



DGK Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C

Dissertationen

Heft 784

Lukas Schack

**Object-based Matching of Persistent Scatterers
to Optical Oblique Imagery**

München 2016

Verlag der Bayerischen Akademie der Wissenschaften

ISSN 0065-5325

ISBN 978-3-7696-5196-6

Diese Arbeit ist gleichzeitig veröffentlicht in:

Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover
ISSN 0174-1454, Nr. 327, Hannover 2016



DGK Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C

Dissertationen

Heft Nr. 784

Object-based Matching of Persistent Scatterers to Optical Oblique Imagery

Von der Fakultät für Bauingenieurwesen und Geodäsie
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades
Doktor-Ingenieur (Dr.-Ing.)
genehmigte Dissertation

von

M.Sc. Lukas Schack

München 2016

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München

ISSN 0065-5325

ISBN 978-3-7696-5196-6

Diese Arbeit ist gleichzeitig veröffentlicht in:
Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover
ISSN 0174-1454, Nr. 327, Hannover 2016

Adresse der DGK:



Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften (DGK)

Alfons-Goppel-Straße 11 • D – 80 539 München
Telefon +49 – 89 – 23 031 1113 • Telefax +49 – 89 – 23 031 -1283 /- 1100
e-mail post@dgk.badw.de • <http://www.dgk.badw.de>

Prüfungskommission

Vorsitzender: Prof. Dr.-Ing. Steffen Schön
Referent: Prof. Dr.-Ing. Christian Heipke
Korreferenten: Prof. Dr.-Ing. Uwe Sörgel
Prof. Dr.-Ing. Monika Sester

Tag der mündlichen Prüfung: 23.09.2016

© 2016 Bayerische Akademie der Wissenschaften, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen

ISSN 0065-5325

ISBN 978-3-7696-5196-6

Abstract

Modern subsidence monitoring for urban areas comprises active spaceborne techniques like synthetic aperture radar (SAR). Specifically, the persistent scatterer interferometry (PSI) has reached market maturity in recent years. This weather independent remote sensing approach allows for large-area observations over long time series and is capable of detecting movements at the scale of some millimeters per year. Recent missions like TerraSAR-X, TanDEM-X, and CosmoSkyMed offer ground resolutions reaching down to 0.6 m, which refined the monitoring granularity from single blocks of buildings towards individual houses and even particular facades. Nevertheless, more advanced monitoring tasks like capturing torsions or other non-rigid deformations demand even finer assignments of individual radar scatterers to their corresponding geometric structures. The spatial resolution of PSI has the potential for establishing such correspondences, but in order to assign individual scatterers to their counterparts at facades a suitable facade representation has to be available. Furthermore, PSI is an opportunistic technique to the effect that it cannot be guaranteed that a particular facade can be used for monitoring. This is due to the fact that the physical nature of many persistent scatterers (PSs) is still unknown, i.e. the exact scattering mechanism at the geometrical structure in general is obscure. Even though simulations revealed some scattering mechanisms at trihedral faces like window corners, the majority of scatterers can be assigned to facades but not to the exact building structure. Thus, sometimes statements about deformation rates are made within millimeter precision, without exactly knowing which structure on the facade was measured.

Airborne oblique imagery delivers optical acquisitions of the scene in a viewing angle comparable to that of SAR sensors. The typically higher spatial resolution and the good interpretability for human operators make this data a very suitable source of information. Geometrical features like window corners, which are known to induce PSs, can be extracted from optical images with high geometric accuracy.

In this thesis, optical oblique images of building faces are used as their representations of the real facade. A method of matching PSs to individual structures of the buildings

mapped in oblique imagery is derived. Even though the viewing direction is similar, the mapping geometries of SAR and optical oblique imagery are very different. To cope with this difference, a feature of many modern multistory buildings is exploited: the regular alignment of windows at the facade. A plethora of office buildings in urban areas are characterized by steady patterns. Those can be extracted both in radar and optical acquisitions of the facade. This is the initial point of this thesis in which a matching algorithm based on graph theory is derived. Regular patterns are grouped in SAR data as well as the optical image by assuming that the facade consists of repeated regular patches of same appearance. One key idea is to perform the subsequent matching by considering *facade objects* instead of single PSs. On this basis prior knowledge is introduced, like the assumption that all regularly aligned scatterers at the facade lie in the same plane. Nevertheless, every PS is assigned to its *individual* counterpart in the oblique image and carries an accuracy information.

The developed object-based approach also allows one to derive topology information from the very precise range-azimuth domain of the SAR data and incorporate it together with a geometrical measure into the matching. This helps mitigating the inaccuracy induced by the projection into the image compared to a solution based on individual PSs. The derived method bases on the iterative application of the Hungarian algorithm and is proven to always yield the optimal solution as long as well defined constraints are met.

With the aid of a variety of case studies, it is shown that a meaningful assignment of PSs to their correspondences in optical oblique images can be achieved. The presented method is robust against changing contrasts on the facade due to shadows, steep aspect angles, and also works on partly occluded building faces due to the object-based approach. In order to derive a statement about the feasibility and quality of assigning individual PSs to structures at the facade, the area of the repeated regular patch is compared to the area of the 95% confidence ellipse of the projected PS position in the image. The latter is found to be usually smaller than the extent of the repeated element at the facade which means that a meaningful assignment was found.

Zusammenfassung

Satellitengestützte Fernerkundungssensoren sind ein wichtiger Bestandteil moderner Methoden zur Überwachung von Bodenbewegungen. Insbesondere die Persistent Scatterer Interferometrie (PSI), die ein spezielles Auswerteverfahren der bildgebenden Radar-Fernerkundung ist, erreichte in den letzten Jahren breite Anwendung. Diese wetterunabhängige Technik ermöglicht großflächige Beobachtungen über lange Zeiträume bei einer Auflösung der Deformationsmessung von bis zu wenigen Millimetern pro Jahr. Moderne Satellitenmissionen wie TerraSAR-X, TanDEM-X und CosmoSkyMed besitzen Bodenaufösungen von bis zu 60 Zentimetern. Im Vergleich zu früheren Sensoren verfeinerte dies die Möglichkeiten der Zuordnung von Streuern zu einzelnen Fassaden. Dennoch benötigen genauere Überwachungsmethoden noch feinere Lokalisierungen der Reflexionsmechanismen. Die räumliche Auflösung der PSI-Methode bietet das Potential dazu. Um jedoch eine Zuordnung einzelner Streuer zu ihren Entsprechungen an Fassaden herzustellen, benötigt man ein gutes Modell ebendieser. Die PSI ist ein opportunistisches Verfahren. Dies bedeutet, dass für eine bestimmte Fassade nicht garantiert werden kann, dass für eine Überwachung günstige Streumechanismen vorhanden sind. Der Grund dafür ist, dass die genauen physikalischen Ursachen für die Entstehung von Persistent Scatterern (PS) in vielen Fällen noch unbekannt sind. Zwar wurden mittels Simulationen einige Streumechanismen, beispielsweise an Fensterecken, erklärt. Für die Mehrzahl an Streuern kann aber nach wie vor nur eine eindeutige Zuordnung zu Fassaden, nicht jedoch zu den genauen geometrischen Strukturen an ebendiesen, gemacht werden. Dies bedeutet, dass oftmals millimetergenaue Aussagen über Deformationen getroffen werden, ohne genau zu wissen, welche Struktur an der Fassade gemessen wurde.

Schrägluftbilder liefern optische Abbildungen der Szene unter einem ähnlichen Blickwinkel wie Synthetische-Apertur-Radar (SAR)-Sensoren. Außerdem weisen sie üblicherweise eine im Vergleich zu SAR höhere räumliche Auflösung sowie eine für den menschlichen Betrachter einfachere Interpretierbarkeit auf. Somit stellen sie eine geeignete Quelle an Informationen über die Fassade dar. Geometrische Merkmale wie Fensterecken, von denen bekannt ist, dass sie Persistent Scatterer verursachen, können mit hoher Genauigkeit extrahiert werden.

In dieser Arbeit werden optische Schrägluftbilder der Fassaden als deren geeignete Abbildung genutzt und eine Methode entwickelt, die Radarstreuer einzelnen Fassadenstrukturen im Bild zuzuordnen. Fensterecken, die aus den optischen Bildern extrahiert wurden, dienen als Partner für die Zuordnung. Auch wenn die Blickrichtungen beider Sensoren ähnlich sind, unterscheiden sich die Abbildungsgeometrien sowie die verwendeten Wellenlängen erheblich. Um dennoch eine Zuordnung etablieren zu können, wird eine Eigenschaft vieler mehrstöckiger Gebäude ausgenutzt: die regelmäßige Anordnung von Fenstern an der Fassade. Diese gitterartigen Muster können sowohl in optischen als auch in SAR-Aufnahmen des entsprechenden Gebäudes erkannt und beschrieben werden. Dies ist der Zusammenhang, bei dem die Methode dieser Arbeit ansetzt. Es wird ein auf Graphmatching basiertes Verfahren entwickelt, welches solch regelmäßige Muster extrahiert und eine 1-zu-1-Zuordnung zwischen jedem PS und seiner Entsprechung im optischen Bild herstellt. Regelmäßige Muster werden sowohl in den SAR-Daten als auch im optischen Bild extrahiert unter der Annahme, dass die Fassade aus sich wiederholenden gleichen Teilen besteht. Bei der anschließenden Zuordnung ist ein zentraler Gedanke, ein *Fassadenobjekt* anstatt viele einzelne PS zu betrachten. Auf dieser Grundlage lässt sich Vorwissen, wie beispielsweise die Annahme, dass alle PS an der Fassade in einer Ebene liegen, einführen. Das Approximieren der Fassade durch eine gemeinsame Ebene wird genutzt, um den Fehlerhaushalt der einzelnen PS zu reduzieren.

Die objekt-basierte Betrachtung des Problems erlaubt die explizite Formulierung der Nachbarschaftsbeziehungen der PS untereinander. Diese wird gemeinsam mit einem geometrischen Maß in das Zuordnungsproblem eingeführt, um eine höhere Robustheit gegenüber einer rein geometrischen Lösung zu erreichen. Die entwickelte Methode basiert auf der iterativen Anwendung des Ungarischen Algorithmus. Es wird bewiesen, dass sich unter den definierten Randbedingungen eine optimale und garantiert konvergierende Lösung ergibt.

Anhand vielfältiger Experimente wird gezeigt, dass eine richtige Zuordnung zwischen PS und ihren Entsprechungen im optischen Schrägluftbild hergestellt werden kann. Das Verfahren ist robust gegenüber sich lokal änderndem Kontrast auf der Fassade sowie steilen Blickwinkeln und kann aufgrund der objekt-basierten Betrachtung erfolgreich teil-verdeckte Fassaden verarbeiten. Um eine Aussage über die Qualität der Zuordnung zu machen, wird die Fläche eines sich wiederholenden, gleichen Fassadenteils mit der Fläche der 95%-Konfidenz-Ellipse eines in das Bild projizierten PS verglichen. Diese ist üblicherweise kleiner als das wiederholte Fassadenelement, was bedeutet, dass eine eindeutige Zuordnung erreicht wurde.

Contents

Abbreviations and notation	9
1. Introduction	11
1.1. Motivation	11
1.2. Objectives and contributions	14
1.3. Structure of the thesis	16
2. Basics	17
2.1. Persistent scatterers	17
2.1.1. Synthetic aperture radar	17
2.1.2. Persistent scatterer interferometry	22
2.2. Oblique optical imagery	24
2.3. The inaccuracy budget of persistent scatterers projected into oblique images	26
2.3.1. Geocoding of SAR acquisitions	27
2.3.2. Projection of single PS object coordinates into oblique images	31
2.4. Graphs	33
2.5. Bipartite graph matching	35
3. State of the art	39
3.1. Recent developments and applications of PSI	39
3.2. Persistent scatterers at facades	42
3.3. Grouping of persistent scatterers	43
3.4. Fusion of SAR data with other sources	44
3.5. Facade extraction and description in optical imagery	45
3.6. Graph matching	47
3.7. Objectives revisited	48
4. Object-based matching of PSs to optical oblique imagery	51
4.1. Regularity at facades	51
4.2. Fundamental assumptions	54
4.3. Grouping persistent scatterer point cloud data	56
4.3.1. Segmenting the persistent scatterer point cloud	57

4.3.2.	Projection of PSs onto the facade plane	60
4.3.3.	Extracting regularity from PS point clouds	64
4.4.	Lattice extraction in optical oblique imagery	67
4.4.1.	Image rectification	69
4.4.2.	Segmenting the facade	71
4.4.3.	Explicitly modeling the window	74
4.5.	Iterative graph-based bipartite matching	78
4.5.1.	Topology information	82
4.5.2.	Optimization scheme	84
4.6.	Improving the accuracy of PS in optical images	88
5.	Experiments	93
5.1.	Data description	94
5.2.	Regular facades in urban area: case study Berlin	95
5.3.	Grouping of persistent scatterers	96
5.3.1.	Grouping of PSs: Park Kolonnaden	97
5.3.2.	Grouping of PSs: Charlottenstraße	100
5.4.	Window corner extraction	102
5.4.1.	Window corner extraction: Park Kolonnaden	103
5.4.2.	Window corner extraction: Charlottenstraße	105
5.4.3.	Window corner extraction: Axel-Springer building	105
5.5.	Iterative graph-based bipartite matching	107
5.5.1.	IGBM: Park Kolonnaden	107
5.5.2.	IGBM: Charlottenstraße	111
5.5.3.	IGBM: Axel-Springer building	111
5.5.4.	IGBM: Ebertstraße	114
5.6.	Balance between geometry and topology	116
5.7.	Application to the whole test area	118
6.	Conclusions and open questions	121
6.1.	Summary and conclusion	121
6.2.	Future work	124
	Bibliography	125
	Appendices	129
A.	Proof of Kónig-Egerváry Theorem	129
B.	Example of the Hungarian algorithm	131

Abbreviations and notation

DEM digital elevation model.

GNSS global navigation satellite system.

GSD ground sampling distance.

IGBM iterative graph-based bipartite matching.

IMU inertial measurement unit.

InSAR interferometric synthetic aperture radar.

LOS line of sight.

NCC normalized cross correlation.

PS persistent scatterer.

PSI persistent scatterer interferometry.

RANSAC random sample consensus.

SAR synthetic aperture radar.

SCR signal-to-clutter ratio.

SNR signal-to-noise ratio.

UTM universal transverse mercator.

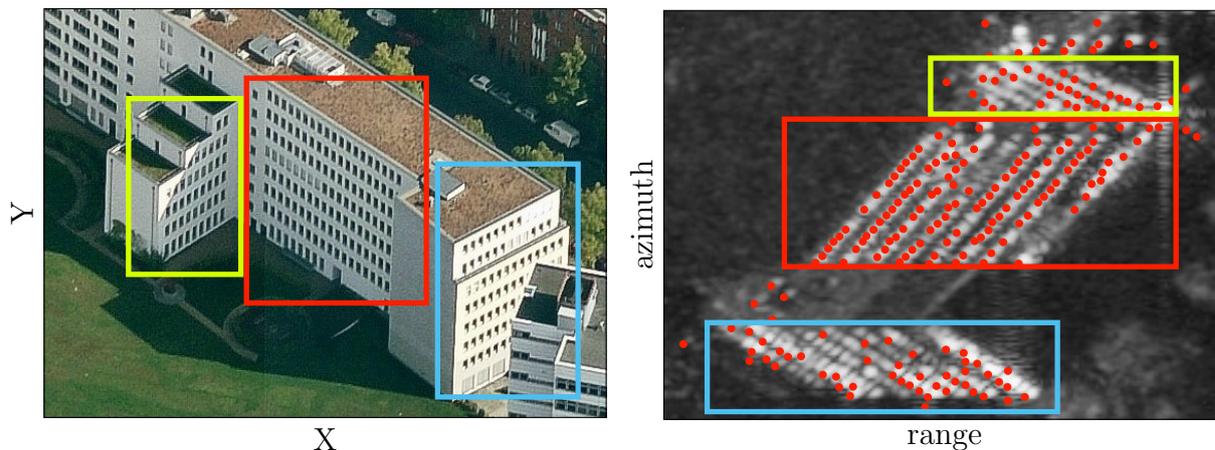
Throughout the thesis, scalars are depicted as lower case or capital Latin or Greek characters, e.g. a , A , or α . Vectors are denoted with an arrow on top of the symbol: \vec{a} , while matrices are written in lowercase or capital bold characters: \mathbf{a} or \mathbf{A} . Sets of any entities are denoted in calligraphic letters, e.g. \mathcal{A} .

1. Introduction

1.1. Motivation

Remote sensing is an indispensable technique in multiple fields of geoscience. The non-contact acquisition of information about the Earth's surface builds the basis for many applications in agriculture, regional planning, mapping, settlement monitoring, landslide detection, and many other applications at the local, regional, and national scale. Especially the weather and sun independent synthetic aperture radar (SAR) systems play an important role for time critical tasks like damage mapping after natural disasters or reliable long-term monitoring of surface deformation areas. In recent years, spaceborne radar sensors have also become widely used for subsidence monitoring as a result of mining. Movements in sensing direction of up to a few millimeters per year can be resolved. In such scenarios, statements about large structures or even building blocks are made. But the geometrical resolutions of modern SAR sensors allow for applications with much more spatial detail since the identification of single small buildings and even parts of facades are now possible. Together with the persistent scatterer interferometry (PSI) processing technique, the assignment of individual scatterers, so called persistent scatterers (PSs), to parts of buildings comes into reach. As a consequence, instead of only detecting rigid movements of complete buildings, more complex deformations like torsions can be monitored. But even though the sensing geometry itself is known, SAR images often require an experienced human interpreter in order to extract information about the nature of the sensed objects. On the one hand, this is due to the used wavelengths which lie in the microwave spectrum, and on the other hand, due to the unknown geometry of the irradiated objects on the surface leading to complex reflection mechanisms. Thus, the physical nature of the object position corresponding to a single pixel in the SAR acquisition, i.e. its equivalent in the object space, is often not known and therefore subject to ongoing research.

Optical remote sensing sensors are also well established. Especially *oblique* imagery has seen a rise in the last decade. For instance, many municipalities acquire optical images of their

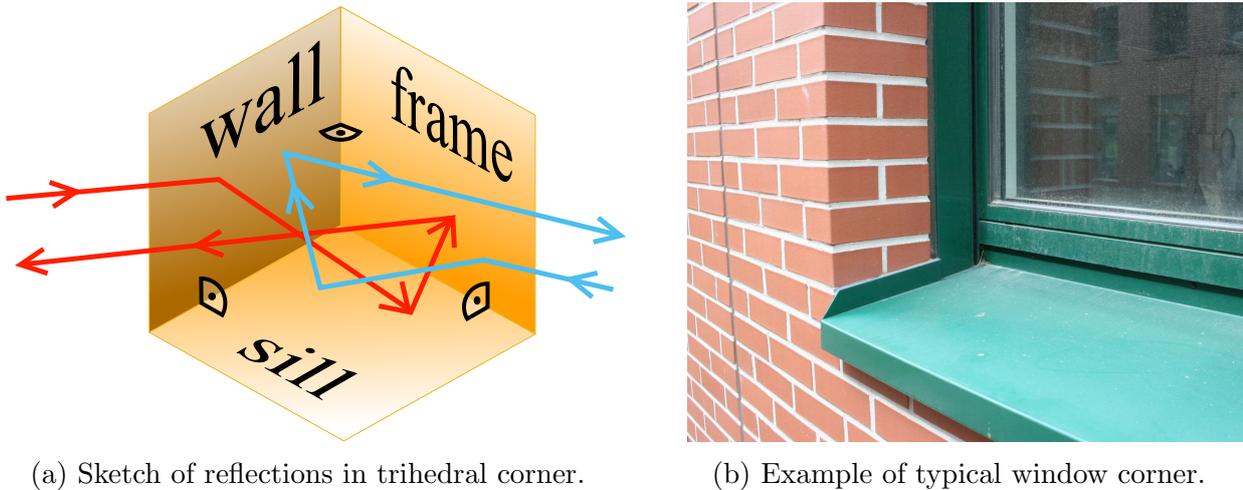


(a) Optical oblique image.

(b) SAR amplitude image with PSs in red.

Figure 1.1: Example for the appearance of facades in optical and SAR data.

inhabited areas in order to use them as a tool for various planning activities, monitor long-term developments or create three-dimensional models of their cities or townships. Compared to SAR products, the interpretation of optical images is easier even for inexperienced operators. Commonly, the ground sampling distance (GSD), i.e. the spatial resolution, of airborne oblique imagery is in the range of 5 cm to 40 cm, while modern SAR sensors reach a resolution on the ground of some 60 cm to 120 cm. Given a SAR acquisition and an optical oblique image of the same facade, the latter typically carries more information about the structure and condition of the building face. This is a central motivation for this thesis: if an assignment of individual PSs to their counterparts in optical oblique imagery is established, it could be used to learn something about the physical nature of the scatterers and may lead the way towards an automatic interpretation of the SAR scene. Finding such correspondences imposes the challenging task of matching PSs to features derived from optical imagery. The sensing geometries and used wavelengths are rather different, which results in two very distinct appearances of the same facade in both sensor types. Figure 1.1 gives an example. In the optical oblique image, three facades of a building complex are shown, marked with colored rectangles. The same facades in the SAR acquisition of Figure 1.1b are consistently marked. In the background, the mean amplitude image is shown as gray values. Bright pixels correspond to high intensities of the sensed signal. The red points mark the result of a PSI processing, yielding temporarily stable scatterers with a high signal-to-noise ratio (SNR). The regular alignment of facade structures is striking in both data sets. Lines of PSs are discernible and comparing them to the facades in the optical image makes the correspondence to the distribution of windows evident. This is



(a) Sketch of reflections in trihedral corner.

(b) Example of typical window corner.

Figure 1.2: (a) Sketch of two threefold reflections in a trihedral corner reflector as it is often formed by window sills, window frames and the wetting building wall. (b) Example window corner which forms a trihedral corner reflector with nearly rectangular planes.

especially true for the facade marked in red. An assignment of PSs to individual windows can be made visually. In fact, the regular alignment of objects at the facades is preserved in both imaging geometries and window corners often induce PSs. The latter results from a scattering mechanism that is already well understood: the threefold reflection in trihedral corners. Three perpendicular planes form structures that reflect the incoming signal back, independently of the direction of the incoming signal, as long as the inlying faces of the trihedral structure are visible for the sensor. Such corner reflectors are often formed by the window sill, the window frame, and the wall of the recessed window. Figure 1.2 shows such a scattering mechanism. A sketch of the reflection model for two rays is depicted in Figure 1.2a. A corresponding real world example of such a window corner is shown in Figure 1.2b.

The aim of this thesis is to develop a method for the automatic identification of the equivalents of PSs in optical oblique images. The principal intricacy to overcome is the inherently different appearance of urban objects in data originating from both sensor types. A key idea to cope with this challenge is to exploit regularities which are discernible in both data domains. The alignment of facade objects like windows or balconies is often very regular and can be described by a lattice in many cases. Graph matching techniques are used in order to match lattice nodes extracted from SAR data to regular point features derived from an optical oblique image of the same facade. This task is formulated as an optimization problem aiming at minimizing the geometrical distance between PSs projected into the image and their matching partners, as well as maintaining the topology of the

grouping. The desired result is a one-to-one assignment of every PS to its equivalent in the optical oblique image. The topology information is used to mitigate possible inaccuracies originating from the SAR domain and the projection into the image. The resulting matching offers the opportunity to systematically investigate the mechanisms leading to the occurrence of PSs at facades which has been found to be dependent on the sensor's line of sight (LOS) and many other sensing parameters. Another application is change detection after natural disasters in a setup where a PS point cloud is available at the time before the event arose. The airborne optical imagery is matched to the PS data and the degree of destruction can be derived.

1.2. Objectives and contributions

In this thesis, a new method for matching PSs to their correspondences in optical oblique imagery is derived. The following three contributions are fundamental for the achievement of this goal. First, the patterns in data of both sensor types are extracted and described. The PS point cloud data is segmented into separate facade clusters followed by a grouping procedure. This information is then transferred into the optical image in order to simplify the pattern recognition. This inter-sensor knowledge transfer allows one to apply rather simple yet very robust methods to extract the patterns in optical images. Once a regularity is found, all PSs and also potential window corners extracted from the optical image are described by a lattice model, respectively. This is the transition from considering individual scatterers towards regarding facade *objects* which allows one to introduce model knowledge.

Relations of data points to each other within one sensor type are then merged into the matching procedure. The goal of this assignment algorithm is to transfer the sensor type internal topology information across the domains and find an optimal solution which minimizes a geometrical, as well as a topological measure between PSs and their counterparts in optical data. The geometric measure is induced by the consideration that the distance between every PS and its counterpart in the image should be as small as possible. On the other hand, the matching partners of neighboring PSs should also be adjacent. This constraint can be formulated as a topological measure which sometimes counteracts the geometrical measure. Furthermore, both terms are mutually dependent on each other and an optimal solution depends on the relative contribution of each term which has to be determined. An algorithm is derived and proven to converge and always yields the optimal solution in terms of the matching, as well as the final position of all PSs in the image.

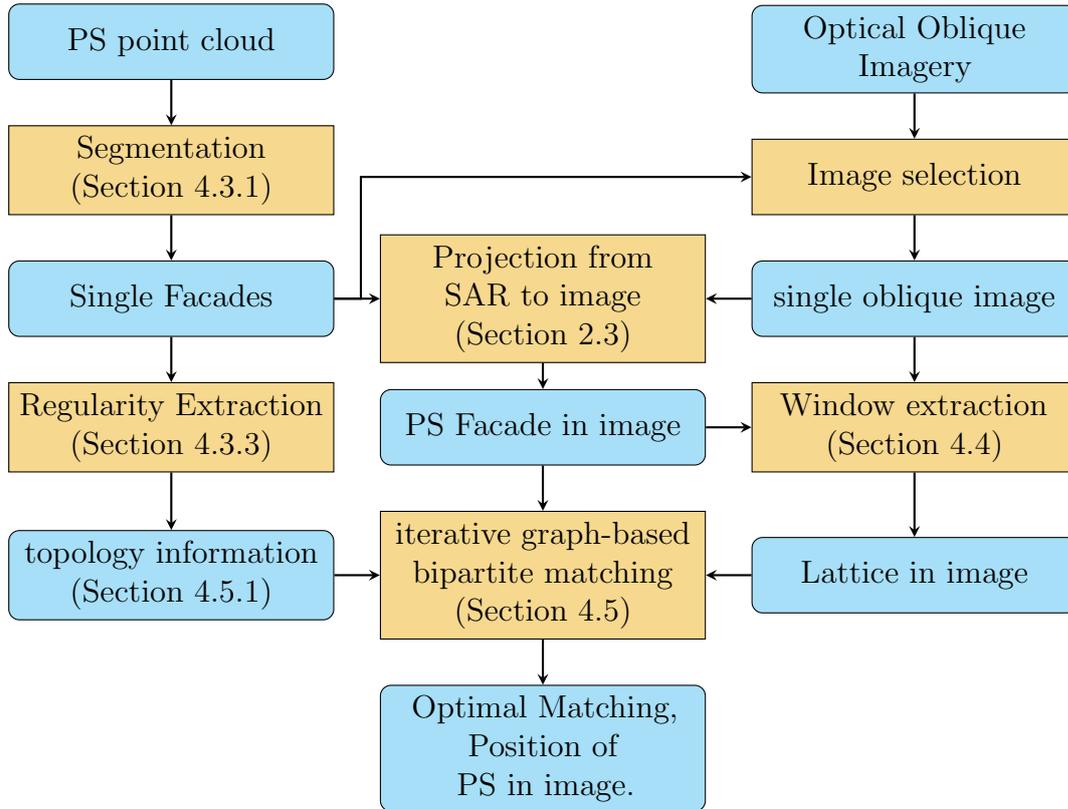


Figure 1.3: Flowchart of the presented method. The sections in brackets contain detailed information about the respective steps.

As the third main contribution of this thesis, investigations are made about how accurate the assignment between PSs and their corresponding optical features can be established. Also, approaches to improve the accuracy of the image coordinates of aggregated PSs compared to simply projecting single scatterers into the image are made. The position uncertainty of the PS position in the optical image is mostly determined by the relatively inaccurate elevation direction. Compared to the other two SAR coordinates range and azimuth the localization accuracy in elevation is worse by a factor of 20. Furthermore, the projection itself carries inaccuracies typically dominated by the exterior orientation parameters of the camera. In order to cope with these influences, assumptions are made like the claim that all regular PSs at the facade lie in a plane. This model knowledge and other considerations then are converted into remedies which reduce the inaccuracy down to the level that individual structures of the facade can unambiguously be identified.

1.3. Structure of the thesis

This thesis is structured as follows. First, in Chapter 2 some basics are presented which form the foundation of the methods derived later. The inaccuracy of the projection of PSs into oblique images and fundamentals of graph matching are of particular interest. Chapter 3 gives insights into related work. This includes recent developments of SAR processing techniques as well as grouping of PS point clouds and approaches to fuse them with data from other sources.

The new method for assigning PSs to their correspondences in optical imagery is then described in Chapter 4. Fundamental assumptions which imply constraints and limitations of the matching approach are explicitly stated in Section 4.2. Figure 1.3 shows the flowchart of the method and also the corresponding sections. The structure of the method chapter basically coincides with the procedure itself. First, the segmentation and grouping of a PS point cloud is described. This result is then projected into the image and used in the extraction of regular window corners in the optical imagery. Both results are then introduced into the derived iterative graph-based bipartite matching (IGBM) algorithm, resulting in an optimal matching and the final positions of the PSs in the image. The influences of each processing step onto the resulting accuracies of PS assignments in the image are highlighted in Section 4.6.

In order to validate the most fundamental propositions derived, experiments are conducted in Chapter 5. Selected case studies show the applicability of the method and investigate their characteristics and limitations. Some violations of made assumptions and their consequences on the result are also presented. A short summary of results and an outlook on open research topics in this field are given in Chapter 6.

2. Basics

This chapter addresses some fundamentals on the creation and characteristics of the used data, namely PS point clouds (Section 2.1) and optical oblique imagery (Section 2.2). Section 2.3 states the inaccuracy budget of the combined coordinate transformation from the SAR domain into the optical image. The basics of graphs and bipartite graph matching are given in Section 2.4 and Section 2.5, respectively.

2.1. Persistent scatterers

Persistent scatterers are the result of the PSI processing technique. Only the main concepts of SAR and PSI are outlined here. A detailed explanation is beyond the scope of this thesis, since the PSI processing is not part of the presented method. Only important aspects, which are crucial for understanding the developed approaches are highlighted. A more comprehensive treatment can be found in Kampes [2006] or Moreira et al. [2013], for instance.

2.1.1. Synthetic aperture radar

A SAR is an active remote sensing system which is mounted on a moving platform, i.e. plane or satellite, and sequentially illuminates the Earth's surface in an oblique angle with electromagnetic radiation. The emitted radiation interacts with the objects on the ground and is partly backscattered to the receiving antenna. For remote sensing purposes, the used wavelength typically lies in between one centimeter up to one meter where a tradeoff between geometrical resolution versus penetration depth has to be taken into account [Moreira et al., 2013]. The emitted signal has to be coherent in order to combine series of measurements into an integrated SAR acquisition. In order to resolve the complex nature of the signal, the received pulses are sampled twice. SAR systems deliver two-dimensional images of the illuminated objects. Figure 2.1 shows the imaging geometry schematically. The sensor moves

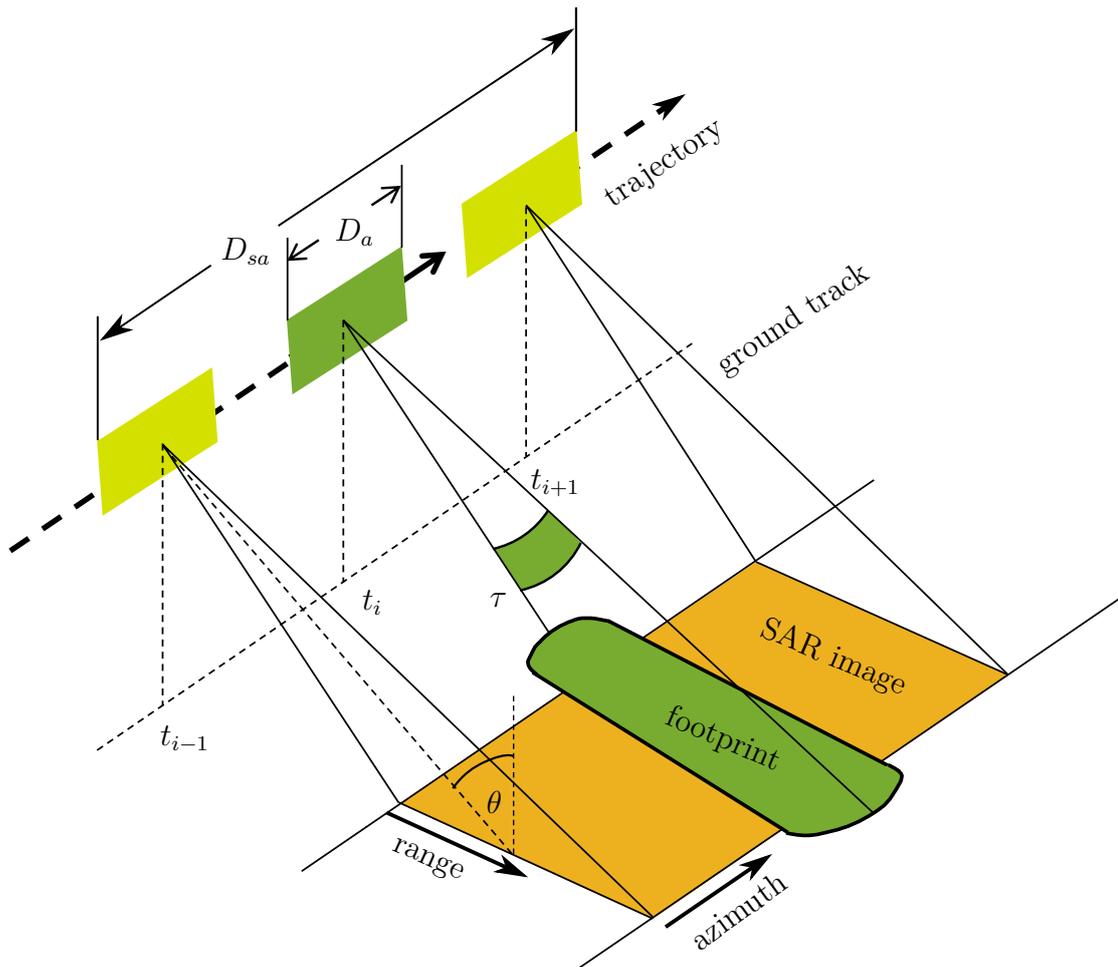


Figure 2.1: Sketch of imaging geometry of a SAR acquisition. The SAR image, depicted as orange area, is formed by coherently integrating the signal from t_{i-1} to t_{i+1} .

along its trajectory, which is denoted as azimuth direction and coincides with the first image coordinate. The emitted pulse of duration τ is directed sideways and nearly orthogonal to the azimuth direction. On the Earth's surface the pulse forms a footprint on the ground which is marked in green in Figure 2.1. The travel time of the signal is directly proportional to the distance between the objects on the ground and the sensor. This direction is denoted as range and coincides with the second image coordinate. The resolution in range direction ρ_r is directly proportional to the length of the pulse τ :

$$\rho_r = \frac{\tau c_0}{2}, \quad (2.1)$$

where c_0 is the speed of the emitted pulse in the atmosphere and the factor $\frac{1}{2}$ results from traversing the distance sensor-object twice. For most applications, the projection of the

range direction and its resolution onto the ground plane is of interest and achieved by incorporating the incidence angle θ as:

$$\rho_{gr} \approx \frac{\tau c_0}{2 \sin \theta}. \quad (2.2)$$

The resolution in azimuth direction is determined by the antenna beam width which is approximately inversely proportional to the real aperture size D_{ra} :

$$\rho_{az,ra} \approx \frac{\lambda r}{D_{ra}} \quad (2.3)$$

where λ is the used wavelength and r is the distance to the sensed target. D_{ra} denotes the length of the real antenna. Incorporating a set of consecutive pulses with overlapping footprints results in a synthetic aperture of length

$$D_{sa} = \frac{\lambda r}{D_{ra}} \quad (2.4)$$

and an improved resolution in azimuth direction of [Bamler & Schättler, 1993]:

$$\rho_{az,sa} \approx \frac{\lambda r}{D_{sa}} = \frac{\lambda r D_{ra}}{2 \lambda r} = \frac{D_{ra}}{2}. \quad (2.5)$$

In the remainder of this thesis only the resolution in azimuth for the synthetic aperture is considered and denoted as ρ_{az} . A detailed derivation of the SAR image formation is out of the scope of this thesis and the reader is therefore referred to Bamler & Schättler [1993]. A notable finding is that the azimuth resolution is range independent and improves with a smaller real aperture. Modern spaceborne sensors like TerraSAR-X offer a further enhancement of the azimuth resolution by electronically steering the beam towards the target scene, resulting in an even longer synthetic aperture [Eineder et al., 2009].

The side-looking imaging geometry of SAR acquisitions can be approximated by a cylindrical projection with the Z -coordinate corresponding to the trajectory of the satellite, i.e. azimuth direction [Soergel, 2010]. The orthogonal distance from the cylinder's axis is given by the range between the sensor and the object. The angle defining the point of the reflection on the circle of equal distances corresponds to the elevation direction. As a consequence of mapping the three-dimensional object into a two-dimensional image, some imaging effects have to be taken into account. For the task of matching PSs to optical imagery, the mixture of signal contributions, denoted as layover, has the largest impact. It is the direct result of not resolving the elevation angle. Figure 2.2 shows this effect

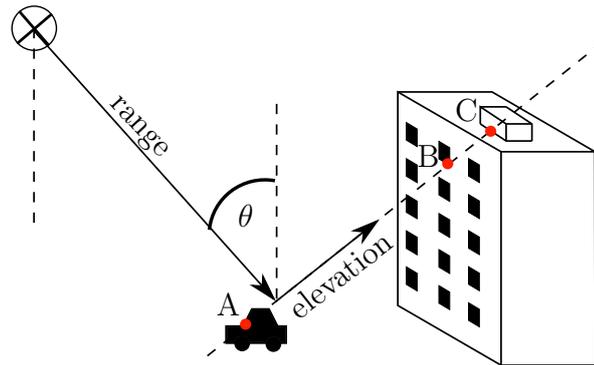
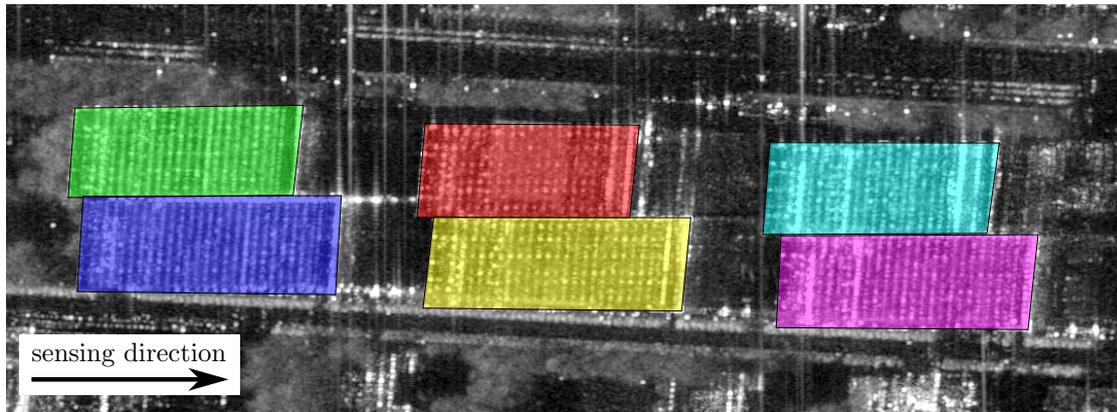


Figure 2.2: Sketch of signal mixing due to layover in one resolution cell of a SAR acquisition. The azimuth direction points into the paper plane. In this example the received signal contains contributions from all three points (A, B, and C).

schematically. The objects on the surface are sensed under the incidence angle θ . Due to the large distance between sensor and scene, the phase fronts are parallel and all points of the same distance from the sensor are approximated as a line. In this sketch, the car (A), a window corner (B), and some structure on the roof (C) have the same distance from the sensor and, therefore, the same range coordinate. This results in a mixture of the integrated signal. A more realistic example is given in Figure 2.3. The areas affected by layover of a housing complex consisting of six multi-storage buildings are shown. In the cyan layover area, for example, a two-story, flat roofed supermarket is fully overlaid with signal contributions originating from the facade of the nearby multi-story building. In order to resolve this two-dimensional mapping of the three-dimensional object, several acquisitions with slightly different sensor positions are exploited. As described in Section 2.1.2, the three-dimensional position of the most dominant scatterer can often be computed. In the following it is assumed that the correct object coordinates can be found by the PSI processing.

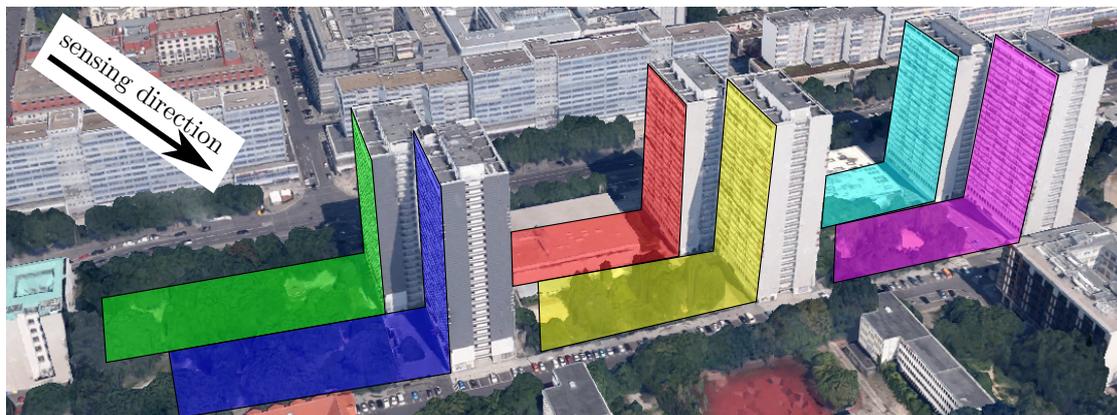
Shadowing and foreshortening are two other effects which are characteristic for the SAR imaging geometry. Specifically high rise buildings lead to shadowing in urban scenes resulting in areas which are invisible to the sensor. This can be mitigated to some extent by combining two different viewing aspects. Given that SAR data from ascending and descending viewing directions are available, this effect can be omitted. Foreshortening occurs when the observed area is bent towards the sensor. Since this thesis solely deals with the appearance of facades in urban areas which are assumed to be vertical, this effect is also no longer taken into account.



