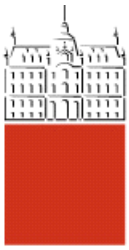
Univerza
v Ljubljani

Fakulteta za
gradbeništvo in
geodezijo

Jamova 2
1000 Ljubljana, Slovenija
telefon (01) 47 68 500
faks (01) 42 50 681
fgg@fgg.uni-lj.si



UNIVERZITETNI ŠTUDIJ
GEODEZIJE
SMER GEODEZIJA

Kandidatka:

URŠA HRIBERNIK

OBOGATITEV 3D MODELOV STAVB Z OKNI IZ PODATKOV AEROLASERSKEGA SKENIRANJA POŠEVNEGA POGLEDA IN FASADNIH TEKSTUR

Diplomska naloga št.: 926/G

ENRICHMENT OF A 3D BUILDING MODEL WITH WINDOWS USING OBLIQUE-VIEW ALS AND FAÇADE TEXTURES

Graduation thesis No.: 926/G

Mentorica:
doc. dr. Mojca Kosmatin Fras

Somentor:
prof. dr. Uwe Stilla
Sebastian Tuttas

Član komisije:
doc. dr. Simona Savšek
izr. prof. dr. Tomaž Ambrožič

Predsednik komisije:
izr. prof. dr. Dušan Kogoj

Ljubljana, 28. 05. 2013

Hribernik, U. 2013. Enrichment of a 3D building model with windows…
Grad. Th. - University Studies. Ljubljana, UL FGG, Department of Geodetic Eng.

Errata

Page	Line	Error	Correction

Ta stran je namenoma prazna.

IZJAVA

Podpisana Urša Hribernik izjavljam, da sem avtorica diplomske naloge z naslovom »Obogatitev 3D modelov stavb z okni iz podatkov aerolaserskega skeniranja poševnega pogleda in fasadnih tekstur «. Izjavljam, da je elektronska različica v vsem enaka tiskani različici. Izjavljam, da dovoljujem objavo elektronske različice v repozitoriju UL FGG.

Ljubljana, 9.5.2013 Urša Hribernik

Ta stran je namenoma prazna.

Bibliographic information and abstract

UDC: 528.7:692.23(043.2)

Author: Urša Hribernik

Supervisor: Assist. Prof. Mojca Kosmatin Fras, Ph.D.

Co-advisors: Sebastian Tuttas, M.Sc.; Prof. Uwe Stilla, Ph.D.

Title: Enrichment of a 3D building model with windows using oblique-view ALS and façade textures

Document type: Graduation thesis – University studies

Notes: 75 p., 7 tab., 52 fig., 8 eq.

Key words: oblique-view ALS, visible façade textures, 3D building model, co-registration, window detection, 3D mapping

Abstract

A wide range of applications using 3D building models exists; such as computer games, city marketing, disaster management, tourist information systems, simulations of noise propagation and surveillance of sustainable construction. Complete acquisition of large urban scenes has become feasible using multi-aspect-oblique-view ALS; however, automated generation of detailed 3D models, the main focus of this thesis, still poses a significant challenge.

To enable enrichment of a 3D building model with windows, the 3D wire-frame building model and the ALS point cloud are first automatically co-registered. The novel approach to window extraction presented in this thesis exploits evidences about window positions in processed oblique-view ALS point cloud and façade image textures. Laser beam penetrates glassy window areas, and thus points found behind a segmented façade plane, projected onto the façade plane, give reliable evidences about intrusion positions. On the other hand, high values of gradient on a texture are usually due to window frames. These two facts are exploited when extracting initial window patches. Additionally, binary masks, obtained by region growing of homogeneous parts of façade textures, are used to eliminate certain façade artefacts and to improve shape of window patches. The assumption, that many windows of the same kind are on the same floor, is used for the refinement procedure. First, façade textures are divided into horizontal blocks, representing floors. Second, a search for non-similar window patch templates within each block is performed. Third, to obtain additional window patch positions, the chosen templates are cross-correlated along the respective block.

Eleven façade planes of an existing 3D wire-frame building model are textured with extracted patches, representing windows and other intrusions. Despite different arrangements of windows, varying window sizes, and relatively strict evaluation method, the method results in 63% detection rate. What is more, the method is mostly data-driven and the detection rate outperforms the method using only oblique-view ALS (Tuttas & Stilla, 2013). The windows are well defined, since the basis for most of window patches are connected components of edges belonging to window frames.

Ta stran je namenoma prazna.

Bibliografsko-dokumentacijska stran in izvleček

UDK: 528.7:692.23(043.2)

Avtorica: Urša Hribernik

Mentor: doc. dr. Mojca Kosmatin Fras

Somentorja: mag. Sebastian Tuttas; prof. dr. Uwe Stilla

Naslov: Obogatitev 3D modelov stavb z okni iz podatkov aerolaserskega skeniranja poševnega pogleda in fasadnih tekstur

Tip dokumenta: diplomska naloga – univerzitetni študij

Obseg in oprema: 75 str., 7 pregl., 52 sl., 8 en.

Ključne besede: ALS poševnega pogleda, fasadne teksture, 3D model stavb, koregistracija, zaznava oken, 3D kartiranje

Izvleček

V zadnjem času se pojavljajo zahteve po čim bolj podrobnih 3D modelih stavb. 3D modeli stavb so oz. bodo uporabni na različnih področjih, kot so npr. računalniške igrice, trženje mest, obvladovanje naravnih nesreč, turistični informacijski sistemi, simulacije prenašanja zvoka in nadzor sonaravne gradnje. Večstransko aerolasersko skeniranje (ALS) poševnega pogleda je omogočilo zajem popolnih podatkov urbanih območij, vendar postopki za samodejno izdelavo 3D modelov še vedno predstavljajo velik izziv. Namen diplomske naloge je izdelava programa, ki samodejno izloči površine oken in z njimi obogati obstoječi 3D model stavb.

Predhodno smo 3D žični model avtomatsko koregistrirali z ALS poševnega pogleda. Postopek zaznave oken predstavljen v diplomski nalogi temelji na informaciji o položajih oken pridobljenih z obdelavo ALS in slikovnih podatkov. Laserski žarki, kateri prodrejo skozi steklene površine, se odbijejo na notranji strani stavb. Te točke, projecirane pod kotom skeniranja na segmentirano fasadno ravnino, predstavljajo površine oken oz. drugih vdolbin. Po drugi strani, prihaja na teh mestih do velike spremembe intenzitetnih vrednosti med zidom in okvirji oken na fasadni teksturi, kar omogoča izločitev robov. Ti dve dejstvi smo izkoristili za pridobitev začetnih okenskih površin. Dodatno smo za eliminiranje tujkov na fasadi in izboljšanje okvirjev oken, uporabili binarno masko narejeno z rastjo regij homogenih delov fasadne teksture (zidov). Predpostavka, da je v enem nadstropju več enakih oken, je služila za pridobitev dodatnih okenskih površin. Najprej so bile fasadne teksture razdeljene na horizontalne nivoje. V vsakem nivoju so bile iterativno določene vzorčne matrike, katere si med seboj niso smele biti preveč podobne. S temi vzorčnimi matrikami smo kasneje izvedli navzkrižno korelacijo vzdolž korespondenčnega nivoja, kar je omogočilo zaznavo dodatnih okenskih površin.

Z zaznanimi okenskimi površinami, katere predstavljajo okna in vdolbine, smo na koncu teksturirali 11 fasad obstoječega 3D žičnega modela stavb. Kljub različni razporejenosti oken, njihovim različnim velikostim in visokim kriterijem ovrednotenja metodologije, je uporabljeni postopek zaznave oken izkazal relativno dobre rezultate (63% pravilnih zaznav in 16% lažnih zaznav). Vrednost pravilne zaznave je bila višja, kot v rezultatih dobljenih po metodi, kjer so bili uporabljeni zgolj podatki ALS (Tuttas & Stilla, 2013). Okna so dobro definirana, saj so osnova za večino povezane komponente robov, katere pripadajo okenskim okvirjem.

Ta stran je namenoma prazna.

Acknowledgement

I would like to thank my supervisor Assist. Prof. Mojca Kosmatin Fras, PhD for academic support. I also thank Prof. Uwe Stilla, PhD for giving me the opportunity to pursue my thesis at the Department of Photogrammetry and Remote Sensing at the Technical University of Munich. This thesis would never be feasible without Sebastian Tuttas who spent countless hours solving issues and trying out new solutions with me. I thank Dorota Iwaszczuk for encouragement and scientific advice. I would sincerely like to thank my parents and my sister for supporting me in my decisions and for giving me the opportunity to gain professional experience abroad. I would especially like to thank my grandmothers and relatives for their moral support. Last but not least, I would like to thank my study colleagues, friends and acquaintances who have genuinely contributed to the development of my personality.

Ta stran je namenoma prazna.

Zahvala

Zahvaljujem se mentorici doc. dr. Mojci Kosmatin Fras za vzpodbudo in strokovno pomoč. Hvala prof. dr. Uweju Stilli za možnost opravljanja diplomske naloge na Oddelku za fotogrametrijo in daljinsko zaznavanje Tehnične univerze v Münchnu in za njegove strokovne nasvete. Diplomska naloga ne bi nastala brez Sebastiana Tuttasa, s katerim sva preždela neštete ure ob reševanju problemov in iskanju novih rešitev. Za vzpodbudo in strokovno usmerjanje se zahvaljujem tudi Doroti Iwaszczuk. Posebej hvala mojima staršema in sestri za razumevanje in brezmejno podporo, ki je omogočila pridobitev strokovnih izkušenj v tujini. Hvala tudi babicama in preostalim sorodnikom za skrb in moralno podporo. Nenazadnje hvala študijskim kolegom, prijateljem in znancem, ki so dodobra prispevali k izoblikovanju moje osebnosti.

Ta stran je namenoma prazna.

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

SEZNAM SLIK

SEZNAM SHEM

SEZNAM TABEL

Ta stran je namenoma prazna.

Abbreviations

ALS	Airborne Laser Scanning
CityGML	City Geography Markup Language
LIDAR/LiDAR/lidar	Light Detection and Ranging
LOD	Level of Detail
NCC	Normalized Cross Correlation
OGC	Open Geospatial Consortium
PC	Point Cloud
RANSAC	Random Sample Consensus

Ta stran je namenoma prazna.

Slovar izrazov

Slovenska terminologija na področju obdelave podob in laserskih podatkov je precej neenotna, zato smo za lažje razumevanje razširjenega povzetka v slovenščini pripravili slovar izrazov.

3D žični model stavb	3D wire-frame building model
Večstransko ALS poševnega pogleda	Multi-aspect-oblique-view ALS
Fasadna slovnica	Façade grammar
Grafični gradnik/Entiteta	Feature
Gručenje po metodi K-povprečja	K-means clustering
Izločitev/Izločen	Extraction/Extracted
Lažne zaznave	False negatives
Mejni okvir	Bounding box
Metoda projekcije gradientov	Projection gradient method
Morfološki operatorji	Morphological operators
Nepravilna zaznava	False positive
Oblak točk	Point cloud
Polnovalovna oblika	Full-waveform
Prag	Threshold
Pravilna zaznava	True positive
Rast regij	Region growing
Samodejen	Automatic
Stopnja podrobnosti detajla	Level of detail
Zaznava	Detection
Zbirna matrika	Accumulator array

Ta stran je namenoma prazna.

1 INTRODUCTION

For more than 20 years, 2D and 3D primitives from aerial images have been used for building extraction. Many primitives were difficult to handle, for example, abundant line segments that did not correspond to meaningful geometric features, leading to the need for a new approach. In the nineties, airborne laser scanning was introduced for commercial use, and thus acquisition of huge scenes, such as cities, became feasible. Point density and accuracy have improved significantly since then. Nowadays, lidar is also mounted to many new acquisition systems, for example, UAVs and mobile mapping systems which enable façade scanning. The new approach to data collection in this thesis is to use lidar set at 45° with respect to the nadir view; thus façade information can be acquired easily and quickly. This is a very big step in urban modeling, since not just geometry but also a great amount of semantics can be extracted when combining this data with imagery.

In the last decade, a massive market for 3D city models has been triggered by improvements in data acquisition methods, enhanced spatial data processing and visualization, the Web 2.0 concept and the proliferation of smart phones. However, most of the applications are specialized, and therefore require the use of specialized 3D city models. Whereas the majority of national mapping agencies do not see acquiring or updating this information as a viable option, Google and Microsoft have revolutionized 3D mapping with Google Maps and Microsoft Virtual Earth (Vosselman & Maas, 2010, p. 169, 170).

1.1 Motivation

Improved 3D city models have become a necessity nowadays. According to the World Health Organization the number of people living in urban areas will drastically increase in the following decades. »By 2030, 6 out of every 10 people will live in a city, and by 2050, this proportion will increase to 7 out of 10 people« This will bring a lot of challenges to urban planning, surveillance, disaster management response, sustainable building construction and thus the need for good, precise 3D maps.

In recent years, vast amounts of data from different sensors have become available. The challenging task that still has to be overcome is automation of procedures for combining data from different modalities and the automatic contextual interpretation of the data. Combined imagery and lidar information can introduce many benefits such as improved detection rate, increased semantic information, 3D scene generation, etc. The concepts of fused data and object recognition is being used also in self-driving cars which is one of the most promising technologies of the next decade.

Moreover, 40% of total energy consumption in the European Union is consumed by buildings. To reduce the EU energy dependency and greenhouse gas emissions more energy efficient buildings are

required (Directive 2010/31/EU, 2010). Fusion of different types of data can provide meaningful information for cost-effective energy savings. Knowing the geometry and semantics of façades from lidar and imagery and subsequently co-registering IR images (Avbelj et al., 2010) can provide information about the exact position of building leakages and encourage better decision making for further regulations and its surveillance.

Precise 3D building models of cities are required when assessing damage after natural disasters. The importance of an automated approach for extracting windows and their sizes would be of significant meaning for insurance companies. In particular, when in Russia shattered glass caused most of the damage due to the shock wave of the fireball meteor (NY Times, 2013). Moreover, the number of floors could be determined from detected windows and used to assess damages after flooding or to prepare evacuation plans.

1.2 Related work

Urban environments usually consist of repetitive and symmetrical structures, which might be of particular interest for façade reconstruction. Many approaches follow the projection gradient method introduced by Lee & Nevatia (2004) which divides a façade rectified image in a rectangular grid assuming most of the edges are due to windows (Recky & Leberl, 2010; Kulkarni, et. al., 2011). Another approach to dividing a façade into a rectangular grid is described in Yakubenko et al. (2010). The choice for the grid scale is made iteratively, considering the size of the cell, existence of a rectangle in the center of each cell, NCC similarity to neighboring cells, and horizontal grid symmetry. Additional grids are found by calculating the median cell of the extracted grid and correlating it with the part of the image without a grid. This method outperforms two other methods used for detecting repeating structures.

Several attempts to reconstruct façade elements have been made, differing in the data type used. Lee & Nevatia (2004) presented an automatic approach using rectified façade textures which not only extracts windows, but also distinguishes between arc-shape and rectangular windows and reconstructs 3D windows using the plane-sweeping method. Recky & Leberl (2010) claim that the afore-mentioned method for windows extraction is not suitable for complex façades that contain a lot of gradient content not belonging to windows. And therefore, they perform k-mean clustering in the CIELAB color space on the blocks divided by the gradient projection method.

Lidar usually penetrates glassy areas where no points are measured (Yang & Wang, 2011), thus one can reconstruct the outer boundary of the area, the accuracy of which depends on the point sampling distance. Especially when reconstructing windows, this can be a great advantage and has been used in

many projects so far, for example, Becker (2009); Pu & Vosselman (2007); Wang, et al. (2012). Boulaassal, et al. (2009) extracted contours that represent edges of windows from the long sides of Dalauney triangles, based on segmented façade plane from a lidar point cloud (PC).

In recent years, many approaches using data from mobile-mapping laser scanning have been implemented, for example, Wang, et al. (2012). Their window detection method is based on coarse assumptions which do not consider windows in the lower levels of the façade due to occlusions. Furthermore, the window width and height and façade spacing are defined for each façade separately, which makes the method not fully automatic. They also report an issue of insufficient data information on the upper levels of high buildings.

Due to occlusions, scanning view limitations, and consequently gaps or varying densities of the PC acquired from terrestrial or mobile-scanning systems, it is not sufficient to consider only geometric features extracted from the PC. Becker (2009) gives a solution for this by combining bottom-up and top-down modeling using high resolution façade images and terrestrial or mobile-mapping lidar scans for façade reconstruction. Aside from geometrical features from the PC, knowledge about the façade appearance and arrangement should also be included. In this particular case, geometric features extracted from the PC refined with 3D image information formed a basis for façade grammar definition—the hierarchy and repetitiveness of features combined with production rules—which enabled the completion of façade structures covered by noisy, inaccurate, or incomplete data. However, it should be noted that knowledge propagation might lead away from the ground truth.

The afore-mentioned limitations of terrestrial and mobile-scanning can also be tackled using multi-looking oblique view airborne laser scanning, which even covers a larger area. Due to the limited point density (less than 10 points per m^2), it is difficult to make use of only gaps in the façade plane to detect window positions. Therefore, Tuttas & Stilla (2011) look for the points which lie behind the segmented plane representing a façade. Additionally, they employ a discrete Fourier Transform to determine spacing between windows in case of detected repetitive structures on a façade. In Tuttas & Stilla (2012) also the reconstruction of windows is presented. An iterative quadrant-based search for window outline edges starting from detected window positions is performed on segmented points representing a façade. The choice of the best window edge is defined as a maximum of the sum of probability density functions calculated for each edge representing the same side of a window. The approach depends on the presence of points behind the façade plane and detected repetitive patterns and assumes a rectangular shape of windows. A more reliable and accurate method for detecting windows could be developed by adding high resolution image information to the scans acquired by multi-looking oblique view airborne laser scanning, which is the main topic of this thesis.

1.3 Problem overview and objective of the thesis

3D city models are already available for certain cities, but are mostly represented only as blocks or LOD 2 textured models and thus lacking detailed information. In an era when a vast amount of data from different sensors is available an automatic procedure for their semantic interpretation which aims to enrich 3D building models is required. The main objective of the thesis is to obtain detailed building geometry from fused data for an existing 3D building model in LOD 2 according to CityGML schema. Furthermore, another research question is also whether window detection performs better when combining oblique view ALS and imagery compared to the approach from Tuttas & Stilla (2013) using only oblique view ALS data.

Within the scope of the thesis, the lidar PC and the 3D wire building model were automatically co-registered and then window detection on fused imagery and ALS data was performed. The latter is based on connected components of edges intersecting rastered ALS points behind the façade plane. In addition, a refinement procedure exploiting similarity of windows within a floor using normalized cross-correlation was carried out. Finally, the methodology was evaluated.

1.4 Thesis structure

The thesis is organized into six main chapters. In the 1st motivation for the work and the state of the art techniques in the field of window detection are described. The 2nd chapter focuses on the theoretical concepts of the approaches used in the 3rd chapter – Methodology. The latter is comprised of three sub-chapters: *Co-registration of a lidar PC and 3D wire frame building model, Data pre-processing and Window detection*. Whereas the first two describe procedures of ALS and image data processing, the *Window detection* consist of all steps required for extraction of window patches. In the 4th chapter data used and the evaluation of results are presented. Conclusions regarding the methodology as carried out in this work are given in the chapter five followed by directions for further work. The last chapter is an abridgement of this work in Slovenian.

2 THEORY

2.1 Airborne laser scanning (ALS)

Data of large urban areas can be easily recorded by using a laser scanner mounted on an aircraft which measures the distance to the ground by emitting laser pulses and detecting their reflection. Combining this information with a given GPS position and IMU orientation in the moment of acquisition delivers 3D coordinates of measured ground points. In addition, the form of the reflected wave can be modeled and material properties of the surface further analyzed. Lidar is an active sensor, largely independent of sunlight, and can therefore be used during the day and at night. Lidar platforms are also often equipped with a digital camera. Simultaneous acquisition of image data brings better scene recognition, especially if lidar point density is very low (Vosselman & Maas, 2010).

2.1.1 Main components of ALS

The lidar system comprises the following components:

The scanner assembly consists of a laser, scanning mechanics and optics mounted over a hole in the aircraft's fuselage. The laser system emits laser pulses towards the Earth's surface. Density of scanning depends on the height and velocity of flight and laser repetition rate.

The GPS antenna is mounted on the top of the aircraft ensuring better reception of a signal from GPS satellites.

The inertial measurement unit (IMU) is closely fixed to the laser scanner. It records acceleration and rotation data. Acceleration data is used together with GPS data to interpolate the platform position on the GPS trajectory with accuracy greater than 10 cm. Rotation data contains information about the platform orientation.

The control and data recording unit stores all acquired data and enables time synchronization and control over the whole system.

Besides the lidar basic components on board the aircraft, GPS ground stations within a distance of 30 km from the aircraft are required for offline differential GPS calculation. This serves for reducing atmospheric disturbances which might influence precise position determination within a decimeter of accuracy. Since many countries already operate their own network of permanent GPS stations, normally no effort is required to set up a new station (Vosselman & Maas, 2010).

2.1.2 Oblique view laser scanning

Airborne laser scanning with a nadir viewing angle can provide information about the Earth's surface, for example, roofs of buildings and terrain. In contrast, mobile or terrestrial scanners can be used to acquire street view scenes, such as façade data. The point density in the latter case is usually higher than in the former. Nevertheless, all three acquisition techniques fail to generate complete data of large

urban scenes. This issue can be solved by using an oblique forward-looking laser scanning approach. As might be seen in the Figure 1 its scanning geometry enables it to acquire complete urban scene information—including roofs as well as façades—and thus, it is a very useful approach for the development 3D building modeling.

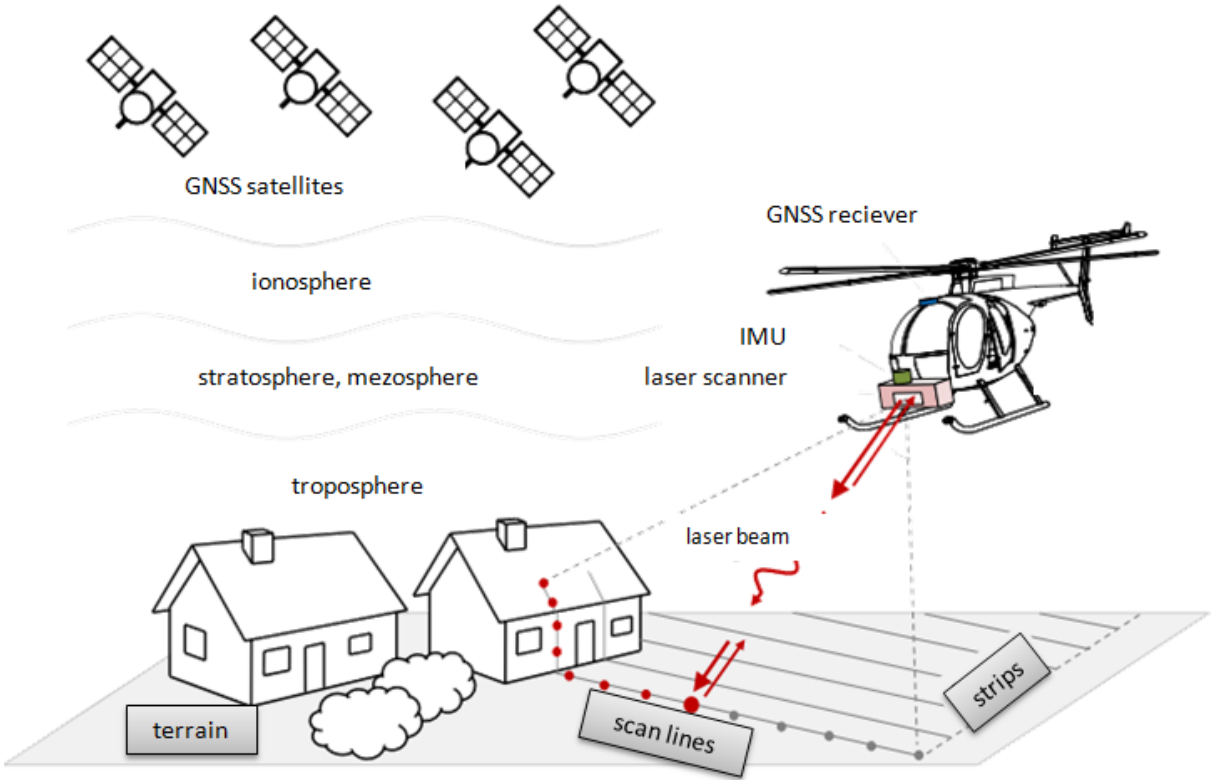


Figure 1: Oblique scanning geometry (Hebel, 2012)

The components of the system are the same as in a nadir-looking laser scanner, except that the laser is inclined at 45° in the forward direction. In our case, a rotating polygonal mirror was used to deflect the laser beam over the surface. This resulted in parallel, uniformly spread scanning lines oriented in one direction. In order to acquire data of the entire scene at least 4 overflights are required.

