes for compression and registration (Friedman & Strans, 2013). They detected the dominant geometry like facade and ground planes and regular fenestrations through Fourier analysis. Since the feature periodicity is salient and facades are extremely planar, they use Principle Component Analysis (PCA) to maintain the facades and ground normal estimation. As the facades are designed like concatenated grid or interlaced structures, Previtali, Scaioni, Barazzetti, Brumana, and Roncoroni (2013) detect these repetitive patterns from face segmentation, element grouping, and structural regularity estimation. They decomposed building facades into planar patches by segmentation and classified the planar segments using prior knowledge. The similarity function and pairwise transformations are computed to group the detected patches. At last, Least Square Optimization are used for quality of structure regularity estimation. Pauly, Mitra, Wallner, Pottmann, and Guibas (2008) detected repeated geometric patterns from point- or mesh-based models. The algorithm is composed of three main steps. First, decompose the input shape into small patches and estimate the patches' similarity transformations. Clustering the all transformations which produce the characteristic lattice patterns. Second, global optimizing the model parameters in the transformation space. A simultaneous registration method is used for aggregating spatial adjacent patches for the final step. This is going to build a larger repetitive pattern.

2.6. Summary

This research intends to develop a methodology for regularities description and detection through those regular and repetitive structural patterns.

3. SYMMETRY DETECTION

3.1. Description of symmetry

The symmetry detection will be applied to detect boundary symmetry, edges symmetry and interior polygon structure symmetry. Following is the axial symmetry in the interior polygon and external building boundary.

An axial symmetry is one of the transformations that maps each point P on the plane maps to another point P' on the plane. So that the line of symmetry “ e ” (see figure 3-1) would be the perpendicular bisector of the segment PP' (“Axial and central symmetry,” 2016). The axial symmetries are inverse isometries, because they preserve the distances between its points and its homologous, but its orientation is inverse, see figure 3-1 left. Another situation is that a figure break in two sections by means of a line, hence two sections are symmetrical regarding the line, see figure 3-1 right (“Symmetry Matching,” 2017.).



Figure 3-1 Axial symmetry

Bend the figure on the traced symmetry axis, it is observed that the points of the opposite parts correspond. For example, let $P = (x_0, y_0)$ be one of the points of opposite part and $P' (x, y)$ the other side, the line function that connecting P to P' is $y=k*x + b$. For calculating the coordinate of the corresponding point P' , following deductions meet the symmetry requirements.

- The middle point of segment PP' , $M \left(\frac{x_0+x}{2}, \frac{y_0+y}{2} \right)$ is in the line $y=k*x + b$, so it can generate equation 3-1.

$$\frac{y_0 + y}{2} = k * \frac{x_0 + x}{2} + b \quad (3 - 1)$$

- The segment PP' is perpendicular to the line “ e ”, that means the equation 3-2 is true.

$$k * \frac{y - y_0}{x - x_0} = -1 \quad (3 - 2)$$

Finally, solving the equation in 3-3 will generate two functions for two unknowns x and y . x and y are the coordinates of P' .

$$\begin{cases} y_0 + y = k * (x_0 + x) + 2 * b \\ k * (y - y_0) = -1 * (x - x_0) \end{cases} \quad (3-3)$$

This research has three symmetry detections. They are external boundary, edges symmetry and interior polygon symmetry. Table 3-1 is the attributes of these three types of symmetry.

Table 3-1 Properties in 2D and 3D datasets

Attributes	External boundary	Edges	Interior polygons
Geometry	<ul style="list-style-type: none"> • Shape • Area 	<ul style="list-style-type: none"> • Start/end points • Length • Orientation 	<ul style="list-style-type: none"> • Shape • Area • Orientation
Spatial distribution		<ul style="list-style-type: none"> • Parallel • Perpendicular • Intersection 	<ul style="list-style-type: none"> • Parallel • Perpendicular • Intersection

3.2. Framework

The framework of detecting symmetry is in figure 3-2. There provided three methods to detect three different symmetry types (in three colour blocks) and the input datasets are raw 3D point clouds. Different indoor mobile laser scanners provide different information. For some point clouds, it includes not only the x, y, z coordinates information, but also room information (I Armeni, Sener, Zamir, & Jiang, 2016). The room information indicates that each point has a mark which belongs to a specific room.

Before detecting different symmetry, we down sampled the point clouds for the too dense points. The segments are generated by surface growing segmentation which extracts the planar surfaces from unstructured laser point cloud data. Down sampling point clouds reduces unnecessary point clouds for the planar surface extraction, moreover, it increases the segmentation speed.

1. Boundary symmetry detection. Boundary symmetry detects building's external boundary symmetrical property. We calculate the boundary centroid and generate the line of symmetry for the boundary symmetry detection. The boundary centroid will also be used as a draft centroid in edges symmetry detection and polygon symmetry detection.
2. Edges symmetry detection. After applying the surface growing segmentation, we extract vertical segments and project them into XY plane. Each segment will be projected into a 2D edges. We keep the edge's start point and end point x, y coordinates. These edges are the original edges of the edges symmetry detection. Then generate original edges' buffer. We over flip the original edges, store the new edges, calculate the intersection ratio and keep the symmetrical edges.
3. Polygon symmetry detection. When room information was provided, horizontal segments whose height are higher than 2.3 meters should be chose. Then we extract the boundary of the segments and digitize every boundary as a polygon. Because the polygons which higher than 2.3 meters may include not only ceiling, but also large droplight, etc. Performing the intersection process to make sure that every segment is corresponding to one room is necessary. Then we detect the polygon symmetry.

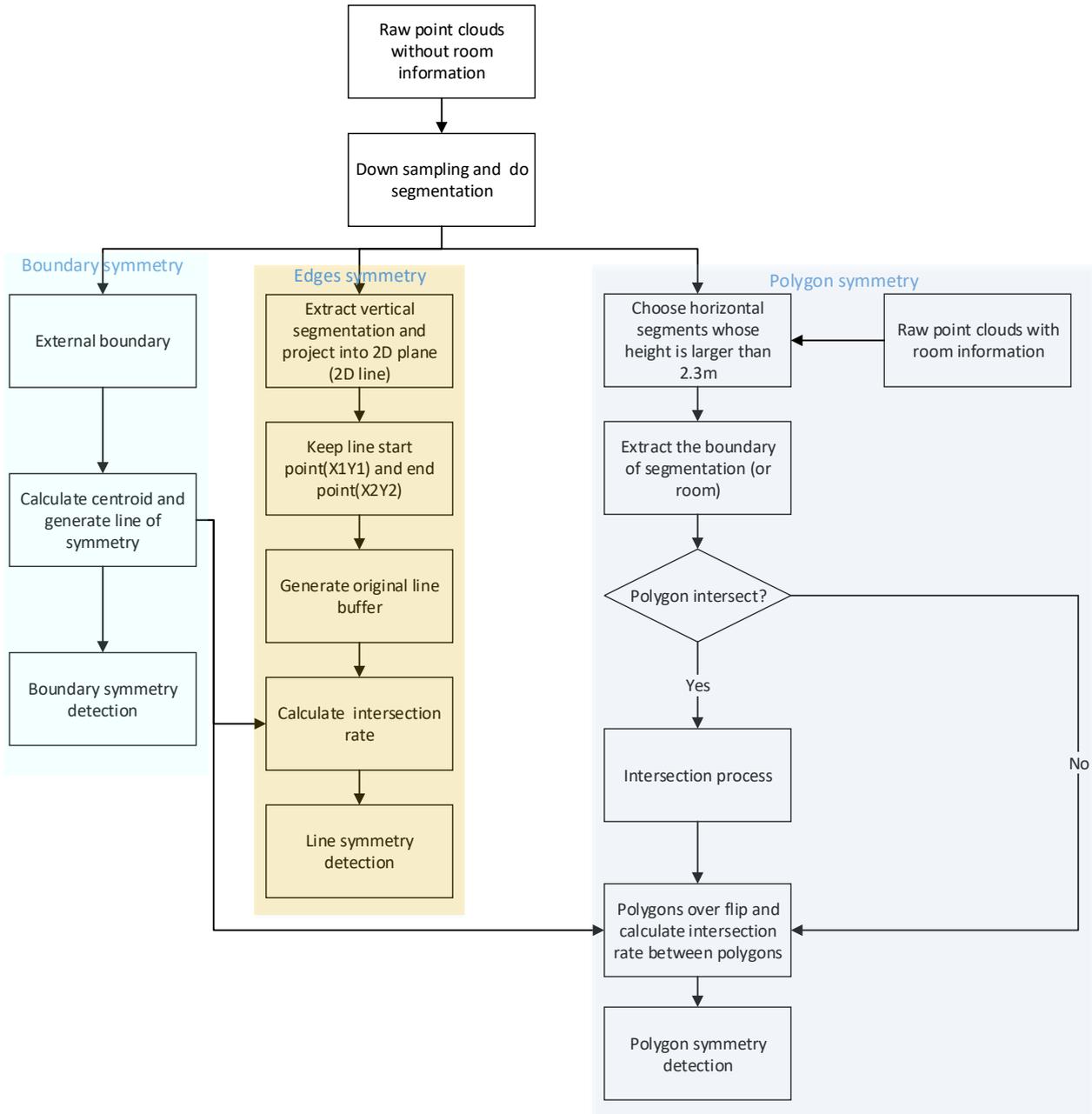


Figure 3-2 Framework of symmetry detection

3.3. Symmetry detection

3.3.1. Boundary symmetry

The external building boundary shape is very valuable for giving the first impression of the fundamental structure of the building and symmetry is a ubiquitous pattern exist in many building external boundaries. With this purpose, here we propose a method to automatically identify the symmetry of the external building boundary. The external boundary of the building per floor can be extracted from the input 2D floor plans or 3D point clouds. The external building boundary is manual digitized from 2D floor. Each corner of the external building boundary is stored in a vertex with x, y coordinates (see figure 3-3 left). The external boundary which extracted from the 3D point clouds is obtained by taking the convex hull of the 3D point clouds. The 3D points have x, y, z coordinates information, keep only the x, y coordinates for each selected boundary points (see figure 3-3 right). The difference between the

2D floor plan vertices and 3D point clouds is that the 3D point clouds have more points consisting of the boundary. Instead of the wall which is a connection of 2 vertices in digitized 2D floor plan, the external building boundary extracted from 3D point clouds is a polygonal chain and one wall is consisting of several points.

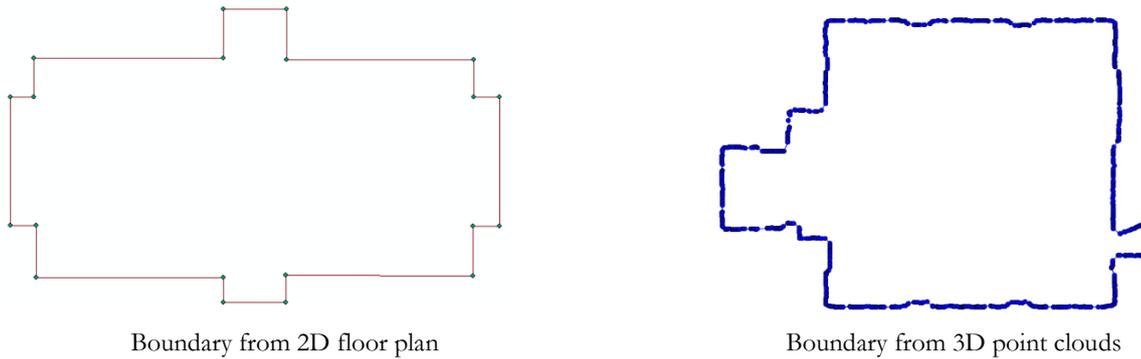


Figure 3-3 Difference of floor boundary

Since the external boundary is a general shape of the building, a simple centroid can be extracted from the average x and y value of all the external building boundary points. Several lines of symmetry candidates are generated by going through the centroid point with different angles (slope k in the line function, the angle is between -89 to 90 degrees). A construction of symmetrical copy equation (equation 3-3) is applied to all the external building boundary points and this will generate a mirrored point dataset. A mirrored external building boundary polygon will be drawn based on the mirrored point dataset. We compare the different part and common part of original polygon and mirrored polygon for each line of symmetry and plot these characteristics in a graph.

To be formal, the external building boundary is represented by a closed polyline $\mathcal{P}: (P_1, \dots, P_n)$. P_i are a series of 2D-points, $i=1 \dots n$. The symmetrical point dataset is represented by a closed polyline $\mathcal{P}_{sym}: (P_{sym1}, \dots, P_{symn})$

Identification of external building boundary symmetry

Input: External building boundary points P_i

Output: Plot of intersection and difference between original boundary polygon and mirrored boundary polygon.

Output: Optimal degree of the line of symmetry

- 1 Calculate boundary centroid based on polyline $\mathcal{P}: (P_1, \dots, P_n)$
- 2 Generate multi lines of symmetry which go through the boundary centroid with different angle (-89 to 90 degrees)
- 3 For $i \leftarrow -89$ to 90 do
 - 4 Apply the equation (3-3) for all the $P_i (i=1 \dots n)$, create a new mirrored point dataset \mathcal{P}_{mir}
 - 5 Generate original polygon based the closed polyline \mathcal{P} and mirrored polygon based on \mathcal{P}_{mir}
 - 6 Calculate intersection area and difference area between the original polygon and mirrored polygon.
 - 7 Plot the areas in the graph.
- 8 Return the optimal line of symmetry

In the symmetry detection, the line of symmetry is indispensable. $y=k*x+b$ is the equation of the line of symmetry. “k” is the slope, “a” is the angle of the corresponding slope “k”. According to the plot graph in figure 3-4, The x and y axis are the local coordinate system (meters), some mirrored polygons generated by the line of symmetry have a large area of intersection with the original polygon, others not. Thus, the larger the intersection area is, the

higher accuracy of the line of symmetric. The smaller the difference area is, the higher accuracy of the line of symmetry.

In Figure 3-4, the first column shows the angle (“a”) of the line of symmetry. In the second column, all the images a) show the original polygon(blue) and mirrored polygon (grey), Images b), the green part in the middle is the intersection of original polygon and mirrored polygon. In images c), the green part is the difference of the original polygon and mirrored polygon. From all the images b) in the second column, the intersection has the largest area when the angle is 0 degrees, the difference has the smallest area when the angle of the line of symmetry is 0 degrees.

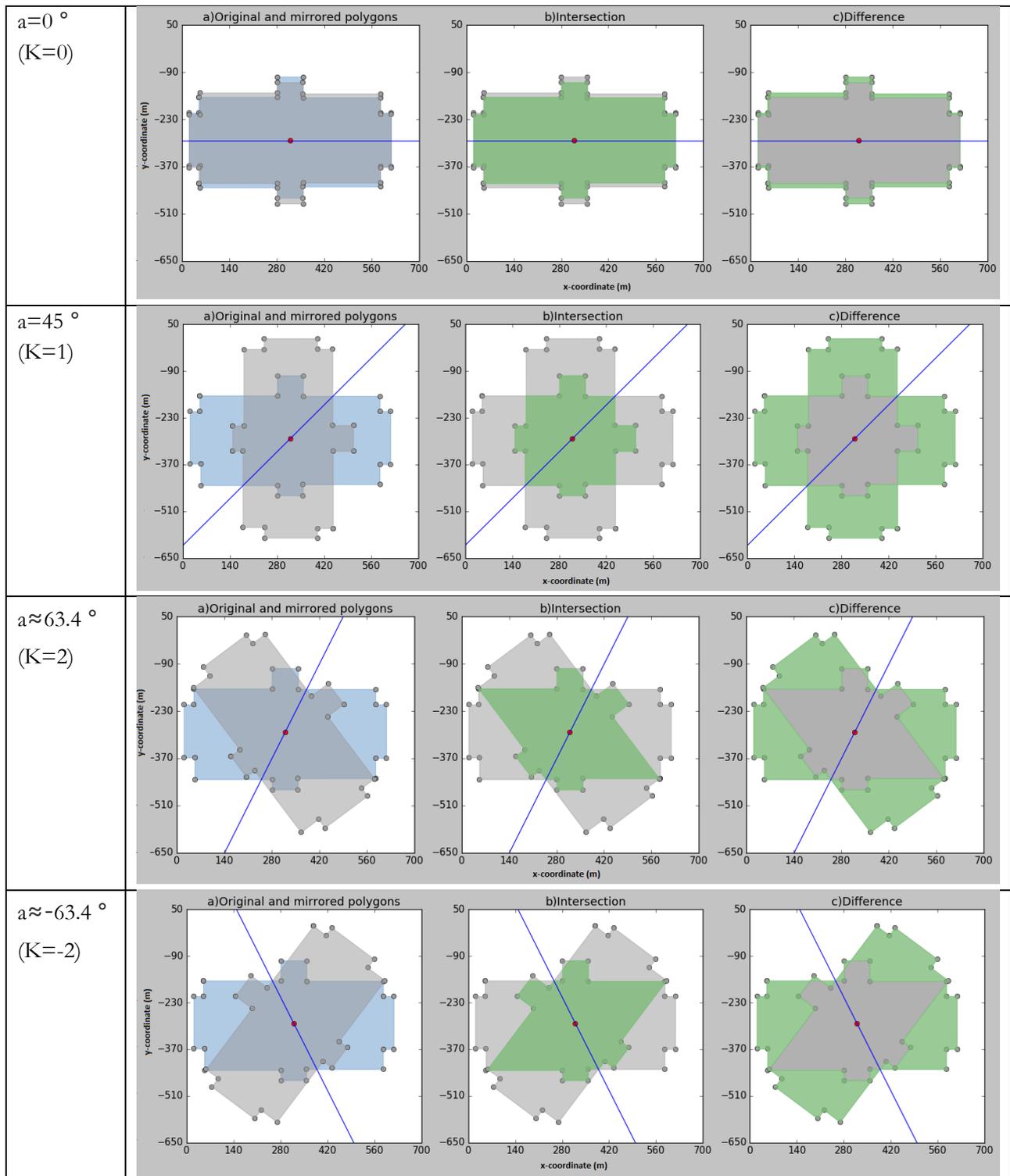


Figure 3-4 Boundary symmetry detection results with different angles

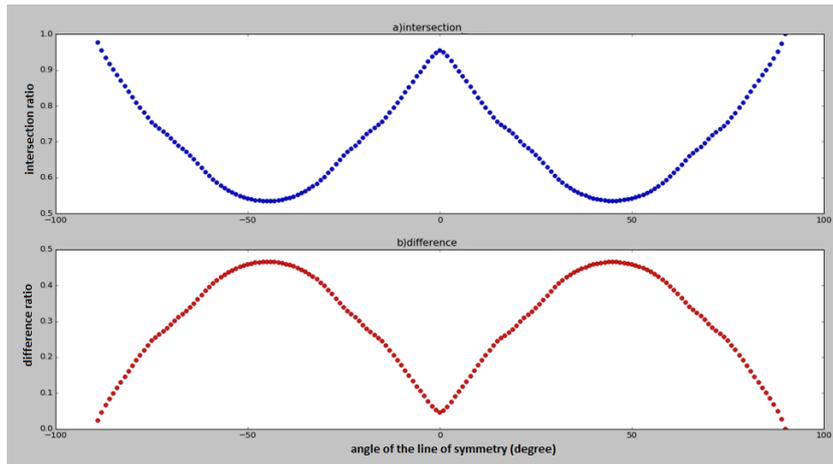


Figure 3-5 Intersection area and difference area for the line of symmetry with angles from -89 to 90 degrees

Figure 3-5 plots the intersection and difference area of external building boundary polygon and its mirrored polygon when the line of symmetry is in degrees of -89 to 90. The x axis is angle of the line of symmetry (degrees), the y axis are the intersection ratio and difference ratio from up to down. From the Figure 3-5, the intersection area is like a “W” shape and it has a peak value in 0 degrees, -89 degrees and 90 degrees. Inversely, difference area showing its minimum value in these 3 degrees. It indicates the optimal mirror line of symmetry is the very likely within these three degrees.