(a) SAR acquisition.



(b) Optical nadir image. (©Microsoft Corporation)



(c) 3D model (©Google Earth).

Figure 2.3: Complex of six multi-story buildings in the city center of Berlin. The colored areas mark areas which are affected by layover. (a) Mean amplitude image of a stack of 79 TerraSAR-X acquisitions. (b) Optical nadir image. The presence of a separate buildings in the layover area marked in cyan can be seen as a white rectangular roof (outline highlighted with red dashed line). (c) A three-dimensional city model of the building complex.

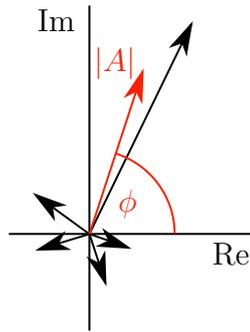


Figure 2.4: Sketch of the complex signal inside one resolution cell. Black vectors symbolize individual scattering contributions. The resulting signal which is the sum of all individual contributions is shown in red. It is dominated by one scatterer with a large amplitude compared to all other scatterers.

2.1.2. Persistent scatterer interferometry

For a given azimuth and range coordinate, the complex valued signal can be expressed as a vector in polar coordinates in which the amplitude corresponds to the length and the phase to the angle with respect to zero. The footprint covered by a distinct azimuth and range coordinate contains many scattering contributions which are integrated into the resulting and measured signal (omitting influences of the atmosphere). Figure 2.4 illustrates the composing of the resulting signal schematically. The black vectors are individual scatterer contributions inside the same resolution cell. The red vector is the summed signal which is received by the sensor. It is dominated by one strong contribution which exhibits a larger amplitude compared to the other scattering contributions. Of course, two strong scatterers with nearly opposed phase values can extinguish each other resulting in a resolution cell with apparently no dominant scatterer present.

The largest advantage of SAR remote sensing is the possibility to analyze the coherent signal over time, and measure movements of the observed objects at the scale of fractions of the used wavelength. In order to do so, a time series of acquisitions covering the scene under investigation is necessary. For TerraSAR-X the orbit repetition time is 11 days which allows for many applications like monitoring of widespread construction sites or anthropogenic subsidence events as a result of mining. The movement of the object in LOS of the sensor can be measured by analyzing the residual phase of at least two acquisitions of the same scene. This is the basic principle of interferometric synthetic aperture radar (InSAR).

Not only does the displacement Φ_{defo} of the object between two acquisitions affect the interferometric phase Φ , but also dissimilar appearances of the object under slightly different viewing directions due to the surface topography Φ_{topo} , different atmospheric conditions

Φ_{atmo} , errors of the estimated orbit position Φ_{orb} , and random errors in the sensor's electronics Φ_ϵ . Thus, the interferometric phase is composed as the sum of these contributions:

$$\Phi = \Phi_{defo} + \Phi_{topo} + \Phi_{atmo} + \Phi_{orb} + \Phi_\epsilon . \quad (2.6)$$

The PSI technique describes the decomposition and computation of the interferometric phase as well as the selection of temporarily stable scatterers where this analysis is feasible. The reduction onto temporarily stable scatterers simplifies the reliable mitigation of atmospheric influences. In the following, the technique is briefly summarized based on the detailed description given in Kampes [2006].

A resolution cell is considered as temporarily stable if the temporal standard deviation of the amplitude is small with respect to the mean amplitude [Ferretti et al., 2001]. This relation can directly be evaluated pixel by pixel through the stack of acquisitions and does not need spatial context which allows the detection of isolated points. A more robust approach is to consider the signal-to-clutter ratio (SCR) which is spatially estimated assuming that a dominant scatterer is surrounded by clutter and that the intensity of clutter can be estimated from resolution cells surrounding the dominant scatterer [Adam et al., 2004].

The phase contribution of the topography can be computed and subtracted by incorporating a digital elevation model (DEM). The resulting interferometric phase is referred to as *differential InSAR* phase. The height of the PSs are also parameters in the PSI processing. The atmosphere is considered to be temporarily uncorrelated but spatially stable. This behavior can be formulated as a high-pass filter over time and low-pass filter over the spatial domain. A similar contribution which cannot be separated from the atmosphere is the orbit estimation. Typically, a linear or periodic model for the scatterer movement is assumed and used to estimate the resulting deformation contribution of the differential phase. The estimation is conducted in a redundant network-like setup which consists of the most dominant scatterers, i.e. those with the smallest standard deviation with respect to the mean amplitude or highest SCR, respectively. After all contributions to the interferometric phase have been estimated for this coarse network, all PSs are appended to it [Gernhardt, 2012].

In order to assign a three-dimensional Cartesian object coordinate to a PS, the precisely known LOS of the satellite is intersected with the DEM plus the estimated height of the PS. Due to the network-like estimation setup, the heights of all PSs are relative with respect to a manually chosen reference point. Thus, a constant offset of the heights has to be taken into account. This geocoding procedure is explained in detail in Gernhardt [2012].

2.2. Oblique optical imagery

Optical oblique imagery has a long history dating back to the beginning of the 20th century [Slama et al., 1980]. The optical axis is pointed to a direction between the horizontal and the vertical axis resulting in an oblique view of the depicted scene. Typically, incidence angles between 40° to 50° in the scene center are used. Therefore, and as opposed to nadir acquisitions, facades of buildings are visible which allows for many applications like widespread texturing of city models [Frueh et al., 2004] or even their creation [Xiao et al., 2010]. A typical GSD lies in the range of 5 cm to few decimeters depending on the camera, the aircraft in use, the flying altitude, and the pixel position in the image [Gerke, 2009]. In most cases blocks of overlapping images are captured during a flight campaign resulting in the possibility of stereo processing. Combined with the exterior orientation measured by a global navigation satellite system (GNSS) and inertial measurement unit (IMU) on board the aircraft, or via bundle adjustment, three-dimensional positions of the depicted objects can be derived.

Optical oblique imagery is well suited to extract information from building facades. Typical oblique camera systems consist of five cameras directed into the four cardinal oblique directions and one nadir view. This means that always at least one image is available where the planimetric angle between any facade's normal vector projected to the ground and the camera's LOS in the ground plane is at most 45° . This is a sufficiently small angle to reliably extract facade objects like windows or balconies, as long as those objects can visually be distinguished from the background of the facade. In order to automate the extraction, approaches based on the gradient or texture of the intensity values or color are applied. In general, optical imagery of facades deliver very strong evidence of geometrical structures.

The tilted view direction leads to a changing GSD across the image. Considering only single facades, this effect can be neglected due to the fact that the area of the image covered by a facade is usually small compared to the overall acquisition, and the GSD can then be assumed to be constant in this area. Apart from that, oblique images do not differ from other types of aerial imagery, e.g. nadir images. Therefore, the standard central projection model can be applied to this type of data. To map object space coordinates onto an image, the orientation Ω of the camera is necessary. This includes the camera position in object space, its alignment in terms of rotations around the axes of the object coordinate system,

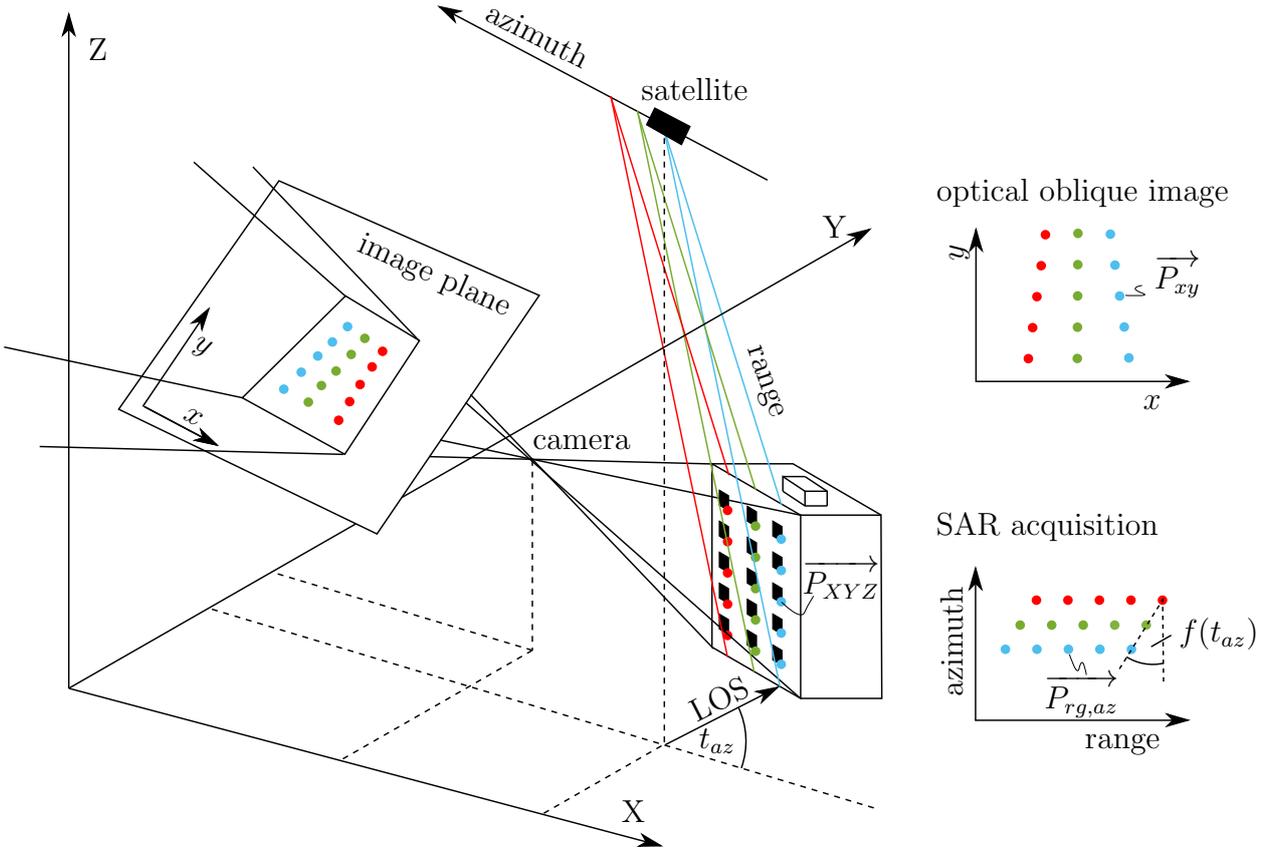


Figure 2.5: A facade mapped in two different imaging geometries. The image taken from the optical camera yields a central perspective. The two-dimensional mapping of the SAR sensor leads to a cylindrical projection.

as well as the interior orientation. The projection from the object point $\vec{P} = \{P_X, P_Y, P_Z\}$ into an image is then performed via the collinearity equations

$$\begin{aligned} x &= x_0 - c \cdot \frac{r_{11}(P_X - X_0) + r_{21}(P_Y - Y_0) + r_{31}(P_Z - Z_0)}{r_{13}(P_X - X_0) + r_{23}(P_Y - Y_0) + r_{33}(P_Z - Z_0)} \\ y &= y_0 - c \cdot \frac{r_{12}(P_X - X_0) + r_{22}(P_Y - Y_0) + r_{32}(P_Z - Z_0)}{r_{13}(P_X - X_0) + r_{23}(P_Y - Y_0) + r_{33}(P_Z - Z_0)}, \end{aligned} \quad (2.7)$$

with r_{ij} elements of a 3×3 rotation matrix, determined by three rotation angles $\{\omega, \phi, \kappa\}$. x and y are the projected image coordinates of the object point \vec{P} . The interior orientation of the camera is determined by the principle point (x_0, y_0) and the focal length c . The exterior orientation describes the camera position (X_0, Y_0, Z_0) and viewing direction $\{\omega, \phi, \kappa\}$ in the three-dimensional object coordinate system. This equation will be revisited in Section 2.3.2 where the variance propagation from object space to image coordinates is described.

2.3. The inaccuracy budget of persistent scatterers projected into oblique images

As a consequence of dealing with two different sensors, one task is to transform all data into a common coordinate system to accomplish the registration. Since the overall aim of this thesis is to find the equivalents of individual PSs in optical data, the target coordinate system is a single image. Figure 2.5 shows the mapping of facade points in the object space to the two imaging geometries. On the left side, the mapping of an optical camera is illustrated. The object point $\overrightarrow{P_{XYZ}}$ is mapped onto $\overrightarrow{P_{xy}}$. The image of the whole facade is subject to a perspective projection which is illustrated in the right hand part of the sketch. In case of the SAR acquisition, the facade is mapped by a cylindrical projection. The angle of shear of the facade mapping is a function of the angle t_{az} between the sensor's LOS projected to the ground and the facade's normal.

A common feature of both sensing geometries is fundamental for the reasoning throughout this thesis and can be seen by comparing the mappings of the facade points in Figure 2.5:

Observation 2.1 (topology preservation). Both the projection of the optical camera as well as the imaging geometry of the SAR sensor are topology preserving.

This means that neighborhood relations of windows are retained in a SAR acquisition and an optical image, respectively. This is a simple consequence of the type of transformations between the SAR domain, object coordinates and the projection into the image. Both the projective transformation of an optical perspective camera as well as the cylindrical projection of a SAR system are homeomorphisms for points on planes: if point B is the right side neighbor of point A and left side neighbor of point C on a plane in the object space, this is also the case in the SAR image as well as the optical image. Furthermore, a characteristic of the SAR sensing geometry in the case of facades is notable. The cylindrical projection always maps scatterers which are vertically aligned, meaning that their planimetric position is equal, into the same azimuth coordinate. This is indicated in Figure 2.5 as red, green, and blue lines of scatterers.

Figure 2.6 shows the relations between the SAR, object, and image coordinate systems as a flowchart. The transformation from SAR to object coordinates is done by geocoding which needs a DEM as well as the imaging geometry parameters of the radar acquisition. Note that the resulting accuracy is not affected by the precision of the DEM. In order to project three-dimensional object coordinates into an optical image, its orientation is needed. The transformation process from SAR to optical image coordinates and its inaccuracy budget is

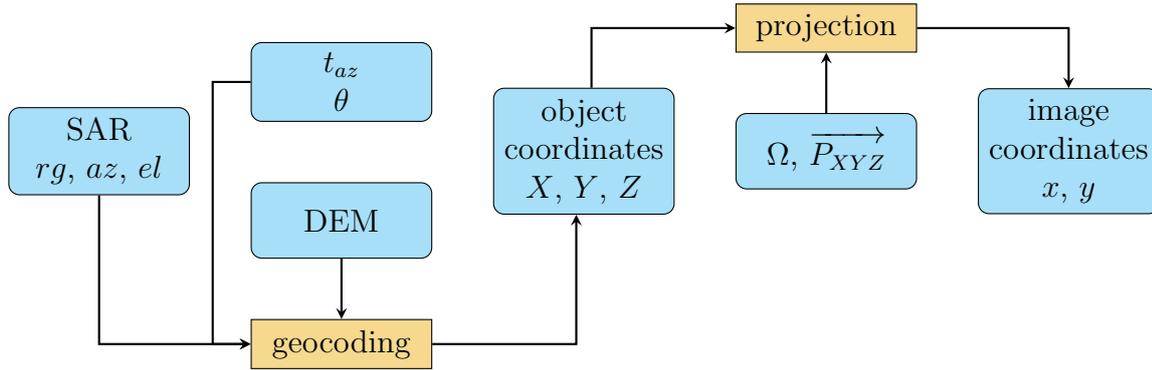


Figure 2.6: Flowchart of the transformation process from SAR domain to image coordinates. Object coordinates are obtained by geocoding. The projection into the image needs the projection parameters.

described in the following two sections. The complete transformation can be split into two separate operations, namely the transformation from SAR to object space and subsequently the projection from object space into the image.

2.3.1. Geocoding of SAR acquisitions

The accuracy of the assignment of PSs to their counterparts in optical imagery depends on the transformations from the SAR domain to object coordinates and from there into the image, as well as the localization precision of a point scatterer in the radar stack. The latter is typically found by assuming an ideal point scatterer which means that the impulse response of the SAR system is a sinc-shaped signal [Bamler & Eineder, 2005]. To detect the exact position with subpixel accuracy, a predefined sinc-function is then correlated with the mean amplitude image of the SAR stack. The estimated error between the predefined sinc and the measured signal is a function of the SNR. Assuming zero mean uncorrelated clutter, the precision for the estimation of the mean of N acquisitions is improved by the factor of \sqrt{N} [Gernhardt, 2012]. The localization precision in range and azimuth of a PS in a stack of N images with a fixed SNR is then:

$$\begin{aligned}\sigma_{rg} &= \frac{\sqrt{3}}{\pi\sqrt{\text{SNR} \cdot N}} \cdot \rho_{rg} \\ \sigma_{az} &= \frac{\sqrt{3}}{\pi\sqrt{\text{SNR} \cdot N}} \cdot \rho_{az} ,\end{aligned}\tag{2.8}$$

SNR	79 acquisitions			30 acquisitions		
	σ_{rg} [m]	σ_{az} [m]	σ_{el} [m]	σ_{rg} [m]	σ_{az} [m]	σ_{el} [m]
10	0.012	0.022	0.269	0.019	0.035	0.436
5	0.016	0.031	0.380	0.027	0.050	0.617
2	0.026	0.048	0.601	0.042	0.078	0.975

Table 2.1: Examples of typical localization precisions of single PSs in SAR coordinates.

where ρ_{rg} and ρ_{az} are the spatial resolutions as described in equations (2.1) and (2.5), respectively. The Cramér-Rao lower bound for the localization precision of the third coordinate, i.e. elevation, of PSs can be approximated by [Bamler et al., 2009]:

$$\sigma_{el} \approx \frac{\lambda r}{4\pi \cdot \sqrt{2} \text{SNR} \cdot N \cdot \sigma_B}, \quad (2.9)$$

where λ denotes the system's wavelength and r the distance from the sensor to the PS under investigation. The accuracy of the elevation direction is a function of the observational basis which can be expressed as the standard deviation σ_B of the baseline B perpendicular to the viewing direction. A wider baseline improves the localization precision in elevation direction. In the case of TerraSAR-X, the absolute orbit positions are limited to a tube with a radius of 250 m which also limits σ_B for a fixed number of acquisitions N [Yoon et al., 2009]. For a given scene and sensor, the SNR is fixed. Since the wavelength can also not be influenced from an operator's perspective, the only possibility to improve the accuracy in elevation direction is to increase the number of acquisitions. Typically, the three SAR coordinate axes are assumed to be uncorrelated which leads to a covariance matrix of the form

$$\Sigma_{\mathbf{RAE}} = \begin{bmatrix} \sigma_{rg}^2 & 0 & 0 \\ 0 & \sigma_{az}^2 & 0 \\ 0 & 0 & \sigma_{el}^2 \end{bmatrix}. \quad (2.10)$$

Table 2.1 gives an example of typical values for the localization precision in range, azimuth and elevation (equations (2.8) – (2.9)) by reference to one ascending SAR stack used in the experiments Chapter 5. The standard deviation of the perpendicular baseline is 156 m and the distance from the sensor to the scene center is 673 308 m. The number of acquisitions is 79 and a SNR of 10 dB according to Gernhardt [2012] can be assumed. Since this very high number of available images in one stack exceeds the typical application case, a more realistic stack size of 30 is assumed as well (all other parameters are unchanged). In all scenarios the resolutions in range and azimuth are set to $\rho_{rg} = 0.6$ m and $\rho_{az} = 1.1$ m. Furthermore, worse SNR conditions as they may occur are taken into account. The resulting numbers show two

issues. First, the localization precision in elevation direction is roughly 20 to 30 times worse than in range and azimuth direction. This leads to a highly anisotropic inaccuracy budget in geocoded three-dimensional coordinates. Second, even for a low SNR and a stack size of 30 images the localization precision can be expected to be smaller than the typical extent of a facade object like a window. Even a three-dimensional absolute inaccuracy of roughly one meter, as is the case for a SNR of 2, is smaller than the expected spacing of horizontally or vertically aligned facade objects. This means that the assignment of single PSs to individual repeated objects at facades should not be prevented by too inaccurate SAR coordinates, but this of course, depends on the specific facade. Note that this is not the case for the inaccuracy induced by the projection from object coordinates into the image which will be treated in Section 2.3.2.

The cylindrical projection of the SAR imaging geometry can be approximated by two rotations $\mathbf{R}_y(\theta)$ and $\mathbf{R}_z(t_{az})$ around the Y and Z axes of the object coordinate system, respectively. Thus, the SAR-coordinate $\overrightarrow{P_{RAE}}$ of a given object point $\overrightarrow{P_{XYZ}}$ can be expressed as

$$f : \overrightarrow{P_{RAE}} = \mathbf{R}_y(\theta) \cdot \mathbf{R}_z(t_{az}) \cdot \overrightarrow{P_{XYZ}} . \quad (2.11)$$

Since $\mathbf{R}_y(\theta)$ and $\mathbf{R}_z(t_{az})$ are rotation matrices, they are orthogonal and the inverse operation can therefore be expressed as

$$g = f^{-1} : \overrightarrow{P_{XYZ}} = \mathbf{R}_z(t_{az})^\top \cdot \mathbf{R}_y(\theta)^\top \cdot \overrightarrow{P_{RAE}} . \quad (2.12)$$

In order to derive an accuracy estimation for geocoded PS in object coordinates, the variances from (2.8) – (2.9) are propagated using the law of variance propagation applied to the transformation g :

$$\Sigma_{\mathbf{XYZ}} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} & \sigma_{XZ} \\ \sigma_{XY} & \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{XZ} & \sigma_{YZ} & \sigma_Z^2 \end{bmatrix} = \mathbf{F} \Sigma_{\mathbf{RAE}} \mathbf{F}^\top , \quad (2.13)$$

with F being the Jacobian matrix of the coordinate transformation g :

$$\mathbf{F} = \begin{bmatrix} \frac{\partial g_1}{\partial P_{rg}} & \frac{\partial g_1}{\partial P_{az}} & \frac{\partial g_1}{\partial P_{el}} \\ \frac{\partial g_2}{\partial P_{rg}} & \frac{\partial g_2}{\partial P_{az}} & \frac{\partial g_2}{\partial P_{el}} \\ \frac{\partial g_3}{\partial P_{rg}} & \frac{\partial g_3}{\partial P_{az}} & \frac{\partial g_3}{\partial P_{el}} \end{bmatrix} , \quad (2.14)$$

symbol	substitution
$\cos(t_{az})$	A
$\sin(t_{az})$	B
$\cos(\theta)$	C
$\sin(\theta)$	D

Table 2.2: Substitution of trigonometric terms.

where g_i denotes the i th entry of the three element vector resulting from applying g . Substituting the trigonometric functions with the expressions of Table 2.2 makes the structure of the resulting covariance matrix $\Sigma_{\mathbf{XYZ}}$ explicit:

$$\begin{aligned} \Sigma_{\mathbf{XYZ}} &= \begin{bmatrix} \sigma_X^2 & \sigma_{XY} & \sigma_{XZ} \\ \sigma_{XY} & \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{XZ} & \sigma_{YZ} & \sigma_Z^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{el}^2 A^2 C^2 + \sigma_{rg}^2 A^2 D^2 + \sigma_{az}^2 B^2 & AB(\sigma_{el}^2 D^2 - \sigma_{el}^2 - \sigma_{rg}^2 D^2 + \sigma_{az}^2) & ACD(\sigma_{el}^2 - \sigma_{rg}^2) \\ & \sigma_{el}^2 C^2 B^2 + \sigma_{rg}^2 B^2 D^2 + \sigma_{az}^2 A^2 & -C^2 B^2 D^2(\sigma_{el}^2 - \sigma_{rg}^2) \\ & sym. & \sigma_{el}^2 D^2 - \sigma_{rg}^2 D^2 + \sigma_{rg}^2 \end{bmatrix} \end{aligned} \quad (2.15)$$

The anisotropy of the *RAE*-space results in correlations between the coordinate axes in the *XYZ* object space. Assuming $\sigma_{rg}^2 = \sigma_{az}^2 = \sigma_{el}^2$ would cause the off-diagonal elements to vanish. Due to the mentioned high discrepancy between the accuracies in azimuth and range on the one side and elevation on the other side, the covariances are always taken into account throughout this thesis. Typical values for $\Sigma_{\mathbf{XYZ}}$ lie in the lower meter domain.

As stated in Section 2.1.2, the absolute PS heights are biased by a constant offset which is a direct result of the PSI processing in which all heights are appended to a manually selected reference point. In the experiments conducted later on, this offset is corrected manually by applying the shift which minimizes the sum of distances between manually selected and projected PSs and their correspondences in oblique images over the whole investigation area. A method to mitigate the systematic offset induced by the reference point automatically is to merge PS point clouds originating from two SAR stacks with opposing viewing directions and common reference points as introduced in Gernhardt [2012]. Furthermore, the influence of the accuracy of the incidence angle θ which changes over the scene, as well as the azimuth angle t_{az} can be omitted since they are already considered in the PSI processing [Adam et al., 2003] and negligible small compared to the influence of the remaining inaccuracy induced by the elevation direction [Gernhardt, 2012].

2.3.2. Projection of single PS object coordinates into oblique images

To derive an estimation for the accuracy of object coordinates of individual PSs projected into the oblique image the variance propagation procedure is continued. In this section, only single scatterers are considered, the extension to a set of PSs is given in Section 4.6.

The Jacobian of the collinearity equations (2.7) (page 25) with respect to object coordinates and projection parameters are denoted with \mathbf{G} :

$$\mathbf{G} = \left[\underbrace{\begin{array}{ccc} \frac{\partial x}{\partial P_X} & \frac{\partial x}{\partial P_Y} & \frac{\partial x}{\partial P_Z} \\ \frac{\partial y}{\partial P_X} & \frac{\partial y}{\partial P_Y} & \frac{\partial y}{\partial P_Z} \end{array}}_{\text{PS-related}} \mid \underbrace{\begin{array}{ccccccccc} \frac{\partial x}{\partial c} & \frac{\partial x}{\partial x_0} & \frac{\partial x}{\partial y_0} & \frac{\partial x}{\partial X_0} & \frac{\partial x}{\partial Y_0} & \frac{\partial x}{\partial Z_0} & \frac{\partial x}{\partial \omega} & \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \kappa} \\ \frac{\partial y}{\partial c} & \frac{\partial y}{\partial x_0} & \frac{\partial y}{\partial y_0} & \frac{\partial y}{\partial X_0} & \frac{\partial y}{\partial Y_0} & \frac{\partial y}{\partial Z_0} & \frac{\partial y}{\partial \omega} & \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \kappa} \end{array}}_{\text{image-related}} \right], \quad (2.16)$$

where $\{x, y\}$ are the image coordinates of the PS. An important fact, which will be taken into account during the derivation of the method, is the difference between two types of inaccuracy budgets. While the uncertainty in SAR and object coordinates is individual for every PS, the projection parameters and their inaccuracies are the same for all PSs. The first type of inaccuracy is denoted as *PS-related*, while the latter is called *image-related* throughout this thesis.

Analogous to the aforementioned geocoding, the covariance matrix of single PSs in two-dimensional image coordinates is derived as

$$\Sigma_{\mathbf{xy}} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} = \mathbf{G} \begin{bmatrix} \Sigma_{\mathbf{XYZ}} & 0 \\ 0 & \Sigma_{\Omega} \end{bmatrix} \mathbf{G}^{\top}, \quad (2.17)$$

where Σ_{Ω} contains all the variances of the interior and exterior orientations. The variances of Ω typically are correlated depending on the specific imaging geometry. As a simplification for the presented method those parameters are assumed to be independent of each other

Parameter	σ	source	impact
PS position range	3 cm	Section 2.3	PS-related
PS position azimuth	3 cm	Section 2.3	PS-related
PS position elevation	70 cm	Section 2.3	PS-related
focal length c	44 μm	camera calibration	image-related
principle point x_0, y_0	13 μm	camera calibration	image-related
camera position X_0, Y_0	10 cm	GNSS/IMU data sheet	image-related
camera position Z_0	20 cm	GNSS/IMU data sheet	image-related
camera orientation ω, ϕ, κ	0.015°	GNSS/IMU data sheet	image-related

Table 2.3: A priori accuracies of parameter set.

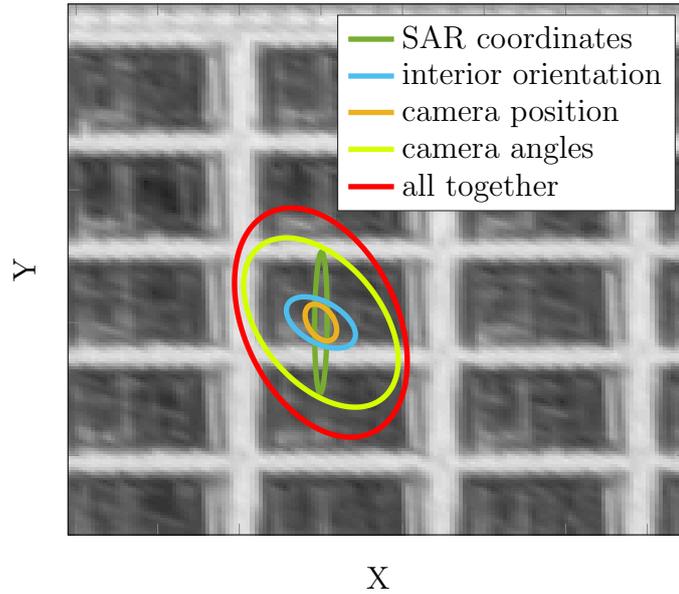


Figure 2.7: Contributions to inaccuracy budget shown as 95% confidence ellipses.

which means that only the main diagonal of Σ_{Ω} carries entries. In order to obtain an impression of the relations with respect to the contribution of $\Sigma_{\mathbf{XYZ}}$ and Σ_{Ω} to the total inaccuracy budget, the variance propagation considering only a single PS is performed for different scenarios. Table 2.3 lists the employed standard deviations. The PS accuracies are derived following the considerations stated in Section 2.3. The cameras of the oblique imagery were calibrated and the resulting accuracies are known. The information for the exterior orientation can be estimated from the sensor specifications of the used GNSS and IMU. Figure 2.7 shows an example in the form of 95% confidence ellipses in the image.

Totally neglecting the inaccuracy budget of the projection into the image results in the green confidence ellipse. In this case, the covariance matrix is solely determined by the inaccuracies of SAR coordinates. The previously described anisotropy becomes clear. The direction of the long semiaxis coincides with the elevation direction of the radar sensor projected into the image which, in this case, is nearly parallel to the Y direction of the image. The inaccuracy budget of the interior orientation (in blue) as well as the camera position (in orange) are comparatively small with respect to the inaccuracy contribution of the camera angles (in yellow). The latter mostly governs the overall inaccuracy budget which is shown in red. A relation to absolute values is given by comparison with the grayvalue image in the background. It shows a section of a regular facade. The window confining wall parts are bright in contrast to the dark windows. Considering the overall inaccuracy budget, the assignment to a distinct window is not reliably possible since the area of the 95% confidence

ellipse is larger than one window. Nevertheless, a method to exploit topology information and model knowledge about the PS distribution in order to reduce this inaccuracy is derived in Chapter 4.

2.4. Graphs

A graph is a common and well established structure to describe and formalize relations between objects or entities of any kind. One example is the investigation of movement patterns in groups of walking persons and individual atypical behavior [Sester et al., 2012]. In such a case, important places of the investigation area are represented as nodes while the paths between them form the set of edges. Another example are network maps of public transport facilities where stations are represented as nodes and the direct connections between them are edges. In general, nodes typically represent objects while edges link nodes and correspond to their bilateral relations. In this thesis, the definition of a graph based on Neuhaus & Bunke [2007] is used and stated as follows:

Definition 2.2 (Graph). A graph \mathcal{G} is defined by the four-tuple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mu, \nu)$, where \mathcal{V} is the set of nodes and \mathcal{E} the set of edges. Each edge connects two nodes. This edge is called *adjacent* to the nodes. Two edges adjacent to the same node are called *adjacent* to each other. μ is a function that labels every node with a weight and ν is a function that labels every edge with an individual weight, respectively.

The cardinality of edges is denoted with $|\mathcal{E}|$ and is the number of edges in the graph, while the cardinality of nodes is denoted with $|\mathcal{V}|$, respectively. Figure 2.8a shows an example of a graph constituted by 7 nodes and 5 edges. The function μ assigns a label (in this example capital letters) to every node while ν attributes a weight to every edge. The minimum cover of a graph is a measure of the graphs' structure in terms of the ratio between nodes and edges. More precisely:

Definition 2.3 (Cover). A cover of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mu, \nu)$ is a set of nodes $\mathcal{K} \in \mathcal{V}$ such that each edge $e \in \mathcal{E}$ has at least one endpoint in \mathcal{K} . A minimum cover of \mathcal{G} is a cover of \mathcal{G} with the smallest number of nodes.

In Figure 2.8a, the nodes $\{A, C, F\}$ form one minimum cover. The concept of a graph's cover is essential for the derivation of the bipartite matching in Section 2.5.

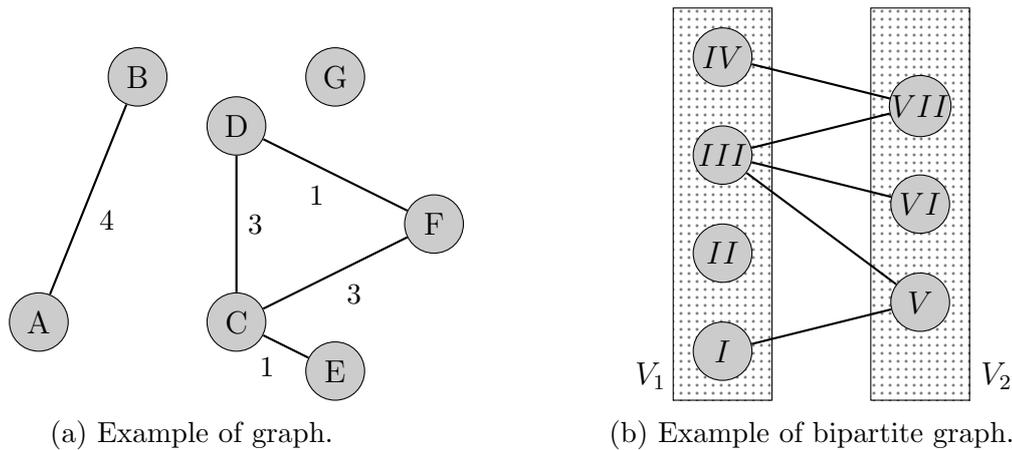


Figure 2.8: (a) A graph with $|\mathcal{V}| = 7$ nodes and $|\mathcal{E}| = 5$ edges. The number next to the edges is the result of the edge labeling function v while the capital letters inside the nodes represent the result of the node labeling function μ . (b) A bipartite graph with the same amount of nodes and edges. The set of nodes is separated into two subsets, $\mathcal{V} = \{\mathcal{V}_1, \mathcal{V}_2\}$, in a way that edges are restricted to relations between these two subsets.

Another application where a graph structure is ideally suitable is the assignment problem. The aim is to find a mapping from one set of entities to another set of entities. This can be represented as a more precisely defined graph, namely a bipartite graph:

Definition 2.4 (Bipartite graph). A bipartite graph is a graph whose set of nodes consists of two disjoint subsets $\mathcal{V} = \{\mathcal{V}_1, \mathcal{V}_2\}$, $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$: $\mathcal{G} = (\{\mathcal{V}_1, \mathcal{V}_2\}, \mathcal{E}, \mu, v)$. The set of edges is restricted to relations between \mathcal{V}_1 and \mathcal{V}_2 .

Figure 2.8b shows an example of a bipartite graph. The set of nodes can be separated according to the definition. Only edges between nodes from different subsets occur.

For the application of the bipartite graph structure in this thesis the disjoint subsets of nodes are specified as the set of all PSs and the set of all window corners extracted from the optical image, respectively. Edges are only allowed between the PSs and window corners. The deviation of suitable edge weights which are used for the matching between PSs and their correspondences in optical imagery is described in Section 4.5.

2.5. Bipartite graph matching

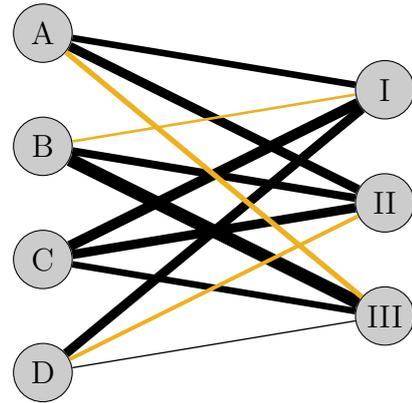
The focus now lies on the application of the bipartite graph structure introduced above to an assignment problem. Given two disjoint sets of elements, the assignment problem asks for a matching between every element from one to an element of the other set. In terms of Definition 2.4, the assignment problem can be formulated as finding a subset of edges $\mathcal{M} \subseteq \mathcal{E}$, of which none are adjacent to each other. More precisely following the definition from Gould [1988]:

Definition 2.5 (Matching). Two distinct edges are *independent* if they are not adjacent. A set \mathcal{M} of pairwise independent edges is called a *matching*.

For many applications, the edges belonging to \mathcal{M} should be selected in a way that the overall sum of their allocated weights is minimized. A vivid example for such problems is the assignment of workers to jobs. The following example is inspired by Burkard et al. [2009]. Imagine a given set of workers and a set of jobs which has to be done. Every worker causes an individual cost to fulfill every job. This bilateral relation can be mapped onto a bipartite graph where one set of nodes are the workers, while the other is the set of jobs. The weighted edges between all workers to all jobs consist of their individual prices. The optimal solution is that all jobs are done with minimum overall costs. Thus, the aim is to select those edges which fulfill these conditions. The situation can also be expressed as a matrix where the rows express the individual prices of every worker while the columns contain the costs for every job. In the following, the two expressions of bilateral relations, namely a bipartite graph and a matrix, are used interchangeably. The matrix in Figure 2.9a shows a cost matrix for four workers and three jobs. The matching minimizing the sum of weights is marked as orange cells. The graphical expression of the same problem is shown in Figure 2.9b. The set of all edges \mathcal{E} is depicted as black lines. The width of the edges corresponds to the individual costs the worker has for the job adjacent to the edge. The result of the bipartite matching which minimizes the overall costs is depicted as orange lines. An important characteristic linking both expressions is the analogy between edges of the matching and the corresponding elements in the cost matrix. Note that every marked cell is *independent*, which means it is the only marked cell in its row and column. The matching is also a *maximum* bipartite matching since no more edges can be added to the subset \mathcal{M} without violating Definition 2.5. Since both subsets of nodes have different cardinality, the node C stays unmatched. Therefore, a maximum bipartite matching for a $m \times n$ cost matrix \mathbf{A} corresponds to $\min(m, n)$ independent cells.

	Job I	Job II	Job III
Worker A	5	7	4
Worker B	2	6	10
Worker C	8	7	5
Worker D	7	3	1

(a) Matrix representation.



(b) Graph representation.

Figure 2.9: The matrix in (a) contains the costs for all worker-job relations. (b) is the graphical representation of the same toy example. The orange lines mark the matching. The thicker the edges, the higher the price for the job performed by the corresponding worker.

The standard algorithm to find a maximum matching exactly was already published by Kuhn [1955] in the mid 50ties and further developed by Munkres [1957] two years later. Kuhn gave it the name *Hungarian algorithm* because his idea was based on previous work by two Hungarian mathematicians named Dénes Kőnig and Jenő Egerváry, namely the following Theorem [Gould, 1988]:

Theorem 2.6 (Kőnig-Egerváry Theorem I). The maximum number $|\mathcal{M}|$ of edges in a matching \mathcal{M} of a bipartite graph $\mathcal{G} = (\{\mathcal{V}_1, \mathcal{V}_2\}, \mathcal{E}, \mu, \nu)$ equals the number of nodes $|\mathcal{C}|$ in a minimum cover of \mathcal{M} .

The proof of this theorem is given in Appendix A.

The Kőnig-Egerváry Theorem can also be restated in the perspective of the matrix representation of the assignment problem. This form creates the basis for the Hungarian algorithm which optimally solves the described assignment problem.

Theorem 2.7 (Kőnig-Egerváry Theorem II). Let \mathbf{A} be a matrix. The maximum number of independent zeros which can be found in \mathbf{A} is equal to the minimum number of lines (either rows or columns) which together cover all the zeros of \mathbf{A} .

Proof. The proof follows directly from the interchangeability of the graph and matrix representation of the bipartite assignment problem. For a given adjacency matrix the corresponding bipartite graph $\mathcal{G} = (\{\mathcal{V}_1, \mathcal{V}_2\}, \mathcal{E})$ can be constructed as follows. The nodes of \mathcal{V}_1 correspond to the rows of the matrix while the nodes of \mathcal{V}_2 to the columns. Therefore,

every matrix element corresponds to one edge between two nodes in disjoint node sets. The bipartite graph and a matching can be constructed by adding an edge between all nodes where the corresponding matrix element is zero. Then a maximum independent set of zeros corresponds to a maximum matching of \mathcal{G} , and following Theorem 2.6, a minimum set of lines covering all the zeros corresponds to a minimum cover of \mathcal{G} . \square

Another fact stated in Flood [1956] is necessary in order to derive the Hungarian algorithm:

Lemma 2.8 (Invariance to constants). If \mathbf{A} is a matrix consisting of non-negative weights a_{ij} the solution of the assignment problem does not change when the elements in \mathbf{A} are replaced by $d_{ij} = a_{ij} - u_i - v_j$, where u_i and v_j are row-wise and column-wise constants, respectively.