2.1.3 3D building model enrichment with multi-looking oblique ALS data

Data from multi-looking oblique view airborne laser scanning can be used for 3D building reconstruction. Once PCs from different stripes are co-registered, segmentation of the PC is performed (Tuttas & Stilla, 2012; Hebel & Stilla, 2010) in order to distinguish planar patches representing buildings from, for example, vegetation, the ground, etc. Having planar patches, one can define planes which form façades or roof of a building. Furthermore, points from the PC which might be found behind the façade plane—inside the building—usually represent points that penetrated glassy surfaces or holes such as windows, and are thus used for window detection.

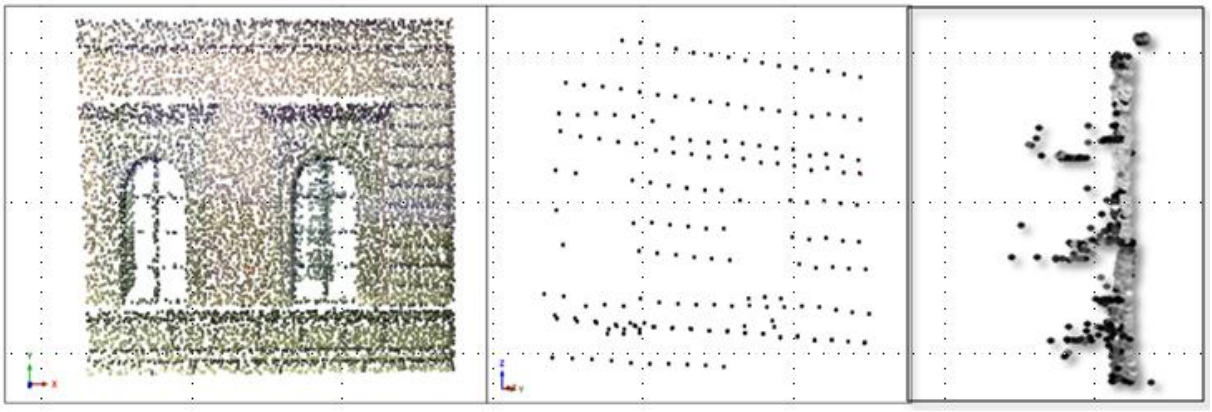


Figure 2: PC of a façade: TLS (right), oblique scanning (middle), the side view of oblique scanning (Tuttas & Stilla, 2012)

Consequently, there is no data from the window positions on the façade plane. These data patterns can be used in the window reconstruction step, as in Tuttas & Stilla (2012). In Table 1 adopted from Yang & Wang (2011) the relation between surface properties and reflection, absorption, and transmission can be found. As might be seen lidar light can be transmitted through glassy windows, whereas a large number of hits from white or colored façades are reflected.

Table 1: Light reflection, absorption, and transmission (Yang & Wang, 2011)

Surface	Specular Reflection	Diffuse Reflection	Absorption	Transmission	Example
Smooth	◯	△	△	×	Polished Metal
Rough	△	◯	◯	×	Dull Metal
Colored	◯	◯	△	×	White Surface
Dark	△	△	◯	×	Black Surface
Transparent	△	△	△	◯	Glass Window

In this table, ◯, △, and × indicate that a large amount, a small amount, and none of light are affected, respectively.

What is more, by inspecting points within a certain orthogonal distance to the façade, other intrusions on the façade can also be found. However, this mainly depends on the PC density and quality of the façade extraction.

2.2 Data fusion and building modeling

In recent years, many different sensors varying in spatial and radiometric resolution, such as SAR, InSAR, hyperspectral sensors, lidar, etc., have enabled relatively quick acquisition of large urban areas. For example, hyperspectral remote sensing can be used in assessing urban materials by exploring chemical or physical changes using spectral information. In contrast, lidar can provide precise geometrical shapes. Different datasets capture the heterogeneous nature of an urban environment and therefore their integration represents a challenging task—also called the registration problem. Manual extraction of large amounts of data is an expensive process calling for automatic procedures. This requires knowledge of certain characteristics of the relevant object in different types of data, such as its shape, size, patterns, tone, shadow outlines, length, height, etc. (Brook, et al., 2012). In the next chapters these issues are tackled.

2.2.1 Comparison between lidar and photogrammetry

Complementary information from lidar and images is beneficial for better description of objects. Whereas lidar can provide accurate information about homogeneous surfaces, accurate break-lines can be extracted from imagery (Habib, 2012). Kaartinen & Hyyppä (2006) (in Vosselman & Maas, 2010) assess building reconstruction from lidar and photogrammetric data and give the following conclusions:

1. Photogrammetric reconstruction was performed with an error rate four times smaller than that of laser scanning.
2. Photogrammetry gives good results in measuring outlines, whereas lidar is good for roof slope and height determination.
3. Processing lidar data takes about 5% of the processing time of photogrammetry, due to the higher automation level of lidar.
4. Quality of results strongly depends on lidar PC density.

In Table 2, a comparison between Lidar and imagery can be found, based on Habib (2012), EM 1110-1-1000 (2002); Díez, et al. (2008).

Table 2: Comparison of lidar and photogrammetry (Habib (2012), EM 1110-1-1000 (2002); Díez, et al. (2008))

	LIDAR	**Photogrammetry**
Energy source	Active	Passive
Geometry	Polar	Perspective
Sensor type	Punctual	Matrix or linear
Measurement of points	Direct without redundancy. Accuracy of information only depends on calibration of system components.	Indirect with redundancy. Images with overlay provide the intersection of the homologous rays (HR).
Semantic information	Extraction of linear, planar primitives due to their geometric relationship. FWF enables determination of surface roughness.	Rich with radiometric information (along break-lines) which enables differentiation between different surfaces.
Sampling	Irregularly located 3D points. Footprint depends on the beam divergence and flying height.	Full areas
Accuracy	Vertical accuracy is better than the planimetric accuracy.	Vertical accuracy is worse than the planimetric accuracy.
Flight plan	More complete. Due to small passes more complex planning of the acquisition campaign.	Considering longitudinal and transverse coverage.
Flight constraints	Less impact of time, daylight, night.	Daylight flying, clear atmosphere
Production range	May be automated, thus a greater production.	Higher need of editing control. Complicated and sometimes unreliable automated matching procedures (especially when dealing with large scale imagery over urban areas).
Budget	25%-33% of photogrammetric compilation budget	
Production	Software: depends on qualified commercial & technical people.	Software for the end-user: slow process of identification and manual extraction. Not reliable if automated, implies editing, especially at large scales.
Quality assurance & quality control	No common standards for quality assurance and quality control.	Straightforward procedures for quality assurance and quality control.
Feature extraction	Linear, planar patches	Points, linear, areal primitives

2.2.2 Establishment of a common reference system

Normally, the relation between data in different projections is given by forward and inverse mapping equations. However, when dealing with projected datasets based on different geodetic coordinate systems, the change of datum also has to be taken into account, as seen in Figure 3.

The German geodetic coordinate system is based on a local Bessel ellipsoid which optimally fits the area. Consequently much data is still in this system projected into the Gauss-Krüger projection. On the other hand, recently much data is obtained by GPS positioning, which uses the global WGS84 ellipsoid. These two ellipsoids differ in size and shape and in their position and orientation in space.

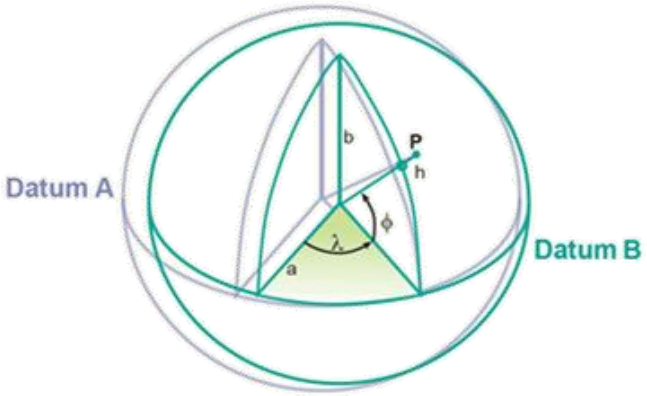


Figure 3: The relationship between two ellipsoids (Knippers, 2009)

The connection between the two ellipsoids, also called datum transformation, is given by the 7-parametric similarity transformation consisting of three rotations, three translations and a scale. In geodesy, the most commonly used transformations are Bursa-Wolf and Molodensky-Badekas. The latter additionally requires coordinates of a centroid at the relevant area, which in this case is the center of rotation. The origin of the Bursa-Wolf transformation is the center of the ellipsoid, defined as follows: the Z-axis points in the direction of the minor axis of a reference ellipsoid passing through the north pole; the X-axis is defined as the intersection of the Greenwich meridian plane and the equatorial plane of the ellipsoid; and the Y-axis is advanced 90° east along the equator (Ilife & Lott, 2008).

Normally rotations in datum transformations are less than 1 second of arc, and therefore a simplified rotational matrix can be applied. Equation 1 shows the Bursa-Wolf transformation using a simplified rotational matrix (Deakin, 2006).

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} + m \cdot \begin{pmatrix} 1 & \varepsilon_z & -\varepsilon_y \\ -\varepsilon_z & 1 & \varepsilon_x \\ \varepsilon_y & -\varepsilon_x & 1 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad \textit{(Eq. 1)}$$

x, y, z… coordinates of the target Cartesian reference system

d_x, d_y, d_z… translation vector

X, Y, Z… coordinates of the source Cartesian reference system

ε_x, ε_y, ε_z… small angles of the linear rotational matrix

m… the scale factor

2.2.3 Automatic co-registration of different kinds of datasets

Co-registration is a procedure aiming to optimally co-align two different datasets in a common reference frame. Dealing with vast amounts of urban data triggers the question of how to automate this procedure. This is not a straightforward task, since different types of data have different characteristics, and thus it is sometimes tricky to find corresponding similarities between different datasets which would ensure their fusion.

The scope of automatic co-registration consists of the following steps (partly adopted from Habib (2006)):

1) ***Search for registration primitives (features)*** – for example, points, lines, or regions, from the involved datasets. The choice of features strongly depends on the properties of datasets. For instance, lines extracted from the scene's lidar PC are not the same as edges extracted from the visible or infrared images of the same scene.

2) ***Matching condition that ensures the correspondence of conjugate primitives***. Once we already have primitives, normally several conditions have to be set in order to eliminate primitives which are not conjugate in both datasets. Different conditions such as proximity, relative positions of the primitives, different similarity measures, etc., can be set.

3) ***A transformation function*** that describes the mapping between the reference frames of the datasets. This depends on data dimensionality (2D, 3D) and differences between the data reference frames (change of orientation, scale, translation, etc.).

4) ***Mathematical constraint between primitives***, minimizing the differences in the position of the conjugate primitives. Its formulation depends on the choice of primitives as well as the utilized transformation function.

2.2.4 Lidar PC co-registration to other datasets

Normally when registering two datasets, interactive selection of tie points that link two datasets is required, making the transformation parameters feasible to calculate. In particular, this approach might be inconvenient when co-registering lidar to other datasets. Individual points of the lidar PC are difficult to interpret on the image. Therefore, either linear or areal features may be more appropriate.

Linear features can be extracted either from the lidar intensity image (Habib, 2004), by the segmentation of the PC or intersection of two planar patches (Habib, 2012). Another solution to derive primitives is to make an intersection of three differently oriented segmented planar patches from the lidar PC and to find the corresponding intersection point in the other lidar dataset (Hebel & Stilla, 2009). Planar patches might be extracted from the lidar PC by a seeded region growing algorithm based on comparison of normal vectors in an organized TIN as in Armenikas et al. (2010).

On the other hand, the 3D Hough transform proposed by Vosselman (1999) can be used for planar patch extraction. It parameterizes each point to the parameter space, representing a sinusoidal surface. A 3D accumulator array is created and the maxima represent planes of the points in the PC. However, Tarsha-Kurdi, et al. (2007) claim that their extended RANSAC algorithm performs much better and faster in lidar plane detection than the 3D Hough transform. Their extended RANSAC calculates a plane from 3 randomly chosen points and assigns all inliers at a given threshold to this plane. It repeats the procedure N times, comparing number of inliers between sequentially defined sets of inliers that contribute to a faster implementation of the algorithm. They introduce a new parameter: the maximum probable number of points belonging to the same plane that makes the original RANSAC more reliable.

To summarize, the choice of proper primitives must be made considering the possibilities of extracted features from lidar depending on the common characteristics of the other dataset.

2.2.5 Lidar PC and 3D building model co-registration

The method used in this study to co-register a lidar PC and a 3D wire building model is described in Hebel & Stilla (2009). A search is performed for a sufficient number of corresponding planes that are close enough to each other and that have small angles between their normals, as seen in Figure 4.

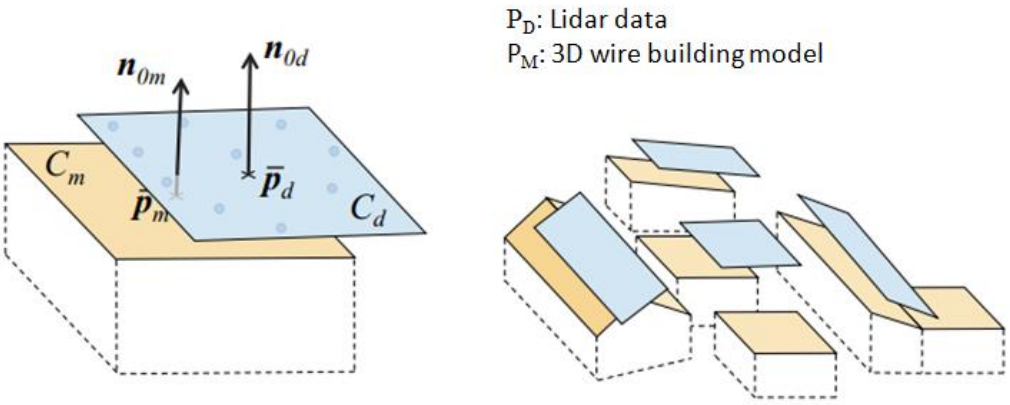


Figure 4: Corresponding planar patches from both datasets (Hebel, 2012)

The relation between two corresponding planes is established by a rigid transformation given by Equation 2.

$$x' = R \cdot x + t \qquad \textit{(Eq. 2)}$$

x… coordinates of the lidar data

x'… transformed coordinates of the lidar data

R… linear rotational matrix (valid only for small angles)

t… translational matrix

After rotational and translational matrices are applied to one of the planes, its corresponding plane should lie on the same plane.

The Hessian form of all planes has to be known as well as their centroid positions. For each pair of homologous planes, the following two conditions should be fulfilled: the centroid of a lidar plane should be on the corresponding plane from the 3D wire building model. Furthermore, the corresponding normal vectors should be the same. This can be written by the following equations:

$$(R \cdot c_D + t - c_M) \cdot n_M = 0 \qquad \textit{(Eq. 3)}$$

$$(R \cdot n_D) \cdot n_M = 1 \qquad \textit{(Eq. 4)}$$

R… 3 by 3 linear rotational matrix (valid only for small angles)

t… translational matrix

c_D… centroid of a plane belonging to the lidar data

c_M… centroid of a plane belonging to the 3D building model

n_D… normal of a plane belonging to the lidar data
n_M… normal of a plane belonging to the 3D building model

Equations 3 and 4 are solved for each pair of corresponding planes. In order to solve all six variables from the rigid transformation, at least three pairs of corresponding planes have to be found. In case of redundancy of corresponding pairs, an over-determined system can be solved by linear least squares. For the weights in the stochastic model, areas of the planes can be used.

2.2.6 Level of Detail (LOD) in building modeling

For easier representation and exchange, an international standard for 3D city and landscape models—CityGML—was adopted by the Open Geospatial Consortium (OGC, 2012). CityGML represents a common semantic information model for representing 3D urban objects through a multi-scale representation. Four different aspects of virtual city models, i.e., semantics, geometry, topology and appearance, are taken into account. CityGML comprises five levels of detail (LODs) which differ from each other in their geometric and thematic characteristics, enabling efficient visualization and quality assessment of 3D building models and data analysis. Whereas the LOD 0 shows the least detail and is represented by a two-and-a-half dimensional Digital Terrain Model (DTM) which can be draped, the most detailed LOD 4 denotes architectural models with interior structures, vegetation and transportation.

For better categorization of the most important geographic features (buildings, DTMs, water bodies, vegetation, and city furniture) of virtual 3D city models, CityGML provides: class definitions, normative regulations and explanations of the semantics. In particular, there is a special class—Buildings—for 3D building models with specifications for each LOD.

LODs 1 to 4 show building volume and a 3D Terrain Intersection Curve, which integrates buildings from different sources with the DTM.

LOD 0 – Buildings can be shown as horizontal, 3-dimensional surfaces of identical height—for example, the footprint of the building and its corresponding roof edge.

LOD 1 – Buildings are represented as blocks and not differentiated in detail (flat roofs). The volumetric and surface parts of the exterior building shell are identical.

LOD 2 – Buildings differ according to their roof, and their boundary surface can be thematically described—for example, wall, roof, ground plate, outer floor, outer ceiling, etc. Also outer building

installations, such as chimneys, dormers, balconies, or outer stairs, can be included in LOD 2 models, but only if their extent exceeds 4*4 m. The positional absolute accuracy should be within 2 m.

LOD 3 – models are architectural models with detailed wall and roof structures, doors, and windows, and their 3D point accuracy is within 0.5 m. Objects are represented as real features. The main difference between the LOD 2 and LOD 3 is that the latter represents openings in a building, i.e., windows, doors as shown in Figure 5. "All openings should be represented as holes by interior ring within the surface they lie on. In case these openings are recognized as doors or windows their outer boundary can be built of the same points as the inner ring."

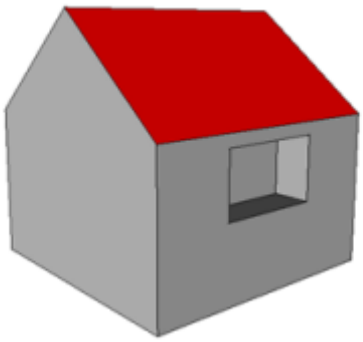


Figure 5: Example of an opening (OGC, 2012)

LOD 4 – has all the characteristics of LOD 3, including the interior structure of a building.

2.3 Image processing

2.3.1 Region growing

Region growing groups pixels whose values are similar to the predefined pixels, e.g. color, intensity, texture taking their connectivity into account. In order to start region growing a criteria that merges sub-regions of pixels into regions has to be set and it stops when the criteria is not fulfilled anymore. When all the regions are extracted additional characteristics might be taken into account, e.g., size, shape of the regions. Different approaches of region growing might be used using other starting criteria, e.g., seed points, arbitrary regions (González & Woods, 2008: p. 763-769).

2.3.1.1 Seed points

Seed points are chosen depending on the region we want to extract. One possibility is to find connected components and erode them to 1 pixel. In our case, façade points were defined due to the prior knowledge about the position of the façade. The second step is to define the connectivity criteria, e.g. 4-connected, 8-connected (González & Woods, 2008: p. 763-769). In other words, this refers to

the number of pixels around the seed point that should be tested for the predefined threshold. Pixels that are close enough to the value of the seed point are kept and the rest are eliminated. This is an iterative procedure employed until a region is formed which does not change within the next iteration.

2.3.1.2 Region splitting and merging

An image is divided into arbitrary, disjoint rectangular regions of a predefined size. The values of the center points of two adjacent rectangles are compared and if their intensity difference is within the threshold they are merged together (Halcon, 2011).

2.3.2 Normalized cross correlation

Normalized cross correlation (NCC) is used in image processing as a metric to evaluate the degree of similarity between two compared images. It is less sensitive to linear changes in the amplitude of illumination in the two compared images than cross correlation. The process of the NCC consists of:

- template matching
- peak finding.

First, a template, displaying a scene which should be found on an image, is chosen. The image has to be larger than the template. The template is moved along the image and for each position the value of the NCC is calculated which ranges between -1 and 1. Value of 1 represents complete similarity, and occurs when the normalized template and the normalized region of the image are identical. The result is a matrix representing NCC values for each position of the template (Gonzales & Woods, 2008: p. 869-72). Depending on the purpose, a search for maxima is performed. In many cases the measure of similarity is the global maximum between the image and the template.

3 METHODOLOGY

The methodology is organized in three main chapters: *Co-registration of a lidar point cloud and a 3D wire frame model*, *Data pre-processing*, and *Window detection*. The diagram in Figure 6 shows intermediate steps of the first two chapters with their respective input data. Three types of data were used as input: an oblique-view-multi-aspect lidar point cloud (PC), a 3D wire-frame building model, and visible façade textures (marked blue in Figure 6). The data marked red in Figure 6 is the output data of the pre-processing steps from chapters 3.1 and 3.2 and was used in the window detection stage, which can be seen in Figure 7.