3.3.2. Polygon symmetry

Only showing the external building boundary information for a building is not enough. People have more interest in the variety of building interior structures. To detect whether the symmetry is also existing in the interior rooms, an approach based on method “identification of external building boundary symmetry” is explained in the subchapters.

3.3.2.1. Refinement of centroid

The centroid plays an important role in the symmetric detection. A draft centroid can be extracted by calculating the average x and y coordinates value of all the external boundary points. However, for some floors, not all the rooms are perfect symmetric. For example, left and right parts of building are symmetric along the corridor, however, up and down parts of building are not. Then the centroid selected by the average value of x, y leans to up or down parts.

To refine the draft centroid, following steps are devised to select the optimal centroid:

1. Calculate a draft centroid first (figure 3-6 blue triangle).
2. Then chose a buffer distance from the draft centroid. A certain number of candidate centroids (figure 3-6 black points) are randomly generated inside the buffer (figure 3-6 light green circle region).
3. Try angles for the line of symmetry, for each candidate centroid, in a range of (-89,90) degrees with 5 degree intervals.
4. Calculate the symmetry from all different centroid candidates and degrees combinations.
5. From the symmetry calculation, set a criterion like largest area or largest number of remaining polygons after polygon symmetry detection for selecting the optimal centroid among the candidates.

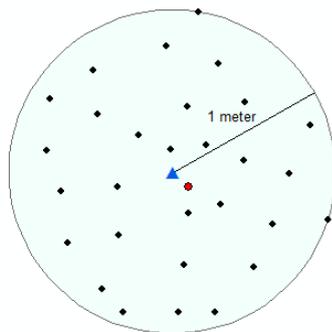


Figure 3-6 Buffer distance (1 meter), buffer region (light green circle region), draft centroid (blue triangle), 29 candidate centroids (black), one optimal centroid (red point)

3.3.2.2. Boundary intersection pre-process

We proposed two approaches to extract building interior boundary polygons. Because different indoor mobile laser scanners provide different information. For some data information, it includes not only the x, y, z coordinates information, but also room information which indicates each point has contributed to a room. The method to extract boundary polygons is projecting all the points in one room into a 2D plane and then drawing the points boundary. Keep the x, y coordinates information of points which are consisting of the boundary. Another approach is that the boundary polygons will be extracted from several segments. The segments are generated by surface growing segmentation which extracts the planar horizontal surfaces (such as ceiling) from unstructured laser point cloud data. Then extract segments boundary and keep the x, y coordinates of points which consisting of the boundary.

Because $k=\text{boundary}(x, y)$ returns the vector of point indices. $x(k), y(k)$ are a series of points coordinates which forming a 2D boundary. When 2 polygons are very close to each other, their boundary may intersect like the situation in figure 3-7 right. Figure 3-7 left is the whole room polygons in one floor. Figure 3-7 right is the close look of two polygons intersection in the red rectangle.

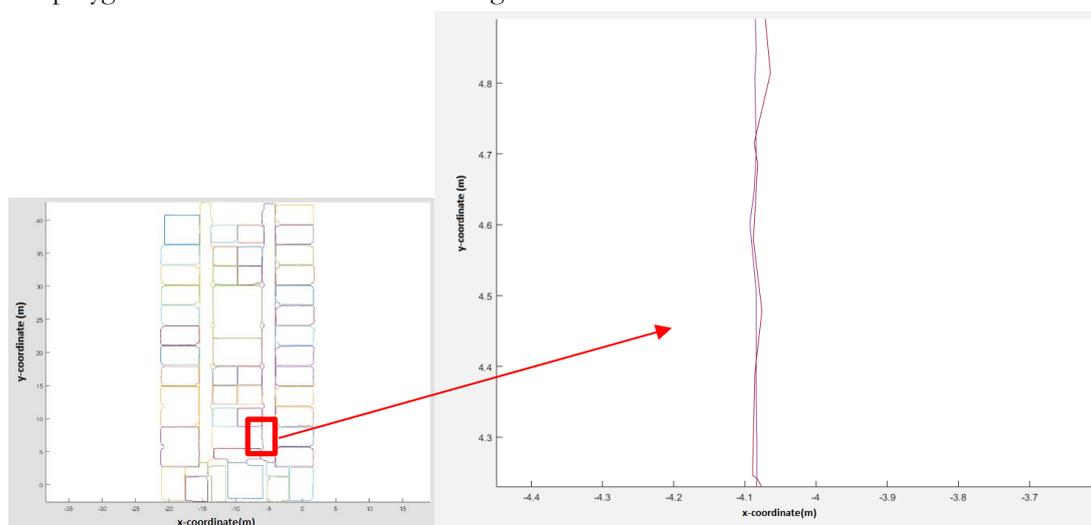


Figure 3-7 Overview of polygons per floor (left) and close look of boundaries intersect (right)

In reality, these boundary polygons shouldn't have intersections. A method to separate them is that if one polygon has intersection part with another polygon and the intersect ratio is equal or larger than a threshold (e.g. threshold = 0.3), then remove the smaller polygon. The intersection ratio defined here is the intersection area divides the smaller polygon area when two polygons have an intersection.

$$\textit{intersection ratio} = \frac{\textit{intersection area}}{\textit{smaller polygon area}}$$

The large area intersection is caused by large furniture inside the room. For example, the big drop light segment has an intersection with ceiling segment when both segments are projected into an XY plane. If the intersect ratio is small than the threshold, delete the intersection region. Because the intersection caused by closed boundary points which belong to two different polygons and the intersection region is very small which can be ignored in the symmetry detection.

3.3.2.3. Polygon symmetry detection method

Extracting the boundary of each segment using 2D alpha shape. Each boundary polygon is recorded by a closed polyline \mathbf{P} : ($\mathbf{Point}_1, \dots, \mathbf{Point}_n$). These different boundary polygons are store in \mathbf{POLY}_{ori} ($\mathbf{P}_1, \dots, \mathbf{P}_n$). \mathbf{P}_i is a polygon, $i=1 \dots n$.

To detect polygon symmetry, selecting a proper centroid is quite important. Because the line of symmetry is the line which crosses the centroid with different angles. We provide two criteria to get the optimal centroid. One is selected by area and the other is by number. Optimal centroid selected by area is designed by setting a test centroid candidates number firstly. For each centroid candidate, doing the polygon symmetry and calculating the symmetrical polygon area. The larger the symmetrical polygon area, the better of the centroid. The centroid which generated by the largest symmetrical area is the optimal centroid. Another optimal centroid selection method is by setting a test centroid candidates first. For each centroid candidate, doing the polygon symmetry and calculating the number of symmetric polygons which kept after polygon symmetry detection. The larger the number, the better of the centroid.

After selecting the centroid, then over flipping all the polygons in one side in \mathbf{POLY}_{ori} ($\mathbf{P}_1, \dots, \mathbf{P}_n$) into another side of the lines of symmetry with angles from -90 to 90. Polygons in the left side will be over flipped into the right side and right side into the left side. Over flipping the original polygons and record all the over flipped polygon into a new set in \mathbf{POLY}_{mir} ($\mathbf{P}_{mir1}, \dots, \mathbf{P}_{mirn}$). For each \mathbf{p}_j in the \mathbf{POLY}_{ori} , detect every \mathbf{P}_{mirj} in \mathbf{POLY}_{mir} to check whether they have an intersection. If they have an intersection, record two intersection ratios. Intersection ratio 1 is the intersection area divides \mathbf{P}_j area. Intersection ratio 2 is the intersection area divides \mathbf{P}_{mirj} area.

$$\textit{intersection ratio 1} = \frac{\textit{intersection area}}{\mathbf{P}_i \textit{ area}}$$

$$\textit{intersection ratio 2} = \frac{\textit{intersection area}}{\mathbf{p}_{miri} \textit{ area}}$$

Both intersection ratios will be smaller than 1 and larger than 0. A method called the symmetric level (level is different thresholds of symmetry) is designed for showing polygons symmetric level even it doesn't have a perfect symmetry. If both intersection ratios are higher than 0.01, the polygons will be stored. 0.01 is to make sure that the polygons are possible to symmetric to another polygon even the symmetric level is low. The level can be decided by the difference of intersection ratio 1 and intersection ratio 2. If the absolute value of the intersection ratio difference is smaller than 0.1, colouring them by the average value of intersection ratio 1 and 2. The higher the average value is, the deeper of the colour. If the absolute value of the intersection ratio difference is larger than 0.1, assigning both of polygons white colour.

If the intersection ratio 1 and intersection ratio 2 are both higher than a threshold (e.g. threshold = 0.7), the polygon pairs will be considered as the symmetric to each other. The threshold is a parameter of symmetry standard to detect whether the two polygons are symmetric to each other. The higher the threshold, the more symmetric of the polygons.

Polygon symmetry detection

Input: 3D building point clouds

Output: Pairwise symmetry polygons level $\mathcal{POLY}_{level} (P_{level1}, \dots, P_{leveln})$.

Output: Pairwise symmetry polygons $\mathcal{POLY}_{sym} (P_{sym1}, \dots, P_{symn})$.

- 1 Pre-process of the raw point clouds
- 2 Extract the boundary of segmentation store in $\mathcal{POLY}_{ori} (P_1, \dots, P_n)$
- 3 for $P_i \leftarrow 1$ to n do
- 4 └ Polygon intersection process, generate P^i

- 5 for centroid candidate number $\leftarrow 1$ to m do
- 6 └ for angle $\leftarrow -89$ to 90 do
- 7 └ for $P^i \leftarrow 1$ to n do
- 8 └ Polygons over flip and calculate intersection ratio between polygons
- 9 └ Keep the different level of symmetry part $\mathcal{POLY}_{level} (P_{level1}, \dots, P_{leveln})$.
- └ Keep the symmetric part $\mathcal{POLY}_{sym} (P_{sym1}, \dots, P_{symn})$.

3.3.2.4. 2D map polygon symmetry detection

The building interior structure is manual vectorised from a 2D map. The corner points of each room are represented by 2D vertices and they are stored in a .txt file in a clockwise order, the information includes room index number, vertex x, and y coordinates.

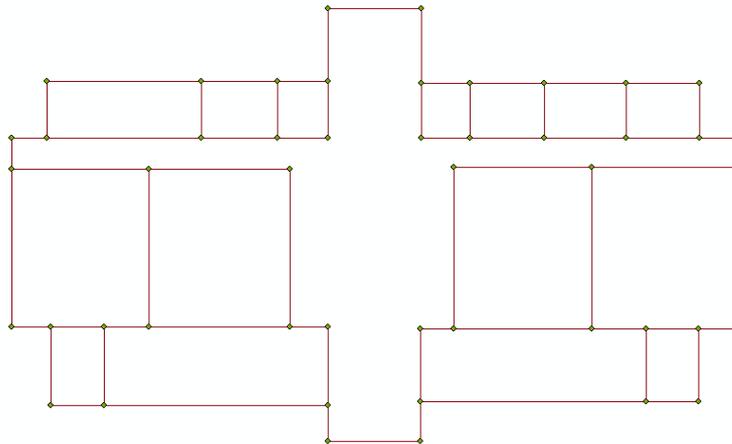


Figure 3-8 Digitized 2D floor map

Figure 3-8 shows the interior structure of the building which vectorised from the original 2D floor plan in figure 1-1 right. The blue points are the room corner points which can be used for generating room polygons. In figure 3-9 a), most of the polygons are symmetric to each other on the two sides of the line of symmetry in 90 degrees, except two on the top left and top right of the room. The top left one is larger than the top right one.

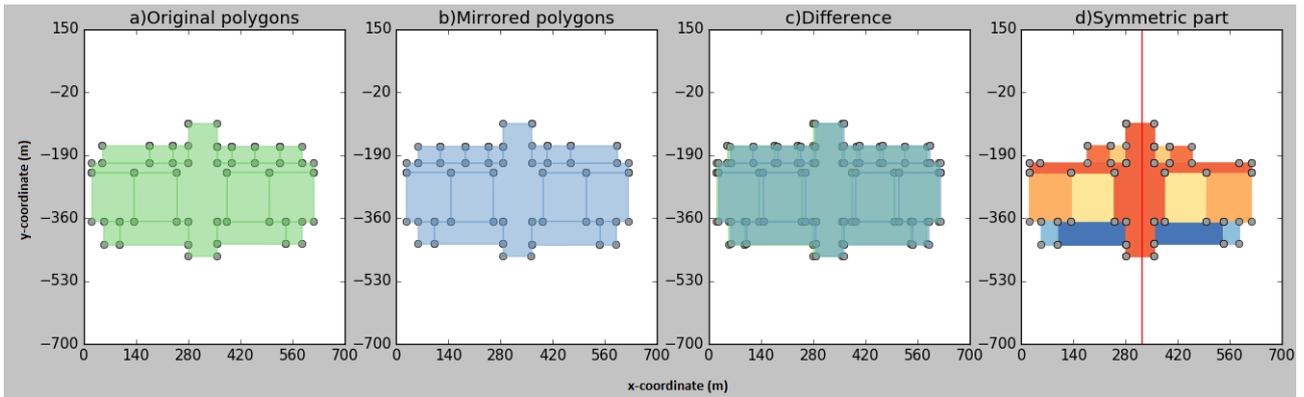


Figure 3-9 Inner building symmetry

Figure 3-9 a) is showing the original polygons. After applying the symmetry detection from the draft centroid, figure 3-9 b), mirrored polygons are generated. Figure 3-9 c) is the overlap of the original and mirrored polygons. A small difference can be seen from there. After polygon symmetry detection (intersection ratio > 0.7), only the symmetrical parts are kept in figure 3-9 d). On the left side and right side of the line of symmetry (red line in figure 3-9 d)), the symmetrical pair polygons are shown in the same colour. The top left and top right polygon which are not symmetrical to each other in the original polygons have been removed from the symmetric part.

3.3.2.5. Room polygon and segments polygon symmetry detection. (Optimal centroid selected by area.)

- Figure 3-10 a) are the original boundary polygons which are extracted from the room information. Figure 3-10 b) is the different level of symmetry parts. The deeper the colour, the more symmetric of two polygons. Most of the polygons have corresponding polygons with a high level of symmetry which shown in a dark red. In the middle bottom of the floor, one polygon is shown in orange colour which means it doesn't have a perfect symmetric polygon in another side of the line of symmetry. The white colour polygons are in the low symmetry level. Figure 3-10 c) is the remaining symmetric part after a polygon symmetry detection. Symmetry comes with a pair. For a better recognition, two symmetric polygons take the same colour located in the two sides of the line of symmetry.

Table 3-1 is the polygon symmetry detection results from figure 3-10. The number of the original polygon is 48. All the polygons' area sum is 890.848 m². The symmetrical area is 853.098 m². Thus, 95.8 % of the room polygons have the symmetry property. The draft centroid (-9.652, 20.063) is the blue triangle in the figure 3-10 b) middle. The optimal centroid selected by the largest symmetrical area is the red point in the figure 3-10 b). Its coordinate is (-9.699, 19.599) which is a little bit lower than the draft centroid. The optimal angle of the line of symmetry is 90 degrees, shown in figure 3-10 b) and 3-10 c) red line.

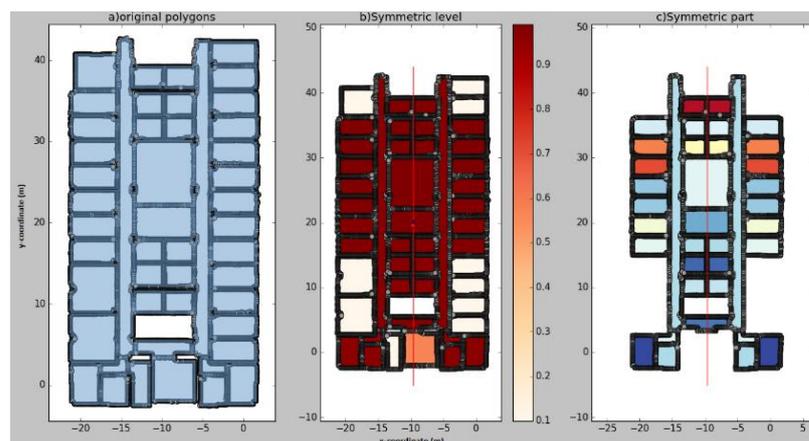


Figure 3-10 Identification of room polygon symmetry by area

Table 3-1 Parameter of room polygon symmetry detection by area

number of original polygons	48
all polygons area (m ²)	890.848
max area (m ²)	853.098
draft centroid	(-9.652, 20.063)
optimal centroid	(-9.699, 19.599)
optimal alpha	90

2. When the boundary polygons are extracted from segments, the shape is more complicated than the polygon which extracted from room information. For example, figure 3-11 a) is the original boundary polygons. The right corridor which are supposed to be one straight polygon becomes two overlapped polygons because of the segmentation. In figure 3-11 b), the up part of right corridor polygon is shown in white colour which means it has a low-level symmetry with the left corridor polygon. In figure 3-11 c), the up part of the right corridor polygon was removed after polygon symmetry detection because it doesn't meet the polygon symmetry threshold (threshold=0.7). Thus, segmentation result has a big influence on the polygon symmetry detection.