Proof. Due to the definition of a bipartite matching, every node can only be matched to, at most, one other node. This corresponds to at most one selection of the corresponding row or column, respectively. Thus, subtracting a row- or column-wise constant cannot interfere with other selections. \square

Applying Theorem 2.7 and the considerations of Lemma 2.8 lets one modify the matrix \mathbf{A} in a way that a minimum number of lines covers all zeros and, thus, a maximum matching is found. The Hungarian algorithm can be described in four steps and is given in Algorithm 1. The weight or cost matrix \mathbf{A} has to be square. If this is not the case, Step 1 adds dummy rows or dummy columns accordingly with equal maximum values. This ensures that if an entity is assigned to a dummy, it can be considered not to be part of the matched set. Step 2 is only conducted once and leads to an updated cost matrix \mathbf{A} that contains at least one zero without changing the relative costs. The following steps are repeated until an optimal solution is found. The König-Egerváry Theorem states that instead of looking for all independent zeros, one can also aim at finding a minimum set of lines covering all zeros in \mathbf{A} . Therefore, Steps 3 and 4 are repeated until n lines are found. Step 4 punishes matrix elements which are covered twice by lines. This is motivated by the fact that such elements cannot be part of the globally optimal matching set, since elements with smaller weights are contained in the set of elements covered by these two lines. This is also the reason why the Hungarian algorithm always yields the optimal result in finite time [Gould, 1988]. A graphical step-by-step example of the Hungarian algorithm is given in Appendix B.

Data: Matrix \mathbf{A} with weights a_{ij}

Result: Optimal solution to weighted assignment problem.

Step 1:

if \mathbf{A} *is square* **then**

$n \leftarrow \dim(\mathbf{A}, 1)$;

else

$\mathbf{A} \leftarrow$ add dummy column(s) or dummy row(s);

end

Step 2:

$\mathbf{A} \leftarrow$ subtract row minima;

$\mathbf{A} \leftarrow$ subtract column minima;

while *no complete and optimal assignment is found* **do**

Step 3:

 Cover all zeros with a minimum number of lines;

$m \leftarrow$ minimum number of lines;

if $m = n$ **then**

 complete and optimal assignment found among independent zeros;

 Stop algorithm;

end

Step 4:

$k \leftarrow$ find smallest element in \mathbf{A} that is not covered by a line of Step 3;

$\mathbf{A} \leftarrow$ subtract k from all uncovered elements and add k to all elements that are covered twice;

end

complete and optimal solution can be found among the zero entries in \mathbf{A} ;

Algorithm 1: Hungarian algorithm for bipartite graph matching.

3. State of the art

Recent developments in remote sensing research which affect the matching of PSs to optical imagery are stated in this chapter in order to present the context of this thesis and highlight its scientific contribution. First, examples for PSI applied to infrastructure monitoring are given in Section 3.1 in order to show that PSI and related techniques are prevailing approaches for urban infrastructure monitoring. Section 3.2 highlights the investigation of PSs, especially at building facades. The exploitation of regular structures in SAR data (Section 3.3), as well as the fusion with optical imagery (Section 3.4), has been performed before to some extent but never with oblique aerial images. Recent approaches for the extraction of facade descriptions are then given in Section 3.5. The mathematical foundation for the method presented in this thesis is graph matching, which has a long history but is also still subject to research (Section 3.6). Finally, in Section 3.7 a discussion of the state-of-the-art is given and the objectives and contributions of this thesis are revisited in this context.

3.1. Recent developments and applications of PSI

InSAR has a long history in geophysical research, like landslide detection or volcano monitoring. These were also the first successful applications, because they could deal with the limited spatial resolution of medium resolution SAR sensors like ERS 1/2 and ENVISAT. The first systematic exploitation of SAR interferometry to a subset of pixels sufficiently coherent over time was given in Ferretti et al. [2001]. A linear movement of 5 mm per year could be detected for a landslide area in Ancona, Italy. Using the same technique with more detailed interpretations, Terranova et al. [2015] systematically investigate the volcanic region around Napoli, Italy. They identify separate regions with different subsidence rates and also relate those movements to their causes like volcano tectonic processes, volume changes in a water reservoir, as well as ongoing subway construction work during the time period of investigation. For many applications, the interpretation of a given PSI result is

a crucial step in order to derive meaningful results which are then the basis for further research. Cigna et al. [2013], for instance, only focus on this interpretation and present a theoretical framework for decision making in the case of landslides, given the PSI results of the investigated area.

The systematic monitoring of urban areas at the scale of single buildings arose with the increased spatial resolution of modern SAR sensors. TerraSAR-X, TanDEM-X, and COSMO-SkyMed are examples for recent spaceborne SAR missions with ground resolutions at the range of 0.5 m to 1.2 m. In this way, the assignment of PSs to individual buildings became possible. Higher ground resolutions also lead to higher PS densities. Gernhardt et al. [2010] first investigated the PS density for stacks of TerraSAR-X High Resolution Spotlight acquisitions. In the city center of Berlin, Germany, a density of up to 120 000 PS/km² was found which corresponds to an average of one PS per $2 \times 2m^2$. Considering typical facades, the distance between consecutive windows is usually larger. That means that as a coarse estimate one can expect a data point for every window as long as these are inducing PSs at all. Lan et al. [2012] show in a case study how the deformation rates and their spatial distributions can be detected for different infrastructure objects. The authors also identify different movement patterns and assign them to individual building types. This segmentation of PSs to single objects remains on the building scale and does not include particular segments of construction. Nevertheless, it is a step towards a semantic interpretation of PS data. An advanced application of PSI is given in Tosi et al. [2013] where long-term measurements of ERS and ENVISAT are used to determine the natural subsidence rate of Venice, Italy. Signals of short term movements induced by anthropogenic interventions like construction sites are separately studied. Those short-term subsidence rates were measured with COSMO-SkyMed and TerraSAR-X over just a few months. This is an example of the broad variety of applications for which PSI has become a standard method, not only in the field of remote sensing, but also in modeling the causes leading to the measured deformations.

In recent years, the deformation estimation was improved in two important ways. First, the orthogonal basis of the satellites position with respect to the sensing target was exploited in terms of resolving the elevation direction and, thus, allowing one to estimate two or even more scattering contributions in one range-azimuth cell. This technique is called tomographic SAR inversion (TomoSAR) and was first described for airborne sensors by Reigber & Moreira [2000]. The expansion to spaceborne sensors in Bamler et al. [2009] and Fornaro et al. [2015] allowed for various new applications since the resulting TomoSAR point clouds are much denser than the PS point cloud derived from the same SAR image stack.

Persistent scatterers can be considered as a special case of TomoSAR points. While for PSI the assumption is that the signal is induced by only one dominating scatterer per resolution cell, TomoSAR assumes several scatterers. This approach mitigates the problems arising with layover effects as discussed in Section 2.1.2. Theoretically, the methods presented in this thesis are independent of the input, thus TomoSAR point clouds and a set of PSs can both be processed. Nevertheless, only PS point clouds are actually considered due to their better availability.

A second very important development is the extent of traditional linear motion estimation towards nonlinear or periodic deformations. Previous approaches all supposed a linear displacement model in order to separate the movement from contributions of the atmosphere and other interfering effects. Crosetto et al. [2015], for instance, extend the standard PSI-technique by explicitly modeling the contribution of thermal expansion of the observed buildings to the deformation. Here the temperature is introduced as an error free observation and constant for the whole scene per acquisition. The authors show that adding this information allows for increasing the amount of detectable PSs due to the explicitly modeled thermal expansion which otherwise had led to incoherence over time. Without explicitly adding the temperature to the estimation process but allowing for a periodicity in the movement model, Gernhardt & Bamler [2012] detect seasonal movements of buildings and bridges. Here the thermal expansion is assumed to be a direct function of time during the year. Another contribution of the aforementioned paper is the fusion of two PSI results derived from image stacks of different viewing directions (ascending and descending, respectively) resulting in a complete data set incorporating facades from all sides of a building. Zhu et al. [2016], finally, combine TomoSAR and the integrated use of different viewing directions into a framework which yields three-dimensional point clouds with a density comparable to LiDAR products, but with a precise deformation information with mm/yr accuracy at every point.

Another tangible example for the applicability of PSI for building monitoring is given in [Ma et al., 2015]. The authors evaluate 44 TerraSAR-X and TanDEM-X strip map acquisitions to monitor the behavior of young concrete buildings. A separation between thermal expansion and subsidence is possible due to the presence of many PSs and the height of some hundreds of meters of the skyscrapers under investigation. Scatterers close to the ground-floor show a movement which can solely be described by subsidence whereas the movement of PSs at the top of the building are verified as the superposition of subsidence and thermal expansion. Furthermore, the thermal expansion coefficient for concrete is derived and found to be in agreement with typical values derived under laboratory conditions. All

of the examples mentioned above are illustrations of a large variety of PSI applications. This technology is already in productive use but still offers great potential towards more detailed analysis of the acquired data, for example, an infrastructure monitoring on the scale of individual building parts.

3.2. Persistent scatterers at facades

All aforementioned examples consider PSs as representative for building blocks or buildings. The assignment to individual facades or, even more detailed, parts of facades was not done so far. Nevertheless, some investigations in that direction already exist. The fundamental motivation that facade structures like window corners induce PSs is given in Bamler et al. [2009]. The authors show that for TerraSAR-X acquisitions in the High Resolution Spotlight mode a trihedral corner with side lengths of 8 cm is sufficient to cause a temporarily stable scattering mechanism. Such corners are often formed by the window sill, the wall and parts of the window frame at many buildings. The first investigations towards assigning PSs to individual parts of facades were conducted by Auer et al. [2010]; Auer [2011] using ray tracing. For this simulation technique a detailed three-dimensional model of the building under investigation is necessary. Knowing the SAR capturing parameters like the sensor's LOS and the wavelength, for instance, the path of the signal can be simulated and compared to the originally obtained image. These investigations confirmed the conjecture that window corners induce threefold reflections resulting in a very strong signal. Similar investigations revealed a fivefold reflection effect, leading to so called ghost scatterers [Auer et al., 2011]. In such cases, the signal is reflected on the ground before and after the scattering in the trihedral corner at the facade. This extends the travel path of the signal resulting in misleading PS coordinates beneath the actual surface level.

Another approach for the interpretation of SAR scenes in urban areas with the help of a SAR simulator is given in Tao et al. [2014]. A LiDAR DEM is fed into the SAR simulator resulting in a virtual SAR scene. Comparing this rendered image with an actual acquisition of the same area allows for outlining individual buildings in the radar image. This corresponds to the interpretation of single scatterers as belonging to individual building facades.

A similar assignment of PSs to facades with the help of building outlines is given in Schunert [2014]. In two case studies, the assignment of scatterers originating from facades to their correspondences in a LiDAR point cloud was investigated. By visual interpretation

of the facades and the respective LiDAR data the author concluded that window corners induced the PSs, which is in accordance with the previously mentioned investigations.

The interpretation of facades in terms of describing TomoSAR point clouds with simple geometrical primitives is shown in Zhu & Shahzad [2014]. First, the point cloud is segmented into clusters of points presumably belonging to the same facade. Vertical planes or curved surfaces are fitted into such sets of facade points. The so derived primitives are then merged into building models which build the basis for precise infrastructure monitoring visualization. Instead of estimating the primitives directly from the data, 2D footprints from a GIS database are incorporated in Zhu et al. [2015] in order to reduce the number of necessary SAR acquisitions for the TomoSAR point cloud inversion. Since the position of the facade in the scene is known, a sparsity constraint is applied, reducing the possible locations of scatterers in the scene and therefore the number of required acquisitions needed to resolve them.

3.3. Grouping of persistent scatterers

One key idea of the approach presented in this thesis is to exploit the regularity of windows at facades. This means that grouping single PSs to more complex objects like lines or lattices, is a crucial step. First attempts in this direction were already made in Stilla et al. [2003] and later improved by Michaelsen et al. [2010]. These approaches are based on Gestalt principles, a perception concept from psychology which was already elaborated in the first half of the last century [Wertheimer, 1938]. These grouping approaches are very data dependent and are not very efficient in terms of completeness and also computational time. Also inspired by Gestalt principles, Schunert & Soergel [2012] later grouped PSs induced by horizontally aligned facade objects in order to improve the three-dimensional coordinates of the whole set. It was shown that the height estimate of the grouped set consisting of N PSs could be improved by the factor of $1/\sqrt{N}$ assuming uncorrelated observations and zero mean noise. Two dimensional building outlines are needed to select those PSs presumably belonging to the same facade.

Another grouping approach was proposed by Auer et al. [2012, 2015]. Their work is based on the Hough transform and Fourier analysis of facade signatures in SAR images, followed by an iterative procedure of completing a lattice structure. The main limitation of this method is that a three-dimensional model of the building under investigation is needed in order to select those PSs belonging to the same facade. Gernhardt et al. [2015] compare the

localization accuracy with a very detailed photogrammetric model of a building with regular facade. Since a very detailed three-dimensional model of the building under investigation is needed, an exhausting capturing with the help of an unmanned aerial vehicle is necessary. The accuracy is then estimated by measuring the deviation of the PSs to a lattice position. In order to estimate the lattice parameters, a voting scheme developed for LiDAR point clouds is applied [Mesolongitis & Stamos, 2012]. A statement about the overall accuracy of PS positions is then made by comparing the lattice parameters to that derived from the photogrammetric model. The difference of the spacing of subsequent structural elements and the one derived from the photogrammetric model lies in the cm domain.

3.4. Fusion of SAR data with other sources

Data fusion between SAR and optical imagery is a popular research topic due to its promising results. The advantages of weather independent microwaves can be paired with the high interpretability of optical acquisitions of the same scene. Especially for change detection, this field offers a plethora of interesting insights. A very general approach for multitemporal data is given in Mercier et al. [2008]. After co-registering two very heterogeneous remote sensing images, as for example, a SAR acquisition and an optical nadir image of the same scene, local context is used in order to infer changed or unchanged areas. This is done by computing local statistics and assuming that in unchanged areas these statistics are correlated to at least some extent even across the sensor domains. This hypothesis is examined for a case where an optical SPOT image was taken before a flood event and changes were inferred with respect to a post-event ERS SAR acquisition.

A more specific approach is presented by Brunner et al. [2010]. The authors derive a simple building model from pre-event optical data and try to fit it to post-event SAR data by simulating the appearance of the particular building in the radar acquisition. This procedure is very prone to errors since the individual steps, the building model estimation, the correct simulation, as well as correctly detecting the deviation of the measured SAR data from the simulated building model cannot be guaranteed to work without problems.

Height information retrieval from fusion of WorldView-2 stereo image pairs and SAR acquisitions is the focus in Xu et al. [2015]. The authors show that the combined use of optical image data and SAR images outperforms the single use of either data alone, compared to a reference height model. The fusion step is performed on the basis of objects after separately calculating building heights by dense matching in case of the optical imagery and

TomoSAR for the SAR data, respectively. The fusion itself is a straightforward averaging procedure in case that both height estimates do not differ too much or the maximum in case of very different heights.

Kuny et al. [2014] and Hammer et al. [2014] produce a shadow and layover mask from a 3D building model derived from optical imagery of an urban area in order to detect changed areas in SAR acquisitions after an earthquake event. The key idea is to identify areas which are assumed to be shadowed by buildings but exhibit bright pixels, i.e. a strong signal. This is interpreted as a collapsed building. A drawback of this approach are high rise objects, like trees, which are present in the real world but not contained in the 3D model. Such an approach could benefit from exact knowledge about the scattering mechanisms by refining the interpretation from detecting completely collapsed houses, towards the possibility of finding partly destroyed buildings.

Instead of incorporating SAR and optical data, Hung et al. [2011] fuse a PSI result with leveling data in order to improve the accuracy of the subsidence measurement. In this study the wide spread characteristic of the PSI processing and the higher spatial resolution are combined with the high-precision sub-millimeter point like measurement of leveling. All of these examples show how SAR products can be combined with other data sources in order to improve the interpretation capabilities of change detection and sometimes make it possible in the first place.

3.5. Facade extraction and description in optical imagery

Even though some interpretation approaches for high resolution SAR data of facades already exist, the majority of attempts to automatically describe building faces are typically encountered in the optical image community. In this section, a short overview over automated facade interpretation in data originating from optical sensors is given.

In Schindler et al. [2008] a method for estimating the camera pose in urban areas by exploiting the regular structure of facades is presented. The method relies on an existing database of geolocated planar surfaces with regular texture. Given an image containing at least one such regular surface a lattice is fitted into the set of SIFT features extracted from the image. All two-dimensional lattices found in the images are then mapped onto the 3D database including the estimation of the position and orientation of the camera. In this approach, the lattice is explicitly modeled and the appearance of each facade can be inferred from one image to the other once the orientation is known. The task in Liu & Liu [2014] is

to detect and describe repetitive facades in oblique aerial images. The facade description is not needed to match a single image to a database but to register a set of airborne images to each other. The vertically and horizontally aligned facade elements are found by patch-wise evaluation of the histograms of edge segments which are projected on hypothesized directions. Projecting horizontal facade element edges onto the actual horizontal direction leads to a sparse histogram. The sparsity is measured by the Gini-index. Then, lattice proposals are expanded greedily. The two last described approaches are methods for detecting regular facades in terrestrial or airborne images, respectively. The description of the regularity is explicitly stated in terms of their spacing and extent, an idea which is also exploited in this thesis.

Reznik & Mayer [2007] also present an approach to extract and describe window outlines from terrestrial images. Window outlines are generated by implicit shape models. In order to do so, the distribution of Förstner points around the center of windows are learned from training data and then used to retrieve similar configurations in unlabeled data. In a second step, the regular alignment of windows at facades is exploited to aggregate individual windows to rows and grids describing the facade. A model selection criterion is then applied in order to choose the configuration which describes all extracted windows with the fewest parameters.

Xiao et al. [2014] use a matrix representation of a single facade. Given a rectified image of the facade, the task is to find a matrix which has a 1 where the repetitive facade element, i.e. windows, are situated and 0 otherwise. This matrix can be decomposed into block matrices which comprise the spacing in two directions as well as one matrix containing the shape of the repetitive window element. This approach needs an initial user interaction when one repetitive element is selected.

Approaches for extracting facade regularities in optical images presented so far are only capable of extracting regularities of the form of simple lattices. Even though this kind of regularity can very often be met in real world scenarios due to the cost effective manner of how office buildings are structured in urban areas, more general forms of regularity are also subject to recent research. One example of differently expressing the regularity is given in Lian & Shen [2015]. The authors extract regular patches of rectified single color images by matching the whole image with template patches comprising the regular structure. These templates are found by aggregating areas of the same color and applying some model knowledge of the expected template size and form. Each template patch is characterized by features derived from the color and the local binary pattern descriptor. In the whole image areas with a high value of this description are then searched in a Markov random

field optimization. The extracted regularity is expressed as segments and their symmetry axes which represent a more implicit form of characterization. A more explicit regularity description is used in the work of Ripperda & Brenner [2009]; Ripperda [2010]. The authors apply a grammar which parses the whole facade out of individual elements like facade parts, windows, doors, or balconies. In order to parse a facade in a given image, probabilities about which grammar rule has to be applied are derived from a set of labeled facades. Depending on the detail of the grammar, arbitrary complex facades can be described at the cost of accordingly more necessary training data.

3.6. Graph matching

Graph matching has a long history dating back to Euler's formulation of *The Seven Bridges of Königsberg* in 1736 [Shields, 2012]. Due to its broad applicability to many problems in engineering, natural science, computer science, and many more, it is still subject to ongoing research. The Hungarian Algorithm and equivalent formulations like the Kuhn-Munkres algorithm [Kuhn, 1955; Munkres, 1957] or similar approaches like the Hopcroft-Karp algorithm [Hopcroft & Karp, 1971] are already optimal and only were improved in terms of computational efficiency. Kao et al. [2001] and Das & Kapoor [2014], for instance, use decompositions of the graph's adjacency matrix for a significant reduction of computational complexity.

Much more substantial developments were made in the field of applying graph matching to remote sensing tasks. Sirmacek & Unsalan [2009] use graph matching for detecting urban areas in optical satellite images. SIFT features are computed for templates of the buildings to search for, as well as for the testing image. After some graph cut operations, individual subgraph matchings are conducted in order to find the best fitting match between the test image and the templates. Elberink & Vosselman [2009] use graph matching to match LiDAR data to a building model. This robust matching is able to handle different appearances of buildings in the data which are not in accordance with the predictions from the model. Robustness is achieved by not only relying on the geometry in the data, but also by using the inherent topology information.

Another example is given in Tuia et al. [2013] where graph matching is used for transfer learning. Here, the classification of land surface usage at different epochs is formulated as two classification tasks in two data domains. The authors derive an unsupervised transformation which maps the classification from one to the other domain. Given the classification result

for the source, all features are represented in a multidimensional space and their centroids are formed. In a similarly segmented target domain, the question is which centroid in the source is mapped to which target. This nonlinear transformation is conducted via graph matching which always guarantees a maximum and optimal matching on the one side but also preserves the structure, and therefore the topology between the centroids.

3.7. Objectives revisited

With the presented background knowledge about recent developments, the objectives of this thesis are revisited and put into context. Even though PSI is a well established technique for subsidence monitoring, especially in urban areas, the assignment of individual PSs is limited to buildings so far. However, the spatial resolution of modern SAR sensors is fine enough to further refine the assignment to building structures. This is the main goal of the thesis at hand. As a source of additional information about buildings, airborne oblique imagery is widely used. In this case, an assignment of PSs to their counterparts in optical images of the same building has to be found. More specifically, it is to be investigated how and with how much accuracy this registration can be done.

These questions belong to the kind of questions often asked in recent years about the physical nature of PSs. Even though the PSI technique has reached practical implementation, many reflection mechanisms of radar rays leading to temporarily stable scatterers remain unknown. As opposed to the investigation techniques based on SAR simulation which need very precise 3D models of the buildings, the goal of this thesis is to rely on single optical images as a strong indication of the actual building structure inducing the PSs. Since it is well known that trihedral structures like window corners typically induce PSs at a facade, they are considered to be meaningful assignment partners for individual PSs. SAR data has already been matched to other data sources like LiDAR models or nadir images, but never to oblique imagery at the scale of individual scatterers.

In achieving this goal, the error prone projection of PSs into the optical image should be mitigated in order to enhance the localization accuracy. For this purpose, two approaches are derived in this thesis. First, the assumption that all PSs at the facade lie in a plane is exploited and, second, the regularity of facade elements is used. The latter is formulated as topology information which should also be preserved during matching. Previous grouping approaches needed explicitly formulated aggregation rules which led to incomplete results. A more general solution is found in this thesis. Besides, many earlier approaches rely on precise

3D models or GIS building outlines. An independent solution would mean independence from such additional data.

Even though techniques for complex facade extraction exist, the approach presented in this thesis remains at the level of simple lattices. This is due to the fact that a simpler model is assumed to be more robust against deviations of the actual PS distribution or image with respect to the prediction of the regularity model.

Not all PSs at a facade always belong to a regular pattern. Those scatterers which do not belong to such regularities should also be considered. Therefore, a geometrical transformation defined by the matching between the set of regular PSs and their counterparts in the image is computed which should reduce a systematic error contribution introduced by a possibly defective interior and exterior orientation of the image. All this leads to a precise technique which can incorporate existing approaches like explicitly modeling the influence of expansion due to temperature changes of monitored building structures, for instance, and combine it with the exact assignment of PSs to parts of buildings.

4. Object-based matching of PSs to optical oblique imagery

In this chapter, a novel method of assigning PSs to features in the optical oblique image is derived. The presented approach is motivated by the fact that optical imagery provides very strong hints on geometrical structures known to induce PSs. The regular alignment in that many facades in urban areas are built is exploited in order to establish the matching. These assignments are then used to minimize the systematic inaccuracies of the projection from PSs into the optical image. This allows to not only establish the assignment of regularly aligned PSs to their counterparts in oblique imagery, but also improves the position of PSs at the facade, which do not belong to the regularity in the image. These positions can then be interpreted in order to learn more about the occurrence of PSs in urban areas.

First, the definition of how *regularity* is used in this method is given in Section 4.1 followed by some crucial assumptions which set the scene for the presented approach (Section 4.2). Then, the method itself is presented in three sections. First, the grouping process of PS point clouds (Section 4.3) and the lattice extraction in optical oblique imagery (Section 4.4) are presented. The matching process itself is subsequently derived in Section 4.5. One application of the resulting matching, namely the estimation of a systematic inaccuracy contribution induced by the camera orientation, is presented in Section 4.6.

4.1. Regularity at facades

Urban areas are often characterized by multistory buildings showing regular patterns of windows at their facades. This regular appearance provides the basis for the presented registration task, since it can be perceived in both types of data, the PS point cloud as well as in optical oblique imagery. In order to exploit this fact systematically, an exact definition of *regularity* as well as the representation in mathematical terms are given in this section.

Due to the very different sensing geometries of SAR sensors and airborne optical imagery, it is an extremely difficult task to find the correspondence of a single PS in the image. Therefore, a single scatterer at the facade is seen as part of the whole facade object. This allows to use topology information of all PSs at the facade and a robust matching of individual scatterers over the data domains. As a very simple yet common facade structure a regular lattice is modeled in the following. In order to reliably derive this object information, the extent of the repetitive pattern has to be large enough to be perceivable as such. Furthermore, the perception of a two-dimensional array of objects as a regular pattern requires some structure of the distances between subsequent elements. The most prominent form of regular alignments is given by equidistant spacings. A lattice \mathcal{L} describes such a regular discrete subgroup and is defined by a basis of lattice spanning vectors and integer repetitions of these [Conway & Sloane, 1999]:

$$\begin{aligned} \mathcal{L} &:= \left\{ a \cdot \vec{t}_1 + b \cdot \vec{t}_2 + \vec{t}_3 \right\} \\ \text{with: } &\{a \in \mathbb{N}_0 \mid 0 \leq a \leq a_U\} \\ &\{b \in \mathbb{N}_0 \mid 0 \leq b \leq b_U\} \text{ ,} \end{aligned} \tag{4.1}$$

where \vec{t}_1 and \vec{t}_2 denote the lattice spanning vectors. \vec{t}_3 can be interpreted as the origin of the lattice. Since all man-made facades are spatially limited, the lattice has an extent defined by the upper bounds a_U and b_U of the integer repetitions a and b , respectively. This description and a restriction for the extent are used to define *regular facades* for this thesis:

Definition 4.1 (Regular facade). A regular facade is a facade with at least 3×3 facade objects which are equally spaced and which can be described by a single lattice \mathcal{L} according to Equation (4.1).

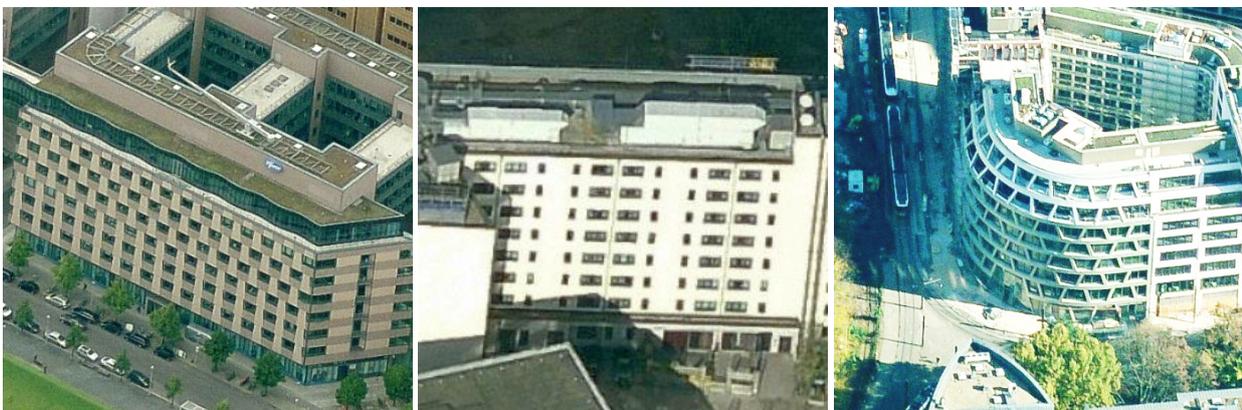
Figure 4.1 shows some examples of facades in the city center of Berlin, Germany, which are regular for a human observer but only some satisfy the definition above. The upper row shows three examples where at least two building faces are in accordance with Definition 4.1. The lower row shows three facades which are not regular following the above mentioned definition. The facade of the building in Figure 4.1d could be composed out of two lattices. Thus, considering the whole facade, this fact infringes on the requirement of having only one set of lattice spanning vectors. Dividing the whole facade into two subsets is a solution to stay consistent with the presented definition. The regularity of the facade in Figure 4.1e is too small. No 3×3 pattern can be found. Of course, a regularity is perceivable for the human interpreter. Nevertheless, a single window is the basis element of the regularity and thus, no sufficiently large lattice can be found. The facade in Figure 4.1f is bent and,



(a) Positive example.

(b) Positive example.

(c) Positive example.



(d) Positive and negative example.

(e) Negative example.

(f) Negative example.

Figure 4.1: Three positive (a) - (c) and three negative examples (d) - (f) of urban regularity according to the derived definition.

therefore, is not a plane and cannot be described by two lattice spanning vectors. Note that even though the whole building face is not a regular facade in cases 4.1d and 4.1f, parts of the facades are in accordance with the above mentioned definition and could be treated as partly regular.

Two important questions at this point are: how often does this type of regularity occur in urban areas and under which conditions do facade elements like window corners induce PSs? The first question is empirically answered for the case study Berlin, Germany in Section 5.2. It is shown that approximately 24% of all facades in the city center of Berlin have been found to be regular according to the aforementioned definition. The second question is partly answered in Gernhardt et al. [2015]. The authors show that the same facade may or may not induce a regular pattern of PSs dependent on the SAR incidence angle. A general rule whether a specific facade is present in the data as a regular pattern currently cannot be derived. Thus, in order to ensure that it is possible to monitor a specific building, one should carefully investigate the sensing geometry in advance.

4.2. Fundamental assumptions

The considerations and methods developed in the following are based on three fundamental assumptions. They are explicitly mentioned for later discussion.

Assumption 1. Facades are regular

Facades in urban areas are characterized by regular and repetitive patterns of facade elements like, for instance, windows or balconies. This regularity can be described by a simple lattice with two constant spacing vectors. It follows that all facade elements lie in the same plane. Furthermore, this plane is assumed to be vertical, i.e. the Z-element of its normal vector is zero.

This assumption is necessary in order to establish the connection between the different sensing geometries. The assignment between a particular facade element in PS data and its counterpart in the optical imagery is established via this required regularity. Of course, this fact limits the approach to the subset of facades showing this strict regularity. In order to successfully extract the regular pattern, the appearance of repeated facade elements has to be discernible from the background of the facade. This is expressed as:

Assumption 2. A facade consists of physically equal and repetitive facade elements.

For any given facade, all facade elements have the same appearance in the SAR acquisitions and optical images, respectively. Furthermore, this repetitive element is discernible from the background of the facade.

This means that every window looks the same in the optical image and that every PS at the facade has undergone the same signal propagation path, respectively. Note that this equivalence is just meant inside one sensor. It is not possible to infer the appearance of a facade object from one sensing domain to the other. The second part simply means that the repetitive element has to be tangible for an automated process. This assumption plays an important role in the process of extracting regular patterns in the oblique imagery which is described in Section 4.4. Here, the regular element needs a gradient in gray values with respect to the facade background, for instance.

Another assumption that is necessary in order to successfully extract the regular pattern in the PS point cloud is:

Assumption 3. At a regular facade, there are more regular PSs than points which are not part of the pattern.

Furthermore, the positions of regular points are significantly more precise in terms of multiples of the lattice defining vectors than the random positions of irregular points.

From this assumption follows that for a given set of PSs the spacing between directly neighbored scatterers are those which occur the most often. This observation forms the initial point of the PS grouping procedure derived in Section 4.3.3.

In order to derive an estimate for the inaccuracy budget of the matched PSs, the contributions originating from the SAR domain and the projection into the image are separated:

Assumption 4. The inaccuracy budget of persistent scatterers in the image consists of a image-related and a PS-related part.

The inaccuracy budget of a single PS is solely determined by the geocoding from SAR to object coordinates and the projection into the oblique image. The former step incorporates the SAR coordinates which are distinct for every PS, whereas the projection is the same for all PSs.

This is a very reduced model of the inaccuracy budget, but captures the largest influences which originate from the elevation coordinate as well as the orientation angles of the optical camera.

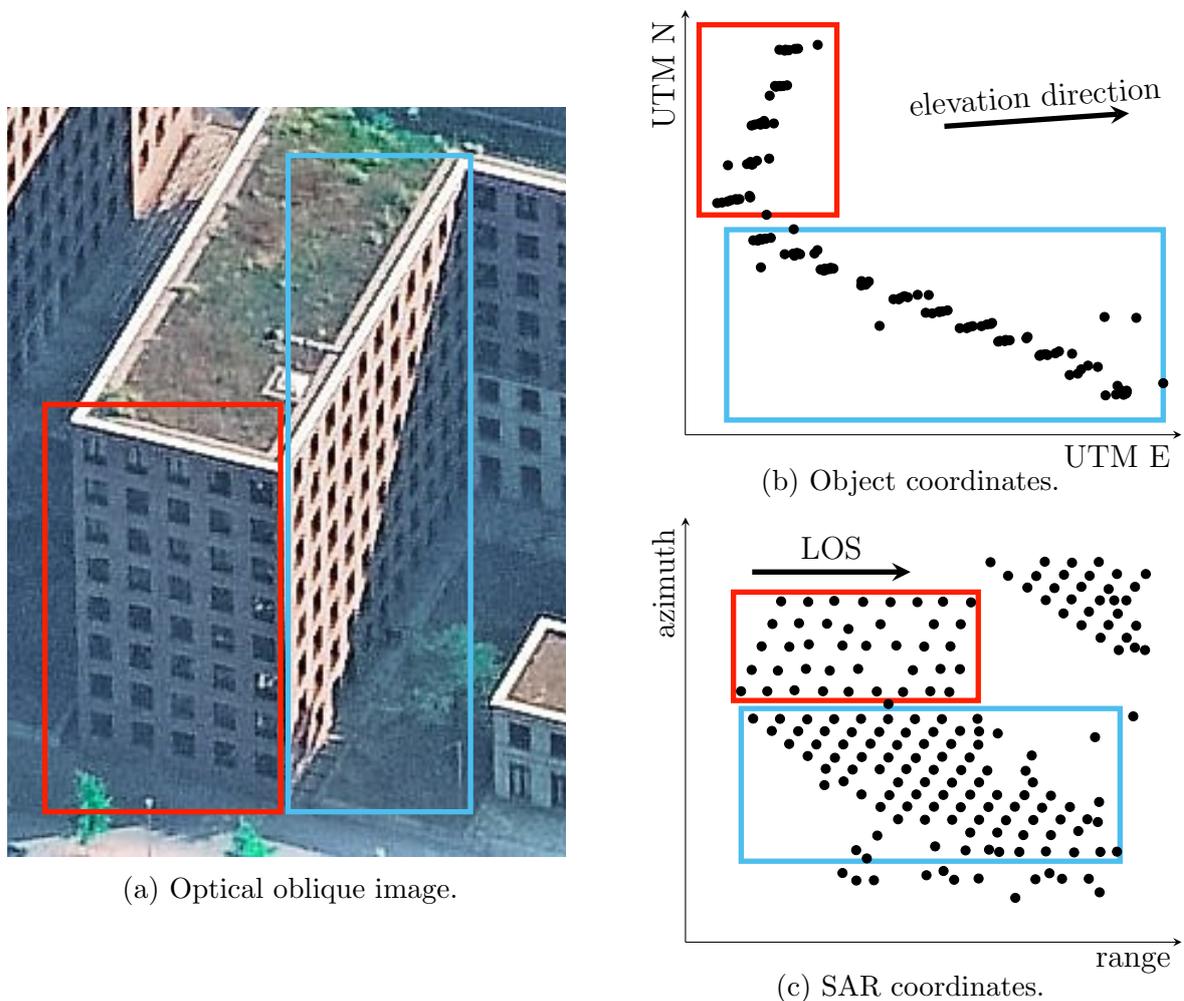


Figure 4.2: Two facades and their appearances in (a) an oblique aerial image, (b) geocoded planimetric UTM coordinates of corresponding PSs, and (c) in the range-azimuth plane. The red and blue rectangle enclose the same facades in their respective domains. The elevation direction and the sensor's LOS are marked with arrows.

4.3. Grouping persistent scatterer point cloud data

The grouping process is based on the observation that in urban areas structures at facades are frequently regularly aligned. In data mapped by a SAR system, this regularity is preserved and leads to salient patterns in the range-azimuth plane. Indeed, this two-dimensional plane is better suited to capture the regularity than incorporating the third coordinate, i.e. elevation, which is significantly less accurate as described in Section 2.3.1, see Figure 4.2. The very different appearances of two facades are depicted in three imaging domains. Figure 4.2a shows an optical oblique image of both building faces. In order to easily find the correspondences in all domains, the facades are marked by a red and blue rectangle. The

planimetric universal transverse mercator (UTM) North and East coordinates of geocoded PSs belonging to these facades are shown in Figure 4.2b. An observation that motivates the following grouping approach can be made here: The anisotropic inaccuracy budget of the elevation direction leads to systematic inaccuracies along the elevation direction. These uncertainties can be fully mitigated by performing the grouping in the range-azimuth plane (Figure 4.2c). By visually comparing the appearance of the red marked facade in Figures 4.2a and 4.2c, one can easily relate single facade elements to each other. In both a 5×8 pattern is discernible (facade marked with red rectangle).

Even though groups of PSs belonging to the same level of the building can better be grouped in the two-dimensional SAR domain, the geocoded three-dimensional coordinates are also necessary in order to perform a segmentation. In the area where the two example facades are adjacent to each other, the assignment of a single facade to one or the other facade is not straightforward. Furthermore, the layover effect may cause PSs induced by roof structures, for instance, to fall into the area of the regular pattern in the range-azimuth plane. Thus, capturing the regular alignment of facade elements requires the segmentation of the total PS point cloud into subsets of PSs belonging to single facades as well as describing the regularity itself. The aim is to find a lattice representation corresponding to (4.1) for every PS subset belonging to a facade. This also means a transition from dealing with single PSs, towards the consideration of facade objects.

4.3.1. Segmenting the persistent scatterer point cloud

In order to assign PSs to parts of buildings, a segmentation into clusters corresponding to single facades is conducted. The problem at hand can be stated in the following form: Given a point cloud of PSs of an urban area, segment it into subsets belonging to individual facades. The total number of facades is not known in advance. A previous version of the following approach can be found in Schack & Soergel [2014].

To solve this task, the structure of PS point clouds in urban areas is exploited. Figure 4.3a shows an example of geocoded PSs in the city center of Berlin. Most PSs originate from facades as the building outline-like structure is perceivable in the top-view of the scene. This finding also matches previous investigations about the distribution of PSs in urban areas [Gernhardt, 2012; Gernhardt et al., 2015]. The taller a facade is, the more PSs are typically induced by it. Since the planimetric position of PSs at consecutive levels is the same, the two-dimensional footprint of the set of all PSs belonging to one facade is very dense along a line which is the building outline. Thus, the local density of PSs in planimetric



Figure 4.3: Facade Segmentation. See text for details.

geocoded coordinates is used as an indication of the presence of facades in the data. A two-dimensional kernel density estimator [Botev et al., 2010] is used to compute the local density. Regions belonging to facades are found by thresholding the density estimation. Figure 4.3b shows an enlarged region of the area. Blueish and yellowish colors indicate areas with high PS density. The outlines of the thresholded areas are drawn in red. Together with the areas corresponding to actual facades some small isolated areas can be identified. These are discarded by sorting out all areas with less than 9 PSs. This number is motivated by the consideration that a facade should have at least a 3×3 lattice to capture the regularity. Inside a segment more than one facade may be present.

Another characteristic of PSs at facades is exploited to further separate adjacent facades. The normal of the plane through the k nearest neighbors of a PS ideally coincides with the facade's normal. This allows for separating adjacent facades like it is the case at building corners. Thus, for every PS the normal of the plane through its k nearest neighbors is computed. The number k is a trade-off between robustness against blunders in the data and robustness against systematically wrong PSs in terms of not belonging to the same facade, especially at the border of adjacent facades. Empirically a value for $10 \leq k \leq 15$ was found to yield good results.

Furthermore, more than one facade with the same normal orientation may be present in one thresholded area. A red arrow marks such a case in Figure 4.3b. In order to further separate these facades, the mutual closeness of PSs is taken into account. PSs belonging to the same facade are situated much closer to each other compared to PSs originating from different facades. This observation is formulated as thresholding over the mutual distances and considering the resulting components as the clustering result. A threshold of 3 m has empirically been found to produce reasonable results. Figure 4.3c shows a segmentation for the enlarged area in SAR geometry. The colors indicate different assignment of the PSs to facades. Again, an arrow marks a group of three facades which are close to each other and two of them with the same normal vector. Finally, Figure 4.3d shows the overall segmentation result in planimetric UTM coordinates. Again, the color marks the membership of a PS to the facade. Note that the colors are assigned randomly to every facade. For the remainder of this thesis, only single facades are considered. Scenes consisting of more than one facade are handled sequentially.