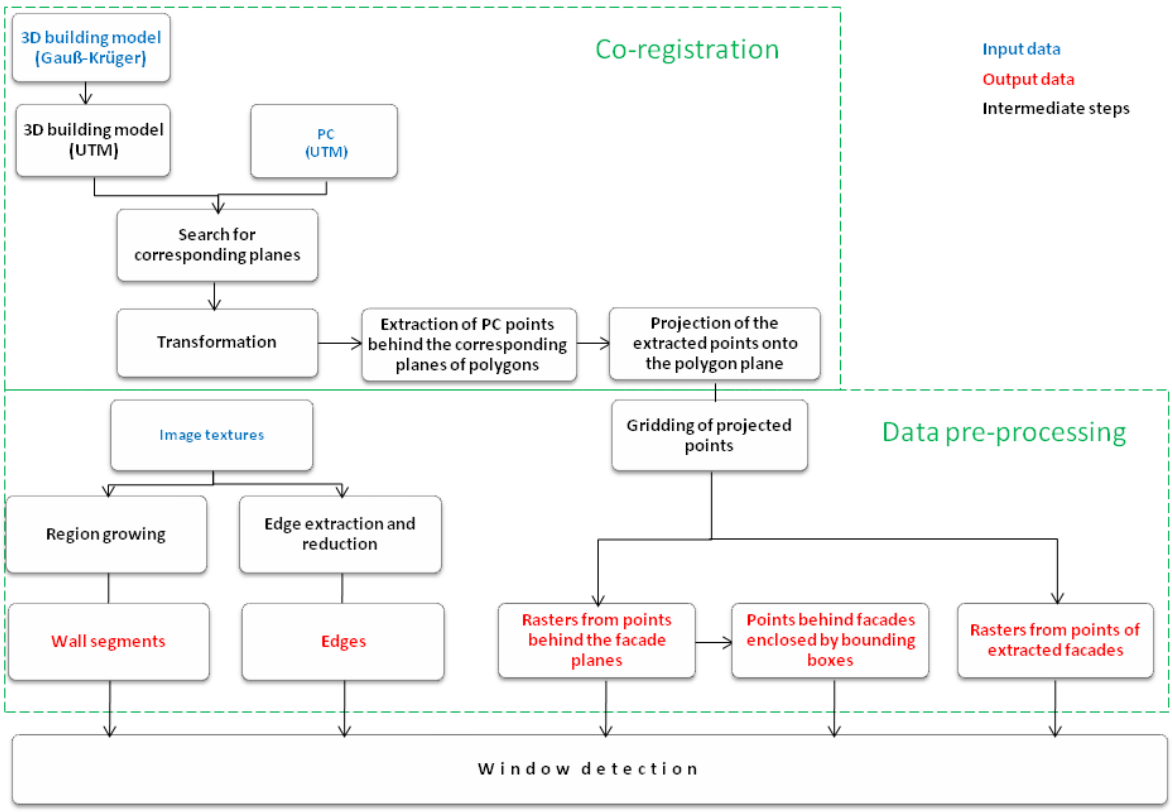


Figure 6: Co-registration of lidar PC and 3D wire-frame model and data pre-processing steps

Data first had to be co-registered to then enrich the 3D building model with detected windows using façade image textures and the oblique lidar PC. The 3D building model was transformed from Gauß-Krüger to UTM. Whereas image textures were already on the 3D building model, co-registration of the oblique lidar PC and 3D building model had to be carried out. Co-registration was based on a rigid transformation whose parameters were obtained from mathematical constraints between planar segments extracted from the PC and their corresponding polygon planes from the 3D building model, described in 3.1.2.

Lidar beam usually penetrates glassy areas (windows), and thus hits might be found inside a building. Therefore, points found behind a façade plane, projected for the angle of 45° to the respective façade plane were assumed potential candidates for intrusion areas. On the other hand, due to the penetration of laser beam through these areas, normally there are no hits from the façade plane in these areas. These two facts were exploited when deriving information about window areas from the oblique-view lidar PC. To enable easier data manipulation, the unorganized PC was gridded. In particular, rasters were created from *points found behind the façade plane*—representing positions of intrusions after being projected onto the polygon plane and those *belonging to extracted PC façade planes*—representing no-data areas at window positions. To enclose window areas of *rastered points behind the façade plane*, bounding boxes were assigned to each connected component. More about PC gridding is shown in 3.2.1.

Whereas the *rasters from points behind façade planes* gave reliable and accurate information about window positions, additional information to obtain exact window frames was required. Since window frames on a façade texture usually have a different color from the wall and glass, notable changes in the intensity values occur at these positions. Thus, edges using Sobel were extracted from visible façade textures. However, intensity values do not only change at the positions of window frames. To eliminate or at least reduce edges not belonging to windows, additional image processing steps were carried out, such as morphological operations and thresholding of gradient profiles shown in 3.2.2.

Furthermore, façade walls in the textures are represented by large areas of very homogeneous color compared to the window areas. Therefore, region growing resulted in significantly larger wall segments compared to the segments representing window areas. The smaller segments were eliminated by thresholding and binary masks were produced from larger segments representing wall areas, as described in 3.2.3. The binary masks, in turn, were used for the refinement of the intermediate results in the window detection stage, as well as for the refinement of the results produced by the block-wise NCC described in 3.3.

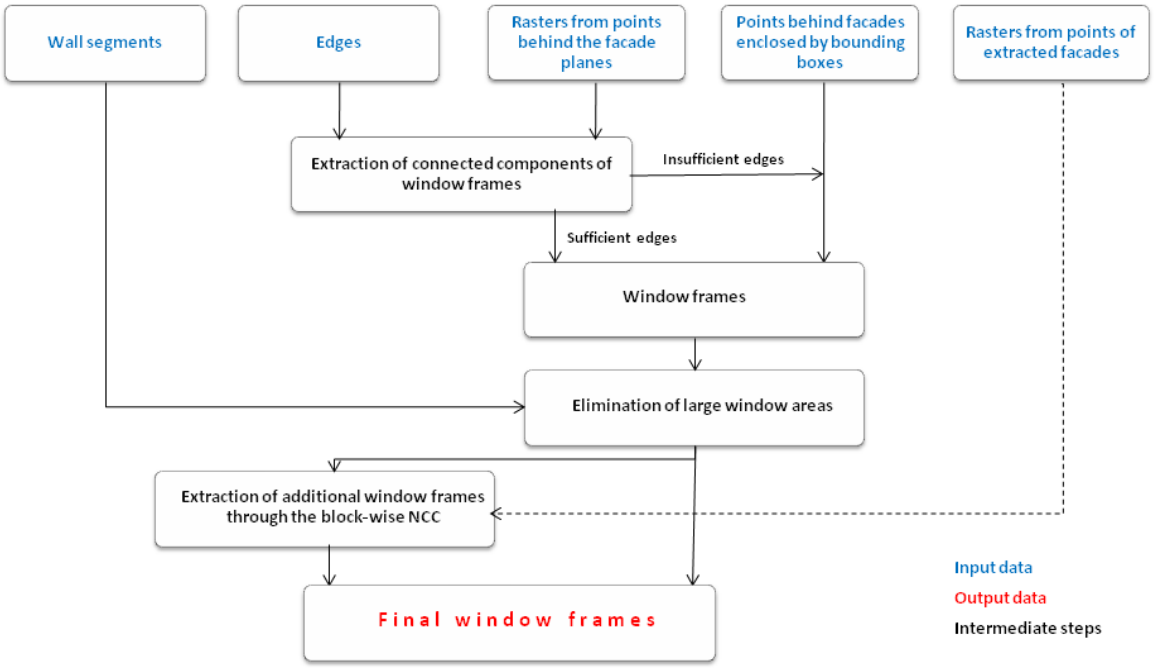


Figure 7: Window extraction steps

The data obtained in the pre-processing steps was used for window extraction as seen in Figure 7 shown in blue. Assuming that extracted edges at the positions of *rastered points behind a façade plane* are mostly caused by the difference in intensity values between window frames and their surroundings, connected components formed by edges, that intersected areas where *points behind windows* were present, were extracted. In case edges were badly represented at a position of points behind a façade plane, a bounding box enclosing *rastered points behind a façade plane* was placed at this position. More can be found in subchapters 3.3.1 and 3.3.2.

The coarse resolution of rasters representing points behind windows caused that two window areas lying close enough to each other, were merged into one. Moreover, high changes in intensity values of non-window objects connected to edges of window frames contributed to detection of too large window areas. Therefore, the binary mask of wall segments obtained from region growing was used to refine this. The application of masks from region growing is described in 3.3.2. This led to detected outer window frames.

Since all the steps, leading to extracted window outlines, were dependant on the presence of points behind windows, there were certain windows not detected at the positions where points behind a façade plane were not found. Assuming that there are many windows in the same floor of the same kind, templates for each floor (block) were chosen from previously detected window patches and the normalized cross correlation was performed with them along the respective block. Maximally three

templates were chosen per block in an iterative procedure, reducing computing cost and preventing false alarms. A rectangle of a size of the template was placed at a position of a maximum obtained by NCC that was within a threshold. This process also called block-wise NCC is described in 3.3.3. The output from this step was combined with previously detected outer window frames, resulting in final extracted window frames as shown in 3.3.4.

3.1 Co-registration of a lidar PC and a 3D wire frame building model

The lidar PC was given in the UTM coordinate system, whereas the 3D wire frame building model was given in the Gauss-Krüger (GK) projection. Therefore, establishment of a common reference system was required, enabling easier data manipulation and interpretation.

3.1.1 Transformation to the UTM reference system

The 3D wire frame building model of the TUM building was transformed from the GK coordinate system to the UTM coordinate system using indirect 7-parametric transformation with parameters for the region of Bavaria.

First, 3D wire building model data were transformed from the GK projection to the Bessel ellipsoid. Second, the geographic coordinates (λ, ϕ, H) on the ellipsoid were transformed to Cartesian coordinates (X, Y, Z) on the ellipsoid using the known ellipsoid parameter set. In the next step, the datum transformation from the origin to the target ellipsoid was done. This is a 3D similarity transformation based on 7 parameters (3 translations, 3 rotations and 1 scale factor) which performs a best fit of the two ellipsoid shapes for the relevant area. For the particular case, Bursa-Wolf transformation was chosen. Cartesian coordinates give no information about position and height on the target ellipsoid, therefore, a transformation to geographic coordinates was performed (λ, ϕ, H). In the last step, the ellipsoidal coordinates were projected into the UTM, transforming longitude and latitude to easting and northing. Concerning the height reference system, both datasets have orthometric heights.

Since parameters for Bavaria were used, whose accuracy ranges around a few meters, both datasets were not sufficiently co-registered. An additional reason for misalignment could be assigned to the position accuracy of both datasets and the generalization of the 3D wire model. The 3D building model was acquired from aerial images, and thus the outlines of building walls are aligned to the outlines of roofs, showing a slightly different geometrical representation from that of the PC. To overcome this issue, co-registration was carried out.

3.1.2 Co-registration

The 3D building model was acquired from aerial images and thus, polygons representing façades of buildings were aligned with roof overhangs. Whereas the PC showed hits from real façade planes. The difference in geometric representation of buildings between the datasets and rather low PC point density, contributed to the choice of planes, being the most suitable primitives for the given task. Since the PC and 3D building model were already within a few meters of each other, a search for corresponding pairs in both datasets was based on proximity and geometrical constraints. Mathematical constraints between corresponding planes were established and parameters of a rigid transformation determined. Since the number of corresponding pairs was higher than the number of unknowns of the system, the over-determined system was solved using least squares.

Plane extraction and a search for corresponding pairs in both datasets were done in two different ways described as Method 1 and Method 2.

3.1.2.1 Search for registration primitives

The 3D wire frame building model and PC were reduced to a centroid local coordinate system.

3.1.2.1.1 Method 1

The Random Sample Consensus Algorithm (RANSAC) was used to get lidar planar patches from the segments representing parts of buildings (Hebel & Stilla, 2009).

RANSAC is an algorithm for robust model fitting of, for example, a plane. First, it defines a plane through three randomly chosen points. Then it searches for points that are within a threshold to the plane and extracts them as a subset, partitioning the data set into inliers and outliers. If the number of inliers, also called consensus set, is above the threshold, a new plane is estimated from them. Otherwise, a new subset is selected and the aforementioned process is repeated. Finally, after minimal subsets have been established, the model with the subset containing the most inliers is chosen as the best-fitting result (Vosselman & Maas, 2010, 67-72). Planar patches were extracted from the segments. For each of them plane coefficients were calculated using singular value decomposition. The plane parameters for polygons were defined using three polygon vertices. In addition, also their corresponding barycenters were calculated.

3.1.2.1.2 Method 2

Method 2 was done following the steps from Tuttas & Stilla (2013). Points which were close enough to the 3D building model—behind and in front of their polygons—were extracted using bounding box. Each plane from the 3D building model was assigned a set of points and RANSAC based on the

normal vector, parallel to the one of its corresponding polygon plane, was performed. Planes were estimated from inliers using principal component analysis.

3.1.2.2 Matching condition between corresponding primitives

3.1.2.2.1 Method 1

In order to find correspondences between planar PC segments and polygons, four conditions were taken into account: proximity, angle between normals, size of the area and wall elimination.

Proximity

All points within 1 m from the barycenter were extracted. These points belonged to different segments. The mode was calculated in order to define the frequency of occurrence of each segment number within a set. The one with most frequent occurrence was chosen as a potential candidate to be a corresponding pair of a polygon.

Angles

For all potential candidates, the angle between the normal of a PC segment and its corresponding polygon plane (assigned in the previous step) were calculated. Angles below 10° and between 170° and 190° were taken as limits.

Area

After applying the above-mentioned constraints, some visually inspected outliers were present in corresponding pairs. All pairs with this character had a small area. Therefore, another constraint of the size of the polygon area was set to be minimally 100 m^2.

Elimination of walls

Due to inconsistency of the 3D wire frame building model obtained by photogrammetric acquisition and the PC, corresponding pairs representing walls were eliminated. Walls of the 3D model do not really represent the plane of the façade but are extended from overhangs.

In Figure 8, all seven extracted PC segments with their corresponding polygon planes might be seen.

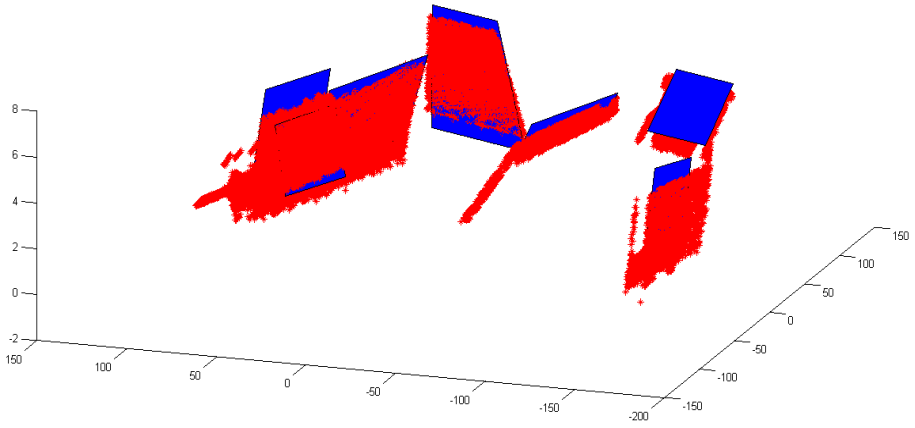


Figure 8: Extracted PC planes and polygons

3.1.2.2.2 Method 2

The search for corresponding pairs took into account almost the same factors as mentioned in Method 1, but it was performed in a slightly different way. The distance between corresponding planes and the angle between their normals were considered and additionally, vertical planes representing walls were eliminated. The transformation parameters were estimated using these corresponding planes. More can be found in Tuttas & Stilla (2013).

3.1.2.3 Transformation function and mathematical constraints between primitives

The mathematical correspondences between primitives and the transformation function were in both cases the same.

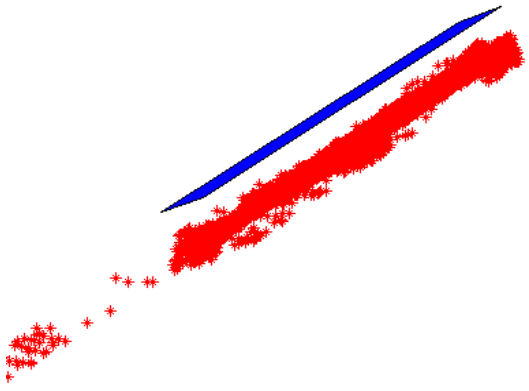


Figure 9: A PC segment and its corresponding polygon plane before co-registration (side view)

In Figure 9, the deviation between a pc segment and its corresponding polygon plane before the transformation might be seen.

The relation between lidar planar patches and polygon planes was established by a 6-parametric rigid transformation shown by Equation 2 in chapter 2.2.5. At least 3 pairs of corresponding planes oriented in approximately orthogonal directions were required to solve 6 unknowns—3 rotations (α_1, α_2, α_3) and 3 translations (t_1, t_2, t_3). Since redundant pairs were found, least squares was used to solve the over-determined system. Equations 5 and 6 represent the functional model of the system. An identity matrix was used for the stochastic model. The system was linear due to the assumption of small rotation angles, and therefore there was no need to approximate it to a first-order Taylor series expansion (Hebel & Stilla, 2009).

$$\begin{bmatrix} c_{D2} * n_{M3} - c_{D3} * n_{M2} \\ c_{D3} * n_{M1} - c_{D1} * n_{M3} \\ c_{D1} * n_{M2} - c_{D2} * n_{M1} \end{bmatrix}^T \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + [n_{M1} \quad n_{M2} \quad n_{M3}] \cdot \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = [n_{M1} \quad n_{M2} \quad n_{M3}] \cdot \begin{bmatrix} c_{D1} - c_{M1} \\ c_{D2} - c_{M2} \\ c_{D3} - c_{M3} \end{bmatrix}$$

(Eq. 5)

$$\begin{bmatrix} n_{D2} * n_{M3} - n_{D3} * n_{M2} \\ n_{D3} * n_{M1} - n_{D1} * n_{M3} \\ n_{D1} * n_{M2} - n_{D2} * n_{M1} \end{bmatrix}^T \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 1 - [n_{M1} \quad n_{M2} \quad n_{M3}] \cdot \begin{bmatrix} n_{D1} \\ n_{D2} \\ n_{D3} \end{bmatrix}$$

(Eq. 6)

α_1, α_2, α_3… rotation angles

t_1, t_2, t_3…translation parameters

n_{D1}, n_{D2}, n_{D3}… normal of a plane belonging to the lidar data

n_{M1}, n_{M2}, n_{M3}… normal of a plane belonging to the 3D building model

c_{D1}, c_{D2}, c_{D3}… centroid of a plane belonging to the lidar data

c_{M1}, c_{M2}, c_{M3}… centroid of a plane belonging to the 3D building model

Linear least squares was used on 7 pairs of planes obtained by automatic extraction. In Figure 10, the result after adjustment is shown. The same PC segment and its corresponding polygon plane are lying on the same plane after the transformation.

Figure 10: The PC segment and its corresponding polygon plane after the co-registration

3.1.3 Extraction of points behind a façade plane and their projection to the 3D building model façade plane

Once the lidar PC had been co-registered to the 3D building wire model, a search for points behind façade planes was done. All points from the PC that were within the bounding box—1 m in front and 3 m behind a polygon plane—were extracted for each plane of interest. Orthogonal distances were calculated from each extracted point to the corresponding PC segment plane representing façade according to the Equation 7.

$$\Delta d = \frac{ax_0+by_0+cz_0+d}{\sqrt{a^2+b^2+c^2}} \qquad \textit{(Eq. 7)}$$

Δd… distance to the plane

a, b, c… façade plane normal

x_0, y_0, z_0… laser point

When projecting extracted points to the respective façade plane, 45° oblique-view geometry of lidar scanning had to be considered. An approximation was done, and points found behind the windows were projected at a 45° angle to the façade plane.

First, they were projected orthogonally to the polygon plane:

$$q_{proj} = q - ((q - p) \cdot n) * n \qquad \textit{(Eq. 8)}$$

q_{proj}…projection of a point on a plane (x_{q_proj}, y_{q_proj}, z_{q_proj})

q…laser point (x_q, y_q, z_q)

p… point on a façade plane (x_p, y_p, z_p)

n… façade plane normal (a, b, c)

In the next step, a shift (Δh) along the z-axis was done for each point for its previous distance to the polygon plane, due to the oblique scanning geometry seen in Figure 11.

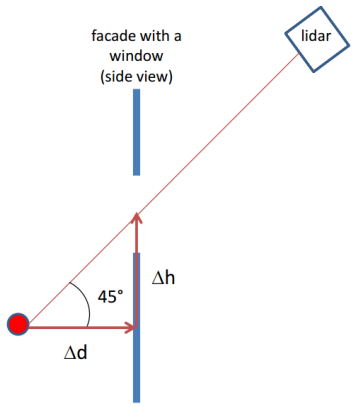


Figure 11: Oblique scanning geometry

After performing this operation, points were projected onto the polygon plane, mostly showing positions of intrusion areas. In order to enable easier comparison to images in an image coordinate system, all extracted points were transformed to the *yz*-plane. First, a translation of PC points projected to a polygon for the distance of the left upper edge of the corresponding polygon was done. Second, rotation of the coordinate system established by the polygon plane normal, normal vector in the *z* direction [0, 0, 1] and their cross product to the *yz*-plane was determined. The polygon plane was rotated to a plane parallel to yz-plane, having one of the corners of the corresponding polygon as the origin of the system, as shown in Figure 12.

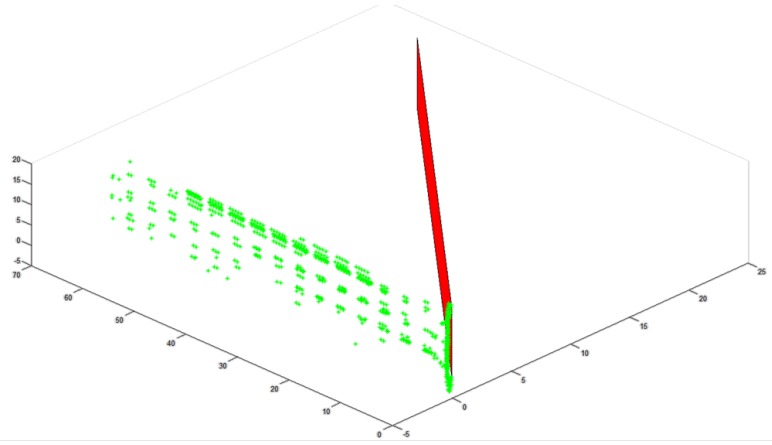


Figure 12: Plane after translation (red) and after rotation (green)

3.2 Data pre-processing

In this chapter all the data pre-processing steps required to obtain the input data for window extraction are explained.

3.2.1 Gridding of the PC

To enable easier window detection from lidar PC, the unorganized lidar PC was gridded. Rasters were produced both from *points lying behind façade planes* and from *points belonging to façade planes*.

For each extracted and projected set of points belonging to a certain polygon, gridding of points in a 2D coordinate system within the limits of the polygon was performed. Each raster cell was assigned 1 if there was a point and 0 if there was no point. Points that were gridded on the left, right or upper edges of a raster were eliminated by setting their values to zero, since they did not represent windows, but rather adjacent façades.

Connected components were found on each raster and their areas were enclosed by bounding boxes. Window positions were mostly unambiguous. However, as seen in
Figure 13, in many cases there was more than one bounding box at the position of one intrusion, preventing the raster from uniquely showing the positions of windows. In some cases more windows were represented by only one bounding box.