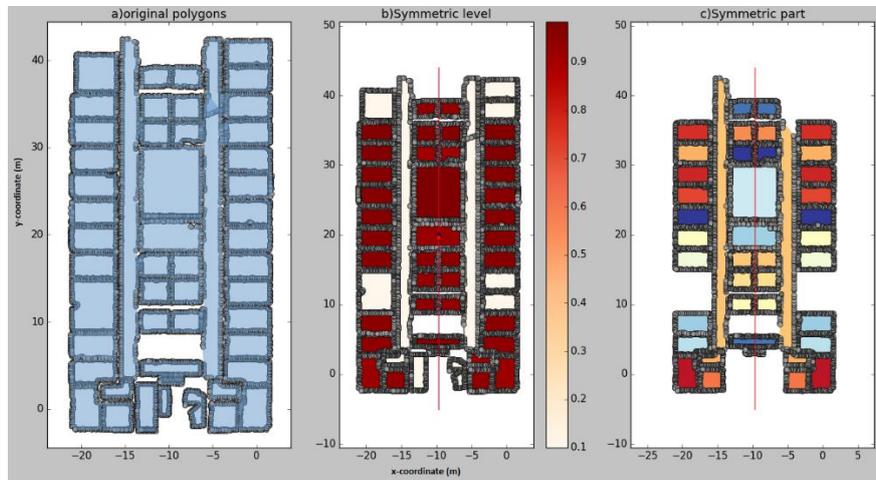


Figure 3-11 Identification of segment polygon symmetry by area

Table 3-2 Parameter of segment polygon symmetry detection by area

number of original polygons	72
all polygons area	783.488
max area	722.296
draft centroid	(-9.652, 20.063)
optimal centroid	(-9.669, 19.537)
optimal alpha	90

3.3.2.6. Room Polygon and segments polygon symmetry detection. (optimal centroid selected by number)

The number of the line of symmetry per floor is not only one in some cases. Two or more lines of symmetry may exist in the floor. However, one method of polygon symmetry detection is by detecting the largest symmetrical polygon area. Another method to detect the line of symmetry is by counting the most number of symmetry polygons during the polygon symmetry detection. The more number of symmetric number of polygons kept after the polygon symmetry detection, the better of the line of symmetry. Figure 3-12 a) is the original polygons. Figure 3-12 b) kept the largest number of polygons after the polygon symmetry detection. However, figure 3-12 c) has only few polygons which meet the symmetry requirement (intersection ratio >0.7) based on the line of symmetry in 0 degrees in the centroid of (-8.661, 19.915).

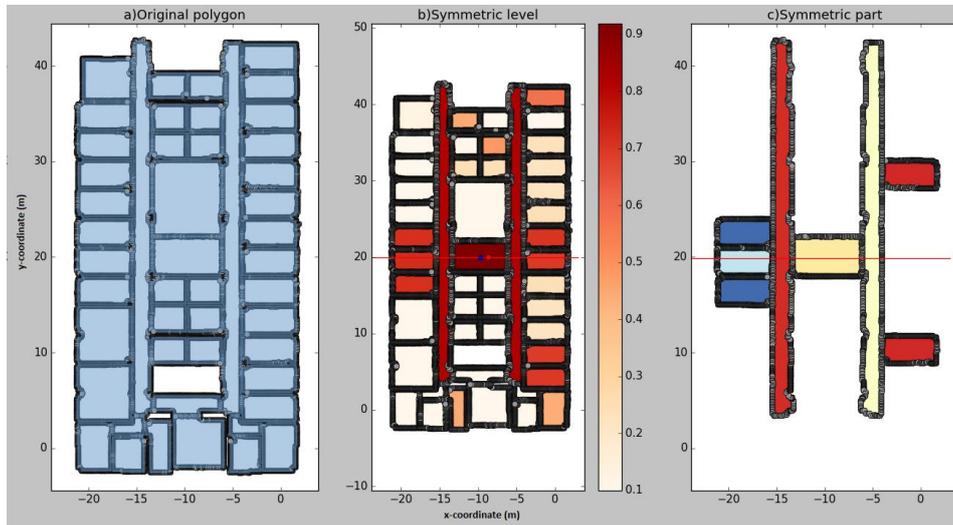


Figure 3-12 Identification of room polygon symmetry by number

Table 3-3 Parameter of room polygon symmetry detection by number

number of original polygons	48
number of symmetric polygons	45
draft centroid	(-9.652, 20.063)
optimal centroid	(-8.661, 19.915)
optimal alpha	0

When the boundary polygons are extracted from segments, select the optimal centroid by the largest number of polygons which are remaining after the segment polygon symmetry detection. The result is as figure 3-13.

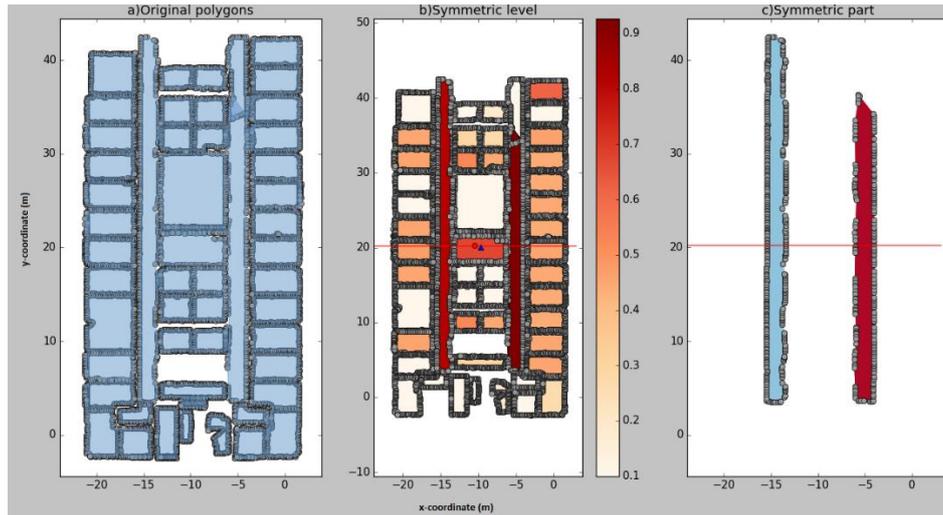


Figure 3-13 Identification of segment polygon symmetry by number

Table 3-4 Parameter of segment polygon symmetry detection by number

number of original polygons	72
number of symmetric polygons	50
draft centroid	(-9.652, 20.063)
optimal centroid	(-10.410, 20.269)
optimal alpha	0

3.3.3. Edges segmentation and projection from 3D point clouds

If the raw point clouds don't have room information, or the point clouds don't include the ceiling points, or all the ceiling are segmented into one large planar segment, the boundary polygons extracted from ceiling points are not possible. Then the edges symmetry could help to detect the building symmetry character. The edges symmetry approach is similar to the polygon symmetry detection. But because the data formats from polygon and edges are different, the processing procedures is a little bit different which was described in the following subchapters. Comparing to the polygon symmetry detection results, the edges symmetry detection result only show the symmetrical edges. It cannot show the edges symmetrical level.

From 3D point clouds extracting the wall segment and project it into a 2D plane. The following steps should be executed to get the figure 3-14:

1. Subsample datasets, the parameter setting is that the minimum distance between two points is 0.1 meters;
2. Segment the subsampled point clouds using surface growing in this experiment.;
3. Compute the normal vector per segment using principal component analysis (PCA);
4. Compute minimum, maximum z values (for the usage of selection of segments height);
5. Select segments which parallel or perpendicular to the x axis;
6. Fit a line to the x, y coordinates of the segment's points (using PCA);
7. Project the segment's points onto the line, find min and max (start and end) points of the edge;
8. Save to a text file: x1, y1, x2, y2, number of points, min z, max z, room number.

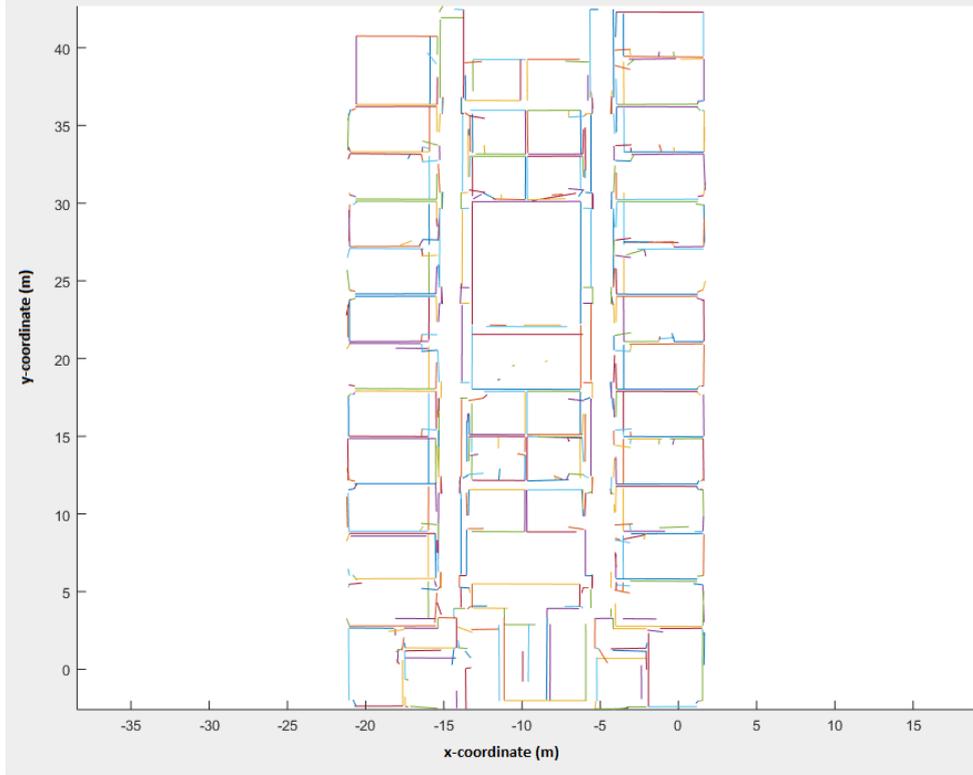


Figure 3-14 Segmentation and projection of 3D point clouds into 2D plane

3.3.3.1. Edge symmetry detection

A sequence of edges is presented in $\mathbf{E}_{ori} (\mathbf{e}_1, \dots, \mathbf{e}_n)$, each \mathbf{e}_i is consisting of the the start point (\mathbf{st}_i) and end points (\mathbf{en}_i). For each edge \mathbf{e}_i , generating a distance of 0.1-meter buffer, the buffer is a polygon which is in the buffer set $\mathbf{B} (\mathbf{b}_1, \dots, \mathbf{b}_n)$. We over flip each segment \mathbf{e}_i and calculate the intersection with all buffer polygons in $\mathbf{B} (\mathbf{b}_1, \dots, \mathbf{b}_n)$. If the \mathbf{e}_i has an intersection with one or multiple buffer polygons, calculate the intersection ratio 1 and intersection ratio 2. Intersection ratio 1 defined in this chapter is the value of the intersection length divides the length of \mathbf{e}_i in \mathbf{e}_{ori} . Intersection ratio 2 is equal to the value of the intersection length divides the length of \mathbf{e}_{miri} in \mathbf{e}_{mir} .

$$\text{intersection ratio 1} = \frac{\text{intersection length}}{e_i \text{ length}}$$

$$\text{intersection ratio 2} = \frac{\text{intersection length}}{e_{miri} \text{ length}}$$

Only if intersection ratio 1 and intersection ratio 2 are both higher than a threshold (e.g. threshold = 0.6), then the \mathbf{e}_i in \mathbf{e}_{ori} and \mathbf{e}_{miri} in \mathbf{e}_{sew} are symmetric to each other. Use the optimal centroid and degree combination to detect line symmetry. The criteria for selecting the best symmetry result is the highest number of \mathbf{e}_i in $\mathbf{E}_{ori} (\mathbf{e}_1, \dots, \mathbf{e}_n)$ which are kept after edges symmetry detection.

Edges symmetry detection

Input: 3D edge start point (\mathbf{st}_i) and end points (\mathbf{en}_i)

Output: Pairwise symmetry edges $\mathbf{E}_{new} (\mathbf{e}_1, \dots, \mathbf{e}_n)$.

- 1 Vectorise the start point (x1 y1) and end point (x2, y2) into an original edges $\mathbf{E}_{ori} (\mathbf{e}_1, \dots, \mathbf{e}_n)$

```

2 For  $i \leftarrow 1$  to  $n$  do
3   Over flip from the one side of the line of symmetry into another side, store new edges in  $E_{mir}$ ;
4   Generate a buffer within 0.1 meter, store buffer polygon in  $B(b_1, \dots, b_n)$ ;
5 For  $j \leftarrow 1$  to  $n$  in  $E_{mir}$  do
6   Calculate two intersection ratios (intersection ratio 1 and intersection ratio 2)
7   If two intersection ratios are both larger than 0.6 then
8     Keep the  $e_i$  in  $E_{new}$ 

```

Figure 3-15 is the result of applying the edge symmetry detection. Figure 3-15 a) is the original edges (blue edges) and mirrored edges (magenta edges) which extracted by the draft centroid (blue triangle in the middle). The blue line in figure 3-15 a) is the line of symmetry cross the draft centroid. Figure 3-15 b) is the original edges (blue edges) and mirrored edges (magenta edges) which extracted by the optimal centroid (red point in the middle). The red line is the line of symmetry which cross the optimal centroid with 90 degrees. The draft centroid is very close to the optimal centroid. Much more overlap can be seen from the figure 3-15 b) when comparing to figure 3-15 a). Figure 3-15 c) is the symmetrical parts which kept after edge symmetry detection. From table 3-5, 134 symmetrical edges out of 247 original edges were remaining after edges symmetry detection. The optimal centroid is near the draft centroid and the optimal alpha is 90 degrees.

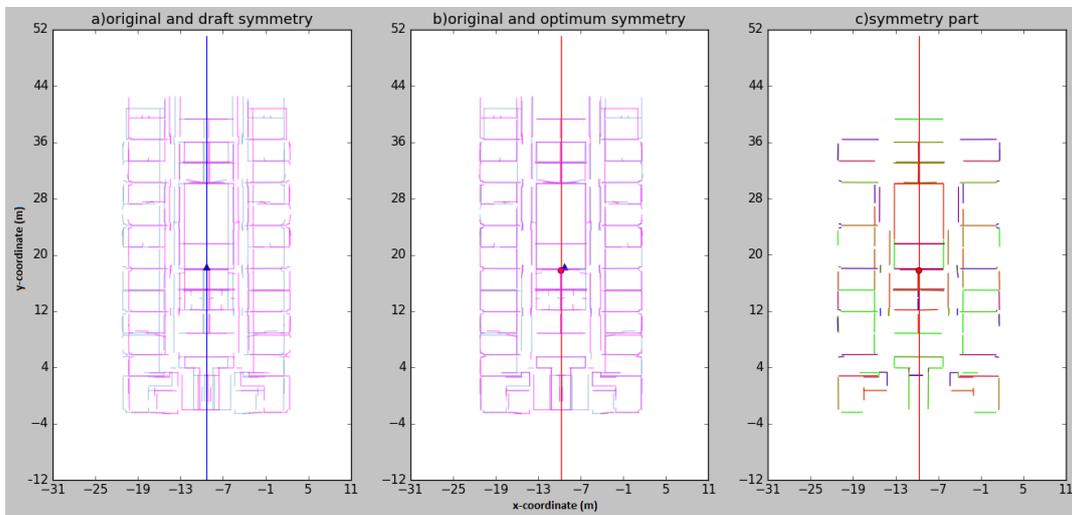


Figure 3-15 Identification of edge symmetry

Table 3-5 Parameter of edge symmetry detection

original edges	247
optimal number of edges	134
draft centroid	(-9.340, 18.346)
optimal centroid	(-9.761, 17.778)
optimal alpha	90

4. REPETITIVE PATTERNS

4.1. Repetitive patterns description

A geometric pattern is consisting of elements which repeated in a discernible regularity (Torii, Sivic, Okutomi, & Pajdla, 2015). Each element has its features. They are retrieved into three main types: orientation, metric and topology (Sithole & Zlatanova, 2016). Different types of features are represented by different visualization variables based on their characteristics. Unit shows their quantitative properties. Table 4-1 lists eleven features.

Table 4-1 Repetitive patterns

ID	Repetitive patterns	Representative	Unit	Type
1	Edge length	Histogram	Meter	Metric
2	Wall thickness	Broken line graph	Meter	Metric
3	Floor center & polygon centroid	Point	x, y coordinates	Orientation
4	Distance from center	Histogram	Meter	Metric
5	Polygon area	Histogram	Square meter	Metric
6	Bounding box area	Histogram	Square meter	Metric
7	Polygon percentage in bounding box	Bar chart	%	Topology
8	Concave corner area	Bar chart	Meter	Metric
9	Orientation	Histogram	Degrees	Orientation
10	Bounding box Length	Bar chart	Meter	Metric
11	Bounding box Width	Bar chart	Meter	Metric

There extracted 11 features and four of them are used to detect the similar rooms inside the indoor environments. Figure 4-1 is the framework of eleven features and four which are used to detect similar rooms.