Even though a coarse segmentation into single facade segments is conducted at this stage, some ambitious cases remain unsolved. In particular, the PSs at the edge between facade and roof pose a challenge since they cannot be discerned with the aforementioned approach. Thus, a finer granulation is needed in order to distinguish PSs belonging to repetitive facade

patterns from scattering mechanisms induced from non-repetitive elements at the facade or from the roof and other parts of the building. To tackle this problem, the following three types of PSs are defined:

- **Regular facade points:** PSs in the facade plane that are part of the regular structure.
- **Irregular facade points:** PSs on the facade but not part of the regular structure. Such scatterers are assumed to be induced by facade elements which are not part of the regularity. The deviation from the facade plane is small enough for the PSs to be classified as belonging to the plane.
- **No-facade points:** PSs that neither lie on the regular lattice nor are situated in the facade plane. These points are identified as outliers in the plane fitting step.

The classification of PSs is achieved by robustly fitting a vertical plane to their three-dimensional object coordinates followed by explicitly estimating the facade regularity as described in the following two sections.

4.3.2. Projection of PSs onto the facade plane

Assumption 1 states that all PSs at the facade lie in one plane. In the following this plane is called *facade plane*. As described in Section 2.3.1 the largest deviation for measured PS object coordinates from this assumption is caused by the elevation coordinate. In the following the plane assumption is exploited twofold. First, the above mentioned segmentation into regular, irregular and no-facade points is performed. Second, the actual positions of PSs belonging to the facade are shifted in elevation direction to the common facade plane. This allows for computing the empirical variance of the mean elevation as an estimator for the accuracy in elevation direction. The main idea is to consider the elevation direction of each PS in the facade plane as a repeated measurement of the same value.

Figure 4.4 shows the situation schematically. Every PS has a distance Δ in elevation direction \vec{el} to the plane which is defined by $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. The projected position $\{\Delta, r, s\}$ on the plane is simply the solution to the equation

$$\vec{P}_{XYZ} + \Delta \vec{el} = \vec{v}_3 + r \vec{v}_1 + s \vec{v}_2 . \quad (4.2)$$

The deviations in elevation direction can be considered as independent and following a Gaussian distribution [Bamler et al., 2009]. The expectation value of the distance to the plane in elevation direction is $\Delta = 0$. In order to simplify the problem, Assumption 1 is

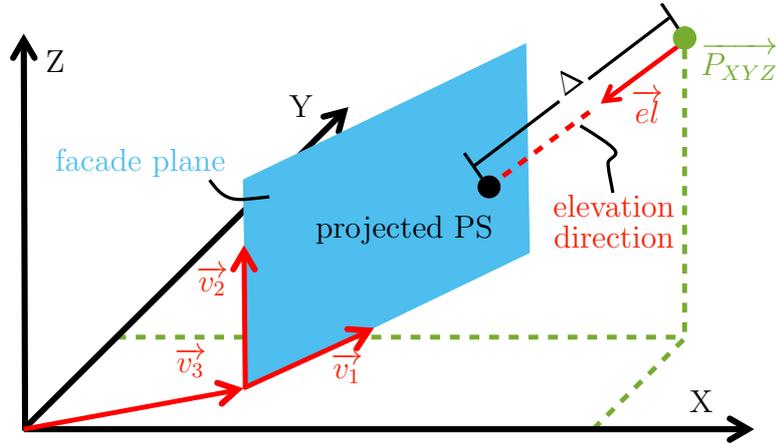


Figure 4.4: Sketch of the plane projection. The PS marked in green at position $\overrightarrow{P_{XYZ}}$ is projected onto the robustly estimated plane (shown in blue) along the elevation direction \overrightarrow{el} . The plane is defined by three vectors $\overrightarrow{v_1}$, $\overrightarrow{v_2}$, and $\overrightarrow{v_3}$.

further exploited by assuming a vertical plane and projecting the elevation direction onto the planimetric plane. The problem at hand is now reduced to fitting a line defined by two parameters which are found by minimizing the distances in the projected elevation direction. This is done as follows. A random sample consensus (RANSAC) inspired scheme [Fischler & Bolles, 1981] is used to robustly select those PSs presumably belonging to the facade. This selection also includes Student's t-test to ensure that the chosen samples come from a normal distribution with mean equal to zero. The threshold for the decision whether a PS is regarded as inlier or blunder is derived from the theoretical variances in elevation direction as derived in Section 2.3.1. Finally, the plane is fitted through all inliers by a least squares adjustment. All inlier PSs are subsequently projected onto this plane.

The benefit of incorporating prior knowledge as, for example, gained by grouping PSs induced by facade objects of the same height, was already theoretically derived in Schunert & Soergel [2012]. Assuming homogeneous PSs, the absence of modeling errors, and zero mean Gaussian noise the estimate of the variance of the Z-coordinate is simply divided by the number of PSs belonging to the group. This method is applied in an analogous way for the plane distance Δ in elevation direction. The empirical mean of all distances Δ_i of N PSs is given by

$$\overline{\Delta} = \frac{\Delta_1 + \Delta_2 + \dots + \Delta_i + \dots + \Delta_N}{N} . \quad (4.3)$$

The standard error of this mean can be seen as an accuracy estimate of all PSs at the facade in elevation direction. It can be assessed by considering all distances Δ_i to be equally precise and independent of each other. Then,

$$\sigma_{\Delta}^2 = N \left(\frac{1}{N} \right)^2 \sigma_{\Delta_i}^2 = \frac{\sigma_{\Delta_i}^2}{N}. \quad (4.4)$$

The empirical variance $\sigma_{\Delta_i}^2$ can directly be computed from the data and measures the scattering of the distribution from the empirical mean. The mean variance σ_{Δ}^2 , however, is an estimator for the deviation of the empirical mean from the mean of the statistical population and can be interpreted as the group accuracy in elevation direction under the assumption that the true position of all PSs is situated on the facade plane. Therefore, the accuracy information in elevation direction for those PSs in the facade plane is updated as

$$\overline{\Sigma_{\text{RAE}}} = \begin{bmatrix} \sigma_{rg}^2 & 0 & 0 \\ 0 & \sigma_{az}^2 & 0 \\ 0 & 0 & \frac{\sigma_{el}^2}{N} \end{bmatrix}. \quad (4.5)$$

Since N is always a positive number larger or equal to 9, because this is the minimum requirement for a regular facade, the resulting covariance information in object space is always smaller than in the case where each PS is seen independently. This point of view constitutes the object-based perspective of the problem. By considering the whole facade object instead of individual and uncorrelated PSs, each point at the facade can be seen as a repeated measurement, and thus allows one to infer the accuracy of the basic population from individual accuracies. Violations of the aforementioned assumption can be caused by PSs that are assigned to the plane but, in reality, are induced by other building structures like the roof, for instance. In such cases, the mean variance underestimates the true inaccuracy budget. Therefore, Equation (4.4) has to be seen as a lower bound for the group accuracy in elevation direction.

Of course, the projection on the facade plane does not make sense for those PSs that are not supposed to be situated on the regular facade pattern. The variance information of such PSs is not considered and the theoretical variance information in elevation direction as described in Section 2.3.1 is applied. In order to obtain a visual impression of the projection step, an example facade is shown in Figure 4.5. The inaccuracies along the elevation direction are striking. The inliers of the RANSAC procedure are marked in red.

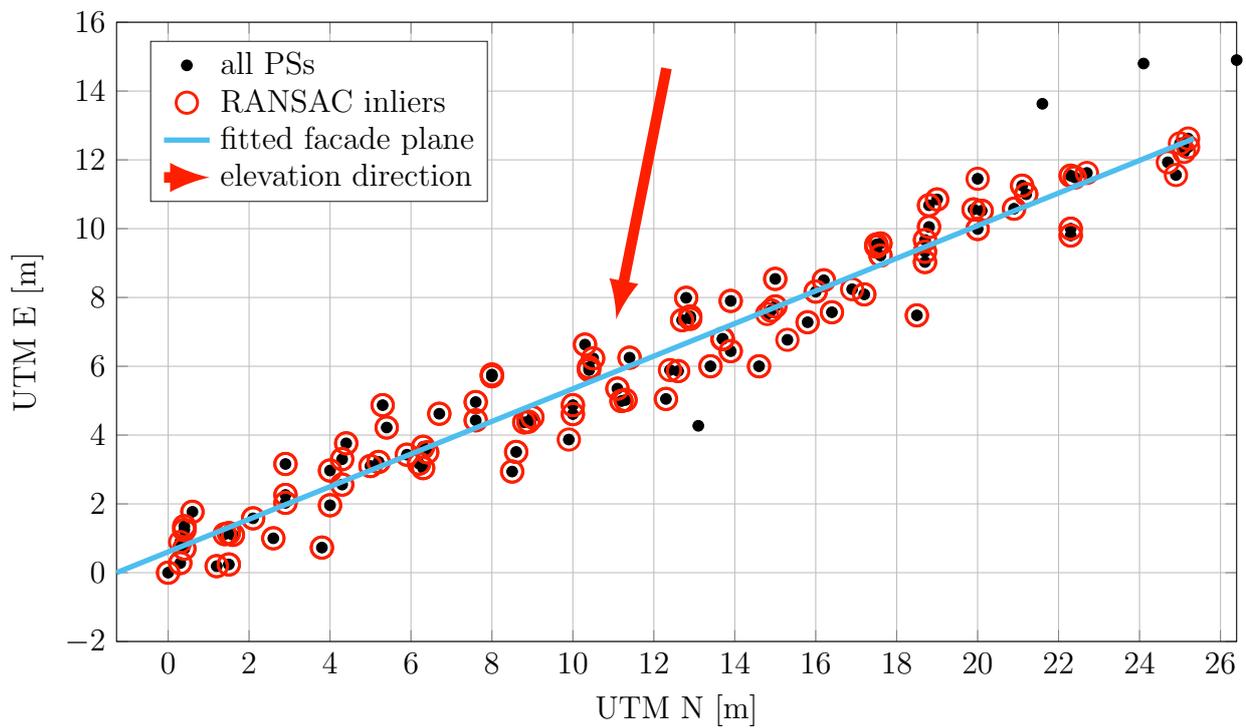


Figure 4.5: Topview of exemplary plane estimation result. All PSs are shown as black points. The inliers of the RANSAC plane fitting are marked with red circles. Red vector: The elevation direction projected to the UTM N - UTM E plane. The robustly estimated plane is shown as a blue line.

4.3.3. Extracting regularity from PS point clouds

At this stage, all PSs belonging to the same facade are segmented accordingly in three-dimensional object coordinates. The facade plane has been estimated and the PSs found to be RANSAC inliers have been projected onto this plane. As discussed before, regularly aligned structures like windows or balconies at a building's facade often induce highly regular patterns in the distribution of PSs in the two-dimensional SAR coordinates. Even though such patterns are very prominent, they are often interrupted by missing data points or interfered by PSs which do not belong to the apparent regularity. Even though such patterns are very prominent, they are often interrupted by missing data points or PSs which do not belong to the apparent regularity. An example can be seen in Figure 4.2c on page 56. Thus, a robust procedure is needed in order to extract and describe such patterns. The discernible patterns are caused by PSs with constant spacings in range and azimuth direction. Therefore, estimating these constant spacings is a key step in order to describe the underlying regularity of the facade. In terms of the lattice definition $\mathcal{L} := \left\{ a \cdot \vec{t}_1 + b \cdot \vec{t}_2 + \vec{t}_3 \right\}$ (see Equation (4.1) on page 52) these values correspond to the horizontal (\vec{t}_1) and vertical (\vec{t}_2) spacing of consecutive facade elements. This spacing occurs between every point and the closest neighbor in horizontal and vertical direction, respectively (with the exception of PSs at the borders). Following Assumption 3, these spacings occur the most often among all possible spacings which is the central idea for the following histogram based voting scheme: All PSs are transformed into a translation space by computing the distances in range and azimuth direction to all other scatterers. Transforming N PSs this way results in N^2 two-dimensional translations. In this translation space, the identity mapping is also included which means that a cluster at $(0, 0)$ with N data points always exist. The search for the lattice defining vectors is split into finding the vector's directions and lengths, respectively. The former is found by a RANSAC inspired robust search procedure which is sketched in Figure 4.6a. The two directions describing the orientation of the underlying regularity are identical with those of two lines through $(0, 0)$. At this point, knowledge about the SAR sensing geometry is exploited: Since the vertically aligned facade structures have the same azimuth coordinate, the \vec{t}_1 direction is already known in advanced (see Section 2.3). The direction collinear with the \vec{t}_2 vector is found by counting the inliers inside a buffer around a line through $(0, 0)$ which is rotated stepwise with rotation interval r . The buffer is defined by the orthogonal distance d from this line. An appropriate value is found by considering the theoretical accuracies of PSs in the range-azimuth plane: d is set to the length of the major semiaxis of the confidence ellipse defined by the covariance matrix of the range-azimuth plane according to Section 2.3.1. Since the \vec{t}_2 direction cannot coincide with \vec{t}_1 , and the latter is known

from the SAR sensing geometry, the directions between $[0, z]$ and $[\pi - z, \pi]$ are excluded from the search. Figure 4.6a shows this excluded area schematically. If Assumption 3 holds, there should be a clear peak in the number of inliers as a function of the corresponding search direction. Figure 4.6b shows a typical example. The data processed here is the same as in the previous example, see Figure 4.5, for instance. The peak of more than 600 inliers is marked in red. Three significantly smaller local maxima can also be identified. Figure 4.6c shows the corresponding result of the transformed space overlaid by the correctly found \vec{t}_2 direction. The cluster-like distribution of translations can be seen. The missing length of the vectors is then found by projecting all inliers onto the line and computing the histogram of their distances to $(0, 0)$. Since the most often occurring spacing is the one closest to the origin, the global maximum excluding the identities is assumed to be the length of the lattice defining vector.

If a regularity is present in the data, it will be strengthened in the translation space. Equal or similar spacings in range and azimuth lead to clusters in the transformed space which are also distributed regularly. The clusters are distributed along a lattice whose spanning vectors are the ones defining the underlying regularity. A number of p PSs in the data which are not part of the pattern will lead to $(N - p)N$ outliers in the transformed space. Considering the buffer of width $2d$ along the search direction and the above mentioned assumption it becomes clear that increasing the sample size by computing the translations between all PSs also increases the ratio between regular points and overall points inside the buffered area. As d is chosen in the magnitude of the localization accuracy of single PSs in the two-dimensional SAR geometry, every regular point will contribute with $(N - p)$ regular translations inside the buffered area. Every irregular point, however, only falls into the buffered inliers area when the original position coincides with a multiple of the correct spacing.

Having found the lattice defining vectors, the extent and the final shift onto the PSs have to be computed. The first is done by dividing the distances between the PSs with smallest and largest range and azimuth coordinates by the spacings in \vec{t}_1 and \vec{t}_2 directions. Then, a bipartite matching as introduced in Section 2.5 is applied between the lattice nodes \mathcal{L} and the PSs. The resulting translation vectors between all matched nodes are averaged and applied to the lattice nodes which corresponds to minimizing the sum of distances between all matching partners (an analogue proof can be found in Section 4.5.2 on page 84). With the extracted regularity, the set of PSs can now be classified into three groups as introduced in Section 4.3.1. Regular facade points are those which are inliers for the RANSAC plane fitting as well as part of the regularity. The latter criterion is determined

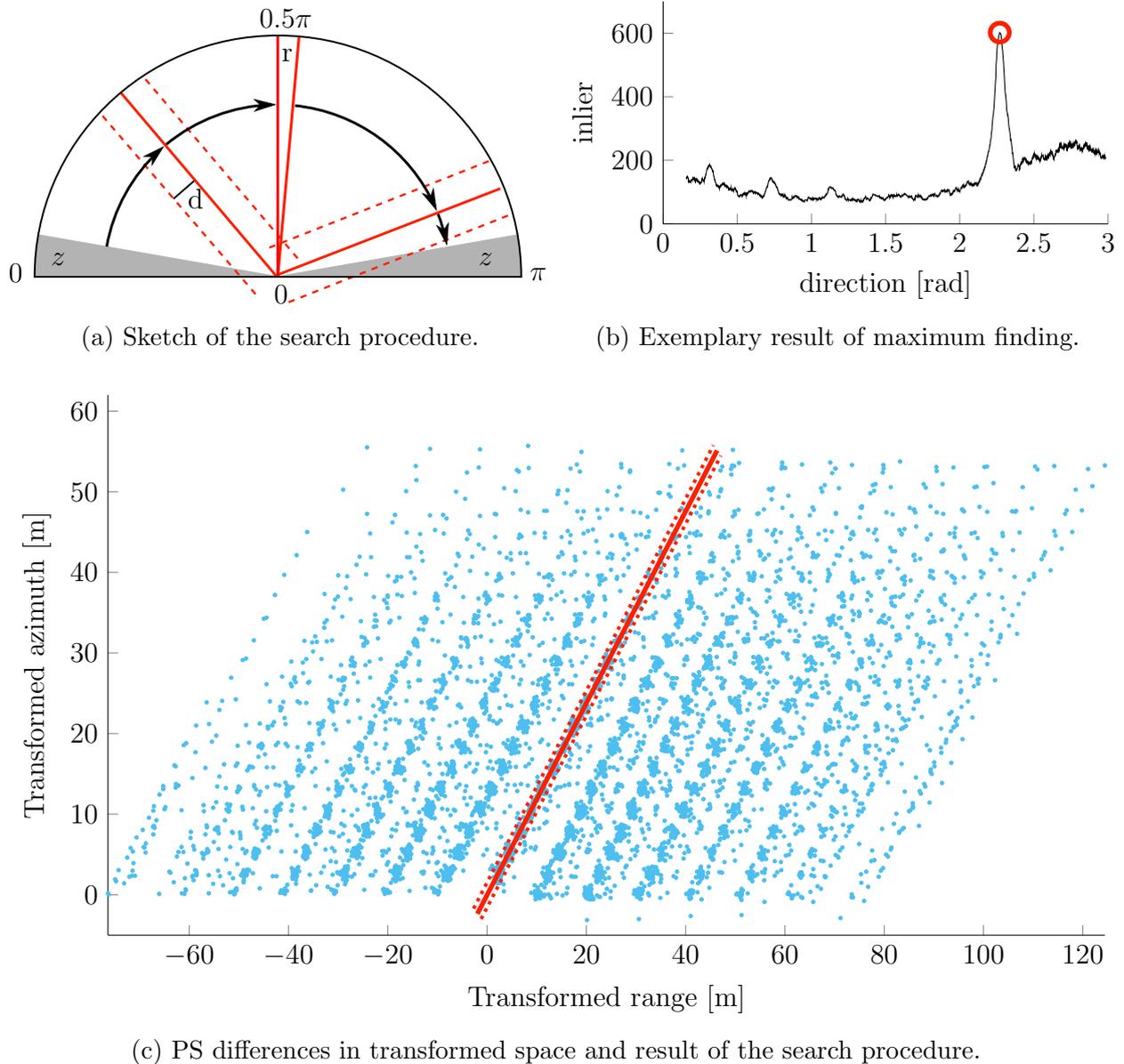


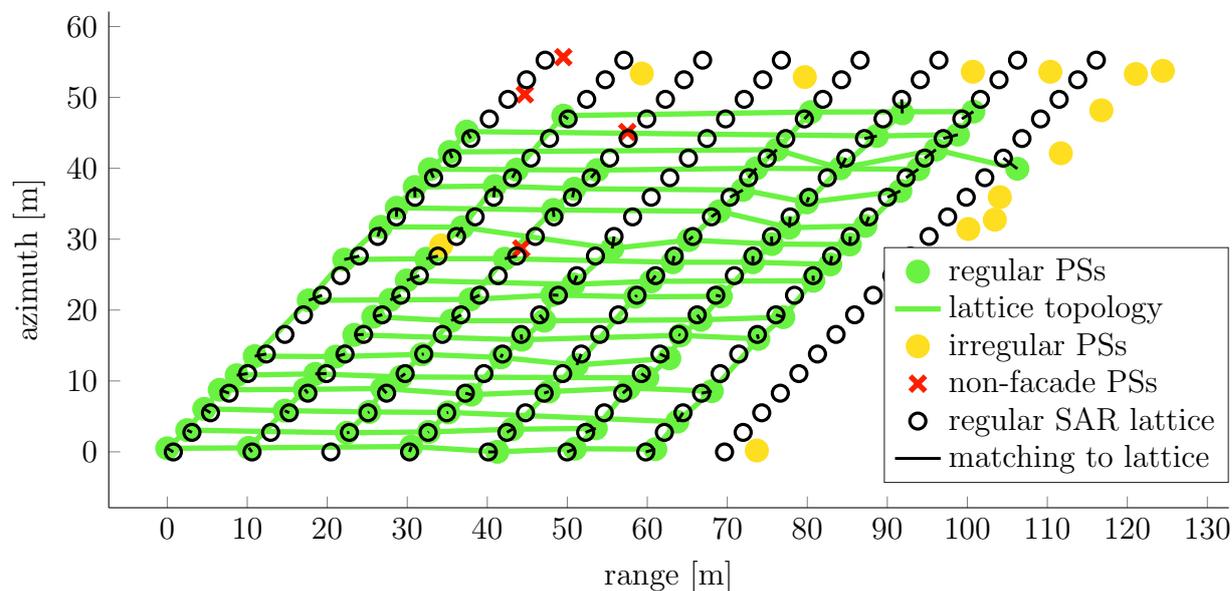
Figure 4.6: Search Procedure to find maximum amount of inliers inside buffer. (a) shows a sketch of the search procedure. The dashed lines mark the buffer around the actual search direction (solid red lines). All directions between z and $\pi - z$ are searched with a sampling of r . (b) Zoom into a result of the search procedure. (c) Blue points are PSs in the translation space. The solid red line is the direction with the maximum amount of inliers. The buffer d is marked with dashed lines.

by thresholding the distances between all PSs and the regular lattice nodes. This threshold is found analogously to the distance d by considering the theoretical inaccuracy budget of a PS in the range-azimuth plane. The second group is formed by irregular facade PSs. Those are inliers of the RANSAC plane fitting but not part of the regular structure. The last group of PSs probably do not belong to the facade as they are not part of the inliers set from the robust plane fitting. Figure 4.7a shows a result of the regularity extraction as well as the classification. The fitted lattice nodes are shown as black circles. The green, yellow and red markers correspond to the regular, irregular, and non-facade PSs, respectively. The matching vectors between the regular PSs and the lattice nodes are shown as black lines. Green lines correspond to the edges connecting adjacent regular PSs. The influence of the projection to the facade plane and the result of the classification in the image can be seen by comparing Figure 4.7b with Figure 4.7c. The first figure shows the image coordinates of the PSs without projecting them to the common facade plane. In Figure 4.7c, however, this projection was done and results in positions which makes clear that the assignment of PSs to individual windows is feasible. The irregular PSs at the border of the facade belong to windows which are correctly classified as not belonging to the dominant regularity.

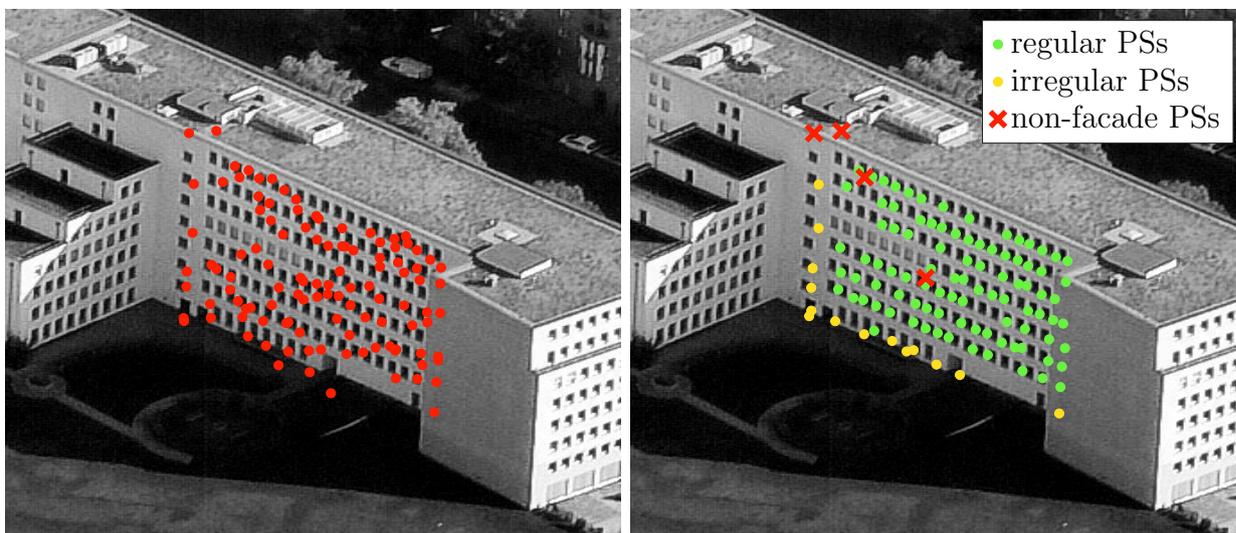
Having assigned each regular PS to a lattice node (black circles in Figure 4.7a), the lattice spanning vectors can also be computed in three-dimensional object coordinates. This is done by considering the lattice representation $\mathcal{L} := \{a \cdot \vec{t}_1 + b \cdot \vec{t}_2 + \vec{t}_3\}$ and averaging all coordinate differences along sets of PSs with the same indices a or b , respectively. These mean lattice spanning vectors in object coordinates are used for the following step of extracting a regular window corner grid from optical oblique imagery.

4.4. Lattice extraction in optical oblique imagery

Regular point features belonging to window corners extracted from the optical facade image are needed as matching partners for the set of regular PSs. The extraction of those corners in optical oblique imagery is conducted for single oblique images showing the facade from an approximately frontal view. Some occlusions by trees or other objects may be present in the data. The extraction of a set of regularly aligned window corners from optical oblique images comprises three steps: First, the original oblique image is rectified in such a way that the horizontal and vertical alignment of facade objects coincides with the horizontal and vertical image coordinate system. This simplifies the subsequent regularity extraction. Second, the facade image is correlated with a template known to include the regular facade element. A voting scheme on differences of local correlation maxima is used in order to



(a) Result of lattice extraction from PSs. All scatterers were classified into regular (green points), irregular (yellow points) or non-facade PSs (red crosses). The derived topology of regular PSs are shown as green lines. The black circles mark the estimated lattice nodes and black lines their matching to the corresponding PSs. The estimate lattice covers all PSs since its extent is estimated before classifying the scatterers.



(b) Projected PSs into oblique image before plane fitting and classification. All PSs are marked as red points.

(c) Projected PSs into oblique image after plane fitting and classification. Green points: regular PSs. Yellow points: irregular PSs. Red crosses: non-facade points.

Figure 4.7: Result of PS grouping and classification.

estimate the spacing in both directions. This regularity in terms of high cross-correlation values captures the spacings of repetitive facade elements. This forms the basis for the third step, namely explicitly modeling the windows at the facade in order to extract that corner which most probably corresponds to the geometrical structure inducing the PSs. The methods presented in the following are partly based on previous versions which were published as Schack & Soergel [2015]; Schack et al. [2015].

4.4.1. Image rectification

Given an aerial oblique image of a building complex and its surroundings, the aim is to extract regular facade structures like repetitive occurrences of windows or balconies. Note that the task is not to find a facade in the aerial image but to extract the regularity of windows or balconies given the coarse area where the facade is situated. This information originates from the PS grouping procedure as described in Section 4.3. This inter-sensor knowledge transformation significantly simplifies the approach since the area of facades under investigation compared to the whole image area only constitutes approximately 3-5% for all scenes in the conducted experiments. This means that no detection of the facade area in the image is necessary, but only the extraction of the regularity. On the left side of Figure 4.8, red points mark PSs in the oblique image. The minimum and maximum image coordinates of the projected scatterers define a bounding box which is broadened by a buffer B . For instance, a typical value for B is 100 pixel. The following steps of the method are only conducted on the bounding box marked with a black dashed line.

The regularity of facade objects proceed along the vertical and horizontal direction in object space coordinates. Extracting the spacings along two given directions is simpler than estimating the spacings together with the directions. Thus, an intermediate aim is to apply a geometrical transformation which aligns the directions of horizontally and vertically consecutive facade objects with the coordinate axis of the image, and therefore to reduce the complexity from four unknowns (two spacings and two directions) to only the two spacings. The main idea is to find dominant lines in the edge image of the facade assuming that they coincide with the horizontal and vertical direction of the facade. Following Assumption 1 (page 54), facades can be approximated by planes. The projection of a plane into an image is described by a homography with eight degrees of freedom. As mentioned above, a single facade covers only a small part of the overall oblique aerial image, and therefore an affine transformation is used to approximate the homography. Having six degrees of freedom an affine transformation is defined by three identical points. This is a crucial advantage over the

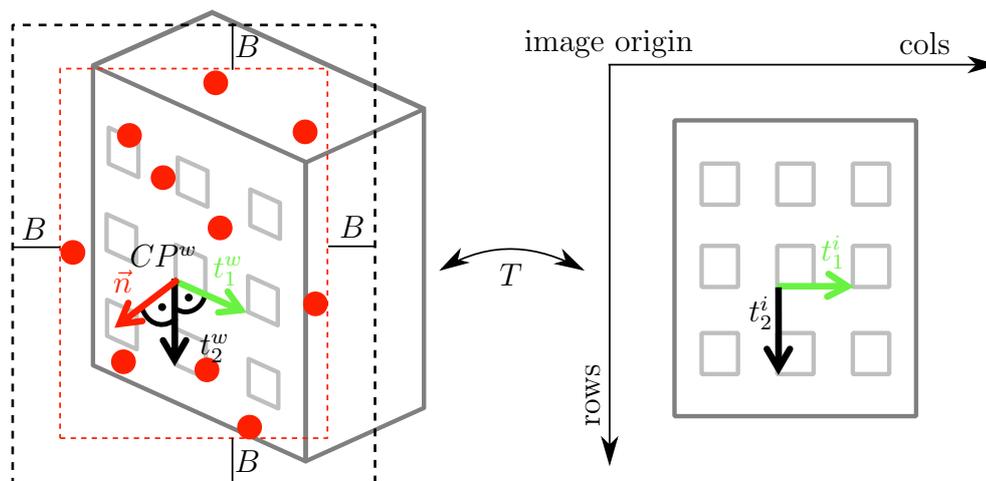


Figure 4.8: Sketch of rectification process. All PS belonging to the facade are marked with red points. Their minimum and maximum image coordinates in the original image plus a buffer B define the region marked with black dashed lines which is transformed into the rectified image via the affine transformation $\mathbf{T}^{\text{affine}}$. In the rectified image \vec{t}_1^i and \vec{t}_2^i are parallel to the image coordinate system.

homography as shown below. The orthogonal coordinate system with its origin in the center of the facade in Figure 4.8 defines the regularity at the facade as well as their orientation in object coordinates. The image coordinates of the origin of the coordinate system CP^w and the two points $CP^w + \vec{t}_1^w$ and $CP^w + \vec{t}_2^w$ are used as the three identical points. The vector \vec{t}_1^w describes the horizontal spacing of consecutive facade elements in object coordinates while the vertical spacing is contained in the vector \vec{t}_2^w . Here, the orientations of the vectors \vec{t}_1^w and \vec{t}_2^w and an arbitrary spacing are sufficient in order to apply the affine transformation. The three identical points are found by setting the origin of the coordinate system to the mean of the PSs projected into the image. Estimates for the horizontal spacing \vec{t}_1^w and vertical spacing \vec{t}_2^w in object coordinates are derived as described in Section 4.3.3. At this stage, the facade regularity can be described as the centroid $\overline{CP^w}$ of all regular PSs and the two vectors $\{\vec{t}_1^w, \vec{t}_2^w\}$ in object coordinates. Due to remaining inaccuracies in the vectors \vec{t}_1^w and \vec{t}_2^w which may be caused by the PS segmentation step, the actual facade orientation in the oblique image is found by identifying salient lines via the Hough transformation [Duda & Hart, 1972]. This approach is motivated by the assumption that both the outline of the building as well as the windows have a high contrast with respect to the background. The edge image is created by applying the Canny filter [Canny, 1986] with standard parameters. Once the directions of the horizontal and vertical alignment of facade objects in the image are found, the affine transformation $\mathbf{T}^{\text{affine}}$ which maps the three points $\{\overline{CP^w}, \vec{t}_1^w, \vec{t}_2^w\}$ to $\{(0, 0), (1, 0), (0, 1)\}$ is computed. The right hand side of Figure 4.8 shows the result

schematically. Extracting the facade regularity in terms of spacings and extent of windows is now considerably simplified since it can be divided into two separate one-dimensional pattern recognition tasks. Following the introduced notation, the spacings parallel to two image axes are denoted as t_1 and t_2 instead of their corresponding vectors \vec{t}_1 and \vec{t}_2 in the non-rectified image. Furthermore, the rectification of the image is assumed to be error-free, and therefore does not influence the inaccuracy budget of the projected PSs into the image. A violation of this assumption would lead to a wrong rectification. Since the image as well as all PSs are projected with the same possibly erroneous transformation, the relative inaccuracy considerations remains unchanged.

4.4.2. Segmenting the facade

The periodicity of facade objects leads to a repetitive appearance of patches also in the rectified image. The facade image can be regarded as a set of aligned patches of one repetitive element plus small deviations from it. One aim is to find this repetitive element and quantify the deviations as a score for the similarity. This view motivates the following approach to capture the regularity and determine its geometry in terms of the spacing in vertical and horizontal direction. More precisely, the spacing between consecutive windows in vertical and horizontal direction is computed by evaluating the peaks of the two-dimensional normalized cross correlation (NCC) function $\gamma(u, v)$. This is a common measure of similarity and compares the resemblance independently of absolute brightness and contrast [Lewis, 1995]:

$$\gamma(u, v) = \frac{\sum_{x,y} (f(x, y) - \bar{f}_{u,v})(t(x - u, y - v) - \bar{t})}{\sqrt{\sum_{x,y} (f(x, y) - \bar{f}_{u,v})^2 \sum_{x,y} (t(x - u, y - v) - \bar{t})^2}} \quad (4.6)$$

where $\mathbf{f}(\cdot)$ is the rectified image and $\mathbf{t}(\cdot)$ the template. (u, v) is the index over the template and (x, y) over the image, respectively. \bar{f} and \bar{t} are the mean values of the corresponding patches. NCC yields a value between -1 and 1 for every pixel where a high value corresponds to a high patch similarity. On the one hand, the NCC is not invariant to rotation, scale and perspective distortions, and may yield false positive results in cases where similar windows occur in the vicinity of the facade under investigation. On the other hand, it is very robust to changing light conditions as imposed by shadows, for instance. Furthermore, due to the very precise approximate values transferred from the SAR domain to the image space, this straightforward method yields reliable results.

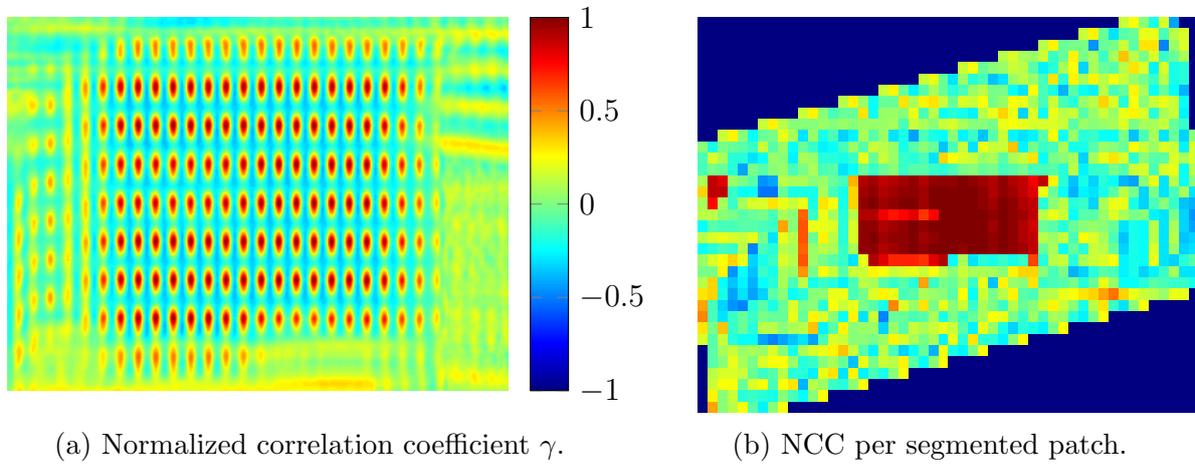


Figure 4.9: Example for lattice estimation via normalized cross correlation. (a) shows an enlarged cut out of the NCC result. The colorbar is valid for all three figures. In (b) the NCC score per patch is shown. The extent of the regular facade is visible. (c) shows the facade segmented into patches colored accordingly to the NCC score γ .

The matching template \mathbf{t} is initialized with the image section from the middle of the facade with a width equal to the spacings found in the PS grouping step and projected into the rectified image. Even though the width does not exactly coincide with the actual spacing in the image, NCC yields the correct result as long as the patch is large enough to contain characteristic features like a window corner. The subsequent processing step derives the spacings t_1, t_2 in both image directions describing the regular pattern, the position of the lattice origin \vec{L}_0 , as well as its extent N_X and N_Y . Furthermore, the NCC score γ as a measure of how similar each patch is compared to a mean patch is returned for every lattice node. The prior rectification step is exploited by separately searching for local maxima \mathcal{Q} along vertical and horizontal direction. Algorithm 2 lists the iterative procedure.

Data: Rectified image \mathbf{f} ; spacings from PS grouping

Result: Position \vec{L}_0 ; spacings t_1 and t_2 ; extent N_X and N_Y ; regularity score γ of every lattice node

Step 0:

Initialize template \mathbf{t} ;

do

Step 1:NCC

$\gamma(\mathbf{f}, \mathbf{t}) \leftarrow$ perform NCC on \mathbf{f} with template \mathbf{t} ;

Step 2: thresholding

$\mathcal{Q} \leftarrow$ local maxima of $(\gamma(f, t) \geq \rho)$;

Step 3: voting

$t_1, t_2 \leftarrow$ maximum voting on $\text{diff}(\mathcal{Q})$;

Step 4: average template

$\mathbf{t} \leftarrow$ compute mean patch with t_1 and t_2 ;

while $\mathcal{Q} \neq \emptyset$;

Step 5: Find extent N_X and N_Y , and position \vec{L}_0

Algorithm 2: Regularity extraction in rectified facade image.

Beginning from the maximum NCC value, the following iterative procedure is performed. The results of the aforementioned initialization (Step 0) and NCC (Step 1) can be seen in Figure 4.9a. Then, Step 2 performs a thresholding on the normalized correlation coefficients. A conservative value of $\rho = 0.8$ was empirically found to yield good results. In order to reduce areas of high correlation coefficients to their center positions, only the local maxima \mathcal{Q} are considered for further processing. Analogous to the grouping procedure of PSs, the most frequent spacing along the X and Y axes is found by computing the differences between subsequent peaks and taking the maximum (Step 3). This is done separately for

the horizontal and vertical direction. In order to improve the template for the following iteration (Step 4), the neighborhood defined by t_1 and t_2 around every peak in \mathcal{Q} is averaged and constitutes the new template \mathbf{t} . The procedure is repeated until no new peak above the threshold is found.

As an intermediate result, one obtains the spacings in the rectified image and a mean patch incorporating the common similarity. With this information, the facade image can be divided into patches. The cross correlation coefficient with respect to the mean patch is computed in order to obtain an estimate of the similarity of each patch to the overall patch which is assumed to be a good approximation of the image of the regular facade element. The result of this is given in Figure 4.9b. The outline of the patches belonging to the regular facade is visible. Step 5 then finds the extent of the regularity and its position in the image in the following way. To decide whether a patch belongs to the regular part of the facade in the image or not, a threshold of the patch correlation score γ can be used, which divides all patches into the two classes. For simplicity, it is assumed that γ follows a bimodal distribution. In such a case, the nonparametric and unsupervised method presented in Otsu [1979] yields the optimal threshold by minimizing the intra-class variances. Furthermore, the local context can be incorporated since a facade element is assumed to never be isolated. This assumption is accounted for by applying a morphological opening and closing operation. Furthermore, a strictly rectangular facade is assumed. Thus, having correctly captured the majority of the facade area allows one to infer occluded parts of the building face. Figure 4.9c shows the result after thresholding. Note that the segmentation at this stage is arbitrary with respect to the windows. In the following step, the window outlines are explicitly estimated.

The presented method is able to explicitly describe the regularity of facades in optical oblique images. The description comprises the spacing in horizontal and vertical direction, the extent of regular facade objects, the position of all regular patches in the image as well as a score of how well a particular patch coincides with the regularity. All this information will be used in the assignment approach as matching partners for PSs.