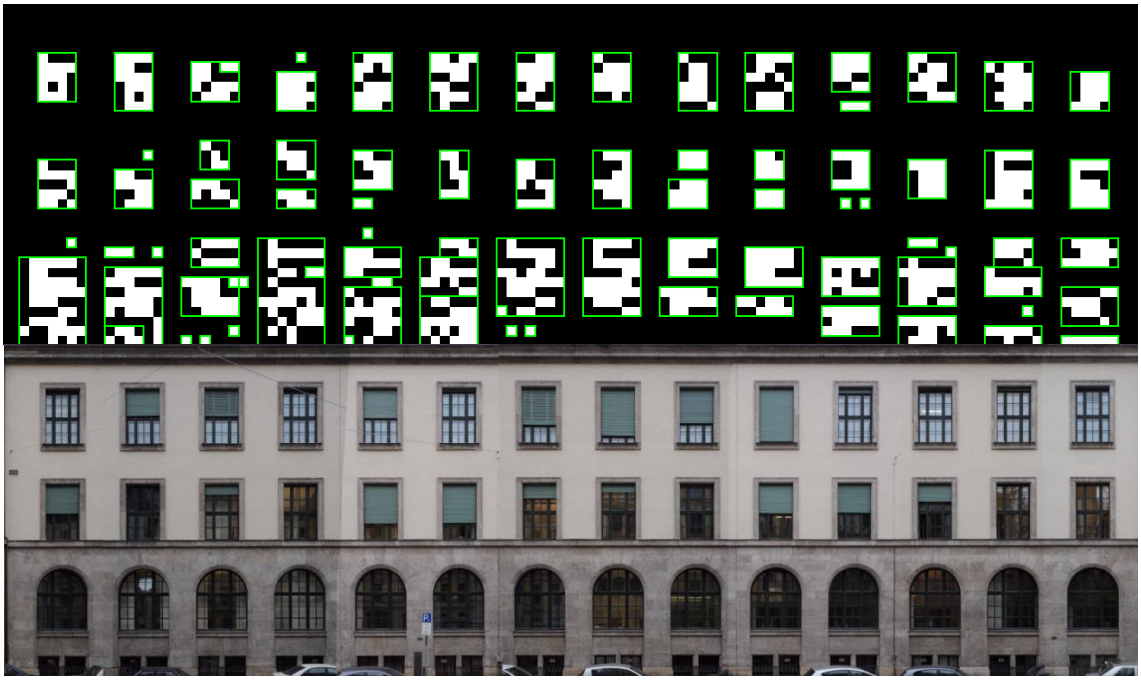


Figure 13: Points behind a façade plane with its corresponding bounding boxes (above) and image texture of the same façade (below)

Additional information about window positions was also extracted from the points which formed the façade plane. In particular, the lack of information about the façade plane, seen as no-data areas in Figure 14, is usually due to windows, doors, extrusions, or holes in the façade plane.

Figure 14: Raster of points belonging to the extracted façade plane

3.2.2 Edge detection

Window frames are usually different in color from their surroundings, causing high values of gradient on their boundaries in a façade texture. Therefore, edges were extracted to complement the information obtained from *rasters of points behind façades* for window extraction.

Window frames are either rectangular or window bars are ordered in a rectangular grid, therefore Sobel edge detector in vertical and horizontal direction was used. Prior to that images were smoothed using a Gaussian filter. However, the intensity values do not noticeably change only at the window frame boundaries, but also on certain wall parts where color changes. Two kinds of unwanted changes were noted:

- *changes in gradient due to small ornaments or details on façades*

These edges were smaller than the ones forming window frames. Therefore, a morphological operator opening was used.

- *changes in gradient due to differently colored parts of façades*

On certain edge detected textures, few very long edges along the façade not representing windows or other intrusions but still connected to edges of windows were caused due to battens on the façade or different façade colors on the upper and lower floors. Since these edges were very long compared to the edges of window frames and were oriented either in the vertical or horizontal direction, gradient profiles were made in both directions on the original image. Gradient profiles are sums of pixels of gradient of an image in either vertical or horizontal direction as shown in Figure 15. Pixel values of a row or a column were set to 0, if a threshold was exceeded.

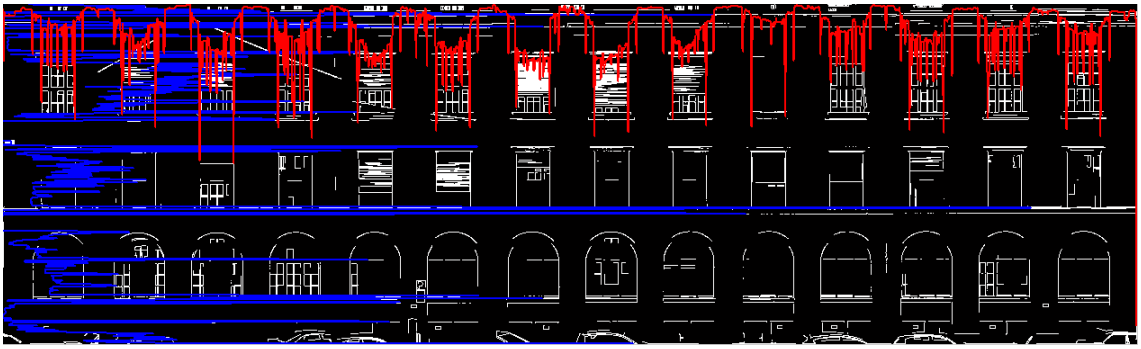


Figure 15: Vertical profiles

3.2.3 Region growing

Wall parts on the visible façade textures were very homogeneous and large compared to the window areas. This fact enabled rather easy separation of wall parts from the rest. Images were segmented using the region growing algorithm as described in 2.3.1.2. First, to avoid stopping the region from growing due to noise, images were smoothed using a median filter. Second, region growing was performed, comparing grayscale values of center points of 3-by-3 neighborhoods for a predefined threshold. If values of centers of adjacent rectangles (neighborhoods) were within the threshold, they were merged together. Since the façades represent the major part of the textures, all segments smaller than a threshold were eliminated. This cut out smaller or heterogeneous parts which are, for example, windows or ornaments. Figure 16 shows the segmented façade texture after area thresholding. It can be seen that most of the window areas were eliminated. A binary mask was created from the wall parts. All the remaining segments were set to 0 and the eliminated parts to 1.

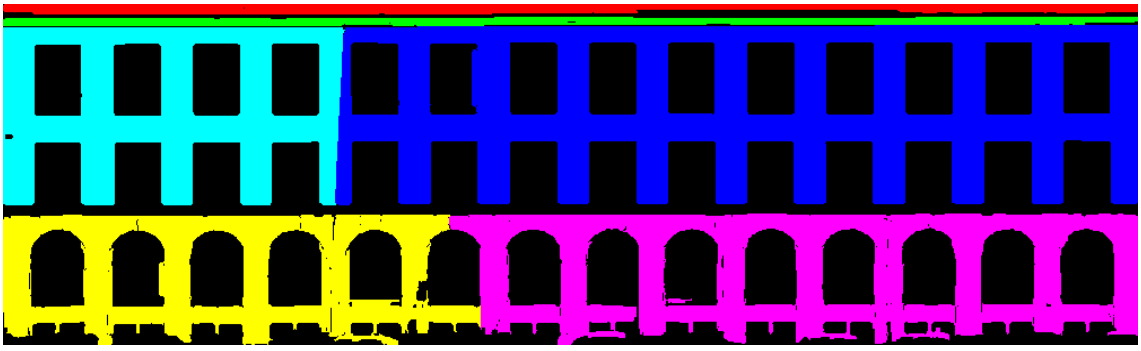


Figure 16: Segmented image using region growing after elimination of small segments

3.3 Window detection

Window detection was mainly based on extracted edges and *rastered points behind façade planes* (rasters). Whereas rasters yielded accurate and reliable positions of windows, extracted edges contributed to improved extracted window frames. Therefore, all connected components of edges at a certain position that were supported by the evidence of presence of points behind a façade, were extracted. All areas where points behind a façade were present, but almost no edges were found, were refined by adding the information of bounding boxes around *rastered points behind façade planes*. Furthermore, to retrieve certain window areas whose evidences were not present in the data, but they were similar to other already extracted windows in a floor, the block-wise NCC was carried out. The fact that several windows in the same floor are usually of the same kind was considered. Façade textures were divided into blocks. Templates were chosen from the previously extracted window areas based on data evidences belonging to a certain block and normalized cross correlation was performed along the respective block. A limited number of maxima in each block that were higher than a threshold were assigned a rectangle of the size of the template. Results of the block-wise NCC were added to the previously detected window areas that resulted in final window areas.

3.3.1 Detection of connected components of intrusions

In Figure 17, the relation between the extracted edges (red and orange) and rastered points behind a façade (light blue) can be seen. The intersecting area marked in red is well connected to other edges representing the window frame, and this fact was exploited to detect the initial window frames. On the other hand, edges that are not connected to the pixels, intersecting points behind the façade, usually do not belong to a window frame and should be eliminated.

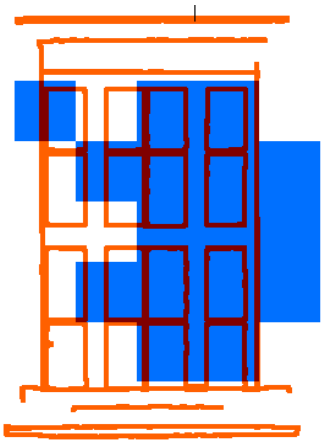


Figure 17: Extracted edges (red and orange) and rastered points behind the façade plane (blue)

To get positions of intrusions, a binary raster representing points behind a façade was multiplied with a binary image representing detected edges for each façade plane separately. This resulted in a raster with a cell value of 0 in areas with no edges or no points behind a façade and of 1 when both conditions were satisfied. The operation eliminated many noisy edges not belonging to windows or other intrusions. On the other hand, edges belonging to windows were split into small unconnected edges or in some cases even into pixels. Therefore, a procedure to retrieve edges that had been previously connected to these intersected parts was required. Thus, connected components were formed on the original edge-detected binary image and compared to the multiplied raster. For any pixel of value 1 (on a multiplied raster) within a connected component, the entire connected component was taken as a potential candidate for a window frame. Otherwise, the connected component was not considered further. A bounding box was formed around each connected component, as seen in Figure 18, to enable easier area comparison in the next step.

Figure 18: Extracted connected components enclosed by bounding boxes

3.3.2 Area comparison and façade region growing

In the previous step, certain information about the intrusion positions was lost due to lack of edges. For example, the gradient values representing window frames were smaller than the threshold and thus, edges were not detected, or they were detected, but the frame edge did not intersect the area where points behind the façade plane were found.

To retrieve intrusion areas where insufficient edges were detected, the area of obtained bounding boxes in the previous step, shown in red, was compared to the area of bounding boxes enclosing areas of *rastered points behind the façade plane*, as shown in green in Figure 19.

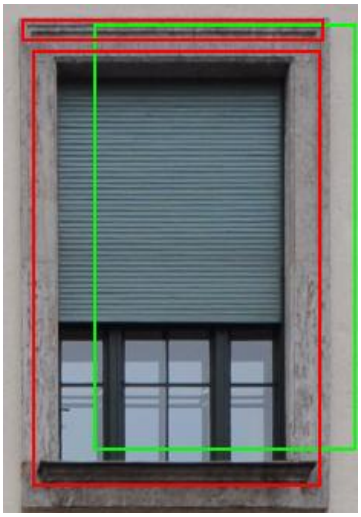


Figure 19: Comparison of the area of points behind the façade plane to the areas enclosing connected components

If the sum of all red bounding boxes (belonging to detected connected components of edges) within a green bounding box (belonging to *rastered points behind the façade plane*) was smaller than 25% of the green bounding box, then the red boxes were replaced by the green. In other words, if connected components of edges were insufficiently represented within an area with points behind the façade plane, then the bounding box of the latter was placed in this area. This resulted in additional intrusion areas, as marked in green in Figure 20. Red bounding boxes belong to the connected components from 3.3.1.

Figure 20: Replacement of small areas of connected components (red) by bounding boxes of points behind the façade (green)

Additional information was required to eliminate or at least reduce certain artifacts, such as oversize bounding boxes of *rastered points behind the façade plane* (1st green bounding box on the left side in Figure 20), oversize bounding boxes caused by edges not belonging to window frames, but connected to them (4th window upper floor in Figure 20), and false alarms of the *rastered points behind the*

façade plane (small green bounding boxes on the façade). Since the façade texture consists of very homogeneous and large wall parts, the binary mask representing wall segments were intersected with the bounding boxes.

3.3.3 Block-wise NCC refinement procedure for detection of additional windows in a floor

The assumption, that windows of the same kind can be found on the same floor, was the basis for the refinement procedure. A procedure that extracted missing window areas was carried out as shown in Figure 21.

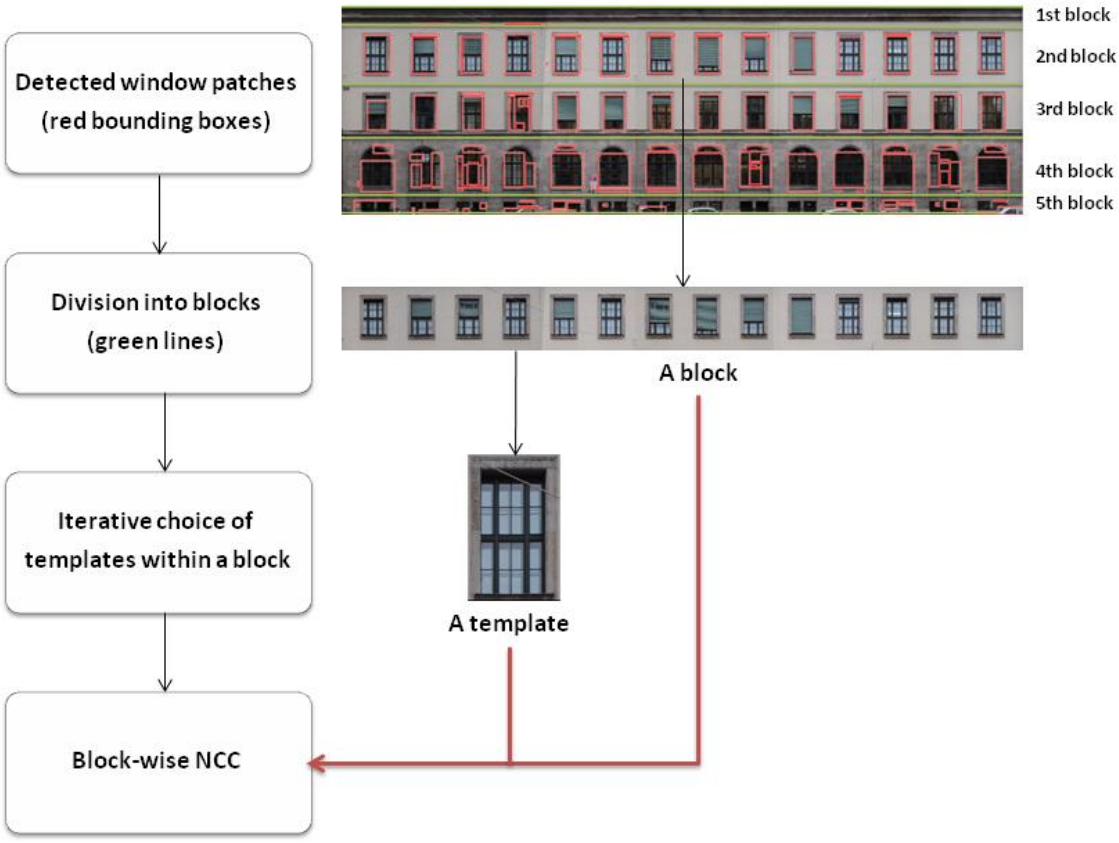


Figure 21: Refinement procedure with block-wise NCC

First, all detected window areas (patches) from the previous step were divided into blocks, described in 3.3.3.1. Then maximally three templates, that were not similar within each other, were found per block, shown in 3.3.3.2, and the NCC was performed with these templates along the respective block. A rectangle of a size of the template was created at restricted number of positions of a maximum that was within a threshold, described in 3.3.3.3.

3.3.3.1 Division into blocks

Assuming that similar windows are arranged into floors, previously extracted window patches were separated according to minimum positions of the vertical profile to further explore their similarity. The vertical profile is the result of row-wise sums of pixels of the binary image of extracted window patches obtained in 3.3.2, shown in green in Figure 22. Red dividing lines are positions of minima that separate blocks. The separation into blocks restricted the area on a façade texture where a search for similar windows was performed. Windows were thus investigated on each level separately.

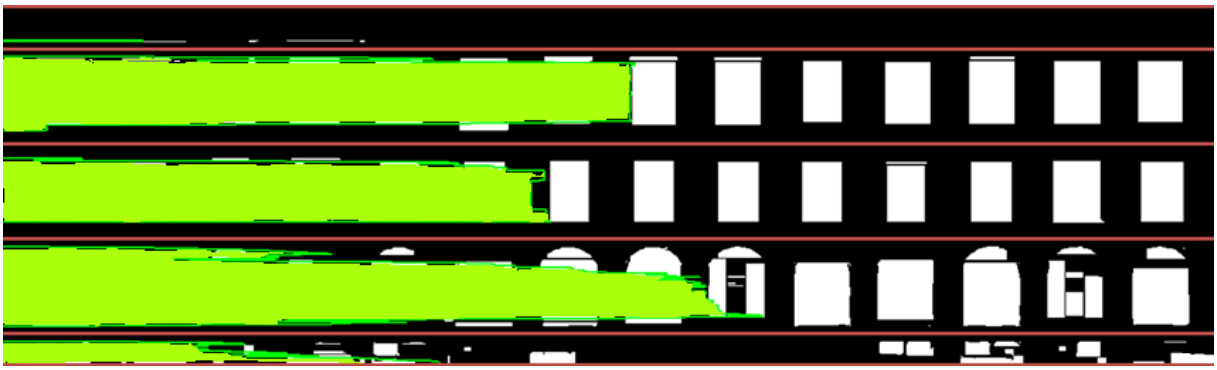


Figure 22: Division into blocks

3.3.3.2 Iterative procedure to find the proper templates for the block-wise NCC

The extracted window patches were assigned image information from façade textures at the same position and were thus potential candidates for templates being used in the block-wise NCC. Performing NCC with templates not representing intrusions would result in a large false alarm rate. Therefore, an iterative procedure to find proper templates for the block-wise NCC was performed, as seen in Figure 23.

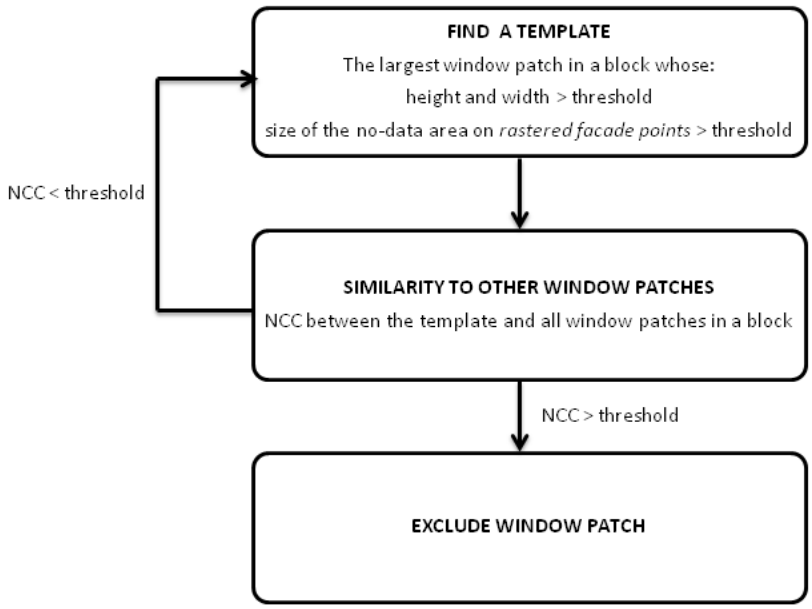


Figure 23: Search for proper templates from a block

The sizes of extracted window patches were compared within a block. The largest window patch in a block that was higher and wider than a threshold was accepted as the template. In addition, a template, which could later be used for the NCC along the whole block, had to have sufficient size of the no-data area on the raster belonging to the points of a façade. When the façade plane is extracted properly, there should be no data at the window positions (i.e., raster values are 0). As might be seen in Figure 24, window areas (red bounding boxes) are not entirely 0 due to limitations in point density, hits from the window bars, inaccuracy of the façade plane extraction, the large grid cells or oversize connected components around the window area. Therefore, only window areas (red bounding boxes in Figure 24) whose no-data size was more than a threshold were considered potential candidates for templates. This eliminated detected window areas containing a lot of wall parts (where most of the values were 1) from being possible templates.

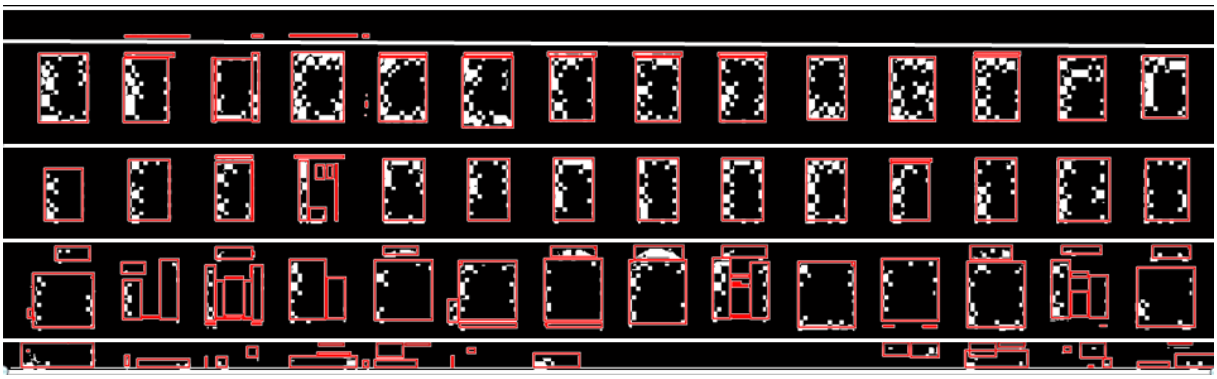


Figure 24: Rastered points belonging to the façade plane at the bounding box positions

Even if windows in one block are of the same type, they do not have to be necessarily similar according to the NCC similarity measure. In other words, differences in illumination, differently lowered blinds on each window, and different scenes behind windows can contribute to different grayscale values of window patches. Thus, performing the NCC with only one template per block might not detect the missing windows. Therefore, a search for more templates within each block was performed based on NCC similarity between window patches.

A template was already chosen from each block according to its size and the size of no-data area, seen in Figure 25.

Figure 25: Example of a template

The NCC was performed between the template (for example the one from Figure 27) and the other extracted window patches from the 3.3.2 (seen in Figure 28 as red bounding boxes) for each block separately. All the previously detected window patches (red bounding boxes in Figure 28) in a block that were sufficiently correlated to the template were excluded from being potential candidates for templates and not considered further. This assured that not very similar templates for the block-wise NCC were chosen per block.

Figure 26: A chosen template in each block is cross-correlated to all areas enclosed by red bounding boxes in a respective block

All window patches within a block which were not sufficiently correlated to the template went to the next iteration. The procedure was repeated. *A new template from each block was chosen. It was the largest of the remaining window patches in a block that met the same three criteria as before, relating to the width and the height of the bounding box along with the no-data area size when multiplied by the façade PC raster. The NCC between new templates and the remaining window patches was performed in each block separately.* The whole procedure was repeated once again.

3.3.3.3 Block-wise NCC

The templates obtained in 3.3.3.2 were cross-correlated to the entire block. The number of maxima in a block was limited to the number of previously found window patches in a block (from 3.3.2). If a maximum above a threshold was found, a rectangle of the same dimensions as the template was created around the maximum.

The results of the block-wise NCC were added to the extracted window patches from 3.3.2.

4 EXPERIMENT

4.1 Data description

In this study, three types of data were used. The existing 3D wire building model of the TUM main campus was added additional information about window positions extracted from a multi-aspect-oblique-view lidar PC and orthorectified façade textures. All codes to perform the methodology steps: *Co-registration of a lidar PC and a 3D wire frame model, Data pre-processing, and Window detection* were programmed using commercial software for technical computing Matlab. The region growing was performed using the image processing software Halcon.