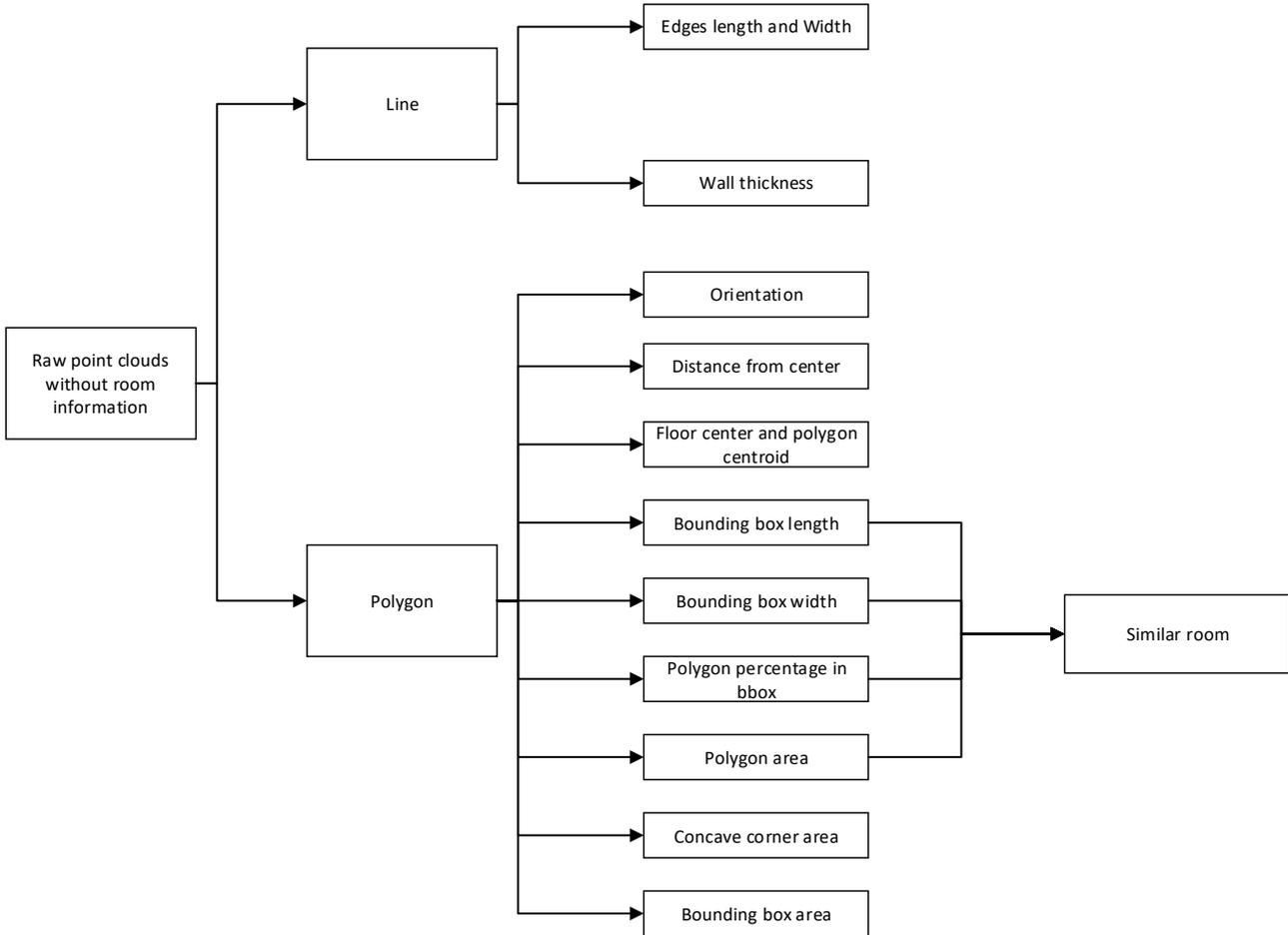
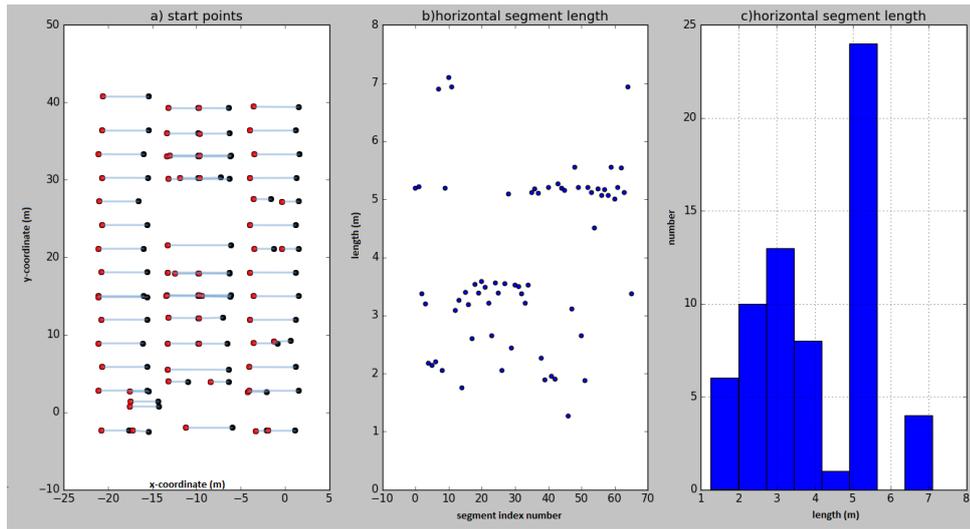


Figure 4-1 Framework of repetitive pattern detection

4.2. Edges length

The method applied to extract edges from figure 3-14 which are parallel to x-axis is shown in figure 4-2 a). Method applied to extract edges from figure 3-14 which are perpendicular to x-axis is shown in figure 4-3 a). Each edge keeps the information of x and y coordinates in local coordinate (meters) of the start point and end point.

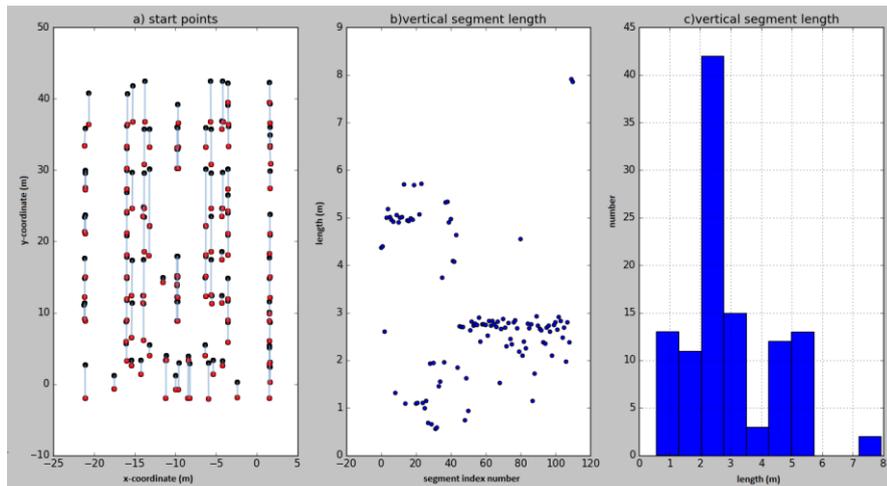
Edge length is the distance between the start point and end point of each segment. From the figure 4-2 b), x axis is the segment index number, y axis is the length of the segments. Most of the points plot around 5 to 5.2 meters and a small number of points are separated around 3 to 3.5 meters. Those smaller segments are very likely to be the cabinets in each room which are hard to filter. See the histogram in figure 4-2 c), x axis is the segment length interval and y axis is the number of segments fall in the interval, a peak value in the interval of 4.8 to 5.5 meters approves that most of the wall length is between 4.8 to 5.5 meters.



Edges parallel to x-axis Edges length plot Edges length histogram

Figure 4-2 Edges length

From the figure 4-3 plot (middle), x axis is the segment index number, y axis is the segment length, a variety of wall length and one peak value are shown. The peak interval is between 2 to 2.8 meters in figure 4-3 (right). The x axis is the length interval and the y axis is the number of the length within the interval. The main length of the edge which perpendicular to x-axis is concentrated in this interval.



Edges perpendicular to x-axis Edges length plot Edges length histogram

Figure 4-3 Edges length

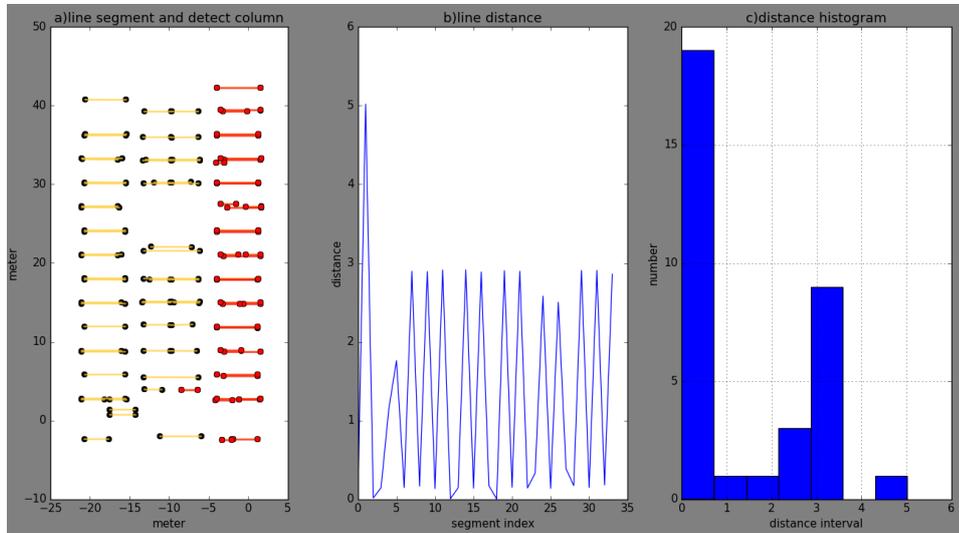
4.3. Wall thickness

When considering the detail of the 2D edges in figure 4-4 left which extracted by the 3D point clouds, a repetitive distance patterns showed in the figure 4-4 b). To detect this pattern, the following steps are designed:

1. Ordering (sort) start points x coordinates from small to large and keep them in list A first.
2. In the list, we use the first (smallest) x coordinate to minus the next one and keep the differences in another list B.
3. In list B, we use outliers' extraction to get the biggest difference value. Because start points in the same column have similar x coordinates. The difference between them is small. However, the values of x

coordinate in different columns have bigger differences. The big difference is utilized for classifying them into different groups.

4. Then, we divide the dataset by outliers` values. They are divided into 3 groups in the test data.
5. In each group, we order each dataset by y coordinates value and keep them in the different lists.
6. Last, we calculate the distance between the first line and the next in each list, plot the distance and create a distance histogram like figure 4-4 right.



Last column edges

Distance between edges

Distance between edges histogram

Figure 4-4 Repetitive distance patterns

The represented data in figure 4-4 (left) calculates the distance between edges in the last column. Figure 4-4 (middle) shows a repetitive distance pattern. They are the distances between edges which can be deduced as room width and wall thickness. The larger value is the room width and the smaller one is the wall thickness. In the figure 4-4 (right) image, two peaks are shown. The first peak is the wall thickness value and the second peak is the room width value. This approach detects the edges which are parallel to x axis. If the edges don't have x or y axis orientation, detect the edges orientation first. Then perform above steps to detect the repetitive distance patterns. The histogram of distance between the edges should be similar to figure 4-4 right.

4.4. Floor center and polygon centroid

Floor Center: X, Y average value of all polygons in one floor (green triangle in the middle of figure 4-5 left).

Centroid: X, Y average values of each polygon boundary points. (red circle in middle of each polygon in figure 4-5 left). Each red point represents a polygon, a polygons distribution graph is simplified.

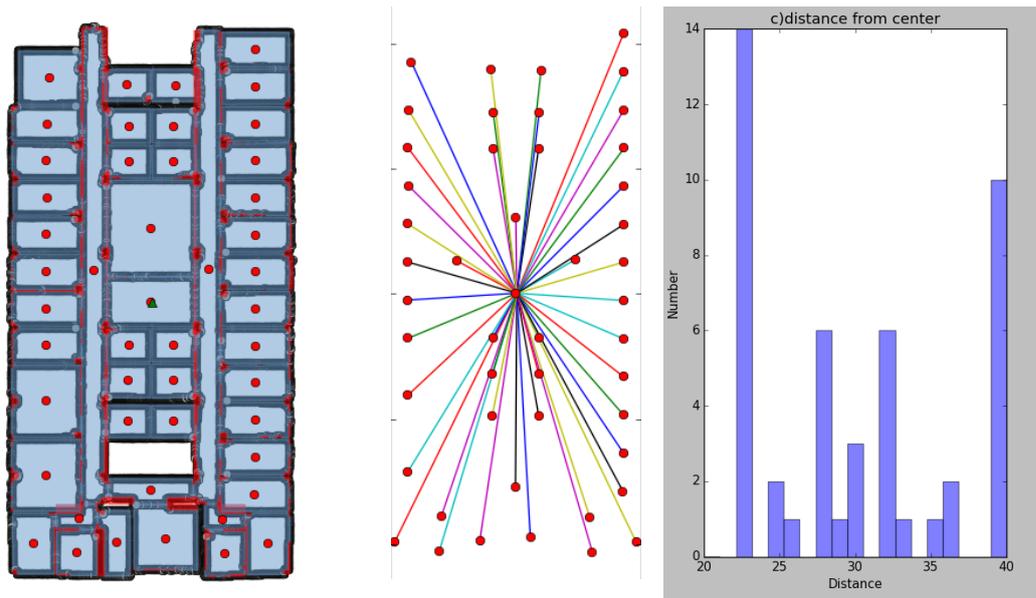


Figure 4-5 Floor center and polygon centroid, center centroid connection, and distance from center

4.5. Distance from center

When connecting all the centroid to a center, a radial pattern can be seen for figure 4-5 middle. Figure 4-5 right is the histogram of distance from polygon centroid to center. From the histogram, the distances also show a thoroughly symmetry pattern with a line of symmetry in the 30 meters.

4.6. Polygon area, Bounding box area, and concave corner area

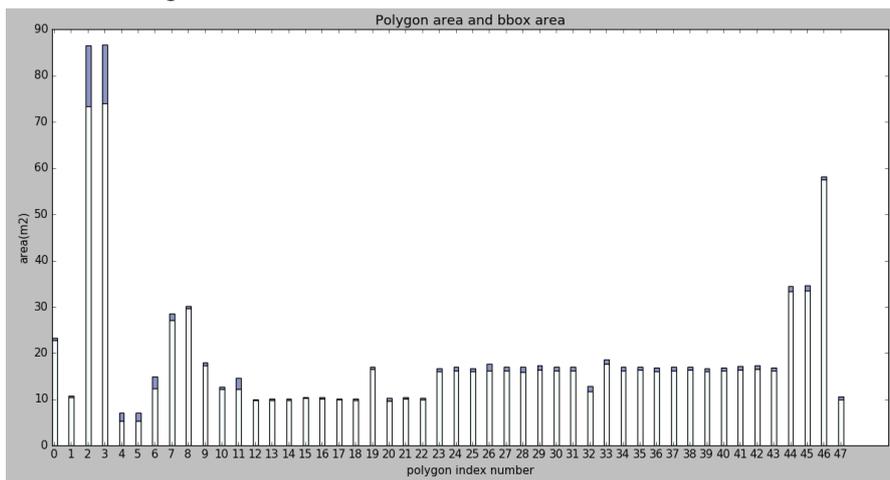


Figure 4-6 Polygon area (white bar), bounding box area is a bit bigger (blue) than polygon area

1. Polygon area

The polygons which extracted from 3D point clouds are different from each other. Each polygon has a different area and this is a special characteristic for each polygon. The white bar length in Figure 4-6 shows the area in different polygons.

2. Bounding box area

A point set bounding box is the box with the smallest measure within which all the point lie in geometry (O'Rourke, 1985). The default bounding box for each polygon is x-axis aligned and y-axis aligned. The bounding box area is the white bar plus the blue part height in Figure 4-6.

3. Concave corner area

When using a bounding box to wrap a polygon, the empty places outside of polygon but inside of bounding box are the concave corners, it has dents at vertex (Torii et al., 2015). See figure 4-5 left, the red area around each polygon is the polygon's concave corner area. The concave corner area is calculated by the difference of bounding box area and polygon area. The blue bar length on the top of white bar is the concave corner area of each polygon shown in figure 4-6.

4. Polygon percentage in bounding box

The polygon percentage in bounding box is the value of polygon area dividing boundary box area. It returns a relative comparison of polygons.

4.7. Arbitrary orientation of point clouds

An x-axis aligned and a y-axis aligned bounding box is useful for x- and y-axis aligned polygon. However, some 3D point clouds are not axis-aligned. Thus, its orientation cannot be detected by axis-aligned bounding box. One method is using the minimum (smallest) area bounding box to calculate subject to no constraints of the orientation (O'Rourke, 1985). The main idea of extracting one polygon's orientation is like following ("Geometry - python: Help to implement an algorithm to find the minimum-area-rectangle for given points in order to compute the major and minor axis length - Stack Overflow," 2012):

- Calculate the convex hull of the input point clouds;
- Calculate each edge's orientation in the convex hull;
- Rotate the convex hull following the edge's orientation.
- Compute the bounding box area with minimum of the rotated convex hull;
- Record each orientation with the corresponded area;
- Select the smallest area bounding box as the input point clouds' minimum area bounding box.

Figure 4-7 (left) is the arbitrary orientation polygons. Figure 4-7 (middle) is an enlarged polygon with its minimum area bounding box. Figure 4-7 (right) is the angle of the polygons' orientation.

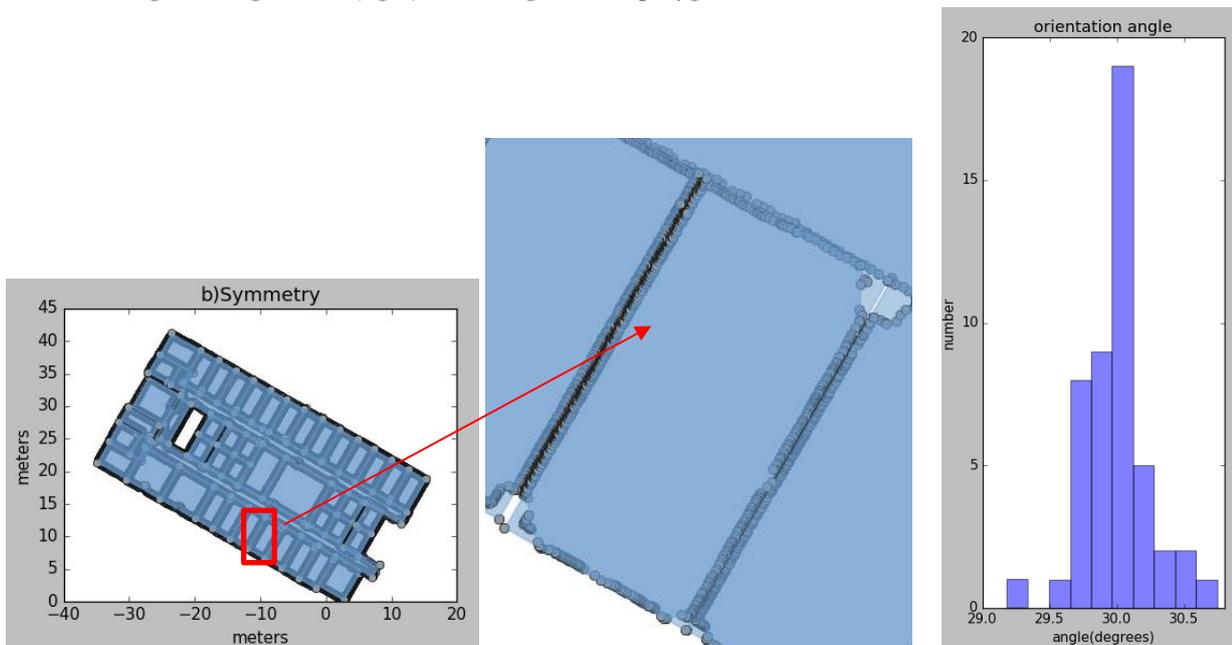


Figure 4-7 Minimum area bounding box overview of floor, detail polygon, and orientation angles

4.8. Bbox Length and Width

1. Bounding box Length: Fit every polygon a corresponding bounding box and then calculate the bounding box length. The bounding box is axis-aligned. The length in the x direction. See Figure 4-8.
2. Bounding box Width: Fit every polygon a corresponding bounding box and then calculate the bounding box width. The width in the y direction.