4.4.3. Explicitly modeling the window

The overall aim of the presented approach is to find correspondences of individual PSs in optical imagery. Recent investigations on simulated SAR scenes and the incurrence of PSs with the help of precise three-dimensional buildings showed that window corners often induce PSs [Auer et al., 2012; Gernhardt et al., 2015]. The sill, a part of the wall and the

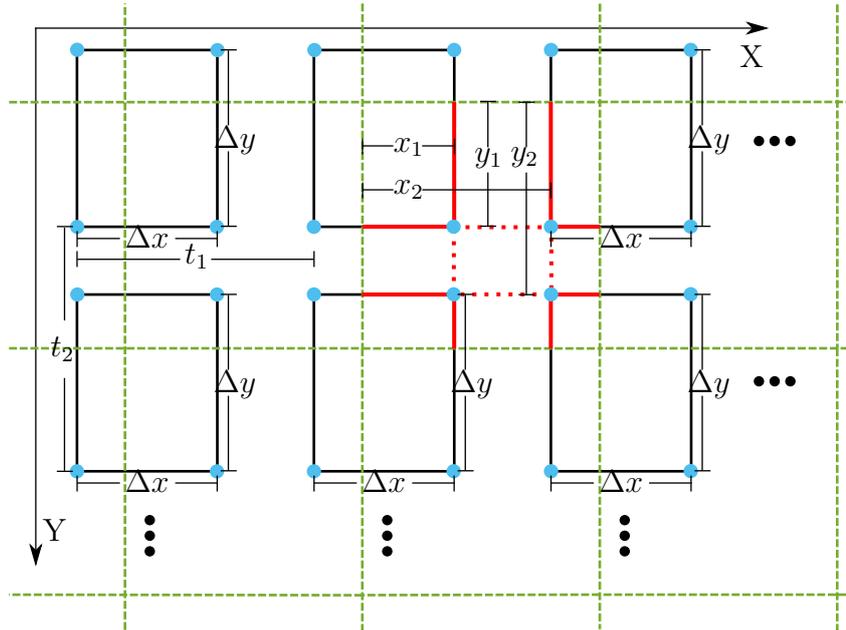
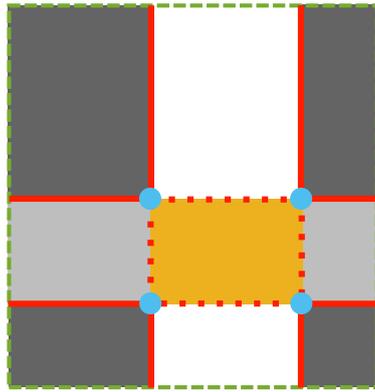


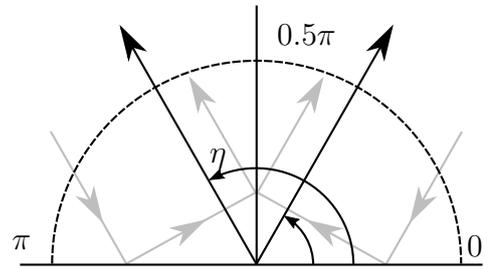
Figure 4.10: Applied window model. Every window has the same height Δy and width Δx . The spacing between identical points of windows, t_1 and t_2 , respectively, is already known from previous processing steps. Independent of the actual segmentation of the facade (shown as green dashed lines), four dominant lines (marked in red) are present in each patch. The unknown parameters are the distances (x_1, x_2, y_1, y_2) from the patch borders to these dominant lines.

window frame form trihedral corner reflectors causing the PSs. These results are picked up again and continued in the form of defining window corners as the matching partners for PSs. Since optical oblique images carry strong evidence for the geometrical position of such window corners, the goal of this processing step is to explicitly model these window corners. Considering the sensing geometry of the SAR sensor allows one to select the correct window corner out of four possibilities per window.

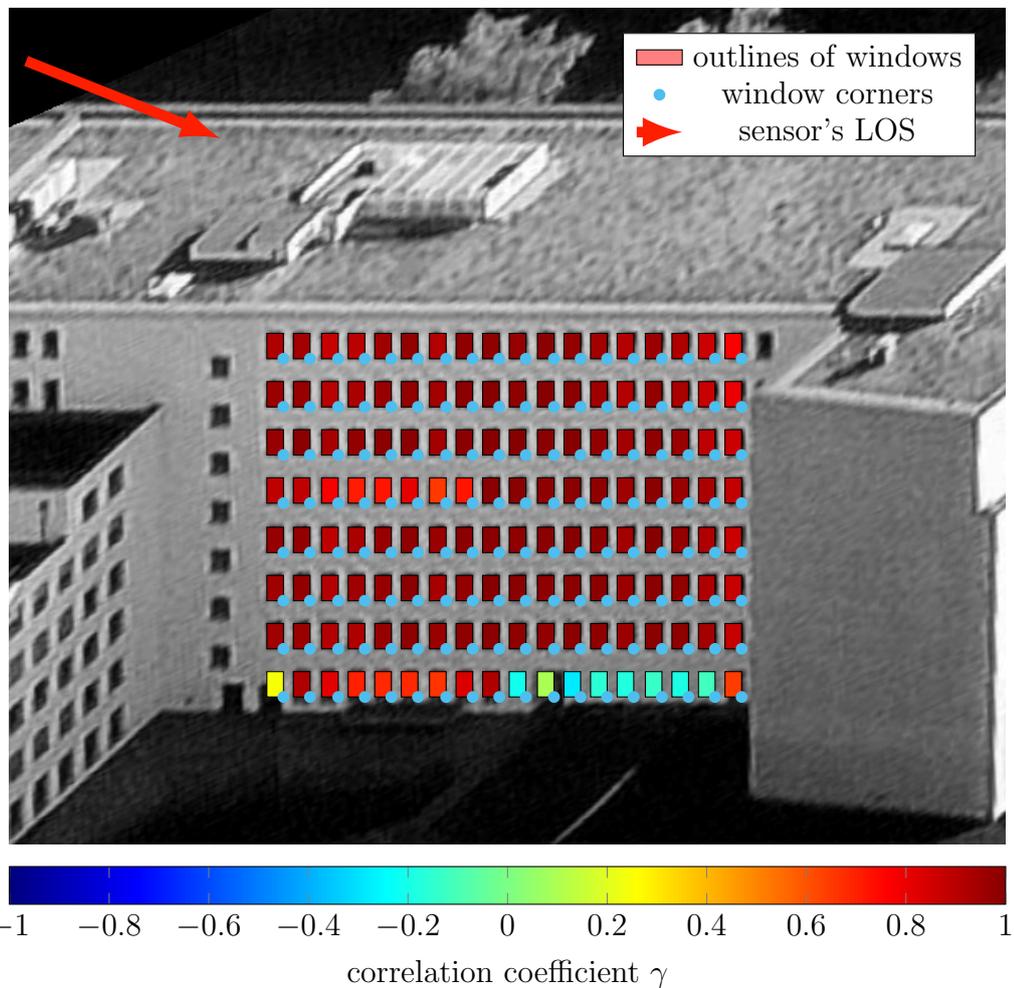
The model describing the window geometry is shown in Figure 4.10. The aim is to extract the distances (x_1, x_2, y_1, y_2) of the window outlines from the patch borders and infer from this the window corners (blue points). Knowing the LOS of the SAR satellite from the meta data lets one determine the window corner from the four candidates which is most likely to induce the PSs. This of course, is only the case under the assumption that the window corners themselves and no reflection mechanism inside the building or at regular facade structures other than the window corners are inducing the PSs. Intermediate results of the processing to extract regular lattices in optical data as described in Section 4.4 are used here. Working on rectified images is continued due to the inherent simplicity. Also, the edge image is used again since the process of fitting an explicit window model into the



(a) Four possibilities for window fragmentation inside single patch. Equally colored areas belong to the same window proposal.



(b) 2D scattering mechanism at window corner. See text for details.



(c) Result of window extraction. Outlines of windows are shown as colored rectangles. The color encodes the correlation coefficient of the corresponding cell. The window corners selected as possible matching partners are shown in blue. The red vector marks the sensor's LOS.

Figure 4.11: Sketch and result of window fitting procedure.

data underlies the Assumption 2 that a rectangular structure is discernible from the overall facade surface.

The segmentation of the facade into same size patches is continued as described in Section 4.4.2. This segmentation does not necessarily have to coincide with the desired segmentation, which would contain one complete window inside each patch. The key idea to become independent of this arbitrary segmentation and correctly detect the four window corners is to search for the most prominent lines in horizontal and vertical directions in the edge image and consider their intersections as window corners. In a subsequent step, the window width and height can be inferred from comparing four possibilities. Figure 4.10 sketches the procedure. The green dashed lines show the segmentation result. The position of the segmentation grid on the facade is arbitrary, but each patch always contains exactly four dominant lines as long as a single window is discernible from the rest of the facade in the edge image. The position of these dominant lines in each segmentation patch is the same and described by x_1 , x_2 , y_1 , and y_2 relative to the patch borders. Due to the rectification, dominant lines can be searched for independently in X and Y direction. Considering a patch size of $N \times M$ pixels, there are $\frac{N+(N-d)^2+M+(M-d)^2}{2} - 3d$ combinations, where d is the minimum distance between the two dominant lines. This gap between the dominant lines is necessary due to the fact that the edge image of the window is often not perfectly rectangular. Empirically, a gap of $d = 3$ has been found to be sufficient. Considering the characteristics of the used oblique imagery in this thesis and the typical facade extent with respect to the overall image size, a patch size of at most $N \times M = 100 \times 100$ pixel occurred in the area under investigation. This means that less than 9,500 combinations have to be checked. The score β_i which evaluates the combination i is simply defined as the number of edge pixels covered by the present dominant line hypotheses $\{x_1, x_2, y_1, y_2\}$, or more precisely

$$\beta_i = |\mathbf{E} \cap \{x_1 \cup x_2 \cup y_1 \cup y_2\}|, \quad (4.7)$$

where \mathbf{E} is the edge image and $|\cdot|$ counts the number of pixels which are covered by the lines defined by $\{x_1, x_2, y_1, y_2\}$. Since this scoring function is a basic logical operation and the search space is limited, an exhaustive search over all combinations can be performed and no approximate heuristic is necessary.

Having found the most dominant lines, their intersections are assumed to be the window corners. Due to the randomness of the segmenting lines (see green dashed lines in Figure 4.10 which do not contain a complete window) the assignment of the intersections to the

semantic window corner, i.e. top left, top right, bottom left, and bottom right, is fourfold ambiguous. This is sketched in Figure 4.11a which shows an enlarged part of Figure 4.10. The four possible window areas in one patch are marked with equally filled rectangles. The previously extracted dominant lines are marked in red and their intersections as blue points. To determine the correct solution, all four possibilities are scored by counting the edge pixels which are covered by the rectangle border lines and dividing this number by the length of the border lines. This method is motivated by the assumption that only edge pixels from the window borders are present in the edge image.

Finally, the sensor's LOS can be projected into the rectified image in order to select the window corner which induces the PSs. As the ray has to hit the interior faces of the trihedral corner which is formed by the window sill, the wall and the frame of the window, only one of the two bottom corners can induce the PSs. Figure 4.11b shows this fact schematically for a reflection in the two-dimensional plane. The y axis separates two corners. The gray arrows symbolize incoming rays and their reflections on the window sill (x axis) and the side wall (y axis) of the corresponding window. All rays entering the corner in an angle between $0 \leq \eta \leq 0.5\pi$ can only be reflected from the lower left corner, while all rays entering in an angle greater than 0.5π can only be reflected from the lower right corner. η can directly be inferred from the sensing geometry projected into the rectified image and the PS inducing corner selected accordingly. At this point, fivefold scattering mechanisms which potentially include threefold reflections in the upper corners are neglected. Figure 4.11c shows an example result with the extracted windows, the sensor's LOS projected into the image, and the assumed correct window corners.

4.5. Iterative graph-based bipartite matching

In this section, a method for establishing a one-to-one assignment between PSs and window corners extracted from the optical image is derived. The previously described grouping results of PS point clouds as well as extracted window corners act as input data. As outlined above, the direct geocoding of PSs into object coordinates is error-prone, mostly due to the inaccurate elevation coordinates of the PSs which is a random factor for each individual scatterer. This was already mitigated by projecting the PSs onto a common plane. Another source of uncertainty for the locations of PSs in the oblique image is the orientation of the camera, as described in Section 2.3.2. This part is image-related and constant for all PSs. In order to mitigate this inaccuracy contribution, the transformation which maps the PSs onto the corresponding window corners can be estimated. This is motivated by assuming that the

window corners induce PSs and that their position in the image can correctly be determined. If the assignment between the regular PSs and the corresponding window corners is correctly established, this can be used to correct the possibly inaccurate interior and exterior camera orientation.

To further make the matching more robust against such uncertainties and to ensure topologic consistency, the regular structure of the facades under investigation is exploited. This can be done because the regular alignment of facade elements like windows is preserved in both mappings of the real world, the SAR acquisitions and the optical oblique image, respectively. Thus, the key idea is to use this prior knowledge about the regularity and exploit the fact that the neighboring properties, i.e. the topology of facade objects, are preserved in both imaging geometries. According to Assumption 4, the inaccuracy budget of the projection of a single PS into an oblique image is composed of a systematic (image-related) and a random (PS-related) contribution. Since the systematic part is the same for all PSs and the topology is preserved, the matching should show this systematic behavior as well. A solely geometry based matching solution which is affected by individual random inaccuracies originating from the SAR domain may lead to an incorrect result. Another argument for incorporating topology information is that some PSs may not fall on a lattice position after the projection onto the facade plane. This may be due to erroneous range or azimuth coordinates. In such cases, the topology information can help to decide which matching partner is the correct one. An example for such a situation is shown in Figure 4.12. Blue triangles mark the window corners which act as possible matching partners for the PSs in green. All PSs roughly lie on a lattice which leads to rather constant matching vectors. Figure 4.12a shows a solution where only the geometry is considered. The closest matching partner for the scatterer in the center of the red rectangle is the window corner to the left of the PSs. This result is in disagreement with all other matching vectors. A rather constant behavior of all matching vectors is expected, since the projection inaccuracy is systematic for all PSs and the largest remaining inaccuracy source. Incorporating this information leads to the solution in Figure 4.12b. In the following, the implementation of the joint use of geometry and topology information is derived.

A solution of the matching process is defined as an optimal bipartite matching between the set of PSs and the set of window corners. *Optimal* here means that a certain score evaluating the matching is minimized. Defining this score is the crucial part where the tradeoff between geometry and topology plays an important role. More precisely, this target function is derived as follows. First, the geometrical distance between all $|\mathcal{M}|$ matched PSs

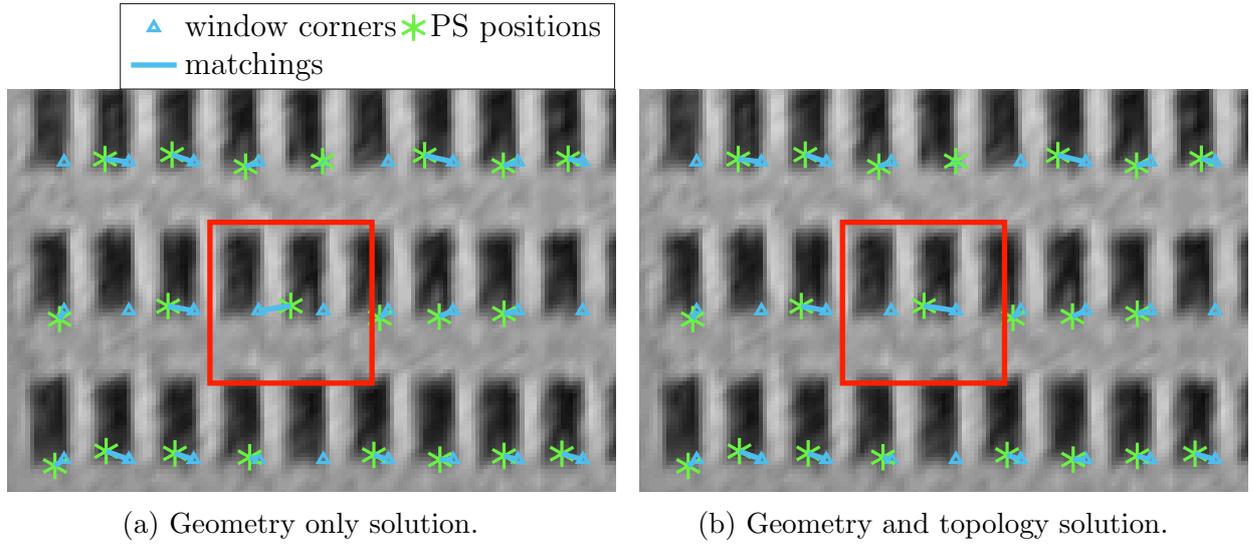


Figure 4.12: Example for the necessity of topology information. In green the projected PSs are shown. The extracted window corners are depicted as blue triangles. The PS inside the red rectangle changes its matching to a more consistent solution if topology information is introduced.

and optical lattice nodes should be small. This term is denoted by d_{geom} . Incorporating the anisotropic inaccuracy budget leads to the Mahalanobis distance

$$d_{geom} = \sum_m^{|\mathcal{M}|} \sqrt{(\vec{L}_m - \vec{P}_m)^\top \Sigma_{\mathbf{xy}}^{-1} (\vec{L}_m - \vec{P}_m)}, \quad (4.8)$$

where \vec{P}_m is the m th PS and \vec{L}_m the matched window corner, both in rectified image coordinates. The inaccuracy metric is captured in the covariance matrix $\Sigma_{\mathbf{xy}}$. Second, the topology of the PSs should be preserved when projected into the optical image. The term capturing this property is denoted by d_{topo} and its derivation is given in Section 4.5.1. Both terms can be adjusted as to how much they contribute to the overall solution. This leads to the general form

$$d^* = \alpha d_{geom} + (1 - \alpha) d_{topo}, \quad (4.9)$$

where α is defined between 0 and 1 and controls the relative contribution.

A flowchart of the strategy derived in this section is shown in Figure 4.13. At this stage, only regular PSs are considered. The aim is to find the optimal correspondence of each of them in the optical image. As outlined above, the systematic inaccuracy contribution originating from the camera orientation should be estimated. The transformation which

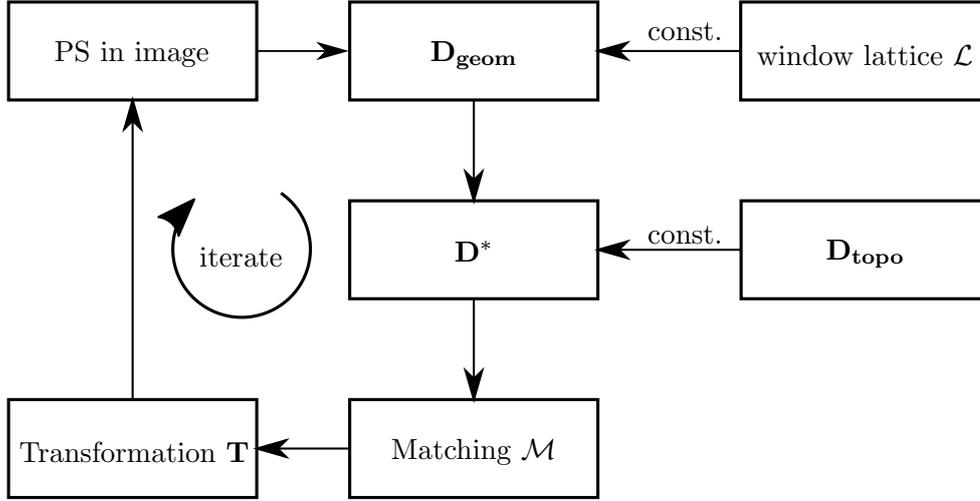


Figure 4.13: Flow chart of the iterative graph-based bipartite matching procedure. The PSs in image coordinates as well as the topography information \mathbf{D}_{topo} are given and are constant during all iterations. The distance matrix \mathbf{D}^* is composed of the geometric and topographic distance matrices. \mathbf{D}^* is subject to standard bipartite matching approach which leads to changed image coordinates of the lattice nodes.

geometrically maps the PSs onto their corresponding window corners depends on the actual matching. Applying this transformation to the PSs coordinates in the image domain results in a potentially changed optimal matching compared to the untransformed PSs. Thus, the transformation and the matching are mutually dependent. This is the reason why the whole procedure is solved iteratively.

The positions of window corners in the image as well as the topologic information from the SAR grouping is constant over all iterations. Given a set of window corners in the rectified image and initial PS positions, the geometrical score d_{geom} for all combinations of the set of lattice nodes \mathcal{L} and all PSs can be computed via the Mahalanobis distance (4.8) and is stored in the matrix \mathbf{D}_{geom} . The topology information is seen as a priori information and is considered as an offset which is added to the geometric score. It can also be computed for all combinations (see Section 4.5.1) and stated in a matrix form \mathbf{D}_{topo} . Combining both contributions results in the combined matrix

$$\mathbf{D}^* = \alpha \mathbf{D}_{\text{geom}} + (1 - \alpha) \mathbf{D}_{\text{topo}} . \quad (4.10)$$

On this matrix, standard bipartite matching is applied. Based on the resulting assignments, the optimal transformation considering the anisotropic inaccuracy budget is computed yielding

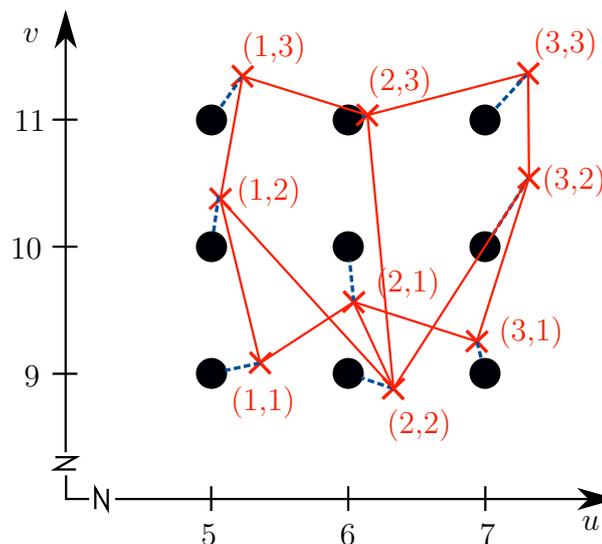


Figure 4.14: Example of an intermediate matching result. Black dots mark window corners which were derived from the rectified optical oblique image. They can be described by their (u, v) indices. Red crosses symbolize PSs projected into the optical image. Their indices (a, b) which are the result of the grouping procedure in the SAR domain are given in red. The matching result of this iteration is shown as blue dashed lines. A contradiction in the lower middle part is present.

updated PS positions. Then, the next iteration starts. In the following, the whole approach and its subproblems are derived in detail. First, the topology term is introduced.

4.5.1. Topology information

Following Observation 2.1 (page 26), the topology of facade objects is preserved in both sensing geometries. A key idea of the presented method is to exploit the topology in order to support the registration of PSs to window corners derived from the optical image. For this purpose, the grouping in the two-dimensional SAR domain is assumed to be correct in terms of being similar to that in the optical image domain. As a result of the grouping of the SAR data, every PS can be expressed as a lattice node. As stated in Equation (4.1), a lattice is defined by two spanning vectors \vec{t}_1 , \vec{t}_2 , and an offset vector \vec{t}_3 . A distinct lattice node is described by integer repetitions a and b of these spanning vectors. Thus, given these three vectors and neglecting random inaccuracies of the PS positions, all PSs can be described by their integer indices a and b .

The regular structure present in rectified optical images of facade buildings can analogously be described by two spacings, as described in Section 4.4. Again, single nodes of this regularity

can be represented as the integer repetition u, v of these spacings. Projecting all PSs into the optical image and matching it one-to-one to optical lattice nodes yields the following information: Every PS has an index (a, b) as well an index (u, v) of the matched lattice node. Figure 4.14 shows a simple example of such an intermediate matching result. The regular lattice of window corners derived from the optical image are marked with black points, while the projected PSs are red crosses. Their (a, b) indices are written directly next to the PSs. The grouping result of the PSs in the SAR domain is shown as red solid lines. A matching is symbolized as blue dashed lines. The topology is consistent except for the two PSs $(a = 2, b = 1, u = 6, v = 10)$ and $(a = 2, b = 2, u = 6, v = 9)$. Modifying the matching for these two PSs would harmonize the topology but increase the geometric distance of the matching.

The topology is stored in the following form. All elements of the matrix \mathbf{D}_{topo} with one column for every PS and one row for every lattice node are set as

$$\mathbf{D}_{\text{topo}}(i, j) = \sqrt{(a_i - u_j + O_X)^2 + (b_i - v_j + O_Y)^2}, \quad (4.11)$$

where O_X and O_Y are offsets which are necessary to establish the overall overlap of both indices and make the score independent of the absolute indices. D_{topo} explicitly stores the correct assignment of PSs to their corresponding partners from the lattice derived from optical imagery. Note that *correct* in this context means the topology information derived from the grouping in the two-dimensional SAR domain. The entry for the i th PS and j th lattice point is zero when $a_i = u_j + O_X$ and $b_i = v_j + O_Y$. The larger the deviation of the indices are, the larger the corresponding score is. In order to enable this kind of index comparisons, two important conditions have to be met. First, the offsets O_X and O_Y between the (a, b) and (u, v) indices have to be correctly computed. This means that in the case that no topology violation exists, $O_X = a - u$ and $O_Y = b - v$ is constant for every PS. In the case of topology violations at some nodes, deviations from these constants mark these particular wrongly assigned lattice nodes. To ensure that the offset of correctly assigned nodes can be estimated in the presence of wrongly assigned nodes, it is assumed that the majority of nodes are assigned correctly. The estimation itself is done by performing a majority voting scheme over the differences $O_X = a - u$ and $O_Y = b - v$. Due to the regular characteristic of the window lattice nodes L , the offsets simply shift the whole lattice. This can be interpreted as an ambiguity concerning the absolute position in the lattice. The sum of distances to the matched lattice nodes is not affected by this offset. Therefore, \mathbf{D}_{topo} is constant regardless of changes in offsets across iterations.

The second consideration is that the set of regular lattice nodes derived from optical imagery has to be at least as large as the amount of PSs in order to guarantee a maximal matching in terms of assigning one optical lattice node to every PS. This can be ensured by properly extending the optical lattice with pseudo nodes which have no support by the optical data but can be inferred. Depending on the deviation of the sensor's LOS between the SAR and optical acquisitions, some parts of the facades could be occluded in one data set but not in the other.

4.5.2. Optimization scheme

The problem at hand is to tune the influence of two measures which are mutually dependent: The geometry as well as the topology information may favor different matchings. Neglecting the topology information at all leads to a bipartite matching which minimizes the geometric distance between the window corners and the PSs projected into the image. On the other hand, neglecting the geometry information and just considering the topology leads to a solution in accordance with the topology derived from the SAR grouping step but has no support from the optical image. Putting these considerations into analytical form leads to the following optimization task.

$$\begin{aligned} \arg \min_{\mathcal{M}, \mathbf{T}} \sum_m^{\mathcal{M}(\mathbf{D}^*)} \sqrt{(\vec{L}_m - \mathbf{T}\vec{P}_m)^\top \boldsymbol{\Sigma}_{\mathbf{xy}}^{-1} (\vec{L}_m - \mathbf{T}\vec{P}_m)} \\ \text{w.r.t. } \mathbf{D}^* = \alpha \mathbf{D}_{\text{geom}}(\mathbf{T}) + (1 - \alpha) \mathbf{D}_{\text{topo}} , \end{aligned} \quad (4.12)$$

where $(\vec{L}_m - \mathbf{T}\vec{P}_m)$ denotes the difference vector in image coordinates between a geometrically transformed PS $\mathbf{T}\vec{P}_m$ and its matched lattice node \vec{L}_m . The covariance matrix $\boldsymbol{\Sigma}_{\mathbf{xy}}$ contains the anisotropic inaccuracy budget mainly resulting from the elevation direction of the SAR sensing system and the exterior orientation of the camera. The matching \mathcal{M} is a function of the matrix \mathbf{D}^* which in turn is dependent on the positions of the PSs in the image. The latter are updated in every iteration by applying the transformation \mathbf{T} . Since all regular PSs are assumed to be situated on the same plane and are mapped into the image plane, a homography is used for \mathbf{T} . For the following deviation, however, the specific type of \mathbf{T} is irrelevant as long as the transformation can be expressed as a matrix multiplication. The contribution of the topology with respect to the geometry is controlled by the parameter α . Assuming for the moment that the *correct* tradeoff between both contributions is known, i.e. that α in Equation (4.12) is given, how can a bipartite matching be found? Two additional

important questions at hand at this point are: Does the solution always converge? And: Is the solution optimal? The answers can be found by assembling the following parts:

Lemma 4.2 (Optimal shift). Given a bipartite matching \mathcal{M} between two point sets \mathcal{L} and \mathcal{P} in a Mahalanobis distance metric, applying a geometrical transformation which optimally maps \mathcal{P} onto \mathcal{L} minimizes the summed Mahalanobis distances over all matches in \mathcal{M} .

Proof. The sum of $|\mathcal{M}|$ Mahalanobis distances is

$$f : \sum_{m=1}^{|\mathcal{M}|} \sqrt{(\vec{L}_m - \mathbf{T}\vec{P}_m)^\top \Sigma_{\mathbf{xy}}^{-1} (\vec{L}_m - \mathbf{T}\vec{P}_m)}, \quad (4.13)$$

where \mathbf{T} is the same transform applied to all $|\mathcal{M}|$ points of the set \mathcal{P} . Let the sets \mathcal{L} and \mathcal{P} of image coordinates be given in homogenous coordinates, then \mathbf{T} can be any geometrical transformation that can be expressed by a 3×3 matrix. The particular transformation which minimizes the sum of Mahalanobis distances can be found by computing the derivative of f with respect to \mathbf{T} and setting it to zero. Since the square root in Equation (4.13) increases strictly monotonically, it can be ignored for finding the minimum. Thus,

$$\frac{\partial f}{\partial \mathbf{T}} : -2 \sum_{m=1}^{|\mathcal{M}|} \vec{P}_m^\top \Sigma_{\mathbf{xy}}^{-1} (\vec{L}_m - \mathbf{T}\vec{P}_m) \stackrel{!}{=} 0. \quad (4.14)$$

It can directly be seen that if a transformation existed which exactly maps all \vec{P}_m to \vec{L}_m the term inside the brackets vanishes and thus, gives a perfect solution for (4.14). Since there are typically more data points available than necessary to determine the parameters of the transformation \mathbf{T} , an approximate solution is needed. In the simplified case that \mathbf{T} comprises only a translation,

$$\mathbf{T}^{\text{translation}} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix},$$

the derivative can be expressed as

$$\frac{\partial f}{\partial \mathbf{T}} : -2 \sum_{m=1}^{|\mathcal{M}|} \overrightarrow{P}_m^\top \Sigma_{\mathbf{xy}}^{-1} \left(\overrightarrow{L}_m - (\overrightarrow{P}_m + \overrightarrow{t}) \right) \stackrel{!}{=} 0 \quad (4.15)$$

$$\Rightarrow \overrightarrow{t} = \frac{\sum_{m=1}^{|\mathcal{M}|} 2\overrightarrow{P}_m^\top \Sigma_{\mathbf{xy}}^{-1} (\overrightarrow{L}_m - \overrightarrow{P}_m)}{\sum_{m=1}^{|\mathcal{M}|} 2\overrightarrow{P}_m^\top \Sigma_{\mathbf{xy}}^{-1}} = \frac{1}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} \overrightarrow{L}_m - \overrightarrow{P}_m. \quad (4.16)$$

This means that averaging the difference vectors of matched points is the optimal solution, independently of the Mahalanobis distance. For more complex transformations like affine or projective mappings between \mathcal{P} and \mathcal{L} , the optimal transformation can be found by minimizing the quadratic residual of $\sum_{m=1}^{|\mathcal{M}|} \sqrt{(\overrightarrow{L}_m - \mathbf{T}\overrightarrow{P}_m)^\top \Sigma_{\mathbf{xy}}^{-1} (\overrightarrow{L}_m - \mathbf{T}\overrightarrow{P}_m)}$. This L_2 solution can be found by least squares adjustment, for instance, and is guaranteed to yield the optimal solution which minimizes the sum of quadratic residuals. \square

Another important point in the line of arguments is the optimality of the bipartite matching as described in Section 2.5. Thus, given a bipartite graph represented as its adjacency matrix, the Hungarian algorithm delivers the optimal assignment in terms of minimum costs. Holding on to this optimal matching and applying the transformation to the PS positions in the image minimizes the Mahalanobis distance according to Lemma 4.2. Recalculating the adjacency matrix again leads to smaller matching costs, or, more formally expressed:

Lemma 4.3 (Updated matching). An optimal matching \mathcal{M}_1 with matching costs c_1 in a bipartite graph $\mathcal{G} = (\{\mathcal{V}_1, \mathcal{V}_2\}, \mathcal{E})$ is given. The transformation which minimizes the summed distances is applied on one set of the bipartite graph, i.e. \mathcal{V}_1 . Computing again an optimal matching will always yield a matching \mathcal{M}_2 with matching costs c_2 in a way that $c_2 \leq c_1$.

Proof. Let \mathbf{T} be the transformation according to Lemma 4.2. Applying this transformation on \mathcal{V}_1 but holding on to the matching \mathcal{M}_1 leads to matching costs which are at most as high as c_1 since the sum of distances of the transformed matching vectors are smaller and the number of matched nodes is the same. Performing the optimal matching on the shifted nodes will always yield smaller or equally high matching costs than c_1 , since otherwise the matching would not be optimal. \square

Putting these parts together leads to the formulation of Algorithm 3, named iterative graph-based bipartite matching (IGBM). The key idea is that due to the mutual dependence

Data: Image coordinates of PSs; topology matrix \mathbf{D}_{topo}
Result: Assignment \mathcal{M} of PSs to lattice nodes; transformation \mathbf{T}
Step 0:
Initialize transformation $\mathbf{T} = \mathbf{I}$;
Initialize matching \mathcal{M} ;
while *solution has not converged* **do**
 Step 1:
 Apply geometric transformation: $\mathcal{P} \leftarrow \mathbf{T}(\mathcal{P})$;
 Step 2:
 $\mathbf{D}^* \leftarrow \alpha \mathbf{D}_{\text{geom}}(\mathcal{P}) + (1 - \alpha) \mathbf{D}_{\text{topo}}$;
 Step 3:
 $\mathcal{M} \leftarrow \text{Hungarian algorithm}(\mathbf{D}^*)$;
 Step 4:
 $\mathbf{T} \leftarrow \text{compute optimal geometric transformation}(\mathcal{M})$;
end

Algorithm 3: Iterative Graph-based Bipartite Matching.

of both parameters, the assignment \mathcal{M} and the transformed PSs positions $\mathbf{T}(\mathcal{P})$, respectively, an iterative procedure delivers a stable solution by alternately fixing one parameter and optimizing the other. This is inspired by the well known technique *expectation maximization* [Dempster et al., 1977]. The IGBM algorithm consists of four iteratively repeated steps and is initialized by setting the transformation \mathbf{T} to the identity matrix. The following steps are repeated until no change in the target function (4.12) occurs anymore. The positions of the PSs in the image are transformed according to \mathbf{T} . The new positions of iteration i are computed as

$$\mathcal{P}^i = \mathbf{T} \cdot \mathcal{P}^{i-1}, \quad (4.17)$$

where \mathcal{P}^i is the set of current positions of all PSs in the image. In step 2 the matrix \mathbf{D}^* is constructed. This is the point where the information from geometry and topology is merged. The geometry term \mathbf{D}_{geom} purely captures the metric distance in image coordinates with regard to the anisotropic inaccuracy budget of the projected PSs, represented as their covariance information Σ_{xy} . The n^{th} column of \mathbf{D}_{geom} is computed as

$$\mathbf{D}_{\text{geom}}^{(n,\cdot)} = \underset{|\mathcal{P}| \times 1}{\mathcal{L}} \cdot \underset{|\mathcal{P}| \times 2}{\Sigma_{\text{xy}}^{-1}} \cdot \underset{2 \times 2}{\vec{P}_n}, \quad (4.18)$$

where $|\mathcal{P}|$ is the number of PSs, \mathcal{L} is the set of all window corners, and \vec{P}_n the image coordinates of the n^{th} PS. Thus, \mathbf{D}_{geom} contains the quadratic Mahalanobis distances between all PSs to all lattice nodes, incorporating the covariance matrix as derived in

Section 4.3.2. The topology information is stored in a matrix \mathbf{D}_{topo} of the same size as \mathbf{D}_{geom} , as described in Section 4.5.1. $\mathcal{M}(\mathbf{D}^*)$ yields the optimal maximum bipartite matching for each iteration. The corresponding matching costs are denoted as $\text{cost}(\mathcal{M}(\mathbf{D}^*))$.

Finally, a statement about this algorithm considering the optimality and convergence can be made:

Theorem 4.4. For a given α , the iterative graph-based bipartite matching Algorithm (3) always converges and returns a local minimum of (4.12).

Proof. The proof is given by combining the aforementioned lemmata. Step 1 applies the transformation which minimizes the sum of distances for the present matching according to 4.2. Since \mathbf{D}_{topo} is constant, due to the regular characteristic of the window lattice, \mathbf{D}^* is always smaller or equally as high as in the previous iteration. From this fact it follows that the optimal matching, which is guaranteed by the Hungarian algorithm, has smaller or equally as high matching costs which again can be minimized by applying the transformation of the present iteration. Therefore, the algorithm always converges to a local minimum. \square

The resulting minimum depends on the initial PS positions as well as on α . The best guess for the first are the planned PSs where the accuracy in elevation direction is improved by the factor of $\frac{1}{\sqrt{N}}$ for N regular PSs. Since the exterior orientation is the result of a bundle adjustment, it is assumed to be unbiased. Therefore, the initial PS positions are given by directly projecting the plane projected PSs into the image. The *correct* parameter α is hard to determine since no ground truth data is available. Therefore, the stability of the IGBM algorithm is investigated in Section 5.6.

4.6. Improving the accuracy of PS in optical images

The IGBM algorithm yields the optimal matching between a set of PSs and a set of window corners extracted from optical oblique imagery. One application of this method is the assignment of individual PSs to building structures which are then subject to monitoring tasks. Another motivation is to learn something about the physical nature of PSs. Therefore the assignment is exploited in order to improve the accuracy of the PS position in the image. In a subsequent step, the image information in the direct vicinity of the corrected PS image coordinate can be investigated. As outlined before, the matching is used to estimate the homography between regular PSs and the corresponding window corners. If the assumption is correct that the regular PSs at the facade are all induced by the same regular structure

PS type	influence on inaccuracy	upper bound (without any object knowledge)	object knowledge	lower bound (all assumptions correct)
regular	increasing	range, azimuth, elevation, camera orientation	range, azimuth, camera orientation	transformation estimation
	decreasing	-	plane projection	plane projection, transformation estimation
irregular	increasing	range, azimuth, elevation, camera orientation	range, azimuth, camera orientation	range, azimuth, transformation estimation
	decreasing	-	plane projection	plane projection, transformation estimation
non-facade	increasing	range, azimuth, elevation, camera orientation	range, azimuth, elevation, camera orientation	range, azimuth, elevation, camera orientation
	decreasing	-	-	-

Table 4.1: Comparison of inaccuracy budgets with and without object knowledge.

which are correctly extracted, then the resulting homography can be used to reduce the inaccuracy contribution of the camera orientation. Of course, this view of the problem only holds under ideal conditions. Therefore, this ideal case is assumed as a lower bound for the accuracy of a PS after applying IGBM and the resulting homography. Table 4.1 lists three different scenarios for all three types of PSs as classified in Section 4.3.2. Regular PSs are those lying in the facade plane and are regularly aligned, irregular scatterers also lie in the plane but are not members of the regular pattern and, lastly, PSs classified as not belonging to the facade neither lie in the plane nor are they regularly aligned in SAR coordinates.

The upper bound of the PS position accuracy in the image is determined by directly performing variance propagation from the SAR domain across object coordinates and the projection into the image as described in Section 2.3. This can be interpreted as neglecting any object knowledge and considering every PS individually. In this case, the inaccuracy is determined by the SAR coordinates and the camera orientation. If object knowledge is introduced in the form of assuming that regular PSs lie in a common plane, then the inaccuracy in elevation direction can be improved as presented in Section 4.3.2. This of course, only holds for the regular and irregular PS types. The lower bound is defined by

	Lindenstraße	Ebertstraße	Charlottenstraße
patch size horizontal [px]	16.25	33.75	45.75
patch size vertical [px]	38.25	46.00	38.52
$A_{95}^{\text{regular}}/A_{\text{patch}}$	0.28	0.04	0.01
$A_{95}^{\text{irregular}}/A_{\text{patch}}$	2.31	0.77	0.60
$A_{95}^{\text{no facade}}/A_{\text{patch}}$	3.78	1.25	0.92
$A_{95}^{\text{regular}} [m^2]$	0.55	0.28	0.11
$A_{95}^{\text{irregular}} [m^2]$	4.59	4.86	4.70
$A_{95}^{\text{no facade}} [m^2]$	7.51	7.89	7.20

Table 4.2: Resulting accuracies for three exemplary facades in relation to the patch size.

applying the homography on the matched PSs. As there are typically more PSs than needed to compute the projective transform describing the homography, the residuals can be used to empirically estimate a covariance matrix. Of course, this does not hold for PSs classified as not belonging to the facade since they do not lie in the plane and therefore cannot be improved by the application of the homography. Figure 4.15 shows an example of 95% confidence ellipses of the lower bounds as described in the rightmost column of Table 4.1. The centers of the confidence ellipses are the final positions of all PS types after applying the IGBM algorithm. In red, the unchanged inaccuracy budget for PSs not at the facade is shown. Irregular PSs are also projected to the facade plane and therefore the inaccuracy contribution of the elevation direction is reduced. Finally in green, the additional mitigation of the systematic error induced by the camera orientation by estimating the covariance matrix after applying the homography is depicted.