4.1.1 Multi-looking oblique view airborne laser scanning PC

The area of acquisition was the main campus of the Technical University of Munich in the district Maxvorstadt, Munich. Lidar data of the TUM main campus was captured in an oblique way from four cross directions during the flight campaign in October 2006. Average height of the acquisition was 400 m. Point density of the PC was 9.7 points/m^2. Pre-processing, co-registration of stripes, and segmentation were done by Marcus Hebel for his Phd dissertation, described in Hebel (2012).

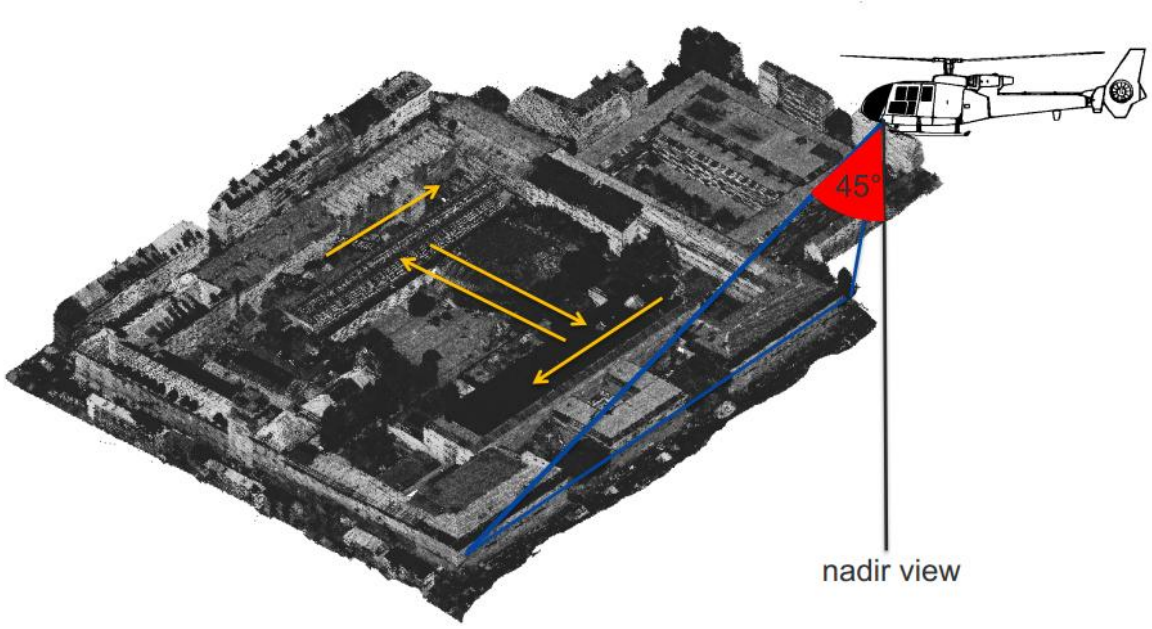


Figure 27: Acquisition geometry of a multi-looking oblique view airborne laser scanning PC

The speciality of this campaign was oblique view scanning geometry, i.e., the laser scanner was inclined at a 45° angle from the nadir view looking in a forward direction. This geometry of acquisition also enables the recording of the façade parts of buildings. In order to acquire the entire scene of the area, the area was overflown in 4 different directions as seen in Figure 28.

The lidar campaign was done by Fraunhofer IOSB with the instrument RIEGEL LMS-Q560 Version 2006. In the Table 3 its characteristics as given by Hebel, 2012 can be found.

Measurement principle	time-of-flight measurement
Wavelength	near IR
Pulse repetition rate	100 kHz
Beam divergence	0.5 mrad
Distance accuracy	2 cm
Scanning mechanism	Rotating polygonal mirror
Angular accuracy of the mirror	0.001°
Scan lines (scan angle is 60°)	100/s
Digitalisation of a signal	Full Wave Form (resolution 1s)
GNSS/INS sensorsystem frequency	200 Hz

Table 3: The lidar instrument RIEGEL LMS-Q560 version 2006 characteristics (Hebel, 2012)

4.1.2 Textures

Visible images that represent the basis of the façade textures were acquired by the camera model Nikon D3 with a focal distance of 24.4 mm. The experiment was carried out on eleven existing façade textures (Tuttas & Stilla, 2013).

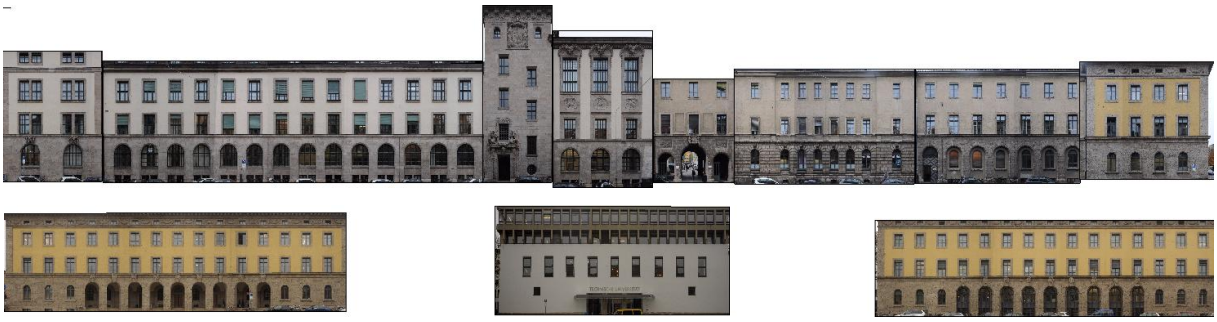


Figure 28: Rectified image textures

4.1.3 The 3D building model

Gauss-Krüger coordinates of vertices representing polygons were manually acquired from aerial images using 3D StereoAnalyst in ArcGIS and generalized to LOD2. Vertices of polygons were sequentially ordered and each polygon's first and last point were differently marked.

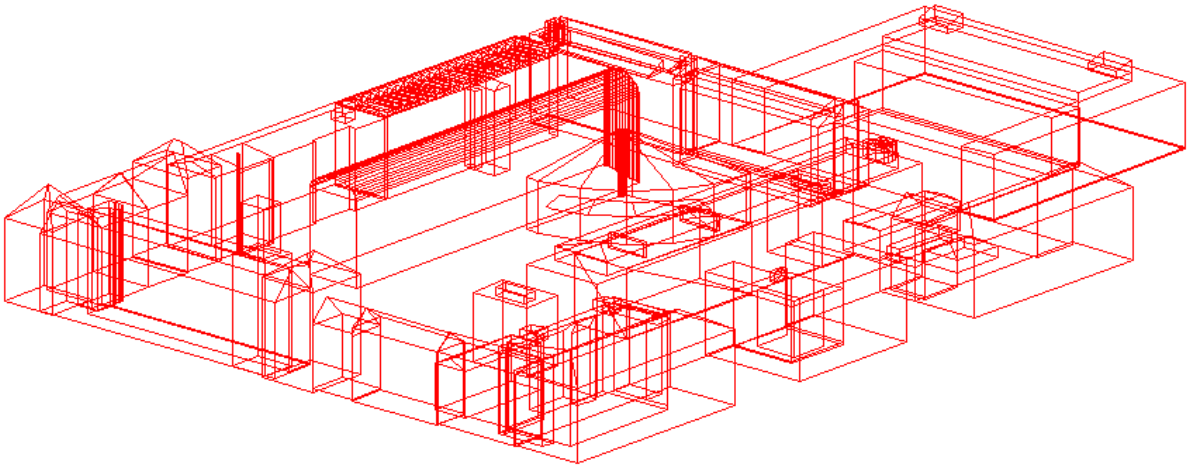


Figure 29: 3D wire frame building model of the TUM main campus area

4.2 Test

The methodology as described in 3.2 was applied.

Co-registration of a lidar PC and a 3D wire frame model

The transformation of the 3D vertices of the building model were transformed using Matlab Toolbox for geodetic transformations from Wasmeier (2011). Plane extraction and a search for corresponding pairs in both datasets were done by Method 1 and Method 2. Method 1 was done by me and Method 2 by Sebastian Tuttas. The second one resulted in better co-registration and therefore it was used in the following steps. Transformation parameters were calculated and applied to the PC using rigid transformation. The 3D building model and the PC were co-registered. All points from the PC that were within the bounding box—1 m in front and 3 m behind a polygon plane—were extracted for each plane of interest. In this work we dealt with 11 façade planes. Only points more than 10 cm behind the extracted PC planes were chosen as possible candidates for window points. Whether points lying behind a window are included depends on façade roughness—the façade is either smooth or has many ornaments, protrusions, or intrusions. More can be found in Tuttas & Stilla (2011). Figure 30 shows the extracted points behind façade planes.

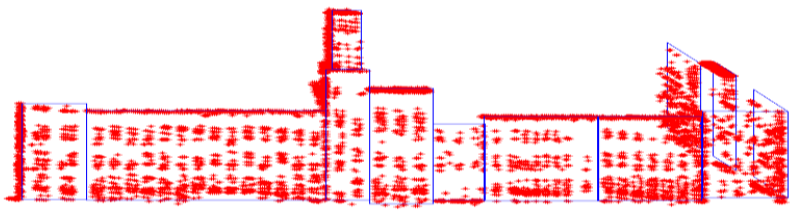


Figure 30: Extracted points behind the façade plane

Then the projection of points onto their corresponding planes was done considering the oblique-view geometry.

Data pre-processing

Rasters were produced both from *points lying behind façade planes* with the cell size of 0.5 m and from *points belonging to façade planes (extracted planes)* with the cell size of 0.2 m. The reason for different cell size of both kinds of rasters—*points behind a façade plane* and *points belonging to a façade plane*—lies in different usage of their information. *Rasters of points behind façades* were created to be used for window detection and thus, grid cells forming an intrusion area had to be as connected as possible (large enough). However, the size of a grid cell had to be sufficiently small to prevent connecting intrusion areas that were next to each other. The size of no-data areas of the *rasters of points belonging to façades* were used as a parameter to choose templates for the block-wise NCC. If the size of a grid cell was set to more than 0.2 m, the size of no-data area on a raster would be smaller, and thus, not so distinguishable from other areas on the raster. On the other hand, a smaller size of a cell grid would lead to a lot of gaps due to the low point density which is cca 10 pts/m^2.

Edges were extracted using Sobel edge detector from façade textures. Morphological operator opening was used to eliminate all objects smaller than 100 pixels on all the textures. In the next step, dilation was used to enhance the remaining edges, mostly belonging to the window frames. Very long edges in vertical and horizontal direction were reduced by thresholding gradient profiles. The threshold was set to 50% of the image width and 50% of its height. However, this method could eliminate most of the edges of interest, if intrusions would be very densely represented, i.e., windows one next to another without separations. In our case, windows were not so densely represented on any of the analyzed façades, and therefore, elimination of edges whose sum exceeded 50% in a certain direction did not cause much loss of the edges belonging to window frames.

Prior to perform region growing for façade extraction, K-means in the CIELAB color space was attempted. In my case it has proven to be not very robust in façade color separation due to the proximity of colors belonging to façade and windows in the CIELAB color space, and the difficulty of defining the proper number of clusters.

To derive binary masks representing wall segments from images, region growing was performed. First, to avoid stopping the region from growing due to noise, images were smoothed using a median filter of size 19. Second, region growing was performed, comparing grayscale values of center points of 3-by-3 neighborhoods for a predefined threshold set to 1. If values of centers of adjacent rectangles

were within the threshold, they were merged together. Since the walls represent the major part of the textures, all segments smaller than 250,000 pixels were eliminated. This cut out smaller or heterogeneous parts which are, for example, windows or ornaments.

Window detection

Connected components of edges were extracted at the areas where *rastered points behind façade*s were found. Additional window areas, where edges were not sufficiently present, were obtained only from the *raster of points behind façades*. All extracted connected components of edges and *rastered points behind façades* were enclosed by bounding boxes. The obtained window patches were intersected with the binary mask produced by region growing. After performing this operation, windows were better separated and shapes of windows better defined. Certain windows were still connected by thin lines seen in Figure 31, resulting in not exact window areas. To eliminate these, the sum of vertical profiles was used, as seen in Figure 31. All sums in vertical and horizontal directions that were smaller than 2% of the entire image were eliminated, causing the lines to disappear.

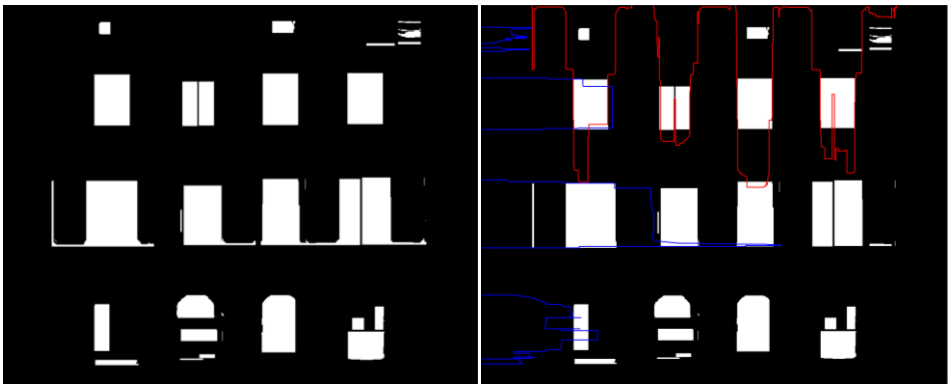


Figure 31: Line connections between window areas (left) and their elimination (right)

All these window patches were represented in a binary image for each façade separately. At this point, most of the window areas were already detected. A few of them were represented by many small bounding boxes instead of just one and a few windows were still missing. Thus the refinement procedure was performed. The window areas that were either undetected or being comprised of many extracted window patches, mostly had an unambiguous representative within already extracted window patches of the same window type in the same level. Therefore, separation into blocks did not cause division of windows, but rather restricted the area where a search for similar windows was performed. An iterative procedure to find proper templates from each block was carried out. The largest window patch in a block that was higher and wider than 100 pixels was accepted as the initial template. Additionally the window patch had to satisfy the condition that the size of no-data area on the *raster of points belonging to façade* was more than 60% of its size. This eliminated detected window areas containing a lot of wall parts (where most of the values were 1) from being possible

templates. Furthermore, the chosen template was compared to other extracted window patches within a block according to the value of the global maxima of NCC similarity. All window patches above 0.8 were excluded from being potential candidates for a template within a respective block. All the rest (that were below 0.8) went to the next iteration as being potential candidates for the 2nd template of the respective block. The whole process was repeated for another two times, resulting in at most three not completely similar templates per block. Each of these templates was cross-correlated to the respective block. The number of maxima in the result of the NCC between a template and its block was limited to the number of the extracted window patches. The threshold of maxima was set to 0.7. A binary raster was created for each template. A rectangle area of ones of the same dimensions as the template was created around each maximum that surpassed the threshold. This resulted in maximally three binary rasters per block. Before combining all results together, two problems had to be tackled.

First, in few cases the NCC similarity between a façade and a window template was high enough and thus, resulted in false alarms on the façade. Therefore, the results after performing the NCC were intersected with the extracted wall segments from the façade region growing that eliminated certain false alarms. In certain cases, the intersection of the result from the block-wise NCC and the façade area extracted using region growing resulted in very tiny or small areas. These areas were eliminated with an additional constraint: all areas smaller in height or width than 100 pixels and all areas whose width and height were both smaller than 200 pixels were eliminated.

Second, another refinement was required due to templates that represented windows and a part of the wall. If the distance between similar intrusions was smaller than half of the part of extracted wall area next to the window in a template, two windows were merged into one big area. If any of three templates produced an area, that was larger than 40% of the image width, the area was compared to the result of other two templates. If any of other two templates resulted in a smaller area, the large area was eliminated, otherwise it was preserved.

Since the intersection with the result of region growing produced not completely rectangular window areas, connected components of the areas were enclosed by bounding boxes, resulting in the final frames of window patches.

4.3 Results

The applied method resulted in well detected windows shown as red bounding boxes in Figure 32.

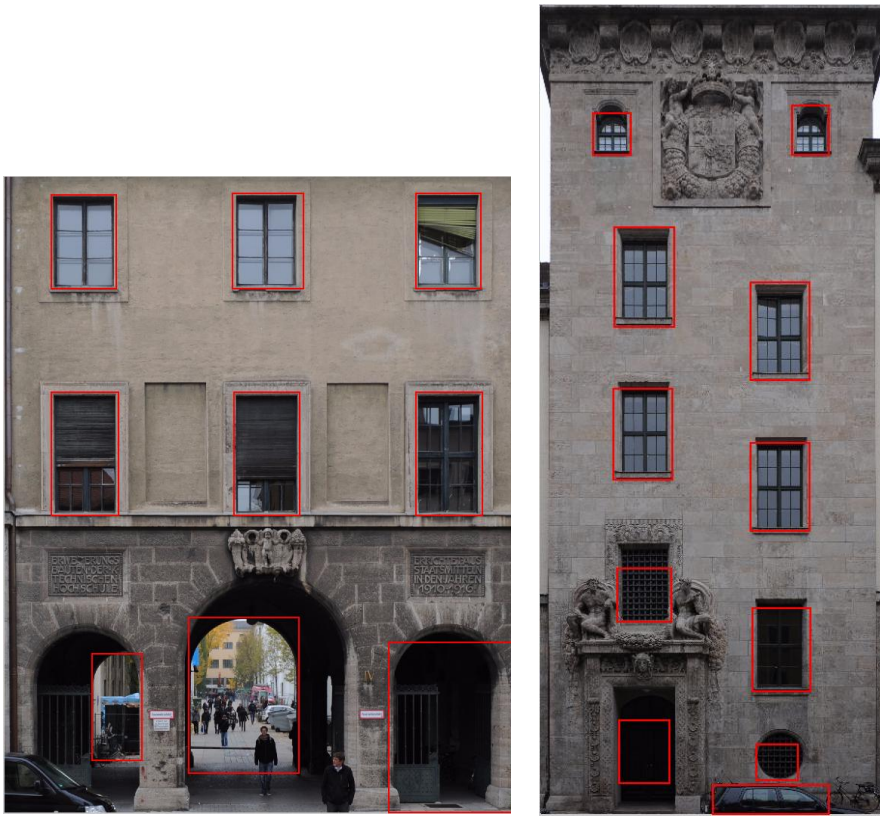


Figure 32: Results of detection

Not only were windows detected, but in most cases also the frames of windows were well defined. For most of the extracted window patches the basis were connected components of edges lying at the same position as *rastered points behind façade planes*. As can be seen in Figure 34, many bounding boxes enclose boundaries where intensity values noticeably change (edges). Additional reason for well defined window frames is the application of binary masks representing wall segments obtained by region growing. Most of the windows in Figure 35 are separated by homogeneous wall parts and were thus extracted by region growing, forming the binary mask. The intersection of binary masks with previously extracted window patches based on connected components of edges and *rastered points behind façades*, eliminated certain previously extracted artifacts on the wall parts and contributed to better defined window frames. Since *rastered points behind façade planes* were the basis for the detection step, other intrusions such as arcades, niches, and doors were also detected, as seen in Figure 34.

Figure 33: Detected windows despite cars on the texture

Furthermore, window frames were well defined even though there were cars on the texture, seen in Figure 33. Since *rastered points behind façade planes* were found also at these areas, connected components of edges found in these areas were extracted. In particular case, connected components of edges belong to cars and not windows. However, this can lead to a wrong window size. A problem might arise if the car is much larger than the window, in which case the resulting detected frame of a window would be too large and possibly shifted from the real window position.

The block-wise NCC proved to be useful, since it detected window patches whose evidence was not present in the data. Furthermore, in cases when a window was preliminary detected by many window patches, it resulted in better window patches, since a larger window patch was placed at the window position, if similarity was sufficient.

The objective of this work was to enrich an existing 3D building model. The final result is shown in Figure 34. The extracted window patches (areas) were textured to the model.

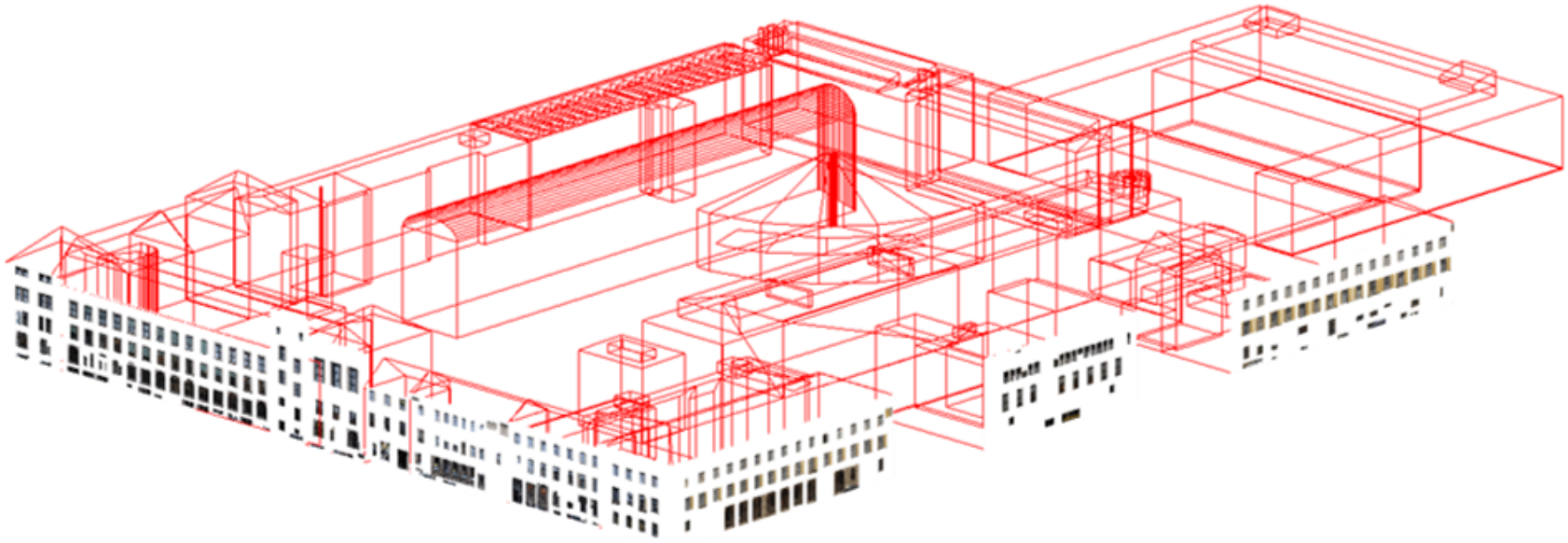


Figure 34: Enriched 3D building model with window areas

4.4 Limitations of the method

In the upper floor of buildings, sufficient points behind the façade might not be found. This might happen in cases where windows lie directly under the roof overhang and, due to the oblique view geometry, are not acquired. Consequently, there was no real window detected, as seen in Figure 35, and therefore this also could not be refined using the block-wise NCC.