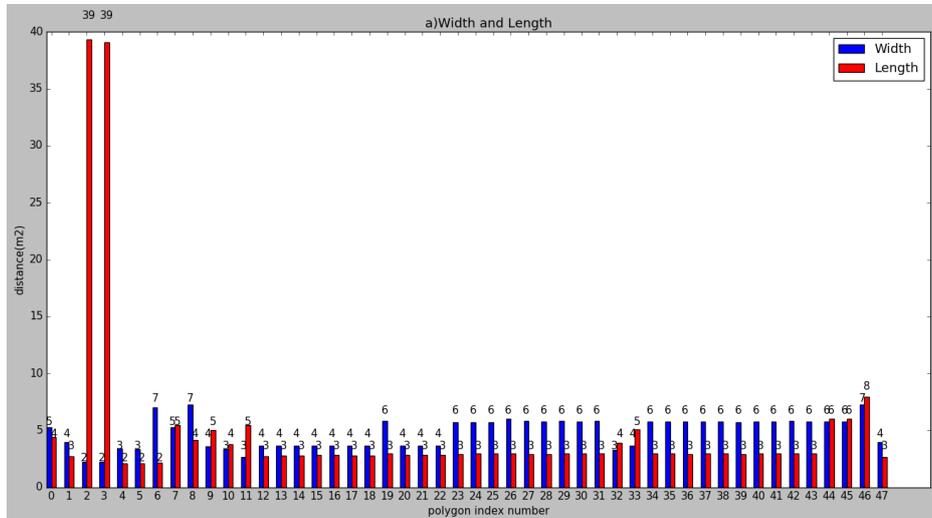


Figure 4-8 Round value of width and length of bounding box

4.9. Identification of similar polygons

After calculating those features in the room environment, some of the features can be applied to explore the similar rooms. Because the rooms are represented by polygons, we devised two methods to find the similar polygons.

4.9.1. Identification of similar polygons by weight

Identification of similar polygons by weight

Input: 3D point clouds

Output: Similar polygons groups

- 1 Calculate several features like bounding box width, bounding box length, polygon areas, polygon percentage in the bounding box;
- 2 Normalize all the features into 0-1;
- 3 Assign each feature a random weight and sum the product of feature values and weights. Each polygon will have a sum value;
- 4 Group the polygons with similar sum value. The indexes of polygons in the same group are stored in the same list;
- 5 Compare the group polygons with the ground truth group polygons;
- 6 Keep the group polygons which are most similar to the ground truth data.

The criteria to select the most similar weight combination is set as:

- Initial a variable count which is equal to 0;
- If the number of group members equal to 1 and there is a same group in the ground truth, count add one;
- If length of group is larger than one, when the ground truth is included or equal to the group classified by weights, then count add one;

- For each weight combination, a unique count value will be generated. The weight combination which corresponding the highest count value is the best result. If the number of weight combination is enough, there will be more than one best result like figure 4-10.

Many features can be calculated to differentiate polygons. However, the width, length, polygon area and polygon percentage in the minimum bounding box are the most prominent features to differentiate different polygons. Thus, the first step is to calculate those features. Since several features have different units and scale. It is important to quantize these features into a standard unit firstly. For rescaling data, all the variable in one feature will in the arrange of 0 to 1. This method allows variables which still have different standard deviations and means but the equal ranges (“Methods for data standardization,” 2017). To normalize a value, the formula is in equation 4-1.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (4-1)$$

The x represents the value of a variable in the original data set, x' is the normalized value. $\max(x)$ and $\min(x)$ are the maximum value and the minimum value of the whole original dataset separately (Juszczak, Tax, & Dui, 2002). After normalizing the dataset, next step is to assign a weight to each feature. For better comparing the weight difference between different experiments, the sum of the weight which assigned to each feature should add up to one. Then the differences of weights can be calculated. There are many combinations of the weights. The more features, the more combinations. A certain number of weights combination has to be chose firstly. For each weight combination, a weight was assigned to a feature and then calculate the product of feature value and weight. We sum all the products of all features and generate a final value for each polygon. Last, we group the polygons with similar sum values. For each different weight combination, different groups will be generated. Because a large number of weights combination are selected randomly, there will show both good results and poor results. If the groups are classified most similar to the ground truth data, keep the weight combination. Otherwise, ignore the weight combination. Figure 4-9 is the ground truth.

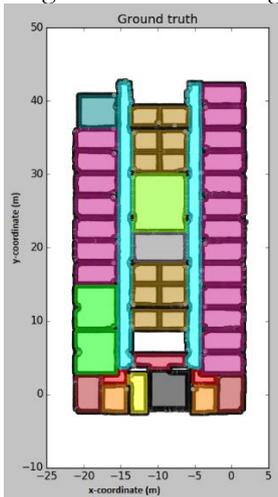


Figure 4-9 Ground truth data

Figure 4-10 are the good classified results which are selected automatically as the closest to the ground truth data among the 100 results. Each weight combination is listed in table 4-2 from left to right. The first column is the weight combination index. The four weights are the bounding box width, length, polygon area, and polygon percentage in the bounding box.

For all image in figure 4-10, the leftmost and rightmost polygons (blue) are well classified into the same group. So do the black polygons in the middle top and middle down. The two corridor shape polygons are grouped into the same class. The difference of each image is that the bottom polygons are classified into different groups depends on the different weights. An interesting phenomenon is that the first three weights are assigned a relatively high weight and the forth is much lower than other three.

For some polygons (e.g. one yellow polygon in the top left in figure 4-10 left), it can be group to yellow group with middle and bottom polygons. It is also reasonable that it was grouped into blue polygons in the figure 4-10 second image. Thus, the advantage of the method is that it produces all possible good results depends on the different weight combinations. The weakness of this method is that it cannot differentiate the small shape difference like “L” and “J” shape located in the left and right bottom. It probably because the polygon area percentage in bounding box was assigned too small weight. A concave corner area feature may provide help when detecting different shapes.

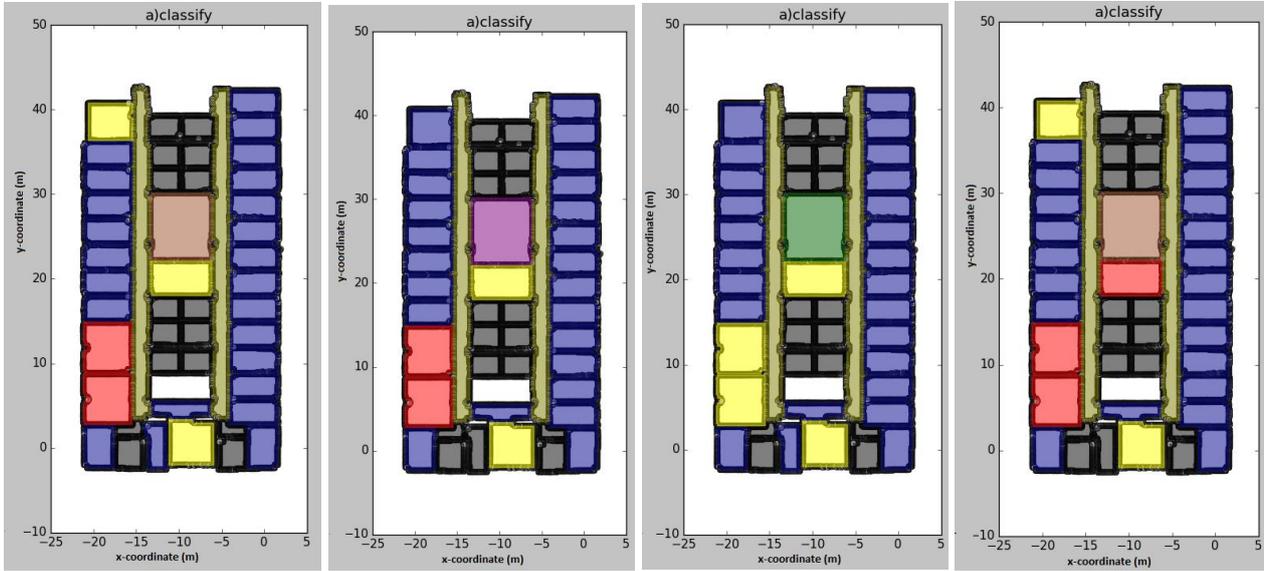


Figure 4-10 Good result of identification of similar polygons by weight in table 4-2

Table 4-2 4 different weight combination

Weight Combination index	Width	Length	Polygon area	Polygon area % in bounding box
7	0.2662336	0.10980497	0.61277759	0.01118384
29	0.33217571	0.06930744	0.58914431	0.00937254
61	0.66367814	0.19210933	0.13172125	0.01249128
94	0.01359538	0.00159885	0.96519845	0.01960732

4.9.2. Hierarchical identification of similar polygons

The second method is using histogram to classify different polygons step by step. To differentiate polygons into different groups by its feature, frequency histogram is an important tool as it analysis the probability distribution of a continuous numerical data (Pearson, 1895). The histogram is constructed by dividing the entire range of values into a series of intervals. The entire range can be calculated by the minimum and maximum value of the entire range. The intervals (bins) are adjacent and they are equal size. Then count the number of values which falls into each interval. The interval can be changed for different dataset. The interval in this research is 1 meter in bounding box width, 1 meter in bounding box length, and 1 square meters in polygon area. The length of rectangle which is erected over the bin is corresponding to the frequency. That indicates how many cases are in each bin. The idea of detecting the similar room in a hierarchical manner is as following:

Hierarchical identifying similar polygons

Input: 3D point clouds

Output: Similar polygons groups

- 1 Classify the polygon by its width firstly;
From the original polygons, each polygon has a specific width. Set a bin, polygons which have the similar width will fall into the same bin. The polygons in the same bin will be classified into the same group. See figure 4-11 b). All polygons have similar width are classified into one group.
- 2 Based on the width classified group, classifying the group inside the width group by its length. That means the group which has the similar width will be classified again by the polygon length;
- 3 Based on the width and length, classifying polygons by the polygon area;
- 4 If more features needed to be added, the next is based on the previous features and it can be done in the same manner.

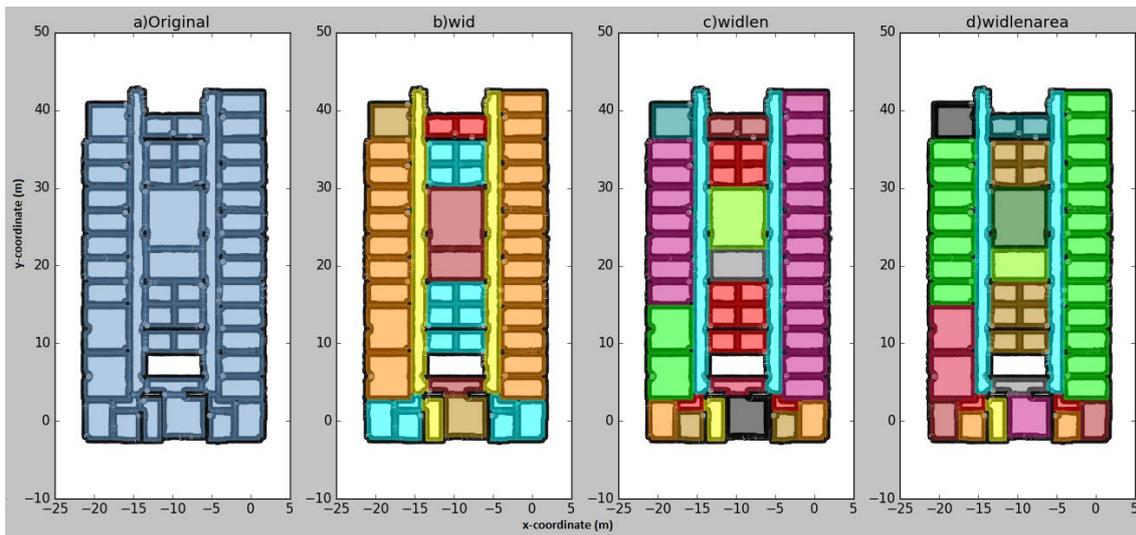


Figure 4-11 Hierarchical identification of similar polygons

In the figure 4-11, figure a) is the original polygons. Figure b) is classified by the bounding box width feature. Polygons with similar width are grouped into same color, same group. Inside the similar width polygons, group the polygons by the polygon bounding box length. Figure c) is result of polygons which classified by width and then length. This step's result is better than the b) result. For example, the two bigger room in the left bottom are differentiated from the orange color in image a) into light green in figure b). Figure d) is classified by bounding box width, then bounding box length and last, polygon area. The red “L” and “J” shape which located in the left and right bottom are differentiated from khaki color polygons in the middle.

Comparing the figure d) result to the ground truth data in figure 4-9, the groups are classified with a 95.83% accuracy. The good aspect of this method is that it produces high accuracy result with a pre-defined features order. However, the poor aspect is that it needs a pre-defined features order.

5. RESULT AND DISCUSSION

5.1. Study area and datasets

5.1.1. 2D floor map data

Two 2D floor plans are used in this research. One is in figure 1-1 right, the other is figure 5-1 left (“April | 2012 | Gideon’s desk...,” 2012). They were searched on google images.

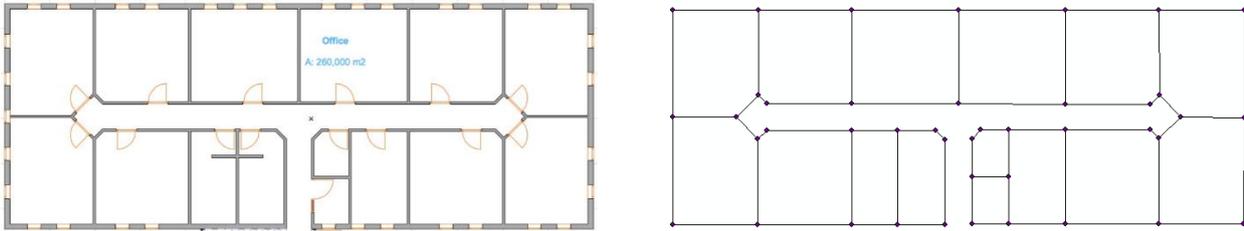


Figure 5-1 Original 2D floor map and its vector data

5.1.2. 3D LASER point clouds data

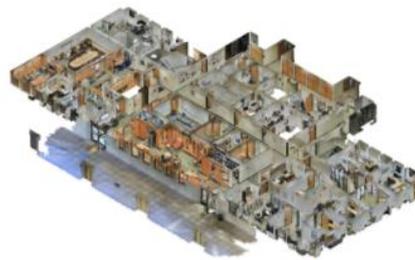
The data is downloaded from Stanford university website (Iro Armeni, Sax, Zamir, & Savarese, 2017). Four 3D experiment datasets were used in this research and they are large-scale indoor parts mainly for educational and office use. The areas show diverse properties in architectural style and appearance. The test area consists of conference rooms, hallway, office, W.C, storage room and staircase, see table 5-1. The building elements in each building area are shown in table 5-2. The reasons for choosing these datasets are as following:

- High accuracy and high point cloud density
- Large scale which has high possibility to explore the regularity inside
- Study areas are the most useful and common places and multiple uses in the daily life. This helps extract the most general regularities which have laser application fields as mentioned in the regularities usage.

Datasets meet the need in both quantitative and qualitative aspects. The 4 3D point clouds datasets are showed in figure 5-1.



Area 1



Area 4



Figure 5-2 3D laser point clouds datasets

Table 5-1 Disjoint space statistics per building area

Name	Area (m ²)	Volume (m ³)	Office	Conf. Room	Lounge	Storage	Hallway	W.C	Staircase	Total Num
Area1	965	2850	32	2	--	2	7	1	1	45
Area4	870	2780	22	3	2	4	12	4	2	49
Area5	1700	5370	42	3	--	4	4	2	--	55
Area6	935	2670	42	2	2	--	7	--	--	53
Total Num	4470	13670	138	10	4	10	29	7	3	202

Table 5-2 Building element statistics per building area

Name	Ceiling	Floor	Wall	Beam	Column	Door	Window	Bookcase	Board
Area1	55	44	233	59	57	86	29	90	27
Area4	68	47	286	3	34	103	32	82	10
Area5	73	65	325	3	72	127	51	209	42
Area6	63	49	246	67	54	93	31	89	29
Total Num	389	284	1533	188	249	539	159	561	137

5.2. Implementation

2D floor map are vectorised manually in ArcMap 10.4.1.

3D point clouds visualization is displayed in Cloud Compare v2.6.2.

Raw point clouds are down sampled in Point Cloud Mapper (PCM). Surface growing segmentation is using the function provided by PCM.

Edges and polygon points are extracted from Matlab R2015a.

Symmetry detection and repetitive pattern detection are executed using python in JetBrains PyCharm Edu 2.0.2.

5.3. Symmetry result and discussion

5.3.1. Building boundary symmetry detection result

In figure 5-3, the image a) is the overlap of both original (blue) and mirrored polygon (grey) based on the line of symmetry of 45 degrees (blue line). The intersection area of the two boundary polygons are shown in image b) in green region. The difference area of the two boundary polygons are shown in image c) green region.

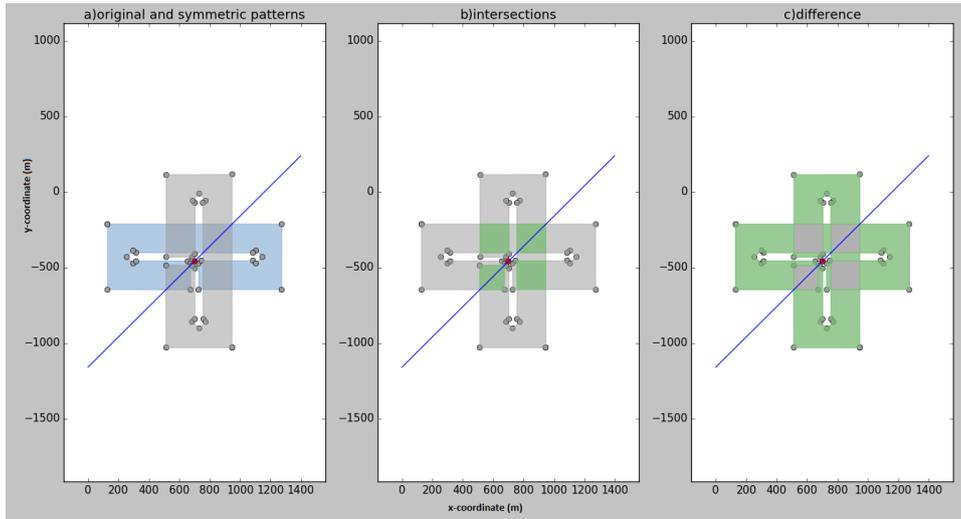


Figure 5-3 Building symmetry detection

For the interior symmetry detection, more than one lines of symmetry exist in the floor. For instance, in figure 5-4, the angle of the line of symmetry is 90 degrees. However, the angle of the line of symmetry in figure 5-5 is 0 degrees. When the angle of the line of symmetry is 90 degrees, more rooms are remaining comparing to the angle of 0 degrees of the line of symmetry.