In order to express these confidence ellipses numerically, Table 4.2 states the area of 95% confidence ellipses in terms of m^2 as well as fractions of the area of the regular element at the facade. The latter number is facade dependent but gives an impression of how well the assignment of regular PSs to individual facade elements can be done. The numbers are given for three case studies which are introduced in the following chapter. The dependency on the individual facade becomes obvious. The facade denoted as Lindenstraße is the one depicted in the previous Figure 4.15. Many windows are situated very close to each other. On such facades the assignment of a PS is more challenging compared to a facade of the kind of Eberstraße (see Figure 5.13a on page 115) or Charlottenstraße (see Figure 5.5a on page 101). The absolute areas of the resulting confidence ellipses are given in the lower part of Table 4.2. Their values do not vary significantly, due to the independence of the facade layout. Incorporating the distances of consecutive facade elements leads to the relative values. A fraction below 1 between the area of the 95% confidence ellipse of the scatterer and the area

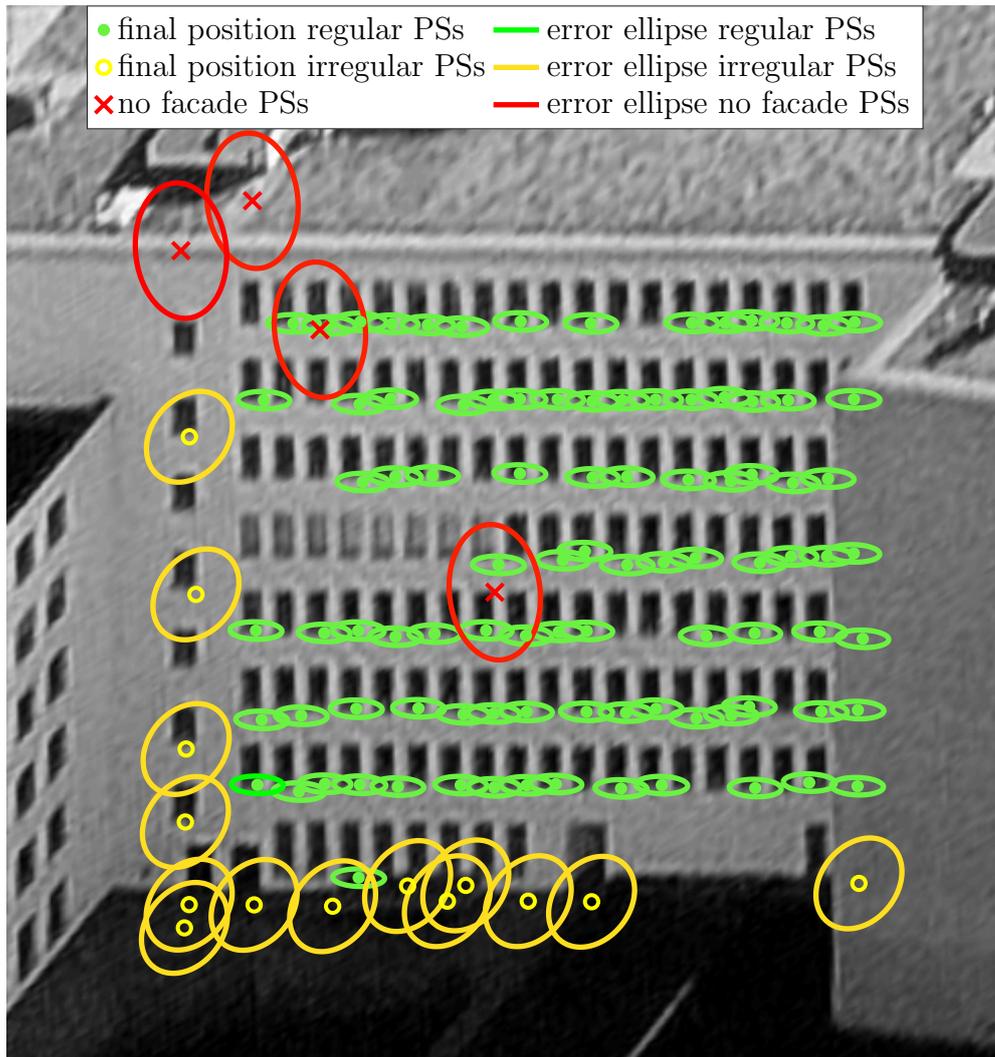


Figure 4.15: 95% confidence ellipses for three types of PS.

of the regular patch is interpreted as small enough to allow the identification of the correct matching partners. In the previous figure (Lindenstraße), this is only the case for the regular PSs and by assuming that the applied homography correctly captures the inaccuracy of the camera orientation. This numerical assessment also makes sense when comparing the corresponding confidence ellipses of the case study Eberstraße in Figure 5.13b on page 115. The confidence ellipses of the irregular PSs in yellow only overlap one of each window corner while slightly shifting the red ellipses would cause to have two identical window corners inside them. This score is meant to act as an indication of the feasibility of unambiguously assigning PSs to their counterparts in optical images.

5. Experiments

In this chapter, the following hypotheses introduced in the previous chapter are investigated.

1. As a very fundamental assumption and in order to establish the relation between SAR data and optical images a regular structure of facades is required.
2. Regular patterns in the data can be made explicit with the presented methods.
3. The presented approach is robust against distortions like occlusions, varying illuminations across the facade, steep aspect angles, and a small portion of regular PSs compared to the total number of scatterers.
4. The IGBM matching between PSs and their correspondences in optical imagery can be established for a variety of facade types. The facade type may vary in size and/or appearance.
5. The transformation yielded by IGBM for two or more facades depicted in the same image is similar since the image-related inaccuracy budget is the same for those facades.

The experiments are conducted on the basis of selected case studies. Those exemplary facades are accordingly selected from the data which is introduced in Section 5.1. First, the existence of regular patterns is verified empirically for a typical urban area in Section 5.2. The presented methods to make such patterns explicit and to extract topology information in PS point clouds as well as optical images are investigated in Section 5.3 and 5.4. Hypotheses about the robustness and applicability on small and large facades are examined. Furthermore, some characteristics and limitations are presented with the help of case studies. The IGBM algorithm is applied on four case studies regarding the aforementioned hypotheses in Section 5.5. The sensitivity of this algorithm with respect to the parameter α which controls the relative influences of geometry and topology on the matching result is investigated in Section 5.6. Finally, the extension to the overall study area is presented in Section 5.7.

	beam 42	beam 57
orbit direction	descending	ascending
mode	High Resolution Spotlight	
resolution	range: 0.59 m	azimuth: 1.10 m
incidence angle	36.09°	41.90°
heading angle	190.57°	350.34°
looking direction	right	
stack size	94	79
time span	Feb. 2008 - Feb. 2012	Feb. 2008 - Mar. 2012
number of PSs	109,521	110,121
covered area	3.8 km ²	

(a) Properties of PS data sets.

values for scene center	campaign 2006	campaign 2010
number of images	86	174
image size	2672 × 4008 px ²	3248 × 4872 px ²
avg. GSD	14.4 cm	12.3 cm
avg. incidence angle	49.95°	49.49°
avg. covered area per image	0.42 km ²	0.47 km ²
covered area	10.9 km ²	21.8 km ²
focal length	170 mm	
avg. flying alt.	1899.4 m	2047.4 m

(b) Properties of optical data.

Table 5.1: Properties of used data.

5.1. Data description

The presented case studies in this section are all situated in Berlin, Germany. Two PS point clouds processed by the PSI processor GENESIS at DLR [Adam et al., 2003] are available. In Table 5.1a, the parameters defining the SAR acquisitions and PSI processing are listed for both stacks. Every PS has a three-dimensional UTM height coordinate as well as the position in the range-azimuth plane. A more detailed description of the used data can be found in Schunert [2014].

The optical oblique imagery was provided by BLOM UK and comprises acquisitions from two flight campaigns. The first, denoted as campaign 2006 took place between October 2006 and April 2008, the latter, named campaign 2010 was acquired between September and October 2010. Table 5.1b lists the characteristics of the imagery. The resulting GSD is in the range of some 12 cm to 14 cm which allows for the identification of individual



Figure 5.1: Overview on available data in the city center of Berlin, Germany. The two PS point clouds are marked as green (ascending) and red (descending) points. The area covered by at least one oblique image is shaded. Map background: © OpenStreetMap contributors.

facade elements like windows and sometimes smaller objects like the actual window corner. Figure 5.1 gives an overview of the distribution of the two PS point clouds as well as their superposition with the oblique imagery. The shading marks the area which is covered by at least one optical oblique image. In cases where one facade is visible in more than one image, the one which depicts the interested area the closest to the scene center is selected.

5.2. Regular facades in urban area: case study Berlin

The presented approach is based on the assumption that facades in urban areas are characterized by regular and repetitive patterns of facade elements like windows or balconies. This expectation appears trivial due to everyday experience. However, to systematically

image orientation	irregular facades	regular facades	ratio regular to total
north	3,149	1,022	24.5%
east	2,017	775	27.8%
south	3,026	852	22.0%
west	2,084	642	23.6%
sum / avg.	10,276	3,291	24.3%

Table 5.2: Result of manually labeling facades.

exploit the present regularity, an investigation is useful in order to analyze their appearance in the oblique imagery. Since the task is to assign PSs to their correspondences in optical images, the regularity which is used in order to establish this matching has to be present in the imagery. Furthermore, previous investigations of Gernhardt et al. [2015] showed that the SAR sensing geometry strongly influences the fact whether PSs are induced by a facade or not. It is assumed that the optical image is more robust against such changes in viewing directions compared to SAR. Therefore, this investigation is only conducted for the optical images and the result is considered as an upper bound estimate for the of the applicability of the assignment approach.

Optical oblique images of the study area and its direct neighborhood were labeled manually. A total of 13,567 individual facades in 62 images from the 2010 campaign were inspected for regularity according to Definition 4.1 (page 52). Furthermore, the patterns were only considered to be regular if they were not interrupted by other structures like a row of balconies inside a lattice of windows because the presented approach to extract window corners from optical images is not capable of capturing such complex patterns.

Table 5.2 lists the labeling results depending on the viewing direction of the respective camera. Approximately 24% of all facades in the city center of Berlin show a regular pattern which can be exploited for the presented matching approach. Thus, applying the developed method one could expect an upper bound of approximately 24% of all facades to be successfully handled by the IGBM algorithm and its preprocessing steps.

5.3. Grouping of persistent scatterers

In order to derive topology information from a set of PSs, the grouping procedure introduced in Section 4.3 is applied to two case studies. The selected examples are chosen in a way that generic characteristics of the grouping approach can be shown.

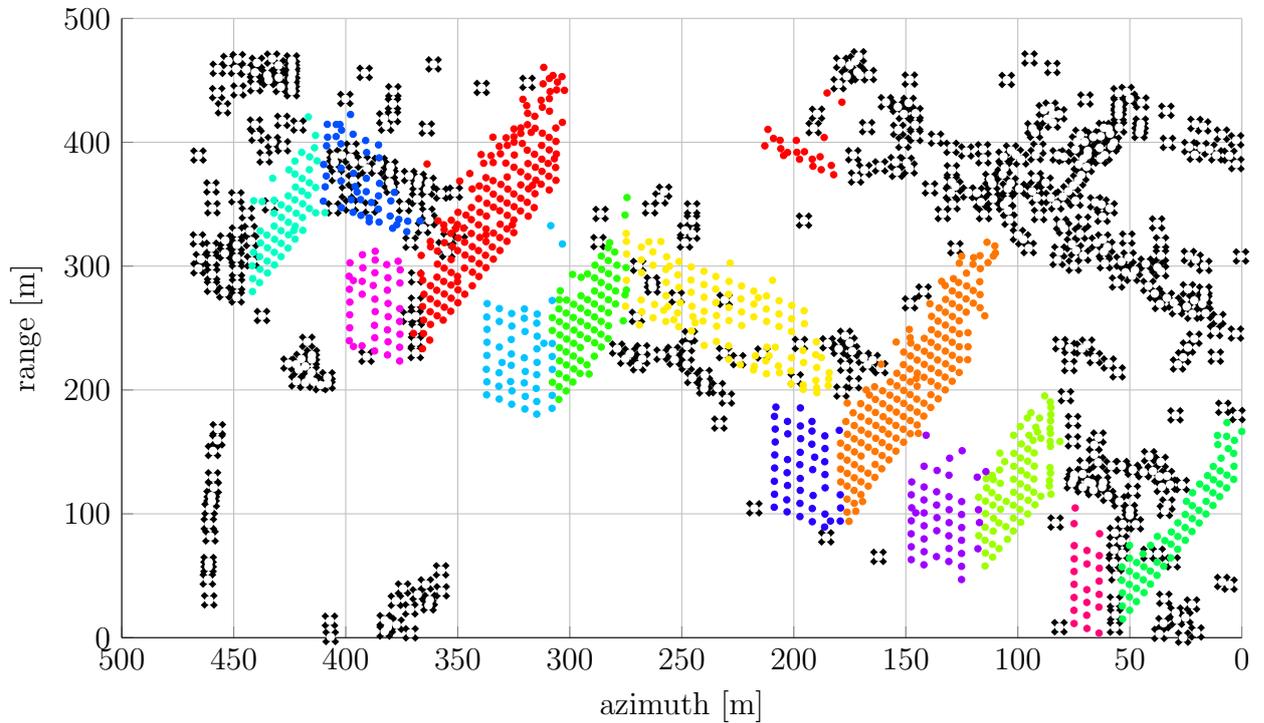
parameter	value
k (normal estimation, page 59)	20
d (RANSAC inlier bound, page 64)	0.5 m
threshold grouping (page 65)	0.5 m
γ (cross correlation threshold, page 73)	0.8
α (page 80)	0.75

Table 5.3: Applied parameters for all following experiments if not stated otherwise.

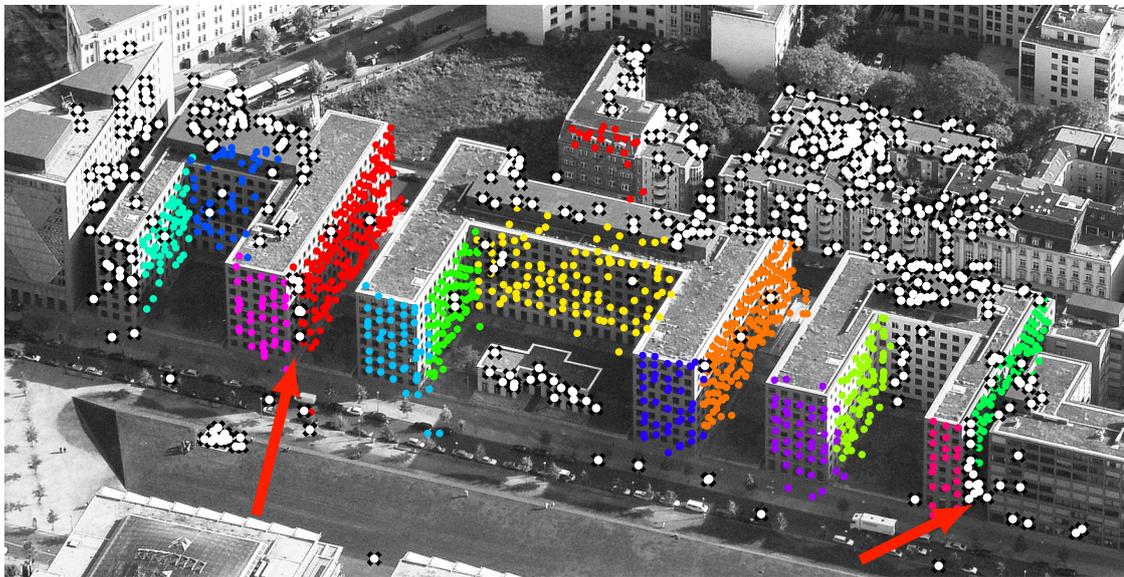
5.3.1. Grouping of PSs: Park Kolonnaden

The building complex Berliner Park Kolonnaden consists of three U-shaped buildings. The west and south facades of the flat roofed buildings are illuminated by the SAR sensor which passes over the scene in a right-looking ascending sensing geometry. This case study is chosen in order to show the applicability of the grouping approach to small (3×7) and widely extended (21×8) facades. Figure 5.2 shows the initial PS distribution at the facades of the complex. As a result of the procedure described in Section 4.3.1, the segmented point cloud is shown in two-dimensional SAR coordinates (Figure 5.2a) and projected into an oblique image of the scene (Figure 5.2b), respectively. The color represents the affiliation to the facade cluster. Since regularity was defined as a lattice with at least 3×3 elements in Section 4.1, all clusters with less than 9 members are discarded. Such PSs are marked with white dots on black crosses in both representations. The clustering of PSs into individual facades is meaningful except for some scatterers, marked with red arrows in Figure 5.2b, at the edge between two adjacent facades. In two cases, such PSs are discarded because their local normal does not coincide with the normal of neither of the adjacent facades. This is due to the fact that the normal is estimated by fitting a plane through the k nearest neighbors. For PSs at the edges, these k neighbors come from both building faces resulting in a false normal direction.

The grouping step is conducted for every facade separately but the results are depicted together in Figure 5.3. Note that the used parameters are all the same for every facade and are given in Table 5.3. If not stated otherwise, those parameters are chosen for all conducted experiments in this chapter. The colors coincide with the facade clustering results of the previous figure. The lattice topology as derived from the grouping step is visualized as edges between the adjacent PSs. Scatterers which are not part of the regularity are plotted as individual crosses, i.e. without an edge to any other PS. Before mapping into the oblique image, the object coordinates are improved by the projection onto the plane. Comparing the positions of PSs in Figure 5.2b with those in Figure 5.3 visually shows the



(a) Segmentation result in SAR coordinates. For a better orientation range and azimuth axes are interchanged with respect to usual illustrations.



(b) Segmentation result in image coordinates.

Figure 5.2: Building complex Park Kolonnaden. Segmentation result as starting point for the PS grouping procedure. Color encodes the affiliation of every PS to its facade cluster. Discarded PSs are marked with white dots on black crosses.



Figure 5.3: Result of the PS grouping approach. The color encodes the affiliation to individual facades. The projected PSs with the resulting grouping is shown. PSs which are not part of the regular pattern, i.e. irregular scatterers, are shown as crosses.

large improvement in terms of the regular alignment of scatterers. Without performing any further processing, an assignment between PSs to individual windows can be visually done. The grouping approach yields meaningful results for widespread facades like those in red, orange, or yellow, as well as smaller facades like the ones oriented towards the reader. Nevertheless, the grouping method failed at extracting the pattern for the facade marked with a red arrow. The plane estimation for this facade leads to a grossly incorrect result. This is due to the fact that the proportion of PSs belonging to the regular facade pattern is small with respect to the overall number of scatterers. This is a violation of Assumption 3 on page 55. Points at the structure on the roof distort the estimated plane in a way that the projected PSs fall on positions far away from the actual facade. This facade is not considered any further.

By visual impression one sees that a conformity between the grouped PSs and the pattern of windows in the image exists. This finding is confirmed by considering the projected confidence ellipses of the grouped PSs. Figure 5.4 compares the confidence ellipse of the projected regular PSs of the facade marked in yellow before and after the plane projection step which is described in Section 4.3.2. The accuracy information applied here is the same as introduced in Table 2.3 (page 31). The red ellipse marks the 95% confidence area of the PS position at the facade without making use of the assumption that all regular facade points lie in one plane, whereas the blue ellipse is smaller. The 95% confidence ellipse after

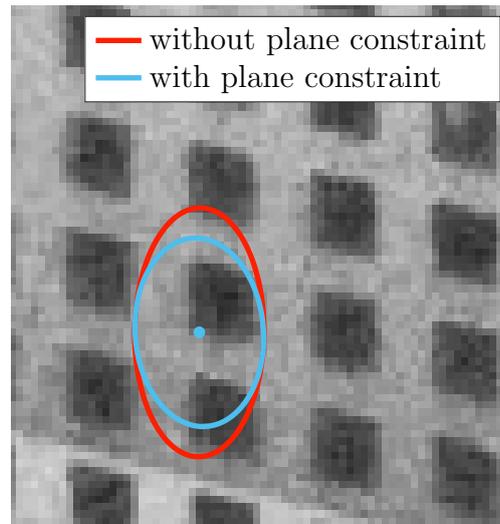


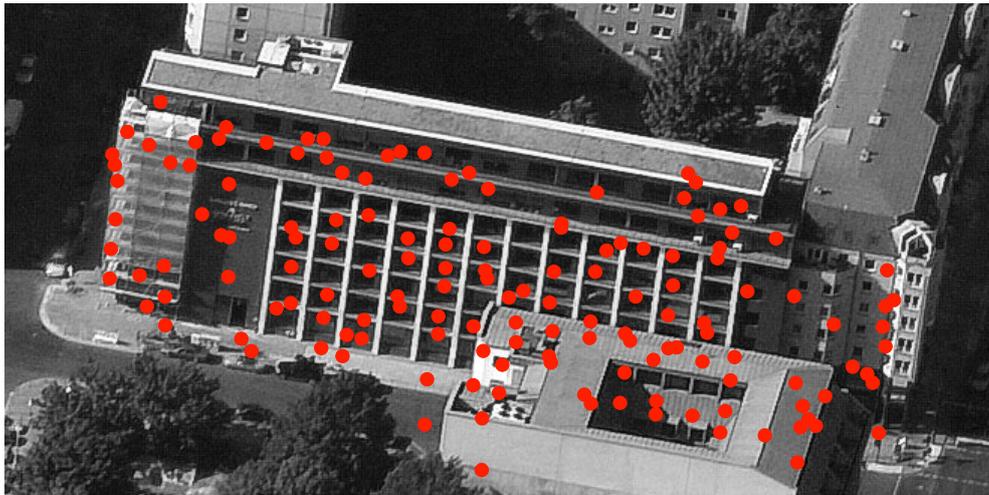
Figure 5.4: Effect of projection to plane on 95% confidence ellipse.

the projection onto the facade, which reduces the inaccuracy budget induced by the elevation direction, roughly covers one facade element. Without applying this prior knowledge the assignment of such a PS to the corresponding window corner is ambiguous since at least two window corners lie inside the red 95% confidence area. A further reduction of uncertainty is later achieved by exploiting the result matching between the PSs and extracted window corners from the optical image.

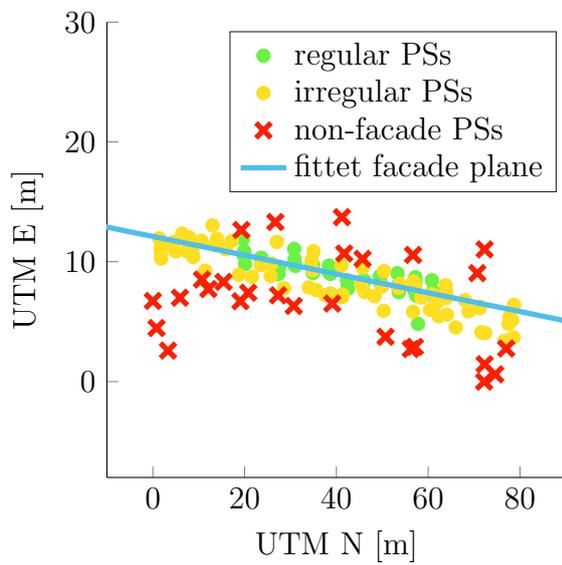
5.3.2. Grouping of PSs: Charlottenstraße

This case study shows the robustness of the PS grouping approach. The facade under investigation poses two challenges for the grouping process. First, the regular part only covers a portion of the whole facade, and second, there are several other PSs present in the facade plane which do not belong to the regular structure. These two properties complicate the identification of regular facade points. As described in Section 4.3, a PS has to lie in the facade plane and it must be possible to describe it as part of a lattice in order to be considered as belonging to the regular facade pattern.

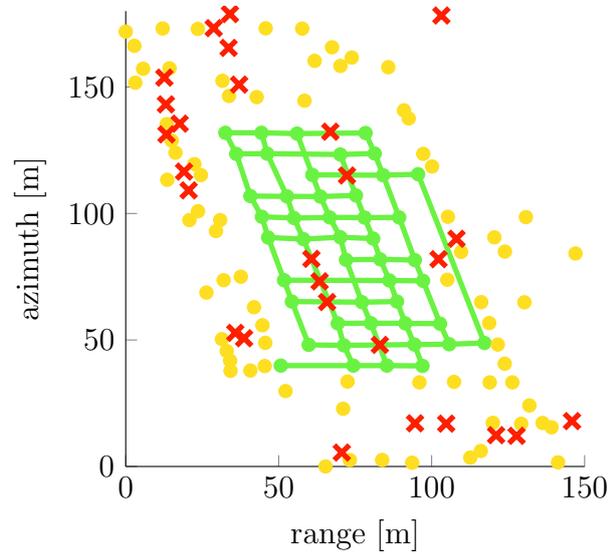
Figure 5.5a shows the initial set of 148 PSs projected into the oblique image. The lower right part of the facade is occluded by another building. On the right side some PSs belong to another building. Also, a scaffolding on the left side covers part of the facade. The regular part of the facade can be seen in the optical image but only rudimentarily in the PS distribution. Some patterns of vertically regular scatterers can be identified. The estimated facade plane is shown as a blue line in Figure 5.5b. In this top view of the scene all 148 PSs



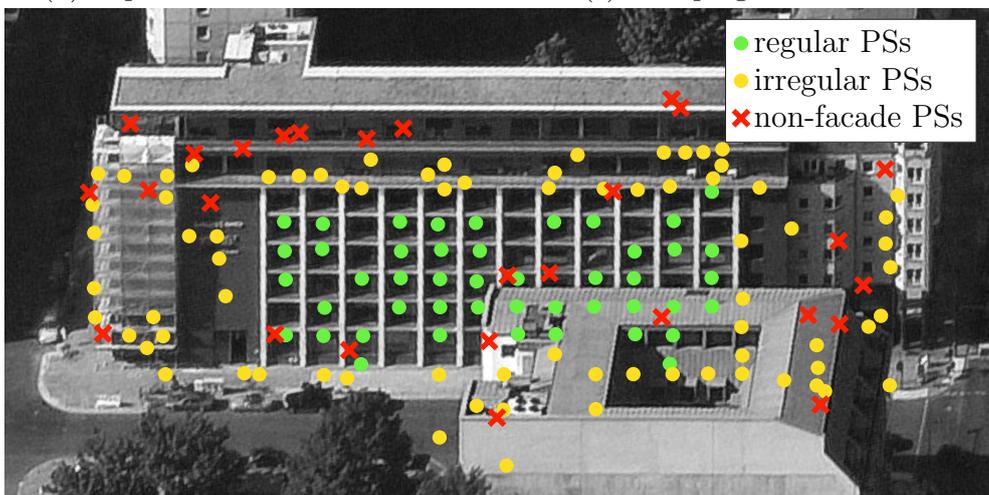
(a) Initial positions of all PSs.



(b) Top view on PSs.



(c) Grouping in SAR coordinates.



(d) Classification result in rectified image.

Figure 5.5: Facade plane estimation and grouping of PSs for the case study Charlottenstraße.

are marked with a symbol corresponding to their classification result as regular, irregular, or non-facade PSs (note that due to visualization purposes, the UTM E and UTM N axes are swapped with respect to the usual layout.). Both the regular PSs (shown in green) as well as irregular facade points (depicted in yellow) are RANSAC inliers. Those points were subject to the grouping procedure in SAR coordinates. The result is shown in Figure 5.5c where the pattern of those PSs belonging to the regularity is strikingly clear. The lattice topology of regular PSs is marked with green lines. An interesting observation can be made by comparing the PS distribution in geocoded coordinates in Figure 5.5b with the one SAR coordinates as seen in Figure 5.5c. In the latter, six PSs classified as not belonging to the facade due to not belonging to the set of RANSAC inliers fall on lattice node positions of the regular pattern. Their planimetric UTM positions, however, suggest that they are too far away from the facade plane. At these positions, the threshold for the deviation in elevation direction from the facade plane is chosen too small.

The resulting PS positions in the oblique image and also the classification result is shown in Figure 5.5d. The pattern of 46 PSs classified as regular facade points directly leaps out. 75 PSs also lie in the facade plane but are not regularly aligned and 27 scatterers originate from the slightly set back roof structures and other parts of the building. This case study showed that a regular pattern can robustly be extracted even when less than one third of all PSs under investigation belong to the regular pattern.

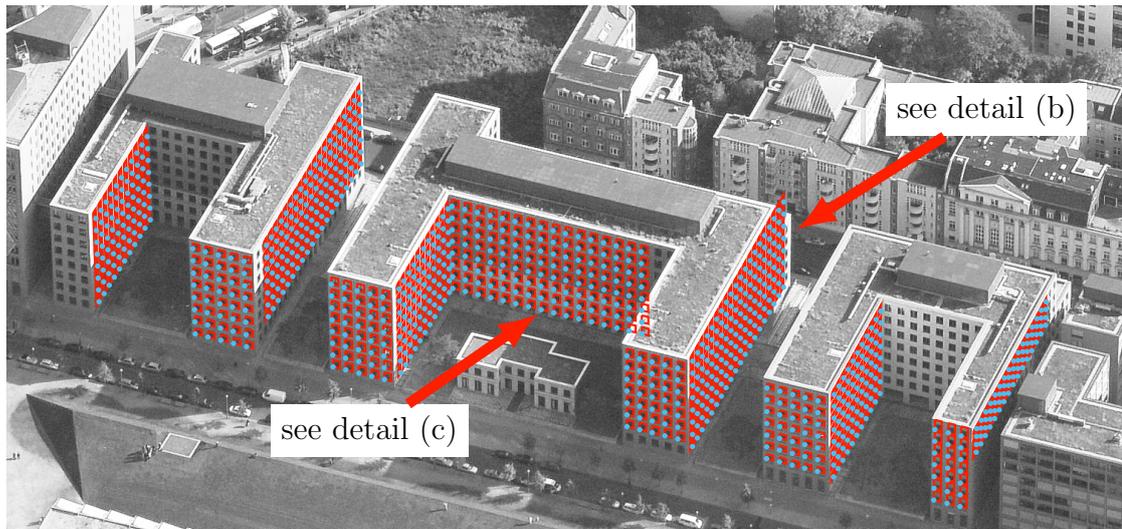
5.4. Window corner extraction

In this section, three case studies are presented which show the variety of facade types which can be successfully handled by the presented method for extracting window corners in oblique optical images as matching partners for PSs. Also, limitations and violations of the underlying assumptions are demonstrated. Two steps are performed. First, the detection of a pattern that coincides with the lattice information obtained in the SAR grouping step and, second, the extraction of the most probable repeated structure which induces the reflection leading to the PSs. The former is achieved by performing cross correlation which is very robust against changing contrasts and can be used due to very precise approximation values for the lattice position in the window as well as its spacings and extent. The latter is done by fitting a rectangular window model into the repeated element and projecting the sensor's LOS into the image which narrows down the possible matching candidates to one corner. The motivation that window corners often induce PSs was given in Section 3.2.

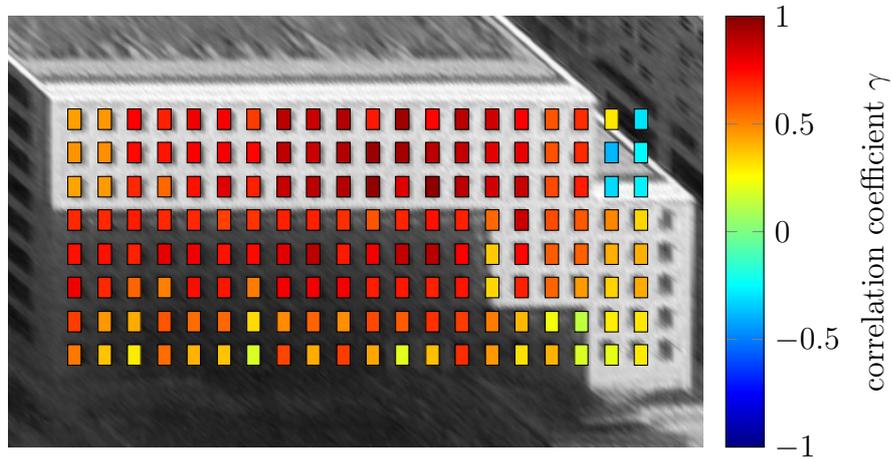
5.4.1. Window corner extraction: Park Kolonnaden

To continue with the previously introduced case study, the building complex Park Kolonnaden is further investigated, and the mentioned robustness of the NCC approach is qualitatively examined. The building complex poses the challenge of very steep aspect angles to the window extraction algorithm. Figure 5.6a shows the result of the window extraction approach. The red rectangles are the detected outlines of the windows. The blue points are those corners which are assumed to be the PS inducing structure of the window inferred from projecting the sensor's LOS into the rectified images. Here, the LOS approximately coincides with the capturing direction of the optical image. With only a few exceptions, the extent as well as the shape of the windows are correctly identified. To obtain a more detailed impression Figures 5.6b and 5.6c present two facades in more detail. Both facades are shown in the rectified image plane where the window extraction process takes place. The color encodes the correlation coefficient which measures the individual similarity to the mean of all window patches. As a general assumption, the pattern is assumed to be of a rectangular form. The upper right area of the facade in Figure 5.6b infringes this assumption. This is reflected in the correlation coefficient and can be taken into account in further processing steps. A similar violation is shown in Figure 5.6c where the lower right area of the facade is occluded by another part of the building. In this case, however, the lattice information can be used to infer on occluded window corners. Those can be matching partners for PSs even though there is no evidence from the optical imagery. These two examples show how missing optical information can be interpreted. Dependent on the actual situation, only weak support for a window corner may imply that parts of the facade are occluded but visible in the SAR scene.

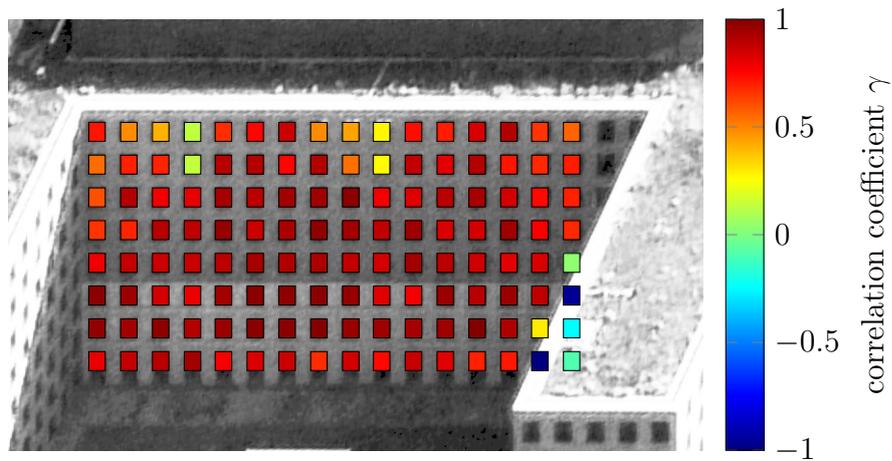
An example for the robustness of the NCC approach can be seen in Figure 5.6b. The shadow of a neighboring facade falls on parts of the building face under investigation, leading to extremely heterogeneous local contrasts. Due to the normalization, NCC is independent of the global contrast. Thus, a globally very small contrast as it can be found in the shaded parts of the facade is sufficient to achieve a high NCC value. Furthermore, the presented approach is able to deal with the steep aspect ratio. The outlines of the windows can correctly be extracted. The extent is nearly correctly estimated. The robustness against varying local contrast can also be seen in Figure 5.6c where the lower part of the facade has a higher contrast with respect to the window due to a brighter facade color.



(a) Extracted windows for building complex Park Kolonnaden.



(b) Extracted windows with correlation coefficient.



(c) Extracted windows with correlation coefficient.

Figure 5.6: Window extraction results for building complex Park Kolonnaden.

5.4.2. Window corner extraction: Charlottenstraße

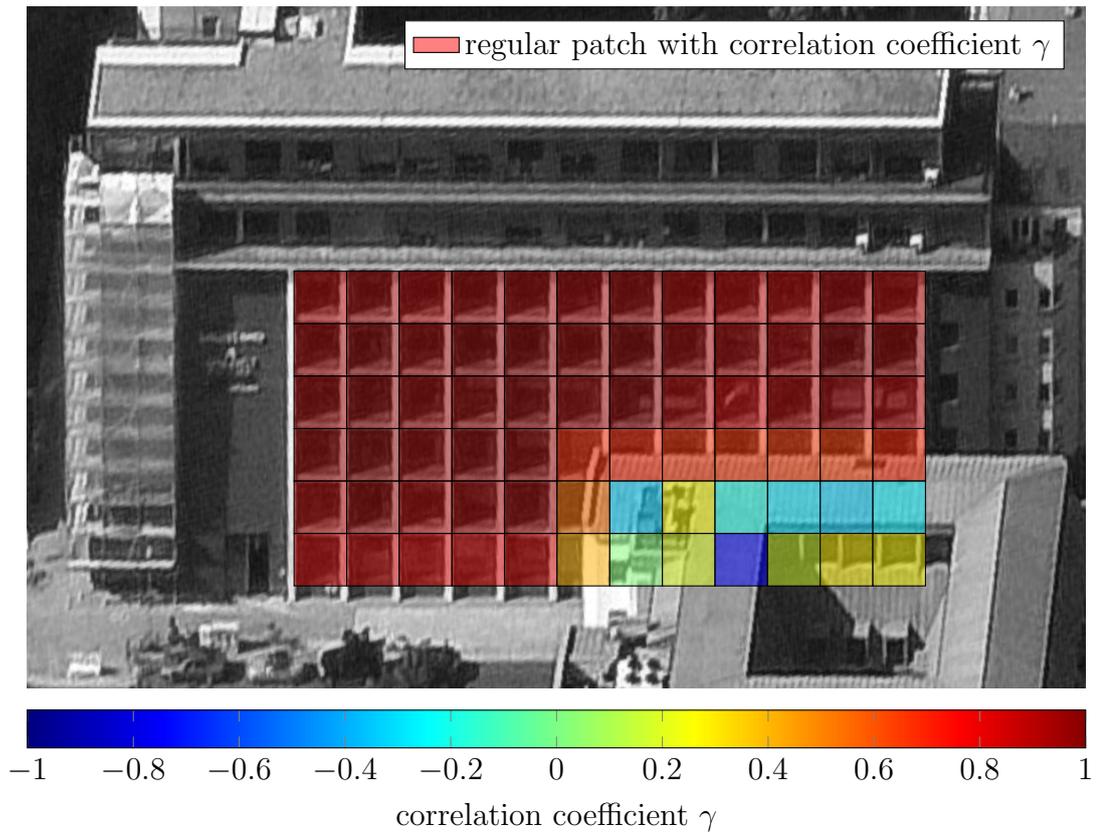
In continuation of the case study Charlottenstraße, Figure 5.7 shows the result of the corresponding window extraction process. In Figure 5.7a, the segmentation of the facade into regular patches can be seen. The color encodes the correlation coefficient between each patch and the mean. Similar to the previous case study, the low values for γ in the lower right corner are caused by another building which occludes a part of the regular facade. Figure 5.7b shows the result of the window estimation. The corners of the red rectangles are the intersection points of the most dominant lines in the edge pixel in horizontal and vertical direction, respectively. The blue points are the selected corners as matching partners and used as input for the IGBM algorithm. The selection of the lower right corner is directly inferred from the sensor's LOS which is shown as a red vector.

This case study shows an example where the simplified assumption that the geometrical structure inducing the PSs can directly be extracted from the optical image is violated. The outlines of the extracted windows correspond to the lines with the most edge pixels in each regular patch. Considering the oblique image of the facade, it can be assumed that the phase center of the reflection mechanism may be more recessed towards the inside of the facade. The shadows on the window sills suggest that the third surface of the trihedral corner is located a few centimeters inside the facade and may not be visible in the oblique image. Nevertheless, the extracted window corners are further used as matching partners for the PSs.

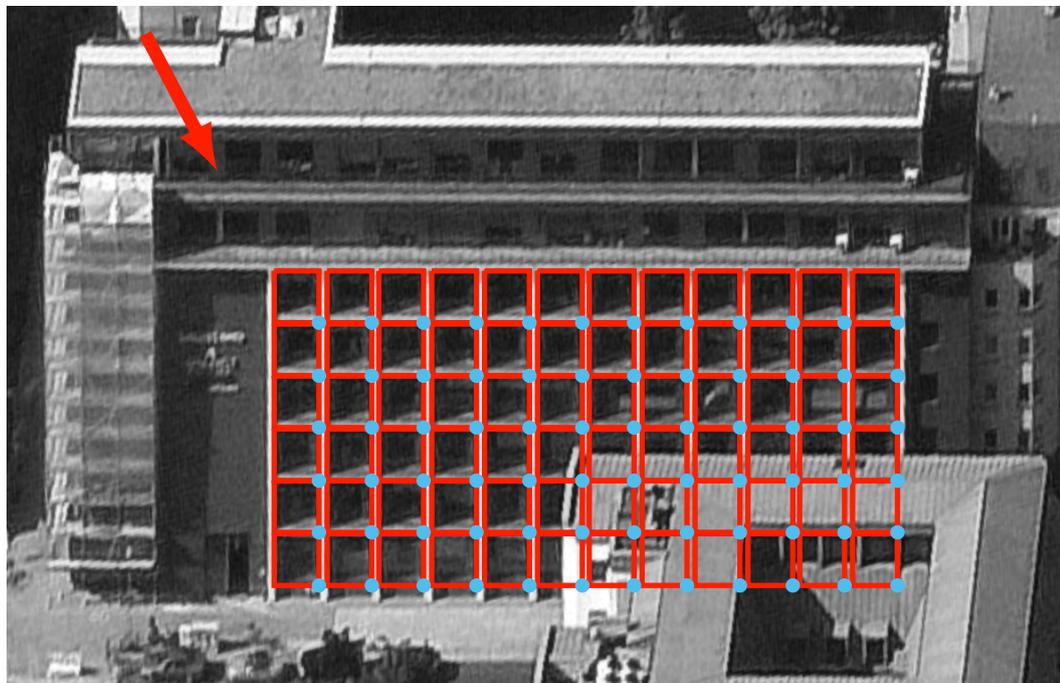
5.4.3. Window corner extraction: Axel-Springer building

The Axel-Springer building is an example of a typical high-rise office building in densely developed city centers. It is presented here as a case study in order to show that the strong lattice model helps to overcome local violations of important assumptions. The tower measures 78 meters in height and has 19 floors. Figure 5.8a shows the regular alignment of braces structuring the facade. The PSs which successfully could be grouped together are shown in green. Irregular scatterers on the facade are marked in blue and red crosses symbolize PSs presumably not induced by the facade.

The window fitting process as described in Section 4.4.2 comprises two major steps. First, the facade is segmented into regular and irregular patches, and second, an explicit window model is estimated. Reflections on the glass panels as well as slightly underestimated lattice spanning vectors lead to violations of the assumption that the optical image of the facade



(a) Segmentation into regular patches.



(b) Extracted window outlines and corners. Sensor's LOS is shown as red arrow.

Figure 5.7: Example of window extraction for case study Charlottenstraße.

consists of repeated patches of the same element. The gray-value image of the facade texture inside the red rectangle shows that this is not always the case. The values of the correlation coefficient in Figure 5.8c show that γ decreases towards the four corners of the facade. Nevertheless, the lattice extent can correctly be extracted since the values for γ in the center regions of the facade are high enough. Also, the window modeling benefits from the assumption that every patch of the facade is a repeated version of the same element. The reflections in the glass panels marked with the red rectangle in Figure 5.8a have a strong effect on the edge image as it can be seen in Figure 5.8b. In the lower part the edge pattern almost vanishes. Nevertheless, due to the introduced model knowledge the window outlines and consequently the window corners can correctly be extracted as shown as blue points in Figure 5.8c.