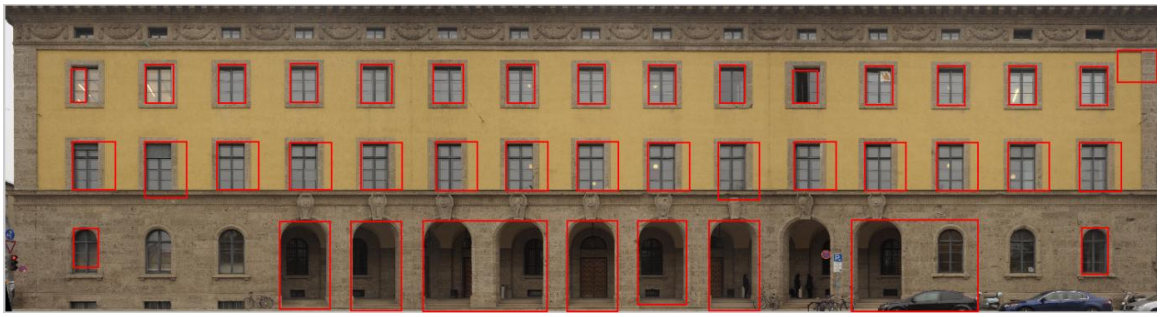


Figure 35: No extracted window patches in the upper floor and extracted window patches containing wall parts in the middle floor

A template chosen for the block-wise NCC, comprised not only of a window area, can lead to detection of window patches containing parts of non-window areas, as seen in the middle row in Figure 35. Due to the high NCC similarity between the template and window areas in the respective block, rectangles of the size of the template were placed at the positions of maxima. This resulted in window patches including not just window area, but also a part of a wall.

Very large window patches, containing many windows, can be extracted if windows are separated by wall parts containing bricks, as seen in Figure 36 . Boundaries of bricks result in high values of gradient connected to the edges representing window frames. If *rastered points behind a façade* are found at these areas, entire connected component of edges (including edges belonging to bricks) is extracted.

Furthermore, a high NCC value between a template and wall parts in a block can lead to extracted window patches on walls. This can be seen in the left and right upper part of the texture in Figure 40. The NCC was performed block-wise, and the template chosen in the upper block was very similar to the wall parts. This produced false alarms on the wall. Since the number of peaks in a block is limited, this prevented to detect more false alarms.

Figure 36: Extracted window patches containing many windows due to connected components formed by window frames and bricks

Oversize window patches, seen in Figure 37, can be extracted also due to niches which are at the same depth as windows in the façade; and therefore, they do not belong to the façade plane points, but to points behind the façade. Consequently, all connected components of edges within these areas are extracted, resulting in window patches not containing only window areas.

Figure 37: Oversize window patches due to niches

4.5 Comparison between windows detected from oblique view ALS and from fused data

Evaluation of window detection is not very straightforward, as it is sometimes hard to define the area of a window, depending on the required generalization. In other words, two windows 10 cm apart might be interpreted as one or two windows. In my case, a window was defined as an intrusion area with separated from other parts by a wall margin.

In Figure 38 the evaluation procedure is illustrated. The detected window was considered true positive if its bounding box was within the area of a window. If there were many bounding boxes within a window frame, only one was taken as the true positive, and the rest were considered false positives. Cases when two or more windows were detected by a bounding box were counted as false positives. Any detection on non-window areas of the façades was treated as a false positive. All undetected windows were accounted to false negatives.

Figure 38: Illustration of the evaluation method (TP=True positives, FP=False positives)

A comparison was done between the results of Tuttas & Stilla (2013) and the results of this method. The latter resulted in many basement windows being detected which were not considered in the evaluation of Tuttas & Stilla (2013). Therefore, in this chapter, an evaluation was done taking the same ground truth as in Tuttas & Stilla (2013), excluding all detections belonging to the level of basement windows.

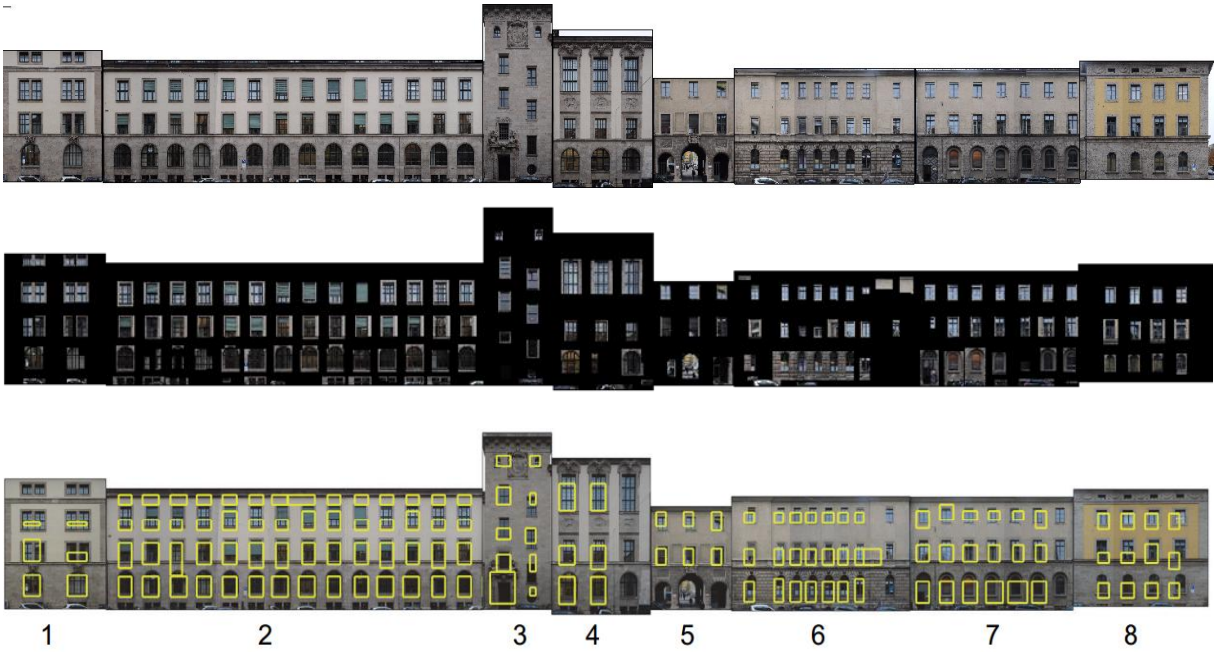


Figure 39: Façade textures (above), windows extracted using fused data (middle), and windows extracted using solely oblique view ALS from Tuttas & Stilla (2013) (below) – southern side of the TUM main campus

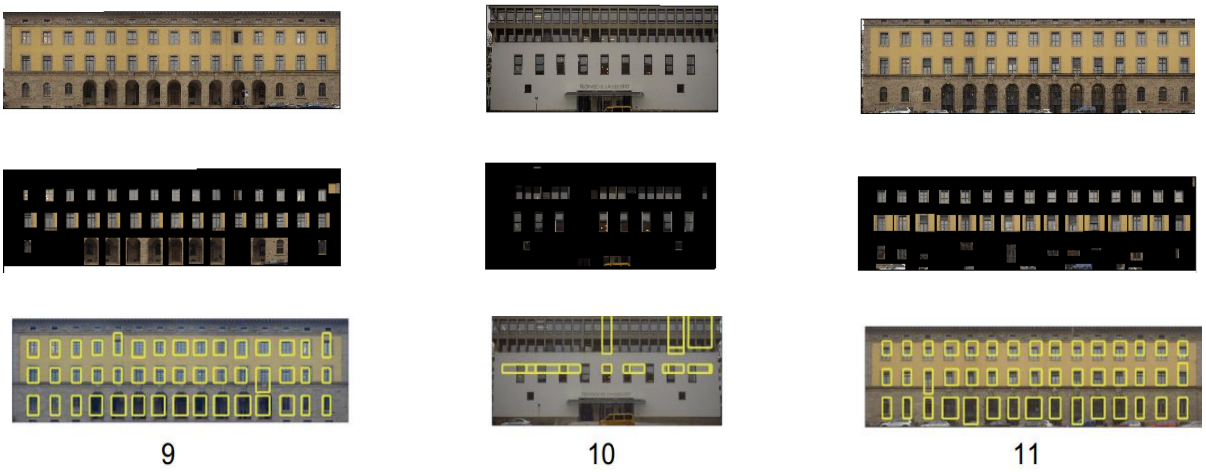


Figure 40: Façade textures (above), windows extracted using fused data (middle), and windows extracted using solely oblique view ALS from Tuttas & Stilla (2013) (below) – eastern side of the TUM main campus

Table 4: Evaluation of the method using solely oblique-view ALS from Tuttas & Stilla (2013)*

Number	1	2	3	4	5	6	7	8	9	10	11	Σ
GT	14	42	9	9	9	24	21	16	60	57	60	321
D	6	56	10	6	6	22	18	12	45	15	45	241
TP	2	42	9	6	6	21	18	12	45	2	45	208
FP	4	14	1	0	0	1	0	0	0	13	0	33
FN	12	0	0	3	3	3	3	4	15	55	15	113
DR	0.14	1	1	0.67	0.67	0.88	0.86	0.75	0.75	0.04	0.75	0.65
FR	0.67	0.25	0.1	0	0	0.05	0	0	0	0.87	0	0.14

Table 5: Evaluation of the method using fused data without considering basement windows*

Number	1	2	3	4	5	6	7	8	9	10	11	Σ
GT	14	42	9	9	9	24	21	16	60	57	60	321
D	14	49	10	10	9	21	21	12	40	30	45	261
TP	8	42	9	9	9	17	19	12	37	30	42	234
FP	6	7	1	1	0	4	2	0	3	0	3	27
FN	6	0	0	0	0	7	2	4	23	27	15	84
DR	0.57	1.00	1.00	1.00	1.00	0.71	0.90	0.75	0.62	0.53	0.70	0.73
FR	0.43	0.14	0.10	0.10	0.00	0.19	0.10	0.00	0.08	0.00	0.07	0.10

* GT=Ground truth, D=Detections, TP=True positives, FP=False positives, FN=False negatives, DR=Detection rate, FR=False alarm rate

The overall detection rate increased from 65% to 73%. What is more, the false alarm rate decreased from 14% to 10% in using the method as carried out in this work.

Eight façades out of eleven were better detected using the approach from this work. In addition, the window frames were in most cases better defined as in Tuttas & Stilla (2013), as seen in Figure 39 and Figure 40.

As reported in Tuttas & Stilla (2013) their method did not perform well on Façade 10 due to many windows lying very close to each other. The detection rate on this façade was improved from 4% to 53%. However, this façade resulted in the worse detection rate of all façades. The reason was not sufficient lidar points behind the façade plane in the upper part of the building.

On the other hand, the method from Tuttas & Stilla (2013) performed better on façades with large number of windows ordered in patterns that repeat along the façade, for example texture 9 and 11 in Figure 40. Furthermore, their approach also gives better results in case when bricks are connected to the window frames as, for example, is seen on texture 6 in Figure 39.

4.6 Evaluation

Another evaluation of detection was done considering all windows on the façade.

Figure 41: Detection results

Table 6: Evaluation of the method using fused data considering basement windows

Number	1	2	3	4	5	6	7	8	9	10	11	Σ
GT	18	70	10	15	9	32	27	16	66	61	66	390
D	16	62	11	13	9	24	24	12	40	30	48	289
TP	8	47	10	9	9	17	21	12	37	30	44	244
FP	8	15	1	4	0	7	3	0	3	0	4	45
FN	11	23	0	6	0	15	7	4	29	31	22	148
DR	0.44	0.67	1.00	0.60	1.00	0.53	0.78	0.75	0.56	0.49	0.67	0.63
FR	0.50	0.24	0.09	0.31	0.00	0.29	0.13	0.00	0.08	0.00	0.08	0.16

GT=Ground truth, D=Detections, TP=True positives, FP=False positives, FN=False negatives, DR=Detection rate, FR=False alarm rate, Number= Façade ID

This resulted in an overall detection rate of 63% and a false alarm rate of 16%. It should be noted that many bounding boxes representing two windows lying very close to each other were evaluated as false positives. Furthermore, many windows were detected by multiple bounding boxes, in which case one box was counted as true positive and the others as false positives.

5 CONCLUSION AND FUTURE WORK

The main objective of this thesis of automatically extracting windows from fused data was achieved. Combining high resolution visible façade textures and oblique viewing ALS outperformed the approach using only data from oblique viewing ALS. The detection rate increased from 65% to 73%, and the false alarm rate decreased from 14% to 10%. The shape and the area of most of the detected windows were well defined, since they were mainly based on connected edges forming a window frame. Although cars and certain occlusions were on the textures, windows on the ground floor were detected. However, these artifacts can lead to a wrong window size. This issue could be resolved by either using only oblique-view information for the window detection in the ground floor, or by first, classifying façade textures into buildings, cars, trees, similar to the one described in Yang & Förstner (2011) and then, adapt the window detection method for not building areas.

Despite the varied composition of windows and different window sizes between the textures, the overall detection rate was 63%. The approach used in this work is mainly data driven. Windows were either detected due to the supporting evidence from the data or found because of a high normalized cross-correlation to a template based on the data evidence. No probability of propagation was included in the work, and thus the fact that urban areas mainly consist of more or less repetitive structures or symmetries should be further exploited. For example, detected windows could form a grammar for procedural modeling. Furthermore, a database could be created from correctly detected windows.

The algorithm performed poorly on façades with bricks. This could possibly be resolved by changing the resolution and eliminating high frequencies on the image, so that the edges from the bricks were not detected anymore. Another solution would be to detect brick areas and at these areas use only information from the oblique-view PC for window detection. Tuttas & Stilla (2013) obtained exact window positions and well defined window frames at bricks' areas using only oblique-view PC.

More specific window frames can be derived from the detected window patches using Hough transform, as seen in Figure 42. In the present case, edges were extracted using Canny, and the accumulator array of the Hough transform was limited to the angles that formed approximately horizontal and vertical lines. Adding certain topological or connectivity constraints might lead to complete window frames. Finding horizontal and vertical lines using Hough on edges detected by Sobel proved to complement those obtained using Canny edge detection.

Figure 42: Hough lines restricted to approx. vertical, horizontal directions at detected window patches

Future work might attempt in the reconstruction step to classify windows in arc-shaped and rectangular windows, following the approach from Lee & Nevatia (2006) or other parametric shapes using Hough transform generalized to arbitrary shapes (Antolovic, 2008). More semantic information could also be attached to the detected intrusions, such as distinguishing between doors, windows and niches.

The accuracy of data co-registration has to be high to assure window detection. Therefore, simultaneous acquisition of different data types would create fewer registration problems due to changes in scenes and observation conditions.

It should be noted that we dealt with very high resolution images without many artifacts—such as occlusions, shadows and changes in illumination—and therefore, the Sobel edge detector performed well. Also, region growing, as carried out in this work, would probably fail unless the façade grayscale values on images were close to each other.

However, it has been proven that there is great potential in combining high resolution visible façade textures with oblique multi-view airborne laser scanning. With improvements in lidar point density, the detection rate would also improve. Furthermore, one could exploit full-waveform information of façades to define façade roughness.

6 RAZŠIRJEN POVZETEK V SLOVENŠČINI

6.1 Uvod

V devetdesetih letih prejšnjega stoletja je aerolasersko skeniranje (ALS) prišlo tudi v komercialno rabo in omogočilo zajem podatkov večjih območij, kot so npr. mesta. Zaradi nadirnega pogleda ALS, je bilo skeniranje fasad stavb do nedavnega onemogočeno. Večstransko ALS poševnega pogleda (multi-aspect-oblique-view ALS) je poseben primer ALS, pri katerem je laserski skener usmerjen pod kotom 45° glede na nadirni pogled. S tem je omogočen zajem podatkov o fasadah stavb iz zraka. Za pridobitev popolne informacije o obravnavanem območju (objektu), je potrebno skeniranje v vseh štirih glavnih smereh. Novost metode, katera je predstavljena v tej diplomski nalogi, je program, kateri na osnovi kombinirane uporabe podatkov ALS poševnega pogleda in ortorektificiranih vidnih podob fasad omogoča samodejno zaznavanje oken.

6.1.1 Motivacija

Zaradi intenzivnega povečanja prebivalstva v urbanih okoljih se je močno povečal pomen izboljšanih 3D modelov stavb. Po podatkih Svetovne zdravstvene organizacije bo do leta 2050 70% ljudi živelo v mestih, kar predstavlja veliko izzivov na področjih kot so: prostorsko načrtovanje, nadzor varnosti v mestih, načrtovanje ukrepov in ugotavljanje škode ob naravnih nesrečah ter sonaravna gradnja objektov.

Veliko podatkov različnih senzorjev je že dostopnih, vendar večina postopkov za njihovo obdelavo in interpretacijo še vedno ni popolnoma avtomatizirana. Kombinirana obdelava različnih vrst podatkov lahko prispeva k izboljšanju zaznavanja in avtomatski interpretaciji obravnavanega območja, kar je osnova za izdelavo detajlnih 3D modelov.

V Evropski uniji porabijo stavbe okoli 40% vse energije (Directive 2010/31/EU, 2010). Izboljšave fizikalnih karakteristik objektov so nujno potrebne za zmanjšanje energetske odvisnosti Evropske unije in zmanjšanje izpustov toplogrednih plinov. Iz ALS in slikovnih podatkov lahko pridobimo geometrične podatke o fasadah, npr. položaje okvirjev oken, ki jih nato hipotetično lahko koregistriramo z infrardečimi posnetki kot npr. v (Avbelj et al., 2010), kateri prikazujejo toplotne izgube stavb. S tem omogočimo boljšo interpretacijo infrardečih posnetkov za nadaljnje oblikovanje zakonodaje na področju toplotnih izgub stavb.

Pomembnost detajlnih 3D modelov se kaže tudi na področju ocene škode, katero povrzročijo naravne nesreče. Postopek za samodejno zaznavo oken bi lahko bil hipotetično uporabljen v oceni škode stekla, katera je bila povzročena zaradi padca meteorita v Rusiji (NY Times, 2013).

6.1.2 Sorodna dela

Urbana okolja so običajno sestavljena iz ponavljajočih in simetričnih struktur, kar se lahko s pridom izkoristi tudi v avtomatski rekonstrukciji fasad. V strokovni literaturi (Recky & Leberl, 2010; Kulkarni, et. al., 2011) je velikokrat uporabljena metoda projekcije gradientov (Lee & Nevatia, 2004), katera razdeli fasadno rektificirano sliko v pravokotno mrežo glede na prisotnost robov v vertikalni in horizontalni smeri.

Objavljenih je bilo že veliko postopkov rekonstrukcije fasad, ki se med sabo razlikujejo glede na tip uporabljenih podatkov. Lee & Nevatia (2004) sta predstavila samodejni postopek na rektificiranih posnetkih fasad, ki ne le da zazna okna, ampak jih tudi klasificira v arkadna (okna z lokom) in pravokotna okna ter določi njihov položaj v prostoru. Recky & Leberl (2010) trdita, da prej omenjeni postopek zaznave oken ni primeren za kompleksne fasade z veliko detajla, saj z uporabo gradientov zaznamo tudi detajle, kateri ne pripadajo oknom. Predlagata, da se bloki, razdeljeni po gradientni metodi, dodatno klasificirajo na okna in zid z gručenjem po metodi K-povprečja v barvnem prostoru CIELAB.

Dejstvo, da laserski žarek večinoma prodre skozi steklene površine in posledično ni odbojev od ravnine fasade, so izkoristili pri rekonstruiranju zunanjih okvirjev oken v več raziskavah, kot npr. Becker (2009), Pu & Vosselman (2007), Wang, et al. (2012).

V zadnjih letih se pri rekonstrukciji fasad vse bolj uveljavlja mobilno lasersko skeniranje. Wang, et al. (2012) predstavijo metodo, ki temelji na precej grobih domnevah. Ne upošteva namreč oken v pritličju, zaradi ovir, ki zakrivajo pročelja v tem delu. Tudi višina in širina oken ter presledki med okni so posebej določeni za vsako fasado, kar onemogoča popolno avtomatizacijo postopka. Poročajo tudi o pomanjkanju podatkov v zgornjih delih višjih stavb, zaradi kota snemanja mobilnega laserskega sistema.

V podatkih terestričnega in mobilnega laserskega skeniranja velikokrat pride do pomanjkanja podatkov na delih fasad. Vzrok temu so razna zakritja zaradi ovir, omejitve zaradi kota skeniranja, oz. spremenljiva gostota oblaka točk. Le te onemogočajo avtomatizacijo postopka rekonstrukcije oken z le geometričnimi entitetami izločenimi iz oblaka točk. Becker (2009) predlaga, da se geometrični gradniki izločeni iz oblaka točk izboljšajo z informacijo s podob. Takšni podatki tvorijo osnovo za

definicijo fasadne slovnice (angl. façade grammar). Fasadna slovnica sestoji iz hierarhije, vzorcev ponavljanja in produkcijskih pravil o tem, kako naj bi se izločeni primitivi razporejali po fasadi. To omogoča določitev oken tudi v predelih, kjer ni podatkov. Vendar pa zgolj statistično prenašanje določenih vzorcev ne vodi vedno do prave rešitve.

Prej omenjene omejitve terestričnega in mobilnega skeniranja se lahko odpravi z uporabo podatkov večstranskega ALS poševnega pogleda, ki med drugim omogoča še hitrejši zajem večjega območja. Zaradi relativno nizke gostote točk (manj kot 10 točk/m^2), je rekonstrukcija oken v območjih brez podatkov v ravnini fasade otežena. Posledično Tuttas & Stilla (2011) za zaznavo oken predlagata uporabo točk, ki so na notranji strani segmentiranih fasad. Le te, projecirane na fasado, podajo dovolj zanesljive podatke o položaju vdolbin in steklenih površin na fasadi. Če so predhodno zaznani ponavljajoči se vzorci na fasadi, se za določitev razmakov med okni uporabi Fourierjeva transformacija. V Tuttas & Stilla (2012) predstavita tudi metodo rekonstrukcije oken. Točke, ki so bile predhodno zaznane kot okna, so izhodišče za iterativni postopek, kateri išče najbližje točke v vsakem kvadrantu na ravnini segmentirane fasade. Položaj izbrane stranice okvirja okna je določena z vsoto funkcij gostote verjetnosti. Metoda zavisi od prisotnosti točk za fasadno ravnino in zaznanih ponavljajočih vzorcev ter predpostavlja pravokotno obliko oken. Bolj zanesljiva in natančna metoda zaznavanja oken bi lahko temeljila na kombinirani uporabi in obdelavi podatkov večstranskega ALS poševnega pogleda in vidnih podob, kar je tudi glavna tema te diplome.