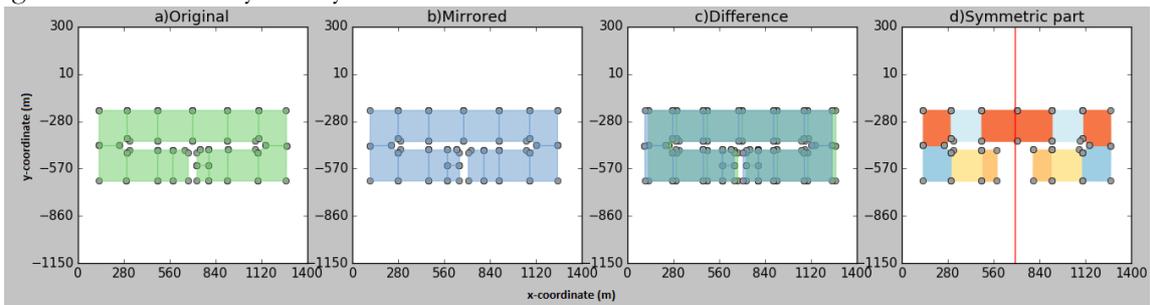


Figure 5-4 Interior symmetry with the line of symmetry in 90 degrees

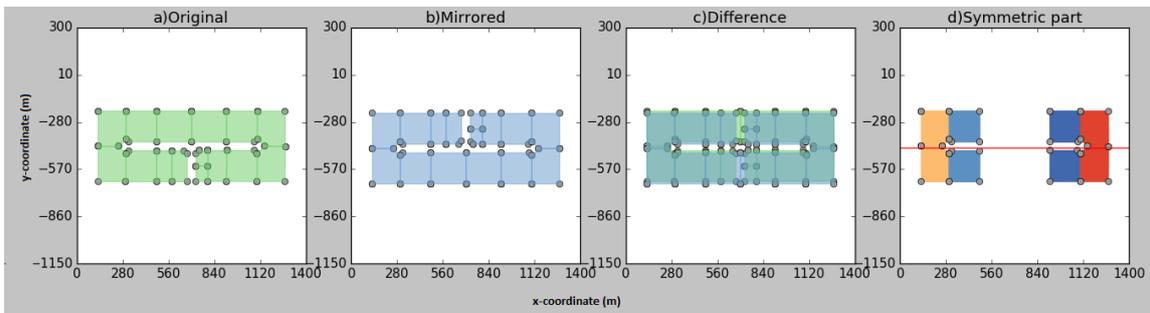


Figure 5-5 Interior symmetry with the line of symmetry in 0 degrees

5.3.2. Polygons symmetry detection result

Dataset 5 is the most complicated dataset, here it is used to test the polygon symmetry detection method. Figure 5-6 a) is the original polygons in dataset 5. Figure 5-6 b) is the polygon symmetric level. Most of colour are light

red when the line of symmetry is 90 degrees comparing to figure 3-10 middle. Figure 5-6 c) is the symmetric part, small number of symmetric polygons are kept after the polygon symmetry detection. Comparing to figure 3-10 right, less polygons in dataset 5 meet the intersection ratio (0.7). Thus, dataset 5 has less symmetry regularity.

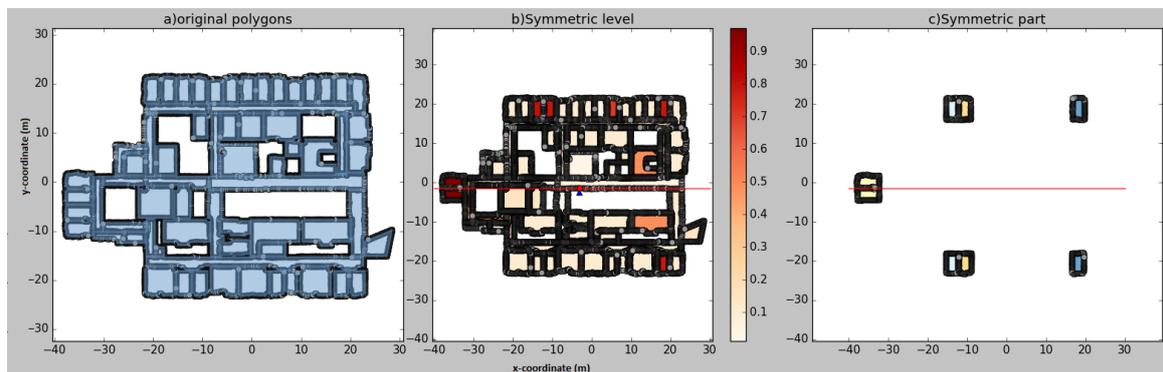


Figure 5-6 Identification of room polygon symmetry for dataset 5

Table 5-3 Parameters of room polygon symmetry detection by area for dataset 5

number of original polygons	67
all polygons area	1663.794
max area	1138.512
draft centroid	(-3.179, -2.361)
optimal centroid	(-3.061, -1.450)
optimal alpha	0

In dataset 5, compare to the room polygons, segment polygons have a more complicated shape and area. From table 5-3 and table 5-4, room polygon number is 67 and the segment polygon number is 119. However, the max symmetric area in room polygon (1138.512 m²) is larger than the segment polygons (917.469 m²). Because the degrees of the line of symmetry are different, the remaining symmetric part are also different. More polygons are symmetric to each other when the line is symmetry is 90 degrees when compare figure 5-6 right and figure 5-7 right.

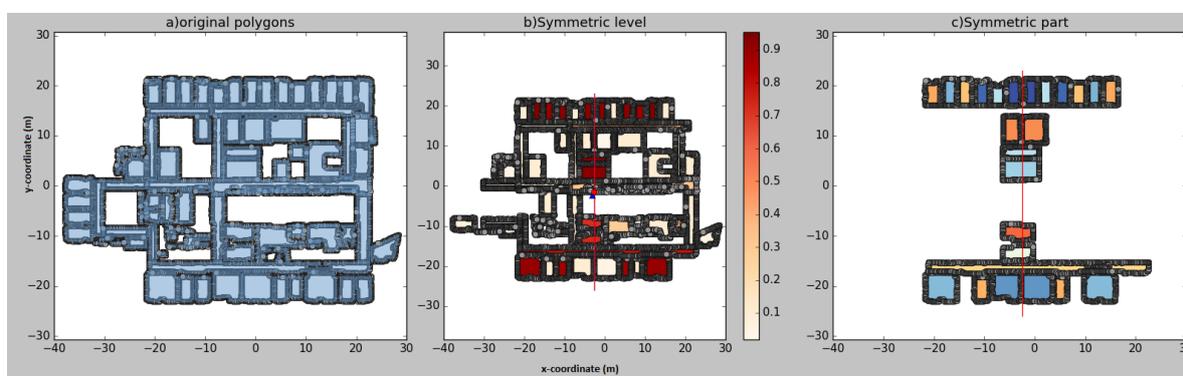


Figure 5-7 Identification of segment polygon symmetry for dataset 5

Table 5-4 Parameters of room polygon symmetry detection by area for dataset 5

number of original polygons	119
all polygon area	1452.305
max area	917.469

draft centroid	(-3.179, -2.361)
optimal centroid	(-2.535, -1.588)
optimal alpha	90

If detecting the polygon symmetry by counting the remaining number of symmetry polygons, in room polygons, the optimal angle is 45 degrees. Because of the line of symmetry, some parts of the dataset 5 on the leftmost and right most have been removed in figure 5-8 d). Only 3 polygons can meet the symmetric requirement in figure 5-8 c).

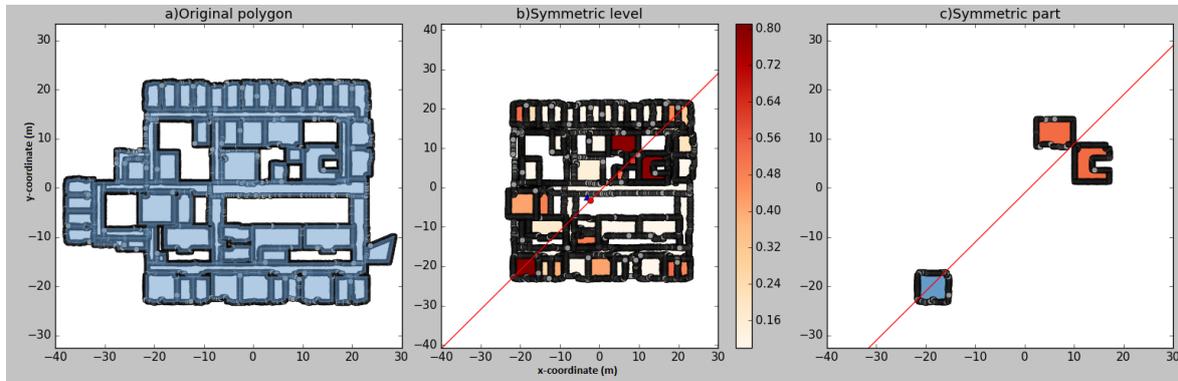


Figure 5-8 Identification of room polygon symmetry for dataset 5

Table 5-5 Parameters of segment polygon symmetry detection by room for dataset 5

number of original polygons	67
number of symmetric polygons	59
draft centroid	(-3.179, -2.361)
optimal centroid	(-2.245, -3.284)
optimal alpha	45

However, in segment polygons, the largest number of polygons (59 out of 119) were kept by the optimal angle of the line of symmetry in 0 degrees in figure 5-9.

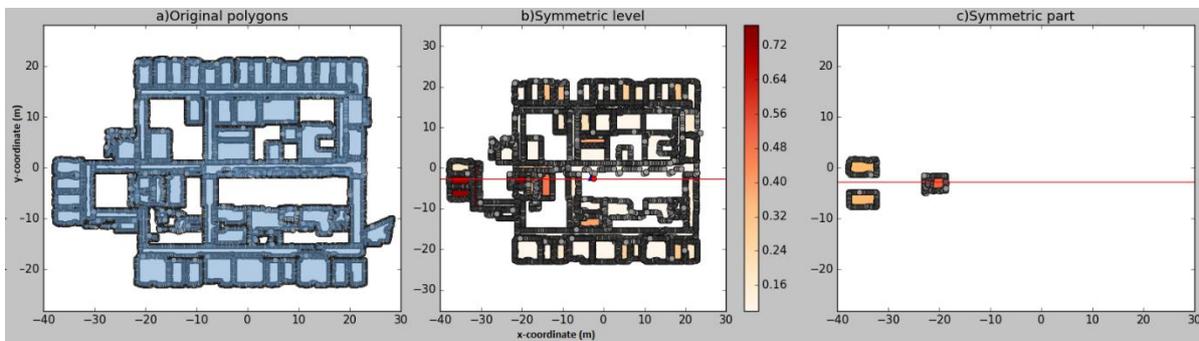


Figure 5-9 Identification of segment polygon symmetry for dataset 5

Table 5-6 Parameters of segment polygon symmetry detection by number for dataset 5

number of original polygons	119
number of symmetric polygons	59
draft centroid	(-3.179, -2.361)

optimal centroid	$(-2.419, -2.654)$
optimal alpha	0

5.3.3. Edges symmetry detection result

3 edges datasets for the edges symmetry detection are shown in figure 5-10.

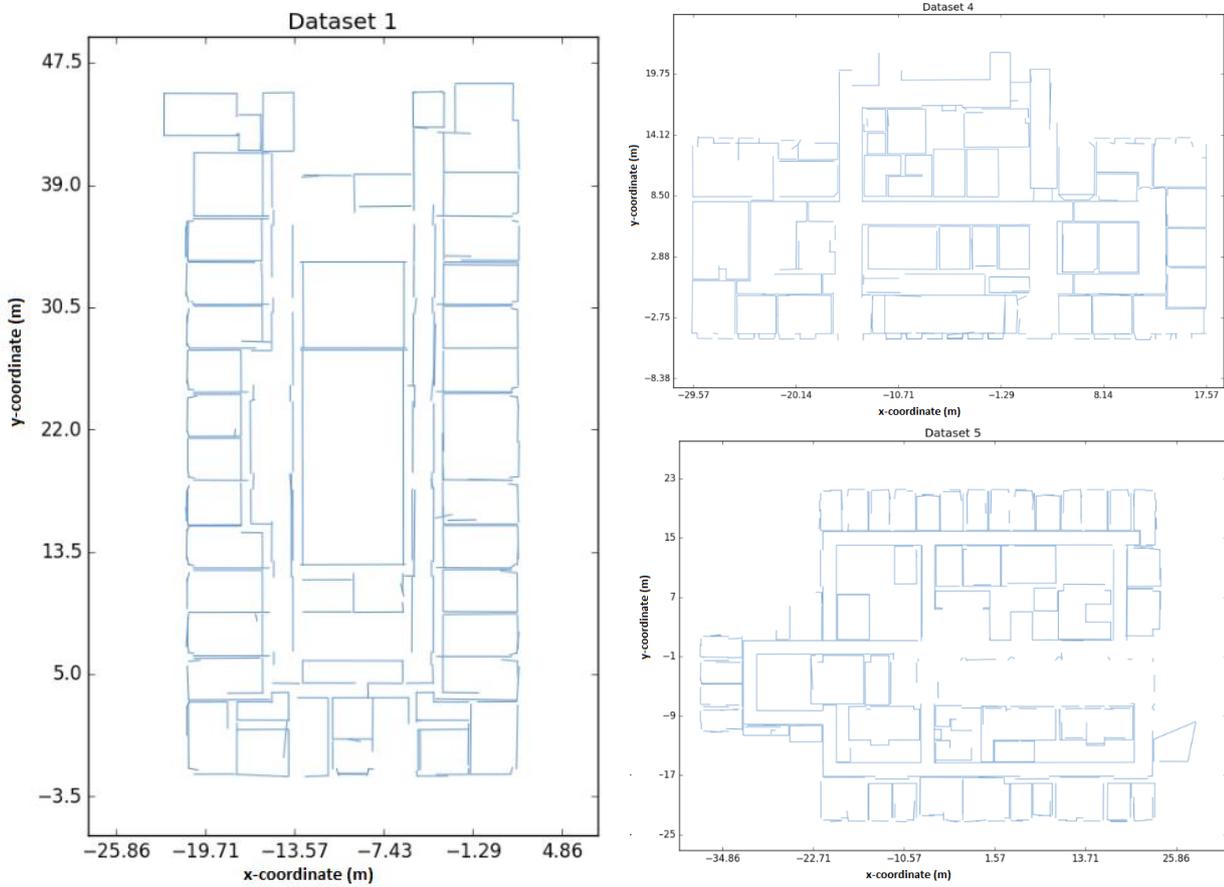


Figure 5-10 Original edges test datasets

In figure 5-11, image a) is the original edges (blue edges) and mirrored edges (magenta edges) which extracted by the draft centroid (blue triangle). Image b) is the original edges (blue edge) and mirrored edges (magenta edges) which extracted by the optimal centroid (red point). In image c), the symmetrical edges are remaining after edges symmetry detection. From the table 5-7, the original edges are 301. After edges symmetry detection, 167 edges were left. The optimal centroid $(-9.563, 20.028)$ is very close to the draft centroid $(-9.979, 19.546)$ and the angle of the line of symmetry is 90 degrees. This method selects 30 centroid candidates and every centroid candidate detects the angle of -89 to 90 degrees (5 degrees interval).

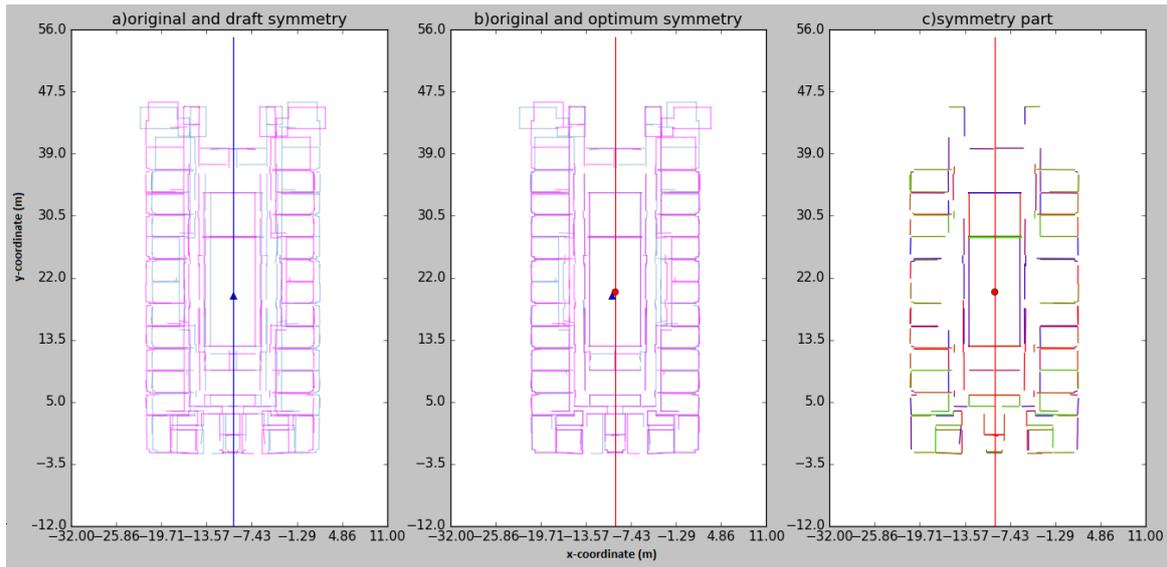


Figure 5-11 Result of edge symmetry detection for dataset 1

Table 5-7 Parameters of edge symmetry detection for dataset 1

original edges	301
optimal number of edges	167
draft centroid	(-9.979, 19.546)
optimal centroid	(-9.563, 20.028)
optimal alpha	90

Figure 5-12 is detecting the edges symmetry from dataset 4. 61 symmetrical edges out of 373 original edges are left after the edges symmetry detection. From figure 5-12 c), not many symmetrical edges are kept.

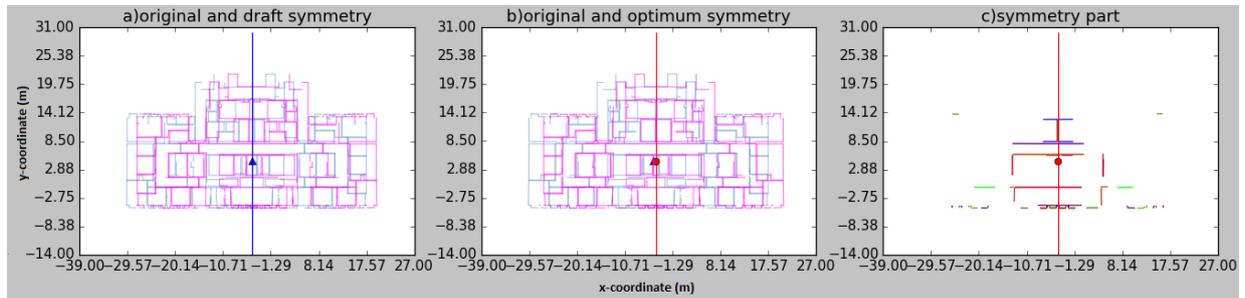


Figure 5-12 Result of edge symmetry detection for dataset 4

Table 5-8 Parameters of edge symmetry detection for dataset 4

original edges	373
optimal number of edges	61
draft centroid	(-5.052, 4.680)
optimal centroid	(-4.564, 4.484)
optimal alpha	90

From figure 5-13 a), dataset 5 has more edges (439 edges) compare to dataset 1 (301 edges) and dataset 4 (373 edges). After the edge symmetry detection, 68 edges are remaining.