5.5. Iterative graph-based bipartite matching

The IGBM algorithm is applied in four case studies. The coherence of the resulting matching vectors for different facades in one image is investigated as well as the influence of different values for the parameter α .

5.5.1. IGBM: Park Kolonnaden

The building complex Park Kolonnaden is characterized by very regular facade structures which can be seen in Figure 5.9a. A total of 15 facades are visible in both data sets. The automated PS grouping and window extraction from the optical image yielded results for 12 facades. IGBM is applied to this subset of facades. Two hypotheses are tested here. First, the behavior for facades with different size is further investigated. Second, for similarly oriented facades, the resulting matching vectors are expected to be equal. This assumption is examined. It is motivated by considering the camera orientation which is constant for one image. Since the PSs also originate from the same processing, the matching vectors from the plane-corrected regular PSs onto the resulting position after applying IGBM should be the same. Violations of this behavior can be due to incorrectly extracted window corners, for instance.

Figure 5.9a shows the result of IGBM. In order to better interpret the figure, the extracted window corners are not shown. Figure 5.9b shows an example for the matching result of a large facade and Figure 5.9c the smallest one. The legend in Figure 5.9b holds for all three

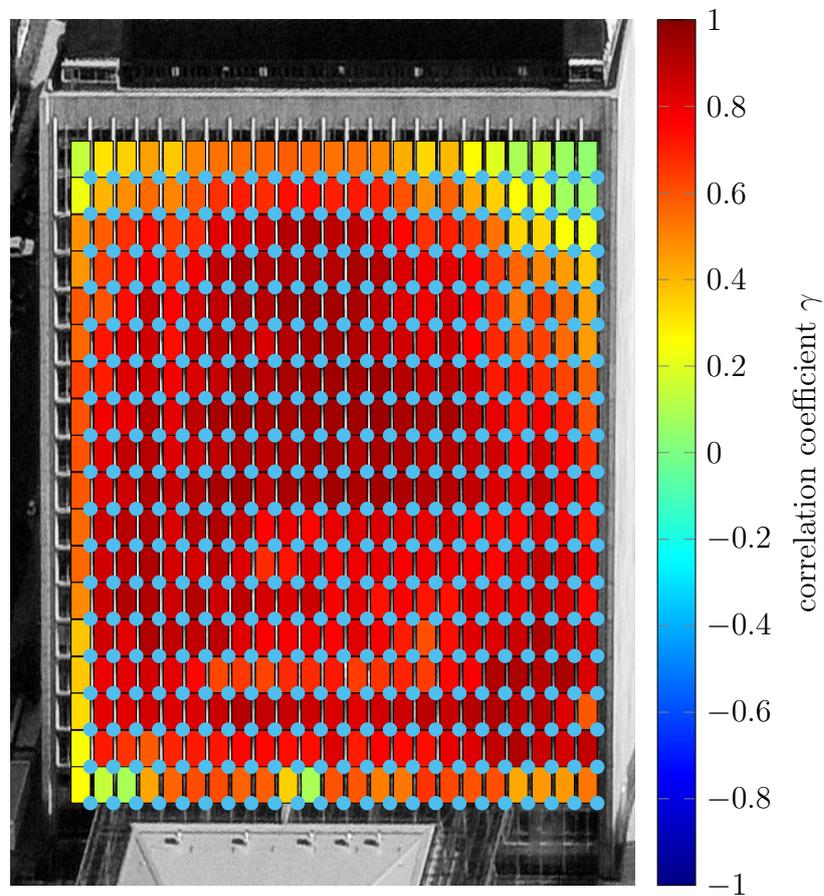
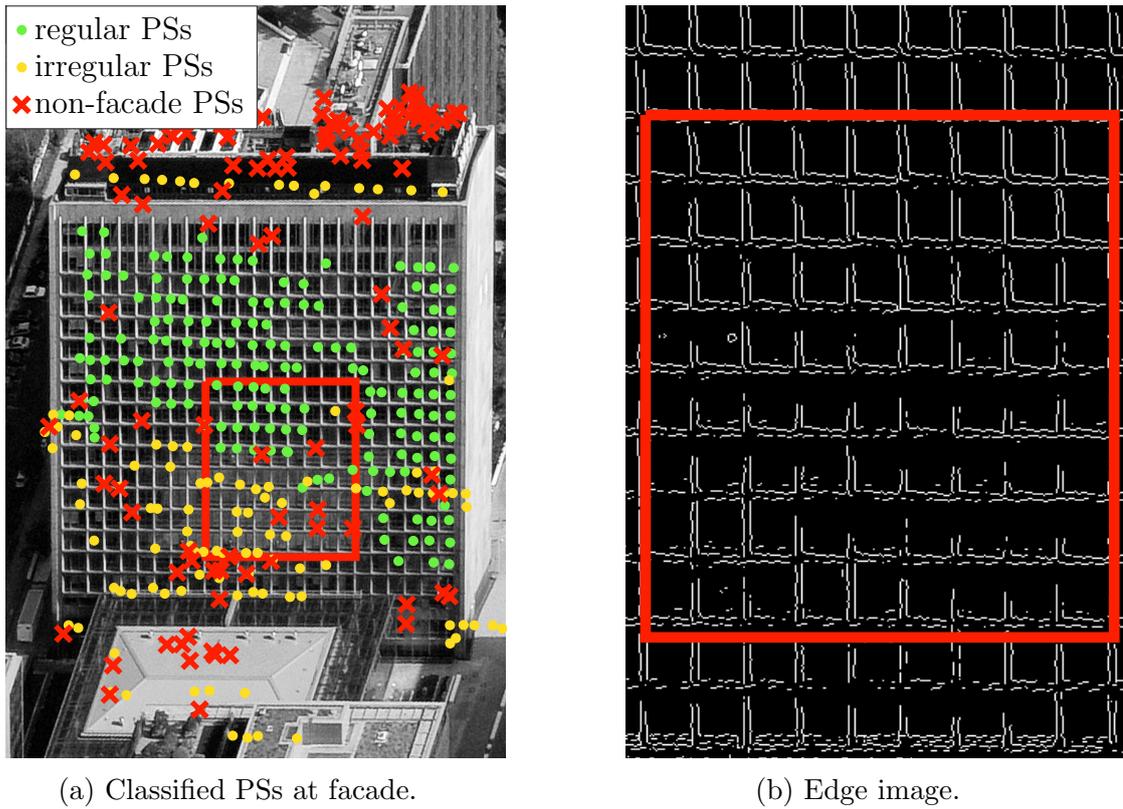
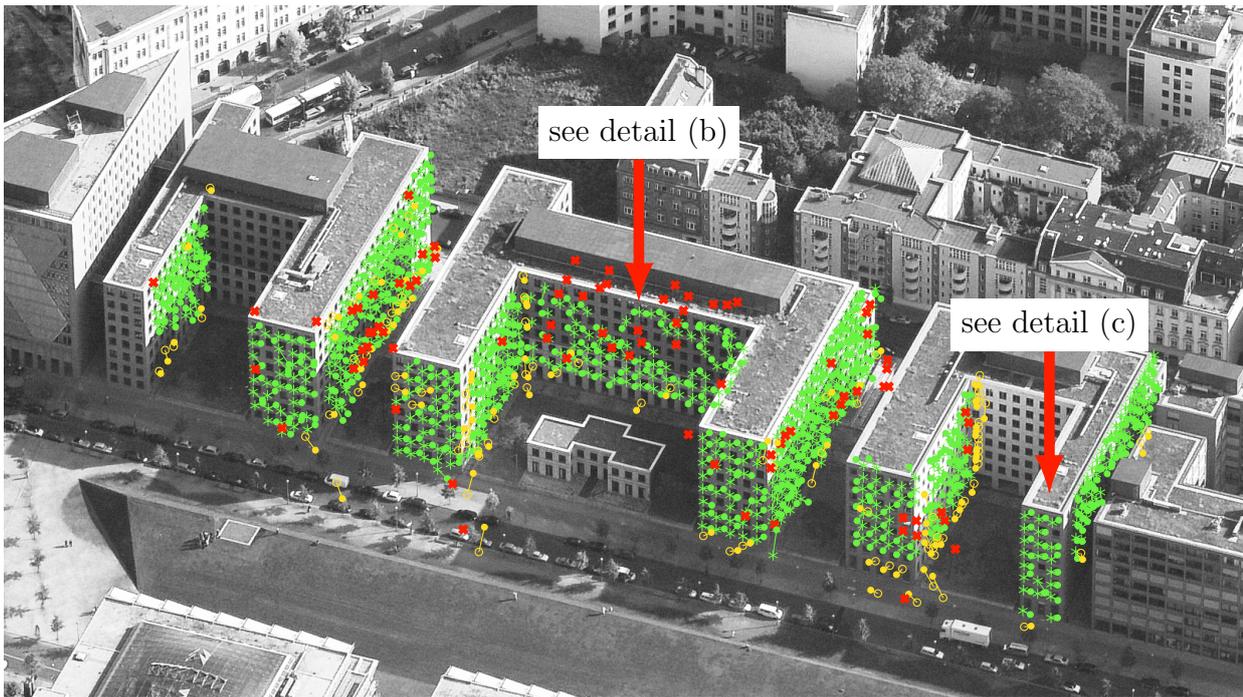
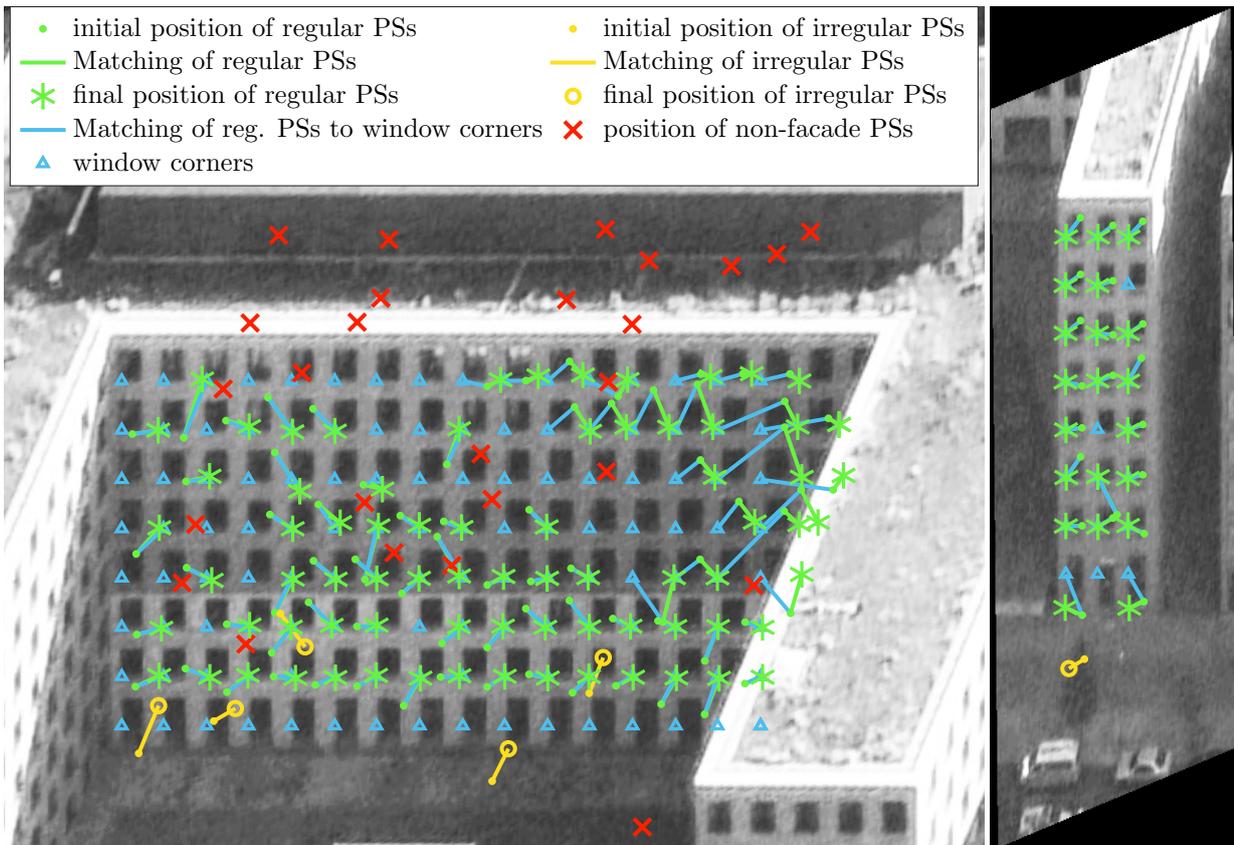


Figure 5.8: Result of window extraction process for case study Axel-Springer building.



(a) Result of IGBM for building complex Park Kolonnaden.



(b) Detail of large facade.

(c) Detail of small facade.

Figure 5.9: IBM result Park Kolonnaden. The legend in (b) is valid for Figures (a) – (c).

illustrations and the following description of symbols and colors is valid for all subsequent figures showing IGBM results. The extracted window corners are shown as blue triangles. PSs classified as regular are shown in green, irregular scatterers are yellow and red crosses mark RANSAC outliers assumed to not be situated at the facade. The initial positions of regular and irregular PSs before the projection onto the facade plane are shown as green and yellow dots, respectively. These coordinates are the result of directly projecting each PS individually into the rectified oblique image and form the set of initial PS positions. The blue lines show the final IGBM matching result. Based on these assignments, the homography is computed and applied on the set of regular and irregular PSs resulting in image coordinates marked as green asterisks (final positions of regular scatterers) and yellow circles (final positions of irregular scatterers), respectively. The resulting matching between the initial and final positions are shown as green and yellow vectors.

The overall assignment of PSs to window corners and the improved position in the image is meaningful. The large facade in Figure 5.9b is an example where the occluded facade leads to missing matching partners for the PSs on the right side of the building's face. The bipartite matching enforces a one-to-one matching resulting in strikingly erroneous matchings in the upper right part of the facade. Nevertheless, the estimated homography between the initial PSs and their matching partners is robust enough to ensure meaningful final positions of the PSs. As an example for small facades Figure 5.9c shows the correct result of a building face consisting of only 21 regular and one irregular PS.

The whole building complex has 6 facades oriented towards the West and 6 facades oriented towards the South. Figure 5.10 shows their averaged matching vectors in image coordinates of the original image, i.e. they are not rectified in order to make them comparable to each other. The projection onto the facade plane is performed prior to computing the matching vectors. The vectors marked in red correspond to the small facades oriented towards the camera in Figure 5.9a. The blue vectors belong to the facades oriented towards the right side of the image. The red vectors are more consistent to each other which is due to the fact that the window corners are easier to determine at these facades. Nevertheless, the expected uniformity of the averaged vectors is not present. A reason might be the initial plane fitting. A slightly wrong plane direction may result in leveraged inaccuracies for the initial PS positions. Here, additional information like building outlines from GIS data bases could stabilize this step.

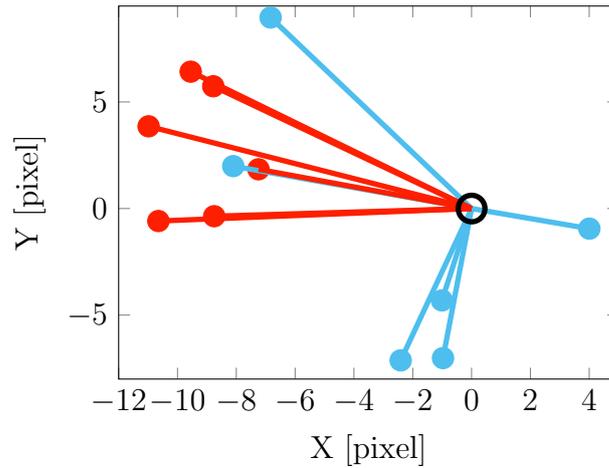


Figure 5.10: Comparison of averaged matching vectors between planned regular PSs to final positions after applying IGBM. Each vector represents the mean matching vector between the initial and final PS positions per facade. Red: Facades oriented towards West. Blue: Facades oriented to South.

5.5.2. IGBM: Charlottenstraße

This case study shows the influence of parameter α on the resulting matching. In Figure 5.11 the result of IGBM applied on the case study Charlottenstraße for two different values of α is shown. Figure 5.11a depicts a configuration where the solution is solely determined by geometry ($\alpha = 1$). Especially at the right side of the regular part of the facade, some matches that do not coincide with the other matching vectors are present. As stated above, this contradicts the assumption that due to the systematic characteristic of the influence of the camera orientation, the matching vectors are expected to be rather constant despite the influence of the elevation direction. A second solution for $\alpha = 0.5$ is shown in Figure 5.11b. Now, the majority of matching vectors are consistent and coincide with the topology derived from the PS grouping in SAR coordinates. Nevertheless, the overall influence of the false matchings on the final PS positions in the image is marginal and significantly smaller than the extent of a window. This fact shows that the method is not very sensitive to changes of α . The sensitivity of IGBM on this parameter is investigated in detail in Section 5.6.

5.5.3. IGBM: Axel-Springer building

In continuation of the case study Axel-Springer building, the IGBM algorithm is applied to its facade consisting of 23 horizontally and 18 vertically aligned regular facade elements. The aim of this experiment is to show the applicability of IGBM for large facades. Figure

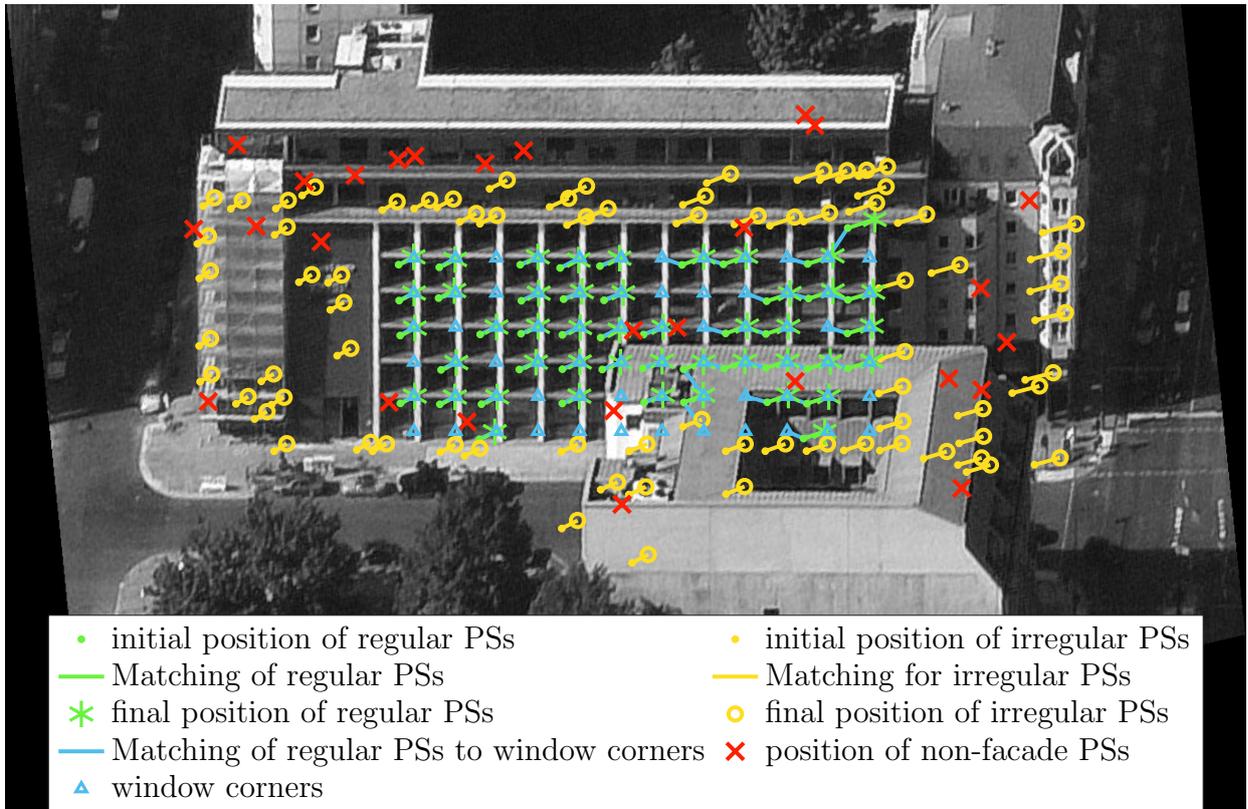
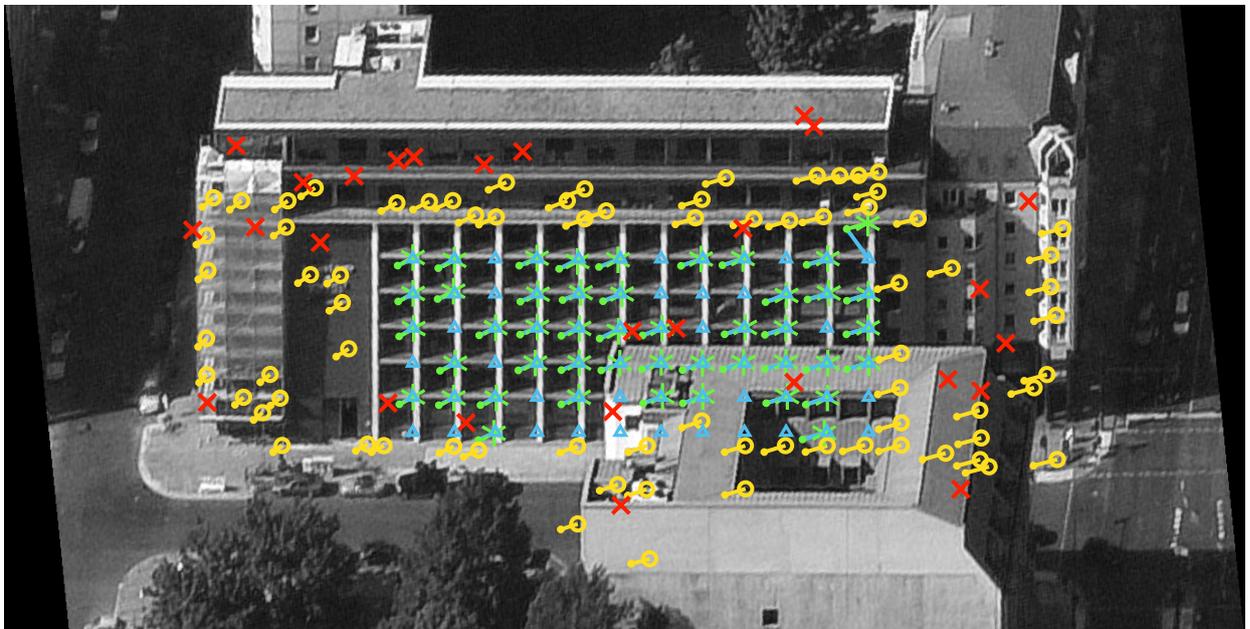
(a) Result of IGBM with $\alpha = 1$.(b) Result of IGBM with $\alpha = 0.5$.

Figure 5.11: Result case study Charlottenstraße.

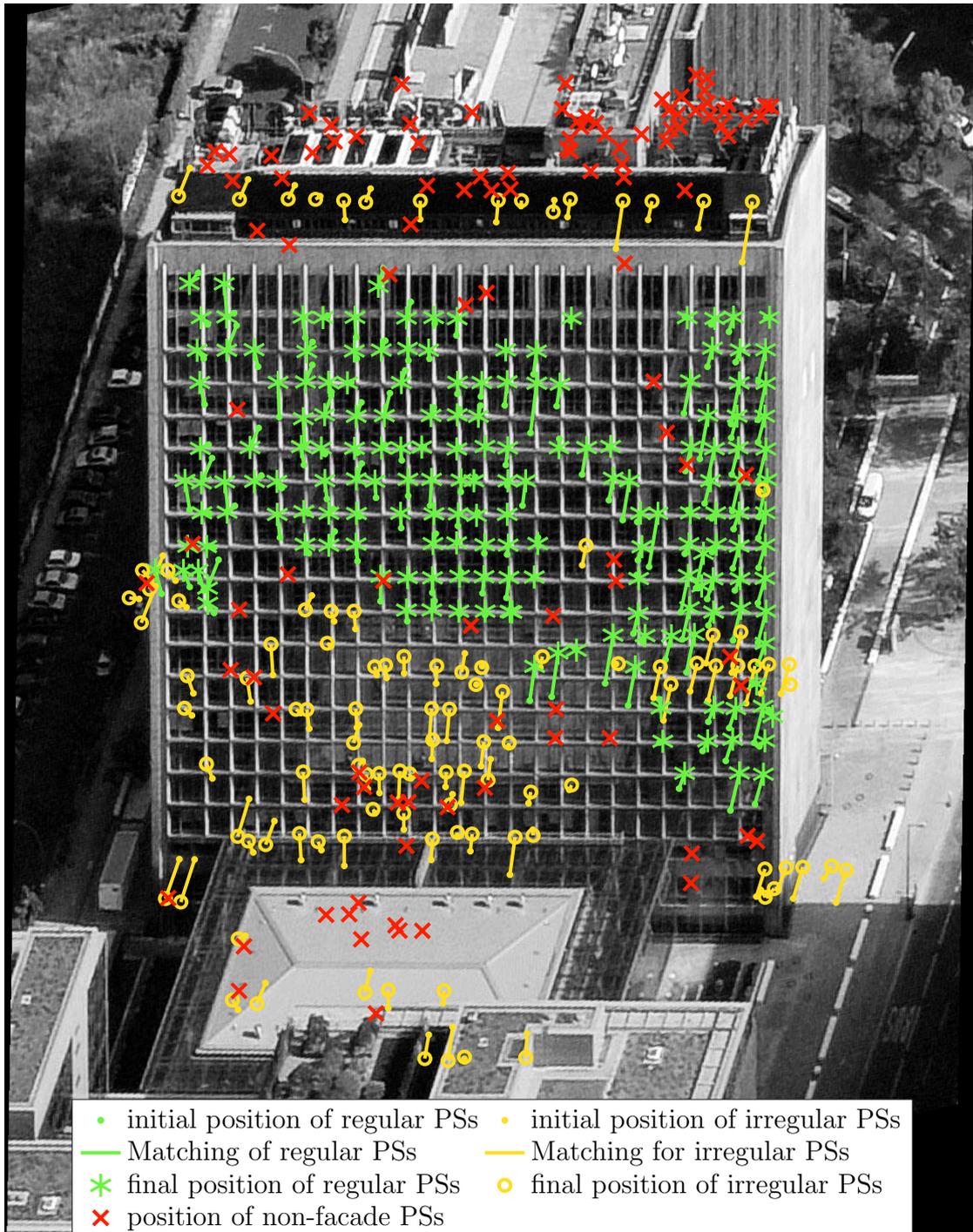


Figure 5.12: Result of IGBM applied to the case study Axel-Springer building.

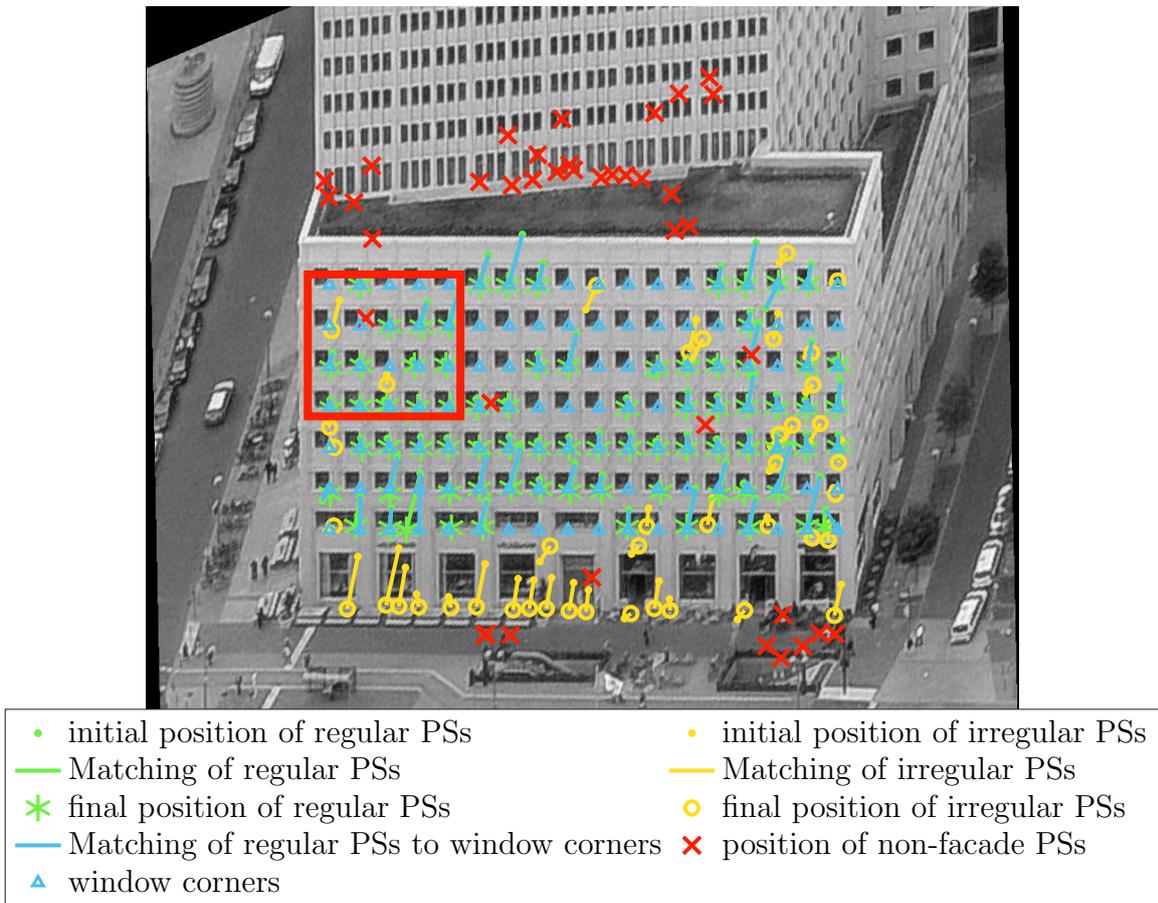
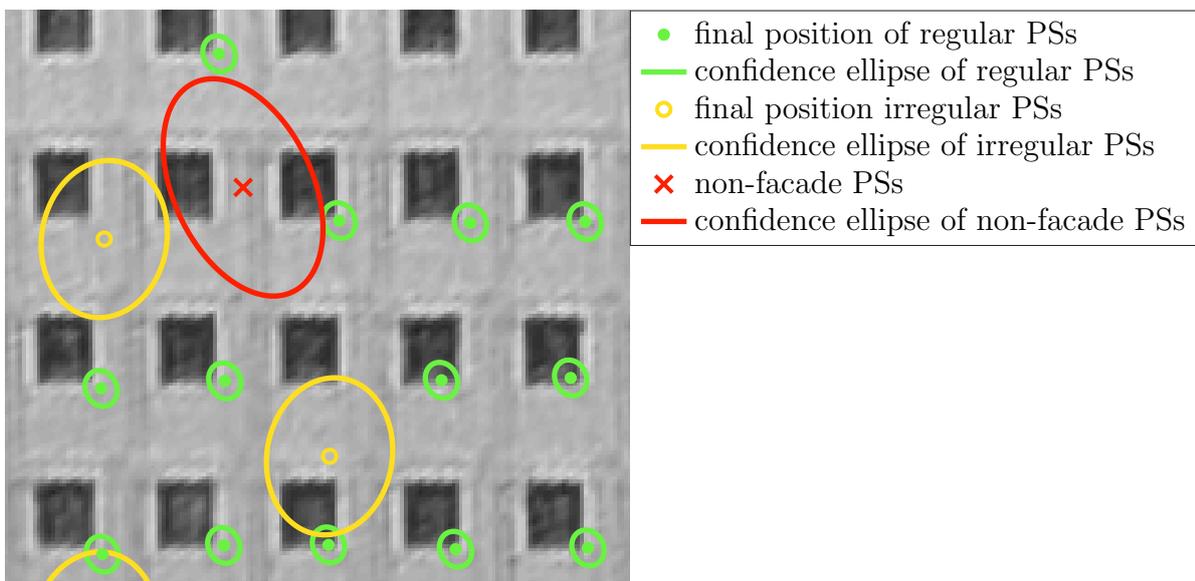
5.12 shows the final result for $\alpha = 0.5$. For better visibility, the extracted window corners are not depicted. They are the same as shown in Figure 5.8c on page 108. The visual interpretation suggests that correct assignments between regular PSs and the corresponding window corners are found. An important observation can be made considering the lower left part of the facade. In this area, no PSs were classified as regular which is due to the exceedance of the threshold in range-azimuth coordinates when the affiliation to the facade plane of each PS is checked. As a result of the window extraction process of the same facade, the lower left corner also had smaller correlation coefficients than the rest of the facade. Since both data sources coincide it can be inferred that the regularity in that part of the window is indeed smaller compared to the rest of the facade.

Another interesting detail can be seen at the roof of the building. The estimated homography mapping the regular PSs onto their correspondences at the facade is also applied to the irregular scatterers. This and the previously performed projection onto the facade plane results in a strikingly regular alignment at the edge between the roof and an elongated structure on top of the roof. This leads to the conclusion that these PSs are induced by some scattering mechanism located at this particular edge. Furthermore, some ghost-scatterers [Auer et al., 2011] seem to be present in the data. They can be identified as irregular PSs in the lower part of the figure. They are inliers of the RANSAC plane estimation but located beneath the facade in the image.

5.5.4. IGBM: Ebertstraße

This case study is shown in order to prove the usefulness of applying the homography inferred from the matching vectors onto irregular PSs. The building Ebertstraße is characterized by very well distinguishable windows. On top of the facade under investigation another facade with a slightly different orientation is situated. Figure 5.13a shows that the classification approach successfully separates these scatterers from those belonging to the facade under investigation. The result of the IGBM algorithm is reasonable. A value of $\alpha = 0.75$ is chosen. The results for irregular scatterers are of special interest. The resulting positions marked with yellow circles often coincide with extracted window corners. This means that even though their positions in the range-azimuth plane lead to an exclusion of the set of regular points, the application of the estimated homography leads to a presumably correct position at the facade. This is another indicator of the robustness of the IGBM method.

Figure 5.13b depicts an enlarged detail of the upper right side of the facade. It shows the three confidence ellipses in relation to each other. In red, the 95% confidence ellipse as a

(a) Result of IGBM algorithm with $\alpha = 0.75$.

(b) 95% confidence ellipses for enlarged area inside red rectangle in (a).

Figure 5.13: IBM result for case study Ebertstraße with 95% confidence ellipses.

result of variance propagation with the parameters introduced in Table 2.3 on page 31 is shown. It includes the inaccuracy budget of the SAR coordinates as well as that of the camera orientation. In yellow, the result of projecting all appropriate PSs onto the facade plane can be seen. Applying prior knowledge about those PSs which are found to be RANSAC inliers of the estimated plane reduces the influence of the elevation direction. Finally, assuming that all assignments are correct, the corresponding homography is estimated and from the resulting matching vectors the green confidence ellipse is computed. Thus, the assignment of individual PSs to repeated elements is possible after applying the projection on the facade (yellow confidence ellipse). A further reduction by estimating a covariance matrix after applying the homography constitutes a lower bound of the possible assignment accuracy. This approach assumes that the window corner extracted from the optical image coincides with the phase center of the radar scattering mechanism. This is not necessarily the case but still a good approximation and allows the assignment of individual PSs to repeated facade elements.

5.6. Balance between geometry and topology

As outlined in Section 4.5.2, the factor α in Equation (4.12) controls the balance between the geometry and topology terms with respect to the optimization problem. In this section, the dependence of the solution in dependency of α is investigated. The sensitivity is determined by qualitatively and quantitatively comparing the solutions of IGBM with varying values for α . Two cases are distinguished. First, a facade with mainly regular and very few irregular PSs or even scatterers not belonging to the facade is investigated. Here, it can be expected that the dependency of the result of the IGBM algorithm on α is small since both the geometry and the topology should strive in the same direction. As a second case, the case study Charlottenstraße is revisited. Here, a higher dependency on α is to be expected due to the fact that more PSs are present which disturb the regular pattern.

Figure 5.14 compares the outcome of the scores measuring the geometrical fit in blue, the topological score in red, and the combined score in dependency on α (black). The scores are normalized since they depend on the number of matched PSs. In Figure 5.14a, no significant change in the scores can be identified except when the geometry is totally neglected, i.e. $\alpha = 0$. A very small jump close to $\alpha = 0.6$ corresponds to the example shown in Figure 4.12 (page 80). Before this jump, the topology holds on to the solution with correct topology. Further increasing α makes the solution more dependent on the geometry which results in a smaller geometry score but larger topology penalty. Figure 5.14b shows an

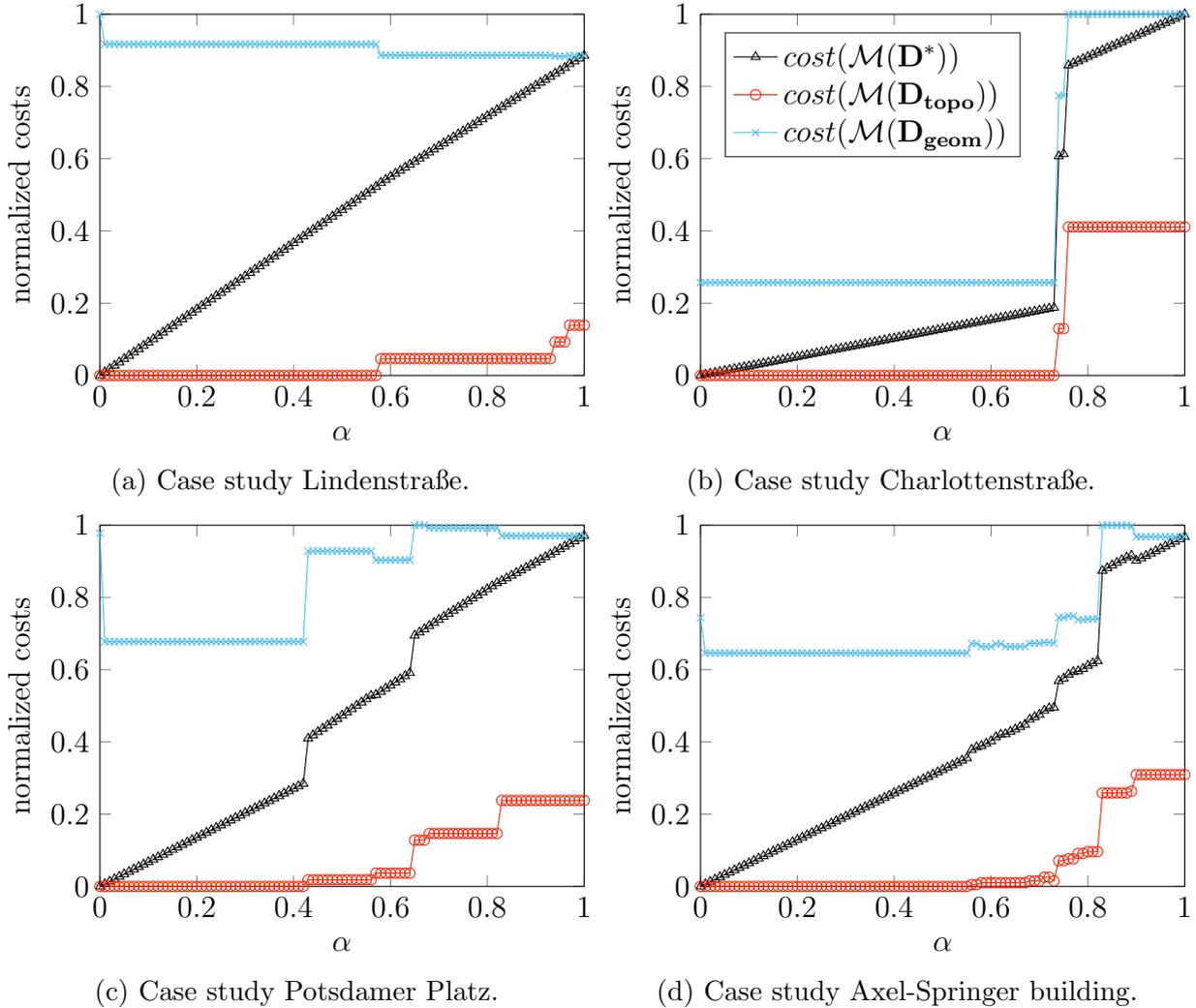


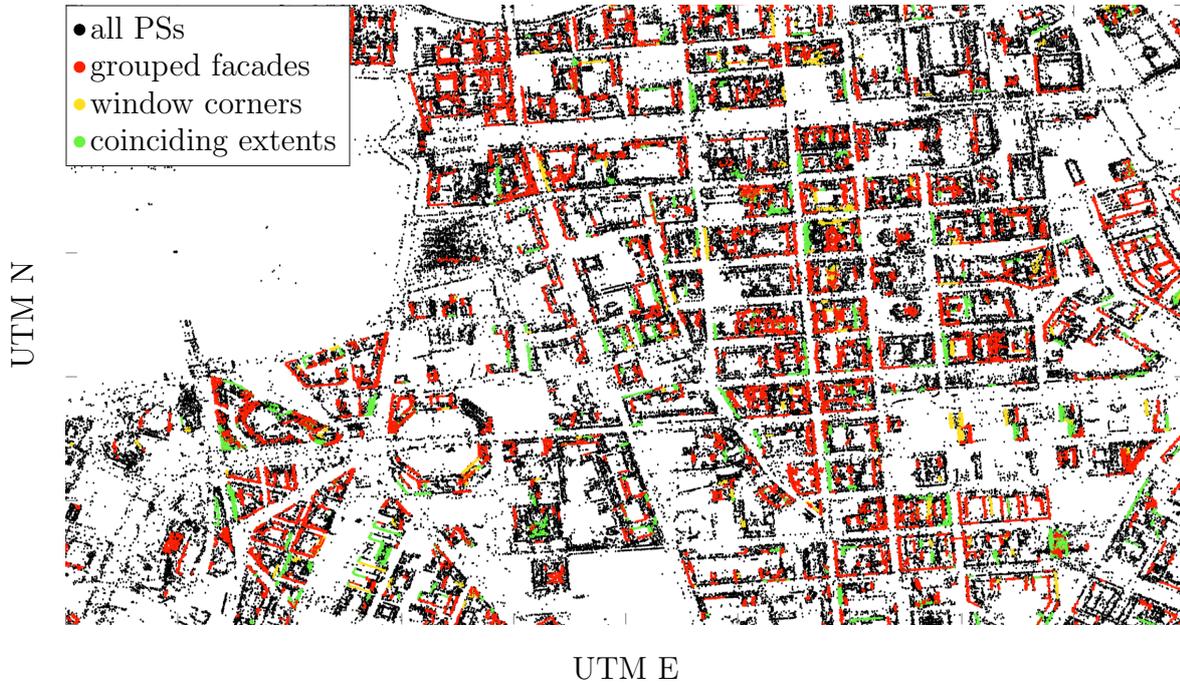
Figure 5.14: Dependency of the IGBM result on α . For each case study, the matching costs only considering the geometry score are shown in blue. The equivalent for topology is depicted in red. $cost(\mathcal{M}(\mathbf{D}^*))$ are the combined matching costs of the linear combination of both in dependency of α . The costs are normalized since they depend on the number of PSs and window corners, respectively.

analogous investigation for the case study Charlottenstraße which was already investigated in aforementioned experiments. In this case, both scores change into the same direction. This means that for such facades the IGBM converges to solutions where the topology and geometry prefer the same result. Also, facades with a mixed form occur as shown in Figures 5.14c and 5.14d. In all cases, the sensitivity is rather low meaning that over broad ranges of α the matching costs do not change. As stated before, the challenge at hand is that due to the lack of ground truth data, no *true* value of α is known at this stage. Nevertheless, only a small number of solutions exists. This allows for visually comparing the solutions and manually selecting an appropriate α .