6.1.3 Opis problema in cilji raziskave

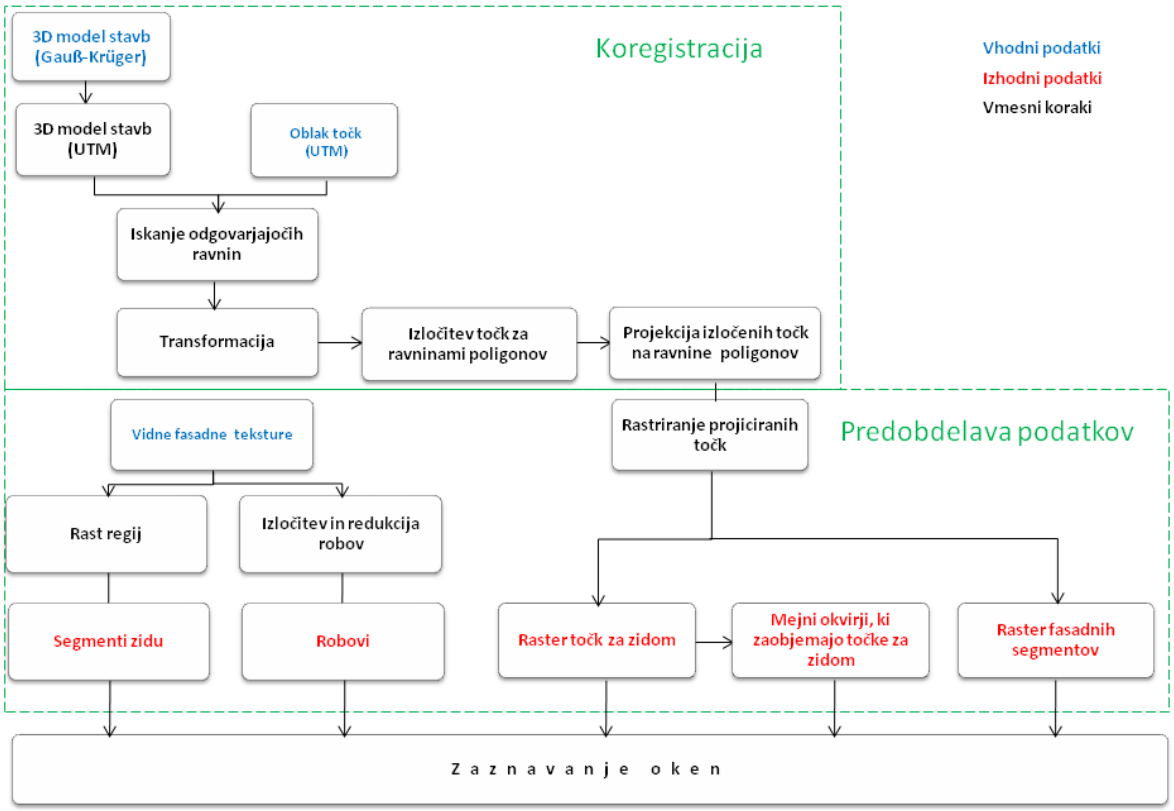
3D modeli stavb so že dostopni za določena mesta, vendar so v njih stavbe večinoma predstavljene le kot kvadri ali teksturirani modeli stopnje podrobnosti detajla 2 (angl. LOD 2) skladno s CityGML. V času, ko je dostopna velika količina podatkov, zajetih z različnimi senzorji, se poraja zahteva po samodejni semantični interpretaciji, katere namen je obogatitev 3D modelov stavb. Glavni cilj te diplomske naloge je samodejna pridobitev detajlne geometrije fasad (oken in vdolbin) iz kombiniranih podatkovnih tipov za izboljšanje obstoječega 3D žičnega modela stavb, kateri je stopnje podrobnosti detajla 2. Dodatno raziskovalno vprašanje je, ali uporaba kombiniranih tipov podatkov (fasadne teksture, aerolaserski oblak točk) privede do boljših rezultatov zaznave oken, kot uporaba samo enega tipa podatkov (aerolaserki oblak točk), kar je opisano v Tuttas & Stilla (2013).

Pred zaznavanjem oken iz kombiniranih vrst podatkov (vidne fasadne teksture in aerolaserski oblak točk poševnega pogleda) smo samodejno koregistrirali 3D žični model stavb in aerolaserki oblak točk. ALS poševnega pogleda namreč prodre skozi steklene površine in posledično pride do odbojev za fasadno ravnino. Te točke projicirane v smeri geometrije skeniranja ALS poševnega pogleda predstavljajo potencialna mesta oken. Na območjih okenskih okvirjev (stik med zidom in steklom) na

fasadni teksturi pride do spremembe intenzitetnih vrednosti. Posledično lahko predpostavimo, da robovi na območjih, kjer so bile zaznane aerolaserske točke za fasadno ravnino in kasneje projecirane na fasadno ravnino, pripadajo oknom. Osnovo za zaznavanje okenskih površin predstavljajo točke aerolaserskega oblaka locirane na notranjih straneh ravnin fasad in povezane komponente robov, katere so izločene iz fasadne podobe. Za izboljšanje postopka smo izkoristili predpostavko, da je v istem nadstropju več oken istega tipa. Prej zaznane površine oken so služile kot vzorčne matrike v postopku normalizirane navzkrižne korelacije, katero smo uporabili za zaznavanje preostalih okenskih površin po posameznih nadstropjih. Rezultati pridobljeni z opisano metodologijo so bili ovrednoteni na podlagi referenčnih lokacij oken.

6.2 Metodologija

Razvit postopek sestoji iz treh glavnih delov: koregistracije oblaka točk in 3D žičnega modela stavb, predobdelave podatkov in postopka zaznave oken. Shema 1 prikazuje vmesne korake koregistracije in predobdelave podatkov, vključno z vhodnimi podatki (označeni modro). Uporabili smo tri vrste podatkov istega območja in sicer: 3D žični model stavb, aerolaserski oblak točk poševnega pogleda z več strani in vidne fasadne teksture. Rdeče so obarvani podatki, ki so bili rezultat koregistracije in predobdelave ter so služili kot vhodni podatki za zaznavo oken.



Shema 1: Potek korakov koregistracije in predobdelave podatkov

Koregistracija

Najprej sem 3D model transformirala iz Gauss-Krügerjeve projekcije v UTM z Matlab kodami za geodetske transformacije (Wasmeier, 2011). Po spremembi projekcije je bilo odstopanje med 3D modelom stavb in oblakom točk preveliko za zaznavo oken. Vzrok temu so bili uporabljeni transformacijski parametri 7-parametrične transformacije za celotno Bavarsko, natančnost georeferenciranja in koregistracije ALS podatkov ter razlika v geometriji obeh vrst podatkov. 3D model stavb je namreč zajet iz zračnih posnetkov in posledično so fasade poravnane z robovi streh, medtem ko ALS podatki predstavljajo odboje od dejanske ravnine fasade. Vidne fasadne teksture so že bile na 3D modelu, zato je bila potrebna le koregistracija 3D modela in oblaka točk.

Avtomatsko koregistracijo 3D žičnega modela in oblaka točk sva izvedla s Sebastianom Tuttasom po dveh različnih metodah v štirih korakih:

- izločitev ravninskih segmentov iz oblaka točk
- iskanje korespondenčnih ravnin (ravninskih segmentov in korespondenčnih poligonov) v obeh podatkovnih setih
- vzpostavitev matematičnega modela med korespondenčnimi ravninami iz obeh podatkovnih setov (Hebel & Stilla, 2011) in določitev transformacijskih parametrov
- 6-parametrična transformacija.

Moja metoda koregistracije se je izkazala za manj natančno, zato je bil za nadaljno obdelavo uporabljen transformiran oblak točk Sebastiana Tuttasa (Tuttas & Stilla, 2013).

Predobdelava

Laser prodre skozi steklo in posledično pride do odboja znotraj stavbe. Te točke, v kolikor so projicirane na fasado glede na geometrijo zajema, podajo položaje oken in drugih vdolbin. Za lažjo obdelavo neorganiziranega oblaka točk je bilo potrebno točke, katere so bile za izločeno fasadno ravnino, projicirati na ravnino poligona glede na geometrijo zajema in nato rasterizirati. Tako sem pridobila *raster točk za fasadno ravnino*. Dodatno se je vsako povezano komponento na tem rastru zaobjelo z mejnim okvirom (pravokotnik najmanjšega obsega okoli določene površine).

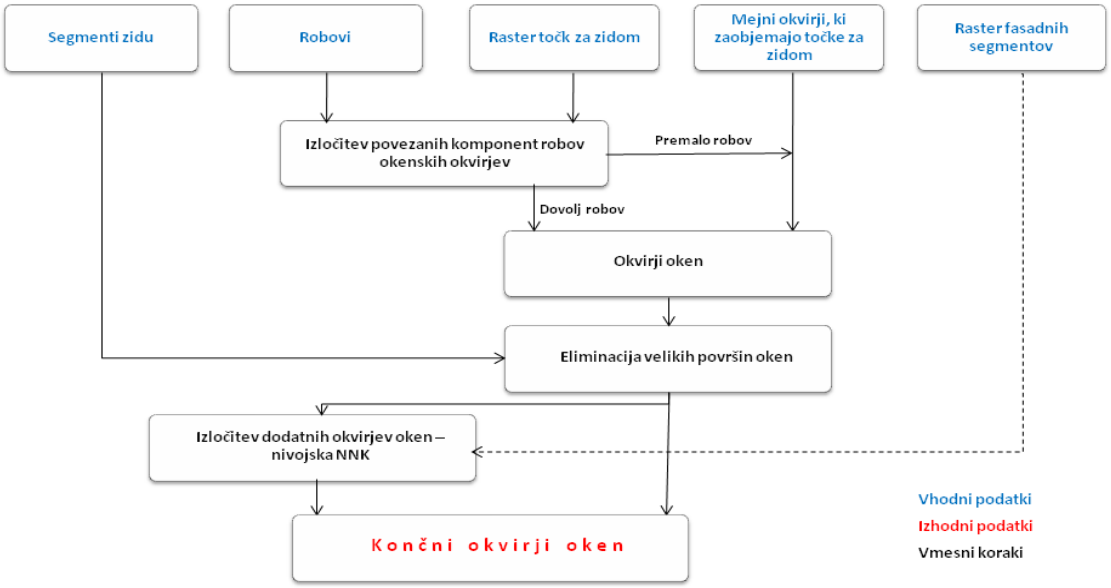
Na območju oken na ravnini fasade po navadi ne pride do odboja laserskega žarka; zato le ta ostanejo brez vrednosti in to dejstvo se lahko izkoristi za dodatno informacijo o položaju oken. Narejeni so bili posebni rastri iz točk, ki so pripadale izločenim fasadnim ravninam. Ta raster se v nadaljevanju imenuje *raster fasadnih točk*.

Na fasadnih teksturah je zid predstavljal večje, homogene površine v primerjavi z okni. To dejstvo sem uporabila pri segmentaciji fasadnih tekstur. Segmentacijo sem izvedla z rastjo regij, ki temeljijo na primerjavi intenzitetnih vrednosti sosednjih regij. Segmenti, katerih velikost je bila nad določenim pragom, so bili uporabljeni za nastanek binarne maske velikosti fasadne teksture z vrednostmi 0. Segmenti, katerih velikost je bila pod pragom, so imeli vrednost 1.

Položaj oken se dokaj zanesljivo pridobi z *rastra točk za fasadno ravnino*. Po drugi strani pa je, zaradi rasterizacije in nizke gostote točk, pridobitev podatkov o položajih okenskih okvirjev iz samo tega vira podatkov otežena. S tem namenom sem dodatno informacijo o okvirjih oken pridobila iz vidnih fasadnih tekstur. Velike spremembe intenzitetnih vrednosti med zidom in okvirji oken omogočajo izločitev robov na teh mestih. Po drugi strani pride do spremembe intenzitentih vrednosti tudi na drugih mestih fasadne teksture. S tem namenom sem robove po izločitvi s Sobelovim filtrom še dodatno obdelala. Morfološki operatorji so bili uporabljeni za eliminacijo majhnih robov, ki so bili posledica detajlov na fasadi. Problem so predstavljali tudi dolgi robovi vzdolž horizontalne in vertikalne smeri fasade, ki so se dotikali oken in so bili posledica sprememb v barvi fasade med nadstropji. Eliminirala sem jih z uporabo pragov gradientnih profilov. To so vrstične vsote robov v vertikalni smeri in stolpične vsote robov v horizontalni smeri. Večina fasad ima določen razmik med okni in posledično manjšo vsoto robov v določeni smeri, kot npr. rob, ki poteka po celotni dolžini ali višini fasade.

Zaznavanje oken

Shema 2 prikazuje glavne korake v zaznavi oken. Podatki, pridobljeni v predobdelavi, so bili uporabljeni za zaznavo oken (označeni modro). Iskanje in uporaba informacij iz različnih virov podatkov je bila mogoča, saj so bili ti med sabo koregistrirani in so bili enake ločljivosti.



Shema 2: Potek zaznave oken

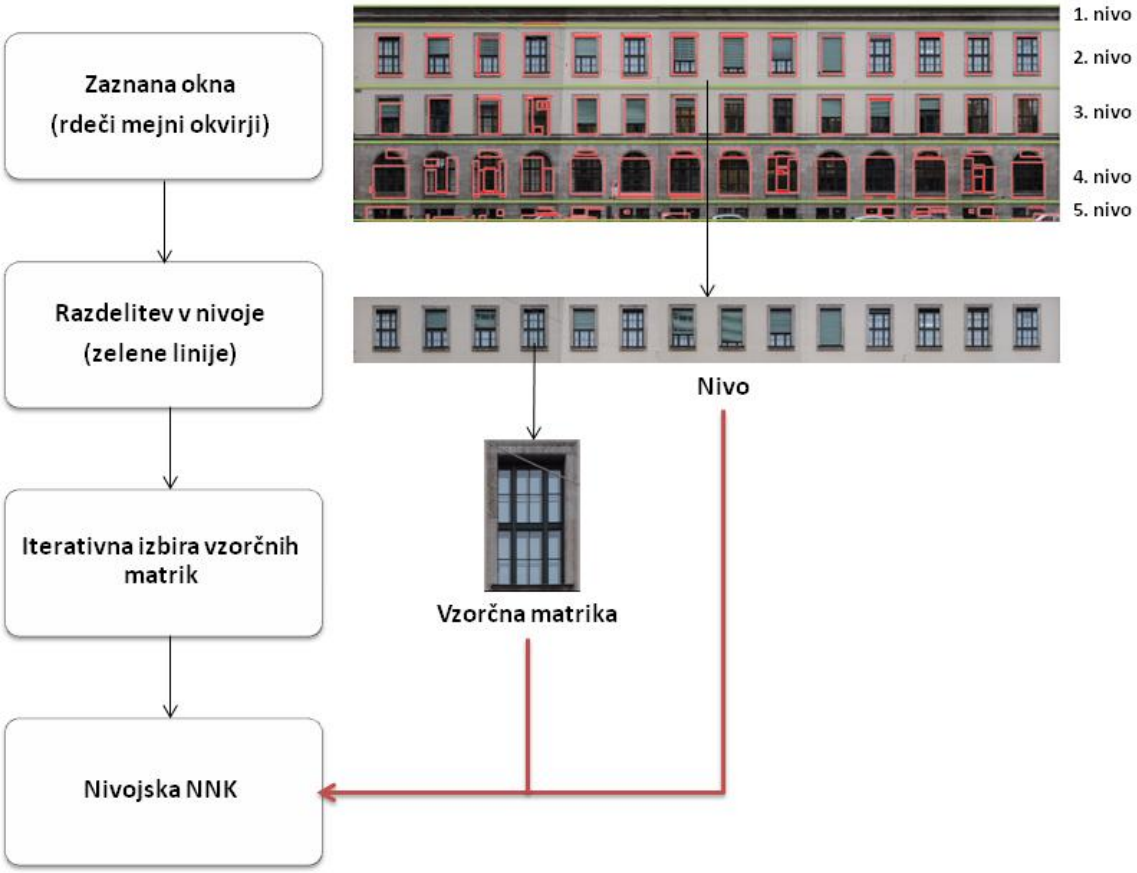
Kot že omenjeno, informacijo o položaju oken podajajo *rastri točk za fasadno ravnino*, medtem ko lahko bolj natančne okvirje oken pridobimo iz izločenih robov. Za zaznavo začetnih okvirjev oken sem združila ti dve informaciji in izločila povezane komponente robov, katere so bile na položaju točk za zidom v rastru. To pomeni, da je moral biti dokaz za prisotnost okna na določeni poziciji, tako na izločenih robovih, kot tudi na *rastru točk za fasadno ravnino*. Za večjo zaokroženost okenskih regij sem vsako zaznano povezano komponento robov omejila z mejnim okvirjem (rdeči mejni okviri na sliki 1).

V primeru pomanjkljivih robov na določenem položaju, kjer so bile prisotne točke za fasadno ravnino, je mejni okvir točk za fasadno ravnino (zeleni mejni okvirji na sliki 1) nadomestil mejne okvirje povezanih komponent robov iz prejšnjega koraka.

Slika 1: Izločene povezane komponente, zaobjete z mejnimi okviri (rdeče) in regije točk za fasadno ravnino (zeleno)

Raster točk za fasadno ravnino predstavlja večinoma le vdolbine in okna, vendar lahko zaradi omejitev pri izločitvi fasade pride do lažnih interpretacij podatkov na fasadi in posledično do izločitve napačnega kandidata za okno. Primer lažnega podatka na fasadi so mali zeleni mejni okvirji na sliki 1. Zaradi rastriranja z velikostjo piksla 0,5 m, lahko pride do spojitve sosednjih oken, kot je lahko vidno na levi strani slike 1 (zeleni mejni okviri). Dodatne omejitve predstavljajo tujki na fasadi, kot npr. žica, ki je povezana s četrtim oknom v zgornjem nadstropju. Le ta namreč predstavlja zadostno spremembo vrednosti intenzitete in je izločena kot rob, ki je povezan z robovi okenskega okvirja. Posledično izločitev povezanih komponent robov na območjih, kjer je prisotnost točk za fasadno ravnino, povzroči prevelike okvirje oken. Za eliminacijo omenjenih lažnih alarmov in prevelikih mejnih okvirjev, sem dobljene rezultate pomnožila z binarno masko segmentov, katero sem pridobila z rastjo regij na fasadni teksturi. Le ta je namreč neodvisna od točk za fasadno ravnino. Poleg tega sem v procesu segmentacije za glajenje uporabljen dovolj velik filter mediane, da se rast regij ni ustavila zaradi tankega elementa, kot je npr. žica in je bil ta del združen skupaj s preostalim zidom. To pomeni, da je bila vrednost maske na tem območju 0 in je eliminirala večje mejne okvirje.

Vsi do sedaj opisani koraki so temeljili zgolj na domnevah, ki podajajo prisotnost oken v uporabljenih podatkih. Kljub veliki količini informacij, ki se jih lahko pridobi iz podatkov ALS poševnega pogleda o položaju oken in iz fasadnih tekstur o okvirjih, še vedno ostaja problem na določenih mestih, kjer ni podatkov. Zaradi relativno nizke gostote podatkov ALS poševnega pogleda po prej omenjenem postopku ne moremo zaznati nekaterih oken.



Shema 3: Izboljšava postopka zaznave s nivojsko NNK

Zato sem uporabila predpostavko, da je v istem nadstropju več oken enakega tipa. Glede na to, da so bile zaznane površine oken že zanesljivo določene, se je le te izrabilo kot vzorčne matrike v procesu normalizirane navzkrižne korelacije. Primerjava je potekala med vzorčnimi matrikami in fasadno teksturo, razdeljeno na nivoje, kot je vidno na shemi 3. S tem sem omejila prostor iskanja.

Že prej zaznana območja oken sem pretvorila v binarni raster z vrednostjo 1 na območjih zaznanih okenskih površin. Binarni raster sem nato razdelila v nivoje glede na minimume vrstične vsot. Okna v enem nadstropju niso vedno samo ene vrste in tudi če so ene vrste, je možno, da vrednost normalizirane navzkrižne korelacije med njimi ni dovolj visoka, zaradi rolet, spuščenih na različne nivoje, ozadij oken, različne osvetljenosti, zato je bil izveden iterativni postopek, ki je našel največ tri različne vzorčne matrike izmed zaznanih oken v vsakem nivoju.

Izbira vzorčne matrike v posameznem nivoju zavisi od:

- velikosti zaznane površine okna,
- velikosti območja brez podatkov na *rastru fasadnih točk* znotraj te površine.

Z namenom preprečitve izbora več podobnih vzorčnih matrik sem z izbrano vzorčno matriko izvedla navzkrižno korelacijo z drugimi že prej zaznanimi površinami oken v istem nivoju. Vsa zaznana okna, katera so bila korelirana za več kot podano mejno vrednost, so bila v nadaljevanju izključena iz procesa. Neizločene zaznane okenske površine v posameznem nivoju so nastopale v drugi iteraciji kot potencialne vzorčne matrike. Za analizo vsakega posameznega nivoja so bile izbrane največ tri vzorčne matrike.

Z izbranimi vzorčnimi matrikami sem v naslednjem koraku izvedla normalizirano navzkrižno korelacijo vzdolž korespondenčnega nivoja z namenom zaznave dodatnih oken, za katere ni bilo dokaza v podatkih ALS poševnega pogleda.

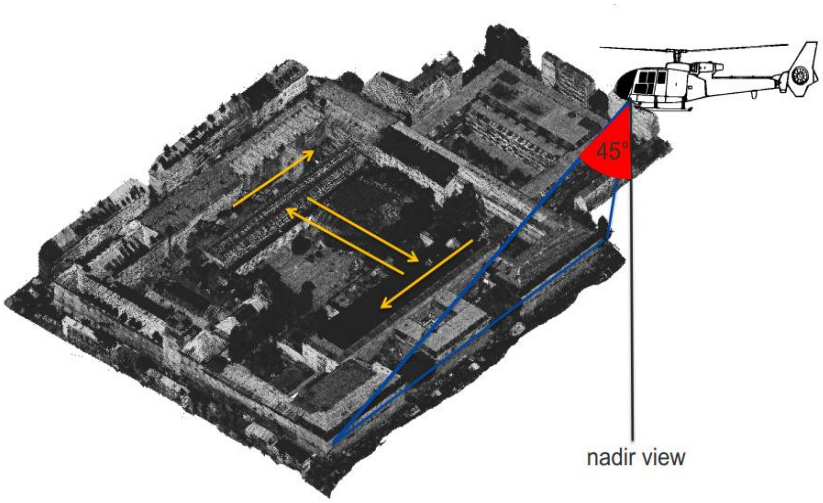
6.3 Uporaba metodologije

6.3.1 Opis podatkov

Uporabljeni so bili tri različne vrste vhodnih prostorskih podatkov: 3D žični model stavb, ortorektificirane fasadne teksture iz vidne podobe in aerolaserski oblak točk poševnega pogleda. Vse postopke razen rasti regij sem sprogramirala v Matlabu, rast regij pa sem izvedla s programom za obdelavo slik Halcon.

Aerolaserski oblak točk poševnega pogleda z več smeri

Območje zajema je bil kampus Tehnične univerze v Münchnu (TUM). Podatki so bili zajeti z aerolaserskim skeniranjem poševnega pogleda v oktobru 2006 in pripadajo Fraunhofer IOSB. ALS poševnega pogleda je ALS z laserjem, usmerjenim pod kotom 45° glede na nadir. Za popoln zajem celotnega območja, je bilo območje preleteno v štirih glavnih smereh, kar je razvidno s slike 2. Gostota oblaka točk je 9,7 točk na m^2. Marcus Hebel je v sklopu svoje doktorske naloge georeferenciral podatke, koregistriral proge različnih preletov in segmentiral oblak točk (Hebel, 2012). Referenčni sistem podatkov je UTM.



Slika 2: ALS poševnega pogleda v več smereh

Ortorektificirane fasadne teksture

Vidne podobe, ki predstavljajo osnovo za teksture, so bile zajete s kamero Nikon D3 z goriščno razdaljo 24,4 mm. Sebastian Tuttas je naredil blokovno izravnavo in rektifikacijo. Za eksperiment je bilo uporabljenih 11 vidnih ortorektificiranih tekstur, ki predstavljajo pročelja južnega in vzhodnega dela kampusa TUM.

3D model stavb

3D model stavb je bil ročno zajet iz zračnih posnetkov z uporabo 3D StereoAnalyst v ArcGISu. Koordinate oglišč poligonov so v Gaus-Krügerju in so urejene za vsak poligon posebej.