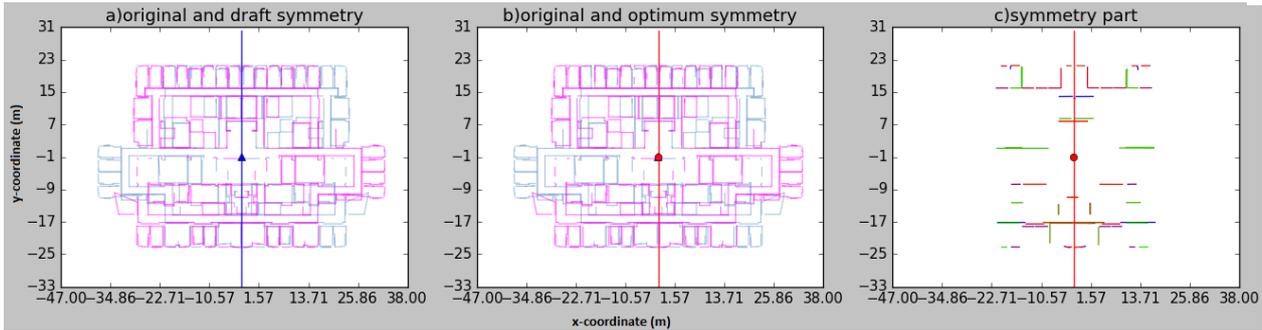


Figure 5-13 Result of edge symmetry detection for dataset 5

Table 5-9 Parameters of edge symmetry detection for dataset 5

original edges	439
optimal number of edges	68
draft centroid	(-2.678, -0.996)
optimal centroid	(-2.446, -1.081)
optimal alpha	90

5.3.4. Symmetry detection discussion

After detecting these three types of symmetry to a variety of datasets, a conclusion is made there.

For all these symmetry detection, the input 3D point clouds datasets are the same. The theory for the detection polygon and edges symmetry is similar. For example, the intersection ratio is used in both detection.

The difference is that for boundary symmetry detection, it returns the boundary symmetry information no matter how is the data quality. Usually the number of the line of symmetry for boundary is more than one. This can be noticed from the plot like figure 3-5.

Polygon symmetry detection detects a 100 % complete dataset and it returns nice interior symmetry information like symmetric polygons, symmetric level and which two rooms are symmetric to each other. However, it needs room information for the segmentation or for extracting the room boundary.

Edges symmetry detection apply the similar theory of polygon symmetry detection. It returns only symmetrical edges (no symmetrical level information). Nevertheless, the advantage of this method is it can be applied to any of the 3D point clouds datasets. Comparing to the boundary symmetry detection, it detects the interior symmetry. Furthermore, it is stable as it will not be influenced by the boundary asymmetry.

Several parameters need to be selected while performing the symmetry detection. All the parameters listed in table 5-10, they are applied for all the four datasets and reasonable results were shown in the above chapters and appendix 1.

Table 5-10 Symmetry detection parameters

Parameters selection:	Boundary symmetry	Polygon symmetry	Edges symmetry
Intersection ratio 1	—	0.7	0.6
Intersection ratio 2	—	0.7	0.6
Intersection ratio range for symmetry level	—	(0.01,1)	—

Difference between intersection ratio 1 & 2	—	0.1	—
The candidate centroids number.	1	30	30
The draft centroid buffer	—	1 meter	1 meter
Angle of the line of symmetry	(-89,90,1)	(-89,90,5)	(-89,90,5)

5.4. Repetitive patterns result and discussion

5.4.1. Identification of similar room by weight result

Figure 5-14 is the ground truth for dataset 5. Figure 5-15 are the results of identification of similar room by weight. The row of polygons in the top and left side are classified into one group in black colour when weight combination indexes are 310 and 699 in figure 5-15. The difference is the big rectangle shape polygon are grouped into yellow class on the left image and they grouped into blue class on the right image. Different index number has different weight combination. 1000 weights combinations are randomly chosen in the research. Because the smaller number of weight combinations causes unstable result and the large number of weight combination has the similar result as 1000 times. The weight combination for index 310 and index 699 are shown in table 5-11.

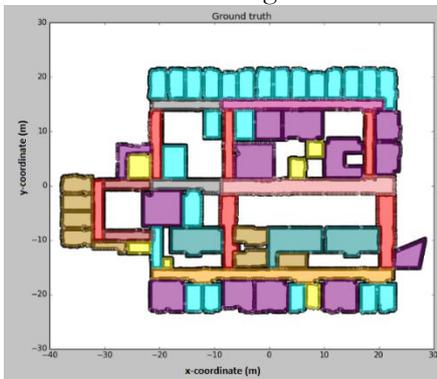


Figure 5-14 Ground truth of dataset 5

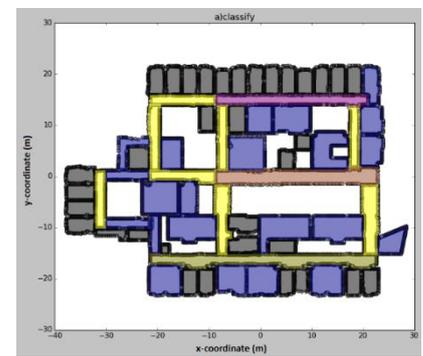
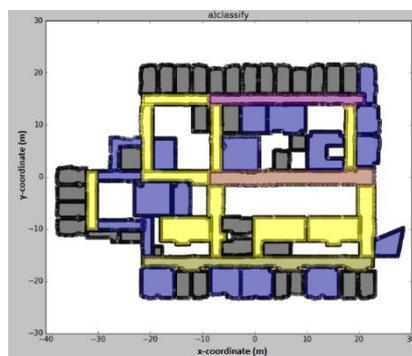


Figure 5-15 Identification of similar room by weight

Table 5-11 Different weight combination

n	Width	Length	Polygon area	Polygon area % in bounding box
310	0.56470608	0.0771415	0.09253949	0.26561293
699	0.6131038	0.16334411	0.0106081	0.21294399

From table 5-11, the width is assigned overwhelming weight comparing to other three features' weights. Higher weights are put on width and length when n is 699. This concentrates more on the shape than the area size. Thus, the three yellow polygons which located in the down row in figure 5-15 left are grouped into blue polygons in figure 5-15 right.

5.4.2. Hierarchical identification result

Dataset 4 have a variety of polygon shapes and sizes. From figure 5-16 b), polygons which has similar width are grouped in one class. At this point, the fire brick colour, pink colour, and cyan colour are the three main groups. When considering length, in figure 5-16 c), multiple groups appeared and it is over grouped. Several similar polygons are grouped into different classes (e.g. fire brick colour polygon in the rightmost is grouped different from its up neighbour and down neighbour in pink colour). Figure 5-16 is even worse over group. This is because the interval of width (1 m), length (1 m), and area (1 m²) is too small for the variety of the shape and size of the

polygons. In this situation, increase the interval of the features will help the group results. For more test dataset, please go to appendix 2.

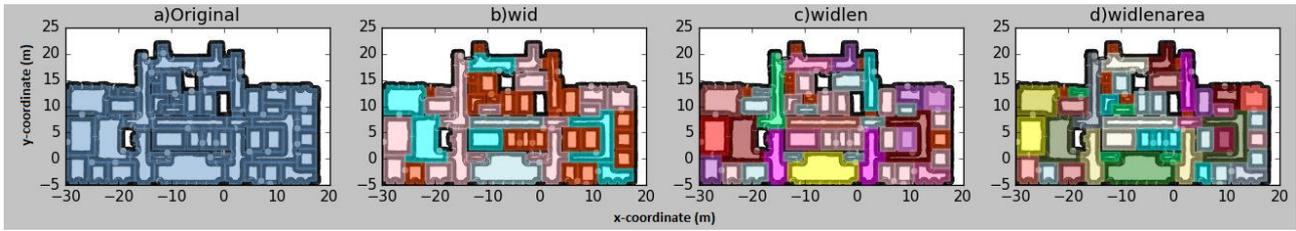


Figure 5-16 Hierarchical identification of similar polygons for dataset 4

5.4.3. Identification of similar polygons discussion

For identification of similar polygons by weight, if the ground truth data is provided, this method returns all the good result automatically. And randomly assigning a feature a weight doesn't need the prior experience which returns all possible good results. The drawback is that if there is no ground truth data, this method will return both good results and bad results. Then, a manual differentiation is needed for selecting the good result from bad. The weight which assigned to each feature has an important influence on the result. Another necessary factor which influences the result is the number of features. This research select 4 most prominent features to detect the similar room. If more features are selected, the results may have a little bit improvement.

For hierarchical identification of similar polygons, it usually returns high accuracy result compare to the result from the first method. Furthermore, the number of groups can be adjusted according to the room complexity. Because usually simple polygons types result in smaller groups than the complex polygons types. This method has two disadvantages. One is that it has to pre-define the features` order. That means a which one should be on the top of the hierarchical level and which feature should be the next. Another shortcoming is that it is easy to get over group. If the interval was set too small, the similar polygons may be grouped into different groups.

Table 5-12 is the comparison of two approaches for identifying similar polygons.

Table 5-12 Advantage and disadvantage of two method of identification of similar polygons

Method	Identification of similar polygons by weight result	Hierarchical identification of similar polygons
Advantages	<ul style="list-style-type: none"> Multiple good results 	<ul style="list-style-type: none"> High accuracy Change group range
Disadvantages	<ul style="list-style-type: none"> Need ground truth 	<ul style="list-style-type: none"> Pre-defined feature order Easy to over group

6. CONCLUSION AND RECOMMENDATIONS

6.1. Conclusions

The main objective of this research is to describe and detect the regularity in the indoor scenes. Symmetry and repetitive patterns are the two main regularities detected in this research.

For symmetry, we detect the boundary symmetry, edges symmetry and polygon symmetry.

- Boundary symmetry returns a general symmetry of the building external boundary.
- Edges symmetry detection detects whether two or more edges are symmetrical to each other.
- Polygon symmetry detection. Both segment polygon and room polygon are extracted and two formats of the criteria for selecting good symmetry results.

The intersection ratio was used in the polygon symmetry detection and edges symmetry detection. The intersection ratio for detecting symmetric polygon is 0.7. If two polygons' intersection ratio are both higher than 0.7, it will be seen as symmetric to each other. The intersection ratio for detecting symmetric edges is 0.6. If two edges' intersection ratio are both higher than 0.6, then they are symmetric to each other in this research.

For the repetitive pattern, 11 repetitive features were explored in this research. Following two approaches are designed to find the similar polygons.

- For identification of the similar polygons by weight: Weights should be normalized from 0 to 1 and the weights combination should sum to one. Sufficient number of random weight combinations are selected for the best results.
- For hierarchical identification of the similar polygons: features are ordered to detect the similar polygons step by step. The feature order tested in this research is bounding box width first, then bounding box length, and last polygon area. The intervals for them are 1 m, 1 m, and 1 m² separately.

6.2. Answer to research questions

1. How can regularities be described and expressed in a concise and comprehensive way from 3D point clouds of indoor scenes?

The symmetry and repetitive patterns from 3D point clouds of indoor scenes are the regularities which described and detected in this research. They are described in a mathematical way in corresponding chapters and can be visualized by colours.

2. Because of heavy occlusion and varying point density, how to improve the identification of the potential regularities?

For each specific regularity, the way of dealing with occlusion and variable point density is different. For instance, we pre-process the polygon intersection when detecting polygon's boundary intersection. We extracting the building boundary for calculating the draft centroid which avoid the variable point density.

3. How can those regularities be detected?

We detect different regularities based on their properties. For example, if it is going to detect whether two objects are symmetric to each other. Try to find the optimal centroid and the optimal angle of the line of symmetry first, and then find out their symmetrical polygons. If it is going to detect the similar polygons, features like bounding box width, bounding box length and polygon area should be extracted.

4. How to evaluate the results?

The symmetry results are visualized by proportional colour. The repetitive features are displayed in the different graphs like a bar chart, histogram. The rooms with similar shape are grouped in same class with same colour. The result will be compared with ground truth data. If no ground truth data was provided, a manual visual is necessary.

5. How to apply regularities for improvement of 3D reconstruction?

For the detected symmetric polygons, it can be used to fill the data gap which caused by occlusion. For the repetitive features like polygon centroid, they can help to compress the big 3D point clouds. For the identified similar polygons, they provide a general ideal of how many different room types per floor.

6.3. Recommendation

By extracting the repetitive features, there may have a simpler way to store the dense 3D point clouds. Like storing the polygons` centroid and bounding box and segment length, a 3D dataset will be stored in a smaller size.

In the method of “Identification of similar room by weight”, only 4 features are included. To improve the result, more weights should be considered.

Hierarchical identification of similar polygons

In this research, the hierarchical level is polygon width first, then length, under the width and length is area. If we change the order of the hierarchical, for example, putting area first and then width and length, then the group result may be different. A more flexible way of detecting the hierarchical features is by performing all the orders combinations and select the best results.

LIST OF REFERENCES

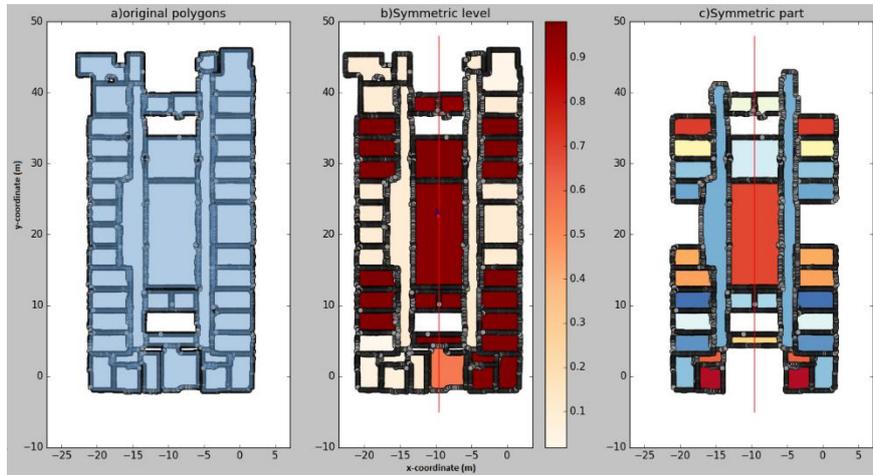
- April | 2012 | Gideon's desk... (2012). Retrieved February 12, 2017, from <https://gtaljaard.wordpress.com/2012/04/>
- Armeni, I., Sax, S., Zamir, A. R., & Savarese, S. (2017). Joint 2D-3D-Semantic Data for Indoor Scene Understanding. Retrieved from https://www.mendeley.com/research-papers/joint-2d3dsemantic-data-indoor-scene-understanding/?userDocumentId=%7Bcaebcbea-169f-4b9a-9ffa-4da66307ea32%7D&utm_campaign=open_catalog&utm_medium=1.16.2&utm_source=desktop
- Armeni, I., Sener, O., Zamir, A., & Jiang, H. (2016). 3D Semantic Parsing of Large-Scale Indoor Spaces. *Buildingparser.Stanford.Edu*. Retrieved from http://buildingparser.stanford.edu/images/3D_Semantic_Parsing.pdf
- Axial and central symmetry. (2016). Retrieved November 1, 2016, from <http://www.sangakoo.com/en/unit/axial-and-central-symmetry>
- Becker, S., Peter, M., & Fritsch, D. (2015). Grammar-Supported 3D Indoor Reconstruction From Point Clouds for “As-Built” Bim. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-3/W4*(July 2016), 17–24. <http://doi.org/10.5194/isprsannals-II-3-W4-17-2015>
- Budroni, A. (2013). *Angela Budroni Automatic Model Reconstruction of Indoor Manhattan-World Scenes from Dense Laser Range Data*. Retrieved from <http://www.ifp.uni-stuttgart.de/publications/dissertationen/c-715.pdf>
- Cruz, U. C. S., Kim, C., & Manduchi, R. (2014). Previously Published Works UC Santa Cruz Author : Planar Structures from Line Correspondences in a Manhattan World, 0–16.
- Durrant-Whyte, H., & Bailey, T. (2006). Simultaneous localization and mapping: part I. *IEEE Robotics & Automation Magazine, 13*(2), 99–110. <http://doi.org/10.1109/MRA.2006.1638022>
- Friedman, S., & Stamos, I. (2013). Online detection of repeated structures in point clouds of urban scenes for compression and registration. *International Journal of Computer Vision, 102*(1–3), 112–128. <http://doi.org/10.1007/s11263-012-0575-y>
- Furukawa, Y., Curless, B., Seitz, S. M., & Szeliski, R. (2009). Manhattan-world stereo. *IEEE International Conference on Computer Vision and Pattern Recognition, 1422–1429*. <http://doi.org/10.1109/CVPR.2009.5206867>
- Geometry - python: Help to implement an algorithm to find the minimum-area-rectangle for given points in order to compute the major and minor axis length - Stack Overflow. (2012). Retrieved January 15, 2017, from <http://stackoverflow.com/questions/13542855/python-help-to-implement-an-algorithm-to-find-the-minimum-area-rectangle-for-gi>
- Golovinskiy, A., Podolak, J., & Funkhouser, T. (2009). Symmetry-aware mesh processing. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5654 LNCS*, 170–188. http://doi.org/10.1007/978-3-642-03596-8_10
- Hsu, J. T., Li-Chang Liu, & Li, C. C. (2001). Determination of structure component in image texture using wavelet analysis. In *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)* (Vol. 2, pp. 166–169). IEEE. <http://doi.org/10.1109/ICIP.2001.958077>
- Ikehata, S., Yang, H., & Furukawa, Y. (2015). Structured Indoor Modeling. *Proceedings of the IEEE International Conference on Computer Vision, 1323–1331*. <http://doi.org/10.1109/ICCV.2015.156>
- Jiang, X. Y., & Bunke, H. (1991). Determination of the symmetries of polyhedra and an application to object recognition (pp. 113–121). Springer Berlin Heidelberg. http://doi.org/10.1007/3-540-54891-2_9
- Juszczak, P., Tax, D. M. J., & Dui, R. P. W. (2002). Feature scaling in support vector data descriptions. *Proc. 8th Annu. Conf. Adv. School Comput. Imaging*, 95–10. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.100.2524>
- Kazhdan, M., Chazelle, B., Dobkin, D., Funkhouser, T., & Rusinkiewicz, S. (2003). A reflective symmetry descriptor for 3D models. *Algorithmica (New York), 38*(1), 201–225. <http://doi.org/10.1007/s00453-003-1050-5>
- Kim, Y. M., Mitra, N. J., Yan, D., & Guibas, L. (2012). Acquiring 3D indoor environments with variability and repetition. *ACM Transactions on Graphics, 31*(6), 1. <http://doi.org/10.1145/2366145.2366157>
- Leica Pegasus: Backpack | C.R.Kennedy Survey Division. (2016). Retrieved November 24, 2016, from