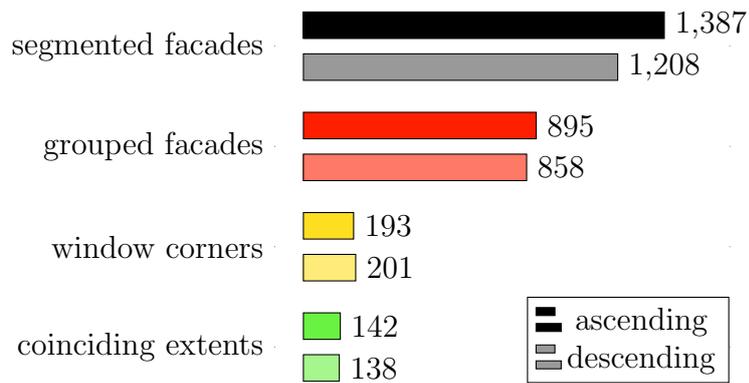
5.7. Application to the whole test area

The previous case studies showed some particular results and characteristics of the developed methods. In this section, the application to the whole area under investigation is shown. In order to do so, the approaches are applied to the whole PS point cloud with the standard parameters which were introduced in Table 5.3 on page 97. After segmenting the whole point cloud into clusters of scatterers belonging to individual facades, the grouping approach is performed followed by automatically extracting appropriate window corners as matching candidates. Then, the extracted extents are compared in order to validate the accordance between the PS grouping result and the lattice extraction from the oblique imagery. If one facade is visible in more than one image, the one where the building face is depicted the closest to the scene center is chosen.

Figure 5.15a shows the result for the whole study area in UTM coordinates. In black the merged PS point clouds from the ascending and descending TerraSAR-X High Resolution Spotlight stacks are shown. Figure 5.15b shows the quantitative result. A total of 2595 facades with more than 9 PSs are segmented and fed into the PS grouping process. Without adapting the parameters individually, a total of 1753 facades can be grouped successfully. This is a ratio of 68% and the corresponding facades are shown in red. After projecting the grouped PSs onto the corresponding facade planes, the automated extraction of windows from optical images is performed. Those facades where the extraction was successful are shown in yellow. The estimated extents from the grouping in SAR coordinates and from the lattice estimation in the optical image are compared. In order to obtain a rough estimate about the correspondence of the PS grouping result with the window extraction from optical imagery, the result is considered to be coinciding when the extent in vertical and horizontal direction vary with not more than one row or column, respectively. Of course, different



(a) Results of complete study area.



(b) Number of successfully processed facades.

Figure 5.15: Application of methods to overall study area.

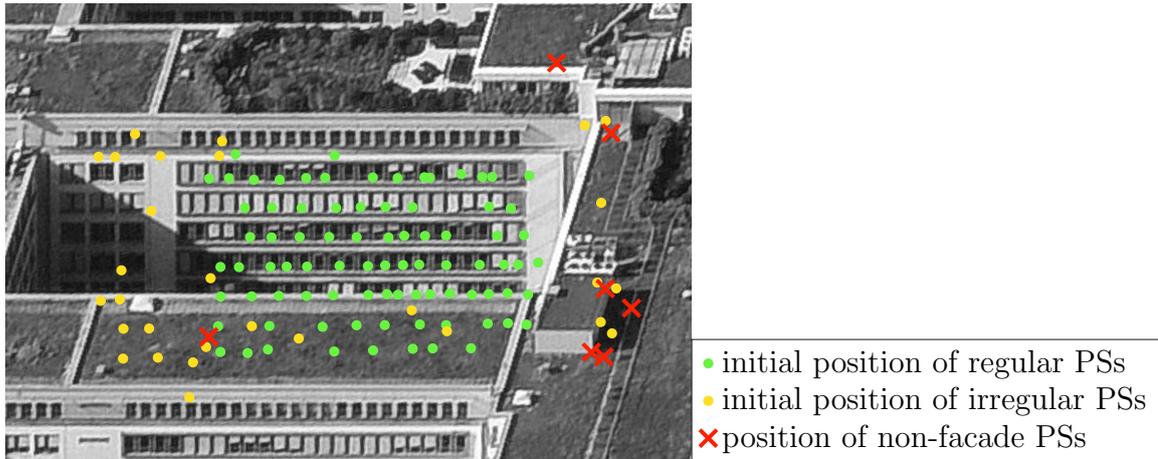


Figure 5.16: Example of facade with regular PS pattern but no extractable window lattice.

viewing directions and the possibly resulting occlusion may also lead to facades with more than one row or column not visible in the data. Nevertheless, the aforementioned criteria is used as a rough estimate for comparing the results of the grouping approaches in SAR data and optical images. These 280 facades with coinciding extents are shown in green. Their proportion is approximately 11% with respect to the total number of segmented facades with at least 9 PSs. There is no significant difference in these numbers between the ascending and descending SAR stacks.

The first conclusion is that more than two thirds of all facades with more than 9 PSs can be considered as being regular. However, for only 22% of those regular facades the automated window extraction yields successful results. That means that even when a pattern is visible in the PS data, for only one fifth of these cases the same can be said considering the optical imagery. Many facades show regular alignments in PS point clouds which cannot be extracted from the optical imagery using the described approach. Figure 5.16 shows an example for such a case. For instance, the appearance of the windows are very heterogeneous due to roller blinds. This is a violation of Assumption 2 stating that the appearance of each repeated element is similar. Furthermore, a small part of the facade is occluded by the building in front of it.

6. Conclusions and open questions

In this chapter, the most important findings are summarized and further research questions are outlined. The following four main contributions are covered in this thesis. First, the grouping of PSs from TerraSAR-X image stacks into lattice representations of the corresponding facade objects. Second, the extraction of window corners from oblique aerial images aided by the results of the grouping in the SAR domain. Third, the mathematically sound matching of both aforementioned data sets. The fourth contribution is the application of the matching result which can be used in order to learn something about the physical nature of PSs at facades. One important step towards this goal is made possible by reducing the systematical inaccuracy of the image coordinates of PSs induced by the exterior orientation of the camera as well as random contributions originating from the PSI processing.

6.1. Summary and conclusion

Grouping of PSs

The challenging task for PS grouping is that not all PSs at the facade belong to the regularity. This requires a robust approach and is taken account by transforming all PSs into a translation space where the influence of outliers is reduced. The grouping process reliably yields results suitable for the subsequent processing steps. Even cases where the regular pattern consists of only a small subset of all PSs can correctly be handled. Having reasonable prior knowledge about the SAR sensing geometry as, for instance, the fact that all vertically aligned scatterers have the same azimuth coordinate, allows one to exhaustively search for the best fitting lattice over all possibilities. The grouping is carried out in the precisely determined range-azimuth coordinates plane to infer topology information in order to stabilize the subsequent matching. Even though the presented grouping result is limited to a very basic form of regularity, it is the first attempt to group PSs in two dimensions for the sake of matching them to counterparts in optical imagery.

Extraction of window corners from optical oblique imagery

The extraction of window corners from optical oblique imagery is motivated by the findings of Bamler et al. [2009] which states that even small trihedral corners with side lengths of 8 cm induce PSs as well as the investigations with the help of simulation techniques of Auer et al. [2010]; Auer [2011] which suggest window corners at facades as the most probable structures inducing PSs. This makes sense in particular for cases where regular PS patterns coincide with alignments of windows at the facade. Optical oblique imagery delivers very strong hints for geometrical features like window corners. The presented approach of extracting such corners is based on standard image processing techniques like NCC and edge detection. These methods yield reliable results and do not need additional training steps. The application of these data driven approaches is made possible by very precise approximation values coming from the SAR data domain. The approximate facade outline is known as the result of a plane fitting in the segmented PS point cloud, for instance. This information reduces the search space significantly. The experiments showed that even for facades imaged under steep aspect angles meaningful results are achieved. The extent of the regular patterns as well as the outlines of the windows are correctly extracted in most cases. Even cast shadows on parts of the facade do not significantly disturb the extraction process. Projecting the SAR sensor's LOS into the image allows for selecting the most probable window corner when only threefold reflections are considered. Nevertheless, the assumption that the geometrical structure inducing the PSs coincide with the window corner which is extracted by considering the intersection points of line segments in the edge image, does not always hold. However, the assignment of PSs to these corners is assumed to be a good approximation of the *correct* phase centers. Window extraction was also applied to the whole data set and yielded a proportion of roughly 22% of all facades with at least 9 PSs.

Matching of PSs to window corners

An algorithm for matching PSs to window corners is developed. Next to the assignment itself, the geometric transformation mapping the PSs to their matched window corners is optimized. The challenge here is that both entities are mutually dependent on each other. To solve this problem, an iterative procedure is derived and proven to converge and yield the optimal solution. The assignment is done without any additional data, as opposed to related work. Instead from building outlines, the facade plane is estimated directly from the data. Furthermore, no three-dimensional building model is needed. The transformation into

the optical imagery is solely determined by the camera orientation and constraints imposed by the topology.

As the largest limitation of the presented method, the regular pattern has to be present and automatically extractable in data coming from both sensor types. This restricts the applicability to only a subset of all facades captured by the sensors. Nevertheless, the presented approach positively answers the research question if PSs can be assigned to individual parts of facades for this subset. The applicability is shown in several case studies reaching from small facades with a regularity of extent 3×7 up to multistory buildings showing a pattern of size 23×18 . The application of IGBM to the overall investigation area showed that the assignment of PSs to window corners was successful in terms of an coinciding extent of the regularity for 11% of all facades with at least 9 PSs. Furthermore, the sensitivity of the matching algorithm on the parameter α is found to be rather small. Typically, certain states were found which are stable over rather large ranges of α .

Improving the image coordinates of matched PSs

In order to answer the question of how precisely the assignment of PSs to their counterpart in optical oblique images can be done, two measures increasing the accuracy are applied. The first is to reduce the inaccuracy in elevation direction by considering all PSs in the facade plane as repeated measurements of samples of this surface. The second approach is to apply a homography transformation which maps the regular PSs to their corresponding window corners. The resulting transformation can also be applied to irregular PSs as long as they are situated in the same plane. The resulting covariance matrix is empirically estimated from the residuals of the transformed PSs and the matched window corners.

The assignment to individual facade elements can be made as even the 95% confidence ellipses of adjacent regular PSs in the image do not overlap. But as shown in Section 4.6 (page 88), even for the case that the assumptions are violated, the 95% confidence ellipses are sometimes smaller than the area of the patch. This depends on the size of the regular structure at the facade. The resulting 95% confidence ellipses cover an area of some 0.1 m^2 to 0.6 m^2 for regular and 4 m^2 to 5 m^2 for irregular PSs.

6.2. Future work

The conducted experiments showed that the parameter α has an influence on the resulting matchings. Besides setting this parameter manually, one could combine the presented method with simulation techniques in order to learn an optimal value for α for specific classes of facade types. Furthermore, the investigated regularity is limited to a very basic lattice description. Even though many facades conform to this type, an extension to more complex patterns would broaden the range of applicability. A grammar as applied by Ripperda & Brenner [2009], for instance, is one way of doing this. All the subsequent processing steps would only need minimal adaptations in order to cope with these more sophisticated types of regularity. The applied covariance matrix for the projection of object coordinates into the image is a simplified model since no correlations were taken into account. As a further extension of the presented approach, those covariances could be modeled based on the geometry of the image acquisition.

First interpretations of irregular PSs were made. A systematic investigation of such points could be performed as a subsequent step. This could open up new applications of detailed building monitoring. The assignment to individual parts of the facade is possible leading to a credible statement on which part is actually monitored. The deformation signal of the PSI product are neglected in this thesis and incorporating it can make deformation modeling possible beyond rigid movements. Here, the combination with the interpretation of optical imagery allows a deeper understanding of the geometrical situation at the monitored facade. Another benefit of the presented method is to use the topology information and infer missing PSs in the PSI processing step. Depending on the used PSI processor, the sensitivity at exactly those positions where the lattice structure predicts a PS could be modified accordingly.

In this thesis, a method for assigning PSs to their correspondences in optical imagery were presented. The challenges here are the highly anisotropic inaccuracy budget of the three-dimensional PS object coordinates as well as a possibly inaccurate exterior camera orientation. Topology information was extracted from the precisely determined range-azimuth SAR coordinates and used twofold. First, the three-dimensional PS coordinates were improved by incorporating model knowledge in terms of assuming scatterers at the facade to be situated in a plane. Second, the optimal matching was found by minimizing a joint score incorporating the geometry as well as the topology. The presented algorithm was proven to yield the optimal solution.

Bibliography

- Adam N, Kampes B, Eineder M (2004) Development of a scientific permanent scatterer system: Modifications for mixed ERS/ENVISAT time series. In: *Proceedings of the 2004 Envisat & ERS Symposium*: 457–465.
- Adam N, Kampes B, Eineder M, Worawatanamateekulb J, Kircher M (2003) The development of a scientific permanent scatterer system. In: Schroeder M, Jacobsen K, Heipke C (eds) *Proceedings of the Joint ISPRS/EARSeL Workshop High Resolution Mapping from Space 2003*
- Auer S (2011) *3D Synthetic Aperture Radar Simulation for Interpreting Complex Urban Reflection Scenarios*. PhD thesis, Technische Universität München, Bayerische Akademie der Wissenschaften, Verlag C.H. Beck. Schriftenreihe C der DGK, no. 660.
- Auer S, Gernhardt S, Bamler R (2011) Ghost persistent scatterers related to signal reflections between adjacent buildings. In: *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*: 1485–1488.
- Auer S, Gisinger C, Bamler R (2012) Characterization of SAR image patterns pertinent to individual facades. In: *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*: 3611–3614.
- Auer S, Gisinger C, Tao J (2015) Characterization of Facade Regularities in High-Resolution SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 53 (5): 2727–2737.
- Auer S, Hinz S, Bamler R (2010) Ray-tracing simulation techniques for understanding high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 48 (3): 1445–1456.
- Bamler R, Eineder M (2005) Accuracy of differential shift estimation by correlation and split-bandwidth interferometry for wideband and Delta-k SAR systems. *IEEE Geoscience and Remote Sensing Letters*, 2 (2): 151–155.
- Bamler R, Eineder M, Adam N, Zhu X, Gernhardt S (2009) Interferometric Potential of High Resolution Spaceborne SAR. *Photogrammetrie - Fernerkundung - Geoinformation*, 2009 (5): 407–419.
- Bamler R, Schättler B (1993) SAR data acquisition and image formation. In: Schreier G (ed) *SAR Geocoding: Data and Systems* chapter 3, (pp. 53–102). Wichmann.
- Botev ZI, Grotowski JF, Kroese DP (2010) Kernel density estimation via diffusion. *Annals of Statistics*, 38 (5): 2916–2957.
- Brunner D, Lemoine G, Bruzzone L (2010) Earthquake damage assessment of buildings using VHR optical and SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 48 (5): 2403–2420.
- Burkard R, Dell’Amico M, Martello S (2009) *Assignment Problems*. Philadelphia: Society for Industrial and Applied Mathematics, revised edition.
- Canny J (1986) A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8 (6): 679–698.
- Cigna F, Bianchini S, Casagli N (2013) How to assess landslide activity and intensity with Persistent Scatterer Interferometry (PSI): The PSI-based matrix approach. *Landslides*, 10 (3): 267–283.
- Conway J, Sloane N (1999) *Sphere Packings, lattices and groups*, volume 338 of *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen*. New York: Springer New York, third edition.
- Crosetto M, Monserrat O, Cuevas-González M, Devanthéry N, Luzi G, Crippa B (2015) Measuring thermal expansion using X-band persistent scatterer interferometry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 100: 84–91.
- Das S, Kapoor K (2014) Fine-Tuning Decomposition Theorem for Maximum Weight Bipartite Matching. In: Gopal, T.V. and Agrawal, Manindra and Li, Angsheng and Cooper S (ed) *Theory and Applications of Models of Computation* (pp. 312–322). Springer International Publishing.

- Dempster AP, Laird NM, Rubin DB (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39 (1): 1–38.
- Duda RO, Hart PE (1972) Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15 (1): 11–15.
- Eineder M, Adam N, Bamler R, Yague-Martinez N, Breit H (2009) Spaceborne spotlight SAR interferometry with TerraSAR-X. *IEEE Transactions on Geoscience and Remote Sensing*, 47 (5): 1524–1535.
- Elberink SO, Vosselman G (2009) Building reconstruction by target based graph matching on incomplete laser data: analysis and limitations. *Sensors*, 9 (8): 6101–6118.
- Ferretti A, Prati C, Rocca F (2001) Permanent Scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 39 (1): 8–20.
- Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24 (6): 381–395.
- Flood MM (1956) The Traveling-Salesman Problem. *Operations Research*, 4 (1): 61–75.
- Fornaro G, Verde S, Reale D, Pauciuolo A (2015) CAESAR: An Approach Based on Covariance Matrix Decomposition to Improve Multibaseline-Multitemporal Interferometric SAR Processing. *IEEE Transactions on Geoscience and Remote Sensing*, 53 (4): 2050–2065.
- Frueh C, Sammon R, Zakhor A (2004) Automated texture mapping of 3D city models with oblique aerial imagery. In: *Proceedings of 2nd International Symposium on 3D Data Processing, Visualization, and Transmission. 3DPVT 2004*: 396–403.
- Gerke M (2009) Dense Matching in High Resolution Oblique Airborne Images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, 38 (part 3/W4): 77–82.
- Gernhardt S (2012) *High Precision 3D Localization and Motion Analysis of Persistent Scatterers using Meter-Resolution Radar Satellite Data*. PhD thesis, Technical University Munich, Verlag der Bayerischen Akademie der Wissenschaften, in consignment at Verlag C.H. Beck. Published in Schriftenreihe C der DGK, no. 672.
- Gernhardt S, Adam N, Eineder M, Bamler R (2010) Potential of very high resolution SAR for persistent scatterer interferometry in urban areas. *Annals of GIS*, 16 (2): 103–111.
- Gernhardt S, Auer S, Eder K (2015) Persistent scatterers at building facades - Evaluation of appearance and localization accuracy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 100: 92–105.
- Gernhardt S, Bamler R (2012) Deformation monitoring of single buildings using meter-resolution SAR data in PSI. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73: 68–79.
- Gould R (1988) *Graph Theory*. Menlo Park: The Benjamin/Cummings Publishing Company, Inc.
- Hammer H, Kuny S, Schulz K (2014) Preparing a new data set for earthquake damage detection in SAR imagery: the Christchurch example II. In: *SPIE Proceedings*. SPIE, 9243: 9243061 – 9243068.
- Hopcroft JE, Karp RM (1971) A $n^5/2$ algorithm for maximum matchings in bipartite graphs. In: *Proceedings of 12th Annual Symposium on Switching and Automata Theory*: 122–125.
- Hung WC, Hwang C, Chen YA, Chang CP, Yen JY, Hooper A, Yang CY (2011) Surface deformation from persistent scatterers SAR interferometry and fusion with leveling data: A case study over the Choushui River Alluvial Fan, Taiwan. *Remote Sensing of Environment*, 115 (4): 957–967.
- Kampes B (2006) *Radar interferometry, Persistent Scatterer Technique*. Springer Netherlands.
- Kao MY, Lam TW, Sung WK, Ting HF (2001) A Decomposition Theorem for Maximum Weight Bipartite Matchings. *SIAM Journal on Computing*, 31 (1): 18–26.
- Kuhn HW (1955) The Hungarian method for the assignment problem. *Naval research logistics quarterly*, 2 (1-2): 83–97.

- Kuny S, Hammer H, Schulz K (2014) Preparing a new data set for earthquake damage detection in SAR imagery: the Christchurch example I. In: *SPIE Proceedings*. SPIE, 9243: 92430X1 – 92430X8.
- Lan H, Li L, Liu H, Yang Z (2012) Complex Urban Infrastructure Deformation Monitoring Using High Resolution PSI. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5 (2): 643–651.
- Lewis JP (1995) Fast Normalized Cross-Correlation. *Vision Interface*, 10 (1): 120–123.
- Lian Y, Shen X (2015) Detecting repetitive elements with accurate locations and shapes from urban facade. In: *Proceedings of IEEE International Conference on Image Processing (ICIP), 2015*: 1920–1924.
- Liu J, Liu Y (2014) Local Regularity-driven City-scale Facade Detection from Aerial Images. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 3778–3785.
- Ma P, Lin H, Lan H, Chen F (2015) Multi-dimensional SAR tomography for monitoring the deformation of newly built concrete buildings. *ISPRS Journal of Photogrammetry and Remote Sensing*, 106: 118–128.
- Mercier G, Moser G, Serpico SB (2008) Conditional copulas for change detection in heterogeneous remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 46 (5): 1428–1441.
- Mesolongitis A, Stamos I (2012) Detection of windows in point clouds of urban scenes. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*: 17–24.
- Michaelsen E, Soergel U, Schunert A, Doktorski L, Jaeger K (2010) Perceptual grouping for building recognition from satellite SAR image stacks. In: *Proceedings of IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS)*: 1–4.
- Moreira A, Prats-Iraola P, Younis M, Krieger G, Hajnsek I, Papathanassiou KP (2013) A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, 1 (1): 6–43.
- Munkres J (1957) Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5 (1): 32–38.
- Neuhaus M, Bunke H (2007) *Bridging the Gap Between Graph Edit Distance and Kernel Machines*. River Edge, NJ, USA: World Scientific Publishing Co., Inc.
- Otsu N (1979) A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9 (1): 62–66.
- Reigber A, Moreira A (2000) First demonstration of airborne SAR tomography using multi-baseline L-band data. *IEEE Transactions on Geoscience and Remote Sensing*, 38 (5): 2142–2152.
- Reznik S, Mayer H (2007) Implicit Shape Models, Model Selection, and Plane Sweeping for 3D Facade Interpretation. In: *Proceedings of Photogrammetric Image Analysis 2007*: 173–178.
- Ripperda N (2010) *Rekonstruktion von Fassadenstrukturen mittels formaler Grammatiken und reversible jump Markov Chain Monte Carlo sampling*. PhD thesis, Leibniz Universität Hannover, Bayerische Akademie der Wissenschaften, Verlag C.H. Beck. Schriftenreihe C der DGK, no. 649.
- Ripperda N, Brenner C (2009) Application of a Formal Grammar to Facade Reconstruction in Semiautomatic and Automatic Environments. In: *Proceedings of 12th AGILE International Conference on Geographic Information Science*: 1–12.
- Schack L, Soergel U (2014) Exploiting Regular Patterns to Group Persistent Scatterers in Urban Areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (10): 4177–4183.
- Schack L, Soergel U (2015) Matching Persistent Scatterers to Optical Oblique Images. In: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*: 52–60.
- Schack L, Soergel U, Heipke C (2015) Persistent Scatterer Aided Facade Lattice Extraction in Single Airborne Optical Oblique Images. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-3/W4 (March): 197–205.

- Schindler G, Krishnamurthy P, Lubliner R, Liu Y, Dellaert F (2008) Detecting and matching repeated patterns for automatic geotagging in urban environments. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 1–7.
- Schunert A (2014) *Assignment of Persistent Scatterers to Buildings*. PhD thesis, Leibniz Universität Hannover, Bayerische Akademie der Wissenschaften, Verlag C.H. Beck. Schriftenreihe C der DGK, no. 728.
- Schunert A, Soergel U (2012) Grouping of Persistent Scatterers in high-resolution SAR data of urban scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73: 80–88.
- Sester M, Feuerhake U, Kuntzsch C, Zhang L (2012) Revealing Underlying Structure and Behaviour from Movement Data. *KI - Künstliche Intelligenz*, 26 (3): 223–231.
- Shields R (2012) Cultural Topology: The Seven Bridges of Königsburg, 1736. *Theory, Culture & Society*, 29 (43): 43–57.
- Sirmacek B, Unsalan C (2009) Urban-area and building detection using SIFT keypoints and graph theory. *IEEE Transactions on Geoscience and Remote Sensing*, 47 (4): 1156–1167.
- Slama CC, Theurer C, Henriksen SW (1980) *Manual of photogrammetry*. American Society of Photogrammetry, 4th edition.
- Soergel U (2010) Review of radar remote sensing on urban areas. In: Soergel U (ed) *Radar remote sensing of urban areas* chapter 1. Springer Netherlands.
- Stilla U, Michaelsen E, Soergel U, Schulz K (2003) Perceptual grouping of regular structures for automatic detection of man-made objects. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 6: 3525–3527.
- Tao J, Auer S, Palubinskas G, Reinartz P, Bamler R (2014) Automatic SAR simulation technique for object identification in complex urban scenarios. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (3): 855–864.
- Terranova C, Ventura G, Vilaro G (2015) Multiple causes of ground deformation in the Napoli metropolitan area (Italy) from integrated Persistent Scatterers InSAR, geological, hydrological, and urban infrastructure data. *Earth Science Reviews*, 146: 105–119.
- Tosi L, Teatini P, Strozzi T (2013) Natural versus anthropogenic subsidence of Venice. *Scientific Reports*, 3 (2710): 1–9.
- Tuia D, Munoz-Marí J, Gómez-Chova L, Malo J (2013) Graph matching for adaptation in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 51 (1): 329–341.
- Wertheimer M (1938) Laws of Organization in Perceptual Forms. In: Ellis W (ed) *A Source Book of Gestalt Psychology* chapter II, (pp. 71–88). London: Paul, Routledge & Kegan.
- Xiao H, Meng G, Wang L, Xiang S, Pan C (2014) Facade repetition extraction using block matrix based model. In: *Proceedings of IEEE International Conference on Image Processing*: 1673–1677.
- Xiao J, Gerke M, Vosselman G (2010) Automatic detection of buildings with rectangular flat roofs from multi-view oblique imagery. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38 (3A): 251–256.
- Xu Y, Ma P, Ng E, Lin H (2015) Fusion of WorldView-2 Stereo and Multitemporal TerraSAR-X Images for Building Height Extraction in Urban Areas. *IEEE Geoscience and Remote Sensing Letters*, 12 (8): 1795–1799.
- Yoon YT, Eineder M, Yague-Martinez N, Montenbruck O (2009) TerraSAR-X Precise Trajectory Estimation and Quality Assessment. *IEEE Transactions on Geoscience and Remote Sensing*, 47 (6): 1859–1868.
- Zhu X, Ge N, Shahzad M (2015) Joint Sparsity in SAR Tomography for Urban Mapping. *IEEE Journal of Selected Topics in Signal Processing*, 9 (8): 1498–1509.
- Zhu XX, Montazeri S, Gisinger C, Hanssen RF, Bamler R (2016) Geodetic SAR Tomography. *IEEE Transactions on Geoscience and Remote Sensing*, 54 (1): 18–35.
- Zhu XX, Shahzad M (2014) Facade reconstruction using multiview spaceborne TomoSAR point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 52 (6): 3541–3552.

Appendices

A. Proof of König-Egerváry Theorem

The König-Egerváry Theorem is central to the bipartite matching problem. The theorem's proof is given after some important relations between matchings and covers are stated. The terminology and notation of Section 2.4 is retained. Two more terms have to be introduced [Gould, 1988]:

Definition A .1 (weak node). An edge is *weak* with respect to a matching \mathcal{M} if it is not in the matching. A node is *weak* with respect to \mathcal{M} if it is only incident to weak edges.

In Figure 2.9b node C is an example for a weak node. The path (I, B, III, A, II, D) is an alternating path:

Definition A .2 (alternating path). An *alternating path* with respect to a matching \mathcal{M} is a path whose edges are alternately in \mathcal{M} and not in \mathcal{M} . It does not matter whether the first edge of the path is in \mathcal{M} or not.

With these definitions the following lemma can be stated:

Lemma A .3. If \mathcal{M} is a matching of a graph \mathcal{G} and \mathcal{K} is any cover of \mathcal{G} , then $|\mathcal{M}| \leq |\mathcal{K}|$.

Proof. Following Definition 2.3 (page 33) \mathcal{K} has to contain at least one node from each edge in \mathcal{M} . Following Definition 2.5 no node in \mathcal{K} can be an endpoint of more than one edge in \mathcal{M} and, thus, $|\mathcal{M}| \leq |\mathcal{K}|$. \square

Corollary A .4. Given a maximum matching \mathcal{M}_{max} and $\tilde{\mathcal{K}}$ is a minimum cover of the same graph \mathcal{G} , then $|\mathcal{M}_{max}| \leq |\tilde{\mathcal{K}}|$.

Proof. This follows directly from Lemma A.3 since any constructable graph with $|\mathcal{M}_{max}| > |\tilde{\mathcal{K}}|$ is not bipartite. \square

Now, joining the above considerations lets one formulate the Kőnig-Egerváry Theorem:

Theorem A .1 (Kőnig-Egerváry Theorem I). The maximum number $|\mathcal{M}|$ of edges in a matching \mathcal{M} of a bipartite graph $\mathcal{G} = (\{\mathcal{V}_1, \mathcal{V}_2\}, \mathcal{E}, \mu, \nu)$ equals the number of nodes $|\mathcal{C}|$ in a minimum cover of \mathcal{M} .

Proof. Assuming a maximum matching \mathcal{M} in a bipartite graph $\mathcal{G} = (\{\mathcal{V}_1, \mathcal{V}_2\}, \mathcal{E}, \mu, \nu)$ and \mathcal{W} the weak nodes in \mathcal{V}_1 with respect to \mathcal{M} . The number of nodes $|\mathcal{M}|$ in the Matching equals $|\mathcal{M}| = |\mathcal{V}_1| - |\mathcal{W}|$. Let \mathcal{S} be those nodes in \mathcal{G} that are connected to some node in \mathcal{W} by an alternating path. The intersection of \mathcal{V}_1 and \mathcal{S} is called \mathcal{S}_1 and the intersection of \mathcal{V}_2 and \mathcal{S} \mathcal{S}_2 , respectively. By definition, no node in $\mathcal{S}_1 - \mathcal{W}$ is weak and, therefore, $\mathcal{S}_1 - \mathcal{W}$ is matched under \mathcal{M} to \mathcal{S}_2 . Thus, $|\mathcal{S}_1| - |\mathcal{S}_2| = |\mathcal{W}|$. Let $\mathcal{C} = (\mathcal{V}_1 - \mathcal{S}_1) \cup \mathcal{S}_2$ be a cover for all edges \mathcal{E} from the graph \mathcal{G} . The number of nodes $|\mathcal{C}|$ in this cover is $|\mathcal{C}| = |\mathcal{V}_1| - |\mathcal{S}_1| + |\mathcal{S}_2|$. Substituting $|\mathcal{W}|$ and $|\mathcal{M}|$ yields $|\mathcal{C}| = |\mathcal{V}_1| - |\mathcal{W}| = |\mathcal{M}|$. Hence, the number of nodes in a maximum matching equals the minimum number of covered nodes for \mathcal{G} . \square

B. Example of the Hungarian algorithm

In order to make the principle of the Hungarian algorithm more graphic an example is given step by step. The toy example outlined in section 2.5 is continued. Four workers have to be assigned to three jobs. Every worker has individual costs to fulfill every job. This can be expressed in the following cost matrix with fictional values:

	Job I	Job II	Job III
Worker A	5	7	4
Worker B	2	6	10
Worker C	8	7	5
Worker D	7	3	1

The Hungarian algorithm in four steps as outlined in Algorithm 1 (page 38) is applied as follows (changes are color-marked):

Step 1: Add dummy rows or columns if cost matrix is not square.

Since the Hungarian algorithm works only on square matrices, a fourth column with dummy entries is added:

	Job I	Job II	Job III	Dummy
Worker A	5	7	4	0
Worker B	2	6	10	0
Worker C	8	7	5	0
Worker D	7	3	1	0

Step 2: Subtract row and column minima.

Considering Lemma 2.8 (page 37) the relative costs for all assignments do not change when row- and column-wise constants are subtracted. Since the row-wise minima are always 0 due to the dummy column this step is skipped. Subtracting the column-minima results in:

	Job I	Job II	Job III	Dummy
Worker A	3	4	3	0
Worker B	0	3	9	0
Worker C	6	4	4	0
Worker D	5	0	0	0

If one had now one zero in every row or column the assignment problem would be solved, simply by assigning the minimal costing jobs to every worker. In this example worker D, for instance, has zero costs for jobs II and III. Thus, the globally minimal solution cannot be derived from this intermediate cost matrix. In order to eliminate such ambiguities Steps 3 and 4 are repeated until a unambiguous assignment can be done.

Step 3: Cover all zeros with a minimum number of lines.

Three lines are necessary to cover all zero elements:

	Job I	Job II	Job III	Dummy
Worker A	3	4	3	0
Worker B	0	3	9	0
Worker C	6	4	4	0
Worker D	5	0	0	0

Following the König-Egerváry Theorem 2.7 (page 36) four lines would be necessary to derive a complete matching. Therefore, additional zeros have to be created.

Step 4: Create additional zeros.

The smallest not covered element in the cost matrix is 3. This number is subtracted from all uncovered elements and added to all elements that are covered twice. The resulting cost matrix is than:

	Job I	Job II	Job III	Dummy
Worker A	0	1	0	0
Worker B	0	3	9	3
Worker C	3	1	1	0
Worker D	5	0	0	3

Now, Step 3 is repeated. The minimum number of lines to cover all zero elements of the cost matrix is now four. Therefore, a complete and optimal assignment is found among the independent zeros. There is only one solution to mark four independent zeros:

	Job I	Job II	Job III	Dummy
Worker A	0	1	0	0
Worker B	0	3	9	3
Worker C	3	1	1	0
Worker D	5	0	0	3

This matching corresponds to the original costs of

	Job I	Job II	Job III	Dummy
Worker A	5	7	4	0
Worker B	2	6	10	0
Worker C	8	7	7	5
Worker D	7	3	1	0

Since Worker C is matched to the Dummy job, he stays jobless and is not part of the matching.

List of Figures

1.1. One facade in two sensing geometries	12
1.2. Threefold reflections at trihedral corners	13
1.3. Flowchart of method	15
2.1. Sketch of SAR imaging geometry	18
2.2. Sketch of signal mixing due to layover	20
2.3. Example for layover	21
2.4. Sketch of complex signal	22
2.5. Sketch of the two imaging geometries	25
2.6. Flowchart of transformations between coordinate systems	27
2.7. Contributions to inaccuracy budget	32
2.8. Example of graphs	34
2.9. Example of graph matching	36
4.1. Examples for regularity in urban areas	53
4.2. Example of facade appearances in optical image and SAR	56
4.3. Facade segmentation	58
4.4. Sketch of projection onto facade plane	61
4.5. Example of facade plane estimation.	63
4.6. Sketch and result of estimating the lattice spanning vectors	66
4.7. Result of PS grouping and classification.	68
4.8. Sketch of image rectification	70
4.9. Normalized cross correlation	72
4.10. Window model	75
4.11. Result of window corner extraction	76
4.12. Example for necessity of topology information	80
4.13. Flow chart of Iterative Graph-based Bipartite Matching algorithm	81
4.14. Example of topology violation	82
4.15. 95% confidence ellipses for PS	91

5.1. Data overview	95
5.2. Segmentation result: Park Kolonnaden	98
5.3. Result of PS grouping: Park Kolonnaden	99
5.4. Effect of projection to plane on 95% confidence ellipse.	100
5.5. Facade plane estimation and grouping of PS: Charlottenstraße	101
5.6. Window extraction: Park Kolonnaden	104
5.7. Window Extraction: Charlottenstraße	106
5.8. Window extraction: Axel-Springer building	108
5.9. IBM result: Park Kolonnaden	109
5.10. Comparison of averaged matching vectors	111
5.11. Result: Charlottenstraße	112
5.12. IBM Result: Axel-Springer building	113
5.13. Ebertstraße: IBM result with confidence ellipses ellipses	115
5.14. Dependency of the IGBM result on α	117
5.15. Application of methods to overall study area.	119
5.16. Typical negative example	120

List of Tables

2.1. Examples of typical localization precisions of single PSs in SAR coordinates.	28
2.2. Substitution of trigonometric terms.	30
2.3. A priori accuracies of parameter set.	31
4.1. Comparison of inaccuracy budgets with and without object knowledge. . . .	89
4.2. Resulting accuracies for three exemplary facades in relation to the patch size.	90
5.1. Properties of used data.	94
5.2. Result of manually labeling facades.	96
5.3. Applied parameters	97

Danksagung

Ich möchte mich bei einer ganzen Reihe von Personen bedanken, die direkt und indirekt an der Entstehung dieser Arbeit beteiligt waren. Als erstes möchte ich mich bei meinen beiden Doktorvätern bedanken. Uwe Sörgel brachte mir sehr viel Vertrauen gerade zu Anfang meiner Zeit am IPI entgegen. Christian Heipke ermöglichte mir nach dem Weggang von Uwe Sörgel und dem mir zugeordneten DFG-Projekt aus Hannover, dass ich in Hannover meine Arbeit fortsetzen konnte. Frau Monika Sester danke ich für die Übernahme des Korreferats sowie wertvolle inhaltliche Vorschläge.

Bei Ralf Schroth und BLOM UK möchte ich mich für das Zurverfügungstellen der verwendeten Schrägluftbilder bedanken. Stefan Auer und Stefan Gernhardt von der TU München haben mir nicht nur die Persistent Scatterer Punktwolken zukommen lassen, sondern gerade zu Beginn meiner Arbeit interessante Hinweise gegeben. Vielen Dank auch an Mareike Dorozynski, die als Hiwi Fachwissen, Kreativität sowie ihre Ausdauer unter Beweis stellte. Allen Kolleginnen und Kollegen am IPI gilt ein großer Dank für das tolle Klima im Institut!

Endlose wertvolle Diskussionen sowie unzählige koffeinhaltige Heißgetränke habe ich mit meinen über die Zeit wechselnden Bürokollegen Alexander Schunert, Martin Reich und Andreas Paul genossen. Martin möchte ich darüber hinaus für das Korrekturlesen danken.

Ganz besonders möchte ich mich bei meiner Familie bedanken. Zum einen bei meinen Eltern, die letztendlich die Weichen dafür stellten, dass diese Arbeit entstanden ist. Zum anderen bei Claudia und Lina. Vielen Dank für Euer Verständnis, wenn ich besonders in den letzten Wochen vor der Abgabe abends erst sehr spät nach Hause kam! Ihr habt mich voll und ganz unterstützt und Ihr seid das Allerwichtigste für mich!

LEBENS LAUF

BERUF SERFAHRUNG

Nov. 2011 – Sep. 2016 wissenschaftlicher Mitarbeiter am Institut für Photo-
grammetrie und GeoInformation,
Leibniz Universität Hannover

STUDIUM

Okt. 2011 Abschluss als Master of Science in Geodäsie und Geo-
informatik (Bewertung: „mit Auszeichnung“)
Masterarbeit:
„Tomographie in hochaufgelösten SAR-Daten“

Okt. 2009 – Sep. 2011 M.Sc. Studium Geodäsie und Geoinformatik
an der Leibniz Universität Hannover
Vertiefung: Geoinformatik

Okt. 2006 – Sep. 2009 B.Sc. Studium Geodäsie und Geoinformatik
an der Leibniz Universität Hannover

AUSLANDSAUFENTHALT

Sep. 2009 – Apr. 2010 ERASMUS Auslandssemester in Valencia, Spanien

FORSCHUNGS-EXPEDITIONEN

Okt. 2010 – Nov. 2010 Kampagne ANT XXVII/1 auf dem Forschungsschiff
„Polarstern“ für das Alfred-Wegener-Institut
Messungen mit einem Fächersonar zur Kartierung des
Meeresbodens

Feb. 2013 – Apr. 2013 Antarktis-Expedition für das Alfred-Wegener-Institut
GPS-Messungen im Rahmen der deutsch-
argentinischen Messkampagne CAV13

ZIVILDIENTST

Aug. 2005 – Apr. 2006 Zivildienst in der Tagespflege Bremen-Grambke

SCHULBILDUNG

Jun. 2005 Abiturprüfung bestanden mit Note 1,5
2002 – 2005 Gymnasium Schulzentrum Bördestraße, Bremen