6.3.2 Preizkus razvite metodologije s podatki

Metodologijo smo uporabili na prej opisanih podatkih. Najprej smo opravili koregistracijo oblaka točk in 3D žičnega modela. Detekcija oken je potekala na 11 fasadah za katere smo naredili vse korake predobdelave in zaznavanja.

6.3.3 Rezultati

Kot je razvidno s slike 3 uporabljena metoda omogoča dobro zaznavo oken; tudi okvirji oken so dobro določeni. Na sliki 3 je jasno viden barvni prehod med okvirji oken in zidom, kar je omogočilo izločitev robov, ki pripadajo okvirjem oken. Dodaten razlog za dobro definirane okvirje oken in majhno število lažnih alarmov na zidu je posledica preseka z binarno masko segmentov zidov pridobljenih z rastjo regij. Okna na sliki 3 so namreč med sabo ločena s homogenimi deli zidu, kateri so izločeni z rastjo regij in tvorijo binarno masko. Ker so točke za fasadno ravnino osnova za zaznavo, so poleg oken zaznane tudi druge vdolbine na fasadi, kot so arkade, vrata in niše (glej sliko 3).

Slika 3: Zaznana okna in druge vdolbine

Okna so bila zaznana tudi navkljub prisotnosti avtomobilov na fasadni teksturi, kateri so zastirali površine oken. Razlog temu so najdene točke za fasadno ravnino na tem območju fasade. Okenski okvirji za avtomobili na sliki 4, so dokaj dobro definirani. Poudariti je potrebno, da v tem primeru povezane komponente robov, katere tvorijo osnovo za obliko okvirjev, pripadajo avtom in ne oknom.

Uporaba nivojske navzkrižne korelacije je omogočila zaznanavo dodatnih oken, za katere ni bilo neposrednih dokazov v podatkih. Poleg tega je ta metoda prispevala tudi k boljši določitvi površine oken, katera so bile predhodno zaznane z večimi mejnimi okvirji.

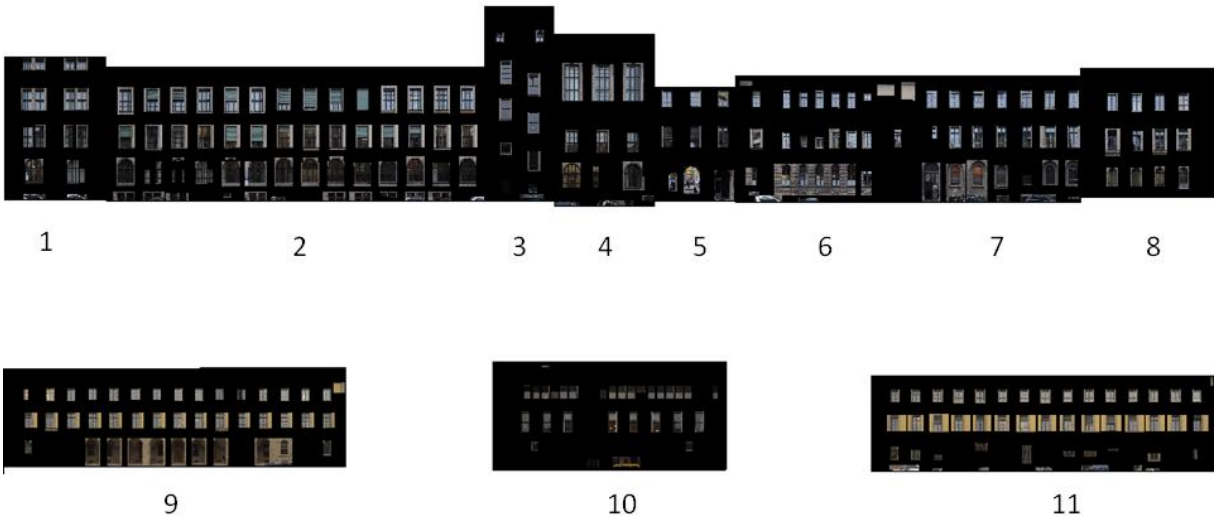
Slika 4: Zaznava oken navkljub avtom pred stavbo

Ker je bil namen te diplomske naloge obogatiti obstoječi 3D model z okni, sem zaznane površine oken v obliki teksture prenesla na 3D žični model (glej str. 46).

6.3.4 Ovrednotenje metodologije

Metodologijo sem ovrednotila po postopku, kateri je prikazan na sliki 5. Vsako okno, ki je bilo zaznano z enim mejnim okvirjem, sem smatrala kot pravilno zaznavo (angl. true positive). Če je bilo na enem oknu več zaznav, je bila ena izmed zaznav privzeta kot pravilna, ostale pa kot nepravilne zaznave (angl. false positives). V primeru, ko sta bili dve ali več oken omejeni z enim mejnim okvirjem, je bil ta primer obravnavan kot nepravilna zaznava (angl. false positive). Vsa nezaznana okna sem obravnavala kot nepravilna (angl. false negatives).

Slika 5: Način ovrednotenja metodologije



Slika 6: Teksturirane zaznane površine oken na 3D modelu stavb

Tabela 1: Ovrednotenje metodologije

Number	1	2	3	4	5	6	7	8	9	10	11	Σ
GT	18	70	10	15	9	32	27	16	66	61	66	390
D	16	62	11	13	9	24	24	12	40	30	48	289
TP	8	47	10	9	9	17	21	12	37	30	44	244
FP	8	15	1	4	0	7	3	0	3	0	4	45
FN	11	23	0	6	0	15	7	4	29	31	22	148
DR	0.44	0.67	1.00	0.60	1.00	0.53	0.78	0.75	0.56	0.49	0.67	0.63
FR	0.50	0.24	0.09	0.31	0.00	0.29	0.13	0.00	0.08	0.00	0.08	0.16

GT= referenčna vrednost, D= zaznave, TP= pravilna zaznava, FP= nepravilna zaznava, FN=lažna zaznava, DR=odstotek zaznave, FR=odstotek detekcije lažnih vrednosti, Number= oznaka fasade

Zaznava je bila ovrednotena na 63% in količina lažnih zaznav na 16%. V tabeli 1 je prikazano ovrednotenje vsake fasade posebej. Uporabljena metoda ovrednotenja je precej stroga, kar je razvidno iz vrednotenja rezultatov na fasadni teksturi št.1 (glej sliko 6 in tabelo 1), kjer je po ovrednotenju 44% pravilnih zaznav in 50% lažnih zaznav. Po drugi strani pa na sliki 6 vidimo, da so bila navkljub relativno nizki vrednosti pravilne zaznave, vsa okna precej pravilno izločena.

6.4 Zaključek in nadaljnje raziskave

Glavni cilj te diplomske naloge, avtomatska zaznava oken iz kombiniranih tipov podatkov, je bil dosežen. Kljub različni razporejenosti oken, njihovim različnim velikostim in visokim kriterijem ovrednotenja metodologije, je uporabljeni postopek zaznave oken izkazal relativno dobre rezultate (63% pravilnih zaznav in 16% lažnih zaznav). Vrednost pravilne zaznave je bila višja, kot v rezultatih dobljenih po metodi, kjer so bili uporabljeni zgolj podatki ALS (Tuttas & Stilla, 2013). Po uporabljeni metodologiji so bila okna zaznana tudi na tistih območjih fasade, katera so bila zakrita z avtomobili; po drugi strani pa lahko te zaznave vodijo do nerealnih velikosti okenskih površin.

Okna na obravnavanih fasadah so bila zaznana bodisi zaradi privzetih predpostavk v podatkih ali zaradi visoke vrednosti normalizirane navzkrižne korelacije z vzorčno matriko. V postopek nisem vključila verjetnosti razporeditve oken po fasadi, saj bi lahko le to privedlo do napačnega intervala razporeditve in posledično napačne zaznave oken; poleg tega pa bi pri tem šlo zgolj za statistično projekcijo, ne pa za dejansko stanje iz podatkov. Nasprotno pa urbana območja sestojijo večinoma iz simetričnih in ponavljajočih se struktur in bi bila vključitev verjetnostne porazdeliteve le kljub temu smiselna. Obravnavana metodologija zaznave oken, bi lahko predstavljala osnovo za slovnico v proceduralnem modeliranju.

Uporabljeni algoritem se je izkazal slabo na fasadah z opekami, saj so se robovi na območjih razmika med opekami, držali robov okenskih okvirjev. To je privedlo do velikega območja zaznave s petimi okni. To bi bilo možno odpraviti s spreminjanjem ločljivosti in eliminacijo visokih frekvenc podobe, kar bi odpravilo robove zaznane med presledki opek.

Bolj geometrično definirane okvirje oken bi se dalo pridobiti s Houghovo transformacijo (glej sliko 7). V tem primeru je bil na zaznanih površinah oken uporabljen detektor Canny in nato Houghova transformacija, katere zbirna matrika (angl. accumulator array) je bila omejena na skoraj vertikalne in horizontalne linije. Kot je vidno na sliki 7 bi bilo potrebno dodati določene povezljivostne kriterije zahtevane za spojitev linij. Poleg tega se je isti postopek, izveden s Sobelovim filtrom, izkazal z dokaj podobnimi linijami, kot v primeru Cannyjevega detektorja.

Slika 7: Houghova transformacija na zaznanih območjih oken

Nadaljnje raziskave bi lahko klasificirale okna na arkadna in pravokotna, kot je opisano v Lee & Nevata (2006). S Houghovo transformacijo bi se lahko poiskale tudi druge parametrične oblike.

Za izvedbo zaznave oken je potrebna natančna koregistracija različnih vrst podatkov. To bi bilo olajšano z istočasnim zajemom različnih tipov podatkov, kar bi posledično olajšalo problem koregistracije.

Za separacijo segmentov zidu sem najprej poskusila gručenje po metodi K-povprečja v CIELAB barvnem prostoru, ki pa se v mojem primeru ni izkazal za robustnega, zaradi bližine barv območja oken in fasade v CIELAB prostoru in težavnosti avtomatske določitve števila gruč.

Kombiniranje aerolaserskega poševnega pogleda in vidnih fasadnih tekstur se je potrdilo kot potencialno zelo primerno za rekonstrukcijo fasad. Z izboljšavami gostote ALS podatkov bi se izboljšala tudi zaznava. Poleg tega bi se za razlikovanje med detajli na fasadi, vdolbinami in ravnino fasade v nadaljnje lahko uporabila tudi polnovalovna oblika zapisa aerolaserskih podatkov.

REFERENCES

Armenakis, C., Gao, Y., Sohn, G. 2010. Co-registration of lidar and photogrammetric data for updating building databases. In: ISPRS Joint Workshop on Core Spatial Databases-Updating, Maintenance and Services-from Theory to Practice, Haifa, Israel, 15-17 March, 2010. ISPRS Archives, 38, Part 4-8-2/W9.

Antolovic, D. 2008. Review of the Hough Transform Method, With an Implementation of the Fast Hough Variant for Line Detection. Technical Report TR663, Indiana University.

Avbelj, J., Iwaszczuk, D., Stilla, U. 2010. Matching of 3D wire-frame building models with image features from infrared video sequences taken by helicopters or UAVs. In: Paparoditis N., Pierrot-Deseilligny M., Mallet C., Tournaire O. (Eds). IAPRS, Saint-Mandé, France, 1-3 September, 2010, 38, Part 3B: p.149-154.

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

Boulaassal, H., Landes, T., Grussenmeyer, P., 2009. Automatic extraction of planar clusters and their contours on building façades recorded by terrestrial laser scanner. International Journal of architectural computing, 7,1: p. 1-20.

Brook, A., Vandewal, M. and Ben-Dor, E. 2012. Fusion of optical and thermal imagery and lidar data for application to 3-d urban environment and structure monitoring. In: Escalante-Ramirez, B. Remote sensing - Advanced techniques and platforms: China, InTech: p. 29-50. http://cdn.intechopen.com/pdfs/37511/InTech-fusion_of_optical_and_thermal_imagery_and_lidar_data_for_application_to_3_d_urban_environment_and_structure_monitoring.pdf (Date of access: 11.03.2013.)

Díez, A., Arozarena, A., Ormeño, S., Aguirre, J., Rodríguez, R., Sáenz, A. 2008. Fusion and optimization of lidar and photogrammetric technologies and methodologies for cartographic production. The internal archives of the photogrammetry, remote sensing and spatial information sciences, Beijing, 37, Part B1.

Deakin, R.E. 2006. A note on the Bursa-Wolf and Molodensky-Badekas transformations. RMIT University. http://user.gs.rmit.edu.au/rod/files/publications/Similarity%20Transforms.pdf (Date of access: 07.03.2013.)

Department of the Army: U.S. Army Corps of Engineers Washington, DC. 2002. Engineering and design - Photogrammetric mapping (EM 1110-1-1000). http://publications.usace.army.mil/publications/eng-manuals/EM_1110-1-1000_sec/EM_1110-1-1000_Sections/c-11.pdf (Date of access: 25.02.2013.)

Directive 2010/31/Eu of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings. 2010. Official journal of the European Union. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:153:0013:0035:EN:PDF (Date of access: 05.03.2013.)

Gonzales, R.C., Woods, R.E. 2008. Digital image processing. Third edition. Pearson education, New Jersey.

Habib, A., Ghanma, M., Mitishita, E. 2004. Co-registration of photogrammetric and lidar data: methodology and case study. Revista Brasileira de Cartografia, 56, 1.

Habib, A.F., Cheng, R.W.T. , Kim, E.M., Mitishita, E. A., Frayne, R., Ronsky, J. L. 2006. Automatic surface matching for the registration of LIDAR Data and MR imagery. ETRI Journal, 28, 2.

Habib, A.F. 2012. LiDAR and photogrammetric data integration. Lecture notes at Advanced Photogrammetric and Ranging Techniques, University of Calgary. http://dprg.geomatics.ucalgary.ca/system/files/AKAM_531_PPT_CH5.pdf (Date of access: 05.03.2013.)

Halcon (Version 10.0.1). 2011. Reference Manual. MVTEC Software GmbH.

Hebel, M., Arens, M., Stilla, U. 2009. Utilization of 3D city models and airborne laser scanning for terrain-based navigation of helicopters and UAVs. In: Stilla U, Rottensteiner F, Paparoditis N (Eds) CMRT09. IAPRS, Paris, France, 3-4 September, 2009, 38, Part 3/W4.

Hebel M., Stilla U. 2009. Automatische Koregistrierung von ALS-Daten aus mehreren Schrägansichten städtischer Quartiere. PFG Photogrammetrie, Fernerkundung, Geoinformation, 2009, 3: p. 261-275.

Hebel, M., Stilla, U. 2010. ALS-aided navigation of helicopters or UAVs over urban terrain. In: Proceedings of: International Calibration and Orientation Workshop EuroCOW 2010: p. 187-192.

Hebel, M. 2012. Änderungsdetektion in urbanen Gebieten durch objektbasierte Analyse und schritthaltenden Vergleich von Multi-Aspekt ALS-Daten. Phd dissertation. Technical University of Munich.

Ilife, J., Lott, R. 2008. Datums and map projections for remote sensing, GIS and surveying, 2nd edition. Whittles Publishing, UK: p. 120-122

Knippers, R. 2009. Coordinate transformations. http://kartoweb.itc.nl/geometrics/Coordinate%20transformations/coordtrans.html (Date of access: 11.03.2013.)

Kulkarni, V., Nagesh, R., Wu, H. 2011. Window detection in frontal façades. Project work at CS294-69 Image Manipulation and Computational Photography. University of Barkley.

Lee, S.C., Nevatia, R. 2004. Extraction and integration of window in a 3D building model from ground view images. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 27 June-2 July 2004, 2, II-113, II-120.

Kramer A. E. 2013. In Russia, Ruins and Property Spared by Meteor, Side by Side. NY Times (17 February 2013). http://www.nytimes.com/2013/02/18/world/europe/in-russia-property-ruined-and-spared-by-meteor-share-space.html?src=recg (Date of access: 22.02.2013.)

OpenGIS® City Geography Markup Language (CityGML) Encoding Standard. 2012. https://portal.opengeospatial.org/files/?artifact_id=47842 (Date of access: 05.03.2013.)

Pu, S., Vosselman, G. 2007. Extracting windows from terrestrial laser scanning. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Espoo, Finland, September 12-14, 36, part 3/W52: p. 320-325.

Recky, M., Leberl, F. 2010. Windows Detection Using K-means in CIE-Lab Color Space. 20th International Conference on Pattern Recognition (ICPR) 23-26 August 2010: p. 356-359.

Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P. 2007. Hough-Transform and Extended Ransac Algorithms for Automatic Detection of 3D Building Roof Planes from Lidar Data. IAPRS 36, Part 3 / W52.

Tuttas, S., Stilla U., 2011. Window detection in sparse point clouds using indoor points. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Munich, Germany, 38, Part 3/W22: p. 131-136.

Tuttas, S., Stilla, U. 2012. Reconstruction of Rectangular Windows in Multi-Looking Oblique View ALS Data. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXII ISPRS Congress, Melbourne, Australia, 25 August – 01 September 2012, I-3.

Tuttas, S., Stilla, U. 2013. Rekonstruktion von Fenstern aus Schrägsicht-ALS Puntkwolken zur Anreicherung von Gebäudemodellen. In: 33. Wissenschaftlich-Technische Jahrestagung der DGPF, 27. Februar – 1. März 2013, Freiburg i. B. DGPF Tagungsband, 22, 2013: p. 363-372.
http://www.pf.bv.tum.de/pub/2013/tuttas_stilla_dgpf13_pap.pdf

Vosselman, G. 1999. Building reconstruction using planar faces in very high density height data. International Archives of Photogrammetry and Remote Sensing, 32, part 3-2W5: p. 87-92.

Vosselman, G. in Mass, H.G. 2010. Airborne and Terrestrial Laser Scanning. Whittles Publishing, Scotland, UK.

Wang, R., Ferrie, F.P., Macfariane, J. 2012. A Method for Detecting Windows from Mobile Lidar Data. Photogrammetric Engineering & Remote Sensing, 78, 11: p. 1129-1140.

Wasmeier, P. 2011. Geodetic Transformations Toolbox. Matlab file exchange: http://www.mathworks.com/matlabcentral/fileexchange/9696-geodetic-transformations-toolbox (Date of access: 10.08.2012.)

World Health Organization. Urban population growth. http://www.who.int/ (Date of access: 25.02.2013.)

Yang, M., Förstner, W. 2011. Regionwise Classification of Building Façade Images. Photogrammetric Image Analysis (PIA2011). Springer: p. 209 - 220.

Yang, S.W., Wang, C.C. 2011. On Solving Mirror Reflection in LIDAR Sensing. IEEE/ASME Transactions On Mechatronics, 16, 2.

.

Other references used

Boulaassal, H., Chevrier, C., Landes, T. 2010. From Laser Data to Parametric Models: Towards an Automatic Method for Building Façade Modelling. In: M. Ioannides (Ed.): EuroMed, LNCS 6436, pp. 42–55, 2010. Springer-Verlag Berlin, Heidelberg.

Mastin, A., Kepner, J., Fisher, J. 2009. Automatic Registration of LIDAR and Optical Images of Urban Scenes. In: . IEEE Conference Computer Vision and Pattern Recognition, 2009.

Böhm, J., Haala, N. 2005. Efficient Integration Of Aerial And Terrestrial Laser Data For Virtual City Modeling Using Lasermaps. ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005". Enschede, the Netherlands.

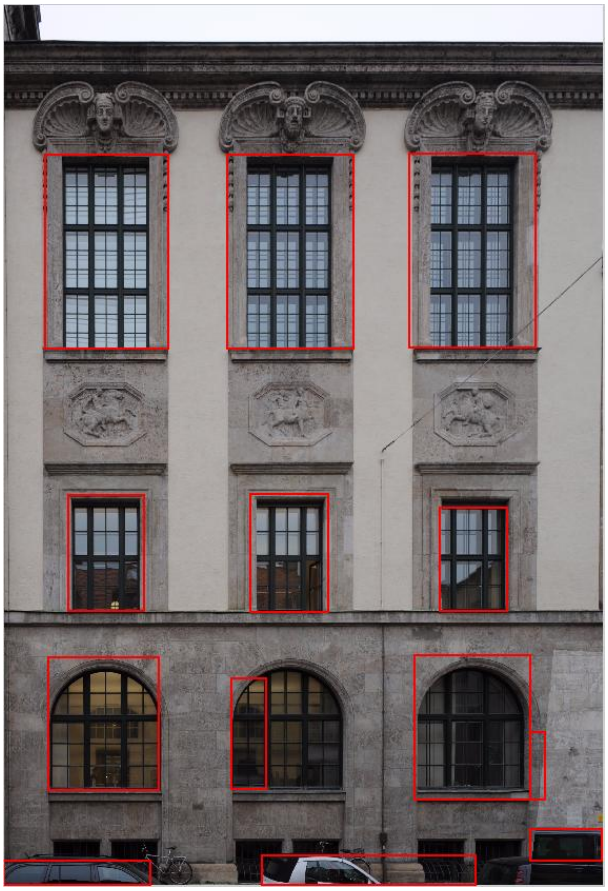
Duda, R. O., Hart, P. E.. 1972.Use of the Hough Transformation to Detect Lines and Curves in Pictures. Communications of Association for Computing Machinery. 15, 1:11-15. http://www.ai.sri.com/pubs/files/tn036-duda71.pdf (Date of access: 10.12.2012.)

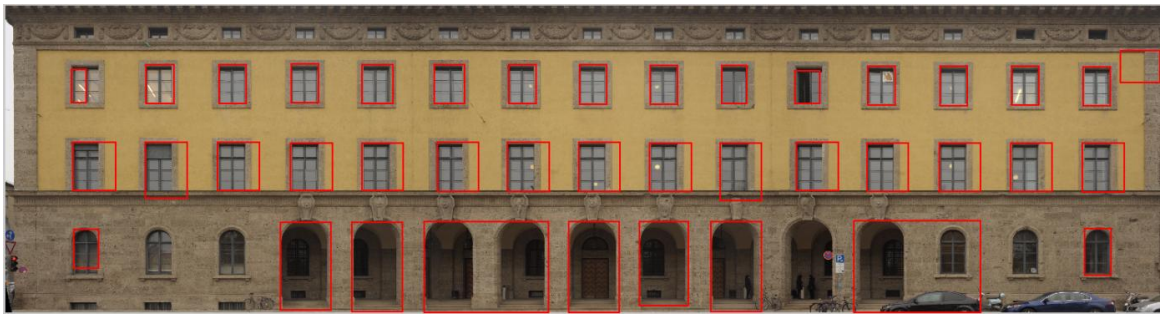
Hansen, W., Gross, H., Thoennessen, U. 2008. Line-Based Registration of Terrestrial and Airborne Lidar Data. ISPRS Archives, 37, Part B3a: p. 161-166.

Koler Povh, T., Turk, G. 2011. Navodila za oblikovanje visokošolskih del na Fakulteti za gradbeništvo in geodezijo in navajanje virov. Ljubljana, Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo.

Wu, C., Frahm, J.M., Pollefeys, M. 2010. Detecting Large Repetitive Structures with Salient Boundaries. In: ECCV'10 Proceedings of the 11th European conference on Computer vision: Part II. Springer-Verlag, Berlin, Heidelberg, Part II, p. 142–155.

APPENDIX A: RESULTS OF WINDOW DETECTION ON 11 FAÇADE TEXTURES (scaled differently)






TECHNISCHE UNIVERSITÄT