- <http://survey.crkennedy.com.au/products/3d-laser-scanning/leica-hds-3d-laser-scanners/leica-pegasus-backpack>
- M3 Trolley - NavVis. (2016). Retrieved October 31, 2016, from <http://www.navvis.com/products/m3-trolley/>
- Mattausch, O., Panozzo, D., Mura, C., Sorkine-Hornung, O., & Pajarola, R. (2014). Object detection and classification from large-scale cluttered indoor scans. *Computer Graphics Forum*, 33(2), 11–21. <http://doi.org/10.1111/cgf.12286>
- Methods for data standardization. (2014). Retrieved January 24, 2017, from https://www.biomedware.com/files/documentation/boundaryseer/Preparing_data/Methods_for_data_standardization.htm
- Mitra, N. J., Guibas, L. J., & Pauly, M. (2007). Symmetrization. *ACM Transactions on Graphics*, 26(3), 63. <http://doi.org/10.1145/1276377.1276456>
- Mitra, N. J., Guibas, L. J., Pauly, M., Mitra, N. J., Guibas, L. J., & Pauly, M. (2006). Partial and approximate symmetry detection for 3D geometry. *ACM SIGGRAPH 2006 Papers on - SIGGRAPH '06*, 25(3), 560. <http://doi.org/10.1145/1179352.1141924>
- O'Rourke, J. (1985). Finding minimal enclosing boxes. *International Journal of Computer & Information Sciences*, 14(3), 183–199. <http://doi.org/10.1007/BF00991005>
- Ochmann, S., Vock, R., Wessel, R., & Klein, R. (2016). Automatic reconstruction of parametric building models from indoor point clouds. *Computers & Graphics*, 54, 94–103. <http://doi.org/10.1016/j.cag.2015.07.008>
- Ochmann, S., Vock, R., Wessel, R., Tamke, M., & Klein, R. (2014). Automatic Generation of Structural Building Descriptions from 3D Point Cloud Scans. *Proceedings of GRAPP 2014 - International Conference on Computer Graphics Theory and Applications, January*. Retrieved from http://cg.cs.uni-bonn.de/aigaion2root/attachments/GRAPP_2014_54_CR.pdf
- Oesau, S., Lafarge, F., & Alliez, P. (2014). Indoor scene reconstruction using feature sensitive primitive extraction and graph-cut. *ISPRS Journal of Photogrammetry and Remote Sensing*, 90, 68–82. <http://doi.org/10.1016/j.isprsjprs.2014.02.004>
- Okorn, B., Xiong, X., Akinci, B., & Huber, D. (2010). Toward Automated Modeling of Floor Plans. *Proceedings of the Symposium on 3D Data Processing, Visualization and Transmission*, 2. Retrieved from [http://swing.adm.ri.cmu.edu/pub_files/2010/5/2009_3DPVT_plan_view_modeling_v13_\(resubmitted\).pdf](http://swing.adm.ri.cmu.edu/pub_files/2010/5/2009_3DPVT_plan_view_modeling_v13_(resubmitted).pdf)
- Pauly, M., Mitra, N. J., Wallner, J., Pottmann, H., & Guibas, L. J. (2008). Discovering Structural Regularity in 3D Geometry. *ACM Transactions on Graphics*, 27(3), 1. <http://doi.org/10.1145/1360612.1360642>
- Pearson, K. (1895). Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 186(0), 343–414. <http://doi.org/10.1098/rsta.1895.0010>
- Petitjean, M. (2007). A definition of symmetry. *Symmetry: Culture and Science*, 18(2–3), 99–119. Retrieved from <http://petitjeanmichel.free.fr/itoweb.paper.SCS.2007.petitjean.pdf>
- Previtali, M., Scaioni, M., Barazzetti, L., Brumana, R., & Roncoroni, F. (2013). Automated Detection of Repeated Structures in Building Facades. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5/W2(November), 241–246. <http://doi.org/10.5194/isprsannals-II-5-W2-241-2013>
- Principle of Repetition; Pattern | Visual Communication Design. (2013). Retrieved March 1, 2017, from <https://visscom.wordpress.com/2013/04/16/principle-of-repetition-pattern/>
- Simari, P., Kalogerakis, E., & Singh, K. (2006). Folding meshes: hierarchical mesh segmentation based on planar symmetry. *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, 111–119. Retrieved from <http://portal.acm.org/citation.cfm?id=1281972> \nhttp://www.cs.jhu.edu/~psimari/publications/2006_sg_p_simari_kalogerakis_singh.htm
- Sithole, G., & Zlatanova, S. (2016). Position, Location, Place and Area: an Indoor Perspective. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 3(4). <http://doi.org/10.5194/isprsannals-III-4-89-2016>
- Spinello, L., Triebel, R., Vasquez, D., Arras, K. O., & Siegwart, R. (2010). Exploiting repetitive object patterns for model compression and completion. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6315 LNCS(PART 5), 296–309. http://doi.org/10.1007/978-3-642-15555-0_22

- Symmetry. (2016). Retrieved November 1, 2016, from <https://en.wikipedia.org/wiki/Symmetry>
- Symmetry Matching. (2017). Retrieved January 31, 2017, from <http://www.topmarks.co.uk/symmetry/symmetry-matching>
- Thomson, C., Apostolopoulos, G., Backes, D., & Boehm, J. (2013). Mobile Laser Scanning for Indoor Modelling. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-5/W2*(November), 289–293. <http://doi.org/10.5194/isprsannals-II-5-W2-289-2013>
- Torii, A., Sivic, J., Okutomi, M., & Pajdla, T. (2015). Visual Place Recognition with Repetitive Structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(11), 2346–2359. <http://doi.org/10.1109/TPAMI.2015.2409868>
- Vosselman, G., Gorte, B., Sithole, G., & Rabbani, T. (2004). Recognising Structure in Laser Scanner Point Clouds. *Information Sciences*, 46(April 2016), 1–6. <http://doi.org/10.1002/bip.360320508>
- Vosselman, G., & Maas, H.-G. (2010). *Airborne and terrestrial laser scanning*. Dunbeath: Whittles Publishing.
- Xiao, J., & Furukawa, Y. (2014). Reconstructing the World's Museums. *International Journal of Computer Vision*, 110(3), 243–258. <http://doi.org/10.1007/s11263-014-0711-y>
- Zabrodsky, H., Peleg, S., & Avnir, D. (1997). Comments on “symmetry as a continuous feature.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(3), 246–247. <http://doi.org/10.1109/34.584101>
- ZEB1 - 3D Laser Mapping. (2016). Retrieved October 31, 2016, from <http://www.3dlasermapping.com/zeb1/>

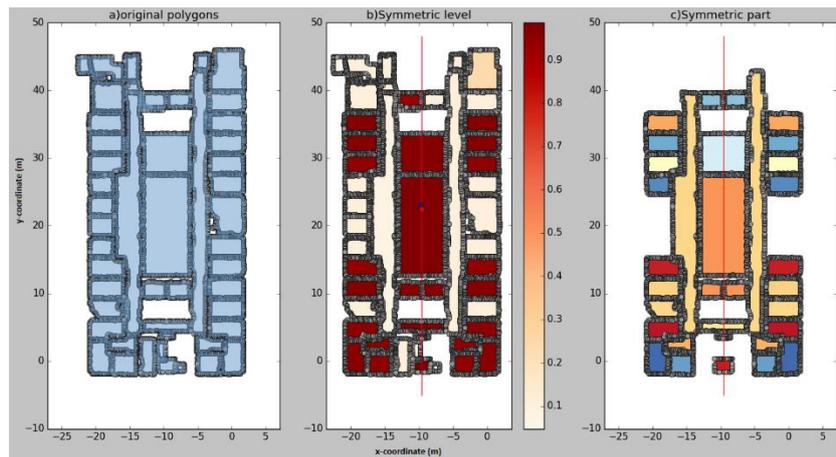
APPENDIX

Appendix 1: Polygon symmetry detection for dataset 1 and dataset 4



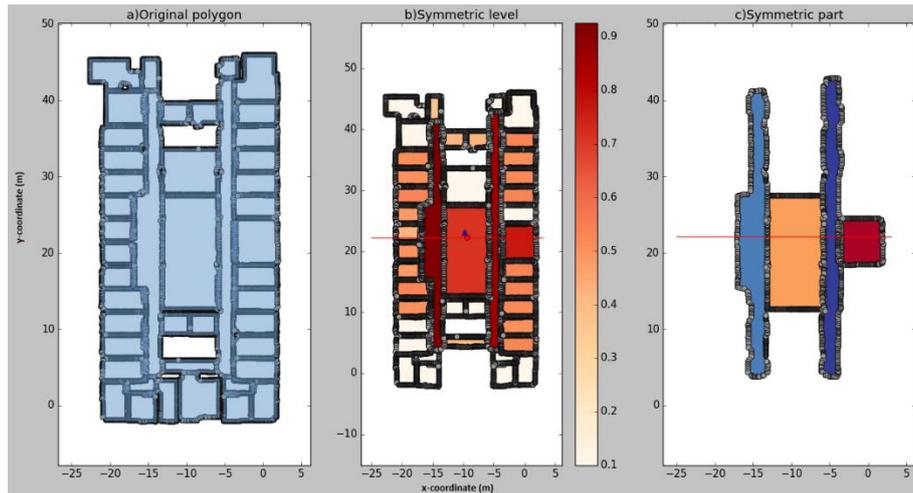
Identification of room polygon symmetry by area for dataset 1

number of original polygons	44
all polygons area	959.461
Max area	932.257
draft centroid	(-9.775, 23.123)
optimal centroid	(-9.609, 22.544)
optimal alpha	90



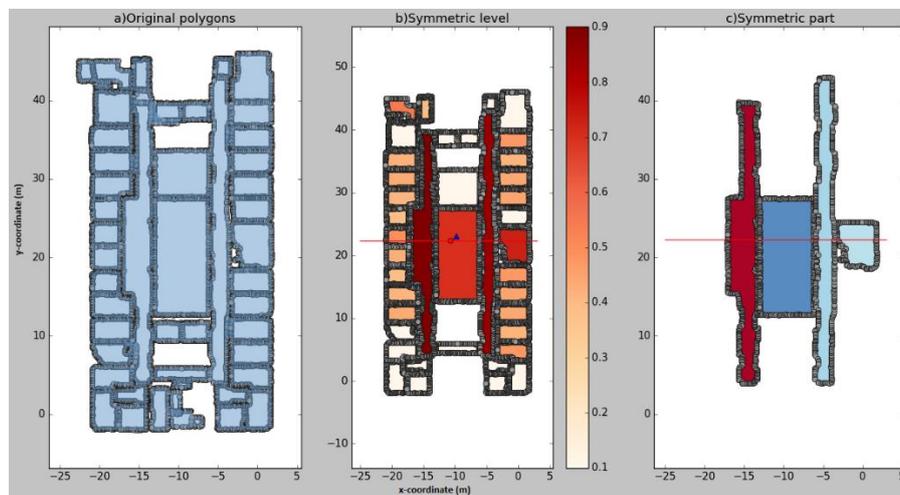
Identification of segment polygon symmetry by area for dataset 1

number of original polygons	56
all polygons area	830.416
max area	747.546
draft centroid	(-9.775, 23.123)
optimal centroid	(-9.622, 22.434)
optimal alpha	90



Identification of room polygon symmetry by number for dataset 1

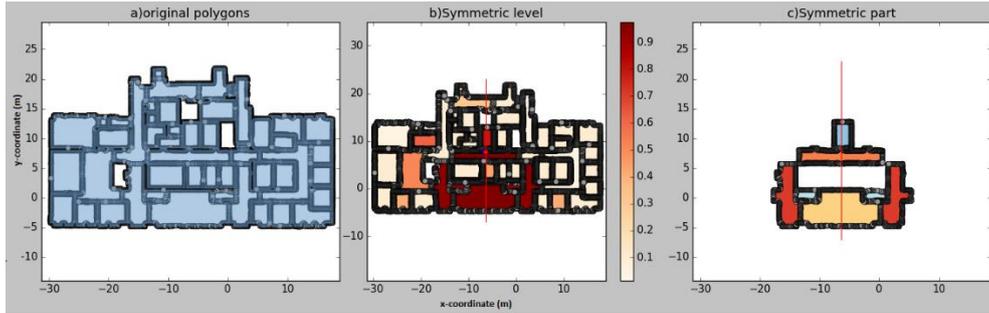
number of original polygons	44
number of symmetric polygons	40
optimal centroid	(-9.371, 22.171)
draft centroid	(-9.775, 23.123)
optimal alpha	0



Identification of segment polygon symmetry by number for dataset 1

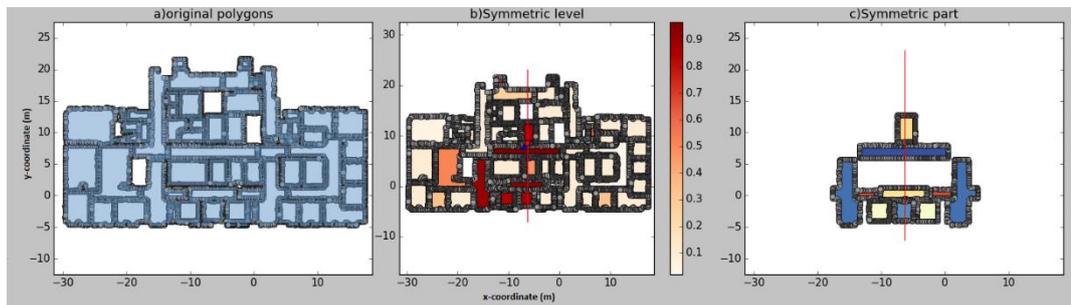
number of original polygons	56
number of symmetric polygons	39
optimal centroid	(-10.678, 22.264)
draft centroid	(-9.775, 23.123)
optimal alpha	0

Dataset 4



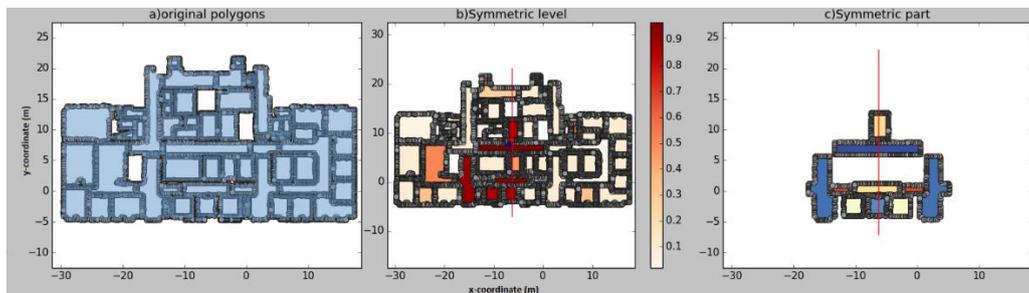
Identification of room polygon symmetry by area for dataset 4

all polygons area	885.264
Max area	805.009
draft centroid	(-7.016, 7.885)
optimal centroid	(-6.345, 7.673)
optimal alpha	90



Identification of segment polygon symmetry by area for dataset 4

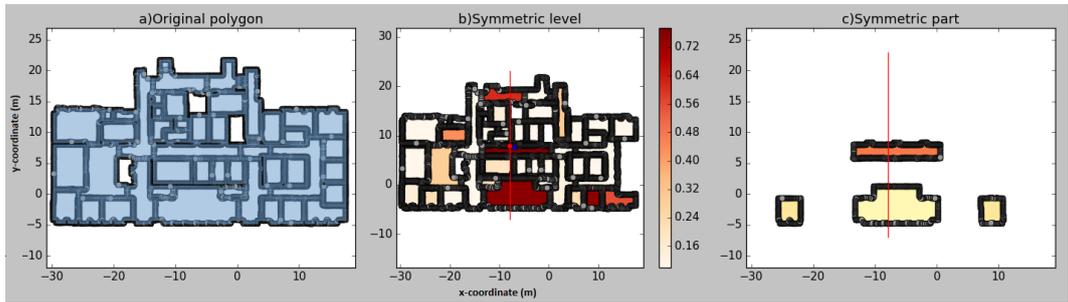
number of original polygons	83
all polygons area	797.461
Max area	681.606
draft centroid	(-7.0169, 7.885)
optimal centroid	(-6.253, 8.550)
optimal alpha	90



Identification of room polygon symmetry by number for dataset 4

number of original polygons	49
number of symmetric polygons	43
optimal centroid	(-7.822, 7.827)
draft centroid	(-7.016, 7.885)

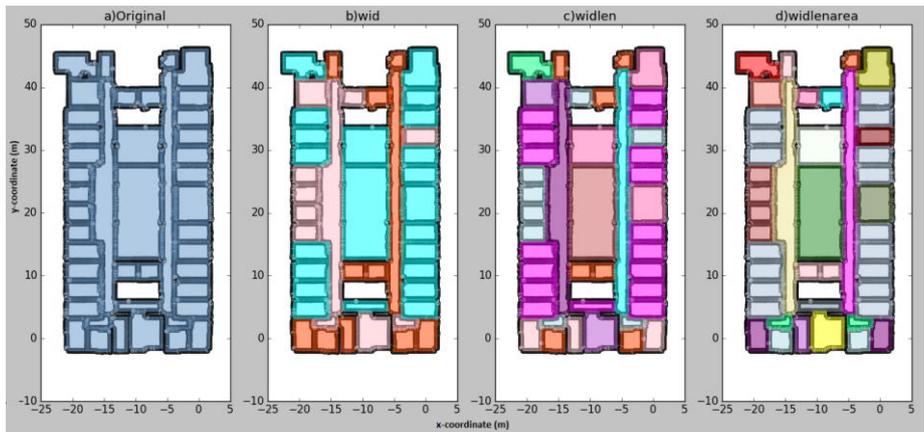
optimal alpha	90
---------------	----



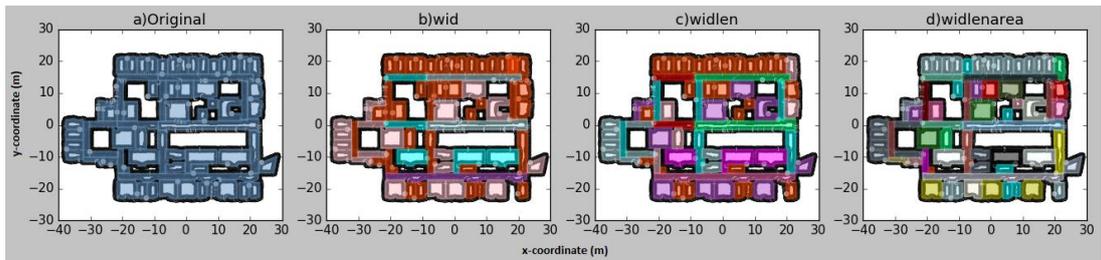
Identification of segment polygon symmetry by number for dataset 4

number of original polygons	83
number of symmetric polygons	52
optimal centroid	(-7.029, 8.345)
draft centroid	(-7.016, 7.885)
optimal alpha	90

Appendix 2: Hierarchical identification of similar polygons for dataset 1 and dataset 5



Hierarchical identification of similar polygons for dataset 1



Hierarchical identification of similar polygons for dataset 5